Glass phases in Optimization Problems

Marc Mézard, LPTMS, Université de Paris Sud, Orsay, France

14th April 2003

Collaborators:

- R. Zecchina (ICTP Trieste)
- G. Parisi (Univ. Roma 1)

+ A. Braunstein (Trieste); G. Biroli (Saclay); S. Franz (Trieste); F. Ricci-Tersenghi (Roma); M. Weigt (Göttingen);

Combinatorial optimization

- Configurations : N discrete variables (e.g. Boolean $s_i \in \{0,1\}$)
- Energy (cost) function $E(\mathcal{C})$; typically computable in $\sim N^b$ operations.
- Optimization Pb: Find $\mathcal{C}^*$ which minimizes $E(\mathcal{C})$.
  Evaluation Pb: Find the cost $E(\mathcal{C}^*)$.
  Decision Pb: Is there a $\mathcal{C}$ with $E(\mathcal{C}) < E_0$?

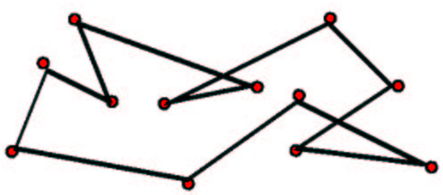
Examples:

- Travelling Salesman Problem
- Assignment
- Spin glass
- Eulerian circuit
- Hamiltonian cycle
- Colouring
- Satisfiability
- ...

'Instance' = 'Sample'

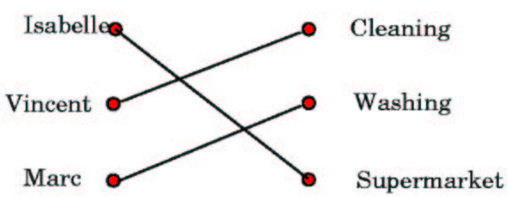
Examples

1) Travelling Salesman Problem



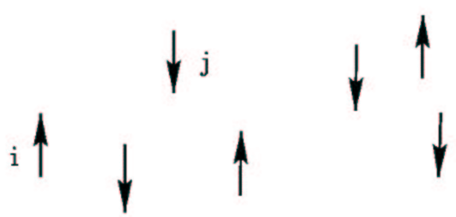
N points
$\mathcal{C}$ = tour; $(N-1)!/2$
$E(\mathcal{C})$ = length

2) Assignment



N persons, N jobs
$\mathcal{C}$ = assignment; $N!$
$E(\mathcal{C}) = -\sum$ affinities

3) Spin glass



N spins
$\mathcal{C}$ = spin configuration; 2^N
$E(\mathcal{C}) = -\sum J_{ij}\sigma_i\sigma_j$

Classification: computational complexity

Worst case analysis of decision problems ;

Main Complexity Classes:

P = polynomial $\leftrightarrow$ tractable, $t < N^\alpha$. Ex: Assignment, Eulerian circuit, Spin glass in $d = 2$, RFIM,...

NP = non-deterministic polynomial (A 'yes' solution can be checked in polynomial time) $\leftrightarrow$ many problems!

NP-complete: the hardest **NP** problems. Problem A is **NPC** iff all problems in **NP** are polynomially reducible to it. (If A is solvable in polynomial time, all problems in **NP** are solvable in polynomial time).

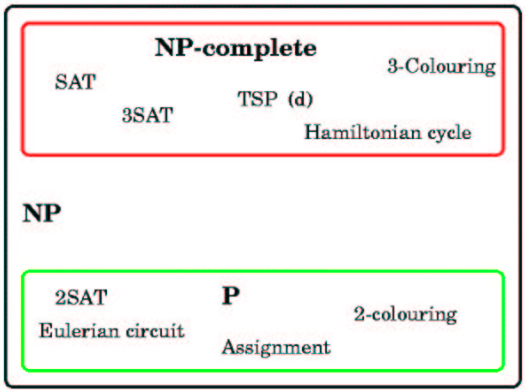
Theorem (Cook, 1971): The SATISFIABILITY problem is NP-complete.

Other NPC: 3SAT, TSP, Hamiltonian cycle, Spin glass in $d \geq 3$,..

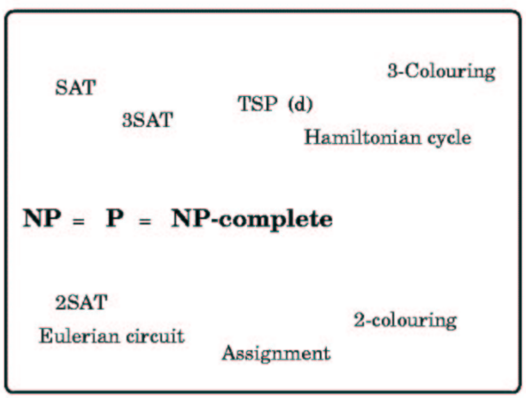
P = NP ?

Complexity map

Conjectured ($P \neq NP$):



Possible ($P = NP$):



SATISFIABILITY

Example:

..a theatrical director feels obligated to cast either his ingénue, Actress Alvarez, or his nephew, Actor Cohen, in a production. But Miss Alvarez won't be in a play with Cohen (her former lover), and she demands that the cast include her new flame, Actor Davenport. The producer, with her own favors to repay, insists that Actor Branislavsky have a part. But Branislavsky won't be in any play with Miss Alvarez or Davenport.[] Is it possible to satisfy the tangled web of conflicting demands? (from G. Johnson, The New York Times 1999).

N Boolean variables: $x_i \in \{0, 1\}$ $\quad i = 1, \ldots, N$

M constraints =clauses like $x_1 \vee x_{27} \vee \bar{x}_3, \quad \bar{x}_{11} \vee x_2, \; \ldots.$

Pb: is there a choice of the Boolean variables such that all constraints are satisfied (SAT)? Generic: conjunctive normal form.

$$(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_2) \wedge (\ldots)$$

Worst-case vs. Typical-case

Computational complexity = worst case analysis.

Experimental complexity = typical case analysis: → class of samples (probability measure on instances).

Ex 1: CuMn at one percent Mn. Properties of the generic sample with $N \gg 1$ variables?

Ex 2: Complexity of the random 3SAT problem. Three variables per clause, chosen randomly in $\{x_1, .., x_N\}$, negated randomly with probability $1/2$:

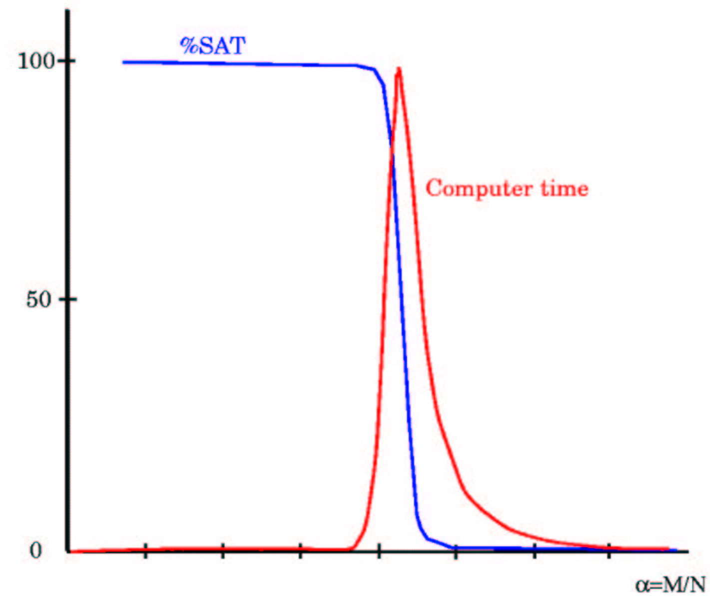
$$(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_3 \vee x_2) \wedge \ldots \wedge (x_9 \vee \bar{x}_8 \vee \bar{x}_{30})$$

Control parameter: $\alpha = \frac{M}{N}$= Constraints/Variables.

Numerically: Threshold phenomenon at $\alpha_c \sim 4.26$.

Mitchell Selman Levesque Kirkpatrick Crawford Auton...; Friedgut Kaporis Kirousis Lalas Dubois Boufkhad...

Threshold phenomenon → Phase transition, complexity much bigger near to the phase transition.



- Easy, and generically SAT, for $\alpha < \alpha_c$
- Hard, in the region $\alpha \sim \alpha_c$
- Easy, and generically UNSAT, for $\alpha > \alpha_c$

Typical-case complexity and Statistical Physics

Two types of questions:

- Properties of a generic sample: Phase diagram? (Theory + experiment)
- Algorithms: improvement near phase transitions?

A first example of application: Simulated annealing

Boltzmann probabilities:

$$P(\mathcal{C}) = \frac{1}{Z} e^{-\beta E(\mathcal{C})}$$

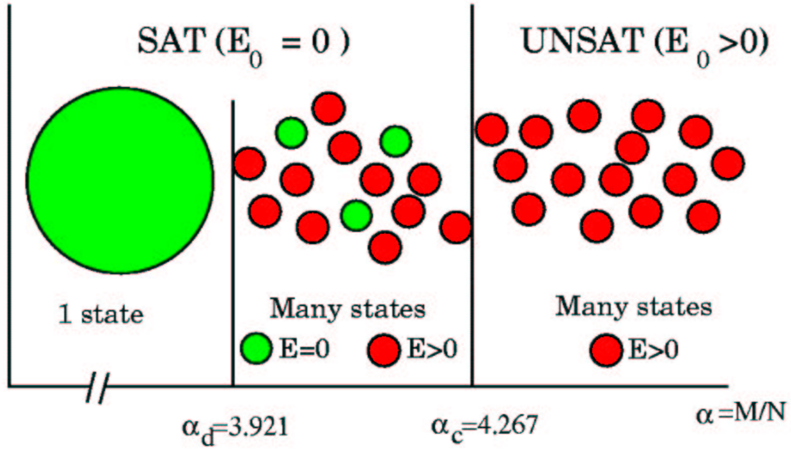
Optimization = Find ground state ($\beta \to \infty$). Particularly difficult close to a $T = 0$ phase transition.

β finite: generalized problem, number of configuration at a given energy; entropy, free-energy, phase transitions etc... Useful from the algorithmic point of view (simulated annealing+...). Useful from an analytic point of view...

Statistical physics of the random 3SAT problem

Monasson, Zecchina, Biroli, Weigt, ..., MM, Parisi, Zecchina: → Phase diagram + New algorithm.

1- Analytic result: Three phases, clustering phenomenon



2- A new class of algorithms

In the Hard-SAT phase $\alpha_d < \alpha < \alpha_c$: how to handle the proliferation of metastable states with $E > 0$: a new message passing algorithm: "survey propagation"

States and complexity

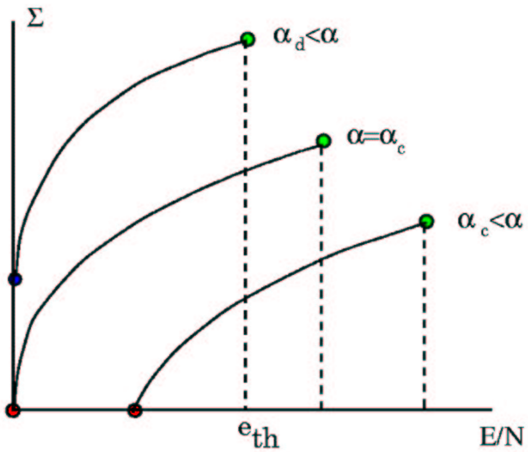
Minimum Energy Cconfigurations: energy cannot be lowered by a finite numberof flips

State: { MEC connected by finite flips }. (Rigorous study on "XORSAT": precise algorithmic definition of states).

Proliferation of states: At $\alpha > \alpha_d$, exponentially large number of states:

$$\mathcal{N}(E) \sim \exp\left(N\Sigma\left(\frac{E}{N}\right)\right)$$

Qualitative behaviour of the complexity Σ:

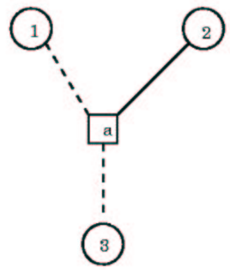


Main steps

- Graphical representation: 3SAT as a random graph
- Elementary message passing procedure (Bethe approximation): an algorithm when there is a single state.
- The many states situation: cavity method $\rightarrow$ passing surveys of elementary messages.

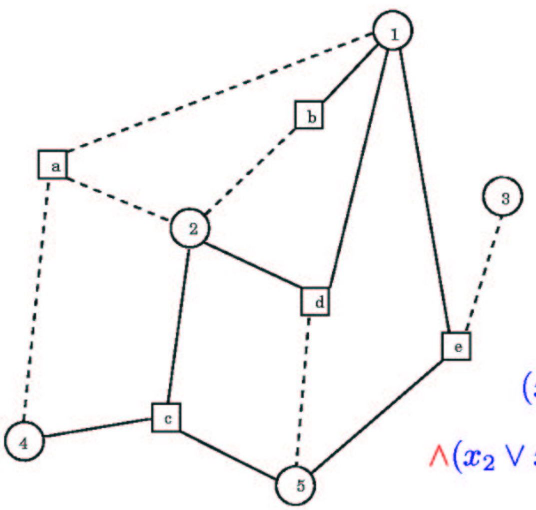
Graphical representation:

One clause a:



Boolean: $\bar{x}_1 \vee x_2 \vee \bar{x}_3$

Ising Spins: $E_a = \frac{1-s_1}{2} \frac{1+s_2}{2} \frac{1-s_3}{2}$



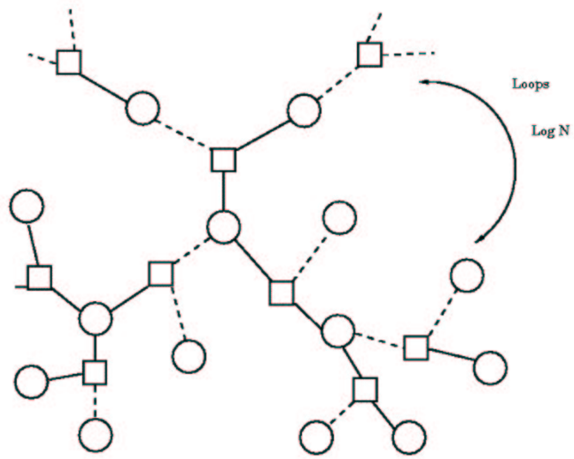
$(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_4) \wedge (x_1 \vee \bar{x}_2)$

$\wedge (x_2 \vee x_4 \vee x_5) \wedge (x_1 \vee x_2 \vee \bar{x}_5)$

$\wedge (x_1 \vee \bar{x}_3 \vee x_5)$

Geometry: tree-like structure

$\rightarrow$ 3-spin interactions on a random hypergraph



Locally tree-like, but loops of order $\log N$.

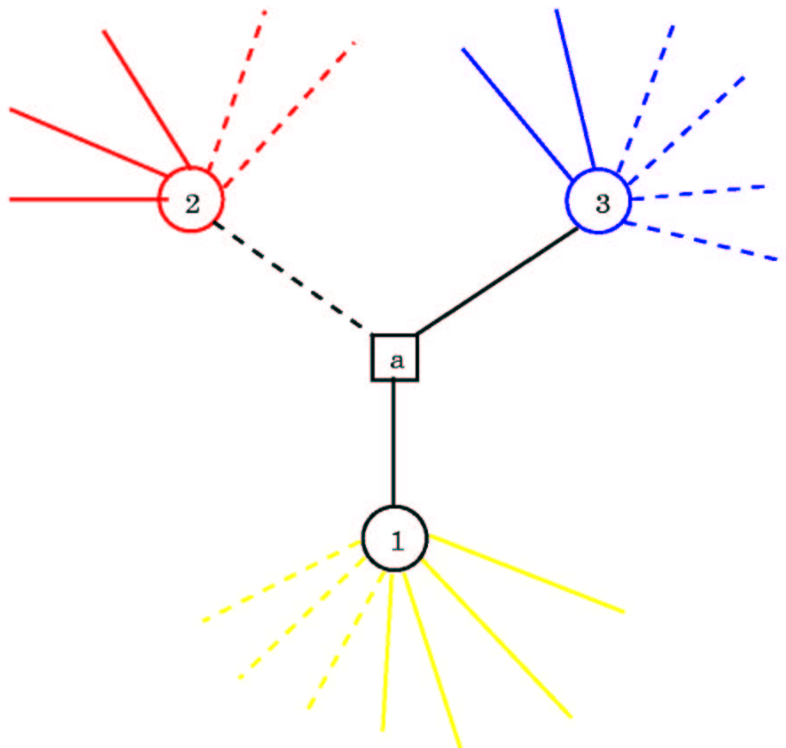
Proba(var. with connectivity$=k$) $= \frac{(3\alpha)^k}{k!} e^{-3\alpha}$.

$\rightarrow$ Possible to use iterative methods (Bethe lattice) to solve the statistical physics problem.

$\rightarrow$ Possible to use iterative methods (message passing) as an algorithm.

Simple message passing: belief propagation

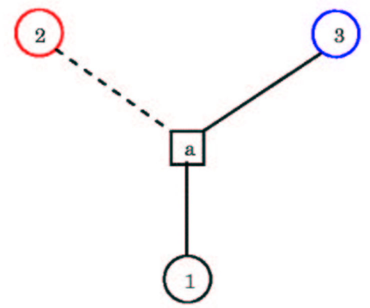
Message= warning sent from a clause to a variable:



Converges and gives the correct answer on a tree: SAT iff no contradictory message

Statistical physics analogue

Bethe approximation:



$$P(s_1, s_2, s_3) \propto e^{-\beta\left[E_a(s_1,s_2,s_3) - h_1^{(a)} s_1 - h_2^{(a)} s_2 - h_3^{(a)} s_3\right]}$$

$$P_{cavity}(s_1) \propto \sum_{s_2,s_3} e^{-\beta\left[E_a(s_1,s_2,s_3) - h_2^{(a)} s_2 - h_3^{(a)} s_3\right]}$$

$$\rightarrow P_{cavity}(s_1) \propto e^{-\beta\left[u_{a\rightarrow 1}\left(h_2^{(a)}, h_3^{(a)}\right) s_1\right]}$$

Proliferation of states

Belief propagation (Bethe approximation) works if one can neglect the correlations between the input fields (tree).

Random 3SAT "locally tree-like": generically, s_2 and s_3 are very far away (distance $O(\log(N))$) $\to$ Uncorrelated if there is a single pure state.

$\to$ OK for $\alpha < \alpha_d$ (Easy-SAT phase).

$\to$ Wrong in the Hard-SAT: Proliferation of local ground states, stable to finite number of spin flips.

State proliferation hypothesis:

$$\mathcal{N}(E) \sim \exp\left(N\Sigma\left(\frac{E}{N}\right)\right)$$

$\Sigma(e)$: complexity.

From belief propagation to survey propagation

Message with many states= Survey of the elementary messages in the states of energy density e:

$$Q_{a\to 1}(u) = C^t \sum_{\alpha} \delta\left(u^{\alpha}_{a\to 1} - u\right) \delta\left(\frac{E^{\alpha}}{N} - e\right)$$

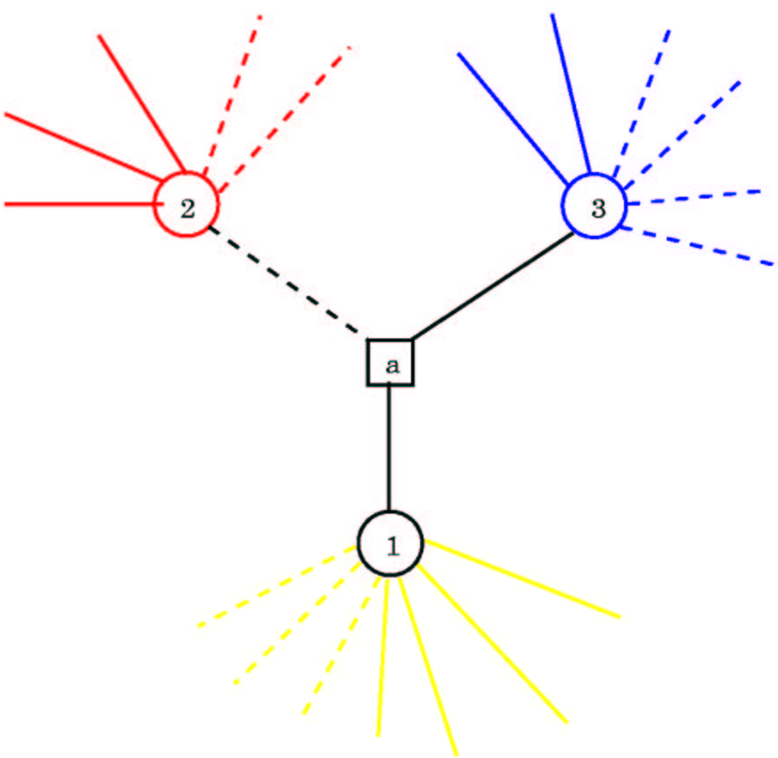
Propagate the surveys along the graph. Converges!

$\to$ Results on the phase diagram and the complexity (generic sample), but also information on a single sample.

$$Q_{a\to 1}(u) = (1 - \eta_{a\to 1})\,\delta(u) + \eta_{a\to 1}\delta(u-1)$$

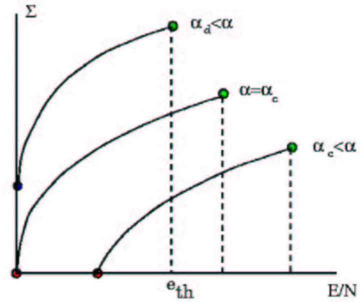
parametrized by a single number $\eta_{a\to 1}$ = probability of a warning being sent.

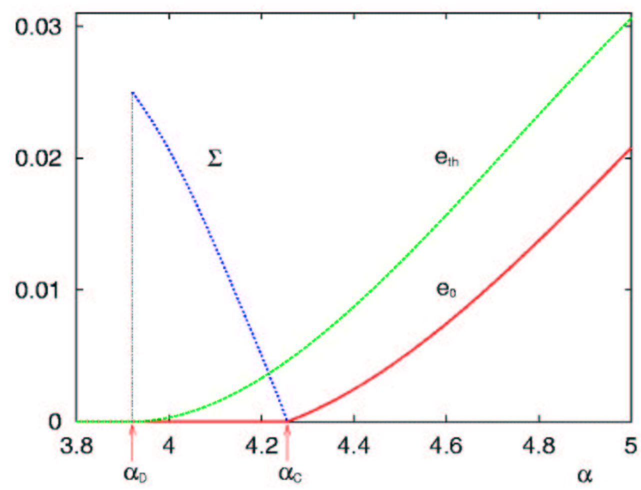
Survey propagation



Thermodynamics and complexity

Qualitative behaviour of the complexity as the function of the number of constraints per variable:



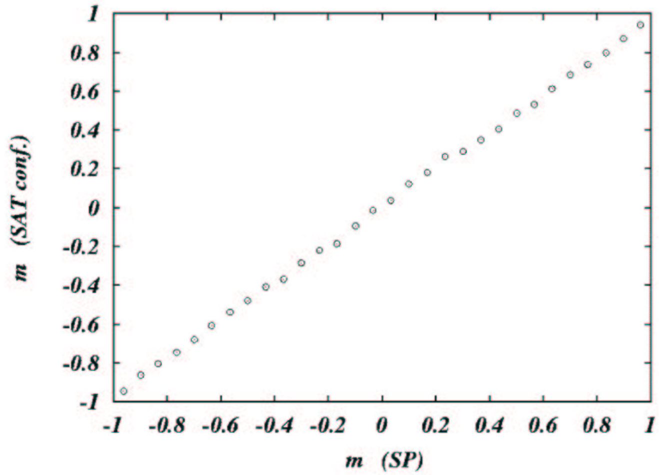


Single sample analysis

Order parameter= Survey of local magnetizations, in all states → Algorithm for the Hard-SAT phase.

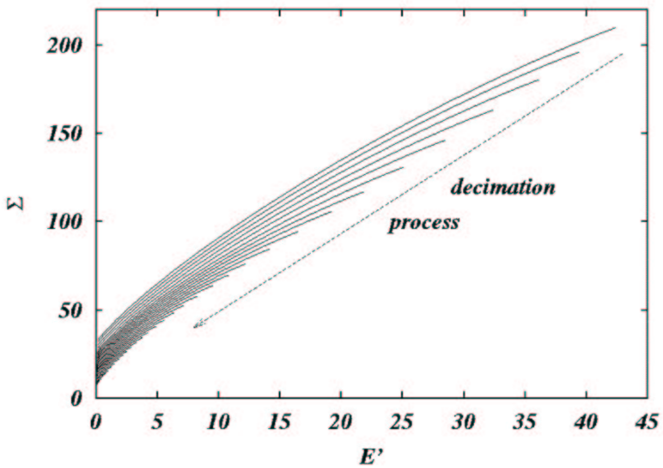
For each variable s_i: fraction of SAT-states where $s_i = 1$.

Check: compute many ground state with a standard algorithm, average the local spin in each state:



Survey Inspired Decimation

An algorithm to solve the Ksat problem: fix the spin which is most biased, rerun the survey propagation, iterate...



Solves the 'large' hard benchmarks of random 3sat at $\alpha = 4.2$ with $N = 1000, 2000$.

Solves typical random 3sat up to $N = 10^7$ at $\alpha = 4.2$. Complexity $O(N \log N)$.

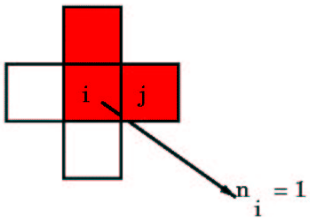
Best known algorithm: "walksat35" stops around $N = 10^4$

Local surveys of magnetic fields → a lot of information. Probably possible to invent many new algorithms.

Digression: unsatisfied glasses

Glass formation in systems with short range repulsion (e.g. hard spheres): geometric frustration. Local densest structure (icosahedral) incompatible with long range crystalline order.

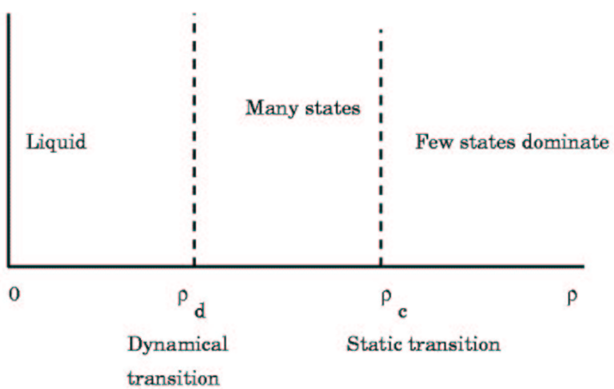
Lattice glass models (Biroli MM, Coniglio et al.,...): density constraint



Many local constraints.

Bethe approximation:

Density constraint: $\sum_j n_j < m$



Summary

- Analytic result on the generic samples of random 3sat: Phase diagram.
- Slowdown of algorithms near to $\alpha_c = 4.267$ due to the existence of a Hard SAT phase at $\alpha \in [3.921, 4.267]$, with exponentially many states.
- The whole construction can be checked versus rigorous computations on the "random XORSAT" (or "3 spin glass") problem. (MM, Ricci-Tersenghi, Zecchina; Cocco, Dubois, Mandler, Monasson).
- Single sample analysis: Survey propagation converges, yields non trivial information on the sample (diversity of sites) $\rightarrow$ Survey Inspired Decimation: a very efficient algorithm for solving random 3sat problems. Also applicable to many "constrained satisfaction problems", e.g. graph colouring (Mulet, Pagnani, Weigt, Zecchina).
- Similar mechanism is involved in the generation of long relaxation times in glasses (mena field).

References

- M. Mézard, G. Parisi and R. Zecchina, *"Analytic and Algorithmic Solution of Random Satisfiability Problems"*, Science 297 (2002) 812. Comment by C.P. Gomes and B. Selman: *"Satisfied with Physics"*, Science 297 (2002) 784.

- M. Mézard, R. Zecchina, *"The random K-satisfiability problem: from an analytic solution to an efficient algorithm"*, cond-mat/0207194, to appear in Phys. Rev. E

- M. Mézard and G. Parisi, *The cavity method at zero temperature*, cond-mat/0207121, to appear in J. Stat. Phys.