Random constraint satisfaction problems: a point of view from physics

Guilhem Semerjian

LPT-ENS

01.09.2010 / Journées MAS

Outline

Constraint satisfaction problems : definitions

n variables $\quad \underline{x} = (x_1, \dots, x_n) \in \mathcal{X}^n \quad$ discrete alphabet $\mathcal{X}$

m constraints $\quad \psi_a(\{x_i\}_{i \in \partial a}) = \begin{cases} 1 & \text{satisfied} \\ 0 & \text{unsatisfied} \end{cases}$

solutions $\quad \mathcal{S} = \{\underline{x} : \psi_a(\underline{x}_{\partial a}) = 1 \ \forall a\}$

- decision problem, is $|\mathcal{S}| > 0$?
- counting problem, what is $|\mathcal{S}|$?
- optimization problem, what is $\max_{\underline{x}} \left[\sum_a \psi_a(\underline{x}) \right]$?

Constraint satisfaction problems : examples

- $\mathcal{X} = \{\text{True}, \text{False}\}$, ψ_a depends on k variables $x_{i_a^1}, \dots, x_{i_a^k}$
  - $\psi_a = \mathbb{1}(z_{i_a^1} \vee \cdots \vee z_{i_a^1} = \text{True})$, with $z_i \in \{x_i, \overline{x}_i\}$

    k-satisfiability problem
  - $\psi_a = \mathbb{1}(x_{i_a^1} \oplus \cdots \oplus x_{i_a^k} = y_a)$, with $y_a \in \{\text{True}, \text{False}\}$

    k-xor-satisfiability problem
- $\mathcal{X} = \{1, \dots, q\}$, $\psi_a(x_i, x_j) = \mathbb{1}(x_i \neq x_j)$
  on the edges $a = \langle i, j \rangle$ of a graph

  q-coloring problem

Worst-case complexity of the decision problem:

- *k*-xor-satisfiability easy for all k
- *k*-satisfiability, *q*-coloring difficult for $k, q \geq 3$

Random constraint satisfaction problems

What about their "typical case" complexities ?

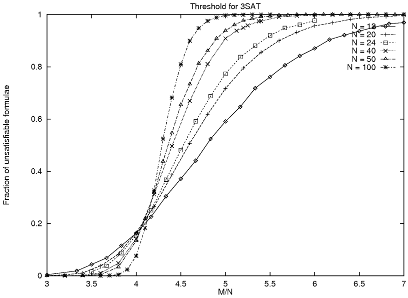
"typical"= with high probability in some random ensemble of instances

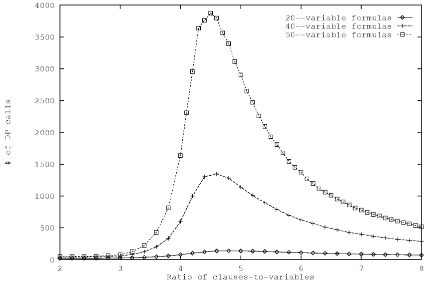
Examples :

- coloring Erdös-Rényi random graphs $G(n, m)$
  choose m edges uniformly at random in the $\binom{n}{2}$ possible ones
- random (xor)satisfiability ensembles
  choose m hyperedges (k-uplets of variables), among $\binom{n}{k}$

Most interesting regime : $n, m \to \infty$ with $\alpha = m/n$ fixed

Phase transition for the unsatisfiability probability :





associated to a peak in the hardness of solving

A few rigorous results for random k-satisfiability and q-coloring :

- existence of a sharp threshold $\alpha_s(k)$ [in fact $\alpha_s(k, n)$] [Friedgut]
- upper and lower bounds on $\alpha_s(k)$
  [Chao and Franco, Frieze and Suen, Achlioptas, Dubois et al]
- asymptotics of $\alpha_s(k)$ at large k [Achlioptas, Moore, Naor, Peres]

But :

- no precise value of $\alpha_s(k)$ for small k
- unsatisfactory understanding of algorithmic difficulty at $\alpha < \alpha_s(k)$

Why physics ?

Statistical mechanics :

- configuration space $\underline{x} = (x_1, \dots, x_n)$
- energy function $E(\underline{x})$
- temperature T
- Gibbs-Boltzmann distribution $\mu(\underline{x}) = \exp[-E(\underline{x})/T]/Z$

Low-temperature statistical physics $\approx$ combinatorial optimization

randomness in the distribution of instances $\approx$ disordered systems

Outcomes of the physics approach

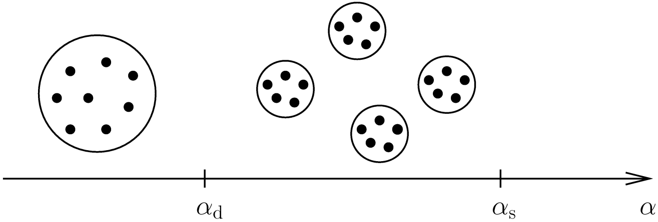
- quantitative estimation of $\alpha_s(k)$
- refined picture of the satisfiable phase
- analysis of known algorithms
- suggestion of new ones

Refined picture of the satisfiable phase

Exponential number of solutions for $\alpha < \alpha_s$, $\sim \exp[ns(\alpha)]$

Clustering transition at another threshold $\alpha_d < \alpha_s$:

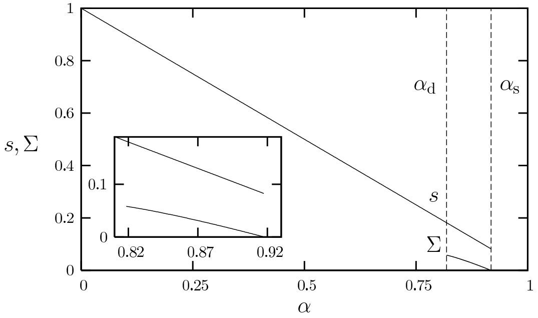
apparition of an exponential number of clusters ($\sim \exp[n\Sigma(\alpha)]$),
each containing an exponential number of solutions ($\sim \exp[n s_{\rm int}(\alpha)]$)



Refined picture of the satisfiable phase

Can be proven rigorously for random xorsat [Creignou, Daudé] [Cocco, Dubois, Mandler, Monasson] [Mézard, Ricci-Tersenghi, Zecchina]

At $\alpha_{\rm d}$, apparition of a 2-core in the hypergraph



$$s(\alpha) = \Sigma(\alpha) + s_{\rm int}(\alpha)$$
$$\Sigma(\alpha_{\rm s}) = 0$$

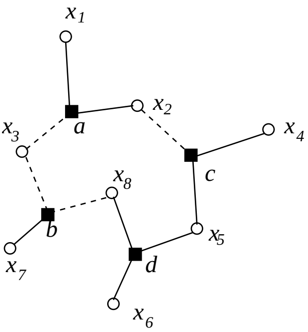
[more complicated picture for sat and col]

Methods

If the formula F has solutions,

define $\mu(\underline{x}) = \frac{1}{Z} \prod \psi_a(\underline{x}_{\partial a})$ uniform measure on $\mathcal{S}$

Factor graph representation of a formula :



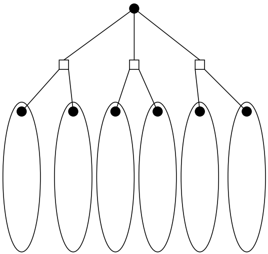
Crucial property : in the $n, m \to \infty$ limit with $\alpha = m/n$ fixed
local convergence of the factor graph to a random Galton-Watson tree

Methods

For a tree factor graph $\mu(\underline{x})$ is a rather simple object

(Belief Propagation is exact)

All marginal probabilities can be easily computed recursively :



Sparse random graphs converge locally to trees

Is it enough for their $\mu(\underline{x})$ to converge locally to the measure on the associated tree ?

It depends... (on the correlations decay)

Methods

- Yes in "replica symmetric" cases

  - Ferromagnetic Ising models — [Dembo, Montanari]
  - Matchings — cf J. Salez's talk
  - Random CSP for $\alpha < \alpha_{\rm d}$

  Allows to compute in particular $\lim \frac{1}{n} \log Z$ entropy of solutions

- No in presence of "replica symmetry breaking" $(\alpha_{\rm d} < \alpha < \alpha_{\rm s})$

  Configuration partitioned in clusters $\Rightarrow$ long-range correlations

  But with $\mu(\underline{x}) = \sum_c w_c \, \mu_c(\underline{x})$, each μ_c has short-range correlations, can be treated as above

  Properties of w_c encode the value of $\Sigma(\alpha)$, hence $\alpha_{\rm s}$

Short bibliography

- M. Mézard, G. Parisi and R. Zecchina, Science **297**, 812 (2002)
- S. Mertens, M. Mézard and R. Zecchina, Random Struct. Alg. **28**, 340 (2006)
- F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian and L. Zdeborova, PNAS **104**, 10318 (2007)
- F. Altarelli, R. Monasson, G. Semerjian and F. Zamponi, in *Handbook of Satisfiability*, IOS press (2009), `arxiv`:0802.1829
- M. Mézard and A. Montanari, *Information, Physics, and Computation*, OUP (2009)