Thresholds for 3-SAT.

Josep Díaz

joint work with L. Kirousis, D. Mitsche and X. Pérez

Threshold for the 3-satisfiability problem (3SAT)

Given n Boolean variables $X = \{x_1, x_2, \ldots, x_n\}$ a Boolean formula ϕ is a conjunction of clauses each of which is a disjuction of literals (variables or their negation).

$\phi = (x_1 \vee \bar{x_2} \vee x_4) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (x_1 \vee \bar{x_3} \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3})$.

A formula ϕ is satisfiable if there exists a *truth assignment* A to the variables so that each clause in ϕ contains at least one "true" literal. $A \models \phi$

$A = (1, 0, 0, 1) \Rightarrow A \models \phi$

The *3-Satisfiability Problem (3SAT)*: given a formula $\phi = C_1 \wedge \cdots \wedge C_m$, where each C_i contais 3 literals, is it satisfiable?

The ratio $r = \frac{m}{n}$ is the density.

Phase transition for 3SAT

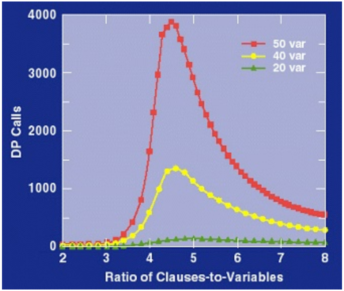
there is a constant r_c such that

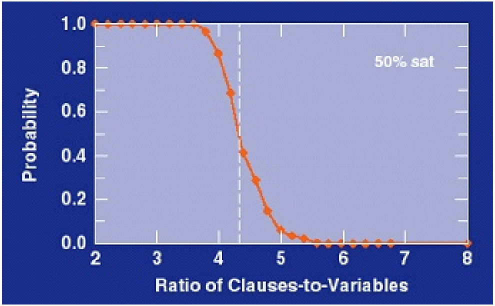
- if r is *away* from r_c, then whp the number of calls to Davis-Putnam is small, while if r is *close* to r_c, the number of calls is large.
- if $r < r_c$, then whp the formula is satisfiable, while if $r > r_c$, whp the formula is unsatisfiable

It has been rigorously settled that for 2-SAT: $r_c = 1$.
Goerdt (1992), Chvátal-Reed (1992),....

Phase transition for 3-SAT

Experimentally:
Chesman, Kanefsky, Taylor (1991) for k-SAT
Mitchell, Selman, Levesque (1991) for 3-SAT





Phase transition 3SAT: The physics approach

Using techniques from statistical physics: *Replica Symmetry Breaking, Cavity method* on very large instances of 3SAT, physics people where able to give **theoretical non-rigorous evidence** that the threshold for 3SAT occurs at

$$r_c = 4.27$$

Mézard, Parisi, Zecchina (2002), Mézard, Zecchina (2002), ...

Phase transition 3SAT: The rigorous approach

Theorem (Friedgut (1997))

There is a sequence $r_c(n)$ such that $\forall \epsilon$:
$\mathbf{Pr}\left[\phi_{r_c(n)-\epsilon} \text{ is SAT}\right] \rightarrow 1$ *and* $\mathbf{Pr}\left[\phi_{r_c(n)+\epsilon} \text{ is SAT}\right] \rightarrow 0$.

Friedgut's theorem says that the transition interval can be made arbitrarily thin. But he doesn't give threshold point (the convergence of $\{r_c(n)\}$).

Question: Does $r_c(n)$ converge? If yes, to what value?

Rigorous approach

Consider a random 3SAT formula ϕ, with $m = rn$ clauses.

Upper bound: $r > r_c = 4.27$ Get a value as low as possible of r (≥ 4.27) such that whp ϕ is not SAT.

Lower bound: $r < r_c = 4.27$ Consider an easy to analyze algorithm. Get a value as large as possible of r ($r \leq 4.27$) such that whp the algorithm produces satisfying assignment for ϕ.

Random Formula

Given n variables, the set of possible clauses has size $2^3\binom{n}{3}$.
We have 4 ways to select a random ϕ:

1. $G_{n,p}$: Each clause is independently selected with probability p to be included in the formula. Notice in this case the number of clauses is a random variable. Therefore to have a $m = rn$ we need a value of $p = \frac{3r}{4n(n-1)} \sim \frac{3r}{4n^2}$.
2. $G_{n,m}$: Exactly $m = rn$ clauses are uniformly, independently and with replacement selected to be included in the formula. Notice in this model, there could be repeated clauses.
3. $G^*_{n,m}$: Exactly $m = rn$ clauses are uniformly, independently and without replacement selected to be included in the formula. Notice in this model, every clause is different.
4. $\mathcal{C}_{n,D}$ the configuration model.

Configuration model

A degree sequence $D = \{d_{ij}\}$ for variables $\{x_1, \dots, x_n\}$, where each d_{ij} tell us how many variables must appear i-times not negated and j-times negated in ϕ.
Given a set of n and a D, a formula ϕ is generated according to $\mathcal{C}_{n,D}$ if the appearance of the n variables in ϕ follows D.

Given $n = 4$ and D: $d_{12} = 2, d_{22} = 1, d_{14} = 1$ and remaining $d_{ij} = 0$, then a possible ϕ is
$(x_1 \vee \bar{x_2} \vee x_3) \wedge (\bar{x_2} \vee x_3 \vee \bar{x_4}) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3}) \wedge (\bar{x_2} \vee x_4 \vee \bar{x_3})$.

For instance only x_1 and x_4 appear 1 time afirmative and 2 times negated.

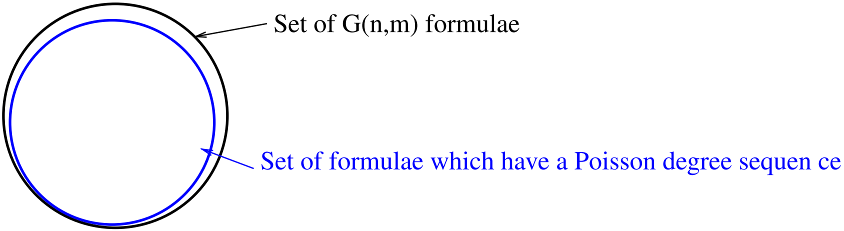
Configuration model

In the setting of SAT, the degree sequence follows a Poisson distribution, where ϕ is given by

$$d_{ij} = \frac{e^{-\mu}(\mu/2)^{i+j}}{i!j!},$$

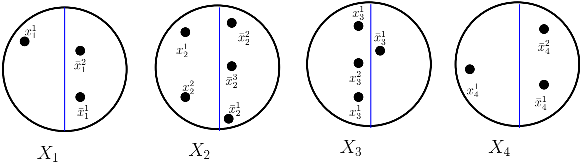
with $\mu = 3r$.
Then, $m = 3\sum_{i,j}(i+j)d_{ij}$.

Dubois, Boufkhad, Mandler (2000), called *typical formula*, the formula with Poisson degree sequence.
They prove that most of the formulae $G_{n,m}$ are typical:

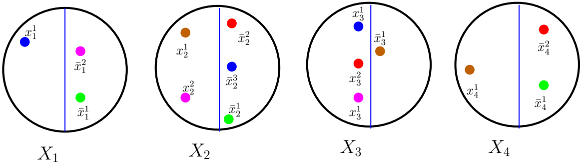


Example:

Given $\{d_{ij}\}$: $d_{12} = 2, d_{23} = 1, d_{31} = 1$ and $X = \{x_1, x_2, x_3, x_4\}$ to form a possible 3SAT formula ϕ:



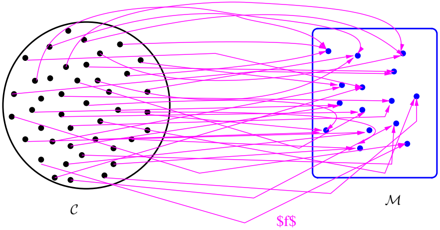
3



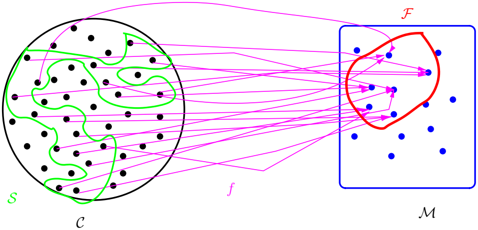
3

$$\phi = (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x_2} \vee x_3) \wedge (x_2 \vee \bar{x_3} \vee x_4).$$

Let $\mathcal{C}$ be the set of configurations on set X of variables and degree sequence D.
Let $\mathcal{M}$ be the set of multiformulae on set X of variables with m



Consider $\mathcal{F} \subset \mathcal{M}$ the set of $G_{n,m}$ formulas.
Let $\mathcal{H} \subset \mathcal{C}$ the set of anti-images of $\mathcal{F}$ ($\mathcal{H} = f^{-1}(\mathcal{F})$)



A property which happens aas in $\mathcal{C}$ also happens aas in $\mathcal{H}$
Which can de transfered to $\mathcal{F}$: For given assignment A, the probability that a ϕ is SAT in $\mathcal{H}$ is the same that ϕ is SAT in $\mathcal{F}$.
So probability that a ϕ is SAT in $\mathcal{C}$ is the same that ϕ is SAT in $\mathcal{F}$.

Relation between models

Let E ad F be events. It is well known:

$$\mathbf{Pr}_m[E] \asymp \mathbf{Pr}_{m*}[E] \leq \mathbf{Pr}_p[E]$$

$$\mathbf{Pr}_m[E] \to 1 \Leftrightarrow \mathbf{Pr}_{m*}[E] \to 1 \Leftrightarrow \mathbf{Pr}_p[E] \to 1 \Leftrightarrow \mathbf{Pr}_C[E] \to 1$$

Status Upper bound to 3-SAT:

$r = 5.1909$ (1983) Franco, Paull (and others)
$r = 5.19 - 10^{-7}$ (1992) Frieze and Suen
$r = 4.758$ (1994) Kamath, Motwani, Palem, Spirakis
$r = 4.667$ (1996) Kirousis, Kranakis, Krizanc.
$r = 4.642$ (1996) Dubois, Boufkhad
$r = 4.602$ (1998) Kirousis, Kranakis, Krizac, Stamatiou
$r = 4.596$ (1999) Janson, Stamatiou, Vamvakari (1999)
$r = 4.571$ (2007) Kaporis, Kirousis, Stamatiou, Vamvakari
$r = 4.506$ (1999) Dubois, Boukhand, Mandler
$r = 4.49$(2008) Díaz, Kirousis, Mitsche, Pérez
$r_c = 4.27$ Experimental threshold (Replica Method)

First moment: Basic technique for upper bounds

Let ϕ be a random formula and $S(\phi)$ the set of its satisfying truth assignments. Using Markov inequality

$\mathbf{Pr}_{m*}[\phi \text{ is sat}] = \mathbf{Pr}_{m*}[|S(\phi)| \geq 1] \leq \mathbf{E}[|S(\phi)|]$.

Must compute $\mathbf{E}[|S(\phi)|]$

Notice that given a truth assignment A and 3 variables x_i, x_j, x_k then *there is only one clause on x_i, x_j, x_k which is not SAT by A*. Therefore, in the $G^*_{n,m}$ model, out of the $8\binom{n}{3}$ clauses only $\binom{n}{3}$ evaluate to 0 under any given A.

$\mathbf{E}[|S(\phi)|] = \sum_{A \in S(A)} \mathbf{Pr}[A \models \phi] = \frac{|\{<A,\phi> \mid A \models \phi\}|}{|\{\phi\}|}$

$\mathbf{E}[|S(\phi)|] = (2(7/8)^r)^n$ to make it < 1 we need

$$r \geq 5.1909$$

5.2 is far above the experimental 4.27, because there could be a few formulas with many sat. truth assignment which contribute too much to $\mathbf{E}[|S(\phi)|]$.

Single Flips

We wish to find r s.t. $\mathbf{Pr}_{m*}[\phi \text{ is SAT}] \to 0$ and $r < 5.2$

Kirousis, Kranakis, Krizanc (1996), Dubois, Boufkhad (1996)
Instead of using $S(\phi)$, we restrict to the class $S^1(\phi)$:

Let $S^1(\phi)$ be the set of assignments $\{A \mid A \in S(\phi)\}$ such that if we modify A to A' by changing a single 0 assignment to 1 then $A' \not\models \phi$
If $A \models \phi$ in the *single flip* sense, we denote $A \models^{sf} \phi$.

Single Flips

$\phi = (x_1 \vee \bar{x_2} \vee x_4) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (x_1 \vee \bar{x}_3 \vee \bar{x}_4) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_3)$:
$S(\phi) = \{(1,1,1,0),(1,1,0,0),(1,0,0,1),(1,0,0,0),(0,1,0,1),$
$(0,0,1,0),(0,0,0,1),(0,0,0,0)\}$.
Take, $A = (1,0,0,1)$. Flipping the second 0 yields $A' = (1,1,0,1)$ and $A' \not\models \phi$, due to $(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_4)$. Such a clause is called a *blocking clause* for the flip.

So a blocking clause for an assignment A is a clause which contains a negated variable, which if we change only the value of that variable from 0 to 1, the clause if not satisfied.

$S^1(\phi) = \{(0010),(0101),(1000),(1001),(1110)\}$

If ϕ is satisfiable then $S^1(\phi) \neq \emptyset$.
Moreover, $S^1(\phi) \subseteq S(\phi)$ and, thus, $|S^1(\phi)| \leq |S(\phi)|$. In fact, $|S^1(\phi)| << |S(\phi)|$ and convergence to 0 is faster and therefore we get a smaller r.

$\mathbf{E}\left[|S^1(\phi)|\right] \leq \left[\left(\frac{7}{8}\right)^r (1 - e^{-\frac{3}{7}r})\right]^n$
Therefore, $r = 4.667$
Kirousis, Kranakis, Krizanc (1996)

Double flips

Kirousis, Kranakis, Krizanc, Stamatiou (1998)
Given a random ϕ and a $A \models^{sf} \phi$, we say A *satisfies ϕ in the double flip sense* $A \models^{df} \phi$ if for variables x_i, x_j with $i < j$ and s.t. $A(x_i) = 0$ and $A(x_j) = 1$, when we modify A to A' by changing only $A'(x_i) = 1$ and $A'(x_j) = 0$ then $A' \not\models \phi$.
Let $S^2(\phi) = |\{A| \models^{df} \phi\}$.

$\phi = (x_1 \vee \bar{x_2} \vee x_4) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (x_1 \vee \bar{x_3} \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3})$.
$S(\phi) =$
$\{(0000), (0001), (0010), (0101), (1000), (1001), (1100), (1110)\}$
$S^1(\phi) = \{(0010), (0101), (1000), (1001), (1110)\}$
$S^2(\phi) = \{(0010), (1001)\}$

Double flips

Kaporis, Kirousis, Stamatiou, Vamkari, Zito (07)
For ϕ in $G^*_{n,m}$,

$$\mathbf{Pr}\left[\phi \text{ is SAT}\right] \leq \mathbf{E}\left[S^2(\phi)\right] =$$

$$= \sum_A \mathbf{Pr}\left[A \vDash \phi\right] \mathbf{Pr}\left[A \in S^1(\phi) | A \vDash \phi\right] \mathbf{Pr}\left[A \in S^2(\phi) | A \in S^1(\phi)\right]$$

They obtained $r = 4.571$, by more accurate computations.

Balanced literals

Dubois, Boufkhad, Mandler (2000)
For each variable x_i in a given a random ϕ:

- if number occurrences of $x_i \geq$ number of occurrences of $\bar{x}_i$, leave all appearances of x_i as they are.
- if number occurrences of $x_i <$ number of occurrences of $\bar{x}_i$, swap all appearances of x_i and $\bar{x}_i$ in ϕ.
  $\phi = (x_1 \vee \bar{x_2} \vee x_4) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (x_1 \vee \bar{x_3} \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3})$.
  $S(\phi) =$
  $\{(0000), (0001), (0010), (0101), (1000), (1001), (1100), (1110)\}$
  after balancing it:
  $\phi' = (x_1 \vee x_2 \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee x_4) \wedge (x_1 \vee x_3 \vee x_4) \wedge (\bar{x_1} \vee \bar{x_2} \vee x_3)$.
  $S(\phi') =$
  $\{(0010), (0101), (0110), (0111), (1001), (1011), (1110), (1111)\}$

Notice $|S(\phi)| = |S(\phi')| \Rightarrow \mathbf{E}\left[|S(\phi)|\right] = \mathbf{E}\left[|S(\phi')|\right]$

Single flips + balancing

However, if $\phi = (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4})$. Then

$S^1(\phi) = \{(011), (101), (110)\}$ but

$S^1(\phi') = \{(111)\}$

So $\mathbf{E}\left[|S^1(\phi')|\right] << \mathbf{E}\left[|S^1(\phi)|\right]$ (exponentially small)

Dubois, Boufkhad, Mandler (2000) starting from a random ϕ in $G_{n,m}$, and modifying ϕ and $S(A)$ according to:

Formula typicallity + balancing + single flips:

$r = 4.506$

Pure literal elimination

Given a random ϕ a literal is said to be pure if its completement does not appear in ϕ.

Pure literal rule: As long as there is a pure literal in ϕ assign value 1 and remove all clauses where it appears.

Broder, Frieze, Upfal (1996), proved that whp, the pure literal rule finds SAT assignments for ϕ in $G_{n,m}$ up to $r = 1.63$, but no further.

Clause typicality

Given any 3SAT formula there are 4 types of clauses

- Type 0: $(\bar{x}, \bar{y}, \bar{z})$
- Type 1: $(x, \bar{y}, \bar{z})$
- Type 2: $(x, y, \bar{z})$
- Type 3: (x, y, z)

A ϕ with n variables and m clauses is said to be *clause typical* if
\# clauses type 0 = \# clauses type 3= $\frac{m}{8}$ and
\# clauses type 1 = \# clauses type 2= $\frac{3m}{8}$

Díaz, Kirousis, Mitsche, Pérez (2008)

1. Use random ϕ with typical degree sequence (Poisson)
2. Define a process over time for the elimination pure literals. This yield a system of ODE, which can be solved by the differential equation method. The result will be a set of formulae which will have *almost* typical degree sequence.
3. Use positive balancing.
4. Use clause typicality to thin the space of formulae resulting from the previous step to formulae with typical clause.
5. Apply single-flips to the obtained space.

$r = 4.4898$

A new approach to upper-bound

Key fact emerging from *Replica method*: For ϕ with $r > 3.92$ $S(\phi)$ is split into clusters. Within the same cluster we can change from one assignment to the other by flipping a single variable.
To move between assignments in two different clusters, we need to flip several variables at the same time.

Give a ϕ consider the set $V(\phi)$ of partial valid assignments where each variable can be assigned a value $\{0, 1, *\}$, and such that each clause gets no $(0, 0, 0)$ or $(0, 0, *)$.
Notice $S(\phi) \subseteq V(\phi)$.
Given any partial or total assignment of ϕ, a literal is constrained in C_i it has value 1 and the remaining literals in the clause have value 0.
If $C_i = (x, \bar{y}, z)$ and $(0, 1, 0) \Rightarrow \bar{y}$ is constrained.
If $(0, 1, *)$ then $\bar{y}$ is not constrained.

Lattice structure for $V(\phi)$

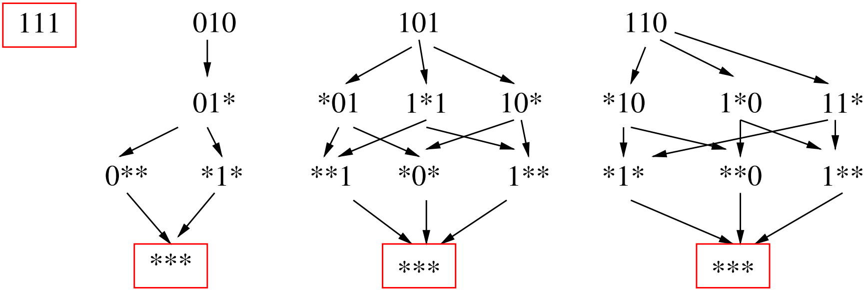
Braunstein, Zecchina (2004); Maneva, Mossel, Wainwright (2005).

Partial valid assignments A_1^* and A_2^*, $A_1^* \rightarrow A_2^*$ if A_2^* has one $*$ more than A_1^*
$(01001010 * 0 * 1) \rightarrow (010 * 1010 * 0 * 1)$
This create a set of lattices of $V(\phi)$ (a lattice for cluster)
The lattices have layers: layer 0 contains $S(A)$.
Layer i contains the p. assignments with i $*$
The minimal elements of each lattice is *unique* and is called the core.

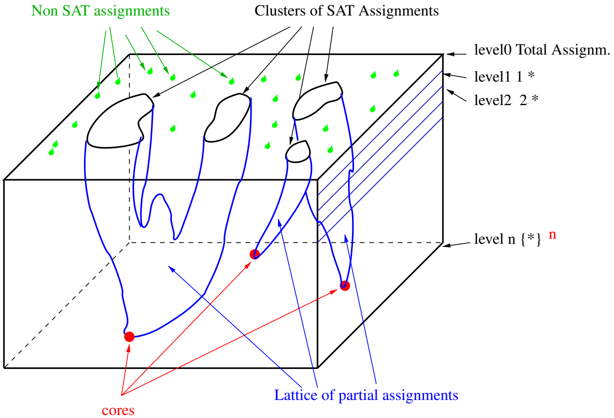
From any $A \in S(\phi)$, choose an unconstrained variable in ϕ and substitute its value by $*$, and continue until it is not possible.

Example of lattices with the core

Let $\Phi = (x_1 \vee \bar{x_2} \vee \bar{x_3}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3}) \wedge (\bar{x_1} \vee x_3 \vee \bar{x_3})$.
The lattices of partial assignments for assignments:
$A_1 = (1, 1, 1)$; $A_2 = (0, 1, 0)$; $A_3 = (1, 1, 0)$.
In boxed red, the core for each lattice.



The world of partial assignments



The world of partial assignments for a 3-SAT formu la (Maneva et all)

Non-trivial core

A core is non-trivial if the core is different from $*^n$.

Achlioptas, Ricci-Tersenghi (2006) have proved that for a random k-SAT ϕ ($k \geq 8$), and densities r near the threshold, whp every SAT assignment has a non-trivial core.
This result is open for $k < 8$.

The explanation of the success of message passing techniques (SP,BP) to find the threshold density of SAT problems, is based on the existence of those non-trivial cores near r_c.

New result and new technique.

Maneva, Sinclair (2008).
For 3-SAT one of the two statements holds for random 3-SAT:

- $r \leq 4.453$ or
- there is a range of densities immediately below the 3-SAT threshold, for which whp there are no non-trivial cores.

Instead of bounding the probability of a random ϕ to be SAT by the expectance of a thinned subspace of $S(\phi)$, they bounded by the expected number of non-trivial cores.
They use a weighted version of the first moment method.
It opens a new and interesting line of research.

Status of the lower bounds to 3SAT threshold

$r_c = 4.27$ Experimental threshold (Replica Method)
$r > 3.52$ Kaporis, Kirousis, Lalas (2003)
$r > 3.52$ Hajiaghayi-Sorkin (2003)
$r > 3.42$ Kaporis, Kirousis, Lalas (2002).
$r > 3.26$ Achlioptas and Sorkin (2001).
$r > 3.145$, Achlioptas (2000).
$r > 3.003$, Frieze, Suen (1992).
$r > 2.99$ Chao, Franco (1986).
$r > 2.66$ Chao, Franco (1986).

General methods for lower bounds to 3SAT threshold

Given a random ϕ in $G_{n,m}$, $m = rn$ consider an *easy to analyze* heuristic, to find a $A \vDash \phi$,

Let r_l denote the lower bound for the density that we try to compute. Prove that for $r < r_l$, the heuristic succeeds whp.

The heuristic *succeeds* if no empty clause is ever generated, (x and $\bar{x}$ are not at the same time in the same clause).

Let $C_i(t)$ be the number of clauses with i literals, at $i = 1, 2, 3$,
At step $t + 1$ and empty clause can be generated only if $\Delta(C_1(t))/\Delta(t) > 1$.
At every step, the algorithm should strive to keep the expected number of new unit clauses less than 1.

The Differential Equation Method (DEM)

T. Kurtz (1970); Karp-Sipser (1981); Wormald (1995).
Given a sequence of random processes, we wish to find properties in the limit:

1. Compute the expected changes in random variables per unit of time,
2. regard the variables as continuous,
3. writte down the ODE suggested by the expected changes
4. use large deviations theorems (Wormald) to show that a.s. the solution to the ODE is close to the values of the variables

The Unit Clause algorithm

Chao, Franco (1986)

UC ϕ
if there is a 1- clause **then**
  select u.a.r. one 1-clause and satisfy it (forced step)
else select u.a.r a x_i and assign u.a.r. T or F (free step)

Analyzing the UC algorithm

The expected number of 1-clauses generated at t is $\frac{C_2(t)}{n-t}$

If $\exists t$ s.t. $\frac{C_2(t)}{n-t} > (1+\epsilon)$, a.s. UC will fail

If $\forall t$ $\frac{C_2(t)}{n-t} < (1-\epsilon)$ UC will succeed with positive probability.

We have to find a value r_l s.t. $\forall t$ $\frac{C_2(t)}{n-t} < (1-\epsilon)$

Analyzing the UC algorithm

Let $\Delta C_i(t) = C_i(t+1) - C_i(t)$, scaling down $x = t/n$
$\mathbf{E}[\Delta C_3(t)] = -\frac{3C_3(t)}{n-t} \Rightarrow c_3'(x) = -\frac{3c_3(x)}{1-x}$
$C_3(0) = rn \Rightarrow c_3(0) = r$
$\mathbf{E}[\Delta C_2(t)] = \frac{3C_3(t)}{2(n-t)} - \frac{2C_2(t)}{(n-t)} \Rightarrow c_2'(x) = \frac{3c_2(x)}{2(1-x)} - \frac{2c_2(x)}{1-x}$
$C_2(0) = 0 \Rightarrow c_2(0) = 0$
Solving and using Wormald's theorem we get $r_l = 8/3 = 2.6$