

Constraint satisfaction networks: Belief propagation and beyond

Marc Mézard

14th December 2003

Collaborators:

- **Riccardo Zecchina** (ICTP Trieste)
- **A. Braunstein** (Trieste), **S. Mertens** (Magdeburg), **M. Weigt** (Göttingen).

Message passing in constraint satisfaction problems (CSP)

- Factor graph representation of CSPs and Belief Propagation (BP)
- **Appearance of clusters**: histograms of the usual messages
- **Focusing on constrained variables**: Warning Propagation (WP)
- Clusters with constrained variables: histograms of warnings = Survey Propagation (SP)
- Satisfiability: the subtle difference between BP and SP...

Constraint satisfaction problems

- **Discrete variables** x_1, \dots, x_N . Each variable $x_i \in \{0, 1, \dots, q - 1\}$.
- **Constraints** $a \in \{1, \dots, M\}$. Each constraint a involves a subset of K variables: $X_a = \{x_{i_1(a)}, \dots, x_{i_K(a)}\}$. Defined by a function $f_a(X_a) \in \{0, 1\}$. Interpretation: $f_a(X_a) = 1$ iff the configuration X_a satisfies constraint a

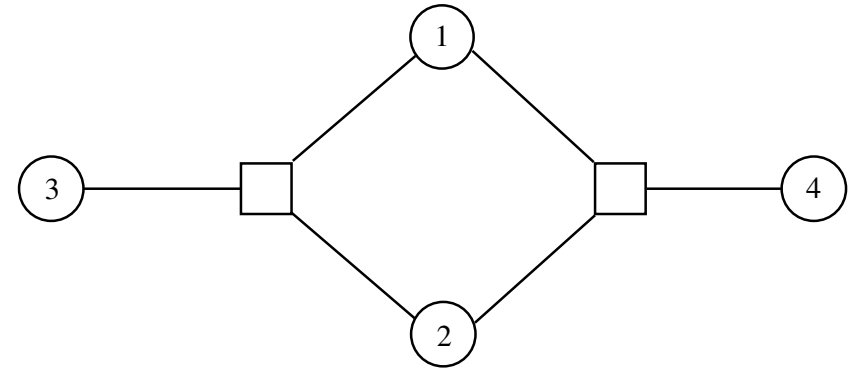
K-Satisfiability: Boolean variables, $q = 2$. Constraint = clause: forbids one configuration of $X_a = \{x_{i_1(a)}, \dots, x_{i_K(a)}\}$ among the 2^K possible ones. (Random K-sat: the forbidden configuration is chosen at random with uniform distribution).

Coloring: q = number of colors. Constraints: for each edge a of the graph to be colored, imposes that the colors of the two vertices at the ends of this edge are different.

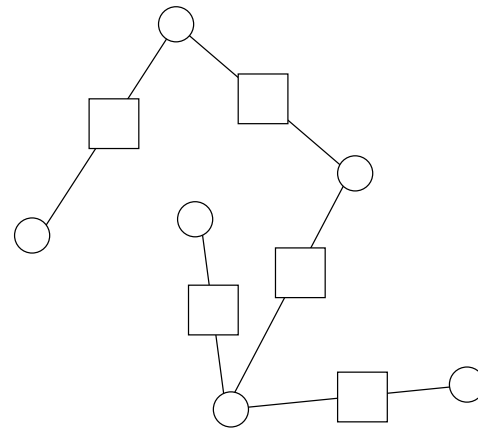
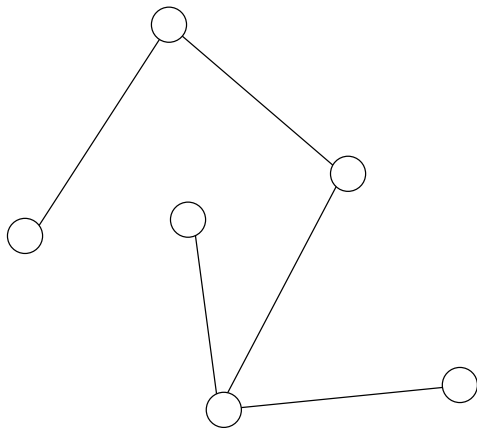
Factor graphs

Satisfiability:

$$(x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_2 \vee \bar{x}_3 \vee \bar{x}_4).$$



Colouring:



Belief propagation (BP)

For a randomly chosen SAT configuration:

- $\mu_{a \rightarrow i}(x_i)$ = probability that constraint a is satisfied, given the value of the variable x_i .
- $\mu_{i \rightarrow a}(x_i)$ = probability that the variable takes value x_i , when constraint a is absent. ($\sum_{x_i \in \{0,1\}} \mu_{i \rightarrow a}(x_i) = 1$)

BP equations = message passing:

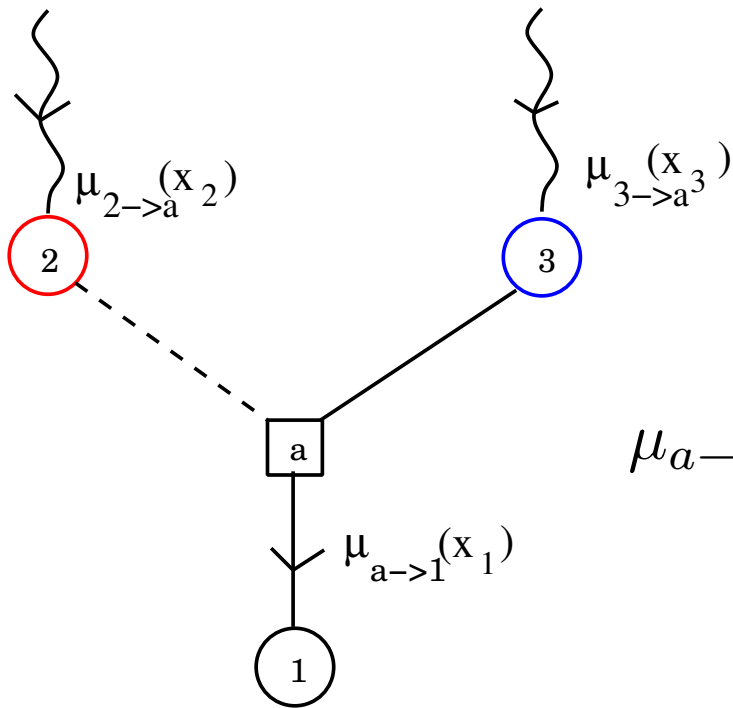
$$\mu_{i \rightarrow a}(x_i) = C_{i \rightarrow a} \prod_{b \in V(i) \setminus a} \mu_{b \rightarrow i}(x_i) ,$$

$$\mu_{a \rightarrow i}(x_i) = \sum_{\{x_j(j \neq i)\}} f_a(X_a) \prod_{j \in V(a) \setminus i} \mu_{j \rightarrow a}(x_j) ,$$

Basic approximation: uncorrelated inputs

Hypothesis: joint 'cavity' probability of x_2, x_3 , in absence of a , factorizes:

$$P^{(a)}(x_2, x_3) \simeq \mu_{2 \rightarrow a}(x_2) \mu_{3 \rightarrow a}(x_3)$$



BP equations:

$$\mu_{a \rightarrow 1}(x_1) = \sum_{x_2, x_3} f_a(x_1, x_2, x_3) \mu_{2 \rightarrow a}(x_2) \mu_{3 \rightarrow a}(x_3)$$

$$\mu_{2 \rightarrow a}(x_2) \propto \prod_{b \in V(2) \setminus a} \mu_{b \rightarrow 2}(x_2)$$

Validity of the BP factorization

$$P^{(a)}(x_2, x_3) \simeq \mu_{2 \rightarrow a}(x_2) \mu_{3 \rightarrow a}(x_3)$$

- In the absence of constraint a , the variables x_2, x_3 should be generically distant. True for random graphs with a finite coordination number (Erdos Renyi, random regular....).
- There should be a single pure state, or the set of SAT configurations used to define the $\mu_{i \rightarrow a}, \mu_{a \rightarrow i}$ should be within the same pure state.

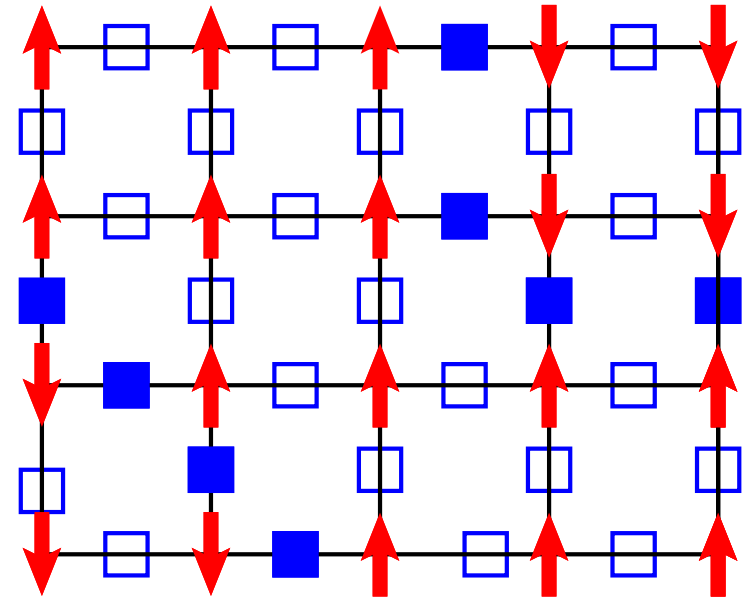
Pure states and clustering

Interacting variables on graphs embedded in space: “locality of correlations” is a concept from statistical physics. e.g. Ising model:

2 SAT configurations (“states”):

“+”: $\forall i : s_i = +1$

“-”: $\forall i : s_i = -1$



Inside one state $P_{\pm}(s_i, s_j) = \delta_{s_i, \pm 1} \delta_{s_j, \pm 1} = P_{\pm}(s_i) P_{\pm}(s_j)$

Global measure $P(s_i, s_j) = \frac{1}{2} (P_+(s_i, s_j) + P_-(s_i, s_j)) \neq P(s_i) P(s_j)$

Pure states and clustering (II)

Ising model at finite inverse temperature $\beta > \beta_c$: Boltzmann measure $P(s_1, \dots, s_N) = \exp(-\beta \sum_a E_a)$. Less trivial, but still two states:

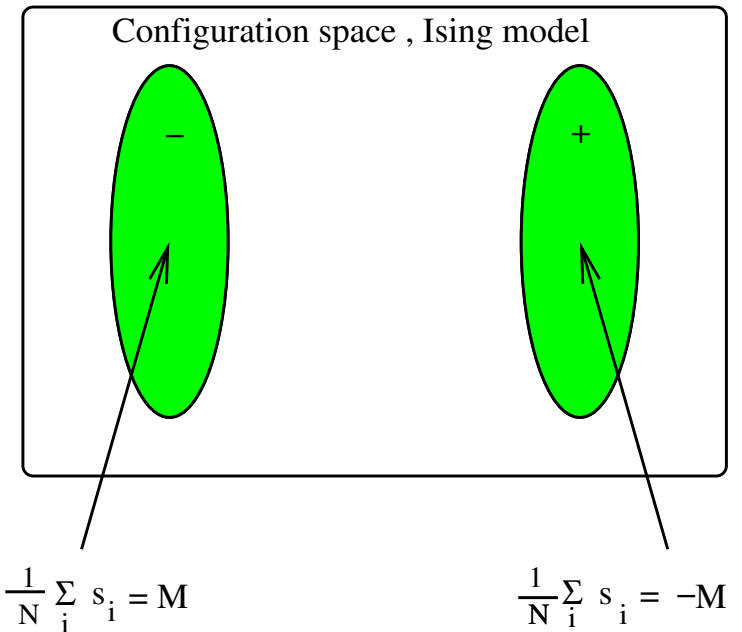
$$\text{"+"} : \frac{1}{N} \sum_i s_i = +M$$

$$\text{"-"} : \frac{1}{N} \sum_i s_i = -M$$

Correlations at large distances $d(i, j) \gg 1$:

Inside one state $P_{\pm}(s_i, s_j) - P_{\pm}(s_i)P_{\pm}(s_j) \simeq \exp[-d(i, j)/\xi]$

Global measure $P(s_i, s_j) - P(s_i)P(s_j) \neq 0$



SAT clusters: surveys of beliefs

Lesson: pure states = fragmentation of phase space = clusters of SAT configurations

If there are several pure states $\alpha = 1, 2, \dots$ in a constraint satisfaction problem: BP messages OK within one given pure state α (factorization property): $\vec{\mu}^\alpha = \{\mu_{a \rightarrow i}^\alpha, \mu_{i \rightarrow a}^\alpha\}$

Convergence to a single $\vec{\mu}^\alpha$: difficult in large graphs.

Survey = two functionals on each edge $a - i$ of the graph:

$P_{a \rightarrow i}[\mu(\cdot)] = \text{Probability that } (\mu_{a \rightarrow i}^\alpha(\cdot) = \mu(\cdot))$

$P_{i \rightarrow a}[\mu(\cdot)] = \text{Probability that } (\mu_{i \rightarrow a}^\alpha(\cdot) = \mu(\cdot))$

when α is chosen at random with uniform probability.

From beliefs to warnings

Functional message passing: heavy!

Simplification: project the beliefs to simpler objects.

Messages $\mu_{a \rightarrow i}(x_i), \mu_{i \rightarrow a}(x_i)$ are probabilities, in $[0, 1]$.

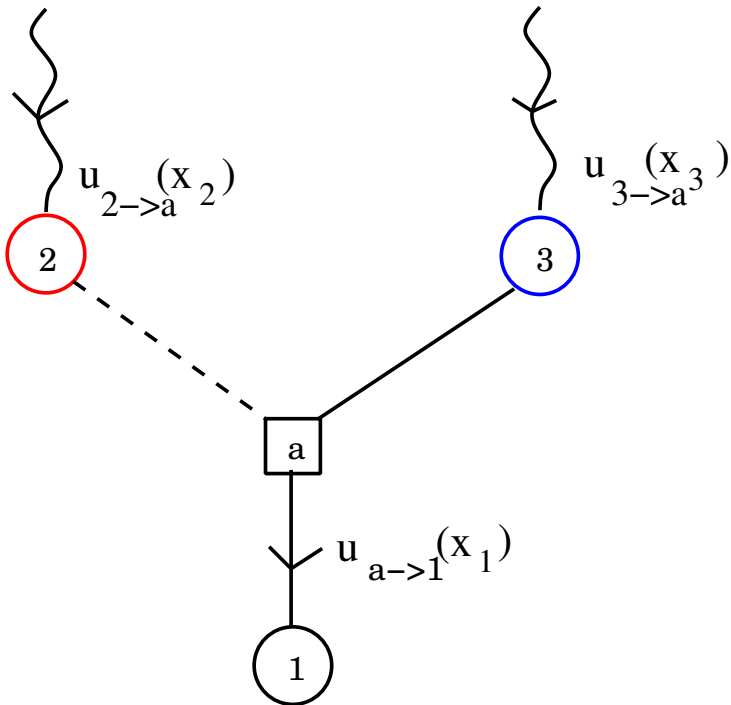
For each $\mu_{a \rightarrow i}(x_i)$, construct the **warning** $u_{a \rightarrow i}(x_i) \in \{0, 1\}$.

$$u = I(\mu) = \begin{cases} 1 & \text{if } \mu_{a \rightarrow i}(x_i) = 0 \\ 0 & \text{if } \mu_{a \rightarrow i}(x_i) > 0 \end{cases}$$

Warning $u_{a \rightarrow i}(x_i) = 1$: “you should not be in state x_i ”.

Warning propagation = projected belief propagation (autonomous).

Warning propagation



$u_{a \rightarrow 1}(x_1) = 0$ **iff:** there exist x_2, x_3 with $f_a(x_1, x_2, x_3) = 1,$

$$u_{2 \rightarrow a}(x_2) = 0, u_{3 \rightarrow a}(x_3) = 0$$

$u_{2 \rightarrow a}(x_2) = 0$ **iff:** $\forall b \in V(2) \setminus a : u_{b \rightarrow 2}(x_2) = 0$

Survey propagation

Survey of the warnings:

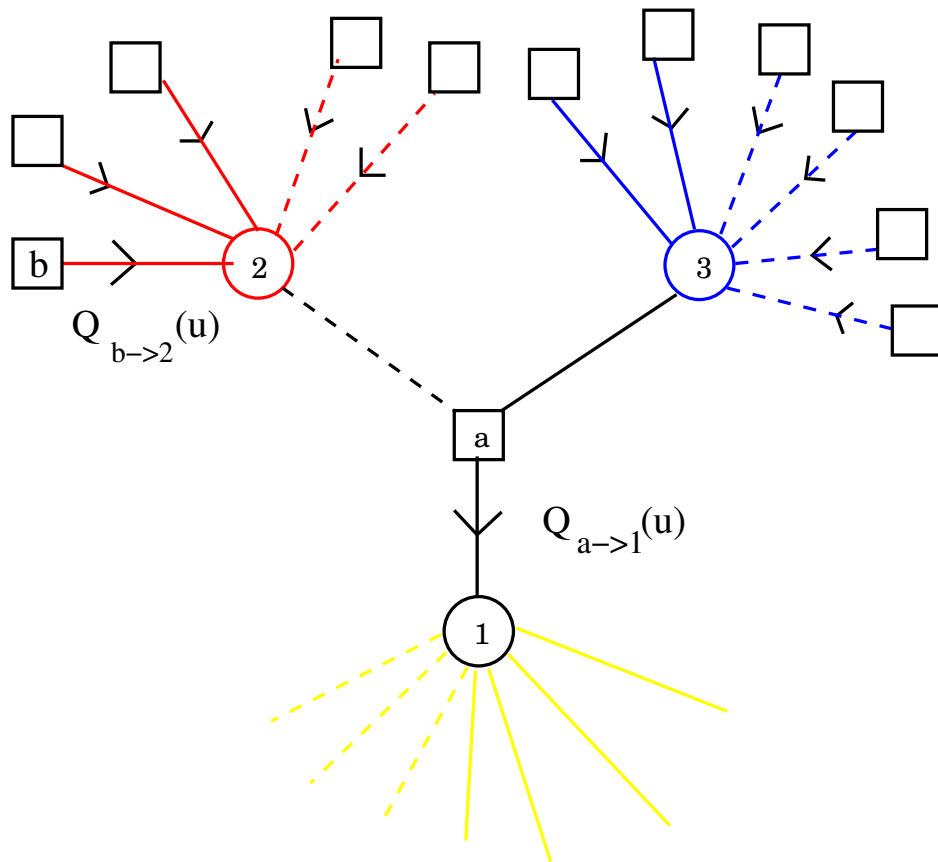
Warning $\vec{u}_{a \rightarrow i} = \{u_{a \rightarrow i}(x), x \in \{0, \dots, q-1\}\}$. Takes 2^q values.

Survey = Probability distribution $Q_{a \rightarrow i}(\vec{u})$, on each edge $a - i$ of the graph (and similarly for $Q_{i \rightarrow a}$).

$Q_{a \rightarrow i}(\vec{u})$ = Probability that $(\forall x_i : u_{a \rightarrow i}^\alpha(x_i) = u(x_i))$

when α is chosen at random among the SAT states with uniform probability.

Survey propagation



$Q_{a \rightarrow 1}(\vec{u})$: known exactly from
joint probability of
incoming warnings.

SP approximation: this joint
probability factorizes.

The subtle difference between BP and SP, II

q_2 : probability that $x_2 = 1$ when a is absent. 3 cases:

No warning at all:

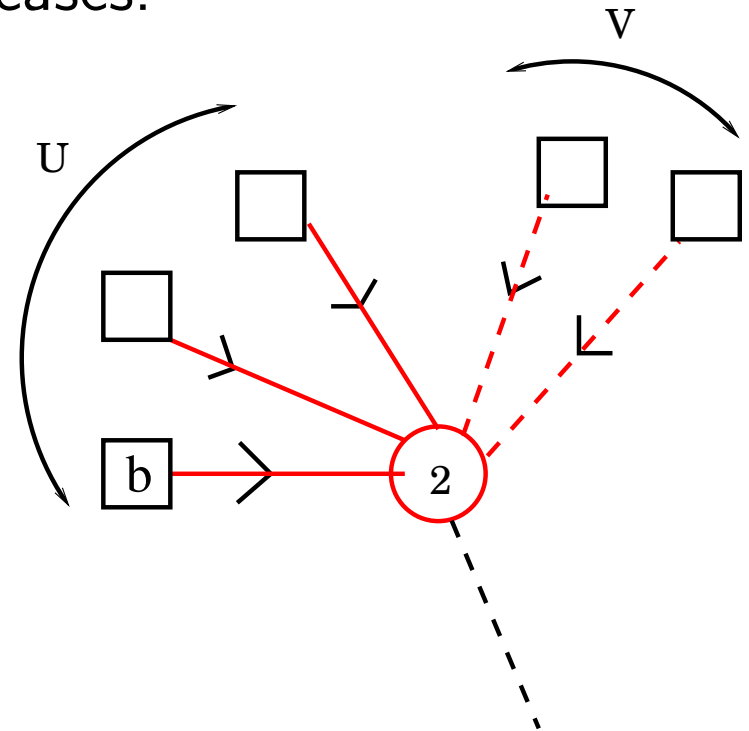
$$\rho_0 = \prod_{b \in U \cup V} (1 - \eta_{b \rightarrow 2})$$

Warning from U , not from V :

$$\rho_- = \prod_{b \in V} (1 - \eta_{b \rightarrow 2}) - \rho_0$$

Warning from V , not from U :

$$\rho_+ = \prod_{b \in U} (1 - \eta_{b \rightarrow 2}) - \rho_0$$



$$\text{SP: } q_2 = \frac{\rho_-}{\rho_+ + \rho_- + \rho_0} \quad ; \quad \text{BP: } q_2 = \frac{\rho_- + \rho_0}{\rho_+ + \rho_- + 2\rho_0}$$

From BP to SP: the joker state

Equations: introduce an extra * state of the variable (“unconstrained”), and use BP in this generalized space \rightarrow SP. Why???

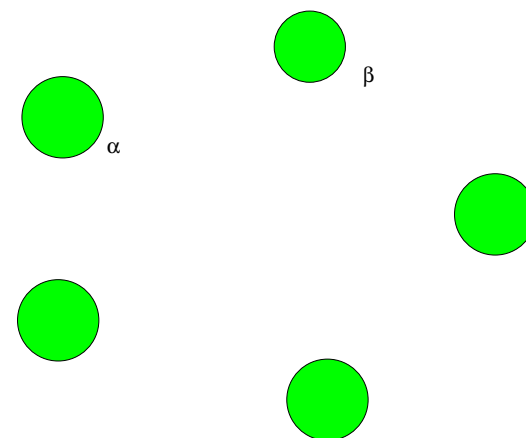
Local belief of variable x_i in cluster α :

If warning for state 0: $\mu_i^{(\alpha)}(x_i) = \delta_{x_i,1}$

If warning for state 1: $\mu_i^{(\alpha)}(x_i) = \delta_{x_i,0}$

Else: $\mu_i^{(\alpha)}(x_i) = a\delta_{x_i,0} + (1 - a)\delta_{x_i,1}$ with $a \in]0, 1[$.

Generalized assignment: $\hat{x}_i \in \{1, 0, *\}$ describes the three cases above. Shrinks each cluster of solutions to a point. Entropy for this generalized problem: counts the clusters (complexity).



Convergence of SP

Experimentally: SP **converges**, on one given instance of random 3-SAT, up to $\alpha = M/N = 4.36$.

→ A set of surveys, $\eta_{a \rightarrow i}$ (one real number for each edge of the hypergraph), such that the SP equations are satisfied for each edge.

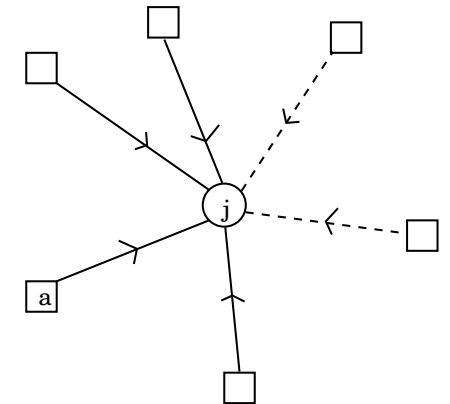
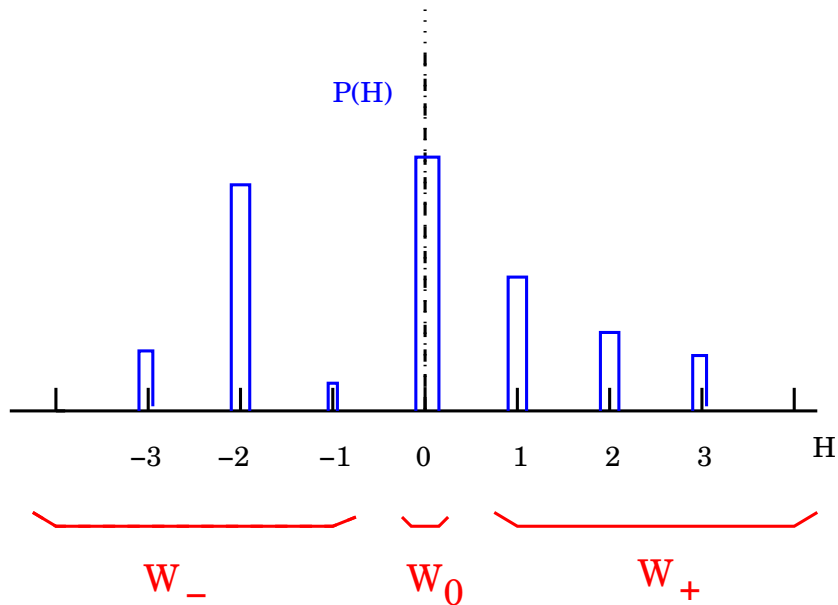
Interpretation:

$\eta_{a \rightarrow i}$: Probability of a warning being sent from a to i , when one picks up at random a cluster of solutions.

Using the surveys: local field

“Local field” on variable j : $H_j = \sum_{a \in V_+(j)} u_{a \rightarrow j} - \sum_{a \in V_-(j)} u_{a \rightarrow j}$

→ Survey of local field, without contrad. $P_j(H)$



Some types of variables

Balanced: $W_+ \simeq W_- \simeq 1/2, W_0 \simeq 0$

Polarized: $W_+ \simeq 1$ or $W_- \simeq -1$

Underconstrained: $W_0 \simeq 1$

Survey Inspired Decimation

Biased variable $W_+^i \simeq 1$: In almost all clusters of solutions, $x_i = 1$.

→ Fix $x_i = 1$

SID algorithm: Iterate:

- Run SP until convergence
- Find most biased variable, i such that $|W_+^i - W_-^i|$ maximal.
- Fix it to $x_i = 1$ if $W_+^i > W_-^i$, to $x_i = 0$ if $W_+^i < W_-^i$, simplify the formula.

Two possible ends: 1) Fix all variables 2) reduce the formula to a stage where all $W_0^i = 1$. Underconstrained or 'paramagnetic' problem, easily solved by e.g. simulated annealing or Walksat.

Solves: 10^7 variables at $\alpha \simeq 4.2 - 4.25$. Time $O(N^2)$, reduced to $O(N)$ by fixing a fraction of the variables.

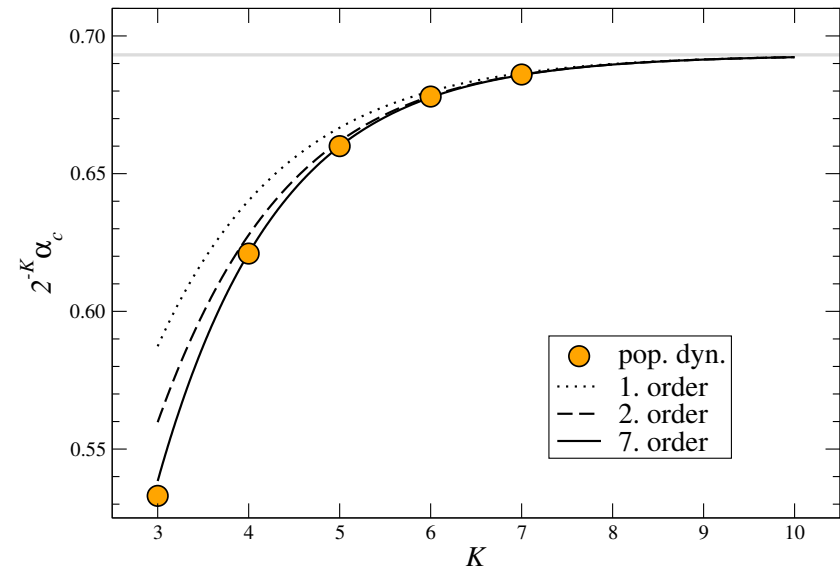
Phase diagram

Thresholds:

α_c : SAT-UNSAT threshold.

α_d : Onset of a non trivial $P(\phi)$

→ clusters with frozen variables.



Large K behaviour:

$$2^{-K} \alpha_c(K) = \ln(2) - \left[\frac{\ln(2)+1}{2} \right] 2^{-K} +$$

$$+ \left[\frac{1}{8} - \frac{\ln(2)}{12} + \frac{3 \ln(2)-2}{8} K - \frac{\ln(2)+2 \ln^2(2)}{8} K^2 \right] 2^{-2K} + \dots$$

Threshold values

K	α_d	α_s	α_c	$\alpha_c^{(7)}$
3	3.93	4.15	4.267	4.307
4	8.30	9.08	9.931	9.938
5	16.1	17.8	21.117	21.118
6	30.5	33.6	43.37	43.372
7	57.2	63	87.79	87.785
8	107.2			176.543
9	201.3			354.010
10	379.1			708.915

$\alpha_s(K)$: Stability threshold. Present solution (1RSB cavity solution) stable to further clustering (higher order RSB) only for $\alpha > \alpha_s$.

SAT-UNSAT threshold α_c always in stable region. α_d in the unstable region.

Discussion

Two main ideas:

- Clustering of SAT configurations: bad for BP. Use histograms of the usual messages.
- In constraint satisfaction problems: simple signature of a cluster, saying which variables are fixed and which ones are constrained. → warning propagation, and histograms of warnings.

Discussion 2

Mathematical results: $\alpha_c(K)$ obtained from SP statistical analysis is an upper bound to exact one for K even (Franz Leone, using Guerra's line of approach). Whole structure (including clustering effect) checked for XORSAT (MM, Ricci-Tersenghi, Zecchina; Cocco, Dubois, Mandler, Monasson).

Mathematical challenges: Prove (or infirm) the validity of the whole structure (including clustering effect), and the values of thresholds $\alpha_c(K)$. Study convergence properties of the various message passing procedures...

Algorithmic challenges: “practical” graphs with hubs and small loops: pre processing, generalized survey propagation,...

References

- A. Braunstein, M. Mézard and R. Zecchina, “*Survey propagation: an algorithm for satisfiability*”, <http://arXiv.org/abs/cs.CC/0212002>
- A. Braunstein, M. Mezard, M. Weigt, R. Zecchina, “*Constraint Satisfaction by Survey Propagation*”, <http://arXiv.org/abs/cond-mat/0212451>
- S. Mertens, M. Mézard and R. Zecchina, “*Threshold values of Random K -SAT from the cavity method*” <http://arXiv.org/abs/cs.CC/0309020>