Lieu du seuil d'(1,2)-QSAT

N. Creignou[1] H. Daudé[2] U. Egly[3] R. Rossignol[4]

[1]LIF, Marseille
[2]LATP, Marseille
[3]TU Wien
[4]Orsay

1er septembre 2010, journées MAS

Plan

Plan

(1,2)-QSAT

Une formule (1,2)-QCNF est une formule quantifiée de la forme

$$\forall X \exists Y \, \varphi(X, Y)$$

- X et Y désignent des ensembles disjoints de variables
- $\varphi(X, Y)$ est une formule 3-CNF telle que chaque clause contienne exactement 1 variable de X et exactement 2 variables de Y
- (1,2)-QSAT est la propriété pour une formule (1,2)-QCNF d'être vraie
- (1,2)-QSAT est dans coNP

Un exemple de formule (1,2)-QCNF

- Soit $X = \{x_1, x_2\}$ et $Y = \{y_1, y_2, y_3\}$
- Alors

$$\forall X\ \exists Y\ (x_1 \vee \bar{y}_1 \vee y_2) \wedge (\bar{x}_2 \vee y_2 \vee y_3) \wedge (\bar{x}_1 \vee \bar{y}_1 \vee \bar{y}_3)$$

désigne la formule

$$\forall x_1 \forall x_2 \exists y_1 \exists y_2 \exists y_3\ (x_1 \vee \bar{y}_1 \vee y_2) \wedge (\bar{x}_2 \vee y_2 \vee y_3) \wedge (\bar{x}_1 \vee \bar{y}_1 \vee \bar{y}_3)$$

Un algorithme exhaustif pour (1,2)-QSAT

- Etant donnée une formule (1,2)-QCNF

$$\forall X \, \exists Y \; \varphi(X, Y)$$

  Regarder si, pour toute assignation des variables X, il existe une assignation des variables Y telle que la formule 3-CNF φ soit vraie

- Complexité

$$O(2^{|X|} \cdot |\varphi|)$$

- Permet de voir (1,2)-QSAT comme la conjonction de $2^{|X|}$ problèmes 2-SAT

(1,2)-QSAT et sa complexité

Soit m le nombre de variables universelles
n le nombre de variables existentielles

- Si m est constant, (1,2)-QSAT est résoluble en temps linéaire
- Si $m = \alpha\lceil \log n \rceil$, (1,2)-QSAT est résoluble en temps polynomial
- Si $m = n$, (1,2)-QSAT est coNP-complet
  (Flögel, Karpinski and Kleine Büning, 1990)

L'intérêt d'(1,2)-QSAT (a priori)

- C'est un problème quantifié, pour lequel les questions usuelles se posent :
  - Quel modèle probabiliste ?
  - Y a-t-il une transition de phase VRAIE/FAUSSE ?
  - Y a-t-il un motif facile-difficile-facile ?

  (Cadoli et al. 97, Gent and Walsh 99, Chen and Interian 05)
- Sa complexité "interpole" entre linéaire et coNP-complet, en fonction de m, le nombre de variables universelles
- Il y a une caractérisation combinatoire simple des formules (1,2)-QSAT

Un modèle de formules (1,2)-QCNF aléatoire

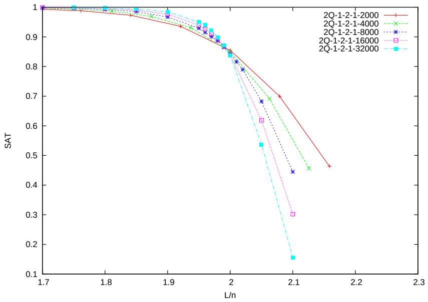
- m est le nombre de variables universelles, $\{x_1, \ldots, x_m\}$
- n est le nombre de variables existentielles, $\{y_1, \ldots, y_n\}$
- Une formule $\forall X \exists Y \varphi(X, Y)$ à $L = c.n$ clauses est tirée au hasard uniforme

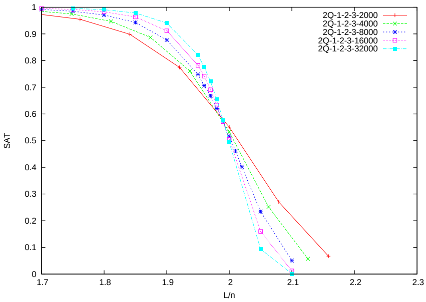
Nombre de clauses possibles : $N = 2^3 \binom{m}{1} \binom{n}{2}$

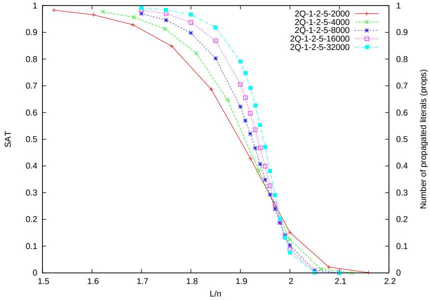
Nombre de formules possibles : $\binom{N}{L}$

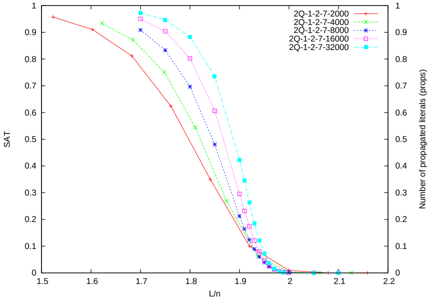
Formule aléatoire : $\mathcal{F}(n, m, c)$

Courbes pour $m = 1, 3, 5, 7$ et divers n



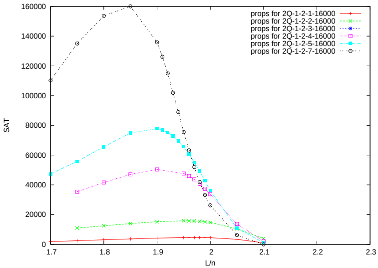


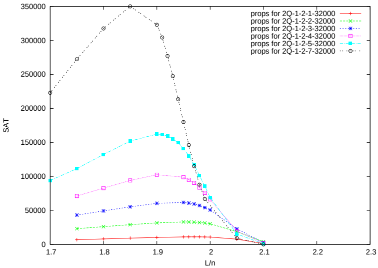




Effort de calcul

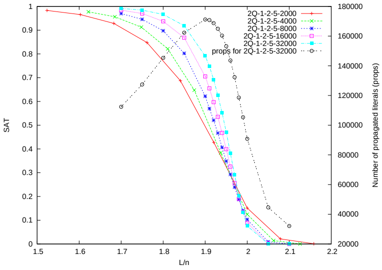
- Effort de calcul : nombre de propagations de littéral par formule
- Pour quelle valeur du ratio l'effort de calcul est-il maximisé ?

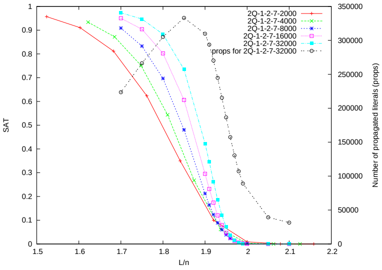




Motif easy-hard-easy

Le pic d'effort de calcul coïncide avec la transition





Plan

(1,2)-QSAT
- Définition et complexité
- Le phénomène de seuil d'(1,2)-QSAT
- Résultats expérimentaux

Résultats

Lien avec 2-SAT

Perspectives

Préliminaires

- Si $c > 2$, la formule est UNSAT(a.g.p)
- Si m est assez petit,

$$m \leq \frac{\log n}{\log 2},$$

il y a un seuil étroit en $c = 2$

- Si $c < 1$, la formule est SAT (a.g.p)
- Si m est assez grand,

$$m \gg \ln n,$$

il y a un seuil étroit en $c = 1$

Un régime intermédiaire quand $m = \lfloor \alpha \ln n \rfloor$

Théorème

Soit $c^*(\alpha)$ la solution de l'équation $\alpha \cdot H(c) = 1$, où

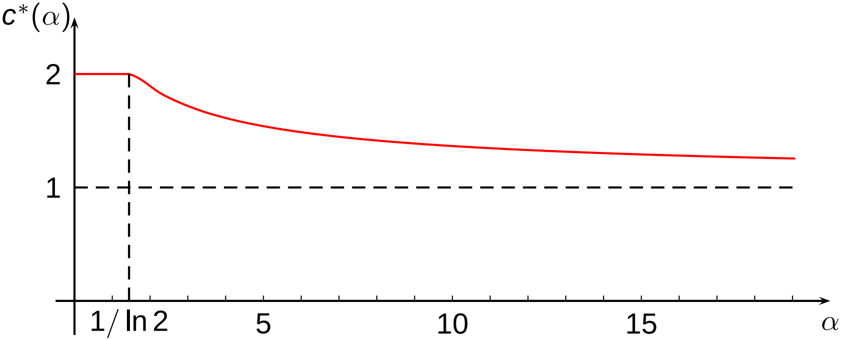
$$H(c) = \ln(c) + \left(\frac{2}{c} - 1\right)\ln(2 - c)$$

Pour tout $\alpha > \frac{1}{\ln 2}$,

- *si $c < c^*(\alpha)$, alors $\mathcal{F}(n, m, c)$ est SAT (a.g.p)*
- *si $c > c^*(\alpha)$, alors $\mathcal{F}(n, m, c)$ est UNSAT (a.g.p)*

Lieu du seuil

Valeur du ratio critique en fonction de α



Plan

Représentation d'une formule 2-CNFpar un digraphe

Formule		Digraphe
$l_1 \vee l_2$	$\rightleftharpoons$	$(\overline{l_1} \rightarrow l_2,\ \overline{l_2} \rightarrow l_1)$

Ex : $(y_1 \vee y_2) \wedge (\overline{y_2} \vee y_3) \wedge (\overline{y_2} \vee \overline{y_3})$ devient :

$$\begin{tikzcd}
\overline{y_1} \arrow[r] & y_2 \arrow[r] \arrow[d] & y_3 \arrow[d] & \\
 & \overline{y_3} \arrow[r] & \overline{y_2} \arrow[r] & y_1
\end{tikzcd}$$

Une formule est UNSAT

$\Updownarrow$

elle contient un cycle contradictoire $y \rightsquigarrow \overline{y} \rightsquigarrow y$

Représentation d'une formule 2-CNFpar un digraphe

Formule		Digraphe
$l_1 \vee l_2$	$\rightleftharpoons$	$(\overline{l_1} \to l_2,\ \overline{l_2} \to l_1)$

Ex : $(y_1 \vee y_2) \wedge (\overline{y_2} \vee y_3) \wedge (\overline{y_2} \vee \overline{y_3})$ devient :

$$\begin{tikzcd}
\overline{y_1} \arrow[r] & y_2 \arrow[r] \arrow[d] & y_3 \arrow[d] & \\
& \overline{y_3} \arrow[r] & \overline{y_2} \arrow[r] & y_1
\end{tikzcd}$$

Une formule est UNSAT

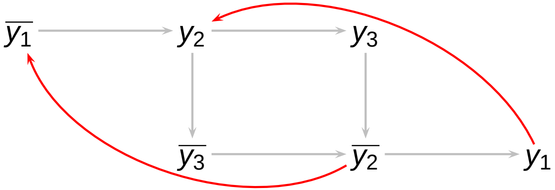
$\Updownarrow$

elle contient un cycle contradictoire $y \rightsquigarrow \overline{y} \rightsquigarrow y$

Représentation d'une formule 2-CNFpar un digraphe

Formule		Digraphe
$l_1 \vee l_2$	$\rightleftharpoons$	$(\overline{l_1} \to l_2,\ \overline{l_2} \to l_1)$

Ex : $(y_1 \vee y_2) \wedge (\overline{y_2} \vee y_3) \wedge (\overline{y_2} \vee \overline{y_3}) {\color{red}\wedge (\overline{y_1} \vee y_3)}$ devient :



Une formule est UNSAT

$\Updownarrow$

elle contient un cycle contradictoire $y \rightsquigarrow \overline{y} \rightsquigarrow y$

Le phénomène de seuil de 2-SAT

- La transition SAT/UNSAT pour 2-SAT est étroite, et le ratio critique vaut 1
  (Chvatal & Reed, Goerdt, 1992)
- La largeur de la fenêtre de seuil est connue (Bollobàs *et al.*, 2001)

Représentation d'une formule (1,2)-QCNFpar un digraphe étiqueté

Formule		Digraphe
$x \vee y_1 \vee y_2$	$\rightleftharpoons$	$(\overline{y_1} \xrightarrow{\overline{x}} y_2,\ \overline{y_2} \xrightarrow{\overline{x}} y_1)$

Instancier x à Vrai $\iff$ supprimer les arcs étiquetés par $\overline{x}$

Donc une formule est UNSAT

$\Updownarrow$

elle contient un cycle contradictoire pur
(i.e sans étiquettes complémentaires)

Représentation d'une formule (1,2)-QCNFpar un digraphe étiqueté

Formule		Digraphe
$x \vee y_1 \vee y_2$	$\rightleftharpoons$	$(\overline{y_1} \xrightarrow{\overline{x}} y_2,\ \overline{y_2} \xrightarrow{\overline{x}} y_1)$

Instancier x à Vrai $\Longleftrightarrow$ supprimer les arcs étiquetés par $\overline{x}$

Donc une formule est UNSAT

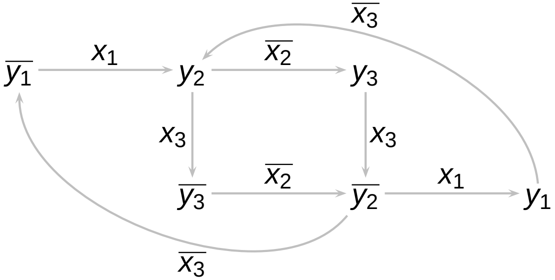
$\Updownarrow$

elle contient un cycle contradictoire pur

(i.e sans étiquettes complémentaires)

Exemple

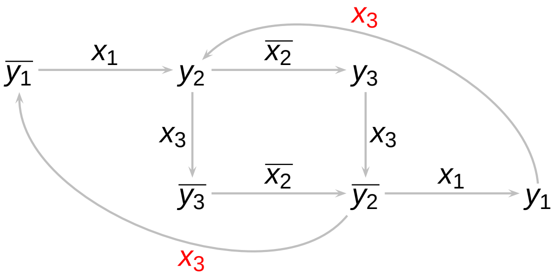
$$(\overline{x_1} \vee y_1 \vee y_2) \wedge (x_2 \vee \overline{y_2} \vee y_3) \wedge (\overline{x_3} \vee \overline{y_2} \vee \overline{y_3}) \wedge (x_3 \vee \overline{y_1} \vee y_2)$$



SAT

Exemple

$$(\overline{x_1} \vee y_1 \vee y_2) \wedge (x_2 \vee \overline{y_2} \vee y_3) \wedge (\overline{x_3} \vee \overline{y_2} \vee \overline{y_3}) \wedge (\overline{x_3} \vee \overline{y_1} \vee y_2)$$



UNSAT

Conséquences

On peut voir (1,2)-QSAT comme la conjonction de 2^m sous-problèmes 2-SAT “entrelacés”.

- On a une solution commune à tous les sous-problèmes 2-SAT si $c < 1$ (a.g.p)
- Pour chaque sous-problème, si $c > 2$, il n’y a pas de solution a.g.p

Plan

Comprendre la structure des solutions

Dans la phase SAT, une formule (1,2)-QSAT correspond à 2^m formules 2-CNF dont les espaces de solutions sont non vides

- La phase $c < 1$ correspond à une intersection non vide des 2^m espaces de solutions. En dehors de cette phase, quelle est la structure des intersections des espaces de solution ?
- Existe-t-il des solutions de (1,2)-QSAT (i.e collection de 2^m points de $\{0,1\}^n$) qui soient de faible diamètre ?
- Peut-on le “voir” expérimentalement ? Repartir de la dernière solution trouvée …
- Lorsque $m \leq \log_2 n$ et c proche de 2, lien avec la sensibilité au bruit de 2-SAT ?

Conclusion

- Quand $m = \lfloor \alpha \ln n \rfloor$, on donne le lieu exact $c^*(\alpha)$ du seuil en fonction de α
- Il faudrait comprendre la nature de la géométrie des solutions de (1,2)-QSAT et en tirer partie expérimentalement. Nécessite de mieux comprendre 2-SAT.

Idée de preuve : Chvatal et Reed (2-SAT)

On peut représenter une formule 2-CNF comme un graphe dirigé sur les variables existentielles

- Toute formule insatisfaisable contient un bicycle (facile à dénombrer)
- Toute formule contenant un snake est insatisfaisable (nb de snakes concentré)

Chvatal et Reed modifiés

On peut représenter une formule (1,2)-QCNF comme un graphe dirigé sur les variables existentielles, avec des arcs étiquetés par les variables universelles

- Toute formule fausse contient un bicycle pur
- Toute formule contenant un snake pur est fausse

Premier moment sur les bicycles et second moment sur des snakes bien choisis.