Phase transition and finite-size scaling in the vertex-cover problem

Alexander K. Hartmann [a,*] Wolfgang Barthel [a] Martin Weigt [b]

[a] *Institut für Theoretische Physik, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany*

[b] *ISI Foundation, Viale S. Severo 65, 10133 Torino, Italy*

Abstract

NP-complete problems play a fundamental role in theoretical computer science. Recently phase transitions were found in such problems, when studying suitably parameterized random ensembles. By applying concepts and methods from statistical physics, it is possible to understand these models and the phase transitions. Here, we consider the vertex-cover problem, one of the six "basic" NP-complete problems. We describe the phase transition and the running time of an exact algorithm around the phase transition. To connect to transitions in physical systems, we apply finite-site scaling and we look for a correlation-length like quantity.

Key words: Combinatorial Optimization, Phase Transitions, NP-complete
PACS: 64.60.-i, 05.20.-y, 02.10.Eb

---

1 Introduction

During the last three decades, more and more problems in physics are treated by applying optimization algorithms[1]. This growth is mainly due to the strong increase in computer power and because of the rapid developments of new sophisticated optimization algorithms. This means that physicist are applying tools from computer science to study physics problems. Recently, knowledge has been transferred also from physics to computer science. One famous example is the invention of the simulated annealing algorithm [2], which allows to solve optimization problems heuristically, by interpreting the function to be optimized as an energy, introducing an artificial temperature and cooling the system down using a Monte Carlo simulation.

* Corresponding Author:
Email address: hartmann@physik.uni-goe.de (Alexander K. Hartmann).

More recently, concepts and methods from statistical physics have been applied to study "classical" problems from theoretical computer science. The most interesting problems are those from the class of *NP-complete* problems [3]. For these problems so far only algorithms have been invented having *worst case* running time which increases exponentially with the size of the problems, hence one calls these problems *hard*. On the other hand, up to now nobody is able to prove that no faster algorithm exists, i.e. an algorithm having a running time growing only polynomially (P) with the system size. This is the famous *P-NP problem*. In order to solve the P-NP problem, one has to understand the source of hardness of such problems.

However, in many cases one is interested in a (suitably defined) *typical* behavior, instead of the worst case behavior. To elucidate typical cases, one has started to study suitably parametrized random ensembles of problems. For these ensembles phase transitions in the solvability have been observed [4], and the hardest to solve instances are often located right at the phase boundary. Hence, to understand the source of hardness, one can study the problems occurring near phase transitions. In this context also analytical tools borrowed from statistical mechanics are very valuable. Combinatorial problems are mapped to equivalent physical models, and methods like the replica trick or the cavity approach [5] are used to calculate and analyse phase diagrams. One main observation is that computational hardness is typically related to the glassiness of the corresponding physical system.

Here, to introduce to this field and to illustrate some crucial phenomena, we present the vertex-cover problem. In the next section, we define the problem. Then, we demonstrate the phase transition and the corresponding change in the running time. Finally, we present new results on the use of critical exponents to describe the phase transition and on the existence of a correlation-length like quantity for this phase transition.

2 The vertex-cover problem

Let $G = (V, E)$ be a graph, V being the set of vertices and $E \subset V^2$ the set of undirected edges. In particular we study an ensemble of Erdös-Rényi random graphs [6], where each graph has $N = |V|$ vertices and $M = cN/2$ randomly selected edges, c being the average connectivity of single vertices. Note that at $c = 1$ a percolation transition occurs.

Let be $V' \subset V$ a subset of vertices of size $X = |V'|$. We call an *edge* $\{i, j\}$ *covered*, iff $i \in V$ or $j \in V$. We call the *graph* G covered iff *all* edges are covered by V' and in this case V' is a vertex cover. The vertex-cover problem (VC) is: *For given graph G and given value X, exists there a vertex cover of size X?* A related optimization problem is: for given graph G and given value X, what is the minimum *energy*, defined as the number of uncovered edges?

3 The phase transition

One can use an exact algorithm [7] to solve the related optimization problem. If the minimum number of uncovered edges is zero, the graph is coverable by X vertices.

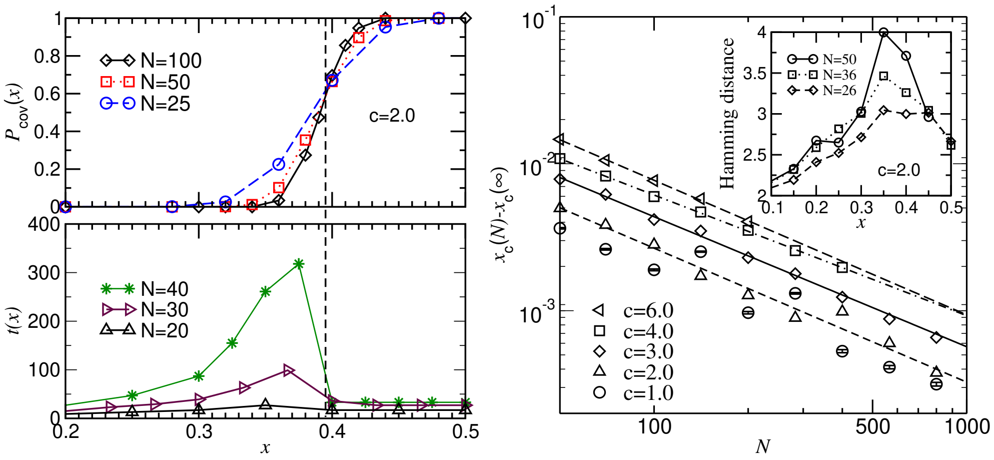

Fig. 1. Left top: Probability that a random graph of average connectivity $c = 2$ has a vertex cover of relative size $x = X/N$. Left bottom: Running time of an exact algorithm to solve VC with given value of x. Right: Finite-size scaling behavior of the critical cover size. Right inset: scaling of the correlation volume as a function of x for different sizes. Error bars are at most of the order of the symbol size.

In Fig. 1, we show the numerically determined probability that a random graph with $c = 2.0$ is coverable using a fraction $x = X/N$ of all vertices. One observes a strong increase near $x_c = 0.3919$. This increase sharpens for larger N, and all curves cross close to x_c. Hence, there appears to be a *phase transition* in the solvability of VC at a critical value x_c. Furthermore, the running time of the optimization algorithm is shown. One can clearly see a peak close to the phase transition, i.e. there the hardest instances are located.

The function $x_c(c)$ determines the phase boundary of the problem. For small connectivities $c \leq e \approx 2.71828$, the phase boundary can be calculated [7] exactly in the replica-symmetric framework. All the exponentially many different minimal vertex covers are collected in one cluster [8]. For larger connectivities this is only an approximation and many solution clusters exist. Interestingly, one can use the phase diagram to calculate analytically the running time of a simple version of the optimization algorithm [9].

4 Finite-size scaling

To elucidate whether the cov-uncov transition has more in common with a physical phase transition, we study the finite-size scaling behavior of the critical value $x_c(N)$

as a function of system size, by fitting it to functions of the form $x_c(\infty) + aN^{-b}$. Such behavior one finds for the critical temperature T_c near second order phase transitions, where the critical exponent of the correlation length determines the shift of T_c. And indeed, for $c > 1$, we find a good scaling behavior, see Fig. 1, while for $c = 1$, at the percolation transition, the data scatters strongly. For $c > 1$, we get exponents $b(c = 2) = 0.91(9)$, $b(3) = 0.88(4)$, $b(4) = 0.82(4)$, and $b(6) = 0.92(11)$. Hence, within error bars, this exponent seems to be universal.

One is tempted to ask: is this scaling behavior connected to a divergence of a correlation length-like quantity, as in conventional second order phase transitions? Since VC is a mean-field model, it does not make sense to speak of length scales. Instead, we study the *correlation volume*, which we define as the minimum number of vertices which changes, if we force one vertex to change its state in a ground state. Indeed, the correlation volume exhibits a peak near the phase transition, see Fig. 1. Unfortunately, calculating the correlation volume requires to enumerate all ground states, so we cannot treat large systems. Hence, using small sizes, we are not able to draw a precise connection to the exponent $b \approx 0.86$ measured above.

References

[1] A. K. Hartmann, H. Rieger, Optimization Algorithms in Physics, Whiley-VCH, Berlin, 2001.

[2] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671.

[3] M. R. Garey, D. S. Johnson, Computers and intractability, Freeman, San Francisco, 1979.

[4] S. Kirkpatrick, B. Selman, Critical behavior in the satisfiability of random boolean expressions, Science 264 (1994) 1297.

[5] M. Mézard, G. Parisi, M. A. Virasoro, Spin glass theory and beyond, World Scientific, Singapore, 1987.

[6] B. Bollobas, Random Graphs, Academic Press, 1985.

[7] M. Weigt, A. K. Hartmann, The number guards needed by a museum – a phase transition in vertex covering of random graphs, Phys. Rev. Lett. 84 (2000) 6118.

[8] W. Barthel, A. K. Hartmann, Clustering analysis of the ground-state structure of the vertex-cover problem, preprint cond-mat/0403193.

[9] M. Weigt, A. K. Hartmann, The typical-case complexity of a vertex-covering algorithm on finite-connectivity random graphs, Phys. Rev. Lett. 86 (2001) 1658.