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

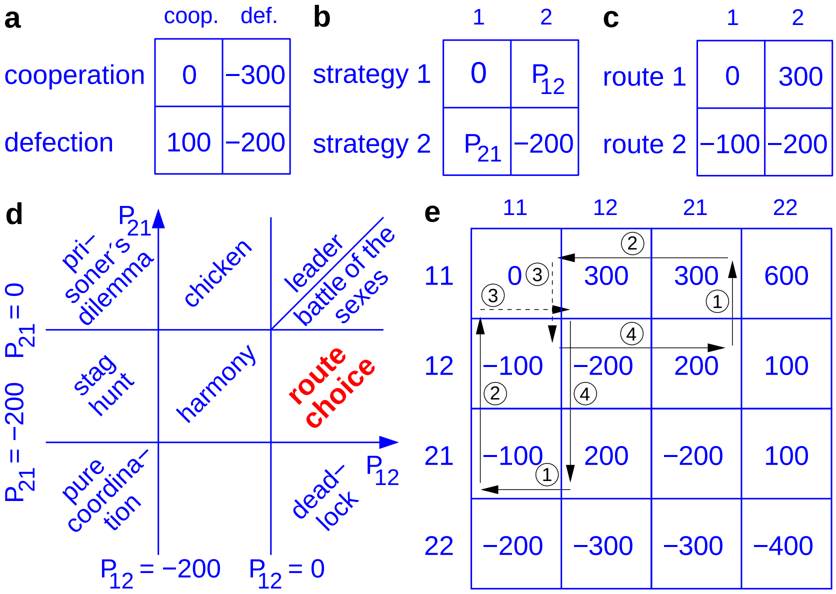

Figure 1: Symmetrical two-person games can be represented by a payoff matrix of the form $\mathbf{P} = (P_{ij})$, where P_{ij} is the success ("payoff") of person 1 in a one-shot game when choosing strategy i and meeting strategy j. The respective payoffs of the second person are given by the symmetrical 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.

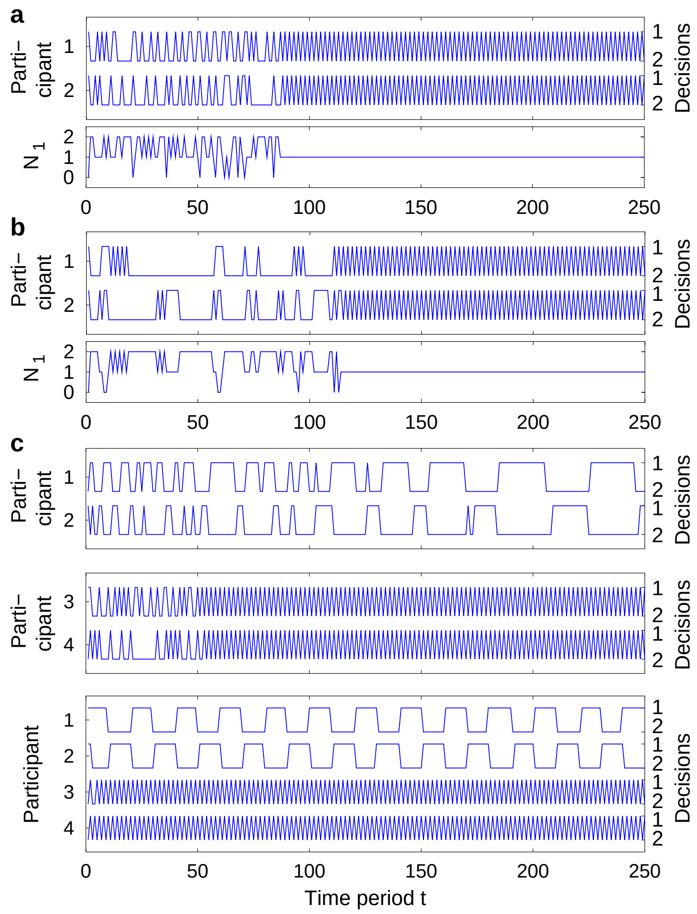

Figure 2: (a) Experimentally observed 1- and 2-decisions of both individuals in a two-person route choice experiment with the parameters specified in Fig. 1c, and corresponding number N_1 of 1-decisions. The system optimum is reached for $N_1 = 1$, the user equilibrium for $N_1 = 2$. Despite the initial 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

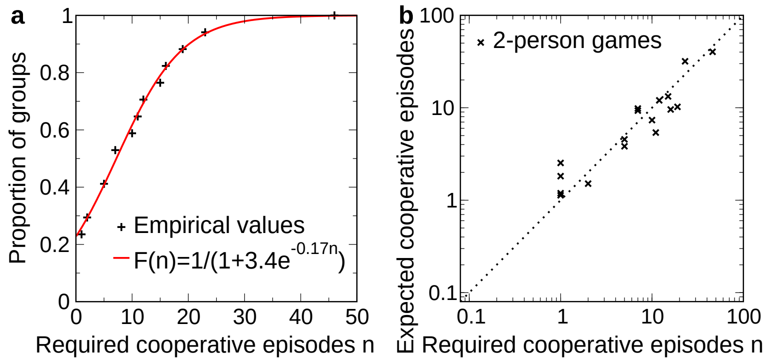

Figure 3: (a) Cumulative distribution of required cooperative episodes until persistent cooperation is established, given that cooperation occurs within 300 time periods (as in 17 out of 24 two-person experiments). The experimental data are well approximated by a logistic curve. (b) Comparison of the 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.

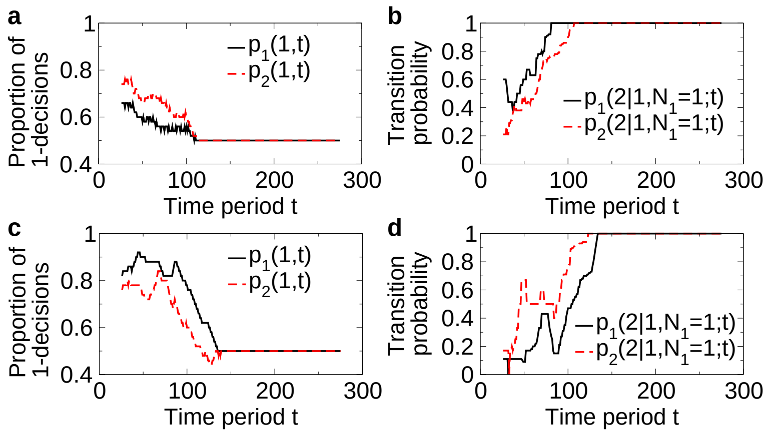

Figure 4: (a) Proportion $p_l(1,t)$ of 1-decisions of both participants l in the two-person route choice experiment displayed in Fig. 2a. (b) Transition probability $p_l(2|1,1;t)$ of person l from route 1 (the "freeway") to route 2, when the other person has chosen route 2, averaged over a time window 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.