On message passing guided algorithms for solving constraint satisfaction problems

Guilhem Semerjian

LPT-ENS
Paris

09.10.09 / Orsay

in collaboration with

Andrea Montanari and Federico Ricci-Tersenghi

[AM, FRT, GS, Allerton 2007, arXiv:0709.1667]
[FRT, GS, J Stat. Mech. P09001 (2009), arXiv:0904.3395]

Outline

Introduction to CSPs

Constraint satisfaction problem :

- N variables $\underline{\sigma} = (\sigma_1, \dots, \sigma_N)$ (e.g. Ising or Potts spins)
- M constraints, each on k variables,

$$\psi_a(\sigma_{i_1^a}, \dots, \sigma_{i_k^a}) = \begin{cases} 1 & \text{if constraint } a \text{ satisfied} \\ 0 & \text{if constraint } a \text{ not satisfied} \end{cases}$$

solution : $\underline{\sigma}$ such that all constraints are simultanously satisfied (zero-energy groundstate)

Two examples with Ising spins variables $\sigma_i \in \{-1, +1\}$:

- XORSAT $\psi_a = 1 \quad \Leftrightarrow \quad \sigma_{i_1^a} \dots \sigma_{i_k^a} = J^a$
- SAT $\psi_a = 1 \quad \Leftrightarrow \quad (\sigma_{i_1^a}, \dots, \sigma_{i_k^a}) \neq (J_1^a, \dots, J_k^a)$

Phase transitions in random CSP

Random instances :

- for each constraint a take the indices $i_a^1 \dots, i_a^k$ uniformly at random
- take the J's ± 1 with equal probability

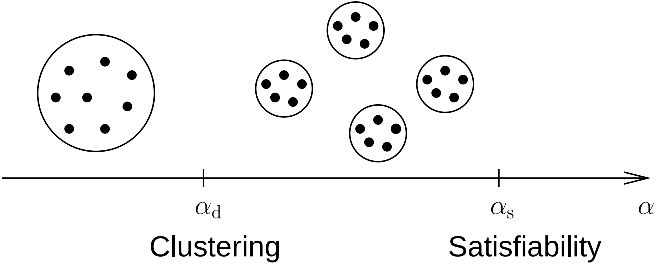
Large size (thermodynamic) limit, $N, M \to \infty$ with $\alpha = M/N$

Satisfiability phase transition :

- for $\alpha < \alpha_{\mathrm{s}}$ there are solutions (with high probability)
- for $\alpha > \alpha_{\mathrm{s}}$ there is no solution (with high probability)

XORSAT :

[Cocco, Dubois, Mandler, Monasson]
[Mézard, Ricci-Tersenghi, Zecchina]



For $\alpha < \alpha_{\rm s}$, there are $\exp[N\omega(\alpha)]$ solutions

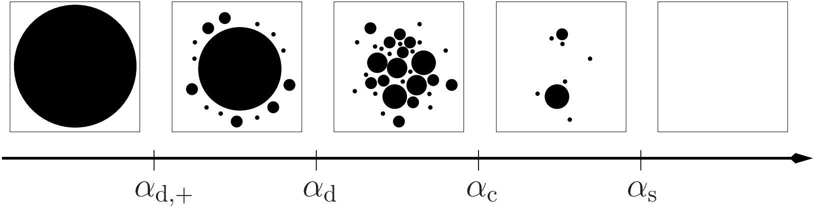
For $\alpha_{\rm d} < \alpha < \alpha_{\rm s}$, splitting of the entropy, $\omega(\alpha) = \omega_{\rm int}(\alpha) + \Sigma(\alpha)$

Clusters of “close” solutions

Phase transitions in random CSP

SAT [Biroli, Monasson, Weigt]
[Mézard, Parisi, Zecchina]

More transitions (for $k \geq 4$)
[Krzakala, Montanari, Ricci-Tersenghi, GS, Zdeborova]



Exponential number of relevant clusters only in $[\alpha_d, \alpha_c]$,
then “condensation”

Solving CSPs

Solving means :

- proving that there is no solution
- or exhibiting one solution

Here we concentrate on the second case
(assuming the existence of solutions)

Two broad class of algorithms :

- Local search (random walk in the space of configurations)
- Sequential construction

Uniform measure over the solutions : $\mu(\underline{\sigma}) = \frac{1}{\mathcal{N}} \prod_a \psi_a(\underline{\sigma}_a)$

Sequential generation from μ : for $t = 1, \dots, N$

- choose $i(t)$ u.a.r. in $\{1, \dots, N\} \setminus \{i(1), \dots, i(t-1)\}$
- draw $\sigma_{i(t)}$ according to $\mu(\sigma_{i(t)} | \sigma_{i(1)}, \dots, \sigma_{i(t-1)})$

At the end, $\underline{\sigma}$ is distributed according to μ

$\Rightarrow$ solves the uniform generation problem $\Rightarrow$ construction problem

BUT : needs an oracle to compute the marginals

Remarks :

- Never fails on a satisfiable instance
- Much more elaborate than the myopic algorithms (UC-like)
- Direct description of the dynamical process seems difficult

No hope to compute exactly $\mu(\sigma_{i(t)}|\sigma_{i(1)},\ldots,\sigma_{i(t-1)})$

$\Rightarrow$ approximative estimation with Belief Propagation

BP is a message passing algorithm :

$\nu_{i\to a}(\tau_i)$: "belief" on the probability of $\sigma_i = \tau_i$ in the absence of a

Local relations between messages, iterated towards a fixed point

- exact on trees
- extremization of the Bethe free energy [Yedidia, Freeman, Weiss]

If variable i has already been fixed, $\nu_{i\to a}(\tau_i) = \delta_{\tau_i,\sigma_i}$

... its approximate implementation...

Pseudocode : for $t = 1, \ldots, N$

- run BP until convergence
- choose $i(t)$ u.a.r. in $\{1, \ldots, N\} \setminus \{i(1), \ldots, i(t-1)\}$
- draw $\sigma_{i(t)}$ from the BP approximation of $\mu(\sigma_{i(t)} | \sigma_{i(1)}, \ldots, \sigma_{i(t-1)})$

On random instances, for which values of α will this avoid contradictions ?

Fails only because BP is approximate

New difficulty of the analysis : mistakes of BP hard to analyze/describe

Let us assume it is “as close as possible” from perfect...
thus try to infer its behavior from the study of the thought experiment

...and its analysis

State of the (perfect) algorithm at step t, $(\sigma_{i(1)}, \dots, \sigma_{i(t)})$, equivalent to:

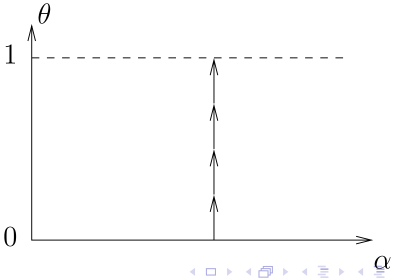
- draw $\underline{\sigma}$ from the uniform measure μ
- draw $V_t = \{i(1), \dots, i(t)\}$ u.a.r.
- keep the spins $\underline{\sigma}_{V_t} = (\sigma_{i(1)}, \dots, \sigma_{i(t)})$ and discard the others

$\Rightarrow$ Reduction of the dynamical process to a series of "static" problems

To understand the move $t \to t+1$, study $\mu(\cdot|\underline{\sigma}_{V_t})$

thermodynamic limit,
$t = \theta N$, $\quad V_t \to V_\theta$

Phase diagram in (α, θ) :



... and its analysis

Characterization of $\mu(\cdot|\underline{\sigma}_{V_\theta})$:

- residual entropy

$$\omega(\theta) = \frac{1}{N} \mathbb{E}_F \ \mathbb{E}_{\underline{\sigma}\sim\mu} \ \mathbb{E}_{V_\theta} \ \ln Z(\underline{\sigma}_{V_\theta})$$

  $Z(\underline{\sigma}_{V_\theta})$: number of solutions compatible with $\underline{\sigma}$ on V_θ

  Remark : similar to Franz-Parisi potential, but

  - potential = number of solutions at a given distance from $\underline{\sigma}$
  - technical difficulty due to the diluted nature of the model

- Possible splitting of the entropy in clusters
- Number of logically implied variables (by UCP)

... and its analysis

$$\omega(\theta) = \frac{1}{N} \mathbb{E}_F \; \mathbb{E}_{\underline{\sigma} \sim \mu} \; \mathbb{E}_{V_\theta} \; \ln Z(\underline{\sigma}_{V_\theta})$$

Technical difficulty : $\mathbb{E}_{\underline{\sigma} \sim \mu}$ and $\mathbb{E}_F$ are not independent

Needs an extension of the cavity method :

- Local convergence of a formula towards a random tree
- Assumptions on the local behaviour of the probability measure μ
- Reduction of the problem to a tree model

On a tree (almost) everything can be computed
with (more or less complicated) recursions

XORSAT

Toy model example because

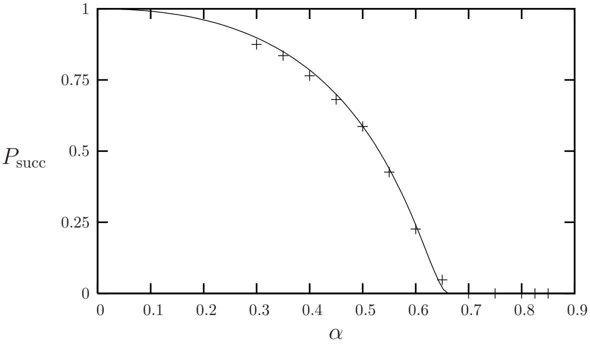
BP guided decimation on XORSAT $\Leftrightarrow$ Unit Clause Propagation

Analysis can be made rigorously by differential equations method

Useful check of the heuristic cavity computation

XORSAT

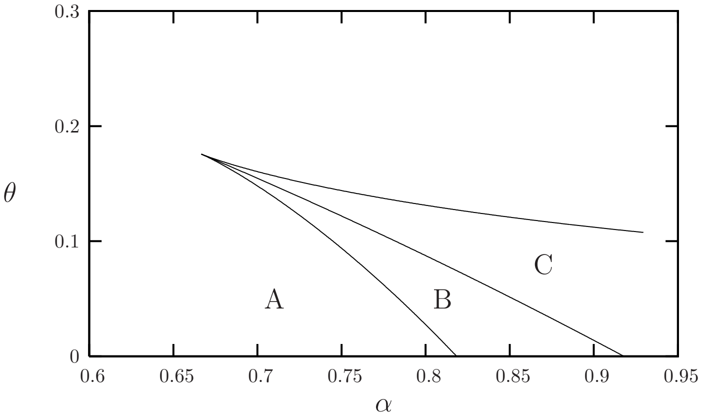
Probability of success for “BP guided decimation” :



symbols : 800 tries, N=20000

line : analytical prediction in the thermodynamic limit [Frieze, Suen] [Deroulers, Monasson]

XORSAT : interpretation from the cavity computation

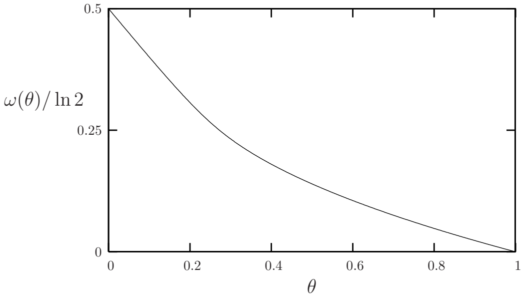


A : no clusters B : clusters, $\Sigma > 0$ C : clusters, $\Sigma = 0$

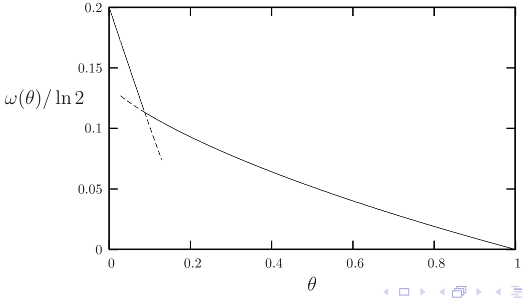
Algorithm fails w.h.p. when it has to cross the phase transition lines

XORSAT : interpretation from the cavity computation

$\alpha < \alpha_*$:



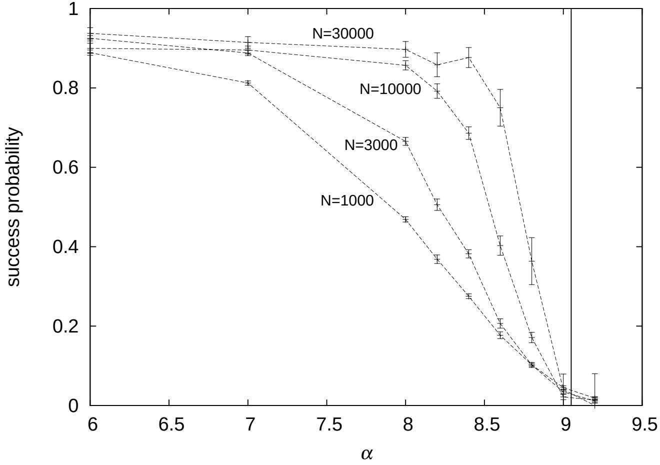
$\alpha > \alpha_*$:



Not equivalent to a myopic algorithm

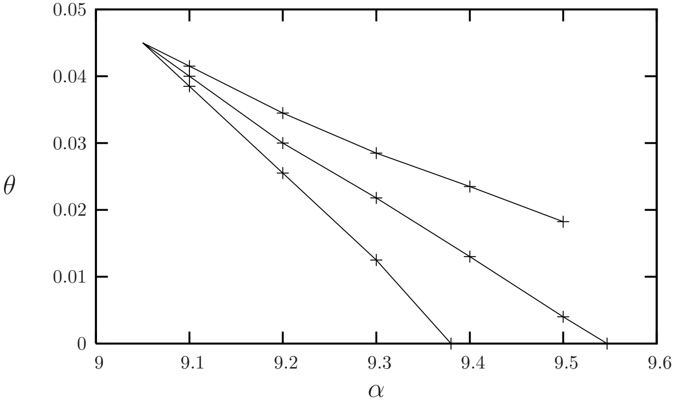
Cavity computation leads to functional equations
which have to be solved numerically

Probability of success for BP guided decimation (4-SAT) :



SAT

Phase diagram from the cavity computation :



same meaning of the lines as in XORSAT

rather good coincidence with the failure of the algorithm

Conclusions

- $k = 3$ qualitatively distinct from $k \geq 4$ for SAT
- at large k, $\alpha_* \sim 2^k/k$, far from the satisfiability threshold $2^k \ln 2$
- other models (e.g. coloring)
- "sorting matters"
- Reinforcement
- Survey Propagation