Statistical Physics of Hard Optimisation Problems

Lenka Zdeborová
Work presented in my thesis


Statistical Physics of Hard Optimisation Problems
Why Optimization?

To save time, energy, money, resources, ....

Vast number of practical applications.
How difficult is optimization?

May be very hard ...
How difficult is optimization?

But in other cases quite easy ...
The Question

Which problems are hard and why?

We want to:

- recognize them
- make them easy
- avoid them
- use them

In particular pertinent combinatorial optimization, where the space of possible solutions is discrete.
My favorite example: Map coloring

**Four color theorem** - Conjectured 1852 *(Francis Guthrie, 1852)*
- Proven 1976 *(Appel, Haken’76)*
- \(O(N^2)\) algorithm *(Robertson, Sanders, Seymour, Thomas’96)*

And is it possible with only **three** colors?
- Not always and moreover **NP-complete** *(Cook’71)*.
Definition of graph coloring

- **State**: each node has a color
- **Rule (energy cost)**: neighbors have different colors
Definition of graph coloring
• State: each node has a color
• Rule (energy cost): neighbors have different colors
Coloring of general graphs

Applications in computer science - scheduling, register allocation ...

NP-complete, but typical case may be easier than the worst case.
Coloring of random graphs  
Pivotal step in understanding typical hardness

• What are random graphs?
  ‣ **Erdős-Rényi random graphs**: $N$ vertices, each couple connected with probability $c/(N-1)$, average degree $c$.
  ‣ **Random regular graphs**: Uniformly at random one of all the graphs in which each vertex has fixed degree.

• Why random graphs?
  ★ Natural definition of what *typical* means.
  ★ Many asymptotic properties can be computed 😊.
Where the really hard problems are

- **Answer #1**: Near to the colorability threshold.  
  *(Cheeseman, Kanefsky, Taylor’91; Mitchell, Selman, Levesque’92)*

- **Answer #2**: When the space of solutions is “clustered”.  
  *(Mezard, Parisi, Zecchina’02)*

- **Answer #3**: When clusters contain “frozen” variables.  
  *(LZ, Krzakala’07)*
Existence of sharp COL/UNCOL threshold

COL/UNCOL threshold at average constraint density $C_S$

- W.h.p. colorable for $C < C_S$ and w.h.p. uncolorable for $C > C_S$
- Proof of existence *(up to a detail, Friedgut, ’97)*
The easy-hard-”easy” pattern

Time of Davis-Putnam branch and bound algorithm needed to decide colorability increases close to the threshold *(Cheesman, Kanefsky, Taylor,‘91).*
Statistical physics approach

Coloring = zero temperature behavior of an anti-ferromagnetic Potts model. Hamiltonian:

\[ \mathcal{H} = \sum_{<ij>} \delta(s_i, s_j) \quad s_i = 1, 2, \ldots, q \]

The random graph - plays the role of quenched disorder.

Tree-like structure of the graph; as N grows the neighborhood of a random vertex is almost surely a tree up to distance \( \log_c N \).

Bethe approximation (mean field) equipped with the proper level of replica symmetry breaking is exact!

Random graphs coloring is exactly solvable via the cavity method (Mezard, Parisi ‘99)
Equations on a tree

\( \psi_{s_i}^{i \rightarrow j} \): Probability that node i takes color s_i when edge (constraint) (ij) is erased from the graph.

Recursive equations on a tree graph (Belief Propagation, BP):

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

Entropy follows from a similar expression.
Main elements of the cavity method

• **Replica symmetric solution:**
  Recursion on a tree graph  
  \( (\text{Mezard, Montanari, '05}) \) - correct only if correlations on distant boundary are on average forgotten on the root (decay of point-to-set correlation functions, not necessarily point-to-point).

• **One-step of replica symmetry breaking:**
  Split the set of solutions into exponentially many clusters such that within each cluster the point-to-set correlations decay. Use RS equations inside every cluster and then average over all of them. \( (\text{Mézard, Parisi, '99}) \)
How to imagine clusters?

• Intuitively: groups of nearby solutions which are in some sense disconnected from each other. Nearby = Hamming distance, overlap.

• In the cavity approach: pure states, extremal Gibbs measures (*Krzakala, Montanari, Ricci-Tersenghi, Semerjian, LZ ’07*)

• Alternative for computer scientist: clusters correspond to fixed points of Belief Propagation equations.
Selection of the most important previous results

- The exact **SAT/UNSAT (COL/UNCOL)** threshold computed.
  
  - K-SAT: Mézard, Zecchina, Parisi, '02,
  - q-COL: Mulet, Pagnani, Weigt, Zecchina, '03

- Prediction of a **glassy (clustered) phase** in the colorable region
  
  - cavity: Mézard, Zecchina, Parisi, '02, Biroli, Monasson, Weigt, '99
  - rigorous: Mézard, Mora, Zecchina, '05, Achlioptas, Ricci-Tersenghi, '06

- **Survey Propagation** algorithm designed
  
  - Mézard, Zecchina, Parisi, '02

  - The equations for the clustered phase used on a single graph.

  - **The best known** heuristic algorithm for large random SAT instances near to the threshold.
A refined analysis of clustering

\[ \Sigma = \frac{\log(\#\text{clusters})}{N} \]

Entropy (size) of a cluster \( s \):
logarithm of the number of solutions belonging to the cluster (divided by the number of variables).

Complexity function \( \Sigma(s) \):
logarithm of the number of clusters of size \( s \)

\[ N(s) = e^{N\Sigma(s)} \]

If \( \Sigma(s) > 0 \), there are exponentially many states of size \( s \).
If \( \Sigma(s) < 0 \), then states of size \( s \) become exponentially rare as \( N \).
A refined analysis of clustering

\[ \Sigma = \frac{\log(\#\text{clusters})}{N} \]

Entropy (size) of a cluster \( s \):
logarithm of the number of solutions belonging to the cluster (divided by the number of variables).

Complexity function \[ \Sigma(s) \):
logarithm of the number of clusters of size \( s \)

\[ N(s) = e^{N\Sigma(s)} \]

If \( \Sigma(s) > 0 \), there are exponentially many states of size \( s \).
If \( \Sigma(s) < 0 \), then states of size \( s \) become exponentially rare as \( N \)

★ We compute the complexity function \( \Sigma(s) \) using the entropic cavity method (Mézard, Palassini, Rivoire, ‘05).
Learning from $\sum(s)$

(with Florent Krzakala)

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 $c=20$
Glassy phase transitions

**Clustering transition**
- The phase space splits into exponentially many states
  \[ c_d(3) = 4, \quad c_d(4) = 8.35, \quad c_d(5) = 12.84 \]

**Condensation transition**
- Entropy dominated by finite number of the largest states.
  \[ c_c(3) = 4, \quad c_c(4) = 8.46, \quad c_c(5) = 13.23 \]

**COL/UNCOL transition**
- No more clusters, uncolorable phase
  \[ c_s(3) = 4.69, \quad c_s(4) = 8.90, \quad c_s(5) = 13.67 \]
Glassy phase transitions

**Clustering transition**
- The phase space splits into exponentially many states
  \[ c_d(3) = 4, c_d(4) = 8.35, c_d(5) = 12.84 \]

**Condensation transition**
- Entropy dominated by finite number of the largest states.
  \[ c_c(3) = 4, c_c(4) = 8.46, c_c(5) = 13.23 \]

**COL/UNCOL transition**
- No more clusters, uncolorable phase
  \[ c_s(3) = 4.69, c_s(4) = 8.90, c_s(5) = 13.67 \]

*Same phenomenology as in the ideal glass transition (ex: p-spin)*

Dynamic (Ergodicity Breaking) transition

Static (Kauzmann) transition
Glassy phase transitions

- **Clustering transition**
  - The phase space splits into exponentially many states
    \[ c_d(3) = 4, \quad c_d(4) = 8.35, \quad c_d(5) = 12.84 \]
  - *Dynamic (Ergodicity Breaking) transition*

- **Condensation transition**
  - Entropy dominated by finite number of the largest states.
    \[ c_c(3) = 4, \quad c_c(4) = 8.46, \quad c_c(5) = 13.23 \]
  - *Static (Kauzmann) transition*

- **COL/UNCOL transition**
  - No more clusters, uncolorable phase
    \[ c_s(3) = 4.69, \quad c_s(4) = 8.90, \quad c_s(5) = 13.67 \]

Moreover: The entropically dominating clusters are 1RSB stable in the colorable phase.
The freezing of clusters

Two types of clusters are found

- **Soft or “unfrozen” clusters**: All variables are allowed at least two different colors in the cluster.
- **Hard or “frozen” clusters**: A finite fraction of variables are allowed only one color in all solutions belonging to the cluster: we say that these variables “freeze”.

**Freezing transition**

Frozen variables appears in all the states.

\[ c_r(3) = 4.66, \quad c_r(4) = 8.83, \quad c_r(5) = 13.55 \]
Why should freezing be relevant for hardness?

1. On large instances we never find frozen solutions (many authors).

2. Only frozen clusters can disappear - incremental algorithms fail (Krzakala, Kurchan ‘07).

3. Compare freezing and performance of algorithms in 3-SAT. (Ardelius, LZ ‘08)

4. “Ultimate” challenge - the totally frozen locked constraint satisfaction problem (LZ, Mézard ’08)
Probability that an unfrozen solution exists in 3-SAT
(with John Ardelius)

![Graph showing probability of unfrozen solution against density of constraints for different values of N.]
Zoom on the rigidity transition in 3-SAT

probability unfrozen

density of constraints

$N=25$

$N=35$

$N=45$

$N=55$

$N=65$

$N=80$

$N=100$
The Locked Problems
(with Marc Mézard)

• **Definition:**
  - A closed loop has to be flipped to go from one solution to another one.

• **Examples:**
  - XOR-SAT on the core,
  - 1-in-3 SAT without leaves.

In locked problems *clusters are point-like.*
Example: 1-or-3-in-5 SAT
Example: 1-or-3-in-5 SAT
Example: 1-or-3-in-5 SAT
Locked problems: clustering = freezing

\[
\begin{align*}
\ell_d &= 3.07 \\
\ell_s &= 4.72
\end{align*}
\]
1-or-3-in-5 SAT

BP-reinforcement

\[ l_d \]

\[ l_s = 4.72 \]

\[ M = 4 \times 10^3 \]
\[ M = 2 \times 10^4 \]
\[ M = 10^5 \]
1-or-3-in-5 SAT

BP-reinforcement

$l_d$, $l_s$

$M=4\times10^3$
$M=2\times10^4$
$M=10^5$
1-or-3-in-5 SAT

stochastic local search

\[ l_d \]

\[ l_s \]

\[ M = 4 \times 10^3 \]

\[ M = 2 \times 10^4 \]

\[ M = 10^5 \]
1-or-3-in-5 SAT

stochastic local search

Easy

Hard

UNSAT

$M = 4 \cdot 10^3$

$M = 2 \cdot 10^4$

$M = 10^5$
Conclusions
Conclusions I: Structure of solutions

- Clustering/Dynamic transition $C_d$
- Condensation/Kauzmann transition $C_C$
- Freezing transition $C_R$
- COL/UNCOL transition $C_S$

average connectivity
Conclusions II: Algorithmic consequences

★ Freezing transition seems crucial

★ Locked problems: The really hard ones?
Many thanks for your attention!

Many thanks to:

LPTMS

Les services de la vie etudiante

CROUS
Supplements
We do not want to arrive here :)
The replica symmetric solution

After addition of node $i$ and all the edges $(ik)$ the entropy changes by $\Delta S^{i\rightarrow j}$, $Z^{i\rightarrow j} = e^{\Delta S^{i\rightarrow j}}$. The total entropy:

$$S = \frac{1}{N} \left( \sum_i \Delta S^i - \sum_{(ij)} \Delta S^{ij} \right)$$

Average over graph ensemble

$$\mathcal{P}(\psi) = \sum_k Q_1(k) \int \prod_{i=1}^{k-1} d\psi^i \mathcal{P}(\psi^i) \delta[\psi - \mathcal{F}(\{\psi^i\})]$$

Solution:

Only the paramagnetic $\psi = (1/q, 1/q, \ldots)$ in the COL phase.

$$s_{RS} = \log q + \frac{c}{2} \log \left( 1 - \frac{1}{q} \right)$$
How to compute the free entropy?

Order Parameter: Probability distribution of fields for every edge. Self-consistent equation:

\[ P^{i \rightarrow j}(\psi) = \frac{1}{Z_{i \rightarrow j}} \int \delta[\psi_{s_i}^{i \rightarrow j} - \mathcal{F}(\{\psi_{s_i}^{k \rightarrow i}\})] e^{m \Delta S_i^{i \rightarrow j}} \prod_{k \in V(i) - j} dP^{k \rightarrow i}(\psi) \]

Free entropy:

\[ N \Phi(m) = \sum_i \log \left( \int e^{m \Delta S_i} \prod_{k \in V(i)} (\psi_{k \rightarrow i}) \right) - \sum_{(ij)} \log \left( \int e^{m \Delta S_{ij}} dP^{i \rightarrow j}(\psi_{i \rightarrow j}) dP^{j \rightarrow i}(\psi_{j \rightarrow i}) \right) \]

+ average over the graph ensemble.

**Numerical Solution** - Population dynamics: Population of populations - very heavy!!! (Mézard, Palassini, Rivoire, 2005)
A refined analysis of clusters

- **Entropy (size) of a cluster $s$:**
  
  logarithm of the number of solutions belonging to the cluster (divided by the number of variables).

- **Complexity function $\sum(s)$:**
  logarithm of the number of clusters of size $s$

  \[ N(s) = e^{N \sum(s)} \]

  If $\sum(s) > 0$, there are exponentially many states of size $s$.
  If $\sum(s) < 0$, then states of size $s$ become exponentially rare as $N$ grows.

- We compute the complexity function using the **cavity method** via a Legendre transform $\Phi(m)$ of $\sum(s)$.

- **Main idea** (Mézard, Palassini, Rivoire, '05): weight each cluster by its size to the power $m$:

  \[
  e^{N\Phi(m)} = \sum_\alpha (e^{Ns_\alpha})^m = \int e^{N[ms + \sum(s)]} \, ds
  \]

  \[ \Phi(m) = ms + \sum(s), \quad \frac{\partial \sum(s)}{\partial s} = -m \]
Solving the functional equations

Analytical

- **Large q limit:** first three orders of development fully analytical.
- **Frozen variables at m=1:**

\[ \eta = \sum_k Q_1(k) \sum_{m=0}^{q-1} (-1)^m \left( \begin{array}{c} q - 1 \\ m \end{array} \right) \left( 1 - \frac{m}{q-1} \eta \right)^k \]

Simple functional equations

- **m=0 + frozen variables:** The results of Mézard, Zecchina, 2002 + Mulet, Pagnani, Weigt, Zecchina, 2002.
- **m=1:** using the results from Mézard, Montanari, 2005
- **Random regular (2-regular) graphs:** all edges equivalent, factorized solution.

Functionals of functionals, but only on soft variables.

- **Frozen variables present:** SP-like equations for frozen variables interconnected with the population dynamics for soft variables.
Solve (mostly numerically) the cavity equations

“Order parameter” $\mathcal{P}[P(\vec{\psi})]$ is a fixed point of functional equation.
+ Work out the many special cases when the equation simplifies
  (m=1, m=0, large q limit, frozen variables, regular graphs ...)

$P[ P(\vec{\psi}) ]$
Solve (mostly numerically) the cavity equations

“Order parameter” \( \mathcal{P}[P(\psi)] \) is a fixed point of functional equation.

+ Work out the many special cases when the equation simplifies
(m=1, m=0, large q limit, frozen variables, regular graphs ...)
Solve (mostly numerically) the cavity equations

“Order parameter” \( \mathcal{P}[P(\vec{\psi})] \) is a fixed point of functional equation.

+ Work out the many special cases when the equation simplifies (\( m=1, m=0, \) large q limit, frozen variables, regular graphs ...)

\[ \mathcal{P}[P(\vec{\psi})] \]

Our results (+ their meaning)
Determining freezing without knowledge of clusters?

- **Whitening** - determines if a solutions belongs to a frozen cluster.
  - In SAT - start with a solution and iteratively turn 0/1 into * if the variable belong only to already satisfied clauses (by definition * satisfies a clause).
  - In general - initialize Warning Propagation equations in a solution and iteratively find a fixed point.

- **Definition of freezing transition n. 2**: All solutions lead to a non-trivial whitening.
  - Equivalent to the definition n. 1 if $N >> 1$, graphs random, 1RSB correct.
  - But might be meaningful in general.

- **Observation**: On large graphs ($N > 20,000$) we are not able to find frozen solutions.
Entropy of 5-colorings of Erdős-Rényi graphs

ER graphs, $q=5$
Unsophisticated algorithm: Walk-COL (A-COL)

(1) *Randomly choose a spin* that has the same color as at least one of its neighbors.
(2) *Change randomly its color.* Accept this change with probability one if the number of unsatisfied spins has been lowered, otherwise accept it with probability $p$.
(3) If there are unsatisfied vertices, go to step (1) unless the maximum running time is reached.

![Graph showing the fraction of unsatisfied spins over time for different values of $c$ and $N$.](image)
Clustering = reason for computational hardness?

- At the clustering (dynamical) transition the Monte-Carlo equilibration time diverges \((\text{Montanari, Semerjian, 2005})\). Thus Monte-Carlo sampling is hard.
- But finding solution is a question of \textit{basin of attraction}, further analysis of the energy landscape needed \((\text{Krzakala, Zdeborova, work in progress})\).
1-or-3-in-5 SAT 010100

\( M=4 \times 10^3 \), \( T=5 \times 10^4 \)
\( M=2 \times 10^4 \), \( T=5 \times 10^4 \)
\( M=2 \times 10^4 \), \( T=5 \times 10^5 \)
\( M=1 \times 10^5 \), \( T=5 \times 10^5 \)

\( l_s = 4.72 \)

stochastic local search
Clusters as connected components in 3-SAT

\[ \frac{\langle \log(\#) \rangle}{N} \]

vs.

Density of constraints

N=25
N=50
N=75
N=100
N=125
N=150

Theory

back