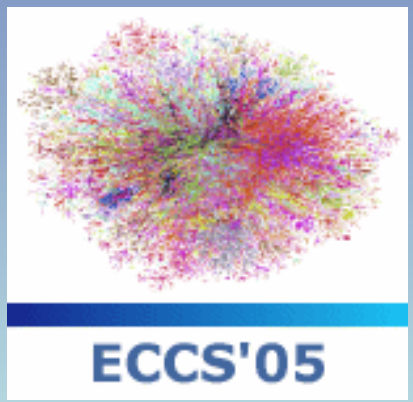
ECSS

European Complex Systems Society

Proceedings of ECCS'05

The European Conference on Complex Systems

Paul Bourgine
François Képès
Marc Schoenauer



PROCEEDINGS of

ECCS'05

European Conference on Complex Systems

Paris, November 14-18, 2005

Towards a science of complex systems

Complex systems, as networks of interactive entities, are studied through a rapidly increasing mass of data in all domains. At the same time, these domains share a lot of new and fundamental theoretical questions. This situation is especially favourable for developing the new science of complex systems in an interdisciplinary way. The ECCS'05 is a step towards this new science.

There are two kinds of interdisciplinarity within complex systems. The first kind begins with a particular complex system and addresses a variety of questions coming from its particular domain and points of view. The second kind addresses issues that are fundamental to complex systems in general. The first kind leads to domain-specific interdisciplinary fields such as cognitive science. The new science of complex system belongs to the] second kind of interdisciplinarity. It starts from fundamental open questions relevant to many domains, and searches for methods to deal with them.

These two kinds of interdisciplinarity are complementary and interdependent: any advance in one is valuable for the other. The science of complex systems will develop through a constantly renewed process of reconstructing data from models with a permanent interaction between the two kinds of interdisciplinarity. The reconstruction of the dynamics of complex systems presents a major challenge to modern science but it is becoming increasingly accessible through an accumulating mass of data, combined with the increasing power of computers, leading to theoretical advances in understanding.

This conference follows the one organized in Torino (Italy) in December 2004 with support from the coordination actions EXYSTENCE and ONCE-CS, funded by the Future and Emerging Technologies' unit of the European Commission. ECCS'05 benefits from the same support and is the first conference in an annual series organized by the new European Complex System Society (ECSS) and its Conference Steering Committee.

We hope that the participants will appreciate the beautiful venue of the conference this year, at the Cité Internationale Universitaire de Paris.

Our special thanks to the staff at CIUP for preparing the ground to this conference. We would also like to thank the sponsors of ECCS'05 for making it possible for all the participants to share their enthusiasm and ideas in the most constructive way.

The ECCS'05 Program Committee,
The ECCS'05 Local Organization Committee,
The ECSS Conference Steering Committee.

Editorial Information

This volume contains the papers that have been presented at ECCS'05, the European Conference on Complex Systems, that took place in Paris, November 14-18, 2005.

In total, 273 papers have been submitted to the conference, and the Program Committee met during 2 full days to build the final program of the conference, accepting 3 types of presentations: 30 papers were accepted as long communications and were allowed 40 minutes talks; 66 papers were accepted as short communications and were given 20 minutes talks; Finally, 95 posters were displayed continuously during the conference, with 3 specific poster sessions for discussions.

The communications (long and short together) were grouped into thematic sessions in the program of the conference, and the same categorization has been used here. Each thematic chapter first includes the long communications for the corresponding theme, followed by the short communications. All posters were allowed a 1-2 pages abstract, and those abstracts are grouped together at the end of this volume.

Because this conference was the first of its kind (hopefully the first of a long series), a few practical details were overlooked during the preparation, and the lack of time later forbid any real-time adjustment. This is why this volume is probably not reaching the high standards of typesetting quality we would have wished. For instance, even though the authors were given a chance to react and modify their papers according to the reviewers? comments, there was not enough time to decently ask authors of long papers accepted as short communications to reduce the length of their papers. We are the only ones to blame, our apologies to the authors, and to the readers.

Nevertheless, this leaves room for large improvements for the future editions of ECCS (starting in 2006, in OXFORD), and we are confident that the readers will in any case enjoy the high scientific quality of the contributions included in this volume.

But whereas all errors and mistakes in the layout of this volume are ours, many people are to be thanked for contributing to make this conference the success it has been. Many thanks to Geneviève Tual, David Chavalarias, Jean-Baptiste Souffron, Bertrand Chardon, blablabla à compléter.

Paul Bourgine, François Képès, Marc Schoenauer,
December 2005.

Steering Committee

Paul Bourgine, Ecole Polytechnique (France)
Jeff Johnson, Open University (UK)
Jürgen Jost, Max Planck Institute-Leipzig (or MPI for Mathematics in the Sciences) (Germany)
François Képès, Epigenomics Project, Genopole® (France)
Cris Moore, Santa Fe Institute and UNM (USA)
Michel Morvan, chair, ENS-Lyon (France)
Grégoire Nicolis, ULB (Belgium)
Sorin Solomon, HUJ/ISI (Italy)

Local Committee

Paul Bourgine, CREA, Ecole Polytechnique / CNRS, Paris (France)
David Chavalarias, CREA, Ecole Polytechnique / CNRS, Paris (France)
François Képès, Chair, Epigenomics Project, Genopole®, Evry (France)
Marie-Jo Lécuyer, CREA, Ecole Polytechnique / CNRS, Paris (France)
Catherine Meignen, Genopole® Recherche, Evry (France)
Hélène Pollard, Genopole® Recherche, Evry (France)
Marc Schoenauer, INRIA Futurs, Orsay (France)
Geneviève Tual, ECSS, (France)

Program Committee

Ozalp Babaoglu, University of Bologna (Italy)
Paul Bourgine, chair, Ecole Polytechnique (France)
Vladimir Batagelj, University of Ljubljana (Slovenia)
Jean-Louis Deneubourg, Free University of Bruxelles (Belgium)
Jean-Pierre Françoise, P.-M. Curie University/Paris VI (France)
Nigel Gilbert, University of Surrey (UK)
Dirk Helbing, Dresden University of Technology (Germany)
Jeffrey Johnson, Open University (UK)
Jürgen Jost, Max Planck Institute-Leipzig (Germany)
Scott Kirkpatrick, Hebrew University of Jerusalem (Israel)
Kristian Lindgren, Göteborg University (Sweden)
John McCaskill, Ruhr University Bochum (Germany)
Eve Mitleton-Kelly, London School of Economics (UK)
Rémi Monasson, ENS (France)
Michel Morvan, ENS-Lyon (France)
Norman Packard, Protolife (Italy)
Denise Pumain, Paris-Sorbonne University (France)
Felix Reed-Tsochas, University of Oxford (UK)
Vincent Schächter, Genoscope (France)
Frank Schweitzer, ETH-Zürich (Switzerland)
Peter Sloot, University of Amsterdam (Netherlands)
Paul Spirakis, RACTI (Greece)
Luc Steels, Free University of Bruxelles (Belgium)
Eörs Szathmary, Collegium Budapest (Hungary)
Alessandro Vespignani, LPT, Orsay (France)

External Reviewers

This year, the Program Committee members were the main reviewers of all submitted papers. However, in some occasions, external reviewers were asked to review some papers, and we would like to thank them here.

Gerard de Zeeuw, University of Amsterdam (Netherlands)
Peter K. Allen , Columbia University (USA)
Pierpaolo Andriani, advanced Institute of Management Research (UK),
Jannis Kallinikos, LSE (UK)

Table of Contents

Biological Modelling

~

Long Communications

Short Communications

Social Modelling

~

Long Communications

Short Communications

Bio Inspired Methods

~

Long Communications

Short Communications

Complex Systems Methods

~

Long Communications

Short Communications

Information Technology Modelling

~

Long Communications

Short Communications

Cognition Modelling

~

Long Communications

Short Communications

Network Modelling

~

Long Communications

Short Communications

Posters

Biological Modelling

Invariant grids: method of complexity reduction in reaction networks

Alexander Gorban[2,4] *ag153@leicester.ac.uk*

Iliya Karlin[1,2] *karlin@lav.mavt.ethz.ch*

Andrei Zinovyev[2,3] *andrei.zinovyev@curie.fr*

[1] ETH, Swiss Federal Institute of Technology, Switzerland
[2] Institute of Computational Modeling SB RAS, Russia
[3] Bioinfomatics service of Institut Curie, France
[4] Centre for mathematical modeling, University of Leicester

Abstract

Complexity in the description of big chemical reaction networks has both structural (number of species and reactions) and temporal (very different reaction rates) aspects. A consistent way to make model reduction is to construct the invariant manifold which describes the asymptotic system behavior. In this paper we present a discrete analog of this object: an invariant grid. Invariant grid is introduced independently from the invariant manifold notion and can serve itself to represent the dynamic system behavior as well as to approximate the invariant manifold after refinement. The method is designed for pure dissipative systems and widely uses their thermodynamic properties but allows also generalizations for some classes of open systems. The method is illustrated by two examples: the simplest catalytic reaction (Michaelis-Menten mechanism) and the hydrogen oxidation.

Keywords: Kinetics; Model Reduction; Grids; Invariant Manifold; Entropy; Nonlinear Dynamics; Mathematical Modeling; Numerical methods

Running title: Method of invariant grid

Corresponding author: Andrei Zinovyev, Institut Curie, Service Bioinformatique, rue d'Ulm, 26, 75248, Paris, France. Tel: +33 1 42 34 65 27; Fax: +33 1 42 34 65 28.

1 Introduction

Reaction networks serve as a good model to imitate and predict behavior of complex systems of interacting components. Modern research faces with constantly increasing complexity of the systems under study: as a good example, nowadays one can observe a boom connected with studies of biochemical processes in a living cell (for recent overviews, see [1],[2]). There is no need to underline emerging needs for the methods of reducing the complexity of system description and behavior.

Complexity in modeling big chemical reaction networks has both structural (number of species and reactions) and temporal (very different reaction rates) aspects, see Fig. 1. In general, it is not possible to disregard the temporal organization of the network when one wants to create a realistic system model. Of course, the rate constants and reaction laws are rarely available completely. This makes extremely desirable the development of methods allowing to reduce the number of system parameters as well as methods for qualitative analysis of chemical reaction networks [2].

The idea of model reduction with respect to slow motion extraction can be introduced as follows: we have a system of ordinary differential equations describing time evolution of n species concentrations (or masses) in time:

$$\frac{d\boldsymbol{x}}{dt} = J(\boldsymbol{x}), \tag{1}$$

Every particular state of the system corresponds to a point in the phase space U and the system dynamics is determined by the vector field $J(x)$, $x \in U$. We construct new (reduced) dynamics

$$\frac{d\boldsymbol{y}}{dt} = J'(\boldsymbol{y}), \tag{2}$$

where $y_i, i = 1..m, m \ll n$ is a new set of variables corresponding to slow dynamics of the initial system (1). By analogy with statistical physics it corresponds to the "macroscopic" description of the chemical system (we observe only effects of slow system changes, comparable in time scale with characteristic times of experimental measurements) as opposite to "microscopic" variables x_i. The reduced system dynamics exists on a m-dimensional manifold (surface) Ω embedded in the n-dimensional phase space and defined by functions $x_i = x_i(y_1, ..., y_m)$.

A consistent way for model reduction is to construct a positively invariant slow manifold Ω_{inv}, such that if an individual trajectory of the system (1) has started on Ω_{inv}, it does not leave Ω_{inv} anymore, i.e. the vector field $J(x)$ in the points of the manifold is tangent to it, Fig. 2a. The 'ideal' picture of the reduced description we have in mind is as follows: A typical phase trajectory, $x(t)$, where t is the time, and x is an element of the phase space, consists of two pronounced segments. The first segment connects the beginning of the trajectory, $x(0)$, with a certain point, $x(t_1)$, on the manifold Ω_{inv} (rigorously speaking, we should think of $x(t_1)$ not on Ω_{inv} but in a small neighborhood of Ω_{inv} but this is inessential for the ideal picture). The second segment belongs to Ω_{inv}. Thus, the manifolds appearing in our ideal picture are "patterns" formed by the segments of individual trajectories, and the goal of the reduced description is to "filter out" this manifold (Fig. 2a).

Usually construction of invariant manifold in the explicit form is difficult. Most of the time one deals with its approximation constructed using some method (see, for overview, [6], [8], [4], [5]). It is formally possible to induce new dynamics on any given manifold Ω, not necessarily invariant, if one introduces a projector operator P of the vector field on the tangent bundle of the manifold Ω: $PJ(x \in \Omega) \in T_x\Omega$. By definition, the manifold Ω is invariant with respect to the vector field J if an only if the following equality is true for each $x \in \Omega$:

$$[1 - P]J(x) = 0, \tag{3}$$

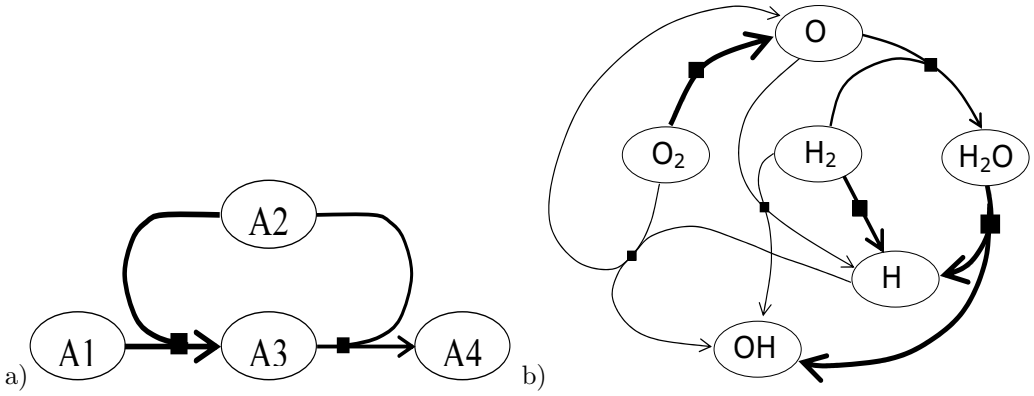


Figure 1: Graphical representation of two model systems considered as examples in this paper: a) Michaelis-Menten mechanism; b) Hydrogen burning model with 6 variables. Here circles represent chemical species, squares represent chemical reactions. Line widths reflect direct reaction rate constants, thicker line corresponds to a slower reaction (in a logarithmic scale). All reactions here are governed by mass action law and supposed to be reversible.

where projector P depends on the point x and on the manifold Ω in the vicinity of x. This equation is a differential equation for functions that define the manifold Ω. Newton method and relaxation method, both iterative, were proposed to find a sequence of corrections to some initial approximation Ω, in such a way that every next approximation has less *invariance* defect $[1-P]J(x)$, see [5]. These corrections can be performed analytically in some cases.

For the case of a complex chemical reaction network, one has to develop a computationally effective method of invariant manifold construction. If one constructs a surface of a relatively low dimension, grid-based manifold representations become a relevant option [8]. In this paper we present such an approach named *method of invariant grids* (MIG). From one hand, grid representation can be refined and converge more and more closely to the invariant manifold. From the other, we define *invariant* grid as an object independent on the manifold itself. Thus, it can be used independently: for example, for visualization of the global system dynamics as it will be shown in the end of this paper.

Invariant grid is an undirected graph which consists of a set of nodes and connections between them. The graph can be represented in two spaces: in the low-dimensional space of the internal (reduced) coordinates where it forms a finite lattice (usually, regular and rectangular or hexagonal), and, simultaneously, it is embedded in the phase space U, thus every node corresponds to a species concentrations vector $\boldsymbol{x}$. Using connectivity of the graph, one can introduce differentiation operators and calculate the tangent vectors and define the projector operator in every node. This is the only place where the connectivity of the graph is used. The node positions in U are optimized such that the invariance condition (3) is satisfied *for every node*. In this paper we propose two algorithms for how to do it, both iterative: of Newton type and a relaxation method. After node positions optimization the grid is called *invariant*.

In this study we consider class of dissipative systems, i.e. such systems for which there exists a global convex Lyapunov function G (thermodynamic potential) which implements the second law of thermodynamics. For example, because of this reason, all reactions on Fig. 1 are reversible. Dissipative systems have the only steady state in the equilibrium point and as the time t tends to infinity, the system reaches the equilibrium state while in the course of the transition the Lyapunov function decreases monotonically. Thermodynamic properties of dissipative systems help a lot: for example, they unambiguously define metrics in the phase space to perform geometrical calculations and also define the choice of projector P almost uniquely (see the next section).

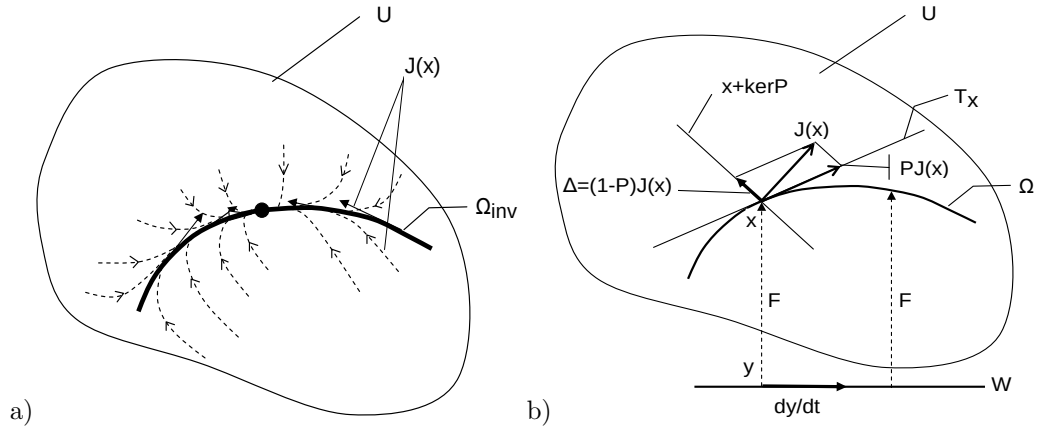


Figure 2: Main geometrical structures of model reduction: U is the phase space, $J(x)$ is the vector field of the system under consideration: $dx/dt = J(x)$, Ω is an ansatz manifold, W is the space of macroscopic variables (coordinates on the manifold), the map $F : W \to U$ maps any point $y \in W$ into the corresponding point $x = F(y)$ on the manifold Ω, T_x is the tangent space to the manifold Ω at the point x, $PJ(x)$ is the projection of the vector $J(x)$ onto tangent space T_x, the vector field dy/dt describes the induced dynamics on the space of parameters, $\Delta = (1 - P)J(x)$ is the defect of invariance, the affine subspace $x + \ker P$ is the plain of fast motions, and $\Delta \in \ker P$. a) Here Ω_{inv} is an invariant manifold (all $J(x \in \Omega_{inv})$ are tangent to Ω_{inv}) and a possible dynamics is shown in its vicinity; b) here Ω is some manifold approximating the invariant manifold ($J(x \in \Omega)$ is not necessarily tangent to Ω), one can use operator P to derive new dynamics (2).

Low dimensional invariant manifolds exist also for systems with a more complicated dynamic behavior so why to study the invariant manifolds of slow motions for a particular class of purely dissipative systems? The answer is in the following: Most of the physically significant models include non-dissipative components in a form of either a conservative dynamics or in the form of external fluxes. For example, one can think of irreversible reactions among the suggested stoichiometric mechanism (inverse process are so unprobable that we discard them completely thereby effectively "opening" the system to the remaining irreversible flux). For all such systems, the method of invariant grids is applicable almost without special refinements, and bears the significance that invariant manifolds are constructed as a "deformation" of the relevant manifolds of slow motion of the purely dissipative dynamics. Example of this construction for open systems is presented below in the last section of the paper. The calculations in the last chapter do not use grid specifics and can be applied not only for grid representation of the invariant manifold, but also for any analytical form of its representation.

2 Dissipative systems and thermodynamic projector

2.1 Kinetic equations

Let us introduce the notions used in the paper (see also [3], [9], [7]). We will consider a closed system with n chemical species $\mathrm{A}_1, \ldots, \mathrm{A}_n$, participating in a complex reaction. The complex reaction is represented by the following stoichiometric mechanism:

$$\alpha_{s1}\mathrm{A}_1 + \ldots + \alpha_{sn}\mathrm{A}_n \rightleftharpoons \beta_{s1}\mathrm{A}_1 + \ldots + \beta_{sn}\mathrm{A}_n, \tag{4}$$

where the index $s = 1, \ldots, r$ enumerates the reaction steps, and where integers, α_{si} and β_{si}, are stoichiometric coefficients. For each reaction step s, we introduce n–component vectors $\boldsymbol{\alpha}_s$ and $\boldsymbol{\beta}_s$ with components α_{si} and β_{si}. Notation $\boldsymbol{\gamma}_s$ stands for the vector with integer components $\gamma_{si} = \beta_{si} - \alpha_{si}$ (the stoichiometric vector).

For every A_i an *extensive variable* N_i, "the number of particles of that species", is defined. The concentration of A_i is $x_i = N_i/V$, where V is the volume.

Given the stoichiometric mechanism (4), the reaction kinetic equations read:

$$\dot{\boldsymbol{N}} = V\boldsymbol{J}(\boldsymbol{x}),\ \boldsymbol{J}(\boldsymbol{x}) = \sum_{s=1}^{r} \boldsymbol{\gamma}_s W_s(\boldsymbol{x}), \tag{5}$$

where dot denotes the time derivative, and W_s is the reaction rate function of the step s. In particular, *the mass action law* suggests the polynomial form of the reaction rates:

$$W_s(\boldsymbol{x}) = W_s^+(\boldsymbol{x}) - W_s^-(\boldsymbol{x}) = k_s^+(T)\prod_{i=1}^{n} x_i^{\alpha_i} - k_s^-(T)\prod_{i=1}^{n} x_i^{\beta_i}, \tag{6}$$

where $k_s^+(T)$ and $k_s^-(T)$ are the constants of the direct and of the inverse reactions rates of the sth reaction step, T is the temperature.

The rate constants are not independent. The *principle of detail balance* gives the following connection between these constants: There exists such a positive vector $\boldsymbol{x}^{\rm eq}(T)$ that

$$W_s^+(\boldsymbol{x}^{\rm eq}) = W_s^-(\boldsymbol{x}^{\rm eq}) \text{ for all } s = 1, \dots, r. \tag{7}$$

For V, $T = const$ we do not need additional equations and data. It is possible simply to divide equation (5) by the constant volume and to write

$$\dot{\boldsymbol{x}} = \sum_{s=1}^{r} \boldsymbol{\gamma}_s W_s(\boldsymbol{x}). \tag{8}$$

Conservation laws (balances) impose linear constrains on admissible vectors $\boldsymbol{x}$:

$$(\boldsymbol{b}_i, \boldsymbol{x}) = B_i = const,\ i = 1, \dots, l, \tag{9}$$

where $\boldsymbol{b}_i$ are fixed and linearly independent vectors. Let us denote as $\boldsymbol{B}$ the set of vectors which satisfy the conservation laws (9) with given B_i:

$$\boldsymbol{B} = \{\boldsymbol{x} | (\boldsymbol{b}_1, \boldsymbol{x}) = B_1, \dots, (\boldsymbol{b}_l, \boldsymbol{x}) = B_l\}.$$

The natural phase space $\boldsymbol{X}$ of the system (8) is the intersection of the cone of n-dimensional vectors with nonnegative components, with the set $\boldsymbol{B}$, and $\dim \boldsymbol{X} = d = n - l$. In addition, we assume that each of the conservation laws is supported by each elementary reaction step, that is

$$(\boldsymbol{\gamma}_s, \boldsymbol{b}_i) = 0, \tag{10}$$

for each pair of vectors $\boldsymbol{\gamma}_s$ and $\boldsymbol{b}_i$.

We assume that the kinetic equation (8) describes evolution towards the unique equilibrium state, $\boldsymbol{x}^{\rm eq}$, in the interior of the phase space $\boldsymbol{X}$. Furthermore, we assume that there exists a strictly convex function $G(\boldsymbol{x})$ which decreases monotonically in time due to (8):

$$\dot{G} = (\boldsymbol{\nabla} G(x), J(x)) \le 0. \tag{11}$$

Here $\boldsymbol{\nabla} G$ is the vector of partial derivatives $\partial G/\partial x_i$, and the convexity assumes that the $n \times n$ matrices

$$\boldsymbol{H}_{\boldsymbol{x}} = \|\partial^2 G(\boldsymbol{x})/\partial x_i \partial x_j\|, \tag{12}$$

are positive definite for all $\boldsymbol{x} \in \boldsymbol{X}$. In addition, we assume that the matrices (12) are invertible if $\boldsymbol{x}$ is taken in the interior of the phase space.

The matrix $\boldsymbol{H}$ defines an important Riemann structure on the concentration space, the thermodynamic (or entropic) scalar product:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{c}} = (\boldsymbol{x}, \boldsymbol{H}_{\boldsymbol{x}} \boldsymbol{y}), \tag{13}$$

This choice of the Riemann structure is unambiguous from the thermodynamic perspective. We use this metrics for all geometrical constructions, for measuring angles and distances in the phase space U.

The function G is the Lyapunov function of the system (5), and $\boldsymbol{x}^{\rm eq}$ is the point of global minimum of the function G in the phase space $\boldsymbol{X}$. Otherwise stated, the manifold of equilibrium states $\boldsymbol{x}^{\rm eq}(B_1, \dots, B_l)$ is the solution to the variational problem,

$$G \to \min \text{ for } (\boldsymbol{b}_i, \boldsymbol{x}) = B_i,\ i = 1, \dots, l. \tag{14}$$

For each fixed value of the conserved quantities B_i, the solution is unique.

For perfect systems in a constant volume under a constant temperature, the Lyapunov function G reads:

$$G = \sum_{i=1}^{n} x_i [\ln(x_i / x_i^{\rm eq}) - 1]. \tag{15}$$

2.2 Thermodynamic projector

For dissipative systems, we keep in mind the following picture (Fig. 2). The vector field $J(x)$ generates the motion on the phase space U: $dx/dt = J(x)$. An ansatz manifold Ω is given, it is the current approximation to the invariant manifold. This manifold Ω is described as the image of the map $F : W \to U$, where W is a space of macroscopic variables, U is our phase space.

The projected vector field $PJ(x)$ belongs to the tangent space T_x, and the equation $dx/dt = PJ(x)$ describes the motion along the ansatz manifold Ω (if the initial state belongs to Ω). The induced dynamics on the space W is generated by the vector field

$$\frac{dy}{dt} = (D_y F)^{-1} PJ(F(y)).$$

Here the inverse linear operator $(D_y F)^{-1}$ is defined on the tangent space $T_{F(y)}$, because the map F is assumed to be immersion, that is the differential $(D_y F)$ is the isomorphism onto the tangent space $T_{F(y)}$.

Projection operators P contribute to the invariance equation (3). Limiting results, exact solutions, etc. only weakly depend on the particular choice of projectors, or do not depend on it at all. However, thermodynamical validity of approximations obtained on each iteration step towards the limit strongly depends on the choice of the projector.

Let *some* (not obligatory invariant) manifold Ω is considered as a manifold of reduced description. We should define a field of linear operators, $\boldsymbol{P}_{\boldsymbol{x}}$, labeled by the states $\boldsymbol{x} \in \Omega$, which project the vectors $\boldsymbol{J}(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega$ onto the tangent bundle of the manifold Ω, thereby generating the induced vector field, $\boldsymbol{P}_{\boldsymbol{x}} \boldsymbol{J}(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega$. This induced vector field on the tangent bundle of the manifold Ω is identified with the reduced dynamics along the manifold Ω. The *thermodynamicity* requirement for this induced vector field reads

$$(\boldsymbol{\nabla} G(\boldsymbol{x}), \boldsymbol{P}_{\boldsymbol{x}} \boldsymbol{J}(\boldsymbol{x})) \le 0, \text{ for each } \boldsymbol{x} \in \Omega. \tag{16}$$

The condition (16) means that the entropy (which is the Lyapunov function with minus sign) should increase in the new dynamics (2).

How to construct the projector P? Another form of this question is: how to define the plain of fast motions $x + \ker P$? The choice of the projector P is ambiguous, from the formal point

of view, but the second law of thermodynamics gives a good hint [3]: the entropy should grow in the fast motion, and the point x should be the point of entropy maximum on the plane of fast motion $x + \ker P$. That is, the subspace $\ker P$ should belong to the kernel of the entropy differential:

$$\ker P_x \subset \ker D_x S.$$

Of course, this rule is valid for closed systems with entropy, but it can be also extended onto open systems: the projection of the "thermodynamic part" of $J(x)$ onto T_x should have the positive entropy production. If this thermodynamic requirement is valid for any ansatz manifold not tangent to the entropy levels and for any thermodynamic vector field, then the thermodynamic projector is unique [13]. Let us describe this projector P for a given point x, subspace $T_x = \mathrm{imP}$, differential $D_x S$ of the entropy S at the point x and the second differential of the entropy at the point x, the bilinear functional $(D_x^2 S)_x$. We need the positively definite bilinear form $\langle z|p\rangle_x = -(D_x^2 S)_x(z,p)$ (the entropic scalar product). There exists a unique vector g such that $\langle g|p\rangle_x = D_x S(p)$. It is the Riesz representation of the linear functional $D_x S$ with respect to entropic scalar product. If $g \neq 0$ then the thermodynamic projector is

$$P(J) = P^{\perp}(J) + \frac{g^{\|}}{\langle g^{\|}|g^{\|}\rangle_x}\langle g^{\perp}|J\rangle_x, \tag{17}$$

where $P^{\perp}$ is the orthogonal projector onto T_x with respect to the entropic scalar product, and the vector g is splitted onto tangent and orthogonal components:

$$g = g^{\|} + g^{\perp};\ g^{\|} = P^{\perp}g;\ g^{\perp} = (1 - P^{\perp})g.$$

This projector is defined if $g^{\|} \neq 0$.

If $g = 0$ (the equilibrium point) then $P(J) = P^{\perp}(J)$.

For given T_x, the *thermodynamic projector* (17) depends on the point x through the x-dependence of the scalar product $\langle|\rangle_x$, and also through the differential of S in x.

2.3 Symmetric linearization

The invariance condition (3) supports a lot of invariant manifolds, and not all of them are relevant to the reduced description (for example, any individual trajectory is itself an invariant manifold). This should be carefully taken into account when deriving a relevant equation for the correction in the states of the initial manifold Ω_0 which are located far from equilibrium. This point concerns the procedure of the linearization of the vector field $\boldsymbol{J}$, appearing in the equation (1). Let $\boldsymbol{c}$ is an arbitrary point of the phase space. The linearization of the vector function $\boldsymbol{J}$ about $\boldsymbol{c}$ may be written $J(\boldsymbol{c} + \delta\boldsymbol{c}) \approx \boldsymbol{J}(\boldsymbol{c}) + \boldsymbol{L_c}\delta\boldsymbol{c}$ where the linear operator $\boldsymbol{L_c}$ acts as follows (for the mass action law):

$$\boldsymbol{L_c x} = \sum_{s=1}^{r} \boldsymbol{\gamma}_s [W_s^+(\boldsymbol{c})(\boldsymbol{\alpha}_s, \boldsymbol{H_c x}) - W_s^-(\boldsymbol{c})(\boldsymbol{\beta}_s, \boldsymbol{H_c x})]. \tag{18}$$

Here $\boldsymbol{H_c}$ is the matrix of second derivatives of the function G in the state $\boldsymbol{c}$, see (12). The matrix $\boldsymbol{L_c}$ in (18) can be decomposed as follows:

$$\boldsymbol{L_c} = \boldsymbol{L'_c} + \boldsymbol{L''_c}. \tag{19}$$

Matrices $\boldsymbol{L'_c}$ and $\boldsymbol{L''_c}$ act as follows:

$$\boldsymbol{L'_c x} = -\frac{1}{2}\sum_{s=1}^{r} [W_s^+(\boldsymbol{c}) + W_s^-(\boldsymbol{c})]\boldsymbol{\gamma}_s(\boldsymbol{\gamma}_s, \boldsymbol{H_c x}), \tag{20}$$

$$\boldsymbol{L''_c x} = \frac{1}{2}\sum_{s=1}^{r} [W_s^+(\boldsymbol{c}) - W_s^-(\boldsymbol{c})]\boldsymbol{\gamma}_s(\boldsymbol{\alpha}_s + \boldsymbol{\beta}_s, \boldsymbol{H_c x}). \tag{21}$$

Some features of this decomposition are best seen when we use the thermodynamic scalar product (13): The following properties of the matrix $\boldsymbol{L}'_{\boldsymbol{c}}$ are verified immediately:

(i) The matrix $\boldsymbol{L}'_{\boldsymbol{c}}$ is symmetric in the scalar product (13):

$$\langle \boldsymbol{x}, \boldsymbol{L}'_{\boldsymbol{c}}\boldsymbol{y}\rangle = \langle \boldsymbol{y}, \boldsymbol{L}'_{\boldsymbol{c}}\boldsymbol{x}\rangle. \tag{22}$$

(ii) The matrix $\boldsymbol{L}'_{\boldsymbol{c}}$ is nonpositive definite in the scalar product (13):

$$\langle \boldsymbol{x}, \boldsymbol{L}'_{\boldsymbol{c}}\boldsymbol{x}\rangle \leq 0. \tag{23}$$

(iii) The null space of the matrix $\boldsymbol{L}'_{\boldsymbol{c}}$ is the linear envelope of the vectors $\boldsymbol{H}_{\boldsymbol{c}}^{-1}\boldsymbol{b}_i$ representing the complete system of conservation laws:

$$\ker \boldsymbol{L}'_{\boldsymbol{c}} = \mathrm{Lin}\{\boldsymbol{H}_{\boldsymbol{c}}^{-1}\boldsymbol{b}_i, i = 1, \ldots, l\} \tag{24}$$

(iv) If $\boldsymbol{c} = \boldsymbol{c}^{\mathrm{eq}}$, then $W_s^+(\boldsymbol{c}^{\mathrm{eq}}) = W_s^-(\boldsymbol{c}^{\mathrm{eq}})$, and

$$\boldsymbol{L}'_{\boldsymbol{c}^{\mathrm{eq}}} = \boldsymbol{L}_{\boldsymbol{c}^{\mathrm{eq}}}. \tag{25}$$

Thus, the decomposition (19) splits the matrix $\boldsymbol{L}_{\boldsymbol{c}}$ in two parts: one part, (20) is symmetric and nonpositive definite, while the other part, (21), vanishes in the equilibrium. The decomposition (19) explicitly takes into account the mass-action law. For other dissipative systems, the decomposition (19) is possible as soon as the relevant kinetic operator is written in a gain–loss form.

3 Invariant grids

In most of the works (of us and of other people on similar problems), analytic forms were required to represent manifolds (see, however, the method of Legendre integrators [14, 15, 16]). However, in order to construct manifolds of a relatively low dimension, grid-based representations of manifolds become a relevant option [8].

The main idea of the method of invariant grids (MIG) is to find a mapping of the finite-dimensional grids into the phase space of a dynamic system. That is, we construct not just a point approximation of the invariant manifold $F^*(y)$, but an *invariant grid*. When refined, it is expected to converge, of course, to $F^*(y)$, but in any case it is a separate, independently defined object.

Let's denote $L = R^n$, G is a discrete subset of R^n. It is natural to think of a regular grid, but this is not so crucial. For every point $y \in G$, a neighborhood of y is defined: $V_y \subset G$, where V_y is a finite set, and, in particular, $y \in V_y$. On regular grids, V_y includes, as a rule, the nearest neighbors of y. It may also include the points next to the nearest neighbors.

For our purpose, we should define a grid differential operator. For every function, defined on the grid, also all derivatives are defined:

$$\left.\frac{\partial f}{\partial y_i}\right|_{y\in G} = \sum_{z\in V_y} q_i(z,y)f(z), i = 1, \ldots n. \tag{26}$$

where $q_i(z, y)$ are some coefficients.

Here we do not specify the choice of the functions $q_i(z, y)$. We just mention in passing that, as a rule, equation (26) is established using some approximation of f in the neighborhood of y in R^n by some differentiable functions (for example, polynomials). This approximation is based on the values of f at the points of V_y. For regular grids, $q_i(z, y)$ are functions of the difference $z-y$. For some of the nodes y which are close to the edges of the grid, functions are defined only

on the part of V_y. In this case, the coefficients in (26) should be modified appropriately in order to provide an approximation using available values of f. Below we assume this modification is always done. We also assume that the number of points in the neighborhood V_y is always sufficient to make the approximation possible. This assumption restricts the choice of the grids G. Let's call *admissible* all such subsets G, on which one can define differentiation operator in every point.

Let F be a given mapping of some admissible subset $G \subset R^n$ into U. For every $y \in V$ we define tangent vectors:

$$T_y = Lin\{g_i\}_1^n, \tag{27}$$

where vectors $g_i (i = 1, \dots n)$ are partial derivatives (26) of the vector-function F:

$$g_i = \frac{\partial F}{\partial y_i} = \sum_{z \in V_y} q_i(z, y) F(z), \tag{28}$$

or in the coordinate form:

$$(g_i)_j = \frac{\partial F_j}{\partial y_i} = \sum_{z \in V_y} q_i(z, y) F_j(z). \tag{29}$$

Here $(g_i)_j$ is the jth coordinate of the vector (g_i), and $F_j(z)$ is the jth coordinate of the point $F(z)$.

The grid G is *invariant*, if for every node $y \in G$ the vector field $J(F(y))$ belongs to the tangent space T_y (here J is the right hand side of the kinetic equations (1)).

So, the definition of the invariant grid includes:

1. The finite admissible subset $G \subset R^n$;
2. A mapping F of this admissible subset G into U (where U is the phase space of kinetic equation (1));
3. The differentiation formulas (26) with given coefficients $q_i(z, y)$;

The *grid invariance equation* has a form of an inclusion:

$$J(F(y)) \in T_y \text{ for every } y \in G,$$

or a form of an equation:

$$(1 - P_y) J(F(y)) = 0 \text{ for every } y \in G,$$

where P_y is the thermodynamic projector (17).

The grid differentiation formulas (26) are needed, in the first place, to establish the tangent space T_y, and the null space of the thermodynamic projector P_y in each node. It is important to realize that the locality of the construction of the thermodynamic projector enables this without a global parametrization.

Let $x = F(y)$ be the location of the grid's node y immersed into U. We have the set of tangent vectors $g_i(x)$, defined in x (28), (29). Thus, the tangent space T_y is defined by (27). Also, one has the entropy function $S(x)$, the linear functional $D_x S|_x$, and the subspace $T_{0y} = T_y \bigcap \ker D_x S|_x$ in T_y. Let $T_{0y} \neq T_y$. In this case we have a vector $\boldsymbol{e}_y \in T_y$, orthogonal to T_{0y}, $D_x S|_x(\boldsymbol{e}_y) = 1$. Then the thermodynamic projector is defined as:

$$P_y \bullet = P_{0y} \bullet + \boldsymbol{e}_y D_x S|_x \bullet, \tag{30}$$

where P_{0y} is the orthogonal projector on T_{0y} with respect to the entropic scalar product $\langle | \rangle_x$.

If $T_{0y} = T_y$, then the thermodynamic projector is the orthogonal projector on T_y with respect to the entropic scalar product $\langle|\rangle_x$.

The general schema of solving the invariance equation (3) to optimize positions of the invariant grid nodes in space is the following:

0) The grid is initialized. For example, one can use spectral decomposition of $(D_x^2 S)_x$ in the equilibrium;

1) Given some node positions, one calculates the tangent vectors in every node of the grid (27), at this stage the connectivity between nodes is used;

2) With set of tangent vectors calculated at the previous step, solve the invariance equation for every node *independently* and calculate a shift δy of every node in the phase space; we propose two algorithms to calculate the shift: *the Newton method with incomplete linearization* and *the relaxation method* (see also [6],[8], [5], [4]).

3) Repeat steps 1) and 2) until some convergence criterion will be fulfilled: for example, all shifts $\delta y_i, i = 1..n$ will be less than a predefined ϵ_{conv}.

4) Update the structure of the grid: for example, add new nodes and extend (extrapolate) or refine (interpolate) the grid. Some strategies for this are described further;

5) Repeat steps 1)-4) until some criterion will be fulfilled: typically, when the nodes reach the phase space boundary or the spectral gap is too small (see further).

The idea of the Newton method with incomplete linearization is to use linear approximation of J in the vicinity of a grid node y (keeping the projector P fixed). At the same time the node is shifted in the fast direction (in $y + \ker P_y$ affine subspace).

For the Newton method with incomplete linearization, the equations for calculation the new node location $y' = y + \delta y$ are:

$$\begin{cases} P_y \delta y = 0 \\ (1 - P_y)(J(y) + DJ(y)\delta y) = 0. \end{cases} \tag{31}$$

Here $DJ(y)$ is a matrix of derivatives of J evaluated at y. Instead of $DJ(y)$ (especially in the regions that are far from the equilibrium) one can use the symmetric operator $\boldsymbol{L}'(y)$ (20), this will provide better convergence towards the "true" invariant manifold.

Equation (31) is a system of linear algebraic equations. In practice, it proves convenient to choose some orthonormal (with respect to the entropic scalar product) basis $\boldsymbol{b}_i$ in $\ker P_y$. Let $r = \dim(\ker P_y)$. Then $\delta y = \sum_{i=1}^{r} \delta_i \boldsymbol{b}_i$, and system (31) takes the form

$$\sum_{k=1}^{r} \delta_k \langle \boldsymbol{b}_i \mid DJ(y)\boldsymbol{b}_k \rangle_y = -\langle J(y) \mid \boldsymbol{b}_i \rangle_y, i = 1...r. \tag{32}$$

This is the system of linear equations for adjusting the node location according to the Newton method with incomplete linearization. We remind once again that one should use the entropic scalar products.

For the relaxation method, one needs to calculate the defect $\Delta_y = (1 - P_y)J(y)$, and the relaxation step

$$\tau(y) = -\frac{\langle \Delta_y | \Delta_y \rangle_y}{\langle \Delta_y | DJ(y)\Delta_y \rangle_y}. \tag{33}$$

Then, the new node location y' is computed as

$$y' = y + \tau(y)\Delta_y. \tag{34}$$

This is the equation for adjusting the node location according to the relaxation method.

4 Grid construction strategy

From all the reasonable strategies of the invariant grid construction we consider here the following two: the *growing lump* and the *invariant flag*.

4.1 Growing lump

The construction is initialized from the equilibrium point y^*. The first approximation is constructed as $F(y^*) = x^*$, and for some initial V_0 ($V_{y^*} \subset V_0$) one has $F(y) = x^* + A(y - y^*)$, where A is an isometric embedding (in the standard Euclidean metrics) of R^n in E.

For this initial grid one makes a fixed number of iterations of one of the methods chosen (Newton's method with incomplete linearization or the relaxation method), and, after that, puts $V_1 = \bigcup_{y \in V_0} V_y$ and extends F from V_0 onto V_1 using the linear extrapolation, and the process continues. One of the possible variants of this procedure is to extend the grid from V_i to V_{i+1} not after a fixed number of iterations, but only after the invariance defect Δ_y becomes less than a given ϵ (in a given norm, which is entropic, as a rule), for all nodes $y \in V_i$. The lump stops growing after it reaches the boundary and is within a given accuracy $\|\Delta\| < \epsilon$.

4.2 Invariant flag

In order to construct the invariant flag one uses sufficiently regular grids G, in which many points are located on the coordinate lines, planes, etc. One considers the standard flag $R^0 \subset R^1 \subset R^2 \subset ... \subset R^n$ (every next space is constructed by adding one more coordinate). It corresponds to a sequence of grids $\{y^*\} \subset G^1 \subset G^2 ... \subset G^n$, where $\{y^*\} = R^0$, and G^i is a grid in R^i.

First, y^* is mapped on x^* and further $F(y^*) = x^*$. Then the invariant grid is constructed on $V^1 \subset G^1$ (up to the boundaries and within a given accuracy $\|\Delta\| < \epsilon$). After that, the neighborhoods in G^2 are added to the points V^1, and the grid $V^2 \subset G^2$ is constructed (up to the boundaries and within a given accuracy) and so on, until $V^n \subset G^n$ is constructed.

While constructing the kth-order grid $V^k \subset G^k$, the important role of the grids of lower dimension $V^0 \subset ... \subset V^{k-1} \subset V^k$ embedded in it, is preserved. The point $F(y^*) = x^*$ (equilibrium) remains fixed. For every $y \in V^q$ ($q < k$) the tangent vectors $g_1, ..., g_q$ are constructed, using the differentiation operators (26) on the whole V^k. Using the tangent space $T_y = Lin\{g_1, .., g_q\}$, the projector P_y is constructed, the iterations are applied and so on. All this is done in order to obtain a sequence of embedded invariant grids, given by the same map F.

4.3 Boundaries check and the entropy

We construct grid mapping of F onto a finite set $V \in G$. The technique of checking whether the grid still belongs to the phase space U of the kinetic system ($F(V) \subset U$) is quite straightforward: all the points $y \in V$ are checked whether they belong to U. If at the next iteration a point $F(y)$ leaves U, then it is pulled inside by a homothety transform with the center in x^*. Since the entropy is a concave function, the homothety contraction with the center in x^* increases the entropy monotonically. Another variant to cut off the points which leave U.

By construction (17), the kernel of the entropic projector is annulled by the entropy differential. Thus, in the first order, the steps in the Newton method with incomplete linearization (31) as well as in the relaxation method (33) do not change the entropy. But if the steps are quite large, then the increase of the entropy may become essential, and the points are returned on their entropy levels by the homothety contraction with the center in the equilibrium point.

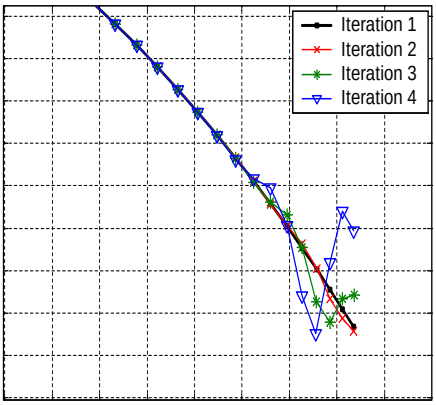


Figure 3: Grid instability. For small grid steps approximations in the calculation of grid derivatives lead to the grid instability effect. Several successive iterations of the algorithm without adaptation of the time step are shown that lead to undesirable "oscillations", which eventually destroy the grid starting from one of its ends.

5 Instability of fine grids

When one reduces the grid spacing in order to refine the grid, then, once the grid spacing becomes small enough, one can face the problem of the *Courant instability* [17, 18, 19]. Instead of converging, at every iteration the grid becomes more and more entangled (see Fig. 3).

A way to avoid such instability is well-known. This is decreasing the time step. In our problem, instead of a true time step, we have a shift in the Newtonian direction. Formally, we can assign the value $h = 1$ for one complete step in the Newtonian direction. Let us extend now the Newton method to arbitrary h. For this, let us find $\delta x = \delta F(y)$ from (31), but update δx proportionally to h; the new value of $x_{n+1} = F_{n+1}(y)$ is equal to

$$F_{n+1}(y) = F_n(y) + h_n \delta F_n(y) \tag{35}$$

where n denotes the number of iteration.

One way to choose the step value h is to make it adaptive, by controlling the average value of the invariance defect $\|\Delta_y\|$ at every step. Another way is the convergence control: then $\sum h_n$ plays a role of time.

Elimination of the Courant instability for the relaxation method can be done quite analogously. Everywhere the step h is maintained as large as it is possible without running into convergence problems.

6 Analyticity and effect of superresolution.

When constructing invariant grids, one must define the differential operators (26) for every grid's node. For calculating the differential operators in some point y, an interpolation procedure in the neighborhood of y is used. As a rule, it is an interpolation by a low-order polynomial, which is constructed using the function values in the nodes belonging to the neighbourhood of y in G. This approximation (using values in the nearest neighborhood nodes) is natural for smooth functions. But we are looking for the *analytical* invariant manifold. Analytical functions have a much more "rigid" structure than the smooth ones. One can change a smooth function in the

neighborhood of any point in such a way, that outside this neighborhood the function will not change. In general, this is not possible for analytical functions: a kind of a "long-range" effect takes place (as is well known) .

The idea is to make use of this effect and to reconstruct some analytical function f_G using a function given on G. There is one important requirement: if the values given on G are values of some function f which is analytical in a neighborhood U, then, if the G is refined "correctly", one must have $f_G \to f$ in U. The sequence of reconstructed function f_G should converge to the "right" function f.

What is the "correct refinement"? For smooth functions for the convergence $f_G \to f$ it is necessary and sufficient that, in the course of refinement, G would approximate the whole U with arbitrary accuracy. For analytical functions it is necessary only that, under the refinement, G would approximate some uniqueness set $A \subset U$. A subset $A \subset U$ is called *uniqueness set* in U if for analytical in U functions ψ and φ from $\psi|_A \equiv \varphi|_A$ it follows $\psi \equiv \varphi$. Suppose we have a sequence of grids G, each next is finer than the previous, which approximate a set A. For smooth functions using function values defined on the grids one can reconstruct the function in A. For analytical functions, if the analyticity domain U is known, and A is a uniqueness set in U, then one can reconstruct the function in U. The set U can be essentially bigger than A; because of this such extension was named as *superresolution effect* [20]. There exist formulas for construction of analytical functions f_G for different domains U, uniqueness sets $A \subset U$ and for different ways of discrete approximation of A by a sequence of refined grids G [20]. Here we provide only one Carleman's formula which is the most appropriate for our purposes.

Let domain $U = Q_\sigma^n \subset C^n$ be a product of strips $Q_\sigma \subset C$, $Q_\sigma = \{z|\mathrm{Im} z < \sigma\}$. We shall construct functions holomorphic in Q_σ^n. This is effectively equivalent to the construction of real analytical functions f in the whole R^n with a condition on the convergence radius $r(x)$ of the Taylor series for f as a function of each coordinate: $r(x) \geq \sigma$ in every point $x \in R^n$.

The sequence of refined grids is constructed as follows: let for every $l = 1, ..., n$ a finite sequence of distinct points $N_l \subset Q_\sigma$ be defined:

$$N_l = \{x_{lj}|j = 1, 2, 3...\}, x_{lj} \neq x_{li} \ \ for \ \ i \neq j \tag{36}$$

The countable uniqueness set A, which is approximated by a sequence of refined grids, has the form:

$$A = N_1 \times N_2 \times ... \times N_n = \{(x_{1i_1}, x_{2i_2}, .., x_{ni_n})|i_{1,..,n} = 1, 2, 3, ...\} \tag{37}$$

The grid G_m is defined as the product of initial fragments N_l of length m:

$$G_m = \{(x_{1i_1}, x_{2i_2}...x_{ni_n})|1 \leq i_{1,..,n} \leq m\} \tag{38}$$

Let us denote $\lambda = 2\sigma/\pi$ (σ is a half-width of the strip Q_σ). The key role in the construction of the Carleman's formula is played by the functional $\omega_m^\lambda(u, p, l)$ of 3 variables: $u \in U = Q_\sigma^n$, p is an integer, $1 \leq p \leq m$, l is an integer, $1 \leq p \leq n$. Further u will be the coordinate value at the point where the extrapolation is calculated, l will be the coordinate number, and p will be an element of multi-index $\{i_1, ..., i_n\}$ for the point $(x_{1i_1}, x_{2i_2}, ..., x_{ni_n}) \in G$:

$$\begin{aligned} \omega_m^\lambda(u, p, l) &= \frac{(e^{\lambda x_{lp}} + e^{\lambda \bar{x}_{lp}})(e^{\lambda u} - e^{\lambda x_{lp}})}{\lambda(e^{\lambda u} + e^{\lambda \bar{x}_{lp}})(u - x_{lp})e^{\lambda x_{lp}}} \\ &\times \prod_{j=1 j \neq p}^{m} \frac{(e^{\lambda x_{lp}} + e^{\lambda \bar{x}_{lj}})(e^{\lambda u} - e^{\lambda x_{lj}})}{(e^{\lambda x_{lp}} - e^{\lambda x_{lj}})(e^{\lambda u} + e^{\lambda \bar{x}_{lj}})} \end{aligned} \tag{39}$$

For real-valued x_{pk} formula (39) simplifyes:

$$\omega_m^\lambda(u, p, l) = 2\frac{e^{\lambda u} - e^{\lambda x_{lp}}}{\lambda(e^{\lambda u} + e^{\lambda x_{lp}})(u - x_{lp})} \times \prod_{j=1 j \neq p}^{m} \frac{(e^{\lambda x_{lp}} + e^{\lambda x_{lj}})(e^{\lambda u} - e^{\lambda x_{lj}})}{(e^{\lambda x_{lp}} - e^{\lambda x_{lj}})(e^{\lambda u} + e^{\lambda x_{lj}})} \tag{40}$$

The Carleman formula for extrapolation from G_M on $U = Q^n_\sigma$ ($\sigma = \pi\lambda/2$) has the form ($z = (z_1, ..., z_n)$):

$$f_m(z) = \sum_{k_1,...,k_n=1}^{m} f(x_k) \prod_{j=1}^{n} \omega^\lambda_m(z_j, k_j, j), \tag{41}$$

where $k = k_1, .., k_n$, $x_k = (x_{1k_1}, x_{2k_2}, ..., x_{nk_n})$.

There exists a theorem [20]:

If $f \in H^2(Q^n_\sigma)$, then $f(z) = lim_{m\to\infty} f_m(z)$, where $H^2(Q^n_\sigma)$ is the Hardy class of holomorphic in Q^n_σ functions.

It is useful to present the asymptotics of (41) for large $|\mathrm{Re} z_j|$. For this purpose, we shall consider the asymptotics of (41) for large $|\mathrm{Re} u|$:

$$|\omega^\lambda_m(u, p, l)| = \left| \frac{2}{\lambda u} \prod_{j=1 j\neq p}^{m} \frac{e^{\lambda x_{lp}} + e^{\lambda x_{lj}}}{e^{\lambda x_{lp}} - e^{\lambda x_{lj}}} \right| + o(|\mathrm{Re} u|^{-1}). \tag{42}$$

From the formula (41) one can see that for the finite m and $|\mathrm{Re} z_j| \to \infty$ function $|f_m(z)|$ behaves like $const \cdot \prod_j |z_j|^{-1}$.

This property (zero asymptotics) must be taken into account when using the formula (41). When constructing invariant manifolds $F(W)$, it is natural to use (41) not for the immersion $F(y)$, but for the deviation of $F(y)$ from some analytical ansatz $F_0(y)$ [21, 22, 23].

The analytical ansatz $F_0(y)$ can be obtained using Taylor series, just as in the Lyapunov auxiliary theorem [24]. Another variant is to use Taylor series for the construction of Pade-approximations.

It is natural to use approximations (41) in terms of dual variables as well, since there exists for them (as the examples demonstrate) a simple and effective linear ansatz for the invariant manifold. This is the slow invariant subspace E_{slow} of the operator of linearized system (1) in dual variables at the equilibrium point. This invariant subspace corresponds to the set of "slow" eigenvalues (with small $|\mathrm{Re}\lambda|$, $\mathrm{Re}\lambda < 0$). In the space of concentrations this invariant subspace is the quasiequilibrium manifold. It consists of the maximum entropy points on the affine manifolds of the form $x + E_{\mathrm{fast}}$, where E_{fast} is the "fast" invariant subspace of the operator of the linearized system (1) at the equilibrium point. It corresponds to the "fast" eigenvalues (large $|\mathrm{Re}\lambda|$, $\mathrm{Re}\lambda < 0$).

Carleman's formulas can be useful for the invariant grids construction in two places: first, for the definition of the grid differential operators (26), and second, for the analytical continuation of the manifold from the grid.

7 Example: Two-step catalytic reaction

Let us consider a two-step four-component reaction with one catalyst A_2 (the Michaelis-Menten mechanism, see Fig. 1a):

$$A_1 + A_2 \leftrightarrow A_3 \leftrightarrow A_2 + A_4. \tag{43}$$

We assume the Lyapunov function of the form

$$S = -G = -\sum_{i=1}^{4} c_i[\ln(c_i/c_i^{\mathrm{eq}}) - 1].$$

The kinetic equation for the four-component vector of concentrations, $\boldsymbol{c} = (c_1, c_2, c_3, c_4)$, has the form

$$\dot{\boldsymbol{c}} = \boldsymbol{\gamma}_1 W_1 + \boldsymbol{\gamma}_2 W_2. \tag{44}$$

Here $\boldsymbol{\gamma}_{1,2}$ are stoichiometric vectors,

$$\boldsymbol{\gamma}_1 = (-1,-1,1,0),\ \boldsymbol{\gamma}_2 = (0,1,-1,1), \tag{45}$$

while functions $W_{1,2}$ are reaction rates:

$$W_1 = k_1^+ c_1 c_2 - k_1^- c_3,\ W_2 = k_2^+ c_3 - k_2^- c_2 c_4. \tag{46}$$

Here $k_{1,2}^{\pm}$ are reaction rate constants. The system under consideration has two conservation laws,

$$c_1 + c_3 + c_4 = B_1,\ c_2 + c_3 = B_2, \tag{47}$$

or $\langle \boldsymbol{b}_{1,2}, \boldsymbol{c}\rangle = B_{1,2}$, where $\boldsymbol{b}_1 = (1,0,1,1)$ and $\boldsymbol{b}_1 = (0,1,1,0)$. The nonlinear system (43) is effectively two-dimensional, and we consider a one-dimensional reduced description. For our example, we chosed the following set of parameters:

$$\begin{array}{l} k_1^+ = 0.3,\ k_1^- = 0.15,\ k_2^+ = 0.8,\ k_2^- = 2.0; \\ c_1^{\rm eq} = 0.5,\ c_2^{\rm eq} = 0.1, c_3^{\rm eq} = 0.1,\ c_4^{\rm eq} = 0.4; \\ B_1 = 1.0,\ B_2 = 0.2 \end{array} \tag{48}$$

The one-dimensional invariant grid is shown in Fig. 4 in the (c_1,c_4,c_3) coordinates. The grid was constructed by the growing lump method, as described above. We used Newton iterations to adjust the nodes. The grid was grown up to the boundaries of the phase space.

The grid in this example is a one-dimensional ordered sequence $\{x_1, \ldots, x_n\}$. The grid derivatives for calculating the tangent vectors g were taken as $g(x_i) = (x_{i+1} - x_{i-1})/||x_{i+1} - x_{i-1}||$ for the internal nodes, and $g(x_1) = (x_1 - x_2)/||x_1 - x_2||$, $g(x_n) = (x_n - x_{n-1})/||x_n - x_{n-1}||$ for the grid's boundaries.

Close to the phase space boundaries we had to apply an adaptive algorithm for choosing the time step h: if, after the next growing step (adding new nodes to the grid and after completing $N = 20$ Newtonian steps, the grid did not converged, then we choose a new step size $h_{n+1} = h_n/2$ and recalculate the grid. The final (minimal) value for h was $h \approx 0.001$.

The location of the nodes was parametrized with the entropic distance to the equilibrium point measured in the quadratic metrics given by the matrix $\boldsymbol{H}_c = -||\partial^2 S(\boldsymbol{c})/\partial c_i \partial c_j||$ in the equilibrium $c^{\rm eq}$. It means that every node is located on a sphere in this metrics with a given radius, which increases linearly with number of the node. In this figure the step of the increase is chosen to be 0.05. Thus, the first node is at the distance 0.05 from the equilibrium, the second is at the distance 0.10 and so on. Fig. 5 shows several important quantities which facilitate understanding of the object (invariant grid) extracted. The sign on the x-axis of the graphs at Fig. 5 is meaningless since the distance is always positive, but in this situation it indicates two possible directions from the equilibrium point.

Fig. 5a,b represents the slow one-dimensional component of the dynamics of the system. Given any initial condition, the system quickly finds the corresponding point on the manifold and starting from this point the dynamics is given by a part of the graph on the Fig. 5a,b.

One of the useful quantities is shown on the Fig. 5c. It is the relation between the relaxation times "toward" and "along" the grid (λ_2/λ_1, where λ_1,λ_2 are the smallest and the next smallest by absolute value non-zero eigenvalue of the system, symmetrically linearized at the point of the grid node). The figure demonstrates that the system is very stiff close to the equilibrium point (λ_1 and λ_2 are well separated from each other), and becomes less stiff (by order of magnitude) near the boundary. This leads to the conclusion that the one-dimensional reduced model is more adequate in the neighborhood of the equilibrium where fast and slow motions are separated by two orders of magnitude. On the end–points of the grid the one-dimensional reduction ceases to be well-defined.

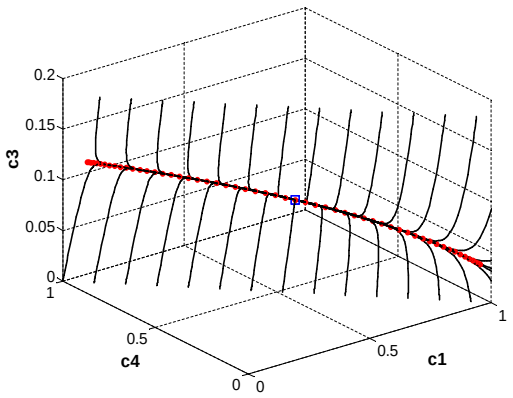


Figure 4: One-dimensional invariant grid (circles) for the two-dimensional chemical system. Projection into the 3d-space of c_1, c_4, c_3 concentrations. The trajectories of the system in the phase space are shown by lines. The equilibrium point is marked by the square. The system quickly reaches the grid and further moves along it.

8 Example: Model hydrogen burning reaction

In this section we consider a more complicated example (see Fig. 1b), where the concentration space is 6-dimensional, while the system is 4-dimensional. We construct an invariant flag which consists of 1- and 2-dimensional invariant manifolds.

We consider a chemical system with six species called H_2 (hydrogen), O_2 (oxygen), H_2O (water), H, O, OH (radicals), see Fig. 1. We assume the Lyapunov function of the form $S = -G = -\sum_{i=1}^{6} c_i[\ln(c_i/c_i^{\rm eq}) - 1]$. The subset of the hydrogen burning reaction and corresponding (direct) rate constants have were taken as:

$$\begin{array}{ll} 1.\ H_2 \leftrightarrow 2H & k_1^+ = 2 \\ 2.\ O_2 \leftrightarrow 2O & k_2^+ = 1 \\ 3.\ H_2O \leftrightarrow H + OH & k_3^+ = 1 \\ 4.\ H_2 + O \leftrightarrow H + OH & k_4^+ = 10^3 \\ 5.\ O_2 + H \leftrightarrow O + OH & k_5^+ = 10^3 \\ 6.\ H_2 + O \leftrightarrow H_2O & k_6^+ = 10^2 \end{array} \tag{49}$$

The conservation laws are:

$$\begin{array}{l} 2c_{H_2} + 2c_{H_2O} + c_H + c_{OH} = b_H \\ 2c_{O_2} + c_{H2O} + c_O + c_{OH} = b_O \end{array} \tag{50}$$

For parameter values we took $b_H = 2$, $b_O = 1$, and the equilibrium point:

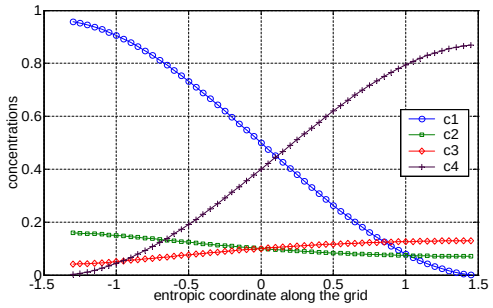
$$c_{H_2}^{\rm eq} = 0.27 \quad c_{O_2}^{\rm eq} = 0.135 \quad c_{H_2O}^{\rm eq} = 0.7 \quad c_H^{\rm eq} = 0.05 \quad c_O^{\rm eq} = 0.02 \quad c_{OH}^{\rm eq} = 0.01 \tag{51}$$

Other rate constants $k_i^-, i = 1..6$ were calculated from $\boldsymbol{c}^{\rm eq}$ value and k_i^+. For this system the stoichiometric vectors are:

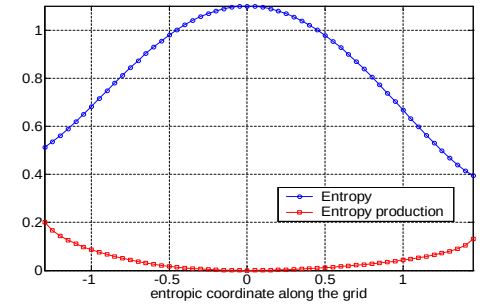
$$\begin{array}{ll} \boldsymbol{\gamma}_1 = (-1,0,0,2,0,0) & \boldsymbol{\gamma}_2 = (0,-1,0,0,2,0) \\ \boldsymbol{\gamma}_3 = (0,0,-1,1,0,1) & \boldsymbol{\gamma}_4 = (-1,0,0,1,-1,1) \\ \boldsymbol{\gamma}_5 = (0,-1,0,-1,1,1) & \boldsymbol{\gamma}_6 = (-1,0,1,0,-1,0) \end{array} \tag{52}$$

The system under consideration is fictitious in the sense that the subset of equations corresponds to the simplified picture of this chemical process and the rate constants do not correspond

a)



b)



c)

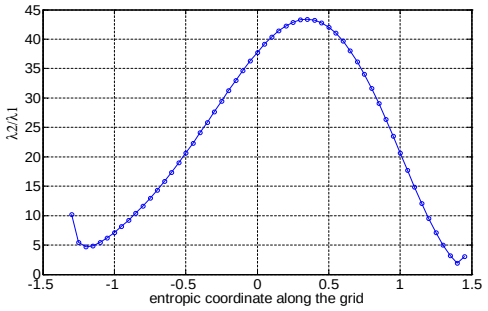


Figure 5: One-dimensional invariant grid for the two-dimensional chemical system. a) Values of the concentrations along the grid. b) Values of the entropy and the entropy production (-dG/dt) along the grid. c) Ratio of the relaxation times "towards" and "along" the manifold. The nodes positions are parametrized with entropic distance measured in the quadratic metrics given by $\boldsymbol{H}_c = -||\partial^2 S(\boldsymbol{c})/\partial c_i \partial c_j||$ in the equilibrium c^{eq}. Entropic coordinate equal to zero corresponds to the equilibrium.

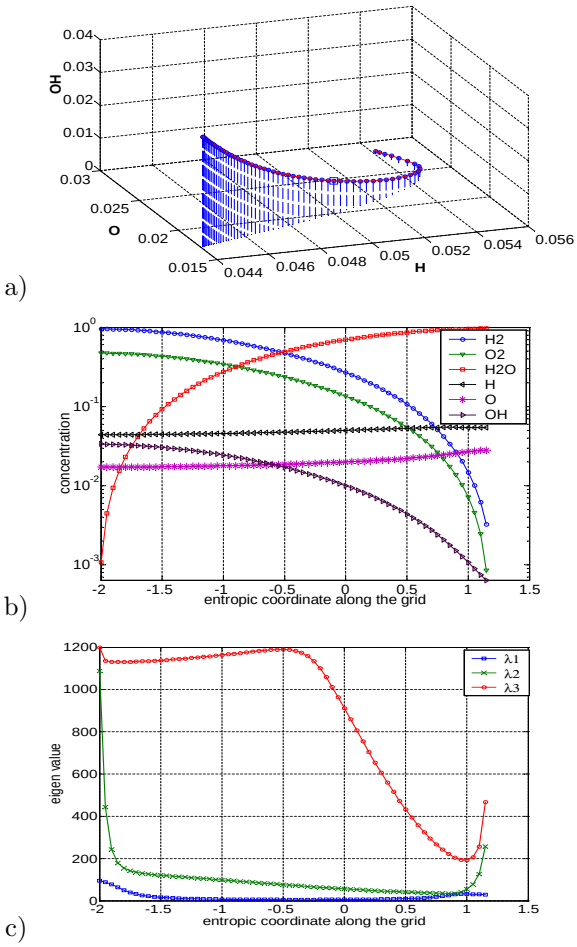


Figure 6: One-dimensional invariant grid for model hydrogen burning reaction. a) Projection into the 3d-space of c_H, c_O, c_{OH} concentrations. b) Concentration values along the grid. c) Three smallest by the absolute value non-zero eigenvalues of the symmetrically linearized system.

to any experimentally measured quantities, rather they reflect only orders of magnitudes relevant real-world systems. In that sense we consider here a qualitative model system, which allows us to illustrate the invariant grids method. Nevertheless, modeling of more realistic systems differs only in the number of species and equations. This leads, of course, to computationally harder problems, but difficulties are not crucial.

Fig. 6a presents a one-dimensional invariant grid constructed for the system. Fig. 6b demonstrates the reduced dynamics along the manifold (for the explanation of the meaning of the x-coordinate, see the previous subsection). In Fig. 6c the three smallest by the absolute value non-zero eigenvalues of the symmetrically linearized Jacobian matrix of the system are shown. One can see that the two smallest eigenvalues almost interchange on one of the grid ends. This means that the one-dimensional "slow" manifold faces definite problems in this region, it is just not well defined there. In practice, it means that one has to use at least a two-dimensional grids there.

Fig. 7a gives a view of the two-dimensional invariant grid, constructed for the system, using the "invariant flag" strategy. The grid was raised starting from the 1D-grid constructed at the previous step. At the first iteration for every node of the initial grid, two nodes (and two

edges) were added. The direction of the step was chosen as the direction of the eigenvector of the matrix A^{sym} (at the point of the node), corresponding to the second "slowest" direction. The value of the step was chosen to be $\epsilon = 0.05$ in terms of entropic distance. After several Newton's iterations done until convergence was reached, new nodes were added in the direction "ortogonal" to the 1D-grid. This time it was done by linear extrapolation of the grid on the same step $\epsilon = 0.05$. Once some new nodes become one or several negative coordinates (the grid reaches the boundaries) they were cut off. If a new node has only one edge, connecting it to the grid, it was excluded (since it was impossible to calculate 2D-tangent space for this node). The process was continued until the expansion was possible (the ultimate state is when every new node had to be cut off).

The method for calculating tangent vectors for this regular rectangular 2D-grid was chosen to be quite simple. The grid consists of *rows*, which are co-oriented by construction to the initial 1D-grid, and *columns* that consist of the adjacent nodes in the neighboring rows. The direction of the columns corresponds to the second slowest direction along the grid. Then, every row and column is considered as a 1D-grid, and the corresponding tangent vectors are calculated as it was described before:

$$g_{row}(x_{k,i}) = (x_{k,i+1} - x_{k,i-1})/\|x_{k,i+1} - x_{k,i-1}\|$$

for the internal nodes and

$$g_{row}(x_{k,1}) = (x_{k,1} - x_{k,2})/\|x_{k,1} - x_{k,2}\|, g_{row}(x_{k,n_k})$$

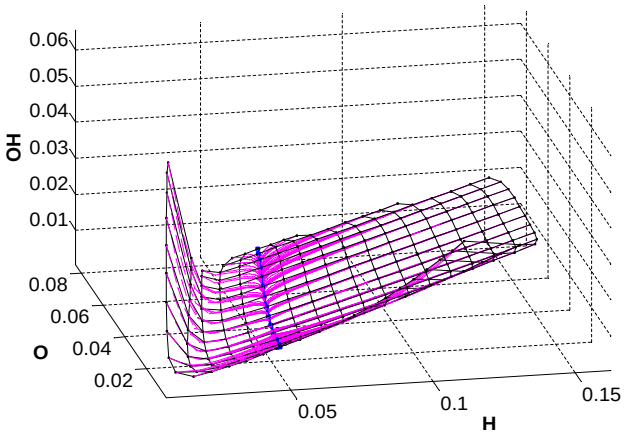
$$= (x_{k,n_k} - x_{k,n_k-1})/\|x_{k,n_k} - x_{k,n_k-1}\|$$

for the nodes which are close to the grid's edges. Here $x_{k,i}$ denotes the vector of the node in the kth row, ith column; n_k is the number of nodes in the kth row. Second tangent vector $g_{col}(x_{k,i})$ is calculated analogously. In practice, it proves convenient to orthogonalize $g_{row}(x_{k,i})$ and $g_{col}(x_{k,i})$.

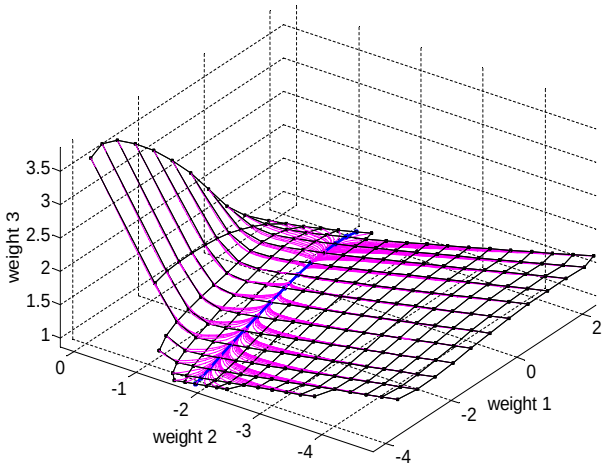
9 Invariant grid as a tool for visualization of dynamic system properties

Usual way of dealing with a system (1) is to define some initial conditions and solve the equation for a given time interval. This gives us one particular trajectory of the system. Can we have a look at the global picture of all possible trajectories or in other words can we visualize the vector field in R^N, defined by $\boldsymbol{J}(x)$? It would be possible if one has two or three species in the system (1). Invariant manifolds and their grid representation allow to do it for higher dimensions, thus they can serve as a data visualization tool. The situation is somewhat close in spirit with data visualization using principal manifolds (for example, see [11]) where one uses two-dimensional manifolds to visualize a finite set of points. Invariant manifolds allow to visualize the global system dynamics on the non-linear manifold of slow motions (i.e., in the space which corresponds to the effects observed in a real-life experiment).

In this section we demonstrate global system dynamics visualization on the model hydrogen burning reaction. Since the phase space is four-dimensional, it is impossible to visualize the grid in one of the coordinate 3D-views, as it was done in the previous subsection. To facilitate visualization one can utilize traditional methods of multi-dimensional data visualization. Here we make use of the principal components analysis (see, for example, [12]), which constructs a three-dimensional linear subspace with maximal dispersion of the othogonally projected data (grid nodes in our case). In other words, the method of principal components constructs in a multi-dimensional space a three-dimensional box such that the grid can be placed maximally



a)



b)

Figure 7: Two-dimensional invariant grid for the model hydrogen burning reaction. a) Projection into the 3d-space of c_H, c_O, c_{OH} concentrations. b) Projection into the principal 3D-subspace. Trajectories of the system are shown coming out from every node. Bold line denotes the one-dimensional invariant grid, starting from which the 2D-grid was constructed.

tightly inside the box (in the mean square distance meaning). After projection of the grid nodes into this space, we get more or less adequate representation of the two-dimensional grid embedded into the six-dimensional concentrations space (Fig. 7b). The disadvantage of the approach is that the axes now do not bear any explicit physical meaning, they are just some linear combinations of the concentrations.

One attractive feature of two-dimensional grids is the possibility to use them as a screen, on which one can display different functions $f(\boldsymbol{c})$ defined in the concentrations space. This technology was exploited widely in the non-linear data analysis by the elastic maps method [10], [11]. The idea is to "unfold" the grid on a plane (to present it in the two-dimensional space, where the nodes form a regular lattice). In other words, we are going to work in the internal coordinates of the grid. In our case, the first internal coordinate (let's call it s_1) corresponds to the direction, co-oriented with the one-dimensional invariant grid, the second one (let us call it s_2) corresponds to the second slow direction. By the construction, the coordinate line $s_2 = 0$ line corresponds to the one-dimensional invariant grid. Units of s_1 and s_2 is the entropic distance.

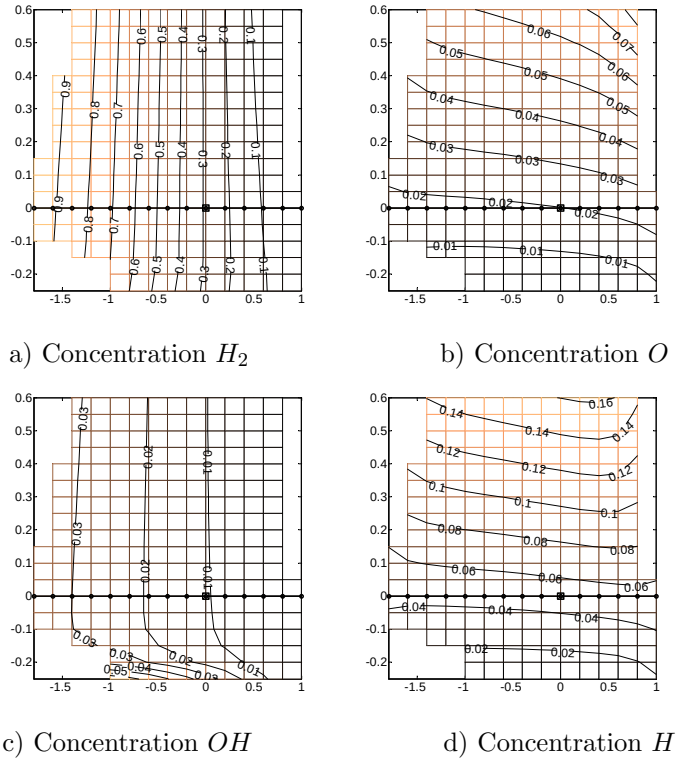
Every grid node has two internal coordinates (s_1, s_2) and, simultaneously, corresponds to a vector in the concentration space. This allows us to map any function $f(\boldsymbol{c})$ from the multi-dimensional concentration space to the two-dimensional space of the grid. This mapping is defined in a finite number of points (grid nodes), and can be interpolated (linearly, in the simplest case) between them. Using *coloring* and *isolines* one can visualize the values of the function in the neighborhood of the invariant manifold. This is meaningful, since, by the definition, the system spends most of the time in the vicinity of the invariant manifold, thus, one can visualize the behavior of the system. As a result of applying this technology, one obtains a set of color illustrations (a stack of information layers), put onto the grid as a map. This enables applying the whole family of the well developed methods of working with the stack of information layers, such as the *geographical information systems* (GIS) methods.

Briefly, this technique of the visualization is a useful tool for understanding of dynamical systems. It allows to see simultaneously many different scenarios of the system behavior, together with different system's characteristics.

Let us use the invariant grids for the the model hydrogen burning system as a screen for visualisation. The simplest functions to visualize are the coordinates: $c_i(\boldsymbol{c}) = c_i$. In Fig. 8 we displayed four colorings, corresponding to the four arbitrarily chosen concentrations functions (of H_2, O, H and OH; Fig. 8a-d). The qualitative conclusion that can be made from the graphs is that, for example, the concentration of H_2 practically does not change during the first fast motion (towards the 1D-grid) and then, gradually changes to the equilibrium value (the H_2 coordinate is "slow"). The O coordinate is the opposite case, it is the "fast" coordinate which changes quickly (on the first stage of the motion) to the almost equilibrium value, and it almost does not change after that. Basically, the slopes of the coordinate isolines give some impression of how "slow" a given concentration is. Fig. 8c shows an interesting behavior of the OH concentration. Close to the 1D grid it behaves like a "slow coordinate", but there is a region on the map where it has a clear "fast" behavior (middle bottom of the graph).

The next two functions which one could wish to visualize are the entropy S and the entropy production $\sigma(\boldsymbol{c}) = -dG/dt(\boldsymbol{c}) = \sum_i \ln(c_i/c_i^{\rm eq})\dot{c}_i$. They are shown on Fig. 9a,b.

Finally, we visualize the relation between the relaxation times of the fast motion towards the 2D-grid and the slow motion along it. This is given on the Fig. 9c. This picture allows to make a conclusion that two-dimensional consideration can be appropriate for the system (especially in the "high H_2, high O" region), since the relaxation times "towards" and "along" the grid are well separated. One can compare this to the Fig. 9d, where the relation between relaxation times towards and along the 1D-grid is shown.



a) Concentration H_2 b) Concentration O

c) Concentration OH d) Concentration H

Figure 8: Two-dimensional invariant grid as a screen for visualizing different functions defined in the concentrations space. The coordinate axes are entropic distances (see the text for the explanations) along the first and the second slowest directions on the grid. The corresponding 1D invariant grid is denoted by bold line, the equilibrium is denoted by square.

10 Invariant manifolds for open systems

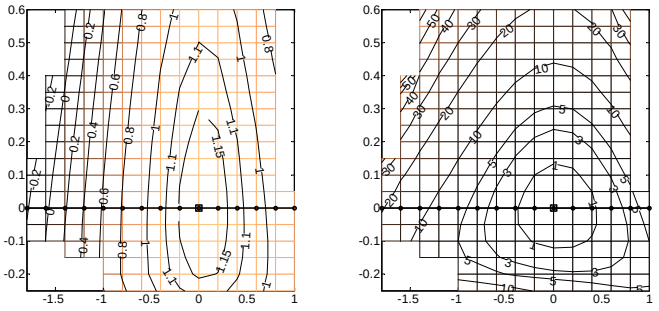
10.1 Zero-order approximation

Let the initial dissipative system (1) be "spoiled" by an additional term ("external vector field" $J_{ex}(x,t)$):

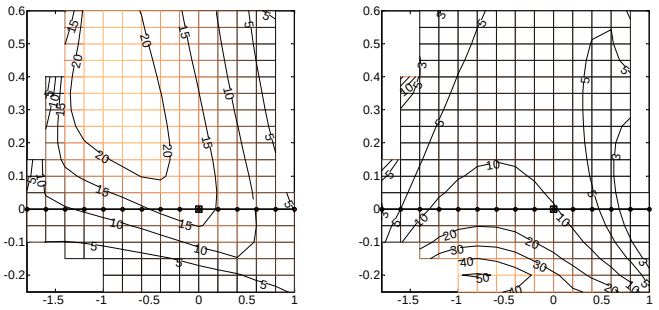
$$\frac{dx}{dt} = J(x) + J_{ex}(x,t), x \subset U. \tag{53}$$

For this new system the entropy does not increase everywhere. In the new system (53) different dynamic effects are possible, such as a non-uniqueness of stationary states, auto-oscillations, etc. The "inertial manifold" effect is well-known: solutions of (53) approach some relatively low-dimensional manifold on which all the non-trivial dynamics takes place [27, 25, 26].

It is natural to expect that the inertial manifold of the system (53) is located somewhere close to the slow manifold of the initial dissipative system (1). This hypothesis has the following basis. Suppose that the vector field $J_{ex}(x,t)$ is sufficiently small. Let's introduce, for example, a small parameter $\varepsilon > 0$, and consider $\varepsilon J_{ex}(x,t)$ instead of $J_{ex}(x,t)$. Let's assume that for the system (1) a separation of motions into "slow" and "fast" takes place. In this case, there exists such interval of positive ε that $\varepsilon J_{ex}(x,t)$ is comparable to J only in a small neighborhood of the given slow motion manifold of the system (1). Outside this neighborhood, $\varepsilon J_{ex}(x,t)$ is negligibly small in comparison with J and only negligibly influences the motion (for this statement to be



a) Entropy b) Entropy Production



c) λ_3/λ_2 relation d) λ_2/λ_1 relation

Figure 9: Two-dimensional invariant grid as a screen for visualizing different functions defined in the concentrations space. The coordinate axes are entropic distances (see the text for the explanations) along the first and the second slowest directions on the grid. The corresponding 1D invariant grid is denoted by bold line, the equilibrium is denoted by square.

true, it is important that the system (1) is dissipative and every solution comes in finite time to a small neighborhood of the given slow manifold).

Precisely this perspective on the system (53) allows to exploit slow invariant manifolds constructed for the dissipative system (1) as the ansatz and the zero-order approximation in a construction of the inertial manifold of the open system (53). In the zero-order approximation, the right part of the equation (53) is simply projected onto the tangent space of the slow manifold.

The choice of the projector is determined by the motion separation which was described above: fast motion is taken from the dissipative system (1). A projector which is suitable for all dissipative systems with given entropy function is unique. It is constructed in the following way. Let a point $x \in U$ be defined and some vector space T, on which one needs to construct a projection (T is the tangent space to the slow manifold at the point x). We introduce the entropic scalar product $\langle|\rangle_x$:

$$\langle a \mid b\rangle_x = -(a, D_x^2 S(b)). \tag{54}$$

Let us consider T_0 that is a subspace of T and which is annulled by the differential S at the point x.

$$T_0 = \{a \in T | D_x S(a) = 0\} \tag{55}$$

If $T_0 = T$, then the thermodynamic projector is the orthogonal projector on T with respect

to the entropic scalar product $\langle|\rangle_x$. Suppose that $T_0 \neq T$. Let $e_g \in T$, $e_g \perp T_0$ with respect to the entropic scalar product $\langle|\rangle_x$, and $D_x S(e_g) = 1$. These conditions define vector e_g uniquely.

The projector onto T is defined by the formula

$$P(J) = P_0(J) + e_g D_x S(J) \tag{56}$$

where P_0 is the orthogonal projector onto T_0 with respect to the entropic scalar product $\langle|\rangle_x$. For example, if T a finite-dimensional space, then the projector (56) is constructed in the following way. Let $e_1, .., e_n$ be a basis in T, and for definiteness, $D_x S(e_1) \neq 0$.
1) Let us construct a system of vectors

$$b_i = e_{i+1} - \lambda_i e_1, (i = 1, .., n-1), \tag{57}$$

where $\lambda_i = D_x S(e_{i+1})/D_x S(e_1)$, and hence $D_x S(b_i) = 0$. Thus, $\{b_i\}_1^{n-1}$ is a basis in T_0.
2) Let us orthogonalize $\{b_i\}_1^{n-1}$ with respect to the entropic scalar product $\langle|\rangle_x$ (1). We thus derived an orthonormal with respect to $\langle|\rangle_x$ basis $\{g_i\}_1^{n-1}$ in T_0.
3) We find $e_g \in T$ from the conditions:

$$\langle e_g \mid g_i\rangle_x = 0, (i = 1, .., n-1), D_x S(e_g) = 1. \tag{58}$$

and, finally we get

$$P(J) = \sum_{i=1}^{n-1} g_i \langle g_i \mid J\rangle_x + e_g D_x S(J). \tag{59}$$

If $D_x S(T) = 0$, then the projector P is simply the orthogonal projector with respect to the $\langle|\rangle_x$ scalar product. This is possible if x is the global maximum of entropy point (equilibrium). Then

$$P(J) = \sum_{i=1}^{n} g_i \langle g_i|J\rangle_x, \langle g_i|g_j\rangle = \delta_{ij}. \tag{60}$$

10.2 First-order approximation

Thermodynamic projector (56) defines a "slow and fast motions" duality: if T is the tangent space of the slow motion manifold then $T = \mathrm{im}P$, and $\ker P$ is the plane of fast motions. Let us denote by P_x the projector at a point x of a given slow manifold.

The vector field $J_{ex}(x,t)$ can be decomposed in two components:

$$J_{ex}(x,t) = P_x J_{ex}(x,t) + (1 - P_x) J_{ex}(x,t). \tag{61}$$

Let us denote $J_{ex\,s} = P_x J_{ex}$, $J_{ex\,f} = (1-P_x)J_{ex}$. The slow component $J_{ex\,s}$ gives a correction to the motion along the slow manifold. This is a zero-order approximation. The "fast" component shifts the slow manifold in the fast motions plane. This shift changes $P_x J_{ex}$ accordingly. Consideration of this effect gives a first-order approximation. In order to find it, let us rewrite the invariance equation taking J_{ex} into account:

$$\begin{cases} (1 - P_x)(J(x + \delta x) + \varepsilon J_{ex}(x,t)) = 0 \\ P_x \delta x = 0 \end{cases} \tag{62}$$

The first iteration of the Newton method subject to incomplete linearization gives:

$$\begin{cases} (1 - P_x)(D_x J(\delta x) + \varepsilon J_{ex}(x,t)) = 0 \\ P_x \delta x = 0. \end{cases} \tag{63}$$

$$(1 - P_x)D_xJ(1 - P_x)J(\delta x) = -\varepsilon J_{ex}(x, t). \tag{64}$$

Thus, we have derived a linear equation in the space $\ker P$. The operator $(1-P)D_xJ(1-P)$ is defined in this space.

Utilization of the self-adjoint linearization instead of the traditional linearization D_xJ operator considerably simplifies solving and studying equation (64). It is necessary to take into account here that the projector P is a sum of the orthogonal projector with respect to the $\langle|\rangle_x$ scalar product and a projector of rank one.

Assume that the first-order approximation equation (64) has been solved and the following function has been found:

$$\delta_1 x(x, \varepsilon J_{ex\,f}) = -[(1 - P_x)D_xJ(1 - P_x)]^{-1}\varepsilon J_{ex\,f}, \tag{65}$$

where D_xJ is either the differential of J or symmetrized differential of J (20).

Let x be a point on the initial slow manifold. At the point $x + \delta x(x, \varepsilon J_{ex\,f})$ the right-hand side of equation (53) in the first-order approximation is given by

$$J(x) + \varepsilon J_{ex}(x, t) + D_xJ(\delta x(x, \varepsilon J_{ex\,f})). \tag{66}$$

Due to the first-order approximation (66), the motion of a point projection onto the manifold is given by the following equation

$$\frac{dx}{dt} = P_x(J(x) + \varepsilon J_{ex}(x, t) + D_xJ(\delta x(x, \varepsilon J_{ex\,f}(x, t)))). \tag{67}$$

Note that, in equation (67), the vector field $J(x)$ enters only in the form of projection, $P_xJ(x)$. For the invariant slow manifold it holds $P_xJ(x) = J(x)$, but actually we always deal with approximately invariant manifolds, hence, it is necessarily to use the projection P_xJ instead of J in (67).

Remark. The notion "projection of a point onto the manifold" needs to be specified. For every point x of the slow invariant manifold M there are defined both the thermodynamic projector P_x (56) and the fast motions plane $\ker P_x$. Let us define a projector Π of some neighborhood of M onto M in the following way:

$$\Pi(z) = x, \quad \text{if } P_x(z - x) = 0. \tag{68}$$

Qualitatively, it means that z, after all fast motions took place, comes into a small neighborhood of x. The operation (56) is defined uniquely in some small neighborhood of the manifold M.

A derivation of slow motions equations requires not only an assumption that εJ_{ex} is small but it must be slow as well: $\frac{d}{dt}(\varepsilon J_{ex})$ must be small too.

One can get the further approximations for slow motions of the system (53), taking into account the time derivatives of J_{ex}. This is an alternative to the usage of the projection operators methods [28].

11 Conclusion

In this paper we presented a method for reducing complexity in complex chemical reaction networks using a consistent approach of constructing invariant manifold for the system of kinetic equations. The method is applicable to the class of dissipative systems (with Lyapounov function) and can be extended to the case of open systems as well.

An attractive feature of the approach is its clear geometrical interpretation. The geometrical approach becomes more and more popular in applied model reduction: one constructs a slow approximate invariant manifold, and dynamical equations on this manifold instead of an approximation of solutions to the initial equations. After that, the equations on the slow manifold can be studied separately, as well as the fast motion to this manifold (the initial layer problem [29]).

The notion of invariant grid may be useful beyond the chemical kinetics. This discrete invariant object can serve as a representation of approximate slow invariant manifold, and as a screen (a map) for vizualization of different functions and properties. The problem of the grid correction is fully decomposed into the problems of the grid's nodes correction which makes it open to effective parallel implementations.

The next step should be the implementation of the method of invariant grids for investigation of high-dimensional systems "kinetics+transport". The asymptotic analysis of the methods of analytic continuation the manifold from the grid should lead to further development of these methods and modifications of the Carleman formula.

12 Acknowledgements

The project is partially supported by Swiss National Science Foundation, Project 200021-107885/1 "Invariant manifolds for model reduction in chemical kinetics" and Swiss Federal Department of Energy (BFE) under the project 100862 "Lattice Boltzmann simulations for chemically reactive systems in a micrometer domain".

References

[1] Hasty J, McMillen D, Isaacs F, Collins J: Computational studies of gene regulatory networks: in numero molecular biology. *Nat Rev Genet* 2001; 4: 268–79.

[2] Endy D, Brent R: Modelling cellular behaviour. *Nature* 2001; 409(6818): 391–5.

[3] Gorban AN, Karlin IV: Thermodynamic parameterization. *Physica A* 1992; 190: 393–404.

[4] Gorban AN, Karlin IV, Zinovyev AYu: Constructive methods of invariant manifolds for kinetic problems. *Phys. Reports* 2004; **396**, 4-6: 197–403. Preprint online: http://arxiv.org/abs/cond-mat/0311017.

[5] Gorban AN, Karlin IV: Method of invariant manifold for chemical kinetics. *Chem. Eng. Sci.* 2003; **58**, 21: 4751–4768. Preprint online: http://arxiv.org/abs/cond-mat/0207231.

[6] Gorban AN, Karlin IV: Invariant Manifolds for Physical and Chemical Kinetics. *Series: Lecture Notes in Physics*, Vol.660. Springer, 2005

[7] Gorban AN, Karlin IV: Methods of nonlinear kinetics. In Encyclopedia of Life Support Systems, Encyclopedia of Mathematical Sciences. EOLSS Publishers, Oxford, 2004. Preprint online: http://arXiv.org/abs/cond-mat/0306062.

[8] Gorban AN, Karlin IV, Zinovyev AYu: Invariant grids for reaction kinetics. *Physica A* 2004; **333**: 106–154. Preprint online: http://www.ihes.fr/PREPRINTS/P03/Resu/resu-P03–42.html.

[9] Gorban AN: Equilibrium encircling. Equations of chemical kinetics and their thermodynamic analysis. Nauka, Novosibirsk, 1984.

[10] Gorban AN, Zinovyev AYu: Visualization of data by method of elastic maps and its applications in genomics, economics and sociology. Preprint of Institut des Hautes Etudes Scientifiques, 2001. Online: http://www.ihes.fr/PREPRINTS/M01/Resu/resu-M01-36.html.

[11] Gorban AN, Zinovyev AYu: Elastic principal graphs and manifolds. *Computing* 2005. In press.

[12] Jolliffe IT: Principal component analysis. Springer–Verlag, 1986.

[13] Gorban AN, Karlin IV: Uniqueness of thermodynamic projector and kinetic basis of molecular individualism. *Physica A* 2004; **336** 3-4: 391–432. Preprint online: http://arxiv.org/abs/cond-mat/0309638.

[14] Gorban AN, Gorban PA, Karlin IV: Legendre integrators, post-processing and quasiequilibrium. *J. Non–Newtonian Fluid Mech.* 2004; **120**: 149–167. Preprint on-line: http://arxiv.org/pdf/cond-mat/0308488.

[15] Ilg P, Karlin IV, Öttinger HC: Canonical distribution functions in polymer dynamics: I. Dilute solutions of flexible polymers. *Physica A* 2002; **315**: 367–385.

[16] Ilg P, Karlin IV, Kröger M, Öttinger HC: Canonical distribution functions in polymer dynamics: II Liquid–crystalline polymers. *Physica A* 2003; **319**: 134–150.

[17] Courant R, Friedrichs KO, Lewy H: On the partial difference equations of mathematical physics. *IBM Journal* (March 1967): 215–234.

[18] Ames WF: Numerical Methods for Partial Differential Equations, 2nd ed. New York, Academic Press, 1977.

[19] Richtmyer RD, Morton KW: Difference methods for initial value problems, 2nd ed. Wiley–Interscience, New York, 1967.

[20] Aizenberg L: Carleman's formulas in complex analysis: Theory and applications. Mathematics and its applications, Vol. 244. Kluwer, 1993.

[21] Gorban AN, Rossiev AA: Neural network iterative method of principal curves for data with gaps. *Journal of Computer and System Sciences International* 1999; **38**, 5: 825–831.

[22] Dergachev VA, Gorban AN, Rossiev AA, Karimova LM, Kuandykov EB, Makarenko NG, Steier P: The filling of gaps in geophysical time series by artificial neural networks. *Radiocarbon* 2001; **43**, 2A: 365 – 371.

[23] Gorban AN, Rossiev A, Makarenko N, Kuandykov Y, Dergachev V: Recovering data gaps through neural network methods. *International Journal of Geomagnetism and Aeronomy* 2002; **3**, 2: 191–197.

[24] Lyapunov AM: The general problem of the stability of motion. London, Taylor & Francis, 1992.

[25] Temam R: Infinite–dimensional dynamical systems in mechanics and physics, ed. 2. Applied Math. Sci., Vol 68. New York, Springer Verlag, 1997.

[26] Constantin P, Foias C, Nicolaenko B, Temam R: Integral manifolds and inertial manifolds for dissipative partial differential equations. Applied Math. Sci., Vol. 70. New York, Springer Verlag, 1988.

[27] Foias C, Sell GR, Temam R: Inertial manifolds for dissipative nonlinear evolution equations. *Journal of Differential Equations* 1988; **73**: 309–353.

[28] Grabert H: Projection operator techniques in nonequilibrium statistical mechanics. Berlin, Springer Verlag, 1982.

[29] Gorban AN, Karlin IV, Zmievskii VB, Nonnenmacher TF: Relaxational trajectories: global approximations. *Physica A* 1996; **231** : 648-672.

Shape spaces in formal interactions

Davide Prandi*[1] , Corrado Priami[1] and Paola Quaglia[1]

Abstract

In recent years formal methods from concurrency theory and process calculi have gained increasing importance in modeling complex biological systems. In this paper propension to biological interaction, as seen by the shape spaces theory, is given a linguistic interpretation. Entities from the living matter are viewed as terms of a formal concurrent language of processes with typed interaction sites. Types are strings, and interaction depends on their distance. Further, the language is associated with syntax-driven rules that permit the inference of the possible computational behaviours of the specified biological system. This approach leads to the use of all the methods and techniques developed in the context of formal languages (e.g. language translation, model checking, ...), opening new ways for studying complex biological systems.

Keywords: shape spaces, Hamming distance, process calculi, mathematical modeling of cellular systems.

[1] {prandi,priami,quaglia}@dit.unitn.it

[1] Dip. Informatica e Telecomunicazioni, Università di Trento, via Sommarive 14, 38050 Povo (TN), Italy

* Corresponding author

1 Introduction

The next century biology research will be strongly influenced by the way in which we will be able to hammer the complexity of systems. After the Human Genome Project we have to face a scaling up of the size of problems. Unfortunately this fast growth in the knowledge is not supported by a corresponding enhancement of the methods and techniques for analysing biological systems. The complexity of the processes we want to model and control is mainly given by the interactions of the constituents of the systems and the consequently emergent behaviours. Therefore the complexity of a biological system is related to the interconnected nature of the problem under analysis. In this context, the classical *reductionist* approach seems no more suitable to handle the current challenges. We need an integrated or systemic view of the investigated phenomena that is hypothesis driven and based on formal/mathematical grounds.

Following these principles biology is moving to the so-called *Systems Biology*. Systems biology is an approach based on systems theory in the applicative domain of biological processes. The basic idea is to view each system as something that has its own behaviour not obtained simply by gluing the behaviour of the systems components of which we already have all the information. We could say that biology is moving towards the organisation of the knowledge acquired with the human genome project and with the high-throughput tools.

The challenge we are now facing is to model, analyse and possibly predict the temporal and space evolution of complex biological systems. The key point seems to find a suitable level of *abstraction* to model the phenomena of interest. On this view we need to build a framework that allows to speed up the understanding of the systems in hand and to exploit the knowledge we are going to discover. Computer science proposed many abstractions to model behaviour and evolution of complex systems over the last decades. We now could adapt such abstraction to the new applicative domains such as e.g. molecular biology or immunology. The basic techniques we should exploit in this strategy are completely different from those used so far in bioinformatics because they lay in the programming languages and modeling field rather than in the classical algorithmic one.

Systems Biology has not to be seen as a "revolution" but rather as a *change* of paradigm. Over the years biologist have understood that they need models for representing and understanding complex biological phenomena. For instance, the wide accepted Gillespie's algorithm [1] is a stochastic model that describes the temporal evolution of biochemical reaction. The programming languages approach allows the integration and organization of different models in an unique picture. Biochemical Stochastic π-calculus [2] integrates Gillespie's algorithm in a programming language leading to the use of computer science theory for analysing biochemical reactions. In this paper we make a step further in this direction enriching Beta-binders [3], a language for describing molecular interactions, with the shape space model [4], a model for representing protein shapes.

The rest of the paper is organized as follows. Section 2 presents process calculi in the context of Systems Biology. In that section we outline a limit of process calculi for Systems Biology and we see how shape spaces can be used for overcoming this limit. Section 3 introduces the shape spaces theory. Next, in Section 4, the Beta-binders language is briefly recalled. For complete formal details, the interested reader is referred to [3]. Here we stick to a graphical and intuitive presentation, and focus on those modifications that allow shape spaces to be natively dealt with in the language. Section 5 concludes the paper with an application of the language to a simple example inspired to the immune system. We show how the phenomenon can be modeled and comment on the behaviour that can be derived by applying the rules of the language. Finally Section 6 concludes the paper and proposes some perspectives.

2 Process Calculi in Systems Biology

Process calculi are formal languages that have been originally developed for modeling distributed systems. They typically allow the abstract description of complex interacting entities in terms of basic parallel components that can either act as stand-alone machines or synchronise and exchange data. Once fixed a language with a limited number of operators, the specification of a system (synonym of *process*) is given as a term that fully defines the way in which the various parts of the system are composed together. They may be sequentialised, (meaning that the operations of one component have to be performed before those of another one), let run in parallel, repeated many times, etc.. The formal language, and hence its sentences, is further associated with syntax-driven rules that permit the inference of the possible computational behaviours of the specified system. Those rules, which can be implemented by an automated tool, allow, e.g., to state that a given process $\mathcal{P}$ transforms into the process $\mathcal{Q}$, written

$$\mathcal{P} \rightarrow \mathcal{Q}.$$

A recent research paper by A. Regev and E. Shapiro points out the analogies between distributed systems and the living matter [5]. Indeed, various description languages in the style of process calculi have already been proposed to model biological behaviours (see, e.g., [2, 6–8]). These languages allow the automatic simulation of all the possible future behaviours of the modeled molecular system, as well as the use of the methods for qualitative and quantitative analysis developed for classical process calculi. The challenge becomes to deeply investigate the relation between biological knowledge and process calculi representation for finding the "best" abstraction. Within this paper we make a step further in this direction grounding process calculi in a well developed biological model.

Classical process calculi assume a *key-lock* model for interactions (think, e.g., of the strict matching between an input and an output over a given channel). Under this assumption only the interaction (a) in Fig 2 is enabled, while (b) is not. This is because "interfaces" of the components 1 and 2 match exactly, while those of 1 and 3 do not. Reactions like the one drawn in (b), however, are quite common in biology [9].

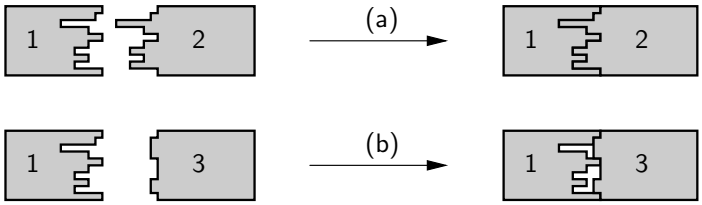


Figure 1: Interaction models

A proposal to relax the key-lock assumption is Beta-binders [3]. In that formalism, processes are encapsulated into boxes with interfaces that are identified by a name and have an associated *type* that represents the interaction capabilities of the box. A type is a set of names, and the interaction is enabled if and only if the types of the interfaces of the two partners are not disjoint. For example $x : \Delta$ and $y : \Gamma$ are two beta binders. The first interface has name x and type Δ and the second one has name y and type Γ. The interaction is allowed if and only if $\Delta \cap \Gamma$ is not empty. This model might be too abstract for a practical biological use, and indeed it was originally chosen by the authors just as a very simple form of typing policy for processes interacting through names. Therefore we introduce here a notion of *affinity* that is finer than

the one expressed by the intersection of the types of interfaces. We formally ground the concept of affinity on *shape spaces* [4], a model introduced in the context of immunology, and we incorporate them into Beta-binders.

3 Shape spaces

Shape spaces [4] were introduced in the context of the theoretical studies of clonal selection in the field of immunology. In this section we generalise the ideas underlying this methodology abstracting as many biological details as possible.

A protein is composed of many different independent structural parts called *domains*. The interaction capability between domains depends on the structural and chemical complementarity of particular portions, called *molecular determinants* or *motifs*. Suppose it is possible to describe the features of a motif by specifying N "shape" parameters. These parameters include geometric quantities which specify the size and the shape of the molecular determinant, and physical characteristics of amino acids comprising the motif (e.g., the charge or the ability to form hydrogen bonds). The N parameters define an N-dimensional vector space that is called *shape space*, say S. A point in S represents a molecular determinant. A function $\mathcal{C} : S \to S$ maps motif shapes to their complements. By defining a metric on S, the distance between two points can be used as a measure of interaction propension between two molecular determinants.

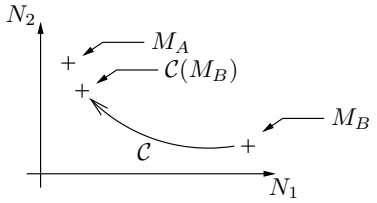


Figure 2: Shape space example

The above intuitions are sketched in Figure 2. We assume that two parameters N_1 and N_2 suffice to describe molecular determinants, leading to a 2-dimensional shape space. Moreover we choose the Euclidean metric, obtaining an *Euclidean* shape space [10].

As an example, let us consider two proteins A and B, with molecular determinants M_A and M_B, respectively. The molecular determinants M_A and M_B are two points in the 2-dimensional Euclidean shape space. The function $\mathcal{C} : N_1 \times N_2 \to N_1 \times N_2$ maps M_B into its complement $\mathcal{C}(M_B)$, and the distance $\varepsilon = \| M_A - \mathcal{C}(M_B) \|$ represents the molecular affinity between the motifs M_A and M_B. In order to get symmetric interaction propension, one can require $\varepsilon = \| M_A - \mathcal{C}(M_B) \| = \| \mathcal{C}(M_A) - M_B \|$.

If the N shape parameters do not contribute equally to the specificity of a motif (e.g. small charge differences could be more important than small differences in geometry), a metric different from the Euclidean one is required. Finding an appropriate metric for measuring the affinity of molecular determinants is a non trivial task in chemistry. The specific choice of the metric, however, does not affect our semantics.

It is computationally difficult to calculate a distance in a high-dimensional continuous space. For this reason the abstract model of the shape spaces is not particularly well suited to a concrete implementation. Computational efficiency is gained by relying on strings and string matching rules to represent the affinity of motifs. Each motif is associated with a string of symbols, and hence a string can be loosely interpreted as

an amino acid sequence. Different symbols represent different values of properties of the amino acids, like, e.g., hydrophobicity or charge.

To effectively compute the interaction propension between motifs it is necessary to define a string matching rule. Choosing the 'right' rule can be hard, and different biological situations might require different matching rules. Indeed quite a bit of distinct rules have been proposed in the literature, like, e.g., the *Hamming* distance and the *Manhattan* distance. The first is given by the number of positions in which two strings differ, while the Manhattan distance between two strings is the sum of the distances between their digits [11]. For example, let us fix the two strings "54" and "84". Their Hamming distance is 1 (they only differ in the leftmost symbol), and their Manhattan distance is 3 (obtained as 8-5+4-4). Yet another definition of distance comes from the so-called *xor rule*. Each symbol in the string is represented as a binary number, and the distance between two strings is computed as the normalised sum of the digits of the xor of the two numbers. This last notion of distance is finer than the Hamming distance, and it is computationally more efficient than the Manhattan rule. The shape spaces that use strings and matching rules are globally called *Hamming spaces*.

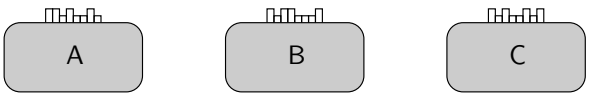


Figure 3: Hamming space example

Figure 3 shows an example of the use of Hamming spaces. Three proteins A, B and C with one motif each are reported in the picture. Each motif is represented by an eight digit string, where each digit can be 1 (drawn as a rectangle) or 0 (drawn as a square). So the motifs of A, B, and C are 11010010, 10110001, and 10100101, respectively. Assume that the function $\mathcal{C} : \{0,1\}^8 \rightarrow \{0,1\}^8$ maps 0s to 1s and 1s to 0s. We adopt the Hamming distance and compare each pair of strings reading the digits from left to right. The mutual interaction propensions of A, B and C are then given by $d_{A\,\mathcal{C}(B)} = 4$, $d_{A\,\mathcal{C}(C)} = 2$ and $d_{B\,\mathcal{C}(C)} = 6$, where $d_{X\,\mathcal{C}(Y)}$ stays for the distance between X and Y. Since the lowest distance value is $d_{A\,\mathcal{C}(C)}$, we can conclude that the two proteins A and C have the greatest interaction propension in the considered set.

For the sake of clarity, in the above example we made two choices that are not quite realistic from a biological point of view. First, we chose to evaluate strings from left to right, while proteins freely float in the living matter and therefore many other different kinds of interaction are possible. A more concrete model would define the value $d_{X\,\mathcal{C}(Y)}$ as the longest stretch of consecutive complementary bits [12]. Second, we assumed a direct map between the distance $d_{X\,\mathcal{C}(Y)}$ and the interaction propension between X and Y. More generally, one would need to define a map between distance and interaction propension [11].

4 Beta-binders graphically

In this section we briefly recall Beta-binders and comment on the generalisation that allows a direct representation of shape spaces. We resort to a graphical presentation of the language and of its rules. The interested reader is referred to [3] for the mathematical details concerning notation and semantics. In the present paper we just point out the single modification that has to be applied to the original formalism to directly render the notion of distance.

Beta-binders builds on the intuition that biological entities have an internal 'process unit' and an 'interface' exposed to the external environment. For example, a protein has a backbone and motifs for interacting

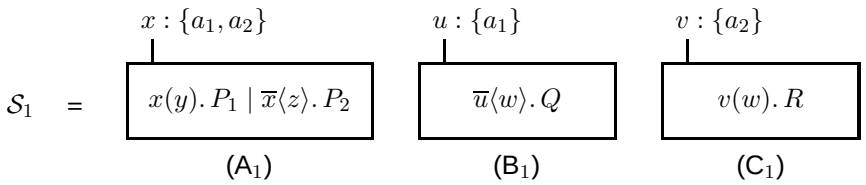
with the environment. A cell has a similar structure: it has a membrane whose proteins act as an interface, and a complex internal structure that responds to the external changes. This interpretation of cell is quite limited if we are studying a single cell, but in the context of the study of cellular populations, e.g. in immunology, this vision is acceptable [13]. Furthermore, the computations internal to cells have a high degree of parallelism and is not surprising that techniques from concurrency theory can be used for representing structural changes of the living matter.

Specifically, the Beta-binders formalism encloses mobile processes [14] into active borders. These borders, that represent the interface of the described entity, are equipped with typed binders which are used for discriminating between allowed and disallowed interactions with the environment. The processes lying within the borders are made up of a limited number of operators, each corresponding to a distinct possible behaviour. Given a denumerable set of names (channels), the basic syntax of internal processes (ranged over by $P, Q, \ldots$) and the semantic meaning associated with the various operators is given as follows:

$\overline{x}\langle y\rangle.\, P$	can output the name y over x and subsequently act as P;
$x(y).\, P$	can perform an input over x, bind the received datum to y, and then act as P;
$P \mid Q$	behaves as P in parallel with Q; the two sub-processes can either run independently or synchronise;
$!\, P$	behaves as $P \mid !\, P$, i.e., it can spawn infinitely many copies of P.

Above, "synchronisation" corresponds to the matching of complementary actions, namely an input and an output over the same channel name. A few more operators are also used. They will be presented later on in this section.

Beta-binders is equipped with an intuitive graphical representation. We now explain the computational rules of the formalism by showing their application to a running example. Firstly consider the following Beta-binders process, denoted as $\mathcal{S}_1$.



System $\mathcal{S}_1$ is composed by three Beta-binders processes, called *boxes*: A_1, B_1, and C_1. These boxes represent sub-components that run in parallel, and their distribution in the space is irrelevant (e.g., one could as well draw box C_1 to the left of A_1). Each box is equipped with a beta binder (i.e. an interface). For example, box A_1 is given the binder $x : \{a_1, a_2\}$, named x and typed by the set of names $\{a_1, a_2\}$.

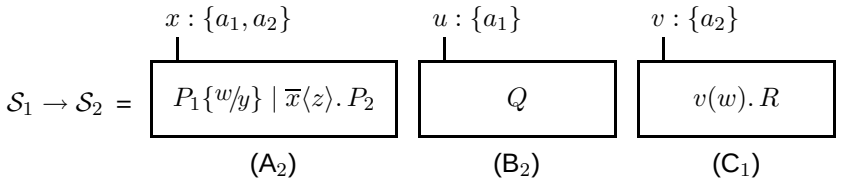
Inter-boxes communication

In system $\mathcal{S}_1$, box A_1 can interact with either B_1 or C_1. This is so because:

- A_1 can perform the input $x(y)$ over the name x of its binder, B_1 can perform the output $\overline{u}\langle w\rangle$ over the name u of its binder, and, since the types of x and u are not disjoint, these input and output can match;

- also, the types of x and of the binder v are not disjoint, and the output $\overline{x}\langle z\rangle$ of A_1 can match the input $v(w)$ of C_1.

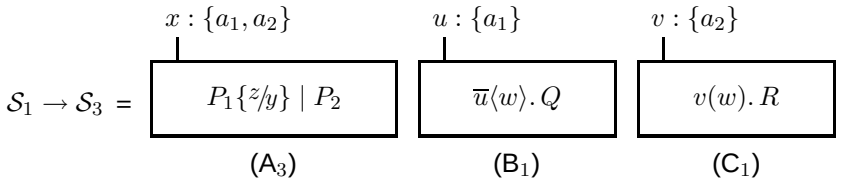
Let us consider the inter-communication between A_1 and B_1. It consumes the actions $x(y)$ and $\overline{u}\langle w\rangle$, in the first and in the second box respectively, and leads to the following configuration.



Notice that the information w flowed from box B_1 of $\mathcal{S}_1$ to box A_2 of $\mathcal{S}_2$, represented by the substitution of the name w for the occurrences of y in P_1, written $P_1\{^w/_y\}$.

Intra-boxes communication

In system $\mathcal{S}_1$, a communication within box A_1 is enabled as well. It would lead to the following configuration.



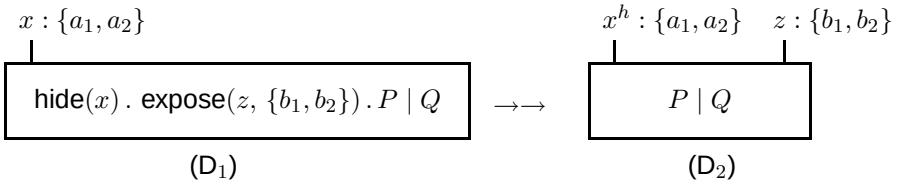
In the transformation $\mathcal{S}_1 \to \mathcal{S}_3$ we observe an internal modification of the leftmost box from A_1 to A_3. The other two boxes remain unaffected. Given the initial system $\mathcal{S}_1$, inter-communication and intra-communication are both allowed to occur. This reflects real biological situations. For example, internal modifications of the structure of a protein are in competition with environmental solicitations, like, e.g., the interaction with an enzyme.

Interface handling

Internal processes are also provided with a limited number of operations for managing box interfaces. The associated syntax and semantics is described below:

$\mathsf{hide}(x) . P$	make the binder x invisible then behaves like P.
	(When made invisible, x is written x^h.)
	If the enclosing box has no x-named binder, it stucks;
$\mathsf{unhide}(x) . P$	make the binder x^h visible then behaves like P.
	(When made visible again, x^h is turned back to x.)
	If the enclosing box has no x^h-named binder, it stucks;
$\mathsf{expose}(x, \Delta) . P$	add a x-named binder typed by Δ then behaves as P.

Consider for instance the box D_1 drawn below.

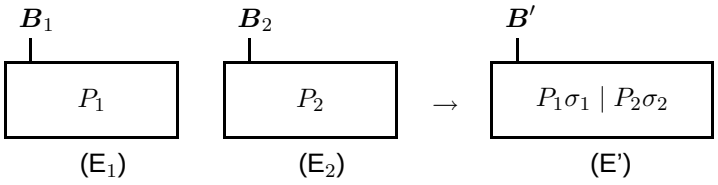


The execution of the prefix hide(x) hides the binder named x and changes its name to x^h. Then the execution of the prefix expose$(z,\ \{b_1, b_2\})$ adds to the box a new beta binder typed by $\{b_1, b_2\}$. As a result of this, after two computational steps, D_1 is transformed into D_2.

Notice that a box may be associated with more than one single binder, as it is the case for D_2 above. Indeed, unless otherwise specified, when talking about the "binder" of any given box, we refer to the set of all its singular binders. For instance, the binder of D_2 is the set composed of the two elements $x^h : \{a_1, a_2\}$ and $z : \{b_1, b_2\}$.

Box joining and splitting

To handle the box structure, Beta-binders is provided with operations for joining boxes together and for splitting one box in two. The *join* operation is parametric w.r.t. a function, called f_{join}, and models different possible ways of merging boxes, each of them depending on a distinct instantiation of f_{join}. Let $\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}'$ stay for box binders, and let σ_1, σ_2 represent name substitutions. Then, under the hypothesis that the actual function f_{join} is defined at $(\boldsymbol{B}_1, \boldsymbol{B}_2, P_1, P_2)$ and that $f_{join}(\boldsymbol{B}_1, \boldsymbol{B}_2, P_1, P_2) = (\boldsymbol{B}', \sigma_1, \sigma_2)$, the general pattern of the join transformation can be graphically rendered as follows:



The above transformation, just like those illustrated before, can be applied to a subset of a bigger system, leaving the rest unaffected. Namely, if the global system were made up of E_1, of E_2, and of some other box E_3, then after the transformation the system would be composed of two boxes: E' and E_3.

The operation that rules the splitting of boxes is dual to the above joining transformation. If $f_{split}(\boldsymbol{B}, P_1, P_2) = (\boldsymbol{B}_1, \boldsymbol{B}_2, \sigma_1, \sigma_2)$ then a box with binder $\boldsymbol{B}$ and internal process $P_1 \mid P_2$ is modified in two boxes: each of them with binder $\boldsymbol{B}_i$ and internal process $P_i\sigma_i$, for $i = 1, 2$.

4.1 Integrating shape spaces into Beta-binders

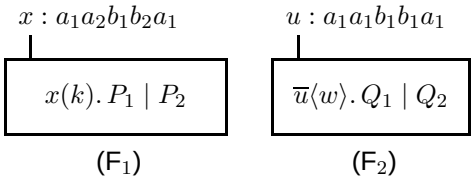
Beta-binders offers a natural ground for the integration of shape spaces. Essentially, in other process calculi interactions only depend on the matching of complementary actions (e.g., of an input and an output over the same channel). In Beta-binders the above requirement is partially relaxed and, at the level of interfaces, an input over $x : \Delta$ can match whichever output over $w : \Gamma$, provided that Δ and Γ share some common element.

Here the definition of types and their management is specialised further so to capture the intuition behind Hamming spaces and string matching rules. In particular, types become *strings of names*, and compatibility of types (originally interpreted as non-empty intersection of sets) becomes *distance between strings*.

As it was observed in Section 3, distinct matching rules can best fit different contexts. Hence we adopt the following abstract definition of distance.

Definition 4.1 *Given two strings of symbols Γ and Δ over the alphabet $\mathcal{A}$, the **distance** function $\rho(\Gamma, \Delta)$ is a map $\mathcal{A}^n \times \mathcal{A}^m \to \mathbb{R}$, where n is the length of Γ, and m is the length of Δ.*

The definition of the distance function leaves the user free to use different matching rules, leaving the rest of the formal system unaffected. Consider for example the following boxes.



We can set to ∞ the distance between strings of different length, and assume symbols be complementary to themselves. Then, adopting the Hamming distance for strings of the same length, we get $\rho(a_1a_2b_1b_2a_1\ ,\ a_1a_1b_1b_1a_1) = 2$. In a quantitative context this value could be directly used for deriving specific stochastic parameters. In a qualitative view, one can say that the interaction between F_1 and F_2 is allowed only if the distance between the types of their binders is lower than a given threshold (see [11] for some examples about this). Indeed, to directly deal with shape spaces, the formal rule defined in [3] for the inter-communication between a box with an elementary binder $x : \Gamma$ and a box with an elementary binder $y : \Delta$ is modified by requiring that:

$$\rho(\Gamma, \Delta) < \text{threshold}$$

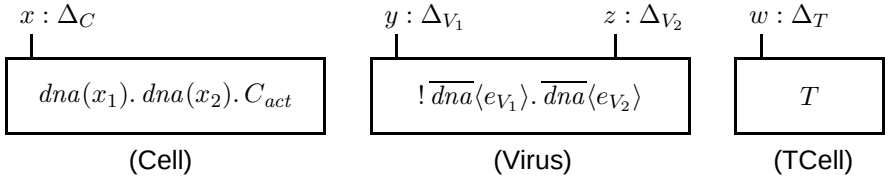
where "threshold" is a suitable user-defined value.

5 A simple model from the immune system

We conclude the paper by showing how the above formalism can be applied to model the interactions of a few key players of the immune system.

As an example, we consider a little system composed of a cell, a virus, and a cytotoxic T cell. First, we represent the three actors as boxes with appropriate binders. Then define suitable instances of the f_{join} and of the f_{split} functions that allow the modeling of cell infection, virus replication, and binding of T cells to infected cells. On passing, we show one of the possible runs of the whole system.

The formal representation of the system is graphically given as follows:



$(S1)$

where $\Delta_C = cabbaba$, $\Delta_{V_1} = vbaaaab$, $\Delta_{V_2} = vabbbabba$, $\Delta_T = tbaaabbbb$, and $C_{act} =!\,\mathsf{expose}(y,\, x_1).\ \mathsf{expose}(z,\, x_2)$. Also, for notational convenience, we use the following shorthands: $C = dna(x_1).\ dna(x_2).\, C_{act}$, and $V = !\,\overline{dna}\langle e_{V_1}\rangle.\,\overline{dna}\langle e_{V_2}\rangle$.

Looking at the above specification of C_{act}, notice that we use a name, rather than a set, as second parameter of the $\mathsf{expose}(\_,\, \_)$ operator. This discrepancy w.r.t. the semantics presented in [3] can be motivated by supposing that some names are taken from a distinguished set in bijection with the set of the typing strings. Under the same assumption, the names e_{V_1} and e_{V_2} transmitted by Virus over dna are to be thought of as encodings of the types Δ_{V_1} and Δ_{V_2}.

Cell represents an eukaryote cell with a site x that expresses its interaction capabilities. The cell machinery is rendered by the input actions over dna that, when consumed, trigger C_{act} and hence the exposition of new binders.

Virus stays for an intracellular parasite. This kind of parasites consist of an outer cell (*capsid*) made up of proteins and of an interior core containing the genome (DNA or RNA). A virus can enter into a cell and, once inside, it uses the cell machinery to duplicate the genome and to synthesize proteins. In this way the virus builds a new capsid and a new core, i.e. it duplicates itself. The newly generated virus can exit the cell, while the originator still infects it. The Virus box is provided with two sites, the one typed by Δ_{V_1} is used to model cell infection, while the site typed by Δ_{V_2} can be recognised by highly specific lymphocytes (which are missing from the present picture).

TCell represents a cytotoxic T cell of the adaptive immune responses. T cells circulate in the body searching for cells that have been infected by external organisms, like viruses. In fact, infected cells display on their surfaces some fragments of the viral proteins. A T cell that recognises an infected cell kills it, so preventing the diffusion of the virus. T cells are highly specific and hence they require a high affinity with the virus fragment displayed by the infected cell.

The binder types we use in our model are strings that encode the representation of the box they belong to. In the above model, the alphabet of the typing strings is $\{c, v, t, a, b\}$. The first symbol of the string encodes the owner of the binder associated with that type: c stays for Cell, v for Virus, and t for TCell. The rest of any typing string, that actually represents the shape of the binder, is made up of as and bs.

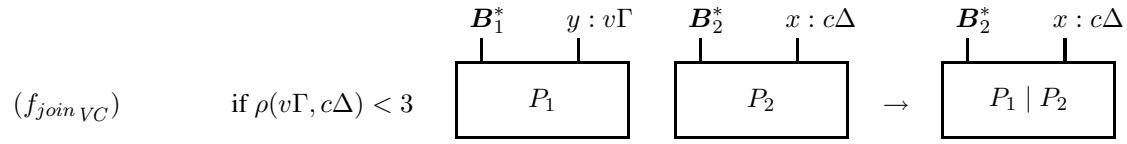
We assume that the complementarity function $\mathcal{C} : \mathcal{A}^m \to \mathcal{A}^m$ behaves as the identity on the elements of the set $\{c, v, t\}$ and maps as to bs and bs to as. Letting $H(\_,\, \_)$ denote the Hamming distance, we now define the distance function as follows:

$$\begin{aligned}\rho(x\Delta, y\Gamma) \quad = \quad & \text{if } m =\mid \Delta \mid = \mid \Gamma \mid \text{ and } \Delta \in \{a,b\}^m \text{ and } \Gamma \in \{a,b\}^m \\ & \text{then } H(\Delta, \mathcal{C}(\Gamma)) \text{ else } max(\mid \Delta \mid, \mid \Gamma \mid).\end{aligned}$$

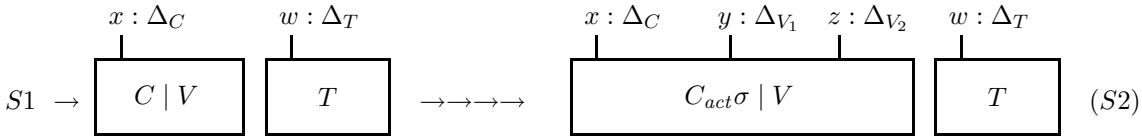
Notice that the first symbol of each of the two strings, representing the class of the binder rather than its shape, is ignored. Also observe that $\rho(\Delta_C, \Delta_{V_1}) = H(abbaba, \mathcal{C}(baaaab)) = 1$, meaning that the affinity between the binder of Cell and the virus binder named y is high.

We now complete the specification of our model by defining the functions that rule the joining and splitting of boxes. In what follows, the metavariables $\boldsymbol{B}^*, \boldsymbol{B}^*_1, \boldsymbol{B}^*_2, \ldots$ are used to denote possibly empty box binders.

The first rule, driven by an instance of f_{join} called $f_{join\,VC}$, can be graphically rendered as follows.



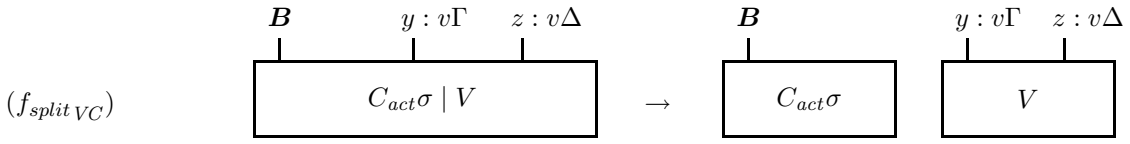
The rule states that, if in the global system there are two boxes which exhibit binders typed by $v\Gamma$ and by $c\Delta$, respectively, and if the distance between these two types is less than 3, then the two boxes can be joined together and the resulting box has the same binder as the $c\Delta$-typed box. When applicable, this rule models *cell infection*. In particular, starting from the global system $S1$, we get:

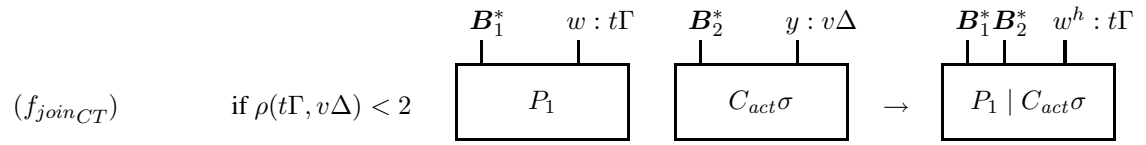


where $\sigma = \{e_{V_1}/x_1\,, e_{V_2}/x_2\}$. In the above, the first computational step is due to the $f_{join\,VC}$ transformation that makes the virus genetic material V enter the cell. The following computational steps correspond to intra-communications over dna and to the subsequent exposition of the received binder types.

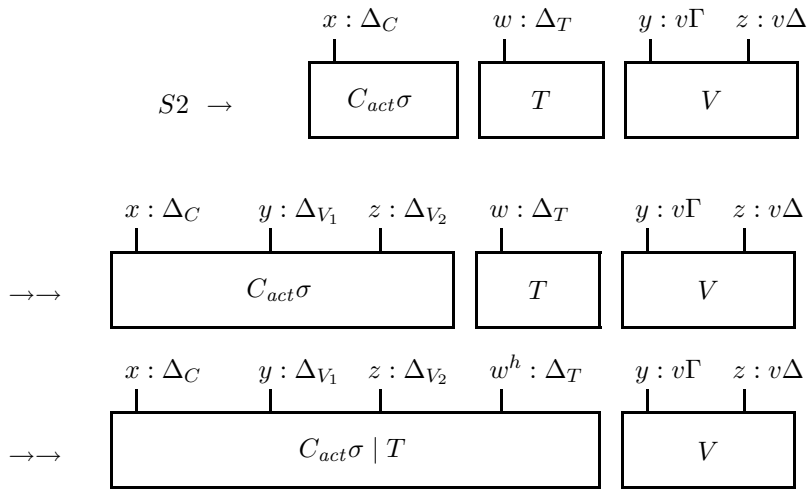
Referring to $S2$, observe that $C_{act}\sigma = !\,\mathsf{expose}(y,\, e_{V_1})\,.\,\mathsf{expose}(z,\, e_{V_2})$ can keep exposing the virus proteins an unlimited number of times. So the cell never gets "consumed" by the virus. As outlined in [3], a slight refinement of the formal model could take care of this aspect and limit the number of possible replications of the process $\mathsf{expose}(y,\, e_{V_1})\,.\,\mathsf{expose}(z,\, e_{V_2})$.

We now define the rules for *virus replication* and for the *binding of T cells* to infected cells, which are driven by $f_{split\,VC}$ and $f_{join_{CT}}$, respectively.





We conclude this section by showing one of the possible computations that can be automatically derived from $S2$. The computation reflects the application of the following sequence of transformations: $f_{split\,VC}$, exposition of the Δ_{V_1} and Δ_{V_2}-typed binders by the infected cell, and $f_{join_{CT}}$ transformation.



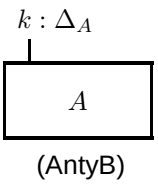
5.1 Compositionality to hammer complexity

One of the main advantages in using formal methods and process calculi theory comes from *compositionality*. Compositionality means that it is possible to develop different pieces of a model separately and then putting all together following mathematical rules. The underlying idea is to see biomolecular (as well as cellular) systems as a set of elementary components from which complex entities are constructed. This introduces a new paradigm with respect to "classical" complex biological system modeling. Indeed immunologists are moving in this direction, leaving differential equations models for agent based model [15] or stochastic stage-structured model [13]. But the new models lack in strong mathematical backgrounds, that seems a mandatory requirement for modeling, analysing and sharing biological knowledge.

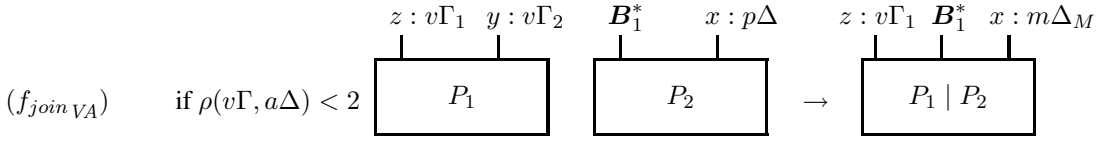
In this section we realise this idea developing a simple model of an antibody and showing how it can be integrated in the model we presented above.

An antibody is a molecule that has a specialised portion for identifying other molecules called *paratope*. Paratope has a defined shape that characterises the molecules that it can interact with. Each foreign molecule (e.g. viruses) presents a certain relief or pattern that can be recognised with various degrees of precision by complementary patterns or paratopes located on antibody molecules. When an antibody A recognises a virus V, it happens that A binds V preventing infection. Moreover the newly generated complex $A - V$ has a new binder that helps phagocytic cell. For instance, the capsule that surrounds pneumococci protects them from phagocytosis. If the appropriate antibodies are present in the body, they combine with the capsule and now the pneumococci can be ingested [16].

The formal representation of an antibody is graphically described below:

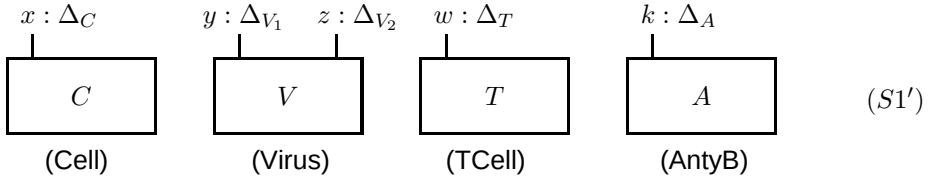


where $\Delta_A = pabbbba$. Also we have to define the function that drives the joining of a virus and an antibody.

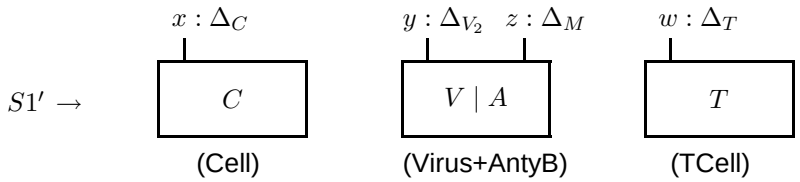


The above rule states that, if there is an antibody that is able to recognise a virus, then the antibody binds it. The new complex is no more able to enter in a cell and moreover a new binder $x : m\Delta_M$ is added. This binder has high affinity with phagocytic cells helping the complex ingestion.

Once defined the antibody model we can extend the system $(S1)$ leading to:



This system can perform the same sequence of transformations depicted in the previous section, without any differences. Moreover the antibody can bind to the virus so preventing cell infection.



Notice that the complex (Virus-AntyB) is no longer able to interact with (Cell). Moreover it is possible to automatically infer all the possible future behaviours of the system, giving a powerful methodology for investigating complex systems.

6 Conclusions and perspectives

We presented a formalism to model complex (biological) systems. The main objective is to determine the suitable abstraction to have a formal description of systems on top of which analysis and simulation can be implemented. We selected here the biological applicative domain and in particular the immune systems because handling these phenomena could help us to control, design and program complex systems as well as to re-construct systems from incomplete data. The main achievement of the present work with respect to the definition of Beta-binders is the inclusion in the formalism of the representation of shape spaces. This enhancement shows how the formalism proposed is flexible in specifying systems at different levels of abstraction. We worked out also the main feature of compositionality showing how a system

can be specified incrementally by simply adding new descriptions to the ones already produced when new information becomes available.

We are confident that this is a first step towards a theory of complex systems whose main brick is formal semantics of programming languages for concurrency (specifically process calculi). The study of process calculi and their definition since the beginning inspired by biological phenomena could lead to new biomimetic computational paradigms and primitives. The main goals of the new specification languages are to implicitly handle complexity in their semantic definition and to model, analyse, simulate and compare different systems. We feel this step very important because the major big step in research has been performed in the past when difficult concepts have been abstracted into good linguistic frameworks. Furthermore, the hierarchical nature of biological systems allows us to exploit the relation macro-world, meso-world and micro-world as a compilation problem.

The impact of the above mentioned strategy is for sure on the life science side by helping biologists in their research, but also on the computer science side defining new techniques able to handle systems more complex than the actual ones. Furthermore the abstraction of biological models in terms of interaction/communication of active entities can help enhancing the understanding of many fields of computer science, besides of course complex systems theory. We mention here just two of them because are of relevant interest nowadays.

Global computing and global computer have similar properties to biological systems because they are made up of autonomous and widely dispersed entities not centrally controlled, they include mobile code and appliances whose configurations vary over time and have incomplete information on the environment in which they work. If we can devise formalisms to handle biological systems, then we have good chances to improve the global computing field as well.

The particular application of immune system design, analysis and simulation could lead to new ideas for developing information security framework that are self-adapting to the new context in which a treat is ongoing.

Summing-up, we hope that the development of the present paper could lead to a larger applicability of the formalism presented in many different fields. In fact the abstraction molecules as processes and interaction as communication can be applied to any system whose evolution step is an interaction of some kind.

References

1. Gillespie D: **A general method for numerically simulating the stochastic time evolution of coupled chemical reactions**. *J Comput Phys* 1976, **22**.
2. Priami C, Regev A, Silverman W, Shapiro E: **Application of a stochastic passing-name calculus to representation and simulation of molecular processes**. *IPL* 2001, **80**.
3. Priami C, Quaglia P: **Beta binders for biological interactions**. In *CMSB '04*, Volume 3082 of *LNBI*. Edited by Danos V, Schächter VV, Springer 2005.
4. Perelson AS, Oster GF: **Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination**. *J Theor Biol* 1979, **81**(4).
5. Regev A, Shapiro E: **Cells as Computations**. *Nature* 2002, **419**.
6. Danos K, Krivine J: **Formal molecular biology done in CCS-R**. In *BioConcur '03* 2005.
7. Regev A, Panina E, Silverman W, Cardelli L, Shapiro E: **BioAmbients: An Abstraction for Biological Compartments**. *TCS* 2004, **325**.

8. Cardelli L: **Brane Calculi**. In *CMSB '04*, Volume 3082 of *LNBI*. Edited by Danos V, Schächter VV, Springer 2005.

9. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: *Molecular biology of the cell (IV ed.)*. Garland science 2002.

10. Smith D, Forrest S, Hightower R, Perelson A: **Deriving shape-space parameters from immunological data**. *J Theor Biol* 1997, **189**.

11. Detours V, Sulzer B, Perelson AS: **Size and Connectivity of the Idiotypic Network Are Independent of the Discreteness of the Affinity Distribution**. *J Theor Biol* 1996, **183**.

12. De Boer RJ, Perelson AS: **Size and connectivity as emergent properties of a developing immune network**. *J Theor Biol* 1991, **149**.

13. Chao D, Davenport M, Forrest S, AS P: **A stochastic model of cytotoxic T cell responses**. *J Theor Biol* 2004, **228**.

14. Milner R: *Communicating and Mobile Systems: the π-calculus*. Cambridge Univ. Press 1999.

15. Seiden P, Celada F: **A Model for Simulating Cognate Recognition and Response in the Immune System**. *J Theor Biol* 1992, **158**.

16. Wood W, Smith M, Watson B: **Studies on the Mechanism of Recovery in Pneumococcal Pneumonia**. *J Exp Med* 1946, **84**:387.

Complex Qualitative Models in Biology: a new approach

P. VEBER [1] M. LE BORGNE [1] A. SIEGEL [1] S. LAGARRIGUE [3]
O. RADULESCU [2]

[1] Projet Symbiose. Institut de Recherche en Informatique et Systèmes Aléatoires, IRISA-CNRS 6074-Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

[2] Institut de Recherche Mathématique de Rennes, UMR-CNRS 6625, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

[3] UMR Génétique animale, Agrocampus Rennes-INRA, 65 rue de Saint-Brieuc, CS 84215 Rennes, France

Abstract. *We advocate the use of qualitative models in the analysis of large biological systems. We show how qualitative models are linked to theoretical differential models and practical graphical models of biological networks. A new technique for analyzing qualitative models is introduced, which is based on an efficient representation of qualitative systems. As shown through several applications, this representation is a relevant tool for the understanding and testing of large and complex biological networks.*

1 Introduction

Understanding the behavior of a biological system from the interplay of its molecular components is a particularly difficult task. A model-based approach proposes a framework to express some hypotheses about a system and make some predictions out of it, in order to compare with experimental observations. Traditional approaches (see [6] for an interesting review) include ordinary differential equations or stochastic processes. While they are powerful tools to acquire a fine grained knowledge of the system at hand, these frameworks need accurate experimental data on chemical reactions kinetics, which are scarcely available. Furthermore, they also are computationally demanding and their practical use is restricted to a limited number of variables.

As an answer to these issues, many approaches were proposed, that abstract from quantitative details of the system. Among others, let us stress the work done on gene regulation dynamics [7], hybrid systems [10] or discrete event systems [4], [3]. The goal of such qualitative frameworks is to enable system-level analysis of a biological phenomenon. This appears as a relevant answer to recent technical breakthrough in experimental biology:

- microarrays, mass spectrometry, protein chips currently allow to measure thousands of variables simultaneously,
- obtained measurements are rather noisy, and may not be quantitatively reliable.

Microarrays for instance, are used for comparing the activity of genes between two experimental settings. A microarray experiment gives differential measure between two experimental settings. It delivers informations on the relative activity of each gene represented on the array. Despite many attempts made to quantified the output of microarrays, the essential output of the technique says, for example, that a gene G is more active in situation A than in situation B.

In this paper, we use a framework developed in [25] for the comparison of two experimental conditions, in order to derive qualitative constraints on the possible variations of the variables. Our main contribution is the use of an efficient representation for the set of solutions of a qualitative system. This representation allows to solve systems with hundreds of variables. Moreover, this representation opens the way to finer analysis of qualitative systems. This new approach is illustrated by solving three important problems:

- checking the accordance of a qualitative system with qualitative experimental data.
- minimally correcting corrupted data in discordance with a model
- helping in the design of experiments

Our main focus here is to show how to use large qualitative models and qualitative interpretations of experimental data. In this respect our work could be used as an extension to what was proposed in [23], where basically the authors propose to analyze pangenomic gene expression arrays in *E.coli*, using simple qualitative rules.

In the first section we establish links between differential, graphical and qualitative models.

2 Mathematical modeling

In this section we show how qualitative models can be linked to more traditional differential models. Differential models are central to the theory of metabolic control [9, 11]. They also have been applied to various aspects of gene networks dynamics. The purpose of this section is to lay down a set of qualitative equations describing steady states shifts of differential models. For the sake of completeness, we rederive in a simpler case results that have been established in greater generality in [25, 22].

2.1 Modeling assumptions

Let us consider a network of interacting cellular constituents, numbered from 1 to n. These constituents may be proteins, RNA transcripts or metabolites for instance. The state vector X denotes the concentration of each constituent.

Differential dynamics X is assumed to evolve according to the following differential equation:

$$\frac{dX}{dt} = F(X)$$

where F is an (unknown) nonlinear, differentiable function. A steady state X_{eq} of the system is a solution of the algebraic equation:

$$F(X_{eq}) = 0.$$

Steady states are asymptotically stable if they attract all nearby trajectories. A steady state is nondegenerated if the Jacobian calculated in that steady state is non-vanishing. According to the Grobman-Hartman theorem, a sufficient condition to have nondegenerated asymptotically stable steady states is $Re(\lambda_i) < -C, C > 0, i = 1, \ldots, n$, where λ_i are the eigenvalues of the Jacobian matrix calculated at the steady state.

Experiment modeling Typical two state experiments such as differential microarrays are modeled as steady state shifts. We suppose that under a change of the control parameters in the experiment, the system goes from one non-degenerated stable steady state to another one. The output of the two state experiment can be expressed in terms of concentration variations for a subset of products, between the two states. We suppose that the signs of these variations were proven to be statistically significant.

Interaction graph The only knowledge we require about the function F concerns the signs of the derivatives $\frac{\partial F_i}{\partial X_j}$. These are interpreted as the action of the product j on the product i. It is an activation if the sign is **+**, an inhibition if the sign is **−**. A null value means no action.

An interaction graph $G(V,E)$ is derived from the Jacobian matrix of F:

- with nodes $V = \{1,\ldots,n\}$ corresponding to products
- and (oriented) edges $E = \{(j,i) | \frac{\partial F_i}{\partial X_j} \neq 0\}$. Edges are labeled by $s(j,i) = \mathrm{sgn}(\frac{\partial F_i}{\partial X_j})$.

The set of predecessors of a node i in G is denoted $\mathrm{pred}(i)$. The interaction graph is actually built from informations gathered in the literature. In consequence in some places it may be incomplete (some interactions may be missing), in others it may be redundant (some interactions may appear several times as direct and indirect interactions). It is an important issue that neither incompleteness nor redundancy do not introduce inconsistencies and this will be addressed in section 5.

Negative diagonal in the Jacobian matrix For any product i, we exclude the possibility of vanishing diagonal elements of the Jacobian $\frac{\partial F_i}{\partial X_i}$. This can be justified by taking into account degradation and dilution (cell growth) processes that can be represented as negative self-loops in the interaction graph, that is for all i, $(i,i) \in E$ and $s(i,i) = -$.

Discussion In our mathematical modeling we suppose that the system starts and ends in non-degenerated stable steady states. Of course this is not always the case for several reasons: the waiting time to reach steady state is too big; one can end up in a limit cycle and oscillate instead of reaching a steady state. All these possibilities should be considered with caution. Actually this hypothesis might be difficult to check from the two states only. Complementary strategies such as time series analysis could be employed in order to assess the possibility of limit cycle oscillations.

Positive self-regulation is also possible but introduces a supplementary complication. In this case for certain values of the concentrations degradation exactly compensates the positive self-regulation and the diagonal elements of the Jacobian vanish (this is a consequence of the intermediate value theorem). We can avoid dealing with this situation by considering that the positive self-regulation does not act directly and that it involves intermediate species. This is a realistic assumption because a molecule never really acts directly on itself (transcripts can be auto-regulated but only via protein products). Thus, all nodes can keep their negative self-loops and all diagonal elements of the Jacobian can be considered to be non-vanishing. Although the positive regulation may imply vanishing higher order minors of the Jacobian, this will not affect our local qualitative equations.

2.2 Quantitative variation of one variable

We focus here on the variation of the concentration of a single chemical species represented by a component X_i of the vector X. Since we have adopted a *static* point of view, we are only interested in the variation of X_i between two non-degenerated stable steady states X_{eq}^1 and X_{eq}^2 independently of the trajectory of the dynamical system between the two states.

Let us denote by $\hat{X}_i$ the vector of dimension n_i obtained by keeping from X all coordinates j that are predecessors of i in the interaction graph. Then, under some additional assumptions described and discussed in [22], we have the following result:

Theorem 2.1
The variation of the concentration of species i between two non-degenerated steady states X^1_{eq} and X^2_{eq} is given by

$$X^1_{eq_i} - X^2_{eq_i} = \int_S - \left(\frac{\partial F_i}{\partial X_i}\right)^{-1} \sum_{k \in \text{pred}(i)} \frac{\partial F_i}{\partial X_k} dX_k \tag{1}$$

where S is the segment linking $\hat{X}^1_{eq_i}$ to $\hat{X}^2_{eq_i}$.

Full proof is given in [22]. The above formula is a quantitative relation between the variation of concentrations and the derivatives $\frac{\partial F_i}{\partial X_j}$. Now our next move will be to introduce a qualitative abstraction of this relation.

2.3 Qualitative equations

We propose here to study Eq. 1 in sign algebra. By sign algebra, we mean the set $\{+,-,?\}$, where **?** represents undetermined sign. This set is provided with the natural commutative operations:

$$\begin{array}{lllll} + + - = ? & + + + = + & - + - = - & + \times - = - & + \times + = + \quad - \times - = + \\ ? + - = ? & ? + + = ? & ? + ? = ? & ? \times - = ? & ? \times + = ? \quad ? \times ? = ? \end{array}$$

Equality in sign algebra $\approx$ is defined as follows:

$\approx$	+	−	?
+	T	F	T
−	F	T	T
?	T	T	T

Importantly, qualitative equality is not an equivalence relation, since it is not transitive. This implies that computations in qualitative algebra must be carried with care. At least two major properties should be emphasized:

- if a term of a sum is indeterminate (**?**) then the whole sum is indeterminate.
- if one hand of a qualitative equality is indeterminate, then the equality is satisfied whatever the value of the other hand is.

A *qualitative system* is a set of algebraic equations with variables in $\{+,-,?\}$. A *solution* of this system is a valuation of the unknowns which satisfies each equation, and such that no variable is instantiated to **?**. This last requirement is important since otherwise any system would have trivial solutions (like all variables to **?**).

Theorem 2.2
Under the assumptions and notations of Theorem 2.1, if the sign of $\frac{\partial F_i}{\partial X_j}$ is constant, then the following relation holds in sign algebra:

$$s(\Delta X_i) \approx \sum_{k \in pred(i)} s(k,i)s(\Delta X_k) \tag{2}$$

where $s(\Delta X_k)$ denotes the sign of $X^1_{eq_k} - X^2_{eq_k}$.

By writing Eq. 2 for all nodes in the graph, we obtain a system of equations on signs of variations, later referred to as *qualitative system* associated to the interaction graph G. This will be used extensively in the next sections.

2.4 Link between qualitative and quantitative

The qualitative system obtained from Eq.2 is a consequence of the quantitative relations that result from Theorem 2.1. So the sign function maps a quantitative variation between two equilibrium points onto a qualitative solution of Eq.2. The converse is not true in general. For a given solution S of the qualitative system, there might be no equilibrium change ΔX in the differential quantitative model, s.t. each real-valued component of ΔX has the sign given by S.

However, some components of the solution vectors are uniquely determined by the qualitative system. They take the same sign value in every solution vector. For such so-called hard components, the sign of any quantitative solution (if it exists) is completely determined by the qualitative system.

We will use the previous properties to check the coherence between models and experimental data. By experimental data we mean the sign of the observed variation in concentration for some nodes. In particular, if the qualitative system associated to an interaction graph G has no solution given some experimental observations, then no function F satisfying the sign conditions on the derivatives can describe the observed equilibrium shift, meaning that either the model is wrong, either some data are corrupted. In the next section, we introduce a simplified model related to lipid metabolism, and illustrate the above described formalism.

3 Toy example: regulation of the synthesis of fatty acids

In order to illustrate our approach, we use a toy example describing a simplified model of genetic regulation of fatty acid synthesis in liver. The corresponding interaction graph is shown in Fig. 1.

Two ways of production of fatty acids coexist in liver. Saturated and mono-unsaturated fatty acids are produced from citrates thanks to a metabolic pathway composed of four enzymes, namely ACL (ATP citrate liase), ACC (acetyl-Coenzyme A carboxylase), FAS (fatty acid synthase) and SCD1 (Stearoyl-CoA desaturase 1). Polyunsaturated fatty acids (PUFA) such as arachidonic acid and docosahexaenoic acid are synthesized from essential fatty acids provided by nutrition; D5D (Delta-5 Desaturase) and D6D (Delta-6 Desaturase) catalyze the key steps of the synthesis of PUFA.

PUFA plays pivotal roles in many biological functions; among them, they regulate the expression of genes that impact on lipid, carbohydrate, and protein metabolism. The effects of PUFA are mediated either directly through their specific binding to various nuclear receptors (PPARα – peroxisome proliferator activated receptors, LXRα – Liver-X-Receptor α, HNF-4α) leading to changes in the trans-activating activity of these transcription factors; or indirectly as the result of changes in the abundance of regulatory transcription factors (SREBP-1c – sterol regulatory element binding-protein–, ChREBP, etc.) [13].

Variables in the model We consider in our model nuclear receptors PPARα, LXRα, SREBP-1c (denoted by PPAR, LXR, SREBP respectively in the model), as they are synthesized from the corresponding genes and the trans-activating active forms of these transcription factors, that is, LXR-a (denoting a complex LXRα:RXRα), PPAR-a (denoting a complex PPARα:RXRα) and SREBP-a (denoting the cleaved form of SREBP-1c. We also consider SCAP – (SREBP cleavage activating protein), a key enzyme involved in the cleavage of SREBP-1c, that interacts with another family of proteins called INSIG (showing the complexity of molecular mechanism).

We also include in the model "final" products, that is, enzymes ACL, ACC, FAS, SCD1 (implied in the fatty acid synthesis from citrate), D5D, D6D (implied in PUFA synthesis) as well as PUFA themselves.

Interactions in the model Relations between the variables are the following. SREBP-a is an activator of the transcription of ACL, ACC, FAS, SCD1, D5D and D6D [20, 13]. LXR-a is a direct activator of the

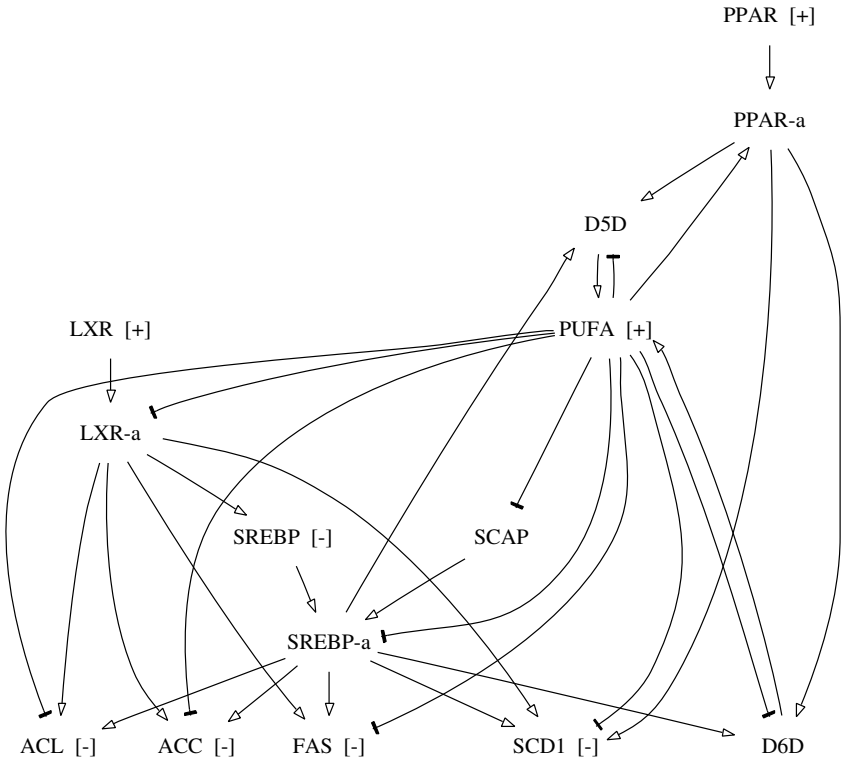


Figure 1: Interaction graph for the toy model. Self-regulation loops on nodes are omitted for sake of clarity. Observed variations are depicted next to each vertex, when available.

transcription of SREBP and FAS, it also indirectly activates ACL, ACC and SCD1 [26]. Notice that these indirect actions are kept in the model because we don't know whether they are only SREBP-mediated.

PUFA activates the formation of PPAR-a from PPAR, and inhibits the formation of LXR-a from LXR as well as the formation of SREBP-a (by inducing the degradation of mRNA and inhibiting the cleavage) [13]. SCAP represents the activators of the formation of SREBP-a from SREBP, and is inhibited by PUFA.

PPAR directly activates the production of SCD1, D5D, D6D [19, 27, 18]. The dual regulation of SCD1, D5D and D6D by SREBP and PPAR is paradoxical because SREBP transactivates genes for fatty acid synthesis in liver, while PPAR induces enzymes for fatty acid oxidation.
Hence, the induction of D5D and D6D gene by PPAR appears to be a compensatory response to the increased PUFA demand caused by induction of fatty acid oxidation.

Fasting-refeeding protocols The fasting-refeeding protocols represent a favorable condition for studying lipogenesis regulation; we suppose that during an experimentation, animals (as rodents or chicken) were kept in a fasted state during several hours. Then, hepatic mRNA of LXR, SREBP, PPAR, ACL, FAS, ACC and SCD1 are quantified by DNA microarray analysis. Biochemical measures also provide the variation of PUFA.

A compilation of recent literature on lipogenesis regulation provides hypothetical results of such protocols: SREBP, ACL, ACC, FAS and SCD1 decline in liver during the fasted state [17]. This is expected because fasting results in an inhibition of fatty acid synthesis and an activation of the fatty

acid oxidation. For the same reason, PPAR is increased in order to trigger oxidation. However, Tobin et al ([28]) showed that fasting rats for 24h increased the hepatic LXR mRNA, although LXR positively regulates fatty acid synthesis in its activated form. Finally, PUFA levels can be considered to be increased in liver following starvation because of the important lipolysis from adipose tissue as shown by Lee et al in mice after 72h fasting ([15]).

Qualitative system derived from the graph As explained in the previous section, we derive a qualitative system from the interaction graph shown in Fig. 1. For ease of presentation, we denote by `A` the sign of variation for species A.

System 1

```
(1)  PPAR-a  = PPAR + PUFA
(2)  LXR-a   = -PUFA + LXR
(3)  SREBP   = LXR-a
(4)  SREBP-a = SREBP + SCAP -PUFA
(5)  ACL     = LXR-a + SREBP-a - PUFA
(6)  ACC     = LXR-a + SREBP-a - PUFA
(7)  FAS     = LXR-a + SREBP-a - PUFA
(8)  SCD1    = LXR-a  + SREBP-a - PUFA + PPAR-a
(9)  SCAP    = -PUFA
(10) D5D     = PPAR-a + SREBP-a - PUFA
(11) D6D     = PPAR-a + SREBP-a - PUFA
```

Observations 1

```
PPAR    = +
PUFA    = +
LXR     = +
SREBP   = -
ACL     = -
ACC     = -
FAS     = -
SCD1    = -
```

In the next section, we propose an efficient representation for such qualitative systems.

4 Analysis of qualitative equations: a new approach

4.1 Resolution of qualitative systems

The resolution of (even linear) qualitative systems is a NP-complete problem (see for instance [29, 8]). One can show this by reducing the satisfiability problem for a finite set of clauses to the resolution of a qualitative system in polynomial time.

Let us consider a collection $C = \{c_1, \ldots, c_n\}$ of clauses on a finite set V of variables. Let $\{+, -, ?\}$ a sign qualitative algebra. In order to reduce the satisfiability problem to the resolution of a qualitative system, let us code $true$ into $+$ and $false$ into $-$. If c is a clause, let us denote by $\bar{c}$ the encoding of c in a qualitative algebra formula. The following encoding scheme provides a polynomial procedure to code a clause into a qualitative formula. :

clause		sign algebra
$a \in V$	$\rightarrow$	$\bar{a}$
$c_1 \vee c_2$	$\rightarrow$	$\bar{c_1} + \bar{c_2}$
$\neg c$	$\rightarrow$	$-\bar{c}$

The satisfiability problem for the set of clauses C is then reduced to finding a solution of the qualitative system:

$$\{\bar{c_i} \approx + \ / \ i = 1, \ldots, n\}$$

So a NP-complete problem can be reduced to the resolution of a qualitative system in polynomial time (with respect to the size of the problem). This shows that solving qualitative systems is a NP-complete problem. For example, the only pair of values which are not solution of $-\bar{a} + \bar{b} \approx +$ are $(+, -)$. This corresponds to the only pair $(true, false)$ that does not satisfy $\neg a \vee b$.

Several heuristics were proposed for the resolution of qualitative systems. For linear systems, set of rules have been designed [8]. This set is complete: it allows to find every solution. It is also sound: every solution found by applying these rules is correct. The rules are based on an adaptation of Gaussian

elimination. However only heuristics exist for choosing the equation and the rule to apply on it. In case of a dead-end, when no more rule can apply, it is necessary to backtrack to the last decision made. As a result programs implementing qualitative resolution are not very efficient in general and only problems of small size can be resolved in reasonable time. For that reason we propose an alternate way to solve qualitative systems (linear or not).

4.2 Qualitative equation coding

Our method is based on a coding of qualitative equations as algebraic equations over Galois fields $\mathbb{Z}/p\mathbb{Z}$ where p is a prime number greater than 2. The elements of these fields are the classes modulo p of the integers. If $\bar{x}$ denotes the class of the integer x modulo p, a sum and a product are defined on $\mathbb{Z}/p\mathbb{Z}$ as follows:

$$\bar{x}+\bar{y}=\overline{x+y} \qquad \bar{x}\times\bar{y}=\overline{x\times y}$$

Galois fields have two basic properties which we use extensively:

- Every function $f:(\mathbb{Z}/p\mathbb{Z})^n \to \mathbb{Z}/p\mathbb{Z}$ with n arguments $\mathbb{Z}/p\mathbb{Z}$ is a polynomial function
- if $\oplus$ denotes the operation $f\oplus g = f^{(p-1)}+g^{(p-1)}$, then every equation system $p_1(X)=0,\ldots,p_k(X)=0$ has the same solutions than the unique equation $p_1\oplus p_2\oplus\ldots\oplus p_k(X)=0$.

The following table specifies how the sign algebra $\{+,-,?\}$ is mapped onto the Galois field with three elements $\mathbb{Z}/3\mathbb{Z}$ is used for that coding.

sign algebra		$\mathbb{Z}/3\mathbb{Z}$
+	$\to$	1
−	$\to$	-1
?	$\to$	0

sign algebra		$\mathbb{Z}/3\mathbb{Z}$
e_1+e_2	$\to$	$\overline{e_1}.\overline{e_2}.(\overline{e_1}+\overline{e_2})$
$e_1\times e_2$	$\to$	$\overline{e_1}.\overline{e_2}$
$e_1\approx e_2$	$\to$	$\overline{e_1}.\overline{e_2}.(\overline{e_1}-\overline{e_2})$

Finally a qualitative system $\{e_1,\ldots,e_n\}$ is coded as the polynomial $\overline{e_1}\oplus\cdots\oplus\overline{e_n}$. A similar coding for the qualitative algebra $\{+,-,0,?\}$ uses the Galois field $\mathbb{Z}/5\mathbb{Z}$ and will not be presented here.

With this coding, every qualitative system has a solution if and only if the corresponding polynomial has a solution without null component. Null solutions are excluded since ? solutions are excluded for qualitative systems. In general we will have to add polynomial equations $X^2=1$ to insure this.

4.3 An efficient representation of polynomial functions

Recall that our purpose is to efficiently solve a NP-complete problem. There is no hope to find a representation of polynomial functions allowing to solve polynomial systems of equations in polynomial time. The coding of a qualitative system as a polynomial equation is obviously polynomial in the size of the system (number of variables plus number of equations). So finding the solution of a polynomial system of equations is itself a NP-complete problem. It is more or less the SAT problem.

Nevertheless, there exists a representation of polynomial functions on Galois fields which gives, in practice, good performances for polynomials with hundreds of variables. This kind of representation was first used for logical functions which may be considered as polynomial functions over the field $\mathbb{Z}/2\mathbb{Z}$. This representation is known as BDD (Binary Decision Diagrams) and is widely used in checking logical circuits [2] and in model checkers as nu-SMV [5].

We present here this representation for the field $\mathbb{Z}/3\mathbb{Z}$. Generalizations to other Galois fields could be treated as well. The starting point is a generalization of Shannon decomposition for logical functions:

$$p(X_1,X)=(1-X_1^2)p_{[X_1=0]}(X)+X_1(-X_1-X_1^2)p_{[X_1=1]}(X)+X_1(X_1-X_1^2)p_{[X_1=2]}(X)$$

where p is a polynomial function with n variables. This decomposition leads to a tree representation of the polynomial function: the variable X_1 is the root and has three children. Each of these is obtained by instantiating X_1 to -1, 0 or 1 in $p(X_1, X)$. This representation is exponential (3^n) as each non constant node has 3 children. It also depends on a chosen order on the variables.

Then a key observation (see [2]), is that several subtrees are identical. They have the same variable as root variable and isomorphic children. If we decide to represent only once each type of tree, then the tree representation is transformed into a direct acyclic graph. With this representation there is no more redundancy among subtrees. The result may be a dramatic decrease in the size of the representation of a polynomial function.

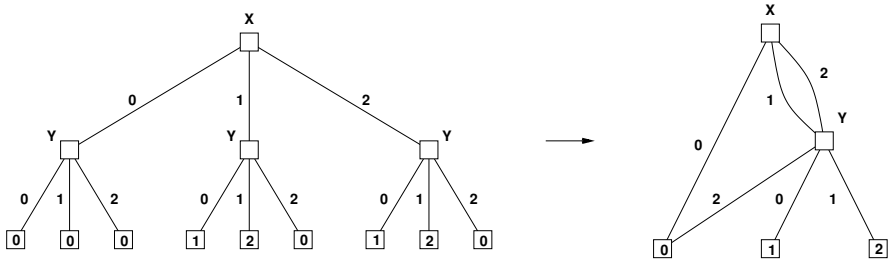


Figure 2: From tree representation to direct acyclic graph for $X^2(Y+1)$. The tree has 13 nodes while the DAG representing the same function has 5 nodes.

A property of the Shannon like decomposition is that many operations on polynomial functions are recursive with respect to this decomposition. More precisely let

$$p^i(X_1, X) = (1 - X_1^2)p_0^i(X) + X_1(-X_1 - X_1^2)p_1^i(X) + X_1(X_1 - X_1^2)p_2^i(X)$$

$i = 1, 2$ be two polynomial functions with $p_\alpha(X) = p_{[X_1=\alpha]}(X)$, $\alpha = 0, 1, 2$. Then for binary operations Δ on polynomial functions,

$$p^1 \Delta p^2 = (1 - X_1^2)(p_0^1 \Delta p_0^2) + X_1(-X_1 - X_1^2)(p_1^1 \Delta p_1^2) + X_1(X_1 - X_1^2)(p_2^1 \Delta p_2^2)$$

This kind of recursive formula leads to an exponential complexity of any computation. Again, it is possible to take advantage of the redundancy by using a cache to remember each operation. This technique is known as memoisation in formal calculus. A 40% cache hit rate is commonly observed.

More complex operations on polynomial functions are also implemented with a recursive scheme and memoisation. Let us just mention quantifier elimination as among the most useful for our purpose.

This representation of polynomial functions on Galois fields has also several drawbacks:

- the memory size heavily depends on the order of variables. The libraries implementing formal computations always have reordering algorithms.
- for each order, there exists polynomial functions which are exponential in memory size.

Nevertheless, in practice, this representation has proved to be very efficient for polynomial functions with several hundred of variables. The computations performed on our toy model and on another real size one used a program named SIGALI which is devoted to polynomial functions on $\mathbb{Z}/3\mathbb{Z}$ representation. Several algorithms were added to this program in order to answer questions of biological interest.

5 Qualitative models and experimental data

In this section, we show how to compute some properties of a qualitative system, and eventually get some insights on the biological model it represents. The algorithms we derive heavily rely on the representation introduced above. Hence, not only they can deal in practice with computationally hard problems efficiently, but also they are expressed in a rather simple and generic fashion.

Let M be a qualitative model represented by its associated interaction graph $G(V,E)$. Recall that V is the set of variables. Let V_O be the set of observed variables, and $o_i \in \{+,-\}$ for $i \in V_O$ the experimental observations. As explained in the previous section, the qualitative system derived from M can be coded as a polynomial function $P_M(X_1,\ldots,X_n)$. Roots of P_M correspond to solutions of the qualitative system.

5.1 Satisfiability of the qualitative system

A property of the coding described above, is that the system has no solution iff P_M is equal to the constant polynomial 1. Alternatively if $P_M = 0$, the qualitative equations do not constraint the variables at all.

Now if some observations o_i for $i \in V_O$ are available, checking their consistency with the model M boils down to instantiating $X_i = o_i$ in $P_M(X_1,\ldots,X_n)$, for all $i \in V_O$, and testing whether the resulting polynomial is different from 1.

We computed the polynomial P_L associated to our toy example (see section 3) and it has roots. Recall that it does not guarantee the existence of some (quantitative) differential model conforming to the interaction graph depicted in Fig. 1. Satisfiability of the qualitative system is only a necessary condition for the model to be correct.

The polynomial obtained by instantiating these observations into P_L is different from 1, meaning that our model does not contradict generally observed variations during fasting.

Large size models might advantageously be reduced using standard graph techniques. First we look for connected components in the interaction graph. A graph with several connected components represents a coherent qualitative model iff each component is coherent. Second, a node without successor except itself appears only in its associated equation. If this node is not observed, its associated qualitative equation adds no constraint on the other nodes. So, at least for satisfiability checking, this node can be suppressed and its qualitative equation removed from the system. This procedure is applied iteratively, until no node can be deleted. The resulting graph leads to a new qualitative system which is satisfiable iff the initial system is satisfiable.

5.2 Correcting data or model

If the qualitative system, given some experimental observations, is found to have no solution, it is of interest to propose some correction of the data and/or the model. By correction, we mean inverting the sign of an observed variable or the sign of an edge of the interaction graph. In the general case, there are several possibilities to make the system satisfiable, and we need some criterion to choose among them. We applied a parsimony principle: a correction of the data should imply a minimal number of sign inversions.

In the following, we show how to compute all minimal corrections for the data. Given $(o_i)_{i\in V_O}$ a vector of experimental observations which is not compatible with the model, we compute all $(o'_i)_{i\in V_O}$ vectors which are compatible with the data and such that the Hamming distance between o and o' is minimal. By Hamming distance, we mean the number of differences between o and o'. The set of such o' vectors might be very large; but again, by encoding it as the set of roots of a polynomial function, we obtain a compact representation.

This procedure can be extended in a straightforward manner to corrections of edges sign in the interaction graph. This is done by considering these signs as variables of the model. For ease of presentation, we only detail data correction.

Input:
P, a polynomial function on variables V
$i \in V$
Output:
C, a polynomial function encoding all minimal corrections
d, minimal number of corrections

if *P is constant* **then**
  if *$P = 0$* **then**
    Result: $C = 0$, $d = 0$
  else
    Result: $C = 1$, $d = \infty$
  end
else
  let P_0, P_1, P_2 be the Shannon decomposition of P with respect to variable X_i,
  and (C_j, d_j) the result obtained by recursively applying the algorithm on P_j and $i+1$ for $j \in \{0,1,2\}$

  let $d'_j = \begin{cases} d_j + 1 & \text{if } i \in V_O \text{ and } o_i \neq j \\ d_j & \text{otherwise} \end{cases}$ and $C'_j = \begin{cases} (X_i - j) \oplus C_j & \text{if } i \in V_O \\ C_j & \text{otherwise} \end{cases}$

  Result: $d = \min\, d'_j$, $C = \prod_{j,\, d'_j = d} C'_j$
end

Algorithm 1: Algorithm for experimental data correction.

Let us illustrate this algorithm on our toy example: during fasting experiments, synthesis of fatty acids tends to be inhibited, while oxidation, which produces ATP, is activated. In particular ACC, ACL, FAS and SCD1 are implied in the same pathway to produce saturated and monounsaturated fatty acids. Expectedly, they are known to decline together at fasting. Suppose we introduce some wrong observation, say for instance an increase of ACL, while keeping all other observations given above. The polynomial obtained from P_L including these new observations is equal to 1, and hence has no solution. Applying algorithm 1, we recover this error. Now if we wrongly change two values, say ACL and ACC to 1, the algorithm proposes a different correction, namely to change the observed value of SREBP to 1, which is more parsimonious.

5.3 Experiment design

It is often the case that not all variables in the system under study can be observed. Biochemical measurements of metabolites can be costly and/or time consuming. By experiment design, we mean here the choice of the variables to observe so that an experiment might be informative.

Let $P_M(X_O, X_U)$ be the polynomial function coding for the qualitative system M. X_O (resp. X_U) denotes the state vector of observed (resp. unobserved) variables. The polynomial function representing the admissible values of the observed variables is obtained by elimination of the quantifier in $\exists X_U\ P_M(X_O, X_U)$. Let $P_M^O(X_O)$ denote the resulting polynomial function.

For some choice of observed variables, it might well be that P_M^O is null, which basically means that the experiment is totally useless. Remark that no improvement can be found by taking a subset of X_O The solution is either to add new observed variables or to chose a completely different set of observed variables.

In order to assess the relevance of a given experiment (namely of a given observed subset), we suggest to compute the following ratio: number of consistent valuations for observed variables versus the total number of valuations of observed variables. A very stringent experiment has a low ratio. An experiment having a ratio value of one is useless.

Again this computation is carried out in a recursive fashion. Let P be a polynomial function rep-

resenting the set of admissible observed values. Let $Rat(p)$ the percentage of solutions of $P(X) = 0$ in the space $(\mathbb{Z}/p\mathbb{Z})^n$, where n is the number of variables X. If P is constant then $Rat(P) = 1$ (resp. $Rat(P) = 0$) if $P = 0$ (resp. $P \neq 0$). Else, let P_1, P_2, P_3 be a Shannon like decomposition of $P(X)$ with respect to some variable of P. Then it is easy to prove:

$$Rat(P) \;=\; (Rat(P_0) + Rat(P_1) + Rat(P_2))/3$$

The relevance of this approach was assessed on our toy example: for each subset O of variables in the model, containing at most four variables, we computed $Rat(P_L^O)$. Expectedly, the lowest ratios (i.e. the most stringent experiments) were achieved observing four variables: either {SCAP, PUFA, PPAR-a, PPAR}, or {SREBP, SCAP, PUFA, LXR-a}, or {SREBP, PPAR-a, PPAR, LXR-a}.

Interestingly, the procedure captures what might be though of as control variables, like PUFA/SCAP, SREBP/LXR-a and PPAR/PPAR-a. The first two pairs control the activation of fatty acids synthesis; the third one controls fatty acid oxidation.

Indeed one can go even further: if we isolate some kind of control variables, we are naturally interested in knowing how they constrain other variables. Achieving this amounts to computing the set of variables which value is constant for all solutions of the system (the so called hard components). Recall that these hard components of qualitative solutions are also important with respect to the hypothetical differential model which is abstracted in the qualitative one. Indeed, all solutions of the quantitative equation for equilibrium change have the same sign pattern on the hard components. Algorithm 2 describes a recursive procedure which finds the set of hard components, together with their value.

Input: P, a polynomial function on variables V
Output:
the set $W \subset V \times \{0,1,2\}$ of hard components, together with their values
a boolean b which is true if P has at least one root

if *P is constant* **then**
  if *$P = 0$* **then**
    return $(\emptyset, true)$
  else
    return $(\emptyset, false)$
  end
else
  let P_0, P_1, P_2 be the Shannon decomposition of P with respect to variable X_i,
  and (W_j, b_j) the result obtained by recursively applying the algorithm on P_j for $j \in \{0,1,2\}$
  let $W = \{(v,v') | v \in V,\ v' \in \{0,1,2\},\ \forall j\ b_j \Rightarrow (v,v') \in W_j\}$
  if *there exists a unique j_0 s.t. b_{j_0} is true* **then**
    add (i, j_0) to W
  end
  return $(W, b_0 \vee b_1 \vee b_2)$
end

Algorithm 2: Determination of hard components

Let us set some of our previously found control variables of the toy example, to a given value, say PUFA to 1, and LXR to -1. Then applying the algorithm 2, the corresponding polynomial has the following hard components:

```
ACL      = -1          FAS    = -1
ACC      = -1          LXR-a  = -1
SCAP     = -1          SREBP  = -1
SREBP-a  = -1          PPAR   = -1
PPAR-a   = -1
```

which expectedly corresponds to the inhibition of fatty acids synthesis.

5.4 Real size system

We have used our new technique to check the consistency of a database of molecular interactions involved in the genetic regulation of fatty acid synthesis. In the database, interactions were classified as behavioral or biochemical.

- a behavioral interaction describes the effects of a variation of a product concentration. It is either direct or indirect (unknown mechanism).
- a biochemical interaction may be a gene transcription, a reaction catalyzed by an enzyme ... Such molecular interactions can be found in existing databases. They need a behavioral interpretation.

All the behavioral interactions were manually extracted from a selection of scientific papers. Biochemical interactions were extracted from public databases available on the Web (Bind [1], IntAct [12], Amaze [16], KEGG [21] or TransPath [24]). A biochemical interaction may be linked to a behavioral interpretation in the database.

The database is used to generate the interaction graph. While behavioral interactions directly correspond to edges in the graph, biochemical interactions are given a simplified interpretation. Roughly, any increase of a reaction input induces an increase of the outputs.

The interaction graph which is built from the database contains more than 600 vertices and more than 1400 edges. It is clear that even though, the obtained graph is not a comprehensive model of genetic regulation of fatty acid synthesis in liver. Anyway our aim is to see how far this model can account for experimental observations, and propose some corrections when it cannot.

We used our technique to check the coherence of the whole model. After reducing the graph with standard graph techniques as described in section 5.1, we found that the model was incoherent. The reduced graph has about 150 nodes. We developed a heuristic to isolate minimal incoherent sub-systems. It turned out that all the contradictions we detected resulted from arguable interpretations of the literature.

6 Conclusion

In this paper we proposed a qualitative approach for the analysis of large biological systems. We rely on a framework more thoroughly described in [25], which is meant to model the comparison between two experimental conditions as a steady state shift. This approach fits well with state of the art biological measurement techniques, which provide rather noisy data for a large amount of targets. It is also well suited to the use of biological knowledge, which is most of the time descriptive and qualitative.

This qualitative approach is all the more attractive that we can rely on new analysis methods for qualitative systems. This new technique is also introduced in this paper and is original in qualitative modeling. It relies on a representation of qualitative constraints by decision diagrams. Not only this has a major impact on the scalability of qualitative reasoning, but it also permits to derive many algorithms in a quite generic fashion.

We plan to validate our approach on pathways which are published for yeast and *E.Coli*. Not only this pathways are of significant size but microarray data for this species are publicly available. Concerning the scalability of the methods, qualitative systems with up to 200 variables are handled within a few minutes.

On the theoretical side, we study applications of our algebraic techniques to network reconstruction, as proposed in [30]. The problem is to infer direct actions between products, based on large scale perturbation data, in order to obtain the most parsimonious interaction graph. Our approach could lead to a reformulation of this problem in terms of polynomial operations. Indeed, finding a minimal regulation network from a minimal polynomial representation has already been described in [14], though it was

applied to a rather different type of network. A similar approach tailored to the framework described in this paper could eventually lead to original and practical algorithms for network reconstruction.

Acknowledgment This research was supported by ACI IMPBio, a French Ministry for Research program on interdisciplinarity.

References

[1] GD Bader, D Betel, and CW Hogue. Bind: the biomolecular interaction network database. *Nucleic Acids Res.*, 31(1):248–50, 2003.

[2] R.E Bryan. Graph-based algorithm for boolean function manipulation. *IEEE Transactions on Computers*, 8:677–691, August 1986.

[3] Chaouiya C, Remy E, Ruet P, and Thieffry D. Qualitative modelling of genetic networks: From logical regulatory graphs to standard petri nets. *Lecture Notes in Computer Science*, 3099:137–156, 2004.

[4] N. Chabrier-Rivier, M.Chiaverini, V. Danos, F. Fages, and V. Schächter. Modeling and querying biomolecular interaction networks. *Theoretical Computer Science*, 325(1):25–44, 2004.

[5] E. Clarke, O. Grumberg, and D. Long. Verification tools for finite-state concurrent systems. in: A decade of concurrency - reflections and perspectives. *Lecture Notes in Computer Science*, 803, 1994.

[6] H. de Jong. Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1):69–105, 2002.

[7] H. de Jong, J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bulletin of Mathematical Biology*, 66(2):301–340, 2004.

[8] J.L. Dormoy. Controlling qualitative resolution. In *Proceedings of the seventh National Conference on Artificial Intelligence, AAAI88', Saint-Paul, Minn.*, 1988.

[9] David Fell. *Understanding the Control of Metabolism*. Portland Press, London, 1997.

[10] Ronojoy Ghosh and Claire Tomlin. Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modelling: Delta-notch protein signalling. *Systems Biology*, 1(1):170–183, 2004.

[11] Reinhart Heinrich and Stefan Schuster. *The Regulation of Cellular Systems*. Chapman and Hall, New York, 1996.

[12] H. Hermjakob, L. Montecchi-Palazzi, C. Lewington, S. Mudali, and al. Intact – an open source molecular interaction database˙ *Nucleic Acids Research*, 32:D452–D455, 2004.

[13] DB Jump. Fatty acid regulation of gene transcription. *Crit. Rev. Clin. Lab. Sci.*, 41(1):41–78, 2004.

[14] R. Laubenbacher and B. Stigler. A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theor. Biol.*, 229:523–537, 2004.

[15] SS Lee, WY Chan, CK Lo, and alt. Requirement of pparalpha in maintaining phospholipid and triacylglycerol homeostasis during energy deprivation. *J Lipid Res.*, 45(11):2025–37, 2004.

[16] C. Lemer, E. Antezana, F. Couche, and alt. The amaze lightbench: a web interface to a relational database of cellular processes. *Nucleic Acids Res.*, 32:D443–D448, 2004.

[17] G Liang, J Yang, JD Horton, RE Hammer, and alt. Diminished hepatic response to fasting/refeeding and liver x receptor agonists in mice with selective deficiency of sterol regulatory element-binding protein-1c. *J Biol Chem*, 277(15):9520–8, Jan 2002.

[18] T Matsuzaka, H Shimano, N Yahagi, and alt. Dual regulation of mouse delta(5)- and delta(6)-desaturase gene expression by srebp-1 and pparalpha. *J Lipid Res.*, 43(1):107–14, 2002.

[19] CW Miller and JM Ntambi. Peroxisome proliferators induce mouse liver stearoyl-coa desaturase 1 gene expression. *Proc Natl Acad Sci U S A.*, 93(18):9443–8, 1996.

[20] TY Nara, WS He, C Tang, SD Clarke, and MT Nakamura. The e-box like sterol regulatory element mediates the suppression of human delta-6 desaturase gene by highly unsaturated fatty acids. *Biochem. Biophys. Res. Commun.*, 296(1):111–7, 2002.

[21] H. Ogata, S. Goto, K. Sato, W. Fujibuchi, and al. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 27:29–34, 1999.

[22] O. Radulescu, S. Lagarrigue, A. Siegel, M. Le Borgne, and P. Veber. Topology and linear response of interaction networks in molecular biology. submitted to Royal Society Interface.

[23] Gutierrez-Rios RM, Rosenblueth DA, Loza JA, Huerta AM, Glasner JD, Blattner FR, and Collado-Vides J. Regulatory network of escherichia coli: consistency between literature knowledge and microarray profiles. *Genome Res.*, 13(11):2435–43, 2003.

[24] F. Schacherer, C. Choi, U. Gotze, M. Krull, S. Pistor, and E. Wingender. The TRANSPATH signal transduction database: a knowledge base on signal transduction networks. *Bioinformatics*, 17(11):1053–1057, 2001.

[25] A. Siegel, O. Radulescu, M. Le Borgne, P. Veber, J. Ouy, and S. Lagarrigue. Qualitative analysis of the relation between dna microarray data and behavioral models of regulation networks. *Biosystems*, submitted 2005.

[26] KR Steffensen and JA. Gustafsson. Putative metabolic effects of the liver x receptor (lxr). *Diabetes*, 53(Supp 1):36–52, Feb 2004.

[27] C Tang, HP Cho, MT Nakamura, and SD Clarke. Regulation of human delta-6 desaturase gene transcription: identification of a functional direct repeat-1 element. *J Lipid Res*, 44(4):686–95, 2003.

[28] KA Tobin, HH Steineger, S Alberti, O Spydevold, and alt. Cross-talk between fatty acid and cholesterol metabolism mediated by liver x receptor-alpha. *Mol Endocrinol*, 14(5):741–52, May 2000.

[29] L. Travé-Massuyès and P. Dague, editors. *Modèles et raisonnements qualitatifs*. Hermes sciences, 2003.

[30] A. Wagner. Reconstructing pathways in large genetic networks from genetic perturbations. *Journal of Computational Biology*, 11:53–60, 2004.

War of attrition with implicit time cost

Anders Eriksson[1], Kristian Lindgren[1] and Torbjörn Lundh[2]
[1] Dept. Physical Resource Theory, [2] Dept. of Mathematics
Chalmers University of Technology, SE-41296 Göteborg, Sweden

Many animals have a formidable arsenal of teeth, hooves or horns, and violent fights among these animals often result in death or serious injury. It is thus perhaps not surprising that there is a wide variety of ways in which animals settle disputes over food, mates or territory without resorting to violence. A common theme in such contests, is that the animals display until one of them gives up, leaving the prize to the animal that endured. It is then safe to assume that a cost can be associated with the length of the display, since otherwise animals would wait indefinitely.

Maynard Smith and Price (1974, J. Theor. Biol., 47, p. 209) pioneered the waiting-game as a model of wars of attrition. In this game, there is a prize worth one unit of fitness, that goes to the winner of the contest. For instance, the prize could be a desirable territory, and the fitness is then the expected number of offspring in the territory. The loser has to settle for a less attractive territory, which entails $k < 1$ fitness units. It is assumed that the contest costs c fitness units per unit of time, and both contestants pay the cost until one of them gives up, losing the contest. Making standard assumptions of the mating structure of the population, e.g. abundance of waiting-times in the population obey the replicator dynamics, this model has been thoroughly investigated in the literature.

The standard waiting-game assumes that all individuals in the population play the same number of games per unit of time. We investigate the co-evolutionary dynamics of a population where players engage in wars of attrition, where the time cost is not explicitly given, but instead depends implicitly on the strategies of the whole population (Eriksson *et al.* 2004, J. Theor. Biol., 230, p. 319). Each player participates in a series of games, where those prepared to wait longer win with higher certainty but play less frequently. The players in the population can be in one of two states: either they are involved in a contest with another player, or they are available for entering a new contest. The activity of the players in the population, during a generation, is modelled as a process that randomly selects pairs of available players to engage in contests. This leads to an implicit time cost, which is higher for players involved in longer games.

The model is characterised by the ratio of the winner's score to the loser's score, in a single game. The fitness of a player is determined by the accumulated score from the games played during a generation. We derive the stationary distribution of strategies under the replicator dynamics. When the score ratio is high, we find that the stationary distribution is unstable, with respect to both evolutionary and dynamical stability, and the dynamics converge to a limit cycle. When the ratio is low, the dynamics converge to the stationary distribution. For an intermediate interval of the ratio, the distribution is dynamically but not evolutionarily stable.

We find that our model has immediate implications for two earlier models that takes implicit costs into account. Hines (1977, J. Theor. Biol., 67, p. 141) proposed a model in which animals forage for food. When an animal finds a piece of food, with a given probability it may consume the food undisturbed, otherwise it enters a war of attrition for the food parcel. Here, it is assumed that engaging in competitions prevents foraging, so that there is a trade-off between the probability of winning a contest and the time spent foraging. It turns out that we can capture the evolutionary dynamics of this model within our model (assuming replicator dynamics), although the original model has four parameters and our model only has one. Cannings and Whittaker (1994, J. Theor. Biol., 167, p. 397) studied a modification of the model by Maynard Smith, similar to the one we present. They suggest a mechanism that implies more games for players that finish faster, but keep the explicit time cost. Unlike our model, their approach is restricted to positive integer waiting-times. Here, we are able to apply our results in the limit of long games and to calculate the stationary distribution analytically.

Finally, we note that the dynamics of the population during a generation modelled here can be useful for studies of game-theoretic problems in general; one example could be the study of the Prisoner's Dilemma game with refusal in which a player may quit a repeated game when encountering a deviation from cooperation. Here, the threat to abandon is equivalent to an outside option, where the value of the option is again implicit: it depends on the composition of the population.

Modeling, inference and simulation of biological networks using Constraint Logic Programming (CLP)

Eric Fanchon (1), Fabien Corblin (1,2) and Laurent Trilling (2)
(1) Institut de Biologie Structurale, CNRS-CEA-U. Joseph Fourier, 41 rue J. Horowitz, 38027 Grenoble Cedex 1. (2) LSR-IMAG, U. Joseph Fourier, BP 53, 38041 Grenoble Cedex 9.

Biology is now entering a new era in which molecular components have to be integrated into a *system* in order to reach new levels of understanding. Our objective consists in developing a computing tool allowing on one hand to infer models from properties which can be incomplete and qualitative, on the other hand to perform simulations or predictions starting from these (partially known) models. Such a tool should allow biologists to specify a network from the available data in order to obtain a class of models consistent with the data. More generally, the tool should be highly flexible to support the *exploration* of model properties in the context of incomplete knowledge.

The concept of interaction network is a fundamental one in systems biology. Our work is based on the "asynchronous multivalued logic networks" proposed by R. Thomas, E. H. Snoussi *et al*. (1,2). This formalism has been used to model genetic, neuronal and immunological networks. Formally, it can be viewed as a discrete abstraction of a special class of Piecewise-Linear Differential Equations (PLDEs). It allows a qualitative analysis of the dynamical behavior of such differential systems. Another benefit of this type of formalism lies in the discreteness which lends itself very well to computational implementations. The interaction graph associated to the PLDE system defines the architecture of the network. The parameters characterize the strength of the (non-linear) interactions. Recently, this formalism has been extended by de Jong *et al.* (3) to take into account the so-called 'singular states' and 'sliding modes'. Singular states are states of reduced dimensionality located at thresholds or intersection of thresholds, and sliding modes are trajectories that slide along a threshold (or intersection of thresholds). This extended formalism is sound in the sense that every continuous trajectory of the original PLDEs is associated to a qualitative (discrete) trajectory of the discrete network.

We show that logic networks of this type can be described formally and exploited via a Constraint Logic Programming (CLP) implementation. The CLP approach rests on the cooperation of solvers on various fields (tree, list, rational, real, boolean). Its advantages are that (i) the implementation is expressed in a very similar way to the *formal specification*, thus guaranteeing the correctness of the implementation, (ii) it is *iterative* - when new information become available, new constraints can be added to reduce further the space of possible models; (iii) many different queries can easily be posed to this formal specification due to its logical form. For example, queries equivalent to simulation (parameters known / computation of behavior) as well as inference of model parameters (information on behavior / computation of parameter values). Situations that are intermediate between simulation and inference are frequent. Indeed, the experimental characterization of *behaviors* (trajectories in phase space) is itself often partial, and a current challenge in the field is to be able to exploit all available partial knowledge to get more precise models.

These principles are applied to the study of adhesion between human endothelial cells. The work is done in collaboration with experimental biologists (4). A submodel extracted from a larger network is presented (2 variables, 7 discrete parameters).

We explain briefly the architecture of the implementation in the declarative language prolog IV (5). A preliminary version of the tool has been published (6) which did not take into account the existence of sliding modes. This was too restrictive and a full implementation is now available. A set of logical predicates defines the discrete *transition rules* corresponding to the type of networks studied (asynchronous multivalued networks with singular states). A given network is described by a set of discrete equations and a set of inequalities between parameters; these entities are derived from the

architecture of the given network (number of nodes/genes and pattern of interactions: activation or repression of gene g_i by gene g_j with threshold θ_{ij}). Observational knowledge is also described by constraints (logical predicates). This can be a direct measurement of a kinetic parameter, or knowledge about the behaviour of the system, such as, for example: "when the system is perturbed and set into state S_p, it returns to stable state S_0 by going through at least one state in which the concentration of such protein P is above such threshold θ". As illustrated in this example, this knowledge can be incomplete. It can nevertheless be formalized into a logical expression (after discretization) and exploited to make deductions about, for example, the possible values of the model parameters. As said in (ii) above, each new observation allows to add a new constraint which, in general, reduces the space of solutions. Likewise, hypotheses can be expressed as formal prolog queries in order to test their consequences. This provides a flexible tool to query model properties or, more generally, properties of a given network architecture.

In the cell adhesion study, we exploit *behavioral* information resulting from the observation of the response of the cell culture after a perturbation. To illustrate the strength and the flexibility of the CLP approach, results will be commented concerning : some general properties of the model; the existence of stationary states; and the use of behavioral information to reduce the space of possible models. In particular, it is shown that imposing the existence of a path from the perturbed state to the adherent state eliminates a large number of models. If times permit, results from larger published models of developmental biology will also be presented: segmentation of the drosophila embryo (7) and the drosophila gap-gene system (8).

(1) E. H. Snoussi, Dyn. Stab. Sys., **4**, 189 (1989).

(2) R. Thomas and M. Kaufman. Multistationarity, the Basis of Cell Differentiation and Memory. II. Logical Analysis of Regulatory Networks in Term of Feedback Circuits. Chaos, **11**, 180-195 (2001).

(3) H. de Jong, J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Qualitative Simulation of Genetic Regulatory Networks Using Piecewise-Linear Models. Bulletin of Mathematical Biology, **66**, 301-340 (2004).

(4) B. Hermant, S. Bibert, E. Concord, B. Dublet, M. Weidenhaupt, T. Vernet and D. Gulino-Debrac. Identification of Proteases Involved in the Proteolysis of Vascular Endothelium Cadherin during Neutrophil Transmigration. The Journal of Biological Chemistry, **278**, 14002-14012 (2003).

(5) A. Colmerauer. Prolog - Constraints Inside, Manuel de Prolog, PROLOGIA, Case 919, 13288 Marseille cedex 09, France (1996).

(6) E. Fanchon, F. Corblin, L. Trilling, B. Hermant and D. Gulino. Modeling the Molecular Network Controlling Adhesion Between Human Endothelial Cells: Inference and Simulation Using Constraint Logic Programming. Computational Methods in Systems Biology 2004, V. Danos and V. Schachter (Eds.), LNBI **3082**, 104–118, Springer-Verlag (2005).

(7) Sanchez L, Thieffry D. A logical analysis of the Drosophila gap-gene system. J. Theor Biol., **211**, 115-141 (2001).

(8) Sanchez L, Thieffry D. Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module. J. Theor. Biol., **224**, 517-537 (2003).

CONCENTRATION AND SPECTRAL ROBUSTNESS OF BIOLOGICAL NETWORKS WITH HIERARCHICAL DISTRIBUTION OF TIME SCALES

A.N.GORBAN AND O.RADULESCU

Abstract. We discuss here the robustness of the relaxation time using a chemical reaction description of genetic and signalling networks. First, we obtain the following result for linear networks: for large multiscale systems with hierarchical distribution of time scales the variance of the inverse relaxation time (as well as the variance of the stationary rate) is much lower than the variance of the separate constants. Moreover, it can tend to 0 faster than $1/q$, where q is the number of reactions. We argue that similar phenomena are valid in the nonlinear case as well. As a numerical illustration we use a model of signalling network that can be applied to important transcription factors such as NFκB or TGFβ.
Keywords: *Complex network; Relaxation time; Robustness; Signalling network; Chemical kinetics; Limitation; Measure concentration*

Recent progress in molecular biology showed that the development and the functioning of living organisms are controlled by large complex networks, such as genetic and signalling networks. These networks are dynamical and their time scales distribution is log-uniform, which means that there is an hierarchy of characteristic times.

Some numerical studies [1] emphasized the robustness of gene networks functioning, with respect to changes of the constants. This is important for modeling: it shows that a precise knowledge of the constants is not needed. It also brings understanding on how nature deals with unavoidable variability: the regulation structures are robust.

Very little is known on the origin of robustness. In our conception there are two intrinsically related sources of robustness. One has to do with size and concentration of measure on high dimensional metric–measure spaces [3, 2]. The second is related to topology and hierarchy of time scales.

We discuss here the robustness of the relaxation time using a chemical reaction description of genetic and signalling networks. Relaxation time is an important issue in chemical kinetics. It is so for practical reasons because it says how long one has to wait until the end of a process. In biology, the reasons are slightly different. A biological system is a hierarchically structured open system. Any biological model is necessarily a submodel of a bigger one. After a change of the external conditions, a cascade of relaxations takes place and the spatial extension of a minimal model describing this cascade depends on time. It is therefore important to know how the relaxation time depends on the size of the model and how robust is this against variations of the kinetic constants.

First, we obtain some results for linear networks. Let us enumerate reactions in the order of their constants decrease: $k_1 > k_2 > \ldots > k_q$. If kinetic constants are all well separated then we can use $\gg$ instead of $>$. The reaction graph is weakly

ANG: University of Leicester, Leicester, LE1 7RH, UK, e-mail: ag153@le.ac.uk.
OR: IRMAR, UMR CNRS 6625, Université de Rennes 1, France,
e-mail: ovidiu.radulescu@math.univ-rennes1.fr.

ergodic (further we omit the adverb "weakly"), if for each two vertices (components) A_i, A_j $(i \neq j)$ we can find such a vertex A_k that oriented paths exist from A_i to A_k and from A_j to A_k. One of these paths can be degenerated: it might be $i = k$ or $j = k$. The reaction constant k_r $1 \leq r \leq q$ is the *ergodicity boundary* if the reaction graph for reactions with constants $k_1, k_2, \ldots, k_r$ is ergodic, but for reactions with constants $k_1, k_2, \ldots, k_{r-1}$ it is not. For the relaxation time τ of the whole system the following estimate holds:

$$\overline{a}\frac{1}{k_r} \geq \tau \geq \underline{a}\frac{1}{k_r}, \tag{1}$$

where $\overline{a}, \underline{a} > 0$ are some positive functions of $k_1, k_2, \ldots, k_{r-1}$ (and of the reaction graph topology [4]).

The well known concept of stationary reaction rates *limitation* by "narrow places" or "limiting steps" should be complemented by the *ergodicity boundary* limitation of relaxation time. It should be stressed that the relaxation process is limited not by the classical limiting steps (narrow places), but by absolutely different reactions. The simplest example of this kind is a catalytic cycle: the stationary rate is limited by the slowest reaction (the smallest constant), but the ergodicity boundary is the reaction constant with the second lowest value. In order to change the slowest relaxation time one should coordinately alter the lowest and the second lowest constant.

In general, for large multiscale systems we observe concentration effects: the variance of the inverse relaxation time (as well as the variance of the stationary rate) is much lower than the variance of the separate constants. Moreover, here we meet a "simplex–type" concentration ([2], pp. 234–236) and the variance of the relaxation time can tend to 0 faster than $1/q$, where q is the number of reactions. For simplest linear reaction mechanisms with random constants k the estimate $\mathrm{Var}(1/\tau) \sim \mathrm{Var}(k)/q^2$ is proven.

We argue that similar phenomena are valid in the nonlinear case as well. As an illustration we use a rather generic model of signalling network that can be applied to important transcription factors such as NFκB or TGFβ. The model consists of five reactions:

(1) $R + F \leftrightarrow C$
(2) $C + K \leftrightarrow F$
(3) $2F^* \rightarrow 2F^* + R$
(4) $R \rightarrow$
(5) $F \leftrightarrow F^*$

F is a transcription factor that forms a complex C with the repressor R. The complex is localized in the cytosol. The signal is represented by a kinase K that phosphorylates the repressor and frees the transcription factor. Nuclear F^* comes from cytoplasmic F by transport and controls the transcription of various genes among which the repressor R.

In order to study the robustness of the system, we have performed five operations. In each one of these operations the reaction constants k_i, k_i^r of the forward and of the reversed reaction i have been divided by a scale factor. The effects of these operations on the value of F^* at stationarity and on the relaxation time τ have been represented in Fig.1 in the presence and in the absence of a signal. Although the constants have been changed on four decades, the relaxation time have large plateaus where it is constant and its total variation is smaller than two decades. Without signal, robustness is less pronounced.

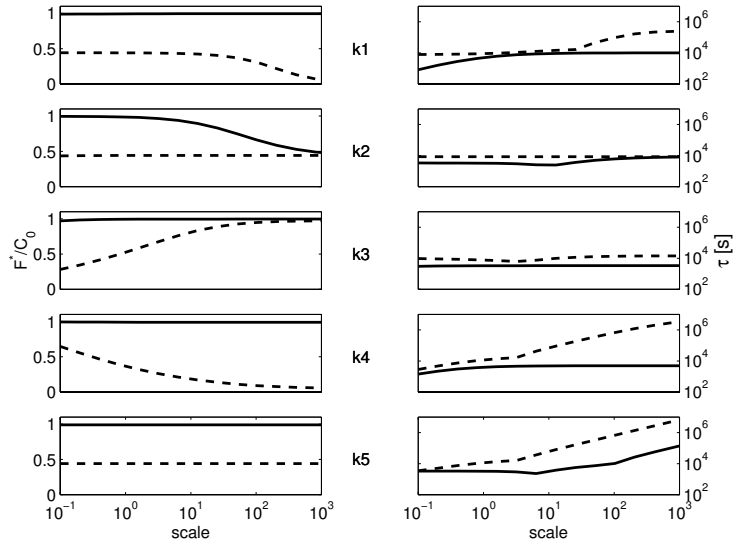


FIGURE 1. Robustness of the steady state and of the relaxation time τ of the signalling network. The response to the signal is the concentration F^* of transcription factor in the nucleus. Dynamics evolves on conserved hyperplanes $F + F^* + C = C_0$ according to the reactions described in the text and to the law of mass action. The presence ($K \neq 0$) and the absence ($K = 0$) of a signal are represented by a continuous and a dotted line, respectively. Each subfigure correspond to a different reaction whose constants were divided by *scale*. Unscaled constants are those of the reference [5].

Obviously, results as general as Eq. 1 do not work in the nonlinear case. For instance, the relaxation time diverges near a saddle-node bifurcation point. In this case there are no concentration effects. Our simple model suggests that with some restrictions, concentration of the relaxation time might work in the nonlinear case as well. We do not know yet which are the restricting conditions and how to connect these effects to topology.

The observed phenomena can give a clue to robustness of relaxation characteristics of multidimensional networks with hierarchical distribution of time scales. It suggests that for systems with wide distributions of reaction rate constants, the relaxation time of the whole system is much more stable than the relaxation times of individual small fragments. In particular, the relaxation time of the whole system is much more stable than the relaxation times of the individual reactions.

REFERENCES

[1] Von Dassow G. et al., The segment polarity network is a robust developmental module, Nature, 406 (2000), 188–192.

[2] Gromov M., Metric structures for Riemannian and non-Riemannian spaces, Birkhauser Boston, Inc., Boston, MA, 1999.

[3] Talagrand M., Concentration of measure and isoperimetric inequalities in product spaces, Inst. Hautes Etudes Sci. Publ. Math. No. 81 (1995), 73–205

[4] Gorban A.N., Bykov V.I., Yablonskii G.S., Essays on Chemical relaxation, Novosibirsk, Nauka Publ., 1986.

[5] Hoffmann A., Levchenko A., Scott M.L., and Baltimore D., "The IκB-Nf-κB signaling module: temporal control and selective gene activation", Science, 298 (2002), 1241–1245.

Dynamics and pattern formation in invasive tumor growth

Evgeniy Khain, Leonard M. Sander
ekhain@umich.edu

One of the most common and clinically aggressive forms of primary brain tumor is Glioblastoma Multiforme (GBM). Despite major advances in the fields of molecular biology and cellular biology, the overall prognosis still remains very poor. One of main reasons for such a high mortality and low success in medical treatment is the fact that GBMs are highly invasive. In-vitro experiments show that a growing tumor consists of two zones: an inner dense proliferative region and an outer less dense invasive region. This is the invasive nature of malignant gliomas that makes treatment to be a very difficult and challenging task.

Malignant brain tumors are complex self-organized multicellular biological systems. Experiments with different types of cells show qualitatively different behavior. For wild type cells, the invasive region grows faster, and tumor remains spherically symmetric. On the other hand, the invasive region grows slower for mutant type cells, and there are indications of symmetry-breaking of spherically symmetric growth. We formulate a continuum model that captures these experimental findings, using two coupled reaction-diffusion equations for cells and nutrient concentrations. When the ratio of nutrient and cell diffusion coefficients exceeds some critical value, the plane propagating front becomes unstable with respect to transversal perturbations. The instability threshold and the full phase-plane diagram in the parameter space are determined. Based on our model, we can explain different patterns by different diffusion constants and proliferation rates of wild and mutant cells: wild type cells diffuse faster, but have lower proliferation rate, compared to mutant type cells.

87.18.Ed Aggregation and other collective behavior of motile cells 87.18.Hf Spatiotemporal pattern formation in cellular populations

Reduction of complexity in dynamical systems:

Tri Nguyen-Huu, Pierre Auge, Christophe Lett, Jean-Christophe Poggiale
tnguyenh@ens-lyon.fr

Realistic ecological models must take into account processes which are going on in different levels: the individual, the population, the community level. This leads to mathematical models involving many variables and parameters, which are usually difficult to handle. In many cases, the time scales associated to processes going on at each level are different. At the individual level, the time scale is typically the day, at the population level, the year and at the community level, the evolutionary time scale. Aggregation methods take advantage of these time scales to build a reduced model governing a few global variables at a slow time scale. We present an application to a spatial model of a host-parasitoid community. We consider a square two-dimensional grid of spatial patches. The initial model (complete model) is described by a huge number of equations (20000 for a 100x100 square grid). We show that when the dispersal process becomes fast in comparison with local interactions, the dynamics of the metapopulation can be described by a two-equation model governing the total insect population densities on the grid (aggregated model). We present numerical simulations of both models. Our results show a good agreement between asymptotic behaviour of the complete model and the aggregated model for small differences in time scales. This allows using the aggregated model to make valid predictions about global host-parasitoid spatial dynamics.

Emergent properties of metabolic systems and the effect of constraints on enzyme concentrations

Sicard Delphine, Dillmann Christine, Fiévet Julie, Talbot Grégoire, Grima Laure & de Vienne Dominique

UMR de génétique végétale, INRA/CNRS/INAPG/UPS, Ferme du Moulon, Gif sur Yvette, 91190, France, Phone : +33-01 69 33 22 42, Fax : +33-01 69 33 23 40, Email : sicard@moulon.inra.fr

Cell functioning and evolution rely on complex metabolic systems constituted of many components that communicate and interact with one another through networks. Several metabolic theories have been developed to predict the emergent properties of metabolic systems. However, the effects of constraints on the properties of such systems and on their evolution under selection have been poorly studied, whereas cell necessarily functions with limited resources. Using both theoretical and experimental approaches, we have studied the effect of constraints on enzymes concentrations and their consequences on metabolic fluxes and fitness.

The theoretical developments were based on the metabolic control analysis, which provides a framework linking enzymatic parameters, such as enzyme activities, to a macroscopic output of the system, the metabolic flux. We analysed the effect of competition for space and energy by introducing an overall cost for producing enzymes or by limiting the range of variation of the enzymes concentration in a pathway. In addition, we studied the effect of co-regulation by introducing correlation between enzyme concentrations. Under those conditions, our modelling revealed new emergent properties of metabolic fluxes. First, the total enzyme concentration allocated to a pathway, which is positively correlated with flux, can respond to selection. Second, competition leads to a distribution of enzyme concentrations within the pathway that maximizes metabolic flux: selection can act to increase low enzyme concentrations and to decrease high enzyme concentrations until an optimal level. Third, co-regulation leads to metabolic flux consistently lower than the one obtained with competition alone, suggesting that co-regulation may be costly. Finally, a biochemical model for hybrid vigour can be derived.

In vitro reconstruction of the first part of glycolysis and *in vivo* analysis of various *Saccharomyces cerevisiae* strains were carried out to test these predictions. *In vitro* experiments allowed us to estimate global enzymatic parameters and confirmed that a distribution of enzyme concentrations that optimizes flux can be predicted. “Test tube genetics” performed by varying *in vitro* enzyme concentrations allowed us to corroborate the metabolic mechanism for hybrid vigour. Finally, proteomics and biochemical analysis of a collection of *S. cerevisiae* strains showed that there is genetic variability at two levels of cell integration, enzyme concentrations and glycolytic fluxes, and that selection can act to increase or decrease flux.

In conclusion, taking into account constraints on enzyme concentrations allowed us to develop new modelling and *in vitro* tools for metabolic optimization, and gave new insight in the understanding of metabolic system evolution.

DISCRETE DELAY MODEL FOR THE MAMMALIAN CIRCADIAN CLOCK

K. Sriram[1], Gilles Bernot[1,2] and François Képès[1,3]

1. Epigenomics Project, Genopole®, 93 Rue Henri Rochefort, Evry France-91000

2. Laboratoire de Méthodes Informatiques, UMR 8042 CNRS/Université d'Evry

3. ATelier de Génomique Cognitive, CNRS UMR8071/Genopole®

A circadian rhythm is an oscillation with a period of approximately 24 hr, which exhibits entrainment to environmental light dark (LD) cycles and shifting of phase by light stimulation. Even though many theoretical models with ordinary differential equation (ODE) have been proposed based on the biochemical mechanisms for circadian rhythms [1], relatively few studies have been carried out with delay differential equations (DDE) [2, 3, 4, 5]. Delayed feedbacks are common and occur naturally in many biological systems and in particular, the regulatory networks of circadian rhythms.

Here, we propose a delay model for the circadian rhythm of the mammals [6] with three dynamical variables that has three delayed positive and negative feedback loops. The delayed positive feedback loops modelled by Michaelis-Menten kinetics that describes saturation behavior. The delayed negative feedback loops are modelled with Hill's type of equation that describes a switch like behavior. The form of interlocked positive and negative feedback loop is modelled along the same lines as that of Smolen *et al.* [5]. In formulating the present model, BMAL1 (B), PER-CRY (P) complex and REV-ERBα (R) protein concentrations are considered as the dynamical variables. The biological circuit is shown in Figure-1. The transcriptional activators CLOCK and BMAL1 form a heterodimer which positively regulates *Per, Cry* and *Rev-Erb*α genes. PER-CRY complex is taken as another dynamical variable because PER and CRY expressions are positively coregulated by BMAL1-CLOCK. Their phases are also similar and they both negatively regulate BMAL1 and CLOCK activity. REV-ERBα, the negative regulator of BMAL1 is taken as the third dynamical variable. Thus broadly, there are three negative and positive feedback loops, with BMAL1-CLOCK acting as positive limb and PER-CRY acting as negative limb. The following are the corresponding delay differential equations:

$$\frac{dB}{dt} = \frac{v_1 k_1^{n1}}{k_1^{n1} + R(t-\delta_3)^{n1}} + \frac{v_2 P_f(t-\delta_2)}{k_2 + P_f(t-\delta_2)} - k_3 B \tag{1}$$

$$\frac{dP}{dt} = \frac{v_3 k_5^{n2}}{k_5^{n2} + P_f(t-\delta_2)^{n2}} + \frac{v_4 B(t-\delta_1)}{k_4 + B(t-\delta_1)} - k_6 P \tag{2}$$

$$\frac{dR}{dt} = \frac{v_5 k_7^{n3}}{k_7^{n3} + P_f(t-\delta_2)^{n3}} + \frac{v_6 B(t-\delta_1)}{k_8 + B(t-\delta_1)} - k_9 R \tag{3}$$

Here, P_f is the free PER-CRY complex (P_f = P - B) and $P_f = 0$ if P < B, to account for the interlocked feedback loops between BMAL1 and PER-CRY complex proteins. $v_{1,2,3,4,5,6}$ are the rates at which the proteins are synthesized and the production rate v_3 of PER-CRY complex increases in the light phase. The other parameters are the Michaelis constants $k_{1,2,4,5,7,8}$, the Hill's coefficients, $n_{1,2,3}$ characterizing the degree

of co-operativity of the repression processes; $k_{3,6,9}$ are the first order degradation constants of B, P and R respectively. In the model, the overall time delay for the positive and negative feedback is approximately one circadian cycle. Delayed BMAL1 activation (13hrs) of PER-CRY complex and REV-ERBα constitutes half of the circadian cycle. Delayed activation and suppression of BMAL1 and REV-ERBα respectively by PER-CRY complex and its own suppression (6hr) constitutes one quarter of the cycle. Repression of BMAL1 by REV-ERBα is assumed to be 6hr, which is another quarter of the circadian cycle. In totality, half of the circadian cycle amounts to delay in positive feedback loop (BMAL1 activation) and the other half is the negative feed back loop (PER-CRY and REV-ERBα put together). The interplay of delayed positive and negative feedback loops contribute to one circadian cycle (Figure-2).

There are also other features exhibited by the model. The model shows entrainment to both shorter and

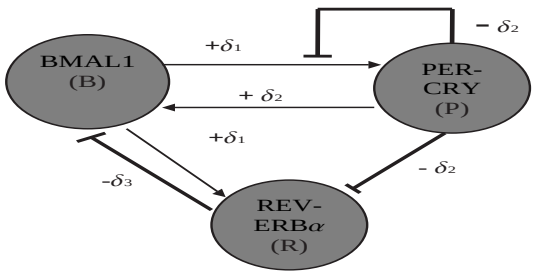


Figure 1: *Schematic representation of the present model for the mammalian rhythm. δ_1 is the delay in the positive feedback from B to initiate the synthesis of PER-CRY protein. The delay δ_1 is also the delayed positive feedback from B to initiate the synthesis of REV-ERBα protein. δ_2 is the delay for to activate and suppress BMAL1 and REV-ERBα protein, respectively. δ_3 is the time delay for REV-ERBα protein to suppress the production of BMAL1.*

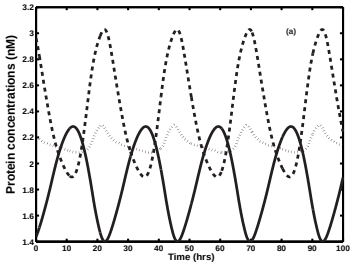


Figure 2: *Sustained oscillations generated by the model. The BMAL1 protein (in black continuous line) is approximately anti phase to proteins PER-CRY protein (black dotted lines) and REV-ERBα (black dashed dotted lines). The time series have been obtained by numerical integration of delay equations 1, 2 and 3 under constant darkness (DD) for the standard parameter set v_s= 4nMh^{-1}, v_d=0.97nMh^{-1}, v_p=1.0nMh^{-1}, v_m=0.7nMh^{-1}, v_r=0.1nMh^{-1}, v_c =1.0nMh^{-1},k_1=0.5nM ,k_2 =2.0nM,k_3 =0.21h^{-1}, k_4 =0.9nM, k_5=0.6nM, k_6 =0.45h^{-1}, k_7 =0.1nM, k_8 =0.1nM , k_9 =0.45h^{-1}, n1=n2=n3=2.0, δ_1=13 hr, δ_2=6 hr, δ_3=6 hr*

longer LD cycles. In all the LD cycles, the oscillator is entrained to 24 hr rhythm for the standard parameter set. When delay δ_2 is varied under LD cycles, the model exhibits phase advance, phase delay and lack of

entrainment, which are linked to physiological disorders. Apart from limit cycle, quasiperiodic and chaotic oscillations are also observed when the delay δ_2 is varied under the influence of constant periodic 12:12 LD cycles. Periodic forcing is known to bring about rich dynamical phenomena [7] and in our model constant periodic forcing with delay brings about a rich bifurcation diagram (Figure-3). The observed complex phenomena such as quasiperiodic and chaotic oscillations are linked to non 24hr sleep-wake syndrome and occurrence of cancer incidence, which may be a direct consequence of improper delayed circadian regulation due to *Per* gene mutation. The effects of mutant phenotype on the circadian period are well simulated by changing the parameters and time delay. The model also uncovers the possible existence of multiple oscillatory networks.

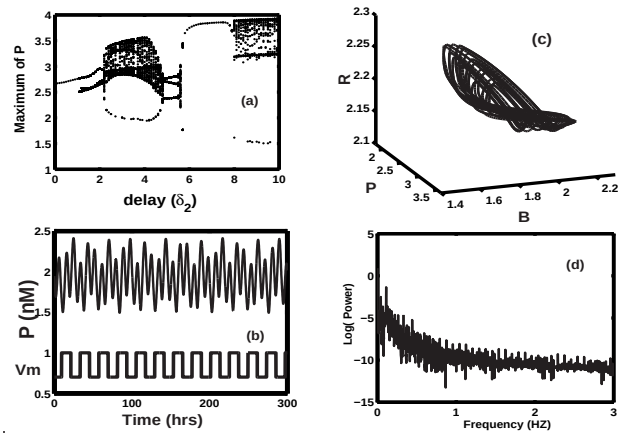


Figure 3: *(a) Bifurcation diagram obtained for the constant LD cycle with delay δ_2 as the parameter. The 12:12 LD cycle simulated with v_m taken a square wave function that is changed from the basal value of 0.7 to 1. (b) Chaotic time series of dynamical variable P, with 12:12 LD cycle for delay δ_2 = 3hrs, (c) the chaotic attractor and (d) the power spectrum. P is the dynamical variable, namely PER-CRY complex. All the other parameters are kept constant with delay δ_1 taken as 12hrs*

References

1. Goldbeter A : Proc R Soc Lond B Biol Sci 1995; 261:319-324
2. Sriram K, Gopinathan MS: J. Theor. Bio. 2004; 231:23-38.
3. Scheper T, Klingenberg D, Pennartz C, Vanpelt J: J. Neurosci. 1999; 19:40-47.
4. Lema MA, Golombek, DA, Echave J: J. Theor. Biol. 2000; 204: 565-573.
5. Smolen P, Baxter DA Byrne JH: J. Neurosci. 2001; 21:6644-6656.
6. Reppert, SW, Weaver, DR: Annu. Rev. Physio. 2001; 63:647-676.
7. Holden AV Chaos: An Introduction., Manchester: Manchester university press, 1985.

Self-organisation and other *emergent* properties in a simple biological system of microtubules.

James Tabony

Commissariat à l'Energie Atomique, Département Réponse et Dynamique Cellulaires, Laboratoire d'Immunochimie, INSERM U548, DSV, CEA-Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France. E-mail tabony@cea.fr; telephone 04 38 78 96 62; fax 04 38 78 96 62

Key words: biological self-organisation, cell biology, microtubules, reaction-diffusion, bifurcations and triggering factors, particle transport and organisation, positional information, replication of form; *complex* systems and *emergent* phenomena.

Running title: Microtubule self-organisation

Abstract.

In biological systems, *emergent* properties develop when numerous individual molecular elements in a population are coupled in a non-linear manner. Under suitable conditions, the formation *in vitro* of a population of microtubules, a major component of the cellular skeleton (cytoskeleton), behaves as a *complex* system and develops a number of *emergent* phenomena. These preparations, which initially contain just two molecular species, a nucleotide and a protein, self-organise by reaction and diffusion and the morphology that develops is determined at a critical moment early in the process by weak external factors such as gravity and magnetic fields. Other *emergent* phenomena; namely, replication of form, generation of positional information, and collective transport and organisation of colloidal sized particles, develop. Microtubules are responsible both for cellular organisation and the transport of sub-cellular particles from one part of the cell to another. Frequently, this behaviour is triggered by some weak internal or external factor. The *in vitro* observations outlined illustrate how in a simple biological system, a *complex* behaviour may give rise to *emergent* phenomena outwardly resembling major biological functions.

Over the last hundred years, enormous advances in biology have been made based upon the general concept of molecular reductionism. This has found its apogee in the molecular basis of DNA and RNA function and more generally with the association of molecular structure with function. The implication is that knowledge of the exact disposition of atoms in a molecule will eventually (if established for a sufficiently large number of proteins and other macromolecules) lead to a description of living systems. Many biologists make substantial efforts to identify individual molecular agents involved with a specific biological function. Molecular reductionism is, however, subject to limitations and most biologists realise that some biological properties or functions somehow arise simply because a large number of interacting molecular species are present. To anyone familiar with non-linear dynamics, this suggests that populations of biological molecules might behave as *complex* systems and develop *emergent* phenomena. Some scientists are asking whether some of the global properties of biological systems can be accounted for in terms of *emergent* properties and even whether life itself should be considered as such.

Until recently, biologists have played little attention as to the manner by which populations of specific biological molecules might behave as *complex* systems. A major *emergent* phenomenon in many *complex* systems is self-organisation. Here, I would like to outline a very simple biological system of a population of microtubules in a test-tube, comprised initially of just two molecular species (a nucleotide and a protein) that behaves as a *complex* system and shows a number of *emergent* phenomena including self-organisation.

1. Self-organisation by reactive processes

The 2nd law of thermodynamics teaches us that at equilibrium order will be progressively and ineluctably lost with time. Two miscible liquids, initially separated from one another, will slowly mix by way of diffusion and convection and the existing order progressively lost. One of the characteristic properties of living systems is order and self-organisation. This naturally raises the question as to the nature of the physical-chemical processes by which order and form spontaneously develop in an initially largely unstructured biological object such as an egg or seed. Biological processes are based upon biochemical and chemical reactions. However, normally solutions of reacting chemicals in a test-tube do not self-organise. Because of this, for very many years, it was not believed possible that solutions of reacting chemicals or biochemicals could self-organise by reactive processes.

Nevertheless, very slowly over the last hundred years, both theoreticians and experimentalists have progressively shown that this is not necessarily the case. Since the late 1930's, some theoreticians (Kolmogorov, Rashevsky, Turing, and Prigogine and co-workers) [1-5] have proposed that some particular types of chemical reaction might, due to being sufficiently far-from-equilibrium, show strongly non-linear reaction dynamics. They predicted that these non-linear dynamics could, in some cases, result in a macroscopic self-organisation of the sample. Some chemical systems originally discovered in the 1920's [6] and 1950's [7] have been shown to self-organise this way [8, 9]. At a molecular level, self-organisation results from a coupling of reaction and diffusion and the patterns that arise are comprised of periodic variations in the concentration of some of the reactants. Such structures are often called reaction-diffusion or Turing structures; the latter after the British mathematician who was one of the first persons to propose such a mechanism in 1952 [3]. Prigogine and co-workers called them 'dissipative' structures [5, 10] because a dissipation of chemical energy is required to drive and maintain the system sufficiently far-from-equilibrium such that self-organisation occurs. It is this flux or dissipation of chemical energy that provides the thermodynamic driving force for self-organisation. Rashevsky, Turing, Prigogine et al, and others, all proposed that biochemical mechanisms of this type might provide an underlying physical chemical explanation for biological pattern formation and morphogenesis. Although such terms were not used at the time, what these theoreticians predicted was that biological self-organisation could arise as an *emergent* phenomenon in a *complex* system by molecular processes of reaction and diffusion.

A different aspect of these systems is the manner by which some reaction-diffusion systems may show bifurcation properties and can hence be sensitive to weak external factors. In the early 1970's, Kondepudi and Prigogine explicitly calculated that the presence of an external factor, such as

gravity, or an electric or magnetic field, at a critical moment early in the process, might determine the self-organised morphology which subsequently develops [11].

The pioneer workers in this field were fully aware of the possible implications that their approach might have towards some problems in biology, and the concepts outlined above have aroused interest and debate at various times over the last 50 years. However, for a variety of reasons, the majority of biologists and chemists have not adopted this approach. Although the main reason is conceptual, another reason is the scarcity of simple experimental systems proven to self-organise this way. For example, in chemistry, it was not until 1990 that a chemical reaction, similar to those first discovered long ago by Bray (1921) [6] and Belousov (1951) [7] was finally accepted as the first example of a Turing-like structure [8, 9].

The same situation has prevailed in biology. Since the work of Turing and Prigogine and co-workers, many authors have compared the morphologies that occur in biological organisms with the mathematical predictions of reaction-diffusion theories. There is a whole body of literature in this area [12-14]. More recently, other workers [15, 16] have demonstrated that the patterns of calcium waves observed *in vivo* in the cytosol arise from reaction-diffusion processes. In spite of these advances, one of the elements lacking has been an example of a simple biochemical system in a test-tube that self-organises this way.

Under suitable conditions, we have found that the *in vitro* formation of microtubules, a major component of the cellular skeleton (cytoskeleton), does behave this way. These preparations, which initially contain just two molecular species (a nucleotide and a protein) self-organise by reaction and diffusion and the morphology that develops is determined at a critical moment early in the process, by weak external factors such as gravity and magnetic fields. This behaviour is not a result of the sum of the properties of individual microtubules and cannot be understood in terms of molecular reductionism. On the contrary, it arises from the collective action of the entire microtubule population in which individual microtubules interact and communicate with one another by way of the chemical trails that they themselves form. These observations illustrate how in a simple biological system, reactive processes give rise to a population of interacting elements that behaves as a *complex* system and shows a number of *emergent* phenomena as a consequence. In addition to self-organisation, and its triggering by weak external factors, a number of other collective phenomena develop; namely, replication of form, generation of positional information, and collective transport and organisation of colloidal sized particles. These *emergent* phenomena, outwardly at least, resemble the major biological properties of microtubules and they may turn out to be of considerable biological significance.

2. Self-organisation in colonies of living organisms

Colonies of living organisms provide many examples of self-organisation [17]. In many cases, structure and organisations develop, not by action at the level of the individual, but by way of the dynamics in which individuals, strongly coupled to one another in a non-linear manner, behave as a collective ensemble. Similar types of morphology often develop in spite of large differences in the nature and size of the individual element. Striped arrangements frequently arise; when they do, they are nearly always the result of an outside external perturbation that induces a directional bias on the actions of the individual. For example, over a distance scale of several centimetres some bacteria colonies form a stationary pattern. Observations at higher magnification show that the pattern is comprised of regions containing differing bacterial densities. At even higher magnification, the individual bacteria are seen to be undergoing a rapid, seemingly random movement. Individual bacteria interact indirectly with one another via trails in the concentration of chemical attractants and repellents that they themselves produce. It is by way of dynamic processes involving the collective movement of many bacteria that the stationary pattern arises. The energy source driving this process is the chemicals consumed by the bacteria. When it runs out, the bacteria stop moving and the pattern disappears.

A similar mechanism is the basis for the self-organisation of ant colonies and other social insects. The behaviour of the population results essentially from the actions of individuals strongly coupled to one another by a form of chemical communication [17, 18]. A moving ant leaves behind itself trails of chemicals known as pheromones that attract or repel other ants. An ant encountering a

trail of attractive pheromone will change its direction to follow this trail. This ant, will in its turn, deposit more pheromone on the trail thus reinforcing it. The progressive reinforcement of these chemical trails leads to the self-organisation of the ant population. Although the rules governing the behaviour of individual ants are relatively simple, the overall behaviour is extremely sophisticated.

One of the advantages of this type of process is that ants rapidly establish the shortest route between a food source and the nest. Consider a situation where there are two food sources close to a population of ants; but where one of the food sources is closer than the other. As ants return to the nest with food, they leave behind themselves chemical trails. These trails are then followed by other ants who in their turn, deposit chemicals that reinforce the original trails. In such a way, progressively more and more ants follow the paths to the food sources. However, because the trail from the closer of the two sources is shorter, it takes less time for an ant to return to the colony. This results in a slightly larger number of ants taking the path to this food source, thus reinforcing the strength of the chemical trail of the shorter path at the expense of the longer path. Hence, progressively more and more ants take the shorter path to the closer food supply until they nearly all follow this route. This illustrates how self-organisation results from the progressive reinforcement of chemical trails by moving objects which themselves produce these trails. If, the two food sources are at approximately equal distance from the nest, then the ants still mostly accumulate on the path to one of the food sources. This comes about because any small factor which early in the process favours the reinforcement of one of the chemical trails over the other, will progressively lead to nearly all the ants using this pathway. Once the reinforcement of one pathway has gone sufficiently far, then the determining factor may be removed without affecting the subsequent behaviour. This is a simple example of a bifurcation due to a weak external factor in a self-organising *complex* system.

3. Microtubules

Microtubules [19, 20] are a major filamentary component of the internal skeleton of cells (cytoskeleton). They have two major cellular roles; they organise the cell interior, and they permit and control the directional movement of intracellular particles and organelles from one part of the cell to another. Microtubules participate in many fundamental cellular functions including the maintenance of shape, motility, and signal transduction. They frequently organise or reorganise in response to weak internal and external stimuli of either physical or biochemical nature. Microtubules are a significant component of brain neurone cells and they make up the mitotic spindles that separate chromosomes during cell division. They play a determining role in the organisational changes that occur during the early stages of embryogenesis. Microtubule organisation is a fundamental cellular property affecting numerous biological functions and the viability of a cell is compromised when it does not occur correctly.

Microtubules are long tubular shaped supra-molecular assemblies with inner and outer diameters of about 16 nm and 24 nm respectively. Although their length is variable, they are often several microns long. The walls of the tube are comprised of a protein, tubulin, and microtubules arise from the self-assembly of this protein by way of reactions involving the hydrolysis of a nucleotide, guanosine triphosphate (GTP), to guanosine diphosphate (GDP). Once microtubules form in this way, they continually grow and shrink by processes in which additional tubulin molecules are added to one end of a microtubule whilst other tubulin molecules are lost from the opposite shrinking end. This process is likewise associated with the hydrolysis of GTP to GDP. The system is hence chemically irreversible and there is a continual consumption and dissipation of chemical energy. Biologists have established in living cells that microtubule organisation and reorganisation results from the chemical dynamics of the reactive processes associated with their formation and maintenance.

Microtubules can be readily formed and studied *in vitro*. A solution of purified tubulin, in the presence of an excess of GTP, when warmed from about 4°C to 36°C, assembles within a few minutes into microtubules. After the microtubules have formed, this reaction continues by processes in which the complex, tubulin-GTP, is added to the growing (+) end of a microtubule and tubulin-GDP is lost from the opposite shrinking (-) end. An unusual and important feature of microtubules is that they posses a reactive polarity, and the reaction dynamics at opposite ends of the microtubule are different. Due to this difference in reactivity, microtubules often grow from one end (+) whilst shrinking from

the other end (-). When the rates of growth and shrinking are comparable, individual microtubules retain the same approximate length but change position at speeds of several μm per minute. This type of behaviour is termed 'treadmilling'. Another type of behaviour termed 'dynamic instability' occurs when individual microtubules either shrink or grow very abruptly. By modifying experimental conditions, such as buffer composition, it is possible to observe *in vitro* a very large range of microtubule reaction dynamics.

A shrinking microtubule is capable of forming a trail of free tubulin. This tubulin is initially liberated in the form of the complex, tubulin-GDP. This progressively diffuses out into the solution. Simultaneously, excess GTP present reconverts the tubulin-GDP to tubulin-GTP. At this point, the tubulin-GTP can be incorporated into the growing ends of neighbouring microtubules. Because the incorporation of tubulin into the growing ends of microtubules increases strongly with tubulin-GTP concentration, neighbouring microtubules will preferentially grow into regions of higher tubulin-GTP concentration whilst avoiding those of lower concentration. Hence, for some types of microtubule reaction dynamics (and rates of tubulin diffusion) neighbouring microtubules can communicate with one another, and modify their rate and direction of growth, by way of the chemical trails that they themselves produce. In this way, a population of microtubules is capable of behaving as a *complex* system. It can self-organise and generate other *emergent* phenomena in a manner that shows many analogies with the way that ants and other social insects self-organise.

4. Microtubule self-organisation and other emergent phenomena

4.a. Self-organisation

Under many conditions, microtubule solutions show neither temporal nor spatial self-organisation. However, in 1987 it was reported that they could show regular damped oscillations of assembly and disassembly [21]. In 1990, we reported experiments under different buffer conditions in which macroscopic self-ordering occurred [22]. When assembled in glass containers, measuring 40 mm by 10 mm by 1 mm, the microtubule solution progressively self-organises over approximately 5 hours to form a series of periodic horizontal stripes of about 0.5 mm separation. Once formed, the striped pattern remains stationary and it is stable for between 48 to72 hours, after which the system progressively runs out of reactants. In each striped band, all the microtubules are very highly oriented with respect to one another. The direction of orientation is at about either 45° or 135° to the direction of the stripe, but adjacent stripes differ from one another in having a different orientation from their neighbours. Hence, the microtubule orientation flips from left to right periodically up the length of the sample container. In addition to this orientational pattern, a pattern of variations of microtubule concentration is also present that coincides with the changes in orientation [23]. The microtubule concentration drops by about 30% and then rises again each time the microtubule orientation flips from acute to obtuse or vice versa (Figure 2). Experiments, that will not be described here, show that self-organisation contains both reactive and diffusive contributions and arises from processes involving the continual growth and shrinking of individual microtubules [24-26].

4.b. Replication of form

The structure is complicated, for each 0.5-mm stripe also contains within it another series of stripes of about 100 μm separation. These, in their turn, contain other sets of stripes of about 20 μm, 5 μm and 1 μm separation [24, 27]. In samples made up in a 15 mm diameter test tube, an additional level of ordering of several mm arises. These large stripes contain the lower levels of organisation already mentioned. Hence, similar types of pattern spontaneously arise over distances ranging from a few microns up to several centimetres. So here, we already see two emergent phenomena; self-organisation and replication of form.

The range of dimension over which these microtubule structures occur is typical of those found in many types of higher organisms. Cells are about 10 μm in size, eggs are often about 1 mm, and a developing mammalian embryo is several centimetres long. Self-organisation also arises when

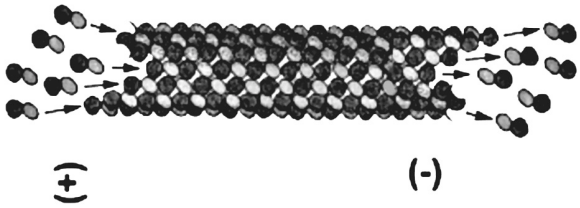


Figure 1 Schematic illustration of a microtubule growth and shrinking". Tubulin-GTP is added to the growing end of a microtubule (+) and tubulin-GDP is lost from the other shrinking end (-). During this process, GTP is hydrolysed to GDP.

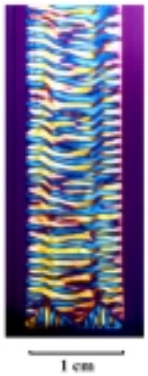


Figure 2. Self-organised microtubule structures as formed in optical cells, 40 mm by 10 mm by 1 mm, positioned vertical. Microtubules were formed by warming a solution containing 10mg/ml of tubulin from 4°C to 36°C in the presence of an excess of GTP. Microtubules form within 2-3 minutes after warming the solution, and the structure shown progressively develops over the next 5-6 hours. The structure once formed is stationary and the solution is stable for about 3 days. The strong optical birefringence indicates that the microtubules are highly aligned. The structure is photographed through crossed polars (0° and 90°) with a wavelength retardation plate at 45°. The retardation plate produces a uniform mauve background. Microtubule orientations of about 45°, such that their birefringence adds to the birefringence of the wavelength plate, produce a blue wavelength shift whereas orientations at about 135° subtract from the birefringence and result in a yellow interference colour. The alternating blue and yellow stripes arise from periodic variations in microtubule orientation from obtuse to acute.

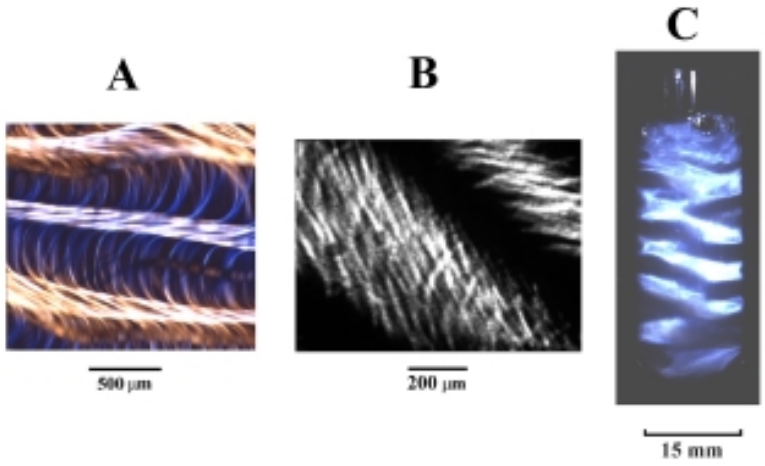


Figure 3. Replication of form. The striped structure, as shown in figure 2, is itself comprised of stripes of smaller periodicity. Photographs A) and B) show one of the individual stripes at higher magnification. Separations of approximately 100 µm and 20 µm are clearly visible. C) is a photograph of the structure that forms in a 15 mm diameter test tube.

samples are prepared in small containers (50-200 μm) of dimensions comparable to those of cells and embryos.

4.c. Positional information

Another feature of these structures and organisations is that they contain a considerable amount of positional information. This is clearly seen in the self-organised morphology, shown in Figure 4, in which the pattern has a clearly defined centre. Moreover the centre of the pattern is positioned in the centre of the sample. Thus, in some way or other, the microtubules have worked out where the centre of the sample is. In addition, the positional information thus produced is expressed and manifested in a clear-cut manner. The generation of positional information is a basic phenomenon underlying embryogenesis and biological pattern formation. Its creation by reactive processes in a simple *in vitro* preparation, initially devoid of it, is an important feature of the observed behaviour.

4.d Dependence of self-organisation on gravity.

The difference in conditions leading to the two different morphologies shown (figures 2 and 4) is merely the orientation of the sample with respect to gravity during self-organisation. Striped morphologies occur when the microtubules are prepared rectangular sample cells that are upright, but concentric circles arise when they are prepared in the same containers lying horizontal, flat down [28]. This fact indicates that gravity in some way intervenes in the self-organising process. Once formed, the structures are stationary and independent of their orientation with respect to gravity. To establish at what moment during self-organisation the sample morphology depends on the gravity direction, we carried out the following simple experiment [24]. Twenty samples of purified tubulin in the presence of GTP (at 4°C) were placed vertical. Samples were simultaneously warmed to 36°C to instigate microtubule formation. Consecutive cells were then turned from vertical to horizontal at intervals of one minute and left in this position for the rest of the self-organising process and examined 12 hours later after the structures had formed. Twenty minutes after instigating microtubule formation, when the last sample was rotated from vertical to horizontal, there are no obvious signs of any striped structure. Since the structures form while the cells are flat, one might expect that they would all form the horizontal pattern. This is the case for samples turned during the first few minutes. However, samples which were upright for six minutes or more all formed striped morphologies similar to preparations that remained vertical all the time. The final morphology depends upon whether the sample container was horizontal or vertical, at a critical time six minutes after instigating assembly, early in the self-organising process. This can be described as a bifurcation between pathways leading to two different morphological states, and in which the direction of the sample with respect to gravity determines the morphology that subsequently forms.

An obvious question is; what would happen if gravity was not present at the bifurcation time? To answer this we carried out an experiment under conditions of weightlessness produced in a free-falling rocket of the European Space Agency. This produced conditions of weightlessness for the first 13 minutes of the self-organising process. We found that, contrary to reference samples assembled on an on-board 1g centrifuge, samples assembled under conditions of weightlessness did not self-organise [27]. This result shows that under the conditions used in this experiment, the presence of gravity at the bifurcation time actually triggers self-organisation. To study the effects of weightlessness, it is not necessary to go to the trouble, expense, and risk-to-life, of carrying out experiments in space. Gravity effects may be substantially reduced in ground-based laboratories using simple inexpensive methods such as clinorotation and magnetic levitation. We have also carried out experiments using these methods and observed behaviour very close to that obtained in space-flight [29]

4.e. Proposed molecular mechanism

In far-from-equilibrium systems that self-organise, bifurcations are associated with an instability in the initially homogenous state. When self-organisation arises from a chemical processes, as in the present case, then this instability will involve reactive elements. For the microtubule case, we would hence expect a chemical instability, involving the relative concentrations of microtubules and free tubulin, to occur close to the bifurcation time. This is the case. Frequently, the kinetics of microtubule self-assembly, after an initial increase due to the formation of microtubules from the tubulin solution, remains at a stationary level. In general, microtubule solutions showing this type of

Figure 4. The morphology that forms is dependent on the gravity direction. A different stationary morphology forms when the sample container is positioned horizontal during self-organisation. For this morphology, the centre of the sample is determined by the centre of the pattern. This illustrates the generation of positional information.

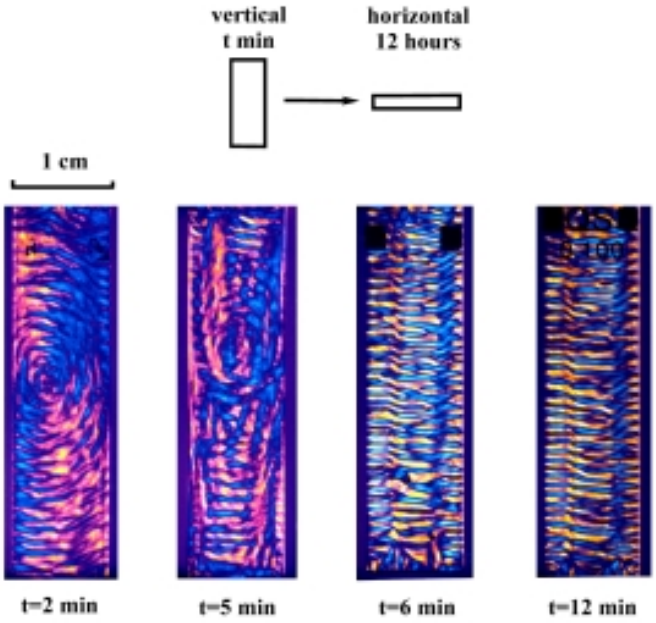


Figure 5. Bifurcation behaviour of self-organised microtubule preparations. The morphology that forms is determined by the gravity direction at a critical time early in the process.. The photographs show the final stationary morphologies for samples rotated from upright to horizontal at different times, t, during the first twenty minutes of self-organisation. Samples that remained vertical for 6 minutes or more formed striped structures as though they had remained vertical throughout the entire period of structure formation

behaviour do not self-organise. However, microtubule preparations that do self-organise do not show this type of assembly kinetics. Instead, after an initial rapid increase, corresponding to the formation of microtubules from tubulin, the microtubule concentration shows an overshoot and progressively decreases over the next 30 minutes to a value about 20% lower than at the maximum [24, 30]. The maximum in the microtubule concentration occurs approximately six minutes after instigating microtubule assembly, and coincides with the bifurcation time when self-organisation is determined by gravity.

Microtubule self-organisation depends not only upon the presence of gravity at an early critical moment. It also depends on other weak external factors, such as magnetic fields, sheering and weak vibrations and geometrical factors [30, 31]. These experiments strongly suggest that any factor, which at the bifurcation time, leads to a privileged direction of microtubule orientation, will trigger self-organisation. This conclusion provides an important clue to the molecular mechanism by which self-organisation occurs. Microtubules are continually growing from one end and shrinking from the other. For appropriate values of reaction dynamics, the shrinking end of a microtubule will leave behind itself a chemical trail of high tubulin-GDP concentration. Excess GTP in the reaction mixture then converts tubulin-GDP back to tubulin-GTP. At this point, the tubulin-GTP is again available either to be incorporated in the growing end of a neighbouring microtubule, or to nucleate with other tubulin-GTP molecules to form a new microtubule. During this time, the tubulin freely diffuses into the surrounding solution. Likewise, growing microtubule ends produce regions depleted in tubulin-GTP. Because reactions rates increase with increasing concentration, neighbouring microtubules will preferentially grow into regions of high tubulin-GTP concentration whilst avoiding those of low concentration. We postulated that for appropriate reaction dynamics, the chemical trails produced by individual microtubules, can modify and determine the direction of growth of their neighbours [27]. Thus neighbouring microtubules will "talk to each other" by depleting and accentuating the local concentration of active chemical. Under such circumstances, the coupling of reaction with diffusion will progressively lead to macroscopic variations in microtubule orientation and concentration.

When the microtubules first form from the tubulin solution, they are in a phase of growth and are distributed uniformly through the solution in an isotropic manner. At this stage, there is almost no disassembly from their shrinking ends. However, the rapid initial growth of the microtubules depletes the concentration of free tubulin in solution and this in turn provokes the partial disassembly of the microtubules. This partial disassembly manifests itself as the 'overshoot' in the assembly kinetics. When partial disassembly starts to occur, just prior to the bifurcation time, it leads to the formation of the chemical trails outlined above. The isotropic arrangement of microtubules is now unstable, for at this time, orienting just a few microtubules will induce their neighbours to grow along the same orientation. Once some microtubules have take up a specific orientation, then neighbouring microtubules will also grow into the same direction. Orientational order will then spread from neighbour to neighbour, and so on. The process mutually reinforces itself with time and leads to self-organisation. Hence, in agreement with experiments, any small factor that at the instability (bifurcation time) directly orients microtubules, or leads to a privileged direction of microtubule growth, will trigger self-organisation.

4.f. Numerical simulations

To investigate whether such an explanation is realistic we carried out computer simulations of a population of growing and shrinking microtubules, incorporating microtubule reaction dynamic consistent with experimental values, [32, 33]. Simulations involving just a few microtubules, demonstrated both the formation of the tubulin trails outlined above and the growth of neighbouring microtubules into these trails, along their direction. When the simulations were extended to a population of about 10^4 microtubules on a two-dimensional reaction space, 100 μm by 100 μm, then after 2-3 hours of reaction time, a self-organised structure comprised of regular bands of about 5 μm separation developed [32, 33]. This structure is comparable with the experimental self-organised structure that arises over a similar distance scale. In addition, the simulations also predict an 'overshoot' in the microtubule assembly kinetics. At the calculated 'overshoot', the simulations predict that the strongly shrinking microtubules result in strong fluctuations of concentration and density (3%). For self-organisation to occur, the algorithm also requires the presence, at this critical moment, of a small asymmetry in the reaction-diffusion process. The asymmetry acts either by directly

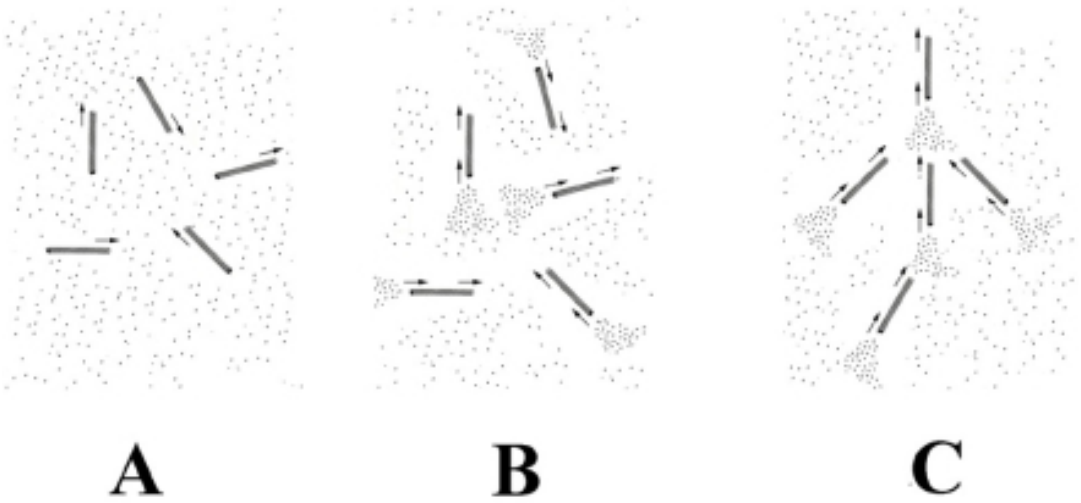


Figure 7 Proposed mechanism for the formation of the self-organised structure. Microtubules are chemically anisotropic, growing and shrinking along the direction of their long axis. This leads to the formation of chemical trails, comprised of regions of high and low local tubulin concentration from their shrinking and growing ends respectively. These concentration trails (density fluctuations) are oriented along the direction of the microtubule. Neighbouring microtubules will preferentially grow into regions where the local concentration of tubulin is highest. In A.), microtubules have just formed from the tubulin solution. They are still in a growing phase and have an isotropic arrangement. In B), microtubule disassembly has started to occur at the bifurcation time. This produces trails of high tubulin concentration from the shrinking ends of the microtubules. In C) microtubules are growing and forming preferentially into these tubulin trails. The isotropic arrangement shown in B) is unstable. Once a few microtubules start to take up a preferred orientation then neighbouring microtubules will also grow into the same orientation. Once started, the process mutually reinforces itself with time and leads to self-organisation. At the instability, any small effect that leads to a slight directional bias will trigger self–organisation. Gravity acts by way of its directional interaction with the macroscopic density fluctuations present in the solution.

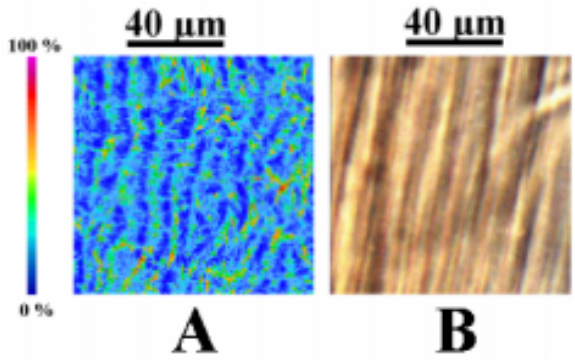


Figure 8. Numerical simulations (A) containing only reactive and diffusive terms predict macroscopic self-organisation comparable with experiment (B).

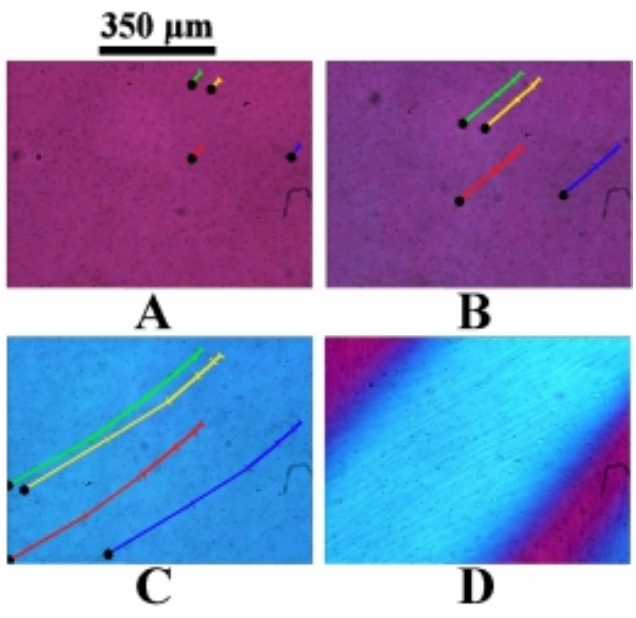


Figure 9. Transport of colloidal polystyrene particles during microtubule self-organisation. Images of the preparation at different times during self-organisation; A), 20 min; B), 40 min; C), 60 min; D), 5 hours. The numerous small dots are polystyrene beads of 1.1 µm diameter. Several have been highlighted and the coloured lines indicate their trajectories. During the first hour of self-organisation, the microtubules orient along the direction indicated by the bead trajectories.

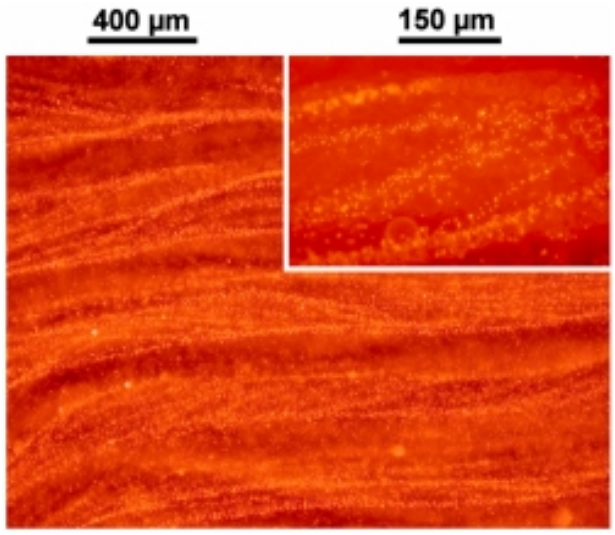


Figure 10Microtubule self-organisation also results in the organisation of colloidal particles. The photograph shows the distribution of 1.0 µm diameter fluorescent polystyrene particles in a self-organised preparation. This pattern coincides with the microtubule pattern. The particle distribution was homogenous prior to self-organisation.

orienting some microtubules, or by making tubulin diffusion faster along one direction than the others. The latter favours the growth of microtubules along this direction and triggers self-organisation by the orientational effect thus produced. Gravity, by interacting with the density fluctuations produced by strongly shrinking microtubules at the 'overshoot' (bifurcation time), gives rise to increased molecular transport along the vertical direction and so triggers self-organisation. Magnetic fields and sheering, on the other hand, act by directly orienting microtubules at the bifurcation time. Hence, gravity and magnetic fields may break the symmetry of the initially homogenous state and thus lead to the emergence of form and pattern. Gravity and magnetic fields can thus intervene in a fundamental cellular process and will indirectly affect other cellular processes that are in their turn dependent upon microtubule self-organisation. Other external factors, such as vibrations, have the same effect. Processes of this type could form a general type of mechanism by which outside environmental factors are transduced into living systems. Such processes may have played a role in the development of life on earth.

4.g .Collective particle transport and organisation

The computer simulations outlined above also suggest an explanation for another emergent property of this system; namely the directional transport and organisation of transport of colloidal sized particles [34]. One of the major biological properties of microtubules is the transport of sub-cellular particles, such as chromosomes and vesicles, from one part of a cell to another. For a variety of reasons, we suspected that the self-organising process could also result in collective particle transport. This turned out to be the case. We observed the following behaviour. When 1 μm diameter colloidal polystyrene particles were added to the initial preparation of tubulin and GTP [34] then about 15 minutes into the self-organising process, all the beads start to move in the same direction at speeds of several μm per minute. The direction of movement corresponds to the direction of microtubule orientation that develops at this time. When self-organisation is complete after about 5 hours, then there is no further particle transport. Particle transport does not occur when self-organisation is not triggered by gravity, or if microtubules are assembled under different reactive conditions such that self-organisation does not occur. Numerical simulations of the self-organised arrangement at different times during the process, show that the parallel fronts of oriented microtubules shown in figure 8A, cross the reaction space at speeds of several μm per minute [32]. These travelling fronts correspond to variations in microtubule concentration of at least 30%. As the microtubule preparation is extremely viscous, they also correspond to waves of differences in viscosity of several thousand poises. Such travelling waves, comprised of variations in concentration and viscosity, would be quite capable of transporting colloidal sized particles along with them.

Moreover, the distribution of particles, which was initially homogenous, takes on a pattern coincident with that of the microtubules. So, in addition to being to transported, the colloidal beads are also themselves organised by the self-organising process [34]. We believe that this comes about in the following way. The speed of particle transport depends on the reaction rate and is strongly dependent on the initial tubulin concentration. During self-organisation, regions of different microtubule concentration develop in the sample. As these develop, the rate of particle movement will not be the same everywhere. Particles will hence tend to accumulate into different regions of space in a manner analogous to that which cars travelling at different speed aggregate into clusters or form traffic jams.

5. Conclusions

Under appropriate conditions, *in vitro* microtubules preparations behave as a *complex* system. They self-organise and show a number of other *emergent* phenomena by way of a reaction-diffusion process, which shows analogies with the way ants and other social insects self-organise. The principal *emergent* properties that develop; self-organisation, collective particle transport, and their triggering by a weak factor, outwardly resemble the major features of microtubule behaviour in living systems. It may turn out that the mechanism and *emergent* properties outlined above are of major biological significance. Although biologists have long known that the self-organising behaviour of microtubules in living systems arises from their reaction dynamics, as yet they do not view this in terms of *emergent* properties in a *complex* system.

The question thus naturally arises as to whether the processes outlined above might also occur *in vivo*; and in particular whether or not they might arise during the cell cycle and the early stages of embryogenesis. One of the characteristic properties of microtubule self-organisation by reaction and diffusion is its dependence on various external factors such as gravity. It is known that cellular functions are modified when cells are cultured under conditions of weightlessness [35, 36]. Recent experiments on cell lines cultured under conditions of weightlessness show a disorganised microtubule network compared to control experiments under normal gravity conditions [37-40]. This latter behaviour is consistent with the *in vitro* observations reported here and raises the possibility that the processes outlined above might occur in living cells.

Rashevsky, Turing, Prigogine and co-workers, first developed their theories as a possible underlying physical-chemical explanation for biological self-organisation during embryogenesis. They predicted a way by which macroscopic chemical patterns could spontaneously develop from an initially unstructured egg. Although there is evidence that microtubule self-organisation by reaction and diffusion occurs during *drosophila* embryogenesis [23, 41, 42], it is too early to affirm whether or not this process plays a role in determining the body plan of the resulting organism. What we can say is that non-linear reaction dynamics can in principle account for biological self-organisation and pattern formation, and that an important cellular component, microtubules, behaves this way in a test-tube.

The overall phenomenological behaviour of the microtubule preparations shows a qualitative resemblance to some aspects of living organisms in the following ways. Firstly, macroscopic ordering appears spontaneously from an initially homogenous starting point. Secondly, the final state depends upon small differences in conditions at a critical moment at an early stage in the process. This is reminiscent of what occurs during biological development, when after a certain stage, cells of identical genetic content take different developmental pathways to form different cell types. Just after bifurcating, a non-linear system could be described in biological vocabulary as being ‘determined but not yet differentiated’.

The mechanism of self-organisation outlined above shows significant differences from the type of reaction-diffusion scheme originally proposed by Turing. In the Turing system, the molecules communicate with one another by diffusion (fast diffusion of the inhibitor and slow diffusion of the activator). In the microtubule system, on the other hand, as for ants, communication occurs essentially by way of the chemical trails that the microtubules produce by their own reactivity. It is a reaction-diffusion system, since without tubulin diffusion at the appropriate rate, self-organisation would not occur. Another difference with the Turing scheme is the reactive anisotropy and heterogeneity of the microtubule system. In a normal reaction-diffusion scheme there is no inherent anisotropy in the reactive process. This is not the case for an individual microtubule. Here, reactive growing and shrinking can lead to chemical trails along only one specific direction. The system has an in-built propensity for symmetry breaking under the effect of a weak external factor. In addition, in a microtubule preparation, chemical reactions can only occur at the ends of individual microtubules, and these ends are often several microns apart. The solution, once microtubules have assembled, is hence chemically heterogeneous and this factor likewise favours self-organisation. It may be that the specific type of mechanism encountered here, based on reactive growth and shortening of tubes or rods, is particularly suited to self-organisation. At present, it is not yet clear whether these processes are widespread in biology or if they are limited to microtubules.

The results outlined above demonstrate how a very simple biological system comprised initially of just a protein and GTP, and without DNA, can show a *complex* behaviour and develop *emergent* phenomena that outwardly resemble certain biological functions. These phenomena, which may be of considerable biological importance, are not the sum of properties of individual molecules, but come about spontaneously as a consequence of non-linear reaction dynamics in a population of strongly interacting elements.

References

1. Kolmogorov A, Petrovsky L, Piskunov N. An investigation of the diffusion equation combined with an increase in mass and its application to a biological problem. Bull. Uni. Moscow. Ser. Int. A1 1937; 6:1-26.
2. Rashevsky N. An approch to the mathematical biophysics of biological self-regulation and of cell polarity. Bull. Math. Biophys. 1940; 2:15-25.
3. Turing AM. The chemical basis of morphogenesis. Phil. Trans. Roy. Soc. 1952; 237:37-72.
4. Prigogine I, Nicolis G. Biological order, structure and instabilities. Q Rev Biophys 1971; 4:107-48.
5. Nicolis G, Prigogine I. Self-organization in nonequilibrium systems : from dissipative structures to order through fluctuations. New York: Wiley, 1977:xii, 491.
6. Bray W. A periodic reaction in homogenous solution and its relation to catalysis. J. Am. Chem. Soc 1921; 43:1262-1267.
7. Belousov BP. A periodic reaction and its mechanism, Sb. Ref. Radiats. Med. Medgiz, Moscow, 1958. Vol. 1.
8. Castets VV, Dulos E, Boissonade J, De Kepper P. Experimental evidence of a sustained standing Turing-type nonequilibrium chemical pattern. Physical Review Letters 1990; 64:2953-2956.
9. Ouyang Q, Swinney H. Transition from a uniform state to hexagonal and striped Turing patterns. Nature 1991; 352:610-612.
10. Nicolis G. Dissipative structures and biological order. Adv Biol Med Phys 1977; 16:99-113.
11. Kondepudi DK, Prigogine I. Sensitivity of non-equilibriim systems. Physica 1981; 107A:1-24.
12. Murray JD. Mathematical biology. Biomathematics ; v. 19. Berlin ; New York: Springer-Verlag, 1989:xiv, 767.
13. Meinhardt H. Models of biological pattern formation. London: Academic Press, 1982:xi, 230.
14. Harrison LG. Kinetic theory of living pattern. Developmental and cell biology series ; 28. Cambridge ; New York: Cambridge University Press, 1993:xix, 354.
15. Lechleiter J, Girard S, Peralta E, Clapham D. Spiral calcium wave propagation and annihilation in Xenopus laevis oocytes. Science 1991; 252:123-6.
16. Dupont G, Goldbeter A. Oscillations and waves of cytosolic calcium: insights from theoretical models. Bioessays 1992; 14:485-493.
17. Camazine S. Self-organization in biological systems. Princeton studies in complexity. Princeton, N.J.: Princeton University Press, 2001:viii, 538 , 8 of plates.
18. Hèolldobler B, Wilson EO. The ants. Cambridge, Mass.: Belknap Press of Harvard University Press, 1990:xii, 732 , [24] of plates.
19. Alberts B. Molecular biology of the cell. New York: Garland Science, 2002:cm.
20. Dustin P. Microtubules. Berlin ; New York: Springer-Verlag, 1984:xviii, 482.
21. Pirollet F, Job D, Margolis RL, Garel J. EMBO J 1987; 6:3247.
22. Tabony J, Job D. Spatial structures in microtubular solutions requiring a sustained energy source. Nature 1990; 346:448-51.
23. Papaseit C, Vuillard L, Tabony J. Reaction-diffusion microtubule concentration patterns occur during biological morphogenesis. Biophysical Chemistry 1999; 79:33-39.
24. Tabony J. Morphological Bifurcations Involving Reaction-Diffusion Processes During Microtubule Formation. Science 1994; 264:245-248.
25. Tabony J, Papaseit C. Microtubule self-organisation as an example of a biological Turing structure. In: Malhotra S, Tuszynski J, eds. Advances in Structural Biology. Vol. 5. Stamford, Connecticut: JAI Press, 1998:43-83.
26. Glade N, Tabony J. Brief exposure to high magnetic fields determines microtubule self-organisation by reaction-diffusion processes. Biophysical Chemistry 2005; 115:29-35.
27. Papaseit C, Pochon N, Tabony J. Microtubule self-organization is gravity-dependent. Proc Natl Acad Sci U S A 2000; 97:8364-8368.
28. Tabony J, Job D. Gravitational symmetry breaking in microtubular dissipative structures. Proc Natl Acad Sci U S A 1992; 89:6948-52.
29. Tabony J. Gravity dependence of microtubule self-organisation. ASGSB Bull 2004; 17:13-25.

30. Tabony J, Glade N, Papaseit C, Demongeot J. Microtubule self-organisation and its gravity dependence. Adv Space Biol Med 2002; 8:19-58.
31. Tabony J, Glade N, Papaseit C, Demongeot J. The effect of gravity on microtubule self-organisation. Journal De Physique Iv 2001; 11:239-246.
32. Glade N, Demongeot J, Tabony J. Numerical simulations of microtubule self-organisation by reaction and diffusion. Acta Biotheoretica 2002; 50:239-268.
33. Glade N, Demongeot J, Tabony J. Comparison of reaction-diffusion simulations with experiment in self-organised microtubule solutions. Comptes Rendus Biologies 2002; 325:283-294.
34. Glade N, Demongeot J, Tabony J. Microtubule self-organisation by reaction-diffusion processes causes collective transport and organisation of cellular particles. BMC Cell Biol 2004; 5:23.
35. Lewis ML. The cytoskeleton, apoptosis, and gene expression in T lymphocytes and other mammalian cells exposed to altered gravity. Adv Space Biol Med 2002; 8:77-128.
36. Hughes-Fulford M. Physiological effects of microgravity on osteoblast morphology and cell biology. Adv Space Biol Med 2002; 8:129-57.
37. Gaboyard S, Blanchard MP, Travo C, Viso M, Sans A, Lehouelleur J. Weightlessness affects cytoskeleton of rat utricular hair cells during maturation in vitro. Neuroreport 2002; 13:2139-42.
38. Lewis ML, Reynolds JL, Cubano LA, Hatton JP, Lawless BD, Piepmeier EH. Spaceflight alters microtubules and increases apoptosis in human lymphocytes (Jurkat). Faseb J 1998; 12:1007-18.
39. Uva BM, Masini MA, Sturla M, et al. Clinorotation-induced weightlessness influences the cytoskeleton of glial cells in culture. Brain Res 2002; 934:132-9.
40. Vassy J, Portet S, Beil M, et al. The effect of weightlessness on cytoskeleton architecture and proliferation of human breast cancer cell line MCF-7. Faseb J 2001; 15:1104-6.
41. Tabony J, Glade N, Demongeot J, Papaseit C. Biological self-organization by way of microtubule reaction-diffusion processes. Langmuir 2002; 18:7196-7207.
42. Tabony J, Glade N, Papaseit C, Demongeot J. Microtubule self-organisation as an example of the development of order in living systems. J. Bio. Phys. Chem. 2004; 4:50-63.

An overview of the quest for regulatory pathway in microarray data

Nizar Touleimat, Florence d'Alche-Buc, Marie Dutreix
marie.dutreix@curie.u-psud.fr

An overview of the quest for regulatory pathway in microarray data
Nizar Touleimat(1,2), Florence dAlché-Buc(2) and Marie Dutreix(1)
1 Institut Curie, UMR 2027 CNRS, Bat 119, Centre Universitaire, 91405 0RSAY 2 Programme Epigenomique, Genopole, LAMI, 523 Place des Terrasses F-91000 EVRY

Complex regulatory networks allow living cells to respond and adapt to a large set of stimuli coming from their environment. Understanding these networks is one of the major challenges to modern molecular biology. The recent developments of biotechnological tools that allow to study global transcriptional response (i.e. variation of expression of all the genes of a cell) open a new field of investigation and knowledge building. However, the tremendious amount of data generated requires new methods of data mining, classification, normalization, comparison, crossing thinking. Here, we will try, through a concrete exemple, to present the different analysis that can be done starting from a define set of genes and looking for the regulatory pathway that control there expression. The first challenge is to define as precisely as possible the set of genes to study. A potentially regulated gene group may be constructed in two different ways : by ranking differentially expressed genes in a supervised analysis comparing to classes of conditions (i.e with or without stimuli) or by clustering genes by their expression profiles in hierarchical classification or kinetic profile similarity. Once a set of genes determined, many questions arise. Has this cluster of genes any biological signification? Are the genes co-regulated and if so, by what? Are they specific to the studied stimulus? Do they share more biological characteristics (i.e. gene locus, function)? In order to answer these questions, we used a simple method based on the systematic analysis of the set of genes and its comparison to the full set of data. Genes within a set can be compared for shared qualities like conserved sequence patterns, common promoter sequences, common functions or molecular processes. The informations are considered as relevant if the frequency of these qualities within the group studied are significantly different from their distribution in the whole set of genes. In some cases, the respective chromosome gene positions can indicate a correlation between genes within the given set. ChIP-on-chip data analysis and transcriptional factor data indicates transcriptional factors that could control the expression of this set of gene. Analysis of the average behavior of our set of genes in other experimental conditions (with one or more experimental parameter changing) using laboratory data as well as published microarray allow to test the persistence of the gene clustering in different conditions. The integration of all this data enables us to find potential common regulators for genes grouped by their transcriptomic response to a stimulus. This method must respect some limits and precautions. As the gene clustering based on the gene expression profiles and intensities is the first step of our methodology, the coherence of the gene groups and the quality of our analysis are bound to be strongly affected by the quality of the gene clustering. Another limit comes from the crossing of our gene clustering with data from other laboratories. Raw data are not always accessible and different normalization methods are applied. The number of genes data available strongly vary from an experiment to another and from a laboratory to another. We will illustrate the full process of analysis starting from a group of genes clustered according to their expression profiles after irradiation with ?-rays of Saccharomyces cerevisiae.

Towards the modelling of the regulation of early haematopoiesis

Sylvie Troncale[1], David Campard[2], Fariza Tahi[1], Abdelghani Hachami[1], Janine Guespin[3] and Jean-Pierre Vannier[2]

[1]LaMI, CNRS-UMR 8042, Université d'Évry Val-d'Essonne,
Génopole, 91 000 Evry, France
[2]Laboratoire M.E.R.C.I, EA 2122, Faculté de Médecine et
de Pharmacie, 76183 Rouen, France
[3]Laboratoire de Microbiologie du Froid, UPRES 2123,
76821 Rouen, France

Abstract

Haematopoietic stem cells (HSCs) can either remain quiescent, with a constant and low rate of self renewal, or become committed to differentiation into haematopoietic cell lineages. The switch between those two fates can be seen as the switch from one stationary state to another in a dynamic process (epigenetic switch). The control of early haematopoiesis is quite complex, involving several factors and feedback circuits. Its study needs modelling, which was not straightforward, since it requires to take into account both consumption and production of resources as well as discrete and continuous time. This was performed using Hybrid Functional Petri Nets (HFPN). The model presented simulates the results obtained *in vitro*, where the quiescent stem cells are rapidly lost, and shows the potentiality of HFPN to model complex systems such as the regulation of haematopoiesis.
In silico simulations can then be performed to study the mechanism responsible for the epigenetic switch.

Key words: haematopoiesis, regulation, cytokine, multi-scale modelling, Hybrid Petri nets, simulation.

1 Introduction

Haematopoiesis is a complex phenomenon leading to the continuous production of all types of mature blood cells. This process is ensured by a population of haematopoietic stem cells (HSCs), which are able by a process called self-renewal to maintain a constant pool *in vivo*. Nevertheless, HSCs do not self-renew *in vitro* despite the development of numerous methods of culture.

The proliferation and the differentiation of HSCs are greatly influenced by cytokines secreted by the cellular micro-environment, suggesting that paracrine loops are essentials in the haematopoietic cell production. Recently, a few evidences indicate that HSCs could control their own determination, by secreting some of their growth factors [**?**]. This constitutes the prerequisite element for formation of autocrine loops, which could be responsible for bi-stability of the HSCs.

Many molecules have been studied (unsuccessfully so far) as candidate autocrine factors, but the role of a cytokine, the interleukin-6 (IL-6), has not yet been considered. IL-6 associated with its cognate receptor (sIL-6R) could be involved in an autocrine feedback. We hypothesise that a potential feedback circuit in the IL-6 network could be responsible for the switch of HSCs between self-renewal and differentiation. In this way, IL-6 activation of gp130 can be considered as the leading step of this epigenetic switch [**?**].

The nature of the haematopoietic system gives hard some experiments. Consequently, it is advisable to first test this hypothesis *in silico*. Thus, we needed first to built an accurate model simulating the regulation of IL-6 during early haematopoiesis. In order to describe this process, several aspects must be considered, such as: cellular dynamics, since a cell can be subject to different evolutions during its life-time, molecular interactions, influences of the interleukins on the cells and *vice versa*.

To model all these aspects, we need a formalism which integrates the notion of production and consumption of resources (for modelling cellular and molecular variation). Moreover, the notions of

both discrete and continuous modelling are necessary to model respectively cellular and molecular interactions and evolutions. Among the formalisms existing in the literature, we focused on Hybrid Functional Petri Nets (HFPN) [?] [?]. These nets can merge all the required elements and they also allow to define functions. They are composed of places, transitions and arcs which can be either discrete or continuous.

2 Biological process

The regulation of HSCs involves numerous growth factors. We focused on IL-6 and the molecules directly associated, since this signalling pathway is known to play a central role in stem cell biology [?]. Receptors involved in recognition of IL-6 are IL-6R and gp130. The assembly of the complete signalling receptor is sequential and hierarchically ordered. IL-6R binds IL-6. IL-6R exists in two forms: a membrane-bound form, called mIL-6R, and a freely-soluble-form, called sIL-6R. The soluble form of the receptor, sIL-6R, is secreted in an autocrine fashion by HSCs.

The complexes IL-6/mIL-6R or IL-6/sIL-6R are recruited by two gp130 subunits (membrane receptor) and the resulting gp130/IL-6/IL-6R complexes are internalized. Gp130 stimulation triggers activation of several intra-cellular pathways, particularly JAK *(Janus kinase)*, which subsequently activates the STAT pathway *(Signal transducer and Activators of Transcription)* [?]. Once the JAK/STAT pathway is activated, it stimulates self-renewal as well as synthesis of the membrane receptors which were internalized (gp130 and mIL-6R). A part of mIL-6R is cleaved in the freely-soluble-form of the receptor. Finally, the JAK/STAT pathway activates SOCS proteins *(Suppressor of Cytokine Signaling)*. These SOCS act as an inhibitor of the JAK/STAT pathway.

Since the regulation of early haematopoiesis by IL-6 involves numerous and complex signalling pathways, it is indispensable to study this process *in silico*, before performing costly experiments.

3 Modelling by HFPN the role of IL-6R in early haematopoiesis regulation

The model presented Figure **??** simulates the role of interleukin-6 in early haematopoisis. This model was built in several steps corresponding to the multiple and interdependent regulation networks that are contitutive of the complex biological process of hematopoieis.

Cellular model

First, we built a sub-model for the cellular evolution which represents evolution of each cell lineage as a function of time. This sub-model is built in discrete time, since cells can be numbered (sub-model in the bottom part of Figure **??**). It contains three discrete places (symbolized by simple circles) representing three biological entities: quiescent stem cells *Pq*, permissive stem cells *Pp*, and cells committed to differentiation *C*. Discrete transitions (symbolized by filled rectangles) represent all the processes that allow a cell to change its state. A *Pp* cell can: leave the cell-cycle to become a *Pq* cell, and inversely (respectively, transitions *T2* and *T1*), symmetrically divide (self-renewal *T3* or differentiation *T6*), asymmetrically divide *(T4)*, and differentiate without division *(T5)*.

Molecular model

The second sub-model is the molecular model representing interactions between cytokines, i.e. association and dissociation of the molecular complexes. Since the formation and dissociation of complexes are continuous phenomena in cells, this sub-model was built in continuous time (top part of Figure **??**).

Each cytokine as well as each complex of interest is modelled by a continuous place (symbolized by double circles). We therefore have places for the molecules IL-6, sIL-6R, mIL-6R, gp130, IL-6/IL-6R and gp130/IL-6/IL-6R. The continuous transitions used in the model (symbolized by

empty rectangles) represent biological processes such as formation (*T7, T8, T9* and *T10*) and dissociation (*T11, T12* and *T13*) of all the complexes.

The final model (presented on Figure **??**) was obtained by modelling biological links between the two submodels (molecular interactions and cellular evolution), such as the autocrine secretion of sIL-6R by *Pp* cells *(T14)*, as well as the kinetics of signalling pathway *(Pv_1-Pv_6, T17-T20)*.

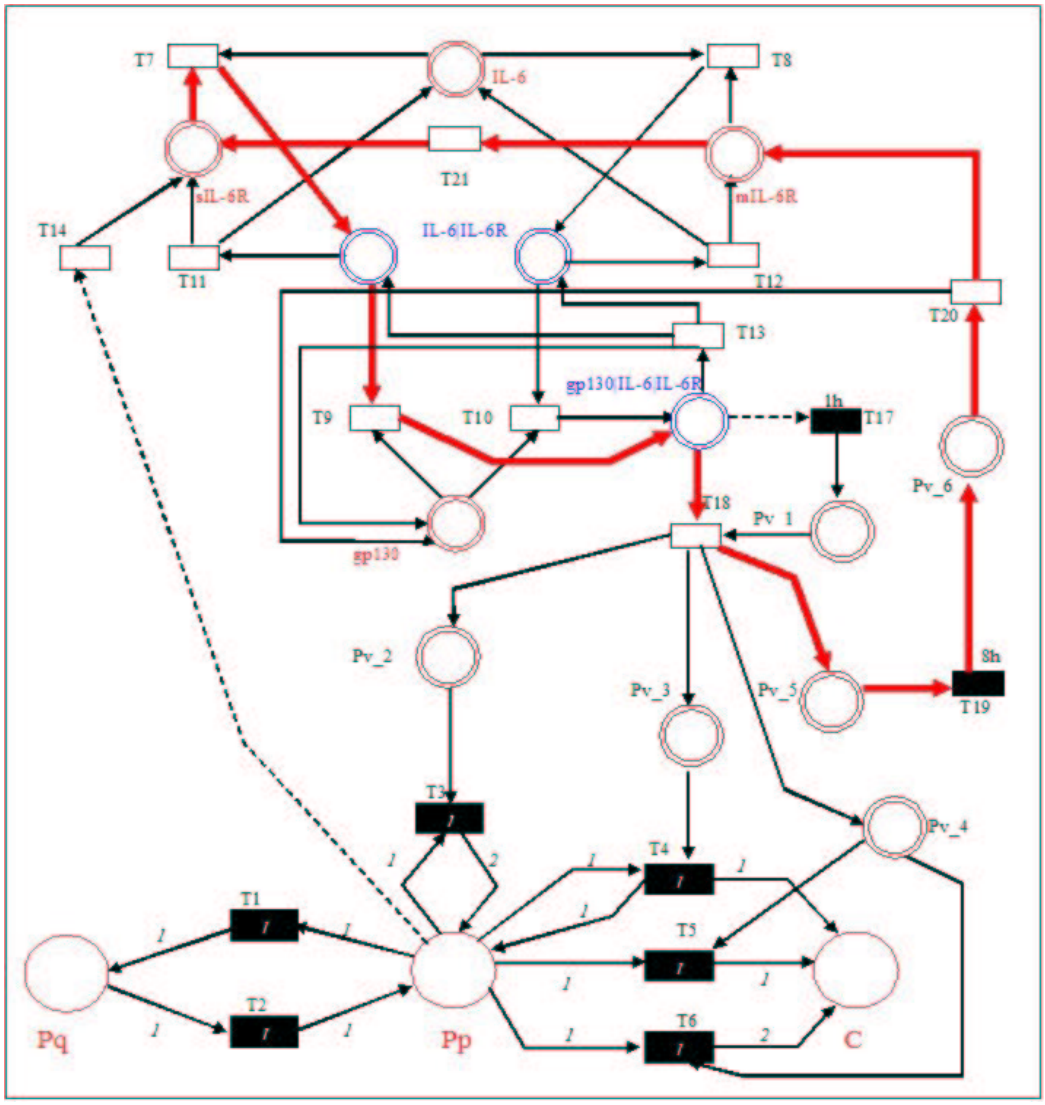


Figure 1: ***Global model.*** The cellular and molecular sub-models were joined thanks to biological links: the autocrine secretion of sIL-6R by permissive cells *(T14)* and the kinetics of signalling pathway *(Pv_1-Pv_6, T17-T20)*. The hypothetical feedback circuit is indicated in boldface arcs.

4 Results and discussion

Once the appropriate model of the IL-6-mediated regulation of the early haematopoiesis was built, a set of simulations was carried out with experimental values [?]. The simulations were done using the *Cell Illustrator* software [?] [?], which implements the HFPN formalism.

We concentrated our work on the study of the evolution of permissive cells (*Pp*) as a function of time. The simulation of the model led to the disappearance of permissive cells in about ten days (Figure ??). This experiment subsequently tested *in vitro*, and the *in vitro* results conform to the *in silico* results [?]. Consequently, our model is adequate for further testing our hypothesis of an epigenetic switch.

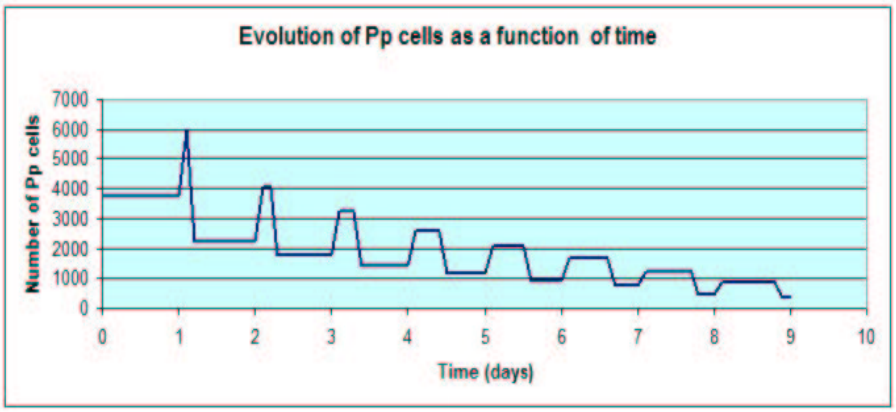


Figure 2: ***Evolution of Pp cells as a function of time.***

5 Conclusion

Haematopoiesis is a complex process involving numerous interdependent circuits in its regulation. To model such a system, we needed integrating different formalisms. The Hybrid Functional Petri Nets (HFPN) are Petri Nets which offer a maximum of functionalities. Results of simulations led to the disappearance of primitive HSC subpopulation. We succeeded to build a model, whose simulation results are in accordance with the *in vitro* results. Simulations can then be performed, in a future work, to study the mechanism responsible for the epigenetic switch.

The agreement of *in silico* and *in vitro* results provides informal validation of our model. Nevertheless, it would be interesting to validate a model in a more formal manner. At present, HFPN only enable simulations. We are therefore interested in developing a system of validation and verification for hybrid Petri nets considering a maximum of functionalities.

References

[1] Behringer D, Kresin V, Henschler R, Mertelsmann R, and Lindermann A. Cytokine and chemokine production by cd341 haematopoietic progenitor cells: detection in single cells. *Br J Haematol*, 97:9–14, 1997.

[2] Thomas R and d'Ari R. Biological feedback. *CRC Press*, 1990.

[3] Matsuno H, Tanaka Y, Aoshima H, Doi A, Matsui M, and Miyano S. Biopathways representation and simulation on hybrid functional petri nets. *In Silico Biology*, 3:389–404, 2003.

[4] Doi A, Fujita S, Matsuno H, Nagasaki M, and Miyano S. Constructing biological pathway models with hybrid functional petri nets. *In Silico Biology*, 4:271–291, 2004.

[5] Heinrich PC, Behrmann I, Haan S, Hermanns HM, Müller-Newen G, and Schaper F. Principes of interleukin (il)-6-type cytokine signalling and its regulation. *Biochem J*, 374:1–20, 2003.

[6] Campard D, Vasse M, Poyer F, Rose-John S, Lamacz M, and Vannier J-P. *Stem Cells (submitted)*, 2005.

Social Modelling

On the Dynamics of Communication and Cooperation in Artificial Societies

A.E. Eiben M.C. Schut* N. Vink

Artificial Intelligence Section
Department of Computer Science
Vrije Universiteit
De Boelelaan 1081,
1081 HV Amsterdam (NL)
{gusz,schut}@cs.vu.nl nvink@illyan.nl

Abstract

This paper compares two fundamental building blocks of complex interaction-based systems: communication and cooperation. We investigate the effectiveness of communication in an environment where the need for cooperation is scalable as well as the available resources. Several aspects of communication are considered: firstly, we compare a centralised with a decentralised communication protocol; secondly, we compare a population that always communicates with one where the entities can (evolutionary) learn to communicate. This work is part of a larger project whose main goal is to investigate the emergence of cooperation and communication in response of (scalable) environmental challenges. Our application context is an artificial society, i.e., a simulation of a societal system that was inspired by the classical SUGARSCAPE that embodies a bottom-up approach to investigate complex effects that do not necessarily have complex causes.

Keywords: communication, cooperation, artificial societies, evolutionary learning.

*Corresponding author, tel: +31.20.598.7668, fax: +31.20.598.7653.

1 Introduction

The building blocks of a complex system with many active entities are *communication* and *cooperation*. Communication is used for information exchange beetween the entities and cooperation is necessary if the entities want to achieve goals that are beyond their own reach. If we are to design complex artificial systems, we have to know the effects of deciding for available ways to communicate with respect to a given problem environment. This paper presented a first step towards such design questions.

We investigate the communication and cooperation properties of a complex system with many interacting entities among three different dimensions.

- Firstly, we compare i) a *centralised* communication protocol – where individuals multicast messages that can be received by any individual – and ii) a *decentralised* communication protocol – where information is transferred directly between agents without a third party (messageboard, or alike).
- Secondly, we vary the available *resources* and the *cooperation threshold*; the resources are needed by the agents to survive, the cooperation threshold determines how strong the pressure is on the agents to cooperate.
- Thirdly, we parameterise the communication protocols (e.g., probability that an agent talks or listens to other agents) and we empirically compare the implications of i) fixing these parameters and ii) letting the agent learn these parameters themselves.

Our agents are equiped with a hardwired mechanism for communication, and learn (by evolution) to use this mechanism. Our notion of cooperation is rooted in the environment. It is interesting to mention that our approach is complementary to some of the classics. Namely, we study the emergence of communication under fixed properties of cooperation (hard-coding its mechanics), while many studies focus on the emergence of cooperation under fixed properties of communication (see for instance [2], which assumes there is none).

Our experiments are conducted in a straightforward artificial society. This society consists of a collective of agents that lives off harvesting sugar resources in the environment. In some situations, agents may be forced to harvest sugar together with other agents. Each agent is able to communicate information about the amount of sugar at its location and may also receive such information from other agents.

Our main research objective can now be specified as follows: To study the development of the agents communication attitude (talk/listen gene distributions) and cooperative behaviour (eating together) under varying levels of cooperation (maximum amount of sugar they can eat alone).

This paper is organised as follows. In Section 2 we explain the concept of an artificial society, the communication protocols that we researched and some earlier work that we did on this topic. Section 3 describes the system that we designed for carrying out the experiments. Section 4 contains the setup of the experiments and presents the obtained results. Section 5 analyses the results. We conclude and present future work in Section 6.

2 Background

2.1 Artificial Societies

Our research can be positioned in a broader context, that of artificial societies. We let artificial societies be agent-based models of social processes [11]. This definition brings with it some notion of agents (the "people" of the artificial society), simulation (models are computationally executed to explore societal phenomena) and social structures (the macroscopic behaviour of a group of interacting individuals).

Epstein and Axtel [11] let an artificial society consist of 1) agents, 2) an environment or space, and 3) rules. An agent then has internal states and behavioral rules, which each can be fixed or flexible. Interactions and changes of internal states depend on rules of behaviour for the agents and the space. Environments can be abstractly defined (e.g., a communication network) or more resemble our own natural environment (e.g., a lattice of resource-bearing sites). The environment is a medium separate from agents, on which the agents operate and with which they interact. Rules can be defined to describe the behaviour of agents and the environment on different interaction levels, i.e., agent-environment (e.g., agents looking for and consuming food), environment-environment (e.g., growing resources), and agent-agent (e.g., combat and trade).

2.2 Communication

As mentioned, our research compares a centralised with a dentralised communication protocol.

Centralised Communication In the centralised approach, communication between agents is supported by a centralised component; this is a component that is accessible by all agents, e.g., communication through a messageboard. A messageboard enables the agents to communicate by facilitating the storage and retrieval of communicated information. All agents wanting to communicate can access the board to post a message to it, or read messages posted on it.

Decentralised Communication In the decentralised approach, there is no central support for communication. Agents that wish to communicate to other agents need to manually find agents to communicate to and exchange the information with them. Many decentralised communication protocols (e.g., gossip, epidemic-based) have been proposed and researched recently. We have implemented the newscast model [12] in the experiments described below that we explain in more detail.

Newscast Communication

The newscast computing model is a fully distributed information propagation protocol for large-scale peer-to-peer computing [12]. The main idea of newscast is that each agent maintains a cache of information items holding the information for and from the agent; the cache also contains the names of all agents that are "friends" with the agent. The cache of names, i.e., IDs and addresses, is used each time a communication is initiated by the agent. Each agent can listen and receive the messages from other agents that have it in their cache. At fixed time intervals, the agent updates the information in its cache and the list of names.

The used metaphore for illustrating the newscast model is the concept of a *news agency*. This agency regularly asks all agents for news. Additionally, the agency provides each agent with news about the other agents in the society.

Each agent has a *correspondent module* that maintains a cache of $c > 0$ newsitems, where c is fixed. A news item contains a timestamp, the agent ID and the message itself (location + sugar amount). Agents regularly exchange their caches by following this procedure (where the local agent is the agent who initiates an exchange with a peer agent):

1. Request a fresh news item from the local agent and merge the item into the cache.
2. Randomly select a peer correspondent by considering its ID as found in the cache.
3. Send and receive each other's caches. Merge received items into the local cache.
4. Since the cache now contains $2c + 1$ cache items, the oldest ones are thrown away to keep the c freshest ones (breaking ties randomly).

2.3 Scalable Environments

As mentioned, the work presented here is part of a larger research project investigating the (communication) responses of an artificial society to scalable environmental challenges (here: cooperation). As such, this work extends earlier work [9] where we also compared the centralised messageboard protocol with the decentralised newscast protocol, but where we fixed the cooperation threshold and available resources on single values. Other work [5] investigated the message removal methods for the messageboard protocol. We investigated the following settings: 1) removal after fixed number of iterations, 2) removal after the message has been listened to, and 3) removal after cooperation took place resulting from the message. No significant differences between methods 1) and 2) were found, whereas method 3) consistently resulted in early extinction of the society. In the work reported here, we used removal method 2).

In related work [4] we investigated the agent's learning capabilities to develop *physical* and *mental* properties. We researched both lifetime and evolutionary learning. The results indicate that the evolutionary approach is able to sustain larger and more stable agent populations as well as maintain a higher degree of individual success compared to the lifetime learning approach. Furthermore, quite unexpectedly, the method used for mental development has a strong effect on the development of the physical features within the very same environment: the individuals bodies evolve to completely different segments of the physical feature space under the two regimes.

In [8] we extended 1) the environment to include a heterogeneous set of resources, and 2) the agent's reproduction mechanism as to where and when reproduction took place. For the first extension, the environment presented here knows only one type of resource (defined in terms of the benefit to the agent utilising this resource), whereas in [8] we researched three different types of resources. For the second extension, agents could reproduce 1) only with their neighbours or with anyone in the environment and 2) during their lifetime or only at the end of their lifetime. The results show that indicate that utilizing reproduction at the end of an agents lifetime and local reproduction (only with neighbours) afforded the agent collective a significantly higher level of performance in its cooperative task.

Finally, in [9] we also present a preliminary theoretical model on the relationship between the decay of value of information that agents act upon and the rate at which agents exchange information with each other. The environments that we investigate turn out to have a very rapid decay of information value: a communicated message may long have lost its value once its reaches its listeners. This puts the results found in these environments in the following perspective: the used communication protocol must be able to remove outdated information fast enough. We later indeed find that the newscast protocol (due to its decentralised nature) is not able to remove outdated information fast enough.

Because our decentralised communication protocol (messageboard) removes information once it has been listened to, this gives it an advantage in these environments already from the outset.

3 System Description

The system in which we conduct our experiments consists of a simulated environment that represents an artificial society (called VUSCAPE) and a set of agents that populates this society.

3.1 JAWAS

The JAWAS[1] simulation platform is comparable with existing social simulation software such as Ascape [13], Repast [6] and Swarm [7]. All settings in JAWAS can be specified either in configuration files or via command line arguments. This enables the user to automate experiments, which substantially speeds up the time needed for, for example, investigating effects of varying experimental parameters (often requiring a large number of runs). Data are automatically saved at specified locations, enabling detailed experimental logging. Thorough statistically based experimental research on complex systems that depend on numerous parameters requires a large number of runs. Facilitating this is one of the main design objectives of JAWAS. It is very easy to add, replace or delete code when changes or extensions to the model need to be implemented. No direct connection is necessary between the program and the graphical user interface. Initial exploration can use the graphical user interface, and for automated experimentation, the system can be run solely with configuration files or from the command line.

3.2 Simulated Environment

The simulated environment is an artificial society called VUSCAPE [5], which is based on SUGARSCAPE [11]. This artificial society concerns a two dimensional grid, wrapped around the edges, where each position corresponds with an area which can contain multiple agents and some amount of sugar. Agents move through the world by vertically or horizontally jumping to another location (cf. moving in SUGARSCAPE). The agents live off the sugar, determining their level of fitness; if an agent's fitness reaches zero, it dies. The major differences between VUScape and SUGARSCAPE concern the implementations of cooperation, communication, explorative behaviour, increased grid-point inhabitance, randomised sugar distribution, and randomised age initialisation. We investigated the effects of these differences experimentally in [5].

Cooperation

Each agent can only harvest a maximum amount of sugar on its own. This amount is set by the maxSugarHarvest (MSH) parameter. If an agent is at a location at which the amount of sugar is over this threshold, it needs other agents to harvest the sugar. If there are more agents at such a location, these agents harvest the sugar together and the sugar is evenly distributed over these agents. In the experiments described below, the cooperation threshold is the same for all agents.

In addition to the MSH parameter, we scale the necessity to cooperate by varying the number of available resources in the environment, called the maximum sugar size (MSS) in VUSCAPE. Based on the settings of MSH (implements the earlier mentioned *cooperation threshold*) and MSS (implements

[1] JAWAS: Java Artificial Worlds and Agent Societies, can be downloaded from http://www.cs.vu.nl/ci/eci/.

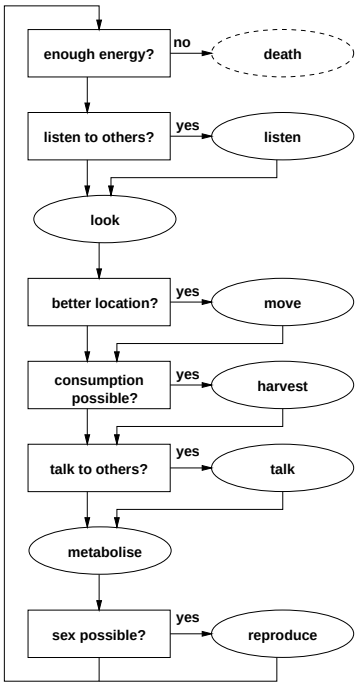


Figure 1: The agent control loop.

the earlier mentioned *available resources*), we can create easier and more difficult environments for the agents to survive.

3.3 Agents

Our agents were based on the classical SUGARSCAPE agent design: prominent features include metabolism, gender, child bearing, death, vision, allow sex and replacement. An agent was able to detect agents and resources for a number of grid-cells determined by its *vision* parameter. It was able to move for a number of grid-cells determined by its *move* parameter.

The control loop of the agent is shown in Figure 1:

- An agent gathers information about the distribution of sugar in the world. This is done by means of listening (to other agents) and looking (at the directly surrounding locations and the current location). Upon completion of this stage, the agent has at its disposal an array of locations and amounts of sugar on these locations.
- Based on this array, the agent makes a decision about its next action. In particular, it chooses a location and moves to this location. This location chosen is always the one containing the largest amount of sugar. If there are more than one such location, the agent chooses a random one.
- Having arrived at the sugar, this sugar is harvested in case the amount is under the cooperation threshold. If the amount is above the cooperation threshold, the agent cooperates immediately

if there are more agents at the location. Otherwise, it multicasts (with some probability) to other agents that it needs help.

- If possible, the agent reproduces and generates offspring. For this, it is (at least) necessary that there is another agent of the opposite sex at the location.

Communication

Agents are endowed with talk and listen capabilities. The talk feature determines whether the agent performs a communicative action itself, namely informing other agents of: 1) the amount of sugar that is on its location, and 2) the coordinates of its location. The listen feature is used in the observation and decision making processes of the agent. By listening, the agent receives information from other agents about amounts of sugar at the locations of those agents.

The listen and talk genes express probabilities and are formally real valued numbers between 0 and 1. They are called *listen preference* and *talk preference*, respectively. These probabilities are used in the control loop of the agent - steps "listen to others?" and "talk to others?" in Figure 1. In these steps, a random real is drawn and if the resultant value is under the (talk or listen) preference, the agent listens or talks to others.

Evolution

Agents underwent an evolutionary process of selection and variation. Agents with a high fitness were selected for and variation of agents was accomplished by *crossover* of agent genotypes. The agent genes involved are the talk and listen preferences and the initial amount of sugar. Crossover happens by reproduction of two agents; this is not subject to individual decisions, nor is there any mate selection. If 1) two agents are on the same location or next to each other, 2) the genders differ, 3) their sugar levels are above the reproduction threshold, 4) they are both in the fertile age range, 5) there is a vacant cell in the vicinity for placing the offspring, agents will always mate and generate offspring.

Reproduction takes place by crossover applied to two parents yielding the child, followed by a mutation operator on the child. The value of a gene in a child is the inherited value (from the wealthiest parent) plus a random number drawn from a Gaussian distribution with zero mean and fixed standard deviation σ. The child receives from each of the parents half of their sugar. The child inherits each of the values for vision, age of death, metabolism, and child bearing independently from one of the parents without change. After mating, each agent has a so-called recovery period, which is the number of cycles after mating that an agent cannot mate.

To illustrate the process of reproduction, consider the following example (without mutation). Two agents are next to each other one with 24 sugar units, a listen preference of 0.7 and a talk preference of 0.55; the other has 16 sugar units, a listen preference of 0.6 and a talk preference of 0.5. A child of these two agents gets its listen and talk preferences from the first agent (0.7 and 0.55 respectively). Its initial sugar amount is the sum of half of the sugar amounts of each of the parents, thus 12 from the first parent and 8 from the second parent - its initial sugar amount is thus 20.

4 Experiments

4.1 Setup

The experimental setup is shown in Table 1. We varied four parameters:

maxSugarHarvest (cooperation)	**maxSugarSize (resources)**	**Communication protocol**	**Communication parameters**
[0 . . . 10]	[1 . . . 10]	messageboard newscast	evolutionary fixed

Table 1: An overview of the experimental setup.

- The cooperation threshold - in VUSCAPE this is implemented by the maxSugarHarvest (MSH) (maximum amount of sugar that agent can harvest on its own). We varied the MSH from 0 to 10 in single increments.
- The number of resources - in VUSCAPE this is implemented by the maxSugarSize (MSS) (amount of sugar that is ditributed in the world). We varied the MSS from 1 to 10 in single increments.
- The communication protocol - this was either the centralised *messageboard* protocol or the decentralised *newscast* protocol.
- The communication parameters - these could be either *fixed* (agents always listen and talk) or *evolutionary* (agents learn to use their communication capabilities by evolution).

This makes a total of $11 \times 10 \times 2 \times 2 = 440$ experimental runs. Each run was done 50 times.

An overview of all experimental parameter values is given in Table 2. Additionally, talk and listen features are inherited from the parent with the most sugar. The mutation sigma is 0.1.

4.2 Results

The obtained results have been included in Figures 2, 3 and 4.

Figure 2 shows the results concerning the *performance* of the research populations[2]. Each datapoint corresponds to an environment with a certain need for cooperation. This need is expressed by the number of available resources (maximum sugarsize MSS) and the cooperation threshold (maximum sugar harvest MSH). Each line should be interpreted as follows: in environments on and under the line, all runs resulted in extinction of the agent population. In environments above the line, the agent population size stabilised at some value. (Note that Figure 2 also contains the results of a benchmark experiment that we conducted without communication.)

Figure 3 and 4 show the results of the experimental runs where the communication parameters were evolutionary. Each graph shows, for some given MSH and MSS, the final talk and listen preferences of these runs. In other words, every datapoint represents the average talk preference and average listen preference of all agents of the population at iteration 2000.

Because we used $10 \times 11 = 110$ different combinations of MSH and MSS, it is infeasible to give separate graphs for each environment; we therefore defined *low* (MSH $\in [0-2]$), *medium* (MSH $\in [3-5]$) and *high* (MSH $\in [6-10]$) cooperation thresholds. The ranges have not been chosen

[2]This graph is very dense presentation of many experimental runs. Each of the 5 graph lines (no communication, messageboard (fixed), newscast (fixed), messageboard (evolutionary) and newscast (evolutionary)) was extracted from the set of experimental runs covering the complete MSH $\times$ MSS range. These graphs can be found in [15].

Experiment		Agent	
numberOfRuns	10	maxSugarHarvest (msh)	**exp**
Scape		singleStep	false
height	50	initAgeZero	false
width	50	minVision	1
runLength	2000	maxVision	1
reseedSugar	true	minSugarMetabolism	1
initialPopulation	400	maxSugarMetabolism	1
sugarSeed.uniqueCell	false	minDeathAge	60
sugarDistributionUnif	3	maxDeathAge	100
sugarGrowBackRate	1.0	sexRecoveryPeriod	0
numberOfSeeds	1000	minReproductionSugar	0
sugarDistributionType	uniform	allowSex	true
maxSugarSize (mss)	**exp**	minInitialSugar	50
Cell		maxInitialSugar	100
allowMultipleAgents	true	preferNearestCell	false

Table 2: An overview of the experimental parameters. Parameters indicated with **exp** are varied in the experiments.

arbitrarily, but are related to the percentage of seeds that can be eaten by the agents when they are alone on a cell. They are chosen such that when the cooperation threshold is low, and when the most sugar is available (MSS = 10), the agents can harvest less than 25% of the seeds without the help of other agents. When the cooperation threshold is medium, the agents can harvest only about 50% of the seeds without the help of other agents. When the cooperation threshold is high, the agents can harvest nearly all seeds without help. Note that the above mentioned percentages are not stable during the simulation; they only apply to the initial situation where all seeds are at their maximum size. (For space restrictions, we did not include the graphs for `mss` = 2, 4, 7, and 9; these can be found in [15].)

5 Analysis

We analyse the results according to how we presented them in the graphs. Firstly, we analyse Figure 2 that says something about the *performance* of the populations. Secondly, we analyse Figures 3 and 4 to find out about the *evolution of communication* (what values did the talk and listen preferences evolve to?) in our experiments.

5.1 Performance

Figure 2 shows the performances of the populations[3] under the investigated communication protocols. Each line in the graph represents a *extinction border* and should be read as such: on all datapoints

[3]Our notion of *performance* only takes into account the *population size*. In [15] we present a more elaborate *welfare function*, but extensive testing (reported in [15]) demonstrated no significant difference between the population size and the population welfare for the environments that we investigate here. Hence, this paper only reports on the population sizes.

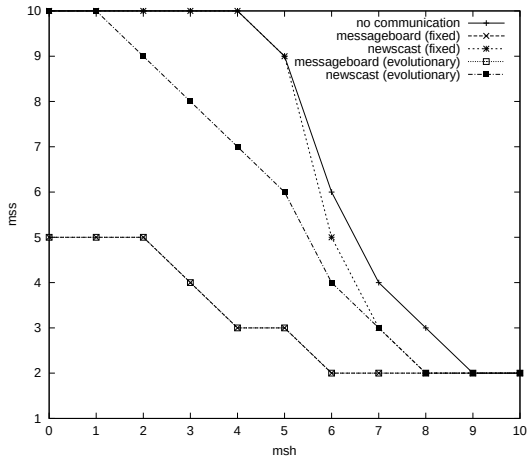


Figure 2: Summary of the performance results of our comparison of protocols. Each datapoint corresponds to an environment with a certain need for cooperation. This need is expressed by the number of available resources (maximum sugarsize mss) and the cooperation threshold (maximum sugar harvest). Each line should be interpreted as follows: in environments on and under the line, all runs resulted in extinction of the agent population.

under and on the graphline the population died out during the runs; on all environments above the line, the population stabilised at some value.

From the data from which we generated Figure 2 (covering the complete MSH $\times$ MSS spectrum), we drew the following conclusions. (Note that these conclusions can only in part be observed in Figure 2; we included them here because of their relevance to the findings we make later.) We found that overall the messageboard protocol with fixed communication parameters performs much better than the newscast protocol with fixed communication parameters. Also for the evolutionary variant, we saw that the messageboard protocol outperforms the newscast protocol. However, the difference is smaller here, because the performance of the newscast protocol with evolutionary parameters performs better that with fixed parameters. In general, we saw that the messageboard performs better in the fixed variant than in the evolutionary variant, whereas the newscast protocol performs better in the evolutionary variant than in the fixed variant.

In Figure 2 the graphs show that the messageboard (with evolution) performs best, followed by the newscast protocol with evolution, then by pure newscast and finally no communication, which performs worst. In more detail:

- *Communication gives an advantage to the agents that use it.*
  In Figure 2 we see that communication improves the life expectancy of agents. All protocol variants we used extend the set of environments in which agent populations are viable. There are no environments in which agents using communication go extinct while agents which do not use communication survive; the set of environments in which agents not using communication survive is a true subset of the set of environments in which agents using communication survive.

- *In environments in which agents can survive without using communication, the added value of communication is relatively small.*
  When we only look at environments in which non-communicating societies survive, we see that

adding communication leads to a relatively small increase in the 'performance' of the agent population. We contemplate that there exists some minimum population size for the population to survive; this may logically follow from the fact that the probability of reproduction depends on the population size. If the population size is too small, the probability of finding a partner for reproduction may also be too small. Based on our results, we expect the minimum population size necessary for survival to be somewhere close to 500.

5.2 Evolution of Communication

Messageboard (Figure 3)

- *Talk preference evolves to higher values when the need for cooperation increases.*
  In figure 3(a) we see that when the need for cooperation is high (`mss` ≤ 6), the runs in which the agent population survive result in higher average talk preference than when the need for cooperation is lower[4] (e.g. `mss` = 10). We observe the same effect in figure 3(b). The effect is not visible in figure 3(c); this is possibly because the selective pressure on higher talk preference is insignificant when the cooperation threshold is too high.

- *Listen preference evolves to higher values when the need for cooperation increases.*
  What we observed for talk preference also holds for listen preference. When looking at figure 3(c) (high cooperation threshold) we see that as the value for `mss` increases, the surviving runs result in higher average listen preferences. The effect is also visible when comparing the corresponding graphs in figures 3(a), 3(b) and 3(c) to each other; when the cooperation threshold is low, the surviving runs result in higher average listen preferences than when the cooperation threshold is higher.

- *Listen preference evolves to higher values than talk preference.*
  In almost all graphs in figure 3 we see that listen preference evolves to higher values than talk preference. In other words, in almost all graphs we see that most points are below the diagonal from $(0, 0)$ to $(1, 1)$.

Newscast (Figure 4)

- *Listen preference evolves to* lower *values when the need for cooperation increases.*
  The results for the newscast protocol show that the average listen preference evolves to *lower* values when the need for cooperation increases. This effect is most clearly observable in figure 4(b). Supposedly there exists some optimal level of use of the newscast protocol which the agents should not exceed. These results indicate that this level of use involves an average listen preference of less than 0.5.

- *Talk preference does not evolve consistently.*
  We observe no effects for the evolution of average talk preference in the results of the newscast protocol. The resulting average talk preferences of the surviving agent populations seem to span the entire possible range for average talk preference. The evolution of talk preference observed may be due to genetic drift. Such genetic drift may explain some observation made in Figure 4(c). In the bottom right corners of the graphs in this figure, we see many experiments in which

[4] We assume that the need for cooperation is lower when more sugar is available. Although more sugar also means larger seeds, and thus less seeds that can be harvested by single agents, we assume that the need for cooperation in order to survive is lower, because when agents harvest, they harvest more sugar.

nearly all agents talk whenever possible, while hardly any agents listen. We can also observe such results for the messageboard protocol, but they are much more prevalent in the results for the newscast protocol.

- *Listen preference evolves to higher values than talk preference.*
  As for the messageboard protocol, in the results of the newscast protocol we see that listen preference generally evolves to higher values than talk preference. Still, the effect is weaker than in the results of the messageboard protocol because of the earlier mentioned findings.

Summary

In both experiments, we observed that *the listen preference evolves consistently to higher values than talk preference*. This indicates that, for some reason, listening is more 'important' to the agents than talking. Although we think that this is because of the induced distribution of talk and listen events, we do not speculate further about what may be this reason - this is subject to future study.

6 Conclusions

The building blocks of a complex system with active entities are communication and cooperation. Communication is used for information exchange beetween the entities and cooperation is necessary if the entities want to achieve goals that are beyond their own reach. If we are to design complex artificial systems, we have to know the effects of deciding for available ways to communicate with respect to a given problem environment. This paper presented a first step towards such design questions.

In a straightforward artificial society, we compared a centralised communication protocol (messageboard) with a decentralised protocol (newscast). The environments we investigated were scalable as we varied the need for cooperation (by including actions to be carried out by multiple agents) and the available resources. In addition we considered hardwired agents (that always communicate) with learning agents (that evolutionary learn to communicate).

The results show that the performance of the messageboard protocol is much better than that of the newscast protocol in the environments that we examined. We observe this with both types of agents. Preliminary further investigation of this result indicates that the ratio between the speed at which information is distributed and the speed at which information looses its value is essential for the success of a given communication protocol [9]. Surprising still, with the newscast protocol, the learning agents outperformed the hardwired agents. Also, with the learning agents, we consistently observe that agents develop a higher preference for listening (receiving information from others) that talking (communicating information to others).

In the long run, we aim to design Emergent Collective Intelligence (ECI). The end goal of ECI research is to combine and exceed achievements in multi-agent systems [1], swarm intelligence [3], and evolutionary computation [10] research via developing synthetic methodologies such that groups of computationally complex agents produce desired emergent collective behaviors resulting from the bottom-up development of certain individual properties and social interactions. Forthcoming results will be published on different scientific forums; for locating them conveniently one can visit: http://www.cs.vu.nl/ci/eci.

References

[1] Avouris, N. and Gasser, L., *Distributed Artificial Intelligence: Theory and Praxis*. Kluwer Academic Publishers. Dordrecht, 1992.

[2] Axelrod, R *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton University Press, New Jersey.

[3] Bonabeau, E., Dorigo, M., Theraulaz, G., *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, Oxford, 1998.

[4] Buresch, T., *Evolutionary versus lifetime learning of agents controllers on a gridworld*, Technical report, Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands, 2004. Available online at http://www.cs.vu.nl/ci/eci/.

[5] Buzing, P.C., Eiben, A.E., and Schut, M.C. *Evolving Agent Societies with VUScape*. In: Banzhaf, W., Christaller, T., Dittrich, P., Kim, J.T., and Ziegler, J., editors. *Advances in Artificial Life, Proceedings of the 7th European Conference on Artificial Life (ECAL 2003)*, Lecture Notes in Artificial Intelligence, volume 2801, Springer Verlag, 2003, pp. 434-441.

[6] Collier, N. *Repast: An extensible framework for agent simulation*. Technical report, Social Science Research Computing, University of Chicago, Chicago, 2000.

[7] Daniels, M. *Integrating simulation technologies with swarm*. In: Proceedings of the Workshop on Agent Simulation: Applications, Models, and Tools, University of Chicago, Chicago, 1999.

[8] Eiben, A.E., Nitschke, G., and Schut, M.C., *Comparison of Reproduction Schemes in an Artificial Society for Cooperative Gathering*. In: Proceedings of the AISB Socially Inspired Computing - Engineering with Social Metaphors Symposium. Technical report. University of Hertforshire. 2004.

[9] Eiben, A.E., Schut, M.C. and Toma, T., *Comparing Multicast and Newscast Communication in Evolving Agent Societies*. In: Beyer, H.-G., et al, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2005)*, ACM Press, 2005, pp. 75-82.

[10] Eiben, A. E. and Smith, J. E., *Introduction to Evolutionary Computing*, Springer-Verlag: Berlin, 2003.

[11] Epstein, J., and Axtell, R. *Growing Artificial Societies*, Brookings Institute Press, 1996.

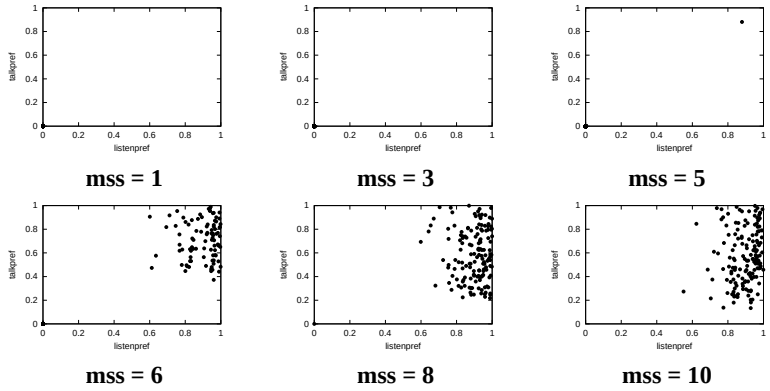
[12] Jelasity, M., and Van Steen, M. *Large-Scale Newscast Computing on the Internet*. Technical Report IR-503, Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands, 2002.

[13] Parker, M. *Ascape: Abstracting complexity*, Technical report, Center on Social and Economic Dynamics, The Brookings Institution, Washington D.C., 2000.

[14] Toma, T. *Communication in Artificial Societies - Effects of Different Communication Protocols in an Artificial Environment*, Technical report, Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands, 2003. Available online at http://www.cs.vu.nl/ci/eci/.

[15] Vink, N. *Exploring Communication Dynamics in VUSCAPE*, Technical report, Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands, 2004. Available online at http://www.cs.vu.nl/ci/eci/.

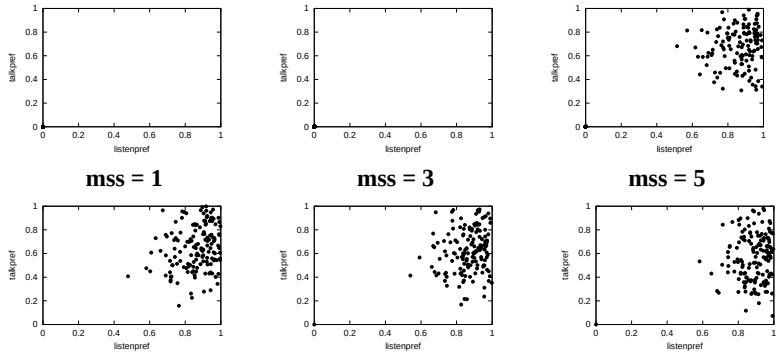
(a) **`maxSugarHarvest`** $\in [0-2]$



mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

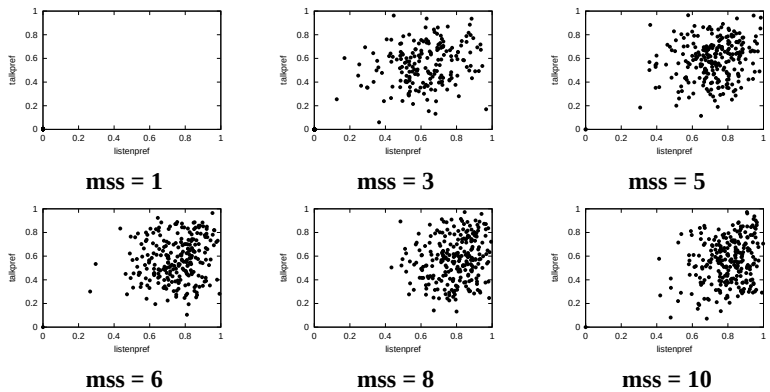
(b) **`maxSugarHarvest`** $\in [3-5]$



mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

(c) **`maxSugarHarvest`** $\in [6-10]$

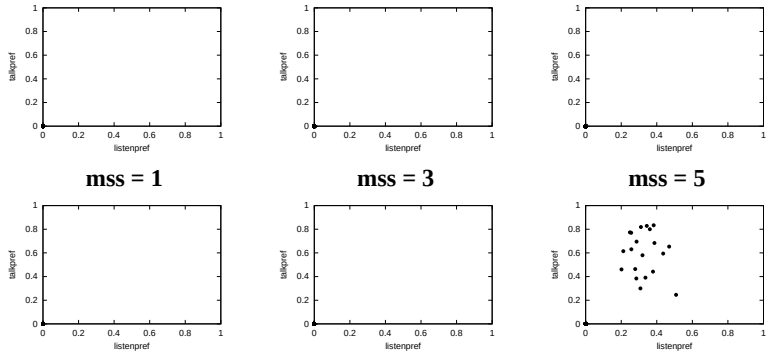


mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

Figure 3: Messageboard results (**mss** = maxSugarSize)

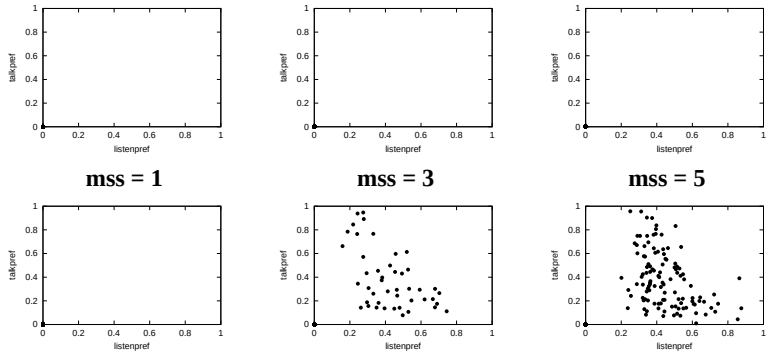
(a) **maxSugarHarvest** $\in [0-2]$



mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

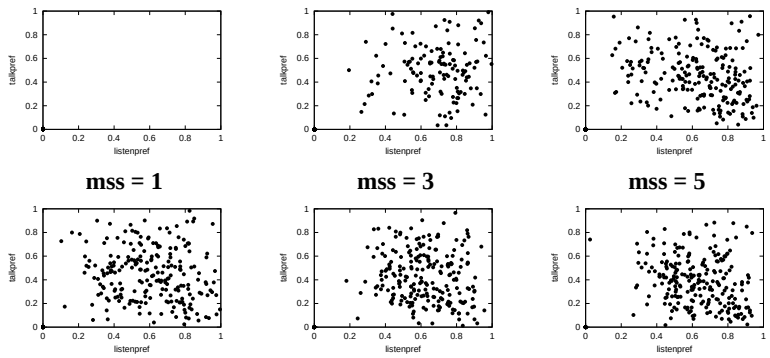
(b) **maxSugarHarvest** $\in [3-5]$



mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

(c) **maxSugarHarvest** $\in [6-10]$



mss = 1 **mss = 3** **mss = 5**

mss = 6 **mss = 8** **mss = 10**

Figure 4: Newscast results (**mss** = maxSugarSize)

Towards a Functional Formalism for Modelling Complex Industrial Systems *

D. Krob † and S. Bliudze ‡
Laboratoire d'Informatique de l'Ecole Polytechnique (LIX)
CNRS & École Polytechnique

Abstract

This paper is devoted to the presentation of the main lines of an unified functional formalism for modelling complex industrial systems, that is to say systems that typically mix a big number of different software and physical devices. Our approach is based on a discrete non standard representation of time. It captures both the hierarchical architecture and the temporal and data multi-scale structures of a complex industrial system. We show in particular in this paper how our formalism allows to recover in an totally unified way various sorts of simple systems such as Turing Machines, elementary conservative physical systems and low level software/physical interfaces (sampler, modulator).

Keywords – Complex system; Hybrid system; Non standard analysis; Physical system; Software system; System modelling; Temporized system; Time scale; Turing machine.

This paper is dedicated to the memory of Imre Lakatos

1 Introduction

In the modern world, complex industrial systems are just everywhere even if they are so familiar for us that we usually forgot their underlying technological complexity. Transportation systems (such as airplanes, cars or trains), industrial equipments (such as microelectronic or telecommunication systems) and information systems are for instance typical examples of complex industrial systems that we are using or dealing with in the everyday life.

"Complex" refers here to the fact that the engineering of these industrial systems relies on incredibly complex technical and managerial processes. Such systems are indeed characterized by the intrinsic difficulty of their design, due both to an important technological heterogeneity and to the large number of sub-systems they involve. To face this huge complexity, engineers developed a number of methodological tools, popularized in the industry under the name of *system engineering* (see [22, 23] for general systems or [30, 34] for software systems), that fundamentally

* This paper was supported by the Ecole Polytechnique and Thales' chair "Engineering of complex systems".

† Corresponding author: Laboratoire d'Informatique de l'Ecole Polytechnique (LIX) – École Polytechnique – 91128 Palaiseau Cedex – France – Tel/Fax: 01 69 33 40 73/30 14 – e-mail: dk@lix.polytechnique.fr

‡ Laboratoire d'Informatique de l'Ecole Polytechnique (LIX) – École Polytechnique – 91128 Palaiseau Cedex – France – e-mail: bliudze@lix.polytechnique.fr

rely on the fact that complex industrial systems can be always recursively decomposed in a series of coupled sub-systems, up to arriving to totally elementary systems which can be completely handled. In such a framework, system engineering provides then methods for helping both the design, the architecture, the progressive integration and the final validation and qualification steps that structure the construction of an industrial complex system.

This methodological environment is however not a fully satisfactory answer to the problems that engineers must permanently solve in practice to handle this complexity. This empirical and operational engineering approach indeed hides the fact that there are basically no theoretical tools for dealing with systems at a global level. The key problem comes here in particular from the fact that the notion of an industrial system in its whole is not very well defined and rather subjective, even if it clearly corresponds to a strong industrial reality. Due to this lack of global formalism, lots of different approaches were developed – without any unified point of view – to take into account different (important, but partial) aspects of complex industrial systems. The direct practical consequence of this situation is the tremendous diversity and heterogeneity of the modelling and simuling tools that the system engineers are using in practice. These tools indeed depend first whether one is dealing with a pure software system (where UML – see [6] or [28] – is for instance still one of the core modelling technique), an embedded (real-time) system (where synchronous languages – see [4] for a comprehensive introduction and a rather complete bibliography – are the typical good design tools ; see however also [19, 20] for different types of approaches to the same subject), a physical system (where Mathlab and Simulink – see [21] – are probably the key modelling and simuling standards) or an hybrid system (different kinds of existing models for this other important class of mixed systems can be found in [2, 13, 15, 25, 29, 37]). Independently of the type of system which is modelled, engineers also use different specification tools according to the part of the life cycle of a system they are dealing with: the initial informal global specification tools (such as Doors – see [35]) are for instance totally disconnected from the more formal tools that can support verification techniques (see [1, 5, 7, 31, 36] for details), but that are only used for specifying the most critical sub-systems, which are themselves altogether totally independent from the IVVQ [1] tools that can be found at the very end of the design and construction cycle of a complex industrial system.

The purpose of this paper is to try to reconnect all these (more or less disconnected) streams by going back to the very fundamentals, that is to say by looking for *an unified definition of an industrial system* from which all these different models could be deduced. Observe that such an approach is clearly in rupture with the usual one which is in fact more oriented to the local fixing of the connection problems existing between the different tools that are used for designing and managing an industrial system (by transforming them into interface design questions). We think however that the key problem is much deeper and comes directly from the fact that there does not really exist any mathematically consistent global point of view on industrial systems (even if some interesting approaches are to be noticed – see for instance [8, 33, 38]). This paper should therefore be seen as a – modest – attempt in this direction.

We indeed propose here a formal definition of a system which intends to capture both continuous and discrete systems which are the two major types of technical sub-systems that compose a given industrial system [2] (see Section 3). The key point on which our approach relies is a common (discrete) modelling of time based on the use of a non standard model of real numbers (see Section 2). Making this (very strong) change allows indeed to take into account in the same way both conservative physical systems and computer systems (see again Section 3 for several

[1] Integration, Verification, Validation, Qualification.

[2] We will in particular not try to integrate human systems – such as company organizations – in our modelling even if such systems are also fundamental for some type of applications (typically information systems).

examples). Moreover our systemic models are always causal (see Section 3.1): non causality appears indeed in our approach as the consequence either of abstraction (i.e. simplifying our model) or standardization (i.e. going back to the usual model of time). Observe also finally that large classes of classical systems such as synchronous, Hamiltonian or dynamical systems can also be recovered within our approach (see Section 4 for some insights on these questions).

2 Time

In order to model all industrial systems in an unified way, the first key problem – as already mentioned above – is to develop a common functional [3] framework that takes both into account continuous and discrete systems (typically physical and computer systems), that is to say systems whose time evolution behavior is represented either continuously or discretely. At this point, two directions can be chosen to construct an unified theory, depending whether one prefers to develop a continuous or a discrete point of view on time. In this paper, we decided to develop the discrete approach since we wanted to keep the usual intuitions on computer systems. The price to pay is then the change of model of real numbers on which time rely in order to capture also continuous systems in a same global framework. Note however that one could also do the converse and deal always with the usual continuous modelling of time which will then lead us to develop a distribution point of view (see [32]) on computer systems, using for instance typically Dirac combs for modelling the discrete entries of a given software system.

2.1 Non standard analysis

To develop a global (discrete) unified framework for dealing both with continuous and discrete industrial systems, we will therefore go back to the 18th century representation of real numbers (see [10, 17, 18]), that is to say to the so-called non-standard model of $\mathbb{R}$. In other words, we will consider in all the sequel that real numbers are given by the set ${}^*\mathbb{R}$ of *non-standard reals* [4] as initially formalized by Robinson [27] (for a comprehensive introduction to non standard reals see the paper of Lindström in [9]). This set (see Figure 1) is a real-closed field that contains all usual real numbers, but also the infinitesimal reals (i.e. the non zero non standard real numbers that have their absolute value strictly less than any $r \in \mathbb{R}_*^+$) – whose set will be denoted by $\mathbb{I}$ in the sequel – and their inverses which are the infinitely great reals (i.e. whose with an absolute value strictly greater than any usual real number $r \in \mathbb{R}$).

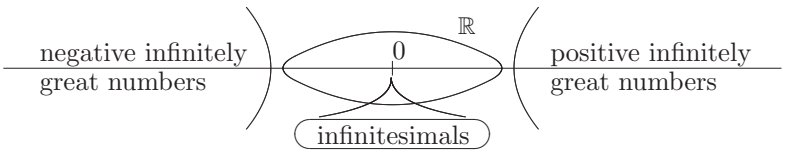


Figure 1: Graphical representation of non-standard real numbers.

Note in particular that one can prove that the field ${}^*\mathbb{R}$ is elementarily equivalent to $\mathbb{R}$, which means that the first order logical properties of $\mathbb{R}$ and ${}^*\mathbb{R}$ (expressed in the logical theory of

[3] "Functional" refers to the fact that systems will only be considered here as input/output functions. We will in particular not try to model the architecture on which relies a system (which would bring us too far ...).

[4] A star on the left of a symbol as in ${}^*\mathbb{R}$ stands always for "non-standard", whereas on the right, it means either that zero is not included, as in $\mathbb{R}^*$ or ${}^*\mathbb{R}^*$ (which denote respectively the usual and the non standard sets of non zero real numbers) or that we speak of the set of words over a given alphabet, as in A^*.

ordered fields) are exactly the same (see [3] and [14] for more model theoretical fundamentals of non standard analysis). Observe also that, among all non-standard real numbers, one can of course consider the set ${}^*\mathbb{Z}$ of *non-standard integers* that, on top of standard integers, contains infinitely great ones, having absolute value greater than any $n \in \mathbb{N}$.

Note finally that two non-standard real numbers x and y are said to be *infinitely close* (which is denoted by $x \approx y$) if and only if $x - y$ is infinitesimal. This last notion allows, for instance, to recover the usual concept of continuity. Indeed, the standard image $f : \mathbb{R} \to \mathbb{R}$ of a non standard function ${}^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ is continuous in a standard point $x_0 \in \mathbb{R}$ if and only if one has

$$\forall\, \delta \in {}^*\mathbb{R}, \quad \delta \approx 0 \;\Rightarrow\; {}^*f(x_0 + \delta) \approx {}^*f(x_0)\,.$$

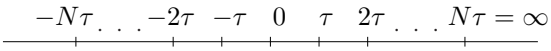
More details about non standard analysis can be found in [11] and [9].

2.2 Time scales

From now on, we will suppose that the time is modelled by ${}^*\mathbb{R}$ throughout all this paper. Let us therefore give the following first fundamental definition.

Definition 2.1 *Let $\tau \in {}^*\mathbb{R}_*^+$ be a strictly positive non-standard real number. The set $\mathbb{T}_\tau = {}^*\mathbb{Z}\,\tau$ will then be called the* time scale *of step $\tau > 0$. Any element $t \in \mathbb{T}_\tau$ will said to be a* moment *on this last time scale.*

A time scale can therefore be seen as a discrete series of clock ticks occuring at times $n \cdot \tau$, with $n \in {}^*\mathbb{Z}$ being a non-standard integer eventually infinitely great.



The following lemma shows then we can recover usual continuous time within this model by considering time scales with infinitesimal steps (i.e. with $\tau \approx 0$).

Lemma 2.2 *Let $\mathbb{T}_\tau$ be a time scale with an infinitesimal step τ and let $r \in \mathbb{R}$ be any standard real number. Then there always exists a moment $t \in \mathbb{T}_\tau$ which is infinitely close to r.*

Proof — Let $\tau \approx 0$ be an infinitesimal non standard real number and let r be a usual standard real number. The non standard version of Archimedes' axiom (see Chapter A.6 of [3] or use the transfer principle as explained in [11]) shows that there exists a non standard integer $N \in {}^*\mathbb{N}$ (here necessarily infinitely great) such that $N < r/\tau \leq N+1$. Hence we have $N\tau < r \leq (N+1)\tau$ and consequently $0 < r - N\tau \leq \tau \approx 0$. Thus putting $t = N\tau$, we have $t \approx r$. ∎

More generally, we can classify possible all time scales $\mathbb{T}_\tau$ into the following three groups according to the nature of their step τ:

- *continuous* time scales when their time step is infinitesimal, i.e. when one has $\tau \approx 0$,
- *discrete* time scales when their time step is a non infinitesimal bounded non standard real number, i.e. when one has $\tau \approx r$ for some strictly positive usual real number $r \in \mathbb{R}_*^+$,
- *infinite* time scales when their time step is an infinitely great non standard real number.

Note that the latter case is not of practical interest as there are essentially only three "standard" moments on an infinite time scale, that is to say $-\infty$, 0, and $+\infty$. Therefore we will concentrate in the sequel of this paper uniquely on the time scales of the first two types.

Let us finally give the following three definitions that we will use for dealing with systems.

Definition 2.3 *A time scale $\mathbb{T}_\tau$ is said to* refine *another time scale $\mathbb{T}_{\tau'}$ – which is denoted by $\mathbb{T}_{\tau'} \preceq \mathbb{T}_\tau$ – if and only if one of the two following equivalent properties holds:*

- $\mathbb{T}_{\tau'} \subset \mathbb{T}_\tau$
- $\exists\ N \in {}^*\mathbb{N},\ \tau' = N\tau$

Definition 2.4 *Given two time scales $\mathbb{T}_\tau$ and $\mathbb{T}_{\tau'}$, we shall call* synchronization points *all the moments that belong both to the two time scales, i.e. to $\mathbb{T}_\tau \cap \mathbb{T}_{\tau'}$.*

Observe that, if $\mathbb{T}_\tau$ refines $\mathbb{T}_{\tau'}$, any moment on $\mathbb{T}_{\tau'}$ is a synchronization point.

Definition 2.5 *A* temporal filter *is a set of time scales $\mathcal{F} = \{\mathbb{T}_\tau\}_{\tau \in T}$ such that $\preceq$ is a total order on $\mathcal{F}$. In other words, $\mathbb{T}_\tau$ and $\mathbb{T}_{\tau'}$ are always comparable by $\preceq$ for any $\tau, \tau' \in T$.*

3 Systems

3.1 Definition

In this section, we give a formal definition of the notion of system which tries to capture the realness of industrial complex systems. Most systems – be that industrial scale technological systems or networks of information processing machines (and probably also certain biological systems) – are too complex to be modelled or analyzed as a whole, but can be treated as the result of the integration of several components. These components tend to be simpler systems that can in their turn be decomposed in the same way. We arrive eventually at the level where all such components are *elementary systems*, i.e. sufficiently simple to be considered independently of their structure. The following key definition is an attempt to capture this as yet intuitive and informal process (on which however relies system engineering in the industry).

Definition 3.1 *A* system *S is recursively defined as the union of the following elements:*

- *an input/output mechanism that consists respectively in:*
  - *an* input channel *x which is capable of receiving – only at moments that belong to a given time scale $\mathbb{T}_{\tau_i}$ called the* input time scale *– data that belong to a given set I, called the* input domain *of S (and also denoted by $In(S)$),*
  - *an* output channel *y which is capable of emitting – only at moments that belong to a given time scale $\mathbb{T}_{\tau_o}$ called the* output time scale *– data that belong to a given set O, called the* output domain *of S (and also denoted by $Out(S)$),*
- *two internal storage mechanisms that consist respectively in:*
  - *an* internal memory *given by a tape indexed by ${}^*\mathbb{N}$ – with a window that can take any value within ${}^*\mathbb{N}$ – which may contain any (non standard) finite number* [5] *of values in a given set M (also denoted by $Mem(S)$), called the* memory domain*,*

[5] Recall that a set is said to be a (non standard) "finite" set if and only if it can be put in bijection with a set of the type $[0, N]$ where N stands for any (either usual or infinitely great) non standard positive integer in ${}^*\mathbb{N}$.

  - *an* internal state set *which is just an arbitrary finite (in the usual standard meaning) set Q (that is also denoted by $State(S)$),*

- *an internal time behavior which is given by*

  - *an* internal time scale $\mathbb{T}_{\tau_s}$ *that refines both the input and the output time scales, i.e. which satisfies to* $\mathbb{T}_{\tau_s} \preceq \mathbb{T}_{\tau_i}$ *and* $\mathbb{T}_{\tau_s} \preceq \mathbb{T}_{\tau_o}$,
  - *an* internal state evolution function $q(t)$ *which maps each element* $t \in \mathbb{T}_{\tau_s}$ *onto some element* $q(t) \in Q$ *(called the value of the internal state at moment* t*),*
  - *an* internal memory evolution function $m(t)$ *which maps each element* $t \in \mathbb{T}_{\tau_s}$ *onto some element* $m(t) \in M^f$ *(called the value of the internal memory at moment* t*)* [6] ,

- *three transition mechanisms that consist respectively in:*

  - *a function* read : $Q \times I \to Q \times M^f$ *that can read a given value on the input channel and write correspondingly a series of values – depending on the status of a state* $q \in Q$ *(updated after the operation) – onto the internal memory,*
  - *a* controller function $\delta : Q \times M \times {}^*\mathbb{N} \to Q \times M \times {}^*\mathbb{N}$ *that allows – as we will see – to replace an element of the internal memory by another one* [7] ,
  - *a function* write : $Q \times M^f \to Q \times O$ *that can read a set of values on the internal memory and write correspondingly a value – that may depend on the status of a state* $q \in Q$ *(updated after the operation) – onto the output channel,*

- *a finite (in the usual standard meaning) set of systems* $Sub(S) = \{S_1, \ldots, S_n\}$*, which are called the* sub-systems *of* S *that are equipped with:*

  - *for each* $k = 1, \ldots, n$*, a function* $\rho_k : Q \times Out(S_k) \to Q \times M^f$ *that reads the output of* S_k*, writes it into the internal memory and changes possibly of internal state,*
  - *for each* $k = 1, \ldots, n$*, a function* $G_k : Q \times M^f \times \bigotimes_{i=1}^{n} Out(S_i) \to Q \times In(S_k)$ *that defines the interactions between the sub-systems,*

  and such that moreover the output and input time scales of all these sub-systems are always refined by the internal time scale of S [8] .

The previous definition is however purely statical and does in particular not give any insight on the time behavior of a given system. To progress in this last direction, we first introduce the notion of instantaneous description of a system.

Definition 3.2 *Let* S *be a system. An* instantaneous description *of* S *is then any quadruple of the type* $d = (t, q, m, i) \in \mathbb{T}_{\tau_s} \times Q \times M^f \times {}^*\mathbb{N}$ *such that one has*

- $t \in \mathbb{T}_{\tau_s}$ *is a moment of the internal time scale* $\mathbb{T}_{\tau_s}$ *of* S,
- $q \in Q$ *is the value* $q(t)$ *of the internal state evolution function of* S *at moment* t,
- $m \in M^f$ *is the value* $m(t)$ *of the internal memory evolution function of* S *at moment* t,

[6] We denote here by M^f the set consisting of all non standard finite (in the non standard meaning) sequences over a set M (i.e. partial functions ${}^*\mathbb{N} \to M$ with non standard finite support).
[7] Depending on the value of a given state of Q which can be also changed by the action of δ.
[8] Forming therefore altogether a temporal filter.

- $i \in {}^*\mathbb{N}$ *is the position of the window of the internal memory of* S *at moment* t [9] .

We are now in position to define the time behavior of a system which just appears as a series of instantaneous descriptions (completed by the time evolutions of the values of the input and output channels that we did not integrated in these descriptions) indexed by its internal time scale and submitted to some natural transition constraints.

Definition 3.3 *Let* S *be a given system. A* time behavior *associated with* S *is then any family* $d(t) = (t, m(t), q(t), i(t))$ *of instantaneous descriptions of* S *where* t *describes the internal time scale* $\mathbb{T}_{\tau_s}$ *of* S *and where one always pass from* $d(t)$ *to* $d(t+\tau_s)$ *by executing one of the following possible transition actions:*

1. *for any* $k = 1, \ldots, n$ *and at every synchronization point* $t + \tau_s$ *between all concerned time scales, update the input of* S_k *and the current internal state of* S *by setting*

$$(q(t+\tau_s), x_k(t+\tau_s)) = G_k\Big(q(t); m(t); y_1(t), \ldots, y_n(t)\Big), \tag{1}$$

   where x_k *denotes the input channel of the sub-system* S_k *and where each* y_i *stands for the output channel of* S_i *correspondingly,*

2. *for any* $k = 1, \ldots, n$ *and at each synchronization point* t *between the internal time scale* $\mathbb{T}_{\tau_s}$ *of* S *and the output time scale of* S_k*, read the output of* S_k *and update both the internal memory of* S *– beginning at the current position of its associated window – and the current value of its internal state by setting*

$$(q(t+\tau_s), m(t+\tau_s)_{i \geq i(t)}) = \rho_k\Big(q(t); y_k(t)\Big), \tag{2}$$

3. *at each synchronization point* t *with the input time scale, perform a* read *operation – depending on the value of the internal state of the system – and update both the internal memory of* S *beginning at the current position of its window and this internal state:*

$$(q(t+\tau_s), m(t+\tau_s)_{i \geq i(t)}) = \mathrm{read}\Big(q(t); x(t)\Big), \tag{3}$$

4. *at each synchronization point* t *with the output time scale, perform a* write *operation by taking into account both the internal state of the system (that is updated after the operation) and the values of the internal memory that begin at the current position of its window:*

$$(q(t), y(t)) = \mathrm{write}\Big(q(t); m(t-\tau_s)_{i \geq i(t-\tau_s)}\Big), \tag{4}$$

5. *at any other moment* t *on* $\mathbb{T}_{\tau_s}$*, update the internal state, the value of the current position of the internal memory of* S *and the current position of its associated window:*

$$(q(t+\tau_s), m(t+\tau_s)_{i(t)}, i(t+\tau_s)) = \delta\Big(q(t), m(t)_{i(t)}, i(t)\Big). \tag{5}$$

Time behaviors may however not be unique. This remark leads us to the following definition (observe that we will mainly consider deterministic systems in the examples of this paper).

Definition 3.4 *A system* S *is said to be* deterministic *if and only* S *possesses exactly one single time behavior. If it is not the case, the system is said to be* non-deterministic.

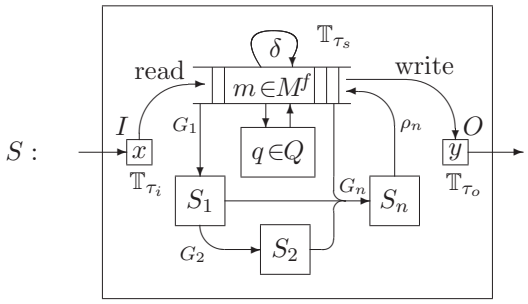


Figure 2: Graphical representation of a system S.

The previous definitions are illustrated by the diagram in Figure 2. For the sake of simplicity and clarity, we will occasionally vary this representation. In particular, we will sometimes omit domains of variables or their names, if the omitted parts are clear from the context.

Note 3.5 One can also consider systems with more than one tape, that is to say with several internal memory variables, which corresponds to saying that M is a direct product of different independent sets. In this case, there has to be a window as above per tape.

Note 3.6 The sub-system graph of S is the graph which is formed by taking $Sub(S)$ as vertices and with edges defined by $(\mathcal{G}_i)_{i=1,\dots,n}$, i.e. such that there is an edge going from S_i to S_j if and only if x_j depends on y_i according to Equation (1). This last graph can contain cycles and does not necessarily have to be connected. Cycles in the sub-system graph represent feedback loops common to the sub-systems that are involved in such cycles.

Note 3.7 Note that we suppose that each time scale $\mathbb{T}$ of an input or an output channel of a sub-system of a given system S is always refined by the internal time scale $\mathbb{T}_{\tau_s}$ of this system, i.e. that one has $\mathbb{T}_{\tau_s} \prec \mathbb{T}$, even if such a time scale is *free for interaction* which means that the corresponding channel is never implied in one of the relations (1).

In order to classify the level of complexity of a system, we now introduce the notion of *order* of a system. A system shall said to be of *order* N if it is constructed using only sub-systems of order $N-1$ and less. Systems of *zero-th order* are called *elementary* systems.

Definition 3.8 *We define the* order *of a system S by setting*

$$ord(S) = \begin{cases} 0 & \text{if } Sub(S) = \emptyset\,, \\ 1 + \max\{ord(S') \,|\, S' \in Sub(S)\} & \text{otherwise}\,. \end{cases}$$

We are now in position to introduce the notion of well defined system (which will in fact be the only kind of systems that we shall consider in the sequel).

Definition 3.9 *A system S is said to be* well defined *if there exists a positive standard integer $N \in \mathbb{N}$ such that $N = ord(S)$.*

[9] Whose value is given by $m(t)$.

Note that each time behavior of a system implicitly defines an input/output relation of the following type (where we took here all the notations of the previous definitions):

$$y(t_0+\tau_o) = \mathcal{F}\Big(x(u), m(t), q(t) \,\Big|\, u, t \in [t_0, t_0+\tau_o[\Big)\,, \tag{6}$$

where t_0 describes the output time scale $\mathbb{T}_{\tau_0}$ of the system and where $u, t \in [t_0, t_0+\tau_o[$ signifies that each of these variables is taken between t_0 and $t_0+\tau_o$ (without reaching this last upper limit) respectively on the time scale $\mathbb{T}_{\tau_i}$ and $\mathbb{T}_{\tau_s}$. For more simplicity, we will rewrite equivalently the input/output relation (6) in the following simplified functional form

$$y = \mathcal{F}(x; q, m)\,. \tag{7}$$

It is easy to see that the function $\mathcal{F}$ is uniquely associated with a given system S if and only if S is a deterministic system. Observe also that $\mathcal{F}$ must obviously (by construction) always be a causal function, i.e. each value $y(t)$ depends only on the values $x(t')$ with $t' < t$.

Note 3.10 In several classical models of systems – such as dynamical systems (see [12]) or synchronous systems (see [4]) – non causality is however a real problem that can eventually occur. Such a situation is indeed always the direct consequence of the collapse between system feedbacks and the hypothesis – which is implicit in usual continuous modelling (and corresponds to standardization with respect to our approach), but totally explicit in synchronous modelling – that no time is required to realize the internal treatments of a given system.

Two main and complementary approaches to system design, namely specification and engineering, can be distinguished. While the latter is interested in full details in the effective way a system can be constructed, the former only deals with formal requirements on the input/output relation of a system. In other words, the system specification approach considers a system as a "black box" with a precise functional behavior, but without trying to know how such a behavior is obtained. This approach is fundamental in the industry: a computer manufacturer will for instance be able to give the specification of a given wireless interface to different suppliers that may realize different wireless computer interfaces from their structural point of view, as soon as the input/output relations specified by the manufacturer are exactly the same. Such situations are modelled by the following definition for equivalent systems.

Definition 3.11 *Two systems S_1 and S_2 are said to be* equivalent *if and only if one both has $In(S_1) = In(S_2)$, $Out(S_1) = Out(S_2)$ and the following conditions, i.e.*

$$\forall\ t \in \mathbb{T}_{\tau_i},\ x_1(t) = x_2(t)\ \Rightarrow\ \forall\ t \in \mathbb{T}_{\tau_o},\ y_1(t) = y_2(t)\,,$$

where (x_1, y_1) and (x_2, y_2) stand respectively for the input and output channels of S_1 and S_2.

Note 3.12 The previous definition has mainly a meaning when S_1 and S_2 are deterministic.

3.2 Elementary systems

We will now study more in details some important classes of elementary systems, i.e. of systems of order 0. We will in particular show that our framework allows already to capture at this level different classical interesting classes of systems of very different nature. Note finally that we will concentrate in this subsection on elementary systems of the following three types (using here a general terminology which is not specially reserved to elementary systems):

1. *software systems* model phenomena observed mostly in information technologies: they are characterized by the fact that their three defining time scales are all discrete,
2. *physical systems* are characterized by the fact that their three defining time scales are all continuous: they are generally used to model real-life physical systems,
3. *hybrid systems* mix finally – by definition – both discrete and continuous time scales.

3.2.1 Elementary software systems

Elementary software systems have only discrete time scales. Their input and output spaces will be called alphabets (which refers to the notion of set of letters or symbols). We assume that an elementary software system is equipped with a tape and a corresponding window that indicates the position of its head (this definition can be generalized to take into account several tapes).

As depicted on Figure 3, elementary software systems are receiving – on an input channel – data within some input alphabet I at a rate given by the input time scale $\mathbb{T}_{\tau_i}$. They are also emitting – on an output channel – data that belong to some output alphabet O at a rate given by the output time scale $\mathbb{T}_{\tau_o}$. Moreover any elementary software system has also the right to perform – at a rate given by its internal time scale $\mathbb{T}_{\tau_s}$ – a number of internal actions controlled by the value of an internal state $q \in Q$, that is to say:

- read an input data x, transform it into a word (depending only on x) on the tape alphabet and write it finally on the tape beginning at the current position of its window,
- change the value of the element of the tape that is obtained by looking on the current position of its window (that can be updated after the operation),
- write an output data y by taking a word w on the tape beginning at the current position of its window and transforming it (depending only on w) into y.

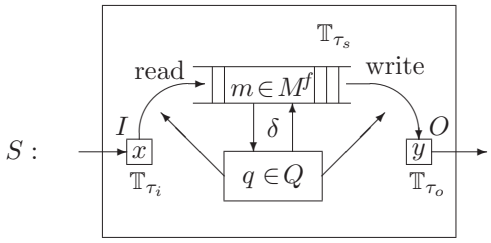


Figure 3: Graphical representation of an elementary software system S.

As we can see, elementary software systems are therefore just a slight generalization of usual Turing machines which is obtained by adding to this classical model a permanent input/output temporal behavior. The reader can indeed easily check that Turing machines (or equivalently recursive functions if one prefers to stay within a functional approach) correspond to the degenerated case of our model where one considers elementary software systems that can only perform a unique read action (or whose input channel will only receive a single input data during all possible moments of time). Note also that as an immediate consequence of this simple observation, we got the undecidability of the existence of a system's output !

Example 3.13 (One element buffer) Let us now present an example of elementary deterministic software system that we will use in the sequel (as a sub-system of an higher order system). Our example consists in a buffer capable of storing only – at each moment of time – one single message out of a message set A. We assume that this buffer has two input channels. On the first input channel, the buffer can only receive either a message $m \in A$ or a distinguished empty message ε. On the second one, it can receive either a write request, that we will denote by '$\uparrow$', or again the distinguished empty message ε. The buffer stores each message it receives on the first channel in a fixed memory cell which is overwritten each time a new non-empty message arrives on the same channel. When the buffer receives a write request, it sends the currently stored message on its output channel. A representation of such a buffer is shown in Figure 4.

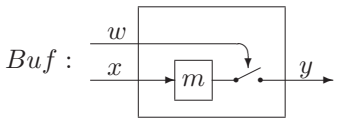


Figure 4: Graphical representation of a one-element buffer.

Let us now describe how to model this simple buffering mechanism by an elementary deterministic software system, denoted by Buf. The input, internal and output time scales of such a system are all discrete with respective time steps $\tau_i = \tau$ and $\tau_s = \tau_o = \tau/2$. In other words, we require the buffer to operate internally on a rate which is twice as fast as its input rate. The input domain of Buf is clearly modelled by $(A \cup \{\varepsilon\}) \times \{\varepsilon, \uparrow\}$ in order to take into account the two entry channels, when the output domain is just equal to $A \cup \{\varepsilon\}$. The memory domain will be equal to A. Finally the internal state set of Buf is defined as $\{r, \varepsilon, \uparrow\}$ (the first state models the reading of the input channel, when the two last ones correspond to the two possible writing decisions on the output channel). Therefore we have:

$$In(Buf) = (A \cup \{\varepsilon\}) \times \{\varepsilon, \uparrow\}, Mem(Buf) = A, Out(Buf) = A \cup \{\varepsilon\}, State(Buf) = \{r, \varepsilon, \uparrow\},$$

The control mechanisms of Buf are now given by the following transition functions:

$$\delta = -, \quad \texttt{read}(r; (x, w)) = \begin{cases} (w; x) & \text{if } x \neq \varepsilon, \\ (w; -) & \text{otherwise}, \end{cases} \qquad \texttt{write}(q; y) = \begin{cases} (r; y) & \text{if } q = \uparrow, \\ (r; \varepsilon) & \text{if } q = \varepsilon, \end{cases} \tag{8}$$

where $-$ means not defined or no action (depending on the situation). Observe that the choice of δ just reflects the fact that the input message is stored in a single cell of the internal memory on which no action can be made. The unique possible time behavior of the buffer consists then just in alternating a `read` and a `write` action at each moment of its internal time scale. ■

3.2.2 Elementary physical systems

Before presenting the notion of elementary physical system, let us first introduce the general framework on which rely this concept. We will indeed suppose that each instance p of a given physical parameter φ (such as mass, distance, kinetic energy, potential energy, torsion energy, temperature, kinetic momentum, etc.) with whom we will deal, can always be both

1. *measured* using a measure function m_φ, which means that one can associate to each such physical quantity p of type φ its measure $m_\varphi(p) \in {}^*\mathbb{R}$,
2. *decomposed into a (non standard) finite sum of infinitesimal quantities*, i.e. a sum of physical quantities p of type φ that have an infinitely small measure $m_\varphi(p) \approx 0$.

The set of all physical quantities of a given type φ is then called the physical domain associated with φ and denoted by $\mathbb{P}_\varphi$. In the same way, the physical infinitesimal domain of type φ – which is denoted by $\mathbb{I}_\varphi$ – consists of all infinitesimal physical quantities of type φ.

Elementary physical systems can now be described exactly in the same way than elementary software systems, i.e. by a mechanism similar to the one given by Figure 3, the only (but fundamental) difference being here that such systems manipulate physical quantities using continuous time scales. An elementary physical system is indeed characterized by the fact that it has

1. continuous input, internal and output time scales,
2. input and output domains that are both equal to the same finite (in the usual meaning) product of physical domains, i.e. both equal to $\otimes_{i=1}^{n} \mathbb{P}_i$ for some finite (standard) positive integer $n \in \mathbb{N}$, where each $\mathbb{P}_i$ stands for a physical domain of a given type,
3. an internal domain which is necessarily equal to $\otimes_{i=1}^{n} \mathbb{I}_i$ where each $\mathbb{I}_i$ stands for the infinitesimal physical domain associated with the physical domain $\mathbb{P}_i$ which is involved both in the corresponding input and the output domain.

For the sake of simplicity, we can of course consider – without any extension of the representation power of our model – that an elementary physical system has a finite (in the usual sense) number of tapes, each of them devoted to some particular infinitesimal physical domain.

Such elementary systems are intended to model real physical systems, considered as transformers of infinitesimal physical quantities. The behavior of an elementary physical system S – in our framework – has indeed to be physically interpreted as follows:

- S receives at each moment of its input time scale (hence infinitely often) a vector x that consists of different physical quantities of given types, i.e. a vector $x \in \otimes_{i=1}^{n} \mathbb{P}_i$ where each $\mathbb{P}_i$ stands for some physical domain; it transforms then each component $x_i \in \mathbb{P}_i$ of x into a (non standard) finite sequence $(x_i^j)_{j=1...N}$ – written on a specific tape of the internal memory of S – of infinitesimal physical quantities within $\mathbb{I}_i$ (i.e. of the same type than $\mathbb{P}_i$) whose sum has the same measure than x_i, i.e. such that

$$\sum_{j=1}^{N} m_i(x_i^j) = m_i(x_i)\,, \tag{9}$$

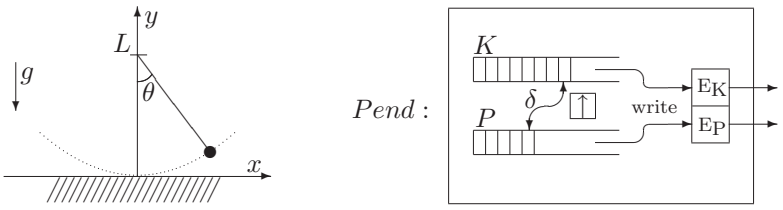
  where m_i stands for the measure function associated with the physical domain $\mathbb{P}_i$,

- S can transform infinitely often, at the rate given by its internal time scale, any infinitesimal physical quantity of a given type written on one of its tapes into another infinitesimal physical quantity of another given type (that can also be stored on another tape),
- S emits at each moment of its output time scale (hence again infinitely often) a vector $y \in \otimes_{i=1}^{n} \mathbb{P}_i$ whose components are obtained by "gluing" altogether sequences of infinitesimal physical quantities (of compatible types) coming from the internal memory of S, by using the reverse process of the initial writing mechanism as described above.

Observe that we can only model in such a way *non dissipative* physical systems since the internal controller of an elementary physical system – modelled by the function δ – is only able to change an elementary physical quantity of known type into another elementary physical quantity of known type ("elementary" being modelled by "infinitesimal" in our approach). Conversely one can also prove that large classical classes of conservative physical systems – such as Hamiltonian

systems (cf. [24]) – can be recovered (as higher order systems) inside our model. We will however not prove here this last result which is quite technical, but rather illustrate on a simple example how to analyze a classical mechanical system as an elementary deterministic physical system.

Example 3.14 (Simple pendulum) Let us now consider a simple pendulum as shown in Figure 5-a. The pendulum consists of a point mass m attached to the point $(0, L)$ by a rigid string of negligible mass and of length L (when hanging freely the pendulum touches the ground).



(a) (b)

Figure 5: Simple pendulum: mechanical (a) and systemic (b) representations

Its motion can be classically described by applying the fundamental principle of dynamics which leads immediately to the following differential equation

$$m\,L\,\dot{\varphi} = -m\,g\,\sin\theta\ , \tag{10}$$

where θ and $\varphi = \dot{\theta}$ stand respectively for the angle formed by the string and the y axis and for the corresponding angular speed. In its turn, Equation (10) is clearly equivalent to the following energy preservation equation (obtained by integrating this last equation)

$$\frac{1}{2}\,m(L\varphi)^2 + mgL\cos\theta = C\,, \tag{11}$$

where the first and the second summand in the left hand side represents respectively the kinetic and the potential energy of the pendulum and where C stands for the initial potential energy of the pendulum, when it is in the point farthest from the y axis with zero angular speed.

An elementary deterministic physical system $Pend$ modelling such a pendulum is shown in Figure 5$-b$. This system takes no input and provides on the output channel a pair of physical quantities that consists respectively in the pendulum's current kinetic and potential energy. Its internal and output time scales are supposed to be the same continuous time scale (with dt as common infinitesimal time step). We define finally the output domain, the memory domain and the internal state set of the system to be respectively equal to

$$Out(Pend) = \mathbb{E}_K \times \mathbb{E}_P\ ,\ Mem(Pend) = \{0, de_K\} \times \{0, de_P\}\ ,\ State(Pend) = \{x, s\} \times \{\uparrow, \downarrow\}\,,$$

where $\mathbb{E}_K$ and $\mathbb{E}_P$ denote respectively the two physical domains which are associated with kinetic and potential energy, where de_K and de_P stand respectively for two infinitesimal quanta of kinetic and potential energy – with a common measure $m_K(de_K) = m_P(de_P) = de \in \mathbb{I}$ – and where the x, s and the arrow states should respectively be interpreted as the two possible internal actions of the pendulum (exchanging energy on its two tapes – see below – or sending physical quantities to the output channel) and as the two possible directions of the pendulum's motion. This system has therefore two tapes K and P, each of them containing a (non standard)

finite number of copies of the corresponding infinitesimal quantum of energy (the other parts of the two tapes being equal to 0). Moreover the memory of our system evolves in such a way that it always contains the same global number (necessarily infinitely great) $N \in {}^*\mathbb{N}$ of energy quanta (which satisfy to the energy conservation condition $N\,de = C$). The system's behavior consists then essentially in taking at each moment of time one quantum of energy from one tape, depending on the direction the pendulum is moving, and putting it on the other one, until the working tape is empty. The corresponding controller function δ is given below.

$$\delta((x,\uparrow),(k,p),(i_K,i_P)) = \begin{cases} \Big((s,\uparrow),(0,de_P),(i_K{-}1,i_P{+}1)\Big) & \text{if } (k,p){=}(de_K,0) \text{ and } i_K{>}1, \\ \Big((s,\downarrow),(0,de_P),(0,i_P{+}1)\Big) & \text{if } (k,p){=}(de_K,0) \text{ and } i_K{=}1, \end{cases}$$

$$\delta((x,\downarrow),(k,p),(i_K,i_P)) = \begin{cases} \Big((s,\downarrow),(de_K,0),(i_K{+}1,i_P{-}1)\Big) & \text{if } (k,p){=}(0,de_P) \text{ and } i_P{>}1, \\ \Big((s,\uparrow),(de_K,0),(i_K{+}1,0)\Big) & \text{if } (k,p){=}(0,de_P) \text{ and } i_P{=}1, \end{cases}$$

where i_K and i_P are the cursors of tapes K and P. The `write` function is then defined as the constructer of the two global physical quantities (i.e. kinetic and potential energy) that can be obtained by summing all infinitesimal quanta that respectively exist on tapes K and P [10]. This last function works only in state $(s,\updownarrow)$ and ends in state $(x,\updownarrow)$ (where $\updownarrow$ stands for any type of arrow). The unique possible time behavior of our (deterministic) system consists hence just in alternating permanently a tape exchange operation with a write operation.

It is then immediate to see that the elementary physical system that we just defined, always satisfies – by construction – to the energy conservation equation

$$E_K + E_P = C \ , \tag{12}$$

where E_K and E_C stand for the measures of the kinetic and the potential energy of the pendulum, that is exactly equivalent to Equation (11). Note however that our approach does not connect the physical quantities that we manipulated (here kinetic and potential energy) with the high level parameters φ and θ that were used in writing down Equation (11).

To fill this gap, we must interpret the pendulum as a new deterministic system *Newpend* of higher order that contains the previous elementary physical system *Pend* as a sub-system. This new system has another sub-system which makes alternatively the two only operations:

- it reads the outputs of *Pend* and transforms them – by applying the two associated measure functions – into (non standard) real values k and p that are stored in its internal memory,
- it takes the two last non standard real values k and p and writes them on its output channel by applying the following transformation

$$\texttt{write}\,(k,p) = \left(\frac{1}{L}\sqrt{\frac{2\,k}{m}}\,,\, \arccos\frac{p}{mgL} \right) ,$$

  which can be expressed within our model by making use of adapted sub-systems, due to the fact that we are only dealing here with analytic transformations (see Section 4).

If one identifies the output channels of this last sub-system to the output channels of *Newpend*, it is then obvious to see that *Newpend* produces on its own output channels the pair (φ,θ) in such a way that the energy conservation Equation (11) is always fulfilled. ■

[10] We can easily assume that the cardinality of these two families of infinitesimal quanta is respectively given by i_K and i_P. Under this hypothesis, the `write` function can be more precisely defined – independently of the value of the current internal state of the system – by setting $\texttt{write}(K,P) = (i_K\,de_K, i_P\,de_P)$.

3.2.3 Elementary hybrid systems

Elementary hybrid systems are systems of order 0 which can transform continuous behaviors into discrete ones (or vice-versa). They can therefore naturally be used for modelling interfaces between software and physical systems. The two following examples illustrate how to interpret within our approach two classical interfaces of this kind – here a sampler and a modulator – that are totally fundamental in practice.

Example 3.15 (Sampler) A sampler is a mechanism that takes a continuous time input function and produces a discrete sequence of samples of its values. It can be modelled by an elementary deterministic hybrid system H_τ which is parameterized by the time step $\tau > 0$ of its discrete internal time scale. The output time scale of this system is also discrete with a double time step $2\,\tau$. On the other hand, the input time scale of H_τ is continuous with an infinitesimal time step $dt = \tau/N$ (where $N \in {}^*\mathbb{N}$ is a given infinitely great non standard integer). The input, output and memory domains of H_τ are all equal to ${}^*\mathbb{R}$, when its internal state set has just two states r and w that should be interpreted as the two possible internal actions of the sampler (reading on the input channel, writing on the output channel), that is to say:

$$In(Sampler) = Out(Sampler) = Mem(Sampler) = {}^*\mathbb{R}\,,\ State(Sampler) = \{r, w\}\ .$$

The control mechanisms of H_τ are now given by the following transition functions:

$$\texttt{read}\,(r; x) = (w; x)\,, \quad \texttt{write}\,(w; y) = (r; y)\,, \quad \delta = -\,, \tag{13}$$

where $-$ means that no action should be done (on the tape). The unique temporal behavior of this deterministic hybrid system is now obvious since H_τ can just make a succession of internal read and write actions. At each moment of its internal time scale which is not a synchronization point with the time scale of the output channel, i.e. at every $(2\,k + 1)\,\tau$ with $k \in \mathbb{N}$, the system uses the `read` function to memorize the value on the input channel inside a fixed cell of its internal memory (such a moment is always a synchronization point with the time scale of the input channel). On the other hand, the systems outputs – with the `write` function – the value currently stored in this cell at each synchronization point with the output time scale, i.e. at every $2\,k\,\tau$ with $k \in \mathbb{N}$, thus producing a discrete sequence out of a continuous one. ■

Example 3.16 (Modulator) The action of a modulator is more or less reciprocal to that of a sampler and consists in converting a discrete sequence of real numbers into a continuous function by making use – at a discrete rate – of a pulse shape $p_\tau(t)$, which will be considered here – for the sake of simplicity (see below for more details on this hypothesis) – as a given function with interval $[0, \tau]$ as support. A modulator can then be modelled by an elementary deterministic hybrid system Mod_{p_τ} which is parameterized by this last continuous function. The input, output and memory domains of such a system are equal to ${}^*\mathbb{R}$, when its internal state set has exactly four states r, b, w and a that correspond to the four allowed internal actions of the system (reading on the input channel, making a blank action – i.e. nothing – , writing on the output channel, adding the value 1 in the internal memory), that is to say:

$$In(Mod) = Out(Mod) = Mem(Mod) = {}^*\mathbb{R}\,,\ State(Mod) = \{r, b, w, a\}\,.$$

The input time scale of Mod_{p_τ} is then a discrete time scale of time step τ, when its internal and output time scales are both continuous with respective time steps $dt_s = \tau/(2\,N)$ and $dt_o = \tau/N$ where N stands for a fixed infinitely great positive integer within ${}^*\mathbb{N}$. All the cells of the internal

memory of Mod_{p_τ} are initialized with 0 values and the control is initially put on the 1-cell. The control mechanisms of this system are then given by the following transition functions:

$$\begin{aligned} \texttt{read}\,(r\,;y) &= (b\,;y)\,, \\ \texttt{write}\,(w\,;(x,y)) &= (a\,;y\cdot p_\tau(x\,dt_o))\,, \\ \delta(b;y;1) &= (w;y;0) \\ \delta(a;x;0) &= \begin{cases} (w;x+1;0) & \text{if } x\neq N-1\ , \\ (r;0;1) & \text{if } x=N-1\ . \end{cases} \end{aligned} \tag{14}$$

Note that this system makes only use of the two first cells of its internal memory: the 0-cell is used to store internal computations (here the sequence of all non standard integers from 0 to $N-1$) when the 1-cell stores the input values. The unique temporal behavior of Mod_{p_τ} consists then to read a value of its input channel (at each possible synchronization point) and store it in the 1-cell of its internal memory, to make a blank operation (i.e. doing nothing during one single internal clock tip) and then to apply alternatively a `write` operation and a δ controlled transition (up to reading a new entry value and re-entering in the same processing cycle).

Note also that though the pulse shape p_τ in the above example gives us a degree of freedom in the way we can modulate the continuous output, the most realistic choice is unfortunately just to take $p_\tau(t)=1$ for any $t\in[0,\tau]$. The examples that correspond to practical situations are indeed obtained if the pulse shape has support bigger than $[0,\tau]$, which requires to add several consecutive input values. This technique would however lead to a more complicated modelling, which explains why we restrained ourselves to the simpler system presented above. ■

3.3 An example of higher order system

To show the potentiality of our approach in practice, we will now model within our framework a simplified version of a radio transmission system (which appears to be an higher order system in the meaning of Definition 3.8).

Example 3.17 (Simplified radio transmission) We will now show how to model a communication transmitter taking messages from a buffer Buf – as described in Example 3.13 (whose notations will be taken here) – and transmitting them over a radio channel. To achieve this aim, we must first introduce an encoder component Enc which reads messages from the buffer, converts them into binary form and encodes the resulting sequence of bits – by blocks of N_b bits – into complex symbols [11]. Whenever the encoder has less than N_b bits to work on, it sends a write request to the buffer. The systemic organization of this component is given in Figure 6.

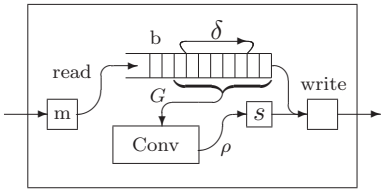


Figure 6: Description of the encoder component.

[11] The simplest example of such a process is the Binary Phase Shift Keying (BPSK) modulation protocol that converts every bit $b\in\mathcal{B}=\{0,1\}$ into the real number $(-1)^b$ (for more details, see for instance [26] or [16]).

The corresponding (non deterministic) system can indeed be defined as follows. We first suppose that the internal time scale of *Enc* has a discrete time step which is N_b times smaller than the rest of the global radio transmission system to which it will belong. The input domain of the encoder is just A, the same set of messages than the buffer to which it is connected. The output domain is taken to be $\mathbb{C} \times \{\varepsilon, \uparrow\}$ in order to model the fact that two types of output information can be sent on two different channels (i.e. a complex encoding or a write request). The memory domain is defined as $\{0,1\}^* \times \mathbb{C}$ since one must be able to store (on two different tapes) both sequences of bits and complex numbers. Finally the internal state set is reduced to two elements q_0 and q_1 (see the explanations below). These elements can be summarized by setting:

$$In(Enc) = A\,,\ Out(Enc) = \mathbb{C} \times \{\varepsilon, \uparrow\}\,,\ Mem(Enc) = \{0,1\}^* \times \mathbb{C}\,,\ State(Enc) = \{q_0,\, q_1\}\ .$$

We also suppose that *Enc* possesses a unique sub-system *Conv* – i.e. that $Sub(Enc) = \{Conv\}$ – which can convert blocks of N_b bits into complex symbols according to a given table and which has the same input time scale than *Enc* (the second state of *Enc* is in particular used to interconnect the reading of a bit on the tape of *Enc* with its sending to *Conv*). The control mechanisms of *Enc* can now be given by means of the following transition functions:

$$\begin{aligned}
\texttt{read}\,(q_0; m) &= (q_0; b \cdot \overline{m})\,, \\
G(q_0; (b, s)) &= (q_1; b_1)\,, \\
\delta(q_1; (b, s); 0) &= (q_0; (b \ll 1, s); 0)\,, \\
\rho(q_0; (b, s); y) &= (q_0; (b, y))\,, \\
\texttt{write}\,(q_0; (b, s)) &= \begin{cases} (q_0; (s, \uparrow)) & \text{if } |b| < N_b\ , \\ (q_0; (s, \varepsilon)) & \text{otherwise}\ , \end{cases}
\end{aligned} \tag{15}$$

where $m \in A$ and $\overline{m} \in \{0,1\}^*$ stand respectively for the message at the input of the encoder and for its binary form and where $(b, s) \in \{0,1\}^* \times \mathbb{C}$ is the current value of the internal memory of the system, with b denoting the sequence of bits that is currently stored on the main tape of the encoder and s being the complex symbol produced by the *Conv* sub-system. We also denoted above the concatenation product by $\cdot$, the first bit of the sequence b by b_1 and the shift of b by one bit by $b \ll 1$. Note finally that we did not represented here – for the sake of clarity – the precise behavior of δ on the main tape of *Enc* (that we modelled just as a single cell that can contain sequences of bits on which concatenation products can be directly applied).

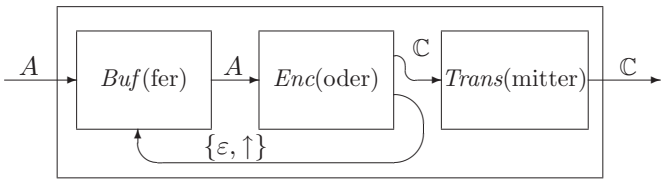


Figure 7: Graphical description of a simplified radio transmission chain.

The transmitter in its turn receives on the input a sequence of complex symbols $(s_k)_{k\geq 0}$ at a discrete rate and generates on the output a radio signal of the form

$$u(t) = \sum_{k=0}^{\infty} s_k\, p(t - t_k)\,,$$

where $p(t)$ is a pulse shape function and t_k is the moment when the k-th complex symbol is sent. This new operation can be realized by a modulator similar to that of Example 3.16. Composing now, as shown in Figure 7, the three components we introduced, one obtains the simplified radio transmitter system that we wanted to model.

Observe that in the above example, the nature of the first two components (that is to say Buf and Enc) is completely discrete, whereas that of the third component ($Trans$) is hybrid (discrete input with continuous output). This should be interpreted within the classification that we introduced in Section 3.2, as Buffer and Encoder are purely logical, whereas Transmitter serves as an interface between software and physical environments. ■

4 Conclusion

We tried to show in this paper that it is possible to construct a general unified theory of systems which may give a common framework to deal both with continuous and discrete systems (that are the two core kinds of systems used in engineering modelling). Note that our theory allows to take into account large classes of classical systems. For instance, (controlled) causal dynamical systems (see [12]) can be recovered within our framework as the standardization of suited physical-like deterministic systems, if one supposes that the observability function and the vector fields implied in the definition of such systems (see formula (III.2) of Section III.2 of [12]) are rationally analytic (which means that their Taylor development has only coefficients that are rational functions of their generic integer parameters) [12] . In the same way, synchronous systems – which are the discrete equivalents of causal dynamical systems – can also be modelled in our framework by software deterministic systems with the same input and output time scales.

However our approach does not reduce to re-interpret classical classes of systems. It also leads to the introduction of new classes of systems (such as the elementary software and physical systems that were discussed in Sections 3.2.1 and 3.2.2) and to lots of new questions (can one develop a Λ-calculus formalism for systems ? what are the "good" system sub-families ? can one construct a complexity theory for systems ? what are really the differences existing between deterministic and non-deterministic systems ? etc.) that should now be studied more in details in order to understand more deeply the notion of system.

Acknowledgements

The two authors would sincerely like to thank both Matthieu Martel and Marc Pouzet for the numerous discussions we had together during the writing of this paper, leading to a number of important suggestions and modifications of point of view.

[12] To recover this family of systems, one should transform the differential equation (III.2) of [12] into a finite difference equation that involves an infinitesimal time step. To see that this last equation expresses the functional behavior of a system, the key problem is then just to prove that any $A(q)$ can be computed by a system when A is a rationally analytic vector field. This last property can then be reduced to prove that (non standard) finite sums and products of non standard real numbers can be realized by a system, which can be done by using physical like systems (cf. Section 3.2.2). The only technical difficulty lies in fact in proving that one can design such a system for realizing the product of two non standard real numbers which can be done by implementing an adapted Euclidian algorithm (two non standard real numbers u and v being respectively infinitely closed to known multiples K and L of a given infinitesimal $1/N \in \mathbb{I}$ with $N \in {}^*\mathbb{N}$, one can find the non standard integer M such that $u\,v \approx M/N$ by computing the quotient of the non standard Euclidian division of $K\,L$ by N).

References

[1] Abrial J.R., *The B-book – Assigning programs to meanings*, Cambridge University Press, 1996.

[2] Alur R., Courcoubetis C., Halbwachs N., Henzinger T.A., Ho P.H., Nicollin X., Olivero A., Sifakis J., Yovine S., *The algorithmic analysis of hybrid systems*, Theor. Comp. Sci., 1995; **138** (1): 3–34.

[3] Barwise J., *Handbook of Mathematical Logic*, Studies in logic and the foundations of mathematics, **90**, North Holland, 1977

[4] Benveniste A., Caspi P., Edwards S.A., Halbwachs N., Le Guernic P., de Simone R., *The Synchronous Languages Twelve years later*, Proc. of the IEEE, Special issue on Embedded Systems, 2003; **91** (1): 64–83.

[5] Bérard B., Bidoit M., Laroussinie F., Petit A., Schnoebelen P., *Vérification de logiciels – Techniques et outils du model-checking*, Vuibert Informatique, 1999.

[6] Booch G., Rumbaugh J., Jacobson I.,, *The Unified Modeling Language User Guide*, Addision Wesley, 1999.

[7] Börger E., Stärk R., *Abstract state machines – A method for high-level system design and analysis*, Springer, 2003.

[8] Cha D.P., Rosenberg J.J., Dym C.L., *Fundamentals of Modeling and Analyzing Engineering Systems*, Cambridge University Press, 2000.

[9] Cutland N., *Nonstandard analysis and its applications*, London Mathematical Society Student Texts, **10**, Cambridge University Press, 1988

[10] D'Alembert, J. Le Rond dit, *Article "Différentiel"*, in Diderot D. et D'Alembert J. Le Rond dit, Eds.: Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers, Briasson, David, Le Breton et Durand, Paris, 1754, Tome **4**, pp. 985–989.

[11] Diener F. Reeb G., *Analyse non standard*, Hermann, 1989

[12] Fliess M., *Fonctionnelles causales non linaires et indtermines non commutatives*, Bull. Soc. Math. France, 1981; **109**: 3–40.

[13] Henzinger T.A., *The theory of hybrid automata*, in Proceedings of the 11th Annual IEEE Symposium on Logic in Computer Science, LICS'96, IEEE Society Press, 1996, pp. 278–292.

[14] Jensen C.U., Lenzing H., *Model Theoretical Algebra*, Gordon and Breach, 1989.

[15] Kowalewski S., Stursberg O., Fritz M., Graf H., Hoffmann I., Preussig J., Remelhe M., Simon S., Treseler H., *A case study in tool-aided analysis of discretely controlled continuous systems: the two tanks problem*, In Proceedings of the 5th International Workshop on Hybrid Systems, 1997, pp. 165–177.

[16] Krob D., Vassilieva E.A., *Performance analysis of modulation with diversity – A combinatorial approach II: Bijective methods*, Disc. App. Maths., 2005; **145**, (3): 403-421.

[17] Lakatos I., *Preuves et réfutations*, Hermann, 1984.

[18] Lakatos I., *Proofs and Refutations*, The British Journal for the Philosophy of Science, 1963-1964; Vol. XIV: 1–25, 120–139, 221–45, 296–342.

[19] Lavagno L., Martin G., Selic B., *UML for real – Design of embedded real-time systems*, Kluwer, 2003.

[20] Marwedel P., *Embedded systems design*, Kluwer, 2003.

[21] Mathworks, *Mathlab and Simulink*; `http://www.mathworks.com`.

[22] Meinadier J.P., *Ingénierie et intégration de systèmes*, Hermès, 1998.

[23] Meinadier J.P., *Le métier d'intégration de systèmes*, Hermès-Lavoisier, 2002.

[24] Meyer K.R., Hall G.R., *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Applied Mathematical Sciences, **90**, Springer Verlag, 1992.

[25] Mosterman P.J., *An overview of hybrid simulation phenomena and their support by simulation packages*, in Hybrid Systems: Computation and Control, Lectures Notes in Computer Science, Springer Verlag, 1999, pp. 165–177.

[26] J. Proakis, *Digital Communications*, 3rd Edition, McGraw-Hill, 1995.

[27] Robinson A., *Non Standard Analysis*, North-Holland, 1966

[28] Rumbaugh J., Jacobson I., Booch G., *The Unified Modeling Language Reference Manual*, Addision Wesley, 1999.

[29] Samad T., Balas G., Eds., *Software-Enabled Control*, John Wiley, 2003.

[30] Satzinger J.W., Jackson R.B., Burd S., Simond M., Villeneuve M., *Analyse et conception de systmes d'information*, Les éditions Reynald Goulet, 2003.

[31] Schneider K., *Verification of reactive systems – Formal methods and algorithms*, Springer, 2004.

[32] Schwartz L., *Théorie des distributions*, Hermann, 1966.

[33] Severance F.L., *System modeling and simulation – An introduction*, John Wiley, 2001.

[34] Sommerville I., *Software Engineering*, Addison Wesley, 6th Edition, 2001.

[35] Telelogic, *Doors*; `http://www.telelogic.com/index.cfm`.

[36] Wordsworth J.B., *Software engineering with B*, Addison-Wesley, 1996.

[37] Zaytoon J., Ed., *Systmes dynamiques hybrides*, Herms, 2001.

[38] Zeigler B.P., Praehofer H., Gon Kim T., *Theory of modeling and simulation – Integrating discrete event and continuous complex dynamic systems*, Academic Press, 2000.

A THEORY-BASED DYNAMICAL MODEL OF INNOVATION PROCESSES

David Lane, Roberto Serra*, Marco Villani, Luca Ansaloni

Department of Social, Cognitive and Quantitative Sciences
Modena and Reggio Emilia University
Via Giglioli Valle 9, I-42100 Reggio Emilia

MODEL OF INNOVATION PROCESSES

Key words: multi-agent based model, innovation processes, economy, design of complex artificial systems

* Corresponding author:
prof. Roberto Serra
Department of Social, Cognitive and Quantitative Sciences
Faculty of Communications and Economics
Modena and Reggio Emilia University
via Giglioli Valle 9; I- 42100 Reggio Emilia - Italy
tel. +39-0522-522433 (operator 522511), fax 522512
e-mail: serra.roberto@unimore.it

Abstract

We present an agent-based model of innovation processes, based upon a theory of innovation by Lane and Maxfield. The theory inspires and constrains the features of the model, thus reducing the *embarasse de richesse* that is one of the major methodological problems of agent-based modeling. Artifacts are produced by agents using recipes; the basic dynamics, absent innovation, is one of production and sales, where the external world supplies "raw materials" and external demand. Depending upon the initial conditions, self-sustaining cycles of production and exchange can emerge among the agents. Innovation – that is, the generation of new recipes, in particular desired directions, called "goals" – results in substantial modification of the system dynamics. Two innovation regimes are introduced: a "lonely" mode, in which each agent tries to introduce new products by itself, and a "relational" mode, in which two agents can improve their reciprocal knowledge and can decide to try to jointly develop a new artifact.

1 Introduction

Modelling social and economic phenomena is a major challenge. In recent years agent-based models (ABMs) have attracted considerable attention, as they allow one to escape from the constraints of the "representative agent" and to account for agent heterogeneity (see e.g. [1-6] and further references quoted therein) . ABMs are well suited to bridge the gap between hypotheses concerning the microscopic behaviour of individual agents and the emergence of collective phenomena in a population composed of many interacting agents. Moreover, by requiring that hypotheses be stated precisely, ABMs may improve the development of the theory itself

Attractive as agent-based models may be, they still face formidable problems. In particular they allow the modeller too much freedom with respect to the design of agents and their interactions. ABMs often have too many variables and parameters. Moreover, while in some cases experimental data for model testing are available, in many interesting cases data are scarce and do not sufficiently constrain the model. In these cases, what is needed to get meaningful results is a principle which limits the degrees of freedom of the modeller. While Ockham's razor is of the outmost importance, it does not seem to be sufficient by itself for the purpose of limiting the *embarasse de richesse* of ABMs.

In this work we consider another approach, namely that of relying upon an existing theory of the phenomenon in order to constrain our modelling options. The theory, which is qualitative in nature, provides the basic entities of the model and predicates about their relationships. The model represents a simplified universe inhabited by (some of) the theoretical entities. Simplification is necessary in order to deal with manageable systems: we do not look for an all-encompassing model, but we rather imagine a family of models which capture different features of the theory.

The phenomenon which is considered in this paper is that of innovation, and the theory we refer to has been initially developed by Lane and Maxfield [7,8], and is under further development within the UE-FET project Iscom (Information Society as a Complex System:), whose approach is based on the interaction between theory development, case studies and model (see the project website www.iscom.unimo.it). We will refer for brevity to this theory as the Lane-Maxfield (LM) theory, while the model will be called I_2M (Iscom Innovation Model).

A very brief outline of some of the major features of the LM theory will be presented in section 2. The theory constrains the model, and the main constraints will be summarized at the end of that section. Section 3 describes the main features of the model. Section 4 describes the characteristics of a software simulator that implements parts of the model, while section 5 presents the results of some simulations. Finally, section 6 contains comments on the results achieved so far and indications for further work.

2 Indications from the theory

In the work of Lane and Maxfield [7,8], artifacts are given *meanings* by the agents that interact with them, but these meanings cannot be understood without taking into account the *roles* which different agents can

play. Thus, artifacts may be given different meanings by different agents, or by the same agents at different times.

Therefore, LM theory can be seen as a theory of the interpretation of innovation. According to LM, a new interpretation of a potential artifact functionality can be put forth in the context of so-called generative relationships. By interacting, a few agents come to invent and share this interpretation, based on the discovery of different perspectives and use of existing artifacts. The generative potential of a relationship may be assessed in terms of the following criteria:

- heterogeneity: the agents are different from each other, they have different features and different goals; the heterogeneity is not so intense as to prevent communication and interaction
- aligned directedness: the agents are all interested in operating in the same region (or in neighbouring regions) of agent-artifact space
- mutual directedness: the agents should be interested in interacting with each other.

Moreover, Lane and Maxfield discuss two further features that deal with organizational issues, i.e. permissions and action opportunities. The former refers to the fact that the agents are authorized to engage in such a relationship, the latter to the possibility of moving from "talk" to action.

Lane and Maxfield further argue that, in a situation where innovations happen at a very fast pace, predicting the future is impossible; so a better strategy would be to identify those relationships that have the potential for generativeness, and to foster them in order to effectively explore the new opportunities they can give rise to. It is therefore very important to be able to estimate the "generative potential" of the existing and prospective relationships.

The LM theory of innovation is highly sophisticated in describing the interactions between different players in innovation processes, and it cannot be entirely mapped onto a specific computer-based model. Therefore, the modelling activity aims at developing models that are based on abstraction of some key aspects of this theory, which is of a qualitative nature.

The basic requirements for the model derived from the theory can be summarized as follows:

1. the meanings of artifacts must be generated within the model itself: since the LM theory claims that new meanings are generated through interactions among agents and artifacts, it would be inappropriate here to resort to an external oracle to decide a priori which meanings are better than others
2. the roles of agents must also be generated within the model: indeed the LM theory claims that also new roles are generated through interactions among agents and artifacts
3. agents must interact with artifacts and with other agents: interacting with artifacts only would prevent the possibility of describing agent-agent relationships
4. an agent should be able to choose the other agents with whom to start a relationship; in general, an agent will be able to handle a finite number of relationships at a time, and it will choose a subset of the other agents as its partners. Agents must be allowed to cut a disappointing or unsatisfactory relationship to look for a better one
5. an agent should be able to have different degrees of interaction with another agent: in this way it will be possible to tune the intensity of these relationships
6. some agent-agent relationships should be generative in character, i.e. they should be able to lead to new attributions (for agents and/or artifacts) and should respect the criteria for generative potential described above.

3 Model description

We introduce a model where innovation happens in an "artificial world" [9], which is inhabited by agents which "produce" artifacts, which in turn can be used by other agents to build their own artifacts, etc. An agent can produce several artifacts for different agents (and it can sell one type of artifact to several different customers).

The meaning of artifacts is just what agents do with them, while the role of agents is defined by which artifacts they produce, with whom, and for whom. The role of agents is also partly defined by the social network they are embedded into. In order to describe these features, and to have them generated in the model, without being enforced *a priori*, we envisage a network of agents that are producers of artifacts, which are in turn used by other agents.

So the "strong ties" between two agents are mediated by a chain of artifacts: for example, agent A1 produces artifact O1 which is used e.g. by agent A2 to produce artifacts O2 and O3, and by agent A3 to produce artifact O4. O2, O3 and O4 are used by other agents, etc. If we describe the agent network, then there is a strong (directed) link from A to B if A produces an artifact which is used by B: the network of agents may well have loops, and an agent may have multiple outgoing and multiple incoming links.

There are also weak ties between two agents ("acquaintances") which refer to the fact that agent A knows something about agent B (e.g. its products, or part of its providers): so the social neighbours of a given agent are a superset of its providers and customers.

A key point is the structure of artifact space. For modelling purposes, we have considered different alternatives: binary coding as it is done in classifier systems, or λ-calculus as in the Alchemy model [10, 11] or numbers (either natural or real). What is required is that the space has an algebraic structure, and that suitable constructors can be defined to build new artifacts by combining existing ones. We concentrated on the real number representation and the use of mathematical operators because it is more compact than the binary representation and simpler than the predicate calculus and the λ-calculus. A full comparative evaluation of the features of the different representations lies beyond the purpose of this paper.

Therefore the agents are "producers" of numbers by combination of other numbers. The mathematical operators which are available to an agent represent its "technology", so a major technology shift might be described by the appearance of new operators, like e.g. the discovery of division in a world were addition and subtraction were the only available technologies.

It is well known that technological constraints limit the set of possible artifacts, a feature which is sometimes described by defining a fitness landscape on the artifact space; however, in our model the value of an artifact must be related to what other agents do, so assigning *a priori* a given fitness would not be a suitable alternative.

A more precise description of the model is given below.

3.1 Entities description

Artifatcs

The artifact Ar_i ($i \in [1,N]$) is described by:

- the identifier $i \in \mathbf{N}$, which indicates in a unique manner the artifact itself
- the integer value $name \in \mathbf{N}$, which identifies the artifact type and is the main value manipulated by the agent recipes
- the identifier of the agent that produced it
- the level of its stock
- the identifiers of the agents that acquire it.

We foresee the possibility to develop in the future a richer representation of the agent type, by resorting to a vector of integers (or real numbers) instead of a single scalar quantity

Agents

The agent Ag_j ($j \in [1,M]$) is described by:

- the identifier $j \in \mathbf{N}$, which indicates in a unique manner the agent itself

- the number of recipes nrecipes$\in \mathbf{N}$ owned by the agent
- the number of goals ngoal$\in \mathbf{N}$ and the list listgoal$\in \mathbf{N}^{ngoal}$, "owned" by the agent
- the "catalogue", i.e. a list of existing recipes, available to the agent
- the list of the agent acquaintances
- the list of the known artifacts
- the real vector vote$\in \mathbf{R}^{M}$, which measures the quality that the agent attributes to each relation (with other agents) it maintains
- the real value activity$\in \mathbf{R}$, which measures the "wealth" of the agent
- a set of parameters that control the agent propensity to innovate, to enlarge its own acquaintances, to participate to joint projects, and so on (in the future, we will sometimes refer to this set of parameters as the agent "style").

Recipes

The third object we define is the production recipe. The k-th recipe of the agent Ag_j, Rc_{jk} ($k \in [1, Ag_j.nrecipes]$) is described by:

- the number ninput$\in \mathbf{N}$ of needed inputs
- the ordered list of identifiers of the input artifacts
- the ordered list of operators, whose number is equal to ninput-1
- the identifier of the product
- the integer value prode$\in \mathbf{N}$, which states the number of tokens created at each recipe production
- the real value coste$\in \mathbf{R}$, which measures the cost required to produce the product associated to that recipe
- the time tlast $\in \mathbf{N}$ elapsed from the last time the recipe has been active, i.e. actually used for production.

It is useful to define a further entity, the recipe support, which is simply a subset of recipe features:

- the number ninput$\in \mathbf{N}$ of needed inputs;
- the ordered list of operators, whose number is equal to ninput-1.

This concept allows the separation between the "fundamental" structure of one recipe (i.e. the recipe support, which stores the construction modalities that are the basis of the recipe itself) and one of its instantiations (which leads from the particular input artifacts to the output one).

External environment

The model presented here is not necessarily a closed one, even if, as we will see, it can behave in a self-sustaining mode, without need of external inputs. However, in order to allow a wider generality of the model, we will assume that there may be an "external world", which represents other agents which are not included in the model (external customers) or artifacts which are not produced within the model (e.g. raw materials). We can call all these entities "environment".

More specifically, we identify:

- a set of "raw materials" (artifacts whose quantities are neither limited nor associated to some particular agent; they can be used as input to production recipes)
- a set of "external reward regions", areas where artifacts are acquired by entities not belonging to the simulated market system.

3.2 Model dynamics

We first provide an overall description of the model behaviour. Agents transform artifacts; in agent-artifact space they are defined by their production recipes and by their relationships with other agents. Each agent can produce more than one product, and for each product there is at least one production recipe. For simplicity we state that at each simulation step only one recipe is active for a given product. Agents have also "goals", that is, particular names that the agent aims to produce. Furthermore, agents are able to participate to joint ventures, to enlarge their acquaintances, to give birth to new agents.

A schematic sketch of the possible agents' actions is given here below.

At each time step (updating is asynchronous):

1) a randomly chosen agent is selected for updating
   a) for each recipe, the agent looks for the required inputs
      I) if the inputs are found (i.e. they have nonzero stock level), the recipes outcome is produced and inserted immediately into the stock, while the corresponding input stocks are decreased
      II) if the inputs aren't found, the agent searches, among the artifacts known to it, another object with the same name, whose stock level is different from zero
         (i) if this artifact is found, it substitutes the old one; return to step (1.a.I)
         (ii) if this artifact isn't found, pass to another recipe; in this case, a counter associated to the unused recipe is incremented (this counter will be used to cancel recipes which have not been used for several times)
   b) the following actions need not be performed at each time step. Depending upon the value of suitable parameters, and upon the level of its strength, the agent may perform one or more of the following (the decision about which ones is taken in a stochastic way):
      I) try to innovate by itself its catalogue, by:
         (i) choosing a goal
         (ii) trying to realise it, i.e. to find a recipe for the goal (or for a product close enough to the goal, see below section 3.2.1)
         (iii) producing it, if a successful recipe is found
      II) try to innovate its catalogue together another agent, by:
         (i) choosing a partner
         (ii) with the partner, choosing a goal
         (iii) with the partner, trying to realise it
         (iv) sharing the obtained results, if the innovation process succeeds
      III) try to enlarge its acquaintances
      IV) cancel the less useful recipes

   c) The agents' activations are updated.

Moreover, new agents can be created, according one of the following rules:

   d) randomly, with a given probability
   e) in order to substitute the dead agents
   f) according to the internal dynamics.

Now we present in detail the actions just sketched above.

3.2.1 Goals

In order to survive (i.e. to keep its strength above a given threshold) it is useful for an agent to find artifacts that could have high sales. Therefore identifying a proper goal (the choice of the artifact that has to be built) is crucial for the agents. When its turn comes, an agent which is reach enough (that is, whose strength is high enough) can try to create a new recipe. Let us consider "lonely" agents, which try to do this basically by themselves; the agent:

- looks at the portion of artifact space which it knows
- updates the set of its goals
- chooses a specific goal
- tries to come close to that goal by using the available operators on the available inputs and recipes
- if, using the available inputs, it comes close to the goal (within a given range) it puts the corresponding recipe in the set of active recipes.

Heuristics for goal setting include:

- locate a region of artifact space which is populated by several artifacts (and that therefore is likely to be able to sustain a high number of them)

- explore an existing reward region looking for higher reward points
- explore new regions of artifact space, e.g. by mutating an existing artifact, by interpolating between some of them, etc
- respond to the request of another agent
- examine another agent's recipes and find some input product which can become a goal for the first agent.

3.2.2 Recipes

Agents transform artifacts by means of their production recipes. Each agent can produce more than one product, and for each product there is one production recipe. A recipe is a function Rc_{ij}:In→Out, where In=$\{x_1,x_2,\ldots,x_{ninput}\}$ is the vector of the names of the input artifacts and Out is the name of the product.

The total transformation is composed by a set of simpler transformations, called operators; the presence (or absence) of particular kinds of operator defines the "technological level" achieved by the recipe. The operators are not limited to the basic numerical operators (addition, subtraction, multiplication, division), but more generally are any sort of functions Op:E⊆$\mathbf{R}^m$→$\mathbf{R}$, like for example exponential, mean, maximum, or function restrictions (where the operand domain is restricted to certain intervals, e.g. addition with names∈ [0,120], multiplication between artifacts with names ≤10, etc.). If we assume that Op=$\{op_1,op_2, \ldots , op_{ninput-1}\}$ is the list of the operators, the result of a recipe could be written as:

$$Out=((((((x_1.op_1.x_2).op_2.x_3).op_3.x_4).op_4.x_5)\ldots).op_{ninput-1}.x_{ninput})$$

Note that, in the present notation, the order of the application of operators is from op_1 to $op_{ninput-1}$. Changes of the order of the input artifacts, as well as changes of the operator's order, generally lead to changes in the recipe output. Of course, changes of the names of the input artifacts also lead to changes in the recipe output. All these alterations nevertheless could be done without modifications of the recipe support, which remains the fundamental structure of the recipe.

Agents are capable of changing the recipe repertoire by:
1) eliminating one recipe and the corresponding product (at no cost)
2) introducing a new recipe, by means of:
   a) the change of the providers (new recipe, same recipe support, same input names)
   b) the change of the order of the inputs (new recipe, same recipe support);
   c) the change of the order of the operators (new recipe, new recipe support)
   d) modification of the choice of operators and possibly of their number (different recipe, different recipe support).

It is possible to introduce new recipes with at least two different approaches:
- acting alone and using only one's own recipes
- initiating a collaboration with another agent and combining the knowledge of both partners.

A recipe may be active as long as it has input artifacts; of course, if one or more of its inputs are absent from the system, the recipe does not operate and its output is not produced. A recipe whose product doesn't have any client for a certain number of time steps is removed and the product disappears.

When a new recipe is first introduced, the corresponding agent inserts its product in its catalogue, making it known to its acquaintances. If at least one customer is found the recipe is activated, and remains active as long as it has at least one customer. The first time no customers are found the recipe does not act, although it remains in the accessible catalogue for some time t_{forg}. Recipes that have been constantly inactive for a certain number of steps are removed ("forgotten").

Recipes which are regarded as useless (e.g. the recipes that don't give a sensible contribution to the agent activity) or even damaging (those that give a negative contribution) may be removed by the owner, in a way

whose details (e.g. parameters) depend upon the agent "style". Moreover, an agent could choose to eliminate the less performing recipes to leave room for more performing ones.

Note also that, in the model described so far, as well as in the simulations shown in section 5, artifact space is "flat", i.e. uniform, although some regions may be unreachable at a certain time because of lack of proper inputs or operators. The model could however also deal in a straightforward way with intrinsic difficulties in reaching some regions of artifact space ("technical barriers"), by modulating the success probability of an innovation using some suitable external function F(x) (e.g. |sin x|): when an innovation is tried, the probability of success is proportional to F(x).

3.2.3 Joint developments of new products

In order to align and coordinate their actions, two or more agents, which have goals which are close in artifact space, can cooperate; if one agent alone is not able to reach the interesting region, the combined efforts of two agents could be successful.

A way to combine the efforts is that of sharing the agents' knowledge (input artifacts, recipes). The agents could create a new recipe (for example combining their recipes by means of genetic operators) which is able to reach the interesting region; after this action, they could divide the obtained recipe into two different recipes in such a way that the product of the first recipe is one of the input artifacts of the second recipe. Finally the first agent becomes the owner of the first recipe and the second agent becomes the owner of the second recipe (one recipe for each agent). In such a way they are able to share the success of the research results.

3.2.4 Artifacts stock level

Artifacts are characterised, among other variables, also by their stock levels, that is the "quantity" of the available tokens. When a recipe uses one artifact the corresponding level decreases, whereas when the appropriate recipe produces the artifact, the corresponding level increases.

Generally speaking, a recipe combines a certain amount of each input product to generate a certain amount of output. A precise quantitative description of these aspects would lead to further complications. Furthermore, we expect that this complication is related more to the specific aspects of the production chain than to the model aims, i.e. the study of innovation processes. On the other hand, we feel it necessary to maintain the fundamental distinction between types (i.e. names of the artifacts) and tokens (i.e. quantities). This is an example of how the theory constrains the model! In order to meet this need while keeping the complications at a minimum, we consider recipes where the quantity of each input is one, while the quantity of the output product is a property of the recipe itself (described by the variable "*prod*").

If the number of produced tokens is equal to the number of inputs, the total number of tokens is conserved; otherwise, this number is allowed to change. Therefore, if the value of this variable is greater than the number of inputs the total number of artifacts (tokens, not types) will tend to increase in time.

The agents control the stock levels: their aim is to avoid stock depletion and to avoid a useless surplus. There are several methods to obtain these aims:

- to allow production only when the stock level is equal to zero
- to allow production only when the stock level is lower than a given value (e.g. *prod*)
- to allow production depending on the artifact sales (e.g. by comparing sales trend with stock levels).

3.2.5 Activation dynamics and agents' wealth

A real variable, called activation, is associated to each agent, and measures its "wealth". At each time step, the agent activation is updated; it receives positive contributions from those products that are sold, while it has to "pay" for all its recipes. Whenever an agent performs a change (e.g. by introducing a new product or changing a recipe) it has to pay some activation in order to be allowed to try. A stochastic mechanism then

determines whether the attempted modification happens or not: this simulates the effort that can be put in R&D to discover new recipes.

The formula that describes activation dynamics is:

$$Ag_j.act(t+1) = Ag_j.act(t) + \sum_{k \in sold.products} n_{jk}(t) - Ag_j.act(t) \cdot \sum_{i \in active.recipes} Rc_{ji}.prod \cdot Rc_{ji}.\cos t - \lambda Ag_j.act(t)$$

[Eq. 1]

where $n_{jk}(t)$ indicates the number of sales of product k (belonging to agent j) at time t. The final term is necessary to make the activation of an agent which never acts vanish.

Of course, different formulae for the activation dynamics can be tested; simulation results will provide indications about the better ones.

Agents whose activation is equal to or smaller than zero are removed from the system. New agents could be generated, according to rules whose definition depends upon the aims of the simulations. For example, some suitable rules are:

- to substitute the removed agents (in order to maintain the agents' number)
- to introduce new agents at each step according to a given probability distribution (in order to simulate the appearance of new agents coming from outside the system)
- to introduce new agents according to internal dynamics, for example depending upon the total system wealth, or the wealth of a particular agent that can generate some spin-off.

3.2.6 Relationships among agents

Agents can use their knowledge about agent-artifact space, which is in general limited (an agent may not know all the existing artifacts, all the existing agents, all the goals and the recipes of the known agents).

An agent can also generate new goals or remove some old ones, with the aim to increase its activation. The definition of a new goal can be based upon the knowledge of the other agents' requirements and of the existing products, and upon the requirements of the external world.

An agent is also endowed with a "social network". There are two different kinds of interactions among agents: i) relationships with providers and customers clients, mediated by artifacts, and ii) relationships with the agent's acquaintances ("weak ties"), in which some exchanges of information can take place.

Each agent has its own measure of the quality of its relationship with other agents, which increases with the frequency of the positive events that take place during their interactions (purchases, sales, exchange of information, etc.), while it decreases with the frequency of negative events (wrong information, mistakes, poor activity income, failure in joint developments, etc.).

A reciprocal high vote could allow two agents to:

- exchange information (acquaintances, goals, operators, ...)
- share (part of) their recipes
- create alliances in order to reach interesting zones of agent-artifact space.

Agents can invest part of their activity, in order to:

- enlarge their acquaintances
- look for new providers
- improve their knowledge of some agent with whom they hold a high-ranked relationship
- increase the number of artifacts they know.

To improve their relationships, an agent can inform other agents about (some of) its desired inputs. If the interaction has been successful, the agent lets its providers "know more" about itself - e.g. other desired

inputs, or its production recipe(s). Moreover, reciprocity leads providers to inform their customers about themselves.

4 The prototype simulator

In order to test the behaviour of the model and the suitability of some mechanisms that have been proposed, we developed a preliminary, simplified version of the simulation environment, which has three main modules:

- a specialised initial condition builder, whose aim is to create the initial network of agents and artifacts
- the simulation engine
- a tool to visualise the main simulation results.

The interaction among the modules allows one to simulate a wide set of situations and conditions. The environment has been developed in C, while the visualization module has been developed by means of scripts that use the R environment. The simulations have been performed on personal computers and on clusters of PCs.

The model outlined above is rather complicated and the presently running software adopts some suitable simplifications, in order to test the appropriateness of some mechanisms that have been proposed. The main differences between the model described in section 3 and the current version of the simulation environment are the following:

- updating is synchronous
- the only arithmetic operators are addition and subtraction, without interval constraints
- no vote is given to the relationships with other agents
- production instantaneously adjusts to the total amount of requests (stocks are always adequate)
- at each time step all the agents have a certain probability to create a new artifact (no more than one per agent)
- new recipes can be generated by modifying the input of an existing recipe (without changing the recipe supports) or by applying an evolutionary mechanism to the existing recipes (a genetic algorithm, which allows changes in the recipe supports)
- no new agents are generated.

4.1 Generation of new goals and new recipes

In order to introduce new products, agents need to have new goals and to develop the recipes to produce them. Two different methods for devising new goals have been tested: “imitation” and “imitation & jump”. We will show in section 5 that only the latter is able to sustain continuous innovation.

In “imitation”, an agent is allowed to choose a goal identical to one of the existing products. Actually, the goal is chosen with uniform probability among the known artifacts. In such a way a simple sampling of artifact space is performed: the higher the number of artifacts which lie in a certain region of artifact space, the higher the probability of finding that region. The underlying heuristics is that clusters of artifacts are probably high reward regions. The agents try to build an artifact similar (within a given threshold) to the selected one (which represents the “goal” of the imitation): if they succeed, the new artifact is added to the set of those that are already present.

Imitation alone is unlikely to enhance the agents’ activities, since the new artifacts have no chance to be purchased. Indeed, the name of the artifact created by imitation can be either the same of the target one (and, in this case, the agents that already utilise this name don’t have any reason to change their previous suppliers) or different from the name of the target (and, in this case, no agent is able to utilise this new kind of artifact). The only possibility for a new artifact to survive is that another agent creates a new recipe, and chooses the artifact itself to build its own product. This event has a very low probability, and in effect with only imitation the initial agents’ activities drop to zero. A way to avoid this fate is that of allowing the agents to change their suppliers in a random way. If this may happen, many agents typically survive (as shown in section 5 below).

Although in this latter case some agents survive, it is observed that the production of new artifacts ends after a finite number of time steps. The problem is indeed related to the definition of new goals: as long as agents try to imitate already existing products, they can't "imagine" new things. It is therefore appropriate to allow agents to change the way in which the goal is chosen, by

- multiplying the name of the artifact, selected by imitation, times a randomly selected value (belonging to a finite interval [min,max])
- adding or subtracting a random value.

In such a way the jump allows the exploration of new regions of artifact space and, as it will be shown in the next section, the combination of imitation and jump allows a continuous generation of new artifacts.

Once a goal has been set, an agent needs to develop the corresponding recipe. Two methods have been devised: exhaustive search and genetics.

In exhaustive search, an agent looks at each of its own recipes and, for each input, it substitutes it with another artifact. The exploration concerns all the artifacts known to the agent, although in the present version not all the combinations of known artifacts are tried: indeed, substitutions in existing recipes are done one at a time. The search halts when the goal has been reached; if this never happens, the closest product is considered and, if its distance from the initial goal is smaller than a threshold, the corresponding recipe is added to the set.

While exhaustive search deals only with the input artifacts, in genetic search also the recipe supports are combined to give rise to new recipes. In particular, the method starts by generating an initial population from a set of recipes (which can either belong to a single agent in the "lonely" mode previously described, or to two collaborating agents). In order to have an initial population, the existing recipes are augmented with some modified versions. These modifications are obtained by changing the input (as in the previous method) with some probability.

The evolution of the population proceeds according to the well known genetic algorithm. In particular, the fitness function of each individual in the population is a decreasing function of its distance from the goal. The main genetic operators are uniform crossover (which has been modified in order to take into account the existence of recipes of different lengths) and mutation (note that this term is used here with a meaning which is very different from that used in previous section). Mutation involves here changing one input artifact with the closest one among those known to the agent.

4.2 Initial conditions

Setting up initial conditions "by hand" is time demanding, so we realised an automatic "initial condition builder". First of all, raw materials are introduced; afterwards, agents are added one at a time. The recipes of the newly added agent can use the already existing artifacts, producing in such a way new artifacts. Networks of this kind are such that older artifacts are likely to be connected to more agents than new one.

The networks generated according to the above mentioned procedure have the common feature that the unsold artifacts (belonging to the last added agents) can't provide any source of activation to their owners, and therefore the latter cant' survive, causing a "domino effect" that destroys the entire net. To avoid this collapse we can either provide an external reward region or introduce a modification of the dynamics (e.g. the possibility of changing a provider) which leads to self-sustaining rings, as shown in the next section.

5 Simulation results

Several simulations have been performed, in order to test the model and to understand some if its behaviours. We do not present statistical results, but rather typical model behaviours. The simulations that are shown below are organized in an order that reflects the introduction of an increasing number of features. In particular, the sequence is the following:

- agents operate without innovation

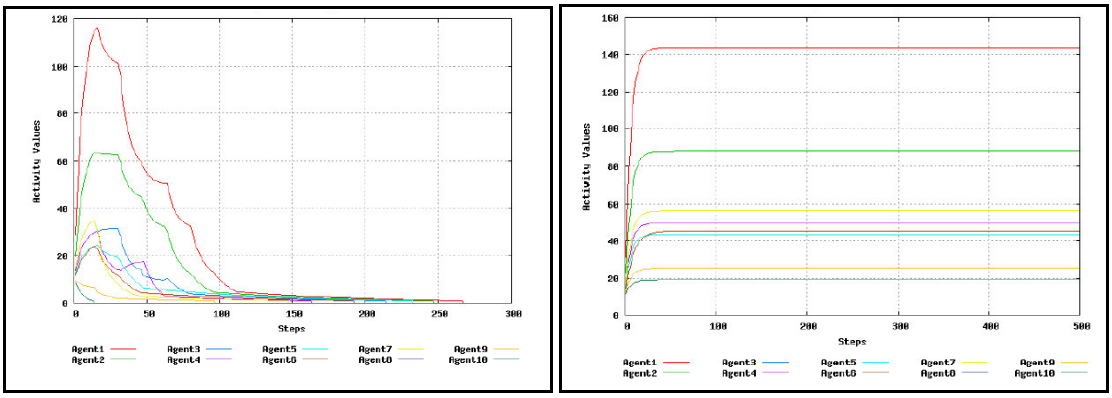
- agents perform innovations, either in suppliers or in goals and recipes, each agent knowing all the existing artifacts
- agents perform innovations, either in suppliers or in goals and recipes, each agent knowing only the artifacts produced by a limited number of other agents.

Different innovation methods have been tested. In particular, when innovation in recipes is performed by applying genetic operators to the set of existing recipes, we have considered both the case of "lonely" agents (which modify only their own recipes) and the case of "collaborative" agents, which share their recipes in genetic search of new ones.

5.1 Activation dynamics without changes

In this paragraph we show the behaviour of the model when agents do not perform any kind of innovation and never change their providers.

As already stressed, the actual initial condition builder generates networks that are never able to survive without external reward. Figure 1a shows the decrease of activities of a net with 10 agents and no self-sustaining loops or reward regions, whereas Figure 1b shows the stabilising action of a region of reward (the initial situation is the same in both simulations). During these runs the agents can't innovate and/or change providers.



(a) (b)

Figure 1: activities vs time in a net with 10 agents. (a) no self-sustaining loops or reward regions, (b) the stabilising action of reward regions.

5.2 Innovation with complete knowledge of existing artifacts

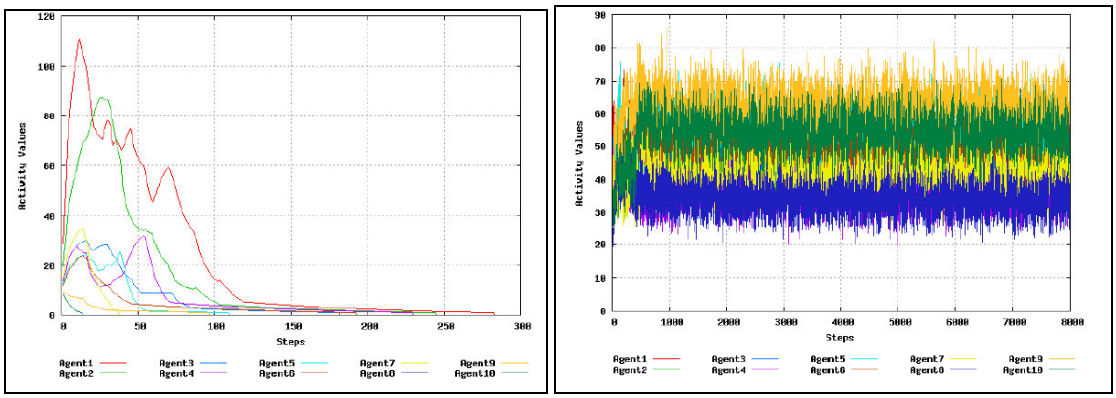
In this section we will consider agents which make use of different innovation methods (new goals, new recipes), and which can also modify their suppliers. The distinguishing feature of the different cases described here is that agents have a complete knowledge of all the artifacts produced by the other agents. In most cases we will show the behaviour of lonely agents which operate only on their own set of recipes, but we will also show the results of allowing "knowledge sharing" between pairs of agents.

5.2.1 Imitation and change of suppliers

In this section we will consider the effects of innovations (of different kinds) which agents are allowed to perform. Let us first consider imitation only. As it has been already anticipated, using imitation without change of suppliers eventually leads to the vanishing of agent activities, no matter which method is used to generate new recipes (see Figure 2a). A way to avoid this fate is that of allowing the agents to change their suppliers in a random way. If this may happen, many or all the agents typically survive (as shown in Figure 2b).

Figure 2b also shows that the system is self-maintaining. In absence of reward regions there are two possible explanations of this kind of behaviour:

1. there are self-sustaining loops;
2. the agents innovate at such a high frequency that they have always sell some artifact.



(a) (b)

Figure 2: agents' activities (no self-sustaining loops nor reward regions): (a) with imitation and without change of suppliers (the new artifacts are not purchased, and the agents die), (b) with change of suppliers the new artifacts have a chance to be purchased, and the network is able to self-maintain.

In effect, given the particular way to build the initial condition, small changes in the network structure (i.e. the change of providers) can create self-sustaining loops. It is possible to test this hypothesis, by deactivating the innovation mechanisms (imitation and change of suppliers) and therefore making the rate of innovations vanish. Figure 3 shows that after such deactivation a situation of stability is achieved, demonstrating in such a way the presence of self-sustaining loops. The removal of agent 1 at step 2000 changes the levels of stability but is not able to make the whole system collapse; further analysis could determine whether there is more than one loop or whether agent 1 is located downstream with respect to a loop.

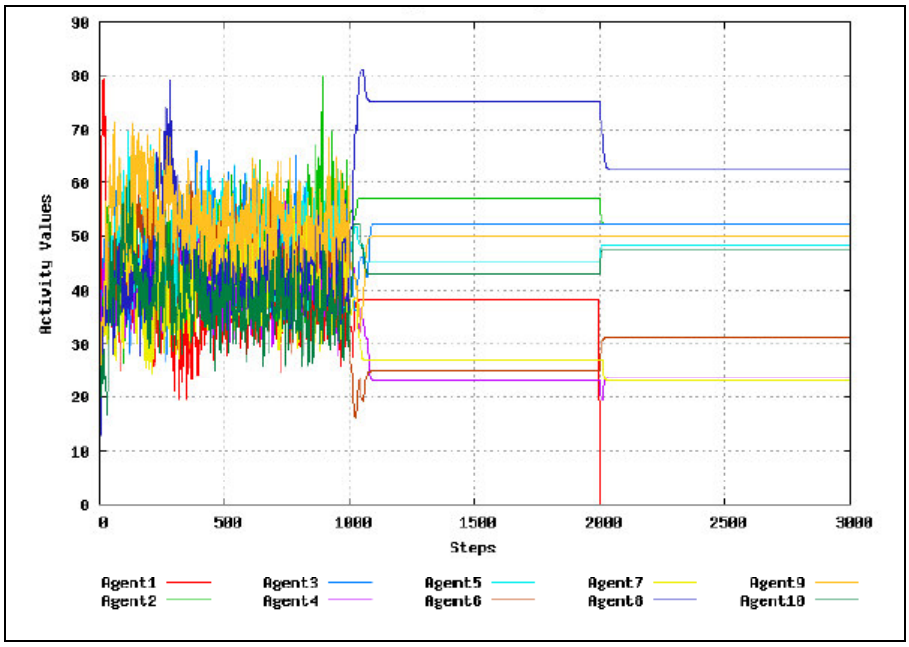
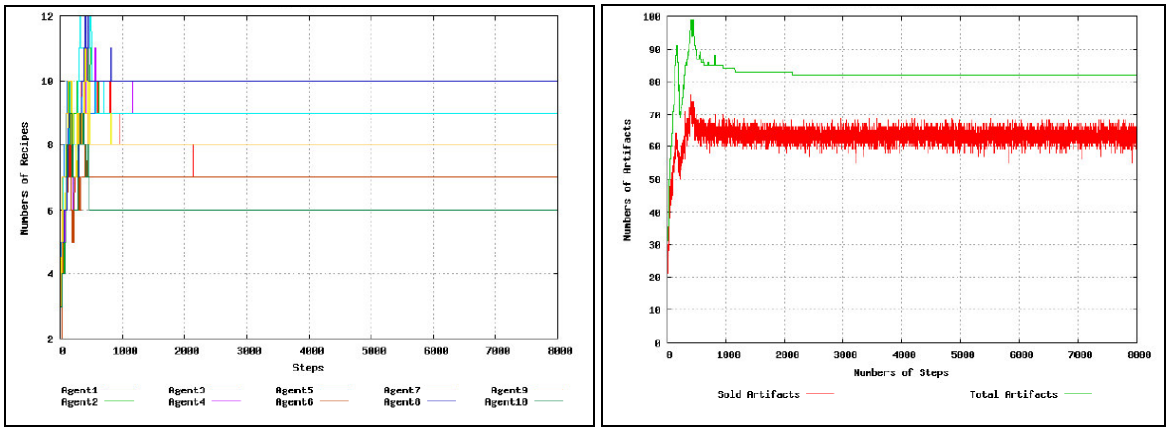


Figure 3: behaviour of the agents' activities under some external changes; step 0-999: imitation and change of suppliers are active; step 1000-3000: the rate of innovations is set equal to zero; step 2000: agent 1 is removed from the simulation.

Imitation and change of suppliers show a particular limitation, which is evident in Figure 4a: after a while, the number of recipes owned by each agent stops growing. This phenomenon indicates that imitation and

change of suppliers don't allow continuous innovation; therefore imitation is not a true "novelty generator", and in a few steps all the possible combinations are generated (Figure 4b).



(a) (b)

Figure 4: simulation with imitation and random change of suppliers: (a) number of recipes owned by each agent (b) total number of artifacts and number of sold artifacts; the distance between the two lines is due to the innovations which are not purchased by the agents.

5.2.2 Jump as a continuous source of novelties

Note that the fact that innovations come to a halt, in the previous simulations, should not be attributed to a limitation in the methods used to generate new recipes but rather to the limitations in the available mechanism to devise new goals. Indeed, new recipes can be generated (see section 3) either by modifying the input of an existing recipe or by applying an evolutionary mechanism (e.g. a genetic algorithm) to the recipes owned by the agent. The simulations presented above made use of the former mechanism; however the results are similar in runs with the second mechanism turned on. The creation of new recipes is not upper limited, and all the numbers $n \in \mathbf{N}^+$ are potentially reachable, but after an initial success we saw that no new artifacts are produced.

Let us then allow agents to "modify" the chosen goal; so now an agent multiplies, with fixed probability, the name of the artifact, selected by imitation, times a randomly selected value (belonging to a finite interval [0.7, 6]). In order to appreciate the influence of this mechanism, it is convenient to compare the behaviour of two simulations, starting from the same initial conditions and differing only in the probability of performing the jump (0 in the reference case and 0.3, respectively). In both simulations the innovation process is stopped at step 3000, in order to check the existence at this time of self-sustaining loops (Figure 5).

The jump introduces very strong effects: the number of recipes owned by each agent during the second simulation is five times greater than that of the first simulation, and the number of artifacts present at step 3000 is larger by a factor of four. Even more interesting, the diameter of artifacts (the distance between the names of the most distant artifacts) and the number of different artifacts' names (the number of different kinds of artifacts present at each step of the simulation) are continuously growing, at least during the simulation time span. The network structure is continuously changing.

Note that the use of the word "jump" here refers to a change in the value of the variable "goal" which is not necessarily a small one. In effect, it has been observed (and could be theoretically justified) that the effects of jumps become sensible when their size exceeds the threshold which defines the maximum distance between a given goal and the closest product which can be actually built (i.e., when innovation is tried, a new product which can be produced using the existing recipes and inputs is actually produced only if its distance from the initial goal is smaller than the threshold). On the other hand, a too large jump coefficient leads to a basically random exploration of a too vast artifact space, which is ineffective.

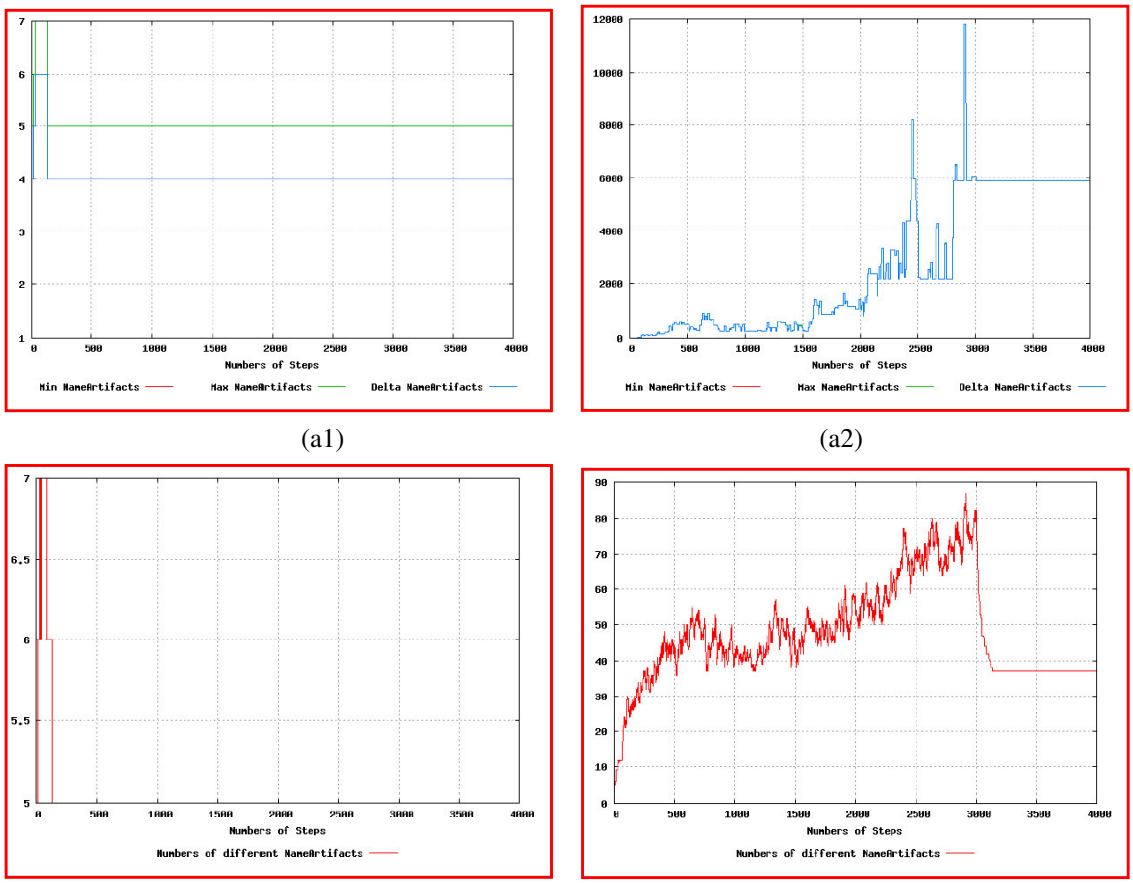


Figure 5: comparison between the behaviours with and without the goal jump; the first column is relative to the simulation without jump, and the second column is relative to the simulation with jump. (row a): diameter of artifacts vs. time; (row b): number of different artifacts' names vs. time;

5.2.3 Response to perturbations

A powerful way to explore the global properties of a system is that of introducing a slight perturbation and following its subsequent dynamics, which can be compared to that of an unperturbed system. For example, we can consider questions like: What is the system response to perturbations? Is the system robust with respect to changes in the structure of the network? What is the typical system behaviour when one or more of its components (agents, artifacts) disappear? In our case we could:

- remove a recipe from an agent;
- remove an entire agent;
- remove a group of agents at the same moment;
- remove agents at randomly selected moments.

Different behaviours to such perturbations are possible: Figure 6 shows some examples. Agents 1, 3, 5, 7, 9 disappear respectively at step 50, 200, 500, 1000 and 1300. There are several consequences, depending upon the particular agent, the time of fault, the number of agents surviving at the moment of disappearance, the current network structure, and so on. Without describing the details, it is interesting to observe that some perturbations have important consequences (elimination of agent 5 and 9), whereas other perturbations have not. The perturbations affect mainly the number of recipes owned by each agent, the number of different kinds of artifacts and the number of artifacts present at each step of the simulation. Also the diameter of artifacts is deeply influenced. These changes modify in an irreversible way the system structure (agents,

artifacts). However, the global properties described in Figure 6 display a strong resilience, i.e. the system recovers after crashes.

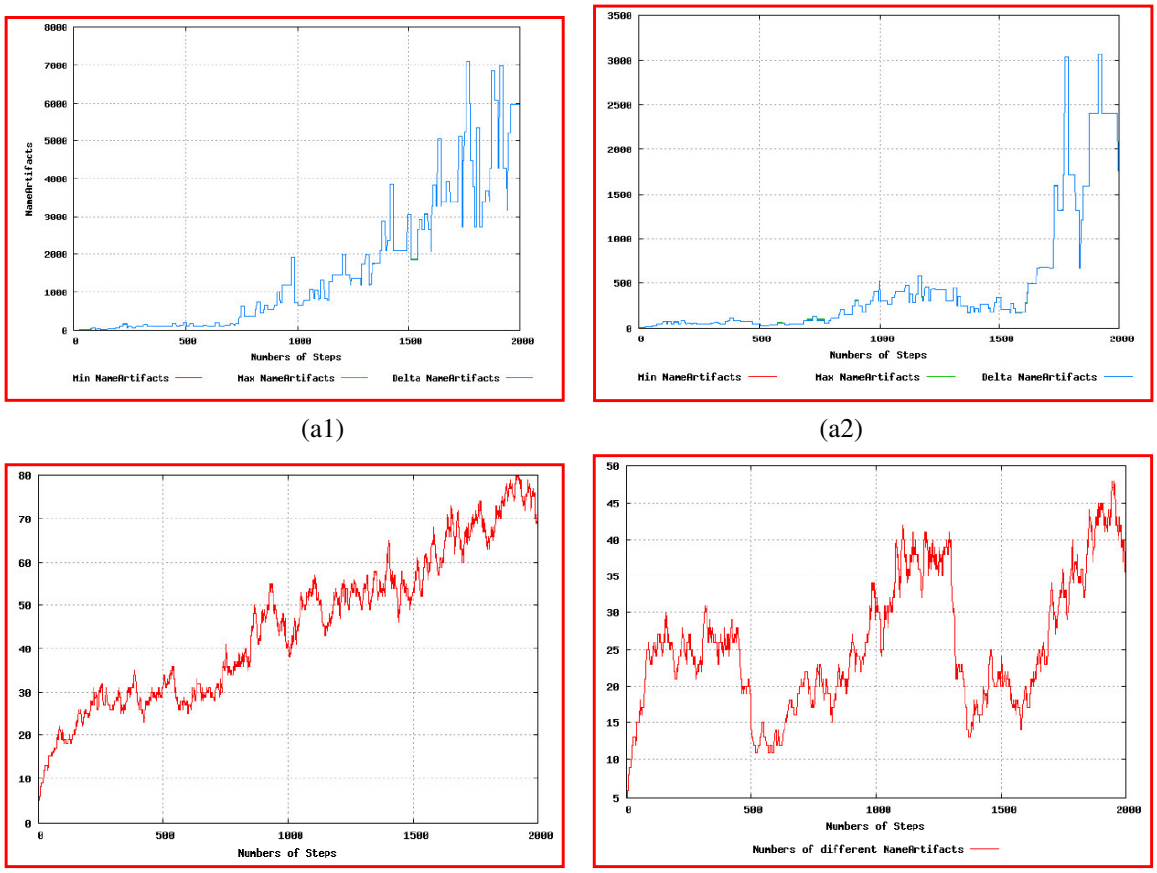


Figure 6: comparison between the behaviours with and without the externally enforced elimination of some agents (agents 1, 3, 5, 7, 9 disappear, respectively at step 50, 200, 500, 1000 and 1300). The first column is relative to the simulation without eliminations, and the second column is relative to the simulation with eliminations. (row a): diameter of artifacts vs. time; (row b): number of different artifacts' names vs. time.

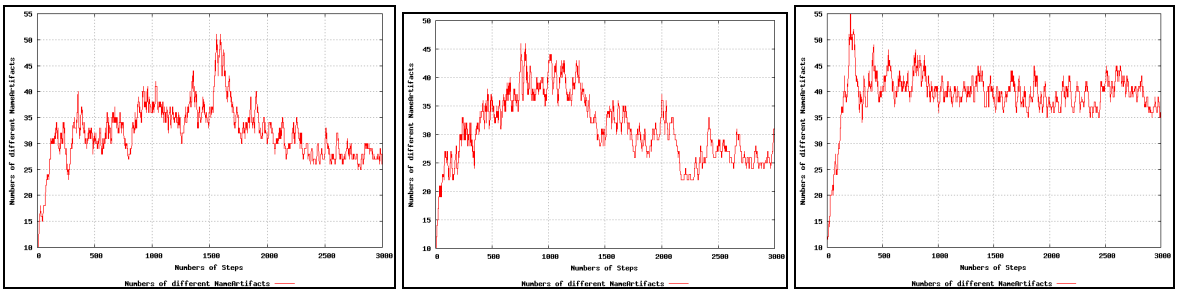
5.2.4 Using genetic search

Three different strategies have been used in these simulations

a) an agent performs genetic search using only its own set of recipes
b) an agent tries to reach its goals using only its own recipes; if it fails, it chooses another agent with which genetic is performed. If the search is successful, both partners own a copy of the successful recipe
c) agents always perform genetic search in pairs.

When joint search is performed, the chosen partner is the best supplier, i.e. the one which is selling most to the agent

Some results of the simulation are shown in Figure 7: Note that strategies a and b give very similar results, probably because, if each agent knows all the artifacts, the need for joint search is small and the two strategies behave in much the same way. An interesting observation, however, is that if joint search is prescribed, the number of different types of artifacts which are generated is higher than in the previous cases, as is also the number of different recipes per agent. Moreover, using the first two methods one observes a decline in the number of different artifacts, which instead seems to remain fairly constant in the third case.



(a) (b) (c)

Figure 7: number of different artifact types; (a), (b) (c) are described in the text

5.3 Innovation with limited knowledge of existing artifacts

While complete knowledge of the "artifacts" is probably useful for an agent, it represents an unrealistic assumption in many real world situations. It is therefore very interesting to explore the behaviour of a system whose agents have incomplete knowledge of the products of other agents. In particular, aspects related to the interaction between agents are expected to be better described in such a framework.

In this case, each agent knows the artifacts produced by some other agents (i.e. its own "network"). The properties of this network should play a major role in the system behaviour, as we will show in the following sections.

We will consider first the case where an agent knows only its suppliers, and we will then analyze the case where the network of acquaintances is enlarged.

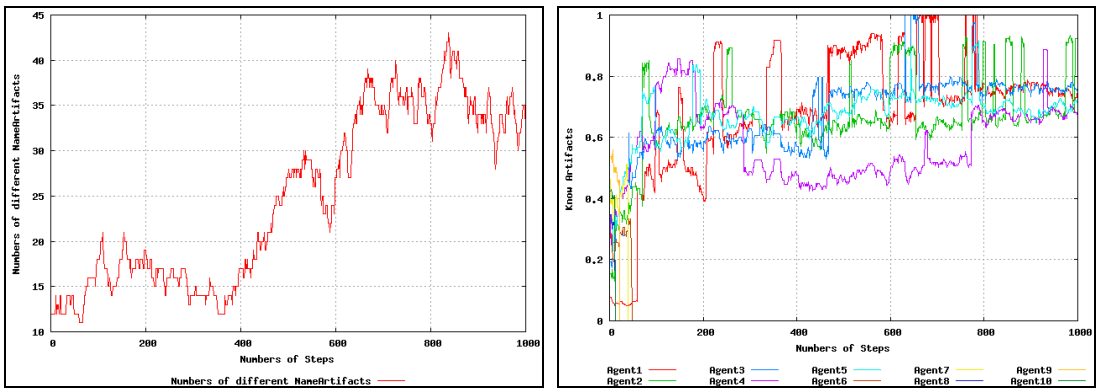
5.3.1 Knowing only one's suppliers

We start with the simplest case where an agent knows only the products of its current suppliers; the agent tries to innovate by

- defining its goal through imitation & jump (where imitation is performed among the known artifacts)
- finding new recipes by genetics, according to strategy b (i.e. it first tries alone, if it fails it cooperates with its best supplier).

It is interesting to note that, if we had used strategy a (i.e. lonely genetics), the system would have died out, while with strategy b it survives. So in this case the two strategies lead to markedly different results, which is likely to be due mainly to the enlargement of an agent's network which is possible only in the latter case.

Note also that the diversity (i.e. number of different artifact types) grows in time and that the network of each agent grows, without however reaching complete knowledge. This is shown in Figure 8b where the fraction of artifacts known to an agent (i.e. the ratio known artifacts / total artifacts) is plotted vs. time. It is particularly interesting to observe that, due to the continuous change which is taking place, even an agent which, at a certain time, knows all the artifacts, is bound to loose this property at subsequent time steps.



(a) (b)

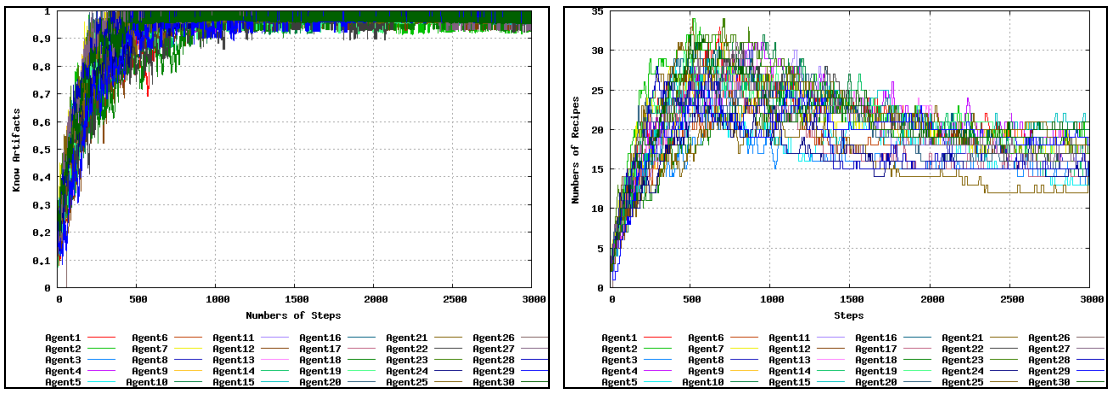
Figure 8: genetic strategy b: knowledge of only one's suppliers: (a) number of different artifacts' names vs. time; (b) fraction of artifacts known to each agent vs. time

5.3.2 Knowing not only an agent's suppliers

A particularly important aspect is the social network of each agent, and the way in which it changes in time. While in section 2 a sophisticated way to handle this aspect has been described, based upon an evaluation of each potential partner, here we limit to a simple way to "go beyond" the knowledge of an agent's suppliers, which has been described above. We will consider the case where an agent knows all the products of its current suppliers and all the products of one or more other agents chosen at random, at each time step, among the set of existing agents (which will be referred to below as "acquaintances").

In order to disentangle this effect from that of enlarging the set of suppliers, due to the combination of recipes described in the previous section, we limit here genetics to the recipes of a single agent. Therefore the agent here is lonely as far as the generation of new recipes is concerned, but it has a wider social network in that it knows the products of some agents which are not among its suppliers.

We have tested three cases which differ in the number of acquaintances per time step; the system is self sustaining and the number of different artifact types grows and then slowly declines in all cases. The most interesting observation is that each agent soon comes to know all the artifacts, and this situation persists in time, even if at each time step a single new acquaintance is considered (Figure 9). This particular behaviour, which may appear surprising, is due to the fact that, in all the simulations above, there is no hard limit on the number of recipes an agent can simultaneously possess. The only mechanisms for eliminating a recipe are indeed i) lack of inputs or ii) lack of sales. What is observed in the previous simulations is that each agent has a large number of recipes, and in such a way the network of suppliers covers the whole set of agents.



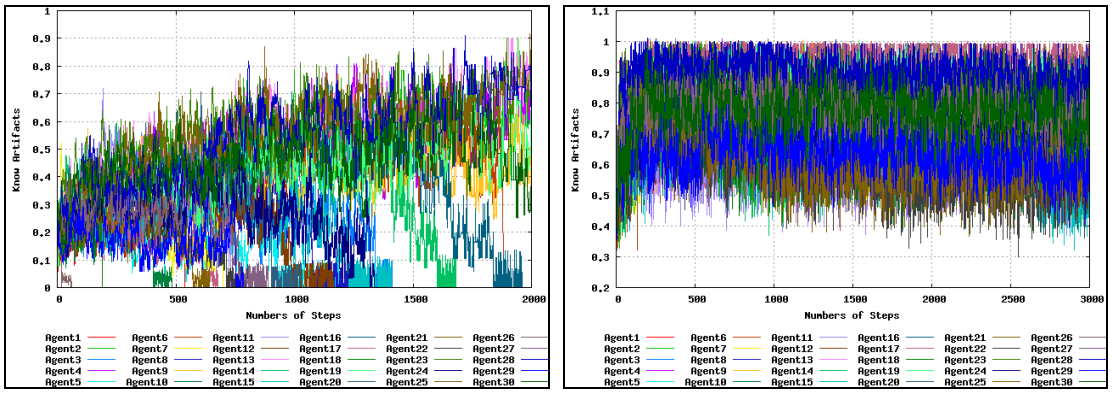
(a) (b)

Figure 9: genetic strategy a with acquaintances: (a) fraction of artifacts known to each agent vs. time; (b) number of recipes owned by each agent vs. time

In order to further investigate this aspect, we therefore introduced a novel mechanism for eliminating recipes: at each time step, there is a small but finite probability of eliminating the less performing recipe, defined as the one which has the lowest volume of sales in the last 15 steps.

Some very interesting phenomena have been observed in this case. The number of agents which are known to each agent increases with the number of acquaintances per time step N_a, as expected, but even in the case of a very high N_a (e.g. 10 new agents in a population of 30) a distribution of the number of known agents is observed: while there is an agent which knows all the other agents and therefore all the artifacts, there are some other surviving agents which know no more than 40% of the agents.

Similar results are obtained for the number of artifacts that are known to the different agents. In Figure 10 relative data are plotted: note that, although in the small N_a case the total number of artifatcts and agents declines, the relative number (known artifacts divided by total number of artifacts) actually increases, thus allowing the system long term survival. Indeed in this case the total number of agents declines, while with high N_a this number remains constant.



(a) (b)

Figure 10: genetic strategy a with acquaintances: f**raction of artifacts known to each agent**: **(a) one acquaintance at each step; (b) ten acquaintances at each step**

It is interesting to consider the number of recipes per agent that, as it may be expected, is a growing function of the number of acquaintances. Even in this case a wide distribution of the number of recipes per agent is observed: in the case N_a=1, the final (i.e. at t=2000) number ranges between 4 and 8 recipes, while if N_a=15 this number is comprised between 3 and 15. The analysis of the distribution of recipes (as well as the distribution of known agents and artifacts) is one of the most interesting future directions of research that can be pursued with this model.

Finally, let us consider the number of different artifact types, which is a growing function of N_a (Figure 11). Note that in the graph a maximum number of different artifacts is reached, which is followed by relaxation to a lower number which persists for a long time. This is an unexpected feature, which is found in many different simulations and which needs further investigation.

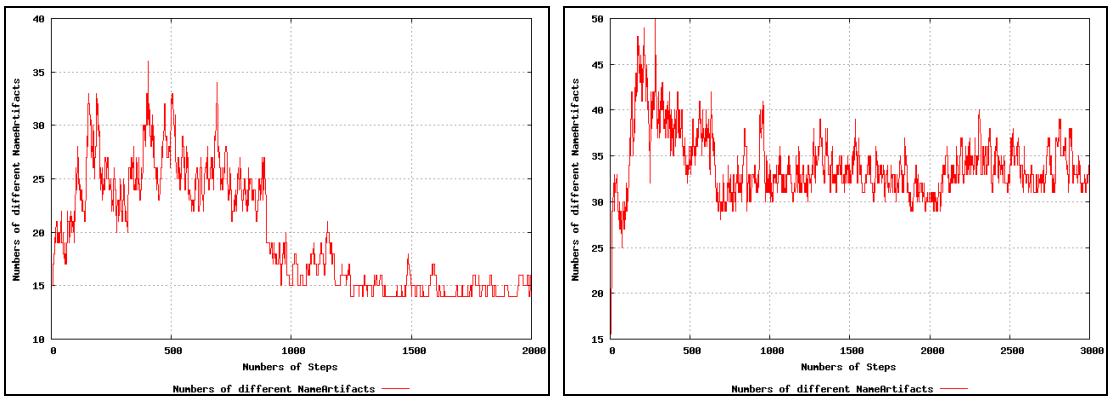


Figure 11: genetic strategy a with acquaintances: number of different artifact types: (a) one acquaintance at each step; (b) ten acquaintances at each step

6 Conclusions

The model under consideration belongs to the class of agent-based models, which are well suited to bridge the gap between a careful description of the behaviour of single agents (including their interactions with another or a few other agents) and the system-level features (which come out of the interaction of many such agents).

The relationship with the innovation theory of Lane, Maxfield and co-workers requires that intentionality plays a key role. And indeed in I_2M the agents' intentions are explicitly represented as goals, i.e. products which the agents want to produce and, for this purpose, they try to develop new recipes. This is a major distinguishing feature with respect to many agent-based models which rely upon chance as a generator of novelty.

Another distinguishing feature is the role of the agents' social interactions (both of the customer-supplier and of the "acquaintance" type). Although only preliminary simulations have been performed concerning this aspect, it is apparent from the results of section 5 that the outcomes are deeply affected by the parameters which describe how this social network evolves.

In the present version of the prototype a very simplified way of modifying the social network has been tested, based upon chance, while one of the major future directions of research would involve intentionality also in the change of the social interactions.

It is also interesting to observe that the tests performed so far have shown that the network of customers and suppliers is actually able to maintain itself, often forming closed loops, and that it is able to evolve giving rise to "perpetual" novelties. Therefore the modelling framework which has been envisaged seems well suited for the exploration of the main goals of innovation modelling in Iscom.

A major line for further development concerns a treatment of agents social interactions more closely related to the LM theory. In particular, in a system like ours it will be possible to measure the generative potential of a binary relationship, and to compare the behaviour of agents with different strategies. Some agents will follow the heuristics of favouring relationships with high generative potential (as suggested by Lane and Maxfield), while others will use different heuristics like e.g. "imitate the most successful agents you know" or "cooperate with your best customer". In this way it will be possible to explore under which conditions one heuristics outperforms the others.

The measure of the generative potential may be the product of three normalized quantities, each of which can be determined in our model

- mutual directedness, which can be measured by the vote given to a relationship

- heterogeneity of the agents, which can be related to the differences between their sets of recipes
- aligned directedness, which is measured by the distance between the agents' goals

Moreover, in the future it will be useful to develop suitable tools in order to analyse directly the structure (i.e. characteristic path length, clustering coefficient, topology, etc) of the network which develops, connecting different agents and artifacts.
It will also be interesting to consider possible ways to simplify the model, at least in particular conditions. Indeed, in spite of our efforts to keep things as simple as possible, the intimate relationship with the LM theory led us to a rather complicated model. Although we do not subscribe the claim that a model with several parameters can describe almost everything [12, 13], we feel that considering slightly simplified versions of the model might provide useful insights about its behaviour.

Finally, we stress that further improvements are under way on the code in order to make it faster, thus allowing simulations which will involve very many agents, which will allow a better exploration of the emergent properties.

Acknowledgements

We wish to express our gratitude to Andrea Ginzburg, who discussed at length with us the features of this model, since the time of its first conception. We are also indebted to Margherita Russo and to many members of the group working on the Iscom project for many helpful discussions and suggestions. This work has been supported by UE-FET initiative under grant IST-2001-35505 Iscom: The Information Society as a Complex System.

References

[1] Epstein JM, Axtell R. : Growing Artificial Societies: Social Science from the Bottom Up. Cambridge (MA), MIT Press, 1996
[2] Gilbert N, Terna P: How to build and use agent-based models in social science, Mind & Society 2000; 1:57-72
[3] Cioffi-Revilla C**.:** Invariance and universality in social agent-based simulations, PNAS 2002; 99: 7314-7316
[4] Ormerod P, Rosewall B.: On the methodology of assessing agent-based evolutionary models in the social sciences, Second Brisbane workshop on complex adaptive systems 2002; www.paulormerod.com/brisbane3PO
[5] Gilbert, N : Agent-based social simulation: Dealing with complexity; 2005: http://www.complexityscience.org/NoE/ABSS-dealing%20with%20complexity-1-1.pdf
[6] Axelrod R, Tesfatsion, L.: A guide for newcomers to agent-based modeling in the social sciences. In Kenneth L. Judd, KL, Tesfatsion, L (editors): Handbook of Computational Economics, Vol.2: Agent-Based Computational Economics. Amsterdam, North-Holland, 2005 (forthcoming)
[7] Maxfield R, Lane D : Foresight, complexity and strategy. In Arthur WB, Durlauf S, Lane D (eds): Economy as an evolving complex system II (chap. 4). Redwood city (CA), Addison Wesley, 1997
[8] Lane D, Maxfield R. : Ontological uncertainty and innovation, Journal of Evolutionary Economics 2005, 15: 3-50
[9] Lane D: Articial worlds and economics (I and II), Journal of Evolutionary Economics 1993, 3: 88-107, 177-197
[10] Fontana W: Algorithmic chemistry, in Langton CG et al (eds): Artificial Life II. Redwood City (CA), Addison Wesley, 1992
[11] FontanaW, Buss LW: The arrival of the fittest: towards a theory of biological organization, Bull. Mathem. Biol. 1994, 56: 1-64
[12] Di Gregorio S, Serra R, Villani M.: Applying cellular automata to complex environmental problems: the simulation of the bioremediation of contaminated soils. Theoretical computer science 1999, 217**:**131-156
[13] Serra R, Villani, M, Colacci, A. (2001): Differential equations and cellular automata models of the growth of cell cultures and transformation foci. Complex Systems 2001, 13 : 347-380

Modeling Firm Skill Set Dynamics as a Complex System[1]

Edoardo Mollona (emollona@cs.unibo.it)
David Hales (dave@davidhales.com)
Department of Computer Science, University of Bologna, Italy.

Key words: Labour, Internal labour markets, Skill, Resource-based view of the firm, Agent-based modelling, Computer simulation,

Correspondence to: Edoardo Mollona, Department of Computer Science, Mura Anteo Zamboni, 7; 40127 Bologna, Italy. Telephone: +39 051 209 4883; Fax: +39 051 209.
Email: emollona@cs.unibo.it

[1] This work is partially supported by the EU within the 6th Framework Programme under contract 001907 (DELIS). We acknowledge the helpful comments received by Dr. Ferdinando Monte and three anonymous referees, errors and omissions remain our responsibility.

Abstract

The article explores emergence and survival of human resource management strategies and organisational types in a knowledge-based job market. The analysis considers a dynamic environment in which skill requirements change rapidly. We built an agent-based model to simulate a market where firms post job offers to fill vacancies and decide how to select and reward employees; employees, bearing skills, select firms comparing job offers. Taking an evolutionary approach, we explore how hiring strategies, which guarantee survival, emerge from interconnected variation, selection and retention processes. The simulation experiments suggest that, as the rate of change of the environment increases, long-term employment and firm-specific knowledge building emerge as the survival strategy.

1. Introduction: Organisational Issues in a Knowledge-Based Economy

A prevalent claim is that we are in a *knowledge economy*. In this work, we take the view that what characterises a knowledge economy is the growing importance of human capital in productive processes [Foss, 2005: 8] and the increasing knowledge intensity of jobs [1]. In addition, an increasingly influential argument is that the division of labour is becoming complex and firms can be viewed as networks of knowledge nodes [2], that is, sets of interacting individuals with key skills and competencies. Such networks crystallises firm-specific knowledge and provide ground upon which firms build their heterogeneity. The fact that the knowledge content of jobs increases raises questions concerning emerging organisational forms.

Hodgson [1], for example, suggests that the lack of managerial control on knowledge-based jobs, especially when knowledge is tacit and cannot be codified, impairs and bounds the appliance of traditional employment contracts [1:193]. Hodgson proposes that the nature of contracts evolves along with the evolution of the distribution of bargaining power. As a matter of fact, employers maintain a *de iure* ownership of produced goods or services and of the physical means of production but these latter have a decreasing impact in a firm's value-creation processes [1: 194]. On the other hand, employees have got ownership on knowledge-based means of production and have and increasing control on production processes [1: 208].

Yet, firms maintain ownership on the mechanisms of knowledge accreditation, which increases rents extracted from knowledge-based jobs. For example, the brand Microsoft allows to extract rents from the jobs of many computer scientists and IBM brand allows extracting rents from the jobs of information system experts and consultants. Along similar lines, Porter Liebeskind [3] advises that firms have institutional capabilities that protect knowledge from expropriation and imitation thereby creating unique knowledge assets [3: 104].

Firm-specificity is a further characteristic of knowledge-based jobs that contributes to influence the evolution of employment relations. Learning processes are largely grounded upon exchange of tacit knowledge [4] [5] in groups of actors working together [6] [7]. Thus, knowledge-based jobs require workers to invest in firm-specific learning; in exchange, workers might want security and long-term employment [1: 248]. On the other hand, if by learning-by-doing processes, workers develop unique ways to perform tasks, the emergence of idiosyncratic jobs makes *internal labour markets* an efficient organisational mode [8].

Capitalising on the resource-based view of the firm [9] [10] [11] [12] [13] [15] [16], we assume that firms have incentives in integrating firm-specific, unique networks of scarce skills which fit the requirements of the competitive environments they are embedded into. We also assume that workers select jobs on the base of wage and security. That is, workers, aware of the idiosyncratic nature of their knowledge-jobs, prefer long-term contracts rather than short-term employment agreements.

We address how the evolving nature of jobs, namely the knowledge content of jobs, produces ambiguous and partially counterbalancing effects on hiring policies and emerging organisational forms. To explore the issue, we are inspired by scholars who have addressed how strategic resources management affects competence exploitation [16] [17] and by [16] that the dynamics of competence adaptation are strictly connected with the processes, for example, hiring or rewarding, that influence organisational demography.

The study presents results from experiments with an agent-based model, the FirmWorld model, which contains three kinds of agents: The Environment, Firms and Employees. The model captures the dynamic of employee skill sets, firm performance and organisational policies adaptation through the process of selective hiring, firing, firm creation and bankruptcy.

We specify a very simplified artificial dynamic "economic environment" in which firms potentially compete for employees in order to recruit a high quality skill set. Each firm has an internal model concerning what an optimal workforce should be. However, these internal models may or may not be correct. Additionally, what constitutes an optimal skill set may change over time dynamically. In this environment firms modify their work force based on hiring and firing policies making offers based on a pay policy.

2. Firms Incentives in a Resource-Based Theory of Rent Generation

According to the Resource-based View of the firm [9], the essence of a firm's characteristics is the bundle of resources that constitutes it and the main aim of firms is to acquire inputs to which rents may accrue [14]. Among the most widely quoted, Barney [11] [12] and [10] provided theoretical arguments explaining the link between resource heterogeneity and rents. Generally, this view sees firms' diversity and rents as generated by exchanges of heterogeneous resources in imperfect factor markets. These markets, generating information asymmetries, allow differences in resource positions among firms to be created and sustained. Rents are results of first-mover advantages: skilled managers reckon the value of resources and acquire them before competitors thereby building resource position barriers. Barney [11] [12] proposed that necessary conditions for inter-firm heterogeneity to be maintained in equilibrium include differences in *luck* or *foresight* among agents, *imperfect factors markets*, and *imperfect imitability* and *substitutability*. Resources are heterogeneous: some are more valuable than others; and imperfect factor markets ensure that agents maintain asymmetric information and different expectations concerning the values of the resources. Lucky managers, or managers with more accurate expectations, acquire before their competitors, valuable resources at a price that does not reflect yet their true value, thereby creating a rent. The latter can be sustained if the valuable resource is offered in limited quantity and/or if complexity and causal ambiguity [13] prevent competitors from recognising how to create value from the deployment of a particular resource. In both cases, either acquiring a valuable resource, or acquiring exclusive know-how concerning a particular process of value creation, the firm generates a rent.

More specifically, necessary conditions for rent creation are fourfold [19]. First, resources must be heterogeneous; this is a necessary condition for Ricardian and monopoly rents to accrue.

Second, imperfect factor markets must create *ex-ante* conditions for the rents not to be offset by the costs of resource acquisition. Information concerning value of resources should be asymmetric among agents to limit competition for resource acquisition. To create a rent, it is necessary that one agent, for reasons of luck or because he has more information than the others, hires the capable applicant at the same salary the other firms pay for their less productive resources.

Third, *ex-post* conditions, such as imperfect imitability and substitutability, allow resource heterogeneity to be sustained. Imperfect substitutability prevents substitute products from

decreasing rents via increases in demand elasticity [19]. Imperfect imitability derives from limitations in input or from cognitive and organisational difficulties in replicating a valuable resource.
Fourth, *imperfect mobility* ensures that valuable resources remain inside the firm. Indeed, the specificity of resources to the firm's asset base links valuable productive factors to the firm. If a resource was identically valuable for many firms, this could be sold in the market. Specific resources, on the other hand, are not tradable because they have a market price that is significantly less than their value for the firm employing it [14] [19]. The difference between the value for the owner and the market price of the resource defines a Paretian rent. This type of rent is also defined as *quasi-rent* [15] because the firm employing the valuable productive factor shares the rent with the productive factor itself [19].

Given the framework proposed above, a number of scholars stressed how firms' competitive advantages hinge upon the ability to build and maintain knowledge-based assets by integrating different skills within an organisation. Prahalad and Hamel [20], for example, focused on the concept of *core competence* as the "*…collective learning in the organization...*" advocating that firms ought to *"...to co-ordinate diverse production skills and integrate multiple streams of technologies…"* [20: 82]. More recently, Conner and Prahalad [21] set the premises to build a theory of why firms exist based on the capability of these latter to integrate knowledge of different actors. Grant [22] puts forward that the strategically most important resource of a firm is knowledge and that the essence of organisational capability is the integration of individual knowledge. Spender [23] proposes to establish on knowledge a dynamic theory of the firm [1996] and Kogut and Zander [24] recommend that firms, by the creation of an identity, facilitates internal processes of learning, knowledge communication and coordination. Finally, Adler [25] explains how trust has unique effective properties for the coordination of knowledge and suggests that community, rather than hierarchy or market, might be an efficient organisational form to integrate knowledge-based jobs.

3. Schumpeterian rents, Evolutionary Approach and Inter-Firm Heterogeneity

Some scholars [26] claim that the resource-based approach to inter-firm heterogeneity is associated with the concept of *appropriation* and competition is reduced to a race for first-mover advantage thereby overlooking intra-organisational processes leading to creation and adaptation of firm-idiosyncratic resource endowments. Indeed, the resource-based view of inter-firm heterogeneity apparently emphasises Ricardian rents and overlooks Schumpeterian rents.

The emphasis on long-lived and durable rents and on an equilibrium analysis explains the lack of interest in Schumpeterian rents. Schumpeterian rents generate dynamic inter-firm differences. In a Schumpeterian framework, a firm builds a rent by finding a new, more profitable, combination of productive factors. This position lasts until competitors are able to imitate it. However, the firm that introduced the innovation can use its advantage to conceive of another, new combination of inputs that puts this firm ahead again in the competition. On the other hand, the imitator, in his imitative attempt, might introduce a different, more advantageous, combination of productive factors. The situation described is not one in which a favourable competitive position exists, protected by imitation, but one where competitive positions evolve dynamically.
In this line of thinking, competitive advantage is the result of the ability to create and update a *situational fit* between combination of resources and environmental demand, playing an ever-changing, *dynamic puzzle game* [27]. Firms build up *flexible-response capabilities* [22] or *dynamic capabilities* [28] [29] [30] [31] in order to respond to dynamic environment by recombining their skill endowments.

As Levinthal [32] suggests, the analysis of Schumpeterian rents dynamics requires an evolutionary approach; firms do not pursue an optimising decision-making behaviour, rather they learn by searching for new, more efficient routines of resource management. Decision makers are *procedurally* rational [33] [34], they use heuristics to decrease the average number of searches and have different sets of choices because of their different histories [5].

4. Method and Approach

4.1 Modelling and Simulation

Modelling and simulation constitute a fundamental element of the research design[1]. Simulation helps rigorously to deduce consequences from modelled assumptions when complexity of modelling makes difficult to obtain closed-form solutions. In addition, simulation allows looking at unfolding organisational and social processes, capturing the behavioural characteristics in transitory states. In this work, we use a computer simulation model as a theoretical laboratory to analyse the circumstances in which different hiring and reward strategies, firms' heterogeneity and rent distribution patterns emerge. Alternative hypothetical, though dormant, trajectories will be activated by modifying the underlying modelled assumptions. This approach has the advantage of creating an appropriate setting to conduct controlled experiments. History can be re-run, showing how small, *ab-initio* modifications in parameter values can be amplified over time, to yield firms with distinct characteristics. Simulation is a unique methodology to perform this journey in history. This kind of method is a form of computational "thought experiment:" in which we ask "what if" questions in an artificial world. However, the ultimate aim is to allow us to develop hypotheses and theories that can then applied to real world phenomena and data. Our ultimate aim is to understand the real world. We use the computer model at this stage to help us to generate and test, in a rigorous and deductive way, candidate ideas.

4.2 Multi-Agent Based Simulation

We used an agent-based model to simulate interaction among employees and firms. Axtell [35], for example, presents a model in which employees spontaneously form firms and then make rational calculations on when and if to leave or stay. In that model firms are not modelled directly and agents act rationally to maximise their own payoffs. In our model firms are represented directly as an agent with internal models of the economy and our employees are far from rational – they greedily and locally attempt to increase their salary without any model of the economy or of the firms.

The processing of using computer simulation models in this way [36] is an emerging paradigm within the social sciences. Increasingly social scientists are using the techniques of multi-agent based simulation (MABS) to explore complex dynamics in artificial social systems [37].

The FirmWorld model should be viewed as an "artificial society" type model (i.e. similar to the SugarScape model [38]). It is not an attempt to capture a specific target such as some real labour market based on a real industry informed by quantitative empirical data. Rather, the FirmWorld model allows use to express formally (computationally) a number of hypotheses about potential processes that may occur in real labour markets but in a stylised and executable manner such that experiments can be performed to deduce the consequences of those hypotheses when they are combined in complex, adaptive systems (CAS). We therefore purposefully present a simplified model in which we hope to capture the kinds of complex dynamics in which we are interested.

The FirmWorld can be contrasted with previous more simplified tag-based skill models [39] in which only a single level agent is modelled and organisations (or groups) emerge from simple interactions. In the FirmWorld, firm agents directly recruit and coordinate their employee agents based on their evolving policies (see next section).

5. The FirmWorld Agent-Based Model

The model contains three kinds of agents: The Environment, Firms and Employees. They are related into a non-strict container hierarchy. The environment contains all other agents. Firms contain employees. However, firms may have no employees and agents may be outside all firms when unemployed.

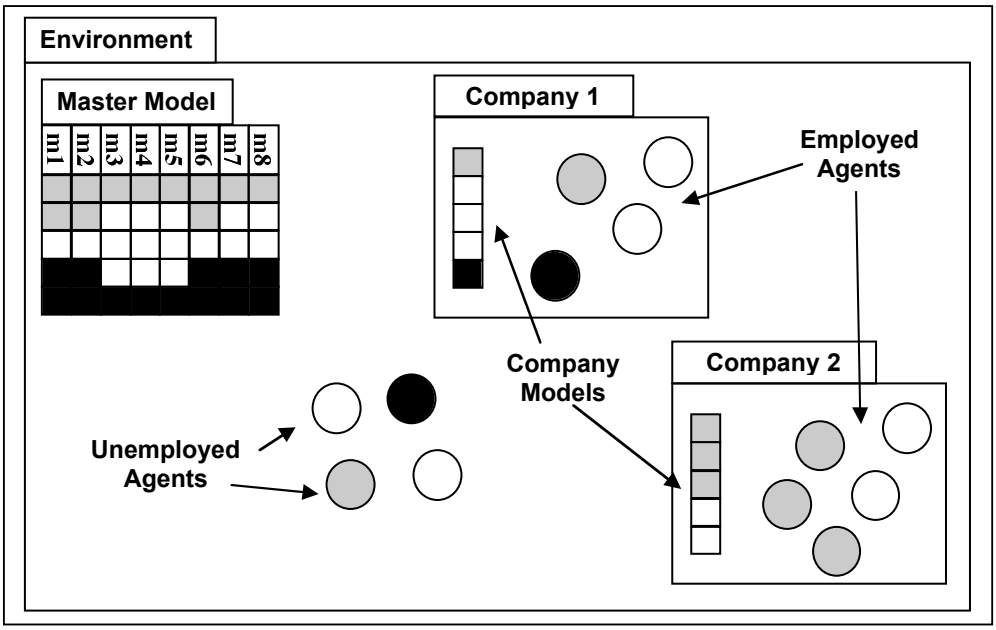
There is a single environment agent; it stores an economic model that represents the actual economy in which the firms and employees reside – we call this the "master model". This model is not directly visible to other agents. Only indirectly, via the receipt of earnings over time, do firms receive information from this model.

Firms consist of a hiring policy, a pay policy, an internal economic model (called a "firm model"), capital and a (possibly empty) set of individual employee agents. A firm with negative capital is considered bankrupt and is closed – making all employees unemployed. We describe the firm agents in more detail below.

Each employee possesses a single skill type from a set. Currently these are fixed and endogenously specified (so employees do not change skills). Each employee has an employment policy – a decision process that allows it to decide if to accept an offer from a firm. Currently, unemployed agents accept any offer but employed agents only accept offers that are higher than their current salary (incorporating a "security bonus" see below) - to this extent, employees can be seen as greedy maximisers. Employees currently have no internal economic model of their own so they cannot calculate their own worth and, hence, rents potentially accruing to employers. In the current implementation of the model, skills are represented by single cardinal values (though they have no ordinal significance). We describe employee agents in more detail below.

Figure 1 gives a schematic of the entire FirmWorld – indicating the major objects and their relationships.

Figure 1



A schematic diagram of the main entities in the system. The environment contains a "master model" giving the optimal set of employee skills for each cycle (here we only see eight cycles m1..m8 we use shades to indicate three skill types). Each firm contains a company model and some employee agents. Each firm attempts to make its workforce match its model by hiring and firing. In this case firm 2 has managed to archive this (it has 3 grey agents and two white agents) but firm 1 is one white agent short. The calculate earnings the workforce is compared to the master model for the given cycle and the distance calculated (see text for details).

6. A Month in the FirmWorld

The model is executed by running it for a fixed number of "cycles". We designate a cycle as a notional month. For all the experiments presented here we run the model for 120 months. At the start of each month, each company considers its internal economic model of the economy and its current employee skill set. If the company detects a shortfall in any given skill, it "advertises" publicly for employees with that skill. All unemployed agents, and some randomly selected proportion of employed agents, approach a randomly selected subset of firms advertising for their skill. The firms respond with salary offers based on their internal economic model, hiring policy and pay policy. This works in the following way: selected employee agent *i* approaches a set of firms F that have advertised for their skill. Agent *i* approaches each company *k* in F and *k* makes a job offer. A job offer consists of a salary amount plus a job status: either permanent or non-permanent. Those taking permanent positions cannot be fired at a future date; this is not the case for non-permanent jobs.

When making a salary offer, a firm uses its economic model to determine how much it believes its earnings would increase if it hires the employee agent and then makes a salary offer determined by its pay policy. The job status type offered (permanent or non-permanent) is determined based on the perceived scarcity of the skill in the market.

After the employee agent *i* has visited each company in F it takes the best offer proposed and revisits each company in F, communicates this highest offer, asking for a further offer. The company may then make a higher second offer or make no further offer.

If *i* is currently unemployed, it accepts the best offer and it becomes an employee of the relevant company. If *i* is already employed it compares the best offer with its current job and moves if a better offer has been made. No employed employee moves without consulting its current employer in the round of offers (this allows a current employer to retain an employee by making the best offer).

After the hiring process firms have their bank balance reduced by their total monthly costs (which equal total salary costs plus other fixed costs). The environment then allocates, for each firm, earnings for the month based on the composition of the workforce in the firm and the economic "master model". That is, we assume that the composition of the workforce (number of employees with each skill) determines the earnings for each firm. Hence, two firms with identical workforces will receive identical earnings.

Currently, the economic "master model" stored by the environment is an exogenously defined optimal employee set (number and skill set) for each month. In our experiments we have modelled both static (never changing) and dynamic (constantly changing) "master models".

The workforce of each firm is compared to the current "master model". For each useful employee (an employee with a required skill) a marginal contribution to the firm's total sales is calculated using both a marginal decreasing return function and a specificity function that adjust marginal productivity by assessing how specific the employee is for a company (see below). The more specific the employee is for a company, the higher his value will be for the company.

After companies have received their income from the economy, they pay their outgoings (salary and fixed costs). Those companies that run-out of capital go bankrupt – they close and all their employees become unemployed. Since our model imposes a fixed number of companies, when a company goes bankrupt, a new one is immediately formed to take its place. The new company copies the characteristics of a successful company (such as internal company model and pay policy) and then changes this slightly with a low probability. This is a kind of "replication" and "mutation". New companies start with some initial capital and zero employees.

Below is an outline algorithm of FirmWorld. In the following sections we describe in a little more detail the behaviour of Employee agents and Company Agents to cover each of the processes described in the sequence.

FirmWorld Outline Algorithm

```
Initialise firms
Initialise employees
Loop for 120 cycles
     Firms fire non-permanent employees they do not want to
keep
     Firms advertise job vacancies
     All Unemployed agents approach some companies for
offers
     Sample of employed agents approach some companies for
offers
```

```
     Companies are awarded income and pay costs and
salaries
     Bankrupt companies dissolved – employees become
unemployed
     New companies formed – copy "gene" of more successful
companies
End Loop
```

7. Employee Agents

Employee agents are relatively simple (see figure 2). They are marked by a single skill, a number fixed for the career of the agent. In the experiments for this paper, in all cases, there are 5 skills represented by the cardinal numbers [1...5]. Associated with an agent's skill is a second value called the skill "specificity factor" (*sf*). This is a real number [$1<sf<2$] representing how specialised the skill is to a particular employer. A high value means the skill is of high value to the current employer but of low value to another employer. This value is not fixed but changes during the career of the employee.

Figure 2

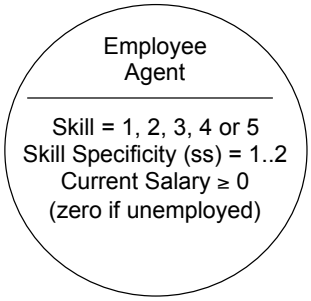


Figure 2 shows the composition of Employee agents – they store a fixed skill, skill specificity and a current salary.

7.1 Specificity of Skills

The specificity factors of all agents start at $sf = 1$. They increase non-linearly (following a convex, learning, curve function) over each month an agent is employed in a given company such that after 4 notional years (48 months) the *sf* value goes from 1 to 2. The *sf* value is not allowed to become greater than 2. However, the value is reduced back to 1 if an agent leaves its current employer. This captures the notion that skills produce value when embedded within a firm-specific network and training is a socialisation process that takes place when a new employee is embedded within a group of incumbent workers [40][41][8]. In equation (1), *sf* is a function of x, that is, the number of months that an employee i stays within the same organisation k. They

become specialised within a company, build firm-specific knowledge and this latter is not transferable to other firms (see figure 3.a).

$$sf = 2 - e^{-\frac{x_{i,k}}{12}} \qquad (1)$$

The specificity potentially adds value to the company. If an employee possesses a required skill then the marginal return generated by the employee is multiplied by the *sf* value. When firms consider employing an agent they consider the specificity value to be 1 even if it is higher for a current employer – since the value is reset to 1 if the employee decides to move.

Figure 3

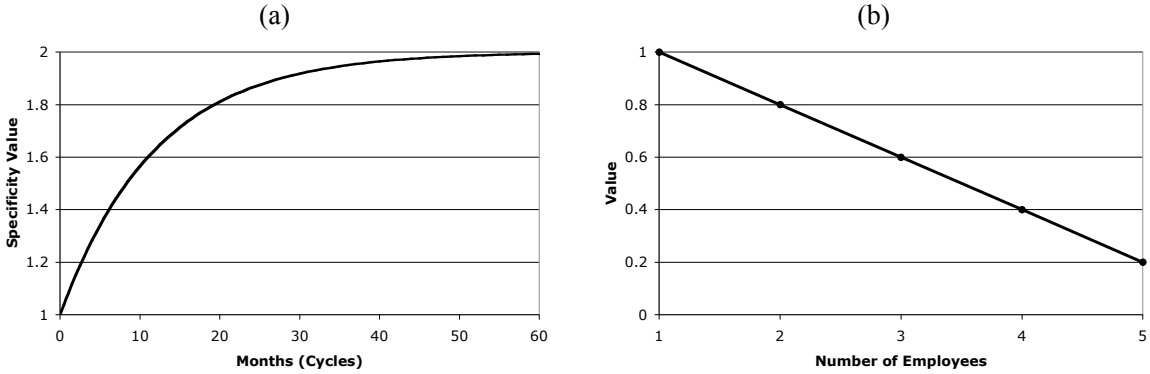


Figure 3a shows a graph of the specificity function ($y = 2\text{-}e^{-x/12}$) where y is the specificity value and x is the number of months with the same employer – this represents a kind of simplified "learning curve". Figure 3b shows a graph of the simple linear marginal return function ($y = 1\text{-}((x\text{-}1)/n)$) - here shown where the number of required employees of skill *i* is n = 5 and the number of employees already in the organisation holding skill *i* is x. A company uses its internal company model to choose the n value and the environment uses the master model.

7.2 The Employee Career History

All agents start out as unemployed but may become employees of a firm through a hiring process as described previously. Through bankruptcy or firing, they may become unemployed again during periods of their career, employees may also move among firms comparing different salaries offered – this way agents may make several career moves during a simulation run (which simulates a notional 10 years).

Agents do not exercise complex decision processes or maintain internal models of the environment, firms or other agents, they simply choose the best jobs offered to them and move to the associated firm.
Over the course of a career (the entire length of a simulation run) their skill never changes but the specificity may change several times. If an agent joins a firm and stays there for many months then its specificity will eventually become 2; in this condition the employee is potentially worth twice the maximum of what it could be worth in any other firm (graph in figure 3.a describes employees' specificity curve).

7.3 Marginal Productivity of Employees

Productivity of each worker marginally decreases as the number of employees hired with the same skill increases. Therefore, we define *max mp* as the maximum marginal productivity that a

worker contributes when hired in a firm. If the economy requires more than one worker with the same skill, each new worker hired with that skill will contribute $mp < max\ mp$. In equation 2, we model marginal productivity as a function of the number of employees hired in a firm k with skill j ($e_{j,k}$) and the total number of employees holding skill j that an organisation k desires to employees according to its company model ($\bar{e}_{j,k}$) (see graph in figure 3.b).

$$mp = 1 - \left(\frac{e_{j,k} - 1}{\bar{e}_{j,k}} \right) \qquad (2)$$

7.4 Marginal Contribution of Employees to Firms' Sales

In our model, the marginal contribution that each worker gives to a firm's total sales depends both on his marginal productivity, as specified in equation (2), and the specificity (*sf*) of a worker in an organisation. Thus, specificity is the second element that influences a worker's contribution to a firm's sales. In the model, to obtain a worker's contribution to sales (*mc*), then, we multiply his marginal productivity, as calculated in equation 2, by the specificity factor *(sf*) which captures his embeddedness within the organisational network. Thus, each hired worker contributes zero if his skill is not required by the economy or if the firm has already the amount of workers required with that skill. If the worker hired bears a skill required by the economy, he contributes with:

$$mc = mp \cdot sf \qquad (3)$$

8. Company Agents

Company (or Firm) agents store a model of their believed optimal skill set called the "company model". This model is a vector giving the number of each kind of skill believed to be optimal and represented by a believed optimal workforce skill set – that is what the firm *believes* would produce optimal results. This may or may not match the *actual* optimal skill set contained in the master model for a given quarter. It is important to realise that firms are not rewarded based on the similarity of their economic model but on their *actual* workforce, which although informed by the model will often not be identical to it since this depends on the hiring process and other policies. However, the firm uses its actions to attempt to minimize the gap between its economic model and the actual workforce composition. The firm tries to recruit employees that match the required skills. Hence if the company model indicated a company needed two employees with skill 3 and 1 employee with skill 5 then the company would "advertise" vacancies for those skills if it did not have enough employees with those skills.

So to summarize we have the interplay of three factors that determine the earnings of a company, its internal economic model (the firm model), its actual workforce composition and the master model (representing the actual optimal workforce composition determined by the economic environment).

8.1 Hiring Employees

A company model may or may not match the master model. If it does then the firm will tend to make "economically rational" decisions when hiring and firing. Obviously, if it bears no relationship to the master model a firm may hire employees that add no value and do not increase their earnings in reality.

In addition to the company model, firms store three real values that potentially affect hiring, firing and salary offers (*ne*, *oe* and *st*). When a salary offer is made to a potential employee *i*, the firm uses its company model and the prospective employee's skill to calculate the value the firm believes the new employee would add (*mc*) excluding any company costs. The offer made is not this full amount *mc* rather it is $mc \cdot ne$. So for $ne < 1$ the offer is less than the believed value and if $ne = 1$ it is identical. The *oe* value is used in a similar way but for "firing" calculations (see later). The *st* value gives a "scarcity threshold" above which a skill is considered "scarce". New employees with scarce skills are offered permanent contracts (see below).

Essentially, then, the company model, combined with *ne*, *oe* and *st*, defines the hiring and firing behaviour of a company; one can think of these three items combined as a kind of "company gene", which, if copied to another company, brings over much of the behaviour.

8.2 Scarcity of Skills and Permanent Contract Offers

When companies make a job offer they make a salary offer and a status offer (permanent or non-permanent). They decide on this latter aspect by assessing whether a skill is "scarce". If it is, then they make a permanent offer.

A firm calculates the binary function of scarcity for a given skill in the following way: a proportion is calculated as the number of companies still advertising for employees with the given skill *after* the recruitment phase, i.e. the proportion of companies still requiring the skill. If this value is *larger* than the internally stored scarcity threshold (*st*) then the company offers a permanent contract.

Hence, companies with low *st* values are characterized by a propensity to offer permanent contracts as they will offer permanent contracts at lower scarcity than those with high *st* values. As stated previously, the *st* value forms part of the company "gene" and is copied by new companies from the more successful companies (based on profit).

Employees on permanent contracts cannot be fired; however, they are more loyal than temporary workers and are much less likely to look for new jobs (probabilistically 75% less likely). The assumption is justified by the fact the employment contract may contain clauses and agreement that disincentives employees from leaving a firm. For example, employment contracts may contain *non-compete* clause, which forbids employees from working for competitors for a given period of time after leaving the firm [3]. In addition, importantly, when employees decide on the "best" job offer they weight a permanent offer by notionally increasing the salary offer by a "security bonus" (currently set to 100% for all employees). This means that a permanent offer is "as good as" a temporary offer of double the salary.

8.3 Firm Financials, Bankrupt and Evolutionary Learning

Firms maintain a bank balance (which is initialised to some positive value for new firms) from which payments are made (fixed costs and salaries) and sales are paid into. Firms' sales are given by the sum of marginal contributions of skills of the workers employed in the firm. Thus, sales of firm *k* are:

$$s_k = \sum_{i=1}^{n} mc_i \qquad (4)$$

Firms total costs (*tc*) are the sum of fixed costs ($\overline{c}$) and salaries, which are variable costs (c) depending on the number of employees hired. We did not consider any economies of scale.

$$tc = \overline{c} + c \cdot n$$

If the balance goes below zero then the company is considered bankrupt. When a firm becomes bankrupt, all employees are made unemployed and a new company is formed to take its place. The new company is not initialised randomly but sets its "gene" (i.e. company model and *ne, oe* and *te* values) by sampling a subset of the population of firms and copying the "gene" of the firm with highest, last cycle, profit. Also with some small probability the "gene" is "mutated" by applying small random changes to the company model and the *ne, oe* and *te* values. This creates a weak evolutionary learning in which profit in the last cycle can be seen as a measure of fitness. The process is weak in the sense that we assume that inertia prevents firms' adaptation. Thus, learning is determined by firms' selection. In this respect, we assume that both company model and hiring policies are elements of core features that firms cannot easily adapt [42][43][44].

The number of firms is kept constant for the purposes of simplicity of analysis and modelling. We wait for a firm to "die" before reproducing a successful one. However, a similar evolutionary process would emerge if high profit firms spontaneously generated copies.

8.4 Firing Employees

Firms periodically reassess the value of their current employees on *non-permanent* contracts using a similar method as for hiring new employees. The only difference is that the calculated value of the employee (*mc*) based on the company model and specificity (as previously described) is multiplied by *oe* instead of *ne*. If $mc \cdot oe <$ current salary then the employee is fired. Firing is only allowed for employees on non-permanent contracts. Hence a company with a high *oe* value is more likely to keep its non-permanent employees than one with a low value.

9. Simulation Experiments

In our initial experiments we made runs for 4 scenarios based on the different combinations of two binary dimensions:

1. Scarce labour (SL=1) v. abundant labour (SL=0)
2. Static economy (FE=1) v. dynamic economy (FE=0)

For 1, scarce labour meant 200 employees, abundant labour 400. For 2, in the static market case the master model was never changed, in the dynamic case the master model was changed slightly with some probability each month.

For all experiments, we fixed the number of firms at 50 and the number of different skill types to 5. Each experiment was run to 120 cycles (notional months). The master model was set to one for each skill type (i.e. the optimal firm would contain 5 employees, one with each skill). Company model vectors were initialised randomly with each skill being set to a uniform randomly drawn integer [0...5]. The *ne* and *oe* real values were drawn from the range [0...1]. New companies were initialised with a bank balance of 50,000 units, maximum marginal productivity of a skill is 1000 and fixed costs of 5000 units per month.

Given these values, the maximum value of sales for a firm would be 2000 x 5 = 10000 units, if it had the perfect skill set, highest specificities and if economy required one worker for each of the

skills. Maximum profit would be 10000 – (costs of) 5000 = 5000 units. This is true in the fixed economy case (FE=1). Where the economy was dynamic (FE=0) then the possible maximum income values will changed randomly over time because the number of employee agents required for each skill in the master model follows a random walk.

Employees were initialised with a randomly selected skill [1...5] and a skill specificity (*sf*) of 0.5. This means that at the start of each simulation run, skills are, probabilistically, distributed evenly over the population.

For each of the 4 scenarios we ran 100 independent runs with different pseudo-random number seeds.

10. Findings

The conducted experiments explore how firms manage with different policies their skill endowments. In the following, we focus on how management policies emerge as we move from a stable to a dynamic scenario, with labour scarcity. The aim of the experiments is to understand what kind of organisations survive in a dynamic economic environment where firms need both to nurture firm-specific skills and accommodate competitive pressures that evolve rapidly and generates ambiguous signals concerning strategic values of different skills.

10.1 Stable economy

In the stable economy, firms' perceptions regarding scarcity and strategic value of different skills converge toward the economy model through a process of imitation. Thus, firms are able to aggregate workers in different categories depending on the value of their skills.
As explained in figure 4, permanent contract includes two clusters of workers. A first cluster, in the upper right corner of the graph, includes valuable scarce skills which have been hired with long-term contracts and receive high wages. A second cluster includes skills which are scarce but produce less value. For this reason, these skills have lower salaries given similar level of specificity.

Figure 4

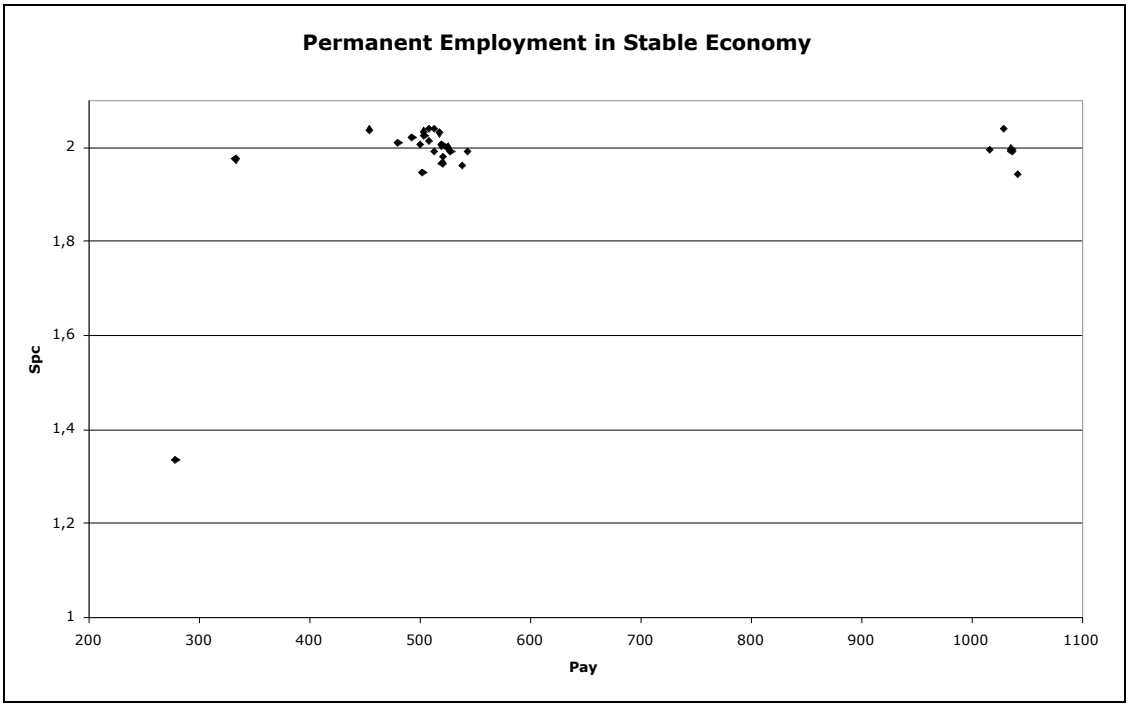
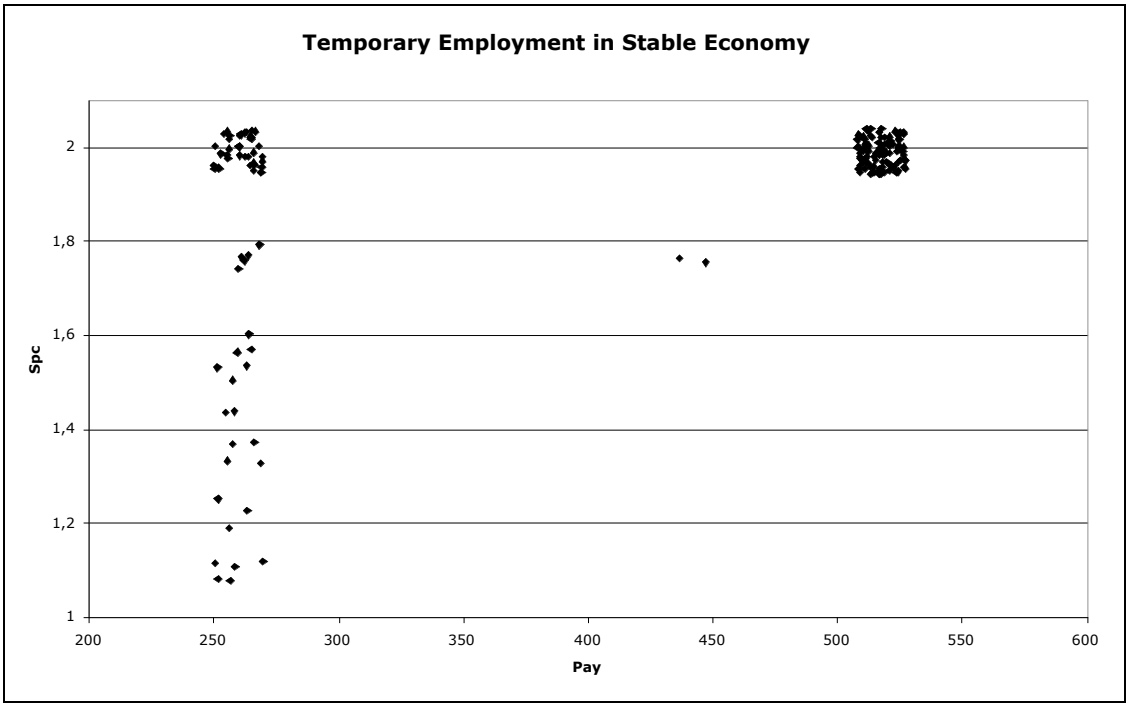


Figure 5 describes clusters of workers holding a temporary contract in the fixed economy. Here again we can recognise two clusters of workers. We can recognise, in the top right corner of the graph, a cluster of workers enjoying high salaries, they are not scarce but their skills are strategic. On the top left corner, we can notice the cluster of workers whose skills are neither strategic nor scarce. As we can see, there is a vertical line of workers below this cluster; along the vertical line are distributed workers that have different level of specificity but the same low salary. This phenomenon is generated by the fact that the workers, holding a temporary contracts, move from one firm to the other, thus, their specificity level is different. However, since their skills are not required by the environment, their wages remain equally low notwithstanding the different level of specificity. These workers are similar to commodities; they move among organisations, these latter are able to lower salaries putting workers in competition.

Simulation experiments suggest, that, in general, with stable economies, firms have a clear idea of which skills they need to hire long term and which skills can be managed with temporary arrangements. Thus, the proportion of permanent contracts is low compared to temporary jobs. Temporary workers move among firms as commodities, scarce strategic skills are locked within firms and produce value by building firm-specific knowledge. Firms, in general, maintain high bargaining power that allows them to appropriate large part of value produced by labour.

Figure 5



10.2 Dynamic economy

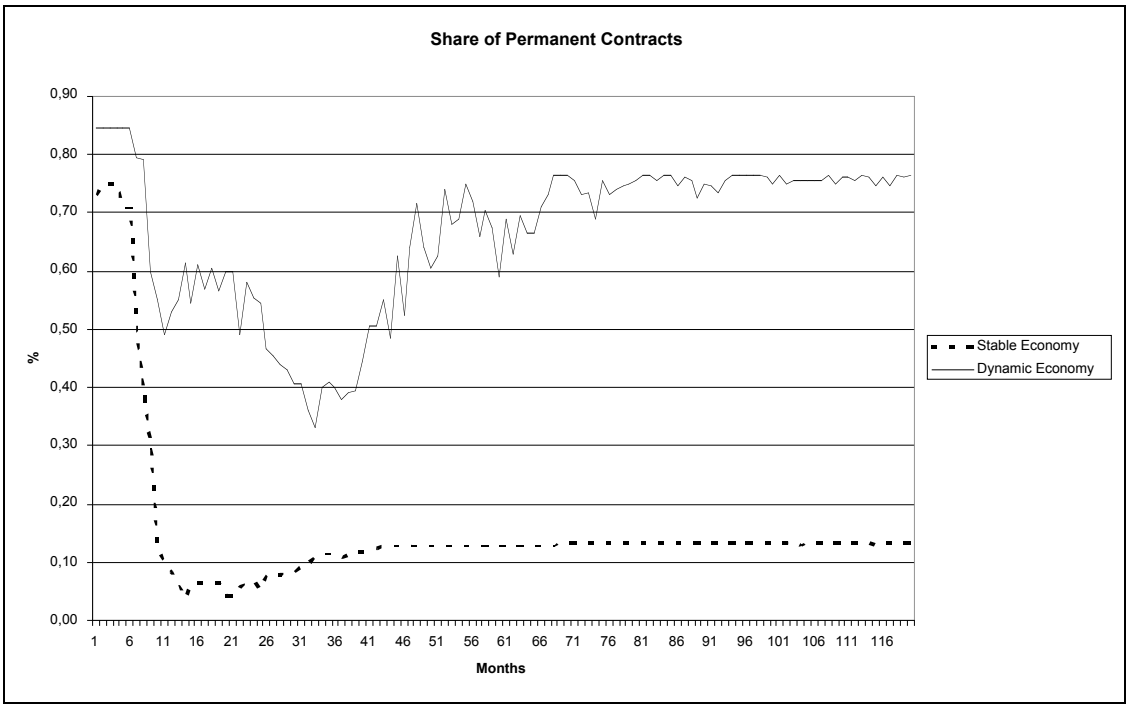
In the dynamic economy, strategic value of skills changes rapidly thereby producing ambiguous signals to firm that adapt their skill endowments. Ex ante, we expected that firms would have preferred to select temporary contracts. The expectation was grounded on two hypotheses.

First, assuming ambiguous information on skill strategic values, we expected a repertoire of very different models among firms, indicating different priorities in term of skills hiring. Different models would lead to different hiring policies; thus, we expected that firms directed their attention to different skills thereby decreasing perceived scarcity of each skill. We expected that lower perceived scarcity favoured temporary rather than long term contracts.

Second, the emergence of a large proportion of temporary contracts was suggested by the evolutionary mechanism built in the model. We expected that those firms, which were initially assigned a high propensity to hire long term, would have been selected out in a simulated environment in which firms need to be more flexible and continuously adapt their skill endowments to evolving competitive environments.

Simulation experiments proved that our expectations were faulted: as described by figure 6, in a dynamic environment the proportion of long term contracts is, on average, significantly higher than in stable economies.

Figure 6

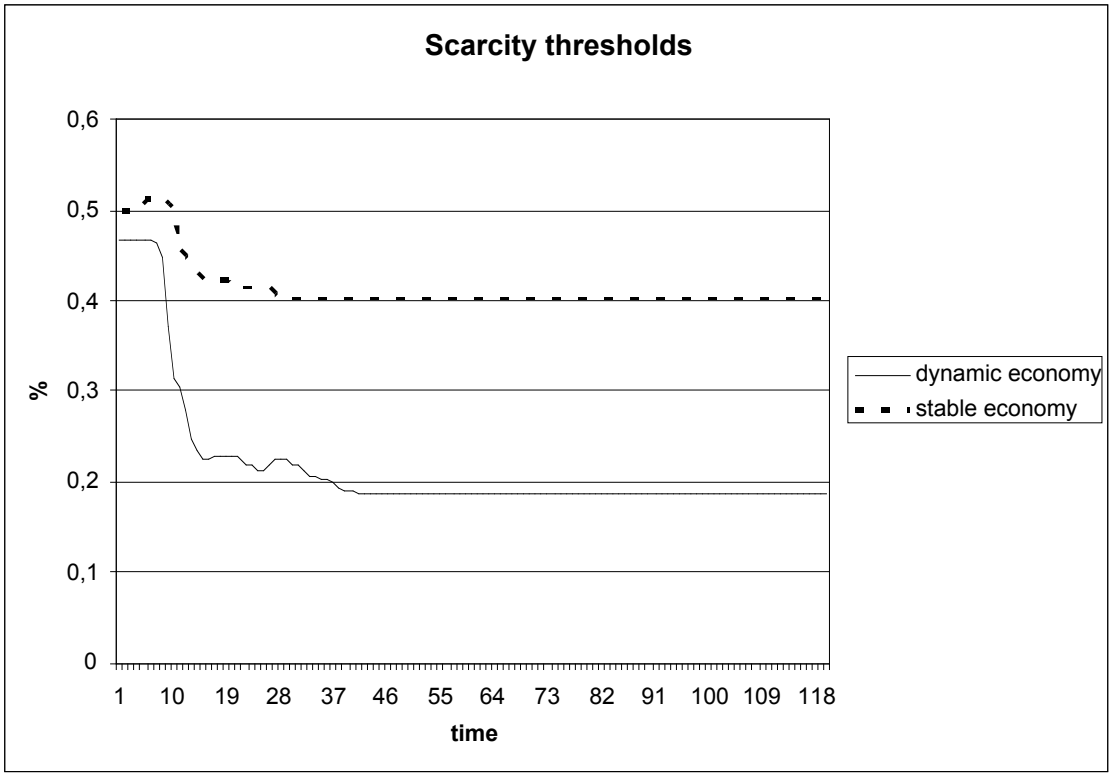


The simulation experiments articulate a counterintuitive lesson that suggests two plausible causes of observed behaviours.

A first mechanism deals with the interaction among individual perceptions and aggregate decision-making. Firms have different perceptions concerning skills' strategic value this leads them to use, at least some, long term contracts. As environments change, skills' value changes as well. Firms cannot lay off their employees so the employees already hired with long term contracts will remain within the firm. In addition, the firm will hire with long term contracts those skills whose strategic value has increased due to change in the status of the economy. As the process continues, firms rush to hire long term workers endogenously generating labour scarcity and perceiving an increasing need for long term contracts.

A second explanation of the spread of long term contracts in the dynamic economy is the evolutionary selection of such a hiring policy in the simulated economy.
If hiring temporary workers was the optimal strategy in a dynamic economic environment, why new firms created did not copy such a strategy? Why the attitude to hire long term is positively selected in the evolutionary process? The reason is that firms hiring long term have superior performances and, thus, are copied by firms newly created. As described in figure 7, as the simulation unfolds, firms operating in a dynamic economy decrease their 'scarcity threshold' compared to firms operating in stable economies. That is, in what we defined as a dynamic economy, surviving firms tend to use permanent contracts more frequently that those firm operating in a stable economy.

Figure 7



In other words, in our simulation, firms have a biased model of which skill is really valuable. Such a perception is honed through the process of going bankruptcy and recreating a firm that copies strategies of best performers. This adjustment delay is long and by the time a company has reshaped its strategy, the environment might have changed again making useless any previous adaptation. In this context, temporary contracts do not generate a sustainable competitive advantage.

On the other hand, by hiring long term skills as they emerge as scarce, firms build a very rich skill endowment, composed of skills whose value change as the simulation unfolds. The key issue is that these skills, by remaining within a firm, build firm-specific knowledge. Had the value of a particular skill, included in a firm's skill endowment, to increase, the demand for that skill will increase in the labour market. Yet, high specificity of the skill decreases its transferability and the firm will be able to maintain the skill in the organisation paying a lower salary compared to the salary the firm would have paid if the skill had been fully transferable to other organisations. This is because, full mobility leads to bargaining processes that increase salaries and decreases rents appropriated by firms. In this respect, lack of specificity and transferability increases the portion of rent appropriated by labour [19]. Figure 8 can help us in explaining this mechanism. By looking at figures 4 and figure 8 together, we can compare emerging clusters of workers holding permanent contracts in both fixed and dynamic economies. In dynamic economies, it is much harder to define well defined clusters. If we look at the left side of the graph, we recognise a vertical line of workers with low wage and increasing specificity, which is not present in graph 4. If we look in the upper part of the graph, we can see a horizontal line of workers with fixed specificity and high wage, which, again, is not present in graph 4. The vertical line of workers on the left is explained by the fact that, differently from the stable economy, firms' lifetime is

shorter, new firms rapidly substitute for failing old firms. In this environment, new entrant firms hire scarce skills in different point in time, this explaining why we can observe different levels of specificity.

The horizontal line in the upper part of the graph has another interesting explanation. As we can see, differently from graph 4, in graph 8, strategic values of skills may change as the simulation unfolds. Thus, firms hire workers bearing skills whose value change along the simulation. As a consequence, firms pay different salaries to workers with the same levels of specificity depending on the strategic value of the skills they bear. Yet, given the high level of specificity and, consequently, low transferability of skills among different firms, job market does not erode firms' rents and the wages paid are much lower value created as described in figure 9.

Figure 8

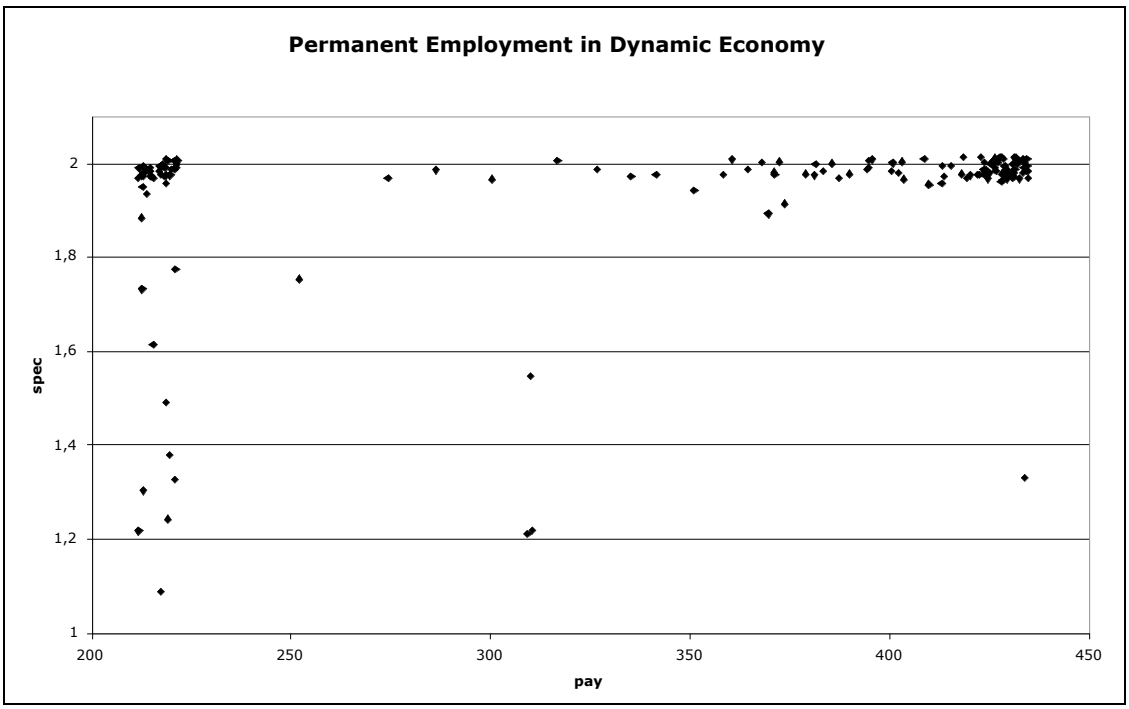
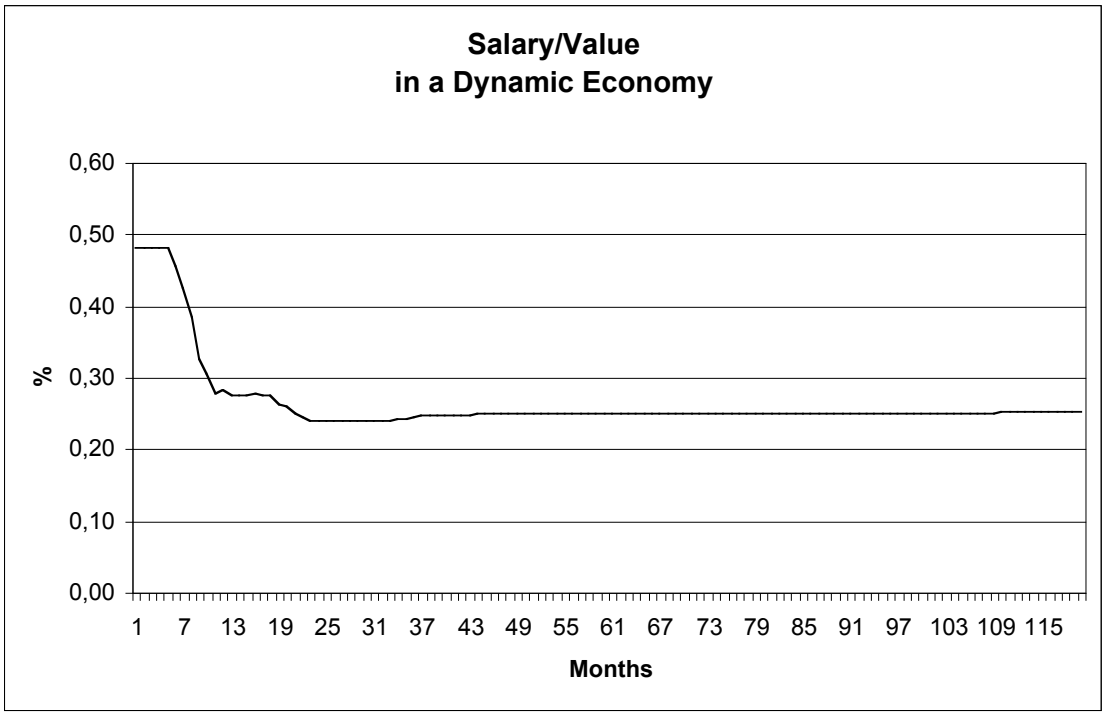


Figure 9



11. Discussion: Learning without Earning

The simulation experiments in our work suggest that mobility favours firms when estimation of expected marginal productivity of workers is not ambiguous. In such a context, where employers are able to discriminate between strategic and non-strategic skills, it is possible to use temporary jobs to increase mobility for non-strategic skills. On the other hand, the use of temporary jobs and mobility may lead to counterintuitive results when estimated expected marginal productivity of workers is ambiguous in highly dynamic environment. When competitive environmental dynamics continuously change the strategic value of different skills, the attempt to continuously adapt to environmental requirements, adopting hiring policies based on temporary contracts, may be suboptimal for two reasons.
First, firms may find themselves hiring skills when the expected marginal productivity of these latter is high and, thus, the market salary is increasing.
Second, a firm, once has paid high wage to hire the worker, might have to discover that competitive environment has changed and the expected marginal productivity of the skill is decreasing.

The results of our work support the idea that in a dynamic environment firms have higher survival performances when they use long-term contracts to build and maintain a repertoire of different firm-specific skills. These skills provide the organisations with the flexibility and adaptability needed to take advantage of emergent opportunities and neutralise threats [45].

In the environment that we describe in our model, firms are fairly inertial and observed organizational change result more from organizational selection than from voluntary adaptation [42][43][44]. The rate of change in the environment is faster than the speed of learning

mechanisms, that is, the rate at which firms are able to hire new skills and extract rents by exploiting accumulation of firm-specific knowledge. Thus, firms are better of when, rather than trying to follow environmental change, maintain a repertoire of skills to face different competitive settings. The situation that our experiments depict calls to mind Hannan and Freeman's hypothesis of structural inertia, according to which *attempts at reorganization increase death rates* [43: 159].
On the other hand, our study suggests that, despite their inertial features, organizations may adapt to evolving environments by exploiting the network-specific nature of organizational learning. In our model, employed workers, when embedded within an organisation, start to learn. If their skill is not strategic, given the competitive context, they accumulate network-specific knowledge but not necessarily their salaries increase. As the environment evolves and their skills become strategic, the specificity of their know-how makes the skills not perfectly tradable in the job market. As a consequence, the emerging idiosyncratic nature of the skills push downward the wage that other firms are ready to offer thereby decreasing the wage that the original employer needs to pay to retain the worker. Employers are thus able to retain portion of quasi-rents because they are not paying the full value of extra output of their firm-specific human capital [46].

Thus, our study tease out two mechanisms – protection of quasi-rents and speed of organisational learning – that explain why trading off flexibility in hiring policies with adaptability, this latter deriving from reallocating a large repertoire of firm-specific skills, may result in a successful strategy. Advantages accrue to firms not only because they are able to fit changing needs of environment with their skill portfolio but also because they are able to protect quasi-rents produced by firm-specific human capital. Thus, while in some authors, for example, [8], idiosyncratic jobs create a small number bargaining situation in which incumbents workers with idiosyncratic know-how opportunistically display a perfunctory cooperation and destroy portions of idiosyncratic efficiencies gains, we stress how job idiosyncrasies may define a 'bilateral monopoly' in which, once relationship is established, both parties lose if it is terminated [47]. Indeed, we focus on the small number situation on the demand side created by idiosyncratic know-how which cannot entirely be transferred to other organisations. In our model, employers' opportunistic behaviour facilitates the acquisition of scarce skills at a salary that allows large rents to be extracted.

Another issue concerns the relations between our findings and results from empirical studies.
In some respect, our argument has a connection with the countercyclical hiring posited by Greer and Ireland as these latter found that firms having high financial performances adopt a countercyclical hiring; that is they hire in downturns when salaries are lower [48].
On the other hand, a number of studies found a positive correlation between variability in employment levels required by economic cycles and use of temporary workers [49]. Our simulation experiments suggest that the empirical relationships between variability in demand of jobs and the hiring of temporary workers may be mediated by the relationship between workers' marginal productivity and their embeddedness within firm-specific networks of skills. Indeed, Davis-Blake and Uzzi [49] also found a negative correlation between jobs requiring firm-specific training and use of temporary workers.

A further issue concerns the relationship between size and the use of temporary workers. Davis-Blake and Uzzi [49], for example, found that large firms are less likely to use temporary workers. They observed a negative relation between size and use of temporary workers thereby advocating the hypothesis that large firms can reallocate employers within the organisation. Observable behaviours in our simulations are coherent with these findings. In addition, simulation experiments help to articulate hypotheses concerning causal relationships among firms' size, use of temporary workers and adaptability. The mentioned empirical study ex-post captures a

relationship between organizational size and hiring of temporary workers: because firms are large, they can adapt by reallocating workers. In our experiments, we design firms of the same size and give them the choice to growth and build large repertoires of skills or remain small and adapt by hiring temporary workers as environment requires different skills. The observed emerging pattern is that firms that decide, in turbulent times, to use permanent, rather than temporary, workers get larger and are more successful than firms that decide to remain small and hire temporary workers. The issue is explored in the graphs in figure 10, 11 and 12. The graphs describe the relationships among age, size and use of permanent contracts in a typical simulation run, in a dynamic economy. In figure 10, the firms that survive longer, approximately 6-7 years, are those that use large proportions of permanent contracts. These firms, as described in figure 11, are the larger ones, approximately ten employees. Figure 10 and 12 suggest that it is unlikely to observe firms that survive longer than 1 year by adopting a small proportion of permanent contracts (around 20%) and maintaining a small size.

Figure 10

Figure 11

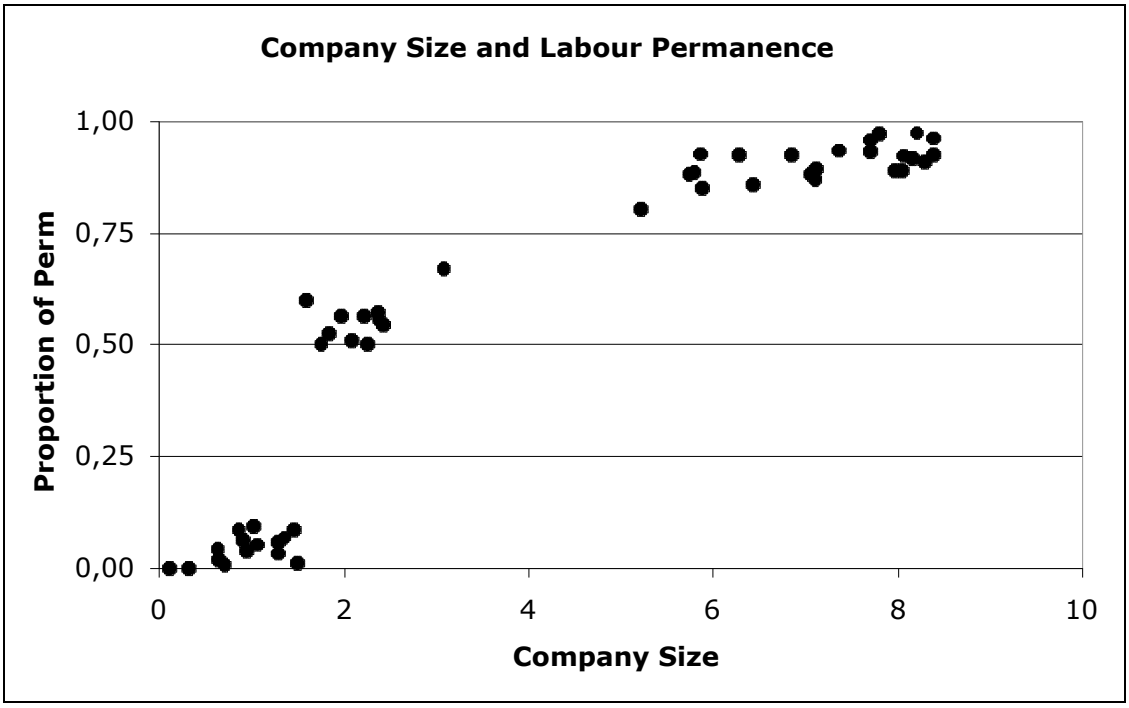


Figure 12

Last issue concerns what kind of organization is likely to emerge in a knowledge economy. If we assume that in a knowledge economy an increasing number of technological disciplines are required and organizations increasingly need to rely on a large number of knowledge specialists interconnected within firm-specific networks [2: 9-10], then it is interesting to speculate on how organizations ought to manage the employment relationship with these specialists.
Davis-Blake and Uzzi, for example, in the mentioned empirical study, found that large firms are more likely to hire independent contractors to have temporary access to specialised skills and services. Again, we suspect that the problem is to assess the extent to which the specialised skills are firm-specific and firm-specificity impacts on marginal productivity. Our study suggests that hiring independent contractors in turbulent environments may work if specialised skills do not need network-specific learning and proposes, instead, the hypotheses that firms in the knowledge economy may have an incentive to have an increasing number of specialists of technological disciplines in-house [50]. Thus, the simulation experiments support the argument that, in turbulent environments, firms, which rely on in-house availability of diverse specialists, better face change because specific skills provide access to cutting-edge knowledge and novel solutions to organizational problems [51][52].

Concluding, in most western countries, even in those in which employment is growing, the proportion of jobs that qualifies as temporary or part-time is dramatically increasing. Interest in this pattern is motivated by the concern of a decreasing quality of the job stock. Indeed, temporary jobs suffer from reduction in real wages, increased inequality in wages, reduced job protection and insurance benefits [53]. A number of studies addressed the welfare implications of temporary employment. Jenkins and Chun-Yan Kuo [54], for example, addressed the social opportunity costs of temporary employment. The angle that we take in our study deals with firm-level strategies and suggests that, in dynamic environments, the use of temporary jobs might results in decreasing survival performances at organisational level.

A line of further work that is worth considering concerns the representation of actors-employees reaction to firms' policies. In particular, in the present model, employees do not control the rate and direction of learning. In a further work, we intend to model with more accuracy workers' decision-making. On the one hand, we have in mind to assign workers the decision on whether or not to invest in firm-specific know-how. To take the decision, we assume that employees consider long-term job security [1: 248] and that, in general, high turnover rates discourage firm-specific learning [55]. On the other hand, we would like to model workers' capability to develop particular skills, evaluating, both by imitating other workers and by interpreting environmental scenario, which skill it is desirable to build up.

12. References

1. Hodgson GM: Economics & Utopia, London, Routledge, 1999.

2. Foss N: Strategy, Economic Organization, and the Knowledge Economy. The coordination of firms and resources, Oxford University Press, 2005.

3. Liebeskind J: Knowledge, Strategy, and the Theory of the Firm. Strategic Management Journal 1996; 17, Winter Special Issue:93-107.

4. Polanyi M: Personal Knowledge, Chicago, IL, University of Chicago Press, 1962.

5. Nelson RR, Winter SG: An Evolutionary Theory of Economic Change, Cambridge, Mass., The Belknap Press of Harvard University Press, 1982.

6. Aoki M: The Partecipatory Generation of Information Rents and the Theory of the Firm; in Aoki M, Gustafsson B, Williamson O (eds): The Firm as a Nexus of Treaties, London, Sage, 1990.

7. Teece DJ, Pisano G: The Dynamic Capabilities of Firms: An Introduction. Industrial and Corporate Change 1994; 3(3):537-536.

8. Williamson O, Watcher ML, Harris JE: Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange. The Bell Journal of Economics 1975; 6(1):250-278.

9. Penrose E: The Theory of the Growth of the Firm, Oxford University Press, ed 3, 1995.

10. Wernerfelt B: A resource-based view of the firm. Strategic Management Journal 1984; 5: 171-180.

11. Barney JB: Strategic factor markets: expectations, luck, and business strategy. Management Science 1986; 32(10).

12. Barney JB: Asset stocks and sustained competitive advantage: a comment. Management Science 1989, 35(12).

13. Reed R, DeFilippi RJ: Causal ambiguity, barriers to imitation and sustainable competitive advantage. Academy of Management Review 1990; 15(1):88-102.

14. Conner KR: A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?. Journal of Management 1991; 17(1):121-154.

15. Mahoney JT, Pandian JR: The resource-based view within the conversation of strategic management. Strategic Management Journal 1992; 13:363-380.

16. Ulrich D, Lake D: Organizational Capabilities, New York, Wiley, 1990.

17. Lado AA, Wilson MC: Human resource systems and sustained competitive advantage: a competency-based perspective. Academy of Management Review 1994; 19(4):699-727.

18. Harrison JR, Carroll GR: Keeping the Faith: A Model of Cultural Transmission in Formal Organizations. Administrative Science Quarterly 1991; 36(4):552-582.

19. Peteraf MA: The cornerstones of competitive advantage: a resource-based view. Strategic Management Journal 1993; 14:179-191.

20. Prahalad CK, Hamel, G: The Core Competence of the Corporation. Harvard Business Review 1990, May-June.

21. Conner KR, Prahalad CK: A Resource-Based Theory of the Firm: Knowledge versus Opportunism. Organization Science 1996; 7(5):477-501.

22. Grant RM: Prospering in Dinamically-Competitive Environments: Organizational Capability as Knowledge Integration. Organization Science 1996; 7(4):37-387.

23. Spender JC: Making knowledge the basis of a dynamic theory of the firm. Strategic Management Journal 1996; 17**:** 45-62.

24. Kogut B, Zander U: What Firms Do? Coordination, Identity, and Learning. Organization Science 1996; 7(5):502-518.

25. Adler PS: Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism. Organization Science 2001; 12(2):215-234.

26. Goshal S, Moran P: Value creation by firms. London Business School, SRLP Working Paper 11/1996, 1996.

27. Bogaert I, Martens R, Van Cauwenbergh A: Strategy as a situational puzzle: the fit of components; in Hamel G, Heene A (eds) : Competence-Based Competition, Chichester, Wiley, 1994.

28. Teece DJ, Pisano GP, Shuen A: Dynamics capabilities and strategic management. Strategic Management Journal 1997; 18(7):509-533.

29. Teece DJ, Pisano GP, Shuen A: Dynamic Capabilities and Strategic Management; in Dosi G, Nelson RR, Winter SG (eds); The Nature and Dynamics of Organizational Capabilities, Oxford University Press, 2000.

30. Pisano GP: Knowledge, integration and the locus of learning: an empirical analysis of process development. Strategic Management Journal 1997; 15:85-100,

31. Pisano GP: In search of dynamic capabilities; in The Nature and Dynamics of Organizational Capabilities; in Dosi G, Nelson RR, Winter SG (eds); The Nature and Dynamics of Organizational Capabilities, Oxford University Press, 2000.

32. Levinthal DA: Strategic management and the exploration of diversity; in Montgomery CA (ed): Resource-Based and Evolutionary Theories of the Firm. Towards a Synthesis, Kluwer Academic Publishers, 1995.

33. Simon HA: A behavioral model of rational choice. The Quarterly Journal of Economics 1955; 69:99-118.

34. Simon HA: Information processing in computer and man. American Scientist 1964, 52(3).

35. Axtell RL: The Emergence of Firms in a Population of Agents, Working paper 99-03-019, Santa Fe, New Mexico, Santa Fe Institute, 1999. www.brook.edu/es/dynamics/papers.

36. Axelrod R: Advancing the Art of Simulation in the Social Sciences; in Conte R, Hegselmann R, Terna P (eds): Simulating Social Phenomena, LNEMS 456, Berlin, Springer, 1997.

37. Hales D: Evolving Specialisation, Altruism and Group-Level Optimisation Using Tags; in Sichman JS, Bousquet F, Davidsson P (eds): Multi-Agent-Based Simulation II, Lecture Notes in Artificial Intelligence 2581, Springer, pp. 26-35, 2002. http://www.davidhales.com.

38. Epstein JM, Axtell R: Growing Artificial Societies: Social Science from the Bottom Up, Cambridge, Massachusetts, MIT Press, Washington, D.C, Brookings Institution Press, Brookings Press, 1996.

39. Hales D, Edmonds B, Norling E, Rouchier J (eds): Multi-Agent Based Simulation III, Proceedings of the 4th International Workshop, MABS 2003, Springer, Lecture Notes in Artificial Intelligence, 2927, Melbourne, Australia, 2003.

40. Doeringer P, Piore M: Internal Labor Markets and Manpower Analysis; Lexington, Mass, Heath, 1971

41. Piore MJ: Fragments of a "Sociological" Theory of Wages. The American Economic Review 1973; 63(2):377-384.

42. Hannan M, Freeman J: The Population Ecology of Organizations. American Journal of Sociology 1977; 82(5):929-964.

43. Hannan M, Freeman J: Structural Inertia and Organizational Change. American Sociological Review 1984, 49:149-164.

44. Hannan M, Freeman J: Organizational Ecology, Cambridge, Mass., Harvard University Press, 1989.

45. Miner A: Idiosyncratic jobs in formal organizations. Administrative Science Quarterly 1987; 32:327-351.

46. Milgrom P, Roberts J; Economics, Strategy & Management, Englewood Cliffs, NJ, Prentice-Hall, 1992.

47. Weaklien DL: (1989) The Employment Contract: A Test of Transaction Cost Theory. Sociological Forum, 1989; 4(2):203-266.

48. Greer CR, Ireland TC: Organizational and financial correlates of a "Contrarian" human resource investment strategy. Academy of Management Journal 1992; 35:956-984

49. Davis-Blake A, Uzzi B; Determinants of Employment Externalization: A Study of Temporary Workers and Independent Contractors. Administrative Science Quarterly 1993; 38(2):195-223.

50. Brusoni S, Prencipe A, Pavit K: Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make?. Administrative Science Quarterly 2001; 46:597-621.

51. Moch MK, Morse EV: Size, centralization and organizational adoption of innovation. American Sociological Review 1977; 42:716-725.

52. Haveman HA: Organizational size and change: diversification in the savings and loan industry after deregulation. Administrative Science Quarterly 1993; 38:20-50.

53. Farber HS: Alternative and Part-Time Employment Arrangements as a Response to Job Loss. Journal of Labor Economics 1999; 27(4), part2:S142-S169.

54. Jenkins GP, Chun-Yan Kuo: On Measuring the Social Opportunity Costs of Permanent and Temporary Employment. The Canadian Journal of Economics 1978; 11(2):220-839.

55. Jovanovic B, Firm-specific capital and turnover. Journal of Political Economy 1979; 87:1246-1260.

56. Cyert RM, Feigenbaum EA, March JM: Models in a behavioral theory of the firm. Behavioral Science 1950: 4.

57. Cohen MD, March JM, Olsen HP: A garbage can model of organizational choice. Administrative Science Quarterly 1972; 17:1-25.

58. Cyert RM, March JM: A Behavioral Theory of the Firm, New Jersey, USA, Prentice-Hall, Inc., 1963.

59. Lomborg B: Nucleus and shield: the evolution of social structure in the iterated prisoner's dilemma. American Sociological Review 1996; 61:278-307.

60. Burgelman RA, Mittman BS: An intraorganizational ecological perspective on managerial risk behavior, performance, and survival: individual, organizational, and environmental effects; in Baum JAC, Singh JV (eds): Evolutionary Dynamics of Organizations, Oxford University Press, 1994.

61. Mezias SJ, Glynn MA: The three faces of corporate renewal: institution, revolution, and evolution. Strategic Management Journal 1993; 14:77-10.

62. Lant TK, Mezias SJ: An organizational learning model of convergence and reorientation. Organization Science 1992; 3(1).

63. Sastry MA: Problems and paradoxes in a model of punctuated organizational change. Administrative Science Quarterly 1997; 42:237-275.

[1] Simulation studies have a long tradition in organisational research, dating back to the seminal works in the area of the behavioural theory of the firm and organisational decision theory [56] some of the most important theoretical pieces in the theory of the firm and organisational theory are based on simulations studies. This is true, for example, for the "Garbage Can" model [57] and for the work leading to the development of "The Behavioral Theory of the Firm" [58]. More recently, simulations have characterised studies in organisational evolution and dynamics, and, in particular, inter-organisational evolution [59], intra-organisational evolution [60] and organisational change [61][62][63].

Behaviour as a Complex Adaptive System: On the role of Self-Organization in the Development of Individual and Collective Behaviour

Stefano Nolfi

Institute of Cognitive Sciences and Technologies
National Research Council (CNR)
Via S. Martino della Battaglia, 44, Rome, Italy
stefano.nolfi@istc.cnr.it

Short Title: behaviour as a Complex Adaptive System

Corresponding Author:
Stefano Nolfi
Institute of Cognitive Sciences and Technologies
Consiglio Nazionale delle Ricerche (CNR-ISTC)
Via S. Martino della Battaglia, 44, 00185, Roma, Italy
voice: ++39-06-44595233
fax: ++39-06-44595243
e-mail:stefano.nolfi@istc.cnr.it

Keywords: behavioural systems, self-organization, adaptive systems

Abstract

In this paper we will explicit the complex system and adaptive nature of behaviour. The complex system nature of behaviour derives from the fact that behaviour and behavioural properties are phenomena that occur at a given time scale and result from several non-linear interactions occurring at a smaller time scale. Interactions occur in time (i.e. consists of a sequence events in which future interactions are constrained by preceding interactions) and might eventually consists of a vector of concurrent interactions. Moreover we argued that behaviour might involve several emergent dynamical processes, hierarchically organized, that affect each others bottom-up and top-down. The adaptive system nature of behaviour derives from the fact that, due to the very indirect relationship between the properties of the interacting elements and the emergent results of the interactions, behavioural systems can hardly be designed while they can be effectively developed through self-organizing methods in which properties emerging from interactions can be discovered and retained through an adaptive process based on exploration and selection. These two claims will be demonstrated in two concrete examples involving mobile robots in which non-trivial individual and collective behaviours have been developed through an evolutionary technique.

1. Introduction

A new research paradigm, that has been called Embodied Cognitive Science [1-4], has recently challenged the traditional view according to which intelligence is an abstract process that can be studied without taking into consideration the physical aspects of natural systems. In this new paradigm, researchers tend to stress (1) situatedness, i.e., the importance of studying systems that are situated in an environment [2-3], (2) embodiment, i.e., the importance of study systems that have bodies, receive input from their sensors and produce motor actions as output [2-3], and (3) emergence, i.e. the importance of viewing behaviour and intelligence as the emergent result of fine-grained interactions (i.e. interactions that occur at small time scales) between the control system of an agent including its constituents parts, the body structure, and the environment. An important consequence of this view is that the agent and the environment constitutes a single system, i.e. the two aspects are so intimately connected that a description of each of them in isolation does not make much sense [5-7].

In section 1 we clarify why behaviour is a complex adaptive system and we discuss how behavioural systems can be developed. After discussing the advantages of self-organizing over design methods, we present two concrete example of effective and robust behavioural system developed through a self-organizing method based on artificial evolution. The first example concerns the development of the control system for an artificial finger that should be able to discriminate objects' shape on the basis of tactile information (section 2). The second example involve the development of the control system of a group of physically assembled robots that should produce coordinated behaviours (section 3). In section 4, we point the hierarchical organization of behaviour. Finally, in section 5, we draw our conclusions.

1.1 Behaviour as a dynamical process resulting from sequences of fine-grained interactions

Behaviour is a dynamical process resulting from the non-linear interactions between an agent (natural or artificial), its body, and the external environment (including the social environment). As we will

see, this implies that behavioural systems (such us mobile robots): (1) are extremely difficult to design from the perspective of an external observer, and (2) can be effectively developed through self-organizing methods (e.g. evolutionary methods) that allow to discover and retain useful behavioural properties emerging from the interactions between agents, their bodies, and the environment.

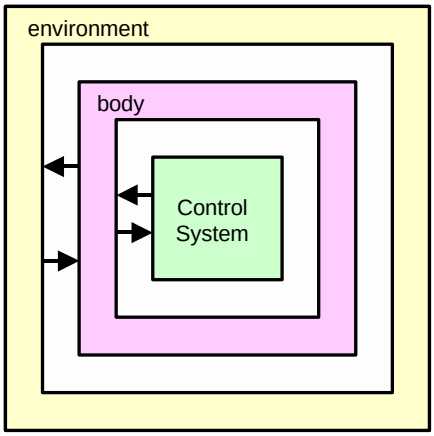


Figure 1. Individual behaviour results from fine-grained non-linear interaction between the agent's control system, its body, and the environment.

At any time step, the environmental structure and the agent/environmental relation influence the body and the motor reaction of the agent that in turn influences the next environmental structure and/or the agent/environmental relation (see Figure 1). Sequences of these form of fine-grained interactions lead to a dynamical process – the behaviour – in which the contributions of the different aspects (i.e. the agent, the body, and the environment) cannot be separated. This implies that even a complete knowledge of the elements governing the interactions provides little insights on the behaviour emerging from the interactions [5-6]. Please notice that we will use the term 'emergence' to indicate a property resulting from a sequence of interactions that can hardly be predicted or inferred from an external observer even on the basis of a complete knowledge of the interacting elements and of the rules governing the interactions.

The relation between the interaction rules and the resulting behaviour is further complicated by the fact that, when interactions are non-linear, small variations at the levels of the rules governing the interactions might translate to very different forms of behaviour due to cumulative and amplifying effects.

1.2 On the advantages of self-organizing over design techniques

From a theoretical point of view, the complex adaptive system nature of behaviour has several important consequences that are far from being fully understood. One important aspect, for instance, is the fact that motor actions partially determine the sensory pattern that agents receive from the environment. By coordinating sensory and motor processes organisms can select favourable sensory patterns and thus enhance their ability to achieve their adaptive goals [8-11].

From a engineering point of view, the complex adaptive system nature of behaviour explains why methods based on explicit design are inadequate for developing behavioural systems and why self-organizing methods (e.g. methods based on evolutionary techniques) might be appropriate instead.

The inadequacy of design methods lay on the fact that they require from the designer an ability to infer the rules governing the interactions between the agent and the environment that will lead to a desired behaviour. Unfortunately, as we pointed out above, the properties of the behaviour that emerges from a sequence of fine grained non-linear interactions between the agent and the environment can hardly be inferred from the structure of the interacting elements and the rules governing the interactions. The inverse problem faced by the designer (i.e. the problem of determining the rules governing the interaction that will lead to a desired behaviour) is at least equally hard.

The advantage of self-organizing methods is indeed the fact that they do not require to identify the relation between the rules governing the interactions and the resulting behaviour. They are based on an evolutionary and/or learning process in which the rules governing the interactions, initially randomly assigned, are progressively modified through a process of random variation and selection. Algorithms with this property include evolutionary, simulator annealing, and reinforcement learning algorithms when: (a) the rules governing the interaction are encoded in free parameters, and (b) variations of free parameters are retained or discarded on the basis of variation of performance observed at the behavioural level (i.e. at the time scale of seconds or more). These characteristics allow these methods to discover and retain useful properties emerging from the several interactions without the need to identify the relation between the rules governing the interaction (and/or the interacting elements) and the resulting behaviour.

The possibility to discover and retain useful properties emerging from the interactions also allow self-organizing methods to come up with solutions that are simple from the point of view of the interaction rules (for examples, see [8-9]). Indeed, while in design methods the effects of the detailed characteristics of the agent and the environment (i.e. inertia, elasticity of materials, detailed characteristics of the shape etc.) cannot be predicted and thus constitute problems to be avoided, in self-organizing methods they constitute possibilities to be exploited.

Two example of how self-organizing methods might be used to develop effective behavioural system and to exploit properties emerging from the interactions will be presented in section 2 and 3.

1.3 Collective behaviour emerge from a large number of interactions

Collective behaviour is a dynamical process resulting not only from the fine-grained interactions between agents, their bodies, and the external environment but also between agents (see Figure 2).

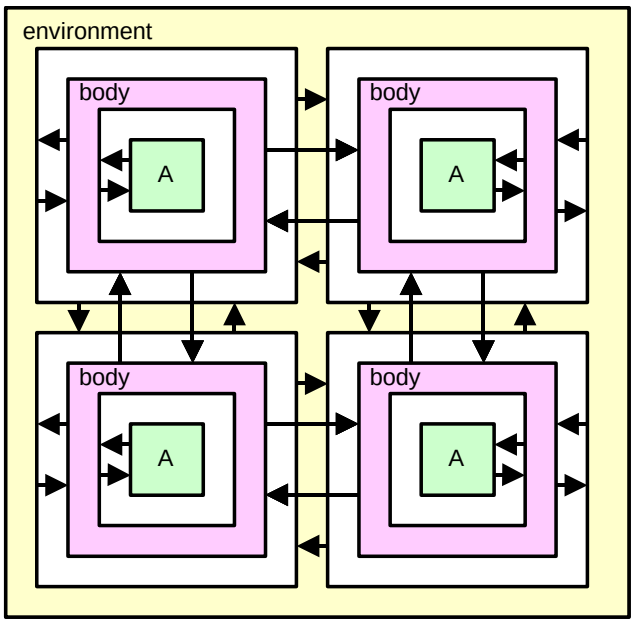


Figure 2. Collective behaviour results from a sequence of several concurrent interactions occurring between agents, their body, and the environment and between agents. The figure schematically represents the situation of four agents in which each individual interacts with two adjacent individuals directly (through physical contact) and indirectly (through environmental modifications that affect other agents' sensors).

The fact that collective behaviour results from a much larger number of fine-grained interactions implies that the relation between the rules governing the interactions and the resulting behaviour is more indirect and more difficult to infer than in the case of individual behaviour. In fact, (a) individual behaviour might be hard to infer or predict on the basis of the rules governing the interactions between

the agents, their body, and the external environment (see previous section), (b) groups' aggregate-level behaviour might be hard to infer or predict on the basis of individual behaviours, and (c) the effects of group level dynamics on individual behaviour might be hard to infer or predict. For these reasons, the problem of designing the interaction rules that lead to a desired collective behaviour might be extremely hard even in simple cases [12-13].

As we mentioned above, however, the indirect relation between the rules governing the interactions and the resulting collective behaviour does not constitute a problem for self-organizing methods. On the contrary the large number of interactions might increase the possibility to identify parsimonious solutions (from the point of view of the complexity of the rules governing the interactions) by exploiting useful behavioural properties emerging from the interactions.

An example of how self-organizing methods might be used to develop effective and robust collective behaviours will be presented in section 3.

2. Evolving the control system of an artificial finger able to discriminate objects with different shapes on the basis of tactile information.

Consider the case of a robot with an artificial finger that has to discriminate objects with different shapes on the basis of rather rough tactile information [10].

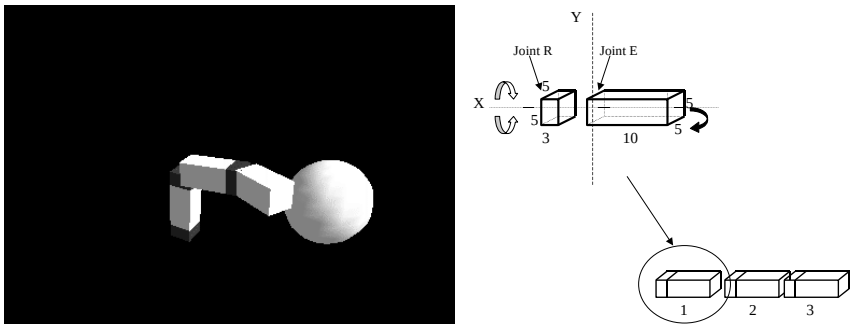


Figure 3. **Left:** The artificial finger and a spherical object. **Right:** A schematic representation of the finger.

The artificial finger consists of 3-segments with 6 degrees of freedom (DOF) and coarse touch sensors (see Figure 3, left). More precisely, the artificial finger consists of a basic structure of two bodies and two joints replicated for three times (see Figure 3, right). These two bodies are connected by means of a joint (i.e. the *Joint E* in Figure 3, right) that allows only one DOF on axis *Y* , while the shorter body is connected at the floor, or at the longer body, by means of a joint (i.e. the *Joint R*) that provides one DOF on axis *X*. In practice, the *Joint E* allows to elevate and to lower the connected segments and the *Joint R* allows to rotate them in both direction. *Joint E and Joint R* are free to moves only in a range between [0 and $\pi/2$] and [$-\pi/2$, $+\pi/2$], respectively. Each actuator is provided with a corresponding motor that can apply a varying force. Therefore, to reach every position in the environment the control system has to appropriately control several joints and to deal with the constraints due to gravity (collisions and physical dynamics was carefully simulated on the basis of Vortex^{TM} libraries).

The sensory system consists of three simple contact sensors placed on each longer body that detect when these bodies collides with obstacles or other bodies and six proprioceptive sensors that provide the current position of each joint. The motor system consists of six motors controlling the corresponding six DOF.

The controller of each individual consists of a neural network with 10 sensory neurons directly connected to 7 motor neurons and 2 internal neurons receiving connections from the sensory neurons and from themselves and projecting connections to the motor neurons. The first 9 sensory neurons encode the angular position (normalized between 0.0 and 1.0) of the 6 DOF of the joints and the state of the three contact sensors located in the three corresponding segments of the finger. The last sensory neuron is a copy of the last motor neuron that encodes the current classification produced by the indi-

vidual (see below). The first 6 motor neurons control the actuators of the 6 corresponding joints. The output of the neurons is normalized between [0, $+\pi/2$] and [$-\pi/2$, $+\pi/2$] in the case of elevation and rotational joints respectively and is used to encode the desired position of the corresponding joint. The motor is activated so to apply a force proportional to the difference between the current and the desired position of the joint. The seventh motor neuron encodes the categorization output (value below or above 0.5 are interpreted as classifications corresponding to a cubic or spherical object respectively).

The connection weights of the neural controllers were evolved. An initial population of different artificial genotype, each encoding the connection weights of a corresponding neural controller, is created randomly. Each connection weight was represented in the genotype by eight bits that were transformed into a number in the interval [–10, +10]. Each robotic finger is then allowed to interact with the environment on the basis of a corresponding, genetically specified, neural controller. The fittest robots are allowed to reproduce by generating copies of their genotypes with the addition of changes (random mutations). This process is repeated for a number of generations.

Evolving individuals were allowed to "live" for 36 epochs, each epoch consisting of 150 actions. At the beginning of each epoch the finger is fully extended and a spherical or a cubic object is placed in a random selected position in front of the finger (the position of the object is randomly selected between the following intervals: 20.0 >= X <= 30.0; 7.5 >= Y <= 17.5; -10.0 >= Z <= 10.0). The object is a sphere (15 units in diameter) during even epochs and a cube (15 units in side) during odd epochs so that each individual has to discriminate the same number of spherical and cubic objects during its “lifetime” Fitness is computed by counting the number or epochs in which individuals correctly categorize the object (i.e. the number of times in which at the end of the epoch the activation of the last motor units is below 0.5 and the object is a cube or is above 0.5 and the object is a sphere). Therefore, individuals are free to determine how to interact with the objects, i.e. the are only selected on the basis of the ability to correct categorizations.

Population size was 100. The best 20 individuals of each generation were allowed to reproduce by generating 5 copies of their genotype with 1% of their bits replaced with a new randomly selected value.

By running 10 replications of the experiment and by evolving individuals for 50 generations we observed that in many of the replications evolved individuals display a good ability to categorize objects and, in some cases, produce close to optimal performance. Figure 4 shows how a typical evolved individual behave with a spherical and a cubic object (left and right sides of the Figure, respectively). As can be seen, first the finger bends on the left side and move to the right so to start to feel the object with the touch sensor of the third segment. Then the finger continues to move on the same direction by slightly moving up when the third segment of the finger touches the object. As a result of this simple motor rules, in the case of spherical objects, the finger keeps moving toward the left side following the curvilinear surface. In the case of cubic objects, instead, it remains stuck in one of the angles by moving back and fourth.

The behaviour emerging from the interactions between the finger and the objects lead to two rather different behavioural outcomes in the case of spherical and cubic objects: (a) a fully extended position of the finger in the case of spherical objects, and (b) a fully bended position of the finger, in the case of cubic objects. These two positions, in turn, provide a straightforward indication of the type of object the finger interacted with. For other example, involving different environment and robots with different morphologies, in which the convergence or the luck of convergence on a limit cycle behaviour can be used to categorize the environment, see [8-10]).

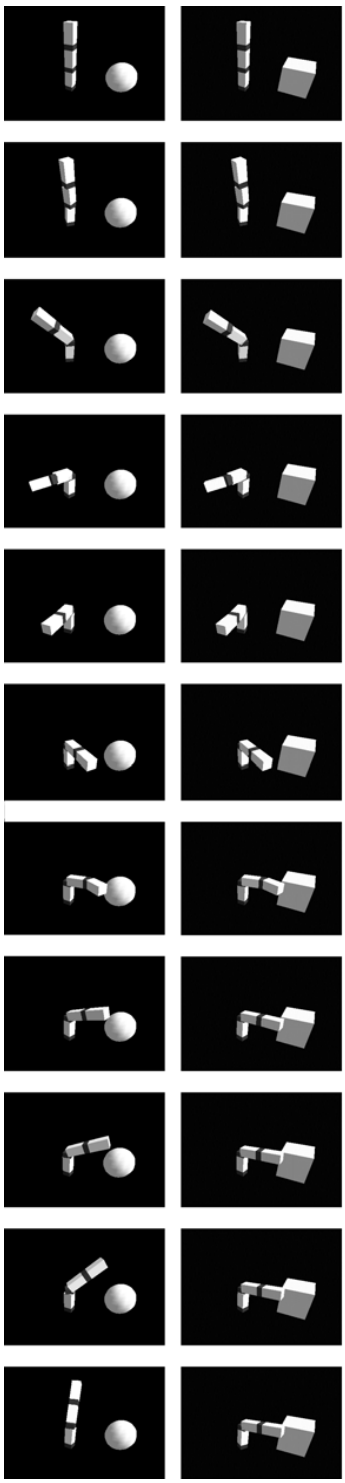


Figure 4. Behaviour of a typical evolved individual during an epoch (150 cycles) in which the object consists of a sphere (left pictures) and of a cube (right pictures). For reason of space, the pictures show the position of the finger each 15 cycles.

Individuals of other replications of the experiments display similar behaviour although the length of the phase with which individuals interact with spherical objects before leaving them varies. The fact that the best performance are observed in cases in which the interaction phase lasts longer (result not shown), demonstrates that the discrimination process is not the result of a single decision but rather the

end result of a sequence of interactions between the finger and the object. A similar temporally-extended decision process has been observed in a different experimental set-up in which evolving agents are asked to catch and avoids objects with different shapes [11].

3. Evolving the control system of a collection of physically assembled robots able to display coordinated collective behaviour

Consider the case of four assembled robots forming a linear structure (Figure 5) that should move and reach a light target [14]. Given that the orientations of individual robots might vary and given that the target might be out of sight, robots should be able to coordinate to choose a common direction of movement and to change their direction as soon as one or few robots start to detect a light gradient.

Figure 5. **Left:** Four robots assembled into a linear structure. **Right**: A simplified simulation of the robots described in the left part of the figure based on Vortex™ libraries.

Each robot [15] consists of a mobile base (chassis) and a main body (turret) with a diameter of 116 mm that can rotates with respect to the chassis along the vertical axis. The chassis has two drive mechanisms that control the two corresponding tracks and teethed wheels. The turret has one rigid and one flexible gripper, that allow robots to assemble together and to grasp objects, and a motor controlling the rotation of the turret with respect to the chassis. Robots are provided with a traction sensor, placed at the turret-chassis junction, that detects the intensity and the direction of the force of traction that the turret exerts on the chassis (along the plane orthogonal to the vertical axis) and light sensors. The robots also have several other sensors (a sound sensors, an omnidirectional camera, accelerometers etc.) that, however, were not used in the experiments reported below.

Robots' controller only have access to local sensory information. In particular, each robot's controller consists of a neural network with nine sensory neurons directly connected to two motor neurons. The first four sensory neurons encoded the intensity of the traction from four different orientations with respect to the chassis (rear, left, front and right). The next four sensory neurons provide information on the light gradient with respect to the chassis. The last neuron consists of a bias unit that is always activated to 1.0. The activation state of the two motor neurons was normalized within [–5, +5] rad/s and was used to set the desired speed of the two corresponding wheels and of motor controlling the degree of freedom between the turret and the chassis. The initial population consisted of 100 randomly generated genotypes that encoded the connection weights of 100 corresponding neural controllers. Each connection weight was represented in the genotype by eight bits that were transformed into a number in the interval [–10, +10]. Each genotype encoded the connection weights of a corresponding neural controllers that was then duplicated four times and embodied into the four robots forming the team (i.e. the team is homogeneous).

By evolving the connection weights of the robots' controller and by selecting the team of four robots on the basis of the distance travelled from its initial position (when the light target was not on sight) and for the distance travelled toward the target light (when the light target was on sight) we observed that evolving individual are able to effectively solve their problem by negotiating a common

direction of movement and by collectively moving toward the light as soon as a light gradient can be detected.

By testing evolved controllers in different conditions we surprisingly observed that they are able to generalize their abilities in new conditions and also to spontaneously produce new unexpected behaviours. More precisely, evolved robots display a capacity to generalize their abilities to: (a) the number of assembled robots, (b) the shape with which robots are assembled together, and (c) the use of flexible rather than rigid links. Moreover, evolved robots also display an ability to: (a) spontaneously produce a collective obstacle avoidance behaviour, (b) dynamically rearrange the physical shape of the team in interaction with the environment to negotiate narrow passages, (c) spontaneously produce a coordinate object pushing/pulling behaviour when assembled to or around an external object.

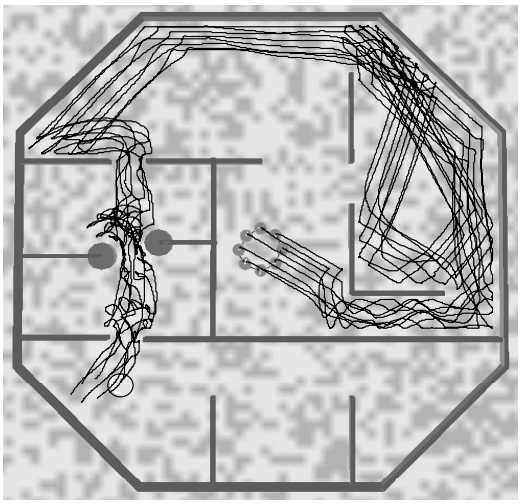


Figure 6. A circular shape structure formed by eight robots assembled through flexible links in a maze with obstacles consisting of walls and cylindrical objects (represented with grey lines and circles). The team of robots starts in the central portion of the maze and reach the light target located in the bottom-left side of the environment (see the light grey circle) by exhibiting a combination of collective obstacle avoidance and collective light approaching behaviour. The irregular lines, that represent the trajectories of the individual robots, provide an indication of how the shape of the assembled robots changes during motion by adapting to the local structure of the environment.

Figure 6, shows the behaviour displayed by eight robots assembled into a circular shape through flexible links (i.e. links that allow two connected robots to modify their relative positions within certain limits) placed in a maze environment with walls and cylindrical obstacles. As shown in the figure the same control system evolved to control four robots assembled into a linear structure generalize to: (1) a team consisting of eight robots assembled into a different shape, (2) robots assembled through flexible links that modify the shape of the assembled structure during motion. The figure also show how robots: (a) produce a collective obstacle avoidance behaviour (as a result of the traction force generated during collisions with obstacles), and (b) rearrange the shape of the team to pass narrow passages.

Figure 7, that shows the behaviour of how 8 robots assembled through flexible links around a cylindrical object. As shown on the figure, the same control system evolved to control four robots assembled into a linear structure generalizes in new conditions and display a coordinate object pushing-pulling behaviour.

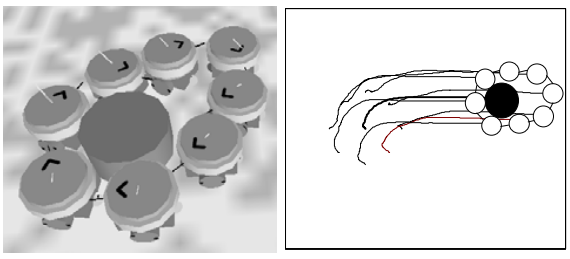


Figure 7: **Left**: Eight robots connected around an object. **Right**: Coordinated object pushing/pulling behaviour exhibited by a team of robots assembled around an external cylindrical object. The empty circles and the full circle indicate the final positions of the robots and of the object, respectively. The thin lines and thick line indicate the trajectory of the robots and of the object.

For a demonstration of how the neural controller evolved in simulation are able to display similar behaviours when embodied and tested in the real physical robots see [16].

4. Behaviour as dynamical system organized hierarchically

In the introduction we pointed out that behaviour is a dynamical process emerging from the interactions between the agents' control systems, the agents' body, and the external environment (eventually including the social environment). The fact that behaviour (even in simple cases such us grasping an object or reaching a target location) is a property that can be observed only at macro time scale (in the range of seconds or minutes) while interactions occur at micro time scales (milliseconds) imply that behaviour emerge from a large number of non-linear interactions not only in the case of collective behaviour but also in the case of individual behaviour. Behaviour is always the result of a sequence of fine-grained interactions (distributed in time) and eventually of a number of concurrent interactions between different agents (distributed in space). Overall this implies that both individual behaviour and collective behaviour are the emergent result of a large number of fine-grained interactions. Although this fact is widely recognized in the case of collective behaviour, it is much less recognized in the case of individual behaviour.

The picture is further complicated by the fact that behaviour might be based on a series of emergent dynamical processes, hierarchically organized, that affect each others bottom-up and top-down (for a similar view, see [17]). More precisely: (a) interactions between properties emerging from a sequence of fine-grained interactions might lead to higher level emergent properties (that typically extend over larger time scales than the interacting properties), (b) higher level properties might affect the interactions between lower level properties.

As an example of a top-down effects of high level properties (emerging from the interaction between the agent and the environment) and the interaction between the agent and the environment consider the example of the discrimination behaviour described in section 2. The behavioural properties emerging from the interactions between the agent control system, its body, and the external object (occurring at a time scale of 100ms) result in two different emerging behaviours (in the case of cubic or spherical objects respectively): (1) the finger remains bended and keeps touching the object, or (2) the finger becomes fully extended by passing over the object. These two emergent properties occur at a time scale of seconds while the interaction between the agent and the environment are mediated by control rules that operate at the time scale of milliseconds.. These two high level properties, in turns, affect the successive lower level interactions mediated by the agent neural controller (i.e. the neural controller produces a categorization output corresponding to "cubic object" or "spherical object" on the basis of the state of the sensors that detect the current angular position of the joints of the finger).

As an example of behaviours organized in three hierarchical levels and in which level 3 properties emerge from the interaction between level 2 properties, that in turn emerge from the interaction between the agent and the environment, consider the case of the collective navigation problem described in section 3. Interactions occurring between the agents and the environment (at a time scale of 100ms) lead to two behavioural properties (that extend at a time scale of seconds): (1) an ability to negotiate and converge on a common direction of movement, and (2) an ability to turn toward the light. The interactions between these two high level properties, in turn, lead to several collective behaviours that occur at larger time scales (i.e. several seconds). More precisely, the interaction between these two behavioural capacity lead to: (a) an ability to collectively approach the light target (even when only few agents detect the light because of their relative distance with respect to the light or because of shadows), (b) an ability to display a collective exploration behaviour and a collective light approaching behaviour, and (c) an ability to combine the two behaviours by avoiding to get stuck in situations

in which these two behavioural capacity, by triggering opposite motor responses, might interfere one with the other.

5. Conclusion

In this paper we pointed out the complex system and adaptive nature of behaviour.

The complex system nature of behaviour derives from the fact that (both in the case of individual and collective behaviour) behaviour and behavioural properties are phenomena that occur at a given time scale and result from several non-linear interactions occurring at a smaller time scale. Interactions occur in time (i.e. consists of a sequence events in which future interactions are constrained by preceding interactions) and might eventually consists of a vector of concurrent interactions. Moreover we argued that behaviour might involve several emergent dynamical processes, hierarchically organized, that affect each others bottom-up and top-down.

The adaptive system nature of behaviour derives from the fact that, due to the very indirect relationship between the properties of the interacting elements and the emergent results of the interactions, behavioural system can hardly be designed while can be effectively synthesized on the basis of a self-organization process (in which properties emerging from interactions can be discovered and retained through an adaptive process based on exploration and selection).

These two claims have been demonstrated in two examples in which non-trivial individual and collective behaviour have been developed through an evolutionary technique.

Acknowledgements

This research has been partially supported by the ECAGENTS project founded by the Future and Emerging Technologies program (IST-FET) of the European Community under EU R&D contract 001940. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that may be made of data appearing in this publication.

References

1. Varela F.J., Thompson E. and Rosch E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
2. Brooks R.A. (1991) Intelligence without reason. In J. Mylopoulos & R. Reiter (Eds.), *Proceedings of 12th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann.
3. Clark A. (1997) *Being There: Putting Brain, Body and World Together Again.* Cambridge, MA: MIT Press.
4. Pfeifer R. and Scheier C. (1999) *Understanding Intelligence*. Cambridge, MA: MIT Press
5. Maturana H.R. & Varela F.J. (1980). *Autopoiesis and cognition: the realization of the living*. Dordrecht: reidel.
6. Maturana H.R. & Varela F.J. (1988). *The three of knowledge: the biological roots of human understanding*. Boston: New Science Library.
7. Beer R.D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72:173-215.
8. Nolfi S. (2002). Power and limits of reactive agents. *Neurocomputing*, 49:119-145.
9. Nolfi S. (2005). Categories formation in self-organizing embodied agents. In H. Cohen & C. Lefebvre (Eds), *Handbook of Categorization in Cognitive Science*. Elsevier.
10. Nolfi S., Marocco D. (2002). Active perception: A sensorimotor account of object categorization. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, J-A. Meyer (eds.) *From Animals to Animats 7: Proceedings of the VII International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press, pp. 266-271.

11 Beer, R.D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209-243.
12. Baldassarre G., Nolfi S., Parisi D. (2003). Evolving mobile robots able to display collective behavior. *Artificial Life*, 9: 255-267.
13. Funes P., Orme B. & Bonabeau E. (2003). Evolving emergent group behaviors for simple humans agents. In P. Dittrich & J. T. Kim (Eds.) Proceedings of the Seven European Conference on Artificial Life (ECAL 2003).
14. Baldassarre G., Parisi D., Nolfi S. (in press). Distributed Coordination of Simulated Robots Based on Self-Organisation, *Artificial Life*.
15. Mondada F., Pettinaro G., Guigrard A., Kwee I., Floreano D., Denebourg J-L & Nolfi S., Gambardella L.M., Dorigo M. (2004). Swarm-bot: A new distributed robotic concept. *Autonomous Robots*. vol. 17 (2-3) : 193-221.
16. Baldassarre G., Trianni V., M. Bonani, Mondada F., Dorigo M. & Nolfi S. (2005). Coordinated motion and in teams of physically linked robots. *Technical Report*. Roma: Institute of Cognitive Science and Technology, CNR.
17. Keijzer F. (2001). *Representation and Behavior*. Cambridge, MA: MIT Press.

Optimization and control of the urban spatial dynamics

Ferdinando Semboloni [1]

Department of Town and Regional Planning, and
Center for the Study of Complex Dynamics,
University of Florence

Short title: Optimization and control

Keywords: Optimization; Control of complex systems; Urban dynamics; Urban planning

Abstract

The urban planning concerns the assignment of a land use to each land cell. This process coexists and may conflict with the complex self-organized dynamics of the urban system, which should be constrained by the plan. The purpose of this study is the identification of a method for attaining the planned goals through the utilization of the self-organized dynamics, and the minimization of the constraints. The paper is organized in three steps. First: the urban plan is defined as a process of optimization. Second: the set of optimal solutions is compared with the configurations resulting by the self-organized dynamics. Third: a method for the convergence of the self-organized dynamics with the optimal configuration is proposed. In conclusion the study shows that planning a complex system may be an hard task, while the control and the utilization of the self-organized dynamics helps in the attainment of a total utility.

Introduction

Two main streams of problems arise for urban strategic planning from the widely recognized self-organizing character of the urban dynamics [1]. From one side the chaotic behavior highly dependent on initial conditions of the self-organizing system makes quite unpredictable the effects of the planning policy [2]. On the other side the urban planning is questioned by the ability of the urban system to steer itself [3]. While classic urban planning seeks to regulate the urban structure with a top-down approach, this last evolves with the interplay of a lot of local actions[4]. Thereby these two processes, even if coexistent in a city, may conflict. One solution to this topical interest problem in urban planning is a just-in-time method, as opposed to a just-in-case, where projects are delimited in both space and time and coupled with a constant reevaluation of the whole sketch [5].

In this paper we choose a different way. We suggest the utilization of this self-organizing character in order to achieve in an easier way the planning objectives. The proposed method includes: first the definition of the plan as an optimization process, second a comparison of the optimal solutions with the configurations emerging from the self-organized dynamics, and third the convergence of the self-organized dynamics with the optimal configuration. In a first step this method is applied to a system with two land uses, and in the second step a more realistic situation is utilized in order to show the possible utilization of the proposed method.

[1]Address: Dipartimento di Urbanistica e Pianifi cazione del Territorio–Via Micheli,2–50121 Firenze–Italy. E-mail: semboloni@urba.arch.unifi .it

Optimization

Even if urban planning is a complex process involving many actors which bargain for urban projects, after all it ends up in the assignment of land uses or activities to land parcels, so that the public utility, however defined, is also attained. The problem of the optimal spatial assignment of activities was formulated by Koopmans and Beckman [6]. Later the optimization process has been proposed in the context of rational urban planning [7][8], as well as in architectonic design [9]. Recently new methods for the research of the global optimum such as genetic algorithm, have been utilized in order to establish the set of efficient alternatives in a multi-objective optimization [10],[11]. This multi-objective optimization method refers to a theory of the urban planning in which a set of efficient alternatives, or Pareto optima, are proposed to the decision makers. In other words the optimization method is utilized as an aid to the decision process, and the efficient solutions are taken as alternatives to be bargained among the decision makers[12].

In order to analyze the relation between the urban plan and the self-organized dynamics, we consider firstly the plan as an optimization process. The land-use planning is an assignment of land uses to land parcells in order a total utility function in which the spatial relations among land-uses are included, is maximized. To formalize the problem, let us suppose a set of possible land uses or states of a cell i of a squared grid having n cells in each side. Hence $s_i = \{h, r, ...\}$, the state of cell i, represents the land use assigned to a cell i (for instance housing, retail etc.) and N_h is the total number of cells in state h which is established as a constraint to the optimization. In addition, for each cell i is established a neighborhood Φ_i, and an utility u_{hi} of land use h in cell i which depends on h, and on the land uses localized in the Φ_i neighborhood, as in the following equation:

$$u_{hi} = f(h, s_{\Phi_i}), \tag{1}$$

where u_{hi} is the utility of land use h in cell i, and f is a function. This function may take many forms. In this paper it has been chosen the simplest, i.e. a linear combination, which simulates a perfect substitution of factors. Other functions with imperfect substitution of factors giving different results, are not analyzed in this paper. According to the linear combination method, the utility is calculated as in the following expression:

$$u_{hi} = \sum_r \left(m_{hr} \sum_{j \in \Phi_i} r_j \right), \tag{2}$$

where $r_j = 1$ if the use r is assigned to cell j, otherwise $r_j = 0$. Further m_{hr}, an element of the matrix $\mathbf{M}$, is the utility for use h to be surrounded by use r ($\sum_r m_{hr} = 1$). In words, equation (2) states that utility depends on the quantity of the land uses in the neighborhood multiplied by a parameter. The total utility of each land use is:

$$U_h = \sum_i u_{hi}. \tag{3}$$

Moreover, a weight is assigned to each land use in order to calculate the total utility U_T, as in the following equation:

$$U_T = \sum_h W_h U_h. \tag{4}$$

This W_h weight ($\sum_h W_h = 1$) represents the economic capacity of the land use h to bid for the land, or the importance assigned to this land use in case a plan is implemented. This is an usual assumption in multi-criteria decision making, and it is the simplest method to compare the otherwise incommensurable vectors of the utilities. The planning problem is stated as the research of the assignment which maximizes the total utility, for each possible set of weights. By using this method, under the assumption in the previous equation (4), one obtains the set of Pareto optimal solutions[13]. For each of these solutions it is impossible to improve the utility of a land use U_h without decreasing the utility of another land use U_r.

In the following experiments the maximization of utility is obtained by using the simulated annealing method[14]. Not only this is a general optimization method which can be applied to each sort of problems, but it is formulated as a simulation of a dynamic process which can be easily compared with the self-organized dynamics. When the simulated annealing method is applied, the process begins by using a random pattern of land uses. Then, at each step one couple of cells is taken at random, the land uses are exchanged, and the total utility U_T of the new pattern is evaluated by using equation (4). A new configuration is considered or not, according to the method of simulated annealing which decreases the temperature thus allowing the system to reach and maintain a stable state in which the energy $E = -U_T$ is minimized, and the total utility is maximized.

Nevertheless, the urban system is usually able to reach some stable state by using its self-organizing character. In this case the maximization of the total utility is not a necessary, even if a possible [15] outcome of the process. We call self-organized dynamics this process, which is presented in the following section.

Self-organized dynamics

A lot of studies in the micro-simulation of the urban dynamics are available. Usually a cellular automaton framework is utilized, even if an agent based modeling approach is emerging as a novel method[16]. For the most part these approaches consider an expanding system from a central seed. Since we are more interested in the internal rearrangement of the system then in its growth, a different approach is utilized which is partially similar to that utilized in the simulated annealing method. In fact beginning from a random pattern of the established land uses, at each step two cells are chosen at random and the land uses are exchanged if this exchange does not decrease the sum of the utilities of the two land uses. Thereby the land use h in cell i is exchanged with the land use r in cell j if:

$$u_{ri} + u_{hj} > u_{hi} + u_{rj}. \tag{5}$$

This method simulates a spatial dynamics in which the local individual utility is maximized, as in a market in which a transaction happens only if a couple of individuals thinks that they will be more satisfied after, than before the transaction. In addition this method mimics the efforts for establishing a solution through a lot of repeated trials, like in the simulated annealing method. But, in this case at each exchange, is maximized the sum of the local utilities instead of the total utility U_T which may both increase or decrease because the utility of the surrounding cells is not considered. The stability of this self-organized dynamics is attained when for each couple of cells the exchange does not increase the sum of the utilities of the two land uses.

The Utopian situation corresponds to the similarity between the optimal configuration and the result of the self-organized dynamics. In this case the utilities obtained with the optimization process and with the self-organized dynamics are equivalent in the steady state. This comparison allows to divide the set of Pareto solution in two subset: the A-set of solutions in which the optimum is similar to the outcome of the self-organized dynamics and the B-set for which there is no similarity. This similarity is evaluated comparing the utility attained by the optimal configuration in relation with the utility attained by the self-organized dynamics. Thereby an index of dissimilarity I^d is defined:

$$I^d = \frac{U_T^1 - U_T^2}{U_T^1}, \tag{6}$$

where U_T^1 is the total utility attained with the optimal configuration and U_T^2 is the total utility attained with the self-organized dynamics. This index ranges from zero, because $U_T^1 \geq U_T^2$, to 1.

Convergence of the self-organized dynamics with the optimal configuration

In case the optimal solution and the self-organized dynamics do not match, one would like to make the self-organized dynamics convergent with the optimal configuration. The classic urban planning approaches this

problem by establishing a feedback for each point of the urban surface. In case the land use in an urban zone does not agree with the established use, and this mismatch is officially observed by the control agency, then a sort of penalty, usually established by the law, is applied in order to reestablish the planned assignment. This is the principle of the dynamic system control, in which there is a master system, the plan, and a slave system, the urban dynamics, and each point of the master is coupled with the corresponding point of the slave system. This method simply constraints the self-organized dynamics into the optimal configuration. The energy spent in this process by the control agency depends on the *distance* between the urban plan and the result of the self-organized dynamics. For this reason often an urban plan is conceived in a way that this distance is reduced in order to make the desired configuration more attainable.

In opposition to the classic urban planning, the proposed method is based on the utilization of the self-organizing characteristic of the urban system and on the minimum number of cells whose established land use is forbidden to change during the simulated dynamics. These pinning cells play the role of control but also of catalysts accelerating positive effects in the urban dynamics [17] without itself being transformed. The identification of the minimum number of pinning cells which is able to make the self-organized dynamics convergent with the optimal configuration is the core of the problem. This aspect has been widely studied in the field of the control of chaotic spatio-temporal systems [18], [19], and in essence consists in the identification of the cells which are *strategic* for the control of other cells in the spatial dynamics.

To establish the minimum number of cells able to control the urban dynamics, these are sorted according to the index of influence. The method utilized to calculate the index of influence of a cell is based on a kind of input-output analysis. In fact in the optimal configuration each cell generates and receives utility from the bordering cells. According with equations (2) and (4), this utility v_{ij} generated by the land use r in cell i in relation to the land use h in cell j is defined in the following way:

$$v_{ij} = W_h m_{hr}. \tag{7}$$

By using the previous equation, the matrix $\mathbf{G}$ of the exchanged utility is calculated. Each element g_{ij} of the matrix is the ratio: utility generated by the cell i in the cell j–total utility generated in cell j by the whole set of cells, as in the following expression:

$$g_{ij} = \frac{v_{ij}}{\sum_k v_{kj}}. \tag{8}$$

In order to consider both the direct and the indirect effects the following matrix is calculated:

$$\mathbf{G}^\star = \mathbf{G} + \mathbf{G}^2 + \mathbf{G}^3 + \dots \tag{9}$$

The index of influence I^i for each cell is then calculated as the difference between the capacity to influence ($\sum_j g^\star_{kj}$) minus the degree of being influenced ($\sum_i g^\star_{ik}$):

$$I^i_k = \sum_j g^\star_{kj} - \sum_i g^\star_{ik}. \tag{10}$$

In the first step t of the ordering process, the first pinning cell k is chosen which corresponds to the maximum of I^i. Than we reason in the following way. If the cell k is established as a pinning site, the control of the other cells immediately dependent by k cell is useless. We would rather that the second pinning cell has an high index of influence, while being the less dependent on the first. For this reason we subtract to the capacity of control of each cell l a share, $g^\star_{kl}/\sum_i g^\star_{il}$, corresponding to the share of control exercised by the k cell on cell l. Thereby in the next step $t+1$ the elements of the matrix $\mathbf{G}^\star$ are recalculated in the following way:

$$g^\star_{lj}(t+1) = g^\star_{lj}(t)\left[1 - \frac{g^\star_{kl}(t)}{\sum_i g^\star_{il}(t)}\right]. \tag{11}$$

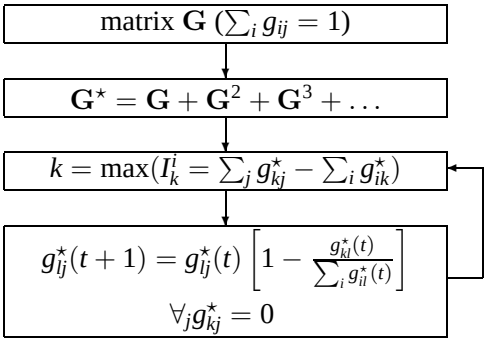


Figure 1. The algorithm for ranking the pinning cells.

Further the row k is set equal zero in order to exclude the cell k to be newly selected as a pinning cell, and the process is repeated until for each cell, the index related to its capacity to influence the land use of the other cells is calculated (figure 1).

Clearly a relation should exist between the number of pinning cells and the similarity between the optimal configuration and that obtained with the self-organized dynamics. In order to analyze this relation, we define D_{pq} as the distance between the configurations p, the plan, and q, the outcome of the self-organized dynamics. This distance is calculated by summing up the cells having a different land use in p and q. By using this distance we consider the following relation:

$$D_{pq} = f(C_q), \tag{12}$$

where C_q is the number of pinning cells, successively taken from the ranked list established by using the previous explained method. Usually D_{pq} decreases with the increase of C_q. In fact in case all the cells are declared as pinning cells, this is like in the classical urban planning, and $D_{pq} = 0$. In turn we can decide the quantity of pinning cells, beginning from the first in the ranked list and evaluate the rate at which the distance decreases, looking for the less number of pinning cells resulting in an acceptable distance. In order to study this and the other aspects in the next section some theoretic experiments are shown which highlight, with an increasing complexity, the relation between optimization and control in the spatial dynamics.

Experiments and discussion

In this section the proposed method is applied to a set of theoretic cases of increasing complexity. In each case, through the variation of the weights the set of optimal solution is obtained. Further the relationship between this set and the corresponding results of the self-organized dynamics is discussed in order to highlight the conditions under which the control of the complex system is the easiest and the most efficient.

In order to study the simplest situation it has been chosen a 10×10 lattice surface and two possible land uses for each cell: the housing (h) and retail (r) land use. In the first three experiments, the total quantities of each land use are constrained to 50. The rules of interaction refer to three paradigmatic cases. *Integration*: each land use is attracted by the other land use. *Segregation*: each land use is attracted by a similar land use. And *integration and segregation*: the housing land use is attracted by the retail land use, while this last is attracted only by itself. The neighborhood Φ_i is limited to the eight bordering cells. The corresponding values s of the elements of the $\mathbf{M}$ matrix are shown in table 1.

The segregation dynamics, has been widely investigated, beginning with the work of Schelling [20], with the purpose of connecting the emergent properties of the resulting pattern with a parameter related

Table 1. Matrix $\mathbf{M}$ in the three cases.

Integration		
Land use	Hous.	Ret.
Housing	0	1
Retail	1	0

Segregation		
Land use	Hous.	Ret.
Housing	1	0
Retail	0	1

Integration-Segregation		
Land use	Hous.	Ret.
Housing	0	1
Retail	0	1

to the degree of segregation. The self-organized dynamics here presented as a first step in the theoretic experiments, is very similar to these previous models. In fact in the first two sets of experiments (*integration* and *segregation*) the two extreme cases with a low and an high degree of segregation are considered. In the third experiment (*integration and segregation*), through the variation of the weights assigned to the two land uses (one of which is totally devoted to integration and the other one to segregation) we explore the different configuration emerging, as from a variation of the degree of segregation. In turn the results of the optimization process may differ from the Schelling model, because the total utility is considered and not only that of the two exchanging land uses.

Integration

In the *integration* experiment, for each set of weights the outcome of the optimization is similar to that obtained with the self-organized dynamics (figure 2), and practically all the optimal configurations belong to the A-set. The resulting total utilities for each set of weights are also similar (figure 3). This effect depends by the equal number of land uses. In fact, because the utility is generated by a couple of different cells, the number of housing cells surrounding a retail cell in the steady state is equal to the number of retail cells surrounding an housing cell. A set of rows horizontally or vertically disposed of alternate land uses, is the configuration that attains this effect, and the utility for each land use is always the same. The Pareto front is in fact represented by only one point, and the optimal configuration roughly coincides with the result of the self-organized dynamics. This is the Utopian situation in which the social utility agrees with the individual utility and in essence a plan is not necessary.

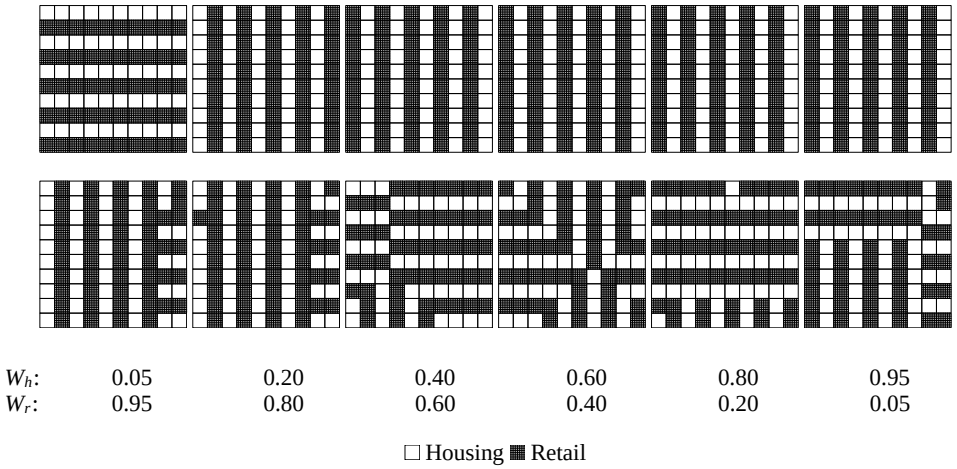


Figure 2. Integration. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. At the bottom of each column the set of weights utilized for housing (W_h), and retail (W_r) is shown.

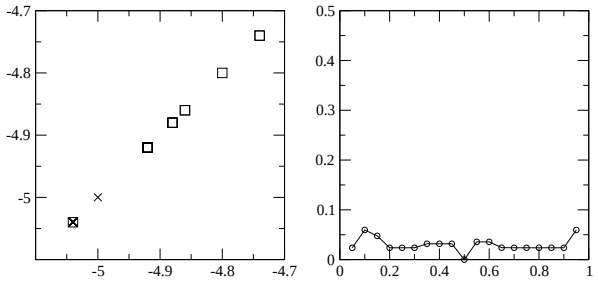
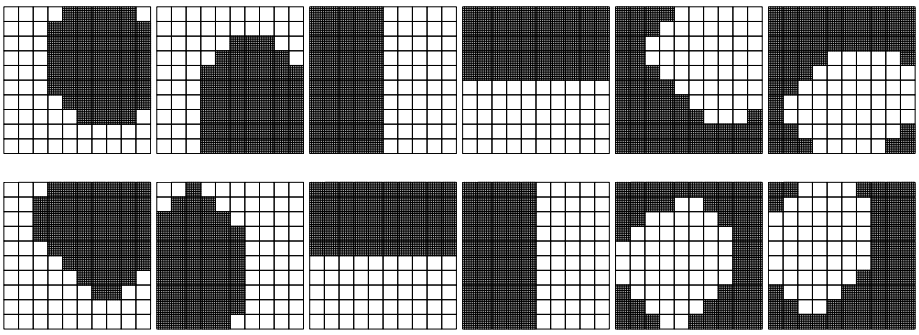


Figure 3. Integration. Left side, the Pareto front. Cross: configurations obtained through optimization. Square: configurations obtained by applying the rules of the self-organized dynamics to a random pattern. X axis: average utility of housing (negative). Y axis: average utility of retail (negative). Right side: the variation of I^d, the index of dissimilarity, in relation to W_h, the weight applied to the housing. X axis: W_h, Y axis: I^d.

Segregation

In the *segregation* experiment the patterns emerging from the applying of the segregation rules are twofold: the concentration of one use in the center and the division of space (horizontally or vertically) (figure 4). In fact due to the finite size of the surface, it is not possible for the two land uses, at the same time, to be circular shaped. The Pareto front is thereby represented by three points: the first two corresponding to the concentration of one use in the center and the third corresponding to the the division(figure 5). When the self-organized dynamics method is applied it produces a pattern which is similar to that produced by the optimization process (figure 4): all the optimal configurations belong to the A-set. Even if the utilities of the two land use are in opposition the situation is similar to the previous one and in essence a plan is superfluous.



W_h:	0.05	0.20	0.40	0.60	0.80	0.95
W_r:	0.95	0.80	0.60	0.40	0.20	0.05

□ Housing ■ Retail

Figure 4. Segregation. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. At the bottom of each column the set of weights utilized for housing (W_h), and retail (W_r) is shown.

Integration and segregation

The results of *integration and segregation* experiment appear similar to those of the two previous experiments: both the concentration of one use in the center and the rows horizontally or vertically disposed of

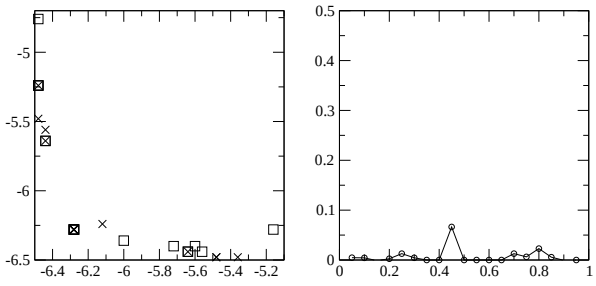


Figure 5. Segregation. Left side, the Pareto front. Cross: configurations obtained through optimization. Square: configurations obtained by applying the rules of the self-organized dynamics to a random pattern. X axis: average utility of housing (negative). Y axis: average utility of retail (negative). Right side: the variation of I^d, the index of dissimilarity, in relation to W_h, the weight applied to the housing. X axis: W_h, Y axis: I^d.

alternate land uses (figure 6). In order to compare more precisely these sets of experiments, let us focus on the differences in the matrices $\mathbf{M}$. This matrix is a-symmetric in this experiment while it was symmetric in the two previous ones. However this difference is only apparent. In fact, according with equation (7), the total utility U_T can also be calculated as the sum of the utilities related to each couple of bordering cells:

$$u^{\star}_{ij,hr} = W_h m_{hr} + W_r m_{rh}, \tag{13}$$

where $u^{\star}_{ij,hr}$ is the utility related to the couple of bordering cells i, where the land use h is located, and j, where the land use r is located, and $U_T = \sum_{ij} u^{\star}_{ij}$. Further the symmetric matrix of interaction $\mathbf{M}^s$ is defined in which:

$$m^s_{hr} = m^s_{rh} = \frac{W_h m_{hr} + W_r m_{rh}}{2}. \tag{14}$$

The total utility can also be calculated by using this symmetric matrix, as in the following equations:

$$u^{\star}_{ij,hr} = 2m^s_{hr}, \text{ and } u^{\star}_{ji,rh} = 2m^s_{rh}. \tag{15}$$

Because each couple is considered twice the total utility calculated with this method equals that calculated with equation (4). Now consider the symmetric interaction matrices $\mathbf{M}^s$ of the *integration* and of the *integration and segregation* experiments in relation to the set of weights: $W_h = 0.95$ and $W_r = 0.05$. As it is straightforward from the first part of table 2 these two matrices are similar as well as the interaction matrices of the *segregation* experiment and of the *integration and segregation* in relation to the set of weights: $W_h = 0.05$ and $W_r = 0.95$ (second part of table 2). This is why the pattern (figure 6) obtained

Table 2. Comparison of the matrices $\mathbf{M}^s$.

$W_h = 0.95$ and $W_r = 0.05$

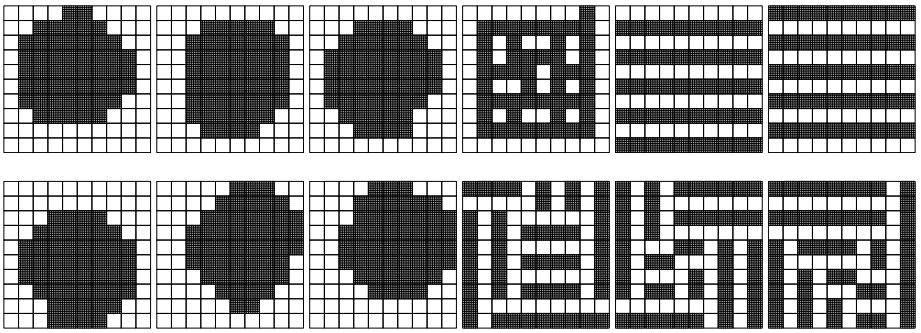
	Integration		Integr.-Segreg.	
Land use	Hous.	Ret.	Hous.	Ret.
Housing	0	0.5	0	0.475
Retail	0.5	0	0.475	0.05

$W_h = 0.05$ and $W_r = 0.95$

	Segregation		Integr.-Segreg.	
Land use	Hous.	Ret.	Hous.	Ret.
Housing	0.05	0	0	0.025
Retail	0	0.95	0.025	0.95

with $W_h = 0.05$ and $W_r = 0.95$ is similar to that obtained in the *segregation* experiment, while the pattern obtained with $W_h = 0.95$ and $W_r = 0.05$ is similar to that obtained with the *integration* experiment. In addition an intermediate pattern (a transition between the two) is obtained when the weigths assigend to the land uses are similar.

The Pareto front is less convex than in the previous experiment, while the outcomes of the self-organized dynamics and of the optimization process are similar (A-set) till the weight assigned to the retail is greater than that assigned to housing (figure 7). Especially in the intermediate regime and when the weight assigned to the housing is greater than that assigned to retail, the optimal solutions belong to the B-set.



W_h:	0.05	0.20	0.40	0.60	0.80	0.95
W_r:	0.95	0.80	0.60	0.40	0.20	0.05

□ Housing ■ Retail

Figure 6. Integration and segregation. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. At the bottom of each column the set of weights utilized for housing (W_h), and retail (W_r) is shown.

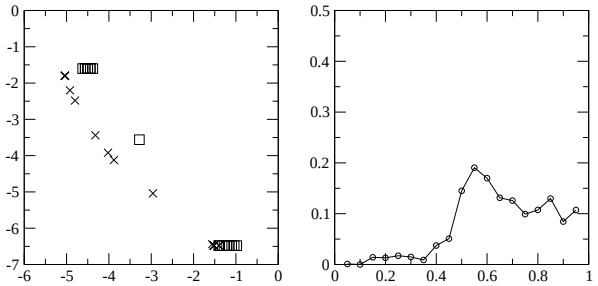
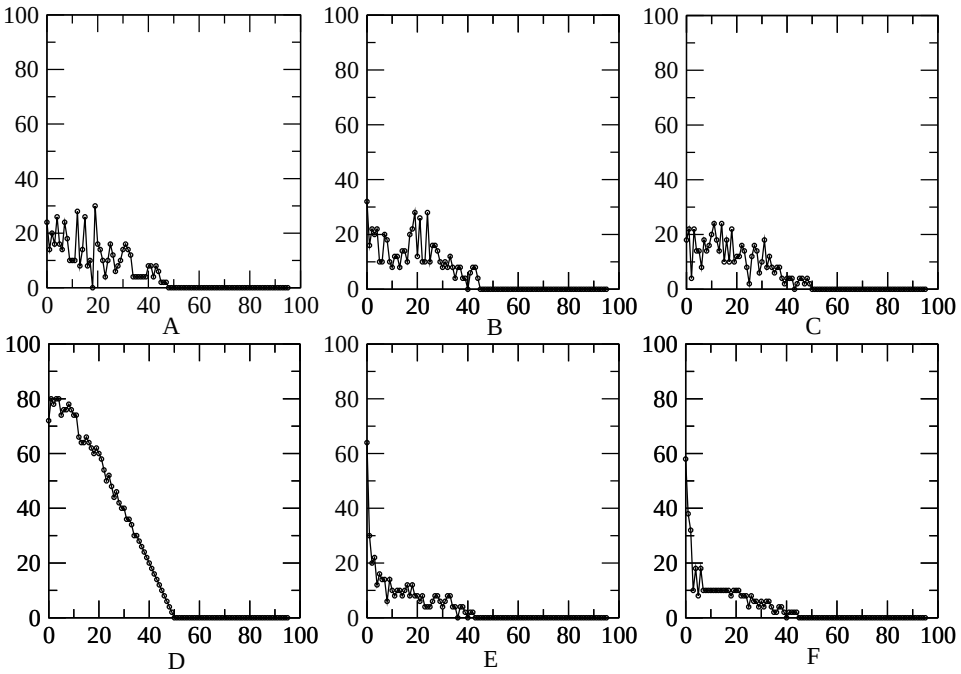


Figure 7. Integration and segregation. Left side, the Pareto front. Cross: configurations obtained through optimization. Square: configurations obtained by applying the rules of the self-organized dynamics to a random pattern. X axis: average utility of housing (negative). Y axis: average utility of retail (negative). Right side: the variation of I^d, the index of dissimilarity, in relation to W_h, the weight applied to the housing. X axis: W_h, Y axis: I^d.

The first time observed dissimilarity between the optimal and the self-organized configuration, results in the applying of the method for the convergence. The outcome has been evaluated by considering the relation: number of pinning cells–distance between optimal and self-organized configuration (D_{pq}). As figure 8 shows, the decrease of the distance is proportional to the number of pinning cells unless the last two cases (E and F, figure 8). But in all the cases the distance is zero only when the number of pinning cells equals 50, which coincides with the number of housing and retail land uses.



	A	B	C	D	E	F
W_h	0.05	0.2	0.4	0.6	0.8	0.95
W_r	0.95	0.8	0.6	0.4	0.2	0.05

Figure 8. Integration and segregation. The varying of distance D_{pq} with the increase of the number of the pinning cells in which the land use has been fixed. X axis: the number of the pinning cells, Y axis the distance D_{pq}.

Integration and segregation, second experiment

Because these results are influenced by the equal number of housing and retail land uses, a similar experiment has been performed with a different number of land uses, and with the inclusion of the open land. The number of land uses has changed as it follows: 50 cells for housing, 10 for retail and 40 for open space, while the matrix $\mathbf{M}$ has been slightly modified, as in table 3. Through this variation the degrees of freedom

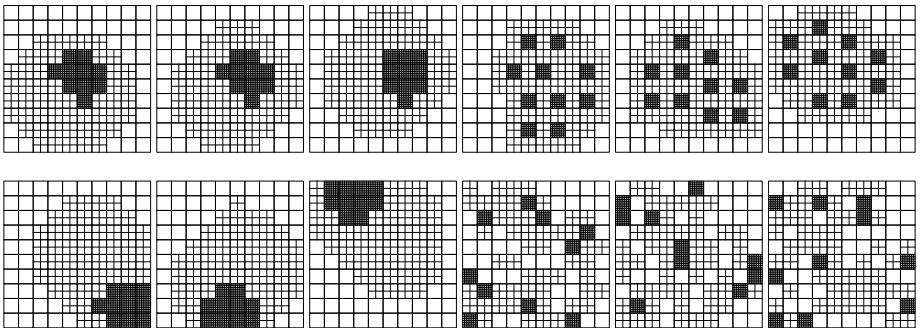
Table 3. Integration and segregation, second experiment. Matrix $\mathbf{M}$.

Integration and segregation		
Land use	Hous.	Ret.
Housing	0.1	0.9
Retail	0	1

of the housing and retail land uses, in the occupation of the surface increase. In fact, as figure 9 shows, the resulting patterns are quite different. The optimal configurations are twofold: the concentration of the retail land use in the center, which is similar to the result of the self-organized dynamics and an homogeneous distribution of retail regularly mixed with the housing land use which differs from the result of the self-organized dynamics. Thereby the Pareto front is reduced to only two points (figure 10). The index of dissimilarity I^d increases with the increase of the weight W_h assigned to the housing land use (figure 10).

In turn a low number of pinning cells is enough to obtain an important decrease of the distance (figure 11). In fact in the first ten pinning cells (see graphs D, E, and F of figure 11) are just located the ten retail land uses, which influence the location of the housing cells.

This last aspect highlights the relation between the self-organized dynamics and the optimization process. These two processes converge when the maximum weight is assigned to a land use which plays a *central* role in the the interaction. The centrality of a land use can be roughly calculated by using the difference: sum by columns minus sum by row in the matrix $\mathbf{M}$ (table 3) i.e. the capability to influence minus the degree of being influenced, as it has been done with the calculation of the influence index. It is easy to conclude that in this experiment the retail land use plays the most central role. Hence the optimal configuration and the outcome of the self-organized dynamics coincide when a big weight is assigned to the retail land use. In turn, a lot of energy has to be spent to compel the system to converge with the optimal configuration, when the maximum weight is assigned to a land use which does not play a central role, in the sense previously defined.



W_h:	0.05	0.20	0.40	0.60	0.80	0.95
W_r:	0.95	0.80	0.60	0.40	0.20	0.05

⊞ Housing ▩ Retail □ Open land

Figure 9. Integration and segregation, second experiment. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. At the bottom of each column the set of weights utilized for housing (W_h), and retail (W_r) is shown.

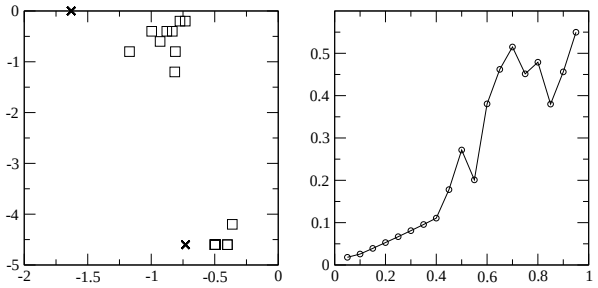
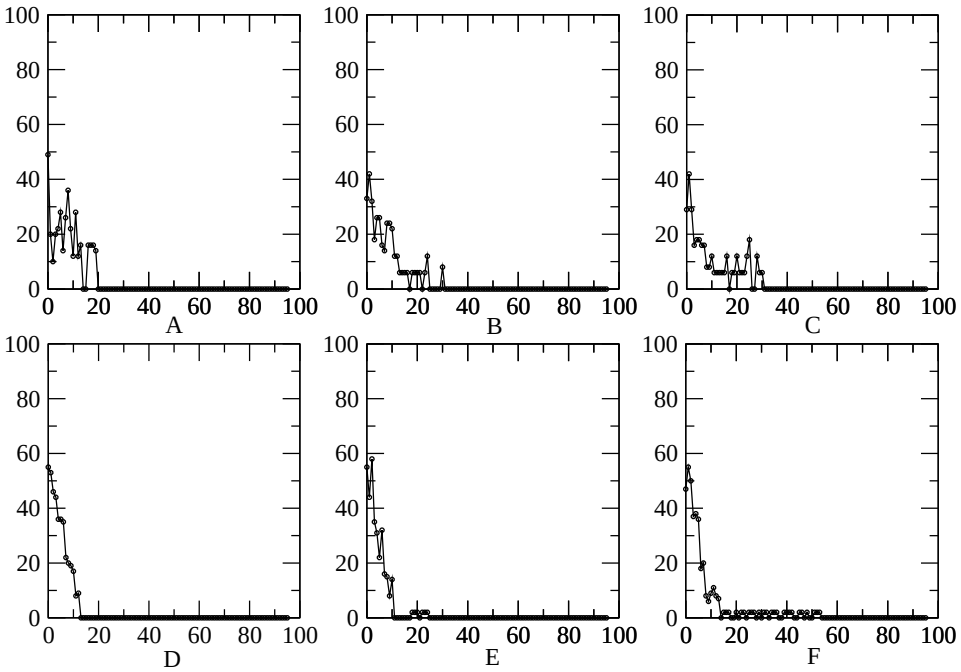


Figure 10. Integration and segregation, second experiment. Left side, the Pareto front. Cross: configurations obtained through optimization. Square: configurations obtained by applying the rules of the self-organized dynamics to a random pattern. X axis: average utility of housing (negative). Y axis: average utility of retail (negative). Right side: the variation of I^d, the index of dissimilarity, in relation to W_h, the weight applied to the housing. X axis: W_h, Y axis: I^d.



	A	B	C	D	E	F
W_h	0.05	0.2	0.4	0.6	0.8	0.95
W_r	0.95	0.8	0.6	0.4	0.2	0.05

Figure 11. Integration and segregation, second experiment. The varying of distance D_{pq} with the increase of the number of the pinning cells in which the land use has been fixed. X axis: the number of the pinning cells, Y axis the distance D_{pq}.

A more realistic experiment

The following experiment is conceived in a way that is less schematic and more similar to the reality. Two new land uses, industry and equipment, are added and each cell can be assigned to one of the following land-uses: housing (30 cells), retail (5 cells), industry (10 cells), open land (53 cells), and equipment (2 cells). The location of equipment is fixed, in order to simulate the existence of some exogenous factors. The matrix $\mathbf{M}$ is shown in table 4. As usual, positive effects occur for housing in case it has in the nearby

Table 4. Matrix $\mathbf{M}$, and number of land uses.

Land use	Housing	Retail	Industry	Open land	Equipment	Number of land uses
Housing	0.3	0.4	0.1	0.2	0.0	30
Retail	0.35	0.5	0.15	0.0	0.0	5
Industry	0.0	0.1	0.2	0.0	0.7	10
Open land	0.0	0.0	0.0	0.0	0.0	53
Equipment	0.0	0.0	0.0	0.0	0.0	2

cells retail and open land uses, and in a lower degree, industry in which workplaces are located. Positive effects occur also for the retail activities in relation to housing and industry which represent in various

degree a demand for retail. Industry utility increases with the contiguity to the equipment.

Figure 12, shows the outcomes of the experiment including also the influence index of each cells, while in figure 13 the Pareto front is reported in three views. The convergence depends on the weigh assigned to the retail land use which at a first insight appears as the most central in the interaction rules. However, to calculate more rigorously the index of centrality of a land use we reason in the following way. Reminding that m_{hr} is the utility for use h to be surrounded by use r, m_{rh} is a measure of the influence of r on h. The more the quantity of land uses h, the more the absolute influence of r. For this reason a matrix $\mathbf{M}'$ is defined which is the transpose of the matrix $\mathbf{M}$, where each element is multiplied by the probability to find the influenced land use use:

$$m'_{rh} = m_{hr}\frac{N_h}{n^2}. \tag{16}$$

Similarly as for matrix $\mathbf{G}$, the direct and indirect influence is calculated as in the following expression:

$$\mathbf{M}^\star = \mathbf{M}' + \mathbf{M}'^2 + \mathbf{M}'^3 + \dots \tag{17}$$

The index of the centrality for the land use h (I_h^c) is given by the sum by row minus the sum by columns, in other words the capacity to influence minus the degree of being influenced:

$$I_h^c = \sum_r m^\star_{hr} - \sum_h m^\star_{hr}. \tag{18}$$

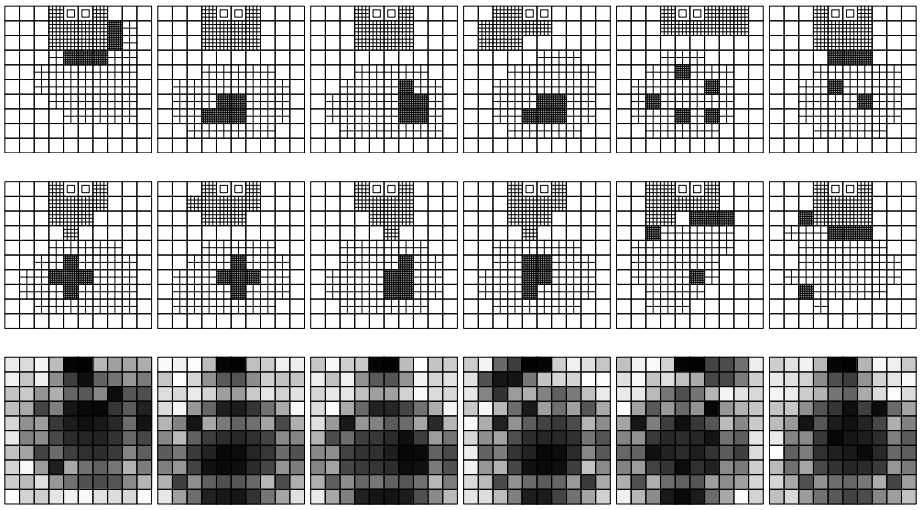
After having established a method for calculating an index of centrality we can forecast the relation between the optimal configuration and the self-organized dynamics. According to the result, shown in the table 5, last column, all the efficient solutions generated with a big weight assigned to the retail land use should coincide with the result of the self-organized dynamics. In fact, as figure 12 shows, the optimal configuration matches the self-organized dynamics when a big weight is assigned to the retail land use. In addition, the consequences of the establishment of pinning cells are shown in figure 14. The distance decreases quickly with the increase of the number of the pinning cells when the weight of the land uses and the centrality index does match (see the cases B and C, figure 14). In case they don't, the central role of the retail land use is utilized to control the other land uses (cases A and F, figure 14, in which the first chosen pinning cells are mostly retail land use). In essence in this case the control of an only limited part of the pinning cells produces a convergence, with a limited amount of energy spent in the control.

Table 5. Elements of the matrix $\mathbf{M}^\star$.

Land use	Housing	Retail	Industry	Open land	Equipment	Centrality index (I^c)
Housing	0.09899	0.01983	0.00020	0	0	-0.22033
Retail	0.13632	0.02316	0.01053	0	0	0.11671
Industry	0.03497	0.00851	0.01861	0	0	-0.03929
Open land	0.06661	0.00120	0.00001	0	0	0.06782
Equipment	0.00247	0.00060	0.07202	0	0	0.07510

A 30×30 grid experiment

Finally and in order to discuss the scalability of the proposed method the spatial grid has been enlarged to 30×30 squared cells. The rules of interactions included in matrix M are the same unless the neighborhood Φ_i utilized to calculate the spatial relations which has been enlarged to the 48 cells included in the square of 7×7 cells around the central cell in question. This enlargement is necessary in order to simulate the long range spatial relations, and it is the only change which assures the scalability of the proposed method.



W_h:	0.05	0.05	0.05	0.50	0.90	0.50
W_r:	0.05	0.50	0.90	0.45	0.05	0.05
W_i:	0.90	0.45	0.05	0.05	0.05	0.45

Figure 12. The more realistic experiment. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. Third row: the rank of pinning cells in relation to the influence index (I^i). The gray scale represents the influence of the cells: the most influent cells are black, the less are white colored. At the bottom of each column the set of weights utilized for housing (W_h), retail (W_r) and industry (W_i) is shown.

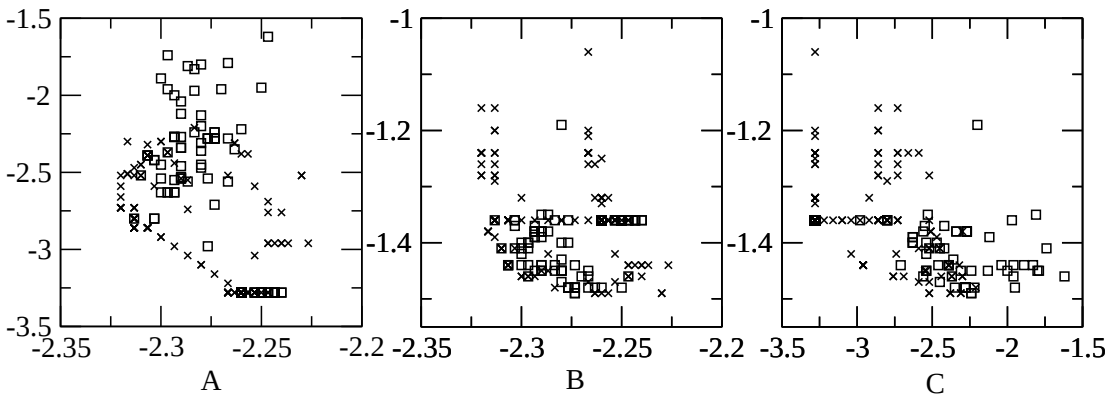
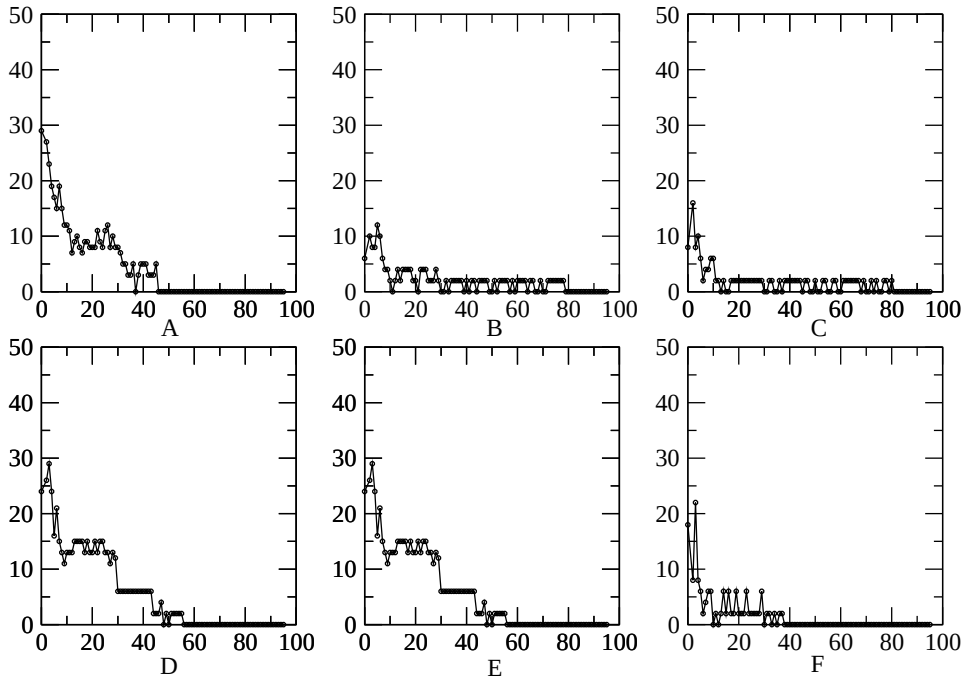


Figure 13. The more realistic experiment. The Pareto front. Cross: configurations obtained through optimization. Square: configurations obtained by applying the rules of the self-organized dynamics to a random pattern. A: X axis: average utility of housing (negative). Y axis: average utility of retail (negative). B: X axis: average utility of housing (negative). Y axis: average utility of industry (negative). C: X axis: average utility of retail (negative). Y axis: average utility of industry (negative).

The number of land uses has been proportionally increased: housing: 270 cells, retail: 50 cells, industry: 90 cells, open land: 470 cells and equipment: 20 cells. Only two cases are shown (figure 15) which utilize the same set of weight as the first and the sixth cases shown in figure 12. The analysis of the convergence



	A	B	C	D	E	F
W_h	0.05	0.05	0.05	0.50	0.90	0.50
W_r	0.05	0.50	0.90	0.45	0.05	0.05
W_i	0.90	0.45	0.05	0.05	0.05	0.45

Figure 14. The more realistic experiment. The varying of distance D_{pq} with the increase of the number of the pinning cells in which the land use has been fixed. X axis: the number of the pinning cells, Y axis the distance D_{pq}.

in relation to the number of the pinning cells is shown in figure 16. The results are similar to the 10×10 experiments, even if a finer tuning of the interaction matrix is requested in order to obtain a pattern less schematic and more similar to that observed in the reality.

Relation with the urban planning and control

In this experimental section we have found that the optimization process and of the self-organized dynamics produces similar results when the weigh assigned to the land uses coincide with the index of centrality of land use. However when this coincidence does not happen, the land use with an big index of centrality can be easily utilized to control the self-organized dynamics. These conclusions have to be related to some experience of urban planning especially in the fields of the regeneration of existing areas[17]. In this case what planners are looking for is the catalyst effects of the urban project[21]. This is very similar to obtain the desired plan with the control of a limited number of cells. The catalyst is in fact like a pinning cell in which planning effort are concentrated in order to stimulate the development in the desired direction. However the identification of the critical point [22] where to concentrate investments is a further possible utilization of the proposed method. In this way the strict zoning control could be relaxed in order to allow the network of local actions to operate more freely.

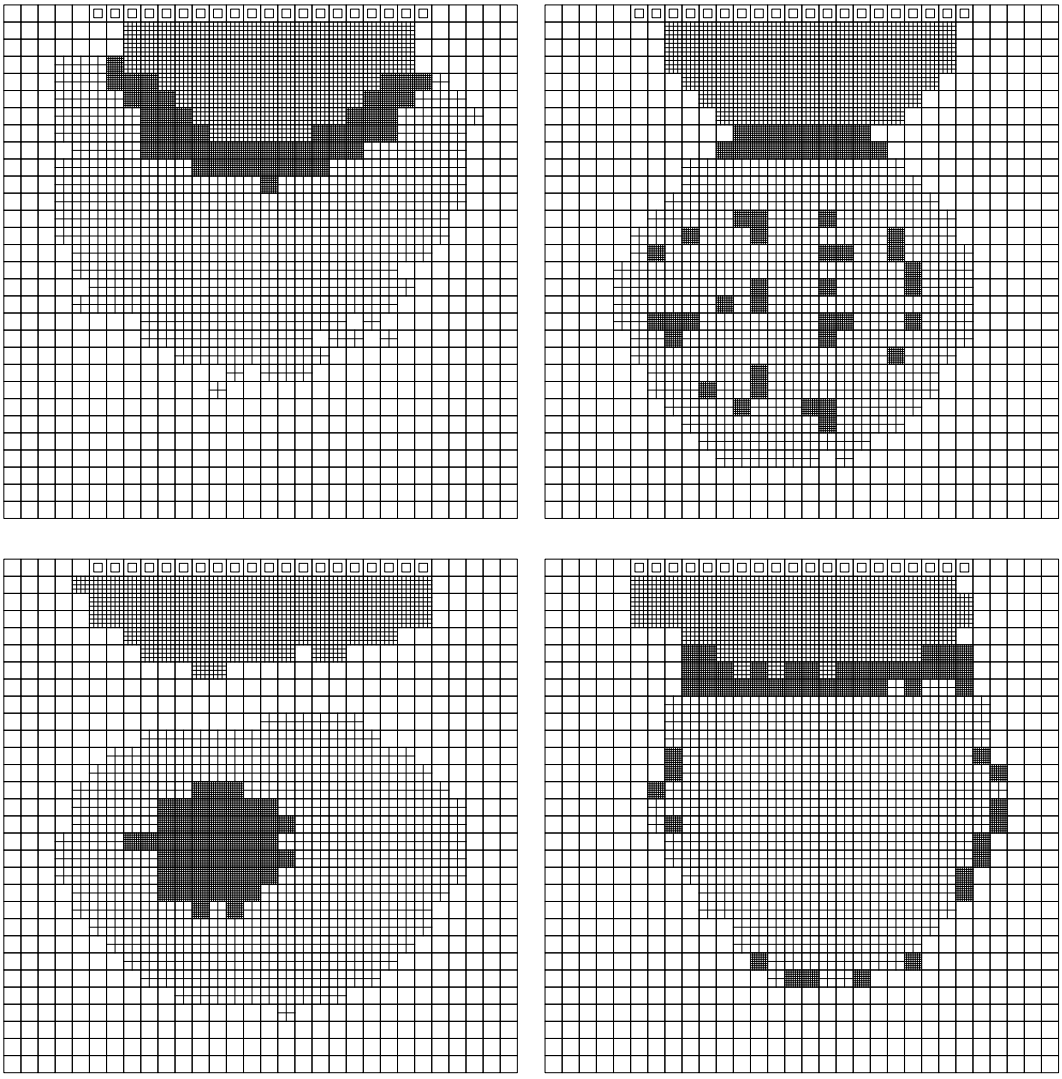


Figure 15. The 30×30 grid experiment. First row. The patterns obtained through optimization. Second row. The patterns obtained by applying the rules of the self-organized dynamics to a random pattern. The two columns of graphs refer to the following set of weights: first column: $W_h = 0.05$, $W_r = 0.05$, and $W_h = 0.9$; second column: $W_h = 0.5$, $W_r = 0.05$, and $W_h = 0.45$. They have to be compared with first and sixth cases in figure 12.

Conclusion

We have shown that optimization and self-organized dynamics can be conceived as similar process. From one side the optimization is a special kind of dynamics, and from the other side the self-organized dynamics is a special kind of optimization. In other words the urban systems are also problem solving. In turn the solving of the problem can be conceived as a dynamic process in which at each step the utility arising from each couple of bordering cells is considered instead of focusing in the utility of only one land use, and

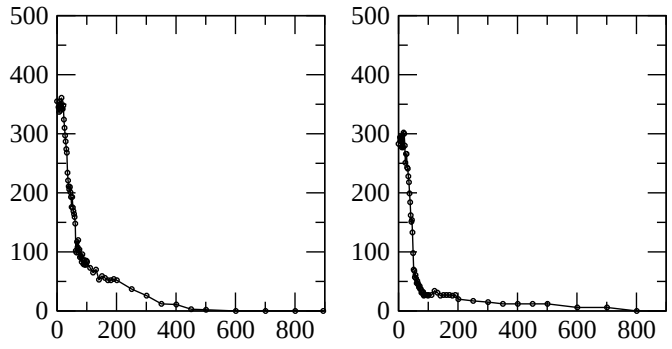


Figure 16. The 30×30 grid experiment. The varying of distance D_{pq} with the increase of the number of the pinning cells in which the land use has been fixed. X axis: the number of the pinning cells, Y axis the distance D_{pq}.

with the utilization of an exogenous factor, i.e. the cooling, which allows the system to reach and maintain the optimal configuration. Sometimes the attractors of the self-organized dynamics and of the optimization process are similar. This is the Utopia in which individual and total utility do coincide. In case this does not happen, as usually, instead of using the control extended to each point of the surface that totally constraints the self-organized dynamics with the plan, we have proposed a method which takes advantage of the self-organizing character of the urban dynamics and minimizes the controlling effort. Finally, this method can be applied to the urban planning practice based on the research of the element able to produce positive catalytic effect on the whole urban structure.

Acknowledgments

I thank Stefano Ruffo for stimulating discussions while this study has been performed and Fabio Schöen for the review of a previous version of the paper. Responsibility for the opinions expressed is solely that of the author.

References

[1] Portugali J: Self-Organization and the City. Berlin, Springer-Verlag, 1999.

[2] Cartwright T J: Planning and chaos theory. J Am Plann Assoc 1991; 57: 44–56.

[3] Alexander C: The Timeless Way of Building. New York, Oxford University Press, 1979.

[4] Batty M: Cities, planning, design, computation and evolution. Environ Plann B 1995; 22: 379–382.

[5] Portugali J, Alfasi N: Planning just-in-time versus planning just-in-case. Cities 2004; 21: 29–39.

[6] Koopmans TC, Beckman MJ: Assignement problems in the location of economic activities. Econometrica 1957; 25: 53–76.

[7] Schlager KJ: A land use plan design model. J Am Inst Plann 1965; 31: 103–111.

[8] Bammi D, Bammi D: Development of comprehensive land use plan by means of a multiple objective mathematical programming model. Interfaces 1979; 9: 50–63.

[9] Michalek JJ, Choudhary R, Papalambros PY: Architectural layout design optimization. Eng Optimiz 2002; 34: 461–484.

[10] Balling RJ, Taber JT, Brown MR, Day K: Multiobjective urban planning using genetic algorithm. J Urban Plan Dev Div 1999; 125: 86–99.

[11] Cheng-Min F, Jen-Jia L: Using a genetic algorithm to generate alternative sketch maps for urban planning. Comput Environ Urban 1999; 23: 91–108.

[12] Harris B. Optimization and design. Environ Plann B 1998, Anniversary Issue: 23–28.

[13] Fonseca CM, Fleming PJ: Multiobjective optimization; in Baeck T, Fogel D, Michalewicz Z (eds): Handbook of Evolutionary Computation. Oxford, IOP Publishing Ltd and Oxford University Press, 1998-2000.

[14] Kirkpatrick S, Gelatt C, Vecchi M: Optimization by simulated annealing. Science 1983, 220: 671–680.

[15] Helbing D, Vicsek T: Optimal self-organization. New J Phys 1999, 1.

[16] Torrens PM, O'Sullivan D: Cellular automata and urban simulation: where do we go from here? Environ Plann B 2001, 28: 163–168.

[17] Attoe W, Logan D: American Urban Architecture: Catalysts in the Design of Cities. Berkeley, University of California Press, 1991.

[18] Grigoriev RO, Cross MC, Schuster HG: Pinning control of spatiotemporal chaos. Phys Rev Lett 1997, 79: 2795–2798.

[19] Junge L, Parlitz U, Tasev Z, Kocarev L: Synchronization and control of spatially extended systems using sensor coupling. Int J Bifurcat Chaos 1999, 9: 2265–2270.

[20] Schelling TC: Models of segregation. Am Econ Rev 1969, 59: 488–493.

[21] Lowe M: The regional shopping centre in the inner city: A study of retail-led urban regeneration. Urban Stud 2005, 42: 449–470.

[22] Chalkley B, Essex S: Urban development through hosting international events: a history of the olympic games. Planning Perspectives 1999, 14: 369–394.

Simulating pedestrians and cars behaviours in a virtual city : an agent-based approach

Arnaud Banos [1] , Abhimanyu Godara [2], Sylvain Lassarre[3]

[1]SET (UMR 5603 CNRS), Université de Pau
Avenue du Doyen Poplawski, 64000 Pau, France
arnaud.banos@univ-pau.fr
[2] Dept. of Civil Engg., Institute of Technology, Banaras Hindu University,
221005 Varanasi, India
abhimanyu_godara@rediffmail.com
[3] GARIG, Institut National De Recherche sur les Transports Et Leur Sécurité,
Arcueil 94114, France
lassarre@inrets.fr

Introduction

Despite the central and fundamental role pedestrian walking plays within the urban transport system, it still remains a badly known transportation mode. Generally speaking, while most of the developed countries have been developing, for the last 40 years, a wide variety of sophisticated methods and tools aimed at studying urban mobility, only a few of them were really designed to deal with pedestrian movement, especially in interaction with the other transportation modes. Noticing the tremendous asymmetry that has been existing for long in the scientific literature between traffic flow and pedestrian movement, Weifeng and his colleagues [Weifeng, 2003] propose several explanations :

"*The reason may be that the pedestrian movement is more complex than vehicular flow. First, pedestrians are more intelligent than vehicles and they can choose an optimum route according to the environment around. Secondly, pedestrians are more flexible in changing directions and not limited to the "lanes" as in vehicular flow. Thirdly, the slight bumping is acceptable and need not be absolutely avoided as in traffic flow models. So the model developed for pedestrian movement should fully consider these differences in order to study the special phenomena in pedestrian movement. It is pity that till now, most pedestrian movement models are established based on the rules used for traffic flow and consider little of the special characteristics of pedestrian movement itself*"[pp. 634].

The key point we will try to defend here, is that pedestrian movement needs not only to be considered as a specific phenomenon. It also needs to be included in a much more global and complex perspective, the urban system as a whole. Pedestrian motion indeed occurs in an ever changing environment, defined by constraints and opportunities, but also nuisances and dangers. The SAMU prototype has been precisely designed to explore the behaviour of pedestrians in interaction with the motorized traffic, in a virtual city where most of the phenomenon can be mastered and studied. This idea of designing "virtual laboratories" [Batty and Torrens, 2001], within which "artificial societies"can be grown for example [Epstein and Axtell, 1995] has become very popular in the recent years and is largely related to two other fields, science of complex systems and agent based modelling. More, it is firmly embedded in a microscopic approach of urban mobility, where the world is represented as closely as possible in a one-to-one way, which means that "*people should be represented as people, cars should be represented as cars and traffic lights should be represented as traffic lights and not as, say, departure rates, traffic streams and capacities respectively*" [Nagel et al., 2000].

The SAMU prototype

While being fundamental steps, most of the micro-simulation approaches of pedestrian movement [Blue and Adler, 1998 ; Helbing et Monnar, 1997 ; Helbing et al., 2001 ; Batty, 2003 ; Haklay et al., 2001 ; Kerridge et al., 2001] have the drawback of relying on an excessive simplification of the urban environment (a corridor, a place or a room). Motivations and goals of pedestrians are also particularly simplified, reduced to the couple destination to reach / obstacles to avoid. Theses limitations encouraged other researchers to explore agent-based models of pedestrian movement [Batty, 2003 ; Haklay et al., 2001 ; Kerridge et al., 2001 ; Schweitzer, 2003 ; Zachariadis, 2005]. In this last family, each pedestrian/agent is defined by a set of capacities and tries to achieve a set of goals, interacting locally with its environment and with other agents.

The prototype SAMU[1] directly relies to that specific field, its originality being defined by its focus on interactions between pedestrian and traffic flows. Developed in NETLOGO[2], SAMU is an hybrid model, combining characteristics of both cellular automata and agent-based models. Cars and pedestrians are indeed defined as agents, situated on an active grid (Figure 1), with which they interact.

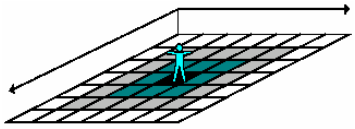


Figure 1 : the basic principle of SAMU

Then, agents have to perform specific tasks, interacting locally with other agents and with their environment. Figure 2 shows the prototype developed in order to observe and test these interactions, as well as emerging parameters, such as speed of cars or proportion of cars/pedestrians collisions.

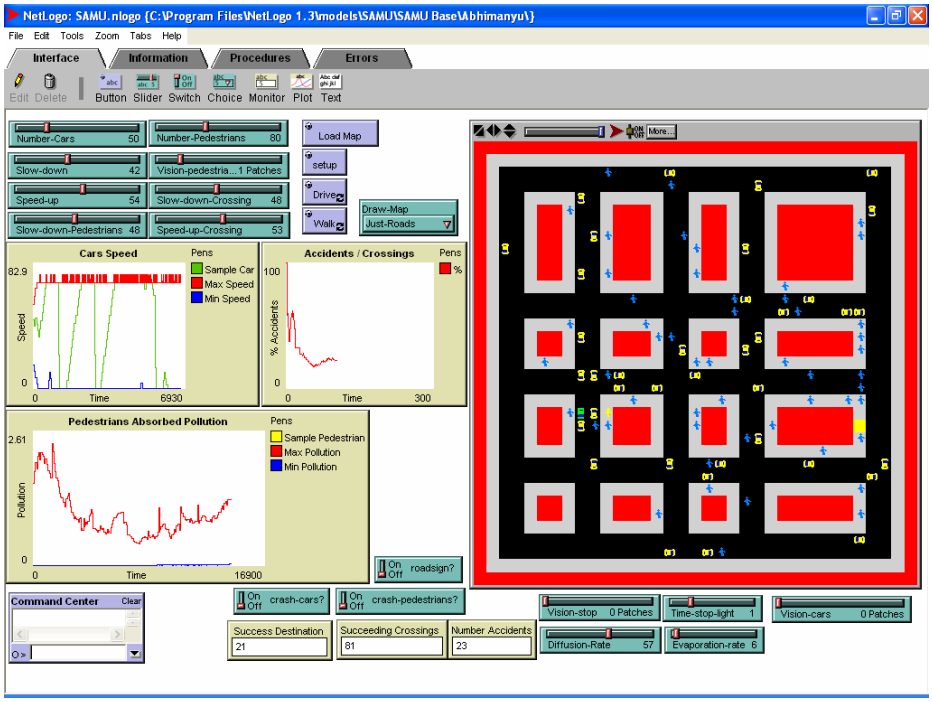


Figure 2 : The SAMU prototype

[1] Simulation Agents et Modélisation Urbaine, http://www.univ-pau.fr/~banos/sma.html
[2] http://www.ccl.sesp.northwestern.edu/netlogo

Model Formulation

Our prototype SAMU, developed in Netlogo, integrates both cellular Automata and Agent-Based Modelling approaches to generate a new model of cars behaviours by appropriately modifying earlier NaSch/ChSch rules [Nagel and Schreckenberg, 1992] to take into account pedestrian and also turning movements. We also consider two-way traffic with turning movements to bring the model closer to the real world.

Following the prescription of the NaSch model, we allow the speed V of each vehicle to take one of the Vmax + 1 integer values V = 0, 1, 2….Vmax. For urban systems we do not want to have Vmax more than 70 km/h. So we are taking maximum speed as 3 (22.5 m/s as each cell is 7.5 m in length as in NaSch model). Suppose Vn is the speed of the nth vehicle at time t while moving in any direction (different from NaSch/ChSch/BML model where vehicles move either towards east or towards north and number of cars is fixed on a given road). To emphasize the effect of turning movements and pedestrians we are considering only unsignalized intersections in our model and also we want each car to stop at the intersection and decide regarding the turning movements to get homogeneity. The above assumption is true considering the fact that drivers become more cautious and reduce their speed at intersections to avoid any kind of collisions with other vehicles. At each discrete time step t → t+1, the arrangement of N vehicles is updated in parallel according to the following "rules" :

Step 1: Acceleration

If Vn < Vmax, the speed of the nth vehicle is increased by one, i.e., Vn→ Vn+1.

Step 2: Deceleration (due to other vehicles, intersections, or pedestrians)

Suppose Dn is the gap in between the nth vehicle and the vehicle in front of it, Pn is the distance between the same car and the closest intersection in front of it (note here intersection not signal, which makes the model more general) where the car has to make a decision about turning, and D is the minimum gap between the car under consideration and the pedestrian in front of it on the road (if any).

Now, If min(Dn, Pn, D) ≤ Vn

and if min(Dn, Pn) < D,

then Vn→min(Dn – 1, Pn)

else Vn→D – 1 (the motivation for this choice comes from the fact that to avoid accident, the car will stop one cell before the pedestrian)

Step 3: Randomization

If Vn > 0, the speed of the car under consideration is decreased randomly by unity (i.e, Vn→Vn – 1) with probability p ($0 \leq p \leq 1$); p the random deceleration probability is identical for all the vehicles, and does not change during updating.

Step 4: Movement

Each vehicle moves forward with the given speed i.e. Xn→Xn + Vn, where Xn denotes the position of the nth vehicle at any time t.

The major changes are made in step 2 which reflects the interaction among vehicles and pedestrians. Step 4 shows there are no more north-bound or east-bound vehicles, the speed of each car being updated simultaneously without any specific classification. While being a work in progress, SAMU already provides an ergonomic platform useful to test the behaviour of the system under different configurations of parameters. Anyway, reaching such a modelling level, without being flooded with microscopic details, requires an ad-hoc procedure. Crucial principles like reductionism and parsimony [Batty and Torrens, 2001] may therefore constitute main guidelines, in our quest for the identification of the micro-specifications sufficient to generate macrostructures of interest [Epstein, 1999].

References

Batty M., 2003 : Agent-based pedestrian modelling, CASA paper n° 61, 26 p.

Blue V., Adler J., 1998 : Emergent fundamental pedestrian flows from cellular automata microsimulation, Transportation Research Record, vol. 1644, pp. 29-36

Drogoul A., 2000 : Systèmes multi-agents distribués, Habilitation à Diriger des Recherches, Université de Paris 6, 117 p.

Epstein J., 1999 : Agent-based computational models and generative social science, Complexity, Vol. 4, N°5

Epstein J., Axtell R., 1996 : Growing artificial societies : social science from the bottom up, Brookings Institution Press, MIT Press, Washington DC

Hacklay M., Thurstain-Goodwin M., O'Sullivan D., Schelhorn T., 2001 : So go downtown : simulating pedestrian movement in town centers, Environment and Planning B, n° 28, pp. 343-359

Helbing D., Molnar P., Farkas I., Bolay K., 2001 : Self-organizing pedestrian movement, Environment and Planning B, n° 28, pp. 361-383

Kerridge J., Hine J., Wigan M., 2001 : Agent-based modelling of pedestrian movements : the questions that need to be asked and answered, Environment and Planning B, vol. 28, pp. 327-341

Nagel K., Schreckenberg M., 1992: Cellular automaton models for freeway traffic. Physics. I , 2, 2221-2229

Nagel Kai, Esser Jorg, Rickert Marcus, 2000 : Large-scale traffic simulations for transportation planning, Annual Reviews of Computational Physics VII, pp. 151-202

Scheiwtzer F., 2003 : Brownian Agents and active particles Springer-Verlag

Weifeng Fang, Lizhong Yang., Weicheng Fan, 2003 : Simulation of bi-direction pedestrian movement using a cellular automata model, Physica A 321, p. 633 – 640

Emergence of Fame

Haluk Bingol

bingol@boun.edu.tr
Department of Computer Engineering, Bogazici University

In many cases complex networks, including social networks, *the accessible is much less than available* since nodes have limited capacity. A model with interacting individuals of fixed memory size is proposed. A person "knows" another person if there is a corresponding pointer in his memory. A simple recommendation model is defined. Memory content is changed by the recommendation process. The model is simulated. When the memory size is much less then the population size, some self-organized patterns are observed. The findings are used to explain the concept of fame.

This model could be a constructive model for the memory of individuals in a population and the collective memory of the population. The model has the potential that can be extend to modeling of cultural changes, marketing and advertisement strategies.

1. Introduction

Suppose we need a dentist or information about an item. What do we do? Our first approach is to check our memory whether we know any. Then, our second move would be to ask our friends for a recommendation. Then, we would go to search engines.

In many cases of complex networks, including social networks, *the accessible is much less than available* since nodes have limited capacity. A node should "know" the other node in order to link to it. Since the number of nodes is, simply, too large, no node should be able "to know" all the other nodes. Only a small fraction of the total nodes can be known by any single node. On the other hand, there is no limit to "be known". As an example consider a web site. The number of outgoing links to different sites is limited by the number of sites the webmaster knows. Clearly, the number of persons an average person knows is much less than the number of people on the world. Similarly, the capacity of the IP table of a router is much less then available IPs.

Suppose a new web site is created and someone discovers it. If she likes the site, starts recommending it to her contacts. The contact that receives the recommendation evaluates it. If he likes the site, he also starts recommending it to his contacts. Recommendation is the fundamental process of how to become "known".

A couple of concepts are involved in these processes including our memory, our social network, the cumulative memory of the social network and a recommendation process. Suppose we are looking for *S*. The more *S* is known, the more likely that it can be reached, since the probability of finding a person that knows it within our social network increases. This discussion calls for fame which is the central concept of this paper. This paper investigates the effect of the size of memory to the fame. A model is defined. Simulation results of different memory sizes and population sizes are investigated. The findings are interpreted. A constructive model for fame is proposed. To the author knowledge, there is no previous such model for fame exists.

1.1. Fame

One pragmatic measure of fame is the number of pages a Google search returns. It is assumed that the more pages return, the more popular the item is. This measure of popularity has an implicit problem since an item can be listed because of some other reason. For example, a singer can be listed because of being a movie actor as well as being a singer. The fame of WWI fighter-pilots is obtained using this approach [1]. It is found that

the probability $p(f)$ of having fame f decreases with a power-law of f, namely $p(f) \sim f^{-\gamma}$ where $\gamma \approx 1.9$. Let a be the achievement of the person which is defined as the number of opponent aircrafts destroyed. It is found that $p(f)$ increases exponentially with achievement a, $p(f) \sim e^{\beta a}$ where $\beta \approx 0.074$. In a similar study, the popularity of scientists in condensed matter and statistical physics is investigated [2]. The findings of this are quite different. $p(f)$ decays exponentially, rather than in power-law fashion, $p(f) \sim e^{-\eta f}$ where $\eta \approx 0.00102$.

The quality of a search engine is to bring the "right" pages, in the "right" order. Although there are many web pages which cover a topic, some web pages become more preferred, that is famous. Google's page ranking algorithm is based on number of in-links to page and the rank of the pages that initiate the link [3]. A page gets more value if it gets more links and it gets links from a "valued" page. Incrementing the value of someone based on the reference of a valuable one reminds the trust mechanisms [4].

1.2. Large Networks

One of the early studies of social networks is done by Milgram which lead to the concept of "small-world". Watts and Strogatz come up with a model for "small-world" [6]. Power-law degree distribution, $p(k) \sim k^{-\gamma}$ where $p(k)$ is the probability of having degree of k, seems to be quite common in many complex networks including internet router, www, e-mails, movie actors as well as co-authorship of scientific papers or citations of scientific papers [7-12]. Number of citations a paper gets could be a measure of its popularity. Number of links a web page gets is another measure of fame. In who-knows-who networks, the more links one gets, the more popular she becomes. Another social network is studied in the co-occurrence in the news. Persons that occur in the news are represented as vertices and there is an edge if two persons occurred in the same news article [13]. The more a person occurs in news, the more famous she becomes.

2. Definitions

Population is made out of persons. The *size* $\boldsymbol{n}$ of the population is defined to be the number of persons in the population. $\boldsymbol{p_i}$ is the i'th person. A person has a *memory* which is a list of persons. The *size* $\boldsymbol{m}$ of the memory is the maximum number of persons that can be stored. The *memory ratio* $\boldsymbol{\rho}$ is defined to be the ratio of the memory size to the population size, that is $\rho = m/n$.

Total memory capacity of the population is the summation of the size of memories of the individual persons that is $n \times m$. A person p_1 *knows* person p_2 if person p_2 is in the memory of person p_1. The *knownness* $\boldsymbol{k_i}$ of a person p_i is the number of persons that "know" the person. If nobody knows a person p_i, that is $k_i=0$, then person is called *completely unknown*. The *fame* $\boldsymbol{f_i}$ of a person p_i is defined as the ratio of the *knownness* to the population size, that is $f_i = k_i/n$. Notice that if a person can not be repeated in the memory of a person, that is no duplication is allowed, then $0 \leq f_i \leq 1$, that is, f_i is normalized.

Person *learns* if he gets a new person in his memory. A person p_1 *remembers* a person p_2, if p_1 selects p_2 among persons stored in his memory. Person *forgets* if he removes a person from his memory. Persons interact by exchanging recommendations. Person p_g, called *giver*, *recommends* person p_0 to person p_t, called *taker*. Steps of recommendation process are i) p_g remembers p_0; ii) p_g gives p_0 to p_t; iii) p_t learns p_0. Assuming that no duplication is allowed and $m<n$, the steps of learning are i) remember some person; ii) forget him in order to obtain an empty slot; iii) learn the new person using this slot. It is important to note that learning, remembering and forgetting have implicit selection processes. Any model based on this concept has to specify these mechanisms.

3. A Simple Recommendation Model

Many models can be built on these concepts but the simplest one is obtained by defining the mechanisms based on pure random selections only. This model is called Simple

Recommendation Model (SRM). The giver and the taker are selected randomly. The giver randomly selects a person from his memory as his recommendation. This is the "remembering" process. The taker checks if she already knows the recommended person. If she remembers the person, then nothing is done. If she does not remember the recommended person, a memory location has to be freed. She randomly selects a person from her memory and "forgets" the person. Then, she "learns" the recommended person by storing the person into this location. Notice that the initial memory configurations are not specified within this model.

4. Approach

In order to find out the relationship between the fame and the memory, the simple recommendation model is implemented in Java with the parameters n, m and the number of reference exchanges. Different combinations of values of n and ρ are used including n=10^2, 10^3, 10^3 and ρ=0.5, 0.3, 0.2, 0.1, 0.05, 0.005. The number of reference exchanges is tried to be as large as possible. As n increases, it has to increase, too. It is started from 10^6 for n=10^2up to 10^{10} for n=10^4 which seems to be large enough. 10 simulation runs are made for different combinations of n and ρ values. Random selections required by the model are implemented by java.lang.Math.random() method which is a pseudo random number generator with a uniform distribution.

5. Observations

Initially, the memories of the persons are filled with persons selected randomly with duplications are allowed. Any two persons are expected to have almost the same knownness, hence every one has almost the same fame initially. At any given time, memory dumps of individuals provide information of about *who-knows-who*. The knownness of person p_i is calculated by counting the persons that know person p_i. The *knownness table* which consists of persons and their knownnesses is obtained. For better visualization, the data is sorted in descending order in knownness and the *knownness graph* is obtained. System comes to its steady state quite fast as seen in Figure 1.

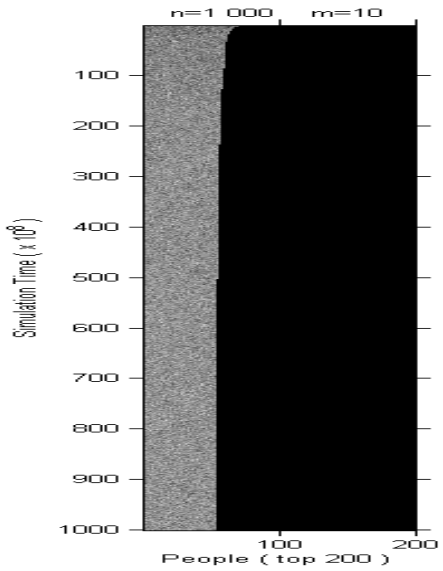


Figure 1. System settles down quite fast.

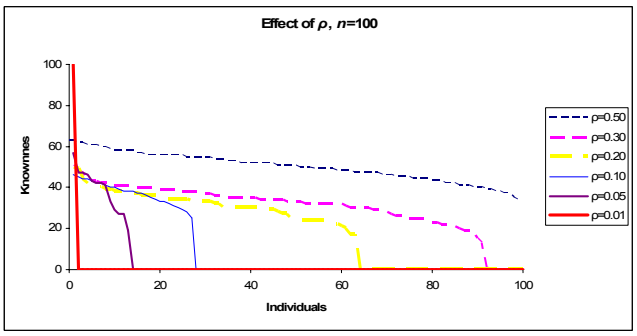


Figure 2. Effect of ρ as ρ goes to 0.

5.1. Effects of ρ

Firstly, the effect of ρ is investigated. Since there is enough memory space to remember everybody for $\rho \geq 1$, the cases of $\rho<1$ are studied.

Case ρ<1. Almost uniform distribution of fame that is observed at ρ=1 disappears as ρ decreases. Some people become more known whereas some become less known. Further decreasing ρ causes some people to be completely forgotten by the population. Figure 2 gives the effect of ρ to the

change in fame. For ρ=0.5, everyone is known in varying degrees. Completely unknown peoples start to appear around. ρ=0.3 and the number of completely unknown people increases as ρ decreases. It is important to notice that if a person becomes completely unknown, then there is no way for her to become known again.

Case *m*=1. As ρ further decreases, it eventually reaches the extreme case in which an individual can remember only one person, that is m=1. n is increased while m=1. For n=100 (ρ=0.01) only 1 person is known by the population where as for $n=10^3$ (ρ=0.001), there are around 50 known people. n=100 case can be explained by the small size of the population. Suppose this one slot is slightly dominated by a person p_i. Then p_i would be recommended more. Due to the rules of recommendation process, p_i replaces the position of other people. This further increases its dominance and as a result other persons would be removed from the memory of the population.

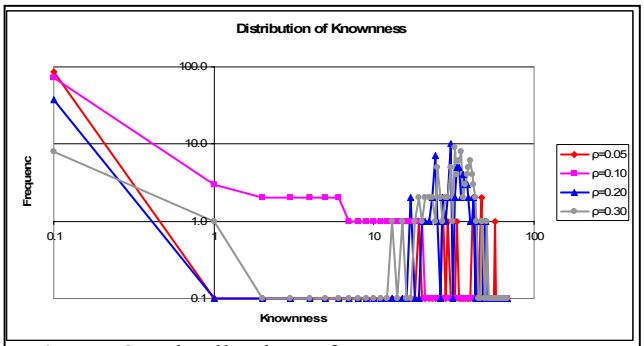


Figure 3. Distribution of Knownness

5.2. Distribution of Fame

One property that needs to be investigated is the distribution of fame. Figure 3 gives the distribution of knownness with respect to ρ. In order to show the pattern around origin, log-log scale is used. The zero values which create problems in log are replaced with an ordinary value, this time 0.1. The distribution of knownness in the initial memory is bell-shaped as expected. For ρ>0.50 both pre and post data are still bell-shaped. Around ρ=0.30, completely unknown persons start to appear. This pattern grows as ρ decreases to 0. The interpretation of the growing component near 0 is that less famous people in the population are increasing.

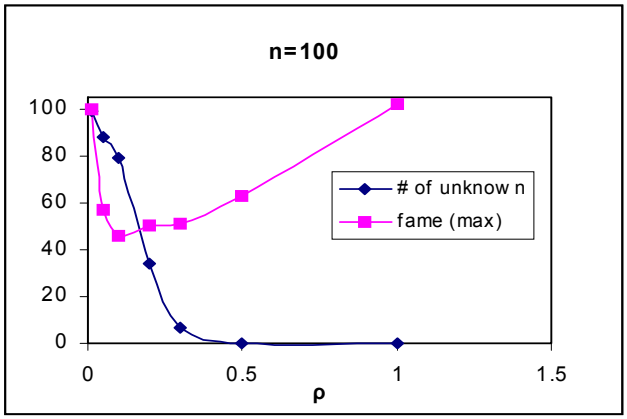


Figure 4. Change of fame.

5.3. Change of Fame

The minimum value of fame is 0 which corresponds to completely unknown case. The number of completely unknown keeps increasing as ρ approaches to 0.

The maximum value of fame has an interesting behavior as seen in Figure 4. The maximum fame slowly decreases as ρ goes from 1.00 to 0.10. It reaches a minimum around ρ=0.10. Then it rapidly increases as ρ approaches from 0.10 to 0. This pattern deserves an explanation. When ρ=1, that is m=100, everybody is known by everybody else so the fame is 1. As ρ decreases, the memory of the individuals decreases. Since no one dominates the memories yet, people are almost evenly distributed in the memories. So the reduction of the maximum fame is due to the decrease of the memory size. But as ρ keeps decreasing, after a certain point some people become completely unknown and some others become the dominating ones. As ρ approaches to the limit of 0, majority of the population become completely unknown and a few people dominate the memories. Those that dominate take all the references relinquished by the unknowns. So the rapid increase of maximum fame can be explained due to this positive feedback.

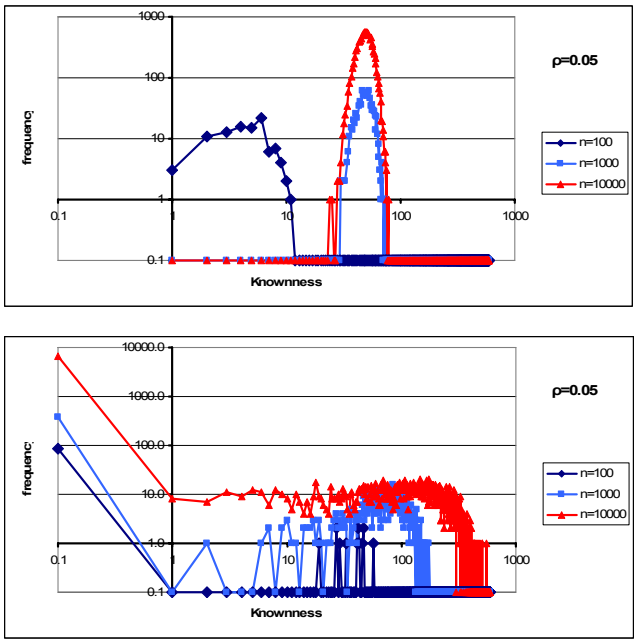


Figure 5. Effect of population size.

5.4. Effect of Population Size to Fame

Similar observations are obtained for values of $n=10^2$, 10^3 and 10^4. Figure 5 compares the distribution of fame $n=10^2$, 10^3 and 10^4 for $\rho=0.05$. In the first graph, the initial distributions are bell-shaped. After recommendation process, the shape of distribution changes as seen in the second graph. In order to better visualized the $\rho=0$ region log-log scale is used and zero values are replaced with 0.1 as in Figure 4.

5.5. Effect of Initial Memory

Population size is a parameter that is considered. Another important parameter is the initial memory configuration. So far the initial memory configuration is obtained by random selection with duplications allowed. One needs to be sure that the patterns observed are not due to this method. It could be the case that some fluctuations of the random number generation favor some individuals in the initial memory configuration and once this happens, recommendation model amplifies them. In order to test this, a) random selection without duplications allowed, b) completely symmetric initial configurations that favor nobody are employed. One of the simplest completely symmetric configuration is obtained by filling the memory of p_i with p_{i+1}, …, p_{i+m}. where $i+m$ is calculated in mod n, similar to model of Watts and Strogatz [6]. In this regular lattice structure every individual is guaranteed to have the knownness of exactly m. For both initial configurations similar patterns are observed.

6. Conclusions

A Simple Recommendation Model which changes the memory of the individuals is defined. The effect of the memory size m of individuals with respect to population size n is investigated. The ratio ρ of memory size to population size is used as parameter.

The value of ρ is changed from 1 to 0. For $\rho \geq 1$, everybody is almost known by everybody else. As ρ decreases, the knownness of some individuals decreases and eventually become 0, that is unknown by the population. As ρ gets close to 0, almost everybody becomes unknown. On the other hand, decreasing number of people becomes much more known. Since the number of people is much larger then the size of the memory of the individuals, ρ values close to 0 are realistic values. In this respect, the model explains why there are very few famous people and millions of people known by very few.

The model can be extended in many ways. The items stored in the memory were again people. The model is valid if some other items are stored, too. Web pages, dentists, people living in a country, scientific papers, radio stations and books could be the items to store. The model can be used to understand disappearing of cultural values such as traditions. Once a tradition becomes known by a few members of the society, it will quite rapidly disappear according to the model.

Culture can be thought as the memory of the entire population. Although the memory size of the individual is small, the total memory of the population, that is, the number of items in the memory of the population is much larger. On the other hand, population memory is less than the total memory capacity of the population that is $n \times m$. Since famous people repeatedly occur in the

memory of the individuals, number of different people known by the population, called effective population memory, is less then the total memory capacity. As ρ gets close to 0, effective population memory rapidly reduces.

"How to become popular" is another related question. Given that the population has already famous people, how do new comers become famous? What percentage of the population initially made to "know" the person by advertisement, so that she becomes famous? In this respect the model can be extended to a model for marketing and advertisement campaigns.

7. References

1. Simkin M. V., Roychowdhury V. P. (2003) Theory of Aces: Fame by chance or merit? *cond-mat/0310049*.

2. Bagrow J. P., Rozenfeld H. D., Bollt E. M., Avraham D. B. (2004) How Famous is a Scientist? - Famous to Those Who Know Us, *cond-mat/0404515*.

3. Brin S., Page L. (1998) The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems* (30), 107-117.

4. Yolum P., Singh M. P. (2004) Engineering Self-Organizing Referral Networks for Trustworthy Services Selection, *IEEE Transactions On Systems, Man, And Cybernetics*, 3, 396–407.

5. Milgram, S. (1967) The small-world problem, *Psychology Today*, **2**, 60-67.

6. Watts, J.D., Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks, *Nature* **393**, 440-442

7. Newman M.E.J. (2003) The structure and function of complex networks, *cond-mat/0303516*.

8. Dorogovtsev, S. N., Mendes, J. F. F. (2001) Evolution of Networks, *cond-mat/0106144v2*.

9. Albert, R., Barabasi, A. L. (2002) Statistical Mechanics of Complex Networks, *In Reviews of Modern Physics* **73**, 47-97, (cond-mat/0106096).

10. Kirlidog, M. Bingol, H. (2003) The shaping of an electronic list by its active members, *ITIRA'03*, 40-48.

11. Newman M.E.J. (2001) The structure of scientific collaboration networks, *Proc. Natl. Acad. Sci.* USA **98,** 404-409.

12. Redner, S. (1998) How Popular is Your Paper? An Empirical Study of the Citation Distribution, *cond-mat/9804163*.

13. Ozgur, A., Bingol H. (2004) Social Network of Co-occurrence in News Articles, *LNCS 3280*, eds. Aykanat C. et al. (Springer, Heidelberg), pp. 688-695.

Acknowledgements. The author would like to thank Albert Ali Salah and Arzucan Ozgur for helping in Matlab programming, Cem Ersoy for invaluable comments. This work was partially supported by Bogazici University Research Projects under the grant number 05A105. An early version of this paper will be presented in ISCIS'05 in October 2005 in Istanbul.

Metamimetic games
Modeling metadynamics in social cognition

David Chavalarias[1],
Center for Research in Applied Epistemology (CREA),
Ecole Polytechnique, Paris, France,
www.polytechnique.edu

Abstract :

Imitation is fundamental in the understanding of social system dynamics but the diversity of imitation rules employed by modelers proves that the modeling of mimetic processes cannot avoid the traditional problem of endogenization of all the choices, including the one of the mimetic rules. Starting from the remark that human reflexive capacities are the ground for a new class of mimetic rules, I propose a formal framework, metamimetic games, that enable to endogenize the distribution of imitation rules while being human specific. The corresponding concepts of equilibrium - counterfactually stable state - and attractor are introduced. Finally, I give an interpretation of social differentiation in terms of cultural co-evolution among a set of possible motivations that departs from the traditional view of an optimization process indexed to criteria that exist prior to the activity of agents.

Keywords :
Social cognition, imitation, cultural co-evolution, differentiation, reflexivity, metacognition, stochastic game theory, endogenous distributions, metamimetic games, counterfactual equilibrium.

---

[1] david.chavalarias@polytechnique.edu ; home page : www.chavalarias.com

✉ : Ecole Polytechnique, 1, rue Descartes, 75005 Paris – France. ☎ : (+0033) 1 55 55 86 23

I. Social system modeling and the grounding problem

I.1 An interdisciplinary convergence on imitation

Recent years have seen a burgeoning of interest in the phenomenon of imitation on the part of researchers in many fields. After the seminal work of Tarde (1890) in sociology and Baldwin (1897) in developmental psychology, Girard (1961) takes imitation as the first foundation stone of his theory of the origin of primitive societies, and Donald (1991) considers that the sophistication of human mimetic skill could have been one of the major evolutionary transitions in hominization.

After three of the finest economists of all time, Smith, Keynes and Hayek gave imitation a central role (Dupuy 2004), imitation is now more and more often taken into account in the study of micro-economic behaviors as information and communication technologies spawn a proliferation and speed-up of feedback loops in business and finance (Orléan 1998 ; Frank 2003 ; Selten & Ostmann 2001). Today, the study of imitation is a dynamic field of research in developmental psychology (Meltzoff 2002 ; Gergely et al. 2002 ; Zelazo & Lourenco 2003) with bridges toward neurology (Arbib 2000 ; Chaminade et al. 2002) and theory of mind (Meltzoff & Gopnik 1993 ; Pacherie 1998). Finally, in artificial intelligence and robotic, imitation is now seriously considered as a means of constructing socially intelligent artificial agents (Brezeal et al. 2005 ; Jansen et al. 2003 ; Zlatev 2000). Evidence that stems from ethology (Tomasello 1999) about the incapacity of animals to imitate in a human way will certainly not weaken this interest for imitation in the study of human behavior.

This convergence of interests has had some consequences in the literature of socio-economic modeling where modelers, to account for the extremely rich structures observed in human societies, more and more often incorporate mimetic processes into formal approaches. But the diversity of imitation rules employed by modelers proves that the introduction of mimetic processes into formal models cannot avoid the traditional problem of the endogenization of all choices, including the choice of the imitation rules.

I.2 Which rules are the "good" ones ?

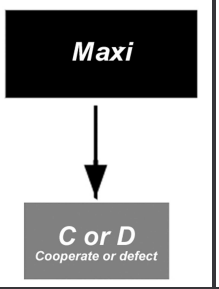
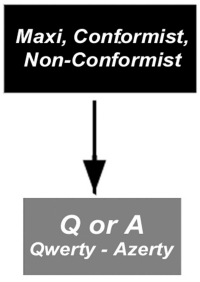
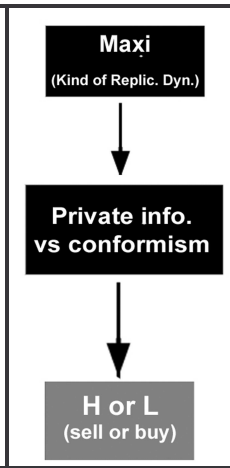
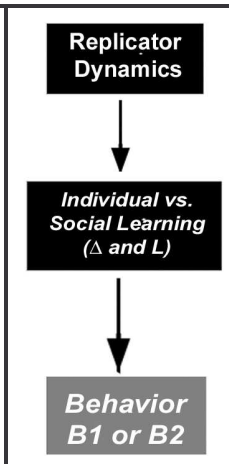
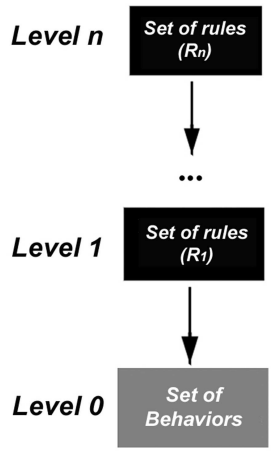
In the literature of social systems modeling, the most frequent types of rules are these : (1) *payoffs-biased imitation,* i.e. imitation of the most successful agents in one's neighborhood (Nowak and May 1992) and (2) *conformism* (Axelrod 1997 ; Galam 1998 ; Orléan 1985, 1998, Rogers 1988). Here, *conformism* is the propensity of individuals to adopt some behavior when it has already been adopted by some of their neighbors, the propensity being relative to the frequency of that behavior in the neighborhood. To a lesser extent, other imitation processes have been studied, among which we can mention (3): *non-conformist*, the propensity of an individual to adopt the behavior of the minority (Arthur 1994 ; Edmonds 1999) and (4) *prestige* (Henrich & Gil-White 2001). This list of imitation rules is far from exhaustive, and we should note that even for *conformism* or *payoffs-biased imitation*, several technical definitions have been proposed both deterministic and probabilistic (Nowak et al. 1994). Moreover, it is also possible to propose models that include several imitation rules, as some authors already did (Boyd & Richerson 1985 ; Henrich & Boyd 1998 ; Janssen & Jager 1999 ; Kaniovski et al. 2000 ; Vriend 2002).

This raises an epistemological question for modelers. Which rule(s) for imitation should be considered given the situation under study?
Let's try be a little more precise. If we represent schematically the models cited above we can remark that all can be described in terms of hierarchies of rules governing a behavior (cf. table 1. for some examples). The proportion of the rules at a given level is determined by the metarules of

the level above. In this representation, rules can be interpreted as a dynamics principles acting at the population level (like the replicator dynamics) or as a decision-making rule used at the individual level. The emergence of patterns at the collective level is thus understood as a selection of a particular distribution on the set of possible rules and meta-rules. Now the question is: *How are these distributions are selected ?*

Some scholars have addressed this question from an evolutionary perspective, assuming that the distribution of imitation rules is shaped by natural selection (see for example Henrich & Boyd 1998 : Table *1-d*). On this model, there is a unique selection rule, indexed to fitness, that drives the entire system from the top.
But the slow dynamics of genetic processes seems to be incompatible with the fast evolution observed in human socio-economic systems (Feldman & Laland 1996, Frank 2003, Gould 1987, Gintis 2003) that varies on the scale of a lifespan. The challenge is thus to identify an evolutionary process that could lead individuals (or firms in the case of economic systems) to choose between several possible imitation rules while interacting with their environment. Rather than considering agents whose imitation rule exists prior to their socio-economic activities, we have to imagine agents who's way of being influenced by others is the result of a historical process, their interactions with their social network.

Table 1 : Schematic representation of models

				
a. : Nowak et May 1992.)	**b. : Kaniovski et al. 2000.**	**c. : Orléan 1998.**	**d : Henrich et Boyd 1998.**	**e : Modelers in the social sciences generally represent agents as a hierarchy of rules, where rules at each level evolve under the dynamics defined by their meta-rules.**

Here a strange loop appears: the distribution of imitation rules in a population is the consequence of interactions among agents ; the consequences of interactions among agents are defined in terms of their imitation rules. We face a grounding problem concerning the dynamics of imitation rules in the true sense of the term. If we define *social cognition* in terms of a collective process of information distributed over all the individuals of a society (Bourgine 2004), themselves processing information using their rules for decision making processes, we have to study the dynamics of this information processing which is a metadynamics relative to the agents' rules. Can dynamics and metadynamics coincide ? This leads us to reformulate our question in terms close to the notion of *operational closure* (Varela 1983) : *Is it possible to endogenize the distribution of metarules of decision-making such that this distribution becomes the outcome of the dynamics it defines ?*

The point of this article is to demonstrate that there exists at least one class of models that allow us to give a yes answer to this question. We will show on the way that considering the specificity of human imitation, an imitation rule can be its own metarule in formal models.

I.3 The specificity of human cognition

Another way to address the question of the endogenization of imitation rules will perhaps come from a recent concern in social systems modeling. The complexity of human social systems has no counterpart in other species. For example, considering group coordination, only insect societies, composed of very simple entities, have social structures involving several thousand members. This feature disappears as soon as the repertoire of behavioral possibilities of a species gets wider, and reappears only when it comes to humans (Bourgine 2004, Wilson 1975).
This remark is noteworthy because it is precisely the modeling of self-organized systems in ethology that has been a precursor for multi-agent modeling in the social sciences. It is clear that the goal for social systems modeling is not to consider humans as cloned insects. What is at stake is rather to find differences between humans and others mammals which make possible the emergence of highly structured social groups while keeping inter-individual heterogeneity. This has led recently some modelers to propose, as a heuristic for social systems modeling, prioritizing models that could be human specific (Alvard 2003, Bowles & Gintis 2003, Fehr & Fischbacher 2003). In the social sciences, a similar heuristic that particularly focuses on imitation, was formulated few decades ago by René Girard[i] (1978):

> *In order to develop a science of man, we must compare human imitation with animal mimetism and separate the modalities specific to humans, if they exist, from mimetic behaviors.*

Following this heuristic, I will quickly evoke some differences between animal and human cognitive capacities that could have a qualitative impact on imitation processes. From numerous studies in psychology, philosophy, neurology and ethology, two elements appear to play a crucial role in human behavior while being apparently out of the reach of non-human cognition.

First, humans are reflexive beings. To give a low level definition of reflexivity, it is the ability to take as an object of cognitive treatment the cognitive treatments themselves by creating new levels of cognitive processing. Emergence of reflexive capacities can be traced through ontogeny through the study of the development of infant cognitive capacities (Zelazo et al. 1996) and the self-triggered loop that should be the elementary component of reflexive processes is closely linked with the constitution of the self (Damasio 1999, Donald 1991, Mounoud 1995). Reflexivity allows us to think of others as we think of ourselves and ourselves from other's point of view and thus develop our social skills. From the point of view of imitation processes, reflexivity makes all the difference since, as Eric Gans (1995) says, *"prehuman imitation is non-reflexive; the subject has no knowledge of itself as a self imitating another"*.

The second difference between animal and human cognitive capacities, closely related to reflexivity, is metacognition (Jacob 1998 ; Sperber 2000, Tomasello 2000), defined here as cognition about cognition. Whether animals have metacognitive capacities is still in debate in the scientific community. Some experiments seem to indicate that great apes and dolphins may have some rudimentary metacognitive capacities (Smith et al. 2003, Rendell and Whitehead 2001), but those are very limited. In particular, there is no evidence that animals can consider learning or imitation processes as object of cognition, and the fact that they do not teach tends to prove the contrary. Moreover, to our knowledge, there is no evidence that animals could voluntarily add a metacognitive level to solve a given problem, although some primates seem to be able to deal with chains of hierarchically organized behaviors (Byrne 1998). This means that animal metacognition, if it exists, is most probably constituted of rigid chains of process monitoring that could as well be hardwired, without requiring reflexivity to monitor their structure.

There is no space here to give more details about these two differences. But I will try to show that taking them into account makes it possible to build a new class of models that may offer an answer to the problem of the endogenization of the distribution imitation rules.

II Reflexivity of imitation rules

To see what metacognition and reflexivity could change in the modeling of mimetic behavior, a more precise definition of an imitation rule is needed. Before this, we have to give a framework for the representation of agents.

II.1. Agents

We will refer to an agent by the pronoun "it." In evolutionary games or in multi-agent modeling, agents are usually defined by a *n-tuple* τ of traits (age, color, opinion, behavior, rules of behavior, etc.) taken from a multi-dimensional set of traits *T*. In the following, we will place these traits into two categories : *modifiable traits* and *other traits*.

- ***Modifiable traits*** are those an agent can change voluntarily like for example a cooperative vs. defective behavior, the colors of clothes it wears, the political party an agent decides to vote for, the learning rules adopted for a given task, the chair where it wants to sit, etc. Most of the time, these kinds of changes take place on small time scales (within a day). The set of modifiable traits of a given agent *A* will be called its strategy and can be represented by an ordered *n*-tuple $s_A=(\tau_1,\ldots,\tau_n)\in S$.
- ***Other traits*** are those that do not entirely depend on agent's will or are immutable. They depend on global dynamics and change generally on large time scales (months, years, lifetime), like social positions, payoffs, reputation, prestige, age, color of eyes, etc.

Agents are usually embedded in a social network and can learn some of the traits of the agents they interact with. For a given agent *A*, the neighborhood Γ_A will be defined as the set of all agents from which *A* can learn some traits.
Agents can then categorize their neighborhood into sub-neighborhoods on the basis of the learnable traits. For example, Boyd and Richerson (1985) consider different types of cultural transmission processes within sub-neighborhoods indexed to age and kinship: vertical transmission from parents to offspring, oblique from elders to younger, and horizontal among peers.

II.2. Imitation rules

Roughly speaking, imitation occurs when an agent decides to adopt a trait observed in one of its neighbors. For example, *A* may want to behave like *B,* wear the same clothes, adopt its opinion or sit where *B* is sitting. The most general definition for an imitation rule is thus a process that takes as input an agent *A* and its neighborhood and gives as output a modifiable trait *A* will try to copy from some of its neighbors.

From this broad definition, two distinct conceptions of imitation can be derived. Following René Girard (1961), *A* may want to be like *B* in some respect because *A* reads the values of things and actions in the eyes of others. In that case, imitation precedes desire, and, from a formal point of view, it can be represented by some kind of conformist rule that governs desires (see Orlean 1985 for an example). The second conception of imitation is teleological and widely used in economics (Frank 2003) and multi-agent modeling (Conte & Paolucci 2001): *A* may want to be like *B* to a certain extent because from *A*'s point of view, *B* is a good model, i.e. *B* is most successful than *A* according some criteria that *A* has adopted *independently* of knowing *B.* What I will propose now is in-between and requires us to refine the above definition of "imitation rule" in a formal perspective.

Definition: ***Imitation rule***
Given an agent A and its neighborhood Γ_A, an imitation rule is a process that :

1. Assigns a value $\nu(B, \Gamma_A)$ in an ordered set (the set of real numbers for example) to each agent B in Γ_A. ν will be called a valuation function.
2. Selects some traits to be copied from the best agents (according to the values given in 1) and defines the copying process.

For example, in payoff-biased imitation, the value assigned to each neighbor is its payoffs. The agent has then to infer which of the traits of the most successful neighbor is responsible for this success and try to copy it. In the case of a *conformist* rule, the value assigned to a neighbor is the size of the group it belongs to, and the traits to be copied are those of the largest group.
The valuation function here is subjective and dynamic, and it plays a role analogous to the utility function in game theory. Two agents can have a different valuation function and can evolve in this respect. The diversity of valuation functions in a population expresses the diversity of points of view.

The goal here is not to be exhaustive, and this definition leaves a lot of things in the shadows, like for example the problem of inferring of the relevant traits. Nevertheless, it is sufficient for our purpose which is to propose a framework for thinking about an endogenous distribution of all kinds of decision-making rules evolving at the cultural time scale.

II.3. Rules as modifiable traits

The fact that human beings have reflexive and metacognitive abilities has some important qualitative consequences when it comes to modeling: agents know in some extent that they are using rules for decision-making. This applies to imitation rules and therefore imitation rules can be viewed as part of the strategy of the agent. They become modifiable traits. We should consequently study systems where imitation rules are modifiable traits by way of applying meta-rules.

This leads us to model agents as hierarchies of rules. A quick argument will convince us of the legitimacy of this representation. Considering the set of all rules an agent A is currently using, we can define a relation $\mathfrak{R}$ in the following way. Let r_1 and r_2 be two rules used by A in its decision-making process, we will say that $r_2 \mathfrak{R}\ r_1$ if the use of r_2 can change the way A uses r_1, which is equivalent to saying that r_2 acts upon r_1. For example, in the model of Orléan (1998, Table 1-c), the top level rule r_2 (a kind of replicator dynamics) can change the rule r_1 used at level one (conformist rule or individual learning). We thus have $r_2 \mathfrak{R}\ r_1$. On any set of rules, $\mathfrak{R}$ defines a partial order that enable us to define a hierarchy of rules acting one on another.

For sake of clarity, we will consider in the following discussion agents such that $\mathfrak{R}$ defines a total order on their set of rules. Moreover, as in most models already mentioned, agents will be defined by a unique type of behavior at the lower level controlled by a hierarchy of rules with a unique rule at each level. We will call such a hierarchy a metamimetic chain. As for models presented in table 1, we can associate a given modifiable trait with a chain of imitation rules that controls its evolution.

Since our aim is to approach some aspects of human collective behaviors, we have to respect what are currently accepted as modeling constraints. In particular, agents must have a bounded rationality. The consequence is that metamimetic chains have to be finite. This leads us to define the maximum length for such chains : the cognitive bound of the agents (c_B). Agents can modify the composition and the length of their metamimetic chain as long as the latter is inferior to c_B.

As for the top-level rule, we face two possibilities. Either we postulate a fixed exogenous rule, which is the option taken by game theory and evolutionary game theory. In that case, top-level

rules are interpreted as genetic determinants or fixed preferences (that have to come from somewhere).
The other possibility is to consider that top-level rules are also modifiable traits, agents can act upon them. From our definition of $\Re$, this is possible only if we assume that r$\Re$ r when r is in a top-level position. We will now show that the above definition of imitation rule allows to illustrate this view and more precisely, that we can give some sense to the fact that $\Re$ is a reflexive relation i.e. $r_k \Re r_k$ for all k.

III. Metamimetic dynamics and endogenization of meta-choices

III.1 Agents as metamimetic chains

In this section, I will explore some possible dynamics in a population of agents described in terms of chains of imitation rules, the relation $\Re$ being reflexive. For reasons of clarity and because they have been extensively studied, we will not evoke other characteristics of the decision-making process like anticipation, learning or memory, although they plays an important part and should be taken into account in future work. What I am trying to catch are the remarkable properties of metamimetic dynamics.

Consider a population of agents defined by metamimetic chains that can deal with a maximum of c_B meta-levels (bounded rationality). Consider that agents have to choose a behavior r_0 among several possibilities (like *C or D*). This behavior has some impact in their environment, like, for example, an influence on their material payoffs or on the densities of the different behaviors in their neighborhood. Assume that agents can change their behavior with metamimetic chains composed of rules taken in a set *R*. Agents will then be defined by a set of modifiable traits[ii] $(r_0,r_1,.. r_n)$ with the constraint $n \leq c_B$, where r_0 is a behavior and $r_j \in R$ for $j>0$ (figure 1).

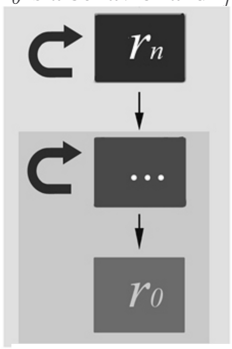


Figure 1 : Schematic representation of a metamimetic chain.

With this representation, the activity of an agent will consist in the ongoing verification of the coherence of its hierarchy of decision-making processes, focusing its attention on different levels at different moments. For example, if the agent focuses its attention to level *k,* since for $j \leq k$ $r_k \Re r_j$ ($\Re$ *is reflexive)* the question it will ask itself will be : "Is the strategy $(r_0,r_1,.. r_k)$ the best one from the point of view of r_k *?"*. If this is not the case, the agent will try to change its strategy for a better one. Except for the *status quo* case we can identify two categories of possible outcomes in the revising of $(r_0,r_1,.. r_k)$:

- 1) Some elements in $(r_0,r_1,.. r_{k-1})$ are modified but the initial rule r_k is still part of the strategy.
- 2) r_k has been modified.

We will now examine these two possibilities. To fix ideas, in our examples, we will consider agents with two different opportunities for action *C* and *D* that have two different material

consequences (monetary payoffs for example). As for the set of rules *R,* we will take the most common rules in the modeling literature: imitation of the most successful agent in terms of material consequences (*payoffs-biased* or *maxi* rule for short) and imitation of the most common traits (*conformist* rule) : *R={maxi, conf.}*. It should be emphasized that payoffs here should not be interpreted as a direct mapping of what could be the utility function of agents. The analog to utility functions is the hierarchical set of valuation functions of a metamimetic chain.

III.2. Endogenous variation of length in metamimetic chains

The standard way of activating a rule at level *n* is to change the trait of the level below. For example in Orléan 1998, an agent *A* with $s_A=(r_0,r_1,r_2)$ (can decide to change its behavior r_0 from *H* to *L* when it takes the point of view of its rule r_1 (for example the conformist rule) (cf. figure 2). But it can also change its rule r_1 when focusing its attention on the rule r_2, i.e. *maxi* if it happens that r_1 is not the most successful rule in its neighborhood (in this article, a random sample in the population).

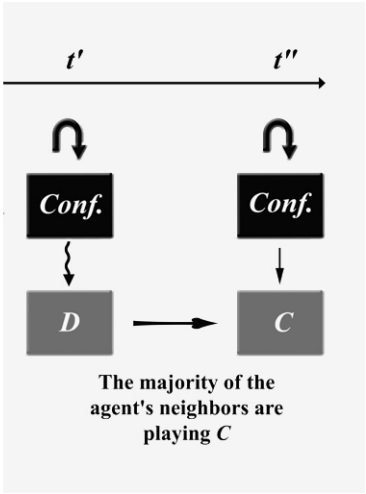


Figure 2: Update of an intermediary modifiable trait. A conformist Agent *A* observes that the majority of agents are playing *C,* and decides to update its behavior to *C (modifiable trait of level 0).*

When we consider agents that can monitor the complexity of their strategy by adapting the composition and the length of their metamimetic chains, this standard way of using rules should be generalized. Let's begin with a simple example.
Assume that after the observation of a conformist agent *B*, *A* with $s_A=(D, maxi)$ comes to the conclusion that to maximize its payoffs, the best thing is to do like *B*. In that case, if its cognitive bound is large enough, the most rational behavior is to adopt the strategy $s_A'=(D, conformist, maxi)$, and act as a conformist as long as this rule is adaptive from the point of view of the *maxi-rule* (cf. figure 3) and can eventually come back to its initial strategy *(* , Maxi)* in a subsequent imitation.

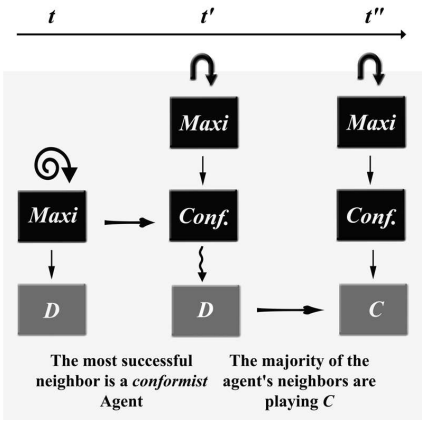


Figure 3: Endogenous variation in the length of the metamimetic chain. At time *t*, a *Maxi* agent *A* has a *conformist* neighbor that is more successful than all agents in Γ_A. *A* will then adopt the conformist rule at its first meta-level, keeping in mind that it is only a means for maximizing its payoffs (second meta-level). Thereafter, it might be that according to this rule, the current behavior is not the best one and has to be changed.

In that case, the complexity of *B's* strategy and the cognitive bound of *A* enable *A* to keep in mind its initial rule r_1. *B's* strategy is a temporary means for achieving the goals defined by r_1. This kind of transition enables the agent to change endogenously the length of its metamimetic chain. We do this kind of mental operation every day every time we decide that the realization of a goal *G'* is the best way to achieve a goal *G*.
More generally, the activity of an agent consists in the ongoing verification of the coherence of its hierarchy of decision-making processes given its environment, focusing its attention on different levels at different moments.
Thus, an agent $A=(r_0,r_1,..\ r_n)$ can focus its attention on level *k,* and ask itself: "Is the strategy $(r_0,r_1,..\ r_k)$ the best one from the point of view of r_k *?"* Then it might happen that *A* is not the best agent in Γ_A according to ν_k. In that case, *A* will try to copy one of its best neighbor and this can lead *A* to modify the length of its metamimetic chain by changing its sub-chain $(r_0,r_1,..\ r_{k-1})$. Thus we have an endogenous variation in the length of metamimetic chains.

III.3. Reflexive update at the limit of the cognitive bound

The metacognitive skill that enables agents to change the structure of their strategy for a more adaptive one has its counterpart also. Take for example an agent *A=(D, maxi)* with a cognitive bound of *1* that finds out that one of its *conformist* neighbors has higher payoffs than any other neighbor, *A* included. To try to be as successful as *B, A* will have no other solution than to *become conformist* (figure 4).

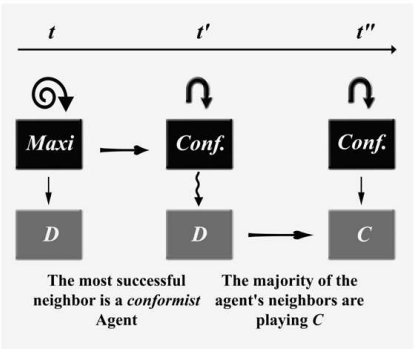


Figure 4: ***Reflexive update at the limit of the cognitive bound.*** At time *t,* a *Maxi* agent *A* has a *conformist* neighbor that is strictly more successful than all other neighbors. Consequently, *A* adopts the *conformist* rule. Thereafter, it might be that according to this rule, the current behavior is not the best one, and has to be changed.

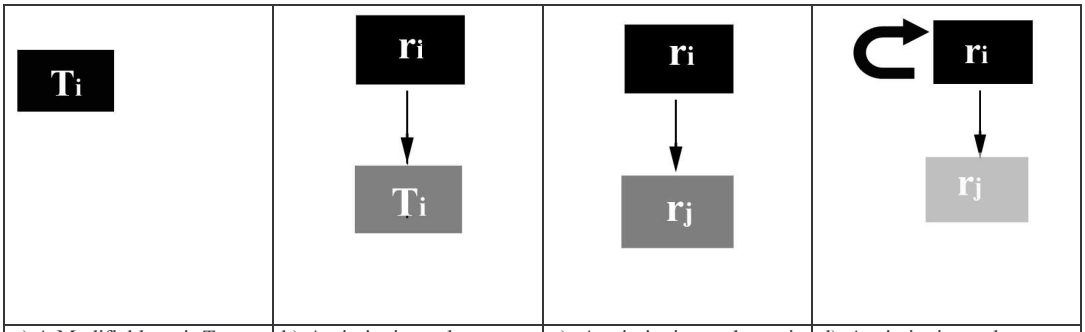
A metamimetic agent will encounter this kind of situation each time the strategy of its model is too complex to manage both new goals and old ones. In that case, *r* will act like its own metarule. In that case, an agent *A* changes reflexively its rule *r* because, **from the point of view of *r***, *r* is not the most adaptive rule. We can say that *r* is not self-coherent in *A's* environment: it prescribes actions that are in conflict with the continued possession of *r.*
In such cases, and when the rule *r* is the top-level one, *A* will have to change its top-level rule. This is possible thanks to our definition of an imitation rule: since imitation rules are modifiable traits, it might happen that the trait to be modified in an imitation process defined by a rule *r* is the rule *r* itself. In this sense, top-level rules are their own metarule. We will say that an imitation rule can *update reflexively by acting on itself as a modifiable trait.*

We can now comment on reflexive updates. People certainly don't have a wired cognitive bound that obliges them to do such clear cut transitions in their decision processes. Nevertheless, it often happens that an activity that was first considered as a means becomes an end in itself. For example it might come to finally take up all our time ; we might forget its primary purpose or we might simply come to like the new activity more than any other. In these examples, the important feature is that the new goals do not come from nowhere but are related to the old ones to some extent. The reflexive mimetic update presented above is a stylized representation of these kinds of transitions when they are triggered by the observation of others.
Viewed at the population level, these kinds of updates define an endogenous dynamics on top-level rules the characteristics of which should be studied.

IV The metamimetic game

IV.1 Definition

The characteristics of imitation rules introduced in the preceding section (figure 5) suggest a new class of models for the modeling of mimetic dynamics**.**



a) A Modifiable trait T_i	b) An imitation rule can act upon a modifiable trait T_i	c) An imitation rule r_j is itself a modifiable trait upon which can act a rule r_i	d) An imitation rule r_i can act reflexively. It is a modifiable trait for itself.

Figure 5 : When imitation rules are modifiable traits, they can be modified by other rules (c) and can be modifiable trait for themselves (d).

Definition: ***Metamimetic game***
Let B be a set of behaviors and R a set of imitation rules, a *metamimetic game* $G=\{N,\Gamma,R,B,C_B\}$ *is an N-player game where each agent A is characterized by a metamimetic chain* $s_A=(r_0,r_1,..\ r_k,)$ with $r_0 \in B$ and $r_j \in R$ for $j>0$.

Moreover, the three following conditions should be satisfied:

*C-I - **Bounded rationality**: the number of meta-levels in a metamimetic chains is finite and bounded for each agent by its cognitive bound c_B ($k \leq c_B$).*
*C-II - **Metacognition**: at all levels in a metamimetic chain, imitation rules are modifiable traits.*
*C-III – **Reflexivity**: imitation rules can update reflexively changing the length of the metamimetic chain in the limit of the cognitive bound of agents. When the cognitive bound is reached, imitation rules may update themselves.*

The study of metamimetic games will thus consist in the study of the evolution in the length and composition of these chains, leading to the emergence of structures at the intra and inter-individual levels.
The main difference with other kinds of games from game theory and evolutionary game theory is that in metamimetic games, there is an endogenous dynamics on the distribution of rules and metarule whatever the dimension of the rule space (even the top level is not a singleton). There is not enough space here, but by writing the master equation of this kind of games, we can demonstrate that contrary to other games of imitation like replication by imitation (Weibull 1995) metamimetic games are not in general reducible to some standard dynamics like the replicator dynamics. In fact, the discrete replicator dynamics (Hofbauer & Sigmund 1988) is reducible to a particular case of metamimetic dynamics where the set of possible top-level rules is a singleton.
We will now make intuitive these differences on a minimal example. This will allow us to introduce the main concepts related to these games.

IV.2. A minimal metamimetic game

Consider the following metamimetic game :

- *Two agents A and B with a cognitive bound of 1 and no memory.*
- *Each agent is in the other's neighborhood.*
- *Two possible actions C and D.*
- *Two possible imitation rules : maxi and conformist*
- *The game is symmetric. When C plays against D, D gives always higher payoffs than C (think of a prisoner dilemma for example).*
- *The game is repeated and at each period, agents change their strategy simultaneously according to their rule.*

The definition of the rules are the following :

- *Maxi* : if your neighbor has higher payoffs than you, copy its rule and use it to update your behavior.
- *Conformist* : If your neighbor has a strategy different from you, copy its rule and then use it to update your behavior.

The state of the game is thus given by the behavior and the imitation rule of each agent like for example : $s=[s_A:(C, maxi) ; s_B: (D, conf.)]$.
There are only 16 possible states and the metamimetic dynamics defines a Markov chain on this set (figure 6).
For example if the initial state is *s=[(C, maxi) ; (D, conf.)],* after one period, A will become conformist because B is the most successful agent, and will change its behavior from *C* to *D* to adopt B's behavior. B will become *maxi* to be like A and will keep on playing *D* because it is the most successful action. In the second period, both agents will have the same behavior and consequently will have the same payoffs. Then only A will change its strategy to be like B and both will end *(D, maxi).* The final state of the game will be *s"=[(D, maxi) ; (D, maxi)].*
We will say that s" is reachable from both states *s=[(C, maxi) ; (D, conf.)]* and *s'=[(D, conf) ; (D, maxi.)]*. More generally, we will say that a state s' is reachable from a state s if and only if a system starting in state s can reach the state s' after a finite number of mimetic transitions.

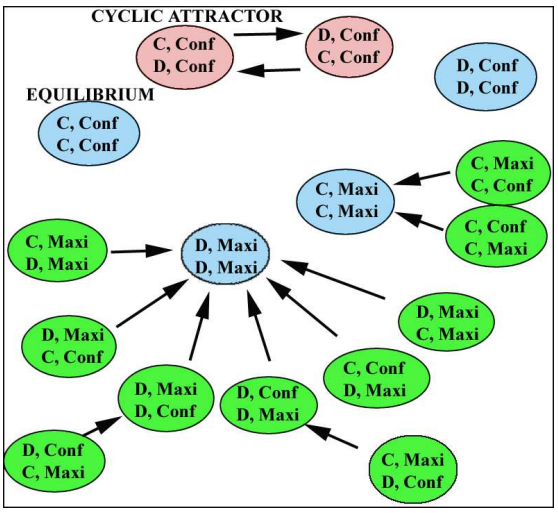


Figure 6 : The dynamics in a minimal metamimetic game. Each arrow represents the mimetic transition that the current state requires (here with a probability 1 everywhere).

IV.3. Metamimetic equilibria

Two types of remarkable subsets of states in figure *6* should be highlighted, equilibria and attractors. They are the remarkable subsets of the underlying Markov chain.

Definitions :

Let $G=\{N,\Gamma,R,B,C_B\}$ be a metamimetic game with $C_B=1$ *and* Ω *the set of all possible states of the game*.

- A set of states $\Sigma=(s^1,\ldots, s^m)$ *is a* ***metamimetic attractor*** *if and only if*
$\forall(s,s')\in \Omega\times\Sigma$*;* s' is reachable from $s \Leftrightarrow s'\in\Sigma$

- A state $s=(s_1,\ldots,s_n)$ is a ***metamimetic equilibrium*** if and only if
$\forall i\in\{1,\ldots,N\},\ \forall k\in\Gamma_i,\ s_i=(r_{0,}^{\ i}\ r_1^{\ i})\neq s_{\ j}=(r_{0,}^{\ j}\ r_1^{\ j}) \Rightarrow v^i(j,\ \Gamma_i)\leq v^i(i,\ \Gamma_i)$

"no agent can find itself better when it imagines itself in the place of one of its neighbors"

where $v^{\,i}$ is the valuation function associated with the imitation rule $r_1^{\ i}$. A metamimetic equilibrium is a particular case of metamimetic attractor. Here the value $v_i(j,\ \Gamma_i)$ can be understood as a counterfactual, it is the well-being that *i* can imagine feeling while being put at the place of *j*. In this sense, a metamimetic equilibrium will also be called a *counterfactually stable state*. This definition can be extended to any value of C_B.

This very simple example has the advantage of making clearly visible the main property of metamimetic games as illustrated in figure *6*: the Markov chain that defines the dynamics is a property of the set of strategies considered, not a dynamics that would be given apart from this set, like for example a *maxi* rule or a replicator dynamics applied to a set of strategies. We will further discuss the choice of this set in paragraph III.3**.** This property is independent of the size of the set of rules and is the key for the expression of a multiplicity of agents' viewpoints. At this stage, we can clearly see the link with the notion of *operational closure* : the dynamics on imitation rules for a given state of the game is the product of the distribution of imitation rules.

Moreover, the goals of an agent at a given moment, defined here as the valuation function of its unique imitation rule, are the outcome of a historical process, its interactions with its environment. We shifted from a perspective where goals are an unchanging property of agents (like the standard maximization of predetermined payoffs in economy) to a perspective were goals are in some

extent chosen by an agent during its life. The distribution of goals in a population should then be understood as the expression of the self-coherence of these goals in the social network and not the expression of the fitness of these goals relatively to pre-given criteria.

To return to the comparison with replicator dynamics, for the latter, the meaning of a transition from a global state *s* to a global state *s'* should be interpreted in terms of a semantics that is external to the system, the fitness function (what is good, what is bad). The different states of the system have no meaning other than the one assigned from the outside. On the contrary in metamimetic games, the meaning of a transition from *s* to *s'* has to be found in *s* itself, it is the expression of the content of imitation rules.

We can sum up these remarks with the following proposition :

Proposition : *Every metamimetic game* $G=\{N,\Gamma,R,B,C_B\}$ *can be associated with a unique matrix* P^0 *that defines the Markov process representing the internal dynamics of the game.* P^0 *defines de metadynamics of the social cognition process.*

The proof of this proposition is straightforward since for each configuration of the social network, an agent's imitation rules define locally the possible transitions (I remark in passing that transitions do not have to be deterministic since for a given agent there could be several models with distinct strategies).

The internal dynamics of a metamimetic game defines some particular distributions corresponding to eigenvectors of P^0. This is a first spontaneous selection among all possible states. In our example, from the sixteen possible states, only six are attractors. As we will see, this selection is sharpened in presence of perturbations.

IV.4. The role of perturbations

In the preceding section, agents were supposed to be mind-readers: they knew perfectly well the strategies of their neighbors. This of course does not pretend to reflect reality since in real setting, people have to infer rules, behaviors and other cognitive components of the decision making process of their neighbors. These inferences about what others think and do are noisy. Moreover agent do not always do what they intend to do. Consequently, there are errors all along the decision-making process due to false perceptions, misunderstanding and mistakes. Following Young (1993, 2001), a more realistic approach would be to suppose that there are some mistakes that constantly perturb the social dynamics. We have then to study metamimetic dynamics in the framework of stochastic game theory (Foster & Young 1990).

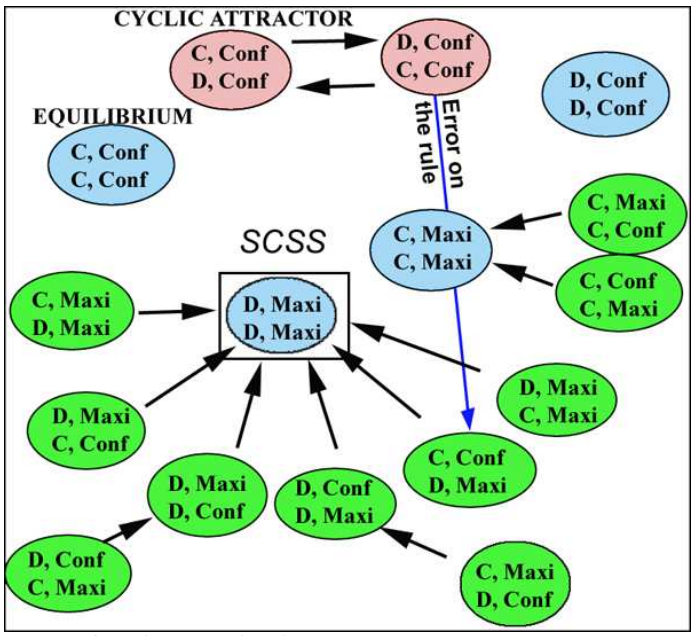


Figure 7: Minimal example of a noisy metamimetic game. An error of one of the agents can lead the system toward a new metamimetic attractor. When the system is constantly perturbated, the only state in the SCSS is the state where all agents are (*D,maxi*). The proportion of this state in the limit distribution when (ε_r, ε_a) tends to zero quickly converges to *1* (right). Here the proportions have been plotted for $0{,}01<\varepsilon_i<0{,}2$. (it should be noticed that since all states are taken with an equal probability in case of mistake, here half of errors are corrected. The real level of noise is then $\varepsilon_i/2$)

For example, as a first approximation we can assume that all mistakes in copying neighbor's strategies are possible and are time-independent and consider that with a probability $1-\varepsilon_r$ (resp. $1-\varepsilon_a$) the agents choose the correct imitation rule (resp. behavior) but with a probability ε_r (resp. ε_a) choose a rule (resp. the behavior) at random in the set *R* (resp. *B*). We obtain a pertubated Markov process P^{ε} that has a unique stationary distribution (figure 7). When the perturbation is small ($||P^{\varepsilon}-P^{0}||_{\infty}<<1$) this stationary distribution is concentrated around a particular subset of attractors of the process defined by P^{0} : the stochastically stable set (Foster & Young 1990). To highlight the fact that this Markov process P^{0} represents endogenous mimetic dynamics, we will call this set the stochastically counterfactually stable set (*SCSS*). In our example, it is straightforward to show that the only state in the *SCSS* is the state [*(D,maxi) ; (D,maxi)*] (figure 8).

The simplicity of our example do not allow us to show all the characteristics of these games in noisy settings. The most important feature here is that the coupling between the internal dynamics and the perturbations operates a second selection on the states privileged by the internal dynamics, the limit distribution being independent of the initial conditions.

Although the role of perturbations would be too long to be exposed in detail here, I will only mention that usually, the SCSS does depend on the structure of the perturbation which here is given by the relation between ε_a and ε_r. Formalizing this dependence is perhaps one of the principal contributions of this approach. In those systems, we have two sources of information : the internal dynamics and the structure of noise in the environment. The relation between these two is an example of structural coupling (Varela 1979).

But the main aspect this simple example highlights is that dynamics of metamimetic games are not optimization of criteria that are arbitrarily assigned to agents and exist prior to their activity. The question here is not to find the "best" strategy, which would first require that the modeler ask herself which definition of best is the best. The question is to find the states such that each agent, through its interactions with its social networks, has found an identity and a strategy that is self-

coherent given its social environment. The question is then to find the social configurations that are maximally stochastically counterfactually stable. This suggests us to study social systems as autonomous systems evolving under phenomena of differentiation among a multiplicity of possible criteria along a process of cultural co-evolution.

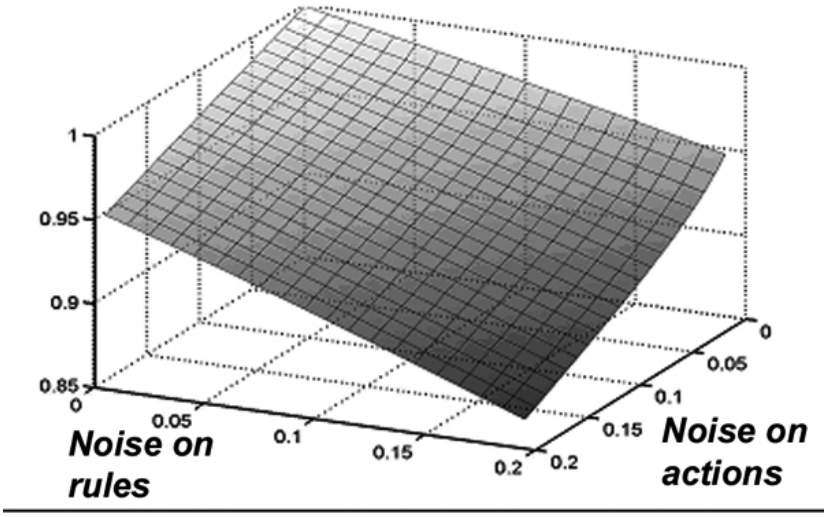


Figure 8: Frequency of the SCSS at the attractor. In our minimal example, the proportion of the SCSS in the limit distribution when (ε_r, ε_a) tends to zero quickly converges to *1*. Here the proportions have been plotted for $0{,}01<\varepsilon_i<0{,}2$. (it should be noticed that since all states are taken with an equal probability in case of mistake, here half of errors are corrected. The real level of noise is then $\varepsilon_i/2$)

V. How to choose the set of imitation rules ?

V.1. Perceptions and computations

Before concluding, I would like to highlight the shift of perspective about social systems modeling that the current approach could bring about. Contrary to standard games where the modeler has first to choose a set of strategies and then a dynamics on this set, in metamimetic games, the set of strategies is sufficient to define the dynamics. For this reason, we must pay a particular attention to this set.

Following Baldwin (1897), who wrote that imitation is a means of selecting stimuli in the environment, it is interesting to think of imitation rules in terms of selective attention: an agent is particularly sensitive to one dimension of its perceptive space and builds from the stimuli detected along this dimension a function that it will use to select the appropriate trait to copy. This suggests that we should not define the set of imitation rules as a list, as is usually done, but as a set generated by some cognitive operators : operators for the selection of a particular dimension in the stimuli space and operators for computation on this space (figure 8).

To choose the appropriate set the modeler then must ask herself, “What can an agent perceive in the situation considered?“ and “What kind of operations can an agent do on these perceptions?”. For example, the fact that an agent is able to act as a conformist means that it can focus its attention on frequencies of traits in a population and find the most frequent trait. The fact that an agent is able to act like a payoffs-maximizor means that it can focus its attention on payoffs and can at least compute the maximum of two scores.

V.2 Regularities of the set of rules

This approach suggests that an important part of the work is to find some regularities of such sets of rules according to what is known in cognitive sciences, in order to built them in a generative way. The first one we can propose as a first approach has been evoked by Gabriel Tarde (1890) in the forewords of the second edition of his *"Laws of Imitation" :* we can reasonably think that there is no social system were an imitation rule is present without its counterpart, the one that prescribes exactly the contrary. If an agent can imagine a rule, it can imagine its contrary. From the formal point of view this means that if an agent can find the maximum between two numerical values, it can also find the minimum, which is formally equivalent to saying that there is a cognitive operator that multiplies numerical values by *–1*.

If we accept this property, the minimal set of rules that contains the *maxi* rule and the conformist rule (figure 9) would be in fact a set with four rules {*maxi, mini,* conformist, anti-conformist}, where the basic percepts are densities of traits and payoffs, and the operators inverse (that multiply by –1 numerical values) and max (that find the max between two numerical values). This means that if one of these rules plays a minor role in the model considered, it should be because the internal dynamics of the system has eliminated it and not because the modeler has decided against this rule a priori. This is possible precisely because the distribution of imitation rules is endogenous once you choose a set.

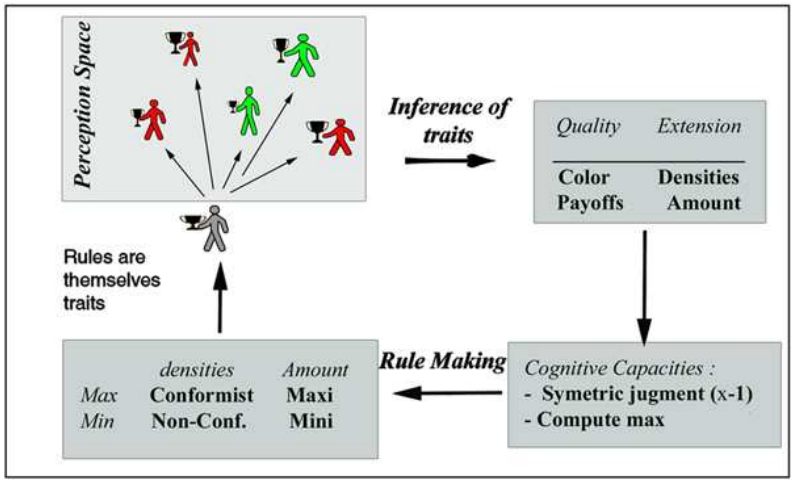


Figure 9: Our approach suggests to define the set of imitation rules in a generative way. Agents are embedded in their environment from which they infer some traits like colors and payoffs. Then, they do some computations on the inferred distributions of these traits : computation of the densities, computation of the maximum, computation of the minimum, etc. These computations are used as a basis for building imitation rules, that are themselves traits agents can try to infer.

V.3. Shifts in the phase space

These remarks raise a fundamental question: How new categories of traits, new dimensions in the perceptive space that will bring the metamimetic dynamics from a Markov chain defined by P^0 *to* an other defined *by* P'^0 emerge?. For example, in a very schematic way is it possible to imagine how a population of mimetic agents can become sensitive to the number of persons that imitate a given agent so that they begin to imagine rules like prestige ? Is it possible to imagine how a population of *conformist/anti-conformist* agents who care only about densities of traits can invent such a thing as money that brings about couples of rules like *maxi/mini* ? Such transitions reflect the shift in the phase space of the whole systems and their study is highly relevant to the study of social systems dynamics. Indeed, it is highly probable that social systems are constantly shifting from phase space in that way as a consequence of innovations. Understanding these kinds of shifts in the phase space is certainly one of the challenges of social systems modeling.

VI Conclusions

As a heuristic for the modeling of human social systems, several scientists have proposed focusing on models that include human specific cognitive capacities. The justification for this is that only such models should be able to explain the huge gap in the complexity of social structures between animal and human societies. Following this heuristic, we proposed including some consequences of reflexivity and metacognition in social systems modeling, i.e. the fact that human beings know to some extent that they are using rules for decision-making and can monitor their use.

When associated with reflections about human imitation, this remark led us to propose a formal framework for the modeling of social systems: *metamimetic games*. In those games, agents imitate in accordance with their preferences, as in most models that deal with imitation, but also form their preferences through imitation. Those games have two remarkable properties:

1) imitation rules can be their own meta-rules. Thus, we escape the regress problem that threatens as soon as the question of grounding the choice strategies is evoked;

2) there is an endogenous metadynamics on imitation rules: the distribution on imitation rules is the product of the dynamics it defines.

With a very elementary example of such a game, I introduced the corresponding concepts of attractors and equilibria: counterfactually stable states. In a noisy setting, we worked in the framework of stochastic game theory (Foster & Young 1990) and proposed to concept of *stochastically counterfactually stable states*.

Unlike other games, the main question in the study of metamimetic games is not to find the "best" strategy. This would require indeed that the modeler knows which definition of "best" is the best. Rather, at issue the study of a phenomenon of differentiation in a population along a process of cultural co-evolution among a set of possible motivations in order to identify the (stochastically) counterfactually stable states that are the states toward which the dynamics converges. Those are states such that the social positions of agents are maximally coherent[iii] from *their own* point of view given what they can do.

This is only the first outline of metamimetic games. As already mentioned, future work will have to develop models adapted to more specific situations and also study the link between metamimetic dynamics and other components of human cognition like perception, inference or learning. An interesting issue to study will be the interactions between mimetic dynamics and other kinds of social dynamics. There always exist some prerequisites for participation in a social activity. For example, to participate in an economic activity, one must be creditworthy. These kinds of prerequisites induce a dynamics in the population – often modeled as a replicator dynamics - that is superimposed on the metamimetic dynamics. We can expect that the understanding of the mutual influence between these different dynamics will be very instructive in understanding the complexity of human social systems.

Acknowledgments :

This paper benefited from discussions at the Centre de Recherche en Epistémologie Appliquée (CREA). For insightful comments, the authors thanks particularly Paul Bourgine, Jean-Pierre Dupuy, Keneth Williford and several anonymous referees. The research was supported in part by the ACI Systèmes Complexes en SHS – « Réseaux, coalition et marchés : une approche interactionniste de la complexité »

.

References

Alvard M. S. (2003), The Adaptive Nature of Culture, *Evolutionary Anthropology*, 136-149

Arthur B. W. (1994), Inductive reasoning and bounded rationality, *Am. Econ. Assoc. Papers and Proc* 84, 406-411.

Arbib M.A., Billard A., Iacoboni M., Oztop M. (2000), *Synthetic brain imaging: grasping, mirror neurons and imitation*, Neural Networks 13 975–997.

Axelrod R. (1997), The Dissemination of Culture: A Model with Local Convergence and Global Polarization, *Journal of Conflict Resolution*, 41, 203-26.

Baldwin, J. M. (1890), *Mental Development in the Child and the Race*.

Bourgine P., What is Cognitive Economics, in *Cognitive Economics*, Bourgine P, Nadal J-P (ed.),Springer, 2004

Bowles S., Gintis H. (2003) The Origins of Human Cooperation, Peter Hammerstein (ed.) *The Genetic and Cultural Origins of Cooperation*, Cambridge: MIT Press.,

Boyd R., Richerson P.J. (1985) *Culture and the Evolutionary process*, Chicago, IL: University of Chicago Press.

Breazeal, C., Buchsbaum, *D*., Gray, J., Gatenby, *D*. & Blumberg, B., (2005) Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots, *Artificial Life* 11(1-2):5 -11

Byrne R.W., Russon A.E., Learning by Imitation: a Hierarchical Approach, *Behavioral and Brain Sciences*, 1998

Chaminade T., Meltzoff A. N. & Decety J. (2002), Does the End Justify the Means? A PET Exploration of the Mechanisms Involved in Human Imitation, *NeuroImage* 15, 318–328.

Conte, R., Paolucci, M., Intelligent Social Learning, *Journal of Artificial Societies and Social Simulation* vol. 4, no. 1,<http: //www.soc.surrey.ac.uk/JASSS/4/1/3.html>, 2001

Damasio R. A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, New York: Harcourt Brace.

Donald M. (1991) *Origins of the Modern Mind*, Cambridge, MA: Harward University Press

Dupuy J.-P. (2004), Economics as symptom, In Lewis, PA, ed. *Transforming Economics: Perspectives on the Critical Realist Project*. London and New York: Routledge, p. 227-251

Edmonds B. (1999) Capturing Social Embeddedness: a constructivist approach, *Adaptive Behavior*, 7:323-348.

Fehr E., Fischbacher U. (2003) The nature of human altruism, *Nature*, Vol 425, october

Feldman M. W., Laland K. N. (1996) Gene-culture coevolutionary theory, *TREE* Vol 11 n°11 Nov 96 Elsevier,

Foster, *D*., Young, P. (1990) Stochastic Evolutionary Game Theory, *Theoretical Population Biology* 38, 219-232.

Frank J., (2003) Natural selection, rational economic behavior, and alternative outcomes of the evolutionary process, *Journal of Socio-Economics,* Vol. 32 - 6 (12), p 601-622.

Galam S. (1998), Sous les chemises de la symétrie, *Pour la Science*.

Gans E. (1995) Mimetic Paradox and the Event of Human Origin, *Anthropoetics - The Electronic Journal of Generative Anthropology* Volume I, number 2.

Gergely, G., Bekkering, H. & Király, I. (2002), Rational imitation in preverbal infants, *Nature*, Vol 415.

Gintis H. (2003) *Game Theory Evolving*, Princeton Univ. Press

Girard R., *Mensonge romantique et Vérité romanesque*, Girard R., Grasset, 1961

Girard R., *Des choses cachés depuis la fondation du monde*, Girard R., Ed. Grasset, 1978

Gould S. J. (1987), *An urchin in the storm,* London Penguin.

Henrich J., Boyd R. (1998), The Evolution of Conformist Transmission and the Emergence of Between-Group Differences, *Evolution and Human Behavior* 19: 215–241.

Hofbauer J., Sigmund K. (1988), *The Theory of Evolution and Dynamical Systems*, London Mathematical Society, Cambridge University Press.

Henrich, J. and Gil-White, F. J. (2001) The Evolution of Prestige, *Evolution and Human Behavior* 22:165-196.

Jacob, P. (1998) Memory, learning and metacognition, *Comptes rendus de l'Académie des sciences*, 321, 2-3, 235-59.

Jansen, B., De Vylder, B., de Boer B., and Belpaeme,T. (2003) Emerging shared action categories in robotic agents through imitation, In Dautenhahn, K. & Nehaniv, C. L. (eds.) *Proceedings of the AISB Second International Symposium on Imitation in Animals and Artifacts*, Aberystwyth: University of Wales, pp. 145-152.

Janssen M., Jager W., (1999) An integrated approach to simulating behavioural processes:A case study of the lock-in of consumption patterns, *Journal of Artificial Societies and Social Simulation* vol. 2, no. 2, <http://www.soc.surrey.ac.uk/JASSS/2/2/2.html>

Kaniovski Y. M., Kryazhimskii A. V. and Young H. P. (2000), Adaptive Dynamics in Games Played by Heterogeneous Populations, *Games and Economic Behavior* 31, 50-96.

Meltzoff A. N. et Prinz W. ed. (2002), *The Imitative Mind: Development, Evolution, and Brain Bases*, Cambridge University Press.

Meltzoff, A. N. & Gopnik, A. (1993) The role of imitation in understanding persons an developing theories of mind. In: Baron-Cohen S., Tager-Flusberg H. & Cohen *D*. (Eds.), *Understanding other minds: perspectives from autism (pp. 335-366).* Oxford Univ. Press.

Mounoud, P., From direct to reflexive self-knowledge: A recursive model. About (self-produced) actions considered as transformations., P. Rochat (Ed.) *The Self in Early Infancy. Theory and Research.* Amsterdam: North Holland. Elsevier Science Publishers., 1995

Nowak, M. A., Bonhoeffer, S. & R. M. May (1994), More Spatial Games, *Int. J. of Bifurcation and Chaos*, Vol.4, n°1, 35-56.

Nowak M. A., May, R.M. (1992) Evolutionary Games and Spatial Chaos, Nowak M. A., May, R.M., *Nature*, 359: (6398) 826-829.

Orléan A. (1998), The evolution of imitation, In Patrick Cohendet, Patrick Llerena, Hubert Stahn et Gisèle Umbhauer (eds.), *The Economics of Networks*, Berlin, Heidelberg et New-York, Springer-Verlag, Chap 15, 325-339.

Orléan, A., Monnaie et spéculation mimétique (1985) In : Paul Dumouchel (ed) *Violence et Vérité, Autour de René Girard*, Colloque Cerisy, Grasset

Pacherie, E. (1998) Représentations motrices, imitation et théorie de l'esprit. In H. Grivois & J. Proust (eds.) *Subjectivité et Conscience d'Agir: Approches Cognitives et Cliniques de la Psychose*, PUF, 1998, pp. 207-248.

Rendell L., Whitehead H. (2001), Culture in whales and dolphins, *Behavioral and Brain Sciences* 24 (2).

Rogers, A. R. (1989) Does biology constrain culture? *American Anthropologist*, 90: 819–831.

Selten, R., Ostmann, A. (2001), Imitation Equilibrium, *Homo oeconomicus,* 43, 111-149.

Smith J.*D*., Shields, W. E. & Washburn *D*. A (2003) The Comparative Psychology of Uncertainty Monitoring and Metacognition, *Behavioral and Brain Sciences*, Volume 26, Issue 3, 261-283

Sperber *D*. (2000), Metarepresentaitons in an evolutionry perspective, *Metarepresentations: A Multidisciplinary Perspective*. Oxford UniversityPress, 2000, pp.117-137

Tarde, G. (1890) *Les Lois de l'Imitation*, Réédition Les Empêcheurs de penser en rond, 2001.

Tarde, G. (1962/1890) *The Laws of Imitation*, translated by E.C. Parsons with introduction by F.Giddings, reprint, Gloucester, MA, Peter Smith.

Tomasello, M., *The Cultural Origins of Human Cognition*., Harvard University Press., 2000

Varela, F.J., *Principles of biological autonomy*, North Holland, New York, 1979.

Varela, F. (1983), L'auto-organisation : de l'apparence au mécanisme, In: *L'auto organisation, de la physique au politique*, J-P Dupuy & P. Dumouchel, Seuil.

Vriend, N. J. (2002) Was Hayek an Ace?, *Soutthern Economic Journal*, 68(4), 810-840.

Weibull J. W. (1995) *Evolutionary Game Theory*, MIT press

Wilson E.O. (1975) *Sociobiology: the new synthesis*., Harward University Press, Cambridge

Young H. P. (2001), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton Univ. Press.

Young H. P. (1993) The Evolution of Conventions, *Econometrica,* Vol. 61, Issue 1, 57-84

Zelazo P.*D*., Lourenco S.F. (2003), Imitation and the dialectic of representation, *Developmental Review* 23: 55–78.

Zelazo P. *D*., Frye *D*. and Rapus T., (1996) Age Related Dissociation Between Knowing Rules and Using Them, *Cognitive Development*, 11, 37-63.

Zlatev J. (2000) The Mimetic Origins of Self-Consciousness in Phylo-, Onto- and Robotogenesis In *Proceedings of the 26th Annual Conference of the IEEE Industrial Electronics Society*, October 2000, Nagoya.

---

[i] « *Pour élaborer une science de l'homme, il faut comparer l'imitation humaine avec le mimétisme animal, préciser les modalités proprement humaines des comportements mimétiques si elles existent » Girard, 1978*

[ii] Also it could be interesting to consider different sets of rules depending on the agents – as in Selten & Ostmann 2001 - and the cognitive level, but this would be superfluous given the present purpose.

[iii] The term coherent should be understood before all from the modeler's perspective. Agents themselves are not looking for coherence but are simply applying their rules. A rule is coherent if its application does not tend to change the rule itself.

Heterogeneity and predictability of global epidemics

Vittoria Colizza[1,*], Alain Barrat[2], Marc Barthélemy[1], and Alessandro Vespignani[1]

August 30, 2005

[1] School of Informatics and Biocomplexity Center, Indiana University, Bloomington, 47406, IN, USA

[2] LPT and UMR du CNRS 8627, Bâtiment 210, Université de Paris-Sud, 91405 ORSAY Cedex - France

Abstract

We investigate the role of the large scale properties of the airline transportation network in determining the global diffusion pattern of emerging disease. We present a stochastic computational framework for the forecast of global epidemics that considers the complete International Air Transport Association 2002 database complemented with census population data. We adopt an information theory approach to analyze quantitatively the level of heterogeneity and predictability of the epidemic pattern and its relation with the network's structure. The level of spatio-temporal heterogeneity of the spreading pattern is globally characterized and found to be a direct consequence of the network statistical complexity. The epidemic pattern predictability is quantitatively determined and traced back to the occurrence of epidemic pathways defining a backbone of dominant connections in the disease spreading. The presented results provide a general framework for the analysis of containment policies and epidemic risk forecast.

Keywords: Complex networks, Epidemiology, Noise and fluctuations.

The mathematical modeling of epidemics has often dealt with the problem of an appropriate description of real populations with complicated age, social and spatial structures and with heterogeneous

*School of Informatics, Indiana University, 901 E. 10th St., Bloomington, IN 47408-3912, USA, tel.+1-812-856-2226, fax +1-812-856-4764, vcolizza@indiana.edu

patterns in the contact network [1, 2, 3, 4, 5, 6]. Recently, the availability of unprecedented computer power has led to numerical approaches relying on agent-based modeling that simulate entire populations and their dynamics at the scale of the single individual and on a minute-by-minute basis [7, 8]. On the other hand, the inherent complex features and emerging properties [9, 10, 11] of the network in which epidemics occur are not the mere juxtaposition of complicated elements used for increased realism in sophisticated epidemic modeling [12]. Indeed, networks' complex properties often imply statistical fluctuations extending over several orders of magnitude and the breakdown of standard homogeneous approaches and results [5, 6].

These considerations are particularly relevant in the study of the geographical spread of epidemics where the various long-range heterogeneous connections typical of modern transportation networks naturally give rise to a very complicated evolution of epidemics characterized by heterogeneous and seeemingly erratic outbreaks [13, 14]. In this context, air-transportation represents a major channel of epidemic diffusion as recently documented for the SARS outbreak [15]. The modeling of global epidemic diffusion via the air transportation network dates back to the seminal paper of Rvachev and Longini [16] capitalizing on previous studies on the russian network [17]. Similar modeling approaches, even if limited by a very partial knowledge of the world-wide transportation network, have been used to study specific outbreaks such as pandemic influenza [18, 19, 20], HIV [21], and SARS [22]. The availability of the complete world-wide airport network dataset (*WAN*) and the recent extensive studies of its topology [23, 24] are finally allowing a full scale computational study of global epidemics. Here, we use the International Air Transport Association (IATA) database [25] containing the world list of airport pairs connected by direct flights and the number of available seats on any given connection for the year 2002. The resulting air-transportation network is therefore a weighted graph, comprising $V = 3,880$ vertices denoting airports and $E = 18,810$ edges whose weight $w_{j\ell}$ represents the passenger flow between airports j and ℓ. This dataset has been complemented by the population N_j of the metropolitan area served by the airport j as obtained by different sources. The final network dataset contains the $3,100$ largest airports, $17,182$ edges (accounting for 99% of the worldwide traffic) and the respective urban population data. The obtained network is highly heterogeneous both in the connectivity pattern and the traffic capacities. In particular the presence of broad statistical distributions and non-linear associations among the various quantities, contrary to linear relations used so far, indicate a possible major impact in the

ensuing disease spreading pattern.

In this work we will consider for the first time a global stochastic epidemic model including the full International Air Transport Association (IATA) [25] database, aiming at a detailed study of the interplay among the network structure, the stochastic features and the infection dynamics in defining the global spreading pattern of epidemics [26]. While previous studies have in general focused on the *a-posteriori* analysis of real case studies of global epidemics, the large scale modeling allowed by the IATA database enables us to address general theoretical issues such as (i) the spatio-temporal statistical properties of the epidemic pattern, (ii) their relation with the complex features of the underlying transportation network and (iii) the reliability of forecasts and outbreak scenarios with respect to the intrinsic stochasticity of disease transmission and traffic flows. The model is analyzed by using an information theory approach that allows the quantitative characterization of the heterogeneity level of the spreading pattern and its predictability in presence of stochastic fluctuations. Results provide a general computational framework for the analysis of containment policies and risk forecast of global epidemic outbreaks. Simulations and reproductions of case studies of real epidemics are also presented.

References

[1] Anderson RM, May RM, *Infectious diseases in humans* (Oxford University Press, Oxford 1992).

[2] Hethcote HW, Yorke JA, *Lect. Notes Biomath.* (Berlin, Springer-Verlag, 1984), vol. 56.

[3] Kretzschmar M, Morris M, *Math. Biosci.*, **133**, 165(1996).

[4] Keeling M, *Proc. R. Soc. Lond. B* **266**, 859(1999).

[5] Pastor-Satorras R, Vespignani A, *Phys. Rev. Lett.* **86**, 3200(2001).

[6] Lloyd AL, May RM, *Science* **292**, 1316(2001).

[7] Chowell G, Hyman JM, Eubank S, Castillo-Chavez C, *Phys. Rev. E* **68**, 066102(2003).

[8] Eubank S, Guclu H, Anil Kumar VS, Marathe MV, Srinivasan A, Toroczkai Z, Wang N, *Nature* **429**, 180(2004).

[9] Albert R, Barabási A-L, *Rev. Mod. Phys.* **74**, 47(2000).

[10] Dorogovtsev SN, Mendes JFF, *Evolution of networks: From biological nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).

[11] Pastor-Satorras R, Vespignani A, *Evolution and structure of the Internet: A statistical physics approach* (Cambridge University Press, Cambridge, 2003).

[12] Ferguson NM *et al*, *Nature* **425**, 681(2003).

[13] Cohen ML, *Nature* **406**, 762(2000).

[14] Cliff A, Haggett P, *British Medical Bulletin*, **69**. 87(2004).

[15] http://www.who.int/csr/sars/en

[16] Rvachev LA, Longini IM, *Mathematical Biosciences* **75**, 3(1985).

[17] Baroyan OV, Genchikov LA, Rvachev LA, Shashkov VA, *Bull. Internat. Epidemiol. Assoc.* **18**, 22(1969).

[18] Longini IM, *Mathematical Biosciences* **90**, 367(1988).

[19] Grais RF, Hugh Ellis J, Glass GE, *European Journal of Epidemiology* **18**, 1065(2003).

[20] Grais RF, Hugh Ellis J, Kress A, Glass GE, *Health Care Management Science*, **7**, 127(2004).

[21] Flahault A, Valleron A-J, *Math. Pop. Studies* **3**, 1(1991).

[22] Hufnagel L, Brockmann D, Geisel T, *Proc. Natl. Acad. Sci. USA* **101**, 15124(2004).

[23] Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A, *Proc. Natl. Acad. Sci. USA* **101**, 3747(2004).

[24] Guimerà R, Amaral LAN, *Eur. Phys. J. B* **38**, 381(2004).

[25] http://www.iata.org

[26] Colizza V, Barrat A, Barthélemy M, Vespignani A, *q-bio.OT/0507029* (2005).

Modelling price competition of retail stores under imperfect information

Margaret Edwards[a], Pablo Jensen[a,b c], Hernán Larralde[d]

[a] *LET-CNRS, Université Lyon-2, France*
[b] *LPMCN, Université Lyon-1 and CNRS, 69622 Villeurbanne Cedex, France*
[c] *Laboratoire de Physique, Ecole Normale Supérieure de Lyon and CNRS, 46 Allée d'Italie, 69007 Lyon, France*
[d] *Centro de Ciencias Físicas, UNAM, Cuernavaca Mor. Apdo. Postal 48-3, C.P. 62251, Cuernavaca, Morelos, MEXICO*

Abstract

We improve the realism of the economic description of store competition through prices. We use multi-agent simulations to study the Nash equilibrium for prices under imperfect consumer and store information. Stores do not know the profit vs. price curve and have to calculate it "on the fly", as consumers buy. Consumers do not know the actual prices and progressively learn them through information exchange with the social network. We show that under many realistic situations, incomplete dissemination of the information for the consumers distorts the estimation of the profit and prevents convergence of the price to the optimum.

Keywords : spatial price competition, Nash equilibrium, multi-agent

---

Introduction

Many economists have studied optimal location and pricing for stores in a one or two dimensional city ([1]-[5]). They supposed that Nash equilibrium would always be found and assumed perfect information. However the understanding of most classical economic models can be improved by linking the global dynamics to the choices taken by individuals in the models. Here we inquire into how the equilibrium state emerges from local laws and interactions. Our purpose is to understand how relaxing the propagation information conditions the convergence of the price towards the Nash equilibrium.

We consider a town with two stores, the position of which is fixed. The prices corresponding to the Nash equilibrium are known. They provide the stores with optimal income. Being in the Nash equilibrium means that for any store, changing its price can only lead to a decrease of its profit. We focus our study on one of the stores. We seek to understand under which conditions of information it can find its optimal price (corresponding to the Nash equilibrium). We consider both the information of the store about the demand and the information of the consumers about the current price of the store.

Modelling choices

General features

We consider a two-dimensional square city represented by *LxL* discrete sites, every of which is occupied by a consumer, and two also by one of the competing stores.

To model the behaviour of the consumers, we assume that each consumer *j* can assign a ``utility" function K_{js} to each store *s* given by:

$$K_{js} = R - p_s - a \cdot d_{js} \qquad (1)$$

where *R* is the ”maximum utility” of buying the product, assumed to be high enough to prevent any negative value for K_{js}. p_s is of the product at store *s*, d_{js} the (Euclidian) distance between consumer *j* and store *s*. A simple way to include the effect of random factors that “smoothen” the consumer's choice among the stores is to use the well-known “logit” weighting factor ([1]). In this approach we assign a probability for consumer *j* to buy at store *s* as:

$$_{js} = \frac{\exp(K_{js}/T)}{\sum_{s'} \exp(K_{js'}/T)} \qquad (2)$$

Implementing the incomplete information of the consumers about the price

We simulate incomplete consumer information and the propagation of the information about the stores prices in the following way. At each change of the price, only the consumer located at the node of the store is informed of the new price. Then at each time step, a certain proportion of the population (*m%*) is randomly picked to exchange information and, based on the updated information, decide whether and where to purchase. Each of the chosen consumers decides at which store to purchase with probability given by Eq (2). Finally, each of these consumers updates his information of the prices to their actual value after buying. The diffusion of information about the new price for *m=10%* appears in Fig.1.

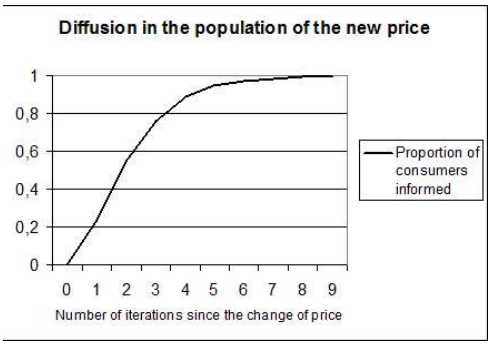


Figure 1. Evolution of the proportion of population informed when 10% exchange information at every time step.

Implementing the incomplete information of the store about the demand

As a first effect, the random choice of consumers at each time step induces fluctuations in the demand for either, even when all consumers are perfectly informed of the prices.

In addition to these fluctuations, the information the consumers have about the prices is changing in time, making the extrapolation of the long term profits from the short term estimates even harder to perform.

Optimization algorithm of the price for the store

We start from the Nash equilibrium prices for both stores. At the beginning of the simulation the price of one of the two stores is randomly changed by ±10%. The store maintains each price over a period of *k* time steps. It then calculates the average profit over this period, to decide on the new price : as long as the profit increases, it keeps the same trend for the price variation. To favour the convergence towards a definite value, we diminish the magnitude of the change in price each time the price change is reversed.

Results

Let us now compare the value after convergence of the prices and profits for different frequencies of change in price (every 5, 10, 20 and 100 iterations) (see Fig.2).

We see that the results improve with the increase of the interval between two changes in price : the average price grows closer to its optimum, and the standard deviation decreases. The profit nears optimal value in the same way. Conversely the results for frequent changes in price are not satisfactory. First the average price does not correspond to the equilibrium value. Second, there is a wide range of price values (determined by minimum and maximum) over the 100 simulations, showing that some simulations lead to very bad estimates of the optimum price.

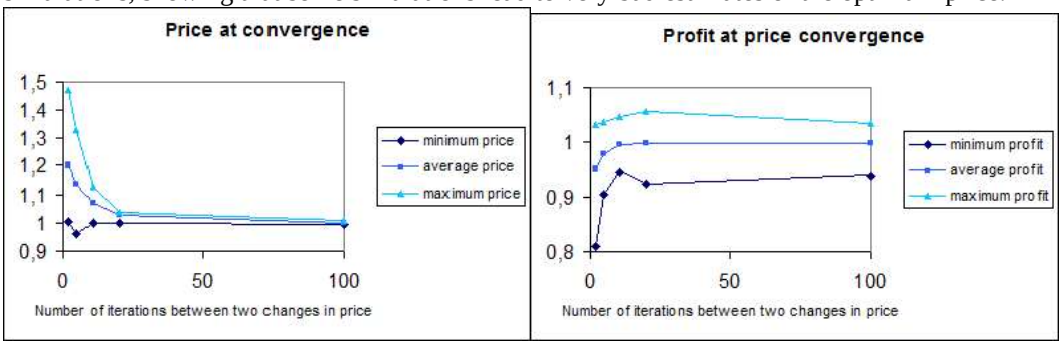
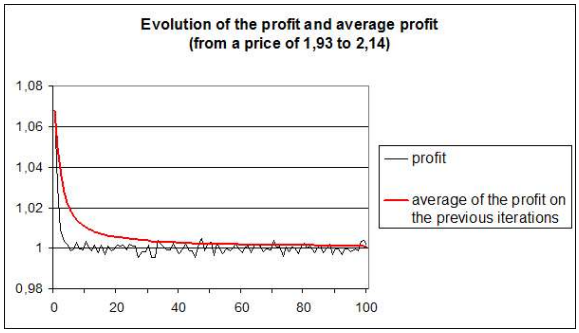


Figure 2. Normalised price (a) and profit (b) after convergence of the price for different frequencies of change in price (10% consumers exchange information and buy at every iteration).

Indeed the information about any change in price takes time to spread. In the first iterations where consumers are still ignorant of the new price the demand does not yet change significantly. Thus, the delay of information among the consumers leads in the first iterations to a distorted evaluation of the profit to be expected for a given price. If the price has increased, the profit is overestimated. On the contrary, if it has decreased, the profit is underestimated. An example of this is shown in the following figure when the price changes from 1,93 to 2,14. The profit is normalised (theoretical value to be attained is therefore 1).

Figure 3. Evolution of the average profit after a change of price from 1,93 to 2,14.

Conclusion



We seek to understand under which conditions a store can find its optimal price by testing how his average profit changes with the price. It appears that when the changes in price are too frequent, the estimation of the profit is distorted by the first iterations. In these iterations consumers are not yet informed of the new price and their demand is not corresponding. Therefore it appears that propagation of the information plays a major role in finding or not the optimal equilibrium price.

References

[1] Anderson, M. Fujita and J. F. Thisse, 1992, *Discrete Choice Theory of Product Differentiation*, MIT Press

[2] Fujita, M. and J. F. Thisse, 2002, *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*, Cambridge University Press

[3] Hotelling, H. ,1929, *Stability and competition*, Economic Journal **39** 41 (1929)

[4] Irmen, A. and J.F. Thisse, 1998, *Competition in Multi-characteristics Spaces : Hotelling Was Almost Right*, Journal of Economic Theory, **78(1)** 76

[5] Veendorp, E.C.H. and A. Majeed, 1995, *Differentiation in a two-dimensional market*, Regional Science and Urban Economics **25** 75

Enabling cooperative behaviour through ICT in organisations

This paper addresses a new way of making sense of behaviour in organisations mediated by information and communication technology (ICT). By studying micro-level interactional processes and relating this to macro-level patterns of interaction, mutually influencing each other, I suggest the importance addressing these issues in projects with the intention of changing behaviour through the employment of ICT.

The empirical case studies are from the Norwegian hospital sector, where the electronic patient records (EPR) is employed in organisational change projects. The first case is a project of moving from oral face-to-face nursing hand-over to hand-over employing EPR for asynchronous communication. The second case is a project of introducing a module of EPR for nurses to document plans, actions and evaluations in their practical work. Both cases have been followed over a period of 10 months. During this time 42 semi-structured interviews with different stakeholders have been carried out. In addition participant observation of everyday practice and different meetings has been carried out comprising approximately 300 hours.

To conceptualize how one actor is interconnected with other actors in the complex social system of organisations, the theory of symbolic interaction and in particular the thoughts of Mead on the fundament of human interaction, is employed [1,2,3]. Behaviour of the individual is constructed taking into account other actors in a continuous and dynamic process. Using ICT, I argue that actors taken into account are distributed in the time and space dimensions. Hence, the interactional processes that are so strong in face-to-face interaction are not absent when it comes to interaction through ICT. This perspective needs empirical investigation, it is argued. From the cases, it is shown that when EPR is employed for nursing hand-over, actors reflect more on the needs of the next nurse reading it than before, therefore including more information, and spending more time documenting. Further, even if these micro-level interactional processes are characterized of local situatedness and unpredictability [4], patterns of interaction emerge that are orderly and random at the same time [5,6].

In ICT change projects behaviour is usually seen through lenses of system theory, controlled through system design, e.g. syntax or sequences for processing information. Here, the action of the individual is connected to other individuals from the perspective of designers and planners of the system. In contrast, in this paper it is argued that it should be realised that the action of the individual is not possible to control, nor is it desirable to do so. The interconnections of actors in complex systems is not something that designers or managers prescribe from the outside, but something that "exists" in the minds of the individual actors continually aligning their actions and interactions, as argued above. This cooperative behaviour is not locked to some definite system, but remains adaptive for further modifications in connection with other agents in the social system. Accordingly, change is seen as being enabled, as through the lenses of complexity theory [7,8,9].

Through the case studies I discuss and show how individual and cooperative behaviour can be understood based on the premise of 'taking others into account' and how the employment of ICT leads to changing patterns of interaction on an overall level working back on the micro level interaction. Further, I show how the failure of addressing these issues jeopardized the success of the projects.

References

[1] Mead, G.H. (1934) Mind, Self & Society, The University of Chicago Press.
[2] Blumer, H. (1969) Symbolic Interactionism, Perspective and Method, University of California Press, Prentice-Hall, Inc.
[3] Blumer, H. (2004) George Herbert Mead and Human Conduct, AltaMira Press.
[4] Bak, P. (1997): “How Nature Works: The Science of Self-Organized Criticality”. Oxford University press.
[5] Orlikowski, W.J. (1996) Improvising Organizational Transformation Over Time: A Situated Change Perspective. Information Systems Research 7, 63-92.
[6] Kauffman, S. (1995): ”At home in the universe: The search for Laws of Self-Organization and Complexity”. Oxford University press:3-69,149-189
[7] Mitleton-Kelly, E. (2003): “Ten Principles of Complexity & Enabling Infrastructures” in Complex Systems and Evolutionary Perspectives of Organisations: The Application of Complexity Theory to Organisations. Edited by Mitleton-Kelly, E. Elsevier, chapter 2.
[8] Shaw, P. (2002): ”Changing Conversations in Organizations: A compexity approach to change”. Routledge
[9] Stacey, R.D. (2003): “Strategic Management & Organisational Dyamics: The Challenge of Complexity”. Fourth Edition, Prentice Hall.

The design of an artificial society

Nigel Gilbert, Stephan Schuster, Lu Yang
n.gilbert@soc.surrey.ac.uk

The New and Emergent World models Through Individual, Evolutionary and Social learning (NewTies) project is concerned with emergence and complexity in socially-inspired artificial systems. It is developing a large computational system consisting of an environment and a population of agents. The main goal of the project is to realize an evolving artificial society whose members are capable of exploring the environment and developing their own images of this environment and the society through interaction and cooperation. A subsidiary objective is to develop the system so that it can be used as an experimental testbed to examine hypotheses originally formulated as explanations of phenomena observed in human societies.

In this presentation, we shall consider what is involved in designing this system of agents and environment. The design problem comes from the need to steer between, on the one hand, producing a system which attempts to simulate so many of the characteristics of humans that understanding the system is impossible or, on the other hand, designing a system that is so simplistic that any inference to human societies is at best tenuous. The issue is made more difficult by our prior assumption that most features of human societies are emergent from social interaction, and that interaction itself is only achieved through the social evolution of the means of communication.

Our solution has been to design an environment containing rather few different types of object, each with carefully chosen properties (e.g. building blocks with which agents can construct either 'roads' to ease their mobility, or 'walls' to provide fortifications) and to give the agents a small number of primitive actions (e.g. move, pick up object, talk). The expectation is that agents will be able to create a rich behavioural repertoire by composing the primitive actions, and that, despite there being only a few different types of object, the environment will be complex enough to encourage the agents to exercise this repertoire.

The presentation will describe the environment and the agents' abilities and explain how and why they were designed in this way. In order to show how the system could eventually be used to examine social theories about human societies, examples will also be given of 'scenarios' that can be created that emulate the circumstances in which 'simple' societies have lived, as described by social anthropologists.

Altruism "For Free" using Tags

David Hales
David Hales, Dept. of Computer Science, The University of Bologna
dave@davidhales.com

Keywords: Altruism, Evolution of cooperation, Tags, Prisoner's Dilemma.

A number of recent models have shown how "tags" (Holland 1993) - arbitrary observable phenotypic markers - can produce altruistic behaviour between initially selfish individuals in an evolving system (Riolo et al 2001). These models use tag and tolerance values associated with each individual such that *those with identical tags are forced to be altruistic* (Roberts et al 2002, Edmonds et al 2003). Here we outline a simpler approach in which individuals play the single round Prisoner's Dilemma (PD) game. Contrary to the previous tag models altruism is demonstrated in the single round game *without forced altruism between those with identical tags* or knowledge of previous interactions. The system is reverse-scalable (the more individuals in the population the quicker altruism emerges) and robust to noise. This *"altruism for free"* property has already been adapted and applied into robotic scenarios (Hales et al 2003), and peer-to-peer networks (Hales 2004) producing light-weight robust algorithms that allow selfish agents to work as cooperative teams in low trust environments without the need for complex trust or market mechanisms.

We demonstrate that tags produce cooperation in the singe-round PD in an agent-based simulation. In the PD agents are paired and play a game by selecting one of two strategies (either cooperate or defect). Depending on what they select they each receive a payoff value. If both cooperate they both get R, if both defect they get P, otherwise the cooperator gets S and the defector gets T. The dilemma arises from the payoff constraints: $T>R>P>S$ and $2R > T+S$. Agents are represented as fixed length bit strings (of length L+1) comprising a tag of length L bits and a single strategy bit. The strategy bit represents a pure strategy, either unconditional cooperation or unconditional defection. Initially the population of agents is set to random bit strings (with each bit decided by a fair coin toss). The following evolutionary algorithm is then applied:

```
LOOP some number of generations
  LOOP for each agent (a) in the population
    Select a game partner agent (b) with matching tag
    Agent (a) and (b) play single round PD
  END LOOP
  Reproduce agents in proportion to their average payoff
END LOOP.
```

In each generation, each agent (a) is selected from the population (of size N) in turn. A game partner is then selected. Partner selection entails the random selection of another agent (b) from the population such that (a) ≠ (b) and tags match identically. If no agent in the population has matching tag to then (a) interacts with a randomly chosen agent. During game interaction (a) and (b) invoke their strategies and receive the appropriate payoff. After all agents have been selected in turn and played a game a new population is asexually reproduced. Reproductive success is proportional to average payoff. The entire population of N agents is replaced using a "roulette wheel" selection method. Mutation is applied to each bit of each reproduced player with probability $M = 0.001$, consequently *tags and strategies* are mutated in reproduced agents. The PD payoffs are parameterised over T such that $T > 1$. The reward R for cooperation is 1. The punishment P for mutual defection and the sucker payoff S for cooperation with a defector are both some small value.

A set of runs to 100,000 generations with a population of size $N = 100$, was executed for various values of T and L. We found high levels of cooperation emerged quickly when L was sufficiently large ($L>16$). The results obtained indicate that very high levels of cooperation can be sustained between selfish greedy optimisering agents in the single round PD via simple tag biasing. There is no requirement for knowledge of past performance or recognition of individual agents

The tag space can be visualised as an L-dimensional hyper-cube with corners representing unique tag values. Agents sharing a tag, share a corner. Mutation produces movement between corners. Game interaction is therefore taking place in an abstract "tag space". Cooperative groups sharing matching tags will form in corners of the hyper-cube. These groups will outperform non-cooperative groups and hence tend to increase in size over generations. However, if mutation introduces defecting agents into a cooperative group they will tend to outperform the cooperators within the group (by suckering them). From this the seeds of the destruction of the group are planted, since as the number of defectors increases within a group the overall fitness of agents within the group decreases. Other more cooperative groups (if they exist) will tend to expand. While this process is occurring, mutation of tag bits will produce a slow migration of agents between corners of the hyper-cube, possibly founding new groups in previously empty corners.

Figure 1 is a visualisation of the process over time taken from a single run. Each line on the vertical axis represents a corner of the hyper-cube (i.e. unique tag value). The horizontal axis represents time in generations. If no agents have a particular tag value in a given generation then the line is left blank (white). Alternatively, if a corner contains all cooperative agents then the line is light grey. For a mixed group in which there are both cooperators and defectors the line is dark grey. For an entirely defective group the line is black. Examination of figure 2 shows the time evolution of groups in tag space. Initially cooperative groups (light grey lines) become invaded by defectors producing mixed groups (darker grey) that very swiftly become entirely defective (black) and then quickly go extinct (white). See Hales (2000) for more results from this model.

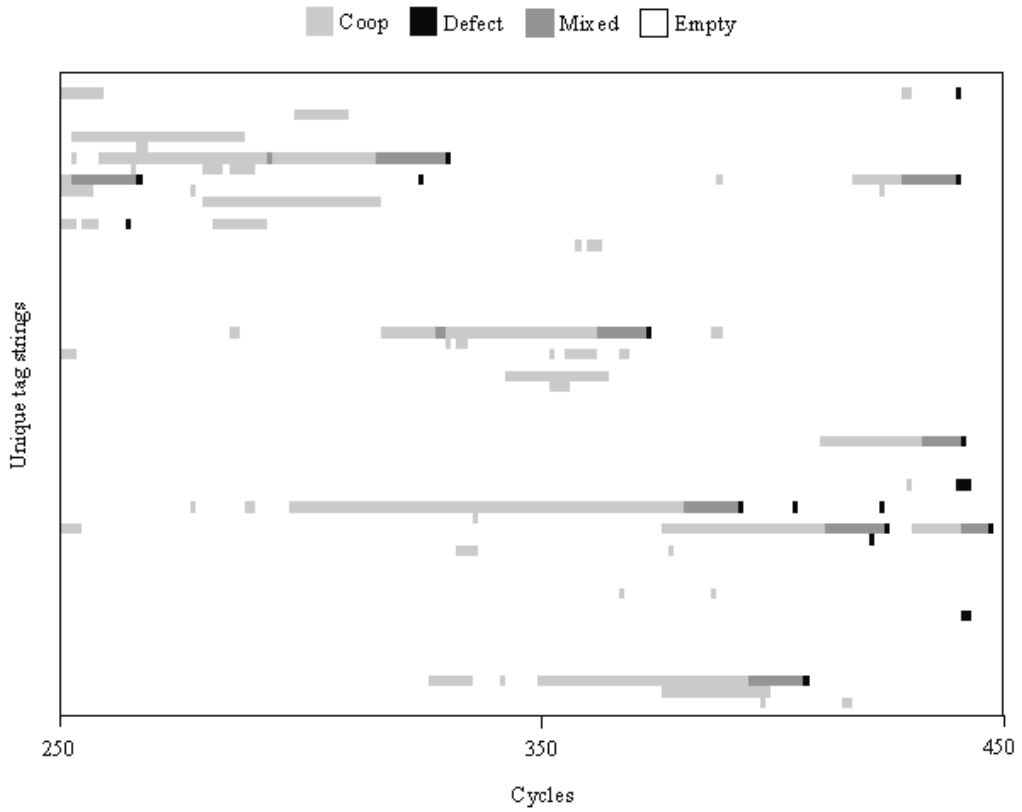


Fig. 1. Visualisation of 200 generations from a simulation run showing cooperative groups coming into and going out of existence. Each line on the vertical axis represents a unique tag value (of which only a subset is shown). If all agents sharing a tag value are cooperative then the line is light grey. If all agents are defectors then the line is coloured black. A mixed group is shown as dark grey.

References

D. Hales, “From selfish nodes to cooperative networks -- emergent link based incentives in peer-to-peer networks”, Proc. of the 4th IEEE International Conference on Peer-to-Peer Computing (P2P2004). IEEE Computer Soc. Press, (2004).

D. Hales, Cooperation without Space or Memory: Tags, Groups and the Prisoner's Dilemma. In Moss, S., Davidsson, P. (Eds.) Multi-Agent-Based Simulation. Lecture Notes in Artificial Intelligence, 1979:157-166. Berlin: Springer-Verlag. (2000).

D. Hales. & B. Edmonds. Evolving Social Rationality for MAS using "Tags", In Rosenschein et al. (eds.) Proceedings of the 2nd International Conference on Autonomous Agents and Multi-agent Systems, (AAMAS03), ACM Press, 497-503. (2003).

B. Edmonds, B. and D. Hales, Replication, Replication and Replication - Some Hard Lessons from Model Alignment. Journal of Artificial Societies and Social Simulation 6(4). (2003).

Holland, J.: The Effect of Lables (Tags) on Social Interactions. Santa Fe Institute Working Paper 93-10-064. Santa Fe, NM (1993)

R. Riolo, M. D. Cohen, R. Axelrod, “Cooperation without Reciprocity”, Nature 414, 441-443, (2001).

Roberts, G. & Sherratt, T. N. (2002) Nature 418, 449-500

Transition to Coherent Oscillatory Behaviour in a Route Choice Game

Dirk Helbing [1,a], Martin Schoenhof [1], Hans-Ulrich Stark [1], and Janusz A. Holyst [2]

[1] Institute for Transport & Economics, Dresden University of Technology, Germany

[2] Faculty of Physics and Center of Excellence for Complex Systems Research, Warsaw University of Technology, Poland

[a]Corresponding author:
Dirk Helbing
Institute for Transport & Economics
Dresden University of Technology
Andreas-Schubert-Str. 23
D-01062 Dresden
Phone: +49 351 463-36802
Fax: +49 351 463-36809
E-mail: helbing1@vwi.tu-dresden.de
URL: www.helbing.org

Short title: Transition to Coherent Oscillatory Behaviour

Abstract

Selfish routing of traffic over alternative routes wastes available street capacities, as individuals tend to generate an equilibrium state (a 'Wardrop' and 'Nash equilibrium') with higher overall travel times than in the optimal state. This system optimum is characterized by coherent oscillatory patterns rather than a stationary behaviour. Here, we study the time-dependent decision behaviour in a day-to-day route choice setting by means of experimental and simulation results. While there is a tendency towards establishing the Nash equilibrium in the beginning, we often find a transition to coherent oscillatory behaviour after a long transient time period. In spite of the complex dynamics leading to co-ordinated oscillations, we have identified mathematical relationships quantifying the observed transition process. Moreover, the main discoveries are reproduced by a reinforcement algorithm, which may help to establish more efficient data traffic on the internet.

Keywords

Game theory; traffic distribution; reinforcement learning

Game theory has been very successful in describing strategic interactions in social, economic, and biological systems, but it has also attracted great attention among theoretical physicists [1]. This includes the minority game [2] and cyclic behaviour in predator-prey or rock-scissors-paper games [3]. Another interesting field is the spontaneous establishment of cooperation in repeated games. The prisoner's dilemma game, for example, reflects many situations, in which individuals are tempted to defect (see Fig. 1). However, cooperation would be better for both and can emerge, if the game is repeated frequently enough, as defection can be punished later on ("shadow of the future") [5, 6]. Apart from future expectations [7, 8], cooperation may be supported by kinship relations [9], reciprocity [10] or similarity [11], small populations [12], spatial interactions [13, 14], or variation in behaviour [15].

The route choice game discussed in the following reflects situations, where the outcome of a decision depends on the independent decisions of many others. It describes the problem of choosing among two alternative routes $i \in \{1, 2\}$ between the same origin and destination. As the travel times are monotonously increasing with road occupancy, we specify the payoffs $P_i(N_i)$ as a function of the number N_i of vehicles on road i by a linearly decreasing function $P_i(N_i) = C_i - D_i N_i$. Experimental results for this setup [16, 17] have shown that groups of many persons tend to establish the Wardrop equilibrium [18] characterized by equal travel times $T_1 = T_2$. This state corresponds to a Nash equilibrium of the one-shot game with $P_1(N_1) = P_2(N_2)$, where no single individual can reach a better payoff by changing the strategy, when all others stick to their strategy. However, street capacity would be better used, if people would establish the system optimum characterized by a maximization of the average group payoff $\overline{P} = [N_1 P_1(N_1) + N_2 P_2(N_2)]/(N_1 + N_2)$. The problem of this usage pattern is that some individuals will get less payoff than in the user equilibrium and less than others, i.e. the system optimum is felt to be "unfair".

Nevertheless, there is a fair and system-optimal solution of the iterated route choice problem: an alternating cooperative usage pattern, where everyone uses the faster road in a certain fraction of cases, while otherwise using the slower road. The question is whether this pattern will actually evolve in the course of time and how coordination would take place. In order to study this experimentally, we have focused on the two-person route-choice game with the payoffs $P_{11} = P_1(2) = 0$, $P_{12} = P_1(1) = 300$, $P_{21} = P_2(1) = -100$, and $P_{22} = P_2(2) = -200$ (see Fig. 1c, d). Altogether we have carried out more than 80 route choice experiments, all with different participants. In the 24 two-person [12 four-person] experiments evaluated here (see Figs. 2 to 4), test persons were instructed to choose between two possible routes between the same origin and destination. They knew that route 1 would correspond to a 'freeway' (which may be fast or congested), while route 2 would represent an alternative route (a 'side road'). Test persons were also informed that, if two [three] participants would choose route 1, everyone would receive 0 points, while if half of the participants would choose route 1, they would receive 100 points on average. but 1-choosers would profit at the cost of 2-choosers. Finally, participants were told that everyone could reach an average of 100 points per round with variable, situation-dependent decisions, and that the (additional) individual payment after the experiment would depend on their cumulative payoff points reached

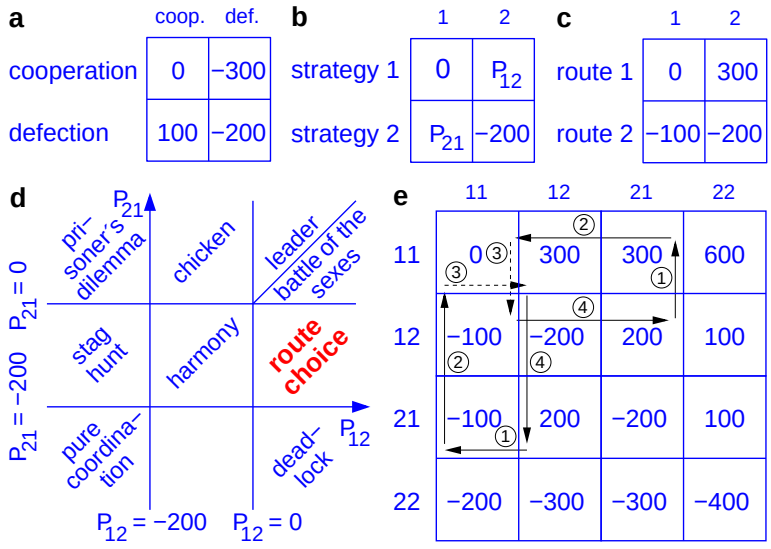


Figure 1: Symmetrical two-person games can be represented by a payoff matrix of the form $\mathbf{P} = (P_{ij})$, where P_{ij} is the success ("payoff") of person 1 in a one-shot game when choosing strategy i and meeting strategy j. The respective payoffs of the second person are given by the symmetrical values P_{ji}. (a) Payoff matrix corresponding to the prisoner's dilemma. (b) General payoff matrix for symmetrical two-person games with two alternatives. The payoffs P_{11} and P_{22} can, for example, be transformed to the values 0 and -200, while the two parameters P_{12} and P_{21} are variable [4]. (c) Payoff matrix $\mathbf{P} = (P_{ij})$ of the one-shot route choice game defined by the conditions $P_{12} > P_{11} > P_{21} > P_{22}$. A strategical conflict results when $P_{12} + P_{21} > 2P_{11}$, so that the system optimum differs from the user equilibrium. Despite some common features, this game has to be distinguished from the minority game [2], as a minority decision for alternative 2 is less profitable than a majority decision for alternative 1. (d) Extended Eriksson-Lindgren scheme of two-person games [4]. (e) Payoff matrix $(P^{(2)}_{(i_1 i_2),(j_1 j_2)}) = (P_{i_1 j_1} + P_{i_2 j_2})$ of the route choice super game with two-period decisions. The analysis of the one-shot two-person route choice game, see c), suggests that the user equilibrium (with both persons choosing route 1) would establish. Once the user equilibrium is reached, no-one can get a higher payoff by changing the decision, if the other person does not change as well. For two-period decisions, see e), the system optimum (strategy 12 meeting strategy 21) corresponds to a user equilibrium, but one person can increase the payoff at the cost of the other (see arrow 1). A change of the other person's decision can punish this egoistic behaviour (arrow 2), which is likely to establish the user equilibrium with payoff 0. In order to leave this state again in favour of the system optimum, one person will have to make an "offer" at the cost of a reduced payoff (arrow 3). If the other person reciprocates this offer (arrow 4), the system optimum is established again. The time-averaged payoff of this cycle lies below the system optimum.

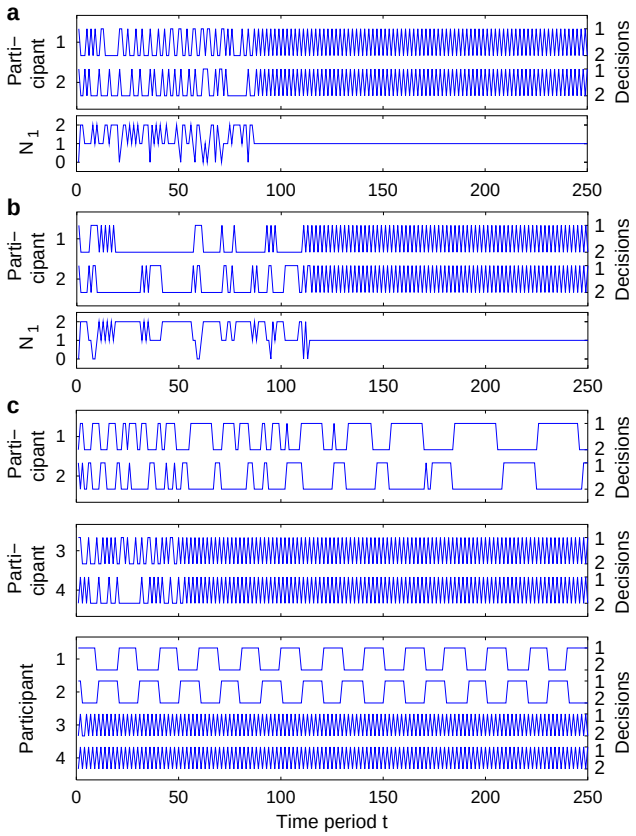


Figure 2: (a) Experimentally observed 1- and 2-decisions of both individuals in a two-person route choice experiment with the parameters specified in Fig. 1c, and corresponding number N_1 of 1-decisions. The system optimum is reached for $N_1 = 1$, the user equilibrium for $N_1 = 2$. Despite the initial preference for route 1 corresponding to a tendency to establish the user equilibrium (see Fig. 4a), route 2 was sometimes checked out in a more or less random way. The irregular changes indicate that most individuals did not have the idea that their average payoff would be maximized by a periodic oscillatory behaviour. However, sooner or later individuals chose routes in a way that a change to route 2 (an "offer") was reciprocated by a cooperative move by the other individual, while in the same iteration the offering individual changed back to route 1. (b) Representative example of route choice decisions simulated with the reinforcement learning model described in the text. For $\nu_l^0 = \nu_l^1 = 0$, no emergent cooperation is found. $\nu_l^0 > 0$ or odd values of n_l produce intermittent breakdowns of cooperation. A small, but finite value of ν_l^1 is important to find a transition to persistent cooperation. Here, we have chosen $\nu_l^1 = 0.08$, $q_l = 1$, $\nu_l^0 = 0$, and $n_l = 2$. (c) Experimentally observed decision behaviour when two groups of two-person experiments afterwards played a four-person game with $C_1 = 900$, $D_1 = 300$, $C_2 = 100$, $D_2 = 100$. Instead of oscillations of period 2, another alternating patterns corresponding to n-period decisions with $n > 2$ emerged in one of the two-person games. After all persons had learnt oscillatory cooperative behaviour, the four-person game just required synchronization (coordination), but not the invention of a cooperative strategy. Therefore, persistent cooperation was quickly established (in contrast to our four-person experiments with new participants). It is clearly visible that the test persons continued to apply similar decision strategies as in the previous two-person experiments.

in 300 rounds (100 points = 0.01 EUR).

The user equilibrium of the 2-person game corresponds to both individuals using route 1 (the 'dominant strategy'), resulting in a payoff of 0. However, in order to reach the system optimum of $(-100 + 300)/2 = 100$ per iteration, one individual has to leave the freeway for one iteration, which yields a reduced payoff of -100 in favour of a high payoff of 300 for the other individual. To be profitable also for the first individual, the other one should reciprocate this "offer" by switching to route 2 in the next iteration, while the first individual returns to route 1. Establishing this oscillatory cooperative behaviour yields 100 extra points on average. If the other individual is not cooperative, both will be back to the user equilibrium of 0 points only, and the uncooperative individual has temporarily profited at the cost of the offering individual (see Fig. 1e). This makes offers for cooperation and, therefore, the establishment of the system optimum unlikely. In spite of this, many experimental time series show a transition to coherent oscillatory behaviour after some time period (see Fig. 2). These cooperative oscillations are to be distinguished from oscillations with reduced system performance due to coordination problems [19] and from cycles in the predator-prey- or rock-paper-scissors games [3], which are predicted by the corresponding game-dynamical equations [20].

The innovation of oscillatory behaviour requires not only a gain in average payoff, but also random changes ("trial-and-error behaviour") and the consideration of multi-period decisions. Instead of just 2 one-period alternative decisions 1 and 2, there are 2^n different n-period decisions. In the two-person route choice game, an encounter of the two-period decision 12 with 21 establishes the system optimum and yields equal payoffs for everyone (see Fig. 1e). Such an optimal and fair solution is not possible for one-period decisions. Yet, the interaction of 12 with 21 ("cooperative episode") is not stable, as individuals can temporarily increase their own payoff by changing their decision to 11 (see Fig. 1e). For this reason, the first cooperative episodes do often not persist (see Fig. 3). However, selfish behaviour can be punished by the other individual by changing to route 1 as well (see Fig. 1e). In this way, persistent cooperation is established after a number of cooperative episodes. In our two-person experiments, the cumulative distribution of required cooperative episodes could be mathematically described by the logistic curve

$$F(n) = 1/[1 + c_N \exp(-d_N n)] \tag{1}$$

with $c_2 = 3.4$ and $d_2 = 0.17$ (see Fig. 3a). Moreover, if the system optimum corresponds to an equal distribution over both alternatives, based on a stochastic model, the expected time interval T until a cooperative episode among $N = N_1 + N_2$ participants occurs can be statistically estimated by the formula

$$T = 2^N \frac{(N/2)!^2}{N!} \prod_{l=1}^{N} \frac{1}{\overline{\nu}_l}\,, \tag{2}$$

where $\overline{\nu}_l$ denotes the average changing rate of individual l until persistent cooperation starts (see Fig. 3b).

Our observations can be qualitatively reproduced by a reinforcement learning model reflecting success- and history-dependent individual decision behaviour [21] (see Figs. 2b

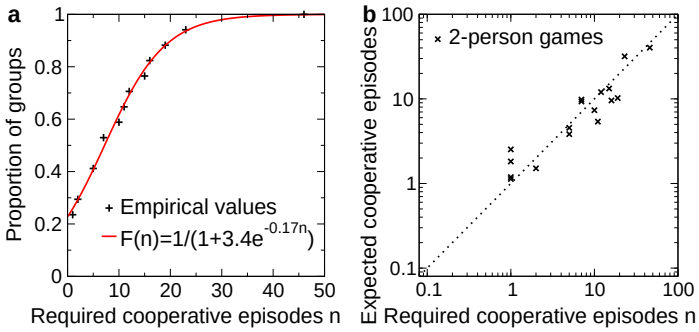


Figure 3: (a) Cumulative distribution of required cooperative episodes until persistent cooperation is established, given that cooperation occurs within 300 time periods (as in 17 out of 24 two-person experiments). The experimental data are well approximated by a logistic curve. (b) Comparison of the required number of cooperative episodes with the expected number of cooperative episodes (approximated as occurence time of persistent cooperation, divided by the expected time interval T until a cooperative episode occurs by chance). The linear regression to the empirical data points supports formula (2), which is also consistent with our 4-person experiments and with the results of our reinforcement learning model.

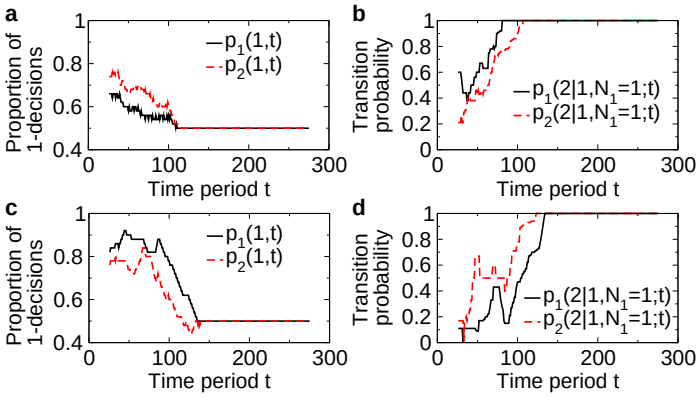


Figure 4: (a) Proportion $p_l(1,t)$ of 1-decisions of both participants l in the two-person route choice experiment displayed in Fig. 2a. (b) Transition probability $p_l(2|1,1;t)$ of person l from route 1 (the "freeway") to route 2, when the other person has chosen route 2, averaged over a time window of 50 time periods. The steep transition from small values to 1 for the experiment displayed in Fig. 2a is characteristic and illustrates the evolution of cooperativeness. (c) Proportion $p_l(1,t)$ of 1-decisions of both participants in the simulated route choice game shown in Fig. 2b. The simulation is based on the reinforcement learning model described in the text. (d) Transition probability $p_l(2|1,1;t)$ of person l for the simulation result shown in Fig. 2b.

and 4). In contrast to mixed strategies, the description of coherent decisions and persistent cooperation requires an almost deterministic model, but some weak stochasticity is needed for the exploration of innovative strategies and the emergence of cooperation. We denote person l's probability to choose decision j at time $t+1$ by $p_l(j|i, N_1; t)$, when i was selected at time t and $N_1(t)$ persons had chosen alternative 1. Moreover, we assume that p_l is either 0 or 1, corresponding to clear (deterministic) preferences. The decision behaviour is assumed to be switched with probability q_l, if the average payoff since the last comparable situation with $i(t') = i(t)$ and $N_1(t') = N_1(t)$ at time $t' < t$ is less than the average individual payoff $\overline{P}_l(t)$ during the last n_l time periods. This replacement of dissatisfactory strategies orients at historical long-term profits and avoids short-sighted changes after temporary losses. Moreover, the decision behaviour is randomly switched with probability

$$\nu_l(t) = \nu_l^0 + \nu_l^1 \max[0, 1 - \overline{P}_l(t)/100] \tag{3}$$

('trial and error behaviour'). $\nu_l^0 \approx 0$ denotes the individual mutation rate in the system optimum, while $\nu_l^1 > 0$ reflects the mutation rate in the user equilibrium. In our simulations, we varied only the parameter ν_l^1, while we chose the simplest possible specification of the other parameters $\nu_l^0 = 0$, $q_l = 1$, $n_l = 2$ and initial conditions $p_l(2|1, N_1; 0) = 0$ and $p_l(1|2, N_1; 0) = 1$. The simulation results reflect many features of our route choice experiments (see Figs. 2b, 4).

Formula (2) gives a good estimate of the time interval T needed for persistent cooperation and its variation with the changing rate and the number N of persons. Route choice experiments confirm that T strongly increases with the system size N. Therefore, spontaneous cooperation is unlikely to emerge in real traffic systems, in accordance with observations. However, cooperation could be rapidly established by means of novel traveller information systems, which would avoid the slow learning process (2). Moreover, while we do not recommend conventional congestion charges, a charge for unfair usage patterns would support the compliance with individual route choice recommendations. It would substitute the inefficient individual punishment mechanism. In systems with many similar routing decisions, a Pareto optimum characterized by coherent oscillations could be spontaneously established by suitable reinforcement mechanisms. This may help to enhance data routing [22]. and to resolve Braess-like paradoxes [23] in networks [24].

For a more detailed analysis see [25].

References

[1] Berg J, Engel A: Matrix Games, Mixed Strategies, and Statistical Mechanics. Phys Rev Lett 1998;81:4999–5002. Ebel H, Bornholdt S: Coevolutionary games on networks. Phys Rev E Stat Nonlin Soft Matter Phys 2002;66:056118. Metzler R, Engel A: Jamming transitions and avalanches in the game of Dots-and-Boxes. Phys Rev E 2002;65:066108. Tomochi M, Kono M: Spatial prisoner's dilemma games with dynamic payoff matrices. Phys Rev E 2002;65:026112. Choe SC, Johnson NF, Hui PM: Error-driven global transition in a competitive population on a network.

Phys Rev E 2004;70:055101. Fort H, Viola S: Self-organization in a simple model of adaptive agents playing 2 x 2 games with arbitrary payoff matrices. Phys Rev E 2004;69:036110.

[2] Challet D, Zhang Y-C: Emergence of cooperation and organization in an evolutionary game. Physica A 1997;246:407–418. Challet D, Marsili M: Relevance of memory in minority games. Phys Rev E 2000;62:1862–1868. Marsili M, Mulet R, Ricci-Tersenghi F, Zecchina R: Learning to coordinate in a complex and nonstationary world. Phys Rev Lett 2001;87:208701. Mansilla R: Algorithmic complexity in the minority game. Phys Rev E 2000;62:4553–4557. Lo TS, Hui PM, Johnson NF: Theory of the evolutionary minority game. Phys Rev E 2000;62:4393–4396. Lo TS, Chan HY, Hui PM,Johnson NF: Theory of networked minority games based on strategy pattern dynamics. Phys Rev E 2004;70:056102. Yuan B, Chen K: Evolutionary dynamics and the phase structure of the minority game. Phys Rev E 2004;69:067106. Chen K, Wang B-H, Yuan B: Adiabatic theory for the population distribution in the evolutionary minority game. Phys Rev E 2004;69:025102.

[3] Szolnoki A, Szabo G: Phase transitions for rock-scissors-paper game on different networks. Phys Rev E 2004;70:037102. Ravasz M, Szabo G, Szolnoki A: Spreading of families in cyclic predator-prey models. Phys Rev E 2004;70:012901. Szabo G, Sznaider GA: Phase transition and selection in a four-species cyclic predator-prey model. Phys Rev E 2004;69:031911.

[4] Eriksson A, Lindgren K: Cooperation in an unpredictable environment; in Standish RK, Bedau MA, Abbass HA (eds): Proceedings of Artificial Life VIII. Sidney, MIT Press, 2002 pp 394–399 (and poster available at http://frt.fy.chalmers.se/cs/people/eriksson.html).

[5] Axelrod R, Hamilton WD: The evolution of cooperation. Science 1981;211:1390–1396.

[6] Axelrod R, Dion D: The further evolution of cooperation. Science 1988;242:1385–1390.

[7] Glance NS, Huberman BA: Dynamics with expectations. Physics Letters A 1992;165:432–440.

[8] Glance NS, Huberman BA: The outbreak of cooperation. Journal of Mathematical Sociology 1993;17(4):281–302.

[9] Queller DC: Kinship is relative. Nature 2004;430:975–976.

[10] Nowak MA, Sigmund K: Evolution of indirect reciprocity by image scoring. Nature 1998;393:573–577.

[11] Riolo RL, Cohen MD, Axelrod R: Evolution of cooperation without reciprocity. Nature 2001;414:441–443.

[12] Nowak MA, Sasaki A, Taylor C, Fudenberg D: Emergence of cooperation and evolutionary stability in finite populations. Nature 2004;428:646–650.

[13] Schweitzer F, Behera L, Muehlenbein H: Evolution of cooperation in a spatial prisoner's dilemma. Advances in Complex Systems 2002;5(2/3):269–299.

[14] Szabo G, Hauert C: Phase transitions and volunteering in spatial public goods games. Physical Review Letters 2002;89:118101.

[15] McNamara JM, Barta Z, Houston AI: Variation in behaviour promotes cooperation in the prisoner's dilemma game. Nature 2004;428:745–748.

[16] Schreckenberg M, Selten R (eds): Human Behaviour and Traffic Networks. Berlin, Springer, 2004.

[17] Helbing D, Schoenhof M, Kern D: Volatile decision dynamics: Experiments, stochastic description, intermittency control, and traffic optimization. New Journal of Physics 2002;4:33.1–33.16.

[18] Wardrop JG: Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers II 1952;1:325–378.

[19] These cooperative oscillations are to be distinguished from the interesting heterogeneity-related oscillations due to coordination problems which are reported by Iwanaga S, Namatame A: The complexity of collective decision. Nonlinear Dynamics, Psychology, and Life Sciences 2002;2:137–158.

[20] Hofbauer J, Sigmund K: The Theory of Evolution and Dynamical Systems. Cambridge, Cambridge University Press, 1988.

[21] Macy MW, Flache A: Learning dynamics in social dilemmas. Proc Natl Acad Sci U S A 2002;99(3):7229–7236.

[22] Wolpert DH, Tumer K: Collective intelligence, data routing and Braess' paradox. Journal of Artificial Intelligence Research 2002;16:359–387.

[23] Braess D: Über ein Paradoxon der Verkehrsplanung [A paradox of traffic assignment problems]. Unternehmensforschung 1968;12:258–268 (For about 100 related references see http://homepage.ruhr-uni-bochum.de/Dietrich.Braess/#paradox).

[24] Cohen JE, Horowitz P: Paradoxial behaviour of mechanical and electrical networks. Nature 1991;352:699–701.

[25] Helbing D, Schoenhof M, Stark HU, Holyst JA: How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoner's dilemma. Advances in Complex Systems 2005;8:87-116.

Noise Sensitivity of Portfolio Selection under Various Risk Measures

I. Kondor[1,*], Sz. Pafka[2], and G. Nagy[2]

[1]Collegium Budapest – Institute for Advanced Study, H-1014 Szentharomsag u. 2, Budapest, Hungary
[2]Risk Management Department, CIB Bank, H-1027, Medve u. 4-14, Budapest, Hungary
*Corresponding author: kondor@colbud.hu

Keywords: estimation noise, random matrix theory, risk management, portfolio optimization

Extended Abstract

The theory of portfolios, initiated by Markowitz, has suffered from the "curse of dimensions" from the very outset. Whereas the computational difficulties associated with the selection of the optimal portfolio have been greatly alleviated by the progress of information technology, the fundamental problem of insufficient input data and the resulting estimation error remain serious stumbling blocks. Even if we disregard the notoriously difficult problem of estimating returns and focus exclusively on the minimal risk portfolio, the amount of information contained in the available finite-length time series is typically far below the amount of information necessary for the reliable determination of the optimal portfolio. As for a portfolio of size N and time series of length T, the number of input data is NT, whereas the number of data needed for the construction of the covariance matrix is $O(N^2)$, we expect that the quality of the estimate essentially depends on the ratio N/T and that the error goes to zero only in the limit of very small N/T. Now the problem is that for typical bank portfolios N/T is never sufficiently small, in fact, it may well be larger than unity, the threshold value where the covariance matrix becomes singular and the portfolio selection problem meaningless.

Over the past decades a large number of different techniques have been developed to tackle this problem and reduce the effective dimension of large bank portfolios, but the efficiency and reliability of these procedures are hard to assess or compare. In this paper we propose a model (simulation)-based approach which can be used for the systematic testing of all these dimensionality-reduction (filtering) techniques [1]. To illustrate the usefulness of our framework, we develop several toy models (including a single-index or market model and a market plus sectors model) that display some of the main characteristic features of empirical correlations and generate artificial time series from them. Then, we regard these time series as empirical data and reconstruct the corresponding correlation matrices which will inevitably contain a certain amount of noise, due to the finite length of the time series. For multivariate normal portfolios and asymptotically large N and T with N/T fixed we derive a simple analytic formula for the relative error in the portfolio.

Now we apply several correlation matrix estimators and dimensionality-reduction techniques introduced in the literature and/or applied in practice. As in our artificial world the only source of error is the finite length of the time series and, in addition, the "true" model, hence also the "true" correlation matrix, are precisely known, we can meaningfully compare the performance of the various noise-reduction techniques. One of our recurrent observations [2] is that the recently introduced filtering technique based on random matrix theory (RMT) [3] performs consistently well in all the investigated cases. Based on this experience, we believe that our simulation-based approach can also be useful for the systematic investigation of several related problems of current interest in finance.

In addition to correlated Gaussian returns, we also consider non-stationary time series of the IGARCH(1,1) type, closely related to the exponentially weighted moving average technique implemented in RiskMetrics [4]. In order to be able to apply the RMT-based filtering technique in this context, we have derived the spectrum of a random covariance matrix where the returns are

exponentially weighted with time [5]. Applying this method to *empirical* data we find that the effect of risk depends on the weight factor, whose optimal value corresponds to the trade off between discarding too many past data, thereby destroying the statistics, or retaining too many, hence including non-stationary effects. We determine the optimal weight factor and find that it is considerably larger than the value advocated in RiskMetrics.

As a further attempt to go beyond the classical mean-variance framework, we have also studied the effect of noise on portfolio selection under some alternative risk measures. In particular, we have studied the case of mean absolute deviation (MAD), as described in ref. [6]. The level surfaces of risk under MAD are polyhedrons (instead of the ellipsoidal iso-risk surfaces corresponding to variance), and this leads to an increased sensitivity to noise. We observe a similar effect under the use of expected shortfall (ES) or conditional value at risk which is strongly promoted in the academic literature as the simplest of the coherent risk measures [7]. In addition, portfolio optimization under this measure has been shown to be reducible to linear programming [8] which might, in principle, allow one to optimize extremely large portfolios at a relatively light computational cost. As we show here, the downside is a strongly increased sensitivity to noise. One might think that this is due to the fact that ES, as a kind of conditional expectation, omits a large amount of input data by concentrating only on those above a (typically high) confidence level. A systematic study of the problem reveals, however, that the enhanced sensitivity remains true even for as low confidence levels as 60% or 50%, where ES can be compared with semi-variance. The fundamental reason of this high noise-sensitivity of ES is not understood at present. At the other extreme, for confidence levels approaching 100%, we have a risk measure that can be called worst loss (WL). Although over-pessimistic, this still has the virtue of coherence. Not surprisingly, WL is found to be very sensitive to noise again.

In the course of our studies of the noise-sensitivity of the risk measures ES and WL we have observed a striking phenomenon. As we have already mentioned, the portfolio selection problem within the mean-variance framework becomes meaningless for $N/T > 1$. The same is true for all the other risk measures studied in this paper. On the other hand, for $N/T < 1$, optimization under the variance and also under MAD always has a solution, even if it may be strongly influenced by noise for N/T not small enough. In contrast to this, the optimization under ES and WL does not necessarily have a solution even under the threshold $N/T = 1$, instead, the existence of a solution becomes a probabilistic issue: it depends on the sample. We have studied this remarkable phenomenon both analytically and numerically. In the case of Gaussian-distributed (or, more generally, elliptically distributed) assets and under the WL risk measure we have been able to derive a closed formula for the probability of the optimization problem to have a solution and found that this probability goes to unity only for N/T going to zero. Similar behaviour is observed in numerical simulations for the ES measure for the entire range of confidence levels we have studied. This puzzling phenomenon is absent if short selling is excluded, which may be the reason why it had not been observed by previous authors.

References:

[1] Elton E, Gruber M: Modern portfolio theory and investment analysis. New York, J. Wiley and Sons, 1995.
[2] Pafka S, Kondor I: Estimated correlation matrices and portfolio optimization. Physica A 2004; 343: 623-634.
[3] Laloux L, Cizeau P, Bouchaud J-P, Potters M: Noise dressing of financial correlation matrices. Physical Review Letters 1999;83:1467-1470, and Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Stanley HE: Universal and nonuniversal properties of cross correlations in financial time series. Physical Review Letters 1999;83:1471-1474.
[4] RiskMetrics Group: Technical document, New York, J. P. Morgan, 1994.
[5] Pafka S, Potters M, Kondor I: Exponential weighting and Random-Matrix-Theory-based filtering of financial covariance matrices for portfolio optimisation. http://lanl.arxiv.org/abs/cond-mat/0402573 2004, submitted to Quantitative Finance.
[6] Konno H, Yamazaki H: Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. Management Science 1991;37:519-531.
[7] Artzner P, Delbaen F, Eber J, Heath RD: Thinking coherently. Risk 1997;10:68-71.
[8] Rockafellar R.,Uryasev S: Optimization of conditional value-at-risk. Journal of Risk 2000;2:21-41.

Reflexivity as a constitutive property of a complex urban system

Sylvie Occelli, Luca Staricco
luca.staricco@polito.it

Applying the notion of complex system to cities and territories has the unwarranted advantage to make it possible to establish several conceptual links between the many facets of urban systems, like openness, self-organization, pro-activity and awareness. While the features concerning openness and self-organization have been extensively scrutinized as constitutive determinants of urban complex systems, those concerning pro-activity and awareness have been generally viewed as external. At most, they have been considered as desirable features of a group of agents, i.e. experts, planners and educated people, entitled to steer the evolution of the urban system towards more desirable and sustainable paths of change. Pro-activity and awareness mean that there exists a reflexive property, which is intrinsically constitutive of an urban system,. i.e. a collective agent constituted by an interacting group of social cognizant agents supported by a physical, social, technological environment. In this regard, Maturana's arguments about the drive of the explanatory search for understanding which characterizes a living being can be understood as an internal drive of the urban system. A Reflexive Urban System (RUS), therefore, is one which: * is able to think about of its knowledge own generative determinants (i.e., the agents involved, the process through which knowledge is yielded, and how it is encoded and decoded in the urban system); * permanently seeks to adapt (improve) those determinants in the course of its evolution. A major challenge for a complexity approach to urban systems, therefore, is that a methodology for a RUS should be viewed as an embedded dimension of the system itself, i.e. it should be an internal endeavour steering the system's own evolution. In order to fully account for the pro-active feature characterizing a RUS, therefore, one cannot be satisfied with an evaluation of the system behavioural performances but has to constructively engage in the knowledge process of building the RUS's own methodology. In addition, as a RUS is a collective entity, the methodological framework cannot help resulting from a collectively determined process. According to this line of reasoning, a claim is made that one major function of applying a complexity approach would be to sustain the systemic coherence necessary to the RUS line of enquire, i.e. to stir its drives in the explanatory search for understanding and help avoiding the constraints and prejudices. Finally it is suggested that underlying a complexity approach for a RUS three major dimensions might be regarded as principal leverages: * recognition, the identification of knowledge needs in order to sustain a pro-active endeavour; * guidance, how to make effective the recognized knowledge needs, in order to get involved in a purposeful line of enquiry ; * capability, the

achievements of social valued outcome for somebody who cares.

A Multi-Level Model for Spatial Dynamics of Systems of Cities through Innovation Processes

Denise Pumain
pumain@parisgeo.cnrs.fr

We develop here an evolutionary theory of urban systems as complex systems characterised by a hierarchical structure (following a Zipfs distribution of city sizes and a more or less regular geographical pattern), a functional diversity and various scaling laws according to different urban functions. These properties are summarised in a multi-agents systems called SIMPOP2, that is a generic model for simulating the dynamics of a variety of settlement systems while they emulate and adopt innovations. The basic conception is that a hierarchic urban system emerges from the spatial interactions between settlements over a long time period. Interactions are driven by competition for promoting and capturing the benefits of innovations (territories, market zones or networks according to each type of urban function). They are submitted to different social and political contexts or functional rules, evolving through time. Most of interactions are defined at the meso level of the urban agglomerations but some of them, as urban governance or innovation, may be acted by agents of a lower level or institutions at higher levels. The model helps identifying which key parameters (at micro, meso or macro level) can lead to the emergence of the three major types of observed settlement systems: regular and dense, but more or less concentrated (European or Asiatic countries) ; dual and primate (developing countries); strongly hierarchised and sparse (countries of the New world). More detailed versions of SIMPOP2 can be adapted for calibrating observed evolutions (cf. EUROSIM).

Complex-city: the shift of urban science from classic to an evolutionary approach

Giovanni A. Rabino
giovanni.rabino@polimi.it

In an effort of interdisciplinary comparison, the aim of this short paper is to illustrate how the new concepts (and methods) of complexity are applied in the science of territory (i.e. cities and towns systems). Being territorial science at the same time a theoretical (e.g. urban geography) and an applied science (e.g. city planning), the presentation is organised in two parts: complexity in urban system analysis and complexity in town planning and design. Note that in the paper all points below are argued referring to real case-studies or experimental applications. Moving from first principles, we describe how urban analysis, namely mathematical models of urban system, leaving from a static functionalist approach (Lowry, 1964), through dynamic modelling, got to a co-evolutionary description of a set of territorial processes (see, for instance, Pumain, 2004). In this framework, it is shown that many classical principles of spatial organisation, such as cities rank-size law or Christaller rule, are better re-interpreted or refined in term of complex phenomena (scale invariance, emergence of order 8230;). As to the methods of complexity used in urban analysis, we focus on two aspects: - fractal morphology, to describe and explain new trends in urbanisation: sprawl and smart growth; - multi-agent modelling, as a powerful tool for catching many non-linear interactions among the various stakeholders in the city. Going to consider urban planning, first of all it is argued that the traditional net divide between urban analysis and planning no more holds. In the light of complexity, analysis and planning, like many other learning - creativity phenomena, are strictly joined in a (individual and social) evolutionary (conscious and intentional) decision process. In this context, we present: -at a general level, the complexification of planning practices (public participation, governance, e-planning, etc.) in relation to a more complex society and innovative technologies; - at a more specific level, some new planning methods and tools (strategic planning, "best practices" methodology, etc.) directly derived from the new evolutionary approach to the science of the city.

Modelling urban networks dynamics with multi-agent systems

Lena Sanders
lena.sanders@parisgeo.cnrs.fr

The paper concerns the geography of system of cities in a dynamic perspective. Already in 1964 Berry wrote a paper called Cities as systems within system of cities, and obviously the evocated systems were complex, their form and their evolution resulting from sets of interactions operating at different scales. The EUROSIM model (a specific version of SIMPOP2 model), built up within a multi-agent framework, integrates accumulated knowledge about urban systems. The purpose is to simulate the evolution of the European system of cities during the past 50 years, and to explore its future dynamics during next half century according to different scenarios.

The hypothesis is that the properties of the urban system at a macro-geographical level emerge from the interactions between entities of lower geographical levels. The choice of the lower level to be formalized is first discussed. In social sciences, the individuals, the households, or the firms are often considered as the elementary entities for modeling a social system. In the Eurosim model, the cities themselves, as collective entities, are represented by agents, and their ontological state has then to be precised.

The functioning of the model is then presented. The cities larger than 200 000 inhabitants constitute the driving layer of the urban system. Depending on its specialization each city is connected to a set of cities through associated networks which are evolving through time. The mechanisms of supply, demand and exchanges which determine the interactions between the cities are formalized using the protocol of communication of the multi-agent system. The dynamics of each city depend on its ability to gain wealth through successful exchanges, while the global growth is driven by innovation. Multiscalar tools for analyzing the outputs of the model are presented and the question of calibration is discussed.

Analysing the resilience of complex resources management systems – a stylised simulation model of human-nature interactions in a river basin

Maja Schlüter[1] & Claudia Pahl-Wostl[2]

[1] *UFZ - Centre for Environmental Research Leipzig-Halle GmbH, Permoserstr. 15, 04318 Leipzig, Germany, Department Ecological Modelling (OESA)*
[2] *University of Osnabrueck, Institute of Environmental Systems Research, 49069 Osnabrueck, Germany*

The uncertainties of global change and the complexity and unpredictability of the dynamics of social-ecological systems demand for new approaches to ecosystem and resources management. Rather than attempting to predict and control natural variability those approaches focus on strengthening the flexibility of a system to cope with unexpected change. The aim is to enhance system resilience and its capacity to adapt. Resilience is the potential of a system to remain in a particular configuration, to maintain its feedbacks and functions and to reorganise following disturbance driven change. Resilience is seen as an important stability property that determines the system's capacity to adapt to and benefit from change. A variety of mechanisms of resilience specific for different systems have been identified, such as genetic and biological diversity, redundancy and modularity, the capacity to learn and store knowledge and experience, to create flexibility in problem solving and balance power among interest groups. In the context of natural resources management a system's capacity to learn from and adaptively respond to stress emerges from interactions between biophysical and social processes. A sound understanding of mechanisms determining the resilience of coupled social-ecological systems is the basis for adaptive management of natural resources.
Simulation models of human-environment interactions are valuable tools to study factors and mechanisms that determine the resilience of a complex resources management system to stresses such as high variability in resources availability, extreme events and long-term changes. We develop a model of a coupled social-ecological system in a river basin to explore the influence of the organisational structure of water management, of cross-scale interactions among actors, information storage and flows, and other factors for the resilience of the coupled system. Special emphasis is put on the role of feedbacks between the human and the environmental system. The aim is to study the interrelationship between system structure and functioning treating the social-ecological system as a complex adaptive system. The model represents a stylized water management system, based on the context of the Amudarya river delta, where water resources are used to sustain agriculture as well as semi-natural fish populations. Besides irrigated agriculture fish is an additional source of income. The model combines a model of a water flows network with an equation-based model of a simplified aquatic ecosystem and an agent-based model of decision making and resources exploitation composed of simple rules. Decision making and information flows are represented at different scales (local, regional, national). The success of individual agents as well as the state of the human-used ecosystem depends on local water availability. Simulation experiments are carried out to test different settings of the agent's ability to obtain information and manage the resource in the face of different levels of variability and uncertainty in water inflow to the region. The resilience (measured as the global and local achievement of production goals, state of the ecosystem, etc.) of the different management regimes is compared. It is expected that the buffer capacities of a water reservoir and the ecosystem are important factors determining the resilience of the system. Access to and transfer of information and learning processes among agents are major factors influencing the capacity of the system to adapt. In the presentation the model concept and first simulation results are presented.

THE EVOLUTION OF FREE/LIBRE AND OPEN SOURCE SOFTWARE LICENSES : A DYNAMIC MODEL

Jean-Baptiste Soufron[*], Jean Sallantin[†]

12 mai 2005

Résumé

We address the understanding of the structure of contracts and their evolution via a dynamic model based on the data gathered through the creation of Free/Libre Open Source contracts and their evolution. We then show how this model can be used in order to facilitate licensing negotiation for players.

Introduction

There has been an increasing interest from lawyers and social scientists for the development of legal tools adapted to share knowledge. To this end, several contractual tools have been proposed for allowing the best uses of intellectual property law, notably Free/Libre Open Source Software licenses (or FLOSS licenses). The legal applicability of these licenses have been well studied but there has been no model proposed to explain their creation and their evolution. In this view, it is of utermost interest to propose tools enabling people to understand the processes at work and use them for their best interest in contractual negotiations. Existing approaches in community findings are either based on expert systems or on classical contractual approach. There has been roughly no attempt to link different licenses between them under a dynamic model. Actually, the various study of FLOSS often concentrates on a very small number of licenses. We give a formal framework for understanding the creation of these contracts and their evolution, using a model based on the realistic approach that contracts are legal answers to problems raised by users, and that new contracts are obtained by combining the clauses of ancient ones or by creating new ones. Suggesting that this leads to the emergence of a legal domain throught the abduction of its terms and the induction of the constraints linking them, we eventually propose to precise this model to help people to better negotiate licenses and contracts. Our main source of data is Sourceforge[1], the world's largest Open Source software online repository providing free hosting to nearly 100 000 projects[2]. This data is summarized and publicly accessible on their website but we precise our study using the various mailing-lists associated with the projects. Eventually we add some data from external FLOSS sources when it is available.

[*]CERSA (Centre de Recherche en Science Administrative), 10 rue Thénard 75005 Paris France, corresponding author: jean-baptiste.soufron@cersa.org

[†]LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier), 161 rue Ada 34392 Montpellier Cedex 5 France

[1]http ://sourceforge.net/

[2]following the data available in April 2005, there were 99 803 projects accounting for 1 068 803.

1 Evolutive contracts

1.1 Law in Text

The first examples of Law as a text were quite complicated. Hamurabi's Code for example, contained 282 different rules, each of them being a legal answer to practical problems. It regulated various issues like the right of the poor to seek redress from wrongs committed by the wealthy or by the nobility, the rights of women, including the rights of women to own property in their own names, and even the right to divorce a husband for grounds which would be recognized even today. But only the questions within the list provided by the Code could receive a solution. It was easy to apply but when any new question were to arise, the last article of the Code precised that the king was the only competent authority to answer it. At that time, Law was a complicated set of norms coupled with a top-down evolution process, difficult to change and to individualize.

1.2 Law in Action

Law today offers many other ways to regulate texts just setting out plain rules for citizens and stating that some authority will answer new questions. People are now allowed to create new answers and contracts are a good example as being law in action, created through a normative process emanating from numerous sources interacting to create Agreements.

Contracts Contracts are Law to the parties[3]. They are agreements containing a combination of clauses, where each clause can be described as an answer to a given legal problem[4]. As parties can create new clauses or adapt existing ones when they face new problems, contracts can be efficiently described as a source of Law. But when applied, these new clauses will be confronted with legal norms governing the validity of contracts. If they get censored by judges, parties will need to modify them or to create new ones once again. Thus, getting their legitimacy from Law and imposing their content to the judiciary process, contracts also take place into the hierarchy of norms regulating human activity and bringing previousness to human relations[5].

Evolution In this model, the dynamics of contracts can be described as a continuous evolution where people add and modify clauses to correct errors in the application of a given contract. Needing to be legally formed, Contracts can first be described as a protection for Individuals, controlling these mandatory Agreements with a certain number of limitations and constraints[6].

[3] According to Article 1101 of the French Civil Code, a Contract is an agreement by which one or several persons bind themselves, towards one or several others, to transfer, to do or not to do something. Then, as Article 1134-1 puts it, Contracts are only lawfully entered Agreements entered and they take the place of the law for those who have made them. Whether they have a specific denomination or not, they are subject to general rules which are the subject matter of this part of the Code but following Article 1135, they are binding not only as to what they express, but also as to all the consequences which equity, usage or statute give to the obligation according to its nature.

[4] F. Terré, P. Simler, Y. Lequette, Droit civil Les obligations, Dalloz, 2002

[5] H. Lécuyer, Le contrat, acte de prévision, Mélanges F. Terré, Dalloz 1998, p. 656

[6] Ch. Jamin, Plaidoyer pour le solidarisme contractuel, in Le contrat au début du XXIème siècle, Etudes offertes à J. Ghestin , L.G.D.J. 2001 ; D. Mazeaud, Loyauté, solidarité, fraternité, la nouvelle devise contractuelle ?, in L'avenir du droit, Mélanges en hommage à F. Terré, Dalloz, 1999 (dir.de) Ch. Jamin et D. Mazeaud, la nouvelle crise du contrat, Dalloz 2003 et (dir. de) L. Grynbaum et M. Nicod.

1.3 Complex contracts

Given that contracts are used on an individual basis, standard economic models imply that they should be highly complicated : rich in the expected number of payoff-relevant contingencies ; variable in the magnitude of payoffs contracted to flow between parties ; and severe in the cognitive load necessary to understand the contract. Yet most real-world contracts are simple on all this points[7]

Complex Systems Instead of being highly complicated, contracts are quite simple texts evolving through dynamics as a complex system controled by individual constraints imposed by the will of the parties, and legal constraints imposed by the Law[8]. Respecting these constraints is difficult but this evolutive mechanism disminish the cost of legal evolution and leads to more adapted rules for a better justice[9]

FLOSS licenses We will limit our analysis to the example of the evolution of FLOSS licenses and we will try to propose a framework highlighting the evolutionnary process at work in contract law. The problem is to understand how players can best negotiate these contracts, and we propose to address four points to understand the dynamics of FLOSS Contracts. First, it is important to describe the mechanism of FLOSS licenses. Then, we will describe the emergence of FLOSS legal domain as a mechanism involving the systematisation of the domain and its application. Finally, we will introduce a tool implementing this model to help writing new licenses.

1.4 FLOSS Principles and Philosophy

Free as Free Speech First of all, Free/Libre and Open Source Software doesn't mean free to use : to briefly precise the principles laid down by the GNU GPL in 1989, a license allows a holder of Intellectual Property Rights to grant the right to use his software to another. The license usually specifies the conditions under which the Software can be used or disseminated to others, payments to the licensor, whether modifications of the software are allowed, the risk and liability each party assumes, representations and warranties, and promises of support and maintenance. Just like any contract, the license must answer a vast number of questions. A source code license grant access to source code for the licensee. Open Source licenses can be defined as a specific set of terms and conditions that meet the requirements summarized by the Open Source Initiative that were only summarized in 1997.

Four requirements FLOSS licenses essentially need to answer four requirements. Each FLOSS license is a different combination of these four rights according to different and specific provisions and the ecology of Rights in an Open Source Software can be representated (Fig. 1). First is the right to access the source code - the right to get the source code and to read it. This right is the main component of any Open Source software since it is a necessary preliminary step to all of them. Every Open Source licenses grant access to the

[7] Ch. Jamin, Économie et droit in Dictionnaire de la culture juridique, D. Alland and S. Rials (dir.), Paris, PUF, 2003

[8] Ch. Jamin, ń Révision et intangibilité du contrat ou la double philosophie de l'art. 1134 du Code civil ż, Droit et patrimoine, mars 1998, p. 58

[9] A. Kronman and R. A. Posner (dir.), The Economics of Contract Law, Boston, Little, Brown and Cy, 1979 ; E. Mackaay, Analyse Économique du droit - I. Fondements, Montréal/Bruxelles, Thémis/Bruylant, 2000 ; E. Mackaay, V. Leblanc, L'économie de la bonne foi contractuelle, in Mélanges Jean Pineau, Benoît Moore (dir.), Montréal, Éditions Thémis, 2003, pp. 421-459

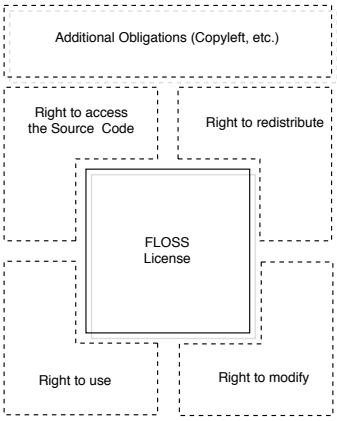


FIG. 1 – The four requirements of FLOSS licenses

Source Code of the Software. The right to modify is the right to change the source code in order to correct or to adapt it. Most Open Source licenses recognizes this right but not all of them. For example the Perl Artistic license states that if if a package is modified so that it changes from a Standard Version, then these changes must be posted in a very specific and detailed way described in the license. The right to redistribute, the right to copy and distribute the Software. Licensors often modify this right in many ways. For example, the Sun Community license allows licensees to distribute software only as long as they make no commercial gain directly from the it. Finally, the right to use is the right to compile the source code and execute it. licenses changes this clause the same way they do with the redistribution clause.

1.5 FLOSS Expansion

The primary issues in choosing license terms and conditions are whether the licensor wants revenues, whether it allows modifications to the source code and whether it allows the source code to be forked in different versions. But most licenses also accept a various number of additional obligations. For example, the Creative Commons NC licenses only authorizes to non-commercial uses of the work.

Users It follows that FLOSS is actually a complex world with a great number of licenses including complex provisions for both parties. One of the first, and the most successful one is the GNU GPL written in 1989, each new license answering new questions following the same legal creation process we are now describing. Since this date, there has been a multiplication of users and since the creation of Sourceforge repository, the number of developers has been growing steadily (Fig. 2).

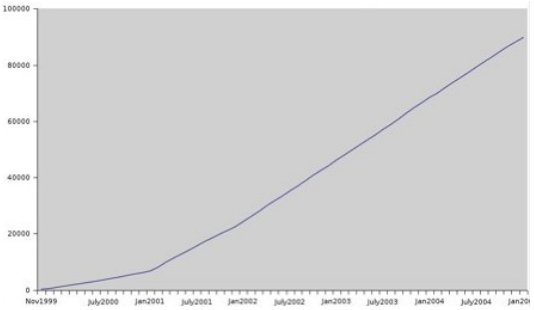


FIG. 2 – FLOSS Developers on Sourceforge between 1999 and 2005 : the number of people using Sourceforge is growing up steadily

licenses But parallel to the multiplication of open source contracts users, there is a multiplication of open source contracts as well, and there are now more than 500 licenses for Software and Content, and new ones are created every day[10]. Since 2002 for example, Creative

[10]Most of them are listed on the website of the Free Software Foundation, on the website of the Open Source Initiative or on the website of Creative Commons. If we restrict ourselves to FLOSS projects recognized by both the FSF and the OSI, there were 56 FLOSS licenses on

In the meanwhile the number of new projects adding every month is stable(Fig. 3).

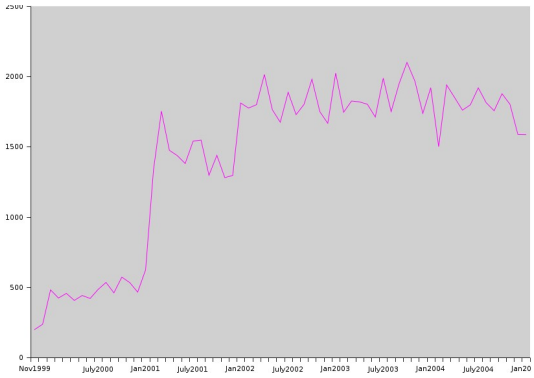


FIG. 3 – New FLOSS projects Sourceforge between 1999 and 2005 : the number of new projects added every month has been stable on Sourceforge since 2002

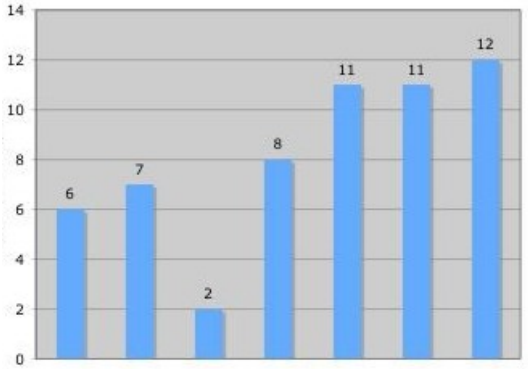


FIG. 4 – New FLOSS licenses between 1998 and 2004 : there are more new licenses created for Sourceforge projects every year

Commons alone has been producing hundreds of licenses for nearly 30 countries. Today, institutions like the CNRS are also on the process of creating their own range of license under the CECILL label. But users are still demanding more details and more precise modifications. Just as the offer for licenses is growing, the demand of users is also going up. When looking at the licenses used on the Sourceforge repository, we need to read the mailing-list associated with each license to know when they were created (Fig. 4). The demand for new contracts is growing steadily as the number of users is growing. But these contracts are not shared equally amongst projects. For example, the Sourceforge repository is mainly dominated by the GPL and the LGPL(Fig. 5). Even when there is more and more new licenses every year, they only apply to a very small number of projects. New contracts do not answer the questions ruled by existing contracts. They are

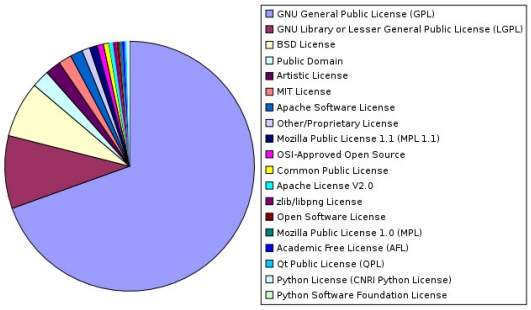


FIG. 5 – Repartition of the more popular FLOSS licenses on Sourceforge in 2005 : the new licenses created every year only account for a very small number of projects

Sourceforge in April 2005.

created by modifying old ones in order to rule more new questions. Logically, they concern less users, but the demand for new contracts is also growing up with the number of new users wanting to fit their new needs. Then, we argue that the multiplication of contracts creates a systematization that will be used to better understand the differences between these contracts, the underlying principles of this structure, its legal efficiency, its consistency and its economical interest. We will now address this question of the emergence of the domain of open source contracts.

2 Emergence of the FLOSS legal domain

Simultaneously to the multiplication of contracts, there is a process of systematization consisting in using new licenses to formulate terms to better formulate the legal problems at stake.

2.1 Systematisation

The terms used within the contracts of the domain allow to abduce a vocabulary for this specific domain. Then, it is possible to induce a normative grind stating the causal and logical constraints of the problem while using the same terms defined in the abduced vocabulary.

Application In this model, a FLOSS license are a singular legal solution to a general legal problem and their proliferation is associated to the need for adaptation to real situations. Thus, a legal problem will be described as the virtual abstraction of a real one. Its answer will be given by a set of questions where answers lead to the different clauses as a specific solution. To sum it up, there is an abstraction process consisting in finding the good set of questions which answers will produce the text of the licenses. In our example of FLOSS contracts, the right to access its Source Code can be seen as a core with every other rights granted by the licensee as optional and potentially modified. Users can also add some other obligations that will add up to the classical rights defining Open Source and Free Software. In the case of the CECILL contract adaptated from the GPL, the right to distribute the software has been divided in two : the right to distribute a modified version of the software and a non-modified one. The CECILL contract will then be a singular solution to a general problem related with the distribution of modified and non-modified versions of a software. This two terms can then become part of the FLOSS legal domain and they will be re-used in following contracts. For example, when people will adapt the CECILL to their own needs.

2.2 Modelizing a class of contracts

Practically, the modelisation of a class of contracts is done by listing every clauses used and by establishing the constraints between them. Each clause will be related to a real use, in the sense that at least one contract is using. We will use a computerized tool called Integre adapted to assist the abduction and induction for a group of lawyers(Fig. 6).

A model of FLOSS licenses As we explained, choosing a FLOSS license means to answer at least the first four different questions we mentionned. But as we can establish constraints between them, each new situation leads to more precise questions, to transformations of the clauses within the existing FLOSS licenses and to the creation of new ones. Asking a specific question means asking another subset of questions, and can exclude from asking another one. For example, accepting to distribute modified versions of a FLOSS Software will first mean to have accepted to let users mo-

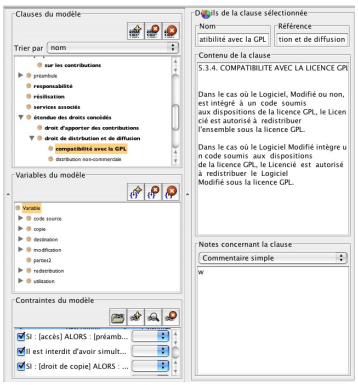


FIG. 6 – Inserting clauses of CECILL-based licenses within integre and detailing constraints between them : the user will hierarchically insert clauses in the first box, he will detail their values in the second one, the constraints in the third one. The text of the clauses are on the left.

dify the Source Code. We can then induce that there is a constraint between these two clauses. Such a work allow to adopt a causal reasoning on the given problem users are confronted to. Each subset of questions will allow us to abduce a doctrinal aspect, and provide the legal referee necessary to contractual freedom. Every question can be seen as a different dimension for solving the problem. Then, summarizing the clauses through an abduction process will allow users to summarize their problem and ask new questions. They will think virtually through this process, and confront its results with reality in their domain : users will refer to the general theory of contracts and to the theory of FLOSS licenses to get answers to their case. Then once again, confronted with reality, these answers will lead them to ask for new questions. When writing new contracts, the goal of users is to be able to attein a better level of legal prediction in and to reduce transfer costs. It is important to describe which elements of theory are guiding users and how they become more precise with new subsets of questions allowing to create adapted new contracts.

Abduction and induction The abduction process produces an ontology, which means an organization of the term of the domain that can be hierarchical or intricated. Their definition and their coordination through induced constraints will guide their application. It can bring to a better judicial safety and it initiates a co-evolutive process between legal solutions and legal problems. The induction process creates the framework for the resolution of the legal problem. Its principle is based on the research for contradiction followed by its acceptation or its refusal. Allowing a contradiction is equal to create a new contract and make evolve the contractual framework (Fig. 7). Thus, every new contract provides new clauses or modified ones and new questions related to new problems.

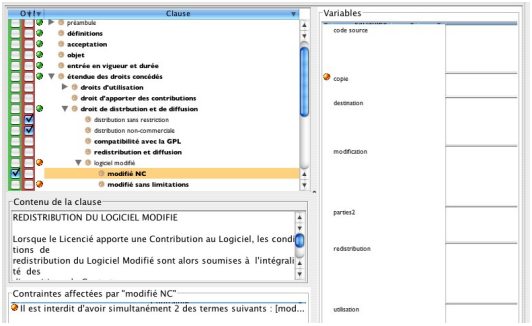


FIG. 7 – Creating a license in Integre by choosing the abduced clauses and following the induced constraints between them : the user decide what clauses to accept or refuse, the values are expressed on the left, the text of the contrat is on the second box and the constraints on the third one

Example In our example, when looking at the mailing-lists of FLOSS projects, most of licenses were created by adding new clauses, or modifying one or several clauses to a previous license. It is a way to go further in the description of the dimensions we previously mentioned. For example, the article 6 of the GPL mandates that any addition to an existing GPL software must itself be released under GPL. But many contributors needed to dynamically link a GPL software library with a classical proprietary software and they felt the GPL was not the correct answer to their problem because of its article 6 since it would mean that they should release their proprietary software under GPL and reveal its source code. Then, the LGPL license was created after the GPL license by modifying its clauses in order to allow people to use external Free Software libraries in their program without being imposed to release the source code of their own software. There were a need to ask new questions to precise the right to use as some people found the solution proposed by the GPL to be not good enough. Thus, users build contracts by putting their clauses together when they answer existing questions, they create new contracts by asking new questions leading to new clauses or modifying the existing ones. For example, people who do not want to allow too many modifications to their software will choose a QPL license which is a modified GPL license stating that additions to the source code must only be submitted as patches and can never be released as binaries. In that case, users felt necessary to ask new questions about the right to modify as the solution brought by the GPL was erroneous in some cases. Also the French license called CECILL is a license drafted by the CNRS and INRIA after the GPL in order to be written in French and more consistent with regards to the French law. A given FLOSS license can be described as a global solution to a given problem, integrating the different sub-solutions to all of it sub-problems. For example, the Creative Commons licenses are exactly built on this model : they provide a webpage with a given subset of questions where users can determine which contract will best correspond to their needs.

3 Conclusion

Evolutionnary model The history of FLOSS licenses tends to show that the number of licenses is always growing. Contrary to the common belief, there is not a need for a unified set of rules, but a need for more clauses describing a greater number of real situations. The more question‘s are answered by these, the more precisely their combination can describe the reality. In other words, FLOSS licenses answer the judicial problems of software distribution by offering an increasing number of news clauses giving birth to a greater number of licenses. All these licenses are not only law in text but they are also law in action, creating a reasoning field that help to decide how to distribute one's software. It does not become more complicated, but the comprehension of the complexity of reality increase as new questions arise and get new answers.

Producing an ideal The dynamics of FLOSS licenses are not about creation but about adaptation. Adaptation to particular situations under the pressure of objectives and constraints weighing on players. And adaptation to their own history under the need to recreate a consistent general theory of law and to give the right answers for OSS questions. But then, it is interesting to see that these new questions always grow along the four dimensions we mentioned, making them more and more visible. The evolution of the system is following an ideal.The description of this

ideal by the system is actual. Users have real needs, they answer them by interpreting an ideal and asking actual questions. Under classical models, these different solutions would be united under a legal theory giving birth to a set of rules like Hamurabi's 282 rules. But following this model, the development of new contracts will be much more easily done when making it easier for users to ask new questions more rapidly, bringing more than 250 and more contracts. In this case, the equivalent of a legal theory would be the product of the relationship between the actual state of the system, the needs of users put in real situations and the ideal solution. If a contract is the product of a number clauses obtained after answering a number of questions, a legal theory of contracts is the product of a number of contracts obtained after a number of real situations.'

Impact on legal theory From a theoretical point of view, the mechanisms of abduction and induction at work here allow contracts to regulate the social game of players. In fact, the number of different contracts is not exploding with the number of its users, and this bottum-up evolutionnary mechanism appears to be of utermost importance for the creation of law. This could be extended : as FLOSS licenses refer and detail a few chosen principles, judges are today referring to human right principles rather than Civil Code articles, leading to new legal solutions on such issues as homosexuality or transsexualism, but also on succession rules or commercial warranties.

Further directions

As a conclusion, this paper considered legal change in FLOSS contracts, focusing on conditions that may determine consolidation or corrosion of their clauses. It would be interesting to study if the stability of one clause is affected by the number of users choosing it, the flow of recent modifications, the institutional threshold of the need for stability and the weights attached to established provisions and recent new situations[11]. It should be needed to highlight the role of exogenous shocks, as different dynamic paths may be produced by a similar shock under different contracts. Plus, FLOSS licenses require varying degrees of consistency and this variable, interacting with other exogenous variable, could generates different patterns of evolution. Also, the present model is considered within a unitary judicial system but there are situations in which contrats have an intrajurisdictional effect, rather than an inter-jurisdictional effect across different judicial branches. Finally, this analysis could be scaled further to include the impact of jurisprudence and other forms of regulations in Law.

Acknowledgment

This work has been partially supported by the Normind Company.

[11] V. Fon, F. Parisi, Judicial Precedents in Civil Law Systems : A Dynamic Analysis, George Mason University School of Law, Law and Economic Working Paper Series 2004

The problem of design in complexity research

Theodore Zamenopoulos and Katerina Alexiou
CASA and Bartlett School of Graduate Studies, UCL

Keywords: design, design systems, complexity research, organization, mathematical structures

Abstract
This paper discusses the contribution of design research problems and abstractions in the formation of a complexity research agenda. Design is a capacity associated with systems that are often characterized as complex - but does design imply a general capacity and class of research problem that is inextricably linked with the complexity of a system? This is a rather uncommon enquiry even though the relation between design and complexity has been explored under two themes: the complexity of design, that is the identification and measurement of the complexity of design objects, processes and problems; and the design of complexity, that is the construction and management of complex (artificial) systems. However, these are mainly applications of complexity concepts and measurements in design research and practice rather an investigation of the meaning of complexity based on the design capacity of systems. The purpose of this paper is to discuss the mathematical basis and problems of a design theory of complexity and demonstrate the uniqueness of design as a distinct problem in the context of a complexity research agenda.

Complexity research is too versatile to be described by an indisputable research agenda. However there are a number of traditions and problems that can characterize the meaning of complexity and the scope of complexity research. For instance, complexity has been identified with the combinatorial capacity of systems; scaling; the capacity to exhibit certain types of critical behaviour or attracted to critical states; the evolutionary capacity of systems, associated with problems of cooperation, competition and reduction of variety; and finally the organizational capacity of systems. For developing the argument of the paper the focus will be on the latter aspect of complexity research.

In this investigation, complexity is associated with the capacity of a system to exhibit a certain type of organization. The motivation for this is that certain types of organized systems, such as the brain, organisms or societies, can exhibit complex functions such as intelligence, life, or governance. The main question is two-fold: first, to identify the organizational conditions that enable such capacities to emerge; and second, to identify the capacity that explains how these organizational conditions are produced and maintained. The production and maintenance of the organizational conditions of a system generally alludes to the capacity of the system to change the structure, behaviour or function of its environment (or its perception of this environment) and through this change to transform itself. Typical examples of such capacities are distinction and intentionality, autonomy and control, creativity and learning, anticipation and - as we claim in this paper - design. Design in particular alludes to the capacity of producing organizational changes in the environment of a system that increase the complexity of the system relative to its environment. Irrespective of where one chooses to embody or allocate this capacity, the special characteristic of design problems -in distinction to other abstractions such as machine, control or evolution- is that the complementary nature of the relation between system-environment is not a given but it is the problem itself.

Goguen and Varela ([1], [2], [3]) have explicitly associated the complementary relation between system and environment with the category-theoretic concept of adjunction. The same adjoint

relation can be implicitly found between allonomy and control, machine and language but also to other organizational concepts such as coordination and subordination, or scaling and variety ([4]). Changes in the system and its environment always preserve the system-environment complementarity. Now the question is whether it is possible to perceive changes where the adjoint relation between levels of organization is not preserved. This class of problems will be generally called design problems. The idea would be to 'push' the system 'far for the adjoint relation' and explore abstractions that underline these capacities and organizations.

In order to fix these ideas, the complementary relation between sets and monoids is analyzed. In particular, a monoid structure is represented in a type of one dimensional cellular automata space. The objects of the structure are realized by natural numbers, whereas the morphisms are realized by mappings between natural numbers. By enabling the state of a cell to play both the role of an object and a morphism between neighbour cells, the operation of composition and coupling is introduced. The paper demonstrates the formation of cells that work as boundaries between inner and outer areas by means of composition and coupling that are not structural preserving. Based on this model, a definition of the design capacity of systems is discussed.

To sum up, the paper identifies the concept of design as a distinct research question in complexity research. It explicitly links general problems of complexity with the specific concept of design and discusses the unique characteristics of design problems. It is hoped that this can be of benefit for both complexity research and design science.

References

[1] Goguen J: Minimal realization of machines in closed categories. Bull Am Math Soc 1972;78:777.
[2] Goguen J, Varela F: Systems and distinctions: duality and complementarity. Int J Gen Systems 1978;5:31-43.
[3] Varela F: Principles of Biological Autonomy. New York, Elsevier North Holland, 1979.
[4] Bar-Yam Y: Multiscale variety in complex systems. Complexity 2004;9:37-45

Corresponding author’s address: CASA, 1-19 Torrington Place, London WC1E 7HB
Tel: + 44 20 7679 5901, Fax: + 44 20 7813 2843, e-mail: t.zamenopoulos@ucl.ac.uk

Bio Inspired Methods

On the Complexity of Physical Problems and a Swarm Algorithm for k-Clique Search in Physical Graphs

Yaniv Altshuler[1], Arie Matsliah[1], and Ariel Felner[2]

[1] Computer Science Department, Technion, Haifa 32000 Israel
{yanival, ariem}@cs.technion.ac.il
[2] Ben-Gurion University, P.O.B. 653 Beer-Sheva 84105, Israel
felner@bgu.ac.il

Abstract. As the complexity of systems increases, so does the need of examining the nature of *complexity* itself. This work discusses the domain of *physical swarm problems*, in which a swarm of mobile agents is employed for solving physical graph problems (where a certain amount of *travel effort* in required for every movement along the graph's edges). A new kind of complexity scheme, suitable for this domain, is discussed by examining a central problem of this domain — the *physical k-clique* problem. In this problem, a swarm comprising of mobile agents travels along the vertices of a physical graph G, searching for a *clique* of size k. Thus, the complexity of the problem is measured in *travel efforts* (instead of in computation resources). In order to share information between the agents, two communication models are discussed — a complete knowledge sharing (referred to as *centralized shared memory*) and a *distributed shared memory* model, where the mobile agents can store and extract information using the graph's vertices. The work presents a search algorithm for the agents, and discusses its performance under each communication model. The major contribution of this work is demonstrating the strength of the distributed shared memory model. Although this model is much easier to implement and maintain, is highly fault tolerant and has high scalability, the quality of the results it produces is very high, compared to the strongest model of complete knowledge sharing.

Keywords — Physical Graphs, K-Clique, Pattern Matching, Swarm Algorithms, Swarm Intelligence, Distributed Knowledge Sharing

1 Introduction

In recent years, much work has been done in the domain of *physical problems* (or problems in *physical graphs*). Such problems concerns a goal which should be achieved, by one or more mobile agents, which travel along the physical environment. When the group of agents comprises several agents it is also called a *swarm*. The domain of *swarm algorithms* and *swarm intelligence* is also a rapidly growing research field, in which the complexity aspect is of great interest (since usually swarms are assume to comprise very simple agents, with limited resources and capabilities).

Traveling within a physical environment however, is very different than usual information access model assumed in orthodox graph theory, since every movement along

an edge of a graph requires a certain amount of *travel effort*. While the travel effort can represents certain amount of time, of fuel, it is immaterial to this model.

However, since movements along the graph no longer assumed to be performed in zero effort, a basic change in the conceptual way in which performance of algorithms is measured, and in which hardness of problem is being formulated, is needed. Hitherto, the complexity of mathematical problems and computer algorithms was mainly measured using the well studied canonical complexity classes such as *P*, *NP*, *P-SPACE*, and others (see [39] for an overview). However, when discussing the domain of physical problems, this scheme should be replaced by some other paradigm, more suitable for the definitions and properties of this domain.

In this work, a new kind of complexity scheme, designed for physical problems, is discussed. Its consequences on the construction of swarm algorithms for physical problems is discussed, and several mechanisms for enhancing the performance of such algorithms, yielding from the complexity scheme are described. This is done by examining a central problem in this domain — the *physical k-clique* problem. In this problem, a swarm of mobile agents travels through a graph G, where its goal is to find a clique of size k (a set of k vertices in which every pair of vertices is connected by an edge in G).

While examining this problem, an algorithm which can be used by a mobile agent is developed, called **PCF** (*Physical Clique Finding*), and its performance examined, as well as several variants of it.

The algorithm is than shown to be an efficient swarm algorithm as well, meaning — be applied successfully in a multi agent environment, where every agent works according to this algorithm. While showing the last, we discusses the calibration of certain parameters of the algorithm, in order to reach optimal performance.

Note that the **PCF** algorithm requires certain cooperation between the agents utilizing it. This cooperation can be achieved by assuming several implicit communication models, using some kind of a shared memory. As an upper bound for the algorithm's performance, we first examine its behavior under the assumption of a *centralized shared memory*, simulating a *completed knowledge sharing* model. As a second step, we examines the algorithm's performance while assuming an allegedly weaker memory model — a *distributed shared memory* model. It is our intention to show that although a system which uses this model is much simpler and easy to implement, the performance which can be achieved over it are not significantly lower than the ones achieved using the superior centralized shared memory model. This is shown using the experimental results in section 7, which demonstrate that while using the proposed algorithm in the distributed shared memory model, the results which are obtained are only slightly inferior than those achieved using a centralized shared memory.

This is surprising since the distributed shared memory model is much simpler to implement and maintain than the centralized shared memory. Moreover, its complexity is far smaller than the stronger model, since its highly scalable — as the graph become larger, small memory units are added to the new vertices. When more agents are added, the communication complexity remains the same, whereas when assuming the centralized shared memory, the communication complexity is $O(k \cdot |V|^2)$.

This fact is perhaps the most meaningful result of this work — although a similar notion was hinted in a few recent works, demonstrating it in a major problem such

as the k-clique problem greatly strengthen this notion. This work will be followed by an additional work, currently under preparation, which presents an extended research concerning ideas and suggestions which appear in this work in part.

1.1 Physical Graphs

A *physical graph* denotes a graph $G(V, E)$ in which information regarding its vertices and edges is extracted using *I/O heads*, or *mobile agents*, instead of the "random access extraction" which is usually assumed in graph theory. These agents can physically move between the vertices of V along the edges of E, according to a predefined, or an on-line algorithm or algorithm.

Moving along an edge e, however, require a certain *travel effort* (which might be a constant time, or alternatively, consumes a constant amount of *fuel*). Thus, the complexity of algorithms which work on physical graphs is measured by the total travel efforts required, which equals the number of edges traveled by the agents. We assume that each edge requires exactly one unit of travel effort.

Physical graphs are conveniently used in order to represent many "real world problems", in which the most efficient algorithm is not necessarily the one whose computational complexity is the minimal, but rather one in which the agents travel along the minimal number edges. For example, the *Virtual Path Layout* problem, concerning the finding of a virtual graph of a given diameter, and its embedding in the physical graph such that the maximum load is minimized (see [38] and [40]).

Problems in physical graphs are thus variants of "regular" graph problems, such as finding a *k-clique* in a graph (description and algorithms can be found in [19]), graph and subgraph isomorphism (description and algorithms can be found in [25, 29, 11, 3, 7]), exploration problems (solved for example by *Breadth First Search* (BFS) and *Depth First Search* (DFS) algorithms [23]), etc., whose input graph is a physical graph. Thus, the complexity of these problems is measured as the number of edges an agent (or agents) solving the problem will travel along.

There is a special type of graph problems which can also be ideally described as physical graph problems. Such problems are those in which a *probabilistic*, or *real time* algorithm is used to return a solution which is not necessarily optimal. While a probabilistic algorithm returns a solution which is correct in a probability of $(1 - \epsilon)$ (for as small ϵ as we require), a real time algorithm can be asked at any stage to return a solution, whereas the quality of the solutions provided by the algorithm improves as time passes. Using such probabilistic or real time algorithms, the computational complexity of many problems can often be reduced from exponential to polynomial (albeit we are not guaranteed of finding the optimal solution). Such algorithms can be found for example in [8, 28, 31] (graph isomorphism) and [32, 21, 16] (distributed search algorithms such as *RTA**, *PBA**, *WS_PBA**, *SSC_PBA** and *BSA**). The physical variants of such problems can be thought of as a swarm of mobile agents, traveling the graph and collecting new information during this process. As time advances, more information is gathered by the agents, causing the quality of the solutions provided by the agents to improve.

Notice that while an algorithm which assumes a random access data extraction (from now on be referred to as *random access algorithm*) may read and write to the

vertices of G at any order, an algorithm which assumes a physical data extraction (referred to as a *physical algorithm*) must take into account the distance between two sequential operations. The reason for this is that the use of a random access algorithm is performed using a processing unit and random access memory, whereas the use of a physical algorithm is actually done in the physical environment (or a simulated physical environment, which maintain the information access paradigm). Thus, a random access algorithm can access any vertex of the graph in $O(1)$, while a physical algorithm is confined to the distances imposed by the physical metric.

For example, for $u, v \in V$, let us assume that the distance between v and u in G is 5. Then if after a 'read' request from u, the algorithm orders a 'write' request to v, this process will take at least 5 time steps, and will consume at least 5 fuel units. Furthermore, depending on the model assumed for the mobile agents knowledge base, this operation may take even longer, if, for example, the agents are not familiar with the shortest path from u to v, but rather know of a much longer path connecting the two.

As can easily be seen from the previous example, while designing physical algorithms, one must take into account an entire set of considerations which are often disregard, while designing random access algorithms. As a result, algorithms which are designed in order to solve physical problems, may often bear only a few resemblance to ones designed for the corresponding random access problems.

1.2 Communication Models

There are many models for communication in multi agent systems. The most trivial model is of "*complete knowledge sharing*", where any new discovery of an agent is immediately shared with all the other agents. Other models restrict the level of communication. Some models allow broadcasting or message exchanging between agents but restrict the amount of data that can be exchanged in each message or restrict the frequency or the number of messages that are allowed. Many times, a cost is associated with each message and the task is to solve the given problem while trying to minimize the cost of the messages involved.

In nature, ants and other insects communicate and coordinate by leaving trails of odor on the ground. The information placed on the ground is called a *pheromone*. Studies on ants (e.g. [9, 10]) show that the pheromone-based search strategies used by ants in foraging for food in unknown terrains tend to be very efficient. It is believed that ants build a network of information with vertices represented by points of encounter between ants and the information is either passed between ants at a vertex with a physical encounter with other ants or via pheromone traces that are left on the ground.

Inspired by nature, a famous and interesting model for communicating in multi-agent systems is that of ant-walk (e.g. [12, 13, 26]). In this model, information is spread to other agents via *pheromones*, i.e., small amounts of data that are written by an agent at various places in the environment (e.g. edges or vertices in the graph) and can be later used or modified by other agents visiting that vertex.

In our model we study two communication models. The first model assumes a *centralized shared memory*, which can be directly accessed by each mobile agent at any time. Notice that this model is equivalent to the *complete knowledge sharing* mentioned

above, which is the most trivial and "strongest" communication model. However, implementing such a model in reality often requires the use of a broadcast capability by the agents, or a fast point to point communication. Thus, there is a strong motivation of presenting other communication model which achieves similar performances to the first.

The second communication model presented in this work is the *distributed shared memory* model. In this model the information network is a graph, and the role of a pheromone is taken by a memory area on each vertex, which each mobile agent can read and modify. The union of these memory areas forms a *distributed shared memory*. Whenever an agent reaches a vertex which contains such memory area, it compares the information located in this memory to the information stored in its database. If the vertex's memory contains new information, the agent incorporates it into its own database, while if the agent had gathered information which is not contained in the vertex's memory, the agent deposits it into the vertex's memory. This process is called a "*data merge*". In this work we assume that all the vertices contain such a memory, although other schemes should be examined as well — requiring only a portion of the vertices to contain such memory module will simplify the implementation of such system and will lower its costs, however the effect of it on the algorithm's performance is yet to be explored.

The distributed shares memory model is especially suitable for large networks (e.g. Internet), in which the vertices are implemented by powerful computers and the edges often have narrow bandwidth and are heavily loaded. This paradigm suggests a distributed group of one or more lightweight autonomous software agents that traverses the network in a completely de-centralized and parallelized manner. Data is spread among the mobile agents via this distributed shared memory. Note that this approach resembles the use of "*large pheromones*", as appears in [42, 41] and elsewhere. However, we believe that the term "distributed shared memory" better describes the nature of the mechanism presented above.

Several works have already been done, considering mobile agents which uses variants of the distributed shared memory model. Such works have shown that such agents are able to cooperate and achieve goals like covering a faulty graph [13], finding an Euler cycle in a graph [26] and solving various combinatorial optimization problems [27].

Following is a communication paradigm between mobile agents assuming a distributed shared memory model in general, and in exploring unknown environments in particular. Each agent has a partial knowledge of the entire environment. It maintains a database with a partial graph that is known to it. Similarly, each vertex holds a database with a partial graph that is 'known' to it, i.e., knowledge that was written to it by the agents. Whenever an agent reaches a vertex, it merges the data known to it with the data that is written in that vertex. The agent then writes the combined data in that vertex and updates its own database according to the new data that was obtained (a *data-merge* operation).

2 Problem

Pattern matching in graphs is the following problem: Given a graph G on n vertices and a graph H on h vertices, find whether G has an induced subgraph isomorphic to H. Many applications of pattern matching can be found in theory and practice of computer science, see e.g. [44] for more details. It is well known that this problem is computationally hard whenever H (and also h) is not constant, but rather a part of the input (a subgraph isomorphism problem), while it has a trivial solution of polynomial complexity if H is a constant graph. Note that the subgraph isomorphism problem is reducible to the *Max-Clique* problem.

This work considers dense physical graphs (graphs which contain many subgraphs isomorphic to the desired pattern) while the goal is to find one of them within as minimal moves on the graph as possible. A graph G on n vertices is called *dense* if the average degree of a vertex in G is $\Omega(n^2)$. Alternatively, we can slightly weaken this condition by requiring that the number of edges in G is $\Omega(n)$, as long as the existence of a large number of the requested patterns can be ensured (see section 7 for more details). The vertices of G are indexed from 1 to n, and the edges of G are indexed from 1 to m, where $m \leq \binom{n}{2}$.

We assume that whenever a vertex $v \in V(G)$ is visited, all edges incident with v are revealed (i.e. their indices are revealed), and naturally v's index is also revealed. Hence, if an edge $e = (u, v)$ exists in G, and some agent visited u and v (in any order), then he can deduce that v is a neighbor of u, even if the agent did not travel on e. If there is a communication between the agents, it is enough that one of the agents visited u and some other agent visited v for this information to be deduced.

One of the artificial examples of a similar model might be an unknown terrain with indexed cities and roads, where the roads are signed with their indices (say with road signs), but their end point indices are not mentioned. However, we do not assume that the graph is planar (i.e. there might be roads, bridges and tunnels, crossing each other in any way).

Similar to ordinary navigation tasks (for example [20–22, 24]), the purpose of each agent employing **PCF** is to reach the "goal vertex" as soon as possible. However, since the goal of the agents is to detect a *k-clique*, the goal vertex is in fact the vertex which when discovered at time t, will complete a *k-clique* in the agent's knowledge base. Thus, there is no specific goal vertex, but rather several *floating goal vertices*, whose identity depends on the following elements :

- The structure of G (namely, V and E).
- The information the agent had collected thus far.
- The information sharing model of the problem (be it a distributed memory, centralized memory, etc').

Note also that this problem is not an ordinary exploration problem (see [4]), where the entire graph should be explored in order for it to be mapped out. Once a requested *k-clique* is found by one of the agents, the problem is terminated successfully. This often occurs while only a small portion of the graph's vertices have been visited.

Another distinction should be made between the problem that is presented in this work and those examined in the field of multi agents routing (see [5, 6]). While multi

agents routing mainly deals with the problem of finding paths in dynamic networks where the structure of the environment is dynamically changing due to load balancing and congestion, the *physical k-clique* problem considers a stable physical environment (somewhat similar to the work of [2] and [1]).

3 Motivation

It is our belief that work on swarm algorithms for physical graph problems is strongly required since physical graphs and networks, which take an essential role in everyday's life, are becoming more and more complex. Thus new techniques for such problems must be composed. Several examples for such networks are the world wide web [30, 17, 18], the physical and logical infrastructure of the internet [17], power grids, electronic circuits [17] — all of which are complex physical environments.

The *physical k-clique* problem has several "real world" applications. For example, tracking the connectivity of a computer of telephone network, which can be utilized in order to improve routing schemes within this network. Another example may be a distributed mechanism which search many databases containing transactions and email correspondences, in order to find groups of people who are maintaining a tight communication between them. This information may be used by intelligence units, tracking potential terrorists.

The reason the *physical k-clique* problem was selected for this research is that the *k-clique* problem, on which the *physical k-clique* problem is based, is known to be a significantly hard problem. While most of known NP-complete problems can be approximated quite well in polynomial (sometimes linear) time (as shown in [45]), this is not the case for the *k-clique* problem. An explanation of why there are no "semi-efficient" algorithms for the *k-clique* problem (and that the best solution is exhaustive search) and why the *k-clique* problem can not be approximated in *any* way, unless $P = NP$, can be found in [46]. Additional details regarding NP-Complete problems can be found in [47].

Since the *physical k-Clique* problem is in fact an instance of the *physical pattern matching* problem, solving it can serve as a first step towards a general pattern matching swarm algorithm. In future works we intend to show that the algorithm presented in this work can be applied to any pattern with few modifications.

4 Basic Physical Clique Finding Algorithm

This sections presents the basic swarm algorithm for finding k-cliques in a physical graph. The algorithm described here assumes that the implicit communication used is that of a centralized shared memory (which simulated complete knowledge sharing), meaning — all the agents share the same knowledge base and can perfectly coordinate their operations.

Notice that this model is equivalent to a centralized algorithm in which all the agents are controlled by a "leader" or a "queen", thus it is the strongest model possible, with respects to knowledge sharing and operation coordination.

4.1 Search Algorithm - Definitions and Requirements

Let $r(t) = (\tau_1(t), \tau_2(t), \ldots, \tau_n(t))$ denote the locations of the n agents at time t. The requested search algorithm is therefore a movement rule f such that for every agent i :

$$f\big(\tau_i(t), N(\tau_i(t)), \mathcal{M}_i(t)\big) \in N\big(\tau_i(t)\big)$$

where for a vertex v, $N(v)$ denotes the neighbors of v, e.g. :

$$N(v) \triangleq \{u \in V \mid (v,u) \in E\}$$

$\mathcal{M}_i$ is the memory for agent i, containing information gathered by it through movement along G and by reading information from the shared memory.

The requested rule f should meet the following goals :

- **Finding a k-Clique** : ensuring that :

$$\exists t_{success} \mid k\!-\!clique \in \mathcal{M}_i(t_{success})$$

  such that this $t_{success}$ is minimal.
- **Agreement on Completion** : within a finite time after the completion of the mission, all the mobile agents must be aware that the mission was completed, and come to a halt.
- **Efficiency** : both in time and space.

The requested rule should also be *fault tolerant*, meaning that even if some of the agents malfunction, the remaining ones will eventually find the target clique, albeit slower.

4.2 The PCF (Physical Clique Finding) Algorithm

For solving the *Physical k-Clique* problem we suggest the **PCF** search algorithm. This algorithm can be described as follows. Each agent holds a knowledge base which contains some information regarding G. Every time an agent enters a vertex v, it performs a data merge process, in which the knowledge base of the agent updates and is updated by the central knowledge base.

The main idea in the search algorithm is exploring the graph in directions that have the highest potential of discovering the desired pattern. In our case, considering only cliques simplifies the arguments because of their perfect symmetry, but it still emphasizes the main ideas of the general case — a generic search algorithm for any given pattern.

This is done by sorting the potential sets of vertices according to the largest clique which is contained in the set. Within the same size of largest clique, the sets are sorted according to the total number of unexplored edges which touch the vertices of the set (unexplored edges are edges $e(v,u)$ whereas at least one the identities of v and u is yet unknown) . As large the number of such edges is, the more likely it is for the set to be expandable to a *k-clique*. In addition, if a *k-clique* was found, the sort criteria will place it on the top of the list, and hence the agents will immediately recognize it. To maintain

the lists compact, the algorithm focuses on certain sets (the top most in the sorted list) and either eliminates them or discovers that they are expandable to a *k-clique*.

Our algorithm uses a *job list*, where each job is associated with a set of vertices that are potentially expandable to a *k-cliques*. The jobs are sorted primarily by the size of their corresponding sets, and secondarily by the number of unexplored edges incident to one of the vertices in the set. In addition to the set of vertices, a job holds an instruction to the agent that takes it. The instruction is of the form "travel on edge e from vertex s", where $s \in V$ is the source vertex, and the agent might need to go to s before he performs the instruction of the job.

Generally, when looking for a pattern H of size h in graph G, every set of $m, m < h$ visited vertices in G that might be completed to a subgraph of G isomorphic to H (i.e. there are enough unexplored edges for every vertex) forms a potential sub-H of size m. While considering cliques as the patterns, we only need to verify that the m vertices of the set form a clique and that every vertex in the set has at least $h-(m+1)$ unexplored edges (which is the minimal requirement in order for this set to be expendable to an $h-clique$.

The list of jobs is updated (if needed) after every move of the agents.

If there are α available agents in a turn, the algorithm assigns the first α jobs in the sorted list to the agents, where the assignments are made in a way that minimizes the total travel distance in L_∞ norm. This is done in order to minimize the travel efforts of the agents. This is an example of a difference between physical problems and "regular" problems — whereas in conventional complexity scheme all the ways of dispersing the jobs among the agents are identical, in the complexity scheme of physical problems we would like to assign jobs to the nearest agents.

Once an agent reaches its goal (the closest vertex of the job assigned to the agent), it writes the new discovered information in the adjacency matrix and becomes available. In the beginning of each turn, if the first job in the list is associated to a list of k vertices the algorithm declares that a clique is found and terminates.

Notice that all the agents use a common job list. In addition, the agents are assumed to broadcast every new information they encounter.

We assume that all distances (i.e. graph edges) are equal to one unit, and that the agents have sufficient memory and computational power. We also assume that that there are many cliques of the required size in G (otherwise, if the number of cliques is too small (e.g. $o\binom{n}{k}$), the optimal algorithm for discovering them would be the exhaustive exploration of G).

The algorithm, used by each agent i appears in figure 1.

4.3 Correctness

The main observation, on which our algorithm is based, is that none of the agents travel on same edge more than once. Hence, even in the worst case scenario, the whole graph will be explored within no more than $|E|$ moves (which is asymptotically similar to the standard *DFS* and *BFS* algorithms). Therefore, our algorithm terminates after at most $O(|E|)$ moves. Since all *k-cliques* are in the top of the job list in some stage, the algorithm described above eventually recognizes them.

Algorithm **PCF** :
While the first job in the common job list is not associated to a set of k vertices
    Assign top job in jobs list to the agent;
    Perform the job;
    Update the agent's data structures;
    Update the agent's jobs list (or the common list);
    Update the list of available agents;
    If (all vertices explored) then
        STOP;
End **PCF**;

Procedure **CREATE DATA STRUCTURES**(i) :
Create an $n \times n$ matrix with entries of type integer;
/ Represents edge indices according to vertex adjacency of G, as discovered by the agents. */*
Create a sorted list of jobs;
/*
Each job is associated with one potential clique. The job list is sorted according to the following criteria :
(a) Size of the potential sub clique
(b) Number of still unexplored edges
(c) The distance of the job from the agents
The distance is computed according to L_∞ norm, that is the min—max *on the travel distance.*
*/
End **CREATE DATA STRUCTURES**;

Procedure **INITIALIZE**$()$:
Initialize the agents' jobs lists;
/*
Note that the common job list may contain large potential sub-cliques, immediately upon initialization.
*/
CREATE DATA STRUCTURES(i);
Choose random starting points on G for the agents;
Place the agents in these starting points;
Start the agents according to the **PCF** algorithm;
End **INITIALIZE**;

Fig. 1. The **PCF** search algorithm for the centralized shared memory model.

5 Physical Clique Finding Algorithm — Distributed Shared Memory

This sections presents the variant of the basic swarm algorithm for finding k-cliques in a physical graph, while assuming the distributed shared memory model.

Notice that the algorithm is *fault tolerant*, meaning that even if some of memory units within the graph's vertices will malfunction, the swarm will still be able to function and complete its mission with only a slight decrease of efficiency.

The general principles of the algorithm are identical to those appear in section 4.2 although the following features were added.

Every time an agent enters a vertex v, the data merge process is performed between the knowledge bases of the agents and this stored within v.

Notice that while in the centralized shared memory model, all the agents use a common job list, in the distributed shared memory model, each agent is equipped with its own list.

Since the agents are allowed to leave data in the vertices of the graph — we assume without loss of generality that the information stored will include all their knowledge base at the time of arrival to this vertex. This is comparable to the assumption made in the centralized shared memory model that the agents broadcast every new information they encounter.

The correctness of the algorithm is derived from the correctness of the centralized **PCF** algorithm, as appears in section 4.3.

The algorithm, used by each agent i appears in figure 2.

5.1 Guaranteeing Operation Diversity

When introducing to a physical graph more than one agent, using the **PCF** algorithm, and assuming a distributed shared memory, one may discover that several agents may travel together, causing the performance of the system to degrade. This problem is caused since all agents follow the same deterministic algorithm, and are initially located at the same vertex. Thus, all agents face the same initial information regarding the graph. As a result, they will necessarily make the same decisions and follow the same path. When m out of n agents visit the same vertices at the same time, for all practical purposes, the performance achieved are equivalent to those achieved when using only $(n - m)$ agents.

In order to avoid this problem, several techniques may be used. First, we may use a disperser protocol, which will choose random values for the first m moves of the agents, thus guaranteeing (with a high probability) that in time step m, no two agents will be located in the same vertex. Thus, the agents will now be able to start activating the **PCF** algorithm.

Another method which could be used is "breaking the symmetry" of agents entering the same vertex. At any time during the operation of the agents, two or more agents may enter the same vertex at the same time. In order to prevent such agents from "uniting" and start traveling together, a mechanism which will force an asymmetry between the agents must be implemented. By using such a mechanism, the agents could be able

Algorithm **PCF** :
While the first job in the agent's list is not associated to a set of k vertices
    Assign top job in jobs list to the agent;
    Perform the job;
    Update the agent's data structures;
    Update the agent's jobs list;
    Update the agent's availability status;
    If (all vertices explored) then
        STOP;
End **PCF**;

Procedure **CREATE DATA STRUCTURES**(i) :
Create an $n \times n$ matrix with entries of type integer;
/ Represents edge indices according to vertex adjacency of G, as discovered by the agents. */*
Create a sorted list of jobs;
*/**
Each job is associated with one potential clique. The jobs list is sorted according to the following criteria :
(a) Size of the potential sub clique
(b) Number of still unexplored edges
(c) The distance of the job from the agent
/
End **CREATE DATA STRUCTURES**;

Procedure **INITIALIZE**$()$:
Initialize the agents' jobs lists;
/
Note that each list will be initialized to contained at most a potential singleton (the vertex the agent is currently located in).
/
CREATE DATA STRUCTURES(i);
Choose random starting points on G for the agents;
Place the agents in these starting points;
Start the agents according to the **PCF** algorithm;
End **INITIALIZE**;

Fig. 2. The **PCF** search algorithm for the distributed shared memory model.

to generate a local order of the agents which are located in the same vertex, and use this sorting in order to ensure that the problem above will not occur. Since the **PCF** algorithm uses "job queues", and since when two agents are entering the same vertex, their knowledge base, and thus also their job queues, must be identical. By using the local order each agent can select the i-th job in its queue (rather than the top job), when i denote its place in the local order. Thus, we guarantee that each agent will select a different destination.

A sorting mechanism, as mentioned in the previous paragraph, can be implemented using the IDs of the graph's edges E. Since we assume each edge (v, u) has a distinct ID, known to agents arriving either from v or from u, this ID can be used in order to create the requested sorting. Let $e_{i,v}$ be the edge through which agent i had entered the vertex v. Thus, all the agents which had entered v examines $ID(e_{i,v})$ for every agents i in v. The IDs of the edges are than sorted (locally, by each agent), and through this sorting, the agents in v are sorted, when each agent i places itself according to the position of $ID(e_{i,v})$ in the ordered IDs of edges through which the agents had entered v. Since the system is not assumed to be synchronized, if two agents are entering v using the same edge e, one of them does this earlier, and thus it will be placed before the rest of the agents which entered v through e.

This problem, of course, does not appear when assuming a centralized shared memory. This is so since the centralized memory model also encapsulates the mechanism which enables each agent to "pop" a task from the global queue, with no chance of other agents selecting the same destination.

The problem of preventing simple agents from performing the same movements and thus decrease the swarm's total performance was discussed in various works which presented multi agents or swarms based systems. Since each problem assumes a different model, both for the environment in which the agents operate and for the capabilities and limitations of the agents, a different mechanism of avoiding this undesired phenomenon must be produced for each problem. For example, see [41, 35, 36].

6 PCF Algorithm — Improvements and Extensions

6.1 Towards a Generic Physical Patterns Finding Algorithm

The main difference between cliques and graph patterns in general is that cliques (or alternatively independent sets) have a perfect symmetry between their vertices. In other words, every pair of vertices in a clique has equivalent elements under some automorphism.

The notion of "potential" is also naturally extensible to the general case. For example, given a pattern H on h vertices and a subset C of c vertices from the graph G, we say that C is potentially expandable to H if :

1. $c < h$.
2. There exists an assignment π to the unknown pairs of vertices.
3. There exists a set C' of $(h - c)$ vertices such that the induced subgraph on $C \cup C'$ under the assignment π is isomorphic to H.

To scale the potential of a certain subset C, we compute the amount of possible assignments under which C is expandable to a set of vertices that induce a subgraph isomorphic to H.

6.2 Communication Agents

In order to further improve the performance of the **PCF** algorithm, working in the distributed shared memory model, the notion of *communication agents* was introduced. Communication agents are agents which instead of following the top task of the queue, decides to chose an arbitrary destination, and move towards it. Whenever such agents arrive to a new vertex, they perform the data merge process, similar to regular agents (referred to as *searching agents*). As discussed earlier, the goal of the searching agents is to gather new data regarding vertices whose potential of being a part of a large clique is high. The goal of the communication agents, however, is quite different — since searching agents only follow their tasks queue, there is a danger that the information collected by some of these agents will not become available to other searching agents, for which it might be extremely relevant. Thus, the random nature of the communication agents' movements is aimed for making sure information gathered by the searching agents proliferate along the graph.

Since the primary task of the communication agents is to spread data and thus perform as many data merge operations as possible, a communication agent may choose the vertex where the contribution of the data merge operation will be as significant as possible, as its destination. Furthermore, when such an agent arrives at a new unexplored vertex, this vertex is added to the general knowledge of the system (immediately in a centralized shared memory model, and partially in the distributed shared memory model).

There arises the question of the exact mechanism which will instruct an agent to become a communication agent. Notice that the simplest mechanism will chose a portion of the agents, upon initialization, and mark them as communication agents. However, there is also the possibility of dynamically allocating communication tasks to agents — whenever an agent will become available for a new job, it could be assign a communication job in some probability p, or a (regular) searching task in probability $(1-p)$. Notice that this allocation will assign one job only — when this agents will become available again, it could be assigned either communication or searching task.

Note that if the dynamic allocation method is used, there exist again a wide range of possibilities for deciding the probability p — should it remain constant throughout the entire mission, or should it change (increasing or decreasing the number of communication agents at any given time) ? Should it depend on the parameters of the environment (the size of the requested clique, the size of the graph, etc.) ?, should it depend on the parameters of the swarm (the number of agents, other enhance mechanisms which are used, etc.) ?

Let p_{comm} denote the probability for an agent to being assigned a communication mission. In order to answer these questions and find the optimal scheme in which the communication agents mechanism should be used, the following alternatives were implemented and tested :

- p_{comm} was first defined as a static number, upon the initialization of the system. This means that the at any given time, the number of agents which served as communication agents was roughly the same. When implementing this alternative, many values for p_{comm} were tested, in order to find the optimal value for this parameter.
- After simulating with static values for p_{comm}, as dynamic scheme was examined, in which the value of p_{comm} changes in time. The value of p_{comm} is defined according to the *simulated annealing* model (see for example [33]), where in each time step t the system has a temperature C_t which represents the "noise" of the system — the chance that an agent will perform a random movement. The temperature of the system changes according to a certain rule :

  $$p_{comm} = e^{-\frac{t}{C_0}}$$

  when C_0 is the initial temperature of the system.
- Similarly to the previous method, the value of p_{comm} is dynamic, where in p_{comm} itself is a function, defined as follows :

  $$p_{comm} = e^{-\frac{t \cdot f}{C_0}}$$

  while f is the utility of the last movement (calculated according to the increase or decrease of the potential of the top tasks in the tasks queue). Thus, when an agent had performed a movement which had improved its knowledge base (for example, by adding vertices to the largest clique known to the agent), it will have much lower probability of becoming a communication agent, than an agent whose last movement yielded no new valuable information.
- Alternative policies for the communication agents should be developed and examined. Since performing a random move does not necessarily guarantee that the agents will reach their goal — spreading the information among the swarm — a more complex *communication movement algorithm* is required. Such algorithm may take advantage of information regarding the positions of the agents themselves, and not only information concerning the graph. This will allow communication agents to ensure the connectivity of the swarm more efficiently.

6.3 Exploratory Agents

Similar to the use of communication agents in the distributed shared memory model, there is also a possibility of using a special kind of agents, in order to enhance the **PCF** algorithm, in the centralized shared memory. Such agents are called *exploratory agents*.

The goal of the exploratory agents is to try and add robustness to the basic algorithm used. Since the main principle of the **PCF** algorithm is sorting all the known sub-cliques according to their size, unexplored edges and proximity to the agents, it can easily be seen that in some "adversary generated" cases, a swarm of agents utilizing this algorithm will face much difficulties in finding a requested $k-clique$. In fact, such swarm may complete the mission even slower than a swarm of agents which decide their next movements randomly ! It is true that such cases are rare, however, a method of preventing them (or at least — restraining their effect) is required.

The allocation of a portion of the agents as exploratory agents achieves exactly that — by following the very simple rule of "*selecting the next goal vertex randomly*", the agents are immune to cases in which the logic of the **PCF** algorithm fall into some corner-case of the problem topology. Thus, the use of exploratory agents adds the needed robustness to the system, albeit slightly decreasing its performance, since less agents are now using the **PCF** algorithm.

When examining the issue of exploratory agents many questions arise, similar to those which concerns the case of communication agents (see section6.2). Should the identity of the exploratory agents be decided upon initialization or should it use dynamic allocation ? Should the number of exploratory agents remain constant throughout the operation of the agents, or should it change (in which case, should it increase or decrease, as time advances ?).

Similar to the case of communication agents, many techniques have been tested, trying to find the best scheme for the use of exploratory agents. The results in section7 provide some answers as to the method of use.

6.4 Meta Data

Throughout this work, we assume that the agents use the shared memory (be it centralized or distributed) in order to store only information which concerns the graph's structure. However, allowing the agents to store other kinds of information, may dramatically improve their search performance, since the essential element which stands in the basis of all swarm algorithms is the (mostly implicit) cooperation between the agents. Thus, storing and using "*swarm meta-data*", which describes the past movements and even future movements which the agents intend to perform, can greatly increase this cooperation and as a result, create a more efficient and organized swarm.

Such meta data may include for example a time stamp for each piece of information deposited into the shared memory. Such time stamps enable the reconstruction of the recent movements the swarm's agents had performed (when "fresh" time stamps indicate a presence of agents in this region, lately).

Apart from examining what kind of meta data can be produced and stored in the shared memory, the specific nature of utilizing this data by the agents should also be investigated. Hitherto, we have not yet fully explored the use of meta-data in the *physical k-clique* problem, although we intend to present such results in future works.

7 Results

The algorithm was tested on *Erdös-Renyi* random graphs $G \sim G(n,p)$ where G has n vertices, and each pair of vertices form an edge in G with probability p independently of each other. In order to ensure that G has enough clique sub-graphs of size k, the value of p should be chosen wisely. Formally, the probability that k vertices of G form a clique is p^k, thus by linearity of expectation and the second moment, G will have at least $\frac{1}{4}\binom{n}{k}p^k$ cliques of size k with probability $1-o(1)$ (for further background on probabilistic arguments see e.g. [43]). Hence we choose the value of p with respect to this formula.

It is essential that the number of desired cliques in the physical graph is of adequate order (meaning — high enough). If, for example, the number of cliques in the graph is order of unity, then the considered problem becomes trivial, since solving it in the physical environment requires nearly full exploration of the graph. This will be the case for a great majority of the problem instances, in view of the fact that we assume the initial placement of the agents on the physical graph to be chosen uniformly at random. Since environment exploration and similar problems were already studied (for example, in [4]), the relevance of such sparse graphs is rather low.

Furthermore, once a graph is fully explored, every classical algorithm might be applied for finding cliques in it. Thus, in some sense the difference between the physical environment which is almost fully explored, and the classical random access model is diminished. Hence the efficiency of algorithms in physical environment should be naturally measured on problems that do not inherently require exploration of the graphs (or in other words, problems that can be decided after acquiring only a partial information on the graph).

Figure 3 contains experimental results of the search algorithm for cliques of size 10 in graphs of sizes 500 and 2000. The number of agents searching the graphs is 5 through 50 agents. The communication model assumed is the centralized shared memory. It can be seen that while adding more agents dramatically improves the search time, the system reach a saturation at some level, after which adding more agents yield a mild contribution to the performance, if any. This can be best observed in the smaller graph, in which enlarging the number of agents from 15 to 50 decreases the search time by less than 30%, where as increasing the number of agents from 5 to 15 decreases the search time by 70%.

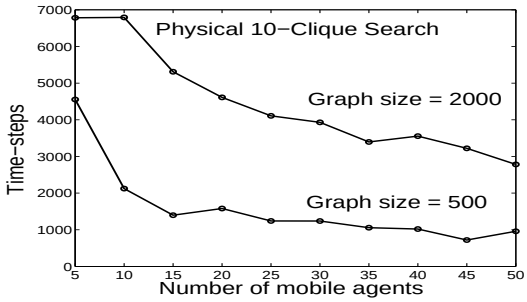


Fig. 3. Results of the Physical 10-Clique problem, with a centralized shared memory model. Notice how the performance of the swarm increases while adding more agents, until it reaches a saturation.

Figure 4 examines the increase in the swarm's search time for larger graphs. Notice that unlike the regular *k-clique* problem, in which the search time grows exponentially with the input graph's size, since we demanded that the number of cliques will be pro-

portional to the graph's size, the search time should increase moderately, or does not increase at all (depending on the exact edge probability used while building the graphs). Note that the search time for 20 agents is consistently smaller than for 10 agents. However, while in smaller graphs the difference is approximately 50%, for larger graphs it shrinks down to 25% for a graph of 4500 vertices. Interestingly, when increasing the number of agents to 30, the performance almost do not improve at all (meaning that the system had reached a state of saturation).

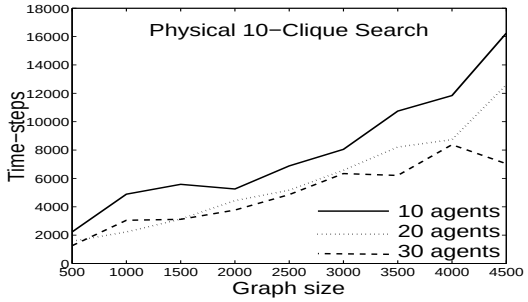


Fig. 4. Results of the Physical 10-Clique problem, with a centralized shared memory model. Notice how the performance of the swarm increases while increasing the number of agents from 20 to 30, whereas this number is increased to 30, the performance almost never changed.

Figure 5 presents the improvement in performance of a swarm working within the distributed shared memory model, as the number of agents increases. Notice that the improvements rate is much faster that in the centralized shared memory model. Figure 6 demonstrates how the performance of the distributed model improves faster than those of the centralized model.

Figure 7 demonstrates how the performance of the distributed model are only about 180% than those of the centralized model. This is most important, since one might suspect that replacing the strongest model possible (this of complete knowledge sharing) with a much simpler one (of distributed memory) might cause a dramatic decrease in performance.

Figure 8 contains a comparison between the centralized model, the distributed model, and the distributed model where 10% of the agents are assigned to be communication agents. Notice that the results show that (at least in this case) adding the communication agents slightly decrease the system's performance.

Figure 9 examines the change in performance in comparison to the size of the input graph. Results for both centralized and distributed models are presented, where an interesting observation can be made — while the performance of the centralized model in the smaller graphs is superior to those of the distributed model, the distributed model is much less vulnerable to the increase in the vertices number. This can be observed as the size of the input graph increases — in the case of the distributed model. the perfor-

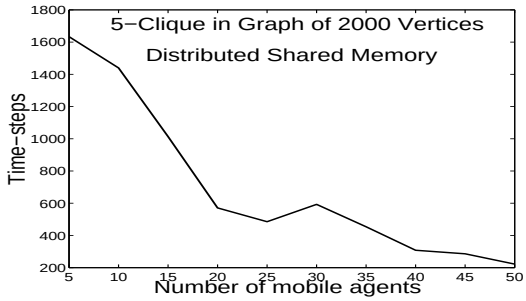


Fig. 5. Results of the Physical 5-Clique problem, in a graph of 2000 vertices, using the distributed shared memory model. Notice the fast improvement of performance as the size of swarm increases.

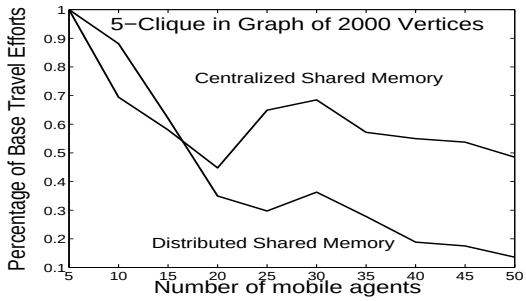


Fig. 6. Results of the Physical 5-Clique problem, in a graph of 2000 vertices — a comparison between the centralized and the distributed models. The data represents the travel efforts as percentage of the travel efforts for 5 agents. Notice how the performance of the distributed model improves faster than those of the centralized model.

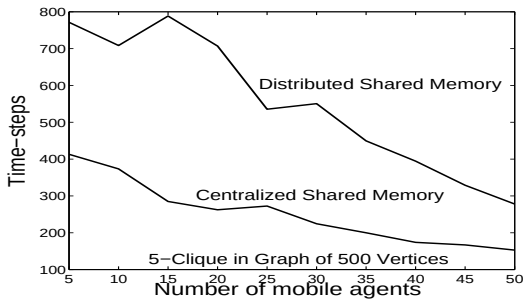


Fig. 7. Results of the Physical 5-Clique problem, in a graph of 500 vertices — a comparison between the centralized and the distributed models. Notice how the performance of the distributed model approaches slowly those of the centralized model. In average, a performance using the distributed model are only 180% than those achieved using the centralized model.

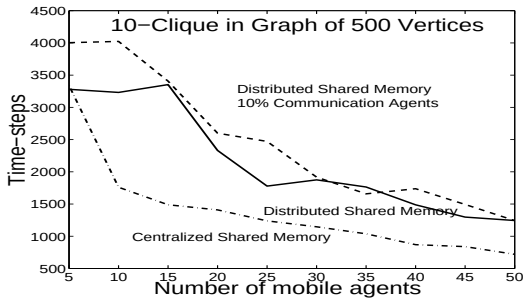


Fig. 8. Results of the Physical 10-Clique problem, in a graph of 500 vertices — a comparison between a centralized, a distributed model and a distributed model with 10% of the agents assigned as *communication agents*.

mance for a graph of 4500 vertices are only 88% more than those of 500 vertices. As the same time, the travel effort for the centralized model increases by 367%.

The reason for this remarkable phenomenon may be the lack of scalability of the centralized model, which may cause its performance to degrade in larger graphs.

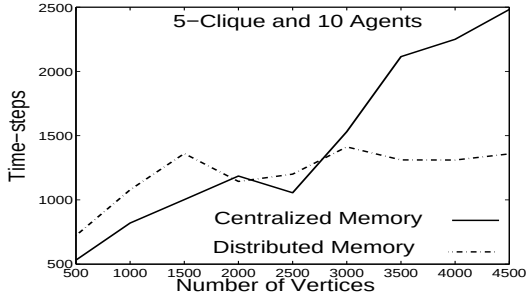


Fig. 9. Results of the Physical 5-Clique problem, in graph of growing number of vertices — a comparison between the centralized memory model and the distributed model. Note that while the performance of the centralized model in the smaller graphs is superior to those of the distributed model, the distributed model is much less vulnerable to the increase in the vertices number — the performance for a graph of 4500 vertices are only 88% more than those of 500 vertices whereas for the centralized model, the performance for a graph of 4500 vertices are 367% more than those of 500 vertices.

8 Observations

As expected, increasing the size of the swarm always improve the swarm's performance. However, there is a state of saturation to which the system can arrive, after which adding more agents will achieve only mild improvements, if any. As small the graph is, the less agents can be added before the system is saturated.

Note that while adding agents to a swarm operating in the centralized memory model do increase its performance, the effect of the same addition of agents on a swarm which operates in a distributed memory model is significantly noticeable.

Regarding the mechanism of communication agents — so far no contribution of such agents were shown, although it may be due to naive communication algorithm, which should be replaced for a more sophisticated one.

Another interesting observation is that unlike orthodox complexity scheme, used for "regular *k-clique* problem", in which the search time grows exponentially with the input graph's size, since we demanded that the number of cliques will be proportional to the graph's size, the search time of the swarms increases only moderately. This is yet another fascinating basic difference between the two complexity schemes.

Furthermore, it was discovered that while the centralized model in more affected by the increase in the graph's size, the effect of it on the distributed model is much more moderate. In fact, when working in relatively large graphs (4500 vertices) the performance of the distributed model were sometimes superior to those of the centralized model. This can be explained by flaws in the centralized algorithm. Such flaws may be inherent in the entire approach of such centralized algorithms, which leaves no room for stochastic influences by the environment, which may help direct the search towards the requested patterns.

This should be combined with the fact that almost in all simulations the performance of the distributed model was always close to this of the centralized model, while the rate of improvement of the distributed model was faster (meaning that for larger swarms and for larger graphs the performance of the two models would become more and more similar). Together, these two observation help demonstrate the basic hypothesis of this work, which is that while the distributed may appear at first simple and limited, it's strength is actually not far from this of the strongest knowledge sharing model possible. Taking into account the scalability, fault tolerance and the simplicity of the implementation of this memory sharing model, emphasizes its advantages even more.

9 Related Work

Hitherto, there have been various works which examine problems in physical graphs, as well as works which uses a distributed shared memory of some sort, have been performed. For example, [6, 14] use mobile agents in order to find shortest paths in (dynamically changing) telecom networks, where the load balancing on the edges of the graph is unstable. Similarly, several widely used Internet routing algorithms (such as BGP [34], RIP [37] etc') propagate shortest path information around the network and cache it at local vertices by storing routing tables with information about the next hop. Another known routing system which uses "ants" and "pheromonoes" is the *AntNet* system [5]. In AntNet, ants randomly move around the network and modify the routing tables to be as accurate as possible.

While these approaches seem similar, there is a great difference between these works and one presented here, both concerning the environment in which the agent operate, the data that is stored at each vertex and the problem to be solved.

First, most of the works mentioned above which concern routing problems, assume that the environment is a telecom network of some sort, which changes dynamically over time. In this work, we assume the graph to be fixed and corresponds to a real physical environment (with Euclidean distances or some other metric), while the difficulty is derived from fact that the graph is not known to the agents.

Another difference is that these algorithms try to help a mobile agent (which for example, represents a phone conversation) to navigate to a certain goal. In other words, given the destination vertex (which could be any vertex of the environment) and the knowledge written in the current vertex, these algorithms try to locally decide where should the agent go to next, while the goal of each agent is merely to minimize the time it takes this particular agent to reach its destination. In the *physical k-clique* problem,

on the other hand, the agents' goal is to find k vertices which form a *k-clique*, while the goal of the agent is to minimize the time it takes the entire *swarm* to find such a clique.

Third, when dynamic graphs are concerned, the accuracy of the data stored within the vertices often degrade over time. Thus, while designing algorithms for dynamic physical graphs, much effort is being put on keeping the data as accurate as possible at all times. This is often achieved by having ensuring that the mobile agents consistently update this information. In such paradigm, the accuracy of the stored data must be analyzed by the agents, which must decide whether this data is still relevant. Since our problem assumes that the data placed at a vertex stays valid at all times, the algorithm presented contain several basic differences from the ones mentioned above.

Essentially, the approach presented in this work can be seen as a generalization of the *next hop lookup tables* mentioned earlier, since according to the partial graph information stored in each vertex, an agent decides on its next move, when this process is continually advancing, until a clique is found.

Similar to this approach, the work of [15] also employed a distributed shared memory (a *distributed blackboard*, or a *distributed database*), while the works of [41] and [42] present a mechanism which find the shortest path within a physical graph, while assuming several communication mechanism which are similar to the concepts which appear in this work.

10 Summary and Future Work

This work had presented the *physical k-clique* problem, where a swarm comprising n mobile agents travels along the vertices of a physical graph, searching for a *clique* of size k. In order to share information between the agents, two communication models were examined. First, a full communication model is assumed, referred to as *centralized shared memory*. Than, a *distributed shared memory* model is examined, which enables the mobile agents to store and extract information using the graph's vertices.

The **PCF** search algorithm, which can be employed by a swarm of mobile agents traveling along the graph's edges, was presented. The algorithm was discussed and experimental results for it were shown.

The results demonstrated that the distributed shared memory model can produce results which are very close to those achieved by using the much stronger centralized shared memory model.

While examining the *physical k-clique* problem, new interesting opportunities for an extended research have emerged. We have already started investigating some of the above, producing more valuable results. Following are various aspects of this ongoing and future research :

1. The current version of the **PCF** algorithm was designed for static physical graphs. The problem of clique finding in dynamic physical graphs, however, bears a great deal of interest. The dynamic version of this problem would concern a graph G which in each time step t, each pair of its vertices u and v such that $\neq ((u, v) \in E_t)$ has a probability p_{on} for $(u, v) \in E_{t+1}$ and each pair v, u such that $(u, v) \in E_t$ has a probability p_{off} for $\neg((u, v) \in E_{t+1})$. Such a problem has many interesting applications as well.

2. As mentioned in previous sections, although the algorithm described in this work was designed in order to find *k-cliques* in a physical graph, we believe that by minor adjustments a generic algorithm for finding a given pattern in a physical graph may be composed. We have starting examining this issue, in hope of presenting such an algorithm in the future.
3. Since the algorithm was designed to be employed by a swarm of mobile agents, it is highly dependent on the proliferation of information (performed by the agents by the data merge processes). One method used in order to enhance this aspect of the algorithm, was the introduction of the communication agents to the system (see section 6.2. However, since this aspect significantly affects the performance of the algorithm, an analysis of the information proliferation process should be performed. Among others, we hope to achieve bounds for the rate in which information is distributed in graphs, as well as for the optimal percentage of communication agents and correlation of this value and for the features of the input graph.
4. As important the use of communication agents will become, more emphasis should be put no designing "smarter" and more efficient agents of this type. More complex heuristics ought to be developed in order to accelerate the performance of the communication agents in gathering the maximal amount of information and distributing it optimally (by estimating where and when it will be sought for).
5. The same goes for exploratory agents, which are used in the centralized shared memory model.
6. In addition to the development of better ways of using the communication and exploratory agents, more research is needed considering the best mechanism of controlling p_{comm} and $p_{explore}$ — the probabilities of becoming such an agent. Extensive search for the optimal parameters defined in sections 6.2 and 6.3 is required, as well as developing new methods of dynamically changing these probabilities.
7. Apart from performing simulations for the algorithm, and in order to fully understand the difficulty of the problem and the performance of the algorithms developed for it, analytic bounds should become available. to correctly appreciate the performance of algorithms for the problem, a lower bound for the shortest search time for a *k-clique*, for a given input graph, should be developed. Such a bound will help put into the right perspective empirical results of the simulations, since it will not be algorithm dependent. On the other hand, upper bounds for the search time for the specific algorithm used should be developed as well, in order to guarantee certain minimal performance of those algorithms.
8. An additional research should also be performed concerning the domain itself — its features and topology. Such research can help us to better foretell for example whether the use of communication and exploratory agents is expected to improve the swarm's performance. Such research can also help us design better variants of the communication and exploratory algorithms, as well as of the **PCF** algorithm itself.
9. The use of dynamic data, as suggested in section 6.4, should be studied.

References

1. I.A. Wagner, M. Lindenbaum, A.M. Bruckstein: "Efficiently Searching a Graph by a Smell-Oriented Vertex Process", Annals of Mathematics and Artificial Intelligence, Issue 24 (1998), pp. 211–223
2. R.C. Arkin, T. Balch: "Cooperative Multi Agent Robotic Systems", Artificial Intelligence and Mobile Robots, MIT/AAAI Press, Cambridge, MA (1998)
3. B.D. McKay: "Practical Graph Isomorphism", Congressus Numerantium, 30, pp. 45–87, (1981).
4. M.A. Bender, A. Fernandez, D. Ron, A. Sahai, S.P. Vadhan: "The power of a pebble: Exploring and mapping directed graphs", In the proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computating, pp. 269–278, Dallas, Texas, May (1998).
5. G. Di Caro, M. Dorigo: "AntNet:Distributed stigmergetic control for communiction networks", Journal of Artificial Intelligence Research, 9:317–365, (1998).
6. S. Appleby, S. Steward: "Mobile software agents for control in telecommunication networks", British Telecom Technology Journal, 12, pp. 104–113, (1994).
7. D.G. Corneil, C.C. Gotlieb: "An efficient algorithm for graph isomorphism", Journal of the Association for Computing Machinery, 17, pp. 51–64, (1970).
8. W.J. Christmas, J. Kittler, M. Petrou: "Structural Matching in Computer Vision Using Probabilistic Relaxation", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 8, pp. 749–764, (1995).
9. F.R. Adler, D.M. Gordon: "Information collection and spread by networks of partolling agents", The American Naturalist, 140(3):373–400, (1992).
10. D.M. Gordon: "The expandable network of ant exploration", Animal Behaviour, 50:372–378, (1995).
11. M.R. Garey, D.S. Johnson: "Computers and Intractability: A Guide to the Theory of NPCompleteness", Freeman & co., New York, (1979).
12. I.A. Wagner, A.M. Bruckstein: "ANTS: agents, networks, trees and subgraphs", Future Generation Computer Computer Systems Journal, 16(8):915–926, 2000.
13. V. Yanovski, I.A. Wagner, A.M. Bruckstein: "Vertex-ants-walk: a robust method for efficient exploration of faulty graphs. Annals of Mathematics and Artificial Intelligence, 31(1–4):99–112, (2001).
14. R. Schnooderwoerd, O. Holland, J. Bruten, L. Rothkrantz: "Ant-based load balancing in telecommunication networks", Adaptive Behavior 5(2), (1996).
15. P. Valckenaers, V. Marik, D.C. McFarlane: "Holonic and multi-agent systems for manufacturing", First International Conference on Industrial Applications of Holonic and Multi-Agent Systems, (1993).
16. J.B.H. Kwa: "BS*: An Admissible Bidirectional Staged Heuristic Search Algorithm", Artificial Intelligence, pp. 95–109, Mar., (1989).
17. S.N. Dorogovtsev, J.F.F. Mendes: "Evolution of networks", Adv. Phys. 51, 1079, (2002).
18. D.J. Watts: "Small Worlds", Princeton University Press, Princeton NJ, (1999).
19. D.S. Hochbaum, O. Goldschmidt, C. Hurken, G. Yu: "Approximation algorithms for the k-Clique Covering Problem", SIAM J. of Discrete Math, Vol 9:3, pp. 492–509, August, (1996).
20. P. Cucka, N.S. Netanyahu, A. Rosenfeld: "Learning in navigation: Goal finding in graphs", International journal of pattern recognition and artificial intelligence, 10(5):429–446, (1996).
21. R.E. Korf: "Real time heuristic search", Artificial intelligence, 42(3):189–211, (1990).
22. L. Shmoulian, E. Rimon: "Roadmap A*: an algoritm for minimizing travel effort in sensor based mobile robot navigation", In the proceedings of the IEEE International Conference on Robotics and Automation, pp. 356–362, Leuven, Belgium, May (1998).
23. R.J. Wilson: "Introduction to Graph Theory", Longman, London, 2nd ed., (1979).

24. A. Stentz: “Optimal and efficient path planning for partially known environments.”, In the proceedings of the IEEE International Conference on Robotics and Automation, pp. 3310–3317, San Diego, CA, May (1994).
25. L.P. Cordella, P. Foggia, C. Sansone, M. Vento: “Evaluating Performance of the VF Graph Matching Algorithm”, Proc. of the 10th International Conference on Image Analysis and Processing, IEEE Computer Society Press, pp. 1172–1177, (1999).
26. V. Yanovski, I.A. Wagner, A.M. Bruckstein: “A distributed ant algorithm for efficiently patrolling a network”, Algorithmica, 37:165–186, (2003).
27. V. Maniezzo, M. Dorigo, A. Colorni: “The ant system: Optimization by a colony of cooperating agents”, IEEE Transactions on Systems, Man and Cybernatics — Part B, 26(1), pp. 29–41, (1994)
28. P. Kuner, B. Ueberreiter: “Pattern recognition by graph matching — combinatorial versus continuous optimization”, Int. J. Pattern Recognition and Artif. Intell., vol. 2, no. 3, pp. 527–542, (1988).
29. J.R. Ullmann: “An Algorithm for Subgraph Isomorphism”, Journal of the Association for Computing Machinery, vol. 23, pp. 31–42, (1976).
30. R. Albert, A.L. .Barabasi,: “Statistical Mechanics of Complex Networks”, Reviews of Modern Physics, vol. 74, January, (2002).
31. K.A. De Jong, W.M. Spears: “Using genetic algorithms to solve NP-complete problems”, in Genetic Algorithms, (J.D. Schaffer, ed.), Morgan Kaufmann, Los Altus, CA., pp. 124–132, (1989).
32. A.A. Toptsis, P.C. Nelson: “Unidirectional and Bidirectional Search Algorithms”, IEEE Software, 9(2), (1992).
33. S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi: “Optimization by Simulated Annealing”, Science, 220, pp. 671–680, (1983).
34. Y. Rekhter, T. Li: “A Border Gateway Protocol”, Request for Comments 1771, T.J Watson Research Center IBM Corporation & cisco Systems, March (1995).
35. Y. Altshuler, A.M. Bruckstein, I.A. Wagner: “Swarm Robotics for a Dynamic Cleaning Problem”, in “IEEE Swarm Intelligence Symposium” (IEEE-SIS05), (2005)
36. Y. Altshuler, V. Yanovski: “Dynamic Cooperative Cleaners — Various Remarks”, Technical report, (2005)
37. G. Malkin: “RIPng Protocol Applicability Statement”, RFC 2081, IETF Network Working Group, January, (1997).
38. O. Gerstel, S. Zaks: “The Virtual Path Layout problem in fast networks”, In Proceedings of the Thirteenth Annual ACM Symposium on Principles of Distributed Computing, pp. 235–243, Los Angeles, California, August, (1994).
39. J.Y. Cai: “Lectures in Computational Complexity”, Department of Computer Sciences, University of Wisconsin Madison, WI 53706 (2003).
40. S. Zaks: “Path Layout in ATM networks”, Lecture Notes in Computer Science, 1338:144-177, (1997).
41. A. Felner, R. Stern, A. Ben-Yair, S. Kraus, N. Netanyahu: “PHA*: Finding the Shortest Path with A* in Unknown Physical Environments”, Journal of Artificial Intelligence Research, vol. 21, pp. 631–679, (2004).
42. A. Felner, Y. Shoshani, I.A.Wagner, A.M. Bruckstein: “Large Pheromones: A Case Study with Multi-agent Physical A*”, Forth International Workshop on Ant Colony Optimization and Swarm Intelligence, (2004).
43. N. Alon, J. H. Spencer: “The probabilistic method”, Wiley-Interscience (John Wiley & Sons), New York, (1992) (1^{st} edition) and (2000) (2^{nd} edition).
44. A. Apostolico, Z. Galil: “Pattern Matching Algorithms”, Oxford University Press, Oxford, UK.

45. V. Vazirani: “Approximation Algorithms”, Springer-Verlag, (2001).
46. S. Arora, S. Safra: “Probabilistic checking of proofs: A new characterization of NP”, Journal of the ACM, (1998).
47. M. Garey, D. Johnson: “Computers and Intractability: A Guide to the Theory of NP-Completeness”, San Francisco, CA: W. H. Freeman, (1979).

Chemotaxis-Inspired Load Balancing

Geoffrey Canright[1] **, Andreas Deutsch[2], and Tore Urnes[1]

[1] TELENOR Reseach and Development, Fornebu, Norway.
geoffrey.canright,tore.urnes@telenor.com
[2] Center for High Performance Computing, Dresden University of Technology, Dresden, Germany.
deutsch@zhr.tu-dresden.de

Abstract

We present an approach to the problem of load balancing on networks of nodes. Our approach is inspired by the phenomenon of negative chemotaxis in living systems. We use a diffusing signal (which is emitted by load, and moves faster than the load) to guide the movement of load towards the balanced state. Our reference system (for comparison) is unguided, diffusing load, moving at the same speed. Our tests show that the chemotaxis system can give large improvements over the reference system in convergence speed, as well as showing much reduced sensitivity to variations in network topology and in initial load distribution.

Short title: Chemotaxis-Inspired Load Balancing
Keywords: Load balancing, diffusion, chemotaxis, adaptivity, bio-inspired engineeering

1 Introduction

The problem of load balancing is quite general. Technological applications include balancing processing jobs among parallel processors [1]; balancing processing among distributed computers in a grid processing system—either peer-to-peer [2], or otherwise [3]; and balancing storage among networked storage devices [4]. There are also non-technological instances of load balancing. One of the very simplest may be taken from physics. That is, *diffusion* is a primitive mechanism which quite reliably seeks to distribute particles so as to move them from non-uniform states to uniform states. Furthermore, this is accomplished in a fully decentralized, even "unintelligent", manner, using only the random thermal motion of the particles.

One finds load balancing also in biological systems. One example is the use of chemical repellents to prevent growth of yeast colonies into one another's space [5]. The fact that bacteria are able to communicate with each other changed our general perception of many single, simple organisms inhabiting our world. As well as releasing the signalling molecules, yeast cells are also able to measure the number (concentration) of the molecules within a population. The term 'Quorum Sensing' (QS) is used to describe the phenomenon whereby the accumulation of signalling molecules enable a single cell to sense the number of bacteria (cell density) and to behave accordingly.

This kind of territorial load balancing also occurs with larger organisms. In most cases, some kind of signal (urine, bird songs) is used to maintain distance between individuals or groups. Hence, in general, load balancing in biological systems is more *active* than the correspondingly passive, physical mechanism of diffusion. We note also that living systems are distinguished from most nonliving ones

** *Corresponding author's address:* Telenor R&D, B6d, Snarøyveien 30, 1331 Fornebu, Norway
Phone: +47-91815638
Fax: +47-96211086
Email: geoffrey.canright@telenor.com

by being able to perform the *opposite* of load balancing. That is—again using active mechanisms—living systems can establish and maintain highly non-homogeneous states; examples include processes such as aggregation, morphogenesis, and self-repair.

We are interested in the problem of designing distributed, self-managing systems for solving technological problems. Also, as members of the BISON [6] project, we focus on biologically-inspired solutions to such problems. In this work, we wish to explore and evaluate a mechanism for load balancing in distributed systems. This mechanism is rather directly inspired by a common biological mechanism for controlling the aggregate behavior of many small entities, namely, chemotaxis. Chemotaxis means the control of movement (taxis) via diffusive chemical signals. It is used in biology both for bringing about homogeneous distributions (negative or repulsive chemotaxis), and for inducing highly non-homogeneous aggregates (positive chemotaxis) [7, 8]. In particular, chemotaxis is essential for guiding cells in biological development, the process of wound healing, and tumour angiogenesis (a process by which a tumour is able to guide the growth of neighbouring vasculature into its own direction). Here we will implement a form of repulsive chemotaxis on a networked system, with the aim of bringing about time-efficient load balancing.

An essential aspect of chemotaxis is that the signal can diffuse faster than the bodies which are guided by the signal. In other words, there are *two distinct time scales* in operation in chemotaxis. This notion of two time scales is central to our thinking. Specifically, we assume that the load (which we specify only as an abstract, real-valued quantity at each node of the network) has a natural rate of movement that is significantly lower than that for the signal. We believe that, when this two-time-scale picture is correct, the use of signal can significantly speed up the movement of load towards a uniform (balanced) distribution. To put this picture into more relevant technological terms, we might suppose that the load is stored data, on the order or megabytes or gigabytes, while the signal is an indicator requiring only a few bytes. Then—possibly with the addition of prioritized queueing—we can assume that signalling data moves over the network much faster than load.

Given then that signal moves faster than load, we expect the effect of the signal to be a speedup of convergence to the balanced state. That is, we can imagine a system without signal. Repulsion or attraction can still be turned on whenever the bodies come into direct contact with one another. For example, the growing yeast colonies can stop growing when they contact others; or, in the attractive case, the amoebae of slime molds [9, 10] can stick together when they come into contact. Conceivably, this mechanism can also achieve the desired distribution—but on a much longer time scale, since the movement of the amoebae is "blind" until they actually contact others. In fact, the (load balancing or aggregating) system, without signal, accomplishes redistribution via *collisions*; hence it works in much the same way as diffusion, with particle movement being completely "blind" between collisions.

Based then on our understanding of the utility of chemotaxis in living systems, we will use "blind" movement of load as our reference system. That is, we will compare a signal-aided load balancing system on a network with a passive load balancing process, which simply uses the network equivalent of plain diffusion. Our aim in making this comparison is simple: we wish to test the hypothesis that the mechanism of chemotaxis (with its two time scales) can significantly enhance the speed of convergence of load to the uniform state, with respect to the *same* system operating without signalling. And we are saying here that the "same" system, without signals, is plain diffusion.

We will make this comparison more precise below. The principal idea is to ensure that the load has the same time scale, or in other words the same speed of movement, both when it is influenced by signal, and when it is not. We call this our "fairness criterion". Given a precise prescription for ensuring this kind of fairness, we can then test whether chemotactic signalling does in fact enhance the approach to convergence for load balancing.

We are not aware of other works which have applied the idea of chemotaxis to load balancing.

The use of diffusion methods for load balancing on networks has however a considerable history. One of the earliest works is that of Cybenko [11], who gave constraints for keeping the load strictly positive, and then derived some stability criteria. Cybenko also looked at linear modifications of the adjacency matrix as a way of speeding up convergence. There has been considerable subsequent work on diffusion-like mechanisms for load balancing [12–18]. A recent survey of diffusive load balancing may be found in [19]. It is in fact the simplicity and decentralized nature of diffusion that makes it attractive for handling large and/or dynamic networks, where optimization based on global information is not practical. Our aim here is to see if chemotaxis can retain the good features of diffusion, while offering significant imporvement in performance and flexibility.

The remainder of this paper is organized as follows. In section 2 we develop equations for our two algorithms—plain diffusion and chemotaxis. Also, we derive a fairness criterion in this section. Then, in section 3, we specify a number of details which are needed to translate our general equations into specific implementations for discrete-time, networked systems. Section 3 gives our basic algorithm for plain, slow diffusion, and then goes on to give a detailed discussion of our studies of *fast* diffusion. This problem merits some attention, since we wish to find a stable method for fast diffusion of signal; and it is known, at least since the work of Cybenko [11], that such a problem is nontrivial for discrete-time systems on a discrete space (network). Finally, in this section, we give the details for the two-component, coupled system that represents chemotactic load balancing. Next, in Section 4, we evaluate the performance of chemotactic load balancing with respect to our reference system (unguided diffusion of slow-moving load). Here we find some encouraging confirmation of our hypothesis—that "smart" (signal-aided) diffusion can give faster, or even much faster, convergence than "dumb" diffusion. Section 5 then addresses the question of the *adaptivity* of the chemotactic system. That is: being both decentralized and "smart", such a system might be expected to also be adaptive—to variations in network topology, initial load distribution, etc. In section 5 we will offer some limited answers to this question. We also include a section (section 6) on the remarkably complex behavior over time that we have observed for this simple, two-component (chemotactic) system. Its behavior stands in marked contrast to the smooth, predictable behavior of slow, single-component diffusion for the same network topology. Finally, in section 7 we offer a summary.

2 Chemotaxis

First we define our terminology. We let ϕ_i represent the load at node i. We will normally enforce the constraint that $\phi_i \geq 0$ for all i and for all times. We will also assign to each node a *capacity* C_i. Then the load balancing problem is to seek to minimize the difference between load and capacity at each node. For this exploratory study, we set all capacities to one, and choose start distributions so that total load equals total capacity for the network. Thus the converged state is a uniform distribution with all $\phi_i = 1$.

Now we address plain diffusion. The physical mechanism of diffusion, when mapped onto a discrete (space and time) system, may be expressed quite simply: nodes simply send out a fraction of what they have, at each time step, in all "directions" (to each neighbor). The fraction is then the "diffusion constant" c; while what the nodes have to send is quantified as $\phi_i - C_i$. Hence, for the simplest version of plain diffusion, a node i with load ϕ_i and capacity C_i will send a small fraction c of its excess load $(\phi_i - C_i)$ to each of its neighbors, independent of node, of neighbor, and of time. Each transfer of load to a neighbor node j can then be captured by the following equation:

$$\Delta\phi_{i\to j} = c \cdot (\phi_i - C_i) \tag{1}$$

This is our basic equation for discrete diffusion. In the next section, we will address a number of modifications to this "most simple" form of diffusion, so as to keep the load at each node positive, to enforce that transferred load is also positive, and also to allow for very fast diffusion without instabilities.

Now we introduce signal. For the signal to be useful, it must (i) move faster than load, and (ii) give information to distant nodes about the load distribution. Hence we want the load itself to "emit" signal, which then diffuses away, bringing useful information to distant nodes. Let us define S_i to be the signal level at node i. Then our emission equation is

$$\Delta S_i^{emit} = c_2 \cdot (\phi_i - C_i) \ . \tag{2}$$

This emission event occurs at every time step. In fact, each equation in this section is to be implemented once in each time step.

It is clear from Eq. (2) that, unless total load equals total capacity, the total signal will grow (in absolute value) steadily over time. Such growth may be useful in dynamic situations where the local signal level (not the gradient) gives information which may guide a decentralized mechnism for deciding how much new load may be accepted. However in this study we focus on the static case; and to avoid divergence of the average signal level, we set total load equal to toal capacity.

Now we want the load to be guided by the signal. Specifically, we let the load respond to local signal gradients, as follows:

$$\Delta \phi_{i \to j} = c_3 \cdot (S_i - S_j) \ . \tag{3}$$

This equation makes the movement of load less blind, and hence (we believe) more smart, than that due to plain diffusion.

Finally, we need to arrange for fast diffusion of signal. The fast versions that we develop for plain diffusion will be used to guide the movement of signal in the two-component case. Hence, the signal diffusion equation may be written analogously to Eq. 1, as follows:

$$\Delta S_{i \to j} = c_4 \cdot S_i \ . \tag{4}$$

Note that no reference to capacity is needed here–signal in fact diffuses blindly.

Equation (1) thus defines our rule for the one-component case—plain diffusion of load. Equations (2—4) give the rules for the two-component case. We require some more work however before we can implement these rules. For one thing, we have, from these four equations, four free parameters. Also, we need a way to ensure that the movement of load is "equally slow" in both the one-component and the two-component case—this is our fairness criterion. Finally, we need to decide how we can implement the notion of two time scales—signal moving faster than load. In physics, we can simply set $c_4 \gg c$; but for our wholly discrete system, ensuring stable fast diffusion is more problematic [11].

In the remainder of this section we will develop a fairness criterion, which will actually pin down two of our four constants. We will then use the next section to develop two stable versions of fast diffusion, involving some modification of Eq. (4). These two algorithms will in effect define our "constant" c_4. Also in that section we will give a simple working definition of slow diffusion, thus determining the last constant c.

In order to develop a quantitative fairness criterion, we need a precise expression for the speed of the load in both the one- and the two-component case. Naively, we would take these to be, respectively, c and c_3; hence our naive fairness criterion sets $c = c_3$.

Let us now test this naive conclusion with more detailed reasoning. For plain diffusion, we define the total load moved *out* from node i, at a given time, to be

$$\Delta\phi_{i,out}^{diff} = ck_i(\phi_i - C_i) \ , \tag{5}$$

where k_i is the degree of node i. We will take this quantity (in some roughly averaged sense) to be a measure of the "speed" of the load: it tells us (again, roughly) how much load the system, using plain diffusion, can push out in a single time step. Hence, for our fairness criterion, we set (here 'CT' means 'chemotaxis')

$$\Delta\phi_{i,out}^{diff} \simeq \Delta\phi_{i,out}^{CT} \ . \tag{6}$$

The RHS of this equation incorporates the dynamics of the signal. That is:

$$\begin{aligned}
\Delta\phi_{i,out}^{CT} &= \sum_{j=nn(i)} \Delta\phi_{i\to j} && (7)\\
&= \sum_{j=nn(i)} c_3(S_i - S_j) && (8)\\
&= c_3 k_i S_i - c_3 \sum_{j=nn(i)} S_j && (9)\\
&= c_3 k_i (S_i - S_{av,i}) \ , && (10)
\end{aligned}$$

where $nn(i)$ means "nearest neighbor of i", and the last line defines the quantity $S_{av,i}$.

Now, at any time t, the signal at node i is the newly emitted signal, plus the signal that was present at time $t-1$, minus any signal diffused away in the time step from $t-1$ to t, plus any diffused in from neighbors in that same time step. That is,

$$S_i^t = \Delta S_i^{emit} + (1 - c_4 k_i) S_i^{t-1} + c_4 \sum_j S_j^{t-1} \ . \tag{11}$$

From the work of Cybenko [11] (and also from our own experiments with fast diffusion—see next section), we know that fast, but still stable, diffusion of signal will be defined roughly by sending out everything in one time step—but not more! That is, fast stable diffusion involves letting the fraction c_4 sent to each neighbor be (roughly, or even exactly) equal to the inverse node degree. Furthermore, we will choose a signal diffusion rule—again, out of consideration for speed and stability—so that the node with more to send chooses its own inverse degree for all transfers (out *and* in) with its neighbors. If we suppose then that node i is in the role of sender (by this definition), we set $c_4 = 1/k_i$ everywhere in (11), thus obtaining

$$S_i^t = \Delta S_i^{emit} + S_{av,i}^{t-1} \ , \tag{12}$$

so that

$$S_i^t - S_{av,i}^t = \Delta S_i^{emit} - (S_{av,i}^t - S_{av,i}^{t-1}) \ . \tag{13}$$

Next we argue that the last two terms in (13) may be of any sign, and have a time average of zero. Hence we will ignore them, in order to find a way of calibrating our chemotaxis rule against simple diffusion. This gives

$$\begin{aligned}
S_i^t - S_{av,i}^t &\approx \Delta S_i^{emit} && (14)\\
&= c_2(\phi_i - C_i) \ ; && (15)\\
& && (16)
\end{aligned}$$

hence the RHS (10) of our fairness test becomes

$$\Delta\phi_{i,out}^{CT} \approx c_3 k_i c_2 (\phi_i - C_i) \ . \tag{17}$$

This equation is readily compared with the LHS of our fairness test, see Eq. (5). We can then give the same "speed" (load-moving capacity) to both plain diffusion and our chemotactic system by setting

$$c = c_3 c_2 \ . \tag{18}$$

Now we can eliminate one unnecessary constant by setting

$$c_2 = 1 \ ; \tag{19}$$

that is, we measure signal in the same units as load and capacity. Our final result for enforcing fairness is then

$$c = c_3 \ . \tag{20}$$

This is the same result as our naive guess. In fact, even more detailed arguments [20] still give the same result. Hence we take (20) as our working fairness criterion for comparing plain diffusion to diffusion aided by signal.

3 Algorithms

To implement the simple equations for plain and signal-aided diffusion of slow-moving load on a given overlay network topology we must develop corresponding algorithms. We start by presenting our algorithm for plain diffusion. This algorithm is used as a reference for comparison with chemotaxis-inspired load balancing throughout the paper. Next, we turn to the development of our algorithm for signal-aided diffusion. We exploit our assumption that signal can move fast, and propose two algorithms for signal diffusion. We also present our simulation model for diffusion and compare the plain diffusion algorithm for load to the fast signal diffusion algorithms. We are then ready to present our algorithm for signal-aided diffusion.

3.1 Plain diffusion

Our equation for plain diffusion, Eq. (1), exhibits two questionable features: negative load is sent whenever a node's load is less than capacity, and a node's load may become negative. Each of these features is either unrealistic or meaningless; hence we introduce simple modifications to Eq. (1) to address this. To remove the possibility of sending negative load we find the net difference $(\phi_i - C_i) - (\phi_j - C_j)$ for each node-neighbor link ij. Then the node with the largest (most positive) difference between load and capacity is chosen as the sending node and only the net, positive quantity of $c \cdot |(\phi_i - C_i) - (\phi_j - C_j)|$ is sent. To prevent a node's load from becoming negative we must ensure that no node sends more load than it has. If a node i has k_i neighbors then the total load sent in one time step is at most $c \cdot k_i \cdot (\phi_i - C_i)$. Hence, if c is chosen to be less than $1/k_i$ for all nodes i then loads will always remain positive.

We recall that we want the load to be slow-moving. Hence we see no reason to implement, for plain diffusion, a "smart" strategy which chooses a local value for c that still satisfies $c < 1/k_i$. Instead, for this study, plain diffusion is given a global value for c—which we set to be less than $1/k_{max}$, where k_{max} is the highest degree of any node in the system. We will however look at "smart" strategies (with locally , time- and space-varying diffusion "constants") for fast diffusion of signal.

3.2 Signal diffusion

Signal and signal diffusion are not restricted in the same ways as load. Specifically, we assume that signal can move quickly, that signal can take on negative values, and that signal diffusion mechanisms need not be mass preserving. After considerable experimentation, exploring a range of algorithms, we came up with two candidate algorithms for fast signal diffusion. The two algorithms are, for historical reasons, named the "version 6" and "version 10" algorithms.

The version-6 algorithm is based on the algorithm for plain diffusion presented above. Though employed by us as a signal diffusion algorithm, it is also suitable for diffusing load in systems that do not have restrictions on how quickly load can move. The diffusion constant c is assigned a "default" value c_{default}. Any node i which discovers that $c_{\mathrm{default}} > 1/k_i$ (where k_i is its degree), will adjust its own c value to be precisely $1/k_i$ in order to avoid negative load values. Hence, two neighbors i and j who have both adjusted their values by this rule will have diffusion constants $1/k_i$ and $1/k_j$, respectively. This gives an asymmetry which can seemingly violate mass conservation. The problem is solved by picking a sending node, and only transferring the net positive difference as in the algorithm for plain diffusion. The effect is that the sending node forces its choice of c on both ends of a given link, making the version-6 algorithm mass conserving. An interesting feature of the version-6 algorithm, that we exploit in our experiments below, is that its speed can be continuously tuned: maximum speed is obtained by setting $c_{\mathrm{default}} = 1$, while decreasing values reduce the speed correspondingly.

The version-10 algorithm is only suitable for signal diffusion because it does not maintain a strictly positive "load" (in our case, signal). There are similarities between the version-6 and version-10 algorithms, but contrary to version-6, version-10 has each sending node i always choosing $1/k_i$ as its diffusion constant. Also, the definition of sending node is modified to allow for the fact that the sent quantity (signal) can be negative.

As noted in the previous section, we require that the total amount of signal in the network be zero. This enables convergence of the signal to a fixed point (zero signal on each node). Knowing this fixed point, we can define the "sending node" of two nodes i and j to be the node which has its signal value farthest from zero, i.e., farthest from the signal value corresponding to the uniform fixed point distribution. A sending node i sends its neighbor node j the amount of signal equal to $(S_i - S_j)/k_i$.

We have conducted simulation experiments that compare the performance of our plain diffusion algorithm with that of our version-6 and version-10 signal diffusion algorithms. Before presenting those results, we first briefly describe the simulation model that we used.

3.3 Simulation model

We used the Peersim simulator framework [21] to conduct all performance evaluations of diffusion algorithms.

We recall that the aim of load balancing is to reach a state of convergence where nodes in a network receive loads according to their capacities. Recall further that, to simplify our simulation model, we require that all nodes have the same capacity, the capacity of 1.0, and that total load equals total capacity. To reach a state of absolute convergence, i.e., when all nodes have load equal to capacity, typically takes a long time since smaller and smaller load amounts are moved as the system under study approaches convergence. Our convergence criterion is therefore more relaxed than absolute convergence.

It is our experience that faster algorithms for diffusion do not approach convergence in a smooth, monotonic fashion. Typically the trend is clearly that of gradually converging smallest and largest load values in the network, but occasionally the trend is broken with small intervals of surges. Taking

this jittery behavior into account, we propose the following definition of convergence.

Definition. *Convergence of load balance protocol.* At the end of each simulation cycle, let min be the smallest load value in the network and max be the largest. Also, let $threshold$ be a small number. For each simulation cycle where $(max - min) < threshold$ holds, a convergence counter is incremented by one. Also, for each simulation cycle where $(max - min) >= threshold$ holds, the convergence counter is decreased by 10 or set to zero if a negative value would result. A load balance protocol simulation is then said to converge when the convergence counter reaches the value of 100. All experiments presented in this paper used a $threshold$ value of 0.1.

Important simulation model parameters are the choice of overlay network topology and start distribution for load on nodes. Unless otherwise specified, all simulation runs reported on in this paper used the same power-law network topology and random start distribution. We feel that the power-law topology and the random start distrubution offer a fairly realistic context for simulations. Our power-law topology consists of $10,000$ nodes, with the most connected node having 2200 neighbors. The power-law topology was generated by the topology generator described in [22]. To generate our random start distribution we simply divide the total load into $10,000$ units, and place one unit at a time on a randomly selected node until all units have been placed. This procedure gives a distribution of initial loads over nodes varying from zero to about 6.

We have used our simulation model to conduct numerous experiments of plain diffusion using multiple instances of both power-law topology and random start distribution. We choose the value $c = c_3 = 1/2500$ as our diffusion constant.

Figure 1 illustrates the differences in speed and behavior between our plain diffusion algorithm and our two algorithms for signal diffusion. For the sample runs shown in Figure 1, plain diffusion took $15,816$ cycles to reach convergence, while the version-6 and version-10 algorithms took 486 and 361 cycles, respectively. The instabilities that are characteristic of fast diffusion algorithms are evident from the sample runs in Figure 1. The largest and smallest values of the random start distribution were six and zero, respectively; the version-6 algorithm exhibited maximum load values as high as 370, while the version-10 algorithms saw an extreme maximum value of 160 and an extreme minumum value of -180. We also note that the largest amount of load transferred over a link was 0.0024 for the plain diffusion algorithm; for each of the two signal diffusion algorithms, the largest value transferred over a link was about 3. Thus we see a difference in this quantity by a factor of about 1000. In short: the "fast" one-component diffusion algorithms have oscillations, sometimes very high load/signal values at nodes, and sometimes very high transfers over the links.

3.4 Signal-aided diffusion

So far we have developed a single algorithm for slow diffusion of load (the plain diffusion algorithm) and two algorithms for the fast diffusion of signal (the version-6 and version-10 algorithms). We now present the practical details of our chemotaxis-inspired algorithm for load balancing: the signal-aided diffusion algorithm.

We recall our basic assumption that signal speed can be very high. Thus, for the signal, we are free to choose the fastest stable algorithm that we can find. After experimenting with half a dozen candidate algorithms, we arrived at the version-6 and the version-10 algorithms as our preferred candidates for fast signal diffusion algorithms.

Early experiments with signal-aided diffusion showed that when diffusion of load responds to signal gradients according to Eq. 3, instabilities often resulted. A closer inspection of those early results in addition to insights offered by the work of Cybenko [11] led us to believe that chemotaxis can be made less prone to instabilities if nodes that contain less load are more constrained in their response

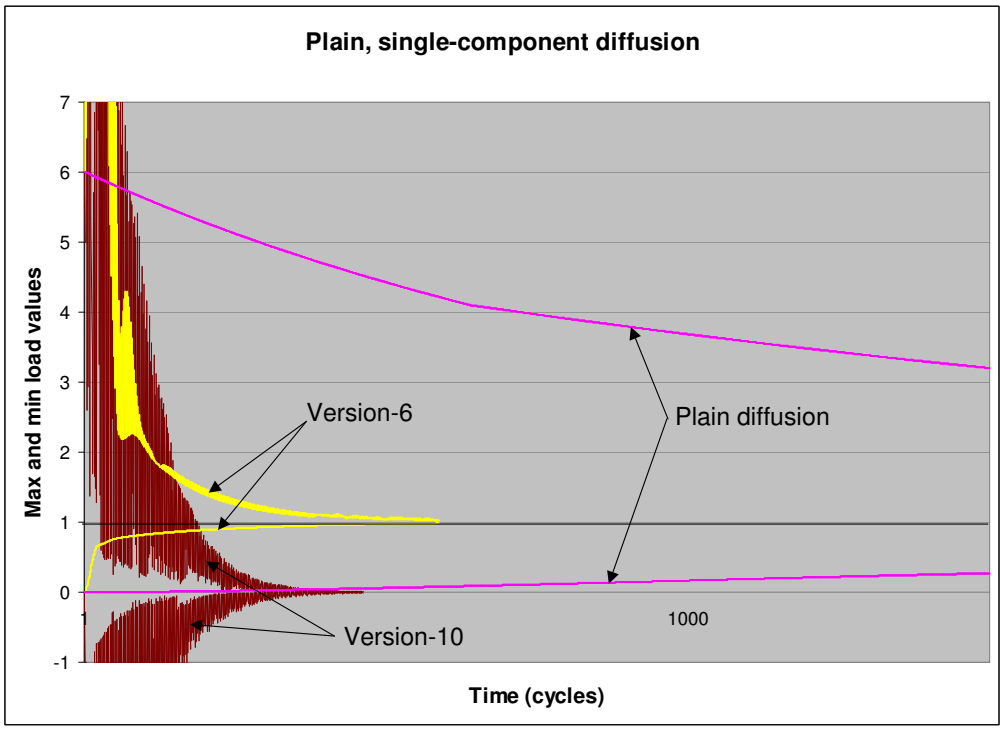


Figure 1: Sample runs of the plain diffusion algorithm and the version-6 and version-10 diffusion algorithms. Largest and smallest values at the end of each cycle are plotted for the first 1500 cycles . All simulation runs used a random start distribution on a 10,000-node scale free topology.

to signal gradients. Our algorithm for Eq. (3) therefore incorporates the following two constraints. First, only nodes with more load than capacity are allowed to send to neighbors. Secondly, the total load sent must be less or equal to the difference between load and capacity of the sending node. The effect of these constraints is that once a node receives more load than capacity, it will maintain a load of at least capacity.

The main focus of our simulation experiments is to compare plain diffusion with signal-aided diffusion. It is therefore important to ensure that comparisons between the two be fair. We showed in section 2 that fairness means that the diffusion coefficient for load must be the same for both plain and signal-aided diffusion, i.e., we choose the same value for c and c_3. Hence, c and c_3 must be network-wide constants. We also recall that c must be smaller than the inverse degree of the most connected node in the topology.

We now have two load balancing algorithms for load that is restricted to move slowly. Plain diffusion simply sends out load in all directions independent of the current load distribution. Signal-aided diffusion, on the other hand, employs a fast-moving signal to guide load in directions that contain less load.

4 Performance

In this section we report on results from simulations using the simulation model defined in the previous section. We start by presenting our results regarding time to reach convergence, then we look at the quantities of load that were moved between nodes.

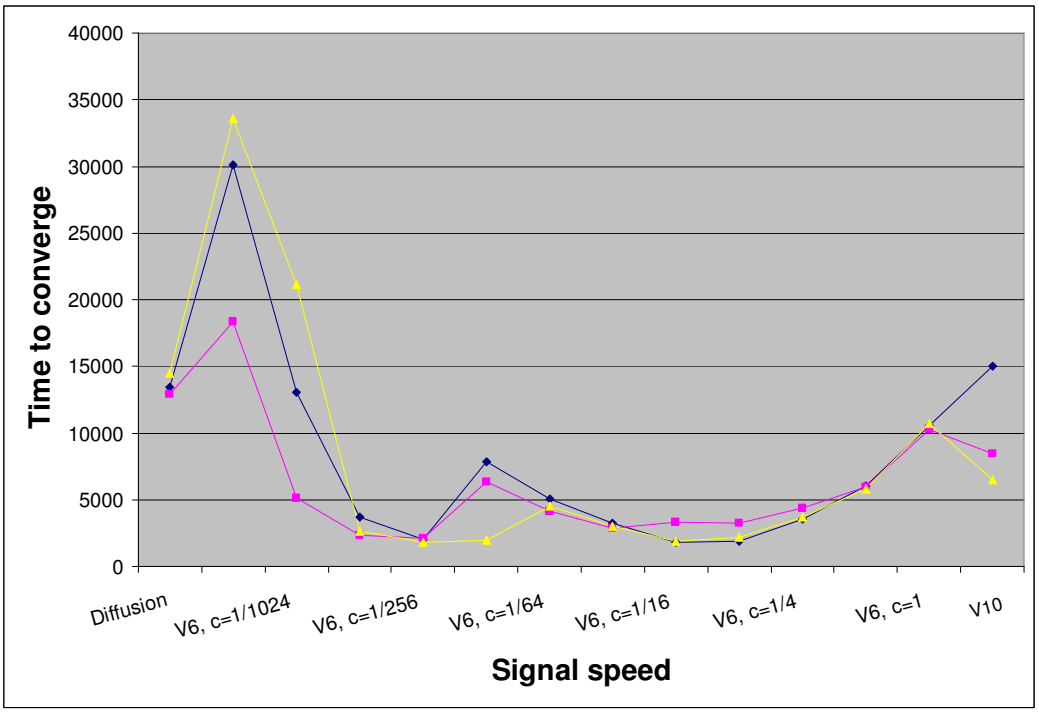


Figure 2: The effect of increasing signal diffusion speed on time to reach convergence for signal-aided load diffusion. Each of the three graphs corresponds to a different instance of our random start distribution. The convergence times for plain diffusion are plotted to the far left. All other plots show convergence times for signal-aided diffusion.

Our convergence time experiments aim at both comparing plain diffusion with signal-aided diffusion, and exploring the effect of different signal speeds on signal-aided diffusion performance. Recall that the Version 6 signal diffusion algorithm allows its speed to be altered by choosing different values for the diffusion constant c_{default}. The fastest signal diffusion speed is always obtained by our Version 10 algorithm. Version 6 with $c_{\text{default}} = 1$ gives the second fastest signal speed. Progressively slower signal speeds are then obtainded by halving the value of c_{default}. We chose Version 6 with $c_{\text{default}} = 1/2048$ as our slowest signal diffusion algorithm.

Figure 2 plots the time to reach convergence for a few sample runs, both for plain diffusion and for signal-aided diffusion with different signal speeds. Each of the three graphs in Figure 2 represents a set of sample runs sharing a particular instance of the random start distribution. Signal speed increases along the horizontal axis. Convergence times for plain diffusion are shown as the left-most plot of each graph. Note that, for this graph and for all subsequent plots showing performance, our results are based on only a small number of typical sample runs, and not on a statistical average over many runs, for each load speed. Also note that with respect to classes of topologies and start distributions, our experiments focus on only scale-free topologies and random start distributions. We revisit these issues in the future work section below and suggest that a wider range of experiment configurations is needed to provide stronger evidence that our algorithm behaves well in most situations.

As can be seen from Figure 2, convergence to a balanced load was achieved for all runs, even when signal diffused very quickly compared to the load. It is also evident that signal-aided diffusion exhibited shorter time to reach convergence than did plain diffusion for most simulation runs.

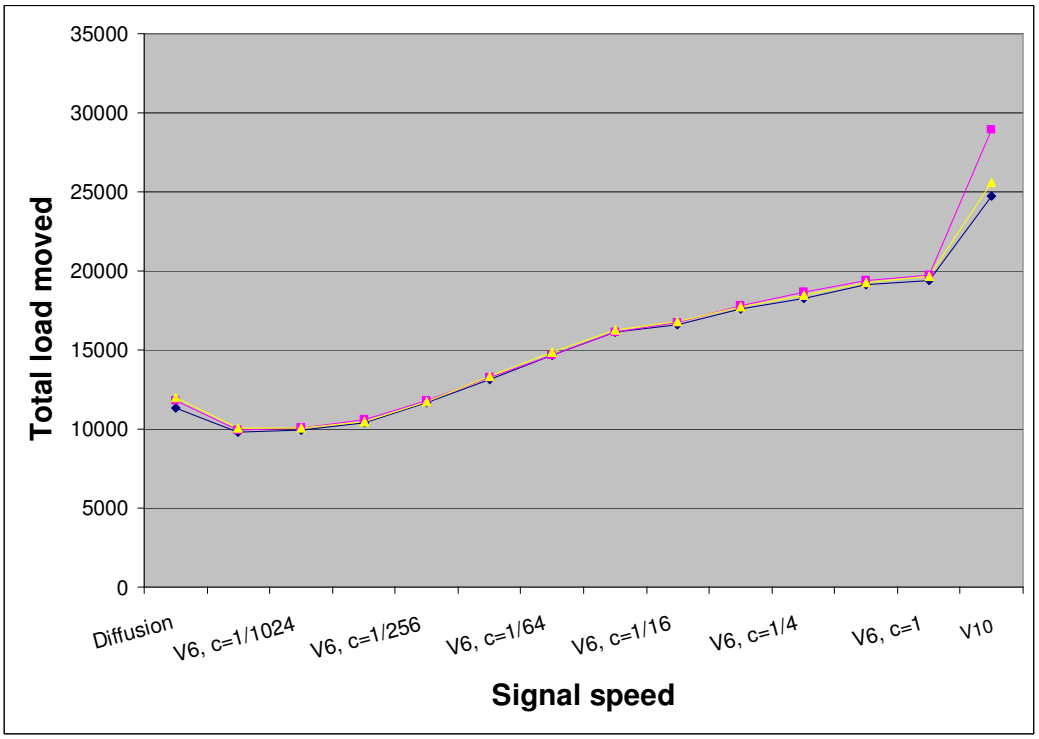


Figure 3: Total load moved to reach convergence as a function of signal speed. Each graph corresponds to a different instance of the random start distribution. Total load moved numbers for plain diffusion are plotted to the far left.

Several signal speeds produced reductions in convergence times of about 80%. (The shortest time to reach convergence for signal-aided diffusion was 1795 cycles; plain diffusion took about $13,000$ cycles to reach convergence.) Interestingly, the shortest convergence times were obtained when signal diffused at medium speeds. Signal-aided diffusion performed worse than plain diffusion when using our slowest signal speed (version 6 with $c_{\text{default}} = 1/2048$).

Figure 3 shows the total load moved for the simulation runs depicted in Figure 2. It is interesting to see if the use of a guiding signal not only reduces the time to reach convergence but also leads to a reduction in the total load moved. Figure 3 tells us that most signal-aided runs appear to move more load in total than does plain diffusion. Also, we can observe a clear indication of an increase in total load moved with increasing signal speed. Only signal-aided runs with slower signal speeds, ie, those using the version-6 algorithm with $c_{\text{default}} < 1/128$, resulted in less load moved, and even then the reduction was modest compared to the increases experienced when using faster signal speeds.

Chemotaxis-inspired load balancing addresses systems where the ability to move load is limited. Therefore, a key metric is the amount of load that is moved during a short interval, e.g., during a single cycle. Figure 4 shows the largest amount of load that was moved in a single cycle during each experiment. From Figure 4 we observe that loads guided by the speedier instances of the version-6 signal diffusion algorithm produced smaller values for the largest single-step load amount moved. The smallest maximum is 0.014. It is interesting to note the relatively poor performance resulting from using the version-10 algorithm to diffuse signal: the maximum load amount moved in a single cycle is about fifty times higher than for the best cases. Relative differences among different instances of the version-6 algorithm are also rather big (up to a factor of seven).

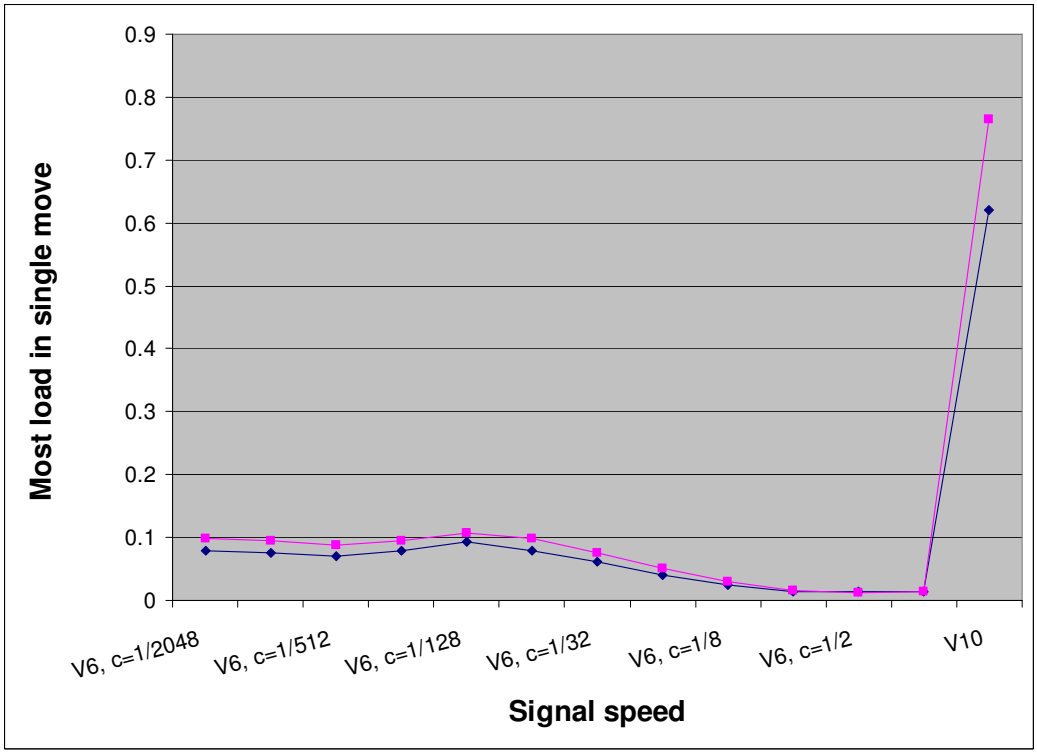


Figure 4: Largest load amount moved during any single cycle, over an entire simulation run, as a function of signal speed. Each graph represents a different instance of the random start distribution.

Though the results for largest single-step load amount moved are not as clear as the convergence time results, our tests suggest that performance gains are possible while still maintaining slow load movement. Also, it is worth noting that best results for the two-component algorithm, with version 6 for the signal, had a maximum link load moved that exceeded that for plain diffusion by a factor of only about 7. This may be contrasted with the factor 1000 increase for this same cost figure that we observed in the previous section (Figure 1), by increasing the *single-component* speed to the same value that *signal* uses in Figure 4. In other words (very roughly), by using chemotaxis instead of brute load speed, we obtain the same speed increase, but with much less penalty in maximum link load moved.

5 Nice properties

One of the main motivations for studying mechanisms from biology is that living systems exhibit "nice properties": they are self-organizing, self-repairing, adaptive, etc. It is of course of great interest to inquire whether engineered systems also have these nice properties. In this section we offer a limited discussion of this kind of question for our chemotactic load balancing system.

We note that both our reference diffusion algorithm and our chemotaxis algorithm are self-organizing systems, in that they make little reliance on global information—nodes only need updated information about their neighbors' load and signal values to carry out the algorithms. Our favored chemotaxis rule—ie, the one based on version-6 diffusion for signal—does rely on one global parameter, namely, the value for c_{default}. On the other hand, Figures 2 and 4 show that these two performance measures are rather insensitive to the value of c_{default}. (The total load moved is somewhat more sensitive.) In

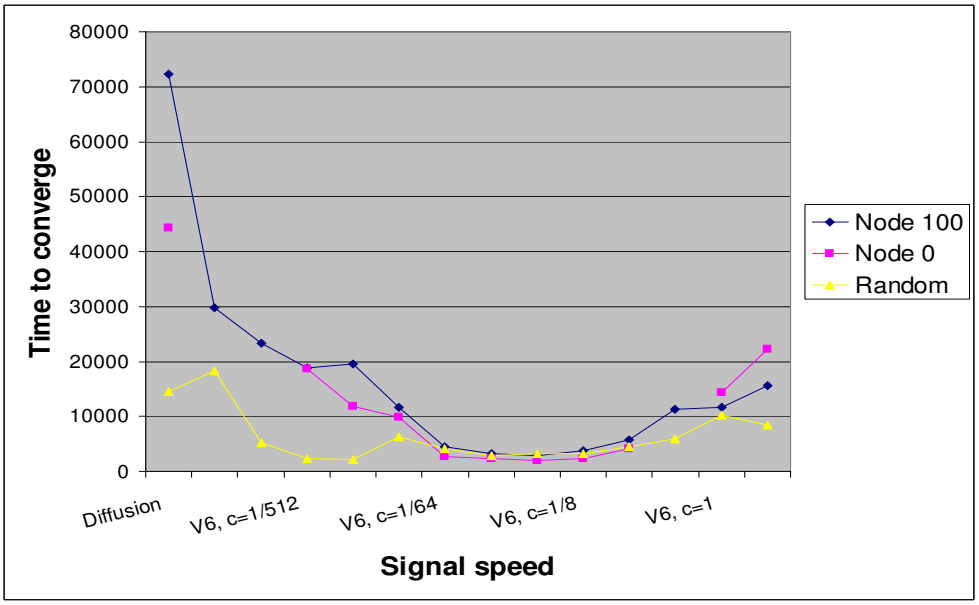


Figure 5: Sample runs showing time to reach convergence from different initial load distribution as a function of signal speed. Plain diffusion is plotted to the far left. Each graph represents a different start distribution. The three start distributions are: all load placed on a poorly connected node (node 100), all load placed on the best connected node (node 0), and random start distribution.

fact, we would argue that this type of insensitivity is in itself a kind of nice property. That is, insensitivity of performance to parameter values allows a self-organized system to find parameters giving good performance, without the necessity of fine tuning. Nevertheless we retain, as a goal for future work, the task of finding a satisfactory, decentralized method for tuning to a good value for c_{default}.

Figure 2 also shows, in a limited way, another type of insensitivity. In this figure, it is clear that version-6 chemotaxis (for most speeds) and plain diffusion are each less sensitive to the particular random start distribution than is version-10 chemotaxis. Figure 5 gives us more insight into the question of this type of sensitivity. In the blue curve of this figure, the initial distribution consists of all the load (10,000 units) being placed at a very poorly connectd node (node 100). The red curve is generated from a start distribution with all load placed at the best connected node (node 0). Finally, the yellow curve comes from a random start distribution, like that used in Figure 2.

Hence Figure 5 shows results for a considerably larger variation of initial distribution than Figure 2. We make three observations from Figure 5. First, version-10 chemotaxis is again more sensitive to initial distribution than is version-6 chemotaxis. Second, plain diffusion is *highly* sensitive to the initial load distribution. We see that the insensitivity and relatively good convergence times shown by plain diffusion for a random initial distribution are lost when we examine more extremely skewed start distributions. Finally, we feel that the insensitivity shown by the "good" range of version-6 chemotaxis is remarkable: its convergence time is virtually indepedent of start distribution, even in the face of such extreme variations. While we do not have a detailed explanation for this nice property, we would claim that it merits the term "adaptivity". In these words, we would say that (at least for this performance measure) version-6 chemotaxis is by far the most adaptive of the three systems studied.

Although most of our test runs have been done using a power-law topology, we have performed a limited series of tests with a random topology. This topology was generated using the Peersim plat-

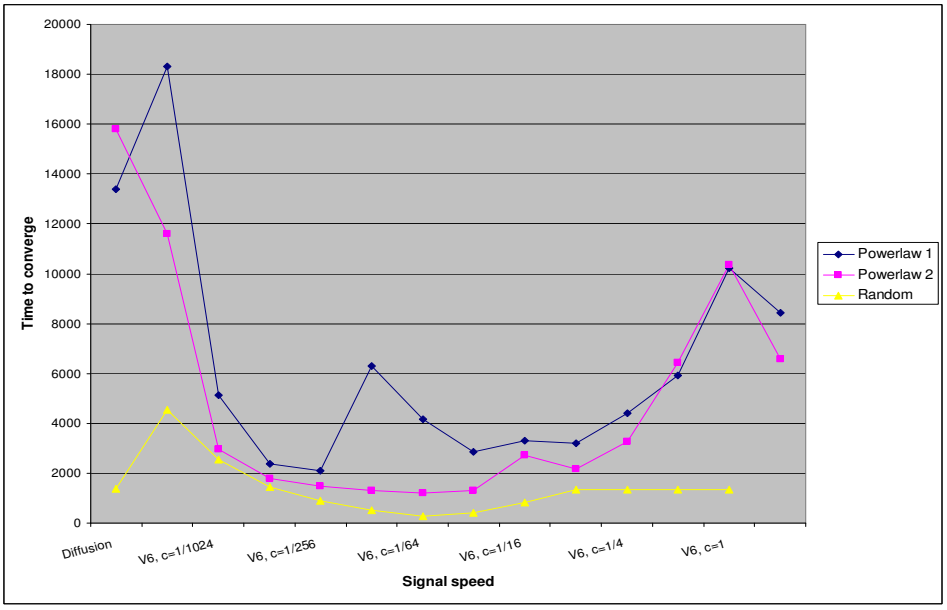


Figure 6: Sample runs showing time to reach convergence on different topologies as a function of signal speeds. Plain diffusion plotted on the far left. Two graphs correspond to different instances of a powerlaw topology, the third graph corresponds to a random topology.

form [22], with a constant node degree ($k = 20$). Figure 6 shows some results for convergence time, comparing the random topology with the power-law topology. This figure thus shows sensitivity (or its inverse, adaptivity) to (rather large) topology variation. The two upper curves are for two distinct instances of the power-law topology, while the bottom (yellow) curve is for the random topology. Start distributions were random, generated for both topologies in the same way as outlined in section 4.

Our observations from this figure are roughly like those from Figure 5. That is: large sensitivity (poor adaptivity) for plain diffusion, and also for version-10 chemotaxis; and relatively low sensitivity for version-6 chemotaxis.

It is not so surprising that plain diffusion adapts poorly to changes in topology or initial distribution. Plain diffusion is after all blind: it treats every "direction" on the network as equal. Intuition says that this blind approach should work best for (i) fairly uniform initial distributions, and (ii) a highly mixing topology such as the random topology; and Figures 5 and 6 are consistent with this. The news from these figures is then that our "smart" system is (at least, based on these tests) truly more adaptive than the blind system: it is only relatively weakly affected by the change from a random to a power-law topology, and virtually unaffected (in its best range) by changing from a fairly uniform start distribution to a highly skewed one.

6 Complex behavior

We devote this section to the illustration of a number of striking behaviors exhibited by our chemotaxis system. These behaviors are not adequately (or at all) revealed by our figures showing performance in sections 4 and 5. We do however see some signs of complex behavior in Figure 1. This figure supports

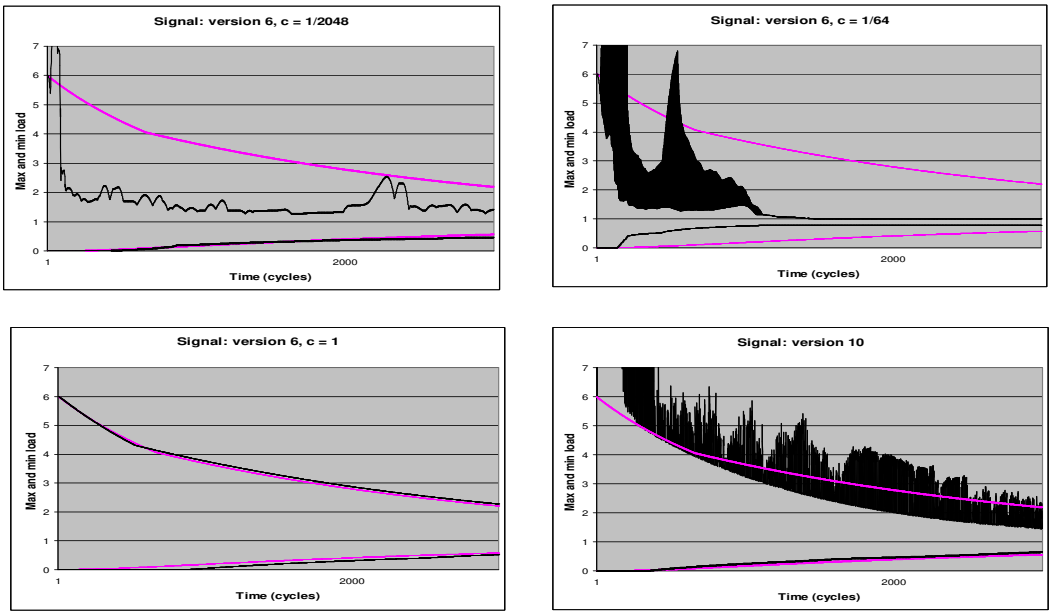


Figure 7: Sample runs showing approach to convergence (largest and smallest load values in the system under study) over time. Only the first 2500 cycles are plotted. Each sample run corresponds to a different signal speed; the chosen signal speeds are: version-6 with $c_{\text{default}} = 1/2048$, $c_{\text{default}} = 1/64$, and $c_{\text{default}} = 1$ in addition to version-10. The time behavior of plain diffusion is included for reference in each plot. All runs use the same power-law topology and random start distribution.

the intuitive notions that we extract from the work of Cybenko [11]. We see that slow systems (plain diffusion in the figure) are stable, while the fastest stable one-component systems (version 10 in the figure) are close to instability. Version-6 one-component diffusion then represents an intermediate case.

Given the fact that the one-component diffusive system, when discretized, can give rise to oscillations and instability, it is to be expected that the two-component system should do the same. In fact, our notion of two time scales corresponds, for continuous systems, to the notion of "stiff" coupled differential equations. It is known, for such systems, that coupling a "fast" system to a "slow", followed by discretizing, can lead to oscillations, and even instability. Our chemotaxis system is such a system—but one for which there is no opportunity to adjust the size of the time step (the common remedy for stiff systems). In this light, it is not surprising that the chemotaxis system exhibits oscillations. Nor is it surprising that the best performance, in terms of convergence, is obtained for intermediate (rather than maximal) values for the signal speed.

Nevertheless there are surprises in store from the study of the two-component system: the behavior is even more complex than our intuition (based on the above) would guess. Figure 7 gives a sample of the kinds of complex behavior we have observed for the chemotactic system. The four plots show the time behavior of the load (in the same way as in Figure 1) for increasing signal speed, reading left to right and top to bottom; also, each of the four plots show plain (one-component) diffusion in pink. The first plot ($c_{\text{default}} = 1/2048$) gave very slow convergence by our rather tight convergence criterion. However, here we see that this system moves extremely rapidly to within 1 unit of convergence, and that the long convergence time is due to the long-lasting, irregular behavior after this first, fast

plunge. Hence, a weaker convergence criterion would rank this value for c_{default} considerably higher in performance.

The second plot ($c_{\text{default}} = 1/64$) is taken from the "good" range of signal speeds. The approach to convergence is strikingly different from that in the first plot. In fact, the two systems are almost complementary: one ($c_{\text{default}} = 1/2048$) shows a *high* rate of approach to convergence when *far* from convergence, and a low rate when close; while, for the other ($c_{\text{default}} = 1/64$) the description is essentially reversed.

Yet another surprise is revealed by the third plot. Here we see behavior for the "fastest" ($c_{\text{default}} = 1$) version-6 chemotaxis—and yet we see no oscillations whatsoever. Furthermore, the behavior is nearly identical to that of unaided diffusion, with $c = 1/2500$. We have no explanation for this behavior. Finally, moving to the fastest signal diffusion rule (version 10), we see in the fourth plot that the oscillations have returned.

This section is purely descriptive—we do not try to explain (yet) the complex behaviors seen here. The main points from this section are then that (i) the coupled system (even leaving out the unstable cases!) can exhibit highly complex behavior over time, on the way to convergence; and (ii) the nature of this complex behavior varies, as a function of signal speed, in a way that is in itself highly complex.

7 Discussion and Conclusions

In this paper, we have presented a biology-inspired mechanism for improving the performance of basic diffusion in load balancing. Diffusion is a widely studied approach that offers advantages of being simple, decentralized, and flexible. Our goal has been then to see if a more active mechanism, taken from biology, could give better results than the passive physics-inspired diffusion approach.

The mechanism that we borrow from biology is *chemotaxis*—a system in which diffusing chemical signals guide the movement of the bodies emitting them. We have allowed the load at nodes on a network to emit a signal, which follows a fast diffusion law for its motion. The motion of the load itself is then guided by gradients of signal. We have used repulsive chemotaxis—load moving down the gradient—to drive the load towards a uniform distribution. We have emphasized that chemotaxis makes sense when there are *two time scales* present: the slow movement of the emitting bodies, coupled to the fast movement of the load. This idea may be technically useful when bandwidth constraints prevent the load from diffusing rapidly, while signal (a few bytes) may not be subject to such constraints.

We have implemented a reference (plain diffusion) algorithm, and (after some experimentation) settled on two algorithms for fast diffusion of signal. One of our fast algorithms in fact allows for the continuous tuning of signal speed; hence we have been able to sample a wide range of signal speeds, always holding the load speed to a fixed (low) value.

Results of our tests show that chemotaxis can give a large improvement in rate of convergence over plain diffusion. These tests were performed on a scale-free topology with a random start distribution. The same tests show that the total load moved by chemotaxis is typically somewhat higher than that for plain diffusion (lower for some signal speeds, but up to a factor of two higher for others). Note that our results, though covering a wide range of signal speed values, are only based on a small number of runs. More work is needed to increase the number runs, each with a distinct instance of scale-free topology and random start distribution.

We have also looked at the maximum value (over a run, and over all links) of the load moved over a single link. This quantity is also higher for chemotaxis, for all signal speeds. Hence, it is not clear that the advantages of chemotaxis will persist in a system with hard quotas on load movement over

the links. We intend to examine this question in future work. (In fact, we have begun such studies, but cannot yet draw any definite conclusions.)

Our best algorithm for chemotaxis ("version 6") has a tunable parameter (c_{default}) which controls the signal speed. This is a global parameter. Hence another item for future work is to find satisfactory ways for the system itself to tune c_{default} to a good working value.

We note in this context that our version-6 chemotaxis rule gives good performance over a rather wide range of c_{default} values. This kind of insensitivity is encouraging—it should make the task of decentralized self-tuning easier. Furthermore, we have observed other kinds of insensitivities for version 6: insensitivity to wide variations in start distribution, and to large differences in network topology. In fact, version-6 chemotaxis is clearly and unambiguously the best choice if this kind of insensitivity (or adaptivity) is important; both plain diffusion and version-10 chemotaxis showed much higher sensitivity to these two environmental perturbations.

The adaptivity that we see for version 6 is the kind of "nice property" that one hopes to achieve from decentralized, bio-inspired technological systems. We have noted the high adaptivity of version-6 chemotaxis. In particular, we find it remarkable that the convergence time for this rule, in the good range of c_{default} values, is virtually unaffected by extreme variations in start distribution (Figure 5)—from fairly uniform, to the most disadvantageously skewed distribution possible. We note here that all three systems compared in this figure (plain diffusion, version 6, and version 10) are decentralized, flexible, 'swarm-intelligence-like' systems. Hence, the argument that nice properties 'simply arise' in such systems cannot help to explain the clear difference in adaptivity between these three. We leave the understanding of this also to future work.

Our chemotaxis system involves the coupling of two discrete dynamical systems, each with its own natural time scale (speed), and typically with a large difference between these two time scales. One does not expect smooth behavior from such systems. In fact, even the one-component (diffusive) system, because it is discrete, is subject to instabilities if it is too fast, and to oscillations when it is stable but near the instability threshold (Figure 1). The two-component system shows highly complex behavior (Figure 7). The good news is that simple precautionary rules, as reported in Section 3, can ensure that stable convergence can be reliably achieved, in spite of the complexity. However, an understanding of the complex behavior exhibited by our two-component system is lacking.

We have mentioned several topics for future work. Besides those mentioned, we plan on exploring other topologies; on collecting more systematic and quantitative results on nice properties such as adaptivity; and on seeking an understanding of the variations in adaptivity that we observe.

One further, outstanding, direction to take is to move the present approach towards practical application. The present work represents a proof of principle: that signalling can strongly enhance the convergence speed of a diffusive load balancing approach. We are interested in seeing whether this principle will survive in practice—for example, in large, dynamic systems such as peer-to-peer, where adaptivity is important, and central control is not practical.

Finally, we note that we have thus far only dealt with *repulsive* chemotaxis. The idea of using attractive chemotaxis should also be explored for possible technological applications. We mention one, speculative, example. Suppose that a distributed file-sharing system could benefit (achieving more efficient searching) if the documents in the system were clustered, meaning that similar files were located close to one another in the network sense. One might then use attractive chemotaxis to bring about this clustering. Of course, since information is multidimensional, one would need several types of signals. In any case, we suggest that attractive chemotaxis is also worth looking into for accomplishing technological ends.

8 Acknowledgments

We thank Niloy Ganguly for numerous helpful discussions. This work was partially supported by the Future & Emerging Technologies unit of the European Commission through Project BISON (IST-2001-38923).

References

[1] K. D. Devine, E. G. Boman, R. T. Heaphy, B. A. Hendrickson, J. D. Teresco, J. Faik, J. E. Flaherty, and L. G. Gervasio. New challenges in dynamic load balancing. *Appl. Numer. Math.*, 52(2–3):133–152, 2005.

[2] A. Oram, editor. *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*. O'Reilly & Associates, Mar. 2001.

[3] I. Foster and C. Kesselman, editors. *The Grid: Blueprint for a Future Computing Infrastructure,*. Morgan Kaufmann, 1999.

[4] E. Lee. Highly-available, scalable network storage. In *40th IEEE Computer Society International Conference (COMPCON95)*, page 397, 1995.

[5] Z. Palkova, B. Janderova, J. Gabriel, B. Zikanova, M. Pospisek, and J. Forstova. Ammona mediates communication between yeast colonies. *Nature*, 390:532–536, 1997.

[6] BISON. http://www.cs.unibo.it/bison/index.shtml.

[7] L. Benov and I. Fridovich. E*scherichia coli* exhibits negative chemotaxis in gradients of hydrogen peroxide, hypochlorite, and N-chlorotaurine: Products of the respiratory burst of phagocytic cells. *Proc. Nat. Acad. Sc. US*, 93(10):4999–5002, 1996.

[8] J. C. Dallon, H. G. Othmer, C. v. Oss, A. Panfilov, P. Hogeweg, T. Höfer, and P. K. Maini. Models of D*ictyostelium discoideum* aggregation. In W. Alt, A. Deutsch, and G. Dunn, editors, *Dynamics of cell and tissue motion*, pages 193–202. Birkhäuser, Basel, 1997.

[9] P. N. Devreotes. *Dictyostelium discoideum*: A model system for cell-cell interactions in development. *Science*, 245:1054, 1989.

[10] Y. Jiang, H. Levine, and J. Glazier. Possible cooperation of differential adhesion and chemotaxis in mound formation of D*ictyostelium*. *Biophys. J.*, 75:2615–2625, 1998.

[11] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal of Parallel and Distributed Computing*, 7:279–301, 1989.

[12] J. Boillat. Load balancing and poisson equation on a graph. *Concurrency: Practice and Experience*, 2:280–313, 1990.

[13] T. Decker, B. Monien, and R. Preis. Towards optimal load balancing topologies. *Lecture Notes in Computer Science*, 1900:277–287, 2001.

[14] R. Diekmann, A. Frommer, and B. Monien. Efficient schemes for nearest neighbor load balancing. *Parallel Computing*, 25(7):789–812, 1999.

[15] Y. F. Hu and R. J. Blake. An improved diffusion algorithm for dynamic load balancing. *Parallel Computing*, 25:417–444, 1999.

[16] M. Jelasity, A. Montresor, and O. Babaoglu. A Modular Paradigm for Building Self-Organizing Peer-to-Peer Applications. In G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, editors, *Engineering Self-Organising Systems: Nature-Inspired Approaches to Software Engineering*, number 2977 in Lecture Notes in Artificial Intelligence, pages 265–282. Springer-Verlag, Apr. 2004.

[17] G. Karagiorgos and N. M. Missirlis. Accelerated diffusion algorithms for dynamic load balancing. *Inf. Process. Lett.*, 84(2):61–67, 2002.

[18] P. Sanders. On the efficiency of nearest neighbor load balancing for random loads. In *Parcella 96, VII. International Workshop on Parallel Proccessing by Cellular Automata and Arrays*, pages 120–127, 1996.

[19] R. Elsässer and B. Monien. Diffusion load balancing in static and dynamic networks. In *Proc. Internat. Workshop on Ambient Intelligence Computing*, pages 49–62, Dec. 2003.

[20] A. Deutsch, N. Ganguly, G. Canright, M. Jelasity, and K. Engø-Monsen. Models for advanced services in AHN, P2P Networks. Bison Deliverable, www.cs.unibo.it/bison/deliverables/D08.pdf, 2003.

[21] G. P. Jesi. Peersim howto: build a new protocol for the peersim simulation framework, November 2004.

[22] G. P. Jesi. Peersim howto 2: build a topology generator, 2004.

Elements about the Emergence Issue
A survey of emergence definitions

Joris Deguet*
Laboratoire Leibniz

Yves Demazeau
Laboratoire Leibniz

Laurent Magnin
Université de Montréal
Canada

*46 avenue Félix Viallet, 38031 Grenoble CEDEX, France, joris.deguet@imag.fr, tel: +33 4 76 57 48 07, fax: +33 4 76 57 46 02

Abstract

Emergence, a concept that first appeared in philosophy, has been widely explored in the domain of complex systems and is sometimes considered to be the key ingredient that makes "complex systems" "complex". Our goal in this paper is to give a broad survey of emergence definitions, to extract a shared definition structure and to discuss some of the remaining issues. We do not know of any comparable surveys about the emergence concept. For this presentation, we start from a broadly applicable approach and finish with more specific propositions. We first present five selected works with a short analysis of each. We then propose a merged analysis in which we isolate a common structure through all definitions but also what we think needs further research. Finally, we briefly describe some perspectives about the emergence engine idea also referred to as emergent engineering.

Keywords survey, emergence, complexity, levels definition

1 Introduction

Emergence, a concept that first appeared in philosophy [1, 2], has been widely explored in the domain of complex systems [3, 4, 5, 6, 7, 8, 9, 10] and is sometimes considered to be the "key ingredient that makes complex systems complex" [11].

On March 9th 2005, we made a basic one-keyword Internet search for "emergence" papers on computer science specific engines and generalist scientific engines. We retrieved impressive amounts of relevant documents:

Table 1

Search Engine	Number of results
ACM	648
IEEE	1450
CiteSeer	8257
ScholarGoogle	372000

From these, we chose to survey five works matching the following criteria:

- Emergence definition is the primary goal
- It contains a significantly different (and possibly contradictory) approach from other selected papers

We chose not to give any introductory example or vague intuition here because it might fall out of the scope of a particular approach. Our goal in this paper is to give a broad survey of emergence definitions, to extract a shared definition structure

and to discuss some of the remaining issues. We do not know of any comparable surveys about the emergence concept.

In this paper, we start from a broadly applicable approach and finish with more specific propositions. We first present five selected works with a short evaluation of each. We then propose a merged analysis in which we isolate a common structure through all definitions but also what we think needs further research. Finally, we briefly describe some perspectives about the emergence engine idea also referred to as emergent engineering.

2 Elements from existing definitions

2.1 Detection and emergence

2.1.1 Concept

The first idea about emergence we present is the work of Bonabeau and Dessalles [7]. As the title suggests, the authors give significant importance to the detection of the phenomenon in their proposition:

> "We propose here a conceptual framework, based on the notion of detection [...] Then we show that emergence is related to complexity shifts. Lastly, we propose to focus on the observer, rather on the emerging system, in order to show that all characterizations of emergence are implicitly connected to the notion of detection."

Given the two following notions:

detector defined as "any device which gives a binary response to its input"

relative complexity $C(S|D,T)$ of a system S "where D is a set of detectors and T a set of available tools that allow to compute a description of structures detected through D" which corresponds to the difficulty to describe the system given T and D.

Emergence happens when between time t and $t + \Delta t$, two events happen:

1. a detector D_k becomes activated
2. $C_{t+\Delta t}(S|T, D_1, \ldots, D_{k-1}, D_k) < C_t(S|T, D_1, \ldots, D_{k-1})$

This property is likely to happen in a hierarchy of detectors as they point out: "When a detector becomes active in such a hierarchy, the active detectors from the lower level that are connected to it can be omitted from the description.[...] Emergence is thus a characteristic feature of detection hierarchies."

2.1.2 Discussion

One widely shared feature of emergence definitions is the existence of levels. This definition is interesting because it defines emergence as internal to an observation device, that must be hierarchically organized. The authors do not assume levels *a priori* in the definition but show that this is a condition *sine qua non* for the complexity discontinuity to happen.

No assumption is made about the system under detection, therefore one can apply this criterion on both artificial and natural systems as long as detection is possible.

This defines a *low-to-high level emergence*.

2.2 The emergence test

2.2.1 Concept

The first definition focused on an observer modeled by a detection apparatus. This makes emergence somehow "subjective" as the complexity measure depends on this apparatus. However, once the observer is defined, emergence only depends on the perceived behavior. The emergence test introduces the consideration of the system's design in addition to its behavior, and therefore moves subjectivity out of the very domain of observation.

Explicitly inspired by Turing's test for *intelligence* [12], Ronald, Sipper and Capcarrère [13, 14] proposed to define an "emergence tag gun" instead of a formal definition.

This *emergence* test involves a system designer and a system observer (both of whom can in fact be one and the same). Then if the following three conditions hold, the emergence tag is conferred:

Design The system has been constructed by the designer by describing *local* elementary interactions between components in a language L_1

Observation The observer is *fully aware* of the design, but describes *global* behavior and properties of the running system, over a period of time, using a language L_2

Surprise The language of design L_1 and the language of observation L_2 are distinct, and the causal link between the elementary interactions programmed in L_1 and the behaviors observed in L_2 is *non-obvious* to the observer, who therefore experiences surprise. In other words, there is a cognitive dissonance between the observer's mental image of the system's design stated in

> L_1 and his contemporaneous observation of the system's behavior stated in L_2.

They describe this question as reposing on how easy it is for the observer to bridge the gap between L_1 and L_2.

2.2.2 Discussion

We think we can consider Bonabeau and Dessalles' D and T as words and syntax of an observation language L_2.

The introduction of the design language L_1 has two important consequences:

1. Emergence happens between the design and the observation. This defines a *design-to-behavior emergence*.
2. Existence of L_1 restricts the application of this criterion to artificial systems i.e. designed by the human hand.

Emergence happens when observation and design appear loosely coupled to the observer. Therefore, the result of ones "tag gun" might differ from another, and the resulting emergence is highly subjective.

This corresponds to Baas' *deducible emergence* [15] where two disjoint levels are linked by a computational process. Indeed, Baas defines Obs^2 (similar to L_2), the "new observational" mechanisms with respect to the observation mechanisms Obs^1 (that are part of L_1) used in the dynamics.

In the field of decentralized artificial intelligence, Demazeau and Müller [16] have made a similar distinction between *internal* and *external* descriptions of agents where internal description refers to the real architecture of an agent and external description refers to its externally perceived behavior.

2.3 Simulation emergence

2.3.1 Concept

Making the parallel between intelligence and emergence as subjective notions defined by tests can lead to controversy. One answer could be to consider that emergence happens when a large number of scientists agree that it does. Another answer is to make the definition objective. Simulation emergence is such an attempt, focused on the simulation domain.

In Darley [17] we find this definition:

> "A true emergent phenomenon is one for which the optimal means of prediction is simulation."

The author defines two means of prediction depending on n the size of a system:

- $s(n)$: the optimal "amount of computation required to simulate a system, and arrive at a prediction of the given phenomenon".
- $u(n)$: stands for "deeper level of understanding", the way we try to avoid computation by "a creative analysis", $u(n)$ is the amount of computation required by this method.

Then the system will be considered as emergent iff $u(n) \geq s(n)$ i.e. direct simulation is optimal relative to the "amount of computation" measure. When decomposed into "steps" the amount of computation is defined as the sum over steps of Kolmogorov complexities.

We link this definition with the *weak emergence* from Bedau [18, 19]:

> "A macro-state P of S with micro-dynamic D is weakly emergent iff P can be derived from D and S's external conditions but only by simulation. [...] for P to be weakly emergent, what matters is that there is a derivation of P from D and S's external conditions and *any* such derivation is a simulation."

2.3.2 Discussion

First, Bedau describes a new relation namely the *micro-to-macro* one, the macro-level being composed of micro-entities. We believe we can join this *micro-to-macro emergence* with the low-to-high one (cf 2.1.1) without a loss of sense.

The key issue is to understand what a simulation is. Among all the ways to derive the phenomenon in a computable manner, some are simulations, others are "shortcuts". Then optimality of simulation is equivalent to the absence of "shortcuts", this is why we decided to present the two definitions together.

Interpreted in the L_1 L_2 framework, this states an irreducible gap between the language of design L_1 and observation L_2 which is optimally filled by going in every details of the system's evolution (i.e. simulation). We note that the emergence "tag gun" used the size of the gap ("ease to bridge"), here the size itself does not matter.

An interesting point is that both authors address the question of emergence's decidability:

- In Bedau's formulation: "One might worry that the concept of weak emergence is fairly useless since we generally have no proof that a given macrostate of a given system is underivable without simulation."

- With Darley's words "Can we determine, for a given system, whether or not it is emergent ?".

Darley suggests that: "for any complex system which is capable of universal computation, we know that the best (only) means of prediction in such a situation is to *run the program* i.e. perform the simulation". Bedau notes that we usually "possess substantial *empirical* support" to assess it is so. Then, even if we have gained in objectivity, we might have encountered an undecidable criterion based on the simulation's definition.

If we reformulate as "the global behavior is optimally obtained by running a system made of interacting micro agents", it provides a natural way to apply the definition to multi-agent based simulations.

This definition might not apply out of the simulation domain[1].

2.4 Downward causation and emergence

2.4.1 Concept

Bedau has defined *weak* emergence with respect to the *strong* emergence based on *downward causation*. This view is illustrated by Timothy O'Connor [20]:

> "to capture a very strong sense in which an emergent's causal influence is irreducible to that of the micro-properties on which it supervenes; it bears its influence in a direct *downward* fashion, in contrast to the operation of a simple structural macro-property, whose causal influence occurs via the activity of the micro-properties which constitutes it."

In [21], Sawyer notes that:

> "In MAS and Alife social simulations, the emergent pattern is fully explained by the microsimulation; that is, reduced to an explanation in terms of agents and their interactions. Such reductionist assumptions imply that higher-level emergent patterns do not have any causal force."

In order to achieve *downward causation*, he proposes that:

1. "as in blackboard systems, the emergent frame must be represented as a data structure external to all of the participating agents"

[1] perhaps an adaptation to problem solving could be: emergent problems are "optimally" solved (resp. derived) by a decentralized system (resp. micro-dynamics' simulation)

2. "all emergent collective structures must be internalized by each agent, resulting in an agent-internal version of the emergent."

3. "This internalization process is not deterministic and can result in each agent having a slightly different representation."

2.4.2 Discussion

The question here is the possibility of *downward causation*.

We believe that L_1 and L_2 are of significant interest to clarify this issue. It sounds natural to us to consider that everything with causal powers in an artificial system lies in the L_1 design language as it must live within algorithm. Thus even if a data structure exists out of the agents at a macro level, it belongs to the design language. Then L_2 to L_1 causal power is impossible.

Until here we might have mixed design/observation with micro/macro as it is often the same: We conceive agents and we are very happy to show their collective behavior to colleagues. However, it can be interesting to distinguish the micro/macro from design/observation.

Sawyer's definition is based on the existence of a macro entity external to micro agents. This existence might provide causal powers to this entity on agents. Therefore it allows a *macro to micro causation* we can consider as *downward* as scale decreases. However, this is different from O'Connor's view as agents do not constitute the macro entity.

Existence of micro as well as macro entities implies they are part of the L_1 which makes the definition based on design only. This makes Sawyer's definition contradictory to Ronald ad al.'s emergence test as L_2 vanishes.

2.5 Grammar emergence

2.5.1 Concept

This last definition of emergence is specific as its scope is limited to systems expressed in a particular grammar model. This model provides intuitive definitions for micro/macro and design/observation levels.

Kubik [22] has proposed an approach based on "the whole is more than the sum of its parts" as inspiration and *Isometric Array Grammars* [23] as a modeling tool.

The key idea is to define a "whole" language and a "sum of the parts" language. From an initial *array* configuration, a language is obtained by rewriting using *isometric* production rules. For a given set of rules P_i, the corresponding language is noted $L(P_i)$.

We can sum up the proposal as follows:

$$\underbrace{L(\underbrace{\bigcup_i P_i}_{Whole})}_{Whole} \underbrace{\supset}_{More} \underbrace{superimposition_i}_{Sum} \underbrace{(L(P_i))}_{Parts}$$

We do not give the definition of the $superimposition$ operator here.

Emergence is the case of an array being in the whole language but not in the sum of parts. The first is obtained by putting all parts together and deriving configurations, the last by deriving configuration for every part separately and putting results together afterward. Putting together is the way we get a macro entity from micro ones, and derivation is the way to get the language (L_2) we observe from the rules (L_1) we designed.

2.5.2 Discussion

When someone hears "the whole is more than the sum of its parts", he or she might reply very fast that a system *is* composed of its parts and therefore cannot be more. To go beyond this triviality, Kubik's elegant idea is to switch micro/macro with design/observation. This makes things comparable as Kubik defines his gap between two set of arrays (similar to L_2 and a L_2'), at the observation level. Unfortunately, the definition is not so homogeneous as putting together is different for arrays and for rules. There is another drawback: Without restrictions on rules, it might be impossible to determine if an array is emergent.

Kubik's idea is close to an informal definition of emergence from [24] stated in the VOWELS framework [25] for multi-agent systems (MAS). This framework suggests a description of such systems as agents (A) in their environment (E), using interactions (I) forming an organisation (O). Then the pseudo equation from [24]:

$$MAS = A + E + I + O + Emergence$$

can be seen as:

$$\underbrace{L(MAS)}_{Whole} \underbrace{\supset}_{More} \underbrace{\sum_{v \in vowels}}_{Sum} \underbrace{(L(v))}_{Parts}$$

with VOWELS as an alternate micro partition of a macro MAS.

3 General framework

3.1 The minimal setting

We chose to survey very different works. However, the following setting is shared by most of emergence definitions:

1. something appears, it is a candidate to the title of *emergent*
2. it happens within the dynamics of a system
3. at least 2 levels/languages are distinguished
4. it satisfies a criterion that makes it an emergent

The first two points describe a system where something pops up, usually called a phenomenon.

The last point describes a criterion that defines the emergent subset of the larger set of things that pop up (we said the *phenomena*), this criterion uses the notion of levels (third point).

3.2 Open issues

Any precise definition requires refinements about the minimal setting. Most of the time, the refinements concern the definition of levels and what kind of criterion we define between them. We come back on these two points but first we want to clarify a prerequisite: The observation of the phenomenon.

3.2.1 Observation

The possibility to perceive the emergent phenomenon is not clear. Actually, we have to consider two issues, perceive the phenomenon and perceive its "emergenceness". We here focus on the phenomenon itself as its "emergenceness" depends on the chosen criterion.

If we consider a phenomenon P, we can wonder what ways we have to observe it. Bonabeau and Dessalles suppose we have a detector. For Ronald and al. the emergent phenomenon is the word of the L_2 language. Sawyer describes agent internalization which seems to be a way for the agent to perceive the phenomenon. Finally, Kubik's phenomena are words.

We can wonder what happens to computability. For example, can we consider that a phenomenon is a computable property of the system's trace? Furthermore, we might wonder if Church-Turing thesis makes the space of "any device which gives a binary response to its input" (cf 2.1.1) equivalent to the space of Turing machines.

Unfortunately, observation is not always clearly defined. This is important if we consider that emergents are a subset of observable phenomena.

3.2.2 Levels and downward causation

One of the main issues about emergence is to clarify what are the different levels in the system. We identified two principal conceptions:

- Design/Observation distinction [13] (close to internal and external descriptions from [16])
- Micro/Macro or Local/Global levels possibly structured into a hierarchy [7, 21]

In the following table, we summarize how these two distinctions are expressed in the presented works:

Table 2

Author	Micro/Macro	Design/Observation
Bonabeau and al.	Hierarchy	Observation only
Ronald and al.	L_1/L_2	L_1/L_2 ($L_1 \cap L_2 = \emptyset$)
Darley	Agents/Phenomenon	Agents/Phenomenon
Bedau	Micro-dynamics/Macro-state	Micro-dynamics/Macro-state
Kubik	Parts/Whole	Rules/Configurations
Sawyer	Agents/Emergent	Design only

One might ask: "Do we always design micro and observe macro?". The definition from Bonabeau and Dessalles do not deal with design. Sawyer claims a macro entity must exist but it is not clear if it must be artificial (and then designed). Kubik makes the distinction between the two relations but still the whole system is designed micro (as the union of rules) and observed macro.

Then we have a macro phenomenon. Based on where observation takes place, Müller [26] distinguished:

Strong emergence: "when the observer of the phenomenon is inside the system, endowing the phenomenon has causal powers." This is very close to Sawyer's emergence and certainly related to the idea of internal description, as the observation mechanism must be inside the system's entities.

Weak emergence: "when it is not, making it an *epi*phenomenon", which corresponds to Ronald and al.'s L_2 language excluding all the design and also to Forrest's definition of *emergent computing* [27].

Internal observation allows causal powers and we are back to the question of causation. Many philosophical works about emergence have stated "downward" causation has a key feature [28, 20]. The impossibility of such a feature is sometimes used to exclude emergence from the ken of artificial systems.

We have seen the definition of Sawyer's downward causation from a macro entity to micro ones. All these entities are part of the design language. Müller [26] suggested that this macro entity where macro phenomena leave their prints might be called the environment. This provides a multi-agent formulation where agents with reduced action/perception (micro) fields interact with a shared environment (macro).

However, this definition is weaker then O'Connor's who required the macro entity to be composed of the micro ones to assess downward causation; In this case we have one single system which can be seen as composed or as a whole. This small modification makes the levels completely different; it results in a radically different notion of emergence. Indeed "downward" causation depends a lot on what we mean by "up" level, "down" level and then "downward".

3.3 Criterion

We have some phenomena generated in a multi-level framework. Some of them are said emergent, according to a defining criterion. We have jointly discussed bidirectional causation and levels because of a direct dependence.

Bonabeau and Dessalles define emergence as a sudden concision of the system's description given by a detection apparatus. Their criterion is explicitly based on a complexity measure and emergence is an irregularity in this complexity's evolution during the system's activity. Ronald and al.'s criterion is surprise. We think we can reformulate this as "how complex it is to describe what we see with respect to some information", i.e. design information. This is interesting because Bonabeau and Dessalles describe emergence as a shift of such a complexity. Both definitions make emergence close to the notion of relative (to some information) descriptive complexity.

For Bedau [18], two criteria for emergent phenomena are:

- "Emergent phenomena are dependent on underlying processes."
- "Emergent phenomena are autonomous from underlying processes."

This autonomy seems difficult to define, especially for artificial systems, because the system runs as designed and its design is available. Autonomy for Bedau is the need for simulation, as simulation is the only way to predict. The terms "algorithmic effort" [18] or "amount of computation" [17] suggests that optimality is relative to some kind of time complexity. Therefore, they make emergence close to relative (to simulation) time complexity.

Kubik gives an alternative to such complexity considerations with a criterion based on a gap between languages. Although, his definition as a whole system

more powerful than the sum of its parts can be considered as a difference of generative power between systems, for a specific phenomenon (array), emergence is a binary criterion.

Finally, Sawyer's definition is based on the presence or absence of downward causation that is hardly a complexity issue or a gradual criterion.

The following table summarizes some properties of the criteria we have seen:

Table 3

Author(s)	Criterion	Binary/Gradual	Complexity
Bonabeau and al.	Complexity shift	Binary	Explicit
Ronald and al.	Surprise	Gradual	Implicit
Darley	$u(n)/s(n)$ Balance	Gradual	Implicit
Bedau	Simulation optimality	Binary	Implicit
Kubik	Set difference	Binary	No
Sawyer	Downward Causation	Binary	No

One problem is how far we can decide whether a given phenomenon is emergent or not (satisfies the criterion). For an observed phenomenon, can we decide of its *emergenceness*?

Bonabeau and Dessalles: The criterion is decidable as far as we have access to the complexity measures before and after a detector's activation.

Ronald and al.: No decidability assumption is made about surprise.

Bedau and Darley: Optimality of simulation might be impossible to decide; usually, empirical support exists.

Kubik: No assumption made on decidability for the two languages.

Sawyer: Causation of a macro phenomenon on micro entities might decidable if the micro/macro is well defined and causation is given a decidable definition.

4 Conclusion and perspectives

With this survey, our goal was to identify a "computer science" emergence definition (the reader interested in a more philosophical approach might consult [29, 30, 31]). We have isolated a minimal setting, small as definitions are significantly different. These differences might fit more or less your intuition of emergence.

By going through these definitions, we have noticed that emphasis is usually put on the criterion proposed. However, for a computational definition, we think the following points should be refined:

- How do we apply levels on existing systems?
- Can we tag a phenomenon as emergent in a computable way?

We might also explore to what extent a specific definition of emergence is linked with definitions of self-organization or complexity and other terms we usually meet in the field of complex systems.

Nonetheless, the reason we wanted a computer definition is the "much from little" idea that Holland has associated to emergence [32]. Then a lazy computer engineer would certainly be emergentist to work little for a great result. Moreover, if little is all we can do, emergence could be a way to go beyond our limits. Thus emergent engineering sounds like an appealing research track.

This idea is already present in [13] and [26]. We can also refer to the "New Emergent World models Through Individual, Evolutionary and Social learning" (NEW TIES) project, the idea of "Emergent Intelligence" from [33] or the ADELFE methodology [34].

In the future, we hope to progress in this direction by using insights provided by definitions and mechanisms suggested by widely accepted emergence examples (social animals, markets), and Holland's inspiration as a goal.

References

[1] Mill JS. System of Logic. John W. Parker; 1843.

[2] Lewes GH. Problems of Life and Mind. Trubner and Company; 1874.

[3] Steels L. Towards a theory of emergent functionality. In: From Animals to Animats SAB; 1992. .

[4] Deneubourg JL, Theraulaz G, Beckers R. Swarm-made architectures. In: Toward a practice of autonomous systems; 1992. .

[5] Gilbert N. Emergence in social simulation. In: Artificial societies: The computer simulation of social life. UCL Press; 1995. .

[6] Bonabeau E, Dessalles JL, Grumbach A. Characterizing emergent phenomena. Revue Internationale de Systémique 1995;.

[7] Bonabeau E, Dessalles JL. Detection and Emergence. Intellectica 1997;2(25).

[8] Bonabeau E, Theraulaz G, Deneubourg JL, Franks NR, Rafelsberger O, Joly JL, et al. A model for the emergence of pillars, walls and royal chambers in termite nests. Philosophical Transactions: Biological Sciences 1998;.

[9] Barabasi AL, Albert R. Emergence of scaling in random networks. Science 1999;.

[10] Deneubourg JL, Lioni A, Detrain C. Dynamics of aggregation and emergence of cooperation. Biol Bulletin 2002;.

[11] Standish RK. On Complexity and Emergence. Complexity International 2001;.

[12] Turing AM. Computing Machinery and Intelligence. Mind 1950;.

[13] Ronald E, Sipper M, Capcarrère MS. Design, observation, surprise! A test of emergence. In: Artificial Life 5; 1999. p. 225–239.

[14] Ronald E, Sipper M. Surprise versus unsurprise: Implications of emergence in robotics. Robotics and Autonomous Systems 2001;.

[15] Baas NA. Emergence, Hierarchies and Hyperstructures. In: Artificial Life 3; 1994. .

[16] Demazeau Y, Muller JP. From reactive to intentional agents. In: Decentralized Artificial Intelligence 2. Elsevier; 1991. .

[17] Darley V. Emergent Phenomena and Complexity. In: Brooks R, Maes P, editors. Artificial Life 4; 1994. p. 411–416.

[18] Bedau MA. Weak Emergence. Philosophical Perspectives: Mind, Causation and World 1997;.

[19] Bedau MA. Downward causation and the autonomy of weak emergence. Principia 2001;.

[20] O'Connor T. Emergent properties. American Philosophical Quaterly 1994;.

[21] Sawyer RK. Simulating emergence and downward causation in small groups. In: MABS; 2001. .

[22] Kubik A. Toward a formalization of emergence. Artificial Life 2003;9.

[23] Dassow J, Freund R, Paun G. Cooperating array grammar systems. International Journal of Pattern Recognition and Artificial Intelligence 1995;.

[24] Demazeau Y. VOYELLES. CNRS; 2001.

[25] Demazeau Y. Steps towards Multi-Agent Oriented Programming. In: 1st International Workshop on Multi Agent Systems; 1997. .

[26] Muller JP. Emergence of Collective Behaviour and Problem Solving. In: ESAW; 2003. .

[27] Forrest S. Emergent Computation. MIT Press; 1990.

[28] Kim J. Downward causation. In: Emergence or reduction? Essays on the prospects of nonreductive physicalism; 1992. .

[29] Ali SM, Elstob CM, Zimmer RM. The Question Concerning Emergence: Implications for artificiality. In: Computing Anticipatory Systems: CASYS - First International Conference; 1998. .

[30] Stephan A. Emergenz: Von der Unvorhersagbarkeit zur Selbstorganisation. Dresden University Press; 1999.

[31] Kim J. Making sense of emergence. Philosophical Studies 1999;.

[32] Holland JH. Emergence: From Chaos to Order. Perseus Books; 1997.

[33] Angeline PJ. Advances in Genetic Programming. MIT Press; 1994. .

[34] Bernon C, Gleizes MP, Peyruqueou S, Picard G. ADELFE, a Methodology for Adaptive Multi-Agents Systems Engineering. In: ESAW 2002; 2002. .

The POEtic Electronic Tissue and its Role in the Emulation of Large-Scale Biologically Inspired Spiking Neural Networks Models

Short title: POEtic tissue

J. Manuel Moreno[1], Yann Thoma[2], Eduardo Sanchez[2], Jan Eriksson[3], Javier Iglesias[3,4], Alessandro Villa[4]

[1]Technical University of Catalunya, Dept. of Electronic Engineering, Barcelona, Spain, moreno@eel.upc.edu
[2]Swiss Federal Institute of Technology Lausanne, Logic Systems Laboratory, Lausanne, Switzerland, {Yann.Thoma, Eduardo.Sanchez}@epfl.ch
[3]Laboratory of Neuroheuristics, Information Systems Department INFORGE, University of Lausanne, Lausanne, Switzerland, jan@lnh.unil.ch, Javier.Iglesias@unil.ch
[4]Laboratory of Neurobiophysics, INSERM U318, University Joseph-Fourier, Grenoble, France, Alessandro.Villa@ufj-grenoble.fr

Corresponding author:
Juan Manuel Moreno Arostegui
Technical University of Catalunya
Department of Electronic Engineering
Campus Nord, Building C4
c/Jordi Girona 1-3
08034 – Barcelona – Spain
E-mail: moreno@eel.upc.edu
Phone: +34 93 401 56 91
Fax: +34 93 401 67 56

Keywords: Artificial tissue, phylogenesis, ontogenesis, epigenesis, learning, spiking neural networks models, STDP rule, programmable hardware, POEtic.

Abstract:
One of the major obstacles found when trying to construct artefacts derived from principles observed in living beings is the lack of actual dynamic hardware with autonomous capabilities. Even if programmable devices offer the possibility of modifying the functionality implemented in the device, they rely on external hardware and software elements to provide its physical configuration. In this paper we shall present a new family of electronic devices, called POEtic, whose architecture has been derived from the basic properties that can be extracted from the three major organisation principles present in living beings: phylogenesis, ontogenesis and epigenesis. We shall demonstrate that the capabilities present in these new programmable devices make them an ideal candidate for the real-time emulation of large-scale biologically inspired spiking neural networks models.

1. Introduction

Even if there is a huge variability in the external features and functions associated with the living beings we can observe on the earth, their organisation is driven by principles that can be grouped around three main axes:

Phylogenesis: Also called evolution, it includes all the mechanisms that, driven by the pressure posed by nature, permit to determine the genetic information for a population of individuals that best fits to a given environment.
Ontogenesis: Ontogenetic mechanisms permit the development of a single individual driven by the information contained in its genome. Apart from developmental capabilities, self-replication and self-repair (what for most living beings means healing abilities) constitute clear examples of ontogenetic processes.
Epigenesis: It includes all the mechanisms that permit a single individual to efficiently interact with its direct environment. Epigenetic mechanisms include those plasticity-oriented processes that, driven by a sensor-actuator loop, permit an organism to modify its internal structure or its behaviour in order to adapt to the specific conditions present in a given environment at any time. Examples of biological subsystems showing epigenetic principles can be found in the central nervous system of mammals and in the immune system.

Taking inspiration from these organisation principles, the main goal of the POEtic project was the development of a flexible hardware substrate showing the basic features that permit living beings to show evolutionary, developmental or learning capabilities. The hardware substrate, in the form of a new electronic device, should permit the construction of electronic tissues able to solve tasks where these bio-inspired features represent a clear advantage over classical techniques.

The paper is organised as follows: In the next section we shall present the overall organisation of the POEtic tissue, describing the details of its main constituent parts. Then we shall introduce the features of a new learning model for spiking neural networks models that, when used in large-scale networks, shows interesting feature extraction capabilities. Once physically implemented into the POEtic devices, it will be demonstrated that these provide an efficient prototyping instrument for neuroscience research. The paper will finish presenting the conclusions and our current work.

2. Overall organisation of the POEtic tissue

The POEtic tissue is organised as a homogeneous bi-dimensional array of POEtic chips, each

one of them being able to implement a given number of cells as required by the application to be handled. The organisation of a single POEtic chip is presented in figure 1.

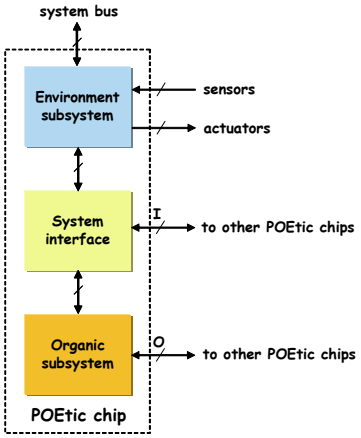


Figure 1. Organisation of a POEtic chip

From a structural point of view the organisation of a POEtic chip is divided in three main sections: the environment subsystem, the organic subsystem and the system interface. The environment subsystem is in charge of managing the interactions with the environment, and also of implementing the phylogenetic mechanisms of the tissue. The organic subsystem manages the physical realisation of the epigenetic and ontogenetic processes to be exhibited by the tissue. Finally, the system interface takes care of the efficient communication between these two subsystems. It also provides the mechanisms that permit the tissue to exhibit scalable properties.

The overall organisation of the resulting tissue is depicted in figure 2.

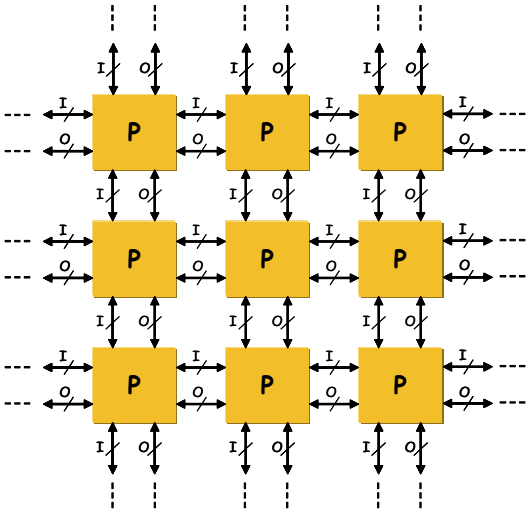


Figure 2. Overall organisation of the POEtic tissue

The squares in figure 2 represent POEtic chips, so that the sample tissue represented in the figure is constituted by 9 POEtic chips (the squares labelled as P) organised as a 3x3 matrix. As it can be deduced from the figure, the local communication between chips is separated in two different sections. The bidirectional lines labelled as I represent those connections associated with the system interface, while the bidirectional lines labelled as O indicate the connections established between the organic subsystems included in every chip. As it will be explained later the connections corresponding to the system interface provide the scalability features required

by the POEtic tissue, meaning that it can be constituted by as many chips as required by the actual application to be tackled. The connectivity between the organic subsystems is established at the routing plane level, and they allow for an effective communication mechanism between cells that are physically implemented in different chips.

Even if the POEtic tissue may be constructed from an arbitrary number of POEtic chips, each of them with their own functional subsystems, the system interface and the choice of the system bus makes it possible to handle the final tissue as a single POEtic chip. The only difference between a single chip and a tissue is the actual size of the organic subsystem, which in the later case is an aggregation of all the organic subsystems present in the tissue.

2.1. The environment subsystem

Figure 3 shows the internal organisation of the environment subsystem.

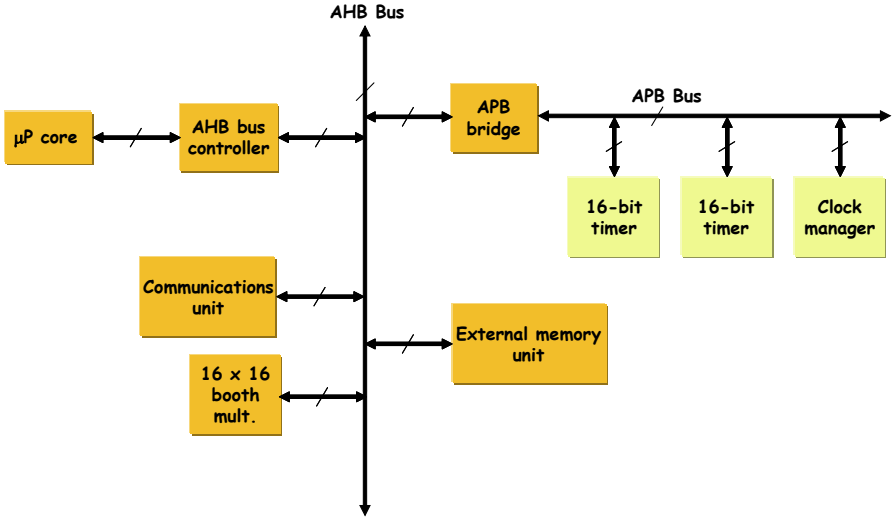


Figure 3. Internal organisation of the environment subsystem

As it can be deduced from figure 3 the architecture of the environment subsystem is structured around a specific microprocessor core. It is a 32-bit custom RISC processor, with dedicated instructions for developing evolutionary algorithms. A pseudo-random number generator is included in the ALU of the processor. The organisation of the environment subsystem is organised around the AHB (Advanced High-performance Bus) bus corresponding to the AMBA specification [1]. Simple peripherals are placed in a separate bus section, called APB (Advanced Peripheral Bus) that interfaces with the AHB bus through a bridge.

All the subsystems included in the POEtic tissue can be managed by the environment subsystem through a careful design of its memory map, whose structure is presented in table 1. The numbers provided in table 1 are specified in hexadecimal format. Even if the organic subsystem of the POEtic tissue is mapped in only one memory section, in fact this section maps the organic subsystems of all the chips that are present in the tissue for a given application.

The first 25 words of the program data section are reserved for the interrupt vectors of the microprocessor. Table 2 summarises the organisation of this interrupt vector table. The content of each of these memory positions is a JUMP instruction that points to the start address of the corresponding interrupt service routine.

The priority of the interrupt sources is directly related to the value of its associated interrupt vector, being thus the internal interrupt 0 the interrupt source with the highest priority.

Section	Start address	End address
Program	0x0000_0000	0x3FFF_FFFF
Data	0x4000_0000	0x7FFF_FFFF
Multiplier	0xC000_0000	0xC000_0003
Communications unit	0xD000_0000	0xD000_0150
Timers	0xE000_0000	0xE000_0007
Clock manager	0xE000_0008	0xE000_000F
Organic subsystem	0xF000_0000	0xFFFF_FFFF

Table 1. Memory map organisation of the POEtic tissue

Interrupt source	Interrupt vector
Main program	0x0000_0000
Timer 0	0x0000_0001
Timer 1	0x0000_0002
Multiplier	0x0000_0003
Clock manager	0x0000_0004
UART 0 TX	0x0000_0008
UART 0 RX	0x0000_0009
UART 1 TX	0x0000_000A
UART 1 RX	0x0000_000B
I2C	0x0000_000B
SPI	0x0000_000C
Parallel port	0x0000_000D
External interrupt 1	0x0000_0010
External interrupt 0	0x0000_0018

Table 2. Organisation of the interrupt vector table of the microprocessor

The communications unit included in the environment subsystem permits to implement an 8-bit bi-directional port, two UARTs, one SPI interface and one I2C interface. The functionality of these interfaces can be programmed by the user to match the requirements of a given application.

The clock manager unit has been added to the environment subsystem of the POEtic tissue in order to facilitate the hardware debugging procedures for the functionality implemented in the organic subsystem. This unit permits to generate a clock signal for the organic subsystem whose frequency is divided with respect to that associated to the system clock. Furthermore, if desired, this unit permits also to stop the clock signal provided to the organic subsystem after a specified number of clock cycles (from 1 to 65535). This feature allows for advancing the state of the organic subsystem edge-by-edge and then observing it (note that the environment subsystem has access through the system interface to the configuration and state of the organic subsystem).

From an architectural point of view the organisation of the external memory unit of the environment subsystem is divided in three main parts: the boot ROM, the program ROM and the data RAM.

The presence of a boot ROM section permits the user to load upon a power up sequence a program that may be transferred to the microprocessor using any one of the peripherals included in the communications unit. This means that the physical architecture of the memory unit of the microprocessor has two possible configurations, as depicted in figure 4.

The organisation depicted in figure 4(a) corresponds to a situation where the program to be executed by the microprocessor is fixed and already stored in a ROM. In this case after the power up sequence the microprocessor starts executing directly from this memory section. Figure 4(b) shows an organisation corresponding to a case where there is just a boot loader program stored in a boot ROM that takes care of capturing through one of the peripherals included in the communications unit the actual program to be executed by the microprocessor. This program is stored in the program ROM section that is physically implemented by means of a Flash or a SRAM unit. In order to permit the microprocessor to physically write the program Rom section during this boot sequence the memory map is slightly changed, so that the program ROM section is mapped in the memory area starting at address 0x6000_0000.

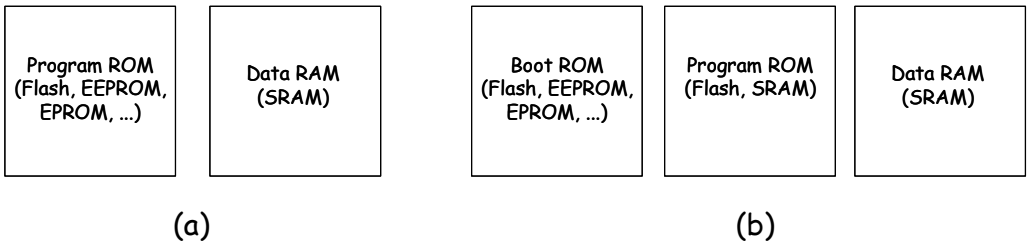


Figure 4. Physical architecture of the external memory unit of the environment subsystem

2.2. The organic subsystem

The organic subsystem is made up of 2 layers, as depicted in figure 5: a two-dimensional array of basic elements, called molecules, and a two-dimensional array of routing units. Each molecule is connected to its four neighbours in a regular structure. Mainly containing a 16-bit look-up table (LUT) and a flip-flop (DFF), it has the capability of accessing the routing layer that is used for inter-cellular communication. This second layer implements a dynamic routing algorithm allowing the creation of data paths between cells at runtime.

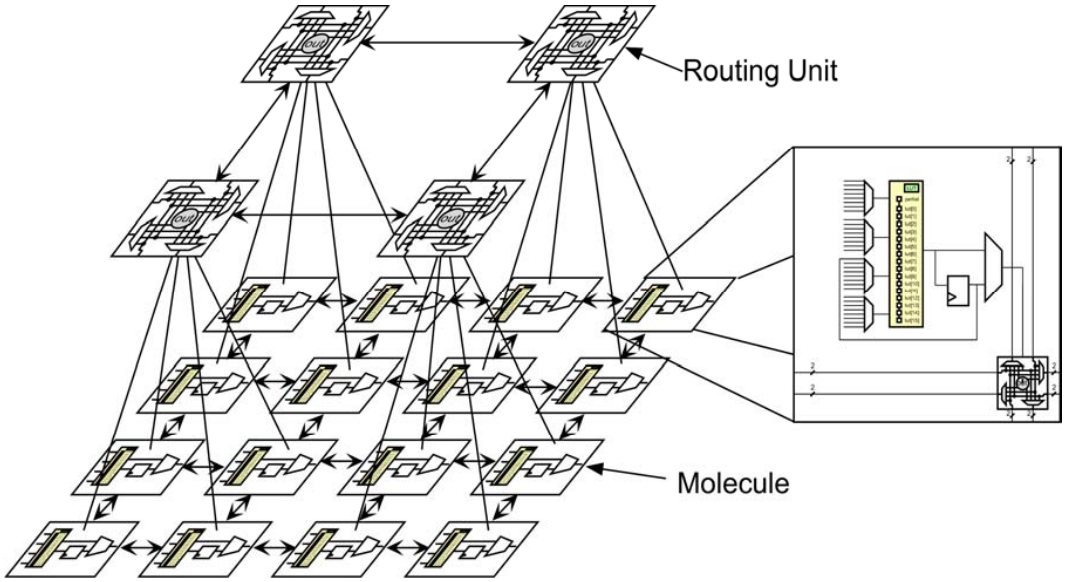


Figure 5. Organisation of the organic subsystem

A molecule is the smallest programmable element of the POEtic tissue. It is mainly composed

of a flip-flop (DFF), and a 16-bit look-up table (LUT) (figure 5). Eight modes of operation are supplied to ease the development of applications that need cellular systems and/or growth and self-repair. The LUT is composed of a 16-bit shift register that can be split in two, used as a shift register, or as a normal look-up table.

A molecule has eight different operational modes, to speed up some operations, and to use the routing plane. The functional modes provided for the molecules are the following:

- In **4-LUT** mode, the 16-bit LUT supplies an output, depending on its four inputs.
- In **3-LUT** mode, the LUT is split into two 8-bit LUTs, both supplying a result depending on three inputs. The first result can go through the flip-flop, and is the first output. The second one can be used as a second output, and is directly sent to the south neighbor (can serve as a carry in parallel operations).
- In **Comm** mode, the LUT is split into one 8-bit LUT, and one 8-bit shift register. This mode could be used to compare a serial input data with a data stored in the 8-bit shift register.
- In **Shift Memory** mode, the 16 bits are used as a shift register, in order to store data, for example a genome. One input controls the shift, and another one is the input of the shift memory.
- In **Input** mode, the molecule is a cellular input, connected to the inter-cellular routing plane. One input is used to enable the communication. When inactive, the molecule can accept a new connection, but won't initiate a connection. When active, a routing process will be launched at least until this input connects to its source. A second input selects the routing mode of the entire POEtic tissue.
- In **Output** mode, the molecule is a cellular output, connected to the inter-cellular routing plane. One input is used to enable the communication. As in Input mode, when inactive the molecule can accept a new connection, but won't initiate a connection. When active, a routing process will be launched at least until this output connects to one target. Another input supplies the value sent to the routing plane, as so to another cell.
- In **Trigger** mode, the 16-bit shift register should contain "000...01" for a 16-bit address system. It is used by the routing plane to synchronize the address decoding during the routing process. One input is a circuit enable, that can disable every DFF in the tissue, and another one can reset the routing plane, and so start a new routing.
- In **Configure** mode, the molecule can partially configure its neighborhood. One input is the configuration control signal, and another one is the configuration shifting to the neighbors.

Long distance inter-molecular communication is possible by the way of switch boxes. Each switch box consists of eight input lines (two from each cardinal direction) and eight corresponding output lines, and is implemented with eight inputs multiplexers. Two outputs are sent into each of the four neighbors of the molecule, as shown in figure 6.

Each output line can be connected to one of the six input lines from the other cardinal directions (no u-turns allowed) or to one of two possible outputs of the molecules (the output or the inverted output).

A molecule is defined by 75 configuration bits. They are configured by loading them in parallel, from the micro-controller. A partial reconfiguration is also possible, a molecule being able to shift configuration bits of its neighbourhood. Actually, when shifting, 76 bits are used, as the value of the flip-flop has to be in the configuration chain, in order to be able to retrieve its value.

The configuration system of the molecules can be seen as a shift register of 76 bits split into 5 blocks: the LUT, the selection of the LUT's input, the switch box, the mode of operation, and an extra block for all other configuration bits. Each block contains, as shown in figure 7, together

with its configuration, one bit indicating, in case of a reconfiguration coming from a neighbour, if the block has to be bypassed. This bit can only be loaded from the micro-processor, and remains stable during the entire lifetime of the organism.

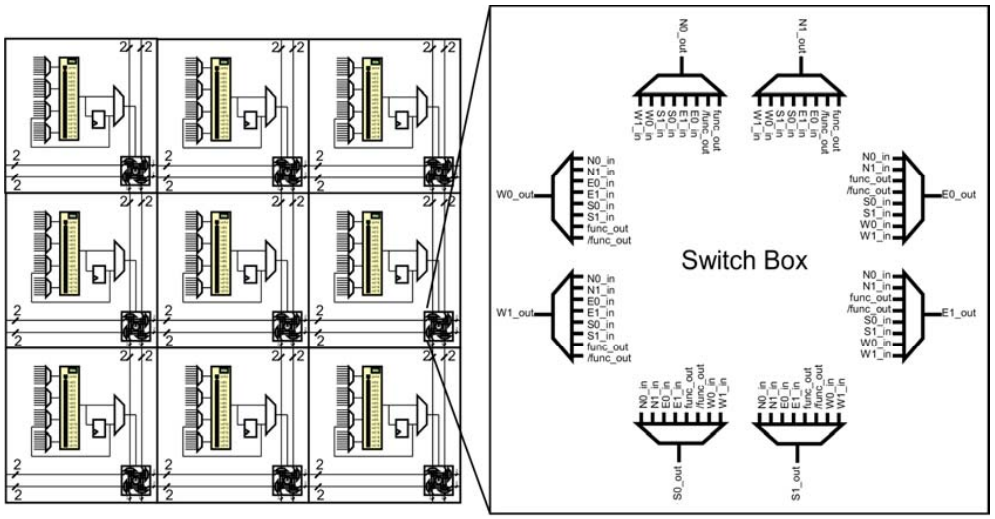


Figure 6. Nine molecules, connected through their switchboxes, and detailed view of a switchbox

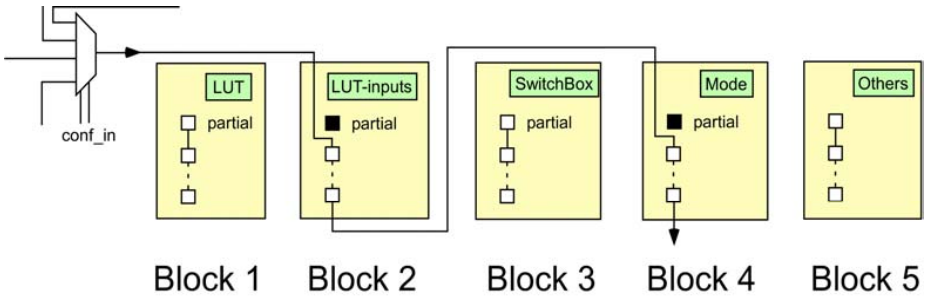


Figure 7. Organisation of the configuration bits for partial reconfiguration

The special configure mode allows a molecule to partially reconfigure its neighbourhood. It sends bits coming from another molecule to the configuration of one of its neighbours. By chaining the configurations of neighbouring molecules, it is possible to modify multiple molecules at the same time, allowing, for example, the synaptic weights in a neuron to be changed.

Three configuration bits are used to define the possible origin of a partial reconfiguration: two bits for selecting the origin, and one bit that enables the partial configuration. In case of a neighbor tries to partially reconfigure the molecule, if this config_partial_enable bit is set to ’1’, then the molecule is partially reconfigured, and it tries to partially reconfigure its neighbors, by chaining the output of the configuration stream. If the config_partial_enable bit is set to ’0’, then no partial reconfiguration is executed, and no signal is sent to the neighbors.

This partial reconfiguration allows for instance to use the configuration bits of a molecule to store information. A maximum of 54 bits can be stored in only one molecule, allowing for efficiently implementing genome storage. By modifying the LUT content, a cell can also modify its behaviour, that is a useful feature for evolvable hardware.

The second plane of the organism subsystem implements a dynamic routing algorithm to allow the circuit to create paths between different parts of the molecular array. The possibility of having a pseudo-static routing has also been added, to ease the development of applications that only need local connections between cells.

The dynamic routing system is designed to automatically connect the cells' inputs and outputs. Each output of a cell has a unique identifier, at the organism level. For each of its inputs, the cell stores the identifier of the source from which it needs information. A non-connected input (target) or output (source) can initiate the creation of a path by broadcasting its identifier, in case of an output, or the identifier of its source, in case of an input. The path is then created using a parallel implementation of the breadth-first search algorithm. When all paths have been created, the organism can start operation, and execute its task, until a new routing is launched, for example after a cell addition or a cellular self-repair.

Our approach has many advantages, compared to a static routing process. First of all, a software implementation of a shortest path algorithm, such as Dijkstra's [2], is very time-consuming for a processor, while our parallel implementation requires a very small number of clock cycles to finalize a path. Secondly, when a new cell is created it can start a routing process, without the need of recalculating all paths already created. Thirdly, a cell has the possibility of restarting the routing process of the entire organism, if needed (for instance after a self-repair). Finally, our approach is totally distributed, without any global control over the routing process, so that the algorithm can work without the need of the central micro-processor.

Every routing unit is composed of a switch box and a finite state machine. The switch box contains five multiplexers that can select the value sent to each of the four neighbors, and to the molecules underneath. The state machine is responsible for correctly configuring the multiplexers, and implements the distributed routing algorithm, by communicating with the other routing units.

The routing algorithm is executed in four phases:

Phase 1: Finding a Master
In this phase, every target or source that wants to and is not connected to its correspondent partner tries to become master of the routing process. A simple priority mechanism chooses the most bottom-left routing unit to be the master, as shown in figure 8. Note that there is no global control for this priority, every routing unit knowing whether or not it is the master. This phase is over in one clock cycle, as the propagation of signals is combinational.

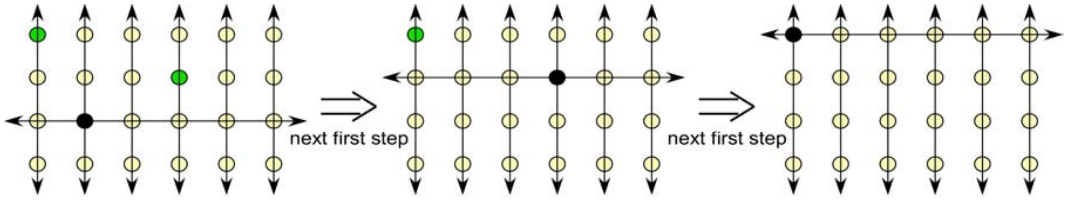


Figure 8. Three consecutive steps of the routing algorithm. The black routing unit will be the master, and therefore will perform its routing

Phase 2: Broadcasting the Address
Once a master has been selected, it sends its address in case of a source, or the address of the needed source in case of a target. It is sent serially, in `n` clock cycles, where `n` is the size of the address. The same path as in the first phase is used to broadcast the address, as shown in figure 9.

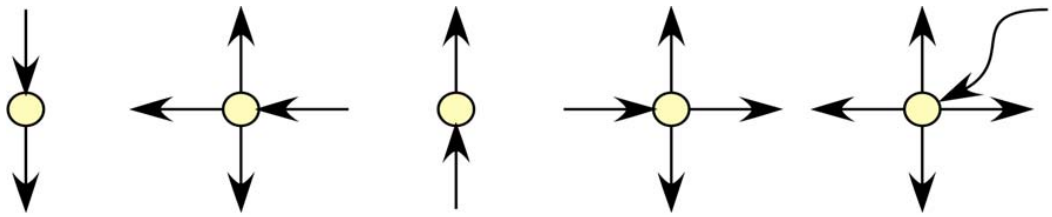


Figure 9. The propagation direction of the address: north → south | east → south, west, and north | south → north | west → north, east, and south | routing unit → north, east, south, and west

Every routing unit, except the one that sends the address, compares the incoming value with its own address (stored in the molecule underneath). At the end of this phase, that is, after `n` clock cycles, each routing unit knows if it is involved in this path. In practice, there has to be one and only one source, and at least one target.

Phase 3: Eliminating sources and targets
In some situations, a source should start a routing process, for instance, in a developmental process. In such a process, it would be useful to have many sources and targets with the same ID. So at this stage, it is possible there is more than one source involved in the routing process. In order to avoid multiple sources, in this phase that lasts only one clock cycle, if a source is at the origin of the routing process, it sends a signal to every other routing unit, to let them know a source is at the origin. Then every other source with the same ID disabled its participation in the current process, and during the next phase, the source will connect to the nearest target.
The same disable is performed in case a target launched the routing process. Every target that is not the master disables its participation to the current process, to ensure that the target that started the process will be the only one connected to a source. In this case, the nearest source will be connected to this target.

Phase 4: Building the Shortest Path
The last phase, largely inspired by [3], creates a shortest path between the selected source and the selected targets. An example involving 8 sources and 8 targets is shown in figure 10, for a densely connected network.

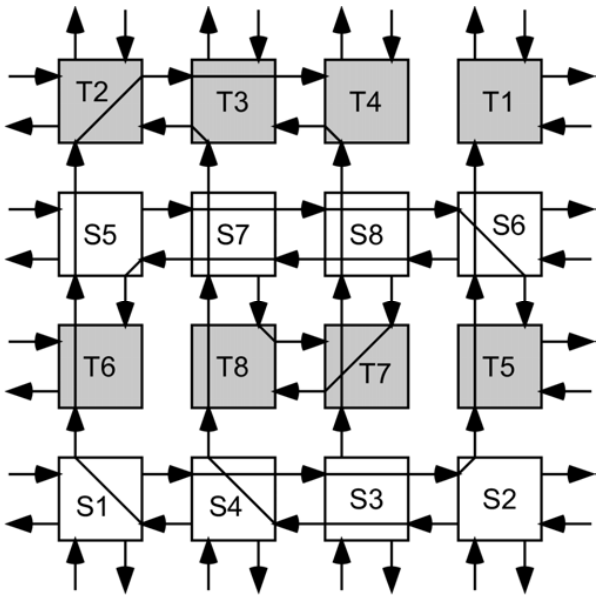


Figure 10. Test case with a densely connected network

A parallel implementation of the breadth-first search algorithm allows the routing units to find the shortest path between a source and many targets. Starting from the source, an expansion process tries to find targets. When one is reached, the path is fixed, and all the routing resources used for the path will not be available for the next successive iterations of the algorithm.

Figure 11 shows the development of the algorithm, building a path between a source placed in column 1, row 2 and a target cell placed in column 3, row 3. After 3 clock cycles of expansion, the target is reached, and the path is fixed, prohibiting the use of the same path for a successive routing.

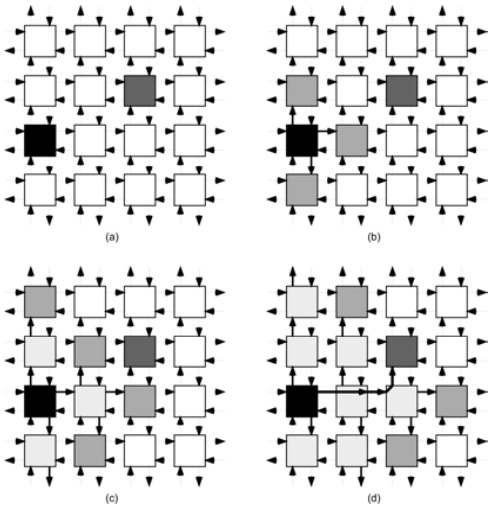


Figure 11. Step (a) one, (b) two, (c) three and (d) four of the path construction process between the source placed in column 1, row 2 and target cell placed in column 3, row 3

Based on addresses, the dynamic routing presented above is very flexible. However, for some applications, this flexibility can become a disadvantage, for example if we only need local communications between cells like a 4-neighborhood.

A second mode of routing has been added for this purpose. A flip-flop in the tissue can be configured by the molecules to choose the mode to use for a specific application. The pseudo-static mode uses the fact that every switch boxes are pass-through after a hardware reset. When in pseudo-static mode, the routing units that are connected to input or output molecules only shift the content of the molecule LUT into the configuration of the switchbox. By this way, in 16 clock cycles, the inter-cellular routing is completed, and the circuit can start its task. The only limitation is that a path between two cells can only be a vertical or a horizontal one, without more complex possibilities (figure 12).

2.3. The system interface

As it has been mentioned previously, the system interface of the Poetic tissue plays a major role in allowing for its scalability features. This means that the physical size of the tissue can be accommodated to the actual needs of a given application without posing specific constraints neither on the system architecture nor in the connectivity pattern among the POEtic chips that constitute the tissue.

The POEtic tissue, as it was presented in figure 1, can be constructed as a bidimensional array constituted by POEtic chips. The connectivity between these chips, as depicted in this figure, is based on two different buses, named organic (O) and interface (I) buses. The signals that constitute the organic bus allow the organic subsystems present in every POEtic chip to

communicate (at a cellular level).

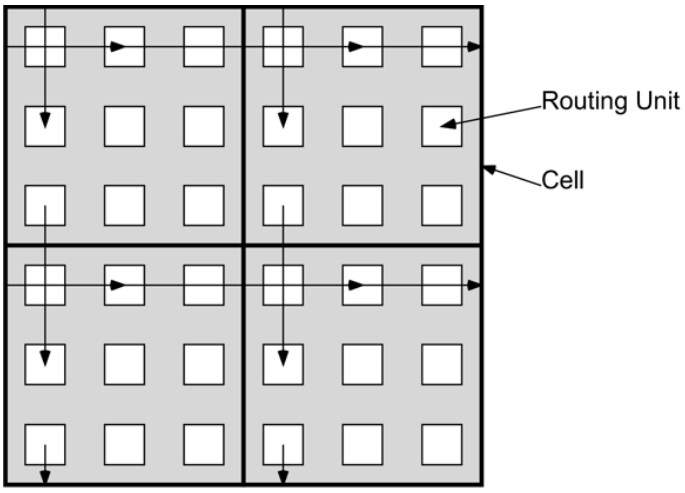


Figure 12. A pseudo-static communication scheme between four cells

The interface bus carries those signals that permit to handle the collection of POEtic chips as a single tissue, so that from a user point of view the tissue has only one environment subsystem and one organic subsystem. This is represented in figure 13.

Regarding the scalability of the environment subsystem, even if every POEtic chip contains a single environment subsystem, only one of them will be active in the tissue. This is accomplished by a specific signal present in every POEtic chip, called master, that indicates (when set to a value '0') that the environment subsystem of a specific chip will be managing the complete tissue.

The 68 signals (32 data lines, 32 address lines, sahbi_hsel, sahbi_hready, sahbi_hwrite and sahbo_hready) that constitute the AHB bus used for the POEtic tissue are connected to all the POEtic chips. This means that the chip identified as a master of the system can access the resources present in any other chip. A specific chip is identified within the array using Cartesian coordinates that correspond to the physical position of the chip in the array. This means that a chip with coordinates (X,Y) is placed in column X and row Y within the array.

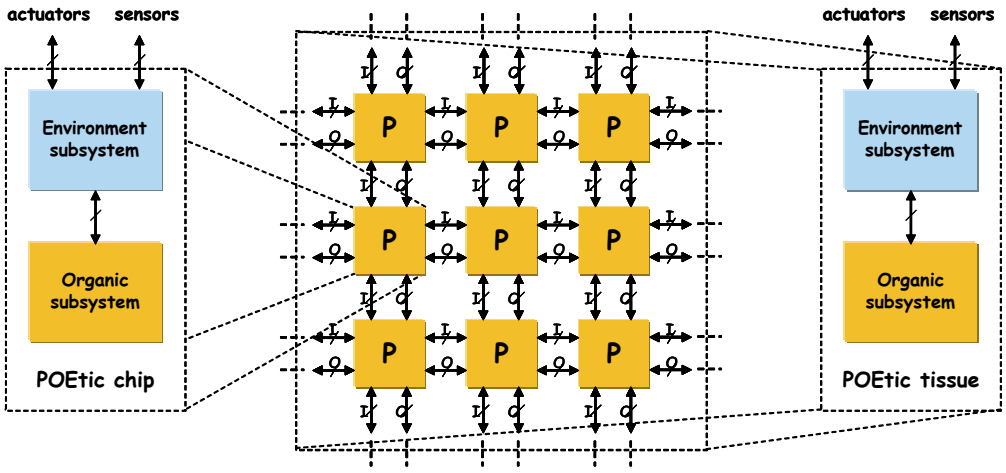


Figure 13. Scalability properties of the POEtic tissue

The coordinates of a given chip are not pre-programmed, but are calculated for a given array configuration during a coordinate propagation phase that should be performed before the tissue is operational. For this purpose every POEtic chip has two inputs, named Xin and Yin, and two outputs, Xout and Yout. The Xin input of a given chip is connected to the Xout output of the chip placed in the same row and in the previous column within the array. The Yin input of a given chip is connected to the Yout output of the chip placed in the same column and in the previous row within the array.

Every POEtic chip receives in serial mode its X coordinate through its Xin input and its Y coordinate through its Yin input. The coordinates are received in serial mode, so that by default the Xin and Yin inputs are in idle state (i.e., with a value '0'), and after one of these input is set to value '1' the POEtic chip should recognise that during the next 4 (in the current version of the POEtic chip the X and Y coordinates are 4-bit wide, but this can be easily extended to any desired size) cycles its X or Y coordinate will be received through the corresponding input. Once a given chip has received its X and Y coordinates it calculates and sends the coordinates for its direct neighbours. The coordinate propagation process is started by the chip whose environment subsystem has been identified as a master. The coordinate propagation process is started when the microprocessor included in the environment subsystem of the master chip performs a write cycle on the address 0xF000_0004 (as it was indicated in table 1, the organic subsystem is mapped in the memory space ranging from 0xF000_0000 to 0xFFFF_FFFF).

Once all the chips have got their actual coordinates within the Poetic tissue it is quite simple for the environment subsystem to access to the organic subsystem present in any chip. In order to access (either in read or write mode) the configuration of a specific molecule present in a POEtic chip placed at coordinates (X,Y) the environment subsystem should perform a read or write access to the memory position 0xF00X_YABC, where:

- **X:** Row where the POEtic chip is placed
- **Y:** Column where the POEtic chip is placed
- **A(3:0)B(3:0)C(3:2):** These 10 bits indicate the address of the molecule within the chip. One POEtic chip contains 144 molecules, and their mapping ranges from 0x002 to 0x091.
- **C(1:0):** These 2 bits indicate which one of the 3 configuration words of the molecule are to be read or written. A value "01" implies the activation of the cs1 signal, a value "10" implies the activation of the cs2 signal, while a value "11" implies the activation of the cs3 signal.

Bearing this in mind, the final organisation of the system interface included in every POEtic chip is that depicted in figure 14.

The wen signal depicted in this figure indicates if the access to the configuration of a given molecule is in read or write mode. The bidirectional configuration data bus is in fact constituted by two independent 32-bit buses, one for read access and the other for write access.

2.4. Physical implementation

The POEtic chip has been implemented and fabricated as an ASIC of 54 mm^2 using a 0.35 μm CMOS process. The chip, whose layout is depicted in figure 15, contains 144 molecules organised as an 8x18 array and the complete environment subsystem explained in previous sections.

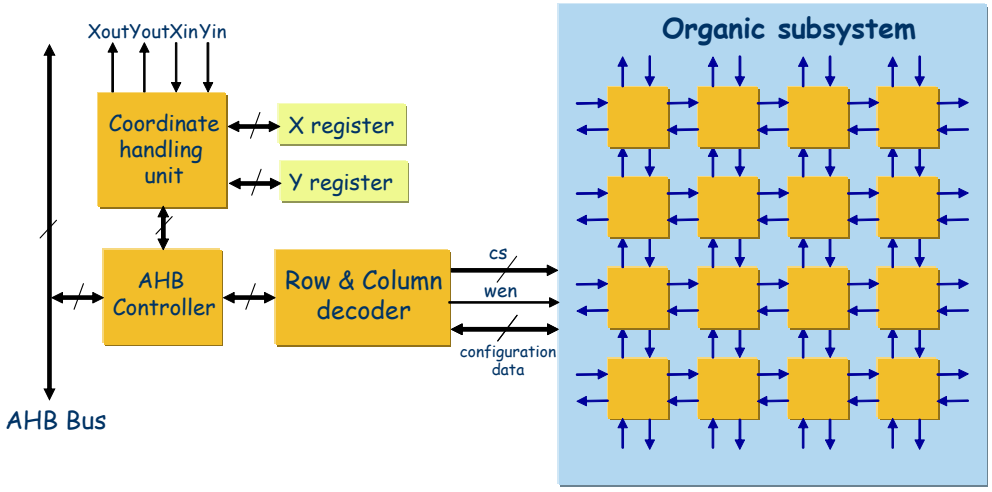


Figure 14. Internal organisation of the system interface

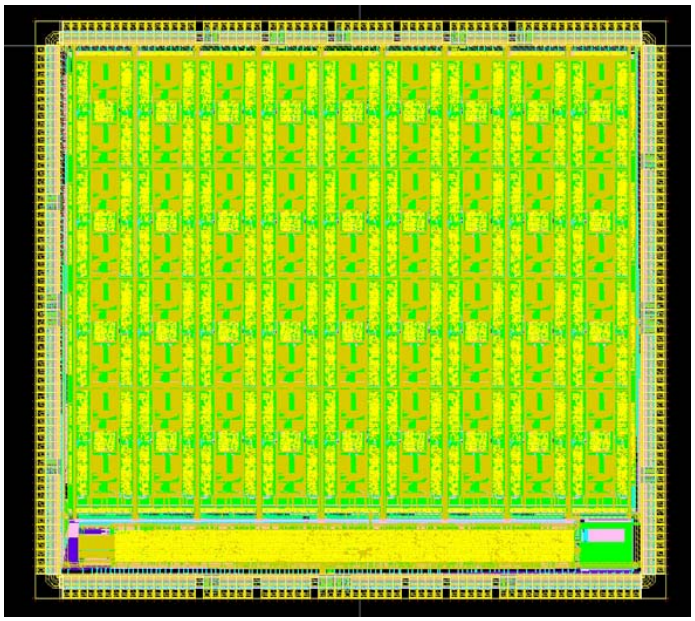


Figure 15. Layout of the POEtic chip

3. Emulation of large-scale spiking neural networks models

The spiking neural network model considered in our approach is that presented in [4]. This model outperforms previous approaches for implementing Spike Time Dependent Plasticity (STDP)-like learning methods when dealing with dynamic input stimuli.

Basically, this model consists in a leaky Integrate-And-Fire scheme, in which synapses can change their weights depending on the time difference between spikes. The outputs of the synapses are added until their result $V_i(t)$ overcomes a certain threshold θ. Then a spike is produced, and the membrane value is reset.
The simplified equation of the membrane value is:

$$V_i(t+1) = \begin{cases} 0 & when\ S_i(t) = 1 \\ k_{mem} \cdot V_i(t) + \sum J_{ij}(t) & when\ S_i(t) = 0 \end{cases} \quad (1)$$

Where $k_{mem}=\exp(-\Delta t/\tau_{mem})$, Vi(t) is the value of the membrane, J_{ij} is the output of each synapse and $S_i(t)$ is the variable which represents when there is a spike.

The goal of the synapse is to convert the spikes received from other neurons in proper inputs for the membrane. When there is a spike in the pre-synaptic neuron, the actual value of the output J_{ij} is added to the weight of the synapse multiplied by its activation variable. But if there is no pre-synaptic spike then the output J_{ij} is decremented by the factor k_{syn}. The output J of the synapse i-j is ruled by:

$$J_{ij}(t+1) = \begin{cases} J_{ij}(t) + (w_{RiRj} \cdot A_{RiRj}(t)) & when\ S_j(t) = 1 \\ k_{syn} \cdot J_{ij}(t) & when\ S_j(t) = 0 \end{cases} \quad (2)$$

Where j is the projecting neuron and i is the actual neuron. R is the type of the neuron: excitatory or inhibitory, A is the activation variable which controls the strength of the synapse, and k_{syn} is the kinetic reduction factor of the synapse. If the actual neuron is inhibitory, this synaptic kinetic factor will reset the output of the synapse after a time step, but if the actual neuron is excitatory, it will depend on the projecting neuron. If the projecting neuron is excitatory the synaptic time constant will be higher than if it is inhibitory. The weight of each synapse also depends on the type of neuron it connects. If the synapse connects two inhibitory neurons, the weight will always be null, so an inhibitory cell can not influence another inhibitory cell. If a synapse is connecting two excitatory neurons, it is assigned a small weight value. This value is higher for synapses connecting an excitatory neuron to an inhibitory one, and it takes its maximum value when an inhibitory synapse is connected to an excitatory cell.
In order to strengthen or weaken the excitatory-excitatory synapses, the variable A will change depending on an internal variable called L_{ij} which is ruled by:

$$L_{ij}(t+1)=k_{act}*L_{ij}(t) + (YD_j(t)*S_i(t)) - (YD_i(t)*S_j(t)) \quad (3)$$

Where k_{act} is the kinetic activity factor, which is the same for all the synapses.
YD is the learning variable that measures, with its decay, the time separation between a pre-synaptic spike and a post-synaptic spike. When there is a spike, YD will have its maximum value in the next time step, but when there is not, its value will be decremented by the kinetic factor $k_{learn,}$ which is the same for all synapses.
When a pre-synaptic spike occurs just before a post-synaptic spike, then the variable L_{ij} increases and the synapse strengthens. This means it reinforces the effect of a pre-synaptic spike in the soma. But when a pre-synaptic spike occurs just after a post-synaptic spike, the variable L_{ij} decreases, the synapse weakens and the effect of a pre-synaptic spike in the soma will descend. For other kind of synapses, the activation variable is always equal to 1.
Regarding the network configuration, 80% of the neurons are excitatory, while the remaining 20% are inhibitory. Each cell makes connections with other neurons within a 5x5 neighbourhood, i.e. 24 neurons. Figure 16 represents this connectivity pattern.

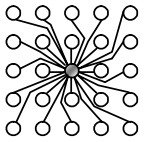


Figure 16. Connectivity of a single neuron.

The parameters that govern the functionality of the neuron block are:

- The membrane path has a resolution of 12 bits, with a range [-2048, 2047], and the threshold is kept fixed to +640.
- The membrane decay function has a time constant value of τ=20.
- The refractory time is set to 1.

The decay block will be used both in the learning and synapse blocks. Its goal is to have a logarithmic decay of the input; it is obtained with a subtraction and controlling the time when it is done depending on the input value. Taking into account that this block is used in many parts of the design, the variable decayed has been called x.
The block diagram is represented in figure 17. First of all, a new value of x should be obtained. It will be the input of a shift register which is controlled by the most significant bit of x and the external parameter *mpar*.
The output of this shift register will be subtracted from the original value of x. This operation will be done when the time control indicates it. The time control is done with the value of a counter that is compared with the result of choosing between the external value *step* or the multiplication of (MSB – *mpar*) by *step*. The decay variable τ depends on the input parameters *mpar* and *step*.

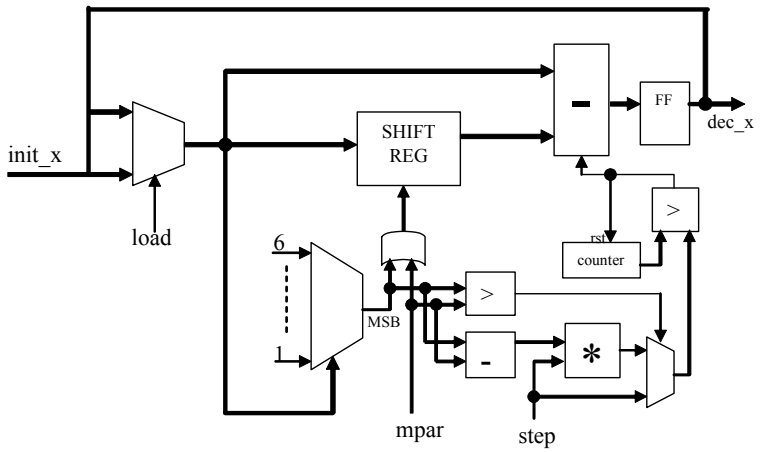


Figure 17. Block diagram of the decay block.

The learning block "measures" the time difference between a spike in the projecting neuron (j) and the actual neuron (i). Depending on these time differences and the types of the neurons, the synapse will be more or less active.

When a spike is produced in the projecting neuron, the variable YD loads its maximum value and starts to decay slowly. Then, if the actual neuron spikes, the value of YD_j is added to the decayed value of the L variable. On the other hand, if a spike is produced first in the actual neuron and after in the projecting neuron, the value of YD_i is subtracted to the decayed value of the L variable.

When the L variable overcomes a certain threshold (L_th), positive or negative, the activation variable (A) increases or decreases respectively, unless it is already in its maximum or minimum. If A is increased, L is reset to the value L-2*L_th, but if it is decreased, then L is reset to L+2*L_th. Figure 18 presents the organisation of this learning block.

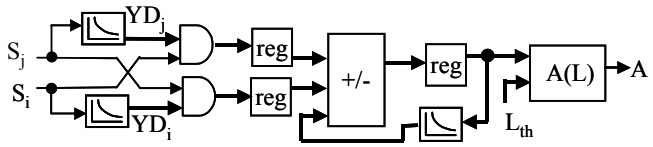


Figure 18. Organisation of the learning block.

The parameters that govern the functionality of the learning block are:

- The YD variable has a resolution of 6 bits and the learning variable (L) of 8 bits. The activation variable (A) can have four states.
- The time constant for the variable YD is τ=20.
- L_th= [-128,127]

To improve the sensitivity of the block for long time differences spikes, the time constant for the variable L is 4000, but it can change depending on the size of the network where the neuron works.

When there are spikes in the actual neuron after the spikes in the projecting neuron, the value of L increases, and the value of A also increases, so the synapse becomes more active.

The goal of the synapse block is to set the value of J (the input value added to the membrane) and it depends on four factors: the synapse activation level (A), the spikes of the projecting neuron (s_j) and the type of the actual neuron and the projecting neuron (r_i and r_j).
For each synapse a certain weight is set. This weight is multiplied by the activation variable (A). For this purpose, a shift register is used, so when A=0, the weight becomes 0, when A=1 the weight rests the same, when A=2 the weight is multiplied by 2 and when A=3 it is multiplied by 4.
This output weight is added to the decayed value of the output J. But the decay curve depends on the type of the actual and the projecting neurons (r_i and r_j).
There are two possible types for each neuron, excitatory and inhibitory, so we should obtain four possible values for the time constant which will decrease the addition. But, when both neurons are inhibitory, the weight of the synapse is always 0, so the J value is also always 0 and therefore it is nonsense to decrease it. Due to this reason, there are only three possible decay time constants.
The three time constants are multiplexed, and the multiplexer is controlled by the types of neurons (r_i,r_j). The multiplexer output controls the decay block, and finally we obtain the J value at the output of this decay block. Figure 19 shows the organisation of the synapse block.

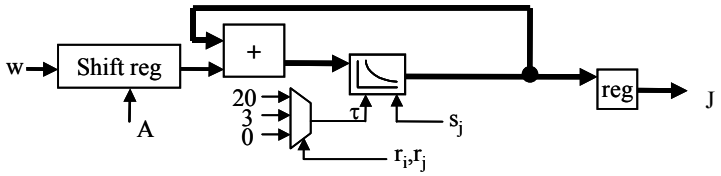


Figure 19. Organisation of the synapse block.

The parameters that govern the functionality of the synapse block are:

- The internal resolution of the block is 10 bits. But the output resolution is of 8 bits, due to the internal value of J is divided by 4 to keep the correct scaling.
- The time constants used by this block are presented in Table 3.

Time Constant (τ)	Projecting Neuron Type (r_j)	Actual Neuron Type (r_i)
20	0	0
0	0	1
3	1	0
0	1	1

Table 3. Time constants for different synapse types.

In this table r=0 means an excitatory neuron, while r=1 indicates an inhibitory neuron.

The high resolution needed for the variables, as well as the number of operations to be performed may pose a serious limitation for the final implementation. Therefore, the first step in the physical realisation of the model has consisted in an evaluation of the minimum resolution to be used in the neuron data path.

In a first attempt the resolution of the parameters has been reduced by two bits and some values and time constants have been changed to keep the correct scaling. Table 4 shows the new values of the internal parameters after this optimisation process. The final organisation resulting from this optimisation process is depicted in figure 20.

Due to the complexity of the design, the simplification of the model is very important to avoid redundancy or to use just the necessary components. For this reason, a further simplification of all the building blocks that constitute the model has been performed [5].

Parameter	New value
Membrane resolution	10
Threshold	+160
Input (J) resolution	6
Weights	[0:8],[64:128],[128:256],[0:0]
YD resolution	4
L resolution	6
Membrane decay time constant	20
YD decay time constant	20
L decay time constant	4000
J decay time constants (00,01,10,11)	20,0,3,0 (keep the same values)

Table 4. Resolution of the parameters for an optimised implementation.

Once the model has been optimised it has been physically translated into the molecules that constitute the basic building blocks of the organic subsystem of the POEtic tissue. Figure 21 shows this physical realisation.

The molecule organisation shown in figure 21 corresponds to the actual structure of the organic subsystem present in the POEtic tissue, which is arranged as an 8x18 array of molecules.

After designing the neuron model the VHDL models developed for the POEtic tissue have been configured and simulated to validate its functionality.

After this validation stage the strategy for the simulation of large-sale SNN models has been considered. Since in its actual implementation the POEtic chip only allows for the implementation of a single neuron and the current number of POEtic chips is far less than 10000 it will be necessary to use a smaller array of POEtic chips whose functionality should be time multiplexed in order to emulate the whole network.

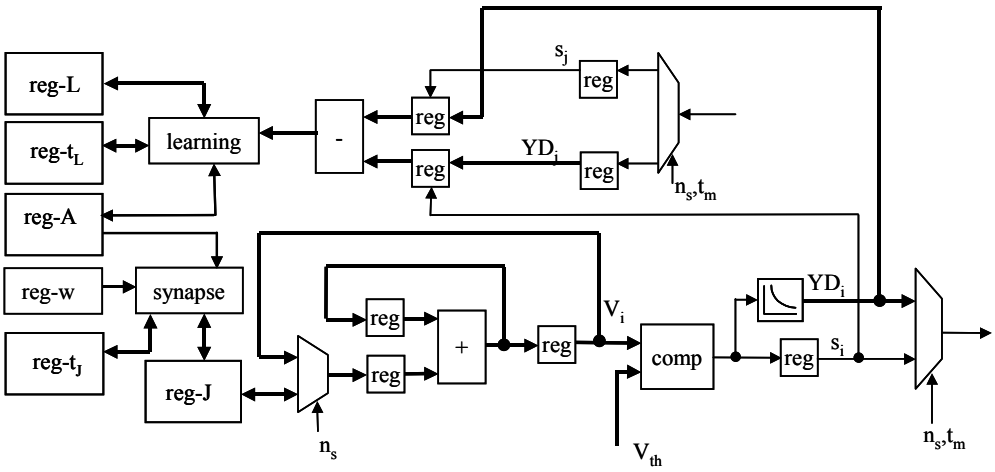


Figure 20. Block diagram for the serial implementation of the neuron model.

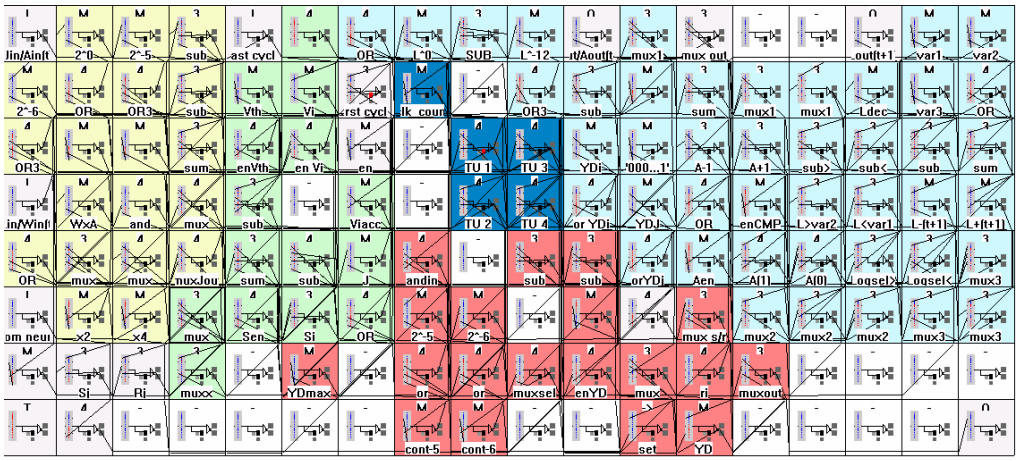


Figure 21. Molecule-level implementation of the neuron model.

This means that every POEtic chip should be able to manage a local memory in charge of storing the weights and learning variables corresponding to the different neurons it is emulating in time.

A 16-neuron network organised as a 4x4 array has been constructed using this principle. This would permit the emulation of a 10000-neuron network in 625 multiplexing cycles. Bearing in mind that each neuron is able to complete a single cycle in 150 clock cycles, this means that the minimum clock frequency required to handle input stimuli in real time (i.e., to process visual input stimuli at 50 frames/second) is around 5 MHz, far less than the actual clock frequency achievable by the organic subsystem of the POEtic tissue.

The visual stimuli will come from an OmniVision OV5017 monochrome 384x288 CMOS digital camera. Specific VHDL and C code have been developed in order to manage the digital images coming from the camera. To test the application, artificial image sequences have been generated on a display and then captured by the camera for its processing by the network.

4. Conclusions

In this paper we have presented a new family of programmable integrated electronic systems, called POEtic, that include features derived from some of the properties present in living beings, like evolution, development, self-repair, self-replication and learning.

The combination of partial and total dynamic reconfiguration, as well as the self-configuration and dynamic routing capabilities make these devices an ideal candidate for the efficient implementation of bio-inspired artefacts.

After describing in detail the different building blocks that constitute the tissue, an implementation approach for the emulation of large-scale spiking neural network models has been presented. The results derived from this implementation demonstrate that an electronic tissue built around these devices will permit the real-time emulation of this kind of models, thus serving as an excellent development and experimentation instrument for neuroscientists.

After receiving the first POEtic chips specific development boards have been constructed to develop applications to be solved using the bio-inspired features offered by the tissue.

Acknowledgements

The work presented in this paper has been funded by the grant IST-2000-28027 (POEtic) of the European Union (FET Proactive initiative on Neuroinformatics for living artefacts) and by grant OFES 00.0529-2 of the Swiss government. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

[1] ARM. Amba specification, rev 2.0. advanced risc machines ltd (arm). http://www.arm.com/armtech/amba_spec, 1999.

[2] Edsger. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.

[3] J. M. Moreno Arostegui, E. Sanchez, and J. Cabestany. An in-system routing strategy for evolvable hardware programmable platforms. In *Proc. 3rd NASA/DoD Workshop on Evolvable Hardware*, pages 157–166. IEEE Computer Society Press, 2001.

[4] J. Eriksson, O. Torres, A. Mitchell, G. Tucker, K. Lindsay, D. Halliday, J. Rosenberg, J.M. Moreno, A.E.P. Villa, "Spiking Neural Networks for Reconfigurable POEtic Tissue", Evolvable Systems: From Biology to Hardware, A.M. Tyrrell, P.C. Haddow, J. Torresen (eds.), pp. 165-173, Springer-Verlag, 2003.

[5] O. Torres, J. Eriksson, J.M. Moreno, A. Villa, "Hardware optimization and serial implementation of a novel spiking neuron model for the POEtic tissue", BioSystems, Vol. 76, No. 1-3, pp 201-208, August_October 2004.

AN ELECTRONICALLY CONTROLLED MICROFLUIDICS APPROACH TOWARDS ARTIFICIAL CELLS

Uwe Tangen*, Patrick F. Wagler*†, Steffen Chemnitz*,
Goran Goranovic‡, Thomas Maeke* and John S. McCaskill*

31st August 2005

* Ruhr-Universität-Bochum, Biomolecular Information Processing (BioMIP), c/o IZB Schloss Birlinghoven, 53754 St. Augustin, Germany

† Fraunhofer Gesellschaft, 53754 St. Augustin, Germany

‡ MEMPHYS-Center for Biomembrane Physics, Physics Department, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

Corresponding author:
Dr. Uwe Tangen, Ruhr-Universität-Bochum, Biomolecular Information Processing (BioMIP), c/o IZB Schloss Birlinghoven, 53754 St. Augustin, Germany, phone +49 2241/141530, email: uwe.tangen@biomip.ruhr-uni-bochum.de

Abstract

This work focuses on the application of on-line programmable microfluidic bioprocessing as a complementation vehicle towards the design of artificial cells. The electronically controlled collection, separation and channel transfer of the biomolecules are monitored by a sensitive fluorescence setup. Two different physical effects, electrophoresis and electroosmotic-flow, are used to allow for a detailed micro-control of fluids in micro-reaction environments. A combination of these two basic electronically controlled input reaction chambers makes combinatorial fluidic-networks and indefinitely sustained biochemical or chemical reaction-networks feasible. Experimental data showing the power of this approach is presented. Not only does this processing power pave the way towards the development of artificial cells (using a technology to complement not yet established autonomous metabolic or replication capabilities) but it also opens up new processes for applications of combinatorial chemistry and lab-on-a-chip biotechnology to drug discovery and diagnosis.

Keywords

Reconfigurable hardware, microfluidics, micro-reactor, PDMS, DNA, microelectrode, fluorescence detection

1 Introduction

Integrated microfluidics holds the promise of allowing complex parallel and pipelined processing of combinatorial families of biomolecules without the serial limitations of external robotic pipetting [1]. They allow one to build up dynamic micro-reactors and biochemical networks that are user-programmable, in space and time. Through operational feedback and multiple reconfiguration, they can also be iteratively optimized in an evolutionary process [2]. The development of integrated applications in biotechnology are limited by the necessity of custom hardware and time-consuming development cycles [3, 4, 5, 6]. Reconfigurability opens new possibilities, as one patented strategy allowing researchers to proceed from simple connected micro-reactor elements to dynamically reconfigurable micro-reactor networks [7], shows. These consist of a combinatorial array of processing and switching elements, which can be externally programmed.

Examples of applications are microscale reactors for evolutionary biotechnology, active microfluidics for μTAS and molecular computing in micro-reactors [8, 9, 10]. Artificial containers such as droplets and vesicles in microfluidic environments represent an important area of research in down-scaling combinatorial chemistry but we have proposed that they can also assist to engineer artificial [11, 12] and minimal [13] cells. The important idea in this respect is to use the electronically controlled environment as a substitute for controlled cell functions like metabolism and/or replication, that cannot yet be autonomously regulated. We analyzed the integration of digital microfluidics using immiscible fluids in closed channels, for droplet generation, fusion and transport. Reactors were fabricated in micro-molded PDMS (polydimethylsiloxane) or bonded silicon to glass. We extended previous work [14] to integrate continuous droplet generation and autonomously regulated serial droplet fusion [15]. In the experiments reported here, we incorporated in our microfluidic technology electroosmotic flow [16] (EOF) in addition to electrophoresis and show its feasibility.

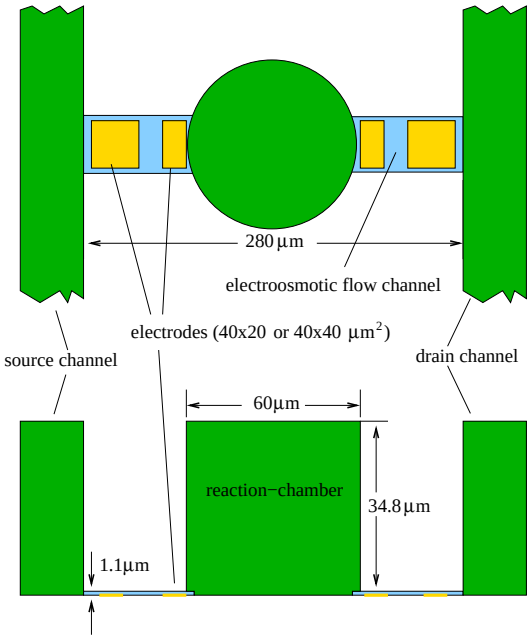


Figure 1: Electronic regulated chamber for an artificial cell connected to chemical supply and drain with two IO-channels of height 1.1μm (measured with profiler P-10 (Fa. KLA-Tencor, USA) and White-Light-Interferometer Wyko NT 1000 Optical Profiler (Veeco Process Metrology)) and regulated by two gold electrodes placed below each I/O. The reaction-chamber, micro-moulded (SU-8 master) in PDMS with size of 34.8μm by 60μm diameter, is effectively decoupled from the hydrodynamic flow of the supply and drain channels, while the programmable electrodes enable adjustable bulk electroosmotic flow (EOF) and electrophoresis. The entire structure can form a basis of a system of programmable microfluidic networks (e. g. Figure 6)

The main unit of our system is an electronically programmable microfluidic chip equipped with micro-electrodes and fluidic channels and chambers, linked to a reconfigurable electronic chip. The micro-electrodes are the actuator components in the microfluidic network, see Figure 1, for the EOF reaction-chamber used in the experiments reported here. Furthermore, the chip contains a standard field-programmable-gate-array (FPGA). The reconfigurable logic device is integrated to control a maximum number of electrodes individually. With this intimate control of biochemical processes a precondition for "live control" comes within reach, that the seamless integration of computer technology and biology occur in both directions simultaneously. This involves both the real time detection and data analysis of the biomolecular system and the active control of biomolecules in a hybrid system [17]. Currently most of our experiments use small DNA oligomers (21 nt, Rh6G–GATGGTCACAGCATGTCTGTS) to test the setup.

Three physical effects are apparent using this technology:

- Electrophoresis
  The electrical field is quite inhomogeneous due to the small electrodes (e.g. 10-40μm) and high field strengths. DNA contains non-aligned permanent dipole moments and a permanent negative charge from the phosphate groups which is partially offset by counter ions from the solvent. Two forces cause electrophoretic molecular transport. The first one is the force of the field on the net charge. This is the same force that drives the ionic constituents of an electrolyte through the solution. The second is the force on the dipole of the molecule: which increases with the divergence of the field. Besides permanent dipole moments DNA also has significant induced dipole moment forces. Note that the strong molecular length dependence of DNA in gels (amplified by chain entanglement) is much less pronounced in free solution. The electrophoretic mobility in free solution in the simplest analysis is the ratio of net charge to hydrodynamic resistance (which follows the Stokes law $F = 6\pi f \eta R$ with an asymmetry correction factor f).

- ElectroOsmotic-Flow (EOF)
  In translation invariant geometry (infinite channels) the Hagen-Poiseuille (H-P) law is valid for a steady-state pressure driven flow of incompressible fluid

$$\triangle p \quad = \quad R_{\mathrm{hyd}} Q \tag{1}$$

  which says that the flow rate Q is proportional to the pressure drop. Electroosmosis, on the other hand, is a non-equilibrium effect, where a liquid is brought to move relative to a charged surface an applied external electric field by acceleration of ions in the solution double layer which screens the surface charge. Surface double-layers are characterized by the ζ-potential. For an ideal EOF at an infinite plate (uniform ζ-potential along the wall, homogeneous E-field, steady-state flow, the Debye layer much thinner than the smallest channel dimension), the electroosmotic velocity is given by

$$u_{eo} = \alpha_{eo} E = \frac{\epsilon \zeta}{\mu} E. \tag{2}$$

  Here α_{eo} is the electroosmotic mobility, ζ is the ζ-potential of the wall, ε and μ are the dielectric constant and dynamic viscosity of the electrolyte, respectively, and E is the applied electric field. Total EOF in a micro-channel is then the sum of individual EOFs

from each wall, each approximated by (2).
In order to obtain stable fluid control into and out of the reaction chambers, Figure 1, it is crucial that pressure driven hydrodynamic flow from the source and the drain channels [3] is suppressed. Flow control can be obtained despite high resistance using electroosmotic flow in many small or a single high aspect ratio channel. To underline this essential mechanism, we first look at an isolated rectangular micro-channel of length L, width W, and height h, with an (infinitely) thin double layer on all walls. The electroosmotic flow rate in it is given by

$$Q_{eo} = u_{eo}hW = \alpha_{eo}V_{\text{eff}}\frac{hW}{L} \propto, \tag{3}$$

where $u_{eo}$is the electroosmotic velocity, α_{eo} the EO mobility, and V_{eff} the electric potential drop inside the channel. Note that this expression implies that electrodes at fixed voltage, placed at either end of a very short electroosmotic channel segment, would produce a very large (diverging) volume flow. Clearly the reason for the problem is that the hydrodynamic resistance of the remaining microfluidic system has not been included. For a high aspect ratio, i.e., $W \gg h$, the hydraulic resistance of the channel is

$$R_{\text{hyd}} = \frac{12\mu L}{h^3 W}\frac{1}{1 - 0.63\frac{h}{W}}, \tag{4}$$

where μ is the dynamic viscosity. The pressure driven flow rate Q_p through the channel is given by (1) with $Q_p \propto h^3$. From (3) and (1) it seems that pressure driven flow will be negligible compared to the EO flow for small values of h, but this only applies for loadless channels.
In the case of electroosmotic channels with a load, e.g. connected to a field free channel segment of hydrodynamic resistance R_L, the lowest order approximation is to use Kirchoffs laws in connection with the low Reynolds number (creeping) flow. Basically, the volume flow for the loadless electroosmotic driven segment is the resting state, and reductions of this flow velocity by the load cause a linear pressure difference across the electroosmotic segment, which can be calculated by (1). Kirchoffs laws imply that the pumped flow velocity satisfies

$$\triangle p_{\text{ext}} + \triangle p_{eo} + \triangle p_L = 0 \tag{5}$$

with the volume flows in the osmotic flow and load channels equated

$$Q_L = \frac{\triangle p_L}{R_L} = Q_{eo} + \frac{\triangle p_{eo}}{R_{eo}} \tag{6}$$

The hydrodynamic resistance is calculated from (4) and the loadless electroosmotic flow from (3). Rearranging, we find

$$Q_L = \frac{R_{eo}Q_{eo} - \triangle p_{\text{ext}}}{R_{eo} + R_L} \tag{7}$$

which for zero external pressure difference reveals a reduction in the loadless osmotic flow rate by a factor of $r = \frac{R_{eo}}{R_{eo}+R_L}$. The underlying additivity assumption is based on a sharp transition from a plug flow profile to a parabolic flow profile at the junction of the driving and load channels.

For our experimental setup, the dimensions of the electroosmotic segments are $W \times L \times h = 40 \times 40 \times 1.1 \mu m^3$, and the total load resistance is dominated by the electroosmotically non-active part in the horizontal channel, the sink resistors are, compared to this, two orders of magnitude smaller and the source does not even contribute to the total load resistance. This leads to $r = 0.18$, implying a six-fold reduction in the electroosmotic induced velocity. For an overall electroosmotic mobility of $4 \times 10^{-4} cm^2/Vs$, an effective voltage of 3.3V, the channel geometry as above, $Q_{eo} = 0.52 \mu L/h$ and without an external pressure difference, the flow in the horizontal channel is given by $Q_L = 0.1 \mu L/h$. In our experiments the non-reversed EOF is directed towards the cathode, indicating negatively charged walls and a positively charged double layer. In PDMS, one common problem is the poorly defined EOF and its dependence on the process used to seal the chip. While it has been reported that an oxidation step is required to produce EOF [18] others claim that no such a step is necessary [19]. The use of multiple materials in a single device, each with a unique ζ-potential, may create non-uniform flows. Currently we cannot distinguish between the EOF created by the SiO_xN_y layer and that created by PDMS walls. The charged site on SiO_2 layers is invariably Si-OH, regardless of the silica type [20] and the presence of silicon nitride also results in predominantly Si-OH charges, with a minority of Si-NH_2 charges shifting the ζ-potential slightly toward positive values, so that the composed layer exhibits roughly similar results. Liu et al. [21] measured the electroosmotic mobility of native and oxidized PDMS/glass devices, both in a range of $4 \cdot 10^{-4} cm^2/Vs$ at neutral pH, while, and in contrast to this, Lin et al. [25] investigated the temporary effect of the plasma modified PDMS and found that the ζ-potential decays to zero within two days. Finally, Bianchi et al. [22] presented a model allowing the determination of each individual ζ-potential in composite micro-channels. Here it remains for future work to clearly identify all electro-physical properties.

- Electrolysis
  The applied voltage of 3.3V implies a highly positive reduction potential. Thus it would be possible to perform quite easily an oxidation of the DNA oligomers used in our experiments, since the redox-potentials of DNA are in the range of a few hundred millivolts [23]. Furthermore, the applied potentials should lead to an electrolysis of water of which the redox-potential is +1.23V (O_2/H_2O). Although it is not easy to produce visible electrolysis in the buffer employed in our experiments, this will become an issue if electrodes are driven with DC for more than 10 seconds. Fortunately, it turned out that the experiments with EOF, even when statically driving the electrodes (at constant voltage) for 10 minutes did not show electrolysis. The major reason might be due to the extreme surface-volume ratio (1.1μm height of the channels, see Figure 1). The EOF produced high fluid velocities inside the shallow channels and thus gas-bubbles would be quickly transported away from the electrical field.

2 Overall experimental setup

To distribute the DNA and buffer solutions on the chip a liquid-handling system was connected, see Figure 2. The microfluidic workstation includes high-performance syringe pumps for ultra-low flow rates (MMT). The detection system is based on laser-induced fluorescence and a special optical system as well as a CCD camera for image acquisition. The electrode design was realized with a standard CAD schematic entry system for PCBs (Board Station Mentor

Graphics). The special computer hardware MereGen^{TM} based on reconfigurable electronics [24] analyzes the inner state of the micro-reactor network and realizes the programmability of the system using feed back loops.

The microfluidic device is equipped with gold-electrodes (size typically 20 x 40μm or 40 x 40μm, 96 per module) and a standard field-programmable gate-array (FPGA, SpartanXL XCS20CSP144 Xilinx) that serves as the driver for the electrodes and as the interface to the MereGen board. The electrodes themselves are driven via the standard output drivers of the FPGA meaning a digital 3.3V signal level. With the three states provided (3.3V, 0V and tristate) fast cycling allows us to emulate effectively arbitrary voltage-levels between both extreme values. The configurable logic of the FPGA also serves as a testing-device for the communication buses and built-in self tests.

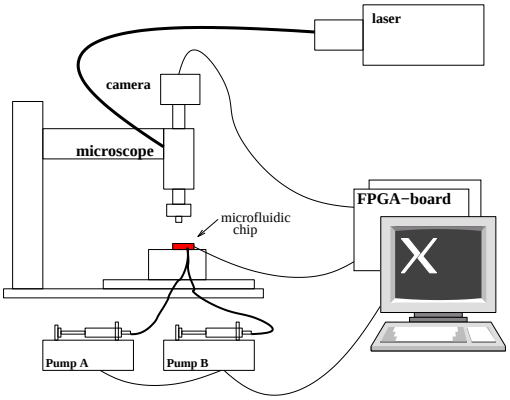


Figure 2: Overall experimental setup. Camera Vosskuehler VDS-QCam 1300QLN (1280*1024, 11 fps), laser argon 514nm, about 100 mW at the chip, Micro-Mechatronic Technologies AG (MMT) pumps (0.2μL/h - 240μL/h with 10μL syringes), custom designed and custom built MereGen^{TM}-FPGA board [24], Linux based VME-bus host-computer (Concurrent Technologies VP 100/01x Pentium III 1.2 Ghz, 512MB), microscope Zeiss Axiovert. The user-friendly control-software is custom-designed.

When molecules are moved, or chemical properties in the fluidic-system changed, fluorescence- or white-light imaging using the camera mounted on the microscope gives the possibility of direct on-line control (see Figure 2). The camera image is preprocessed on the MereGen^{TM} board. Arbitrary regions selected either automatically, via image processing, or manually may be evaluated, statistically processed and subjected to control rules. Depending on the complexity of these calculations, at least the hardest real-time-critical components can be realized in hardware. Introducing low-level control-structures paves the way to connect this work with computer-science on hybrid-systems.

A custom-developed software and operating system, with a user-friendly graphical interface, gives the experimenter an easy access to the underlying hardware. A nice feature of this software is the remote operating capability.

3 Experimental results

Two different solutions in two Hamilton syringe pumps were connected to the two inlets: one solution contained just buffer solution and the other Rhodamine-6G-labeled single stranded DNA with a length of 21 nt. A pump rate chosen between 0 and 2 μL/h was maintained stably in the supply and drain channels. The buffer solution employed was a freshly prepared histidine buffer (50mM, 59μS/cm, pH 7.7) to minimize the electrostatic screening effects (Debye length) [25].

The first task of the experiment was to fill the reaction chamber, see Figure 1, with the labeled DNA. The pump rates were adjusted such that a slow net flow from bottom to top (drain to supply) was maintained if no electrodes were active. At the beginning of the

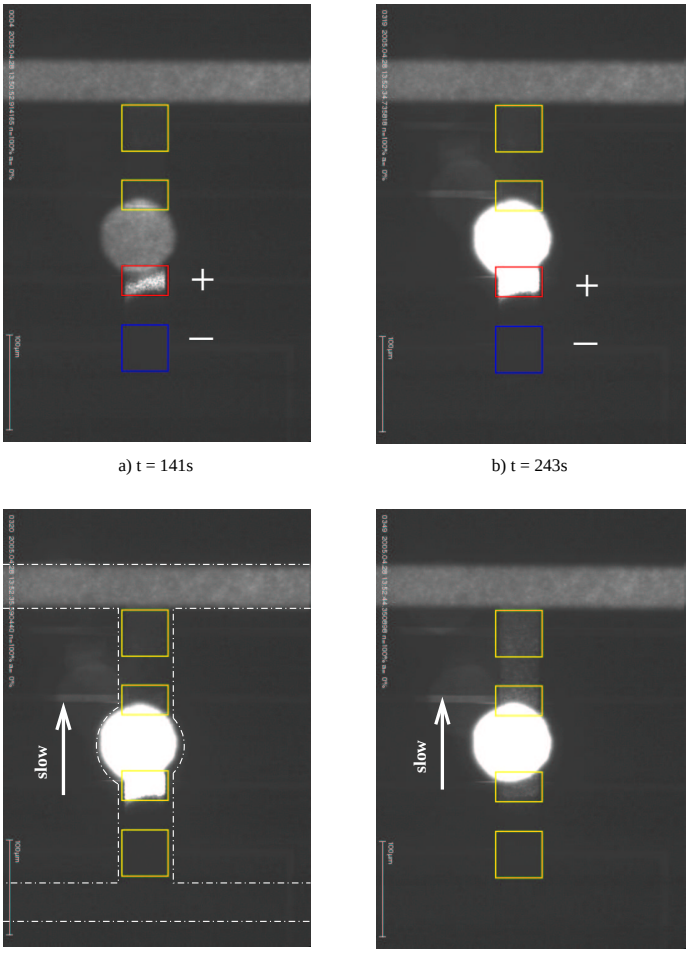


Figure 3: Experimental demonstration of electronically programmable cell filling. Four samples at different times of a filling experiment are shown. The upper horizontal channel (depicted in part c as dashed white lines) is filled with Rh6G $5 \cdot 10^{-6}$ labeled 21mer primer Rev500 batch #33436N in Histidin-buffer 50mM, 59μS/cm, pH 7.7 and excited with laser-light at 514nm with a power of 100mW. At the beginning of the experiment no fluorescence could be observed in the reaction chamber. Two different physical effects are contributing to filling the reaction-chamber: electrophoresis (the positive electrode, 3.3V, a and b) attracts the negatively charged DNA and electroosmotic-flow which drags fluid from the upper supply-channel against the weak background flow from the lower drain-channel upwards. After releasing the electrodes (tristate, high-impedance), material from the reaction-chamber is slowly washed out again. Because of the considerable difference in transport-capacity of the input/output channels and the storage capacity (see Figure 1) of the reaction chamber, this process is quite slow. With an appropriate presetting of the electrode potentials, even this slow dilution can be prevented.

experiment, a faint fluorescence in the DNA-supply ($5 \cdot 10^{-6}$M) channel was visible. The filling procedure startet with setting the second electrode from below to 3.3V, and the bottom electrode to 0V, as shown in Figure 3-a. The two field effects, EOF and electrophoresis, help each other in this case. The negatively charged DNA is attracted from the supply-channel and in addition a net EOF flow is generated which pumps fluid from the supply-channel towards the drain-channel. On the other hand, the electrical field is polarized such that the DNA is repelled from the bottom electrode and attracted by the upper electrode. This results in the observed increase of concentration of DNA, see Figure 3-a-c. Note also the times involved in this filling procedure. Even after more than 200 seconds no electrolysis was observed. Immediately after releasing the electrodes, the slow volumetric flux (from bottom to top of the image) washes the material from the formerly positive electrode back into the reaction chamber (see Figure 3-d). Despite the background flow, the time needed to wash out all the material from the reaction chamber is considerable as a result of the extreme height-ratio between reaction chamber and feeding channels. Further experiments showed (data not shown here) that, with suitably activated electrodes, even this weak outwash process can be drastically reduced and that additionally a reaction in the chamber might be fed continuously.

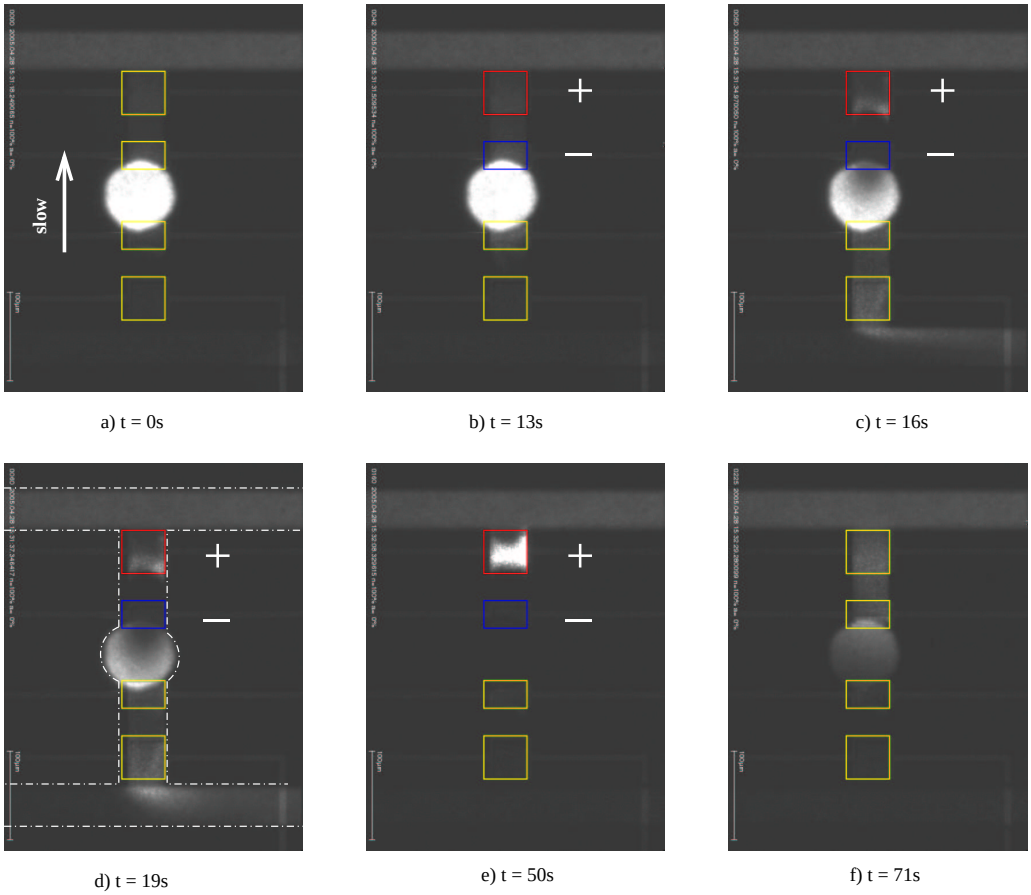
The second task was to drain the reaction chamber again. Two possible draining directions are available - pumping the material back into the supply-channel (data not shown) which works remarkably fast (typically about 20 seconds are needed to reduce the fluorescence intensity to background level again), and pumping the material against the default flow-pressure into the drain channel, see Figure 4. A time-series of six images is presented. The experiment differed a little from the filling-experiment in that the concentration of DNA in the supply-channel was chosen ten times smaller, $5 \cdot 10^{-7}$M. Everything else remained the same. In this case, the top electrodes are operated, reversing the electroosmotic-flow and expelling the DNA from the reaction-chamber into the drain-channel. After about 50 seconds, the fluorescence intensity in the reaction-chamber was again indistinguishable from the background. As a side effect, the top electrode additionally attracts DNA from the supply channel. This DNA can either be used as a kind of pre-charging for the next reactor filling step or be actively pumped out, after the reaction chamber has been emptied (data not shown). If only the electrodes are deactivated, the high concentration of DNA diffuses along the filling channel in both directions, part of which can be seen in Figure 4-f.

4 From simple single reactors to networks and programmability

The remarkable local control of fluids and concentrations of charged biomolecules shown, and the opportunity to combine this with water-droplets in oil or vesicles in water, immediately raises hopes for strong programmability and combining these features in large networks.

Hybrid-System Programming and Control

Programming in these systems, in the traditional computer-science sense of fixed serial algorithms, is a challenge in several aspects: it has to cope with the continuous versus discrete dichotomy, it involves essentially a stochastic description and it has to cope with reproducibility and error-prone information processing. Even linearization around working-points is difficult in many cases. The first attempts to describe optimal control-strategies with continuous and discrete systems - termed then as Hybrid Automata - were presented by Alur et al. [26] in 1993, as a model and specification language for systems consisting of a discrete program within a continuously changing environment. A major result is that even the very



a) t = 0s b) t = 13s c) t = 16s

d) t = 19s e) t = 50s f) t = 71s

Figure 4: Experimental demonstration of electronically programmable cell draining. Pumping material out of the reaction-chamber is shown in the sequence of six photographs taken at different times. The starting situation is similar to the end-situation shown in Figure 3-d. The concentration of the DNA in the supply-chain has been reduced by a factor of ten ($5 \cdot 10^{-7}$M) before filling the reaction chamber. The two physical forces, electrophoresis and electro-osmosis, empty the reaction-chamber (rapidly considering the huge size of the reaction-chamber in comparison to the shallow input/output-channels) in seconds. A side effect of electrostatic forces is the uptake of DNA-material from the supply-channel (upper channel, sketched as white dashed lines in d). This side effect can be used to create packets of charged material and transport this into a following channel-system. As can be seen in Figure 3, a slow default stream upwards cleans up the reaction chamber after the experiment. Additional suitable electrode control can drastically clean the reaction-chamber (data not shown).

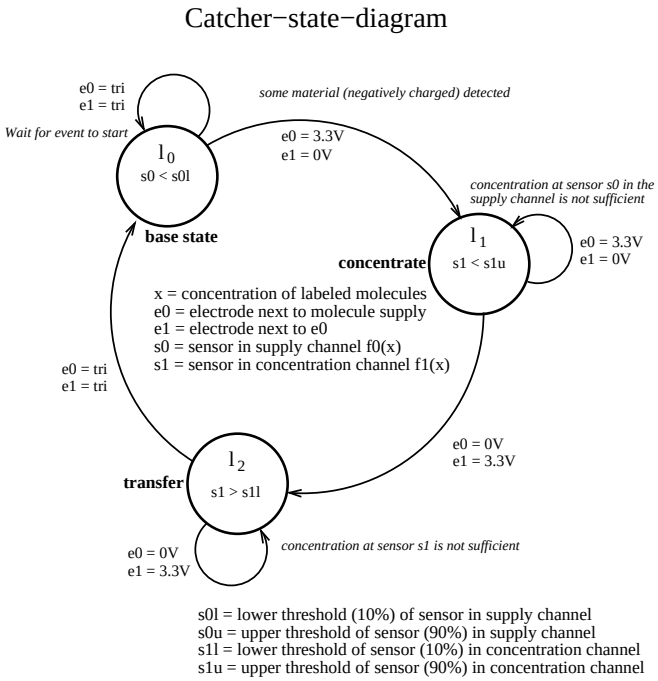


Figure 5: A state-machine diagram (notation used from [26]) of one basic control-module called 'the catcher'. Due to the on-line detection system actively controlled procedures are possible. This one serves for attracting charged material and generating an event to a following stage of the control-network. Such control loops form the link to bridge the gap between microfluidic experimentation and global programming of biochemical processes with software. Noise, detection-problems and other real-world effects can thus be reduced.

simple and restricted linear hybrid automata verification is already undecidable, although three special cases for which verification is possible are presented. In addition to being a hybrid automata, programmability and control of the chip has to cope with stochastic processes because whether molecules are detected or not depends stochastically on their concentration and on surface-properties. This concentration can be either a stochastic continuous variable (10^{14} molecules or more can be seen as continuous reaction-fronts, that may have arbitrary shapes) or discrete when oil-droplets, vesicles or artificial cells are considered. Bujorianu and Lygeros [27] developed a model for general stochastic hybrid systems (GSHS) as a generalization of piecewise-deterministic Markov processes. With the help of dynamic programming, they model GSHS and solve the derived differential equations. Abate et al. [28] describe stochastic approximations of deterministic hybrid systems, again using Markov processes.

Although hybrid-systems research has its origins in classical engineering disciplines, Piazza and Mishra [29] elaborate on questions of stability of hybrid-systems in Systems Biology. Two description-languages in the realm of hybrid-systems seem to be suitable, CHARON [30] and HYSDEL [31], specifically developed as formal specification languages to program and control hybrid-systems. Using the notation of Alur [26], a typical control-element used is shown in Figure 5 and will be discussed here briefly.

Low level regulatory elements

A whole series of low-level control-elements have been devised. The state-machine diagram of one, the catcher, is shown in Figure 5. Two electrodes (actors) and two sensors, suitably placed in the vicinity of the actors are needed. The sensors can produce two signaling events: a drop in intensity of the fluorescent signal below a certain threshold and surpassing intensity above a second threshold. Both thresholds are programmable, even on-line. If the catcher element has captured a sufficient amount of charged material, it signals a following stage (another low-level control element or optionally a more abstract computational entity) that its purpose is finished. The major challenge for incorporating this into control-strategies lies in the non-linear, stochastic and error-prone functional relationship between the concentration of material x and the threshold-parameterized output of sensors s0 and s1. In linear hybrid systems, this relationship is defined as being linear and the proofs of optimal control are based on this assumption. It remains to be experimentally proven that, at least for certain operating conditions, these ill-behaved relationships can be approximated by linear functions or that the current research on optimal control including stability analysis can be extended to these types of functions.

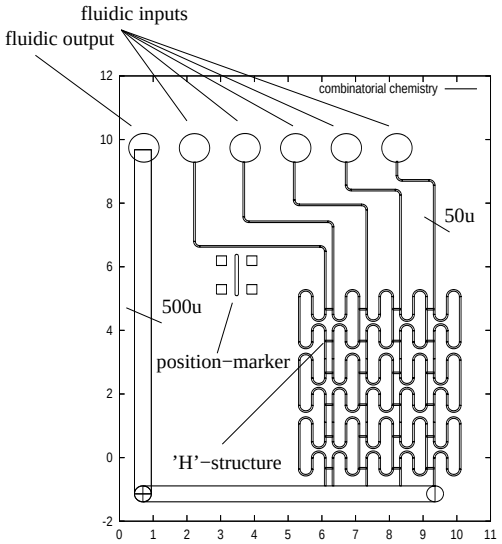


Figure 6: Test-network used to prove the feasibility of combinatorial chemical reactions in microfluidics. The difficult flow-control of this system limited us in the past from doing further research on this topology. However, using the actively controllable reaction-chambers shown in Figure 3 and 4, in a network presented here (graph of a current fluid-design with an n-way fluid-crossbar) combinatorial chemistry in micro-flow-systems will be feasible. Material from each of the five inputs can be brought into contact with each other. The loops shown are used for delay purposes, to allow a synchronized flow behavior. The common output channel not only serves to reduce the number of fluid-connectors but also as a flow-stabilization device (passive regulation).

Combinatorial reaction networks

With these two basic elements, EOF-reaction-chamber and low-level control-elements, the path is open for combining them to create efficient networks. One of the main obstacles in using networks of flow-coupled reaction-chambers, see Figure 6 for a design we tested, is to control the flow. Not only did it prove to be difficult to control the flow in the connecting channel of the H-structures (because of the imperfections which always exist in the fluid-connections and the production process of the imprinted channels in PDMS), it could also be shown that moving DNA packets between the channels do have an effect on the local net flow rates of the fluids in the system (data not shown). With this new strategy, the major flows can now be largely decoupled and independent local cyclic flow-systems can be established and controlled. This latter capability will open up a range of possibilities for evolving artificial cells. More practically at this stage, the role of the specially pulse-free (and expensive) external low-flow-pumps can be reduced to just feeding the system and letting the electrode-systems control the local flow (via EOF).

5 Discussion and conclusion

The aspect of functional complementation of artificial cells that has been investigated here is that of electronically regulated and controllable micro-compartments. To support an artificial cell, chemicals have to be introduced to and removed from these compartments and chemical reactions should be sustained for hours. These long-lasting reactions (the metabolic and replication apparatus in true artificial cells is enzyme-free and thus chemical reactions might take very long times) need to be protected from the outside world and they need to be supplied with the necessary chemicals. Our experimental observation that the transport of charged material worked better for smaller entrance paths, has led to a significant enhancement using EOF. Two physical effects help to alleviate the problems of transport. The strong (at least reciprocal cubic) dependency of the hydrodynamic resistance on channel dimension can be used to decouple channels in microfluidic networks efficiently and drastically reduces fluid-noise in the reaction-chambers. The second effect is the electroosmotic-flow (EOF) that utilizes the ubiquitous ionic-double-layer to produce a net fluid transport under electric fields. Combined, these effects with two-phase systems will further enhance the combinatorial complexity of biochemical reactions which can be regulated in such systems.

The most fascinating potential of this new tool-box is a network of reaction chambers with arbitrary connection-topologies and detailed on-line control. Such systems can benefit from and provide new challenges for the decades of experience and knowledge in computer-science and dynamical system-control. These types of reactor networks can bridge the gap between biochemistry, control-theory, computer-science and "hard" artificial-life research.

Acknowledgments This work was conducted as part of the European Union integrated project PACE (EU-IST-FP6-FET-002035). Microfabrication, according to our design, of the EOF-reactor by Martina Jünger (RUB) and Thomas Palutke (FhG) is greatly appreciated.

References

[1] Sanders GHW, and Manz A: Chip-based microsystems for genomic and proteomic analysis. TrAC Trends in Analytical Chemistry 2000; 19: 364-378

[2] McCaskill JS, Wagler P: From reconfigurability to evolution in construction systems: spanning the electronic, microfluidic and biomolecular domains, Lect. Notes in Comp. Sci. 2000; 1896: 286-299.

[3] Ekins R, Chu FW: Microarrays: their origins and applications. Trends in Biotechnology 1999; 17: 217-218.

[4] Service RF: Will patents fights hold DNA chips hostage, Science 1998; 282: 397.

[5] Edman C, Raymond D, Wu D, Tu E, Sosnowski R, Butler W, Nerenberg M, Heller M; Electric field directed nucleic acid hybridization on microchips. Nucl. Acids. Res. 1997; 25(24): 4907-4914.

[6] Schmidt K, Foerster P, Bochmann A, McCaskill JS, in "Microreaction Technology - First International Conference on Microreaction Technology", Springer-Verlag (Berlin Heidelberg New York), 1998; 233-244.

[7] McCaskill JS, Maeke T, Wagler P, in US Patent Publication,"Configurable Microreactor Network." Application-Number 09/935,740, 2000.

[8] van Noort D, Wagler P, McCaskill JS: Hybrid poly(dimethylsiloxane)-silicon microreactors used for molecular computing, Smart Materials and Structures, 2002; 11(5): 756-760.

[9] Wagler P, Gohlke M, van Noort D, McCaskill JS: Three-dimensional microfluidic systems for computation with DNA molecules. In 12. Micromechanics Europe workshop MME 2001. 2001. Cork, Ireland.

[10] McCaskill JS: Optically programming DNA computing in microflow reactors, Biosystems, 2001; 59(2): 125-138.

[11] Rasmussen S, Chen L, Nilsson M, Abe S: Bridging non-living and living matter, Artificial Life, 9. 2003; 269-316.

[12] Rasmussen S, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF and Bedau MA: Transitions from non-living to living matter. Science, 2004; 303: 963-965.

[13] Noireaux V, Libchaber A: A vesicle bioreactor as a step toward an artificial cell assembly. PNAS, 2004; 101(51): 17669-17674.

[14] Tan Y-C, Fisher JS, Lee AI, Cristini V, Lee AP: Design of microfluidic channel geometries for the control of droplet volume, chemical concentration, and sorting. LabChip, 2004; 4: p. 292-298 DOI: 10.1039/b403280m.

[15] Wagler P, Tangen U, Maeke T, Chemnitz S, Heymann M, Jünger M, Palutke T and McCaskill JS: Dynamical multiphase structures in the microsystem approach to artificial cells, to appear in microTas 2005, Boston

[16] Brask A, Goranovic G, Jensen MJ and Bruus H: A novel electroosmotic pump design for nonconducting liquids: theoretical analysis of flow-rate-pressure characteristics and stability, J. Micromech. Microeng. 2005; 15: 883-891.

[17] Wagler PF, Tangen U, Maeke T, Mathis HP, McCaskill JS: Microfabrication of a biomodule composed of microfluidics and digitally controlled microelectrodes for processing biomolecules. Smart Materials and Structures, 2003; 12(5): 757-762.

[18] Duffy DC, Cooper McDonald J, Schueller OJA, Whitesides GM: Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). Anal. Chem., 1998; 70(23): 4974-4984.

[19] Ocvirk G, Munroe M, Tang T, Oleschuk R, Westra K, and Harrison DJ: Electrokinetic control of fluid flow in native poly(dimethylsiloxane) capillary electrophoresis devices. Electrophoresis, 2000; 21(1): 107-115.

[20] Kirby BJ and Hasselbrink EF Jr.: Zeta potential of microfluidic substrates: 1. theory, experimental techniques, and effects on separations. Electrophoresis, 2004; 25: 187-202.

[21] Liu Y, Fanguy JC, Bledsoe JM, Henry CS: Dynamic coating using polyelectrolyte multilayers for chemical control of electroosmotic flow in capillary electrophoresis. Anal. Chem. 2000; 72(24): 5939-5944.

[22] Bianchi F, Wagner F, Hoffmann P, and Girault HH: Electroosmotic flow in composite microchannels and implications in microcapillary electrophoresis systems. Anal. Chem., 2001; 73(4): 829-836.

[23] Mao C, Sun W, Shen Z, Seeman, NC: A nanomechanical device based on the B-Z transition of DNA; Nature. 1999; 397: 144-146.

[24] Tangen U, Maeke T, McCaskill JS: Advanced simulation in the configurable massively parallel hardware: MereGen. in "3rd Caesarium 2000", Springer-Verlag. 2001; 107-108.

[25] Lin C-H, Fu L-M, Lee K-H, Yang R-J, Lee G-B: Novel surface modification methods and surface property analysis for separation of DNA bio-molecules using capillary electrophoresis. In microTAS, Squaw Valley, California USA, 2003; 1081-1084.

[26] Alur R, Courcoubetis C, Henzinger TA, Ho P-H: Hybrid automata: an algorithmic approach to the specification and verification of hybrid systems. Hybrid Systems, Lecture Notes in Computer Science, Springer-Verlag. 1993; 736: 209-229.

[27] Bujorianu ML, Lygeros J: General stochastic hybrid systems: modelling and optimal control. 43rd IEEE Conference on Decision and Control December 14-17, 2004

[28] Abate A, Ames AD, Sastry SS: Stochastic approximations of hybrid systems. To appear at 24th American Control Conference 2005 (ACC'05), Portland, OR, 2005

[29] Piazza C, Mishra B: Stability of hybrid systems and related questions from systems biology (In honor of Professor Pravin Varaiya on his 65th birthday) Systems & Control: Foundations & Applications, (Ed. T. Basar), Birkhauser, 2005. (In Press)

[30] http://www.cis.upenn.edu/group/mobies/charon

[31] http://control.ee.ethz.ch/~hybrid/hysdel

Container growth and replicator dynamics in pre-biotic chemistry

Olof Görnerup[1], Martin Nilsson Jacobi[1] and Steen Rasmussen[2,3]
[1] Department of Physical Resource Theory, Chalmers University of Technology, Sweden
[2] Self-Organizing Systems, EES-6, Los Alamos Laboratory, USA
[2] University of Copenhagen, IMBG, Denmark

At the transition from non-living to living matter, the distinction between genotype and phenotype, as well as genome replication and ontogeny, is not clearly defined. In primitive organisms, without central control of genome replication, a conflict between selfishly reproducing genes and genes useful for the replication of the whole organism occurs. This raises the question of how, and when, such systems can evolve into contemporary organisms with well-defined separation between the genotype and phenotype, and a coordinated replication.

We study the evolutionary dynamics of systems consisting of self-assembling container aggregates that contain populations of self-replicating information carrying molecules: proto-genes. The aggregates can be viewed at primitive proto-organisms. Their genome consists of an evolving population of proto-genes, which in a steady state may form a quasi-species. The aggregates themselves grow by successively incorporating new building blocks. Eventually the aggregates become unstable and spontaneously divide, whereby a replication of the proto-organism has occurred. The production of new building blocks (e.g. amphiphilic polymers) is controlled by the proto-genes (e.g. through an electron charge transfer process). A strain's ability to self-replicate and its chemical properties critical to the growth of the aggregate are assumed to be uncorrelated. Certain strains of proto-genes are efficient as self-replicators, whereas other strains are more active in the production of new building blocks, and thereby contribute to the reproduction of the container. The evolution of the system as a whole is then characterized by a conflict reminiscent of group selection. The central question is under which conditions selections favors co-existence of selfish genomes and genomes that are active in the growth of the aggregate.

This work was funded by PACE (Programmable Artificial Cell Evolution), a European Integrated Project in the EU FP6-IST-FET Complex Systems Initiative, and by EMBIO (Emergent Organisation in Complex Biomolecular Systems), a European Project in the EU FP6 NEST Initiative.

Behavior transitions provided by dynamical features of recurrent neural network - a case study of complex phenomena in behavior based robotics

Martin Hülse[1], Steffen Wischmann[1], and Frank Pasemann[1]

Fraunhofer Institute for Autonomous Intelligent Systems, Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Abstract. Complex phenomena like bistability, periodic and quasi-periodic oscillations, and chaos can be already observed in small artificial recurrent neural networks (RNN). We utilize this rich reservoir of dynamical properties for behavior control of autonomous mobile robots. Using a special evolutionary program, the ENS^3 (evolution of neural systems by stochastic synthesis), recurrent neural network structures of general type for robot control are developed. Structure and size of the evolving RNN are open to the evolutionary process, and parameters like synaptic strengths are optimized simultaneously.

The ENS^3 can also be utilized to extend or couple already existing RNNs to achieve additional functionality. In such a way different behavioral (sub)-functionality can be non-linearly integrated in one network solving a global robot task. On the one hand, this strategy allows an incremental evolution of complex control structures solving robot tasks including more and more subtasks and their effective coordination. On the other hand, such incrementally evolved RNNs give us a wide variety of empirical setups to investigate multifunctionality and robust behavior changes in complex systems provided by non-linear coupled neural systems.

As a first simple example of this method we present an incrementally evolved recurrent network generating a motivational driven robot behavior. Following an ALife approach the robot has to maintain the level of its internal energy reservoir. This energy reservoir decrease over time, but can be refilled from energy sources located in the environment. To maintain its energy level the robot has to switch between different behaviors. According to the dynamical features of the underlying network we observed that the transition from one behavior to another bases on the switching from periodic attractors to a domain of bistability and the otherway around. This attractor switching leads to increasing fluctuations on the macroscopic level, i.e. the observable robot-environment interaction. Therefore, we claim that behavior transitions are emergent phenomena generated by the attractor switching of the controller and the boundary conditions given by the robot-environment interaction.

Evolving artificial "brains": a biomimetic Evolutionary Neuro-Genetic Algorithm (ENGA) and its testing in simple games

PETER ITTZÉS, ZOLTAN SZATMÁRY, SZABOLCS SZÁMADÓ AND EÖRS SZATHMÁRY

Collegium Budapest (Institute for Advanced Study)

Running head: ENGA

2 Szentharomsag utca, H-1014 Budapest, Hungary
Phone: +36 1 2248300, Fax: +36 1 2248310
Email: szathmary@colbud.hu

Abstract

The evolution of autonomous communicating agents is a hot area in recent research. We have developed a software framework, called ENGA, to evolve flexible neuronal networks capable of solving first simple but then more and more complicated communication tasks. We took a biomimetic approach, that is, the design of the system was inspired by important biological phenomena such as brain ontogenesis, neuron morphologies, and indirect genetic encoding. We exposed our neuronal networks to simple coordination and task-allocation games to test the potential of the system. Neuronal networks were selected and were allowed to reproduce as a function of their performance in the given task. The selected neuronal networks in all scenarios were able to solve the communication problem they had to face. Although this result tells little about the full potential of the biomimetic approach, it is a promising start, on which we intend to build by introducing our neuronal networks to more and more complicated communication scenarios. The most striking feature of the model is that it works with highly indirect genetic encoding—just as brains do.

Key words

Biomimetic approach, neural networks, indirect encoding, origin of communication, artificial brain

Introduction

Emergence of communication systems in populations of agents is a field of active investigation [1-3]. Numerous approaches have been used and many of them include neural networks, evolutionary computing or both. Though these got their inspiration from biology, the actual level of biological realism varies, in some cases leaning toward extreme reductionism. The fact, that artificial systems up to the present day have proven to be inferior to natural systems in many aspects of communication, may suggest that biological phenomena should be imitated more closely. Of course, biology is a vast field ranging from biochemistry to evolutionary biology, and one would not want to

incorporate it all in a huge simulation. However it would be of great help to experiment with such a system and be able to identify which parts of it can safely be reduced to simpler mechanisms or representations while at the same time identify vital components without which the simulation would fail.

We have developed a software framework called Evolutionary Neurogenetic Algorithm (ENGA in short) which offers researchers a fine control over biological detail in their simulations. Our original intent was to create software with much potential for variability. That is, we wanted a piece of software which is general enough to allow for a wide range of experimentation but appears as a coherent system and does not fall apart into a loose set of unrelated pieces of code. This required careful specification and design; especially in partitioning it into modules and the specification of interfaces in a programme that has grown to about 10,000 lines of C++ code. A product release is expected in the future for the benefit of the community.

In such a short communication it is impossible to acknowledge all researchers of all important input fields to this paper. We have been especially influenced by evolutionary robotics, such as the work by Baldassare et al. [1], and by the evolutionary approach to neuronal networks with indirect encoding by Rolls and Stringer [4]. Our model is a recombinant of these approaches, with some key new elements, such as topographical network architecture.

Below we provide a brief enumeration of the supposedly most important properties of neuronal networks, genetic codes and development processes. This enumeration is far from complete but we have taken it into account when designing our system.

Genome

Scalability. If our goal is to develop an evolutionary scenario, which leads to a complexification process, then we have to be very careful with the coding system behind the neural system. The encoded phenotype information has to be indirect to avoid the negative effect of genetic linkage [5] and exponential growth of the genome size. Nature

has shown that the number of genes is orders of magnitudes less than the number of structural units of the phenotype. Even with 100 trillion neural connections in the human brain, there are only about 30 thousand active genes in the human genome [6].

Neuronal networks (phenotype)

Complex Network Properties. Recent research revealed that large-scale brain networks have interesting complex network properties. Scale-free and small-world properties, functional segregation (modularity) and functional integration might be necessary to achieve complex cognitive functions [7].

Degeneracy. A degenerate system, unlike a fully redundant one, is extremely adaptable to unpredictable changes [8].

Lamination. It was shown that lamination (as in the cortical layers) might be advantageous to support more complex information processing in the cortex [9].

Genotype-phenotype mapping and development

Synapse Development. If the scaling of a network causes problems, week or indirect mapping has to be applied [10]; it can be a graph-generating algorithm or a partly stochastic connection building process [11]. There is even a new subdiscipline of evolutionary computation that engages in the field of genotype-phenotype mapping in artificial systems called artificial embryogeny (AE) [6].

Learning. A learning phase in the development process can change the fitness landscape, and it can be as effective as a Lamarckian strategy at improving search [12]. Because synapse development is a non-deterministic process, learning plays a key role in synapse selection, too.

Synapse selection, pruning. One hypothesis for the importance of pruning is that if the synapses at first are overgrown and later pruned, the memory performance of the network is maximized in certain situations [13]. Another possible explanation for the importance of pruning arises from the developmental process. If the synapses join another neuron in

a stochastic way, optimal wiring should emerge after learning based pruning. Redundancy is a precondition to this process [14].

Specification and design

When specifying a piece of software we delineate what features we expect from the programme and how these features should behave. Specification involves decision making, and decisions are better not taken on an arbitrary basis. We took biomimesis as our guiding principle and finally a set of core assumptions and a complementing set of potential variabilities have emerged. The core assumptions summarize ideas we consider fixed throughout all simulations. Some of the most important such assumptions are listed in the following sections.

Population and agents

There is a population of agents. Agents have a genotype, a brain and a body. Bodies contain sensors and effectors and as such they provide the necessary input and output facilities for brains. Bodies reside in a simulated or real physical environment.

The genetic system

In general, information stored in a genotype will be used during phenotype creation, i.e. during ontogenesis. Genotype-phenotype mapping is usually a nontrivial task and decisions involving this process can severely influence the performance and capabilities of a simulation. If we design a coding system, which is not open-ended, evolution will stop at the boundary of the genetic space. One way to create an open-ended coding is to use a marker-based encoding scheme that was inspired by the structure of the real genetic code [15]. The other promising way is to encode information of the network in a tree structure [10]. In our system we chose a hierarchical tree-based representation. Genetic operators (point mutation, deletion, duplication, and recombination) work on this tree structure and alter the values of the nodes or change the structure of the tree. Duplication

ensures the open-ended behaviour of the system, and is able to scale up the search space, if it is required in a new situation. To reduce the negative effect of genetic linkage [5], the use of recombination is necessary. The optimal combination of these operators is an open question and it depends on the coding system and on the fitness landscape of the given task.

Neuronal networks

Neuron classes. Neurons are classified. Neurons belonging to the same class share the same functional and morphological properties, whereas neurons belonging to different classes may differ in these properties. Similarly, synapses belong to classes, too. We call them 'interclasses' because they represent functional properties that are shared by all synapses connecting a neuron belonging to a particular class A to a neuron belonging to class B (A and B can be the same class).

Topography. Unlike the majority of approaches to evolve neural networks we maintain the full topographic information of our networks, that is, neurons are situated in a layered neural space. Topographical information in a neural network can have a number of advantages. First, the interpretation of the structure can be easier. Second, developmental processes can model biological processes of neuron and synapse growth more accurately. Models which acknowledge spatial information in biological systems yield various scale-free and small-world network attributes like the ones that are common in brain structure [7].

Neuron Morphology. It was suggested that more structured neuron morphology is required for more complex cortical functions [16]. A more complex surface can facilitate more complex connection structure and if different neuron morphologies are encoded in the genome, the search process can change the network structure by changing morphology. Representation of neural networks in a topographic space allowed us to equip neurons with spatial dimensions. In our simulations neuron morphologies were only coarsely detailed and reduced to circles on each layer.

Tasks

In each simulation there is a task defined. Agents gain a so-called score (fitness) value according to their proficiency in the task. This fitness value can be used by death and reproduction processes (see Death-and-reproduction). Tasks can be solitary or social, reflecting the number of agents required to solve them. In case of social tasks the fitness of an agent may depend on the performance of its fellows as well.

Death and reproduction processes

There is a death process operating in the population. It takes out members from the population forever. Parallel to the death process a reproduction process is going on, replenishing the population by allowing some members of the population to reproduce. It is the responsibility of the reproduction process to stabilize the size of the population (this is necessary because of finite computing resources). The reproduction process may include a mating process (if the reproduction is sexual), selecting pairs from the population.

Lifecycle of agents

Following the initialization of the population agents start their lifecycle. It begins with the birth phase in which a body and a brain is generated. Only brain generation is affected by the agent's genotype. Brain ontogenesis consists of two steps: neurogenesis and synaptogenesis. The actual life of an agent can be separated into two parts: childhood and adulthood. Learning takes place during both ages, but pruning is present only during childhood and only adult members of the population can reproduce. The separation line between childhood and adulthood is dependent on the actual simulation and may completely disappear.

Ontogenesis of the brain. Development of the brain is controlled by the genotype in a highly indirect manner. By indirect we mean that no part of the genotype corresponds to individual neurons and synapses, and that only gross statistical properties of the brain are

encoded in the genes. Individual brains are sampled from this statistical description. The neurogenesis phase goes on as follows. Neurons are situated in a layered topographic neural space mimicking real cortical layers but low-level biological mechanisms shaping the cortex such as concentration gradient dependence in neurogenesis are not present in the simulation. Instead, neuron classes define a probability density function over the neural space from which individual neuron soma positions are sampled. Neurons possess morphologies, i.e. there is some function over neural space describing their dendritic arborization. This morphology is applied to each neuron relative to its sampled soma position. Synaptogenesis exploits two mechanisms, just as in biology: a long-range one (called projections) and a subsequent short-range one (lock-and-key mechanism). Each neuron class has an associated list of projections. Projections are probability density functions over neural space. They can be defined either in absolute coordinates or in coordinates relative to a neuron's soma. When a neuron's efferents are to be determined, putative synapse locations are sampled from its projections (determined by its neuron class). Neurons having dendritic arborization near these putative locations become candidate efferents. Then the short-range mechanism selects from competing candidates at each synapse location. The short range lock-and-key mechanism mimic the receptor-ligand based binding mechanisms present in real synaptogenesis. Locks and keys are 30 long strings of bits. Every candidate postsynaptic neuron's lock is matched to the presynaptic neuron's key. Binding probability is then a decreasing function of the Hamming distance between the key and the lock. The complete ontogenetic algorithm is depicted in Fig. 1.

Potential for variation

Potential for variation arises within the boundaries of the core assumptions. For example, many kinds of activity models and dynamics can be imagined: rate code or spikes as activity model and summation of synaptic contributions fed through various transfer functions as activity dynamic. Of course, variations of core concepts are not necessarily compatible with each other. If, for example, a synapse model expects the presynaptic neuron activity to be rate coded then somehow it must be ensured that the presynaptic

neuron class really defines its activity model as rate code. ENGA provides some mechanisms to enforce such compatibilities but eventually it is always the person designing the simulation who is responsible for the overall coherence.

Programme architecture

The architecture is of layered nature so that lower modules do not know about the existence of higher modules (Fig. 2). This allows easy modifiability of higher levels without the need to modify lower levels. Moreover, each layer exposes an interface that can be used by any client, even those deviating from the original purpose of simulating evolution of embodied communicating agents. The genetic module for example can be used in any evolutionary computation, not only those evolving artificial neural networks. We may as well talk about a multilevel software framework consisting of several modules that can be used individually or in combination with others to produce various kinds of evolutionary and neural computation related simulations.

Test tasks: simple coordination and task-allocation games

We have used simple 2x2 matrix games with one-way pre-play communication as test tasks: a coordination and a task-allocation game. There are two reasons we have chosen these games. First, because of their simplicity these scenarios are easy to implement. Second, these games are well studied in the economics literature, and it is shown both theoretically and empirically that (one-way) cost-free, pre-play communication (so called ‘cheap talk’) can resolve coordination problems in both coordination and task-allocation games [17-18]. Thus it is reasonable to expect that our agents will evolve the use of signals under these scenarios. To test further the flexibility of the neuronal architecture of our agents we have introduced two environments, which have an influence on the pay-off matrices.

Accordingly we have four different scenarios:

I. Control. The fitness of a given agent is independent from the choice of behaviour of the other agent in the game.
II. Coordination game. Agents have to pick the same response, corresponding to the given environment to get high fitness.
III. Task-allocation game. Agents have to pick complementary responses to get high fitness.
IV. Coordination or task-allocation depending on the environment.

There are: two kinds of environment, $E \in \{-1, 1\}$, a continuous signal S which can take values in [-1,1], and a continuous response function R which can take values in [-1,1]. For signal and response functions only the -1 and 1 values are taken as correct signal or response, all other values are taken as nonsense signal (NS) or nonsense response (NR), respectively. Nonsense signal and response are not rewarded in any of the following games.

I. Control

The fitness of a given agent is independent of the response of the other agent in the game; it only depends on the environment. In case of $E(1)$ the good response is $R(1)$, in case of $E(-1)$ the good response is $R(-1)$.

II. Coordination game

A pair of agents plays a coordination game in which they have to choose a given response conditional on the environment. (i) If both agents pick the correct response they both get a high reward; (ii) if only one of them picks the correct response then one gets a higher reward than the other; (iii) if both of them picks the wrong response then they both get low rewards.

III. Task-allocation game

A pair of agents plays a task-allocation game in which they have to choose complementary responses to get high fitness. If they pick the same response they get low reward regardless of the environment.

IV. Coordination/Task allocation game

A pair of agents plays a coordination or a task-allocation game depending on the state of the environment. In case of *E*(1) they play a task-allocation game, in case of *E*(-1) they play a coordination game.

Steps of the game

1. Mother Nature picks an environment: one of the environments is being selected randomly.
2. Observation: one randomly chosen agent out of each pair can observe the environment, that is, he has correct information about the type of the environment in which the game is situated.
3. Communication: the observer may or may not give any of the signals (*S*) available to it.
4. Picking the response: after the observer gave his signal both agents pick a response (*R*) simultaneously. The fitnesses of the agents are calculated from the pay-off matrices according to the game and to the given environment.

Steps of the simulation

1. The simulation is initialised with *M* naive agents.
2. Each turn of the simulation consists of two stages: playing the game (each agent plays with every other agent) and reproduction.
3. Reproduction: The agent with the least fitness dies. Reproduction is sexual, thus we pick a pair of individuals to produce an offspring that will fill the vacant

position. The chance of an agent being picked to reproduce is proportional to its fitness. After reproduction we return to stage one.

4. The simulation is iterated for *t* turns.

Results

Our agents have achieved a better-than-random performance in all scenarios which suggests that they have managed to communicate sufficient amount of information to solve the task under question. In case of the control probably only as a side-effect, given that signals are cost-free (i.e. not selected against even if they confer no benefit to its users).

The average fitness as a function of time in each population is depicted in Fig 3. One can see that the slowest increase in fitness can be observed in the population which played either a coordination or a task-allocation game, depending on the environment. Finally, it is worth mentioning the topology of the evolved networks. Figs 4 and 5 depict an 'early', and an 'advanced' neuronal network taken at the 10th and at the 750th generation out of the population playing the 4th scenario (coordination or task allocation game depending on the environment). As we can see selection had a profound effect on the topology. While the first network failed to make connections to decision output and is doomed to low fitness values, the later 'advanced' network utilizes both visual and audio input to determine the values of decision and signal outputs.

Discussion

The results of the games analysed in this paper suggest that the basic idea—the biomimetic approach—behind our neuro-genetic system is correct, and the highlighted design principles of the system are sufficient at least in this simple task. The other goal of the test was to prove that the developed software is working properly is useful for scientific experimentation. The core of our neuro-genetic system is the stochastic ontogenetic process and the indirect tree-based coding behind it. Evolution of agents in

these scenarios found the right solution in a relatively short period, and the length of the search time was proportional to the complexity of particular game. The topology and size of the networks both changed in the simulations. It is important to emphasize that we have not taken advantage of any grossly unnatural (but sometimes very effective) algorithms such as back propagation.

Our results show that indirect encoding and a genetically controlled but stochastic ontogenetic process together can provide an appropriate framework for evolving communicating agents in simple scenarios. Our next goal is to introduce our neuronal networks to more and more complicated selective scenarios, by which we hope to obtain better understanding of the evolution of, among other things, symbolic communication.

Acknowledgement

This paper was supported by the IP ECAgents (http://ecagents.istc.cnr.it) in the 6th Framework Programme. Comments by Chrisantha Fernando are gratefully acknowledged.

1. Baldassarre G, Nolfi S, Parisi D: Evolving mobile robots able to display collective behavior. Artif Life 2003;9: 255-267.
2. Steels L: The Talking Heads Experiment.Words and Meanings. Antwerpen, 1999.
3. Kvasnicka V, Pospichal J: An emergence of coordinated communication in populations of agents. Artif Life 1999;5: 319-342.
4. Rolls ET., Stringer SM. On the design of neural networks in the brain by genetic evolution. Prog Neurobiol 2000;61: 557-579.
5. Calabretta R, Di Ferdinando A, Wagner GP, Parisi D: What does it take to evolve behaviorally complex organisms? BioSystems 2003;69: 245-262
6. Stanley KO, Miikkulainen R: A taxonomy for artificial embryogeny. Artif Life. 2003;9: 93-130.

7. Sporns O, Chialvo DR, Kaiser M, Hilgetag, CC: Organization, development and function of complex brain networks. Trends Cogn Sci 2004;8: 418-425.
8. Tononi G, Sporns O, Edelman, GM: Measures of degeneracy and redundancy in biological networks. Proc Natl Acad Sci 1998; 96: 3257-3262.
9. Montagnini A, Treves, A. (2003) The evolution of mammalian cortex, from lamination to arealization. Brain Res Bull 2003;60: 387-393.
10. Zhang B, Mühlenbein H: Genetic programming of minimal neural nets using Occam's razor. Proc. of the Fifth Int. Conf. on Genetic Algorithms, 1993: 342-349.
11. Balakrishnan K, Honavar V: Evolutionary design of neural architectures: A preliminary taxonomy and guide to literature. Technical report 1995; CS-TR-95-01, AI. Lab, Dept. of Computer Science, Iowa StateUniversity.
12. Gruau F, Whitley, D: Adding learning to the cellular development of neural networks: evolution and the Baldwin effect. Evol Comput 1993;1: 213-233.
13. Chechik G, Meilijson I, Ruppin E: Synaptic pruning in development: a computational account. Neural Comput 1998;10: 1759-1777.
14. Kerszberg M, Dehaene S, Changeux JP. Stabilization of complex input-output functions in neural clusters formed by synapse selection. Neural Networks 1992;5: 403-413.
15. Fullmer B, Miikkulainen: Using marker-based genetic encoding of neural networks to evolve finite-state behaviour. Proc. of the First European Conference on Artificial Life. Paris. 1991.
16. Elston GN, Benavides-Piccione R, DeFelipe J: The Pyramidal Cell in Cognition: A Comparative Study in Human and Monkey. J. Neurosci. 2001;21: RC163 1-5.
17. Cooper R, DeJong DV, Forshyte R, Ross, TW: Communication in Coordination Games. Quart J Econ 1992;107: 739-71.
18. Hurkens S, Schlag, KH. Communication, Coordination, and Efficiency in Evolutionary One-Population Models. J Econ. Literature 1999;

LEGENDS

Fig 1. Neurogenesis and long range mechanism of synaptogenesis
(a) Probability distribution of positions. (b) Sampled soma positions. (c) Morphology added to soma. (d) Projection (on layer 2) from a presynaptic cell marked with blue. (e) Putative synapse location (X mark) is sampled from projection. (f) Candidate postsynaptic cells (marked with magenta) are determined.

Fig. 2. The architecture of the programme
Modularity of the different component is apparent.

Fig. 3. Average fitness as a function of time in all 4 scenarios
Orange: control, green: coordination game, red: task allocation game, blue: coordination or task allocation game depending on the environment.

Fig. 4. Topology of an early, incompetent neuronal network
Colour code highlights the flow of information of the visual and audio input from the environment. When these channels join on a particular neuron, mixed colour indicates joint information processing.

Fig. 5. Topology of a late, advanced neuronal network
Colour code as in previous figure.

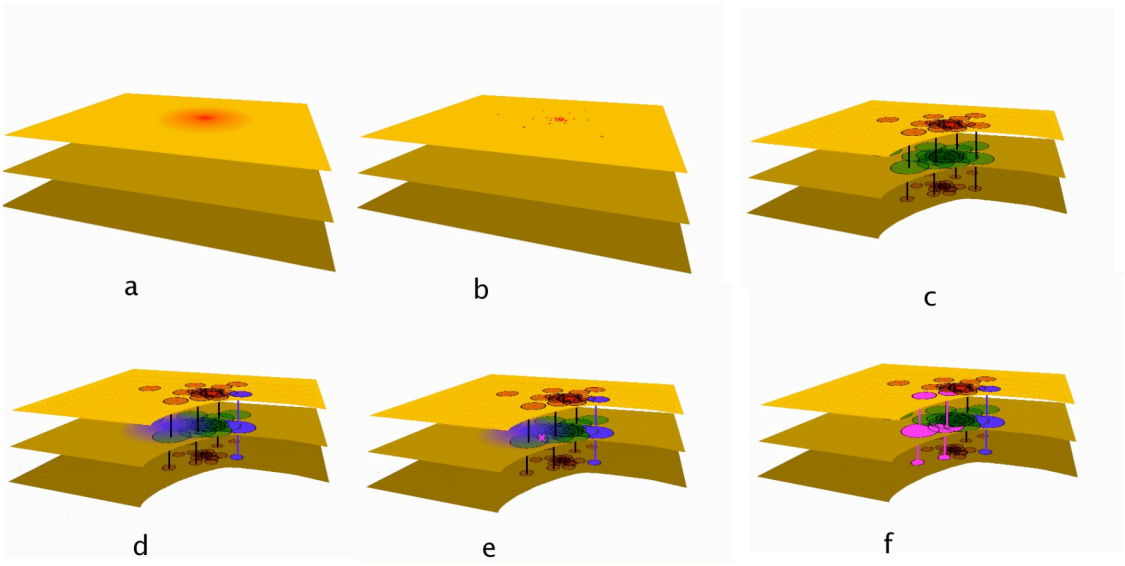


Figure 1.

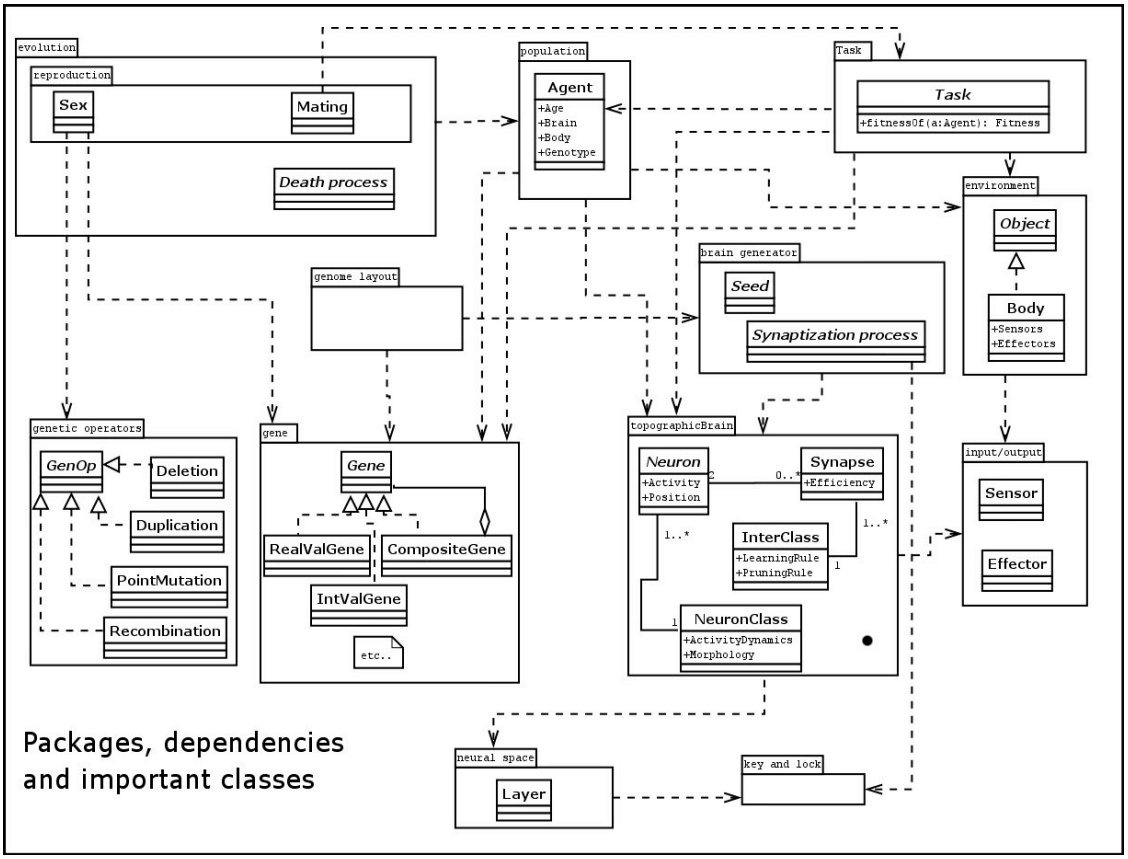

evolution
reproduction
Sex
Mating
Death process
population
Agent
+Age
+Brain
+Body
+Genotype
Task
Task
+fitnessOf(a:Agent): Fitness
environment
Object
Body
+Sensors
+Effectors
brain generator
Seed
Synaptization process
genome layout
genetic operators
GenOp
Deletion
Duplication
PointMutation
Recombination
gene
Gene
RealValGene
CompositeGene
IntValGene
etc..
topographicBrain
Neuron
+Activity
+Position
Synapse
+Efficiency
InterClass
+LearningRule
+PruningRule
NeuronClass
+ActivityDynamics
+Morphology
input/output
Sensor
Effector
neural space
Layer
key and lock
Packages, dependencies
and important classes

Figure 2.


average fitness per runs
3500
3000
2500
2000
1500
1000
500
0
average fitness
1
51
101
151
201
251
301
351
time

Figure 3.

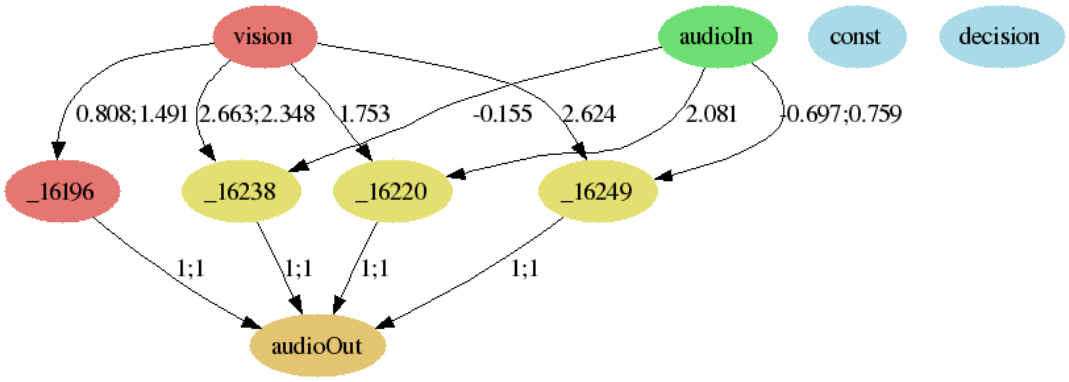

vision
audioIn
const
decision
0.808;1.491
2.663;2.348
1.753
-0.155
2.624
2.081
0.697;0.759
_16196
_16238
_16220
_16249
1;1
1;1
1;1
1;1
audioOut

Figure 4.

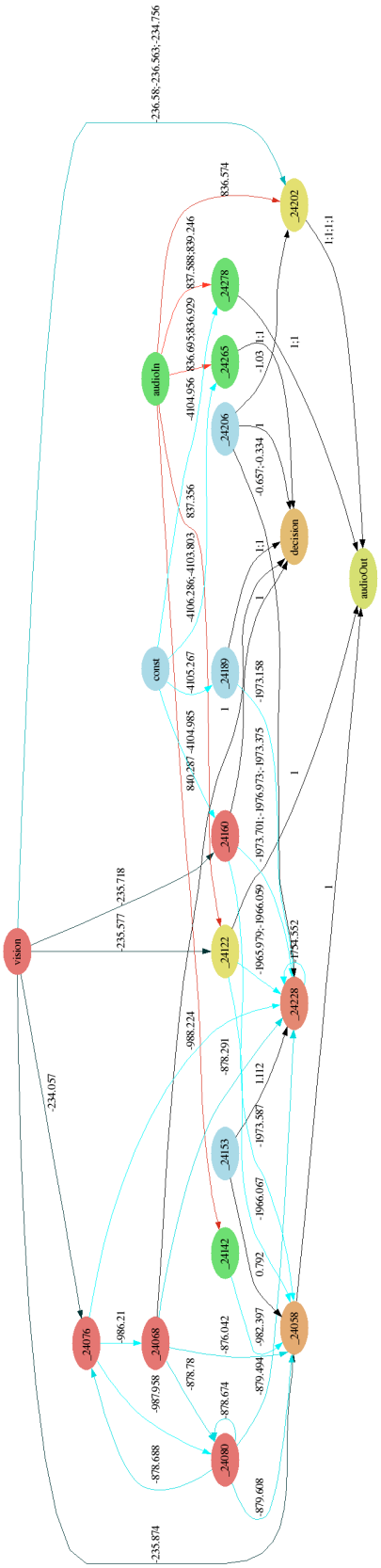

-236.58;-236.563;-234.756
836.574
24202
1;1;1;1
837.588;839.246
24278
836.695;836.929
audioIn
24265
1;1
1.03
1;1
-4104.956
24206
-0.657;-0.334
1
837.356
decision
audioOut
-4106.286;-4103.803
1;1
const
-4105.267
24189
-1973.158
1
-4104.985
840.287
-1973.375
-1976.973;-1973.701;
24160
1
1
235.718
235.577
-1966.059
-1965.979;
24122
-1754.552
24228
1
-988.224
-878.291
1.112
-234.057
-1973.587
24153
-1966.067
24142
0.792
-986.21
-982.397
-876.042
24076
24068
24058
-987.958
-878.78
-879.494
-878.674
-878.688
24080
-879.608
-235.874

Figure 5.

Complex Systems Methods

Message passing algorithms for non-linear nodes and data compression

Stefano Ciliberti* and Marc Mézard†

Laboratoire de Physique Théorique et Modèles Statistiques,

Université de Paris-Sud, bâtiment 100, 91405, Orsay Cedex, France

Riccardo Zecchina‡

ICTP, Strada Costiera 11, I-34100 Trieste, Italy

Abstract

The use of parity-check gates in information theory has proved to be very efficient. In particular, error correcting codes based on parity checks over low-density graphs show excellent performances. Another basic issue of information theory, namely data compression, can be addressed in a similar way by a kind of dual approach. The theoretical performance of such a Parity Source Coder can attain the optimal limit predicted by the general rate-distortion theory. However, in order to turn this approach into an efficient compression code (with fast encoding/decoding algorithms) one must depart from parity checks and use some general random gates. By taking advantage of analytical approaches from the statistical physics of disordered systems and SP-like message passing algorithms, we construct a compressor based on low-density non-linear gates with a very good theoretical *and* practical performance.

*Electronic address: `ciliberti@roma1.infn.it`

†Electronic address: `mezard@lptms.u-psud.fr`

‡Electronic address: `zecchina@ictp.trieste.it`

I. INTRODUCTION

Information theory and statistical physics have common roots. In the recent years, this interconnection has found further confirmation in the analysis of modern error correcting codes, known as Low Density Parity Check (LDPC) codes [1–3], by statistical physics methods [4–6]. In such codes, the choice of the encoding schemes maps into a graphical model: the way LDPC codes exploit redundancy is by adding bits which are bound to satisfy some random sparse linear set of equations – the so called parity checks. Such equations are used in the decoding phase to reconstruct the original codeword (and hence the message) from the corrupted signal. The parity checks can be represented by a graph indicating which variables participate in each check. The randomness and the sparsity of the parity checks is reflected into the characteristics of the associated graphs, a fact that makes mean-field type statistical physics methods directly applicable. Thanks to the duality between channel coding and source coding [3], similar constructions can be used to perform both lossless and lossy data compression [7, 8]. It must be mentioned that while there exist some practical algorithms for the lossless source coding which are very efficient [9], much less work has been done so far for the lossy case [10]. In this work we present a new coding technique which consists in a generalization of parity-check codes for lossy data compression to non-linear codes [11]. This issue has been addressed recently using methods from the statistical physics of disordered systems: a non-linear perceptron [12] or a satisfiability problem [13] have been used to develop practical source coding schemes. Other recent advances in the field can be found in [14–16].

We consider the following problem. We have a random string of uncorrelated bits $x_1, \dots x_M$ ($\mathrm{prob}(x_a = 0) = \mathrm{prob}(x_a = 1) = 1/2$) and we want to compress it to the shorter coded sequence $\sigma_1, \dots \sigma_N$ (encoding). The rate R of the process is $R = N/M$. Once we recover the message (decoding), we may have done some errors and thus we are left in principle with a different sequence $x_1^*, \dots x_M^*$. The number of different bits normalized by the length of the string is the distortion,

$$D = \frac{1}{M} \sum_{a=1}^{M} (1 - \delta(x_a, x_a^*)) \ . \tag{1}$$

The Shannon theorem states that the minimum rate at which we can achieve a given distortion D is $R_{min} = 1 - H_2(D)$, where $H_2(D)$ is the binary entropy of the probability

distribution used to generate the x_a's. This R_{min} is the value one has to compare to in order to measure the performance of a compression algorithm. This problem is a simple version of the lossy data compression problem. In real applications one is often more interested in compression through quantization of a signal with letters in a larger alphabet (or continuous signal [17]). Here we shall keep to the memoryless case of uncorrelated unbiased binary sources, and we hope to be able to generalize the present approach to more realistic cases in the near future.

A. Constraint Satisfaction Problems

The underlining structure of the protocol that we are going to introduce is provided by a constraint satisfaction problem (CSP). A CSP with N variables is defined in general by a cost function

$$E[\sigma] = \sum_{a=1}^{M} \varepsilon_a \left(\{\sigma_i \in a\} \right) \ , \tag{2}$$

where the *function node* $\varepsilon_a(\sigma_1, \ldots \sigma_{K_a})$ involves a subset of K_a variables randomly chosen among the whole set. The problem is fully defined by choosing some ε_a and an instance is specified by the actual variables involved in each constraint. A useful representation for a CSP is the factor graph (Fig. 1, left). The two kinds of nodes in this graph are the variables and the constraints. For the sake of simplicity we shall take the functions ε_a to have values in $\{0, 2\}$ (the clause is resp. SAT or UNSAT) and the variables σ_i will be Ising spins ($\sigma_i = \pm 1$). A somewhat special case of CSP which has been studied extensively occurs when $K_a = K$ for any a, and we shall restrict ourselves to this case in what follows. One can then show that the degree of a given variable is Poisson distributed with mean value KM/N and that the typical length of a loop is $\mathcal{O}(\log N / \log(KM/N))$. We are interested in the limit $M, N \to \infty$, where $\alpha \equiv M/N$ is fixed and plays the role of a control parameter. Once a problem is given in terms of its graph representation, an instance corresponds to a given graph chosen among all the legal ones with uniform probability. We will go back to this problem in section II. We first show how a CSP can be used as a tool for compressing data.

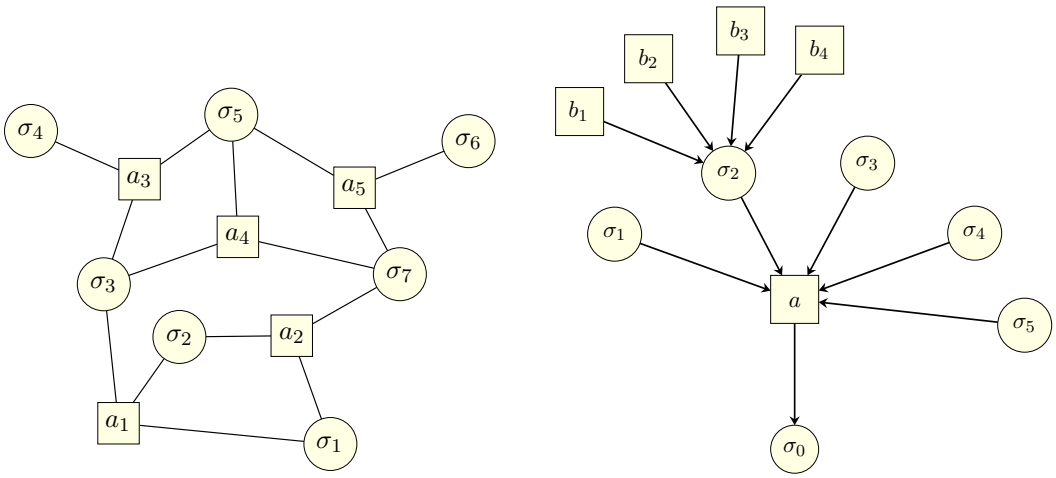


FIG. 1: Left: Factor graph for a constraint satisfaction problem with $N = 7$ variables (circles) and $M = 5$ constraints (squares). In this example the connectivity of the function nodes is kept fixed, that is $K_a = K$ in the notations of the text. Right: Message passing procedure ($K = 6$): The probability $q_{a\to 0}(u_{a\to 0})$ depends on all the probabilities $q_{b_i\to i}(u_{b_i\to i})$, $i = 1, \dots 5$ (see text).

B. Encoding

We use the initial word $x_1, x_2, \dots x_M$ to construct a set of M constraints between N boolean variables $\sigma_1, \sigma_2, \dots \sigma_N$. The factor graph is supposed to be given. Each constraint a involves K_a variables and is defined by two complementary subsets of configurations $\mathcal{S}_0^a$ and $\mathcal{S}_1^a$. The value of x_a controls the role of the subsets as follows: If $x_a = 0$ then all the configurations $\{\sigma_1 \dots \sigma_{K^a}\} \in \mathcal{S}_0^a$ satisfy the constraint, *i.e.* $\varepsilon_a = 0$; all the configurations $\{\sigma_1 \dots \sigma_{K^a}\} \in \mathcal{S}_1^a$ do not satisfy the constraint and thus $\varepsilon_a = 2$. If $x_a = 1$ the role of the two subsets is exchanged. This defines the CSP completely and we can look for a configuration of σ_i that minimizes the number of violated constraints. This is the encoded word and the rate of the process is $R = 1/\alpha$.

C. Decoding

Given the configuration $\sigma_1, \sigma_2, \dots \sigma_N$, we compute for each node a whether the variables $\sigma_{i_1^a}, \dots \sigma_{i_{K_a}^a}$ are in $\mathcal{S}_0$ (in this case we set $x_a^* = 0$) or in $\mathcal{S}_1$ (leading to $x_a^* = 1$). Since a cost

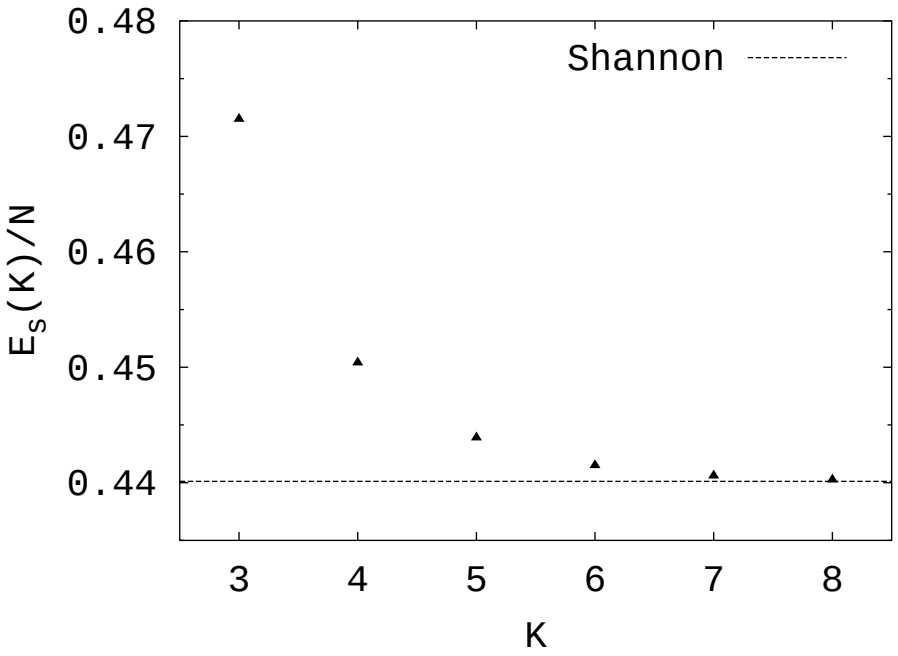


FIG. 2: The ground state energy density for the XORSAT problem at $\alpha = 2.0$ versus K. We also show the corresponding Shannon bound as computed from rate-distortion theory, $E_{Sh} = 2\alpha D_{Sh}$, where D_{Sh} is such that $1 - H_2(D_{Sh}) = 1/\alpha$.

$\varepsilon = 2$ is paid for each UNSAT clause, the energy is related to the distortion by

$$D = \frac{E}{2M} = \frac{E}{2\alpha N} = R\frac{E}{2N} \ . \tag{3}$$

D. The Parity Source Coder

In order to give an example, we apply these ideas to the case where the CSP is a parity check (the optimization problem is then called XORSAT). In other words, the cost of the constraint a is written as

$$\varepsilon_a = 1 + (-1)^{x_a} \prod_{i \in a} \sigma_i \ . \tag{4}$$

This Parity Source Coder (PSC) is the counterpart of the LDPC codes in error-correcting codes and then it seems natural to expect a good performance from it.

In Fig. 2 we show that the ground state energy of the XORSAT problem, that is the theoretical capacity of the PSC, quickly approaches the Shannon bound as K increases [18]. A PSC with checks of degree $K \gtrsim 6$ has then a theoretical capacity close to the optimal one. The problem here is that there does not exist any fast algorithm that can encode a

string (for example, the SID algorithm that we discuss in the next section is known not to converge).

We will thus investigate a new kind of constraint satisfaction problem, using non-linear nodes, with (almost) the same good theoretical performance as the XORSAT problem, but with a fast encoding algorithm. Before we do this, we introduce in the next section the general formalism used in the study of non-linear nodes.

II. THE GENERAL FORMALISM

Given a CSP at some α, we are interested in knowing whether a *typical* instance of the problem is satisfiable (i.e. all the constraints can be satisfied by one global configuration) as the number of variables goes to infinity. Generally speaking, there will be a phase transition between a SAT regime (at low α the problem has a small number of constraints and thus is solvable – at least in principle) and an UNSAT regime where there are too many constraints and one cannot find a zero-energy configuration. In this UNSAT regime, one wants to minimize the number of violated constraints. This kind of problem can be approached by message passing algorithms.

Due to the large length of typical loops, the local structure of a typical graph is equivalent to a tree, and this is crucial in what follows. We introduce the *cavity bias* $u_{a\to i} \in \{-1, 0, +1\}$ sent from a node a to a variable i. A non-zero message means that the variable i is requested to assume the actual value of $u_{a\to i}$ in order to satisfy the clause a. If $u_{a\to i} = 0$ the variable i is free to assume any value. It is clear that this message sent to i should encode the information that a receives from all the other variables attached to it. In order to clarify this point, we refer to Fig. 1 (right), that is we focus on a small portion of the graph. As the graph is locally tree-like, the variables σ_i, $i = 1, 2, ..., K-1$, are only connected through clause a if N is large enough. If a is absent, the total energy of the system can be written as $E^N(\sigma_1, ...\sigma_{K-1}) = A - \sum_{i=1}^{K-1} h_i\sigma_i$. We have used here the assumption that the probability of these σ_i factorizes. This is again motivated by the local tree structure, but we shall go back to this point. After clause a is added, $E^{N+1}(\sigma_0, \sigma_1, ...\sigma_{K-1}) = E^N(\sigma_1, ...\sigma_{K-1}) + \varepsilon_a(\sigma_0, \sigma_1, ...\sigma_{K-1})$ and the variables rearrange in order to minimize the total cost. The minimization then defines $\tilde{\varepsilon}(\sigma_0)$

from

$$A - \sum_{i=1}^{K-1} |h_i| + \tilde{\varepsilon}(\sigma_0) = \min_{\sigma_1,\dots\sigma_{K-1}} E^{N+1}(\sigma_0, \sigma_1, \dots\sigma_{K-1}) \ . \tag{5}$$

This $\tilde{\varepsilon}(\sigma_0)$ is then the cost to be paid for adding one variable with a fixed value σ_0. Without losing generality, it can be written as

$$\tilde{\varepsilon}(\sigma_0) \equiv \Delta_{a\to 0} - \sigma_0 u_{a\to 0} \ , \tag{6}$$

where $u_{a\to 0}$ is the cavity bias acting on the new variable and $\Delta_{a\to 0}$ is related to the actual energy shift by

$$\Delta E \equiv \min_{\sigma_0,\dots\sigma_{K-1}} \left[E^{N+1}(\sigma_0, \sigma_1, \dots \sigma_{K-1}) - E^N(\sigma_1, \dots \sigma_{K-1})\right] = \Delta_{a\to 0} - |u_{a\to 0}| \ . \tag{7}$$

Given that ε_a can be 0 or 2, depending on the set of fields $h_1, \dots h_{K-1}$ one has four possibilities:

$$\tilde{\varepsilon}(+1) = 0 \quad \text{and} \quad \tilde{\varepsilon}(-1) = 0 \ \Rightarrow \ u = 0 \ , \quad \Delta = 0 \tag{8}$$

$$\tilde{\varepsilon}(+1) = 0 \quad \text{and} \quad \tilde{\varepsilon}(-1) = 2 \ \Rightarrow \ u = +1 \ , \Delta = 1 \tag{9}$$

$$\tilde{\varepsilon}(+1) = 2 \quad \text{and} \quad \tilde{\varepsilon}(-1) = 0 \ \Rightarrow \ u = -1 \ , \Delta = 1 \tag{10}$$

$$\tilde{\varepsilon}(+1) = 2 \quad \text{and} \quad \tilde{\varepsilon}(-1) = 2 \ \Rightarrow \ u = 0 \ , \quad \Delta = 2 \ . \tag{11}$$

In other words, a non-zero message u is sent from clause a to variable σ_0 only if the satisfiability of clause a depends on σ_0. A null message ($u = 0$) can occur in the two distinct cases (8) and (11).

The main hypothesis we have done so far consisted in assuming that the two variables are uncorrelated if they are distant (the energy is linear in the σ_i's). It turns out that, in a large region of the parameter space [19, 20], including the regime we are interested in, this is not correct. This is due to the fact that the space of solutions breaks into many disconnected components if α is greater than a critical value. In order to deal with this case, one has to introduce, for each directed link, a probability distribution of the cavity biases, namely $\mathsf{q}(u) \equiv \eta^+\delta_{u,+1} + \eta^-\delta_{u,-1} + (1 - \eta^+ - \eta^-)\delta_{u,0}$. The hypothesis of no correlation holds if the phase space is restricted to one component. The interpretation of $\mathsf{q}_{a\to i}(u_{a\to i})$ is the probability that a cavity bias $u_{a\to i}$ is sent from clause a to variable i when one component is picked at random [21]. According to the rules (8,9,10,11), and with the topology of Fig. 1

(right) as a reference for notations, the Survey Propagation (SP, [22]) equations are then

$$\eta^{+}_{a\to 0} \propto \text{Prob}\left[\left\{\left(u_{b\to i_1}\right)_{b\in i_1\setminus a},\dots\left(u_{b\to i_{K-1}}\right)_{b\in i_{K-1}\setminus a}\right\}\,\middle|\,\left(\tilde{\varepsilon}_a(+1)<\tilde{\varepsilon}_a(-1)\right)\right]e^{-y\Delta E}\ , \quad (12)$$

$$\eta^{-}_{a\to 0} \propto \text{Prob}\left[\left\{\left(u_{b\to i_1}\right)_{b\in i_1\setminus a},\dots\left(u_{b\to i_{K-1}}\right)_{b\in i_{K-1}\setminus a}\right\}\,\middle|\,\left(\tilde{\varepsilon}_a(+1)>\tilde{\varepsilon}_a(-1)\right)\right]e^{-y\Delta E}\ , \quad (13)$$

$$\eta^{0}_{a\to 0} \propto \text{Prob}\left[\left\{\left(u_{b\to i_1}\right)_{b\in i_1\setminus a},\dots\left(u_{b\to i_{K-1}}\right)_{b\in i_{K-1}\setminus a}\right\}\,\middle|\,\left(\tilde{\varepsilon}_a(+1)=\tilde{\varepsilon}_a(+1)\right)\right]e^{-y\Delta E}\ , \quad (14)$$

where the energy shift ΔE is given in eq. (7) and is non-zero only when the constraint is UNSAT for any value of σ_0 (cf. eq. (11)). The crucial reweighting term $\exp(-y\Delta E)$ thus acts as a "penalty" factor each time a clause can not be satisfied. This term is necessary in the UNSAT regime which we explore here (while simpler equations with $y=\infty$ are enough to study the SAT phase).

A. Encoding: SP and decimation

For each fixed value of y, the iterative solution for the SP equations can be implemented on a single sample [22], *i.e.* on a given graph where we know all the function nodes involved. The cavity probability distributions $\mathsf{q}(u)$'s are updated by picking up one edge at random and using (8,9,10,11). This procedure is iterated until convergence. This yields a set of messages $\{\eta^{-}_{a\to i},\eta^{0}_{a\to i},\eta^{+}_{a\to i}\}$ on each edge of the factor graph which is the solution of the SP equations. This solution provides very useful information about the single instance that can be used for decimation. As explained in [23], the finite value of y used for this purpose must be properly chosen. Given the solution for this y, one can compute the distribution $P(H_i)$ of the total bias $H_i=\sum_{a\in i}u_{a\to i}$ on each variable. One can then fix the most biased variables, *i.e.* the one with the largest $|P(H_i=+1)-P(H_i=-1)|$, to the value suggested by the $P(H_i)$ itself. This leads to a reduced problem with $N-1$ variables. After solving again the SP equations for the reduced problem, the new most-biased variable is fixed and one goes on until the problem is reduced to an "easy" instance. This can be finally solved by some conventional heuristic (*e.g.* walksat or simulated annealing). A significant improvement of the decimation performance can be obtained by using a backtracking procedure [23, 24]: At each step we also rank the fixed variables with a strongly opposed bias and unfix the most "unstable" variable with finite probability. The algorithm described here is called Survey Inspired Decimation (SID) and its peculiar versions have been shown to be very useful in

many CSP problems recently. Unfortunately, the basic version of SID does not work for the XORSAT problem because of the symmetric character of the function nodes that is reflected in a large number of unbiased variables (some improvements appear to be possible [25]).

B. Statistical analysis: Theoretical performance

One can perform a statistical analysis of the solutions of the SP equations by population dynamics [21]. The knowledge of the function node ε_a allows to build up a table of values of u for each configuration of the local fields $h_1, \ldots, h_{K-1}$ (this is done according to the minimization procedure described above). We then start with some initial (random) population of $\vec{\eta}_i \equiv \{\eta_i^+, \eta_i^0, \eta_i^-\}$, for $i = 1, \ldots N$. We extract a Poisson number p of neighbors and p probabilities $\vec{\eta}_{i_1}, \ldots \vec{\eta}_{i_p}$. According to these weights, p biases $u_1, \ldots u_p$ are generated and their sum computed, $h_i = \sum_{a=1}^{p} u_a$. Once we have $K-1$ of these fields we perform the minimization in (5) and compute the new probability for u_a according to the rules (8,9,10,11). The whole process is then iterated until a stationary distribution of the cavity biases is reached. This method for solving the SP equations is very flexible with respect to the change of the choice of the node, because this choice just enters in the calculation once at the beginning of the algorithm in order to initialize some tables. One can also study problems with many different types of nodes in a given problem: in this case, one among them is randomly chosen each time the updating is performed. We finally stress that once the probability distribution of the cavity biases is known the ground state energy of the problem can be computed according to the formalism introduced in [21]. The expression for the energy in term of the probability distributions as

$$E(\alpha) = \max_y \Phi(\alpha, y) \ , \tag{15}$$

$$\Phi(\alpha, y) = \frac{(-1)}{y} \left\{ [1 + (K-1)\alpha] \overline{\log A_p(y)} - (K-1)\alpha \overline{\log A_{p+1}(y)} \right\} \ , \tag{16}$$

$$A_p(y) \equiv \sum_{u_1, \ldots u_p} \mathsf{q}_1(u_1) \cdots \mathsf{q}_p(u_p) \exp\left(y \Big| \sum_{a=1}^{p} u_a \Big| - y \sum_{a=1}^{p} |u_a| \right) \ , \tag{17}$$

the overline standing for an average over the Poisson distribution and the choice of the distributions $\mathsf{q}_1, \ldots \mathsf{q}_p$ in the population. Recalling that $R = 1/\alpha$, the average distortion (*i.e.*, the theoretical capacity) of a compressor based on this CSP is computed through equations (3) and (15).

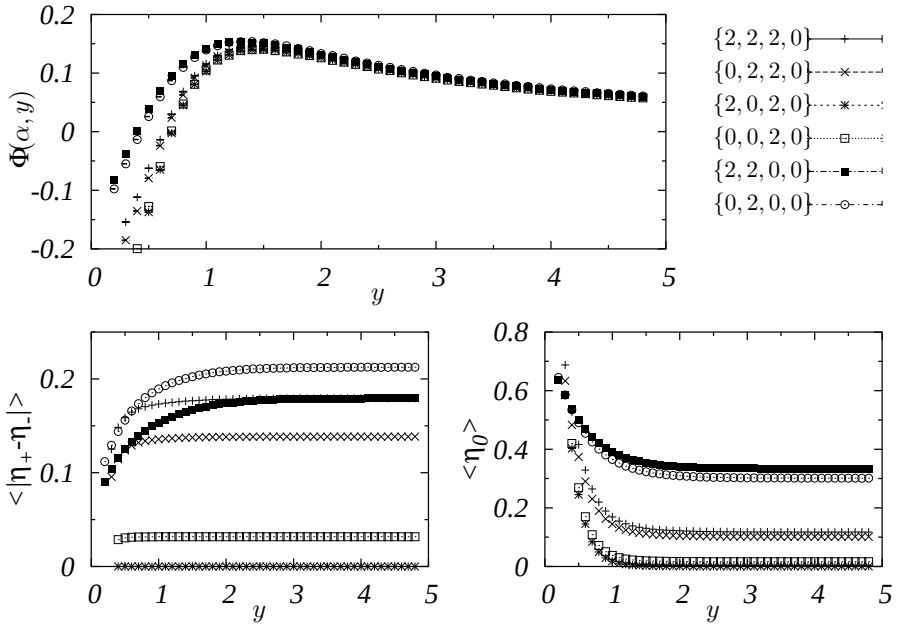


FIG. 3: Top: The free energy $\Phi(\alpha, y)$ of a number of symmetric nodes with $K = 7$ at $\alpha = 1.4$, classified according to the rule given in the text. Bottom: The asymmetry $|\eta_+ - \eta_-|$ of these nodes (left) is measured as well as the "paramagnetic degree" $\eta_0 = 1 - \eta_- - \eta_+$ (right). Note that the usual 7-XOR node corresponds to the $\{2, 0, 2, 0\}$ case in this notation.

Let us study here as an example a family of function nodes whose energy is fully invariant under permutations of the arguments. These nodes can be classified according to the energy of the node for a given value of the "magnetization" $m \equiv \sum_{i=1}^{K} \sigma_i$. We keep to odd K and label a particular node by the sequence of values $\{\varepsilon(m = K), \varepsilon(m = K - 2), \ldots, \varepsilon(m = 1)\}$. In Fig. 3 we report the results of the population dynamics algorithm for some types of nodes at $K = 7$. According to eq. (15), the ground state energy $E(\alpha)$ corresponds to the maximum over y of the free energy $\Phi(\alpha, y)$ represented in the top plot. The theoretical performance of these nodes are quite close to the PSC case (the XOR node, characterized by the sequence $\{2, 0, 2, 0\}$).

III. NON-LINEAR NODES

We consider in this section the function nodes that give the best performance for compression, both form the theoretical point of view (theoretical capacity close the parity-check nodes) and from the algorithmic aspect (the SID algorithm at finite y is found to converge in

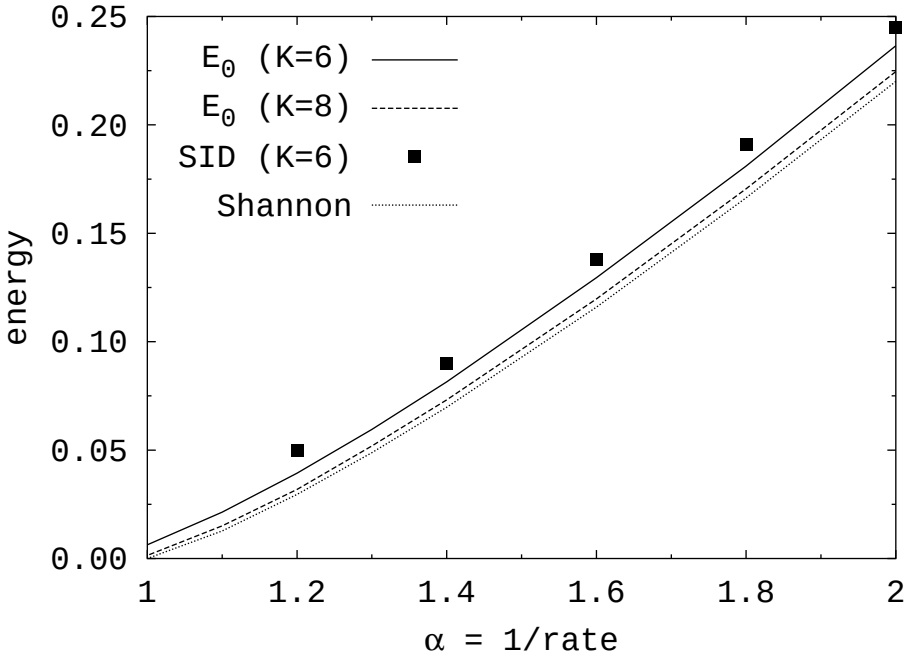


FIG. 4: The ground state energy for a system with 30 non-linear nodes. We also plot the Shannon bound and the performance of the SID algorithm for the $K = 6$ case.

the UNSAT regime, thus giving an explicit encoding algorithm). These non-linear function nodes are defined as follows. We recall that the output ε_{XOR} of a K-XORSAT node is given by twice the sum modulo 2 of the input bits. We label each configuration $\vec{\sigma} \in \{0,1\}^K$ with an integer $l \in [1, 2^K]$ and consider a random permutation π of the vector $\{1, 2, \ldots, 2^K\}$. Then, we can associate the output of the *random node* by letting $\varepsilon^{(\pi)}(l) = \varepsilon_{XOR}(\pi(l))$. In this way we are left with a random but balanced output which can be different from XOR. Also, it is clear that they are not more defined by a linear formula over the boolean variables.

We can take advantage of the formalism introduced in the previous section in order to study the theoretical performance of these new function nodes. In particular, we have used 10 to 30 different random nodes to build the factor graphs. In order to improve the performance, we have also forbidden "fully-canalizing" nodes, that is nodes whose SAT character depends on just one variable.

In Fig. 4 we show our results. The ground state energy is shown to quickly approach the Shannon bound as K increases and for any α. As an example, at $\alpha = 2$ (corresponding to a compression rate $R = 1/2$) the difference between the $K = 8$ value and the theoretical limit is $\simeq 2\%$. This looks very promising from the point of view of data compression.

Furthermore, the results obtained from the SID (same plot) show that in this case the algorithm does converge and its performance is very good. It should be noted that when K becomes large the difference between SP and the Belief Propagation (BP) becomes small (this can be seen for instance in the analysis of [18]), so in fact BP does also provide a good encoding algorithm for $K \gtrsim 6$. At fixed K, the time needed to solve the SP equations is $\mathcal{O}(N \log N)$. The actual computational time required by the decimation process is of the same order and it slightly depends on the details of the SID algorithm (*e.g.* the number of variables fixed at each iteration, whether or not a backtracking procedure is used, which kind of heuristic is adopted to solve an "easy" reduced instance, the proper definition of the latter, etc...). The dependence on K is exponential. Thus, even if increasing K is good from the theoretical point of view, it turns out to be very difficult to work at high K. In practice, it takes a few hours to compress a string of $N = 1000$ bits at $K = 6$ by using our general purpose software. We think that some more specified code would lead to a better performance.

To conclude, we have shown how the methods of statistical mechanics, properly adapted to deal with a new class of constraint satisfaction problems, allow to implement a new protocol for data compression. The new tool introduced here, the non-linear gates, looks very promising for other practical applications in information theory.

Acknowledgments

We thank A. Montanari, F. Ricci-Tersenghi, D. Saad, M. Wainwright and J. S. Yedidia for interesting discussions. S. C. is supported by EC through the network MTR 2002-00307, DYGLAGEMEM. This work has been supported in part by the EC through the network MTR 2002-00319 STIPCO and the FP6 IST consortium EVERGROW.

---

[1] R. G. Gallager, Low Density Parity-Check Codes (MIT Press, Cambridge, MA, 1963)

[2] D. A. Spielman, in Lecture Notes in Computer Science **1279,** pp. 67-84 (1997).

[3] D. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, (2003).

[4] N. Sourlas, Nature **339,** 693-694 (1989).

[5] A. Montanari, Eur. Phys. J. **23,** 121 (2001).

[6] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing : An Introduction*, (Oxford University Press, 2001).

[7] F. Jelinek, IEEE Trans. Inform. Theory, **15,** No. 6, 584 (1969).

[8] E. Martinian, J. S. Yedidia, MERL report 2003.

[9] J. Ziv, A. Lempel, IEEE Trans. Inform. Theory **23,** 337 (1977); J. Cleary, I. Witten, IEEE Trans. Commun. **32,** 396 (1984); F. Willems, Y. Shtarkov, T. Tjalkens, IEEE Trans. Inform. Theory **41,** 653 (1995).

[10] T. Berger, J.D. Gibson, IEEE Trans. Inform. Theory, **44,** No. 6, 2693 (1998).

[11] Some of the results presented here have appeared in S. Ciliberti, M. Mézard, R. Zecchina, Phys. Rev. Lett. **95,** 038701 (2005).

[12] T. Hosaka, Y. Kabashima, H. Nishimori, Phys. Rev. E **66,** 066126 (2002).

[13] D. Battaglia, A. Braunstein, J. Chavas, R. Zecchina, Phys. Rev. E **72,** 015103 (2005).

[14] J. Garcia-Frias, Y. Zhao, IEEE Commun. Lett., 394 (Sept 2002).

[15] G. Caire, S. Shamai, S. Verdù, in Proc. of IEEE Inform. Theory Workshop, 291, Paris (2003);

[16] Y. Matsunaga, H. Yamamoto, in Proc. of IEEE Intern. Symp. on Inform. Theory, 461, Lausanne (2002).

[17] R. M. Gray, D. L. Neuhoff, IEEE Trans. Inform. Theory, **44,** 2325 (1998).

[18] S. Ciliberti, M. Mézard, `cond-mat/0506652.`

[19] M. Mézard, G. Parisi, R. Zecchina, Science **297,** 812 (2002).

[20] G. Biroli, R. Monasson, M. Weigt. Eur. Phys. J. **B14,** 551 (2000).

[21] M. Mézard, R. Zecchina, Phys. Rev. E **66,** 056126 (2002).

[22] A. Braunstein, M. Mézard, R. Zecchina, in *Random Structures and Algorithms* **27**, 201-226 (2005)

[23] D. Battaglia, M. Kolàr, R. Zecchina Phys. Rev. E **70,** 036107 (2004).

[24] G. Parisi, `cond-mat/0308510`.

[25] M. Wainwright, private communication.

UNDERSTANDING FRACTAL ANALYSIS? THE CASE OF FRACTAL LINGUISTICS

HERBERT F. JELINEK
School of Community Health,
Charles Sturt University, Australia
CAMERON L. JONES
Centre for Mathematical Modeling, School of Mathematical Sciences
Swinburne University of Technology, Australia
MATTHEW D. WARFEL
Mennonite Central Committee – Bolivia
Casilla 59 Camiri, Bolivia
CECILE LUCAS; CECILE DEPARDIEU; GAELLE AUREL
I.U.P. Genie Physiologique – Informatique et D.E.S.S. DESSTAUP.
University Poitiers, France

Field: CS questions
Key words: Fractal analysis, nonlinear science, communication, linguistics
Communicating Author:
Herbert F. Jelinek
School of Community Health
Charles Sturt University
Albury, 2640
Australia
PH: +61 2 60516946
FAX: +61 2 60516772
EMAIL: hjelinek@csu.edu.au

ABSTRACT

Terms such as ‘self-similarity’, ‘space filling’ ‘fractal dimension’, and associated concepts have different meanings to different people depending on their background. We examine how methodology in fractal analysis is influenced by diverse definitions of fundamental concepts that lead to difficulties in understanding fundamental issues. The meaning of terms associated with fractal analysis need to be clarified if this method is to be useful in diverse disciplines. It is our premise that communications that are result focused constitute a danger in perpetuating misconceptions of terms due to the concise nature of the writing and the reliance on references to fill in the procedural and conceptual gaps. Communicating effectively requires a sound understanding of the terminology and a clear and meaningful presentation. We address here communication and the nature of scientific discourse, ‘*fractal linguistics*’

INTRODUCTION

The literature shows that the fractal dimension (*D*) of an object reveals something about the natural world not otherwise apparent. This, combined with the seemingly simple procedures involved in fractal analysis, has led to the popularity of this procedure for pattern analysis. The interpretation of such analysis, however, is not always straightforward and confounding this are a number of misconceptions associated with terminology and meaning. This is particularly true for people interested in applications of fractal analysis but lacking a deep understanding of the underlying mathematical theory. In this paper we attempt to correct some of these misconceptions and suggest ways which will avoid their repetition. For language to have meaning in a community requires the specialists to take responsibility for the dissemination of knowledge that enables others to gain control of the discourse (subject matter) and contribute in a meaningful way. The lack of clarity can be attributed to people in specialised fields forming a specific linguistic boundary, *fractal literacy*, and expecting researchers in a different research field with different subject literacy to follow the discourse. The ideas underlying fractal analysis are inherently alien if compared, say, to the mathematics underlying simple principles in physics. The transfer of ideas from the theoretical domain of science to the applied is never easy but in the case of fractal analysis it appears to have been particularly difficult. The problem has been confounded in the case of fractals by the increasing specialization of science.

At the bases of this discussion is that there appears to be little agreement on the meaning of the terms used in fractal analyses and, more importantly, how these terms are used in context of describing the results of fractal analyses. As an example consider the question whether biological forms are fractal.[1,2] Strictly speaking, a fractal only describes forms that are strictly self-similar and infinite. Natural objects and, their representations on computers are not fractals, yet are often so described. Are natural objects space-filling? Like the term ‘fractal’, ‘space-filling’ can have various meanings. With regards to fractal theory, space-

filling is an attribute of fractals and reflects that the recursive nature of the fractal tends to a space-filling limit. An example is the Peano curve that if drawn to the limit of infinity has infinite length and reaches every point of the delimited plane it is drawn on.[3] Biological forms do not have this property as they do not possess infinite length. Thus biological forms are not strictly space-filling. However space-filling can be used in a difference sense as discussed by Murray.[1] Here space-filling is viewed as a process, a structure such as a plant is involved in, to optimise coverage.

Can we then use fractal analysis to discuss forms in nature? As the magnitude of published literature indicates, many seem to think this is possible. What is often not clearly stated is that 'fractals' can only be used as '*models*' for biological shapes because natural objects lack several characteristics that define fractals.[3] The most common characteristic cited for fractals is the exact repetition of detail at every observation scale, its *strict self-similarity*.[4] However the construction of true fractal objects can be randomized. These objects possess *statistical* self-similarity and resemble more closely objects in nature. In the literature though, there has been a tendency to equate statistical self-similarity (which natural objects do possess to differing degree) with strict self-similarity and to assume that estimates of statistical self-similarity are a test of fractality.

D is a parameter that describes the relationship between measured size and the measuring scale. Common examples from the literature that use *D* to describe relationships in space and time, include heart rate irregularities, grazing effects on pastoral lands or structural attributes of blood vessel systems and neurons.[5-10] It is our observation however, that conclusions drawn from fractal research remain at best tentative due to a lack of a generally comprehensible description of fractal theory and its relationship to the associated analysis procedures.

Concepts in fractal research such as the fractal dimension are not strictly defined and much of the terminology is used loosely. Fractal research and discussion is characterised by the repetition of definitions and procedures that were initially ill defined and intended by the authors to be vague.[7] Thus, communicating effectively requires a sound understanding of the terminology and a clear and meaningful presentation. We address here communication and the nature of scientific discourse, '*fractal linguistics*'.

Communication

"A fractal set is a set in metric space for which the Hausdorff-Besicovitch dimension *D* is greater than the topological dimension D_T".[7, p361]

The sentence above is the most often quoted citation found in journal articles to describe fractals, even though Mandelbrot stated, on the next page of his book, that this definition is rigorous, but also tentative.[5]

The definition is an example of communication that may be mathematically rigorous, but uses language that is not accessible to non-specialists. Terms such as '*set*', '*metric space*', '*Hausdorff-Besicovitch dimension*' and are not very informative and meaningful to novices to fractal analysis and possibly to non-mathematicians, bearing in mind that defining what a dimension is, has involved the most brilliant mathematicians since the ancient Greeks.[3]

How can fractal analysis be better understood?

Once a group of the scientific/professional community is established it becomes obvious that a type of professional socialisation takes place. This socialisation leads to communication barriers in terms of the discourse adopted by a specific group. Thus scientific community in terms of research area and proficiency in fractal analysis can be divided into three main categories: 1. those in the know who already posses a solid knowledge base in fractal and scaling theory.[7,14,15] Thus the use of specific language by a group establishes an identity for this group with a specialist domain of knowledge and expertise. 2. those with a solid knowledge base in other disciplines and an understanding of mathematics or physics, [16-18 3] those that bridge the gap between 1. and 2. that is, those that develop tools for fractal analyses and aim at describing these clearly).[19] Category 3 should primarily consist of category 1 researchers that aim to make the theoretical basis more accessible to applied research. Despite the efforts of the third group the gap between groups 1 and 2 is still considerable. In order to have control of the ways language is used to create meaning in fractal analysis throughout the community requires the specialists to take responsibility for the dissemination of knowledge that enables others to gain control of the discourse (subject matter) and contribute in a meaningful way. The lack of clarity lies in the fact that many in the first category, being entrenched within a specific linguistic boundary, *fractal literacy*, find it hard to communicate the necessary information to researchers in a different research field with different subject literacy. Papers in the second category reflect that the authors can acquire the necessary literacy and know how fractal analysis is applied but perpetuate that same inaccuracy in terminology. Thus, it is this section where inherent uncertainties in definitions are most easily propagated. One likely reason for this is that the communicator adopts the new language base but may not have a correct understanding of the terms. Should a correct understanding exist then it is important to communicate this explicitly This latter requirement is important as many researchers in specialised fields do not have familiarity with expertise outside their special interests yet are, in many instances, referred to publications from the first category above. The third category is meant to provide the tools and structure necessary to bridge the gap between the first and the second category. This third category is not well presented.[3,20-22] It is our assertion therefore that category 1 and 2 researchers have an obligation to disseminate their work in a manner understandable to all in addition to their publications in specialist journals. These communications certainly exist. Indeed Mandelbrot, who is known as the father of fractal geometry and a mathematician, published a paper that is aimed for the uninitiated, as did Jürgens and co-workers, albeit this latter paper is more complicated.[23,24] Yet, why are the most active uses of fractal analysis in physics?[25]

The problems which surround fractal research can be illustrated by the case of diffusion limited aggregates (DLAs). These are the result of a random growth process and could serve as a model for many such processes, for example, in biology. Most of the analysis of DLAs has been undertaken by physicists who have a good grasp of the theory. The inherent complexities of DLAs, however, have meant that there is considerable disagreement about the interpretation of these objects. Is a DLA self-similar?[25,27] Is a DLA a multifractal?[26,28-30] These publications and their authors are well known in the field of physics and thus their results are widely cited. All papers cited here involved physicists and apart from one paper, were published in physics journals and written for physicists. This may explain why fractal analysis has not been as successful in the biological sciences as it has in engineering and physics. What is lacking in the field of fractal analyses, is for researchers in the field of fractal theory to articulate clearly how they themselves learned both theory and application, *the fractal discourse.* It is important to disseminate to the rest of the community how meaning is constructed and communicated in their field.

Fractal linguistics

The popularization of fractals and chaos has led to widespread interest in applying these ideas in a range of scientific disciplines. What is not immediately apparent to scientists entering this field is that there is no consensus on a number of the most basic issues, even amongst experts. Terms, such as those mentioned at the beginning of this paper, and procedures used in fractal analysis need to be understood before fractal analysis becomes an effective tool. It is the World Wide Web that can be of use here as many sites exist that address '*basic*' issues. We exemplify our stance in the next section. Teaching the skills of fractal literacy: a language base for novices.

A Language Base for Novices

This section cannot cover all the terms that are required for a good language base. It rather aims at correcting some common misconceptions. What then is a fractal and how is fractal analysis performed? We start by describing some of the properties displayed by fractals. Fractals do not have inherently smooth surfaces no matter what magnification is used to examine the object.

How can a fractal be described?

A common, yet 'mathematically' inappropriate term used to describe the surface of objects, is their '*roughness*'. The term roughness is used in common language to describe what is, in fractal linguistics, an object's space filling capacity or surface irregularity. Yet, roughness and fractal dimension do not describe the same feature. Roughness, is a measure of the average variation about a mean and is not related to scale

or changes in scale of measurement. *D* quantifies the variation in length, area or volume with changes in the size of the measuring scale.[16]

How can the attributes of fractals and natural objects be described more appropriately? Three of these are *characteristic length*, *self-similarity* and *complexity*. Any non-fractal form can be approximated by a simple shape with the same characteristic length. Thus, a sphere can approximate the Earth. The existence of a characteristic implies a smooth surface In case of the Earth, the highest mountain is much smaller than the diameter. Another way of looking at this is to say that the surface of the Earth becomes smoother as the magnification is decreased. Fractal objects cannot be represented by any combination of shapes with characteristic lengths. A cloud is such an object. Thus for many forms, Euclidean geometry and its associated Euclidean dimension suffice to characterise these. Ideal fractal objects have no characteristic length. Consider the representation of several steps in the construction of a Koch curve (Fig.1). Even at the fourth step of construction, it becomes apparent, that no characteristic scale exists for the Koch curve and that it can only be approximated by the use of a number of spheres with different sizes and not by one simple shape.

PUT FIG 1 HERE

Fig.1 also demonstrates the self-similarity of fractal images as any change in scale will show more detail as magnification is increased but will not lose detail if the magnification is decreased. Fractal analysis is then the procedure that compares the size of the outline or mass of the object for each scale of the measurement. This can be done using disks, or squares. If this change of size is constant with change of scale on a double logarithmic plot, the form approximates a fractal and the gradient of the line through these points is proportional to the fractal (self-similarity) dimension (Fig. 2).

PUT FIG 2 HERE

The gradient is mathematically defined by $D = \frac{\log N(e)}{\log(1/e)}$, which transforms into $N(e)^D = 1$, a power law relationship, where the exponent '*D*' is the fractal dimension (for the Koch curve in Fig. 2, $D = 1.246$). The mathematical definition of the fractal states that the size of the image scales indefinitely by a constant proportion. The fractional part of the exponent indicates the complexity of the object. If the fractional part of the exponent is 0, *D* would be equal to 1 and equal to the Euclidean dimension for a line. If the line becomes more space filling the fractional part of the exponent increases towards 2.

Complexity is another common term used to describe the surface irregularities or the intricacy of a branching structure. *D* indicates whether the structure contains a degree of self-similarity, which in turn can be an indication of the underlying biological process that leads to the observed pattern. Self-similarity as

Mandelbrot points out, is a simple design principle that can be independent of genetic determination and only dependent on a consecutive scaling with changes of magnitude. The dimension exponent can indicate how a structures' branches scale from the parent to the daughter branches as can be observed in the lung bronchi for instance.[8] *D* can also be an indicator of how space-filling a structure is. If *D* has a value of 1.2 than the structure is not as space filling as if *D* was 1.4. Thus the feature that makes fractal analysis interesting is measuring the range of the statistical self-similarity across scaling levels and the associated estimate of *D.* Statistical self-similarity can be an indicator of a growth process in tubular structures such as blood vessels, lung bronchi or neurons. The fractal dimension, provided the structure displays 'statistical self-similarity' can be used to infer physiological function such as flow of air in lungs or electrotonic properties in neurons as the diameter exponent in this special case equals the fractal dimension in some cases.[7]

Accuracy or precision is important in fractal analysis as with any other measurement. Using Euclidean objects is the easiest way forward as the dimension of these is known. However because we want to ascertain the fractal dimension of fractal-like structures it is important to use test images that are similar to the structures to be investigated. In general *fractal* objects such a as the Koch snow flake or a DLA (diffusion limited aggregate) is used for such calibration[25] However a fractal object is the result of repeated transformations of a geometrical figure that leads to strictly self-similar pattern with no scaling limit. In practice only prefractals can be used to calibrate the measurement system. These are approximations to true fractals – examples of true fractals shown at a finite number of iterations – since truly infinite cases are not possible in the real world. The '*Koch curve*' shown in Fig. 1 is not a true fractal, as none of the iteration levels depict the case of infinity. Any incomplete representation of an ideal fractal, as is the case with computer representations, is referred to as prefractal.[31] In practice when prefractals are used as test images, accuracy in the strict sense cannot be ascertained as the image is not completely represented on the screen. However analytical estimates using many fractal analysis algorithms agree with the theoretical dimension value.

Once a biological object is analysed and a constant slope is obtained for the log-log representation of scale versus size as a function of scale, the object is said to be self-similar with a certain fractal dimension. This broad use of the term *self-similar* has led to biological forms being often identified as fractal. Biological forms are not fractals, as they are not characterised by identical patterns at different scales (strict or linear self-similarity).[7,33] Biological forms are at best statistically self-similar over a limited range.[32]
This misconception may have arisen from the use of the term self-similarity and scale-invariance to indicate both strict self-similarity for ideal fractals and statistical self-similarity if biological forms are described.[3,34-36] Similarly, obtaining *D* for an object does not indicate that this object is fractal. Fractality is determined by the relationship between the *observation scale* and the *measured size* of the

object, which has to be constant without a limit. Once some of the terminology becomes clear, the application of the algorithms still remains a mystery.

The Black Box

Prior to the advent of fractal analysis, the relationship between scale and length of coastlines was discussed by Richardson.[37] He used callipers of different size (scale) to determine the length of the coast of Britain. What he found was that the coast line length increased as the scale was decreased according to a power law relationship. The *exponent*, is then the parameter that quantifies this relationship. Thus, whether the object is fractal or not did not have any bearing on the aim of the analysis. [37-40] Currently many different methods such as box-counting, dilation, mass-radius as well as the calliper method used by Richardson are used to determine similar relationships. These methods are all analytical tools that estimate the relationship between a scale of measurement and the size/mass of the object being measured. This is now referred to as fractal analysis and the exponent is the fractal dimension. What is a dimension? There exists disagreement among physicists and mathematicians, what constitutes a dimension and what measures can be included in this term. Apart from the Hausdorff dimension other so called fractal dimensions such as the Minkowski and the Kolmogorov dimension are not really dimensions in a mathematical sense.[3,7,16,41] However, in the literature authors often refer to these two measures as dimensions possibly because of trying to simplify what is really quite a complex area in mathematics. Analytical tools, such as the dilation method, and box-counting method, are based on the Minkowski measure and Kolmogorov measure respectively. These two measures are in themselves approximations of the mathematically rigorously defined fractal dimension (Haudorff dimension) and are only equal for strictly self-similar objects.[21,35] Indeed books and articles containing the terms ‘fractal analysis’ in their titles often do not point out that a strict application of the rules of mathematics to the procedures of fractal analysis is not possible. One only needs to observe the intricate definition of the Hausdorff dimension, which is not usable in practice and the shortcomings of the analysis procedures in estimating the Hausdorff dimension using for instance the box-counting method.[32, 41,42]

Having decided to undertake fractal analysis, we need to consider several procedural steps. The first question to ask is whether or not it is important that the image is fractal. No! Fractal analysis procedures are applicable to non-fractal objects (which all biological objects are). Next is the question of what aspect of the structure is to be analysed, especially with respect to the necessity of representing the structure as a 2D object on the computer screen. Is a surface the same as a boundary or border? For instance, when considering a structure such as a particle aggregate, many authors have a different idea of what a 2D image versus 3D object is or what a surface, a boundary or border is. What is mass in 2D?[41,43] What does it mean to analyse a surface when some people see a surface like a relief map, while others view it as the boundary or perimeter of a 3D object? Understanding these terms is critical as techniques for analysing the

scaling characteristics are very different.[44-48] Image silhouettes, for example, have been used in the analyses of fractal-like characteristics in sludge aggregates by several investigators.[47-49] These silhouettes can be presented as sectioned boundaries or as the boundary of the silhouette. Estimates of the dimension of silhouette boundaries are consistently smaller than those of sectioned boundaries.[48] Associated with this is the significant problem of determining the 'appropriate' fractal dimension. What type of fractal does the image represent? Pfeifer suggests several 'prototypes' of fractals including mass and porous fractals.[44] This is important as different fractal analysis procedures have been suggested, depending on the prototype the object is akin to. Note, though that there is no agreement with this step and some authors use the boundary of an object regardless if its length or mass with respect to scale is investigated.[50-52]

Having decided that the image to be analysed does not have to be fractal but that some type of scaling rule may be implicit in the morphology of the structure, the next step is to decide on the type of algorithm to be used.[44,53] Now it becomes important to extend or working definition of self-similarity to include the broader concept of *self-affinity*.[3] Self-affine images scale differently in the *x* and *y* directions. When the object scales equally in the *x* and *y* direction, self-affinity is the same as self-similarity.[3,54]

Before starting the analysis, some image manipulation or preparation of the sample, such as sectioning a 3D object versus representing it as a silhouette, may be required and may have an effect on the estimate of *D*.[5,43] The image needs to be available for the specific fractal analysis software for analysis. Here the placement of the image and the size of the image within the acquisition screen may also influence the estimated *D*. Many of the fractal analysis applications then include a choice of the number of starting points for the box placement and selection of the range of box sizes and whether the image is to be rotated.[4,5,55-57] These options need to be investigated with all software as both the extent of statistical self-similarity and the magnitude of *D* are influenced by these procedures. Further, application of conventional fractal techniques such as box-counting are inappropriate for analysing self-affine fractals.[58] However, using a section parallel to the nominal surface orientation, the resulting boundary lines may be statistically self-similar or scale-invariant. These boundary lines can then be analysed to characterise the surface. The boundary of the silhouette is an approximation of this sectional boundary as discussed previously.

Using 'fractal' test images to ascertain the efficacy of the procedure may also be misleading because of how this aspect of fractal analysis is often described. The accuracy of fractal analysis methods is by its very nature difficult to determine. Consider the determination of *D* for a branching fungus. The control image needs to be of similar shape to the image tested.[59] DLAs (diffusion limited aggregates) are chosen often for this task.[25,52] Yet physicists are not in agreement that a DLA is a fractal or what it's exact dimension is.[23,33,43] If the dimension of a boundary has to be determined, such as for a particle aggregate, the *Koch curve* can be used as a control image. The precise *D* for this curve is 1.264 but as the curve cannot be

represented accurately, that is, with an infinite amount of detail, the estimate of *D* will in many cases deviate from the theoretical. This can be avoided if a level 4 or 5 representation of the curve is used (see Fig. 1). In addition there is a pixelation/staircase effect that is dependent on the size of the pixels (resolution of the screen) that influences the complexity of the border. In classification tasks for instance, the robustness of the procedure is important and this can even be tested by using Euclidean images such as a square or circle. The question is not whether the images are fractal but if a there is difference between images in their scaling behaviour.[60,61]

What fractal analysis procedure is appropriate has been briefly mentioned above. Despite arguments in the literature that both mass and boundary fractal dimensions should be measured using image boundaries, investigators have often used these fractal analysis techniques to measure different types of images.[5,18] Others used the box-counting technique on silhouette areas to characterise the mass of flocs and granules.[48] The resulting dimension, though perhaps not a theoretical mass fractal dimension, was found to be effective in differentiation tasks.[43] Fractal analyses methods such as the dilation technique may also be applied to image boundaries (including both perimeters and pore space boundaries) of a 2D image.[37,49] However, the resulting fractal dimension is greater than the fractal dimension of the perimeter as the pore boundaries provide an increase in the space-filling ability of the surface.

Once the data has been obtained as a scaling relationship between the logarithm of the scale of observation and the logarithm of the size of the image, the extent of statistical self-similarity and the fractal dimension can be determined. A decision that needs to be made here is whether to include all data points or whether there is one, two or more clearly defined linear segments apparent in the plot.[3,6,20,51] Inherent in the linearity of the plot is the selection of the different scales of the measuring device such as a box or circle. If the measuring scale is very much larger than the size of the object, the box always covers the object and the number of boxes required to cover the object, regardless of the size, will always be 1 and the slope or *D* will be 1. Alternatively, if the size of the box is smaller than the "line width" of the object, then one is measuring the area of a solid object and the slope or *D* is 2.[62] Having chosen the appropriate size for the boxes, the log-log plots that represent the scaling characteristics of biological forms still do not show a constant relationship between the scale of measurement and the size/mass of the object. The points where this occurs are the upper and lower cut-off points between which statistical self-similarity occurs. One can remove data points one at a time until the squared correlation coefficient approaches some previously defined number (eg 0.995), or utilise a combination of curve-fitting tests and curvilinearity of residuals to identify the largest range over which the image displays statistical self-similarity.[22,63,64] Curvilinearity can be tested by fitting the data to first and second order polynomials and comparing using a chi square test the better fit or use the error in Y.[22,63] If regression analysis is used, one needs to consider that converting linear data to log-log data changes the profile of the distribution of the data points. This can be corrected for by weighting the log-log data points.[65] However, any method involving linear regression

may not be suited to measure complexity and an alternative may be more appropriate.[41] The final outcome is an estimate of the fractal dimension for a biological object and an indication of the range of statistical self-similarity.

The above considerations have attempted to provide a start to sound understanding of fractal analysis in that they clarify the meaning behind terms commonly used in this field and aspects of the procedure.

CONCLUSION

Being familiar and understanding the basic terms used in any analysis procedure sets up a linguistic domain that is imperative to ensure that the results are meaningful. This paper concentrated on several that are relevant to fractal analysis. The process of selecting an appropriate fractal analysis technique with reference to image characteristics was also outlined. This then represents together with the section on ’language base for the novice’ and ‘the black box’ an attempt to make this methodology available and understandable to scientists within the field of fractal analysis and in other field of research.

Acknowledgments

MDW would like to acknowledge the support of the National Science Foundation, Environmental and Ocean Systems Program (Grant No. BES-9312825). HJ was in receipt of a Faculty Seed Grant.

REFERENCES

1. Murray, JD: Use and abuse of fractal theory in neuroscience. J Comp Neurol 1995; 361:369-371.
2. Panico, J, Sterling, P: Retinal neurons and vessels are not fractal but space filling. J Comp Neurol 1995; 361:479-490.
3. Peitgen, H-O, Jürgens, H, Saupe, D: Fractals for the Classroom. Part One, Springer-Verlag, New York, 1991
4. Montague, PR, Friedlander, MJ: Morphogenesis and territorial coverage by isolated mammalian retinal ganglion cells, J Neurosci. 1991; 11(5):1440-1457.
5. Jelinek, HF, Fernandez, E: (1998) Neurons and fractals: how reliable and useful are calculations of fractal dimensions? J Neurosci Meth 1998; 81:9-18.
6. Vicsek, T: Fractal Growth Phenomena, 2 ed. Singapore, World Scientific, 1992.
7. Mandelbrot, BB: The Fractal Geometry of Nature, W.H. Freeman and Co., N.Y., 1987.
8. Bassingthwaighte, JB, Liebovitch, LS, West, BJ: Fractal Physiology, Oxford University Press, Oxford, 1994.
9. Loehle, C, Li, B: Statistical properties of ecological and geological fractals. Eco. Mod. 1996; 85:271-284.
10. Stanley, HE, Ostrovsky, N: On Growth and Form: Fractal and Non-Fractal Patterns in Physics. Nijhoff, Dordrecht, 1986.
11. Sernetz, M, Wübbeke, J, Wlczek, P: Three-dimensional image analysis and fractal characterization of kidney arterial vessels. Physica A, 1992. 19:13-16.
12. Jelinek, HF, Spence, I: Categorization of cat retinal ganglion cells using fractal dimension: the case of the gamma, epsilon and delta cells, in Australian Neuroscience Soc. 1995; 134.
13. Avnir, D, Biham, O, Lidar D, Malcai, O: Is the Geometry of Nature Fractal. Science, 279, 39
14. Hausdorff, F. (1919) Dimension und äußeres Maß, Math Ann 1998; 79:157-179.
15. Taylor, J: The measure theory of random fractals. Math Proc Camb Phil Soc 1986; 100:383-407.
16. Cox, BL, Wang, JSY: Fractal surfaces: measurement and applications in the earth science. Fractals 1993; 1(1): p. 87-115.
17. Smith, Jr., TG., Behar, TN, Lange, GD, Sheriff, Jr., WH, Neale, EA: (1989) A fractal analysis of cell images. J Neurosci Meth 1989; 27:173-180.
18. Roach, DE, Fowler, AD: Dimensionality analysis of patterns: fractal measurements. Computers Geosciences, 1993; 19(6):849-869.
19. Bernston, GM, Stoll, P: Correcting for finite spatial scales of self-similarity when calculating the fractal dimension of real-world structures. Proc. R. Soc. Lond. B. 1997; 264:1531-1537.
20. Hastings, HM, Sugihara, G: Fractals: a user's guide for the natural sciences. Oxford University Press, Oxford, 1993.
21. Lauwerier, H: Fractals: Images of Chaos. Penguin Books. Princeton University Press, New Jersey,

1991.

22. Russ, J : Fractal Surfaces, Plenum Press, New York, NY., 1994.
23. Mandelbrot, BB: Fractals – geometry of nature. Science 1990; 22-27.
24. Jürgens, H, Peitgen, H-O, Saupe, D: The language of fractals Sci Am 1990; 40-47.
25. Witten, A, Sander, LM: Diffusion Limited Aggregation. Physical Rev. B, 1981; 27(9):5686-5697.
26. Nittman, J, Stanley, HE, Touboul, E, Daccord, G: Experimental evidence for multifractality. Phys Rev Lett 1987; 58:612-622.
27. Mandelbrot, BB: Plane DLA is not self-similar; it is a fractal that becomes increasingly compact as it grows? Physica A.1992; 191:95-107.
28. Viscek, F, Family, F, Meakin, P: (1990) Multifractal geometry of diffusion-limited aggregates. Europhys Lett 1990; 12:217-222.
29. Lam, CH:Finite size effects in diffusion-limited aggregation. Phys Rev B 1995; 52:2841-2847
30. Fernandez, E, Bolea, JA, Ortega, G, Louis, E: Are neurons multifractals? J Neurosci Meth 1999; 89: 151-157.
31. Feder, J: Fractals, Plenum Press, New York, 1988.
32. Schroeder, M: Fractals, Chaos and Power Laws, W.H. Freeman and Co. U.S.A., 1991.
33. Shenker, OR: Fractal geometry is not the geometry of nature. Stud Hist Phil Sci 1994; 25(6):967-981.
34. Takayasu, H: Fractals in the Physical Science,. Manchester University Press, Manchester and New York, 1990.
35. Peitgen, H-O, Saupe, D: The Science of Fractal Image, Springer-Verlag, Berlin, 1988.
36. Pfeifer, P: Is nature fractal? Letter to Science. 1998; 279:784.
37. Richardson, LF: The problem of contiguity: an appendix to statistics of deadly quarrels. General Systems Yearbook. 1961; 6:139-187.
38. Bellouti, MM., Alves, M, Novais, JM, Mota, M: Flocs vs. granules: differentiation by fractal dimension Water Res 1997; 31 (5): 1227-1231.
39. Pfeifer, P, Welz, U, Wipperman, H: Fractal surface dimension of proteins: lysozyme. Chemical Physics Letters, 11985; 13 (6): 535-540.
40. Reilly, S, Clark, NN: Computer based general shape description with harmonics and a roughness index. Proceedings of the Institution of Mechanical Engineers. Part E, J of Proc Mech Eng 1991; 205 (E2): 103-111.
41. Sandau, K, Kurz, H: Measuring fractal dimension and complexity - an alternative approach with an application J Microscopy 1997; 186(2):164-176.
42. Falconer, K: Fractal geometry, mathematical foundations and applications. Wiley, New York, 1990.
43. Warfel, MD. Characterization of particles from wastewater and sludge treatment facilities by size and morphology M.Sc. Thesis, Cornell University. 1998.
44. Pfeifer, P: (1987) Characterization of surface irregularity. In: Preparative Chemistry Using

Supported Reagants. P. Laszlo (ed.). Academic Press, San Diego, CA. 1987.

45. Akashi, T, Kojima, E, Ichikawa, E: Applied fractal method to measurement of flocculation in wastewater treatment processes. In: Control and Instrumentation. Proceedings of the 20th International Conference on Industrial Electronics, 1994; 2:1259-64.
46. Namer, J, Ganczarczyk, JJ: (1994) Fractal dimensions and shape factors of digested sludge particle aggregates Water Pollution Res J 1994; 29 (4):441-455.
47. Jiang, Q, Logan, E: Fractal dimensions of aggregates from shear devices. Journal of the American Water Works Association, 1996; 88:100-113.
48. Ganczarczyk. JJ: A correlation between external surface characteristics and settling velocity of activated sludge flocs. Proceedings of the International Association of Water Quality, 1995; 8 (1), 4-7.
49. Jiang, Q, Logan, B: Fractal dimensions of aggregates determined from steady-state size distributions. Environmental Science and Technology, 1991; 25 (12):2031-2038.
50. Smith, TG, Lange, GD, Marks, WB: Fractal methods and results in cellular morphology – dimensions, lacunarity and multifractals. J Neurosci Meth 1996; 69(2):126-136.
51. Fernandez, E, Jelinek, HF: Use of fractal theory in neuroscience: methods, advantages and potential problems. Meth 2001; 24:309-321
52. Caserta, F, Eldred, WD., Fernandez, E, Hausman, RE., Stanford, LR., Bulderev, SV., Schwarzer, S, Stanley, HE: Determination of physiologically characterized neurons in two and three dimensions. J. Neurosci. Meth. 1995; 56:133-144.
53. Warfel, MD: http://www.cee.cornell.edu/~mdw/index.html 1997.
54. Li, DH, Ganczarczyk, JJ: Fractal geometry of particle aggregates generated in water and wastewater treatment processes. Environmental Science and Technology, 1989; 23 (11):1385-1389.
55. Jones, C, Jelinek, HF: Wavelet Packet Fractal Analysis of Neuronal Morphology Meth 2001; 24(4):347-358.
56. Forest, SR, Witten, TA: Long-range correlations in smoke particle aggregates. J Physics A Math and Gen 1979; 12:L109-L117.
57. Droppo, IG, Flannigan, DT, Leppard, GG, Liss, SN: Microbial floc stabilization and preparation for structural analyses by correlative microscopy. Water Science and Technology, 1996; 34 (5-6), 155-162.
58. Soddell, JA, Seviour, RJ: A comparison of methods for determining the fractal dimensions of colonies of filamentous bacteria Binary 1994; 6, 21-31.
59. Jelinek, HF, Elston, GN: Dendritic branching of pyramidal cells in the visual cortex of the nocturnal owl monkey: a fractal analysis Fractals 2003; 11(4):1-5.
60. Mandelbrot, BB, Pfeifer, P, Biham, O, Lidar, DA, Avnir, D: (1998) Editorial and Letters: Science, 1998; 279: 783.
61. Tsonis, AA, Eisner, JB, Avnir. D: Editorial and Letters: Fractality in Nature. Science, 1988;

279:784.

62. Bourke, P: Fractal dimension calculator user manual. http://www.mhri.edu.au/~pdp/frctals/fractdim.html 1993.
63. Nonnenmacher, TF: Digital image analysis of self-similar cell profiles Biomedical computing, 1994; 37:131-138.
64. Soille, P, Rivest, J-F: On the validity of fractal dimension measurements in image analysis. J. Vis. Com Image Repr. 1996; 7(3):217-229.
65. Struzik, ZR: From coastline length to inverse fractal problem: the concept of fractal metrology. PhD Thesis. University of Amsterdam. 1996.

Figure 1

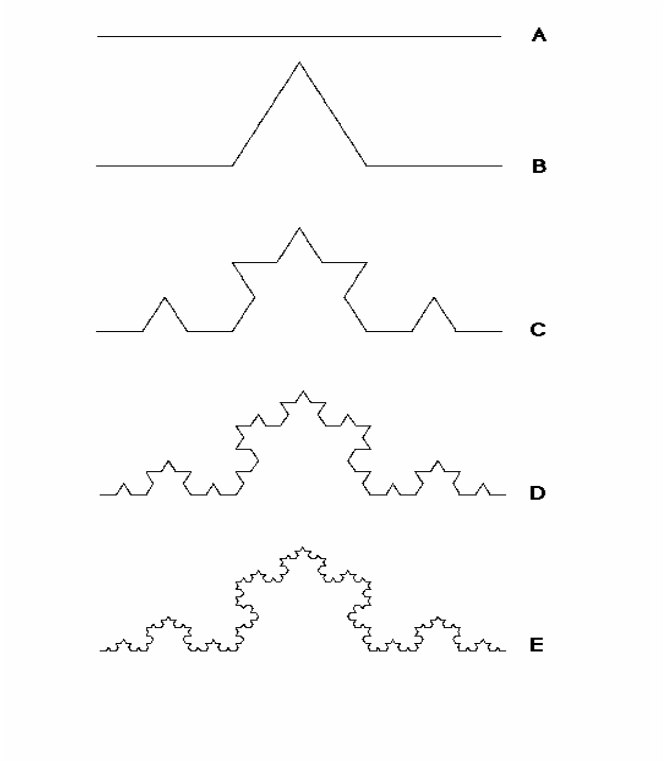
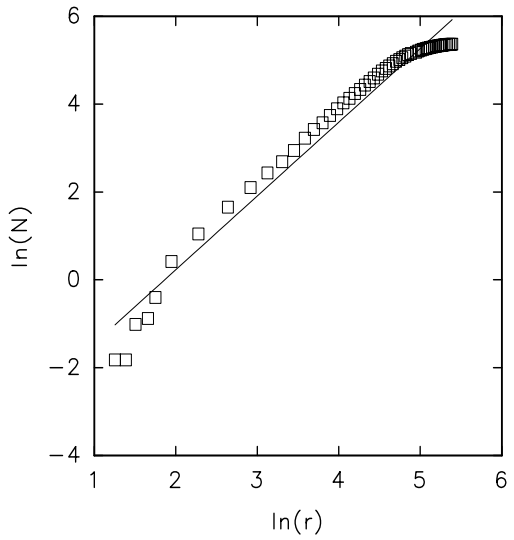


Figure 2



Legends

Figure 1

The Koch curve displaying the iteration process over several generations.

Figure 2

Double logarithmic relationship between the measuring scale (*1/e*) and the size of the image *N(e)*. The regression line through the data points indicates the range of statistical self-similarity and its gradient is proportional to the scaling exponent.

Figure 2

Method of applying mass-radius analyses to quantify the branching complexity of a fungal colony. The mean of 5 separate analyses of this image, using different centres of origin returned a mean Df=1.691 ± 0.017 r^2=0.967.

AMBIGUITY IN ART

Igor Yevin

Mechanical Engineering Institute, Russian Academy of Sciences

Short title: Ambiguity in Art

Address for corresponding:

4, Bardina, Moscow, 117324 Russia.
Phone: 7(095) 5760472 Fax: 7(095) 1357769
E-mail: yevin@list.ru

Key words: ambiguity, multi-stable perception, pattern recognition, art, criticality
Abstract. Non-linear theory proposed different models perception of ambiguous patterns, describing different aspects multi-stable behavior of the brain. This paper aims to review the phenomenon of ambiguity in art and to show that the mathematical models of the perception of ambiguous patterns should regard as one of the basis models of artistic perception. The following type of ambiguity in art will be considered. Visual ambiguity in painting, semantic (meaning) ambiguity in literature (for instance, ambiguity which V.Shklovsky called as "the man who is out of his proper place"), ambiguity in puns, jokes, anecdotes; mixed (visual and semantic) ambiguity in acting and sculpture. Complexity theory of the brain revealed that the human brain as a complex system is operating close to the point of instability and ambiguity in art must be regarded as important tool for supporting the brain near this critical point that gives human being possibilities for better adaptation.

NON-LINEAR MODELS PERCEPTION OF AMBIGUOUS PATTERNS

In perception psychology, multi-stable perception of ambiguous figures is often considered as a marginal curiosity. Nevertheless, this phenomenon is one of the most investigated in psychology. The first description of ambiguity was given by Necker in 1832. The most known examples of ambiguous figures are specially designed patterns such Necker' cube, "young girl-old lady" and so on. But visual and semantic ambiguity is very often connected also with that the available visual or semantic information is not sufficient by itself to provide the brain with its unique interpretation. The brain uses past experience, either its own or that of our ancestors to help interpret coming insufficient and therefore ambiguous information. Many patterns in our every day life, in a way, are ambiguous patterns, but using additional information, we usually resolve or avoid ambiguity [1]. Nikos Legothetis recently shown, that resolution of ambiguity is an essential part of consciousness job [2].
This paper aims to review and to familiarize with the present state the phenomenon ambiguity in art and to show that the mathematical models of the perception of ambiguous patterns should regard as the basic models of artistic perception. Ambiguous patterns are examples of two-state, bimodal systems in psychology. When we perceive ambiguous figure, like the fourth picture in the row on Figure 1, the perception switches between two interpretations, namely "man's face" or "kneeling girl" because it is impossible for the brain to recognize both interpretations simultaneously. Just like for any bifurcative state, it is impossible for ambiguous figure to predict what namely interpretation will appear first. G.Caglioti from Milan Politectic Institute firstly paid attention, that ambiguous figures are cognitive analogue of critical states in physics.
Various authors pointed out that perception of ambiguous figures possess non-linear properties, and that multistable perception could be modeled by catastrophe theory methods [3, 4, 5]

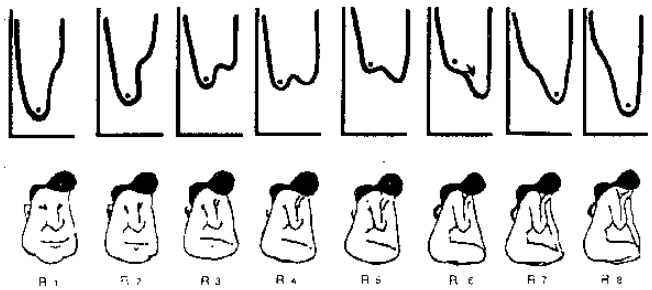


Figure 1. Ambiguous patterns are two-state systems. Their perception one can model by using elementary catastrophe "cusp"

The switch between two interpretations could be described by elementary catastrophe "cusp"

$$x^3 - bx - a = 0$$

where *a* and *b* are control parameters and x is the state variable. The first parameter *a* is called the *normal factor* and quantitatively describes the change in bias in the drawing in a "shape space" from a man's face to a woman's figure. Because this model may be used also for description of perception double meaning situations, it is reasonable to develop the idea of "shape space" on "meaning space" firstly introduced by Ch.Osgood [6].

The second parameter *b* is called the *splitting factor* or *bifurcation factor* and describes how much the amount of details is presented in the ambiguous figure.

The state variable *x* is presented as a scale from +10 ("looks a lot like a man's face") to˜ 10 ("looks a lot like kneeling girl"). For this model we could formally represent potential function $V = \frac{1}{4}x^4 + b\frac{1}{2}x^2 + ax$

which depicted on Figure 1, and consider catastrophic jump from one image to another as non-equilibrium phase transition. It is worth to note, that unlike to physical sciences, where potential function usually deduces from fundamental laws or standard theories, in mathematical models in psychology and others "soft sciences" potential function is hypothesized and really is considered as potential energetic function, which should be minimized. In this case it might be also considered as Lyapunov function in Hopfield's model of pattern recognition.

Actually, during the viewing of ambiguous figures, perception lapses into sequence of alternations, switching every few seconds between two or more visual interpretations.

Ditzinger and Haken offered an approach to the description of such oscillation under recognition of ambiguous figures [7]. Each pattern is described in this model as a vector in the space of quantitative parameters. There is a procedure for selecting non-correlated parameters, which enable to reduce an information volume. The most informative parameters are the order parameters (all they peculiarities occur near critical points, as in the case of order parameters near phase transition [7]).

Pattern recognition procedure is the following. First, pattern-prototypes are stored in the computer memory. Then, the pattern that should be recognized is inputted. The recognition dynamics is built in such a way, that its vector evolves in a parameter space to the most similar pattern stored in the computer memory.

The prototype patterns are encoded by $V_{i_i}\ (i = 1,..., M)$. The components of every vector encode the features of the patterns. We can decompose each prototype pattern into their pixels and denote a pixel by its index j. To each pixel we attribute a gray (or a color) value v_j. The set of all gray values v_j forms a vector $V_{i_i}\ (v_1, v_2,..., v_N)$. It is assumed that all these vectors are linearly independent. A pattern to be recognized is encoded by a vector $Q(0)$ and is inputted in a computer memory at $t = 0$. A dynamic of pattern recognition is constructed so, that the initial vector *Q*(*t*), is pulled into one of prototype patterns V_k with which it mostly coincides.

Recognized pattern is presented as the linear combination of prototype patterns

$$Q(t) = \sum_{j=1}^{M} d_i(t) V_i + \xi(t)$$

Where $d_i(t)$ is the order parameter, characterizing the degree to which a pattern is recognized, and $\xi(t)$ is a residual, uncorrelated with V_i (it is proved, that $\lim_{t\to\infty}\left|\xi(t)\right| = 0$).

The dynamic of pattern recognition is described as a gradient process in networks with only *M* neurons according to

$$\dot{d}_i(t) = \lambda d_i - (B + C) d_i \sum_{j \neq i}^{M} d_j^2 - C d_i^3,$$

$$\lambda_i > 0, B > 0, C > 0, d_i(0) = V_i^{'} Q(0)$$

This system has only the attractors of the type (0, 0,..., $d_k \neq 0$,...0). It can be shown that they must be either saddle points or nodes, but not limit circles (oscillations).

Figure 2. Image ambiguity: "young girl" – "old lady".

Ditzinger and Haken offered synergetic model of the perception of ambiguous patterns, describing dynamical features of such perception. It is based on the model of pattern recognition described above, and the model of the saturation of attention. The recognition of ambiguous patterns is reduced to inputting only two patterns-prototypes (e.g., "young girl" and "old lady") into computer memory with the order parameters d_1 and d_2. In this case the dynamics of pattern recognition is described in the following way:

$$\dot{d}_1 = d_1(\lambda_1 - A d_1^2 - B d_2^2)$$

$$\dot{d}_2 = d_2(\lambda_2 - B d_1^2 - A d_2^2)$$

$$\dot{\lambda}_2 = g(1 - \lambda_2 - d_2^2)$$

$$\dot{\lambda}_1 = g(1 - \lambda_1 - d_1^2)$$

where the overdot means $\frac{d}{dt}$, λ_1 and λ_2 are time dependent attention parameters, and *A*, *B*, and g are constants. The last two equations describe the saturation of attention in the

perception of prototype patterns. As analysis shows, the oscillation of perception occurs when the appropriate relations between constants are satisfied [7]. The recognition of ambiguous patterns has very profound and various analogies with numerous artistic phenomena. This model perception of visual ambiguous patterns also could be applied on the case of meaning ambiguity, because meaning perception also includes such phenomena as saturation of attention and the concept of the order parameter [8].

VISUAL AMBIGUITY IN ART

Let us first consider specially designed visual ambiguity in art. Painting by Giuseppe Arcimboldo "*The Librarer*" is an example one of the first pattern of such type ambiguity in painting. At first sight we recognize face, but a closer look reveals just an arrangement of different books.

Figure 3. Giuseppe Arcimboldo "*The Librarer*"

The most famous example of ambiguity in painting is, of course, *Mona Lisa* by Leonardo. In *The Story of Art* Ernest Gombrich said:
"Even in photographs of the picture we experience this strange effect, but in front of the original in the Paris Louvre it is almost uncanny. Sometimes she seems to mock at us, and then again we seem to catch something like sadness in her smile."
"This is Leonardo's famous invention the Italians call "*sfumato*" - the blurred outline and mellowed colors that allow one form to merge with another and always leave something to our imagination. If we now turn to the "*Mona Lisa*", we may understand something of its mysterious effect. We see that Leonardo has used the means of his "*sfumato*" with the utmost deliberation. Everyone who has ever tried to draw or scribble a face knows that what we call its expression rests mainly in two features: the corners of the mouth, and the corners of the eyes. Now it is precisely these parts which Leonardo has left deliberately indistinct, but letting them merge into a soft shadow. That is why we are never quite certain in which mood Mona Lisa is really looking at us. Her expression always seems just elude us" [9, p.228].
The ambiguity of Mona Lisa's smile one can compare with ambiguous images like "young girl - old lady". The oscillation in the perception of that painting can be described by Ditzinger-Haken's model.

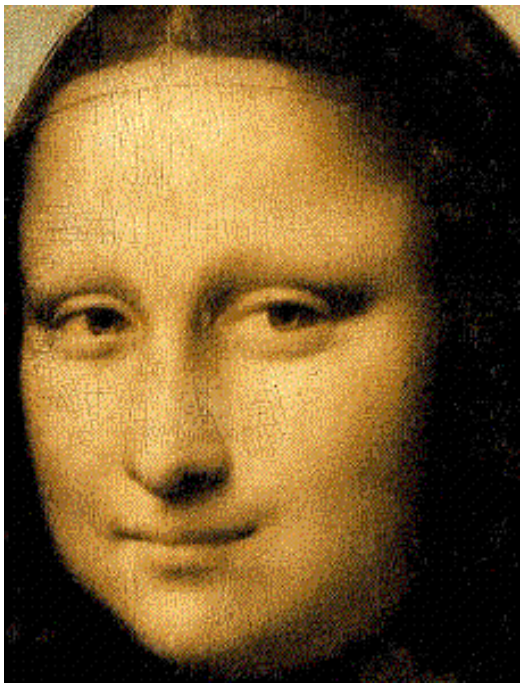


Figure 4. Ambiguity of *Mona Lisa's* smile.

Figure 5 gives an example other kind of visual ambiguity, when the human face is designed from different figures. This is the painting *Disappearing Bust of Voltaire* by Salvador Dali.

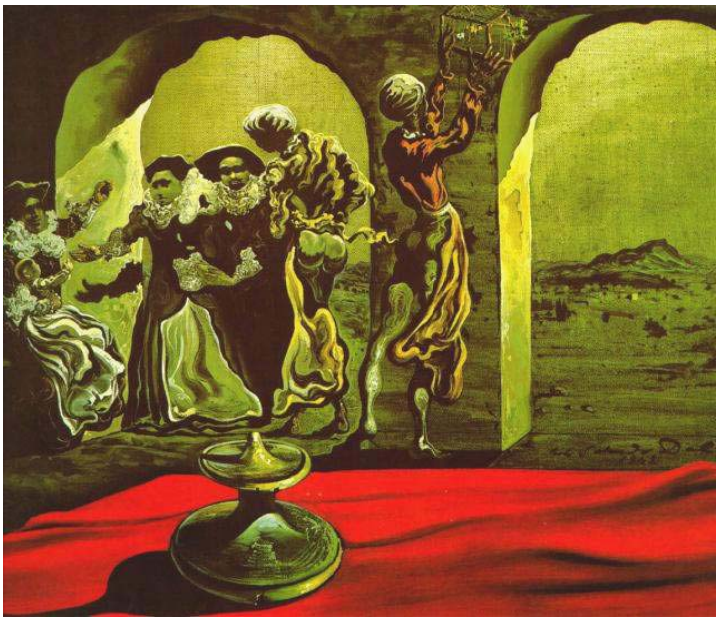


Figure 5. Ambiguity of Voltaire bust in Salvador Dali's painting *Disappearing Bust of Voltaire*

SEMANTIC AMBIGUITY OF VISUAL SCENES

Let us consider the following painting by J.Vermeer [11].

Figure 6. Jan Vermeer. *Girl with a Pearl Earring*

Many art historians consider this portrait as a masterpiece of ambiguity, because it offers huge amount of meaning interpretations. Her face is simultaneously somewhat sad and joyful, erotical and distant, submissive and yet dominant.
There is a humorous book called “Captions Courageous” by B.Reisner and H.Capplow attempting reinterpretation of famous masterpieces in painting – with more or less wit [12]. This possibility to create new interpretations for famous paintings which are perceived as comic is connected with insufficient information [1].

Figure 7. Complex visual scenes usually allow different meaning interpretations. This famous painting "*Birth of Venus*" by S.Bottichelli may be headlined as "Slip into this, it's a raid"(from the book "*Captions Courageous*" by B.Reisner and H.Kapplow [12]).

SEMANTIC AMBIGUITY IN PLOT DEVELOPMENT AND IN COMIC SITUATIONS

A significant type of ambiguity in art means the possible existence in artwork (most often in position of main hero) of two different states, one of them may be hidden until a certain time. A commonplace example of this form of instability exists in numerous book and movie plots in which a spy or Secret Service agent is hiding his identity while maneuvering about in hostile camp. At any moment, he may be unmasked, and the agent's task is to extend his secret identity as long as possible.

In well-known American movie "ROBOCOP" the main character is simultaneously a robot, incarnating an idea pitiless and perfect machine of revenge, and a human being, capable on deep and tender feelings.

Another, less- banal example, ambiguity of social nature - what V.B.Shklovsky describes as "the man who is out of his proper place" - is also widely presented in art [13]. According to Victor Shklovsky, "…The novels of olden times were based on the incidental falling out of a man from his society. By almost fatal position of an illegal birth or a loss of documents has been motivated an introduction of a hero into the world which has been a destiny of millions. But a hero has been rescued from that world with the help of recovered documents. A hero of old novel - is the one who is not in his proper place, a hero of one society finding himself in an absolutely different society." [13]. "The History of Tom Jones, a Foundling" by Henry Fielding, is an example of such a novel. The main character Hlestakov in the play by N.Gogol "Inspector General" obviously one may also describe using this kind of ambiguity.

In Apuleius's "Golden Ass" the main character is, of course, out of his proper place because the ass in reality is a man. The plots of such tales like "The Ugly Duckling" by H. Andersen and "The Beauty and the Beast" also are of the same type of ambiguity, sustained over the entire period of the plot.

In the majority of the novels by Agatha Kristy we deal with semantic ambiguity, as almost any character of these novels could appear as the murderer. This state of semantic ambiguity is skillfully supported by the author down to an outcome of the plot: "You know that I never deceive. I simply speak something such, that it is possible to interpret double" -once confessed A.Kristy.

Without ambiguity of natural languages, the existence of poetry is impossible. According to A.N.Kolmogorov, entropy of language H contains two terms: meaning capacity h_1 - capability to transmit some meaning information in a text of appropriate length, and flexibility of language h_2 - a possibility to transmit the same meaning by different means [14]. Namely h_2 is a source of poetic information, and the ambiguity of language is one of the causes of it's flexibility. Languages of science usually have $h_2 = 0$, they exclude ambiguity, and cannot be used as a material for poetry. Rhythms, rhymes, lexical and stylistic norms of poetry will put some restrictions on a text. Measuring that part of the ability to carry information spent on those restrictions (denoted as β), A.N.Kolmogorov formulated the law, according to which poetry is possible: if $\beta < h_2$. If the language has $\beta \geq h_2$, than poetry is impossible. We know that the brain resolves a visual ambiguity by means of oscillation. A semantic ambiguity (the ambiguity of meaning) is a result of ambiguous words or whole sentence. Semantic ambiguity, wide spread in comic situations, also resolves by oscillations.

Ambiguity of humor is often a clash of different meanings. It involves double or multiple meanings, sounds, or gestures, which are taken in the wrong way, or in incongruous ways.

Here is D.D.Minayev's epigram:

"I am a new Byron" - you proclaim.
I can agree with you:
The British poet was lame
The rhymes of yours are also lame."

The method used in this epigram is connected with a comparison based on different distant meanings (Byron was the lame, and a vain poet was also a lame, but in his rhymes). The situation described in this epigram is common to a lot of semantically ambiguous comic situations, which contain two states. One state we should call a state with high social status. This position is honorable and sometimes brings profit. The second state we should call a state with low social status. Everybody avoids occupying it. In the aforesaid example, the state with the high social status ("a good poet") we connect with words "a new Byron". Another poet is trying to get this state. But the author of the epigram unexpectedly transfers a poet to the second state with a low social status. This state we connect with the words "the rhymes of yours are also lame". Such an unexpected leap is achieved by using the same word ("lame") for totally different states.

So, a feeling of comic is very often connected with sudden transition from a state of high social status to a state of low social status, or the other way round. Is it a single transition? Does it happen only once? Not, of course. It is a multistable perception of meaning. The rhythmical, repeating nature of laughter (ha-ha-ha, etc.) shows that such transitions are repeated. Evidently, a laughing person mentally oscillates every time from the state of high social status to the state of low social status and *vice versa*, by comparing them. As a result, the rhythmical laughter is generated by the nervous system.

Let us consider also the following anecdote about Sherlock Holmes and Dr.Watson

Sherlock Holmes and Dr. Watson are going camping. They pitch their tent under the stars and go to sleep. Sometime in the middle of the night Holmes wakes Watson up.

"Watson, looks up at the stars, and tell me what you deduce."

Watson says, "I see millions of stars, and if there are million of stars, and if even a few of those have planets, it's quite likely there are some planets like Earth, and if there are a few planets like Earth out there, there might also be life."

Holmes replied: "Watson, you idiot, somebody stole our tent".

We see that Watson and Holmes offered two different semantic interpretations of the same visual picture of star sky and if Watson gave namely one of possible interpretation of picture of star sky, Holmes paid attention on semantic context of this picture and connected it with their rest position.

The origin of the oscillatory character of laughter should be connected with the fundamental property of the distributed neuron set, i.e. as the oscillation occurring in the perception of ambiguous patterns. According to Ditzinger-Haken's model of recognizing of ambiguous patterns, stable limit cycles can be formed in systems of usual nonlinear differential equations for those variables, which describe the visual perception (e.g. attention). Evidently, this is the common characteristic of distributing neuron sets. That's why it is manifested not only in evolutionary low stages (the ancient visual-morphologic structure of nervous and psychological activity of a human being), but also in its latest stages as well (in the semantic-analytical structures of the left cerebral hemisphere).

Comic situations are very often connected with polysemantic, i.e. semantically ambiguous, situations. Another situation of perception of ambiguous patterns occurs in a parody of a famous person by some actor. On one hand, we can recognize the manners, gestures, style and voice of that famous person. On the other hand, we see quite a different person. The same method is used in literary and poetic parodies. Every time we are dealing with a bimodal, double-meaning situation. As a result, we have the oscillation of perception, and laughter is one of the external manifestations of this oscillation.

One can assume that in ambiguous comic situations oscillations occur between two semantic images. The phenomena of synchronization are typical for a self-organizing process in an active medium (and the nerve substance is an active medium). From that, we can conclude that the period of oscillation between semantic patterns coincides with the period of outward macroscopic oscillations, manifested as laughter with the duration of about 0.1 sec. This value is much smaller than the oscillation period, which occurs when recognizing ambiguous figures (1-5 sec.).

Why does laughter occur in the perception of double-meaning situations, and not in the visual perception of ambiguous patterns? We can explain this by essentially different periods of the

corresponding oscillations. In the visual perception this period is approximately equal to t=10 sec., and in the perception of the ambiguity of meaning - this period is about t=0.1 sec. That difference could be explained by the fact that a much smaller mass of nerve substance is involved in creating semantic patterns, compared with constructing visual patterns. This is because visual information is processed in the massive and ancient visual cortex, and semantic patterns are interpreted in compact Broke-Vernike zone in the left brain hemisphere. Anecdotes, jokes and sketches deliberately are created as short as possible (laconic), in order to reduce the time needed for the saturation of attention in the process of recognition.

MIXED AMBIGUITY

Ambiguity of sculpture

We have considered visual ambiguity in painting (see also [11]) and semantic ambiguity in jokes, anecdotes and puns. Let us consider mixed (visual and semantic) ambiguity, taking an example from sculpture art. Sculpture involves an ability to depict representatives of living nature (most often humans and animals) from materials of inanimate nature (wood, stone, bronze, etc).

In creativity of different sculptures one can observe a prevalence of one of these phases with respect to another. In Michelangelo's works we see triumph of alive and even spiritual under inert matter of stone. E.Gombrich wrote in book “The Story of Art”: “While in “*The Creation of Adam*” Michelangelo had depicted the moment when life entered the beautiful body of a vigorous youth, he, now, in the “ *Dying Slave*”, chose the moment when life was just fading, and the body was giving way to the laws of dead matter. There is unspeakable beauty in this last moment of final relaxation and release from the struggle of life - this gesture of lassitude and resignation. It is difficult to think of this work as being statue of cold and lifeless stone…”.

It is interesting to note, that ambiguity of sculpture art influences on literature, because the plots of some works of arts in literature are based on the idea of animated statue - that is, the transition "inanimate-animated" (such as opera "Don Giovanni" by Mozart, "Copper Horseman", "Stone Guest" by A.Pushkin and, of course, in ancient legend about sculptor Pygmalion).

Ambiguity of dolls

In the essay “Dolls in system of culture” Yu.Lotman marks ambiguous (as well sculpture) nature of this cultural phenomenon closely connected to ancient opposition alive and dead, spiritual and mechanical. At the same time, as against a sculpture, the doll demands not contemplation but play. It serves as a certain stimulator provoking creativity [15].

Ambiguity of acting

Like any human being, an actor has in his everyday life some set of rather stable physiological and psychological personal properties: sex, appearance, timbre of voice, gait, temper, and so on. The acting involves ability to create a second phase, a "role" phase, different from the original physiological and psychological nature of the actor. In other words, a bimodal "actor-role" state created may be compared with ambiguous patterns, for instance, the pattern where we see in turn "young girl" or "old lady". One may say that in this case young girl will "play the role" of old lady and *vice versa*.

In acting, one can observe the existence of two polar types of actors:

1. An actor as a bright, brilliant individuality, eccentric person with the original appearance, and so on (Alain Delon, Arnold Schwarzenegger). It is rather easy to make a parody of such actors;
2. An actor with prominent outstanding abilities for transformation and reincarnation (Laurence Olivier, Alec Guiness). In that case, it is very difficult to make a parody.

Yu.Lotman note, that in the cinema more, than at the theatre the spectator sees not only role, but also actor [15, p.658]. Observing play of the famous actor we alternately focus our

attention or on guise (image) of actor familiar to us on other movies, or on peculiarities of a role, which the actor plays. Such oscillation of attention is the reason, that with the reference to acting we use a word “play”.
In the case of acting the prototypes are, for instance, "Laurence Olivier" (the image of actor) and "Othello" (the image of character). Therefore, according to the common law of perception of ambiguous patterns, the oscillation of our attention takes place, and we see in turn either an actor or his role.
Just as like bimodal nature of sculpture art begets plots about animated statue, bimodality of actor art gives a possibility to use a phase transition called "character invasion" for plot development [16].
The main hero of the film "A Double Life" plays the role of Othello for so long time that it begins to affect to his psychic activity, making him more and more jealous of his beloved, and like the stage character, he strangles her and then kills himself. In the film "Jesus of Montreal" the actor playing the role of Jesus Christ becomes transformed into a Christ-like figure [16].
As a rule, all bimodal metastable states in the end of movies turn into stable, onemodal states as a result of bifurcation.

CONCLUSION

In ordinary speech, and especially in scientific communication, in general we try to avoid ambiguity. By contrast, in humor, one of the aims is to create ambiguous situations to provoke laughing. And in art as a whole ambiguity is an indispensable, necessary part.
“…art is supposed to have multiple meanings. It is self-defeating to increase one aspect of meaning. The more a single meaning dominates a work, the less it is a work of art. Something that has one and only one meaning – no matter how interesting or important that meaning is -- is no longer a work of art” [17, p.46].
Understanding ambiguous nature of art gives an ability to formulate problem computer simulation of artistic creativity as computer making of ambiguous images and meanings (for instance, by means of genetic algorithms) [18].
Synergetics and the theory of complexity revealed that the human brain operate near unstable point, because only near criticality the human brain could create new forms of behavior. Ambiguity in art is an important tool maintaining the brain near this unstable, critical point.

References:

1 Kruse P, Stadler M: Ambiguity in Mind and Nature: Multistable Cognitive Phenomena. Berlin, Springer, 1995.
2. Legothetis NL: Vision: A window on consciousness. Scientific American. 1999; 11: 69 75
3 Poston T, Stewart I: Nonlinear model of multistable perception. Behavioral Science.1978; 23 (5): 318 334.
4 Stewart IN, Peregoy PL: Catastrophe theory modeling in psychology. Psychological Bulletin, 1983; 94(21): 336 362.
5. Ta'eed LK, Ta'eed O,.Wright JE: Determinants involved in the perception of Necker cube: an application of catastrophe theory. Behavioral Science. 1988; 33: 97 115
6. Osgood Ch, Suci G, Tannenbaum P: The Measurement of Meaning. Urbana, University of Illinois Press, 1957.
7. Haken H: Principles of Brain Functioning. Berlin, Springer, 1996.
8. Wildgen W: Ambiguity in linguistic meaning in relation to perceptual multistability; in: Kruse P, Stadler M (ed): Ambiguity in Mind and Nature. Berlin, Springer, 1995.
9. Gombrich E: The Story of Art. New York, Phaidon, 1995.
10. Caglioti G: Dynamics of Ambiguity. Berlin, Springer, 1992.
11. Zeki S: Inner Vision. Oxford, Oxford University Press, 1999.
12. Reisner B, Kapplow H: Captions Courageous. New York, Abeland-Schuman, 1954.
13. Shklovsky VB: Tetiva: Moscow, Iskusstvo, 1967. (In Russian)
14. Kolmogorov AN: Theory of Poetry. Moscow, Nauka, 1968, pp.145-167 (in Russian)

15. Lotman Yu: About Art. S.Peterburg, Iskusstvo, 1998. (In Russian)
16. Neuringer C, Willis R: The cognitive psychodynamics of acting: Character invasion and director influence. Empirical Studies of the Arts.1995; 13; 1:.47 54
17. Martindale C: The Clockwork Muse. New York, Basic Books,1990.
18. Yevin I: Phase transition as a fundamental model of creativity. Proceedings of International Conference on Informatics and Control. St. Petersburg, Russia, June 9-13, 1997, pp.502-505

Reconstructing the rules of 1D cellular automata using closure systems *

José L. Balcázar, Gemma C. Garriga, and Pablo Díaz-López

LARCA Research Group
Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Campus Nord, Jordi Salgado 1-3
08034 Barcelona, Spain
{balqui,gcasas,pdiaz}@lsi.upc.edu

Abstract. We consider the problem of identifying the rules conforming the local map of a cellular automaton; we explore the capabilities of a closure-based algorithm for this task. The algorithm has been previously proven to identify an optimal Horn-like formula true for the data, in a very precise mathematical sense. A key property of the algorithm is its ability to handle a sequential structure on the data and lift it to the Horn-like rules, thus making it apt to compare the rules it obtains with the ones that originated the data. The outcomes of the experimentation are described.

Keywords: cellular automaton, propositional Horn logic, association rules.

1 Introduction

Closure systems form a very basic mathematical concept, related to many applications. In the field of Formal Concept Analysis, closure systems have been used widely to represent knowledge and to infer rules from data; and some extensions of this work, such as the notion of confidence-based association rules, became recently a cornerstone of the field of Data Mining. In the recent years, the authors have developed appropriate notions of closure systems for the analysis of structured data, notably in the form of sequential itemsets. Up to now, such systems, and similar ones, have been used to model various sorts of data found in diverse application domains; most notably, web-based and user interaction data, and other technical information. We have also constructed, under the DELIS Integrated Project of the Complex Systems Initiative, an implementation in the form of a research toolkit, the ISSA system, that includes algorithms for the analysis of sequential closure systems, one of which is our generalization of deterministic association rules to sequential data. Here we describe the application of this sort of analysis to cellular automata.

Cellular automata are computational systems based on sets of simple rules, introduced back in the 1940's. They operate on a space divided into cells, organized into a regular structure (usually, low-dimensional rectangular or hexagonal grids or tori) with a clear, uniform notion of neighborhood of each cell. Cells may change state among a (usually small) number of states, according to the so-called *local map*, the set of rules that govern the evolution of the system along discrete time steps. Each rule in the local map specifies the change of state

* This work is partially supported by the European Commission - Fet Open project DELIS IST-001907 Dynamically Evolving Large Scale Information Systems and by MCYT TIC 2002-04019-C03-01 (MOISES).

(or absence of change) of a cell on the basis of the configuration of states of cells in the neighborhood of the cell itself.

For instance, in an one-dimensional grid, the state of a cell at a given time will depend on the states of a number of cells in the previous moment: the cell itself, the k cells to its right, and the k cells to its left. These $2k + 1$ cells constitute the neighborhood. These 1D cellular automata are frequently employed to create "textures" on a 2D space or image, by choosing one row of pixels (usually the topmost or the leftmost one) and drawing successive rows of pixels according to the evolution of the cellular automaton; some famous geometries and many interesting evolutions can be obtained in this way, including models of physical phenomena, such as heat-flow and turbulence, as well as computer-generated 2D textures, some of which reproduce extremely well human perceptions of real textures. Let us just mention as a simple example the Pascal triangle of (the parity of the value of) binomial coefficients, showed in figure 1. It is readily verified that this picture is obtained as the evolution of a very simple cellular automaton whose cells can be in two states (ON and OFF, say), and where the local map specifies that a cell is ON exactly if, in the previous time step, exactly one of the two immediate neighbors was ON. The start configuration has exactly one cell ON.

Fig. 1. Pascal's triangle of binomial coefficients obtained with a cellular automaton

Using the AND boolean function instead of the parity for combining the values of the two neighbors gives Wolfram's well-known AND automaton; slightly more involved rules are able to create extremely sophisticated behavior, including self-reproduction. Even insisting on limited 2D grids with just two states per cell and a neighborhood relation that considers, for each cell, the eight immediate cells surrounding it, the very simple and famous rules of Conway's LIFE already give rise to a fascinating system of which a major fact has been proved: it can simulate arbitrary computations and has full Turing computational power. And, with a similar neighborhood and the apparently small extension of cells with 3 states, the amazing evolutions of Brian's Brain are well worth admiration. The Modern Cellular Automata web page provides plenty of material to read, learn, or just enjoy with the visualization of cellular automata at work.

Here, we focus our attention on the problem of learning the set of hidden rules that run the evolution of a cellular automaton. Starting from the sequence of evolutions through the time and considering that the set of rules that generated such evolution is not known, we would like to discover an approximation of those rules. There are some previous works along this sort of analysis. Several of them attempt at modeling textures generated by a sweeping 1D automaton, by identifying "coherent structures" along the spatiotemporal distributions provided by the evolving system, constructing filtering systems for detecting specific phenomena in these evolutions (see [1] and the references there).

We rather take a somewhat different standpoint that the problem of synthesizing the unknown local map rules from the behavior of the system can be seen as a learning problem

that can be addressed with the techniques of knowledge discovery. Previous works (some of them very recent, which shows the timeliness of our study) include a few cases of statistical analysis (such as fitting hidden Markov models via expectation-maximization [2] or applying a Minimum Description Length approach to approximate a stochastic form of cellular automata with probabilistic local maps [3]), some proposals based on genetic algorithms (such as [4] and [5]) and, closer to our case, the use of Data Mining techniques for rule induction [6]. In this paper, we will study the sequence of evolutions of a cellular automaton by means of a recent variant of association rules that is particularly well-tailored to the study of evolving systems, namely, those obtained from a notion of closure operator on sequentially structured data proposed recently by the present authors [7].

Specifically, we proved there that a novel closure system defined in [8], appropriately employed, can extract from sequential data a family of implicational rules that can be mathematically characterized in terms of Horn logic for a propositionalization of the sequential setting, as the empirical Horn approximation of the data, that is, the set of Horn rules that minimally describe the given set of evolutions. Our question now is to what extent these rules can uncover the hidden function that governs the evolution of a cellular automaton. Given that the ISSA system alluded to above implements the corresponding algorithmics so that we are indeed able to operate on real data through these conceptual mechanisms, we are in a position to gather some experimental evidence of the strengths and weaknesses of this data analysis method for the task of reconstructing rules of cellular automata.

The setting of our work is as follows: given is a (large enough) cellular automata, constructed by ourselves so that the rules that govern it are known; the local evolution of a small neighborhood is extracted for each cell, and the data obtained in this form is fed into our ISSA system. Properly handled, the outcome is a set of rules. Then we compare these rules to the ones that actually were used to construct the cellular automaton. Success is defined when the two sets of rules are logically equivalent; but there is absolutely no guarantee of such success, since ISSA was conceived for the analysis of extremely different data.

2 Horn rules, closure systems, and sequences

Assume a standard propositional logic language with a finite set of propositional variables. A literal is either a propositional variable, called a positive literal, or its negation, called a negative literal. A clause is a disjunction of literals. A clause is *Horn* if and only if it contains at most one positive literal. A Horn formula is a disjunction of Horn clauses. A *model* is a complete truth assignment, i.e. a mapping from the variables to $\{0, 1\}$. A set of models is Horn if there is a Horn formula which axiomatizes it, in the sense that it is satisfied precisely by the models in the set.

Then, the minimal Horn set of models including a given set is known as the empirical Horn approximation. A Horn formula axiomatizing it can be constructed as the conjunction of all the Horn formulas that are true of all the models of the original given set.

A Horn formula defines a closure system on the variables. The closure of a set of variables is formed by all those variables that are consequences of those in the set through the implications indicated in the Horn formula. Dually, from a closure system, there are a number of ways of obtaining implications that actually have been proved to correspond to a Horn formula axiomatizing the empirical Horn approximation.

2.1 Identifying Horn rules from a set of sequences

In this section we revisit the results presented in [7]. There, a notion of deterministic association rules is defined from the Galois connection framework of [8]. The set of all the rules obtained from this process turns out to define exactly the natural extension of the notion of empirical Horn approximation to a set of sequences.

More specifically, we are given an input set of sequences $S = \{s_1, \ldots, s_n\}$, where each sequence in this set S is defined to be an ordered list of sets of variables $s_i = \langle (V_1)(V_2) \ldots (V_m) \rangle$; with this notation we mean that the set of variables V_i occurs before the set V_j for $i < j$.

This set S can be transformed into a lattice of closed sets of sequences (see [8, 7]); and from there, the work in [7] derives a proper notion of generator for each closed set. The idea is that each generator will correspond to the antecedent of a rule, and its closure to the implied the consequent. Based on this formalization, it is possible to derive the notion of association rules that deterministically hold for all the sequences $s_i \in S$. These association rules have the form $S' \rightarrow s$, where S' is a set of sequences (i.e. the generators), and s is a single sequence being the consequent.

The main property of these rules is that they hold in all the input sequences s_i, that is, for each $s_i \in S$ containing all the sequences S', it holds that $s \subseteq s_i$ as well. It can be proven that these proposed rules can be formally justified by a purely logical characterization, namely, a natural notion of empirical Horn approximation for ordered data. The algorithmic procedure to come up with such Horn implications is discussed in [7].

3 Horn rules in cellular automata

Whereas the natural application of Horn rules for sequential data is in the data mining realm of the analysis of ordered transactional data, here we tackle the somewhat more challenging problem of using it for the analysis of a very different form of information: the evolutions of cellular automata. Below we discuss in some more depth the reasons why our method could encounter difficulties in the analysis we develop. We use the ISSA implementation of the method of analysis described in the previous section. This implementation offers the additional possibility of discarding those Horn rules whose frequency of apparition in the data is below a user-tunable threshold. Several values for this threshold have been used in the experiments.

3.1 Details of the setting

Our results in this paper are still rather preliminary. We have taken the following initial working assumptions:

- we limit ourselves to two-state 1D automata where the state of each cell depends only on the states of the two neighbors in the previous generation, which is a simplification that still leaves in the picture well-known complex systems such as Wolfram's AND automaton and also the parity automaton that is able to trace the self-similarity structure of the Pascal triangle;
- we assume the initial generation to be a random initialization (which is always the same for each size to guarantee reproducibility);

– each step is encoded by the previous and current states of the left neighbor (l or L), of the right neighbor (r or R) and of the cell itself (c or C), so that each piece of data that the algorithms receive have a form similar to [(l,c,R),(l,C,r)], meaning that at some evolution of the system, at one particular spot, the cell goes from state False (c) to True (C), and at the same time the left neighbor was and remains in state False (l), and the right neighbor changes from True (R) to False (r).

Under these circumstances, we analyze the data repeatedly adjusting a number of parameters:

– the number of cells,
– the boolean function that updates the state,
– the number of generations, and
– the threshold of significancy, the internal parameter of ISSA mentioned above, whose operation indicates the system that a configuration which appears with an empirical probability below the threshold should be disregarded.

3.2 Results

There are a number of reasons why ISSA must be expected to have difficulties in finding the successful rules. First, it is not informed that the cell data of the second 3-tuple depends on the left and right data on the first, nor that all the other states depend on information that is not available to it. Second, although actually the next piece of data it receives corresponds to the right neighbor configuration and pieces of data are correlated, the algorithm is not informed of this fact.

Third, and more seriously, ISSA works in a purely propositional form so that the rules it is able to output are restricted to Horn clauses over partial orderings labeled by the literals l, c, r, L, C, R. Thus, in principle, ISSA does not have explicitly enough expressive power to say that the state of the cell becomes, say, the parity of the states of the neighbors; as we see in a moment, ISSA provided us with the surprise of finding by itself a way of expressing the necessary correlations.

Fourth, and finally, given that the initial configuration is random, there is no guarantee that all the potential instances of the rule employed (that is, all possible combinations of states of the left and right neighbor) appear in the evolution of the system with a frequency above the significancy threshold.

Of all these difficulties, it turned out that the analysis power of ISSA did not seem to be affected by the first two; and that it gave us a way of encoding, in a form developed by the program itself, the rules of the automaton, under the appropriate values of the parameters. The fourth one, though, was the most relevant one: if the significancy threshold is set too high, the description of the rules easily misses cases that were not frequent enough in the specific evolution under analysis, whereas when it is lowered too much, we start finding correlations that do not happen often enough to make sure that counterexamples are found, and thus we get as output rules that are, essentially, statistical noise.

As for the effect of the other parameters, they are as expected: more cells give better results (that is, better chances of success with respect to the random initialization), whereas more generations immediately lead the system astray into repeated failures. We found similar results for all the boolean functions (modulo symmetries due to permutation of left and right or due to negation).

As examples of how ISSA describes the rules behind the automaton, we indicate its results on a large automaton working under a disjunctive rule (that is, the state of a cell is the OR of the two states of the neighbors in the previous generation):

- if 'l c' then 'r c', meaning that if the cell is False and in the previous generation its left neighbor is False, then in the previous generation the right neighbor was False as well
- dually, if 'r c' then 'l c'
- if 'l C' then 'R C', meaning that if the cell is True and in the previous generation its left neighbor is False, then in the previous generation the right neighbor was True
- dually, if 'r C' then 'L C'

Similarly, for the AND function, we get rules such as: if 'L c' then 'r c', meaning that if in the previous generation the left neighbor was True, and the cell is False, then in the previous generation the right neighbor was False as well.

4 Conclusions

Under appropriate parameter settings, ISSA is able to extract reasonable rules from simple 1D cellular automata evolutions, even though it was originally designed for searching for very different correlations in very different datasets. Forthcoming work would consist in experimenting with more complex cellular automata (larger neighborhoods, 2D systems) just enough to glean intuitions, and then try and extend some of the ISSA features into enough first-order logic to transcend the propositional limitation, which is now by far the most restrictive facet.

It must be mentioned as well that these experiments have given us progress not only about cellular automata, but also in understanding the deep consequences of our proposals for the analysis of sequential data based on closure systems. Some of the details of the ISSA implementation were actually motivated by the experimentation we ran on cellular automata data.

References

1. Shalizi C.R., Haslinger R., Rouquier J.B., Klinkner K.L., Moore C. Automatic filters for the detection of coherent structure in spatiotemporal systems. arXiv:nlin.CG/0508001.v1. 2005.
2. Turin W. Fitting probabilistic automata via the EM algorithm. Communications in Statistics – Stochastic Models 12, 405–424. 1996,
3. Craiu R.V., LeeT.C.M. Pattern generation using likelihood inference for cellular automata. IEEE Transactions on Image Processing (to appear). 2005.
4. Richards F.C., Meyer T.P. and Packard N.H. Extracting cellular automaton rules directly from experimental data. Phys. D, 45:1-3, 189–202. 1990.
5. Yang Y.X., Billings S.A. Neighborhood detection and rule selection from cellular automata patterns. IEEE Transactions on Systems, 30:6, 840–847. 2000.
6. Li X. and Gar-On Yeh A. Data mining of cellular automata's transition rules. International Journal of Geographical Information Science, 18:8, 723–744. 2004.
7. Balcázar J.L., Garriga G.C. On Horn Axiomatizations for Sequential Data. Proceedings of the 10th Int. Conference on Database Theory, 215–229. 2005.
8. Garriga G.C. Towards a Formal Framework for Mining General Patterns from Structured Data. Proc. of the 2nd. Multi-relational Datamining, in KDD Int. Conf. 2003.

Statistical Physics of Boolean Control Networks

L. Correale, M. Leone, A. Pagnani, M. Weigt, R. Zecchina
leone@isiosf.isi.it

We explore the static behavior of large Boolean control networks with methods recasting the Boolean regulation problem into a constraint satisfaction problem. Our analysis includes a modified version of the leaf-removal procedure as well as results coming from the belief and survey propagation algorithms. This allows to explore the complex solution-space structure of the problem. We find a phase transition from simple to complex regulatory control, and identify relevant regulatory variables which select the fixed points of the network within the global solution space.

In the last two decades a large wealth of new data about the organization of gene-regulatory networks (RNs) is being collected It has become clear that in many cases biological functions cannot be identified at the level of single genes and proteins. Constructing a detailed biochemical model of an entire cell by analyzing each gene and its interactions with others one by one appears to be a formidable task. Even in the case of a relatively simple cell like yeast, the number of genes is of the order of 6000. To make progress in the understanding of the combinatorial aspects of genes regulation it appears crucial to model these systems on a coarse-grained level, and to encode the interactions in a so-called RNs: nodes of this network represent individual genes, and directed connections genetic interactions. Such models can be expected to give insight into the collective large-scale behavior of the gene as well as more general regulatory mechanism, and to infer emergent cooperative phenomena which cannot be understood at the level of single components.

A possible way of (coarse grain) modeling RNs is to consider them as bipartite factor graphs made of N genes that are regulated by a set M function nodes. Gene expression is modeled by variables taking a discrete set of activity levels $x_i in level set 0, 1, ..., q. A function node directed from i_1, .., i_k (genes, proteins, regulatory regions, etc.) to factor "i" in the graph implies that the first K factors plays the function F_i(x_j in I(i)) determines the expression level x_i. One function node regulates one gene. External in degree zero, appearing only as regulating variables. The binary case of q = 1 is well-known as Boolean network protein interaction networks.$

Despite the large wealth of work on the subject, which however mainly addressed the behaviour of boolean control networks under rather artificial updating dynamics, the issue of fully understand the nature, organization, stability and accessibility of the fixed points of this problem is still a largely open question, since so far numerical techniques have been based mainly on exhaustive enumeration or monte-carlo approaches. The first allowing only for the study of much smaller regulatory subnetworks, the latter failing to extensively explore the solution space structure. The fixed points problem can however be exactly mapped in a Constrained Satisfaction problem, and studied with techniques explicitly developed and optimized in the fields of combinatorial optimization and Bayesian inference.

We extensively explored the fixed points phase space of the K=1, K=2, K=3 and mixed models, in the case of large random regulatory networks with poissonian and scale-free outgoing link degree distributions. For certain Boolean nodes types (namely the XOR boolean function in the K=2 case and all but pure AND and OR functions in the K = 3 case), a region in the phase space is present where the space of solution clusterizes in a large number of well separated thermodynamical states.

The onset of this region cam be studied in some particularly simple cases with a modified leaf-removal procedure, but a correct characterization of it can only be done exploiting a multi-state clustering hypothesis for the phase space, that can be efficiently studied with the Survey Propagation algorithm. The nature of this clusterized phase and the reasons why some boolean functions show a more dramatic behaviour than others with this respect

is quantified and explained.

A precise hierarchy of nodes exhibiting this behaviour is presented and justified. Some conjectures on the universality of the clustering phenomenon are presented and fortified with numerical results.

The meaning of the clustering phenomenon in terms of (biological) control properties is proposed through the application of previous theoretical results to real large gene networks.

Algebraic Hierarchies of cellular automata classes

Marianne Delorme, Jacques Mazoyer, Guillaume Theyssier
LIP, ENS Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France
e-mail : *Marianne.Delorme, Jacques.Mazoyer, Guillaume Theyssier@ens-lyon.fr*

31st August 2005

Keywords : Cellular Automata, Complexity, Computation, (Intrinsic) Universality, Algebraic Classifications, Methods.

Cellular automata are a formal model of locally interacting systems. They are syntactically simple but can present extremely complex behaviors, which make them suitable to study complex systems in general. Many classifications have been proposed in literature [1], often relying on the observation of dynamics. In a first part, we present more recent approaches of algebraic nature based on notions of sub- or quotient systems. A second part is dedicated to new results concerning these algebraic tools. Actually this framework allows to set formal definitions for intuitive global notions and to prove new positive results but also, more interestingly, negative ones. More precisely, we show that modifying local rules may be more powerful in some sense than increasing the number of states; then we illustrate by the construction of an infinite lattice that dynamical universality is more powerful than usual computation universality.

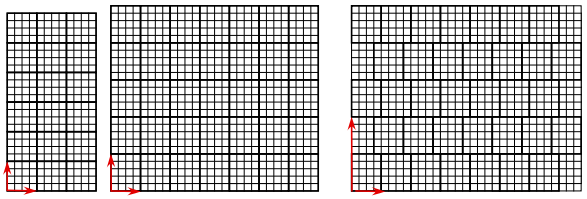


Figure 1: Squar tiling, rectangular tiling and shifted rectangular tiling.

Our approaches are to define "natural" but tractable comparison criteria of orbits (also called space-time diagrams in the case of cellular automata) and then to derive comparison criteria of sets of orbits inducing comparisons on cellular automata themselves.
Let us explain it in case of dimension 1 for sake of simplicity. The bi-dimensional underlying graphs of space-time diagrams are tiled by one (and only one) tile. Actually only bi-periodic tilings characterized by two periods - horizontal and vertical - are considered. If there is no constraint on the periods couple, one gets what will be called a "shifted rectangular tiling", if periods are equal to the sides of the tile, one gets "rectangular tilings" and if only square tiles are allowed, one gets "square tilings" (see Figure 1) [2]. Using theses tilings, the collection of space-time diagrams of a given cellular automaton can be transformed into the collection of space-time diagrams of a new cellular automaton, states of which are the obtained colored tiles. This new automaton is said to be constructed by "shifted rectangular (or rectangular, or square) grouping".

Two standard ways, according to the existence of some injection or some surjection between sets of states, allow to compare automata rules: A is said to be a sub-automaton of B [3] (resp. a quotient-automaton of B) if it is isomorphic to the restriction of B to a subset of its set of states (resp. if it is isomorphic to a cellular automaton obtained from B by identification of some states).
Then two relations can be defined over the set of cellular automata which happen to be pre-orders: $A \leq_s B$ (resp. $A \leq_q B$) if "some grouping of A is a sub-automaton (resp. a quotient-automaton) of some grouping of B. These pre-orders (actually six) induce corresponding equivalence relations and orders on the classes.

Intuitively $A \leq_s B$ means, in some sense, that each global phenomenon of A can be isomorphically reproduced by means of B (see Figure 2) and $A \leq_q B$ that each global phenomenon of A can be reproduced by means of B in splitting states of A.

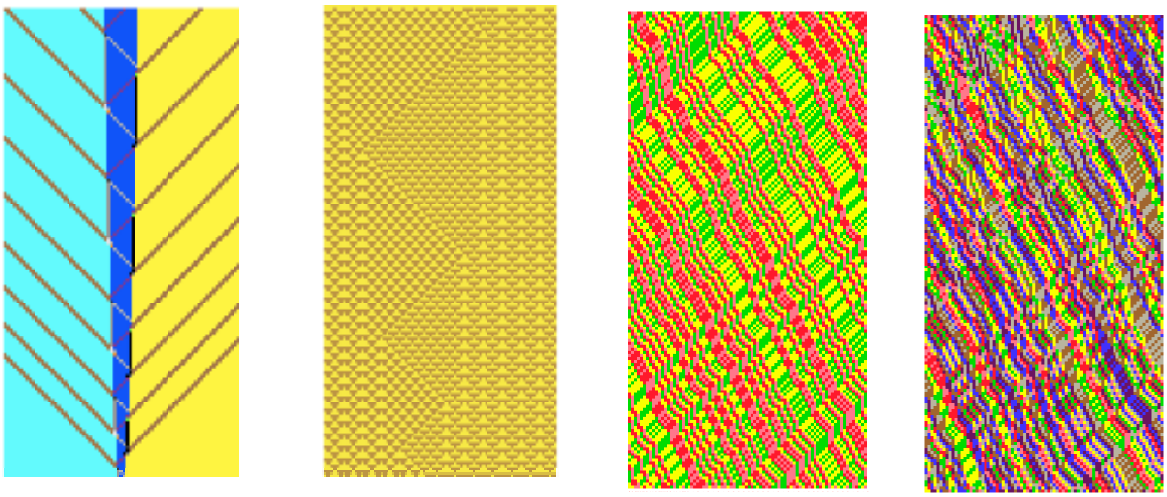


Figure 2: The two left diagrams show two automata A and B. A has 9 states and it is a sub-automaton of B square grouped by 4. The two right diagrams show two automata C and D. C has 2 states and it is a quotient-automaton of D: both states of C are split into 4 states to obtain D but D is not the product of C by another automaton.

Figure 3 represents significant elements of the sub-automaton order structures with square and rectangular groupings. The first important point is that there is a maximum in the case of the rectangular grouping and not in the case of the square one. The second important point is that the maximum contains the set of intrinsic cellular automata of literature [4, 5, 6], i.e. cellular automata which are able to simulate any cellular automaton. Let us observe that if intrinsic universality had already been considered, it is now well formalized in the present algebraic framework; this allows proofs of non-universality but also pertinent comparisons with other notions such as Turing-universality. We also observe that, in both cases, the orders are infinite in width and height, and that there are infinite increasing bounded chains. Understanding the existence of such chains is easy and interesting: if one wants to exhibit some global behavior depending on a parameter n, one needs, for large n, a great amount of states, but this is no more necessary at the upper bound because a new mechanism is introduced which allows to encode the parameter value in the initial configuration.

The same idea can be applied in case of Turing universality with two independent parameters: the number of heads and the ability of a head to make successive zigzags. That allows us to prove that the classes of cellular

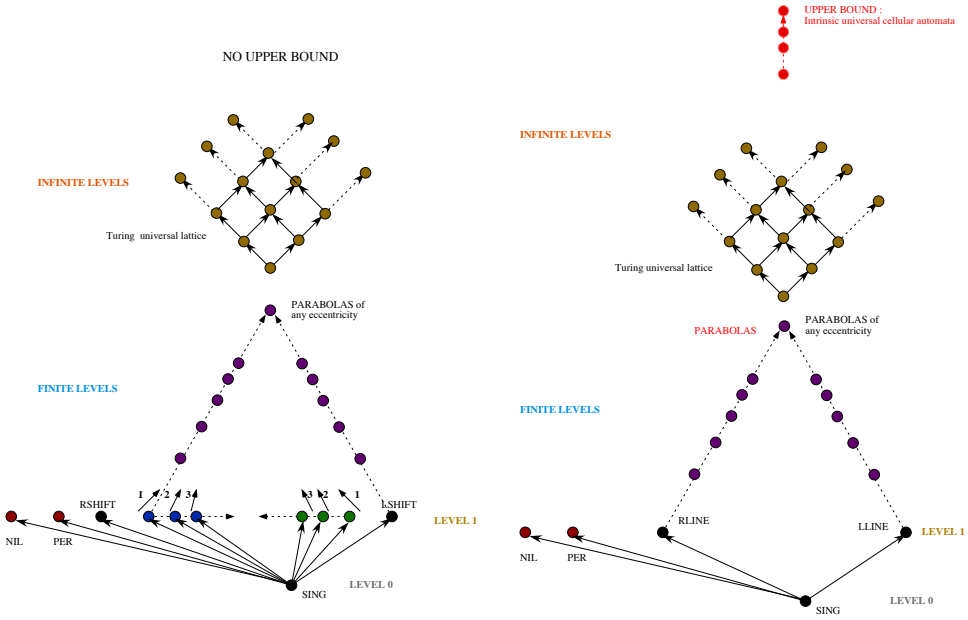


Figure 3: Sub-automaton orders with square and rectangular groupings.

automata simulating Turing universal machines have a structure of lattice (for all sub-automaton orders and orders mixing quotient and sub-automaton). Actually, algebraic hierarchies split Turing universality in an infinite number of classes (inside non trivial order structures) while intrinsic universality is represented in a single class. Moreover, in algebraic classifications corresponding to rectangular and shifted rectangular grouping, the maximum class (of the intrinsic universal cellular automata) is at infinite distance of every other class, especially of these Turing universal classes.

References

[1] Wolfram S: Theory and applications of cellular automata, World Scientific, Singapore, 1986.
[2] Ollinger N: Automates cellulaires : structures, PhD thesis, Ecole Normale Supérieure de Lyon, 2002.
[3] Mazoyer J, Rapaport I: Inducing an order on cellular automata by a grouping operation, in Proceedings of STACS'98, pp 128–227, Springer Verlag, 1998.
[4] Albert J, Čulik K: A simple universal cellular automaton and its one-way and totalistic version, Complex System, Vol 1, pp 1–16, 1987.
[5] Durand B, Róka Zs: The Game of Life : universality revisited, in Cellular Automata : a parallel model, M. Delorme and J. Mazoyers Eds, Kluwer, pp 51–76, 1999.
[6] Ollinger N: The intrinsic universality problem of one-dimensional cellular automata, in Proceedings of STACS'03, pp 376–381, Springer Verlag, 2003.

On stability of computations by cellular automata

Bruno Durand* and Andrei Romashchenko**

Key words: cellular automata, reliable computations, Toom's rule

Abstract. In this paper, we study stability of computations in the presence of random faults (noise). We focus on homogeneous models such as cellular automata. We present a new proof of stability of Toom's 2-dimensional automaton. The arguments are based on the methods from the famous "Positive Rates" paper by P. Gács. The advantage of our construction is that it explains precisely how errors spread in the computational array and how they are stabilized. Also we show that the same technique can be used to prove correctness of a 3-dimensional fault tolerant computational array.

1 Introduction

In most models of computation, programs are not stable towards faults: if a single step is corrupted, then the result of the computation is erroneous. A task of theoretical and practical interest is to implement reliable computations on faulty devices: the problem is to construct models which provide reliable computations even if some elementary step are faulty. Different approaches were proposed, mainly several variants of faulty circuits and cellular automata – see a survey in [1]. In the present paper we focus on computations based on faulty cellular automata.

The first step towards implementation of stable computations on faulty cellular automata, is the following basic problem: construct a cellular automaton that has at least 2 stable configurations. Here we say that a configuration is stable, if with high probability most cells of the array keep the initial state (corresponding to this configuration) for a long (or even infinite!) time despite random perturbation. Note that this problem is interesting for its significance in physics and thus many authors motivate this kind of work by study of phase transition instead of computation.

* Laboratoire d'Informatique Fondamentale de Marseille, CNRS, Université de Provence, 39, rue Joliot-Curie, F-13453 Marseille Cedex 13, France fax: +33 4 91 11 36 02 Bruno.Durand@lif.univ-mrs.fr

** Institute for Information Transmission Problems, RAS, 19 Bolshoy Karetny, 101447, Moscow, Russia, fax: +7 095 209 0579 anromash@mccme.ru The work was supported in part by Russian Science Support Foundation and RFBR grants 03-01-00475, 358.2003.1.

The first result in this area belongs to Andrei Toom. In [2] he presented a very simple two-dimensional stable cellular automaton. Despite apparent simplicity of Toom's cellular automaton (which has only two states) the original proof of its stability was very technical. Several other proofs of Toom's result were then given [3–6], but they all are quite non-trivial. One of the goals of this paper is to provide a simple proof that explains precisely what happens with faults and their consequences in the array of Toom's automata.

Toom's construction was used later in [4] in order to simulate 1-dimensional array of cellular automata by a 3-dimensional array of *stable* automata in real time ("in real time" means that one step of computation of the 1-dimensional *simulated* cellular automata is simulated by one step of the 3-dimensional *simulating* cellular automata). Later, Peter Gács improved this result and implemented fault tolerant computations on 2-dimensional [7] and 1-dimensional [8] arrays of faulty cellular automata. To achieve these results Gács developed interesting technique, which might be useful in other situations. However, the construction by Gács is extremely difficult and long, and is really understood by a very small number of people. This is a serious obstacle to further progress in this field: the problem is fundamental and the construction should be used for many related problems but it cannot be as it is. The author himself notes in [1, 8] that one of the most important open problems in this area is to simplify the proof of his results, i.e, to find an understandable construction of self-correcting computations based on faulty cellular automata.

In this paper we deal with the first (and the easiest) part of this construction. Using ideas from [8], we present a new simplified proof of stability of the Toom 2-dimensional automaton, and also correctness of the three-dimensional reliable cellular array proposed in [4]. Our main intention is to split the proof into a few independent steps and provide a simple proof for each of them. Our goal is to explain what happens with faults and their consequences: how long 'islands' of disturbed cells can survive and how they shrink.

Our paper is organized as follows: we first discuss in all detail the most simple case: Toom's rule for an infinite 2-dimensional cellular automaton. In Section 3 we explain what happens on finite arrays of faulty automata (torus). In Section 4 we use our construction to explain the embedding of computations into 3-dimensional fault tolerant cellular automaton.

2 Toom's rule on the infinite plane.

We start with a 2-dimensional cellular automaton. Let us consider a 2-dimensional array of cells, the infinite plane $\mathbb{Z} \times \mathbb{Z}$. There is a fixed (finite) automaton in each cell of the space. This automaton has two states: 'alive' and

'dead'. We identify the state of the cell with the state of the automaton it contains. All cells update synchronously as follows: the state of any cell at time $t+1$ is defined as majority of states of itself, its North and its East neighbors at time t, see Fig. 1. The majority is well defined since there are 2 state possibilities for

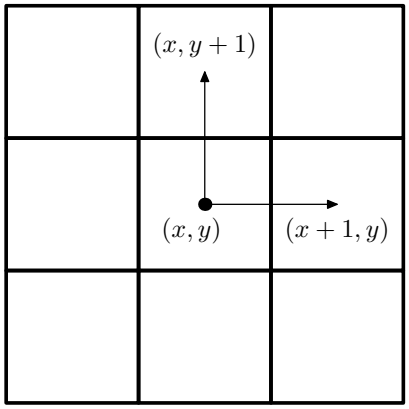


Fig. 1. Toom's rule

each of 3 cells. We call this automaton rule *Toom's rule* since it is inspired by Toom's work [2].

Note that if we have a finite 'island' of n alive points on the infinite plane of dead points, Toom's rule will shrink this island in time at most n. To see this, we circumscribe a triangle that we call 'Toom's triangle' around the island, see Fig. 2. This (virtual) triangle has the property to contain all alive cells, be isocele and minimal.

In course of time, Tooms's rule transforms the island of alive cells. At each step we circumscribe around the island new Toom's triangle. How does the island vary on time, and what happens with the Toom triangle? Obviously, the vertical leg of the triangle may not move to the left, and the horizontal leg triangle may not move downwards. But the hypotenuse moves to the south-west (each step it must shift at least by 1). Thus the triangle shrinks to nothing in a finite number of steps. Moreover, the number of steps required to eliminate the island is not greater than the initial size of Toom's triangle, which is not greater than the number of alive cells in the island.

The arguments above are quite trivial. But we stress that the idea of shrinking of Toom's triangle is the main issue of our arguments. In the sequel we shall

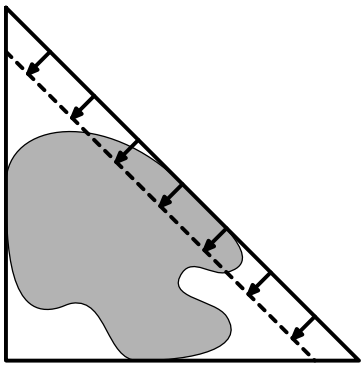


Fig. 2. Toom's triangle

generalize these arguments for a non-trivial case, when the evolution of Toom's triangle becomes more complicated due to random faults.

Now, let us introduce faults. We model them by a random transformation. At any time step, any dead cell might (randomly) come to live. We call such events 'miracles'. Our probabilistic model is as follows: we assume that miracles at different cells and times are independent and that the probability to get a miracle at point a at time t is some (small) $\epsilon > 0$. Beware that miracles are not symmetric: they consist of spontaneous life but not of spontaneous death. This non-symmetry strengthen our result because in order to prove that a dead configuration is stable, we cannot use in our model the spontaneous death of alive cells.

Initially, i.e., at time $t = 0$, all cells are dead.

What is to be proved looks very simple: we want to prove that if ϵ is small enough, then any cell at any time is dead with high probability. Our ideas come from reading of [1, 8, 9].

An Informal plan of our proof is:

1. Separate all miracles into 'islands' of different levels. Islands are sets of miracles in space $\times$ time. An n-level island of miracles should be of size about $q_n \times q_n$ in space and about q_n in time (we specify the values q_n later); an n-level island should contain enough miracles (at least 2^n). Islands of the same rank should be separated from each other in space and time, so than they cannot interfere with each other. The first step of our proof is the fact that such ranking of the set of miracles is possible with probability 1.

2. Show that an n-level island of miracles disappears without consequences in time $\mathcal{O}(q_n)$, without interacting with other islands of the same rank. If there were no lower rank errors, the proof would be trivial: we could circumscribe Toom's triangle around a given island and note that it shrinks as when there are no miracles (the vertical and the horizontal legs may not move left and down respectively, the hypotenuse moves South-West at speed at least 1). In the presence of lower rank errors, the vertical and the horizontal sides *may* move left and down respectively. What we prove is that this movement is sufficiently slow. The speed of the hypotenuse may also be a bit less than 1, but we prove that it remains high enough. This is how we obtain that the triangle vanishes in time $\mathcal{O}(q_n)$.

2.1 Notations and definitions

We denote by $Space$ the set $\mathbb{Z} \times \mathbb{Z}$, and by $Time$ the set $\mathbb{Z}_+$.

Define *a random perturbation* as $\xi(a)$, $a \in Space \times Time$. Let $Prob[\xi(a) = 1] = \epsilon$ and $Prob[\xi(a) = 0] = 1 - \epsilon$. We assume that $\xi(a)$ and $\xi(b)$ are independent for $a \neq b$. Intuitively, when $\xi(a) = 1$ there is "noise" and a miracle may occur in the cell and time addressed by a.

Now we are ready to define Toom's cellular automaton in a faulty situation. For any point $(x, y, t) \in Space \times Time$ the state of a cell $s(x, y, t)$ is 1 (alive) or 0 (dead).

1. $s(x, y, 0) = 0$ for any x, y, i.e., initially all cells are dead;
2. for any $t > 0$ the following transition rule is used: let $a = (x, y, t)$, $a_0 = (x, y, t-1)$, $a_e = (x+1, y, t-1)$, $a_n = (x, y+1, t-1)$; then

$$s(a) = \max\{s'(a), \xi(x, y, t)\},$$

where $s'(a) = majority\{s(a_0), s(a_e), s(a_n)\}$. In other words, at a noisy point (i.e., if $\xi(a) = 1$) a cell is made alive, independently of its neighborhood. In a normal point (i.e., in a point a such that $\xi(a) = 0$) Toom's rule is applied.

A point $a \in Space \times Time$ is called *a miracle*, if Toom's rule is not applied at this point, i.e., $s(a) \neq s'(a)$. Beware that if the point is noisy and becomes alive, it is not sure that it is a miracle. It could be alive due to Toom's rule, and in this case, even if it is noisy, it is not called a miracle.

We denote by $M \subset Space \times Time$ the set of all miracles. The main idea of the proof is to split M into 'islands' of miracles of different size.

We use l_∞-norm to measure the distance between points in the space-time. For two points $a = (x, y, t)$ and $a' = (x', y', t')$ in $Space \times Time$ we denote $\mathrm{dist}(a, b) = \max\{|x - x'|, |y - y'|, |t - t'|\}$. For $A, B \subset Space \times Time$ denote

$$\mathrm{dist}(A, B) = \min_{a \in A, b \in B} \mathrm{dist}(a, b).$$

The size (diameter) of a set $S \subset Space \times Time$ is defined as $\max\limits_{a,b \in S} \mathrm{dist}(a, b)$

Let us fix now the constants q_n: we choose $C = 10000, q_0 = 3, q_n = Cn^2 q_{n-1}$. The exact values of this parameters are not very important. However, it is essential that $\sum(q_n/q_{n+1}) < \infty$.

Definition 1 (0-level islands and semi-islands).
A) We call a 0-level semi-island *every singleton containing a miracle.*
B) We call a 0-level island *any* 0*-level semi-island* S *such that* $\mathrm{dist}(S, M \setminus S) > q_1/5$.
C) The union of all 0*-level islands is denoted* M_0.

In general, M_0 is a proper subset of M.

Suppose that for all $k < n$ we have defined k-level semi-islands, k-level islands, and M_k. Then we define by induction semi-islands and islands of level n and the set M_n:

Definition 2 (n-level semi-islands and islands).
A) A nonempty set $S \subset M \setminus M_{n-1}$ *is called an* n-level semi-island *if*
(1) S *is of size at most* q_n*;*
(2) S *contains at least two disjoint* $(n-1)$*-level semi-islands.*

B) An n*-level semi-island* S *is an* n-level island *if*
(3) $\mathrm{dist}(S, M \setminus (S \cup M_{n-1})) > q_{n+1}/5$
C) We denote by M_n *the union of* M_{n-1} *and all* n*-level islands.*

Note that an n-level island is defined as an n-level semi-island that is isolated from other semi-island of level n and higher. Note also that any n-level semi-island contains at least 2^n miracles.

Definition 3 (birth time). *Let* S *be an* n*-level island. The* birth time *of* S *is* $\min\{t : \exists x, y\ (x, y, t) \in S\}$.

2.2 The structure of M

We say that an n-level semi-island $S \subset M$ is *minimal*, if any proper subset of S is *not* a n-level semi-island. Obviously, any n-level semi-island contains

a subset that is a minimal n-level semi-island. Further we prove a statement concerning minimal semi-islands. In the following proposition we let $M_{-1} = \emptyset$ to simplify the notation.

Proposition 1. *For any m, for all $a \in M \setminus M_{m-1}$, the miracle a belongs to a* minimal *m-level semi-island S, which contains exactly 2^m miracles and is of size at most $q_m/3$.*

Proof: From Definition 2, condition (2), it follows that any n-level semi-island contains at least 2^n points. Thus, if some n-level semi-island contains exactly 2^n points, it is minimal.

We prove the proposition by induction. The inductive base $m = 0$ holds, as a 0-level semi-island is just a singleton, and $q_0 = 3$.

Assume that the statement is true for all $n < m$; let $a \in M \setminus M_{m-1}$. We must prove that a belongs to some minimal m-level semi-island. From the inductive hypothesis, there exists a minimal $(m-1)$-level semi-island S_1 such that $a \in S_1$; the set S_1 consists of 2^{m-1} points and is of size at most $q_{m-1}/3$.

Claim: *There exists a point $b \in M \setminus M_{m-2}$ such that $q_{m-1}/3 < \mathrm{dist}(b, S_1) < q_m/4$.*

Assume the contrary is true, i.e., for any point $b \in M \setminus M_{m-2}$ if $\mathrm{dist}(b, S_1) \leq q_m/4$, then we have $\mathrm{dist}(b, S_1) \leq q_{m-1}/3$. In this case we can extend S_1 and get an $(m-1)$-level island, which implies a contradiction. More precisely, define

$$\tilde{S}_1 = \{b \in M \setminus M_{m-2} \ : \ \mathrm{dist}(b, S_1) \leq q_{m-1}/3\}.$$

The set $\tilde{S}_1$ is of size less than q_{m-1}. Hence, condition (1) from Definition 2 holds for $\tilde{S}_1$ with $n = m - 1$.

Further, $S_1 \subset \tilde{S}_1$, so (2) of Definition 2 is also true for $\tilde{S}_1$.

From our assumption we get that the distance between $\tilde{S}_1$ and $M \setminus (M_{m-2} \cup \tilde{S}_1)$ is at least $(q_m/4 - q_{m-1}/3)$, which is greater than $q_m/5$. Thus, condition (3) from Definition 2 is true for $\tilde{S}_1$. Hence, the set $\tilde{S}_1$ is a $(m-1)$-level island, $a \in M_{m-1}$, and we get a contradiction. The claim is proved.

From this Claim we get a cell $b \in M \setminus M_{m-2}$, which is at distance at least $q_{m-1}/3$ and at most $q_m/4$ from S_1. By inductive hypothesis, there exists a minimal $(m-1)$-level semi-island $S_2 \ni b$; the set S_2 consists of 2^{m-1} points and is of size at most $q_{m-1}/3$. As the distance between S_1 and b is at least $q_{m-1}/3$, the sets S_1 and S_2 are disjoint.

Set $S = S_1 \cup S_2$. By definition, S contains two disjoint $(m-1)$-level semi-islands and consists of $2^{m-1} + 2^{m-1} = 2^m$ miracles. Also, S is of size at most $(2/3 q_{m-1} + q_m/4) < q_m/3$. Thus, S is an m-level semi-island of size less than $q_m/3$. It is minimal, since S consists of 2^m points. □

Definition 4. *Let u be a point in $Space \times Time$. Denote by $\epsilon_n(u)$ probability of the event that there exists at least one n-level semi-island S such that* $\mathrm{dist}(u, S) \leq q_{n+1}$.

Proposition 2. *There exists a $\gamma > 1$ such that for a small enough ϵ (from the definition of perturbation ξ) for any $u \in Space \times Time$ we have $\epsilon_n(u) \leq (\gamma\epsilon)^{2^n}$.*

Proof: Let us fix a point u in the space-time. We are interested in the probability of the event that there exists an n-level semi-island S such that $\mathrm{dist}(u, S) \leq q_{n+1}$. Note that such a semi-island S should be inside of the $2q_{n+1}$-neighborhood of u.

As any n-level semi-island contains a *minimal* n-level semi-island, it is enough to bound the probability of the following event: in the $2q_{n+1}$-neighborhood of u there exists a minimal n-level semi-islands.

We are going to count the total number of all minimal n-level semi-islands in the given neighborhood of u. Each n-level semi-island contains at least 2^n miracles. Hence, if the number of such semi-islands is L_n, then $\epsilon_n \leq L_n \cdot (\epsilon)^{2^n}$.

To make the arguments more clear, we shall count not the number of semi-islands but length of their descriptions. More precisely, we show that a minimal semi-island S in the given area can be uniquely identified (while the point u is known) by some description of length $\mathcal{O}(2^n)$. We suppose that any description is just a string of 0's and 1's. As there exist at most 2^l different descriptions of length l, there are at most $2^{\mathcal{O}(2^n)}$ different semi-island that have such a description. So, our bound for the length of descriptions implies

$$\epsilon_n \leq 2^{\mathcal{O}(2^n)} \cdot \epsilon^{2^n} < (\gamma\epsilon)^{2^n}$$

for some constant γ, which does not depend on ϵ and n.

We should define more formally what a complexity of a semi-island is. For the reader acquainted with the notion of Kolmogorov complexity [10] we could say that this is plain Kolmogorov complexity. But actually a more restrictive and simple definition is enough for our proof:

Definition 5. *Denote by $\mathcal{I}_n$ the set of all minimal n-level semi-islands in $(2q_{n+1})$-neighborhood of u. We shall say that there exist discriptions of length l_n for all n-level semi-islands in the given area, if there exists a surjective mapping*

$$\mathcal{F} : \{0,1\}^{l_n} \to \mathcal{I}_n.$$

Intuitively, $\mathcal{F}$ is a rule, which maps a description (a string of one's and zero's of length l_n) to the corresponding semi-island.

Lemma 1. *Let u be a point in the space-time. Then for all minimal semi-islands in $2q_{n+1}$-neighborhood of u there exist descriptions of length $l_n = D(\sum\limits_{k\le n} \frac{\log k}{2^k})2^n$ for some constant D.*

In the proof we shall explain the description rule in intuitive terms. We believe it should be quite clear how to define the corresponding mapping $\mathcal{F}$ formally.

We prove this lemma by induction. The base is trivial: a 0-level semi-island is just a singleton. There are $\mathcal{O}(q_1^3)$ points in the $(2q_1)$-neighborhood of u, and we can provide descriptions of length $(3\log q_1 + \mathcal{O}(1))$.

Let us prove the inductive step. Let S be a minimal n-level semi-island in the q_{n+1}-neighborhood of u. From the definition of a semi-island it follows that S is a union of two disjoint semi-island of level $(n-1)$. Denote this islands S' and S''. It is not hard to see that we can choose two points u', u'' in the space-time such that

1. $\mathrm{dist}(u', S') \le q_n$,
2. $\mathrm{dist}(u'', S'') \le q_n$,
3. in each coordinate, the difference between u and u', and between u and u'' is a multiple of q_n.

Note that there are only $\mathcal{O}((q_{n+1}/q_n)^3) = \mathcal{O}((n+1)^6)$ possible positions for u' and u''.

To identify S given u, it is enough to identify u' and u'', and then identify S' given u' and S'' given u''. In a word, our description of S consists of 4 parts: description of u' given u, description of u'' given u, and descriptions of S' and S'' given u' and u'' respectively.

To describe each of two points u', u'', it is enough $\log((n+1)^6) + \mathcal{O}(1)$ bits of information. Further, by the inductive assumption, to identify S' given u' or S'' given u'' we need strings of length

$$(\sum_{k\le n-1} D\frac{\log k}{2^k})2^{n-1}.$$

In the whole, the description of S requires

$$\mathcal{O}(\log((n+1)^6)) + 2D(\sum_{k\le n-1} \frac{\log k}{2^k})2^{n-1} \le D(\sum_{k\le n} \frac{\log k}{2^k})2^n$$

digits (if D is taken large enough). The lemma is proven.

To prove the proposition it remains to note that the series $\sum \frac{\log k}{2^k}$ converges. Thus, any minimal n-level semi-island S in q_{n+1}-neighborhood of u has a description of length $O(2^n)$, and we are done.

□

Remark 1. The fact that $Space \times Time$ has dimension 3 is not important for the proof above. The same argument would work for any finite dimension.

Remark 2. Proposition 2 implies that $\sum\limits_{i<\infty} \epsilon_i = \mathcal{O}(\epsilon)$ if ϵ is small enough.

Corollary 1. *If ϵ is small enough, with probability 1 every miracle belongs to some n-level island.*

Proof: Let $u \in M$. By Proposition 1 there are two possibilities: either (i) $u \in M_n$ for some $n \geq 0$, i.e., u belongs to some n-level island, or (ii) u belongs to an n-level semi-island for every $n \geq 0$. From Proposition 2 it follows that (ii) holds with probability 0, since $\epsilon_n \to 0$ as $n \to \infty$. □

We shall say that the set of miracles M is *standard*, if each miracle belongs to some n-level island.

2.3 Evolution of an n-level island

Every alive cell is alive because it is a consequence of some miracles. Let us define this more formally.

Definition 6. *An explanation path from an alive point $a \in Space \times Time$ to an n-level island S is a sequence $a_0, a_1, \ldots, a_m$ of alive points in $Space \times Time$ such that*

1. $a_0 = a,\ a_m \in S$
2. *for any $k < m$ if a_k has coordinates (x, y, t) then a_{k+1} is one of three points $(x, y, t-1)$, $(x+1, y, t-1)$, $(x, y+1, t-1)$.*

Note that some a_i in the sequence above might belong to other islands but S.

If there exists an explanation path from an alive point a to an n-level island S, we say that a is a consequence of S.

Obviously, for a *standard* set of miracles M any alive point is a consequence of some island of miracles (or, maybe, of many islands).

Remark 3. Let a sequence of points $a_0, a_1, \ldots, a_m$ be an explanation path, and $a_i = (x_i, y_i, t_i)$ for $i \leq m$. Then $t_i - t_j \geq \max\{x_j - x_i, y_j - y_i\}$ for any $i < j$.

There may be more than one explanation path from a point a to an island S. Further we select explanation paths with some special properties.

Definition 7 (Space-greedy explanation path). *Let $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to some island S. This path is called* space-greedy *if the following conditions hold for any $i < m$:*

– $a_i \notin S$.
– *Let $a_i = (x, y, t)$. This point is a consequence of S but not its member. Hence, at least two of points $(x, y, t-1)$, $(x+1, y, t-1)$, $(x, y, t+1)$ are alive, and, moreover, at least one of them is a consequence of S. For a space-greedy path, if $(x+1, y, t-1)$ or $(x, y+1, t-1)$ is a consequence of S, then a_{i+1} must be one of these two points. Only if none these points is a consequence of S, $a_{i+1} = (x, y, t-1)$.*

And two similar definitions:

Definition 8 (South-most explanation path). *Let a sequence of points $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to an island S ($a_m \in S$). This path is called* south-most *if for any $i < m$ the following conditions hold:*

– $a_i \notin S$.
– *If one of points $(x, y, t-1)$, $(x+1, y, t-1)$ is alive, then a_{i+1} must be one of these two points. Otherwise $a_{i+1} = (x, y+1, t-1)$,*

Definition 9 (West-most explanation path). *Let a sequence $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to some island S ($a_m \in S$). This path is called* west-most *if for any $i < m$ the following conditions hold:*

– $a_i \notin S$,
– *If one of points $(x, y, t-1)$, $(x, y+1, t)$ is alive, then a_{i+1} is one of these two points. Only otherwise $a_{i+1} = (x+1, y, t-1)$.*

Note that if u is a consequence of some n-level island S, then there exist a space-greedy, a west-most and a south-most explanation paths from u to S.

Definition 10. *A point $a \in Space \times Time$ is called a* proper consequence *of an n-level island S, if the following conditions hold*

1. *a is not a consequence of any island of level higher than n.*
2. *if a is a consequence of another n-level island S', then the birth time of S is not greater than the birth time of S'.*

Lemma 2. *Assume the set of miracles M is standard, $a \in M$, and a is a proper consequence of an n-level island S. Assume a is also a consequence of another n-level island S'. Then for any consequence b of S' we have $\mathrm{dist}(b, S) > 10q_n$.*

Proof of lemma: Let $b = (x, y, t)$ be a consequence of S', and assume $\mathrm{dist}(b, S) \le 10q_n$. Denote by t_0 the birth time of S. Then

$$t - t_0 \le 11q_n. \tag{1}$$

The point b is a consequence of S', so there exists an explanation path $b_0, b_1, b_2, \ldots, b_m$ from b to S' ($b = b_0$ and b_m is point in S'). Let $b_m = (x_m, y_m, t_m)$. As $\mathrm{dist}(b, S) \leq 10q_n$ and $\mathrm{dist}(S, S') > q_{n+1}/5$ (the distance between any two n-level islands must be large), we get $\mathrm{dist}(b, b_m) > q_{n+1}/5 - 10q_n$. Hence, by Remark 3,

$$t - t_m > q_{n+1}/5 - 10q_n. \tag{2}$$

From inequalities (1) and (2) we have $t_m < t_0$, so the birth time of S' is less then the birth time of S. Hence, by Definition 10, item (2), a is not a proper consequence of S. Thus, we get a contradiction. □

Theorem 1. *Suppose that the set of miracles M is standard. Let a be a proper consequence of an n-level island S. Then $\mathrm{dist}(a, S) < 8q_n$. Consequently, for any point $b \in S$ the distance between a and b is less than $9q_n$.*

Proof of theorem: We prove the theorem by induction. Inductive basis $n = 0$ is trivial. Let us deal with inductive step. We should prove that an explanation path from S to its proper consequence a cannot be too long. It is enough to prove that if a is a proper consequence of S and $\mathrm{dist}(a, S) < 10q_n$, then we have $\mathrm{dist}(a, S) < 8q_n$.

Lemma 3. *Let a be a consequence of S. Let $P = (a_0, a_1, \ldots, a_k)$ be a space-greedy explanation path from $a = a_0$ to some $a_k = a' \in S$. Let $a = (x, y, t)$ and $a_k = (x', y', t')$. Denote $T = t - t'$ and $L = (x' - x) + (y' - y)$. Suppose $T > 5q_{n-1}$. Then $L \geq 9T/10$.*

Informally, this lemma says that the hypotenuse of Toom's triangle around S moves to the south-west with average speed greater than $9/10$.

Proof of lemma: Obviously, every next point in the path P has the time coordinate less by 1, so $T = k$. Further, any a_i either has the same space coordinates as a_{i+1}, or one of its space coordinates is shifted by 1. To get the Lemma we should prove that there are at most 10% of such points a_i in the path P, that a_i and a_{i+1} have the same space coordinates.

Let $a_i = (x_i, y_i, t_i)$ and $a_{i+1} = (x_i, y_i, t_i - 1)$ have the same space coordinates. The points $(x_i + 1, y_i, t_i - 1)$ and $(x_i, y_i + 1, t_i - 1)$ are not consequences of S, because P is space-greedy. As a_i is alive, two variants are possible:

1. a_i is a miracle and, hence, it belongs to some island S';
2. at least one of points $(x_i + 1, y_i, t_i - 1)$, $(x_i, y_i + 1, t_i - 1)$ is alive and, hence, is a proper consequence of some island $S' \neq S$.

In both cases we say that S' *supports* the path P at the point a_i.

As a is a proper consequence of S, from Lemma 2 it follows that the level of S' is less than n.

For each number $i < k$ such that the point a_i has the same space coordinate as a_{i+1}, we fix an island $S' = S'(a_i)$ as above. We should answer two questions:

- How many points a_i of the path P can support one k-level island S'?
- How many k-level islands can support the path P at different points?

Answer to the first question is simple: by the inductive hypothesis of Theorem 1, it is less than $9q_{k-1}$ (all proper consequences of a k-level islands S' are at the distance at most $9q_k$ from any point in S').

To answer the second question some calculations are required. Assume that S' and S'' are two k-level islands that support the path P at points a_i and a_j respectively. By the inductive hypothesis, $\mathrm{dist}(a_i, S') < 8q_k$ and $\mathrm{dist}(a_j, S'') < 8q_k$. The same time, the distance between any two k-level islands is lower-bounded: $\mathrm{dist}(S', S'') > q_{k+1}/5$. Hence, the distance between a_i and a_j is quite large. At least, $\mathrm{dist}(a_i, a_j) > q_{k+1}/10$ (a very rough bound). From Remark 3 we get $|t_i - t_j| = |i - j| > q_{k+1}/10$. Upperbound the number of all points a_i on the path that are supported by k-level islands: it is less than $\lceil T/(q_{k+1}/10)\rceil \cdot 9q_k$. Sum up this value for all $k < n$:

$$\sum_{k<n} \lceil 10T/q_{k+1}\rceil \cdot 9q_k \le (1 + 1/5) \cdot 90 \sum_{k<n} Tq_k/q_k < 120T \sum_{k<\infty} (q_k/q_{k+1})$$

(the first inequality follows from the condition $T > 5q_{k+1}$ for all $k < n$). Further, in the definition of q_n the constant C was chosen so that

$$\sum_{k=0}^{\infty} (q_k/q_{k+1}) \le 1/2000.$$

Lemma follows from this bound immediately.

Lemma 4. *Let a be a consequence of S, and $P = (a_0, a_1, \ldots, a_k)$ be a southmost explanation path from $a = a_0$ to some $a' = a_k \in S$. Let $a = (x, y, t)$ and $a_k = (x', y', t')$, and denote $T = t - t'$, $L' = (y' - y)$. Suppose that $T > 5q_n$. Then $L' \le T/10$.*

Informally, this lemma says that the horizontal leg of Toom's triangle around S moves to the south with average speed less than $1/10$.

Proof of lemma: very similar to the proof of Lemma 3. We are interested in points a_i on the path P such that a_i and a_{i+1} have different Y-coordinates. We should prove that there are at most 10% of such points in P.

Let $a_i = (x_i, y_i, t_i)$ and $a_{i+1} = (x_i, y_i + 1, t_i - 1)$. Then the points $(x_i + 1, y_i, t_i - 1)$ and $(x_i, y_i, t_i - 1)$ are not consequences of S, because P is south-most. Hence, either a_i is a miracle (and belongs to some island S'), or at least one of point $(x_i + 1, y_i, t_i - 1)$, $(x_i, y_i, t_i - 1)$ must be alive and be a proper consequence of some islands $S' \neq S$. In these cases we say that S' *supports* the path P at the point a_i.

How many points a_i can support one k-level island S'? By the inductive hypothesis of Theorem 1, the number of such points is less than $9q_k$.

How many k-level islands can support the path P at different points? Assume that S' and S'' are two k-level islands that support P at points a_i and a_j respectively. By inductive hypothesis, the distances $\mathrm{dist}(a_i, S')$ and $\mathrm{dist}(a_j, S'')$ are both less than $8q_k$. The same time $\mathrm{dist}(S', S'') > q_{k+1}/5$. Hence, $|i - j| > q_{k+1}/10$.

The number of all points a_i supported by k-level islands is at most $\lceil T/(q_{k+1}/10)\rceil \cdot 9q_k$. Sum up this value for all $k < n$:

$$\sum_{k<n} \lceil 10T/q_{k+1}\rceil \cdot 9q_k \leq (1 + 1/5) \cdot 90T \sum_{k<n} q_k/q_{k+1} \leq T/10.$$

Lemma is proved.

Also a statement symmetrical to Lemma 4 holds:

Lemma 5. *Let a be a consequence of S, and $P = (a_0 = a, a_1, \ldots, a_k)$ be a west-most explanation path from a to S, where $a = (x, y, t)$ and $a_k = (x', y', t')$. Let $T = t' - t$ and $L'' = (x' - x)$. If $T > 5q_n$ then $L'' \leq T/10$.*

Informally, this lemma says that the vertical leg of Toom's triangle around S moves to the west with average speed less than $1/10$.

Now we can prove the theorem. The idea is simple: for each time t draw Toom's triangle around the set of all consequences of the island S. The lemmas above imply that as the time coordinate increase, the hypotenuse of this triangle goes to the south-west with average speed at least $9/10$, and the vertical and horizontal legs go to the west and south respectively with small average speed (at most $1/10$). Hence, the triangle must shrink in time $\mathcal{O}(q_n)$.

The same idea can be expressed in other terms. Let some point a is a consequence of an error island S. Then there are south-most, west-most and space-greedy explanation paths from a to S. On one hand, these three paths diverge when we go to the past; on the other hand, the ends of all three paths must be inside S, see Fig. 3. Hence, the distance in time and space between a and S

cannot be to large, so these explanation paths have no possibility to diverge too far.

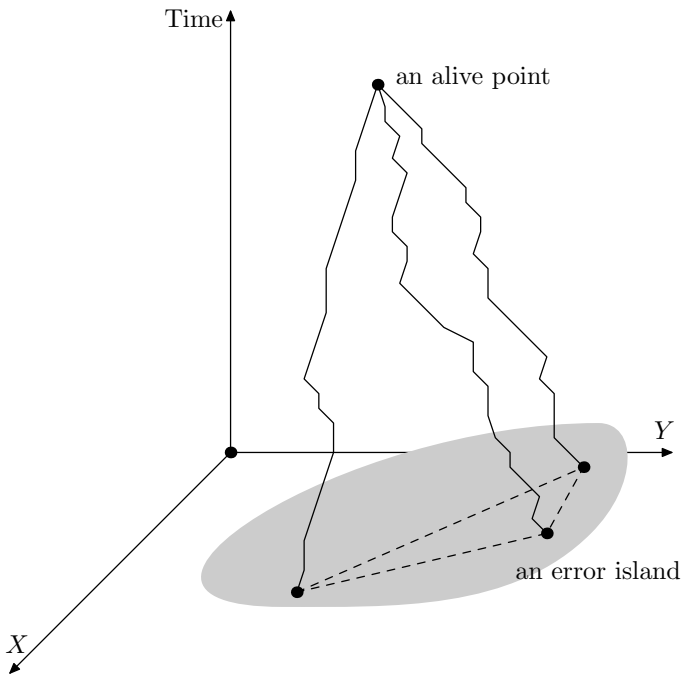


Fig. 3. Three explanation paths from an alive point to an error island

Let us do the precise computations. Denote by t_{first} and t_{last} the minimal and the maximal times of miracles in S; denote by $x_{\min}$, $x_{\max}$, $y_{\min}$, $y_{\max}$ the minimal and the maximal space coordinates of miracles in S respectively.

Let $a = (x, y, t)$ be a consequence of S. First, we prove that $t - t_{last} < 8q_n$. If $t - t_{last} \leq 5q_n$, there is nothing to prove. Assume that contrary and employ the lemmas above. From Lemma 3 it follows that for a space-greedy explanation path from a to $a' = (x', y', t') \in S$ the difference $(x' + y') - (x + y)$ is quite large comparative to $(t - t')$. More exactly, we have $(x_{\max} - x) + (y_{max} - y) > 9/10(t - t_{last})$. Further, from Lemma 4 the south-most explanation path goes from a to $a'' = (x'', y'', t'') \in S$ such that $(y'' - y)$ is quite small comparative to $(t - t'')$. More precisely, $(y_{min} - y) < 1/10(t - t_{first})$. Similarly, from Lemma 5 we get $(x_{min} - x) < 1/10(t - t_{first})$. It is not hard to check that these three inequality above imply $t - t_{last} < 5q_n$ (a very rough bound).

Thus, we have proved that for any proper consequence $a = (x, y, t)$ of S the bound $t - t_{last} < 5q_n$ holds. Let us fix any explanation path from a to S. Denote by $a' = (x', y,', t') \in S$ the last point in this explanation path. Then

$$t - t' \leq (t - t_{last}) + (t_{last} - t_{first}) \ll 8q_n.$$

From Remark 3, $x' - x \leq t - t'$ and $y' - y \leq t - t'$. Hence, $\mathrm{dist}(a, S) < 8q_n$. □

Corollary 2. *Any point $a \in Space \times Time$ is dead with large probability $\hat{\epsilon}$ ($\hat{\epsilon} \to 1$ as $\epsilon \to 0$).*

Proof of corollary: First of all, we may assume that the set of all miracles M is standard (Corollary 1).

Let a be an alive point. As M is standard, from Lemma 2 it follows that a is a proper consequence of some island S. From Theorem 1 it follows that the point a is in the $8q_n$-neighborhood of S.

From Proposition 2 we get that probability of the event 'a is on the distance less then $8q_n$ from some n-level island' is less than $\epsilon_n \leq (\gamma\epsilon)^{2^n}$. Hence, a is alive with probability less than the sum $\sum_{i<\infty} \epsilon_i = \mathcal{O}(\epsilon)$, see Remark 2.

The Corollary above implies a more general statement. Until now, we allowed only 'one-way' noise: a dead cell could become spontaneously alive, but not visa-versa. Let us consider a more traditional model. Let we have a 2-D cellular automaton, each cell has two states. At each cell normally Toom's rule is applied, but with small probabilities a cell can randomly change its state (from 'alive' to 'dead' or from 'dead' to 'alive'). Such a probabilistic automaton is usually called *a small perturbation of Toom's rule*. Again, we suppose that random transitions at different cells and at different moments of time are independent. Combining the result above with the standard arguments [11], we get that any such automaton has at least two different invariant measures: 'most cells are alive' and 'most cells are dead'.

3 Toom's rule on torus.

In this section we discuss the behavior of Toom's automaton on a finite space-time. Let p be a positive integer, and $Space$ be the torus $\mathbb{Z}_p \times \mathbb{Z}_p$, where $\mathbb{Z}_p$ is the set of integers modulo p. The time scale will be also finite: $Time = \{0, \ldots, T\}$. The definition of Toom's faulty automaton from Section 2 can be obviously used for this finite variant of $Space \times Time$. Moreover, as any small area on the torus is equivalent to an area on the plane, the behavior of Toom's automaton on a torus and on the plane is quite similar. Further we explain how the arguments from Section 2 can be applied in the new situation.

3.1 Notation and definitions

Define random perturbation $\xi(a)$ and the evolution of the automaton $s(x, y, t)$ word for word as in Section 2. The only difference is that now x, y run over $\mathbb{Z}_p$ and $t \in \{0, \ldots, T\}$.

For $u, u' \in \mathbb{Z}_p$ denote

$$|u - u'|_p = \min_{k \in \mathbb{Z}} |u - u' + kp|.$$

Define the distance between points in the space-time as follows: for $a = (x, y, t)$ and $a' = (x', y', t')$ in $Space \times Time$

$$\mathrm{dist}(a, b) = \max\{|x - x'|_p, |y - y'|_p, |t - t'|\}.$$

As usually, for $A, B \subset Space \times Time$ denote

$$\mathrm{dist}(A, B) = \min_{a \in A, b \in B} \mathrm{dist}(a, b),$$

and call *size of a set* $S \subset Space \times Time$ its diameter defined as $\max\limits_{a,b \in S} \mathrm{dist}(a, b)$.

The old definition of the n-level semi-islands and the n-level islands can be used now for the finite variant of space-time (employing the defined above notion of distance on the finite $Space \times Time$).

3.2 Adopting the proof for a torus.

All arguments of Section 2 work for n-level islands in the finite space-time if only their size is small comparative to the size of torus. We shall suppose that all islands of level lower than n_0 are small enough:

Definition 11. *Let us fix the size p of the space-time. Denote by n_0 the maximal integer such that $q_{n_0} \leq p$.*

The following analogs of Proposition 1 and Proposition 2 hold for the finite case:

Proposition 3. *For any $m < n_0$, for all $a \in M \setminus M_{m-1}$, the miracle a belongs to a* minimal *(by inclusion) m-level semi-island S. Such a set S contains exactly 2^m miracles and is of size at most $q_m/3$.*

Proposition 4. *There exists $\gamma > 1$ such that for small enough ϵ (from the definition of perturbation ξ) for any $u \in Space \times Time$ and any $n < n_0$*

$$\epsilon_n(u) \leq (\gamma\epsilon)^{2^n},$$

where ϵ_n is defined as in Definition 4.

The proofs are exactly the same as in the infinite case.

Corollary 3. *Let*

$$Space \times Time = \mathbb{Z}_p \times \mathbb{Z}_p \times \{0, \dots, T\}.$$

Then every miracle belongs to an n-level island for some $n < n_0$ with probability $(1 - \epsilon_{n_0}) \geq 1 - (\gamma\epsilon)^{2^{n_0}}$ (for some $\gamma > 1$, and ϵ from the definition of perturbation ξ).

Proof: Let $a \in M$. By Proposition 3, if $a \notin M_n$ for all $n < n_0$, then a belongs to some n_0-level semi-island. But probability of this event is at most ϵ_{n_0}. □

We say that the set of miracles on the finite space-time is *standard* if each miracle belongs to an n-level island for some $n < n_0$.

Theorem 2. *Let*

$$Space \times Time = \mathbb{Z}_p \times \mathbb{Z}_p \times \{0, \dots, T\}.$$

Suppose that the set of miracles M is standard. Let a be a proper consequence of an n-level island S. Then $\mathrm{dist}(a, S) < 8q_n$.

This theorem can be proved by the same arguments as Theorem 1 in Section 2. From Theorem 2 we get a corollary:

Corollary 4. *Let*

$$Space \times Time = \mathbb{Z}_p \times \mathbb{Z}_p \times \{0, \dots, T\}.$$

Then each point $a \in Space \times Time$ is alive with probability $p > 1 - p^2 T \epsilon_{n_0} - \mathcal{O}(\epsilon)$, if ϵ is small enough.

Proof: First of all note that the total number of points in $Space \times Time$ with a *positive* time coordinate is p^2T. Hence, with probability at least

$$(1 - p^2 T \epsilon_{n_0})$$

there is no n_0-level semi-islands in $Space \times Time$. Further, if there is no n_0-level semi-islands, then each miracle belongs to some n-level island for $n < n_0$, i.e., the set of miracles is standard. Then every alive point is a proper consequence of some n-level island. From Theorem 2 it follows that a proper consequence of an n-level island S must be in $8q_n$-neighborhood of S. Hence, each point is alive with probability at most

$$\sum_{i<n_0} \epsilon_i = O(\epsilon).$$

□

Remark 4. We defined $n_0 = n_0(p)$ as the maximal integer such that $q_{n_0} < p$. It is not hard to see that for a fixed ϵ the value $\epsilon_{n_0} = \epsilon_{n_0}(p)$ tends to zero (as $p \to \infty$) faster than any polynomial in p. Thus, if we want to guarantee that each point in space-time is alive with probability $\mathcal{O}(\epsilon)$, we can let T grow faster than any polynomial in p.

4 Implementing fault-tolerant computations on a 3D cellular automaton

In this section we use the technique developed above, to prove the result from [4]: we explain how to construct a 3-D cellular automaton, which simulates a given 1-D cellular automaton in spite of perturbation (i.e., assuming that any cell at any moment with a small probability can randomly change its state). We assume that in the simulated automaton the state of a cell depends on its own state and the states if its closest neighbors on the previous step.

In our model, we have an infinite space $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$ (if we want to simulate an infinite 1-D array of cellular automata) or $\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$ (if we simulate a finite 1-D array). There is a fixed finite automaton in each cell of the space. Each vertical column in the space, i.e., each family of cells

$$C_{u,v} = \{(u, v, z) \in Space\}$$

is supposed to simulate a 1-D array of cellular automata. One cell of our automaton corresponds to one cell of the simulated automaton. Denote by $s(x, y, z, t)$ the state of the cell with coordinates (x, y, z) at time t.

In the beginning, all columns should be synchronized and represent the initial configuration of the simulated array. If there is no random mistakes, all columns should synchronously simulate the computation of the simulated array. In this case for each $t > 0$ the state $s(x, y, z, t)$ is a function of its closest neighbors in the column:

$$s(x, y, z, t + 1) = Trans(s_{-1}, s_0, s_{+1}),$$

where

- $s_{-1} = s(x, y, z - 1, t)$,
- $s_0 = s(x, y, z, t)$,
- $s_{+1} = s(x, y, z + 1, t)$,

and $Trans$ is the transitions rules of the simulated automaton. Thus, if there is no faults, we just simulate a bunch of 1-D arrays, and they all are working synchronously.

To make the simulation working in presence of faults, each cell should observe also its neighbors in nearby columns. We compose the transition mapping F with Toom's rule. More exactly, the following rule is used:

$$s(x, y, z, t+1) = Trans(s_{-1}, s_0, s_{+1}),$$

where

- s_{-1} is majority of the triple $s(x, y, z-1, t)$, $s(x+1, y, z-1, t)$, $s(x, y+1, z-1, t))$,
- s_0 is majority of the triple $s(x, y, z, t)$, $s(x+1, y, z, t)$, $s(x, y+1, z, t))$,
- s_{+1} is majority of the triple $s(x, y, z+1, t)$, $s(x+1, y, z+1, t)$, $s(x, y+1, z+1, t))$.

We assume that random faults in all cells and any moments of time are independent. Our aim is to prove that if probability of a random mistake at any given point and any time is small enough, then for all x, y, z, t the state $s(x, y, z, t)$ with high probability presents the correct value of the cell z of the simulated array at moment t.

4.1 Dead and alive cells again

To prove correctness of the automaton above we should investigate the behavior of the 'spoiled' cells, i.e., those cells whose state is not correct due to random faults. We define a very simple automaton with only two states: 'alive' and 'dead'. We define it so that at any moment the set of all alive points of the new automaton covers the set of all 'spoiled' cells of the original automaton. In other words, if the original automaton has a spoiled cell at (x, y, z) at moment t, then the new automaton (being disturbed with the same faults) must have an alive point at (x, y, z) at the same time t. Note that converse is not true, i.e, the new automaton can have an alive cell (x, y, z) at moment t even though the corresponding cell of the original automaton is not spoiled.

Let us describe the new automaton. Initially, all cells of the new automaton are dead. The cells are updated by the following rule: *The cell (x, y, z) is alive at moment $(t+1)$ if at least one of the majorities s_{-1}, s_0, s_{+1}, where again*

- s_{-1} is majority of the triple $s(x, y, z-1, t)$, $s(x+1, y, z-1, t)$, $s(x, y+1, z-1, t))$,
- s_0 is majority of the triple $s(x, y, z, t)$, $s(x+1, y, z, t)$, $s(x, y+1, z, t))$,
- s_{+1} is majority of the triple $s(x, y, z+1, t)$, $s(x+1, y, z+1, t)$, $s(x, y+1, z+1, t))$.

Besides the deterministic rule above, at any moment any dead cell can be randomly made alive. As in the previous sections, we call these events 'miracles'. Miracles at different cells and moment of time are independent.

More precisely, let us define the perturbation as a random function $\xi(x, y, z, t)$, which has two values (0 and 1). We suppose $\xi(x, y, z, t) = 1$ with a small probability $\epsilon > 0$, and values of this function at different points of space-time are independent. If $\xi(x, y, z, t) = 1$, we make the cell (x, y, z) of the automaton alive at time t; otherwise (if $\xi(x, y, z, t) = 0$) we apply the deterministic rule above.

We will prove that if ϵ is small enough, then any cell at any time is dead with high probability. Clearly, this result implies that the fault tolerant simulation of a 1D array of cellular automata (defined above) is adequate (i.e., each cell at any moment with high probability has a correct value).

The proof of the result above follows the same plan as our proof of Theorem 1 and Theorem 2. We should just update the definitions to deal with 4-D space-time. Further we explain how the required modifications can be done.

4.2 Adopting the arguments for 3-D space

Now $Space$ is the set $\mathbb{Z}^3$ or $\mathbb{Z}_p^3$. Proposition 1 holds for the new definitions (the old proof is valid). The same is true for Proposition 2 (see Remark 1). The only non-trivial modifications are required in the proof of Theorem 1. We should adopt the definition of explanation paths to 3-D space. This can be done as follows.

Definition 12. *An explanation path from an alive point $a \in Space \times Time$ to an n-level island S is a sequence $a_0, a_1, \ldots, a_m$ of points in $Space \times Time$ such that*

1. $a_0 = a,\ a_m \in S$
2. *for any $k < m$ if a_k has coordinates (x, y, z, t) then a_{k+1} is one of the points $(x, y, z', t-1)$, $(x+1, y, z', t-1)$, $(x, y+1, z', t-1)$, where $z' \in \{z-1, z, z+1\}$ (nine variants in total).*

If there exists an explanation path from an alive point a to an n-level island S, we say that a is a consequence of S.

Definition 13. Space-greedy explanation path: *Let $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to some island S. This path is called* space-greedy *if the following conditions hold for any $i < m$:*

- $a_i \notin S$.

– *Let $a_i = (x, y, z, t)$. This point is a consequence of S but not its member. Hence, at least one of points $(x, y, z', t-1)$, $(x+1, y, z', t-1)$, $(x, y, z', t+1)$ $(z' \in \{z-1, z, z+1\})$ is alive and at least one of them is a consequence of S. For a space-greedy path, if one of points $(x+1, y, z', t-1)$ or $(x, y+1, z', t-1)$ is a consequence of S, then a_{i+1} must be one of these two points. Only if none these points is a consequence of S, $a_{i+1} = (x, y, z', t-1)$.*

Definition 14. South-most explanation path*: Let $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to an island S ($a_m \in S$). This path is called* south-most *if for any $i < m$ the following conditions hold:*

– $a_i = (x, y, z, t) \notin S$.
– *If one of points $(x, y, z', t-1)$, $(x+1, y, z', t-1)$ $(z' \in \{z-1, z, z+1\})$ is alive, then a_{i+1} must be one of these two points. Otherwise $a_{i+1} = (x, y+1, z', t-1)$,*

Definition 15. West-most explanation path*: Let $a_0, \ldots, a_m$ be an explanation path from an alive point $a = a_0$ to some island S ($a_m \in S$). This path is called* west-most *if for any $i < m$ the following conditions hold:*

– $a_i = (x, y, z, t) \notin S$,
– *If one of points $(x, y, z', t-1)$, $(x, y+1, z', t)$ $(z' \in \{z-1, z, z+1\})$is alive, then a_{i+1} is one of these two points. Only otherwise $a_{i+1} = (x+1, y, z', t-1)$.*

Based on this definitions, we can apply the arguments from the proof of Theorem 1 given in Section 2. We omit the details.

Note that in the finite case the computation on a zone p can be simulated during super-polynomial time, see Remark 4.

5 Conclusion

In this paper we presented a detailed proof of Toom's theorem, which says that Toom's 2-dimensional cellular automaton has stable nontrivial global state. We showed how the same ideas help to implement reliable computations based on a 3-dimensional array of faulty cellular automata. We proved that our construction allows to simulate any polynomial time algorithm on faulty cell automata, and the result is correct with probability $1 - O(\epsilon)$ if any cell at any step of computation is corrupted with probability ϵ. We stress that the construction works for any small enough ϵ, for any size of the array (and polynomial number of steps of the computation).

In our opinion (which is, of course, quite subjective), the new proof is more easy to understand than the previous ones. More specifically, we understand better the consequences of faults on computations.

We hope that our work will help to use these methods in other problems concerning cellular automata and related areas.

References

1. Gács P: Lectures on Reliable Cellular Automata. http://www.cd.bu.edu/faculty/gacs/papers/ca-lectures.pdf
2. Toom A: Stable and Attractive Trajectories in Multicomponent Systems, *Multicomponent systems, Advances in Probability*. 1980; 6: 549–576.
3. Berman P, Simon J: Investigations of fault-tolerant networks of computers. 20th STOC. 1988: 66–77.
4. Gács P, Reif J: A simple three-dimensional real-time reliable cellular array, *Journal of Computer and System Sciences*. 1988; 36,2: 125–147.
5. Gács P: A New Version of Toom's Proof. *Tech. report*, Department of Computer Science, Boston University, TR-95-009, Boston. MA 02215, 1995.
6. Gács P: A Toom rule that increases the thickness of sets. *Journal of Statistical Physics*. 1990; 59, 1/2: 171–193.
7. Gács P: Self-Correcting Two-dimensional Arrays, *Randomness in Computation*, Advances in Computing Research (a scientifi c annual). JAI Press, Greenwich, Comm. 1989; 5: 223–326.
8. Gács P: Reliable Cellular Automata with Self-Organization. *Journal of Statistical Physics*. 2001; 103, 1/2: 45–267.
9. Grey L-F: A Reader's Guide to Gács' "Positive rates" Paper. *Journal of Statistical Physics*. 2001; 103,1/2: 1–44.
10. Li M, Vitanyi P: Kolmogorov Complexity and its Applications, 2nd Ed., Springer-Verlag, New York, 1997.
11. Toom A: Contours, convex sets, and cellular automata. Text of a course at the 23-rd Colloquium of Brazilian mathematicians. IMPA, 2001.

Universality of Two Dimensional Sandpiles

E. Goles[a], A. Gajardo[b1]

[a]Centre for Mathematical Modeling and Departamento de Ingeniería Matemática,
Universidad de Chile, UMR 2071-CNRS,
Casilla 170/3-Correo 3, Santiago, Chile

Institute for Complex Systems (ISCV),
Av. Artillería 600B, C^o Artillería, Valparaiso, Chile

[b]Departamento de Ingeniería Matemática,
Universidad de Concepción,
Casilla 160-C, Concepción, Chile

Abstract. We present a negative result for the Universality of Sandpiles in dimension two. Universality is taken in the sense of Banks -which consists into embed a logical circuit in the cellular space. We prove that in this context it is not possible to cross information, giving by this way a strong argument to say that Sandpiles are not Universal in dimension two; at least for the usual neighborhoods. Nevertheless, if a neighborhood of radius two is used, the Universality is possible, which is proved.

Key words: Sandpiles, complexity, Discrete Dynamical System, Cellular Automata, Calculability.

1 Introduction

Universality of dynamical systems is understood as the ability of "simulating" a Turing Machine. There is not agreement in the community on the meaning of "simulation", but many researchers agree that it must be strong enough to imply the existence of undecidable problems related with the dynamics of the system. The first universal Cellular Automaton (CA) was the one proposed by von Neumann [1], when the first formal definition of CAs was given. He defined Cellular Automata and showed how a universal and self reproducing Turing Machine can be emulated through it. He proved at the same time the computing ability of CAs and the existence of automatic systems with the ability of self reproduction, which was his actual goal.

Intended as the ability of computing, Universality was looked for in simple systems. Banks optimized the result of von Neumann defining a Universal CA in dimension two to only two states and the same neighborhood [2]. The smaller neighborhood for dimension two (of size three) was studied by Gajardo et al [3],

[1]Tel. (56-41) 207042, FAX. (56-41) 522055, anahi@ing-mat.udec.cl

where Universality of a three state CA is proved. Another particularly simple example of universal CA is the famous Game of Life [4]. But the simplest one is the elementary CA with rule 110, proved by Cook [5]. Other dynamical systems was studied from this point of view, see for example [6, 7].

In these days, Universality of a dynamical system is viewed as an index which talks about the predictability of the system, where predictability is measured under computer science scope. More precisely, the computational complexity of predicting whether a given pattern will or will not appear for a given initial configuration is studied. If the system is universal, this problem is intractable. This means that in order to know the behavior of the system there is not any method better than observing the system it self.

In order to simulate a Turing Machine with a CA -as the Sandpile,- the most used method consists in embedding an infinite logical circuit in the initial configuration in such a way that the logical circuit is computed when the system runs. The logical circuit is able to compute the evolution of the Turing Machine over some initial word. In this way, the problem of deciding whether some given pattern will or will not appear in the CA evolution when it starts over a given finitely described initial configuration is undecidable. This method was introduced by Banks. Embedding a logical circuit in a two dimensional CA is worked out by defining configurations that emulate wires which transport boolean values, logical gates, wire bifurcations and wire crossovers.

Sandpiles were defined by Bak et al in 1987 [8] as an example of a dynamical system presenting both power law distribution of *avalanches* size and a property which he named *self organized critically*. This property establishes that the system always evolves to a set of states (an attractor) which is critical in the sense that avalanches are frequent. Bak observed this for real piles of sand, and introduced a simplified discrete version with the same property, known as Sandpile and defined as follows.

Definition [2D-Sandpile] *Let $N \subset \mathbb{Z}^2$ be a finite set called neighborhood. The elements of $\mathbb{Z}^2$ will be called cells. The neighborhood of a cell i will be the set $\{i+j | j \in N\}$. A finite number of tokens, $x_i(0) \geq 0$, is assigned to each cell. For $t \geq 0$ the system evolves under the following rule applied synchronously to each cell.*

> *If $x_i(t)$ exceeds $|N|$, then the cell "fires" and all its neighbors increase its number of tokens by the number of times that $|N|$ divides $x_i(t)$, while the number of tokens of i decreases by $|N|$ times this number.*

Goles et al [9, 10, 11] studied Sandpiles from this approach. They shown that the system is Universal for an arbitrary graph of degree at most three. Moore et al [12] generalized this result to the sandpile over a three dimensional square

grid. Banks's method cannot be applied for dimension one, and Moore asserted that logical circuits cannot be computed with a Sandpile in dimension one by using computational complexity arguments. Moore also conjectured that the same is true for dimension two, due to the apparent impossibility of the system of crossing information. In [13] we define formally what a wire crossover is and we prove that it cannot exist in a two dimensional Sandpile. In the following we describe the main steps of the proof.

The Crossover and the Firing Graph

A device as a logical gate or a Crossover is defined by a configuration over a finite portion of $\mathbb{Z}^2$. In order to study the Crossover one may consider, without loss of generality, that it is defined over a finite $n \times n$ square, i. e., it is an assignment of tokens to each cell of a $n \times n$ square.

A configuration is said to be *quiescent* if and only if each cell has strictly less than $|N|$ tokens. When a token is added to a cell of a quiescent configuration, the Sandpile evolves to a new quiescent configuration, in this case we say that an *avalanche* was produced.

A Crossover is defined as a configuration such that: 1) if a token is added to a given cell on its West border, an avalanche is produced, a token falls on its East border and no token falls on its South side. 2) the analogous happens when we add a token on a particular site on its North border.

The Crossover satisfies our intuition about crossing information in the sense that a token appears in the East side if and only if a token is added on a given cell of the West side, and a the analogous happens for the North to South direction.

An important property of configurations defined over finite regions is that if one token is added to the cells of the border, then each cell fires at most one time -if a token is added in the middle of the square it is possible that several cells fire several times. This allows us to introduce the following concept.

Definition 1 (Firing Graph) *Let us consider a quiescent configuration* $c : \{1, .., n\}^2 \to \{0, .., 3\}$ *and a cell* (i, j) *on the boundary of* c*. We define its Firing Graph (see Figure 1) as the directed graph* $G = (V, E)$ *where:*

$V =$ *the set of cells in* $\{1, 2, .., n\} \times \{1, 2, .., n\}$ *that fire if a token is added to the cell* (i, j) *in* c*, and*
E *is defined by* $(u, v) \in E \Leftrightarrow u$ *and* v *are neighbors and* u *fires before* v*.*

Some direct properties of G are the following.

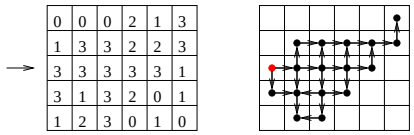


Figure 1: A configuration and its Firing Graph.

- It is connected.
- It has no cycles.
- Only the vertex (i, j) is a source of G (a vertex with in-degree equal to 0).
- If all the neighbors of u are in V then so is u.
- If the in-degree of u is k then u has at least $|N| - k$ tokens in the initial configuration (i.e., $c(u) \geq |N| - k$).

The main results

Theorem 1 *There does not exist a wire crossover for the Sandpile in $\mathbb{Z}^2$ with either a planar neighborhood or with the Moore neighborhood:* $\{(i, j) | \max\{|i|, |j|\} = 1\}$.

The idea of the proof consists in to show that it is possible to suppose that the Firing Graphs associated to the North-South and West-East avalanches G_{ns} and G_{we} respectively are vertex disjoint. This immediately shows the impossibility of the crossover over a planar grid. For the Moore neighborhood we observe that if a vertex of G_{ns} do not fire during the West-East avalanche, then it has more predecessors in G_{ns} than neighbors in G_{we}. But this property makes impossible the connexity of both graphs, carring to a contradiction.

This theorem discards the Banks method for proving the universality of this system. But it do not prove decidability of problems associated with 2D-Sandpiles. On the other hand, if we consider a neighborhood of radius two, the Sandpile becomes universal as the following theorem establishes.

Theorem 2 *The Sandpile over $\mathbb{Z}^2$ with the von Neumann neighborhood of radius $k \geq 2$*
($N = \{0\} \times \{-k, .., -1, 1, .., k\} \cup \{-k, .., -1, 1, .., k\} \times \{0\}$) is Universal.

The proof consists in to define the basic devices that are used to construct logical circuits. Figure 2 shows the devices for this automaton.

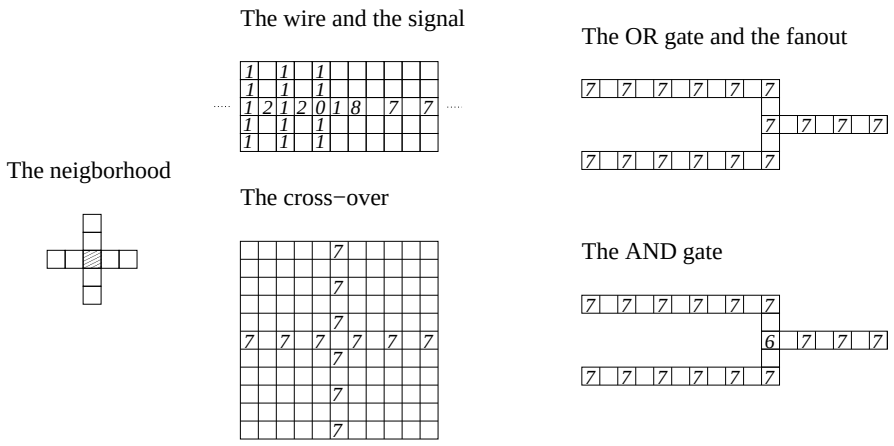


Figure 2: The devices that show the Universality of the Sandpile on a von Neumann neighborhood of radius 2. The wire is a path of cells with seven tokens at distance two. The signal is a cell with eight tokens. When the signal propagates, the wire is distorted.

Open questions

In this work we suppose that a crossover is emulated by a finite configuration which needs to be stable, and that signals are perturbations that propagates over a stable background.

It may be possible to conceive signals propagating over an unstable and periodic background. In this context we could define the different devices as finite patterns which do not interact with the periodical environment in the following way. A set of finite patterns: $\{d_i : A_i \to S\}_{i=1}^{n}$ is a *set of devices* if and only if there exists a periodic configuration $p : \mathbb{Z}^2 \to S$ such that for every i, the configuration g_i defined by $g_i(x) = d_i(x)$ if $x \in A_i$ and $p(x)$ in other case, satisfies that $T^k(g_i)|_{\mathbb{Z}^2 \setminus A_i} = T^k(p)|_{\mathbb{Z}^2 \setminus A_i}$ for every $k \in \mathbb{N}$.

On the other hand, the impossibility of crossing information is an obstacle To apply the Banks method, but it is not known whether it is an obstacle for Universality it self. We think that it is in fact an impediment, but a formal proof of this may require mathematical tools that are not yet developed.

References

[1] Burks AW: Essays on Cellular Automata, Univ Illinois Press, 1970.

[2] Banks ER, Information Processing and Transmission in cellular automata, PhD of the Mass Inst of Tech, Cambridge, Mass, 1971.

[3] Gajardo A, Goles E: Universal cellular automaton over a hexagonal tiling with 3 states. Int J Algebra and Computation 2001; 11: 335-354.

[4] Berlekamp E, Conway JH, Guy R:Winning Ways for Your Mathematical Plays, New York, Academic Press, 1982.

[5] Cook M: Universality in Elementary Cellular Automata. Complex Systems 2004;15: 1-40.

[6] Blondel VD, Bournez O, Koiran P, Papadimitriou CH, Tsitsiklis JN: Deciding stability and mortality of piecewise affine dynamical systems. Theoret Comput Sci 2001;255: 687-696.

[7] Gajardo A, Moreira A, Goles E: Complexity of Langton's ant. Discrete Applied Mathematics 2002;117; 41-50.

[8] Bak P, Tang C, Wiesenfeld K: Self-organized criticality - An explanation of 1/f noise. Phys Rev Lett 1987;59; 381-384.

[9] Goles E, Kiwi M: Games on line graphs and sand piles. Theoret Comput Sci 1993;115; 321-349.

[10] Goles E, Margenstern M: Sand pile as a universal computer. Internat J Modern Phys C 1996;7; 113-122.

[11] Goles E, Margenstern M: Universality of the chip firing game. Theoret Comput Sci 1997;172; 121-134.

[12] Moore C, Nilsson M: The computational complexity of sandpiles. J Statist Phys 1999;96; 205-224.

[13] Gajardo A, Goles E: Complexity of two dimensional Sandpiles. unpublished.

Hierarchical Organization in Smooth Dynamical Systems

Martin Nilsson Jacobi
Complex Systems Group
Department of Energy and Environmental Research
Chalmers University of Technology, 41296 Gothenburg, Sweden
mjacobi@chalmers.se

August 31, 2005

Abstract

We define and characterize hierarchical structures in smooth dynamical systems. Smooth projective maps from the original phase space to a phase space with lower dimensionality define transitions between levels in a dynamical hierarchy. It is required that each level describes a self-contained deterministic dynamical system. We show that a necessary and sufficient condition for a projective map to be a transition between levels in the hierarchy is that the kernel of the differential of the map is tangent to an invariant manifold with respect to the flow. The implications of this condition are discussed in detail. We also show how the projections can be defined using a quotient manifold construction with a Lie group invariant under the flow of the dynamical system. The relation to symmetries of the dynamics is demonstrated. Finally these results are used to define functional components on different levels, interaction networks, and dynamical hierarchies.

Acknowledgment: This work was funded by PACE (Programmable Artificial CellEvolution), a European Integrated Project in the EU FP6-IST-FET Complex Systems Initiative, and by EMBIO (Emergent Organisation in Complex Biomolecular Systems), a European Project in the EU FP6 NEST Initiative.

Combinatorial auctions: From a statistical-mechanics analysis to efficient message-passing algorithms

Michele Leone, Mauro Sellitto, and Martin Weigt*
Institute for Scientific Interchange, Viale S. Severo 65, I-10133 Torino, Italy

September 28, 2005

Keywords: combinatorial auctions, statistical physics, combinatorial optimization, disordered systems, typical-case analysis, message-passing algorithms

Abstract: In this note, we present a statistical-physics framework for combinatorial auctions, i.e. multi-item auctions where bidders bid on combinations of items. The combinatorial problem is represented as a lattice-gas model, such that methods from the statistical physics of complex, disordered systems can be applied. In a minimal probabilistic setting, we find a phase transition from an easily solvable to a harder phase, where the solution space becomes clustered. In addition, the reformulation as a statistical-physics model allows the introduction of new and efficient message passing algorithms for single CA instances.

Auctions are a popular economic institution allowing to sell a variety of commodities [1]. Today, the large diffusion of e-commerce and the use of the Internet as a world-wide market place has brought about fundamental changes in the use of auctions. In situations, in which the number of objects to be sold is large, standard single-item auction protocols are clearly inappropriate. Moreover, objects often exhibit complementary features, such that potential buyers are interested more in a given package of items rather than in separate single objects. Multi-item auctions, in which bidders are allowed to bid on combinations of goods, so-called *combinatorial auctions* (CA), were first motivated by the problem of airport slot allocations (takeoff/landing rights) and radio spectrum licenses, and are now widely used [2, 3]. Theoretical models of CA are interesting to be studied both from an analytical and an algorithmic point of view as prototype examples of new web-based market mechanisms.

In the simplest setting, the CA problem can be formulated as follows: A set $\mathcal{A}$ of objects (goods) is to be sold, and N players (bidders) are given. Every player $i \in \{1, ..., N\}$ submits a sealed bid $\{\mathcal{A}_i, \nu_i\}$, in which he expresses his preference for a package $\mathcal{A}_i \subset \mathcal{A}$ of goods and the price ν_i he is willing to pay for it. The CA is thus a combinatorial optimization problem consisting in determining a collection of winning bids that *maximizes the total auctioneer's revenue* under the condition that no good can be sold twice, i.e. that $\mathcal{A}_i$ and $\mathcal{A}_j$ are disjoint for any two winning bids. Compared to standard single-item auctions, CA have two distinguishing features making them more challenging for theoretical and algorithmic approaches: *(i)* The highest bid is not guaranteed to win. It can be overcome by a collection of various lower bids containing partially the same objects, but giving a higher total revenue. *(ii)* The CA problem is NP-complete, as follows easily in the above setting due to the equivalence to the maximum weighted independent set problem [4]. This is, however, a worst-case result and does not necessarily imply the impossibility of finding optimal solutions in real-life CAs. It is thus important to develop a complementary *typical-case* scenario by considering suitable ensembles of CA instances as a first benchmark test.

As a starting point into this direction, we will focus on a simple probabilistic model of CAs where each player submits a single randomly drawn bid. This choice is motivated by

*corresponding author: weigt@isiosf.isi.it

the following points: *(i)* It allows for a detailed analytical treatment within a statistical-mechanics description. The auctioneer's goal of maximizing his revenue is reduced to finding the ground state of an equivalent hard-sphere lattice-gas model with random chemical potentials [5]. *(ii)* It is conjectured to retain the same level of computational complexity as the most general case. In this sense, any insight into the reasons of computational hardness in the simplified setting can be translated to more general settings. *(iii)* It allows to extend the theoretical description via statistical-mechanics tools to an algorithmic treatment of single CA instances via efficient message-passing procedures. This may bridge the gap between a theoretical analysis on the basis of a probabilistic CA ensemble and the need for fast algorithms in practical applications [5].

To be more precise, the model includes N players and $M = \alpha N$ goods. Each player chooses his package independently by selecting goods with probability z/M. The probability that a player desires ℓ objects is thus, for $M \gg 1$, given by the Poisson distribution $e^{-z}z^{\ell}/\ell!$ of mean z. Analogously, the probability that a good is contained in k bid packages is given by $e^{-z/\alpha}(z/\alpha)^k/k!$. The price is also drawn randomly according to some arbitrary distribution $p(\nu|\ell)$ which may depend on the package size ℓ. The model can be represented graphically in two different ways, cf. Fig. 1: *(i)* The *factor-graph* representation consists of a bipartite graph. Nodes are bidders and goods, and an edge signifies that a good is element of a bidder's package. *(ii)* The *conflict-graph* representation contains only the bidders as nodes. Two of them are connected whenever their bids are in conflict, i.e. whenever their packets contain at least one common good. This conflict graph naturally has the characteristics of a *small-world network*: It has short distances $\mathcal{O}(N)$ inside each connected component, and it has a non-trivial clustering coefficient due to objects wanted by more than two players.

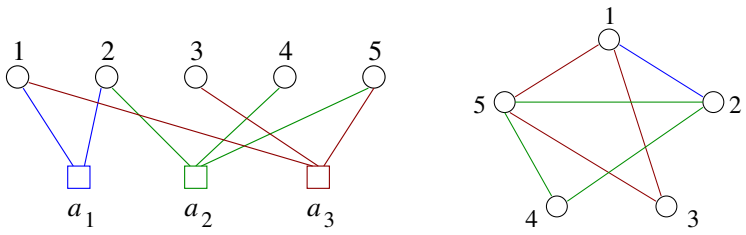


Figure 1: Factor-graph and conflict-graph representation of a combinatorial auction. Circles are bidders, squares present items.

The conflict graph is the starting point of the statistical-mechanics analysis. Let us consider a *gas of hard particles on this graph*: Each node can be position of up to one particle, described by the occupation number $x_i \in \{0, 1\}$, and is subject to a local chemical potential ν_i (being equal to the value of the bid of player i). In this representation, the presence of a particle ($x_i = 1$) will be interpreted as a winning bid, an empty node corresponds to a loosing bid. The total revenue is consequently given by $\sum_{i=1}^{N} \nu_i x_i$. We still have to implement the constraint that no good can be sold more than once: It is obviously equivalent to the statement that no neighboring nodes in the conflict graph can be occupied simultaneously, i.e. $x_i x_j = 0$ for all $(i, j) \in E$, with E denoting the edge set of the conflict graph. The resulting hard-sphere lattice-gas model can be rephrased in the following partition function

$$\Xi = \sum_{\{x_i\}\in\{0,1\}^N} \exp\left(\beta \sum_{i=1}^{N} \nu_i x_i\right) \prod_{(i,j)\in E} (1 - x_i x_j)$$

where the revenue is coupled to the formal inverse temperature β. The last product implements the hard-sphere constraint, whereas a positive β assigns a higher weight to configurations corresponding to a higher revenue of the CA. Consequently, the maximal revenue is

given by

$$\mathcal{R} = \lim_{\beta\to\infty} \frac{\partial}{\partial\beta} \log \Xi \; .$$

Some of the results of this analysis are summarized in Fig. 2.

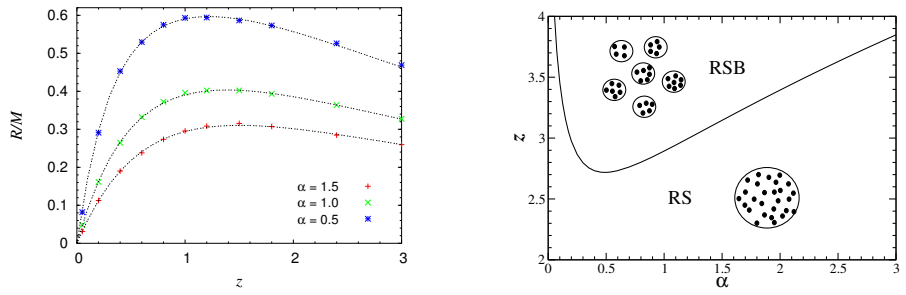


Figure 2: *Left:* Maximized revenue for different values of $\alpha = M/N$, as a function of the average package size z. Analytical results are compared to simulated annealing on random CA instances. Prices are fixed to $\nu_i = 1$. For small z, many goods are not contained in the packages, for high z many conflicts appear. This explains the revenue maximum at an intermediate value of z. *Right:* Phase diagram of the same model. Below the line, all solutions of maximal revenue are contained in one single cluster inside the configuration space $\{0,1\}^N$. Finding one seems to be simple (easy phase). At the line, the model undergoes a phase transition to a clustered solution space, and local cost minima appear. Finding an optimal solution becomes computationally more demanding (hard phase). It is currently under vivid discussion, in how far these local minima generally trap local search algorithms.

Technically, the statistical mechanics analysis is based on the so-called cavity method, cf. [5], which can be reformulated as a message passing algorithm [6, 7] using the ideas of [8]. In the easy phase, this can be realized via the so-called warning- or belief-propagation procedure, whereas message passing in the hard phase requires the application of the survey-propagation algorithm. Technical details of this approach, together with a comparison to standard approaches as simulated annealing or linear programming, go beyond the scope of this note, and will be presented in a separate publication.

References

[1] V. Krishna, *Auction Theory* (Academic Press, 2002).

[2] P. Cramton, Y. Shoham, and R. Steinberg (eds.), *Combinatorial Auctions* (MIT Press, to app. 2005).

[3] S. de Vries and R. Vohra, INFORMS J. of Comp. **15**, 284-309 (2003).

[4] M.R. Garey and D.S. Johnson, *Computers and Intractability* (Freeman, 1979).

[5] A.K. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems* (Wiley-VCH, 2005).

[6] F.R. Kschischang, B.R. Frey, and H.A. Loeliger, IEEE Trans. Info. Theo. **47**, 498 (2001).

[7] J.S. Yedidia, W.T. Freeman, and Y. Weiss, Technical Report TR-2001-22 of the Mitsubishi Technical Laboratory (2001).

[8] M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).

Flows of information as the driving force behind chemical pattern formation[1]

Kristian Lindgren and Anders Eriksson

Physical Resource Theory – Complex Systems
Chalmers University of Technology
SE-412 96 Göteborg, Sweden

Keywords: self-organisation, information theory, pattern formation, second law, information flow

Pattern formation in chemical systems is a dynamical process that has been extensively studied in the litterature since the original work by Turing. Thermodynamic analysis of this type of self-organising systems is of importance in order to understand some of the constraints that self-organisation meets, primarily the need of free energy as a driving force behind formation of spatial patterns.

We have developed an information-theoretic framework that goes one step further in the analysis of physical constraints in chemical pattern formation, see Lindgren et al (2004)[2]. The formalism is based on a thermodynamic information quantity (via statistical mechanics), and this makes it possible to connect an information-theoretic characterisation of a spatial pattern with the free energy driving the system. In this way, a consistent picture of the pattern formation process in terms of free energy being transformed into information in a spatial pattern and eventually destroyed by entropy production when reactions and diffusion processes tries to bring the system towards equilibrium.

In our analysis the information in the pattern is decomposed into contributions from both different positions and different length scales. The overall picture we get is an inflow of information at large length scales, due to the inflow of chemical free energy. Information then flows down in length scale (and also across space), where accumulation at certain positions is connected with the pattern formation. Information is lost from the system at the finest length scales. The whole process is summarised in a continuity equation for information.

In the present paper we investigate the possibility to use this formalism to make predictions on how pattern formation may depend on the structure of the driving force, i.e., the inflow of free energy. Preliminary results indicate that the information flow is generally going in the direction described above – from larger to smaller length scales – which may be viewed as generalised ”second law” of information destruction. If the characteristic length scale of the free energy inflow is reduced below the length scale of the patterns in the system, the flow will not be able to support the structures built up and neither will new structure emerge unless that happends on a smaller length scale. In the same way as ambient heat has too low energy quality to drive a physical process, a chemical free energy inflow of too low length scale characteristics may be insufficient to support pattern formation.

---

[1] This work was funded by PACE (Programmable Artificial CellEvolution), a European Integrated Project in the EU FP6-IST-FET Complex Systems Initiative, and by EMBIO (Emergent Organisation in Complex Biomolecular Systems), a European Project in the EU FP6 NEST Initiative.

[2] Lindgren, K., Eriksson, A., and Eriksson, K.-E., ”Flows of information in spatially extended chemical dynamics,” in *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, Pollack, J., Bedau, M., Husbands, P., Ikegami, T, and Watson, R. A. (MIT Press, 2004).

Toward a multi-scale approach for spatial modelling and simulation of complex systems

Thi Minh Luan Nguyen , Christophe Lecerf , Ivan Lavallée

Abstract— **Complex systems are composed of many heterogeneous elements organized in a hierarchical way, whose mutual interactions make emergent collective behaviors to appear at the highest levels of observation. In some kind of complex systems, especially in biology as shown by the integrative physiology theory [6], space and geometry have a significant role in the simulation results. In this paper we expose a formalized method for modelling and simulation of complex system, going from structural modelling to dynamic simulation while integrating geometrical information in behavior study. Our solution relies on three kind of concepts and techniques: hierarchical graphs for modelling the system structure and organization, Zeigler's formalisms for the specification of agents [18] and a space aware Multi Agent System for agent-based simulation. It is shown how complex system simulation benefits from the combination of agent-based simulation and DEVS.**

Keywords— **complex system, hierarchical graph, agent-based simulation, DEVS, geometry, multilevel and multi-scale analysis, integration.**

I. Introduction

Although there is not a widely accepted definition of complex systems, it is commonly recognized that they are formed of many heterogeneous elements organized in a hierarchical way whose mutual interactions make emergent collective behaviors to appear at the highest levels of observation.

Moreover in biology, as shown by the integrative physiology theory [6], space and time appear both in the speed of signal propagation (humoral, electrical or chemical), and in the changes of spatial relationships between elements (embryology). This aspect is often neglected but, although oversimplifying is acceptable in the first approximation, space has a significant meaning in biology. We believe that the future belongs to models that can integrate and use geometrical data. We propose here a method and a set of formalisms for studying complex systems that goes from structure to behaviors, taking into account space and geometry.

A complex system is composed of a set of components, each of them being itself a set of sub-components, in which various interactions between different levels of organization take place. Their basic properties are presented in [4], we mention here the most important ones:
- A hierarchy in the structural organization may be described.
- Feedback circuits exist in the functional organization of the system, and in a hierarchy between these circuits as well.
- The system exhibits some emergent properties.
- The dynamics is typically non-linear.

One of the most important tasks in studying complex systems is the representation of their structure. Using the modelling techniques summarized by Jennings in [13] (decomposition, abstraction and organisation), modelling a complex system becomes more tractable and a hierarchical graph (detailed in the next section) can be used to represent the structure and the communication inside these systems better than types of models such as equation-based or cellular automata models. Indeed, equation-based models give a (mathematically) formalized, synthetic comprehension of the studied phenomenon, but they are difficult to improve because of the absence of modularity, and only offer the vision at the macroscopic level behavior because they use aggregate parameters. Cellular automata is a framework to explore the dynamics of complex systems whose components are distributed spatially. Time and space are represented in a discrete way. But neither heterogeneity of the complex system components, nor continuous variations can be easily studied with the cellular automata formalism.

Multi-agent systems are a better candidate for modelling the structure of complex systems. This approach aims to represent a complex system with a set of interacting autonomous entities, the agents. This technique has an important role to play since we wish to study the system behavior at a macroscopic level and we know that is the result of interactions at microscopic level. The hierarchy as well as the geometry are well supported there. On the contrary to equation-based models, multi-agent models focus on the constituent entities of the system rather than aggregate variables representing the average variation of entities. As a consequence, observations are related to population rather than on individual entities but, unfortunately, the modularity of multi-agent method hides a lack of formalism.

L. Nguyen (lnguyent@univ-paris8.fr) and I. Lavallée (ivan.lavallee@univ-paris8.fr), Laboratoire de Recherche en Informatique Avancée, Université Paris 8 et EPHE, 41 rue G. Lussac, 75005 Paris, France

C. Lecerf (christophe.lecerf@ema.fr), Ecole des Mines d'Ales - Site EERIE Parc Scientifique Georges Besse F 30035 Nimes Cedex 1, France

II. A FORMALIZED METHOD

We have developped our study of complex systems on an interdisciplinary approach, using both mathematical, biological and computer sciences concepts. In short, we propose a method for studying system that goes from structure to behaviors, as symbolized in fig 1. The first step is, from the real system, to describe the system structural and functional organization using a hierarchical graph. Then, the behavior of each node (class of components) will be described by Zeigler formalism [18]. Lastly, a space aware multi-agent system will be used as an exploration environment to observe the system dynamics.

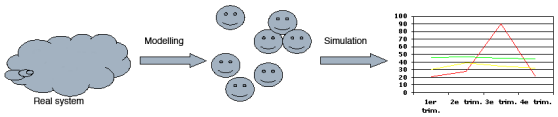


Fig. 1. From sructural modelling to behavior simulation

A. Hierarchical graph

Graph is a structure that is used in the modelling of very diverse situations and, expressed as graphs, many usual problems could be brought back to traditional problems of the graph theory: shortest path, cycle detection, connected components, etc. From a modelling point of view, the graph appears as a coupling intermediate between the physical system and its associated mathematical model. A hierarchical graph offers the ability to represent and view multiple levels in the structure and in the functional organization of the system.

A hierarchical graph (see [9] for a detailed definition) has two types of nodes: atomic and complex.

- The atomic node is the leave of the hierarchy and does not have any internal state.
- On the contrary, the complex nodes have an internal state, which is another hierarchical graph.

A complex node contains a hierarchical graph which may contain other complex nodes in a recurrent scheme. In other words, any complex node in the hierarchical graph may be a component of another complex node of a higher level. Using hierarchical graphs for modelling (i) the structure and (ii) the organization of the system enables multi-scale analysis and viewpoints on both the structure and the organization. Classical graph tools such as cycle detection appear to be useful for the analysis of the system. In the end, a unique complex node represents the whole complex system, assumed that each of its components is possibly a complex node: the entire system is just the one in the highest level in the hierarchy.

B. Zeigler formalisms

This set of formalisms was chosen thanks to its capacity to integrate heterogeneous models, its coupling possibilities and its hierarchical decomposition feature.

Ziegler's set of formalisms (DEVS, DESS, DTSS, DEV&DESS, [18]), allows the system dynamics specification in a modular and hierarchical way that is based on the definition of two types of models: atomic models and coupled models. Atomic models are used to specify elementary input/output behaviors. Coupled models are defined by specifying how basic models (atomic models or coupled models at the lower level) interconnect. A coupled model can then be considered as the basic model of a higher level coupled model. In parallel, Zeigler has also developed the concept of abstract simulator. An abstract simulator represents an algorithmic description of implicit instructions for generating DEVS models behaviors. Moreover, the separation between modelling and simulation does not compel us to redefine simulators for newly defined models. In addition, recent work shows that DEVS can "encapsulate" equation-based model ([18], [8]). This formalism is thus well adapted to the specification of the multi-agents models when the constituted models can be expressed in DEVS. Note that Zeigler formalisms deal only with the behavior of the studied systems, the geometrical information cannot be taken into account.

DEVS is a formalism introduced by Zeigler in 1976. This formalism is based on an abstract mathematical object called system, which can be approximated with an automaton. Basically, a system is described by a time base, input, state, output and function for determining the next state and output for a given state. Two founding types, atomic and coupled, are described.

B.1 Atomic model

The atomic model is the basic element of DEVS (see fig 2), it has the following structure:

$A =< X, Y, S, \delta_{int}, \delta_{ext}, \lambda, t_a >$

- X : input set which is the value of input events;
- Y : set of output value;
- S : set of state;
- δ_{int} : internal transition functions. It is used to describe state transition due to internal events;
- δ_{ext} : transition functions due to external events;
- λ : output function which generate external events at the output;
- ta : time advance function;

At any time, the system is in state S. In the absence of external event, system remains on current state during the time given by the time advance function ta. On the

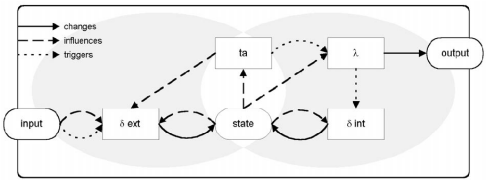


Fig. 2. Internal structure of atomic models ([15])

contrary, it receives external event X by its input port, and the external transition function δ_{ext} will then specify how system changes due to this effect. Then, an event Y which is generated by output function λ is sent to output port. Based on current state, value of external event and the one of time advance function, the next state S is computed. From the outside, this model looks like a black box.

However, a biological system does not contain only such a simple component. In fact, it is composed of many complex components, which are described with sets of subcomponents organized in many levels. Zeigler introduced the coupled model type to fit these requirements.

B.2 Coupled model

A coupled model is composed by a set of components (which are atomic or coupled models, see fig 3) and the coupling of these components. It is defined by a set of input, output ports, a set of constituted components, and coupling among these components. Coupled models allow hierarchical modelling and from a higher level a coupled model can be expressed as an atomic model [18]. The hierarchical aspects of biological system can therefore be naturally modelled with DEVS.

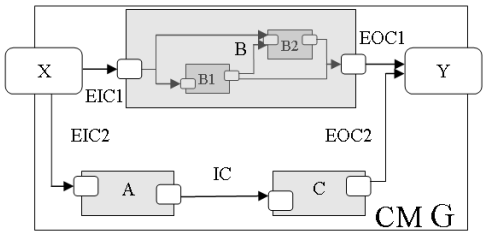


Fig. 3. Coupled models

Coupled model has the following structure:
$C =< X, Y, N, M_d, EIC, EOC, IC, Select >$

- X : set of input ports and values;
- Y : set of output ports and values;
- N : subcomponents list;
- M_d : for each $d \subset N$, M_d is a component described in form of atomic model;
- EIC : external input coupling connect external input to component input;
- EOC : external output coupling connect component output to external output;
- IC : internal coupling connect component output to component input;
- $Select$: the tie breaking function to arbitrate the occurrence of simultaneous events;

Let us consider a coupling component, which consists of a set of atomic components M_d where $d \subset N$. At time t, an atomic component d is in state S_d since e_d (time passed since the last change state of d). The time during which each component d must remain in state S_d if no external event occurred is $ta_d(S_d)$. As a result, a component d will stay at S_d for $\sigma_d = ta_d(S_d) - e_d$. An internal event δ_{int} is scheduled for the component d at $t + \sigma_d$. Suppose that ta is the time scheduled for the first internal event then ta is the smallest value of all $ta_d(S_d)$, that means $ta = Min\{(ta_d(S_d))/d \subset N\}$. The priority list $Select$ allows us to choose among various components having the same σ_d. The atomic component chooses executes its output function and sends the result to all it's influenced neighbours. Then, this component starts the internal transition function δ_{int}, and changes state. We can explore the effects of an arriving external event on an atomic model in the same way. These behavioral components are inter-connected to exchange information through their input/output ports (also called detectors and effectors). Due to the recursion sheme, such a component can be considered in turn like a basic element in a larger model.

Furthermore, Zeigler formalisms do not only model discrete event system, but also deal with continuous and hybrid system thanks to DEV&DESS. This is an extension of DEVS that includes DEVS, DESS and DTSS. Consequently, it is possible to specify some system components by Differential Equations and the others by Discrete Event or Discrete Time systems, the different parts being in interaction to constitute the whole system dynamics.

In biological systems, not only do the dynamic processes vary in time, but also does the topology (see [5]). Clearly, a perturbation of the topology of biological system will affect its evolution both in its (re-)organization and its dynamics. These two aspects are indissociably linked, so that the dynamics may be considered as a consequence of the topological, geometrical and dynamical coupling of the processes involved. Unfortunately, the geometical information cannot be represented in Zeigler's formalisms.

C. Multi-agents systems and simulation

Multi-agent systems (MAS), developed within the framework of distributed artificial intelligence, represent a promising tool to model the dynamics of space aware systems. MAS allow us to represent hierarchy and ge-

ometrical informations, but then lack a formal specification, making mathematical demonstrations impossible. Indeed, there are numerous formal specifications for MAS, for example Object-Z [12], Petri nets ([2], [16]), the colored Petri nets [3], etc. These formalisms are used to describe architectures, behaviors, etc of MAS and agents, but the hierarchical concept is not supported by any of them. However, as hierarchy is one of the most important of integrative physiology fundamental concepts [6], we had to combine Zeigler formalisms ans MAS in order to circumvent both the lack of formalism in MAS and the lack of geometrical representation in DEVS.

Agents are implemented to have internal data representation (memory or state). They possess also means for modifying their internal data representation (perception) and means for modifying their environment (behavior). Different types of agent and their concrete implementation can be found in ([14], [1]). For our study purpose, we used exclusively situated reactive agents. A situated agent lives in an environment where the space is explicitly described.

A multi-agent system is made up of a set of agents evolving in a common environment. Situated MAS are generally made of elementary memoryless agents with a defined position in time and space. Reactive situated agents perform their actions as a consequence of the perception of signals coming either from other agents or from the environment, and are sensitive to the spatial relationships that determine constraints and abilities for actions as well as privileged cooperation relationships. The environment in which agents are situated can reproduce a physical space.

C.1 MAS decomposition

In the outline of Duboz [8], let us make a formal description of a MAS with DEVS. For this purpose, we consider our model according to the four dimensions identified by Yves Demazeau in his methodology "vowels" [7]: "Agents", "Environment", "Interactions" and "Organization"

C.2 Organization

An organization is considered as a configuration that describes how its members act on each other to achieve the goal. In this context, the DEVS coupled models allow to integrate various agents in order to form composed agent (called group of agents). The whole task is then divided in a set of secondary tasks, which are distributed to the group members of the MAS. The MAS hierarchical organization is naturally described by the definition of the atomic and composed agents. The MAS recursion scheme allows us to represent the hierarchical nature of the functional organization of any biological system according to its hierarchical graph model.

C.3 Agent

C.3.a Atomic agent. First, we use a DEVS model to describe atomic agents. Based on this elementary agent, higher-level agents called composed agents are built.

$DEVS$ = {X, Y, S, δ_{ext}, δ_{int}, λ,}

$DEVS$ model for an atomic agent

X : sensors set

Y : effectors set

S : agent possible states set

C.3.b Composed agent. We adopt a recursive definition of composed agent based on atomic agent and composed agent of a lower level.

composed agent $\rightarrow$ composed agent | atomic agent

The DEVS coupled model for composed agent

$N = \langle X, Y, D, \{M_d\}, \{I_d\}, \{Z_d\}\rangle$

where X is the set of input events; Y is the set of output events; D is an index for the components of the coupled agent, $\{M_d\}$: set of constituted agent and $\forall d \in D$, M_d is a basic agent (that is, an atomic or composed agent), I_d is the set of influences of agent d (that is, the agents that can be influenced by outputs of agent d), and $\forall j \in I_d$, Z_{dj} is the d to j translation function. We can see that composed agents are defined as a set of basic components (atomic or coupled) interconnected through the agent's interfaces. The translation function is in charge of converting the outputs of an agent into inputs for the others where N describe a composed agent.

C.4 Interaction

The basic interactions between agents are realized by exchanging messages with their environment via sensors and effectors.

Perception is represented in DEVS by the arrival of external events, which cause state changes of at least one component. We consider the perception of an agent as the change of its internal state due to an external event (external stimulus).

Action refers to pro-action and reaction. Regarding agent A as a coupled DEVS, the set of all external transition functions that do not receive events from coupled model input ports, plus all internal transition functions, define the autonomous behavior of this agent (pro-action). All transition functions driven by external events define the reactional behaviour of the agent.

C.5 Environment

Within the framework of integrative physiology [6], the operation of biological systems strongly depends on the geometrical distribution of the constituent entities, and the environment of MAS, i.e. the structure in which agents evolve, will take into account this information. Generally, the environment may be [10]:

- an interaction medium
- a space in which agents can move around;
- a place where resources are available.

For our study purpose, "environment" corresponds to a space in which the agents have an explicit position. It is considered as a surface divided into cells (see fig 4). Environment is viewed as a collection of n $(n >= 1)$ cells and these cells have a definite size (in a two dimensions environment, a cell is defined by its height and its width, as consequence the environment size $t = (n \times height) \times (n \times width)$. These cells form a matrix whose size is determined by grid parameters. Each cell can contain one or more elements. The cells not hosting any entity have value 0. The value of the other cells corresponds to the density of the elements that they contain.

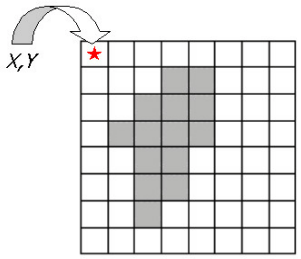


Fig. 4. Environment matrix

Environment constitutes an essential part of situated MAS but only a few works were devoted to their modelling [10]. We believe that it is important to conceive a MAS with a geometrical space representation because an explicit definition of the spatial structure of agent environment allows the definition of distance and adjacency among situated agents. Our solution provides a model that can take into account not only the system hierarchical nature but also the spatial relationship between agents, and even the changes in the system geometry.

III. Application

In order to illustrate the potential of our approach for complex system simulation, we have made an application in the biological neural network field, a case study being the hippocampus.

The hippocampus is part of the cerebrum, and it's one of the area of the brain that deals with memory. The hippocampus plays an essential role in many normal physiological functions, such as information processing, learning and memory formation, as well as in several physiopathological conditions, such as epilepsy and Alzheimer's disease.

We have developed a simple example simulating the hippocampus tissue, using a space aware MAS (a 2D matrix for a hippocampal slice) and the Hodgkin Huxley model [11] for neuron-agents implemented through the following DEV&DESS model:

$HH = \{X, Y, S, ta, \delta_{int}, C_{int}, \lambda, \delta_{ext}, f\}$

The neuron behavior is considered as a hybrid process, the internal evolution is continue, emission and reception of action are discrete. The system has two output: a discrete output "action potential" and a continue output "potential". The continue state variable potential V is used to take into account internal potential evolution. The neuron will produce an "action potential" valued V_{AP} when $V \geq V_\theta$. V is then reset to resting potential V_{rest}.

$$S = S^{discr} \cup S^{cont}$$
$$S^{discr} : \{state | state = \{active, passive\}\}$$
$$S^{cont} : \{V | V \in R\}$$
$$\delta_{ext}(V, x, t) = V + \frac{dV}{dt} + f(w_x)$$

$$\delta_{int}(state, V) \quad \begin{array}{ll} \text{if } state = active & \text{then } state = passive \\ \text{else} & state = active \\ & V = Vrest \end{array}$$

$\lambda(state)$:

if $state = active$	then make an impulsion
otherwise	nothing

$ta(state, V)$:

if $state = passive$	then $ta(state, V) = +\infty$	if $V < V_\theta$
	$ta(state, V) = 0$	if $V \geq V_\theta$
otherwise	$ta(state, V) = t_{ref}$	

$C_{int}(V, x, t)$

=true	if $V_0 \geq \theta$
=false	otherwise

With Huxley-Hodgkin model, f is described by:

$$f = V' = (\overline{g_{Na}} m^3 h(V_{Na} - V) + \overline{g_K} n^4 (V_K - V) + g_L(V_{rest} - V) + I_{inj}(t))/C_m$$

Results of the simulations are available in the form of videos at http://oss.ephe.sorbonne.fr/∼ntmluan/index.htm.

IV. Conclusion

With our method, the integration of well-known approaches (hierarchical graphs, DEVS and agent based simulation) gives a complete process for studying the dynamics of complex systems made up of interacting parts, whatever the field of the considered system. Once defined and build, running such a model relies on instantiating agents population, letting the agents interact in the space aware environment, thus leading to a simulation while monitoring what happens.

In our proposed method, the system modelling process is based on the decomposition of a given real system into various inter-connected elements using hierarchical graphs to represent the system structural organization. Each element is represented by a node, which can be described by a subgraph on a different hierarchical level, and the connection is represented by an edge, forming a multi-scale graph all together. From the behavioral point of view, a hierarchical DEVS formalism is used to describe the behavior of components that are implemented as agents in a situated MAS. At the lowest level, an atomic DEVS component corresponding to an atomic node describes the behavior of an agent in the situated MAS. At the higher level, a coupled DEVS describes a system as a network of coupled components whose connections denote how components influence each other, according to the graph of interactions that represents the organization. Moreover, an explicit definition of the spatial structure of agents environment allows the definition of distance and adjacency among situated agents. Thus, our proposed solution provides a model that can take into account not only the system hierarchical nature but also the spatial and geometrical relationship between agents that we believe has a significant meaning.

With Ziegler's formalisms, agent-based models can be combined with equation-based models because, within an individual agent, behavioral decisions may be done by evaluating theses equations [17].

We have presented in this paper a general modelling and simulation method based on (i) hierarchical graphs enabling multiscale analysis, (ii) DEV&DESS bringing a hierarchical formalism for agent specification, and (iii) a situated MAS reflecting the system geometry. The attractiveness of this method lies on its ability to be used in various domains, and thus to reduce the model building cost.

References

[1] F. Brazier, B.D. Keplicz, N.R. Jennings and J. Treur. (1995). *Formal Specification of Multi-Agent Systems: a Real-World Case*. First International Conference on Multi-Agent Systems (ICMAS'95). San Francisco, 25-32.

[2] I. Bakam, F. Kordon, C. Le Page and F. Bousquet (2001). *Formalization of a spatialized multi-agent system using coloured Petri nets for the study of a hunting management system*. J.L. Rash et al. (Eds): FAABS 2000, Lecture Notes in Artificial Intelligence, 1871: 123-132.

[3] K. Jensen (1997). *Coloured Petri nets: basic concepts, analysis methods and practical use*. Berlin: Springer.

[4] Center for the study of complex systems, *The Study of Complex Systems* http://www.pscs.umich.edu/CSCS/complexity.html.

[5] G. A. Chauvet (1993). *Hierarchical functional organization of formal biological systems: a dynamical approach. I. An increase of complexity by self-association increases the domain of stability of a biological system* Phil Trans Roy Soc London B 339: 425-444.

[6] G. A. Chauvet (1996). *Theoretical systems in Biology: Hierarchical and Functional Integration* volume I, II, III. Oxford : Pergamon.

[7] Y. Demazeau (1995). *From interactions to collective behaviors in agent-based systems*. In Proceeding of First European Conference on Cognitive Science.

[8] R. Duboz (2004). *Intégration de modèles hétérogènes pour la modélisation et la simulation de Systèmes complexes: Application la modélisation multi-échelles en Ecologie marine*. Thèse. Ecole doctorale de l'Université du Littoral - Côte d'Opale.

[9] G. Engels and A. Schürr (1995). *Encapsulated Hierarchical Graphs, Graph Types, and Meta Types*. Departement of computer science, Leiden university.

[10] J. Ferber (1995). *Les systèmes multi-agents: vers une intelligence collective*. InterEditions, Paris.

[11] A. L. Hodgkin and A. F. Huxley (1952). *A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve*. Journal of Physiology. 500-544.

[12] V. Hilaire, A. Koukam, P. Gruer, J-P Muller (2000). *Formal Specification and Prototyping of Multi-agent Systems*. Lecture Notes in Computer Science (vol 1972: 114-127).

[13] N. R. Jennings (2001). *An Agent-based Approach for Building Complex Software Systems*. ACM 44(4): 35-41.

[14] P. Paruchuri, A. R. Pullalarevu and K. Karlapalem (2000). *Multi agent simulation of unorganized traffic*. Proceedings of the first international joint conference on Autonomous agents and multiagent systems. Bologna, Italy. 176 - 183.

[15] A. M. Uhrmacher and B. Schattenberg (1998). *Agent in discrete event simulation*. European Simulation Symposium "Simulation in Industry - Simulation Technology: Science and Art" (ESS'98), Nottingham, SCS Publications, 129-136.

[16] F. Vernadat, A. Lanusse and P. Azéma (1994). *Modélisation par réseaux de Petri d'un langage acteur : Application à la vérification de systèmes multi-agents*. Actes des 2èmes Journées Francophones IAD-SMA.

[17] H. V. D. Parunak and R. S., Rick L. Riolo (1998). *Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide*. Proceedings of Multi-agent systems and Agent-based Simulation, Springer.

[18] B.P. Zeigler, H. Praehofer and T.G. Kim (2000). *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. Academic Press.

PARALLEL vs. SEQUENTIAL THRESHOLD CELLULAR AUTOMATA: Comparison and Contrast

Predrag T. Tošić * and Gul A. Agha
Open Systems Laboratory, Department of Computer Science,
University of Illinois at Urbana-Champaign,
1334 Thomas Siebel Center for Computer Science, 201 N. Goodwin, Urbana, IL 61801, USA
Contact author's phone: +1-217-244-1976 Fax: +1-217-333-9386
Electronic mail: {p-tosic, agha}@cs.uiuc.edu

ABSTRACT

Cellular automata (CA) *are an abstract model of a distributed dynamical system, as well as of* fine-grain parallelism *in computing. In a classical cellular automaton, all the nodes execute their operations in parallel and* in perfect synchrony. *We consider herewith the sequential version of* CA, *called* SCA, *and compare those* SCA *with the classical,* parallel CA. *In particular, we show that there are* 1D CA *with very simple node update rules that cannot be simulated by any comparable* SCA, *irrespective of the node update ordering. Consequently, the granularity of the basic* CA *operations and, therefore, the* fine-grain parallelism *of the classical, synchronous* CA, *insofar as the "interleaving semantics" is concerned, turns out to be* not fine enough. *We also study in some detail the properties of the cellular automata whose nodes update their states according to the* Majority *update rule. Finally, we share some thoughts on how to extend the presented results, and, in particular, we try to motivate the study of* genuinely asynchronous cellular automata.

Keywords: analysis and dynamics of complex networks, cellular automata, discrete dynamical systems, configuration space properties, communication models

1. Introduction and Motivation

Cellular automata (CA) were originally introduced as an abstract mathematical model that can capture the behavior of biological systems capable of self-reproduction [24]. Subsequently, CA have been extensively studied in a great variety of application domains, mostly in the context of complex physical or biological systems and their dynamics (e.g., [16, 36, 37, 38, 39]). However, CA can also be viewed as an abstraction of massively parallel computers (e.g, [11]). Herein, we study a particular simple yet nontrivial class of CA from the parallel and distributed computing perspectives. In particular, we pose - and partially answer - some fundamental questions regarding the nature of cellular automata's parallelism.

It is well known that CA are an abstract architecture model of *fine-grain parallelism,* in that the elementary operations executed at each node are rather simple and hence comparable to the basic operations performed by the computer hardware. In a classical, parallel CA, all the nodes execute their operations in parallel and *in perfect synchrony,* that is, *logically simultaneously:* in general, the state of a node x_i at time step $t+1$ is some simple function of the states of the node x_i and a set of its pre-specified neighbors at time t.

We consider herewith the sequential version of CA, that we shall abridge to *SCA* in the sequel. We shall compare SCA with the perfectly synchronous *parallel* (or *concurrent*) CA. In particular, we will show that there are *1D CA* with very simple state update rules that cannot be simulated by any comparable SCA, irrespective of the node update ordering. While the

*Contact author.

result would be trivial if one considers a single (S)CA computation, we argue that the result is both nontrivial and important when applied to *all possible inputs* (starting configurations) and to the entire classes of CA and SCA. Hence, the granularity of the basic CA operations, insofar as the (im)possibility of simulating their concurrent computation via appropriate sequential interleavings of these basic operations, turns out not to be quite *fine enough.* Furthermore, we will characterize in some detail the possible computations of certain types of *threshold cellular automata.* Last but not least, we will also share some thoughts on how to extend the work presented in this paper. In particular, we will try to motivate the study of *genuinely asynchronous cellular automata,* where the asynchrony applies not only to the local computations at individual nodes, but also to *communication* among different nodes.

An example of asynchrony in the local node updates (i.e., asynchronous computation at different "processors") is when, for instance, the individual nodes update one at a time, according to some random order. This is a kind of asynchrony found in the literature, e.g., in [19, 20]. It is important to understand, however, that even in the case of what is referred to as *asynchronous cellular automata* (ACA) in the literature, the term *asynchrony* there applies to local updates (i.e., computations) *only,* but not to communication, since a tacit assumption of the globally accessible global clock still holds. We prefer to refer to this kind of (weakly asynchronous) (A)CA as to *sequential cellular automata,* and shall consistently keep the term *asynchronous cellular automata* only for those CA that do not have a global clock (see Section 4).

Throughout, we use the terms *parallel* and *concurrent* as synonyms. Many programming languages experts would strongly disagree with this convention. However, a complete agreement in the computer science community on what exactly *concurrency* means, and how it relates to *parallelism,* is lacking. According to *Chapter* §12 of [31], "concurrency in the programming language and parallelism in the computer hardware are independent concepts. [...] We can have concurrency in a programming language without parallel hardware, and we can have parallel execution without concurrency in the language. In short, *concurrency refers to the potential for parallelism*" (italics ours). Clearly, our convention herein does not conform to the notions of concurrency and parallelism as defined in [31]. In contrast, [29] uses the term *concurrent* "to describe computations where the simultaneously executing processes can interact with one another", and *parallel* for "[...] computations where behavior of each process is unaffected by the behavior of the others". [29] also acknowledges that many authors do not discriminate between *'parallel'* and *'concurrent'.* We shall follow this latter convention throughout and, moreover, by a *parallel (concurrent) computation* we shall mean actions of several processing units that are carried out *logically* (if not necessarily *physically*) *simultaneously.* That is, when referring to parallel (or, equivalently, concurrent) computation, we shall always assume *perfect synchrony.*

1.1. Capturing Concurrency via Sequential Interleavings

While our own brains are massively parallel computing devices, we seem to (consciously) think and approach problem-solving rather sequentially. In particular, when designing a parallel algorithm or writing a parallel computer program, we prefer to be able to understand such an algorithm or program at the level of sequential operations or executions. It is not surprising, therefore, that a great deal of research effort has been devoted to interpreting parallel computation in the more familiar, sequential terms. One of the most important contributions in that respect is the (nondeterministic) sequential *interleaving semantics* of concurrency (see, e.g., [10, 12, 18, 22, 23]).

When interpreting concurrency via interleaving semantics, a natural question arises: *Given*

a parallel computing model, can its parallel execution always be captured by such sequential nondeterminism, so that any given parallel computation can be faithfully reproduced via an appropriate choice of a sequential interleaving of the operations involved? For most theoreticians of parallel computing the answer is apparently *"Yes"* - provided that we simulate concurrent execution via sequential interleavings at a sufficiently high level of granularity of the basic computational operations.

We shall illustrate the concept of *sequential interleaving semantics* of concurrency with a simple exercise. Let's consider the following question from a sophomore parallel programming class: *Find an example of two instructions such that, when executed in parallel, they give a result not obtainable from any corresponding sequential execution sequence.*

A possible answer: Assume $x = 0$ initially and consider the following two programs

$$x \leftarrow x + 1;\ x \leftarrow x + 1$$

vs.

$$x \leftarrow x + 1 \ || \ x \leftarrow x + 1$$

where "||" stands for the parallel, and ";" for the sequential composition of instructions or programs, respectively. Sequentially, one *always* gets the same answer: $x = 2$. In parallel (when the two assignment operations are executed synchronously), however, one gets $x = 1$. It appears, therefore, that no sequential ordering of operations can reproduce parallel computation - at least not at the granularity level of high-level instructions as above.

The whole "mystery" can be readily resolved if we look at the possible sequential executions of the corresponding machine instructions:

LOAD x, $*m$	LOAD x, $*m$
ADD x, #1	ADD x, #1
STORE x, $*m$	STORE x, $*m$

There certainly exist choices of *sequential interleavings* of the six machine instructions above that lead to *"parallel"* behavior (i.e., the one where, after the code is executed, $x = 1$). Thus, by refining granularity from the high-level language instructions down to the machine instructions, we can certainly preserve the "interleaving semantics" of concurrency.

As a side, we remark that it turns out that the example above does not require finer granularity quite yet, if we allow that some instructions be treated as no-ops. Indeed, if we informally define $\Phi(P)$ to be the *set of possible behaviors of program P*, then the example above only shows that, for $S_1 = S_2 = (x \leftarrow x + 1)$,

$$\Phi(S_1||S_2) \ \not\subseteq \ \Phi(S_1; S_2) \ \cup \ \Phi(S_2; S_1) \tag{1}$$

However, it turns out that, in this particular example, it indeed is the case that

$$\Phi(S_1||S_2) \ \subseteq \ \Phi(S_1; S_2) \ \cup \ \Phi(S_2; S_1) \ \cup \ \Phi(S_1) \ \cup \ \Phi(S_2) \tag{2}$$

and no finer granularity is necessary to model $\Phi(S_1||S_2)$, assuming that, in some of the sequential interleavings, we allow certain instructions not to be executed at all.

However, one can construct more elaborate examples where the property (2) does not hold. The only way to capture the program behavior of parallel compositions of the form $\Phi(P_1||P_2)$ via sequential interleavings in such cases would then be to find a finer level of granularity.

We address in this work the (in)adequacy of the sequential interleavings semantics when applied to CA where the individual node updates are considered to be elementary operations. It is tacitly assumed that the complete node update operation includes, in addition to computing the local update function on appropriate inputs, also the necessary *reads* of the neighbors' values

preceding the local rule computation, as well as the *writes* of one's new value following the local computation. These points will become clear once the necessary definitions and terminology are introduced in Section 2; see also discussion in Sections 4 and 5.

In particular, we will show that the perfect synchrony of the classical CA's node updates causes the interleaving semantics, as captured by the SCA and NICA sequential CA models (see Section 2), to fail rather dramatically even in the context of the simplest *nonadditive* CA node update rules.

2. Cellular Automata and Types of Their Configurations

We will follow [11] and formally define classical (that is, synchronous and concurrent) CA in two steps: we shall first introduce the notion of a *cellular space,* and then define a *cellular automaton* over an appropriate cellular space.

Definition 1 *A Cellular Space,* Γ*, is an ordered pair* (G, Q)*, where*

- G *is a regular undirected Cayley graph that may be finite or infinite, with each node labeled with a distinct integer; and*
- Q *is a finite set of states that has at least two elements, one of which being the special quiescent state, denoted by* 0*.*

We denote the set of integer labels of the nodes (vertices) in Γ by L. That is, L may be equal to, or be a proper subset of, the set of all integers.

Definition 2 *A Cellular Automaton* A *is an ordered triple* (Γ, N, M)*, where*

- Γ *is a cellular space;*
- N *is a fundamental neighborhood; and*
- M *is a finite state machine such that the input alphabet of* M *is* $Q^{|N|}$*, and the local transition function (update rule) for each node is of the form* $\delta : Q^{|N|+1} \rightarrow Q$ *for CA with memory, and* $\delta : Q^{|N|} \rightarrow Q$ *for memoryless CA.*

The fundamental neighborhood N specifies which of the near-by nodes provide inputs to the update rule of a given node. In the classical CA, Γ is a regular graph that locally "looks the same everywhere"; in particular, the local neighborhood N is the same for each node in Γ.

The local transition rule δ specifies how each node updates its state (that is, value), based on its current state and the current states of its neighbors in N. By composing together the application of the local transition rule to each of the CA's nodes, we obtain *the global map* on the set of (global) configurations of a cellular automaton.

Insofar as the CA "computer architecture" is concerned, one important characteristic is that the memory and the processors are not truly distinguishable, in stark contrast to *Turing machines, (P)RAMs*, and other standard abstract models of digital computers. Namely, each node of a cellular automaton is both a processing unit and a memory storage unit; see, e.g., the detailed discussion in [33]. In particular, the only *memory content* of a CA is a tuple of the current states of all its nodes. Moreover, as a node can "read" (but not "write") the states of its neighbors, we can view the architecture of classical CA as a very simplistic, special case of *distributed shared memory* parallel model, where every *processor* (that is, each node) "owns" one cell (typically, one bit) of its *local memory,* physically separated from other similar *local memories* - yet this local memory is *directly accessible* (for *read* accesses) to some of the other *processors.* In particular, the "reads" from any *memory cell* are restricted to an appropriate neighborhood of that shared value's *owner processor,* while the "writes" are restricted to the owner processor *alone.*

Since our main results in this paper will pertain to a comparison and contrast between the

classical, concurrent threshold CA and their sequential counterparts, we formally introduce two types of the sequential CA next. First, we define SCA with a *fixed* (but arbitrary) sequence specifying the order according to which the nodes are to update. We then introduce a kind of sequential automata whose purpose is to capture the "interleaving semantics", that is, where *all* possible sequences of node updates are considered at once.

Definition 3 *A Sequential Cellular Automaton (SCA)* S *is an ordered quadruple* (Γ, N, M, s), *where* Γ, N *and* M *are as in* Definition 2, *and* s *is an arbitrary sequence, finite or infinite, all of whose elements are drawn from the set* L *of integers used in labelling the vertices of* Γ. *The sequence* s *is specifying the sequential ordering according to which an SCA's nodes update their states, one at a time.*

However, when comparing and contrasting the concurrent CA with their sequential counterparts, rather than making a comparison between a given CA with a *particular* SCA (that is, some concrete SCA with a particular choice of the update sequence s), we would like to compare the parallel CA computations with the computations of the corresponding SCA for *all* possible sequences of node updates. For that purpose, the following class of nondeterministic sequential cellular automata is introduced:

Definition 4 *A Nondeterministic Interleavings Cellular Automaton (NICA)* I *is defined to be the union of all sequential automata* S *whose first three components,* Γ, N *and* M *are fixed. That is,* $I = \cup_s (\Gamma, N, M, s)$, *where the meanings of* Γ, N, M, *and* s *are the same as before, and the union is taken over* all *(finite and infinite) sequences* $s : \{1, 2, 3, ...\} \to L$, *where* L *is the set of integer labels of the nodes in* Γ.

We next introduce some terminology from physics that we find useful for characterizing *all possible computations* of a parallel or a sequential cellular automaton. To this end, a *(discrete) dynamical system* view of CA is helpful. A *phase space* (also called *configuration space*) of a dynamical system is a directed graph where the vertices are the *global configurations* (or *global states*) of the system, and directed edges correspond to the possible direct transitions from one global state to another.

As for any other kind of dynamical systems, we can define the fundamental, qualitatively distinct types of global configurations that a cellular automaton can find itself in. We first define these qualitatively distinct types of dynamical system configurations for the parallel CA, and then discuss how these definitions need to be modified in the case of SCA and NICA.

The classification below is based on answering the following question: starting from a given global configuration, can a cellular automaton return to that same configuration after a finite number of parallel computational steps?

Definition 5 *A fixed point (FP) is a configuration in the phase space of a CA such that, once the CA reaches this configuration, it stays there forever. A cycle configuration (CC) is a state that, once reached, will be revisited infinitely often with a fixed, finite temporal period of* 2 *or greater. A transient configuration (TC) is a state that, once reached, is never going to be revisited again.*

In particular, a FP is a special, degenerate case of a recurrent state with period 1. Due to deterministic evolution, any configuration of a classical, parallel CA is either a FP, a *proper* CC, or a TC. Throughout, we shall make a clear distinction between FPs and proper CCs.

On the other hand, if one considers SCA and NICA, so that *arbitrary* node update orderings are permitted, then, given the underlying cellular space and the local update rule, the resulting phase space configurations, due to the nondeterminism that results from different choices of possible sequences of node updates ("sequential interleavings"), are more complicated. In a

particular SCA, a cycle configuration is any configuration revisited infinitely often - but the period between different consecutive visits, assuming an arbitrary sequence s of node updates, need not be fixed. We call a global configuration that is revisited only finitely many times (under a given ordering s) *quasi-cyclic.* Similarly, a *quasi-fixed point* is an SCA configuration such that, once the SCA's dynamic evolution reaches this configuration, it stays there "for a while", and then leaves. For example, a configuration of an SCA can simultaneously be both an FP and a quasi-CC, or both a quasi-FP and a CC (see the related discussion and examples in [34])).

For simplicity, heretofore we shall refer to a configuration C of a NICA as a *(weak) fixed point* if there exists some infinite sequence of node updates s such that C is a FP in the usual sense when the automaton's nodes update according to the ordering s. A *strong fixed point* of a NICA automaton is a configuration that is fixed (stable) with respect to *all* possible sequences of node updates. Similarly, we consider a configuration C' to be a cycle state, if there exists an infinite sequence of node updates s' such that, if NICA nodes update according to s', then C' is a cycle state of period 2 or greater in the usual sense (see *Definition* 5). In particular, in case of the NICA automata, a single configuration can simultaneously be a weak FP, a CC and a TC; see reference [34] for a simple example.

3. 1D Parallel vs. Sequential CA Comparison for Simple Threshold Rules

After the introduction, motivation and the necessary definitions, we now proceed with our main results and their meaning. Technical results are given in this section. Discussion of the implications and relevance of these results, as well as some possible generalizations and extensions, will follow in Section 4.

We will compare and contrast the classical, parallel CA with their sequential counterparts, SCA and NICA, in the context of the simplest nonlinear and nonaffine local update rules possible, namely, the CA in which the nodes locally update according to *linear threshold functions.* Moreover, we choose these threshold functions to be *symmetric,* so that the resulting (S)CA are also *totalistic* (see, e.g., [11] or [38]). We show the fundamental difference in the configuration spaces, and therefore possible computations, between the parallel threshold CA and the sequential threshold CA: while the former can have temporal cycles (of length two), the computations of the latter always either converge to a fixed point, or otherwise fail to finitely converge to any recurrent pattern whatsoever.

For simplicity, but also in order to indicate how dramatically the sequential interleavings of NICA fail to capture the concurrency of the perfectly synchronous CA, we restrict the underlying cellular spaces to *one-dimensional* Γ. We formally define the class of 1D (S)CA of a finite radius below:

Definition 6 *A* 1D (sequential) cellular automaton *of radius* $r \geq 1$ *is a* (S)CA *defined over a one-dimensional string of nodes, such that each node's next state depends on the current states of its neighbors to the left and to the right that are no more than* r *nodes away. For the* (S)CA with memory, *the next state of a node also depends on that node's own current state.*

Thus, for the Boolean (S)CA *with memory* defined over a one-dimensional cellular space Γ, each node's next state depends on exactly $2r + 1$ input bits, while in the *memoryless* (S)CA case, the local update rule is a function of $2r$ input bits. The underlying 1D cellular space is a string of nodes that can be a finite line graph, a ring (corresponding to the "circular boundary conditions"), a one-way infinite string, or, in the most common case, Γ is a two-way infinite string (or "line").

We fix the following conventions and terminology. Throughout, only *Boolean* CA, SCA and NICA are considered; in particular, the set of possible states of any node is $\{0, 1\}$. The phrases

"monotone symmetric" and "symmetric (linear) threshold" functions / update rules / automata are used interchangeably. Similarly, "(global) dynamics" and "(global) computation" are used synonymously. Unless stated otherwise, *CA* denotes a classical, parallel cellular automaton, whereas a cellular automaton where the nodes update sequentially is always denoted by *SCA* (or *NICA,* when appropriate). Also, unless explicitly stated otherwise, (S)CA *with memory* are assumed. The default infinite cellular space Γ is a two-way infinite line. The default finite cellular spaces are finite rings. The terms "phase space" and "configuration space" are used synonymously throughout, as well, and sometimes abridged to *PS* for brevity.

3.1. On the Existence of Cycles in Parallel and Sequential Threshold Cellular Automata

We begin by defining *linear threshold functions* and *simple threshold functions,* and the CA, SCA and NICA whose updates rules are restricted to such threshold functions.

Definition 7 *A Boolean-valued linear threshold function of m inputs, $x_1, ..., x_m$, is any function of the form*

$$f(x_1, ..., x_m) = \begin{cases} 1, & \text{if } \sum_i w_i \cdot x_i \geq \theta \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where θ is an appropriate threshold constant, and $w_1, ..., w_m$ are arbitrary (but fixed) real numbers, called weights.

In general, weights w_i in the definition above can be both positive and negative. This is esp. common in the neural networks literature, where negative weights w_i indicate an *inhibitory effect* of, e.g., one neuron on the firings of another, near-by neuron. In most studies of discrete dynamical systems, however, the weights w_i are required to be nonnegative - that is, only *excitatory effects* of a node on its neighbors are allowed; see, e.g., [5, 6, 34, 36, 37].

Definition 8 *A* threshold automaton (threshold (S)CA) *is a (parallel or sequential) cellular automaton whose node update rule δ is a Boolean-valued linear threshold function.*

Therefore, given an integer $k \geq 0$, a k-threshold function, in general, is any function of the form as in *Definition* 8 with $\theta = k$ and an appropriate choice of weights w_i, $i = 1, ..., m$. Heretofore we consider *monotonically nondecreasing* Boolean threshold functions only; this, in particular, implies that the weights w_i are always nonnegative. We also additionally assume δ to be a *symmetric function* of all of its inputs. That is, the (S)CA we analyze have *symmetric, monotone Boolean functions* for their local update rules. We refer to such functions as to *simple threshold functions*, and to the (S)CA with simple threshold node update rules as to *simple threshold (S)CA* [5, 6, 34].

Definition 9 *A simple threshold (S)CA is a cellular automaton whose local update rule δ is a monotone symmetric Boolean (threshold) function.*

In particular, if all the weights w_i are positive and equal to one another, then, without loss of generality, we may set them all equal to 1; obviously, this normalization of the weights w_j may also require an appropriate adjustment of the threshold value θ.

Throughout, whenever we say a *threshold automaton* or a *threshold (S)CA,* we shall mean *a simple threshold automaton (threshold (S)CA),* unless explicitly stated otherwise. That is, the 1D threshold (S)CA studied in the sequel will have the node update functions of the general form

$$\delta(x_{i-r}, x_{i-r+1}, ..., x_i, ..., x_{i+r-1}, x_{i+r}) \;=\; \begin{cases} 1, & \text{if } \sum_{j=-r}^{r} x_{i+j} \;\geq\; k \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where k is a fixed integer from the range $\{0, 1, ..., 2r+1, 2r+2\}$.

Due to the nature of the node update rules, cyclic behavior intuitively should not be expected in such simple threshold cellular automata. This is, generally, (almost) the case, as will be shown below. We argue that the importance of the results in this subsection largely stems from the following three factors:

- the local update rules are the simplest nonlinear totalistic rules one can think of;
- given the rules, the cycles are not to be expected - yet they do exist, and in the case of synchronous parallel CA *only;* and, related to that observation,
- it is, for this class of automata, the parallel CA that have more diverse possible dynamics than any of their sequential counterparts, and, in particular, while qualitatively there is nothing among the possible sequential computations that is not present in the parallel case, the parallel threshold CA do exhibit a particular qualitative behavior that cannot be reproduced by any threshold SCA or NICA.

The results that follow hold for two-way infinite 1D (S)CA, as well as for finite (S)CA with the circular boundary conditions (i.e., for the (S)CA whose cellular spaces are finite rings).

Lemma 1 *The following dichotomy holds for (S)CA with $\delta = MAJ$ and $r = 1$:*

(i) Any 1D parallel CA with $r = 1$, the MAJORITY *update rule $\delta = MAJ$, and an even number of nodes, has finite temporal cycles in the phase space (PS); the same holds for the two-way infinite* 1D MAJORITY CA.

(ii) 1D Sequential CA with $r = 1$ and $\delta = MAJ$ do not have any temporal cycles in their phase spaces, irrespective of the sequential node update ordering s.

Proof. To show (i), we exhibit an actual two-cycle. Consider either an infinite 1D CA, or a finite one, with circular boundary conditions and an even number of nodes, $2n$. Then the configurations $(10)^\omega$ and $(01)^\omega$ in the infinite case ($(10)^n$ and $(01)^n$ in the finite ring case) form a 2-cycle. A proof of part (ii) based on a straight-forward case analysis can be found in our prior publication [34]. □

We remark that, insofar as the infinite SCA as in *Lemma* 1 are concerned, a nontrivial temporal cycle configuration cannot exist even in the limit.

Part (ii) of *Lemma* 1 above can be readily generalized: even if we consider local update rules δ other than the MAJORITY rule, yet restrict δ to *monotone symmetric (Boolean) functions,* that is, the simple threshold functions [5, 6, 34], such sequential CA still do not have any proper cycles.

Theorem 1 *For any Simple Threshold 1D Sequential CA A with $r = 1$, and any sequence s of the node updates, the phase space PS(A) is temporal cycle-free.*

Proof. Since $r = 1$ and $2r + 1 = 3$, there are only *five* simple threshold functions on three inputs. Two of those five functions are utterly trivial (the constant functions 0 and 1). The "at-least-1-out-of-3" simple threshold function is the Boolean *OR* on three inputs; similarly, the "at-least-3-out-of-3" simple threshold function is the Boolean *AND*. It is straight-forward to show that the CA (sequential or parallel, as long as they are *with memory*) with $\delta \in \{OR, AND\}$ cannot have temporal cycles. The only remaining simple threshold update rule on three inputs is $\delta = MAJ$, for which we have already argued that the corresponding parallel CA have

temporal two-cycles, but all the corresponding SCA (and therefore the NICA) have cycle-free configuration spaces. □

Similar results to those in *Lemma* 1 and *Theorem* 1 hold for the 1D CA with radius $r = 2$:

Lemma 2 *The following dichotomy holds for (S)CA with $\delta = MAJ$ and $r = 2$:*

(i) There are 1D parallel CA *with $r = 2$ and $\delta =$* MAJ *that have finite temporal cycles.*

(ii) Any 1D SCA *with $r = 2$ and $\delta = MAJ$, for any sequential order s of the node updates whatsoever, has a cycle-free configuration space.*

The proof of this Lemma can be found in our earlier work [34].

Generalizing Lemmata 1 and 2, part (i), we have the following

Corollary 1 *For all $r \geq 1$, there exists a* monotone symmetric CA *A such that A has finite temporal cycles in the phase space.*

Namely, given any $r \geq 1$, a parallel CA with $\delta = MAJ$ and $\Gamma =$ *infinite line* has at least one two-cycle in the *PS*: $\{(0^r1^r)^\omega, (1^r0^r)^\omega\}$. If $r \geq 3$ is odd, then such a threshold automaton has at least two distinct two-cycles, since $\{(01)^\omega, (10)^\omega\}$ is also a two-cycle. Analogous results hold for the more general *simple threshold CA* defined on finite 1D cellular spaces, provided that such automata have sufficiently many nodes, that the number of nodes is appropriate (see [35] for more details), and assuming circular boundary conditions. Moreover, the result extends to many finite and infinite CA in the higher dimensions, as well; in particular, *threshold CA* with $\delta = MAJ$ that are defined over the $2D$ Cartesian grids or hypercubes also have two-cycles in their respective phase spaces.

It turns out that the two-cycles in the *PS* of concurrent CA with $\delta = MAJ$ are actually the only type of (proper) temporal cycles such cellular automata can have. Indeed, for any *symmetric linear threshold update rule* δ, and any *finite* regular Cayley graph as the underlying cellular space, the following general result holds [11, 15]:

Proposition 1 *[15] Let a parallel simple threshold* CA *$A = (\Gamma, N, M)$ be given, where Γ is any* finite *cellular space, and let this cellular automaton's* global map *be denoted by F. Then for all configurations $C \in PS(A)$, there exists a finite time step $t \geq 0$ such that $F^{t+2}(C) = F^t(C)$.*

In particular, this result implies that, in case of any *finite* simple threshold CA, for any starting configuration C_0, there are *only two possible kinds of orbits:* upon repeated iteration, the computation either converges to a fixed point configuration after finitely many steps, or else it eventually arrives at a two-cycle.

It is almost immediate that, if we allow the underlying cellular space Γ to be infinite, if computation from a given starting configuration converges after any finite number of steps at all, it will have to converge either to a fixed point or to a two-cycle (but never to a cycle of, say, period three - or any other finite period). The result also extends to finite and infinite SCA, provided that we reasonably define what is meant by a single computational step in a situation where the nodes update one at a time. The simplest notion of a single computational step of an SCA is that of a single node updating its state. Thus, a single parallel step of a classical CA defined on an infinite underlying cellular space Γ includes an infinite amount of sequential computation and, in particular, infinitely many elementary sequential steps. Discussing the implications of this observation, however, is beyond the scope of this work.

Additionally, in order to ensure some sort of convergence of an arbitrary SCA (esp. when the underlying Γ is infinite), and, more generally, in order to ensure that *all the nodes* get a chance to update their states, an appropriate condition that guarantees *fairness* needs to be specified. That is, an appropriate restriction on the allowable sequences s of node updates is

required. As a first step towards that end, we shall allow only *infinite* sequences s of node updates through the rest of the paper.

For SCA defined on finite cellular spaces, one sufficient fairness condition is to impose a fixed upper bound on the number of sequential steps before any given node gets its "turn" to update again. This is the simplest generalization of the fixed permutation assumption made in the work on *Sequential Dynamical Systems* (SDSs); see, e.g., [5, 6, 7, 8]. In the infinite SCA case, on the other hand, the issue of fairness is nontrivial, and some form of *dove-tailing* of sequential individual node updates may need to be imposed. In the sequel, we shall require from the sequences s of node updates of SCA and NICA to be fair in a sense defined below, without imposing any further restrictions or investigating how are such fair sequences of node updates to be generated in a distributed setting. For our present purposes, the following simple notion of fairness will suffice:

Definition 10 *An infinite sequence* $s : N \to L$ *is* fair *if (i) the domain* L *is finite or countably infinite, and (ii) every element* $x \in L$ *appears infinitely often in the sequence of values* $s_1 = s(1),\ s_2 = s(2), s_3 = s(3), ...$

Let $s : N \to L$ be an arbitrary infinite sequence of elements from some domain L. Let $s^{[q]}$ denote the q-tail of s, i.e., $s^{[q]} = (s_{q+1}, s_{q+2}, s_{q+3}, ...)$, We state the following alternative characterizations of fair sequences:

Lemma 3 *Let an infinite sequence* $s : N \to L$ *be given, where the set* L *is countable. Then the following four properties are all equivalent to one another:*

(i) s *is* fair;

(ii) $\forall n \in N,\ s^{[n]}$ *is* fair;

(iii) $(\forall x \in L)(\forall n \in N)(\exists n' \in N)(n' > n\ \wedge\ s(n') = x)$

(iv) $\forall n \in N,\ \ s^{[n]} : \{n+1, n+2, ...\} \to L$ *is* onto.

Now that we have defined what we mean by a *single step* of a sequential CA, as well as adopted some reasonable notion of *fairness,* we have set the stage for the following generalization of *Proposition* 1 to both finite and infinite 1D CA and 1D SCA:

Proposition 2 *Let a parallel CA or a sequential SCA be defined over a finite or infinite 1D cellular space, with a finite rule radius* $r \geq 1$. *Let this cellular automaton's local update rule be a simple threshold function. Let's also assume, in the sequential cases, that the fairness condition from* Definition 10 *holds. Then for any starting configuration* $C_0 \in PS(A)$ *whatsoever, and any finite subconfiguration* $C \subseteq C_0$, *there exists a time step* $t \geq 0$ *such that*

$$F^{t+2}(C) = F^t(C) \tag{5}$$

where, in the case of fair SCA, the Eqn. (5) *can be replaced with*

$$F^{t+1}(C) = F^t(C) \tag{6}$$

In the case of $\delta = MAJ$ (S)CA, a computation starting from any *finitely supported*[a] initial configuration necessarily converges to either a FP or a two-cycle [15]:

Proposition 3 *Let the assumptions from* Proposition 2 *hold, and let the underlying threshold rule be* $\delta = MAJ$. *Then for all configurations* $C \in PS(A)$ *whatsoever in the finite cases, and for all configurations* $C \in PS(A)$ *such that* C *has a* finite support *when* $\Gamma(A)$ *is infinite, there exists a finite time step* $t \geq 0$ *such that* $F^{t+2}(C) = F^t(C)$. *Moreover, in the sequential cases with* fair *update sequences, there exists a finite* $t \geq 0$ *such that* $F^{t+1}(C) = F^t(C)$.

[a] Also sometimes called *compactly supported;* see, e.g., [15]. A global configuration of a cellular automaton defined over an infinite cellular space Γ is said to be *compactly supported* if all except for at most finitely many of the nodes are *quiescent* (i.e., in state 0) in that configuration.

Furthermore, if *arbitrary* infinite initial configurations are allowed in *Propositions* $2-3$, and the dynamic evolution of the full such global states is monitored, then the only additional possibility is that the particular (S)CA computation fails to finitely converge altogether. In that case, and under the fairness assumption in the case of SCA, the limiting configuration $lim_{t\to\infty}F^t(C) = C^{lim}$ can be shown to be a fixed point.

To summarize, if the computation of a SCA starting from some configuration C converges at all (that is, to *any* finite recurrent structure), it actually has to converge to a fixed point.

To convince oneself of the validity of *Proposition* 2, two basic facts have to be established. One, convergence to finite temporal cycles of length three or higher is not possible. Indeed, *Proposition* 1 establishes that the only possible long-term behaviors of the finite simple threshold CA are (i) the convergence to a fixed point and (ii) the convergence to a two-cycle. The only possibility for fair finite SCA is the convergence to a fixed point. If infinite cellular spaces are considered, it is straight-forward to see that the only new possibility is that the long-term dynamics of a (S)CA fails to (finitely) converge altogether. In some cases with infinite Γ such divergence indeed takes place - even when the starting configuration is finitely (compactly) supported: consider, e.g., the $\delta = OR$ CA and the starting configuration ...00100... on the two-way infinite line.

Two, in the sequential cases (that is, for the simple threshold SCA and NICA), temporal two-cycles are not possible. That is, a generalization of *Lemmata* 1, 2 and *Theorem* 1 to arbitrary finite $r \geq 1$, and arbitrary symmetric threshold update rules, holds. This generalization is provided by an appropriate specialization of a similar result in [6] for a class of sequential graph automata called *Sequential Dynamical Systems* that we have already mentioned. In particular, part (ii) in the *Theorem* 2 below and its proof are directly based on [6]:

Theorem 2 *The following dichotomy holds:*

(i) All 1D (parallel) CA *with any odd* $r \geq 1$, *the local rule* $\delta =$ MAJ, *and cellular space* Γ *that is a finite ring with an even number of nodes, or a two-way infinite line, have finite cycles in their phase spaces. The same holds for* arbitrary *(even or odd)* $r \geq 1$ *provided that* Γ *is either a finite ring with a number of nodes divisible by* $2r$, *or a two-way infinite line.*

(ii) Any 1D SCA *with* any *monotone symmetric Boolean update rule* δ, *for any finite* $r \geq 1$, *defined over a finite or infinite 1D cellular space, and for* an arbitrary *sequence* s *(finite or infinite, fair or unfair) as the node update ordering, has a cycle-free phase space.*

Remark: There are also CA defined over finite rings and with even $r \geq 2$ such that the number of nodes in those rings is not divisible by $2r$ yet temporal two-cycles exist. However, a more detailed discussion on what properties the number of nodes in such CA has to satisfy is required; we leave this discussion out, however, for the sake of clarity and space constraints.

Proof.

Part (i): Consider first $\Gamma =$ *infinite line.* For the special case when $r = 2$, consider the configurations $(1100)^\omega$ and $(0011)^\omega$; it is easy to verify that these two configurations form a cycle for the corresponding parallel CA. Similar reasoning readily generalizes to arbitrary $r \geq 2$. Thus, the "canonical" temporal two-cycle for 1D MAJORITY CA defined over an infinite line with $r \geq 1$ is $\{(1^r0^r)^\omega, (0^r1^r)^\omega\}$, with the obvious modification for the finite CA with $2n$ nodes (and assuming the circular boundary conditions).

Part (ii) (proof sketch): The proof of this interesting property is based on a slight modification of a similar result in [6] for the aforementioned SDSs. A *simple symmetric SDS* is a sequentially updating graph automaton with (possibly different) k-threshold update rules at different nodes, and with the node update ordering given by a *fixed* permutation of the nodes. The central idea of the proof is to assign nonnegative integer *potentials* to both nodes and

edges in the *functional graph* of the given SCA. In this functional graph, for any two nodes x_i and x_j, the unordered pair $\{x_i, x_j\}$ is an edge if and only if these two nodes provide inputs to one another, i.e., in the 1D SCA case, if and only if *distance*$(x_i, x_j) \leq r$ (that is, assuming the canonical labelling of the nodes, so that the consecutive nodes always get labeled by the consecutive integers, iff $|i - j| \leq r$). The potentials are assigned in such a way that, each time a node changes its value from 0 to 1 or vice versa, the overall potential of the resulting configuration is strictly less than the overall potential of the configuration before the node flip. Since all individual node and edge potentials are initially nonnegative, and since the total potential of any configuration (that is, the sum of all individual node and edge potentials in this configuration) is always bounded, the fact that each "flip" of any node's value strictly decreases the overall potential by integer amounts implies that, after a finite number of node flips, an equilibrium where no nodes can further flip is reached; this equilibrium will be a fixed point configuration. □

To summarize, simple threshold CA, depending on the starting configuration, may converge to a fixed point or to a temporal two-cycle; in particular, they may end up "looping" in finite (but nontrivial) temporal cycles. In contrast, the corresponding classes of SCA (and therefore NICA) can never cycle. We also observe that, given any sequence of node updates of a finite threshold SCA, if this sequence satisfies the fairness condition from *Definition* 10, then it can be shown that the computation of such a threshold SCA A is guaranteed to converge to a fixed point (sub)configuration on any finite subset of the nodes in $\Gamma(A)$.

The cycle-freeness of the threshold SCA and NICA holds irrespective of the choice of a sequential update ordering; moreover, extending to infinite SCA, temporal cycles cannot be obtained even *in the limit.* Hence, we conclude that no choice of a "sequential interleaving" can capture the perfectly synchronous parallel computation of simple threshold CA. Consequently, the "interleaving semantics" of NICA fails to capture the synchronous parallel behavior of the classical CA even for this, simplest nonlinear class of totalistic CA update rules.

3.2. Characterizing Configuration Spaces of 1D (S)CA *with* $\delta =$ MAJORITY

Next, we specifically focus on $\delta =$ MAJORITY 1D CA, and characterize the configuration spaces of such threshold cellular automata. In particular, in the $\Gamma =$ *infinite line* case, we show that the cycle configurations are rather rare, that fixed point configurations are quite numerous (there are uncountably many of them) yet still relatively rare in a precise mathematical sense to be discussed below, and that *almost all* configurations of these threshold (S)CA are transient.

In the sequel, for the finite 1D (S)CA circular boundary conditions will be assumed by default. Thus, the cellular spaces in this section will be either infinite lines, or finite rings. In the case of SCA, the fairness condition based on *Definition* 10 will be assumed. Also, when we refer to FPs of NICA, we mean *weak* fixed points. (Recall that a weak FP is a configuration such that there exists an infinite sequence of individual node updates that satisfies the fairness condition and so that, with respect to this sequence, the particular configuration is a "proper" FP - but the same configuration may be, for example, a proper TC with respect to other sequences of node updates).

We begin with some observations about the nature of various configurations in the (S)CA with $\delta = MAJ$ and $r = 1$. We shall subsequently generalize several of these results to arbitrary $r \geq 1$. We first recall that, for such (S)CA with $r = 1$, two adjacent nodes of the same value are stable. That is, 11 and 00 are stable subconfigurations. Consider now the starting subconfiguration $x_{i-1}x_ix_{i+1} = 101$. In the parallel case, at the next time step, $x_i \to 1$. Hence, no FP configuration of a parallel CA can contain 101 as a subconfiguration. In the sequential

case, assuming fairness, x_i will eventually have to update. If, at that time, it is still the case that $x_{i-1} = x_{i+1} = 1$, then $x_i \to 1$, and $x_{i-1}x_ix_{i+1} \to 111$, which is stable. Else, at least one of x_{i-1}, x_{i+1} has already "flipped" into 0. Without loss of generality, let's assume $x_{i-1} = 0$. Then $x_{i-1}x_i = 00$, which is stable; so, in particular, $x_{i-1}x_ix_{i+1}$ will never go back to the original 101. By symmetry of $\delta = MAJ$ with respect to 0 and 1, the same analysis applies to the subconfiguration $x_{i-1}x_ix_{i+1} = 010$. In particular, the following properties hold:

Lemma 4 *A fixed point configuration of a 1D-(S)CA with $\delta = MAJ$ and $r = 1$ cannot contain subconfigurations* 101 *or* 010*. Similarly, a cycle configuration of such a 1D-(S)CA cannot contain subconfigurations* 00 *or* 11*.*

Proof. In any configuration that contains 101 as a subconfiguration, at the very next parallel update the 0 in between the two 1s will flip to 1, regardless of how many other nodes are present, and what are their current states. Analogous argument applies to configurations that contain 010. Hence, FPs of CA with $\delta = MAJ$ and $r = 1$ are solely made of consecutive blocks of two or more 1s and/or similar blocks of two or more 0s.

As for the claim about the cycle configurations, notice that 00 and 11 are stable subconfigurations. Without loss of generality, assume a CC of a parallel MAJORITY CA contains a block of two or more consecutive 0s. Consider, say, the node adjacent to the rightmost 0 in that block. Let's denote that node by x_j. This node x_j is, by assumption, in the state 1. There are two cases to consider. If our node's right neighbor, x_{j+1}, is in the state 0, then, at the very next parallel step, the node x_j, that has at least two out of three of its input bits equal to 0, will itself flip to 0. Since x_j is adjacent to $x_{j+1} = 0$, it will have joined an expanded stable block of zeros, and consequently x_j will remain at 0 thereafter. Hence, the starting configuration cannot be a CC.

The other possibility is that x_{j+1} is also in the state 1; that is, the configuration is of the form $...x_{j-2}x_{j-1}x_jx_{j+1} = ...0011....$ Then the block of (at least) two consecutive 1s is stable, and so is the block of two or more 0s to the left from it. That such a configuration cannot be a cycle state now follows by induction: proceeding moving along the line (or ring) of nodes from the assumed block of zeros to the right, either eventually a block of the form ...010 or ...101 is encountered, in which case this configuration is transient, or else no such subconfiguration exists, in which case the entire configuration is made solely of the stable blocks of two or more 0s and similar stable blocks of 1s, and thus this configuration must be a fixed point. Either way, the assumption that this configuration was actually a cycle configuration cannot hold. □

Lemma 5 *The FPs of the 1D-(S)CA with $\delta = MAJ$ and $r = 1$ are precisely of the form $(000^* + 111^*)^*$. The CCs of such 1D-CA may exist only in the parallel case, and the temporal cycles are precisely of the form $\{(10)^*, (01)^*\}$. The TCs of CA are all other configurations, that is, precisely the configurations that contain both (i)* 000^* *or* 111^* *(or both), and (ii)* 101 *or* 010 *(or both) as their subconfigurations. In addition, the CCs in the parallel case become TCs in all corresponding sequential cases.*

The claim of Lemma 5 follows by the result of Lemma 4 and an elementary case analysis.

We observe that, for $r \geq 2$, there exist cycle configurations that actually contain *stable subconfigurations.* Similarly, there exist FPs that are characterized by *spatial periodicity,* and are not entirely made of the consecutive stable blocks of 0s and/or 1s. Likewise, giving a similar characterization for the higher dimensional cellular spaces is also not as straight-forward as the results in the Lemmata above. Therefore, the nice and clean partition of the configurations into FPs, CCs and TCs obtained in this section is attributable to the peculiarity of the 1D cellular spaces, as well as the assumption that the rule radius is $r = 1$.

However, some generalizations to arbitrary (finite) rule radii r can be readily deducted.

For instance, given any such $r \geq 1$, the finite subconfigurations 0^{r+1} and 1^{r+1} are stable with respect to $\delta = MAJ$ update rule applied either in parallel or sequentially; consequently, any configuration of the form $(0^{r+1}0^* + 1^{r+1}1^*)^*$, for a finite or infinite CA with an appropriate number of nodes, is a fixed point. This characterization, only with a considerably different notation, has been known for the case of configurations with *compact support* for a relatively long time; see, e.g., Chapter 4 in [15]. On the other hand, fully characterizing CCs (and, consequently, also TCs) in case of finite or infinite (parallel) CA is more complicated than in the simplest case with $r = 1$. For example, for $r \geq 1$ odd, $\{(10)^*, (01)^*\}$ is a two-cycle, whereas for $r \geq 2$ even, each of $(10)^*$, $(01)^*$ is a fixed point. However, for all $r \geq 1$, the corresponding (parallel) CA are guaranteed to have some temporal cycles, namely, given r, the set of states $\{(1^r0^r)^*, (0^r1^r)^*\}$ forms a two-cycle.

Back to the $r = 1$ case, we also establish the following property:

Lemma 6 *Given any (finite or infinite) simple threshold (S)CA* with memory *and with the rule radius $r = 1$, one of the following two properties always holds:*

(i) this simple threshold (parallel or sequential) cellular automaton does not have any proper temporal cycles and cycle configurations at all; or else

(ii) if there are cycle configurations in the PS of this CA, then none of those cycle configurations has any incoming transients.

Proof. Let's assume a threshold CA with $r = 1$ has a cycle configuration. Then the update rule δ of this CA cannot be either Boolean AND or Boolean OR and, consequently, since $r = 1$, it follows that it must be the case that $\delta = MAJ$.

Now, if any CC of this CA actually had an incoming transient, then there would exist a predecessor configuration of this cycle configuration such that this predecessor configuration is transient. Let C' denote this transient configuration, and let C denote the cycle configuration in question (so that $F(C') = C$ where F, as before, stands for this cellular automaton's global map). By Lemmata 4 and 5, configuration C' must contain both stable and unstable subconfigurations. In particular, C' contains either a stable block of the form 00... or of the form 11...; however, this implies that $F(C')$ also contains such a stable block and, consequently, by Lemma 4, $F(C')$ cannot be a cycle configuration. Therefore, it follows that any predecessor of a cycle configuration cannot be a TC; hence, such a predecessor itself also has to be a cycle configuration, and the claim of the Lemma follows. □

We strongly suspect that the property in *Lemma* 6 actually holds for *arbitrary* rule radii $r \geq 1$, but do not have a proof - nor do we know of a counterexample - as of yet:

Conjecture: Given any (finite or infinite) simple threshold parallel or sequential CA with any rule radius $r \geq 1$, exactly one of the following two properties always holds:

(i) either this simple threshold (S)CA does not have proper cycles and cycle configurations, or else

(ii) if there are cycle configurations in the *PS* of this cellular automaton, then none of those temporal cycles has any incoming transients.

Next, we show that the fixed points of simple threshold cellular automata may be quite numerous when $\delta = MAJ$. Infinite sequential and parallel MAJ CA alike have infinitely many FPs, and this property holds for every rule radius $r \geq 1$. Moreover, the cardinality of the set of FPs, in the case of $\delta = MAJ$ and the (countably) infinite cellular spaces, equals the cardinality of the entire *PS*:

Theorem 3 *An infinite* 1D-(S)CA *with* δ = MAJ *and any* $r \geq 1$ *has* uncountably many *fixed points.*

Proof. For the notational convenience, let us consider one-way infinite (S)CA. Similar proof can be constructed for the usual, two-way infinite (S)CA on the line.

Let us consider FPs of the form $1^{r+k_1}0^{r+k_2}1^{r+k_3}...$ with all k_i being integers such that $k_i \geq 1$ $(i = 1, 2, 3, ...)$. Let a string of $r + k_i$ *consecutive* 1s or 0s as above be mapped into decimal digit $k_i - 1$ $(mod\ 10)$. We now construct a mapping from a subset of the set of all FPs of such an automaton to the real numbers in the unit interval $[0, 1]$. Let the length of the m-th block of consecutive 0s or 1s be denoted by L_m. Then $L_m \geq r + 1$ gets mapped into $L_m - 2$ $(mod\ 10)$. For instance, if $r = 1$, then $11100111110000111111... = 1^30^21^50^411...$ gets mapped to $0.1032...$, and $1^{15}0^{28}1^7...$ gets mapped to $0.365...$, etc.

It is immediate that this mapping is constructed so that it is *onto* the real line unit interval $[0, 1]$, which has uncountable many "points" (real numbers), that is, this interval is of cardinality $2^{\aleph_0}$. Since the set of infinite 1D (S)CA configurations that includes all configurations made of stable blocks only (and no other configurations) is, in general (for arbitrary $r \geq 1$) a (proper) subset of the set of all fixed points of such an infinite (S)CA, it follows that any such (S)CA with $\delta = MAJ$ has at least as many FPs as there are real numbers in the unit interval on the real line. Therefore, an infinite 1D-(S)CA with $\delta = MAJ$ and any $r \geq 1$ has *uncountably many* fixed points. □

The above result is another evidence that "not all threshold (S)CA are born equal". It suffices to consider only 1D, infinite CA to see a rather dramatic difference. Namely, in contrast to the $\delta =$ MAJORITY CA, the CA with memory and with $\delta \in \{OR, AND\}$ (i) do not have any temporal cycles, and (ii) have *exactly two* FPs, namely, 0^ω and 1^ω. Other threshold CA may have temporal cycles, as shown in the previous subsection, but they still have only a finite number of FPs.

We return to the analysis of the *PS* of MAJORITY 1D infinite (S)CA. We have just shown that there are uncountably many FPs of such sequential and parallel automata. However, the FPs are, when compared to the transient states, still but "a few and far in between". To see this, the basics of probability theory are needed. In particular, let's assume that a random global configuration is obtained by picking each bit (i.e., each site's value) to be either 0 or 1 at random, with equal probability, and so that assigning a bit-value to one site is independent of the bit assignment to any of the other sites. Then the following result holds:

Lemma 7 *If a global configuration of an infinite threshold CA is selected at random, that is, by assigning each node's value independently and according to a toss of a fair coin, then, with probability* 1*, this randomly picked configuration will be a transient state.*

Moreover, the "unbiased randomness", while sufficient, is certainly not necessary. In particular, assigning bit values according to outcomes of tossing a coin with a fixed bias also yields transient states being of probability one.

Proposition 4 *Let p be any real number such that $0 < p < 1$, and let the probability of a site in a global configuration of an infinite simple threshold (S)CA being in the state* 1 *be equal to p. If a global configuration of this threshold automaton is selected at random according to p, then, with probability* 1*, this randomly selected configuration will be transient.*

Proof. Since the cellular space is assumed infinite, and since the probability p of a randomly selected node being in state 1 is fixed, and hence bounded away from both 0 and 1, the following properties hold of a randomly selected infinite configuration:

- with probability 1, any finite substring of 0s and 1s appears *somewhere* in this infinite configuration;

- in particular, stable blocks 1^{r+k} and 0^{r+l} (for some $r, l \geq 1$) appear with probability 1 somewhere in the infinite configuration;

- the same holds of any unstable finite subconfiguration, such as, e.g., 10101 or 01010101.

Consequently, since such a randomly selected configuration contains unstable subconfigurations with probability 1, it follows that, with probability 1, it cannot be a fixed point. Moreover, for any fixed integer m, each of the strings of 0s and 1s of length m appear infinitely often in such a random configuration. In particular, with probability 1, somewhere in the configuration, some finite unstable subconfiguration, let's denote it C_u, is squeezed in between some two stable subconfigurations, say, C_{s1} and C_{s2}. Therefore, regardless of whether the nodes update synchronously in parallel, or sequentially according to an arbitrary fair sequence s, the nodes in C_u will eventually get their turn to update, and at least one of those nodes is going to flip, causing it to join the (expanded) stable block, either C_{s1} or C_{s2}. It therefore follows that the overall (infinite) configuration C that contains the concatenation $C_{s1} \cdot C_u \cdot C_{s2}$ as its finite subconfiguration, cannot be a cycle configuration. □

In case of the finite threshold (S)CA, as the number of nodes, N, increases, the fraction of all 2^N global configurations that are TCs also grows. In particular, under the same assumptions as in *Lemma* 7 and *Proposition* 4 above, in the limit, as $N \to \infty$, the probability that a randomly picked configuration, C, is a transient state approaches 1:

$$lim_{N\to\infty}\ Pr(random\ C\ is\ transient)\ =\ 1 \tag{7}$$

Thus, a fairly complete characterization of the configuration spaces of threshold (S)CA over 1D cellular spaces can be given. In particular, in the infinite threshold CA cases, *almost every* configuration is a TC. However, a striking contrast between the MAJORITY CA on the one, and the CA with any other threshold rule on the other hand, remains: the former have uncountably many FPs, whereas all other simple threshold CA can have only finitely many FPs.

4. Discussion and Future Directions

The results in Section 3 show that the very existence of temporal two-cycles in simple threshold CA can be ascribed entirely to the assumption of *perfect synchrony* of the parallel node updates. In the actual engineering, physical or biological systems that can be modeled by CA, however, such perfect synchrony is usually hard to justify. In particular, when CA are applied to modeling of various complex physical or biological phenomena (such as, e.g., crystal growth, forest fire propagation, information or gossip diffusion in a population, or signal propagation in an organism's neural system), one ought to primarily focus on the underlying CA behaviors that are, in some sense, *robust.* This robustness may require, for instance, a *low sensitivity to small perturbations* in the initial configuration. From this standpoint, temporal cycles in the parallel threshold CA are, indeed, an idiosyncrasy of the perfect synchrony, that is, a peculiarity that is anything but robust. Likewise, it makes sense to focus one's qualitative study of the dynamical systems modeled by the threshold CA to those properties that are *statistically robust* [3]. It can be readily argued in a rigorous, probabilistic sense that the typical, statistically robust behavior of threshold (S)CA computations is a relatively short transient chain, followed by convergence to a fixed point. In particular, the non-fixed-point temporal cycles of the threshold CA with $r = 1$, for example, not only lack any nontrivial basins of attraction (in terms of the incoming transient 'tails'), but are themselves statistically negligible for all sufficiently large finite, as well as for all infinite CA.

We now briefly discuss some possible extensions of the results presented thus far. In particular, we are considering extending our study to *non-homogeneous threshold CA,* where not all the nodes necessarily update according to *one and the same* threshold update rule. Likewise, we are interested in exploring the implications of heterogeneity of the underlying network

structure, that is, what are the effects on the possible behaviors of threshold automata whose underlying cellular spaces are not necessarily one-dimensional, or, more generally, not even regular graphs. We remark that the two particular classes of graph automata defined over arbitrary (not necessarily regular, or Cayley) *finite* graphs, namely, the sequential and synchronous dynamical systems (SDSs and SyDSs, respectively), and their various phase space properties, have been extensively studied; see, e.g., [5, 6, 8, 30] and references therein.

Another future direction is to consider other communication models in cellular automata. We argue that the classical parallel CA can be viewed, if one is interested in node-to-node interactions among the nodes that are not close to one another, as a class of computational models of *bounded asynchrony.* Namely, if nodes x and y are at distance k (i.e., k nodes apart from each other), and the radius of the CA update rule δ is r, then any change in the state of y can affect the state of x *no sooner,* but also *no later* than after about $\frac{k}{r}$ (parallel node update) computational steps.

In the most general setting, we would like to consider various types of *asynchronous cellular and graph automata,* where the nodes are not assumed any longer to update in unison and, moreover, where no global clock is assumed. We again emphasize that such cellular automata would entail what can be viewed as *communication asynchrony,* thus going beyond the kind of asynchrony in local computations at different nodes *only,* that has been studied since at least 1984 [19, 20].

What would, then, such *genuinely asynchronous* CA be like? How do we specify the local update rules, that is, computations at different nodes, given the possible "communication delays" in what was originally a multiprocessor-like, rather than a distributed system-like, parallel model? In the classical, parallel case where a perfect communication synchrony is assumed, any given node x_i of a 1D CA of radius $r \geq 1$ updates according to

$$x_i^{t+1} = f(x_i^t, x_{i_1}^t, ..., x_{i_{2r}}^t) \tag{8}$$

for an appropriate local update rule $\delta = f(x_i, x_{i_1}, ..., x_{i_{2r}})$, whereas, in the asynchronous case, the individual nodes would update according to

$$x_i^{t+1} = f(x_i^t, x_{i_1}^{t_1}, ..., x_{i_{2r}}^{t_{2r}}) \tag{9}$$

We observe that t in *Eqn.* (8) pertains to *the global time,* which of course in this case also coincides with the node x_i's (and everyone else's) *local time.* However, in case of *Eqn.* (9), each t_j (for $j \in \{1, ..., 2r\}$) pertains to an appropriate *local time,* in the sense that each $x_{i_j}^{t_j}$ denotes the node x_{i_j}'s value that was most recently received by the node x_i. That is, $x_{i_j}^{t_j}$ is a local view of the node x_{i_j}'s state, as seen by the node x_i. Thus, the nonexistence of a global clock has considerable implications. The challenge arises, how to meaningfully relate these different local times, so that one can still mathematically analyze such ACA - yet without making the model's description too complicated, that is, while staying away from introducing the explicit *sends* and *receives,* message buffers, etc.? Yet, if we want to study *genuinely* asynchronous CA models (rather than arbitrary sequential models with global clocks), changes along the indicated general lines in the definition of the node update rules seem unavoidable.

We point out that this, genuine (that is, communication) asynchrony in CA (see *Eqn.* (9)) can also be readily interpreted in the nondeterministic terms: at each time step, a particular node updates by using its own current value, and also nondeterministically choosing the current or one of the past values of its neighbors. Such a "past value" of a node x_{i_j} used by the node x_i would be only required not to be any "older" than the particular value of x_{i_j} that x_i

had used as its input on its most recent previous turn to update. That is, insofar as what are the current inputs to any given node's update function δ, there is a natural nondeterministic interpretation of the fact that the nodes have different clocks.

Many interesting questions arise in this context. One is, what kinds of the phase space properties remain *invariant* under this kind of nondeterminism? Given a triple (Γ, N, M), it can be readily shown that the fixed points are invariant with respect to the *fair* node update orderings in the (synchronized) sequential CA, and, moreover, the FPs are the same for the corresponding parallel CA. On the other hand, as the results in Section 3 indicate, neither cycle configurations nor transient configurations are invariant with respect to whether the nodes are updated sequentially or concurrently (and, in case of the former, in what order). It can be readily observed that, indeed, the FPs are also invariant for the asynchronous CA and graph automata, as well - provided that all the nodes have reached their respective states corresponding to the same fixed point global configuration, and that they all *locally agree* what (sub)configuration they are in, even if their individual local clocks possibly disagree with one another. Therefore, earlier results in [5] on the FP invariance for sequential and parallel graph automata are just special cases of this, more general result.

Theorem 4 *Given an arbitrary asynchronous cellular or graph automaton, any fixed point configuration is invariant with respect to the choice of a node update ordering, provided that each node x_i has an up-to-date knowledge of the current state of its neighborhood, N_i.*

In addition to studying invariants under different assumptions on asynchrony and concurrency, we also consider a broad qualitative comparison-and-contrast of the asynchronous CA that we propose, with the parallel CA and the sequential SCA and NICA. Such a study would shed more light on those emerging collective behaviors that are solely due to network delays.

5. Summary and Conclusions

We have presented in this paper some early steps in studying cellular automata when the unrealistic assumptions of *perfect synchrony* and *instantaneous unbounded parallelism* are dropped. Motivated by the well-known model of the sequential interleaving semantics of concurrency, we have tried to apply this metaphor to parallel CA, thereby motivating the study of sequential cellular automata (SCA) and the sequential interleavings cellular automata (NICA). In particular, we have undertaken a comparison and contrast between the sequential SCA/NICA and the classical, parallel CA models when the node update rules are restricted to *simple threshold functions.* Concretely, we have shown that, even in some very simplistic cases, the sequential "interleaving semantics" of NICA fails to capture concurrency of the parallel CA. One lesson is that, simple as they may be, the basic local operations of the classical CA cannot always be considered atomic. It then appears reasonable - indeed, necessary - to consider a single local node update to be made of an ordered sequence of the finer elementary operations:

- Fetching all the neighbors' values ("receiving" or "reading shared variables");
- Updating one's own state according to the update rule δ (that is, performing the local computation); and
- Informing the neighbors of the update, i.e., making available one's new state/value to the neighbors ("sending" or "writing a shared variable").

Motivated by the early results on the sequential and parallel threshold CA, and some of the implications of those results, we next consider various extensions. The central idea is to introduce a class of *genuinely asynchronous CA* (ACA) and to formally study their properties. Our hope is that the models along the lines of ACA would lead to some significant future insights

into the fundamental issues related to bounded vs. unbounded asynchrony, formal sequential semantics for parallel and distributed computation, and, on the dynamical systems side, to the identification of many of those parallel CA, SCA, NICA and/or ACA phase space properties that are solely or primarily due to the assumed communication model, that is, the (a)synchrony in both the local node updates and the inter-node interactions.

To conclude, we find appropriate extensions of the basic CA model to provide a simple, elegant and useful framework for a high-level study of various global qualitative properties of distributed, parallel and real-time systems at an abstract, yet mathematically elegant and comprehensive level.

Acknowledgments: This work was supported in part by the ONR MURI Grant, contract number N00014-02-1-0715, as well as the DARPA IPTO TASK Program, contract number F30602-00-2-0586. The first author would also like to thank the ECCS'05 organizers and European Complex Systems Society for the partial travel support. Many thanks to Reza Ziaei for valuable discussions related to the work presented in this paper.

References

1. M. Anthony. "Threshold Functions, Decision Lists, and the Representation of Boolean Functions", NeuroCOLT Techn. Report Series (NC-TR-96-028), January 1996
2. S. Amoroso and Y. Patt. "Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures", *J. of Computer and System Sciences (JCSS),* vol. 6, pp. 448-464, 1972
3. W. Ross Ashby. *"Design for a Brain",* Wiley, 1960
4. C. Barrett, H. B. Hunt III, M. V. Marathe, S. S. Ravi, D. J. Rosenkrantz and R. E. Stearns. "Predecessor and Permutation Existence Problems for Sequential Dynamical Systems", Los Alamos National Laboratory Report, LA-UR-01-668, 2001
5. C. Barrett, H. Hunt, M. Marathe, S. S. Ravi, D. Rosenkrantz, R. Stearns, and P. Tosic. "Gardens of Eden and Fixed Points in Sequential Dynamical Systems", *Discrete Math. and Theoretical Comp. Sci. (DMTCS),* Proc. AA DM-CCG, pp. 95-110, 2001
6. C. Barrett, H. B. Hunt III, M. V. Marathe, S. S. Ravi, D. J. Rosenkrantz, and R. E. Stearns. "Reachability problems for sequential dynamical systems with threshold functions", *Theoretical Computer Science (TCS),* vol. 295, issues 1-3, pp. 41-64, Feb. 2003
7. C. Barrett, H. Mortveit, and C. Reidys. "Elements of a theory of computer simulation II: sequential dynamical systems", *Applied Mathematics and Computation,* vol. 107 (2-3), 2000
8. C. Barrett, H. Mortveit, and C. Reidys. "Elements of a theory of computer simulation III: equivalence of sequential dynamical systems", *Appl. Math. and Comput.,* vol. 122 (3), 2001
9. C. Barrett and C. Reidys. "Elements of a theory of computer simulation I: sequential CA over random graphs", *Appl. Math. and Comput.,* vol. 98 (2-3), 1999
10. I. Czaja, R. J. van Glabbek, U. Goltz. "Interleaving semantics and action refinement with atomic choice", in "Advances in Petri Nets", (G. Rozenberg, ed.), LNCS 609, Springer-Verlag, 1992
11. Max Garzon. *"Models of Massive Parallelism: Analysis of Cellular Automata and Neural Networks",* Springer, 1995
12. R. J. van Glabbek, U. Goltz. "Equivalences and refinement", Proc. LITP Spring School Theoretical CS, La Roche-Posay, France (I. Guessarian, ed.), LNCS 469, Springer-Verlag 1990
13. E. Goles, F. Fogelman, D. Pellegrin. "Decreasing energy functions as a tool for studying threshold networks", *Discr. Appl. Math.,* vol. 12, pp. 261 - 277, 1985
14. E. Goles, S. Martinez. "Exponential transient classes of symmetric neural networks for synchronous and sequential updating", *Complex Systems,* vol. 3, pp. 589 - 597, 1989

15. E. Goles, S. Martinez. *"Neural and Automata Networks: Dynamical behavior and Applications",* Math. and Its Applications series (vol. 58), Kluwer, 1990
16. E. Goles, S. Martinez (eds.). *"Cellular Automata and Complex Systems",* Nonlinear Phenomena and Complex Systems series, Kluwer, 1999
17. H. Gutowitz (ed.). *"Cellular Automata: Theory and Experiment",* The MIT Press / North-Holland, 1991
18. C. A. R. Hoare. *"Communicating Sequential Processes",* Prentice Hall, 1985
19. T. E. Ingerson and R. L. Buvel. "Structure in asynchronous cellular automata", *Physica D: Nonlinear Phenomena,* Vol. 10, Issues 1-2, pp. 59 - 68, January 1984
20. S. A. Kauffman. "Emergent properties in random complex automata" (ibid.)
21. R. Milner. *"A Calculus of Communicating Systems",* Springer-Verlag Lecture Notes in Computer Science (LNCS), 1980
22. R. Milner. "Calculi for synchrony and asynchrony", *Theoretical Computer Science (TCS),* vol. 25, pp. 267 - 310, 1983
23. R. Milner. *"Communication and Concurrency",* Prentice-Hall, 1989
24. J. von Neumann. *"Theory of Self-Reproducing Automata",* edited and completed by A. W. Burks, Univ. of Illinois Press, Urbana, 1966
25. C. Nichitiu and E. Remila. "Simulations of Graph Automata" Proc. MFCS'98 Satellite Workshop on Cellular Automata, Brno, Czech Republic, Aug. 1998
26. P. Orponen. "On the computational complexity of discrete Hopfield nets", Proc. 20th Int'l Colloqium on Automata, Languages and Programming (ICALP '93), Springer LNCS series, vol. 700, pp. 215 - 226, 1993
27. P. Orponen. "Computational complexity of neural networks: a survey", *Nordic J. Comp.,* vol. 1 (1), pp. 94 - 110, 1994
28. P. Orponen. "Computing with truly asynchronous threshold logic networks", *TCS,* vol. 174 (1-2), pp. 123 - 136, 1997
29. J. C. Reynolds. *"Theories of Programming Languages",* Cambridge Univ. Press, 1998
30. C. Reidys. "On acyclic orientations and sequential dynamical systems", *Advances Appl. Math.,* vol. 27, 2001
31. R. Sethi. *"Programming Languages: Concepts and Constructs"* (2nd ed.), Addison-Wesley, 1996
32. K. Sutner. "Computation theory of cellular automata", MFCS98 Satellite Workshop on CA, Brno, Czech Rep., 1998
33. P. Tosic. "A Perspective on the Future of Massively Parallel Computing: Fine-Grain vs. Coarse-Grain Parallel Models", Proc. 1st ACM Conf. on Computing Frontiers (CF'04), Ischia, Italy, April 2004
34. P. Tosic, G. Agha. "Concurrency vs. Sequential Interleavings in 1-D Threshold Cellular Automata", APDCM Workshop within Int'l Parallel & Dist. Processing Symp. (IPDPS), Santa Fe, New Mexico, USA, April 2004 (in Proc. IEEE IPDPS '04)
35. P. Tosic, G. Agha. "Characterizing Configuration Spaces of Simple Threshold Cellular Automata", ACRI'04, Amsterdam, The Netherlands, Oct. 25-28, 2004; in Springer-Verlag LNCS series, vol. 3305, pp. 861 - 870
36. S. Wolfram. "Twenty problems in the theory of CA", *Physica Scripta,* vol. 9, 1985
37. S. Wolfram (ed.). *"Theory and applications of CA",* World Scientific, Singapore, 1986
38. S. Wolfram. *"Cellular Automata and Complexity (collected papers)",* Addison-Wesley, 1994
39. S. Wolfram. *"A New Kind of Science",* Wolfram Media, Inc., 2002

A Spin Glass Model of Human Logic Systems

Fariel Shafee
Physics Department
Princeton University
Princeton, NJ 08540

In this paper, we model logic networks interacting among themselves and also with the environment.

We assume that although each agent is rational in the sense that he tries to maximize her utilities, the utilities themselves are not absolute and depend in part on slightly varying genetic make-ups leading to innate differences in preferences, varying acquired tastes depending on local non homogeneity among interacting agents and the environment, and also slight differences in perception and vision.

The agents acquire knowledge by interacting with nature and other agents. The preferences serve as the "truths taken for granted" or axioms for an agent's logic system and he uses the knowledge to enhance his utilities.

We can assume perception as an agent's connecting or interacting points with the environment and other agents. However, if we model an agent's perception of the environment as a shadow of a higher dimensional manifold, slight variations in the projection will lead to differing ideas about the environment. An agent will interact with the environment based on the projection he has knowledge of, and the interaction will have effects on another agent's projection, which may not coincide with the first projection. Hence, although both agents act on the "truth" they perceive, one's actions might have conflicting results on another agent's universe.

Also, agents' tastes or preferences dictate to some extent how they want to "change" the environment, or what their visions of the future are. One agent might have an inclination toward one of many possible states of nature while the second agent might prefer another. However, since both agents must interact with the same environment they are connected to, conflicts arise from different visions about the future associated with the same environment.

This situation can be modeled in a way similar to the many-universe hypothesis [1] .We assume that the environment is "collapsing to a future." The collapse can occur to one of many superposed states and the agents bid on different states. A similar model with quantum collapse with bidding has been proposed [2, 3].

The bidding also involves an agent's calculation of risk factors associated with investing his time and energy in a certain "stock." Since each agent has only a finite amount of time and energy, one can calculate upto a finite number of leading terms in the many variables associated when calculating risks. Each agent decides where to truncate the calculation, depending on his weighed preferences.

We connect the logic networks [4] in a model akin to the spin model [5]. Here, we can draw similarities with thermodynamics where a micro-system (in this case the perceptory organs - coupled to the neural network) is placed in conjunction with a macro-system (which is the environment). These logic networks placed in a society may be visualized as a small thermodynamic spin lattice where the axioms (i, j) residing in different networks are coupled to one another by coupling constants J_{ij} , and this lattice itself is coupled to a bigger lattice, which is

the environment. However, since we can define the environment lattice to be huge compared to the neural network lattice, we can take average values for interaction purposes and couple the lattice with the neural network lattice with some multidimensional coupling factor.

In each of the lattices, spins are at a quantum level described as "states" which can coexist in many orthonormal superpositions. However, when the smaller lattice interacts with the bigger lattice, the coupling causes the environment lattice to collapse to a certain value. This value will depend on the probabilistic coefficients of the wave functions and most of the time it would yield the expected value. So an average person will end up with an average set of axioms. Each of the agents has a certain set of axioms to start with. Again, these agents are coupled with one another in a lattice. We argue that the agents are not connected with one another with a random coupling constant J_{ij} , but some rules are defined, and also that these J_{ij}s are updatable according to the specific state of the entire network.

1. J_{ij} is not symmetric, i.e., $J_{ij} \neq J_{ji}$. The value of J_{ij} depends on i possessing axioms that necessitate the existence of j. So i will be coupled to j more strongly if i possesses axioms that require the existence of j. Now each agent will have the following behaviors in the game:

1. Each agent i will tend to change its neighbors' axioms if the neighbors' axioms contain contradictions of its axioms. The frequency and strength of flipping would depend on a coupling constant C_{ij} . C_{ij} depends on the following:

a. The evolution of the logical code developed in i that contains that particular axiom, i.e. the networking of the certain axiom in the logic network of i. More accurately, the number of decisions produced by i that reflect the use of the particular axiom. In other words, i will tend to flip a neighbors anti-axiom with more effort if the axiom has become an important part of its network, and any future attempt of j's flipping it would cost i dearly.

b. The determination of the number of the contradictory axiom in the neighboring agents' logic network. An increased frequency of anti-axioms in is network would increase the possibility of an anti-axiom to be used in a future decision.

c. The effect of j's decision on its environment (might be caused by physical distance between the two agents). The total strength of coupling between i and j ($i \rightarrow j$) would be $- J^{n}_{ij} (state^{n}_{i} \rightarrow state^{n}_{j}) + C^{m}_{ij} (state^{m}_{i} \rightarrow state^{m}_{j})$, where n and m are states or registers representing axioms.

We can implement this scheme by linking two agents with appropriate logic gates. The value of this coupling could be described as feelings of agent i towards agent j. This total coupling will, at a macroscopic level, cause i to play for or against j, i.e. collaborate with j or work against the existence of j.

2. An axiom will flip if the effect of the neighbor's having the same axiom state causes the coupling to exceed a flipping energy. The flipping energy depends on certain axioms' connectivity with other axioms in the agent's cognitive network.

3. All agents must possess an axiom we shall call self preservation or preservation of the network in random probability. We label this axiom P. Agents containing neither of these axioms cannot contribute to the existence of the agent or to the network. In that case those agents' axioms, if few, will be flipped by other agents; or they will self-destroy. A later paper will discuss the effect of agents possessing destructive axioms in their cognitive network, or any critical number that will bound the fraction of agents with self destructive axioms in a network. However, in this paper, we assume that all agents possess P.

The probability of obtaining a certain state of an axiom by another agent can now be written down in a simplified version by the formula:

$$F(A) = \sum C_{ijA}\, f_1(\text{flip}) - \sum J_{ijA}\, f_2(\text{stabilize}) - f_3(1/R) + K < A >$$

Here, A is the axiom, R is a resistance factor for flipping the axiom depending on how entangled the axiom is in the agents own cognitive network, K is the coupling of the agent with nature, and <A > is the expectation probability of the axiom in nature. We can see that this formula is very similar to the formula for a classical neural network, except that the coupling constants, unlike the weight factors in a regular neural network, do not sum up to 1. Also, instead of adding a term in <A >, it might be more realistic to add a term F(< A >). Here F(A) is a switching function that takes on a value of either -1 or 1, depending on whether the RHS exceeds a certain threshold. So every time A or Ã (not-A) flips, the C's and the J's are interchanged, and R is updated to a new value that needs to be updated with the accumulation of new axioms that get entangled with the flipped clause.

A cost of entropy term must be added depending on whether the agent has an axiom involving taking entropy into account or an axiom that ignores entropy in its network. Entropy has been defined variously in different contexts, usually as a measure of information needed about the state of the system, which can be related to the complexity of the system [6, 7], whereas information itself has been related to symmetry-breaking [8]. In a macro-system the dynamics guides the system to a state of least free energy, which is related to energy (U) and the entropy (S) by

$$F = U - T S$$

T being the temperature.

In condensed matter physics T is found to be analogous to inverse time, and hence entropy behaves like energyX time, which is a reasonable guide to the introduction of this quantity, because it is necessary to reduce the utility (F) by a quantity which involves the complexity of the system that the agents have to unravel to evaluate the different terms from the interactions of the axioms, and also the inverse time (rate or speed) involved in the process. This opens up the possibility of highly intricate games. The nearly chaotic nature of even small world examples with variable spin-like constituents in simulation examples have been reported recently [9], but the role of entropy, or of ignoring its role by a class of agents using an "ignore entropy axiom" remains to be studied thoroughly.

REFERENCES

1. H. Everett in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. B. S. DeWitt and N. Graham (Princeton U. P., Princeton, 1973)
2. F. Shafee, quant-ph/0502111 (2005)
3. F. Shafee, quant-ph/0507106 (2005)
4. F. Shafee, quant--ph/0203010 (2002).
5. D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett*. **35**, 1792, (1975)
6. E. Ott, *Chaos in Dynamical Systems*, (Cambridge U.P.. 1993) p. 138
7. M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*. (Springer-Verlag, N.Y., 2nd ed., 1997).
8. J.D. Collier, *Culture and Science* **7**, 247 (1996).
9. F. Shafee, physics/0506161 (2005)

Information Tech. Modelling

Sampling of networks with traceroute-like probes

Alain Barrat [a,*], Ignacio Alvarez-Hamelin [a], Luca Dall'Asta [a], Alexei Vázquez [b], Alessandro Vespignani [a,c]

[a] *Laboratoire de Physique Théorique (UMR 8627 du CNRS), Bâtiment 210, Université de Paris-Sud, 91405 ORSAY Cedex France*

[b] *Nieuwland Science Hall, University of Notre Dame, Notre Dame, IN 46556, USA.*

[c] *School of Informatics and Department of Physics, Indiana University, Bloomington, IN 47408, USA*

Abstract

A large part of the recent development of the interest in complex networks has been triggered by the observation of particular characteristics of real world networks, such as the small-world properties or the heavy-tailed distributions of degrees. Many datasets are however the result of an incomplete sampling of the underlying real networks, and it has been argued that sampling procedures might introduce uncontrolled biases in the statistical properties of the sampled graph. In this paper, we explore this issue in the case of the Internet, which is generally mapped from a limited set of sources by using `traceroute`-like probes. The origin of the biases introduced by such a sampling process is investigated and related with the global topological properties of the underlying network. We complement the analytical discussion with a throughout numerical investigation of simulated mapping strategies in network models with different topologies.

Key words: Network Sampling, Traceroute, Internet exploration, Topology inference

* Corresponding author: A. Barrat, LPT, Bâtiment 210, Université de Paris-Sud, 91405 ORSAY Cedex France; email: Alain.Barrat@th.u-psud.fr; Tel: +33 1 69 15 82 22; Fax: +33 1 69 15 82 87

Preprint submitted to Elsevier Science *10 August 2005*

1 Introduction

A significant research and technical challenge in the study of large information networks is related to the lack of highly accurate maps providing information on their basic topology. This is mainly due to the dynamical nature of their structure and to the lack of any centralized control resulting in a self-organized growth and evolution of these systems. A prototypical example of this situation is faced in the case of the physical Internet. The topology of the Internet can be investigated at different granularity levels such as the router and Autonomous System (AS) level, with the final aim of obtaining an abstract representation where the set of routers (ASs) and their physical connections (peering relations) are the vertices and edges of a graph, respectively. In the absence of accurate maps, local views are obtained by evaluating a certain number of paths to different destinations by using specific tools such as `traceroute` or by the analysis of BGP tables. At first approximation these processes amount to the collection of shortest paths from a source vertex to a set of target vertices, obtaining a partial spanning tree of the network. The merging of several of these views provides the map of the Internet from which the statistical properties of the network are evaluated.

This strategy has led to the obtention of various maps of the Internet [1–5] which have been used for the statistical characterization of the network. Defining $\mathcal{G} = (V, E)$ as the sampled graph of the Internet with $N = |V|$ vertices and $|E|$ edges, it is quite intuitive that the Internet is a *sparse* graph with a much lower number of edges than in a complete graph: $|E| \ll N(N-1)/2$. Moreover, the average distance, measured as the shortest path, between vertices is very small. This is the so called *small-world* property, that is essential for the efficient functioning of the network. Most surprising is the evidence of a skewed and heavy-tailed behavior for the probability that any vertex in the graph has degree k defined as the number of edges linking each vertex to its neighbors. In particular, the degree distribution appears to be approximated by $P(k) \sim k^{-\gamma}$ with $2 \leq \gamma \leq 2.5$ [6]. Evidence for the heavy-tailed behavior of the degree distribution has been collected in several other studies at the router and AS level [7–11] and have generated a large activity in the field of network modeling and characterization [12–16].

The obtained maps are however undoubtedly incomplete. Along with technical problems such as the instability of paths between routers and interface resolutions [17], typical mapping projects are run from relatively small sets of sources whose combined views are missing a considerable number of edges and vertices [11,18]. In particular, the various spanning trees are specially missing the lateral connectivity of targets and sample more frequently vertices and links which are closer to each source, introducing spurious effects that might seriously compromise the statistical accuracy of the sampled graph.

These *sampling biases* have been explored in numerical experiments of synthetic graphs generated by different algorithms [19–22]. Very interestingly, it has been shown (numerically and analytically) that apparent degree distributions with heavy-tails may be observed even from homogeneous topologies such as in the classic Erdös-Rényi graph model [19,20,23]. These studies thus point out that the evidence obtained from the analysis of the Internet sampled graphs might be insufficient to draw conclusions on the topology of the actual Internet network.

This issue may be tackled through a mean-field statistical analysis and extensive numerical study of shortest path routed sampling, considered as the first approximation to `traceroute`-sampling (see Section 2), in different networks models. We recall in Section 3 the theoretical arguments leading to an approximate expression for the probability of edges and vertices to be detected. The analytical study provides a general understanding of which kind of topologies yields the most accurate sampling. In particular, the map accuracy depends on the underlying network *betweenness centrality* distribution; the heavier the tail the higher the statistical accuracy of the sampled graph.

Numerical investigation of maps obtained varying the number of source-target pairs on networks models with different topological properties provides support to the analytical analysis. In particular, we consider networks with degree distribution with poissonian, Weibull and power-law behavior. We study the fractions of discovered vertices and edges as a function of the degree (Section 4), stressing the agreement with the theoretical predictions, as well as the degree distributions obtained in the sampled graph (Section 5). Single source mapping processes are shown to face serious limitations: even the targeting of the whole network results in a very partial discovery of its connectivity. On the contrary, the use of multiple sources promptly leads to obtained maps fairly consistent with the original sample.

In Section 6, we also inspect quantitatively the portion of discovered network in different mapping strategies for the deployment of sources that however impose the same density of probes to the network. A region of low efficiency (less vertices and edges discovered) is found, depending on the relative proportion of sources and targets. This low efficiency region however corresponds to the optimal estimation of the network average degree. This finding calls for a "trade-off" between the accuracy in the observation of different quantities and hints to possible optimization procedures in the `traceroute`-driven mapping of large networks.

2 Network models and `traceroute`-like processes

In a typical `traceroute` study, active sources deployed in the network send `traceroute` probes to a set of destination vertices. Each probe collects information on all the vertices and edges traversed along the path connecting the source to the destination [17]. By merging the information collected on each path it is then possible to reconstruct a partial map of the network. The edges and the vertices discovered by each probe will depend on the "path selection criterium" used to decide the path between a pair of vertices. In the real Internet, many factors, including commercial agreement, traffic congestion and administrative routing policies, contribute to determine the actual path, which may differ even considerably from the shortest path. Despite these local, often unpredictable path distortions, a reasonable first approximation of the route traversed by `traceroute`-like probes is the shortest path between the two vertices. This assumption, however, is not sufficient for a proper definition of a `traceroute` model in that equivalent shortest paths between two vertices may exist. For the sake of simplicity, we can thus define three selection mechanisms defining different ideal-paths that may account for some of the features encountered in real Internet discovery:

- Unique Shortest Path (USP) probe. In this case the shortest path route selected between a vertex i and the destination target T is always the same independently of the source S (the path being initially chosen at random among all the equivalent ones).
- Random Shortest Path (RSP) probe. The shortest path between any source-destination pair is chosen randomly among the set of equivalent shortest paths. This might mimic different peering agreements that make independent the paths among couples of vertices.
- All Shortest Paths (ASP) probe. The selection criterium discovers all the equivalent shortest paths between source-destination pairs. This might happen in the case of probing repeated in time (long time exploration), so that back-up paths and equivalent paths are discovered in different runs.

We will generically call $\mathcal{M}$-path the path found using one of these measurement or path selection mechanism. Actual `traceroute` probes contain a mixture of the three mechanisms defined above. We do not attempt, however, to account for all the subtleties that real studies encounters, i.e. IP routing, BGP policies, interface resolutions and many others. In fact, in the real mapping process, many effective heuristic strategies are commonly applied to improve the reliability and the performances of the sampling. However, it turns out that the different path selection criteria (p.s.c.) have only little influence on the general picture emerging from our results. Moreover, the USP procedure clearly represents the worst case scenario since, among the three different methods, it yields the minimum number of discoveries. For this reason, if not otherwise

specified, we will report the USP data to illustrate the general features of our synthetic exploration. The interest of this analysis resides properly in the choice of working in the most pessimistic case, being aware that path inflations should actually provide a more pervasive sampling of the real network.

More formally, the experimental setup for our simulated `traceroute` mapping is the following. Let $G = (V, E)$ be a sparse undirected graph with vertices (ver tices) $V = \{1, 2, \cdots, N\}$ and edges (links) E. Then let us define the sets of vertices $\mathcal{S} = \{i_1, i_2, \cdots, i_{N_S}\}$ and $\mathcal{T} = \{j_1, j_2, \cdots, j_{N_T}\}$ specifying the random placement of N_S sources and N_T destination targets. For each ensemble of source-target pairs $\Omega = \{\mathcal{S}, \mathcal{T}\}$, we compute with our p.s.c. the paths connecting each source-target pair. The sampled graph $\mathcal{G} = (V^*, E^*)$ is defined as the set of vertices V^* (with $N^* = |V^*|$) and edges E^* induced by considering the union of all the $\mathcal{M}$-paths connecting the source-target pairs. The sampled graph is thus analogous to the maps obtained from real `traceroute` sampling of the Internet.

In our study the parameters of interest are the densities $\rho_T = N_T/N$ and $\rho_S = N_S/N$ of targets and sources. In general, `traceroute`-driven studies run from a relatively small number of sources to a much larger set of destinations. For this reason, it is appropriate to work with the density of targets ρ_T while still considering N_S instead of the corresponding density. In many cases, an appropriate quantity representing the level of sampling of the networks is $\epsilon = N_S N_T/N$: it represents the density of `traceroute` probes in the network and therefore a measure of the load provided to the network by the measuring infrastructure.

In the following, our aim is to evaluate to which extent the statistical properties of the sampled graph $\mathcal{G}$ depend on the parameters of our experimental setup and are representative of the properties of the underlying graph G. The analytical insights of Section 3 will be complemented by a numerical investigation of the `traceroute`-like exploration process on various graph models endowed with very well-defined topological properties, so as to give a clear result on which kind of topologies are related to good sampling performances and vice-versa. Starting from this first investigation, further studies could deal with more realistic models such as those created using Internet topology generators [13,12]. In particular, we will consider two main classes of graphs.

A) Homogeneous graphs in which the degree distribution $P(k)$ has small fluctuations and a well defined average degree. In this context, the *homogeneity* refers to the existence of a meaningful characteristic average degree that represents the typical value in the graph. The most widely known model for homogeneous graphs is given by the classical Erdös-Rényi (ER) model [24]: in such random graphs $G_{N,p}$ of N vertices, each edge is present in E independently with probability p. The expected number of edges is therefore

$|E| = pN(N-1)/2$. In order to have sparse graphs one thus needs to have p of order $1/N$, since the average degree is $p(N-1)$. Erdös-Rényi graphs are typical examples of homogeneous graphs, with degree distribution following a Poisson law. Since $G_{N,p}$ can consist of more than one connected component, we consider only the largest of these components.

Another important characteristic discriminating the topology of graphs is the clustering coefficient c_i that, giving the fraction of connected neighbors of a given node i, measures the local cohesiveness of nodes. The average clustering coefficient $C = \frac{1}{N}\sum_i c_i$ provides an indication of the global level of cohesiveness of the graph. This number is generally very small in random graphs that lack of correlations. In many real graphs however, the clustering coefficient appears to be very high and opportune models have been formulated to represent this property, both for homogeneous and heterogeneous graphs. In particular, we consider the construction algorithm proposed by Watts and Strogatz for small-world networks [28]: starting from a regular network (e.g. a one-dimensional lattice with connections to the $\overline{k}$ nearest neighbors along the chain), each link is rewired with a certain probability p. The resulting degree distribution has a shape similar to the case of Erdös-Rényi graphs, peaked around its average value. The clustering coefficient, however, is large if $p \ll 1$, making this network a typical example of clustered homogeneous network.

B) Heterogeneous graphs for which $P(k)$ is a broad distribution with heavy-tail and large fluctuations, spanning various orders of magnitude. The prototype of a scale-free graph is the growing network model by Albert and Barabási (BA) [29]. The preferential attachment mechanism (each new node is connected to m already existing nodes chosen with a probability proportional to their degree) yields a connected graph of $|V| = N$ nodes with $|E| = mN$ edges, having a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma = 3$, and small clustering coefficient. Another growing model has been introduced by Dorogovtsev, Mendes and Samukhin (DMS) [30]: at each time step, a new node is introduced and connected to *the two extremities of a randomly chosen edge*, thus forming a triangle. A given node is thus in fact chosen with a probability proportional to its degree, which corresponds to the preferential attachment

Table 1
Main characteristics of the graphs used in the numerical exploration.

	ER	ER	WS	BA	DMS	RSF	Weibull
N	10^4	10^4	10^4	10^4	10^4	10^4	10^4
$\|E\|$	10^5	5.10^5	10^5	4.10^4	2.10^4	22000	55000
$\overline{k}$	20	100	20	8	4	4.4	11
C	0.002	0.01	0.52	0.006	0.74	0.067	0.12
k_{max}	40	140	26	334	346	3500	2000

rule. The resulting graphs have a large clustering coefficient (≈ 0.74) along with a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma = 3$.

Such graphs can be considered as particular since they are constructed with the preferential attachment mechanism, and we also consider random graphs with given broad degree distributions. In the literature, different definitions of heavy-tailed like distributions exist. While we do not want to enter the detailed definition of heavy-tailed distribution we have considered two classes of such distributions: (i) *scale-free* or Pareto distributions of the form $P(k) \sim k^{-\gamma}$ (RSF), and (ii) Weibull distributions (WEI) $P(k) = (a/c)(k/c)^{a-1} \exp(-(k/c)^a)$. The scale-free distribution has a diverging second moment and therefore virtually unbounded fluctuations, limited only by eventual size-cut-off. The Weibull distribution is akin to power-law distributions truncated by an exponential cut-off which are often encountered in the analysis of scale-free systems in the real world. Indeed, a truncation of the power-law behavior is generally due to finite-size effects and other physical constraints. Both forms have been proposed as representing the topological properties of the Internet [8]. We have generated the corresponding random graphs by using the algorithm proposed by Molloy and Reed [31]: the vertices of the graph are assigned a fixed sequence of degrees $\{k_i\}$, $i = 1, \ldots, N$, chosen at random from the desired degree distribution $P(k)$, and with the additional constraint that the sum $\sum_i k_i$ must be even; then, the vertices are connected by $\sum_i k_i/2$ edges, respecting the assigned degrees and avoiding self- and multiple-connections. The parameters used are $a = 0.25$ and $c = 0.6$ for the Weibull distribution, and $\gamma = 2.3$ for the RSF case.

The main properties of the various graphs are summarized in Table 1. In all numerical studies we have used networks of $N = 10^4$ vertices. It is noteworthy that the maximum value of the degree (k_{max}) is of the same order as the average for homogeneous graphs, but much larger for heterogenous ones.

3 Mean-field theory of simulated mapping process

We begin our study by recalling briefly the mean-field statistical analysis of the simulated `traceroute` mapping done in [32]. The aim is to provide a statistical estimate for the probability of edge and vertex detection as a function of N_S, N_T and the topology of the underlying graph.

Let us define the quantity $\sigma_{i,j}^{(l,m)}$ that takes the value 1 if the edge (i, j) belongs to the selected $\mathcal{M}$-path between vertices l and m, and 0 otherwise. For a given set of sources and targets $\Omega = \{\mathcal{S}, \mathcal{T}\}$, the indicator function that a given edge (i, j) will be discovered and belongs to the sampled graph is simply $\pi_{i,j} = 1$ if the edge (i, j) belongs to at least one of the $\mathcal{M}$-paths connecting the source-

target pairs, and 0 otherwise. We can obtain an exact expression for $\pi_{i,j}$ by noting that $1-\pi_{i,j}$ is 1 if and only if (i,j) does not belong to any of the paths between sources and targets, i.e. if and only if $\sigma_{i,j}^{(l,m)} = 0$ for all $(l,m) \in \Omega$. This leads to

$$\pi_{i,j} = 1 - \prod_{l \neq m} \left(1 - \sum_{s=1}^{N_S} \delta_{l,i_s} \sum_{t=1}^{N_T} \delta_{m,j_t} \sigma_{i,j}^{(l,m)} \right), \tag{1}$$

where $\delta_{i,j}$ is the Kronecker symbol and selects only vertices belonging to the set of sources or targets.

Starting from the above exact formula, it is interesting to look at the process on a statistical ground by studying the average over all possible realizations of the set $\Omega = \{\mathcal{S}, \mathcal{T}\}$, identified by $\langle \cdots \rangle$. An uncorrelation assumption allows to obtain the average discovery probability of an edge as

$$\langle \pi_{i,j} \rangle \simeq 1 - \prod_{l \neq m} \left(1 - \rho_T \rho_S \left\langle \sigma_{i,j}^{(l,m)} \right\rangle \right), \tag{2}$$

where we take advantage of neglecting correlations by replacing the average of the product of variables with the product of the averages. In the case of the ASP probing, $\left\langle \sigma_{i,j}^{(l,m)} \right\rangle$ is just one if (i,j) belongs to one of the shortest paths between l and m, and 0 otherwise. In the case of the USP and the RSP, on the contrary, only one path among all the equivalent ones is chosen. If we denote by $\sigma^{(l,m)}$ the number of shortest paths between vertices l and m, and by $x_{i,j}^{(l,m)}$ the number of these paths passing through the edge (i,j), the probability that the `traceroute` model chooses a path going through the edge (i,j) between l and m is $\left\langle \sigma_{i,j}^{(l,m)} \right\rangle = x_{i,j}^{(l,m)} / \sigma^{(l,m)}$.

The standard situation we consider is the one in which $\rho_T \rho_S \ll 1$ and since $\left\langle \sigma_{i,j}^{(l,m)} \right\rangle \leq 1$, we have

$$\prod_{l \neq m} \left(1 - \rho_T \rho_S \left\langle \sigma_{i,j}^{(l,m)} \right\rangle \right) \simeq \prod_{l \neq m} \exp\left(-\rho_T \rho_S \left\langle \sigma_{i,j}^{(l,m)} \right\rangle \right), \tag{3}$$

that inserted in Eq.(2) yields

$$\langle \pi_{i,j} \rangle \simeq 1 - \exp\left(-\rho_T \rho_S b_{ij} \right), \tag{4}$$

where $b_{ij} = \sum_{l \neq m} \left\langle \sigma_{i,j}^{(l,m)} \right\rangle$. In the case of the USP and RSP probing, the quantity b_{ij} is by definition the edge *betweenness centrality* $\sum_{l \neq m} x_{i,j}^{(l,m)} / \sigma^{(l,m)}$ [25,26], sometimes also refereed to as "load" [27] (In the case of ASP probing, it is a closely related quantity). Indeed the vertex or edge betweenness

is defined as the total number of shortest paths among pairs of vertices in the network that pass through a vertex or an edge, respectively. If there are multiple shortest paths between a pair of vertices, the path contributes to the betweenness with the corresponding relative weight. The *betweenness* gives a measure of the amount of all-to-all traffic that goes through an edge or vertex, if the shortest path is used as the metric defining the optimal path between pairs of vertices, and it can be considered as a non-local measure of the *centrality* of an edge or vertex in the graph.

The edge betweenness assumes values between 2 and $N(N-1)$ and the discovery probability of the edge will therefore depend strongly on its betweenness. In particular, for edges with minimum betweenness $b_{ij} = 2$ we have $\langle \pi_{i,j} \rangle \simeq 2\rho_T \rho_S$, that recovers the probability that the two end vertices of the edge are chosen as source and target. This implies that if the densities of sources and targets are small but finite in the limit of very large N, all the edges in the underlying graph have an appreciable probability to be discovered. Moreover, for edges with high betweenness the discovery probability approaches one. A fair sampling of the network is thus expected. In most realistic samplings, however, we face a very different situation. While it is reasonable to consider ρ_T a small but finite value, the number of sources is not extensive ($N_S \sim \mathcal{O}(1)$) and their density tends to zero as N^{-1}. In this case it is more convenient to express the edge discovery probability as

$$\langle \pi_{i,j} \rangle \simeq 1 - \exp\left(-\epsilon \widetilde{b_{ij}}\right), \tag{5}$$

where $\epsilon = \rho_T N_S$ is the density of probes imposed to the system and the rescaled betweenness $\widetilde{b_{ij}} = N^{-1} b_{ij}$ is now limited in the interval $[2N^{-1}, N-1]$. In the limit of large networks $N \to \infty$ it is clear that edges with low betweenness have $\langle \pi_{i,j} \rangle \sim \mathcal{O}(N^{-1})$, for any finite value of ϵ. This readily implies that in real situations the discovery process is generally not complete, a large part of low betweenness edges being not discovered, and that the network sampling is made progressively more accurate by increasing the density of probes ϵ.

A similar analysis can be performed for the discovery probability of vertices, leading to the average

$$\langle \pi_i \rangle \simeq 1 - (1 - \rho_S - \rho_T) \exp\left(-\rho_T \rho_S b_i\right), \tag{6}$$

where b_i is the vertex betweenness centrality, that is limited in the interval $[0, N(N-1)]$ [25–27]. The betweenness value $b_i = 0$ holds for the leafs of the graph, i.e. vertices with a single edge, for which we recover $\langle \pi_i \rangle \simeq \rho_S + \rho_T$. Indeed, this kind of vertices are dangling ends discovered only if they are either a source or target themselves. As discussed before, the most usual

setup corresponds to a density $\rho_S \sim \mathcal{O}(N^{-1})$ and in the large N limit we can conveniently write

$$\langle \pi_i \rangle \simeq 1 - (1 - \rho_T) \exp\left(-\epsilon \widetilde{b_i}\right), \tag{7}$$

where we have neglected terms of order $\mathcal{O}(N^{-1})$ and the rescaled betweenness $\widetilde{b_i} = N^{-1} b_i$ is now defined in the interval $[0, N-1]$. This expression points out that the probability of vertex discovery is favored by the deployment of a finite density of targets that defines its lower bound.

We can also provide a simple approximation for the effective average degree $\langle k_i^* \rangle$ of vertex i discovered by our sampling process. Each edge departing from the vertex contributes proportionally to its discovery probability, yielding

$$\langle k_i^* \rangle = \sum_j \left(1 - \exp\left(-\epsilon \widetilde{b_{ij}}\right)\right) \simeq \epsilon \sum_j \widetilde{b_{ij}}. \tag{8}$$

The final expression is obtained for edges with $\epsilon \widetilde{b_{ij}} \ll 1$. Since the sum over all neighbors of the edge betweenness is simply related to the vertex betweenness as $\sum_j b_{ij} = 2(b_i + N - 1)$, where the factor 2 considers that each vertex path traverses two edges and the term $N - 1$ accounts for all the edge paths for which the vertex is an endpoint, this finally yields

$$\langle k_i^* \rangle \simeq 2\epsilon + 2\epsilon \widetilde{b_i}. \tag{9}$$

Finally, the analysis allows to compute the edge redundancy $r_e(i, j)$ of an edge (i, j), defined as as the number of probes passing through the edge (i, j). This quantity is indeed written for a given set of probes and targets as

$$r_e(i,j) = \sum_{l \neq m} \left(\sum_{s=1}^{N_S} \delta_{l,i_s} \sum_{t=1}^{N_T} \delta_{m,i_t} \sigma_{i,j}^{(l,m)} \right). \tag{10}$$

Averaging over all possible realizations and assuming the uncorrelation hypothesis, we obtain

$$\langle r_e(i,j) \rangle \simeq \sum_{l \neq m} \rho_T \rho_S \left\langle \sigma_{i,j}^{(l,m)} \right\rangle = \rho_T \rho_S b_{ij} \ . \tag{11}$$

This result implies that the average redundancy of an edge is related to the density of sources and targets, but also to the edge betweenness. For example, an edge of minimum betweenness $b_{ij} = 2$ can be discovered at most twice in the extreme limit of an all-to-all probing. On the contrary, a very central edge of betweenness b_{ij} close to the maximum $N(N-1)$, would be discovered with

a redundancy close to $(N-1)$ by a `traceroute`-probing from a single source to all the possible destinations.

Similarly, the redundancy $r_n(i)$ of a vertex i, intended as the number of times the probes cross the vertex i, can be obtained:

$$\langle r_n(i) \rangle \simeq 2\epsilon + \rho_S \rho_T b_i \ . \tag{12}$$

In this case, a term related to the number of traceroute probes ϵ appears, showing that a part of the mapping effort unavoidably ends up in generating vertex detection redundancy.

The present analysis shows that the measured quantities and statistical properties of the sampled graph strongly depend on the parameters of the experimental setup and the topology of the underlying graph. The latter dependence appears through the key role played by edge and vertex betweenness in the expressions characterizing the graph discovery. The betweenness is a nonlocal topological quantity whose properties change considerably depending on the kind of graph considered. This allows an intuitive understanding of the fact that graphs with diverse topological properties deliver different answer to sampling experiments.

4 Numerics

The analytical findings of the previous section may be tested and used as guidance in the numerical analysis of simulated mapping experiments of network models. In particular we will consider the graph topologies defined in Section 2. Let us first consider the case of homogeneous graphs (ER and WS model): the vertex and edge betweennesses are homogeneous quantities and their distributions are peaked around their average values $\overline{b}$ and $\overline{b_e}$, respectively, spanning only a small range of variations. These values can thus be considered as typical values. We can thus use Eq. (5) and (7) to estimate the order of magnitude of probes that allows a fair sampling of the graph. Indeed, both $\langle \pi_{i,j} \rangle$ and $\langle \pi_i \rangle$ tend to 1 if $\epsilon \gg \max\left[\overline{b}^{-1}, \overline{b_e}^{-1}\right]$. In this limit all edges and vertices will have probability to be discovered very close to one. At lower value of ϵ, obtained by varying ρ_T and N_S, the underlying graph is only partially discovered. Fig. 1 shows for the WS model the behavior of the fraction N_k^*/N_k of discovered vertices of degree k, where N_k is the total number of vertices of degree k in the underlying graph, and the fraction of discovered edges $\langle k^* \rangle / k$ in vertices of degree k. N_k^*/N_k naturally increases with the density of targets and sources, and it is slightly increasing with k. The latter behavior can be easily understood by noticing that vertices with larger degree have on average

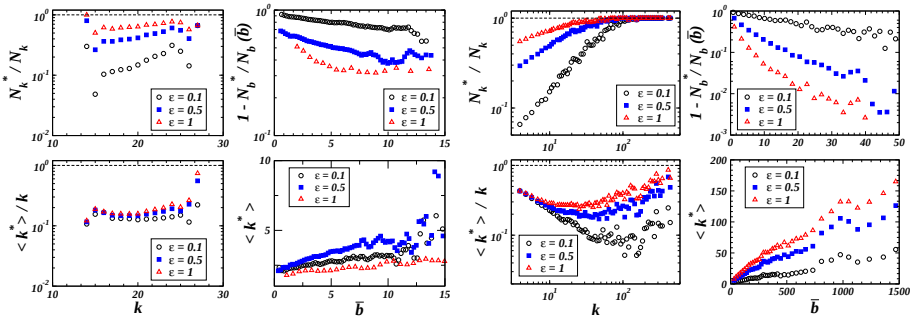


Fig. 1. Frequency N_k^*/N_k of detecting a vertex of degree k, frequency N_b^*/N_b of detecting a vertex of betweenness b and proportion of discovered edges $\langle k^* \rangle / k$ as a function of the degree and as a function of the betweenness for the WS (4 graphs on the left) and the BA (4 graphs on the right) models. The exploration setup considers $N_S = 2$ and increasing probing level ϵ obtained by progressively higher density of targets ρ_T.

a larger betweenness. On the other hand, the range of variation of k in homogeneous graphs is very narrow and only a large level of probing may guarantee very large discovery probabilities. Similarly the behavior of the effective discovered degree can be understood by looking at Eq. (9). Indeed the initial decrease of $\langle k^* \rangle / k$ is finally compensated by the increase of $\overline{b(k)}$.

The situation is different in graphs with heavy-tailed connectivity distributions (BA, DMS, RSF and WEI models), with an appreciable fraction of vertices and edges with very high betweenness [33]. In particular, in scale-free graphs the site betweenness is related to the vertices degree as $\overline{b(k)} \sim k^\beta$, where β is an exponent depending on the model [33]. Since in heavy-tailed degree distributions the allowed degree is varying over several orders of magnitude, the same occurs for the betweenness values, and the tail of the distribution is broader the broader the connectivity distribution. In such a situation, even in the case of small ϵ, vertices whose betweenness is large enough ($b_i \epsilon \gg 1$) have $\langle \pi_i \rangle \simeq 1$. Therefore all vertices with degree $k \gg \epsilon^{-1/\beta}$ will be detected with probability one. This is clearly exemplified for the BA model in Fig. 1 where the discovery probability N_k^*/N_k of vertices with degree k saturates to one for large degree values. Consistently, the degree value at which the curve saturates decreases with increasing ϵ. A similar effect is appearing in the measurements concerning $\langle k^* \rangle / k$. After an initial decay (Fig. 1) the effective discovered degree is increasing with the degree of the vertices. This qualitative feature is captured by Eq. (9) that gives $\langle k^* \rangle / k \simeq \epsilon k^{-1}(1 + \overline{b(k)})$. At large k the term $k^{-1}\overline{b(k)} \sim k^{\beta-1}$ takes over and the effective discovered degree approaches the real degree k. Fig. 1 also displays the frequency N_b^*/N_b and the discovered degree of vertices with betweenness b, showing in a more direct way the qualitative agreement with the analytical predictions.

In Fig. 2 we also report the behavior of the average vertex redundancy as

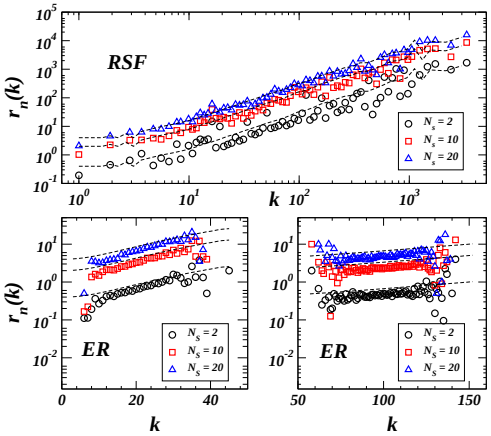


Fig. 2. Average vertex redundancy as a function of the degree k for RSF (top) and ER (bottom) model ($N = 10^4$). For the ER model, two blocks of data are plotted, for $\overline{k} = 20$ (left) and for $\overline{k} = 100$ (right) The target density is fixed ($\rho_T = 0.1$), and $N_S = 2$ (circles), 10 (squares), 20 (triangles). The dashed lines represent the analytical prediction $2\epsilon + \rho_S \rho_T \overline{b(k)}$ in perfect agreement with the simulations.

a function of the degree k for both homogeneous (ER) and heterogeneous (RSF) graphs. For both models, the behaviors are in good agreement with the mean-field prediction, showing the tight relation between redundancy and betweenness centrality. In the case of heavy-tailed underlying networks, the vertex redundancy typically grows as a power-law of the degree, while the values for random graphs vary on a smaller scale. This behavior points out that the intrinsic hierarchical structure of scale-free networks plays a fundamental role even in the process of path routing, resulting in a huge number of probes iteratively passing through the same set of few hubs. On the other hand, for homogeneous graphs the total number of vertex discoveries is quite uniformly distributed on the whole range of connectivity, independently of the relative importance of the vertices.

5 Degree distribution measurements

A very important quantity in the study of the statistical accuracy of the sampled graph is the degree distribution. Fig. 3 shows the cumulative degree distribution $P_c(k^* > k)$ of the sampled graph defined by the ER model for increasing density of targets and sources. Sampled distributions are only approximating the genuine distribution, however, for $N_S \geq 2$ they are far from true heavy-tail distributions at any appreciable level of probing. Indeed, the distribution runs generally over a small range of degrees, with a cut-off that sets in at the average degree $\overline{k}$ of the underlying graph. In order to stretch the distribution range, homogeneous graphs with very large average degree

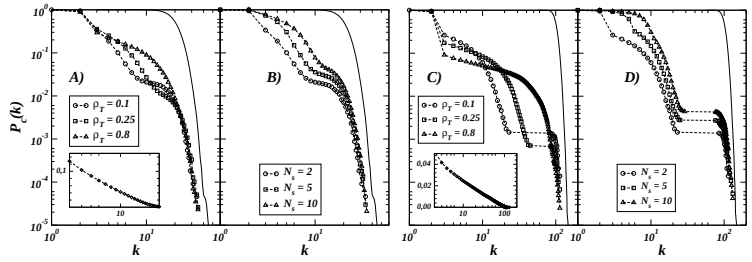


Fig. 3. Cumulative degree distribution of the sampled ER graph for USP probes. Figures A) and B) correspond to $\overline{k} = 20$, and C) and D) to $\overline{k} = 100$. Figures A) and C) show sampled distributions obtained with $N_S = 2$ and varying density target ρ_T. In the insets we report the peculiar case $N_S = 1$ that provides an apparent power-law behavior with exponent -1 at all values of ρ_T, with a cut-off depending on $\overline{k}$. The insets are in lin-log scale to show the logarithmic behavior of the corresponding cumulative distribution. Figures B) and D) correspond to $\rho_T = 0.1$ and varying number of sources N_S. The solid lines are the degree distributions of the underlying graph. For $\overline{k} = 100$, the sampled cumulative distributions display plateaus corresponding to peaks in the degree distributions, induced by the sampling process.

$\overline{k}$ must be considered; however, other distinctive spurious effects appear in this case. In particular, since the best sampling occurs around the high degree values, the distributions develop peaks that show in the cumulative distribution as plateaus. Note that, in the case of RSP and ASP model, the obtained distributions are closer to the real one since they allow a larger number of discoveries.

Only in the peculiar case of $N_S = 1$ an apparent scale-free behavior with slope -1 is observed for all target densities ρ_T, as analytically shown by Clauset and Moore [20,23]. Also in this case, the distribution cut-off is consistently determined by the average degree $\overline{k}$. The present analysis shows that in order to obtain a sampled graph with apparent scale-free behavior on a degree range varying over n orders of magnitude we would need the very peculiar sampling of a homogeneous underlying graph with an average degree $\overline{k} \simeq 10^n$; a rather unrealistic situation in the Internet and many other information systems where $n \geq 2$.

Since, in heterogeneous graphs, vertices with high degree are efficiently sampled with an effective measured degree that is rather close to the real one, the degree distribution tail is fairly well sampled, while deviations should be expected at lower degree values. This is indeed what we observe in numerical experiments on graphs with heavy-tailed distributions (see Fig. 4). Despite both RSF and WEI underlying graphs have a small average degree, the observed degree distribution spans more than two orders of magnitude. The distribution tail is fairly reproduced even at rather small values of ϵ. The data shows clearly that the low degree regime is instead under-sampled. This undersampling can either yield an apparent change in the exponent of the degree

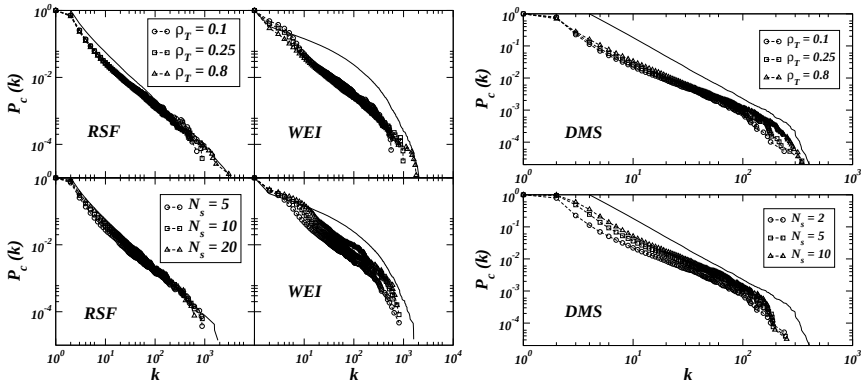


Fig. 4. Cumulative degree distributions of the sampled RSF, WEI and DMS graphs for USP probes. The top figures show sampled distributions obtained with $N_S = 5$ and varying density target ρ_T. The figures on the bottom correspond to $\rho_T = 0.25$ and varying number of sources N_S. The solid lines are the degree distributions of the underlying graph.

distribution (as also noticed in [21] for single source experiments), or, if N_S is small, yield a power-law like distribution for an underlying Weibull distribution. Furthermore, as Fig. 4 shows, an increase in the number of sources starts to discriminate between scale-free and Weibull distributions by detecting a curvature in the second case even at small values $\rho_T = 0.25$. It is, however, fair to say that while the experiments clearly point out a broad and heavy-tailed distribution, the distinction between different types of heavy-tailed distribution needs an adequate level of probing.

In conclusion, graphs with heavy-tailed degree distribution allow a better qualitative representation of their statistical features in sampling experiments. Indeed, the most important properties of these graphs are related to the heavy-tail part of the statistical distributions that are indeed well discriminated by the `traceroute`-like exploration. On the other hand, the accurate identification of the distribution forms requires a fair level of sampling that it is not clear how to determine quantitatively in the case of an unknown underlying network. We will discuss the implications of these results in real Internet measurements in Sec. 7.

6 Optimization of mapping strategies

In the previous sections we have shown that it is possible to have a general qualitative understanding of the efficiency of network exploration and the induced biases on the statistical properties. The quantitative analysis of the sampling strategies, however, is a much harder task that calls for a detailed study of the discovered proportion of the underlying graph and the precise

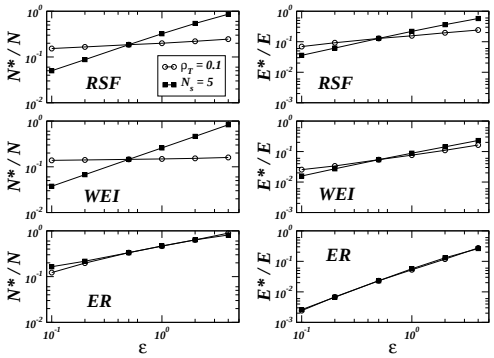


Fig. 5. Behavior of the fraction of discovered nodes and edges in explorations with increasing ϵ, for RSF, WEI and ER graphs. For each underlying graph studied we report two curves corresponding to larger ϵ achieved by increasing the target density ρ_T at constant $N_S = 5$ (squares) or the number of sources N_S at constant $\rho_T = 0.1$ (circles). Curves similar to ER are obtained for WS, and to RSF for BA and DMS.

deployment of sources and targets. In this perspective, very important quantities are the fraction N^*/N and E^*/E of vertices and edges discovered in the sampled graph, respectively. Unfortunately, the mean-field approximation breaks down when we aim at a quantitative representation of the results. The neglected correlations are in fact very important for the precise estimate of the various quantities of interest. For this reason we performed an extensive set of numerical explorations aimed at a fine determination of the level of sampling achieved for different experimental setups.

In Fig. 5 we report the proportion of discovered nodes and edges in the numerical exploration of the graph models defined previously for increasing level of probing ϵ. The level of probing is increased either by raising the number of sources at fixed target density or by raising the target density at fixed number of sources. As expected, both strategies are progressively more efficient with increasing levels of probing. In heterogeneous graphs, it is also possible to see that when the number of sources is $N_S \sim \mathcal{O}(1)$ the increase of the number of targets achieves better sampling than increasing the deployed sources. On the other hand, it is easy to perceive that the shortest path route mapping is a symmetric process if we exchange sources with targets. This is confirmed by numerical experiments in which we use a very large number of sources and a density of targets $\rho_T \sim \mathcal{O}(1/N)$, where the trends are opposite: the increase of the number of sources achieves better sampling than increasing the deployed targets.

In Fig. 6, we report the behavior of E^*/E and N^*/N at fixed ϵ and varying N_S and ρ_T. Very interestingly, the curves show a structure allowing for local minima and maxima in the discovered portion of the underlying graph: at fixed levels of probing ϵ, different proportions of sources and targets may achieve different levels of sampling. This hints to the search for optimal strategies in

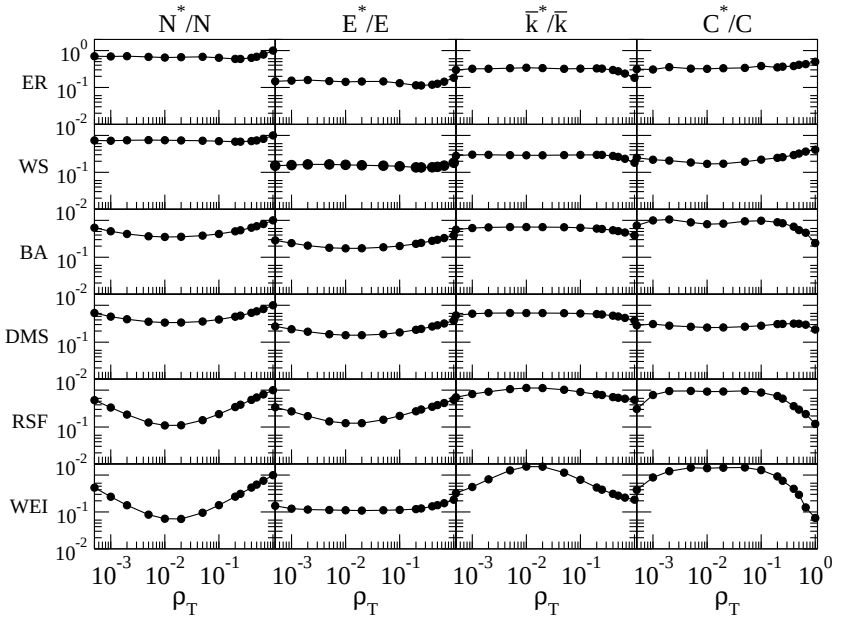


Fig. 6. Explorations with fixed ϵ (here $\epsilon = 2$): Behavior as a function of ρ_T of the fraction of discovered vertices N^*/N and edges E^*/E, of the normalized average degree $\overline{k}^*/\overline{k}$ and of the fraction of the normalized average clustering coefficient C^*/C. Since $\epsilon = \rho_T N_S$, the increase of ρ_T corresponds to a lowering of the number of sources N_S.

the relative deployment of sources and targets. The picture, however, is more complicate if we look at other quantities in the sampled graph. In Fig.6 we show the behavior at fixed ϵ of the average degree $\overline{k}^*$ measured in sampled graphs normalized by the actual average degree $\overline{k}$ of the underlying graph as a function of ρ_T. The plot shows also in this case a symmetric structure. By comparing the data of Fig.6 we notice that the symmetry point is of a different nature for different quantities: the minimum in the fraction of discovered edges corresponds to the best estimate of the average degree. This implies that at the symmetry point the exploration discovers less edges than in other setups, however, achieving a more efficient sampling of the effective degree for the discovered vertices. A similar problem is obtained by studying the behavior of the ratio C^*/C between the clustering coefficient of the sampled and the underlying graphs: the best level of sampling is achieved at particular values of ϵ and N_S that are conflicting with the best sampling of other quantities.

The evidence purported in this section hints to a possible optimization of the sampling strategy. The optimal solution, however, appears as a trade-off strategy between the different level of efficiency achieved in competing ranges of the experimental setup. In this respect, a detailed and quantitative investigation of the various quantities of interest in different experimental

setups is needed in order to pinpoint the most efficient deployment of source-target pairs depending on the underlying graph topology. While such a detailed analysis lies beyond the scope of the present study, an interesting hint comes from the analytical results of Section 3: since vertices with large betweenness have typically a very large probability of being discovered, placing the sources and targets preferentially on low-betweenness vertices (the most difficult to discover) may have an impact on the whole process. The usual correlation between connectivity and betweenness thus indicates that the exploration of a real network could be improved by a massive deployment of sources using low-connectivity vertices.

7 Conclusions and outlook

The rationalization of the sampling biases at the statistical level provides a general interpretative framework for the results obtained from the numerical experiments on graph models. The sampled graph clearly distinguishes between homogeneous and heavy-tailed topologies. This is due to the exploration process that statistically focuses on high betweenness vertices, thus providing a very accurate sampling of the distribution tail. In graphs with heavy-tails, such as scale-free networks, the main topological features are therefore easily discriminated since the relevant statistical information is encapsulated in the degree distribution tail which is fairly well captured. Quite surprisingly, the sampling of homogeneous graphs appears more cumbersome than those of heavy-tailed graphs. Dramatic effects such as the existence of apparent power-laws, however, are found only in very peculiar cases. In general, exploration strategies provide sampled distributions with enough signatures to distinguish at the statistical level between graphs with different topologies.

This evidence might be relevant in the discussion of real data from Internet mapping projects. Indeed, data indicate the presence of heavy-tailed degree distribution both at the router and AS level. The present discussion indicates that it is very unlikely that this feature is just an artifact of the mapping strategies. The upper degree cut-off at the router and AS level runs up to 10^2 and 10^3, respectively. A homogeneous graph should have an average degree comparable to the measured cut-off, which is hardly conceivable in a realistic perspective (for instance, it would require that nine routers over ten would have more than 100 links to other routers). In addition, the major part of mapping projects are multi-source, a feature that readily washes out the presence of spurious power-law behavior. On the contrary, heterogeneous networks with heavy-tailed degree distributions are sampled with particular accuracy for the large degree part, generally at all probing levels. This makes very plausible, and a natural consequence, that the heavy-tail behavior observed in real mapping experiments is a genuine feature of the Internet.

On the other hand, it is important to stress that while at the qualitative level the sampled graphs allow a discrimination of the statistical properties, at the quantitative level they might exhibit considerable deviations from the true values such as size, average degree, and the precise analytic form of the heavy-tailed degree distribution. For instance, the exponent of the power-law behavior appears to suffer from noticeable biases. In this respect, it is of major importance to define strategies that optimize the estimate of the various parameters and quantities of the underlying graph. In this paper we have shown that the proportion of sources and targets may have an impact on the accuracy of the measurements even if the number of total probes imposed to the system is the same. For instance, the deployment of a highly distributed infrastructure of sources probing a limited number of targets may result as efficient as few very powerful sources probing a large fraction of the addressable space [34,35]. The optimization of large network sampling is therefore an open problem that calls for further work aimed at a more quantitative assessment of the mapping strategies both on the analytic and numerical side.

Acknowledgments

We are grateful to M. Crovella, P. De Los Rios, T. Erlebach, T. Friedman, M. Latapy and T. Petermann for very useful discussions and comments. This work has been partially supported by the European Commission Fet-Open project COSIN IST-2001-33555 and contract 001907 (DELIS).

References

[1] The National Laboratory for Applied Network Research (NLANR), sponsored by the National Science Foundation. (see http://moat.nlanr.net/)

[2] The Cooperative Association for Internet Data Analysis (CAIDA), located at the San Diego Supercomputer Center. (see http://www.caida.org/home/).

[3] Topology project, Electric Engineering and Computer Science Department, University of Michigan (http://topology.eecs.umich.edu/).

[4] SCAN project at the Information Sciences Institute (http://www.isi.edu/div7/scan/).

[5] Internet mapping project at Lucent Bell Labs (http://www.cs.bell-labs.com/who/ches/map/).

[6] Faloutsos M, Faloutsos P, Faloutsos C: On Power-law Relationships of the Internet Topology. ACM SIGCOMM '99, Comput Commun Rev 1999; 29: 251-262.

[7] Govindan R, Tangmunarunkit H: Heuristics for Internet Map Discovery. Proc. of IEEE Infocom 2000; Volume 3, IEEE Computer Society Press: 1371–1380.

[8] Broido A, Claffy KC: Internet topology: connectivity of IP graphs. San Diego Proceedings of SPIE International symposium on Convergence of IT and Communication. Denver, CO. 2001

[9] Caldarelli G, Marchetti R, Pietronero L: The Fractal Properties of Internet. Europhys Lett 2000; 52: 386.

[10] Pastor-Satorras R, Vázquez A, Vespignani A: Dynamical and Correlation Properties of the Internet. Phys Rev Lett 2001; 87: 258701. Vázquez A, Pastor-Satorras R, Vespignani A: Large-scale topological and dynamical properties of the Internet. Phys Rev E 2002; 65: 066130.

[11] Chen Q, Chang H, Govindan R, Jamin S, Shenker SJ, Willinger W: The Origin of Power Laws in Internet Topologies Revisited. Proceedings of IEEE Infocom 2002, New York, USA.

[12] Medina A, Matta I: BRITE: a flexible generator of Internet topologies. Tech. Rep. BU-CS-TR-2000-005, Boston University, 2000.

[13] Jin C, Chen Q, Jamin S: INET: Internet topology generators. Tech. Rep. CSE-TR-433-00, EECS Dept., University of Michigan, 2000.

[14] Dorogovtsev SN, Mendes JFF: Evolution of networks: From biological nets to the Internet and WWW, Oxford University Press, Oxford, 2003.

[15] Baldi P, Frasconi P, Smyth P: Modeling the Internet and the Web: Probabilistic methods and algorithms, Wiley, Chichester, 2003.

[16] Pastor-Satorras R, Vespignani A: Evolution and structure of the Internet: A statistical physics approach, Cambridge University Press, Cambridge, 2004.

[17] Burch H, Cheswick B: Mapping the internet. IEEE computer 1999; 32: 97-98.

[18] Willinger W, Govindan R, Jamin S, Paxson V, Shenker S: Scaling phenomena in the Internet: Critically examining criticality. Proc Natl Acad Sci USA 2002; 99: 2573–2580.

[19] Lakhina A, Byers JW, Crovella M, Xie P: Sampling Biases in IP Topology Measurements. Technical Report BUCS-TR-2002-021, Department of Computer Sciences, Boston University (2002).

[20] Clauset A, Moore C: Accuracy and Scaling Phenomena in Internet Mapping. Phys Rev Lett 2005; 94: 018701.

[21] Petermann T, De Los Rios P: Exploration of Scale-Free Networks - Do we measure the real exponents? Eur Phys J B 2004; 38: 201-204.

[22] Guillaume J-L, Latapy M: Relevance of Massively Distributed Explorations of the Internet Topology: Simulation Results. Proc. Infocom 2005 (to appear).

[23] Achlioptas D, Clauset A, Kempe D, Moore C: On the Bias of Traceroute Sampling; or, Power-law Degree Distributions in Regular Graphs. cond-mat/0503087, to appear in STOC 2005.

[24] Erdös P, Rényi P: On random graphs I. Publ Math Debrecen 1959; 6: 290.

[25] Freeman LC: A Set of Measures of Centrality Based on Betweenness. Sociometry 1977; 40: 35-41.

[26] Brandes U: A Faster Algorithm for Betweenness Centrality. J Math Soc 2001; 25: 163-177.

[27] Goh K-I, Kahng B, Kim D: Universal Behavior of Load Distribution in Scale-Free Networks. Phys Rev Lett 2001; 87: 278701.

[28] Watts DJ, Strogatz SH: Collective dynamics of small-world networks. Nature 1998; 393: 440-442.

[29] Barabási A-L, Albert R: Emergence of scaling in random networks. Science 1999; 286: 509-512.

[30] Dorogovtsev SN, Mendes JFF, Samukhin AN: Size-dependent degree distribution of a scale-free growing network. Phys Rev E 2001; 63: 062101.

[31] Molloy M, Reed B: A critical point for random graphs with a given degree sequence. Random Struct Algorithms 1995; 6: 161. Molloy M, Reed B: The size of the giant component of a random graph with a given degree distribution. Combinatorics, Probab Comput. 1998; 7: 295.

[32] Dall'Asta L, Alvarez-Hamelin I, Barrat A, Vázquez A, Vespignani A: Statistical theory of Internet exploration. Phys Rev E 2005; 71: 036135.

[33] Barthélemy M: Betweenness Centrality in Large Complex Networks. Eur Phys J B 2004; 38: 163-168.

[34] http://www.netdimes.org/ ; Shavitt Y, Shir E: DIMES: Let the Internet Measure Itself. preprint cs.NI/0506099.

[35] http://www.tracerouteathome.net/

Traffic dynamics in scale-free networks

Attila Fekete*
Collegium Budapest, Institute for Advanced Studies
Szentháromság u. 2., H-1014 Budapest, Hungary
Tel: (+36 1) 224 8338 Fax: (+36 1) 224 8310

Gábor Vattay†
Collegium Budapest, Institute for Advanced Studies
Szentháromság u. 2., H-1014 Budapest, Hungary

Ljupco Kocarev‡
University of California, Institute for Nonlinear Sciences
9500 Gilman Drive, La Jolla, CA 92093, San Diego, USA

We study traffic dynamics in growing scale-free networks. Both the scale-free structure of the network and the adaptive nature of the dynamics which controls traffic in the network are considered in the model. The model is investigated with computer simulations and analytically for the case of scale-free tree. For the scale-free tree, an exact formula and its power law approximation of the complementary cumulative distribution function (CDF) of link load (edge betweenness) is presented. We examine whether the scaling properties of the network affect the performance of the transport mechanism and estimate the average number of competing transport mechanisms at bottlenecks. We find that bottlenecks tend to appear on the periphery of the network as the performance increases. Various bandwidth allocation strategies are compared. We show that the best performance is achieved when capacity is distributed proportionally to the expected load of links. We demonstrate that it is necessary to study both the topology and the dynamics of the transport mechanism to understand the whole system.

Keywords: scale-free, network, TCP, Internet, simulations

*Electronic address: fekete@complex.elte.hu
†Electronic address: vattay@elte.hu
‡Electronic address: lkocarev@ucsd.edu

I. INTRODUCTION

The statistical properties of complex networks has been investigated extensively in the physics community in recent years [1–4]. With the increasing computing power of modern computers, analysis of large-scale networks and databases has become possible. It has been shown that the degree statistics of many natural and artificial networks follows power law. Examples for such networks vary from social interconnections and scientific collaborations [5] to the world-wide web [6] and the Internet [7, 8]. These networks are usually referred to as *scale-free* networks.

The first mathematical model of complex networks, the random graph theory was developed by Erdős and Rényi (ER) [9]. In this model, the number of nodes is fixed and connections are established randomly. Although the ER model leads to rich theory, it fails to predict the power law distributions observed in scale-free networks. Barabási and Albert (BA) proposed a more suitable model of these networks [10, 11]. The BA model is also based on the random graph theory, but, in addition, it involves two key principles: (a) *growth*, that is, the size of the network increases during development, and (b) *preferential attachment*, that is, new network elements are connected to higher degree nodes with higher probability.

The concepts of graph theory are used throughout this paper. A graph consists of vertices (nodes) and edges (links). Edges are ordered or un-ordered pairs of vertices, depending on whether ordered or un-ordered graph is considered, respectively. The order of a graph is the number of vertices it holds, while the degree of a vertex counts the number of edges adjacent to it. A path is also defined by the most natural way: it is a vertex sequence, where any two consecutive elements form an edge. The graph is called connected, if for any vertex pair there exists a path which starts from one vertex and ends at the other.

The study of complex networks usually deals with the structural properties of networks, like degree distribution, shortest path distribution, degree–degree correlations, clustering, etc. Furthermore, some complex networks also involve a dynamical system which governs traffic in the network. The matter of importance in such systems is the performance of the dynamics. Therefore, exploring the influence of network structure upon traffic dynamics is essential. Moreover, one should be interested in distributing the available resources to obtain the best performance for a given network structure.

From this point of view "betweenness" is the most important attribute. Betweenness measures the number of shortest paths passing through a certain network element. *Node betweenness* has been studied recently by Goh, Kahng, and Kim [12], who argued that it follows power law in scale-free networks, and the exponent $\delta \approx 2.2$ is independent from, in a certain range, the degree distribution. Szabó, Alava, and Kertész [13] used rooted deterministic trees to model scale-free BA trees, and found scaling exponent $\delta_t = -2$.

The study of complex networks usually deals with the structural properties of networks, like degree distribution, shortest path distribution, degree–degree correlations, clustering, etc. Furthermore, a complex network may also involve a dynamical system which governs traffic in the network. In this paper we study scale-free networks with embedded flow dynamics. The dominant algorithm which controls the data traffic in the Internet is the Transmission Control Protocol (TCP) [14]. For the detailed analysis of TCP mechanism we refer to Ref. [15]. Since TCP performance affects overall network performance, TCP modeling is an important issue that has been attracted research interests during the last years. Traditional approaches to performance evaluation packet networks have normally relied on attempts to describe as closely as possible the dynamics of network elements over a discrete state-space. A new class of semi-analytical models has recently been introduced in the networking arena, and today appears to be the most promising approach for scalable and accurate performance analysis of large IP networks. This new approach, that is often called 'fluid models', adopts an abstract deterministic description of the average network dynamics through a set of ordinary differential equations, thus neglecting the short term, packet-by-packet description of the stochastic network dynamics. We will present a simple model, which considers both the scale-free structure of the Internet and the adaptive nature of the underlying dynamics using fluid models. We stress that the main goal of this paper is to study the TCP-like (adaptive) dynamics on growing scale-free networks, not to model the Internet.

II. THE NETWORK MODEL

It has been shown that the structure of autonomous systems (AS) in the Internet is scale-free tree [16]. An AS is a large segment of the Internet, which usually belongs to one organization, for example to a university, a large company, a national office, etc. In order to keep our model analytically tractable, we model the whole Internet with a simple scale-free rooted BA tree, extended with initial attractiveness [17]. Shortcuts, correlations with the geographical distribution of the population [18], and other details are neglected in our model.

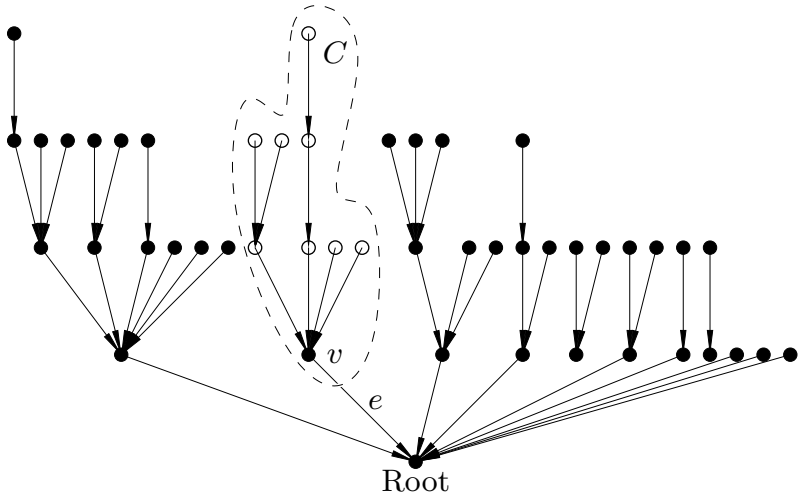


FIG. 1: Schematic illustration of the evolving network at time τ. Vertex v, connected to the network at τ_e, denotes the root of cluster C. Variable $n = |C| - 1$ denotes the number of nodes in C, except v (marked by $\circ$'s).

A. The BA model

The construction of the network proceeds in discrete time steps according to the BA model. Let us denote time with $\tau \in \mathbb{N}$. Initially, at $\tau = 0$, the graph consists only of a single vertex without any edges. Then, in every time step, a new vertex is connected to the network with a single, *directed* edge. Note that the initial vertex is distinguished from all the other, since it has only incoming connections; we refer to it as *root vertex*. The target of the new edge is selected randomly from the present vertices of the graph. The probability that a new vertex connects to an old one is proportional to the attractiveness of the old vertex v. Attractiveness is defined as

$$A(v) = a + q,$$

where parameter $a > 0$ denotes the initial attractiveness and q is the in-degree of vertex v. The scaling properties of the network can be smoothly controlled by parameter a: it has been shown that the probability distribution of in-degrees is $P(q) \sim (q + a)^{-(2+a)}$ [17]. Note that the special case $a = 1$ practically repeats the original BA model. Indeed, except for the root node, the attractiveness of every vertex becomes equal to its degree if $a = 1$; this is exactly the definition of the attractiveness in the BA model [10]. On the other hand, the model tends to Poisson-type ER graph if $a \to \infty$, since preferential attachment disappears in the limit, and $P(q) \sim e^{-q}$.

We refer to a connected sub-graph as a *cluster* in this paper (Fig. 1). To calculate the number of shortest paths passing through a given edge, it is sufficient to know the size of the cluster attached to the given edge n. If the size of the network is N, then from elementary combinatorics it follows that the number of shortest paths, that is the betweenness, or shortly the load of the particular edge is

$$L = (n + 1)(N - n - 1). \tag{1}$$

The probability distribution of cluster size for any finite N can be given exactly:

$$\mathbb{P}_N(n) = \frac{N - \alpha}{N - 1} \frac{1 - \alpha}{(n + 1 - \alpha)(n + 2 - \alpha)}, \tag{2}$$

where $0 \le n < N - 1$, and $\alpha = 1/(1 + a)$. The details of the calculations will be published elsewhere.

For $1 \ll N$ and $1 \ll n \ll N$ Eq. (2) can be approximated with

$$\mathbb{P}_N(n) \approx (1 - \alpha) \frac{1}{n^2},$$

where the scaling exponent $\nu = 2$ is independent of α, therefore it is universal in the class of evolving scale-free trees.

B. Betweenness

The probability distribution of betweenness L can be given by the following transformation formula of random variables:

$$\mathbb{P}_N(L) = \sum_{n=0}^{N-2} \delta_{L,(n+1)(N-n-1)} \mathbb{P}_N(n).$$

However, this expression is difficult to handle. An alternative description of a random variable is the complementary cumulative distribution function (CDF), defined as $F^c(x) = \mathbb{P}(L \geq x)$, that is the probability that the value of random variable exceeds x.

From (1) it is obvious that the load of an edge exceeds L if and only if $n_L \leq n < N - (n_L + 1)$, where

$$n_L = \left\lceil \frac{N-2}{2} - \frac{N}{2}\sqrt{1 - \frac{4L}{N^2}} \right\rceil,$$

and $\lceil \cdot \rceil$ denotes the ceil function. It immediately follows that the complementary CDF of the load is

$$F_N^c(L) = \sum_{n=n_L}^{N-n_L+2} \mathbb{P}_N(n) = \frac{N-\alpha}{N-1} \frac{(1-\alpha)(N-2n_L-1)}{(n_L+1-\alpha)(N-n_l-\alpha)}. \tag{3}$$

If $N \ll L \ll N^2/4$ then the complementary CDF can be approximated by the following power law

$$F_N^c(L) \approx (1-\alpha) N \frac{1}{L}.$$

Finally, the expectation value of the edge betweenness is calculated:

$$\begin{aligned}\mathbb{E}_N[L] &= \sum_{L=0}^{\infty} L\mathbb{P}_N(L) = \sum_{n=0}^{N-2} (n+1)(N-n-1)\mathbb{P}_N(n) \\ &= (1-\alpha)\frac{(N-\alpha)(N+1-2\alpha)}{N+1}\left[\Psi(N-\alpha) - \Psi(1-\alpha)\right] - (1-\alpha)(2N+1-2\alpha),\end{aligned}$$

where $\Psi(z) = d\left[\ln\Gamma(z)\right]/dz$ is the digamma function [19]. Since $\Psi(z) \sim \ln z$ as $z \to \infty$, therefore

$$\mathbb{E}_N[L] = (1-\alpha) N \ln N + O(N),$$

if $N \to \infty$.

III. THE MODEL OF NETWORK DYNAMICS IN THE INTERNET

Data is transferred between source and target computers through intermediate *routers*. Before transmission data is cut into smaller units, called *packets*. This way, if some part of the file is lost or gets corrupted, then only the damaged or lost parts should be retransmitted, not the whole file. The TCP algorithm administers the departure, arrival and retransmission of packets.

Interactions of different TCP flows inevitably cause congestion in the network. Packets are temporarily queued in buffers, but when a buffer is full, incoming packets are dropped by routers. When a packet successfully reaches its destination, the receiver sends an acknowledgement (ACK) packet back to the source. The elapsed time between packet departure and ACK arrival is called *round-trip time*, T_{RTT}.

An important feature of the TCP algorithm is that it can adapt its throughput to the changing network conditions. The throughput, that is, the amount of bits transferred per unit time, is increased when arriving ACK indicates successful transmission, and decreased when missing ACK implies congestion.

In this section the fluid approximation of the TCP algorithm, the Additive Increase–Multiplicative Decrease (AIMD) model is discussed, supposing that the topology of the network does not change. Modelling dynamics on a fixed topology is legitimate when the time scales describing the development of the network topology and the dynamical process superposed to the network differ widely. A good example is Internet traffic, whose modelling requires time resolutions from milliseconds up to a day [20–22], compared with the months required for significant topological changes [23].

Detailed description of AIMD model can be found in [24]. The TCP standard is given in [14].

A. The AIMD model

Let us suppose that N_{TCP} number of TCPs are operating in the network and their throughput is denoted by $X^{(1)}(t), X^{(2)}(t), \ldots, X^{(N_{\mathrm{TCP}}(t))}$. A heuristic, but reasonable assumption of the AIMD model is that between consecutive packet loss events the development of throughput $X^{(i)}$, $0 < i \leq N$ can be approximated by the following differential equation [25]:

$$\frac{dX^{(i)}(t)}{dt} = \frac{P}{T_{\mathrm{RTT}}^{(i)}(t)^2}, \tag{4}$$

where $T_{\mathrm{RTT}}^{(i)}(t)$ denotes the round-trip time of the ith TCP connection at time t, and P is the packet size. In fixed topology, round-trip times may vary due to queuing delays. If queuing delays are negligible, however, then round-trip times are constants $T_{\mathrm{RTT}}^{(i)}(t) \equiv T_{\mathrm{RTT}}^{(i)}$, and (4) can be solved:

$$X^{(i)}(t) = X^{(i)}(0) + \frac{P}{{T_{\mathrm{RTT}}^{(i)}}^2} t. \tag{5}$$

Note that (4) applicable only if the packet loss ratio is modest ($< 1 - 2\%$) [26].

Eqs. (4) and (5) are valid only between packet loss events t_n, which occur when the total throughput on edge e first reaches capacity C_e of that particular edge:

$$\sum_{i \in I_e} \left(X_n^{(i)} + \frac{P}{{T_{\mathrm{RTT}}^{(i)}}^2} \Delta t_{n+1} \right) = C_e, \tag{6}$$

where I_e denotes the set of TCPs which share edge e, $X_n^{(i)} = X^{(i)}(t_n)$ is the throughput of the ith TCP at t_n, and $\Delta t_n = t_n - t_{n-1}$ is the elapsed time between the nth and the previous congestion events. The first moment when (6) holds is

$$\Delta t_{n+1} = \min_e \left(\frac{C_e - \sum_{i \in I_e} X_n^{(i)}}{\sum_{i \in I_e} P / {T_{\mathrm{RTT}}^{(i)}}^2} \right). \tag{7}$$

At congestion events, some TCPs that share the congested link lose packets. The AIMD model provides that the packet losses can be modeled by a stationary stochastic process; the owners of the lost packets are selected randomly and independently. The probability p_s that a TCP flow experience packet loss is called *synchronization parameter*.

According to the TCP congestion control mechanism, those TCPs which lose packets halve their throughput. The schematic time evolution of the total throughput and the throughput of a chosen TCP is shown in Fig. 2.

B. The model of TCP connections

The hosts of the source and the destination of the TCP connections are located randomly in the Internet. The actual location of the connections might be influenced by many factors including the importance and the availability of the computers, the user's language, behavior and preference, etc.

We assume in our model that TCP connections are established randomly in the network, and the distribution of both the source and the destination of the TCPs are homogeneous. That is, every pair of nodes may establish a directed TCP connection with the same, uniform probability, $p = \frac{1}{N-1}$. Therefore, the average number of TCP connections is $\mathbb{E}[N_{\mathrm{TCP}}] = N$ in the network. Moreover, data transfers are considered to be persistent in our model. For a more realistic model, one should take finite file sizes and the heterogeneous TCP connections into consideration.

IV. DISCUSSIONS

The model we outlined in the previous sections was studied with extensive numerical simulations. First, we validate the analytic results that we obtained in Sec. II. Then, the influence of the network topology on the performance is discussed.

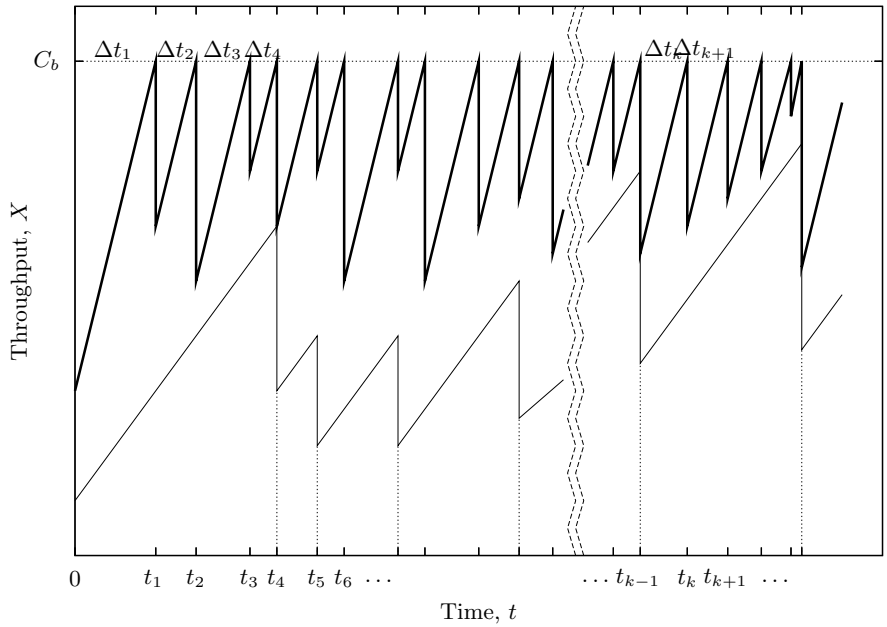


FIG. 2: Schematic time evolution of the total throughput (thick line) and the throughput of one TCP (thin line) in the AIMD model for constant round-trip time. Packet losses occur at $t_1, t_2, \ldots$ moments, when the total throughput reaches the capacity of the bottleneck link C_b. Time intervals between consecutive packet loss events are $\Delta t_1, \Delta t_2, \ldots$.

A. Validation of edge betweenness distribution

The exact formula (3) presented in Sec. II B shows that possible values of the edge betweenness are $L = (N-1), 2(N-2), 3(N-3), \ldots, [N/2](N-[N/2])$, that is, the CDF is constant between the above integer values.

Simulations confirm the validity of Eq. (2). With computer simulations $N = 10^5$ node random networks were generated with $\alpha = 0$ (ER), $\alpha = 1/3$, $\alpha = 1/2$ and $\alpha = 2/3$ (BA) parameter values. Empirical distributions, obtained from 100 independent simulations, formula (3) and power law approximations are compared in Fig. 3.

The expected staircase structure of the distribution can be clearly seen. The power law approximation fits the complementary CDF in the range $N \ll L \ll N^2/4$ accurately.

B. Influence of the network structure on TCP performance

We study in this section how the structure of the topology (α) affects the "performance" of the TCPs in the network. Let us define performance first: if N_{TCP} number of TCPs are operating in the network where bandwidths $\{C_e\}$ have been allocated to the links, then the performance of the ith TCP, $Q^{(i)}$, is defined as the time average of its throughput $X^{(i)}(t)$:

$$Q^{(i)} = \bar{X}^{(i)} \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t X^{(i)}(u)\, du.$$

Let us consider m_e number of TCPs, which utilize a bottleneck link and let us suppose that the competing TCPs share the available bandwidth equally. It has been shown [24] that the expected performance of such TCPs is

$$\mathbb{E}\left[Q^{(i)}\right] = \left(1 - \frac{p_s}{2}\right) \frac{C_e}{m_e},$$

where C_e denotes the capacity of the bottleneck link, and p_s is the synchronization parameter, introduced in the AIMD model above. For the sake of simplicity, let the synchronization parameter be so small that only one packet is

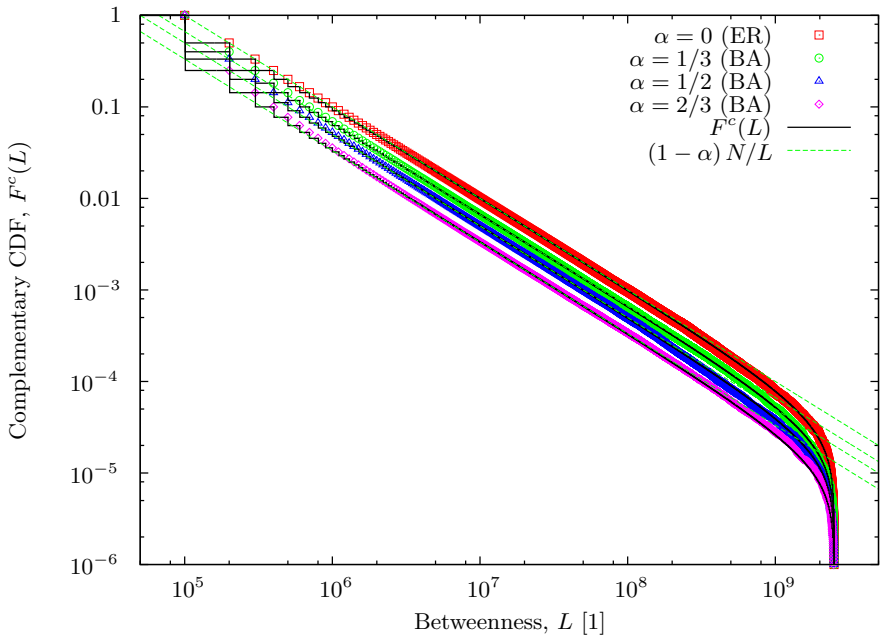


FIG. 3: Complementary CDF of edge betweenness, obtained from 100 realizations of $N = 10^5$ node networks, is shown at $\alpha = 0$ (squares), $\alpha = 1/3$ (circles), $\alpha = 1/2$ (triangles), and $\alpha = 2/3$ (diamonds) parameter values. Equation (3) (solid line) and power law approximation $(1-\alpha)\,N/L$ (dotted line) are also plotted.

dropped at every congestion epoch: $p_s m_e \approx 1$. With this assumption

$$\mathbb{E}\left[Q^{(i)}\right] = \left(1 - \frac{1}{2m_e}\right)\frac{C_e}{m_e}. \tag{8}$$

The performance of the network is obviously influenced by the bandwidth distribution $\{C_e\}$. For the performance of different network structures to become comparable, the average bandwidth $\bar{C} = 1/(N-1)\sum_e C_e$ is fixed. Furthermore, the limited amount of capacity is distributed with the same strategy in networks with different scaling parameter α.

In case of homogeneous TCP distribution, the expected number of TCPs which share a given link is proportional to the betweenness L_e, that is the number of shortest paths passing through the particular link:

$$\mathbb{E}\left[m_e\right] = \frac{\mathbb{E}\left[N_{\mathrm{TCP}}\right]L_e}{N(N-1)} = \frac{L_e}{N-1},$$

where the expected number of TCPs of out model, $\mathbb{E}\left[N_{\mathrm{TCP}}\right] = N$ is substituted. In this section, mean field approximation is applied for distributing capacity, that is capacity is allocated proportionally to the edge betweenness: $C_e = C_0 L_e$. The normalization coefficient C_0 can be given by:

$$C_0 = \frac{\bar{C}}{\mathbb{E}_N\left[L\right]} \approx \frac{\bar{C}}{(1-\alpha)\,N\ln N} + O(1/N)$$

Using the above equations, the following formula can be obtained from Eq. (8) for the expected performance of the ith TCP:

$$\mathbb{E}\left[Q^{(i)}\right] \approx \left(1 - \frac{1}{2m}\right)\frac{\bar{C}}{(1-\alpha)\ln N}, \tag{9}$$

where m is the number of TCPs, including the ith TCP, which share the bottleneck link.

Network performance can be characterized by the complementary CDF of TCP performance: $F^c(Q) = \mathbb{P}\left(Q^{(i)} > Q\right)$. On Fig. 4 CDF of TCP performance is shown on normal-log plot for $N = 10^4$ node networks with $\alpha = 0,\ 1/3,\ 1/2,$

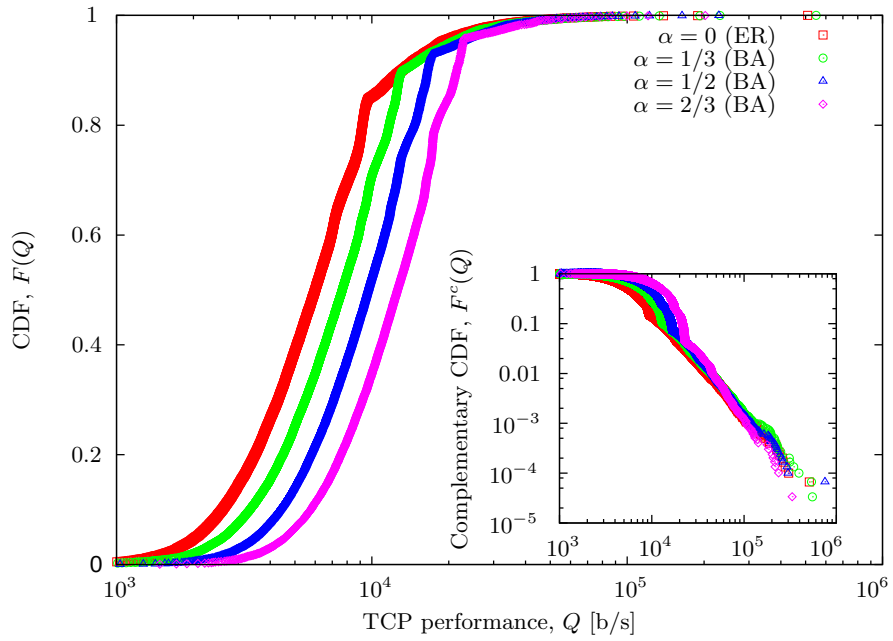


FIG. 4: CDF of TCP performance, obtained from three realizations of $N = 10^4$ node networks, is shown on normal-log plot at $\alpha = 0$ (squares), $\alpha = 1/3$ (circles), $\alpha = 1/2$ (triangles) and $\alpha = 2/3$ (diamonds) parameter values. Average capacity is set to $\bar{C} = 10^5[b/s]$ for every network. Simulation lasted for $100N$ congestion epochs. Inset: Complementary CDF of TCP performance on log-log plot.

2/3 parameter values. Empirical distributions were obtained from simulations running for $100N$ congestion epochs on three realizations of random networks of each α values. The mean of the link capacity was set to $\bar{C} = 10^5[b/s]$. Inset shows complementary CDF of TCP performance on log-log plot.

A point of inflection can be observed in Fig. 4 at every α parameter. The behavior of the CDF is markedly different below and above the point of inflection. The sharp difference in the CDF implies that TCPs can be divided into two categories according to whether their performance is over or below the point of inflection.

The tail of the complementary CDF of TCP performance, above the point of inflection (see inset of Fig. 4) consists of TCPs whose throughput is much higher than the expected performance (9). These TCP operate in the core of the network, where every link along the path of their connection has large bandwidth. Moreover, they either hardly need to compete with other TCPs for the available bandwidth, or they win the competition at congestion epochs. The relative number of such TCPs is approximately 15% if $\alpha = 0$, and it is decreasing with the growth of parameter α.

Below the point of inflection performances of TCPs are limited by low-bandwidth links, located on the periphery of the network, and by congested bottlenecks inside the network. If we suppose that the point of inflection approximately equals the expected throughput of TCPs at bottleneck links $\mathbb{E}\left[Q^{(i)}\right]$, then the number of TCPs competing at bottlenecks can be estimated from Eq. 9. In Table I the location of the point of inflection Q_I and the estimated number of TCPs at bottleneck link is shown for networks with different scaling parameter α. Estimates show that as scaling parameter α increases bottlenecks tend to form on the outer links where only 1–2 TCP share the links.

Overall performance of the networks, measured by the average TCP performance

$$Q = \frac{1}{N_{\mathrm{TCP}}} \sum_{i=0}^{N_{\mathrm{TCP}}} Q^{(i)},$$

is also shown in Table I. We found that the overall performance Q also increases with parameter α. It follows from above that the scaling properties of the topology influence the TCP performance. It is reasonable to suppose that the interaction between the topology and the dynamical system is mutual, that is the evolution of the network can by influenced by the dynamical system to reach optimum TCP performance as well.

TABLE I: Table shows average TCP performance Q, the point of inflection of the CDF of TCP performance Q_I and the estimated number of TCPs at bottlenecks for networks with different scaling parameter α.

α	$Q[b/s]$	$Q_I[b/s]$	m
0	7785	9655	4.52
0.3333	9259	13057	2.52
0.3891	9346	14456	2.68
0.5	11184	17200	2.40
0.6666	13790	23118	1.72

C. Performance of other bandwidth distribution strategies

In this section different bandwidth distribution scenarios are compared. The topology of the network and the average capacity is kept fixed, and only link capacities are changed in simulations. The scaling parameter of the topology is chosen to be $\alpha = 1/2$ for numerical simulations. Besides the mean field bandwidth distribution strategy discussed in the previous section, the following scenarios are considered:

Uniform Capacity is the same for every link: $C_e = \bar{C}$,

Minimum Capacity is proportional with the following minimum: $\min(q_A, q_B)$,

Maximum Capacity is proportional with the following maximum: $\max(q_A, q_B)$,

Product Capacity is proportional with the following product: $q_A \cdot q_B$,

where q_A and q_B denote the in-degrees of the nodes which compose a particular link. Uniform scenario presented as a reference. It can be considered as the worst case scenario, when no information is available on the details of the network. Minimum, maximum and product strategies take the local structure of the network into account, and the more connection the link possess, the more capacity they allocate for the particular link. The difference between the three strategies is whether they prefer loosely, moderately or highly connected links.

The complementary CDF of link capacities, obtained as the result of the above bandwidth distribution strategies, are compared in Fig. 5. CDF of the uniform strategy is degenerated, and the structure of the network is not taken into consideration this case. The maximum strategy prefers the lower bandwidths at the cost of a cutoff at about $10^6 b/s$ capacity. The minimum strategy also prefers lower bandwidths at the cost of high bandwidths, but no cutoff exists. The complementary CDF of minimum strategy also resembles the mean field distribution with a different scaling exponent. The product strategy prefers the mid-range, and it underestimates both the low and the high capacity range, compared to the mean field strategy.

Simulation results of the CDF of TCP performance is shown in Figure 6 for the above mentioned bandwidth distributions. The performance of mean field strategy is clearly the best. The next two best performing strategies, the minimum and the product perform almost the same, although they prefer completely different bandwidth ranges. It follows that the whole bandwidth range must be taken into consideration in any bandwidth distribution strategy to reach the optimum network performance. The performance of the maximum strategy is worse considerably than the previous two. Finally, the uniform bandwidth distribution is the worst of all: its performance is just a few percent of the mean field scenario's performance. The network where this strategy is applied is heavily congested, since the bottlenecks form in the core of the network.

Network performances, that is average TCP performances are shown in Table II for the different bandwidth distribution strategies. The measured network performances confirm the qualitative analysis of Fig. 6. Table II shows that mean field bandwidth allocation strategy is almost twice more effective than the second, minimum strategy, and it is more than twice as good as the product strategy. The performance of a network with maximum bandwidth distribution strategy is just about one fifth of performance of the same network when mean field strategy is used. Moreover, the performance of uniform scenario is even less then the third of the second worst, maximum strategy.

V. CONCLUSIONS

A complex model of the network embedded with dynamics has been studied in this paper. Both the scale-free structure of the network and the TCP dynamics, which controls traffic, are considered in the model. The topology of the network has been modeled by a growing scale-free random graph model, where the scaling properties of the

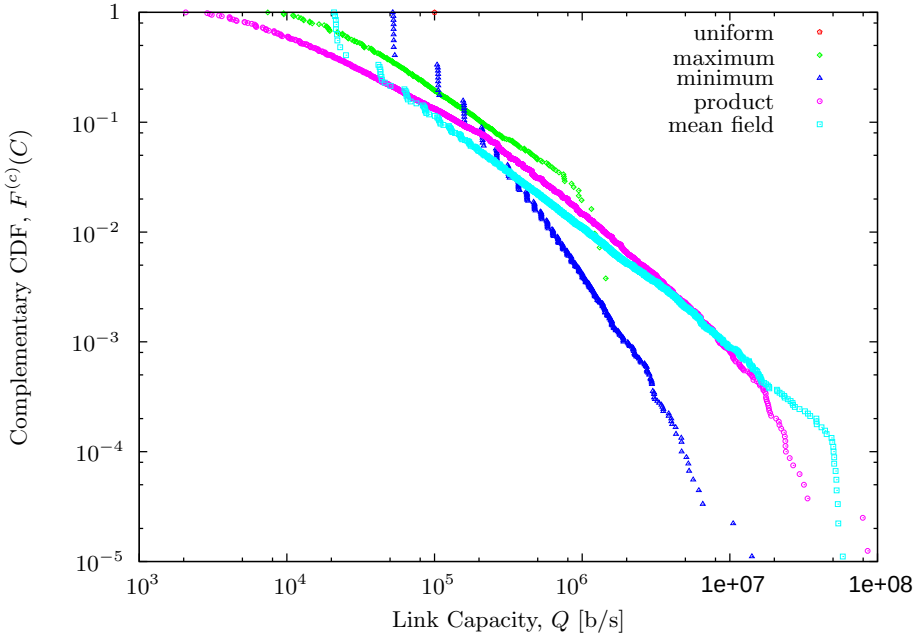


FIG. 5: Comparison of the complementary CDF of link capacity is shown for different bandwidth distribution strategies on log-log plot. Data is obtained from 10 realizations of $N = 10^4$ node networks with scaling parameter $\alpha = 1/2$. Average capacity is set to $\bar{C} = 10^5[b/s]$ for every network. The following scenarios are considered: uniform (pentagons), maximum (diamonds), minimum (triangles), product (circles), and mean field (squares).

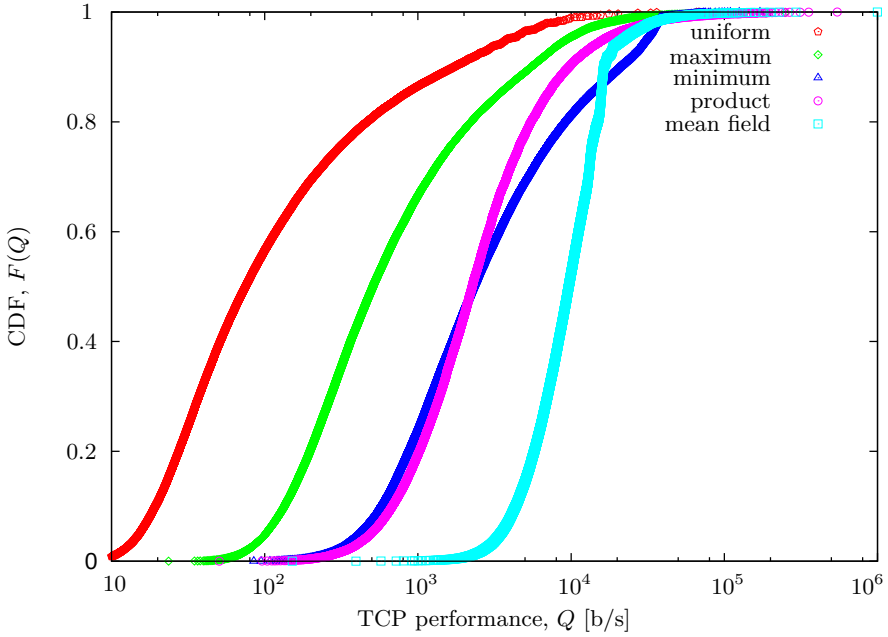


FIG. 6: Comparison of the CDF of TCP performance is shown for different bandwidth distribution strategies on normal-log plot. Data is obtained from 10 realizations of $N = 10^4$ node networks with scaling parameter $\alpha = 1/2$. Average capacity is set to $\bar{C} = 10^5[b/s]$ for every network. Simulation lasted for $100N$ congestion epochs. The following scenarios are considered: uniform (pentagons), maximum (diamonds), minimum (triangles), product (circles), and mean field (squares).

TABLE II: Table shows network performance for different bandwidth distribution strategies.

Strategy	$Q[b/s]$
Uniform	740.79
Maximum	2391.94
Minimum	6574.69
Product	5279.5
Mean field	11284.6

network can be changed with parameter α. The TCP dynamics has been approximated by the fluid AIMD model. We have assumed in the model that TCP connections are distributed homogeneously in the network, that is TCP connections are established between every pair of nodes with the same probability. It follows that the expected number of TCPs which utilize a link is proportional to the link load (edge betweenness), that is the number of shortest paths passing through the particular link.

We can summarize the main conclusions of this paper as follows:

- For the case of scale-free tree we analytically computed conditional cluster size distribution, total cluster size distribution, complementary CDF of link load (edge betweenness), and the expectation value of the link load (edge betweenness), see Section II B. The exact formula and the approximation for the complementary CDF of link load have been validated by numerical simulations in Section IV A.

- The purpose of TCP connections is to transfer data in the network. The performance of a TCP connection is measured by its throughput, that is its average transfer rate. We have investigated in Section IV B whether TCP performance is influenced by the scaling properties of the network. For the comparison of different networks the bandwidth allocation strategy has been fixed to the mean field strategy. It has been shown that the TCP performance increases as the scaling parameter α increases. It follows that the network topology influences the performance of the TCP.

- From the analysis of the CDF of TCP performance we have estimated the number of TCPs at bottleneck links. We have found that the number of TCPs at bottlenecks decreases as parameter α increases. It follows that bottlenecks move to the periphery of the network, when network performance is higher. This is understandable, since a bottleneck in the core of the network can reduce the performance of more TCP than a bottleneck on the periphery of the network.

- We have investigated TCP performance in networks which were built on various bandwidth allocation strategies in Section IV C. We have found that mean field strategy performs about twice as well than the minimum and the product strategies, five times as well than the maximum, and it is more than fifteen times as good as the uniform strategy. These results indicate that the mean field bandwidth distribution strategy provides the optimum TCP performance.

Acknowledgement

The authors thank the partial support of the National Science Fund Hungary (OTKA T37903), the National Office for Research and Technology (NKFP 02.032.04) and the EU IST FET Complexity EVERGROW Integrated Project.

[1] Strogatz SH, Nature **410**, 268 (2001).
[2] Albert R, Barabási A-L, Rev. Mod. Phys. **74**, 47 (2002).
[3] Dorogovtsev SN, JFF Mendes, Adv. Phys. **51**, 1079 (2002).
[4] Newman MEJ, SIAM Rev. **45**, 167 (2003).
[5] Newman MEJ, Phys. Rev. E **64**, 016131, 016132 (2001).
[6] Barabási A-L, Albert R, Jeong H, Physica A **281**, 69 (2000).
[7] Faloutsos M, Faloutsos P, Faloutsos C, Comput. Commun. Rev. **29**, 251 (1999).
[8] Pastor-Satorras R, Vázquez A, Vespignani A, Phys. Rev. Lett. **87**, 258701 (2001).
[9] Erdős P, Rényi A, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[10] Barabási A-L, Albert R, Science **286**, 509 (1999).
[11] Barabási A-L, Albert R, Jeong H, Physica A **272**, 173 (1999).
[12] Goh K-I, Kahng B, Kim D, Phys. Rev. Lett. **87**, 278701 (2001).
[13] Szabó G, Alava M, Kertész J, Phys. Rev. E **66**, 026101 (2002).
[14] Postel J, RFC 793, Internet Engineering Task Force, SRI International (1981), URL `http://www.ietf.org/rfc/rfc793.txt`.
[15] Jacobson V, Comp. Comm. Rev. **18**, 314 (1988).
[16] Caldarelli G, Marchetti R, Pietronero L, Europhys. Lett. **52**, 386 (2000).
[17] Dorogovtsev SN, Mendes JFF, Samukhin AN, Phys. Rev. Lett. **85**, 4633 (2000).
[18] Yook S-H, Jeong H, Barabási A-L, in *Proc. Natl. Acad. Sci. U.S.A.* (Natl. Acad. Sci., Washington, DC, 2002), vol. 99 of *Appl. Phys. Sci.*, p. 13382.
[19] Abramowitz M, Stegun IA, eds., *Handbook of Mathemathical Functions*, vol. 55 of *Appl. Math. Ser.* (National Bureau of Standards, 1964).
[20] Crovella ME, Bestavros A, IEEE/ACM Trans. Net. **5**, 835 (1997).
[21] Willinger W, Taqqu MS, Sherman R, Wilson DV, IEEE/ACM Trans. Net. **5**, 71 (1997).
[22] Sole RV, Valverde S, Physica A **289**, 595 (2001).
[23] Siganos G, Faloutsos M, Faloutsos P, Faloutsos C, IEEE/ACM Trans. Net. **11**, 514 (2003).
[24] Baccelli F, Hong D, in *Proceedings of IEEE Infocom*, IEEE Computer Society (IEEE, New York, 2002), vol. 1, p. 229.
[25] Ott TJ, Kemperman JHB, Mathis M, in *DIMACS Workshop on Performance of Realtime Applications on the Internet* (1996).
[26] Figueiredo DR, Liu B, Misra V, Towsley D, Comp. Net.: Int. J. Comp. Telecomm. Net. **40**, 339 (2002).

Long PaperID169, 2005 European Conference on Complex Systems
Paris, 14-18 November 2005.

Title: Emergent Group-Level Selection in a Peer-to-Peer Network

Short Title: Group Selection in P2P

Author: David Hales, Dept. of Computer Science, The University of Bologna, Mura Anteo Zamboni 7, 40127 Bologna, Italy.
Email: dave@davidhales.com, Tel: +39 051 19980461, Fax: +39 051 2094510

Keywords: Design of Complex Systems, Robustness, Network Dynamics, Emergent Structure.

Emergent Group-Level Selection in a Peer-to-Peer Network [1]

David Hales
Department of Computer Science, University of Bologna, Italy.
dave@davidhales.com

Abstract. Many peer-to-peer (P2P) applications benefit from node specialisation. For example, the use of supernodes, the semantic clustering of media files or the distribution of different computing tasks among nodes. We describe simulation experiments with a simple selfish re-wiring protocol (SLAC) that can spontaneously self-organise networks into internally specialized groups (or "tribes"). Peers within the tribes pool their specialisms, sharing tasks and working *altruistically* as a team – or "tribe" even though their individual behaviour is selfish. This approach is scalable, robust and self-organising. These results have implications and applications in many disciplines and areas beyond P2P systems.

1. Introduction

Open Peer-to-Peer (P2P) networks (in the form of applications on-top of the internet) have become very popular for file sharing applications (e.g. Kazaa[2], Gnutella[3], Bittorrent[4]). However, can such technology be applied to other computing tasks? For example consider a system in which some nodes have lots of free storage, some high bandwidth and others non-firewalled connections to the network. Those nodes could cooperate to provide a data back-up service – something that no individual node could provide. Obviously, in such a situation, if there is demand for a back-up service we would wish the nodes to, somehow, get together and provide the service – but how?

One solution (and currently, it would seem, the only viable one for deployable applications) is to code the process of specialisation, coordination and cooperation into the protocol directly for each different kind required. So for example, where semantic clustering of media files is required for file sharing, protocols exist that implement it[5] (Handurukande et al 2004). Where systems require supernodes (Montresor 2004), again, these are implemented directly. There are tow problems with this approach; firstly, for every kind of specialisation required a programmer must envisage this *a priori*, design a protocol then implement and test it. Secondly, since this process is complex enough on its own, it is generally assumed that nodes will follow the protocol – it is rare to find protocols robust to node failure, noise or malicious behaviour, such as free riding, although this is, to a certain extent, true within the BitTorrent system (Cohen 2003).

Additionally, it is also rare that nodes can spontaneously change their specialism if they come to recognise that they might be able to do better following a different role. The specialsim of the node tends to be hard-coded or relies on user level switches. This kind of approach limits the ability of the system to automatically adapt to changing task scenarios – however see Montresor (2004) in which supernodes are dynamically allocated to improve performance.

Ideally, we would like a more general approach that could be applied to a range of different task domains with minimal tuning. We would like the approach to offer dynamic specialisation and re-specialisation if nodes come to recognise they could do better playing another role and have the ability

---

[1]This work partially supported by the EU within the 6th Framework Programme under contract 001907 (DELIS).
[2]The Gnutella home page: http://www.gnutella.com
[3]The Kazaa home page: http://www.kazaa.com
[4]The BitTorrent home page: http://www.bittorrent.com. See also Cohen (2003) for a description of the way BitTorrent works.
[5] For example see the MLdonkey system: http://mldonkey.org

to do so or if the task domain changes requiring different kinds of skills to be combined. In addition, we want the system to be able to deal with freeriders and errant or malicious nodes but also to support altruistic cooperation between specialists when this is required for job completion. Finally, we require this to be as scalable, self-organising and robust as possible.

In this paper we do not claim to have addressed all these issues to the level of deployment, what we propose is, we claim, the beginnings of an approach that may allow us to address these issues. In the simulated scenarios so far implemented our results are very encouraging and we plan to continue this line of work.

In the following sections, we state our assumptions concerning behaviour in open peer-to-peer systems then we introduce the SLAC algorithm in general terms. We follow by formulating a minimal task domain scenario called the SkillWorld, to which we wish to subject a simulated P2P network running SLAC. We then describe how we apply SLAC within SkillWorld and present some experiments and results. We interpret the results and describe a "typical history" in the SkillWorld.

At the end of the paper we summarise what we have observed and what it means. We claim that the results indicate a process that has possibly profound implications and applications beyond just P2P systems.

2. Behavioural Assumptions in Open Networks

How do nodes behave in open P2P networks? Of course, the simple answer is, assuming nodes are autonomous: *anyway they like to behave!*

Given this fact, how then do we proceed to devise protocols that will lead to desired system-level functions? Obviously, we have to begin by making assumptions about the *likely behaviour* of other nodes in the network. Such assumptions should be as realistic as possible but also simple enough to be practically computable and transferable between a number of domains. Assumptions made here are essentially the axioms of a kind of mini *social theory* which then informs the design of peer software.

Many approaches (often unconsciously) inherit assumptions from previous social sciences (e.g. economics, socio-biology, sociology). For example, if we assume nodes will behave *"rationally"* in the context of *classical game theory*, then, we compute "Nash equilibrium" - as some researchers do – we inherit our assumptions from game-theory which is a body of knowledge assuming perfect rationality and perfect information. The basic approach is to assume that all individuals have perfect knowledge of the game being played and all possible outcomes along with infinite computational time and common knowledge that all individuals are the same in these respects. Given these assumptions it is sometimes possible to analytically derive the "Nash Equilibria" of the game being played. The idea is that given the previous classical assumptions any system will find and stay in a Nash Equilbria. However, it is unclear that such assumptions hold in dynamic open P2P networks and the derivation of such equilbria within dynamic topologies and changing populations is currently beyond state-of-the-art analytical techniques.

In the context of socio-biological models (Trivers 1971, Maynard-Smith 1982), which are based on the evolution of behaviours of interacting animals over time, the assumption is that behaviours (or strategies) reproduce in proportion to their average fitness (utility or score) such that fitter behaviours become more numerous over time. Additionally such models assume that mutation in the form of random changes in behaviour also take place. This *evolutionary game theory* approach allows for an ecology of behaviours to evolve over time. In addition, there is no requirement that agents have perfect rationality or perfect information – just enough, such that better performing strategies tend to increase in the population. For biological systems this occurs via Darwinian evolution where utility equates to fitness. However, P2P networks don't evolve in a Darwinian fashion. Nodes don't reproduce and it is unclear what "fitness" means in this context.

We have shown in recent work that results from evolutionary models *can* be applied in networks if we allow nodes the ability "copy and re-wire" within the network to improve their own situation (Hales 2004, Hales 2005a, 2005b). This latter innovation demonstrates it is possible to import work originally modelled in a conventional evolutionary framework into a dynamic network model. Nevertheless, in the absence of any deductive proof of the equivalence of evolution and the re-wire rules it is necessary to implement and test previous mechanisms to determine if the properties of interest can be carried over into networks.

Summary of assumptions concerning open P2P networks:

1. Nodes are in the network for what they can get out of it
2. Nodes modify their behaviours to improve their individual benefit
3. Nodes have limited knowledge about other peers and the network in general

The first assumption would appear to be plausible within open P2P networks. In the currently popular file sharing networks the majority of uses download and run peer client software (and hence join the network) in order to get something (e.g. to download a movie or a music file). It certainly is true that some people would join for other reasons. For example, a user may join to feel "part of an online community" (Strahilevitz 2003) or to distribute only their own content - not downloading. Some could aim to damage the functionality of the network by distributing malicious content. However, we argue that neither of these motivations informs the majority of the nodes. In any case, most functions would be *enhanced* by purely altruistic behaviour (such as distributing content without downloading) and we *conjecture* that there are at least as many pure altruistic as pure malicious nodes in working networks.

The second assumption is more problematic – *who says nodes within a given P2P network change behaviours to improve their benefit?* Our argument here is rather speculative - if not conjectural. We start from the assumption of autonomy and argue that the function of peer client software is *ultimately* under the control of the user. For example, users may change operating system or client software settings (e.g. limiting upload speeds), download new versions of a peer client (e.g. incorporating ways to improve download success and rates) or simply hack their own code if they have the required skills. Of course, a hacked client can be distributed to others if it appears to have desirable properties and will tend to be adopted if it delivers those properties to others. We therefore claim that currently, this kind of process is occurring at the user level – via the adoption of various clients and the control of various node-level settings. The problem hidden in this assumption is that the space of available behaviours that each user can choose from varies over time and is also dependent on the knowledge of the user, the kind of network connection, form of operating system and many other related factors. However, we note that similar assumptions have provided some insight into human socio-cultural phenomena at least as complex as the socio-cultural phenomena of P2P systems (Binmore 1998).

Perhaps a more plausible way of thinking about the second assumption is to interpret the space of all available clients in a given P2P domain as the space of behaviours a user can select from – that is, a user may change clients programs, say from edonkey to BitTorrent because edonkey is too slow for the content they required. The user has infact changed protocol and network completely – but that need not matter to them, and in fact the interpretation then is an ecology of different networks with users switching between them.

Alternatively, the assumption that behaviour can change regularly within a single network can be seen as a design proposal to be incorporated into a new protocol, rather than an interpretation of existing protocols.

The third assumption would appear to be a necessary one in any large and highly dynamic system – it is not practical or possible to collate accurate global statistics in most such systems.

3. The SLAC Algorithm

In previous work we showed how a simple "copy and re-wire" rule (or protocol or algorithm) could produce high-levels of cooperation within simulated P2P networks performing collective tasks. We named this algorithm "SLAC" because it uses *Selfish Link and behaviour Adaptation to produce Cooperation*. We showed that nodes in a network could emerge cooperation within the single-round Prisoner's Dilemma (PD) game, under, what we argue, are plausible assumptions about the kinds of behaviour we find in P2P systems. We also demonstrated that the same results could carry over into a more realistic file-sharing P2P task domain (Hales 2004).

The basic algorithm assumes that peer nodes have the freedom to change behaviour (i.e. the way they handle and dispatch requests to and from other nodes) and drop and make links to nodes they know about. In addition, it is assumed nodes have the ability to discover other nodes randomly from the network, compare their performance against other nodes and copy the links and (some of) the behaviours of other nodes.

As discussed above, we assume that nodes will tend to use their abilities to selfishly increase their own utility in a greedy and adaptive way (i.e. if changing some behaviour or link increases utility then nodes will tend to select it).

Over time nodes engage in some activity and generate some measure of utility U (this might be number of files downloaded or jobs processed etc, depending on the domain).

```
DO periodically forever
  select a random node j from the network
  compare utility of this node (i) with node j
  IF utility of j is higher (Uj >= Ui)
    drop all current links (clear view of i)
    copy links of node j (copy view from j to i)
    add link to j (add to view i a link to j)
    copy behavioural strategy of j
    with a low probability (mutation rate 1)
      drop all current links (clear view of i)
      add a link to a randomly chosen node
    with a low probability (mutation rate 2)
      change the behavioural strategy randomly
   END IF
END DO
```

Figure 1
The generic SLAC algorithm.
Each node executes this algorithm.

Periodically, each node (*i*) compares its performance against another node (*j*), randomly selected from the population. If *Ui* < *Uj* node *i* drops all current links and copies all node *j* links and adds a link to *j* itself. Also, periodically, and with low probability, each node adapts its behaviour and links in some randomized way using a kind of "mutation" operation. Mutation of the links involves removing all existing links and replacing them with a single link to a node randomly drawn from the network. Mutation of the behaviour involves some form of randomized change - the specifics being dictated by the application domain (see later).

Previous "tag" models, from which SLAC was developed (Hales 2005c) have indicated that for good scalability properties the rate of mutation applied to the links needs to be higher, than that applied to the behaviour, by about one order of magnitude. In the context of the algorithm show in figure 1 this means that "mutation rate 1" >> "mutation rate 2".

When applied in a suitably large population, over time, the algorithm follows a kind of evolutionary process in which nodes with high utility tend to replace nodes with low utility with nodes periodically changing behaviour and moving in the network. However, as will be seen, this does not lead to the dominance of selfish behaviour, as might be intuitively expected, because a form of incentive mechanism emerges via a kind of ostracism in the network. The process can also be viewed as a kind of "cultural group selection" process (see later discussion).

4. The SkillWorld Scenario

In order to determine if the SLAC approach can support specialisation within tribes we construct a abstract and minimal simulated task domain that requires nodes to perform specialized tasks cooperatively in order to satisfy their individual needs. We call the task domain SkillWorld and it is an adaptation of a sociologically inspired scenario originally given in Hales 2002.

The SkillWorld consists of a population of N nodes. Each node may have zero or more links (up to a maximum of 20) to other nodes. Links are undirected such that the entire population can be considered as an undirected graph G with each vertex being a node and each edge being a link. Each vertex (or node) is composed of three state variables – a "skill type" $s \in \{1,2,3,4,5\}$, an "altruism flag" $a \in \{0,1\}$ and a satisfaction score or "utility" $u \in R$ (where R is a positive real number).

Periodically, with uniform probability, a node *i* is selected from the population N. A "job" *J* is then generated marked with a randomly chosen skill *sJ*. The skill is selected, again randomly with uniform probability, from the domain {1,2,3,4,5}. Job *J* is then passed to node *i*. If node *i* posses the correct matching skill (i.e. if $s_i = sJ$) then node *i* may process the job itself without any help from other nodes. For successfully processing a job *J* the receiving node gains one unit of credit: $u \leftarrow u + 1$.

This process of generating and passing jobs to nodes represents user-level requests for services – such as, for example, searching for a particular file, performing some processing task or storing some data. In the SkillWorld we don't represent the actual jobs to be done, rather, we represent the skill required to perform the job. In our minimal scenario, each job only requires one skill to be completed.

But what if node *i* receives a job for which it does not have the correct skill (i.e. if $s_i \neq sJ$) ? In this case *i* passes the job request to each neighbour in turn until all have been visited or one of them, *j*, agrees to process the job *J*. A neighbour *j* will only agree to process *J* if its skill matches ($s_j = sJ$) and the altruism flag is set ($a_j = 1$). If j does agree to process the job then this *costs j* a quarter unit of utility ($uj \leftarrow uj - 0.25$) yet *increases* the utility of *i* by one unit ($ui \leftarrow ui + 1$).

What this means is that node *i* looks for an altruistic neighbour with the correct skill to process job *J*. If *i* finds such a neighbour (*j*) it increases its utility *as if* it had completed the job itself whereas *j decreases its utility*. This reflects the notion that *j* is altruistically processing *J* for the benefit of *i* and that users are happy when jobs submitted to their nodes are completed but are not happy when jobs from other nodes use their node resources with no immediate benefit to themselves.

5. SLAC in the SkillWorld

We apply the SLAC algorithm within SkillWorld by making the node skill types and the altruism flags into evolvable state variables such that they are copied from more successful nodes (based on utility) and mutated occasionally with low probability.

Although SLAC has previously been demonstrated as successful in promoting cooperation in both a Prisoner's Dilemma playing scenario (Hales 2005) and a simple file-sharing scenario (Hales 2004) it has not yet been applied within a scenario requiring intra-group (or tribe) specialisation *in addition* to altruism. We are therefore asking a lot from a simple algorithm: to self-organise the population into altruistic yet internally specialised tribes that pass and process jobs using their various skills.

The SkillWorld is the simplest scenario we could think of that captures a process of specialisation for this initial investigation. We have a small number of skills (five in these simulations) and we only pass jobs to immediate neighbours. Each node and job is related to a single skill only (rather than a subset of skills which would seem more realistic). Also we assume nodes can change skills at will (randomly via mutation). This latter assumption might not hold if skills relate to physical or unchangeable characteristics of nodes like storage or bandwidth for example. However, at this stage we leave more realistic scenarios with multi-hop passing and more complex skill set arrangements to future work.

In order to measure the success of SLAC we take a simple measure - the proportion of submitted jobs that are completed. We can infer that a network in which the majority of jobs submitted are completed is sustaining internally cooperative and specialised tribes since the only way to complete most jobs is for nodes to pass them to altruistic neighbours with required skills.

5.1 Some Experiments and Results

Initially we ran a set of simulation experiments in which we initialised all nodes in the population with uniformly randomly selected skills, altruism flags and links. We experimented with a number of network sizes determining for each how many cycles before the high performance was achieved (a single cycle is the time unit by which all nodes will have executed the SLAC algorithm at least once, on average).

In order to measure the success of SLAC we take a simple measure - the percentage of submitted jobs that are completed (PCJ). We can infer that a network in which the majority of jobs submitted are completed is sustaining internally cooperative and specialised groupings (or tribes) since the only way to complete most jobs is for nodes to pass them to altruistic neighbours with the required skills.

We categorized "high performance" as a PCJ > 90%, we found that in the simulations this was the highest stable value reached, and ran simulations until this value was obtained - recording the number of cycles required. Hence, if SLAC was working well in the SkillWorld we would hope that within a small number of cycles the PCJ would become high.

We used a mutation rate of 0.001 on skill type *s* and altruism flag *a* (shown in figure 1 as "mutation rate 2"). Mutation on the links (shown in figure 1 as "mutation rate 1") was an order of magnitude higher (0.01). We carry over this assumption – that the mutation rate on the links should be higher than that on the "strategies" – from previous experimental work comparing several different scenarios and models (Hales 2005c). We fixed the maximum number of links between nodes to 20. Links are undirected and therefore symmetric. If an operation results a node requiring a new link and it already has the maximum then a random link is discarded by the node and the new link accepted. Using this method

nodes never refuse new links but may often lose old ones. This adds to the noisy and dynamic nature of the scenario.

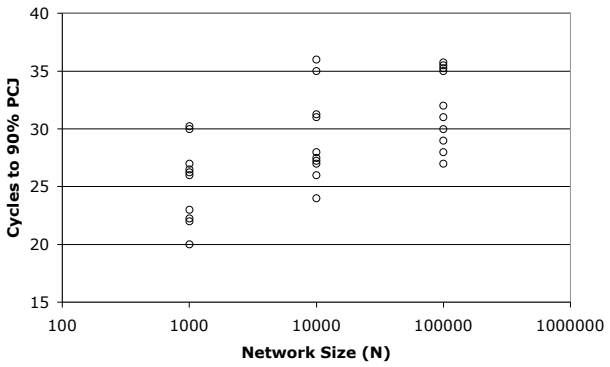


Figure 2
Number of cycles to high performance for different network sizes. When PCJ > 90% this means that over 90% of all jobs submitted to nodes are completed. Note: overlapping circles have identical values.

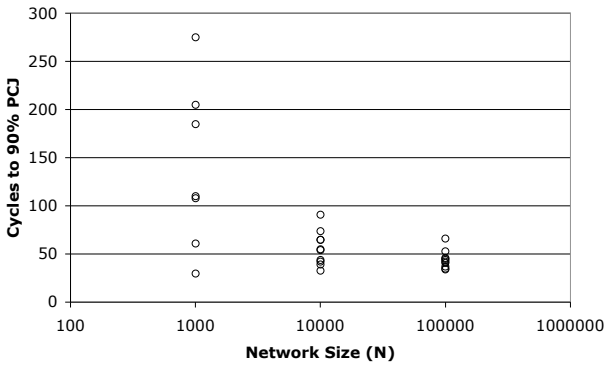


Figure 2b
Number of cycles to high performance for different network sizes when all nodes are initialised selfish (a = 0). This can be compared to the random initialisation results in figure 2. Note that there is a reverse scaling cost here. The results for N = 1000 are worse than shown since three outliers at about 1000 cycles are not shown here.

Figure 2 shows results from 30 individual simulation runs. Each point is a different run showing the first cycle at which the PCJ > 90%. As can be seen, high performance is attained within a few tens of cycles even for networks of size $N = 10^5$. Notice that there appears to be a very slight upward trend in cycles as N increases, however, this is negligible – the results therefore indicate close to *zero scaling cost*. This highly desirable property was also evidenced in a previous application of SLAC to a simulated fire-sharing scenario (Hales 2005). Figure 2b shows results under the same conditions except that all

nodes are initialised to be selfish (a = 0). This gives a kind of "worst case scenario" as far as altruism evolving. It is important to show that the system can escape from this, since this demonstrates that even if a complete failure of node altruism should occur (either through chance or malicious attacks) then the system can recover relatively quickly. We notice here the reverse scaling properties that we originally noticed and analysed in a previous "tag" model (Hales 2000). Essentially, with bigger populations, there is more likelihood of the chance formation of a small altruistic tribe. This then goes on to "seed" the population with altruism[6].

Interestingly, it was found that for populations where N < 1000 high performance was *not* produced even when runs were extended to several thousand cycles. Intuitively this is consistent with the "group selection" hypothesis concerning how SLAC operates. With small populations there are not enough nodes to form enough competing groups (or tribes) so evolution can not operate at the group level.

5.2 History in the SkillWorld – Tribal Dynamics

One way to convey the dynamics of a typical SkillWorld simulation run is to describe a typical "history" in narrative form – this method is sometimes used in computational sociology, particularly in work with artificial societies (Epstein and Axtell 1996, Axelrod 1995) carried over from more traditional sociological methods of explanation. In the rest of this section we give such a "typical history". Although we will make general points we will also refer to a specific single simulation, run given in Figure 3, to illustrate our analysis.

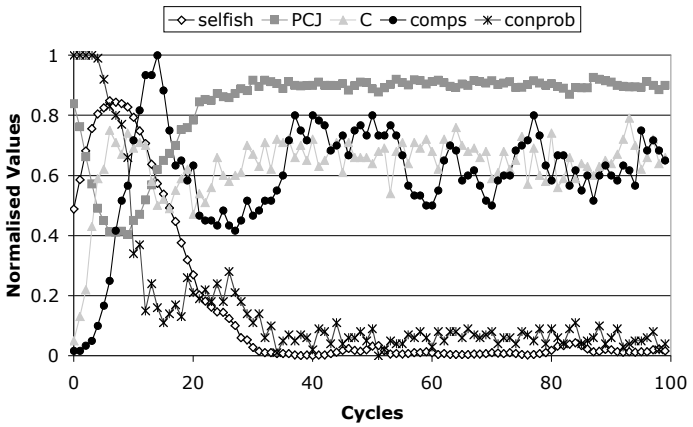


Figure 3

The time series of a typical single run in SkillWorld (N=1000). Shown are the number of selfish nodes as a proportion of the entire population (selfish), the proportion of completed jobs submitted (PCJ), the clustering coefficient (C), the number of components in the population (comps, which is normalised by dividing by 60) and the average probability that a route exists between any two nodes (conprob).

Initially, the SkillWorld is a random graph, all nodes are connected via a few hops and clustering is low. Skills and altruism are randomly scattered. Very quickly, the graph breaks into a population of many disconnected components because nodes quickly re-wire themselves to better performing nodes.

[6] See Hales (2000) for a more detailed explanation of this reverse-scaling cost including the beginnings of an analytical treatment.

The better performing nodes are initially the non-altruists who exploit their groups (or tribes) selfishly. However, this is a non-sustainable strategy since this exploitation causes nodes to leave their exploited tribes and join tribes in which there is less exploitation – nodes in tribes with less exploiters in them do better (higher utility) because they are cooperating as a team. The tribes dominated by non-altruists quickly "wither away" as nodes leave. When no nodes are left then the tribe no-longer exist – in this way *tribes die*, even though *nodes do not die*. This emergent property of the birth and death of tribes lays the ground for evolution to operate that the group (tribe) level.

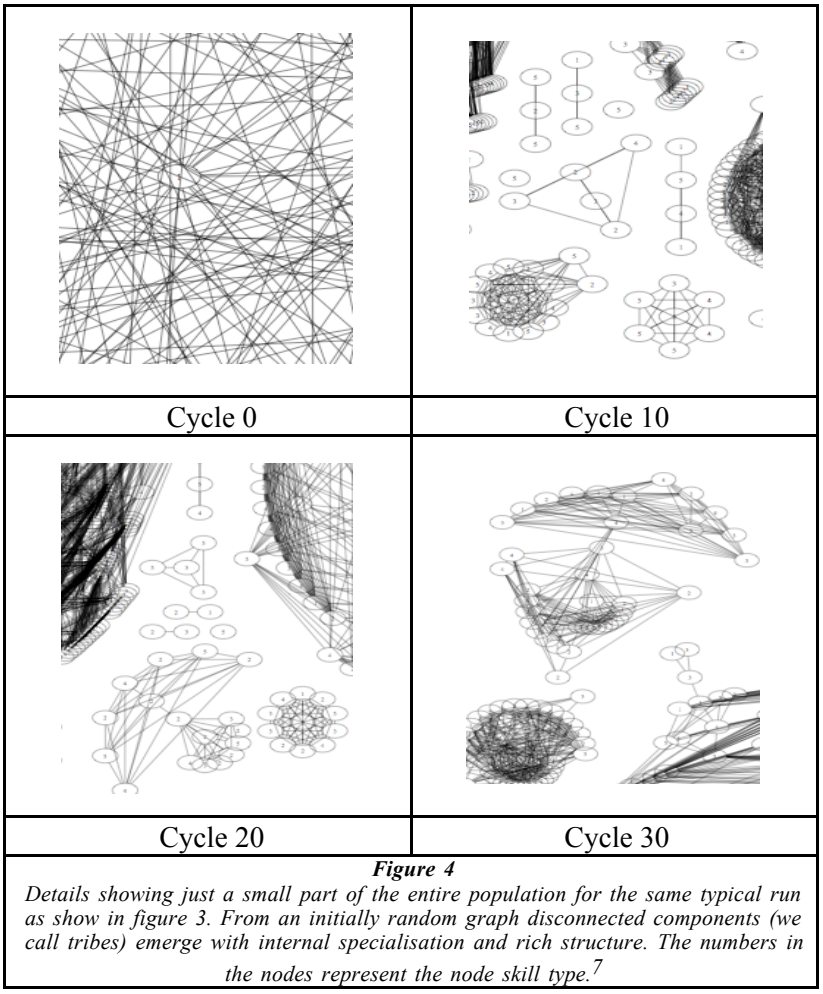


Figure 4
Details showing just a small part of the entire population for the same typical run as show in figure 3. From an initially random graph disconnected components (we call tribes) emerge with internal specialisation and rich structure. The numbers in the nodes represent the node skill type.[7]

Figure 3 indicates the above process occurring in the first 10 cycles or so. Notice that the number of selfish nodes peaks, and the proportion of completed jobs (PCJ) bottoms out, at about cycle 10. The number of components (i.e. tribes) increases in the early phase peaking just before cycle 20 (representing a peak of 60 components).

Altruistic tribes function well and grow as more nodes join, new tribes are occasionally formed as nodes randomly, through mutation, split from a tribe. As altruistic tribes grow larger they eventually become "infected" or "invaded" by a non-altruist node – either by mutation of an existing member node

[7] Full sized pictures can be found at http://www.davidhales.com/esoa05pics.

or the entering of a new node to the tribe. When this happens the tribe is quickly destroy via dispersion since a non-altruist will exploit the tribe selfishly and this will lead to many more nodes quickly copying that node until the tribe "dies" because all nodes leave it – because a tribe dominated by selfish nodes gives lower utility to *all nodes* within it than one dominated by altruists.

Figure 3 shows, from about cycle 20 onward, the above process occurring. A decrease in the number of components (comps) and an increase in completed jobs (PCJ) are correlated with a decrease in the number of selfish nodes (selfish). This is because altruistic tribes grow in size – reducing the total number of components (comps) and reducing selfish nodes (selfish). By about cycle 30 selfishness is very low and completed jobs (PCJ) reaches a high level. Notice that the dynamic nature of the formation and dissolution of the tribes is reflected in the variation of the number of components over time (comps) after PCJ goes high.

History in the SkillWorld is the history of the formation, growth and destruction of tribes. From the simple rules of the SLAC algorithm an *evolutionary process emerges at the tribal or group level*. Essentially one can think of this evolution as the competition between tribes to retain nodes to continue to exist. This process is in constant flux due to mutation and movement, no equilibrium state is attained and no tribe lasts forever. As long as new altruistic tribes are created at least as rapidly as they are destroyed then altruism can survive.

Figure 4 shows a small detail of snapshots of the population over time (space does not permit full size snapshots). As can be seen, tribes quickly emerge and grow, producing various structures and sizes with internally specialised nodes.

5.3 Tribal Structures

Within the SkillWorld, tribes with different structures and skill mixes will support different levels of utility – a highly connected tribe with an even mix of skills would produce better results than a tribe missing some skill. Hence, selection at the tribe level (group selection) will tend to operate to structure the tribes into more optimal structures of skill types. We would therefore expect to see tribes composed of nodes possessing each skill type linked together such that a node receiving a job can either process it directly or will be directly linked to a node with the appropriate skill willing to do the job. In the SkillWorld then, we have tribe level selection not only operating to control selfishness but also to tune the internal (organisational) structure of the tribe.

We find this particularly exciting since we believe that by increasingly the complexity of the task domains and giving nodes a little more freedom to hop more than one link within their tribe it should be possible to evolve tribes with *complex organisational structures* tuned to performing in the given task domain. Moreover, since the tribes are constantly evolving they should be able to *change their structure dynamically to address a change in the task domain*[8].

6. Conclusion

We have demonstrated that the SLAC algorithm can be applied in a scenario (the SkillWorld) requiring node specialisation in addition to the suppression of selfish behaviour. When the algorithm is executed the network quickly divides into competing "tribes" (disconnected components). An evolutionary process then emerges at the group level selecting efficient internally specialised tribes – which deliver high levels of service with respect to user submitted jobs at the nodes.

We adapted the SkillWorld scenario from a previous model developed for the purposes of social scientific theorising (Hales 2002). The previous "tag-based" model relied on mean-field mixing (with no population structure) and followed a conventional evolutionary process.

---

[8] Further experiments not detailed here, demonstrate that even when all skills in the population are initialised to the same single type – the network quickly adapts into an even skill spread due to mutation on the skills and selection at the tribe level.

Our belief that the SLAC algorithm works via a kind of group selection occurring at the level of the "tribe" gave us the *a priori* expectation that it would select tribes that could perform well in the SkillWorld. In this sense, *dare we claim the beginning of a "proto-theory"* allowing us to make some modest qualitative predictions?

More generally, we claim that this paper demonstrates concretely within a dynamic network the emergence of what has been termed a "meta-state transition" (MST) within evolution (Heylighen 1992). It has been argued that the emergence of life itself and major steps in biological evolution (e.g. multi-cellular organisms) and social evolution (e.g. large complex societies) occur over such MST's. In this context we advance our results as possibly of great theoretical insight.

It is important to understand that the concept of *the "tribe" is actually a theoretical construct* we use to help to explain and understand the *emergent phenomena* produced by the SLAC algorithm over time. The tribes are not "programmed" into the nodes *a priori* but rather emerge from the interplay of task domain, interaction and the SLAC algorithm. We use the concept of "tribes" because we believe it to be valuable in beginning to understand, control and theorise about what is occurring in SLAC networks. However, since the tribes are emergent we do not *begin* with a "theory of tribes" rather we observe, experiment and induce knowledge about them. As discussed below, this does not preclude, but, in fact, should support, the formation of an analytical theory – we hope.

Since the nodes do not die or model genetic operators, the tribe level selection process can be viewed as a kind of artificial *cultural group selection process*. What is quite extraordinary is that *such a simple node level algorithm (SLAC) based on a few plausible assumptions about preferential attachment can lead to such complex and useful group level evolutionary dynamics*.

A key issue however, is that, although SLAC is simple to implement the dynamics are complex and currently it is not know how analytical tools can be applied to truly understand, predict and prove the properties of SLAC. So far the only "proofs" we have are in the form of "existence proofs" demonstrated by empirical analysis of simulation runs. Such "proofs" are not watertight and can always be questioned given anomalous results from future simulation studies (rather like experiments in the natural sciences). We have some confidence in the general results from SLAC-like algorithms however (such as those based purely on "tags") since there have been a number of replications of those results from multiple independent implementations using different languages, machines and programmers (Edmonds and Hales 2003). However, none of this offers predictive insight into the process as a good analytical model would. What we currently have is a kind of "toolbox" of algorithmic heuristics that appear to be reasonably robust over some minimal task domains and scenarios.

However, currently, the only way to apply these methods to new domains is to simulate and experiment – copying and adapting heuristics that worked previously in similar domains. Perhaps this is not so far away from the edit / compile / debug cycle of good old-fashioned software engineering (GOFSE). This could bode well for future progress.

Acknowledgements

Thanks go to Mark Jelasity for pointing out some of the recent models that bear close comparison to this one. Also, along with Ozalp Babaoglu, and Alberto Montresor, for writing clear and readable papers about P2P systems that have helped me in beginning this line of work.

References

Axelrod (1984) The Evolution of Cooperation, Basic Books, New York.

Axelrod, R. (1995) A model of the emergence of new political actors. In G. N. Gilbert and R. Conte (eds.), Artificial Societies. UCL Press, London

Binmore, K. (1998), Just Playing, Game Theory and the Social Contract, Volume II, Cambridge, Mass., and London: The MIT Press.

Cohen, B. Incentives Build Robustness in BitTorrent (2003). Presented at the 1st Workshop on the Economics of Peer-2-Peer Systems, June 5-6, Berkley, CA, 2003 (Available at: http://www.sims.berkeley.edu/ research/conferences/p2pecon/).

Davis, L.(1991) Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York.

Di Marzo Serugendo, G. "Engineering Emergent Behaviour: A Vision", Invited Talk. Multi-Agent-Based Simulation III. 4th International Workshop, MABS 2003 Melbourne, Australia, July 2003, D. Hales, B. Edmonds, E. Norling, J. Rouchier (Eds), LNAI 2927, Springer-Verlag, 2003.

Edmonds, B. and Hales, D. (2003) Replication, Replication and Replication - Some Hard Lessons from Model Alignment. Journal of Artificial Societies and Social Simulation 6(4).
Epstein J.M. and Axtell R. (1996) Growing Artificial Societies - Social Science from the Bottom Up, Cambridge MA, MIT Press.
Eytan Adar and Bernardo A. Huberman (2000) Free Riding on Gnutella. First Monday Volume 5, No. 10, (http://www.firstmonday.dk/issues/issue5_10/adar/index.html).
Hales, D. (2000), Cooperation without Space or Memory: Tags, Groups and the Prisoner's Dilemma. In Moss, S., Davidsson, P. (Eds.) Multi-Agent-Based Simulation. Lecture Notes in Artificial Intelligence, 1979:157-166. Berlin: Springer-Verlag.
Hales, D. (2002) Evolving Specialisation, Altruism and Group-Level Optimisation Using Tags. In Sichman, J. S., Bousquet, F. Davidsson, P. (Eds.) Multi-Agent-Based Simulation II. Lecture Notes in Artificial Intelligence 2581:26-35 Springer-Verlag. Berlin.
Hales, D. (2004) From selfish nodes to cooperative networks - emergent link based incentives in peer-to-peer networks. In Proc. of the 4th IEEE International Conference on Peer-to-Peer Computing (P2P2004). IEEE Computer Soc. Press, 2004.
Hales, D. (2005a) Self-Organising, Open and Cooperative P2P Societies – From Tags to Networks. Presented at the 2nd Workshop on Engineering Self-Organising Applications (ESOA 2004). To be published by Springer, 2005.
Hales, D. (2005c) Change Your Tags Fast! – a necessary condition for cooperation? Presented at the workshop on Multi-Agents and Multi-Agent-Based Simulation (MAMABS'04) workshop 2004. To be published by Springer, 2005.
Hales, D. and Edmonds, B. (2005) Applying a socially-inspired technique (tags) to improve cooperation in P2P Networks. To appear in IEEE Transactions on Systems, Man, and Cybernetics, Part A, 2005.
Handurukande, S. A.-M. Kermarrec, F. Le Fessant and L. Massoulié (2004) Exploiting Semantic Clustering in the eDonkey P2P Network. Presented at the 11th ACM SIGOPS European Workshop (SIGOPS), Leuven (Belgium), Sep 2004.
Hardin, Garrett (1968) "The Tragedy of the Commons,", Science, 162:1243-1248.
Heylighen F. (1992) "Evolution, Selfishness and Cooperation", Journal of Ideas, Vol 2, # 4, pp 70-76.
In Proceedings of the 4th International Conference on Peer-to-Peer Computing, pages 202-209, Zurich, Switzerland, August 2004. IEEE Press.
Jelasity, M., Montresor, A., and Babaoglu, O. (2004) A modular paradigm for building self-organizing peer-to-peer applications. , In Di Marzo Serugendo, G.; Karageorgos, A.; Rana, O.F.; Zambonelli (eds.) Engineering Self-Organising Systems - Nature-Inspired Approaches to Software Engineering. Lecture Notes in Artificial Intelligence 2977, Springer, Berlin.
Maynard-Smith, John (1982) Evolution and the Theory of Games (Cambridge: Cambridge University Press).
Montresor, A. (2004) A Robust Protocol for Building Superpeer Overlay Topologies.
Qixiang Sun & Garcia-Molina (2004) SLIC: A Selfish Link-based Incentive Mechanism for Unstructured Peer-to-Peer Networks. In Proceedings of the 24th IEEE international Conference on Distributed Systems (March 2004). IEEE computer Society.
Roberts, G. & Sherratt, T. N. (2002) Nature 418, 449-500
Strahilevitz, Lior, (2003) Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks . Virginia Law Review, Vol. 89, 2003 http://ssrn.com/abstract=329700
Trivers, R. (1971) The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35-57.
Watts, D. (1999) Small Worlds: The Dynamics of Networks between Order and Randomness. Princeton University Press. Princeton, New Jersey.

Measuring the Dynamical State of the Internet: Large Scale Network Tomography via the ETOMIC Infrastructure

Gábor Simon[a,b] József Stéger[a,b] Péter Hága[a]
István Csabai[a,b] Gábor Vattay[a,b]

[a] Department of Physics of Complex Systems, Eötvös Loránd University, Budapest, Hungary;
[b] Collegium Budapest Institute for Advanced Study, Budapest, Hungary

Short Title
Measuring the Dynamical State of the Internet

Key Words
Complex networks, Internet, Network tomography, Queueing delay distributions.

Abstract
In this paper we show how to go beyond the study of the topological properties of the Internet, by measuring its dynamical state using special active probing techniques and the methods of network tomography. We demonstrate this approach by measuring the key state parameters of Internet paths, the characteristics of queueing delay, in a part of the European Internet. In the paper we describe in detail the ETOMIC measurement platform that was used to conduct the experiments, and the applied method of queueing delay tomography. The main results of the paper are maps showing various spatial structure in the characteristics of queueing delay corresponding to the resolved part of the European Internet. These maps reveal that the average queueing delay of network segments spans more than two orders of magnitude, and that the distribution of this quantity is very well fitted by the log-normal distribution.

Corresponding Author
Dr. Gábor Simon
Department of Physics of Complex Systems,
Eötvös Loránd University
H-1518 Budapest Pf.: 32, Hungary.
Tel:+36 1 3722896, Fax: +36 1 3722866
E-mail gaba@complex.elte.hu

Introduction

Internet is a complex network of computers and routers connected by direct wired or wireless links, where information is transmitted in discrete sized packets. Since 1970 the Internet evolved rapidly and in a decentralized manner from a few interconnected local area networks into a highly heterogenous global network spanning continents. The Internet provides various services and applications that became part of our every-day activity including web browsing, file transfer, multimedia, Internet-telephony, and others. In the future it is expected that the reliance of the population on the firm functioning of the Internet and on the new emerging applications will increase dramatically, thus it is of utmost importance to study and to understand the properties of this huge network at various levels of abstraction. By exploring the characteristics of the present-day Internet, and the various phenomena associated to its traffic it may be possible to extrapolate that knowledge to predict the properties of the future Internet and to foresee its problems.

Traditionally the Internet was studied by various engineering approaches that concentrated mostly on practical aspects, however recently with the accumulation of large datasets, it was recognized that the Internet can also be studied from a complex systems perspective, which represents a different level of abstraction. Complex systems [1] are large networks of interacting agents, where the emphasis of the studies are on the global properties, characterizing the system as a whole, by neglecting the fine details of the interactions. It is believed that microscopically vastly different complex systems have some universal features that are independent of the details. For example collective phenomena of congestion buildup are present both in the Internet and highway-traffic, showing analogous dynamics on a global scale [2]. Also both of these systems are characterized by self-similarity [3,4], and 1/f noise in the congested state [5-7].

Most of the existing work so far addressed questions concerning the topological properties of the Internet [8,9] or related topics like tolerance to attacks or failures [10]. These studies established interesting characteristics like the power-law degree distributions of the Internet connectivity graph on the autonomous system and router level. Although as the network evolves this continues to be an open issue and a hot topic, nevertheless it is essential to extend the study of the Internet beyond the discovery of topological properties towards the characterization of its dynamical state. Steps have already been taken in this direction; for instance in references [11,12] who revealed universal scaling in traffic intensity fluctuations.

From the point of view of the current, most widely used data transfer protocols like the variants of TCP, the most relevant state variables of the network are the loss-rates and the delays encountered by data packets on a path, since these quantities control the transmission rate of the transfer protocol. There are two different techniques to measure loss-rate and delay over the Internet, these are active probing and passive monitoring. Active probing involves injecting of probe packets into the network and analyzing the properties of the received probe-stream. This technique is flexible, has wide range of applicability, however its extensive usage may impose unnecessary load on the network. On the contrary passive monitoring does not impose any load, but it relies on the fact that the user must have an access to the network element under investigation, which is generally not the case. In the past years several measurement platforms have been developed for conducting active measurements over the Internet (e.g. Surveyor, Felix, AMP [13]), however these platforms can provide only end-to-end information between the participating nodes, where the measured characteristic can not be resolved on the parts of the end-to-end path. Also most of the existing active probing tools rely on extra cooperation of the routers in the path to process their packets. As the Internet continues to evolve towards more decentralized and heterogenous administration, in the future the cooperation of the network elements can be foreseen to be limited to the basic process of just storing and forwarding incoming probe packets.

The solution to these problems is provided by network tomography, which is a special class of active-probing measuring techniques, that is able to resolve the end-to-end delay statistics [14,15] and packet loss rates [16,17] to internal segments of the paths. In general a tomography measurement made from a single source to a set of receivers admits the determination of the delay statistics and loss-rates on each segment of an underlying logical tree that is spanned by the source and receiver nodes, and the branching nodes (nodes where the path of probes destined for different receivers diverge). By increasing the number of sources and receivers involved in a tomography measurement, the portion of the network for which state information can be resolved grows dramatically. Initially network tomography techniques were developed for the use with multicast probes [18-21], which requires the extra cooperation of the routers to support multicast functionality, however later these approaches were also extended to the case of using unicast probes [16,14,17], and performing the measurements from multiple sources [22,23], which makes unicast network tomography the most general tool to measure spatially resolved characteristics of an uncooperative Internet.

The main idea of unicast network tomography is to use back-to-back packet pairs, where each packet of a pair is destined to different receivers. As the packets of a pair traverse their paths, they experience the same network conditions on the common segment from the source to the branching node, which brings correlation into the time-series of the end-to-end characteristics. This inherent correlation property of such probe streams is the key to resolve the internal characteristics from the end-to-end measurements.

The delay experienced by a packet over an Internet path sums up from two non-negative components, a constant propagation delay and a time-varying component due to queueing in the buffers of routers. In this paper we are concerned with large-scale inference of queueing delay distributions in the Internet by performing extensive unicast network tomography measurements. The large-scale study of queueing delay distributions is motivated by the fact that this observable carries vast amount of information about traffic properties and the state of congestion on the measured path. By resolving the queueing delay distributions from end-to-end measurements we can draw a map of congestion of the network segments, analyze spatial structure, and identify highly congested or faulty segments.

The rest of the paper is organized as follows. In the following section we describe ETOMIC, the measurement platform, where the experiments were conducted. Next we proceed with the presentation of the results of our preliminary large-scale tomography measurement, while the detailed description of the queueing delay tomography method is given in the last section.

The ETOMIC Measurement Platform

To perform unicast queueing delay tomography in a real network environment poses several challenges. First in order to be able to measure true end-to-end delay, source and receiver nodes need to be synchronized to a common clock-reference, and must stay in the synchronized state during the measurements. Second, the measuring infrastructure has to be very precise in order to be able to resolve the microsecond-scale queueing delay components associated to high-bandwidth (multi-Gigabit) links. The precision of commercial workstations are insufficient for this task, thus to achieve sub-microsecond precision, a hardware solution is inevitable.

In the subproject of the European Union sponsored EVERGROW[24] Integrated Project, we are developing a state of the art high-precision, synchronized measurement platform, the Evergrow Traffic Observatory Measurement InfrastruCture (ETOMIC) [25]. This platform among others provides the ability to perform large-scale delay and loss tomography based on unicast probing techniques, and will be generally available and open to the public. Currently ETOMIC consists of 15 measuring nodes deployed at different locations in various European countries (See Table I., while in the future we plan to extend this number to 50 participating nodes. The measurement nodes and the network experiments with them are managed through

a central management system that is accessible to the researchers through a web-based graphical user interface [26].

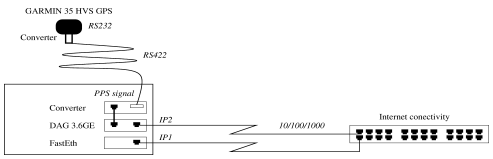


Fig. 1. The measuring node is a standard PC equipped with a DAG 3.6GE network interface card, connected to a GPS unit.

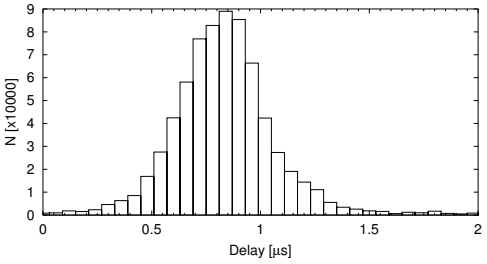


Fig. 2. One-way delay measurement between two gps-synchronized ETOMIC nodes over an empty link. The FWHM of the histogram indicates that the attainable precision of end-to-end delay measurements within ETOMIC is around 0.5µs.

Table 1. The list of current ETOMIC nodes

ABBREVIATION	IP ADDRESS	LOCATION
SICS	193.10.64.81	Stockholm, Sweden
TELI	217.209.228.122	Stockholm, Sweden
ERIC	192.71.20.150	Stockholm, Sweden
UNAV	130.206.163.165	Pamplona, Spain
ASTN	134.151.158.18	Birmingham, England
HUJI	132.65.240.105	Jerusalem, Israel
OVGU	141.44.40.50	Magdeburg, Germany
ROME	141.108.20.7	Rome, Italy
UNIV	193.6.205.10	Budapest, Hungary
COLB	193.6.20.240	Budapest, Hungary
ELTE	157.181.172.74	Budapest, Hungary
UPAR	193.55.15.203	Paris, France
SALZ	212.183.10.184	Salzburg, Austria
UBRU	193.190.247.240	Brussels, Belgium
CRET	147.27.14.7	Chania, Greece

The schematics of an ETOMIC measurement node is displayed in fig. 1. It is based on standard PC hardware, but also includes an Endace DAG 3.6GE card as the network monitoring interface, which is specifically designed for precise active and passive measurements [27]. These DAG cards provide very accurate time-stamping of the probe packets, with a time-resolution of 60 ns, and also advanced capabilities for transmission. A burst composed of several packets can be transmitted with precise user-defined inter packet timings. The measuring nodes are synchronized by GPS (Garmin GPS 35 HVS), that provides a PPS (pulse per second) reference signal directly to the DAG card. The accuracy of one-way delay measurements between two ETOMIC nodes is found to be ≈ 0.5 µs, which is mainly limited by the performance of the GPS receivers. This result was obtained in a lab experiment before the deployment of the nodes, where we connected two ETOMIC nodes by an empty link, and transferred a long stream of probe packets between them. Figure 2 shows the histogram of the measured delay, where the bin size reflects the time resolution of the DAG-cards.

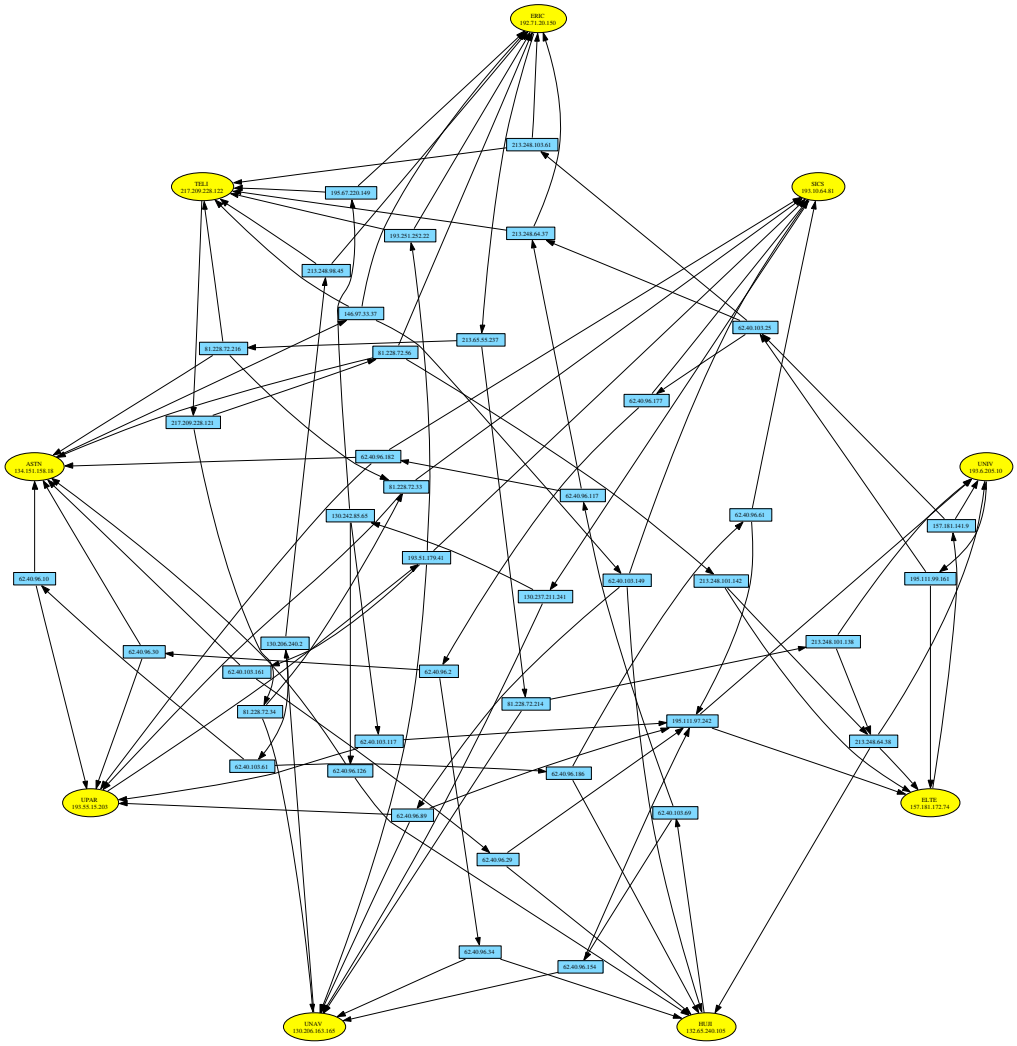


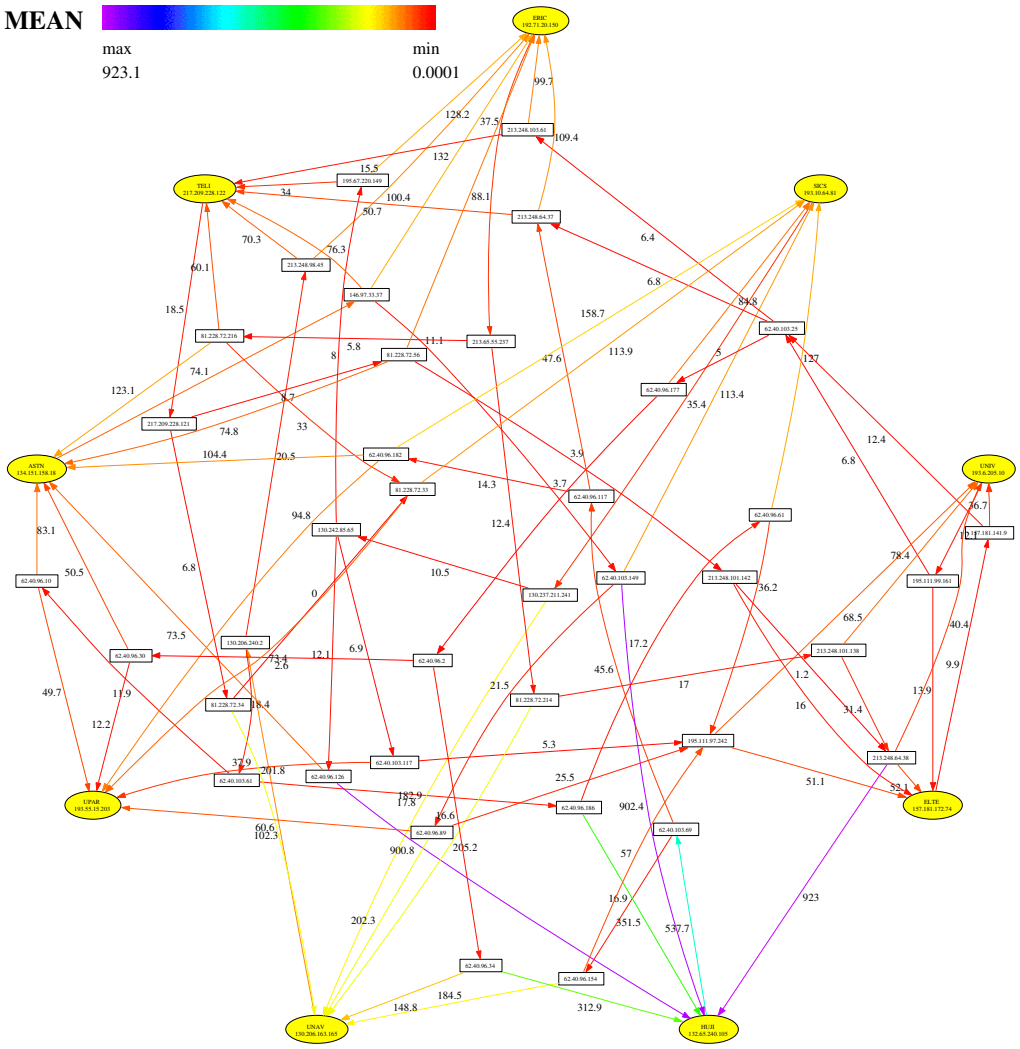
Fig. 3. The connection topology between 9 ETOMIC measurement nodes. The ellipse shaped nodes on the edge are ETOMIC measurement nodes with abbreviations given in Table I., while box-shaped nodes in the interior of the graph are branching nodes. The arrows indicate the direction of probe packet flow on a given network segment. The boxes with an IP address of 62.40.X.X are nodes within GÉANT.

Figure 3 depicts the connection topology between 9 measuring nodes, obtained by extensive traceroute measurements, where we extracted from the detailed traces the segments for which characteristics can be resolved by network tomography measurements. This arrangement involves 104 network segments and 42 branching nodes (among these 19 is situated in the GÉANT multi-Gigabit European academic network [28]) with link speeds ranging from 2 to 15 Gb/s.

3. Large-scale Queueing Delay Tomography

Here we report a large-scale tomography measurement that involved 9 ETOMIC measurement nodes, and a connectivity graph consisting of 38 branching nodes and 93 network segments. Among these 9 ETOMIC nodes 8 were simultaneously sources of outgoing back-to-back probe pairs, and receivers of the incoming probe packets, while one of the nodes was only used as a receiver. Each of the source nodes sent probe pairs consisting of small sized UDP packets to all the possible pairs of receivers in a round-robin fashion with an inter-pair time of 1 ms, and repeated this process many times. This procedure finally resulted

in data sets, each containing two time-series of end-to-end delays with an approximate length of 10000 elements.



Fif. 4. The connectivity graph colored and labeled by the mean queueing delay, given in units of μs, for each network segment.

These data sets comprised the input to the tomography method described in sec. 4 that yielded as an output the queueing delay distributions resolved for each segment contained in the connectivity graph. Since a given segment can be a part of different end-to-end paths, this fact enabled to test the consistency of the results, as well as the averaging of the distributions obtained from different data sets, but attributed for the same segment. For better visualization of the results we extracted the mean and the standard deviation of the queueing delay distributions. These dynamical state parameters are shown on top of the connectivity graph in figs. 4 and 5.

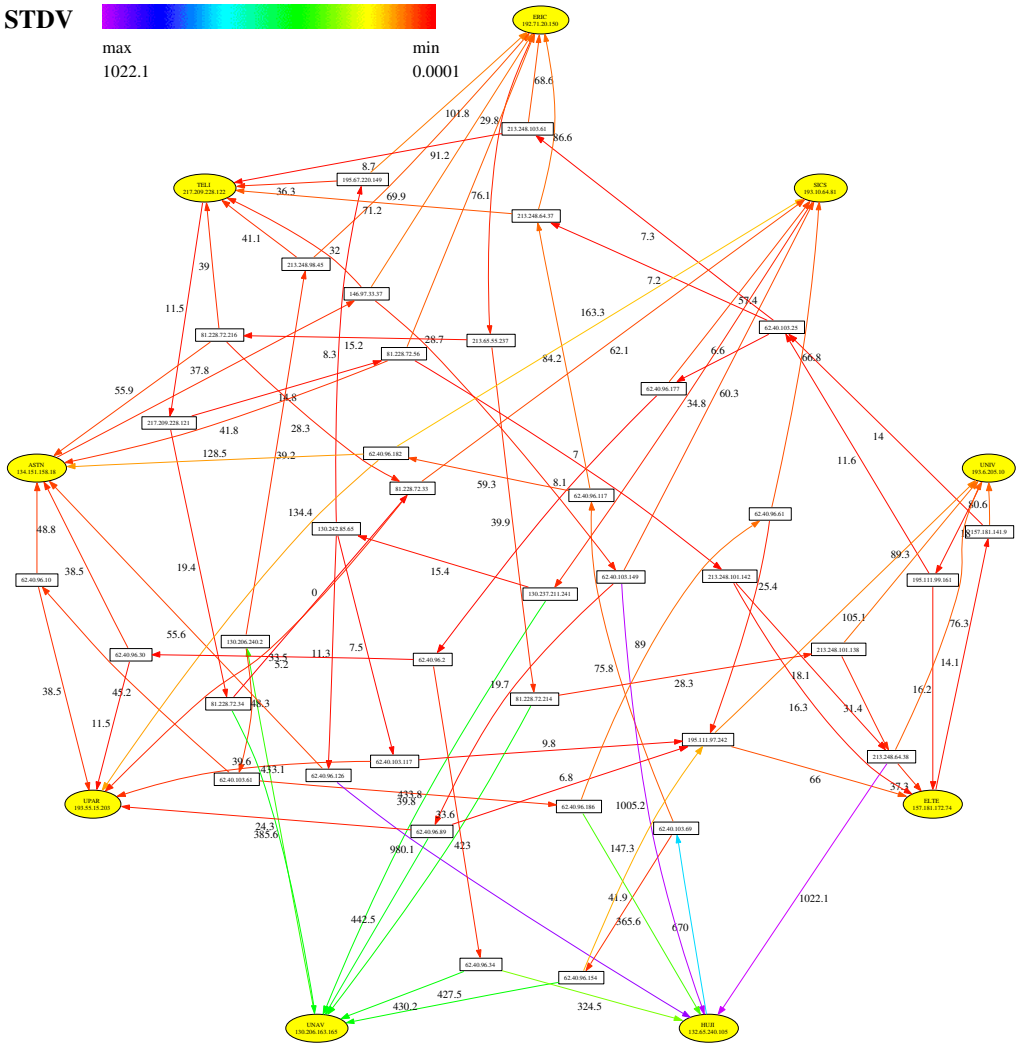


Fig. 5. The connectivity graph colored and labeled by the standard deviation of the queueing delays, given in units of μs, for each network segment.

3.1. Analysis of the results

The results of figs. 4 and 5 reveal some interesting structure. First of all the state variables of the segments span three orders of magnitude, ranging from the error limit of the delay measurements (≈ 0.5 μs), to an average queueing delay of ≈ 1ms, that characterizes a segment which connects GÉANT to the Hebrew University in Jerusalem. The results also reveal an interesting geographical feature, namely that the segments originating or ending in ETOMIC nodes that are located on the south (HUJI, UNAV), are characterized by the highest average and standard deviation of the queueing delays.

As a general feature it can be observed that for all end-nodes incoming segments are characterized by higher values of average queueing delays then outgoing segments. This result can be interpreted by a reasoning that the amount of data downloaded from the Internet to an organization is usually higher, than the amount downloaded from the servers of the organization by clients situated elsewhere in the Internet. Looking at the spatial arrangement of the state variables, one can see that the internal segments that are connections between branching nodes constitute a core which is characterized by the smallest values of the state variables. This is not surprising, since these are network segments in the gigabit backbone.

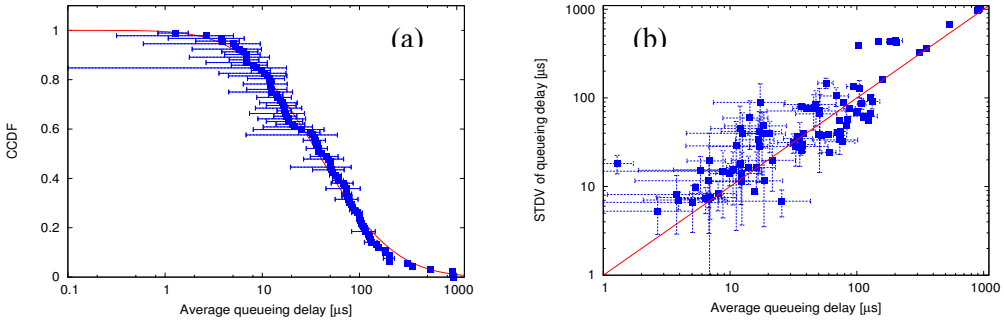


Fig. 6. In (a) the complementary cumulative distribution of the average queueing delays are given by filled boxes with error-bars, while the continuous line is a fit given by equation (1). In (b) the standard deviations as a function of the average queueing delays are plotted with error-bars for all the resolved network segments of fig. 4. Here the continuous line marks the diagonal.

To analyze further the results in fig. 6(a) we plot the complementary cumulative distribution function of the average queueing delays on the different segments, that is closely fitted by

$$g(x) = \frac{1}{2}\left[1 - erf\left(\frac{\ln x - m}{\sigma\sqrt{2}}\right)\right], \tag{1}$$

where x is the average queueing delay of a network segment in microseconds, while σ and m are the fitting parameters. The implication of the very good fit with $\sigma \approx 1.42$ and $m \approx \ln(37.8\mu s)$ in fig. 6(a) is that the average queueing delay of the different segments follows a log-normal distribution.

$$P(x) = \frac{1}{\sigma\sqrt{2}x} e^{-\frac{(\ln x - m)^2}{2\sigma^2}}. \tag{2}$$

It is interesting that despite the features and spatial structures identified earlier, the data follows a smooth continuous function without segregation of the average queueing delays into clearly visible groups. In fig. 6(b) we plot the standard deviations as a function of the average values of the queueing delay. The figure indicates that the majority of the segments, with a data point near the diagonal, are characterized by a queueing delay distribution that is close to being exponential, while the segments with a large deviation above the diagonal may be characterized by a self-similar traffic-flow and heavy-tailed distributions.

4. Description of the Method for Queueing Delay Tomography

In this section we provide the detailed description of the queueing delay tomography method that was applied on the measured end-to-end data in obtaining the results of the previous section.

To achieve a unique labeling of the different network segments of a measurement tree we use the arrangement illustrated in the example of fig. 7. The source node is situated on the top of the hierarchy, labeled with "0", while branching nodes are put on a given level according to their "distance" from the source. This distance is measured by the number of branching nodes a probe packet must pass to reach the target. Different nodes on the same hierarchical level are enumerated in the incremental order from the left to right. Using this scheme one can identify each node by a unique index, while the segments, indicated by arrows in fig. 7, can be identified by the indices of their target nodes. We denote the one-way delay experienced

by the n-th probe on the i-th segment by $\tilde{X}_i(n)$, and the corresponding end-to-end delay from the source to the i-th receiver node by $\tilde{Y}_i(n)$. The queueing delays associated to these quantities are denoted by $X_i(n)$ and $Y_i(n)$. By dropping the n index we indicate the time-series of the respective quantities.

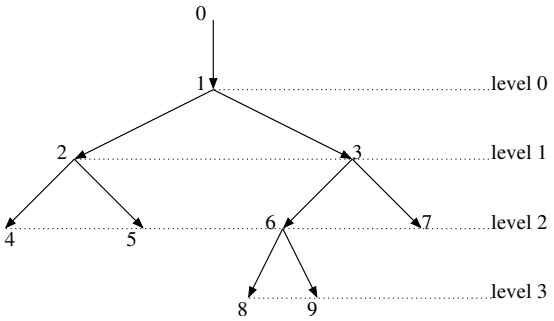


Fig. 7. The labeling of the different nodes and network segments of the measurement tree.

4. 1. Inference in the two-leaf tree

Here we describe the method of queuing-delay distribution inference in the basic case of the two-leaf tree, which is comprised of a source, two receivers, and a single branching node (see the part of fig. 7 until level 1). The generalization to larger trees will be treated in the next subsection. The known quantities are the time series of end-to-end queueing delays Y_2 and Y_3 obtained by minimum filtering the $\tilde{Y}_2$ and $\tilde{Y}_3$ time-series

$$Y_2(n) = \tilde{Y}_2(n) - \min(\tilde{Y}_2), \qquad Y_3(n) = \tilde{Y}_3(n) - \min(\tilde{Y}_3). \tag{3}$$

These are measured by a big number of packet pairs, where the packets in a pair are sent back-to-back from the source, and are destined to the two distinct receivers. We count only those pairs where each of the packets has reached its destination, and assign the number of successful pairs by N. Above we made the assumption that in every end-to-end delay measurement at least one probe will experience no queueing, thus the minimum can be identified with the propagation delay on the given path, which can be subtracted to yield the delay component due to queueing. Measurement studies of references [29,30], as well as our own experience shows that in the current Internet link utilizations are low enough, so that the above assumption is justified. Using the labeling introduced in the previous subsection we have the trivial relations

$$Y_2(n) = X_1(n) + X_2(n), \qquad Y_3(n) = X'_1(n) + X_3(n), \tag{4}$$

where $X'_1(n)$ stands for the queueing delay on segment 1, experienced by that particular probe in the n-th pair, which was destined towards receiver node (3). From this point we will assume perfect correlation of the delays experienced by the two packets in a pair on the common segment $X_1(n) = X'_1(n)$. In practice this ideal situation is approached if for all n, $|X'_1(n) - X_1(n)| << X1(n)$. The goal is to estimate the distribution of the unknown X_1, X_2, and X_3 time-series, based on the knowledge of Y_2 and Y_3.

This can be achieved by introducing the quantized versions of the queueing delays. By using the quantization rule $Y_i^d(n) = jq$, if $(j-1/2)q < Y_i(n) \le (j+1/2)q$, one can map the end-to-end time-series Y_i to their quantized versions Y_i^d, that take values from the set $\{0, q, 2q, \ldots Bq\}$. With the similar quantization we can also introduce the quantized versions of the unknown time-series of X_i^d. We choose the quantization parameters q and B so that they satisfy the relation $\mathbf{max}(Y_2, Y_3) < (B+1)q$, which ensures that all the possible values of Y_i and X_i will be included in one of the bins defined by the quantization rule, and thus probabilities can be

assigned to the bins. We denote the probability of a queueing delay falling in the j-th bin on the i-th segment by $P_{i,j}$, and for each $i \in \{1, 2, 3\}$ we have $\sum_{j=0}^{B} P_{i,j}=1$.

The reason behind the introduction of the discrete quantities above was that using the $P_{i,j}$ quantities one can express the probabilities of finding the quantized end-to-end queueing delays in a given bin $P(l,m) \equiv P(Y_2^d=lq, Y_3^d=mq)$, which can be estimated from the measured data. Assuming temporal and spatial independence of the probe delays these probabilities can be expressed by the convolution

$$P(l,m) = \sum_{k \in H} P_{1,k} P_{2,(l-k)} P_{3,(m-k)}, \tag{5}$$

where the set H is given by $\{B \geq k \geq 0\} \cap \{B \geq (l-k) \geq 0\} \cap \{B \geq (m-k) \geq 0\}$. Labeling the observed variables by $\mathbf{Y} \equiv \{Y_2^d(n), Y_3^d(n)\}$, $(n \in \{1 \ldots N\})$ and the set of unknown probabilities by $\boldsymbol{\theta} \equiv \{P_{i,j}\}$ finally the unknown probabilities $P_{i,j}$ can be determined from the property that their true value maximizes the log likelihood function

$$\log L(\mathbf{Y} \mid \boldsymbol{\theta}) = \sum_{n=1}^{N} \log P(Y_2^d(n), Y_3^d(n)), \tag{6}$$

where the probabilities $P(Y_2^d(n), Y_3^d(n))$ are given by equation (5) evaluated at the known quantized end-to-end queueing delay pairs. Although the log-likelihood-function (6) can not be maximized analytically, there are several numerical procedures to accomplish this goal. Reference [14] suggests to apply the expectation-maximization (EM) algorithm [31].

The EM algorithm is an iterative method to find the maximum likelihood estimate of the unknown probabilities in the presence of some suitably defined hidden or unobserved variables. In our case the hidden variables are the quantized queueing delays on the internal segments of the two-leaf tree, that we denote by $\mathbf{X} \equiv \{X_1^d(n), X_2^d(n), X_3^d(n)\}$, $(n \in \{1 \ldots N\})$. Marginalizing over the hidden variables the condition to maximize equation (6) can be expressed as, for $\forall (i,j)$

$$\frac{\partial \log L(\mathbf{Y} \mid \boldsymbol{\theta})}{\partial P_{i,j}} = \frac{\partial}{\partial P_{i,j}} \left[\sum_{n=1}^{N} \log \sum_{x \in \mathbf{X}(n)} P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta}) + \sum_{b=1}^{3} \lambda_b \left(1 - \sum_{j=0}^{B} P_{b,j} \right) \right] = 0, \tag{7}$$

Above the λ_b quantities are Lagrange-multipliers to account for the normalization constraint of the $P_{i,j}$ probabilities. Performing the derivations we arrive at

$$\begin{aligned}
&\sum_{n=1}^{N} \frac{1}{\sum_{x \in \mathbf{X}(n)} P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta})} \frac{\partial}{\partial P_{i,j}} \left(\sum_{x \in \mathbf{X}(n)} P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta}) \right) - \lambda_i = \\
&\sum_{n=1}^{N} \left[\sum_{x' \in \mathbf{X}(n)} \left(\frac{P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta})}{\sum_{x \in \mathbf{X}(n)} P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta})} \frac{\partial}{\partial P_{i,j}} \log P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta}) \right) \right] - \lambda_i = \\
&\sum_{n=1}^{N} \left[\sum_{x' \in \mathbf{X}(n)} P(\mathbf{X}(n) \mid \mathbf{Y}(n), \boldsymbol{\theta}) \frac{\partial}{\partial P_{i,j}} \log P(\mathbf{Y}(n), \mathbf{X}(n) \mid \boldsymbol{\theta}) \right] - \lambda_i = 0.
\end{aligned} \tag{8}$$

In the above equation the joint probabilities $P(\mathbf{Y}(n),\mathbf{X}(n)\,|\,\boldsymbol{\theta})$ can be factorized as $P(\mathbf{Y}(n)\,|\,\mathbf{X}(n))P(\mathbf{X}(n)\,|\,\boldsymbol{\theta})$. Since $P(\mathbf{Y}(n)\,|\,\mathbf{X}(n))$ does not depend on $\boldsymbol{\theta}$ thus its derivative with respect to $P_{\mathrm{i,j}}$ is zero. Using the assumption of spatial and temporal independence we have

$$P(\mathbf{X}(n)\,|\,\boldsymbol{\theta}) = \prod_{i=1}^{3} P(X_i^d(n)), \tag{9}$$

where $P(X_{\mathrm{i}}^{\mathrm{d}}(n))$ is the probability that the n-th probe experienced a quantized queueing delay of X^d on the i-th segment. Substituting equation (9) into (8) we may write

$$\begin{aligned}&\sum_{n=1}^{N}\left[\sum_{x'\in\mathbf{X}(n)} P(\mathbf{X}(n)\,|\,\mathbf{Y}(n),\boldsymbol{\theta})\frac{\delta(X_i^d(n)=jq)}{P_{i,j}}\right]-\lambda_i = \\ &\frac{1}{P_{i,j}}\sum_{n=1}^{N} P(X_i^d(n)=jq\,|\,\mathbf{Y}(n),\boldsymbol{\theta})-\lambda_i = 0,\end{aligned} \tag{10}$$

After rearranging and using the normalization constraints $\sum^{\mathrm{B}}_{\mathrm{j}=0}P_{\mathrm{i,j}}=1$, finally we arrive at the following system of coupled equations. For $\forall(i,j)$

$$P_{i,j} = \frac{1}{N}\sum_{n=1}^{N} P(X_i^d(n)=jq\,|\,\mathbf{Y}(n),\boldsymbol{\theta}), \tag{11}$$

where $P(X_i^d(n)=jq\,|\,\mathbf{Y}(n),\boldsymbol{\theta})$ stands for the conditional probability of $X_i^d(n)=jq$, given the observed values $Y_2^d(n), Y_3^d(n)$ of the n-th packet pair. According to Bayes-law these conditional probabilities can be expressed as

$$\begin{aligned}P(X_1^d(n)=jq\,|\,Y_2^d(n)=lq,Y_3^d(n)=mq,\boldsymbol{\theta}) &= \frac{P_{1,j}P_{2,(l-j)}P_{3,(m-j)}}{P(l,m)},\\ P(X_2^d(n)=jq\,|\,Y_2^d(n)=lq,Y_3^d(n)=mq,\boldsymbol{\theta}) &= \frac{P_{1,(l-j)}P_{2,j}P_{3,(m-l+j)}}{P(l,m)},\\ P(X_3^d(n)=jq\,|\,Y_2^d(n)=lq,Y_3^d(n)=mq,\boldsymbol{\theta}) &= \frac{P_{1,(m-j)}P_{2,(l-m+j)}P_{3,j}}{P(l,m)},\end{aligned} \tag{12}$$

where $P(l,m)$ is given by equation (5). The system of coupled equations (5,11,12) can be solved iteratively. One starts by assigning initial values for all the $P_{\mathrm{i,j}}$ probabilities (e.g. the evenly distributed probabilities $P_{i,j}^0 = 1/(B+1)$), then equations (11) together with (5) and (12) provide the improved estimates $P'_{i,j}$. This process is iterated until a suitable criterion for convergence is met. We terminate the iterations if for $\forall(i,j)$

$$\left|P'_{i,j}-P_{i,j}\right|B \le \varepsilon \tag{13}$$

where ε is a chosen small number, e.g. 0.0001.

4. 2. Inference in arbitrary trees

For an arbitrary large tree it can be shown that the queueing delay distribution on any segment can be either estimated directly by the inference algorithm developed for the two-leaf tree, or can be generated by numerical deconvolution of two directly inferred distributions. To resolve all the network segments by a minimal approach one must perform packet-pair measurements in a way that all the branching nodes and all the receivers are visited at least in one of the measurements. We demonstrate this in the following example.

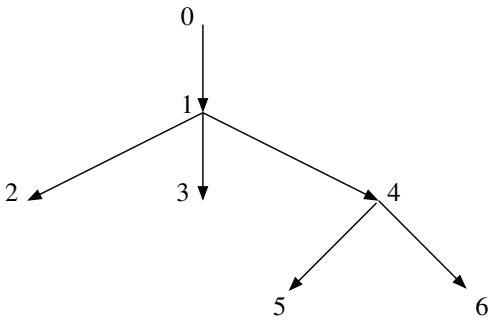


Fig. 8. An example of a measurement tree with 4 receiver nodes.

Consider the tree shown in fig. 8. Performing the back-to-back packet pair measurements to the receiver pairs of (2,3) and (5,6) in a round robin fashion, and applying the inference algorithm developed for the two-leaf tree estimates directly the probabilities $P_{1,\mathrm{j}}$, $P_{2,\mathrm{j}}$, $P_{3,\mathrm{j}}$ and of $P_{(0\to4),\mathrm{j}}$, $P_{5,\mathrm{j}}$, $P_{6,\mathrm{j}}$ ($j \in \{0,1,\ldots B\}$). Here we denoted by $(0\to4)$ the common segment originating in 0 and ending in branching node 4. The probabilities $P_{(0\to4),\mathrm{j}}$ can be expressed by the convolution

$$P_{(0\to4),j} = \sum_{k\in G} P_{1,(j-k)} P_{4,k}, \tag{14}$$

Where the set G is given by $\{B\geq k\geq 0\}\cap\{B\geq(j-k)\geq 0\}$. This convolution can also be viewed as a matrix operation

$$P_{(0\to4),j} = \sum_{k=0}^{B} A_{j,k} P_{4,k}, \tag{15}$$

where the matrix element is $A_{j,k} = P_{1,(j-k)}$, if $k \in G,$ and $A_{j,k} = 0,$ if $k \notin G,$. This way the problem of numerical deconvolution to obtain the remaining probabilities on segment 4 can be mapped to a matrix inversion problem. For this task the best suited numerical method is the non-negative least squares (NNLS) algorithm [32], which guarantees that the resulting probabilities will be strictly non-negative.

In reference [33] we have investigated the performance of the inference method described so far in extensive realistic simulation studies, as well as in controlled LAN experiments with ETOMIC measurement nodes. These investigations revealed impressive agreement of the estimated and the real queueing delay distributions, and an error in the estimation of the first moments being less then the value of the chosen bin size.

5. Conclusion

This paper presented dynamical state measurements of a part of the European Internet, conducted via the ETOMIC measurement infrastructure, using special active probing techniques and the methods of network tomography. This very precise and fully synchronized

infrastructure meets the requirements needed to perform large-scale unicast tomography measurements, and can be viewed as the prototype of network testbeds, that will be able to operate in the uncooperative Internet of the future. In the paper we have investigated a particularly important state variable of the Internet, the characteristics of queueing delay distributions on different network segments, which provides information about traffic properties and the state of congestion in the network. As the main results we presented maps of the averages and standard deviations of queueing delay, and identified in them various structures. We find that the average queueing delay of network segments spans three orders of magnitude, and its distribution function closely follows a log-normal distribution.

Acknowledgements
The authors thank the partial support of the National Office for Research and Technology (NKFP 02 032 04) and the European Union IST FET COMPLEXITY EVERGROW Integrated Project.

References

1 R. Albert, A.-L. Barabási: Statistical Mechanics of Complex Networks. Rev. Mod. Phys. 2002; 74: 47-97.
2 S. Gábor, I. Csabai: The analogies of highway and computer network traffic. Physica A 2002; 307: 516-526.
3 W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Willson: On the self-similar nature of ethernet traffic (extended version). IEEE/ACM Trans. Networking 1994; 2(1): 1-15.
4 M. E. Crovella, A. Bestavros: Self-similarity in World Wide Web traffic: Evidence and possible causes. IEEE/ACM Trans. Networking 1997; 5(6): 835-846.
5 T. Musha, H. Higuchi: 1/fluctuation of a traffic current on an expressway. Jpn. J. Appl. Phys. 1976; 15(7): 1271-1275.
6 I. Csabai: 1/f noise in computer network traffic. J. Phys. A-Math. Gen. 1994; 27(12): 417-421.
7 M. Takayasu, H. Takayasu, T. Sato: Critical behaviors and 1/f noise in information traffic. Physica A 1996; 233(3-4): 824-834.
8 M. Faloutsos, P. Faloutsos, C. Faloutsos: On power-law relationships of the Internet topology. Comput. Commun. Rev. 1999; 29(4): 251-262.
9 R. Govindan, H. Tangmunarunkit: Heuristics for Internet Map Discovery. Proceedings of the IEEE Infocom, March 2000.
10 R. Albert, H. Jeong, A.-L. Barabási: Error and attack tolerance of complex networks.Nature 2000; 406(6794): 378-382.
11 A. Vazquez, R. Pastor-Satorras, A. Vespignani: Large-scale topological and dynamical properties of the Internet. Phys. Rev. E 2002; 65(6): 066130.
12 M. A. de Menezes, A.-L. Barabási: Fluctuations in Network Dynamics. Phys. Rev. Lett. 2004; 92(2): 028701.
13 Active Measurement Project, http://amp.nlanr.net/.
14 M. Coates, R. Nowak: Network tomography for internal delay estimation. In Proc. Of IEEE International Conference on Acoustics, Speech and Signal Processing, Salt Lake City, May 2001.
15 N. Duffield, J. Horowitz, F. L. Presti, D. Towsley: Network Delay Tomography from End-to-end Unicast Measurements, Springer Verlag - Lecture Notes in Computer Science, Berlin, 2001.
16 M. Coates, R. Nowak: Network loss inference using unicast end-to-end measurement. In Proc. ITC Conf. IP Tra_c, Modeling and Management, Monterey, CA, September 2000.
17 N. Duffield, F. L. Presti, V.Paxson, D. Towsley: Inferring link loss using striped unicast probes. In Proc. of IEEE Infocom 2001, Anchorage, AK, April 2001.
18 R. Cáceres, N. G. Duffield, S. B. Moon, D. Towsley: Inference of internal loss rates in the mbone, In Proc. of IEEE/ISOC Global Internet, December 1999.

19 R. Cáceres, N. G. Duffield, J. Horowitz, D. Towsley, T. Bu: Multicast-based inference of network-internal characteristics: Accuracy of packet loss estimation, In Proc. of IEEE Infocom, New York, March 1999.
20 N. Duffield, F. L. Presti: Multicast inference of packet delay variance at interior network links. In Proc. of IEEE Infocom 2000, Tel Aviv, Israel, March 2000.
21 F. L. Presti, N. Du_eld, J. Horowitz, D. Towsley: Multicast-based inference of network-internal delay distributions. IEEE/ACM Transactions on Networking, December 2002.
22 T. Bu, N. Duffield, F. L. Presti, D.Towsley: Network tomography on general topologies. In Proc. of ACM Sigmetrics, Marina del Rey, CA, 2002.
23 M. Rabbat, R. Nowak, M. J. Coates: Multiple source, multiple destination network tomography. In Proc. of IEEE Infocom, Hong Kong, March 2004.
24 EVERGROW homepage, http://www.evergrow.org/.
25 ETOMIC homepage, http://www.etomic.org/.
26 D. Morató, E. Magaña, M. Izal, J. Aracil, F.Naranjo, F. Astiz, U. Alonso, I. Csabai, P. Hága, G. Simon, J. Stéger, G. Vattay: The european traffic observatory measurement infrastructure (etomic): A testbed for universal active and passive measurements, In Proc. of Tridentcom, Trento, Italy, February 23, 2005
27 ENDACE homepage, http://www.endace.com/.
28 GÉANT homepage, http://www.geant.net/.
29 S. Moon, P. Skelly, D. Towsley: Estimation and removal of clock skew from network delay measurements. In Proc. of Infocom, New York, March 1999.
30 V. Paxson: On calibrating measurements of packet transit times. In Proc. of SIGMETRICS, Madison, WI, June 1998.
31 G. J. McLachlan, T. Krishnan: The EM algorithm and extensions. John Wiley, New York, 1997.
32 C. L. Lawson, B. J. Hanson: Solving Least Squares Problems, Prentice-Hall, Englewood Cli_s, NJ, 1974.
33 G. Simon, P. Hága, G. Vattay, I. Csabai: A flexible tomography approach for queueing delay distribution inference in communication networks. In Proc. of IPS-MoMe, Warsaw, Poland, March 14-15 2005.

A Simulation Study of Network Discovery Strategies*

Felix Eberhard, Alexander Hall†
ETH Zürich

Thomas Erlebach‡
University of Leicester

Abstract

Due to its fast, dynamic, and distributed growth process, it is hard to obtain an accurate map of the Internet. In many cases such a map—representing the structure of the Internet as a graph with nodes and links—is a prerequisite when investigating properties of the Internet. A common way to obtain such maps is to make certain local measurements at a small subset of the nodes and then to combine these in order to "discover" (an approximation of) the actual graph. Each of these measurements is potentially quite costly. It is thus a natural objective to minimize the number of measurements which still discover the whole graph. We consider this problem for a specific type of measurements and compare four simple greedy strategies in an experimental analysis. Our results show that one can discover accurate information about the structure of large and complex networks using a surprisingly small number of queries.

KEYWORDS: Internet discovery, simulation experiments, complex networks, random graphs.

1 INTRODUCTION

An important aspect in the study of complex networks is the methodology that is used to obtain information about the nodes and links of an unknown network. Before the structure of a network can be analyzed and interpreted, one needs to measure the network in order to discover its nodes and links. In many cases, measuring the network is a nontrivial task, and obtaining accurate and complete information is a challenging problem. The questions arising include: What type of measurements should be carried out? How many measurements are needed for different types of networks? What is the best strategy for minimizing the number of measurements needed to get an accurate reconstruction of the network?

We are interested in the problem of discovering the presence and absence of links in an unknown communication network. The prime example of a communication network whose structure is difficult to determine is the Internet. Owing to its large scale and distributed growth, there is no easy method to obtain an accurate and complete map of the Internet. We consider a model of network discovery in which the set of nodes of the network is known in advance and where a measurement at a node v yields the set of all edges on shortest paths between v and any other node of the graph. We refer to a measurement carried out at a node v as a query at v. This model is motivated by approaches to discovering the Internet (on the router level or on the level of autonomous systems) that are based on traceroute experiments [1, 2] from selected sources or on the analysis of BGP routing tables at

*Work partially supported by European Commission - Fet Open project DELIS IST-001907 "Dynamically Evolving Large Scale Information Systems."

†E-mail: alex.hall@gmail.com

‡Corresponding author. Department of Computer Science, University of Leicester, University Road, Leicester LE1 7RH, UK. Phone: +44-(0)116-2523411. Fax: +44-(0)116-2523604. E-mail: t.erlebach@mcs.le.ac.uk

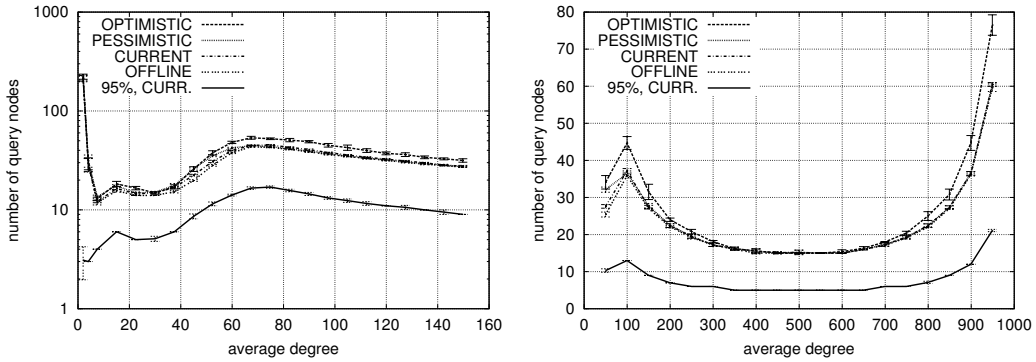


Figure 1: $G_{n,p}$: Erdős-Rényi random graphs on $n = 1{,}000$ nodes. Between a pair of nodes an edge is present with probability p. This parameter is varied. In the charts the expected degree $p \cdot (n-1)$ of a node is given on the abscissa.

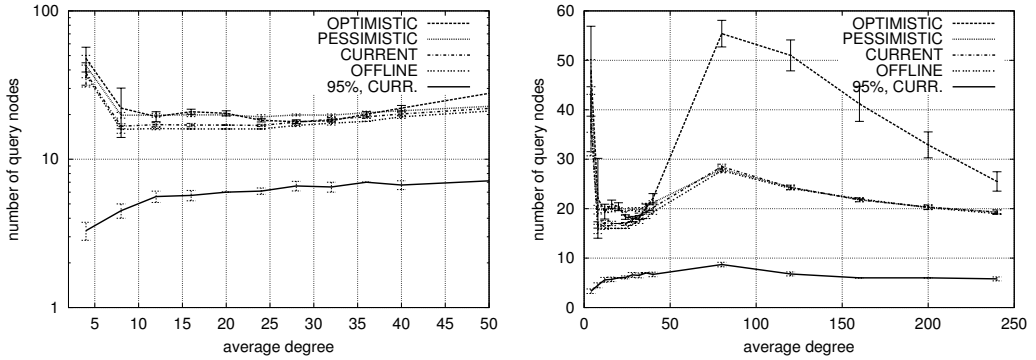


Figure 2: Barabási-Albert random graphs. The average degree of a node is given on the abscissa.

selected BGP routers [3]. A query at v in our model roughly corresponds to carrying out traceroute experiments at v, or obtaining the BGP routing table at v. We implement several discovery strategies and study experimentally the number of queries they need to discover different types of networks.

2 MODEL AND SIMULATION SETUP

Following [4], we model the network discovery problem as follows. The network is represented as an undirected graph with n nodes. A link between two nodes of the network corresponds to an edge between these two nodes in the graph. If there is no link between two nodes u and v, we say that there is a non-edge between u and v. A query at a node v discovers all edges and non-edges whose endpoints have different distance from v. A network discovery strategy makes queries until all edges and non-edges have been discovered. For the selection of the next query, the results of all previous queries can be taken into account. The goal is to use as few queries as possible to discover all edges and non-edges. The difficulty in selecting good queries arises from the fact that the amount of information discovered by a query may depend on the parts of the network structure that are still unknown.

We have implemented several strategies for selecting the next query vertex: Strategy CURRENT selects the next query so as to maximize the number of newly discovered non-edges under the as-

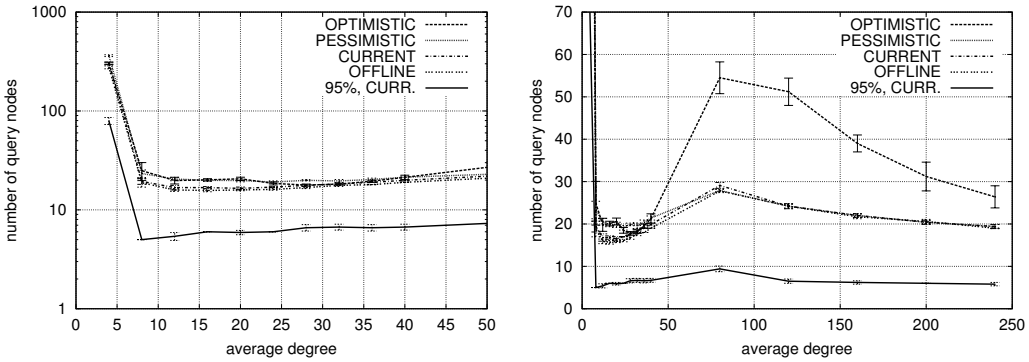


Figure 3: Dorogovtsev-Mendes-Samukhin random graphs. The average degree of a node is given on the abscissa.

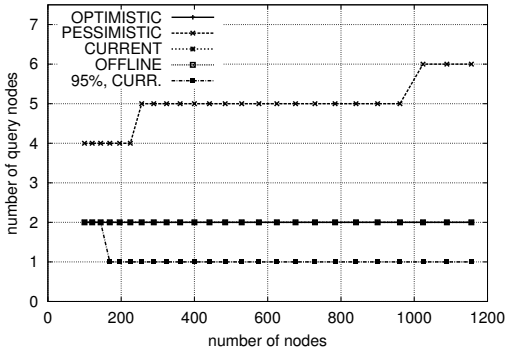


Figure 4: Grids of increasing sizes from 10×10 to 34×34. The number of nodes is given on the abscissa.

sumption that all edges have already been discovered. Strategies OPTIMISTIC and PESSIMISTIC compute upper and lower bounds on the number of newly discovered edges and non-edges at each potential query vertex and select the query that maximizes the upper or lower bound, respectively. For comparison, we have also implemented a strategy OFFLINE that knows the network and computes, using a greedy set-cover heuristic, a small set of queries to discover the whole network. Since OFFLINE has full information, it can be expected to perform better than the three other strategies, and we thus employ it as a benchmark strategy.

In our simulation experiments, we use various types of network graphs (Erdős-Rényi random graphs [5], Barabási-Albert scale-free random graphs [6], Dorogovtsev-Mendes-Samukhin scale-free random graphs [7], and grid graphs) and run the implemented discovery strategies on them.

3 RESULTS

First, we compare the different strategies with each other concerning the number of queries required to discover the whole graph in each of the different random graph models. We generate graphs with $n = 1{,}000$ nodes and varying average degree, and we record the average number of queries that each of the strategies requires. The results are shown in Figures 1–3. The charts on the left-hand side cover smaller degrees and use a logarithmic scale on the vertical axis, whilst those on the right-hand side

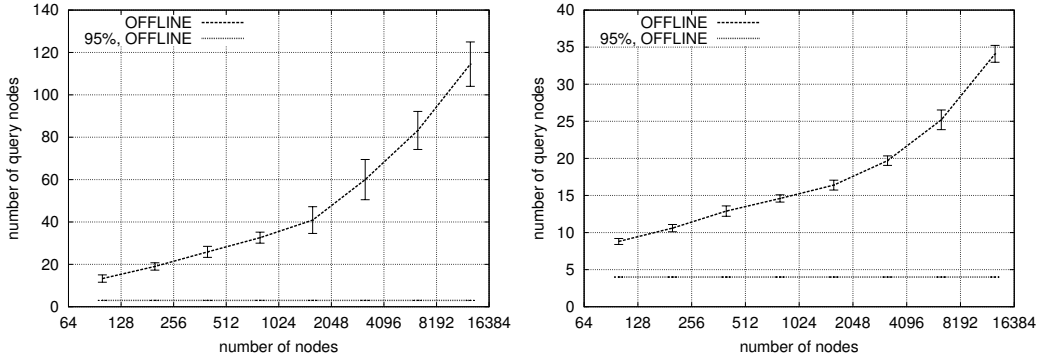


Figure 5: Barabási-Albert random graphs. The average degree is fixed to 4 for the left chart and to 10 for the right chart. The number of nodes is increased in exponential steps.

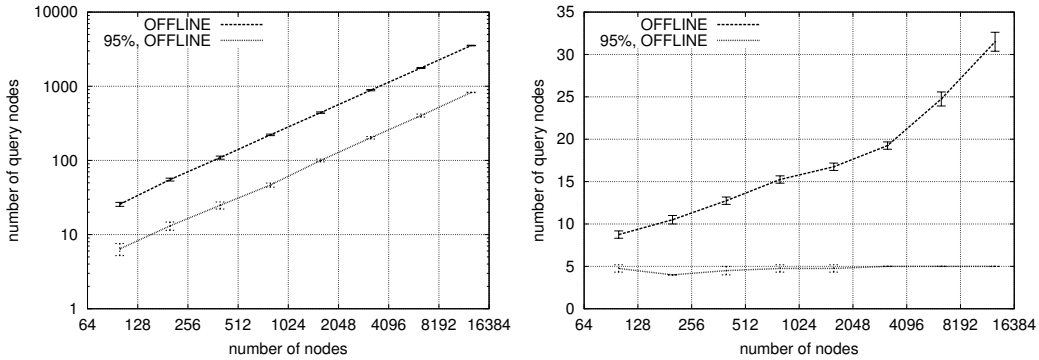


Figure 6: Dorogovtsev-Mendes-Samukhin random graphs. The average degree is fixed to 4 for the left chart and to 12 for the right chart. The number of nodes is increased in exponential steps.

cover a larger range of degrees and use a linear scale. We find that the number of queries required by PESSIMISTIC, CURRENT and OFFLINE is roughly the same and slightly smaller than that of OPTIMISTIC. This indicates that a good selection of queries can be made in spite of the initial lack of information about the graph structure. We also find that the absolute number of queries required to discover a network is often surprisingly small. For networks with 1,000 nodes and average degree ranging from 10 to 250, only 10–60 queries are usually sufficient to discover the whole network. This holds for all three types of random graphs.

Grid graphs were the only graphs for which OPTIMISTIC outperformed PESSIMISTIC (see Figure 4; here, we have varied the number of nodes). This is due to the special property of grids that two queries (in two adjacent corners) are sufficient to discover the whole graph, independent of the number of nodes.

In the charts of Figures 1–4 we also show the number of queries required by CURRENT in order to discover 95% of the edges and 95% of the non-edges (labeled "95%, CURR."). It is a striking observation that more than 10 queries are rarely needed for this purpose.

We have also studied how the number of queries required to discover scale-free random graphs changes if the size of the network grows. The results for strategy OFFLINE are shown in Figures 5 and 6. Each of the charts shows the number of queries needed for networks with increasing number of nodes and fixed average degree. As expected, one finds that when the number of nodes is increased,

the number of queries required to discover the network increases as well. However, when checking how many queries suffice to discover 95% of the edges and 95% of the non-edges of a graph (labeled "95%, OFFLINE"), we notice that for Barabási-Albert graphs this number appears to be a constant smaller than 10 as the number of nodes varies from 100 to 12,800. This indicates that extremely few queries can be sufficient to discover a huge portion of an unknown network, independent of the size of the network. For Dorogovtsev-Mendes-Samukhin random graphs with average degree 4, however, the number of queries for 95% discovery grows linearly with the number of nodes. Only when the average degree is larger (achieved by varying the generation process so that each new node is made adjacent to the endpoints of several randomly chosen edges), the number of queries required for 95% discovery appears to be a constant for this model as well; see the chart on the right-hand side of Figure 6, where the average degree is fixed to 12.

4 CONCLUSION

Our simulation results show that, under the query model considered, it is possible to discover accurate information about the structure of large and complex networks using a moderately small number of queries. If the goal is to discover 95% of the edges and 95% of the non-edges, it even appears that a constant number of queries is often sufficient, independent of the network size. In future work, we would like to investigate this effect from a more theoretical perspective. Furthermore, we plan to study the influence of weaker query models on the number of queries required to discover a network.

References

[1] Dall'Asta L, Alvarez-Hamelin I, Barrat A, Vázquez A, Vespignani A: A statistical approach to the traceroute-like exploration of networks: theory and simulations. *Workshop on Combinatorial and Algorithmic Aspects of Networking (CAAN'04)*, LNCS 3405, Springer, pages 140–153, 2005.

[2] DIMES Project. http://www.netdimes.org/

[3] University of Oregon Route Views Project. http://www.routeviews.org/

[4] Beerliova Z, Eberhard F, Erlebach T, Hall A, Hoffmann M, Mihaľák M, Ram LS: Network Discovery and Verification. *31st International Workshop on Graph-Theoretic Concepts in Computer Science (WG'05)*, LNCS, Springer, 2005. To appear.

[5] Erdős P, Rényi A: On Random Graphs I, *Publ. Math. Debrecen*, 6:290–297, 1959.

[6] Barabási AL, Albert R: Emergence of scaling in random networks, *Science*, 286:509–512, 1999.

[7] Dorogovtsev SN, Mendes JFF, Samukhin AN: Size-dependent degree distribution of a scale-free growing network, *Phys. Rev. E*, 63:062101, 2001.

Evolutionary Game Theory with Applications to Adaptive Routing*

Simon Fischer† Berthold Vöcking‡

Abstract

One of the most important problems in large communication networks like the Internet is the problem of routing traffic through the network. Current Internet technology based on the TCP protocol does not route traffic adaptively to the traffic pattern but uses fixed end-to-end routes and adjusts only the injection rates in order to avoid congestion. A more flexible approach uses load-adaptive rerouting policies that reconsider their routing strategies from time to time depending on the observed latencies. In this manuscript, we survey recent results from [1, 2] about the application of methods from evolutionary game theory to such an adaptive traffic management.

Key words: Evolutionary game theory, Wardrop model, adaptive routing, stale information, analysis and dynamics of complex networks

1 Introduction

Recently, game theoretical analysis of the Internet has attracted a growing amount of interest. One of the models studied in this context is the Wardrop model [3] in which each of an infinite number of selfish users controls an infinitesimal amount of traffic that is to be routed through a network. Due to a lack of central coordination, agents strive to minimise their latency selfishly. Classical game theory predicts that agents will assign their traffic according to a Nash equilibrium, i. e., in a way such that no agent has an incentive to change their routing strategy unilaterally. For this model, a number of interesting results have been found. For example, Beckmann et al. [4] shows existence and essential uniqueness of Nash equilibria. Roughgarden and Tardos [5] prove bounds on the so-called *price of anarchy* which is the worst-case ratio between the social welfare (e. g., the average latency) at a Nash equilibrium and the social welfare at an optimal assignment. Other results show how to impose taxes on the agents such that social optimum and Nash equilibrium coincide [6, 7].

Nash equilibria are interesting from a practical point of view as they represent stable and fair allocations. Classical game theory, however, relies on several assumptions that do not seem to be practical. In particular, players are assumed to have full and accurate information about the game and also about the behaviour of the opponents and must act completely rationally. It is questionable whether these assumptions are satisfied when the game under study should model the Internet. Quite obviously, participants in the Internet have incomplete and inaccurate knowledge. Furthermore, they have bounded rationality.

*Supported in part by the EU within the 6th Framework Programme under contract 001907 (DELIS) and by DFG grant Vo889/1-2.

†*Corresponding author;* address: RWTH Aachen, Computer Science 1, D-52056 Aachen, Germany; phone +49 241 80 21108; fax +49 241 80 22216; email: fischer@cs.rwth-aachen.de

‡RWTH Aachen, Germany, voecking@cs.rwth-aachen.de

A more reasonable assumption might be that players learn how to perform well in the game by experience. This is where evolutionary game theory (see e. g. [8, 9]) comes into play. Here, all of the above assumptions are dropped and instead it is assumed that the game is played repeatedly against random opponents. Over time, agents have the chance to optimise their personal cost by reacting to simple observations. Describing a process in which agents adapt their behaviour to the overall situation based on the observed payoff, we can formulate a system of differential equations. Our hope is that the solution concepts of differential equations like attractors or asymptotically stable rest points coincide with Nash equilibria. This is not the case for games in general. In order to ensure this, a refinement of Nash equilibria, *evolutionary stability*, is required. We can show that for our class of games, Nash equilibria possess this property, thus enabling us to show that load adaptive rerouting policies converge towards Nash equilibria. This provides additional motivation to the well-studied concept of Nash equilibria and the many analyses based on this concept that have been performed so far. In addition, we give bounds on the speed of convergence. One practically very important aspect in adaptive routing is the effect of stale information. In applications, the latency information that rerouting decisions are based on is typically not up to date. It is well known, that basing routing decisions on stale information can cause oscillation effects and seriously harm network performance. In a simplified model, we show how this can be avoided using a class of adaptive rerouting policies that is not too greedy.

2 The Model

2.1 Selfish Routing

For any of a set of *commodities* $i \in \{1, \dots, k\}$, a fraction of r_i agents wants to route an equivalent amount of flow from source s_i to sink t_i using paths from the set of paths $\mathcal{P}_i$ connecting these two nodes. Let $\mathcal{P} = \cup_{i\in[k]}\mathcal{P}_i$ where $[k] = \{1, \dots, k\}$. A *feasible flow* on this network is a positive real-valued vector $(f_P)_{P\in\mathcal{P}}$ that satisfies the flow demands, i. e., $\sum_{P\in P_i} f_P = r_i$ for all $i \in [k]$. Given a flow vector f, the flow on an edge $e \in E$ is $f_e = \sum_{P\ni e} f_P$. The latencies in the network are specified by strictly increasing functions $\ell_e : [0, 1] \mapsto \mathbb{R}_0^+$. Given a flow vector f, the latency on edge $e \in E$ is given by $\ell_e(f_e)$ and the latency of path $P \in \mathcal{P}$ is given by $\ell_P(f) = \sum_{e\in P} \ell_e(f_e)$. A flow vector f is said to be at a *Nash equilibrium* if for any commodity $i \in [k]$ and every pair of paths $P, P' \in \mathcal{P}_i$ with $f_P > 0$ it holds that $\ell_P(f) \le \ell_{P'}(f)$. In other words, at a Nash equilibrium, no agent has an incentive to change their routing strategy.

In a scenario where routing is not controlled by a central authority but by individually and selfishly acting agents, the concept of Nash equilibria makes sense. Though the concept of Nash equilibria seems to be a natural and appealing stability concept for the above scenario, it neglects the question of how the agents can come to such a state. In fact, this requires very accurate knowledge about the network topology, latency functions, the demands of other commodities, and the behaviour of the other agents as well as full rationality. These assumptions are quite obviously not satisfied if the network under study is the Internet. The following section will present some very simple processes which, in contrast to the assumptions of classical game theory, require only very local knowledge and almost no computational effort at all and will turn out to result in Nash equilibria.

2.2 Rerouting Dynamics

Consider a large population of agents in a routing network, each agent choosing one of the possible routing paths. If the number of agents in this scenario is infinite, then a vector $(x_P)_{P\in\mathcal{P}}$ specifying the fractions of agents using each individual path $P \in \mathcal{P}$ is actually equivalent to a feasible network flow.

Now assume that every agent wants to optimise their personal latency having virtually no knowledge about the flow and demands of other agents. One reasonable behaviour for an agent would be to reconsider their routing strategy from time to time, at Poisson rates, say, and revise their strategy based on the observed performance. An agent might then pick another routing path at random (e. g., with probability proportional to the population share currently using this path) and compare the latency of the own path with the latency of the other. If the other path turns out to be worse than the current routing strategy, nothing happens. However, if the other path offers an improvement with respect to latency, the agent might switch to the new path with a probability proportional to the size of the latency gain. If we take this process to the fluid limit, i. e., letting the number of agents go to infinity and identifying random variables describing the change of the population shares in one step with their expectation values, we obtain an expression for the change rate of the population shares:

$$\dot{f}_P = \lambda_i \cdot f_P \cdot (\bar{\ell}_i - \ell_P) \qquad \text{for } i \in [k], P \in \mathcal{P}_i, \tag{1}$$

where $\dot{}$ indicates the derivative with respect to time, $\bar{\ell}_i$ is the average latency of commodity i, and λ_i is some factor that accounts for proportionality factors needed to ensure that probabilities do not exceed 1 etc. Note that the solution orbit $\{\xi | \exists t \geq 0 : \xi = f(t)\}$ of this system of differential equations is independent of the scale of the vector $(\lambda_i)_{i\in[k]}$ as long as all $\lambda_i > 0$. Scaling all λ_i by the same factor scales the speed at which orbits are traversed by this factor. Equation (1) has several appealing properties and has therefore been studied extensively in the evolutionary game theory literature. It is known as the *replicator dynamics* (for a survey of this and other dynamics, see, e. g., [8]).

A natural generalisation of this dynamics that preserves the aspect of local control is a class of rerouting policies that consists of two steps. Again, agents are activated at Poisson rates. Once activated, an agent performs two steps:

1. Sampling: Pick a path Q at random with probability σ_Q. In the most simple case we have $s_Q = 1/m$ where m is the number of paths of the agent's commodity. For the replicator dynamics we have $\sigma_Q = f_Q$, i. e., the probability to sample path Q is proportional to the fraction of agents using it.

2. Migration: Migrate from the current path P to path Q with probability $\mu(\ell_P, \ell_Q)$. For the class of *better response dynamics* we have $\mu(\ell_P, \ell_Q) = 1$ if $\ell_Q < \ell_P$ and $\mu(\ell_P, \ell_Q) = 0$ otherwise. For the replicator dynamics we have $\mu(\ell_P, \ell_Q) = \max\{(\ell_P - \ell_Q) \cdot \lambda, 0\}$ where we choose λ small enough such that the probability is bounded from above by 1.

Altogether, we can specify the rate r_{PQ} at which agents move from path P to path Q and finally the time derivatives of the population shares. For all commodities $i \in \{1, \ldots, k\}$ and all paths $P, Q \in \mathcal{P}_i$ we have

$$r_{PQ} = f_P \cdot \sigma_Q \cdot \mu(\ell_P, \ell_Q) \quad \text{and} \quad \dot{f}_P = \sum_{Q\in\mathcal{P}_i} (r_{QP} - r_{PQ}). \tag{2}$$

Let us remark that by the Picard-Lindelöf-Theorem [10] a unique solution to this system of differential equations exists if the right-hand sides are Lipschitz continuous. We therefore require the latency functions ℓ_e, $e \in E$ as well as σ and μ to be Lipschitz continuous. However, even for linear latency functions, Equation (1) contains cubic terms thus rendering an analytic solution impossible.

3 Equilibria and Convergence

The first natural question is whether the above dynamics actually converge towards Nash equilibria in the long run. More formally, we want to show that Nash equilibria are global attractors of our system of differential equations. Since the replicator dynamics is not *innovative*, i. e., $f_P(t) = 0$ implies that $f_P(t') = 0$ for all $t' \geq t$, the replicator dynamics can never discover such unused paths. We therefore assume for the rest of the paper that there are no unused links in the initial population.

For the purpose of showing convergence, *evolutionary stability* has been introduced as an equilibrium concept which is stricter than the concept of Nash equilibria. For the single-commodity case, evolutionary stability can be characterised as follows [9].

Definition 1 (evolutionary stable). *A flow vector $f \in \Delta$ is called* evolutionary stable *iff (1) it is a Nash equilibrium and (2) for all best replies $\tilde{f}$ to f, $\tilde{f} \neq f$ it holds that $\tilde{f} \cdot \ell(\tilde{f}) > f \cdot \ell(\tilde{f})$.*

In our scenario, a *best reply* to a flow vector f corresponds to a flow vector $\tilde{f}$ that uses only minimum latency paths with respect to the latency vector ℓ induced by f. Since Nash equilibria are not in general unique, but only unique with respect to the edge-flows $(f_e)_{e \in E}$, we say that a flow vector f is *essentially evolutionary stable* if condition (2) above holds for all best replies $\tilde{f}$ that differ from f for at least one edge $e \in E$ (instead of one path $P \in \mathcal{P}$). Then we can show the following lemma.

Lemma 1 ([1]). *For single-commodity networks, Nash equilibria are essentially evolutionary stable.*

Given this property, we can prove convergence of the replicator dynamics towards Nash equilibria for the single-commodity case using standard techniques of evolutionary game theory [9]. For the multi-commodity case and for general dynamics of the form of Equation (2), the proof is an application of Lyapunov's second method [10] in conjunction with a Potential function introduced by Beckmann etal. [4].

Theorem 2 ([1, 2]). *In terms of edge flows $(f_e)_{e \in E}$, the replicator dynamics (1) (provided that $f_P(0) > 0$ for all $P \in \mathcal{P}$) and, more generally, all dynamics of the form of Equation (2) (provided that σ_Q is always positive and σ and μ are continuous) converge towards a Nash equilibrium.*

From the computer scientists' perspective we are interested in the time until our dynamics reach equilibria. Clearly, in the continuous fluid limit model, equilibria cannot be reached exactly but merely approximated. Considering the single-commodity case, we define approximate equilibria as follows. Let $\mathcal{P}_\epsilon$ be the set of paths that have latency at least $(1+\epsilon) \cdot \bar{\ell}$, i. e., $\mathcal{P}_\epsilon = \{P \in \mathcal{P} \mid \ell_P(f) \geq (1+\epsilon) \cdot \bar{\ell}\}$ and let $f_\epsilon := \sum_{P \in \mathcal{P}_\epsilon} f_P$ be the fraction of agents using these paths. A population f is said to be at an *ϵ-approximate equilibrium* if and only if $f_\epsilon \leq \epsilon$.

Note that by definition of the replicator dynamics, we cannot, in general, expect to reach a state where $f_\epsilon = 0$ since a population share using a path with constant high latency

will never completely vanish. Similarly, it can take arbitrarily long, until many agents are within a factor of $(1+\epsilon)$ of the minimum latency path since initially there may be an arbitrarily small fraction of agents on this link. Our definition takes care of these two effects.

In order for the time to reach an approximate equilibrium to be meaningful, the dynamics must not depend on the scale by which we measure latency. We use the parameter λ to normalise our dynamics and choose $\lambda = 1/\bar{\ell}$. For single-commodity networks the replicator dynamics with $\lambda = 1/\bar{\ell}$ converges to an ϵ-approximate equilibrium within time $\mathcal{O}(\epsilon^{-3} \cdot \ln(\ell_{\max}/\ell^*))$ where ℓ^* is the optimal average latency and $\ell_{\max}$ is the maximum latency of a path over all possible flows. Somewhat weaker bounds on the speed of convergence in multi-commodity games can be found in [1] as well.

4 Stale Information

One of the most important problems in load adaptive routing is the fact that information about latency, delay, or bandwidth may be out of date by the time it is gathered. It is well-known that this can cause oscillation effects and seriously harm performance. Mitzenmacher [11] introduced the bulletin board model to study these effects. In this model, all information relevant to the rerouting process is stored on a centralised bulletin board that is accessible to all agents. However, information on the bulletin board is not always up to date but merely updated every T units of time. The bulletin board can be a model for a scenario where latency information is broadcasted to the agents at intervals or where this information is stored on a server from which it can be polled by the agents. Let us remark, that this is a purely theoretical model which, however, exhibits the effects we want to study.

It is clear that two conditions can cause the policy to oscillate. First, if small changes in the flow on an edge can cause a large change in latency, then agents must migrate to this edge very carefully in order not to overshoot the balanced state. Hence, we consider only latency functions of bounded slope, i. e., we consider some β such that $\ell_e'(x) \le \beta$ for all $e \in E$, $x \in [0,1]$. Second, agents must not move to fast if the observed latency difference between the two considered paths is small as this could otherwise cause the same effect. We consider a number $\alpha > 0$ such that $\mu(\ell_1, \ell_2) \le \alpha(\ell_1 - \ell_2)$ for $\ell_1 \ge \ell_2$ and $\mu(\ell_1, \ell_2) = 0$ if $\ell_1 < \ell_2$. We call rerouting policies satisfying this property *α-smooth*. Finally, let L be the length of the longest path in the network. Given these properties we can show that our policies converge towards a Nash equilibrium provided that the bulletin board is updated frequently enough.

Theorem 3 ([2]). *If the update frequency $1/T \ge 4\,L\,\alpha\,\beta$ and σ_P assigns non-zero probabilities to all paths $P \in \mathcal{P}$, then the solution of dynamics (2) converges towards a Nash equilibrium in the bulletin board model.*

We can also show that the parameters that go into our upper bound are actually necessary in the following sense. We say that a function $\mathbf{x}(\cdot)$ *oscillates* if for some $\tau > 0$ and some t_0 it holds that $\mathbf{x}(t_0) = \mathbf{x}(t_0 + n\tau)$ for all $n \in \mathbb{N}$. A differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ *oscillates* if there exists a boundary condition $\mathbf{x}(0) = \mathbf{x_0}$ such that the solution orbit oscillates.

Theorem 4. *For any α, β, and L with $\alpha\,\beta\,L \ge 1$ there exist a network in which the longest path has length L, latency functions whose slope is bounded by β, and an α-smooth migration rule such that for any $T > 8/(\alpha\,\beta\,L)$ the differential equation related to the dynamics oscillates.*

5 Conclusion

We have argued that the assumptions of full information and unbounded rationality made by classical game theory in order to motivate the central concept of Nash equilibria are not realistic with respect to routing in large networks. Instead, we have described a simple class of rerouting policies that, in the fluid limit, converge towards Nash equilibria. These policies do not rely upon these assumptions but merely on very simple observations and little computation. Thus, we have strengthened the motivation of Nash equilibria.

We have also given upper bounds on the time of convergence towards approximate equilibria. Furthermore, we have studied the effects of stale information which often imposes significant performance degradation on practical applications. We have shown bounds on the necessary update time for the information depending on smoothness parameters of the network such that these effects can be avoided.

References

[1] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing. In *Proc. 12th Ann. European Symp. on Algorithms (ESA)*, number 3221 in Lecture Notes in Comput. Sci., pages 323–334, Bergen, Norway, September 2004. Springer-Verlag.

[2] Simon Fischer and Berthold Vöcking. Adaptive routing with stale information. In *Proc. 24th Ann. ACM SIGACT-SIGOPS Symp. on Principles of Distributed Computing (PODC)*. ACM, July 2005.

[3] John Glen Wardrop. Some theoretical aspects of road traffic research. In *Proc. of the Institute of Civil Engineers, Pt. II*, pages 325–378, 1952.

[4] M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics and Transportation*. Yale University Press, 1956.

[5] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[6] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. How much can taxes help selfish routing? In *Proc. 4th ACM Conference on Electronic Commerce*, pages 98–107, 2003.

[7] Lisa Fleischer. Linear tolls suffice: New bounds and algorithms for tolls in single source networks. In *Proc. 31st Int. EATCS Coll. on Automata, Languages and Programming (ICALP)*, pages 544–554, 2004.

[8] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.

[9] Jörgen W. Weibull. *Evolutionary Game Theory*. MIT press, 1995.

[10] William Boyce. *Elementary Differential Equations and Boundary Value Problems*. John Wiley & Sons, 1965.

[11] Michael Mitzenmacher. How useful is old information? In *Proc. 16th Ann. ACM SIGACT-SIGOPS Symp. on Principles of Distributed Computing (PODC)*, pages 83–91, 1997.

Measuring Preferential Attachment in a Hyper-Textual Dictionary Reference Network: Eksi Sözlük*

Amac Herdagdelen[1], Eser Aygun[2], Haluk Bingol[1]

[1] Department of Computer Engineering, Bogazici University
Phone: +90 212 359 45 23-24, Fax: +90 212 287 24 61
Istanbul, TR-34342, Turkey
yalanciborsaci@yahoo.com, bingol@boun.edu.tr
[2] Department of Computer Engineering, Istanbul Technical University
Istanbul, TR-34469, Turkey
aygunes@itu.edu.tr

Abstract. Existing methods used to analyze the relation between preferential attachment and node degree make use of time dependent measures, which result in limited ability to analyze the temporal characteristics of networks. We introduce time independent measures, which allow us to analyze the networks' preferential attachment behavior in a more precise manner. The two different methodologies are compared on a new complex network data: Eksi Sözlük, which spans the whole lifetime (six years) of a complex network with very precise recordings of the node and edge addition events (i.e. one minute). The relation between the likeliness to receive new links and the present degree of a node is found to be linear. Analyses suggest that time independent measures are better in capturing the dynamics of the network and in some cases, provide results that are very hard to obtain by existing methodologies.

Keywords. Preferential attachment, network growth models, temporal analysis, network dynamics, link acquisition.

* This work was partially supported by Bogazici University Research Projects under the grant number 05A105.

1 Introduction

One of the most profound discoveries in complex network studies was realizing that the structure and dynamics of many real world networks do not follow a completely random but rather organized behavior ([1]). Most of the complex networks seem to share some basic properties such as power-law degree distribution, small average path lengths, and high clustering. Among them, the power-law degree distribution has attracted a considerable attention because it is a significant deviation from random behavior and it serves as a basis for the scale-free network concept ([1], [2]). In this study, we focus on the dynamics that lead to power-law degree distributions.

The ubiquity of power-law degree distribution led the scientists to the research of finding out why networks from so diverse origins show the same characteristics and understanding its implications. A line of this research is building up analytical models and comparing them with the real-world data. The original Barabási-Albert (BA) model ([1]) and its generalizations (see [3] for a discussion), growing network model ([4]), growth and deactivation model ([5]), and local models proposed by [6] are among these studies. It is worth noting that in these modeling efforts, an important aim is to be able to model the dynamics of the network, hence they are called growth models and they do not only try to capture the structural properties but also the dynamic evolution of the networks in time.

An important concept introduced by the BA model is the preferential attachment hypothesis and it has deep connections with the power-law distribution. According to the hypothesis, a node with a relatively high degree will be more likely to acquire new links in the future. Almost all scale-free models either assume the existence of preferential attachment and explicitly incorporate it ([1], [3], [4]) or expect it to emerge from the interaction between the growth and dynamics of the network ([5], [6]). This common adoption is of no coincidence because as [4] and [7] analytically show that for the BA model and its generalizations, the nature of the resulting degree distribution depends on the nature of preferential attachment.

Our motivation for conducting this study was observing the fact that although a number of studies form the theoretical relation between the nature of preferential attachment and power-law degree distribution (i.e. being scale-free), the empirical validations that confirm such a tendency is really present in real-world complex networks are relatively low in number. However, there are some studies which provide consistent results showing that there is indeed a preferential attachment phenomenon in some certain complex networks ([8], [9], [10], [11], and [12]). Carrying out such a study is not trivial because for a through analysis of the preferential attachment, the network growth data should contain the exact timestamps of link and node addition events. Any problem with the "quality" of the data in that sense will cause degradation in the reliability of the results. In [8], for instance, the time resolution of the data is one year, which means that the relative ordering of the link additions during each year is unknown. In [11], on the other hand, the full relative ordering of the link additions is known but the actual dates are missing and this fact results in the assumption that all of

the link additions are uniformly distributed in the time line. A detailed discussion for these two and the other studies is made in Section 2.2.

In this study, we tried to measure and see if there is a preferential attachment tendency in a complex network which comes from a slightly different domain than the networks analyzed in the previous studies and has some features missing in other ones (i.e. more precise time resolution). While doing so, we adopted new methodological approaches to provide measures that are more reliable. The main contribution of the new methods for preferential attachment measurement is that they can be classified as time independent measures in contrast to the presently used time dependent measures. Time dependency means that the measures are specific to a certain time interval along the network's lifetime and it is impossible to integrate two different measures obtained from two different time intervals. In contrast, time independent measures provide values that are independent of the specific time interval on which they were calculated and this enables us to integrate different measures obtained in different times. The most obvious application for such measures is carrying out the analyses on very short consecutive time intervals and then integrating the results altogether.

2 Theory and Previous Studies

2.1 Theory

Directed Graph (G): A directed graph G is defined as $G = (V, E)$, where V and E are the set of vertices (nodes) and the set of edges (links) of G, respectively. For two vertices $v_{source}, v_{target} \in V$, an edge $e \in E$ is defined as $e = (v_{source}, v_{target})$.

In-Degree, Out-Degree, Total Degree ($k_i^{\downarrow}$, $k_i^{\uparrow}$, $k_i^{\updownarrow}$): For a vertex v_i, out-degree is the number of all edges leaving v_i, in-degree is the number of all edges pointing to v_i, and total degree is the sum of out-degree and in-degree of v_i. These are represented by $k_i^{\uparrow}$, $k_i^{\downarrow}$, and $k_i^{\updownarrow}$, respectively. In this paper, in-degree is called as degree, and represented by k_i, for simplicity.

Degree Distribution ($P(k)$): Degree distribution of a graph shows the probability of having a certain degree k for a randomly selected node of the graph.

The power-law degree distribution, which is a characteristic property of scale-free networks, is defined as $P(k) = ck^{-\lambda}$ where c is a normalization coefficient to ensure $\sum_k P(k) = 1$ and $\lambda > 0$ is called scaling factor of the distribution. In **Figure 1**, the power-law distribution is compared with a typical exponential ($P(k) = ce^{-mk}$) distribution.

The following temporal definitions are used to represent an evolving network in a formal way:

Graph History (Γ): A graph history Γ is defined as $\Gamma = (\Xi, \mathrm{E}, \tau^{\Xi}, \tau^{\mathrm{E}})$, where Ξ and E are the set of all vertices and the set of all edges created in Γ, respectively, τ_i^{Ξ} is the creation time of vertex ξ_i, and τ_i^{E} is the creation time of edge ε_i.

Recent Degree Increase ($rdi_{t,i}$): For a vertex ξ_i and duration Δt, recent degree increase at time t is the number of all edges pointing to ξ_i that are created in the time interval $[t-\Delta t, t)$, and it is represented by $rdi_{i,t}$.

Snapshot Graph (G_t): A snapshot graph G_t of a graph system $\Gamma = (\Xi, \mathrm{E}, \tau^{\Xi}, \tau^{\mathrm{E}})$ at time t is defined as $G_t = (V_t, E_t)$, where $V_t = \{\xi_i \in \Xi : \tau_i^{\Xi} < t\}$ and $E_t = \{\varepsilon_i \in \mathrm{E} : \tau_i^{\mathrm{E}} < t\}$. For a snapshot graph G_t, the in-degree, out-degree and total degree of a vertex $\xi_i \in V_t$ is represented by $k_{t,i}^{\downarrow}$, $k_{t,i}^{\uparrow}$, and $k_{t,i}^{\updownarrow}$ respectively. Similarly, in-degree is called as degree, and represented by $k_{t,i}$, for simplicity.

Age ($a_{t,i}$): In a graph system, the age of a vertex ξ_i at time t is defined as $a_{t,i} = t - \tau_i^{\Xi}$.

Preferential Attachment Measures ($l_{t,i}$, $l_t^{degree}(k')$, and $l_t^{age}(a')$): $l_{t,i}$ is the number of new links that vertex ξ_i is expected to gain in the time interval $[t, t+\Delta t)$. By using this measure, it is possible to construct preferential attachment tendency as functions of degree and age. $l_t^{degree}(k')$ gives the expected number of new links to be acquired for a random node with a specified degree. It is defined as:

$$l_t^{degree}(k') = \frac{\sum_{x \in X} l_{t,x}}{|X|} \quad (X = \{x \mid k_{t,x} = k'\}) \tag{1}$$

Similarly $l_t^{age}(a')$ is defined as:

$$l_t^{age}(a') = \frac{\sum_{y \in Y} l_{t,y}}{|Y|} \quad (Y = \{y \mid a_{t,y} = a'\}) \tag{2}$$

Although we cannot directly determine these tendencies analytically for a given data, it is possible to come up with empirical computations, which may serve as approximate measures. The empirically computed values are represented by $\hat{l}_{t,i}$, $\hat{l}_t^{\deg ree}(k')$, and $\hat{l}_t^{age}(a')$, correspondingly, and details of their computation is given in Section 3.2.

2.2 Relevant Studies

It is possible to present the relevant studies under two headings. First group is the modeling efforts for complex networks, and the second is the search for empirical evidence supporting or falsifying these models.

Models: In this study, we focused on two network growth models. The first model is the original BA model ([1]). According to BA model, the network starts from a small set of fully connected core nodes, and new links are formed only between new arriving nodes and the existing nodes. Whenever a new node is introduced to the graph, it forms a link with vertex v_i according to the following probability:

$$p(k_i) = \frac{k_i^{\gamma}}{\sum_j k_j^{\gamma}} \tag{3}$$

where the parameter γ is introduced to account for the generalizations of BA model ([3]), and in the original case $\gamma = 1$.

The original BA model assumes a linear dependency between the likeliness to acquire a new connection and the present number of connections. This linearity is of paramount importance as it is shown in [4] and [7] analytically and in [1] empirically that in the presence of sub-linear dependency (i.e. $\gamma < 1$) the degree distribution becomes a stretched exponential; and in the case of super-linear dependency (i.e. $\gamma > 1$), a "winner takes it all" situation is observed and one node acquires all new coming links resulting in a star-like topology. A prediction of BA model is a positive correlation between the age of a node ($a_{t,i}$) and its degree ([5]). The reason is intuitively obvious that the older a node is, the more time it had to acquire links, which results in a higher probability of getting new links. Another important point about the BA model is that it generates an undirected graph (i.e. all edges are symmetric).

The second model is the growth and deactivation model described in [5]. The growth of the network is described by directed links (unlike the BA model). An important concept introduced by this model is the active/inactive node difference. At a given time, there exist a constant number m of active nodes, and an active node always receives an incoming link from a newly arriving node. An inactive node, on the other hand, cannot receive any links. Whenever a new node is introduced to the system, it forms new links pointing to the current active nodes, it is tagged as an active node, and one of the active nodes is deactivated randomly so the number of active nodes is kept constant. The deactivation probability for each active node v_i is inversely proportional to its current number of incoming links:

$$P_i^{deactivation} \propto (k_i + a)^{-1} \tag{4}$$

where a is a constant bias. A strong simplification of the model is that once a node is deactivated there is no way for it to be re-activated and receive new links. There are a number of predictions associated with this model. First, it is analytically shown that the average increase in the degree of a node is linearly proportional to the present degree ($l_{t,i} \propto k_{t,i}$), which leads to the linear dependence $l_t^{degree}(k') \propto k'$. This dependence is not explicitly incorporated into the model but is an emergent property of degree-dependent deactivation dynamics ([5]). Secondly, in sharp contrast to the BA model, this model predicts a negative correlation between the age and probability of acquiring a new link. No matter how high degree a node has, as time passes the probability of being deactivated at some point increases. The dependence between the age and probability

of getting a new connection is shown to be compliant with power-law. We defined the term recent degree increase ($rdi_{t,i}$) in correspondence with this study. While the original study does not contain such a concept, we believe the recent degree increase of a node is in close relation with its status of being active or inactive. Active nodes tend to have high recent degree increases; inactive nodes tend to have low (actually zero).

Empirical Studies: In [8], preferential attachment in two different networks is investigated. The networks are the co-authorship networks in fields of mathematics and neuroscience between years 1991 and 1998. The authors use a measure very similar to $l_t^{degree}(k')$ but the normalization according to the number of nodes that have degree k' is not used and the measurements are carried on a fixed time interval of one year. The findings suggest that the relation between degree of a node and its expected rate of link acquisition is linear. However, as Barabási et al. indicated themselves, the data is not complete but only a recent portion of the all co-authorship network is available. Another problem is the time resolution of the network data. Only the years of collaborations are known so the relative ordering of the link formations during each year is missing. Methodologically, the time period of analysis (Δt) has to be kept relatively short because as the nodes continue to acquire new links, their actual degrees increase while the records used for analyzing the relation is fixed to the analysis of time t. Using longer periods may introduce unwanted bias in this sense, and using shorter periods create fluctuations in the results that are hard to eliminate.

The second study relevant to our work is [11], which also focuses on two networks of co-authorships, one in physics the other in biology and medicine. The data used in this study contains the entire relative ordering of node and link additions but the exact timestamps are missing. Therefore, it is assumed that the links are introduced uniformly in the timeline. This study confirms that the preferential attachment tendency is linearly dependent on the current number of links of the nodes. The methodology is very similar to the one in [8] but the normalization of the expected number of new links according to the number of nodes is employed so the measure calculated is virtually the same with $l_t^{degree}(k')$. This normalization is important because for small values of k', there may be many nodes with k' links, but for high values of k' there are only a few of them if not zero at all. Just summing these values without considering how many nodes contribute to it may introduce some bias in the calculation of tendency to acquire new links, favoring the lowly connected nodes because their number is supposed to be higher than the highly connected ones. However, this kind of normalization does not help to eliminate the problem associated with the length of the time period analyzed.

[10] follows the same methodology, which is adopted in [8] but extends the analyses to different networks from different origins, namely, science citation network, Internet, actor collaboration and scientific co-authorship. The time resolution of the analysis is one year for all of the networks. For the first two a linear dependence for preferential attachment is observed while for the latter two, the dependence is sub-linear.

Of the other two studies, [9] is of significance because it extends the empirical findings to a protein-protein interaction network that has a very diverse origin of the other networks considered so far. [12], on the other hand, employs similar techniques to [8] and [10] but uses a more complete dataset: The full citation network of Physical Review journals, spanning a period from 1893 until 2003. The time resolution of the

analysis is again one year. Both studies confirm a linear dependency between likeliness to acquire a new link and the current number of links.

3 Data and Methodology

3.1 Data

The network that is analyzed in this study is constructed by using the data crawled from "Ekşi Sözlük" (literal translation from Turkish is Sour Dictionary) web site. Ekşi Sözlük (which will be called Sözlük, shortly) is technically a collaborative hypertext dictionary in operation since 15 February 1999, and it is gained a wide popularity among the Internet users in Turkey ([13]).

Sözlük is a dictionary in which one can find explanations and definitions of almost any concept one can think of. In the Sözlük's jargon, a concept for which information can be found is called a "title" (literal translation of "başlık" from Turkish). Each individual definition, explanation, or information of any kind is called an "entry". There may be several entries placed under a title. What makes Sözlük different from any other plain text based dictionary is that it contains hyper-textual references to other titles. For instance, in an entry below the title "Stargazer", which is the name of a masterpiece song of a hard rock band ([14]), the name of the band "Rainbow" can be a hypertext. When clicked on, it may lead to the title "Rainbow", where entries about both the band and the literal definition of a rainbow can be found.

Sözlük is also a collaborative work of its registered users, who are called "susers" (stands for Sözlük USER), and are eligible to add new titles or enter new entries under existing titles.

The graph history $\Gamma_{Sözlük}$ is constructed by using the data crawled from Eksi Sözlük web site. The crawling process consisted of fetching each individual entry, parsing it for possible cross-references, and recording their timestamps. This operation took 168 computing hours and the resulting data is given in **Table 1**.

The graph history corresponding to Sözlük, $\Gamma_{Sözlük} = (\Xi, \mathrm{E}, \tau^{\Xi}, \tau^{\mathrm{E}})$, is constructed as follows: There is a one-to-one correspondence between the vertices of $\Gamma_{Sözlük}$ (i.e. elements of Ξ) and the titles of Sözlük. For an edge $\varepsilon_i = (\xi_{source}, \xi_{target})$, $\varepsilon_i \in \mathrm{E}$ if and only if there is at least one cross-reference from the title *source* to the title *target*. τ^{Ξ} and τ^{E} are constructed such that τ_i^{Ξ} of ξ_i is the timestamp of the first entry of title i, or the timestamp of the first cross-reference pointing to title i, whichever is earlier; τ_i^{E} of edge $\varepsilon_i = (\xi_{source}, \xi_{target})$ is the timestamp of the first entry that contains a cross-reference from title *source* to title *target*. From now on, the snapshot graphs of $\Gamma_{Sözlük}$ will be referred as G_t, for simplicity.

Unlike most of the other complex network examples, Sözlük data spans the whole lifetime of the network since the first day and first node. In addition, the time resolution of the data is very precise: One day for the first two years, and one minute for

subsequent years. Another important characteristic of Sözlük network is that it is not only growing but it is also evolving because of the moderation activities, yet we were able to crawl only the visible entries but not the removed ones.

3.2 Methodology

Our analysis of the data is composed of three parts. In the first part, we employ the methodology used in [11], and for time t we empirically calculate the preferential attachment measures of all nodes ($\hat{l}_{t,i}$) by calculating the current degrees ($k_{t,i}$), and the number of new links they acquired during time interval $[t, t+\Delta t)$ ($\Delta k_{t,i}$).

By the definition, $\Delta k_{t,i}$ is an empirical calculation of the preferential attachment measure $l_{t,i}$, thus $\hat{l}_{t,i} = \Delta k_{t,i}$. Accordingly, preferential attachment tendency as a function of degree is calculated as

$$\hat{l}_t^{\deg ree}(k') = \frac{\sum_{x \in X} \Delta k_{t,y}}{|X|} \quad (X=\{x \mid k_{t,x} = k'\}) \tag{5}$$

and preferential attachment tendency as a function of age is calculated as

$$\hat{l}_t^{age}(a') = \frac{\sum_{y \in Y} \Delta k_{t,y}}{|Y|} \quad (Y=\{y \mid a_{t,y} = a'\}) \tag{6}$$

To keep the bias as low as possible during these analysis, it is necessary to use a relatively short time period (i.e. Δt), which is one day for our data, and one year for [8] and [11]. This part of analysis provides us the time dependent measures.

In the second part, we introduce another measurement method that makes use of the whole temporal data, and provide a time independent measure for the preferential attachment tendency. This way, it is possible to tackle with the bias problem faced in [8] and [11] by using very short time periods (e.g. one week) to calculate the number of new links acquired and then summing them altogether for a longer period (e.g. one year). The major problem for coming up with a time independent measure is that, while having a certain number of links, say 100, for year 2001 may provide a relative advantage to acquire new links, in year 2005, having 100 links may not mean that much because all of the nodes are continuously acquiring new links. Previous studies limited themselves with relatively short periods (i.e. one year) to count the new number of new links, because it is assumed that the bias introduced during one year is negligible. What we have to do is to come up with new measures which aim to reflect the same tendency as $l_t^{degree}(k')$ and $l_t^{age}(k')$ and is independent of the absolute value of k' but its relative magnitude. This way, it will be possible to integrate the different preferential attachment measures of different snapshot graphs of Sözlük instead of focusing on only one.

The method we adopt is to normalize the measures linearly for each snapshot graph G_t, such that they fall in interval $[0,1]$. The normalization process for a measure χ is as follows:

$$\chi_{t,i}^{norm} = \frac{\chi_{t,i} - \chi_t^{\min}}{\chi_t^{\max} - \chi_t^{\min}} \quad (7)$$

where χ is can be substituted with k, rdi, a, and Δk. The corresponding symbols $k_t^{\min}$ ($k_t^{\max}$), $rdi_t^{\min}$ ($rdi_t^{\max}$), $a_t^{\min}$ ($a_t^{\max}$), and $\Delta k_t^{\min}$ ($\Delta k_t^{\max}$) are the smallest (highest) degree, recent degree increase, age, and the number of new links values of the nodes of graph, respectively. Substituting the new normalized values in Equation (**5**) and (**6**), we obtain the new functional representations of preferential attachment tendency:

$$\hat{l}_t^{\deg ree}(k') = \frac{\sum_{x \in X} (\Delta k_{t,x})^{norm}}{|X|} \quad (X=\{x \mid k_{t,x}^{norm} = k'\}) \quad (8)$$

$$\hat{l}_t^{age}(a') = \frac{\sum_{y \in Y} (\Delta k_{t,y})^{norm}}{|Y|} \quad (Y=\{y \mid a_{t,y}{}^{norm} = a'\}) \quad (9)$$

The time-independent measure of preferential attachment tendency as a function of degree and age is obtained by averaging the time-independent measures calculated for successive snapshot graphs of $\Gamma_{Sözlük}$:

$$\eta^{\deg ree}(k') = \frac{\sum_{t \in T} \hat{l}_t^{\deg ree}(k')}{|T|} \quad (T=\{t_0,\ t_0 + \Delta t,\ t_0 + 2\Delta t, ...,\ t_{end}\}) \quad (10)$$

$$\eta^{age}(a') = \frac{\sum_{t \in T} \hat{l}_t^{age}(a')}{|T|} \quad (T=\{t_0,\ t_0 + \Delta t,\ t_0 + 2\Delta t, ...,\ t_{end}\}) \quad (11)$$

where $t_0 > 0$ is a given time for starting the analysis and t_{end} is the timestamp of the last event represented in $\Gamma_{Sözlük}$.

The third part of our analysis aims to present an example application for the new time-independent measures. The results we obtain are used to compare and evaluate two network growth models (BA and growth and deactivation) to see which one provides a better explanation for the data. The methodological tools used are the correlation and partial correlation analysis. By correlation analysis, we calculate the pairwise Pearson correlation coefficients between $k_{t,i}^{norm}$, $\hat{l}_{t,i}^{norm}$, $rdi_{t,i}^{norm}$, $a_{t,i}^{norm}$ obtained for each node for each time step t; and carry out a significance test to see whether the observed correlations are indeed significant. By partial correlation analyses, we measure the correlation between normalized age $a_{t,i}^{norm}$, and the $\hat{l}_{t,i}^{norm}$ when the effect of current number of links $k_{t,i}^{norm}$ is removed. In this case, BA model predicts a zero correlation (given the degrees of two nodes are equal, the age of a node does not have an effect on the link acquisition), but growth and deactivation model predicts a negative correlation as discussed in Section 2.2. Also the partial correlation between $k_{t,i}^{norm}$ and $\hat{l}_{t,i}^{norm}$ is

calculated when the effect of the recent degree increase ($rdi_{t,i}^{norm}$) is removed. BA model predicts a positive partial correlation because according to the dynamics, even if recent degree increase values of two nodes are the same, the more connected one has greater chances to acquire new links. On the contrary, growth and deactivation model predicts a negative correlation because the nodes with low recent degrees are the inactive nodes, and they cannot receive new links anymore. Similarly, the partial correlation between $rdi_{t,i}^{norm}$ and $\hat{l}_{t,i}^{norm}$ is calculated when the effect of the node degree ($k_{t,i}^{norm}$) is removed. BA model does not predict a strong positive correlation, whereas growth and deactivation model predicts a strong positive correlation because the active nodes, which are more likely to receive new links, tend to have high recent degree increase values independent of their overall degree.

4 Results

Degree Distribution: The degree distribution of the snapshot graph $G_{t_{end}}$ is given in **Figure 2**. We tried also to fit a power-law model, and the scaling factor of the model is calculated as 2.63.

4.1 Preferential Attachment, Degree and Age Dependency

Unnormalized measures: For the calculation of $\hat{l}_t^{degree}(k')$ and $\hat{l}_t^{age}(a')$ values, Δt is set as one year, and t_0 is chosen as equal to $t_{end} - \Delta t$. Degree and age values were normalized in interval $[0,1]$ with a resolution of 0.01. The plot of the values is given in **Figure 3**. The robust best fitting lines for degree and age measures are also superimposed on the corresponding plots. In this configuration, the root mean square error of the robust best line fitting for degree data $RMSE_{degree}$ is 0.060 and for age data $RMSE_{age}$ is 0.227.

Normalized measures: For calculating the normalized preferential attachment measures $\eta^{degree}(k')$ and $\eta^{age}(a')$, Δt was set as one week, and the starting time of analysis, t_0, was set as equal to $t_{end} - T$, where T is four years. Thus, a snapshot graph for each week between the years 2003 and 2004 was constructed to calculate corresponding time independent measures $\hat{l}_{t,i}^{norm}$ and $a_{t,i}^{norm}$, which in turn provided the time independent measures $\eta^{\deg ree}(k')$ and $\eta^{age}(a')$. The plot of these values with respect to the normalized degree and age values are given in **Figure 4**. The robust best fitting lines for those values are also superimposed on the corresponding figures. Note that, fitting of the age data excludes values before 0.2 because of the "today effect", which will be discussed later. In this second configuration, the root mean square errors, $RMSE_{degree}^{norm}$ is 0.050 and $RMSE_{age}^{norm}$ is 0.192.

4.2 Interdependency Between Age, Degree and Recent-Degree

In Section 4.1, it is explained how the normalized measured are obtained. Shortly, for each interval of one week $[t+\Delta t)$, $k_{t,i}^{norm}$, $\hat{l}_{t,i}^{norm}$, $a_{t,i}^{norm}$, and $rdi_{t,i}^{norm}$ values are calculated for each node in the corresponding graph G_t. Grouping this quadruple for each individual node, a list was formed and a random sub-sampling was carried on this list. As a result, a subset of 2370 elements was selected. The pairwise Pearson correlation coefficients of the measures are given in **Table 2**.
Partial correlation coefficients as defined in Section 3.2, are given in Table 3.

5 Discussion

In accordance with the previous results, we were able to confirm a linear dependency between the preferential attachment tendency and the degree of the nodes. A linear model provides a better robust fit for the normalized measures than the unnormalized ones. ($RMSE_{degree}^{norm} = 0.050 \; < \; RMSE_{degree} = 0.060$). Assuming the actual relation is indeed linear, this fact suggests that the normalized method is better in capturing the linearity. However, as it can be seen visually in **Figure 4**, for larger degree values, the preferential attachment relation follows a non-linear form. The same effect is not observed for the unnormalized measures. It is an open question whether this observation is a side effect of the normalization or the preferential attachment really follows a non-linear relationship for large degree values. We were not able to answer it with the current findings.
Another interesting result was obtained by measuring the relation between the age and preferential attachment of a node. With the normalized measures, we observed that the relation between age and number of new links could be modeled by an exponential model ($l_{t,i} \propto e^{9.751 a_{t,i}}$). That is, the likeliness to acquire new links for a node increases, as the node gets older. Interestingly, for smaller age values the relation adopts a power law (scaling factor is 0.448) and follows an inverse relation: In the interval [0,0.2], the younger a node the more probable that it will receive a new link in the future. Our account for this phenomenon is the effect of a facility called “Today’s Titles” that lists the titles, which have received new entries during that day. Apart from manual search and following hyperlinks, this facility is the only way to see a group of titles and acts as a buffer in the sense once a new title is created it is immediately put in this list and becomes visible to the other susers for a brief time. As time passes the probability that, it will be removed from the list (because of new comers) increases. We call this effect as “today effect” and it serves as a good example for the benefits of using more precise measurement methods. While it is possible to visually observe the same affect for unnormalized measures, we were unable to fit a meaningful model (i.e. neither exponential nor power-law) to the data in interval [0,0.2], which suggests that the unnormalized measures are not capable of reflecting dynamic preferences in order of weeks and days.
The pairwise correlation coefficients confirm the findings of previous analyses and show that the correlations observed between the measures are indeed statistically

significant. However, we should note that none of the correlations are not strong (but barely moderate), which indicates that the real dynamics of the Sözlük network depends also on other factors which were not addressed here.

The partial correlation analyses enabled us to test the different predictions of BA model and growth and deactivation model on the data.

The first finding is that when the effect of the current degree is removed no statistically significant correlation between the age and number of new links is observed ($r = -0.016,\ p > 0.450$). This finding is in accordance with the predictions of BA model. The age of a node does not seem to have an effect on acquiring new links (neither positive nor negative) independent of the current degree.

Secondly, we found out that when the effect of the recent degree increase is removed, the correlation between degree and number of new links decrease, but remain statistically significant ($r = 0.356,\ p < 0.01$). A similar but weaker effect is observed also in the partial correlation of recent degree increase and number of new links, controlled for node degree ($r = 0.132,\ p < 0.01$). Based on these results, we can claim that both recent degree increase and node degree has a moderate effect on the link acquisition, and this effect is independent to some degree because the remaining partial correlations are still significant.

Considering these findings, our conclusion is that both models have their own merits in predicting the growth of the network. In the short time scale, "Today's Titles" facility and "today effect" is explained by the growth and deactivation model (i.e. power-law relation between age and likeliness to acquire new links) whereas in the longer time scales, the independency of age and new link acquisition probability can be explained by the BA model.

6 Conclusion

In this study, we introduced a new complex network data, which spans the whole lifetime of the network and contains very precise timestamps of the node and edge addition events. Using this dataset, we focused on the empirical validation of the linear dependency of preferential attachment on node degree. We introduced a new time independent measurement method which enabled us to independently analyze very short intervals and then integrating them altogether to provide measures for longer periods of time and capturing the dynamics more properly. The results suggest that the time independent method is more successful in extracting the linear relation between degree and preferential attachment for small degrees. The ability to analyze very short time periods also revealed an interesting relation (i.e. today effect) between the likeliness to acquire new links and age of a node, which was impossible to observe by the existing methods. As an application, the normalized measures were used to compare and evaluate two different network growth models on the present data.

References

1. Barabási, A.-L., Albert, R., Emergence of Scaling in Random Networks, *Science*, Vol. 286, pp. 509-512, 1999.
2. Newman, M.E.J., The structure and function of complex networks, *SIAM Review,* Vol. 45, pp. 167-256, 2003.
3. Pennock, D.M., Flake, G.W., Lawrence, S., Glover, E.J., Giles, C.L., Winners don't take all: Characterizing the competition for links on the web, *National Academy of Sciences*, Vol. 99, Issue 8, pp. 5207–5211, 2002.
4. Krapivsky, P.L., Redner, S., Leyvraz, F., Connectivity of Growing Random Networks, *Physical Review Letters*, Vol. 85, No. 21, pp. 4629-4632, 2000.
5. Klemm, K., Eguiluz, V.M., Highly clustered scale-free networks, *Physical Review E,* Vol. 65, 036123, 2002.
6. Vazquez, A., Growing Networks with Local Rules: Preferential Attachment, Clustering Hierarchy and Degree Correlations, *Physical Review E*, Vol. 67, 056104, 2003.
7. Dorogovtsev, S.N., Mendes, J.F.F., Evolution of Reference Networks With Aging, *Physical Review E*, Vol. 62, pp. 1842-1845, 2000.
8. Barabási, A.-L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T., Evolution of the Social Network of Scientific Collaborations, *Physica A*, Vol. 311, pp. 590-614, 2002.
9. Eisenberg, E., Levanon, E.Y., Preferential Attachment in the Protein Network Evolution, *Physical Review Letters*, Vol. 91, 138701, 2003.
10. Jeong, H., Neda, Z., Barabási, A.-L., Measuring Preferential Attachment for Evolving Networks, *Europhysics Letters*, Vol. 61, pp. 567-572, 2003.
11. Newman, M.E.J., Clustering and Preferential Attachment in Growing Networks, *Physical Review E*, Vol. 64, No. 025102, 2001.
12. Redner, S., Citation statistics from more than a century of physical review, *physics/0407137*, 2004.
13. Eksi Sözlük, http://www.eksisozluk.com.
14. Rainbow, Rising, Polydor, 1976.

Table 1. Characteristics of Sözlük Data

Node Count	Link Count	Title Count	Entry Count	First Entry Date	Last Entry Date	Suser Count
1,543,328	5,016,632	1,543,328	2,414,296	15.02.1999	10.10.2004	29,712

Table 2. Pearson correlation coefficients of the measures, statistically significant correlations (i.e. $p < 0.01$) are printed in bold.

	$k_{t,i}^{norm}$	$\hat{l}_{t,i}^{norm}$	$a_{t,i}^{norm}$	$rdi_{t,i}^{norm}$
$k_{t,i}^{norm}$	-	-	-	-
$\hat{l}_{t,i}^{norm}$	**0.427**	-	-	-
$a_{t,i}^{norm}$	**0.366**	**0.144**	-	-
$rdi_{t,i}^{norm}$	**0.406**	**0.283**	**0.088**	-

Table 3. Partial correlation coefficients of the measures, statistically significant correlations (i.e. $p < 0.05$) are printed in bold.

Correlated Measures	Control measures	Partial Correlation Coefficients
$a_{t,i}^{norm}$, $\hat{l}_{t,i}^{norm}$	$k_{t,i}^{norm}$	-0.015
$k_{t,i}^{norm}$, $\hat{l}_{t,i}^{norm}$	$rdi_{t,i}^{norm}$	**0.356**
$rdi_{t,i}^{norm}$, $\hat{l}_{t,i}^{norm}$	$k_{t,i}^{norm}$	**0.133**

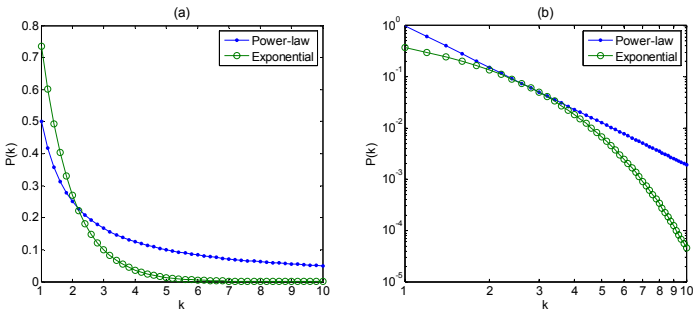


Figure 1. Power-law and exponential distributions: (a): Normal scale, (b): Log-log scale

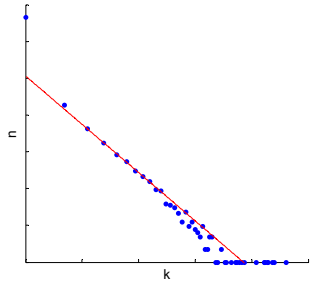


Figure 2. Log-log plot of the degree distribution of the resulting graph, and its best power-law fit

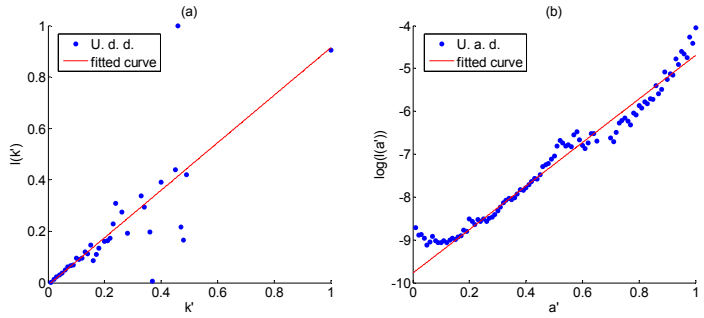


Figure 3. Number of new links as a function of (a): Degree, (b): Logarithm of age

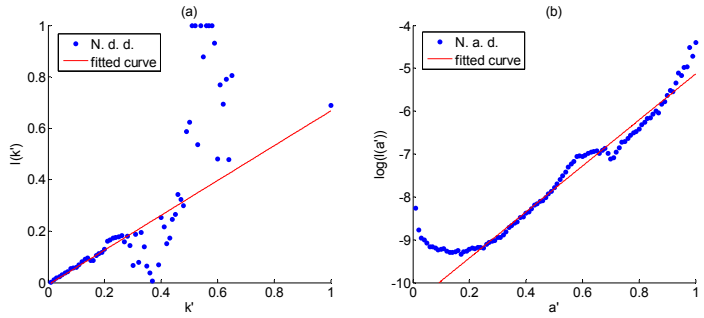


Figure 4. Number of new links as a normalized function of (a): Degree, (b): Logarithm of age

Medusa, a functional model of Internet substructure

Scott Kirkpatrick, Shai Carmi, Eran Shir
kirk@cs.huji.ac.il

We consider the Internet at the level of its sub-networks (called Autonomous Systems, or ASes). All previous studies have used the connection degree as the indicator variable to decompose the network into what one hopes will be nodes with distinct functions or roles. We consider instead a longer-ranged indicator of connectivity, obtained by k-pruning, which removes all sites with less than k neighbors until no such sites remain. Increasing k from 1 in steps of 1 separates any network into "k-shells," leaving at each stage a k-core, and defining a "k-crust" as the union of the k-shells lying outside of a particular core. The construction is unique, the k-core is maximal, and experiments show that the k-cores are indeed k-connected (a hypothesis which is proven for Erdos-Renyi graphs and plausible for dense scale-free graphs such the Internet IP-level and AS graphs. There are similarities and important differences with the "jellyfish" model introduced for the AS-graph by Faloutsis, hence our title for this model, coming from the Eastern Mediterranean. Its characteristics are a core which is the maximal non-vanishing k-core, a scale free region (the successive k-crusts) in which information flows steadily from the periphery towards the core, but also can propagate laterally for unlimited distances through peer connections, and a community of dependent nodes which project directly from the outside world into the core, without taking advantage of the scale-free region. We are in the process of exploring the extent to which this structure distinguishes different models of the Internet's formation, is a basis for projecting its evolution, and suggests new approaches to routing information in the Internet. This work was stimulated by the availability of Internet maps of unprecedented resolution from the DIMES and EVERGROW projects.

Atomic Selfish Routing in Networks: A Survey*

Spyros Kontogiannis[†,‡] Paul Spirakis[†]

May 8, 2005

Abstract

In this survey we present some recent advances in the *atomic congestion games* literature. Our main focus is on a special case of congestion games, called *network congestion games*, which is of particular interest for the networking community. The algorithmic questions that we are interested in have to do with the *existence* of pure Nash equilibria, the *efficiency* of their construction when they exist, as well as the *gap* of the best/worst (mixed in general) Nash equilibria from the social optima in such games, typically called the *Price of Anarchy* and the *Price of Stability* respectively.

1 Introduction

Consider a model where selfish individuals (henceforth called **players**) in a communication network having varying service demands compete for some shared resources. The quality of service provided by a resource decreases with its *congestion*, ie, the amount of demands of the players willing to be served by it. Each player may reveal its actual, unique choice of a subset of resources (called a *pure strategy*) that satisfies his service demand, or he may reveal a probability distribution for choosing (independently of other players' choices) one of the possible (satisfactory for him) subsets of resources (called a *mixed strategy*). The players determine their actual behavior based on other players' behaviors, but they do not cooperate. We are interested in situations where the players have reached some kind of stable state, ie, an equilibrium. The most popular notion of equilibrium in non-cooperative game theory is the *Nash equilibrium*: a "stable point" among the players, from which no player is willing to deviate unilaterally. In [23], the notion of the *coordination ratio* or *price of anarchy* was introduced as a means for measuring the performance degradation due to lack of players' coordination when sharing common goods. A more recent measure of performance is the *price of stability* [2], capturing the gap between the best possible Nash Equilibrium and the globally optimal solution. This measure is crucial for the network designer's perspective, who would like to propose (rather than let the players end up in) a Nash equilibrium (from which no player would like to defect unilaterally) that is as close to the optimum as possible.

A realistic scenario for the above model is when *unsplittable* traffic demands are routed selfishly in general networks with load-dependent edge delays. When the underlying network

*This work was partially supported by the EU within 6th Framework Programme under contract 001907 (DELIS).

† Research Academic Computer Technology Institute, Riga Feraiou 61, 26222 Patra, Greece. {kontog,spirakis}@cti.gr

‡ Computer Science Department, University of Ioannina, 45110 Ioannina, Greece. kontog@cs.uoi.gr

consists of two nodes and parallel links between them, there has been an extensive study on the existence and computability of equilibria, as well as on the price of anarchy. In this survey we study the recent advances in the more general case of arbitrary *congestion games*. When the players have identical traffic demands, the congestion game is indeed isomorphic to an *exact potential game* ([29], see also theorem 1 of this survey) and thus always possesses a *pure Nash equilibrium*, ie, an equilibrium where each player adopts a pure strategy. We shall see that varying demands of the players crucially affect the nature of these games, which are no longer isomorphic to exact potential games. We also present some results in a variant of congestion games, where the players' payoffs are not resource-dependent (as is typically the case in congestion games) but *player-specific*.

1.1 Roadmap

In section 2 we formally define the congestion games and their variants considered in this survey. We also give some game-theoretic definitions. In section 3 we present most of the related work in the literature, before presenting in detail some of the most significant advances in the area. In section 4 we present some of the most important results concerning unweighted congestion games and their connection to the *potential games*. In section 5 we study some complexity issues of unweighted congestion games. In section 6 we presents Milchtaich's extension of congestion games to allow *player-specific* payoffs, whereas in section 5.2 we study some existence and computability issues of PNE in weighted congestion games. Finally, in section 7 we study the price of anarchy of weighted congestion games. We close this survey with some concluding remarks and unrsesolved questions.

2 The Model

Consider having a set of resources E in a system. For each $e \in E$, let $d_e(\cdot)$ be the **delay** per player that requests her service, as a function of the total usage (ie, the *congestion*) of this resource by all the players. Each such function is considered to be *non-decreasing* in the total usage of the corresponding resource. Each resource may be represented by a pair of points: an entry point to the resource and an exit point from it. So, we represent each resource by an arc from its entry point to its exit point and we associate with this arc the **charging cost** (eg, the delay as a function of the load of this resource) that each player has to pay if he is served by this resource. The entry/exit points of the resources need not be unique; they may coincide in order to express the possibility of offering a *joint service* to players, that consists of a sequence of resources. We denote by V the set of all entry/exit points of the resources in the system. Any nonempty collection of resources corresponding to a directed path in $G \equiv (V, E)$ comprises an **action** in the system.

Let $N \equiv [n]$[1] be the set of players, each willing to adopt some action in the system. $\forall i \in N$, let w_i denote player i's **traffic demand** (eg, the flow rate from a source node to a destination node), while $\mathcal{P}^i \equiv \{a^i_1, \ldots, a^i_{m_i}\} \subseteq 2^E \setminus \emptyset$ (for some $m_i \geqslant 2$) is the collection of actions, any of which would satisfy player i (eg, alternative routes from a source to a destination node, if G represents a communication network). The collection $\mathcal{P}^i$ is called the *action set* of player i and each of its elements contains at least one resource. Any $n-$tuple $\varpi \in \mathcal{P} \equiv \times_{i=1}^{n} \mathcal{P}^i$ is a **pure strategies profile**, or a **configuration** of the players. Any

[1] $\forall k \in \mathbb{N},\ [k] \equiv \{1, 2, \ldots, k\}$.

real vector $\mathbf{p} = (\mathbf{p^1}, \mathbf{p^2}, \ldots, \mathbf{p^n})$ s.t. $\forall i \in N,\ \mathbf{p^i} \in \Delta(\mathcal{P}^i) \equiv \{\mathbf{z} \in [0,1]^{m_i} : \sum_{k=1}^{m_i} z_k = 1\}$ is a probability distribution over the set of allowable actions for player i, is called a **mixed strategies profile** for the n players.

A **congestion model** $\left((\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\right)$ typically deals with players of identical demands, and thus the resource delay functions depend only on the *number* of players adopting each action ([31, 29, 12]). In the more general case, ie, a **weighted congestion model** is the tuple $((w_i)_{i\in N}, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E})$. That is, we allow the players to have different (but fixed) demands for service (denoted by their weights) from the whole system, and thus affect the resource delay functions in a different way, depending on their own weights. The *weighted congestion game* $\Gamma \equiv \left(N, E, (w_i)_{i\in N}, E, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\right)$ associated with this model, is the game in strategic form with the set of players N and players' demands $(w_i)_{i\in N}$, the set of shared resources E, the action sets $(\mathcal{P}^i)_{i\in N}$ and players' cost functions $(\lambda^i_{\varpi^i})_{i\in N, \varpi^i\in\mathcal{P}^i}$ defined as follows: For any configuration $\varpi \in \mathcal{P}$ and $\forall e \in E$, let $\Lambda_e(\varpi) = \{i \in N : e \in \varpi^i\}$ be the set of players wishing to exploit resource e according to ϖ (called the **view** of resource e wrt configuration ϖ). We also denote by $x_e(\varpi) \equiv |\Lambda_e(\varpi)|$ the *number* of players using resource e wrt ϖ, whereas $\theta_e(\varpi) \equiv \sum_{i\in\Lambda_e(\varpi)} w_i$ is the **load** of e wrt to ϖ. The **cost** $\lambda^i(\varpi)$ **of player** i **for adopting strategy** $\varpi^i \in \mathcal{P}^i$ in a given configuration ϖ is equal to the cumulative delay $\lambda_{\varpi^i}(\varpi)$ of all the resources comprising this action:

$$\lambda^i(\varpi) = \lambda_{\varpi^i}(\varpi) = \sum_{e\in\varpi^i} d_e(\theta_e(\varpi))\,. \tag{1}$$

On the other hand, for a mixed strategies profile $\mathbf{p}$, the **(expected) cost of player** i **for adopting strategy** $\varpi^i \in \mathcal{P}^i$ wrt $\mathbf{p}$ is

$$\lambda^i_{\varpi^i}(\mathbf{p}) = \sum_{\varpi^{-i}\in\mathcal{P}^{-i}} P(\mathbf{p}^{-i}, \varpi^{-i}) \cdot \sum_{e\in\varpi^i} d_e\left(\theta_e(\varpi^{-i} \oplus \varpi^i)\right) \tag{2}$$

where, $\varpi^{-i} \in \mathcal{P}^{-i} \equiv \times_{j\neq i}\mathcal{P}^j$ is a configuration of all the players except for i, $\mathbf{p}^{-i} \in \times_{j\neq i}\Delta(\mathcal{P}^j)$ is the mixed strategies profile of all players except for i, $\varpi^{-i} \oplus a$ is the new configuration with i definitely choosing the action $a \in \mathcal{P}^i$, and $P(\mathbf{p}^{-i}, \varpi^{-i}) \equiv \prod_{j\neq i} p^j_{\varpi^j}$ is the occurrence probability of ϖ^{-i} according to $\mathbf{p^{-i}}$.

Remark: We abuse notation a little bit and consider the player costs $\lambda^i_{\varpi^i}$ as functions whose exact definition depends on the other players' strategies: In the general case of a mixed strategies profile $\mathbf{p}$, equation (2) is valid and expresses the expected cost of player i wrt $\mathbf{p}$, conditioned on the event that i chooses path ϖ^i. If the other players adopt a pure strategies profile ϖ^{-i}, we get the special form of equation (1) that expresses the exact cost of player i choosing action ϖ^i.

Remark: Concerning the players' private cost functions, instead of charging them for the *sum* of the expected costs of the resources that each of them chooses to use (call it the SUM-COST objective), we could also consider the maximum expected cost over all the resources in the strategy that each player adopts (call it the MAX-COST objective). This is also a valid objective, especially in scenarios dealing with bandwidth allocation in networks. Nevertheless, in the present survey we focus our interest on the SUM-COST objective, unless stated explicitly that we use some other objective.

A congestion game in which all players are indistinguishable (ie, they have the traffic demands and the same action set), is called **symmetric**. When each player's action set $\mathcal{P}^i$ consists of sets of resources that comprise (simple) paths between a unique origin-destination pair of nodes (s_i, t_i) in (V, E), we refer to a **(multicommodity) network congestion game**. If additionally all origin-destination pairs of the players coincide with a unique pair (s, t) we have a **single commodity network congestion game** and then all players share exactly the same action set. Observe that in general a single-commodity network congestion game is not necessarily symmetric because the players may have different demands and thus their cost functions will also differ.

2.1 Dealing with Selfish behavior.

Fix an arbitrary (mixed in general) strategies profile $\mathbf{p}$ for a congestion game $\big((w_i)_{i\in N}, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\big)$. We say that $\mathbf{p}$ is a **Nash Equilibrium (NE)** if and only if

$$\forall i \in N,\ \forall \alpha, \beta \in \mathcal{P}^i,\quad p^i_\alpha > 0 \Rightarrow \lambda^i_\alpha(\mathbf{p}) \leqslant \lambda^i_\beta(\mathbf{p})\,.$$

A configuration $\varpi \in \mathcal{P}$ is a **Pure Nash Equilibrium (PNE)** if and only if

$$\forall i \in N, \forall \alpha \in \mathcal{P}^i,\quad \lambda^i(\varpi) = \lambda_{\varpi^i}(\varpi) \leqslant \lambda_\alpha(\varpi^{-i} \oplus \alpha) = \lambda^i(\varpi^{-i} \oplus \alpha)\,.$$

The **social cost** $\mathrm{SC}(\mathbf{p})$ in this congestion game is

$$\mathrm{SC}(\mathbf{p}) = \sum_{\varpi\in\Pi} P(\mathbf{p}, \varpi) \cdot \max_{i\in N}\{\lambda_{\varpi^i}(\varpi)\} \tag{3}$$

where $P(\mathbf{p}, \varpi) \equiv \prod_{i=1}^n p^i_{\varpi^i}$ is the probability of configuration ϖ occurring, wrt the mixed strategies profile $\mathbf{p}$. The **social optimum** of this game is defined as

$$\mathrm{OPT} = \min_{\varpi\in\Pi} \left\{ \max_{i\in N} [\lambda_{\varpi^i}(\varpi)] \right\} \tag{4}$$

The **price of anarchy** for this game is then defined as

$$\mathcal{R} = \max_{\mathbf{p} \text{ is a NE}} \left\{ \frac{\mathrm{SC}(\mathbf{p})}{\mathrm{OPT}} \right\} \tag{5}$$

2.2 Potential Games.

Fix an arbitrary game in strategic form $\Gamma = \big(N, (\mathcal{P}^i)_{i\in N}, (U^i)_{i\in N}\big)$ and some vector $\mathbf{b} \in \mathbb{R}^n_{>0}$. A function $\Phi : \mathcal{P} \to \mathbb{R}$ is called:

- an **ordinal potential** for Γ, if $\forall \varpi \in \mathcal{P},\ \forall i \in N, \forall \alpha \in \mathcal{P}^i,\ \mathrm{sign}\left[\lambda^i(\varpi) - \lambda^i(\varpi^{-i} \oplus \alpha)\right] = \mathrm{sign}\left[\Phi(\varpi) - \Phi(\varpi^{-i} \oplus \alpha)\right]$,
- a **b-potential** for Γ, if $\forall \varpi \in \mathcal{P},\ \forall i \in N, \forall \alpha \in \mathcal{P}^i,\ \lambda^i(\varpi) - \lambda^i(\varpi^{-i} \oplus \alpha) = b_i \cdot [\Phi(\varpi) - \Phi(\varpi^{-i} \oplus \alpha)]$, or
- an **exact potential** for Γ, if it is a $\mathbf{1}-$potential for Γ.

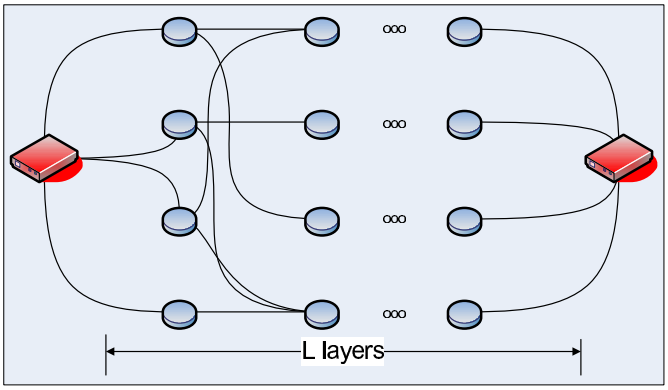


Figure 1: An example of a layered network.

2.3 Configuration Paths and Discrete Dynamics Graph.

For a congestion game $\Gamma = \left((w_i)_{i\in N}, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\right)$, a **path** in $\mathcal{P}$ is a sequence of configurations $\gamma = (\varpi(0), \varpi(1), \ldots, \varpi(k))$ s.t. $\forall j \in [k],\ \varpi(j) = (\varpi(j-1))^{-i} \oplus \pi_i$, for some $i \in N$ and $\pi_i \in \mathcal{P}^i$. γ is a **closed path** if $\varpi(0) = \varpi(k)$. It is a **simple path** if no configuration is contained in it more than once. γ is an **improvement path** wrt Γ, if $\forall j \in [k],\ \lambda^{i_j}(\varpi(j)) < \lambda^{i_j}(\varpi(j-1))$ where i_j is the unique player differing in its strategy between $\varpi(j)$ and $\varpi(j-1)$. That is, the unique defector of the j^{th} move in γ is actually willing to make this move because it improves its own cost. The **Nash Dynamics Graph** of Γ is a directed graph whose vertices are configurations and there is an arc from a configuration ϖ to a configuration $\varpi^{-i} \oplus \alpha$ for some $\alpha \in \mathcal{P}^i$ if and only if $\lambda^i(\varpi) > \lambda^i(\varpi^{-i} \oplus \alpha)$. The set of best replies of a player i against a configuration $\varpi^{-i} \in \mathcal{P}^{-i}$ is defined as $BR_i(\varpi^{-i}) = \arg\max_{\alpha\in\mathcal{P}^i}\{\lambda^i(\varpi^{-i}\oplus\alpha)\}$. Similarly, the set of best replies against a mixed profile $\mathbf{p^{-i}}$ is $BR_i(\mathbf{p^{-i}}) = \arg\max_{\alpha\in\mathcal{P}^i}\{\lambda^i_\alpha(\mathbf{p^{-i}}\oplus\alpha)\}$. A path γ is a **best-reply improvement path** if each defector jumps to a best-reply pure strategy. The **Best Response Dynamics Graph** is a directed graph whose vertices are configurations and there is an arc from a configuration ϖ to a configuration $\varpi^{-i} \oplus \alpha$ for some $\alpha \in \mathcal{P}^i \setminus \{\varpi^i\}$ if and only if $\alpha \in BR_i(\varpi^{-i})$ and $\varpi^i \notin BR_i(\varpi^{-i})$.

A (finite) strategic game Γ possesses the **Finite Improvement Property (FIP)** if any improvement path of Γ has finite length. Γ possesses the **Finite Best Reply Property (FBRP)** if every best-reply improvement path is of finite length.

2.4 Isomorphism of Strategic Games.

Two games in strategic form $\Gamma = \left(N, (\mathcal{P}^i)_{i\in N}, (U^i)_{i\in N}\right)$ and $\tilde{\Gamma} = \left(N, (\tilde{\mathcal{P}}^i)_{i\in N}, (\tilde{U}^i)_{i\in N}\right)$ are called **isomorphic** if there exist bijections $g : \times_{i\in N}\mathcal{P}^i \mapsto \times_{i\in N}\tilde{\mathcal{P}}_i$ and $\tilde{g} : \times_{i\in N}\tilde{\mathcal{P}}_i \mapsto \times_{i\in N}\mathcal{P}^i$ s.t. $\forall\varpi \in \times_{i\in N}\mathcal{P}^i, \forall i \in N,\ U^i(\varpi) = \tilde{U}^i(g(\varpi))$ and $\forall\tilde{\varpi} \in \times_{i\in N}\tilde{\mathcal{P}}_i, \forall i \in N,\ \tilde{U}^i(\tilde{\varpi}) = U^i(\tilde{g}(\tilde{\varpi}))$.

2.5 Layered Networks.

We consider a special family of networks whose behavior wrt the price of anarchy, as we shall see, is asymptotically equivalent to that of the parallel links model of [23] (which is actually a 1-layered network): Let $\ell \geqslant 1$ be an integer. A directed network $G = (V, E)$ with

a distinguished source - destination pair (s,t), $s,t \in V$, is an **ℓ-layered network** if every (simple) directed $s-t$ path has length exactly ℓ and each node lies on a directed $s-t$ path. In a layered network there are no directed cycles and all directed paths are simple. In the following, we always use $m = |E|$ to denote the number of edges in an ℓ-layered network $G = (V, E)$.

3 Related Work

3.1 Existence and tractability of PNE.

It is already known that the class of unweighted (atomic) congestion games (ie, players have the same demands and thus, the same affection on the resource delay functions) is guaranteed to have at least one PNE: actually, Rosenthal ([31]) proved that any potential game has at least one PNE and it is easy to write any unweighted congestion game as an exact potential game using Rosenthal's potential function[2] (eg, [12, Thm1]). In [12] it is proved that a PNE for any unweighted single-commodity network congestion game[3] (no matter what resource delay functions are considered, so long as they are non-decreasing with loads) can be constructed in polynomial time, by computing the optimum of Rosenthal's potential function, through a nice reduction to min-cost flow. On the other hand, it is shown that even for a symmetric congestion game or an unweighted multicommodity network congestion game, it is PLS-complete to find a PNE (though it certainly exists).

The special case of single-commodity, parallel-edges network congestion game where the resources are considered to behave as parallel machines, has been extensively studied in recent literature. In [14] it was shown that for the case of players with varying demands and uniformly related parallel machines, there is always a PNE which can be constructed in polynomial time. It was also shown that it is NP-hard to construct the best or the worst PNE. In [17] it was proved that the fully mixed NE (FMNE), introduced and thoroughly studied in [27], is worse than any PNE, and any NE is at most $(6+\varepsilon)$ times worse than the FMNE, for varying players and identical parallel machines. In [26] it was shown that the FMNE is the worst possible for the case of two related machines and tasks of the same size. In [25] it was proved that the FMNE is the worst possible when the global objective is the sum of squares of loads.

[13] studies the problem of constructing a PNE from any initial configuration, of social cost at most equal to that of the initial configuration. This immediately implies the existence of a PTAS for computing a PNE of minimum social cost: first compute a configuration of social cost at most $(1+\varepsilon)$ times the social optimum ([18]), and consequently transform it into a PNE of at most the same social cost. In [11] it is also shown that even for the unrelated parallel machines case a PNE always exists, and a potential-based argument proves a convergence time (in case of integer demands) from arbitrary initial configuration to a PNE in time $\mathcal{O}\left(mW_{\text{tot}} + 4^{W_{\text{tot}}/m+w_{\max}}\right)$ where $W_{\text{tot}} = \sum_{i\in N} w_i$ and $w_{\max} = \max_{i\in N}\{w_i\}$.

[28] studies the problem of weighted parallel-edges network congestion games with player-specific costs: each allowable action of a player consists of a single resource and each player has its own private cost function for each resource. It is shown that: (1) weighted (parallel-edges network) congestion games involving only two players, or only two possible actions for all the

[2]For more details on Potential Games, see [29].

[3]Since [12] only considers unit-demand players, this is also a symmetric network congestion game.

players, or equal delay functions (and thus, equal weights), always possess a PNE; (2) even a single-commodity, 3-player, 3-actions, weighted (parallel-edges network) congestion game may not possess a PNE (using 3-wise linear delay functions).

3.2 Price of Anarchy in Congestion Games.

In the seminal paper [23] the notion of coordination ratio, or price of anarchy, was introduced as a means for measuring the performance degradation due to lack of players' coordination when sharing common resources. In this work it was proved that the price of anarchy is 3/2 for two related parallel machines, while for m machines and players of varying demands, $\mathcal{R} = \Omega\left(\frac{\log m}{\log\log m}\right)$ and $\mathcal{R} = \mathcal{O}\left(\sqrt{m \log m}\right)$. For m identical parallel machines, [27] proved that $\mathcal{R} = \Theta\left(\frac{\log m}{\log\log m}\right)$ for the FMNE, while for the case of m identical parallel machines and players of varying demands it was shown in [22] that $\mathcal{R} = \Theta\left(\frac{\log m}{\log\log m}\right)$. In [9] it was finally shown that $\mathcal{R} = \Theta\left(\frac{\log m}{\log\log\log m}\right)$ for the general case of related machines and players of varying demands. [8] presents a thorough study of the case of general, monotone delay functions on parallel machines, with emphasis on delay functions from queuing theory. Unlike the case of linear cost functions, they show that the price of anarchy for non-linear delay functions in general is far worse and often even unbounded.

In [32] the price of anarchy in a multicommodity network congestion game among infinitely many players, each of negligible demand, is studied. The social cost in this case is expressed by the total delay paid by the whole flow in the system. For linear resource delays, the price of anarchy is at most 4/3. For general, continuous, non-decreasing resource delay functions, the total delay of any Nash flow is at most equal to the total delay of an optimal flow for double flow demands. [33] proves that for this setting, it is actually the class of allowable latency functions and not the specific topology of a network that determines the price of anarchy.

4 Unweighted Congestion Games: [31, 29]

In this section we present some fundamental results connecting the classes of unweighted congestion games and (exact) potential games [31]. Since we refer to players of identical (say, unit) weights, the players' cost functions are

$$\lambda^i(\varpi) \equiv \sum_{\varepsilon \in \varpi^i} d_e(x_e(\varpi)),$$

where, $x_e(\varpi)$ indicates the *number* of players willing to use resource e wrt configuration $\varpi \in \mathcal{P}$. The following theorem proves the strong connection of congestion games with the exact potential games.

Theorem 1 ([31, 29]) *Every (unweighted) congestion game is an exact potential game.*

Proof: Fix an arbitrary (unweighted) congestion game $\Gamma = \left(N, E, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\right)$. For any configuration $\varpi \in \mathcal{P}$, consider the function $\Phi(\varpi) = \sum_{e \in \cup_{i\in N}\varpi^i} \sum_{k=1}^{x_e(\varpi)} d_e(k)$, which we shall call **Rosenthal's potential**. We can easily show that Φ is an *exact potential* for Γ: For

this, consider arbitrary configuration $\varpi \in \mathcal{P}$, an arbitrary player $i \in N$ and an alternative (pure) strategy $\alpha \in \mathcal{P}^i \setminus \{\varpi^i\}$ for this player. Let also $\hat{\varpi} = \varpi^{-i} \oplus \alpha$. Then,

$$\begin{aligned}
\Phi(\hat{\varpi}) - \Phi(\varpi) &= \sum_{e \in \cup_j \hat{\varpi}^j} \sum_{k=1}^{x_e(\hat{\varpi})} d_e(k) - \sum_{e \in \cup_j \varpi^j} \sum_{k=1}^{x_e(\varpi)} d_e(k) \\
&= \sum_{e \in \cup_i \hat{\varpi}^i \setminus \varpi^i} \left[\sum_{k=1}^{x_e(\varpi)+1} d_e(k) - \sum_{k=1}^{x_e(\varpi)} d_e(k) \right] + \sum_{e \in \cup_i \varpi^i \setminus \hat{\varpi}^i} \left[\sum_{k=1}^{x_e(\varpi)-1} d_e(k) - \sum_{k=1}^{x_e(\varpi)} d_e(k) \right] \\
&= \sum_{e \in \hat{\varpi}^i \setminus \varpi^i} d_e(x_e(\varpi)+1) - \sum_{e \in \varpi^i \setminus \hat{\varpi}^i} d_e(x_e(\varpi)) \\
&= \sum_{e \in \hat{\varpi}^i} d_e(x_e(\hat{\varpi})) - \sum_{e \in \varpi^i} d_e(x_e(\varpi)) = \lambda^i(\hat{\varpi}) - \lambda^i(\varpi)
\end{aligned}$$

where, we have exploited the fact that $\forall e \in E \setminus (\varpi^i \cup \hat{\varpi}^i)$ and $\forall e \in \varpi^i \bigcap \hat{\varpi}^i$ the load of each of these resources (ie, the number of players using them) remains the same. Additionally, $\forall e \in \hat{\varpi}^i \setminus \varpi^i,\ x_e(\hat{\varpi}) = x_e(\varpi) + 1$ and $\forall e \in \varpi^i \setminus \hat{\varpi}^i,\ x_e(\hat{\varpi}) = x_e(\varpi) - 1$. ■

Remark: The existence of a (not necessarily exact) potential for any game in strategic form directly implies the existence of a PNE for this game. The existence of an exact potential may help (as we shall see later) the *efficient* construction of a PNE, but this is not true in general.

More interestingly, Monderer and Shapley [29] proved that every (finte) potential game is isomorphic to an unweighted congestion game. The proof presented here is a new one provided by the authors of the survey. The main idea is based on the proof of Monderer and Shapley, yet it is much simpler and easier to follow.

Theorem 2 ([29]) *Every finite (exact) potential game is isomorphic to an unweighted congestion game.*

Proof: Consider an arbitrary (finite) strategic game $\Gamma = \left(N, (Y^i)_{i \in N}, (U^i)_{i \in N}\right)$ among $n = |N|$ players, that admits an exact potential $\Phi : \times_{i \in N} Y^i \mapsto \mathbb{R}$. Suppose that $\forall i \in N, Y^i = \{1, 2, \ldots, m_i\} \equiv [m_i]$ for some finite integer $m_i \geqslant 2$ (players having a unique allowable action can be safely removed from the game) and let $Y \equiv \times_{i \in N} Y^i$ be the actions space of the game. We want to construct a proper unweighted congestion game $C = \left(N, E, (\mathcal{P}^i)_{i \in N}, (d_e)_{e \in E}\right)$ and a bijection $\varpi : Y \mapsto \mathcal{P}$ (recall that $\mathcal{P} \equiv \times_{i \in N} \mathcal{P}^i$ is the action space of C) such that

$$\forall i \in N, \forall \mathbf{s} \in Y,\ \lambda^i(\varpi(\mathbf{s})) = U^i(\mathbf{s})\,. \tag{6}$$

First of all, observe that for any (unweighted) congestion game C, we can express the players' costs as follows:

$$\begin{aligned}
\forall i \in N, \forall \varpi \in \mathcal{P},\ \lambda^i(\varpi) &= \sum_{e \in \varpi^i} d_e(x_e(\varpi)) \\
&= \sum_{e \in \varpi^i \bigcap (E \setminus \cup_{j \neq i} \varpi^j)} d_e(1) + \sum_{e \in \cup_{k \neq i} \varpi^i \bigcap \varpi^k \bigcap (E \setminus \cup_{j \neq i,k} \varpi^j)} d_e(2) + (7) \\
&+ \cdots + \sum_{e \in \bigcap_{k \in N} \varpi^k} d_e(n)
\end{aligned}$$

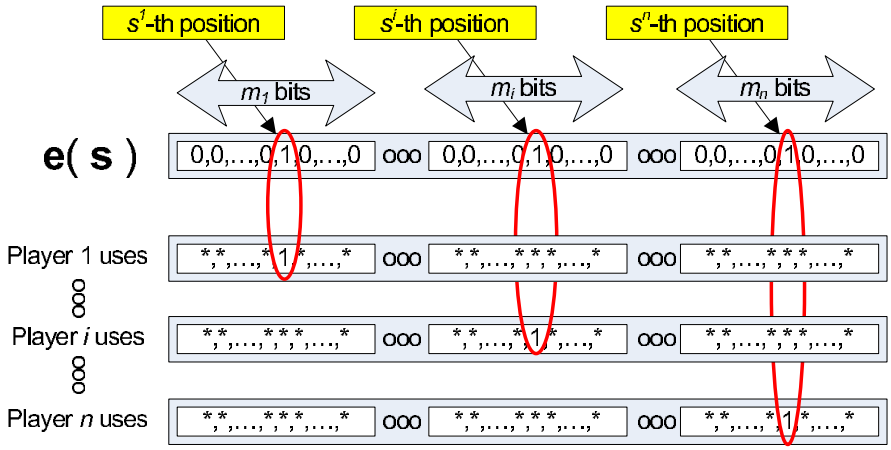


Figure 2: The set of resources representing the binary encodings of the configurations in the potential game.

We proceed by determining the set of shared resources $E = E_1 \cup E_2$ for our congestion game. We construct two sets of resources: The first set E_1 contains resources that determine exactly what the players in the potential game actually do. The second set E_2 contains resources that represent what all other players (except for the one considered by the resource) should not do. More specifically, fix an arbitrary configuration $\mathbf{s} \in Y$ of the players wrt the game Γ. We construct the following vector $\mathbf{e}(\mathbf{s}) \in E_1$ which is nothing more than the binary representation of this configuration (see figure 2):

$$\forall k \in N, \forall j \in [m_k],\ \mathrm{e}_j^k(\mathbf{s}) = \begin{cases} 1 & \text{if } s^k = j \\ 0, & \text{otherwise} \end{cases}$$

Fix now an arbitrary player $i \in N$. We define n more resources whose binary vectors represent what the other players (ie, other than player i) in the potential game *should not* do (see figure 3):

$$\forall k \in N, \forall j \in [m_k],\ \mathrm{e}_j^k(\mathbf{s}, i) = \begin{cases} 0 & \text{if } k \neq i \text{ and } s^k = j \\ 1, & \text{otherwise} \end{cases}$$

It is easy to see that $\mathbf{e}(\mathbf{s}, i)$ actually says that none of the players $k \neq i$ agrees with the profile $\mathbf{s}^{-\mathbf{i}} \in Y^{-i}$. So, we force this binary vector $\mathbf{e}(\mathbf{s}, i)$ to have definitely 0s exactly at the positions where $\mathbf{s}^{-\mathbf{i}}$ has 1s for any of the players except for player i, and we place 1s anywhere else. For $m = m_1 + m_2 + \cdots + m_n$, let $E_1 = \{\mathbf{e}(\mathbf{s}) \in \{0,1\}^m : \mathbf{s} \in Y\}$ and $E_2 = \{\mathbf{e}(\mathbf{s}, i) \in \{0,1\}^m : (i \in N) \wedge (\mathbf{s}^{-i} \in Y^{-i})\}$. The set of resources in our congestion game is then $E = E_1 \cup E_2$.

We now proceed with the definition of the action sets of the players wrt C. Indeed, $\forall i \in N,\ \mathcal{P}^i = \{\pi_1^i, \ldots, \pi_{m_i}^i\}$, where $\forall j \in [m_i],\ \pi_j^i = \{\mathbf{e} \in E : e_j^i = 1\}$. Now the bijective map that we assume is almost straightforward: $\forall \mathbf{s} = (s^i)_{i \in N} \in Y,\ \varpi(\mathbf{s}) = (\pi_{s^i}^i)_{i \in N}$.

Some crucial observations are the following: First, $\forall \mathbf{s} \in Y$, the resource $\mathbf{e}(\mathbf{s})$ is the only resource in E_1 that is used by *all* the players when the configuration $\varpi(\mathbf{s})$ is considered. For any other configuration $\mathbf{z} \in Y \setminus \{\mathbf{s}\}$, the resource $\mathbf{e}(\mathbf{s})$ is *not* used by at least one player, assuming $\varpi(\mathbf{z})$. Similarly, $\forall \mathbf{s} \in Y,\ \forall i \in N$, the resource $\mathbf{e}(\mathbf{s}, i)$ is the only resource in E_2 that is *exclusively used* by player i, assuming the configuration $\varpi(\mathbf{s})$.

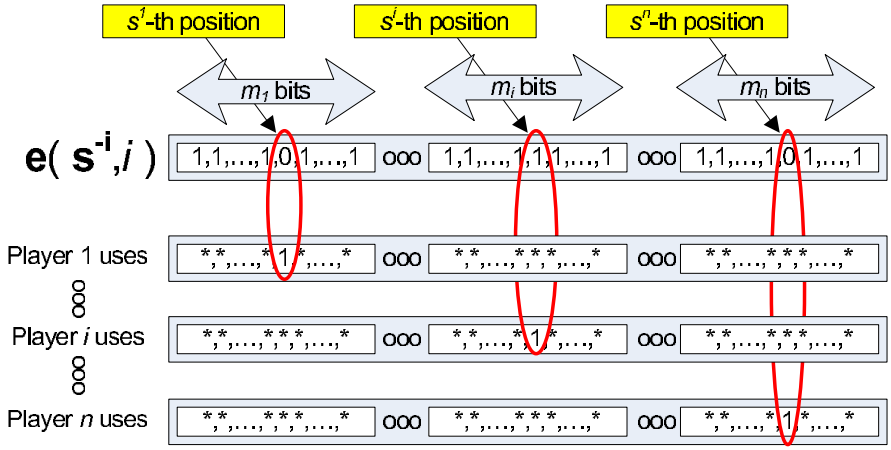


Figure 3: The set of resources representing the binary encodings of the forbidden configuration for the other players in the potential game.

Now it is very simple to set the resource delay functions d_e in such a way that we assure equality (6). Indeed, observe that $\forall i \in N, \forall \mathbf{s} \in Y, \forall \alpha \in Y^i \setminus \{s^i\}$,

$$\begin{aligned} U^i(\mathbf{s}) - U^i(\mathbf{s^{-i}} \oplus \alpha) &= \Phi(\mathbf{s}) - \Phi(\mathbf{s^{-i}} \oplus \alpha) \Leftrightarrow \\ Q^i(\mathbf{s^{-i}}) \equiv U^i(\mathbf{s}) - \Phi(\mathbf{s}) &= U^i(\mathbf{s^{-i}} \oplus \alpha) - \Phi(\mathbf{s^{-i}} \oplus \alpha)\,. \end{aligned}$$

That is, the quantity $Q^i(\mathbf{s^{-i}})$ is invariant of player i's strategy. Recall the form of player i's cost given in equality (7). The charging scheme of the resources that we adopt is as follows:

$$\forall \mathbf{e} \in E,\ d_e(k) = \begin{cases} Q^i(\mathbf{s^{-i}}), & \text{if } (k=1) \wedge (\mathbf{e} = \mathbf{e}(\mathbf{s}, i) \in E_2) \\ 0, & \text{if } (1 < k < n) \vee (k = 1 \wedge \mathbf{e} \in E_1) \\ \Phi(\mathbf{s}), & \text{if } (k = n) \wedge (\mathbf{e} = \mathbf{e}(\mathbf{s}) \in E_1) \end{cases}$$

Now it is easy to see that for any player $i \in N$ and any configuration $\mathbf{s} \in Y$, the only resource used exclusively by i wrt to $\mathbf{s}$ which has non-zero delay, is $\mathbf{e}(\mathbf{s}, i)$. Similarly, the only resource used by all the players that has non-zero delay is $\mathbf{e}(\mathbf{s})$. All other resources charge zero delays, no matter how many players use them. Thus,

$$\forall i \in N, \forall \mathbf{s} \in Y,\ (7) \Rightarrow \lambda^i(\varpi(\mathbf{s})) = d_{\mathbf{e}(\mathbf{s},i)}(1) + d_{\mathbf{e}(\mathbf{s})}(n) = Q^i(\mathbf{s^{-i}}) + \Phi(\mathbf{s}) = U^i(\mathbf{s})\,.$$

■

Remark: The size of the congestion game that we use to represent a potential game is at most $(|N| + 1)$ times larger than the size of the potential game. Since an unweighted congestion game is itself an exact potential game, this implies an essential equivalence of exact potential and unweighted congestion games.

5 Existence and Complexity of Constructing PNE

In the present section we deal with issues concerning the existence and complexity of constructing PNE in weighted congestion games. Our main references for this section are

[12, 24, 15]. We start with some complexity issues concerning the construction of PNE in unweighted congestion games (in which a PNE always exists) and consequently we study existence and complexity issues in weighted congestion games in general.

Fix an arbitrary (weighted in general) congestion game $\left(N, E, (\mathcal{P}^i)_{i\in N}, (w_i)_{i\in N}, (d_e)_{e\in E}\right)$ where the w_i's denote the (positive) weights of the players.

A crucial class of problems containing the weighted network congestion games is PLS [21] (stands for *Polynomial Local Search*). This is the subclass of total functions in NP that are guaranteed to have a solution because of the fact that *"every finite directed acyclic graph has a sink."* A problem Π in PLS is given by

(a) a set of instances $I = \Sigma^*$;

(b) $\forall x \in I$, a set of *feasible solutions* $F_x \subseteq \Sigma^{poly(|x|)}$;

(c) a *polynomial oracle* c which, given $x \in I$, $s \in \Sigma^{poly(|x|)}$ it determines whether $s \in F_x$ and if so, it computes an integer function $c(x, s)$ (considered to be the *payoff* of s);

(d) $\forall x \in I, s \in F_x$, a *neighborhood* of s in F_x, $N_x(s) \subseteq F_x$, and a polynomial function g which either returns a solution $s' \in N_x(s)$ s.t. $c(x, s')$ is *preferable* to $c(x, s)$ (eg, $c(x, s') < c(x, s)$ for a minimization problem), or returns **no** if no such solution exists in the neighborhood of s in F_x.

An instance of Π is then as follows: Given any $x \in I$, find a local optimum in F_x, ie, a solution $s \in F_x$ s.t. $g(s) = \mathbf{no}$.

The problem of constructing a PNE for a weighted congestion game is in PLS, in the following cases: (1) for any unweighted congestion game, since it is an exact potential game (see theorem 1), and (2) for any weighted network congestion game with linear resource delays, which admits (as we shall prove in theorem 6) a $\mathbf{b}-$potential with $b_i = \frac{1}{2w_i}, \forall i \in N$, and thus finding PNE is equivalent to finding local optima for the optimization problem with state space the action space of the game and objective the potential of the game. On the other hand, this does not imply a polynomial-time algorithm for constructing a PNE, since (as we shall see more clearly in the weighted case) the improvements in the potential can be very small and too many. Additionally, although problems in PLS admit a PTAS, this does not imply also a PTAS for finding ε-approximate PNE (approximation of the potential does not imply also approximation of each player's payoff).

5.1 Efficient Construction of PNE in Unweighted Congestion Games

In this subsection we shall prove that for unweighted single commodity network congestion games a PNE can be constructed in polynomial time. On the other hand, even for multicommodity network congestion games it is PLS complete to construct a PNE. The main source of this subsection is the work of Fabrikant, Papadimitriou and Talwar [12].

Theorem 3 ([12]) *There is a polynomial time algorithm for finding a PNE in any unweighted single-commodity netowrk congestion game.*

Proof: Fix an arbitrary unweighted, single-commodity network congestion game $\Gamma = (N, E, \mathcal{P}_{s-t}, (d_e)_{e\in E})$ and let $G = (V, E)$ be the underlying network. Recall that this game admits Rosenthal's exact potential $\Phi(\varpi) = \sum_{e\in E} \sum_{k=1}^{x_e(\varpi)} d_e(k)$. The algorithm computes the

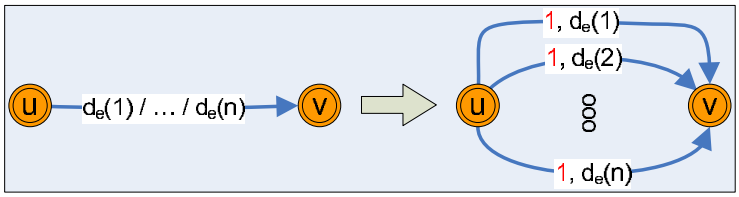


Figure 4: The reduction to a min-cost flow problem [12].

optimum value of $\Phi : \mathcal{P} \mapsto \mathbb{R}$ in the action space of the game. The corresponding configuration is thus a PNE.

The algorithm exploits a reduction to a min-cost flow problem. This reduction is done as follows: We construct a new network $G'(V', E')$ with the same set of vertices $V' = V$, while the set of arcs (ie, resources) is defined as follows (see also figure 4): $\forall e = (u, v) \in E$, we add n parallel arcs from u to v, $e'(1), \ldots, e'(n) \in E$, where these arcs have a *capacity* of 1 (ie, they allow at most one player to use each of them) and fixed delays (when used) $\forall k \in [d_e(n)], d'_{e'(k)} = d_e(k)$. Let $\mathbf{f}^* \in \mathbb{R}_{\geqslant 0}^{|E|}$ be a minimizer of the min-cost flow problem

$$
\begin{aligned}
\min \quad & \sum_{e' \in E'} d'_{e'} \cdot f_{e'} \\
s.t. \quad & \sum_{e'=(s,u) \in E'} f_{e'} - \sum_{e'=(u,s) \in E'} f_{e'} = n \\
& \sum_{e'=(u,t) \in E'} f_{e'} - \sum_{e'=(t,u) \in E'} f_{e'} = n \\
\forall v \in V \setminus \{s,t\}, \quad & \sum_{e'=(v,u) \in E'} f_{e'} = \sum_{e'=(u,v) \in E'} f_{e'} \\
\forall e' \in E', \quad & 0 \leqslant f_{e'} \leqslant u_{e'}
\end{aligned}
$$

This problem can be solved in polynomial time [1] (eg, in time $\mathcal{O}(m \log k(m + k \log k))$ using scaling algorithms) where $m = |E|$ and $k = |V|$ are the number of arcs (ie, resources) and the number of nodes in the underlying network respectively.

It is straightforward to see that any min-cost flow in G' is integral and it corresponds to a configuration in Γ that minimizes the potential of the game (which is then a PNE). ■

On the other hand, the following theorem proves that it is not that easy to construct a PNE, even in an unweighted multicommodity network congestion game. We give this theorem with a sketch of its proof:

Theorem 4 ([12]) *It is PLS-complete to find a PNE in unweighted congestion games of the following types: (i)General congestion games. (ii) Symmetric congestion games. (iii) Multicommodity network congestion games.*

Proof: We only give a sketch of the first two cases. The PLS completeness proof of case (iii) is rather complicated and therefore is not presented in this survey. The interested reader may find it in [12].

We first prove the PLS completeness of general congestion games and consequently we show that an arbitrary congestion game can be transformed into an equivalent symmetric

congestion game. In order to prove the PLS completeness of a congestion game, we shall use the following problems:

NOTALLEQUAL3SAT: Given a set N of $\{0,1\}-$variables and a collection C of clauses s.t. $\forall c \in C, |c| \leqslant 3$, is there an assignment of values to the variables so that no clause has all its literals assigned the same value?

POSNAE3FLIP: Given an instance (N, C) of NOTALLEQUAL3SAT with positive literals only and a weight function $w : C \mapsto \mathbb{R}$, find an assignment for the variables of N, s.t. the total weight of the unsatisfied clauses and the totally satisfied (ie, with all their literals set to 1) clauses cannot be decreased by a unilateral flip of the value of any variable.

It is known that POSNAE3FLIP is PLS complete [34].Given an instance of POSNAE3FLIP, we shall construct a congestion game $\Gamma = (N, E, (\mathcal{P}^v)_{v\in N}, (d_e)_{e\in E})$ as follows: The player set of the game is exactly the set of variables N. $\forall c \in C$, we construct two resources $e_c,\ e'_c \in E$ whose delay functions are $d_e(k) = w(c) \cdot \mathbb{I}_{\{k=3\}}$ and $d_{e'}(k) = w(c) \cdot \mathbb{I}_{\{k=3\}}$. That is, resource e_c (or e'_c) has delay $w(c)$ only when all the 3 players it contains actually use it. Each player $v \in N$ has exactly two allowable actions indicating the possible **true** or **false** values of the corresponding variable: $\mathcal{P}^v = \{\{e_c \in E : v \in c\}, \{e'_c \in E : v \in c\}\}$. Smaller clauses (ie, of two literals) are implemented similarly. Clearly, a flip of a variable corresponds to the change in the strategy of the corresponding player. The changes in the total weight due to a flip equal the changes in the cumulative delay over all the resources. Thus, any PNE of the congestion game Γ is a local optimum (ie, a solution) of the POSNAE3FLIP problem, and vice versa.

We now proceed to show that any unweighted congestion game can be transformed into a *symmetric* congestion game. Fix again an arbitrary congestion game $\Gamma = (N, E, (\mathcal{P}^v)_{v\in N}, (d_e)_{e\in E})$. We construct an equivalent symmetric congestion game $\hat{\Gamma} = \left(N, \hat{E}, \hat{\mathcal{P}}, (\hat{d}_e)_{e\in\hat{E}}\right)$ as follows: First of all we add to the set of resources n new distinct resources: $\hat{E} = E \cup \{e_i\}_{i\in N}$. The delays of these resources are: $\forall i \in N,\ \hat{d}_{e_i}(k) = M \cdot \mathbb{I}_{\{k\geqslant 2\}}$ for some sufficiently large constant M. The old resources maintain the same delay functions: $\forall e \in E, \forall k \geqslant 0,\ \hat{d}_e(k) = d_e(k)$. Now each player has the same action set $\hat{\mathcal{P}} = \times_{i\in N}\{a \cup \{e_i\} : a \in \mathcal{P}^i\}$. Observe that (by setting the constant M sufficiently large) in any PNE of $\hat{\Gamma}$ each of the distinct resources $\{e_i\}_{i\in N}$ is used by exactly one player (these resources act as if they have capacity 1 and there are only n of them). So, for any PNE in $\hat{\Gamma}$ we can easily get a PNE in Γ by simply dropping the unique new resource used by each of the players. This is done by identifying the "anonymous" players of $\hat{\Gamma}$ according to the unique resource they use, and match them with the corresponding players of Γ. ∎

5.2 Existence and construction of PNE in Weighted Congestion Games

In this section we deal with the existence and tractability of PNE in weighted network congestion games. First we show that it is not always the case that a PNE exists, even for a weighted single-commodity network congestion game with only linear and 2-wise linear (ie, the maximum of two linear functions) resource delays. Recall that, as discussed previously, any unweighted congestion game has a PNE, for any kind of non-decreasing delays, due to the existence of an exact potential for these games. This result was independently proved by [15] and [24], based on similar constructions. In this survey we present the version of [15] due to its clarity and simplicity.

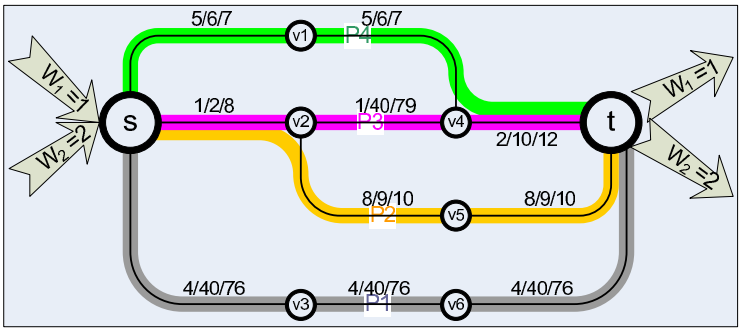


Figure 5: A weighted single-commodity network congestion game that may have no PNE. Consider two players with demands $w_1 = 1$ and $w_2 = 2$. The notation $a/b/c$ means that a load of 1 has delay a, a load of 2 has delay b and a load of 3 has delay c.

Lemma 5.1 ([15]) *There exist instances of weighted single-commodity network congestion games with resource delays being either linear or 2-wise linear functions of the loads, for which there is no PNE.*

Proof: We demonstrate this by the example shown in figure 5. In this example there are exactly two players of demands $w_1 = 1$ and $w_2 = 2$, from node s to node t. The possible paths that the two players may follow are labeled in the figure. The resource delay functions are indicated by the 3 possible values they may take given the two players. Observe now that this example has no PNE: there is a simple closed path $\gamma = \langle (P3, P2), (P3, P4), (P1, P4), (P1, P2), (P3, P2) \rangle$ of length 4 that is an improvement path (actually, each defecting player moves to its new best choice, so this is a best-reply improvement path) and additionally, any other configuration not belonging to γ is either one, or two best-reply moves away from some of these nodes. Therefore there is no sink in the Nash Dynamics Graph of the game and thus there exists no PNE. Observe that the delay functions are *not* player-specific in our example, as was the case in [28]. ■

Consequently we show that there may exist no exact potential function for a weighted single-commodity network congestion game, even when the resource delays are identical to their loads. The next argument shows that theorem 1 does not hold anymore even in this simplest case of weighted congestion games.

Lemma 5.2 ([15]) *There exist weighted single-commodity network congestion games which are not exact potential games, even when the resource delays are identical to their loads.*

Proof: Let $\Gamma = \left((w_i)_{i\in N}, (\mathcal{P}^i)_{i\in N}, (d_e)_{e\in E}\right)$ denote a weighted single commodity network congestion game with $d_e(x) = x$, $\forall e \in E$. Recall the definition of players' costs for a configuration (eq. (1)). Let's now define the quantity $I(\gamma, \lambda) = \sum_{k=1}^{r} \left[\lambda^{i_k}(\varpi(k)) - \lambda^{i_k}(\varpi(k-1))\right]$, where i_k is the unique player in which the configurations $\varpi(k)$ and $\varpi(k-1)$ differ. Our proof is based on the fact that Γ is an (exact) potential game if and only if every simple closed path γ of length 4 has $I(\gamma, \lambda) = 0$ ([29, Thm2.8]).

For the sake of contradiction, assume that every closed simple path γ of length 4 for Γ has $I(\gamma, \lambda) = 0$, fix arbitrary configuration ϖ and consider the path $\gamma = (\varpi, x = \varpi^{-1} \oplus \pi_1, y = \varpi^{-(1,2)} \oplus (\pi_1, \pi_2), z = \varpi^{-2} \oplus \pi_2, \varpi)$ for some paths $\pi_1 \neq \varpi^1$ and $\pi_2 \neq \varpi^2$. We shall

demonstrate that $I(\gamma,\lambda)$ cannot be identically 0 when there are at least two players of different demands. So, consider that the first two players have different demands: $w_1 \neq w_2$. We observe that

$$\lambda^1(x) - \lambda^1(\varpi) = \sum_{e\in\pi_1} \theta_e(x) - \sum_{e\in\varpi^1} \theta_e(\varpi) = |\pi_1 \setminus \varpi^1| \cdot w_1 + \sum_{e\in\pi_1\setminus\varpi^1} \theta_e(\varpi) - \sum_{e\in\varpi^1\setminus\pi_1} \theta_e(\varpi)$$

since the resources in $\varpi^1 \cap \pi_1$ retain their initial loads. Similarly we have:

$$\begin{aligned}
\lambda^2(y) - \lambda^2(x) &= \sum_{e\in\pi_2\setminus\varpi^2} [\theta_e(x) + w_2] - \sum_{e\in\varpi^2\setminus\pi_2} \theta_e(x) \\
&= |\pi_2 \setminus \varpi^2| \cdot w_2 + \sum_{e\in\pi_2\setminus\varpi^2} \theta_e(x) - \sum_{e\in\varpi^2\setminus\pi_2} \theta_e(x) \\
\lambda^1(z) - \lambda^1(y) &= \sum_{e\in\varpi^1\setminus\pi_1} \theta_e(z) - \sum_{e\in\pi_1\setminus\varpi^1} [\theta_e(z) + w_1] \\
&= \sum_{e\in\varpi^1\setminus\pi_1} \theta_e(z) - \sum_{e\in\pi_1\setminus\varpi^1} \theta_e(z) - |\pi_1 \setminus \varpi^1| \cdot w_1 \\
\lambda^2(\varpi) - \lambda^2(z) &= \sum_{e\in\varpi^2\setminus\pi_2} \theta_e(\varpi) - \sum_{e\in\pi_2\setminus\varpi^2} \theta_e(\varpi) - |\pi_2 \setminus \varpi^2| \cdot w_2
\end{aligned}$$

Thus, since $I \equiv I(\gamma,\lambda) = \lambda^1(x) - \lambda^1(\varpi) + \lambda^2(y) - \lambda^2(x) + \lambda^1(z) - \lambda^1(y) + \lambda^2(\varpi) - \lambda^2(z)$, we get:

$$\begin{aligned}
I &= \sum_{e\in\pi_1\setminus\varpi^1} [\theta_e(\varpi) - \theta_e(z)] + \sum_{e\in\pi_2\setminus\varpi^2} [\theta_e(x) - \theta_e(\varpi)] + \\
&+ \sum_{e\in\varpi^1\setminus\pi_1} [\theta_e(z) - \theta_e(\varpi)] + \sum_{e\in\varpi^2\setminus\pi_2} [\theta_e(\varpi) - \theta_e(x)]
\end{aligned}$$

Observe now that

$$\begin{aligned}
&\forall e \in \pi_1 \setminus \varpi^1, \quad \theta_e(\varpi) - \theta_e(z) = \theta_e(\varpi) - \theta_e(\varpi^{-2} \oplus \pi_2) = w_2 \cdot \left(\mathbb{I}_{\{e\in\varpi^2\setminus\pi_2\}} - \mathbb{I}_{\{e\in\pi_2\setminus\varpi^2\}}\right) \\
&\forall e \in \pi_2 \setminus \varpi^2, \quad \theta_e(x) - \theta_e(\varpi) = \theta_e(\varpi^{-1} \oplus \pi_1) - \theta_e(\varpi) = w_1 \cdot \left(\mathbb{I}_{\{e\in\pi_1\setminus\varpi^1\}} - \mathbb{I}_{\{e\in\varpi^1\setminus\pi_1\}}\right) \\
&\forall e \in \varpi^1 \setminus \pi_1, \quad \theta_e(z) - \theta_e(\varpi) = \theta_e(\varpi^{-2} \oplus \pi_2) - \theta_e(\varpi) = w_2 \cdot \left(\mathbb{I}_{\{e\in\pi_2\setminus\varpi^2\}} - \mathbb{I}_{\{e\in\varpi^2\setminus\pi_2\}}\right) \\
&\forall e \in \varpi^2 \setminus \pi_2, \quad \theta_e(\varpi) - \theta_e(x) = \theta_e(\varpi) - \theta_e(\varpi^{-1} \oplus \pi_1) = w_1 \cdot \left(\mathbb{I}_{\{e\in\varpi^1\setminus\pi_1\}} - \mathbb{I}_{\{e\in\pi_1\setminus\varpi^1\}}\right)
\end{aligned}$$

Then, $I = (w_1 - w_2) \cdot \big[|(\pi_1 \setminus \varpi^1) \cap (\pi_2 \setminus \varpi^2)| + |(\varpi^1 \setminus \pi_1) \cap (\varpi^2 \setminus \pi_2)| - |(\varpi^1 \setminus \pi_1) \cap (\pi_2 \setminus \varpi^2)| - |(\pi_1 \setminus \varpi^1) \cap (\varpi^2 \setminus \pi_2)|\big]$, which is typically not equal to zero for a single-commodity network. It should be noted that the second parameter, which is network dependent, can be non-zero even for some cycle of a very simple network. For example, in the network of figure 5 (which is a simple 2-layered network) the simple closed path $\gamma = (\varpi(0) = (P1, P3), \varpi(1) = (P2, P3), \varpi(2) = (P2, P1), \varpi(3) = (P1, P1), \varpi(4) = (P1, P3))$ has this quantity equal to -4 and thus no weighted single commodity network congestion game on this network can admit an exact potential. ∎

Our next step is to focus our interest on the ℓ-layered networks with resource delays identical to their loads. We shall prove that any weighted ℓ-layered network congestion game

with these delays admits at least one PNE, which can be computed in pseudo-polynomial time. Although we already know that even the case of weighted ℓ-layered network congestion games with delays equal to the loads cannot have any exact potential[4], we will next show that $\Phi(\varpi) \equiv \sum_{e\in E}[\theta_e(\varpi)]^2$ is a $\mathbf{b}$-potential for such a game and some positive n-vector $\mathbf{b}$, assuring the existence of a PNE.

Theorem 5 *For any weighted ℓ-layered network congestion game with resource delays equal to their loads, at least one PNE exists and can be computed in pseudo-polynomial time.*

Proof: Fix an arbitrary ℓ-layered network (V, E) and denote by $\mathcal{P}$ all the $s-t$ paths in it from the unique source s to the unique destination t. Let $\varpi \in \mathcal{P}^n$ be an arbitrary configuration of the players for the corresponding congestion game on (V, E). Also, let i be a user of demand w_i and fix some path $\alpha \in \mathcal{P}$. Denote $\hat{\varpi} \equiv \varpi^{-i} \oplus \alpha$. Observe that

$$\begin{aligned}
\Phi(\varpi) - \Phi(\hat{\varpi}) &= \sum_{e\in E} \left(\theta_e^2(\varpi) - \theta_e^2(\hat{\varpi})\right) \\
&= \sum_{e\in \varpi^i\setminus\alpha} \left(\theta_e^2(\varpi) - \theta_e^2(\hat{\varpi})\right) + \sum_{e\in \alpha\setminus\varpi^i} \left(\theta_e^2(\varpi) - \theta_e^2(\hat{\varpi})\right) \\
&= \sum_{e\in \varpi^i\setminus\alpha} \left([\theta_e(\varpi^{-i}) + w_i]^2 - \theta_e^2(\varpi^{-i})\right) + \sum_{e\in \alpha\setminus\varpi^i} \left(\theta_e^2(\varpi^{-i}) - [\theta_e(\varpi^{-i}) + w_i]^2\right) \\
&= \sum_{e\in \varpi^i\setminus\alpha} \left(w_i^2 + 2w_i\theta_e(\varpi^{-i})\right) - \sum_{e\in \alpha\setminus\varpi^i} \left(w_i^2 + 2w_i\theta_e(\varpi^{-i})\right) \\
&= w_i^2 \cdot \left(|\varpi^i \setminus \alpha| - |\alpha \setminus \varpi^i|\right) + 2w_i \cdot \left(\sum_{e\in \varpi^i\setminus\alpha} \theta_e(\varpi^{-i}) - \sum_{e\in \alpha\setminus\varpi^i} \theta_e(\varpi^{-i}) \right) \\
&= 2w_i \cdot \left(\sum_{e\in \varpi^i\setminus\alpha} \theta_e(\varpi^{-i}) - \sum_{e\in \alpha\setminus\varpi^i} \theta_e(\varpi^{-i}) \right) = 2w_i \cdot \left[\lambda^i(\varpi) - \lambda^i(\hat{\varpi})\right] ,
\end{aligned}$$

since, $\forall e \in \varpi^i \cap \alpha$, $\theta_e(\varpi) = \theta_e(\hat{\varpi})$, in ℓ-layered networks $|\varpi^i \setminus \alpha| = |\alpha \setminus \varpi^i|$, $\lambda^i(\varpi) = \sum_{e\in\varpi^i} \theta_e(\varpi) = \sum_{e\in\varpi^i\setminus\alpha} \theta_e(\varpi) = \sum_{e\in\varpi^i\setminus\alpha} \theta_e(\varpi^{-i}) + w_i|\varpi^i \setminus \alpha| + \sum_{e\in\varpi^i\cap\alpha} \theta_e(\varpi)$ and $\lambda^i(\hat{\varpi}) = \sum_{e\in\alpha} \theta_e(\hat{\varpi}) = \sum_{e\in\alpha\setminus\varpi^i} \theta_e(\varpi^{-i}) + w_i|\alpha \setminus \varpi^i| + \sum_{e\in\varpi^i\cap\alpha} \theta_e(\varpi)$. Thus, Φ is a $\mathbf{b}-$potential for our game, where $\mathbf{b} = (1/(2w_i))_{i\in N} > \mathbf{0}$, assuring the existence of at least one PNE.

We proceed with the construction of a PNE in pseudopolynomial time. Wlog assume now that the players have integer weights. Then each player performing any improving defection, must reduce his cost by at least 1 and thus the potential function decreases by at least $2w_{\min} \geqslant 2$ along each arc of the Nash Dynamics Graph of the game. Consequently, the naïve algorithm that, starting from an arbitrary initial configuration $\varpi \in \mathcal{P}^n$, follows any improvement path that leads to a sink (ie, a PNE) of the Nash Dynamics Graph, cannot contain more than $\frac{1}{2}|E|W_{\text{tot}}^2$ defections, since $\forall \varpi \in \mathcal{P}^n$, $\Phi(\varpi) \leqslant |E|W_{\text{tot}}^2$. ∎

A recent improvement, based essentially on the same technique as above, generalizes the last result to the case of arbitrary multicommodity network congestion games with linear resource delays (we state the result here without its proof):

Theorem 6 ([16]) *For any weighted multi-commodity network congestion game with linear resource delays, at least one PNE exists and can be computed in pseudo-polynomial time.*

[4]The example at the end of the proof of lemma 5.2 involves the 2-layered network of figure 5.

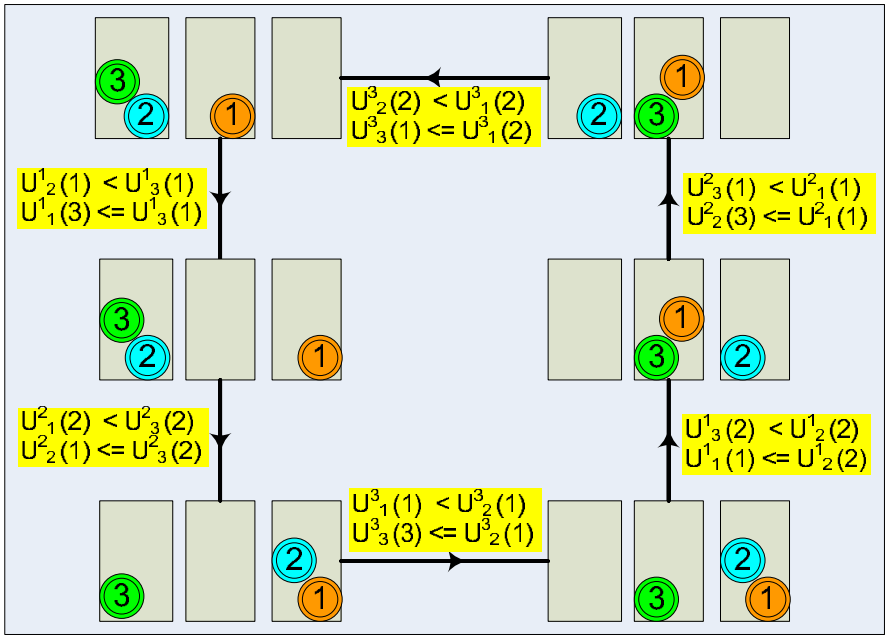


Figure 6: An example of a 3-players, 3-strategies unweighted congestion game with a best-reply cycle.

6 Congestion Games on Parallel Links with Player Specific Payoffs [28]

In his work, Milchtaich studies a variant of the classical (unweighted) congestion games, where the resource charging functions are not universal, but player-specific. On the other hand, he makes two simplifying (yet crucial) assumptions:

(1) each player may choose only one resource from a pool E of resources (shared to all the players) for his service (ie, this is modeled as the *parallel-links* model of Koutsoupias and Papadimitriou [23]), and

(2) the received payoff is *monotonically non-increasing* with the number of players selecting it. Although they do not always admit a potential, these games always possess a PNE.

In his paper, Milchtaich proves that unweighted congestion games on parallel links with player-specific payoffs, involving only two strategies, possess the Finite Improvement Property (FIP). It is also rather straightforward that any 2-players unweighted congestion game on parallel links with player-specific payoffs possesses the Finite Best Reply improvement path Property (FBRP).

Milchtaich also gave an example of an unweighted congestion game on 3 parallel links with 3 players, for which there is a best-reply cycle (although there is a PNE). The example is shown in figure 6. In this example, we only determine the necessary conditions on the payoff functions of the players for the existence of the best-response cycle (see figure). It is easily verified that this system of inequalities is feasible, and that configurations $(3, 1, 2)$ and $(2, 3, 1)$ are PNE for the game.

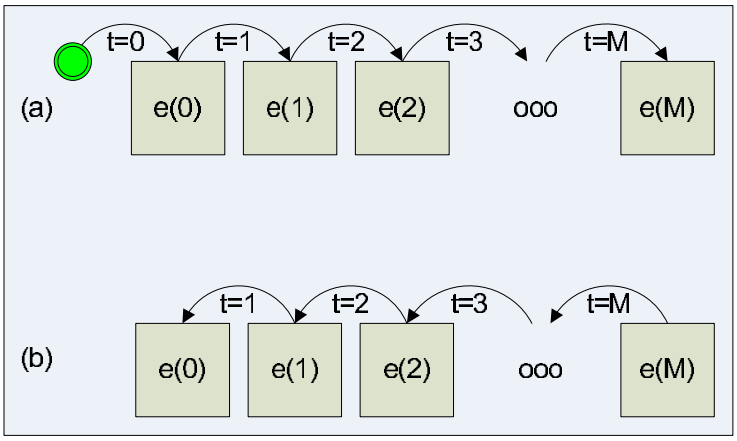


Figure 7: The two types of best-reply improvement paths considered by [28].

Theorem 7 ([28]) *Every unweighted congestion game on parallel links with player-specific, monotonically non-increasing payoffs of the resources, possesses a PNE.*

Proof: First of all, we need to prove the following lemma that bounds the lengths of best-reply paths of a specific kind.

Lemma 6.1 *Let $\langle e(0), e(1), \ldots, e(M)\rangle$ be a sequence of (possibly repeated) resources from E.*

type–(a) path: *Let $\langle \varpi(1), \varpi(2), \ldots, \varpi(M)\rangle$ be a best-reply improvement path of the game, s.t. $\forall t \geqslant 1, \forall e \in E$,*

$$x_e(t) = \begin{cases} x_e(t-1), & \text{if } e \neq e(t-1), e(t) \\ x_e(t-1)+1, & \text{if } e = e(t) \\ x_e(t-1)-1, & \text{if } e = e(t-1) \end{cases}$$

Then $M \leqslant |N|$.

type–(b) path: *Let $\langle \varpi(1), \varpi(2), \ldots, \varpi(M)\rangle$ be a best-reply improvement path of the game, s.t. $\forall t \geqslant 1, \forall e \in E$,*

$$x_e(t) = \begin{cases} x_e(t-1), & \text{if } e \neq e(t-1), e(t) \\ x_e(t-1)+1, & \text{if } e = e(t-1) \\ x_e(t-1)-1, & \text{if } e = e(t) \end{cases}$$

Then, $M \leqslant |N| \cdot (|E|-1)$.

Proof: (a) Let $\forall e \in E, x_e^{min} \equiv \min_{0 \leqslant t \leqslant M} x_e(t)$. By definition of the best-reply sequence, we observe that $\forall e \in E, \forall t \geqslant 1, x_{e(t)}(t) = x_{e(t)}(t-1)+1 = x_e^{min}+1$. That is, a resource reaches its maximum load exactly when it is the next node in the sequence of resources, reaches its minimum load in the very next time step, and remains at its minimum load until before it appears again in this sequence (see case (a) of figure 7). But then, the unique deviator in each move goes to a resource that reaches its maximum load, whereas all the other resources are at their minimum loads at the same time. This implies that there is no chance that the

specific player will move away from this new resource that he chose, till the end of the best reply-path under consideration. More formally, $\forall t \geqslant 1$, let $i(t)$ be the unique deviator that moves from $e(t-1)$ to $e(t)$. Then,

$$\forall e \in E, \forall t+1 \leqslant t' \leqslant M,\ U_{e(t)}^{i(t)}\left(x_e(t')\right) \geqslant U_{e(t)}^{i(t)}\left(x_{e(t)}^{\min}+1\right) \geqslant U_e^{i(t)}\left(x_e^{\min}\right) \geqslant U_e^{i(t)}\left(x_e(t')\right)$$

and so, $i(t)$ cannot move away from $e(t)$ till the end of the best-reply improvement path. Since this holds for all players, this path may have length at most $|N|$.
(b) Let again $\forall e \in E, x_e^{\min} \equiv \min_{0 \leqslant t \leqslant M} x_e(t)$. We observe that $\forall 1 \leqslant t \leqslant M, x_{e(t)}(t) = x_{e(t)}^{\min}$ and $x_{e(t-1)}(t) = x_{e(t-1)}^{\min} + 1$. Due to the best-reply moves considered in the path, if $i(t)$ is again the unique deviator at time t, then $\forall 1 \leqslant t \leqslant M$,

$$U_{e(t)}^{i(t)}\left(x_{e(t)}(t-1)\right) = U_{e(t)}^{i(t)}\left(x_{e(t)}^{\min}+1\right) < U_{e(t-1)}^{i(t)}\left(x_{e(t-1)}(t-1)+1\right) = U_{e(t-1)}^{i(t)}\left(x_{e(t-1)}^{\min}+1\right)$$

which implies that $\forall t' > t$ player $i(t)$ residing at a resource $e \neq e(t)$ could never prefer to deviate to $e(t)$ rather than deviate to (or, stay at, if already there) resource $e(t-1)$. That is, player $i(t)$ can never go back to the resource from which it defected once, till the end of the best-reply path. Thus, each player can make at most $|E|-1$ moves and so, $M \leqslant |N| \cdot (|E|-1)$.
■

The proof of the theorem proceeds by induction on the number of players. Trivially, for a single player we know that as soon as he is placed at its best reply, this is also a PNE (there is nothing else to move). We assume now that any unweighted congestion game on parallel links with player-specific payoffs and $|\tilde{N}| < n$, $\tilde{\Gamma} = \left(\tilde{N}, E, (U_e^i)_{i \in \tilde{N}, e \in E}\right)$, possesses a PNE. We want to prove that this is also the case for any such game with n players. Let $N = [n]$ and $\Gamma = \left(N, E, (U_e^i)_{i \in N, e \in E}\right)$ be such a game. We temporarily pull player n out of the game and let the remaining $n-1$ players continue playing, until they reach a PNE (without player n) $\mathbf{s} \in E^{n-1}$. That the PNE $\mathbf{s}$ exists for the pruned game $\tilde{\Gamma} = \left([n-1], E, (U_e^i)_{i \in [n-1]}, e \in E\right)$ holds by inductive hypothesis. Now we let player n be assigned to a best-reply resource $e \in BR_n(\mathbf{s})$, and we thus construct a configuration $\varpi(0)$ as follows: $\varpi^n(0) = e$; $\forall i \in [n-1], \varpi^i(0) = s^i$.

Consequently, starting from $\varpi(0)$, we construct a best-reply *maximal* improvement path of type (a) (see previous lemma) $\Pi(a) = \langle \varpi(0), \varpi(1), \ldots, \varpi(M) \rangle$, which we already know that is of length at most n. We claim that the terminal configuration of this path is a PNE for Γ. Clearly, any player that deviated to a best-reply resource in $\Pi(a)$ does not move again and is also at a best-reply resource in $\varpi(M)$ (see the proof of the lemma). So we only have to consider players that have not defected during $\Pi(a)$. Fix any such player i, residing at resource $e = \varpi^i(0) = s^i$. Observe that in $\varpi(M)$ the following holds:

$$\forall e \in E, x_e(M) = \begin{cases} x_e(0), & \text{if } e \neq e(M), \\ x_e(0)+1, & \text{if } e = e(M) \end{cases}$$

Observe that if some player $i : s^i(M) = e(M)$ that has not moved during $\Pi(a)$ is not at his best reply in $\varpi(M)$, then we can set $e(M+1) \in BR_i(\varpi^{-i}(M))$ and thus augment the best-reply improvement path $\Pi(a)$, which contradicts its maximality assumption. Consider now any player $i \in e \neq e(M)$ that has not moved during $\Pi(a)$. This player is certainly at a best-reply resource $e \in BR_i(\varpi^{-i}(M))$ since this resource has exactly the same load as in $\mathbf{s}$ and any other resource has at least the load it had in $\mathbf{s}$. So, $\varpi(M)$ is a PNE for Γ since every player is at a best-reply resource wrt $\varpi(M)$. This completes the proof of the theorem. ■

Remark: Observe that the proof of this theorem is constructive, and thus also implies a path of length at most $|N|$ that leads to a PNE. But this is not necessarily an improvement path, when all players are considered to coexist all the time, and therefore there is no justification of the adoption of such a path by the (selfish) players. Milchtaich [28], using an argument of the same flavor as in the above theorem, proves that from an arbitrary initial configuration and allowing only best-reply defections, there is a best-reply improvement path of length at most $|E| \cdot \binom{|N|+1}{2}$. The idea is, starting from an arbitrary configuration $\varpi(0)$, to let the players construct a best-reply improvement path of type–(a), and then (if necessary) construct a best-reply improvement path of type–(b) starting from the terminal configuration of the previous path. It is then easily shown that the terminal configuration of the second path is a PNE of the game.

The unweighted congestion games on parallel links and with player-specific payoffs are **weakly acyclic** games, in the sense that from any initial configuration $\varpi(0)$ of the players, there is at least one best-reply improvement path connecting it to a PNE. Of course, this does not exclude the existence of best-reply cycles (see example of figure 6). But, it is easily shown that when the deviations of the players occur sequentially and in each step the next deviator is chosen randomly (among the potential deviators) to a randomly chosen best-reply resource, then this path will converge almost surely to a PNE in finite time.

6.1 Allowing different weights to the players

Milchtaich proposed a generalization of his variant of congestion games, by allowing the players to have distinct weights, denoted by a weight vector $\mathbf{w} = (w^1, w^2, \ldots, w^n) \in \mathbb{R}^n_{>0}$. In that case, the (player-specific) payoff of each player on a resource $e \in E$ depends on the load $\theta_e(\varpi) \equiv \sum_{i:e\in\varpi^i} w_i$, rather than the number of players willing to use it.

For the case of weighted congestion games on parallel links with player specific payoffs, it is easy to verify (in a similar fashion as for the unweighted case) that:

- If there are only two available strategies then FIP holds.
- If there are only two players then FBRP holds.
- For the special case of resource specific payoffs, FIP holds.

On the other hand, there exists a 3-players, 3-strategies game that possesses no PNE. For example, see the instance shown in figure 8, where the three players have essentially two strategies each (a "LEFT" and a "RIGHT" strategy) to choose from, while their third strategies give them strictly minimal payoffs and can never be chosen by selfish moves. The rationale of this game is that, in principle, player 1 would like to avoid using the same link as player 3, which in turn would like to avoid using the same link as player 2, which would finally want to avoid player 1.

The inequalities shown in figure 8(b) demonstrate the necessary conditions for the existence of a best-reply cycle among 6 configurations of the players. It is easy to verify also that any other configuration has either at least one MIN strategy for some player (and in that case this player wants to defect) or is one of $(2,2,1), (3,3,3)$. The only thing that remains to assure, is that $(2,3,1)$ is strictly better for player 2 than $(2,2,1)$ (ie, player 2 would like to avoid player 1) and that $(3,3,1)$ is better for player 3 than $(3,3,3)$ (ie, player 3 would like to

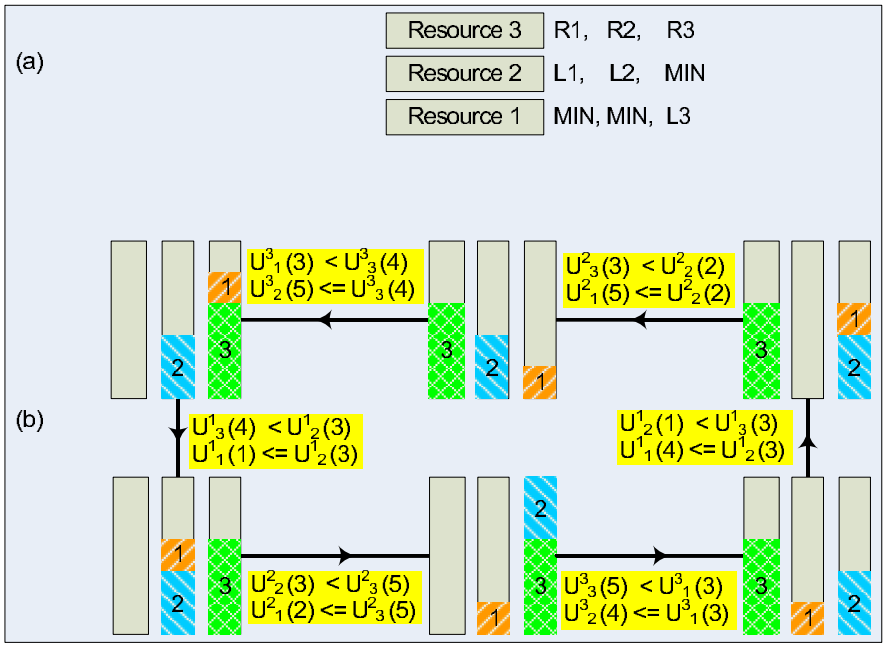


Figure 8: A 3-players weighted congestion game on 3 parallel links with player-specific payoffs without a PNE [28]. (a) The LEFT-RIGHT strategies of the players. (b) The best-reply cycle.

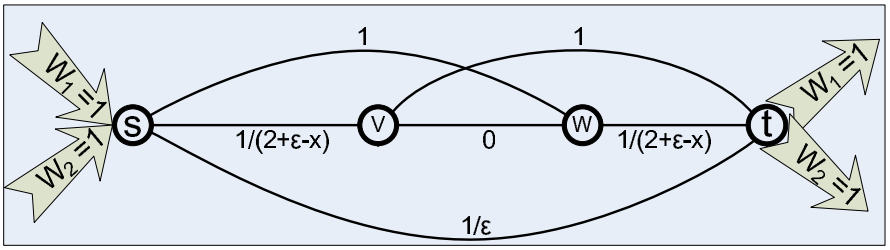


Figure 9: An example of a single-source network congestion game without a PNE ([32]).

avoid player 2). The feasibility of the whole system of inequalities can be trivially checked to hold, and thus this game cannot have any PNE since there is no sink in its Dynamics graph.

7 The Price of Anarchy of Weighted Congestion Games: [15]

In this section we focus our interest on weighted ℓ-layered network congestion games where the resource delays are identical to their loads. Our source for this section is [15]. This case comprises a non-trivial generalization of routing through identical parallel channels.

The main reason why we focus on this specific category of resource delays is that selfish unsplittable flows can have unbounded price of anarchy even for linear resource delays. In [32, p. 256] an example is given where the price of anarchy is unbounded (see figure 9). This example is easily converted in an ℓ-layered network. The resource delay functions used are either constant or M/M/1-like delay functions. But we can be equally bad even with linear resource delay functions: Observe the following example of figure 10. Two players, each of

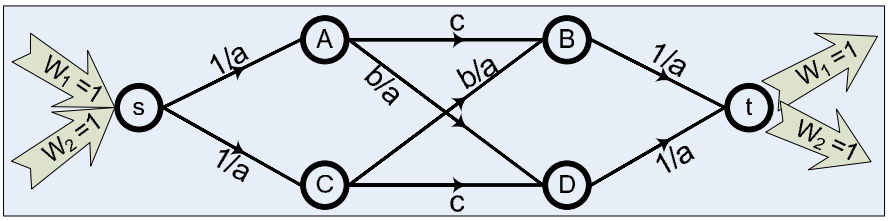


Figure 10: Example of an ℓ-layered network with linear resource delays and unbounded anarchy.

unit demand, want to move selfishly from s to t. The edge delays are shown above them. We assume that $a \gg b \gg 1 \geqslant c$). It is easy to see that the configuration (sCBt,sADt) is a PNE of social cost $2+b$ while the optimum configuration is (sABt,sCDt) whose social optimum is $2+c$. Thus, $\mathcal{R} = \frac{b+2}{c+2}$.

In the following, we restrict our attention to ℓ-layered networks whose resource delays are equal to their loads. Our main tool is to interpret a strategies profile as a flow in the underlying network.

7.1 Flows and mixed strategies profiles.

Fix an arbitrary ℓ-layered network $G = (V, E)$ and n distinct players willing to satisfy their own traffic demands from the unique source $s \in V$ to the unique destination $t \in V$. Again, $\mathbf{w} = (w_i)_{i\in[n]}$ denotes the varying demands of the players. Fix an arbitrary mixed strategies profile $\mathbf{p} = (p_1, p_2, \ldots, p_n)$. A *feasible flow* for the n players is a function $\rho : \mathcal{P} \mapsto \mathbb{R}_{\geqslant 0}$, s.t. $\sum_{\pi\in\mathcal{P}} \rho(\pi) = W_{\text{tot}} \equiv \sum_{i\in[n]} w_i$, ie, all players' demands are actually met. We distinguish between *unsplittable* and *splittable* (feasible) flows. A flow is unsplittable if each player's traffic demand is satisfied by a unique path of $\mathcal{P}$. A flow is splittable if the traffic demand of each player can be routed over several paths of $\mathcal{P}$.

We map the mixed strategies profile $\mathbf{p}$ to a flow $\rho_{\mathbf{p}}$ as follows: For each $s-t$ path $\pi \in \mathcal{P}$, $\rho_{\mathbf{p}}(\pi) \equiv \sum_{i\in[n]} w_i \cdot p_i(\pi)$. That is, we handle the *expected load traveling along π according to* $\mathbf{p}$ as a splittable flow, where player i routes a fraction of $p_i(\pi)$ of its total demand w_i along π. Observe that, if $\mathbf{p}$ is actually a pure strategies profile, the corresponding flow is then unsplittable. Recall now that for each edge $e \in E$,

$$\theta_e(\mathbf{p}) \equiv \sum_{i=1}^{n} \sum_{\pi\ni e} w_i p_i(\pi) = \sum_{\pi\ni e} \rho_{\mathbf{p}}(\pi) \equiv \theta_e(\rho_{\mathbf{p}})$$

denotes the expected load (and in our case, also the expected delay) of e wrt $\mathbf{p}$. As for the expected delay along a path $\pi \in \mathcal{P}$ according to $\mathbf{p}$, this is

$$\theta_\pi(\mathbf{p}) \equiv \sum_{e\in\pi} \theta_e(\mathbf{p}) = \sum_{e\in\pi} \sum_{\pi'\ni e} \rho_{\mathbf{p}}(\pi') = \sum_{\pi'\in\mathcal{P}} |\pi \cap \pi'| \rho_{\mathbf{p}}(\pi') \equiv \theta_\pi(\rho_{\mathbf{p}})$$

Let $\theta^{\min}(\rho) = \min_{\pi\in\mathcal{P}}\{\theta_\pi(\rho)\}$ be the minimum expected delay among all $s-t$ paths. From now on for simplicity we drop the subscript of $\mathbf{p}$ from its corresponding flow $\rho_{\mathbf{p}}$, when this is clear by the context. We evaluate flow ρ using the objective of *maximum latency*, which is defined as

$$L(\rho) \equiv \max_{\pi:\rho(\pi)>0} \{\theta_\pi(\rho)\} = \max_{\pi:\exists i,\ p_i(\pi)>0} \{\theta_\pi(\mathbf{p})\} \equiv L(\mathbf{p}) \tag{8}$$

$L(\rho)$ is nothing but the *maximum expected delay paid by the players*, wrt $\mathbf{p}$. From now on, we use ρ^* and ρ^*_f to denote the optimal unsplittable and splittable flows respectively.

In addition, we sometimes evaluate flow ρ using the objective of *total latency*, which is defined as

$$C(\rho) \equiv \sum_{\pi \in \mathcal{P}} \rho(\pi)\theta_\pi(\rho) = \sum_{e \in E} \theta_e^2(\rho) = \sum_{e \in E} \theta_e^2(\mathbf{p}) \equiv C(\mathbf{p}) \tag{9}$$

The second equality is obtained by summing over the edges of π and reversing the order of the summation.

7.2 Flows at Nash equilibrium.

Let $\mathbf{p}$ be a mixed strategies profile and let ρ be the corresponding flow. For a ℓ-layered network with resource delays equal to the loads, the cost of player i on path π is $\lambda_\pi^i(\mathbf{p}) = \ell w_i + \theta_\pi^{-i}(\mathbf{p})$, where $\theta_\pi^{-i}(\mathbf{p})$ is the expected delay along path π if the demand of player i was removed from the system:

$$\theta_\pi^{-i}(\mathbf{p}) = \sum_{\pi' \in \mathcal{P}} |\pi \cap \pi'| \sum_{j \neq i} w_j p_j(\pi') = \theta_\pi(\mathbf{p}) - w_i \sum_{\pi' \in \mathcal{P}} |\pi \cap \pi'| p_i(\pi') \tag{10}$$

Thus, $\lambda_\pi^i(\mathbf{p}) = \theta_\pi(\mathbf{p}) + \left[\ell - \sum_{\pi' \in \mathcal{P}} |\pi \cap \pi'| p_i(\pi')\right] w_i$. Observe now that, if $\mathbf{p}$ is a NE, then $L(\mathbf{p}) = L(\rho) \leqslant \theta^{\min}(\rho) + \ell\, w_{\max}$. Otherwise, the players routing their traffic on a path of expected delay greater than $\theta^{\min}(\rho) + \ell\, w_{\max}$ could improve their delay by defecting to a path of expected delay $\theta^{\min}(\rho)$. We sometimes say that a flow ρ corresponding to a mixed strategies profile $\mathbf{p}$ is a NE with the understanding that it is actually $\mathbf{p}$ which is a NE.

7.3 Maximum Latency versus Total Latency.

We show that if the resource delays are equal to their loads, a splittable flow is optimal wrt the objective of maximum latency if and only if it is optimal wrt the objective of total latency. As a corollary, we obtain that the optimal splittable flow defines a NE where all players adopt the same mixed strategy.

Lemma 7.1 *There is a unique feasible splittable flow ρ which minimizes both $L(\rho)$ and $C(\rho)$.*

Proof: For every feasible flow ρ, the average latency of ρ cannot exceed its maximum latency:

$$C(\rho) = \sum_{\pi \in \mathcal{P}} \rho(\pi)\theta_\pi(\rho) = \sum_{\pi : \rho(\pi) > 0} \rho(\pi)\theta_\pi(\rho) \leqslant L(\rho)\, W_{\text{tot}} \tag{11}$$

A splittable flow ρ minimizes $C(\rho)$ if and only if for every $\pi_1, \pi_2 \in \mathcal{P}$ with $\rho(\pi_1) > 0$, $\theta_{\pi_1}(\rho) \leqslant \theta_{\pi_2}(\rho)$ (e.g., [6], [30, Section 7.2], [32, Corollary 4.2]). Hence, if ρ is optimal wrt the objective of total latency, for all paths $\pi \in \mathcal{P}$, $\theta_\pi(\rho) \geqslant L(\rho)$. Moreover, if $\rho(\pi) > 0$, then $\theta_\pi(\rho) = L(\rho)$. Therefore, if ρ minimizes $C(\rho)$, then the average latency is equal to the maximum latency:

$$C(\rho) = \sum_{\pi \in \mathcal{P} : \rho(\pi) > 0} \rho(\pi)\theta_\pi(\rho) = L(\rho) W_{\text{tot}} \tag{12}$$

Let ρ be the feasible splittable flow that minimizes the total latency and let ρ' be the feasible splittable flow that minimizes the maximum latency. We prove the lemma by establishing that the two flows are identical.

Observe that $L(\rho') \geqslant \frac{C(\rho')}{W_{\text{tot}}} \geqslant \frac{C(\rho)}{W_{\text{tot}}} = L(\rho)$. The first inequality follows from Ineq. (11), the second from the assumption that ρ minimizes the total latency and the last equality from Eq. (12). On the other hand, it must be $L(\rho') \leqslant L(\rho)$ because of the assumption that the flow ρ' minimizes the maximum latency. Hence, it must be $L(\rho') = L(\rho)$ and $C(\rho') = C(\rho)$. In addition, since the function $C(\rho)$ is strictly convex and the set of feasible splittable flows forms a convex polytope, there is a unique flow which minimizes the total latency. Thus, ρ and ρ' must be identical. ∎

Corollary 7.1 *A flow ρ minimizes the maximum latency if and only if for every $\pi_1, \pi_2 \in \mathcal{P}$ with $\rho(\pi_1) > 0$, $\theta_{\pi_1}(\rho) \leqslant \theta_{\pi_2}(\rho)$.*

Proof: By Lemma 7.1, the flow ρ minimizes the maximum latency if and only if it minimizes the total latency. Then, the corollary follows from the the fact that ρ minimizes the total latency if and only if for every $\pi_1, \pi_2 \in \mathcal{P}$ with $\rho(\pi_1) > 0$, $\theta_{\pi_1}(\rho) \leqslant \theta_{\pi_2}(\rho)$ (eg, [30, Section 7.2], [32, Corollary 4.2]). ∎

The following corollary states that the optimal splittable flow defines a mixed NE where all players adopt exactly the same strategy.

Corollary 7.2 *Let ρ_f^* be the optimal splittable flow and let $\mathbf{p}$ be the mixed strategies profile where every player routes its traffic on each path π with probability $\rho_f^*(\pi)/W_{\text{tot}}$. Then, $\mathbf{p}$ is a NE.*

Proof: By construction, the expected path loads corresponding to $\mathbf{p}$ are equal to the values of ρ_f^* on these paths. Since all players follow exactly the same strategy and route their demand on each path π with probability ρ_f^*/W_{tot}, for each player i,

$$\theta_\pi^{-i}(\mathbf{p}) = \theta_\pi(\mathbf{p}) - w_i \sum_{\pi' \in \mathcal{P}} |\pi \cap \pi'| \, \frac{\rho_f^*(\pi')}{W_{\text{tot}}} = (1 - \tfrac{w_i}{W_{\text{tot}}}) \, \theta_\pi(\mathbf{p})$$

Since the flow ρ_f^* also minimizes the total latency, for every $\pi_1, \pi_2 \in \mathcal{P}$ with $\rho_f^*(\pi_1) > 0$, $\theta_{\pi_1}(\mathbf{p}) \leqslant \theta_{\pi_2}(\mathbf{p})$ (eg, [6], [30, Section 7.2], [32, Corollary 4.2]), which also implies that $\theta_{\pi_1}^{-i}(\mathbf{p}) \leqslant \theta_{\pi_2}^{-i}(\mathbf{p})$. Therefore, for every player i and every $\pi_1, \pi_2 \in \mathcal{P}$ such that the player i routes its demand on π_1 with positive probability, $\lambda_{\pi_1}^i(\mathbf{p}) = \ell w_i + \theta_{\pi_1}^{-i}(\mathbf{p}) \leqslant \ell w_i + \theta_{\pi_2}^{-i}(\mathbf{p}) = \lambda_{\pi_2}^i(\mathbf{p})$. Consequently, $\mathbf{p}$ is a NE. ∎

7.4 An Upper Bound on the Social Cost.

Next we derive an upper bound on the social cost of every strategy profile whose maximum expected delay (ie, the maximum latency of its associated flow) is within a constant factor from the maximum latency of the optimal unsplittable flow.

Lemma 7.2 *Let ρ^* be the optimal unsplittable flow, and let $\mathbf{p}$ be a mixed strategies profile and ρ its corresponding flow. If $L(\mathbf{p}) = L(\rho) \leqslant \alpha \, L(\rho^*)$, for some $\alpha \geqslant 1$, then*

$$\text{SC}(\mathbf{p}) \leqslant (\alpha + 1) \, \mathcal{O}(\tfrac{\log m}{\log \log m}) \, L(\rho^*) \, ,$$

where $m = |E|$ denotes the number of edges in the network.

Proof: For each edge $e \in E$ and each player i, let $X_{e,i}$ be the random variable describing the actual load routed through e by i. The random variable $X_{e,i}$ is equal to w_i if i routes its demand on a path π including e and 0 otherwise. Consequently, the expectation of $X_{e,i}$ is equal to $\mathbb{E}\{X_{e,i}\} = \sum_{\pi:e\in\pi} w_i p_i(\pi)$. Since each player selects its path independently, for every fixed edge e, the random variables $\{X_{e,i},\ i \in [n]\}$, are independent from each other.

For each edge $e \in E$, let $X_e = \sum_{i=1}^{n} X_{e,i}$ be the random variable that describes the actual load routed through e, and thus, also the actual delay paid by any player traversing e. X_e is the sum of n independent random variables with values in $[0, w_{\max}]$. By linearity of expectation,

$$\mathbb{E}\{X_e\} = \sum_{i=1}^{n} \mathbb{E}\{X_{e,i}\} = \sum_{i=1}^{n} w_i \sum_{\pi\ni e} p_i(\pi) = \theta_e(\rho)\,.$$

By applying the standard Hoeffding bound[5]with $w = w_{\max}$ and $t = \exp\,\kappa \max\{\theta_e(\rho), w_{\max}\}$, we obtain that for every $\kappa \geqslant 1$,

$$\mathbb{P}\{X_e \geqslant \exp\,\kappa \max\{\theta_e(\rho), w_{\max}\}\} \leqslant \kappa^{-\exp\,\kappa}\,.$$

For $m \equiv |E|$, by applying the union bound we conclude that

$$\mathbb{P}\{\exists e \in E : X_e \geqslant \exp\,\kappa \max\{\theta_e(\rho), w_{\max}\}\} \leqslant m\kappa^{-\exp\,\kappa} \tag{13}$$

For each path $\pi \in \mathcal{P}$ with $\rho(\pi) > 0$, we define the random variable $X_\pi = \sum_{e\in\pi} X_e$ describing the actual delay along π. The social cost of $\mathbf{p}$, which is equal to the expected maximum delay experienced by some player, cannot exceed the expected maximum delay among paths π with $\rho(\pi) > 0$. Formally,

$$\mathrm{SC}(\mathbf{p}) \leqslant \mathbb{E}\left\{\max_{\pi:\rho(\pi)>0}\{X_\pi\}\right\}\,.$$

If for all $e \in E$, $X_e \leqslant \exp\,\kappa \max\{\theta_e(\rho), w_{\max}\}$, then for every path $\pi \in \mathcal{P}$ with $\rho(\pi) > 0$,

$$\begin{aligned} X_\pi = \sum_{e\in\pi} X_e &\leqslant \exp\,\kappa \sum_{e\in\pi} \max\{\theta_e(\rho), w_{\max}\} \\ &\leqslant \exp\,\kappa \sum_{e\in\pi} (\theta_e(\rho) + w_{\max}) \\ &= \exp\,\kappa\,(\theta_\pi(\rho) + \ell w_{\max}) \\ &\leqslant \exp\,\kappa\,(L(\rho) + \ell w_{\max}) \\ &\leqslant \exp\,(\alpha+1)\kappa\,L(\rho^*) \end{aligned}$$

The third equality follows from $\theta_\pi(\rho) = \sum_{e\in\pi} \theta_e(\rho)$, the fourth inequality from $\theta_\pi(\rho) \leqslant L(\rho)$ since $\rho(\pi) > 0$, and the last inequality from the hypothesis that $L(\rho) \leqslant \alpha\,L(\rho^*)$ and the fact that $\ell w_{\max} \leqslant L(\rho^*)$ because ρ^* is an unsplittable flow. Therefore, using Ineq. (13), we conclude that

$$\mathbb{P}\left\{\max_{\pi:\rho(\pi)>0}\{X_\pi\} \geqslant \exp\,(\alpha+1)\kappa\,L(\rho^*)\right\} \leqslant m\kappa^{-\exp\,\kappa}\,.$$

[5]We use the standard version of Hoeffding bound ([19]): Let $X_1, X_2, \ldots, X_n$ be independent random variables with values in the interval $[0, w]$. Let $X = \sum_{i=1}^{n} X_i$ and let $\mathbb{E}\{X\}$ denote its expectation. Then, $\forall t > 0$, $\mathbb{P}\{X \geqslant t\} \leqslant \left(\frac{\exp\,\mathbb{E}\{X\}}{t}\right)^{t/w}$.

In other words, the probability that the actual maximum delay caused by $\mathbf{p}$ exceeds the optimal maximum delay by a factor greater than $2\exp(\alpha+1)\kappa$ is at most $m\kappa^{-\exp\kappa}$. Therefore, for every $\kappa_0 \geqslant 2$,

$$\begin{aligned} \mathrm{SC}(\mathbf{p}) \leqslant \mathbb{E}\left\{\max_{\pi:\rho(\pi)>0}\{X_\pi\}\right\} &\leqslant \exp(\alpha+1)L(\rho^*)\left(\kappa_0 + \sum_{k=\kappa_0}^{\infty} km k^{-\exp k}\right) \\ &\leqslant \exp(\alpha+1)L(\rho^*)\Big(\kappa_0 + 2m\kappa_0^{-\exp\kappa_0+1}\Big) \end{aligned}$$

If $\kappa_0 = \frac{2\log m}{\log\log m}$, then $\kappa_0^{-\exp\kappa_0+1} \leqslant m^{-1}$, $\forall m \geqslant 4$. Thus, $\mathrm{SC}(\mathbf{p}) \leqslant 2\exp(\alpha+1)(\frac{\log m}{\log\log m} + 1)L(\rho^*)$. ■

7.5 Bounding the Coordination Ratio.

Our final step is to show that the maximum expected delay of every NE is a good approximation to the optimal maximum latency. Then, we can apply Lemma 7.2 to bound the coordination ratio for our selfish routing game.

Lemma 7.3 *For every flow ρ corresponding to a mixed strategies profile $\mathbf{p}$ at NE, $L(\rho) \leqslant 3L(\rho^*)$.*

Proof: The proof is based on Dorn's Theorem [10] which establishes strong duality in quadratic programming[6]. We use quadratic programming duality to prove that for any flow ρ at Nash equilibrium, the minimum expected delay $\theta^{\min}(\rho)$ cannot exceed $L(\rho_f^*) + \ell\, w_{\max}$. This implies the lemma because $L(\rho) \leqslant \theta^{\min}(\rho) + \ell\, w_{\max}$, since ρ is at Nash equilibrium, and $L(\rho^*) \geqslant \max\{L(\rho_f^*), \ell\, w_{\max}\}$, since ρ^* is an unsplittable flow.

Let Q be the square matrix describing the number of edges shared by each pair of paths. Formally, Q is a $|\mathcal{P}| \times |\mathcal{P}|$ matrix and for every $\pi, \pi' \in \mathcal{P}$, $Q[\pi,\pi'] = |\pi \bigcap \pi'|$. By definition, Q is symmetric. Next we prove that Q is positive semi-definite[7].

$$\begin{aligned} x^T Q x &= \sum_{\pi\in\mathcal{P}} x(\pi) \sum_{\pi'\in\mathcal{P}} Q[\pi,\pi']x(\pi') \\ &= \sum_{\pi\in\mathcal{P}} x(\pi) \sum_{\pi'\in\mathcal{P}} |\pi \bigcap \pi'| x(\pi') \\ &= \sum_{\pi\in\mathcal{P}} x(\pi) \sum_{e\in\pi} \sum_{\pi':e\in\pi'} x(\pi') \\ &= \sum_{\pi\in\mathcal{P}} x(\pi) \sum_{e\in\pi} \theta_e(x) \\ &= \sum_{e\in E} \theta_e(x) \sum_{\pi:e\in\pi} x(\pi) \\ &= \sum_{e\in E} \theta_e^2(x) \geqslant 0 \end{aligned}$$

[6] Let $\min\{x^TQx + c^Tx : Ax \geqslant b, x \geqslant \mathbf{0}\}$ be the primal quadratic program. The Dorn's dual of this program is $\max\{-y^TQy + b^Tu : A^Tu - 2Qy \leqslant c, u \geqslant \mathbf{0}\}$. Dorn [10] proved strong duality when the matrix Q is symmetric and positive semi-definite. Thus, if Q is symmetric and positive semi-definite and both the primal and the dual programs are feasible, their optimal solutions have the same objective value.

[7] A $n \times n$ matrix Q is positive semi-definite if for every vector $x \in \mathbb{R}^n$, $x^TQx \geqslant 0$.

First recall that for each edge e, $\theta_e(x) \equiv \sum_{\pi: e \in \pi} x(\pi)$. The third and the fifth equalities follow by reversing the order of summation. In particular, in the third equality, instead of considering the edges shared by π and π', for all $\pi' \in \mathcal{P}$, we consider all the paths π' using each edge $e \in \pi$. On both sides of the fifth inequality, for every edge $e \in E$, $\theta_e(x)$ is multiplied by the sum of $x(\pi)$ over all the paths π using e.

Let ρ also denote the $|\mathcal{P}|$-dimensional vector corresponding to the flow ρ. Then, the π-th coordinate of $Q\rho$ is equal to the expected delay $\theta_\pi(\rho)$ on the path π, and the total latency of ρ is $C(\rho) = \rho^T Q \rho$.

Therefore, the problem of computing a feasible splittable flow of minimum total latency is equivalent to computing the optimal solution to the following quadratic program: $\min\{\rho^T Q \rho : \mathbf{1}^T \rho \geqslant W_{\text{tot}}, \rho \geqslant \mathbf{0}\}$, where $\mathbf{1}/\mathbf{0}$ denotes the $|\mathcal{P}|$-dimensional vector having $1/0$ in each coordinate. Also notice that no flow of value strictly greater than W_{tot} can be optimal for this program. This quadratic program is clearly feasible and its optimal solution is ρ^*_f (Lemma 7.1).

The Dorn's dual of this quadratic program is: $\max\{zW_{\text{tot}} - \rho^T Q \rho : 2Q\rho \geqslant \mathbf{1}z, z \geqslant 0\}$ (e.g., [10], [5, Chapter 6]). We observe that any flow ρ can be regarded as a feasible solution to the dual program by setting $z = 2\,\theta^{\min}(\rho)$. Hence, both the primal and the dual programs are feasible. By Dorn's Theorem [10], the objective value of the optimal dual solution is exactly $C(\rho^*_f)$[8].

Let ρ be any feasible flow at Nash equilibrium. Setting $z = 2\,\theta^{\min}(\rho)$, we obtain a dual feasible solution. By the discussion above, the objective value of the feasible dual solution $(\rho, 2\,\theta^{\min}(\rho))$ cannot exceed $C(\rho^*_f)$. In other words,

$$2\,\theta^{\min}(\rho)\, W_{\text{tot}} - C(\rho) \leqslant C(\rho^*_f) \tag{14}$$

Since ρ is at Nash equilibrium, $L(\rho) \leqslant \theta^{\min}(\rho) + \ell\, w_{\max}$. In addition, by Ineq. (11), the average latency of ρ cannot exceed its maximum latency. Thus,

$$C(\rho) \leqslant L(\rho)\, W_{\text{tot}} \leqslant \theta^{\min}(\rho)\, W_{\text{tot}} + \ell\, w_{\max}\, W_{\text{tot}}$$

Combining the inequality above with Ineq. (14), we obtain that $\theta^{\min}(\rho)\, W_{\text{tot}} \leqslant C(\rho^*_f) + \ell\, w_{\max}\, W_{\text{tot}}$. Using $C(\rho^*_f) = L(\rho^*_f)\, W_{\text{tot}}$, we conclude that $\theta^{\min}(\rho) \leqslant L(\rho^*_f) + \ell\, w_{\max}$. ∎

The following theorem is an immediate consequence of Lemma 7.3 and Lemma 7.2.

Theorem 8 ([15]) *The price of anarchy of any ℓ-layered network congestion game with resource delays equal to their loads, is at most* $8 \exp\left(\frac{\log m}{\log \log m} + 1\right)$.

A recent development which is complementary to the last theorem is the following which we state without a proof:

Theorem 9 ([16]) *The price of anarchy of any unweighted, single commodity network congestion game with resource delays $(d_e(x) = a_e \cdot x, a_e \geqslant 0)_{e \in E}$, is at most* $24 \exp\left(\frac{\log m}{\log \log m} + 1\right)$.

[8] More specifically, the optimal dual solution is obtained from ρ^*_f by setting $z = 2\theta^{\min}(\rho^*_f)$. Since $L(\rho^*_f) = \theta^{\min}(\rho^*_f)$ and $C(\rho^*_f) = L(\rho^*_f) W_{\text{tot}}$, the objective value of this solution is $2\theta^{\min}(\rho^*_f) W_{\text{tot}} - C(\rho^*_f) = C(\rho^*_f)$.

8 The Pure Price of Anarchy in Congestion Games

In this last section we overview some recent advances in the *Pure Price of Anarchy* (PPoA) of congestion games, that is, the worst-case ratio of the social cost of a PNE over the social optimum of the game.

The case of linear resource delays has been extensively studied in the literature. The PPoA wrt the total latency objective has been proved that it is $\frac{3+\sqrt{5}}{2}$, even for weighted multicommodity network congestion games [4, 7]. This result is also extended to the case of mixed equilibria. For the special case of identical players it has been proved (independently by the papers [4, 7]) that the PPoA drops down to 5/2. When considering identical users and single-commodity network congestion games, the PPoA is again 5/2 wrt the maximum latency objective, but explodes to $\Theta(\sqrt{n})$ for multicommodity network congestion games ([7]). Earlier it was implicitly proved by [15] that the PPoA of any weighted congestion game on a layered network with resource delays identical to the congestion, is at most 3.

9 Conclusions

In this survey we have presented some of the most significant advances concerning the atomic (mainly network) congestion games literature. We have focused on issues dealing with existence of PNE, construction of an arbitrary PNE when such equilibria exist, as well as the price of anarchy for many broad subclasses of network congestion games.

We highlighted the significance of allowing distinguishable players (ie, players with different action sets, or with different traffic demands, or both) and established some kind of "equivalence" between games with unit-demand players on arbitrary networks with delays equal to the loads and games with players of varying demands on layered networks.

Still, there remain many unresolved questions. The most important question is a complete characterization of the subclasses of games having PNE and admitting polynomial time algorithms for constructing them, in the case of general networks.

Additionally, a rather recent trend deals with the network-design perspective of congestion games. For these games [3, 2, 20] the measure of performance is the **price of stability**, ie, the ratio of the *best* NE over the social optimum of the game, trying to capture the notion of the gap between solutions proposed by the network designer to the players and the social optimum of the game (which may be an unstable state for the players). This seems to be a rather intriguing and very interesting (complementary to the one presented here) approach of congestion games, in which there are still many open questions.

References

[1] Ahuja R., Magnanti T., Orlin J. *Network Flows: Theory, Algorithms and Applications.* Prentice Hall, 1993.

[2] Anshelevich E., Dasgupta A., Kleinberg J., Tardos E., Wexler T., Roughgarden T. The price of stability for network design with fair cost allocation. In *Proc. of the 45th IEEE Symp. on Foundations of Computer Science (FOCS '04)*, page to appear, 295–304.

[3] Anshelevich E., Dasgupta A., Tardos E., Wexler T. Near-optimal network design with selfish agents. In *Proc. of the 35th ACM Symp. on Theory of Computing (STOC '03)*, pages 511–520, 2004.

[4] Awerbuch B., Azar Y., Epstein A. The price of routing unsplittable flow. In *Proc. of the 37th ACM Symp. on Theory of Computing (STOC '05)*, page to appear, 2005.

[5] Bazaraa M.S., Sherali H.D., Shetty C.M. *Nonlinear Programming: Theory and Algorithms (2nd edition)*. John Wiley and Sons, Inc., 1993.

[6] Beckmann M., McGuire C.B., Winsten C.B. *Studies in the Economics of Transportation*. Yale University Press, 1956.

[7] G. Christodoulou and E. Koutsoupias. The Price of Anarchy of Finite Congestion Games. In *stoc05*, pages to appear, available from `http://cgi.di.uoa.gr/~elias/publications/paper--ck05.ps`, 2005.

[8] Czumaj A., Krysta P., Voecking B. Selfish traffic allocation for server farms. In *Proc. of the 34th ACM Symp. on Theory of Computing (STOC '02)*, pages 287–296, 2002.

[9] Czumaj A., Voecking B. Tight bounds for worst-case equilibria. In *Proc. of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA '02)*, pages 413–420, 2002.

[10] Dorn W.S. Duality in quadratic programming. *Quarterly of Applied Mathematics*, 18(2):155–162, 1960.

[11] Even-Dar E., Kesselman A., Mansour Y. Convergence time to nash equilibria. In *Proc. of the 30th International Colloquium on Automata, Languages and Programming (ICALP '03)*, pages 502–513. Springer-Verlag, 2003.

[12] Fabrikant A., Papadimitriou C., Talwar K. The complexity of pure nash equilibria. In *Proc. of the 36th ACM Symp. on Theory of Computing (STOC '04)*, 2004.

[13] Feldmann R., Gairing M., Luecking T., Monien B., Rode M. Nashification and the coordination ratio for a selfish routing game. In *Proc. of the 30th International Colloquium on Automata, Languages and Programming (ICALP '03)*, pages 514–526. Springer-Verlag, 2003.

[14] Fotakis D., Kontogiannis S., Koutsoupias E., Mavronicolas M., Spirakis P. The structure and complexity of nash equilibria for a selfish routing game. In *Proc. of the 29th International Colloquium on Automata, Languages and Programming (ICALP '02)*, pages 123–134. Springer-Verlag, 2002.

[15] Fotakis D., Kontogiannis S., Spirakis P. Selfish unsplittable flows. In *Proc. of the 31st International Colloquium on Automata, Languages and Programming (ICALP '04)*, pages 593–605. Springer-Verlag, 2004.

[16] Fotakis D., Kontogiannis S., Spirakis P. Symmetry in network congestion games: Pure equilibria and anarchy cost. In *progress (submitted for publication)*, 2005.

[17] Gairing M., Luecking T., Mavronicolas M., Monien B., Spirakis P. Extreme nash equilibria. In *8th Italian Conference on Theoretical Computer Science (ICTCS'03)*. Springer-Verlag, 2003.

[18] Hochbaum D., Shmoys D. A polynomial approximation scheme for scheduling on uniform processors: Using the dual approximation approach. *SIAM J. Comput.*, 17(3):539–551, 1988.

[19] Hoeffding W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[20] Johari R., Tsitsiklis J. Efficiency loss in a network resource allocation game. *Mathematics of Operations Research*, page to appear, 2004.

[21] Johnson D., Papadimitriou C., Yannakakis M. How easy is local search? *J. of Computer and System Sciences*, 37:79–100, 1988.

[22] Koutsoupias E., Mavronicolas M., Spirakis P. Approximate equilibria and ball fusion. *ACM Transactions on Computer Systems*, 36:683–693, 2003.

[23] Koutsoupias E., Papadimitriou C. Worst-case equilibria. In *Proc. of the 16thAnnual Symposium on Theoretical Aspects of Computer Science (STACS '99)*, pages 404–413. Springer-Verlag, 1999.

[24] Libman L., Orda A. Atomic resource sharing in noncooperative networks. *Telecommunication Systems*, 17(4):385–409, 2001.

[25] Lücking T., Mavronicolas M., Monien B., Rode M. A new model for selfish routing. In *Proc. of the 21stAnnual Symposium on Theoretical Aspects of Computer Science (STACS '04)*, page to appear . Springer-Verlag, 2004.

[26] Lücking T., Mavronicolas M., Monien B., Rode M., Spirakis P., Vrto I. Which is the worst-case nash equilibrium? In *26th International Symposium on Mathematical Foundations of Computer Science (MFCS'03)*, pages 551–561. Springer-Verlag, 2003.

[27] Mavronicolas M., Spirakis P. The price of selfish routing. In *Proc. of the 33rd ACM Symp. on Theory of Computing (STOC '01)*, pages 510–519, 2001.

[28] Milchtaich I. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13:111–124, 1996.

[29] Monderer D., Shapley L. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[30] Papadimitriou C., Steiglitz K. *Combinatorial Optimization: Algorithms and Complexity.* Prentice-Hall, Inc., 1982.

[31] Rosenthal R.W. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[32] Roughdarden T., Tardos E. How bad is selfish routing? *J. Assoc. Comput. Mach.*, 49(2):236–259, 2002.

[33] Roughgarden T. The price of anarchy is independent of the network topology. In *Proc. of the 34th ACM Symp. on Theory of Computing (STOC '02)*, pages 428–437, 2002.

[34] Schäffer A., Yannakakis M. Simple local search problems that are hard to solve. *SIAM J. on Computing*, 20(1):36–87, 1991.

Bayesian Inference in densely connected networks applied to CDMA

Juan P. Neirotti and David Saad

The Neural Computing Research Group, Aston University, Birmingham, B4 7ET, UK

Graphical models provide a powerful framework for modelling statistical dependencies between variables [1]. Message passing techniques are typically used for inference in graphical models that can be represented by a sparse graph. Iterative message passing is guaranteed to converge to the globally correct estimate when the system is tree-like; there are no such guarantees for systems with loops.

Two inherent limitations seem to prevent the use of message passing techniques in densely connected systems: 1) Their high connectivity implies an exponentially growing computational cost. 2) The existence of an exponential number of loops that render the method inconsistent. However, a new approach was suggested [2] for extending Belief Propagation (BP) techniques to densely connected systems. In this approach, messages are grouped together, giving rise to macroscopic random variables drawn from a different Gaussian distribution of varying mean and variance for each node.

In a separate development [3], BP was extended to Survey Propagation (SP). This new algorithm has succeeded in solving hard computational problems [3], far beyond other existing approaches.

Inspired by the extension of BP to SP we have extended the approach of [2], designed for inference in densely connected systems, in a similar manner by including an average over multiple pure states. The method is general and could be applied to a number of problems. However, for showing the advantages with respect to the method [2], we apply it to the problem of signal detection in CDMA.

Code Division Multiple Access [4] is based on spreading the signal by using K individual random binary spreading codes of spreading factor N. We consider the large-system limit $N \to \infty$, $K \to \infty$ with $\beta = K/N \sim \mathcal{O}(1)$. The received aggregated, modulated and corrupted signal is of the form:

$$y_\mu = \frac{1}{\sqrt{N}} \sum_{k=1}^{K} s_{\mu k} b_k + \sigma_0 n_\mu$$

where b_k is the bit transmitted by user k, $s_{\mu k}$ is the binary spreading chip value, n_μ is the Gaussian noise variable drawn from $\mathcal{N}(0,1)$, and y_μ the received message. The goal is to get an accurate estimate of the vector $\mathbf{b}$ for all users given the received message vector $\mathbf{y}$ by approximating the posterior $P(\mathbf{b}|\mathbf{y})$.

A solution can be obtained by averaging over the various solutions, inferred from the same data, in a similar manner to the SP approach. Meanwhile, the messages in the current case are more complex.

Using Bayes rule one obtains the BP equations:

$$P^{t+1}(y_\mu|\mathbf{b}_k, \{y_{\nu\neq\mu}\}) \propto \sum_{\mathbf{b}_{l\neq k}} P(y_\mu|\mathbb{B}) \prod_{l\neq k} P^t(\mathbf{b}_l|\{y_{\nu\neq\mu}\}), \quad P^t(\mathbf{b}_l|\{y_{\nu\neq\mu}\}) \propto \prod_{\nu\neq\mu} P^t(y_\nu|\mathbf{b}_l, \{y_{\sigma\neq\nu}\}). \quad (1)$$

An explicit expression for the likelihood is required for deriving the posterior

$$P(\mathbb{B}\,|\,\mathbf{y}) = \frac{\prod_{\mu=1}^{N} P(y_\mu\,|\,\mathbb{B})}{\mathsf{Tr}_{\{\mathbb{B}\}} \prod_{\mu=1}^{N} P(y_\mu\,|\,\mathbb{B})}. \quad (2)$$

The likelihood is derived from the noise model (assuming zero mean and variance σ^2)

$$P(y_\mu\,|\,\mathbb{B}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(\mathbf{y}_\mu - \mathbf{\Delta}_\mu)^{\mathsf{T}}\, \mathbb{I}\, (\mathbf{y}_\mu - \mathbf{\Delta}_\mu)}{2\sigma^2} \right\}, \quad (3)$$

where $\mathbf{y}_\mu = y_\mu \mathbf{u}$, $\mathbf{u}^{\mathsf{T}} \equiv \overbrace{(1,\, 1,\, \cdots,\, 1)}^{n}$ and $\mathbf{\Delta}_\mu \equiv \frac{1}{\sqrt{N}} \sum_{k=1}^{K} s_{\mu k} \mathbf{b}_k$. Understanding the correlation between the replicated solutions is at the heart of the new approach. An explicit expression for the statistical

dependence between solutions is required for obtaining a closed set of update equations. We assume a dependence of the form

$$P^t\left(\mathbf{b}_k \mid \{y_{\nu\neq\mu}\}\right) \propto \exp\left\{ h^t_{\mu k} \sum_{\mathrm{a}=1}^{n} b^{\mathrm{a}}_k + \frac{1}{2n} g^t_{\mu k} \left(\sum_{\mathrm{a}=1}^{n} b^{\mathrm{a}}_k \right)^2 \right\} , \tag{4}$$

Using Eqs.(1), one can then calculate the expected value of b^{a}_k as a message passing algorithm [5]

$$\widehat{m}^{t+1}_{\mu k} = \left(\sigma^2 + \beta\left(1 - Q^t\right) + \beta R^t\right)^{-1} \left(\frac{y_\mu \mathbf{s}_\mu}{\sqrt{N}} - \beta\left(\mathbb{P}_\mu - \mathbb{I}/K\right)\mathbf{m}^t_\mu \right)_k \tag{5}$$

$$m^t_{\mu k} = \tanh\left(\sum_{\nu\neq\mu} \operatorname{artanh}\left(\widehat{m}^t_{\nu k}\right) \right) \qquad m^t_k \simeq \tanh\left(\sum_{\mu=1}^{N} \widehat{m}^t_{\mu k} \right) \tag{6}$$

where $\mathbb{P}_\mu \equiv (1/K) s_{\mu k} s_{\mu l}$, $\mathbb{I} \equiv \delta_{kl}$, $m^t_{\mu k}$ are the messages at time t from b nodes to y nodes and $\widehat{m}^t_{\mu k}$ are the messages at time t from y nodes to b nodes, respectively. The variables Q^t and R^t are related to the diagonal and off-diagonal elements of the covariance matrix of the macroscopic messages $\mathbf{\Delta}_\mu$.

The main difference between equation (5) and the equivalent equation in [2] is the emergence of an extra term in the prefactor, βR^t, reflecting correlations between different solutions groups (replica). This extra degree of freedom can be used to minimise the bit error probability at each time step using an EM like algorithm [6]. The bit error probability is the mean value of the discrepancy between the true bits sent (b_k) and the estimates ($\operatorname{sgn}\left(m^t_k\right)$). When the bit error probability is optimised with respect to R^t one obtains straightforwardly that $R^t = (\sigma_0^2 - \sigma^2)/\beta$. If the noise estimate is identical to the true noise, the term vanishes and one retrieves the expression of [2]; otherwise, an estimate of the difference between the two noise values is required for computing the prefactor of Eq. (5). Using the received signal to calculate the variance of the noise we can rewrite the update equation for $\widehat{m}_{\mu k}$, equation (5), as

$$\widehat{m}^{t+1}_{\mu k} = \left\{ \frac{1}{N} \sum_{\mu=1}^{N} y_\mu^2 - \beta Q^t \right\}^{-1} \left(\frac{y_\mu \mathbf{s}_\mu}{\sqrt{N}} - \beta \left(\mathbb{P}_\mu - K^{-1}\mathbb{I}\right) \mathbf{m}^t_\mu \right)_k \tag{7}$$

where no estimate on σ_0 is required.

This transforms the inference algorithm into a highly practical technique as it obviates the need for a prior belief of the noise level. The inference algorithm merely requires an iterative update of equations (7,6) and converges to a reliable estimate of the signal. The computational complexity of the algorithm is of $\mathcal{O}(NK^2)$ (reducing back to $\mathcal{O}(K^2)$ once the noise has been estimated).

To test the performance of our algorithm we studied the CDMA signal detection problem under typical conditions. Error probability of the inferred signals has been calculated for a system load $\beta = 0.25$, where the true noise level is $\sigma_0^2 = 0.25$ and the estimated value is $\sigma^2 = 0.01$, as shown in Figure 1(a). In this scenario we expect the original algorithm [2] to fail due to the discrepancy between the two noise levels (empty squares). The solid line represents the expected theoretical results, knowing the exact values of σ_0^2 and σ^2, while circles represent simulation results obtained via the suggested *practical* algorithm, where no such knowledge is assumed. The results presented are based on 10^5 trials per point and a system size $N = 2000$ and are superior to those obtained using the original algorithm [2].

Another performance measure one should consider is $D^t \equiv \frac{1}{K}\left|\mathbf{m}^t - \mathbf{m}^{t-1}\right|^2$. It provides an indication to the stability of the solutions obtained. The inset of Figure 1(a) shows that the results obtained using our algorithm (circles) converge to a reliable single solution in stark contrast to the results obtained by the original algorithm (squares) [2]. The physical interpretation of the difference between the two results is related to a replica symmetry breaking phenomena.

To analyse the critical properties of the system we studied the asymptotic regime of the error per bit probability P_b, for different values of σ_0^2 (Figure 1(b)). For low values of the noise variance

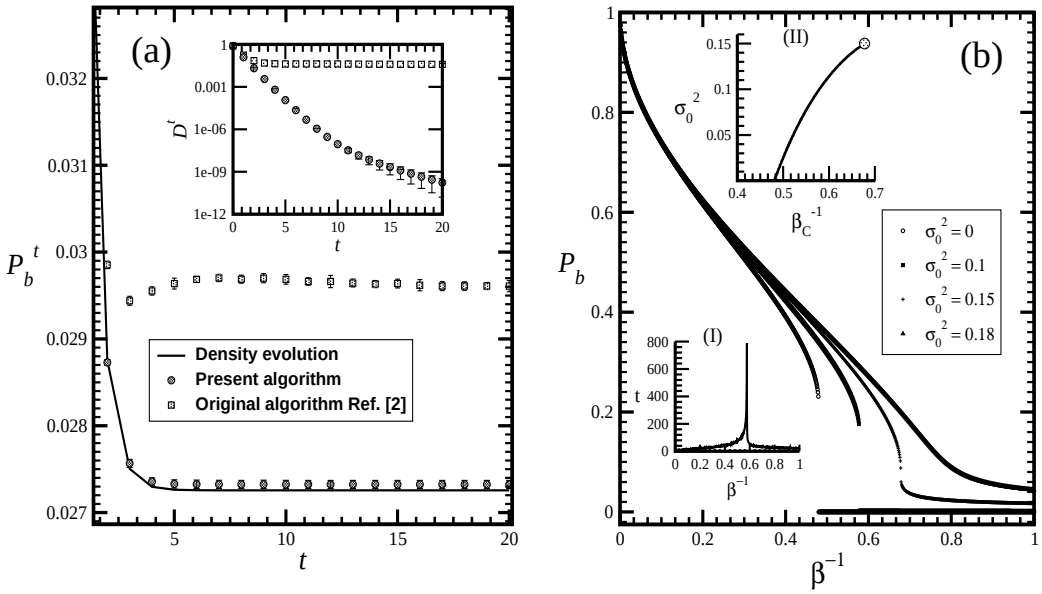


Figure 1: (a) Error probability of the inferred solution as a function of time. The system load $\beta = 0.25$, true and estimated noise levels $\sigma_0^2 = 0.25$ and $\sigma^2 = 0.01$, respectively. Squares represent results obtained by the original algorithm [2], solid line the dynamics obtained from our equations; circles represent results obtained from the suggested *practical* algorithm. Variances are smaller than the symbol size. In the inset we present the measure of convergence D of the obtained solutions, as a function of time. (b) Error probability as a function of β^{-1} for several values of σ_0^2. Below $\sigma_0^2 \simeq 0.15$ all curves are discontinuous for a critical value β_C^{-1}. The number of iterations needed to reach the steady state at the criticality diverges, as it is shown in inset (I) for $\sigma_0^2 = 0.1$. For values of $\sigma_0^2 > 0.15$ all curves are analytical. The critical line is presented in inset (II).

and large values of β, P_b is a decreasing function of β^{-1}. For a given value of $\beta_C\left(\sigma_0^2\right)$ there is a discontinuity in P_b. At this point, the number of iterations needed to reach the steady state diverges (Figure 1(b), inset (I)). The last noise value for which there is a non-analytical behaviour in P_b is $\sigma_0^2 \simeq 0.15$. At this value of the noise parameter, P_b becomes a continuous curve with a singularity in its first derivative. The critical line is presented in Figure 1(b), inset (II).

In summary, we provided a new message passing method for inference in densely connected systems, that delivers an improved detection algorithm for CDMA and could be applicable to a number of problems. We also provide an analysis of the critical properties of the multi-user detection problem by studying the properties of the asymptotic regime of the error per bit probability.

Acknowledgement

Support from EVERGROW, IP No. 1935 in FET, EU FP6 is gratefully acknowledged.

References

[1] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann Publishers, San Francisco, CA (1988)
[2] Y. Kabashima, J. Phys. A **36** 11111 (2003)
[3] M. Mézard, G. Parisi and R. Zecchina, Science **297** 812 (2002)
[4] S. Verdú, *Multiuser Detection*, Cambridge University Press UK (1998)
[5] J. P. Neirotti and D. Saad, Europhys. Lett. *in press* (2005).
[6] A. P. Dempster, N. M. Laird and D. B. Rubin, J. Royal Stat. Soc. B **39**, 1 (1977). .

On emergent phenomena in everyday activities taking place in AmI spaces

Ioannis D. Zaharakis[1], Achilles D. Kameas[1,2]

[1] Computer Technology Institute,
Research Unit 3, Design of Ambient Information Systems Group
N. Kazantzaki str, University Campus, GR-26500 Patras, Hellas

[2] Hellenic Open University
23 Sahtouri str, GR-26222 Patras, Hellas

Email: {jzaharak, kameas}@cti.gr

Corresponding author:

Dr. I. D. Zaharakis
Computer Technology Institute
Research Unit 3, Design of Ambient Information Systems Group

Email: jzaharak@cti.gr

Address:
N. Kazantzaki str,
University Campus,
GR-26500 Patras,
Hellas

Tel.: +30 2 610 960200
Fax: +30 2 610 960490

On emergent phenomena in everyday activities taking place in AmI spaces

Ioannis D. Zaharakis[1], Achilles D. Kameas[1, 2]

[1] Computer Technology Institute
Research Unit 3, Design of Ambient Information Systems Group

[2] Hellenic Open University, Hellas

Email: {jzaharak, kameas}@cti.gr

Abstract

This work aims at a) identifying the forthcoming changes in our everyday life due to the ever-increasing level of complexity that inculcates our interactions with the devices surrounding us, b) introducing a bio-inspired world model (framework) that deals with different perspectives of the interrelations developed in symbiotic ecologies where people and artefacts coexist, and c) proposing a high level architectural scheme of an AmI space reflecting the basic ingredients of the future indoors/outdoors applications based on Swarm Intelligence and Complexity Science.

Keywords: Complex Systems, Emergent behaviour, Ambient Intelligence, Swarm Intelligence, Ubiquitous Computing applications

Introduction

The vision of Ambient Intelligence (AmI) implies a seamless environment of computing, advanced networking technology and specific interface ([1], [2]). In one of its possible implementations, technology becomes embedded in everyday objects such as furniture, clothes, vehicles, roads and smart materials, and people are provided with the tools and the processes that are necessary in order to achieve relaxing interactions with this environment. The AmI environment can be considered to host several Ubiquitous Computing (UbiComp) applications, which make use of the infrastructure provided by the environment and the services provided by the AmI objects therein. An important characteristic of AmI environments is the merging of physical and digital space (i.e. tangible objects and physical environments are acquiring a digital representation); hence, the AmI objects differ from traditional objects in that they can communicate with other AmI objects and can interact with the environment. Of special interest is the information that AmI objects process, which can be descriptions of the context of use, data to be used for a task, guidelines on how to perform a task, messages to be sent or that have been received from other objects. The result of information processing is a set of services, that is, a set of abilities that appear in the digital space. This work builds upon the envisaged structure of AmI environment as one populated by thousands of communicating tangible objects and virtual entities. At a minimum, AmI environment will contain network infrastructure and it will be available making anytime, anyplace (within boundaries of acceptability) interaction among and with these objects feasible.

Research issues and requirements

The heterogeneity of AmI objects makes necessary the development of middleware systems on top of which UbiComp applications can function transparently with respect to the infrastructure [3]. The dynamic nature of UbiComp applications and the mobility of AmI objects, force the middleware to use services and capabilities with changing availability [4]. To ease the development of such applications it is necessary to decouple application composition from context acquisition and representation, and at the same time provide universal models and mechanisms to manage context [5].

Key research challenges have to focus in services availability including both services aimed at end users as well as machine to machine services, and to deal with dynamic composability and adaptability, context awareness, autonomy and semantic interoperability. Essentially, new research issues arise concerning i) the system complexity emerging by the thousands local interactions between people and artefacts, ii) the need for flexible and dynamic system architecture capable to evolve and adapt to new situations and configurations, iii) the context dependence of the exchanged information, and iv) the human involvement and especially new, more natural, human-machine interaction schemes. These features make necessary to the development of a framework that will help us deal with the complexity of using UbiComp applications within AmI environment. The framework must be capable to reflect the available services as well as the potential use of the participating objects. Although the available services may somehow be exhibited, the potential use of the objects emerges mainly from the interactions of the humans with the devices and these interactions are not only time-dependent but also space- or context-dependent.

The next sections set the scene of a near future everyday living/working environment and describe an engineering approach inspired by biological structures capable to deal with phenomena arising in such an environment.

Proposed conceptual framework

A living/working AmI space comprising of many heterogeneous objects with different capabilities and provided services could be considered that is populated by a *heterogeneous swarm*. The swarm will comprise different typologies of societies, and so it will be heterogeneous also from the provided services point of view. Such differences will contribute to the overall capabilities of the system. As a general principle, the services should be as simple as possible. Ideally, the composition should emerge based on previous interactions and on the context (time and place) they took place.

Traditional Artificial Intelligence (AI) focused on addressing intelligence as an individual phenomenon considering (intelligent) agents with cognitive states which maintain a (partial) model of the world they inhabit in and a (partial) model of the others. A radically different approach is based on the belief that intelligent behaviour is inextricably tied to its cultural context and cannot be understood in isolation. Indeed, many natural systems can be described in terms of many individually “simple” components, interacting in “simple” ways and influencing their neighbours, and yet, are able to exhibit “complex” overall system level behaviour; those systems that exhibit this “emergent” globally complex behaviour from simple components are referred to as “complex systems” [6]. In contrast to traditional AI, Swarm Intelligence (SI) is defined as the emergent collective intelligence of groups of simple (unsophisticated) agents and, as an engineering approach, offers an alternative way of designing intelligent systems, in which autonomy, emergence, and distributed functioning replace control, pre-programming, and centralization ([7], [8]).

Typically, when focusing on situated social systems in dynamic and non-deterministic environments, it is very hard (if not aimless) to embody into each organism complete models of the environment and of the others. Alternatively, no explicitly represented world models could be considered ([9], [10]). All the necessary information is out there changing dynamically; thus the world is the model itself. All we need is the means to capture, qualify and exploit the information that surrounds us. In order to deal with the collective behaviour of large ecologies in situated domains, a recent approach is the analysis and synthesis of small pieces of primitive behaviours that result from individual interactions.

Inspired by the biological social systems (ecologies), the analysis of artificial swarm systems could range in different levels depending on the desired granularity, e.g., as single behaviour building blocks [11], as neural networks that can learn and evolve [12], as sensor networks owning limited power, computational capacities and memory [13], etc. Independently of the analysis level, the computational capabilities (and the intelligence) of the

ecology are distributed over the central "nervous" system, the peripheral system, the materials of the ecology's body and the physical phenomena created by the interaction of the ecology with its environment. Putting such entities into a UbiComp environment could lead to extelligent ecologies, where knowledge and tools are diffused in the environment [14], underlying thus the corporal literacy of the ecology, meaning the awareness of the extelligence and the working knowledge of all senses. This will pave the way for the generation of theory and technology of synthetic phenomenology (of the resulting ecology) meaning the understanding of the own self and its relation with the surrounding world. Drawn from the above, the proposed high level architectural scheme that consistently reflects bio-inspired self-aware emergent symbiotic AmI space ecologies consists of the following fundamental ingredients (Figure 1): i) *basic building blocks* including sensors, hardware resources, software modules, artefacts, etc., ii) *ecologies*, that are groups of building blocks, their interrelationships and the associated environment, iii) *ambient knowledge*, that provides a means to emergent consensus as a substitute for social and cultural memory, and iv) *people* divided into the classes of developers (building block developers, hardware designers, artefact manufacturers, application developers, etc.) and end-users.

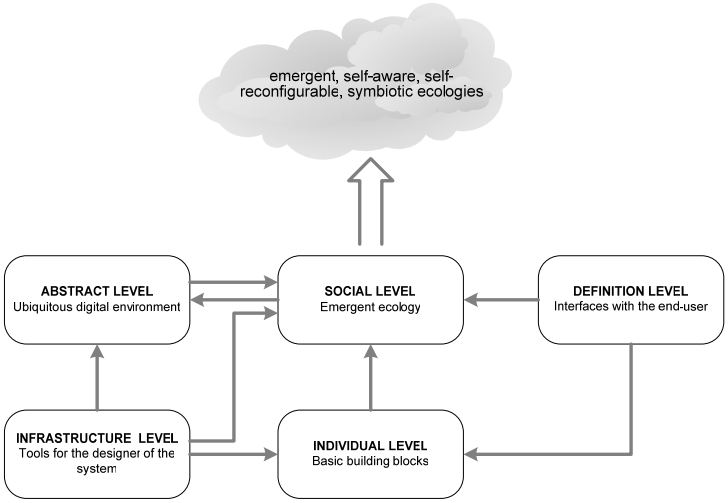


Figure 1. *High level architecture of the system. Individual and Social levels correspond to the basic building blocks and ecologies, respectively. Abstract level encloses the social memory of the ecologies; such knowledge must be transferred to the ecologies implicitly e.g. as stimuli of the environment, since individuals and consequently the emergent ecologies do not contain any knowledge representation scheme neither reasoning mechanism. Infrastructure level provides system designers with the appropriate tools to develop a system. Definition level is the user interface with the final user.*

Engineering emergent phenomena

An especially complex task is to model and build autonomous interactive entities that could form extelligent ecologies exhibiting corporal literacy and leading to a synthetic phenomenology approach. The task is additionally complicated by considering that the resulting ecologies will operate into a ubiquitous environment and will be driven by autonomy, local perceptions and interactions, emergence, and distributed functioning. An important aspect on this focus is that although the entities will not have explicitly represented models of the world or of the others the emergent ecologies will unfold coherent collective behaviour based only on the entities' own agenda of actions and their intrinsic inclination to preserve their own goals.

Realizing the potential benefits of the UbiComp applications populated by simple autonomous entities will require improvements in currently available technology platforms and a translational research paradigm from basic-research findings. Hence the driving force behind the whole idea focuses on the adaptation of concepts from complex biological systems and novel fabrication technology platforms to build truly innovative swarm artificial systems

for emerging real-life applications. Technological challenges posed by this approach and the investigative methodology to overcome them are described below.

1. *Basic building blocks development*

Several levels of abstraction are possible for the formulation of the basic building block. Immediately, the engineer strives with the questions on i) which should be the basic building block, ii) what structural and functional properties it should encompass, iii) how it could interact with the others, and iv) how it could be realised. From a technology development point of view, an essential plan is needed which initially centres about the basic building block and considers as such every self-sustained digital (h/w or s/w) artefact with certain functionality that can operate without the contribution of others.

2. *Engineering emergent behaviour*

In dynamic environments, an individual must be reactive, that is, it must be responsive to events that occur in its environment, where these events affect either the individual's goals or the assumptions, which underpin the procedures that the individual is executing in order to achieve its goals [15]. In order to apply the well-established primitive behaviours approach in swarm societies that can learn and evolve component-oriented principles and practices could be employed. Synthetic behaviour control mechanisms could be developed based on bio-inspired approaches like spiking neural networks. These behaviour control mechanisms responsible for the arbitration and/or the composition of the primitive behaviours could also be subject of learning and evolution. The individuals may exhibit varying behaviour – capable of perceiving/exploring their environment, selectively focusing attention, initiating and completing several tasks. The learning and evolution could be studied and investigated at both the individual and social levels. In this case, the focal point must be the components of behaviour control mechanisms. The outcome could contribute to a novel dynamic and adaptive architecture of swarm systems that exploits the global effects through local rules/behaviour.

3. *Engineering collective behaviour*

Developing a robust swarm system, capable of exhibiting emergent intelligent collective behaviour is a non-trivial task. The nature of the social/collective behaviour sought and the environment that allows efficient development of ecologies requires research. In building a swarm system communication plays a pivotal role and this explains the profuse number of publications in this area. A flexible and light-weight approach is the indirect (stigmergic) communication. The essence of stigmergy is that the individual modifies a local property of the environment, which subject to environmental physics, should persist long enough to affect the individual's behaviour later in time. It is the temporal aspect of this phenomenon, which is crucial for emergent collective behaviour (collaborative exploration, building and maintenance of complex insect nest architectures etc) in societies of ants, agents and robotics. Thus, the individuals could be provided with the proper periphery (actuators/sensors) enabling them to emit/perceive electromagnetic signals emulating thus the biological "quorum sense" signals. Such a quorum sense communication may be based on an application-specific vocabulary that will be encoded in the signal. The specifics of the temporal modulation aspect of this "quorum sense signal" will come from theoretical biology and existing simulation studies.

Summary

As everyday objects are being enhanced with sensing, processing and communication abilities, the near future of our everyday living/working is indicated by a high degree of complexity. The emergent complexity concerns the machine-machine and human-machine interactions as well as the provided services aimed at end users and at other machines. Into this rapidly changing AmI environments new requirements and research issues arise, and the need for a conceptual and analysis framework is apparent. This work attempts to introduce a

bio-inspired word model that draws features from natural systems and applies them into symbiotic ecologies inhabited by both humans and artefacts. Furthermore, it introduces a high-level architecture of AmI spaces that encloses the fundamental elements of bio-inspired self-aware emergent symbiotic ecologies.

References

[1] Disappearing Computer initiative: http://www.disappearing-computer.net/

[2] ISTAG website: http://www.cordis.lu/ist/istag.htm

[3] Kameas A, Bellis S, Mavrommati I, Delanay K, Colley M, Pounds-Cornish A: An architecture that treats everyday objects as communicating tangible components, in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications,* (*PERCOM2003*), Texas-Fort Worth, 2003.

[4] Kameas A, Mavrommati I, Markopoulos P: Computing in tangible: using artifacts as components of Ambient Intelligent Environments. In *"Ambient Intelligence: The evolution of Technology, Communication and Cognition",* (G. Riva, F. Vatalaro, F. Davide and M. Alcaniz, eds), IOS press, pp 121-142, 2004.

[5] Christopoulou E, Goumopoulos C, Zaharakis I, Kameas A: An Ontology-based Conceptual Model for Composing Context-Aware Applications, *Workshop on Advanced Context Modelling, Reasoning and Management, Sixth International Conference on Ubiquitous Computing* (*UbiComp04*), Nottingham, England, 7-10 September 2004.

[6] Bullock S, Cliff D: *Complexity and emergent behaviour in ICT systems*, HP Labs report HPL-2004-187, 2004.

[7] Bonabeau E, Dorigo M, Theraulaz G: *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999.

[8] Kennedy J, Eberhart R: *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.

[9] Brooks R A: Intelligence without Reason. In *Proceedings of 12th Int. Joint Conf. on Artificial Intelligence*, Sydney, Australia, August 1991, pp. 569—595, (1991).

[10] Brooks R A: Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

[11] Lund H H, Marti P: Physical and Conceptual Constructions in Advanced Learning Environments. *Interaction Studies Journal,* 5(2): 271-301, 2004.

[12] Lund H H: Neural Building Blocks. In *1st International IEEE EMB Conference on Neural Engineering*, Capri, Italy, March 20-22 2003. IEEE Press.

[13] Akyildiz I F, Su W, Sankarasubramaniam Y, Cayirci E: A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114, 2002.

[14] Stewart I, Cohen J: *Figments of Reality: the origins of the curious mind.* Cambridge University Press, 1997.

[15] Wooldridge M: Intelligent Agents, in G. Weiss (ed.) *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, MIT Press, pp. 27-77, 1999.

Cognition Modelling

TOWARDS AN ECONOMIC THEORY OF MEANING AND LANGUAGE

Gábor Fáth and Miklos Sarvary*

Abstract: We present a model in which abstract concepts of a language acquire meaning as the result of competition between heterogeneous interacting agents in a community. We argue that bounded rationality requires individuals to use a reduced number of abstract concepts to represent the rich reality of the world. The meanings of these concepts emerge as a trade-off between two objectives: (i) agents want to use concepts that are best adopted to their idiosyncratic preferences and characteristic distribution of choice alternatives, (ii) agents seek to share concepts to benefit from communication. Agents play a non-cooperative game, whose Nash equilibrium determines the collective meanings of concepts in the population, constituting together the community's language. Analysis of the possible Nash equilibria and the evolutionary game dynamics shed light on interesting theoretical questions such as the origins of meaning, the coherence of language, the language-culture relationship, and Whorf's hypothesis on linguistic relativism.

Keywords: evolution of language, mental representation, bounded rationality, game theory, Nash equilibrium, dimension reduction, PCA, agent heterogeneity

1 Introduction

Human language is a unique trait that clearly sets us apart from animals. Part of the human language system is biological, i.e., hard wired in us by millions of years of evolution. Speech organs, for instance, clearly belong to this category. Other parts of the language system are the result of more or less conscious cognitive processes such as learning, or deliberate social interactions. The creation of new words by a community is the result of complex interactions between its members. The subject of this paper is language in this latter sense, i.e., language as a conscious/social process.

Roughly speaking, language consists of words and rules. Words are linguistic signals referring to concepts, the collection of which constitutes the so-called mental dictionary or mental lexicon. We use the term "word" to denote a *listeme*: any string of linguistic elements (e.g., morphemes, words, or composite expressions) that is associated with a particular meaning. (In this general sense idioms, for instance, are also listemes.) This paper focuses on the emergence of the mental dictionary, i.e., it seeks to explain how words acquire meaning.

*Gábor Fáth is a Senior Research Fellow at the Research Institute for Solid State Physics and Optics of the Hungarian Academy of Sciences, P. O. Box 49, Budapest, 1525, Hungary, Tel: +36 1 392 2222, E-mail: fath@szfki.hu. Miklos Sarvary is Associate Professor of Marketing at INSEAD, Bd. de Constance, 77305, Fontainebleau, France, Tel: +33 1 60 71 26 05, E-mail: miklos.sarvary@insead.edu. This work was supported by INSEAD Foundation. Gábor Fáth also acknowledges support from the Hungarian Scientific Research Fund (OTKA) under grant Nos. T43330 and T47003 and by the Hungarian National R&D Program (NKFP) under grant No. 2/051/2004.

Our central thesis is that meanings emerge in language through an *economic* process, in the sense of competition between agents. We study intelligent maximizing agents with well-defined preferences facing a series of decisions. Agents have an overwhelming amount of information (physical perceptions) concerning their decision alternatives. However, with restricted cognitive resources and limited communication bandwidth, agents have to use a reduced number of abstract concepts (words) to represent the rich reality of the world. In this environment, agents' "choices" (of the meanings) of concepts is driven by the objectives of keeping representation error minimal and communication effective. The former means using concepts that provide the "best" description of the world in the sense of leading to optimal decisions given the agents' preferences. Since preferences may differ across agents, optimal concepts will differ too. However, as agents want to communicate with each other (e.g., to gather information from each other or to cooperate) they also need to share their representations (concepts). Trading off these two objectives (representation and communication) agents essentially play a non-cooperative game (Language Game), whose Nash equilibrium is a set of partially shared meanings that we interpret as "language". We are interested in the conditions under which such an equilibrium exists, as well as the level of coherence in the resulting words/concepts. We analyze the properties of the Language Game's evolutionary dynamics, and discuss the multiplicity and stability of its equilibria. Our ultimate goal is to understand the emerging mapping between agents' perceptions and the (abstract) concepts (words) of language. In particular, an important question we investigate is how individual preferences influence agents' mental representations of the world.

The analytic results show that the Language Game always has at least one Nash equilibrium, although it may not be unique. In the resulting language(s), concepts have partial coherence across agents, the strength of which is an increasing function of the exogenous benefit from communication. However, although the average coherence of concepts across agents increases monotonically with the strength of communication, the coherence of concepts can differ significantly. Analysis of the Language Game's evolutionary dynamics shows that it always has at least one fixed point that is also a Nash equilibrium. This equilibrium is also dynamically stable against small collective deviations (an attractor). However, not all attractors of the Language Game's dynamics are Nash equilibria and not all Nash equilibria are dynamically accessible. The number of potential equilibria increases with the number of agents. Which concepts emerge in equilibrium is sensitive to initial conditions, i.e., language exhibits strong path-dependence. These patterns have interesting implications for the interpretation of natural languages, their differences and relative stability across cultures and the psychology of mental representations.

2 Relevant Literature

There are two literature streams that formally model the language formation process and, in particular, the endogenous emergence of meaning in certain signals. In the "cheap-talk" literature of rational game theory [1, 2, 3], agents are properly defined in terms of preferences and interact strategically in a specific game with incomplete information. The purpose of language construction is to better coordinate on an efficient outcome of this game. It is shown that under certain conditions (non-costly) signals become associated with specific meanings and the use of the so-constructed language helps coordination. In models of evolutionary biology [4, 5, 6, 7] agents are boundedly

rational and use different heuristics. The association of specific signals to given meanings emerge through imitation or reproduction. It is shown that a shared mental lexicon (mapping between signals and objects of reality) can dynamically emerge if heuristics leading to successful communication have a higher chance to survive.

Our framework differs from these literature streams in several ways. First, in contrast with existing models, in our case, the social process of language construction involves explicit conflict of interest between agents. We assume that language has a strong influence on decision making (Rubinstein [8], Chapter 4). In fact, optimal decision making - in the sense of optimal choice between alternatives - is the driving force behind language formation. Language is used to describe choice alternatives with a limited number of concepts. As a result, agents with different preferences may not want to attribute the same meanings to the concepts. Yet, if they want to communicate with other members they have to largely agree on the meanings. These, potentially opposing objectives result in a competitive game, that we call Language Game. A related issue is that traditional evolutionary approaches assume that the language construction process is "unconscious", in the sense that agents' heuristics in associating specific signals to given objects survive through imitation or reproduction. In other words, in these models agents do not actively promote certain meanings to be used. In reality however, communities do construct their language in every day's life and their members do so consciously. They add words to the mental lexicon and modify the meanings of existing ones. Consistently with this picture, in our framework, agents are strategic in influencing the formation of a common language.

Second, we seek to build a *general* model of language, in the sense of language being context-independent. In other words, we are interested in the emergence of concepts/words that are used across a large number of different contexts. In existing evolutionary approaches, concepts do not really emerge endogenously. Rather, concepts and signals are given and the issue is to reach agreement over the mental lexicon providing the mapping between them [9, 10]. In other words, these models do not explicitly address *which* concepts should be named by signals (words) and why concepts may be similar or different across agents and/or communities. As Rubinstein [8] points out in his reflections on language, even in cheap-talk games, the relevant concepts are already defined by the underlying game. As such, the messages are limited to certain well-defined objects. Rubinstein argues [8]: "A persuasive explanation of the emergence of linguistic concepts requires a much more general setting" (p. 34.). Our goal is precisely to model such a general (context-independent) setting.

Two recent papers are closest to our work in addressing the endogenous emergence of meanings with conflict of interest between agents. Battigalli and Maggi [11] construct a model of language to build a theory of contract incompleteness. A contract uses language to partition the set of events and associate it to the contracting parties' obligations. A more "precise" - hence more costly - language results in more complete contracts. The paper explains various forms and degree of contract incompleteness by the cost of the language used. In a different formulation, Cremer *et al.* [12] develop a model where parts (departments) in an organization use different partitions of the space of signals to develop an internal code adapted to each department's objectives (see also, Wernerfelt [13]). Conflict between departments may arise when internal codes differ. The focus of the paper is the organizational structures that emerge as a result of such conflict or the need to make conflict disappear.

Our work has several points of departure from these papers. Instead of contracts and

organizational structures, our focus is the process of language formation itself and the features of the resulting equilibrium languages. In this respect, in our framework, the meanings of concepts are not discrete in a language community. As in reality, people do not completely agree or completely disagree with each other on the meaning of a word and this does not (completely) prevent communication. Words may have somewhat different meanings for different people even if there is rough agreement between them. Existing approaches do not allow for this eventuality. In contrast, our model captures this continuous aspect of meaning in the mental dictionary at the outset and - as will become clear later - even in equilibrium. In fact, an interesting question we ask is: to what extent is there agreement between agents on the meaning of words in a language?

Our model structure also differs fundamentally from the above papers'. In particular, we do not define words as a partition of the state space, but rather as weighted averages of the states. As will become clear later, this allows us to make a distinction between *language* and *culture*, that existing models have a hard time to incorporate.

Finally, our model is closely related to psychology, which is concerned by the mental representation of concepts and language's link to decision making in general. Psychologists broadly see mental representations as a result of a clustering procedure that gives birth to hierarchies of categories with more abstract concepts belonging to higher level categories [14, 15]. Objects and concepts are shown to belong to categories to various degrees, ranked by how "typical" members they are of the given category. More recent research in categorization acknowledges that such categorical structures are "ad hoc" and are closely linked to decision making in that they are defined by the decision maker's goal [16, 17]. Our framework is consistent with this view, in the sense that "categories" emerge endogenously in a language equilibrium, partially guided by decision making. Specifically, concepts are assumed to be weighted averages of the elementary (microscopic) signals that agents perceive about the world (we assume that these perceptions are identical across agents). Naturally, the optimal weighting scheme will reflect similarities and contrasts between the signals that we can associate with the "objective" structure of the world. However, the weights will also reflect the agent's preference structure, which we assume to be idiosyncratic, i.e. heterogenous across agents. In sum, our framework explicitly describes how mental representations are influenced by three fundamental "inputs" for language: (i) the correlation structure of the physical world, (ii) agents' (decision makers') individual preferences and (iii) the rate of communication. In this way, it also sheds light on numerous debates in psycholinguistics, such as the well-known Whorf hypothesis on linguistic relativism or the relationship between culture and language.

3 The Language Game

We define "language" as the Nash equilibrium of the Language Game. In the Language Game, agents choose a finite set of concepts (words) trading off two objectives: (i) to best "represent" their vast perceptions (available data) about the world and (ii) to share the concepts across each other, i.e., to associate - to the extent possible - the same meaning to any given word.

So defined, the Language Game has three important characteristics. First of all, language should provide an *accurate representation* of the complex world. The underlying assumption is that agents regularly face decisions, that is, choices between alternatives. Efficient decision making requires that choice alternatives be evaluated as accurately

as possible. The structure of language providing the necessary "mental representation" stems from two principal sources. Language needs to reflect the objective structure of the natural world, and at the same time, the subjective structure of the agents' preferences. For example, a hiker may prefer to use different symbols to draw the map of a region than a botanist, not because they perceive nature differently, but because they would like to distinguish different objects: while the hiker only needs to tell apart forest from clearing, the botanist needs to represent the dominant types of plants. Conversely, the botanist may be less interested in the details of the terrain's morphology. Their preferences are different.Traditionally, cognitive sciences tend to focus on how perceptions of reality influence the development of languages. *In contrast, in this paper, we are primarily concerned with the effect of agents' heterogenous preferences.*

The second characteristic is that the complex perceptions of agents are to be represented by a *finite set of optimally chosen concepts*. This is motivated by our assumptions that agents are boundedly rational and their communication bandwidth is limited. Both of these drive agents (in an evolutionary sense) towards a well-organized and compressed representation of the world. Bounded rationality, which can manifest itself in the form of limited memory space, time constraints on learning the correct use of concepts, or rapidly increasing processing cost as a function of model complexity, cannot let the number of concepts (words) to rise too high. Similarly, biology limits human communication bandwidth to a relatively low rate of words per second. Thus, efficient communication also requires the compression of information along a small number of highly significant concepts. Therefore, the typical size and organization of the human mental lexicon arises from the balance between two competing goals: to increase representational accuracy and decrease model complexity (see Chater and Vitnyi [18] for a short introduction on the importance of model simplicity). In order to simplify our model we will assume that the number of concepts in the language is fixed, and will only concentrate on the optimal choice of these concepts.

Finally, the third characteristic of a Language Game is that agents need to align their concepts if they want to profit from *communication*. The underlying, natural assumption is that the more concepts are similar (words have similar meanings), the more they are useful to transfer information about the state of the world (choice alternatives, etc.) between agents. We will assume that society imposes an exogenous "pressure" to communicate, which will contribute to agents' utility according to the actual alignment of their concepts. Notice, that if agents are different in terms of preferences, then the different aspects of language (accurate mental representation and communication) put agents in conflict with one another. A priori, each agent would prefer to use a different language for efficient decision making. Since agents also want to benefit from communication, it is in their interest to deform the collectively shared meanings in such a way that these better reflect their own preferred view of the world. Language, defined as the collection of concepts, emerges as an equilibrium of the Language Game.

3.1 Mental representation with a finite number of concepts

We consider I agents, each restricted to use a number K of concepts only. Agents divide the world into a number X of *decision contexts* in which they evaluate decision alternatives according to their personal preferences. We assume that alternatives are characterized by their objective (physical) attributes $\boldsymbol{a} = \{a_1, \ldots, a_D\}$, that are common knowledge. In other words, agents are homogenous in terms of their perceptions about

the world. Agent i's *objective payoff* from choosing alternative $\boldsymbol{a}$ in context x is denoted by $\pi_i^{(x)}(\boldsymbol{a})$. In the following, we assume for simplicity that the objective payoff is a *linear* function of the attributes

$$\pi_i^{(x)}(\boldsymbol{a}) = \boldsymbol{\omega}_i^{(x)} \cdot \boldsymbol{a}, \tag{1}$$

in which the vector of coefficients $\boldsymbol{\omega}_i^{(x)}$ will be called the agent's *preference vector in context* x. The collection of vectors $\{\boldsymbol{\omega}_i^{(x)}\}_{x=1}^{X}$ defines the agent's objective (biological or otherwise acquired) preferences in all possible decision contexts. These parameters are assumed to be fixed in the model. In other words, each agent's preference structure is described by a very large number XD of fixed parameters. As opposed to $\boldsymbol{a}$, the parameters $\{\boldsymbol{\omega}_i^{(x)}\}_{x=1}^{X}$ are agent specific, i.e., we consider *agent heterogeneity in preferences*.

If they knew their preference vectors, the agents would be able to calculate the objective payoffs of their decision alternatives in the assumed linear world, and make the best possible decisions whenever they face a decision problem. However, although the preference vectors $\boldsymbol{\omega}_i^{(x)}$ are well defined in theory, we have good reason to suppose that the numbers they represent are not directly available for the agents. On the one hand, the agents need a good deal of experimentation (learning) to find out how a given attribute impacts their payoffs. As an example think about food allergy, which may be an innate condition in the patient, but requires a painful "learning process" through bad choices of food alternatives to recognize and diagnose. When the number of attributes and contexts is large, there is a problem with the sparsity of the "training samples" (cf., the "powerty of stimulus" argument in linguistics [18]). On the other hand, a perfect knowledge of $\boldsymbol{\omega}_i^{(x)}$ would require a detailed understanding of the effect of all physical attributes on the payoffs: a total number of XD parameters in the linear model. We can assume that this is beyond the agent's mental capacity, or the required cognitive complexity is too costly for the agent.

What remains for the agent to reduce the cost of complexity and to have a model which better generalizes from sparse data is to try to invent a simplifying scheme, a so called *mental representation*. The mental representation is an *approximate* mapping from microscopic attributes to payoffs. Given a decision alternative $\boldsymbol{a}$ and using the mental representation, the agent arrives at an *approximate payoff*, $\tilde{\pi}_i^{(x)}(\boldsymbol{a})$, which is not equal but close to the exact payoff $\pi_i^{(x)}(\boldsymbol{a})$. We assume that there exists some learning mechanism, which improves the agent's mental representation by collecting experience on previous choices and their success rate. Eventually, as a result of learning, the approximate payoffs get as close to the objective ones as allowed by the structural constraints of the mental representation. The mental representation gets optimized within its limits.

In the following we suppose that the agent's mental representation has a fixed architecture with a number of free parameters to optimize for. In particular, we assume a two-level hierarchical organization, in which the processing of the input $\boldsymbol{a}$ is done in two steps: (i) evaluating the alternative along concepts, and (ii) weighting the concept scores to arrive at the associated payoff. The structure of this mental representation is depicted on Fig. 1. It defines three layers: the *input layer*, where the microscopic attributes of decision alternatives are perceived, the *middle layer* composed of abstract concepts, and the *output layer* representing the context- and agent-specific payoffs associated with alternatives.

The hierarchical structure depicted in Fig. 1 formally resembles a linear, two-level (concept vectors, mental weights) neural network. Note, however, that it is not a mi-

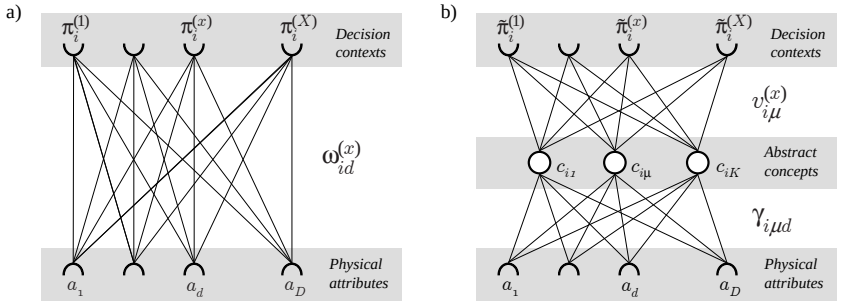


Figure 1: Evaluating decision alternatives under bounded rationality using a finite number of concepts. (a) Structure of reality reflecting idiosyncratic preferences, (b) structure of the mental model (both assumed linear).

croscopic neural network model of a cognitive function, but a phenomenological (macroscopic) model of a generic decision making strategy under bounded rationality. This structure is based on research in psychology, asserting that the human mind is a "feature detector" that can only perceive the aspect of reality which it has a concept for [19].

Each concept μ in the middle layer is a real valued function $\boldsymbol{a} \to c_{i\mu}(\boldsymbol{a})$ (the agent index i represents that the meaning of concepts may vary from agent to agent). We assume that the number of concepts, i.e., the size of the (mental) dictionary, is the same fixed number K for all agents with $K << D$. Thus, the first step implies *dimension reduction*, a mapping $\mathbb{R}^D \to \mathbb{R}^K$. The second step involves calculating the approximate payoff in a given decision context as a function of the concept scores, $\tilde{\pi}_i^{(x)} = p_i^{(x)}(c_{i1}, c_{i2}, \ldots, c_{iK})$, where $p_i^{(x)}$ is an appropriate function $\mathbb{R}^K \to \mathbb{R}$ called the *mental model* of agent i in context x. Each agent possesses a number X of such mental models, one for each decision context.

At the lowest level on Figure 1, agents are homogeneous and have identical perception of reality. In the highest layer they are heterogeneous and have individual payoff functions based on their individual preferences. The middle layer with abstract concepts shows partial coherence, whose measure, as we will see, is determined by the strength of social interactions. In other words, agents more or less agree in the meanings of the concepts, but there is no perfect consensus. This is the layer we intend to monitor across society for observing the emergence and evolution of a shared language.

In general, both the concept functions $c_{i\mu}$ and the mental models $p_i^{(x)}$ can be nonlinear functions. However, in order to keep the model simple, in the following, we will assume that all these functions are *linear*. This "linear mind" assumption and the former "linear reality" assumption are crude approximations but ensure some analytical results and thus useful insight. Accordingly, we assume that concepts are linear mappings defined by *concept vectors* $\{\boldsymbol{\gamma}_{i\mu}\}_{\mu=1}^K$, i.e.,

$$c_{i\mu}(\boldsymbol{a}) = \boldsymbol{\gamma}_{i\mu} \cdot \boldsymbol{a}, \tag{2}$$

and that the mental models are linear, too,

$$\tilde{\pi}_i^{(x)}(\boldsymbol{a}) = \sum_{\mu=1}^{K} v_{i\mu}^{(x)} c_{i\mu}, \tag{3}$$

where the coefficients $v_{i\mu}^{(x)}$ will be called *mental weights*. Putting the two steps together, the *approximate payoff* is bilinear in the $\boldsymbol{\gamma}$ and v parameters

$$\tilde{\pi}_i^{(x)}(\boldsymbol{a}) = \sum_{\mu=1}^{K} v_{i\mu}^{(x)} \boldsymbol{\gamma}_{i\mu} \cdot \boldsymbol{a}. \tag{4}$$

We emphasize that, as opposed to $v_{i\mu}^{(x)}$, the concepts are context-independent (there is no superscript (x) for $c_{i\mu}$), i.e., they have the same "meaning" in all decision contexts for a given agent.

Why is this cognitive architecture preferred? In this scheme the complexity, i.e., the number of variables defining an agent is $(X + D)K$, which can be much less than the total number of parameters describing the world, XD. However, there is a price to pay for this "bounded rationality": due to the reduction of dimensionality, $K << D$, the approximate payoff $\tilde{\pi}_i^{(x)}$ deviates from the objective payoff $\pi_i^{(x)}$. In fact we can think of K as the optimal number of concepts which arises endogenously in the trade-off of precision vs. model (representation) complexity. Nevertheless, as the exact value of K is not essential to our analysis, we simplify the model by assuming that K is exogenously fixed.

In our model, standard grammatical categories like nouns and adjectives are confounded. Each decision alternative receives a score on a concept, which can be associated with a noun or an adjective alike. For example, a stool will receive a high score on the "chair-ness" concept. This feature is consistent with empirical work on people's mental representations [14, 16, 15], which shows that, while people have a tendency to cluster things in distinct categories, membership in a category is not rigid but can be represented with a so-called "graded structure" where a member is measured on how typical it is for that category. (This is also the fundamental concept behind "fuzzy logic".)

Finally, a three-layer structure with one concept layer is clearly a simplified model of language, which could be better described with many layers, each representing a different level of abstraction. However, we can easily replace the lowest (physical) layer by a layer of concepts that are (for all practical purposes) completely agreed upon by agents. For example, there is strong agreement between people on the meanings of concepts/words like "chair", "table", "fork". There is much less agreement however, on the meanings of abstract words like "truth", or "God". We are interested in understanding how the meanings of such abstract concepts emerge.

Let us illustrate the model with a further example. Consider a headhunter seeking candidates (alternatives) for job openings (contexts). The headhunter possesses a large amount of raw data about the candidates in the form of CVs, test results, photos, recommendation letters, certificates, interview recordings, etc. (physical attributes, $\boldsymbol{a}$), and would like to use these to direct the right candidate to the right job. In theory, each job is associated with a complicated function (objective payoff), mapping candidates (described by $\boldsymbol{a}$) to payoff of the candidate for the given job ($\pi^{(x)}(\boldsymbol{a})$). Without going into the intimate details of this complex relationship, the headhunter can "summarize" the candidate profiles along a small number of suitably chosen concepts such as "level of education", "expertise", "communication skills", "physical appearance", etc, which she can (easily) extract from the attributes. Moreover, she classifies openings into typical job categories (say, musician, scientist, executive, etc.) for which she already possesses a weighting scheme (mental weights) along the concepts used. The question is how to find the most efficient set of abstract concepts to minimize her efforts but maximize her

matchmaking efficiency. Obviously, headhunters should optimize their concepts for the general genre of cliental they work for. Those specialized in hunting movie actors will apply rather different concepts (language) than those recruiting corporate executives.

3.2 Representation error

Each agents' goal is to find the best possible set of concepts and mental weights that minimize the error of the mental representation under the constraint that only a finite number K of concepts can be used. The natural measure of agent i's *representation error* is the variance of the payoff deviation over all decision contexts,

$$E_i^{\mathrm{REP}} = \sum_{x=1}^{X} \left\langle \left[\pi_i^{(x)}(\boldsymbol{a}) - \tilde{\pi}_i^{(x)}(\boldsymbol{a}) \right]^2 \right\rangle_{ix} . \tag{5}$$

Such a quadratic error function equally penalizes positive and negative deviations of the predicted (approximate) payoffs from the objective (exact) ones.

In Eq. (5), $\langle f(\boldsymbol{a}) \rangle_{ix} = \int f(\boldsymbol{a})\, \rho_i^{(x)}(\boldsymbol{a})\, d\boldsymbol{a}$ denotes average over the occurrences of alternatives, which is characterized by the probability density $\rho_i^{(x)}(\boldsymbol{a})$. In a realistic setup the distribution of alternatives can be context and agent dependent. Some alternatives may occur with different probabilities (maybe with zero probability) for some agents and/or in some decision contexts. In the following we omit context dependence, but keep a possible agent dependence, and assume $\rho_i^{(x)}(\boldsymbol{a}) = \rho_i(\boldsymbol{a})$ for all x. This simplifies the forthcoming analysis without losing essential features. We assume that the attributes are centered and their correlation structure is represented by the context-independent *covariance matrix* $\boldsymbol{A}_i$,

$$\langle a_d \rangle_{ix} = 0, \quad \langle a_d a_{d'} \rangle_{ix} = [\boldsymbol{A}_i]_{dd'} \quad \forall x. \tag{6}$$

Using Eq. (6) it is easy to see that the representation error becomes

$$E_i^{\mathrm{REP}} = \sum_{x=1}^{X} \left(\boldsymbol{\omega}_i^{(x)} - \sum_{\nu=1}^{K} v_{i\nu}^{(x)} \boldsymbol{\gamma}_{i\nu} \right) \cdot \boldsymbol{A}_i \left(\boldsymbol{\omega}_i^{(x)} - \sum_{\nu=1}^{K} v_{i\nu}^{(x)} \boldsymbol{\gamma}_{i\nu} \right) . \tag{7}$$

The agent's goal is to minimize his/her error E_i^{REP} by optimally choosing the concept vectors $\boldsymbol{\gamma}_{i\nu}$ and mental weights $v_{i\nu}^{(x)}$. Recall that $\boldsymbol{\omega}_i^{(x)}$ and $\boldsymbol{A}_i$ are assumed fixed in the model.

As formulated so far, concept vectors and mental weights are both dynamic variables. However, it is reasonable to think about concepts as "slow" variables, changing noticeably on the scale of decades or centuries, partly because they are shared across agents. In contrast, the mental weights are agent-specific and adapt to the existing concepts in months or years. Certainly, learning to use correctly a concept is much faster than inventing a new, collectively successful concept. Thus, in the following, we optimize the mental weights, assuming that they accommodate to the slow variables very shortly, and only keep concept vectors as dynamic variables.

Given the $\boldsymbol{\gamma}$ vectors the optimal value of $v_{i\mu}^{(x)}$ follows from the solution to the equation $\partial E_i^{\mathrm{REP}} / \partial v_{i\mu}^{(x)} = 0$. From this condition, and assuming that the $K \times K$ symmetric matrix

$$G_{i\nu\mu} = \boldsymbol{\gamma}_{i\nu} \cdot \boldsymbol{A}_i \boldsymbol{\gamma}_{i\mu} \tag{8}$$

is invertible (note that the concepts are not necessarily orthogonal), the optimal mental weights turn out to be:

$$v_{i\mu}^{(x)} = \sum_{\nu=1}^{K} [G_i^{-1}]_{\mu\nu} \boldsymbol{\gamma}_{i\nu} \cdot \boldsymbol{A}_i \boldsymbol{\omega}_i^{(x)}. \tag{9}$$

Writing this back to Eq. (7) we can write the error now as a function of the $\boldsymbol{\gamma}$'s only:

$$E_i^{\mathrm{REP}} = \sum_{x=1}^{X} \boldsymbol{\omega}_i^{(x)} \cdot \boldsymbol{A}_i \boldsymbol{\omega}_i^{(x)} - \sum_{x=1}^{X} \sum_{\mu,\nu=1}^{K} \left(\boldsymbol{\omega}_i^{(x)} \cdot \boldsymbol{A}_i \boldsymbol{\gamma}_{i\mu} \right) [G_i^{-1}]_{\mu\nu} \left(\boldsymbol{\omega}_i^{(x)} \cdot \boldsymbol{A}_i \boldsymbol{\gamma}_{i\nu} \right). \tag{10}$$

The first term is an uninteresting constant which can be neglected. We introduce a *representation utility* U_i^{REP} as the negative of the second term. This can be cast in a more compact form:

$$U_i^{\mathrm{REP}}(\boldsymbol{\Gamma}_i) = \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^T \boldsymbol{A}_i^T \boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{\Gamma}_i \boldsymbol{G}_i^{-1} \right), \tag{11}$$

where Tr is the trace of the matrix, $\boldsymbol{\Gamma}_i \equiv [\boldsymbol{\gamma}_{i1} | \boldsymbol{\gamma}_{i2} | \ldots | \boldsymbol{\gamma}_{iK}]$ is the agent's *language matrix* formed from the concept vectors as columns, and

$$[\boldsymbol{B}_i]_{dd'} = \sum_x \omega_{id}^{(x)} \omega_{id'}^{(x)} \tag{12}$$

is the agent's *preference matrix*. The representation utility U_i^{REP} is a function of the agent's language matrix $\boldsymbol{\Gamma}_i$.

The representation utility is maximal if the concept vectors are chosen optimally. However, even before trying to solve this optimization problem it is immediately clear that the solution cannot be unique. Indeed, due to the linearity of the model the representation utility is invariant for a redefinition of the concept vectors in any (possibly nonorthogonal) ways, provided that the new vectors span the same K dimensional subspace.

Lemma 1 *Let $\boldsymbol{R}$ be an arbitrary real, nonsingular, $K \times K$ matrix. For the transformation $\widetilde{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}\boldsymbol{R}$ the representation utility is invariant:*

$$U^{\mathrm{REP}}(\widetilde{\boldsymbol{\Gamma}}) = U^{\mathrm{REP}}(\boldsymbol{\Gamma}). \tag{13}$$

Proof: Since $\boldsymbol{R}$ is not singular, $\boldsymbol{R}^{-1}$ exists. The metric tensor transforms as

$$\widetilde{\boldsymbol{G}}_x = \boldsymbol{R}^T \boldsymbol{G}_x \boldsymbol{R}, \tag{14}$$

and its inverse becomes

$$\widetilde{\boldsymbol{G}}_x^{-1} = \boldsymbol{R}^{-1} \boldsymbol{G}_x^{-1} (\boldsymbol{R}^T)^{-1}. \tag{15}$$

Using this and the cyclic property of the trace,

$$\begin{aligned} U^{\mathrm{REP}} &= \mathrm{Tr}\left[\widetilde{\boldsymbol{\Gamma}}^T \boldsymbol{A}^T \boldsymbol{B} \boldsymbol{A} \widetilde{\boldsymbol{\Gamma}} \widetilde{\boldsymbol{G}}^{-1} \right] = \mathrm{Tr}\left[\boldsymbol{R}^T \boldsymbol{\Gamma}^T \boldsymbol{A}^T \boldsymbol{B} \boldsymbol{A} \boldsymbol{\Gamma} \boldsymbol{R} \boldsymbol{R}^{-1} \boldsymbol{G}^{-1} (\boldsymbol{R}^T)^{-1} \right] \\ &= \mathrm{Tr}\left[\boldsymbol{\Gamma}^T \boldsymbol{A}^T \boldsymbol{B} \boldsymbol{A} \boldsymbol{\Gamma} \boldsymbol{G}^{-1} \right], \end{aligned} \tag{16}$$

as claimed. □

In natural languages concepts are not always fully independent, but there is a tendency to describe reality along more or less uncorrelated dimensions (synonyms are

exceptions and we do not consider this possibility here). For example, identifying positions in 2D space we can use concepts like "left–right" and "front–rear" or alternatively "North–South" and "East–West", but we hardly use correlated pairs like "North-South" and "North-West–South-East", although this would be theoretically possible. (In this example, the word-pairs like "left–right" are understood as a single concept for a coordinate axis.) The likely reasons are the mental difficulty to process correlated variables and/or the increased sensitivity for noise of the mental model when concepts trying to span the relevant subspace are strongly correlated. It seems that real mental models involve a (nonlinear) cost term related to the correlation of concepts, and this cost term brings about an effective "repulsion" for concepts.

Having the liberty of Lemma 1 to choose a basis freely in the optimal subspace we can mimic this effect by requiring that the concept vectors be exactly uncorrelated. The correlation of concept μ with concept ν is defined as $\langle c_\mu c_\nu \rangle_x = \langle (\boldsymbol{\gamma}_\mu \cdot \boldsymbol{a})(\boldsymbol{\gamma}_\nu \cdot \boldsymbol{a}) \rangle_x = \boldsymbol{\gamma}_\mu \cdot \boldsymbol{A} \boldsymbol{\gamma}_\nu$, thus the agent's language is uncorrelated if

$$\boldsymbol{G}_i = \boldsymbol{\Gamma}_i^T \boldsymbol{A}_i \boldsymbol{\Gamma}_i = \mathbf{1}. \tag{17}$$

This condition can be used as a constraint in the optimization problem. This constraint reduces the degeneracy of the optimum stated by Lemma 1, but not fully, as will be discussed in the sequel. Given the constraint we have $\boldsymbol{G}_i^{-1} = \mathbf{1}$, and the representation utility simplifies to

$$U_i^{\mathrm{REP}}(\boldsymbol{\Gamma}_i) = \mathrm{Tr}[\boldsymbol{\Gamma}_i^T \boldsymbol{A}_i^T \boldsymbol{W}_i \boldsymbol{\Gamma}_i]. \tag{18}$$

where we have introduced for later convenience the generically non-symmetric matrix

$$\boldsymbol{W}_i \equiv \boldsymbol{B}_i \boldsymbol{A}_i, \tag{19}$$

the so-called *world matrix*. $\boldsymbol{W}_i$ represents in a concise form agent i's overall relationship to the world. It encompasses the agent's perception of structure in the occurrences of decision alternatives ($\boldsymbol{A}_i$) and his/her subjective preferences ($\boldsymbol{B}_i$).

3.3 Communication between agents

So far, we have only talked about how using a finite number of concepts affects decision making. Agents have a second important objective, namely to communicate. We assume that communication between pairs of agents occurs on the level of concepts. In other words, agents cannot communicate the values of the large number of physical attributes associated with an alternative, but can only provide the corresponding – relatively small number of – concept scores. The basic idea is that communication cannot operate on the level of attributes due to limited bandwidth, nor on the level of payoffs due to substantial heterogeneity in preferences, which is anticipated by the agents. For example, it is not possible to describe all details of a flower to a person trying to purchase flowers on the phone, but it is not too informative to say "I like it" or "it is beautiful" either unless agents have very similar preferences (tastes).

It is obvious that if agents i and j use a somewhat different definition (different meaning) for concept μ, then their communication involving this concept introduces some misunderstanding. We can assume that misunderstanding, in general, implies disutility for the agents, whose amount depends on how different the two concept vectors $\boldsymbol{\gamma}_{i\mu}$ and $\boldsymbol{\gamma}_{j\mu}$ are. The lowest order (bilinear) measure of the misunderstanding error is related to the overlap of concept vectors. Thus $\boldsymbol{\gamma}_{i\mu} \cdot \boldsymbol{\gamma}_{j\mu}$ can be used as a reasonable

measure of the average communication benefit for the agents. When the two concepts are exactly identical, $\boldsymbol{\gamma}_{i\mu} = \boldsymbol{\gamma}_{j\mu}$, communication utility is maximal. We write i's benefit from communication with agent j as

$$U_{ij}^{\mathrm{COM}} = \sum_{\mu} C_{ij\mu}\, \boldsymbol{\gamma}_{i\mu} \cdot \boldsymbol{\gamma}_{j\mu}, \tag{20}$$

with $C_{ij\mu}$ denoting the importance of concept μ in the communication of agents i and j. In the following, we restrict our attention to the simple case $C_{ij\,\mu} = c/(I-1)$ for all agent pairs identically, where $I-1$ is the number of agents to communicate with, and c is the exogenous *rate of communication* in the community. The denominator is introduced to obtain a meaningful limit when $I \to \infty$.

The contribution of all communications to agent i's utility is

$$U_i^{\mathrm{COM}} = c\,\frac{1}{I-1}\sum_{j\neq i}^{I} \mathrm{Tr}(\boldsymbol{\Gamma}_i^T \boldsymbol{\Gamma}_j). \tag{21}$$

As is formulated above, communication benefit is a symmetric function, $U_{ij}^{\mathrm{COM}} = U_{ji}^{\mathrm{COM}}$. Indeed, it is reasonable to postulate that the benefit of communication is distributed symmetrically between the two agents involved. Communication is typically a role game in which the roles of being a sender (speaker) or a receiver (hearer) interchanges from time to time. Both sender and receiver can benefit in a single communication act: depending on the content of the message the sender can generate profit by influencing the receiver, or the receiver can have benefit by getting information from the sender. On the long run, benefit accumulates on both sides.

Collecting the representation and communications terms together the overall utility for language of agent i is $U_i = U_i^{\mathrm{REP}} + U_i^{\mathrm{COM}}$, which takes the form

$$U_i \;=\; \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^T \boldsymbol{A}_i^T \boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{\Gamma}_i\right) + c\,\frac{1}{I-1}\sum_{j\neq i}^{I} \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^T \boldsymbol{\Gamma}_j\right), \tag{22}$$

with the nonlinear constraint

$$\boldsymbol{\Gamma}_i^T \boldsymbol{A}_i \boldsymbol{\Gamma}_i = \mathbf{1}. \tag{23}$$

As defined by Eqs. (22) and (23), we have a coupled and constrained maximization problem for the individual language matrices $\boldsymbol{\Gamma}_i$, $i = 1, \ldots, I$.

Again, even without solving the problem, it is clear that the maximizing solution will not be unique. Let us introduce the shorthand notations $\boldsymbol{\Gamma}_{\mathrm{all}} \equiv (\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_I)$ (all concept vectors in the problem) and $\boldsymbol{\Gamma}_{-i} \equiv \boldsymbol{\Gamma}_{\mathrm{all}} \backslash \boldsymbol{\Gamma}_i$ (all concept vectors but those of agent i) for later convenience. Clearly $\boldsymbol{\Gamma}_{\mathrm{all}} = (\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_{-i})$ for any i. As the following lemma asserts, a *collective orthogonal rotation* in the subspaces spanned by the concept vectors leaves all U_i invariant:

Lemma 2 *An identical collective rotation of the concept vectors for all agents*

$$\forall\, i \quad \boldsymbol{\Gamma}_i' = \boldsymbol{\Gamma}_i \boldsymbol{O}, \qquad \boldsymbol{O}\boldsymbol{O}^T = \mathbf{1}, \tag{24}$$

leaves U_i and the constraints invariant,

$$U_i(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_{-i}) = U_i(\boldsymbol{\Gamma}_i', \boldsymbol{\Gamma}_{-i}'), \qquad \boldsymbol{\Gamma}_i'^T \boldsymbol{A}_i \boldsymbol{\Gamma}_i' = \mathbf{1}. \tag{25}$$

Proof: It follows from Lemma 1 that U^{REP} is invariant, so it remains to prove that U^{COM} is also invariant. This boils down to show that $\sum_\mu \boldsymbol{\gamma}_{i\mu} \cdot \boldsymbol{\gamma}_{j\mu} = \mathrm{Tr}(\boldsymbol{\Gamma}_i^T \boldsymbol{\Gamma}_j)$ is invariant. Indeed, we have

$$\mathrm{Tr}(\boldsymbol{\Gamma}_i'^T \boldsymbol{\Gamma}_j') = \mathrm{Tr}(\boldsymbol{O}^T \boldsymbol{\Gamma}_i^T \boldsymbol{\Gamma}_j \boldsymbol{O}) = \mathrm{Tr}(\boldsymbol{\Gamma}_i^T \boldsymbol{\Gamma}_j), \tag{26}$$

where we have used the cyclic property of the trace and that $\boldsymbol{O}$ is orthogonal.

As for the constraint, we have

$$\boldsymbol{\Gamma}_i'^T \boldsymbol{A}_i \boldsymbol{\Gamma}_i' = \boldsymbol{O}^T \boldsymbol{\Gamma}_i^T \boldsymbol{A}_i \boldsymbol{\Gamma}_i \boldsymbol{O} = \boldsymbol{O}\boldsymbol{O}^T = \mathbf{1}, \tag{27}$$

where Eq. (23) and the orthogonality of $\boldsymbol{O}$ was used. □

The only possibility to get rid of this rotational degeneracy, and fix the concept vectors unambiguously is to add some nonlinearity to the model. We can introduce a further cost term (complexity or structural cost) associated with the distribution of coefficients connecting the concepts to the physical attributes. A standard choice is

$$Q(\boldsymbol{\Gamma}_i) = -\epsilon \sum_\mu q(\boldsymbol{\gamma}_{i\mu}), \qquad q(\boldsymbol{\gamma}) = \sum_{d=1}^{D} \gamma_d^4, \tag{28}$$

but any similar nonlinear function could do just as well. In Eq. (28) ϵ is positive and infinitesimally small, thus it does not deform the subspace itself. Such a cost eliminates the artificial degeneracy arising from the linearity of our model and can be used to select from otherwise degenerate configurations. A possible interpretation is that it is easier to evaluate a concept, which only depends on a small number of relevant attributes, than one, which has more or less equivalent "loadings" on many. In neurophysiological terms a smaller number of synapses are required to approximate this representation. The weak links can be cut without committing large error, thus instead of working with a fully connected (semantic) network, a sparse network can be used as a good approximation. The specific form of Q in Eq. (28) is the standard "quartimax" rotation criterium, widely applied in the theory of PCA and factor analysis [20]. In fact, most of our results can be presented without considering such a nonlinear term – the exceptions will be pointed out in due course.

3.4 Measuring the coherence of meanings

When agents' preferences and probability densities for alternatives differ they will end up using concepts with different meanings. However, as the importance (benefit) of communication increases (c increases), there is a pressure on agents to share coherent meanings. The coherence will only be perfect at $c = \infty$. For a quantitative measure of coherence we introduce two definitions.

Definition 1 (Average meaning and coherence of concepts) *The* average meaning *of concept μ is the population average of the individual concept vectors*

$$\bar{\boldsymbol{\gamma}}_\mu = \frac{1}{I} \sum_{j=1}^{I} \boldsymbol{\gamma}_{j\mu}. \tag{29}$$

The measure of its coherence *is the length, $|\bar{\boldsymbol{\gamma}}_\mu|$.*

When coherence is perfect, i.e., all agents have identical concept vectors for concept μ the coherence (order-) parameter has unit length $|\bar{\gamma}_\mu| = 1$. In the opposite limit $|\bar{\gamma}_\mu| = 0$ we can speak about complete disorder.

It is also useful to define a scalar parameter, which measures the *overall* coherence of the language as a whole.

Definition 2 (Coherence of language) *The overall coherence L of a language is defined to be*

$$L(\mathbf{\Gamma}_{\text{all}}) = \left(\frac{1}{I^2} \sum_{i}^{I} \sum_{j>i}^{I} \text{Tr}\left(\mathbf{\Gamma}_i^T \mathbf{\Gamma}_j\right) \right)^{1/2}. \tag{30}$$

Note that in the limit of large populations, $I \to \infty$, L can be written as a quadratic function of the concept coherences,

$$L(\mathbf{\Gamma}_{\text{all}}) = \left(\frac{1}{2} \sum_{\mu=1}^{K} \bar{\gamma}_\mu^2 \right)^{1/2} + \mathcal{O}\left(\frac{1}{I}\right), \tag{31}$$

and, thus, measures naturally the coherence of all the concept vectors in the language.

As it was discussed above, $U_i^{\text{REP}} + U_i^{\text{COM}}$ is invariant for a collective rotation of the concept vectors, and the infinitesimal nonlinear term Q was introduced to lift this degeneracy. It is important to note that the $\bar{\gamma}_\mu$ vectors are not invariant for such collective rotations, but L as defined in Eq. (30) is.

Lemma 3 *An identical collective rotation of the concept vectors for all agents*

$$\forall\, i \quad \mathbf{\Gamma}_i' = \mathbf{\Gamma}_i \boldsymbol{O}, \qquad \boldsymbol{O}\boldsymbol{O}^T = \mathbf{1}, \tag{32}$$

leaves L invariant,

$$L(\mathbf{\Gamma}_{\text{all}}) = L(\mathbf{\Gamma}'_{\text{all}}). \tag{33}$$

Proof: As seen in the proof of Lemma 2, each term in Eq. (30) is invariant in itself. □

The fact that collective rotations leave the agents' utilities and the overall coherence of language invariant asks for an interpretation of the invariant subspace spanned by the concept vectors. It is tempting to interpret this subspace as "culture" [21, 22]. Effectively, this interpretation says that we can call two agents culturally identical if one can predict exactly the behavior of the other in all possible decision problems. This doesn't mean that agents would make identical decisions, since their preferences may be different (heterogeneity). However, if their concept subspaces are identical, they can understand/predict each other accurately. In contrast, if the subspaces are misaligned, there is always some prediction error (misunderstanding or cultural incommensurability) between the agents. This is a useful working definition of "culture" because it allows – as in the real world – for the existence of different "languages" within the same culture, i.e., different basis vectors spanning the same subspace. The individual concepts are different across these languages but alternatives can be described identically in each. One could consider a stricter definition of culture, which requires the identity of preferences as well. With this stricter definition however, the left and right political parties in a country would belong to different cultures, which is a somewhat uncomfortable interpretation. The definition of culture has to allow for different preferences.

4 Equilibrium Languages

4.1 Single agent or identical agents

Let us first investigate the properties of our model Eq. (22-23) in the case when there is no communication between agents, i.e., $c = 0$. This is the problem of isolated agents who develop a mental representation of their world on their own. It also corresponds to the case when agents are identical. A solution which maximizes U_i under the constraint is trivially a Nash equilibrium of the Language Game. This limit can serve as a benchmark in the analysis of the more interesting case when multiple, heterogeneous agents interact.

Proposition 1 *Without social interactions the optimal (equilibrium) language solves a Principal Component Analysis (PCA) problem. The optimal concepts span the most significant subspace of the world matrix $\boldsymbol{W}_i = \boldsymbol{B}_i\boldsymbol{A}_i$, and within this subspace minimize the nonlinear cost function Q (if present).*

Proof: Let us consider first the case $Q = 0$. In the lack of interactions the utility of agent i reads

$$U_i(\boldsymbol{\Gamma}_i) \;\; = \;\; \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^T\boldsymbol{A}_i^T\boldsymbol{W}_i\boldsymbol{\Gamma}_i\right) - \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^{\mathrm{T}}\boldsymbol{A}_i\,\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}\right). \tag{34}$$

where the second term is the constraint added with Lagrange multipliers. The $K \times K$ symmetric matrix $\boldsymbol{\Lambda}$ is a compact form of the Lagrange multipliers for all the $K(K+1)/2$ components of the constraint. The maximization problem for $\boldsymbol{\Gamma}_i$ in Eq. (34) is a Principal Component Analysis (PCA) problem [23]. Indeed, varying with respect to $\boldsymbol{\Gamma}^T$ we obtain the condition of extremum

$$\boldsymbol{A}_i^T\boldsymbol{W}_i\boldsymbol{\Gamma}_i = \boldsymbol{A}_i\,\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}. \tag{35}$$

Assuming that the symmetric covariance matrix $\boldsymbol{A}_i = \boldsymbol{A}_i^T$ is nonsingular and thus invertible, Eq. (35) is equivalent to

$$\boldsymbol{W}_i\boldsymbol{\Gamma}_i = \boldsymbol{\Gamma}_i\boldsymbol{\Lambda}. \tag{36}$$

This latter states that the K-dimensional subspace spanned by the concept vectors (language matrix) is an invariant subspace of the world matrix $\boldsymbol{W}_i$. The only K-dimensional invariant subspaces are the ones spanned by K of the eigenvectors of $\boldsymbol{W}_i$. The remaining question is how to choose the eigenvectors to maximize the utility.

Let $\lambda_1 \geq \ldots \geq \lambda_n \geq \ldots \geq \lambda_D \geq 0$ denote the eigenvalues of the world matrix in decreasing order, and $\boldsymbol{w}_n$, $n = 1, \ldots, D$, the associated eigenvectors, $\boldsymbol{W}_i\boldsymbol{w}_n = \lambda_n\boldsymbol{w}_n$. The above ordering is possible, since the eigenvalues of $\boldsymbol{W}_i$ are all real and non-negative. (The fact that both $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are symmetric, positive definite is enough to prove this.) If the subspace is spanned by the eigenvectors $\boldsymbol{w}_{n_1}, \boldsymbol{w}_{n_2}, \ldots, \boldsymbol{w}_{n_K}$, i.e., the language matrix $\boldsymbol{\Gamma}$ is constructed from these vectors as columns, the utility in Eq. (18) becomes

$$U_i^{\mathrm{REP}} = \mathrm{Tr}\left(\boldsymbol{\Gamma}_i^T\boldsymbol{A}_i^T\boldsymbol{W}_i\boldsymbol{\Gamma}_i\right) = \mathrm{Tr}\left(\boldsymbol{\Lambda}\right) = \sum_{\mu=1}^{K}\lambda_{n_\mu}, \tag{37}$$

where we have used Eq. (36). This is maximal if $n_\mu = \mu$, i.e., if the eigenvectors chosen in the language matrix are the ones with the K largest eigenvalues. Thus the utility

maximizing language arises as the PCA problem of the world matrix. Having the optimal concepts determined, the mental weights of the representation adapt according to Eq. (9).

Assuming that $\boldsymbol{W}_i$ is full rank, all the eigenvalues are positive, and for $K < D$ the optimal utility $U_i^{\rm opt}$ is strictly below the theoretical maximum ${\rm Tr}\boldsymbol{W}_i = \sum_{\mu=1}^{D} \lambda_{i\mu}$. The representation error due to the reduction of dimensionality is the weight of omitted eigenvectors

$$E_i^{\rm REP} = \sum_{\mu=K+1}^{D} \lambda_{i\mu}. \tag{38}$$

The particular solution we have found is not the only solution which maximizes $U_i^{\rm REP}$ and complies with the constraints. As shown by Lemma 2, any rigid rotation within the subspace spanned by these eigenvectors produces another degenerate solution. (Note that since $K < D$, any configuration that only differs from the reference configuration by a relabeling of the basis vectors or by sign flips of some of the vectors can also be attained by a suitable rigid rotation.) This infinite degeneracy is, however, lifted by the nonlinear cost term Q, when it is added. When Q is infinitesimally small, the solution remains in the "most significant subspace", but the basis vectors get determined, at least up to discrete transformations such as sign flips and relabeling. For finite Q even the subspace gets deformed, nevertheless the nonlinearity helps fixing the optimal concepts in an unambiguous manner. □

The PCA-based optimal mental representation discussed above may be reached by practically any (myopic) utility maximizing learning dynamics. Note that without agent-agent interactions this is a (constrained) optimization problem on a fixed landscape which is quadratic (without the Q term) and thus smooth. There is no danger that the dynamics get stuck in suboptimal local maxima [24].

4.2 Nash equilibria for multiple agents

Let us investigate now agent interactions, and set $c > 0$ and $I > 1$. Communication between agents will deform their individual mental representations away from the PCA solution. The first question we ask is whether this coupled and constrained system has any Nash equilibrium. The proof of existence will boil down to the following fundamental observation:

Proposition 2 *The Language Game is a potential game. There exists a multi-agent potential $V : \mathbb{R}^{IKD} \to \mathbb{R}$ such that for any two strategy configurations $(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_{-i})$ and $(\boldsymbol{\Gamma}'_i, \boldsymbol{\Gamma}_{-i})$ we have*

$$U_i(\boldsymbol{\Gamma}'_i, \boldsymbol{\Gamma}_{-i}) - U_i(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_{-i}) = V(\boldsymbol{\Gamma}'_i, \boldsymbol{\Gamma}_{-i}) - V(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_{-i}). \tag{39}$$

Proof: The proof goes by an explicit construction of the potential

$$\begin{aligned} V(\boldsymbol{\Gamma}_{\rm all}) &= \sum_{i=1}^{I} \left[U_i^{\rm REP}(\boldsymbol{\Gamma}_i) - Q(\boldsymbol{\Gamma}_i) \right] + \sum_{i=1}^{I} \sum_{j>i}^{I} U_{ij}^{\rm COM}(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_j) \\ &= \sum_{i=1}^{I} \left[{\rm Tr}\left(\boldsymbol{\Gamma}_i^{\rm T} \boldsymbol{W}_i \boldsymbol{\Gamma}_i \right) - Q(\boldsymbol{\Gamma}_i) \right] + \frac{c}{I-1} \sum_{i=1}^{I} \sum_{j>i}^{I} {\rm Tr}\left(\boldsymbol{\Gamma}_i^{\rm T} \boldsymbol{\Gamma}_j \right), \end{aligned} \tag{40}$$

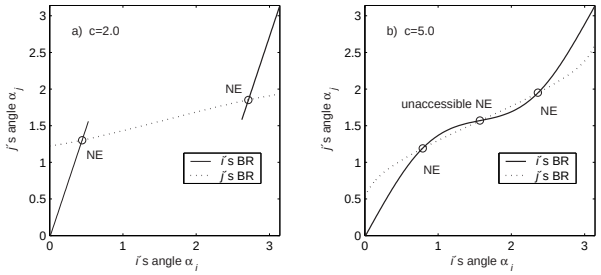


Figure 2: An example: best response (BR) curves for $D = 2$, $X = 1$, $K = 1$ and two agents i and j. The preference vectors are $\boldsymbol{\omega}_i = [1, 0]$, $\boldsymbol{\omega}_j = [0, 1.2]$ and the agents' concept vectors are parameterized by the polar angles $0 \leq \alpha_i, \alpha_j \leq \pi$, resp., $\boldsymbol{\gamma}_i = [\cos\alpha_i, \sin\alpha_i]$, $\boldsymbol{\gamma}_j = [\cos\alpha_j, \sin\alpha_j]$. (a) Two NEs for $c = 2.0$; (b) Three NEs for $c = 5.0$. The one in the middle is a special saddle point of V, and thus unaccessible dynamically.

where we assume that the nonlinear cost term Q is also present in the utility. Obviously the single agent terms are identical on the left and right side of Eq. (39). As for the two-agent terms (COM), the equality holds provided that the interaction is symmetric, i.e., $U_{ij}^{\mathrm{COM}} = U_{ji}^{\mathrm{COM}}$, which is satisfied by the form in Eq. (21). □

A potential can always be constructed if a game is based on symmetric pair interactions [25]. These games are sometimes also called *partnership games* [26]. Note that it follows from the general expression Eq. (39) that the differential forms

$$\frac{\partial U_i}{\partial \boldsymbol{\Gamma}_i} = \frac{\partial V}{\partial \boldsymbol{\Gamma}_i}, \qquad \frac{\partial^2 U_i}{\partial \boldsymbol{\Gamma}_i \partial \boldsymbol{\Gamma}_j} = \frac{\partial^2 V}{\partial \boldsymbol{\Gamma}_i \partial \boldsymbol{\Gamma}_j} \tag{41}$$

also hold for any i and j. This becomes useful in the sequel. Now it is easy to prove that:

Proposition 3 *The Language Game always has at least one Nash equilibrium.*

Proof: The constraints in Eq. (23) make the multi-agent potential V have compact support, and as such it necessarily takes its global maximum at a point $\boldsymbol{\Gamma}^*_{\mathrm{all}} = (\boldsymbol{\Gamma}^*_1, \ldots, \boldsymbol{\Gamma}^*_I)$. (Again, without Q the global maximum would be infinitely degenerate, but Q resolves this degeneracy.) It is easy to see, however, that $\boldsymbol{\Gamma}^*_{\mathrm{all}}$ is a Nash equilibrium. Indeed, no agent has an incentive to deviate from this by choosing $\boldsymbol{\Gamma}_i \neq \boldsymbol{\Gamma}^*_i$, since

$$U_i(\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}^*_{-i}) - U_i(\boldsymbol{\Gamma}^*_i, \boldsymbol{\Gamma}^*_{-i}) = V(\{\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}^*_{-i}\}) - V(\{\boldsymbol{\Gamma}^*_i, \boldsymbol{\Gamma}^*_{-i}\}) \leq 0, \tag{42}$$

where the first equality is assured by Proposition 2, and the second inequality by the fact that $(\boldsymbol{\Gamma}^*_i, \boldsymbol{\Gamma}^*_{-i})$ is a global maximum of V. □

In general, the Nash equilibrium (NE) is not unique. It is typical to have configurations, which are NEs, although they do not maximize the potential V. It is obvious, however, by Eq. (41) that they should necessarily correspond to local extrema of V. This condition is of course not sufficient. Not all local extrema of V are NEs. An example showing the appearance of more than one NEs is shown in Fig. (2) for a simple case: $D = 2$, $X = 1$, $K = 1$ and two agents, $I = 2$. Nash equilibria, which correspond to saddle points of V cannot be accessed dynamically (see later).

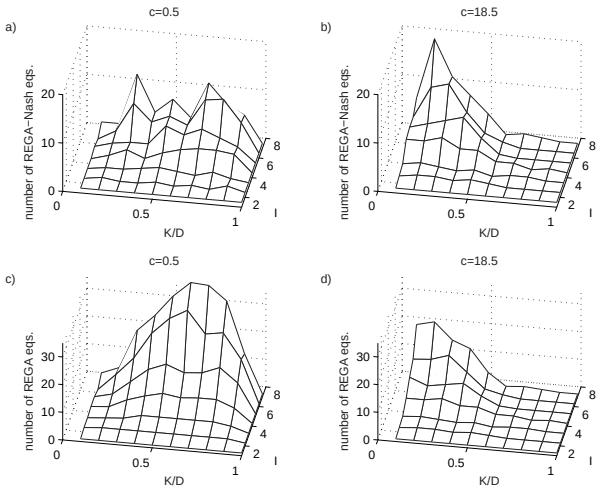


Figure 3: Average number of dynamically accessible Nash equilibria (REGA-Nash equilibria) and stable fixed points (not necessarily NE) of the adjustment dynamics (REGA equilibria) as a function of the relative intelligence K/D and the number of agents I. Subplots (a,c) refer to small communication rate, $c = 0.5$; (b,d) to large communication rate, $c = 18.5$. In the simulations the agent preferences are assumed to be iid Gaussian random vectors, and the structure of alternatives is set $\boldsymbol{A}_i = \mathbf{1}$.

Similar examples can be constructed in higher dimensions and for more agents. The stability of these NEs with respect to the game dynamics will be investigated in Section 5. Some of these NEs turn out to be unstable and thus inaccessible under reasonable evolutionary dynamics. (The possible existence of dynamically inaccessible NEs is a well-known fact of evolutionary game theory. The Folk theorem of evolutionary game theory asserts that under a wide class of dynamics all attractors are NEs, but the converse do not hold. See Cressman [27] and Hofbauer and Sigmund [28] for a formal discussion.) However, even after the omission of these, the non-uniqueness of language equilibria prevails. Thus, our model predicts a strong path-dependence in language evolution, in which the timing and ordering of the appearance of new contexts can play a significant role, and which can be a source of cultural heterogeneity.

The number of dynamically accessible Nash equilibria (to be called REGA-Nash equilibria in the sequel) is a function of the basic model parameters such as the number of agents, the number of concepts, the communication strength, the (heterogeneous) world matrices the agents possess, etc. Figure 3(a-b) illustrates the case when the world matrices are randomly distributed in the population. Random agent properties imply that the number of equilibria is a random variable too. Figure 3(a-b) plots the average number of equilibria determined by a series of simulations. Although there remain considerable fluctuations in the points, especially for small c, the overall picture is rather clear. The number of Nash equilibria can be rather large and typically increases rapidly with the number of agents. This creates a severe coordination problem. An exception for this rule is the region of large c and large K/D. This is the limit when agents are highly intelligent and benefit a lot from communication. In this region we have found that it is typical to have only one or two Nash equilibria, even when the number of agents is rather large. Note that a large communication benefit in itself is not enough

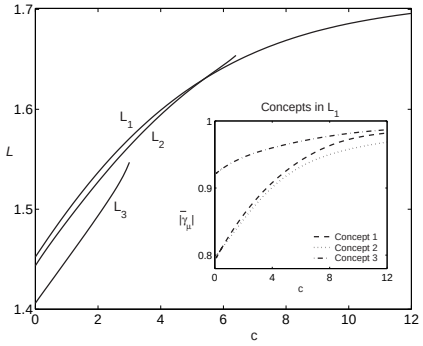


Figure 4: Coherence of language L and individual concepts $|\bar{\gamma}_\mu|$ as a function of the communication strength c for $D = X = 8$, $K = 3$, $I = 4$, and randomly generated, then fixed preferences. L_1, L_2 and L_3 are three different language equilibria appearing for these parameters. Inset shows the coherence of concepts for language L_1.

to reduce the number of NEs, since this only implies that concept vectors of different agents be highly parallel without explicitly defining what they should be. In the case of $K << D$ (low intelligence), the number of Nash equilibria seems to proliferate even when agent utility is dominated by the communication benefit. Our simulation results, although very limited in scope, may indicate a qualitative change in the behavior of the model ("a phase transition") on the phase plain intelligence vs communication rate – the investigation of which is left for a future study.

Many of the equilibria we have found only exist in a certain range of c, and disappear (become unstable) at some critical values. Figure 4(a) demonstrates this by showing the overall language coherence $L = L(c)$ in equilibrium. All curves in the figure denote a different Nash equilibrium. As is seen from Fig. 4(a), the coherence in each equilibrium increases monotonically as a function of c. This is intuitive, since we expect that more communication, i.e., a higher value for the external communication rate, c implies enhanced coherence of the language utilized. In fact we can prove the following:

Proposition 4 *In (each possible) equilibrium the overall coherence $L = L(c)$ of the language is a monotonically increasing function of the communication rate, c.*

Proof: Let us rewrite the interaction part of the potential with the help of L^2. Applying Eq. (30) we get

$$\frac{c}{I}\sum_{i=1}^{I}\sum_{j>i}^{I}\mathrm{Tr}\left(\mathbf{\Gamma}_i^{\mathrm{T}}\mathbf{\Gamma}_j\right) = cI\,L^2. \tag{43}$$

With this the potential has the formal structure

$$V(\mathbf{\Gamma}_{\mathrm{all}}) = V_0(\mathbf{\Gamma}_{\mathrm{all}}) + cI\,L^2(\mathbf{\Gamma}_{\mathrm{all}}), \tag{44}$$

where V_0 represents all single-agent terms, and the c term collects all communication terms. At a dynamically accessible NE, $\mathbf{\Gamma}^*_{\mathrm{all}}$, the potential V necessarily takes its local maximum. As c changes the equilibrium configuration adapts analytically (except for bifurcation points, which are beyond our consideration). The two terms in V compete: for small c it is the maximum of V_0 which determines $\mathbf{\Gamma}^*_{\mathrm{all}}$, whereas for large c it is

the second term. As c increases the balance of importance swings towards the second term, and hence $\mathbf{\Gamma}^*_{\text{all}}$ gets closer and closer to the individual maximizing configuration of this term, meaning that L^{2*} monotonically increases. The rigorous formulation of this argument is delegated, in form of a lemma, to the Appendix. □

Proposition 4 ensures that language as a whole becomes more coherent when the rate of communication increases, at least until the equilibrium (which is followed analytically as a function of c) exists. However, as Fig. 4 illustrates the landscape structure of V can be such that local maxima arise and disappear by varying c. Thus, certain language equilibria may lose stability and disappear in a bifurcation process for some critical value of c. This occurs to L_2 and L_3 in the example presented in Fig. 4.

Although we have found that language as a whole becomes more coherent for increasing communication rate, it is not obviously true for every single concept in the language. The system is coupled in an intimate way, and the concepts themselves get determined by the nonlinear term Q. Even though we could not prove this rigorously, in all our simulations we have found that for all μ, $\bar{\gamma}_\mu$ increases monotonically with c (see the inset of Fig. 4 for an illustration).

5 Social Dynamics

So far, we have only discussed the Nash equilibria of the Language Game. However, we can view language as a dynamic, evolutionary problem in which agents perpetually adapt their mental representations to the changing environment and to each other. Adaptation occurs through a trial-and-error procedure in which the test configuration the agent considers necessarily deviates from his actual (reference) representation. Since the parameter space is enormous, a large (random) change very likely makes the representation worse, and thus will be rejected eventually. It is reasonable to assume a search heuristic, which concentrates on small (local) deformations and a set of obvious discrete transformations, while ad hoc, larger scale deformations of the representation are only tested very rarely. Evolutionary dynamics acts as an equilibrium selection method, which can solve the coordination problem related to multiple equilibria found above.

5.1 REGA dynamics

In the following, we assume a myopic adjustment dynamics in which agents slowly deform their concepts in order to optimize them in a local sense to the natural (perceptual) and social environment, maintaining the assumed constraints. A possible continuous time evolution in this spirit is along the steepest ascent of the utility (*gradient adjustment dynamics*), which reads

$$\frac{\delta \mathbf{\Gamma}_i}{\delta t} = \text{const}\, \boldsymbol{P}_i \frac{\partial U_i}{\partial \mathbf{\Gamma}_i}, \tag{45}$$

where $\boldsymbol{P}_i$ is an adequate projector, which projects the bare gradient $\partial U_i/\partial \mathbf{\Gamma}_i$ (meant by components) into the tangent space allowed by the constraint Eq. (23). It is customary to call $\boldsymbol{P}_i \partial U_i/\partial \mathbf{\Gamma}_i$ the *projected gradient*, which assures that $\mathbf{\Gamma}_i(t)$ continues to respect the constraints for all t.

The gradient dynamics is a local search heuristic, which continuously deforms the concept vectors. However, as dictated by the nature of the problem, it seems reasonable to complement this continuous dynamics with a very specific discrete part, namely *sign*

flips, $\boldsymbol{\gamma}_{i\mu} \to -\boldsymbol{\gamma}_{i\mu}$, and *relabeling transformations*, $\boldsymbol{\gamma}_{i\,\mu} \to \boldsymbol{\gamma}_{i\,R\mu}$, where R is a permutation operator for concepts. Sign flips correspond to using, for instance, "fastness" instead of "slowness", or a (directed) South-North axis instead of a North-South axis. Relabeling, in turn, permutes the associations, which relate concepts to signals used in communication. Configurations attainable by such discrete transformations are readily available for the (boundedly rational) agent, and are assumed to be tested perpetually with some finite probability during gradient adjustment. The flipped or permuted configuration is accepted and replaces the reference configuration, if it increases the agent's utility, whereas it is discarded and the continuous part of the dynamics continues with the reference configuration, if not.

The basic idea is that, as concepts slowly deform due to continuous adjustment, the agent may realize that a proper sign flip or relabeling can vastly improve his/her communication efficiency while leaving his representation error intact. The above discrete processes, which will be referred to collectively as *rematching transformations* (rematching concepts and their linguistic signals used in communication), help to avoid spurious configurations/fixed points, which could be amended trivially by adequately permuting (relabeling) the player's concepts. We will refer to the above dynamics (continuous and discrete parts together) as the "rematching enabled gradient adjustment" (REGA) dynamics. Note that during a REGA iteration step the K-dimensional language subspace only changes slightly. Although rematching transformations make seemingly large configuration changes, they keep the subspace invariant. As such, these rematching transformations can still be considered "myopic" adjustments because they occur within the K-dimensional subspace defined by the concept vectors.

5.2 Fixed points and stability

Let us consider now the potential fixed points of the REGA dynamics.

Proposition 5 *From all initial conditions the REGA dynamics of the Language Game converges to a fixed point.*

Proof: The multi-agent potential V acts as a Lyapunov function for REGA, in the sense that V increases in all iteration steps. This is trivial for the continuous (gradient) part, and also holds for the discrete (rematching) part by Eq. (39). As such the $t \to \infty$ limit of the dynamics is necessarily a fixed point. □

Fixed points of the REGA dynamics can be sorted according to their stability properties. In particular, we define REGA equilibria, as the *stable* fixed points of the dynamics:

Definition 3 (REGA Equilibrium) *A certain choice of concept vectors by agents in the society will be called a REGA equilibrium if this configuration is (1) a fixed point of the REGA dynamics, (2) it is asymptotically stable against infinitesimal individual and collective deviations, and (3) stable against individual and collective rematching transformations.*

Proposition 6 *There exists at least one REGA equilibrium of the Language Game.*

Proof: We are going to prove that the global maximum $\boldsymbol{\Gamma}^*_{\rm all}$ of the multi-agent potential V defined in Eq. (40) satisfies the definition of the REGA equilibrium.

By definition a REGA equilibrium is a fixed points of the REGA dynamics, i.e., a point where the projected gradient $\boldsymbol{P}_i \partial U_i / \partial \boldsymbol{\Gamma}_i$ vanishes for all i. However, by Eq. (41)

this condition is equivalent to the first order condition for local maximum of V. The global maximum $\mathbf{\Gamma}^*_{\rm all}$ of V, which necessarily exists since the support of V is compact, is a point which satisfies this. This is a fixed point of the REGA dynamics. The global potential V may have several local extrema, which are all fixed points of the REGA dynamics. Some of these may be asymptotically stable some may be unstable with respect to infinitesimal deviations. Only the stable fixed points are attainable by the evolutionary dynamics, and only these will be called REGA equilibria.

In order to prove the local stability of $\mathbf{\Gamma}^*_{\rm all}$ against infinitesimally small, not necessarily unilateral but supposedly collective deviations, we have to show that the associated *bordered Hessian* is negative semi-definite at that point [29]. The bordered Hessian is a supermatrix, composed from the second derivatives of the individual utilities $\partial^2 U_i/\partial\mathbf{\Gamma}_i\partial\mathbf{\Gamma}_j$ as a submatrix (the ordinary Hessian), and submatrices formed by the first derivatives of the constraints. However, as follows from Eq. (41), the Hessian piece is the same as the Hessian of the global potential. It follows that the bordered Hessian in question is identical to the bordered Hessian of the potential at $\mathbf{\Gamma}^*_{\rm all}$. Since $\mathbf{\Gamma}^*_{\rm all}$ is the (global) maximum of the potential (satisfying all constraints) its bordered Hessian in necessarily negative semi-definite, proving the assertion.

Lastly, we have to prove that no player can improve his utility by any rematching of his concepts. Let $\tilde{\mathbf{\Gamma}}^*_{\rm all}$ denote a configuration obtained by arbitrary, independent rematching of the concepts. Since $\mathbf{\Gamma}^*_{\rm all}$ is the global maximum rematching cannot improve the global potential

$$V(\tilde{\mathbf{\Gamma}}^*_{\rm all}) \leq V(\mathbf{\Gamma}^*_{\rm all}). \tag{46}$$

On the other hand, as rematching leaves invariant the representation error $U_i^{\rm REP}(\tilde{\mathbf{\Gamma}}^*_i) = U_i^{\rm REP}(\mathbf{\Gamma}^*_i)$ and the complexity cost $Q_i(\tilde{\mathbf{\Gamma}}^*_i) = Q_i(\mathbf{\Gamma}^*_i)$ we can write

$$\begin{aligned} U_i(\tilde{\mathbf{\Gamma}}^*_i, \mathbf{\Gamma}^*_{-i}) - U_i(\mathbf{\Gamma}^*_i, \mathbf{\Gamma}^*_{-i}) &= U_i^{\rm COM}(\tilde{\mathbf{\Gamma}}^*_i, \mathbf{\Gamma}^*_{-i}) - U_i^{\rm COM}(\mathbf{\Gamma}^*_i, \mathbf{\Gamma}^*_{-i}) = \\ & V(\tilde{\mathbf{\Gamma}}^*_{\rm all}) - V(\mathbf{\Gamma}^*_{\rm all}) \leq 0 \end{aligned} \tag{47}$$

where for the last inequality we have used Eq. (46). □

The REGA equilibrium concept differs from the Nash equilibrium concept in two respects. First, a huge amount of otherwise possible global strategy options are excluded from the ones "tested" due to bounded rationality, and second, the equilibrium is required to be stable against collective local deviations too, and not just against unilateral deviations. The rationale behind excluding most of the parameter space from the accessible strategies is its practical infinity with respect to the capacities of the human mind, $K/D \to 0$. Search for better response is necessarily heuristic, focusing essentially on local improvements. The emergence of drastically new successful concepts is necessarily a non-systematic, trial and error mechanism, which has much lower probability.

Note also that none of the two equilibrium concepts implies the other. There are Nash equilibria which are fixed points of the dynamics, but not REGA equilibria because they are unstable against collective deviations – examples are shown in Fig. 2(b). Such points are necessarily saddle points, which cannot be attained dynamically from generic initial conditions. Being exactly on the stable manifold of a saddle point has zero probability, and even if this were the case initially, any noise in the dynamics would finally drive the system away from such a fixed point.

Corollary 1 *A Nash equilibrium of the Language Game that is not a REGA equilibrium at the same time (Nash-only equilibrium), is dynamically inaccessible and thus, cannot be interpreted as language.*

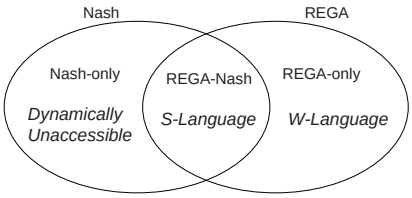


Figure 5: Equilibria and their interpretation as language.

As such, we should focus on REGA equilibria. As demonstrated by the simulation results of Fig. 3(c,d) their number can be rather high. These can additionally satisfy the requirements for a Nash equilibrium too, but this is not necessary. REGA equilibria, which are not Nash equilibria (REGA-only) are unstable against large-scale (global) unilateral deviations. For such fixed points there are agents who would have better response than the actual. However, as assumed above, locating a better response globally is a low-probability event due to the large barrier separating the two configurations. Consequently, the system can stay in such a "metastable" equilibrium for a very long time.

Definition 4 (Weak form of language) *A (metastable) REGA-only equilibrium of the Language Game is called a "Weak form of language" (W-language).*

Finally, equilibria which are both REGA and Nash (REGA-Nash equilibria) have superior stability properties and play a distinguished role. The following definition formally describes such an equilibrium and a proposition states their existence in the Language Game.

Definition 5 (Strong form of language) *A REGA-Nash equilibrium of the Language Game is called a "Strong form of language" (S-language).*

Proposition 7 *The Language Game has at least one REGA-Nash equilibrium (S-language).*

Proof: In the proof of Proposition 3 we have seen that $\mathbf{\Gamma}^*_{\rm all}$, the global maximum of the multi-agent potential V, is a Nash equilibrium, and at Proposition 6 that it is a REGA equilibrium. Consequently, $\mathbf{\Gamma}^*_{\rm all}$ satisfies the requirements of the proposition. □

The different kinds of equilibria, their relationship, and interpretation as language is depicted in Fig. 5. As our simulation results indicate there can be a large number of REGA-Nash equilibria, especially when the population is very heterogeneous and communication has little significance (c small). See Fig. 3 for a quantitative analysis. For less heterogeneity and larger communication strength we have found that almost all REGA equilibria are indeed REGA-Nash equilibria.

The stability properties of the different kinds of equilibria are summarized in Fig. 6. The possible perturbations are categorized according to their scale (local vs. global) indicating the size of the deviation tested, and according to the number of agents they involve (single agent vs. multiple agents). By definition, Nash equilibria are stable against single agent perturbations irrespective whether they are small (local) or large (global). Also, our definition of the REGA equilibrium implies stability with respect to any kind of (single or multiple agent) small (i.e., local) perturbations. As we see REGA-Nash equilibria are the most stable, only their multi-agent global stability remains undetermined.

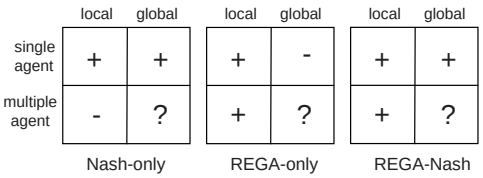


Figure 6: Stability properties of equilibria.

6 Discussion and Conclusion

We have presented a model based on the assumption that agents make choices between alternatives by first describing them on a finite set of concepts. In turn, agents will have an incentive to promote concepts that better fit their preferences. This feature of our model resonates to an old problem in linguistics and philosophy called *linguistic relativism*, which essentially states that the language we speak will influence (or even determine) our thoughts and judgements. As Whorf [30] writes: "We dissect nature along lines laid down by our native language." (See also Wittgenstein [31].) The debate is not whether linguistic relativism is true or false but rather to what extent it is true. Quoting Paul Kay, Ross [32] advances a moderate view: "[Although it may be correct that the languages people speak mold their thought] It is unlikely that the various languages of the world are so different that the ways their speakers think is incommensurable".

Our framework helps predict when linguistic relativism may be important. In our model, language is determined by two key inputs: the structure of the physical world (captured in $\boldsymbol{A}$) and agents' subjective preferences (captured in $\boldsymbol{B}$). As we mentioned however, the three-layer structure of Figure 1 represents the formation of concepts at only one level of abstraction. If we were to study the formation of more abstract concepts, then the bottom layer (the $\boldsymbol{a}$-s) would be concepts themselves that emerged from a previous Language Game. This structure suggests sequential development of more abstract layers in language: simple concepts describing the physical environment emerge first and provide the basis for the evolution of more abstract concepts in a series of subsequent Language Games. The consequence is that linguistic relativism will depend on the concepts' level of abstraction or complexity. When concepts name simple objects that we all perceive identically (because of our biological design), the objective structure of reality will have a major influence on them, and this layer of the equilibrium language is likely to be similar across isolated societies. At higher levels of abstraction, the $\boldsymbol{a}$-s of our model are concepts themselves and as such are less likely to be identical across isolated groups of people as the path-dependence of the previous Language Game has already introduced idiosyncratic structure in them. As we move away from the actual perceptual basis for concepts, heterogenous preferences and the path-dependent nature of language evolution may lead to very different concepts across isolated groups of people. For example, while we have no problem naming furnitures across highly different cultures and make the translation easily between their languages in this domain, our notions of abstract things such as "God" or "ethics" are hard to translate across different cultures.

The well-known Gavagai problem nicely illustrates this point: a foreigner landing on an isolated island tries to infer the meaning of the word "gavagai" shouted by a local pointing to a rabbit crossing the path. While, *a priori* he has no reason to associate the

meaning of "gavagai" with rabbit (it could also be "dinner" or "white" among others) this is the most intuitive meaning. This is because the structure in the data is strong and our differences in terms of payoffs of alternatives are relatively small when it comes to describing the material world around us. Clearly, if the local were to say "gavagai" when pointing to a weird statue representing a creature combining human and animal features, the translation would be extremely hard, if at all possible.

Data in two domains, colors and numbers, also seems to be consistent with the above pattern. Color is a relatively concrete domain that is directly linked to perceptions, i.e., it has low level of abstraction. It turns out that natural languages differ wildly in the (number of) basic colors that they name. Yet, research shows (see, e.g., Ross [32]) that this does not lead to linguistic relativism: people remember and use colors to the same degree of sophistication across cultures. The situation seems to be the opposite when it comes to numbers, which represent a domain with an arguably higher level of abstraction. In a recent study by Peter Gordon of Columbia University [33], it is shown that people using languages that do not have words for numbers higher than 2 have problems comparing and remembering quantities. Linguistic relativism definitely applies in this - more abstract - domain. In sum, empirical patterns suggest that linguistic relativism is likely to be strong in abstract domains of language, while it is weak in domains that are closer to our perceptions, such as concepts describing our physical environment. Our model provides an explanation for why this may happen.

The model is also consistent with psychology's view on how concepts are structured in our mind, namely *categorization theory* and *graded structures* [16]. It is intuitive that language should reflect the structure of the world. This results from humans' innate capacity to categorize things based on similarity, resulting in categories with well defined hierarchies called graded structures. It is also known however, that categories are 'ad hoc' in the sense that similarities are evaluated in relation to goals. As such, objects may potentially belong to any category but with a different weight. In our model, concepts are built from reality in the same way: all components of $\boldsymbol{a}$ may contribute to any of the concepts but their weights will differ wildly depending on people's preferences (goals).

An interesting new insight from our model is the interpretation of the relationship between language and culture. We interpret language as the collection of concepts while culture is the subspace defined by the concepts. Different sets of concepts can define the same subspace, but concepts defining different subspaces cannot be mapped into each other without a large error. This structure suggests that cultural difference doesn't come from disagreement between groups of people or differences in preferences, but rather from the fact that cultures dissect the world along different dimensions. In other words, the concepts used by one culture cannot be mapped accurately into the concepts of the other (there is no, or little possibility for translation) resulting in poor communication across cultures. As mentioned earlier this is only likely for communication in abstract domains (e.g. ethics or esthetics). Furthermore, the model is also consistent with the fact that for any given culture, there might be dozens of languages with notable differences in their concepts. Similarly, the model also accounts for the fact, that individuals may slightly disagree on the meanings of the words within the same language. Despite these differences, communication is possible because choice alternatives can be identically described with multiple sets of concepts covering the same sub-space. In other words, accurate translation and, as a result efficient communication is possible across these languages.

Finally, our study of language's dynamic evolution also provides insights with regard

to empirical observations on the evolution of societies. Archeologists suspect that the so-called *cultural explosion* dating back to some 50 thousand years is somehow linked to the evolution of language (see, e.g., Mithen [34]). At about this date, human evolution seems to become much faster and proper cultures with abstract concepts (e.g. cults or religions) appear in the data. This pattern is consistent with our model. Our analysis of the Language Game's dynamics shows that concepts gradually emerge over time. Once an equilibrium is reached, it represents a consensus across the members of society, which is instantly available for the next generation. This consensus provides a layer for the development of more abstract concepts. While the speed of evolution before language was probably determined by the speed of biological evolution (as evidenced by the gradual increase in brain size, for example), after the cultural explosion, the speed of evolution is primarily determined by the speed of the dynamic *social process* governing the emergence of concepts. This speed may well exceed that of biological evolution.

As our goal was to develop a rather general theory of language, we had to introduce many technical simplifications. We considered a linear language system with well-defined constraints and we haven't explicitly modelled different levels of abstraction. In our analysis, we have only considered pure strategy equilibria, a strictly fixed number of concepts and we used one particular adjustment dynamics. In various extensions we have tried to explore the importance of some of these assumptions, while we argued for the strong validity of others. For example, we have ruled out mixed strategies as they cannot be interpreted for language. We have also explored other dynamics and found that our convergence results hold for other potential improving dynamics too as is generally believed (see Ermoliev and Flam [35], Hofbauer and Sigmund [26]). On a more general level, we have entirely ignored syntax from our analysis based on the broadly accepted argument that grammar is an "innate" capacity of humans, a result of biological evolution [36]. It is hard to imagine however, that the social evolution of concepts happens independently from grammar. We leave this and other interesting questions related to language for future research. However, we believe that to the extent language formation/usage is a social process, economics is likely to have an important role in explaining the related phenomena.

Appendix: Proof of Proposition 4

The proof boils down to the following Lemma:

Lemma 4 *Let $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ denote variables satisfying some constraints $g_j(\boldsymbol{x}) = 0$, $j = 1, \ldots, m$. Let $f(\boldsymbol{x}, c) = a(\boldsymbol{x}) + c\, b(\boldsymbol{x})$, with a, b continuously differentiable and $c \in \mathbb{R}$ a parameter, be a function that takes its local maximum at a point $\boldsymbol{x}^* = \boldsymbol{x}^*(c) = \mathrm{argmax}_x f(\boldsymbol{x}, c)$. Introducing $b^* = b(\boldsymbol{x}^*(c))$, we have*

$$\frac{db^*}{dc} \geq 0. \tag{48}$$

Proof: Let us introduce a set of lagrange multipliers $\boldsymbol{\lambda} = \{\lambda_j\}_{j=1}^m$ to treat the constraints, and define the Lagrangian function

$$L(\boldsymbol{\lambda}, \boldsymbol{x}) = a(\boldsymbol{x}) + c\, b(\boldsymbol{x}) + \sum_j^m \lambda_j g_j(\boldsymbol{x}). \tag{49}$$

Using the chain rule we can write

$$\frac{db^*}{dc} = \frac{\partial b}{\partial \boldsymbol{x}^*} \cdot \frac{\partial \boldsymbol{x}^*}{\partial c}. \tag{50}$$

and the task is to calculate the vector $\partial \boldsymbol{x}^*/\partial c$. A straightforward way is to proceed by a series expansion around a point $c_0 \to \boldsymbol{x}_0^*, \boldsymbol{\lambda}_0^*$. When c is perturbed $c = c_0 + \epsilon$, the local maximum and the value of the Lagrange multipliers shift too, $\boldsymbol{x}^* = \boldsymbol{x}_0^* + \epsilon\,\boldsymbol{\alpha}$, $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_0^* + \epsilon\,\boldsymbol{\beta}$. In order to determine $\boldsymbol{\alpha} \equiv \partial \boldsymbol{x}^*/\partial c$ and $\boldsymbol{\beta}$, we introduce general deviation variables $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ as

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x}_0^* + \epsilon\,\boldsymbol{\alpha} + \boldsymbol{\eta} \\ \boldsymbol{\lambda} &= \boldsymbol{\lambda}_0^* + \epsilon\,\boldsymbol{\beta} + \boldsymbol{\xi} \end{aligned} \tag{51}$$

and require that at the solution of the constrained optimization problem $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ should vanish; this will provide the necessary equations for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The Lagrangian can be expanded in a Taylor series in $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$. It is enough to consider terms $\mathcal{O}(\boldsymbol{\eta})$ and $\mathcal{O}(\boldsymbol{\xi})$ which should vanish. Moreover, these terms can be further expanded in ϵ. The coefficient of the $\mathcal{O}(\boldsymbol{\eta})$ term reads

$$\left[\frac{\partial a}{\partial \boldsymbol{x}} + c_0\,\frac{\partial b}{\partial \boldsymbol{x}} + \sum_j \lambda_{0j}^* \frac{\partial g_j}{\partial \boldsymbol{x}}\right] + \epsilon\Big[\frac{\partial b}{\partial \boldsymbol{x}} + \boldsymbol{G}_{\boldsymbol{x}}\boldsymbol{\beta} + \boldsymbol{L}_{\boldsymbol{x}\boldsymbol{x}}\boldsymbol{\alpha}\Big] + \mathcal{O}(\epsilon^2) \tag{52}$$

where $[\boldsymbol{G}_{\boldsymbol{x}}]_{ji} = \partial g_j/\partial x_i$ is the matrix of the first order derivatives of the constraints and $[\boldsymbol{L}]_{ii'} = \partial^2 L/\partial x_i \partial x_{i'}$ is the Hessian matrix at the point c_0. The first square bracket vanishes as this is the first order condition at c_0. The second square bracket implies

$$\frac{\partial b}{\partial \boldsymbol{x}} + \boldsymbol{G}_{\boldsymbol{x}}\boldsymbol{\beta} + \boldsymbol{L}_{\boldsymbol{x}\boldsymbol{x}}\boldsymbol{\alpha} = 0. \tag{53}$$

Similarly, the coefficient of the $\mathcal{O}(\boldsymbol{\xi})$ term can be expanded in ϵ, and we obtain

$$\boldsymbol{G}_{\boldsymbol{x}}\boldsymbol{\beta} = 0. \tag{54}$$

Equations (53) and (54) gives

$$\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\alpha} \end{pmatrix} = -\begin{pmatrix} \mathbf{0} & \boldsymbol{G}_{\boldsymbol{x}} \\ \boldsymbol{G}_{\boldsymbol{x}}^T & \boldsymbol{L}_{\boldsymbol{x}\boldsymbol{x}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \partial b/\partial \boldsymbol{x} \end{pmatrix} \tag{55}$$

where the hipermatrix is usually called the "bordered Hessian". This is necessarily negative semi-definite since we are at a local maximum. This allows us to express $\partial b^*/\partial c$ as a quadratic form

$$\frac{\partial b^*}{\partial c} = -\begin{pmatrix} \mathbf{0} \\ \partial b/\partial \boldsymbol{x} \end{pmatrix}^T \begin{pmatrix} \mathbf{0} & \boldsymbol{G}_{\boldsymbol{x}} \\ \boldsymbol{G}_{\boldsymbol{x}}^T & \boldsymbol{L}_{\boldsymbol{x}\boldsymbol{x}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \partial b/\partial \boldsymbol{x} \end{pmatrix} \geq 0 \tag{56}$$

which is thus necessarily non-negative. □

Proposition 4 follows as a direct corollary.

References

[1] Blume A., Kim Y.-G., Sobel J. Evolutionary stability in games of communication. *Games Econ. Behav.* 1993; 5(4):547–75.

[2] Blume A., DeJong D. V., Kim Y.-G., Sprinkle G. B. Experimental evidence on the evolution of meaning of messages in sender-receiver games. *Am. Econ. Rev.* 1998; 88(5):1323–40.

[3] Lachmann M., Számadó S., Bergstrom C. T. Cost and conflict in animal signals and human language. *P. Natl. Acad. Sci.* 2001; 98(23):13189–13194.

[4] Steels L. Self-organizing vocabularies. In Langton C., editor, *Proceedings of Artificial Life V* Nara, Japan 1996;.

[5] Nowak M. A., Plotkin J. B., Krakauer D. C. The evolutionary language game. *J. Theor. Biol.* 1999; 200:147–62.

[6] Nowak M. A. Evolutionary biology of language. *Phil. Trans. R. Soc. Lond. B* 2000; 355:1615–1622.

[7] Pinker S. Survival of the clearest. *Nature* 2000; 404:441–442.

[8] Rubinstein A. *Economics and Language*. Cambridge University Press Cambridge, UK 2000.

[9] Hurford J., Studdert-Kennedy M., Knight C. *Approaches to the Evolution of Language*. Cambridge University Press Cambridge 1998.

[10] Plotkin J. B. Nowak M. A. Language evolution and information theory. *J. Theor. Biol.* 1999; 205:147–59.

[11] Battigalli P. Maggi G. Rigidity, discretion and the cost of writing contracts. *Am. Econ. Rev.* 2002; 92:798–817.

[12] Cremer J., Garicano L., Prat A. Codes in organizations. Working Paper, MIT and CEPR 2003.

[13] Wernerfelt B. Organizational languages. *J. Econ. Manage. Strat.* 2004; 13(3):461–472.

[14] Armstrong S. L., Gleitman L. R., Gleitman H. On what some concepts might not be. *Cognition* 1983; 13:263–308.

[15] Croft W. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Univ. of Chicago Press Chicago 1991.

[16] Barsalou L. Ad hoc categories. *Mem. Cognition* 1983; 11(3):211–227.

[17] Barsalou L. Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *J. Exp. Psychol. Learn.* 1985; 11(4):629–654.

[18] Chater N. Vitnyi P. Simplicity: a unifying principle in cognitive science? *Trends Cogn. Sci.* 2003; 7:19–22.

[19] Mérő L. *Ways of thinking. The limits of rational thought and artificial intelligence.* World Scientific New Jersey 1990.

[20] Jennrich R. I. A simple general procedure for orthogonal rotation. *Psychometrika* 2001; 66:289–306.

[21] Fáth G. Sarvary M. A renormalization group theory of cultural evolution. *Physica A* 2005; 348:611–629.

[22] Fáth G. Sarvary M. Cultural evolution in a population of heterogeneous agents. In Namatame A., Kaizouji T., Aruka Y., editors, *Economics and Heterogeneous Interacting Agents* Lect. Notes Econ. Math. Springer 2005;.

[23] Jolliffe I. T. *Principal Component Analysis.* Springer New York 1986.

[24] Baldi P. Hornik K. Learning in linear networks: a survey. *IEEE T. Neural Networ.* 1995; 6:837–858.

[25] Monderer D. Shapley L. S. Potential games. *Games Econ. Behav.* 1996; 14:124–143.

[26] Hofbauer J. Sigmund K. *Evolutionary Games and Population Dynamics.* Cambridge University Press Cambridge 1998.

[27] Cressman R. *Evolutionary dynamics and extensive form games.* MIT Press Cambridge, Mass. 2003.

[28] Hofbauer J. Sigmund K. Evolutionary game dynamics. *Bull. Am. Math. Soc.* 2003; 40(4):479–519.

[29] Simon C. Blume L. *Mathematics for Economists.* W. W. Norton & Company New York 1994.

[30] Whorf B. L. Science and linguistics. In Carroll J., editor, *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf* pages 213–14. MIT Press Cambridge, MA 1956;. (reprinted).

[31] Wittgenstein L. *Philosophical Investigations.* Blackwell Oxford 1953.

[32] Ross P. E. Draining the language out of color. *Sci. Am.* 2004; April.

[33] Gordon P. Numerical cognition without words: Evidence from amazonia. *Science* 2004; 306:496–499.

[34] Mithen S. *The Prehistory of the Mind.* Thames and Hudson London 1996.

[35] Ermoliev Y. M. Flam S. Learning in potential games. Working Papers ir97022, International Institute for Applied Systems Analysis 1997.

[36] Pinker S. *The Language Instinct: How The Mind Creates Language.* Morrow New York, NY 1995.

Self-Organizing Communication In Language Games

A. Baronchelli[†], M. Felici[†], E. Caglioti[‡], V. Loreto[†] and L. Steels[♯]

† Physics Dept. and INFM-SMC *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
‡ Mathematics Dept. *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
♯ SONY CLS, 6 rue Amiot, 75005 Paris, FRANCE

Keywords:
Language games, Communication, Self-organization, Negotiation dynamics, Shared conventions, Social bookmarkings.

Abstract:
From the point of view of semiotic dynamics language is an evolving complex dynamical system. In this perspective, unrevealing the mechanisms that allow for the birth of shared conventions is a major issue. Here we describe a very simple model in which agents negotiate conventions and reach a global agreement without any intervention from the outside. In particular we focus on the possibility of predicting on which of the several competing conventions the agreement is reached. We find from simulations that early created conventions are favored in the competition process and this advantage can be quantified. Beyond the specific results presented here, we think that this paper provides an example of a new way of investigating language features where simple models allow for the investigation of precise problems and, possibly, for analytical approaches.

Corresponding Author:
Andrea Baronchelli
Physics Dept. and INFM-SMC *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
Tel: +39 06 4991 3450, e-mail: Andrea.Baronchelli@roma1.infn.it

Recently the study of the self organization and evolution of language and meaning has led to the idea that language can be seen as a complex dynamical system [1]. In this perspective, the theoretical tools developed in complex systems science acquire a central role for the study of the self generating structures of language systems.

Language is a semiotic system that relates signs (words, grammatical constructions) to the world through the intermediary of conceptualization. Such a semiotic system can then be used for communication in the sense, for example, that signs can be used to draw attention to objects of the physical world. Once relaxed the hypothesis of staticity of a language, a natural and very interesting question is how new conventions, developed in local interactions among few individuals, can become stable in the whole population. Said in different words, the problem is to determine the behavior of a system of many components, endowed with an individual structure, that interact with each others. The analogy with problems of n-body systems of interacting particles, well known in physics, is then obvious and stimulating.

Here we discuss an extremely simple multi-agent model, the Naming Game [Baronchelli A, Felici M, Caglioti E, Loreto V and Steels L: unpublished 2005], in which agents play pairwise games in order to negotiate conventions, i.e. associations between forms and meanings. In this way the attention is focused only on cultural spreading, without resorting to any evolutionary issues [2, 3].

In particular we deal with a population of N agents whose aim is to agree on the name to give to a certain object (the presence of a single object corresponds to the denial of homonymy). Each agent is characterized by its *inventory*, i.e. a list of words that can be dynamically updated. Agents have empty inventories at time $t = 0$ and at each time step ($t = 1, 2, ..$) two players are picked at random to play an interaction: one of them plays as *speaker* and the other as *hearer*. Their interaction obeys the following rules:

- The speaker randomly extracts a word from its inventory, or, if its inventory is empty, invents a new word.
- If the hearer has the word selected by the speaker in its inventory, the interaction is a success and both players maintain in their inventories only the winning word, deleting all the others.
- If the hearer does not have the word selected by the speaker in its inventory, the interaction is a failure and the hearer updates its inventory adding the new word.

The model is able to describe the emergence of a communication system where a unique form (or name) is assigned by all the individuals to the same meaning (here the object). This is clearly showed in Figure 1, where the evolution of total number of words present in the system is shown. At

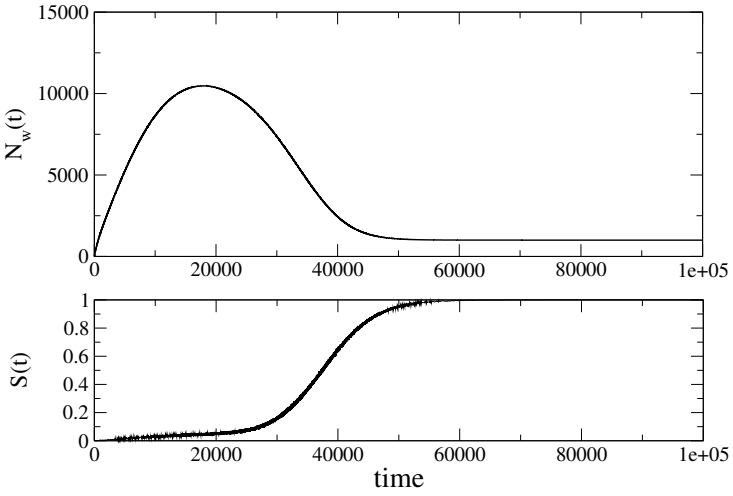


Figure 1: **System evolution:** We show here the evolution in time of the total number of words $N_w(t)$ and the success rate $S(t)$ for a population of $N = 1000$ agents. As it is clear the population is able to build up an effective communication system in which $N_w = N$, i.e. each agent has only one word. Moreover the success rate in this state is equal to 1 indicating that all interactions are successful.

the beginning the curve grows due to the invention process. It follows a further period of growth in which agents perform unsuccessful interactions. While increasing in size, however, inventories correlate with each other, so that at a certain point the number of words, having reached a maximum, starts decreasing. This is due to successful communications between agents, that increase in their number till the system reaches a final convergence state in which all the agents have the same unique word, thus being able to perform only successful interactions. In Figure 1 it is reported the success rate curve too. Obtained averaging over several different runs of the process, this curves shows the probability that an interaction at a given time is successful. As just discussed, the system evolves to a situation of convergence through an intermediate state in which words are eliminated while the success probability increases. It is important noting that the developed communication system is not only effective (every agent understands all the others) but also efficient (no memory or computational resources are wasted for comprehension in the final state).

Given that the population does, at the end, agree on a convention, an interesting question is whether we can predict on *which* convention the agreement takes place. In fact, the same process in which agents negotiate with each other can be seen as a process in which different conventions compete to survive. According to the rules of our model all different words are equiv-

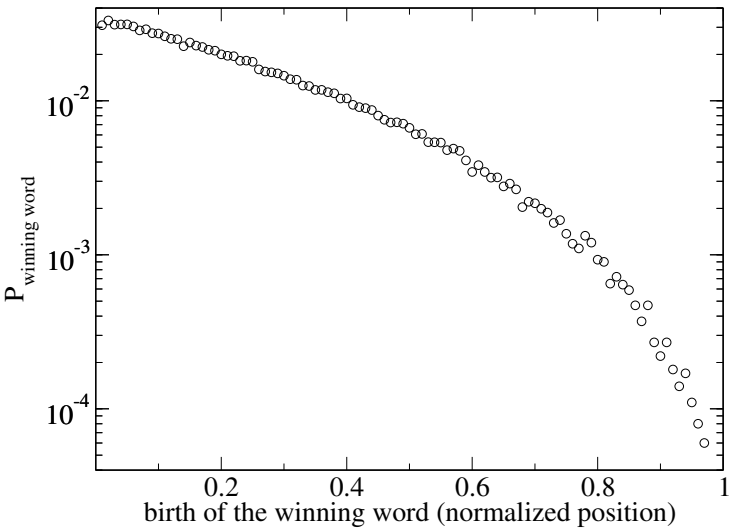


Figure 2: **Winning word probability distribution (order):** We investigate how the probability for a word to dominate is affected by its invention order. We label the first invented word with 1, the second invented word with 2 and so on till the last invented one. Then labels are normalized with the tag of the last invented word. Early invented words are more likely to become dominant since they have more chances to propagate at the beginning. Data are obtained for a population of $N = 1000$ agents by averaging results from 10^5 runs.

alent. The only feature that could differentiate them is their invention moment. In Figure 2 we investigate the role of the invention *order*. It emerges that the probability for a word to become the one the agents will agree upon strongly depends on the moment of its creation, indeed. To investigate the role of creation order we label the first invented word with cardinal number 1, the second invented word with 2, and so on till the last invented word. Then we divide the label of each word by the label of the last invented one so that results from different simulation runs become comparable. Performing several runs and taking memory of the label of the winning words, we are then able to estimate the probability that the winning word has a given label. From Figure 2, it clearly emerges that early invented words have bigger chances to dominate. This can be explained considering that the sooner a word is invented the higher are its opportunities to propagate. Moreover at the beginning of the game agents have small inventories, and this reduces the probability that an existing word is not played by a speaker who holds it. Finally, in Figure 3 we analyze the role of invention *time*. Obviously, the advantage of early invented words is found again, but interestingly in the domain of time this advantage can be quantified. Indeed, in Figure 3 it is

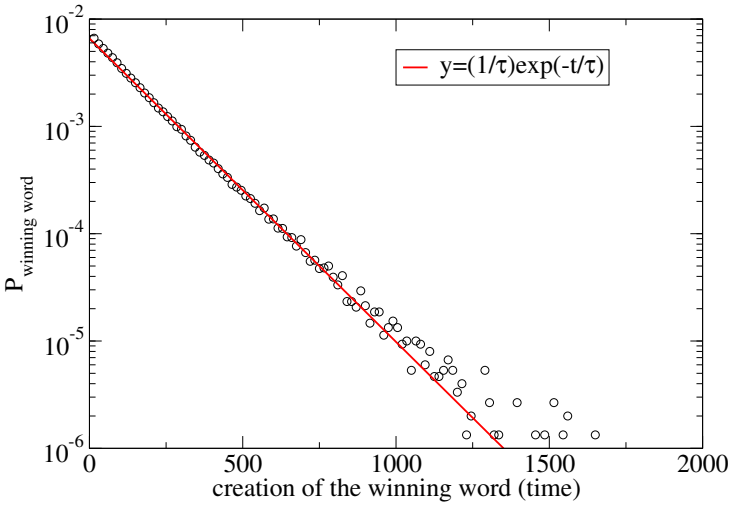


Figure 3: **Winning word probability distribution (time):** We plot here the probability that a word created at a given time becomes the dominating one. Data, relative to a population of $N = 1000$ agents, are well fitted by an exponential distribution (here $\tau \simeq 150$).

shown that data from simulations are well fitted by an exponential distribution. In Figure 2 this behavior was not found due to the presence of a cutoff (the last invented word) that is absent here.

In conclusion, we have seen how the view of language as an evolving systems casts several interesting issues that can be addressed in the framework of complex science methods. A profitable approach consists, in our opinion, in the definition and study of simple models that allow for precise investigations of specific problems. Then, in this perspective, the definition of a reasonable model is in itself an important goal. Here we have presented an interesting model in which agents negotiate conventions according to elementary rules and manage to reach a global agreement. We have also shown that, even though all words (or conventions) are equivalent in the model, the moment of their invention affects (exponentially) their probability of becoming dominant.

References

[1] Steels L: Language as a complex adaptive system. In Shoenauer M (ed): Proceedings of PPNS VI. Lecture Notes in Computer Science, Berlin, Germany, September 2000. Springer-Verlag.

[2] Hurford J: Biological evolution of the Saussurean sign as a component of the language acquisition device, Lingua 1989; 77:187-222.

[3] Nowak MA, Plotkin JB and Krakauer D: The evolutionary language game, J. Theor. Biol. 1999; 200:147-162.

When language breaks into pieces. A conflict between communication through isolated signals and language.

Ramon Ferrer i Cancho[1,*]

(1) Dip. di Fisica. Università "La Sapienza". Piazzale A. Moro 5. 00185 Roma, Italy.

Running tittle: When language breaks into pieces.

Abstract

Here we study a communication model where signals associate to stimuli. The model assumes that signals follow Zipf's law and the exponent of the law depends on a balance between maximizing the information transfer and saving cost the cost of communication. We study the effect of tuning that balance on the structure of signal-stimulus associations. The model starts from two recent results. First, the exponent grows as the weight of information transfer increases. Second, a rudimentary form of language is obtained when the network of signal-stimulus associations is almost connected. Here we show the existence of a sudden destruction of language once a critical balance is crossed. The model shows that maximizing the information transfer through isolated signals and language are in conflict. The model proposes a strong reason for not finding large exponents in complex communication systems: language is in danger. Besides, the findings suggests that human words may need to be ambiguous to keep language alive. Interestingly, the model predicts that large exponents should be associated to decreased synaptic density. It is not surprising that the largest exponents correspond to schizophrenic patients since, according to the spirit of Feinberg's hypothesis, decreased synaptic density leads to schizophrenia. Our findings suggests that the exponent of Zipf's law is intimately related to language and that it could be used to detect anomalous structure and organization of the brain.

Keywords: Zipf's law, communication, human language, syntax, symbolic reference, schizophrenia

(*) Contact information.
Ramon Ferrer i Cancho. Dip. di Fisica, Università "La Sapienza". Piazzale A. Moro 5. 00185 Roma, Italy.
e-mail: ramon@pil.phys.uniroma1.it
Phone: +39 064455705 / +39 064452045 (ext. 46-50)
FAX: +39 064463158

1 Introduction

The XX century witnessed the birth and development of information theory [1, 2], a theoretical framework devoted to the study of communication systems. Recently, various new models have been introduced to explain the organization of word frequencies in human language using an information theory approach [3, 4, 5]. Word frequencies in human language obey a universal regularity, the so-called Zipf's law [6]. If $P(f)$ is proportion of words whose frequency is f in a text, we obtain

$$P(f) \sim f^{-\beta} \tag{1}$$

where we typically have $\beta \approx 2$. Eq. 1 is a way of defining Zipf's law. The models mentioned above assume a system where signals from a set S communicate about stimuli from a set R. Signals are equivalent to words and stimuli are the basic ingredients of word meaning. For instance, the word 'dog' is associated to visual stimuli (e.g. the shape of a dog), auditive stimuli (e.g. barking),... All these stimuli are elicited by the word 'dog' [7]. Stimuli are sometimes called objects or events in the origins of language literature (e.g [8, 9]). Those models assume a set of n signals $S = \{s_1, ..., s_i, ..., s_n\}$ and a set of m stimuli $R = \{r_1, ..., r_j, ..., r_m\}$. Signals link to stimuli and connections are defined by an $n \times m$ binary matrix $A = \{a_{ij}\}$ where $a_{ij} = 1$ if s_i and r_j are linked and $a_{ij} = 0$ otherwise.

According to Shannon's standard theory [1], the goal of communication through isolated signals is maximizing $I(S, R)$, the information transfer between S and R. One of the most important contributions of the models above is that Zipf's law with non-extremal exponents can not be explained by maximizing $I(S, R)$ alone, which would lead to $\beta \to \infty$. Zipf's law with exponents close to the typical values are obtained when $I(S, R)$ is maximized with a further constraint. $H(S)$, the entropy of signals has been shown to be, as far as we know, the best candidate for that constraint [3, 10, 5]. It is known in psycholinguistics that the availability a word is positively correlated with its frequency. The higher the frequency of a word, the higher its availability. That is the so-called word frequency effect [11]. That frequency dependent availability concerns both the speaker and the hearer of a conversation. Imagine we have n words (or signals). When all words are equally likely, that is, when all words have frequency $1/n$, all words are taking the smallest frequency possible. In that case, $H(S) = \log n$, where $\log n$ is the maximum value of $H(S)$ [2]. In contrast, when a word has frequency one (which implies that the remaining words have frequency zero), $H(S) = 0$, which is the minimum value of $H(S)$ [2]. $H(S)$ is a measure of the cost of communication. The higher the value of $H(S)$ the higher the cost (and the lower the word availability). Notice that computers do not have the same information access and retrieval constraints of human brains. In general, information is accessed at a very high speed and frequency effects, when present, are not so heavy as those imposed by the human brain. One can, in general neglect $H(S)$ in many computer problems but not in brain word access and retrieval.

If we restrict ourselves to Shannon's classic information theory, the goal of

a communication system is maximizing the function

$$\Omega_0 = I(S, R) \tag{2}$$

If we take into consideration the cost of communication, we may write

$$\Omega = \lambda I(S, R) - (1 - \lambda)H(S) \tag{3}$$

as the function that a natural communication system should maximize [3, 10, 5]. λ is a parameter controlling the balance between maximizing the information transfer and minimizing the cost of communication. We assume $\lambda \in [0, 1]$. We have $\Omega_0 = \Omega$ when $\lambda = 1$. Ω_0 is suitable for computer or robotic problems where $H(S)$ can be neglected and Ω (with $\lambda < 1/2$ [10, 5]) is specially suitable for brain based communication systems. Ω seems, a priori, a better choice than Ω_0 for natural communication systems.

We do not claim that Ω is the best function for natural communication systems but there are some results supporting its usefulness:

- Maximizing Ω, Zipf's law is obtained for a particular value of λ. If one replaces $H(S)$ in Eq. 3 by the the effective lexicon size, namely, the number of signals with at least one association with stimuli, Zipf's law is not obtained [3, 5]. Vocabulary size is an important factor for the cost of communication [12] but does not seem to be essential for Zipf's law. Zipf's law is still reproduced if $I(S, R)$ is replaced by $H(R|S)$ in the model in [3] but not in the model in [5].

- The exponent of Zipf's law in single author text satisfies $\beta \in [1.6, 2.42]$ [10]. Maximizing Ω in a system following Zipf's law (i.e. searching the value of β maximizing Ω) can explain the interval of variation of β in human language [10].

If one considers texts from a single author [13, 14] and does concentrate on words of a certain type (e.g. nouns) [15, 4], the extremes of the interval of variation of β correspond to schizophrenic patients [10]. The aim of the present paper is deepening our understanding of what may happen when β is large and in particular, what may be happing in schizophrenics with that β. We will show that that language breaks into pieces when the balance between maximizing $I(S, R)$ and minimizing $H(S)$ favours too much the former. More precisely, we will show that the network of signal-interactions becomes suddenly disconnected when λ takes a critical value in a communication system following Zipf's law.

2 The model

Maybe the simplest approach for reproducing Zipf's law for word frequencies is combining two assumptions. First,

$$P(k) \sim k^{-\beta}, \tag{4}$$

where $P(k)$ is the probability that a signal has k connections. Second, $p(s_i) \sim \mu_i$, where $p(s_i)$ is the probability of using s_i and

$$\mu_i = \sum_{j=1}^{m} a_{ij}. \tag{5}$$

Eq. 4 and $p(s_i) \sim \mu_i$ give Eq. 1. Various models recover Zipf's law when maximizing Ω without the constraint in Eq. 4 for a critical value of λ [3, 5].

Going further, we assume

$$p(s_i) = \frac{\mu_i}{M}, \tag{6}$$

where

$$M = \sum_{i=1}^{m} \mu_i \tag{7}$$

is the total amount of links. Assuming Eq. 6 has the virtue of simplicity and allowing one to explain the interval of variation of β in humans. Interestingly, Eq. 6 makes some important assumptions that need to be explicited. To that aim, let us start from a general assumption about $p(s_i, r_j)$, the joint probability of s_i and r_j, namely

$$p(s_i, r_j) = \frac{a_{ij} p(r_j)}{\omega_j}, \tag{8}$$

where $p(r_j)$ is the probability of the j-th stimulus and

$$\omega_j = \sum_{k=1}^{n} a_{kj} \tag{9}$$

is the number of links of that stimulus.

If we assume

$$p(r_i) = \frac{\omega_i}{M}, \tag{10}$$

and replace it in Eq. 10 Eq. 8 we obtain

$$p(s_i, r_j) = \frac{a_{ij}}{M}. \tag{11}$$

Replacing Eq. 11 into $p(s_i) = \sum_{j=1}^{m} p(s_i, r_j)$ we recover Eq. 6. The models in [10, 5] assume Eq. 10 (hence assume Eq. 6 and 11). In contrast, the model in [3] assumes that $p(r_j)$ is constant for each j and considers a particular case, i.e. $p(r_j) = 1/m$.

We define $H(R|S)$ as the conditional entropy of stimuli when signals are known. We may write Eq. 3 as

$$\Omega = -\lambda H(R|S) - (1 - \lambda) H(S) \tag{12}$$

as in the model in [3] when $p(r_j)$ is constant. That is not the case of the present article and related models [10, 5].

The assumption $p(r_j) \sim \omega_j$ means that the probability of each stimulus is dictated by the structure of signal-stimulus associations. In other words, the probability of perceiving $p(r_j)$ in the 'real world' is neglected. One may think that is a very radical assumption but in fact, human language is a communication tool allowing one to detach from the here and now. Nonetheless, displaced reference, our ability to talk about something that is distant in time or space, is a salient feature of human language [16, 17], Displaced reference is not uniquely human since bees have it [18]. Because of displaced reference, we can talk of 'dogs' even when there is no 'dog' in front of us. It seems wise to assume that talking about present stimulus is not the rule of human language and it seems that in some cases such as schizophrenia, the detachment from the here and know could be extreme. Various core aspects of schizophrenia such as false believes, hallucinations [19] and various cognitive impairments, including attention problems [20], suggest that interacting with the 'real world' is difficult. In fact, schizophrenics seem optimal candidates for $p(r_j) \sim \omega_j$. Schizophrenics speakers are a very special case in the results that will follow. We will return to them in the discussion.

For the present article, we assume a communication system following Zipf's law by means of Eq. 6. The distribution of links per signal is given by $P = \{P(1), ..., P(k), ..., P(m)\}$ and the distribution of links per stimulus is given by $Q = \{Q(0), ..., Q(k), ..., Q(n)\}$, where $Q(k)$ is the probability that a stimulus has k links. We are assuming that $Q(k)$ is defined for $k = 0$ while $P(k)$ does not because we allow unlinked stimuli but do not allow unlinked signals. Here we take the simplest distribution for Q as in [10], that is

$$Q \sim binomial\left(\frac{\langle k\rangle_P}{m}, n\right), \tag{13}$$

where $\langle ...\rangle_P$ is the expectation operator over P. Thus, $\langle k\rangle_P$ is the mean signal degree. We may define the information theory measures that matter in the calculation of Ω assuming $p(r_j) \sim \omega_j$ (or $p(s_i) \sim \mu_i$) for any pair of P and Q. The calculation of Ω is straightforward once we know [10, 5]

$$H(S) = \log M - H(R|S) \tag{14}$$

$$H(R) = \log M - H(S|R) \tag{15}$$

where $M = n\langle k\rangle_P = m\langle k\rangle_Q$ and

$$H(R|S) = \frac{\langle k\log k\rangle_P}{\langle k\rangle_P} \tag{16}$$

$$H(S|R) = \frac{\langle k\log k\rangle_Q}{\langle k\rangle_Q}. \tag{17}$$

The present model integrates two recent results. The first result is that β^*, the value of β maximizing Ω grows with λ till $\lambda = \lambda^*$. Then, we have $\beta \to \infty$ for $\lambda > \lambda^*$ [10]. The behavior of β^* is illustrated in Fig. 1. It can be shown that

$\lambda^* < 1/2$ and a heuristic argument suggests the existence of a discontinuity at $\lambda = \lambda^*$ [10]. The idea is very simple. Eq. 3 can be written as

$$\Omega = (2\lambda - 1)\lambda H(S) - \lambda H(S|R) \tag{18}$$

knowing that $I(S, R) = H(S) - H(S|R)$ [2]. Eq. 18 indicates that maximizing Ω minimizes $H(S)$ if $\lambda < 1/2$ and maximizes $H(S)$ if $\lambda > 1/2$. $\beta \to \infty$ minimizes $H(S|R)$ and maximizes $H(S)$ (recall Eq. 17 and 14), so $\beta \to \infty$ is expected for $\lambda > 1/2$. Since maximizing Ω for $\lambda = 0$ gives a finite value of β^* (Fig. 2), β must diverge for $0 < \lambda \leq 1/2$. Notice that maximizing Ω for $\lambda = 0$ is equivalent to minimizing $H(S)$, the signal entropy.

The second result is that a communication system gets a rudimentary form of language if the bipartite network of signal-stimulus associations is connected or almost connected [9]. Roughly speaking, connectedness is the possibility of starting from a signal (or a stimuli) and reaching the remaining signals and stimuli of the network crossing the links of the network. Fig. 3 A and B show, respectively, an almost connected and a disconnected bipartite networks. Almost connectedness means that a wide majority of vertices (e.g. 90%) lay in the largest connected component [9]. When exponents are close to the real ones, it has been shown that Zipf's law provides almost connectedness under a general set of conditions [9]. Connectedness is intimately related to two essential traits that researchers have identified as essential aspects of human language: syntax and symbolic reference [21]. Signal-stimulus associations allow one to define signal-signal associations. More importantly, the network of signal-stimulus association specifies allowed and forbidden signal-signal associations. Taking the example of words, we can explain why the syntactic combination of "drive cars" is a sensible combination in the sentence "John drives cars" and why it is not the combination "drives onions" in the sentence "John drives onions". The combination of 'drive' and 'car' in "John drives cars" exemplifies the relationship between a verb and its argument. As in [9], we adopt the convention that two signals (or two words) s_i and s_k can be combined syntactically if and only if they are linked to at least one common stimulus, that is, if $\xi > 0$ where

$$\xi_{ik} = \sum_j a_{ij} a_{jk} \tag{19}$$

The idea behind $\xi_{ij} > 0$ is that s_i and s_k must be semantically compatible. If s_i ='drive' and s_k ='car' we would have $\xi > 0$ and if s_i ='drive' and s_k ='onion' we would have $\xi = 0$ (Fig. 4). If two signals are linked to the same stimulus it does not mean that the signals are synonyms since stimulus here are not meanings but components of meaning. The meaning of 'drive' is linked among others, to the visual, tactile,... experiences of driving, the objects that can be driven,... whereas 'car' is associated to the visual shape of a car, the action of driving, ... The fact that 'drive' and 'car' share one or more stimuli does not mean that 'drive' and 'car' are synonyms. When the network of signal-object associations is connected, we have that for every signal there is at least another signal sharing stimuli. We could also define a network of signal-signal

associations defined by a binary $n \times n$ matrix $B = \{b_{ik}\}$ where $b_{ik} = 1$ if $\xi_{ik} > 0$ and $b_{ik} = 0$ otherwise. B is a rudimentary syntactic network where vertices are words and two words are linked if the can be combined syntactically [22]. The properties of real syntactic networks have been studied at the global [22] and sentence level [23]. The small-word phenomenon and heterogenous degree distribution have been reported at the global level. In a system following Eq. 4 with $\beta \approx 2$, the signal degree distribution in B has a power tail with the same exponent [9], which is consistent with the degree distribution of real syntactic networks [22].

In Peirce's view, there are three ways in the which words and objects of the 'world' can associate: iconically (by similarity), indexical (by spatial or temporal cooccurrence) or symbolically (by convention) [24]. According to Deacon, an essential aspect of symbolic reference is that real words not only evoke stimuli (or meanings) but also other words [24]. Deacon tried to define symbolic reference but his proposal has been criticized due its lack of precision [25, 26]. Taking the idea of 'signals evoking other signals', Ferrer i Cancho and colleagues have defined symbolic reference as connectedness in the network of signal-stimulus associations [22]. The definition is not ambiguous and relies on standard concepts of graph theory [27]. When a network is connected, one may start from a certain signal and reach its first neighbours (stimulus) and from them one can get to the second neighbours (signals). One can continue from 2nd neighbours to 3rd, 4th, and so on till all the signals and stimulus in the network have been reached. We define L, the normalized size (in number of vertices) of the largest connected component, as

$$L = \frac{l}{n+m}, \tag{20}$$

where l is the number of vertices in the largest connected component and $n+m$ is the total amount of vertices. L is a measure of the expressive power of the rudimentary language emerging from signal-object associations. If $L = 1$ then all signals can be combined in a grammatically correct discourse. If $L < 1$ then that is possible only for a fraction of signals. We will show that L is controlled by λ.

3 Results

For each value of λ,

- We obtained β^*, value of β maximizing Ω, exploring $\beta \in [0, 10]$ with a resolution $\epsilon = 0.1$.
- We calculated the mean value of L in random bipartite network where signal degree follows Eq. 4 with $\beta = \beta^*$. Links with stimuli are formed chosing stimuli at random (all stimuli are equally likely so Eq. 13 follows). Means were calculated over 1000 replicas.

Fig. 5 shows the evolution of a small network of signal-stimulus associations as λ grows. At a critical value of λ, the size of the largest connected component falls

abruptly. In general, L falls abruptly to a small value for $\lambda = \lambda^*$ (Fig. 6). λ^* is the point where β diverges and $I(S,R)$ and $H(S)$ reach their maximum value [10]. The steepness of the fall grows with n. Fig. 7 illustrates what happens to L, $I(S,R)$ and $H(S)$ at the same time.

4 Discussion

We have seen that a communication system maximizing Ω undergoes an abrupt transition to disconnectedness for $\lambda > \lambda^*$. We have seen that the transition is caused by a sudden transition from a finite value of β to $\beta \to \infty$, where the chance that a stimulus has two links vanishes as m grows. The disconnection of the network when $\beta \to$ is easy to understand. In general, a unipartite graph with N vertices and M edges cannot be connected if $M < M^*$ where $M^* = N - 1$ is the number of edges of a tree of N vertices [27]. Thus, a bipartite graph with $N = n + m$ vertices cannot be connected if $M < n + m - 1$. In other words, connectedness is not possible if $\langle k \rangle_P < (n + m + 1)/n$. When $\beta \to \infty$, we have $\langle k \rangle_P = 1$ and $\langle k \rangle_P < (n + m + 1)/n$ holds trivially provided $m > 1$. In sum, connectedness is impossible for $\beta \to \infty$ and $m > 1$.

In Section 2, we have reviewed a heuristic argument suggesting the transition from a highly connected phase to disconnectedness in our model is discontinuous. Discontinuous phase transitions are widespread in nature. For instance, the melting of ice into water or the transformation of boiling water into vapour are discontinuous in normal circumstances. In a communication context, the models in [3, 5] show a continuous phase transition between no communication and a perfect communication phase when Ω is minimized with no constraint on P. There, the presence of Zipf's law in the vicinities of the phase transition is the hallmark of a continuous phase transition. In contrast, the phase transition from disconnectedness to connectedness in a classic Erdös-Rényi graph [28, 29] is continuous [30, 31]. The hallmark of continuous phase transition in classic unipartite graphs is a power distribution of connected component sizes [30], which is related to a critical branching process [32] at the threshold for connectedness. Other examples of continuous phase transitions are the transition from resistivity to superconductivity (continuous in the absence of an external magnetic field) and the conversion of iron from paramagnetic to ferromagnetic form [33]. In a communication context, the model examined here shows not only apparently shows a discontinuous transition to disconnectedness but also to maximum information transfer and maximum cost for $\lambda = \lambda^*$ [5].

The divergence of β for $\lambda = \lambda^*$ is accompanied by a jump to maximum information transfer (Fig. 7). Increasing λ increases $I(S,R)$ but decreases the size of the largest connected component (the significance of the decrease depends on the size of the system). At the point where the $I(S,R)$ is maximum, L is minimum. In a communication system maximizing Ω, communication using isolated signals and language are in conflict. Human speakers may need to regulate λ in order to maximize information transfer but avoid reducing the size of the largest connected component too much. Interestingly, the regulation of the size

of the largest connected component can be done indirectly because increasing $I(S,R)$ also increases $H(S)$, the cost of communication. Word ambiguity may not be a mere defect but a requirement for connectedness and thus language. Our findings suggests a possible scenario for the origins of language. Reducing λ (giving more weight to minimizing $H(S)$) maximizes the chance of connectedness. The emergence of connectedness could be a side effect of saving the cost of communication.

A theory of word frequencies needs answering different questions:

1. Why do words arrange themselves according to Zipf's law (Eq. 1)?
2. Why do humans choose some particular values of β?
3. Why is there variation in β?
4. What are the limits of that variation?
5. What is the link between Zipf's law and human language?

Many answers have been proposed for Questions 1-2 [5]. As far as we know, Question 1 and 2, have only been answered assuming that words are used according to their meaning in [3, 4, 5]. Choosing values of β near 2 could be an optimal solution for a conflict between maximizing the information transfer and saving the cost of communication [3, 5]. Questions 3 and 4 have begun to be addressed in [10, 4]. The idea is that the lower bound and the upper bounds of β are obtained when maximizing Ω for $\lambda = 0$ and $\lambda = \lambda^*$, respectively. The present article sheds new light on Questions 3, 4 and 5. As for 3, variation in β my be due to the chance of connectedness. As for 4, it has been argued that the variation of β is constrained by the fact that maximizing Ω for $\lambda \in [0,1]$ gives a narrow interval of exponents [10]. It has been argued that the interval of variation of β excludes $\beta \to \infty$ because the maximum cost, i.e. $H(S) = \log n$, is paid in that case. The argument has some drawbacks. $H(S) = \log n$ is a slow growing function of n. In practice, significant difference in $\log n$ for two different systems can only be obtained if the respective values of n differ in at least one order of magnitude. In order to explain why $\beta \to \infty$ is not found, one has to argue that speakers, in general, are very sensitive to the variation of $\log n$, which we do not know. Instead, one may propose a stronger argument: $\beta \to \infty$ is not found because the chance of connectedness is 0 for $m > 1$ (as seen above). That is a compelling reason for not finding large β in human language. We do not mean that large β is impossible to attain in humans, but it would be surprising to find it in a system combining words through semantic constraints. In sum, the present article puts another step forward in the construction of a theory of word frequencies. As for 5, our work suggests that the exponent of Zipf's law is an important factor for the presence or absence of language.

The largest values of β than have been found up to now in single author text samples correspond to schizophrenic patients in the acute phase of the illness [10, 35]. One of the most salient features of schizophrenia is 'disorder of thought' [36]. Disorder of thought may be described as disturbances in the

structure, organization and coherence of thought that are reflected in reduced intelligibility and increased disorganization of speech that is difficult, if not impossible, for the listener to comprehend [37]. Our model makes two relevant predictions for the case of schizophrenics. First, the chance of being on the edge of an abrupt transition grows with the value of β, so schizophrenics with large exponents may be threatened by an apparently discontinuous phase transition where language breaks into pieces. Second, if n is small, the decrease in the size of the largest connected component with λ (and therefore β) is significant (recall Fig. 6). The larger the value of β, the smaller the size of the largest connected component. Both predictions are apparently consistent with the appearance of thought disorder in schizophrenia. It is hard to imagine how a schizophrenic can construct a coherent discourse if the size of the largest connected component has dramatically decreased.

The network of signal-stimulus associations is an emergent structure of the neural substrate. Integrating stimuli of various kinds with words implies connecting distant neural tissues. In order to have an example of mind, visual and temporal stimuli tend to be related to occipital and temporal areas of the human brain [7]. It is reasonable to think that the density of synapsis has an influence on the largest connected component of the network of signal-stimulus associations. Thus, β, specially for small n, can be seen as an indicator of the size of the largest connected component, which would be in turn an indicator of the density of the neural substrate. The link density of the network of signal-stimulus associations is $\delta = \langle k \rangle_P / m$. It can be easily seen that $\langle k \rangle_P$ decreases with β (see Appendix). For large m and $\beta > 2$ we have (see Appendix)

$$\langle k \rangle \approx \frac{1-\beta}{2-\beta}. \tag{21}$$

If our hypothetical correspondence between β and synaptic density (or size in words of the largest connected component) was correct, one would expect that the smallest synaptic density would be for the largest values of β, which corresponds to schizophrenic patients in the acute phase [35]. Interestingly, it has been speculated that excessive synaptic pruning occurs in schizophrenia, which may lead to psychosis when it reaches a threshold [19, 38, 39]. Our work is consistent with the spirit of Feinberg's hypothesis, relating the onset of schizophrenia to a critical decrease in synaptic density [40]. We do not mean that a critically low synaptic density is the only possible cause of schizophrenia and that reduced synaptic density must always originate through the exact mechanisms that Feinberg proposed. Instead, we claim it is not surprising that large exponents belong to schizophrenic patients since those exponents predict a decreased synaptic density, which is an important factor that may lead to schizophrenia [41, 19]. Our work suggests that the exponent of Zipf's law could be used to detect synaptic density alterations and more importantly, disconnected brain areas.

The model presented here indicates a track for other related cases. While schizophrenics with large exponents seem to face the problem of the destruction of connectedness, children seem to face an inverse problem, i.e. the development

of connectedness. The relatively short time elapsed from the single-word to multiple-word utterances (of the order of several months [34]), suggests that the emergence of syntactic communication in children could be a phase transition to connectedness in the network of word syntactic interactions. According to our model of a rudimentary form of language, that transition would be an epiphenomenon of a transition to connectedness in the network of signal-stimuli associations. Whether the presumable phase transition would be continuous or not would depend of the presence or not of a special signature: scaling in the distribution of connected component sizes in the network of word syntactic interactions. We know that the network of syntactic interactions of adults is (almost) connected [22] but the signature above may be found in children at a critical time. To sum up, our findings open new research prospects and support that Zipf's law is an essential aspect of human language.

Appendix

We assume k is a random discrete variable whose probability is

$$P(k) = ck^{-\beta} \tag{22}$$

where β is a constant and

$$c = \frac{1}{\sum_{k=1}^{m} k^{-\beta}} \tag{23}$$

is a normalization term. $\langle k \rangle$, the mean value of k is

$$\langle k \rangle = c \sum_{k=1}^{m} k^{1-\beta}. \tag{24}$$

We can approximate $\langle k \rangle$ replacing summations by integrals and write

$$\langle k \rangle \approx \frac{\int_1^m k^{1-\beta} dk}{\int_1^m k^{-\beta} dk}. \tag{25}$$

Solving the integrals we obtain

$$\langle k \rangle \approx \frac{(1-\beta)(m^{2-\beta}-1)}{(2-\beta)(m^{1-\beta}-1)}. \tag{26}$$

For $m \to \infty$ and $\beta > 2$, we get

$$\langle k \rangle \approx \frac{1-\beta}{2-\beta}. \tag{27}$$

Acknowledgments

We are grateful to Vito Servedio for helpful comments. Discussions with Toni Hernández and Brita Elvevåg have had a great influence for the arguments of

that article. This work was funded by the ECAgents project (RFC), funded by the Future and Emerging Technologies program (IST-FET) of the European Commission under the EU RD contract IST-1940. The information provided is the sole responsibility of the authors and does not reflect the Commission's opinion. The Commission is not responsible for any use that may be made of the data appearing in this publication.

References

[1] C. E. Shannon, Bell Systems Technical Journal **27**, 379 (1948).

[2] R. B. Ash, *Information Theory* (John Wiley & Sons, New York, 1965).

[3] R. Ferrer i Cancho and R. V. Solé, Proc. Natl. Acad. Sci. USA **100**, 788 (2003).

[4] R. Ferrer i Cancho, Physica A **345**, 275 (2004), doi:10.1016/j.physa.2004.06.158.

[5] R. Ferrer i Cancho, Submitted to Eur. Phys. J. B (2005).

[6] G. K. Zipf, *Human behaviour and the principle of least effort. An introduction to human ecology* (Hafner reprint, New York, 1972), 1st edition: Cambridge, MA: Addison-Wesley, 1949.

[7] F. Pulvermüller, *The neuroscience of language. On brain circuits of words and serial order* (Cambridge University Press, Cambridge, 2003).

[8] M. A. Nowak, Phil. Trans. R. Soc. Lond. B **355**, 1615 (2000).

[9] R. Ferrer i Cancho, O. Riordan, and B. Bollobás, Proc. R. Soc. Lond. Series B **272**, 561 (2005).

[10] R. Ferrer i Cancho, Eur. Phys. J. B **44**, 249 (2005).

[11] A. Akmajian, R. A. Demers, A. K. Farmer, and R. M. Harnish, *Linguistics. An Introduction to Language and Communication* (MIT Press, 1995).

[12] R. Köhler, Theor. Linguist. **14**(2-3), 241 (1987).

[13] R. Ferrer i Cancho and R. V. Solé, J. Quantitative Linguistics **8**(3), 165 (2001).

[14] M. A. Montemurro, Physica A **300**, 567 (2001).

[15] V. K. Balasubrahmanyan and S. Naranan, J. Quantitative Linguistics **3**(3), 177 (1996).

[16] N. Chomsky, Paper presented at the Univ. de Brasilia, Nov. 26 (1996).

[17] C. F. Hockett, *A course in modern linguistics* (McMillan, New York, 1958).

[18] K. von Frisch, Scientific American **207**, 79 (1962).

[19] K. T. Mueser and S. R. McGurk, The Lancet **363**, 2063 (2004).

[20] B. Elvevåg and T. E. Goldberg, Critical Reviews in Neurobiology **14**, 1 (2000).

[21] C. Knight, M. Studdert-Kennedy, and J. R. Hurford, in *The evolutionary emergence of language* (Cambridge University Press, Cambridge, 2000), pp. 303–322.

[22] R. Ferrer i Cancho, R. V. Solé, and R. Köhler, Physical Review E **69**, 051915 (2004).

[23] R. Ferrer i Cancho, Physical Review E **70**, 056135 (2004).

[24] T. W. Deacon, *The Symbolic Species: the Co-evolution of Language and the Brain* (W. W. Norton & Company, New York, 1997).

[25] J. Hurford, The Times Literary Supplement (October 1998).

[26] R. Hudson, Journal of Pragmatics (August 1999).

[27] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics (Springer, New York, 1998).

[28] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[29] B. Bollobás, *Random graphs*, Cambridge Studies in Advanced Mathematics, vol. 73 (Cambridge University Press, 2001), 2nd Edition.

[30] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[31] V. E. Stepanov, Theory Probab. Appl. **15**, 187 (1970).

[32] T. Harris, *The theory of branching processes* (Prentice-Hall, N.J, 1963).

[33] J. Binney, N. Dowrick, A. Fisher, and M. Newman, *The theory of critical phenomena. An introduction to the renormalization group* (Oxford University Press, New York, 1992).

[34] C. J., H. Davis, and M. Macken, in A. Lock and C. R. Peters, eds., *Handbook of human symbolic evolution* (Blackwell, Oxford, 1999), pp. 686–746.

[35] R. G. Piotrowski, V. E. Pashkovskii, and V. R. Piotrowski, Automatic Documentation and Mathematical Linguistics **28**(5), 28 (1995), first published in Nauchno-Tekhnicheskaya Informatisiya, Seriya 2, Vol. 28, No. 11. pp. 21-25, 1994.

[36] B. Elvevåg and T. Goldberg, CNS Spectrums **2** (1997).

[37] E. Breuler, *Dementia Praecox or the group of schizophrenia* (International University Press, New York, 1911/1950).

[38] G. M. Innocenti, F. Ansermet, and J. Parnas, Mol. Psychiatry **8**, 261 (2003).

[39] M. S. Keshavan, S. Anderson, and J. Pettegrew, J. Psychiatric Res. **28**, 239 (1994).

[40] I. Feinberg, J. Psychiatric Res. **17**, 319 (1982).

[41] P. J. Harrison, Current Opinion in Neurobiology **7**, 285 (1997).

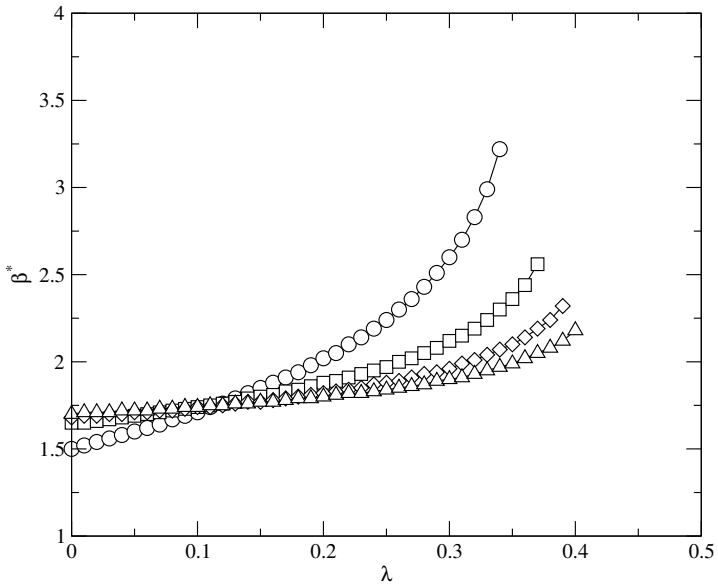


Figure 1: β^*, the value of β maximizing Ω for $n = m = 10$ (circles), $n = m = 10^2$ (squares), $n = m = 10^3$ (diamonds) and $n = m = 10^4$ (triangles). β is the exponent of Zipf's law, Ω is the energy function that communication maximizes, n is the number of signals and m is the number of stimuli. λ tunes the balance between information transfer and cost of communication. When $\lambda = 0$ communication is totally balanced towards saving the cost of communication whereas it is totally balanced towards information transfer when $\lambda = 0$.

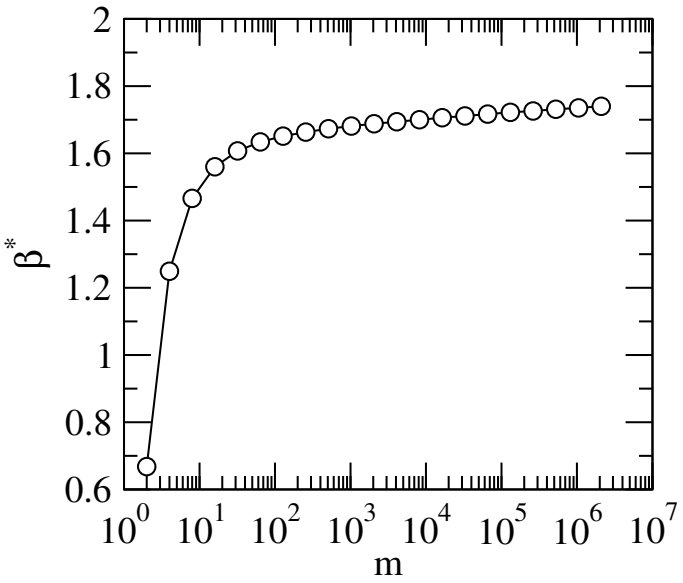


Figure 2: β^*, the value of β minimizing $H(S)$ versus m. $H(S)$ is the signal entropy and m is the number of stimuli.

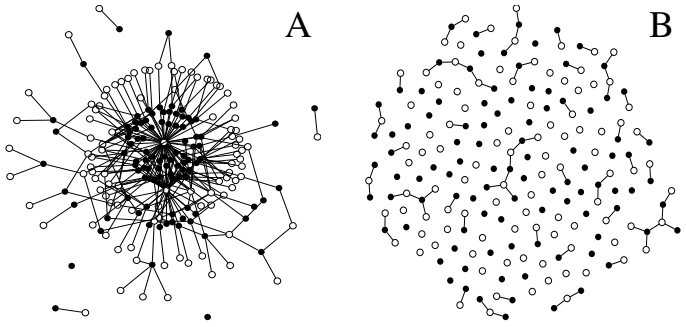


Figure 3: Two bipartite networks. White and black are used for each vertex partition. A. An almost connected networks. B. A disconnected network.

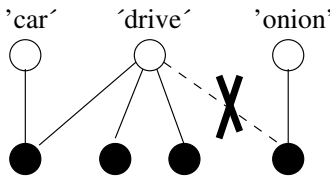


Figure 4: A possible implementation of the constrains of the verb 'drive' with possible arguments 'car' and 'onion'. White circles are words and black circles are stimuli. 'car' is an allowed argument of the verb 'drive' and therefore there is a link between 'drive0 and a stimulus associated to 'car'. 'onion' is not a valid argument of 'drive', so no stimulus linked to 'drive' is linked to 'onion'. ξ_{ik}, the number of shared stimulus by the pair (s_i, s_k), is 1 for ('drive','onion') and 0 for ('drive','onion').

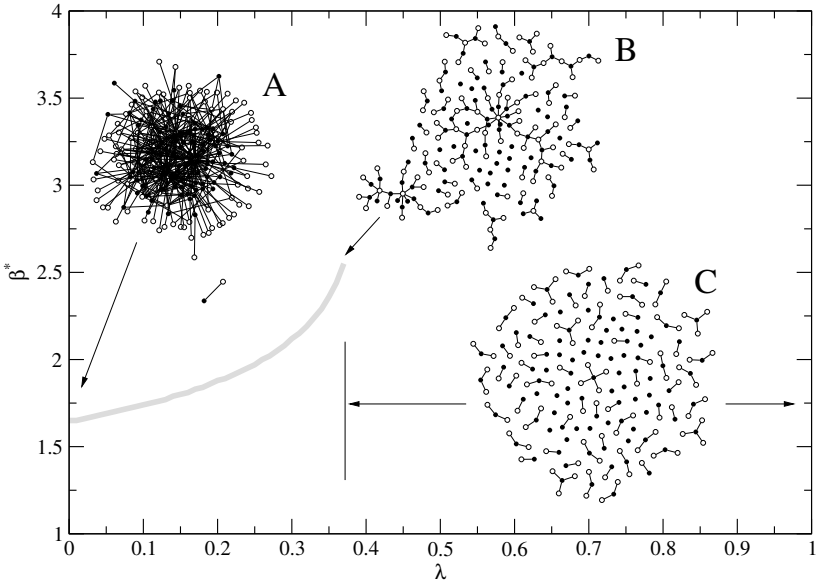


Figure 5: The evolution of β^* versus λ (gray curve) and the structure of network of signal-stimulus associations with $n = m = 100$. White and black circles indicate, respectively, signals and stimuli. The curve for β^* ends at the point of divergence at $\lambda = \lambda^* \approx 0.37$. A, B and C are examples of the kind of the topologies found for $\lambda = 0$, $\lambda = \lambda^*$ and $\lambda > \lambda^*$.

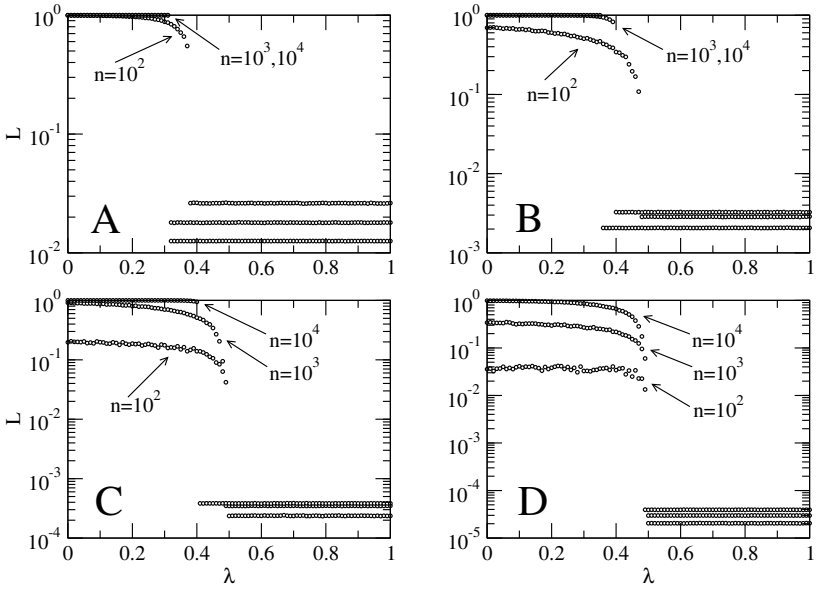


Figure 6: L, the normalized size of the largest connected component versus λ, the parameter controlling the balance between $I(S,R)$ and $H(S)$ in Ω. The connected component size is measured in vertices. n is the number of signals and m is the number of objects. A. $m = 10^2$. B $m = 10^3$. C. $m = 10^4$. D. $m = 10^5$.

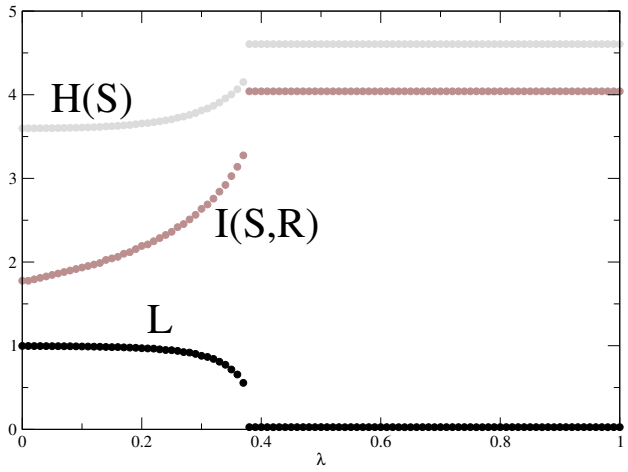


Figure 7: An example of the behavior of L (black), the normalized size in vertices of the largest connected component, $I(S,R)$ (dark gray), the information transfer and $H(S)$ (light gray), the signal entropy, versus λ, the parameter regulating the balance between maximizing $I(S,R)$ and $H(S)$ in Ω. A sudden change of behavior is found for $\lambda \approx 0.37$. $n = m = 100$ was used.

The self-organization of combinatorial vocalization systems

Pierre-Yves Oudeyer
*Sony CSL Paris, 6, rue Amyot, 75005 Paris, France
http://www.csl.sony.fr/ py

Abstract

In previous papers, we presented a system which showed how a society of agents could self-organize a shared discrete vocalization system, starting from holistic inarticulate vocalisations. The originality of the system was that: 1) it did not include any pressure for communication; 2) it did not include any social capacity (agents did not play a language game for example); 3) it pre-supposed neither linguistic capacities nor the existence of conventions. We present here an extension of the system which shows how rules of sound combination as well as patterns of combinations can self-organize and be shared by the society of agents. This illustrates how phonotactics might have bootstrapped.

1 Introduction

Human vocalizations have a complex organization. They are discrete and combinatorial: they are built through the combination of units, and these units are systematically re-used from one vocalization to the other. These units appear at multiple levels (e.g. the gestures, the coordination of gestures, the phonemes, the morphemes). While the articulatory space that defines the physically possible gestures is continuous, each language only uses a discrete set of gestures. While there is a wide diversity of the repertoires of these units in the world languages, there are also very strong regularities (for example, the high frequency of the 5 vowel system /e,i,o,a,u/). The way the units are combined is also very particular : 1) not all sequences of phonemes are allowed in a given language (this is its phonotactics), 2) the set of allowed phoneme combinations is organized into patterns. This organization into patterns means that for example, one can summarize the allowed phonemes of Japanese by the patterns "CV/CVC/VC", where "CV" for example defines syllables composed of two slots, and in the first slot only the phonemes belonging to a group that we call "consonant" are allowed, while in the second slot, only the phonemes belonging to the group that we call "vowels" are allowed.

It is then obvious to ask where this organization comes from. There are two complementary kinds of answers that must be given (Oudeyer, 2003). The first kind is a functional answer stating which is the function of systems of speech sounds, and then showing that systems having the organization that we described are efficient for achieving this function. This has for example been proposed by (Lindblom, 1992) who showed that discretenesss and statistical regularities can be predicted by searching for the most efficient vocalization systems. This kind of answer is necessary, but not sufficient : it does not say how evolution (genetic or cultural) might have found this optimal structure. In particular, naive Darwinian search with random mutations (i.e. plain natural selection) might not be sufficient to explain the formation of this kind of complex structures : the search space is just too large (Ball, 2001). This is why there needs a second kind of answer stating how evolution might have found these structures. In particular, this amounts to show how self-organization might have constrained the search space and helped natural selection. This can be done by showing that a much simpler system spontaneously self-organizes into the more complex structure that we want to explain.

(Oudeyer, 2005) has shown how a system of this kind, based on the coupling of generic neural devices which were innately randomly wired and implanted in the head of artificial agents, could self-organize so that the agents develop a shared vocalization system with discreteness, combinatoriality and statistical regularities. The originality of the system was that: 1) it did not include any pressure for communication; 2) it did not include any social capacity (agents did not play a language game for example); 3) it pre-supposed neither linguistic capacities nor the existence of conventions. We present now an extension of this system which gives an account of the formation of rules of sound combination as well as of patterns of sound combinations. This amounts to the formation of phonotactics. The extension is based on the

addition of a map of neurons with temporal receptive fields. These are initially randomly pre-wired, and control the sequential programming of vocalizations. They evolve with local adaptive synaptic dynamics.

2 The system

We are going to make a summary of the architecture presented in details in (Oudeyer, 2005), before presenting the extension. The system is composed of agents which are themselves composed of an artificial brain connected to an artificial vocal tract and an artificial ear. Agents can produce and hear vocalizations. As described in (Oudeyer, 2005), one can model each component from the most abstract to the most realistic manner. In this paper, our goal is to explore the principles of the formation of phonotactics and of phonological patterns, rather than to build a realistic predictive model. Thus, we will use the most abstract version of the components presented in (Oudeyer, 2005). In particular, this means that agents produce two-dimensional vocalizations (one articulatory dimension and one temporal dimension). We use only one space to represent vocalizations: the perceptual space is bypassed and only the motor space is used. So, we pre-suppose that agents can translate a vocalization from the perceptual space to the motor space, which is acceptable since in (Oudeyer, 2005) we showed how this mapping could be learnt by the agents. The articulatory dimension that we use is also abstract, but one could imagine that it represents the place or the manner of constriction for example. Finally, the agents are put in a virtual space in which they wander randomly, and at random times they generate vocalizations which are heard by themselves as well as the closest agent.

The brain of the agent is organized into two neural maps: 1) one "spatial" neural map coding for static articulatory configurations; 2) one "temporal" neural map coding for the sequences of activations of the neurons in the static neural map (this constitutes the extension of the system presented in (Oudeyer, 2005)).

2.1 The spatial neural map

The spatial neural map contains neural units N_i which have broadly tuned gaussian receptive fields. We denote $f_{p,i}$ the centre of the gaussian, which we call its "preferred vector" since it corresponds to the stimulus which activates maximally the neural unit. All the neural units have initially a random preferred vector, following a uniform distribution. Each neural unit codes for an articulatory configuration, defined by the value of its preferred vector. If the neural unit is activated by the agent and a GO signal is sent to the neural map, then there is a low-level control system which drives the articulators continuously from the current configuration to the configuration coded by the activated neuron. A vocalization is thus here a continuous trajectory in the articulatory space, produced by the successive activation of some neural units in the spatial neural map, combined with a GO signal. As we will see later on, this activation is controlled internally by the temporal neurons.

As we explained earlier, we use only one space to represent vocalizations. Thus, when an agent produces a vocalization, defined by its trajectory in the articulatory space, the agents that can perceive this vocalization gets directly the trajectory in the articulatory space. The perception of one vocalization produces changes in the spatial neural map. The continuous trajectory is segmented in small samples corresponding to the cochlea time resolution, and each sample serves as an input stimulus to the spatial neural map. The receptive fields of each neural units adapt to these inputs by changing their preferred vector (the width of the gaussian does not evolve). For each input, the activation of each N_i is computed, and their receptive field updated so that if the same stimulus comes again next time, it will respond a little bit more (this is weighted by their current activation). Basically, adaptation is an increase in sensitivity to stimuli in the environment.

2.2 The temporal neural map

In (Oudeyer, 2005), the production of vocalizations was realized by activating randomly neurons in the spatial map. There was no possibility to encode the order in which the neurons were activated, and as a consequence agents ended by producing vocalizations in which all phoneme combinations were allowed (but of course only the phonemes that appeared as a result of the self-organization of the neural map were used). On the contrary, we will use here a temporal neural map which can encode the order of activations of spatial neurons, and is used to activate the spatial neurons.

Each temporal neuron is connected to several spatial neurons. A temporal neuron can be activated by the spatial neurons through these connections. The tuning function of temporal neurons has a temporal dimension: their activation depends not only on the amplitude of the activation of the spatial neurons to which they are connected, but depends also on the

order in which they are activated, which itself depends on the particular vocalization which is being perceived.

As stated in the first paragraph, the temporal neurons are also used to activate the spatial neurons. The internal activation of one temporal neuron, coupled with a GO signal, provokes the successive activation of the spatial neurons to which it is connected. Here, the temporal pattern is regular, and only one neuron is activated at the same time. In this paper, each temporal neuron will be connected to only two spatial neurons, which means that a temporal neuron will code for a sequence of two articulatory targets (the order is coded by some internal parameters of the temporal neuron). This will allow us to represent easily the temporal neural map (but this is not crucial to the result). When an agent decides to produce a vocalization, it activates randomly one temporal neuron and sends a GO signal.

Initially, a high number of temporal neurons are created (500), and are connected randomly to the spatial map with random values of their internal parameters. Using many neurons makes that basically all possible sequences of activations of spatial neurons are encoded in the initial temporal neural map. The plasticity of the temporal neurons is different from the plasticity of spatial neurons. The parameters of temporal neurons stay fixed during the simulations, but they can die. As a consequence, what changes in the temporal neural map is the number of surviving neurons. The neuron death mechanism is inspired from apoptosis (Ameisen, 2000), and fits with the theory of neural epigenesis developed by (Changeux, 1983). The theory basically proposes that neural epigenesis consists of an initial massive generation of random neurons and connections, which are afterwards pruned and selected according to the level of neurotrophines they receive. Neurotrophins are provided to the neurons which are often activated, and prevent them from automatic suicide. We apply this principle of generation and pruning to our temporal neurons, and depending on their mean activity level. There is a $vitalThreshold$ constant which defines the level of activity below which the neuron is pruned. This threshold remains the same for all neurons in the map. The value of this threshold is chosen so that there is not enough potential activity for all the neurons to stay alive: stability arises at the map level only after a certain amount of neurons have been pruned.

2.3 The coupling of perception and production

The crucial point of this architecture is that the same neural units are used both to perceive and to produce vocalizations, both in the spatial and in the temporal neural map. As a consequence, the distribution of targets which are used for production is the same than the distribution of receptive fields in the spatial neural map, which themselves adapt to inputs in the environment. This implies for example that if an agent hears certain sounds more often than others, he will tend to produce them also more often than others. The same phenomenon applies also to the order of the articulatory targets used in the vocalizations. If an agent hears certain combinations often, then this will increase the mean level of activation of the corresponding temporal neurons, which in turn increases their chance of survival and so increases the probability that they will be used to produce the same articulatory targets combinations. These coupling create positive feed-back loops which are the basis of the self-organization that we will now describe.

3 The dynamic formation of phonotactics and patterns of combinations

In these simulations, we use a population of 20 agents. As initially the preferred vectors of the spatial neurons are random, and as there is a massive number of random temporal neurons, agents produce vocalizations which are holistic and inarticulate: the continuum of possible articulatory targets is used, and nearly all possible sequences of targets are produced. The initial state of both neural map in two agents is represented on figure 1: the spatial map is represented on the x-axis, which shows the preferred vectors, and is also represented on the y-axis, which shows the same information. The temporal map is represented by the small segments in the middle of the figure, which all correspond to a point (x, y) for which x corresponds to an existing preferred vector, and y to another existing preferred vector. The x coordinate of a temporal neuron corresponds to the first target that it encodes, and the y coordinate corresponds to the second target that it encodes. The length of the segment represents the level of neurotrophins that each neuron possess.

After several hundred time steps, as we have shown and explained in details in (Oudeyer, 2005), we observe a clustering of the preferred vectors of the spa-

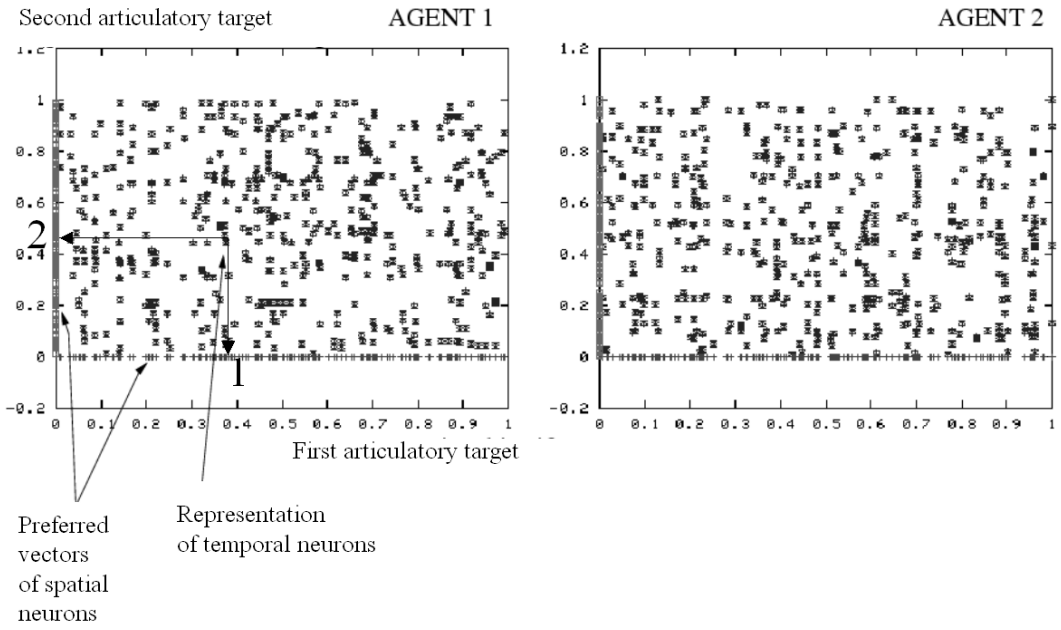


Figure 1: The neural maps of two agents at the beginning of the simulation. The neural maps of one agent is represented on the left, and the neural maps of the other agent are represented on the right. The spatial map is represented by its preferred vectors plotted on the x-axis and also plotted on the y-axis. The temporal neural map is represented by small segments whose center has its x and y corresponding to preferred vectors of the spatial neural map. The x coordinate of a temporal neuron corresponds to the first target that it encodes, and the y coordinate corresponds to the second target that it encodes.

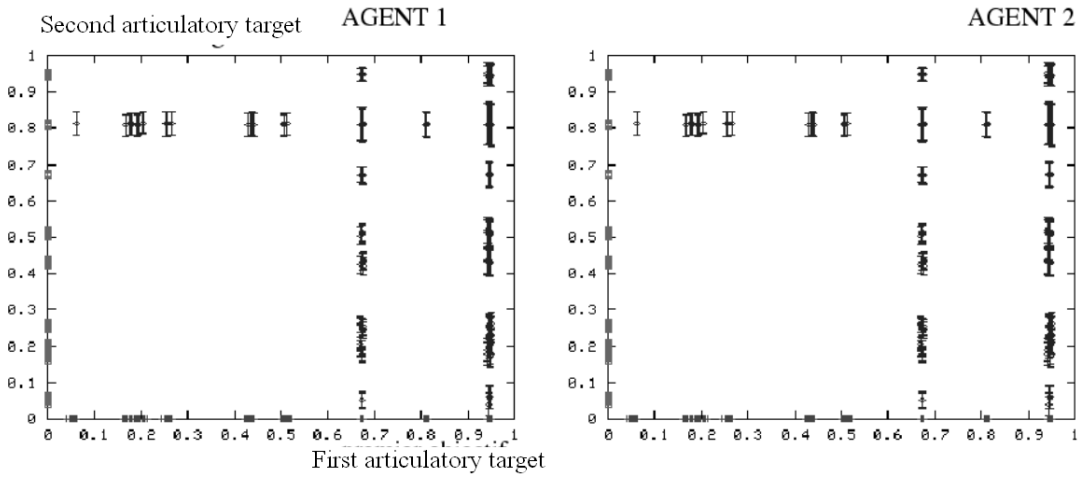


Figure 2: The neural maps of the same two agents after 1000 interactions. We observe: 1) that the preferred vectors of the spatial neural map are now clustered, which means that vocalizations are now discrete: phonemic coding has appeared; 2) that many temporal neurons have died and the surviving ones are organized into lines and columns: this means that phonotactic rules have appeared and moreover that the repertoire of vocalization can be organized into patterns.

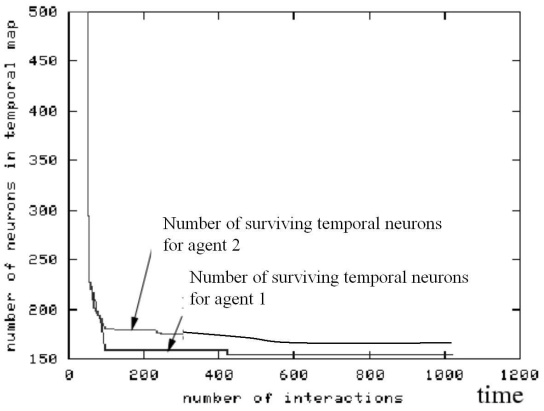


Figure 3: Evolution of the number of surviving temporal neurons corresponding to the temporal neural map of the two agents in figure 2. We observe that there is a first phase of massive pruning, followed by a stabilization which corresponds to a convergence of the system.

tial map. Figure 2 shows an example of the neural maps after 1000 interactions in two agents. Moreover, the clusters are the same for all the agents of the same simulation, and different for agents of different simulations. This shows that now the vocalizations that they produce are discrete: the articulatory targets that they use belong to one of several well defined clusters, and so the continuum of possible targets has been discretized.

Moreover, if we observe the temporal map, we discover that there remains only temporal neurons coding for certain articulatory target sequences. This means that some sequences of targets made possible by the clusters are not produced any more. All the agents of the same population share not only the same clusters in the spatial map, but they also share the same surviving temporal neurons. This means that rules of phoneme sequencing have appeared, which are shared by all the population. In brief, this is the self-organization of phonotactics. Yet, this is not all that we can observe from the temporal neural map. We also see that the surviving temporal neurons are organized into lines and columns. This means that the set of allowed phoneme sequences can be summarized by patterns. If we call the phonemes associated with the eight clusters of the spatial map $p_1, p_2, ..., p_8$, then we can summarize the repertoire of allowed sequences by: $(p_6, *)$, $(p_8, *)$ et $(*, p_7)$ where $*$ means "any phoneme in $p_1, ..., p_8$". The repertoire is thus organized into patterns, in a manner similar for example to the "CV/CVC/VC" organization of syllables in Japanese. As figure 3 shows, this state is the final state of the system: there is a crystallization of the repertoire of vocalizations.

References

J.C. Ameisen. *La Sculpture du vivant. Le suicide cellulaire ou la mort cratrice.* Seuil, 2000.

P. Ball. *The self-made tapestry, Pattern formation in nature*. Oxford University Press, 2001.

J.P. Changeux. *L'homme neuronal*. Fayard, 1983.

B. Lindblom. Phonological units as adaptive emergents of lexical development. In Ferguson, Menn, and Stoel-Gammon, editors, *Phonological Development: Models, Research, Implications*, pages 565–604. York Press, Timonnium, MD, 1992.

P-Y. Oudeyer. *L'auto-organisation de la parole*. PhD thesis, Université Paris VI, 2003.

P-Y. Oudeyer. The self-organization of speech sounds. *Journal of Theoretical Biology*, 233(3): 435–449, 2005.

Stability conditions in the evolution of compositional languages: issues in scaling population sizes

Paul Vogt

Induction of Linguistic Knowledge group, Tilburg University, the Netherlands
Language Evolution and Computation unit, University of Edinburgh, UK.

Abstract

This paper investigates the effect of scaling the population size in a simulation studying the emergence and evolution of compositionality in languages. The simulations are based on multi-agent systems that play language games in order to communicate, invent and learn language. The language games are integrated with the iterated learning model that simulates a population turnover, where the population contains adults and children. Experiments show that when the population size is increased, after an initial decrease in performance, the results show an important improvement when the population size is increased further. These results are explained by a hypothesised trade off between increasing difficulties in achieving a conventionalised system and an increased likelihood of finding structures that emerge by chance when the population size increases.[1]

Keywords: Language evolution, compositionality, language games, iterated learning model

1 Introduction

In the past decade an increasing number of models on the evolution of language have been studied, see, e.g., [1, 2] for overviews. Most of these model are based on the assumption that language is a *complex adaptive dynamical system* [3]. The idea behind this assumption is that *language* itself adapts culturally to its users as if it is a complex dynamical system, which contrasts the nativist approach in which language users adapt to use and acquire language. Central to this paradigm is that language is thought to arise through self-organisation as a result of cultural interactions among individuals and individual learning.

One important area of research has focused on the emergence and evolution of compositional languages, i.e. languages with a simple grammar in which parts of its expressions map onto parts of its meanings. The models that assume language is a complex adaptive dynamical system can – with respect to language transmission – be divided into two classes: those with a vertical transmission of language (such as the *iterated learning model* [4]) and those with a horizontal transmission (such as the *language game model* [5]). In the iterated learning model (ILM), language evolves and is transmitted from one generation to the next by allowing the adult agents to speak only to children agents who learn from the adults. In the language game model, there is typically no generational turnover, so all agents are within one generation and all agents are equally likely to speak and hear. In a recent study by Vogt [6], the two models have been combined in a simulation where there was a generational turnover as in the ILM, but where all agents could speak to each other as in the language game model. In the standard ILM, compositional structures can only emerge when the experimenter imposes a *bottleneck* on the transmission, i.e. the children only learn from a subset of the adults' languages. Vogt has shown that this bottleneck need not be imposed when children are also allowed to speak (i.e. when there is a combination of horizontal and vertical language

[1]Address for correspondence: ILK / Computational Linguistics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Email: *paulv@ling.ed.ac.uk*, phone: +31 13 4663578.

S -> aba/$[1_r, 0_g, 0_b, 1_s]$
S -> A/rgb B/s
A -> a/$[1_r, 0_g, 0_b]$
A -> db/$[0_r, 0_g, 1_b]$
B -> ka/$[1_s]$

Table 1: An example grammar. The grammar consists of simple rewrite rules, where non-terminal nodes (S, A and B) rewrite to a word-meaning pair (rules 1, 3–5) or to a compositional structure (rule 2). In between the square brackets are the rules' meanings, which are categorical features that relate to some feature dimension (r, g, b and s), such as the red, green and blue components of the RGB colour space and a shape feature.

transmission), because the children implicitly face a bottleneck when producing utterances about previously unseen meanings [6].

The limitation of the studies carried out so far on the emergence of compositionality is that the populations were rather small. For instance in ILMs, most studies were carried out with only 2 agents (but see [7, 8] for population sizes up to 6 and 20 agents respectively). In Steels' language model, the studies also contained only 2 agents, and in Vogt's [6] study only 6 agents were present at each time. Although studies with more agents were carried out (e.g., in [9] there were 50 agents), in such studies each agent was restricted to learn only from a limited number of other agents.

In the recently started New Ties project (see http://www.new-ties.org), we aim at evolving language and culture in societies with populations of more than 1,000 agents [10]. This paper reports on a pilot study done with the model introduced in [7]. In particular, the study investigates the effect of increasing the population size up to 100 agents on the emerge of compositionality in simulations with the combined vertical and horizontal language transmission presented in [6].

2 The model

The model is implemented in a toolkit, called THSim [11], which simulates the Talking Heads experiment [12]. The aim of the experiment is for a population of agents to develop a shared compositional communication system to communicate about the coloured geometrical shapes that make up the population's world. As mentioned, the current model combines the iterated learning model with the language game model. As part of the learning mechanisms, some grammar induction techniques are adapted that allow agents to acquire compositional structures. Below follows a very brief description of the model, for more details, consult [6, 7].

The simulation implements a multi-agent system, where at each instance in time the population of agents contains a given number of adults and children. After the population has played a given number of language games, the adults are removed, the children become adults and new children are introduced. This process then repeats. Effectively, this population flow implements the iterated learning model. In each language game, both the speaker and hearer are selected randomly from the entire population. So, unlike in the standard versions of the ILM, where all speakers are adults and all hearers are children such that the language is transmitted vertically (see, e.g., [4, 7]), the language in this model is transmitted horizontally as in [6].

Each agent starts their lifetime without any categories or linguistic knowledge. All categories and linguistic knowledge are acquired by playing language games. The linguistic knowledge is stored by the agents in a private grammar. These grammars contain two types of rules: holistic rules (such as the first rule in Table 1) and compositional rules (such as the second rule in Table 1). In holistic rules, no part of an expression is analysed in terms of any part of its meaning. In compositional rules, parts of an expression can be analysed in terms of distinct meaning parts. (In the current model, the agents can only acquire compositional rules that take two constituents.)

In the current study, the agents play a particular variant of the language game, called the *guessing game*. A guessing game is played by two agents: a speaker and a hearer. The aim of the

game is for the hearer to guess the object the speaker is referring to using a verbal expression.

In a nutshell, the game is organised as follows: Both agents look at a scene displayed on one of the monitors. This scene (or context) consists of 8 randomly sampled coloured geometrical shapes. The agents then form categories such that each object is categorised distinctively from each other object in the context. If categorisation fails, the agents can construct new categories based on the feature descriptions of the particular objects. The speaker selects one object as the target and tries to produce an expression describing this target. The production is based on the rules that are in the grammar. If no expression can be produced, a new expression is created by constructing a random string of characters, which is used to construct (a) new rule(s). The hearer then tries to interpret the expression by searching its own grammar for a composition of rules that parses the expression and that is consistent with one of the categories that relate to the current context. The hearer then points at the object it guesses is the target. If this is the correct guess, the speaker acknowledges this success. Otherwise, the speaker points at the intended target, thus providing corrective feedback.

Both speaker and hearer can find multiple ways to either produce or interpret an expression. In such cases, the agents select the way that has been most effective in the past based on the weights that are attached to the rules. These weights are adapted based on the outcome of the game. If the game is a success, the weights of used rules are increased, while the weights of competing rules that could also be selected are decreased. If the game is a failure, the weights of used rules are decreased and the hearer adopts the expression with the meaning of the intended target. While adopting the expression, the hearer first tries to see if it can be generalised by breaking up the expression and meaning in two parts, thus forming a compositional rule. If this is impossible, the hearer adopts the rule holistically.

Breaking up an expression and meaning can only be done if the hearer had previously heard one or more expression-meaning pairs that share some similarity with this new pair. For instance, consider the following simplified example. If an agent previously heard the expression-meaning pair `ab-11` and then hears the expression-meaning pair `ac-10`, it can break up these pairs forming rules such as `S -> A/x B/y`, where `A/x -> a/1`, `B/y -> b/1` and `B -> c/0`. After a hearer has broken up the expression-meaning, it will perform two post-operations to generalise the language even further and to remove any redundancies.

Summarising, the guessing game implements communicative interactions between two agents. When parts of the game fail, the agents can construct or otherwise acquire new categories or linguistic knowledge. All linguistic knowledge is mediated by weights indicating the effectiveness of used knowledge. When such weights are high, the rules have been used effectively. When agents need to select among two or more competing rules, they select one using the *winner takes all* principle. This way, effective rules will be reselected more frequently and the individual grammars will tend to become shared across the population through self-organisation in a similar way ant paths are formed.

3 Experimental results

As mentioned in the introduction, this paper investigates the scalability of the current model. To this aim, a number of simulations were done where the population size N was increased from 10 agents to 50 agents with steps of 10, and one additional set of simulations with $N = 100$. At each moment, half of the population were adult agents, and the other half were child agents. In all simulations, the model was run for 100 iterations (or generations) of 214,600 guessing games each. (This odd number was derived for earlier experiments, but since one simulation takes 1–3 weeks to process on a modern PC, no attempt was made to redo the experiments with a nicer round value.) For all different population sizes, 10 trials were carried out for statistical analysis.

At the end of each simulation, the final population was tested on the language they evolved. During this test phase, all agents in the population were presented with the same 200 situations (a context with 8 objects, including a given target). In each situation, each agent had to produce

N	C	> 0.6	> 0.7	> 0.8	> 0.9	CA
10	0.82 ± 0.03	10	10	9	0	0.91 ± 0.03
20	0.81 ± 0.02	10	10	6	0	0.84 ± 0.04
30	0.58 ± 0.25	6	3	3	1	0.63 ± 0.13
40	0.41 ± 0.31	3	3	3	1	0.52 ± 0.21
50	0.47 ± 0.36	4	4	4	3	0.50 ± 0.26
100	0.76 ± 0.36	8	8	7	7	0.70 ± 0.29

Table 2: This table summarises the results. Column 1 gives the population size N, column 2 the gives the level of compositionality C reached at the end of the simulation, averaged over 10 trials plus their standard deviations. Columns 3–6 give the number of trials which yielded a level of compositionality higher than the value indicated in the top row. The final column presents the average level of communicative accuracy CA reached at the end of the simulations together with their standard deviations.

an expression about the target, which in turn each other agent had to interpret. All learning was switched off during the test periods from which two measures were calculated: *compositionality* and *communicative accuracy*. Compositionality C measures the proportion of expressions that were produced or interpreted using a compositional rule. Communicative accuracy CA measures the rate with which agents could successfully interpret each other.

Table 2 summarises the results of this experiment. The second column provides the level of compositionality C reached at the end of the simulations, averaged over the 10 trials plus their standard deviations. As the results show, after an initial downward trend from 10 to 40 agents, the results improve substantially for larger populations. Note that for increasing population sizes, the standard deviation also increases. This is due to the fact that for larger populations, compositionality emerges at a stable level in some cases, but not in all cases. Moreover, different levels of compositionality are achieved in different trials.

In order to evaluate the levels achieved, columns 3–6 provide the frequency with which different trials of the simulations achieved a given threshold. For instance, for a population size of 30, 6 trials ended with a compositionality greater than 0.6, 3 ended with $C > 0.8$, and only 1 ended with $C > 0.9$. Strikingly, for a population size of 100, 7 out of 10 trials achieved compositionality greater than 0.98; in 1 trial $C \approx 0.8$ and in the 2 other trials $C < 0.04$. So, again these results show that although compositionality initially is affected by increasing the population size, the results improve when the population size becomes larger.

The final column of Table 2 show that communicative accuracy is pretty much correlated with the level of compositionality in the language. So, given the current model, when no compositionality emerges, accuracy in communication is very low, whereas when compositionality is high, accuracy is high as well.

It is yet unclear why the performance appears to be better for larger populations. As the simulations take a long time to run (1 – 3 weeks per trial on a single computer), it is hard to set up good experiments that can investigate what exactly is going on. So, let me speculate a little on what could be going on. First it is important to realise that – with this model – it has been shown that when children speak early in life, they face an implicit bottleneck (i.e. they may need to communicate about previously unseen objects), which puts pressure on the formation of compositionality [6]. This is true for all population sizes. However, when the population is really large (i.e. $N = 100$ agents), the agents initially invent a lot of variety in the language, because different agents invent different expressions to talk about objects. Obviously, this does not improve the ease of learning the language. However, as a result of the variation, the probability that co-varying alignments occur in the signals by chance increases, because the alphabet is of limited size and there is a tendency to create short words. In addition, the environment also has a relatively high level of co-variation in different feature dimensions, so the likelihood that co-varying structures in signals co-occur with co-varying structures in the meaning space is high as well. As the agents try to exploit these

co-varying structures to form compositional structures, it is likely that such structures will arise rapidly [7]. This way, the – by chance – increased occurrence of co-varying structures soon becomes internalised in compositional structures. Furthermore, since these structures become internalised in multiple agents, they become successful in communication. Due to the positive feedback loop of the guessing games, these items are reinforced; and because the agents prefer to reuse effective items, they are used even more frequently. "Hence, the more success, the more use and the more use the more success" [3] and a conventionalised system emerges through self-organisation. This property is further enhanced, because compositional rules are more generally applicable than holistic ones, and thus tend to be used more frequently.

This way, the reason why the results are so good for large populations can – in part – be explained by the increased likelihood of finding regularities in expressions. However, this does not explain why initially the results show a downward trend when the population size increases. Apparently, the positive effect of larger populations is not yet sufficient for the somewhat smaller population sizes (i.e. $N < 50$) to achieve high results. Perhaps this is due to difficulties in arriving at shared conventions in larger populations where there is a lot of variation in the language, while there is insufficient covariation to overcome these difficulties. So, there seems to be a *trade off*: On the one hand, the larger the population, the greater the difficulty in establishing a conventionalised system. On the other hand, the larger the population, the more likely it is that regularities in the signals emerge by chance, thus giving rise to compositionality. The more compositionality in the system, the more it will be used, but – more importantly – there will occur less variation in the language, which will make compositional systems easier to learn.

Although the above may provide a suitable explanation, other explanations may prove more viable. For instance, it is possible that for some of the smaller populations the simulations have run too long and that there is somewhere an optimum in the number of language games per iteration. However, in simulations where the experiments have been run for a number of language games T per iteration proportional to $N \log N$, where N is the population size and where for $N = 100$, $T = 214,600$, the results were poorer for $N \leq 50$ than the ones presented in Table 2. (Note that the relation $T \propto N \log N$ was found to hold for studies using the language games for lexicon formation [13].) Perhaps the optimum is somewhere in between, but this is hard to assess given the computational complexity of the current model. Other explanations, e.g., based on pressures to improve the communication systems for larger populations may prove viable as well, though still do not explain the initial downward trend found. It may also turn out that, when repeating the simulations more frequently, the differences appear to be less significant, which is possible since the simulations have been repeated only 10 times and different trials can lead to different results. Further research is planned to investigate possible explanations using a less computationally expensive version of the current model.

4 Conclusions

In this paper the effect of increasing the population size in simulations regarding the emergence of compositional structures in language is investigated. The simulation is based on a model that integrates the language game model of the Talking Heads experiment [12] with the iterated learning model [4]. In the model, which is described in detail in [7], agents communicate with each other and while doing so, they construct compositional structures whenever they can, but otherwise incorporate expressions holistically. Where earlier versions of the ILM required a bottleneck on the transmission of language in order to arrive at compositional structures, the current model does not as was also shown in [6]. Moreover, where earlier versions of the ILM proved difficult to scale up in terms of population size, the current study shows that – after an initial decline – the level of compositionality can increase substantially when the population size is increased, though the emergence of a stable compositional system is not guaranteed.

From the results, it is hypothesised that there appears to be a trade off between increased difficulties at arriving at a shared communication system and an increased likelihood of finding

co-varying regularities in the signal-meaning space (thus giving rise to compositionality) when the population size increases. It seems that when the population size increases sufficiently the second aspect becomes strong enough to drive the emergence of compositionality, which – once it gets off the ground – is attracted to an effective and highly compositional language through self-organisation. Hence, it may be the case that languages evolve easier in large(r) populations.

References

[1] Briscoe EJ (Ed): *Linguistic evolution through language acquisition: formal and computational models*. Cambridge: Cambridge University Press 2002.

[2] Cangelosi A, Parisi D (Eds): *Simulating the Evolution of Language*. London: Springer 2002.

[3] Steels L: **The synthetic Modeling of Language Origins**. *Evolution of Communication* 1997, **1(1)**:1–34.

[4] Kirby S, Smith K, Brighton H: **From UG to Universals: linguistic adaptation through iterated learning**. *Studies in Language* 2004, **28**(3):587–607.

[5] Steels L: **Constructivist Development of Grounded Construction Grammars**. In *Proceedings Annual Meeting of Association for Computational Linguistics*. Edited by Daelemans W 2004.

[6] Vogt P: **On the acquisition and evolution of compositional languages: Sparse input and the productive creativity of children**. *Adaptive Behavior* 2005, **13(4)**:325–346.

[7] Vogt P: **The emergence of compositional structures in perceptually grounded language games**. *Artificial Intelligence* 2005, **167(1–2)**:206–242.

[8] Smith K, Hurford JR: **Language Evolution in Populations: extending the Iterated Learning Model**. In *Advances in Artificial Life: Proceedings of the* 7^{th} *European Conference on Artificial Life*, Lecture Notes in Computer Science: Lecture Notes in Artifical Intelligence. Edited by Banzhaf W, Christaller T, Ziegler J, Dittrich P, Kim JT, Heidelberg: Springer-Verlag 2003:507–516.

[9] Gong T, Ke J, Minett JW, Wang WSY: **A computational framework to simulate the co-evolution of language and social structure**. In *Artificial Life IX Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*. Edited by Pollack J, Bedau M, Husbands P, Ikegami T, Watson RA, The MIT Press 2004:214–219.

[10] Vogt P, Divina F: **Language evolution in large populations of autonomous agents: issues in scaling**. In *Proceedings of AISB 2005: Socially inspired computing joint symposium*. Edited by Edmonds B, Gilbert N, Gustafson S, Hales D, Krasnagor N, AISB 2005:80–87.

[11] Vogt P: **THSim v3.2: The Talking Heads simulation tool**. In *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL)*. Edited by Banzhaf W, Christaller T, Dittrich P, Kim JT, Ziegler J, Springer Verlag Berlin, Heidelberg 2003.

[12] Steels L, Kaplan F, McIntyre A, Van Looveren J: **Crucial factors in the origins of word-meaning**. In *The Transition to Language*. Edited by Wray A, Oxford, UK: Oxford University Press 2002.

[13] Kaplan F: **L'émergence d'un lexique dans une population d'agent autonomes**. *PhD thesis*, Laboratoire d'infromatique de Paris 6 2000.

Network Modelling

Weighted networks: empirical results and models

Marc Barthélemy[1], Alain Barrat[2], Alessandro Vespignani[1]

1 *School of Informatics and Biocomplexity Center, Indiana University, Eigenmann Hall, 1900 East Tenth Street, Bloomington, IN 47406, USA*
[2] *Laboratoire de Physique Théorique (UMR du CNRS 8627), Batiment 210, Université de Paris-Sud 91405 Orsay, France*

Abstract

We review the main tools which allow for the statistical characterization of weighted networks. We then present two case studies, the airline connection network and the scientific collaboration network, which are representative of critical infrastructures and social system, respectively. The main empirical results are (i) the broad distributions of various quantities and (ii) the existence of weight-topology correlations. These measurements show that weights are relevant and that in general the modeling of complex networks must go beyond topology. We propose a model which provides an explanation for the features observed in several real-world networks. This model of weighted network formation relies on the dynamical coupling between topology and weights, considering the rearrangement of weights when new links are introduced in the system. Finally, we discuss the effects of spatial constraints on the evolution of weighted networks.

Keywords: Complex Networks, Traffic.

6 May 2005

1 Introduction

Networked structures arise in a wide array of different contexts such as technological and transportation infrastructures, social phenomena, and biological systems. These highly interconnected systems have recently been the focus of a great deal of attention that has uncovered and characterized their topological complexity [1–4]. Along with a complex topological structure, real networks display a large heterogeneity in the capacity and intensity of the connections—the weight of the links. For example, in ecology the diversity of the predator-prey interaction is believed to be a critical ingredient of ecosystems stability [5,6], and in social systems, the weight of interactions is very important in the characterization of the corresponding networks [7]. Similarly, the Internet traffic [3] or the number of passengers in the airline network [4,8,9] are crucial quantities in the study of these systems.

In this paper we review a set of metrics combining weighted and topological observables that allows to characterize the complex statistical properties of the strength of edges and vertices and to investigate the correlations among weighted quantities and the underlying topological structure. Specifically, we present results on the scientific collaboration network and the world-wide air-transportation network, which are representative examples of social and large infrastructure systems, respectively. The measures on weighted networks [9–12] have shown that they can exhibit additional complex properties such as broad distributions and non-trivial correlations of weights that do not find an explanation just in terms of the underlying topological structure. The heterogeneity in the intensity of connections may thus be very important in real-world systems and cannot be overlooked in their description. Motivated by these observations, we review also a model for weighted networks we have recently proposed [13], which naturally produces topology-weight correlations and broad distributions for various quantities. Finally, we present an analysis of the effect of spatial constraints on the evolution of weighted networks.

2 Tools for the characterization of weighted networks

We briefly review the different tools which allow for a first statistical characterization of weighted complex networks.

- *Weights*

  The properties of a graph can be expressed via its adjacency matrix a_{ij}, whose elements take the value 1 if an edge connects the vertex i to the vertex j, and 0 otherwise (with $i, j = 1, ..., N$, where N is the size of the network). Weighted networks are usually described by a matrix w_{ij} specifying the weight on the edge connecting the vertices i and j ($w_{ij} = 0$ if the nodes i and j are not connected). In the following we will consider only the

case of symmetric positive weights $w_{ij} = w_{ji} \geq 0$.

- *Degree and weight distributions*
  The standard topological characterization of networks is obtained by the analysis of the probability distribution $P(k)$ that a vertex has degree k. Complex networks often exhibits a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $2 \leq \gamma \leq 3$. Similarly, a first characterization of weights is obtained by the distribution $P(w)$ that any given edge has weight w.

- *Weighted degree: Strength*
  Along with the degree of a node, a very significative measure of the network properties in terms of the actual weights is obtained by looking at the vertex *strength* s_i defined as [9]

$$s_i = \sum_{j \in \mathcal{V}(i)} w_{ij}, \tag{1}$$

  where the sum runs over the set $\mathcal{V}(i)$ of neighbors of i. The strength of a node integrates the information both about its degree and the importance of the weights of its links, and can be considered as the natural generalization of the degree. When the weights are independent from the topology, we obtain that the strength of the vertices of degree k is $s(k) \simeq \langle w \rangle k$ where $\langle w \rangle$ is the average weight. In the presence of correlations we obtain in general $s(k) \simeq Ak^{\beta}$ with $\beta = 1$ and $A \neq \langle w \rangle$ or $\beta > 1$.

- *Weighted clustering*
  The clustering coefficient c_i measures the local cohesiveness and is defined for any vertex i as the fraction of connected neighbors of i [14]. The average clustering coefficient $C = N^{-1} \sum_i c_i$ thus expresses the statistical level of cohesiveness measuring the global density of interconnected vertex triplets in the network. Further information can be gathered by inspecting the average clustering coefficient $C(k)$ restricted to the class of vertices with degree k. The topological clustering, however, does not take into account the fact that some neighbors are more important than others. In order to solve this incongruity we introduce a measure of the clustering that combines the topological information with the weight distribution of the network. The *weighted clustering coefficient* is defined as [9]

$$c^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij} a_{ih} a_{jh}. \tag{2}$$

  This quantity $c^w(i)$ counts for each triple formed in the neighborhood of the vertex i the weight of the two participating edges of the vertex i. In this way we are not just considering the number of closed triangles in the neighborhood of a vertex but also their total relative weight with respect to the vertex' strength. The factor $s_i(k_i - 1)$ is a normalization factor and

ensures that $0 \leq c_i^w \leq 1$. Consistently, the c_i^w definition recovers the topological clustering coefficient in the case that $w_{ij} = const$. It is customary to define C^w and $C^w(k)$ as the weighted clustering coefficient averaged over all vertices of the network and over all vertices with degree k, respectively. In the case of a large randomized network (lack of correlations) it is easy to see that $C^w = C$ and $C^w(k) = C(k)$. In real weighted networks, however, we can face two opposite cases. If $C^w > C$, we are in presence of a network in which the interconnected triples are more likely formed by the edges with larger weights. On the contrary, $C^w < C$ signals a network in which the topological clustering is generated by edges with low weight. In this case it is obvious that the clustering has a minor effect in the organization of the network since the largest part of the interactions (traffic, frequency of the relations, etc.) is occurring on edges not belonging to interconnected triples. The same may happen for $C^w(k)$, for which it is also possible to analyze the variations with respect to the degree class k.

- *Weighted assortativity: Affinity*

  Another quantity used to probe the networks' architecture is the average degree of nearest neighbors of a vertex i

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in \mathcal{V}(i)} k_j \ , \tag{3}$$

  where the sum runs on the nearest neighbors vertices of each vertex i. From this quantity a convenient measure to investigate the behavior of the degree correlation function is obtained by the average degree of the nearest neighbors, $k_{nn}(k)$, for vertices of degree k [15]. In the presence of correlations, the behavior of $k_{nn}(k)$ identifies two general classes of networks. If $k_{nn}(k)$ is an increasing function of k, vertices with high degree have a larger probability to be connected with large degree vertices. This property is referred in physics and social sciences as *assortative mixing* [16]. On the contrary, a decreasing behavior of $k_{nn}(k)$ defines *disassortative mixing*, in the sense that high degree vertices have a majority of neighbors with low degree, while the opposite holds for low degree vertices.

  In the case of weighted networks an appropriate characterization of the assortative behavior is obtained by the *weighted average nearest neighbors degree*, defined as

$$k_{nn,i}^w = \frac{1}{s_i} \sum_{j=1}^{N} a_{ij} w_{ij} k_j. \tag{4}$$

  In this case, we perform a local weighted average of the nearest neighbor degree according to the normalized weight of the connecting edges, w_{ij}/s_i. This definition implies that $k_{nn,i}^w > k_{nn,i}$ if the edges with the larger weights are pointing to the neighbors with larger degree and $k_{nn,i}^w < k_{nn,i}$ in the

opposite case. The $k^w_{nn,i}$ function thus measures the effective *affinity* to connect with high or low degree neighbors according to the magnitude of the actual interactions. As well, the behavior of the function $k^w_{nn}(k)$ (defined as the average of $k^w_{nn,i}$ over all vertices with degree k), marks the weighted assortative or disassortative properties considering the actual interactions among the system's elements.

- *Disparity*

  For a given node i with degree k_i and strength s_i different situations can arise. All weights w_{ij} can be of the same order s_i/k_i. In contrast, the most heterogeneous situation is obtained when one weight dominates over all the others. A simple way to measure this "disparity" is given by the quantity Y_2 introduced in other context [17,18]

$$Y_2(i) = \sum_{j \in \mathcal{V}(i)} \left[\frac{w_{ij}}{s_i} \right]^2 \tag{5}$$

  If all weights are of the same order then $Y_2 \sim 1/k_i$ (for $k_i \gg 1$) and if a small number of weights dominate then Y_2 is of the order $1/n$ with n of order unity. This quantity was recently used for metabolic networks [19] which showed that for these networks one can identify dominant reactions.

3 Empirical results

3.1 Weighted networks data

Prototypical examples of weighted networks can be found in the world-wide airport network (WAN) [8,9] and the scientific collaboration network (SCN) [20,21]. In the airport network each given weight w_{ij} is the number of available seats on direct flight connections between the airports i and j. For the WAN, we analyze the International Air Transportation Association (IATA) (www.iata.org.) database containing the world list of airports pairs connected by direct flights and the number of available seats on any given connection for the year 2002. The resulting air-transportation graph comprises $N = 3880$ vertices denoting airports and $E = 18810$ edges accounting for the presence of a direct flight connection. The average degree of the network is $\langle k \rangle = 2E/N = 9.70$, while the maximal degree is 318.

In the SCN the nodes are identified with authors and the weight depends on the number of co-authored papers [20,9]. We consider the network of scientists who have authored manuscripts submitted to the e-print archive relative to condensed matter physics (xxx.lanl.gov/archive/cond-mat) between 1995 and 1998. Scientists are identified with nodes and an edge exists between two scientists if they have co-authored at least one paper. The resulting connected

network has $N = 12722$ nodes, with an average degree (i.e. average number of collaborators) $\langle k \rangle = 6.28$ and maximal degree 97. For the SCN we follow the definition of weight introduced in Ref. [20]: The intensity w_{ij} of the interaction between two collaborators i and j is defined as $w_{ij} = \sum_p \delta_i^p \delta_j^p / (n_p - 1)$ where the index p runs over all papers, n_p is the number of authors of the paper p, and δ_i^p is 1 if author i has contributed to paper p, and 0 otherwise. This definition seems to be rather objective and representative of the scientific interaction: It is large for collaborators having many papers in common but the contribution to the weight introduced by any given paper is inversely proportional to the number of authors.

3.2 Results

3.2.1 Topological properties

The topological properties of the SCN network and other similar networks of scientific collaborations have been studied in Ref. [20] and we report on Fig. (1A) the degree distribution showing a relatively broad law.

As shown in Fig. (1B), the topology of the WAN exhibits both small-world and scale-free properties as already observed in different dataset analyses [10,8]. In particular, the average shortest path length, measured as the average number of edges separating any two nodes in the network, shows the value $\langle \ell \rangle = 4.37$, very small compared to the network size N. The degree distribution, on the other hand, takes the form $P(k) = k^{-\gamma} f(k/k_x)$, where $\gamma \simeq 2.0$ and $f(k/k_x)$ is an exponential cut-off function that finds its origin in physical constraints on the maximum number of connections that a single airport can handle [4,8]. The airport connection graph is therefore a clear example of heterogeneous network showing scale-free properties on a definite range of degree values.

3.2.2 Strength distribution

The probability distribution $P(s)$ that a vertex has strength s is heavy tailed in both networks and the functional behavior exhibits similarities with the degree distribution $P(k)$ (see Fig. 1). A precise functional description of the heavy-tailed distributions may be very important in understanding the network evolution and will be deferred to future analysis. This behavior is not unexpected since it is plausible that the strength s_i increases with the vertex degree k_i, and thus the slow decaying tail of $P(s)$ stems directly from the very slow decay of the degree distribution.

3.2.3 Topology-weight correlations

In Fig. 2 we report the behavior obtained for both the real weighted networks and their randomized versions, generated by a random re-distribution

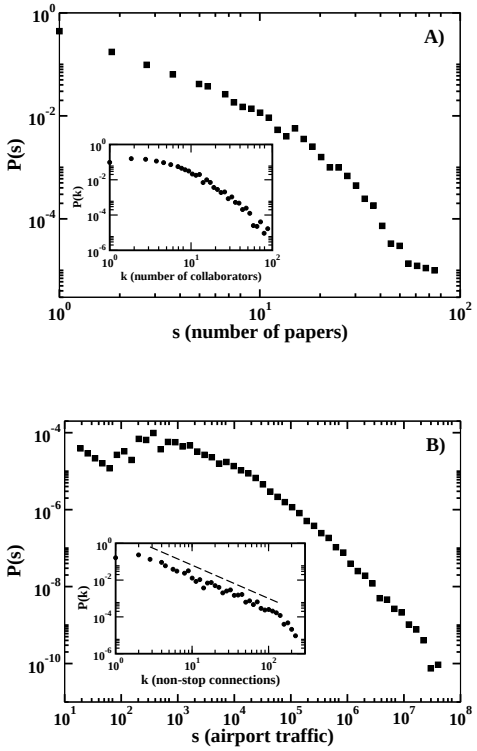


Fig. 1. **A** Degree and strength distribution in the scientific collaboration network. The degree k corresponds to the number of co-authors of each scientist and the strength represent its total number of publications. The distributions are heavy-tailed even if it is not possible to distinguish a definite functional form. **B** The same distributions for the world-wide airport network. The degree is the number of non-stop connections to other airports and the strength is the total number of passengers handled by any given airport. In this case, the degree distribution can be approximated by the power-law behavior $P(k) \sim k^{-\gamma}$ with $\gamma = 1.8 \pm 0.2$. The strength distribution has a heavy-tail extending over more than four orders of magnitude.

of the actual weights on the existing topology of the network. For the SCN the curves are very similar and well fitted by the uncorrelated approximation $s(k) = \langle w \rangle k$. Strikingly, this is not the case of the WAN. Fig. 2B clearly shows a very different behavior for the real data set and its randomized version. In particular, the power-law fit for the real data gives an "anomalous" exponent $\beta_{\mathrm{WAN}} = 1.5 \pm 0.1$. This implies that the strength of vertices grows faster than their degree, i.e. the weight of edges belonging to highly connected vertices tends to have a value higher than the one corresponding to a random assignment of weights. This denotes a strong correlation between the weight and the topological properties in the WAN, where the larger is an airport, the more traffic it can handle.

The fingerprint of these correlations is also observed in the dependence of the weight w_{ij} on the degrees of the end point nodes k_i and k_j. For the WAN

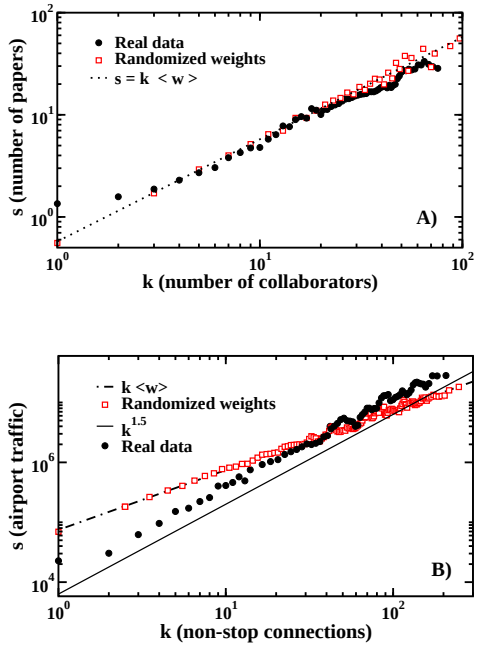


Fig. 2. Average strength $s(k)$ as function of the degree k of nodes. **A** In the scientific collaboration network the real data are very similar to those obtained in a randomized weighted network. Only at very large k values it is possible to observe a slight departure from the expected linear behavior. **B** In the world airport network real data follow a power-law behavior with exponent $\beta = 1.5 \pm 0.1$. This denotes anomalous correlations between the traffic handled by an airport and the number of its connections.

the behavior of the average weight as a function of the end points degrees can be well approximated by a power-law dependence $\langle w_{ij} \rangle \sim (k_i k_j)^{\theta}$ with an exponent $\theta = 0.5 \pm 0.1$. This exponent can be related to the β exponent by noticing that $s(k) \sim k(kk_j)^{\theta}$, resulting in $\beta = 1 + \theta$, if the topological correlations between the degree of connected vertices can be neglected. This is indeed the case of the WAN where the above scaling relation is well satisfied by the numerical values provided by the independent measurements of the exponents. In the SCN, instead, $\langle w_{ij} \rangle$ is almost constant for over two decades confirming a general lack of correlations between the weights and the vertices degree. In this case $\theta = 0$ and the relation $\beta = 1 + \theta$ also holds.

3.2.4 Weighted clustering and assortativity

We present the results [9] obtained for both the SCN (see Fig.3) and the WAN (see Fig.4) by comparing the regular topological quantities with the weighted ones introduced above. In the figures we report the relative difference of the values obtained for teh topological and weighted quantities. It is striking to observe that for large degree values we observe up to 100% relative difference

signalling a strong difference in the clustering and correlation properties if we take into accoutn the weighted nature of the networks.

- *SCN*
  - (i) The measurements tell us that the SCN has a monotonously decaying spectrum $C(k)$. This implies that hubs present a much lower clustered neighborhood than low degree vertices which can be interpreted as the evidence that authors with few collaborators usually work within a well defined research group in which all the scientists collaborate together (high clustering). Authors with a large degree, however, collaborate with different groups and communities which on their turn do not have often collaborations, thus creating a lower clustering coefficient.
  - (ii) The inspection of $C^w(k)$ shows generally that for $k \geq 10$ the weighted clustering coefficient is larger than the topological one. This implies that authors with many collaborators tend to publish more papers with interconnected groups of co-authors and is a signature of the fact that influential scientists form stable research groups where the largest part of their production is obtained.
  - (iii) Furthermore, the SCN exhibits an assortative behavior in agreement with the general evidence that social networks are usually denoted by a strong assortative character [16]. Finally, the assortative properties find a clearcut confirmation in the weighted analysis with a $k^w_{nn}(k)$ strikingly growing as a power-law as a function of k.
- *WAN*

  A different picture is found in the WAN, where the weighted analysis provides a richer and somehow different scenario.
  - (i) This network also shows a decaying $C(k)$, consequence of the role of large airports that provide non-stop connections to very far destinations on an international and intercontinental scale. These destinations are usually not interconnected among them, giving rise to a low clustering coefficient for the hubs.
  - (ii) We find, however, that $C^w/C \simeq 1.1$, indicating an accumulation of traffic on interconnected groups of vertices.
  - (iii) The weighted clustering coefficient $C^w(k)$ has much more limited variation in the whole spectrum of k. This implies that high degree airports have a progressive tendency to form interconnected groups with high traffic links, thus balancing the reduced topological clustering. Since high traffic is associated to hubs, we have a network in which high degree nodes tend to form cliques with nodes with equal or higher degree, the so-called *rich-club phenomenon* [22].
  - (iv) The topological $k_{nn}(k)$ does show an assortative behavior only at small degrees. For $k > 10$, $k_{nn}(k)$ approaches a constant value, a fact revealing an uncorrelated structure in which vertices with very different degrees have a very similar neighborhood. The analysis of the weighted $k^w_{nn}(k)$, however, exhibits a pronounced assortative behavior in the whole

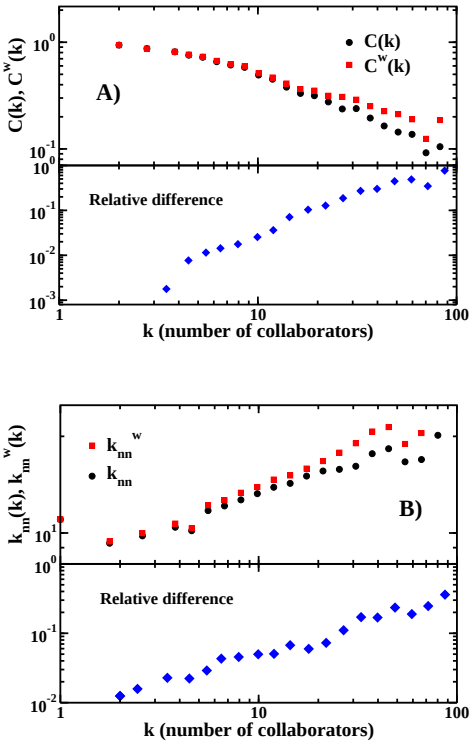


Fig. 3. Comparison of topological and weighted quantities for the SCN. **A** The weighted clustering separates form the topological one around $k \geq 10$. This marks a difference for authors with larger number of collaborators. **B** The assortative behavior is enhanced in the weighted definition of the average nearest neighbors degree. It is worth remarking the large relative difference (up to 50-100%) of the weighted and topological quantities (lower part of the figures).

k spectrum, providing a different picture in which high degree airports have a larger affinity for other large airports where the major part of the traffic is directed.

4 Modeling weighted networks

4.1 The need for weight-topology coupling

Previous approaches to the modeling of weighted networks focused on growing topologies where weights were assigned statically, i.e. once and for ever, with different rules related to the underlying topology [23,24]. These mechanisms, however, overlook the dynamical evolution of weights according to the topological variations. We can illustrate this point in the case of the airline network. If a new airline connection is created between two airports it

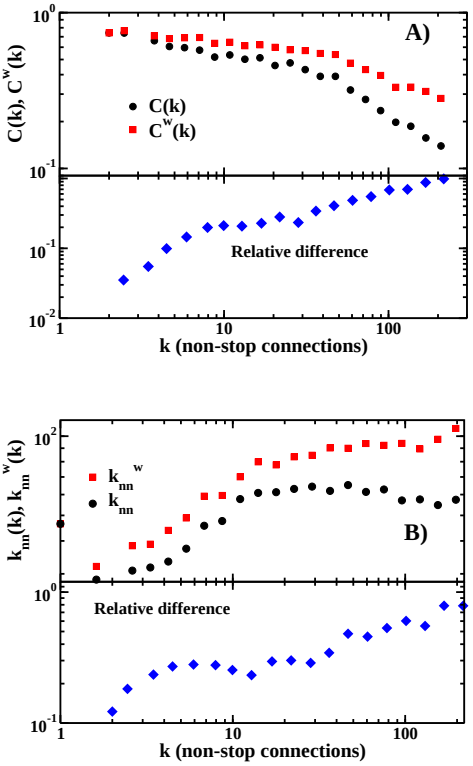


Fig. 4. Topological and weighted quantities for the WAN. **A** The weighted clustering coefficient is larger than the topological one in the whole degree spectrum. **B** $k_{nn}(k)$ is reaching a plateau for $k > 10$ denoting the absence of marked topological correlations. On the contrary $k^w_{nn}(k)$ exhibits a more definite assortative behavior. Also in this case the relative difference of the weighted and topological quantities are of the order of 100% (lower part of the figures).

will generally provoke a modification of the existing traffic of both airports. In general, it will increase the traffic activity depending on the specific nature of the network and on the local dynamics. In the following, we review a model that takes into account the coupled evolution in time of topology and weights. Instead of drawing randomly the weights, an alternative consists in coupling the evolution of the weights and of the topology and allowing the dynamical evolution of weights during the growth of the system. This mimics the evolution and reinforcements of interactions in natural and infrastructure networks.

The model dynamics starts from an initial seed of N_0 vertices connected by links with assigned weight w_0. At each time step, a new vertex n is added with m edges (with initial weight w_0) that are randomly attached to a previously existing vertex i according to the probability distribution

$$\Pi_{n\to i} = \frac{s_i}{\sum_j s_j}. \tag{6}$$

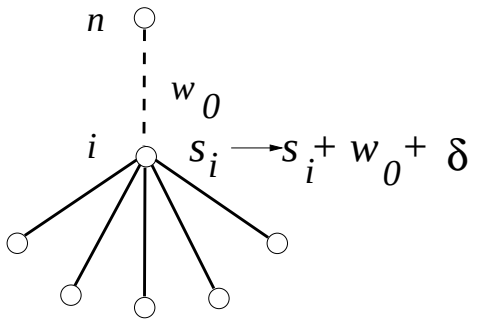


Fig. 5. Illustration of the construction rule. A new node n connects to a node i with probability proportional to $s_i / \sum_j s_j$. The weight of the new edge is w_0 and the total weight on the existing edges connected to i is modified by an amount equal to δ.

This rule of "busy get busier" relaxes the usual degree preferential attachment, focusing on a strength driven attachment in which new vertices connect more likely to vertices handling larger weights and which are more central in terms of the strength of interactions. This weight driven attachment (Eq. (6)) appears to be a plausible mechanism in many networks. In the Internet new routers connect to more central routers in terms of bandwidth and traffic handling capabilities and in the airport networks new connections are generally established to airports with a large passenger traffic. Even in the SCN this mechanism might play a role since an author with more co-authored papers is more visible and open to further collaborations.

The presence of the new edge (n, i) will introduce variations of the existing weights across the network. In particular, we consider the local rearrangements of weights between i and its neighbors $j \in \mathcal{V}(i)$ according to the simple rule

$$w_{ij} \to w_{ij} + \Delta w_{ij}, \tag{7}$$

where

$$\Delta w_{ij} = \delta \frac{w_{ij}}{s_i}. \tag{8}$$

This rule considers that the establishment of a new edge of weight w_0 with the vertex i induces a total increase of traffic δ that is proportionally distributed among the edges departing from the vertex according to their weights (see Fig. 5), yielding $s_i \to s_i + \delta + w_0$. At this stage, it is worth remarking that while we will focus on the simplest model with $\delta = const$, different choices of Δw_{ij} with heterogeneous δ_i or depending on the specific properties of each vertex (w_{ij}, k_i, s_i) can be considered [25,26,29]. Finally, after the weights have been updated the growth process is iterated by introducing a new vertex with the corresponding re-arrangement of weights.

The model depends only on the dimensionless parameter δ (rescaled by w_0),

that is the fraction of weight which is 'induced' by the new edge onto the others. According to the value of δ, different scenarios are possible. If the induced weight is $\delta \approx 1$ we mimic situations in which an appreciable fraction of traffic generated by the new connection will be dispatched in the already existing connections. This is plausible in the airport networks where the transit traffic is rather relevant in hubs. In the case of $\delta < 1$ we face situations such as the SCN where it is reasonable to consider that the birth of a new collaboration (co-authorship) is not triggering a more intense activity on previous collaborations. Finally, $\delta > 1$ is an extreme case in which a new edge generates a sort of multiplicative effect that is bursting the weight or traffic on neighbors.

The network's evolution can be inspected analytically by studying the time evolution of the average value of $s_i(t)$ and $k_i(t)$ of the i-th vertex at time t, and by relying on the continuous approximation that treats k, s and the time t as continuous variables [1,2,13]. The behavior of the strength and the degree are easily obtained and one has in the long time limit

$$s_i(t) \simeq (2\delta + 1)k_i(t) \tag{9}$$

This proportionality relation $s \sim k$ implies $\beta = 1$ but the prefactor is different from $\langle w \rangle$ which indicates the existence of correlations between topology and weights. This relation (9) is particularly relevant since it states that the weight-driven dynamics generates in Eq. (6) an effective degree preferential attachment that is parameter independent. This highlights an alternative microscopic mechanism accounting for the presence of the preferential attachment dynamics in growing networks.

The behavior of the various statistical distribution can be easily computed and one obtains in the large time limit $P(k) \sim k^{-\gamma}$ and $P(s) \sim s^{-\gamma}$ with

$$\gamma = \frac{4\delta + 3}{2\delta + 1}. \tag{10}$$

This result shows that the obtained graph is a scale-free network described by an exponent $\gamma \in [2, 3]$ that depends on the value of the parameter δ. In particular, when the addition of a new edge doesn't affect the existing weights ($\delta = 0$), the model is topologically equivalent to the Barabasi-Albert model [30] and the value $\gamma = 3$ is recovered. For larger values of δ the distribution is progressively broader with $\gamma \to 2$ when $\delta \to \infty$. This indicates that the weight-driven growth generates scale-free networks with exponents varying in the range of values usually observed in the empirical analysis of networked structures [1–3]. Noticeably the exponents are non-universal and depend only on the parameter δ governing the microscopic dynamics of weights. The model therefore proposes a general mechanism for the occurrence of varying power-law behaviors without resorting on more complicate topological rules and variations of the basic preferential attachment mechanism.

Similarly to the previous quantities, it is possible to obtain analytical expressions for the evolution of weights and the relative statistical distribution [13]. The probability distribution $P(w)$ is in this case also a power-law $P(w) \sim w^{-\alpha}$ where $\alpha = 2 + \frac{1}{\delta}$. The exponent α has large variations as a function of the parameter δ and $P(w)$ moves from a delta function for $\delta = 0$ to a very slow decaying power-law with $\alpha = 2$ if $\delta \to \infty$. This feature clearly shows that the weight distribution is extremely sensible to changes in the microscopic dynamics ruling the network's growth.

In summary, the networks generated by the model display power-law behavior for the weight, degree and strength distributions with non-trivial exponents depending on the unique parameter defining the model's dynamics. These results suggest that the inclusion of weights in networks modeling naturally explains the diversity of scale-free behavior empirically observed in real networked structures. Strikingly, the weight-driven growth recovers an effective preferential attachment for the topological properties, providing a microscopic explanation for the ubiquitous presence of this mechanism.

4.2 Effect of spatial constraints

Another important aspect lies in the embedding of some networks in real space. For instance, most people have their friends and relatives in their neighborhood, transportation networks depend on distance, and many communication networks devices have short radio range [33–37]. A particularly important example of such a "spatial" network is the Internet which is a set or routers linked by physical cables with different lengths and latency times [38,3,39]. More generally, many networks can be seen as embedded in a given parameter space: each node carries a certain value of the parameter, and links are established with a certain cost depending on these values. One can for example think of social dimensions and social distances measured by salary, socio-professional category differences. If the cost of a long-range link is high, most of the connections starting from a given node will go to the closest neighbors in the embedding space. Long-range links, on the other hand, correspond usually to connections towards already well-connected nodes (hubs). This seems natural in the case of the air transportation network for instance: short connections go to small airports while long distance flights are directed preferentially towards large airports (*i.e.* well connected nodes). These spatial constraints can have important consequences on the topology of the resulting network [40].

So far, most models and simple mechanisms have focused on the strong heterogeneity of weights and topology, as well as the existence of some level of correlations between these aspects of complex networks. Not all features observed in real networks are however reproduced: in particular, the "traffic" as measured by the strength of a vertex grows only linearly with its degree,

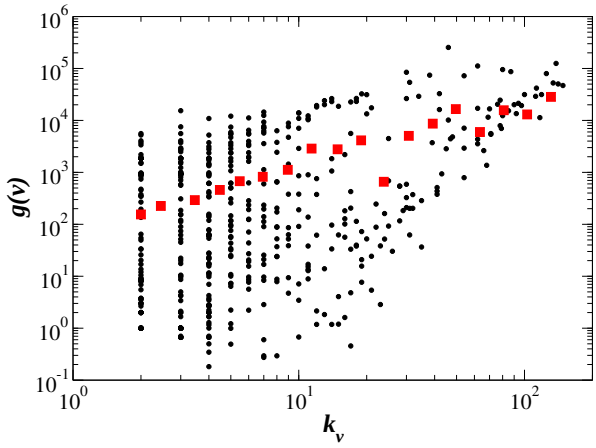


Fig. 6. Scatter-plot of the betweenness centrality versus degree for nodes of the North-American air-transportation network. The red symbols correspond to the average BC versus degree.

(except if non-linearities are introduced by hand [25]) while a non-linear behavior is observed in real-world [9]. The centrality fluctuations can also become anomalously large as it was observed in the airport network [8] (figure 6). One may therefore wonder if these interesting and non-trivial features can result from the combination of weight, topology and space. To this end, we define a simple model of growing network combining these ingredients which can be seen as an investigating tool. This simple model is obtained as the embedding of the weighted growing network introduced above in a two-dimensional space. Spatial constraints are translated into a preference for short links, and combined with the coupling between the evolution of the network and the dynamical rearrangement of the weights. We thus start with an initial seed of N_0 vertices randomly (with uniform distribution) located on a 2d disk (of radius L) and connected by links with assigned weight w_0. At each time step, a new vertex n is randomly placed on the disk. This new site is connected to m previously existing vertices, choosing preferentially nearest sites with the largest strength. More precisely, a node i is chosen according to the probability

$$\Pi_{n\to i} = \frac{s_i e^{-d_{ni}/r_c}}{\sum_j s_j e^{-d_{nj}/r_c}}, \tag{11}$$

where r_c is a typical scale and d_{ni} is the euclidean distance between n and i. This rule of *strength driven preferential attachment with spatial selection*, generalizes the preferential attachment mechanism driven by the strength to spatial networks. Here, new vertices connect more likely to vertices which correspond to the best interplay between Euclidean distance and strength. We then update the weight according to the same rule as for the model without spatial effect $w_{ij} \to w_{ij} + \delta \frac{w_{ij}}{s_i}$.

The model thus contains two relevant parameters: the ratio between the typical scale and the size of the system $\eta = r_c/L$, and the ability to redistribute weights, δ. When $\eta \gg 1$ space is irrelevant and in the opposite case a small value $\eta \ll 1$ indicates that long links are costly. When spatial constraints exist, the network properties are modified and the effects are summarized in the following.

- *Effect of spatial embedding on topology-traffic correlations*
  Spatial constraints induce strong nonlinear correlations ($\beta > 1$) between topology and traffic as measured by the degree and the strength, respectively. The reason for this behavior is that spatial constraints favor the formation of regional hubs and reinforces locally the preferential attachment, leading for a given degree to a larger strength than the one observed without spatial constraints. The existence of constraints such as spatial distance selection induces some strong correlations between topology (degree) and non-topological quantities such as weights or distances.

- *Effect of space embedding on centrality*
  Spatial constraints also induce large betweenness centrality fluctuations. While hubs are usually very central, when space is important central nodes tend to get closer to the gravity center of all points. Correlations between spatial position and centrality compete with the usual correlations between degree and centrality [28], leading to the observed large fluctuations of centrality at fixed degree. Generally speaking when spatial constraints exist central nodes are correlated with barycentric considerations. This phenomenon can be understood with the following argument. For a spatial network—the extreme case being a lattice—the shortest path between two nodes is simply the Euclidean geodesic. If the two endpoints are far away, the probability that the shortest path passes near the barycenter of all nodes is very high. In other words, this implies that the barycenter (and its neighbors) will have a large centrality. This point is illustrated in Fig. 7 in the simple case of a one-dimensional lattice. This general consideration applies also in the case of more complex networks such as the spatial network considered here and indeed as shown in Figure 8 the average distance $\langle d(G, C)\rangle$ between the barycenter G and the most central nodes decreases when spatial constraints become important. As expected, as spatial constraints become more important, the most central nodes get closer to the barycenter of the network.

- *Effect of space embedding on clustering and assortativity*
  Spatial constraints implies that the tendency to connect to hubs is limited by the need to use small-range links. This leads to an almost flat behavior for the assortativity. Connection costs also favor the formation of cliques between spatially close nodes and thus increase the clustering coefficient.

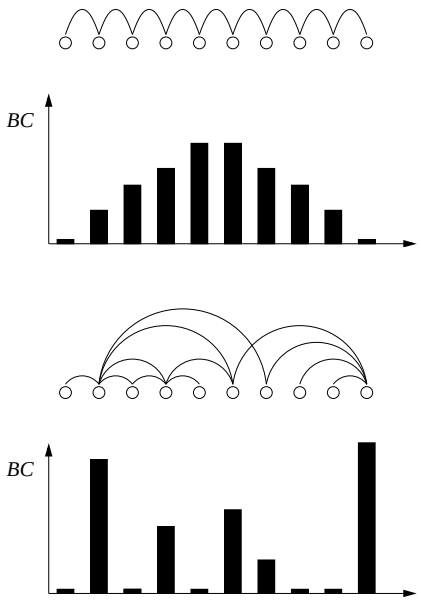


Fig. 7. **A** Betweenness centrality for the (one-dimensional) lattice case. The central nodes are close to the barycenter. **B** For a general graph, the central nodes are usually the ones with la

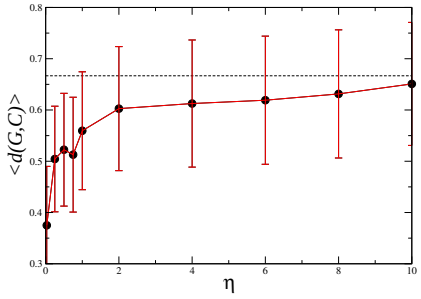


Fig. 8. Average distance between the barycenter G of all nodes and the 10 most central nodes (C) versus the parameter η (Here $\delta = 0$, $N = 5,000$ and the results are averaged over 50 configurations). When space is important (ie. small η), the central nodes are closer to the gravity center. For large η, space is irrelevant and the average distance tends to its value computed for a uniform distribution $\langle r \rangle_{unif} = 2/3$ (dotted line).

5 Conclusions and perspectives

A more complete view of complex networks is thus provided by the study of the interactions defining the links of these systems. The weights characterizing the various connections exhibit complex statistical features with highly varying distributions and power-law behavior. In particular we have considered the specific examples of the scientific collaboration and world-wide airport networks where it is possible to appreciate the importance of the correlations between weights and topology in the characterization of real networks prop-

erties. Indeed, the analysis of the weighted quantities and the study of the correlations between weights and topology provide a complementary perspective on the structural organization of the network that might be undetected by quantities based only on topological information. The weighted quantities thus offer a quantitative and general approach to understanding the complex architecture of real weighted networks. The empirical results shows that purely topological models are inadequate and that there is a need for modelels going beyond pure topology. The model we have presented is possibly the simplest one in the class of weight-driven growing networks. A novel feature in the model is the weight dynamical evolution occurring when new vertices and edges are introduced in the system. This simple mechanism produces a wide variety of complex and scale-free behavior depending on the physical parameter δ that controls the local microscopic dynamics. While a constant parameter δ is enough to produce a wealth of interesting network properties, a natural generalization of the model consists in considering δ as a function of the vertices degree or strength. Similarly, more complicated variations of the microscopic rules may be implemented to mimic in a detailed fashion particular networked systems [25,26,29,31]. In particular, we have shown that including spatial effects in this simple model of weighted networks leads to a large variety of behavior and interesting effects such as the appearance of large centrality fluctuations. This study sheds some light on the importance and effect of different ingredients such as spatial embedding or diversity of interaction weights in the structure of large complex networks and we believe that this attempt of a network typology could be useful in the understanding and modeling of real-world networks.

Acknowledgments. M.B. is on leave of absence from the CEA-Centre d'Etudes de Bruyères-Le-Chatel. We thank IATA for making the airline commercial flight database available to us. We also thank M.E.J. Newman for conceding us the possibility of using the scientific collaboration network data (see [32]). A.B. and A.V. are partially funded by the European Commission - Fet Open Project COSIN IST-2001-33555.

References

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] S.N. Dorogovtsev, & J.F.F. Mendes, *Evolution of networks: From biological nets to the Internet and WWW*, Oxford University Press, Oxford (2003).

[3] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A statistical physics approach* (Cambridge University Press, Cambridge, 2004).

[4] L.A.N. Amaral, A. Scala, M. Barthélemy, and H.E. Stanley, Proc. Natl. Acad. Sci. (USA) **97**, 11149 (2000).

[5] S.L. Pimm, *Food Webs*, The University of Chicago Press (2nd edition, 2002).

[6] A.E. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor, Nature **426**, 282 (2003).

[7] M. Granovetter, American Journal of Sociology, **78** (6) 1360-1380 (1973).

[8] R. Guimerà , S. Mossa, A. Turtschi, and L.A.N. Amaral, Preprint cond-mat/0312535 (2003); R. Guimerà and L.A.N. Amaral, Eur. Phys. J. **38** 381 (2004).
[9] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci. (USA) **101** 3747 (2004).
[10] C. Li and X. Cai, Preprint cond-mat/0309236 (2003).
[11] C. Li and G. Chen, Preprint cond-mat/0311333 (2003).
[12] D. Garlaschelli, S. Battiston, M. Castri, V.D.P. Servedio, and G. Caldarelli, Preprint cond-mat/0310503 (2003).
[13] A. Barrat, M. Barthélemy, and A. Vespignani, Phys. Rev. Lett. **92** 228701 (2004).
[14] D.J. Watts and S.H. Strogatz, Nature **393**, 440 (1998).
[15] R. Pastor-Satorras, A. Vázquez and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
[16] M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[17] B. Derrida and H. Flyvbjerg, J. Phys. A **20** 5273 (1987).
[18] M. Barthélemy, B. Gondran, and E. Guichard, Physica A **319** 633 (2003).
[19] E. Almaas, B. Kovacs, T. Viscek, Z.N. Oltvai, and A.-L. Barabási, Nature **427**, 839 (2004).
[20] M. E. J. Newman, Phys. Rev. E **64**, 016131 (2001); M. E. J. Newman, Phys. Rev. E **64**, 016132 (2001).
[21] A.-L. Barabasi, H. Jeong, R. Ravasz, Z. Neda, T. Vicsek, and A. Schubert, Physica A **311**, 590 (2002).
[22] S. Zhou & R.J. Mondragon, e-print cs.NI0303028 (2003).
[23] S.H. Yook, H. Jeong, A.-L. Barabasi, and Y. Tu, Phys. Rev. Lett. **86**, 5835 (2001).
[24] D. Zheng, S. Trimper, B. Zheng and P.M. Hui, Phys. Rev. E **67**, 040102 (2003).
[25] A. Barrat, M. Barthélemy, and A. Vespignani, LNCS **3243** (2004) 56.
[26] A. Barrat, M. Barthélemy, and A. Vespignani, Phys. Rev. E (in press, 2005). cond-mat/0406238.
[27] M. Barthélemy, A. Barrat, and A. Vespignani, JSTAT in press (2005).
[28] M. Barthélemy, Eur. Phys. J. B **38**, 163 (2003).
[29] R.V.R. Pandya. Preprint cond-mat/0406644 (2004).
[30] A.-L. Barabasi and R. Albert, Science **286**, 509 (1999).
[31] Bo Hu, Xin-Yu Jiang, Jun-Feng Ding, Yan-Bo Xie, Bing-Hong Wang. Preprint cond-mat/0408125 (2004).
[32] http://www-personal.umich.edu/∼mejn/collaboration/
[33] A. Helmy, online archive: cs.NI/0207069.
[34] G. Nemeth and G. Vattay, cond-mat/0211325.
[35] S.P. Gorman and R. Kulkarni, submitted to Environment and Planning Journal B, (2003).
[36] M.T. Gastner and M.E.J. Newman, condmat/0407680
[37] M.T. Gastner and M.E.J. Newman, condmat/0409702

[38] A. Lakhina, J.B. Byers, M. Crovella, and I. Matta, technical report, online version available at : http://www.cs.bu.edu/techreports/pdf/2002-015-internet-geography.pdf

[39] S.-H. Yook, H. Jeong, and A.-L. Barabasi, Proc. Natl. Acad. Sci. (USA) **99**, 13382 (2002).

[40] M. Barthélemy, Europhys. Lett. **63**, 915 (2003).

BOUNDED RATIONALITY AND REPEATED NETWORK FORMATION*

Sylvain Beal*, Nicolas Quérou*

31st August 2005

Abstract: We define a finite-horizon repeated network formation game with consent, and study the differences induced by different levels of individual rationality. We prove that perfectly rational players will remain unconnected at the equilibrium, while nonempty equilibria are possible when, following Neyman (1985), players are assumed to behave as finite automata. We define two types of equilibria, namely the Repeated Nash Network (RNN), in which the same network forms at each period, and the Repeated Nash Equilibrium (RNE), in which different networks may form. We state a sufficient condition under which a given network may be implemented as a RNN. Then, we provide structural properties of RNE. For instance, players may form totally different networks at each period, or the networks within a given RNE may exhibit a total order relationship. Finally we investigate the question of efficiency. We characterize efficient outcomes and prove that the sets of Bentham and Pareto efficient outcomes are identical.

Key words: Repeated network formation game, Two-sided link formation costs, Bounded rationality, Automata.

JEL classification: C72

*CREUSET, University of Saint-Etienne, 42023 Saint-Etienne, France. E-mail adresses : sylvain.beal@univ-st-etienne.fr (corresponding author) and nicolas.querou@univ-st-etienne.fr. Tel : (+33) (0)4 77 42 19 70. Fax : (+33) (0)4 77 42 19 50.

*We thank Hans Haller and Philippe Solal for helpful conversations. We also have benefited from comments of seminars participants at Institut Henri Poincaré, EUREQua and at SING I, SAET and 16th Stony Brook conferences. The present work is supported by the French Ministry of Research under ACI SCSHS-2004-04.

1 Introduction

Both network structures and rationality of agents play a significant role in determining the outcome of many economic economic relationships. A vast and recent literature examines how network structure affects economic outcomes.[1] The literature on bounded rationality has become more and more important since its introduction by Simon (1955). Our aim is to study the process of network formation within a dynamic framework in two cases related to different levels of rationality for economic agents. In the first one, the agents are perfectly rational. In the second one, some aspects of their rationality are limited.

We consider a group of agents who are initially unconnected who form or remove links with each other. A link can be removed unilaterally but agreement by both agents is needed to form a link. Precisely, a player pays an amount $c > 0$ to seek contact with an opponent and the link forms if the opponent behaves likewise. An agent's payoff is determined as in Gilles and Sarangi's (2004) model with consent. Agents receive the same value from all direct and indirect connections. The cost to a player of creating or maintaining a link is greater than the reward of a single direct connection as in Watts (2002).

Since agreement is required to form links, it is crucial to distinguish an action profile, which lists the wishes or efforts of the players, from the induced network. In fact, several distinct action profiles may lead to an identical network. We focus on a particular subset of action profiles called *cost-efficient action profiles*. In such an action profile, if no link connects two players i and j, then neither i nor j seeks contact with the opponent to create the link. In the static network formation game, only cost-efficient action profiles are likely to define Nash equilibria (NE) since a player incurs a cost for seeking contact with an opponent. For the same reason, only cost-efficient action profiles are likely to be Bentham or Pareto efficient. Precisely, we show that in a NE, perfectly rational players must choose the cost-efficient action profile that induces the empty network. In the finitely repeated network formation game, perfectly rational players also remain unconnected. The unique NE consists in a sequence of cost-efficient action profiles such that the empty network forms in each period. In the current paper, we choose to focus on a finite horizon of play in which a network is built in each period. If players are individuals or firms with a finite lifetime, it seems reasonable to suppose that the relationships they establish stop after a certain amount of time. It

[1] Different examples may be found in consumer's theory (Ellison, Fudenberg, 1995), labor market (Calvó-Armengol, Jackson, 2004), industrial organization (Bolton, Dewatripont, 1994), or in game theory (Ellison, 1993).

is assumed that players create a network in each period. Thus there are many types of possible behaviors for a given player, which depend on the opponents' behavior. Networks may be entirely rebuilt from one period to another, some former links may be maintained or removed, or new links may be created.

In this paper, we also limit players with respect to the complexity of the strategies they can implement. We model this by assuming that they use finite automata with a bounded number of states to play their strategies as in Neyman (1985). Such players are boundedly rational in the sense that they are limited in their ability to count the number of periods played and thus may not anticipate the opponents' behaviors against their own actions. However they have the ability to optimize their payoff against the opponents' strategies given that their choice is reduced to strategies played by finite automata. We also restrict the analysis to cost-efficient action profiles. A player cannot use an established contact with an unconsenting opponent in a period as a message for a possible agreement in a subsequent period. Notice that any network is induced by a unique cost-efficient action profile. Therefore, the restriction to cost-efficient action profiles does not limit the architecture of static networks that may form in an equilibrium outcome. We show that if the size of automata is smaller than the duration of the game, then the set of NE of the repeated game is not reduced to outcomes filled with empty networks, as it is the case with perfectly rational agents. We do not explore whether equilibrium outcomes are affected when the analysis is extended to all action profiles. Nevertheless, cost-efficient equilibria induce a large variety of sequences of networks, which deserve attention.

We distinguish two types of equilibria. In the first one, the same network is formed in all periods. We refer to such an equilibrium as a Repeated Nash Network (RNN). We provide a sufficient condition for the existence of RNNs based on any static network (proposition 3). Moreover, we give a practical test that determines if nonempty RNNs do exist (proposition 4). In the second one, we define a Repeated Nash Equilibrium (RNE) where different networks may form in the different periods of the game. The set of RNE includes the set of RNN as a special case. We show that there exist structural relationships between the different networks that form within a given RNE. We study the intertemporal consistency between networks and identify several properties. Proposition 5 exhibits some sequences of networks that cannot be achieved as an outcome of a nonempty RNE. For instance, a sequence of expanding connected networks cannot constitute an equilibrium outcome, or it is not possible that all players remain isolated during the last two periods. Nevertheless, these restrictions allow for several RNE with nonempty outcomes. The networks within a given RNE may exhibit a

total order relationship, the smaller network being formed in the last stage. In particular, there exist RNE in which sequences of contracting networks form, or in which sequences of networks can expand for all but the last period in which connections brutally run low. We also show that there are RNE in which players forget links (proposition 6). Precisely, two directly connected players in one period are not directly connected in another period. In spite of the restrictions on the intertemporal consistency between networks in equilibrium outcomes, we prove that any network can emerge in the outcome of a RNE (proposition 7). We also investigate the question of efficiency for both Bentham and Pareto criteria. In the finitely repeated game, the structure of efficient strategy profiles is closely related to the structure of static efficient networks (propositions 8 and 9). In addition, we prove that the sets of Bentham-efficient networks (strategy profiles) and Pareto-efficient networks (strategy profiles) are identical. The set of efficient strategy profiles of the repeated game is most often slightly reduced when players are assumed to be boundedly rational (proposition 10).

Some other papers are concerned with the structural properties of NE in repeated games with finite automata.[2] Papers relating to repeated games are mainly concerned with the set of average payoffs that can be achieved in RNE. In the current work, we are more interested in the structure of RNE than in the induced average payoffs. The differences with other papers studying a dynamic network formation[3] are mostly related to the aims of the studies. These papers are concerned with the formation of a static network as a result of several steps of a dynamic process. Limiting networks are studied in Watts (2001), and learning or stochastic stability are used to identify limiting equilibria in Bala, Goyal (2000). By contrast, we consider a finitely repeated game that consists in the formation of a static network in each step of the process. We are interested in understanding the influence of different levels of rationality on equilibrium structures in this finite-horizon repeated setting.

The rest of the paper is organized as follows. Section 2 presents preliminaries and notations. We also determine the set of Nash networks in the static game and in the finitely repeated game for the case of perfectly rational players. Then the machine game is introduced and studied in section 3. We start with results on RNN and continue with results on structural properties of more elaborated RNE. Results dealing with the efficiency of networks and strategy profiles are in section 4. Once again, we distinguish results according to players' rationality. Section 5 concludes. Proofs not given in the body

[2]See Rubinstein (1986), Abreu, Rubinstein (1988) and Piccione, Rubinstein (1993) among others.

[3]We refer the reader to Bala, Goyal (2000), Currarini, Morelli (2000), Dutta, Ghosal, Ray (2005), Goyal, Vega-Redondo (2005), Jackson, Watts (2002) and Watts (2001, 2002).

of the paper appear in an appendix.

2 Preliminaries and notations

2.1 Static one-period game

Let $G = (I, A, \pi)$ be a finite n-player game in normal form. The set $I = \{1, \ldots, n\}$ is the player set. For any $i \in I$, A_i is player i's set of actions and $A = \prod_{i \in I} A_i$ is the set of action profiles. Let a_{-i} be the actions chosen by all players except i and $A_{-i} = \prod_{j \neq i} A_j$. Player i's payoff function is $\pi_i : A \longrightarrow I\!R$. An action profile a is a NE of G if for any $i \in I$ and any $a'_i \in A_i$,

$$\pi_i(a) \geq \pi(a'_i, a_{-i}).$$

2.2 Network

The n players are connected in some network relationships. We limit our discussion to non-directed networks on the player set I. As in Jackson, Wolinsky (1996), two players are either related to each other or not, but it cannot be that one is related to the second without the second being related to the first. We write ij to describe the *link* between two players i and j.

Let $g_I = \{ij : i, j \in I, i \neq j\}$ be the set of all potential links. Any set of links $g \subseteq g_I$ defines a *network*. We apply the convention that $g = g_I$ is the *complete* network and that $g = g_0 = \{\emptyset\}$ is the *empty* network. Any (spanning) subset $g' \subset g$ is called a *subnetwork* of g.

A *path* between players i and j in a network g is a sequence of distinct players $i_1, \ldots, i_K$ such that $i_k i_{k+1} \in g$ for each $k \in \{1, \ldots, K-1\}$ where $i_1 = i$ and $i_K = j$. Two such players are said to be connected. Player i is in a *cycle* of network g if there is a path with $K \geq 3$ players such that $i_1 = i_K = i$.

Let $n_i(g) = \{j \in I | ij \in g\}$ be the set of *neighbors* (or direct connections) of player i. Let $N_i(g)$ be the set of players to whom player i is connected in network g. Obviously, $n_i(g) \subseteq N_i(g)$. A network g is connected if there is a path between any two players. Alike, network g is said to be k-connected if there does not exist a set of $k-1$ links whose removal disconnects the network. If g is not connected, its connected subnetworks are called *components*. A connected acyclic network (or 1-connected network) is called a *tree* and a non connected network whose distinct components are trees is called a *forest*.

Let $l_{i,j}(g)$ be the *distance* between two players i and j in network g. If i and j are connected, $l_{i,j}(g)$ is the number of links in the shortest path be-

tween i and j. By convention, if i and j are not connected, $l_{i,j}(g) = \infty$. Let $L_i(g) = \max_{j \neq i} l_{i,j}(g)$ be player i's *eccentricity* in network g. The *diameter* of network g is $L(g) = \max_{i \in I} L_i(g)$. The last two definitions apply to any component $g' \subset g$.

For any two distinct players $i, j \in I$, $g + ij = g \cup \{ij\}$ is the network obtained adding link ij to network g. Likewise, let $g - ij = g \backslash \{ij\}$ be the network obtained removing link ij from network g. The intersection $g \cap g'$ defines the set of links that networks g and g' have in common.

2.3 Link formation cost and inefficient links

In this section, we present a non-cooperative model of costly network formation with consent. We assign a network $g(a) \subseteq g_I$ to every action profile $a \in A$. Each player $i \in I$ has an action set $A_i = \{(a_{ij})_{j \neq i} : a_{ij} \in \{0, 1\}\}$. Player i seeks contact with player j if $a_{ij} = 1$. Link ij forms if both players seek contact. The network induced by a is given by

$$g(a) = \{ij \in g_I : a_{ij} = a_{ji} = 1\}.$$

If player i seeks contact with j, then he supports a cost $c > 0$. As in a wide range of models of costly network formation, player i's payoff consists in his value of the network minus a cost c for any attempt he made to create links. We assume that the value of network $g(a)$ for player i only depends on the number $\#N_i(g(a))$ of players to whom i is connected where $\#$ gives the dimension of the set. Thus, the distance between two connected players does not matter. This is true of many networks such as Internet, economic partnership or subcontracting. This results in the following payoff function

$$\pi_i(a) = v\#N_i(g(a)) - c\sum_{j \neq i} a_{ij}, \tag{1}$$

which can be seen as a particular case of the class of payoff functions investigated by Gilles, Sarangi (2004). Following Watts (2002), we assume $c > v > 0$, that is, creating a link is more costly than the reward of a single direct connection. In other words, player i needs some indirect connections to obtain a positive payoff. We also fix an upper bound $2v > c$ to avoid trivial cases due to an immeasurable cost of creating a link. The next example will help discussing these assumptions.

Example 1
Consider $I = \{1, 2, 3, 4, 5, 6\}$ and the network that forms as a result of the following players' choices:

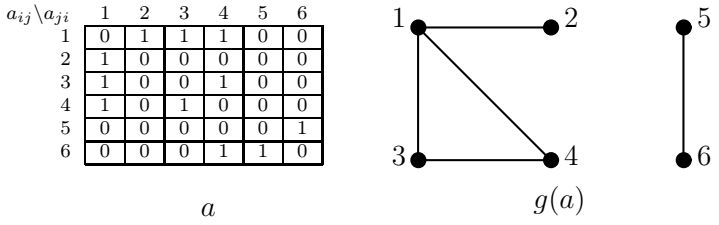


Figure 1.

Network $g(a)$ consists in two components. The value of $g(a)$ is identical for players 1, 2, 3 and 4 in the left component and their payoffs only differ in the cost supported. Player 2's payoff is thus larger than 1's payoff since 2 intends to form a link with 1 while 1 intends to form links with 2,3 and 4. The subset $\{1, 3, 4\} \subset I$ defines a cycle in $g(a)$. Since the distance between two connected players has no influence on payoffs, it is not the interest of player 1 to seek contact with 3 since they are already connected by player 4. Intuitively, link 13 is superfluous. Now consider the right component of $g(a)$ involving players 5 and 6. This component is a tree, but $g(a)$ is not a forest since the right component includes a cycle. The assumption $c > v$ implies that both 5 and 6 would gain to remain isolated as it is the case in all components of diameter 1. Remark also that 6 seeks contact with 4. The link 64 fails to form since $a_{46} = 0$. As player 6's attempt is not drawn on $g(a)$ even if it affects his payoff, we will distinguish the action profile a and its induced network $g(a)$. □

Two identical networks $g(a) = g(b)$ may correspond to distinct action profiles a and b. Precisely, $g(a) = g(b)$ for two action profiles $a \neq b$ across the player set I if both networks have the same set of links and if at least one player seeks the creation of a link but does not receive the consent of the opponent. We will focus on a particular form of action profiles, namely the *cost-efficient* action profiles.

Definition 1 *An action profile a is cost-efficient if there is no $(i, j) \in I^2$ such that $a_{ij} \neq a_{ji}$.*

Let A^{ce} be the set of cost-efficient action profiles. In any cost-efficient action profile, if a link between players i and j fails to form, neither i nor j seeks contact to create it. The only cost-efficient action profile that enables the empty network to form is denoted $a_0 = (a_{1,0}, \ldots, a_{n,0}) \in A^{ce}$ where $a_{i,0} = (0, \ldots, 0)$. Clearly, the action profile a in example 1 is not cost-efficient.

If the payoff function is given by (1) and $2v > c > v > 0$, player i may obtain a larger payoff if he removes some links. Abusing notations, if $ij \in g(a)$, let $a_i(j^-)$ be the action that differs from a_i only by $a_{ij} = 0$.

Definition 2 *A link $ij \in g(a)$ is superfluous for player i in network $g(a)$ if $\pi_i(a_i(j^-), a_{-i}) > \pi_i(a)$. This inequality is satisfied in two cases:*

1. *$N_i(g(a) - ij) = N_i(g(a))$, i.e. link ij belongs to a cycle;*
2. *$n_j(g(a)) = \{i\}$, i.e. player i is j's single neighbor.*

In case 1, the removal of ij does not alter player i's connection set. This is the case of link 13 in example 1. In case 2, link ij increases the value of the network but costs to player i more than it yields since it provides him with a single connection and $c > v$. In example 1, this is the case of link 56 for both players 5 and 6. Let $d_i^1(a)$ (resp. $d_i^2(a)$) denote the total gain for player i that results from removing all superfluous links in case 1 (resp. case 2). We have

$$d_i^1(a) = c\left(\#\{j \in I : N_i(g(a) - ij) = N_i(g(a))\} - \mathcal{C}_i\right), \text{and}$$

$$d_i^2(a) = (c - v)\#\left\{ij \in g(a) : n_j(g(a)) = \{i\}\right\},$$

where $\mathcal{C}_i$ is the number of cycles with i as unique common player. Now let $d_i(a) = d_i^1(a) + d_i^2(a)$ be the maximal gain for player i if he switches from action a_i against a_{-i}.

For any $i \in I$, we define $b_i \in A_i$ as player i's best response against a_{-i}, that is

$$b_i \in \arg\max_{a_i' \in A_i} \pi_i(a_i', a_{-i}).$$

When $a \in A^{ce}$, b_i is the action that satisfies

$$\pi_i(b_i, a_{-i}) - \pi_i(a) = d_i(a),$$

since a player cannot create links by his own will. All these definitions and notations will be used in the rest of the paper to determine whether action profiles are NE or efficient. Now we begin by characterizing Nash networks.

Proposition 1 *The only Nash network of G is the empty network $g_0 = g(a_0)$.*

A direct consequence of proposition 1 is that any player $i \in I$ can secure a null payoff against any opponents' behaviors by choosing $a_{i,0}$. The *minmax* payoff of each player is then 0. Now we introduce the finitely repeated game.

2.4 Finitely repeated game and network formation

In the static game G, players will not create any link even if they would be better off in some nonempty networks. In this section we assume that the players are involved in a T-period repeated game G^T that consists in

$T < \infty$ repetitions of game G in period $t = 1, 2, \ldots, T$. Throughout the paper, we assume $T > 3$. It is natural to consider a dynamic process of network formation. Such a framework fits with many economic situations in which relationships between agents may evolve with time.

We take the view that at the beginning of period t, all players observe a^{t-1} and not just $g(a^{t-1})$. If network $g(a)$ in example 1 forms in period $t-1$, then all players know at the beginning of period t that player 6 has contacted player 4 even if link 64 fails to form.

Thus, a history of play $h^t = (a^1, \ldots, a^{t-1})$ at period t records the action profiles chosen by each player in periods $1, \ldots, t-1$. Let H^t denote the set of histories at period t and $H = \{\emptyset\} \cup (\bigcup_{t=2}^{T} H^t)$ denote the set of all possible histories of play. A repeated game strategy s_i is a sequence $s_i = \{s_i^t\}_{t=1}^{T}$ where $s_i^t : H^t \longrightarrow A_i$ models player i's action played at period t as a function of the $t-1$ previous action profiles. For any $i \in I$, S_i is the set of strategies for player i. Let $s = (s_1, \ldots, s_n)$ denote a strategy profile and $h^{T+1}(s) = (a^1, \ldots, a^T)$ be the repeated game outcome induced by $s \in S$, where $S = \prod_{i \in I} S_i$ is the set of all strategy profiles.

Player i's payoff function $\tilde{\pi}_i : S \longrightarrow I\!R$ from playing G^T is evaluated according to the average payoff

$$\tilde{\pi}_i(s) = \frac{1}{T} \sum_{t=1}^{T} \pi_i(a^t), \tag{2}$$

and $\tilde{\pi} = (\tilde{\pi}_1, \ldots, \tilde{\pi}_n)$. A strategy profile $s \in S$ is a NE of the repeated game G^T if, for each $i \in I$ and $s_i' \neq s_i$,

$$\tilde{\pi}_i(s) \geq \tilde{\pi}_i(s_i', s_{-i}).$$

Since a network is built in each period of the repeated game, it is useful to introduce the notion of *repeated network*.

Definition 3 *A strategy profile s^* induces a repeated network based on network $g(a^*)$ if s^* induces a repeated game outcome $(a^1, \ldots, a^T)$ that satisfies $g(a^t) = g(a^*)$, $\forall t = 1, \ldots, T$.*

A repeated network is simply a network that forms in all periods as a result of players' actions. We want to highlight the robustness of a network being formed period after period. This notion is intuitively related to the robustness of a given set of relationships. We may think about situations in which trust is established among agents on a long-term basis. A Repeated Nash Equilibrium (RNE) is a NE of the repeated game. We need to define a Repeated Nash Network (RNN).

Definition 4 *A strategy profile s^* is a RNN based on network $g(a^*)$ if the induced outcome is a repeated network based on $g(a^*)$ and if it is a RNE of G^T.*

Thus the set of RNE includes as a special case the set of RNNs. When the horizon is finite, the only RNE induces the formation of the only Nash network in all periods, *i.e.* players always remain isolated.[4]

Proposition 2 *Assume $T < \infty$. The only RNE is the one in which the empty network g_0 forms in all periods.*

In the present section the agents are assumed to be perfectly rational. This results in an extreme conclusion: players have an incentive to remain unconnected. Now, we are going to relax the assumption of perfect rationality in order to understand the resulting differences on the types of relationships that are likely to appear.

3 Machine game in a finite horizon setting

In this section, we assume that players use finite automata with a limited number of states to play their strategies. We also focus on the subset of cost-efficient action profiles. This restriction may be justified by the fact that only cost-efficient action profiles are likely to define Nash networks in the static game. Moreover, for any network structure, the corresponding cost-efficient action profile is the most efficient action profile that induces the network (see section 4). In other words, only cost-efficient action profiles are likely to define efficient networks. We begin with the study of RNNs. Then, we examine RNE which are not RNNs. We will prove that the structure of both types of equilibria becomes non degenerate.

3.1 Machine game

Following Neyman (1985), we focus on the repeated network formation game G^T in which player $i \in I$ chooses a *finite automaton* M_i to play his strategy. A finite automaton M_i for player i is a four-tuple $(Q_i, q_i^1, \lambda_i, \mu_i)$ where

1. Q_i is the finite set of states in M_i, with $\#Q_i = m_i$;
2. q_i^1 is the initial state;
3. $\lambda_i : Q_i \longrightarrow A_i$ is the output function, which plays action $\lambda_i(q_i) \in A_i$ whenever M_i is in state q_i;
4. $\mu_i : Q_i \times A_{-i} \longrightarrow Q_i$ is the transition function. In a given period, if M_i is in state $q_i \in Q_i$ and players $-i$ choose $a_{-i} \in A_{-i}$, then the next state of the machine is $\mu_i(q_i, a_{-i}) \in Q_i$.

[4] When the game is infinitely repeated, we can prove a folk theorem like result for RNE and RNN. Such a result includes a very large panel of structures even if, for instance, there cannot be a RNN based on a star network.

We assume that player i's strategy space is limited to the set $\mathcal{M}_i$ of all automata of size (the number of states in the machine) $1 < m_i < T$. Before proceeding to the results, we must apply these notations to the repeated network formation game. A strategy profile of the machine game is an n-tuple $(M_1, \ldots, M_n)$ of automata. We also use the notation (M_i, M_{-i}) instead of $(M_1, \ldots, M_n)$. Abusing notations, we keep up writing a^t for the action profile chosen in period t instead of using the notation $(\lambda_1(q_1^t), \ldots, \lambda_n(q_n^t))$.

3.2 Existence of nonempty RNN

The main goal of this section is to provide a sufficient condition for a nonempty network to be sustained as a RNN. Before stating the result, notice that there cannot be a RNN based on a star network. In fact, the central player obtains a negative payoff as he supports a cost for creating a link with each opponent.

Proposition 3 *Consider any network $g(a^*)$ such that $a^* \in A^{ce}$ and, for any $i \in I$,*

$$0 \leq d_i(a^*) \leq \min \left\{ \frac{v}{2} \# N_i(g(b_i, a^*_{-i})), 2\pi_i(a^*) \right\}. \tag{3}$$

Then there exists a RNN $(M_1, \ldots, M_n)$ such that $g(a^*)$ forms in each period.

Proof. Suppose that condition (3) is satisfied. Choose any $a^* \in A^{ce}$. We show that there exists a RNN $(M_1^*, \ldots, M_n^*)$ that induces $g(a^*)$ in all periods by constructing the required automata. For any player $i \in I$, let us consider the trigger strategy s_i defined for $t = 1$ by $s_i^1(\emptyset) = a_i^*$, and for $t > 1$, by

$$s_i^t(h^t) = \begin{cases} a_i^* & \text{if } a_{-i}^\tau = a_{-i}^*, \forall \tau = 1, \ldots, t-1, \\ a_{i,0} & \text{otherwise.} \end{cases}$$

This strategy is implemented by the two-state automaton M_i^1 represented below:

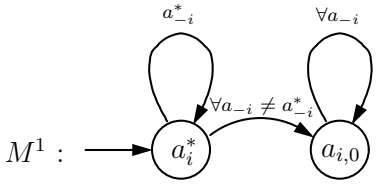


Figure 2.

If the opponents play M_{-i}^1, then player i's payoff from playing M_i^1 is

$$\tilde{\pi}_i(M_i^1, M_{-i}^1) = \pi_i(a^*) \geq 0, \tag{4}$$

since the outcome of the machine game is assumed to be a repeated network.

Notice that any deviation from M_i^1 by player i releases a definitive punishment by players $j \neq i$. Recall that each player i is limited to automata

of size $1 < m_i < T$. This prevents player i from using the standard best response against M^1_{-i} that consists in playing a^*_i to create $g(a^*)$ until the last stage and then playing b_i in round T. In fact such a strategy requires at least a T-state automaton. As a consequence, a deviation by player i must occur in a period $t < T$ and implies $T - t \geq 1$ periods of punishment. This also implies that if player i has an incentive to deviate, then this deviation must occur as late as possible in the game. However, as we will see below, this does not exactly amount to say that if player i has an incentive to deviate, he will do so in stage $T - 1$.

Any deviation in stage $t < T$ is followed by some punishment stages. Thus, player i aims at minimizing the cost of seeking contacts in the actions he plays in periods $t+1, \ldots, T$ against $a_{-i,0}$ since the value of the network in each of these periods is null. Even if the amount paid for seeking contacts in the deviating action is less than a^*_i, it may be very costly. Therefore, it may not be the interest of player i to play this action at stage $t+1$ and thereafter. This is why all possible deviations for player i's may be grouped into the two following cases:

1. He plays b_i in periods $T - 1$ and T,[5]
2. He plays b_i in period $k \leq T - 2$ and uses a $(k + 1)$th state that plays $a_{i,0}$ for the remaining stages. Clearly, player i's highest incentive to deviate is in period $k = T - 2$.

We now consider these two possibilities.

Case 1.
The deviation is implemented by the following $(T - 1)$-state automaton:

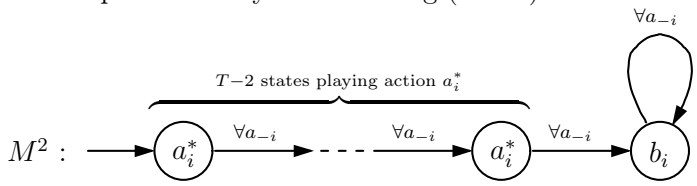


Figure 3.

Using M^2_i, player i mimics a full cooperation to form network $g(a^*)$ up to period $T-2$, and then plays b_i to obtain the best payoff against $a^*_{-i} = a^{T-1}_{-i}$.[6]

[5]Recall that using $T-2$ states for playing a^*_i and one more state for playing b_i prevents player i from playing $a_{i,0}$ in stage T as a response to $a_{-i,0}$. Moreover, player i has not interest in deviating at stage $T - 1$ to any action a_i which is less costly than b_i. This is due to the fact that using a_i removes at least two more players from i's set of connections than b_i. Therefore, $\pi_i(a_i, a^*_{-i}) + \pi_i(a_i, a_{-i,0}) < \pi_i(b_i, a^*_{-i}) + \pi_i(b_i, a_{-i,0})$ due to $2v > c$.

[6]Notice that action profiles considered in the deviation tests are not cost-efficient. In fact removing links from a^*_i by playing b_i induces an action profile $(b_i, a^*_{-i}) \notin A^{ce}$ that is not cost-efficient.

He also plays b_i in period T against the punishment $a_{-i,0}$ since

$$b_i = \arg \max_{a_i \in \{b_i, a_i^*\}} \pi_i(a_i, a_{-i,0}) = \arg \min_{a_i \in \{b_i, a_i^*\}} -c \sum_{j \neq i} a_{ij}.$$

In words, player i's machine has not enough states to use another action than b_i or a_i^*. Moreover, the cost of seeking contacts in b_i cannot be more expensive than in a_i^* since the assumption $a^* \in A^{ce}$ prevents the deviating player from creating links unilaterally. Player i obtains the payoff

$$\tilde{\pi}_i(M_i^2, M_{-i}^1) = \frac{(T-2)\pi_i(a^*) + \pi_i(b_i, a_{-i}^*) + \pi_i(b_i, a_{-i,0})}{T}, \tag{5}$$

which has to be compared to (4). It is not the interest of player i to switch from M_i^1 to M_i^2 if and only if

$$\begin{aligned}
& \pi_i(a^*) \geq \frac{(T-2)\pi_i(a^*) + \pi_i(b_i, a_{-i}^*) + \pi_i(b_i, a_{-i,0})}{T} \\
\iff \quad & \pi_i(a^*) \geq \frac{v}{2} \# N_i(g(b_i, a_{-i}^*)) - c \sum_{j \neq i} a_{ij}^{*d} \\
\iff \quad & \pi_i(a^*) \geq \pi_i(b_i, a_{-i}^*) - \frac{v}{2} \# N_i(g(b_i, a_{-i}^*)) \\
\iff \quad & d_i(a^*) \leq \frac{v}{2} \# N_i(g(b_i, a_{-i}^*)). \qquad (6)
\end{aligned}$$

Case 2.
The deviation is implemented by the following $(T-1)$-state automaton:

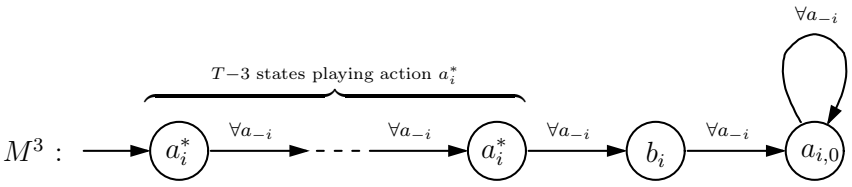


Figure 4.

Using M_i^3, player i simulates a full cooperation to create network $g(a^*)$ up to period $T-3$, and then plays in stage $T-2$ the action b_i against $a_{-i}^* = a_{-i}^{T-1}$. He uses a new state playing $a_{i,0}$ for the last two stages. Thus, player i obtains the average payoff

$$\tilde{\pi}_i(M_i^3, M_{-i}^1) = \frac{(T-3)\pi_i(a^*) + \pi_i(b_i, a_{-i}^*) + 2\pi_i(a_{i,0}, a_{-i,0})}{T},$$

which also needs to be compared to (4). It is not the interest of player i to switch from M_i^1 to M_i^3 if and only if

$$\begin{aligned} \pi_i(a^*) &\geq \frac{(T-3)\pi_i(a^*) + \pi_i(b_i, a^*_{-i}) + 2\pi_i(a_{i,0}, a_{-i,0})}{T} \\ \iff \quad d_i(a^*) &\leq 2\pi_i(a^*). \end{aligned} \tag{7}$$

Combining the two cases, player i will not deviate from strategy M_i^1 if and only if

$$d_i(a^*) \leq \min\left\{\frac{v}{2}\#N_i(g(b_i, a^*_{-i})), 2\pi_i(a^*)\right\},$$

the condition stated in the proposition. By definition, $d_i(a^*) \geq 0$ so that inequality (3) guarantees that players obtain at least the minmax payoff. Thus, the strategy profile $(M_1^*, \ldots, M_n^*)$ is a RNN based on $g(a^*)$. ■

Remark that the condition in proposition 3 does not depend on whether T is large or not since the average payoff function does not include a discount parameter. This condition implies that RNNs must be based on networks that are not too much over-connected if players are very rancorous. Indeed, the gain that each player obtains may not be too small. In static network theory, the fact that networks be too much over-connected is important too, mostly for efficiency considerations (Calvó-Armengol, 2003). The condition in proposition 3 also implies that a player must not support a too large share of the cost needed to connect his component. For instance $g(a^*)$ cannot contain a star subnetwork since a player would be directly connected to each opponent and obtains in average less that the minmax payoff.

Condition (3) is necessary and sufficient to prevent any deviation from (M_i^1, M_{-i}^1). However, it is not a necessary condition to achieve a RNN based on $g(a^*)$. To see this, notice that M^1 induces the hardest possible punishment, either for the duration or for the loss of static game payoffs per period. One can think of an automaton that induces a less significant loss of payoff for a single period of punishment. If all players use such a machine and none of them has an incentive to deviate, then the necessary and sufficient condition to achieve a RNN based on network $g(a^*)$ must be less restrictive than (3). We limit our result to the sufficient condition because it is difficult to determine the action for which the threat of punishment is minimal.

Let $g(l)$, $l \in A^{ce}$, be an n-player line network: for any $i \in I\backslash\{h, j\}$, $n_i = 2$ and $n_h = n_j = 1$ and all players have $n-1$ connections. The line begins and ends with players h and j. The next proposition gives two practical tests to determine whether the set of RNN contains nonempty repeated networks. In the first test, it is enough to check if a repeated line network is not a RNN

to make sure that any other nonempty repeated network is not a RNN. This test proves useful if one wants to check that there cannot be a RNN based on a large and complex static network. In the second test, it is sufficient to look at the number of players to guarantee that some nonempty RNNs do exist.

Proposition 4 *(i) Suppose that the repeated network based on $g(l)$ cannot be a RNN. Then the only RNN of G^T induces the empty network g_0 in all periods.*

(ii) Suppose $n \geq 5$. Then there exists a nonempty RNN of G^T.

Proof. *(i)* Suppose that there is no automata profile such that the repeated network based on $g(l)$ is a RNN. Firstly, observe that if the value to a player of any network $g(a)$ is more than v, then there is at least one player whose cost is $2c$ or more. Given that, $g(l)$ is the network in which the minimal payoff across the player set is maximal across the set of all networks. Secondly, the players in the line network are connected with a minimal number of links and the cost for creating links is distributed such that a player pays at most $2c$, which cannot be less in a connected network. In other words, the network $g(l)$ is the unique architecture that satisfies the maxmin criterion. That is,

$$\min_{i \in I} \pi_i(l) = \max_{a \in A^{ce}} \min_{i \in I} \pi_i(a). \tag{8}$$

It follows from (8) that if a player has an incentive to deviate from the repeated network based on $g(l)$, then there is at least one player who can do so in a repeated network based on any network $a \in A^{ce}$. The assumption that the repeated network based on $g(l)$ is not a RNN implies that the only RNN of G^T induces a Nash network in all periods. Therefore, the empty network must form in all periods.

(ii) We next prove that $g(l)$ is sustained as a RNN (whatever $2v > c > v$) if $n \geq 5$. We are going to use the sufficient condition of proposition 3. In the cost-efficient action profile that induces network $g(l)$, all players except the first and the last of the line seek the creation of two links (the cost to each of these players in the network is $2c$), while they benefit from connections with all opponents (in fact, the value of $g(l)$ is $v(n-1)$ for all players). For players h and j, the value of $g(l)$ remains $v(n-1)$ as $g(l)$ is connected but they only create a single link. As a consequence,

$$\min_{i \in I} \pi_i(l) = v(n-1) - 2c,$$

In $g(l)$, only the neighbors of h and j have a superfluous link. Let i be h's neighbor. It follows that $b_i = l_i(h^-)$ which implies that $g(b_i, l_{-i})$ consists in

an $(n-1)$-player line of extermities i and j and an isolated player h. Player i has $d_i(l) = c - v$. Therefore, $d_i(l) \geq d_j(l)$ for any player $j \neq i$. By condition 3, there is a RNN based on $g(l)$ if

$$c - v \leq \min\left\{\frac{v}{2}(n-2), 2(v(n-1) - 2c)\right\}.$$

The ineguality $c - v < v(n-2)/2$ is satisfied when $n \geq 4$ and the ineguality $c - v < 2(v(n-1) - 4c)$ is satisfied when $n \geq 5$. Thus, the condition $n \geq 5$ is enough to guarantee the existence of nonempty RNN. ■

So far, two sufficient conditions for the existence of nonempty RNNs are provided. The second one has been stated in the most general form to keep the exposition as simple as possible. We may notice that any line network with at least 5 players can be sustained as a RNN. The results of this section enable to make precise the cases in which nonempty RNNs do exist. But not much is said about the structural properties of RNE that are not RNNs. This is the aim of the next section.

3.3 Structural properties of RNE

In this section, we are mainly concerned with the structural properties of RNE of G^T. We identify a property satisfied by any RNE. This property has a crucial impact on the intertemporal consistency between networks that form in the outcome of RNE. In propositions 5 and 6, we use graph theory to characterize these restrictions and to represent the sequences of networks that can be achieved at equilibrium. We also offer economic interpretations.

We begin with an intuitive property satisfied in any RNE of G^T. There is a key argument in the analysis of the structure of RNE. To see this, suppose that player i uses an action a_i^k in a period $k < T$ that is more beneficial than a_i^T against actions a_{-i}^T used by the opponents in stage T. It is the interest of player i to play this action in the last stage since he cannot be punished in a forthcoming period and a $(T-1)$-state automaton can do it. Therefore, players must not have such an opportunity to deviate in any RNE.

Lemma 1 *Consider any RNE (M_i^*, M_{-i}^*) of the machine game G^T with outcome $(a^{*1}, \ldots, a^{*T})$. Then there is no network $g(a^{*k})$ that forms in period $k < T$ such that for any player $i \in I$, $\pi_i(a_i^{*k}, a_{-i}^{*T}) > \pi_i(a^{*T})$.*

Proof. The proof is by contradiction. Consider a RNE (M_i^*, M_{-i}^*) of G^T, a player i and a period $k < T$ in which network $g(a^{*k})$ forms such that $\pi_i(a_i^{*k}, a_{-i}^{*T}) > \pi_i(a^{*T})$. We prove that player i has an incentive to deviate from M_i^* towards the following $(T-1)$-state automaton M_i:

1. $Q_i = \left\{q_i^{s_1}, \ldots, q_i^{s_t}, \ldots, q_i^{s_{T-1}}\right\}$, $m_i = T-1$;
2. $q_i^1 = q_i^{s_1}$;
3. $\lambda_i(q_i^{s_t}) = a_i^{*t} \in h^{T+1}(M_i^*, M_{-i}^*) = ((a_i^{*1}, a_{-i}^{*1}), \ldots, (a_i^{*T}, a_{-i}^{*T}))$, $\forall t \leq T-1$;
4. $\mu_i(q_i^t, \lambda_{-i}(q_{-i}^t)) = \begin{cases} q_i^{s_{t+1}} & \text{if } t \leq T-2 \\ q_i^{s_k} : \lambda_i(q_i^{s_k}) = a_i^{*k} & \text{if } t = T-1 \end{cases}$

The output function indicates that for each $t < T$ the action used by M_i^* against M_{-i}^* in period t is played when in state $q_i^{s_t}$. The transition function of M_i mimics the sequence of actions played by M_i^* against M_{-i}^* for all but the last period. In any period $t < T$, the deviation towards M_i keeps player i's payoff unchanged. In stage T, M_i transits to the state $q_i^{s_k}$ that implements action a_i^{*k} used by M_i^* against M_{-i}^* in stage k. Using M_i, player i obtains the following average payoff:

$$\tilde{\pi}_i(M_i, M_{-i}^*) = \tilde{\pi}_i(M_i^*, M_{-i}^*) + \frac{\pi_i(a_i^{*k}, a_{-i}^{*T}) - \pi_i(a^{*T})}{T},$$

which is larger than $\tilde{\pi}_i(M_i^*, M_{-i}^*)$ since $\pi_i(a_i^{*k}, a_{-i}^{*T}) > \pi_i(a^{*T})$ by assumption. This contradicts the initial assumption that (M_i^*, M_{-i}^*) is a RNE. ■

The idea that a player may use a former state to deviate in stage T without being punished is central in the question of the architecture of RNE. Lemma 1 states restrictions on the intertemporal consistency between static networks that form within a given RNE of G^T. Unfortunately, these restrictions are described in terms of payoff. We are more interested in structural properties of the networks induced by such RNE. We specify some of these properties in points *(i)* and *(ii)* of the next proposition. The third point is related to both proposition 2 and lemma 1.

Proposition 5 *The outcome $(a^1, \ldots, a^T)$ induced by any RNE $(M_1, \ldots, M_n)$ of game G^T must have the three following features:*

(i) there is no connected network $g(a^t) \subset g(a^T)$, $\forall t < T$,

(ii) in network $g(a^T)$, there is no player i with eccentricity $L_i(g(a^T)) = 1$,

(iii) if there is a period $t < T-1$ such that $g(a^t) \neq g_0$, then it cannot be the case that $g(a^{T-1}) = g(a^T) = g_0$.

Proof. *(i)* The proof is by contradiction. Consider any RNE $(M_1, \ldots, M_n)$ of G^T in which, for some $t \in \{1, \ldots, T-1\}$, a connected network $g(a^t) \subset g(a^T)$ forms (see definition page 4). The assumptions $g(a^t)$ connected and $g(a^t) \subset g(a^T)$ imply of course that $g(a^T)$ is also connected. By definition 2, there is a player i who has some superfluous links in $g(a^T)$ that he does

not have in $g(a^t)$. This player is able to play a_i^t in stage T against a_{-i}^T by a mechanism similar to that in proof of lemma 1. In network $g(a_i^t, a_{-i}^T)$ player i maintains a connection with all opponents since $g(a^t)$ is a connected subnetwork of $g(a^T)$ (in fact, $\#N_i(g(a_i^t, a_{-i}^T)) = \#N_i(g(a^T)) = n-1$). Moreover the cost to player i of forming links in $g(a_i^t, a_{-i}^T)$ is smaller since he seeks less contacts in a_i^t than in a^T. Thus, if player i chooses the machine that simulates what plays M_i against M_{-i} for all but the last period and then transits to the state used in period t, he obtains a larger payoff in stage T. Player i obtains a larger average payoff, which implies that (M_i, M_{-i}) is not a RNE. This contradicts the initial assumption.

(ii) By contradiction, consider a RNE $(M_1, \ldots, M_n)$ in which $L_i(g(a^T)) = 1$ for a player $i \in I$. Such an eccentricity means that player i is directly connected with each opponent. This implies that player i obtains the worst possible stage payoff $\pi_i(g(a^T)) = (n-1)(v-c)$. Since $(M_1, \ldots, M_n)$ is a RNE, we know that $\tilde{\pi}_i(M_1, \ldots, M_n) \geq 0$. Thus player i obtains a positive payoff in some periods, that is, he does not seek contact with each opponent in these periods. Formally, there exists $t < T$ such that $L_i(g(a^t)) > 1$. Let a_i^t be the action played by player i in such a period. As in the proof of lemma 1, player i is able to deviate from M_i towards a $(T-1)$-state machine that simulates what plays M_i against M_{-i} in the first $T-1$ periods and then transits in stage T to the state playing action a_i^t. Using this altered strategy, player i must obtain in stage T a payoff $\pi_i(a_i^t, a_{-i}^T) > \pi_i(a^T)$ as he seeks less contacts and $\pi_i(a^T)$ is the worst payoff in the game. All other stage payoffs being identical, the deviating strategy yields player i a larger average payoff than M_i. This contradicts the fact that $(M_1, \ldots, M_n)$ is a RNE.

(iii) Consider any RNE $(M_1, \ldots, M_n)$ for which $g(a^{T-1}) = g(a^T) = g_0$ and for some periods $t \leq T-2$, $g(a^t) \neq g_0$. Let

$$t^* = \max_{t \leq T-2} \{t : g(a^t) \neq g_0\}$$

be the most remote period in which a nonempty network forms. By proposition 1, at least one player i is such that $d_i(a^{t^*}) > 0$. Suppose that i chooses to deviate from M_i towards a (t^*+1)-state machine that mimics M_i's behavior against machines M_{-i} up to period t^*-1, then plays action b_i in stage t^* and transits to a (t^*+1)th state playing action $a_{i,0}$ until the end of the game (see proof of lemma 1). Clearly, such a deviation yields player i a larger payoff, which implies that $(M_1, \ldots, M_n)$ is not a RNE. We conclude that two empty networks cannot form in the last two stages of a nonempty RNE. ■

These results lead to some conclusions. Result *(i)* has several interpretations. Firstly, the only minimal network that is likely to be connected is the last that forms. Secondly, there may be other connected networks in

previous periods but this result implies that these networks must contain the last one. We may say that the formation of a connected network (if it occurs) has to be progressive. A connected network may form quickly in the process but it will be over-connected. If a connected network forms in a given period $t < T$ (of a RNE) and another one forms in the final period, the last network is more beneficial to all players and strictly more to one of them. Even a link formation process by a player generates an externality on the set of direct neighbors, this effect would be gradually internalized by some players. Thirdly, one may also interpret the first result in proposition 5 as the impossibility that the outcome of a RNE consists in a sequence of connected networks that extends as time goes by. This once again emphasizes that too much over-connected networks fail to form in a RNE.

Result *(ii)* shows that in the last network that forms, a player cannot create a direct link with all opponents.[7] For instance, the complete network and the star network are two such networks, and cannot form in the last stage of a RNE. Recall that point *(i)* does not prevent players from creating a connected network in stage T. By point *(ii)*, if players are all connected in the last network induced by a RNE, any of them avoids the burden to seek contact with all others. This shows how players learn to divide the task of connecting the network up among themselves. A consequence of this result is that the diameter of the network formed in the final stage must satisfy $L(g(a^T)) > 1$. This could be interpreted as the absence of an extreme small-world effect as observed by Milgram (1967). Notice that the star network is the only tree of diameter 2. Thus, by point *(ii)*, if $L(g(a^T)) = 2$ then $g(a^T)$ is not a tree. In words, a small-world effect (diameter 2) in stage T is possible only with inefficient networks (see section 4 for efficiency considerations).

Result *(iii)* displays that if the players have established relationships in the $T-2$ first periods, some old connections remain or new links form in at least one of the two last stages. In other words, if players create links in early periods, they must maintain some former links or create new links in at least one of the last two periods. For example, relationships between individuals in a connected population of agents never completely disappear with time. Consider a market represented by a network of firms. Links model competition between firms and the finite horizon of T periods indicates the lifespan of the product. A firm leaving the market is symbolized by an isolated vertex. By result *(iii)*, if some firms have competed in the market in some of the first $T-2$ years, then the market cannot become empty of firms in the final years.

In proposition 5, we provide necessary conditions that any RNE has to fulfill. In the next proposition we are going to state two sufficient structural properties for a given strategy profile to be a RNE.

[7]However, it is possible that a player obtains less than the minmax payoff in stage T.

Proposition 6 *Consider a strategy profile $(M_1, \ldots, M_n)$ with outcome $(a^1, \ldots, a^T)$ such that, for any $i \in I$ and any $t \leq T$, $\pi_i(a^t) \geq 0$, and*

$$\pi_i(a^T) \geq \max \left\{ d_i(a^{T-1}) - c \min_{a_{ij}^1, \ldots, a_{ij}^{T-1}} \sum_{j \neq i} a_{ij}^t ; \max_{k<T-1} d_i(a^k) - \sum_{t=k+1}^{T} \pi_i(a^t) \right\}. \quad (9)$$

If either,

(i) $g(a^t) \cap g(a^{t'}) = \emptyset$, *for any* $t, t' \in \{1, \ldots, T\}$ *or*

(ii) there is a permutation $p : \{1, \ldots, T-1\} \longrightarrow \{1, \ldots, T-1\}$ *such that* $g(a^T) \subseteq g(a^{p(1)}) \subseteq g(a^{p(2)}) \subseteq \cdots \subseteq g(a^{p(T-1)})$,

then $(M_1, \ldots, M_n)$ is a RNE of G^T.

Proof. Assume that for any $i \in I$ and any $t \leq T$, $\pi_i(a^t) \geq 0$. The proof has two parts. Firstly, we show that in both situations *(i)* and *(ii)*, a player has no incentive to deviate as in the proof of lemma 1.

(i) The assumptions that any action profile in the outcome is cost-efficient and that for any $t, t' \leq T$, $g(a^t) \cap g(a^{t'}) = \emptyset$ implies that for any $i \in I$, $N_i(g(a_i^t, a_{-i}^{t'})) = N_i(g(a_i^{t'}, a_{-i}^t)) = \{\emptyset\}$.[8] As player i may still intend to create some links, this implies that $\pi_i(a_i^t, a_{-i}^T) \leq 0$ and $\pi_i(a_i^{t'}, a_{-i}^T) \leq 0$. By assumption, we then have $\pi_i(a_i^t, a_{-i}^T) \leq \pi_i(a^T)$ and $\pi_i(a_i^{t'}, a_{-i}^T) \leq \pi_i(a^T)$. This means that there is no period $t < T$ such that player i benefits from using the automaton constructed in the proof of lemma 1.

(ii) The assumption that for any $t \leq T$, $g(a^T) \subseteq g(a^t)$ implies that $g(a^T) \cap g(a^t) = g(a^T)$. Thus, for any player $i \in I$, $g(a^T) = g(a_i^t, a_{-i}^T) \subseteq g(a^t)$. This relation can be rewritten as $N_i(g(a^T)) = N_i(g(a_i^t, a_{-i}^T)) \subseteq N_i(g(a^t))$. We also know that the cost to player i in a^t is not less than in a^T. Therefore, for any $i \in I$ and for any $t \leq T$, $\pi_i(a^T) \geq \pi_i(a_i^t, a_{-i}^T)$. This implies once again that there is no period $t < T$ such that player i benefits from using the automaton constructed in the proof of lemma 1.

Secondly, we prove that condition (9) guarantees that $(M_1, \ldots, M_n)$ is a RNE. Each machine M_i in $(M_1, \ldots, M_n)$ is assumed to include a punishment state playing $a_{i,0}$ to threaten any deviation by an opponent as the one in machine M^1 in the proof of proposition 3. It is not the interest of player i to deviate in stage T in both situations *(i)* and *(ii)*. Thus, two cases similar to the one in the proof of proposition 3 must be considered.

[8] As $m_i < T$, only $T-1$ totally different networks may form. We thus assume that the empty network g_0 forms in at least two periods. This does not contradict condition *(i)* since trivially $g_0 \cap g_0 = \{\emptyset\}$.

Case 1.
Player i can deviate by using a $(T-1)$-state machine that plays b_i in stage $T-1$ and $a_i = \arg\min_{t<T} c\sum_{j\neq i} a_{ij}^t$ in the last period (remember that player i cannot use a new state playing $a_{i,0}$ is stage T since $m_i < T$). Player i will not choose such an automaton if and only if

$$\begin{aligned} & \pi_i(a^{T-1}) + \pi_i(a^T) \geq \pi_i(b_i^{T-1}, a_{-i}^{T-1}) - c\min_{t<T}\sum_{j\neq i} a_{ij}^t \\ \iff \quad & \pi_i(a^T) \geq d_i(a^{T-1}) - c\min_{t<T}\sum_{j\neq i} a_{ij}^t. \end{aligned} \tag{10}$$

Case 2.
Player i can deviate in period $k < T-1$ by playing b_i and then using a (possibly) new state playing $a_{i,0}$ until the end of the game. It is not the interest of player i to behave like this if and only if

$$\sum_{t=k}^{T} \pi_i(a^t) \geq \pi_i(b_i^k, a_{-i}^k) \iff \pi_i(a^T) \geq d_i(a^k) - \sum_{t=k+1}^{T} \pi_i(a^t). \tag{11}$$

Condition (9) in proposition 6 results from the combination of (10) and (11). By condition (9) we know that these deviations are not profitable to player i. This concludes the proof. ■

In words, if the payoff that each player obtains in the last round is large enough, then two kinds of interesting structures are likely to emerge. These structures are antagonistic. In the first one, at each round a new static network is built. These networks have no common link with any network formed in previous periods. As a consequence, if any two players are direct neighbors in a given period, this is the first period for which they are direct neighbors and they will never be directly connected once again. In a sense, we can refer to such an equilibrium as one with forgotten neighbors. This may highlight that players prefer a variety of one-period direct neighbors than long term direct relationships. In economic relationships, such a pattern of behavior is quite common. In a trading market, some buyers prefer visiting a variety of sellers than establishing a long term relationship with a particular seller (the so-called searchers).

In the second one, networks share an identical subnetwork and the structure allows for a total order relationship among networks that form within a RNE. The sequence of static networks corresponding to the equilibrium may reveal a contraction phenomenon. At each new period, the network that forms may be more and more restricted. This is the case of many economic situations. Consider for example several firms competing in a new market. A link between two firms can represent the investment supported by both firms to differentiate their product from that of the other firm. As time goes

by, least competitive firms are not strong enough to face competition. They either stop investing to differentiate their product (cuts links but preserve some) or leave the market (remain isolated). The competition network may retract progressively. Another interpretation is that more and more efficient networks may form within a RNE (see next section). The network should be k-connected in the initial period. The contraction process may lead to the formation of a tree network in the last stage. The total order may also exhibit an expansion phenomenon from the initial period to period $T-1$ and then a network contained in all others forms in the last round. If all networks in the sequence are connected, one can interpret such a phenomenon as $T-1$ periods of test before stage T in which agents choose the most valuable configuration.

In proposition 6 each player's payoff is assumed to be positive at each round. Despite the drastic restrictions on the intertemporal consistency between networks formed in any RNE, the next result shows that it is possible that any static network based on a cost-efficient action profile forms within a RNE. In the next proposition, we give a sufficient condition on the number of players for which any static network can appear in the outcome of a RNE.

Proposition 7 *Fix $n \geq 11$. There exists a RNE $(M_1, \ldots, M_n)$ whose outcome contains any network $g(a^*) \subseteq g_I$, $a^* \in A^{ce}$, at least once.*

Proof. Fix $n \geq 11$. We prove that any network $g(a^*) \subseteq g_I$, $a^* \in A^{ce}$, can form in the first period of a RNE. Denote by $g(l')$ the network that consists of the line network $g(l)$ (see the proof of proposition 4) except that player h is isolated. Let $g(l'')$ be the network isomorphic to $g(l')$ with player j being the isolated player.

We proceed in two steps. In a first step, we exhibit an n-tuple of automata $(M_1, \ldots, M_n)$ such that any network $g(a^*) \notin \{g(l'), g(l'')\}$ forms in the initial period and the (n-player) maxmin line network $g(l)$ forms in all remaining periods. Any deviation from M_i will release a definitive minmax punishment. The choice to form $g(l)$ in each period $t > 1$ has been already justified in the proof of proposition 4. In such a network the player who obtains the worst payoff is better than in any other network. Therefore, the condition on the number of players that prevents deviations is smaller if $g(l)$ forms from period 2 than if any other network forms.

However, by lemma 1, $g(l')$ and $g(l'')$ cannot form within the outcome of a RNE in which $g(l)$ forms in the mast stage. In fact, $\pi_i(l_i', l_{-i}) > \pi_i(l)$ for the neighbor in $g(l)$ of the isolated player h in $g(l')$. this is also true of $g(l'')$. As a consequence, in a second step, we deal with the cases $g(a^*) = g(l')$ and $g(a^*) = g(l'')$. We show that $g(l')$ and $g(l'')$ can be sustained as a RNN.

In a first step, consider any network $g(a^*) \notin \{g(l'), g(l'')\}$ and suppose that each player i chooses the following automaton:

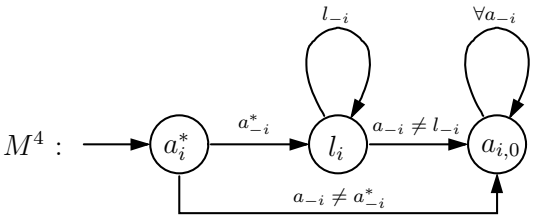


Figure 5.

As it is described in the first paragraph of the proof, the strategy profile $(M_1^4, \ldots, M_n^4)$ outputs network $g(a^*) \notin \{g(l'), g(l'')\}$ in period 1 and thereafter $T-1$ line networks $g(l)$. Player i obtains the average payoff

$$\tilde{\pi}_i(M_i^4, M_{-i}^4) = \frac{\pi_i(a^*) + (T-1)\pi_i(l)}{T}.$$

Any deviation induces a definitive minmax punishment. Thus, deviating in the first period yields player i the average payoff $\pi_i(b_i, a_{-i}^*)/T$. It is not the interest of player i to deviate in stage 1 if and only if

$$\frac{\pi_i(a^*) + (T-1)\pi_i(l)}{T} \geq \frac{\pi_i(b_i, a_{-i}^*)}{T} \iff (T-1)\pi_i(l) \geq d_i(a^*) \qquad (12)$$

Observe that $d_i(a)$ is maximal for a network $a \in A^{ce}$ in which player i seeks contact with each of the $n-1$ opponents whereas only one contact is needed to connect the network. Such a network being cost-efficient, player i keeps the network connected by removing $n-2$ superfluous links and saves $(n-2)c$. That is, for any action profile $a^* \in A^{ce}$,

$$\max_{a \in A^{ce}} d_i(a) = (n-2)c \geq d_i(a^*)$$

From the proof of proposition 4, we also know that the minimal payoff obtained by a player in $g(l)$ is $(n-1)v - 2c$. This implies that

$$\pi_i(l) \geq (n-1)v - 2c,$$

for any $i \in I$. Thus, condition (12) holds if

$$(T-1)((n-1)v - 2c) \geq (n-2)c \iff c \leq \frac{(n-1)(T-1)}{n+2T-4}v.$$

As $2v > c$ and $T > 3$, the reader can check that $n \geq 11$ is enough to guarantee that player i has no incentive to deviate. Next, consider a deviation in stage $t > 1$. The outcome $(a^*, l, \ldots, l)$ satisfies the necessary condition of lemma 1, that is, for any $i \in I$, there is no stage $k < T$ such that $\pi_i(a_i^k, l_{-i}^T) > \pi_i(l^T)$. In other words, players cannot benefit from deviating in stage T. Since we have also proved that $n \geq 11$ garantees that players

will not deviate in stage 1, it remains to test deviations in stages $2, \ldots, T-1$. The line network is formed in each of these stages. Thus, the condition that prevents deviations in these periods is identical to the one in proposition 3 for a RNN based on $g(l)$. To see this, recall that if a player has an incentive to deviate, then he does it late in the game because of the definitive punishment. Condition (3) in proposition 3 is satisfied for the RNN based on a line network when $n \geq 11$. Therefore, any network $g(a^*) \notin \{g(l'), g(l'')\}$ is likely to form in a RNE when $n \geq 11$.

In a second step, assume that each player i chooses the following two-state automaton:

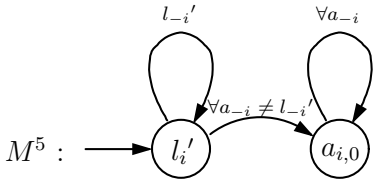


Figure 6.

The strategy profile $(M_1^5, \ldots, M_n^5)$ leads to the formation of a repeated network based on $g(l')$. Once again, any deviation releases a definitive minmax punishment. This situation is a particular case of proposition 3. Therefore, it is enough to check that condition (3) is satisfied to show that $(M_1^5, \ldots, M_n^5)$ is a RNN. As $n \geq 11$, remember that only the direct neighbors of the first and last players in the line component have an incentive to deviate from l'. Let i be one of these two players. We need

$$\begin{aligned}
& d_i(l') \leq \min\left\{\frac{v}{2}\#N_i(g(b_i, l_{-i}')), 2\pi_i(l')\right\} \\
\iff \quad & c - v \leq \min\left\{(n-2)\frac{v}{2}, 2((n-1)v - 2c)\right\} \\
\iff \quad & c - v \leq (n-2)\frac{v}{2} \\
\iff \quad & n \geq \frac{2c}{v}, \qquad (13)
\end{aligned}$$

which always holds when $n \geq 11$. We conclude that network $g(l')$ can also form in an equilibrium of G^T. The networks $g(l'')$ and $g(l')$ being isomorphic, the same conclusion applies to $g(l'')$. ■

There is a contrast between this result and previous results of this section. On one hand, any static network can form if one considers just a particular period of the repeated game. On the other hand, the T-period outcomes of

RNE must satisfy restrictive conditions. In the initial period, players can set links, which form any network (proposition 7). But the formation of this first network then prevents players from creating some other networks in future periods (lemma 1 and proposition 5). For instance, if the network that forms in stage 1 is connected, then by proposition 5 this network cannot be a subnetwork of the network that forms in stage T. Thus, the initial network conditions the structure of the whole outcome of a RNE.

4 Efficiency

In this section, we characterize the set of efficient action profiles for both Bentham and Pareto criteria (see the appendix for the proofs). Next, we examine this question in the finitely repeated game and compare the results with those of the static case. Again, we consider two cases (players being perfectly rational or not) and we study the differences in the corresponding sets of efficient strategy profiles. Precisely, we give a condition on the duration of the game for which boundedly rational players can implement strictly less efficient strategy profiles than perfectly rational players. We denote by $W(a) = \sum_{i \in I} \pi_i(a)$ the welfare resulting from a.

Definition 5 *An action profile $a^* \in A$ is*

- *Bentham-efficient if $a^* \in \arg\max_{a \in A} W(a)$;*
- *Pareto-efficient if for any $i \in I$ and any $a \neq a^*$,*

$$[\pi_i(a) > \pi_i(a^*)] \Longrightarrow [\exists j \neq i | \pi_j(a) < \pi_j(a^*)].$$

We say that an action profile is efficient if it is efficient either in the sense of Bentham and/or in the sense of Pareto.

Proposition 8 *Suppose $n > 4$. An action profile $a^* \in A$ is efficient if and only if (i) it is cost-efficient and (ii) $g(a^*)$ is a tree.*

Thus in our model, the sets of Bentham and Pareto efficient action profiles do coincide. This is quite noteworthy. Remark that the star network with $n > 4$ players is efficient for both Bentham and Pareto criteria although the player in the center of the star obtains less than the minmax payoff. This shows that a static efficient network is not always sustained as a RNN. The efficiency of strategy profiles in the repeated game is defined with respect to the average payoff. Let A_E be the set of efficient action profiles and S_E be the set of efficient strategy profiles. We use the subscripts $_{BE}$ and $_{SE}$ for the Bentham and Pareto criteria. The set S_E is characterized in the following proposition.

Proposition 9 *A strategy profile s^* is efficient if and only if it induces an efficient action profile in each period. Formally,*

$$S_E = \{s \in S : a^t \in A_E, a^t \in h^{T+1}(s), t \leq T\}.$$

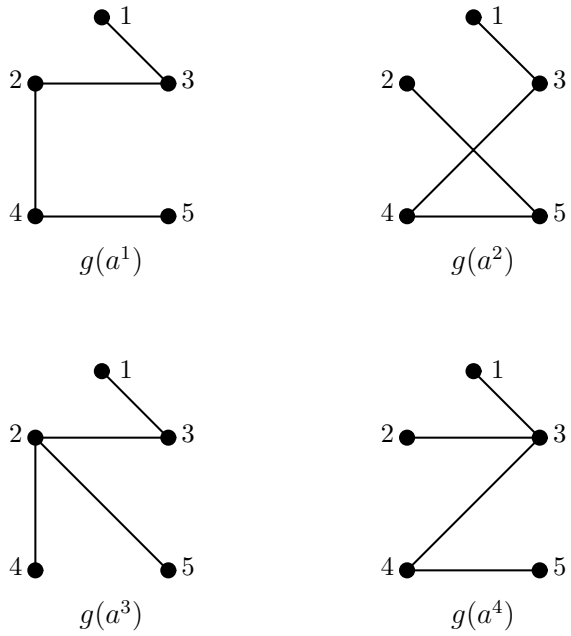
Proof. Firstly we prove that $S_{BE} = S_E$. A strategy profile s^* in which a Bentham-efficient action profile occurs in each period has a maximum average welfare $W(s^*)$ since the welfare is maximal is each period. Thus, the set of Bentham-efficient strategy profiles consists in all strategy profiles that induce T Bentham-efficient action profile, that is $S_{BE} = S_E$.

Secondly, we prove that any Pareto-efficient strategy profile induces a Pareto-efficient action profile in each period. We proceed by contradiction. Consider a Pareto-efficient strategy profile $s \in S_{PE}$ in which at least one non Pareto-efficient action profile a does occur. It is sufficient to suppose that there is only one such action profile, and that it occurs at round t. Thus, by definition, we know that there exists an action profile $p \in A$ such that, for any i, $\pi_i(p) \geq \pi_i(a)$, and $\pi_j(p) > \pi_j(a)$ for some j. Now consider the strategy profile s' that induces p at round t, and the same action profiles than s at the other rounds. Then we check easily than $\tilde{\pi}_i(s') \geq \tilde{\pi}_i(s)$, and $\tilde{\pi}_j(s') > \tilde{\pi}_j(s)$. This contradicts the Pareto-efficiency of s. By proposition 8, s' also induces a Bentham-efficient action profile in each period. We conclude by contradiction that $S_{PE} = S_{BE}$. ■

The results regarding the question of efficiency show that in our model Bentham-efficient and Pareto-efficient action profiles or strategy profiles have identical structures provided that the number of players is not too small. This correspondence between Bentham-efficiency and Pareto-efficiency is discussed in Jackson (2003) for the case of static network formation games.

A population of perfectly rational players can implement any efficient strategy profile in S_E. Now, let $\mathcal{M}_E$ denote the set of efficient strategy profiles of the machine game. A particular efficient strategy profile $s \in S_E$ may induce T different efficient networks. Thus, when players are assumed to use finite automata of limited size as in section 3, one can ask whether n automata with at most $T-1$ states may form T different efficient networks. In example 2, we show that combining actions played in the first $T-1$ stages may not allow for the creation of the Tth efficient network in the final period. We give a sequence of T line networks that cannot be implemented by automata with at most $T-1$ states. From this example, we will state a general proposition.

Example 2 Fix $N = \{1, 2, 3, 4, 5\}$, $T = 4$ and assume that players used strategies that can be implemented by automata with 2 or 3 states. Consider the sequences of four trees represented below and suppose that each action profile a^t, $t = 1, 2, 3, 4$, is cost-efficient.



The strategy profile that induces the sequence of networks (a^1, a^2, a^3, a^4) is efficient. Remark that player 2 has a totally distinct set of neighbors in each period. To implement the four actions a_2^t, $t = 1, 2, 3, 4$, player 2 must use an automaton with 4 states which does not belong to his strategy set. It follows that there does not exist an automata profile with 2 or 3 states that implement (a^1, a^2, a^3, a^4). □

Next, we generalize the idea of example 2. Each player i has $2^{n-1} - 1$ different nonempty set of neighbors in the set of all networks with n players. Player i has also the same number of different nonempty set of neighbors in the set of all trees with n players. This remark leads to the following result.

Proposition 10 *Suppose* $2^{n-1} - 1 \geq T$. *Then* $\mathcal{M}_e \subsetneq S_e$.

Proof. Suppose $2^{n-1} - 1 \geq T$. We construct a sequence of cost-efficient action profiles $(a^1, \ldots, a^T)$ that induce T trees and such that for any two periods t and t', $n_i(g(a^t)) \neq n_i(g(a^{t'}))$. There exists such a sequence since $2^{n-1} - 1 \geq T$. It follows that $a_i^t \neq a_i^{t'}$ such that player i's automaton has to implement T different actions. This can be done only if player i's machine has at least T states. Such a strategy does not belong to $\mathcal{M}_i^{1T}$. Thus, there is no automata profile of size $1 < m_i < T$ for each $i \in N$ that induce $(a^1, \ldots, a^T)$. We conclude that $\mathcal{M}_e \subsetneq S_e$. ■

This result contrasts with those found on RNE. On one hand, the set of RNE is more important when players are boundedly rational (propositions 2 and 3). On the other hand, boundedly rational players can implement less efficient strategy profiles than perfectly rational players (proposition 10).

5 Conclusion

Within a finite-horizon repeated game framework we study the problem of (dynamic) network formation when the players are either perfectly rational, or boundedly rational in the sense of Neyman (1985), and by a restriction to a subset of action profiles. We prove that the set of RNE is reduced to the empty network when the agents are perfectly rational, while this set is much more elaborate when the complexity of their strategies is limited. Then we identify structural properties of RNNs and RNE. In the case of RNNs, we prove that each network that is sustained as a RNN cannot be too much over-connected, and that each player cannot bear a too important share of the cost needed to connect the network. This highlights a lack of robustness of architectures such as star networks. In the case of RNE, we prove that the networks induced in any period satisfy some structural properties. Within a RNE, players may prefer to set links with totally different partners at each round, or the networks may retract progressively. Bounded rationality has a noticeable influence both on the existence of (non trivial) equilibria and on the dynamics of network formation. Assuming a limited ability to implement link formation seems reasonable since it is consistent with well-known economic behaviors (searchers in a trading market for instance). Finally, we make some comparisons between the sets of (Bentham and Pareto) efficient strategy profiles. In this part, the nature of results is reversed. Under some condition between the duration of the game and the number of players, it is shown that more rational players will implement a larger number of efficient strategy profiles. One of the main assumptions of the present work is that consent is needed to form links. A possible extension to this paper would be to see what happens if this assumption is relaxed. In particular, the resulting networks may be directed with the consequence that information is only one-way flow. This is left for future research.

Appendix

Proof. (proposition 1) Firstly, we show that in any network $g(a) \neq g_0$, at least one player has an incentive to deviate. Secondly, we prove that $g_0 = g(a_0)$ is a Nash network.

Consider any network $g(a) \neq g_0$. Two cases must be studied:

1. Suppose that a is not cost-efficient, then $\exists i, j \in I$ such that $a_{ij} = 1 \neq a_{ji} = 0$. It is the interest of player i to choose action $a_i(j^-)$ that only

differs from a_i by $a_{ij} = 0$. Player i saves c while $g(a_i(j^-), a_{-i}) = g(a)$. He obtains $\pi_i(a_i(j^-), a_i) = \pi_i(a) + c$, which implies that $g(a)$ is not a Nash network.

2. Suppose that a is cost-efficient. If $g(a)$ is a cyclic network, then there are at least two players i and j such that $N_i(g(a) - ij) = N_i(g(a))$. Link ij is superfluous. We have $d_i^1(a) > 0$, which implies that $g(a)$ is not a Nash network. If $g(a)$ is an acyclic network, then there is a player i whose set of connections satisfies $\#N_i(g(a)) = 1$. Let $N_i(g(a)) = \{j\}$, then link ij is superfluous for player j. Therefore $d_j^2(a) > 0$, which implies $g(a)$ is not a Nash network.

The empty network $g_0 = g(a_0)$ is the unique network that does not fit with any of the two previous cases. In $g(a_0)$, there is no player i who has an incentive to deviate from $a_{i,0}$ since he cannot create links alone and would support a cost c for any such attempt. Then the cost-efficient empty network is the only Nash network of G. ■

Proof. (proposition 2) Consider network $g(a) \neq g_0$ and suppose $g(a^T) = g(a)$. By proposition 1 and a backward induction argument, there is at least one player say i who has interest in altering his action in the last period. Player i's opponents anticipate this behavior and also remove some links. As a consequence, the empty network necessarily forms in the last stage. A similar process leads to the formation of the empty network in all periods. ■

Proof. (proposition 8) The proof is divided in two parts, one for each criterion.

Bentham-efficiency

($\Longrightarrow$) Suppose $b \in A$ is Bentham-efficient. Firstly, b must be cost-efficient. Otherwise there are players $i, j \in I$ with $b_{ij} = 1 \neq b_{ji} = 0$ such that player i saves c if he plays $b_i(j^-) = 0$. Player i's altered action does not remove any link and maintains other players' payoffs. Remark that $g(b)$ must be acyclic, otherwise there is a player i in $g(b)$ who has some superfluous links or equivalently $d_i^1(b) > 0$. If i removes one such link, say with j, he obtains $\pi_i(b_i(j^-), b_{-i}) = \pi_i(b) + c$ and $\pi_h(b_i(j^-), b_{-i}) = \pi_h(b)$, $\forall h \neq i, j$ since $b_i(j^-)$ keeps unchanged all other players' connections (and payoffs).

Secondly, we have to show that $g(b)$ must be a tree. To see this, we prove that the welfare of a tree is larger than in any other type of network. In a tree, the n vertices must be connected by exactly $n-1$ links. The formation of a link costs c to two players. Thus, the total cost of a tree $g(b)$ is $(n-1)2c$. The value of $g(b)$ for each player is $(n-1)v$. The welfare of any tree $g(b)$

induced by a cost-efficient action profile b is [9]

$$W(b) = n(n-1)v - (n-1)2c = (n-1)(nv - 2c). \tag{14}$$

Now consider any non connected network $g(a)$. By definition, $g(a)$ is splitted in $K > 1$ components, which are connected subnetworks. If $g(a)$ is candidate to be Bentham-efficient, then each subnetwork $g(a_k)$, $k \in \{1, \ldots, K\}$, satisfies condition 1 of the proposition, that is, $g(a)$ is a forest of K trees induced by cost-efficient action profiles. The component $g(a_k)$ has $\#I_k = n_k$ vertices (or players). The welfare of $g(a_k)$ is

$$W(a_k) = (n_k - 1)(n_k v - 2c),$$

and the total welfare of a is

$$W(a) = \sum_{k=1}^{K} (n_k - 1)(n_k v - 2c).$$

By definition of I_k,

$$nv - 2c > n_k v - 2c \Longrightarrow \sum_{k=1}^{K} (n_k - 1)(nv - 2c) > \sum_{k=1}^{K} (n_k - 1)(n_k v - 2c).$$

Furthermore, $\sum_{k=1}^{K} n_k - 1 = n - 1$ implies that

$$(n-1)(nv - 2c) > \sum_{k=1}^{K} (n_k - 1)(nv - 2c),$$

and we conclude that $W(b) > W(a)$.

($\Longleftarrow$) Suppose b satisfies the two conditions listed in the statement of proposition 8. By the previous calculation, the welfare in b is larger than in any other action profile.

Pareto-efficiency

($\Longleftarrow$) By the first part of the proof, any tree induced by a cost-efficient action profile is Bentham-efficient. And it follows from definition 5 that any Bentham-efficient network is Pareto-efficient. Therefore, any cost-efficient action profile that induces a tree is a Pareto-efficient.

[9] Such a network ranges from the n-player line network to the n-player star network. These two networks exhibit extreme situations according to their diameter. A line network has a diameter of $n - 1$, the largest diameter among trees, while a star network has a diameter of 2, the smallest diameter among trees.

($\Longrightarrow$) Firstly, any Pareto-efficient action profile must be cost-efficient for the same reason than in the first part of the proof. Secondly, to show that any Pareto-efficient action profile induces a tree, we argue by contradiction. Consider any Pareto-efficient action profile p for which $g(p)$ is not a tree. Note that $g(p)$ cannot be k-connected, $k > 1$, since the existence of some superfluous links would contradict the fact that p be Pareto-efficient. Then, suppose that the Pareto-efficient action profile p induces a non connected network. We group all possibilities in three cases according to the structure of $g(p)$:

Case 1.
Network $g(p)$ consists in at least 2 connected components with at least 2 players. Each component must be a tree to avoid cycles (and superfluous links). Let K be the total number of components in the forest $g(p)$. Construct the network $g(p')$ that consists in connecting the K components all together with the creation of $K-1$ links (see for instance the first part of the proof). Since $2v > c$, it is easy to check that the resulting network yields all players a larger payoff than $g(p)$. This proves that $g(p)$ cannot be Pareto-efficient.

Case 2.
Network $g(p)$ consists in a connected component and $k > 1$ isolated players. The component must be a tree to avoid cycles. The method described in case 1 is still beneficial to all players whenever at least 2 isolated players can be linked to the main tree component.

Case 3.
Network $g(p)$ consists in a connected component and a single isolated player. The component must also be a tree to avoid cycles. If the connected component includes all but one player denoted h, it follows that the creation of link ih with any player i in the component is beneficial to all players except i, who loses $v - c$. Nonetheless, it is possible to connect player h to all other opponents in a way that increases everyone's payoff. Let i be a player such that $\#n_i(g(p)) = 1$. Player i must exist since the connected component is a tree. Precisely, let $n_i(g(p)) = \{j\}$. Construct the network $g(p') = g(p) - ij + hj + hi$. The reader can check that all players except h have the same number of direct neighbors in $g(p')$ than in $g(p)$ and benefit from the additional connection with player h. Player h creates two links but is connected to at least 4 players. Thus, his payoff is larger than that of a isolated player. Therefore all players obtain a larger payoff, which implies that p cannot be Pareto-efficient.

Thus, we conclude by contradiction that any Pareto-efficient action profile induces a tree. ■

References

[1] Abreu Dilip, Rubinstein Ariel (1988), "The Structure of Nash Equilibrium in Repeated Games with Finite Automata", *Econometrica* **56**:1259-1282.

[2] Bala Venkatesh, Goyal Sanjeev (2000), "A Non Cooperative Model of Network Formation", *Econometrica* **68**:1181-1229.

[3] Bolton Patrick, Dewatripont Mathias (1994), "The Firm as a Communication Network", *Quarterly Journal of Economics* **109**:809-839.

[4] Calvó-Armengol Antoni (2003), "Stable and Efficient Bargaining Networks", *Review of Economic Design*, **71**:411-428.

[5] Calvó-Armengol Antoni, Jackson Matthew O. (2004), "The Effects of Social Networks on Employment and Inequality", *American Economic Review*, **94**:426-454.

[6] Currarini Sergio, Morelli Massimo (2000), "Network Formation with Sequential Demands", *Review of Economic Design*, **5**:229-249.

[7] Dutta Bhaskar, Ghosal Sayantan, Ray Debraj (2005), "Farsighted Network Formation", *Journal of Economic Theory*, **122**:143-164.

[8] Ellison Glenn (1993), "Learning, Local Interaction and Coordination", *Econometrica*, **61**:1047-1071.

[9] Ellison Glenn, Fudenberg Drew (1995), "Word of Mouth Communication and Social Learning", *Quarterly Journal of Economics* **110**:93-125.

[10] Gilles Robert P., Sarangi Sudipta (2004), "Social Network Formation with Consent", *Working Paper*, Department of Economics, Virginia Tech and Department of Economics, Louisiana State University.

[11] Goyal Sanjeev, Vega-Redondo Fernando (2005), "Network Formation and Social Coordination", *Games and Economic Behavior* **50**:178-207.

[12] Jackson Matthew O. (2003), "The Stability and Efficiency of Economic and Social Networks", in *Advances in Economic Design*, edited by Murat Sertel and Semih Koray, Springer-Verlag, Heidelberg.

[13] Jackson Matthew O., Watts Alison (2002), "The Evolution of Social and Economic Networks", *Journal of Economic Theory* **106**:265-295.

[14] Jackson Matthew O., Wolinsky Asher (1996), "A Strategic Model of Social and Economic Networks", *Journal of Economic theory* **71**:44-74.

[15] Milgram Stanley (1967), "The Small World Problem", *Psychology Today*, **2**:60-67.

[16] Neyman Abraham (1985), "Bounded Complexity Justifies the Cooperation in the Finitely Repeated Prisoner's Dilemma", *Economics Letters* **19**:227-229.

[17] Piccione Michele, Rubinstein Ariel (1993), "Finite Automata Play a Repeated Extensive Game", *Journal of Economic Theory*, **61**:160-168.

[18] Simon Herbert A. (1955), "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, **69**:99-118.

[19] Watts Alison (2001), "A Dynamic Model of Network Formation", *Games and Economic Behavior* **34**:331-341.

[20] Watts Alison (2002), "Non-Myopic Formation of Circle Networks", *Economic Letters* **74**:277-282.

Spreading on networks: a topographic view

Geoffrey S Canright and Kenth Engø-Monsen
Telenor R&D

Abstract. We apply our previously developed method of "topographic" analysis of networks to the problem of epidemic spreading. We consider the simplest form of epidemic spreading, namely the "SI" model. We argue that the *eigenvector centrality* of a node is a good indicator of that node's *spreading power*. From this we develop seven specific predictions. In particular, we predict that each *region* (as defined by our approach) will have its own S curve for cumulative adoption over time, and we describe the various phases of the S curve in terms of motion of the infection over the region. Our predictions are well supported by simulations. In particular, the significance of regions to epidemic spreading is clear. Finally, we develop a mathematical theory, giving partial support to our picture. The theory includes a precise quantitative definition of the spreading power of a node, and some approximate analytical results for epidemic spreading.

Short title: Spreading on networks

Keywords: epidemic spreading, eigenvector centrality, networks, regions, topography

1 Introduction

The general phenomenon of spreading over a network is ubiquitous. Examples include the spreading of a disease (in which case the network may be a social network, or a computing/communications network exposed to a virus); the spreading of gossip (social network; although here we will ignore the interesting property that gossip gets modified as it spreads); or the spreading of innovation (again a social network). In each of these cases, it is a common simplification to assign to each node only one of two possible states: 'uninfected' or 'infected'. If you are uninfected ('susceptible'), you are deemed liable to be infected by any infected neighbors. Correspondingly, if you are infected, you remain so for the duration of the experiment—and you remain capable of infecting any or all of your neighbors. Of course, on some appropriate time scale, nodes become 'immune' to the infection: a human develops antibodies, a machine gets antivirus software, the gossip becomes boring, or the innovation becomes outmoded. We focus on a shorter time scale here, so that we can ignore the state of acquired immunity. The process under study here may then be equally termed 'infection', 'adoption', or 'epidemic spreading'; we will use all of these terms freely. The technical name for our model of spreading is 'SI', since the nodes have the two states Susceptible or Infected. .

Since spreading takes place over the links of a network, it is clear that the topology of the network can have a profound influence on the spreading process. In particular, we believe that the best understanding of spreading will come from a perspective which is based on a view of the whole network, and an understanding of that network's structure. In earlier work [1], we have presented an approach to the analysis of network structure which is applicable to any network with symmetric (undirected) links. We also suggested that the analysis should be useful for the understanding of spreading over such a network. In this work, we elaborate on that suggestion, and then test it, in two ways. First, we report on a series of simulations, carried out on network structures which were obtained from empirically measured undirected networks. These simulations offer strong support to our qualitative picture of spreading. Secondly, we develop a mathematical definition of the spreading power of a node, as well as a mathematical description of the SI epidemic spreading process. While we cannot obtain exact answers to either of these formulations, we do find significant support for our qualitative arguments in the approximate results that we obtain.

Authors' address: Telenor R&D, B6d, Snarøyveien 30, 1331 Fornebu, Norway
Phone: +47-91815638 (GSC), +47-41560883 (KEM)
Fax: +47-96211086
Email: geoffrey.canright@telenor.com, kenth.engo-monsen@telenor.com

Our approach departs from previous work in that we focus on both the *time* and *spatial progression* of the epidemic spreading. We take a spatial resolution which is not microscopic, but rather at the level of 'neighborhoods'—connected subgraphs with roughly the same spreading power. More traditional approaches (reviewed in [6]) start from the 'well-mixed' approximation, that every node can infect every other with some probability, at all times. This approach may be said to have no network perspective; or, it may be said to postulate a graph with extremely good mixing—such as a random graph of high degree, or a complete graph. The review of Newman [6] also discusses more recent work, involving a network perspective. All such work is based on whole-graph properties, such as the node degree distribution; also, these approaches have focused on obtaining whole-graph results, either over time [7,8], or focusing especially on the infected fraction at very long times [9]. This latter question is of course only interesting for models more complex than the SI model; and indeed most work is directed towards the behavior of the SIS model (where nodes lose their infection after some time, and so become Susceptible again), or the SIR model (where nodes, after losing their infection, go through a refractory period).

Brauer [10] has examined the SI model for the case that the nodes (organisms, especially humans) are born and die. Because of the addition of these dynamic features, the steady infection rate is not necessarily 100%. This work uses the well-mixed approximation, which gives rise to coupled ordinary differential equations.

A work which is perhaps closest to the present work is that of Wang et al [11]. Their model is SIS, in that nodes can be "cured"; but it is based on a fully microscopic view of the network. In fact, their time evolution operator is the same as that we develop in Section 4.3.2, with two differences. One is their addition of the "curing" term. This term is simply a multiple of the unit matrix, and so does not change the dominant eigenvector—which remains that of the adjacency matrix A. Because their model is SIS, the long-time infection fraction is not obvious, and must be solved for. The second difference in the time evolution operator of Wang et al is that they neglect the cross terms—ie, those arising from multiple transmissions to an infected node. This approximation is valid for low infection fraction—while (as we discuss in detail below) it *may* also be good even as the infection fraction becomes large. Wang et al report simulations which offer some support for this statement.

We emphasize that our work, like that of Wang et al [11], uses the full adjacency matrix A in modelling the time evolution of the infection. Thus we start from a microscopic foundation. However, we will quickly appeal to a '*mesoscopic'* picture, in which it is meaningful and useful to speak of neighborhoods and their properties. As far as we know, our work is unique in this regard.

We conclude this introduction with a brief review of the network analysis presented in [1]. Our strategy there was to choose eigenvector centrality as a useful measure of well-connectedness. Eigenvector centrality (EVC) has the desirable property that—since it depends on the properties of the *neighborhood* of a node, and not just of the node itself—it is rather 'smooth' over the graph (or network; we use these terms interchangeably). This is in contrast to the related quantity *degree centrality*, which simply counts the links leaving a node and so is completely local. The smoothness of the EVC allows one to think in terms of the 'topography' of the graph. That is, if a node has high EVC, its neighborhood (from smoothness) will also have a somewhat high EVC—so that one can imagine EVC as a smoothly varying 'height', with mountains, valleys, mountaintops, etc. We define precisely the notion of a *region* of the graph: all those nodes for which a steepest-ascent path terminates at the same local maximum of the EVC belong to the same region. That is, a region is a mountain, and it is defined by its peak. This definition gives a unique decomposition of any graph into one or more regions, with each node (almost always) belonging to a single region. Links connecting regions are also viewed to play a special role in the network; in the topographic analogy, they correspond to 'saddles'.

In the next section we will give our reasons for believing that this method of analysis is useful for understanding spreading. We emphasize here that the term 'region' has, everywhere in this paper, a precise mathematical meaning; hence, to discuss subgraphs which lie together in some looser sense, we use terms such as 'neighborhood' or 'area'.

2 Qualitative arguments

We wish to evaluate the nodes in a network in terms of their "spreading power". That is, we know that some nodes play an important role in spreading, while others play a less important role. One need only imagine the extreme case of a star: the center of the star is absolutely crucial for spreading of infection over the star; while the leaf nodes are entirely unimportant, having only the one aspect (common to every node in any network) that they can be infected.

Clearly, the case of the star topology has an obvious answer to the question of which nodes have an important role in spreading (have high spreading power). The question is then, how to generate equally meaningful answers for general and complex topologies, for which the answer is not at all obvious? In this section we will propose and develop a qualitative answer to this question.

Our basic assumption (A) is simple, and may be expressed in a single sentence:

Eigenvector centrality (EVC) is a good measure of spreading power. ***(A)***

We will test this idea via simulations (Section 3) and some theory (Section 4). First however we will explore the implications of this idea—coupling it to our earlier work, which uses EVC to define well-connected regions of any given network.

First we recall the definition of, and motivation for, EVC. A simple definition of centrality, which is certainly related to spreading power, is a node's degree centrality, ie, its number of neighbors:

$$k_i = \sum_{j=nn(i)} , \tag{1}$$

where *nn(i)* means "nearest neighbors of node *i*". This quantity is however not smooth at all—there is no necessary correlation between a given node's degree and that of its neighbors.

Seeking a more smooth measure, we can give each node a centrality score which is simply the average of its neighbors' scores:

$$x_i = \frac{1}{k_i} \sum_{j=nn(i)} x_j \ . \tag{2}$$

As shown in [1], this choice gives a node importance weight which is *too* smooth: normal solutions to (2) give equal scores x_i to every node *i.* Hence we discard this idea.

Eigenvector centrality involves one small change from (2): one defines the centrality of node *i* as being proportional to the sum of (but not the average of) *i*'s neighbors' centralities:

$$e_i = \frac{1}{\lambda} \sum_{j=nn(i)} e_j \ . \tag{3}$$

Eq. (3), in contrast to (1), makes a node's centrality dependent on its neighbors' centralities—hence giving a 'smooth' centrality measure—but it also gives nontrivial solutions [unlike (2)]—because, since λ is the same for all nodes, Eq. (3) does not completely cancel out the boost in centrality from having many neighbors [as (2) does]. Eq. (3) can be rewritten as

$$Ae = \lambda e \ , \tag{4}$$

where *e* is the vector of centrality scores, and *A* is the network's adjacency matrix. Eq. (4) reveals the motivation for the name 'eigenvector centrality': the EVC of e_i of node *i* is the *i*'th component of a chosen eigenvector *e* of the adjacency matrix *A*. To ensure that all centrality scores are positive, one takes the *principal* eigenvector of *A*—that is, the one corresponding to the largest eigenvalue λ_{max}.

Thus we see that, because a node's EVC depends on that of its neighbors, the EVC values over a network may be thought of as 'smoothly varying' over the network. That is, a node with very high EVC cannot be surrounded by nodes with very low EVC. Of course, it is true that EVC tends to be positively correlated with a simpler measure of centrality, namely node degree. In fact, one might say that the principal difference between the two measures is that EVC is constrained by its definition to

be smooth, while node degree centrality is not. This difference can however be nontrivial. For instance, a node with high degree, surrounded by many leaf nodes, and linked only tenuously to the bulk of a large and well-connected network, will have a low EVC, in spite of its high degree. The point is that EVC is sensitive to properties of neighborhoods, while node degree is not.

Thus, in short, there are no isolated nodes with high EVC. That is, a node with high EVC is embedded in a neighborhood with high EVC. (There can however be relatively isolated nodes with *low* EVC, as this situation is self-consistent.) Now if we take our basic assumption A to be true, then there are no isolated nodes with high spreading power. Instead, there are neighborhoods with high spreading power. (But there can be isolated nodes with low spreading power.)

We then suppose that an infection has reached a node with modest spreading power. Suppose further that this node is not a local maximum of EVC; instead, it will have a neighbor or neighbors of even higher spreading power. The same comment applies to these neighbors, until one reaches the local maximum of EVC/spreading power.

Now, given that there are neighborhoods, we can discuss spreading in terms of neighborhoods rather than in terms of single nodes. It follows from the meaning of spreading power that a neighborhood characterized by high spreading power will have more rapid spreading than one characterized by low spreading power. Furthermore, we note that these different types of neighborhoods (high and low) are smoothly joined by areas of intermediate spreading power (and speed).

It follows from all this that, if an infection starts in a neighbourhood of low spreading power, it will tend to spread to a neighbourhood of higher spreading power. That is: spreading is faster *towards* neighborhoods of high spreading power, because spreading is faster *in* such neighborhoods. Then, upon reaching neighbourhood of the nearest local maximum of spreading power, the infection rate will also reach a maximum (with respect to time). Finally, as the high neighborhood saturates, the infection moves back ‘downhill’, spreading out in all ‘directions’ from the nearly saturated high neighborhood, and saturating low neighborhoods.

We note that this discussion fits naturally with our earlier work, which describes network topology in topographic terms. That is, the smoothness property of EVC allows one to think of a smoothly varying ‘height’ for nodes—so that one can define mountains, their peaks, valleys, etc. Putting the previous paragraph in this language, then, we get the following: infection of a hillside will tend to move uphill, while the infection rate grows with height. The top of the mountain, once reached, is rapidly infected; and the infected top then efficiently infects all of the remaining adjoining hillsides. Finally, and at a lower rate, the foot of the mountain is saturated.

We see now that this qualitative picture addresses nicely the various stages of the classic S curve of innovation diffusion [12]. The early, flat part of the S is the early infection of a low area; during this period, the infection moves uphill, but slowly. The S curve begins to take off as the infection reaches the higher part of the mountain. Then there is a period of rapid growth while the top of the mountain is saturated, along with the neighboring hillsides. Finally, the infection rate slows down again, as the remaining uninfected low-lying areas become infected.

One might object that this picture is too simple, in the following sense. Our picture gives an S curve *for a single mountain.* Yet we know from earlier analyses that a network is often composed of several regions (mountains). The question is then, why should such multi-region networks exhibit a single S curve?

Our answer here is that such networks need not necessarily exhibit a single S curve. That is, our arguments predict that each region—defined around a local maximum of the EVC—will have a single S curve. Then—assuming that each node belongs to a single region, as occurs with our preferred rule for region membership—the cumulative adoption/infection curve for the whole network is simply the sum of the adoption curves for each region. These latter single-region curves will be S curves. Thus, depending on the relative timing of these various single-region curves, the network as a whole may, or may not, exhibit a single S curve. For example, if the initial infection is from a peripheral node which is close only to one region, then that region may take off well before neighboring regions. On the other hand, if the initial infection is in a valley which adjoins several mountains, then they may all exhibit

takeoff roughly simultaneously—with the result being a sum of roughly synchronized S curves, hence a single S curve.

Let us now summarize and enumerate the predictions we take from this qualitative picture.

- **a.** Each region has an S curve.
- **b.** The number of takeoff/plateau occurrences in the cumulative curve for the whole network may be more than one; but it will not be more than the number of regions in the network.
- **c.** For each region—assuming (which will be typical) that the initial infection is not a very central node—growth will at first be slow.
- **d.** For each region (same assumption) initial growth will be towards higher EVC.
- **e.** For each region, when the infection reaches the neighborhood of high centrality, growth "takes off".
- **f.** An observable consequence of (e) is then that, for each region, the *most* central node will be infected at, or after, the takeoff—but not before.
- **g.** For each region, the final stage of growth (saturation) will be characterized by low centrality.

3 Simulations

We have run simple simulations to test our qualitative picture. As noted above, we have implemented an SI model. That is, each node is Susceptible or Infected. Once infected, it remains so, and retains the ability to infect its neighbors indefinitely. We have used a variety of sample networks, extracted from data obtained in previous studies. Results reported here will be taken from simulations performed on: (i) seven distinct snapshots [2] of the Gnutella peer-to-peer file-sharing network, taken in late 2001; (ii) a social network of students at the Høgskolen i Oslo, measured in 2004 in the course of a study [3] on the spreading of innovation; (iii) a snapshot of the collaboration graph of our own R&D department; (iv) a snapshot of the collaboration graph for the Santa Fe Institute [4]. We will denote these graphs as (i) g1—g7; (ii) hio; (iii) fou (Norwegian for R&D); and (iv) sfi.

Our procedure is as follows. We initially infect a single node. Each link *ij* in the graph is assumed to be symmetric (consistent with our use of EVC), and to have a constant probability *p,* per unit time, of transmitting the infection from node *i* to node *j* (or from *j* to *i*), whenever the former is infected and the latter is not. (We continue to use the terms 'infection' and 'adoption' interchangeably.) All simulations were run to the point of 100% saturation. Thus, the ultimate *x* and *y* coordinates of each cumulative infection curve give, respectively, the time needed to 100% infection, and the number of nodes in the graph.

3.1 Gnutella graphs

The Gnutella graphs, like many self-organized graphs, have a power-law node degree distribution, and are thus well connected. Consistent with this, our analysis returned either one or two regions for each of the seven snapshots. We discuss these two cases (one or two regions) in turn. Each snapshot is taken from [2], and has a number of nodes on the order of 900—1000.

3.1.1 Single region

The snapshot termed 'g3' was found by our analysis to consist of a single region.

Figure 1 shows a typical result for the graph g3, with link infection probability $p = 0.05$. The upper part of the figure shows the cumulative S curve of adoption, with the most central node becoming infected near the 'knee' of the S curve. The lower part of Figure 1 shows a quantity termed μ —that is, the average EVC value for newly infected nodes at each time. Our qualitative arguments say that this quantity should first grow, and subsequently fall off, and that the main peak of μ should coincide with the greatest growth in the S curve. We see, from comparing the two parts of Figure 1, that the most central nodes are infected roughly between time 2 and time 20—coinciding with the period of

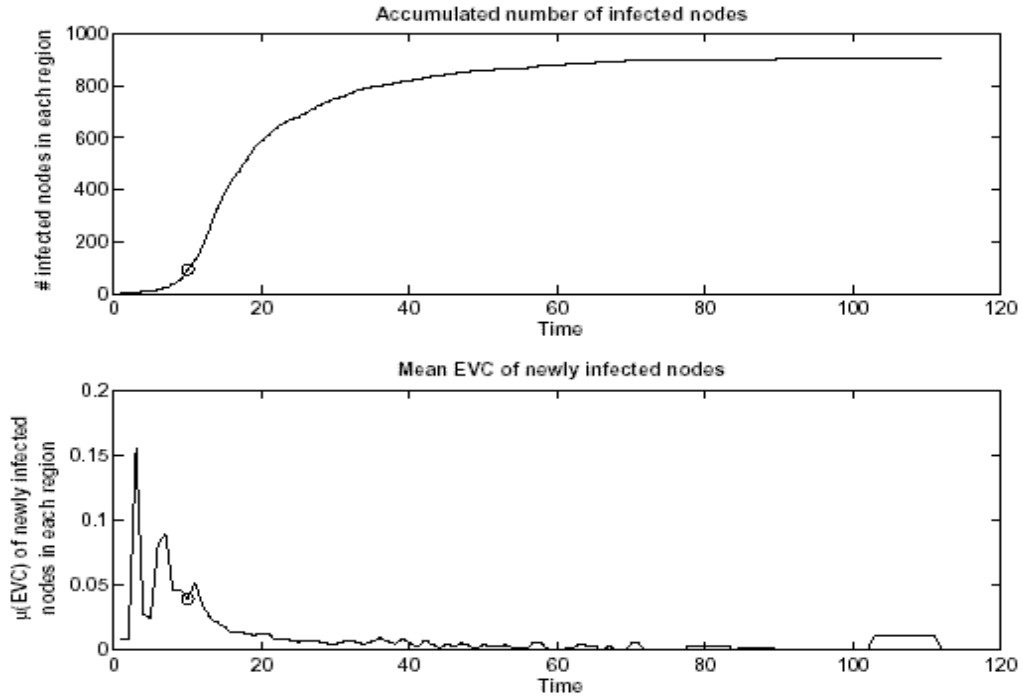


Figure 1. (top) Cumulative adoption for the Gnutella graph g3. The circle marks the time at which the most central node is infected. (bottom) Average EVC ('height') of newly infected nodes at each time step.

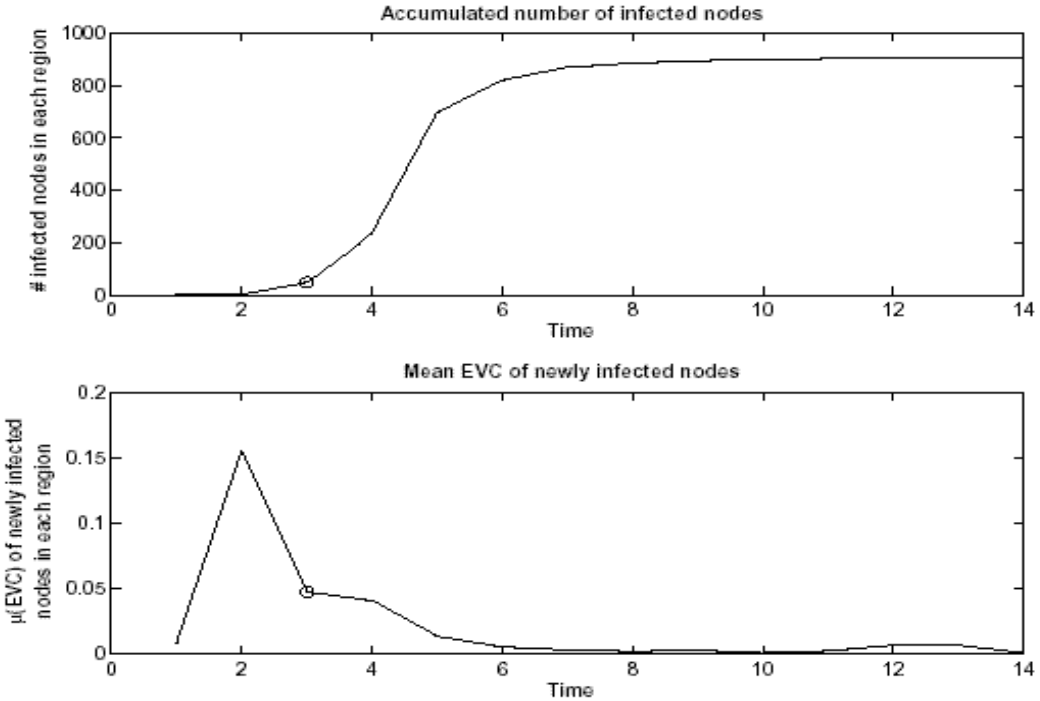


Figure 2. Same simulation as Figure 1, except for: (i) $p = 0.6$, and (ii) new random numbers determining the infection events.

maximum growth in the S curve. Thus, this figure supports all of our predictions a—g above, with the minor exception that there are some fluctuations superimposed on the growth and subsequent fall of μ over time.

It is interesting to note that this picture is rather insensitive to the probability parameter *p*. For example, Figure 2 shows results for the same graph and same initial node, with the probability now $p = 0.6$. We see that the main effect of this much higher *p* is simple, and offers few surprises: the time scale is of course much compressed, with the expected result that the cumulative curve is less smooth. We note that even the extreme case of setting $p = 1$ gives a picture very much like that of Figure 2.

3.1.2 Two regions

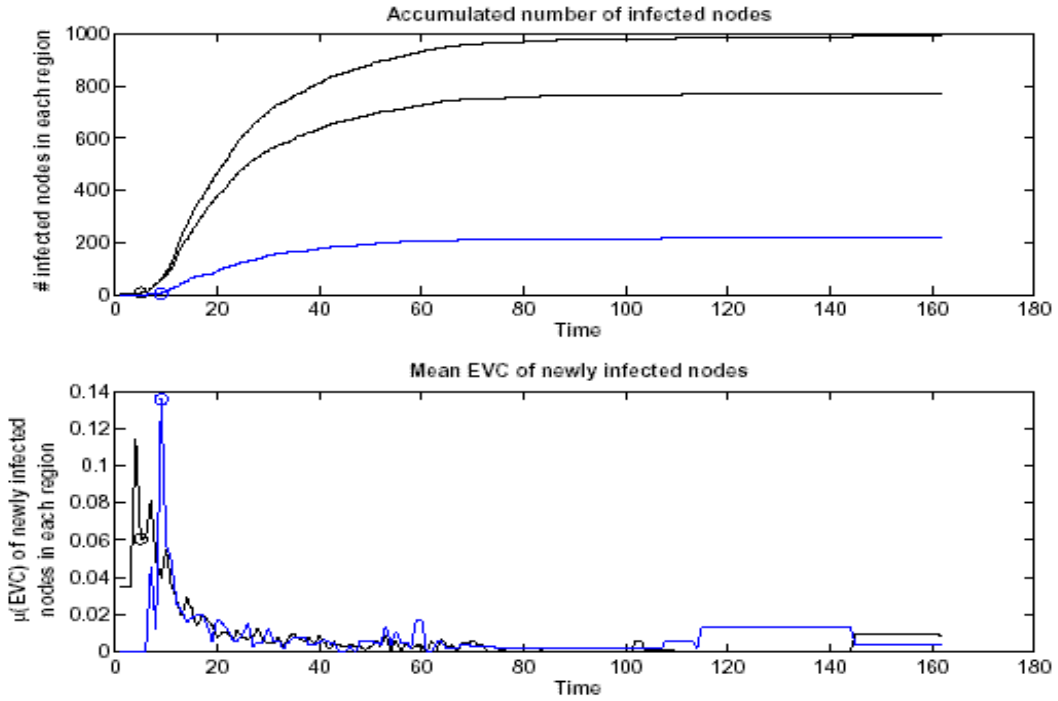


Figure 3. (top) Cumulative infection curves, and (bottom) μ curves, for a two-region graph g1, with $p = 0.04$. The upper plot shows results for each region (black and blue), plus their sum (also black). The lower plot gives μ curves for each region.

Figure 3 shows typical behavior for the graph g1, which consists of two regions. We see that the two regions go through takeoff roughly simultaneously. The result is that the sum of the regional infection curves is a single S curve. We also see that each region behaves essentially the same as did the single-region graph g3. For instance, the new centrality (μ) curves for each region first rise, and then fall, with their main peak (before time 20) coinciding with the period of most rapid growth for the respective regions (and for the whole graph). These results thus add further support to our predictions a—g.

3.2 Student network, HiO

For this graph we found a largest connected component (LCC) consisting of 249 nodes [3]. We call this LCC the 'hio graph'. Our analysis found three regions for the hio graph. However, the spreading behavior on this graph is qualitatively like that for the g1 graph: the three regions tend to take off roughly simultaneously. The effect is also the same: the cumulative infection curve for the entire graph is a clean S curve, composed of the sum of three smaller clean S curves. Also, each region behaves essentially as seen for the graph g3.

3.3 Telenor R&D

We have formed a collaboration graph for the researchers working at Telenor R&D. For this graph, we analyzed the largest connected component, consisting of 137 nodes. Our analysis gave a single region for this graph. Spreading behavior was much like that seen in Figure 1. One difference is that the S curve is less smooth—an effect of the small size of the fou graph. Another difference is that, for many simulations (but not all), the time period of the rise of the S is disproportionately large compared to the time needed for 100% of saturation. Also, in such cases, the time of infection of the most central node tends to fall rather late after the onset of the rise (the 'knee' of the S). We show an example of this behavior in Figure 4.

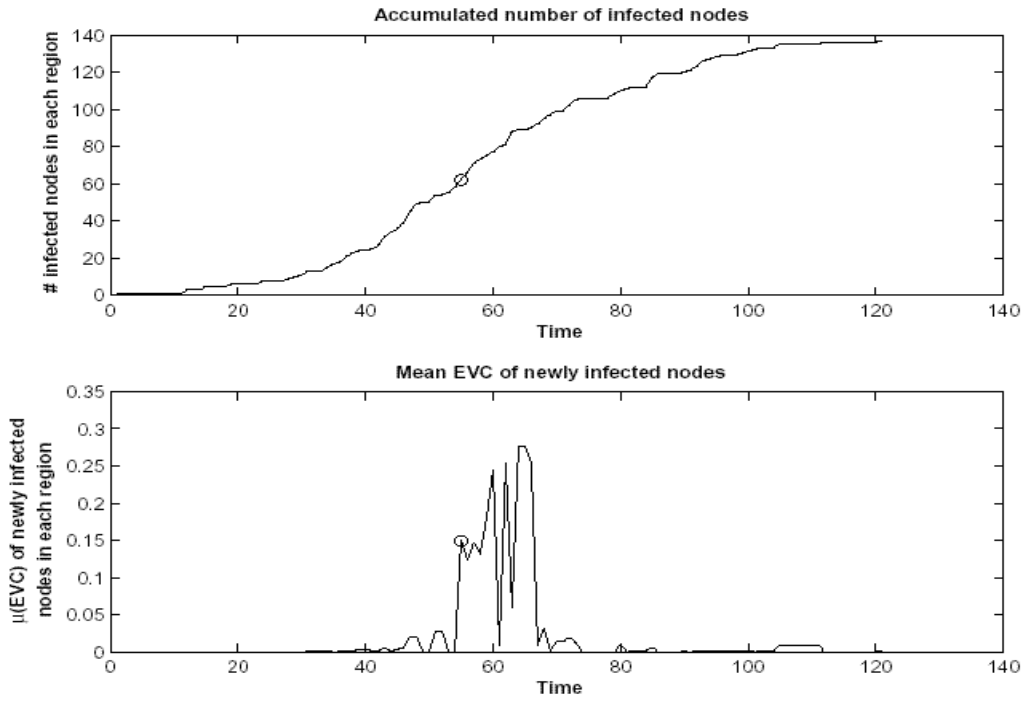


Figure 4. Spreading for the single-region fou graph, with $p = 0.04$. Note that the S curve is somewhat 'flat' on a time scale of 100% saturation.

The differences between Figures 1 and 4 are interesting; however both figures are fully consistent with our predictions a—g. We note also an interesting correlation here, which is not surprising: steeper S curves tend to be associated with earlier infection times (relative to the knee) for the most central node.

3.4 Santa Fe Institute

The sfi graph also gives three regions under our analysis. The spreading behavior varied considerably on this graph, depending both on the initially infected node, and on the stochastic outcomes for repeated trials with the same starting node.

Figure 5 shows an untypical case for this graph. The aspect that is untypical here is that the whole-graph cumulative infection curve resembles (somewhat) a single, smooth, S curve. This result was obtained however for the rather artificial starting condition that the first infected node was the most central node in the largest region—hence in the entire graph. The result is that this region takes off immediately. (Our prediction c thus does not hold; but its *assumption* has been violated by our infecting the most central node first.) The next largest region is however infected fairly soon thereafter, so that its takeoff is not clearly seen in the total curve. Finally, the third region takes off

considerably later. But, because it is small, and its takeoff occurs before the blue region is fully saturated, the takeoff of this third region is also not clear from inspection of the total infection curve.

From observing the μ curves, we see that that of the blue region is as predicted. The red region is similar but not visible on the scale of the figure. The largest (black) region's μ curve lacks the initial rise in centrality—but this is to be expected, as the infection began at the top.

Now we move to a more typical case for the sfi graph. Figure 6 shows the behavior for the same infection probability, but with a randomly chosen start node.

The interesting feature of this simulation is that the cumulative curve shows very clearly two takeoffs, and two plateaux. That is, it resembles strongly the sum of two S curves. And yet it is easy to see how this comes about, from our region decomposition: the blue and red curves take off roughly simultaneously, while the largest (black) region takes off only after the other regions are saturated.

The μ curves show roughly the expected behavior—the qualification being that they are rather noisy. Nevertheless the main peak of each μ curve corresponds to the main rise of the corresponding region's S curve.

We reiterate that the behavior seen in Figure 6 is much more typical for this graph than that seen in Figure 5.

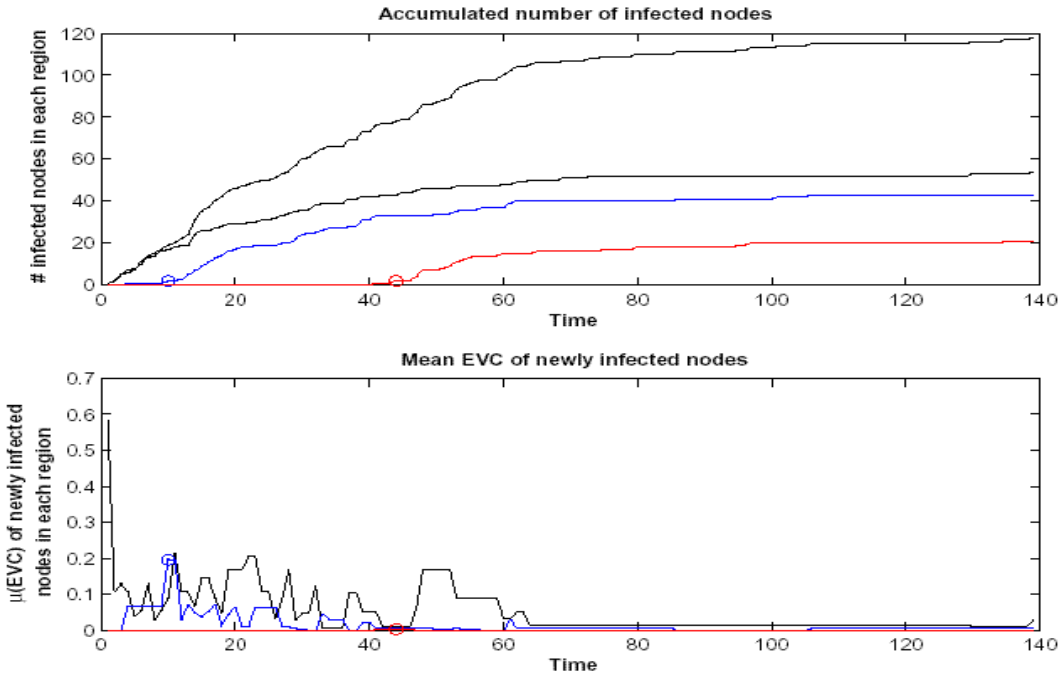


Figure 5. Spreading on the sfi graph, $p = 0.04$. The start node is the most central node in the largest of the three regions (black); hence its infection time is not marked.

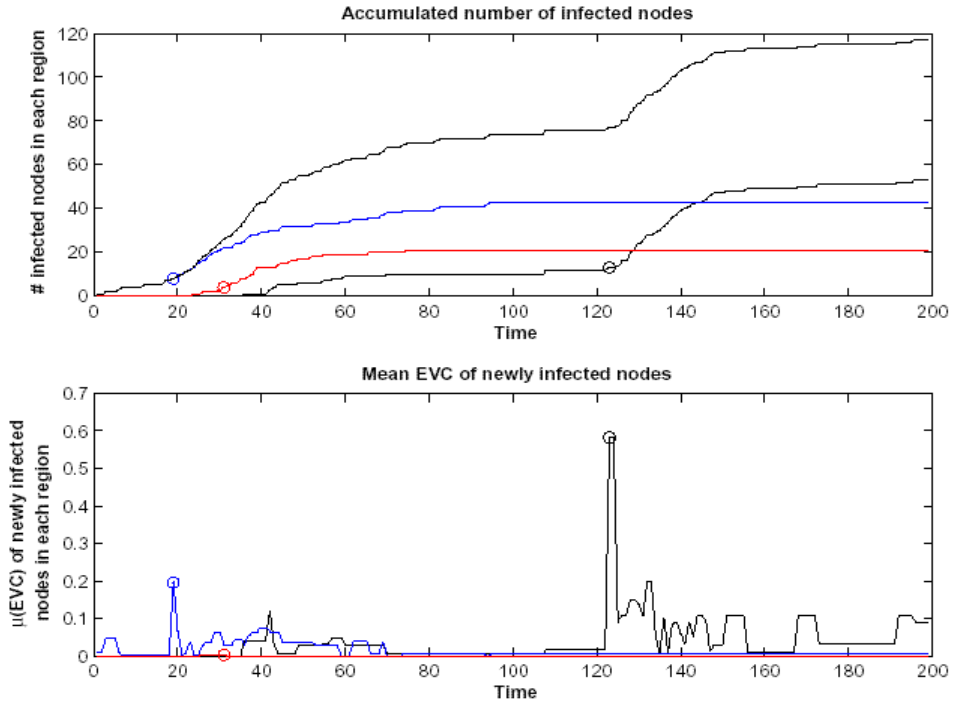


Figure 6. Same as Figure 5, except a randomly chosen start node.

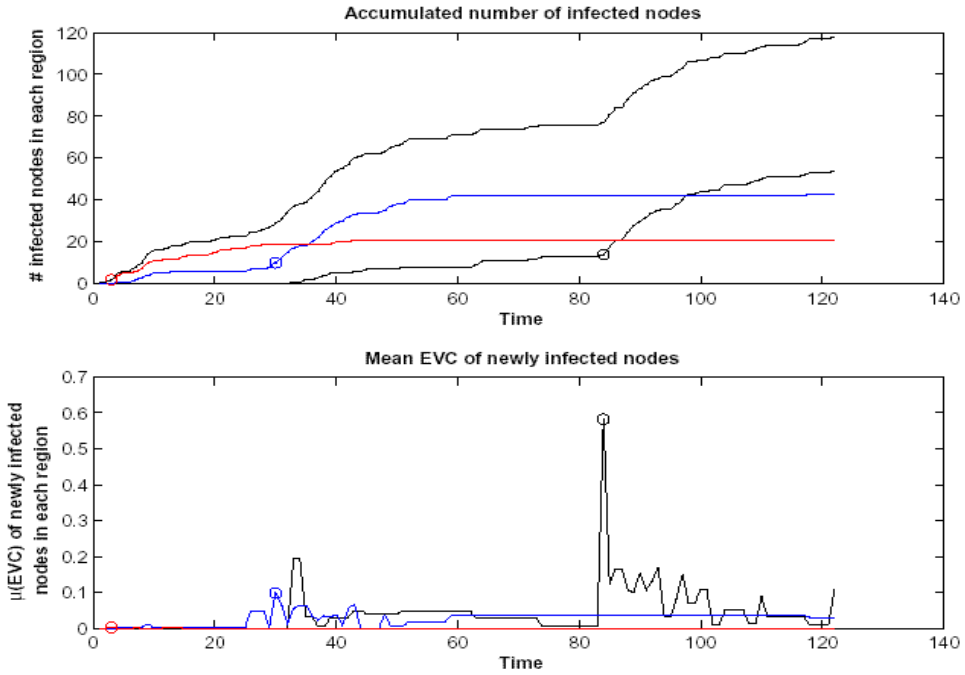


Figure 7. Spreading on the sfi graph, $p = 0.05$. In this case, the existence of three regions in the graph is clearly seen in the total cumulative infection curve.

We examine yet one more example from the sfi graph. Figure 7 shows a simulation with a different start node from Figures 6 or 5, and with $p = 5\%$. The message from Figure 7 is clear: the cumulative adoption curve shows clearly three distinct S curves—takeoff followed by plateau—one after the other. It is also clear from our regional adoption curves that each region is responsible for one of the S curves in the total adoption curve: the smallest (red) region takes off first, followed by the blue region, and finally the largest (black) region. In each case, the time of infection for the most central node of the region undergoing takeoff lies very close to the knee of the takeoff. And, in each case, the peak of the μ curve coincides roughly with the knee of the takeoff.

We note that the behavior seen in Figure 7 is neither very rare nor very common. The most common behavior, from over 50 simulations with this graph, is most like that seen in Figure 6; but we have seen behavior which is intermediate to that in Figures 5 and 6 (ie, neither clearly one nor clearly two takeoffs), and also behavior which is intermediate to that of Figures 6 and 7. In particular, multiple simulations with the same start node and probability as that for Figure 7 have yielded two clear takeoffs (as in Figure 6), three clear takeoffs (as in Figure 7), and intermediate behavior. It is interesting to note that the behavior for this start node, with $p = 1$ (hence with deterministic behavior), gives a cumulative adoption curve which is best described as showing between two and three takeoffs.

Thus the behavior of spreading on the sfi graph gives the strongest confirmation yet of our prediction b: that, for a graph with r regions, one may see up to r, but not more than r, distinct S curves in the total cumulative adoption curve. All of our observations, on the various graphs, are consistent with our predictions; but it is only in the sfi graph that we clearly see all of the (multiple) regions found by our analysis to be present in the graph.

One very coarse measure of the well-connectedness of a network is the number of regions in it, with a high degree of connectivity corresponding to a small number of regions, and with a large number of regions implying poor connectivity. By this very coarse criterion, the hio graph and the sfi graph are equally well connected. Based on our spreading simulations, reported here, we would say that in fact the sfi graph is less well connected than the hio graph. We base this statement on the observation that, for the hio graph, we never found cases where the different regions took off at widely different times—the three regions are better connected to one another than is the case for the sfi graph.

Examination of the sfi graph itself (Figure 8) renders this conclusion rather obvious: the three regions are in fact connected in a linear chain. Thus, it is not surprising that we can, in some simulation runs, clearly see the takeoff of each region, well separated in time from that of the other regions. The hio graph, in contrast, has many links between each pair of regions of the three. We note finally in this context that the Gnutella graphs are very well connected; those that resolve to two regions always have numerous links between the two.

Finally, we note that the sfi graph, while clearly conforming to our assertion b (we see up to three regions), does not wholly agree with assertion a (that each region will have a clear S curve). For instance, in Figure 6, the black region has a small ‘premature’ takeoff around time 40, giving a visible plateau before the ‘main’ takeoff after time 120. Similarly, both the blue and the black regions show such ‘premature’ takeoffs in Figure 7. These observations are not exceptional: the blue region in particular is prone to such behavior, showing two stages of growth in over half of the simulations we have run on the sfi graph. This weakens the support for assertion a; we would then say that a single S curve for a single region is a rule that is followed in most cases, but not all. However these two-stage takeoffs do not contradict the other predictions on our list. For instance, there is always a corresponding ‘premature hump’ in the μ curve for the region in question.

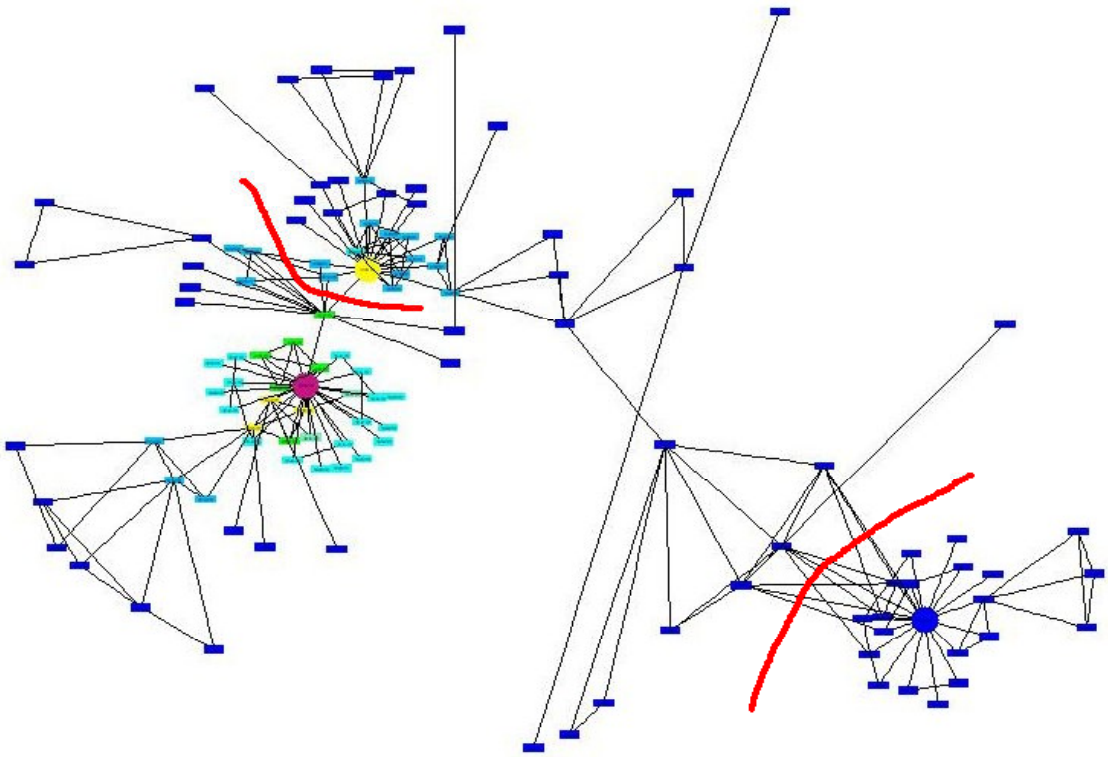


Figure 8. A visualization of the collaboration graph for the Santa Fe Institute. Colors of nodes code for eigenvector centrality, with warm colors implying higher EVC. Also, the most central node in each region is denoted with a round circle. Thick red lines show boundaries between regions. The leftmost region (with a magenta center) is the largest, and has black curves in Figures 5—7; the rightmost region is the smallest (red curves in Figures 5—7); and the middle region has blue curves in Figures 5—7. Hence, in the text, we refer to these three regions, respectively, as the 'black region', the 'blue region', and the 'red region'.

It is clear from Figure 8 where the 'premature takeoffs' arise. The black region includes a subcluster which is easily identified visually in Figure 8 (close to the next region). This subcluster is connected to the remainder of the black region by a single link. These same statements hold true for the blue region—it has a subcluster, joined to it by a single link, but close to the rightmost region. Hence, in Figure 7, for example, the infection starts on the 'far side' of the red region. Then it spreads to the adjacent subcluster of the blue region, but does not reach the most central nodes of that region until after about 20 more units of time (about $1/p$). Soon after the center of the blue region is infected, the subcluster of the black region begins to be infected; but again there is a delay (about twice $1/p$) before reaching the main center of the black region.

We have explored other rules for assigning nodes to regions [1]. For instance, if we simply use shortest distance rather than steepest ascent as our criterion, then the two abovementioned subclusters move: the blue subcluster moves to the red region, and the black subcluster becomes a set of 'border nodes' equidistant from both centers. We have not used the shortest distance rule in this paper, because it tends to place too many nodes in general in border regions. Possibly, one could obtain cleaner S curves for the three sfi regions using the shortest distance rule. However, this rule for defining regions ignores the view that the EVC should be viewed as a height function; hence it is less topographically motivated, and so may also give poorer results in many cases.

4 Towards a quantitative theory

In this section, we will outline a quantitative theory expressing our ideas. We use the word "outline" as we do not obtain exact analytical results for either our definition of node spreading power, or for our dynamic equations for the SI spreading process. Hence we leave the further development of this theoretical outline to future work.

4.1 Definitions

We return to our basic question: how can we quantify the "spreading power" of nodes in a network? If we can do this, then we can test our basic assumption: that eigenvector centrality is a good measure of spreading power.

We propose first to examine nodes pairwise. Consider the pair *ij*. We seek an "infection coefficient" *C(i,j)* which quantifies the ease with which an infection may be spread from *i* to *j* (or vice versa, since the links are symmetric). Clearly, *C(i,j)* should be large if there are many short paths from *i* to *j*. That is: we do not wish only to consider shortest paths; as long as $p < 1$, infection can spread from *i* to *j* via any of the many paths from *i* to *j*. However, longer paths should receive a lower weight in *C(i,j)*.

Next we note that the *(i,j)* element of the matrix A^h contains an integer giving the number of paths of length *h* between nodes *i* and *j* (where *A* is the adjacency matrix). Hence, the collection of matrix elements $A^h(i,j)$ appear to offer the information we need to evaluate *C(i,j)*. There is however one difficulty with using the elements $A^h(i,j)$: they include paths which are self-retracing. Our aim in defining *C(i,j)* is to include paths which might be travelled by an infection moving from *i* to *j* (or the reverse). We do not believe it makes sense to include self-retracing paths in computing *C(i,j)*: for one thing, such paths include paths in which *j* is infected from *i*, and then infected again. Furthermore, since *C(i,j)* must be symmetric, we also exclude paths which revisit the start node before finally reaching the end node. Finally, we generalize this logic to claim that the definition of *C(i,j)* should also exclude paths which revisit *any* intermediate node—hence *all* self-retracing paths should be excluded. (As we see below, self-retracing paths have zero effect on the evolution of the spreading.)

Thus, for each path length *h*, we do not want to use the matrix element $A^h(i,j)$ to count possible infecting paths. Instead, we want the quantity $NSR^h(i,j)$, which is simply a notation for the quantity $A^h(i,j)$ corrected to include only non-self-retracing paths, of length *h*, between *i* and *j*. Here it is important to clarify a question regarding our notation: for A^h, the '*h*' is in fact an (integer) exponent; while for the matrix NSR^h, the '*h*' is only a superscript—we do *not* claim that NSR^h is simply the *h*'th power of NSR^1.

In fact, we know of no general method for computing $NSR^h(i,j)$ for all *h*. Clearly, $NSR^1 = A$; and $NSR^2 = A^2 - Diag(A^2)$. However we know of no general expression for larger *h*.

Given an expression for NSR^h, we build *C(i,j)* as follows. The sum

$$\sum_{h=1}^{\max} NSR^h(i,j)$$

is an integer, counting the total number of distinct, non-self-retracing paths between *i* and *j*. (Note that, given the restriction to non-self-retracing paths, there will be an upper limit "max" to the sum over number of hops *h*.) However, this sum gives equal weight to paths of all lengths. Instead we want a weight function *f(h)* in the sum which gives diminishing weight to paths of increasing length *h*. At this point we do not claim to have a conclusive argument which can determine the choice of weight function *f(h)*—other than that it be monotonically decreasing with *h*. The extreme case [$f(1) = 1$, with $f(h) = 0$ for $h>1$] is as uninteresting as the equal-weight case, since it only gives weight to one-hop paths, and so gives the result that a node's spreading power is simply its degree.

For simplicity, we set $f(h) = w^h$, where w is some positive weight which is less than one. One may argue that this choice best reflects the case in which each link has a probability $p < 1$ (per unit time) of transmitting the infection. This gives the following expression for the infection coefficient:

$$C(i,j) = \sum_{h=1}^{\max} w^h NSR^h(i,j) \quad . \tag{5}$$

We take (5) as our working definition of infection coefficient. Now we wish to use this definition in order to obtain a definition of spreading power for a node. The idea is again simple: If a node has a high infection coefficient with respect to many other nodes, then it has a high spreading power. And in this case we see no need for a decreasing weight function; to find the spreading power $S(i)$ of node i, we simply will add up the infection coefficients involving node i. That is:

$$S(i) = \sum_j C(i,j) = \sum_j \sum_{h=1}^{\max} w^h NSR^h(i,j) \quad . \tag{6}$$

Note that the sum over j can be unconstrained, since, by excluding self-retracing paths, we have forced $C(i,i) = 0$.

4.2 An approximation

Equation (6) gives an expression for the spreading power $S(i)$ of node i. However, lacking a general expression for NSR^h, we cannot test our assumption—namely, that EVC has a strong positive correlation with $S(i)$.

We can however say something about this connection, if we make a gross oversimplification of our expression for $S(i)$. That is, we include self-retracing paths, and so replace NSR^h with A^h. We will use $S^a(i)$ to denote this approximation.That is, we take

$$S^a(i) = \sum_j \sum_{h=1} w^h A^h(i,j) \quad . \tag{7}$$

We leave the upper limit of the h sum unspecified. (It is unbounded if one considers all self-retracing paths.)

Now we gather together all the coefficients $S^a(i)$ into the single vector S^a. Also we define $\mathbf{1}$ to be a vector, of length equal to the number of nodes in the graph, with all entries equal to one. Then

$$\sum_j A^h(i,j) = \left(A^h \mathbf{1}\right)_i \quad .$$

Hence

$$S^a = \sum_{h=1} w^h A^h \mathbf{1} \quad .$$

Next we decompose the vector $\mathbf{1}$, using as a basis the set of eigenvectors $\underline{\alpha}$ of A:

$$\mathbf{1} = \sum_\alpha \sigma_\alpha \underline{\alpha} \quad , \tag{8}$$

where σ_α is the scalar product of $\mathbf{1}$ and $\underline{\alpha}$. This gives, in turn,

$$S^a = \sum_\alpha \sum_{h=1} \left(w\lambda_\alpha\right)^h \sigma_\alpha \underline{\alpha} = \sum_\alpha g_\alpha \underline{\alpha} \quad . \tag{9}$$

The quantity g_α is simply the scalar product of $\underline{\alpha}$ with the vector S^a. Now we note that the principal eigenvector of A is simply the vector of eigenvector centrality scores. Let us denote this vector as $\underline{\pi}$. If our basic assumption is true (and if S^a were a good approximation to S), then the overlap g_α should be

positive, and large in absolute value compared to the other values g_α. That is, S^a should be composed "mostly" of π.

Now we note that

$$g_\alpha = \sum_h (w\lambda_\alpha)^h \sigma_\alpha .$$

This reveals two reasons why g_π should be the largest of all the overlaps g_α. (i) The eigenvalue λ_π is the largest of all the eigenvalues. (ii) The coefficient σ_α is simply the algebraic sum of the components of the vector α. However the components of π are *all positive*—while, for all other eigenvectors, they are of mixed sign. Hence we expect the overlap σ_π to be, in general, larger than the others. For instance, if we normalize the set of α with the L1 norm, then σ_π is equal to 1, while all of the other σ_α are less than one.

These arguments would be persuasive, had they been obtained using S rather than S^a**.** Instead, they must be viewed as suggestive at best. It is interesting to note that the value of the parameter w plays no role in the argument.

4.3 Modelling the spreading process

Now we develop a mathematical model of the spreading process. Not surprisingly, the adjacency matrix A plays a central role in our model; hence this mathematical model throws some light on our notions of the role of eigenvector centrality in spreading. We begin our modelling with the deterministic case, p =1.

4.3.1 *p* =1

The case of SI spreading with p = 1 has much in common with the calculation of eigenvector centrality. To make them as similar as possible, we note that the result of an iterative EVC calculation, using the power method, is independent of start vector. Therefore, we take the start vector for the EVC calculation to be the same as that for spreading: one unit of weight at node J, and zero elsewhere. We call this start vector x^0.

To calculate EVC by the power method, starting with x^0, one simply multiplies by the adjacency matrix A in each iteration. The matrix A is thus the "time evolution operator" for this process: $x^{t+1} = Ax^t$. After many iterations, there is large weight at every node (assuming one does not rescale the weights). These weights all grow with each subsequent iteration; but the *relative* values of the weights converge to a set of constants, giving a weight vector growing in length but not changing in direction. This convergent vector (regardless of length) is the eigenvector e whose entry e_i is the EVC for node i. Also, given our start condition, all the weights will be nonnegative integers at all times, and positive integers after some finite time.

Now we consider SI spreading for the same start condition x^0. The time evolution equation is as follows:

$$x^{t+1} = \mathrm{sgn}\left[(1 + A)x^t\right]. \tag{10}$$

That is, (i) the A operator of the EVC evolution equation is replaced with (1+A); and then (ii) the result is truncated using the sgn operator—which maps any positive integer to 1, and 0 to 0.

The first difference (i) reflects the fact that, each time a node sends infection to its neighbors, it retains its infected state—whereas, with the EVC calculation, nodes send out all of their weight at time t, so that their weight at t+1 depends entirely on inputs from neighbors. This difference is fairly trivial: the dominant eigenvector of (1+A) is identical to the dominant eigenvector of A. Hence, without the sgn operator, both processes would converge to the same distribution.

The second difference—the truncation operator—reflects the fact that a node, once infected, cannot become "more infected" as a result of repeated transmissions from its neighbors. This difference of

course has a dramatic effect—the convergent distribution for the spreading process is the saturated state: a vector of all 1's. We note, in this context, that a major source (but not the only one) of repeated transmissions to a given node j is infection paths that are self-retracing. For example, one time step after node j is infected, all of j's neighbors will be infected; and after one more time step, these neighbors will in turn, under the action of the A part of $(1+A)$, retransmit to j. The sgn operator nullifies the effects of all transmissions to a given node after the first; hence, self-retracing infection paths have zero effect under the SI time evolution operator.

Finally, we note that Equation (10) calls for truncating after every time step. However, it is also true that

$$x^t = \mathrm{sgn}\left[(1+A)^t x^0\right] . \tag{11}$$

That is: since all positive integers map to 1, one can find the distribution at any time t by applying the $(1+A)$ operator t times, without truncating, and then finally truncating only once, at the end of the run. Before this final truncation, for large t, the distribution will approach the same as that for eigenvector centrality—then the truncation operator throws away all this information, and simply places a 1 at every infected node.

This observation suggests that one might observe a "piling up" of weight at nodes of high EVC, if one were to modify the spreading process in such a way that it were possible to pile up weight. We then note that *probabilities* have this property—that is, repeated time steps, when each link has probability p for transmission, will steadily increase the probability that an uninfected node will be infected by its infected neighbors. Hence we are motivated to look at the case of spreading for $p < 1$.

4.3.2 $p < 1$

To discuss the probabilistic case, we consider a series of experiments, with fixed network topology, fixed p, and fixed start vector x^0. For many such experiments, we can discuss the probability of a given node j being infected at time t. As noted above, this probability (denoted u^t_j, to avoid confusion with the parameter p) will grow with time. An exact theory of these experiments will hold the probability u^t_j less than 1 for every node, for all time. However any approximate theory—especially when the question of interest is the rate of growth of infection probability with time—must take care to avoid spurious effects due to the approximations used.

An exact theory may be stated as follows. Define r^t_j to be the probability that node j receives a transmission at time t. (The timing convention here is that, if node j is not infected at time t, and receives a transmission at time t, then it is infected at time t+1.That is, we can think of the event r^t_j of receiving a transmission as occurring slightly 'after' time t.) The probability r^t_j depends of course on the infection probabilities u^t_k for the neighbors k of node j. Specifically: neighbor k has probability $(u^t_k)p$ of transmitting to j at time t. Hence the probability of *no* neighbor transmitting to j is

$$\prod_{k=nn(j)} \left(1 - u^t_k p\right) \ ;$$

and so the probability of at least one neighbor transmitting to j is

$$r^t_j = 1 - \prod_{k=nn(j)} \left(1 - u^t_k p\right) \ . \tag{12}$$

Now we look at u^t_j. We note that node j will be infected at time t+1, unless two conditions are fulfilled: (i) it was not infected at time t, *and* (ii) it did not receive a transmission at time t. Hence the probability that j is not infected at t+1 is $(1 - u^t_j)(1 - r^t_j)$. Thus the probability that j is infected at t+1 is

$$u^{t+1}_j = 1 - \left(1 - u^t_j\right)\left(1 - r^t_j\right) = u^t_j + r^t_j - u^t_j r^t_j \ . \tag{13}$$

Equations (12) and (13), along with the starting condition $u^0_j = x^0_j = \delta_{j=J}$, give the dynamical rule for the growth of probability in a system with all links having transmission probability p per unit time.

These equations are of course easier to simulate than to solve! (Examples of such simulation were given in Section 3.) However we will seek to learn what we can from them.

First we note that the RHS of (13) is of the form (noninteracting case) + (corrections). The corrections come in, once again, because we cannot infect a node twice. For example, the term $u_j^t r_j^t$ is a correction of this sort: the naïve rule (ie, one ignoring problems of double-counting) would be that the infection probability at t+1 is simply the first two terms, ie, the probability of infection at t, plus the probability of transmission at t. Also, the naïve version of (12) would ignore all terms of higher order than p in the product, giving

$$r_j^t \approx \sum_{k=nn(j)} u_k^t p = \left(pAu^t\right)_j \ . \tag{14}$$

Combining the two naïve forms of (12) and (13) gives

$$u^{t+1} \approx u^t + pAu^t = \left(1+pA\right)u^t \ . \tag{15}$$

Thus we see that our naïve time evolution operator (ignoring double counting) is simply (1+pA). Thus the naïve version for p = 1 is simply (1+A)—the same as is obtained from the time evolution operator from Section 4.3.1, if we ignore the sgn operator.

We note in this context that the 'system matrix' (time evolution operator) of Wang et al [11] is simply our 'naïve' time evolution operator, corrected for a uniform decay rate for the infected state. Thus their system operator, our naïve time evolution operator, and the adjacency matrix A all have the same dominant eigenvector, the components of which are the nodes' EVC values.

Thus we find that the probability vector may be written as

$$u^t = \left(1+pA\right)^t x^0 \ + (\text{corrections}) \ . \tag{16}$$

The operator (1+pA) has the same dominant eigenvector as does A itself. Hence repeated multiplication by (1+pA) will drive a vector towards a distribution consistent with eigenvector centrality: each node's weight will be proportional to its EVC score. However, at large t, we know that the (corrections) must also be large; so we cannot draw any clear conclusion about the large-t case from Equation (16). We do however believe that Equation (16) supports assertion d. in our list of assertions a—g of Section 2. That is, for small t and small p—which thus characterizes the early, flat part of the S curve—we expect the corrections to be small.

We have carried out a perturbation expansion, in powers of p, for the coupled equations (12)—(13) above. We expanded to terms of $O(p^3)$. This allows us to look at the vector u^t up to time t = 3, and hence within a radius of three hops from the start node J. Our results are somewhat cluttered, and will not be given here. However, they conform (as they must) to the form of Equation (16); and furthermore, we see that, for each power of A, the leading (lowest order in p) term is simply $(pA)^t$—the "naïve" term. That is, the correction terms, at each radius from J, are of higher order in p than the naïve term. Thus the naïve term dominates—and, as we pointed out above, the naïve term tends to drive the probability vector towards higher EVC. Thus, our perturbation expansion also supports our prediction d.

This observation, although gleaned from our perturbation expansion, in fact holds in general. That is, paths between i and j which are short have a higher probability of infecting j from i than longer paths—for instance, a path of length 4 acquires a factor p^4, while a path of length 3 has a corresponding factor p^3. Hence, even without extending our series expansion in p to higher order, we know that the leading term *at a given radius from start node J* comes from the "naïve" part of (16)—which is, in turn, the part which tends to build up weight in the same pattern as the eigenvector centrality [the dominant eigenvector of (1+pA)]. In other words: we are stating the obvious—that a node j is most likely to be infected from J over a shortest path between them—and then drawing a less obvious conclusion: that the (corrections) in (16), for the infection $J \to j$, are likely to be smaller than the leading term. This is certainly true for very small p; and yet it is also true for the other extreme of p = 1—where only shortest paths play a role. Hence we speculate that the naïve term in (16) dominates for any p. If this is true, then infection probabilities will approach a distribution proportional to eigenvector centrality, as long as those probabilities do not approach their upper limit

of 1 too closely. That is, we again find support for our claim that infection will move towards higher EVC. Finally, we note that our simulations show qualitatively the same behavior (for the same start node) over the whole range of p, from a few per cent to one. This observation supports our speculation that the dominant term in (16) is the same for all p.

5 Discussion and future work

We have applied our method of structural analysis for undirected graphs, developed in [1], to the problem of gossip-like spreading on a network. We believe that our topographic picture of the structure of a network, based on using eigenvector centrality (EVC) as a height function over the network, with mountains, peaks, slopes, and valleys, is an excellent starting point for an understanding of spreading. In this work, we have built from this starting point, and developed a set of qualitative arguments which yield seven specific predictions (Section 2). Our picture, in short, is that an initial infection on the side of a mountain will run 'up' the mountain, while the rate of infection of new nodes grows with height. This is a self-reinforcing process, so that infection rate takes off at some point high up on the mountain, and the whole top is saturated quickly; finally the remaining hillsides are saturated at an ever decreasing rate. These predictions are tested, and convincingly confirmed, in a series of simulations that we have run on various social networks in our possession.

To supplement these qualitative arguments and simulations, we have developed a mathematical theory of two things: the definition of the spreading power of a node, and the dynamics of simple SI spreading. In each case, exact solutions are not possible, due to the problem that double infections must not be counted. However, in each case, we have shown that ignoring the double-counting problem gives an approximation which supports our basic claim (spreading power may be approximated by EVC). In particular, we present arguments why the correction terms due to double counting are likely to be small compared to those which ignore double counting; and these latter terms support the claim that infection probability is positively correlated with EVC.

There is much work remaining to be done on the theoretical side. The connection between our definition of spreading power $S(i)$ for node i, and the eigenvector centrality EVC for the same node, needs to be tested further. To do this, one needs ways to calculate the number of non-self-retracing paths at distance h. This could be done numerically for the smaller graphs we have studied, and compared with EVC. We also have some (unpublished) analytical results for EVC on certain types of tree graphs; it would be interesting to compare these values to values for NSR^h.

It would be of interest to apply our topographic approach to understanding spreading to other cases. For one thing, our analysis suggest simple ways of *hindering* spreading: one could immunize (permanently) key nodes in the network, such as highly central nodes in one region, or nodes whose links provide bridges between regions. These strategies for 'immunizing' a network could be tested via further simulations. Also, one can study ways for modifying the topology of the network, towards the goal of hindering or helping spreading. In the former case one might seek to isolate regions even further; in the latter, one might look for small modifications that cause two or more regions to fuse into one. Again, such strategies may be tested by simulations. Other, perhaps more realistic, cases to be studied, in the light of our topographic picture, include the case in which nodes are infected from 'outside' the graph at some steady rate; or the case where nodes lose their infection status after some time (SIS), perhaps after a refractory period (SIR).

We note that the sfi graph offered the most extreme behavior, stemming from the weak coupling both between and within its three regions. Thus we find the sfi graph to be the least well connected, in terms of criteria derived from our structural analysis, and from our observations of spreading. We have argued in [1] that the property of being poorly connected is related to poor mixing, and thus to a small eigenvalue gap (difference between the dominant and second eigenvalues of A). It would be of interest to test these ideas with the set of graphs studied here. If the gap is small in poorly connected (by our definition) graphs, then many things will be relatively sensitive to small changes in topology: their EVC values (from the dominant eigenvector); their topography, as obtained by our analysis; and their spreading behavior. This too could be tested.

More generally, we wish to deepen and render more quantitative our notion of the well-connectedness of a graph. Clearly, our coarse starting point (number of regions) give useful information in itself; but much more remains to be done. Besides making a connection to the eigenvalue gap, one could seek to quantify the degree of inter-region connectedness. Furthermore, one should seek connections and correlations between these different measures.

Finally, we note that our method of network structure analysis suggests a method of graph visualization. Figure 8 is an example of this: nodes in each region are placed close together on the page, so that the regions (and the height function defining them as 'mountains') are visually clear. Figure 8 represents work in progress. We hope to be able to present improvements on, and refinements of, this visualization approach in future work. One very attractive goal is to display epidemic spreading simulations on a network visualized as in Figure 8—in the form of snapshots, or a movie. Then one can hope to see our predictions a—g, not in the form of static plots as in this paper, but in the form of time development of the infection over the topography (as displayed in 2D) of the graph.

Acknowledgments. We thank Michelle Girvan for sharing data on the SFI collaboration graph with us, and Mark Burgess for helpful discussions. This work was partially supported by the Future & Emerging Technologies unit of the European Commission through Project BISON (IST-2001-38923), and partially supported by the EU within the 6th Framework Programme under contract 001907 (DELIS).

References

[1] Canright G, Engø-Monsen K: Roles in Networks. Science of Computer Programming 2004;53: 195—214.

[2] Jovanovic MA, Annexstein FS, Berman KA: Scalability issues in large peer-to-peer networks - a case study of Gnutella. Technical Report, University of Cincinnati (2001).

[3] Canright G, Engø-Monsen K, Weltzien Å, Pourbayat F: Diffusion in social networks and disruptive innovations. IADIS e-Commerce 2004 proceedings. Lisbon 2004.

[4] Girvan M, Newman MEJ: Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA 2002;99: 8271—8276.

[5] Bonacich P: Factoring and weighting approaches to status scores and clique identification. Journal of Mathematical Sociology 1972;2; 113—120.

[6] Newman MEJ: The structure and function of complex networks. SIAM Review 2003;45: 167—256.

[7] Pastor-Satorras R, Vespignani A: Epidemic Spreading in Scale-Free Networks. Phys. Rev. Lett 2001;86: 3200—3203.

[8] Pastor-Satorras R, Vespignani A: Epidemic dynamics and endemic states in complex networks. Phys. Rev. E 2001;63: 066117.

[9] Newman MEJ: Spread of epidemic disease on networks. Phys. Rev. E 2002;66: 016128.

[10] Brauer F: A model for an SI disease in an age-structured population. Discrete and Continuous Dynamical Systems 2002;B2: 257—264.

[11] Wang Y, Chakrabarti D, Wang C, Faloutsos C: Epidemic spreading in real networks: an eigenvalue viewpoint. Proceedings, 22nd Symposium on Reliable Distributed Systems (SRDS 2003), 25—34.

[12] Rogers EM: Diffusion of Innovations, ed 3. Free Press, New York, 1983.

Lightweight centrality measures in networks under attack *

Giorgos Georgiadis † Lefteris Kirousis ‡†

Abstract

In this paper we study deliberate attacks on the infrastructure of large scale-free networks. These attacks are based on the importance of individual vertices in the network in order to be successful, and the concept of centrality (originating from social science) has been already utilized in their study with success. Some measures of centrality however, as is betweenness, have disadvantages that do not facilitate the research in this area. We show that with the aid of scale-free network characteristics such as the clustering coefficient we can get results that balance the current centrality measures, but also gain insight in the workings of these networks.

Keywords:
dynamics of scale-free networks, centrality, clustering coefficient, network attack

1 Introduction

Recently there has been an increase of interest in many natural and artificial large scale networks. For example, it is estimated that the network of web pages currently consists of several billions of vertices [1]. Many companies owning a search engine would like to know the specific characteristics of this

*This research was partially supported by the EU within the 6th Framework Programme under contract 001907 "Dynamically Evolving, Large Scale Information Systems" (DELIS) and was also partially supported by the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II), and particularly the Program *Pythagoras*.

†Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece. Email address: {georgiad, kirousis}@ceid.upatras.gr

‡Research Academic Computer Technology Institute, P.O. Box 1122, 26110, Patras, Greece

network for purposes of page indexing and maybe to predict, up to a point, its future behavior. In general, a network consists of objects, of different kind in each area of interest, which are represented by vertices, and connections between them, represented by edges. Many theoretical results exist due to graph theory of discrete mathematics, which handles such objects. This kind of modelling is possible for a variety of large scale networks, both natural occurring and artificial, such as networks of acquaintances, citation, food chains, infections, proteins [2, 3, 4, 5, 6, 7] or networks of power grids, internet infrastructure, web pages, and so on [8, 9, 10, 11, 12, 13, 14].

Especially in the case of social networks of any kind, they have been studied by scientists in the social science for quite a while, with methods such as questionnaires and personal interviews. One persistent question was that of the centrality of an individual in such a network, or how well "connected" this person is in his environment. For example a measure of this connectivity is a vertice's degree, that is the number of its immediate neighbours. The size of these kind of networks is in the order of several tens or in some cases several hundreds of vertices and so the research is not directly applicable to the large networks arising, mainly technological. This is just one of the problems scientists face with the large scale networks.

Another problem is the nature of these networks. The most prominent model in graph theory until recently have been the random graph model introduced by Erdős and Rényi [15]. In this model, any two vertices have equal probability to be connected by an edge. This model is very well studied and many results exist on it, but unfortunately does not describe our observations in real world networks. In many real world networks there exist a percent of vertices that appear to be better connected to the rest of vertices. Furthermore, during the network growth, they gain easier connections and certainly not with equal probability to the other vertices. One of the unique characteristics of these networks, that distinguishes them from previous studied networks, is the power law form of the distribution of the vertex degrees.

In this study we address the problem of network attack. We are assuming the existence of an adversary that wants to harm a network, by directly attacking and removing the vertices comprising it. He has the ability to measure some variables of the network in order to make educated guesses as to what vertex should target next. In our experiments we measure the efficiency of strategies based on such measures as to the computational time needed to target a vertex on the specific network and the end result after the attack. We care about measures that produce most harm with little effort. The strategies we are using will be based on the centrality measures taken

from traditional social network research.

In this paper we take into account previous similar studies and we compare our findings with theirs. We propose an attack strategy that is a trade-off between worst and best strategy so far and has significant and unique advantages. We also offer insight to the workings in power-law graphs and indicate future research areas.

The paper is organized as follows. In Section 2 we review the fundamental concepts needed in our study, along with a separate discussion on the most widely used measures of centrality. In Section 3 we introduce the use of centrality measures as attack strategies. Our experiments can be found in Section 4 along with our analysis on the results.

2 Fundamental Concepts and Definitions

In this section we define the graph-related concepts that we will use in our experiments and analysis, along with the main notions of centrality that are of interest to us.

2.1 Graph-related concepts

Throughout this paper we represent a network as an undirected, unweighted graph $G(V, E)$, where V is the set of vertices (i.e. computers) and E is the set of edges (i.e. communication links). Their sizes are $|V| = n$ and $|E| = m$ respectively. The degree k_v of a vertex v is the number of edges originating from or ending to vertex v. We are interested only in graphs generated by the preferential attachment procedure, first proposed by Barabási and Albert [16], which we will briefly describe here. The iterative creation process consists of 4 steps:

Step 0: Initially the graph has n_0 vertices and no edges.

Step 1: Add a new vertex v to the graph.

Step 2: Create l edges, each time connecting the new vertex v to a vertex w, with probability proportional to this vertice's degree:

$$p_w = \frac{k_w}{\sum_{u \in V} k_u}$$

Step 3: Repeat steps 1 and 2 for $(n - n_0)$ rounds.

The end result of this procedure is a graph of n vertices and $m\,(n - n_0)$ edges, with vertex degree distribution $P\,(k)$ that follows a power law, with exponent $\gamma = 3$ ($P\,(k) \propto k^{-3}$). We will call such a graph a Barabási-Albert network, or BA network for short. Of course there are many models with creational procedures that generate graphs with power law degree distributions (like the Watt's "small worlds" [17, 18]), but we feel that the classical preferential attachment model describes complex networks generation in a more general way. Other than that, our specific results may differ in other models but the essence of our insights should still apply.

Other concepts that will be used are that of the 1-neighbourhood and 2-neighbourhood of a vertex. Having a vertex v as center, its 1-neighbourhood $\Gamma_1\,(v)$ consists of all vertices at distance 1, i.e. its direct neighbours. Such neighbours from now on will be called *first-neighbours* and of course it holds that $|\Gamma_1\,(v)| = k_v$. Similarly, the 2-neighbourhood of a vertex v consists of all vertices at distance exactly 2 (from now on *second-neighbours*) and it holds that $|\Gamma_2\,(v)| \leq \sum_{w \in \Gamma_1(v)} k_w$. The inequality in the above expression stands for the fact that some first-neighbours of v may have common neighbours, thus limiting the number of (unique) vertices in $\Gamma_2\,(v)$. This phenomenon is called *clustering* [19, 13, 20] and is not only possible but characteristic of power law graphs. This relation between a vertex and its first- and second-neighbours lead to the emergence of several structures in the graph, most common of which is the triangle. In a triangle, three vertices are joined by three edges, one for each pair of vertices. The existence of triangles is characteristic of a power law graph, and its the feature that makes them so popular in different disciples: for example, in social science, two of one's friends have greater probability of knowing each other than two random-picked strangers.

2.2 Standard Centrality Measures

2.2.1 Degree Centrality

The degree centrality measure gives the highest score of influence to the vertex with the largest number of first-neighbours. This agrees with the intuitive way to estimate someone's influence from the size of its immediate environment. The degree centrality is traditionally defined analogous to a vertice's degree, normalized over the maximum number of neighbours this vertex could have. Thus, in a network of n vertices, the degree centrality of

vertex i, C_i^D, is defined as:

$$C_i^D = \frac{k_i}{n-1} \tag{1}$$

The normalization in the region $[0, 1]$ is used here to make the centrality of different vertices comparable, and also independent of the size of the network.

2.2.2 Closeness Centrality

This notion of centrality focuses on the idea of communication between different vertices. The vertex which is "closer" to all vertices gets the highest score. In effect, this measure indicates which one of two vertices needs fewer steps in order to communicate with some other vertex. Because this measure is defined as "closeness", the inverse of a vertice's mean distance from all others is used. Hence, if C_i^C is the closeness centrality, and d_{ij} the shortest distance between vertices i and j in terms of edge steps:

$$C_i^C = \frac{n-1}{\sum\limits_{j \in V} d_{ij}} \tag{2}$$

Again, this measure is normalized in the region $[0, 1]$. Additionally, it should be stated that the distance between two disconnected vertices must be a predefined very large value and not infinite, if it is desirable to discern among low closeness scores.

2.2.3 Betweenness Centrality

Betweenness centrality refines the concept of communication, introduced in closeness centrality. Informally, betweenness centrality of a vertex can be defined as the percent of shortest paths connecting any two vertices that pass through that vertex. The normalized version divides this value with the maximum possible betweenness centrality, that is all possible shortest paths in a completely connected graph. If C_i^B is the betweenness centrality of vertex i, (u, i, v) is the set of all shortest paths between vertices u and v passing through vertex i and (u, v) is the set of all shortest paths between vertices u and v, then:

$$C_i^B = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} \frac{|(u,i,v)|}{|(u,v)|}}{(n-1)(n-2)} \tag{3}$$

This definition of centrality explores a vertice's ability to be "irreplaceable" in the communications of two random vertices. It is of particular interest in the study of network attacks, because at any given time the removal of the maximum betwenness vertex seems to cause maximum damage in terms of connectivity and mean distance in the network. Its main disadvantage is that the summation operator practically means that it needs global information about the network, in order to compute a single vertice's betweenness, and that is simply not possible in many contexts. For the same reason it is expensive in computing time to compute a vertice's score, although this disadvantage was significantly improved recently [21, 22]. The importance of betweenness centrality as an attack strategy is further discussed below.

3 Centrality Measures As Attack Strategies

It has been shown in the past that the "random vertex hit" strategy performs poorly [23, 24], due to the hierarchical effect these networks present, i.e. a random vertex has increased probability to be one of the less connected vertices, since there are so many of them. So it is desirable to use a strategy that achieves better results, and such strategies could be based upon a vertex measure that can profile each vertice's potential by its value only. A number of publications exist [23, 25, 26, 24, 27], addressing the question of which strategy is best in achieving maximum destructive result with less vertex hits, the most extensive of which, to our knowledge, is [23]. Summarizing the results, the comparison is based upon two axes: different strategies and recalculation of measures. The different strategies studied are vertex deletion based upon degree centrality scores and upon betweenness centrality scores, and it is clearly shown that betweenness produces better results. The recalculation of the involved measures refers to the recalculation of degree/betweenness centrality after each vertex deletion.

In this section we introduce a strategy that balances the advantages and disadvantages of the above mentioned strategies. We will not study the closeness centrality, as it has the same basic flaws as betweenness and none of its advantages. Furthermore we will focus on the recalculated versions, since the distribution of these measures may vary significantly between deletions.

3.1 How to measure destructive power?

We are interested in the destruction of the network under consideration. Ideally that would meant the isolation of each vertex, but it can be argued that it is enough to break the network to a sufficiently large number of connected

components. We choose to examine only the size of the largest component, as a particulary small largest component would mean that the network has degenerated into many small connected components. Additionally, we can measure directly the impact of vertex deletions in the hierarchical structure by examining what happens to the largest component: a successful attack would probably target this component and shrink its size dramatically. Another reason is that this technique has been used successfully during previous studies [23], and its use will make our results directly comparable. We use specifically a normalization over the largest component size with the initial network size, in order to produce a percentage comparable between different size networks.

Since we start with a connected network, it would take some time before it becomes disconnected, and during that time the size of the largest component would not carry significant information. Thus, in addition to the largest component size, we use the mean shortest path length of the network, and specifically its mean inverse. The mean shortest path length is the mean length of all shortest paths in the network, between all pairs of vertices. If by d_{uv} we denote the length of the shortest path between vertices u and v, then the mean shortest path length l, in a network of n vertices, is

$$l = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} d_{uv}}{n\,(n-1)}.$$

The mean inverse of shortest path length l^{-1} is defined as

$$l^{-1} = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} \frac{1}{d_{uv}}}{n\,(n-1)}$$

In practice we use the mean inverse of shortest path length because by doing that we nullify the effects of disconnected vertices and their "infinite" distance. Increasing mean value of this measure means that average distances in the network are increasing, and this subsequently means that the attack in the network produces quantifiable, destructive results. Clearly, since we use the mean inverse of this measure, we expect it to decrease with time.

3.2 Standard Centrality Measures Explained

As it was already mentioned, the random vertex hit strategy has practically no effect to the network's integrity, and that is because it cannot take

into consideration its hierarchical structure. This is exactly where degree based attack succeeds. By targeting the highest degree vertices first, it attacks directly the global network connectivity. It must be pointed out that not all properties of Barabási-Albert networks are known. Initially it was believed that high degree vertices were connected with other high degree vertices preferably over lower degree vertices (assortative mixing) [28]. Recent studies [29] show that Barabási-Albert networks are rather neutral on this property, and in some cases even show the opposite behavior (disassortative mixing), i.e. high degree vertices prefer lower degree vertices to connect to. We believe this observation can explain the success, albeit partial, of this strategy, as in the disassortative mixing the deletion of highest degree vertex would effect many vertices: but since this is not a predominant phenomenon the effectiveness of this strategy would be limited.

On the other hand, the betweenness based strategy seems ideal, especially with the performance metrics used (mean shortest path length, largest component size). By definition, the betweenness measures the ability of a vertex to be irreplaceable in shortest paths throughout the network. So when this vertex is removed, inevitably all shortest paths depended on it will be removed also, and equally long or longer paths would take their place. This has an obvious impact in the mean shortest path length, which is constantly non-decreasing, at least as long as a unique giant component exists. Such high betweenness vertices, which connect many others with shortest paths, would be initially located in the largest component as most vertices would be located there. Therefore, the failure of these paths effects also the largest component size, since multiple failures may produce disconnected vertices. Similar arguments can be used with the closeness centrality.

3.3 Proposed Strategies

We propose a family of strategies based, in part, on edge degree. Although a formal definition of edge degree does not exist, we experimented with several possible definitions, all based on vertex degree. Specifically, an edge's degree has some connection with the endpoint vertices of this edge. As was the case in [23] we settled with the edge degree being the product of the endpoint vertices' degrees, as it followed closely our intuition on the importance of edges. If $e = (w, u)$ an edge with endpoints w and u, having degrees k_w and k_u respectively, its edge degree $k_e^{\Gamma_1}$ is defined as:

$$k_e^{\Gamma_1} = k_w \cdot k_u$$

The first strategy which uses the edge degree to select vertices does so

by first selecting the edge with maximum degree, and then the vertex of this edge with maximum (vertex) degree. In case of multiple edges/vertices with same (maximum) degree, we choose uniformly at random. Note that this strategy examines the immediate neighbourhood of each endpoint vertex, and scores higher edges having endpoints with large 1-neighbourhoods. From now on we will refer to this strategy as "1-neighbourhood edge degree" strategy.

The second strategy defines edge degree as the product of the 2-neighbourhoods of its endpoint vertices. This 2-neighbourhood edge degree $k_e^{\Gamma_2}$ of an edge $e = (w, u)$ is defined formally as:

$$k_e^{\Gamma_2} = \sum_{i \in \Gamma_1(w)} k_i \cdot \sum_{j \in \Gamma_1(u)} k_j$$

The vertex selection is exactly the same as before: choose the edge with maximum degree and then the endpoint vertex with maximum degree. We will refer to this strategy as "2-neighbourhood edge degree" strategy.

The third proposed strategy is based on the "2-neighbourhood edge degree", as defined above. The main difference is that it penalizes the existence of triangles in which the edge is present. Specifically, it divides the above computed edge degree by the number of triangles that this edge participates in, plus one to avoid division by zero. Thus, if T is the number of triangles involving the edge in question as a side of the triangle, the formal definition of the alternative edge degree is:

$$K_e^{\Gamma_2} = \frac{\sum_{i \in \Gamma_2(w)} k_i \cdot \sum_{j \in \Gamma_2(u)} k_j}{T + 1}$$

We will refer to this strategy as "2-neighbourhood edge degree with penalty".

4 Experiments

For the experiments we used networks of 1500 vertices, created by the BA procedure mentioned in Section 2. The parameters of importance are the size of the initial network (before the procedure starts adding vertices) and the degree of each added vertex. We used degree 5 for each new vertex and we kept the initial network small, consisting of 5 vertices connected with random edges. Each edge between two vertices had 0.5 probability of existing, so as to differentiate the vertices for the growing procedure. We kept the initial network intentionally small because larger (initial) networks

create larger gaps between high degree and low degree vertices during the network growth. As a result, highly central vertices are fewer and more easily recognizable by any targeting strategy and are diminished quickly, leaving no time for the various strategies to produce different results.

The results are shown in figures 4, 5, 6 and 7, located at the appendix. The inverse mean shortest path length, the size of the largest component and the clustering coefficient are measured after each vertex deletion and shown in separate illustrations. For each of these parameters, five different data sets exist, corresponding to the five strategies under study (highest degree, betweenness, 1-neighbourhood edge degree, 2-neighbourhood edge degree and 2-neighbourhood edge degree with penalty). Their values at each deletion step are the average of 50 experiments with different networks of 1500 vertices. Of the monitoring parameters, the easiest to read is the size of the largest component and its transition is shown in magnification in fig. 6. It is easy to see the relation between the various strategies, as each one, having done a preliminary work, performs better or worse than the others during the transition.

4.1 Drilling into experimental results

The inverse mean shortest path length initially declines, meaning that distances inside the network begin to grow in general. At some point this trend is reversed because during the deletion process the connected components become quite small and the distances inside them are even smaller than in the initial network. Thus the inverse length continues to increase as connected components are cut into smaller pieces and this continues until they stop breaking up. At this point the inverse length is at its maximum value and almost all significant vertices are gone, as subsequent deletions leave the components at roughly the same size. From this maximum point on, the network continues to shrink with almost constant rate. The five strategies differ mainly in their ability to break the already small connected components to even smaller ones, leading to higher maximum points, as is shown in fig. 4 in appendix.

The size of the largest component is a more straightforward measure. After each deletion the size of a component is reduced by one and, at least initially, the deleted vertex is selected from the largest component. As more "central" vertices are deleted, critical paths collapse and the largest component breaks into smaller pieces. Figure 5 in appendix shows clearly that there is an early stage where the strategies built up tensions by deleting important vertices, a transition phase where very important vertices are

gone and each deletion breaks the largest component in small pieces, followed by a slow shrinking of the largest component. The transition phase is where the various strategies compete, and betweenness is the most successful in making the transition in fewer deletions. However, comparable to the degree strategy which performs poorly, our proposed strategies bridge the gap with betweenness by up to 23%, 29% and 55% for 1-neighbourhood, 2-neighbourhood and 2-neighbourhood with penalty edge degree respectively.

The clustering coefficient during the early stage is decreasing by orders of magnitude, meaning that the deleted vertices, tagged as central by the various strategies, contribute greatly to the global clustering coefficient (fig. 7 in appendix). During the transition phase it appears fluctuating due to the shrinking of the largest component and the increase in the number of components, and in the last phase it is completely wiped out as triangles do not practically exist. The betweenness stands out, since during the transition phase it creates a seesaw effect on the clustering coefficient, never destroying all triangles in the connected components. Although we have no solid evidence, we feel that this observation is the key to understanding the role and the success of betweenness, and to replicate its behaviour in other measures.

In order to understand why the proposed strategies work as they do, we focus on a high degree edge and examine its specific characteristics (fig. 1). Just by looking at the high degree edge alone, one can argue that it connects high degree vertices, therefore is important for the communication of $(k_w - 1)$ vertices (at the one endpoint) with another $(k_u - 1)$ vertices (at the other endpoint). So its deletion alone would probably effect many vertices and the distances between them. As for the highest degree endpoint (which will eventually be deleted), one must keep in mind that high degree vertices don't usually connect to other high degree vertices. On one hand, deleting high degree vertices is a successful enough strategy on its own (see degree centrality strategy), but with maximum edge degree we ensure that the high degree vertex to be deleted will be connected to the highest possible degree vertex (given its not a common phenomenon) and the deletion will effect a greater number of vertices. On the other hand, since high degree vertices don't connect often, this filter differentiates adequately otherwise equal vertices (i.e. when degree centrality is used).

The 2-neighbourhood edge degree strategy operates in an similar way. The same arguments as above are still valid here, i.e. a high degree edge connects more vertices than a low degree one. The main difference is that we are now talking about vertices that are part of the 2-neighbourhood of the one endpoint vertex which connect with the vertices of the 2-neighbourhood of

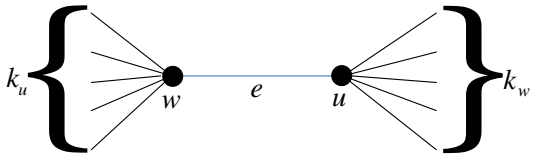


Figure 1: 1-neighbourhoods of two connected vertices

the other endpoint vertex. This may be more reliable than the 1-neighbourhood of the endpoints since there seems to exist a light disassortative mixing. This means that high degree vertices connect to lower degree, thus their influence dies out quickly as we move further from their center. By using 2-neighbourhoods we favor vertices that their influence two steps away is still strong. The downside of this strategy is its slightly larger computational load compared to the 1-neighbourhood edge degree, but this is still far from that of betweenness centrality. Furthermore it uses semi-local information for its computation, which, we estimate, should not be a problem in most practical uses.

The mechanism behind the alternative 2-neighbourhood edge degree strategy is somewhat different. Obviously the same arguments of the two previous strategies are still valid here. An instance that is handled differently is shown in fig. 2. Normally the 2-neighbourhood edge degree of this edge would be

$$\sum_{i \in \Gamma_2(w)} k_i \cdot \sum_{j \in \Gamma_2(u)} k_j$$

but since it participates in a triangle due to the common neighbour of both endpoint vertices, this edge degree is divided by 2. Thus edges that connect two "smaller" vertices in terms of 2-neighbourhoods can have larger edge degree and be selected instead. The situation is even worse if the edge participates in more triangles, as in fig. 3, for its edge degree would be even smaller. The edge degree can gradually increase, if vertices comprising the triangles become selected for the deletion process, and the triangles collapse.

This edge degree with penalty measures the size of two 2-neighbourhood connecting through one edge, as was the case in the previous 2-neighbourhood edge degree. But it also considers the importance of alternative paths

between these two 2-neighbourhoods. There is no doubt that selecting one endpoint of an edge participating in many triangles will destroy these triangles also, but vertices connecting same size 2-neighbourhoods with no triangles are more important to the whole network, and this is expressed by this measure and verified by our results.

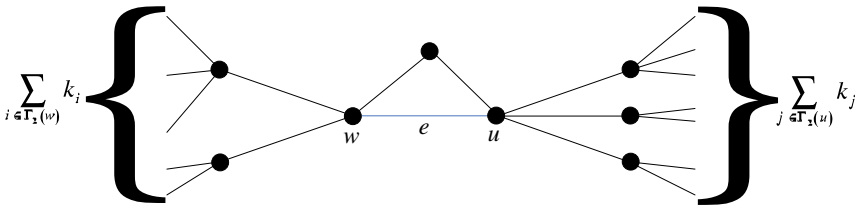


Figure 2: 2-neighbourhoods of two connected vertices

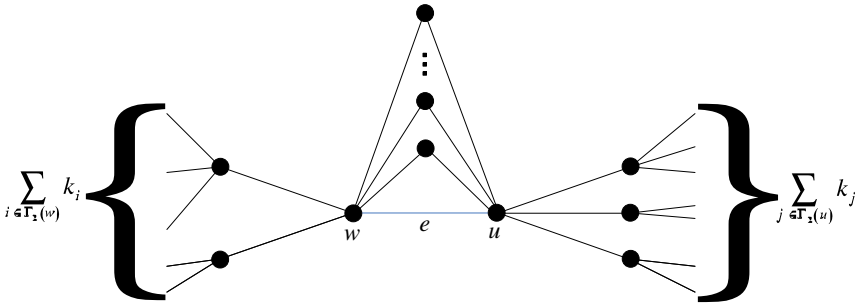


Figure 3: 2-neighbourhoods of two connected vertices with triangles

4.2 Algorithm Complexities

Theorem 1. *The worst case time complexities of the proposed strategies are* $O(m)$, $O(m\sqrt{n})$ *and* $O(mn)$ *for 1-neighbourhood, 2-neighbourhood and*

2-neighbourhood with penalty edge degrees, respectively. Furthermore the average case time complexity is $O(m)$ for all strategies.

Proof. The 1-neighbourhood edge degree just multiplies two integers, namely the vertex degrees of the edge endpoints, for all edges. The query of a vertice's degree is an $O(1)$ operation in the LEDA enviroment we are using [30], so the total cost of computing 1-neighbourhood edge degree for all edges is $O(m)$ in any case.

The 2-neighbourhood edge degree queries for each endpoint vertex, the vertex degree of all its neighbours and sums it, multipling the two endpoint sums, and does this for all edges. The iteration of all neighbouring vertices of a vertex has guaranteed asymptotic complexity on the number of actual neighbours. So in the worst case, this computation has $O(mk_{max})$ complexity, where k_{max} is the maximum degree in the network. Specifically for the Barabási-Albert network there exist an analytic solution [31] for the degree $k_i(t)$ of a vertex i at timestep t, as

$$k_i(t) = l\sqrt{\frac{t}{t_i}} \tag{4}$$

where l is the number of edges per new vertex and t_i is the timestep when vertex i was added to the network. After n timesteps the maximum degree in the network is

$$k_{max} = O\left(\sqrt{n}\right)$$

And so the worst case time complexity is $O(m\sqrt{n})$. But since the mean degree in the network is $\overline{k} = 2l$ (as can be easily seen), the average case time complexity is $O(m)$.

The third strategy is computed as above, but for one endpoint of the edge, we scan its neighbour's neighbour lists to find the other endpoint (indicating the existence of triangles). So its worst case complexity is $O\left(m{k_{max}}^2\right)$ and thus $O(mn)$. Similarly its average case complexity is $O(m)$. □

5 Conclusions

We have studied three novel strategies in network attack and compared them with two traditional approaches, degree and betweeness centrality, both with its own merits and flaws. These strategies have proven to be simple enough to implement, with low computational cost, and yet efficient compared to the best strategy. In addition to their value as attack strategies, they can help to shed light to the inner workings of a power law network. One of

the great difficulties in their study is our ignorance as to what measures are important to the behaviour of these networks. Our experiments link the degree-degree correlations among vertices with their centrality in the network. Furthermore, to the extent of our knowledge, it is the first time that the clustering effect is linked to the centrality of a vertex. Although we know this is responsible for the "denseness" of power law networks, its exact role remains unclear. Our third strategy indicates that it plays a major role in conjunction with other phenomena, such as the degree-degree correlations. It would be of interest to study several models of networks, other than the BA model, that show documented assortative or disassortative behavior and models that have known clustering coefficient distributions, in order to explore further these effects of our strategies. Furthermore, it is the subject of future research to determine whether the utilization of other network structures, similar to the triangles we are using in this study, will help bridge the gap between local strategies and global ones, as is betweenness. This development will not only help us to study larger networks but will also reveal the role of individuals in such a vast network.

References

[1] Baldi P, Frasconi P, and Smyth P. *Modeling the Internet and the Web: Probabilistic Methods and Algorithms*. Wiley, 2003.

[2] Jeong H, Tombor B, Albert R, Oltvai N, and Barabási A. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.

[3] Liljeros F, Edling R, Amaral N, Stanley E, and Aberg Y. The web of human sexual contacts. *Nature*, 411:907–908, 2001.

[4] Mariolis P. Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 56:425–439, 1975.

[5] Pimm L. *Food Webs*. University of Chicago Press, 2nd edition, 2002.

[6] Podani J, Oltvai N, Jeong H, Tombor B, Barabási A, and Szathmáry E. Comparable system-level organization of archaea and eukaryotes. *Nature Genetics*, 29:54–56, 2001.

[7] Jones J and Handcock M. An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, 270(1520):1123–1128, June 2003.

[8] Albert R, Jeong H, and Barabási A. Diameter of the world wide web. *Nature*, (401), 1999.

[9] Chen Q, Chang H, Govindan R, Jamin S, Shenker S, and Willinge W. The origin of power laws in internet topologies revisited. In *IEEE Infocom 2002*, 2002.

[10] Faloutsos M, Faloutsos P, and Faloutsos C. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[11] Huberman A. *The Laws of the Web: Patterns in the Ecology of Information*. The MIT Press, 2001.

[12] Lawrence S and Giles L. Searching the world wide web. *Science*, 280:98–100, 1998.

[13] Ravasz E and Barabási A. Hierarchical organization in complex networks. *Phys. Rev. E*, 67(026112), 2003.

[14] Watts J. A simple model of global cascades on random networks. In *Proc. Natl. Acad. Sci. USA*, volume 99, pages 5766–5771, 2002.

[15] Erdős P and Rényi P. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–291, 1959.

[16] Barabási A and Albert R. Emergence of scaling in random networks. *Science*, 286:509–512, October 1999.

[17] Watts J and Strogatz H. Collective dynamics of 'small-world' networks. *Nature*, pages 393–393, 1998.

[18] Watts J. *Small Worlds*. Princeton Uni. Press, 1999.

[19] Newman J. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[20] Bornholdt S and Schuster G, editors. *Handbook of Graphs and Networks*. Wiley-VCH, 2002.

[21] Brandes U. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[22] Newman J. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64(016132), 2001.

[23] Holme P, Kim J, Yoon No, and Han K. Attack vulnerability of complex networks. *Phys. Rev. E*, 65(056109), 2002.

[24] Cohen R, Erez K, Ben-Avraham D, and Havlin S. Resilience of the internet to random breakdowns. *Phys. Rev. Let.*, 85:4626–4628, 2000.

[25] Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, and Wiener J. Graph structure in the web. *Computer Networks*, 33(309), 2000.

[26] Callaway S, Newman J, Strogatz H, and Watts J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85:5468–5471, 2000.

[27] Cohen R, Erez K, Ben-Avraham D, and Havlin S. Breakdown of the internet under intentional attack. *Phys. Rev. Let.*, 86:3682–3685, 2001.

[28] Krapivsky L and Redner S. Organization of growing random networks. *Phys. Rev. E*, 63(066123), 2001.

[29] Zhuang-Xiong H, Xin-Ran W, and Han Z. Pair correlations in scale-free networks. *Chinese Physics*, 13(3):273–278, 2004.

[30] Mehlhorn K and Nahe S. *LEDA : A Platform for Combinatorial and Geometric Computing*. Cambridge University Press, Cambridge, 2000.

[31] Barabási A, Albert R, and Jeong H. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.

Appendix

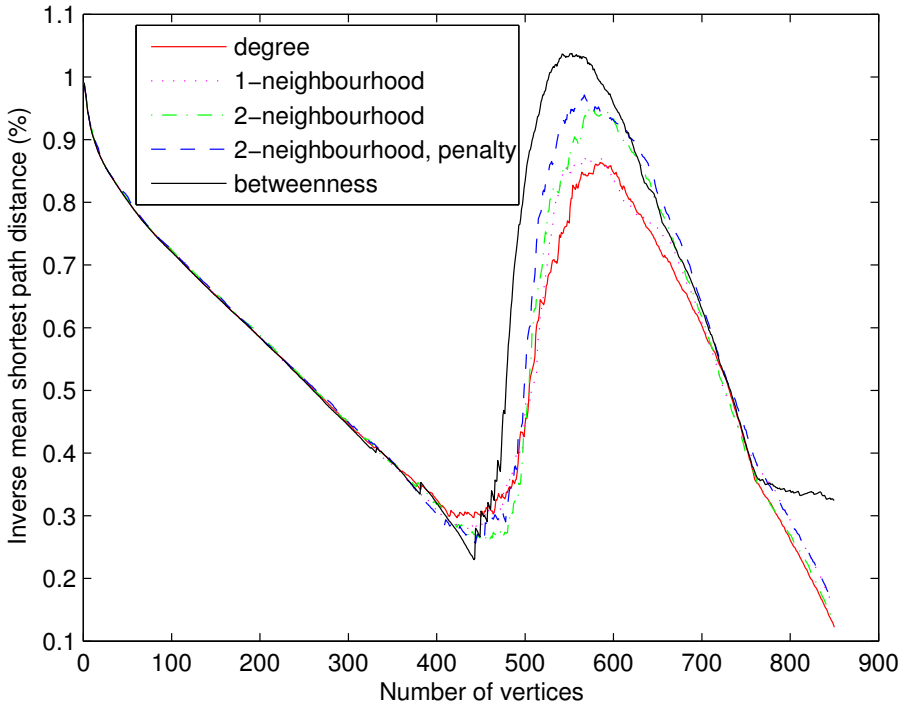


Figure 4: Inverse mean shortest path length shown as percent of the initial length, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.

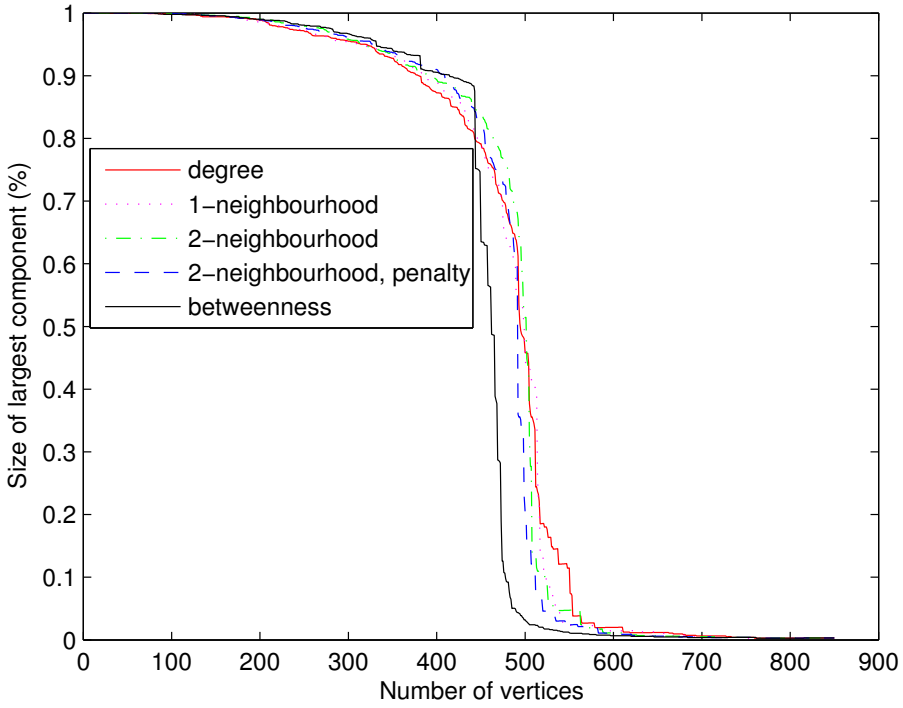


Figure 5: Size of largest connected component shown as percent of the initial size, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.

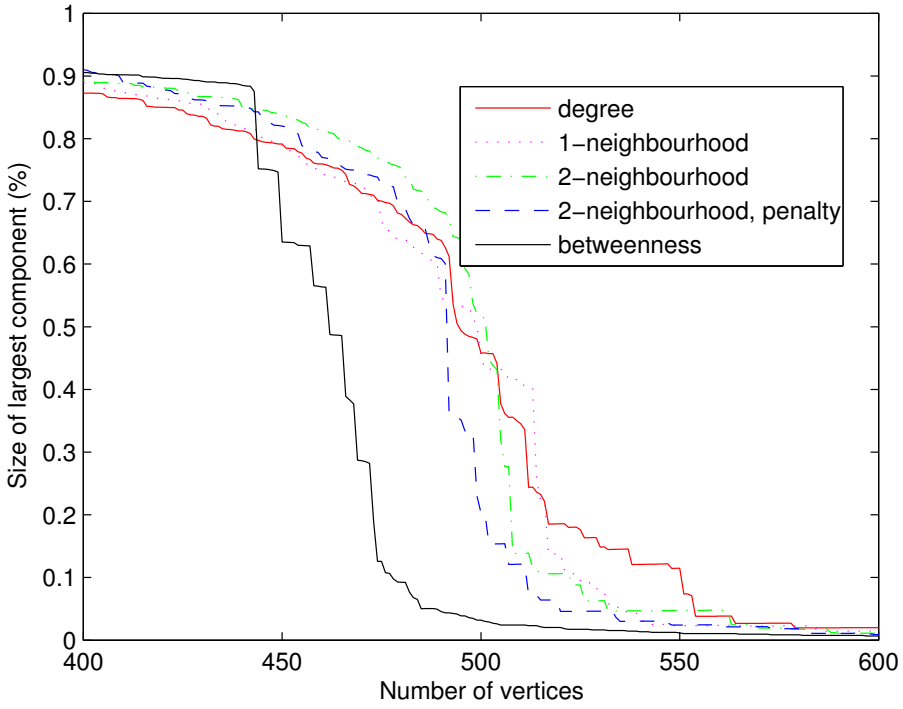


Figure 6: Size of largest connected component shown as percent of the initial size, as vertices are sequentially deleted. Detail of the transition.

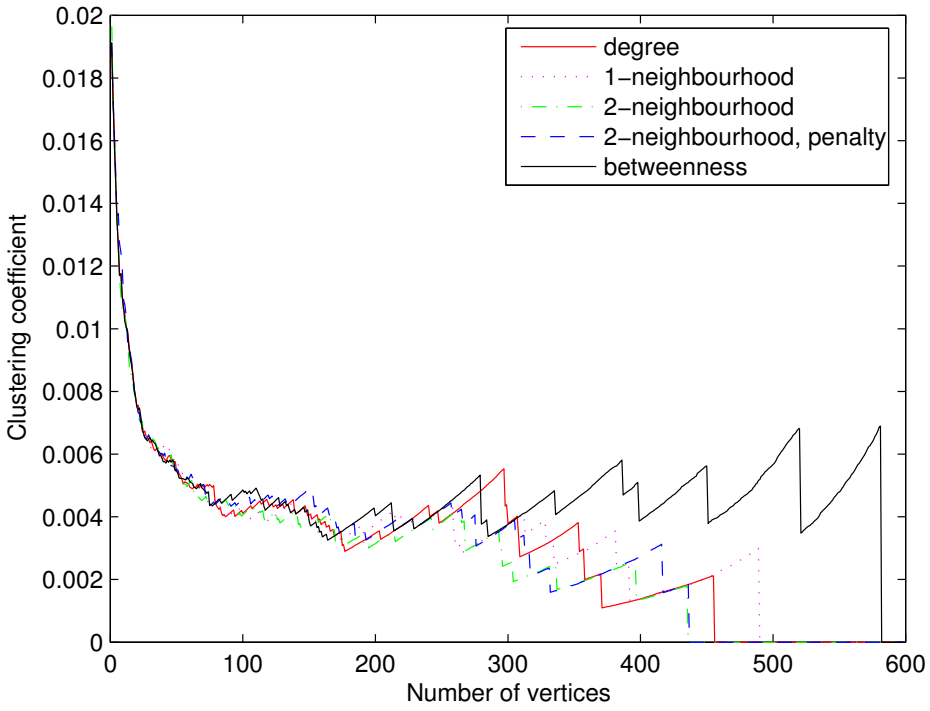


Figure 7: Clustering coefficient of network, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.

Is selection optimal for scale-free small worlds?

Zs. Palotai[1], Cs. Farkas[2], A. Lőrincz[1*]

August 29, 2005

[1]Dept. of Information Systems, Eötvös Loránd University, Budapest, Hungary
[2]Dept. of Computer Science and Engr., University of South Carolina, SC, USA

Abstract

The 'No Free Lunch Theorem' claims that for the set of all problems no algorithm performs better than random search and thus, selection can be advantageous only on a limited set of problems. In this paper we investigate how the topological structure of the environment influences algorithmic efficiency. We study the performances of algorithms, using selective learning, reinforcement learning, and their combinations, in random, scale-free, and scale-free small world (SFSW) environments. The learning problem is to search for novel, not-yet-found information. We ran our experiments on a large news site and on its downloaded portion. Controlled experiments were performed on this downloaded portion: we modified the topology, but preserved the publication time of the news. Our empirical results show that the selective learning is the most efficient in SFSW topology. In non-small world topologies, however, the combination of the selective and reinforcement learning algorithms performs the best.

Keywords: scale-free small world, selection, NFL theorem, evolution, Internet
Running head: Selection in SFSW

*Corresponding author: Eötvös Loránd University, Department of Information Systems, Pázmány Péter sétány 1/c, Budapest, H1117, Hungary. Email: andras.lorincz@elte.hu, tel.: +36 1 209 0555 / 8473, fax: + 36 1 381 2140

1 Introduction

According to the 'No Free Lunch Theorem' (NFL Theorem) [1] there is no performance difference between optimization or search algorithms if we test the algorithms on every possible problem. Therefore, it seems that improved performances of specific algorithms, such as selection, are consequences of specific problem properties. Finding these properties will aid the development of optimized solutions. Recent research shows that evolving structures, both natural and artificial (like the Web), exhibit scale-free or scale-free small world properties [2, 3]. In [4] the authors show that a particular evolutionary algorithm has better performance in artificial scale-free environment than in lattice, small world, or random environments. That is, the structure of the environment has an impact on the efficiency of the algorithms.

We have designed a controllable experiment and compared the performances of different algorithms on different structures. We aimed to keep the complexity of experiments and the number of designer specifiable parameters minimal. We considered the structure of the environment, the algorithms, and the fitness values. Our choice for the structure was the World Wide Web (WWW). WWW is considered the largest source of rapidly changing data. The WWW has a scale-free small world (SFSW) structure [2, 5].

The everyday usage of the Web is the search for novel information. Therefore, we have a natural reward function; the number of novel documents that the algorithms can find. We consider this property as one of the most important components of our work: *We did not specify the temporal and structural details of the reward system.*

We ran controlled experiments on a time stamped and downloaded portion of a large WWW news site. This allowed us to preserve the temporal structure of the rewards, but also supported the modification of the underlying connectivity structure, i.e., how the novel information can be found.

Our agents were Web crawlers, or foragers. Crawlers travel from link to link foraging new, not-yet-seen information. Our agents used the simplest possible algorithms. Selective learning concerned the selection and memorization of good links. We have also used a simple version of reinforcement learning (RL). For a review on RL see, e.g., [6]. RL was used alone, and was also combined with selective learning. Our choice of RL is motivated by its structural and algorithmic simplicity and that RL is concerned with the optimization of the expected value of long-term cumulated profit.

For a recent review on evolutionary computing, see [7]. For reviews on related evolutionary theories and the dynamics of self-modifying systems see [8, 9] and [10, 11], respectively. Similar concepts have been studied in other evolutionary systems where organisms compete for space and resources and cooperate through direct interaction (see, e.g., [12] and references therein.)

Hybrid algorithms have a long history. A well-known example is the TD-Gammon program of Tesauro [13]. The author applied MLP function approximators for value estimation in RL. Reinforcement learning has also been used in concurrent learning problems like ours: robots had to learn to forage together

via direct interaction [14]. Another combinations of the algorithms concerns evolutionary learning embedded into the framework of RL to improve decision making [15, 16, 17, 18].

It is important to note, that communication and competition among our foragers are indirect. Only the first submitter of a document may receive positive reinforcement and this is the only interaction among the crawlers. Our work is different from other studies using combinations of genetic, evolutionary, function approximation, and reinforcement learning algorithms, in that i) it does not require explicit fitness function, ii) we do not have control over the original environment, iii) we can change the environment in a reproducible fashion, iv) collaborating individuals use value estimation under 'evolutionary pressure', and v) individuals work without direct interaction with each other. The crawler system is a self-assembling system, which is made of adaptive components, and the communication between components is kept as little as possible.

The paper is organized as follows. We review the related web crawler tools, including those [19, 20, 21] that our work is based upon, in Section 2. We describe our algorithms and the forager architecture in Section 3. This section contains the necessary algorithmic details imposed by the task, the search of the Web. We present our experiments on the Web and the controlled simulations in Section 4. Discussions can be found in Section 5. Conclusions are drawn in Section 6.

2 Related work

There are important problems that have been studied in the context of crawlers. Angkawattanawit, Rungsawang [28], and Menczer [29] study topic specific crawlers. Risvik et al. [30] address research issues related to the exponential growth of the Web. Cho and Gracia-Molina [31], Menczer [29] and Edwards et. al [32] study the problem of different refresh rates of URLs (possibly as high as hourly or as low as yearly).

An introduction to and a broad overview of topic specific crawlers are provided in [28]. They propose to learn starting URLs, topic keywords and URL ordering through consecutive crawling attempts. They show that the learning of starting URLs and the use of consecutive crawling attempts can increase the efficiency of the crawlers. The used heuristic is similar to the weblog algorithm [21], which also finds good starting URLs and periodically restarts the crawling from the newly learned ones. The main limitation of this work is that it is incapable of addressing the freshness (i.e., modification) of already visited Web pages.

Menczer [29] describes some disadvantages of current Web search engines on the dynamic Web, e.g., the low ratio of fresh or relevant documents. He proposes to complement the search engines with intelligent crawlers, or web mining agents to overcome those disadvantages. He introduces the InfoSpider architecture that uses genetic algorithm and reinforcement learning, also describes the MySpider implementation of it, which starts from the 100 top pages of AltaVista. Our

weblog algorithm uses local selection for finding good starting URLs for searches, thus not depending on any search engines. Dependence on a search engine can be a suffer limitation of most existing search agents, like MySpiders. Note however, that it is an easy matter to combine the present algorithm with URLs offered by search engines.

Risvik and Michelsen [30] overview different dimensions of web dynamics and show the arising problems in a search engine model. The main part of the paper focuses on the problems that crawlers need to overcome on the dynamic Web. As a possible solution the authors propose a heterogenous crawling architecture. The main limitation of their crawling architecture is that they must divide the web to be crawled into distinct portions manually before the crawling starts. A weblog like distributed algorithm – as suggested here – my be used in that architecture to overcome this limitation.

Cho and Garcia-Molina [31] define mathematically the freshness and age of documents of search engines. They propose the Poisson process as a model for page refreshment. The authors also propose various refresh policies and study their effectiveness both theoretically and on real data. They present the optimal refresh policies for their freshness and age metrics under the Poisson page refresh model. The authors show that these policies are superior to others on real data, too. Although they show that in their database more than 20 percent of the documents are changed each day, they disclosed these documents from their studies. Their crawler visited the documents once each day for 5 months, thus can not measure the exact change rate of those documents. While in our work we definitely concentrate on these frequently changing documents.

3 Forager architecture

There are two different kinds of agents: the foragers and the reinforcing agent (RA). The fleet of foragers crawls the web and sends the URLs of the selected documents to the reinforcing agent. The RA determines which forager should work for the RA and how long a forager should work. The RA sends reinforcements to the foragers based on the received URLs.

We employ a fleet of foragers to study the competition among individual foragers. A forager has simple and limited capabilities, like a stack for a limited number of starting URLs and a simple, content based URL ordering. The foragers compete with each other for finding the most relevant documents. In this way they efficiently and quickly collect new relevant documents without direct interaction.

At first we present the basic algorithms, followed by the algorithms for the reinforcing agent and the foragers.

3.1 Algorithms

Our constraints on finding the minimal set of algorithms were as follows: The algorithms should (i) allow the identification of unimportant parameters, (ii)

support the specialization of the individuals (the foragers), (iii) allow the joining of evolutionary learning and individual learning, (iv) minimize communication as much as possible. We shall return to these points in Section 5.

Weblog algorithm and starting URL selection

A forager periodically restarts from a URL randomly selected from the list of starting URLs. The sequence of visited URLs between two restarts forms a path. The starting URL list is formed from the 10 first URLs of the weblog. In the weblog there are 100 URLs with their associated weblog values in descending order. The weblog value of a URL estimates the expected sum of rewards during a path after visiting that URL. The weblog update algorithm modifies the weblog before a new path is started. The weblog value of a URL already in the weblog is modified toward the sum of rewards ($sumR$) in the remaining part of the path after that URL:

$$newValue = (1 - \beta)\, oldValue + \beta\, sumR,$$

where β was set to 0.3. A new URL has the value of actual sum of rewards in the remaining part of the path. If a URL has a high weblog value it means that around that URL there are many relevant documents. Therefore it may worth it to start a search from that URL.

Without the weblog update algorithm the weblog and thus the starting URL list remains the same throughout the searches. The weblog algorithm is a very simple version of evolutionary algorithms. Here, evolution may occur at two different levels: the list of URLs of the forager is evolving by the reordering of the weblog. Also, a forager may multiply, and its weblog, or part of it may spread through inheritance. This way, the weblog algorithm incorporates the basic features of evolutionary algorithms. This simple form shall be satisfactory for our purposes.

Reinforcement Learning based URL ordering

A forager can modify its URL ordering based on the received reinforcements of the sent URLs. The (immediate) profit is the difference of received rewards and penalties at any given step. Immediate profit is a myopic characterization of a step to a URL. Foragers have an adaptive continuous value estimator and follow the *policy* that maximizes the expected long term cumulated profit (LTP) instead of the immediate profit. Such estimators can be easily realized in neural systems [6, 22, 23]. Policy and profit estimation are interlinked concepts: profit estimation determines the policy, whereas policy influences choices and, in turn, the expected LTP. (For a review, see [6].) Here, choices are based on the greedy LTP policy: The forager visits the URL, which belongs to the *frontier* (the list of linked but not yet visited URLs, see later) and has the highest estimated LTP.

In the particular simulation each forager has a $k(= 50)$ dimensional probabilistic term-frequency inverse document-frequency (PrTFIDF) text classifier

[24], generated on a previously downloaded portion of the Geocities database. Fifty clusters were created by Boley's clustering algorithm [25] from the downloaded documents. The PrTFIDF classifiers were trained on these clusters plus an additional one, the $(k+1)^{th}$, representing general texts from the internet. The PrTFIDF outputs were non-linearly mapped to the interval [-1,+1] by a hyperbolic-tangent function. The classifier was applied to reduce the texts to a small dimensional representation. The output vector of the classifier for the page of URL A is $\mathbf{state}(\mathbf{A}) = (state(A)_1, \ldots, state(A)_k)$. (The $(k+1)^{th}$ output was dismissed.) This output vector is stored for each URL.

A linear function approximator is used for LTP estimation. It encompasses k parameters, the *weight vector* $\mathbf{weight} = (weight_1, \ldots, weight_k)$. The LTP of document of URL A is estimated as the scalar product of $\mathbf{state}(\mathbf{A})$ and $\mathbf{weight}$: $value(A) = \sum_{i=1}^{k} weight_i\, state(A)_i$. During URL ordering the URL with highest LTP estimation is selected. (For more details, see, [21].)

The weight vector of each forager is tuned by temporal difference learning (TD-learning) [26, 22, 23]. Let us denote the current URL by URL_n, the next URL to be visited by URL_{n+1}, the output of the classifier for URL_j by $\mathbf{state}(\mathbf{URL_j})$ and the estimated LTP of a URL URL_j by $value(URL_j) = \sum_{i=1}^{k} wegiht_i\, state(URL_j)_i$. Assume that leaving URL_n to URL_{n+1} the immediate profit is r_{n+1}. Our estimation is perfect if $value(URL_n) = value(URL_{n+1}) + r_{n+1}$. Future profits are typically discounted in such estimations as $value(URL_n) = \gamma value(URL_{n+1}) + r_{n+1}$, where $0 < \gamma < 1$. The error of value estimation is

$$\delta(n, n+1) = r_{n+1} + \gamma value(URL_{n+1}) - value(URL_n).$$

We used throughout the simulations $\gamma = 0.9$. For each step $URL_n \rightarrow URL_{n+1}$ the weights of the value function were tuned to decrease the error of value estimation based on the received immediate profit r_{n+1}. The $\delta(n, n+1)$ estimation error was used to correct the parameters. The i^{th} component of the weight vector, $weight_i$, was corrected by

$$\Delta weight_i = \alpha\ \delta(n, n+1)\ state(URL_n)_i$$

with $\alpha = 0.1$ and $i = 1, \ldots, k$. These modified weights would improve value estimation in stationary and observable environments (see, e.g, [6] and references therein), but were also found efficient in large Web environments [21].

Without the reinforcement learning based URL ordering update algorithm the weight vector remains the same throughout the search.

Document relevancy

A document or page is possibly relevant for a forager if it is not older than 24 hours and the forager has not marked it previously. The selected documents are sent to the RA for further evaluation.

Multiplication of a forager

During multiplication the weblog is randomly divided into two equal sized parts (one for the original and one for the new forager). The parameters of the URL ordering algorithm (the weight vector of the value estimation) are either copied or new random parameters are generated. If the forager has a URL ordering update algorithm then the parameters are copied. If the forager does not have any URL ordering update algorithm then new random parameters are generated.

3.2 Reinforcing agent

A reinforcing agent controls the 'life' of foragers. It can start, stop, multiply or delete foragers. RA receives the URLs of documents selected by the foragers, and responds with reinforcements for the received URLs. The response is 100 (a.u.) for a relevant document and -1 (a.u.) for a not relevant document. A document is relevant if it is not yet seen by the reinforcing agent and it is not older than 24 hours. The reinforcing agent maintains the score of each forager working for it. Initially each forager has 100 (a.u.) score. When a forager sends a URL to the RA, the forager's score is decreased by 0.05. After each relevant page sent by the forager, the forager's score is increased by 1.

When the forager's score reaches 200 and the number of foragers is smaller than 16 then the forager is multiplied. That is a new forager is created with the same algorithms as the original one has, but with slightly different parameters. When the forager's score goes below 0 and the number of foragers is larger than 2 then the forager is deleted. Note that a forager can be multiplied or deleted immediately after it has been stopped by the RA and before the next forager is activated.

Foragers on the same computer are working in time slices one after each other. Each forager works for some amount of time determined by the RA. Then the RA stops that forager and starts the next one selected by the RA.

3.3 Foragers

A forager is initialized with parameters defining the URL ordering, and either with a weblog or with a seed of URLs. After its initialization a forager crawls in search paths, that is after a given number of steps the search restarts and the steps between two restarts form a path. During each path the forager takes 100 steps, i.e., selects the next URL to be visited with a URL ordering algorithm. At the beginning of a path a URL is selected randomly from the starting URL list. This list is formed from the 10 first URLs of the weblog. The weblog contains the possibly good starting URLs with their associated weblog values in descending order. The weblog algorithm modifies the weblog and so thus the starting URL list before a new path is started. When a forager is restarted by the RA, after the RA has stopped it, the forager continues from the internal state in which it was stopped.

The URL ordering algorithm selects a URL to be the next step from the frontier URL set. The selected URL is removed from the frontier and added to the visited URL set to avoid loops. After downloading the pages, only those URLs (linked from the visited URL) are added to the frontier which are not in the visited set.

In each step the forager downloads the page of the selected URL and all of the pages linked from the page of selected URL. It sends the URLs of the possibly relevant pages to the reinforcing agent. The forager receives reinforcements on any previously sent but not yet reinforced URLs and calls the URL ordering update algorithm with the received reinforcements.

4 Experiments

We conducted an 18 day long experiment on the Web to gather realistic data. We used the gathered data in simulations to compare the weblog update (Section 3.1) and reinforcement learning algorithms (Section 3.1). In the Web experiment we used a fleet of foragers using combination of reinforcement learning and weblog update algorithms to eliminate possible biases on the gathered data. First we describe the experiment on the Web then the simulations. We analyze our results in the next section.

4.1 Data collection on the Web

We ran the experiment on the Web on a single personal computer with Celeron 1000 MHz processor and 512 MB RAM. We implemented the forager architecture (described in Section 3) in Java programming language.

In this experiment a fixed number of foragers were competing with each other to collect news at the CNN web site. The foragers were running in equal time intervals in a predefined order. Each forager had a 3 minute time interval and after that interval the forager was allowed to finish the step started before the end of the time interval. We deployed 8 foragers using the weblog update and the reinforcement learning based URL ordering update algorithms (8 WR foragers). We also deployed 8 other foragers using the weblog update algorithm but without reinforcement learning (8 WL foragers). The predefined order of foragers was the following: 8 WR foragers were followed by the 8 WL foragers.

We investigated the link structure of the gathered Web pages. As it is shown in Fig. 1 the links have a power-law distribution ($P(k) = k^{\gamma}$) with $\gamma = -1.3$ for outgoing links and $\gamma = -2.57$ for incoming links. That is the link structure has the scale-free property. The clustering coefficient [27] of the link structure is 0.02 and the diameter of the graph is 7.2893. We applied two different random permutations to the origin and to the endpoint of the links, keeping the edge distribution unchanged but randomly rewiring the links. The new graph had 0.003 clustering coefficient and 8.2163 diameter. That is the clustering coefficient was smaller than the original value by an order of magnitude, but the diameter is

almost the same. Therefore we can conclude that the links of gathered pages form *scale-free small world* structure.

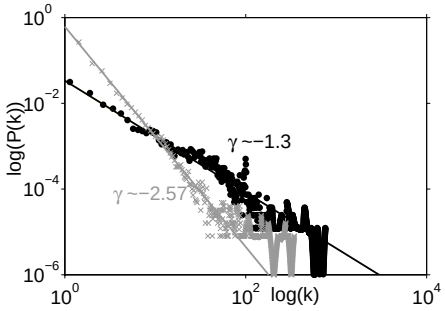


Figure 1: **Scale-free property of the Internet domain**. Log-log scale distribution of the number of (incoming and outgoing) links of all URLs found during the time course of investigation. Horizontal axis: number of edges ($\log k$). Vertical axis: relative frequency of number of edges at different URLs ($\log P(k)$). Dots and dark line correspond to outgoing links, crosses and gray line correspond to incoming links.

The data storage for simulation is a central issue in our experiments. Pages are stored with 2 indices (and time stamps). One index is the URL index, the other is the page index. Multiple pages can have the same URL index if they were downloaded from the same URL. The page index uniquely identifies a page content and the URL from where the page was downloaded. For any foragers, at each page download we stored the followings (with a time stamp containing the time of page download):

1. if the page is relevant according to the RA then store 'relevant'
2. if the page is from a new URL then store the new URL with a new URL index and the page's state vector with a new page index
3. if the content of the page is changed since the last download then store the page's state vector with a new page index but keep the URL index
4. in both previous cases store the links of the page as links to page indices of the linked pages
   (a) if a linked page is from a new URL then store the new URL with a new URL index and the linked page's state vector with a new page index
   (b) if the content of the linked page is changed since the last check then store the page's state vector with a new page index but same URL index

4.2 Simulations

For the simulations we implemented the forager architecture in Matlab. The foragers were simulated as if they were running on one computer as described in the previous section.

Simulation specification

During simulations we used the Web pages that we gathered previously to generate different environments (note that the links of pages point to local pages (not to pages on the Web) since a link was stored as a link to a local page index):

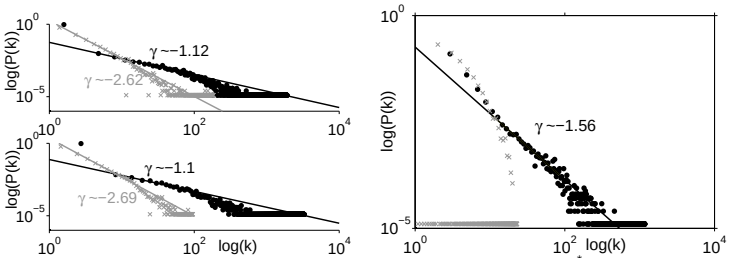
- Simulated documents had the same state vector representation for URL ordering as the real pages had
- Simulated relevant documents were the same as the relevant documents on the Web
- Pages and links appeared at the same (relative) time when they were found in the Web experiment - using the new URL indices and their time stamps
- Pages and links are refreshed or changed at the same relative time as the changes were detected in the Web experiment – using the new page indices for existing URL indices and their time stamps
- Simulated time of a page download was the average download time of a real page during the Web experiment.

We generated 4 different environments for the simulations:

1. **SFSW**: each simulated page had exactly the same links as the original page had on the Web (a simulated page linked those simulated pages, page indices of those pages, which were linked by the original Web page).
2. **SF1**: in each second the new simulated pages had the same number of links as the original pages on the Web. A new simulated page linked to simulated pages selected by the preferential attachment algorithm from the existing simulated pages.
3. **SF2**: the previous algorithm applied for the **SF1** environment.
4. **SFRandom**: similar to the **SF1**, but the linked simulated pages are selected from a uniform distribution of the pages.

The SFSW environment has exactly the same scale-free and small world properties as the web environment downloaded by the web foragers. The SF1 and SF2 environments have 10 times smaller clustering coefficient than the SFSW environment has. These environments have scale-free degree distributions, although those are slightly different from the web environment (see Fig. 2(a)).

The SFRandom environment, also, has 10 times smaller clustering coefficient then SFSW. The SFRandom environment has scale-free outgoing link degree distribution. But because of the uniform selection of linked documents the incoming link degree distribution is exponential (see Fig. 2(b)). With the above given constraints this environment is the most random in the sense that all of the free parameters (linked documents) were selected from the uniform random distribution.



(a) SF1,SF2 (b) SFRandom

Figure 2: **Degree distribution of the environments.**
Dots and dark line correspond to outgoing link distribution. X-s and gray line correspond to incoming link distribution. (a) Upper: degree distributions of SF1 environment. Lower: degree distributions of SF2 environment. (b) : degree distributions of SFRandom environment.

We conducted simulations with four different kinds of foragers in each environment:

1. **WR foragers** used both the weblog update and the reinforcement learning based URL ordering update algorithms.

2. **WL foragers** used only the weblog update algorithm without URL ordering update. Each WL forager had a different weight vector for URL value estimation – during multiplication the new forager got a new random weight vector.

3. **RL foragers** used only the reinforcement learning based URL ordering update algorithm without the weblog update algorithm. RL foragers had the same weblog with the first 10 URLs of the gathered pages – that is the starting URL of the Web experiment and the first 9 visited URLs during that experiment.

4. **Fix foragers** did not use the weblog update and the reinforcement learning based URL ordering update algorithms. These foragers had fixed starting URLs and fixed weight vectors, but the latter was different for each Fixed forager.

In each case, initially there were 2 foragers and they were allowed to multiply until reaching the population of 16 foragers. The simulation for each type of foragers were repeated 3 times with different initial weight vectors for each forager. The variance of the results show that there is only a small difference between simulations using the same kind of foragers, even if the foragers were started with different random weight vectors in each simulation.

Simulation measurements

The first thing that we should note concerns the efficiency as a function of the number of crawlers. On a single computer, and under the time sharing method we applied, and without direct competition between the different crawlers, we found that bipartition gives rise to a transient decrease of the efficiency of the new crawlers, but it quickly recovers. Within the limits of the number of crawlers that we studied (between 2 and 24), performance of the fleet is a slowly increasing function of the number of crawlers. We fixed the number of the crawlers and this slow dependence did not enter our considerations. Table 1 shows the investigated parameters during simulations.

Table 1: Investigated parameters

downloaded	number of downloaded documents
sent	number of documents sent to the RA
relevant	number of found relevant documents
found URLs	number of found URLs
download efficiency	ratio of relevant to downloaded documents in 3 hour time window throughout the simulation.
sent efficiency	ratio of relevant to sent documents in 3 hour time window throughout the simulation.
exploration	ratio of found URLs to downloaded at the end of the simulation
freshness	ratio of the number of current found relevant documents and the number of all found relevant documents [31]. A stored document is current, up-to-date, if its content is exactly the same as the content of the corresponding URL in the environment.
age	A stored current document has 0 age, the age of an obsolete page is the time since the last refresh of the page on the Web [31].

Parameter 'download efficiency' is relevant for the site where the foragers should be deployed to gather the new information. Parameter 'sent efficiency' is relevant for the RA. Note that during simulations we are able to immediately and precisely calculate freshness and age values. In a real Web experiment this is impossible, because of the time needed to download and compare the contents of all of the real Web pages to the stored ones.

5 Discussion

We observed, that the efficiency of the algorithms depends strongly on the weight vectors. As we have mentioned above, the number of foragers had slight effects on the efficiency. This observation is supported by the fact that upon bipartition the weight vectors of the descendant foragers are similar, causing the descendant foragers to follow similar paths and to spoil the performance of the other. This is the reason that efficiency shows a transient decrease, but as a result of adaptation it quickly disappears.

The dependence of the efficiency on the number of starting points is not too much different. Two foragers could use 20 different starting points, whereas 16 foragers could use 160 different starting points, but the efficiency was only slightly influenced by this order of magnitude difference. That is, the number of starting points and the number of foragers are less important than the weight vectors in our simulations indicating that 20 starting points or so, which can adapt, can efficiently shatter the structure. Note that other parameters, e.g., larger number of crawlers, more than one computer can change this picture. The weak dependence, however, is most advantageous for our purpose namely to be confident that we can study structural dependencies.

The measured parameter values are presented in Fig. 3. The figure contains the values for each measured parameter for each type of forager in each type of environment. From the subfigures we can conclude the followings:

It can be seen in the top left and top middle subfigures that freshness and age values of different foragers are changing in the SFSW environment (marker x) while in the other 3 environments the values are almost the same for the different type of foragers. RL and Fix foragers can get trapped in clustered environments, can not easily escape (find outgoing links leaving the cluster) from clusters containing less relevant documents. The weblog algorithm provides a way to go to the found best clusters and in this way to escape from the worse clusters. This can be the reason why the RL and Fix foragers in the SFSW environment performs worse than WL and WR foragers.

In the top right subfigure it can be seen that finding new documents is the hardest in the SFRandom environment for all foragers. In this environment the pages has exponential incoming link degree distribution which means that there are relatively more small degree pages than in the other 3 scale-free environments. It is harder to get to pages which links to many pages. Therefore it is harder for the foragers to find not yet seen documents.

It can be seen in the middle left and middle subfigures that finding relevant or possibly relevant documents is the easiest in the SFSW environment for all foragers. This environment is more clustered than the other three environments. When a forager finds a relevant document in a cluster then it finds other relevant documents in that cluster while it can not escape from the cluster. The other 3 environments are less clustered. Foragers can get out from clusters easier by following the links and therefore have to forage the entire environment for new relevant documents.

In the bottom left subfigure it can be seen that the download efficiency is

the best in the SFSW environment for all foragers. This is because of the high number of found relevant documents compared to the other three environments.

In the bottom middle subfigure it can be seen that the WR forager's sent efficiency is the same in all environments. Although these foragers found less relevant documents in not SFSW environments but also sent back less documents to the Reinforcing Agent. The other three foragers sent more less documents to Reinforcing Agent in the not SFSW environments, therefore their sent efficiencies are the worst in the SFSW environment. Although these foragers also found the most relevant documents in the SFSW environment, compared to the other 3 environments.

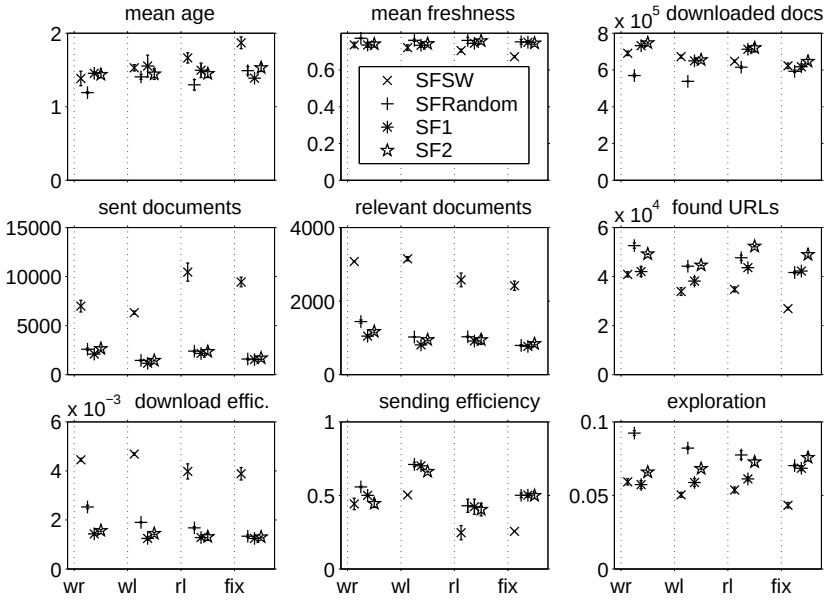


Figure 3: **Measured parameter values.**
The subfigures show the investigated parameters. Each subfigure contains the parameter values for the four type of foragers in four columns as shown below the bottom subfigures. The 4 different markers correspond to the measured parameter values in the 4 environments as shown in the top middle figure legend. On each marker an error bar shows the standard deviation of the corresponding parameter values for the 3 simulations. Mean age is in hours in the upper left subfigure.

It can be seen in the middle right subfigure that finding new URLs is the hardest in the SFSW environment because of its clustered nature. The foragers checks the same URLs for changed documents in the clusters therefore they can collect many new relevant documents. In the other three environments

foragers do not get trapped as much in the clusters and they search the whole environment continuously.

Now, consider Table 2, which contains some of the data of Fig. 3. The number of sent relevant documents is somewhat larger for the WL foragers than for WR foragers in SFSW environments (3149 and 3075, respectively, i.e., the difference is about 2.3%). This slight difference changes sign and becomes much more pronounced in all other environments. For example, in the SFRandom environment the numbers are 1029 and 1441, i.e., the difference is about 28%, but in the opposite direction. It is important to note that, if these foragers compete with each other, then the slight 2.3% difference or the larger 28% difference both enter the argument of an exponential because of bipartitions. Thus, even slight differences can become large, and may give rise to overwhelming population differences. Such competitive runs are under way. The issue becomes more pronounced for real situations, where time is not shared on a single computer and all foragers may search for food at all times.

Our results and the results of Annunziato et al. [4] can not be compared directly. The most obvious reason is that their investigations were restricted to SF, SW and random structures, but they did not study the SFSW structure, which plays a central role in our work. Artificial studies are, however, desirable, because in such examples one can finely gauge the different components and may find the necessary and sufficient conditions of our findings.

Table 2: Quantitative results for algorithm–structure pairs

	Weblog		Weblog and RL		RL	
	SFSW	SFR	SFSW	SFR	SFSW	SFR
No of sent docs	6313	1443	6985	2585	10455	2396
No of relevant docs	3149	1023	3075	1441	2575	1029
Sending efficiency	0.5035	0.7106	0.4425	0.5574	0.2463	0.4299
No of found URLs	33882	44217	40888	52636	34759	47668

We consider the following findings important: Bipartition gives rise to certain transient disadvantages if the descendants are similar. They will have to share the food until they start to learn. Still, selective learning can be more effective than selective learning combined with other methods, or than other methods alone if the environment is SFSW. It seems that selection fits the SFSW structure and vice versa. This could be a good reason for the abundance of emergent SFSW structures in nature. This can be understood through the NFL Theorem if it is read backwards: If we find that a particular algorithm is more efficient than random search or than any other algorithm, then this winning algorithm 'knows' (fits) the most the underlying problem among the investigated algorithms. Based on our computer simulations it seems that highly clustered SFSW structures and simple selective learning algorithms match each other.

6 Conclusions

We investigated algorithms using evolutionary, reinforcement learning, and combined evolutionary and reinforcement learning strategies. We experimented with environments having different degree distributions and clustering coefficients. We generated different topological environments, using data collected during real Web search. Our study focused on the task of searching for new relevant documents. We found that in the scale-free small world environment the evolutionary weblog update algorithm performs the best. It outperformed a reinforcement learning based algorithm and the combinations of these two algorithms. We conjecture that the highly clustered nature and the small diameter of the environment match simple selection over other more sophisticated learning schemes. However, when the scale-free nature of the environment was kept but the small diameter of the environment was increased simply by restructuring the environment, then other algorithms performed better than the simple selection. We found in the 3 not small world environments that the combination of the weblog and reinforcement learning algorithms are the best. That is, when the diameter of the world becomes larger, then estimation of the long-term cumulated reward becomes important. Moreover, the combined algorithm showed the smallest performance variation both on the scale-free small world and on scale-free environments.

7 Acknowledgements

This material is based upon work supported by the European Office of Aerospace Research and Development, Air Force Office of Scientific Research, Air Force Research Laboratory, under Contract No. FA8655-03-1-3036. This work is also supported by the National Science Foundation under grants No. INT-0304904 and No. IIS-0237782. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the European Office of Aerospace Research and Development, Air Force Office of Scientific Research, Air Force Research Laboratory.

References

[1] Wolpert DH, Macready WG: No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation 1*, 1 (April 1997), 67–82.

[2] Barabási A, Albert R, Jeong H: Scale-free characteristics of random networks: The topology of the world wide web. *Physica A 281* (2000), 69–77.

[3] Albert R, Barabási A: Statistical mechanics of complex networks. *Reviews of Modern Physics 74* (2002), 47–91.

[4] Annunziato M, Huerta R, Lucchetti M, Tsimring LS: Artificial life optimization over complex networks. In *Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS* (March 2004).

[5] Kleinberg J, Lawrence S: The structure of the web. *Science 294* (2001), 1849–1850.

[6] Sutton R, Barto A: *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, 1998.

[7] Eiben AE, Smith JE: *Introduction to Evolutionary Computing.* Springer, 2003.

[8] Fryxell J, Lundberg P: *Individual Behavior and Community Dynamics.* Chapman and Hall, London, 1998.

[9] Clark C, Mangel M: *Dynamic State Variable Models in Ecology: Methods and Applications.* Oxford University Press, Oxford UK, 2000.

[10] Kampis G: *Self-modifying Systems in Biology and Cognitive Science: A New Framework for Dynamics, Information and Complexity.* Pergamon, Oxford UK, 1991.

[11] Csányi V: *Evolutionary Systems and Society: A General Theory of Life, Mind, and Culture.* Duke University Press, Durham, NC, 1989.

[12] Pachepsky E, Taylor T, Jones S: Mutualism promotes diversity and stability in a simple artificial ecosystem. *Artificial Life 8*, 1 (2002), 5–24.

[13] Tesauro GJ: Temporal difference learning and td-gammon. *Communication of the ACM 38* (1995), 58–68.

[14] Mataric MJ: Reinforcement learning in the multi-robot domain. *Autonomous Robots 4*, 1 (1997), 73–83.

[15] Stafylopatis A, Blekas K: Autonomous vehicle navigation using evolutionary reinforcement learning. *European Journal of Operational Research 108* (19998), 306–318.

[16] Moriarty D, Schultz A, Grefenstette J: Evolutionary algorithms for reinforcement learning. *Journal of Artificial Intelligence Research 11* (1999), 199–229.

[17] Tuyls K, Heytens D, Nowe A, Manderick B: Extended replicator dynamics as a key to reinforcement learning in multi-agent systems. In *ECML 2003, LNAI 2837* (2003), N. L. et al., Ed., Springer-Verlag, Berlin, pp. 421–431.

[18] Kondo T, Ito K: A reinforcement learning with evolutionary state recruitment strategy for autonomous mobile robots control. *Robotics and Autonomous Systems 46* (2004), 11–124.

[19] Kókai I, Lőrincz A: Fast adapting value estimation based hybrid architecture for searching the world-wide web. *Applied Soft Computing 2* (2002), 11–23.

[20] Lőrincz A, Kókai I, Meretei A: Intelligent high-performance crawlers used to reveal topic-specific structure of the WWW. *Int. J. Founds. Comp. Sci. 13* (2002), 477–495.

[21] Palotai Z, Gábor B, Lőrincz A: Adaptive highlighting of links to assist surfing on the internet. *Int. J. of Information Technology and Decision Making 4* (2005), 117–139.

[22] Szita I, Lőrincz A: Kalman filter control embedded into the reinforcement learning framework. *Neural Computation 16* (2004), 491–499.

[23] Schultz W: Multiple reward systems in the brain. *Nature Review of Neuroscience 1* (2000), 199–207.

[24] Joachims T: A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization. In *Proceedings of ICML-97, 14th International Conference on Machine Learning* (Nashville, US, 1997), D. H. Fisher, Ed., Morgan Kaufmann Publishers, San Francisco, US, pp. 143–151.

[25] Boley D: Principal direction division partitioning. *Data Mining and Knowledge Discovery 2* (1998), 325–244.

[26] Sutton R: Learning to predict by the method of temporal differences. *Machine Learning 3* (1988), 9–44.

[27] Watts DJ, Strogatz SH: Collective dynamics of 'small-world' networks. *Nature 393*, 6684 (1998), 440–442.

[28] Angkawattanawit N, Rungsawang A: Learnable topic-specific web crawler. In *Hybrid Intelligent Systems* (2002), A. Abraham, J. Ruiz-del-Solar, and M. Köppen, Eds., IOS Press, pp. 573–582.

[29] Menczer F: Complementing search engines with online web mining agents. *Decision Support Systems 35* (2003), 195–212.

[30] Risvik KM, Michelsen R: Search engines and web dynamics. *Computer Networks 32* (2002), 289–302.

[31] Cho J, Garcia-Molina H: Effective page refresh policies for web crawlers. *ACM Transactions on Database Systems 28*, 4 (2003), 390–426.

[32] Edwards J, McCurley K, Tomlin J: An adaptive model for optimizing performance of an incremental web crawler. In *Proceedings of the tenth international conference on World Wide Web* (2001), pp. 106–113.

Correlation Model of Worm Propagation on Scale-Free Networks

Zoran Nikoloski[1], Narsingh Deo[2], Ludek Kucera[1]

October 1, 2005

[1] *Department of Applied Mathematics, Faculty of Physics and Mathematics, Charles University, Malostranske Nam. 25, 118 00 Praha 1, Czech Republic, e-mail: nikoloski@kam.mff.cuni.cz*

[2] *School of Computer Science, University of Central Florida, Orlando Fl, 32816, USA*

Abstract

The problem of network worms is worsening despite increasing efforts and expenditure on cyber security. Worm propagation is a random process that creates a complex system of interacting agents (worm copies) over the propagation medium—a scale-free graph, representing real-world networks. Understanding the propagation of network worms on scale-free graphs is the first step towards devising effective techniques for worm quarantining. After presenting the drawbacks of existing mean-field models, we develop a pair-approximation (correlation) model of worm propagation that employs the salient network characteristics—order, size, degree distribution, and *transitivity*. Inclusion of the transitivity shows significant improvement over existing pair-approximation models. The validity of the model is confirmed by comparing the numeric solution of the model to results from our individual-based simulation. Our model demonstrates that the network structure has considerable impact on the propagation dynamics when the worm uses local propagation strategies.

Keywords: propagation, network worm, scale-free graph, correlation model

1 Introduction

Network worms represent a serious threat to confidentiality, integrity, and availability of computer resources on the Internet. The existing automated network-security solutions (*e.g.*, anti-virus software, firewalls, and intrusion detection systems) and human-dependent counter measures (e.g., software patching, traffic blocking) have been deemed inadequate for timely detection and control of worm propagation [10, 42, 43]. Since the problem of network worms is worsening every year despite increasing efforts and expenditure on cyber-security [10],

devising techniques for controlling their propagation is of great practical importance [29]. An important first step in developing control strategies is (1) to understand the dynamics of worm propagation and (2) how worm propagation is affected by the network structure.

Cyber attacks employ *malicious mobile-code* (*MMC*)—a program designed to perform a malicious action by propagate copies of itself to computers, having a known vulnerability, on a network. MMCs can be classified into three broad categories based on the extent of human intervention required for their propagation [24], namely: autonomous, human-dependent, and hybrid. Furthermore, based on their actions, MMCs may be grouped into three classes: Trojan horses, computer viruses, and network worms. A *Trojan horse* is a program that is given (fraudulently) the same name as a legitimate piece of software, but, when executed, it performs a malicious act. A *computer virus* is an MMC that modifies resident programs to perform malicious actions on a single computer. Like Trojan horses, computer viruses require human intervention to propagate on a network.

A *network worm* is a stand-alone program that propagates autonomously by sending copies of itself to other computers on the network. A computer is considered *compromised* if it hosts a replica of the worm. A worm *propagates* by sending a copy of *autonomous MMC* (*AMMC*) to a host, having the exploited vulnerability detected through *scanning* and *probing*. Therefore, the worm propagation is a random process that generates a complex system of interacting agents (AMMCs) over the propagation medium—a scale-free network. A recent study [47] has shown that: (1) probing attempts are in the range from one to three millions per day, (2) sources and destinations of intrusion are uniformly distributed among *Autonomous Systems*, (3) scanning of ports other than 80 (HTTP), characteristic for Code Red and Nimda, constitute 40% to 80% of all intrusion attempts, and (4) while vertical and, in particular, horizontal scans are prevalent [37], other methods such as coordinated and stealthy scans are widely used.

The size of an AMMC is relatively small (less than a hundred KB), and the scan and probe pieces are negligibly small (a kilobyte and a dozen bytes, respectively) [35, 36, 41]. The code of Sapphire Worm, for example, consisted of only 376 bytes [25]. Thus, the only immediately observable effect of an attack on a network is an increase in the routing-related requests [11, 22], as the worm keeps probing different hosts. Due to the short time required for their propagation, worms can inflict considerable damage to the networks. For instance, CodeRed Worm [9] infected 150,000 computer systems in 14 hours. The damage inflicted by Nimda [8] to 86,000 computer systems, has been estimated to $13 billion [10]. On January 24, 2003, the Sapphire Worm, taking advantage of a known vulnerability in Microsoft's SQL Server 2000, spread across different networks, including Bank of America's network of 13,000 ATMs [25]. Since a malicious worm has the capacity for global impact on today's network-dependent society, developing models of propagation and control is a first step towards a comprehensive network-security solution.

An important question in modeling worm propagation on a scale-free net-

work is whether the population of computers (whose communication is captured by the network) is to be represented in groups based on average degree and risk of transmission or whether the computers and their communication are to be explicitly simulated. The latter level of detail provides a powerful tool as it allows the structure of the network to be included in the control techniques. Between the extremes of networks whose represented by a complete graph (see Section 3) and individual-based simulation on scale-free networks, it is possible to divide the population of computers into groups with similar characteristics—*e.g.*, degree (number of contacts)—and derive a set of equations describing the propagation dynamics. Thus, the contact structure can be captured by *stratifying* the modeled population.

Our contribution here is a mathematical model of worm propagation that makes use of information about network structure—order, size, degree distribution, and clustering coefficient (transitivity). We compare the results of the model with an individual-based simulation of worm propagation on scale-free graphs that model the Internet (obtained from the Oregon Route View project). We point out that, although the introduction of the Internet has arguably made the assumption of sparseness no longer valid, the idea of *locality* (especially in the case of analytical modeling of local propagation strategies) is still applicable, and, therefore, used in our model.

The paper is organized as follows: In Section 2, we discuss adequate models for the propagation medium—the Internet, and define two abstractions of the Internet topology: the Microscopic Internet graph and the Macroscopic Internet graph. In Section 3, we present a comprehensive survey of the existing models for worm propagation and control by (1) categorizing the models into seven classes and (2) identifying their advantages and disadvantages. Derivation of a pair-approximation SIS and SIR model for worm propagation on scale-free graphs is presented in Section 4. Finally, to test the accuracy of our approach, in Section 5 we present a comparative empirical study of our pair-approximation model, the mean-field model, and the model of propagation on Erdös-Renyi graphs versus the results of the individual-based simulation.

2 Adequate Model for the Propagation Medium

Due to the diversity of exploited vulnerabilities, network worms can propagate on various types of networks. A (physical) *network* is a collection of interconnected computers, each with its distinct IP address. Such a network can be represented by a connected, undirected graph $G = (V, E)$, with the nodes as computers and the edges as the (physical) communication links (*e.g.*, wire, optical cable). Since G is connected, communication between any arbitrary pair of nodes u and v, takes place through a u, v-path in G. Although graph G, just defined, represents a physical network, it may also represent a logical (or virtual) network. For example, in an *e-mail graph*, a node v would represent a user and the (directed) edges emanating from v would go to all the individuals in the e-mail *address book* of v. Likewise, in a *webgraph* each node u would rep-

resent a web page and each (directed) edge emanating from $u$would represent a hyperlink coming out of web page u.

Despite the differences in what they model, these graphs have the following similar characteristics: (1) degree distribution, (2) clustering coefficient, and (3) average distance. We note that a scale-free random graph model cannot be accepted as a valid representation of a real-world network if it satisfies only these three characteristics. To clarify these characteristics, we give the following definitions:

Definition 1: *The degree distribution gives the probability that a node, chosen uniformly at random, is of degree d.*

Empirical studies of real-world networks have demonstrated that the degree distribution falls in the class of so-called scale-free (power-law) probability distributions, such that $P(d(v) = d) = d^{-f}$ ($f$is the exponent of the power-law degree distribution).

Definition 2: *Given a graph $G = (V, E)$ and a node $v \in V$ of degree $d(v)$, the clustering coefficient C_v of node v is defined as the ratio between the total number of edges incident on all pairs of neighbors of v and the number of edges in a clique formed by the neighbors of v.*

The *clustering coefficient* of G is the average of clustering coefficients over all nodes. The clustering coefficient of G has values in the range $0 \leq C \leq 1$. There is yet another measure of clustering in graphs, called *transitivity* [28].

Definition 3: *Transitivity is the ratio between the number of triangles and the total number of paths of length three.*

Definition 4: *Average distance of G is the mean over all shortest distances between any connected nodes.*

The Internet can be modeled on two levels: microscopic and macroscopic. In the *Microscopic Internet graph*, nodes stand for routers and hosts, while edges represent communication links. The *Macroscopic Internet graph* can be thought of as a contraction of the Microscopic Internet graph: here, each node represents an Autonomous System (which incorporates a number of routers). To simplify the analysis, parallel edges and loops (having negligible influence in modeling propagation) will be deleted from the *Macroscopic Internet graph.* Two nodes in the *Macroscopic Internet graph* are adjacent if there is at least one pair of routers (belonging to different Autonomous Systems) that can communicate. Note that both, the Microscopic and the Macroscopic Internet graphs are undirected. Faloutsos *et al.* [14] studied both graphs, and concluded that the degree distribution follows a power-law. In the Microscopic Internet graph, the exponent of the power-law f had a value of 2.48, while in the Macroscopic Internet graph, the exponent ranged between $f = 2.15$ and $f = 2.2$ (studies were

performed between 1997 and the end of 1998). Govindan and Tangmunarunkit [15] mapped the connectivity of nearly 150,000 router interfaces, confirming the power-law exponent of $f = 2.3$. The studies of Yook *et al.* [48] conducted between 1997 and 1999 showed that the Macroscopic Internet graph has a clustering coefficient in the range from 0.18 to 0.3 and an average distance between 3.70 and 3.77.

3 Existing Models of Worm Propagation

Due to the strong analogy between network worms and infectious diseases, epidemiological models have been widely used in modeling worm's propagation. Since a worm propagates along the edges of a network, we will use graph-theoretic terms to describe the existing epidemiological models of propagation. Epidemiological models are based on two simplifications [16]: (1) At any given time t, each node can be in one of a finite number of states, *e.g.* susceptible, quarantined-susceptible, removed-susceptible, infectious, quarantined-infectious, removed-infectious, and detected. The choice of which states to include in a model depends on the characteristics of the particular worm being analyzed and the purpose of the model; and (2) Translation of the worm transmission mechanism into a probability that a node will infect another node. In a similar way, transitions between other states of the model are described by simple probabilities. Epidemiological models can be analyzed analytically or by means of simulation.

The propagation takes place on a graph G with n nodes and m edges. Let $S(t)$ denote the number of susceptible nodes at time t, $Q_s(t)$ be the number of quarantined-susceptible nodes, $R_s(t)$ be the number of removed-susceptible nodes, $I(t)$ be the number of infectious nodes, $Q(t)$ be the number of quarantined-infectious, and $R(t)$ denote the number of removed nodes. The fraction of nodes in a particular state is represented by the lower case letter. Let β denote the rate at which susceptible nodes are infected. Most models of propagation assume β is constant, averaging out the differences in processor speed, network bandwidth, and location of the infectious node. The existing models also assume that a node cannot be infected multiple times.

Susceptible-Infectious (SI) model: In this class of models, once a susceptible node becomes infectious, it does not change its state. These models can be used in the study of the *worst-case propagation*, when automated and human counter-measures are not available. Let the average degree of an infectious node be $\overline{d}$, and the fraction of infectious nodes at time t be $i(t)$. The expected number of susceptible neighbors that can be infected by a given infectious node is $\overline{d}(1 - i(t))$. Since there are $I(t)$ infectious nodes, the total rate of newly-infected nodes is $\beta\overline{d}(1 - i(t))\, i(t)$. The general SI model is described by the differential equation (1):

$$\frac{di(t)}{dt} = \beta\overline{d}(1 - i(t))\, i(t), \tag{1}$$

with boundary conditions: $i(0) = \frac{I(0)}{n} > 0$ and for all $t \geq 0$, $i(t) + s(t) = 1$. The solution of equation (1) for the fraction of infectious nodes is the *logistic curve*: $i(t) = \frac{i(0)\ e^{\beta' t}}{1 - i(0) + i(0)\ e^{\beta' t}}$, where $\beta' = \beta \overline{d}$. The *S-shaped curve* describing the fraction of infectious nodes has three regions: (1) *slow start*, when only few nodes are infected at every time step, (2) *exponential growth*, when the number of newly-infected nodes grows exponentially, and (3) *equilibrium state*, when the number of infectious nodes assumes some value around which it fluctuates steadily.

If the worm propagates on the complete graph on n nodes, K_n, where $\overline{d} = (n-1)$, the model (1) can asymptotically be written as:

$$\frac{di(t)}{dt} = \beta (1 - i(t)) I(t), \tag{2}$$

with boundary conditions: $i(0) = \frac{I(0)}{n} > 0$ and for all $t \geq 0$, $i(t) + s(t) = 1$. One then has that $i(t) = \frac{i(0)\ e^{\beta (n-1) t}}{1 - i(0) + i(0)\ e^{\beta (n-1) t}}$. Staniford *et al.* [35] applied model (2) to fit the data collected by the Chemical Abstracts Services from the propagation of CodeRed I Worm, and estimated the product $\beta (n-1)$ for Code Red I to be 1.8. However, they used the number of scanned nodes, which is much larger than the number of infectious nodes, thus, leading to erroneous conclusions. Weaver [41] and Wagner *et al.* [38] used this model to study four local propagation strategies: hit-list, topological, permutation, and local-subnet, although we must note that the complete graph as underlying topology is inappropriate for studying such local strategies.

Zou *et al.* [50] used a modification of model (2) to analyze a *trend-detection* mechanism based on the traffic-anomaly created by worms. The detection system is composed of distributed ingress and egress sensors for worm activity. When the monitoring system receives a surge of illegitimate scans, a Kalman filter is activated to estimate the parameter β. Since in the early stage the propagation exhibits exponential growth with constant, positive rate, the model can be described as $I(t) = (1 + \beta n\, dt)\ I(t-1)$. The authors derived a bias-correction formula for estimation of the number of infectious nodes at time t, $I(t)$, from the number of observed infectious nodes $Z(t)$: Let σ be the average number of scans sent by an infectious node. After time interval dt, the expected number of scans observed by u monitors is $\beta u I(t) \sigma dt / 2^{32}$ (assuming the Internet is a complete graph), while the probability that any of the $I(t) - Z(t-1)$ infectious nodes are observed is $1 - \left(1 - u/2^{32}\right)^{\sigma dt}$. When the estimate of β starts oscillating around a positive constant value, the worm has been detected. Yet, it is not evident how the topology might affect Zou *et al.*'s detection mechanism.

In Erdös-Renyi random graphs with edge-density p, the expected degree of

a node is $p(n-1)$. The propagation on these graphs can be described as:

$$\frac{di(t)}{dt} = \beta p(n-1)(1-i(t))\, i(t), \tag{3}$$

with solution: $i(t) = \dfrac{i(0)\ e^{\beta p(n-1)t}}{1-i(0)+i(0)\ e^{\beta p(n-1)t}}$.

Susceptible-Infectious-Susceptible (SIS) model: In this class of models, an infectious node recovers at some rate, and thus it becomes susceptible again. These models can be used in the study of worm's propagation when some computers are temporarily turned off but are not patched (*e.g.*, the case of Code Red I worm). Let the average degree of an infected node be $\overline{d}$, and the rate at which an infectious node recovers be γ. The rate of newly-infected nodes is proportional to the expected fraction of susceptible neighbors, the number of infected nodes, and the probability β. The rate at which infectious nodes recover is proportional to the number of infectious nodes and rate γ. The system of differential equations (4) describes the general SIS model:

$$\frac{di(t)}{dt} = \beta\overline{d}(1-i(t))\, i(t) - \gamma i(t) \tag{4}$$

with boundary conditions $i(0) = \dfrac{I(0)}{n}$, and for all $t \geq 0$, $i(t) + s(t) = 1$. From equation (4), $\dfrac{di(t)}{dt} < 0$ if and only if $s(t) < \dfrac{\gamma}{\beta\overline{d}} = \delta$. Thus, the worm "dies out" if the initial fraction of susceptible nodes is below the *epidemic threshold* $\dfrac{\gamma}{\beta\overline{d}}$. The solution of (4) gives a functional form for the fraction of infectious nodes: $i(t) = \dfrac{(1-\delta)\, i(0)}{i(0)+(1-\delta-i(0))\ e^{-(\beta'-\gamma)t}}$, where $\beta' = \beta\overline{d}$. If the worm propagates on the complete graph on n nodes, K_n, where $\overline{d} = (n-1)$, the model (4) can asymptotically be written as [34]:

$$\frac{di(t)}{dt} = \beta(1-i(t))\, I(t) - \gamma i(t), \tag{5}$$

with solution $i(t) = \dfrac{(1-\delta)\, i(0)}{i(0)+(1-\delta-i(0))\ e^{-(\beta(n-1)-\gamma)t}}$.

Solomon [34] studied a modification of model (5) where the rate γ is a weighted average of the rate γ_1 (for computers not running anti-virus software), applicable to the fraction of infectious nodes, and the rate γ_2 (for computers running the most recent version of anti-virus software), applicable to the fraction of susceptible nodes, *i.e.*, $\gamma = \gamma_1 i(t) + \gamma_2(1-i(t))$. With this modification Solomon found that the necessary effectiveness of the anti-virus software (described by the rate γ) should be 0.5 in order to stop the propagation before it achieves exponential growth.

Kephart *et al.* [18, 19] employed model (4) to study the effects of three topologies on the propagation of *viruses*: Erdös-Renyi random graphs, regular

lattices of degree eight, and hierarchical random graphs. For the Erdös-Renyi random graphs with $\overline{d} \geq 5$, simulation results coincide with the predictions of the model. The simulation study of propagation on 100-by-100 lattice demonstrates quadratic growth, in contrast with the exponential growth characteristics for the complete graph and Erdös-Renyi graphs. The hierarchically-clustered random graphs in this study are generated as follows: given a rooted tree of height h, in which every node has a degree $(d+1)$ (*i.e.*, it has d successors), the nodes of the graph are the leaves of the tree. Two nodes, u and v, are made adjacent with probability $P(h(w))$ proportional to the height of node w—the first common ancestor of nodes u and v. In his simulation, Kephart used $P(h(w)) = \alpha p^{h(w)}$, where parameter p is used to control the degree of localization (when p tends to 0, the graph is composed of isolated nodes, while when p approaches 1, the topology of the hierarchically-clustered random graph is asymptotically that of the Erdös-Renyi random graphs). Here, the propagation shows sub-exponential growth. Further simulation studies conducted by Kephart [19] shows that sparsely-connected (random) graphs inhibit the propagation.

Previous models are limited in their accuracy due to their simplistic treatment of timing factors, such as *infection delay*—the length of time between the instant of worm's arrival at a node and the instant when this node becomes infectious to its neighbors. Model (4) could be altered to incorporate the infectious delay, as follows [40]:

$$\frac{di(t)}{dt} = \beta \overline{d} e^{-\gamma\varepsilon} (1 - i(t)) i(t - \varepsilon) - \gamma i(t), \tag{6}$$

where $i(t-\varepsilon) = 0$ for $t < \varepsilon$. At time $t \geq \varepsilon$, the fraction of infectious nodes is the same as the fraction of infectious nodes at time $(t-\varepsilon)$, since all nodes infected between $(t-\varepsilon)$ and t are delayed. The term $e^{-\gamma\varepsilon}$ accounts for the transfer of a node from infectious to susceptible state during the delay period. Equation (6) belongs to the class of non-linear delayed differential equations, which can be solved under the assumption $i(t-\varepsilon) = i(t)$. Wang *et al.* [40] support their analytical solution with simulation similar to that of *Kephart et al.* [18], and show that the epidemic threshold depends not only on the average degree, but also on the infection delay. In addition, Kim *et al.* [20] performed a simulation study of the propagation on a subgraph of the Internet, using a constant delay equal to the average round-trip time obtained from real-life traffic.

Pastor-Satorras *et al.* [30] modified model (4) to study the effects of the scale-free Barabasi-Albert topology on the propagation with rate of recovery $\gamma = 1$. Since a scale-free degree distribution is not concentrated around its mean value, the model must include differential equation for every group of nodes of degree k:

$$\frac{di_k(t)}{dt} = \beta k (1 - i_k(t)) \Theta\left(\{i_k(t)\}_{d_{\min}}^{d_{\max}}\right) - i_k(t), \tag{7}$$

where $\Theta\left(\{i_k(t)\}_{k=d_{\min}}^{d_{\max}}\right)$ describes the probability that a susceptible node of

degree k is adjacent to an infectious node. For a scale-free network, the probability that an edge is incident on a node of degree k is $kP(k)/\overline{d}$. The average probability that an edge is incident on an infectious node is then $\Theta(t) = \frac{1}{\overline{d}} \sum_{k=d_{\min}}^{d_{\max}} kP(k) i_k(t)$. The conclusion of this model is that scale-free topologies do not have epidemic-threshold. The authors also argued that the cut-off in the scale-free distribution forces a non-zero epidemic threshold. We point out that the result of this study is limited to scale-free topologies without degree-correlations. Contrary to this result, the simulation study of Eguiluz *et al.* [13] demonstrates that in the so-called *structured scale-free networks*, where adjacent nodes share large number of common neighbors, there exists a non-zero threshold even in the limit of large n.

While the results of the presented studies are valuable, a model, where nodes that have recovered and are no longer susceptible, could better approximate the realistic propagation of a worm when human counter-measures are in place.

Susceptible-Infectious-Removed (SIR) model and its variations: In this class of models, an infectious node can be removed (*i.e.*, it can no longer spread the worm). This model can be used to study the effects of software patching and traffic blocking. At any time t, a node can be susceptible, infectious, or removed. Let γ be the rate at which infectious nodes are removed. Using analogous arguments as in the previous section, the general SIR model can be written as:

$$\begin{aligned} \frac{di(t)}{dt} &= \beta\overline{d}(1 - i(t)) i(t) - \gamma i(t), \\ \frac{dr(t)}{dt} &= \gamma i(t), \end{aligned} \tag{8}$$

with boundary conditions: $i(0) = \frac{I(0)}{n} \geq 0$, $r(0) = \frac{R(0)}{n} \geq 0$, and for all $t \geq 0$, $i(t) + s(t) + r(t) = 1$. The epidemic threshold for SIR models is analogous to the one in SIS models. Zou *et al.* [49] used a modification of the system (8) to determine the effect of human counter-measures (on removing both susceptible and infectious nodes) and the decreasing rate $\beta(t)$. This so-called *two-factor* model assumes complete graph as underlying topology, and a constant fraction of the removed-infectious nodes at any time t:

$$\begin{aligned} \frac{di(t)}{dt} &= \beta(t)(1 - r(t) - r_s(t) - i(t)) i(t) - \frac{dr(t)}{dt}, \\ \frac{dr(t)}{dt} &= \gamma i(t), \\ \frac{dr_s(t)}{dt} &= \mu(1 - r(t) - r_s(t) - i(t))(r(t) + i(t)), \\ \beta(t) &= \beta(0)(1 - i(t))^{\eta}. \end{aligned} \tag{9}$$

It is unclear, however, how the parameters have been chosen in order to fit the data from the Code Red I worm propagation.

Boguna *et al.* [5] studied the SIR model, with the probability $\gamma = 1$, on scale-free topologies. Using the notation introduced in previous sub-section, the

model can be formulated as follows:

$$\begin{aligned} \frac{di_k(t)}{dt} &= \beta k \left(1 - i_k(t)\right) \Theta\left(\{i_k(t)\}_{d_{\min}}^{d_{\max}}\right) - i_k(t), \\ \frac{dr_k(t)}{dt} &= \gamma i_k(t), \end{aligned} \tag{10}$$

which can be solved if one assumes that $i(0)$ is very small in the beginning of the propagation. Pastor-Satorras *et al.* [30] conducted a simulation study to investigate the effects of node-immunization (*i.e.* node-removal) on the propagation, before the worm is introduced in the network. They demonstrated that random immunization is inefficient in slowing down the propagation; however, immunization targeted at nodes of highest degrees can significantly inhibit the growth of propagation. While the latter result seems interesting, the authors argue that detecting nodes of high degrees in scale-free networks is a difficult problem.

Similarly, the simulation study of Wang *et al.* [39] examines the effects of immunization of nodes on the propagation on two topologies: rooted trees and clustered networks (composed of cliques inter-connected with small number of edges). The simulation's parameter is the *propagation fan out*—number of nodes to which the worm can send replicas at each time step. The time needed for the worm to propagate from one node to another is assumed to be one time tick. The first set of simulation is conducted on networks where no immunized nodes exist to determine the number of times a node is re-infected (called *re-infection count*). Two types of immunization are simulated—random and selective. Random immunization performs better on rooted trees as there is only one path between any two nodes; thus, it is possible to cut off an entire sub-tree of the network, which is not the case with the clustered network. For the case selective immunization in rooted trees, nodes with highest re-infection counts were chosen (note, these nodes coincide with nodes with largest degrees). In the case of clustered networks, two strategies are used: first based on the re-infection count, and second on the weighted sum of the inter-cluster and inner-cluster degrees for every node. The first strategy was able to contain the propagation, but results in a higher propagation rate. The second could slow down the propagation rate, but was unable to contain the propagation.

The principal disadvantage of the studies in [39] and [30] is that immunization is static, *i.e.*, a fraction of nodes is immunized before the worm starts propagating. In reality, the counter-measures should be dynamic in nature to play important role in slowing down the propagation of the worm.

Susceptible-Infectious-Detected-Removed (SIDR) model: This model was analyzed by Williamson *et al.* [45] in order to determine the effectiveness of the behavior-blocking approach called virus throttling [44]. Virus-throttling is an automatic mechanism for slowing a worm's propagation. Here, a node can be in one of the four states: susceptible, infectious, detected (in which the virus has been detected and cannot actively spread further), and removed. The model assumes complete graph as underlying topology. The model involves two stages: in the first stage, prior to the release of the virus signature, nodes progress from

susceptible to infectious state at some rate β. In the second stage, after some time from the start of the propagation, the virus is detected at some rate γ. Two quantities are studied: the number of infectious nodes and the duration of propagation. The model incorporates virus throttling by dividing the nodes into two groups—throttled and un-throttled. If a throttled node is infected, it does not spread the virus, and immediately enters the detected state. The result of this study show that when more than half of the nodes have throttles, even a late signature will result in a small outbreak.

Susceptible-Infectious-Removed-Susceptible (SIRS) model: Wang *et al.* [40] used a modification of SIS model (4) to study the node's vigilance against infection: Once an infectious node is removed, it remains in this state for a length of time ν, called *vigilance period*, after which the removed node becomes susceptible again. Here, the susceptibility of a node is modeled via a parameter ϕ that takes values between 0 (indicating complete susceptibility) and 1 (indicating immunity). The model is described by the non-linear delay differential equation (11):

$$\frac{di\left(t\right)}{dt}=\beta\overline{d}\left(1-i\left(t\right)-\int_{t-\nu}^{t}i\left(t\right)\right)i\left(t\right)-\gamma i\left(t\right) \tag{11}$$

whose solution shows that the number of infectious nodes decreases as the vigilance period increases. It is worth noting the node's vigilance has no impact on the epidemic threshold.

Compartmental epidemiological models: Compartmental epidemiological models are used with stratified population. The topology in this models is the Macroscopic Internet graph, where every node represents a dense region—Autonomous System (AS). These models can be used to study intra-AS propagation, with the assumption that within an AS (with n_j nodes) the worm propagates as on a complete graph K_{n_j}. The infectious attempts can then be modeled as being external or internal to an AS. If the macroscopic Internet graph has k nodes, the SI compartmental model can be written as:

$$\frac{di_j\left(t\right)}{dt}=\left[\sum_{l=1}^{k}\beta\frac{n_l}{N}i_l\left(t\right)\right]\left(1-i_j\left(t\right)\right), \tag{12}$$

where $1\leq j\leq k$. Here, the parameter N denotes the total number of IP addresses. Serazzi *et al.* [33] used model (12) to derive equations for the bandwidth consumption at each node. For the SIR compartmental model, Liljenstam *et al.* [23] obtained:

$$\begin{aligned}\frac{di_j\left(t\right)}{dt}&=\left[\sum_{l=1}^{k}\beta\frac{n_l}{N}i_l\left(t\right)\right]\left(1-i_j\left(t\right)\right)-\gamma i_j,\\ \frac{dr_j\left(t\right)}{dt}&=\gamma i_j\left(t\right),\end{aligned} \tag{13}$$

where, again, $1\leq j\leq k$. Liljenstam *et al.* [23] used model (13) to study the destabilizing effects of worm propagation on the network infrastructure,

since the compartmental approach allows for inclusion of limited details about communication protocols. In this simulation study, the scan traffic is modeled by using a combination of the average scan rate, individual infection rates, and size of address space for each AS.

Discrete-time approximation models: Chen *et al.* [6] developed a deterministic approximation model of propagation on a complete graph K_n. If σ is the average scanning rate, with the assumption that the total number of nodes is 2^{32}, the average number of newly-infected nodes at step $(t+1)$ is $(S(t) - I(t)) \left[1 - \left(1 - 1/2^{32}\right)^{\sigma I(t)}\right]$. If the probability of removal is γ, in the next time step $\gamma I(t)$ nodes will become susceptible. Thus, the propagation can be described by a system of recurrences for the number of infectious and susceptible nodes.

4 Pair-approximation model on Scale-Free Networks

The existing epidemiological models on scale-free graphs [30, 49] do not explicitly give the system of differential equations for the propagation dynamics. The comparative studies include either simulation of worm's propagation on a macroscopic level or a system of differential equation for propagation on Erdös-Renyi and regular graphs. Thus, in all models described in Section 3, it is not evident how a realistic, scale-free network structure might affect the worm propagation.

Worm propagation is a random process that takes place on networks, such as: the Internet, World Wide Web, e-mail network, modeled as large scale-free random graphs. Using the salient features of the underlying scale-free graphs, here, we develop a realistic model of worm's propagation and techniques for dynamic quarantining. Cast in the SI framework, our model can be used to study the worst-case propagation and determine the optimal time for undertaking preventive action. On the other hand, cast in the SIR framework, this model can be used in the study of quarantining techniques against network worms.

Our model of worm's propagation belongs to the class of *pair-approximation network models*. The benefit of this class of models is that it can incorporate the spatial structure that the existing epidemiological models of propagation ignore. A survey of pair-approximation models is given by Rand [32]. In the pair-approximation model, the variables are the fractions of pairs of nodes in certain states. Usually, these equations contain higher-order correlations (*e.g.*, triples of nodes in certain states) which are approximated by the lower-order correlations. For the most part, previous work on pair-approximation models describes processes on regular-lattices. Our model extends the work by Earnes and Keeling [12] (for triangle-free networks) and Bauch [2] (for dynamic partnerships), and makes pair-approximation applicable to various scale-free topologies.

Next, we present the derivation of the system of differential equations describing the propagation in the SIS framework. Let $N(u)$ be the neighborhood of a node u, $p_t(i_u)$ the probability that, at time t, node u is infectious, and let

$p_t(s_u, i_v)$ be the joint probability that two adjacent nodes u and v are susceptible and infectious, respectively. The time evolution of the state of a single node in the SIS epidemic process can be written in the following form:

$$\begin{aligned} &\frac{dp_t(i_u)}{dt} = \beta \sum_{v \in N(u)} p_t(s_u, i_v) - \gamma p_t(i_u), \\ &p_t(s_u) + p_t(i_u) = 1. \end{aligned} \tag{14}$$

One can also develop an equation for the time evolution of $p_t(s_u, i_v)$ which in turn involves higher-order correlations. To solve this problem, we resort to some approximation scheme: For instance, the SIS epidemiological model assumes that $p_t(s_u, i_v) = p_t(s_u)\, p_t(i_v)$, and thus neglects correlation between states of nodes. In our approach, $p_t(i_u)$ and $p_t(s_u, i_v)$ are kept as variables of interests while the higher-order correlations are expressed, via some appropriate approximation, in terms of these quantities. The time evolution of $p_t(s_u, i_v)$ can be derived by using the Kolmogorov forward equation:

$$\begin{aligned} \frac{dp_t(s_u, i_v)}{dt} = &-(\beta + \gamma)\, p_t(s_u, i_v) - \beta \sum_{w \in N(u) - v} p_t(i_w, s_u, i_v) \\ &+ \beta \sum_{w \in N(v) - u} p_t(s_u, s_v, i_w) + \gamma p_t(i_u, i_v). \end{aligned} \tag{15}$$

Let Λ_a be the set of integers, representing the degrees of the neighbors for all nodes of degree a, and $[ab]$ be the number of edges incident on nodes of degrees a and b. Furthermore, let X, Y, and Z represent a state of a node (*e.g.*, susceptible and infectious).

Given a node u of degree a and a node v of degree b, define

$P_t(X_a, Y) = \frac{1}{a} \sum_{d(u)=a, d(v) \in \Lambda_a} P_t(x_u, y_v)$, and

$P_t(X_a, Y_b, Z) = \frac{1}{[ab]} \sum_{d(u)=a, d(v)=b, d(w) \in \Lambda_b} P_t(x_u, y_v, z_w)$.

Remark:

$P_t(X_a, Y) = \sum_{k \in \Lambda_a} P_t(X_a, Y_k)$ and $P_t(X_a, Y_b, Z) = \sum_{k \in \Lambda_b} P_t(X_a, Y_b, Z_k)$.

Let $E[X_a]$ denote the number of nodes of degree a in state X, $E[X_a Y_b]$ the number of pairs of nodes of degree a, in state X, adjacent to nodes of degree b, in state Y, and $E[X_a Y_b Z_c]$ denote the number of triples where a node of degree b, in state Y, is adjacent to a node of degree a and a node of degree c, in state X and Z, respectively. By multiplying equation (14) by n_a, the number of nodes of degree a in the graph G, one can obtain the following equation:

$$\frac{dE[I_a]}{dt} = \beta \sum_{k \in \Lambda_a} E[S_a I_k] - \gamma\, E[I_a], \tag{16}$$

where $E[S_a] + E[S_b] = n_a$. Similarly, one can transform equation (15) to obtain:

$$\frac{dE[S_a I_b]}{dt} = -(\beta + \gamma)\, E[S_a I_b] - \beta \sum_{k \in \Lambda_a} E[I_k S_a I_b] + \beta \sum_{k \in \Lambda_b} E[S_a S_b I_k] + \gamma E[I_a I_b]. \tag{17}$$

Let φ_{abc} denote the transitivity among nodes of degree a, b, and c, *i.e.*, the ratio of the number of 3-cycles to the number of connected triples whose nodes are of degree a, b, and c. To approximate the third moment $E[X_a Y_b Z_c]$, one can use the definition of multiplicative moments of two variables [17, 32] and the transitivity φ_{abc}, to find:

$$E[X_a Y_b Z_c] = \frac{b-1}{b} \frac{E[X_a Y_b] E[Y_b Z_c]}{E[Y_b]} \left((1 - \varphi_{abc}) + \varphi_{abc} \frac{n^2}{2m} \frac{E[X_a Z_c]}{E[X_a] E[Z_c]} \right). \tag{18}$$

Similarly, one can derive formulae for the other second moments and appropriate approximation of the third moments to obtain the following pair-approximation for the SIS framework:

$$\begin{aligned}
&\frac{dE[I_a]}{dt} = \beta \sum_{k \in \Lambda_a} E[S_a I_k] - \gamma E[I_a], \\
&\frac{dE[S_a]}{dt} = \gamma E[I_a] - \beta \sum_{k \in \Lambda_a} E[S_a I_k], \\
&\frac{dE[S_a S_b]}{dt} = -\beta \left(\sum_{k \in \Lambda_a} E[I_k S_a S_b] + \sum_{k \in \Lambda_b} E[S_a S_b I_k] \right) + \gamma \left(E[S_a I_b] + E[I_a S_b] \right) \\
&\frac{dE[S_a I_b]}{dt} = -(\beta + \gamma) E[S_a I_b] - \beta \left(\sum_{k \in \Lambda_a} E[I_k S_a I_b] - \sum_{k \in \Lambda_b} E[S_a S_b I_k] \right) \\
&+ \gamma E[I_a I_b], \\
&\frac{dE[I_a I_b]}{dt} = \beta \left(E[S_a I_b] + E[I_a S_b] + \sum_{k \in \Lambda_a} E[I_k S_a I_b] + \sum_{k \in \Lambda_b} E[I_a S_b I_k] \right) \\
&- 2\gamma E[I_a I_b].
\end{aligned} \tag{19}$$

Now, the system of differential equations (19) can be numerically solved by using the approximation given in equation (18). Note that our model differs from the one presented in [2] and [12] since we take into consideration the transitivity φ_{abc}, which turns out to have a significant effect on the outcome of the model (see Section 5).

Model (19) can be altered to obtain the system of differential equations (20),

describing the dynamics of propagation in the SIR framework:

$$
\begin{aligned}
&\frac{dE\left[I_a\right]}{dt} = \beta \sum_{k\in\Lambda_a} E\left[S_a I_k\right] - \gamma E\left[I_a\right], \\
&\frac{dE\left[S_a\right]}{dt} = -\beta \sum_{k\,\in\,\Lambda_a} E\left[S_a I_k\right], \\
&\frac{dE\left[R_a\right]}{dt} = \gamma E\left[I_a\right], \\
&\frac{dE\left[S_a S_b\right]}{dt} = -\beta \left(\sum_{k\in\Lambda_a} E\left[I_k S_a S_b\right] + \sum_{k\,\in\,\Lambda_b} E\left[S_a S_b I_k\right] \right), \\
&\frac{dE\left[S_a I_b\right]}{dt} = -\left(\beta+\gamma\right) E\left[S_a I_b\right] - \beta \left(\sum_{k\in\Lambda_a} E\left[I_k S_a I_b\right] - \sum_{k\,\in\,\Lambda_b} E\left[S_a S_b I_k\right] \right), \\
&\frac{dE\left[S_a R_b\right]}{dt} = -\beta \sum_{k\in\Lambda_a} E\left[I_k S_a R_b\right] + \gamma E\left[S_a I_b\right], \\
&\frac{dE\left[I_a I_b\right]}{dt} = \beta \left(E\left[S_a I_b\right] + E\left[I_a S_b\right] + \sum_{k\in\Lambda_a} E\left[I_k S_a I_b\right] + \sum_{k\in\Lambda_b} E\left[I_a S_b I_k\right] \right) \\
&-2\gamma E\left[I_a I_b\right], \\
&\frac{dE\left[I_a R_b\right]}{dt} = \beta \sum_{k\in\Lambda_a} E\left[I_k S_a R_b\right] + \gamma \left(E\left[I_a I_b\right] - E\left[I_a R_b\right] \right).
\end{aligned}
\tag{20}
$$

5 Pair-approximation Model vs. Individual-based Simulation

Our goal is to test the accuracy of the pair-approximation model (19) in comparison to: (1) the individual-based simulation of the worm propagation on a Macroscopic Internet graph (on n nodes and average degree $\overline{d}$), and (2) the standard SIS model (which ignores correlation) on two topologies: the complete graph on n nodes (model (5)) and the Erdös-Renyi graph on n nodes with average degree $\overline{d}$ (model (4)).

The empirical study is conducted on Macroscopic Internet graphs. To obtain the Macroscopic Internet graphs, we used the data for inter-connectedness of the Internet on the Autonomous System level collected by the University of Oregon Route View Project and made available by NLANR (National Laboratory of Applied Network Research). We considered snapshots of the Internet of various order and size, shown in Figure 1. The pre-processing step consists of determining parameters for model (19): for given degrees a, b, and c, the number of adjacent nodes of degree a and b, the set Λ_a, and the transitivity φ_{abc} are determined.

Next, we developed an individual-based simulation of the stochastic propagation process on a Macroscopic Internet graph. The individual-based simulation has two advantages: First, the propagation process and the underlying topology can be controlled to simulate different scenarios. Second, this simulation

Date	Order	Size
08.11.1997	3015	5156
02.04.1998	3522	6324
03.07.1998	3797	6936
02.10.1998	4180	7768
14.01.1999	4517	8376
02.04.1999	4885	9276
02.07.1999	5357	10328
02.10.1999	5861	11313
02.01.2000	6474	12572
03.04.2000	7246	14629
02.07.2000	7956	15943
02.10.2000	8836	17823
02.01.2001	9048	18172
16.03.2001	10515	21455

Figure 1: Macroscopic Internet graphs used in simulations

provides very precise and detailed information about the propagation dynamics without any biases which might be present in real data. The individual-based simulation combines Monte Carlo simulation of events, taking place at given rates, with an event-scheduler that determines the order in which events happen in the system. The scheduler is implemented as a priority queue. The system is composed of nodes that can be either susceptible or infectious. There are two types of events that can take place: infection and curing. If a node u is infectious, it attempts infection of each of its neighbors at rate β. Node u might be cured, and, thus, become susceptible, at rate γ. Let as assume that u has been cured at time t, and has been re-infected at time $t + dt$. Any infection generated by the node u in the time interval $[t, t + dt)$ can be discarded by the scheduler. The event of u attempting infection of an already infectious neighbor v at time t is also discarded by the scheduler.

We simulated worm propagation in the Susceptible-Infectious framework in order to determine the time the worm takes to infect all nodes of a given graph G. Simulations were performed on ten Macroscopic Internet graphs (the results for four graphs, identified by the top entry in the left-most column appear in Figure 2). To determine how choice of the initial node influences the propagation, we first determined the labels of three nodes with smallest degrees and three nodes with highest degree, shown in the first and second column of each table in Figure 2. The rest of the entries show the average time over 100 simulations for the worm to propagate on all nodes by starting from a pre-specified initial node and spreading with infectious rate β. The general results of the experiments can be summarized as follows:

1. *The time to propagate to all nodes decreases with the increase of the degree of the initial node.*
   As an infectious node of higher degree has bigger pool of susceptible nodes, it gives the worm the possibility to establish a considerable fraction of infectious nodes in the early stages of the propagation. On average, the

propagation to all nodes of G initiated from a node of maximum degree takes time by 5% shorter compared to the propagation that starts from a node of minimum degree.

2. *The time to propagate to all nodes increases with the increase of the order of the graph.*
   Given two graphs G_1 and G_2, $|V(G_1)| < |V(G_2)|$, whose degree distributions follow the same scale-free distribution, have diameters $D(G_1)$ and $D(G_2)$, respectively, such that $D(G_1) < D(G_2)$. Therefore, on a graph with greater diameter the worm takes longer time to infect all nodes.

3. *The time to propagate to all nodes does not strictly decreases with the increase of the infectious rate β.*
   In other words, there is a value of the infectious rate β at which the function $t(\beta)$ has a local minimum, as shown in Figure 3. According to the simulation results shown in Figure 2, the value of $\beta = 1.5$ seems to be invariant and depends only on the exponent of the scale-free degree distribution of the graph G. The reason for such behavior is that, at the local minimum, rapidly-building correlations between the states of adjacent nodes hinder the propagation by lowering the number of available susceptible nodes. This observation is of *particular interest* as it provide the means to "control" the propagation of a fast-spreading worm by reducing its rate to the threshold value.

Figures 4 and 5 below, show the number of infectious nodes as a function of time, comparing the three deterministic models with two results of the stochastic individual-based simulation. Neither the mean-field model nor the Erdös-Renyi (or a $\overline{d}$-regular graph on n nodes) satisfactorily predicts the level of propagation (*i.e.*, the number of infectious nodes at a given time). The second-order Runge-Kutta numerical solution of the proposed pair-approximation model (18), (19) with results of the second pre-processing task as input, performs matches the results of the individual based simulation.

The model on $\overline{d}$-regular graphs underestimates the equilibrium level because it does not include the nodes of high degrees (*i.e.*, the *core* of the scale-free graph). The mean-field model consistently over-estimates the number of infectious nodes because the correlations and graph structure, that may inhibit propagation, are ignored. In contrast, our pair-approximation model includes both nodes of various degrees and correlations between states of nodes, and gives an accurate representation of the stochastic propagation process.

Remark: Two numerical algorithms for solving the system of differential equations—Euler's method and Runge-Kutta method—were compared. For small values of dt, such as 0.001 used in the simulation, even the Euler's method produces good results.

6 Conclusion

Developing an accurate model for the worm propagation is of critical importance not only for understanding better the worm's behavior but also for devising techniques to contain such cyber attacks. The existing studies include either simulation of worm propagation on a macroscopic level or a system of differential equation for propagation on Erdös-Renyi and regular graphs. Moreover, in all existing models of worm propagation, it is not evident how the network structure might affect the dynamics of the stochastic propagation process. Our contribution here is twofold: (1) a model of propagation on a scale-free graph G which takes as input: the number of nodes, number of edges, number of edges incident on nodes of certain degrees, and transitivity of the graph, and (2) implementation of an individual-based simulation for worm propagation that can be cast in different epidemiological frameworks. The accuracy of the model is tested by comparing the numerical solution (by second-order Runge-Kutta method) of the pair-approximation model to the results from the individual-based simulation on scale-free Macroscopic Internet graphs. Our model has the potential to be used in developing realistic techniques for propagation control—topic of an ongoing research.

References

[1] Albert R, Jeong H, Barabasi AL: Diameter of the World Wide Web. Nature 1999; 401: 130 - 131.

[2] Bauch CT: A Versatile ODE Approximation to a Network Model for the Spread of Sexually Transmitted Diseases. Journal of Mathematical Biology 2002; 45: 375 - 395.

[3] Berk VH, Gray RS, Bakos G: Using Sensor Networks and Data Fusion for Early Detection of Active Worms, in: Proc. of SPIE conference on Sensors, and Command, Control, Communications and Intelligence, 2002.

[4] Berk VH, Bakos G: Designing a framework for Active Worm Detection on Global Networks, in Proc. of IEEE International Workshop on Information Assurance, 2003.

[5] Boguna M, Pastor-Satorras R, Vespignani A, Epidemic Spreading in Complex Networks with Degree Correlations, in: Pastor-Satorras R(ed): Lecture Notes in Physics, 2003; 625: 127 - 147.

[6] Chen Z, Gao L, Kwiat K: Modeling the Spread of Active Worms, in: Proc. of IEE INFOCOM, 2003.

[7] Chen LC, Carley KM: The Impact of Network Topology on the Spread of Anti-Virus Countermeasures, CASOS Working paper, available at: http://www.casos.ece.cmu.edu, 2003.

[8] CERT Coordination Center, Advisory CA-2001-26, available at: http://www.cert.org/advisories/CA-2001-26.html 2001.

[9] CERT Coordination Center, Advisory CA-2001-19, available at: http://www.cert.org/advisories/CA-2001-19.html 2002.

[10] Computer Security Institute, Ninth Annual Computer Crime and Security Survey, available at: http://i.cmpnet.com/gocsi/db_area/pdfs/fbi/FBI2004.pdf 2004.

[11] Cowie J, Ogielski AT, Premore BJ, Yuan Y: Global Routing Instabilities Triggered by Code Red II and Nimda, available at: http://www.renesys.com/projects/bgp_instability 2001.

[12] Earnes KTD, Keeling MJ: Modeling Dynamic and Network Heterogeneities in the Spread of Sexually Transmitted Diseases, in: Proc. of National Academy of Sciences, 2002; 99: 13330 - 13335.

[13] Eguiluz VM, Klemm K: Epidemic Threshold in Structured Scale-free Networks. Physics Review Letters, 2002; 89: 108701.

[14] Faloutsos M, Faloutsos P, Faloutsos C: On Power-Law Relationships of the Internet Topology, in: Proc. of SIGCOMM, 1999.

[15] Govindan R, Tangmunarunkit H: Heuristics for Internet Map Discovery, in: Proc. of the IEEE INFOCOM, 2000.

[16] Hethcote HW: Mathematics of Infectious Diseases. SIAM Review, 2000; 42: 599 - 653.

[17] Keeling MJ: The Effects of Local Spatial Structure on Epidemiological Invasions, in: Proc. of the Royal Society of London B, 1999; 266: 859 - 867.

[18] Kephart JO, White SR: Directed-Graph Epidemiological Models of Computer Viruses, in: Proc. of IEEE Symposium on Security and Privacy, 1991; 343.

[19] Kephart JO: How Topology Affects Population Dynamics, in: Langton CG (ed), Artificial Life III, Addison-Wesley, 1994.

[20] Kim K, Radhakrishnan S, Dhall SK: Measurement and Analysis of Worm Propagation on Internet Network Topology, in: Proc. of IEEE Conference on Computer Communications and Networks, 2004.

[21] Kumar R, Raghavan P, Rajagopalan S, Tomkins A, Upfal E: Random Graph Models for the Web Graph, in: Proceedings of FOCS, 2000; 57 - 65.

[22] Lan K, Hussain A, Dutta D: Effects of Malicious Traffic on the Network, in: Proc. of PAM, 2003.

[23] Liljenstam MN, Berk, VH, Gray RS: Simulating Realistic Network Worm Traffc for Worm Warning System Design and Testing, in: Proc. of ACM Workshop on Rapid Malcode, 2003.

[24] Mirkovic J, Martin J, Reiher P: A Taxonomy of DDoS Attacks and DDoS Defense Mechanisms. ACM SIGCOMM Computer Communication Review, 2004; 34: 39 - 53.

[25] Moore D, Paxson V, Savage S, Shannon C, Staniford S, Weaver N: Inside the Slammer Worm, available at: http://www.computer.org/security/v1n4/j4wea.htm 2003.

[26] Moore D, Shannon C, Voelker GM, Savage S: Internet Quarantine: Requirements for Containing Self-Propagating Code, in: Proc. of IEEE INFOCOM 2003.

[27] Moore D, Network Telescopes, Presentation at DIMACS Large-scale Internet Attacks Workshop, available at: http:// www.caida.org 2003.

[28] Newman MEJ, Strogatz SH, Watts DJ: Random Graph Models of Social Networks, in: Proc. of National Academy of Science of the USA, 2002; 99: 2566 - 2572.

[29] Nikoloski Z, Deo N: The Game of Cops and Robbers on Graphs: A Model for Quarantining Cyber Attacks. Congressus Numerantium, 2003; 162: 193 - 215.

[30] Pastor-Satorras R, Vespignani A: Epidemics and immunization in scale-free networks, in: Bornholdt S, Schuster HG (ed): Handbook of Graphs and Networks: From the Genome to the Internet, Wiley-VCH, 2002; 113 - 132.

[31] Pennock DM, Flake GW, Lawrence S, Glover EJ, Giles CL: Winners don't Take All: Characterizing the Competition for Links on the Web, in: Proc. of National Academy of Sciences of the USA, 2002; 99: 5207 - 5211.

[32] Rand DA: Correlation Equations and Pair Approximation for Spatial Ecologies, in: McGlade J (ed): Advanced Ecological Theory, Principles and Applications, Blackwell Science, 1999; 100 - 142.

[33] Serazzi G, Zanero S: Computer Virus Propagation Models, in: Calzarossa MC, Gelenbe E (ed): Lectures Notes in Computer Science, 2004; 2965: 26 - 50.

[34] Solomon A: Epidemiology and Computer Viruses, available at: http://vx.netlux.org/lib/static/vdat/epepidem.htm 1990.

[35] Staniford, S, Paxson V, Weaver N: How to Own the Internet in Your Spare Time, in: Proc. of the USENIX Security Symposium, 2002; 149 - 167.

[36] Vivo M, Vivo G, Koeneke R, Isern G: Internet Vulnerabilities Related to TCP/IP and T/TCP, Internet Security Attacks at the Basic Level. Operating Systems Review, 1998; 32: 4 - 15.

[37] Vivo M, Carrasco E, Isern G, Vivo G: A Review of Port Scanning Techniques. ACM SIGCOMM Computer Communication Review, 1999; 29 41 - 48 .

[38] Wagner A, Dubendorfer T, Plattner B, Hiestand R: Experiences with Worm Propagation Simulations, in: Proc. of ACM workshop on Rapid Malcode, 2003; 34 - 41.

[39] Wang C, Knight JC, Elder MC: On Computer Viral Infection and the Effect of Immunization, in: Proc. of the Annual Computer Security Applications Conference, 2000.

[40] Wang Y, Wang C: Modeling the Effects of Timing Parameters on Virus Propagation, in: Proc. of ACM Workshop on Rapid Malcode, 2003; 61 - 66.

[41] Weaver N: Potential Strategies for High-speed Active Worms: A Worst Case Analysis, available at: http://www.cs.berkeley.edu/∼nweaver/worms.pdf (2002).

[42] Weaver N, Paxson V, Staniford S, Cunningham R: A Taxonomy of Computer Worms, in: Proc. of ACM Workshop on Rapid Malcode, 2003.

[43] Weaver N, Paxson V, Staniford S, Cunningham R: Large Scale Malicious Code: A Research Agenda, available at: http://www.cs.berkeley.edu/∼nweaver/large_scale_malicious_code.pdf 2003.

[44] Williamson MM: Throttling Viruses: Restricting Propagation to defeat Malicious Mobile Code, in: Proc. of Annual Computer Security Applications Conference, 2002; 61.

[45] Williamson MM, Leveille J: An Epidemiological Model of Virus Spreading and Cleanup, in: Proc. of Virus Bulletin Conference, 2003.

[46] Wu J, Vangala S, Gao L, Kwiat K: An Effective Architecture and Algorithm for Detecting Worms with Various Scan Techniques, in: Proc. of the Network and Distributed System Security Symposium, 2004.

[47] Yegneswaran V, Barford P, Ullrich J: Internet Intrusions: Global Characteristics and Prevalence, in: Proc. of ACM SIGMETRICS International Conference on Measurement and Modeling of computer systems, 2003;, 138 - 147.

[48] Yook S-H, Jeong H, Barabasi AL: Modeling the Internet's Large-scale Topology, in: Proc. of National Academy of Sciences of the USA, 2002; 99: 13382 - 13386.

[49] Zou CC, Gong W, Towsley D: Code Red Worm Propagation Modeling and Analysis, in: Proceedings of ACM conference on Computer and Communications Security, Washington, 2002; 138 - 147.

[50] Zou CC, Gong W, Towsley D, Gao L: Monitoring and Early Detection for Internet Worms, in: Proc. of ACM Conference on Computer and Communication Security, 2003.

[51] Zou CC, Gong W, Towsley D: Worm Propagation Modeling and Analysis under Dynamic Quarantine Defenses, in: Proc. of ACM Workshop on Rapid Malcode, 2003.

AS graph 02.10.1998		beta				
		0.2	0.5	0.9	1.5	1.8
min degree	node					
1	47	18.3675	7.89627	4.74167	3.20504	3.55996
2	17	18.0627	7.91097	4.6643	3.19542	3.46717
3	22	18.4302	7.76855	4.69197	3.14913	3.33457
max degree	node					
590	5	17.158	7.73703	4.45511	3.01981	3.34582
524	12	16.9676	7.37089	4.50048	2.96612	3.39495
355	10	17.2982	7.48106	4.5871	3.10994	3.2987

1pt

AS graph 02.07.1999		beta				
		0.2	0.5	0.9	1.5	1.8
min degree	node					
1	63	18.3734	7.98559	4.81368	3.24062	3.42445
2	19	18.5593	8.17793	4.70665	3.19895	3.48956
3	15	18.5207	7.93869	4.74685	3.18794	3.343
max degree	node					
1193	2	17.4871	7.43152	4.44526	2.98155	3.39874
674	10	17.6827	7.58919	4.57659	3.11	3.41901
588	7	17.5386	7.70924	4.63347	3.10909	3.38447

1pt

AS graph 02.07.2000		beta				
min degree	node	0.2	0.5	0.9	1.5	1.8
1	15	19.8676	8.42701	5.11502	3.42114	3.71483
2	18	20.1391	8.62913	5.15767	3.41545	3.56204
3	23	19.4788	8.44671	5.05909	3.3409	3.5603
max degree	node					
1772	2	18.9615	8.0676	4.97258	3.29313	3.48871
961	9	19.2357	8.28964	4.90184	3.33247	3.49154
802	7	18.9133	8.24647	4.96538	3.34996	3.47815

1pt

AS graph 16.03.2001		beta				
min degree	node	0.2	0.5	0.9	1.5	1.8
1	44	21.0673	8.91663	5.35035	3.57629	3.76316
2	37	21.385	8.90863	5.35318	3.47916	3.74718
3	34	20.6299	8.93965	5.33861	3.54236	3.64725
max degree	node					
2277	2	20.1728	8.44475	4.98278	3.37857	3.58247
1231	13	20.3132	8.67817	5.248	3.46303	3.62322
899	15	20.4644	8.87768	5.2025	3.42517	3.69419

Figure 2: Time to propagate to all nodes of a Macroscopic Internet graph for five different values of the parameter β.

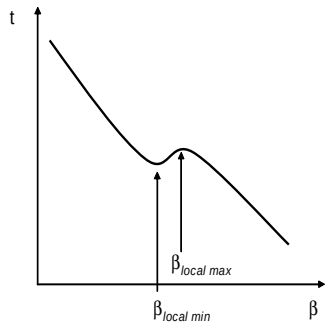


Figure 3: Time to infect all nodes as a function of the rate β is not a strictly decreasing function

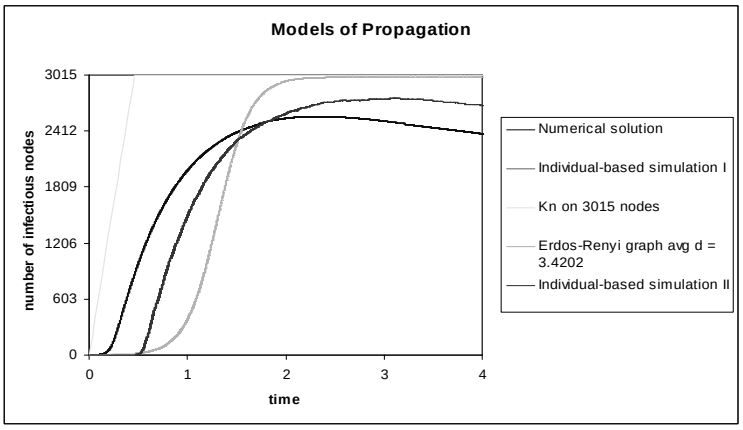


Figure 4: Susceptible-Infectious-Susceptible models of propagation—numerical solution of pair-approximation model, individual-based simulation of propagation on an Internet graph $n = 3015$ and $m = 5156$, propagation on complete graph $n = 3015$, propagation on Erdos-Renyi random graphs with $\overline{d} = 3.4202$; parameters of propagation $\beta = 1.8, \gamma = 0.05$

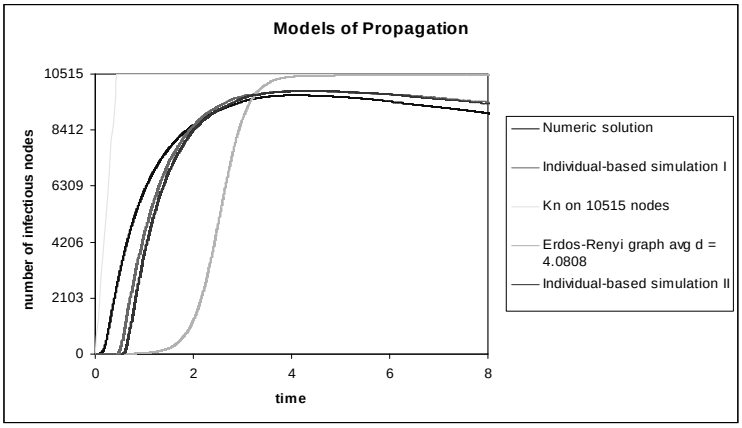


Figure 5: Susceptible-Infectious-Susceptible models of propagation—numerical solution of pair-approximation model, individual-based simulation of propagation on an Internet graph $n = 10515$ and $m = 21455$, propagation on complete graph $n = 10515$, propagation on Erdos-Renyi random graphs with $\overline{d} = 4.0808$; parameters of propagation $\beta = 0.9, \gamma = 0.02$

Analysis and visualization of large scale networks using the k-core decomposition

Ignacio Alvarez-Hamelin[1], Luca Dall'Asta[1], Alain Barrat[1] and Alessandro Vespignani[2]

[1] Laboratoire de Physique Théorique CNRS, Bât 210, Université de Paris-Sud, 91405 ORSAY Cedex France.

[2]School of Informatics and Department of Physics, University of Indiana, Bloomington, IN 47408, USA.

1 Introduction

In recent times, the possibility of accessing, handling and mining large-scale networks datasets has revamped the interest in their investigation and theoretical characterization along with the definition of new modeling frameworks. In particular, mapping projects of the World Wide Web (WWW) and the physical Internet offered the first chance to study topology and traffic of large-scale networks. Gradually other studies followed describing population networks of practical interest in social science, critical infrastructures and epidemiology [1, 3, 7, 10]. The study of large scale networks, however, faces us with an array of new challenges. The definitions of centrality, hierarchies and structural organizations are hindered by the large size of these networks and the complex interplay of connectivity patterns, traffic flows and geographical, social and economical attributes characterizing their basic elements.

In this paper, we propose the use of the k-core decomposition to study the hierarchical properties of large scale Internet maps. The k-core decomposition [4] consists in identifying particular subsets of the network, called k-cores, each one obtained by recursively removing all the vertices of degree smaller than k, until the degree of all remaining vertices is larger than or equal to k. Larger values of coreness clearly correspond to vertices with larger degree and more central position in the network's structure. The k-core decomposition therefore provides a probe to study the properties of the network's regions of increasing centrality. Here we analyze Internet networks at both the Autonomous Systems and router level. In both cases we find that k-cores are always made by a single connected component, indicating the presence of a hierarchy of well defined regions of which it is possible to investigate the statistical properties. Strikingly, the various distributions and quantities analyzed appear to be invariant in the various k-cores. This characteristic appears extremely important in providing an operative definition of a topological self-similarity of scale-free graphs and prompts to the k-core decomposition as a suitable transformation equivalent to a scale-change in the topological space of networks non-embedded in the geometrical space.

Motivated by the previous finding we developed a visualization algorithm based on the k-core decomposition that allows the identification of networks' fingerprints, according to properties such as hierarchical arrangement, degree correlations and centrality, etc. The distinction between networks with seemingly similar properties is achieved by inspecting the different layouts generated by the visualization algorithm. The running time of the algorithm grows only linearly with the size of the network, granting the scalability needed for the visualization of very large scale networks. We apply the proposed visualization algorithm to real Internet maps and several computer generated graphs aimed at their modeling. The visualization algorithm appears to be a convenient tool able to clearly pinpoint the differences of Internet maps obtained at different granularities and with different experimental techniques. In addition, the inspection of computer generated networks provides a first approach to models validation. The presented visualization algorithm is publicly available [8].

Keywords: network analysis, k-cores, Internet, visualization

2 k-core decomposition

In this section, first we introduce the definition of k-core decomposition, then we show how the application of this decomposition can shed light on important hierarchical properties of graphs, and finally we present a visualization that aids to identify networks.

Let us consider a graph $G = (V, E)$ of $|V| = n$ vertices and $|E| = e$ edges, the definition from [4] of k-cores is the following

Definition: *A subgraph $H = (C, E|C)$ induced by the set $C \subseteq V$ is a k-*core *or a core of order k iff $\forall v \in C : \texttt{degree}_H(v) \geq k$, and H is the maximum subgraph with this property.*

A k-core of G can therefore be obtained by recursively removing all the vertices of degree less than k, until all vertices in the remaining graph have degree at least k. It is shown by V. Batagelj and M. Zaversnik [4] that the complexity of this decomposition is e, the number of edges.

2.1 Analyzing AS and IR graphs

We apply the k-core decomposition to Internet's maps. The autonomous system level is represented by collected routes of *Oregon route-views* [9] project, called AS from May 26, 2001. For the router level, we use the graph obtained by an exploration of CAIDA project [6] between April 21st and May 8th, 2003.

Figure 1 displays the cumulative degree distribution for the first k-cores, for two maps of Internet (upper plots); namely, the probability $P_>(d)$ that any vertex in the networks has a degree larger than d. Strikingly, the shape of the distribution, i.e. a power-law with an exponential cut-off, is not affected by the decomposition. In particular the exponent of the power-law is robust although the range of variation of the degree decreases. This feature defines a striking property of statistical self-similarity of the network and the generated k-cores, which resemble one with each other under the opportune rescaling of the average degree.

Another relevant quantity is the *clustering coefficient* that measures the local group cohesiveness and is defined for any vertex j as the fraction of connected neighbors of j [11]: $cc_j = 2 \cdot n_{\texttt{link}}/(d_j(d_j - 1))$, where $n_{\texttt{link}}$ is the number of links between the d_j neighbors of j. The study of the clustering spectrum $cc(d) = \frac{1}{n_d} \sum_{j/d_j=d} cc_j$, allows e.g. to uncover hierarchies in which low degree vertices belong generally to well interconnected communities (high clustering coefficient), while hubs connect many vertices that are not directly connected (small clustering coefficient). Figure 1 presents $cc(d)$ in the lower plots. Also in this case the shape of the spectrums is preserved as the network is recursively pruned of its low-degree vertices.

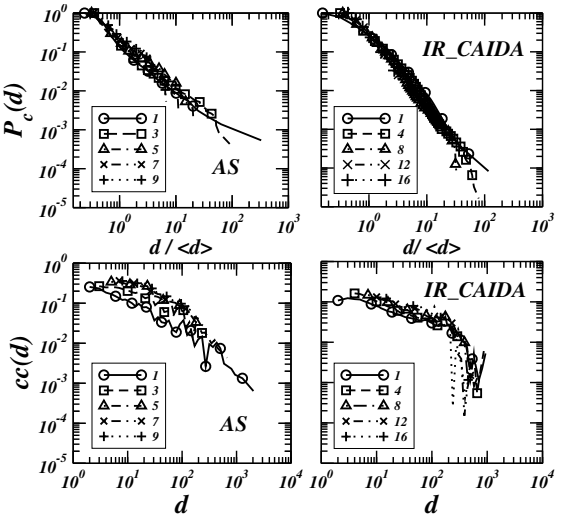


Figure 1: Top: cumulative degree distribution of the various k-cores. The degree is normalized by the average degree of each k-core. Bottom: clustering coefficient spectrum of the various cores.

2.2 Network fingerprints

Motivated by the previous analysis we propose a visualization algorithm based on k-core decomposition that places vertices in 2 dimensions, the position of each vertex depending on its coreness and on the coreness of its neighbors (see also [5]). A color code allows for the identification of core numbers, while the vertex's original degree is immediately provided by its size that depends logarithmically on the degree (see Figs. 2 and 3). For the sake of clarity, our algorithm represents a small percentage of the edges, chosen uniformly at random. The aim is to provide a clear visualization of the hierarchies, coreness shells and self-similarity observed in the context of statistical analysis presented in the previous section. As mentioned, in the most general situation, indeed, the recursive removal of vertices having degree less than a given k can break the original network into various connected components, each of which might even be once again broken by the subsequent decomposition. While this is not the case in the Internet maps that we have analyzed, we cannot exclude this possibility and a central role in our visualization method is played by the possibility of a multi-components representation of the k-core decomposition. The complete description of the algorithm can be found in [2]. The most remarkable fact is that the IR is populated at all levels, and it also has high degree nodes in low k-shells,

while the AS has few nodes in higher k-shells and it has only high degree nodes in higher k-shells.

3 Conclusions

In this paper we have presented the application of the k-core decomposition to the analysis and visualization of large scale networks. The k-core decomposition allows the progressive pruning of large networks and the identification of subgraphs of increasing centrality. The study of these subgraphs and their statistical properties uncover the main hierarchical layers of the network and allows for their statistical characterization. Strikingly, we observe for heterogeneous graphs such the Internet at the Autonomous System (AS) and Router level a statistical self-similarity of the topological properties for cores of increasing centrality. We propose a general visualization algorithm that allows for the graphical distinction of the k-core hierarchy along with the degree of the vertices and their relation with the hierarchical position of the neighbors. The obtained results show the possibility of gaining clear insights on the architecture of many real world and synthetic networks. Networks with different topological properties and structural arrangement can be distinguished and the hierarchical arrangements of the elements rationalized. The present visualization strategy may be also used for determining if a certain model fits or not with the real data, providing a further interesting tool for models validation.

Acknowledgments: This work has been partially funded by the European Commission - Fet Open project COSIN IST-2001-33555 and contract 001907 (DELIS).

References

[1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–, 2000.

[2] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani. k-core decomposition: a tool for the visualization of large scale networks. *arxiv.org*, cs.NI/0504107, 2005.

[3] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley. Classes of small world networks. *Proc. Natl. Acad. Sci. (USA)*, 97:11149–11152, 2000.

[4] V. Batagelj and M. Zaversnik. Generalized Cores. *arxiv.org*, cs.DS/0202039, 2002.

[5] M. Baur, U. Brandes, M. Gaertler, and D. Wagner. Drawing the AS Graph in 2.5 Dimensions. In *"12th International Symposium on Graph Drawing, Springer-Verlag editor"*, pages 43–48, 2004.

[6] Cooperative Association for Internet Data Analysis, Router-Level Topology Measurements. `http://www.caida.org/tools/measurement/skitter/router_topology/`.

[7] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of networks: From biological nets to the Internet and WWW*. Oxford University Press, 2003.

[8] LArge NETwork VIsualization tool. `http://xavier.informatics.indiana.edu/lanet-vi/`.

[9] University of Oregon Route Views Project. `http://www.routeviews.org/`.

[10] R. Pastor-Satorras and A. Vespignani. *Evolution and structure of the Internet: A statistical physics approach*. Cambridge University Press, 2004.

[11] D. J. Watts and S. H. Strogatz. Collective Dynamics of Small-World Networks. *Nature*, 393:440–442, 1998.

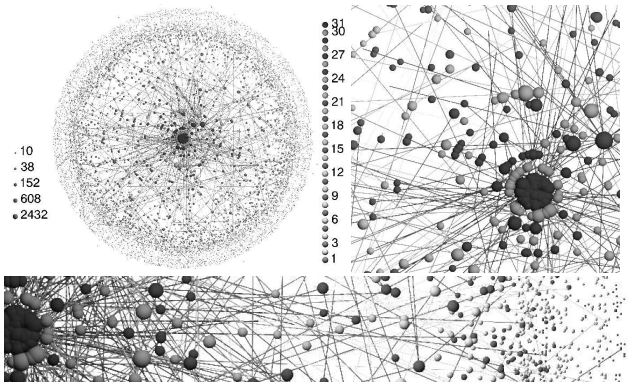


Figure 2: Graphical representation of the AS+ graph. The color code for the coreness is given on the right of the representation, while the legend for the degree of the vertices is given on the left, showing the maximum degree node. The following figure use the same legend.

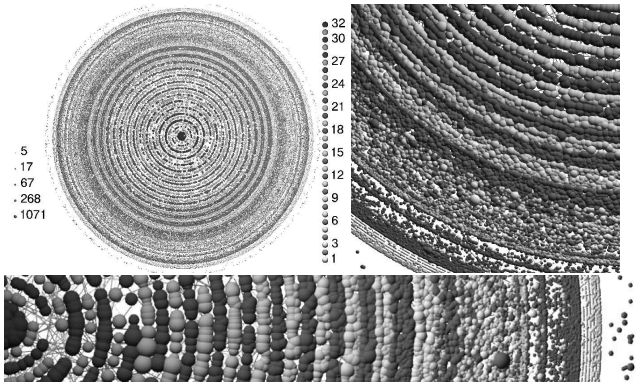


Figure 3: Graphical representation of the CAIDA IR graph.

DATA STREAM COMPUTATION FOR MONITORING STATISTICS OF MASSIVE WEBGRAPHS

LUCIANA S. BURIOL*, DEBORA DONATO*, STEFANO LEONARDI*
AND TOBIAS MATZNER+

Abstract. We are interested in computing properties of large graphs, as the webgraph, using data stream algorihtms. In this work we report results on computing the indegree rarity distribution of a graph obtained as a stream of edges. We implement a rarity algorithm proposed in the literature and show experimentally that the results approximate very well the optimal value with very limited use of memory and time. Moreover, considering some structure in the stream, we present results for the algorithm adapted for maintaining the rarity distribution of the number of cliques of size three.

1. Introduction

Data stream algorithms aim to maintain the underlying information of a stream of data, using small memory space. The data is processed on the fly, as it is generated, or it can also be read from second memory devices. Typical applications of data stream algorithms are originated from massive datasets such as network traffic measurements, telephone call records, biological datasets and atmospheric observations. In these applications is unnecessary or impractical to read data multiple times. In many cases, the data is not even stored. This paper focuses on a "new" natural application for data streams. We are interested in using data stream algorithms for monitoring statistical and topological properties of large graphs such as the webgraph. By webgraph we mean the directed graph generated from the link structure of webpages: each webpage is a node and each hyperlink is an arc in this graph. The graph read in a streaming fashion considers each edge as an item and the stream is not required to be structured.

Several theorical results have been proposed in this new research field, some of them have not yet been implemented and experimented, some of them are not practical. In this paper we observe how a data stream algorithm behaves in practice for computing the indegree rarity distribution of a graph over the arc arrivals. More specifically, we maintain the distribution of the number of nodes that has a given indegree over the total number of different nodes seen in the stream so far. We use the algorithm proposed by Datar and Muthukrishnan [5] and show experimentally that the results are very close to the optima even when a low precision is requested. The original algorithm proposes the use of min-wise hash functions, whereas we

use universal hashing [5]. This decision is due to the fact that computing min-wise hashing consumes about two orders of magnitude more time than universal hashing without providing better results in practice for the graphs we have tested.

When considering a specific structure in the data stream, other properties can be computed. For example, reading the stream in an adjacency list fashion, the same rarity algorithm can be used for estimating the density of minors such as small bipartite cliques.

The indegree of webpages is an important measure of their popularity. The experimental observation of the indegree distribution has been the subject of seminal works aimed to characterize the structure of the web-graph [2, 3]. This study has also revealed a surprising number of dense subgraphs, specifically bipartite cliques, of moderately small size [8], considered as cores of hidden web communities.

In the next section we present the α-rarity algorithm of Datar and Muthukrishnan [5]. Section 3 describes the adaptations of the α-rarity algorithm for computing indegree rarity distribution, as well as for computing minors of small size. In section 4 we present experimental results for rarity of indegree e bipartite cliques of size three `(k3,3)` and for the minors `k(1,3)` and `k(2,3)`. We generalize the set of all minors mentioned above using the term `k(i,3)`, where i denotes the number of nodes in the graph that points to each node of a triple (set of three nodes). Comparison with the results of an optimal computation shows excellent practical results of our implementations.

2. Estimating rarity over data stream windows

We use the $\alpha - rare$ algorithm of Datar and Muthukrishnan [5] for driving our experiments. Consider a stream of items a_i generated in a universe `U`=[1,..,n]. A stream is a set of m elements $a_1, a_2, ..., a_m$ such that $a_i \in U$. An item i is called $\alpha - rare$ if it appears exactly α times in the stream. Let's call $\#\alpha - rare$ the number of elements that appear exactly α times in the stream. Likewise, $\#distinct$ denotes de number of distinct items in the stream. The α-rarity ρ_α is defined as the ratio $\rho_\alpha = \frac{\#\alpha - rare}{\#distinct}$. In other words, the α-rarity of a stream is the measure of number of items that repeat exactly α times in the stream.

The algorithm proposed by Datar and Muthukrishnan [5] for computing the α-rarity of a stream uses min-wise hash functions. Min-wise independent permutation families are defined in [4]. The referred algorithm uses only *O(log N + log u)* space, and *O(log log N)* per item processing time. It estimates ρ by $\hat{\rho} \in [1 \pm \epsilon]\rho + \epsilon\rho$ for a given fraction ϵ, with hight probability. The algorithm uses $h = 2\epsilon^{-3}p^{-1}log\tau^{-1}$ hash functions and two $|h|$-vectors, `min` and `count`, in main memory. Each position i of the vector `min` contains the minimum value found so far by the min-wise hash h_i, whereas `count` maintains, for each position i, the number of times that the current minimum

min-wise value was found. For each value of α, $\hat{\rho}$ is computed as the ration between the number of counters that have exactly value α and h.

A slightly different algorithm is proposed for computing the α-rarity of a windowed stream. E.g, Considering a fix window size equal to `W`, the algorithm maintains the α-rarity of the last `W` items seen in the stream.

3. Computing the rarity distribution of indegree and `k(i,3)`

In this sections we describe how the not-windowed algorithm described in the previous section is adapted to compute the α-rarity algorithm for computing the indegree and `k(i,3)` rarity distributions of a graph.

Consider an arbitrary scan of a digraph G=(V,E). The items of the stream, in this case, are the list of edges. The α-rarity of the stream can be understood as the percentage of nodes that has indegree α. With the underlying data stored for estimating α, we can compute the α_i- rarity for any value of i. The rarity distribution can be computed for a complete stream, or for the window of the last `W` items seen in the stream.

Reading the stream in an adjacency list fashion, the same rarity algorithm can be used for approximating the density of minors, such as small bipartite cliques. Such kind of structured data stream can be found naturally on some applications. For example, during a crawling process, each current fetched page is parsed and all outgoing links of this page identified. In this case G is also read in a streaming fashion, but considering all outgoing links of a node $i \in V$ in sequence. The lists of outgoing edges are not required to be in any specific order, as well as the edges intern to each list. So, for each node u, for each outgoing edge $(\overrightarrow{u,a}) \in OUT(u)$, triples are calculated considering node a and all combinations two by two of the head-node of the edges seen so far in $OUT(u)$. E.g, triples `(a,b,c)` are calculated for nodes $b, c \in OUT(u)$ considering edges $(\overrightarrow{u,b})$ and $(\overrightarrow{u,c})$ previously located in $OUT(u)$ than $(\overrightarrow{u,a})$. So, the overall number of triples (T) of the graph is the sum of the combination three by three of head-nodes of the outgoing list of each node $u \in V$, e.g., $T = \sum_{i=1}^{i} \frac{d_i*(d_i-1)*(d_i-2)}{6}$ where $d_i = |OUT(i)|$ is the outdegree of the node i. We require to store in main memory the whole outgoing adjacency list of the current node.

4. Experimental Results

In this section we describe the experimental results we performed using the α-rarity algorithm applied in maintaining the indegree and `k(i,3)` distributions.. The algorithms were coded in g++ version 3.3.2. The experiments were conducted in a Intel Pentium IV, with 1GB RAM, running Mandrake 9.0.

Due to the excessive computational time spent by min-wise hash functions, we use universal hash functions instead. We used the hash function (`hash31`) and the random number generator (`prng_int`) from the online available codes from the MassDAL group of Rutgers

(*http://www.cs.rutgers.edu/ muthu/massdal-code-index.html*). We use an optimized version of Jerry Zhao's implementation [10] of an approximate restricted min-wise independent permutation family proposed by Alon et al. [1].

We conducted our experiments on streams of Wikipedia graphs. A graph of this type is generated from the link structure of the online and free-content encyclopedia Wikipedia (*www.wikipedia.org*). Following the definition of a webgraph, each article is a node, and each hyperlink is a link in the graph. One graph is extracted for each language. We generate streams of edges of the wikipedia graphs following their generation on time. Due to space restrictions, we limited the presentation of experimental results in this extended abstract to the `wikiEN` and `wikiPT` graphs. The graphs were obtained from an old dump of July 2004. Some comments are added about the experimental results on the other three graphs. Graph `wikiPT` contains 8,131 nodes and 48,168 edges, while graph `wikiEN` is two orders of magnitude larger containing 286,754 nodes and 4,065,530 edges.

Figure 1 presents results for the rarity for the unbounded case, using 100 (right) and 1000 (left) hash functions. The lines are plot for a logarithmic number of indegree values. For a good approximation, a larger number of hash functions are required. But we observed, that even with a small number of hash functions, the results are close to the optima. The plot omits results for indegree higher than 63 for the sake of clarity of the figure, but a complete plot would present additional lines on the bottom of the figure, appearing on increasing order of the number of edges processed.

For the windowed case, similar quality of results were found, but spending more time, as it was expected.

We also found good approximation when using the α-rarity algorithm for computing the rarity distribution of `k(i,3)` on the graph. Results for i=1,2,3 are plot in Figure 2. The plot is in log scale to be able to visualize all three distributions clearly on the same plot. Usually the number of `k(1,3)` $\gg$ `k(2,3)` $\gg$ `k(3,3)`. The difference between this values decrease with the increase of i. Observe, for example, the precision on results between the estimated and exact computation of `k(1,3)` and `k(2,3)`. Since `k(1,3)` is found many more times than `k(2,3)`, the results are more accurate. For values of $i > 4$ we did not plot for the sake of clarity of the plot, but the precision on the results decrease with the increase of i. As expected, we have less precision for computing $\hat{\rho}$ of α-rare elements that occur less frequently.

5. Concluding remarks

In this paper we use in practice data stream algorithms for computing statistical and topological properties of large graphs. We presented experimental results for the α-rarity algorithm applied on webgraphs for computing the rarity distribution of indegree and `k(i,3)` and obtained very good approximations. For the windowed case, applied for the indegree distribution,

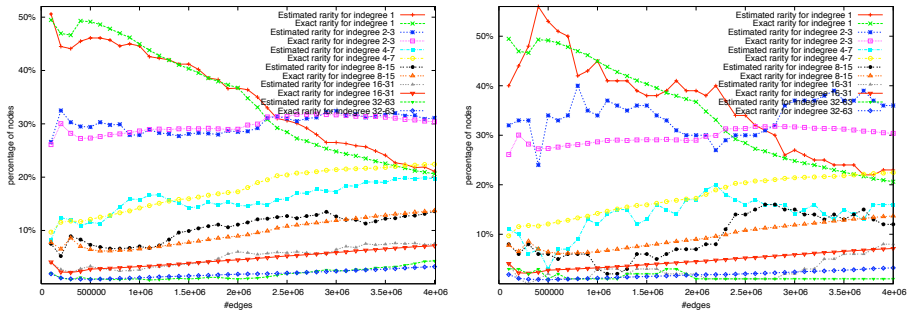


FIGURE 1. Estimated and exact indegree rarity distribution computed for edges arrivals of graph `wikiEN`. The estimation makes use of 1000 (graph on the left) and 100 (graph on the right) universal hashing functions. Values are presented to α up to 63, presented as $\log_2$ plot. This plot presents the percentage of nodes with a given indegree (y-basis) considering the amount of edges processed so far (x-basis). Results are plot every 100,000 items processed.

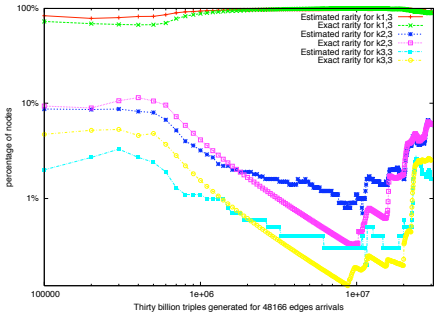


FIGURE 2. Plot in log scale of the estimated and exact `ki,3` rarity distribution, for i=1,2 and 3, computed for edges arrivals of graph `wikiPT`. The estimation makes use of 1000 universal hash functions. This plot presents the percentage of triples pointed by exactly i nodes (y-basis) considering the amount of triples seen so far (x-basis). The triples are computed accordingly with the edges arrivals. Results are plot every 10,000 triples processed.

we observed again good approximation in a reasonable time. For the `k(i,3)` estimation we obtained good approximations, but spending long time. That happens because, in this case, all triples obtained are hashed by the `#h` hash functions. For the `wikiPT` graph, we observe a total of 624 triples generated for each edge processed.

We conclude that using universal hashing by this algorithm speed up a lot the codes, maintaining good approximations.

As further work, we would like to test other algorithms that estimates interesting statistical and topological properties of webgraphs. Moreover, dynamic aspects of webgraphs also could be explored, as edges being inserted and removed over time. The α-rarity algorithm does not have solution for deletions. But a recent publications [6, 7] support also deletions.

6. Acknowledgements

We are very thankful to Jerry Zhao for providing the first version of the min-wise hash functions and for the suggestions for its optimization. We also thanks S. Muthukrishnan for several helpful discussions.

References

[1] N. Alon, O. Goldreich, J. Hastad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Journal of Random structures and Algorithms*, 3(3):289–304, 1992.

[2] A.L. Barabasi and A. Albert. Emergence of scaling in random networks. *Science*, (286):509, 1999.

[3] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, S. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *Computer Networks*, 33:309–320, June 2000.

[4] A.Z. Broder, M. Charikar, A.M. Frieze, and M. Mitzenmacher. Min-wise independent permutations. *Proc. of STOC*, pages 327–336, 1998.

[5] M. Datar and S. Muthukrishnan. Estimating rarity and similarity over data stream windows. *LNCS*, 2461:323–334, 2002.

[6] Irina Rozembaum G. Cormogode, S. Muthukrishnan. Summarizing and mining inverse distributions on data streams via dynamic inverse sampling. *Proccedings of the 31st VLDB Conferenct*, 2005.

[7] C. Sohler G. Frahling, P. Indyk. Sampling in dynamic data streams and applications. *21st Annual Symposium on Computational Geometry*, 2005.

[8] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the web for emerging cyber communities. pages 403–416, 1999.

[9] D. Sivakumar Z. Bar-Yosseff, R. Kumar. Reductions in streaming algorithms, with an application to counting triangles in graphs. *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 623–632, 2002.

[10] J. Zhao. An implementation of min-wise independent permutation family. 2005. http://www.icsi.berkeley.edu/ zhao/minwise/.

(*) Dipartimento di Informatica e Sistemistica, Universitá di Roma "La Sapienza", Via Salaria 113, 00198 Roma, Italy

E-mail address: {buriol,donato,Stefano.Leonardi}@dis.uniroma1.it

(+) Fakultät für Informatik, Universität Karlsruhe, Karlsruhe, Germany

E-mail address: tobias.matzner@rechnerpost.de

Distributed Algorithms for Data Propagation in Deeply Networked Wireless Sensor Devices

Ioannis Chatzigiannakis Sotiris Nikoletseas Paul Spirakis

Research Academic Computer Technology Institute and Patras University
P. 0. Box 1122, 261 10 Patras, Greece.
Email: {ichatz,nikole,spirakis}@cti.gr

Abstract

Wireless sensor networks are comprised of a vast number of ultra-small fully autonomous computing, communication and sensing devices, with very restricted energy and computing capabilities, which co-operate to accomplish a large sensing task. Such networks can be very useful in practice in applications that require fine-grain monitoring of physical environment subjected to critical conditions (such as inaccessible terrains or disaster places).

Very large numbers of sensor devices can be deployed in areas of interest and use *self-organization and collaborative methods* to form deeply networked environments. Features including the huge number of sensor devices involved, the severe power, computational and memory limitations, their dense deployment and frequent failures, pose *new design and implementation aspects*. The efficient and robust realization of such large, highly-dynamic, complex, non-conventional environments is *a challenging algorithmic and technological task*.

In this paper we present certain important aspects of the design, deployment and operation of distributed algorithms for data propagation in wireless sensor networks and discuss some characteristic protocols, along with an evaluation of their performance.

Keywords: Wireless sensor networks, distributed algorithms, data propagation, performance evaluation

This work has been partially supported by the IST Programme of the European Union under contract number IST-2004-001907 (DELIS) and the Programme PYTHAGORAS under the European Social Fund (ESF) and Operational Program for Educational and Vocational Training II (EPEAEK II).

1 Introduction

Recent dramatic developments in micro-electro-mechanical (MEMS) systems, wireless communications and digital electronics have already led to the development of small in size, low-power, low-cost sensor devices. A vast number of such sensor devices that integrate sensing with wireless network interfaces, that collect and disseminate information about the physical environment, are deployed in areas of interest (e.g. inaccessible terrains, disaster places, etc.) for fine grained monitoring in different classes of applications [2]. Some typical services provided by the network are [5]: (i) *Periodic Sensing* (the sensor devices constantly monitor the physical environment and continuously report their sensors' measurements to

a control center), (ii) *Event driven* (to reduce energy consumption, sensor devices monitor silently the environment and communicate to report when certain events are realized) and (iii) *Query based* (sensor devices respond to queries made by a supervising control center). Recently, new applications have been proposed, that require different approaches for disseminating sensor data to the control center, such as *Target Tracking* (where sensors exchange sensor readings in order to detect the movement pattern of a detected target) [11] or *Area Surveillance* (where sensors are equipped with video capturing devices) [17, 10].

The quality of the services provided can be measured in terms of (i) *delivery rate* (or *success rate*) that corresponds to the ratio of packets delivered to the control center over all packets generated by the sensors that correspond to a particular event, (ii) *energy dissipation rate* that captures the energy dissipated by the sensors in the process of propagating packets towards the control center and (iii) *propagation delay* (or *latency*), the time that elapsed from the realization of a particular event, to the final delivery of the message reporting it, to the control center. The importance of each of the above metrics depends on the nature of the application since there are inherent trade-offs between success rate, energy and latency. Trying to minimize *energy dissipation rate*, in an attempt to extend the lifetime of the network, by possibly forcing sensors to alternate between sleep and awake time periods [7, 22], inevitably results in increased *propagation delays*.

The efficient and robust realization of such large, highly-dynamic, complex, non-conventional, deeply networked environments is *a challenging algorithmic and technological task*. An approach for propagating information in such networks is to use routing techniques similar to those for mobile ad-hoc networks ([20]), however, the huge number of sensor devices involved, the severe power limitations, their dense deployment and frequent failures, pose *new design and implementation aspects* which are essentially different not only with respect to distributed computing and systems approaches but also to ad-hoc networking techniques.

We emphasize the following characteristic differences between sensor networks and ad-hoc networks: (i) the number of sensor particles in a sensor network is extremely large compared to that in a typical ad-hoc network, (ii) sensor networks are typically prone to faults and (iii) the limitations in energy, computational power and memory are much more severe in sensor networks. Because of faults as well as energy limitations, sensor nodes may (permanently or temporarily) join or leave the network. This leads to highly dynamic network topology changes. Because of the these rather unique characteristics of sensor networks, efficient and robust distributed protocols and algorithms should exhibit the following critical properties:

Scalability. Distributed protocols for sensor networks should be highly scalable, in the sense that they should operate efficiently in extremely large networks composed of huge numbers of nodes. This feature calls for an urgent need to prove by analytical means and also validate (by large scale simulations) certain efficiency and robustness (and their trade-offs) guarantees for asymptotic network sizes.

Efficiency. Because of the severe energy limitations of sensor networks and also because of their time-critical application scenaria, protocols for sensor networks should be efficient, with respect to both energy and time.

Fault-tolerance. Sensor particles are prone to several types of faults and unavailabilities, and may become inoperative (permanently or temporarily). Various reasons for such faults include physical damage during either the deployment or the operation phase, permanent (or temporary) cease of operation in the case of power exhaustion (or energy saving schemes,

respectively). The sensor network should be able to continue its proper operation for as long as possible despite the fact that certain nodes in it may fail.

2 An Abstract Model for Wireless Sensor Networks

Sensor networks are comprised of a vast number of ultra-small homogenous sensors, which we here call *particles*. Each particle is a fully-autonomous computing and communication device, characterized mainly by its available power supply (battery) and the energy cost of computation and transmission of data. Such particles (in our model here) cannot move.

We adopt here (as a starting point) a two-dimensional (plane) framework: A *wireless sensor network* (a set of grain particles) is spread in an area (for a graphical presentation, see Fig. 1). Usually the deployment of particles is done in a rather random manner (such as when particles are dropped by an airplane over the area of interest. In variations of this basic model, we may include the possibility of a (more or less) structured deployment (possibly done by humans or robots). Let n be the number of sensor particles in the area.

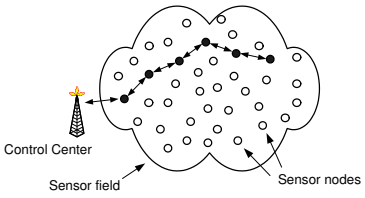


Figure 1: A Wireless Sensor Network

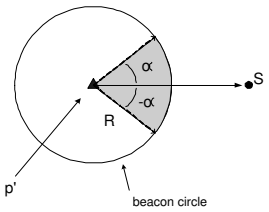


Figure 2: Directed transmission of angle α

There is a single point in the network area, which we call the sink S, that represents a control center where data should be propagated to. In variations of this basic model, there might be multiple sinks, which may be static or moving.

The particles are equipped with a set of monitors (sensors) for light, pressure, temperature etc. Each particle has a *broadcast* (digital radio) *beacon mode* which can be also a directed transmission of angle α around a certain line (possibly using some special kind of antenna, see fig. 2). The transmission range (which we denote by $\mathcal{R}$) can vary while the transmission angle (let it be α) is fixed and cannot change throughout the operation of the network (since this would require a modification or movement of the antenna used). Note that the protocols we study in this work can operate even under the broadcast communication mode (i.e. $\alpha = 2\pi$).

We believe that this model depicts accurately enough the technological specifications of real wireless sensor systems. Similar models are being used by other researchers in order to study sensor networks (see [12, 18]). The above assumptions suggest a strong model, that however does not trivialize the problem; we believe that even assuming such a model, the design of efficient distributed algorithms is still a challenging task. In contrast to [14, 15], our model is weaker in the sense that *no geolocation abilities* are assumed (e.g. a GPS device) for the particles leading to more generic and thus stronger results. In [13] a thorough comparative study and description of wireless sensor systems is given, from the technological point of view.

3 Distributed Algorithms for Data Propagation

Because of the complex nature of a sensor network (that integrates various aspects of communication and computing), protocols, algorithmic solutions and design schemes for all layers of the networking infrastructure are needed. Far from being exhaustive, we mention the need for frequency management solutions at the physical layer, for Medium Access Control (MAC) protocols to cope with multi-hop transmissions at the data link layer. The interested reader may use the excellent survey by Akyildiz et al [1] for a detailed discussion of design aspects of all layers of the networking infrastructure.

We focus in this paper on distributed algorithms for the network layer. We believe that a complementary use of rigorous analysis and large scale simulations is needed to fully investigate the performance of data propagation protocols in wireless sensor networks. In particular, asymptotic analysis may lead to provable efficiency and robustness guarantees towards the desired scalability of protocols for sensor networks that have extremely large size. On the other hand, simulation allows to investigate the effect of a great number of detailed technical specifications of real devices, a task that is difficult (if possible at all) for analytic techniques which, by their nature, use abstraction and model simplicity.

Any distributed algorithm solving the data propagation problem must satisfy the following three properties:

- **Correctness**. The distributed algorithm must guarantee that data arrives to the position S, given that the whole network exists and is operational.
- **Robustness**. The distributed algorithm must guarantee that data arrives at enough points in a small interval around S, in cases where part of the network has become inoperative.
- **Efficiency**. If the distributed algorithm activates k particles during its operation then Π should have a small ratio of the number of activated over the total number of particles $r = \frac{k}{N}$. Thus r is an energy efficiency measure of Π.

We below present two representative state-of-the-art protocols that try to avoid flooding the network, achieving good performance (with respect to time and energy) and robustness.

3.1 The Probabilistic Forwarding Protocol (PFR)

The PFR (Probabilistic Forwarding) protocol [6] is inspired by the probabilistic multi-path design choice for the Directed Diffusion paradigm mentioned in [14]. Its basic idea of the protocol (introduced in [6]) is to minimize energy consumption by *probabilistically favoring certain paths of local data transmissions towards the sink.*

The protocol *avoids flooding* by favoring (in a probabilistic manner) data propagation along sensors which lie "close" to the (optimal) transmission line, ES, that connects the sensor node detecting the event, E, and the sink, S. This is implemented by locally calculating the angle $\phi = (\widehat{EPS})$, whose corner point P is the sensor currently running the local protocol, having received a transmission from a nearby sensor, previously possessing the event information. If ϕ is equal or greater to a predetermined threshold ($\phi_{threshold}$), then p will transmit (and thus propagate the event information further). Else, it decides whether to transmit with probability equal to $\frac{\phi}{\pi}$. Because of the probabilistic nature of data propagation decisions and

in order to prevent the data propagation process from early failing, we initially use (for a short time period which we evaluate) a flooding mechanism that leads to a sufficiently large *"front"* of sensors possessing the data under propagation. When such a "front" is created, we perform probabilistic forwarding.

Note that transmission along this line is energy optimal. However it is not always possible to achieve this optimality, basically because certain sensors on this direct line might be inactive, either permanently (because their energy has been exhausted) or temporarily (because these sensors might enter a sleeping mode to save energy). Further reasons include (a) physical damage of sensors, (b) deliberate removal of some of them (possibly by an adversary in military applications), (c) changes in the position of the sensors due to a variety of reasons (weather conditions, human interaction etc). and (d) physical obstacles blocking communication.

Essentially, PFR captures the intuitive, deterministic idea "if my distance from ES is small, then send, else do not send". This idea was enhanced by random decisions (above a threshold) to allow some local flooding to happen with small probability and thus to cope with local sensor failures.

Performance Evaluation. In [6] the authors prove the correctness of the PFR protocol, by using a geometric analysis: PFR always propagates data to the sink, under ideal network conditions (no failures), thus it is provably correct. Using properties of stochastic processes, it is shown that the protocol is very energy efficient. Also, when part of the network is inoperative (which is more realistic, because sensors are prone to faults), the protocol manages to propagate data very close to the sink, thus it is robust.

Note that the number of steps in the forwarding phase of the protocol depends on the $\phi_{threshold}$ of the protocol as it can be seen from the analysis in [6]. For $\phi_{threshold} = 134^o$ the number of flooding steps must be at least $180\sqrt{2}$ for correctness reasons. We can increase the $\phi_{threshold}$; this will increase also the number of flooding steps. This also implies a tradeoff between energy efficiency and robustness.

The energy efficiency of the PFR protocol is $\Theta\left(\left(\frac{n_0}{n}\right)^2\right)$ where $n_0 = |ES|$ and $n = \sqrt{N}$, where N is the number of particles in the network. For $n_0 = |ES| = o(n)$, this is $o(1)$. In order to prove the energy efficiency of PFR, let consider the area around the ES line, whose particles participate in the propagation process. The number of active particles is thus, roughly speaking, captured by the size of this area, which in turn is equal to $|ES|$ times the maximum distance from $|ES|$ (where maximum is over all active particles).

This maximum distance is clearly a random variable. To calculate the expectation and variance of this variable, the authors in [6] basically "upper bound" the stochastic process of the distance from ES by a random walk on the line, and subsequently "upper bound" this random walk by a well-known stochastic process (i.e. the "discouraged arrivals" birth and death Markovian process, see e.g. [16]).

In order to evaluate the robustness of PFR lets consider particles "very near" to the ES line. Clearly, such particles have large ϕ-angles (i.e. $\phi > 134^o$). Thus, even in the case that some of these particles are not operating, the probability that none of those operating transmits (during the probabilistic phase 2) is very small. In particular, in [6] it is shown that PFR manages to propagate the crucial data across lines parallel to ES, and of constant distance, with *fixed* nonzero probability (not depending on n, $|ES|$).

3.2 The Local Target Protocol (LTP)

We now present the LTP protocol [8] for wireless sensor networks. The basic idea of the protocol is to try to *search* for all active neighboring particles and in the sequence use the information retrieved in order to *forward* (i.e. propagate) the data towards the neighbor that is closer to the sink. In this protocol, each particle p' that has received *info($\mathcal{E}$)* from p (via, possibly, other particles) does the following:

> **Phase 1: The Search Phase.** It uses a periodic low energy broadcast of a beacon in order to discover a particle nearer to $\mathcal{S}$ than itself. Among the particles returned, p' selects a unique particle p'' that is "best" with respect to progress towards the sink, that is, the particle p''_E that among all particles found achieves the bigger progress on the $p'S$ line, should be selected.
>
> **Phase 2: The Direct Transmission Phase.** Then, p' sends *info($\mathcal{E}$)* to p'' and sends a *success* message to p (i.e. to the particle that it originally received the information from).
>
> **Phase 3: The Backtrack Phase.** If consecutive repetitions of the *search phase* fail to discover a particle nearer to $\mathcal{S}$, then p' sends *fail* message to the particle that it originally received the information from.

In the above procedure, propagation of *info($\mathcal{E}$)* is done in two steps; (i) particle p' locates the next particle (p'') and transmits the information and (ii) particle p' waits until the next particle (p'') succeeds in propagating the message further towards $\mathcal{S}$. This is done to speed up the backtrack phase in case p'' does not succeed in discovering a particle nearer to $\mathcal{S}$.

Note that one can estimate an a-priori upper bound on the number of repetitions of the search phase needed, by calculating the probability of success of each search phase, as a function of various parameters (such as density, search angle, transmission range). This bound can be used to decide when to backtrack.

Performance Evaluation. In [8], the "hops" efficiency of LTP is evaluated as a ratio of the number of transmissions required to reach the sink Sover the "optimal" (direct to S) transmissions needed to reach Sin the *ideal* case in which particles alwasy exist in pair-wise distances $\mathcal{R}$ on the vertical line from p to $\mathcal{S}$. Remark that $h_{opt} = \left\lceil \frac{d(p,\mathcal{S})}{\mathcal{R}} \right\rceil$, where $d(p,\mathcal{S})$ is the distance of p from the sink $\mathcal{S}$. Clearly, the number of hops (transmissions) needed characterizes the energy consumption and the time needed to propagate the information $\mathcal{E}$ to the sink.

In the case where the protocol Π is randomized, or in the case where the distribution of the particles in the cloud is a random distribution, the number of hops h and the efficiency ratio C_h are random variables and one wishes to study their expected values.

To enable a first step towards a rigorous analysis of smart dust protocols, [8] makes the following simplifying assumption: *The search phase always finds a* p'' (of sufficiently high battery) in the semicircle of center the particle p' currently possessing the information about the event and radius $\mathcal{R}$, in the direction towards $\mathcal{S}$. Note that this assumption on always finding a particle can be relaxed in the following ways: (a) by repetitions of the search phase until a particle is found. This makes sense if at least one particle exists but was sleeping during the failed searches, (b) by considering, instead of just the semicircle, a cyclic sector

defined by circles of radiuses $\mathcal{R} - \Delta\mathcal{R}$, $\mathcal{R}$ and also take into account the density of the smart cloud, (c) if the protocol during a search phase ultimately fails to find a particle towards the sink, it may *backtrack*.

[8] also assumes that the position of p'' is uniform in the arc of angle $2a$ around the direct line from p' vertical to $\mathcal{S}$. Each data transmission (one hop) takes constant time t (so the "hops" and time efficiency of our protocols coincide in this case). It is also assumed that each target selection is stochastically *independent* of the others, in the sense that it is always drawn uniformly randomly in the arc $(-\alpha, \alpha)$.

The above assumptions may not be very realistic in practice, however, they can be relaxed and in any case allow to perform a first effort towards providing some concrete analytical results.

Lemma 3.1 ([8]) *The expected "hops efficiency" of the local target protocol in the a-uniform case is*

$$E(C_h) \simeq \frac{\alpha}{\sin\alpha}$$

for large h_{opt}. Also

$$1 \leq E(C_h) \leq \frac{\pi}{2} \simeq 1.57$$

for $0 \leq \alpha \leq \frac{\pi}{2}$.

Proof: Due to the protocol, a sequence of points is generated, $p_0 = p$, p_1, $p_2, \ldots, p_{h-1}$, p_h where p_{h-1} is a particle within $\mathcal{S}$'s range and p_h is part of the sink. Let α_i be the (positive or negative) angle of p_i with respect to p_{i-1}'s vertical line to $\mathcal{S}$. It is:

$$\sum_{i=1}^{h-1} d(p_{i-1}, p_i) \leq d(p, \mathcal{S}) \leq \sum_{i=1}^{h} d(p_{i-1}, p_i)$$

Since the (vertical) progress towards $\mathcal{S}$ is then $\Delta_i = d(p_{i-1}, p_i) = \mathcal{R}\cos\alpha_i$, we get:

$$\sum_{i=1}^{h-1} \cos\alpha_i \leq h_{opt} \leq \sum_{i=1}^{h} \cos\alpha_i$$

From Wald's equation for the expectation of a sum of a random number of independent random variables (see [21]), then

$$E(h-1) \cdot E(\cos\alpha_i) \leq E(h_{opt}) = h_{opt} \leq E(h) \cdot E(\cos\alpha_i)$$

Now, $\forall i$, $E(\cos\alpha_i) = \int_{-\alpha}^{\alpha} \cos x \, \frac{1}{2\alpha} dx = \frac{\sin\alpha}{\alpha}$. Thus

$$\frac{\alpha}{\sin\alpha} \leq \frac{E(h)}{h_{opt}} = E(C_h) \leq \frac{\alpha}{\sin\alpha} + \frac{1}{h_{opt}}$$

Assuming large values for h_{opt} (i.e. events happening far away from the sink, which is the most interesting case in practice since the detection and propagation difficulty increases with distance) we have (since for $0 \leq \alpha \leq \frac{\pi}{2}$ it is $1 \leq \frac{\alpha}{\sin\alpha} \leq \frac{\pi}{2}$) and the result follows.

■

In order to further study the performance of LTP, [8] investigates the possibility where the protocol carries out more than one search phase and now assumes that these sequential phases always return *two points* p'', p''' each uniform in $(-\alpha, \alpha)$. They call this variation of LTP the "min-two uniform targets" protocol (M2TP). Now, the protocol selects the best of the two points, with respect to the local (vertical) progress. This is in fact an optimized version of the Local Target Protocol.

In a similar way as in the proof of the previous lemma, the authors prove the following result:

Lemma 3.2 ([8]) *The expected "hops efficiency" of the "min two uniform targets" protocol in the a-uniform case is*

$$E(C_h) \simeq \frac{\alpha^2}{2(1-\cos\alpha)}$$

for $0 \leq \alpha \leq \frac{\pi}{2}$ *and for large* h.

Now remark that

$$\lim_{\alpha\to 0} E(C_h) = \lim_{\alpha\to 0} \frac{2\alpha}{2\sin a} = 1$$

and

$$\lim_{\alpha\to \frac{\pi}{2}} E(C_h) = \frac{(\pi/2)^2}{2(1-0)} = \frac{\pi^2}{8} \simeq 1.24$$

Thus, [8] proves the following:

Lemma 3.3 ([8]) *The expected "hops" efficiency of the min-two uniform targets protocol is*

$$1 \leq E(C_h) \leq \frac{\pi^2}{8} \simeq 1.24$$

for large h *and for* $0 \leq \alpha \leq \frac{\pi}{2}$.

Remark that, with respect to the expected hops efficiency of the local target protocol, the min-two uniform targets protocol achieves, because of the one additional search, a relative gain which is $(\pi/2 - \pi^2/8)/(\pi/2) \simeq 21.5\%$.

4 Some Recent Work

In [9], the problem of *energy-balanced* data propagation in wireless sensor networks is studied. The energy balance property guarantees that the average per sensor energy dissipation is the same for all sensors in the network, during the entire execution of the data propagation protocol. This property is important since it prolongs the network's lifetime by avoiding early energy depletion of sensors.

They propose a *new algorithm* that in each step decides whether to propagate data one-hop towards the final destination (the sink), or to send data directly to the sink. This randomized choice balances the (cheap) one-hop transimssions with the direct transimissions to the sink, which are more expensive but *"bypass" the sensors lying close to the sink*. Note that, in most protocols, these close to the sink sensors tend to be overused and die out early.

By a detailed analysis they *precisely estimate* the probabilities for each propagation choice in order to guarantee energy balance. The needed estimation can easily be performed by

current sensors using simple to obtain information. Under some assumptions, they also derive a *closed form* for these probabilities.

The fact (shown by the analysis) that direct (expensive) transmissions to the sink are needed only rarely, shows that their protocol, besides energy-balanced, is *also energy efficient.*

In [3], the authors propose a new energy efficient and fault tolerant protocol for data propagation in smart dust networks, the Variable Transmission Range Protocol (VTRP). The basic idea of data propagation in VTRP is the varying range of data transmissions, ie. they allow the transmission range to increase in various ways. Thus data propagation in the protocol exhibits high fault-tolerance (by bypassing obstacles or faulty sensors) and increases network lifetime (since critical sensors, ie. close to the control center are not overused). As far as we know, it is the first time varying transmission range is used.

In [19] *extended versions* of two data propagation protocols are presented: the Sleep-Awake Probabilistic Forwarding Protocol (SW-PFR) and the Hierarchical Threshold sensitive Energy Efficient Network protocol (H-TEEN). These non-trivial extensions aim at improving the performance of the original protocols, by introducing *sleep-awake periods* in the PFR protocol to save energy, and introducing a *hierarchy of clustering* in the TEEN protocol to better cope with large networks areas. b) They have *implemented* the two protocols and performed an *extensive experimental comparison* (using simulation) of various important measures of their performance with a focus on energy consumption. c) They investigate in detail the *relative advantages and disadvantages* of each protocol and discuss and explain their behavior. d) In the light above they propose and discuss a possible *hybrid combination of the two protocols* towards optimizing certain goals.

Recently, [4] propose a novel and efficient energy-aware distributed heuristic, which they refer to as EAD, to build a special rooted broadcast tree with many leaves that is used to facilitate data-centric routing in wireless microsensor networks. EAD algorithm makes no assumption on local network topology, and is based on residual power. It makes use of a *neighboring broadcast scheduling* and *distributed competition among neighboring nodess.*

EAD basically computes a tree with many leaves. With the transceivers of all leaf nodes being turned off, the network lifetime can be greatly extended. In [4] EAD scheme is implemented and an extensive simulation experiments is conducted to study the its performance. The experimental results indicate clearly that EAD scheme outperforms previous schemes, such as LEACH among other protocols.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, *A survey on sensor networks*, IEEE Communications Magazine (2002), 102–114.

[2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, *Wireless sensor networks: a survey*, Journal of Computer Networks **38** (2002), 393–422.

[3] T. Antoniou, A. Boukerche, I. Chatzigiannakis, G. Mylonas, and S. Nikoletseas, *A new energy efficient and fault-tolerant protocol for data propagation in smart dust networks using varying transmission range*, 37th Annual Simulation Symposium (ANSS 2004), 2004, IEEE Press, pp. 43–52.

[4] A. Boukerche, X. Cheng, and J. Linus, *Energy-aware data-centric routing in microsensor networks*, ACM Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2003) (Paris, France), 2003, pp. 42–49.

[5] A. Boukerche, R.W.N. Pazzi, and R.B. Araujo, *A supporting protocol to periodic, event-driven and query-based application scenarios for critical conditions surveillance*, 1st International Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS 2004), Springer-Verlag, 2004, Lecture Notes in Computer Science, LNCS 3121, pp. 137–146.

[6] I. Chatzigiannakis, T. Dimitriou, S. Nikoletseas, and P. Spirakis, *A probabilistic forwarding protocol for efficient data propagation in sensor networks*, 5th European Wireless Conference on Mobile and Wireless Systems beyond 3G (EW 2004), 2004, pp. 344–350.

[7] I. Chatzigiannakis and S. Nikoletseas, *A sleep-awake protocol for information propagation in smart dust networks*, 3rd International Workshop on Mobile, Ad-hoc and Sensor Networks (WMAN 2003), 2003, IPDPS Workshops, p. 225.

[8] I. Chatzigiannakis, S. Nikoletseas, and P. Spirakis, *Smart dust protocols for local detection and propagation*, 2nd ACM International Annual Workshop on Principles of Mobile Computing (POMC 2002), 2002, pp. 9–16.

[9] C. Efthymiou, J. Rolim, and S. Nikoletseas, *Energy balanced data propagation in wireless sensor networks*, 4th International Workshop on Mobile, Ad-hoc and Sensor Networks (WMAN 2004), 2004, IPDPS Workshops, p. 225.

[10] W. Feng, B. Code, C. Kaiser, M. Shea, and W. Feng, *Panoptes: Scalable low-power video sensor networking technologies*, ACM Multimedia, 2003, pp. 562–571.

[11] C. Gui and P. Mohapatra, *Power conservation and quality of surveilance in target tracking sensor networks*, 10th ACM/IEEE Annual International Conference on Mobile Computing (MOBICOM 2004) (Philadelphia, USA), 2004, pp. 129–143.

[12] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, *Energy-efficient communication protocol for wireless microsensor networks*, 33rd IEEE Hawaii International Conference on System Sciences (HICSS 2000), 2000, p. 8020.

[13] S.E.A. Hollar, *Cots dust*, Msc. thesis, Engineering-Mechanical Engineering, University of California, Berkeley, USA, 2000.

[14] C. Intanagonwiwat, R. Govindan, and D. Estrin, *Directed diffusion: A scalable and robust communication paradigm for sensor networks*, 6th ACM/IEEE Annual International Conference on Mobile Computing (MOBICOM 2000), 2000, pp. 56–67.

[15] B. Karp, *Geographic routing for wireless networks*, Ph.D. dissertation, Harvard University, Cambridge, USA, 2000.

[16] L. Kleinrock, *Queueing systems, theory*, vol. I, John Wiley & Sons, 1975.

[17] W. Kumwilaisak, Y.T. Hou, Q. Zhang, W. Zhu, C.C.J. Kuo, and Y.Q. Zhang, *A cross-layer quality of service mapping architecture for video delivery in wireless networks*, IEEE Journal on Selected Areas in Communications **21** (2003), no. 10, 1685–1689.

[18] A. Manjeshwar and D.P. Agrawal, *TEEN: A routing protocol for enhanced efficiency in wireless sensor networks*, 2nd International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing (WPIM 2002), 2002, IPDPS Workshops, p. 195b.

[19] S. Nikoletseas, I. Chatzigiannakis, H. Euthimiou, A. Kinalis, T. Antoniou, and G. Mylonas, *Energy efficient protocols for sensing multiple events in smart dust networks*, 37th Annual Simulation Symposium (ANSS 2004), 2004, IEEE Press, pp. 15–24.

[20] C.E. Perkins, *Ad hoc networking*, Addison-Wesley, Boston, 2001.

[21] S.M. Ross, *Stochastic processes*, 2nd ed., John Wiley and Sons, Inc., 1995.

[22] C. Schurgers, V. Tsiatis, S. Ganeriwal, and M. Srivastava, *Topology management for sensor networks: Exploiting latency and density*, 3rd IEEE/ACM Annual International Symposium on Mobile Ad Hoc Networking & Computing (MOBIHOC 2002), 2002, pp. 135–145.

Variability of the infection time in scale-free networks

Pascal Crépey,[1,2,*] Fabián Alvarez,[1,2,†] and Marc Barthélemy[3,‡]

[1]*INSERM, Unité de Recherche "Épidémiologie, Systèmes d'Information et Modélisation" (U707), Paris, F-75012, France*

[2]*Université Pierre et Marie Curie, Faculté de Médecine Pierre et Marie Curie, Unité de Recherche "Épidémiologie, Systèmes d'Information et Modélisation", Paris, F-75012, France*

[3]*CEA-Centre d'Etudes de Bruyères-le-Châtel, Département de Physique Théorique et Appliquée BP12, 91680 Bruyères-Le-Châtel, France*

Abstract

Recent studies suggest that a large number of natural and artificial networks are characterized by very large degree fluctuations. This result means that a non-negligible number of nodes are extremely well connected while the majority just have a few links. The effect of such large fluctuations can be dramatic as illustrated by the fact that infectious agents can spread on these networks even for a very small value of the transmission probability.

Another consequence is that random immunization is inefficient for this kind of networks. Consequently, containment protocols and vaccination of traced contacts become our only defense but are unfortunately difficult to use at a large scale and in this context, an efficient method for detecting epidemics during their early stage becomes imperative. These efforts, now broadly labeled as "syndromic surveillance" are the centers of attention of public health agencies concerned with bioterrorism-related diseases. More precisely, an important point is to be able to determine and characterize specific nodes in the network which display interesting features in regards of an early detection system. In the present work, we focus on two specific features: (i) a small average infection time t_{inf} and (ii) low fluctuations around that time t_{inf}.

We analyze and compare the behaviors of the infection time obtained for the usual random Erdos-Renyi graph and the Barabasi-Albert scale-free network. We first compare the patterns obtained on both kind of networks and we then describe the variations of the infection time with the degree and with the topological distance to the initially infected node.

Keywords: scale-free networks;epidemic spreading;variability

*Corresponding author: ESIM - INSERM U707, CHU Saint-Antoine, 27 rue Chaligny, 75571 Paris cedex 12, France
Tel : +33 1 44 73 86 69; Fax : +33 1 44 73 84 54; Electronic address: pascal.crepey@u707.jussieu.fr
†Electronic address: fabian.alvarez@u707.jussieu.fr
‡Electronic address: marc.barthelemy@cea.fr

I. INTRODUCTION

Numerous studies have shown that many natural and artificial networks (e.g. air transportation, sexual contact, coauthorship, ...) have a distribution of their degree k which follows a power law ($P_k \sim k^{-\gamma}$) [1–5]. This implies the absence of a characteristic scale of their degree distribution, hence their name "scale-free network" (SFN) [6, 7].

Classically, epidemic modelisation is based on the homogeneous mixing paradigm. All individuals have the same probability to become infected and no heterogeneity in the contact network is taken into account.

However, superspreading events like the ones that appeared in the onset of the recent SARS outbreak [8], cannot be explained by using the previous paradigm. Consequently, they make very difficult the realistic estimation from initial data of the outbreak behavior [8, 9]. A solution to this issue could lie on the particular degree distribution of SFN which renders some features of social networks by bringing on a non-negligible number of highly connected nodes, called hubs or superspreaders. Therefore, from a public health point of view, studying the spreading of epidemics on SFN is all the more appropriate.

This issue, whose applications extend to new emergent diseases and bioterrorist threats, justify detailed studies of the incidence of the connectivity distribution at the initial stage of the epidemics. In particular, given the stochastic nature of epidemic processes, we focus our study on the characterization and the understanding of their variability. The variability plays an important role in the accuracy of models. Thus, it has to be quantified to assess the meaningfulness of simulations with respect to real outbreaks.

Concerning the epidemiological modeling, the simplest approach is to consider that infected individuals (I) may infect susceptible (S) ones with probability λ, which will then remain infected (SI model). This approach, in spite of its simplicity, allows to easily outline the initial growth of epidemic outbreaks.

Using a numerical approach, we analyzed the evolution of epidemics generated by different sets of initial parameters. Hence, we compared epidemic variability on a SFN (Barabási-Albert (BA) [10]) to an homogeneous network model (Erdös-Renyi (ER) [11]) in order to highlight the singularities due to the connectivity distribution.

II. INFECTION TIME FLUCTUATIONS

We consider that a good picture of the predictability of epidemic path is given by the variation of t_{inf} on each node. This information tells us how systematically nodes will be infected at regular moments during the outbreak. In this study, we characterize it by computing the t_{inf} coefficient of variation, $CV(t_{inf}) = \frac{\sqrt{\langle(t_{inf}-\langle t_{inf}\rangle)^2\rangle}}{\langle t_{inf}\rangle}$, on a set of a thousand of outbreaks simulated on the same network.

In Fig. 1, we show nodes $\langle t_{inf}\rangle$ (left panel) and $CV(t_{inf})$ (right panel) for an ER and a BA network. In those plots, symbols aligned vertically represent nodes with a given degree k. By comparing Fig. 1 panels, we clearly notice that while the averaged t_{inf} on BA is lower than on ER network, its corresponding $CV(t_{inf})$ is larger. From a practical point of view, it means that, on one hand, BA nodes are infected more quickly (smaller t_{inf} values), but on the other hand, the moment t_{inf} will vary more from an epidemic to another (greater $CV(t_{inf})$), and consequently will be less predictable.

Figure 1 also reveals the tendency of $CV(t_{inf})$ to increase with the nodes degree, and the high values reached by high degree nodes on BA network.

On BA networks, low degree nodes are large majority. Their wide range of $CV(t_{inf})$ values tells us that the degree may not be the most relevant discriminant and that values computed over the same degree (lighter symbols) are not representative when dealing with small degree nodes. Consequently, variations of t_{inf} can be less easily predicted for low degree nodes. On the contrary, higher degrees nodes have values following a more clearly drawn slope, giving $CV(t_{inf})$ as a function of their degree more meaning.

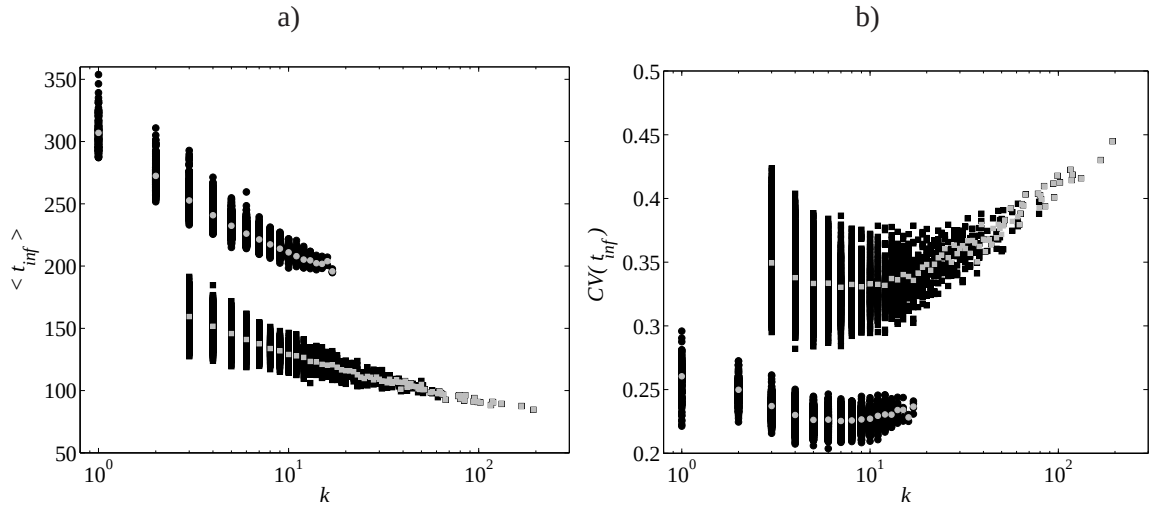


Figure 1: Infection time of network nodes as a function of their degree. **(a)** Averaged infection time: Each black symbol represents t_{inf} for a single node averaged over 10^3 outbreaks. Gray symbols show t_{inf} averaged over nodes with the same degree. **(b)** Infection time coefficient of variation: each black symbol stands for $CV(t_{inf})$ computed over 10^3 outbreaks. Gray symbols show $CV(t_{inf})$ computed over nodes with the same degree. For both panels, ■ are values on BA network, and • are values on ER network; results are computed on a single network ($N = 10^4$, $\langle k \rangle = 6$); initially infected node degree $k_0 = 6$; $\lambda = 0.01$.

III. CONCLUSIONS

In this short paper, we address the concern about the reliability and the efficiency of detection sites. A good site candidate has to be on the outbreak path as soon and as surely as possible. Here, we point out that, despite their low delay before infection (t_{inf}), superspreading nodes exhibit high variations of their moment of infection. In other words, we may not be able to predict an accurate and reliable time of infection for them. As a consequence, high degree nodes should be used with caution in the set-up of an early epidemic detection system.

We also draw attention to the differences between ER and BA networks concerning variability. As depicted on Fig. 1b, the homogeneous nature of ER networks seems to prevent singularities in the spreading, and thus tend to lower the $CV(t_{inf})$ of their nodes. Infection velocities, which can be deduced from Fig. 1a, also differ and are in concordance with previous studies [12]. These spreading disparities confirm that outbreaks on the two topologies behave distinctly in many points and ignoring their differences may cause problems.

In particular, our results stress that the epidemic variability is amplified on scale-free BA network compared to homogeneous ER network. This phenomenon has a practical importance for the modeling of epidemic control strategies, especially during the beginning of the spreading process which is a highly unstable period due to high degree nodes. In particular, this study suggest that this high variability should be taken into account in planning and designing control and containment strategies.

Acknowledgments

P.C. acknowledges financial support from A.C.I. Systèmes Complexes en Sciences Humaines et Sociales, and F.A. from FRM (Fondation pour la Recherche Médicale).

[1] Barabási A-L, Albert R, Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web. *Phys A* 2000; 281:69–77.
[2] SN Dorogovtsev, JFF Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW.* Oxford University Press, Oxford, 2003.
[3] Jeong H, Mason SP, Barabási A-L, Oltvai ZN. Lethality and centrality in protein networks. *Nature* 2001; 411:41–2.
[4] Pastor-Satorras R, Vespignani A. *Evolution and Structure of the Internet: A Statistical Physics Approach.* Cambridge University Press, Cambridge, 2004.
[5] Schneeberger A, Mercer CH, Gregson SA, Ferguson NM, Nyamukapa CA, Anderson RM, Johnson AM, Garnett GP. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sex Transm Dis* 2004; 31:380–7.
[6] Amaral LAN, Scala A, Barthélemy M, Stanley HE. Classes of small-world networks. *PNAS* 2000; 97:11149–52.
[7] Barabási A-L, Albert R. Emergence of scaling in random networks. *Science* 1999; 286:509–12.
[8] Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D, Murray M. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003; 300:1966–70.
[9] Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, Leung GM, Ho L-M, Lam T-H, Thach TQ, Chau P, Chan K-P, Lo S-V, Leung P-Y, Tsang T, Ho W, Lee K-H, Lau EMC, Ferguson NM, Anderson RM. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 2003; 300:1961–6.
[10] Albert R, Barabási A-L. Statistical mechanics of complex networks. *Rev Modern Phys* 2002; 74:40–97.
[11] Erdös P, Renyi A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* 1960; 5:17–61.
[12] Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys Rev Lett* 2004; 92:178701.

Partitioning networks into classes of mutually isolated nodes *

J. Díaz[1], A.C. Kaporis[2], L.M. Kirousis[2,3], and X. Pérez[1]

[1] Universitat Politècnica de Catalunya, Departament de Llenguatges i Sistemes Informàtics, Campus Nord – Ed. Omega, 240, Jordi Girona Salgado, 1–3, E-08034 Barcelona, Catalunya, {diaz, xperez}@lsi.upc.edu

[2] University of Patras, Department of Computer Engineering and Informatics, GR-265 04 Patras, Greece, {kaporis, kirousis}@ceid.upatras.gr

[3] Research Academic Computer Technology Institute, P.O. Box 1122, GR-261 10 Patras, Greece

Abstract. Following the work of Newman, we model complex networks with random graphs of a given degree distribution. In particular we study the case where all vertices have exactly the same degree (random regular graphs). We survey some recent results on the problem of partitioning such a graph into the smallest possible number of classes of mutually isolated vertices, known in graph theory as the colorability problem. We also describe the solution of an open problem about 5-regular graphs.

1 Introduction

It is not an exaggeration to state that Random Graph Theory is undergoing a revolutionary change. Graph theorists have definitely been expelled from the paradise that Erdős and Rényi [6, 7] had created for them. According to this classical theory, a random graph is the product of repeated Bernoulli trials: for each pair of vertices, decide with a given probability, if they will be connected by an edge or not. Classical random graphs however, despite the deep and complicated mathematics devised for their study, are not of much interest from the complex systems point of view. Local decisions that affect the global picture, emergent properties, self-organization, robustness, resilience, to name only a few issues of interest to complexity science, do not pertain to classical graph theory.

As is well known by now, this state of affairs radically changed just before the turn of the century, when researchers tried to study from a graph-theoretical point of view existing complex systems, live or artificial. The pioneering work of e.g., Watts and Strogatz [12, 13], and Barabási and Albert [4] gave a new twist to classical graph theory. Random graphs according to this new paradigm have properties, like power law distribution of the degrees, or freedom from scale, that reflect the interaction patterns of many complex systems.

* Research partially supported by Future and Emerging Technologies programme of the EU under contract 001907 *"Dynamically Evolving, Large-Scale Information Systems (DELIS)"*. The 1st author was also partially supported by the Distinció de la Generalitat de Catalunya per a la promoció de la recerca, 2002. The 2nd and 3rd authors are also partially supported by European Social Fund (ESF), Operational Program for Educational and Vacational Training II *(EPEAEK II)*, and particularly *Pythagoras*. Part of the research of the 3nd author was conducted while visiting on a sabbatical the Departament de Llenguatges i Sistemes Informàtics of the Universitat Politècnica de Catalunya.

One approach to such models was the one undertaken by Newman et al. [10]. In this approach, instead of modelling a random graph with an edge generating process with a specified attachment rule at each step (generative models), one assumes instead that the graph is uniformly random conditional on its degree sequence. Certainly, this static approach is more restricted in comparison to the generative approach. However, it is often more suitable for the analysis of graph properties like clustering, distribution of component sizes, resilience, thresholds etc.

In this paper, we restrict our study to an important subclass of random graphs with a given degree sequence: that of random regular graphs, ie graphs where all vertices have the same degree. A comprehensive account of the pioneering work on random regular graphs can be found in the paper by Wormald [14]. The present paper concerns a much studied graph parameter, that of the chromatic number of a graph. The chromatic number is defined to be the smallest number of colors needed to color the vertices of a graph, given that no two adjacent vertices get the same color. In network terminology, the chromatic number is the smallest number of classes that one can partition the nodes of a network, so that each class comprises only of mutually isolated nodes.

Molloy [9] proved that 6-regular graphs have chromatic number at least 4, asymptotically almost surely (a.a.s.) with respect to the number of their vertices n. Achlioptas and Moore [1] proved that 4-regular graphs have chromatic number 3 with constant probability. Subsequently, Achlioptas and Moore [2] showed that a.a.s., the chromatic number of a d-regular graph ($d \geq 3$) is k or $k+1$ or $k+2$, where k is the smallest integer such that $d < 2k \log k$. Shi and Wormald [11] showed that a.a.s. the chromatic number of a 4-regular graph is 3, that a.a.s. the chromatic number of a 6-regular graph is 4 and that a.a.s. the chromatic number of a 5-regular graph is either 3 or 4. They also showed that a.a.s. the chromatic number of a d-regular graph, for all other d up to 10, is restricted to a range of two integers. Their proofs were algorithmic.

The above results leave open the question of whether the chromatic number of a 5-regular graph can take the value 3 with constant probability (or perhaps even a.a.s.). The difficulty of devising a rigorously analyzable algorithm that provides a 3-coloring for a 5-regular graph with constant probability was explained by research in physics. Krząkała et al. [8] provided a non-rigorously analyzable algorithm that a.a.s. finds a 3-coloring for a 5-regular graph. However, they also showed that the space of assignments of three colors to the vertices (legal or not, i.e. with no two adjacent vertices with the same color or not) consists of clusters of *legal* color assignments inside which one can move from point to point by steps of small Hamming distance. Additionally they showed that to go from one cluster to another by such small steps, it is necessary to go through assignments of colors that grossly violate the requirement of legality (high-energy color assignments). Moreover, the number of clusters that contain points with energy that is a local, but not global, minimum is exponentially large. As a result, local search algorithms are easily trapped into such local minima (metastable states). Non-local search algorithms however are usually not amenable to rigorous analysis.

The above considerations left as the only plausible alternative to try to prove that 5-regular graphs are 3-colorable with constant probability in an analytic way. A technique that has been used towards similar ends is the Second Moment Method. The basic ingredient of this method is the fact that if X is a non-negative random variable (r.v.) then the probability that X is positive is bounded from below from the ratio of the square of its

first moment to its second moment. We call this ratio the *Moment Ratio*. Symbolically:

$$\mathbf{Pr}\left[X > 0\right] \geq \frac{(\mathbf{E}\left[X\right])^2}{\mathbf{E}\left[X^2\right]}. \tag{1}$$

This technique was used [3] to solve the long-standing open problem of computing the two possible values of a random graph. In this work the authors considered as X the r.v. that gives the number of balanced 3-colorings of a graph (balanced are the colorings where there is an equal number of vertices of each color). The same r.v. however proved not to work for the case of 5-regular graphs.

In a recent work [5], the r.v. that counts *stable* balanced 3-colorings was considered. These are balanced 3-colorings with the property that for no single vertex v can one change its color without the appearance of an edge with the same color at its endpoints.

It turned out that for stable balanced 3-colorings the variance diminishes (as expected) but the square of the expectation diminishes in a lesser degree and as a result the Moment Ratio becomes asymptotically positive. Thus a non-empirical proof that 5-regular graphs are 3-colorable with positive probability was obtained.

References

1. D. Achlioptas and C. Moore. Almost all graphs with degree 4 are 3-colorable. *Journal of Computer and Systems Sciences* 67(2), 441–471, 2003.
2. D. Achlioptas and C. Moore. The chromatic number of random regular graphs. In: *Proc. 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX) and 8th International Workshop on Randomization and Computation (RANDOM)* (Springer, LNCS, 2004) 219–228.
3. D. Achlioptas and A. Naor. The two possible values of the chromatic number of a random graph. In: *36th Symposium on the Theory of Computing (STOC)*, 587–593, 2004.
4. A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286, 509–512, 1999.
5. J. Díaz, G. Grammatikopoulos, A.C. Kaporis, L. Kirousis, X. Pérez and D.G. Sotiropoulos. 5-Regular graphs are 3-colorable with positive probability. To appear in: *Proc. 13th Annual European Symposium on Algorithms (ESA)* (Springer, LNCS, 2005).
6. P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6, 290–297, 1959.
7. P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 17–61, 1960.
8. F. Krząkała, A. Pagnani and M. Weigt. Threshold values, stability analysis and high-q asymptotics for the coloring problem on random graphs. *Phys. Rev. E* 70, 046705 (2004).
9. M. Molloy. *The Chromatic Number of Sparse Random Graphs.* Master's Thesis, University of Waterloo, 1992.
10. M.E.J. Newman, S.H. Strogatz and D.J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64, 026118, 2001.
11. L. Shi and N. Wormald. *Colouring random regular graphs* Research Report CORR 2004-24, Faculty of Mathematics, University of Waterloo, 2004.
12. D.J. Watts *Small Worlds* Princeton University Press, 1999
13. D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442, 1998.
14. N.C. Wormald. Models of random regular graphs. In: J.D. Lamb and D.A. Preece, eds., *Surveys in Combinatorics* (London Mathematical Society Lecture Notes Series, vol. 267, Cambridge U. Press, 1999) 239–298.

Clustering and Robustness in Networks

Y Grondin & D J Raine[1]
Centre for Interdisciplinary Science, Department of Physics, University of Leicester

Clustering and Robustness in Networks

[1]D J Raine, Department of Physics and Astronomy, University Road, Leicester, LE1 7RH, United Kingdom

Keywords: Clustering, Robustness, Network structure.

Extended Abstract.
The robustness of a network is defined as the integrity of the network to attack through the deletion of network nodes. Such attacks can be random or targeted: in the former case nodes are selected for removal at random; in the latter one usually chooses to remove the most highly connected nodes at each stage, either systematically or with some enhanced probability. As nodes are removed the size of the largest connected component is tracked. In a robust network the size of this component decreases slowly as nodes are removed; a fragile network rapidly breaks into sub-networks.

The connectivity distribution of a network p_k is the probability that a randomly chosen node of the network has degree k, that is has k links to or from the node. Ever since Albert *et al* [1] showed that the robustness of networks depends on the distribution of connectivity and the mode of attack, random or targeted, there have been numerous attempts [2, 3, 4] to understand the relationship between the architecture of a network and its robustness.

The impression from the seminal paper of Albert *et al* [1], which compares Erdös-Rényi random networks [5], with a Poisson degree distribution, with a class of scale-free networks [6], with their power law distribution of connectivity, is that robustness is associated with the connectivity distribution.

Newman [7] argues that the robustness to node removal is related to what is referred to as the assortative behaviour, or assortativity, of the network. This is defined as the preference for nodes of large degrees of connectivity to be attached to each other. It is measured by the assortative coefficient, r. To define r, let e_{ij} be the joint probability distribution of the degrees of the nodes at the ends of a randomly chosen link, not counting this link itself in the nodal degrees [8]. Then r ($-1 \le r \le 1$) is given by

$$r \propto \frac{\sum_{ij} ij(e_{ij} - q_i q_j)}{\left(\sum_k k^2 q_k - \sum_k k q_k\right)},$$

where the normalised 'remaining degree' distribution [4, 6] q_k is

$$q_k = \frac{(k+1)\,p_{k+1}}{\sum_j j p_j}.$$

The coefficient r is positive for assortative networks and negative otherwise. Thus sociological networks ($r > 0$) appear to be more robust than biological or physical networks ($r < 0$). However, since both random and scale-free networks have the same assortativity but react differently to node removal, the assortative behaviour cannot be the only criterion for the robustness of a network.

By investigating networks having the same assortativity but different degrees of clustering, we show here that robustness depends also on the cluster coefficient of a network.

Clustered networks
The cluster coefficient is defined as the degree of clustering of a node averaged over all the nodes of the network. It is given by

$$\langle C \rangle = N^{-1} \sum_i \frac{T_i}{2k_i(k_i - 1)},$$

with T_i the number of triangle from node i, k_i the degree of connectivity of node i, and N the number of nodes in the network.

We begin by constructing networks with a fixed connectivity, k, and assortativity, of an ordered network of nodes arranged on a circle. We vary the cluster coefficient via a 'cross-rewiring' operation. It leads to a sequence of networks with decreasing degrees of clustering. We call these networks clustered k-regular networks, or kC-networks for short.

Robustness
We extract networks along the rewiring process and test their robustness to the removal of nodes at random or to the continued targeted (systematic) removal of the most highly connected nodes. In the latter case, if several nodes have the same degree of connectivity, one is targeted at random.

In the case of an attack on the most connected nodes of a kC-network, the critical value at which the network breaks down increases as the clustering coefficient diminishes. This is somewhat counter-intuitive as one might expect that a network with larger cluster coefficient provides more redundancies of pathways between any pair of nodes. The results show the opposite, because rewiring links to form clusters reduces the number of links associated with the large scale resilience of the network.

This behaviour is confirmed when the nodes are removed at random: kC-networks with a very low cluster coefficient are more resilient to node failure than their highly clustered counterparts.

Aside from the fact that these networks are more or less robust according to the clustering coefficient, they always break down when the most connected nodes are removed. On the contrary, the random removal of nodes from a network with a small value of C (<0.275) does not break it down but leads to a steady erosion of the largest component.

Note there exists a value of C ($0.370 < C < 0.470$) for which the network behaves in the same way to random or targeted removal.

Random and scale-free networks
Random and scale-free network present a range of connectivity that requires us to set a constraint on the cross-rewiring operation in order to maintain the assortativity constant.

We study the behaviour of random and scale-free networks for various clustering coefficients as above. The random network is an Erdös and Rényi network, where pairs of nodes are connected with probability, p. The scale-free network is generated using preferential attachment in the context of growing networks. By construction, both networks have an assortative coefficient $r = 0$.

For both networks, there is still an effect of the clustering coefficient on their resilience. Networks with smaller clustering coefficients are more robust than their counterparts, in either random or targeted removal of the nodes. The main difference from our previous case is that the random removal of nodes no longer has a critical effect: the main connected component loses only few nodes at a time and does not fall apart.

As observed previously [1], the critical values at which the network breaks down is lower for the scale-free network than for the random one.

Along with assortativity, the clustering coefficient is an important parameter in determining the architectural robustness of networks. Networks with smaller clustering coefficient are more robust than networks with larger one. As a consequence, the small world property, which is associated with large values of C and which is a feature of many real networks [9, 10], gives rise to more fragile networks.

[1] Albert R, Jeong H, Barabasi A-L: Error and attack tolerance of complex networks, Nature 2000; 406: 378
[2] Dunne J A, Williams R J, Martinez N D: Network structure and biodiversity loss in food webs: robustness increases with connectance. Proc Natl Acad Sci USA 2002; 99: 12917
[3] Stelling J, Klamt S, Bettenbrock K, Shuster S, Gilles E D: Metabolic network structure determines key aspects of functionality and regulation. Nature 2002; 420: 190
[4] Callaway D S, Newman M E J, Strogatz S H, Watts D J: Network robustness and fragility: percolation on random graphs. Phys Rev Lett 2000; 85: 5468.

[5] Erdös P, Rényi A: On the evolution of random graphs, Publ Math Inst Hung Acad Sci 1960; 5: 17.
[6] Barabasi A-L, Albert R: Emergence of scaling in random networks. Science 1999; 286: 509.
[7] Newman M E J: Assortative mixing in networks, Phys Rev Let 2002; 89: 208701.
[8] Callaway D S, Hopcroft J E, Kleinberg J M, Newman M E J, Strogatz S H: Are randomly grown graphs really random? Phys Rev E 2001; 64: 041902.
[9] Amaral L A N, Scala A, Barthélémy M, Stanley H E: Classes of small world networks. Proc Natl Acad Sci USA 2000; 97: 11149.
[10] Wagner A, Fell D A: The small world inside large metabolic networks. Proc Roy Soc London Ser B 2001; 268:1803.

A Generative Model of Power Law Distributions with Optimizing Agents with Constrained Information Access

Extended Abstract

Laszlo Gulyas

AITIA International Inc.
H-1039, Budapest, Czetz Janos utca 48-50.
Fax: +36 1 4538081, e-mail: lgulyas@aitia.ai

Short Title: A Generative Model of Power Laws

Abstract

US metropolitan areas show an intriguing empirical regularity. According to statistical surveys, the population of a city is inversely proportional to its rank (by population). This type of power law relationships are of special interest to complex systems studies, because they are surprisingly common in natural and man-made systems. Therefore, a number of general, abstract mechanisms and generative models have been proposed to explain the occurrence of power law distributions, independent of the particular domain they may occur in. These models make different assumptions of the system, its actors and their dynamics. In this paper a new generative model is presented, based on utility maximizing behavior of agents using limited local information. This model is capable of generating power law distributions given sufficient heterogeneity is present in the distribution of information among the agents. The model is related to the city formation model of Simon that has become more widely known in its variant by Albert and Barabási as the 'preferential attachment' model. The main difference between these models and the one presented in this paper is that the former models require global access to information (i.e., the arriving new agent or node has to assess the distribution of links/size of the whole pre-existing population), while our model operates with limited information and utility maximization subject to this set of information.

Keywords

Power Law, Scale-Free Distribution, Constrained Information Access, Limited Information, Heterogeneity, Generative Model

Introduction

US metropolitan areas show an intriguing empirical regularity. According to statistical surveys, the population of a city is inversely proportional to its rank (by population). If x_q denotes the size of the q^{th} largest city, then x_q equals $a \cdot r^{-1}$, where a is the size of the largest city. This type of power law relationship is surprisingly common in many complex systems as testified by a vast and growing body of work in the literature of various disciplines. [3][4][5][6][7][8][9][10][11][14]

Power laws are of special interest to complex systems studies. A number of general, abstract mechanisms and generative models have been proposed to explain the occurrence of power law

distributions, independent of the particular domain they may occur in. These models make different assumptions of the system, its actors and their dynamics. In this paper a new generative model is presented, based on utility maximizing behavior of agents using limited local information. This model is capable of generating power law distributions given sufficient heterogeneity is present in the distribution of information among the agents. The model is related to the city formation model of Simon [12][13] that has become more widely known in its variant by Albert and Barabási as the 'preferential attachment' model. [1][2] The main difference between these models and the one presented in this paper is that the former models require global access to information (i.e., the arriving new agent or node has to assess the distribution of links/size of the whole pre-existing population), while our model operates with limited information and utility maximization subject to this set of information.

The Model

Let x_1^t, x_2^t, …, $x_L^t \in \mathbb{N}$ be positive integers for $t \geq 0$ and $L \in \mathbb{N}$, so that $x_l^0 = 1$, for all $l \in [1, L]$. At each time t, we take a random sample of x_l^t's without replacement. S^t will stand for the size of the sample, a random positive integer drawn uniformly from $[1,L]$: $S^t \in \mathbf{U}[1, L]$, for all $t \geq 0$. Z^t will stand for the sample itself:

$$(1) \qquad Z^t = \left\{ x_l^t \text{ such that } l \in \mathbf{U}[1, L] \right\} \text{ where } |Z^t| = S^t, \text{ for all } t \geq 0.$$

Let μ^t denote the largest element in the sample:

$$(2) \qquad \mu^t = \arg\max_{l \in Z^t} x_l^t, \; t \geq 0.$$

We will call μ^t as the *selected element* of the sample. If μ^t is not unique, we take the one with the lowest or highest index l, or simply a random 'maximum'. The dynamics of the system is the following:

$$(3) \qquad x_l^{t+1} = \begin{cases} x_l^t & , \text{if } l \neq \mu^t \\ x_l^t + \gamma & , \text{if } l = \mu^t \end{cases}, \text{ for all } t \geq 0, \text{ where } \gamma \in \mathbb{N}.$$

That is, the selected element of the sample is increased by a constant γ, while all the other elements remain the same.

We are interested in the properties of x_l^t, in the limit of $t \to \infty$. Especially, we are interested in the following 'frequency distribution': $FD^t(y) = \sum_{l \in [1,L]} \chi\left(x_l^t \geq y\right)$, where $\chi(.)$ is the membership function, and thus $FD^t(y)$ gives the number of x_l^t's that are at least the size of y at time t.

Numerical Results

Numerical simulations suggest that $\lim_{t \to \infty} FD^t(y) \sim y^{\alpha}$ with $\alpha \approx -2$, as shown by Figure 1. The figure shows the emerged power law distribution for L=100, 1000 and 10000 after t=1000·L iterations on a log-log scale. The linear approximation of the L=1000 case is also included for convenience, together with its goodness-of-fit (R^2) value.

Figure 2 provides a summary of the sensitivity analysis. The sensitivity analysis consisted of creating 10 simulation runs with different pseudo random number sequences for each parameter combination (L=100, 1000, 10000, 100000). The resulting frequency distribution of each individual run was saved after t=1000000 iterations and linear regression was applied to their log-log scale transformations. The figure summarizes the goodness-of-fit (R^2) values obtained. It is clear that the results are independent of the particular pseudo random sequence used. On the other hand, the fit is quite

naturally dependent on the number of iterations (t). The apparent drop in goodness-of-fit at L≥10000 is due to the insufficient number of iterations.

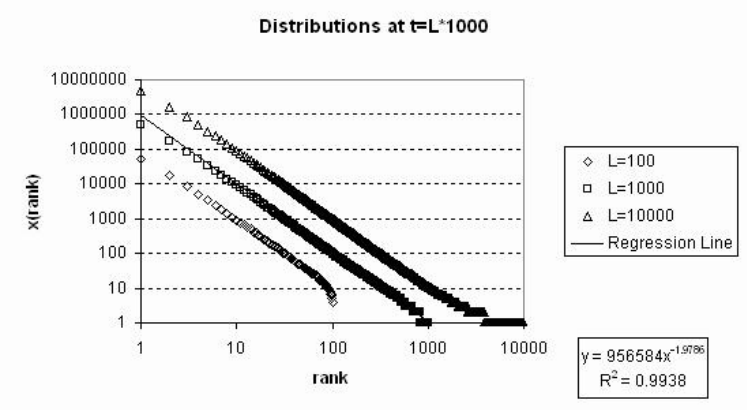


Figure 1: The emerging power law distributions for L=100, 1000 and 10000.

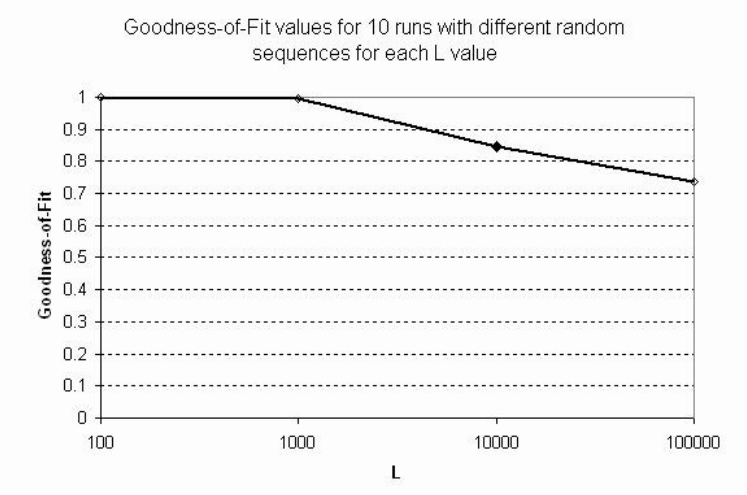


Figure 2: Sensitivity analysis of the results. Goodness-of-fit values for 10 runs with different pseudo random sequences for L=100, 1000, 10000 and 100000.

Conclusions

This paper presented a novel generative model of power law distributions. The driving force behind the process resulting in the scale-free distribution is the maximizing behavior of agents, acting on limited local information that is distributed unevenly among them. We believe that this entirely local approach has a great potential to explain the occurrence of power law distributions in social systems with bounded rational agents, where the assumption global information access is unfeasible.

Acknowledgements

The author gratefully acknowledges the partial support of the Informatics Cooperative Research and Education Center, Loránd Eötvös University, Budapest and the GVOP-3.2.2-2004-07-0005/3.0 grant of the Hungarian National Office for Research and Technology. He would also like to thank Prof. George Kampis at Loránd Eötvös University, Budapest and Prof. Peter Erdi at the Center for Complex Systems Studies, Physics Department, Kalamazoo College, MI for their help.

References

[1] Albert, R. and Barabási, A.-L.: "Statistical mechanics of complex networks", *Reviews of Modern Physics*, Vol. 74:47, 2002.

[2] Barabási, A.-L. and Albert, R.: "Emergence of scaling in random networks", *Science*, Vol. 286, pp. 509-512, 1999.

[3] Cederman, L.-E.: "Modeling the Size of Wars: From Billiard Balls to Sandpiles", *American Political Science Review*, Vol. 97, pp. 135-150, 2003.

[4] Champernowne, D.: "A model of income distribution", *Economic Journal*, Vol. 63, pp. 318-351, 1953.

[5] Faloutsos, M., Faloutsos, P. and Faloutsos, C.: "On power-law relationships of the Internet topology", In *Proceedings of ACM SIGCOMM 1999 Conference*, pp. 251-261. 1999.

[6] Gabaix, X.: "Zipf's law for cities: an explanation", *Quarterly Journal of Economics*, Vol. 114, pp. 739-767, 1999

[7] Gulyás, L.: "Generation of Robust Networks with Optimization under Budget Constraints", In *Proceedings of The 5th International Workshop on Emergent Synthesis (IWES'04)*, Budapest, Hungary, 2004.

[8] Mega, M. S., Allegrini, P., Grigolini, P., Latora, V., Palatella, L., Rapisarda, A. and Vinciguerra, S.: "Power-Law Time Distribution of Large Earthquakes", *Physical Review Letters*, Vol. 90:18, pp. 188501, 2003.

[9] Mitzenmacher, M.: "A brief history of generative models for power law and lognormal distributions", *Internet Mathematics*, Vol. 1:2, pp. 226-251, 2003.

[10] Neukum, G. and Ivanov, B. A.: "Crater size distributions and impact probabilities on Earth from lunar, terrestrial-planet, and asteroid cratering data", *Hazards due to comets and asteroids* (Ed. Gehrels, T.), University of Arizona Press, pp. 359-390. 1994.

[11] Pareto, V.: *Cours d'Economie Politique*, Droz, Geneva, Switzerland. 1896.

[12] Simon, H. A.: "On A Class of Skew Distribution Functions", *Biometrika*, Vol. 52, pp. 425-440, 1955.

[13] Simon, H. A. and Bonini, C. P.: "The Size Distribution of Business Firms", *American Economic Review*, Vol. 48, pp. 607-617, 1958.

[14] Willis, J. C.: *Age and Area: A Study in Geographic Distribution and Origin of Species*, Cambridge University Press, Cambridge, UK. 1922.

Universal scaling of inter-node distances in complex networks

Janusz A. Hołyst*, Julian Sienkiewicz, Agata Fronczak, Piotr Fronczak, Krzysztof Suchecki and Piotr Wójcicki

August 31, 2005

Abstract

We have studied dependence of distances between nodes in various networks and degrees of such vertices. We have observed that the mean distance between two nodes of degrees k_i and k_j equals to $\langle l_{ij}\rangle = A - B\log(k_i k_j)$. The relation holds for the following systems: Erdős-Rényi random graphs, scale-free Barabási-Albert models, science collaboration networks, biological networks, Internet Autonomous Systems and public transport networks. A simple heuristic theory for this scaling law is presented. Corrections due to the network clustering coefficient and node degree-degree correlations are taken into account.

Keywords: complex networks, distances, scaling, correlations

In [1, 2] an analytical model for average path lengths in random uncorrelated networks was considered and it was shown that the shortest path length between nodes i and j possessing degrees k_i and k_j can be described as

$$l_{ij}(k_i,k_j) = \frac{-\ln k_i k_j + \ln\left(\langle k^2\rangle - \langle k\rangle\right) + \ln N - \gamma}{\ln\left(\langle k^2\rangle/\langle k\rangle - 1\right)} + \frac{1}{2}, \tag{1}$$

where $\gamma = 0.5772$ is the Euler constant, whereas $\langle k\rangle$ and $\langle k^2\rangle$ correspond to the first and the second moments of node degree distribution $P(k)$. It follows that the mean distance between two nodes is lineary dependent on the logarithm of their degree product

$$\langle l_{ij}\rangle = A - B\log(k_i k_j). \tag{2}$$

Below we show that the relation (2) can also be obtained from a simple model of branching trees exploring the space of a random network [3, 4] (see Fig. 1). Let us consider a path from a randomly chosen node i to a randomly chosen node j in such a network. Following a random direction of a randomly chosen edge one approaches a node j with the probability $p_j = k_j/(2E)$, where $2E = N\langle k\rangle$ is a double number of links. It means that in average one needs $M_j = 1/p_j = 2E/k_j$ of random trials to arrive at the node j. Now let us consider a branching process represented by the tree T_i (see Fig. 1) that starts at the node i where an average branching factor is κ (all loops are neglected). If the distance between the node i and the surface of the tree equals to x then in average there are $N_i = k_i\kappa^{x-1}$ nodes at such a surface and there is the same number of links ending at these nodes. It follows that in average the tree T_i touches the node j when $N_i = M_j$ i.e. when

$$k_i k_j \kappa^{x-1} = N\langle k\rangle. \tag{3}$$

*Corresponding author: Faculty of Physics, Warsaw University of Technology, ul. Koszykowa 75 , 00-662 Warszawa, *Tel.:* +48 22 6607133; *fax:* +48 22 6282171; *email address:* jholyst@if.pw.edu.pl

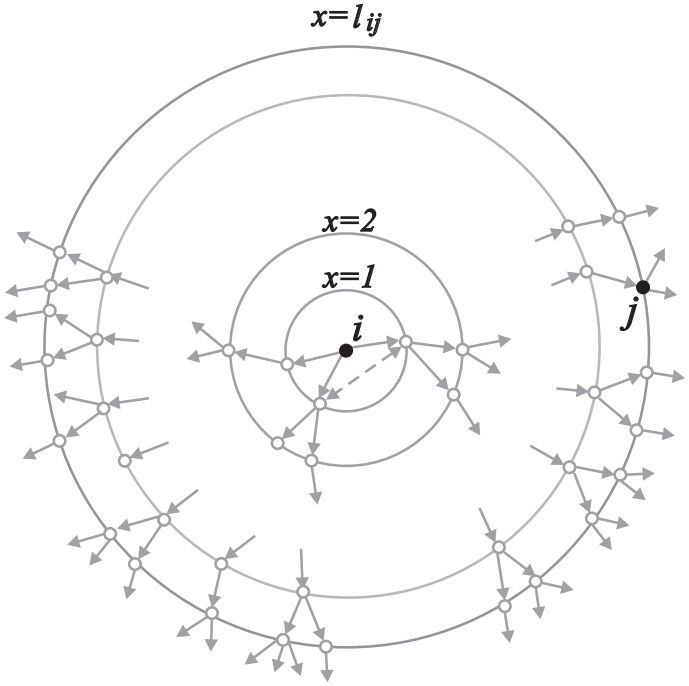


Figure 1: Tree formed by a random process, starting from the node i and approaching the node j.

Since the mean distance from the node i to the node j is $\langle l_{ij} \rangle = x$ thus we get the scaling relation (2) with

$$A = 1 + \frac{\log(N\langle k \rangle)}{\log \kappa} \quad \text{and} \quad B = \frac{1}{\log \kappa}. \tag{4}$$

The result (4) is in agreement with the paper [5] where the concept of generating functions for random graphs with arbitrary degree distributions has been used.

One has to take into account that in the above considerations we have assumed that there are no degree-degree correlations, we have neglected all loops and we have treated the branching level x as a continuum variable to fulfill the relation (3). Assuming that the branching factor κ can be expressed as $\langle k^2 \rangle / \langle k \rangle - 1$ [6], one can see that the differences between the results (1) and (4)are small, at least for the case when $N \to \infty$ and κ is finite.

Fig. 2 present mean distances $\langle l_{ij} \rangle$ between pairs of nodes i and j as a function of a product of their degrees $k_i k_j$ for the following systems: Erdős-Rényi random graphs, Barabási-Albert evolving networks, biological networks [7, 8, 9], social networks [10, 11], Internet Autonomous Systems [12] and selected networks for public transport in Polish cities [14, 15]. The relation (2) is very well observed over several decades of data points, although among the systems mentioned there are both scale-free and single scale networks, with either negligible or very high clustering coefficient, assortative [16], disassortative or uncorrelated.

Although the scaling (2) works well for distances averaged over all pairs of nodes specified by a given product $k_i k_j$, there can be large differences if one changes k_i while keeping $k_i k_j$ constant. The Fig. 3 presents the dependence of average path length $\langle l_{ij} \rangle$ on k_i, for a fixed product $k_i k_j$ in the case of several networks from different classes. One can see that although the *Astro* network is assortative (short-range

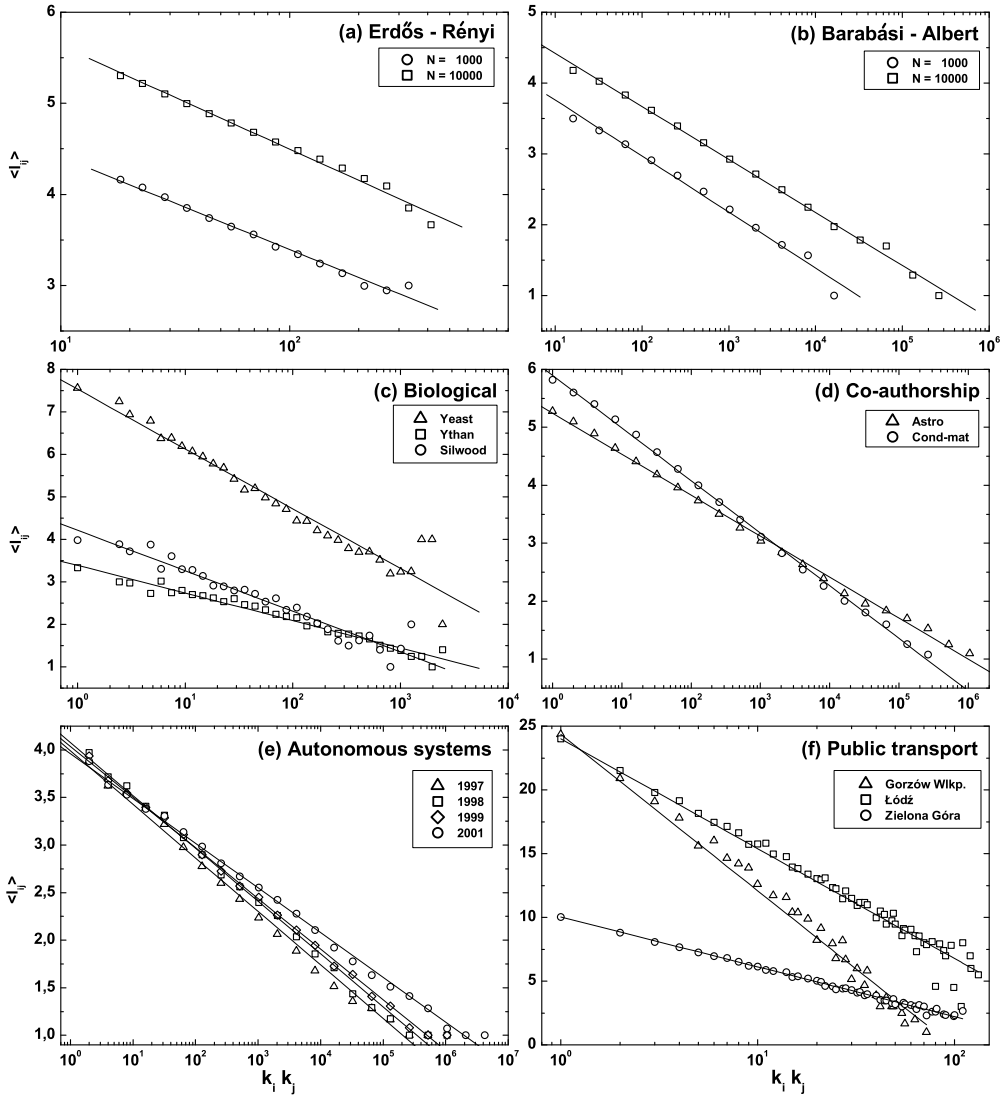


Figure 2: Mean distance $\langle l_{ij} \rangle$ between pairs of nodes i and j as a function of a product of their degrees $k_i k_j$. **(a)** *Erdős-Rényi* random graphs: $\langle k \rangle = 8$ and $N = 1000$ (circles) $N = 10000$ (squares), **(b)***Barabási-Albert* networks: $\langle k \rangle = 8$ and $N = 1000$ (circles) $N = 10000$ (squares), **(c)** Biological networks: *Silwood* (circles), *Yeast* (triangles), *Ythan* (squares), **(d)** Co-authorship networks: *Astro* (triangles), *circles* (circles), **(e)** Internet Autonomous Systems: *Year 1997* (triangles), *Year 1998* (squares) *Year 1999* (diamonds), *Year 2001* (circles), **(f)** Public transport networks in Polish cities: *Gorzów Wlkp.* (triangles), *Łódź* (squares), *Zielona Góra* (circles) In **(a)**, **(b)**, **(d)** and **(e)** data are logarithmically binned with the power of 2, in case of **(c)** with the power of 1.25 and in case of **(f)** data are not binned.

attraction), pairs of nodes with similar degrees are in average further away than different degree pairs (long-range repulsion). For the disassortative network AS [16] the behavior is opposite. For uncorrelated networks (Erdős-Rényi, Barabási-Albert), the average path length is constant given the product $k_i k_j$ fixed [4].

In [3] we compare (4) to results from real networks and numerical simulations. In fact our approximate approach (4) fits very well to random Erdős–Rényi graphs and BA models but the corresponding coefficients A and B for real networks are different from results of our simple theory. The formulas (4) can be improved by taking into account effects of loops and node degree-degree correlations.

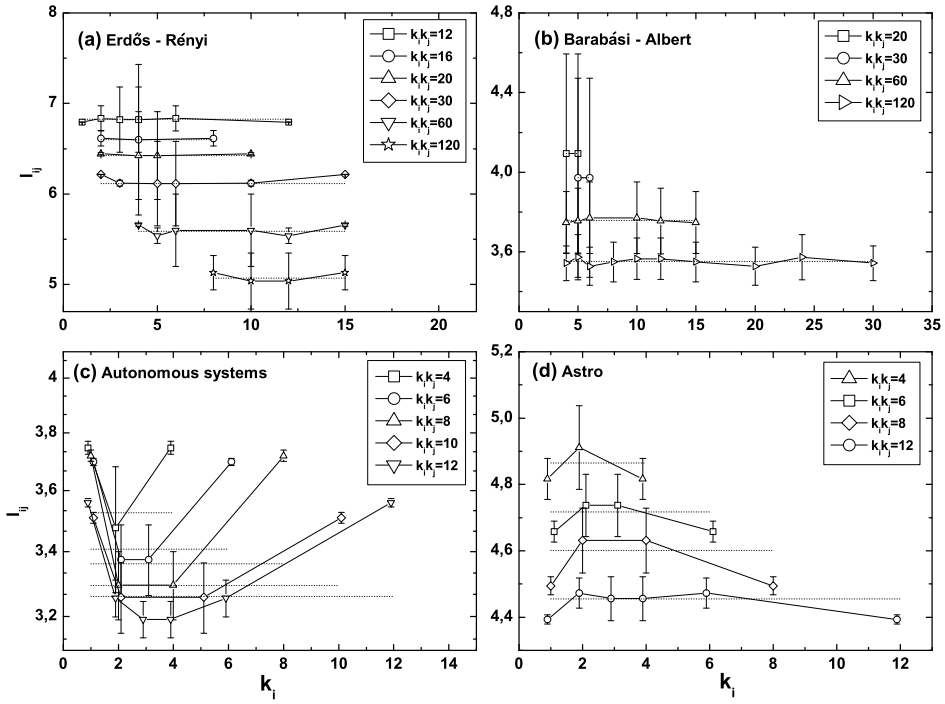


Figure 3: Dependence of average path length on k_i, for fixed $k_i k_j$ product. The lines connecting the symbols are there for clarity. The bars show point weight, meaning relative numbers of pairs ij. The horizontal lines are weighted averages over k_i and represent average path lengths for the given product $k_i k_j$. Note: the very small shifts on k_i axis between data for different $k_i k_j$ are artificially introduced to make the weight bars not overlap.

The influence of loops of the length three can be estimated as follows [3]. Let us assume that in the branching process forming the tree T_i two nodes from the nearest neighborhood of the node i are *directly* linked (the dashed line at Fig.1). Such a situation can occur at any point of the branching tree T_i and corresponding links are useless for further network exploration by the tree T_i. It follows the *effective* contribution from both connected nodes to the mean branching factor of the tree T_i is decreased. Assuming that clustering coefficients of every node are the same, the corrected factor for the branching process equals to $\kappa_c = \kappa - c\kappa$ where c is the network clustering coefficient. This equation is not valid for the branching process around the node i where $\kappa_i' = \kappa - c(k_i - 1)$. A similar situation arises around the node j. Replacing k_i and k_j with $\langle k \rangle$ in κ_i' and κ_j' one gets

$$k_i k_j [\kappa(1-c')]^2 [\kappa(1-c)]^{x-3} = N\langle k\rangle, \tag{5}$$

where $c' = c(\langle k\rangle - 1)/\kappa$. It follows that instead of (4) we have

$$A' = 3 + \frac{\log(N\langle k\rangle) - 2\log[\kappa(1-c')]}{\log[\kappa(1-c)]}, \quad B' = \frac{1}{\log[\kappa(1-c)]}. \tag{6}$$

Now, let us consider the presence of degree correlations [4]. Such correlations mean that average degrees $k_i^{(nn)}$ of nodes in the neighborhood of a node i depend on the degree k_i. Let us assume that this relation can be written as

$$\kappa_i \equiv k_i^{(nn)} - 1 = D k_i^{\phi-1} \tag{7}$$

If ϕ is larger than one then the network is assortative, i.e. high degree nodes are mostly connected to other high degree nodes and similarly low degree nodes are connected to other low degree nodes. Such a situation occurs for example in networks describing scientific collaboration [16]. If ϕ is smaller than one, then the network is disassortative and high degree nodes are mostly connected with low degree nodes what is typical for the Internet Autonomous Systems [16]. If we neglect higher order correlations then Eq.3 should be replaced by

$$k_i k_j \kappa_i \kappa_j \kappa^{x-3} = N\langle k\rangle \tag{8}$$

Taking into account Eq. 7 we can replace parameters A and B given by the Eq. 4 with

$$A_\phi = A + 2 - 2B\log D \quad \text{and} \quad B_\phi = \phi B \tag{9}$$

In conclusions we have observed a universal path length scaling for different classes of real networks and models. The mean distance between nodes of degrees k_i and k_j is a linear function of $\log(k_i k_j)$. The scaling holds over many decades and does not depend on network degree distributions, clustering coefficients or degree-degree correlations. We have showed that a simple model of random tree exploring the network explains such a scaling behavior. In an extended version of the model clustering effects and first order degree-degree correlations have been introduced to improve theoretical predictions for real world systems. In our opinion a better agreement between theoretical results and experimental data may be obtained taking into account higher order correlations.

Authors acknowledge the State Committee for Scientific Research in Poland for financial support under grant No. 134/E-365/6.PRUE/DIE239/2005-2007. This work has also been partially supported by European Commission Projects: CREEN FP6-2003-NEST-Path-012864 and MMCOMNET FP6-2003-NEST-Path-012999.

References

[1] Fronczak A, Fronczak P, Hołyst JA: How to calculate the main characteristics of random uncorrelated networks. E-print 2005: cond-mat/0502663.

[2] Fronczak A, Fronczak P, Hołyst JA: Average path length in random networks. Phys. Rev. E 2004; 70, 056110.

[3] Hołyst JA, Sienkiewicz J, Fronczak A, Fronczak P, Suchecki K: Universal scaling of distances in complex networks. Phys. Rev. E 2005; 72, 026108.

[4] Hołyst JA, Sienkiewicz J, Fronczak A, Fronczak P, Suchecki K: Scaling of distances in correlated complex networks. Physica A 2005; 351, 167.

[5] Motter AE, Nishikawa T, Lai YC: Range-based attack on links in scale-free networks: Are long-range links responsible for the small-wrold phenomenon? Phys. Rev. E 2002; 66, 065103(R).

[6] Newman MEJ, Strogatz SH, Watts DJ: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E 2001; 64, 026118.

[7] Data for Foodwebs (*Silwood* and *Ythan*) [8] have been taken from the website http://www.cosin.org /extra/data/foodwebs/ whereas data for protein interaction network (*Yeast*) [9] has been taken from http://www.nd.edu/ networks/database/index.html.

[8] Garlaschelli D, Caldarelli G, Pietronero L: Universal scaling relations in food webs. Nature 2003; 423, 165.

[9] Jeong H, Mason SP, Barabási A-L, Oltavi ZN: Lethality and centrality in protein networks. Nature 2001; 411, 41.

[10] Scientific collaboration network [11] data have been collected from two publicly available databases of papers http://arxiv.org/archive/astro-ph (*Astro*) and arxiv.org/archive/cond-mat (*Cond-mat*) for the period 1995-2001.

[11] Newman MEJ: Scientific collaboration networks. I. Network construction and fundamental results. Phys. Rev. E 2001; 64, 016131.

[12] Data for the Internet [13] has been taken from www.cosin.org/extra/data/internet/.

[13] Pastor-Satorras R, Vespignani A: Evolution and Structure of the Internet: A Statistical Physics Approach. Cambridge Univ. Press 2004.

[14] In those networks vertices are bus- and tramstops while an edge exists if at least one public transport line crosses two stops. [15].

[15] Sienkiewicz J, Hołyst JA: Statistical anlysis of 22 public transport networks in Poland. E-print 2005: physics/0506074.

[16] Newman MEJ: Assortative Mixing in Networks. Phys. Rev. Lett. 2002; 89, 208701.

[17] Watts DJ, Strogatz SH: Collective dynamics of 'small-world' networks. Nature 1998; 393, 440.

On Small-World generating Models *

Michael Kaufmann, Katharina A. Lehmann and Hendrik Post

Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 14, 72076 Tübingen
email: mk / lehmannk / post@informatik.uni-tuebingen.de

[...] the best material model for a cat is another, or preferably the same cat. - N. Wiener and A. Rosenblueth [24]

Abstract. What exactly is a *small-world*? Watts and Strogatz [23] define every network with a high clustering coefficient and a low diameter to be a small-world. We will show here that for this classic definition there are some counter-intuitive examples where either false-negative or false-positive classifications of network models occur.
To bring forth a new definition for small-world generating network models, we will first introduce a slightly varied small-world network model. This model is based on a regular grid graph and an added $G(n,p)$[13] random graph. We will then give an upper bound for the diameter of the generated networks dependent on p and n. This upper bound is generalized to combinations of a so-called 'locally clustered' graph family with a $G(n,p)$ graph. On the basis of this general method we propose a new definition for small-world generating models.

1 Introduction

The 'small-world effect' has long been a part of folklore. It describes the fact that most of us are tightly knit into small social clusters while on the other hand we need just a short chain of acquaintances to connect us to any other human on the world. Milgram estimated the number of persons in such a chain to be around six [19] which is why this observation is also known under the title 'six degrees of separation'. The first formal approach to explain this astonishing result was made by Watts and Strogatz in a seminal paper [23] in which they gave a rough definition of small-world networks and presented a model for their generation. They defined a small-world to be every network with a high clustering coefficient and a low diameter.

Following their publication, several real-world networks such as the WWW or file-sharing communities were analyzed and shown to be small-worlds (e.g. [1, 2, 11, 14]). A second research area deals with network-based processes on small-world networks, like the behavior of neural networks on small-worlds [17] or disease spreading in small-worlds [21]. Other directions of research tried to find

* Partially supported by DFG-Grant Ka812/11-1

more rigorous analytical results on the properties of either the classic small-world model or on variants of the small-world model that were easier to analyze or captured new aspects of small-worlds [20, 12, 16, 15, 3, 10, 4, 7].

In the classical model of Watts and Strogatz, n vertices are placed equidistantly on a ring and every vertex is connected with its k next neighbors. Every edge has a probability of p' to be rewired, i.e., one of the endpoints is fixed and the edge is rewired to a new, randomly drawn target vertex [23]. For $p' = 1$ a certain kind of random graph emerges. A variant proposed by Watts together with Newman [20] is based on the same basic ring graph but instead of rewiring, a special kind of random graph is added to this basic graph. It is not quite clear how this random graph is generated but the number of its edges is restricted to pkn which makes it impossible that a proper $G(n, p)$ random graph instance is added for $k < n - 1$. Watts and Newman state that there is a single threshold value for p' such that the scaling of the diameter changes from a linear behaviour to a logarithmic scaling.

We see three main objections regarding the classic small-world model [23] and the variant in [20]:

Random Graph Component In the classic small-world model, the rewired random edges are not building an instance of the commonly known and analyzed random graph families $G(n, m)$ or $G(n, p)$, not even for $p = 1$ [8, 6]. In a $G(n, m)$ instance, a fixed number of m random edges is drawn between n vertices, in a $G(n, p)$ instance every possible edge between the vertices exists with probability p. The diameter of these graph classes is $O(\log n)$. In the classic small-world model [23], not all possible instances of the $G(n, m)$ or $G(n, p)$ sets can be realized by the described process. This is mainly due to the fixation of one of the endpoints in the rewiring process. Nonetheless, it is claimed in [23] that the emerging random graph for $p = 1$ has a diameter that scales with $O(\log n)$. In the variant of [20], the type of the added random graph is unclear. Thus, the scaling of the diameter of this random graph component is possibly not given by $O(\log n)$ but the calculated threshold value depends heavily on this assumption.

Clustering Coefficient and Diameter as Indicators for Small-Worldness The model Watts and Strogatz proposed for generating small-world networks is based on the rewiring process in which edges are generated randomly. The main problem with the classic definition of small-world networks is that it cannot determine whether some random process was involved in the generation of a given network. We will illustrate this point:

It is very easy to construct a network with a high clustering coefficient and a small diameter by first building a balanced tree, e.g., a quaternary tree, and then adding edges between those leaves of the tree that have only distance 2 to each other (Fig. 1). The balanced tree will then provide a diameter of $O(\log n)$ for the whole network and more than half of the vertices are leaves with a clustering coefficient of 1, which generates an average clustering coefficient of > 0.5 for the whole graph. This network model is based on a hierarchical tree that

connects locally dense clusters with each other. For example, the power grid of the U.S. may rather be based on a hierarchical backbone that spans the network rather than that it is the result of a process where random edges are added to a locally clustered network [23]. This example shows that the clustering coefficient identifies networks as small-worlds that are generated by another network model, and thus it identifies **false-positives**.

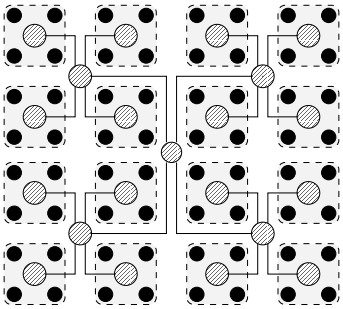


Fig. 1. Hybrid graphs of small cliques and a balanced tree as a counterexample for the classic definition of small-worlds: Each block of four vertices constitutes a clique in which every vertex has a clustering coefficient of 1. These blocks are combined to a connected graph by a quaternary tree where every vertex has exactly four children and most of the vertices in the tree have a clustering coefficient of 0. The clique component consists of 64 vertices, the added tree component consists of 21 vertices and thus the combined graph shows an average clustering coefficient of at least 0.75. The diameter of this graph is determined by the diameter of the tree component which scales with $O(\log n)$.

Also, networks with a small clustering coefficient can be considered as small-world networks. We want to illustrate this point with some citations: Regular grid graphs with a degree of $2d$ show a clustering coefficient of 0. Nonetheless, also those graphs constitute a small-world if they are combined with random graphs: Kleinberg bases his searchable small-worlds on regular grid graphs with a clustering coefficient of 0 [15, 16]. Networks representing mostly hetero-sexual relationships contain very little triangles. Rather, the graph forms nearly a bipartite graph between males and females [5]. Nonetheless these networks are normally classified as small-worlds because of their combination of mainly local relationships with some additional long range contacts [18, 22]. Chung et al. generalize the idea behind the clustering coefficient to include more models into the general framework of small-world generating models [10, 4]: Their small-world model combines a so-called (k, l)-*local* graph with a power-law random graph. A graph is a (k, l)-local graph if for each edge $e = (v, w)$, v and w are connected by at least k edge-disjoint paths of at most length l. Summarizing, the intu-

ition is that not all small-worlds can be detected with the help of the clustering coefficient: it gives **false-negatives**.

Missing Generalizability The small-world model of Watts and Strogatz is attractive because of its simplicity. On the other hand it lacks some extendability to create more practical small-world generating models [10, 4].

Our goal is to build a generalized framework for small-world generating models that removes the objections given above. To deal with the first point we will simplify the small-world model of Newman and Watts by adding a $G(n, p)$[13] instance to a regular, d-dimensional grid. We will give an upper bound on the diameter of this basic hybrid graph. The notion of the *clustering coefficient* will be semantically replaced by the notion of *locally clustered* graph families. Then, the upper bound on the diameter of our basic hybrid graphs can be generalized to any combination of a *locally clustered* graph family with certain random graph families. Thus, we offer possible building blocks for constructing small-world generating models that are based on two components: a local and a random component. This allows for more complex small-world generating models that may be the basis for the simulation of real-world complex systems. We will further use the building-block framework to give a new definition of small-world generating models that incorporates all classic small-world models.

The paper is organized as follows: In Sec. 2 we give some basic definitions needed in the concourse of the article. Sec. 3 is structured into three subsections: Subsec. 3.1 introduces the main model, Subsec. 3.2 gives the upper bound for this model and in Subsec. 3.3 we generalize the analysis to any combination of locally clustered graph families with $G(n, p)$. In Sec. 4 the new definition for small-world generating models is introduced and discussed. Sec. 5 concludes with a summary and discussion of the results.

2 Definitions

A graph family $G(n)$ in this article denotes any set of graphs generated by the same algorithm and parameterized by the number of vertices in it. For non-random graph families and a fixed set of parameters only one specific graph is generated. For graph families generated partly by probabilistic processes, $G(n)$ is defined as the set of all possible realizations. Statements about $G(n)$ are then interpreted as statements about expected characteristics of this set. We will use the notation $G(n)$ interchangeably for the set or a specific realization of this set.

A regular d-dimensional, equilateral grid (hypercubical lattice) $G_d(n)$ is defined as a set of vertices, placed on integer positions in d dimensions. $a \in \mathbb{N}$ denotes the number of vertices placed in each of the d dimensions. The number of vertices in this grid is then given by $n = a^d$, where every possible position - identified by a d-dimensional vector $(1 \leq b_1 \leq a, 1 \leq b_2 \leq a, \ldots, 1 \leq b_d \leq a)$ - is occupied with one vertex. The degree $deg(v)$ of a vertex is defined as the number of incident edges and equals the number of direct neighbors of v. Every

vertex v is connected by an edge to those vertices that differ in their position by exactly one in exactly one dimension from the position of v, i.e., every vertex has at most degree $2d$. For these grids, the graph theoretic distance $d(v, w)$ of any two vertices v, w, i.e., the minimal number of traversed edges to walk from v to w, coincides with the Manhattan distance $d_M(v, w)$ of these vertices which is defined by:

$$d_M(v, w) = \sum_{1 \leq i \leq d} |b_i(v) - b_i(w)| \tag{1}$$

The diameter $D(G)$ of any graph G is defined as the maximal distance of any two vertices within the graph. The diameter $D(G_d(n))$ is given by the maximal Manhattan distance of any two vertices in $G_d(n)$ and can be calculated by:

$$D(G_d(n)) = \sum_{1 \leq i \leq d} a - 1 = d(a - 1) \tag{2}$$

A graph is *connected* if there is a way from every vertex v to any other vertex w.

The clustering coefficient $C(v)$ of a vertex v is defined as the number of edges $e(v)$ between direct neighbors of v and the maximal possible number of edges between direct neighbors [23]:

$$C(v) = \frac{e(v)}{deg(v)(deg(v) - 1)} \tag{3}$$

The clustering coefficient $C(G)$ of a graph G is defined as the average clustering coefficient of G's vertices.

A $G(n, p)$ random graph is defined as an instance of all possible graphs with n vertices where every of the $\binom{n}{2}$ edges exists with probability p [13]. A $G(n, m)$ random graph is defined as an instance of all possible graphs with n vertices and exactly m edges, drawn uniformly at random from all possible edges.

We will use the following theorem on the diameter of random graphs $G(n, p)$ [8]:

Theorem 1. *If* $pn/\log n \to \infty$ *and* $\log n/\log(np) \to \infty$ *then* $D(G(n, p))$ *is asymptotically equal to* $\log n/\log(np)$ *with high probability.*

Note that this theorem implicitly includes that the random graph is connected with high probability. To simplify the following proofs we will use a stricter version of the theorem and require additionally that $p \geq (\log n)^{1+\epsilon}/n$.

3 A Framework for Small-World generated Models

3.1 A first Starting Point

As argued in the introduction, the Watts-Strogatz- and the Newman-Watts-model suffer from some problems. We replace their models by a simplified version composed of a random graph $G(n, p)$ and a regular d-dimensional grid in the

following way: The basic regular graph is the d-dimensional grid of n vertices, where each vertex is connected to its $2d$ next neighbors, combined with a $G(n,p)$ random graph on the same n vertices. We will denote by $G_d(n,p)$ a graph from our model, which is given by the combination of a $G_d(n)$ regular grid and a random graph $G(n,p)$.

The remaining part of this section gives answers to the following question: How does the diameter of regular networks combined with a small set of random edges scale?

Since the basic network is a d-dimensional grid, the diameter of it without any added random edges will scale with $a-1$ for a fixed dimension d: $D(G_d(n)) = d \cdot (a-1)$. If the added random graph has a probability of $(\log p)^{1+\epsilon}/n$ then the combined graph will have a diameter that is dominated by the diameter of the random graph and thus is asymptotical to at most $\log n/\log(np)$ (Theorem 1).

What happens in the regime where p lies below $(\log p)^{1+\epsilon}/n$? When will the diameter of the combined graphs scale at most (poly-) logarithmically?

In Theorem 5 we will give a detailed upper bound for the diameter of the combined graph for a given number of random edges.

3.2 The Diameter of $G_d(n,p)$-Graphs

For the above given model of a graph $G_d(n,p)$ the following lemma holds:

Lemma 2. *For $p = \frac{1}{cn}$, $c \in \mathbb{R}^+$ the diameter of $G_d(n,p)$ is asymptotically bounded by at most*

$$d \cdot \left(\left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil - 1 \right) \cdot \left(\frac{\log n}{(1+\epsilon)\log\log n - \log 2} + 1 \right) \tag{4}$$

Proof. The proof proceeds in four steps:

1. To prove the lemma we partition $G_d(n,p)$ into n_S connected d-dimensional equilateral subgraphs $S_i, 1 \leq i \leq n_S$ with a side length l such that each subgraph contains at least $s = l^d \geq c(\log n)^{1+\epsilon}$ vertices (Figure 2).
2. For any a, we can only build $\lfloor a/l \rfloor$ full subgraphs per dimension. n^* denotes the number of all vertices contained in a full subgraph. We will show that the $n - n^*$vertices that are not contained in any full subgraph build a vanishing fraction of all vertices for $n \to \infty$. We will thus base our proof on a reduced regular d-dimensional grid of size n^* that contains only the full subgraphs.
3. We construct a supergraph $G_S(n_S) = (S, E')$ where each vertex $v_i \in S$ uniquely represents the subgraph S_i for $1 \leq i \leq n_S$. $e = (v_i, v_j) \in E'$ iff there is at least one random edge from any vertex in S_i to any vertex in S_j. We will prove that Theorem 1 is applicable on $G_S(n_S)$.
4. Then we will expand $G_S(n_S)$ to gain a bound on the diameter of the original but reduced graph $G_d(n^*,p)$. The diameter of $G_d(n^*,p)$ is bounded by the product of the diameter of the subgraphs $D(S_i)$ and the diameter $D(G_S)$. We will show that there are numerous partitions of $G_d(n,p)$ into n_S subgraphs. Especially, for any pair of vertices v, w there is at least one partition of

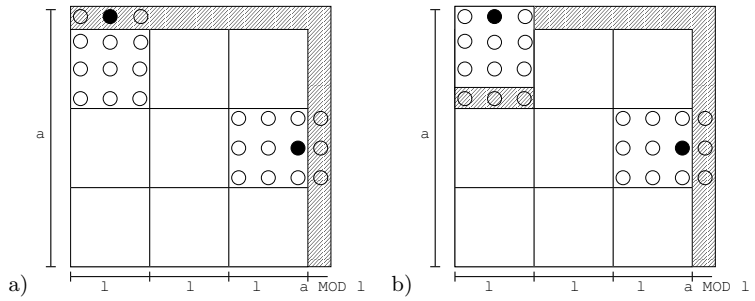


Fig. 2. Valid partitions for a 2-dimensional grid with side length a. Full equilateral subgraphs with side length l may be placed arbitrarily as long as their number is maximal. Therefore numerous partitions exist and for each pair of vertices numerous partitions can be found where both are contained in full subgraphs.

$G_d(n)$ such that both, v and w, are contained in full subgraphs. Since every supergraph based on a possible partition obeys Theorem 1, we will therefore have shown that the whole graph $G_d(n,p)$ obeys Lemma 2 and the case is proven.

We will start by partitioning a $G_d(n,p)$ graph. Let $S_i, 1 \leq i \leq n_S$ denote an equilateral subgraph that has a side length of $l = \left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil$ in each dimension. The number s of vertices contained in one (full) subgraph is bounded by:

$$c \cdot (\log n)^{1+\epsilon} \leq s = \left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil^d < 2^d \cdot c \cdot (\log n)^{1+\epsilon} \tag{5}$$

We will now partition $G_d(n,p)$ into the subgraphs as shown in Figure 2. Obviously, incomplete subgraphs exist if a/l is not integer. The leftover vertices can be placed arbitrarily between full subgraphs as indicated in Figure 2 b). For simplicity we will consider instead of $G_d(n)$ a smaller hypercube $G_d(n^*)$ containing all full subgraphs. Note that now a^* with $\sqrt[d]{n}^* = a^* \leq a$ is the maximal integer smaller than a that is a multiple of l. Let $q = a^*/l$ denote the number of subgraphs in one dimension.

The relative fraction of vertices not contained in full subgraphs is approaching 0 for $n \to \infty$:

$$\frac{n - n^*}{n} \leq \frac{(l \cdot (q+1))^d - (l \cdot q)^d}{(l \cdot q)^d} \tag{6}$$

$$= \left(\frac{q+1}{q}\right)^d - 1 \tag{7}$$

Since $q \to \infty$ for $n \to \infty$, the relative fraction of ignored vertices is asymptotically 0. Note that $n_S = \frac{n^*}{s} \geq \frac{n}{2^d \cdot c \cdot (\log n)^{1+\epsilon}} \to \infty$. Thus for $n \to \infty$ we may

safely use

$$n \geq n^* > n/2 \tag{8}$$

In $G_d(n^*, p)$ there are s^2 possible random edges between any vertex from subgraph S_i and any vertex from subgraph S_j. Each of these edges exists independently with probability p. It follows that for G_S the probability p_S is exactly $\frac{s^2}{cn}$.

We will now prove that Theorem 1 can be applied to $G_S(n)$. A basic observation is that for $n \to \infty$, also $n_S \to \infty$. Additionally, we must show that $\frac{p_S n_S}{\log n_S} \to \infty$ and $\frac{\log n_S}{\log(n_S p_S)} \to \infty$ for $n_S \to \infty$.

Regarding, that for all $n_S > 1$, $n^* > n/2$ (eq. 8) the following two equations hold:

$$\frac{p_S \cdot n_S}{\log n_S} = \frac{s^2}{cn} \cdot \frac{n^*}{s} \cdot \frac{1}{\log \frac{n^*}{s}} \tag{9}$$

$$\geq \frac{s}{2c(\log n^* - \log s)} \tag{10}$$

$$\geq \frac{(\log n)^{1+\epsilon}}{2\log n - 2\log s} \tag{11}$$

such that $\frac{p_S \cdot n_S}{\log n_S} \to \infty$ for $n \to \infty$ and

$$\frac{\log n_S}{\log(p_S \cdot n_S)} = \frac{\log \frac{n^*}{s}}{\log \frac{s \cdot n^*}{cn}} \tag{12}$$

$$\geq \frac{\log n/2 - \log(2^d \cdot c(\log n)^{1+\epsilon})}{\log(2^d (\log n)^{1+\epsilon})} \tag{13}$$

such that also $\frac{\log n_S}{\log(p_S \cdot n_S)} \to \infty$. By theorem 1 we know that thus G_S has a diameter asymptotical to $\frac{\log n_S}{\log(p_S \cdot n_S)}$. Regarding that $n^*/n > 1/2$ this is bounded by

$$D(G_S) = \frac{\log n_S}{\log(p_S \cdot n_S)} \tag{14}$$

$$\leq \frac{\log n}{\log \frac{s}{2c}} \tag{15}$$

$$\leq \frac{\log n}{(1+\epsilon)\log\log n - \log 2} \tag{16}$$

$$\leq \frac{\log n}{\log\log n} \tag{17}$$

Where the last inequality is valid for all n with $\epsilon \log\log n > \log 2$.

We will now expand $G_S(n)$ in order to get an upper bound for the diameter of $G_d(n, p)$.

Let v and w be two vertices in the original graph $G_d(n, p)$. First, we will reduce $G_d(n, p)$ to $G_d(n^*, p)$ in such a way that v and w are contained in $G_d(n^*, p)$.

Then we know that there is a path from subgraph S_i containing v to subgraph S_j containing w with a length of no more than $D(G_S)$. This path is denoted by $(e_1, e_2, ..., e_k)$, the sequence of edges to traverse to walk from S_i to S_j.

To use this path in the original graph $G_d(n,p)$, we will first have to walk from vertex v to that vertex v' from S_i that is attached to e_1. This will at most take $D(S_i) = d \cdot (\left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil - 1)$ steps. For every entered subgraph S_x on the way to subgraph S_j, an additional distance of $D(S_x)$ has at most to be added to get from the random edge entering the subgraph to the edge leaving this subgraph. Thus, the distance of v, w in the original graph $G_d(n,p)$ is asympotically given by at most

$$D(S_i) \cdot (D(G_S)+1) \leq d \cdot \left(\left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil - 1 \right) \cdot \left(\frac{\log n}{(1+\epsilon)\log\log n - \log 2} + 1 \right) \tag{18}$$

With this, Lemma 2 is proven. □

In the following we want to discuss what happens if the degree of the underlying grid graph is enlarged.

As stated in Lemma 2, the diameter of a $G_d(n,p)$ graph is asymptotically at most $D(S_i) \cdot (D(G_S) + 1)$. Let $G_d(n,k,p)$ denote an extended regular grid, in which every vertex is connected to its k next neighbors, combined with an additional $G(n,p)$ graph. The diameter $D(S_i)$ depends on the degree of the vertices in the underlying grid graph. Thus, if we want to reduce the diameter of the $G_d(n,k,p)$ graph we just have to add some more edges to the grid. For example, $D(S_i)$ is reduced to 1 if for $p = \frac{1}{c \cdot n}$ we add edges from every vertes to its $c \cdot (\log n)^{1+\epsilon}$ next neighbors. The combined graph $G_d(n, (\log n)^{1+\epsilon}, p)$ has now at most the diameter of G_S.

3.3 Generalizing the Small-World Model

In this section we will generalize Lemma 2 in two ways:

1. The probability p of the added random graph $G(n,p)$ can be as small as $\frac{1}{f(n) \cdot n}$ as long as $\frac{1}{(\log n)^{1+(\epsilon/2)}} \leq f(n) \leq \frac{1}{n^{1-\delta}}$ for some constants $\delta, \epsilon > 0$ and $n \to \infty$.
2. The basic regular d-dimensional grid can be replaced by certain graph families. This was already indicated at the end of section 3.2.

These two extensions lead finally to our generalized theorem on the diameter of small-worlds generated by our model.

Generalizing the random graph component At first, we explain in which range p can be chosen, such that the proof-technique can still be applied. In section 3.2, we kept $p = 1/cn$. For smaller $p = \frac{1}{f(n) \cdot n}$, the size s of the subgraphs has to be chosen larger such that Theorem 1 can be applied. Let again n_S denote the size of the supergraph.

For simplicity we assume that $n_S = n/s \in \mathbb{N}$ and $p \cdot s \cdot n = (\log n)^{1+\epsilon} \in \mathbb{N}$. The general case follows the argumentation above.

The number of nodes in each subgraph will be chosen such that $s = \frac{(\log n)^{1+\epsilon}}{p \cdot n} = f(n) \cdot (\log n)^{1+\epsilon}$. Again, Lemma 2 requires the validity of

$$\frac{p_S n_S}{\log n_S} \to \infty \tag{19}$$

and

$$\frac{\log n_S}{\log(n_S p_S)} \to \infty \tag{20}$$

As before $p_S = s^2 \cdot p$. We first analyze the condition given in equation (19):

$$\begin{aligned} \frac{p_S n_S}{\log n_S} &= s^2 \cdot p \cdot \frac{n}{s} \cdot \frac{1}{\log n_S} && (21)\\ &> s \cdot p \cdot \frac{n}{\log n} && (22)\\ &= (\log n)^{\epsilon} && (23) \end{aligned}$$

which tends to infinity for increasing n. The second condition (20) simplifies to

$$\begin{aligned} \frac{\log n_S}{\log(n_S p_S)} &= \frac{\log\left(\frac{n}{s}\right)}{\log\left(\frac{n}{s} \cdot s^2 \cdot p\right)} && (24)\\ &= \frac{\log\left(\frac{\frac{n}{f(n)}}{(\log n)^{1+\epsilon}}\right)}{\log(\log n)^{1+\epsilon}} && (25)\\ &= \frac{\log\left(\frac{n}{f(n)}\right)}{\log(\log n)^{1+\epsilon}} - 1 && (26) \end{aligned}$$

which tends to infinity for $f(n) \leq \frac{1}{n^{1-\delta}}, \delta > 0$. Therefore both conditions are met and Theorem 1 can be applied to G_S. If $f(n)$ is too low than there will be many random edges per vertex such that the subgraph size $s < 1$. To avoid this, we restrict $f(n) \geq \left(\frac{1}{(\log n)^{1+(\epsilon/2)}}\right)$ in order to guarantee a reasonable size for the subgraphs.

We summarize this result in

Lemma 3. *For any function* $\frac{1}{(\log n)^{1+(\epsilon/2)}} \leq f(n) \leq \frac{1}{n^{1-\delta}}, \epsilon, \delta > 0$ *and* $p = \frac{1}{f(n) \cdot n}$*, we can partition the grid graph within a* $G_d(n,p)$ *graph into* $n_S = \frac{n}{s}$ *subgraphs* S_i *of size* $s = f(n) \cdot (\log n)^{1+\epsilon}$ *such that* $G_d(n,p)$ *shows a diameter of asymptotically at most (Eq. 26)*

$$\underbrace{d \cdot \left(\left\lceil \sqrt[d]{c \cdot (\log n)^{1+\epsilon}} \right\rceil - 1\right)}_{D(S_i)} \cdot \underbrace{\left(\frac{\log(n/f(n))}{\log(\log n)^{1+\epsilon}}\right)}_{D(G_S)} \tag{27}$$

Possible replacements of the regular grid graph In the general proof we have used the following two properties of regular grid graphs: First, regular grid graphs are partitionable for every n into $\Theta(n/s(n))$ subgraphs of size $s(n)$ for any function $s(n) \leq n$ such that each of these subgraphs is a connected graph. The second property used is that for any pair of vertices v, w there must be at least one partition such that v and w are contained in any of the subgraphs.

To abstract from this special graph family to all graph families with these to properties we introduce the following definition:

Definition 4. Let $G_L(n)$ be a graph family with the following two properties:

1. $G_L(n)$ is partitionable for every n into $\Theta(n/s(n))$ subgraphs of size $s(n)$ for any function $s(n) \leq n$ such that each of these subgraphs is a connected graph
2. For any pair of vertices v, w and every n there must be at least one partition as described such that v and w are contained in proper subgraphs

$G_L(n)$ is called a *locally clustered* graph family.

Furthermore, a graph family can be *restricted locally clustered* with respect to some function $s(n) \leq n$ if for every n and every pair v, w $G_L(n)$ is partitionable into $\Theta(n/s(n))$ connected subgraphs of size $s(n)$ such that v and w are contained in proper subgraphs.

The notion of (restricted) local clusters in a graph thus can simply be interpreted as that every vertex in $G_L(n)$ can directly or indirectly reach at least $s(n)$ other vertices.

Classical small-world models are either based on the 1-dimensional ring lattice [23, 12, 7] or on d-dimensional regular grids [20, 16] and thus are based on *locally clustered* graph families. k-next neighborhood graphs in which n vertices are distributed uniformly in a unit-square and where every vertex is connected to its k geometrically next neighbors are also a *locally clustered* graph family. The proof for this statement is kind of lengthy, so the interested reader will find it in the appendix in Sec. 6.

Note that every graph family $G_L(n)$ is *restricted locally clustered* for at least $s(n) = 1$. Let $s_{max}(n)$ be that function $s'(n)$ that has the fastest growth of all functions $s(n)$ for which $G_L(n)$ is restricted locally clustered. If now $s_{max}(n) = k,\ k \in \mathbb{N}$ for $G_L(n)$ and $G_L(n)$ replaces the regular grid then it is clear that the size of the subgraphs is also at most k to obey $\Theta(n/s)$. This implies that p of the added random graph must be at least $O\left(\frac{(\log n)^{1+\epsilon}}{n}\right)$ in order to achieve a supergraph that obeys Theorem 1. It follows that the diameter is reduced to the diameter of a random graph because we added a random graph with the wanted diameter. This is certainly not a very interesting combination of graph classes. We will discuss this point further in Sec. 4.

We conclude this section with a theorem on the diameter of generalized small-world models combining a locally clustered graph family with a thin random graph:

Theorem 5. *Let $G_L(n,p)$ denote the combination of instances of a locally clustered graph family $G_L(n)$ and a $G(n,p)$ graph where $p = \frac{1}{f(n)\cdot n}$, $\frac{1}{(\log n)^{1+(\epsilon/2)}} \leq f(n) \leq \frac{1}{n^{1-\delta}}, \epsilon > 0, \delta > 0$. $D(s(n,p))$ denotes the maximal diameter of any subgraph of $G_L(n,p)$ with size $s(n,p) = \frac{(\log n)^{1+\epsilon}}{p\cdot n}$, $\epsilon > 0$, the diameter of $G_L(n,p)$ is asymptotically at most:*

$$D\left(s\left(n, \frac{1}{f(n)\cdot n}\right)\right) \cdot \underbrace{\left(\frac{\log(n/f(n)}{\log(\log n)^{1+\epsilon}}\right)}_{D(G_S)} \tag{28}$$

As we will discuss in the next section, the surprise in the whole small-world discussion lies in the fact that both graph components alone will have a much higher diameter than $O(\log n)$. Nonetheless, in our small-world model that can be based on any *locally clustered* graph family, also the use of cliques (complete graphs) is possible. There, the diameter is still bound from above by Theorem 5 but this bound is not tight: the diameter is 1. This brings us to the questions: What are small-worlds? We will discuss this question in the next section.

4 A new definition for small-world network models

We have shown in the introduction that the classic definition for small-worlds is somewhat erroneous and misleading. Our impression is that this stems from the following: The small-world generating process proposed by Watts and Strogatz required a random process for creating short-cuts [23]. But the classification of networks as small-worlds was not based on the recognition of this process but on a combination of structural measurements, i.e., the high clustering coefficient and the small diameter. This combination of network characteristics is not able to differentiate between those networks that include a local network and a random network and networks generated by other processes. In summary, the small-world network generation was process-oriented whereas the classification of small-world networks was result-oriented and there is no direct one-to-one matching between both sets of networks.

The drastic interest in the small-world phenomenon based on the classic small-world model seems to be based on two effects:

1. There is no centralized organization of the small-world network: every vertex builds random edges independently from others
2. The hybrid graph of a local and a random network component has a significantly lower diameter than the minimum of the diameter of boths components

To cover the first point there needs to be a *process-oriented classification of small-world networks* because the result may not always tell which generation process was behind it. The second effect excludes all those network models as small-world models in which the small diameter is an inbuilt feature of one component: If we have n unconnected vertices and add a dense random graph to it,

there is no surprise that the hybrid graph's diameter scales with $O(\log n)$. Also, the addition of some random edges to a clique will show the same diameter as the clique alone.
These two aspects lead us to the following new definition for small-world generating models:

Definition 6. A small-world model is defined as any combination $G_{LR}(n)$ of a *restricted locally clustered* graph family $G_L(n)$ and a random graph family $G_R(n)$ where the diameter $D(G_{LR}(n))$ is at most scaling poly-logarithmically and where the following relations hold for $n \to \infty$:

$$\frac{D(G_L(n))}{D(G_{LR}(n))} \to \infty \quad \text{and} \quad \frac{D(G_R(n))}{D(G_{LR}(n))} \to \infty \tag{29}$$

A small-world network is then a network that is generated by a small-world generating model - in real-world systems one may rather speak of small-world generating processes. Note that for *restricted locally clustered* graph families not any combination with a $G(n,p)$ graph may be appropriate. If it is a *locally clustered* graph family then it can be combined with any $G(n,p)$ under the conditions given in Theorem 5.

This definition removes the above given problems and captures - in our opinion - both effects that are the basis for the small-world phenomenon. It includes all classical small-world models because they are based on *(restricted) locally clustered* graphs [23, 15, 16, 12, 10, 4]. Note that these models may not be classified as small-world generating models for all combinations of their parameters. For the classic Watts-Strogatz-model or the Newman-Watts-model a very high k, e.g., $k = n-1$ would result in a clique as the local component. The above given definition restricts the models to those cases where the small-world effect is the result of the combination of both components and not of one of the components alone. The restriction on those cases reflects a similar decision made in the classic definition where the number of local edges per vertex k was required to be much smaller than n and also $p \ll 1$. We want to discuss two small-world generating models a bit more detailed with respect to the above given definition: In the case of the Kleinberg-small-world model the random graph component has not yet been analyzed with respect to its diameter to our knowledge. Kleinberg uses a random graph in which for each vertex v, q edges are added. Every vertex has a position in a 2-dimensional grid and edge $e = (v, w)$ is drawn with a probability proportional to $(d(v, w))^{-2}$ [16]. Kleinberg shows that a small-world network emerges for $q = 1$. Although there is no analysis available on the diameter of this random graph family, it seems quite likely that the described component with only one random edge per vertex is not even connected and will thus show a diameter of ∞.

The other small-world model to be analyzed was given by Chung et al. [10, 4]. In their basic model, the random graph component is a power-law random graph. Following [9], this component alone has almost surely a diameter of $O(\log n)$. On the other hand, Chung et al. have shown that for (k, l)-local graphs with certain properties, the resulting hybrid graph has a diameter of, e.g., $O(\log \log n)$. This is

clearly a new, emerging characteristic of the hybrid graph that is not dominated by one of the components alone. Thus, in the tradition of Watts and Strogatz we regard only the second, specialized cases as small-world generating models. We just want to mention here that it might be possible to make the power-law random graph sparser in the general Chung-small-world model but this analysis has not yet been conducted.

We summarize that these classic small-world models are represented by the above given definition.

5 Discussion

In this paper we have proposed a general framework for the design of small-world generating models: We have shown that they can be combined of (restricted) locally clustered and random graph models. This provides high flexibility in tuning a model to simulate a given real-world complex system. We have given a generalized theorem that describes an upper bound for the diameter of these hybrid graphs in dependence of the structure of both, the local and random component. Based on this framework we have proposed a very broad definition of small-world generating models, incorporating all classic small-world models.

Watts and Strogatz have provided us with the first formal model for generating small-world networks. The beauty of their model lies in its simplicity and clarity. Although we have questioned the usefulness of measuring the clustering coefficient as an indicator for small-worldness of real-world networks we have not given an alternative with which real-world networks can be identified as small-worlds. Admittedly, this is much more complicated in our framework than in the simple model of Watts and Strogatz. The only way to decide whether a real-world network is a small-world network in the above given definition is to analyze its generating process. If this is based on a local and a random component than the network should be regarded as a small-world if the diameter is short and each component alone has a high diameter.

Chung and Lu proposed to partition a given real-world network in a local and a global, random component [4]. They provide an approximative partitioning algorithm which works fine but is depending on two parameters that may not always be known in advance.

The partitioning into two components is also our suggestion: If there is any additional information about the network this could be used to partition it into its local and random component. For example, if the building of edges in a real-network is associated with a cost this could be regarded as a measure of distance. Then, both components can be separated and analyzed. Here, we can rehabilitate the clustering coefficient: If the global - presumably random - component shows a high clustering coefficient then it can be safely concluded that with high probability this component is not the result of a classic random process.

Of course, our definition of small-world generating components is somewhat influenced by our personal impression of what small-worlds really are. Since we are aware of that problem we want to conclude our discussion with the

introductory quotation of Wiener and Rosenblueth in "The Role of Models in Science":

> *[...] the best material model for a cat is another, or preferably the same cat.* - N. Wiener and A. Rosenblueth [24]

References

1. L. Adamic. The small world web. In *Proceedings of ECDL'99 - Lecture Notes on Computer Science*, pages 443–452, 1999.
2. L. Adamic and E. Adar. How to search a social network, 2004.
3. L.A.N. Amaral, A. Scala, M. Barthelemy, and H.E. Stanley. Classes of small-world networks. *PNAS*, 97(21):11149–11152, 2000.
4. R. Andersen, F. Chung, and L. Lu. Analyzing the small world phenomenon using a hybrid model with local network flow. In *Algorithms and Models for the Web-Graph; Third International Workshop, WAW 2004*, LNCS 3243, pages 19–30, 2004.
5. P.S. Bearman, J. Moody, and K. Stovel. Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, 110:44–91, 2004.
6. B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics 73. Cambridge University Press, London, 2nd edition, 2001.
7. B. Bollobás and F. Chung. The diameter of a cycle plus a random matching. *SIAM Journal on Discrete Mathematics*, 1:328–333, 1998.
8. B. Bollobás and O. M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks (S. Bornholdt, H.G. Schuster (eds.))*, pages 1–34, 2003.
9. F. Chung and L. Lu. The average distances in random graph with given expected degrees. *PNAS*, 99(25):15879–15882, 2002.
10. F. Chung and L. Lu. The small world phenomenon in hybrid power law graphs. In *Complex Networks (E. Ben-Naim, H. Frauenfelder, Z. Toroczkai (eds.))*, pages 91–106, 2004.
11. P.S. Dodds, R. Muhamad, and D. J. Watts. An experimental study of search in global social networks. *Science*, 301:827–829, 2003.
12. S.N. Dorogovtsev and J.F.F. Mendes. Exactly solvable analogy of small-world networks. *Europhys. Lett.*, 50, 2000.
13. E. N. Gilbert. Random graphs. *Anual Math. Statist.*, 30:1141–1144, 1959.
14. A. Iamnitchi, M. Ripeanu, and I. Foster. Small-world file-sharing communities. In *Proceedings of the IEEE INFOCOM 2004*, 2004.
15. J. Kleinberg. Navigation in a small world. *Nature*, 406:845, 2000.
16. J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.
17. L.F. Lago-Fernandez, R. Huerta, F. Corbacho, and J.A. Siguenza. Fast response and temporal coherent oscillations in small-world networks. *Phys. Rev. Letters*, 84:2758–2761, 2000.
18. F. Liljeros. Sexual networks in contemporary western societies. *Physica A*, 338:238–245, 2004.
19. S. Milgram. The small world problem. *Psychology Today*, 1:61–67, 1967.
20. M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Phys. Lett. A*, 263:341–346, 1999.

21. M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 60:7332–7342, 1999.
22. M.C. Gőpfert and D. Robert. The web of human sexual contacts. *Nature*, 411:907–908, 2001.
23. D. J. Watts and S. Strogatz. Collective dynamics of small world networks. *Nature*, 393:440–442, 1998.
24. N. Wiener and A. Rosenblueth. The role of models in science. *Philosophy of Science*, 12(4):316–321, 1945.

6 Appendix: *k*-next neighborhood graphs are *locally clustered*

The k-next neighbor graph family or in short, knn-graphs, belongs to the class of *locally clustered* network families as we will show in this section. The proof is applied only to 2-dimensional knn-graphs defined as:

A k-next neighborhood graph $G(n,k)$ is any possible instance of n vertices, uniformly distributed in a two-dimensional unit-square, where every vertex is connected with its k geometrically next vertices.

Note that the relation is not symmetric and therefore the knn-graph is a directed graph and thus we differentiate between the outgoing and ingoing edges of a vertex.

In order to prove the given property we will proceed in the following steps:

1. Bound the distance to next neighbors from above and below.
2. Prove that a knn-graph is highly likely connected.
3. Show that a generic partition procedure yields the required n/s subgraphs.

6.1 A bound for the maximum distance of nearest neighbors

Let the knn-disk of any vertex v be defined as the minimal disk which contains all of its k next neighbors. Note that the disc-radius is equal to the maximum distance of any connected nearest neighbor to v.

The probability for any vertex v to be placed in some area A within the unit square is exactly A. Thus, the placement of vertices into a given area is a Bernoulli trial with $p = A$ and $q = 1 - A$. Therefore the Chernoff bound may be applied to yield an upper bound for the diameter. Moreover, the expected radius is given by $\bar{r} = \sqrt{\frac{k}{\pi \cdot n}}$. The result for the upper bound is given in

Lemma 7. *Let $\hat{r} = \sqrt{\hat{c}} \cdot \bar{r}$ denote a knn-disk radius with $\hat{c} = 3 + \sqrt{8}$. Further let $k \geq \log n$.*
With high probability ($Pr[...] \geq 1 - 1/n$), no disk with radius $\hat{r}$ around any vertex v exists that does not contain at least k vertices.

Proof. Let D_v denote a knn-disk around v with an expected number of vertices lying in that disk equal to $\bar{k} = c \cdot k$. X_k denotes the number of vertices lying inside of D_v. Now we apply a relaxed version of the Chernoff inequality for independent Bernoulli trials. With $\mu = c \cdot k$ and $\delta = 1 - \frac{1}{c}$

$$Pr[X_k < (1-\delta)\mu] < e^{-\frac{1}{2}\mu\delta^2} = e^{-\frac{ck}{2}(1-\frac{1}{c})^2} < n^{-\frac{c}{2}(1-\frac{1}{c})^2} \tag{30}$$

For $c = \hat{c} = 3 + \sqrt{8}$ we yield

$$Pr[X_k < (1-\delta)\mu] < \frac{1}{n^2} \tag{31}$$

Hence, the probability that there is a knn-disk with radius larger than $\hat{r} = \sqrt{3 + \sqrt{8}} \cdot \bar{r}$ in a knn-graph is $< 1/n$.

The interpretation of this result is that it is almost impossible for $n \to \infty$ that any knn-disk exists with a radius larger than $\hat{r}$. Therefore in our following theorems, we consider the radius of knn-discs to be bound by $\hat{r} = \sqrt{\hat{c}} \cdot \bar{r}$.

Note that theses equations are only valid for disks that do not intersect with the unit squares border. If a vertex v_c were to be positioned in a corner of the unit disc, a factor of 2 would apply to the radius.

6.2 Connectedness of knn-graphs

We will now show that a knn-graph is **whp** connected.

The proof for the following lemma will be ommited. Here we will just sketch it shortly: As can be seen in Fig. 6.2, in every unconnected knn-graph the pair of closest vertices lying in different components have an angle of at least 120° in which none of their k next neighbors is placed. A simple stochastic argument shows that the probability for this is given by $(2/3)^k$. Equating this with the probability bound of $1/n$ and solving the equation to k yields the needed k such that with high probability not even one vertex with the above mentioned property exists. This leads to the following lemma:

Lemma 8. *A knn-graph is connected with high probability for* $k \geq \frac{\log n}{\log(3/2)}$

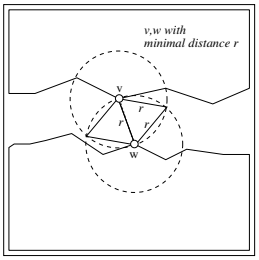


Fig. 3. v and w are two vertices from different connected components of a *knn*-graph having minimal euclidian distance to each other. Two circles are drawn around v and w, respectively, with a radius that equals the euclidian distance between v and w. The figure shows that none of the k-next neighbors of neither v nor w can exist in the intersection of these circles without contradicting the condition that v and w are the pair of vertices from different connected components with minimal euclidian distance.

Note that the probability for an unconnected *knn*-graph is smaller than $1/n$ since the existence of a vertex with the above given property is only necessary for an unconnected graph but certainly not sufficient.

Having argued that knn-graphs are connected, we will now show that connected commensurate partitions can be found.

6.3 Constructing the partition

The following procedure constructs partitions as required by the definition for *locally clustered* graph families. The size of the subgraphs is depending on the added random graph family. Nevertheless, if the random graph is given, one can easily calculate the fixed subgraph of size s. The definition requires that for each pair of vertices a partition into $\Theta(n/s)$ subgraphs must exist, so that both vertices are included in full subgraphs.

For each pair v, w of vertices we construct slightly different partitions. For each of them, we start with a geometric partition, based on squares containing at least $4/\pi \cdot s$ vertices. The exact positions for the squares are chosen such that both vertices are contained in full subgraphs. Beside this requirement the positions of the squares are arbitrary as long as the number of squares placed completely inside the unit square is maximal. Note, that a constant relative fraction of vertices may exist, that is not contained in any subgraph. Each of the squares covers an area A_s so that with high probability at least $4/\pi \cdot s$ vertices are geometrically contained in each of them. The area is given by $A_s > 4/\pi \cdot \pi \cdot \hat{r}^2$, where $\hat{r}$ denotes the maximal expected knn-disc radius (Lemma 7).
The maximal (centered) circle (Fig. 6.3) within each A_s contains only vertices from the same connected component. Otherwise at least one vertex would have an arc of more than 120° without any knn-edge which is highly unlikely as was already shown in Lemma 6.2. The area of this circle covers $\pi/4$ of A_s. We expect therefore that at least $\pi/4 \cdot 4/\pi \cdot s = s$ vertices from the same connected component for each A_s. As we explained before, for each constructed partition a constant fraction of vertices can be disregarded.

The overall result of this section is summarized by

Lemma 9. *The family of k-nearest neighbor graphs $G_k(n, k)$ on a point set in a 2-dimensional Euclidean space is locally clustered.*

Figure 6.3 shows an example for a partition for $s = 2$.

Note that the expected diameter $D(S_i)$ of the subgraphs is expectedly scaling with $O(\sqrt[d]{s})$ as it is the case with grid graphs.

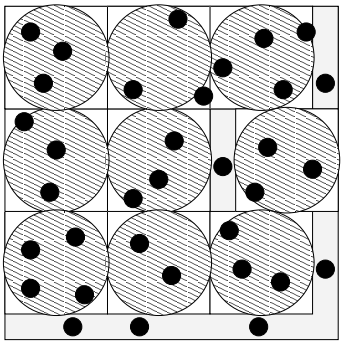


Fig. 4. This figure shows the result of the partitioning procedure for $s = 2$ as described in 6.3. Each square contains more than $4/\pi \cdot s \approx 2.5$ vertices. Each circle within any quadratic region contains at least $s = 2$ vertices that must form a connected subgraph. Note that the distribution of points is only schematic.

Counting loops in random graphs and real-world networks

E. Marinari [1], R. Monasson [2] and G. Semerjian [3]

[1] *Dipartimento di Fisica and INFN, Università di Roma La Sapienza, P. A. Moro 2, 00185 Roma, Italy.*

[2] *CNRS-Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75005 Paris, France,*

[3] *Dipartimento di Fisica and SMC-INFM,*

Università di Roma La Sapienza, P. A. Moro 2, 00185 Roma, Italy.

(Dated: July 25, 2005)

Data gathering in fields as diverse as social sciences, biology or Internet measurements, has provided an impressive amount of knowledge on the topology of the underlying interaction networks (i.e. graphs) in these domains. A crucial direction of research is now to identify characteristic features of these networks, in particular in view of (in)validating proposed modelings.

These features can be roughly classified in two categories. "Local" ones, for instance the connectivity distribution and the clustering coefficient, can be efficiently computed on any graph, even if very large. However the most distinguishing features might be "global", involving large patterns of the network. Identifying and counting these patterns becomes a very demanding task when their sizes increase, prohibiting in general the use of exact counting algorithms. Among these global features, a very natural one is the distribution of the lengths of the circuits (closed loops) in the graphs. Whereas it is quite easy to measure the number of short loops (triangles for instance), this becomes very hard when the loops studied have a length of the order of the size of the network.

We have developed in [1] an alternative approach to this counting problem, based on its reformulation in terms of a statistical mechanics model, which is treated within the Bethe approximation. The outcomes of this method are of two types. On the one hand, it yields an efficient approximate counting algorithm based on a message passing procedure, with a computational cost linear in the size of the graph. On the other hand, we have also obtained results on the typical number of long circuits in ensembles of random graphs, in particular in situations where usual probabilistic methods fail because of large fluctuations in these numbers.

---

[1] E. Marinari, R. Monasson and G.Semerjian, `cond-mat/0507525`.

On the Propagation of Congestion Waves in the Internet

József Stéger[a,b] Péter Vaderna[a,c] Gábor Vattay[a,b]

[a]Dept. of Physics of Complex Systems Eotvos University, Budapest
[b]Collegium Budapest, Institute for Advanced Study
[c]Traffic Analysis and Network Performance Laboratory Ericsson Hungary Ltd.

Keywords: Computer networks, TCP/IP protocol, congestion propagation

Abstract

Traffic modeling of communication networks such as Internet has become a very important field of research. A number of interesting phenomena are found in measurements and traffic simulations. One of them is the propagation of congestion waves opposite to the main packet flow direction. The purpose of this short paper is to model and analyze packet congestion on a given route and to give a possible explanation to this phenomenon.

Corresponding author:
József Stéger
Department of Physics of Complex Systems,
Eötvös Loránd University
H-1518 Budapest Pf.: 32, Hungary.
Tel: +36 1 3722896, Fax: +36 1 3722866
E-mail: steger@complex.elte.hu

1 Introduction

In the recent past many aspects of computer data traffic on the Internet have been investigated. Signatures of long-range correlations [1], scaling [2], chaos [3] and phase transitions [4] have been found. In Ref. [5] $1/f^{\alpha}$ noise has been observed in the time series of round trip times similar to those observed in highway traffic measurements [6]. Computers connected to a network communicate via data flows consisting of discrete data packets. The flow of these packets along a network path is quite similar to a one-dimensional granular flow of matter through a pipe. The basic concepts of the kinetic model of this traffic has been proposed by Antoniou et al. [7]. Since the slowing down and acceleration of packet flow in computer networks are very similar to those of cellular automaton models of cars in highway traffic [8], it has been argued that a valid analogy exists between these subjects which later has been demonstrated quantitatively [9, 10]. One of the most peculiar features of car traffic and flow of granular media is the propagation of density waves [11]. In car traffic stop-and-go type congestion waves propagate against the direction of the flow. In Ref. [12] the authors made an attempt to observe the propagation of congestion in the routers of a real computer network. They studied the spatio-temporal correlations of the level of congestion in routers and showed that congestion can propagate from a heavily loaded router to one of its empty neighboring routers.
Today's computer networks can be studied with the help of network simulators, which enable us to assemble any computer network configuration, to use the most commonly used packet sending mechanism TCP/IP protocol and to emulate its behavior without building the system from hardware components. These simulators are developed to reproduce the behavior of computers and routers accurately by implementing their real transport layer protocols in the simulators. Due to some minor differences between the real protocols and their implementation in the simulator, minor details of simulation results may differ from the corresponding outcome of real networks, which is well documented in the engineering literature [14]. In our study we use the popular Berkeley Network Simulator (`ns-2.26` [13]).
In this paper we study the network traffic generated in a unidirectional ring of identical routers connected. This way we can study the propagation of congestion in an isolated, clean setup, where the effects of inhomogeneity and the complex topology of the real Internet does not interfere with the basic mechanism creating the congestion wave. We show that this system drives itself in a self-organized way into a critical congested state, where the system is overloaded both the position of the congested router and the rate of the packet sending activity at the sites propagate against the direction of the packet flow. The profile of the congestion wave is reconstructed and its speed is measured. Then a simple model of the system is introduced which is able to explain the main features of the congestion wave such as shape and speed.

2 The model

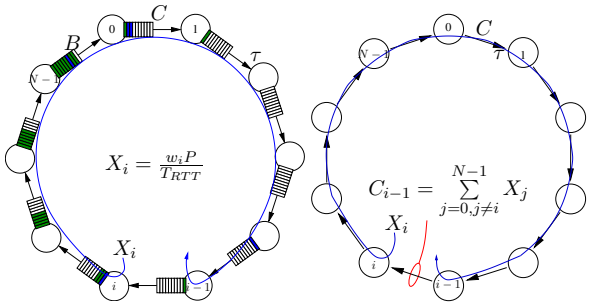


Figure 1: The ring structure. **Left,** Network simulator: TCP agents at all sites i continuously transfer data trough the network with lines of capacity C and time delay τ. Unprocessed packets queue in buffers of size B. The stream of data can traverse in clockwise direction only to arrive at the destination node terminal at site $i-1$. ($C = 10^7$ bits/s, $\tau = .031$ s, $B = 300$ packets, $P = 4416$ bits, $N = 10$). **Right,** Continuous model: Each TCP is modeled with its sending rate X_i and its traffic flow traverses the ring topology. On the link connecting sites $i-1$ and i all the sending rates of traversing TCP flows are summed up to yield the total link utilization.

In our model system a ring is formed by identical routers, which can forward packets in clockwise direction as depicted in the left part of Fig. 1. Incoming data flow in a router, which is a mixture of injected packet and the background traffic, can temporarily exceed the capacity of the outgoing line. To avoid data loss in this situation routers contain finite storage buffers. Computers are instructed to send data to their anti-clockwise neighbors, so that the packets traverse the longest possible route in the ring. The dynamics of the data traffic of computers is controlled by the

TCP/IP protocol [15]. This protocol ensures that the data packet-sending rate is decreased whenever congestion occurs and that it is increased when there is an available unused capacity in the system.
After establishing connection between two computers over the network TCP algorithm regulates the packet-sending rate. First a single packet is sent out. Upon receiving it the receiver acknowledges the arrival of the packet by sending back a small size acknowledgement packet (ACK). The time elapsed between the sending out of a packet and receiving the corresponding ACK is called round trip time (RTT). The TCP maintains an internal variable, the Congestion Window (w), which is used to control the number of packets sent out when the ACK is received. Starting with $w = 1$ it is increased according to $w \mapsto w + 1/w$ each time an ACK is received. Two new packets are sent out if the congestion window crosses an integer value and only a single packet otherwise. Assuming constant RTT during this process, the congestion window and the number of packets out in the network are increased linearly in time, until a packet is lost somewhere in the network, indicating congestion. As a response the packet-sending rate should be decreased, the TCP drops the value of the congestion window $w \mapsto \beta w$ ($\beta < 1$) and is silent in response to ACKs until the number of still unacknowledged packets decreases to the integer part of the new, reduced value of the congestion window. After that the packet-sending algorithm returns to the original linear increase phase described above.
In case the congestion window variable is large we can neglect its granularity and can treat it as a continuous variable. The sending rate X (bit/sec) can be estimated as the amount of data sent within an RTT, $X = Pw/T_{RTT}$. Neglecting the change of RTT on the scale of RTTs the sending rate satisfies the following pair of equations:

$$\frac{dX}{dt} = \frac{P}{T_{RTT}^2(t)}, \tag{1}$$

$$X(t_+) = \beta X(t_-) \quad \text{at packet loss.} \tag{2}$$

We present the results of our simulation study carried out with the network simulator. Fig. 2 shows the spatiotemporal diagram of the congestion wave occurring in the network simulator. One can see that after a short initial transient (up to 500 s) the pattern remains stable and propagates in anti-clockwise direction. In this respect it resembles the congestion propagation in car traffic. The speed of the congestion wave pattern is almost constant.

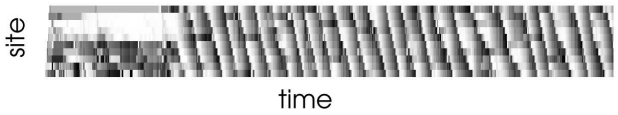


Figure 2: Spatiotemporal diagram of congestion propagation. The horizontal axis is the time (covering 2000 seconds) and the site index (i) is on the vertical axis. In this simulation the number of sites was $N = 10$. The shade of the figure represents the sending rate X_i. Light patches indicate very low sending rate due to high congestion.

Representing the sending rates $X_{i'+\lfloor\langle i\rangle\rfloor}(t)$ in co-moving coordinates i' relative to the center of mass we recover the shape of the traveling wave pattern. Averaging the new series in time the profile of the front emerges as in Fig. 3 left.

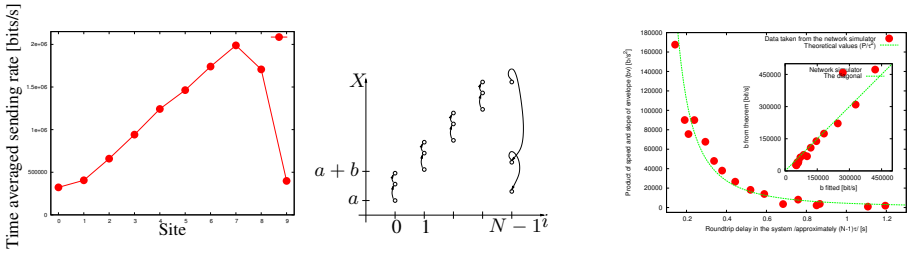


Figure 3: **Left,** The shape of the traveling wave profile, determined by averaging the time series in co-moving spatial coordinates. **Middle,** Sending rate evolution in time. $N = 5$, $l = 2$, X_4 decreases until it reaches the value of X_0. **Right,** Testing the relation $vbT_{RTT}^2/P = 1$ for various system sizes in `ns`. Values of b, T_{RTT} and v have been obtained by averaging over samples of 5000 sec. **Inset:** Testing Eq. 6 in `ns`. Average values of b are plotted against the results of (6) with q^l obtained from the simulation.

While the continuous equations constitute gross simplification of the original TCP dynamics, the main properties of the traveling wave can be recovered from them with some additional assumption made on the packet loss process as we show next.
The bandwidth $C_{i-1}(t)$ utilized on the link connecting nodes $i-1$ and i is the sum of sending rates of TCPs whose traffic flows through that link. In our case the flows of all TCPs traverse that link except the one starting at node i and

ending at node $i-1$:

$$C_{i-1}(t) = \sum_{j=0, j\neq i}^{N-1} X_j(t) = \sum_{j=0}^{N-1} X_j(t) - X_i(t), \tag{3}$$

where site $i = N$ is identified with site $i = 0$ due to periodicity, see Fig. 1. Congestion and packet loss occur in the system whenever one of the bandwidths $C_i(t)$ reaches the link capacity C. According to (3) the largest link utilization $C_i(t)$ is at site $i = i^* - 1$ where i^* is the site where the sending rate $X_{i^*}(t)$ is the lowest. In principle all the TCP flows traversing the congested link can lose packets, so only the TCP at site i^* is immune. Our observation is that the TCP flow starting at the actual congested link (with sending rate X_{i^*-1}) experiences the packet loss almost surely.
The mechanism described above is responsible for the emergence of the congestion wave in the system. The TCP at site $i^* - 1$ suffers packet losses repeatedly until its sending rate X_{i^*-1} becomes smaller than X_{i^*}. From then on X_{i^*-1} will be the lowest in the system, link utilization C_{i^*-2} will be the highest and TCP at site $i^* - 2$ suffers the packet losses. This way congestion propagates site by site anticlockwise in the system. After several rounds of congestion propagation the propagating front of Fig.3 emerges.
The shape of the front is linear with a sharp drop connecting its ends. We can determine the parameters of the linear front in our model. Let the nodes forming the linear part of the front range from 0 to $N-1$ and let the minimum sending rate first be at the 0th node

$$X_i(0) = a + bi, \tag{4}$$

as it is shown in Fig. 3 middle that illustrates the evolution of a system containing $N = 5$ nodes. The packets start to be dropped at site $N-1$. We start our description at the first packet drop, which occurs when $C_{N-1}(0) = \sum_{i=0}^{N-1} X_i(0) - X_0(0) = C$ holds. This initial condition gives the first condition for the initial shape of the front

$$C = (N-1)\left(a + \frac{N}{2}b\right). \tag{5}$$

Immediately after the packet drop the sending rate at site $N-1$ decreases to $X_{N-1}(0+) = \beta X_{N-1}(0-)$, while the rest of the sending rates stay unchanged. Then the sending rates increase in a uniform manner with an amount $X'_i = X_i + (1-\beta)X_{N-1}/(N-1)$ until the next packet loss occurs. In particular, from before the first packet loss until the second packet loss the sending rate at site $N-1$ changes to $X'_{N-1} = qX_{N-1}(0)$, where $q = \beta + \frac{1-\beta}{N-1}$. This process is then repeated l times until the sending rate $q^l X_{N-1}$ becomes lower than the actual value of X_0. This way the wave moves one site to the left, while its linear shape is preserved as the sending rates of all nodes, except $N-1$, increase with b. From Eq. 1 one can calculate the time needed for this $T_p = bT_{RTT}^2/P$. Accordingly the speed of the congestion wave is $v := 1/T_p$ (measured in site/sec.).
The formula derived for the speed of the congestion waves can be tested against the data produced by the network simulator. In the simulation one can measure the average round trip time observed by the TCPs, the mean slope of the linear part of the front b and the speed of the front. On the main part of Fig. 3 right one can see that the measured values satisfy the relation $vbT_{RTT}^2/P = 1$ very well. With the help of (5) and (4) the slope of the front in the model can also be directly expressed as

$$b = \frac{2C(1-q^l)}{(N-1)(N-2)(q^l+1)}, \tag{6}$$

where l should be determined independently. Our mathematical model allows several positive integer values of l with an upper bound due to the positiveness of b. We found numerically that only the largest possible l value is stable against small random perturbations of the wave front. Systems started at lower l values always shift towards a greater value of l. In the network simulator we always observed the realization of the most stable (the highest possible l) solutions of the model. In the inset of Fig. 3 right we compare the measured values of b with (6). We again find good agreement.

3 Conclusion

As a summary we showed that congestion waves are formed naturally in the data traffic of computer networks. The mechanism behind the wave formation is that packet losses occur most likely in computers nearest to the site of the actual congestion and other computers sharing the congested link increase their sending rates, moving the site of the congestion one site downstream. This basic mechanism is quite general and can create congestion moving against the direction of the data traffic in more complicated geometries. A formula for the speed of the congestion wave has been derived in a simple ring topology and network simulations have confirmed it. Such formulas can be developed for more complicated geometries, which is our next research goal.

Acknowledgements

The authors thank the partial support of the National Science Fund Hungary (OTKA T37903), the National Office for Research and Technology (NKFP 02_032_04) and the EU IST FET COMPLEXITY EVERGROW Integrated Project.

References

[1] W. E. Leland, M. S. Taqq, W. Willinger and D. V. Wilson, Proceedings of ACM SIGCOMM, San Francisco, California, Ed. Deepinder P. Sidhu, p. 183–193 (1993); K. Fukuda, H. Takayasu and M. Takayasu, Physica **A 287**, 289-301 (2000);

[2] S. Valverde, R. V. Sole, Physica **A 312**, 636 - 648 (2002), J. Yuan, Y. Ren and X. Shan, Phys. Rev. **E 61**, 1067 (2000)

[3] A. Veres, M. Boda, Proceedings of IEEE INFOCOM Tel Aviv, March 2000, p. 1715–1723 (2000) http://www.ieee-infocom.org/2000/papers/74.pdf

[4] M.Takayasu, H.Takayasu, and K.Fukuda, Physica **A 277**, 248-260 (2000)

[5] I. Csabai, J. Phys. **A27**, L417 (1994);

[6] T. Musha and H. Higuchi, Japan Appl. Phys. **15** 1271 (1976)

[7] I. Antoniou, V. V. Ivanov and Yu. L. Kalinovsky, Physica **A 308**, 533 (2002)

[8] D. Chowdhury, L. Santen and A. Schadschneider, Statistical physics of vehicular traffic and some related systems. Physics Reports, 329(4-6):199-329 (2000); B. S. Kerner and P. Konhuser. Structure and parameters of clusters in traffic flow. Physical Review E, 50(1):54 (1994); K. Nagel and M. Paczuski. Emergent traffic jams. Physical Review E, 51:2909 (1995) O. Biham, A. Middleton, and D. Levine. Self-organization and a dynamical transition in traffic-flow models. Physical Review A, 46:R6124 (1992)

[9] G. Simon, I. Csabai The analogies of highway and computer network traffic Physica **A 307** 516 - 526 (2002)

[10] T. Huisinga, R. Barlovic, W. Knospe, A. Schadschneider, M. Schreckenberg, Physica **A 294**, 249 (2001)

[11] K. Nagel and M. Schreckenberg, J. Physique **I 2** 2221 (1992); K. Nagel and H. J. Hemann, Physica **A199** 254 (1993); M. Leibig, Phys. Rev. **E 49** 184 (1994)

[12] K.Fukuda, H.Takayasu, and M.Takayasu, Fractals, **7**(1) 23-31 (1999)

[13] S. McCanne and S. Floyd. Ns (network simulator). Available via http://www-nrg.ee.lbl.gov/ns. (1995)

[14] S. Floyd and V. Paxson, Difficulties in simulating the Internet, IEEE/ACM Transactions on Networking, **9**(4) 392–403, (2001)

[15] Jacobson, V., *Congestion Avoidance and Control*, Proceedings of SIGCOMM '88, August 1988, pp. 314-329.

[16] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. Physical Review E, 51(2):1035 (1995); A. Schadschneider. Statistical physics of traffic flow. Physica **A 285**, 101 (2000);

Towards Peer-to-Peer Web Search (Extended Abstract)

Gerhard Weikum, Holger Bast, Geoffrey Canright, David Hales,
Christian Schindelhauer, Peter Triantafillou *

Contact author: Gerhard Weikum, Max-Planck Institute for Computer Science,
Stuhlsatzenhausweg 85, Saarbrücken, Germany, Email: weikum@mpi-inf.mpg.de

The peer-to-peer (P2P) computing paradigm is an intriguing alternative to Google-style search engines for querying and ranking Web content. In a network with many thousands or millions of peers the storage and access load requirements per peer are much lighter than for a centralized Google-like server farm; thus more powerful techniques from information retrieval, statistical learning, computational linguistics, and ontological reasoning can be employed on each peer's local search engine for boosting the quality of search results [1, 2, 10–12, 26]. In addition, peers can dynamically collaborate on advanced and particularly difficult queries. Moroever, a peer-to-peer setting is ideally suited to capture local user behavior, like query logs and click streams, and disseminate and aggregate this information in the network, at the discretion of the corresponding user, in order to incorporate richer cognitive models.

The DELIS project is aiming at a P2P system where each peer has a full-fledged Web search engine, including a crawler and an index manager. The crawler may be thematically focused or crawl results may be postprocessed so that the local index contents reflects the corresponding user's interest profile. With such a highly specialized and personalized "power search engine" most queries should be executed locally, but once in a while the user may not be satisfied with the local results and would then want to contact other peers. A "good" peer to which the user's query should be forwarded would have thematically relevant index contents, which could be measured by statistical notions of similarity between peers [3, 4]. Both query routing and the formation of "statistically semantic" overlay networks could greatly benefit from collective human inputs in addition to standard statistics about terms, links, etc.: knowing the bookmarks and query logs of thousands of users would be a great resource to build on [8, 14]. Note that this notion of Web search includes ranked retrieval and thus is fundamentally much more difficult than Gnutella-style file sharing or simple key lookups via distributed hash tables (DHTs) [24]. Further note that,

* The authors are with the Max-Planck Institute for Computer Science in Saarbrücken, Telenor in Oslo, the University of Bologna, the Heinz-Nixdorf Institute in Paderborn, and the University of Patras. The work presented in this paper is partially supported by the EU within the 6th Framework Programme under contract 001907 "Dynamically Evolving, Large Scale Information Systems" (DELIS).

although query routing in P2P Web search resembles earlier work on metasearch engines and distributed information retrieval [17, 18], it is much more challenging because of the large scale and the high dynamics of the envisioned P2P system with thousands or millions of computers and users. Finally, the P2P setting poses great challenges also for network-efficient top-k query processing [25, 16], decentralized and other advanced forms of link analyis [6, 7, 13, 14, 21], distributed gathering and dissemination of statistics about data, load, and user behavior [16, 19, 20], and the creation of self-organizing overlay networks [9, 15, 22, 23, 27].

A system architecture for the envisioned solution is currently prototyped, as an experimental platform within the DELIS project, under the name *Minerva* [5]. This system has all the characteristics and poses the challenges of a complex system. The autonomy of peers and the diversity of different behavioral patterns can be understood only by analyzing and controlling the system at different levels, ranging from the underlying physical network and the virtual overlay network layers to the level of intelligent search, query routing, and collaboration strategies of the individual peers. For cost-efficient solutions it is crucial to consider benefit and cost factors at all levels. Finally, a deep understanding mandates studying such complex systems at different scales in terms of time and space, for example, the short-term interactions of a peer with its immediate neighborhood, triggered by query routing and query execution, on one hand, and the long-term, long-range evolution of the entire system, to organize itself into effective and robust semantic overlay structures, on the other hand.

References

1. Holger Bast, Debapriyo Majumdar. Understanding Spectral Retrieval via the Synonymy Graph, ACM SIGIR Conf. on R&D in Information Retrieval, Salvador, Brazil, 2005.
2. Holger Bast, Ingmar Weber: A Framework for Ranked Retrieval based on Rank Aggregation, Int. Workshop on Web Information Retrieval and Integration, Tokyo, Japan, 2005.
3. Matthias Bender, Sebastian Michel, Gerhard Weikum, Christian Zimmer: Bookmark-driven Query Routing in Peer-to-Peer Web Search, ACM SIGIR Workshop on Peer-to-Peer Information Retrieval, Sheffield, UK, 2004
4. Matthias Bender, Sebastian Michel, Peter Triantafillou, Gerhard Weikum, Christian Zimmer: Improving Collection Selection with Overlap Awareness in P2P Search Engines, ACM SIGIR Int. Conf. on R&D in Information Retrieval, Salvador, Brazil, 2005.
5. Matthias Bender, Sebastian Michel, Peter Triantafillou, Gerhard Weikum, Christian Zimmer: Minerva: Collaborative P2P Search, Demo Paper, Int. Conf. on Very Large Data Bases (VLDB), Trondheim, Norway, 2005.
6. Klaus Berberich, Michalis Vazirgiannis, Gerhard Weikum: T-Rank: Time-aware Authority Ranking, Int. Workshop on Algorithms and Models for the Web Graph, Rome, Italy, 2004.
7. Geoffrey Canright and Kenth Engo-Monsen. Roles in Networks. Science of Computer Programming 53: 195–214, 2004.

8. Hang Cui, Ji-Rong Wen, Jian-Yun Nie, Wei-Ying Ma: Query Expansion by Mining User Logs. IEEE Transactions on Knowledge and Data Engineering 15(4), 2003
9. David Hales: From Selfish Nodes to Cooperative Networks – Emergent Link-based Incentives in Peer-to-Peer Networks. IEEE Conf. on Peer-to-Peer Computing, Zurich, Switzerland, 2004
10. Thomas Hofmann: Unsupervised Learning by Probabilistic Latent Semantic Analysis, Machine Learning 42(1/2): 177–196, 2001
11. Georgiana Ifrim, Martin Theobald, Gerhard Weikum: Learning Word-to-Concept Mappings for Automatic Text Classification, Int. Workshop on Learning in Web Search, Bonn, Germany, 2005.
12. Amy N. Langville, Carl D. Meyer: A Survey of Eigenvector Methods of Web Information Retrieval. The SIAM Review, 47(1):135-161, 2005.
13. Amy N. Langville, Carl D. Meyer: Deeper Inside PageRank. Internet Mathematics, 1(3):335-400, 2004.
14. Julia Luxenburger, Gerhard Weikum: Query-log based Authority Analysis for Web Information Search, Int. Conf. on Web Information System Engineering (WISE), Brisbane, Australia, 2004
15. Peter Mahlmann, Christian Schindelhauer: Peer-to-Peer Networks based on Random Transformations of Connected Regular Undirected Graphs, ACM Symp. on Parallelism in Algorithms and Architectures (SPAA), Las Vegas, 2005
16. Sebastian Michel, Peter Triantafillou, Gerhard Weikum: KLEE: Internet-scale Distributed Top-k Query Algorithms, Int. Conf. on Very Large Data Bases (VLDB), Trondheim, Norway, 2005.
17. Weiyi Meng, Clement T. Yu, King-Lup Liu: Building efficient and effective metasearch engines. ACM Computing Surveys 34(1): 48-89, 2002.
18. Henrik Nottelmann, Norbert Fuhr: Evaluating different methods of estimating retrieval quality for resource selection. ACM SIGIR Conf. on R&D in Information Retrieval, Toronto, Canada, 2003.
19. Nikos Ntarmos, Peter Triantafillou: AESOP: Altruism-Endowed Self Organizing Peers, 3rd Int. Workshop on Databases, Information Systems and Peer-to-Peer Computing, Toronto, Canada, 2004.
20. Nikos Ntarmos, Peter Triantafillou: SeAl: Managing Accesses and Data in Peer-to-Peer Data Sharing Networks, 4th IEEE Conf. in Peer-to-Peer Computing, Zurich, Switzerland, 2004.
21. Josiane Xavier Parreira, Gerhard Weikum: JXP: Global Authority Scores in a P2P Network, 8th Int. Workshop on Web and Databases (WebDB), Baltimore, USA, 2005.
22. Josiane Xavier Parreira, Sebastian Michel, Gerhard Weikum: p2pDating: Real Life Inspired Semantic Overlay Networks for Web Search, ACM SIGIR Workshop on Heterogeneous and Distributed Information Retrieval, Salvador, Brazil, 2005.
23. Christian Schindelhauer, Gunnar Schomaker: Weighted Consistent Hashing, ACM Symp. on Parallelism in Algorithms and Architectures (SPAA), Las Vegas, USA, 2005
24. Ion Stoica, Robert Morris, David Liben-Nowell, David R. Karger, M. Frans Kaashoek, Frank Dabek, Hari Balakrishnan: Chord: a scalable peer-to-peer lookup protocol for internet applications. IEEE/ACM Transactions on Networking 11(1): 17-32, 2003
25. Martin Theobald, Gerhard Weikum, Ralf Schenkel: Top-k Query Evaluation with Probabilistic Guarantees, Int. Conf. on Very Large Data Bases (VLDB), Toronto, Canada, 2004
26. Martin Theobald, Ralf Schenkel, Gerhard Weikum: Efficient and Self-Tuning Incremental Query Expansion for Top-k Query Processing, ACM SIGIR Conf. on R&D in Information Retrieval, Salvador, Brazil, 2005.
27. Peter Triantafillou, Chryssani Xiruhaki, Manolis Koubarakis, Nikos Ntarmos: Towards High Performance Peer-to-Peer Content and Resource Sharing Systems, 1st Int. Conf. on Innovative Data Systems Research (CIDR), Asilomar, USA, 2003.

Resource allocation on sparse graphs

K.Y. Michael Wong[1], David Saad[2] and Zhuo Gao[1,3]

[1]The Hong Kong University of Science & Technology, Clear Water Bay, Kowloon Hong Kong
[2]The Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK
[3] Department of Physics, Beijing Normal University, Beijing 100875, China.

Key words: Statistical physics, inference, message passing, resource allocation, sparse graphs

Optimal resource allocation is a well known problem in the area of distributed computing [1, 2] to which significant effort has been dedicated within the computer science community. The problem itself is quite general and is applicable to other areas as well where a large number of nodes are required to balance loads/resources, such as reducing internet traffic congestion [3]. The problem has many flavours and usually refers, in the computer science literature, to finding practical algorithmic solutions to the distribution of computational load between computers connected in a predetermined manner. Many of the solutions are heuristic and focus on practical aspects (e.g., communication protocols).

The problem we are addressing is more generic and is represented by nodes of some computational power that should carry out some tasks. Both computational powers and tasks will be chosen at random from some arbitrary distribution. The nodes are located on a randomly chosen sparse graph of some connectivity. The goal is to migrate tasks on the graph such that demands will be satisfied while minimising the migration of (sub-)tasks. Decisions on messages to be passed are carried out locally. We focus here on the satisfiable case where the total computing power is greater than the demand, and where the number of nodes involved is very large. We analyse the problem using both the Bethe approximation and the replica method [4] of statistical physics. The latter will not be discussed in this extended abstract but can be found in [5]. Insights gained from the analysis give rise efficient message passing distributed algorithms for solving the problem with a modest computational cost. The approach is based on passing local information between nodes to facilitate decisions about the movement of tasks.

The Bethe approximation: We consider a typical resource allocation task on a sparse graph of N nodes, labelled $i = 1, .., N$. Each node i is randomly connected to c other nodes[1] and has a capacity Λ_i randomly drawn from a distribution $\rho(\Lambda_i)$. The objective is to migrate tasks between nodes such that each node will be capable of carrying out its tasks. The *current* $y_{ij} \equiv -y_{ji}$ drawn from node j to i is aimed at satisfying the constraint $\sum_j \mathcal{A}_{ij} y_{ij} + \Lambda_i \geq 0$, representing the 'revised' assignment for node i, where $\mathcal{A}_{ij} = 1/0$ for connected/unconnected node pairs i and j, respectively. To illustrate the statistical mechanics approach to resource allocation, we consider the load balancing task of minimising the energy function (cost) $E = \sum_{(ij)} \mathcal{A}_{ij} \phi(y_{ij})$, where the summation (ij) runs over all pairs of nodes, subject to the above constraints; $\phi(y)$ is a general even function of the current y.

When the connectivity c is low, the probability of finding a loop of finite length on the graph is low, and the Bethe approximation describes well the local environment of a node. In the approximation, a node is connected to c branches in a tree structure, and the correlations among the branches of the tree are neglected. In each branch, nodes are arranged in generations. A node is connected to an ancestor node of the previous generation, and another $c - 1$ descendent nodes of the next generation.

We derive a recursion relation for calculating the free energy of the system, the average asymptotic cost and the current distribution. Similar results have been obtained from the replica approach [5]. Although both derivations have been formulated for general cost functions, we concentrate on the particularly simple case of $\phi(y) = y^2/2$, where one can compare the obtained solutions with known results.

[1]Although we focus here on graphs of fixed connectivity, one can easily accommodate any connectivity profile within the same framework; the algorithms presented later are completely general.

Distributed algorithms: The local nature of the recursion relation we obtained points to the possibility that the network optimisation can be solved by message passing approaches, which have been successful in problems such as error-correcting codes [6] and probabilistic inference [7]. The major advantage of message passing is its potential to solve a global optimisation problem via local updates, thereby reducing the computational complexity. For example, the computational complexity of quadratic programming for the load balancing task typically scales as N^3, whereas capitalising on the network topology underlying the connectivity of the variables, message passing scales as N. An even more important advantage, relevant to practical implementation, is its distributive nature; it does not require a global optimiser, and is particularly suitable for distributive control in evolving networks.

However, in contrast to other message passing algorithms which pass conditional probability estimates of *discrete variables* to neighbouring nodes, the messages in the present context are more complex, since they are *functions* of the current y. We simplify the message to 2 parameters, namely, the first and second derivatives of these functions. For the quadratic load balancing task, it can be shown that the message functions are piecewise quadratic with continuous slopes. This makes the 2-parameter message a very precise approximation. The message passed from node j to i, (A_{ij}, B_{ij}) becomes

$$A_{ij} \leftarrow -\mu_{ij}, \quad B_{ij} \leftarrow \Theta(-\mu_{ij}) \left[\sum_{k \neq i} \mathcal{A}_{jk} (\phi''_{jk} + B_{jk})^{-1} \right]^{-1}, \tag{1}$$

$$\text{where} \quad \mu_{ij} = \min \left[\frac{\sum_{k \neq i} \mathcal{A}_{jk} [y_{jk} - (\phi'_{jk} + A_{jk})(\phi''_{jk} + B_{jk})^{-1}] + \Lambda_j - y_{ij}}{\sum_{k \neq i} \mathcal{A}_{jk} (\phi''_{jk} + B_{jk})^{-1}}, 0 \right], \tag{2}$$

with ϕ'_{jk} and ϕ''_{jk} representing the first and second derivatives of $\phi(y)$ at $y = y_{jk}$ respectively. The forward passing of the message from node j to i is then followed by a backward message from node j to k for updating the currents y_{jk} according to $y_{jk} \leftarrow y_{jk} - \frac{\phi'_{jk} + A_{jk} + \mu_{ij}}{\phi''_{jk} + B_{jk}}$. For the quadratic load balancing task considered here, an independent exact optimisation is available for comparison. The Kühn-Tucker conditions for the optimal solution yields

$$\mu_i = \min \left[\frac{1}{c} \left(\sum_j \mathcal{A}_{ij} \mu_j + \Lambda_i \right), 0 \right]. \tag{3}$$

Numerical results: We exploit both the theoretical framework developed using methods of statistical physics and the message passing techniques mentioned above to study properties of the resource allocation problem with a quadratic cost function. The iterative solution of the free energy is obtained numerically using the recursion relation obtained from the Bethe approximation. We generate 1000 nodes at each iteration with capacities randomly drawn from the distribution $\rho(\Lambda) = \mathcal{N}(\langle \Lambda \rangle, 1)$, and each is being fed by $c-1$ nodes randomly drawn from the previous iteration.

Figure 1(a) illustrates the current distribution for various average capacities. The distribution $P(y)$ consists of a delta function component at $y = 0$ and a continuous component whose breadth decreases with average capacity. The fraction of links with zero currents increases with the average capacity. Hence at a low average capacity, links with nonzero currents form a percolating cluster, whereas at a high average capacity, it breaks into isolated clusters. As shown in Fig. 1(b), both the analytic results and the message passing algorithm Eq.(1) yield excellent agreement with the iteration of Eq.(3). Besides the case of $c = 3$, Fig. 1(b) also shows the simulation results of the average energy for $c = 4, 5$, using both Eqs. (1) and (3). We see that the average energy decreases when the connectivity increases. This is because the increase in links connecting a node provides more freedom to allocate resources. When the average capacity is 0.2 or above, an exponential fit $\langle E \rangle \sim \exp(-k\langle \Lambda \rangle)$ is applicable, where k lies in the range 2.5 to 2.7. Remarkably, multiplying by a factor of $(c-2)$, we find that the 3 curves collapse in this regime of average capacity, showing that the average energy scales as $(c-2)^{-1}$ in this regime (inset).

Further properties of the optimised networks have been studied by simulations, and will be presented elsewhere. Here we merely summarise the main results: (a) When the average capacity drops below

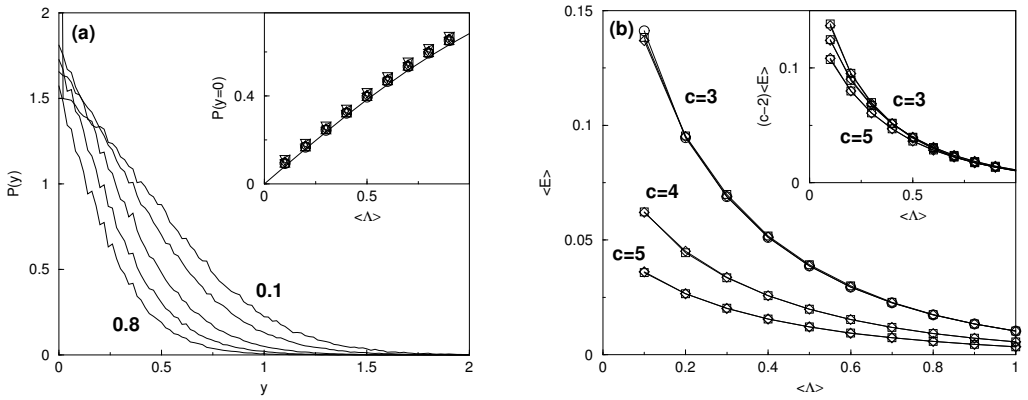


Figure 1: Results for system size $N = 1000$ and $\phi(y) = y^2/2$. (a) The distribution $P(y)$ obtained by iterating the recursive equations to steady states for $\langle \Lambda \rangle = 0.1$, 0.2, 0.4, 0.6, 0.8 from right to left. Inset: $P(y=0)$ as a function of $\langle \Lambda \rangle$. Symbols: $c=3$ (◯) and (□), $c=4$ (◇) and (△), $c=5$ (◁) and (▽); each pair obtained from Eqs. (1) and (3) respectively. Line: $\mathrm{erf}(\langle \Lambda \rangle/\sqrt{2})$. (b) Mean cost $\langle E \rangle$ as a function of $\langle \Lambda \rangle$ for $c = 3, 4, 5$. Symbols: results obtained by iterating the recursive equations to steady states (◯), Eq.(1) (□), and Eq. (3) (◇). Inset: $\langle E \rangle$ multiplied by $(c-2)$ as a function of $\langle \Lambda \rangle$ for the same conditions.

0.1, the energy rises above the exponential fit applicable to the average capacity above 0.2. (b) The fraction of links with zero currents increases with the average capacity, and is rather insensitive to the connectivity. Remarkably, except for very small average capacities, the function $\mathrm{erf}(\langle \Lambda \rangle/\sqrt{2})$ has a very good fit with the data. Indeed, in the limit of large $\langle \Lambda \rangle$, this function approaches the fraction of links with both vertices unsaturated, that is, $[\int_0^\infty d\Lambda \rho(\Lambda)]^2$. (c) The fraction of unsaturated nodes increases with the average capacity, and is rather insensitive to the connectivity. In the limit of large average capacities, it approaches the upper bound of $\int_0^\infty d\Lambda \rho(\Lambda)$, which is the probability that the capacity of a node is non-negative. (d) The convergence time of both the Bethe recursion equations and Eq. (3) follows a power-law dependence on the average capacity when the average capacity is 0.2 or above; the exponent is ranging from -1 for $c=3$ to -0.8 for $c=5$ for Eq. (3), and being about -0.5 for $c=3, 4, 5$ for Eq. (1). When the average capacity decreases further, the convergence time deviates above the power laws.

Summary: We have studied a prototype problem of resource allocation on sparsely connected networks using the replica method and the Bethe approximation. The resultant recursion relation leads to a message passing algorithm for optimising the average energy, which significantly reduces the computational complexity of the global optimisation task and is suitable for online distributive control. The suggested 2-parameter approximation produces results with excellent agreement with the original recursion relation. We have considered the simple but illustrative example of a quadratic cost function, where both Bethe recursion equations and message passing algorithm show remarkable agreement with the exact result. The suggested simple message passing algorithm can be generalised to more realistic cases of nonlinear cost functions and additional constraints on the capacities of nodes and links. This constitutes a rich area for further investigations with many potential applications.

Acknowledgements This work is partially supported by research grants HKUST6062/02P and DAG04/05.SC25 of the Research Grant Council of Hong Kong, and by EVERGROW, IP No. 1935, IST Priority, EU FP6.

References

[1] L. Peterson, B.S. Davie, *Computer Networks: A Systems Approach*, Academic Press, San Diego CA, (2000)
[2] Y.C. Ho, L. Servi and R. Suri, *Large Scale Systems* **1** 51 (1980)
[3] S. Shenker, D. Clark, D. Estrin and S. Herzog, *ACM Computer Communications Review* **26** 19 (1996)
[4] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing*, OUP, Oxford, UK (2001)
[5] K.Y.M. Wong, D. Saad and Z. Gao, submitted (2005)
[6] M. Opper and D. Saad *Advanced Mean Field Methods*, MIT press (2001)
[7] D.J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, CUP UK(2003)

Poster Session

Creativity Patterns in Art Perception

Adam R., Goldenberg J., Adi-Japh E, Mazursky D., Solomon S.

Adam R., Cognitive Science Program, Hebrew University of Jerusalem.
Goldenberg J. and Mazursky D., School of Business Administration, Hebrew university of Jerusalem.
Adi-Japha E., School of Education, Bar-Ilan University.
Solomon S., the Racah Institute for Physics, Hebrew university of Jerusalem.
Correspondences should be send to: ruthinka@gmail.com.

Goldenberg J., Mazursky D. and Solomon S. searched for regularities in creative ideas and creative thinking. They studied those subjects mainly in the field of marketing: advertisements and new products innovations. They found creativity templates, meaning well-defined, objectively verifiable, and generalizable abstract patterns. From those templates, creativity emerges. The templates describe the relations between the components or the attributes of a product. Creativity templates appear in most creative advertisements and new products. For instance, 89% of the award-winning advertisements contained templates. When templates guide the innovation process, the ideas created are, in all probability, creative.

Our aim was to check whether creativity templates exist and can achieve similar results in other fields. We chose the most creative and subjective field: art, and to be more specific, creativity perception in photography. What makes a photograph creative?

In this study we used 23 photographs that were regarded on the web as "the best pictures of 2003". Fifteen subjects from various countries chose the five most creative photographs from the set of photographs, and explained why, in their opinion, those pictures were creative. We were interested only in the components referred to by viewers when judging art, and not in the whole photograph or all of its objects. Thus,we searched for patterns in subjects' explanations.

In total, we had 72 explanations. Only 35 referred to the elements in the photographs. Only these 35 explanations could be translated to the formal way which is essential for detecting creativity templates. The other explanations were unclear or described the subjects' feelings while seeing the photographs. From the 35 explanations, 43% contained patterns. We found three types of patterns: Diagonal Link, Properties Tension, and Unification. We will describe briefly each pattern, and illustrate it.

Diagonal Link (DL) pattern is a transitive connection: a connection between two components generated by the connection of every one of them to intermediate component. Twenty percent of the 35 explanations contained DL. It is embedded in one frequently seen creativity template known as the Replacement Template. We will demonstrate the DL by a photograph showing a ship sailing in icy water at sunrise. One of the subjects who chose this photograph as creative explained his choice in the following sentence: "The back light is illuminating the fog in a very intense manner so that it almost appears as fire". In this explanation the back light affects the fog, the back light is red, and the transitive result is a 'fire-like' red fog.

Properties Tension (PT) appears when the main components in a photograph have contrasting properties. This pattern appeared in twenty percent of the explanations. In one of the photographs one could see the head of a black cat with yellow eyes in a golden wheat field. One subject's explanation was: "...the eyes of the cat look bad and the field of grain looks kind". The main components in this photograph (the cat's eyes and the field) had the same color, but contrasting properties (bad versus kind).

Unification pattern appears when two unrelated objects joined to form one hybrid object, which looks like the first object but functions as the second. It was found in 11.4% of the explanations. One photograph captured a child in mid air jumping onto a pile of dead leaves. One of its Unification explanations was: "A child jumps into a pile of dead leaves as if it were a swimming pool". A child jumped onto an object that looks like a pile of dead leaves but functions as a pool.

Three fixed formulations of the elements in a photograph were found to cause creativity perception. Those are primary results and our patterns need to be further validated to be called creativity templates. However, those results show that art perception is not as arbitrary as might be imagined.

Agent Based Modeling of Consumer Behavior

Iqbal Adjali, Ben Dias, Robert Hurling
iqbal.adjali@unilever.com

Understanding the drivers for consumer behavior and the underlying mechanisms of market dynamics is of key importance to marketing practitioners, market strategists and competition and regulatory bodies. Furthermore the cross-disciplinary nature of this area has attracted the interest of researchers in diverse fields such as economists, social scientists, psychologists and computational scientists. Marketers usually ask questions like 8220;What are the chances for success when a new product is launched (e.g. brand extension)?8221; or 8220;How does word-of-mouth affect sales and the long-term prospects of a brand?8221; To answer these questions, traditional marketing models, based on equilibrium statistics and macro variables such as market share and price elasticity, have usually been employed with limited success. As a powerful computational method, agent based modeling has the potential to address many of the shortcomings of traditional techniques. Specifically, ABM (i) allows the dynamic nature of markets to be modeled, (ii) treats consumers as individuals therefore retaining the richness of information at the micro level and (iii) allows consumer interactions and social networks to be explicitly modeled. Here we describe the architecture for an agent based consumer behavior model drawing from the marketing and the behavioral sciences literature. We present some simulation results exploring and comparing different consumer behavior models using real, individual-based market data.

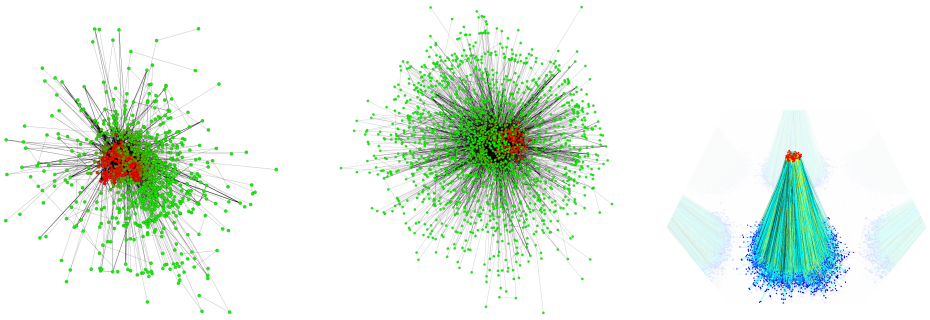
Analytic Visualizations and their Applications for the Autonomous System Graph*

Vinay Aggarwal†, Anja Feldmann†, Marco Gaertler‡, Robert Goerke‡, Yuval Shavitt§, Eran Shir§, Dorothea Wagner‡, Arne Wichmann†

Abstract

The graph of the Autonomous Systems, i. e., collection of computer devices under the same administrative authority that establishes global connection in the Internet, is an instance of a complex system. Out of the vast range of issues that have been addressed in the context of this network, we focus on revealing structural information via visualization.

In the following, we present several examples: excerpts from the temporal evolution of the AS graph between 2001 and now, overlay networks, i.e., comparing Gnutella communication and AS peering relations, and the comparison of different sources for AS path.



(a) Gnutella communications (b) Random communications (c) Multiple data sources

We used the technique of [2] as the basis of our analytic visualizations. In Figure (a) and (b), the induced communication network of Gnutella and of a random overlay network, respectively, are shown. Although, both the Gnutella network and the random network (uniformly drawn from the IP space) are non-correlated with the underlying AS graph [1], the two networks significantly differ in certain characteristics. This fact is apparent from the visualizations.

The two data sets obtained from the DIMES project and Oregon Routeview mapping the AS graph share more than 40% of their edges. However, combining both sets yields a more representative picture. The hierarchical layout of this combined set (Figure (c)) is used to study the differences between the two sets. A surprising result is the fact, that both sets equally influence the global shape. These examples show that the technique of [2] has promising potential to analyse various different aspects of the AS graph.

References

[1] Vinay Aggarwal, Stefan Bender, Anja Feldmann, and Arne Wichmann. Methodology for estimating network distances of gnutella neighbors. In *GI Jahrestagung*, volume 2, 2004.

[2] Michael Baur, Ulrik Brandes, Marco Gaertler, and Dorothea Wagner. Drawing the as graph in 2.5 dimensions. In *Proceedings of the 12th International Symposium on Graph Drawing (GD'04)*, volume 3383 of *Lecture Notes in Computer Science*, pages 43–48, 2005.

*Work partially supported by European Commission - Fet Open project COSIN - COevolution and Self-organisation In dynamical Networks - IST-2001-33555, by European Commission - Fet Open project DELIS - Dynamically Evolving Large Scale Information Systems - Contract no 001907 and 6th FP EVERGROW projects

†TU München, Institut für Informatik, München, Germany. http://www.net.informatik.tu-muenchen.de/

‡Universität Karlsruhe (TH), Faculty of Informatics, Karlsruhe, Germany. http://i11www.iti.uka.de/

§EE-Systems dept., Tel Aviv University, Israel. http://www.netdimes.org/

Are epidemics on scale-free networks predictable?

Fabián Alvarez [1,2] Pascal Crépey [1,2] Marc Barthelémy [3]

[1] INSERM, U707, F-75012 Paris, France
[2] Université Pierre et Marie Curie, F-75005 Paris, France
[3] CEA, 91680 Bruyères-Le-Châtel, France

The scale-free topology has been widely described in many artificial and natural networks. It is characterized by a broad distribution of individual connectivities which can take values over several orders of magnitude.

In this work, we study how these large connectivity fluctuations affect the sensitivity to the noise of an epidemic scenario and consequently its predictability. We investigate this problem by inspecting thoroughly the time evolution of the number of infected at the beginning of the outbreak. We study numerically the variability of epidemic outbreaks spreading on scale-free networks and compare these results with simulations on random homogeneous networks, in which connectivities are normally distributed around their average value <k>. Propagation on random networks corresponds to the usual assumption of homogeneous mixing in classical epidemiology and provides a useful reference model to assess the effect of degree fluctuations.

We consider the usual compartmentalization of individuals into 3 categories: susceptible (S), infected (I) and immunized (R), and the time evolution of the system is described by the usual epidemiological schemes SI, SIS or SIR. In this work, we will focus on the temporal evolution of the prevalence (number of infected in the population). For a fixed set of the parameters of the model (infection scheme, infection rate, number of index cases), we simulate numerically a large number of outbreaks (up to 10^5), and we analyze different averages: average over a number e of outbreaks on the same network, average of a single outbreak on r different networks, and average of e outbreaks on r networks. We study quantities such as the coefficient of variation of the prevalence (CV_i), the time to peak of CV_i, and the distribution of the prevalence doubling time. In the light of the numerical results for these quantities, we discuss the influence of the different sources of noise on the variability of the epidemic scenario.

In particular, our results show that on scale-free networks, a peak of the CV_i is reached in the very first stages of the epidemic outbreaks, before its exponential growth. At this peak, the CV_i is of order 2-3 in contrast with a CV_i of order 1 (1.0 to 1.5) obtained for random homogeneous networks. These value are calculated when "dying" outbreaks are discarded (SIS and SIR models), and then may be larger if all outbreaks are considered. We also verified that the results are robust for large ranges of the models' parameters.

Our results suggest that in the situation of an emerging disease spreading on a scale-free network, the epidemic scenario is very sensitive to the topology of the network. Furthermore, these large fluctuations make the prediction of the disease time evolution -based solely on the topology of the network and on the first infected cases- almost impossible.

Are epidemics on scale-free networks predictable?

Fabián Alvarez [1,2] Pascal Crépey [1,2] Marc Barthelémy [3]

[1] INSERM, U707, F-75012 Paris, France
[2] Université Pierre et Marie Curie, F-75005 Paris, France
[3] CEA, 91680 Bruyères-Le-Châtel, France

The scale-free topology has been widely described in many artificial and natural networks. It is characterized by a broad distribution of individual connectivities which can take values over several orders of magnitude.

In this work, we study how these large connectivity fluctuations affect the sensitivity to the noise of an epidemic scenario and consequently its predictability. We investigate this problem by inspecting thoroughly the time evolution of the number of infected at the beginning of the outbreak. We study numerically the variability of epidemic outbreaks spreading on scale-free networks and compare these results with simulations on random homogeneous networks, in which connectivities are normally distributed around their average value <k>. Propagation on random networks corresponds to the usual assumption of homogeneous mixing in classical epidemiology and provides a useful reference model to assess the effect of degree fluctuations.

We consider the usual compartmentalization of individuals into 3 categories: susceptible (S), infected (I) and immunized (R), and the time evolution of the system is described by the usual epidemiological schemes SI, SIS or SIR. In this work, we will focus on the temporal evolution of the prevalence (number of infected in the population). For a fixed set of the parameters of the model (infection scheme, infection rate, number of index cases), we simulate numerically a large number of outbreaks (up to 10^5), and we analyze different averages: average over a number e of outbreaks on the same network, average of a single outbreak on r different networks, and average of e outbreaks on r networks. We study quantities such as the coefficient of variation of the prevalence (CV_i), the time to peak of CV_i, and the distribution of the prevalence doubling time. In the light of the numerical results for these quantities, we discuss the influence of the different sources of noise on the variability of the epidemic scenario.

In particular, our results show that on scale-free networks, a peak of the CV_i is reached in the very first stages of the epidemic outbreaks, before its exponential growth. At this peak, the CV_i is of order 2-3 in contrast with a CV_i of order 1 (1.0 to 1.5) obtained for random homogeneous networks. These value are calculated when "dying" outbreaks are discarded (SIS and SIR models), and then may be larger if all outbreaks are considered. We also verified that the results are robust for large ranges of the models' parameters.

Our results suggest that in the situation of an emerging disease spreading on a scale-free network, the epidemic scenario is very sensitive to the topology of the network. Furthermore, these large fluctuations make the prediction of the disease time evolution -based solely on the topology of the network and on the first infected cases- almost impossible.

Environmental uncertainty and language complexity

P Andras, J Lazarus, J G Roberts
j.lazarus@ncl.ac.uk

We examine how the complexity of language is influenced by the level of uncertainty (= risk) in the environment (Andras, Roberts and Lazarus 2005), when language is used to communicate cooperative intentions between interacting agents. Since cooperation increases with risk (measured as both environmental adversity and uncertainty) in a number of biological contexts (Andras and Lazarus 2005), and cooperation requires communication between interacting agents, the question arises of whether language complexity might vary adaptively as a function of risk. An example is the evolution of human language in the small cooperative groups of early hunter-gatherers.

We model a world of simple agents that own resources that are used for survival and to produce new resources, environmental risk being expressed as the variance in the reproduction of these resources (Andras, Roberts and Lazarus 2003). Agents produce offspring and the number of offspring depends on the amount of resources owned by the agent. New agents inherit the language of their parent with some random mutations.

Agents communicate with each other about their intention to cooperate by sharing resources. The language contains the following semantic elements: no interest in communicating; start of communication; intend to communicate further; want to engage in sharing; lost interest; share action; not share action. Agents remember previous encounters and if they meet again an agent with whom they have previously cooperated then communication symbols likely to lead to cooperation increase in probability in the current encounter (and vice versa).

As risk increased in the simulations cooperation increased, and cheating and non-cooperation decreased, as found in nature (Andras and Lazarus 2005). Overall, there was no clear relationship between language complexity (i.e. the length of communication sequences between agents) and environmental risk, although communication sequences were shortest in the most risky environments.

The lack of a clear relationship between language complexity and environmental risk may have been because the model language was too simple, varying between only 4 and 6 elements at the outset. For a richer language which we plan to investigate - we predict that communications will be shorter as risk increases (as the data for our highest risk level suggested) since: (1) there was a positive correlation between cooperation level and risk, and (2) there was a negative correlation between cooperation level and language complexity. Those who cooperated had shorter communication strings than those who cheated or failed to cooperate. This is because if a cooperator meets an agent for whom it has a memory biased towards cooperation then it has a higher probability of producing positive communication symbols (those encouraging cooperation) and therefore moves with fewer communication steps into an interaction that is likely to be cooperative. Thus cooperative agents, by positive feedback, build an increasingly cooperative relationship with each other. Cooperation thus saves on communication effort.

The results have implications for the design and use of communication systems under conditions of uncertainty, and for the role of environmental uncertainty for example, in foraging - in shaping the evolution of human communication patterns.

Andras, P and Lazarus, J (2005) Cooperation, Risk and the Evolution of Teamwork. In: Teamwork: Multi-Disciplinary Perspectives, edited by N Gold. Basingstoke: Palgrave. Macmillan. Pp.56-77.

Andras, P., Roberts, G., and Lazarus, J. (2003) Communication Complexity, Environmental Risk, and Cooperation. In Alonso, E., Kudenko, D. and Kazakov, D. (eds.)

Adaptive Agents and Multi-Agent Systems, Springer-Verlag, Berlin, pp.49-65.

Andras, P, Roberts, G, and Lazarus, J (2005) Environmental risk. In Proceedings of the Joint Symposium on Socially Inspired Computing - AISB 2005, pp.88-94.

Keywords

multi-agent systems, communication, language, cooperation, environmental uncertainty

Amino acid evolution: an alternative hypothesis

Peter Andras[a], Alina Andras[b], Csaba D. Andras[c,d]
[a]School of Computing Science,
University of Newcastle, Newcastle upon Tyne, NE1 7RU, UK
[b]Institute for Ageing and Health,
University of Newcastle, Newcastle upon Tyne, NE4 6BE, UK
[c]Department of Chemical Engineering,
Budapest University of Technology and Economics, Budapest, Hungary
[d]Department of Technology and Natural Sciences,
Sapientia – Hungarian University of Transylvania, Miercurea Ciuc, Romania

The understanding of the evolution of amino acid usage in proteins of living organisms is a fundamental issue in the context of theories about the origins of life. The commonly accepted view is that life started with a few amino acids and newer ones were added to the amino acid library of organisms during evolution. Here we propose an alternative hypothesis, suggesting that life might have started with an initial expansion of amino acids followed by gradual decrease of the number of amino acids used in proteins.

It has been recently reported that the replacement frequencies of amino acids are asymmetrical [1], e.g., during the independent evolution of human and chimpanzee genomes the replacement of leucin codons by phenylalanine codons happened almost twice as many times than the reverse replacement. This phenomenon can be interpreted according to the hypothesis that early amino acids (i.e., amino acids produced in origins-of-life experiments [2] or found in meteorites [3]) are replaced by newer amino acids, which emerged as living systems evolved [4]. The data presented by Jordan et al. [1] show that indeed the amino acids which loose in terms of frequency are those which can be considered early amino acids, and those which gain are the other amino acids.

We propose an alternative hypothesis for the interpretation of asymmetric amino acid replacement frequencies. We hypothesise that in the early stages of life there were many amino acids composing proteins of early organisms and during evolution, due to optimisation pressures, the set of amino acids used in proteins was reduced to the currently known 22 amino acids [1,5], which participate in protein formation in extant organisms. In our view the asymmetrical replacement frequencies reflect the effects of this optimisation process.

Plants are able to synthesize more than 200 amino acids that are not directly incorporated into proteins during the translation of mRNA molecules. The existence of natural tRNA molecules with anticodons of length 4 and possibly more (e.g., stop-supressor tRNAs, +1/-1 frameshift tRNAs) indicate that the coding potential of tRNA-s is much larger than the number of genetically encoded amino acids. The discovery in the 70s of the rare protein forming amino acid seleno-cysteine [5], and the recent discovery of the genetically encoded pyrrolysine in archaea [5] show that it is possible that some non-early amino acids might be on the track of disappearance from the list of genetically encoded amino acids.

The cornerstone of our hypothesis is the assumption that simplification of the amino acid lexicon may lead to more efficient living systems incorporating more complex and more functional proteins. This assumption is based on the finding that codon usage is more restricted in more complex organisms [6]. Our assumption is also supported by recent findings that the number of genes of complex organisms is much smaller than their expected number, and that the complex behaviour of these

organisms is achieved by complex regulatory combination of a relatively small number of genomic components [7,8] (i.e., genes and regulatory segments of the DNA).

According to our hypothesis life started with an initial expansion of amino acids and early organisms built their proteins using a wide range of amino acids encoded by codons of length k, with $k \geq 3$. During evolution the optimisation pressures selected organisms using shorter codons and fewer amino acids resulting in the currently dominant set of 20 amino acids and length 3 codons. Remnants of early living systems are the rarely used genetically encoded amino acids (seleno-cysteine and pyrrolysine), tRNA molecules with anticodons of length 4 or more, and viruses that produce frameshift and stop-supressor tRNAs [9]. Our hypothesis also provides an explanation for the observed code simplification of mitochondria, implying that these highly specialised organelles evolved further the simplification of the tRNA lexicon resulting in their reduced set of tRNAs and increased number of unused codons [9].

Our interpretation suggests that asymmetric replacement frequencies do not indicate the replacement of early amino acids by newer ones, but instead indicate the optimisation of the genetic code and the set of encoded amino acids towards a functionally more complex and more efficient form of life based on fewer amino acids and simpler, less ambiguous encoding of these amino acids.

1. Jordan, I.K., Kondrashov, F.A., Adzhubei, I.A., Wolf, Y.I., Koonin, E.V., Kondrashov, A.S. & Sunayev, S. *Nature*, **433**, 633-638 (2005).
2. Miller, S.L. & Orgel, L.E. *The Origins of Life on the Earth*. Englewood Cliffs, NJ, Prentice Hall (1974).
3. Pizzarello, S. *Origins of Life and Evolution of the Biosphere*, **34**, 25-34 (2004).
4. Brooks, D.J., Fresco, R.F., Lesk, A.M., & Singh, M. *Molecular Biology and Evolution*, **19**, 1645-1655 (2002).
5. Cobucci-Ponzano, B., Rossi, M., & Moracci, M. *Molecular Microbiology*, **55**, 339-348 (2005).
6. Duret, L. & Mouchiroud, D. *Proceedings of the National Academy of Sciences*, **96**, 4482-4487 (1999).
7. Andras, P. & Andras, C.D. *Medical Hypotheses*, **64**, 678-688 (2005).
8. Mattick, J.S. & Gagen, M.J. *Molecular Biology and Evolution*, **18**, 1611-1630 (2001).
9. Söll, D. & RajBhandary U.L. (eds.) *tRNA Structure, Biosynthesis and Function*. (ASM Press, Washington, DC, 1995).

Towards adaptive self-aware software

Peter Andras (1), Bruce Charlton (2)
(1) School of Computing Science
(2) School of Biology / Psychology
University of Newcastle
Newcastle upon Tyne, NE1 7RU, UK

Abstract

Self-aware adaptive software is a saint Grail of computer science. Recent advances are very promising, and software systems have several features required for being self-aware and adaptive (e.g., introspection and self-modification). However no truly self-aware adaptive software system exists currently. We analyse biological and social systems as abstract communication systems, building an analogy with software systems (i.e., systems of communicating objects). We highlight three critical features of natural systems: the asymmetry of true/false, the growth by correctness checks, and error-induced adaptation based on self-monitoring. Based on the established analogy we propose to build software systems with non-terminating proofs of correctness, to achieve self-aware adaptive software.

Since the beginning of computer science researchers aimed to build systems adaptive self-aware software systems that are comparable with animals in terms of self-awareness and adaptability [1, 2]. Recent progress in this direction has been significant. In particular, object oriented thinking led to systems with features required for adaptive self-awareness. Such features include the ability for introspection and self-modification (i.e., reflection [3]), the ability for dynamic integration, re-use of components and use of integration patterns (i.e., component-based programming [4] and use of design patterns [5]), and the ability for improved self-monitoring and action selection based on self-monitoring (i.e., aspect oriented programming [6]).

However, self-aware adaptive software systems comparable to natural systems are not yet available. In our view to make such systems possible we need further paradigmatic changes in software theory and design. We suggest that the needed change is to aim to build programs and software systems with non-terminating proofs of correctness.

Biological systems, like cells or animals, are composed of a very large number of constituents. These components interact with each other in well-defined ways forming patterns of interactions, which define the biological system [7]. In case of social systems [8] humans interact with other humans forming patterns of human communications obeying to a variety of regularities. We can see these systems as a self-aware, adaptive, self-reproducing and expanding abstract communication system made of patterns of communications obeying a set of rules, which define the system [9]. This leads to an analogy with object oriented software systems, made of a multitude of objects, which constitute the system by producing communication patterns between themselves, according to the rules of the software system [9].

Analysing natural systems in terms of abstract communication systems reveals critical features of them, which are closely linked with their adaptive self-aware nature. These features are: (1) the asymmetry of

true/false or the Popper principle, i.e., communications can be proven to be false (incorrect) by not leading to further continuation of the system, while the opposite never can be proven; (2) growth by correctness checks, i.e., the system growth by generating new communications that check the validity of earlier communications; (3) self-monitoring and error-induced adaptation, i.e., the system monitors itself by correctness checking communications and in case of errors (large scale halting of continuation of communications) the blueprint of the system is modified such that errors are avoided in the future. In our view these features are needed in truly self-aware adaptive software systems, and by analogy with natural systems, such features may be achievable if software systems aimed to have non-terminating proofs.

Acknowledgement: The authors thank Matthew Pocock for very useful discussions.

References

1.. Landauer C, Bellman, KL: Self-modelling systems. In: LNCS 2614 2001; pp 238-256.

2. Ungar, D, Smith, RB: SELF: The power of simplicity. LISP and Symbolic Computation 1991; 4: 187-205.

3. Bracha, G, Ungar D: Mirrors: Design principles for meta-level facilities of object-oriented programming languages. In: Proceedings of OOPSLA'2004; pp 331-344.

4. Jarzabek, S, Knauber, P: Synergy between component-based and generative approaches. ACM SIGSOFT Software Engineering Notes 1999; 24: 429-445.

5. Gamma, E, Helm, R, Johnson, R, and Vlissides, E: Design Patterns: Elements of Reusable Object-Oriented Software. Reading, MA, Addison-Wesley, 1995.

6. Katara, M, Katz, S: Architectural views of aspects. In: Proceedings of AOSD 2003; pp 1-10.

7. Miller, JG: Living Systems. McGraw-Hill, 1978.

8. Luhmann, N: Social Systems. Palo Alto, Stanford University Press, 1996

9. Andras, P, Charlton, BG: Self-aware software – Will it become a reality ? In: Babaoglu, O et al (eds) SELF-STAR, LNCS 3460 2005; pp 229-259.

How complexity theory may explain influence of music

Svetlana Apjonova, Igor Yevin
yevin@online.ru

Recent researches revealed that music reduces the degree of chaos in brain waves. A.Loskutov, A.Hubler, and others carried out a series of studies concerning control of deterministic chaotic systems. It turned out, that carefully chosen tiny perturbation could stabilize any of unstable periodic orbits making up a strange attractor. V.Bondarenko shown in computer experiments a possibility to control a chaotic behavior in neural network by external periodic pulsed force or sinusoidal force. Low-dimensional outputs are observed when the frequency of the external force is close to delta-, theta-, alpha-, and beta frequencies. We suggest that music acts on the brain near these eigenfrequencies of self-excited oscillations in the neural network to suppress chaos. We explain the structure of music tonalities using concept of attractor network model. Three stable steps of tonality: tonic, median, and dominant are keynotes or attractors of neural network model. There are many reasons to believe that among four types of instinctive behaviors and four frequencies of the brain exists one-to one correspondence: delta rhythm – food behavior, theta – fear, alpha – sex, and beta – aggression. This hypothesis does not contradict with available empirical date. Delta rhythm is the main rhythm for all newborns both animals and humans (infants up to one years of age). Undoubtedly, the food instinct is the main instinctive behavior for all newborns. Therefore, delta rhythm may be connected with the pleasure taken from food. Theta rhythm one usually connects with waking behavior in different species which are pivotal to the animals survival. For instance, it is predation for cats, exploration for rats, and apprehension for rabbits. Undoubtedly, the basic for all such kinds of pivotal behaviors is food behavior. There is never-ending conflict and tension between food searching and avoiding predators which is the strongest natural stressor that wild animals experience. Theta rhythm is also dominant rhythm for human infants by 13 years old. We may suggest that all such kinds pivotal behaviors are accompanied by fear because namely during food searching and during childhood animals and humans are the least defended and protected from different threats outside. The source of theta oscillations is the hippocampus which has strong reciprocal connections with amygdala - the centre of fear emotion. Therefore the fear is the next in importance and the next in order instinctive behavior and it may be connected with alarm, threat, and fear feelings. Completion of the alpha rhythm occurs at the puberty period. Children before 12-13 years old have rather weak alpha waves. Adult humans have maximal alpha rhythm when he or she is relaxing with close eyes. But relaxing state with closed eyes usually accompanies sexual enjoyment. Beta rhythm is associated with focused attention toward external stimuli, alert mental activity with increasing muscle tension and raising blood pressure. We suppose that beta rhythm is connected with aggression. This assumption is supported by A.Blood and R.Zatorre experiment that music involves the brain regions such as ventral stratum, midbrain, amygdala that are connected with such biologically relevant stimuli as food, sex and others.

How complexity theory may explain influence of music

Svetlana Apjonova, Igor Yevin
yevin@online.ru

Recent researches revealed that music reduces the degree of chaos in brain waves. A.Loskutov, A.Hubler, and others carried out a series of studies concerning control of deterministic chaotic systems. It turned out, that carefully chosen tiny perturbation could stabilize any of unstable periodic orbits making up a strange attractor. V.Bondarenko shown in computer experiments a possibility to control a chaotic behavior in neural network by external periodic pulsed force or sinusoidal force. Low-dimensional outputs are observed when the frequency of the external force is close to delta-, theta-, alpha-, and beta frequencies. We suggest that music acts on the brain near these eigenfrequencies of self-excited oscillations in the neural network to suppress chaos. We explain the structure of music tonalities using concept of attractor network model. Three stable steps of tonality: tonic, median, and dominant are keynotes or attractors of neural network model. There are many reasons to believe that among four types of instinctive behaviors and four frequencies of the brain exists one-to one correspondence: delta rhythm – food behavior, theta – fear, alpha – sex, and beta – aggression. This hypothesis does not contradict with available empirical date. Delta rhythm is the main rhythm for all newborns both animals and humans (infants up to one years of age). Undoubtedly, the food instinct is the main instinctive behavior for all newborns. Therefore, delta rhythm may be connected with the pleasure taken from food. Theta rhythm one usually connects with waking behavior in different species which are pivotal to the animals survival. For instance, it is predation for cats, exploration for rats, and apprehension for rabbits. Undoubtedly, the basic for all such kinds of pivotal behaviors is food behavior. There is never-ending conflict and tension between food searching and avoiding predators which is the strongest natural stressor that wild animals experience. Theta rhythm is also dominant rhythm for human infants by 13 years old. We may suggest that all such kinds pivotal behaviors are accompanied by fear because namely during food searching and during childhood animals and humans are the least defended and protected from different threats outside. The source of theta oscillations is the hippocampus which has strong reciprocal connections with amygdala - the centre of fear emotion. Therefore the fear is the next in importance and the next in order instinctive behavior and it may be connected with alarm, threat, and fear feelings. Completion of the alpha rhythm occurs at the puberty period. Children before 12-13 years old have rather weak alpha waves. Adult humans have maximal alpha rhythm when he or she is relaxing with close eyes. But relaxing state with closed eyes usually accompanies sexual enjoyment. Beta rhythm is associated with focused attention toward external stimuli, alert mental activity with increasing muscle tension and raising blood pressure. We suppose that beta rhythm is connected with aggression. This assumption is supported by A.Blood and R.Zatorre experiment that music involves the brain regions such as ventral stratum, midbrain, amygdala that are connected with such biologically relevant stimuli as food, sex and others.

Bayesian Reconstruction of Particle Size Dynamic Distributions of Particulate Polydisperse Systems from in vitro Drug Dissolution Data

Ana Barat, Heather Ruskin, Martin Crane
abarat@computing.dcu.ie

The direct deterministic modelling of in vitro drug dissolution is not a trivial problem, because of the complexity of the various physical phenomena involved in the dissolution process, which should be accounted for when setting models. The most typical experimental data of the in vitro dissolution are obtained in practice by dissolution testing in the USP (United States Pharmacopoeia) Apparatus II. The hydrodynamic conditions in the USP Apparatus are time-dependent and very heterogeneous in space, which result in a quite undesirable feature: high variability of the dissolution profiles. Some kinds of drug delivery systems like disintegrating or particulate polydisperse systems have an extremely complex behaviour in the USP because of the variable distribution of the solid particle sizes combined with heterogeneous flow conditions in the apparatus, thus many aspects of their dissolution are subject to uncertainty. In this work, we create a range of models based on Inverse Monte Carlo simulations in a Bayesian context, that are capable to extract knowledge from experimental data (time series of dissolved quantities). At different time-steps of the dissolution, we consider the parameters of the particle size distribution as random variables. A deterministic function based on classical equations for drug dissolution (or based on less classical particle-tracking techniques in the apparatus) is used to assess the relation between the particle size distribution and then the amount of dissolved drug in the solution. The inaccuracy in the dynamics of the model is considered stochastic and represented by some noise factors. In reporting simulated results to the experimental dissolution results, we re-update the unknown size distributions in order to obtain their reconstruction at the end of the simulation, within a previously specified error range. This method is at experimental stage, but can prove to be very useful in the field of drug dissolution: the time series of reconstructed particle size distribution can be used for solving the direct dissolution problem for different sets of parameters in optimisation research.

A Novel Medical Diagnosis System

Barna Laszlo Iantovics

The purpose of the study consists in the development of an open, large-scale heterogeneous medical diagnosis system capable of solving a large variety of difficult medical diagnosis problems. We propose the endowment of the expert systems specialized in medical diagnosis with the agents' capabilities. We name these agents *expert system agents* [1]. The expert system agents can solve cooperatively difficult diagnosis problems in a more flexible way than the expert systems. In this paper, we propose a cooperative heterogeneous medical diagnosis system. Cooperating diagnosis systems seems to be well suited for medical diagnosis in many medical domains [2]. The proposed system is composed from a set $MDS = \{M_1, ... M_k\}$ of *agents* (expert system agents, doctors) specialized in medical diagnosis with different specializations in medical domains. The proposed system can solve problems the solving of which require knowledge from more medical domains. Each agent member of the system can overtake problems for solving.

In the following, we describe briefly how an overtaken problem P is solved. An agent M ($M \in MDS$) receives the problem. If M is capable it processes the problem obtaining the result R. If R doesn't represent the solution of the problem, then M will find an agent capable to processes the result R (R represents a new problem). The problem P results transmission from an agent to another agent continues until the problem is solved. If an agent can't processes a received problem, then it must transmits the problem in the received form to a capable agent to processes it. The finding of an agent capable to processes a problem is based on the problem allocation described in the paper [3]. The parameters of a problem's announcement to the agents, and the parameters of the received responses to the problem announcement are also analyzed. An agent N transmits a problem announcement to more agents. From the agents who answer to the announcement, N must choose the best-fitted agent to processes the problem. In the choosing of the best-fitted agent, N analyses the parameters of the received responses. In the response to a problem announcement each agent indicates its capability and capacity to process the problem. The capability of an agent consists in the problems that can be solved by the agent [1]. For example, an expert system agent can be specialized in diagnosis in more medical domains. The capacity of an agent consists in the amount of problems that can be solved by the agent in deadline [1].

As an example, we consider the problem P (a cardiology and urology related illness). An expert system agent G specialized in general medicine, a doctor C specialized in cardiology and a doctor U specialized in urology can solve the problem. The problem P solving can be described as follows: $G(P) \Rightarrow C(R) \Rightarrow U(V) \Rightarrow S$. Processing P by G, the result R is obtained. Processing R by C the result V is obtained. Processing V by U the solution S is obtained. R represents symptoms of the patient's illnesses and the observations elaborated by G related to the patient's illnesses. V represents the cardiology related illness identified by C, the patient's illnesses symptoms and the observations elaborated by G. The result S represents the identified two illnesses (the problem P solution). The urology related illness is identified by U.

The main advantage of the proposed medical diagnosis problem solving is the flexible and precise solving of a large variety of difficult medical diagnosis problems, the solving of which require knowledge from more medical domains. The knowledge necessary to the diagnosis problems solving are not specified in advance, the diagnosis system members must discover cooperatively the problems solving. The agents' (human and artificial) capabilities and capacities are efficiently used. If an agent is not capable of processing a problem (he doesn't have the necessary capability and/or capacity) then it will transmit the problem to an agent capable to processes it. There are no uncertainties in a problem distribution from an agent to another agent. The parameters of the responses to the problem announcement contain knowledge that can be used in the decision elaboration.

References

1. Iantovics BL: Intelligent Agents, PhD dissertation, Babes-Bolyai University, Cluj-Napoca, 2004.
2. Vesnenko AI, Popov AA, Pronenko MI: Topo-typology of the structure of full-scaled clinical diagnoses in modern medical information systems and technologies. Cybernetics and Systems Analysis 2002; 38:6.
3. Iantovics BL: A New Task Allocation Protocol in Distributed Multiagent Systems, 4th International Conference On Education, Training and Information/Communication Technologies, Sovata, 2005, 5-10.

[1]

The Effects of Topology on the dynamics of Naming Games

Andrea Baronchelli‡, Luca Dall'Asta†,
Alain Barrat† and Vittorio Loreto‡§

May 8, 2005

Abstract

The study of language as the result of a complex adaptive system allows to understand how a population of agents can develop a coherent set of linguistic conventions on the basis of simple interactions and negotiating processes. Among the language games used to study the evolution of language, Naming Games (Steels, 1996) are conceived as simple models reproducing the dynamics by which a word or a minimal set of words are selected as a commonly shared vocabulary to identify one or more objects. We consider a minimal model of naming game in which all interactions among pairs of individuals, one selected as speaker and the other as hearer, are focused to communicate the name of a single object. At each step of dynamics two agents discuss on the name of the objects, updating their inventory of words on the basis of the result of the game. A two-agents game has a success if the hearer understands the word pronounced by the speaker, i.e. if they both possess it in their words inventories. The type of dynamics favors the spreading of a word having high rate of success, with a final state in which all individuals share a single common word to identify the object. We study the effects of topological constraints on the dynamics embedding the system in graphs, where the agents can interact only with their neighbors. The emergence of a common linguistic convention for the object's name seems to be solid with respect to different topological structures: regular (1D and 2D lattices), homogeneous random graphs and heterogeneous scale-free graphs. However, the temporal scale of the dynamical process as well as the evolution of spatial structures (groups of agents sharing a words) show a variety of interesting different behaviors, that are analyzed using both analytical and numerical techniques. In particular, the dynamics on low dimensional lattices is characterized by diffusive phenomena that are explained using a simple master equation's approach.

*† Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210 Université de Paris-Sud 91405 ORSAY Cedex - France
‡Dipartimento di Fisica, Universitá "La Sapienza", P.le A. Moro 2, Roma, Italy

The structure of large social networks

Dominik Batorski
db@uw.edu.pl

The last few years have seen a burst of interest in the properties of large networked systems such as the Internet, the WWW, and social and biological networks. The main goal of our research was to identify structural patterns of relations in a large social network. We examined the networks of acquaintanceship and communication between the users of Gadu-Gadu the most popular Instant Messenger in Poland. This kind of computer-mediated communication is used mainly to communicate with people known from off-line world. For that reason the research in question, although carried out on the Internet, was aimed at identifying the patterns of relations typical also for the social networks outside the Internet. At the moment Gadu-Gadu has 3.4 million users. The number of registered relations exceeds 75 millions.

Using the whole network data on relations from Gadu-Gadu buddy lists and communication between users we focus on main statistical properties of complex networks that have received attention recently. We report that social networks show the characteristics of small-world networks most pairs of vertices in network are connected by a short path through the network, and scale-free networks highly skewed degree distributions. We consider also degree correlations; it has been observed that the degrees of adjacent vertices in social networks are positively correlated.

In the last part of the paper we describe the possible consequences of measured structural properties of network for social processes that take place in social networks.

Managing as Designing : how designers can help managers in designing their organization as complex environments ?

Brigitte Borja de Mozota
bborjade@u-paris10.fr

Keywords: design ,complexity, design management ,design thinking , organizational design,strategy

Jim Hackett president CEO Steelcase Inc Organizations tend to drive for simplicity when attempting to solve complex,systemic problems. These practices are sweeping the rich opportunities of complexity under the carpet. Jim contends that successful organizations will embrace systems thinking,approaching their system of business as a design to be tested against the fitness of other businesses Source Institute of Design Strategy Conference May 2005

Summary : -organizations as complex systems -design as system thinking -rethinking design thinking with complexity theory -beyond the representations the importance of 'sense making' -design as organizational design

1. Organizations as complex systems Managers are facing complexity at three stages : 1.the complexity by itself :They see the real world as complex . Their perceived reality always remains incomplete and unfinished.The 'all' and the 'parts' are linked in inextricable and entangled interactions .The product is producing what organization is produces it in a spiral of cause and effect. Different logics co-exist in dialogical manner. Reality considered as complex is difficult to understand and to control .

2.the level of representation. Managers as observers who perceive the reality as complex build their representation of the phenomenon observed. But this is the major difficulty since how can they represent something that basically they dont understand? For this issue, they have to go back to the methods of representation and to understand their representation systems. A representation system is a set of elements including the past,the present and the future.

Our representations are not the image of the reality but a construction of our minds. In this universe of complexity, managers have to practice 'an ecology of mind' which means to adapt their way of thinking the environment and to consider themselves as a part of the complexity they have to control.

3.The level of the finality Since our representations are the source of our actions,it is urgent to understand and to control the paradigm of the system and to broaden its scope. This systemic modelisation is an intellectual tool , a mental construction , a method to construct a finality. It is the model that provides the finality and 'makes sense' to the whole system.

2. Design as system thinking

But managers are trained to think analytically in order to decide and control. Designers on the contrary are encouraged to adopt a holistic approach to any situation, to adopt a system thinking.(fig1)

Descartes System thinking Evidence is what makes a thing real. A representation is not true in itself; its pertinence comes from the intention of the person that formulates the representation Analysis :difficulties should be divided in as many parts necessary in order to understand them separately Global approach of interactions recommends to link together the elements of a system in order to be able to understand them in the globality of their interactions Hierarchical structure: you have to conduct your thinking by startign from the things simple to the more compses by orderign them together the concept of teleology that places the finality of the system is the essence of its dynamic.a system does not find its meaning in a preexisting structure but in a project that induces its behavior Number:

one has to number the whole in order to be sure not to forget something. Representation is always 'limited rationality'; it consists of a non exhaustive selection of aggregates that are thought pertinent to the situation one wants to shed light on Fig 1 : 'Discours de la méthode' versus systems thinking (adapted from Genelot 2001)

The process of constructive thought and action called 'Design thinking' belongs to the center of basic design education today because -design thinking is multidisciplinary and applicable to any subject, -design thinking integrates imagination and analytical thinking -design thinking emphasizes constructive thinking over factual retention -design thinking requires ongoing definition, representation and assessement: it is a continuous learning experience. -design thinking links information to experience and reasonable action . -design thinking encourages objective assessment (source Charles Burnette www.Idesignthinking.com 2005) Systems perception of design recognize its universal role as a core human activity. Design as a systemic activity is well summarized by Herbert Simon definition: 'Everyones designs who devises courses of action aimed at changing existing situations into preferred ones'.

This view typically integrates design into dynamic systems and in so doing recognizes the highly complex multi faceted nature of designing. It requires that hard systems methodologies of the reducionist perspective be augmented with soft system methodologies (Broadbent 2005 for a list of definitions)

Designers are creating systems in organizations whether brand identity systems or global systems of customer experience relationship. Design creates also value in a systemic way .Design creates value to the global value chain: substantial value as well as financial value . (Borja de Mozota 2003)

But this design thinking is targeted towards an end, a project. In summary , design thinking is sipmple system thinking.

3. Rethinking design with complexity theory Complexity science explores non linear dynamic systems which balance both order and chaos have emergent properties and are adaptive in nature (Sardar and Adams 2003).

But there is a dsitinction between simple systems and complex systems.Complex systems are exploring the principles of complexity such as richness, interconnections, iteration, emergence, holism and fluctuations.

For David Thomson (2005) ,the application of complexity theory in organizations is providing new insight into strategic organisatioanl management. Stacey (2000) believes that where a system shows traits of complexity strategic management takes on a new meaning.It becomes a domain of adaptive responses rather than control.

But much design is about successive design projects and design management literature is about control . However if you are truly holistic in your design thinking then you have to let go of control. And rethink design as a complex system

Many tools exist to help organizations pilot a course of action and communicate a plan however these tools are often prescriptive and do not account for the reality of the complex responsive processes at work.

A key skill of the designer is the ability to visualise in a form such that others can comprehend its meaning.With product, graphic, web design, design dsicipline project based - we can show people a physical model, a rough, a representation.

With strategy and strategic design , designers have to present a conceptual model. Conceptual mapping is one technique that holds promise for modelling design strategy in complex organisational systems. A new field of research opens for information designers :dealing with conceptual models invention for improving strategic decisions in a dynamic emergent way . This complexity theory background is another way to explain the cognitive shift the design profession has to face from an 'activity based' profession to a 'knowledge

based' profession . Developing 'global design' thinking as a sense making' decision environment on top of each design discipline - product, packaging, graphic, web,interface - with new specific conceptual tools.

4. Behind the representations 'sense making' "La pensée n'est pas faite pour servir la logique ; elle se sert de la logique" Edgar Morin

Hence the major responsibility in the governance of complex organizations is the search for and the expression of 'sense making'. For this issue , information and communication and knowledge are fundamental keys to pilot complex systems .

This responsiblity demands vigilance because the proliferation of signs in our present world intensified by NTIC have lead to a loss of meaning. So how to manage information and communication in order to make sense in complex situations ?

Give a shared meaning , put the communication media into a coherent framework of shared finality this is the objective . Designers have created graphic signs that are giving sense to corporate strategy values. But what complexity tells us is the importance of sense making for all the parts of the systems .Which means probably adapt the sign and design outputs to the level of logics of the different actors. The designer as transmitter has to refer to the mind of the users at different logics level.And to learn more on Karl Weick theory and research on representations and 'sense making'.

'User Oriented Design' or human centered design is a new trend of design model (Veryzer Borja de Mozota 2005). But it is rather focused on understanding users externally. The systems should be broadened to internal users. This is already taking place in brand management where the objective of brand equity building has emphasized the neccessity of managing brand value both externally and internally.

Another level of design responsibility in building representations that make sense is innovation. Because innovation is by essence a complex process uncertain,it does not go very well with traditional management tools .Innovation requires approaches that come from complex thinking.The challenge is to conduct a process which results are not predetermined.One can create the conditions to innovate but not ex ante the contents of innovation. The exercise of mental mobility proposed by complex thinking is one of the best openings for innovation . Design thinking is pertinent here because innovation means transgressing existing models , cross fertilization of ideas , cognitive process for piloting organizational learning rather than rationalizing .

The change in strategic thinking from a planification view to a managerial view of does not deny complexity .Practices change in order to be able to take into account complexity. There is a progressive evloution of the concept of strategy towards the integration of complexity explains why more and more MBAs are partnering with design schools or why design education is more and more research and PhD based

5. Rethinking organizational design The organization is weaving together different logics of invention not of conformity.It has to invent practices for self-organization .Any organization learns , memorizes and builds the intelligence of its evolution ,

Students in management seem to be developing the topic of organizational design in an unwieldy direction.As the variables multiply ,they tend to simplify the reality and say less and less about more and more. If you look into the definition of organizational design you begin to see the obstacles that management has put in its own path . Consider a definition of organizational design : 'the choosing of structures and associated managerial processes to enable an organization to operate effectively'.

On one hand organizational design creates nothing but one the other hand organizational design creates everything since organizational design will have a fundamental framing effect on people's expectations and perceptions . If designing the organization is both nothing and everything ,one needs to dig deeper into the process.

When people try to make sense of the world around them, they follow a progression towards abstraction ,from perception to form naming .And at this stage the reification of things dictate behavior .Labels are seen as intrinsic and imposed rather than emergent and discovered.

At each step in the sequence of compounded abstraction details get lost; the concrete is replaced by the abstract and design options gets foreclosed. Managing as designing means in part the monitoring ,containing and reversing of compounded abstraction .

The necessity to act this way becomes clearer when we connect the image of compounded abstraction more directly to the complexities of organizational life. If managers keep imposing machine metaphors and mechanistic assumptions onto events in a effort to stabilize them,predict them,and control them ,then categories, stereotypes,schemas,routines seem like useful tools.This is a pervasive scenario in organizational design where management expertise has become the creation and control of constants ,uniformity and efficiency .

The problem is that version of managerial expertise is no longer what we need. Instead the need has become the understanding and coordination of variability, complexity and effectiveness.(Weick 2004)

This means that managers need to create designs that mix together perceptual and conceptual modes of action .Design can be seen as a mental concept of human relationships in a world of exploding complexity and diversity .

Conclusion : The reality of the firms may be only the reality of the models its various actors have been building .There is probably not a good ,a better or a true model by itself but on the contrary multiple and lively projects by which everyone represents the firm he or she contributes to build by its action .And these rich ,intelligible ,communicative and dynamic respresentations are in permanent transformations :they have sense and they are sense making.

It is not because eveyone is active at every level that a firm builds sense . Sense making emerges when people are able at their level of responsibility to understand and interpret the level of logics superior to them. This progression in the sense making elaboration supposes an improvement in the richness of the contexts of interpretation . This asks the question of the cultural openess of the firm and of the richness of the profound intentions whether individual and collective , nourishing themselves reciprocically .

Firms are not to produce 'de l'organise' mais de l'organisant' organized routines but rather organizing schemes. Firms are active actors of a building that does not pre-exist. Managing in complexity might be enriching our representations of the firm and enriching our reasoning on the models of the firm .

References :

Borja de Mozota Brigitte , 2003, Design Management , Allworth Press,NYK,278 pages

Broadbent John ,2005 ,Towards an evolutionary ecology of design , evolutionary considerations , Sixth International conference of the European Academy of Design , EAD6, University of the Arts,Bremen , March 29-31

Génelot Dominique, 2001, Manager dans la complexité ,Insep consulting, Paris

Morin Edgar, Le Moigne Jean Louis, 1999, L'intelligence de la complexité, L'Harmattan, Paris

Sardar Z , Abrams I, 2003, Introducing chaos ,Cambridge iconBooks

Townson David, 2005, Mapping design strategy in complex organisational systems, Sixth International conference of the European Academy of Design ,EAD6, University of the Arts,Bremen , March 29-31

Veryzer Robert, Borja de Mozota Brigitte , 2005 , The impact of User-Oriented Design on New Product Development : An examination of Fundamental relationships,Journal of

Product Innovation Management ,volume 22 ,p 128-143.

Weick Karl ,2004, in Managing as Designing , Boland Richard Collopy Fred, editors, Stanford University Press

Measuring graph symmetry: discussion and applications

Carlos Bousoño-Calzón
cbousono@tsc.uc3m.es

Apart form its aestetical appeal, symmetry is considered as an essential tool for understanding complex systems[1]: It has been traditionaly used in the realm of theoretical physics, to study differential equations or dynamical properties [2], to name just a few. Recently, since network and graph theoretic tools grow as unifying concepts for complexity and universal properties studies [5], it seems natural to ask for the symmetry of networks and its implications.

In order to address this target, a reference to algebraic graph theory is deserved [6,7]. Nevertheless, there is an additional need of effective computational tools to tackle real applications. In this line of research, two measures for graph symmetry have been proposed recently which have rendered useful in order to predict dynamical behaviour [3] or robustness of a network in the face of attack [4].

This short paper elaborates on this kind of measure. Although the both aformentioned measures are essentially equivalent for a subset of graphs, their application to graph spectra (for Erdos or scale free random models [5]) raises some validity questions which are empirically explored in this proposal in order to propose a better symmetry measure. Some applications so as to illustrate its potential utility are developed in toy problems.

References:

[1] G. Mack: Universal dynamics, a unified theory of complex systems. Emergence, life and death. arXiv: hep-th/0011074 v1 9 Nov-2000. [2] M. Golubitsky, I Steward: The symmetry perspective. Birkhauser, Berlin, 2002. [3] A. Oropesa, C. Bousoño: A structural symmetry measure for complex systems: a neural network case. CSIMTA Int. Conf., pp: 119-122, France, 2004. [4] A. Dekker, B. Colbert: The symmetry ratio of a network, CATS2005. Australia, 2005. [5] R. Albert, A. Barabási: Statistical mechanics of complex networks. arXiv: cond-mat/0106096 v1 6 Jun 2001. [6]N. Biggs: Algebraic graph theory, Cambridge Univ. Press, 1993. [7] F. Chung: Spectral graph theory, CBMS-92, AMS, 1997.

A generic model simulating two temporalities of evolution in the European system of cities

Poster presented by Anne Bretagnolle and Jean-Marc Favaro (UMR Géographie-cités and university Paris I)

SIMPOP2 and EUROSIM are two versions of a same generic multi-agent systems model. The model is built in cooperation between geographers (UMR Geographie-cités) and computer scientists (LIP6). The model simulates the evolution of a network of interacting cities. The main interaction is the competition for resources (population, goods, information, innovation). The cities produce and invent new resources and exchange them on a spatialised market. Cities are heterogeneous agents that differ according to their functional specialisation (central market places, territorial capitals and production of non central manufacturing goods or services). They are also different according to their strategy, as defined by an agent called "governance". Two applications are presented here:

- the first one (model SIMPOP2) is theoretical and aims at simulating the emergence and the evolution of the European system of cities, from the end of the Middle Ages to nowadays. Two main features are to be reproduced by the model: 1) the persistency of a hierarchical configuration both at macrogeographic level of the system as a whole and in the individual trajectories of cities; 2) the dramatic increase of contrasts in city sizes since the first industrial revolution. The key parameters that will be experimented are: i) the succession of innovation cycles of different temporal scales and their more or less rapid diffusion within the system of cities; ii) the increasing spatial range of exchanges due to the progress in communication technologies.
- the second application (model EUROSIM) refers to the European urban system during the last 50 years. Urban functions and parameters are described more precisely, in order to reproduce as accurately as possible the operating and evolution of the system during the last decades. In that case, simulations are used for testing scenarios and making predictions at a 50 years horizon, according to migration policies, demographic expectations from different sources, and the evolution of the international boundaries barrier effects as a function of enlargement of European Union.

Complexity in living organisms : mosaic structures

Georges Chapouthier, Chapouthier,
chapout@ext.jussieu.fr

The present thesis (Chapouthier, G., 2001, L'homme, ce singe en mosaïque, Odile Jacob, Paris), which is compatible with Darwinian theory and arguments, endeavours to provide original answers to the question of why the evolution of species leads to beings more complex than those existing before. It is based on the repetition of two main principles alleged to play a role in evolution towards complexity, i.e. "juxtaposition" and "integration". Juxtaposition is the addition of identical entities. Integration is the modification, or specialisation, of these entities, leading to entities on a higher level which use the previous entities as units. Several concrete examples of the process will be given. At the genetic level there is silent duplication and integration of introns, which explains the origin of complex organs. At the anatomical level, the application of the two principles can be found in unicellular organisms (that develop into "juxtaposed organisms", e.g. Gonium, and then into "integrated organisms", e.g. Volvox). In more complex didermic species, juxtaposition produces colonies of polyps and the integration of these polyps produces intagrated siphonophores. In tridermic organisms, the juxtaposition of metamers produces the earthworm, whereas integration leads to the bee, the octopus or the chimpanzee. At higher levels, juxtaposition and integration become social, with the grouping of identical individuals (crowds) or the specialisation (or integration) of their roles (colonies of bees or societies of primates). For these structures, where integration at one level leaves the units at a lower level in a state of relative autonomy, the metaphor of the "mosaic" could be used. In complex living beings, as in a mosaic, the properties of a given level, taken as a whole, leave the autonomy of the component parts intact. Several examples will be given, specifically in the human brain and the functioning of thought, where essential functions such as language or memory have a mosaic structure.

Complexity in living organisms : mosaic structures

Georges Chapouthier, Chapouthier,
chapout@ext.jussieu.fr

The present thesis (Chapouthier, G., 2001, L'homme, ce singe en mosaïque, Odile Jacob, Paris), which is compatible with Darwinian theory and arguments, endeavours to provide original answers to the question of why the evolution of species leads to beings more complex than those existing before. It is based on the repetition of two main principles alleged to play a role in evolution towards complexity, i.e. "juxtaposition" and "integration". Juxtaposition is the addition of identical entities. Integration is the modification, or specialisation, of these entities, leading to entities on a higher level which use the previous entities as units. Several concrete examples of the process will be given. At the genetic level there is silent duplication and integration of introns, which explains the origin of complex organs. At the anatomical level, the application of the two principles can be found in unicellular organisms (that develop into "juxtaposed organisms", e.g. Gonium, and then into "integrated organisms", e.g. Volvox). In more complex didermic species, juxtaposition produces colonies of polyps and the integration of these polyps produces intagrated siphonophores. In tridermic organisms, the juxtaposition of metamers produces the earthworm, whereas integration leads to the bee, the octopus or the chimpanzee. At higher levels, juxtaposition and integration become social, with the grouping of identical individuals (crowds) or the specialisation (or integration) of their roles (colonies of bees or societies of primates). For these structures, where integration at one level leaves the units at a lower level in a state of relative autonomy, the metaphor of the "mosaic" could be used. In complex living beings, as in a mosaic, the properties of a given level, taken as a whole, leave the autonomy of the component parts intact. Several examples will be given, specifically in the human brain and the functioning of thought, where essential functions such as language or memory have a mosaic structure.

Peer-to-peer data management: the SP2+SP6 perspective

Giovanni Cortese„ Stefano Leonardi, Friedhelm Meyer auf der Heide, Christian Schindelhauer
schindel@upb.de

The goal of this workshop is to enlight the contribution of complex systems research towards the goal of a new generation of information architectures built on a self-organizing P2P architectures.

The Web bears the potential of being the world's largest encyclopedia and knowledge base, but we are very far from being able to exploit this potential. Search-engine technologies provide support for organizing and querying information. Yet, the current systems mainly support the simple mass-user queries. While advanced information queries too often require excessive manual pre- and postprocessing.

Approaches based on Collaborative Web information search in an Internet-scale peer-to-peer (P2P) system bear the potential of overcoming the shortcomings of today's Google-style search engine technology. In the approach proposed within the DELIS project every user (peer) has a full-fledged search engine that indexes a small portion of the Web, according to the interest profile of the user. Thus the requirements to the peer-to-peer overlay network are higher than in the usual key lookup or file sharing scenario. Consistent Hashing, aka. Distributed Hash Tables (DHT), is the state of the art data structure for such peer-to-peer networks. Because of the more involved task of evaluating documents and various locality issues new designs are necessary. Therefore work on the design of peer-to-peer networks is an important part of the scope of this workshop.

From a different perspective, Delis is studying large-scale IT infrastructures which, in order to be managed or used, pose challenging data management requirements. Such infrastructures include GRIDs, large networks, peer-to-peer applications etc. Their management requires (at least) the ability of building and querying large resource directories, of collecting and processing (summarising) lots of usage and performance data, also mining such data (in real time or near-real time) for detection of significant episodes (e.g. routing failures, denial of service attacks etc).

Access to management data should be decentralised, thus allowing many components in the system to efficiently query or subscribe to information. Also, management information should be organized and indexed to allow users (or applications) to issue queries at a higher level than supported by current frameworks (e.g. SNMP, HTML/XML), using query languages that support taxonomical reasoning and eventually the full power of ontology languages such as Owl. Again, the peer-to-peer approach based on DHTs with its promises for load balancing is a starting point for our research, to be explored for building large-scale, self-managing data repositories, however requiring a number of extensions.

Emphasis in the workshop will be given to understanding the dynamics and behavior of such a network requires analyses at different levels and scales of the overall network.

The Inter-disciplinary Analysis of Multidimensionality of Complex Systems' Evolution and the Method of its Topological Estimation

Victor F. Dailyudenko
selforg@newman.bas-net.by

The inter-disciplinary analysis of complex systems (CS) is implemented from standpoint of enlarging their degrees of freedom that is shown to be the intrinsic feature of their functionating. The method of nonlinear analysis of data describing evolution of CS is developed. Using temporal localization along phase trajectories of the attractor, we achieve the essential reduction of computation time and required experimental data at computational analysis of the attractor's topological dynamics that allows the algorithm to be realized even for higher-dimensional cases (much more than a hundred degrees of freedom). The numerical simulations confirm reliability of the developed algorithm and its high efficiency. So, the algorithm is applicable for the sake of statistical characterization of CS under investigation. In particular, within tasks of modeling of turbulent flows the method can be applied for the automatic detection of appearance of turbulence in technical devices. Time series (TS) obtained from CS are essentially nonlinear [1,2] and often lead to a multidimensional attractor in a relevant phase space [2]. Namely, it occurs at investigation of highly-developed turbulence where applying three-dimensional models (as in the model of Lorenz) is not enough for description of complex processes and higher-order modes become important for increasing reliability [1]. Multidimensional attractor also arises at modeling of nonlinear delayed feedback described by delay differential equations [2], those are widely applied for modeling of nonlinear processes in optics and laser physics, medicine and population dynamics, nonlinear vibrating systems and various self-sustained oscillations, in tasks of automatic control. But it is worthy to note that the main problem of numerical analysis of TS in such high-dimensional cases is that the computation complexity of fractal-topological algorithms essentially increases with enlarging a dimension m, as well as a quantity of required experimental data N. Namely, the computation complexity increases exponentially for the box-count algorithm and almost linearly for the Grassberger - Procaccia algorithm (GPA) with growth of m at expense of growing a number of computation operations. Again, additional increase of computation complexity results from respective growth of N (N increases exponentially for the GPA with enlarging m). For this reason, implementation of such algorithms for multidimensional attractors is cumbersome and even impossible for high m. So, with comprehensive inter-disciplinary analysis of multidimensional systems, the main purpose of this work is to develop a method allowing reduction of N and computation time as well as being insensible to growing m on these characteristics. This is attained by elaborating a topological method based on temporal localization in relation to points of the attractor. The most conventional methods of fractal-topological analysis imply just the spatial localization, i.e. investigation of distribution of points on the attractor basin based on estimating the quantity of hits into the m-dimensional cell with a size l. In particular, for the GPA this cell can be considered as a ball with a center in an attractor point. Similar approach of spatial localization (but with fixed number of nearest points) is used in a "nearest neighbor" method. In contrast to these methods, we show that temporal localization provides more convenient realization of topological analysis with essential reduction of required experimental data and computation time and makes these characteristics practically independent on dimensionality within some restricted range of changing m, that is the development of the approach couched in [3].

References 1 Guckenheimer, J, Holmes, P: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Berlin, Springer - Verlag, 1983 2 Farmer, JD: Chaotic

attractors of an infinite-dimensional dynamical system. Physica D 1982; 4: 366-393. 3 Dailyudenko, VF: Characterization of the topological structure and stability for a vector map derived from a delay differential equation. Nonlinear Phenomena in Complex Systems (Minsk). 2000; 3: 231 - 241

[1]

Centrality and vulnerability in weighted complex networks with spatial constraints

Luca Dall'Asta†, Alain Barrat†,
Marc Barthélemy‡ and Alessandro Vespignani†§

May 8, 2005

Abstract

Structure and functionality of real weighted networks are results of a complex interplay between topological and weights properties. The present work is devoted to understand the influence that topology and weights have on the vulnerability of weighted networks under intentional damage. In particular, weighted networks with spatial constraints are considered, the world-wide air-transportation network being a typical example. Since different attacking strategies can lead to very different conclusions on the vulnerability of a network, we first characterize some relevant topological and weighted centrality measures. These quantities are then used as selection criteria for the removal of vertices, in order to find which measure of centrality is the most effective. We also apply the same analysis on a recent model of growing weighted network with spatial constraints, that has been shown to reproduce some particular characteristics observed in air-transportation networks. By comparing the results, we gain useful information for a better understanding of the properties of growing weighted networks with spatial constraints.

*† Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210 Université de Paris-Sud 91405 ORSAY Cedex - France
‡ CEA-Service de Physique de la Matière Condensée, 91191 Bruyeres-Le-Chatel, France
§School of Informatics and Department of Physics, Indiana University, Bloomington, IN 47408, USA

1

1

Centrality and vulnerability in weighted complex networks with spatial constraints

Luca Dall'Asta†, Alain Barrat†,
Marc Barthélemy‡ and Alessandro Vespignani†§

May 8, 2005

Abstract

Structure and functionality of real weighted networks are results of a complex interplay between topological and weights properties. The present work is devoted to understand the influence that topology and weights have on the vulnerability of weighted networks under intentional damage. In particular, weighted networks with spatial constraints are considered, the world-wide air-transportation network being a typical example. Since different attacking strategies can lead to very different conclusions on the vulnerability of a network, we first characterize some relevant topological and weighted centrality measures. These quantities are then used as selection criteria for the removal of vertices, in order to find which measure of centrality is the most effective. We also apply the same analysis on a recent model of growing weighted network with spatial constraints, that has been shown to reproduce some particular characteristics observed in air-transportation networks. By comparing the results, we gain useful information for a better understanding of the properties of growing weighted networks with spatial constraints.

*† Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210 Université de Paris-Sud 91405 ORSAY Cedex - France
‡ CEA-Service de Physique de la Matière Condensée, 91191 Bruyeres-Le-Chatel, France
§School of Informatics and Department of Physics, Indiana University, Bloomington, IN 47408, USA

1

1

Centrality and vulnerability in weighted complex networks with spatial constraints

Luca Dall'Asta†, Alain Barrat†,
Marc Barthélemy‡ and Alessandro Vespignani†§

May 8, 2005

Abstract

Structure and functionality of real weighted networks are results of a complex interplay between topological and weights properties. The present work is devoted to understand the influence that topology and weights have on the vulnerability of weighted networks under intentional damage. In particular, weighted networks with spatial constraints are considered, the world-wide air-transportation network being a typical example. Since different attacking strategies can lead to very different conclusions on the vulnerability of a network, we first characterize some relevant topological and weighted centrality measures. These quantities are then used as selection criteria for the removal of vertices, in order to find which measure of centrality is the most effective. We also apply the same analysis on a recent model of growing weighted network with spatial constraints, that has been shown to reproduce some particular characteristics observed in air-transportation networks. By comparing the results, we gain useful information for a better understanding of the properties of growing weighted networks with spatial constraints.

*† Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210 Université de Paris-Sud 91405 ORSAY Cedex - France
‡ CEA-Service de Physique de la Matière Condensée, 91191 Bruyeres-Le-Chatel, France
§School of Informatics and Department of Physics, Indiana University, Bloomington, IN 47408, USA

[1]

Centrality and vulnerability in weighted complex networks with spatial constraints

Luca Dall'Asta†, Alain Barrat†,
Marc Barthélemy‡ and Alessandro Vespignani†§

May 8, 2005

Abstract

Structure and functionality of real weighted networks are results of a complex interplay between topological and weights properties. The present work is devoted to understand the influence that topology and weights have on the vulnerability of weighted networks under intentional damage. In particular, weighted networks with spatial constraints are considered, the world-wide air-transportation network being a typical example. Since different attacking strategies can lead to very different conclusions on the vulnerability of a network, we first characterize some relevant topological and weighted centrality measures. These quantities are then used as selection criteria for the removal of vertices, in order to find which measure of centrality is the most effective. We also apply the same analysis on a recent model of growing weighted network with spatial constraints, that has been shown to reproduce some particular characteristics observed in air-transportation networks. By comparing the results, we gain useful information for a better understanding of the properties of growing weighted networks with spatial constraints.

*† Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210 Université de Paris-Sud 91405 ORSAY Cedex - France
‡ CEA-Service de Physique de la Matière Condensée, 91191 Bruyeres-Le-Chatel, France
§School of Informatics and Department of Physics, Indiana University, Bloomington, IN 47408, USA

Complex Systems and Cognition: the incoherent dynamics implementation using multi-agents systems

Leonardo Lana de Carvalho, Salima Hassas
llanadec@cnrs.liris.fr

This paper presents an approach of addressing cognition from the perspective of complex systems by using the multi-agents paradigm. Whitakers notion of cognition is based on Maturana and Varelas theory and stipulates that cognitive systems are a consequence of complex systems capable of self-adaptation (Whitaker, 1997). Rocha and Hordijk (2004) suggest representation or incoherent dynamics as inert structures: « organisms employ inert structures to store initial conditions used to construct appropriate dynamical (active bio-chemical) components as well as reproduce other organisms such as themselves ». For a given dynamics, « ... we can treat certain components as incoherent, if their dynamics are irrelevant for those time scales. We can further treat them as memory, if such dynamically incoherent components are used as information to specify sets of initial conditions for the original dynamics. » (p-11). According to authors it would be enough to stop the reactivity of the system. Evolving Cellular Automate and Genetic Algorithm, however « the information they store is still not separated from the dynamics »; « CA were able to create static memory stores, [but] these are still reactive with the embedding dynamics » (Rocha and Hordijk, 2004, p-25). Agent-based models demonstrated their effectiveness on several computation problems. The variation of the population size provides the swarm with mechanisms that improves its self-adaptability and causes the emergence of a more robust self-organized behavior, resulting in a higher efficiency on searching peaks and valleys over dynamic search landscapes (Fernandes et all, 2004). To go further than the Rocha and Hordijks work, we make the hypothesis that Incoherent Dynamics (ID) is what makes the emergence of representations possible in complex systems, by making cognitive the computational system. This emergence of representations takes place like an opposition of the system to its environment, breaking its reactivity while being embodied. To illustrate this idea, a metaphor of natural ants collective sorting was implemented. Agents were designed like simple automatas which obey linear functions to move randomly in the environment by leaving a central point. As expected, after the agent-environment interactions the behavior of agents is no more linear, but exhibits emergent complexity. We observe that small modifications in automated functions of moving causes a greater or minor dispersion of the amount of f̈ood. Ahead this emergent property, dispersion-centralization, characterizes a new system, called here Order. A dispersed order is got with agents who cover small distances before verifying the food existence. A centralized order is got with agents who make it after covering long distances. Thus we have implemented different orders on one same environment and an order after the emergence and stabilization of another. It is instructive to recognize which relationship exists between the two emergent orders. We stress that relationship between orders are not reducible to agent-environment level. With the studies of the relationships between orders, it was possible to observe an opposing emergent system, that is, completely not programmed. We had observed that a dispersed order can always be isolated in a certain environmental site where was embodied a previous order. The emergence of opposition is obtained by heterogeneity of agents. It is necessary that the implementation of heterogeneity takes place after the stabilization of the first observed order which corresponds to the centralization phenomenon. Indeed the emergence of opposition opens a new way in approaching Incoherent Dynamics. Isolated orders are incoherent systems because they are essential for maintaining the global pattern. They have their own dynamics, they can be controlled by the global pattern itself and they could be used as an internal representation of an external system able to promote hetero-

geneity effects into agents. Thus, it seems that we have a bridge between the computability and the cognitivity. However, the heterogeneity of agents is still directly programmed. To generate ID based systems able of solving problems representatively and non reactively, it will be essential to have more profound levels of heterogeneity in the agents behaviors. We have two ways. Ascend multiples levels of emergence changing agent more insignificant or conceiving not only situated agents but embodied agents.

Percolation for Power Control (Abstract)

Emilio De Santis
Department of Mathematics
University La Sapienza of Rome

Fabrizio Grandoni and Alessandro Panconesi
Department of Computer Science
University La Sapienza of Rome

May 6, 2005

Loosely speaking, the problem of power control for wireless communication networks is that of adjusting the transmission power of the devices in a network in such a way that two conflicting objectives are met. On the one hand the communication radius of each node should be as small as possible, for power consumption grows (at least) quadratically with the radius. On the other, the radius should be large enough to ensure global connectivity or the emergence of a giant component. Energy efficient protocols ensuring good connectivity properties are particularly desirable in the case of ad-hoc wireless communication, for these networks have severe resource constraints. This problem can be formulated mathematically in different (not necessarily equivalent) ways. One of the most interesting models is the so-called *nearest-neighbor* model. Here we have N points uniformly distributed at random in the unit square. Each node u in the network decides a number $k(u)$ and sets its transmission radius in order to reach the closest $k(u)$ neighbors. Links are then established according to different rules. Some authors consider the resulting oriented network. If node u is within the ball of radius $r(v)$ centered at v, then there is a direct link $v \to u$, meaning that v can send a message to u but the reverse is not necessarily true [3]. In this case strong connectivity is analyzed. Others remove the orientations altogether [1]. In this paper uv is an undirected edge if u is inside v's ball, and viceversa. This model is the most realistic from the point of view of wireless applications since the communication primitives of standards such as IEEE 802.11 and Bluetooth rely on ack messages, and therefore a communication link exists only when the two nodes are both within radius of each other. In any case the three models are equivalent from the point of view of the question studied in this paper.

Assuming N points uniformly distributed within a region is the standard assumption that is made when analyzing this kind of problems. In some cases this is quite realistic. Sensor networks for instance are small, inexpensive devices that can be deployed massively within a certain region, often at will. But the point is that proofs and properties proved in this model oftentimes do not really need the uniformity assumption, but something weaker that makes the results and the properties more widely applicable than originally envisaged.

The question we consider in this paper is, how large should $k(u)$ be *on average* for connectivity and/or the emergence of a giant component?

Xue and Kumar [1] show that, if $k(u) = C$ for all u, then C must grow like $\log N$. Note that this does not imply that the average degree must be $\Omega(\log N)$, since values for the $k(u)$ need not be equal. Indeed, Kucera [3] shows that on average $k(u)$ can be constant, even though some $k(u)$ will have to grow like $\log N$. This result however is existential in flavor, since the protocol used in [3], although distributed, is not local, for a node can explore a linear-size component of the network and this requires linearly many communication rounds. In wireless networks, and distributed computing in general, it is crucial to bound

the number of communication rounds of the protocol. In this paper we show that an exponential speed-up is possible by exhibiting a distributed protocol for power control that uses $O(\log^4 N)$ communication rounds. The average $k(u)$ is constant as in Kucera's protocol and the maximum $k(u)$ is $O(\log N)$.

To prove the result we establish along the way a useful fact. Namely, that there is a protocol to generate a giant component where each node has constant degree, and this protocol uses no communication at all. Moreover the giant component is uniformly spread within the area. This is a useful property because the main application envisaged so far for sensor networks is monitoring of a certain area by means of a connected network, and a giant component is enough for this purpose. The result we prove is a typical percolation result. We prove that there is a universal threshold K, that is, independent of N, such that a giant component arises if $k(u) \geq K$ for all u.

From the technical point of view we reduce our problem to bond percolation. Well-known results of percolation theory show that for independent, bond percolation a unique infinite component arises with probability one as the bond probability p exceeds a threshold $p_c \in (0, 1)$. In a finite box, we will have a unique giant component, while the remaining components will have size $O(\log^2 N)$. To use a metaphor, we will have land (sites that are on) and water (sites that are off). Land is organized in a unique super-continent of linear size and a set of very small islands, of poly-logarithmic size. These islands lie in seas (water between the border and the super-continent) or in lakes inside the super-continent. It can be proven that each sea and each lake, together with the islands they contain, are of poly-logarithmic size. All this happens with probability going to one as long as N goes to infinity, provided that the bond probability exceeds a threshold.

In the original graph this translates to a probabilistic test for membership in the giant component. If a node is in a component of size $\Omega(\log^3 N)$, then it can safely assume to be part of the giant component. Otherwise it will increase its transmission radius to reach a new node, until, within $O(\log N)$ such iterations, it will be connected with the super-continent.

References

[1] Feng Xue and P. R. Kumar, The number of neighbors needed for connectivity of wireless networks, Wireless Networks, Vol 10, No. 2, pp.169-181, 2004.

[2] Grimmett, G. Percolation. New York: Springer-Verlag, 1989.

[3] L.Kucera, Low Degree Connectivity in Ad-hoc Networks, DELIS Tech Report.

Science and Engineering of Business Systems

Kemal A. Delic
kemal.delic@hp.com

We are surrounded by the systems involving people, technologies and processes better known as the "business systems". They function well despite all deficiencies, imperfections and lack of coordination while creating the wealth and moving the world's economy forward. However, it seems that the practice of business systems is well ahead of the science of business systems.

Thinking about world's economy as the business ecosystem, implies dealing with large-scale systems, saturated with complexity, high uncertainty and rapid dynamics. As such, the large-scale business systems should represent an excellent playing field for the multiple, multi-faceted scientific disciplines and scientists, while, for the various reasons and causes, they are or disregarded or sometimes even carefully avoided.

Analysis of the complex system "TE-TA-P" [telomeres-telomerase-proliferation] coupled to experimental data in cancer cells.

Despite the difficulty for analyzing the physiopathological mechanisms which are involved, the practical importance of biological oscillations justifies the development of their study because they condition the response to an external stimulation depending on the time when it is applied to the system. The Te-TA-P network includes 3 oscillating parameters : - length of telomeres (chromosome endings) - telomerase activity (telomere repair)- cell proliferation rate. The Te-TA-P system is paradigmatic among oscillatory biological systems, because it is a key both for normal tissular development (characterized by progressive cell senescence due to definitive telomere erosion), and for long-term survival of stem cells and tumor cells (characterized by telomere repair which allows unlimited proliferative capacity). The maintenance of the system TE-TA-P at equilibrium is linked to the cancerous character of the cells. One can foresee that pushing the cells out of this equilibrium by light impulsions on a component of the system will durably change the proliferation dynamics. Our hypothesis is that such bifurcation of the dynamics, incompatible with persistent high proliferation, will induce the slow-down of mitoses and/or increase apoptosis, resulting in a negative balance of tumor cell growth.

The purpose of our experimental work on tumor cultured cells is 1) to analyse the interactions of the 3 variables in the Te-TA-P network by simultaneous quantification in long time series, 2) to modelize the metabolic loops that link them in order to predict the fluctuations, and finally 3) to control the activity of this system by external forcing.

This approach of the cell proliferation dynamics already brought us important results. It bears important applications on practical problems such as tumor regression and intermittent resistance to anticancer drugs.

Criteria for coalition formation

Jean-Louis Dessalles
ParisTech ENST (École Nationale Supérieure des Télécommunications)
dessalles@enst.fr - http://www.enst.fr/~jld

Like many other species, Human beings establish stable social bonds that are not necessarily based on kin. This ability is an essential component of sociality, and may be a crucial factor for the evolutionary emergence of the human species (Dunbar 1996). However, the criteria used by individuals to choose their social partners is poorly understood. In particular, no general law has been proposed to delineate what would count as a valid preference criterion in a strict Darwinian framework.

Social groupings based on mutual partner choice are sometimes said to be political. A typical example is the formation of political coalitions among chimps to take or keep control over the group (de Waal 1982). Another example is offered by babblers, those little birds living in the desert, which form coalitions to defend sheltering bushes against other babbler coalitions (Zahavi & Zahavi 1997). The human species could be characterised by the complexity of political bonds: friendship networks, alliances, supporters, partisans, leaders, factions, all emerge in a political context in which individuals are prone to choose each other. These social phenomena have often been described as resulting from psychological factors such as esteem, feeling of security, dominance. They have also been considered as cultural constructs emerging from a general willingness to cooperate, in which case the partner choice chiefly relies on reliability (Axelrod 1984).

We want to explore an alternative account of coalition formation, which is that individuals choose their partners according to their potential usefulness for the fate of the coalition. This approach departs from studies based on cooperation in several aspects.
- social bonds do not depend on some utilitarian trade of goods or services.
- the criteria for partner choice are expected to be in part biologically determined, and not based on pure rational choice.
- potential partners are expected to display qualities corresponding to those criteria, even if the display is costly (Dessalles 1999; Gintis, Smith & Bowles 2001).
- criteria for social bonding are required to be positively correlated with the success of the coalition.

The latter requirement is the main result we arrived at, both theoretically and through computer simulation (Dessalles 1999). In a political settings, isolated individuals have little chance of success, and they must choose their partners with discernment to resist coalitions formed by others.

In the present paper, we will explore the coexistence of several criteria. It may be in the interest of individuals to choose partners who have qualities different from their own. The result may be seen as a marketplace on which individuals advertise various competences and gauge each other's corresponding qualities. The marketplace metaphor is somewhat misleading, though, as no exchange of goods or services is necessary. The goods are the individuals themselves that may join to form or maintain coalitions.

References

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basics Books.

de Waal, F. B. M. (1982). *Chimpanzee politics: power and sex among apes*. Baltimore: The John Hopkins Univ. Press, ed. 1989.

Dessalles, J-L. (1999). "Coalition factor in the evolution of non-kin altruism". *Advances in Complex Systems* 2(2), 143-172.
http://www.enst.fr/~jld/papiers/pap.evol/99091402.pdf

Dunbar, R. I. M. (1996). *Grooming, gossip, and the evolution of language*. Cambridge: Harvard University Press.

Gintis, H., Smith E. A. & Bowles S. (2001). "Costly Signaling and Cooperation". *Journal of Theoretical Biology* 213, 103-119.
http://faculty.washington.edu/easmith/JTB2001.pdf

Zahavi, A. & Zahavi A. (1997). *The handicap principle*. New York: Oxford University Press.

Coarse-graining and continuum physics

Antonio DiCarlo
adicarlo@mac.com

I claim that an adequate continuum treatment of gross physics should be able to interact with the models of atomistic physics at small scales, and have a word to say on collective multiscale phenomena, which are of paramount interest in, say, materials physics and biophysics. Continuum physics as it stands today, however, is unfit for this task, being hampered by deeply rooted misconceptions, which make it blind to key scale issues.

In any physical theory, matter forms patterns in space that evolve in time. In an atomistic theory, molecules do not exist as matter: they are "mere" patterns. However, they are often the interesting objects, to be recognized among a terrible host of more ephemeral patterns on smaller space-time scales. A finer theory would deprive atoms of their material status, making them patterns of quarks, etc. The same conceptual scheme holds at supra-molecular levels including all of continuum physics: selected patterns of micro-matter appear as macro-matter in coarser grained theories. In continuum physics, matter is represented by spacetime averaged densities of molecular quantities, taken on a suitable meso-scale; fields of substantial markers are contrived according to rules in the candidate model of the system at the same scale, and the patterns that they form on larger scales are the objects one is interested in observing.

In complex systems, it is of the essence to relate phenomena happening at widely scattered scales, both in time and space, converting information from finer to coarser levels of description (and vice versa). A special difficulty resides in mechanics, which filters off more and more phenomena under up-scaling: only a tiny portion of the information contained in a detailed, microscopic knowledge of motions and interactions is subsumed by the corresponding macroscopic mechanical quantities, the rest being (partly) summarized in other macroscopic theories. While, in principle, all of chemistry is contained in quantum mechanics, macroscopic mechanics ignores all of it, and the coupling between mechanical and chemical phenomena (which can be of paramount importance: think of molecular motors in our own muscles) has to be rebuilt by the modeller.

Conventional continuum mechanics - supposedly representing the phenomena of interest at the macroscale - has too few sockets where to plug in the vast amount of macro-useful information surfacing from the depths of microscales. Too much of what comes from the bottom gets lost, or is bluntly lumped into constitutive black boxes. In the standard model of continuum thermomechanics, the body-points labeled by meso-markers possess as their only attributes position and temperature. Such a model is by far insufficient when dealing with complex systems, and has to be extended and generalized in two different directions. First, position is to be supplemented by extra order parameters, providing meso-averaged information on the presence of different molecular species and their conformations. Second - and more innovative - a mesoaverage of the stress-free, chemically determined, molecular conformation is to be tracked: in fact, the key for developing a comprehensive theory is in the ability to compare the actual conformation (stressed in general) with the corresponding stress-free prototype, which is also evolving in time because of the changing chemical conditions.

The format I call material remodelling comprises new balances, which govern the time evolution of material properties, such as the relaxed (i.e., stress-free) metric of body elements, or their elastic stiffness. Therefore, it is able to capture the gross mechanical effects of phenomena such as conformational changes and structural rearrangements. I advocate its potential usefulness in bridging the gap between continuum physics and the physics of the invisible worlds underlying the continuum model.

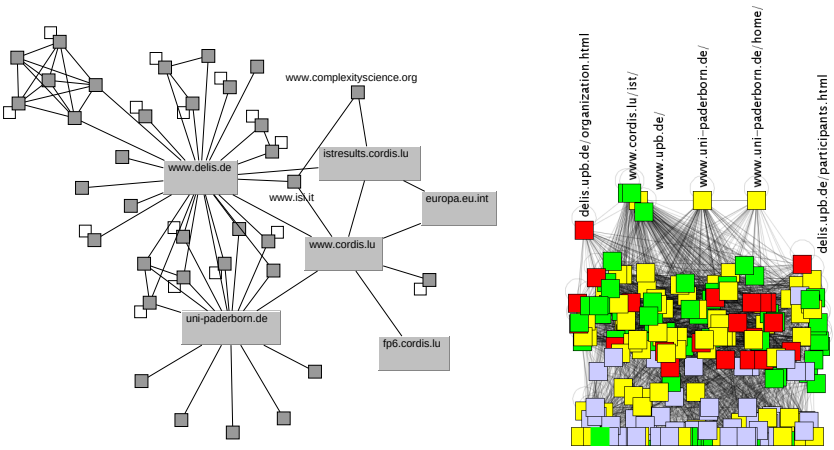
What are Complex Systems? What is DELIS?*

Debora Donato†, Marco Gaertler‡, Robert Goerke‡,
Stefano Leonardi†, Dorothea Wagner‡

Abstract

"What are Complex Systems?" – Commonly, the quest for information is accomplished by various search engines adapted to specific tasks such as Google or Yahoo! for web search and Citeseer for publication lookup. The results of these tools are usually lists that are sorted according to one relevance concept for the items. Nowadays, search queries become more complex and ambiguous. Therefore a simple evaluation of a query may not produce satisfactory results to the user.

In the following, we present visualizations enriched with search-relevant information, e.g., pagerank [2] or authority score [1] for the sites. As an example, we study the topics *complex system* and, in particular, *delis*.



The above visualizations show a cleaned web-crawl started from the DELIS homepage. The left hand figure represents the domain-structure, i.e., sites with the same domain are grouped together and their sizes reflect the number of sites contained. In contrast, the right hand figure visualizes the pagerank value on the y-axis (color reflects domain).

Using different analyses in combination with their visualizations allows for a more individual and user-friendly exploration of search spaces.

References

[1] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[2] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Manuscript, 1999.

*Work partially supported by European Commission - Fet Open project COSIN - COevolution and Self-organisation In dynamical Networks - IST-2001-33555, by European Commission - Fet Open project DELIS - Dynamically Evolving Large Scale Information Systems - Contract no 001907

†Universitá di Roma La Sapienza, Dipartimento di Informatica e Sistemistica, Roma, Itlay. http://www.dis.uniroma1.it/

‡Universität Karlsruhe (TH), Faculty of Informatics, Karlsruhe, Germany. http://i11www.iti.uka.de/

Connectivity and Routing in Poisson Small-World Networks

Moez Draief, Ayalvadi Ganesh
M.Draief@statslab.cam.ac.uk

In recent work, Jon Kleinberg considered a small-world network model consisting of a d-dimensional lattice augmented with shortcuts.The probability of a shortcut being present between two points decays as a power, $r^{-\alpha}$ of the distance r between them. Kleinberg showed that greedy routing is efficient if $\alpha = d$ and that there is no efficient decentralized routing algorithm if $\alpha \neq d$. The results were extended to a continuum model by Franceschetti and Meester. In our work, we extend the result to more realistic models constructed from a Poisson point process, wherein each point is connected to all its neighbours within some fixed radius, as well as possessing random shortcuts to more distant nodes as described above.

ABSTRACT FOR A SHORT PRESENTATION

The New Ties project: 3 dimensions of adaptivity and 3 dimensions of complexity scale-up

A.E. Eiben, N. Gilbert, A. Lörincz, B. Paechter, P. Vogt
www.newties.org

New Ties is an FP6/IST/FET open STREP project (started in fall 2004) seeking **N**ew and **E**mergent **W**orld Models **T**hrough **I**ndividual, **E**volutionary, and **S**ocial Learning.

The project is concerned with emergence and complexity in socially-inspired artificial systems. It investigates large systems consisting of an environment and an inhabitant population. The main goal of the project is to realize an evolving artificial society capable of exploring the environment and developing its own image of this environment and the society through cooperation and interaction. The “physical” environment is based on virtual grid worlds that are sufficiently complex and demanding so that communication and cooperation are necessary to adapt to the given tasks. The population's “toolkit" to develop advanced skills bottom-up consists of

1. individual learning,
2. evolutionary learning, and
3. social learning.

One of the main innovations of this project is social learning interpreted as passing knowledge explicitly via a language, which is evolved in the simulation, to others in the same generation. This has a synergetic effect on the learning processes and enables the society to rapidly develop an "understanding" of the world collectively. If the learning process stabilises, the collective must have formed an appropriate world map. Then the collective mind can be investigated to learn how the agents perceive the environment, including themselves, and what skills and procedures they have developed to adapt successfully. This could yield new knowledge and surprising perspectives about the environment and the survival task. The project represents a significant scale-up beyond the state-of-the-art in three dimensions:

1. the inner complexity of inhabitants,
2. the size of the population, and
3. the total processing power for duration of the simulations.

To achieve and explore highly complex organisms and behaviours, very large populations will be studied. This will make the system at the macro level complex enough to allow significant behaviours (cultures etc) to emerge in separate parts of the system and to interact. To enable this, a large distributed computing infrastructure is being set up, together with a shared platform to allow very large scale experiments in a p2p fashion.

The short presentation will outline the vision behind the project, the main objectives, the approach to be followed, and the expected outcomes. By the time of the conference we will have technical results that will be also briefly summarized (and given in more details by other presentations).

Monotonicity and Almost-Monotonicity in Biological Systems

G.A. Enciso E.D. Sontag

7 May 2005

A dynamical system is a type of continuous mathematical model which is frequently used to describe the behavior of physical systems over time. Dynamical systems are powerful modeling tools, and they have been successfully used for many applications, but they are conspicuously hard to study, especially when there is a high number of variables involved. One increasingly important application of dynamical systems modeling is the description of molecular biological processes at the cell level. It allows the modeler to incorporate a large amount of information about the process, such as what proteins influence which genes and at what rates these influences take place. A drawback for the modeler is precisely the fact that often there is a very large number of variables (proteins, genes, etc) involved. On the other hand, and for all the apparent complexity of the models, the behavior of the solutions over time is often relatively simple: the convergence of all solutions towards one or two points, or a stable oscillatory behavior, seem the rule rather than the exception. Moreover, the direct influence of one variable on another is often consistent: if a protein influences the production of a given gene, say, it often does so in order to consistently inhibit it or to consistently promote it.

A concept that allows to exploit these characteristics is that of a monotone dynamical system. In its simplest form, a monotone system is one in which all the influences among the variables are promoting, and no inhibitory influences are observed. But many systems with inhibitory reactions can also be monotone, as long as the indirect effect from any given variable to another is consistent along any path in the network. A necessary condition for a system to be monotone is for it to have consistent direct influences, which is commonly satisfied for molecular biological systems as described above. Also, monotone systems have a very stable behavior over time. By considering monotone systems with inputs and outputs and introducing a negative feedback, we were able to describe the behavior of systems that are not monotone ('almost-monotone') in terms of that of monotone systems, and to give sufficient conditions for such a system to converge globally towards

an equilibrium. This approach can potentially be used for a formal study of complex, non-monotone, large scale dynamical systems using techniques from monotone systems theory. We give a simple example of a testosterone dynamics model with delay, and show even in the presence of arbitrary delays all the solutions of the system converge.

RNA Secondary structure prediction

Stéfan Engelen and Fariza Tahi

Laboratoire La.M.I. - UMR 8042.
Université d'Evry-Val d'Essonne, CNRS, Genopole
E-mail: {sengelen,tahi}@lami.univ-evry.fr

Structural RNA are important for regulatory, catalityc or structural roles in the cells. They can act alone to catalyse RNA processing. They can also form protein-RNA complexes like the ribonuclease P which has a role in tRNA processing. The secondary structure is composed of all the Watson-Crick pairings, AU, GC, and the Wobble pairing GU. The consecutive pairings form helices and could be interleaved to make pseudoknots. The knowledge of secondary structure is essential to understand the relations between structure and function of the RNA.
One approach to predict the secondary structure of RNA is to search covarying residues which maintain the Watson-Crick pairings. This approach is called the comparative approach and consists to retrieve mutation informations from homologous sequence alignments.

We have developped an algorithm called P-DCfold implementing the comparative approach. It is based on the "divide and conquer" approach which allows a low complexity in time. This method segments the prediction in under less complex problems which underline the fact that the RNA secondary structures are also segmented in under fields. The helices are searched recursively from the "most significant" to the "less significant". Thanks to this approach and to the criteria we use (criteria of minimum length of searched helices and number of the compensatory mutations in these helices), the algorithm allows to obtain predictions of very good quality, since almost all the structure helices are predicted and no false positive helices are selected. P-DCfold algorithm allows also to predict the pseudoknots appearing in the structure. It searches for all kinds of pseudoknots in a complexity in time of $O(n^2)$ when algorithms existing in litterature are of complexities higher than $O(n^4)$ for the prediction of only certain kinds of pseudoknots.
An important point when using the comparative approach for the predicion of a RNA secondary structure is how to choose the homologous sequences to use for the comparison. More precisely, how to choose the ones which allow to obtain good predictions. A preprocessing of available homologous sequences makes it possible, by assigning scores to these sequences. We expound an algorithm, called SSCA, which measures the interest of a sequence. The measurement is based on evolutionary model in helices regions. This algorithm is in complexity in time of ($O(n^2)$).

Thus we obtain a complete system which implement efficiently the comparative approach for the RNA secondary structure prediction and which goes to the prediction of the tertiary structure, since pseudoknots can be considered as elements of the tertiary structure instead of the secondary one.

The efficiency of our system concerns its complexity in time ($O(n^2)$), its complexity in space (less than $O(n^2)$) and above all the quality of the predictions. It has been tested in several RNAs: tmRNA, RNaseP, 5S RNA, U1 RNA, SRP RNA, tRNA, 16S RNA and 23S RNA. In all these examples, more than 80 percent of the helices have been well predicted as well as all the pseudoknots and the predictions have been done in immediate time for small RNAs and less than 5 seconds for longuest ones (16S and 23S RNA, respectively of about 1500 and 2400 nucleotides).

Finally, several applications of the system could be considered: prediction of unknown structures, search for characteristic structures ("hairpin" and "hammerhead" structures), prediction of structures of certain regulating areas (5 ' and 3' areas of genes, introns ...), etc.

Studying decentralized collective change of behaviours : the example of phase transitions in elementary cellular automata

Nazim Fates
Nazim.Fates@ens-lyon.fr

Complex systems are usually defined using the well-known formula that the ẅhole is more than the sum of its parts.̈ This informal idea expresses the possibility for a system defined as a collection of simple components in interaction to exhibit a c̈omplexb̈ehaviour even when each component behaviour is very-well understood.

Cellular automata (CA) are a well-known model for defining such systems : they are constituted by a collection of finite state automata that are regularly located alond a line (1D) or a grid (2D). These automata, or cells, interact with their neighbours and change their state according to a transition function. In this work, we shall only consider elementary cellular automata (ECA), i.e., two-state one-dimensional CA that change their state according to their own state and the state of their two immediate neighbours. Classically, CA are updateed sycnrhonously : the transition function is applied to all cells simultaneously. The model we here investigate does not use this hypothesis of perfect synchrony : instead each cell has propability α, called the synchrony rate, to be updated and if it is not updated it stays in the same state.

The question we here investigate is to know whether it is possible that the CA (collective) behaviour undergoes a brutal change while the synchrony rate is continuously increased. In a previous work [1] we showed that some ECA had indeed the striking property of displaying two different behaviours when the synchrony rate was lower or higher than a critical value α_c. This article aims at finding an explanation for such a phenomenon.

Using a methodolgy very close to the one used in Grassbergerś work [2], we conducted two sets of experiments to determine if the change of behaviour was was a directed percolation phenomenon (i;e., a particular king of second order phase transition). The first set of experiments aimed at determining α_c for which the transition occurs and teh second set aimed at calculating the critical exponents.

Analysis of the results allowed us to deduce good agreement with the values predicted for directed percolation. Moreover, observation of space-time diagrams also show similarities of patterns observed with other experiments that involved directed percolation (e.g., in hydrodynamics see work of Pomeau).

To conclude, we will discuss why such phase transitions are important. Indeed,they provide an striking example where the synchronization of elements is a key feature to understand the collective behaviour of *complex system*. Such mechanisms could provide a way to explain some phenomena in the biological domain : it is for example striking to see how the amiboe *Dictyostelium Discoidum* brutally change their behviour form an individual one to a collective one when the yare submitted to a stress. Modelling such organisms with cellular automata may thus help us understand some of their most interesting features.

refs: [1] An Experimental Study of Robustness to Asynchronism for Elementary Cellular Automata, Nazim Fatès, Michel Morvan To appear in Complex Systems

[2] Sycnhronization of coupled systems with spatiotemporal chaos Peter Grassberger Physical Review E vol. 59 Number 3

Traffic distribution in scale-free networks

A Fekete, G Vattay, L Kocarev
lkocarev@ucsd.edu

In this paper we study one of the most relevant example of scale-free networks with embedded flow dynamics: the Internet. The dominant algorithm which controls the data traffic in the Internet is the Transmission Control Protocol (TCP). We will present a simple model, which considers both the scale-free structure of the Internet and the adaptive nature of the underlying dynamics. Since links mean real physical boundaries for the traffic flow in the Internet, bottlenecks are formed at links. It follows that edge betweenness should be investigated. An exact analytic solution is presented for its distribution, which apply for even finite networks. Finally, the performance of the TCP model is compared in different network scenarios using simulation methods.

Thresholds for the emergence of cooperation between signals

Ramon Ferrer i Cancho, Vittorio Loreto
ramon@pil.phys.uniroma1.it

Little is known from the mathematical point of view about the reasons why signals start to cooperate spontaneously for communicating. Human words combine to form phrases and sentences. Often, the meaning of a single word is incomplete or too vague. A combination of two or more words is needed in many cases. Recent mathematical models have studied the conditions turning combinatorial communication advantageous over isolated signals in the context of biological evolution (Nowak et Krakauer, 1999; Nowak et al., 2000). Some of the models indicate that the presence of noise in the communication channel may eventually turn signal combinations advantageous over isolated signals (Nowak et Krakauer, 1999), which is not surprising from the point of view of Shannon's noisy coding theorem (Shannon, 1948; Plotkin et Nowak,2000). Besides, those models make little use of information theory and use a particular pay-off function for defining the reliability of a communication system.

Here we start from the communication framework developed in (Ferrer i Cancho, 2005b-d) on a noiseless channel. We assume that biological evolution cannot choose cooperation between signals unless the information transferred by a combinatorial system is larger than the amount transmitted by isolated signals. Thus, we study the topological properties of the network of signal-object associations under which cooperation is advantageous and under which it is not using an information theory approach. We find the existence of domains where cooperation does not have any advantage and domains where it has. We identify some wide mathematical conditions for belonging to one or the other domain. Previous models (Nowak Krakauer, 1999; Nowak et al., 2000) have shown the existence of thresholds in the number of signals or stimuli for the emergence of cooperation. Our model indicates that the topology of the network of signal-stimulus associations (e.g. the network density) cannot be neglected for the emergence of cooperation. The presence of phases and thresholds between them suggests that cooperation may be not accessible to any species.

We show that maximizing the information transferred by individual signals, decreases the chances of advantageous cooperation. Our results support the hypothesis that cooperation has higher chances of being advantageous in degenerated communication systems, which confirms a possible track for the origins of human language language(Ferrer i Cancho, et al. 2005). Being a bad communicator with individual signals should not be seen as a very negative feature since it translates into higher chances of developing advantageous cooperation.

Clusters of computations for a linear transition system

Vladimir Filatov, Rostislav Yavorskiy, Nikolay Zemtsov[1]

Department of Mathematical Logic
Faculty of Mechanics and Mathematics
Moscow State University
Moscow, 119992, Russia

Let M be a state transition system (S, τ) where S is a state space, τ is a transition function. We study the case when $S = \mathbb{R}^n \times \mathbb{Z}^m \times \Omega$, where Ω is a finite set. That is every state of the studied system is characterized by a real-valued n-vector $\vec{x} \in \mathbb{R}^n$; an integer-valued m-vector $\vec{v} \in \mathbb{Z}^m$; and one variable of a finite range $w \in \Omega$. A *simple linear update* is defined by the following three rules: $\vec{x} := A\vec{x} + \vec{a}$, $\vec{v} := B\vec{v} + \vec{b}$, and $w := c$. Here, A is an $n \times n$ matrix with real coefficients, B is an $m \times m$ matrix with integer coefficients; $\vec{a}$, $\vec{b}$ and c are constants of the corresponding types. *Linear updates* are built from the simple updates by means of `do-in-parallel` blocks, conditionals and nondeterministic choice (see [1, 2]).

We use the idea of predicate abstraction to define the basic equivalence relation on the set of computations [3]. Consider a set of predicates $P_1, \ldots, P_r$ over the state space S. Then

$$s_1 \approx s_2 \text{ iff } (P_1(s_1) \equiv P_1(s_2)) \wedge \cdots \wedge (P_r(s_1) \equiv P_r(s_2)), \qquad s_1, s_2 \in S.$$

Consider now two computations $\mathbf{s}' = (s'_0, s'_1, \ldots, s'_l)$ and $\mathbf{s}'' = (s''_0, s''_1, \ldots, s''_l)$ of length l. We define them to be equal if for each step i the corresponding states are equivalent:

$$\mathbf{s}' \approx \mathbf{s}'' \text{ iff } \forall i \in \{0, 1, \ldots, l\}(s'_i \approx s''_i).$$

The maximal possible number of different equivalence classes is $(2^r)^l$, where r is the number of the predicates, l is the length of computations. This is not feasible even for small values of r and l.

In our work we consider other definitions of equivalence relations on the set of all computations, each of which is a weakening of the relation defined above. The main tasks here are the following. First, to find a reasonable definition of equivalence with rather small number of the equivalence classes (see [1]). Second, to get a representative for every equivalence class or to prove that the class is empty.

References

[1] V. Filatov and R. Yavorskiy. *Scenario based analysis of linear computations.* International workshop on Abstract State Machines, March 2005, pp.167–174.

[2] E. Börger and R. Stärk, **Abstract State Machines. A method for High-Level System Design and Analysis.** Springer-Verlag 2003.

[3] W. Grieskamp, Yu. Gurevich, W. Schulte, and M. Veanes. *Generating Finite State Machines from Abstract State Machines.* ACM Software Engineering Notes 27:4 (2002), 112-122.

[1] `E-mail: VolodyaFilatov@yandex.ru, rey@mi.ras.ru, Nikolay.Zemtsov@mtu-net.ru`

A note on fixed points of generalized ice pile models

E. Formenti* B. Masson*

Extended abstract

Ice Pile Models (IPM) were introduced in [1] to study the lattice of integers partitions. They are also a perused model for studying the behavior of systems governed by self-organised criticality (*SOC systems*). IPM can be seen as an extension of Sand Pile Models (SPM). SPM are a simple model based on a single local rule (*vertical rule*): a sand grain falls on its right if the difference its sandpile and the one on its right is bigger than a certain amount of grains. In the case of IPM, one more rule is added (*horizontal rule*): a grain slides to the right when there is a plateau.

The horizontal rule is not a local one. For this reason, in [4], the authors have introduced a variant for bounding the radius of the horizontal rule. They denote IPM(k), an IPM system in which a grain can slide to the right of at most k grains. This is the variant that we are going to discuss in the present paper.

Both SPM and IPM have been mathematically characterized(see [3] for a rather complete survey of known results). In particular, it has been proved that both systems have fixed point behavior and explicit formulas for the fixed points have been found. The issue is that these results are given only for a special initial configuration: all grains are concentrated in a single pile (column).

In [2], the authors exhibited a fast algorithm for computing fixed points of SPM under general initial conditions. They also claimed that it could be easily adapted to IPM. The present work aims at giving a formal proof to that claim.

References

[1] T. Brylawski. The lattice of integer partitions. *Discrete mathematics*, 6:201–219, 1973.

[2] E. Formenti and B. Masson. On computing fixed points for generalized sandpiles. *International Journal of Unconventional Computing*, 2005. To appear.

[3] E. Goles, M. Latapy, C. Magnien, M. Morvan, and H. D. Phan. Sandpile models and lattices: a comprehensive survey. *Theoretical Computer Science*, 322:383–407, 2004.

[4] E. Goles, M. Morvan, and H. D. Phan. Sand piles and order structure of integer partitions. *Discrete Applied Mathematics*, 117:51–64, 2002.

*Laboratoire I3S, Université de Nice-Sophia Antipolis,2000, route des lucioles, Sophia Antipolis Cedex, FRANCE. Email: {enrico.formenti,benoit.masson}@i3s.unice.fr

Clustering Data Streams - A Survey

Gereon Frahling, Christian Sohler
eccs@frahling.de

In modern computer society huge data sets occur in the form of data streams. The data traffic at an internet backbone router, the measurement of data (e.g. from satellite systems), and financial stock data are prominent examples of data streams. However, data stream algorithms are important as well in scenarios where random access to a huge data set is possible, but in fact too expensive. All huge data sets stored on hard disks or archive tapes can be accessed much faster sequentially.

The theoretical model of data streaming algorithms addresses problems of huge data sets consisting of n items which arrive in a sequential order. The computer has very limited local memory (the memory usage is bounded by a polynomial in $\log n$) and computation time for each element seen in the stream is limited (the computation time is again bounded by a polynomial in $\log n$). Since a data streaming algorithm is not able to store the whole data set, it must maintain statistics about the data seen so far. To analyse huge data sets it is often a good idea to group the vast number of items into a small set of clusters. In a second step an algorithm can analyse the clusters in local memory. The method of clustering can furthermore be a good tool to reduce the complexity of a data set and make it ascertainable for human beings.

In this survey we will recapitulate several theoretical results for clustering data streams. To cluster data one first has to provide a distance function between the items within the data stream. All clustering algorithms in this paper assume that this distance function is a metric. Some algorithms furthermore assume that the items are assigned to a point within a Euclidean space, and the distance between items is measured by the distance within this Euclidean space.

Having a distance function we have to decide which properties the clustering should have. The k-Median problem asks for k centers and an assignment of each item to a center, such that the sum of distances of each item to it's nearest center is minimized. In the k-Means problem the sum of the squares of these distances is minimized. The k-Center problem asks for a similar assignment to k centers, but the goal is to minimize the largest assignment distance. The Facility Location Problem is a relaxation to the k-Median problem, where the number of centers is unrestricted, but there is an additional constant cost for each center.

We will summarize recent theoretic results on how to obtain approximate clusterings, when the items are given in a data stream.

A model for the genetic code

L Frappat, A Sciarrino, Paul Sorba
frappat@lapp.in2p3.fr

Among the numerous and important questions offered to the physicist by the sciences of life, the ones relative to the genetic code present a particular interest. An attempt for modelizing the genetic code can be seen as a step further in the understanding of the complexity of the genetic code processes.

The heart of the genetic code machinery is the double helix shaped DNA macromolecule, constituted by two linear chains of four types of nucleotides characterized by their bases: adenine (A) and guanine (G) deriving from purine, and cytosine (C) and thymine (T) coming from pyrimidine. The genetic information is transmitted via the messenger ribonucleic acid or mRNA through the transcription in which the A, G, C, T bases in the DNA are associated respectively to the U, C, G, A bases, U denoting the uracile base. The translation process then associates to an ordered triplet of nucleotides (or codon) an amino-acid. Following the universal eukariotic code, 61 of such codons can be connected in an unambiguous way to the twenty amino-acids (except the three codons UAA, UAG and UGA, the role of which is to stop the biosynthesis), showing a degeneracy of the genetic code, the codons being arranged into sextets, quadruplets, triplets, doublets and singlets, each multiplet corresponding to a specific amino-acid (note that degeneracy is primarily found in the third position of the codon). Such a picture suggests to look for an underlying symmetry able to describe the observed structure.

The starting point of the mathematical framework we have developed is to consider the four nucleotides as basic states of a certain representation of some algebra, and to construct a codon as a state in the tensor product of three such representations. The complementarity rule in the DNA–mRNA transcription may suggest to assign a "quantum number" with opposite values to the couples (A,T/U) and (C,G). The distinction between the purine bases (A,G) and the pyrimidine ones (C,T/U) can be algebraically represented in an analogous way. Morevover, the fact that a codon is an ordered triplet of nucleotides imposes to construct the composite states as pure states (i.e. that cannot be linear combinations of other states). Such a structure is provided by using quantum enveloping algebras $\mathcal{U}_q(\mathcal{G})$, in the limit $q \to 0$, called crystal basis. One is then naturally led to consider the fundamental representation $(1/2, 1/2)$ of the quantum enveloping algebra $\mathcal{U}_{q\to 0}(sl(2) \oplus sl(2))$. Note that we do not deal anymore with a Lie algebra (which are the structure usually emerging in physics).

A "charge" Q of a dinucleotide state can be defined: the dinucleotide states are split into two octets with respect to the charge Q: the eight strong dinucleotides associated to the quadruplet or sextets of codons satisfy $Q > 0$, while the eight weak dinucleotides associated to the doublets or singlets of codons satisfy $Q < 0$. The model does not gather codons associated to one particular amino-acid in the same irreducible multiplet. However, it is possible to construct an operator $\mathcal{R}$ out of the algebra, acting on the codons, that will describe the various genetic codes: two codons have the same eigenvalue under $\mathcal{R}$ if and only if they are associated to the same amino-acid. Remarkably, the various genetic codes share the same basic structure.

It is a nearly universal phenomenon that alternative synonymous codons inside a multiplet are generally not used with equal frequency (codon usage). Considering branching ratios of probabilities of usage (i.e. frequencies of usage in the limit of very large number of codons) for codons differing only by their third nucleotide (XYZ and XYZ$'$, Z $\neq$ Z$'$), correlations has been remarked for biological species belonging to the vertebrate class when the dinucleotides XY correspond to quartets. One may wonder why the model is

able to explain the observed correlations. A possible answer is that the model incorporates information which is typical of the organisation of the code, and that such information is in itself relevant to determine the correlations. One can also deduce, in the crystal basis model, that the sum of usage probabilities of the codons with C and A in the third position for the quartets and/or sextets is independent of the biological species for the vertebrate class ("sum rules"). A comparison between the theoretical predictions and the experimental data shows a satisfactory agreement.

Finally, a set of relations between the physical-chemical properties of the amino-acids, based on the content of the irreducible representations of the dinucleotides and of the codons coding for the amino-acids, can be derived and compared with the experimental data (of course, not all the physical-chemical properties are supposed to follow the scheme above; some of them are essentially given by the specific chemical structure of the amino acid itself).

Genetic Self-Assembly: Many Simple or a Few Complex?

Rudolf M. Füchslin, Thomas Maeke, Uwe Tangen, John S. McCaskill
rudolf.fuechslin@biomip.ruhr-uni-bochum.de

Keywords: Self-assembly, inductive generalization, evolvable logic, circuit design, genetic algorithm, evolution, multiplier

In a recent paper, we investigated a genetically encoded system of a small number of different types of self-assembling components and showed that, if these components were equipped with type-dependent logical functions, scalable circuitry for several problems can be evolved. The evolution of scalable circuitry requires inductive generalization, which means in this context the inductive derivation of a solution to a problem with a potentially infinite number of instances from a limited set of test examples. To effectuate this by the mechanisms of evolution is known to be difficult; one of the investigated problems, digital multiplication, has been intensively studied in recent years, where, until employing self-assembling components, only the evolutionary design of specific small multipliers has been achieved.

This work investigates the relation between the evolutionary efficiency and the number of different types of self-assembling components and the complexity of the logic they carry respectively. We address the question for the optimal size and complexity for an evolving self-assembling system with respect to different problems. This sheds light on the evolutionary role of self-assembly in biology and is of relevance for the design of complex systems in nano- and bionanotechnology.

A way to characterize complex cellular automata and those able to perform density classification

Gabriele Anna Rosa, Gervasi Stefania
University of Calabria, Italy

In this work we consider particular mathematical models useful in studying complexity: the cellular automata (CA) [1]. They are constituted by a great number of connected entities which interact at a local level, showing, at a global level, emergent behaviors and self-organization. CA were introduced by von Neumann and Ulam as simple models for the study of some biological processes. CA have been used as models for systems like bugs' colonies, immune systems, economic systems, etc., because they show the emergent computation and complex behavior typical of these systems.
A CA is a discreet dynamical system in which the space is represented by a n-dimensional regular grid. Each element of the grid (called cell) can be considered as the basic element of CA and it contains a datum a_i^t that represents the state of the i^{th} cell at time t. A CA configuration (c_t) is the set of the states in which all the cells are at a particular time t. The state of each cell evolves following a rule that depends both on the state of the cell itself and on the state of a certain number of cells that compose the neighborhood of the cell. The behavior or the evolution of a CA, starting from an assigned initial configuration, is the series of configurations $\{c_t\}_{t\geq 0}$. CA can show different types of behavior: periodic, chaotic and complex. The most interesting one is the complex behavior, in which self-organized complex structures interacting with each other emerge in the evolution. The problem of finding this kind of CA among all the possible rules, is not so simple. Once the number of states (k) and the number of cells that compose the neighborhood (m) are fixed, the number of all the possible rules is k^{k^m}. For one-dimensional CA with $k=2$ and $m=3$ (elementary CA) it is possible to do an exhaustive examination of all the rules. Nevertheless, by increasing k and m, the high power of the rules' space make this examination difficult to be undertaken. For this reason, the use of genetic algorithms (GA) can be viewed as an interesting alternative approach. Various experimentations in this direction have been done [2], [3], [4].
The genetic algorithms, introduced by John Holland [5], are inspired by the natural evolution. They are used a lot in the context of evolutionary computation, particularly for the optimization and the automatic learning. The main idea is to evolve a population of candidate solutions for a particular problem using some operators (mutations and cross over) that are inspired to genetics and natural selection.
In this work the genetic algorithms have been used for exploring the rules' space of elementary CA and multi-states CA ($k>2$) with the aim to characterize, by different fitness functions, the rules that show complex behaviors and the rules able to perform density classification [3], [4].
The density classification problem consists in determining some rules that evolve towards an homogeneous final configuration constituted by the state with the higher concentration in the initial configuration.
Rules have been evolved using GA. The input entropy, able to select complex rules [2], was used as fitness function in some experiments. In other experiments, another fitness function [4] was specifically chosen in order to select rules which perform density classification. The evolved rules have been subsequently evaluated by computing, for each rule, some parameters taken from the literature [6] by broadening their applicability to multi-state CA.
Finally, some considerations have been drawn about the relationships between the values of the parameters described above and the specific characteristics of the analyzed rules.

References

[1] Wolfram S: Universality and complexity in cellular automata. Physica D 1984; 10: 1-35.
[2] Bilotta E, Lafusa A, Pantano P: Searching for complex CA rules with GAs. Complexity 2003;8(3):56-57.
[3] Packard NH: Adaptation toward the edge of chaos; in Kelso JAS, Mandell AJ, and Shlesinger MF (eds): Dynamic Patterns in Complex Systems. World Scientific, 1988, pp 293-301.
[4] Mitchell M, Crutchfield JP and Hraber PT: Evolving Cellular Automata to Perform Computations: Mechanisms and Impediments. Physica D 1994; 75: 361-391.
[5] Holland JH: Adaptation in Natural and Artificial Systems, ed 2, MIT Press, 1992.
[6] Oliveira GMB, de Oliveira PPB, Omar N: Definition and Application of a Five-Parameter Characterization of One-Dimensional Cellular Automata Rule Space. Artificial Life 2001; 7(3): 277-301.

A way to characterize complex cellular automata and those able to perform density classification

Gabriele Anna Rosa, Gervasi Stefania
University of Calabria, Italy

In this work we consider particular mathematical models useful in studying complexity: the cellular automata (CA) [1]. They are constituted by a great number of connected entities which interact at a local level, showing, at a global level, emergent behaviors and self-organization. CA were introduced by von Neumann and Ulam as simple models for the study of some biological processes. CA have been used as models for systems like bugs' colonies, immune systems, economic systems, etc., because they show the emergent computation and complex behavior typical of these systems.
A CA is a discreet dynamical system in which the space is represented by a n-dimensional regular grid. Each element of the grid (called cell) can be considered as the basic element of CA and it contains a datum a_i^t that represents the state of the i^{th} cell at time t. A CA configuration (c_t) is the set of the states in which all the cells are at a particular time t. The state of each cell evolves following a rule that depends both on the state of the cell itself and on the state of a certain number of cells that compose the neighborhood of the cell. The behavior or the evolution of a CA, starting from an assigned initial configuration, is the series of configurations $\{c_t\}_{t\geq 0}$. CA can show different types of behavior: periodic, chaotic and complex. The most interesting one is the complex behavior, in which self-organized complex structures interacting with each other emerge in the evolution. The problem of finding this kind of CA among all the possible rules, is not so simple. Once the number of states (k) and the number of cells that compose the neighborhood (m) are fixed, the number of all the possible rules is k^{k^m}. For one-dimensional CA with $k=2$ and $m=3$ (elementary CA) it is possible to do an exhaustive examination of all the rules. Nevertheless, by increasing k and m, the high power of the rules' space make this examination difficult to be undertaken. For this reason, the use of genetic algorithms (GA) can be viewed as an interesting alternative approach. Various experimentations in this direction have been done [2], [3], [4].
The genetic algorithms, introduced by John Holland [5], are inspired by the natural evolution. They are used a lot in the context of evolutionary computation, particularly for the optimization and the automatic learning. The main idea is to evolve a population of candidate solutions for a particular problem using some operators (mutations and cross over) that are inspired to genetics and natural selection.
In this work the genetic algorithms have been used for exploring the rules' space of elementary CA and multi-states CA ($k>2$) with the aim to characterize, by different fitness functions, the rules that show complex behaviors and the rules able to perform density classification [3], [4].
The density classification problem consists in determining some rules that evolve towards an homogeneous final configuration constituted by the state with the higher concentration in the initial configuration.
Rules have been evolved using GA. The input entropy, able to select complex rules [2], was used as fitness function in some experiments. In other experiments, another fitness function [4] was specifically chosen in order to select rules which perform density classification. The evolved rules have been subsequently evaluated by computing, for each rule, some parameters taken from the literature [6] by broadening their applicability to multi-state CA.
Finally, some considerations have been drawn about the relationships between the values of the parameters described above and the specific characteristics of the analyzed rules.

References

[1] Wolfram S: Universality and complexity in cellular automata. Physica D 1984; 10: 1-35.
[2] Bilotta E, Lafusa A, Pantano P: Searching for complex CA rules with GAs. Complexity 2003;8(3):56-57.
[3] Packard NH: Adaptation toward the edge of chaos; in Kelso JAS, Mandell AJ, and Shlesinger MF (eds): Dynamic Patterns in Complex Systems. World Scientific, 1988, pp 293-301.
[4] Mitchell M, Crutchfield JP and Hraber PT: Evolving Cellular Automata to Perform Computations: Mechanisms and Impediments. Physica D 1994; 75: 361-391.
[5] Holland JH: Adaptation in Natural and Artificial Systems, ed 2, MIT Press, 1992.
[6] Oliveira GMB, de Oliveira PPB, Omar N: Definition and Application of a Five-Parameter Characterization of One-Dimensional Cellular Automata Rule Space. Artificial Life 2001; 7(3): 277-301.

Probing the robustness of the clustering

David Gfeller[1], Jean-Cédric Chappelier[2], and Paolo De Los Rios[1]

[1] *Laboratoire de Biophysique Statistique, SB/ITP and*

[2] *School of Computer and Communication Sciences,*

Ecole Polytechnique Fédérale de Lausanne, CH-1015, Lausanne, Switzerland

(Dated: May 4, 2005)

Keywords: analysis of complex networks, clustering

Many complex systems can be represented as complex networks. Nevertheless the very large size of such systems often does not allow any possible mental or graphical representation. It has been pointed out that a complex network consists in many cases of several densely interconnected clusters, with only a few connections to the rest of the network. The identification of these clusters is therefore a key point to understand the relevant organization of the network and to reduce the size and complexity of the system. Here we introduce a new method to probe the robustness of the clusters. This is particularly important since often clustering algorithms produce a hard clustering even for the nodes that "lie between" clusters, whose classification is very arguable.

Our method is based on the introduction of noise over the weight of the edges. After running the clustering algorithm for several realizations of the noise, we can assess a probability on the edges of connecting two nodes belonging to the same cluster. This probability represents how much the grouping of two nodes in a cluster is reasonable. For example nodes connected together by edges with probability close to 1 can be assigned to the same cluster with high confidence. On the other hand, nodes connected with a probability close to 0.5 are rather *unstable* with respect to the cluster structure.

Furthermore we introduce a general measure of the robustness of the clusters, defined as the *clustering entropy*. It allows to discriminate networks with well-defined clusters from networks that, although topologically very similar, do not have a significant cluster structure.

We tested our method on several real world examples. In a linguistic network based on the synonymy relation, the unstable nodes corresponded to ambiguous words. In sociological networks they represented individuals connected to different groups. In all those cases, the clustering entropy of the real network was much lower than the one of a randomized network where the degree of the nodes had been conserved. We finally note that the method can be applied with any clustering algorithm, provided that the algorithm allows weights on the edges.

Robust Cooperation in Multi-Agent Systems

Robert Ghanea-Hercock, robert.ghanea-hercock@bt.com

Abstract

In large-scale multi-agent systems (MAS) the ability of agents to form coalitions of trusted partners and reputation networks is vital to autonomous operation. This paper investigates the dynamic formation of cooperative communities within a simulated MAS. In particular this work considers the requirements for stable high-trust coalitions to self-organise and survive, while some percentage of agents in the population is defecting. A model of dynamic group formation is presented which enables the rapid formation of a self-organising cooperative agent community. The results presented indicate that by utilizing a self-reinforcing cooperative trust model, a very high degree of resilience to perturbation and defection can be achieved. A number of critical parameters are investigated which indicate a phase transition in the formation and resilience of the cooperative structures that emerge.

Keywords: Autonomous Agents, Trust, Adaptive Behaviour

1 Introduction

The aim of this paper is firstly to investigate the possible strategies a MAS community may adopt in order to build a community of cooperating agents with a high degree of communal trust. A significant body of work has also emerged in agent based network authentication systems [Helmer et al. 1998], and reputation development between agent brokers in e-commerce systems [Braynov & Sandholm 1999].

1.1 The Problem

While agent reputation systems [Abdul-Rahman & Hailes 2000], and [Yu & Singh 2002] have been extensively reviewed there is still a lack of understanding in terms of how processes of trust and reputation evolve dynamically over time. I This work therefore follows the concept of social processes of trust construction as set out by Yu and Singh [2003]

2 Dynamics of Trust

An agent's interaction in any complex environment requires a continuous reassessment of the degree of trust it should assign to external agents and events. Agents shift their probability of cooperation or defection based on the *expected behaviour* of the majority of its neighbours, i.e. if the majority of neighbours play defect then each agent will increase the probability that it defects, and the same for cooperation.

2.1 Experimental Model

The REPAST agent simulation platform [http://repast.sourceforge.net] was used to simulate a number of agents.

3 Results

The key result obtained from this method of cooperative social group formation is the degree of resilience and resistance to defection achieved. A series of experiments were then applied to the set of agents after such clusters had emerged in order to determine the degree to which they could resist defection by internal and external agents.

4 Conclusions

The formation of cooperating agent communities is a fundamental requirement for the large-scale deployment of MAS. In this work a flexible approach is adopted that accommodates dynamic variation in the degree of trust allocated to an agent.

The key behaviour of the agents to enable them to form cooperative groups is that they shift their probability of cooperation or defection based on the *expected behaviour* of the majority of its neighbours, i.e. if the majority of neighbours play defect then each agent will increase the probability that it defects, and the same for cooperation.

The result is that a group of agents using such a dynamic trust model becomes highly resistant to defection by external or internal agents, once a sufficiently large cooperative group has formed. However, below the critical group size the agents are susceptible to defection strategies, which then dominate the system

What Models for Complex Systems?

Sica Giandomenico
g.sica@polimetrica.org

Following Bar-Yam ["Overview: The Dynamics of Complex Systems"], one approach to the study of complex systems begins from an understanding of the relationship of systems to their descriptions. In this context, our enquiry focuses on the construction of rigorous easy-to-use models to describe complex systems. In particular, we consider a model as defined by a mathematical structure, where the expression "mathematical structure" aims to indicate a combination of different kinds of sets through an arbitrary order. The concept of "set" and the modalities to construct sets are defined through the following axioms and corollary ("Elementary Set Theory"): Axiom I (extensionality): A set is completely determined by its elements and by the composition relations among the same elements. Element of a set is always a set. The composition relations are not commutative. Axiom II (construction): It is possible to construct a set through the execution of the algorithm $<$ Df1, Df2, Df3$>$, where Df1 represents the operation of elements definition, Df2 represents the operation of composition relations definition, Df3 represents the operation defining the membership relation among elements and set. Corollary: Given two sets A and B between which there exists a composition relation r, it's always possible to determine a set C, having A and B as elements, ordered through the composition relation r. Having defined the mathematical structure, in order to complete our task, we have to call the different sets and composition relations (namely, create a label for the variety of sets and composition relations): as result, we obtain a model, mathematically exact and linguistically precise. Our goal is to introduce this technique in order to construct models able to describe the different typologies of systems (e.g. biological, physical, economical, social ones) keeping intact, in the complexity, the precision degree.

What Models for Complex Systems?

Sica Giandomenico
g.sica@polimetrica.org

Following Bar-Yam ["Overview: The Dynamics of Complex Systems"], one approach to the study of complex systems begins from an understanding of the relationship of systems to their descriptions. In this context, our enquiry focuses on the construction of rigorous easy-to-use models to describe complex systems. In particular, we consider a model as defined by a mathematical structure, where the expression "mathematical structure" aims to indicate a combination of different kinds of sets through an arbitrary order. The concept of "set" and the modalities to construct sets are defined through the following axioms and corollary ("Elementary Set Theory"): Axiom I (extensionality): A set is completely determined by its elements and by the composition relations among the same elements. Element of a set is always a set. The composition relations are not commutative. Axiom II (construction): It is possible to construct a set through the execution of the algorithm $<$ Df1, Df2, Df3$>$, where Df1 represents the operation of elements definition, Df2 represents the operation of composition relations definition, Df3 represents the operation defining the membership relation among elements and set. Corollary: Given two sets A and B between which there exists a composition relation r, it's always possible to determine a set C, having A and B as elements, ordered through the composition relation r. Having defined the mathematical structure, in order to complete our task, we have to call the different sets and composition relations (namely, create a label for the variety of sets and composition relations): as result, we obtain a model, mathematically exact and linguistically precise. Our goal is to introduce this technique in order to construct models able to describe the different typologies of systems (e.g. biological, physical, economical, social ones) keeping intact, in the complexity, the precision degree.

What Proteins Are Made From? Informational Way To Protein Alphabet

Alexander Gorban[*1], Mikhail Kudryashev[2], and Tatiana Popova[3]

[1]*Centre for Mathematical Modeling, University of Leicester, UK*
[2]*Krasnoyarsk State University, Krasnoyarsk, Russia*
[3]*Institute of Computational Modeling, Akademgorodok, Krasnoyarsk, Russia*

Keywords. Amino acid, protein, classification, relative entropy, alphabet reduction

What proteins are made from, as the working parts of the living cells protein machines? To answer this question, we need a technology to disassemble proteins onto elementary functional details and to prepare lumped description of such details. Our hypothesis is that informational approach to this problem is possible. We propose a way of hierarchical classification of amino acids that makes the primary structure of protein *maximally non-random*. In order to formalize this idea we follow [1,2] and analyse *frequency dictionaries* of short protein fragments, and *relative entropies* of such frequency dictionaries. The *entropic optimality principle* is formulated and applied for amino acids classifications for various databases of primary protein sequences. In contrary to the widespread *MaxEnt* approach (that is, of maximal disorder), we deal with the principle of maximal order. The following properties of amino acids binary informational classifications are studied 1) the *existence and uniqueness* of optimal classification for given frequency dictionary, 2) structure of the set of highly informative classifications in the vicinity of the optimal one, 3) *stability/instability* of the optimal classification with respect to variations of the frequency dictionary, 4) similarity between classifications constructed on the basis of 2-letter words frequencies and those constructed on the basis 3, 4 and 5-letter words frequencies. We compared informational binary classifications of amino acids with classifications obtained by other methods. Amino acids groupings mentioned in most of reviewed papers do have moderate similarity with two types of Hydrophobic/Polar classification while informational classifications is shifted to Charged/Uncharged property. Classification of [3] is the only one to be rather close to informational classifications. Detailed statistic data are published in preprint [2]. The binary informational classification gives us "the first letter of protein alphabet". Algorithms of *hierarchical information classification* are developed in order to find the following "letters". On each level of hierarchy we find the optimal classification that is independent (with given accuracy) of classifications obtained on the previous levels. The second level is surprisingly independent of all known "usual" amino acids classifications. Below the example of the first and the second classifications is presented for E-coli proteome (number of proteins is 5797 [4]):

1st classification: 1st class: A,C,D,E,F,G,I,K,L,M,N,P,S,T,V; 2nd class: H,Q,R,W,Y.
2nd classification 1st class: A,E,K,L,M,P,Q,R,T,V,W; 2nd class: C,D,F,G,H,I,N,S,Y.

Binary classification tree: Zero level (one class) {A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y}
First level (two classes) {A,C,D,E,F,G,I,K,L,M,N,P,S,T,V}, {H,Q,R,W,Y}
Third level (four classes) {A,E,K,L,M,P,T,V}, {C,D,F,G,I,N,S}, {Q,R,W}, {H,Y}.

The next step of the research program should be the informational classifications analysis of 2 and 3-symbol elements of primary structures presented as a sequence of 2 and 3-letters elements and so on.

References

1. Bugaenko N.N., Gorban A.N., Sapozhnikov A.N. Minimum entropy principle and classification of symbols for revealing statistical regularities in text. In: *Proc. of the first Int. Conf. on Bioinformatics of Genome Regulation and Structure*, 1998; Vol. 2: 280-282; http://www.bionet.nsc.ru/bgrs/author/gorban.html .
2. Gorban A.N., Kudryashev M., Popova T., On the Way to Protein Alphabet: Informational Classification of Amino Acids in Comparison to Other Classifications, e-print: http://arxiv.org/abs/q-bio.BM/0501019 .
3. Murphy LR, Wallqvist A, Levy RM. Simplified amino acid alphabets for protein fold recognition and implications for folding. *Protein Eng.* **2000**, 13(3), 149-52.
4. *Swiss Institute of Bioinformatics website*, Expert Protein Analysis System http://expasy.org/

[*] *University Road, Leicester, LE1 7RH, UK, ag153@le.ac.uk*

What Proteins Are Made From? Informational Way To Protein Alphabet

Alexander Gorban[*1], Mikhail Kudryashev[2], and Tatiana Popova[3]

[1]*Centre for Mathematical Modeling, University of Leicester, UK*
[2]*Krasnoyarsk State University, Krasnoyarsk, Russia*
[3]*Institute of Computational Modeling, Akademgorodok, Krasnoyarsk, Russia*

Keywords. Amino acid, protein, classification, relative entropy, alphabet reduction

What proteins are made from, as the working parts of the living cells protein machines? To answer this question, we need a technology to disassemble proteins onto elementary functional details and to prepare lumped description of such details. Our hypothesis is that informational approach to this problem is possible. We propose a way of hierarchical classification of amino acids that makes the primary structure of protein *maximally non-random*. In order to formalize this idea we follow [1,2] and analyse *frequency dictionaries* of short protein fragments, and *relative entropies* of such frequency dictionaries. The *entropic optimality principle* is formulated and applied for amino acids classifications for various databases of primary protein sequences. In contrary to the widespread *MaxEnt* approach (that is, of maximal disorder), we deal with the principle of maximal order. The following properties of amino acids binary informational classifications are studied 1) the *existence and uniqueness* of optimal classification for given frequency dictionary, 2) structure of the set of highly informative classifications in the vicinity of the optimal one, 3) *stability/instability* of the optimal classification with respect to variations of the frequency dictionary, 4) similarity between classifications constructed on the basis of 2-letter words frequencies and those constructed on the basis 3, 4 and 5-letter words frequencies. We compared informational binary classifications of amino acids with classifications obtained by other methods. Amino acids groupings mentioned in most of reviewed papers do have moderate similarity with two types of Hydrophobic/Polar classification while informational classifications is shifted to Charged/Uncharged property. Classification of [3] is the only one to be rather close to informational classifications. Detailed statistic data are published in preprint [2]. The binary informational classification gives us "the first letter of protein alphabet". Algorithms of *hierarchical information classification* are developed in order to find the following "letters". On each level of hierarchy we find the optimal classification that is independent (with given accuracy) of classifications obtained on the previous levels. The second level is surprisingly independent of all known "usual" amino acids classifications. Below the example of the first and the second classifications is presented for E-coli proteome (number of proteins is 5797 [4]):

1st classification: 1st class: A,C,D,E,F,G,I,K,L,M,N,P,S,T,V; 2nd class: H,Q,R,W,Y.
2nd classification 1st class: A,E,K,L,M,P,Q,R,T,V,W; 2nd class: C,D,F,G,H,I,N,S,Y.

Binary classification tree: Zero level (one class) {A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y}
First level (two classes) {A,C,D,E,F,G,I,K,L,M,N,P,S,T,V}, {H,Q,R,W,Y}
Third level (four classes) {A,E,K,L,M,P,T,V}, {C,D,F,G,I,N,S}, {Q,R,W}, {H,Y}.

The next step of the research program should be the informational classifications analysis of 2 and 3-symbol elements of primary structures presented as a sequence of 2 and 3-letters elements and so on.

References

1. Bugaenko N.N., Gorban A.N., Sapozhnikov A.N. Minimum entropy principle and classification of symbols for revealing statistical regularities in text. In: *Proc. of the first Int. Conf. on Bioinformatics of Genome Regulation and Structure*, 1998; Vol. 2: 280-282; http://www.bionet.nsc.ru/bgrs/author/gorban.html .
2. Gorban A.N., Kudryashev M., Popova T., On the Way to Protein Alphabet: Informational Classification of Amino Acids in Comparison to Other Classifications, e-print: http://arxiv.org/abs/q-bio.BM/0501019 .
3. Murphy LR, Wallqvist A, Levy RM. Simplified amino acid alphabets for protein fold recognition and implications for folding. *Protein Eng.* **2000**, 13(3), 149-52.
4. *Swiss Institute of Bioinformatics website*, Expert Protein Analysis System http://expasy.org/

[*] *University Road, Leicester, LE1 7RH, UK, ag153@le.ac.uk*

MODELING TUMOR GROWTH AS THE EVOLUTION OF A BIOLOGICAL COMPLEX SYSTEM WITH VARIABLE FRACTAL DIMENSIONS .

Caterina Guiot[†1,2,*], Pier Paolo Delsanto[2,3,*], Nicola Pugno[4]

and Thomas S. Deisboeck[5,*]

Affiliations: [1]Dip. Neuroscience, Università di Torino, Italy; [2]INFM, sezioni di Torino Università e Politecnico, Italy; [3]Dip. Fisica, Politecnico di Torino, Italy; ; [4]Dip. Ing. Strutturale e Geotecnica, Politecnico di Torino, Italy ; [5]Complex Biosystems Modeling Laboratory, Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA 02129, USA.

* cVit supported by the NIH-National Cancer Institute through the Integrative Cancer Biology Program (CA113004), the Division of Engineering and Applied Sciences at Harvard University and IBM Corporation.

[†]**Corresponding Author:**

Caterina Guiot, Ph.D., Dip. Neuroscienze, 30, C. Raffaello, 10125 Torino

tel: +39.11.670.7710, fax: +39.11.670.7708, e-mail: caterina.guiot@unito.it

Universal growth laws, as proposed by West and collaborators for all living organisms (West et al, *Nature* **413**: 628-631,2001), may be extended to describe the growth of tumors *in vivo, provided the* scaling exponent *p* (assumed by West et al. to be equal to 3/4) varies according to the tumor and its vascular network evolution (Guiot et al, *J. Theor. Biol.* **225**: 147-283, 2003, Delsanto et al, *Appl Phys Lett*. **85**: 4225-4227,2004).

In fact, the idea of a fractal topology of the tumor vasculature has already been proposed by Baish (*Microvasc. Res.* **51**: 327-346,1996) and Baish & Jain (*Cancer Res.* **60**:3683-3688, 2000), and in-vivo estimates of the fractal dimensions of planar vascular networks based on the box-counting method were performed. The evolution of cancer topology during the growth of tumors implanted in mice has been studied by Gazit et al. (*Microcirculation* **4**:395-402,1997) ,in the chick embryo

by Vico et al (*J. Theor. Biol.* **195**:525-532,1998) and after delivering docetaxel to cultured HUVEC cells in Matrigel by Guidolin et al (*Microvasc. Res.*, **67**: 117-124, 2003) .
In a completely different context, Carpinteri and Pugno (*J. Appl. Mech.* **69**: 854-856, 2002) have developed universal scaling laws for energy dissipation during the fragmentation of solids, by assuming a self-similar size distribution of fragments. It is interesting to note that, according to an interpretation based on their analysis, the exponent *p* should be strongly related to the fractal nature of cancer topology and thus susceptible of independent measurements. It could possibly also be applied for diagnostic purposes to mark the emergence of a functionally effective neo-angiogenetic structure.

MODELING TUMOR GROWTH AS THE EVOLUTION OF A BIOLOGICAL COMPLEX SYSTEM WITH VARIABLE FRACTAL DIMENSIONS .

Caterina Guiot[†1,2,*], Pier Paolo Delsanto[2,3,*], Nicola Pugno[4]

and Thomas S. Deisboeck[5,*]

Affiliations: [1]Dip. Neuroscience, Università di Torino, Italy; [2]INFM, sezioni di Torino Università e Politecnico, Italy; [3]Dip. Fisica, Politecnico di Torino, Italy; ; [4]Dip. Ing. Strutturale e Geotecnica, Politecnico di Torino, Italy ; [5]Complex Biosystems Modeling Laboratory, Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA 02129, USA.

* cVit supported by the NIH-National Cancer Institute through the Integrative Cancer Biology Program (CA113004), the Division of Engineering and Applied Sciences at Harvard University and IBM Corporation.

[†]**Corresponding Author:**

Caterina Guiot, Ph.D., Dip. Neuroscienze, 30, C. Raffaello, 10125 Torino

tel: +39.11.670.7710, fax: +39.11.670.7708, e-mail: caterina.guiot@unito.it

Universal growth laws, as proposed by West and collaborators for all living organisms (West et al, *Nature* **413**: 628-631,2001), may be extended to describe the growth of tumors *in vivo, provided the* scaling exponent *p* (assumed by West et al. to be equal to 3/4) varies according to the tumor and its vascular network evolution (Guiot et al, *J. Theor. Biol.* **225**: 147-283, 2003, Delsanto et al, *Appl Phys Lett*. **85**: 4225-4227,2004).

In fact, the idea of a fractal topology of the tumor vasculature has already been proposed by Baish (*Microvasc. Res.* **51**: 327-346,1996) and Baish & Jain (*Cancer Res.* **60**:3683-3688, 2000), and in-vivo estimates of the fractal dimensions of planar vascular networks based on the box-counting method were performed. The evolution of cancer topology during the growth of tumors implanted in mice has been studied by Gazit et al. (*Microcirculation* **4**:395-402,1997) ,in the chick embryo

by Vico et al (*J. Theor. Biol.* **195**:525-532,1998) and after delivering docetaxel to cultured HUVEC cells in Matrigel by Guidolin et al (*Microvasc. Res.*, **67**: 117-124, 2003) .

In a completely different context, Carpinteri and Pugno (*J. Appl. Mech.* **69**: 854-856, 2002) have developed universal scaling laws for energy dissipation during the fragmentation of solids, by assuming a self-similar size distribution of fragments. It is interesting to note that, according to an interpretation based on their analysis, the exponent *p* should be strongly related to the fractal nature of cancer topology and thus susceptible of independent measurements. It could possibly also be applied for diagnostic purposes to mark the emergence of a functionally effective neo-angiogenetic structure.

Vasomotion in arteriolar networks

Martin H. Kroll
martin.kroll@med.va.gov

The microvasculature forms an interconnecting network of vessels that supplies nutrients to the tissues in a timely manner and removes waste products. It serves as a conduit for blood and provides peripheral vascular resistance (necessary to maintain adequate blood pressure). The arterioles react to the accumulation of waste products from the tissues, dilating and constricting, resulting in rhythmic patterns of blood flow known as vasomotion. The intricate geometric structure, biochemical negative feedback, flexible adaptation to changing demands, and distribution of multiple units demonstrating synchronization characterize the microvasculature as a complex system. Alteration of structure induces dynamic changes, which may become altered by disease. We studied the inherent rhythm in blood flow in the microvasculature using spectral analysis. Blood flow was measured in the skin of the forehead using the Periflux PF3 Laser Doppler flow meter (Perimed, Sweden) in male and female subjects between the ages of 20 and 60. The data was collected at 8 Hz. Spectral analysis was performed by means of fast Fourier transforms using Matlab v. 6.5. The raw data (blood flow vs. time) demonstrated rare sinusoidal behavior. Spectral analysis (power spectra) revealed multiple frequencies. Distinct peaks occurred between 0 and 0.2 Hz, with no peaks above 0.4 Hz. All studies showed a fundamental frequency of 0.000977 Hz (a period of 17 minutes). Each subject demonstrated numerous frequencies between the fundamental frequency and 0.16 Hz. Division of the higher frequencies by the fundamental frequency resulted in unique series of integers, such as 1,3,6,12,15,18,22,101; 1,6,11,13,18,23,26,28,101; 1,5,9,15,18,21,23,28,30,101; 1,4,8,13,16,18,22,28,31,101. The original data set could be reconstructed by the summation of a series of sine waves, indicating a highly ordered deterministic system. Spectra from subjects with diabetes or hypertension demonstrate a decreased number of frequencies below 0.1 Hz and an increased power of those between 0.11 and 0.17 Hz. Combining architectural structure, nonlinear response, time-delay and the dissipative use of energy with the presence of a complex periodic system (the frequencies are commensurate with the fundamental frequency) demonstrated by the spectra analysis defines the microvasculature as a nonlinear, dissipative system with limit cycles providing the source for the periodic behavior. Only a few elements in the series of frequencies represent a harmonic series. The non-harmonic components offer a means to study highly complex systems.

Analysis of the vasomotion spectra reveals that each person produces distinct sharp frequencies. Disease causes distinctive alteration in pattern with a decrease in the number of dominant frequencies and a preponderance of frequencies between 0.11 and 0.17 Hz. Such a change in pattern may be the result of a simplification of the relationships among the component variables of the system. Spectral analysis of blood flow in the microvasculature offers a unique tool in understanding the complex behavior of blood delivery at the tissue level and may be useful in assessing the impact of disease prior to end-organ tissue damage.
keywords: analysis of complex networks, microvasulature, multiple limit cycle behavior

Adaptive rational modelling of complex systems

Wouter Hendrickx, Tom Dhaene
wouter.hendrickx@ua.ac.be

Complex computer simulations take a very long time to compute. In many engineering applications such simulations depend on several tuning parameters. Abstractly we therefore see a simulator as an unknown system

$$H : R^d \rightarrow R^p$$

which maps the tuning parameters onto certain output values of the simulation. Frequently one wants to have an overview of the outputs of the simulator as a function of the inputs. To run the simulator at a huge number of input points would be too expensive. Therefore, we try to use adaptive rational modelling techniques in order to minimise the number of simulations needed to get an accurate approximation of the system's global output behaviour.

We have developed a MATLAB framework for adaptively building rational models for systems which depend on 1 to 4 input parameters. The framework starts with an initial set of simulation outputs for several random inputs. Based on these input/output pairs (*samples*) we try to approximate the system's global behaviour by constructing several interpolants through the samples. To assess the quality of the models, we calculate the difference between each pair of models. Models which are most similar to each other are considered more accurate than those that deviate from the others. In an adaptive loop, new simulations are run at the inputs where the most accurate models differ the most. Afterwards, the sample set is augmented with these newly found input/output pairs and the process starts all over again.

In the sequel we present the results of tests on both artificial and real-life examples. We will demonstrate how the framework succeeds in improving accuracy by adding more and more samples.

A Collaborative Open Architecture for Data Collecting and Interpretation on Complex Artificial Systems

Silviu Ionita, Ionel Bostan, Petre Anghelescu, Alin Mazare
ionis@upit.ro

A tremendous effort to manage the complex systems in a very dynamic society is currently made. Understanding the complex systems and the investigation of the related processes requires a large amount of knowledge and the appropriate techniques for their representation and reasoning. The key issue in the knowledge-based model building on complex systems is the capability of acquisition and interpretation of high quantity and multi-resolution data. In this paper is proposed an architecture for data collecting and fusion on the large scale heterogeneous distributed systems in order to make useful analysis for decision making. The real time data availability and the high capability of their processing are essential for the efficient and robust control of complex systems. The main steps of our research are reported in the following. First, we identify the systems and subsystems and investigate the available and prospective communication infrastructure. We also provide an analysis of the major communication standards and candidate technologies for data acquisition and preprocessing. Second, an original classification of data is given according with criteria of level of resolution, real time availability, relevancy, degree of uncertainty, etc. These aspects are very important for data processing and interpretation. Third, with the development of network-based enabling technologies the novel concept of web-Centric society was introduced to coin a novel organizational paradigm that is information-exchange driven and knowledge empowered. Recent developments in distributed artificial intelligence and network technologies have enabled technological paradigms such as holonic and global organizations to become a reality. A worthwhile approach is based on dynamical systems theory and their relation to Multi-Agent Systems (MAS) whose objective is to support a so-called "going concern". Under these circumstances, the proposed collaborative open architecture is designed as a multi-agent monitoring heterarchy with a large capability in gathering, grabbing and processing the data with the goal of their interpreting in benefit of the large scale distributed systems. From an evolutionary perspective, these are powerful communication systems that receive data from their environments; find regularities in the data extract the information and then identify internal models that process this information in an optimal manner with respect to goal achievement. These distributed systems exhibit self-organization at higher aggregate system levels via information exchange through interactions between the individual agents. Communication between the distributed entities results in knowledge distribution across the overall system in a manner that optimizes goal achievement. This results in emergent and adaptive properties of the overall distributed system that are not exhibited by the individual component entities. All in all, some very suggestive examples and simulation results are presented from the relevant study cases focused on the complex systems like the urban traffic road, power distributed networks, communication networks. A special attention is dedicated of the data interpretation in order to evaluate the interdependency in terms of so called side-effects of different sub systems.

Conceptual analysis of the complexity of socio-technical systems

Bjørn Jespersen, Maarten Franssen, Maarten Ottens, Peter Kroes
Section of Philosophy, Delft University of Technology, The Netherlands
contact: b.t.f.jespersen@tbm.tudelft.nl

The purpose of our research is to develop a notion of complexity that suits the needs of a theory intended to provide a philosophical account of complex engineering systems such as infrastructures. We wish to make a contribution both to fundamental philosophical reflections on the concept of complexity as well as to philosophy of engineering.

We propose to conceptualise infrastructures as so-called 'socio-technical systems'. Such systems join entities belonging to both the physical and the intentional realms and are, therefore, characterised by being inherently hybrid.

We intend to capture hybridity by means of type theory. Thus, a system is hybrid if it includes constituents from two or more types. We are dealing with systems some of whose constituents are themselves systems (hence subsystems), making the overarching systems higher-order systems.

By 'system' we mean a unit that organises anywhere between one and infinitely many constituents into a coherent whole. By 'hybrid system' we mean a system whose constituents are drawn from two or more different kinds of entities.

By 'complex' we mean anything that has parts, constituents or components (ranging from one, i.e., the complex entity itself, to infinitely many). Every system is a complex; but it is not obvious that every complex is a system, for a complex may fail to form a coherent whole. We treat 'structure' as synonymous with 'system'.

According to what seems to be a widespread view (2), complexes are sets. We deny this, however. This is not to say that set theory might not be suitable for modelling complexes. But at least sets cannot be identified with complexes. A system is necessarily a complex, a complex is necessarily distinct from a set, which is invariably simple; hence, systems cannot be sets. It is a contentious matter, of course, how similar or dissimilar the modelling may be from what is being modelled; but our ambition is to study the complexity of socio-technical systems by studying the complexity of their formal renderings. Therefore, we prefer our modelling tools to possess complexity. (A neighbouring theme, not to be broached here, is whether, ultimately, various degrees of complexity, or indeed complexity itself, reside in the model or in what it models.)

Since complexes are distinct from sets, complexes are not mappings (functions), graphs or networks, as these are all set-theoretic entities. If systems are identical to networks, networks to graphs, and graphs are identified with their functions, then systems are functions; but a function f is a set (of all the pairs $<x \in X, y \in Y>$ such that $f(x)=y$), and sets are simple, hence systems come out simple. But systems must be composite, because, as we are assuming, they have constituents that are structured in a particular way, or in particular ways.

In conceptualising higher-order systems, we invoke a ramified type theory incorporating a simple type theory; the latter will be known from the typed λ-calculus. The simple type theory encompasses ground and functional types. The ramified type hierarchy encompasses higher-order objects as well as functions with domain or range in higher-order objects. The particular type theory underlying our modelling is the one of Transparent Intensional Logic (1); it takes over from Russell the notion of ramified type hierarchy and from Frege and Montague the idea of functions (rather than sets or relations) as basic. The (relevant) higher-order objects are related to functions by the fact that they are either the procedure, or operation, of functional formation or of functional application. These operations will also be known from the lambda-calculi.

Thus, the sort of complexity we are after is a matter of procedural, or operational, complexity. Our hypothesis is: the more operations a given system plays host to, the more complex the system is. But the complexity of a system is a function of at least one more factor; namely, the number of different types represented by instances in the system. Thus, a system encompassing n types of entities is less complex than a system with $n+1$ types. Since socio-technical systems are hybrid, instances of at least two different types occur in them. Assume that simple-type objects are first-order objects, and ramified- type objects are higher-order objects. By distinguishing between first-order objects and higher-order objects, we can say, furthermore, that a system that contains both first-order and higher-objects is more complex than one containing only higher-order objects. This is just to say that hybridity induces complexity. Each system contains as a constituent at least one higher-order object, because, trivially, it contains itself as a constituent. Since socio-technical systems count other systems among their constituents, and must be one level higher up than their constituent(s) of the highest order, they are at least third-order objects.

We imagine systems as being input/output mechanisms that transform a cluster of input into an output that may well differ significantly in kind from the input. The inner organisation of a system is crucial; *qua* procedure it processes input in a particular sequence. Apart from minimal systems, a system contains two or several sub-procedures organised in such a way that the output of one procedure serves as input for another. The notion of function comes in handy when modelling this: the value of one function is the argument of another function. (This idea will be known from, e.g., composite functions, $f \circ g$, where the domain of f includes the range of g.)

The following arithmetic example serves to illustrate the point that there is a logical dependence between input and output as the first input 'trickles down' the procedure. Consider $7+5=4\times3$. First, apply the addition function to the pair (7,5) to obtain a number m. Second, apply the multiplication function to the pair (4,3) to obtain a number k. Third, apply the identity function (as defined over positive integers, say) to the pair (m,k) to obtain a truth-value; the truth-value is True iff $m=k$. Now, it is arbitrary which of the two calculations (addition or multiplication) is carried out first, but the application of the identity function can get off the ground only when, in this case, a sum and a product have been obtained so as to provide the required arguments.

Socio-technical systems contain three categorically distinct kinds of entities: technical entities (instruments, devices, constructions); social entities (various prescriptions in the form of rules, regulations, laws, conventions); intentional and intervening entities (bearers of agency, such as individual humans, groups of people, legal bodies). Our fundamental conception of socio-technical systems is that they are irreducible either to purely social or purely technical systems. In particular, a socio-technical system is not a system or an aggregate made up of two major subsystems, one social, the other technical. [One can form the set of all technical components of any given socio-technical system, but not necessarily its technical subsystem; this is so if the elements of that set fail to form a coherent whole.)

Socio-technical systems owe their complexity to the fact that they, typically though not necessarily, contain subsystems, as well as the fact that their constituents are drawn from (at least) three different types of entities. Standard examples of socio-technical systems would be the infrastructural systems of transportation, energy supply and telecommunications. But while we claim that every infrastructural system is a socio-technical system, the converse does not hold; some socio-technical systems are, intuitively speaking, 'too small' to qualify as infrastructural.

A minimal socio-technical system: minimal by containing only one agent (but two roles: being the one to order the pushing of the button; being the one to push the button when ordered to), one rule (when ordered by the appropriate authority to push the button, push the button), and two technical devices (button incl. hook-up, and a missile on a launch-pad). By 'role' we mean an individual concept (as expressed by 'the unique *F*'). We allow that the same individual may occupy two roles or more simultaneously.

This system transforms a missile located on a launch-pad into a launched missile, i.e., trades one property of the missile for another property. The procedure is one for bringing about this change in property. Formally, the product of the procedure is a function taking the property of being a missile placed on a launching pad to the property of being a launched missile.

This research on the complexity and hybridity of socio-technical systems is part of a general research project on socio-technical systems that falls under the project Next Generation Infrastructure (NGI) through the Besluit Strategische Investeringen in de Infrastructuur.

(1) P. Tichý, *The foundations of Frege's logic*, Berlin/New York: De Gruyter, 1988.

(2) M.E.J. Newman, 'The structure and function of complex networks', *http://arxiv.org/PS cache/cond-mat/pdf/0303516.pdf*, 25 March 2003.

Conceptual analysis of the complexity of socio-technical systems

Bjørn Jespersen, Maarten Franssen, Maarten Ottens, Peter Kroes
Section of Philosophy, Delft University of Technology, The Netherlands
contact: b.t.f.jespersen@tbm.tudelft.nl

The purpose of our research is to develop a notion of complexity that suits the needs of a theory intended to provide a philosophical account of complex engineering systems such as infrastructures. We wish to make a contribution both to fundamental philosophical reflections on the concept of complexity as well as to philosophy of engineering.

We propose to conceptualise infrastructures as so-called 'socio-technical systems'. Such systems join entities belonging to both the physical and the intentional realms and are, therefore, characterised by being inherently hybrid.

We intend to capture hybridity by means of type theory. Thus, a system is hybrid if it includes constituents from two or more types. We are dealing with systems some of whose constituents are themselves systems (hence subsystems), making the overarching systems higher-order systems.

By 'system' we mean a unit that organises anywhere between one and infinitely many constituents into a coherent whole. By 'hybrid system' we mean a system whose constituents are drawn from two or more different kinds of entities.

By 'complex' we mean anything that has parts, constituents or components (ranging from one, i.e., the complex entity itself, to infinitely many). Every system is a complex; but it is not obvious that every complex is a system, for a complex may fail to form a coherent whole. We treat 'structure' as synonymous with 'system'.

According to what seems to be a widespread view (2), complexes are sets. We deny this, however. This is not to say that set theory might not be suitable for modelling complexes. But at least sets cannot be identified with complexes. A system is necessarily a complex, a complex is necessarily distinct from a set, which is invariably simple; hence, systems cannot be sets. It is a contentious matter, of course, how similar or dissimilar the modelling may be from what is being modelled; but our ambition is to study the complexity of socio-technical systems by studying the complexity of their formal renderings. Therefore, we prefer our modelling tools to possess complexity. (A neighbouring theme, not to be broached here, is whether, ultimately, various degrees of complexity, or indeed complexity itself, reside in the model or in what it models.)

Since complexes are distinct from sets, complexes are not mappings (functions), graphs or networks, as these are all set-theoretic entities. If systems are identical to networks, networks to graphs, and graphs are identified with their functions, then systems are functions; but a function f is a set (of all the pairs $<x \in X, y \in Y>$ such that $f(x)=y$), and sets are simple, hence systems come out simple. But systems must be composite, because, as we are assuming, they have constituents that are structured in a particular way, or in particular ways.

In conceptualising higher-order systems, we invoke a ramified type theory incorporating a simple type theory; the latter will be known from the typed λ-calculus. The simple type theory encompasses ground and functional types. The ramified type hierarchy encompasses higher-order objects as well as functions with domain or range in higher-order objects. The particular type theory underlying our modelling is the one of Transparent Intensional Logic (1); it takes over from Russell the notion of ramified type hierarchy and from Frege and Montague the idea of functions (rather than sets or relations) as basic. The (relevant) higher-order objects are related to functions by the fact that they are either the procedure, or operation, of functional formation or of functional application. These operations will also be known from the lambda-calculi.

Thus, the sort of complexity we are after is a matter of procedural, or operational, complexity. Our hypothesis is: the more operations a given system plays host to, the more complex the system is. But the complexity of a system is a function of at least one more factor; namely, the number of different types represented by instances in the system. Thus, a system encompassing n types of entities is less complex than a system with $n+1$ types. Since socio-technical systems are hybrid, instances of at least two different types occur in them. Assume that simple-type objects are first-order objects, and ramified- type objects are higher-order objects. By distinguishing between first-order objects and higher-order objects, we can say, furthermore, that a system that contains both first-order and higher-objects is more complex than one containing only higher-order objects. This is just to say that hybridity induces complexity. Each system contains as a constituent at least one higher-order object, because, trivially, it contains itself as a constituent. Since socio-technical systems count other systems among their constituents, and must be one level higher up than their constituent(s) of the highest order, they are at least third-order objects.

We imagine systems as being input/output mechanisms that transform a cluster of input into an output that may well differ significantly in kind from the input. The inner organisation of a system is crucial; *qua* procedure it processes input in a particular sequence. Apart from minimal systems, a system contains two or several sub-procedures organised in such a way that the output of one procedure serves as input for another. The notion of function comes in handy when modelling this: the value of one function is the argument of another function. (This idea will be known from, e.g., composite functions, $f \circ g$, where the domain of f includes the range of g.)

The following arithmetic example serves to illustrate the point that there is a logical dependence between input and output as the first input 'trickles down' the procedure. Consider $7+5=4\times3$. First, apply the addition function to the pair (7,5) to obtain a number m. Second, apply the multiplication function to the pair (4,3) to obtain a number k. Third, apply the identity function (as defined over positive integers, say) to the pair (m,k) to obtain a truth-value; the truth-value is True iff $m=k$. Now, it is arbitrary which of the two calculations (addition or multiplication) is carried out first, but the application of the identity function can get off the ground only when, in this case, a sum and a product have been obtained so as to provide the required arguments.

Socio-technical systems contain three categorically distinct kinds of entities: technical entities (instruments, devices, constructions); social entities (various prescriptions in the form of rules, regulations, laws, conventions); intentional and intervening entities (bearers of agency, such as individual humans, groups of people, legal bodies). Our fundamental conception of socio-technical systems is that they are irreducible either to purely social or purely technical systems. In particular, a socio-technical system is not a system or an aggregate made up of two major subsystems, one social, the other technical. [One can form the set of all technical components of any given socio-technical system, but not necessarily its technical subsystem; this is so if the elements of that set fail to form a coherent whole.)

Socio-technical systems owe their complexity to the fact that they, typically though not necessarily, contain subsystems, as well as the fact that their constituents are drawn from (at least) three different types of entities. Standard examples of socio-technical systems would be the infrastructural systems of transportation, energy supply and telecommunications. But while we claim that every infrastructural system is a socio-technical system, the converse does not hold; some socio-technical systems are, intuitively speaking, 'too small' to qualify as infrastructural.

A minimal socio-technical system: minimal by containing only one agent (but two roles: being the one to order the pushing of the button; being the one to push the button when ordered to), one rule (when ordered by the appropriate authority to push the button, push the button), and two technical devices (button incl. hook-up, and a missile on a launch-pad). By 'role' we mean an individual concept (as expressed by 'the unique *F*'). We allow that the same individual may occupy two roles or more simultaneously.

This system transforms a missile located on a launch-pad into a launched missile, i.e., trades one property of the missile for another property. The procedure is one for bringing about this change in property. Formally, the product of the procedure is a function taking the property of being a missile placed on a launching pad to the property of being a launched missile.

This research on the complexity and hybridity of socio-technical systems is part of a general research project on socio-technical systems that falls under the project Next Generation Infrastructure (NGI) through the Besluit Strategische Investeringen in de Infrastructuur.

(1) P. Tichý, *The foundations of Frege's logic*, Berlin/New York: De Gruyter, 1988.

(2) M.E.J. Newman, 'The structure and function of complex networks', *http://arxiv.org/PS cache/cond-mat/pdf/0303516.pdf*, 25 March 2003.

Conceptual analysis of the complexity of socio-technical systems

Bjørn Jespersen, Maarten Franssen, Maarten Ottens, Peter Kroes
Section of Philosophy, Delft University of Technology, The Netherlands
contact: b.t.f.jespersen@tbm.tudelft.nl

The purpose of our research is to develop a notion of complexity that suits the needs of a theory intended to provide a philosophical account of complex engineering systems such as infrastructures. We wish to make a contribution both to fundamental philosophical reflections on the concept of complexity as well as to philosophy of engineering.

We propose to conceptualise infrastructures as so-called 'socio-technical systems'. Such systems join entities belonging to both the physical and the intentional realms and are, therefore, characterised by being inherently hybrid.

We intend to capture hybridity by means of type theory. Thus, a system is hybrid if it includes constituents from two or more types. We are dealing with systems some of whose constituents are themselves systems (hence subsystems), making the overarching systems higher-order systems.

By 'system' we mean a unit that organises anywhere between one and infinitely many constituents into a coherent whole. By 'hybrid system' we mean a system whose constituents are drawn from two or more different kinds of entities.

By 'complex' we mean anything that has parts, constituents or components (ranging from one, i.e., the complex entity itself, to infinitely many). Every system is a complex; but it is not obvious that every complex is a system, for a complex may fail to form a coherent whole. We treat 'structure' as synonymous with 'system'.

According to what seems to be a widespread view (2), complexes are sets. We deny this, however. This is not to say that set theory might not be suitable for modelling complexes. But at least sets cannot be identified with complexes. A system is necessarily a complex, a complex is necessarily distinct from a set, which is invariably simple; hence, systems cannot be sets. It is a contentious matter, of course, how similar or dissimilar the modelling may be from what is being modelled; but our ambition is to study the complexity of socio-technical systems by studying the complexity of their formal renderings. Therefore, we prefer our modelling tools to possess complexity. (A neighbouring theme, not to be broached here, is whether, ultimately, various degrees of complexity, or indeed complexity itself, reside in the model or in what it models.)

Since complexes are distinct from sets, complexes are not mappings (functions), graphs or networks, as these are all set-theoretic entities. If systems are identical to networks, networks to graphs, and graphs are identified with their functions, then systems are functions; but a function f is a set (of all the pairs $<x \in X, y \in Y>$ such that $f(x)=y$), and sets are simple, hence systems come out simple. But systems must be composite, because, as we are assuming, they have constituents that are structured in a particular way, or in particular ways.

In conceptualising higher-order systems, we invoke a ramified type theory incorporating a simple type theory; the latter will be known from the typed λ-calculus. The simple type theory encompasses ground and functional types. The ramified type hierarchy encompasses higher-order objects as well as functions with domain or range in higher-order objects. The particular type theory underlying our modelling is the one of Transparent Intensional Logic (1); it takes over from Russell the notion of ramified type hierarchy and from Frege and Montague the idea of functions (rather than sets or relations) as basic. The (relevant) higher-order objects are related to functions by the fact that they are either the procedure, or operation, of functional formation or of functional application. These operations will also be known from the lambda-calculi.

Thus, the sort of complexity we are after is a matter of procedural, or operational, complexity. Our hypothesis is: the more operations a given system plays host to, the more complex the system is. But the complexity of a system is a function of at least one more factor; namely, the number of different types represented by instances in the system. Thus, a system encompassing n types of entities is less complex than a system with $n+1$ types. Since socio-technical systems are hybrid, instances of at least two different types occur in them. Assume that simple-type objects are first-order objects, and ramified- type objects are higher-order objects. By distinguishing between first-order objects and higher-order objects, we can say, furthermore, that a system that contains both first-order and higher-objects is more complex than one containing only higher-order objects. This is just to say that hybridity induces complexity. Each system contains as a constituent at least one higher-order object, because, trivially, it contains itself as a constituent. Since socio-technical systems count other systems among their constituents, and must be one level higher up than their constituent(s) of the highest order, they are at least third-order objects.

We imagine systems as being input/output mechanisms that transform a cluster of input into an output that may well differ significantly in kind from the input. The inner organisation of a system is crucial; *qua* procedure it processes input in a particular sequence. Apart from minimal systems, a system contains two or several sub-procedures organised in such a way that the output of one procedure serves as input for another. The notion of function comes in handy when modelling this: the value of one function is the argument of another function. (This idea will be known from, e.g., composite functions, $f \circ g$, where the domain of f includes the range of g.)

The following arithmetic example serves to illustrate the point that there is a logical dependence between input and output as the first input 'trickles down' the procedure. Consider $7+5=4\times3$. First, apply the addition function to the pair (7,5) to obtain a number m. Second, apply the multiplication function to the pair (4,3) to obtain a number k. Third, apply the identity function (as defined over positive integers, say) to the pair (m,k) to obtain a truth-value; the truth-value is True iff $m=k$. Now, it is arbitrary which of the two calculations (addition or multiplication) is carried out first, but the application of the identity function can get off the ground only when, in this case, a sum and a product have been obtained so as to provide the required arguments.

Socio-technical systems contain three categorically distinct kinds of entities: technical entities (instruments, devices, constructions); social entities (various prescriptions in the form of rules, regulations, laws, conventions); intentional and intervening entities (bearers of agency, such as individual humans, groups of people, legal bodies). Our fundamental conception of socio-technical systems is that they are irreducible either to purely social or purely technical systems. In particular, a socio-technical system is not a system or an aggregate made up of two major subsystems, one social, the other technical. [One can form the set of all technical components of any given socio-technical system, but not necessarily its technical subsystem; this is so if the elements of that set fail to form a coherent whole.)

Socio-technical systems owe their complexity to the fact that they, typically though not necessarily, contain subsystems, as well as the fact that their constituents are drawn from (at least) three different types of entities. Standard examples of socio-technical systems would be the infrastructural systems of transportation, energy supply and telecommunications. But while we claim that every infrastructural system is a socio-technical system, the converse does not hold; some socio-technical systems are, intuitively speaking, 'too small' to qualify as infrastructural.

A minimal socio-technical system: minimal by containing only one agent (but two roles: being the one to order the pushing of the button; being the one to push the button when ordered to), one rule (when ordered by the appropriate authority to push the button, push the button), and two technical devices (button incl. hook-up, and a missile on a launch-pad). By 'role' we mean an individual concept (as expressed by 'the unique *F*'). We allow that the same individual may occupy two roles or more simultaneously.

This system transforms a missile located on a launch-pad into a launched missile, i.e., trades one property of the missile for another property. The procedure is one for bringing about this change in property. Formally, the product of the procedure is a function taking the property of being a missile placed on a launching pad to the property of being a launched missile.

This research on the complexity and hybridity of socio-technical systems is part of a general research project on socio-technical systems that falls under the project Next Generation Infrastructure (NGI) through the Besluit Strategische Investeringen in de Infrastructuur.

(1) P. Tichý, *The foundations of Frege's logic*, Berlin/New York: De Gruyter, 1988.

(2) M.E.J. Newman, 'The structure and function of complex networks', *http://arxiv.org/PS cache/cond-mat/pdf/0303516.pdf*, 25 March 2003.

Fractal Analysis of Microglial Morphology

A Karperien, C. Lucas, C. Depardieu, G Aurel, H.F Jelinek
hjelinek@csu.edu.au

Field: Particular complex systems Keywords: fractal analysis, nonlinear systems, fractal dimension, microglia classification, pathology

Microglial morphology is an important clue to microglial activity and thus may be an important indicator in several pathologies, including dementia and schizophrenia, but is rather difficult to quantify. Structural complexity has been assessed using fractal analysis, offering a way to quantitatively describe otherwise difficult to characterize forms and may be useful in the study of the role of microglia in human disease. We used box-counting dimension to ascertain the utility of fractal analysis in the classification of microglia cells and have found statistically significant differences between subpopulations of these cells ($P < 0.05$).

Introduction Microglial form depends on context and is a continuum in a changing context. Not only is form different for subpopulations of microglia; microglia change form over minutes to hours; microglia move; they change complexity with age; they become activated in many different ways; and they are influenced by the local and overall physiological environments. Nevertheless, perhaps form can be assessed during an interval as manifested by their morphology and divided into ramified, bushy or reactive. While much about microglial function can be inferred from staining and morphological features, several studies argue that nearly imperceptibly altered microglia may underlie chronic, subtle pathology in schizophrenia and other brain diseases not always detectable by histological methods. If this is the case, detecting problems prior to overt damage will be a priority. Yet it is currently difficult to detect subtle abnormalities in microglia. Fractal analysis may be of use here to identify these subtle changes in morphology associated with functional states.

Method Three exclusive operational definitions for microglia structure were applied: ramified=having usually many processes; bushy=having usually shorter, thicker processes; and reactive=having few relatively very short, thick extensions with large round to amorphous cell body. Cells were randomly selected from within 15 image areas from a random sample of the set of microglia images, unsorted with respect to species, staining and gender. Cells were then classified into one of the above categories, then blindly and randomly matched as pairs from within each area. This gave 29 ramified and bushy pairs and 7 bushy and reactive pairs. Contours of cells were analysed using FracLac as a plug in for ImageJ, with a maximum box size set at 47

Results For the ramified (mean=1.42; standard deviation=0.004) and bushy (mean=1.40; standard deviation=0.086) pairs, the mean difference was 0.015 (standard deviation of the differences=0.152). Using a two-tailed t-test, a p value of 0.5 and r=0. 0.3950 was obtained. For the bushy (mean=1.43; standard deviation=0.0901) and reactive (mean=1.27; standard deviation= 0.1394) pairs, at $p = 0.0143$; r=0.51 for these pairs. There is also significant difference between ramified and reactive, $p < 0.05$ (mean=1.42; standard deviation=0.004 mean=1.27; standard deviation= 0.1394).

Discussion This investigation looked at how well the Db differentiates individual microglia in a general sense, looking not at absolute values of the Db but at differences in differently classified cells within one environment. The major finding was that the Db differentiated cells classified as bushy from cells classified as reactive, and ramified from reactive but did not clearly differentiate those classified as bushy from those classified as ramified when multiple paradigms were considered together and a consistent relative difference between categories was looked for. The Db proved useful for measuring differences between different functional states and thus may provide novel insights into understanding

the complex development of pathology associated with schizophrenia or dementias, which have been reported to have microglia involvement.

Visualising Interactions in Complex Design

René Keller, Claudia M. Eckert, P. John Clarkson
Engineering Design Centre
University of Cambridge
Trumpington Street
Cambridge, CB2 1PZ
United Kingdom
{rk313, cme26, pjc10}@cam.ac.uk

Designing a product is complex in many ways. Four layers, which can be sources of complexity in design, can be identified. First, the product itself can be complex as it might have many components that are highly interrelated and linked in various ways. Second, the process of designing the product can consist of many interlinked tasks with probabilistic outcomes that can cause costly iteration. Third, the organisation that designs the product can be considered complex, as it consists of a large number of multidisciplinary teams. Fourth, the relation of the product to its environment can be complex. In addition representations used to interact with the product can also be highly complex.

We concentrate on how to visualise the first two aspects of complexity, the complexity of the product architecture and the complexity of the corresponding design process. Both, products and processes can be modelled as networks, consisting of items (components and tasks) and interactions and dependencies between them. Networks that show an appropriate level of detail can easily become very huge, consisting of hundreds or thousands of items connected in various ways. The key for successful design is to have appropriate interfaces and visualisations of this information so that the designer can effectively make use of the models.

Two main representations for network data exist: Adjacency matrices and node edge diagrams. Both representations have advantages and disadvantages when showing such data. While matrices are very compact representations that let users easily extract information about direct linkages between items, networks proved superior behaviour when indirect interactions between items are to be visualised. However, for very large and complex models, further advanced information visualisation techniques such as filtering and fisheye techniques are necessary in order not to overwhelm the user with the amount of available information.

We will introduce network and matrix representations that make use of advanced information visualisation techniques and are used in two methods developed at the Engineering Design Centre Cambridge to support the design of complex products. The signposting methodology allows a dynamic approach to planning and analysing the design process used to design complex products. The basic idea is to dynamically update the state of the design and present the designer at each state of the design using a signpost of how to proceed with the design process. Simulations of such design processes allow further analyses in order to detect bottlenecks and to find optimal task execution orders. The other method that makes extensive use of network visualisations is the CPM (Change Prediction Method) to predict direct and indirect design changes. When a component of a complex product has to be changed (e.g. because of a customer request), this change can have knock-on effects on other components that then equally affect further components. Using the CPM method allows designers to predict the impacts of changes to directly and not directly connected components.

In both examples it is important for the designer to get detailed information about a particular case as well as a global overview over the entire information space. In this paper we will introduce displays that visualise complex process and product data taking advantage of enhanced information visualisation techniques and that have shown to be highly beneficial to the design process.

Evolutionary influence of the protein network topology on gene organisation in artificial organisms

Carole Knibbe[1], Guillaume Beslon[1], Jean-Michel Fayard[2]

Prisma – INSA Lyon – France
BF2I – INSA Lyon – France

Although molecular biology provided us with huge data about individual genes, their roles cannot be fully understood unless we put them into the broader context of their interactions. Multiple links indeed exist between genes: they can share the same ancestral sequence, they can be neighbours on a chromosome, they can be coregulated or regulate each other's expression level, their proteins can interact physically or be involved in the same metabolic pathway. Those various points of view lead to different networks - different but not independent, notably because they are involved in a common evolutionary story. Understanding these structures therefore requires to understand how they build up and interact at the evolutionary time scale.

In particular, are genes randomly distributed on the chromosomes, or does their organisation depend on their functional interactions? It is difficult to answer this question directly for living cells, notably because the notion of gene function is subjective in this context. We therefore chose an artificial life approach, using the fuzzy logic formalism as a generic framework for the functional level.

In our model, called aevol, each organism is able to perform abstract processes with various degrees, depending on the set of interacting functional elements ("proteins") its genome encodes. The genome is a binary string that contains coding sequences (genes) separated by non-coding sequences. Each gene is translated into a protein, able to either realize or inhibit a fuzzy set of processes. Two proteins can have the same process in their fuzzy sets, implying that for the organism, the degree of the process depends on their interaction. Thus, the organism's phenotype is the logic combination of the sets of its proteins. Like in the genetic algorithms used in computer science, selection and variation mechanisms allow the phenotypes to become closer and closer to an arbitrary optimal fuzzy set. When the fittest organisms reproduce, their genomes are replicated with eventual random errors, at the local scale (punctual mutations, small insertions, small deletions) or at the scale of large genomic segments (duplications, translocations, large deletions, inversions).

Since the fitness computation does not include genomic criteria like genome length or gene order, the genome is free to self-organise. Similarly, the functional network is not predefined, and the number of proteins can evolve, as well as the strengths of their interactions. We can however set the maximal interaction potential of the proteins - that is to say the maximal number of processes they can be involved in -, thereby limiting the network's average connectivity. This feature is expected to have evolutionary consequences, since deleting a highly connected gene, for instance, has a much more deleterious effect than deleting an "independent" gene.

To investigate the evolutionary influence of the functional network's topology on the genomic level, we carried out experiments with five different maximal interaction potentials, during 30,000 generations. In most cases, the gene organisation alternates shuffling phases with steady states throughout the evolution, and these steady states seem to depend on the functional network topology. If the maximal interaction potential is high, hence if many genes are highly connected, genes tend to distribute regularly along the chromosome after a shuffling phase. On the contrary, if the interaction potential is low, genes tend to form clusters. For intermediate values of the parameter, gene distribution either remains random, or alternates both steady states.

This emerging relation between the genomic and the functional levels could be the result of selective pressures for both robustness and evolvability, requiring an intermediate effect for the large-scale mutations – not too high, but also not too small. When genes are highly connected, this intermediate effect can be achieved by affecting one gene only, whereas when genes are more independent, affecting several of them at the same time can speed up the discovering of new solutions.

An evolutionary process could therefore make links emerge between different organisation levels in our artificial complex system. To confirm these results, additional experiments will be carried out with extreme values of the parameter, while further analysis of the impact of the mutational events should help us testing our explanatory mechanism.

Morphometrica

Romulo Krafta
krafta@ufrgs.br

Agents and Space in a Urban Model

Key words: spatial dynamics simulation, urban morphology; urban growth

It is reported efforts towards the construction of a model for urban spatial dynamics simulation, based on multi-agents and space. The underlying idea is to have urban space producers and consumers operating in a two-layer, two-circuit model. The first one holds urban space and its successive transformations; one second layer contains agents related to space; one first circuit simulates space production, and a second one simulates space consumption. Relationship between layers is represented as objective spatial features that agents are submitted to and subjective meanings agents attach to each spatial feature. While space works always in the same way, meanings vary according to each agent's background and context. Relationship between circuits are represented by means of a market game in which producers try to maximize their profits by gambling with their risks, whereas consumers try to foresee the spatial distribution of local externalities that maximizes their utilities and investments.

Reliable Broadcasting

Reliable Broadcasting in an Automotive Scenario

Design overview

DELIS

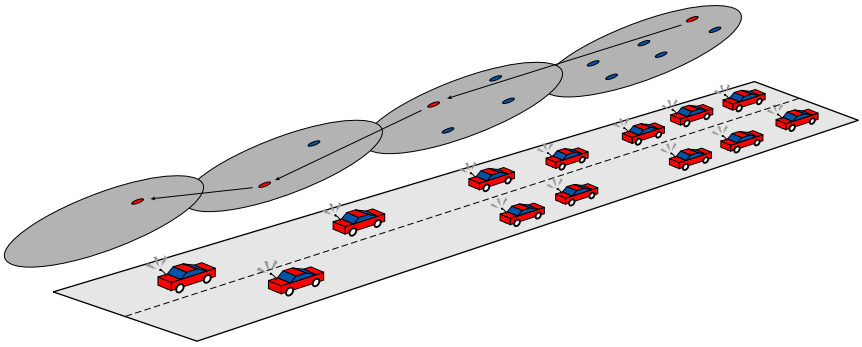
Dynamically Evolving, Large-scale Information Systems

Information Society Technologies

University of Paderborn, Algorithms and Complexity & Wrocław University of Technology

Abstract

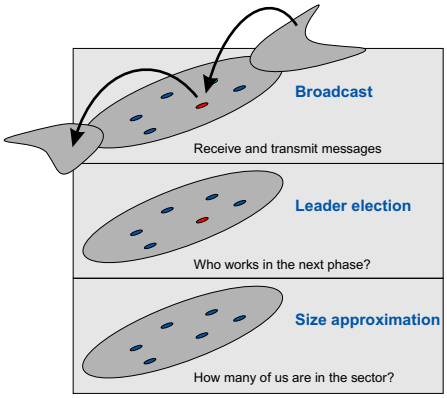
We propose a dynamic, ad-hoc communication network consisting of mobile units that can warn about traffic jams on highways.
Our goal is to provide a practical, low cost solution. Therefore we consider very simple wireless communication hardware, without collision detection, with very small bandwidth and a probabilistic model of link failure.
We provide a complete system architecture, consisting of three fine-tuned algorithms which allow the stations to self-organize and transmit traffic jam warnings.



Architecture

We consider a street conceptually divided into sectors. Whenever a traffic jam is detected, a warning message is issued in the appropriate sector and is sent from sector to sector down the road.

The system operates with the help of three layered algorithms. A broadcast algorithm is responsible for transmitting information about the traffic jams between sectors. Leader election and size approximation algorithms help the nodes to self-organize and provide a framework for the broadcast algorithm.



The wireless transceivers work with one frequency thus a time-division protocol is applied. The time is statically divided between algorithms and sectors, so that nodes in sectors lying within interference range do not transmit at the same time.

The link unreliability model causes every successful (without collision) transmission to fail to reach a node with a constant probability. This event is independent for each station, thus each transmission is heard only by a fraction of nodes within transmission range.

Broadcast

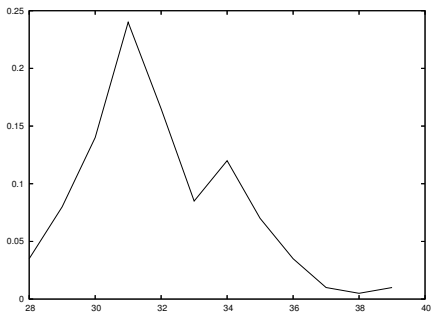
To avoid problems caused by link unreliability, the broadcast algorithm resends each message several times. This way, a message makes its way to the next sector with high probability.

Basing on the analysis of the features of the algorithm it follows that a message will travel some distance within a given time with a constant probability, even with information losses due to link unreliability. The probability for a message to reach distance D within time T is given by

$$c^D \left(\sum_{t=D}^{T} \binom{t-1}{D-1} p_r^D (1-p_r)^t \right) ,$$

with some appropriate constant c (dependent on p_r) and link unreliability given by $(1-p_r)$ (i.e. each station hears a transmission successfully with probability p_r).

Besides the theoretical analysis, the framework has been evaluated within a simulation of a highway with a cellular automaton. The figure shows the probability distribution of the message transmission time on a distance of 27 sectors, corresponding to 6.75 km road distance. The link unreliability has been chosen to be $p_r = 0.9$. The most common transmission time of 31 slots corresponds to 3.1 seconds.



Framework

The size approximation and leader election algorithms are fine-tuned for best operation with small node numbers. According to our experiments, the best known algorithms in terms of asymptotical runtime have much too large constants and thus are not usable in our practical setting.

We have chosen a simple leader election algorithm in which every node transmits in every round with probability $1/n$. For node quantities in the order of 100 it has performed best, using only $log_2 \delta^{-1}$ rounds for a successful leader election with probability greater than $1-\delta$.

For size approximation a family of algorithms has been proposed. A common feature of all these algorithms is that they divide they runtime into several phases. In each phase every active node (which has not successfully transmitted yet) chooses one round and transmits in it. The number of successful transmissions in the whole run is counted and, after the last phase, describes the number of nodes in a sector.

The task of choosing the lengths of phases and their number is challenging, when trying to both obtain a small runtime and high probability for exact estimation of the number of nodes. We provide an approach for constructing phase length sequences. This sequence assures that the exact number of nodes will be returned with a given probability. Obtained results show that for practical scenarios a runtime of cn with $c < 10$ can be achieved. This outperforms known algorithms for small n.

When runtime is even more crucial, there exists also the possibility to use only one round of length n and approximate the actual number of stations basing on the number of successful transmissions within this single round.

Authors

Jarosław Kutyłowski
University of Paderborn
Fürstenalle 11
33102 Paderborn, Germany
Phone: +49 5251 606458
Fax: +49 5251 606482
Email: jarekk@upb.de

Filip Zagórski
Wrocław University of Technology
Janiszewskiego 14a
50-370 Wrocław, Poland
Phone: +48 71 3202105
Fax: +48 71 3202109
Email: filipz@im.pwr.wroc.pl

DELIS: Integrated European Project founded by the "Complex Systems" Proactive Initiative within the Sixth Framework Programm

Peanuts: A Top-Down Peer-to-Peer Network

Peter Mahlmann* Christian Schindelhauer

We introduce Peanuts – a top-down Peer-to-Peer network combining the benefits of reliable random graphs and semantic search trees. The main goal of Peanuts is to overcome the restricted query languages of almost all current peer-to-peer networks. These restrictions are a consequence of the use of distributed hash tables where any semantic relationship of the data is lost. In a top-down approach such semantic relationship is preserved since the peers are assigned to the data and not vice versa.

Peanuts allows nontrivial lookups, like range queries, neighborhood search, and estimation of the popularity of prefixes. Furthermore, the maintenance of the network structure is kept local so that periodic handshake operations suffice to maintain the network structure. Simplifying, Peanuts supports three kinds of locality what results in the possibility to support non-trivial queries:

Network Locality: Lookup operations can be performed with small latency.

Information Locality: Closely related data elements are stored on network-wise close peers.

Interest Locality: Peers can choose on providing lookup service and data storage for certain data. If peers choose to provide certain data, the network structure will improve their lookup of related data.

Random graphs are Peanuts main component to build arbitrary search trees supporting these kinds of locality. There is a *Reliable Backbone* which is a random graph consisting of all participating peers. This Backbone represents the root of the semantic search tree. The backbone network should be simple and reliable under churn. Therefore, the backbone is realized by a random network. A random network has several advantages compared to deterministic structured networks. For example the maintenance costs in highly dynamic networks are minimal, since there is no predetermined neighborhood. In [1] we introduce a simple scheme to maintain such random networks. The main component of this scheme is a simple periodically performed local link exchange, guaranteeing connectivity and establishing expander graphs with high probability.

The peers of the backbone are recursively assigned into sub-sets, each representing one child node of the current tree node. Instead of the usual approach of hashing data onto peers, we hash peers onto data. The assignment of peers to tree nodes is done using weighted consistent hashing [2], an extension of consistent hashing to support non-uniform load distributions. In Peanuts, the number of peers assigned to a tree node depends on its load, i.e. the amount respectively popularity of data stored on this nodes subtree. This way unbalanced trees can be handled efficiently.

Peanuts is currently implemented in cooperation within the EU-project DELIS to provide an efficient peer-to-peer based data structure for the implementation of a distributed Web search engine.

References

[1] Peter Mahlmann, Christian Schindelhauer, Peer-to-Peer Networks based on Random Transformations of Connected Regular Undirected Graphs, 17th ACM Symposium on Parallelism in Algorithms and Architectures, (SPAA 2005), Las Vegas, NV, USA, July 17-20, 2005

[2] Christian Schindelhauer, Gunnar Schomaker, Weighted Consistent Hashing, 17th ACM Symposium on Parallelism in Algorithms and Architectures 2005, (SPAA 2005), Las Vegas, NV, USA, July 17-20, 2005

*{mahlmann,schindel}@upb.de, Department of Computer Science and Heinz Nixdorf Institute, University of Paderborn, Germany. Partially supported by the DFG-Sonderforschungsbereich 376 and by the EU within 6th Framework Programme under contract 001907 “Dynamically Evolving, Large Scale Information Systems” (DELIS).

Tubes

Alexandre Makarovitsch
malex@club-internet.fr

This is a journey in complex systems. Starting with a tube made of a "virtual material", and applying a set of couples of operators - cut/paste, fold/unfold, pierce/patch - at each clock tic, systems more and more complex appear at each generation. The research has the ambition to model such innovative shapes generation and try to provide a new look into the complex systems generation from simpler shapes, using simple operators. The research is pursued within the Institute of Applied Mathematics of the Université Catholique de l'Ouest, the POLIS initiative.

One single molecule to access another scale? PAI-1, Microenvironment and Cancer cell migration.

M. Malo, F. Delaplace, G . Barlovatz Meimon
gbm@lami.univ-evry.fr

The cancer pathology is a complex process caused by a cellular dysfunction leading to a whole organ or even organism vital perturbation. To better understand this process we need to understand each of the levels involved and what allows the change of scales.

A matricellular protein, called PAI-1* has been identified as able to induce cell adhesion, reorganization of actin cytoskeleton and morphological changes, and to promote cell migration. PAI-1 participates in a regulation network linking the extracellular matrix and the cell. The molecular bridge formed by this protein and others is able to transduce biochemical and biomechanical signals to the cell. The cell response can be evaluated in terms of mRNA rate constants. It can further be assessed by the evaluation of migration speed. We are using the game theory and the game network theory to model this dynamic at the cell scale.

PAI-1, found closely around the most invasive tumors represents an independent factor of bad prognosis. Cancer cells, having undergone the epithelial-mesenchymal transition (EMT), can use PAI-1 for their migration. However, if the environnement is very rich in PAI-1, the cells undergo a second transition, the so-called mesenchymal-amaeboid transition (MAT). Then their migration becomes amaeboid very different from the previous one by the independence towards integrins and proteases activity. This amaeboid state is characterized by a complete reorganisation of actin (with actin rings) and by the activation of the Rho GTPases transduction pathways. Furthermore, in the presence of this particular (i.e. PAI-1 enriched) environment, the cells down-regulate their PAI-1 mRNAs rate constants.

From the cellular level to the tissue level. Here the regulation network occurs at the cellular/microenvironmental level and includes a PAI-1 concentration threshold. However, the consequences of a unique cell MAT can be its escape from the initial tumour, i.e. metastatic migration. If this occurs, the moving cancer cell will meet a microenvironment in which PAI-1 progressively decreases, and the cell could return to the mesenchymal condition through an inverse transition (form amaeboid to mesenchymal behaviour). As it is a cancer cell, it will proliferate, produce more PAI-1 (a characteristic of the most invasive cells) and perhaps again undergo MAT. The PAI-1 negative feed-back in terms of RNAs, described earlier, could help this regulation to occur. We have perhaps identified one of the conditions for a cell to undergo MAT and return.

A "grain of sand" The regulation of a cancer cell behaviour by its PAI-1 microenvironment could then be considered as a cause leading to consequences at the organ or the organism level. And its critical transition point is that one single molecule of PAI-1 could do the job! In avalanches, it is known that at the critical state the output is not proportional to the input, that is, the system is highly nonlinear and gives rise to "non-obvious" effects. Could one molecule of PAI-1 play the role of a "grain of sand" in the biological avalanche of cancer? Is the suggestion of self-organised criticality useful to explain such scale changes?

*Plasminogen Activator Inhibitor type 1

One single molecule to access another scale? PAI-1, Microenvironment and Cancer cell migration.

M. Malo, F. Delaplace, G . Barlovatz Meimon
gbm@lami.univ-evry.fr

The cancer pathology is a complex process caused by a cellular dysfunction leading to a whole organ or even organism vital perturbation. To better understand this process we need to understand each of the levels involved and what allows the change of scales.

A matricellular protein, called PAI-1* has been identified as able to induce cell adhesion, reorganization of actin cytoskeleton and morphological changes, and to promote cell migration. PAI-1 participates in a regulation network linking the extracellular matrix and the cell. The molecular bridge formed by this protein and others is able to transduce biochemical and biomechanical signals to the cell. The cell response can be evaluated in terms of mRNA rate constants. It can further be assessed by the evaluation of migration speed. We are using the game theory and the game network theory to model this dynamic at the cell scale.

PAI-1, found closely around the most invasive tumors represents an independent factor of bad prognosis. Cancer cells, having undergone the epithelial-mesenchymal transition (EMT), can use PAI-1 for their migration. However, if the environnement is very rich in PAI-1, the cells undergo a second transition, the so-called mesenchymal-amaeboid transition (MAT). Then their migration becomes amaeboid very different from the previous one by the independence towards integrins and proteases activity. This amaeboid state is characterized by a complete reorganisation of actin (with actin rings) and by the activation of the Rho GTPases transduction pathways. Furthermore, in the presence of this particular (i.e. PAI-1 enriched) environment, the cells down-regulate their PAI-1 mRNAs rate constants.

From the cellular level to the tissue level. Here the regulation network occurs at the cellular/microenvironmental level and includes a PAI-1 concentration threshold. However, the consequences of a unique cell MAT can be its escape from the initial tumour, i.e. metastatic migration. If this occurs, the moving cancer cell will meet a microenvironment in which PAI-1 progressively decreases, and the cell could return to the mesenchymal condition through an inverse transition (form amaeboid to mesenchymal behaviour). As it is a cancer cell, it will proliferate, produce more PAI-1 (a characteristic of the most invasive cells) and perhaps again undergo MAT. The PAI-1 negative feed-back in terms of RNAs, described earlier, could help this regulation to occur. We have perhaps identified one of the conditions for a cell to undergo MAT and return.

A "grain of sand" The regulation of a cancer cell behaviour by its PAI-1 microenvironment could then be considered as a cause leading to consequences at the organ or the organism level. And its critical transition point is that one single molecule of PAI-1 could do the job! In avalanches, it is known that at the critical state the output is not proportional to the input, that is, the system is highly nonlinear and gives rise to "non-obvious" effects. Could one molecule of PAI-1 play the role of a "grain of sand" in the biological avalanche of cancer? Is the suggestion of self-organised criticality useful to explain such scale changes?

*Plasminogen Activator Inhibitor type 1

Modeling Reflective, Anticipatory, Complex Adaptive Systems

Peter McBurney
p.j.mcburney@csc.liv.ac.uk

KEYWORDS: Anticipatory systems, Economic systems, Multiagent systems, Performatives, Rational expectations, Reflexivity.

Robert Rosen [1985] defined "anticipatory systems" as those where some or all entities have models of the system itself, which they use to predict system properties. Most living and human social systems are anticipatory in this sense. I define "reflective systems" to be those where some or all entities exchange predictions of system properties with other entities. The challenge in modeling such systems is that reflective messages passed between the entities in the system may modify not only each others' predictions of system properties, but also their current behaviours, and thus modify the system itself and its properties.

Modern economies are reflective, with considerable discussion by participants on the future developments of the economy, and thedirect exchange of predictions and forecasts. The theory of rational expectations (Muth 1961) is an attempt within economics to model the impact of such discussion on the revision of agents' beliefs and actions. However, economic theory itself may have a reflective impact on economic phemonena, as is shown by the progress of the Black-Scholes equation in financial markets (MacKenzie 2003). These performative aspects of utterances makes control of such systems difficult if not impossible, because the impacts of utterances and interventions (on participants and on the system) need to be taken into account when intervening. An example is shown, in the setting of interest rates by central banks, such as the US Federal Reserve System or the Bank of England. Central bankers need to allow not only for the real economic impact that increases or decreases in interest rates will have on underlying economic activity, but also for their impacts on the expectations of economic actors, such as investors and stock analysts. In order to manage these expectations, central banks in some countries (eg, the UK, the USA) have recently begun to make public the minutes of their deliberations, so as to better inform other economic actors as to the Banks' reasoning processes (Kohn 2005). It is not obvious that better-informed actors are less susceptible to major fluctuations in expectations, but this is certainly the view of central bankers.

How should reflective, anticipatory complex adaptive systems be modeled? Current multi-agent models of complex human systems have typically used only cognitively-simple agents, such as those playing the minority game in simulations of financial markets, e.g. Johnson et al. [2001]. These agents are usually not equipped to undertake sophisticated anticipatory reasoning or engage in much joint reflective activity. However, within the computer science agent technology community, considerable efforts have been devoted to the design and implementation of intelligent software agents and to the design of sophisticated, human-like, communications languages for their interactions (Luck et al. 2003). The computational technologies therefore now exist to enable effective modeling of large-scale anticipatory, reflective systems.

In this paper, I present a layered multi-agent architecture for modeling such systems. The architecture comprises intelligent software agents who communicate using sophisticated agent interaction protocols, with the layers being (in ascending order):

- Layer 1: The underlying socio-economic phenomena
- Layer 2: The anticipatory phenomena (ie, the ability of some agents to engage in prediction of system properties)
- Layer 3: The reflective phenomena (ie, the ability of some agents to engage in discussion about system properties).

Activities of each layer may impact those above and below it, in defined ways. Agents

are connected to other agents at each layer, with possibly different connections at different layers. Not all agents may be active at all layers. In this paper, I present this architecture in detail and describe an implementation of it in the domain of malaria epidemiology.

Developing a domain independent model of Emergence

Diane M. McDonald, George R.S. Weir
Diane.McDonald@strath.ac.uk

The complex and rapidly changing demands of the Information Society and the recognition of the power of community learning has led to the observation that learning communities could be a pivotal tool for achieving more complex aims and objectives than learning alone. For example, the use of learning communities to stimulate learning, creativity and economic capacity is emerging as a potential tool to increase economic success. These Complex Learning Communities (CLC) are radically heterogeneous and non-linear with multiple drivers and desired products (McDonald 2005a). This complex nature means that emergence may occur; new properties - be it behaviour or characteristics, at the individual actor level or collective properties of the system as a whole - may emerge from the interactions within the CLCs. For example, relevant emergence may include social capital (Daniel et al 2003), eLiteracy (McDonald McGill 2005) and creativity (Cavalletti 2003). The intrinsic unpredictability of Complex Systems and hence CLCs means, however, that specific emergence is not guaranteed; it is dependent on the dynamics of interaction and initial conditions. Indeed other unwanted emergence may occur for example, spamming or digital exclusion. If successful CLCs are to be seeded and their effectiveness maximised, the factors which may contribute to the desired emergence need to be identified and the dynamics of their interaction understood. Emergence itself is however not yet fully understood in all its many forms (Kubik 2003). While one of the principle drivers of Complexity Science is the vision of seeking out generalisations based on the well-researched instances of emergence and using those understandings to make sense of less researched complex phenomena, current conceptual models of emergence are either too simplistic to be useful when examining real life social systems (Funtowicz and Ravetz 1994) or do not encompass the full range of possibilities (Gross and Jeffries 2001). The identification of emergence, its potential range, the generation of desired emergence and the effect of intervention, in contexts such as CLCs, are still difficult to achieve. Improved domain independent conceptual understanding of emergence and its dynamics at a meta level is required before the pivotal role of emergence can be incorporated into understanding, seeding and intervention of complex social systems such as CLCs. This paper reports on the development of a domain independent meta classification of emergence from examination of well-researched domain specific emergence, which is then used to develop instruments to examine emergence in a specific domain CLCs and the issues encountered. The methodology employed consists of traditional literature review and synthesis combined with focus groups and peer review to assess the validity of the new conceptual model and inform instrumentation requirements analysis. This forms part of a general research strategy being developed which applies the instrumentation developed from the meta classification to investigate emergence in real CLCs, through a case study approach. This latter investigation is reported elsewhere (McDonald 2005b). The ultimate research goal is to develop a framework for understanding, seeding and managing CLCs. The results reported in this paper support the need for new tools and instruments to analyse and seed emergence in social systems such as CLCs and lend credence to the hypothesis that a domain independent meta classification of emergence comprising seven factors will enable appropriate investigatory and seeding frameworks to be developed in domains such as CLCs. Thus, it offers a promising way forward to improve detection and understanding of emergence in complex social systems, capitalising on existing research in other domains. Its novelty is two-fold: (i) it introduces an improved conceptual model of emergence which addresses the fundamental problem of extracting, at a meta level, generalisations of emergence and applying them to gain insight

in a different domain and (ii) it develops instrumentation for investigation of emergence in complex social systems. Future work includes further refinement of instrumentation and its application to other complex systems with the aim of developing understanding of how the various forms of emergence arise, thus ultimately developing an ontology of emergence in complex social systems.

References 1. Cavalletti, M., 2003, 'Creative Communities in Networked Hybrid Spaces'. Art and Consciousness in the Post-biological Era Consciousness Reframed: 5th CAiiA International Research Conference. 2. Daniel, B.K., McCalla, G. Schwier, R. 2003, Social Capital in Virtual Learning Communities and Distributed Communities of Practice. The Canadian Journal of Learning Technology, 29(3), pp113-139. 3. Funtowicz, S. and Ravetz, J.R. 1994, 'Emergent Complex Systems', Futures 26(6) pp560-582. 4. Gross, D. and Jeffries, D., 2001, 'Complexity Beyond Agent Based Models'. Complexity International, 8. 5. Kubik, A., 2003, 'Towards a Formalization of Emergence'. Artificial Life 9(1). 6. McDonald, D.M., 2005a, Complex Learning Communities. Proceedings of eSociety 2005. Qawra, Malta. 7. McDonald, D.M., 2005b. Emergent Characteristics of Complex Learning Communities. Proceedings of Complexity, Science and Society 2005. Liverpool, UK. 8. McDonald, D.M. and McGill, L., 2005, 'eLiteracy as an Emergent Phenomenon'. Proceedings of eLiteracy 2005. Glasgow, UK.

Modeling of the Exocytotic Process by Chemical Kinetic Formalism

Aviv Mezer, Esther Nachliel, Menachem Gutman, Uri Ashery
aviv@hemi.tau.ac.il

The exocytotic process in neurons and neuroendocrine cells consists of a consecutive reactions between well-defined proteins. Even that the partial reactions had been investigated by variety of methods, a comprehensive description, based on chemical kinetics had not been attempted. In the present study, we have, created, a comprehensive kinetic model that reconstructs the physiological process using a standard chemical kinetic formalism. the reactions between the synaptic proteins were transformed into a set of coupled, non-linear ordinary differential equations where the rate constants and some of the reactants concentrations are adjustable parameters. Upon integration over time the reactions equations, reconstructed the experimental signal. The model can perfectly reconstruct the kinetics of exocytosis, the calcium-dependent priming and fusion processes and the effects of genetic manipulation of synaptic proteins. The model suggests that fusion occurs from two parallel pathways and assigns precise, non-identical synaptic protein complexes to the two pathways. In addition, it provides a unique opportunity to study the dynamics of intermediate protein complexes during the fusion process, a possibility that is hidden in most experimental systems. We have used the Genetic Algorithm analysis to achieve high level of accuracy and to find a single global minimum, over a multi dimensional parameter space. Our study demonstrates that complex biological processes can be mathematically modeled and gain high predictive power, up to the level of serving as research tools. It is our intention to expand the model from the level of a comprehensive description of the whole exocytotic process, to the level of cell physiology.

Towards evaluation methodology of p2p systems for complex network management scenarios

Federico Morabito, Giovanni Cortese, Fabrizio Davide
federico.morabito@telecomitalia.it

Network Management In the management domain, we support a vision where information will no longer be exchanged only along a manager-agent hierarchy, (as it is today based on the SNMP and Policy architectures defined by IETF), which represent the standards for monitoring and configuration of networks, respectively. Instead, we build on communicating in a peer-2-peer fashion among nodes. A new monitoring architecture should then support a scalable, data centric communication model, allowing efficient many-to-many information exchange. Also, the information model should be much more flexible than today, allowing peers to communicate even when data schemas have not been agreed in advance as in the currently used standard-based approach. System Model In order to support a flexible management infrastructure, we are working on a distributed system model for data management, built on a peer to peer overlay infrastructure. Current system architecture is based on DHT, and we are designing/ evaluating technologies for Storage/ query including a) algorithms for efficient multi-dimensional range query processing b) algorithms for distributed storage and query of information using richer data models based on ontologies.

Classic approach for p2p system evaluation The performance evaluation is a relevant issue to be integrated in the design and deployment phase of the functional prototype. Proper evaluation mechanisms of non-functional qualities of different designs are needed. The classical approaches for evaluation of p2p system try to measure traditional network parameters, for example link capacity of nodes, physical distance of nodes, packet loss, cost of maintaining the respective overlay structure. The main limitation of the current approach is that they do not show the complex aspects of the system model, but they try to monitor the efficiency from a communications and network point of view. In this sense, some examples of used metrics are average bandwidth overhead (induced for maintain the overlay structure), the latency (dued to the physical topology) and the bandwidth stretch. Furthermore, these metrics do not consider the application scenarios in which the p2p approaches are used. For example, in a network management scenario, an important issue to be measured and monitored is the ability to propagate the information despite of changing pattern of the overlay topology (robustness against dynamics).

Innovative approaches for complex systems Recent works have criticized the intrinsically property of uniform randomness for large complex networks, revealing the hyperlinked structure of the WWW and complex topology of the Internet Over the complex networks, also the applications built using overlay mechanisms, can be treated as a dynamic systems: for example a network management infrastructure over Internet should take in account the dynamics of nodes, data and resources over a complex, wide-area topology and should maintain a statistics overall view of the system behaviour. As result of the integration between complex networks and applications with fast dynamics (for example, a data management application will require data replication mechanisms, data migration techniques), we need innovative approaches and metrics for evaluation, deeply dependent on the applications dynamics and deployment scenarios. For this purpose, we define an evaluation framework composed by a testbed, a prototype for advanced network management and a measurement methodology. The evaluation framework will permit the monitor and measurement of dynamic parameters for investigating robustness property and reliability, using a graph theoretical approach. Some examples of parameters that can be used in the evaluation of the robustness property for a prototype for network management are degree distribution, average path length, clustering coefficient. The proposed methodology will aim at

evaluating some critical parameters, using a statistical load characterization that models the application scenarios. The approach will permit the real-time evaluation of critical parameters for the management applications, bringing together the complex aspects of the underlying network and the dynamics of the specific applications. The testbed will permit the emulation of a distributed, large scale environment and the final deployment in an experimental setup.

A structured approach for modelling of integrated systems in biology

Nicolas Parisey[1*], Marie Beurton-Aimar[12*], S. Randall Thomas[3*]
[*]All authors contributed equally to this work

[1]Mitochondrial Physiopathology Lab. U688 - University Bordeaux 2
[2]LaBRI UMR 5800 - University Bordeaux 1
[3]LAMI UMR 8042 - University Evry

Data and information-flow processes applied to the description of living systems present a new challenge to computational biology. The level of complexity attained by biological knowledge and the development of ever more sophisticated informatics tools for the modelling and simulation of biological processes have revealed the need for methods to structure the manipulated data (where "data" is taken in its most general sense) in a more formal way. Structured approaches like Object Oriented Methods (OOM) allow the definition of data types and relations among these types in order to qualify them semantically. Thus structured, it becomes possible to design computer programs that can automatically interpret the information.

Data types and relations between them can be mapped to arborescences described by a structuration language. Following the emergence of XML (eXtended Markup Language), which is such a language, two XML-languages for biology, SBML (Systems Biology Markup Language) and CellML have attracted a growing community of users in the bioinformatics/ bio-modelling/ systems biology community.

XML provides a means to supplement a text file with any relevant information describing its contents, its origin, its intended use, etc. That is, an XML file is a self-describing ascii file. XML is a structured language with marks (called `tags`) which wrap the file's contents with defining concepts or object classes. SBML [1] and CellML [2] are augmented versions of XML that specify `tags` for descriptions of mathematical models of biological processes and specifically for metabolic networks, kinetic equations, and so on.

The principle of XML is to define a *schema* which specifies the nested `tags` and their hierarchical tree structure. Users are free to define the `tags` as they see fit. However, when a community of users wishes to share files with a common purpose and with a minimum of ambiguity, they agree on a standardized `tag` structure, fixing its definition in a shared formal description, of which there are two standard types: Data Type Definition (DTD) files, and XML Schema. Many thousands of such specifications have been developed for a myriad of applications, of which SBML and CellML are two typical examples. They both

have specific DTD and XML *Schema* definitions, which have gone through several revisions, the latest being available on the respective websites. It is the standardization of these publically available definitions that makes it possible for anyone to design a software program taking one or both of these description formats as input (or producing them as an output file).

As with any standardized language, new situations arise that are not well captured by the current schemas, which leads to continual pressure to extend them to the new situations. Such extensions are a natural part of the evolving use of such systems, but they bring with them the inevitable problem of backward compatibility with previous versions.

Our particular concern is with the modelling of integrated biological systems at not only the cellular and sub-cellular levels (which are essentially the focus of present versions of SBML and CellML) but also of multi-cellular systems, tissues, and even organs. We acutely feel the need to standardize the model-descriptions of such systems using a markup language, but the current versions of SBML and CellML are inadequate to this task. We have thus formed a French Working Group on Markup Languages for Integrated Systems Modelling to characterize these shortcomings and to suggest relevant extensions and improvements to the developers of these two existing MLs (with whom we have established a dialogue).

We present two specific but simple examples drawn from our work (in very different fields) which typify the features we most miss in SBML and CellML: 1) epithelial transport, with transcellular and paracellular transport of water and several solutes; and 2) mitochondrial metabolism modelling. Extension of SBML and CellML to cover these two cases would in fact render them adequate for a much larger class of models which are currently outside their scope.

References

[1] M Hucka and et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, Mar 2003.

[2] Autumn A. Cuellar, Catherine M. Lloyd, Poul F. Nielsen, David P. Bullivant, David P. Nickerson, and Peter J. Hunter. An overview of cellml 1.1, a biological model description language. *SIMULATION: Transactions of The Society for Modeling and Simulation International*, 2003.

Why Does BitTorrent Work So Well?

Simon Patarin and David Hales

Keywords: BitTorrent, computational sociology, altruism

BitTorrent (BT) is currently the most popular file-sharing peer-to-peer (P2P) clients [1]. Several reports have shown that it accounts for the majority of peer-to-peer traffic, up to one third of the global Internet traffic. However, leaving the media attention aside for one moment, we ask the question: Why does BT work so well?

BT attempts to build robustness to freeloading (i.e. downloading without uploading) by implementing a tit-for-tat-like strategy (TFT) within its protocol. It is often believed that this strategy alone is responsible for the high levels of cooperation found. The TFT strategy was championed in Axelrod's book "the evolution of cooperation" [2]. BT works by groups of peers (called swarms) with an interest in downloading a specific file coordinating and cooperating to speed-up the process [3]. Each peer in the swarm stores pieces of the file. Cooperating peers download and upload required pieces. If a peer stops uploading it will tend to be "choked" by other peers, meaning they stop uploading to it. This implements the TFT-like process. So-called "seeder" peers store the whole file: if a swarm contains no seeders, it may lead to a situation in which pieces of the file are missing from the swarm as a whole. Since seeders have nothing left to gain, the system requires some altruistic behavior from peers.

We argue that the TFT strategy is not an adequate explanation of the high levels of cooperation found within BT because: (i) the TFT strategy can be bettered by other less cooperative strategies; (ii) identity can be faked by modifying the client thus circumventing TFT; (iii) unconditional altruism is required for BT to operate in any case. Given that such loopholes exist: why is the system not dominated by freeloaders?

Hypothesis: Group Selection We hypothesize that BT may resist freeloaders and support altruism, at least in part, in a way that has not been previously fully comprehended. Ironically, this process relies on what is commonly believed to be a weakness of BT — the lack of integrated meta-data search. One consequence of this is to partition the BT network into numerous isolated swarms — often with several independent swarms for an identical file — which is one of the necessary conditions for a kind of novel group selective process that has been recently identified in similar simulated systems both in the context of computational sociology and simulated P2P file sharing.

Essentially, if users move between swarms (leave one swarm and enter another) based on the quality of the service they receive, then this process means that swarms containing many freeloaders will tend to "die" as peers leave the swarm for better swarms. Swarms that contain altruists will tend to grow since they support a quality service. Similar models have been advanced in computational sociology [4], [5], [6].

A further implication of the hypothesis is that, given the choice, users may choose unconditional altruism rather than the more restrictive reciprocal approach [7] because of the same group selective process has been shown to select for pure altruism — peers acting for the benefit of the group to their own individual cost.

Conclusions An awareness of some results from social scientific work, particularly the emerging area of computational sociology, can help to inform the analysis of existing working systems "in the wild". This is valuable because such systems currently demonstrate some of the desirable features required of future software systems. Following the empirical line, one way to test our hypothesis would be to implement and distribute a modified BT client that allows users to select pure altruism over the more restrictive reciprocal protocol currently implemented.

REFERENCES

[1] "Bittorrent," official website, http://bittorrent.com/.
[2] R. Axelrod, *The evolution of cooperation*. N.Y.: Basic Books, 1984.
[3] B. Cohen, "Incentives build robustness in bittorrent," in *1st Workshop on the Economics of Peer-2-Peer Systems*, 2003.
[4] D. Hales, "Cooperation without space or memory: Tags, groups and the prisoner's dilemma," in *Multi-Agent-Based Simulation, LNAI*, no. 1979, 2000, pp. 157–166.
[5] R. Riolo, M. D. Cohen, and R. Axelrod, "Cooperation without reciprocity," *Nature*, no. 414, pp. 441–443, 2001.
[6] D. Hales, "From selfish nodes to cooperative networks – emergent link based incentives in peer-to-peer networks," in *P2P2004*, 2004.
[7] R. Trivers, "The evolution of reciprocal altruism," *Q. Rev. Biol*, no. 46, pp. 35–57, 1971.

Simon Patarin and David Hales are with the Department of Computer Science of the University of Bologna, Mura Anteo Zamboni, 40127, Bologna, Italy. e-mail: patarin@cs.unibo.it, dave@davidhales.com

Scaling laws in urban systems (France, South Africa, United States of America)

Fabien Paulus, Céline Vacchiani-Marcuzzo
fpaulus@parisgeo.cnrs.fr

The aim of this poster is to bring evidence about scaling laws in urban systems. Scaling laws are providing a theoretical framework for understanding the dynamics in complex systems. They are used as a methodological device for investigating the location process of activities in the cities. The relationship between urban activities and the sizes of cities where they are located give insights about the interactions between economic activities and dynamics of urban systems.

Contrary to Christallers central place theory explaining the urban hierarchy by population oriented services, we take into account the whole set of urban activities, including specialisation in national or global networks, to build an evolutionary theory on cities. A representation, more suggestive and more precise than rank-size plots, consists to examine the direct statistic relation between the urban activities and the population of each city. Empirical results are obtained from data sets on employment figures in France and USA and number of foreign firms in South Africa. The slope of the adjustment gives a scaling measurement of the dependence of the location of activities to the size of the cities: if the coefficient is equal to 1, it reveals a simple proportionality compared to the population, if it is larger than 1, the relation is superlinear, i.e. there is an effect of selection of the cities, in favour of the largest ones; if the coefficient is lower than 1, the relation, sublinear, indicates on the contrary a selection of the smallest towns. We can still interpret the slope as a measurement of the elasticity of the presence of one kind of activities according to the size of the cities. It is an indicator of the power of the centrality for a function or a given activity, whose significance would approach the concept of economy of agglomeration, or of increasing returns with the size. We prefer an evolutionary interpretation that links the value of the scaling parameter and the stage of specific activities within innovation cycles. Data about evolution of employment on a rather long period (France, 1962-99) confirm the validity of such hypothesis.

Scaling laws in urban systems (France, South Africa, United States of America)

Fabien Paulus, Céline Vacchiani-Marcuzzo
fpaulus@parisgeo.cnrs.fr

The aim of this poster is to bring evidence about scaling laws in urban systems. Scaling laws are providing a theoretical framework for understanding the dynamics in complex systems. They are used as a methodological device for investigating the location process of activities in the cities. The relationship between urban activities and the sizes of cities where they are located give insights about the interactions between economic activities and dynamics of urban systems.

Contrary to Christallers central place theory explaining the urban hierarchy by population oriented services, we take into account the whole set of urban activities, including specialisation in national or global networks, to build an evolutionary theory on cities. A representation, more suggestive and more precise than rank-size plots, consists to examine the direct statistic relation between the urban activities and the population of each city. Empirical results are obtained from data sets on employment figures in France and USA and number of foreign firms in South Africa. The slope of the adjustment gives a scaling measurement of the dependence of the location of activities to the size of the cities: if the coefficient is equal to 1, it reveals a simple proportionality compared to the population, if it is larger than 1, the relation is superlinear, i.e. there is an effect of selection of the cities, in favour of the largest ones; if the coefficient is lower than 1, the relation, sublinear, indicates on the contrary a selection of the smallest towns. We can still interpret the slope as a measurement of the elasticity of the presence of one kind of activities according to the size of the cities. It is an indicator of the power of the centrality for a function or a given activity, whose significance would approach the concept of economy of agglomeration, or of increasing returns with the size. We prefer an evolutionary interpretation that links the value of the scaling parameter and the stage of specific activities within innovation cycles. Data about evolution of employment on a rather long period (France, 1962-99) confirm the validity of such hypothesis.

The importance of parallel and anti-parallel alignment in the collective motion of self-propelled particles

F. Peruani,, M. Baer, A. Deutsch
peruani@mpipks-dresden.mpg.de

Collective motion constitutes a challenging example of self-organization. Applications range from robotics to biology. In particular, emergent large-scale patterns mediated by merely local interactions of individual system components are observed in an amazingly huge number of biological systems of different complexity (e.g. herds, fish schools, bird flocks, swarms of social insects, amoebae and bacteria). Despite the fact that in each case the interactions between individuals are of a different nature, it is possible to determine common requirements for self-organization. Here, we focus on the implications of alignment and analyze 'ferromagnetic' (1, 2) and 'liquid crystal'-like (3) alignment mechanisms. The latter allows parallel and anti-parallel alignment, while the former only admits parallel alignment. For the two types of alignment mechanisms, we observe phase-transitions related to the noise introduced into the system. In the case of parallel and anti-parallel alignment, we also observe spontaneous symmetry breaking of the rotation symmetry. We present numerical evidence of this phase transition. In addition, we show that the effect of attractive and repulsive forces lead to different responses depending upon the type of alignment mechanism.

1) T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. 75, 1226 (1995) 2) G. Gregoire and H. Chate, Phys. Rev. Lett. 92, 25702 (2004) 3) F. Peruani, M. Baer and A. Deutsch (unpublished)

Keyword: particular complex systems

Strong emergence in a population of agents

Denis Phan, Jean-Louis Dessalles
jean-louis.dessalles@enst.fr

Characterising emerging properties is a crucial challenge when studying complex systems. The problem is especially challenging in case of "strong emergence" (Müller 2002), when the agents involved in the emerging phenomenon are able to detect it and are influenced by it. Strong emergence is often observed in populations of human agents. The identification of emerging phenomena by the agents themselves induces a feedback from the observation to the process. As there is a coupling between the process level (the system) and the observation level due to the agents themselves, emergence is immanent in the system. In order to simulate cases of strong emergence, the simulated agents must be given means to detect the collective phenomena of which they are part. Our claim is that algorithmic complexity may provide such means.

Various characterisations of emergence have been proposed (see Bonabeau et al., 1995 for a review). One definition seems to capture the common intuition underlying most accounts: emerging phenomena correspond to a sudden decrease of relative algorithmic complexity (RAC) (Bonabeau Dessalles 1997). If S is the current state of the system and Di is a set of available detection tools, then emergence occurs whenever the RAC, i.e. the shortest description of S using a subset of Di, paradoxically decreases when the subset is augmented.

Ct+dt(S | D1, ..., ,Dk-1, Dk) $<$ Ct(S | D1, ...,Dk-1)

Taking an additional description tool Dk into account should normally increase the complexity, as Dk must be computed on S. However, when some Di are redundant with Dk, the complexity of the minimal description may drop by an amount that characterises the quality of the emergence.

To simulate strong emergence, agents must be given ways to assess RAC. In this paper, we explore such a possibility in an experiment about the emergence of classes (Axtell, Epstein, Young 2001). In that experiment, agents play an iterative Nash bargaining tournament, where they may make high, low or medium claims. When agents are assigned arbitrary tags such as green and red, the system may evolve in such a way that a dominant high-claiming class, the red, say, emerges together with a submissive low-claiming class, the green. This is not a case of strong emergence, though, as agents have no way to shift to a binary class-based decision when choosing a claim.

We designed an extension of Axtell et al. experiment, where individuals display a combination of T tags and may include the tags observed on the opponent in their decision to play high, low or medium. Agents are biased to favour simplicity, i.e. they base their decision on tags whenever some simple tag combination offers a more concise description of the behaviours hold in memory. Not only is the emergence of classes preserved, as in Axtell et al. experiment, but also the decision by individual agents to shift to a class-based decision offers a genuine case of strong emergence.

Our model may offer a new insight into how abstract entities like group or collective behaviour should be understood. Individuals are expected to assign behavioural characteristics to collective entities, though such entities may have no central control, as soon as it reduces the algorithmic complexity of their own decision process. Thanks to this hypothesised bias for simplicity, such decisions behave like self-realising anticipations.

References

Axtell, R., Epstein J. M. Young H. P. (2001). "The emergence of classes in a multi-agent bargaining model". In: S. N. Durlauf H. P. Young (Eds), Social Dynamics. Cambridge, MA: MIT Press, ed. 2004, 191-212. http://www.brookings.edu/es/dynamics/papers/classes/classes

Bonabeau, E., Dessalles J-L. Grumbach A. (1995). "Characterizing emergent phenomena (1): a critical review". Revue Internationale de Systémique 9(3), 327-346.

Bonabeau, E. Dessalles J-L. (1997). "Detection and emergence". Intellectica 25(2), 85-94. http://www.enst.fr/ jld/papiers/pap.cogni/98072401.html

Müller, J-P. (2002). Des systèmes autonomes aux systèmes multi-agents: interactions, émergence et systèmes complexes. Rapport présenté pour l'obtention de l'HDR.

Complexity, Networks and the Modernization of Antitrust

Cristina Poncibo'
cristinaponcibo@tin.it

Title: Complexity, Networks and the Modernization of Antitrust

Author: Cristina Poncibò, Department of Economics and Legal Studies, University of Eastern Piedmont, Alessandria, Italy Lagrange post-doc scholarship, CRT Foundation, Turin, Italy, e-mail: cristinaponcibo@tin.it

Key Words: Complexity Science, Antitrust, Networks

Abstract:

One perspective for understanding competition that has garnered increased attention by those in antitrust is the field of science known as complexity science. Incorporating insights and relying on metaphors from population ecology, evolutionary biology, systems theory, chaos and the study of networks, the science of complexity attempts to describe and explain how systems and their occupants, including industries and firms, evolve and compete against one another over time through adaptation, co-evolution and other dynamic processes. Insights from complexity science are also being applied to describe and better understand the evolution and competition taking place in various sectors and industries of the new economy. Although the integration of knowledge gained through the study of the science of complexity to antitrust remains to be more fully developed, initial findings have yielded some provocative insights. For example, networks may be different from antitrust markets, encompassing multiple markets or even industries. This challenges standard neoclassical antitrust economics, whose focus is the relevant market, where the primary test of membership is substitutability. Within some networks, huge market power may be accrued, even when the size of the dominating firm is relatively small within the system, a possibility that cuts against the concentration thesis of neoclassical antitrust economics, which only recognizes market power in the context of very large market shares. The network phenomenon calls attention to the importance to competition of relationships of power and influence that tend to be more familiar to business people and political scientists than to neoclassical economists. The question of whether complexity science can contribute to antitrust has also been brought into sharp focus through the workings of the antitrust modernization in order to explore whether industries involving significant technological innovation should be treated differently under the antitrust laws and, whether there are features of the modern economy that warrant special antitrust treatment - whether harsher or more lenient.

Complexity in Neuroscience: How to relate the Digital aspects of Brain function with the Analog-driven Mind Processes?

Walter Riofrio
Neuroscience and Behavior Division, Laboratory of Pathology,
Faculty of Veterinary Medicine and Animal Sciences.
Universidad Peruana Cayetano Heredia.
E-mail: wriof@terra.com.pe

Abstract

If we consider that mind (especially, consciousness and self-reference) is more than a collection of mechanisms (we are not just robots or zombies), this question arise: what more "it is added" at the brain function (the mechanisms) to produce the mental phenomena?

The so called "hard problem" in the studies about the consciousness, alert us that always we are confronted with the subjective component of it. But, if we are not so ambitious in the scope of our investigations, the possibility of success it may be more probable.

The dynamical development of living forms, with the management of matter and energy from the environment, will generate the basic forms of transmitted information processes.

In this respect, the emergence of mental properties and processes from the normal activity of the brain will be a central focus of our presentation. So, with this clarification, we propose an investigation about the most fundamental forms of mental processes.

An important issue, it is to ask how the brain structures the data that comes from external and internal to organism. And with these has the possibility or capacity to represent information.

In this respect our first question is: Which are the conditions that make it possible the emergence of "meaning" or "semantic" in the information coming from the data received in the brain?

In our presentation, we centre in a region of information management that code analogic signals to digital signs: auditive perceptions.

With the final purpose to open the debate about if neuron are able to pass the analogic information to digital information and again to analogic information; and incorporate the 'meaning' capable to be expressed in an integrated spatial-temporal network organizations by neurons (for us, they certainly do continuously).

Evolving cell phone families

José Salgado, Ricardo Gama
miguelsal@ipg.pt

Cell phone networks is usually refered as a good example of a social complex system. Indeed, we are inquiring students, between 12 and 18 years old, and using their cell phone contact numbers and their cell phone contact statistics to meet the following investigation questions: - how does the cell phone network evolve with age? - Is there a peak of frequent contacts? That is, does students contact more people as they become old? - Does the "cell phone family" becomes regional with age? That is, we expect that students first contacts are local (students from the same school) and then with increasing age and social interactions they become more scatered. - Does the networks go beyond the portuguese borders? What are the countries that these networks reach? The scope of this project goes beyond this questions. In fact we would like to study later on the role that time plays in the network of each individual. This study may be crucial to the understanding of the portuguese society structure strenght and the understanding of the youth portuguese place in europe. Although our study is not finished yet, we would like to present our results.

Comparison between parallel and serial dynamical behaviour of boolean networks.

Lilian Salinas, Eric Goles.
Deparment of Mathematics Engeneering, University of Chile.
Casillla 170 correo 3, Santiago, Chile.
lsalinas@dim.uchile.cl, secrepriv@conicyt.cl

Abstract

In this article we study some aspects about the parallel and serial dynamical behaviour of a boolean network.

We study the relationship between the structure of the associated graph and the attractors of a boolean network. We show that for boolean networks with an arborescent associated graph, the length of the dynamical cycles are a power of two and under certain conditions the attractors are only fixed points.

Furthermore, we study the robustness of the dynamical behaviour of a boolean network with an arborescent associated graph, for serial and parallel updates. We show that, if we choose a good order of update then the dynamical behaviour is the same in both cases, and if the associated graph does not have loops the converse is also true.

On the other hand, we are interested in the robustness of the set of attractors of a boolean network for serial and parallel updates. It is well known that the set of fixed points of a boolean network is the same for both types of update. In this article we prove that a dynamical cycle of a boolean network without loops with parallel update is not a dynamical cycle for the same network with serial update and viceversa.

Keywords: Boolean network, attractor, robustness, parallel update, serial update.

Comparison between parallel and serial dynamical behaviour of boolean networks.

Lilian Salinas, Eric Goles.
Deparment of Mathematics Engeneering, University of Chile.
Casillla 170 correo 3, Santiago, Chile.
lsalinas@dim.uchile.cl, secrepriv@conicyt.cl

Abstract

In this article we study some aspects about the parallel and serial dynamical behaviour of a boolean network.

We study the relationship between the structure of the associated graph and the attractors of a boolean network. We show that for boolean networks with an arborescent associated graph, the length of the dynamical cycles are a power of two and under certain conditions the attractors are only fixed points.

Furthermore, we study the robustness of the dynamical behaviour of a boolean network with an arborescent associated graph, for serial and parallel updates. We show that, if we choose a good order of update then the dynamical behaviour is the same in both cases, and if the associated graph does not have loops the converse is also true.

On the other hand, we are interested in the robustness of the set of attractors of a boolean network for serial and parallel updates. It is well known that the set of fixed points of a boolean network is the same for both types of update. In this article we prove that a dynamical cycle of a boolean network without loops with parallel update is not a dynamical cycle for the same network with serial update and viceversa.

Keywords: Boolean network, attractor, robustness, parallel update, serial update.

Tracing Experience as a potential support for meaning negociation between human and computer agents

Arnaud Stuber, Salima Hassas, Alain Mille
CexAS Team – LIRIS
Université Claude Bernard Lyon 1
Bâtiment Nautibus – 8 boulevard Niels Bohr
69622 Villeurbanne
email: {arnaud.stuber,salima.hassas,alain.mille}@liris.cnrs.fr

In this study, we consider a group of human actors, who work together to accomplish a collective task. These actors use a common computer environment as a medium for their cooperation to manage reports of their activities in the context of their collective task.
The general goal of the project is to help each actor to achieve his/her task by providing a contextual access to his/her individual experience, to the experience of peer actors, and finally to the collective experience.
The specific problem treated in this paper is to assist one actor to reuse his/her past experience in the context of his/her current task. Individual experience is represented by individual use traces, which are observations of the real usage situated in the activity. This assistance requires some common meanings between the actor and the system in order to allow their mutual understanding.
The fundamental issues are to model the experience captured by observing actors' actions in the environment, and to have a way to reuse this experience with a minimal cost for an actor. A first way to ease information retrieval is to personalize the assistance. However it is always limited by the numerous interactions which are required to the actor, before the assistant starts retrieval. Our proposal is to use the principles of the emergence of language to achieve this negociation, and by this way to take into account past negociations. Here, the emergence of language is chosen to obtain a shared communication system between one actor and his/her personal assistant. By the same way, this system could be extended to the communication between personal assistants in order to assist cooperation.
To construct individual experience bases, we use the trace-based reasoning paradigm, an extension of the case-based reasoning paradigm to unstructured cases where problem and solution descriptors are not defined beforehand. Traces are captured by observing users' actions, and represented in the system by use traces; a reasoning by analogy is done to retrieve past traces. A use trace is an alternate sequence of state and transition, describing respectively the stable state of the environment and the action performed by the actor.
In order to allow the emergence of meaning in the traces, we apply the mechanisms of emergence of language considering the interactions between agents. An *alter ego* agent is associated to each actor to manage the interactions; it has to assist the reuse of the individual experience, to allow experience sharing and exchange and to participate in the emergence of a collective experience. The interactions are then grouped in two categories : between one actor and his/her assistant, and between actors mediated by the system, and thus between alter ego agents. The system we consider is an hybrid system, and our paper focuses on the subsystem constituted of an actor and his/her assistant. The same principles could be applied to the system as a whole.
To allow emergence of language, a symbolic representation of the use traces is required. A formal grammar for use traces defines the elements of this representation; associated to a given abstraction, they constitute the patterns for the language. The meanings of the language are the meanings of the assistance, and they are defined using the patterns. During actor's actions, some pattrens are recognized in the current trace, and a ascendant process allows to finally identify some meanings. The recognized meanings are used to retrieve past traces, which are proposed to the actor. The actor can interact with these proposals, and these interactions constitute some support for negociations, which are taken into account to perform the language games necessary to emergence of language. By this way, patterns and meanings are modified and can be reused in forthcoming negociations.
A prototype is presented to illustrate our proposal, some experiments are to come.

Complex Systems Perspectives and Inter-Discilinary Curriculum-a real challenge and opportunity for the Romanian Higher Education System

Marta-Christina Suciu
suciuchristina@yahoo.com

Keywords organised within a particular complex systems-education: global knowledge-based society;complex systems; interdisciplinary curriculum.

Main Topic :education

The global knowledge-based society and creative economy are very dynamic and increasingly complex systems. Change seems to be the single constant of the universe. Connectivity and networks of interactive entities are more and more relevant for this new type of society called also a network society. Complex systems questions ask for new ways to manage a s̈uper-complexs̈ociety as knowledge-based society is also called.Traditionally ätoms;̈ individuals and tangible things were the key determinants of wealth. Networks involve interactions and connectivity and focus more on intangible assets such as: intellectual capital, knowledge, innovation and creativity considered new sources of wealth in the context of knowledge-based society and creative economy.

Complex and adaptative systems have to learn constantly. Traditionally adaptative systems deal with n̈egative feed-back;̈ complex adaptative systems in a network society seem to ask for p̈ositive feed-backs̈ince the systems might function more effectively in a process of long-life-learning; society itself is considered to be a l̈earning society.̈ The complex systems approach provide a solid foundation for understanding better the complexity of the knowledge-based society. There are multitude of problems that have to be addressed mostly in a country like Romania. It becomes more difficult to manage complexity applying traditional ways of learning seen mostly as convergent ways.

In the highly complex education system there may be various combinations of the different approaches to teach using a mix of convergent and divergent teaching strategies, methods and styles. Traditionally there is a tendency to teach s̈tep by step;̈ p̈iece by piece;̈ b̈it by bitj̈ust partially information related to a particular disciplines. By applying the perspective of complex systems in education there is a strong need to focus more on interdisciplinary and cross-disciplinary approaches. The curriculum can be integrated-mostly in the higher education system-around topics that reflect the patterns, interactions, and interdependencies of different fields. This is a real challenge and a great opportunity for Romanian higher education system!

Competitive Adaptive Lotka-Volterra Systems with Complex Behavior

Claudio Tebaldi
claudio.tebaldi@polito.it

We study a NxN dimensional Lotka-Volterra system which describes competition among N species and includes behavioral adaptation, i. e. a learning mechanism. The existence of reduced models, where N appears as a parameter, is discussed, depending on the level of symmetry of the system. Such models give full account of equilibria and their stability in the complete system and are effective also in describing the time dependent regimes, even chaotic ones, for a large range of parameter values. Relevant questions, as species survival/exclusion and different kind of coexistence (stable equilibria, periodic oscillations, strange attractors, sincronization) are addressed focusing on the role of adaptation

Functioning-dependent structures

Michel Thellier, Camille Ripoll, Patrick Amar, Guillaume Legent, Vic Norris

Abstract

Numerous studies have shown that proteins involved in metabolic or signalling pathways are often distributed non-randomly as multimolecular assemblies. Such assemblies range from quasi-static, multi-enzyme complexes (such as the fatty acid synthase or the (-oxo acid dehydrogenase systems) to transient, dynamic protein associations. Multi-molecular assemblies may comprise proteins but also nucleic acids, lipids, small molecules and inorganic ions. Such assemblies may interact with membranes, skeletal elements and/or cell organelles. They have been termed *metabolons*, *transducons* and *repairosomes* in the case of metabolic pathways, signal transduction and DNA repair, respectively, or, more generally, *hyperstructures*.

Although channelling is sometimes challenged, most authors have assumed that, in many molecular assemblies, intermediates are channelled from each enzyme to the next without diffusion of the intermediates into the surrounding cytoplasm. Potential advantages of channelling are:

1. reduction in the size of the pools of intermediates (a point however contested by some authors),
2. protection of unstable or scarce intermediates by maintaining them in a protein-bound state,
3. avoidance of an "underground" metabolism in which intermediates become the substrates of other enzymes and
4. protection of the cytoplasm from toxic or very reactive intermediates.

In the case of transient, dynamic multi-molecular assemblies, certain only form in an activity-dependent manner due, for example, to an association between enzymes that only occurs when they are engaged in transporting or transforming substrates or transducing a signal. We have proposed to term *functioning-dependent structure* (FDS) any such type of dynamic, multi-molecular structure; in other words, an FDS assembles when functioning and disassembles when no longer functioning and thus is created and maintained by the very fact that it is in the process of accomplishing a task.

The formation of an FDS adds protein-protein interactions to the classical protein-substrate (or protein-signal) interactions and is therefore likely to generate novel kinetic behaviours. Given the extremely diverse mechanisms at work in the variety of biological systems quoted above, it is difficult to devise a single approach that would be appropriate to model them all. Therefore, in the following, we begin by determining whether the metabolite-induced association of two enzymes into an FDS (a so-called metabolite-induced metabolon) may, under steady-state conditions, confer to the overall FDS kinetic features that the individual enzymes do not have; this particular two-enzyme FDS is fairly straightforward to model, which is why we have chosen it. We then discuss how the type of approach used here to analyse the metabolite-induced metabolon may be used to analyse other varied assemblies. Finally, we speculate on the relevance of such concepts to the debate on the nature of life.

A Simulation Environment for Emergent Properties

Heather R. Turner, Susan Stepney, and Fiona A. C. Polack
Department of Computer Science, University of York, Heslington, York, YO10 5DD, UK
{turner,susan,fiona}@cs.york.ac.uk
This research is funded by the EPSRC and BAE Systems as part of a CASE Studentship.

Abstract: We propose a *multi-layer* architecture for simulating emergent properties. This is implemented as a form of cellular automaton at the lowest layer, with *mobile processes* to represent objects at multiple upper layers. This architecture supports *multiple levels of emergence.*

1 Introduction

Many definitions of emergence share the theme of existence of distinct *levels*, between which the emergence occurs. They might correspond to, for example, a change of spatial or temporal scale, be used only to simplify a description of behaviour, or they may be essential in *identifying* emergent properties. Regardless of the form of the emergence, levels always seem to be present, and can lead to a hierarchical structure of emergence.

Researchers tend to focus on the generation of emergence in a particular environment, or on simulating a particular natural behaviour. We provide a broader simulation architecture to study emergence, and emergence hierarchies as phenomena in general, whilst also permitting specific, constrained, simulations.

2 Overview of Architecture

The architecture is based on a series of communicating layers of abstraction. The lowest layer is implemented as a variant cellular automata (VCA), that need not be strictly finite, regular, and/or deterministic. This layer encodes environmental information, and captures absolute spatial reference. It forms the substrate of the emergent architecture. In keeping with similar work by Capcarrere [1], we call this the *environmental layer.*

The next layer is constructed from *mobile processes*, communicating with other processes in the same layer, the layer below, and the layer above, via mobile channels. We implement these mobile processes using features of the occam-π language [2]. This *primary mobile layer* uses rules without explicit spatial reference; it relies on *relative* spatial information transmitted from the lower layer. Additional layers of mobile processes can be added, either corresponding to levels in the emergence hierarchy, or purely for implementation convenience.

In their most basic form, the mobile processes simply *tag and track* emergent features at the lower levels, allowing the emergent features to be treated as objects in their own right at the higher level. But they may additonally encompass some of the logic of the simulation, providing information and instruction to lower level processes.

This use of higher level structuring allows us to engineer *specific* emergent behaviour. The system is designed with higher level objects, then rules are systematically *migrated* down to lower level objects. By migrating rules to the lowest VCA layer, we can remove the higher layers and achieve a primitive mono-layer system for implementation.

3 Case Studies

We are now using our occam-π simulation environment to study some simple models: a basic 1D model of platelet flow and blood clot formation [3]; emergent properties in a 2D environment, incorporating rule combination and migration [4]; unbounded growth of diffusion limited aggregation.

References

[1] Capcarrere MS: An evolving ontogenetic cellular system for better adaptiveness. *BioSystems* **76**:177–189, 2004.

[2] Barnes FRM, Welch PH: Communicating mobile processes. *Communicating Process Architectures*. IOS Press, 2004.

[3] Polack FAC et al.: An architecture for modelling emergence in CA-like systems. In *ECAL 2005*, LNCS. Springer, 2005.

[4] Turner HR, Stepney S: Rule Migration: Exploring a design framework for modelling emergence in CA-like systems. In *Workshop on Unconventional Computing*, ECAL 2005.

Long Abstract: Uprichard, E. and Byrne, D.

Studying complex social change:
Linking levels and meaning through adult and child personal reflections

Complex social systems are recognized as multi-level and multi-dimensional entities. However, understanding the ways in which the multiple multi-dimensional levels interact together and produce social change at the local level is methodologically problematic. So far, the dominant mode of representation of social systems as complex systems has been through iconic modelling. Iconic models are founded on algorithms that derive complex emergence from the iterative progression of either non-linear equation sets or 'game' rules in simulations. However, whilst such approaches certainly have their uses, they do not address crucial aspects of the processes of complex causation in social systems. There are three main reasons for this. First, they have a limited capacity for capturing generative dynamics that occur across the multiple hierarchical levels. Most models only account for two or three levels whereas aggregate social form is derived from the multiple nonlinear interactions of multiple levels; the interactions are interdependent across and between levels. If we are to construct representations of the social then somehow, we also need to construct models that acknowledge multiple non-linear multi-level interactions.

Second, whilst many models are based on empirical data, few use this data directly within the model and instead take the form of hypothetical predictive propositions about particular systems rather than working with what is known to be real about the system. We stress the importance of exploring modes or representations in which actual data is directly integrated. Traditionally, data would be understood in this sort of context as the quantitative products of measurement processes. We absolutely agree that social measurements are relevant but we think that measurements alone will not resolve the issue. Third, measured data and the structures of explanation we construct with it are not adequate either as account of the nature of complex systems incorporating human agency or as modes of establishing such accounts. We also need data forms, which can convey meaning and the potential for social action. Narratives have this capacity.

Hence, this paper argues that personal narratives are a way of responding to this methodological challenge because narratives allow access to the multi-level and multi-dimensional meaning that underlies human agency. This argument is illustrated in relation to studying urban change - where cities and urban regions are understood as complex systems - by drawing on first-hand empirical examples from Byrne and Doyle's examination of people's understanding of 'post-industrial transformation' in South Shields (UK) from the past through the present to the future, and Uprichard's work on children's understanding of York (UK) and Dijon (France) in the present and the future. In turn, the paper proposes that the use of adult and child personal narratives must be part of the methodological repertoire of approaches to studying and understanding complex social change.

Analysis of branched-chain amino acid biosynthesis by a Thomas network approach

Aurélie Urbain, Pierre Renault, S.Dusko Ehrlich, Jean-Michel Batto

INRA - Institut National de la Recherche Agronomique

Jean-Michel Batto, UR895 – Domaine de Vilvert 78352 Jouy-en-Josas, FRANCE
Email : jean-michel.batto@jouy.inra.fr

Keywords : Subsystem comparisons, Network design, ILV regulation, Prokaryotes

Cellular growth depends upon the ability of the cell to synthesize new molecules from nutrients present in the environment. It includes a lot of different processes such as increase in cell mass, catabolism and anabolism of various components, duplication of the bacterial chromosome as well as transcription and translation of genes. During the phase of exponential growth, cells divide at a constant rate depending on availability of nutrients and on their ability to form new membrane and proteins. Because of the complexity of biology encompassing all these interactions by a holistic approach is a difficult task. For example, *Bacillus subtilis* biosynthesis of amino acids requires 120 genes [1].
Our goal is to better understand biological mechanisms involved in cell growth by building a predictive computational model. So we prefer to adopt a step-by-step approach and work on targeted mechanisms. We focused on the prediction of one biological mechanism in one species, hoping to achieve a general understanding of the mechanism (holding true for many, and ideally all, species). Thus, we are working on a local/transversal approach with the aim of combining genetic and enzymatic levels of regulation in order to describe production of essential aminoacid. Our model aims to describe the production of Isoleucine Leucine and Valine (ILV) during growth of the widely studied bacteria *B.subtilis*, *L.lactis* and *E.coli*.
We are attempting to draw a first draft of a common regulation pattern in the absence of the requested wet lab data, using a regulation network modeling. In fact, we can build a topological view of interaction for the studied regulation, but we lack in-vitro view of the metabolite-protein interactions. As the topological model can only support a set of untested hypotheses, we use the formal logic of René Thomas [2-4] to describe the differential equations governing the studied process. However, the space of possibilities is huge. For example a model with 4 interactants and 12 direct activation and feedback regulations yields to a set of 1022 networks. To reduce this space we use a model checker (NuSMV [5]) associated with SMBioNet [6] to represent the global regulatory network as a graph.
The reduction of the space is based on a systematic validation of each graph. A selected biological regulatory graph would describe a true regulation network as it corresponds to a biological fact. In the prospect of describing the differences between the selected networks, we have designed a protocol that allow the classification of the set of networks. We used the length of the graph, the existence of a dead-end and the number of loops for traveling between René Thomas state's. Preliminary results show a correlation between length of the networks and the growth rate. Our next objective is to test variants of a studied formal model and deduce biological property of the subset we study.
This approach should be relevant to a number of complex systems. As it combines the contribution of different levels of regulation, our *in silico* model will be a fast and easy-to-use tool to test hypothesis for wet lab biologists at the early stage of their research projects.

1. *Mader, U., et al., Transcriptome and proteome analysis of Bacillus subtilis gene expression modulated by amino acid availability. J Bacteriol, 2002. 184(15): p. 4288-95.*
2. *Thomas, R. and M. Kaufman, Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits. Chaos, 2001. 11(1): p. 180-195.*
3. *Thieffry, D. and R. Thomas, Qualitative analysis of gene networks. Pac Symp Biocomput, 1998: p. 77-88.*
4. *Thomas, R., Logical analysis of systems comprising feedback loops. J Theor Biol, 1978. 73(4): p. 631-56.*
5. *A. Cimatti, E.C., F. Giunchiglia and M. Roveri, NuSMV : a new symbolic model verifier. Computer science, 1999. 1633: p. 495-499.*
6. *Bernot, G., et al., Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. J Theor Biol, 2004. 229(3): p. 339-47.*

The Uncertainty in Modelling Complex Systems

Steve Whittle, Eshan Rajabally, John Dalton, Simon Snape
s.whittle2@lboro.ac.uk

One of the great challenges in complex system development is the control of design properties throughout the lifecycle. Modelling plays a key role in this activity but current approaches do not adequately support an integrated approach. In particular, there is a failure to provide proper traceability of design properties throughout the product breakdown structure. Furthermore, there is inadequate management of the impact of uncertainties in modelling activity throughout the lifecycle. A system model may give accurate results but absolute fidelity cannot be demonstrated for most practical applications due to our incomplete understanding of the world. Typically, certifying that a model is appropriate for a specific purpose is an informal, ad hoc procedure. This paper will explain the benefits of formally capturing this procedure and present the need for the development of a decision support capability to advance the effective modelling of increasingly complex systems. The technical aspects of the developed methodology will be introduced describing the modelling framework representation, which enables model and design decision traceability and the capture of emergent property flow, and a Bayesian Belief Net based methodology to capture the reasoning associated with justifying the dependability of models.

Furthermore, the link to Design Margins and Critical Design Features (A sub-set of top level product attributes / requirements that are agreed as being critical to project success) is explained and an approach to determine and manage both will be introduced.

Coordinated Action of a Large Scale Robotic System through Self-Organizing Processes

Steffen Wischmann and Martin Hülse

Fraunhofer Institute for Autonomous Intelligent Systems,
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
{steffen.wischmann, martin.huelse}@ais.fraunhofer.de

Abstract. Within the context of *Artificial Life* and *Evolutionary Robotics* we are facing the problem of how to coordinate the behavior of large scale distributed robotic systems ($n > 100$). Following the ideas of decentralized control and self-organizing processes we present an example of how to implemented such principles in a homogeneous group of robots controlled by artificial recurrent neural networks (RNN). It is important to emphasize that we are aiming at situated and embodied real world robots acting in dynamically changing environments.

The behavior of the described overall system is determined by three kinds of *structural coupling*. Every agent within the group is controlled by a RNN. By following a modular neurodynamics approach three evolved basic behaviors are encoded within the RNN that (i) have to be coupled in a way to maintain stable behavioral patterns which (ii) have to be robust against dynamically changing environmental conditions. Furthermore, agents are able to interact with each other through acoustic signaling. Hence, due to this local interaction, each agent (iii) is coupled to others in its surrounding in a way that it can influence the behavior of others as well as itself can be influenced by others. Here, we want to focus on the latter aspect. On a macroscopic level we show how a group of robots coordinates individual foraging and homing behavior in order to transfer collected energy to a nest while avoiding obstacles. Within the control architecture of each single robot a neural pattern generator is implemented that is able to generate very low frequent internal rhythms (period lengths up to two million time steps are possible). This pattern generator determines whether the robot displays foraging or homing behavior. Each robot is able to communicate its behavior switching through acoustic signaling that in turn can reset the pattern generator of perceiving robots. To some extend this reset of internal neural oscillators through external stimuli is inspired by biological systems, for example the flashing of fireflies during mating.

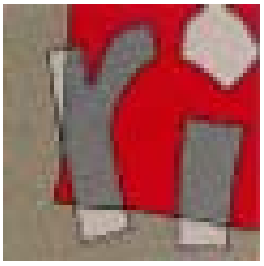
We show that it is possible to almost perfectly synchronize the behavior of a population containing up to 150 autonomous robots through simple local interactions, whereby at the beginning their behavioral patterns are completely out of phase. Furthermore we can demonstrate that is also possible to get individual rhythms in-phase even if the period length of individual internal rhythms are different among the population.

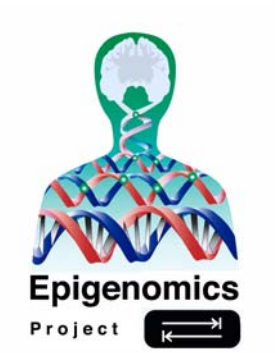
Incremental and unifying modelling formalism for biological interaction networks

Anastasia Yartseva, Hanna Klaudel, Francois Kepes
iartseva@gmail.com

This article introduces a new unifying incremental formalism for the modelling of biological regulatory networks. The semantics of the model is given through the translation into the multivalued logical formalism. A methodology for constructing such models is presented on a classical benchmark: the lambda phage genetic switch system. The main advantage of the presented approach is to provide an intermediate level of representing the biological data about interaction networks. From one side, this level enables the expression of the knowledge in a form intuitive for a biologist. From the other side, the knowledge is represented as a formal structured model. This formal model can be translated automatically in existing formalisms enabling the complex study of the dynamical properties of the biological system.

Index





Supported by the Future and Emerging Technologies programme of the Information Society Technologies programme of the European Commission.