Phase Transitions in Random Constraint Satisfaction Problems

Florent Krzakala
(ESPCI, Paris)

In collaboration with:
L. Zdeborová (LPTMS Paris)
J. Kurchan (ESPCI Paris)
G. Semerjian (ENS Paris)
A. Montanari (Stanford)
F. Ricci-Tersenghi (Roma)

Find these slides on my webpage:
www.pct.espci.fr/~florent/
Constraint satisfaction problems (CSP)

- Boolean Satisfiability, Traveling Salesman Problem,
- Vertex coloring (generalization of map coloring)

Many applications: Scheduling, Register allocation ...
Constraint satisfaction problems (CSP)

- Boolean Satisfiability, Traveling Salesman Problem,
- Vertex coloring (generalization of map coloring)

Many applications: Scheduling, Register allocation ...

Only four colors are needed to color a map (planar graph)
Constraint satisfaction problems (CSP)

- Boolean Satisfiability, Traveling Salesman Problem,
- Vertex coloring (generalization of map coloring)

Many applications: Scheduling, Register allocation ...

Only four colors are needed to color a map (planar graph)
Coloring random graphs

N=100 nodes, M=218 links, average connectivity c = 2M/N = 4.36

In general, finding a q-coloring is a NP-Hard problem!
Consider the following Potts Hamiltonian:

\[ H = \sum_{\langle ij \rangle} \delta(s_i, s_j) \]

where \( s_i = 1, 2, \ldots, q \).
Consider the following Potts Hamiltonian:

\[ H = \sum_{\langle ij \rangle} \delta(s_i, s_j) \]

where \( s_i = 1, 2, \ldots, q \).

A random graph is locally tree-like (“mean field” model) with large loops (frustration) of size \( \log(N) \).

\[ \text{Mézard, Parisi, Zecchina, Science (2002)} \]
Random graphs: sharp threshold in colorability/satisfiability

- Existence proven rigorously (Friedgut, 1997)
- Location
  - empirically - exhaustive search
  - rigorously - bounds, (Achlioptas, Naor, Peres, 2005)
  - physics - “exact” calculation via the cavity method

Mézard, Parisi, Zecchina 2002, Mézard, Zecchina, 2002
Mulet, Pagnani, Weigt, Zecchina, 2002
FK, Pagnani, Weigt 2004
The location of the COL/UNCOL transition

3-coloring Erdős-REnyi random graphs

- Numerical estimates: $c \sim 4.7$
- Rigorous bounds: $4.03 < c < 4.99$
- Cavity prediction
The location of the COL/UNCOL transition

3-coloring Erdös-renyi random graphs

- Numerical estimates: $c \approx 4.7$
- Rigorous bounds: $4.03 < c < 4.99$
- Cavity prediction: $4.687(2)$
The location of the COL/UNCOL transition

3-coloring Erdös-renyi random graphs

- Numerical estimates: $c \sim 4.7$
- Rigorous bounds: $4.03 < c < 4.99$
- Cavity prediction: $4.687(2)$

$q$-coloring Erdös-renyi random graphs

- Rigorous bounds: $2q \log q - \log q - 1 \geq c_q \geq 2q \log q - \log q + o(1)$
- Cavity prediction:
The location of the COL/UNCOL transition

3-coloring Erdös-renyi random graphs

• Numerical estimates : \( c \sim 4.7 \)
• Rigorous bounds : \( 4.03 < c < 4.99 \)
• Cavity prediction : \( 4.687(2) \)

\[
2q \log q - \log q - 1 \geq c_q \geq 2q \log q - \log q + o(1)
\]

q-coloring Erdös-renyi random graphs

• Rigorous bounds: \( 2q \log q - \log q - 1 \geq c_q \geq 2q \log q - \log q + o(1) \)
• Cavity prediction: \( c_q \sim 2q \log q - \log q + 1 \)
The location of the COL/UNCOL transition

3-coloring Erdös-renyi random graphs

- Numerical estimates: \( c \sim 4.7 \)
- Rigorous bounds: \( 4.03 < c < 4.99 \)
- Cavity prediction: \( 4.687(2) \)

\[ 2q \log q - \log q - 1 \geq c_q \geq 2q \log q - \log q + o(1) \]

The cavity predictions agree perfectly with rigorous work!
Computational complexity

- **Worst case**: NP-complete (Cook, 1971)
- **Average case**
  - empirical evidence - backtracking (Cheeseman, Kanefsky, Taylor, 1991)
  - physics explanation - clustering of the space of solutions?

Phase transitions:
a possible mechanism for the hardness?
Outline

The phase space of the random coloring problem

• Description of the cavity method

• The phase diagram and the phase transitions of the q coloring problem

• A brief discussion of algorithmic implications
The Cavity Method

Parisi, Mézard, Virasoro ‘87
Parisi, Mézard ‘00
Parisi, Mézard, Zechinna ‘02
The iterative solution on a tree

Coloring = anti-ferromagnetic Potts model at zero temperature

Recursive equations on a tree (Belief propagation):

\[
\psi_{s_i}^{i \rightarrow j} = \frac{1}{i \rightarrow j} \prod_{k \in V(i) - j} \sum_{s_k} (1 - \delta_{s_i s_k}) \psi_{s_k}^{k \rightarrow i}
\]

\[
= \frac{1}{i \rightarrow j} \prod_{k \in V(i) - j} (1 - \psi_{s_i}^{k \rightarrow i})
\]

\(\psi_{s_i}^{i \rightarrow j}\) is the set of probabilities that the spin i takes the color q in absence of the spin j
The replica symmetric solution on a graph

A random graph being locally tree-like, assume a “fast” decay of correlations, then the RS solution should be correct.
The replica symmetric solution on a graph

A random graph being locally tree-like, assume a “fast” decay of correlations, then the RS solution should be correct.

Rigorously proven for regular random graphs for $c < q-1$...

... and believed to be correct even beyond (until $c \sim q \log q$).

Bandyopadhyay, Gamarnik '05
The replica symmetric solution on a graph

A random graph being locally tree-like, assume a “fast” decay of correlations, then the RS solution should be correct.

Rigorously proven for regular random graphs for $c < q-1$.

... and believed to be correct even beyond (until $c \sim q \log q$).

Solution of the self-consistent equation:
Only the “paramagnetic” $\psi = (1/q, 1/q, \ldots)$ in the COL phase.

This leads to the following entropy:

$$s_{RS} = \log q + \frac{c}{2} \log \left( 1 - \frac{1}{q} \right)$$
When is the tree solution expected to be correct?

★ If we are in a “paramagnetic phase”.

★ From spin glass theory, we expect however a transition to a spin glass phase.

1) Continuous transition (divergence of the Spin glass susceptibility) like in the Sherrington-Kirkpatrick model

2) Discontinuous transition (like in the p-spin or the Random Energy model model)
The phase space splits into an exponential number $N$ of components. Define the complexity (or configurational entropy) $\Sigma$ as $\mathcal{N} = e^{N \Sigma}$.

The complexity can be computed using a “modified” partition sum:

$$\sum_{\alpha} z_{\alpha}^m = \sum_{\alpha} \left( \sum_{s \in \alpha} e^{-\beta E(s)} \right)^m = \int df \, e^{-N(\beta m f(\beta) - \Sigma(f))} = e^{-\beta m N \Phi(\beta, m)}$$
Replica symmetry breaking

The phase space splits into an exponential number $\mathcal{N}$ of components. Define the complexity (or configurational entropy) $\sum$ as $\mathcal{N} = e^{N \sum}$.

The complexity can be computed using a "modified" partition sum:

$$\sum z^m_{\alpha} = \sum_{\alpha} \left( \sum_{s \in \alpha} e^{-\beta E(s)} \right)^m = \int df \, e^{-N(\beta m f(\beta) - \Sigma(f))} = e^{-\beta m N \Phi(\beta, m)}$$

The "Replicated" free energy is the Legendre transform of the complexity

$$-\beta m \Phi(m, \beta) = -\beta m F(\beta) + \Sigma(F)$$
The replica symmetry breaking recursion

Order Parameter:
Probability distribution $P^{i \rightarrow j}(\psi)$ of fields for every edge.

Self-consistent equation:

$$P^{i \rightarrow j}(\psi) = \frac{1}{Z^{i \rightarrow j}} \delta[\psi^{i \rightarrow j} - F(\{\psi^{k \rightarrow i}_{s_i}\})] e^{m \Delta S^{i \rightarrow j}} \prod_{k \in V(i) - j} dP^{k \rightarrow i}(\psi)$$

See:
Mézard, Palassini, Rivoire '05
The replica symmetry breaking recursion

Order Parameter:
Probability distribution $P^{i \rightarrow j}(\psi)$ of fields for every edge.

Self-consistent equation:
\[
P^{i \rightarrow j}(\psi) = \frac{1}{Z^{i \rightarrow j}} \delta[\psi^{i \rightarrow j} - \mathcal{F}(\{\psi^{k \rightarrow i}\})] e^{m\Delta S^{i \rightarrow j}} \prod_{k \in V(i) - j} dP^{k \rightarrow i}(\psi)
\]

Numerical Solution - Population dynamics: very heavy!!!

See: Mézard, Palassini, Rivoire ‘05

Simplifications: $m=0$, $m=1$, regular graphs, hard fields...

Large $q$ expansion
The results and the phase diagram
Learning from $\Sigma(s)$

Example of 6-coloring, connectivities 17, 18, 19, 20 (from top).
6 coloring of regular random graph

very low connectivity
6 coloring of regular random graph  
connectivity $c=17$
6 coloring of regular random graph

connectivity $c=18$
6 coloring of regular random graph

connectivity $c = 19$
6 coloring of regular random graph

connectivity $c=20$
The phase transitions

★ **Clustering (dynamical) transition**
  - Monte Carlo equilibration time diverges.
  - Point-to-set correlation functions diverge.
  - The replica symmetric entropy still correct, no non-analyticity.
  - Entropy dominated by exponentially many states.

★ **Condensation (Kauzmann) transition**
  - Discontinuity in the second derivative of entropy.
  - Entropy dominated by few of the largest states.

★ **Rigidity transition**
  - Frozen variables appears in the dominating states.
  - Minimal rearrangements diverge.
Large number of colors

0

$0$ −

$q \log q$

\{ SP nontrivial, Clustering, Rigidity \}

$\{ \begin{align*}
& c_{SP} \sim q (\log q + \log \log q + 1 - \log 2) \\
& c_r \sim q (\log q + \log \log q + 1) \\
& c_g \sim 2q \log q - \log q - 2 \log 2 \\
& c_q \sim 2q \log q \quad \log q \quad 1
\end{align*} \}

$2q \log q$
Large number of colors

\[ q \log q \]

- SP nontrivial: \[ c_{SP} \approx q (\log q + \log \log q + 1 - \log 2) \]
- Clustering: \[ c_r \approx q \left( \log q + \log \log q + 1 \right) \]
- Rigidity: \[ c_{\text{rig}} \approx 2q \log q - \log q - 2 \log 2 \]
- Col/Uncol: \[ c_q = \frac{2q \log q - \log q - 1}{\log q} \]
A final word on dynamics
“Wet toes” algorithm

Arkless strategy for flood victims

You are on a rugged landscape that is being flooded
“Wet toes” algorithm

Arkless strategy for flood victims

Water goes up. When your toes are wet step back on the land!
“Wet toes” algorithm

Arkless strategy for flood victims

And wait until your toes get wet again...
“Wet toes” algorithm

Arkless strategy for flood victims

Sooner or later you’ll find yourself on a smaller island...
“Wet toes” algorithm

Arkless strategy for flood victims

Then even a smaller one...
“Wet toes” algorithm

Arkless strategy for flood victims

Until eventually you’ll drown (if you can’t swim!)
“Wet toes” algorithm

Arkless strategy for flood victims

Finally, all land will be flooded!
“Wet toes” algorithm

Add links one by one and use a local algorithm to solve contradictions.
• This simple algorithm goes beyond the dynamical and the condensation transition for q=3 & 4
• For 3-coloring it allows to color in linear time until at least 4.55 (compares well with SP)
Algorithmic implications

1) + 2) Monte Carlo like (simulated annealing, random walker search) algorithms work. Easy to find and uniformly sample solutions.

3) Monte Carlo like algorithms do not equilibrate: some solutions can be found, but sampling is hard.

   *Belief Propagation* (RS) gives exact marginals. Solutions without frozen variable can be reached with local algorithm.

4) BP fails, *Survey Propagation (1RSB)* works! (Mézard, Zecchina, 2002)

5) No solutions anymore
Conclusions

- Determination of the phase diagram of the random coloring problem. A rich "glassy"-like phenomenology is found:
  - dynamical transition
  - clustering transition
  - "rigidity/freezing" transition

- Discussion of the algorithmic implication, importance of the rigidity transition

- Future works:
  - Numerical check of static and dynamic predictions?
  - Towards a rigorous formulation?
  - New algorithm?
References

Florent Krzakala (ESPCI Paris), Andrea Montanari (ENS Paris + Standford), Federico Ricci-Tersenghi (Rome), Guilhem Semerjian (ENS Paris), Lenka Zdeborová (Orsay):
Gibbs States and the Set of Solutions or Random Constraint Satisfaction Problems,
cond-mat/0612365, to be published in PNAS.

Lenka Zdeborová, Florent Krzakala:
Phase Transitions in the Coloring of Random Graphs,
arXiv:0704.1269

Florent Krzakala, Jorge Kurchan:
A landscape analysis of random constraint satisfaction problems,
cond-mat/0702546