Satisfiability and Percolation in Two Dimensions

J. M. Schwarz and A. A. Middleton

Dagstuhl Seminar, September 17, 2003

[cond-mat/0309240]

Playing with P-NPC & physics

Exploring how ideas from statistical physics can be applied to solving hard problems with a *finite-dimensional* structure.

- Phase transitions (cf. w/ mean field)
- Percolation
- Rare region arguments
- Thermodynamic state

Exploring problem structures and algorithms.

Faster algorithms $\Rightarrow$ Better physics (?)

Ensemble of formulas

Given N Boolean variables $X = \{x_1, x_2, \ldots, x_N\}$ (with literals $Y = \{x_1, \ldots, x_N, \overline{x}_1, \ldots, \overline{x}_N\}$), K-SAT formulas are conjunctive compositions of disjunctive clauses $c_m = \vee_{l=1}^{K} y_l^m$:

$$F = \wedge_{m=1}^{M} \vee_{l=1}^{K} y_l^m$$

Ensemble of formulas

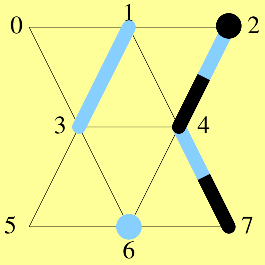
$$F = \wedge_{m=1}^{M} \vee_{l=1}^{K} y_l^m$$

We study the ensemble of formulas:

- that are *two-dimensional*
- with $K = 1, 2$ [$\alpha = M_2/N$, $\gamma = M_1/N$]
- whose clauses connect *nearest neighbor* variables.
- have literals with prob. $1/2$ of being negated variables.
- *Note:* spin glass in (correlated) field.

Example

$(x_1 \vee x_3) \wedge (x_2 \vee \overline{x}_4) \wedge (x_4 \vee \overline{x}_7) \wedge (\overline{x}_2) \wedge (x_6)$



How about that SAT/UNSAT transition?

Absent, since

two loops are much easier in finite d.

For any $\alpha = M_2/N$, $\exists$, with prob. 1, unsatisfiable subgraphs (not true in MFT.)

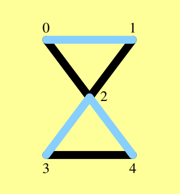
How about that SAT/UNSAT transition?

Absent, since

two loops are much easier in finite d.

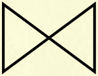
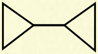
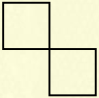
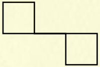
For any $\alpha = M_2/N$, $\exists$, with prob. 1, unsatisfiable subgraphs (not true in MFT.)

Not just a simple frustrated cycle, but two "*forcing triangles*":



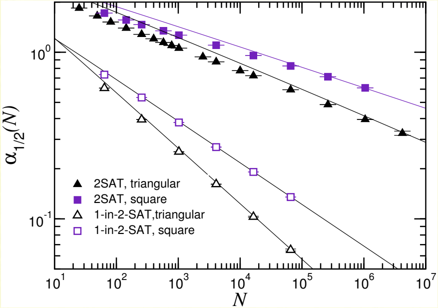
Density of UNSAT subgraphs

Standard enumeration:

Graph	Density	Graph	Density
	$\frac{\alpha^6}{10\,368} + O(\alpha^7)$		$\frac{25\alpha^7}{279\,936} + O(\alpha^8)$
	$\frac{\alpha^8}{65\,536} + O(\alpha^9)$		$\frac{\alpha^9}{16\,384} + O(\alpha^{10})$

Crossover location

$\alpha_{1/2}(N = L^2)$ is where $1/2$ of the graphs are satisfiable.



In general, for $\rho_r = c_r \alpha^r$, $\alpha_{1/2}(L) = (c_r^{-1/r} \ln 2) N^{-1/r}$.

What to do?

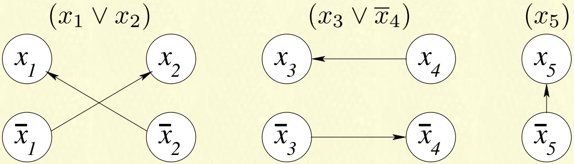
Small *double* loops make the SAT transition *coarse*.

So look at the big loops:

- Use implication digraphs to untangle logical structure.
- *Large* contradictory logical structures appear at an $\alpha_S > \alpha_c$, where α_c is edge percolation.

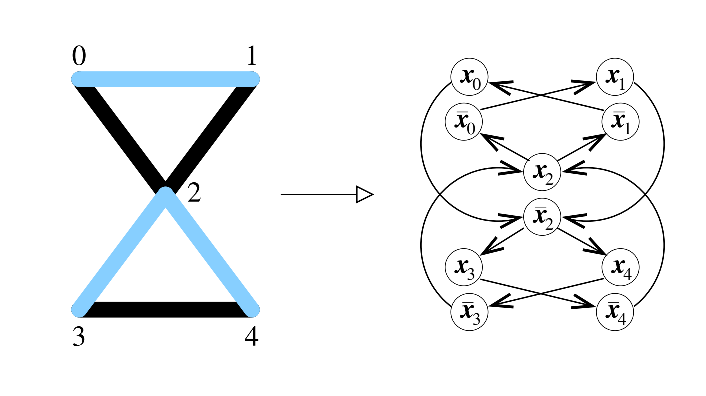
Digraphs

Implication digraphs replace 2-clauses with 2 implications and 1-clauses with a single implication:



Construct set of directed edges among the *literals* for a given formula.

Two forcing triangles

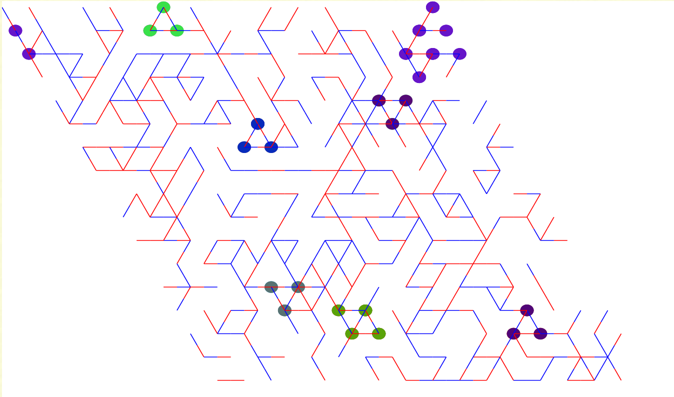


Extracting structure

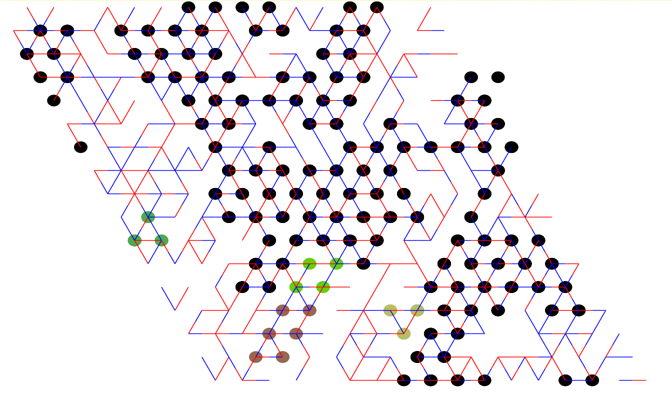
Can find *strongly connected components with contradictions* (CSCs) in time linear in $(\alpha + 1)N$.

For drawing purposes, project onto *variables*.

Extracting structure: “small” α



Extracting structure: larger α



Decomposing structure

- CSCs are logically independent.
- So can study 2SAT or MAX2SAT using single CSCs.
- Seems this can be treated as percolation problem:
  - When spans, large component to logical structure.
  - Checks show *consistent with* standard percolation
    - Cluster distribution and finite-size scaling.
    - $\tau = 2.02(5)$, $\nu = 1.32(3)$ [cf. $187/91$ and $4/3$.]

Percolation

Further support:

- Pfeiffer & Reiger [PRE **67**, 056113 (2001)] studied uncorrelated directed edges. Quite consistent with standard percolation.
- Bollobas, Borgs, Chayes, Kim & Wilson [Rand. Struct. Alg. **18**, 201 (2001)] relate implication digraphs to standard percolative graphs, to study width of 2SAT transition in MFT. This mapping *works* in finite-d, but it doesn't directly address CSC percolation.

Running times for MAXSAT

- Planar MAX2SAT is NP-hard!
- Adapted Borchers & Furman code [DPLL after initial heuristic - see J. Comb. Optim. **2**, 299 (1999)] to get exact solutions.
- Timing measured in # of backtracks.
- We find that the median or exp(mean of log of running time) for a cluster of size s scales as

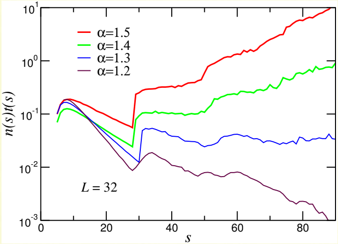
$$t^*(s,\alpha) \sim \exp(s/s_\tau(\alpha))$$

  where s_τ is only weakly dependent on α.
- Combine with cluster distribution $n(s) \sim \exp(-s/s_\xi(\alpha))$, where s_ξ diverges as α approaches α_c.

Dynamic divergence

Define α_G, $\alpha_c < \alpha_G < \alpha_S$, where the exponential growth of running time exceeds the exponential decay of the cluster size distribution.

$$\alpha_G \approx 1.3$$

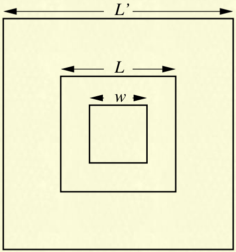


Dynamic divergence

- For $\alpha < \alpha_G$, can solve in time $\propto L^2$.
- For $\alpha > \alpha_G$, median solution time should be L^{2s_ξ/s_τ}, while mean time diverges (dominated by system size clusters.)

How many thermodynamic states?

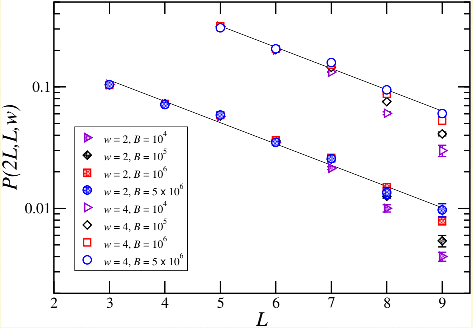
1. Solve for the ground state in area L^2 system.
2. Solve again over area $(L')^2$.
3. Compare the solutions in a window of area w^2.
4. Repeat for many samples $\Rightarrow P(L', L, w)$.
5. Solved WMAXSAT (uniform distribution of weights) to resolve degeneracies.



Exponential convergence to a single state

Use $\alpha = 1.7$, $\gamma = 0.2$, above CSC percolation.

B is max. # of backtracks (bounds run time, take $B \to \infty$).



Some speculation, single state systems

- Solve for ground states of many overlapping subsamples of linear dimension $\ell \gg \xi$.
- Patch together the solutions (agreement in overlapping areas).
- This will fail, of course, in rare cases.

Wrapping up . . .

- Finite dimensional ensemble of 2SAT problems.
- No SAT/UNSAT transition. Quantitative understanding.
- Percolation of logical structure *after* bond percolation.
- Rare regions can dominate running time in sub-percolative phase. (Dynamic transition.)
- Unique ground state in limit of large systems. [Heuristic algorithm.]

Paleodictyon Nodosum