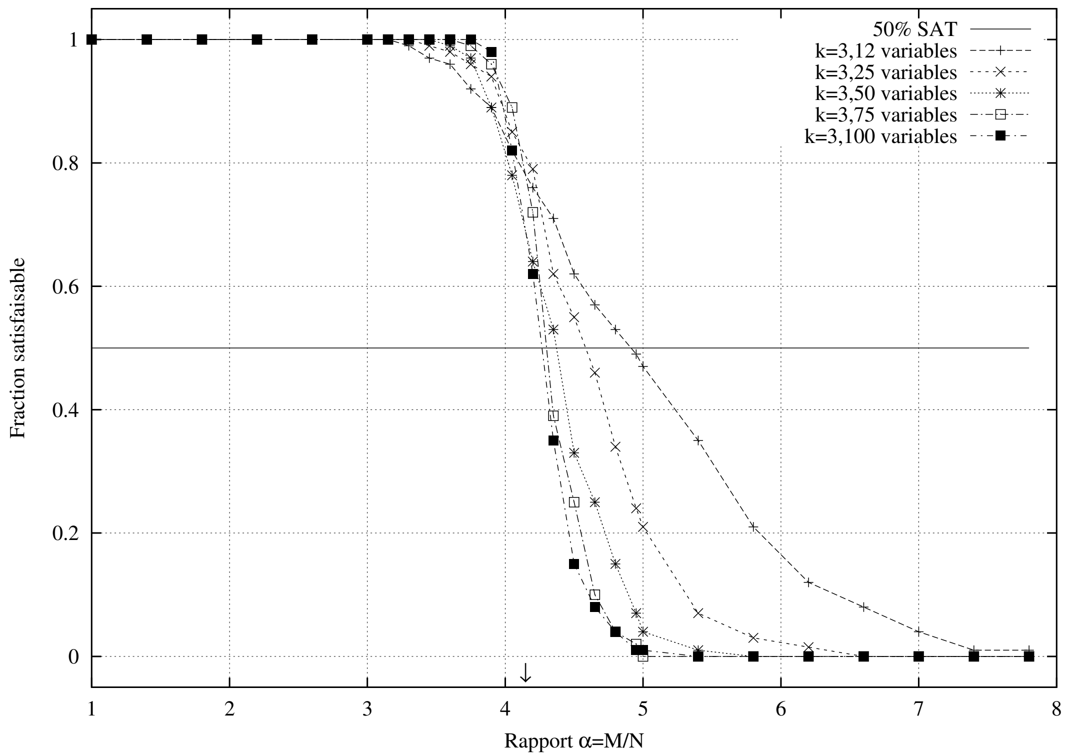
Combinatoire analytique appliquée à la satisfiabilité des formules 3-SAT.

Vincent Puyhaubert

20 Janvier 2003



Modèle choisi

$x_1, \ldots, x_n$ des variables booléennes.

$$\Phi = C_1 \wedge C_2 \wedge \cdots \wedge C_m$$

avec $C_i = a \vee b \vee c$ et $a, b, c \in \{x_i, \neg x_i\}$

Eventuellement plusieurs fois la meme clause, mais des littéraux de variables distinctes dans une meme clause (pas de $x_1 \vee \neg x_1 \vee x_4$ ou $x_1 \vee x_1 \vee x_2$).

Nombre de clauses : $8\binom{n}{3}$ Nombre de formules : $\left(8\binom{n}{3}\right)^m$

Conjecture : Soit $r = m/n$. Il existe r_3 tel que si $m, n \to +\infty$ avec r constant, alors :

- $\Pr(\Phi$ satisfiable$) \longrightarrow 1$ si $r < r_3$
- $\Pr(\Phi$ satisfiable$) \longrightarrow 0$ si $r > r_3$

Valeur expérimentale ≈ 4.25

Bornes inférieures	
2.9^*	Chao and Franco (1986,1990)
$2/3^*$	Chvátal and Reed (1992)
1.63	Broder et al. (1993)
3.003^*	Frieze and Suen (1996)
3.145	Achlioptas (2000)

Bornes supérieures	
5.191^*	Franco and Paull (1983)
5.081	El Mafthoui and Fernandez de la Vega (1993)
4.762^*	Kamath et al. (1995)
4.643^*	Dubois and Boufkhad (1997)
4.602	Kirousis et al. (1998)
4.596	Janson et al. (1999)
4.506	Dubois et al.

Esperance du nombre de solutions

Exemple $(n = 4, m = 8)$:

$$\begin{aligned}\Phi = \quad & (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_3) \\ & \wedge(\neg x_1 \vee x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3 \vee \neg x_4) \\ & \wedge(\neg x_2 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_4)\end{aligned}$$

Ensemble des solutions :

$x_1 x_2 x_3 x_4$

1 1 1 1

0 1 1 1   1 0 1 1   1 1 0 1   1 1 1 0

0 0 1 1   0 1 0 1   0 1 1 0   1 0 0 1   1 0 1 0   1 1 0 0

0 0 0 1   0 0 1 0   0 1 0 0   1 0 0 0

0 0 0 0

Lemme : *Pour tout entier k non nul, on a $1 \leq k$.*

Corollaire :

$$| \, \Phi \text{ satisfiable} \, | \leq | \, (\Phi, A) \text{ tels que } \Phi \text{ soit satisfaite par } A \, |$$

$$\boxed{\Pr(\Phi \text{ satisfiable}) \leq \frac{| \, (\Phi, A) \text{ tels que } \Phi \text{ soit satisfaite par } A \, |}{|\Phi|}}$$

Remarque : X v.a. à valeurs dans $\mathbb{N}$, alors :

$$\Pr(X > 0) \leq E(X)$$

A une affectation des n variables, $C = \pm x_i \vee \pm x_j \vee \pm x_k$ une clause. Une fois les 3 variables choisies, il y a :

– 1 seule façon de contredire A en choisissant les signes des variables.

– 7 façons de la satisfaire.

- Nombre de clauses satisfaisant A : $7\binom{n}{3}$
- Nombre de formules satisfaites par A : $\left(7\binom{n}{3}\right)^m$
- $\mid (\Phi, A)$ tels que A soit solution de $\Phi \mid = 2^n \left(7\binom{n}{3}\right)^m$

$$\boxed{\Pr(\Phi \text{ satisfiable}) \leq \left(2\left(\frac{7}{8}\right)^r\right)^n}$$

Pour $r > \ln(2)/\ln(8/7) \approx 5.19$, elle tend vers 0.

Ensemble des solutions :

$x_1 x_2 x_3 x_4$

1 1 1 1

0 1 1 1   1 0 1 1   1 1 0 1   1 1 1 0

0 0 1 1   0 1 0 1   0 1 1 0   1 0 0 1   1 0 1 0   1 1 0 0

0 0 0 1   0 0 1 0   0 1 0 0   1 0 0 0

0 0 0 0

Affectations partielles :

- A 4 variables : toutes les solutions
- A 3 variables : $-\,1\,1\,1 \quad 0\,-\,1\,1 \quad -\,0\,0\,1$

  $0\,0\,-\,1 \quad 0\,0\,1\,- \quad 0\,0\,-\,0$

  $0\,0\,0\,-$
- A 2 variables : $0\,0\,-\,-$

Impliquants premiers

$$\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ 0 & - & 1 & 1 \end{array}$$ est un impliquant premier car :

- la formule est déjà satisfaite sans même donner de valeur aux autres variables
- cette affectation est minimale au sens où un minimum de variables sont affectées (si l'on en retire une, la formule n'est plus satisfaite).

$$\begin{aligned} \Phi = \; & (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_3) \\ & \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3 \vee \neg x_4) \\ & \wedge (\neg x_2 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_4) \end{aligned}$$

Toute formule satisfiable a au moins un impliquant premier et donc :

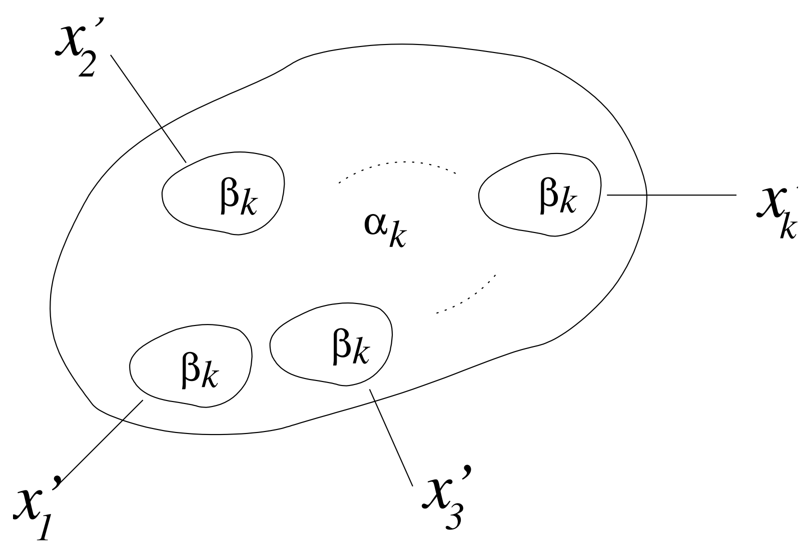
$$|\ \Phi \text{ satisfiable}\ | \leq |\ (\Phi, I) \text{ tels que } I \text{ est un impliquant premier pour } \Phi\ |$$

I une affectation partielle de k variables $x'_1, \ldots, x'_k$ parmi les n.

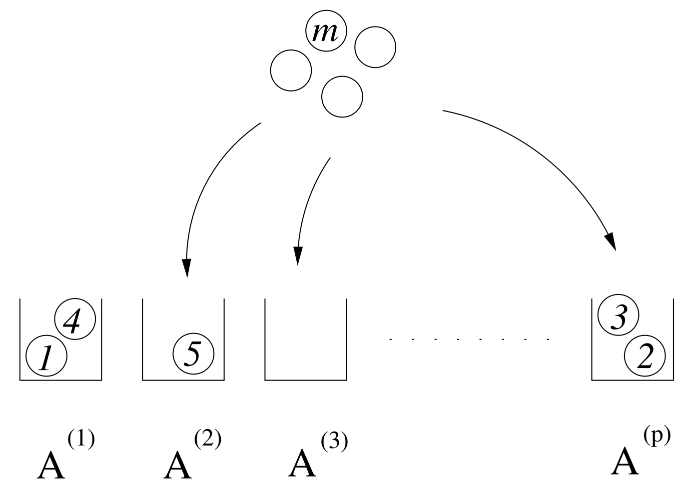
$$\Phi = C_1 \wedge \ldots \wedge C_m$$

- $\forall i \in [\![1 ; m]\!]$ C_i contient au moins un littéral rendu positif par I.
- $\forall j \in [\![1 ; k]\!] \quad \exists i \quad C_i = (\pm x'_j \vee a \vee b) \quad a, b$ fausses ou non affectées

Lemme important : Les k ensembles ont le même cardinal β_k et sont disjoints deux à deux.



$A^{(1)}, \ldots, A^{(p)}$ des sous-ensemble de $\mathbb{N}$.



F_m : nombre de façons de jeter m boules de sorte que dans l'urne i, il y ait au final un nombre de boules qui appartienne à $A^{(i)}$.

$$F_m = \sum_{n_1+\cdots+n_l=m} \binom{m}{n_1;\cdots;n_p} \left[\left[n_1 \in A^{(1)} \right]\right] \cdots \left[\left[n_p \in A^{(p)} \right]\right]$$

$$\boxed{F(z) = \sum_{m\geq 0} F_m \frac{z^m}{m!} = \prod_{i=1}^{p} \left(\sum_{j\in A^{(j)}} \frac{z^j}{j!} \right)}$$

Dans le cas qui nous intéresse :

$$F(z) = (e^z)^{\alpha_k} \left((e^z)^{\beta_k} - 1\right)^k$$

avec $\alpha_k \leq \frac{1}{3}k^2(3n-k)$ et $\beta_k \leq \frac{1}{2}(2n-k)^2$.

Lemme : *Si $F(z) = \sum F_m \frac{z^m}{m!}$ et que F_m est un réel positif pour tout m, alors, pour tout réel s strictement positif :*

$$F_m \leq m! \frac{F(s)}{s^m}$$

Corollaire : *Soit I_k nombre de formules pour lesquelles I est un impliquant premier. On a la majoration valable pour tout réel s positif par :*

$$I_k \leq \frac{m!}{s^m} e^{\frac{s}{3}k^2(3n-k)} \left(e^{\frac{s}{2}(2n-k)^2} - 1\right)^k$$

$$|\Phi \text{ satisfiable}| \leq \sum_{k=0}^{n} \binom{n}{k} 2^k \frac{m!}{(s_k)^m} e^{\frac{s_k}{3} k^2 (3n-k)} \left(e^{\frac{s_k}{2}(2n-k)^2} - 1\right)^k$$

On pose $k = \alpha n$, $u_\alpha = s_{\alpha n} n^2$ et $m = rn$. Alors :

$$\boxed{\Pr(\Phi \text{ satisfiable}) \leq \sum_{\alpha \in \left\{0, \frac{1}{n}, \ldots, 1\right\}} f(\alpha)^n}$$

avec $$f(\alpha) = \frac{2^\alpha}{\alpha^\alpha (1-\alpha)^{1-\alpha}} \frac{e^{\frac{u_\alpha}{3}\alpha^2(3-\alpha)} \left(e^{\frac{u_\alpha}{2}(2-\alpha)^2} - 1\right)^\alpha}{{u_\alpha}^r}$$

et $$u_\alpha \left(\frac{\alpha^2(3-\alpha)}{3} + \frac{(2-\alpha)^2}{2} \frac{e^{\frac{u_\alpha}{2}(2-\alpha)^2}}{e^{\frac{u_\alpha}{2}(2-\alpha)^2} - 1} \right) = r$$

Pour $r = 4.89$, le maximum de f sur $[0, 1]$ est strictement inférieur à 1.

Solutions maximales

$\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ 1 & 0 & 0 & 1 \end{array}$ est une solution maximale pour Φ car si l'on change n'importe lequel de ses 0 en 1, la nouvelle affectation obtenue n'est plus une solution.

$$\begin{aligned}\Phi = \quad & (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_3) \\ & \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3 \vee \neg x_4) \\ & \wedge (\neg x_2 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_4)\end{aligned}$$

Toute formule satisfiable a au moins une solution maximale et donc :

$$\mid \Phi \text{ satisfiable} \mid \leq \mid (\Phi, A) \text{ tels que } A \text{ solution maximale } \Phi \mid$$

A une affectation des n variables $x_1, \ldots, x_n$ dont k variables sont affectées à la valeur 0 (par exemple les k premières).

$$\Phi = C_1 \wedge \ldots \wedge C_m$$

– $\forall i \in [\![1\,;\, m]\!]$ C_i est l'une des $7\binom{n}{3}$ clauses satisfaite par A

– $\forall i \in [\![1\,;\, k]\!] \quad \exists j \quad C_j = (\neg x_i \vee a \vee b)$ avec a, b rendues fausses par A.

Lemme bis : Les k ensembles ont à nouveau le même cardinal $\binom{n-1}{2}$ et sont à nouveau disjoints deux à deux.

$$F(z) = \sum F_m \frac{z^m}{m!} = (e^z)^{7\binom{n}{3} - k\binom{n-1}{2}} \left((e^z)^{\binom{n-1}{2}} - 1 \right)^k$$

Notation : Si $F(z) = \sum F_m z^m$, on note $F_m = [z^m]F(z)$.

$$|\,(\Phi, A \text{ max})| = \sum_{k=0}^{n} \binom{n}{k}\, m!\, [z^m] e^{z\left[7\binom{n}{3} - k\binom{n-1}{2}\right]} \left(e^{z\binom{n-1}{2}} - 1 \right)^k$$

Lemme : $[z^m]f(z) + g(z) = [z^m]f(z) + [z^m]g(z)$

$$\boxed{\ |\,(\Phi, A \text{ max})| = m!\ [z^m]e^{z\ 7\binom{n}{3}} \left(2 - e^{z\binom{n-1}{2}}\right)^n}$$

En utilisant le lemme précédent, et en posant $u = z\binom{n-1}{2}$, $\forall\delta > 0$:

$$\boxed{\Pr(\Phi \text{ satisfiable}) \leq \left(\left(\frac{3\,r}{8\,e}\right)^r \frac{e^{7\delta/3}\left(2 - e^{-\delta}\right)}{\delta^r}\right)^n}$$

Cette expression est minimale pour δ vérifiant :

$$\delta\left(\frac{7}{3} + \frac{e^{-\delta}}{2-e^{-\delta}}\right) = r$$

La valeur $r = 4.643$ est la plus petite constante telle que l'expression sous la puissance soit strictement plus petite que 1 avec un tel δ.

Formules typiques

Idée : Jeter une fraction négligeable de formules ayant beaucoup de solutions.

Nouveau modèle : On pioche $3m$ variables parmi les n.

$\omega_{l,p}$: fraction de variables ayant l occurences dont p positives.

$$\kappa_{l,p} = \frac{1}{2^l}\binom{l}{p}\frac{(3r)^l}{l!}e^{-3r} \qquad \left(r = \frac{m}{n}\right)$$

$$\forall l, p \quad \forall \epsilon > 0 \qquad \Pr\left(|\omega_{l,p} - \kappa_{l,p}| > \epsilon\right) \underset{n\to\infty}{\longrightarrow} 0$$

Soit $x_{\max}$ un entier et $\epsilon > 0$. Une formule Φ est dite typique ssi :

$$\forall \leq p \leq l \leq x_{\max} \qquad |\omega_{l,p}(\Phi) - \kappa_{l,p}| \leq \epsilon$$

Alors, en vertu de ce qui précède, une fois fixés $x_{\max}$ et ϵ, les formules non typiques forment un ensemble négligeable de formules quand n tend vers l'infini.

Corollaire :

$$\Pr(\Phi \text{ satisfiable}) \leq \frac{|\,(\Phi \text{ typique}, A \text{ max})|}{|\Phi|} + \circ(1)$$

$$\begin{aligned}\Phi = \quad & (x_1 \vee \neg x_2 \vee \neg x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_3) \\ & \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_2 \vee x_4) \wedge (x_2 \vee x_3 \vee \neg x_4) \\ & \wedge (x_2 \vee x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_4)\end{aligned}$$

<u>Ensemble des solutions :</u>

$x_1 x_2 x_3 x_4$

1 1 1 1

0 1 1 1   1 0 1 1   1 1 0 1   1 1 1 0

0 0 1 1   0 1 0 1   0 1 1 0   1 0 0 1   1 0 1 0   1 1 0 0

0 0 0 1   0 0 1 0   0 1 0 0   1 0 0 0

0 0 0 0