Where the *really* hard problems are...

A highly-biased review of the “statistical physics” picture of random constraint satisfaction problems

Florent Krzakala (ESPCI Paristech)

In collaboration with
Lenka Zdeborová (LANL)
Where the \textit{really} hard problems are...

A highly-biased review of the “statistical physics” picture of random constraint satisfaction problems

Florent Krzakala (ESPCI Paristech)

In collaboration with
Lenka Zdeborová (LANL)

And using many results from people in the audience !!!

Monday, November 23, 2009
Where the \textit{really} hard problems \textit{really} are...

A highly-biased review of the “statistical physics” picture of random constraint satisfaction problems

Florent Krzakala (ESPCI Paristech)

In collaboration with
Lenka Zdeborová (LANL)

And using many results from people in the audience !!!
This Talk

- A bit of history...
- Clustering: The “physicist” scenario and exact simulations of impossible-to-simulate models.
- Hard problems and the energy landscape: where the really hard problems really are.
A bit of history....
Random CSPs

Many constraint satisfaction problems display a similar behavior on random instances.
Random CSPs

Many constraint satisfaction problems display a similar behavior on random instances.

- A SAT/UNSAT transition as the ration of clauses per variables is increased.
- A peak in the difficulty of checking satisfiability close to the sat transition.
Random CSPs

Many constraint satisfaction problems display a similar behavior on random instances.

A SAT/UNSAT transition as the ration of clauses per variables is increased.

A peak in the difficulty of checking satisfiability close to the sat transition.
Random CSPs

Many constraint satisfaction problems display a similar behavior on random instances.

![Graph showing SAT/UNSAT transition and peak difficulty close to sat transition.](image)

![Graph with curves for PSAT and comp. time, showing median computation time.](image)
Random CSPs

Many constraint satisfaction problems display a similar behavior on random instances.

3-COL

[Graph showing percent satisfiable vs. ratio of clauses to variables for different numbers of variables (50, 40, 30, 20, 10).]

[Graph showing median computational cost for different numbers of variables (N = 100, 71, 50, 100) and different values of M (60, 70, 80, 90, 100).]

Monday, November 23, 2009
Where the Really Hard Problems Are

Peter Cheeseman
RIACS*

Bob Kanefsky
Sterling Software
Artificial Intelligence Research Branch
NASA Ames Research Center, Mail Stop 244-17
Moffett Field, CA 94035, USA
Email: <last-name@ptolemy.arc.nasa.gov>

William M. Taylor
Sterling Software

Abstract

It is well known that for many NP-complete problems, such as K-Sat, etc., typical cases are easy to solve; so that computationally hard cases must be rare (assuming P = NP). This paper shows that NP-complete problems can be summarized by at least one "order parameter", and that the hard problems occur at a critical value of such a parameter. This critical value separates two regions of charac-

In this paper we show that for many NP problems one or more "order parameters" can be defined, and hard instances occur around particular critical values of these order parameters. In addition, such critical values form a boundary that separates the space of problems into two regions. One region is underconstrained, so the density of solutions is high, thus making it relatively easy to find a solution. The other region is overconstrained and very unlikely to contain a solution. If there are solutions in this overconstrained region, then they have such a deep local minimum (strong basin of attraction) that any
Where the really hard problem are

Where the Really Hard Problems Are

Peter Cheeseman
RIACS*

Bob K.
Sterling
Artificial Intelligence
NASA Ames Research Center
Moffett Field,
Email: < last-name@ame

Abstract

It is well known that for many NP-complete problems, such as K-Sat, etc., typical cases are easy to solve; so that computationally hard cases must be rare (assuming P = NP). This paper shows that NP-complete problems can be summarized by at least one "order parameter" and that the hard and easy cases are on opposite sides of a boundary that separates the space of problems into two regions. One region is underconstrained, so the density of solutions is high, thus making it relatively easy to find a solution. The other region is overconstrained and very unlikely to contain a solution. If there are solutions in this overconstrained region, then they have such deep local minimum (strong basin of attraction) that any

The results reported above suggest the following conjecture:

All NP-complete problems have at least one order parameter and the hard to solve problems are around a critical value of this order parameter. This critical value (a phase transi-
Where the really hard problem are

Peter Cheeseman
RIACS*

Bob K.
Sterling
Artificial Intelligence
NASA Ames Research
Moffett Field,
Email: <last-name>

Abstract
It is well known that for many NP-complete problems, such as K-Sat, etc., typical cases are easy to solve; so that computationally hard cases must be rare (assuming P = NP). This paper shows that NP-complete problems can be summarized by at least one "order parameter" and that this order parameter varies continuously from easy to hard.

The results reported above suggest the following conjecture:

All NP-complete problems have at least one order parameter and the hard to solve problems are around a critical value of this order parameter. This critical value (a phase transi-

cut on early in the search. Really hard problems occur on the boundary between these two regions, where the probability of a solution is low but non-negligible. At this point there are typically many local minima corresponding to almost solutions separated by high "energy barriers". These almost solutions form deep local minima that may often trap search methods that rely on local information.
The energy landscape

Cost

Solution

Quasi-solution

Landscape

Config
The energy landscape

Solution

Quasi-solution

Cost

Config

Landscape
The energy landscape

Cost

Solution

Quasi-solution

Landscape

Config

Monday, November 23, 2009
The energy landscape

Cost

Landscape

Solution

Quasi-solution

Config

Monday, November 23, 2009
The energy landscape

Cost

Landscape

Solution

Quasi-solution

Config
The energy landscape

Cost

Landscape

Solution

Quasi-solution

Config
The energy landscape

Cost

Solution

Quasi-solution

Landscape
The energy landscape

Cost

Solution

Quasi-solution

Landscape

Config

Monday, November 23, 2009
Where the really hard problem are

Answer #1: Near to the colorability threshold.

(Cheeseman, Kanefsky, Taylor’91; Mitchell, Selman, Levesque’92)
The clustering transition and the glassy phase

Analytic and Algorithmic Solution of Random Satisfiability Problems

M. Mézard,¹ G. Parisi,¹,² R. Zecchina¹,³*

We study the satisfiability of random Boolean expressions built from many clauses with $K$ variables per clause ($K$-satisfiability). Expressions with a ratio $\alpha$ of clauses to variables less than a threshold $\alpha_c$ are almost always satisfiable, whereas those with a ratio above this threshold are almost always unsatisfiable. We show the existence of an intermediate phase below $\alpha_c$, where the proliferation of metastable states is responsible for the onset of complexity in search algorithms. We introduce a class of optimization algorithms that can deal with these metastable states; one such algorithm has been tested successfully on the largest existing benchmark of $K$-satisfiability.
Glassy landscape: many minima

- $E=\ldots$
- $E=2$
- $E=1$
- $E=0$
Glassy landscape: many minima

The energy landscape is glassy and has many "trapping" minima in the whole clustered phase.
“Hard SAT” phase?
“Hard SAT” phase?

The “$q \log q$” problem

- Consider $q$ color (with $q$ large enough) and a large random graph of average degree $c$
- W.h.p this graph is colorable if $c < 2q \log q$
- However, no algorithm is able to do so efficiently (polynomial) for $c > q \log q$!

Talk of L. Kirousis

D. Achlioptas et al. Nature 2005

$q \log q$

$2q \log q$

Average degree $c$
"Hard SAT" phase?

The "$q \log q$" problem

- Consider $q$ color (with $q$ large enough) and a large random graph of average degree $c$
- W.h.p this graph is colorable if $c < 2q \log q$
- However, no algorithm is able to do so efficiently (polynomial) for $c > q \log q$!
"Hard SAT" phase?

The “$q \log q$” problem

- Consider $q$ colors (with $q$ large enough) and a large random graph of average degree $c$
- W.h.p this graph is colorable if $c < 2q \log q$
- However, no algorithm is able to do so efficiently (polynomial) for $c > q \log q$!
“Hard SAT” phase?

The “$q \log q$” problem

- Consider $q$ color (with $q$ large enough) and a large random graph of average degree $c$
- W.h.p this graph is colorable if $c < 2q \log q$
- However, no algorithm is able to do so efficiently (polynomial) for $c > q \log q$

No one has ever seen the solution of, say 5-coloring, for large enough $c$ and $N=10^6$

---

Talk of L. Kirousis

D. Achlioptas et al. Nature 2005
"Hard SAT" phase ?

- $2^{K/K \log K}$
- $2^{K \log 2}$
- $q \log q$
- $2q \log q$

Easy $\rightarrow$ HARD $\rightarrow$ UNSAT $\rightarrow$ UNCOL $\rightarrow$ Easy

Average degree $c$
“Hard SAT” phase?

\[ 2^{K/K \log K} \]

\[ 2^K \log 2 \]

\[ q \log q \]

\[ 2q \log q \]

\[ \alpha \]

Average degree \( c \)
Where the really hard problem 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” and the landscape glassy
(Mezard, Parisi, Zecchina’02)
So... everything is fine?
So... everything is fine?

No!!! We have a major problem!!!
The “Hard” phase is (sometime) easy!

3-col: clustered phase for $c>4$... but a rigorous simple algorithm works until $c=4.03$ !!
(Achlioptas & Moore ’03)
The “Hard” phase is (sometime) easy!

3-col: clustered phase for $c>4$... but a rigorous simple algorithm works until $c=4.03$ !!
(Achlioptas & Moore ’03)

3-col: monte carlo annealing works well much beyond $c=4$ (Saad, Van Mourik 02’)
The “Hard” phase is (sometime) easy!

3-col: clustered phase for $c > 4$... but a rigorous simple algorithm works until $c = 4.03$ !!
(Achlioptas & Moore ’03)

3-col: monte carlo annealing works well much beyond $c = 4$ (Saad, Van Mourik 02’)

RANDOM GRAPH COLORING: STATISTICAL PHYSICS …

PHYSICAL REVIEW E 66, 056120 (2002)
The “Hard” phase is (sometime) easy!

3-col: clustered phase for $c>4$... but a rigorous simple algorithm works until $c=4.03$!!
(Achlioptas & Moore ’03)

3-col: Monte-Carlo annealing works well much beyond $c=4$ (Saad & Van Mourik 02’
The “Hard” phase is (sometime) easy!

3-col: clustered phase for \( c > 4 \ldots \) but a rigorous simple algorithm works until \( c = 4.03 \) !!
(Achlioptas & Moore ’03)

3-col: Monte-Carlo annealing works well much beyond \( c = 4 \) (Saad & Van Mourik 02’)

K-SAT: Walk-SAT and variant can finds solutions in the clustered phase (M. Alava et al. PNAS ‘08, Ardelius & Aurell ‘06)
The “Hard” phase is (sometime) easy!

3-col: clustered phase for $c>4$... but a rigorous simple algorithm works until $c=4.03$ !!
   (Achlioptas & Moore ’03)

3-col: Monte-Carlo annealing works well much beyond $c=4$ (Saad & Van Mourik 02’)

K-SAT: Walk-SAT and variant can finds solutions in the clustered phase (M. Alava et al.
   PNAS ’08, Ardelius & Aurell ’06)

Q-col: The same happens in coloring (Zdeborova & FK ‘07, Kurchan & FK ‘07)
Is something wrong?

- The predictions of clustering is made with the replica/cavity method, which is not rigorous...

- ... so we need to check the cavity predictions.

- Rigorous technics tell us there IS clustering, but proofs work only for $q$ and $K$ large enough... and for $c$ and $\alpha$ larger than the predicted ones.

- It is thus important to check the cavity predictions in simulations.

Mora, Zecchina, Daudé, Mézard, Achlioptas, Ricci-Tersenghi, Coja-Oghlan
Puzzles....

- Prediction of clustering transition...
- ... but rigorous proofs exist only for large enough q (color) or K(in SAT).

Why is the clustered phase sometime so easy if the landscape is so glassy and complex?
Clustering:
Simulations of impossible-to-simulate models
A closer look to clustering

Clusters can be seen as fixed points of Belief propagation.

In the cluster phase, there are many such fixed points, and SP is doing statistics over them!
A closer look to clustering

Clusters can be seen as fixed points of Belief propagation
In the cluster phase, there are many such fixed points, and SP is doing statistics over them!
A closer look to clustering

Clusters can be seen as fixed points of Belief propagation
In the cluster phase, there are many such fixed points, and SP is doing statistics over them!

Proofs?

All rigorous works concentrate on fixed point of BP with frozen variables

But at the onset of clustering, variables are NOT frozen!
A closer look to clustering

Clusters can be seen as fixed points of Belief propagation
In the cluster phase, there are many such fixed points, and SP is doing statistics over them!

Proofs?
All rigorous works concentrate on fixed point of BP with frozen variables
But at the onset of clustering, variables are NOT frozen!

Simulations?
The cavity prediction is for “typical solutions”
Perfect sampling is almost impossible numerically!!!!!
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with $N$ nodes and $M$ links.

1) Color randomly the $N$ nodes
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with N nodes and M links.

1) Color randomly the N nodes

2) Put the M links randomly such that the planted configuration is a proper coloring
Consider the 3-coloring problem with $N$ nodes and $M$ links.

1) Color randomly the $N$ nodes
2) Put the $M$ links randomly such that the planted configuration is a proper coloring
3) Now, we have created a problem for which we know the solution
Faster than-the-clock simulations

Do this procedure for $c < c_c$. According to Lenka’s talk yesterday, this procedure is equivalent to the picking a typical solution at random in a random graph.

This has actually been proven rigorously in some cases

- Montanari and Semerjian ‘06, for XOR-SAT
- Achlioptas and Coja-Oghlan ’08, for q-coloring but only (so far) for $c_q < c_c$
Faster than-the-clock simulations

Do this procedure for $c < c_c$. According to Lenka’s talk yesterday, this procedure is equivalent to the picking a typical solution at random in a random graph.

This has actually been proven rigorously in some cases

- Montanari and Semerjian ‘06, for XOR-SAT
- Achlioptas and Coja-Oghlan ’08, for q-coloring but only (so far) for $c_q \leq c_c$

With $c_q = 3.83$ (q=3) versus $c_c = 4$
$c_q = 7.81$ (q=4) versus $c_c = 8.46$
$c_q = 12.52$ (q=5) versus $c_c = 13.23$
$c_q = 17.70$ (q=6) versus $c_c = 18.44$
$c_q = 29.15$ (q=8) versus $c_c = 29.90$
$c_q = 41.70$ (q=10) versus $c_c = 42.53$

and $c_q \rightarrow c_c$ for $q \rightarrow \infty$ to leading order
Faster than-the-clock simulations

Do this procedure for $c < c_c$. According to Lenka’s talk yesterday, this procedure is equivalent to the picking a typical solution at random in a random graph.

This has actually been proven rigorously in some cases

- Montanari and Semerjian ‘06, for XOR-SAT
- Achlioptas and Coja-Oghlan ’08, for q-coloring but only (so far) for $c_q < c_c$

With $c_q = 3.83$ (q=3) versus $c_c = 4$
$c_q = 7.81$ (q=4) versus $c_c = 8.46$
$c_q = 12.52$ (q=5) versus $c_c = 13.23$
$c_q = 17.70$ (q=6) versus $c_c = 18.44$
$c_q = 29.15$ (q=8) versus $c_c = 29.90$
$c_q = 41.70$ (q=10) versus $c_c = 42.53$

and $c_q \rightarrow c_c$ for $q \rightarrow \infty$ to leading order

Use them in simulations!
Testing the clustering predictions

**Prediction:** beyond the so-called “clustering” threshold, a non-trivial and non-factorized fixed point of BP is obtained if one starts from an equilibrium configuration

\[
\psi_{\text{factorized}} = \left( \frac{1}{q}, \frac{1}{q}, \ldots, \frac{1}{q} \right)
\]

\[
M = \frac{1}{qN} \sum_{\text{graph}} \sum_{c=1}^{q} \frac{\psi_{BP}^{c,i} - \frac{1}{q}}{1 - \frac{1}{q}}
\]
Testing the clustering predictions

**Prediction:** beyond the so-called “clustering” threshold, a non-trivial and non factorized fixed point of BP is obtained if one starts from an equilibrium configuration

\[
\psi_{factorized} = \left( \frac{1}{q}, \frac{1}{q}, \ldots, \frac{1}{q} \right)
\]

\[
M = \frac{1}{qN} \sum_{graph} \sum_{c=1}^{q} \left( \psi_{BP}^{c,i} - \frac{1}{q} \right) \frac{1}{1 - \frac{1}{q}}
\]

Simulation with N=10^6
Cavity predictions for the clustering transition can be tested with great accuracy!
Cavity predictions for the clustering transition can be tested with great accuracy!

Clustering is definitively present!
Cavity predictions for the clustering transition can be tested with great accuracy!

Clustering is definitively present!

In order to close the gap between rigorous proofs and the “exact” cavity predictions it is necessary to look on non-frozen variables.
Cavity predictions for the clustering transition can be tested with great accuracy!

Clustering is definitively present!

In order to close the gap between rigorous proofs and the “exact” cavity predictions it is necessary to look on non-frozen variables.

Maybe this can be done via reconstruction on tree?

cf Montarani & Mézard ’06, FK, Montanari, Ricci-Tersenghi, Semerjian & Zdeborova ’07 Sly ’08, Maneva ’09...
Algorithms:
Where the really hard problems really are...
Landscapes: canyons, mountains and valleys...
A question of basins of attraction

Canyon dominated vs. Valley dominated

EASY vs HARD
How to compute the feature of the landscape?
First concentrate on the energy where large barriers appear
First concentrate on the energy where large barriers appear.
Then take a point at random on this line
Then take a point at random on this line
Compute the size of the “valley” below this point for different energies.
Compute the size of the "valley" below this point for different energies
Now you know everything about the typical valley starting from this energy
Now you know everything about the typical valley starting from this energy
And you know where the dynamics will end!
And you know where the dynamics will end!
This can be done using the “following state” formalism

FK & Zdeborova arxiv:0909.3820
Zdeborova Talk yesterday
This can be done using the “following state” formalism

FK & Zdeborova arxiv:0909.3820  Zdeborova Talk yesterday

...and also by planting!
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with N nodes and M links.

1) Color randomly the N nodes
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with N nodes and M links.

1) Color randomly the N nodes
2) Put the M links randomly such that the planted configuration has a finite fraction of mistake
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with N nodes and M links.

1) Color randomly the N nodes
2) Put the M links randomly such that the planted configuration has a finite fraction of mistake
3) Now, we have created a problem for which we know a typical configuration at this energy
The Planted Ensemble in the coloring problem

Consider the 3-coloring problem with \( N \) nodes and \( M \) links.

1) Color randomly the \( N \) nodes

2) Put the \( M \) links randomly such that the planted configuration has a finite fraction of mistakes

3) Now, we have created a problem for which we know a typical configuration at this energy

Now simply use BP initialized in this configuration for different temperatures!

(temperature allows to select the desired energy)
Valleys
3-XOR-SAT with L=3

Canyons
4-coloring of 9-regular random graphs
Valleys

3-XOR-SAT with L=3

Canyons

4-coloring of 9-regular random graphs
When do we go from one to the other?

Canyon dominated vs. Valley dominated

EASY vs HARD

Zero energy states

Positive energy states

Monday, November 23, 2009
When do we go from one to the other?

Canyon dominated vs. Valley dominated

EASY vs HARD

When the typical solutions are frozen (BP marginal fully biased) then we are in a valley dominated landscape!
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?
Take-Home message

Why does a glassy landscape is able to break ergodicity but yet allow some solutions to be found easily?

The focus will be on future developments and interactions between physics and computer science, thus brave conjectures and challenges will be very appropriate.

Thank you,
Elitza and Matteo
Where the really hard problem 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” and the landscape glassy
(Mezard, Parisi, Zecchina’02)
Where the really hard problem really 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” and the landscape glassy  
(Mezard, Parisi, Zecchina’02)

Answer #3: When all or most of solutions are frozen so that the landscape is truly complicated....  
(FK & Zdeborova “07-09)
A set of questions and conjectures

**Monte Carlo:**
- Is it possible to sample efficiently beyond the clustering transition?
- It is possible to find solution beyond the clustering transition?
- It is possible to find solution in the hard valleys-dominated region?

**Local algorithms:**
- Can they find solutions beyond clustering?
- It is possible to find solution in the hard valleys-dominated region?

**Message passing:**
- Can they find solutions beyond clustering?
- It is possible to find solution in the hard valleys-dominated region?
- Is it possible to find solution when all solutions are frozen?
A set of questions and conjectures

Monte Carlo:
- Is it possible to sample efficiently beyond the clustering transition?
- It is possible to find solution beyond the clustering transition?
- It is possible to find solution in the hard valleys-dominated region?

Conjectures
NO

Local algorithms:
- Can they find solutions beyond clustering?
- It is possible to find solution in the hard valleys-dominated region?

Message passing:
- Can they find solutions beyond clustering?
- It is possible to find solution in the hard valleys-dominated region?
- Is it possible to find solution when all solutions are frozen?
A set of questions and conjectures

**Monte Carlo:**
- Is it possible to sample efficiently beyond the clustering transition?
- Is it possible to find solution beyond the clustering transition?
- Is it possible to find solution in the hard valleys-dominated region?

**Local algorithms:**
- Can they find solutions beyond clustering?
- Is it possible to find solution in the hard valleys-dominated region?

**Message passing:**
- Can they find solutions beyond clustering?
- Is it possible to find solution in the hard valleys-dominated region?
- Is it possible to find solution when all solutions are frozen?

Conjectures

- NO
- YES
A set of questions and conjectures

Monte Carlo:
- Is it possible to sample efficiently beyond the clustering transition? **NO**
- It is possible to find solution beyond the clustering transition? **YES**
- It is possible to find solution in the hard valleys-dominated region? **NO**

Local algorithms:
- Can they find solutions beyond clustering? **NO**
- It is possible to find solution in the hard valleys-dominated region? **NO**

Message passing:
- Can they find solutions beyond clustering? **NO**
- It is possible to find solution in the hard valleys-dominated region? **NO**
- Is it possible to find solution when all solutions are frozen? **NO**
A set of questions and conjectures

**Monte Carlo:**
- Is it possible to sample efficiently beyond the clustering transition? **NO**
- Is it possible to find solution beyond the clustering transition? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**

**Local algorithms:**
- Can they find solutions beyond clustering? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **YES**

**Message passing:**
- Can they find solutions beyond clustering? **NO**
- Is it possible to find solution in the hard valleys-dominated region? **NO**
- Is it possible to find solution when all solutions are frozen? **NO**
**A set of questions and conjectures**

**Monte Carlo:**
- Is it possible to sample efficiently beyond the clustering transition? **NO**
- Is it possible to find solution beyond the clustering transition? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**

**Local algorithms:**
- Can they find solutions beyond clustering? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**

**Message passing:**
- Can they find solutions beyond clustering? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**
- Is it possible to find solution when all solutions are frozen? **NO**
### A set of questions and conjectures

<table>
<thead>
<tr>
<th><strong>Monte Carlo:</strong></th>
<th><strong>Conjectures</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it possible to sample efficiently beyond the clustering transition?</td>
<td>NO</td>
</tr>
<tr>
<td>It is possible to find solution beyond the clustering transition?</td>
<td>YES</td>
</tr>
<tr>
<td>It is possible to find solution in the hard valleys-dominated region?</td>
<td>NO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Local algorithms:</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Can they find solutions beyond clustering?</td>
<td>YES</td>
</tr>
<tr>
<td>It is possible to find solution in the hard valleys-dominated region?</td>
<td>NO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Message passing:</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Can they find solutions beyond clustering?</td>
<td>YES</td>
</tr>
<tr>
<td>It is possible to find solution in the hard valleys-dominated region?</td>
<td></td>
</tr>
<tr>
<td>Is it possible to find solution when all solutions are frozen?</td>
<td></td>
</tr>
</tbody>
</table>
### A set of questions and conjectures

**Monte Carlo:**
- Is it possible to sample efficiently beyond the clustering transition? **NO**
- Is it possible to find solution beyond the clustering transition? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**

**Local algorithms:**
- Can they find solutions beyond clustering? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **NO**

**Message passing:**
- Can they find solutions beyond clustering? **YES**
- Is it possible to find solution in the hard valleys-dominated region? **YES**
A set of questions and conjectures

Monte Carlo:

- Is it possible to sample efficiently beyond the clustering transition?
- Is it possible to find solution beyond the clustering transition?
- Is it possible to find solution in the hard valleys-dominated region?

Local algorithms:

- Can they find solutions beyond clustering?
- Is it possible to find solution in the hard valleys-dominated region?

Message passing:

- Can they find solutions beyond clustering?
- Is it possible to find solution in the hard valleys-dominated region?
- Is it possible to find solution when all solutions are frozen?

Conjectures

- NO
- YES
- NO
- YES
- NO
- YES
- YES

Dall’Asta, Ramezanpour and Zecchina ’08
Zdeborova ’08 (private communication)
A set of questions and conjectures

Monte Carlo:
- Is it possible to sample efficiently beyond the clustering transition?  
  Conjectures: NO
- It is possible to find solution beyond the clustering transition?  
  Conjectures: YES
- It is possible to find solution in the hard valleys-dominated region?  
  Conjectures: NO

Local algorithms:
- Can they find solutions beyond clustering?  
  Conjectures: YES
- It is possible to find solution in the hard valleys-dominated region?  
  Conjectures: NO

Message passing:
- Can they find solutions beyond clustering?  
  Conjectures: YES
- It is possible to find solution in the hard valleys-dominated region?  
  Conjectures: YES
- Is it possible to find solution when all solutions are frozen?  
  Conjectures: NO

Dall’Asta, Ramezanpourand & Zecchina ’08
Zdeborova ’08 (private communication)
Example: Studying Monte-Carlo annealings starting from equilibrium

XOR-SAT problems (Parity-check)

\[ \mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2} \]

**Prediction:**
Monte Carlo cooling and heating follow a well defined line
Example:

Studying Monte-Carlo annealings starting from equilibrium

XOR-SAT problems (Parity-check)

\[ H(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2} \]

\( N = 200\,000 \) spins

**Prediction:**
Monte Carlo cooling and heating follow a well defined line

Monday, November 23, 2009
Example:

Studying Monte-Carlo annealings starting from equilibrium

XOR-SAT problems (Parity-check)

\[ \mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk}S_i S_j S_k}{2} \]

\( N=200,000 \) spins

**Prediction:**
Monte Carlo cooling and heating follow a well defined line
Example:

Studying Monte-Carlo annealings starting from equilibrium

Monte Carlo cooling and heating follow a well defined line

XOR-SAT problems (Parity-check)

\[ \mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2} \]

N=200 000 spins

Temperature

**Prediction:**

Monte Carlo cooling and heating follow a well defined line
Example:

Studying Monte-Carlo annealings starting from equilibrium

XOR-SAT problems (Parity-check)

\[
\mathcal{H} (\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}
\]

\(N=200,000\) spins

**Prediction:**

Monte Carlo cooling and heating follow a well defined line