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.

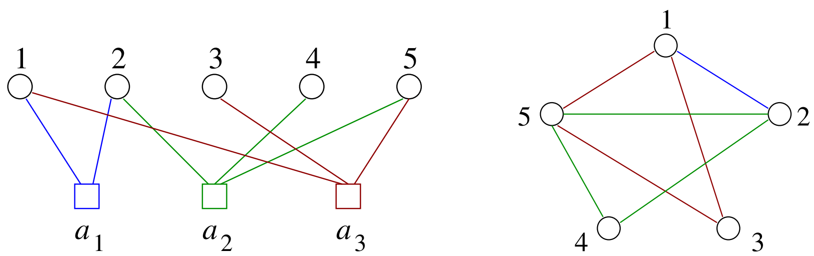

Figure 1: Factor-graph and conflict-graph representation of a combinatorial auction. Circles are bidders, squares present items.

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.

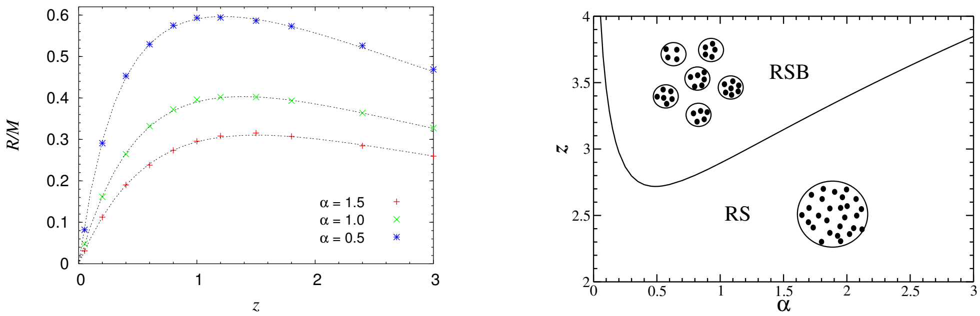

Figure 2: *Left:* Maximized revenue for different values of $\alpha = M/N$, as a function of the average package size z. Analytical results are compared to simulated annealing on random CA instances. Prices are fixed to $\nu_i = 1$. For small z, many goods are not contained in the packages, for high 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).