

Transitions de phase et complexité en informatique

Les transitions de phase, omniprésentes en physique de la matière condensée, se rencontrent aussi, de manière suprenante, en informatique. L'existence de phénomènes critiques a des conséquences dramatiques sur la complexité, c'est-à-dire le temps de résolution, de nombreux problèmes d'optimisation ou de décision. Les méthodes et concepts de la physique statistique des systèmes désordonnés et hors équilibre apportent un éclairage nouveau sur le fonctionnement dynamique des algorithmes.

Imaginez un instant être le chef du protocole, responsable de l'organisation du prochain bal de l'ambassade. Le prince vous ordonne d'inviter l'ambassadeur du Pérou ou d'exclure celui du Qatar. La reine vous demande de convier ce dernier ou le représentant de la Roumanie, ou encore les deux à la fois. Quant au roi, d'une humeur massacrant, il serait par trop incommodé de la présence simultanée des diplomates péruvien et roumain. Êtes-vous en mesure de satisfaire tous les caprices de la famille royale ?

Cet exemple, dû à Brian Hayes, est un cas particulier du problème de la « satisfaisabilité » (SAT), d'importance considérable tant sur le plan pratique que théorique. Il s'agit de savoir si un ensemble de contraintes logiques portant sur des variables booléennes (ayant seulement deux valeurs possibles, vrai ou faux) admet ou non des solutions. D'un point de vue formel, le chef du protocole doit résoudre un problème SAT comportant trois variables booléennes p, q, r [indiquant si les ambassadeurs de chaque pays seront présents (vrai) ou absents (faux)] et trois contraintes, appelées clauses, qui s'écrivent comme $(p \text{ ou } \bar{q})$, $(q \text{ ou } r)$, $(\bar{p} \text{ ou } \bar{r})$. La barre désigne la

négation logique ; par exemple \bar{p} est vraie si l'ambassadeur péruvien n'est pas invité. Quelques secondes de réflexion suffisent dans le cas présent à passer en revue les $8 = 2^3$ listes d'invités *a priori* possibles et à en sélectionner une satisfaisante. Mais on ne connaît pas de méthodes efficaces pour résoudre les grands problèmes SAT. Même avec les meilleurs algorithmes disponibles, les temps de résolution peuvent augmenter dramatiquement (exponentiellement) avec la taille des problèmes, c'est-à-dire leurs nombres de variables et de clauses. SAT est un paradigme de problème computationnel difficile et, à ce titre, joue un rôle central en informatique théorique (encadré 1). Comprendre pourquoi SAT est difficile et améliorer les performances des algorithmes aurait aussi des retombées bénéfiques pour la résolution des problèmes d'optimisation rencontrés fréquemment dans les applications industrielles.

3-SAT ALÉATOIRE ET SEUIL DE TRANSITION

Avec ces objectifs en tête, les informaticiens ont entrepris il y a quelques années l'étude d'une version particulièrement intéressante de SAT, la 3-SAT aléatoire. Il s'agit d'un modèle simplifié, mais reproduisant les caractéristiques essentielles des problèmes SAT survenant dans la pratique et qui dépend seule-

ment de deux paramètres : le nombre N de variables booléennes et le rapport α du nombre de clauses divisé par N . Chaque clause est de longueur trois, d'où la dénomination 3-SAT : elle contient trois variables ou leurs négations logiques. Les αN clauses sont tirées au hasard, indépendamment les unes des autres. Les injonctions de la famille royale correspondent à un tirage aléatoire de problème 2-SAT, portant sur $N = 3$ variables et avec un rapport $\alpha = 1$.

Intuitivement, α représente un nombre de contraintes par degré de

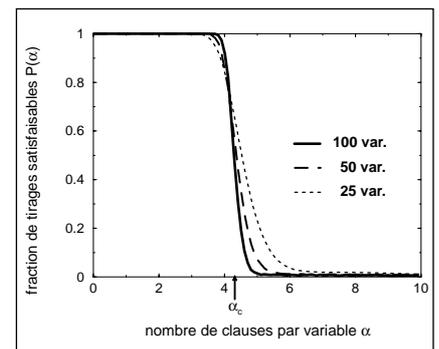


Figure 1 - Probabilité $P(\alpha)$ qu'un ensemble de αN clauses formant un problème 3-SAT soit satisfaisable pour différentes tailles N (nombre de variables). Cette probabilité est obtenue numériquement en générant, pour chaque valeur de α et N , 10 000 tirages de problèmes 3-SAT et en mesurant la fraction de ceux-ci qui admet au moins une solution. Dans la limite de taille N infinie, la valeur $\alpha_c \simeq 4.3$ sépare la région satisfaisable ($\alpha < \alpha_c$, $P = 1$) de la région non satisfaisable ($\alpha > \alpha_c$, $P = 0$). Une telle transition existe aussi pour le cas de 2-SAT aléatoire, où toutes les clauses contiennent deux variables ; le seuil vaut alors exactement $\alpha_c = 1$.

– Laboratoire de physique théorique de l'ENS (LPTENS), UMR 8549 CNRS, université Paris 11, ENS, 24 rue Lhomond, 75231 Paris cedex 05.

Encadré 1

NOTION DE COMPLEXITÉ EN INFORMATIQUE

Les tâches computationnelles sont classées selon leurs complexité, c'est-à-dire le nombre d'opérations élémentaires (additions, substitutions, comparaisons, ...) qu'un algorithme doit effectuer pour les résoudre. Il est d'usage de considérer qu'un problème dont la complexité augmente polynômialement avec sa taille, à savoir le nombre de variables ou paramètres qui servent à le définir, est facile, alors qu'une croissance plus forte, par exemple exponentielle, est synonyme de problème **difficile**.

De manière générale, pour montrer qu'un problème est facile, il faut trouver explicitement un algorithme qui le résout en un nombre d'opérations polynômial. Un exemple important dans la pratique est celui du tri. On se donne une liste de N nombres qu'il faut classer par ordre croissant. Les meilleurs algorithmes sont capables d'exécuter cette tâche en un nombre d'opérations proportionnel à $N \log N$ lorsque N est grand, et cela pour n'importe quelle liste. Il a été démontré que la 2-satisfaisabilité, où chaque clause comporte exactement deux variables booléennes, fait partie de cette classe P des problèmes polynômiaux **faciles**.

La 3-satisfaisabilité par contre, ainsi que des centaines de problèmes d'optimisation, comme le célèbre voyageur de commerce (consistant à chercher la tournée la plus courte passant par N villes), ont jusqu'ici résisté à des décennies d'efforts et de recherches d'algorithmes de résolution polynômiaux. Dans la pratique, les seuls algorithmes que l'on connaisse pour résoudre ces problèmes peuvent nécessiter un nombre exponentiel d'opérations. Pire, les informaticiens théoriciens ont démontré que, si un seul d'entre eux était facile, alors tous les autres le seraient également, ce qui rend cette éventualité peu plausible. De tels problèmes sont appelés NP -complets. Cependant, d'un point de vue rigoureux, démontrer l'impossibilité de résoudre facilement tous les problèmes NP -complets fait partie des sept questions fondamentales ouvertes en mathématiques énoncées au congrès Clay de Paris en mai 2000. Il n'est donc pas formellement exclu qu'il existe un algorithme totalement inédit capable de résoudre facilement 3-SAT !

liberté. Il est légitime de penser que, à N fixé, un problème 3-SAT aléatoire admettra des solutions si α est petit (problème sous-contraint), et sera de plus en plus difficile à satisfaire au fur et à mesure que α augmente (problème sur-contraint). C'est bien ce que les informaticiens ont observé mais avec une surprise de... taille ! La figure 1 montre la probabilité $P(\alpha)$ qu'un problème 3-SAT aléatoire soit satisfaisable pour différentes tailles N . Lorsque N augmente, un changement abrupt survient à une valeur critique $\alpha_c \simeq 4.3$: les problèmes ayant moins de $\alpha_c N$ clauses sont presque toujours satisfaisables, alors que ceux comportant plus de $\alpha_c N$ clauses n'ont presque jamais de solution.

Cette transition s'accompagne d'une explosion de la difficulté de résolution, appelée aussi complexité.

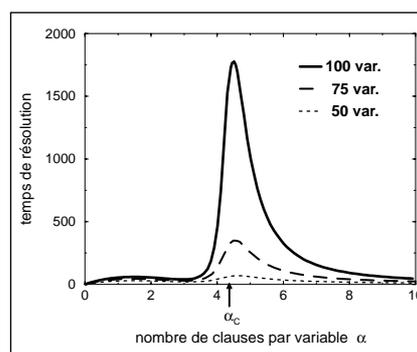


Figure 2 - Temps de résolution de 3-SAT aléatoire en fonction du nombre de clauses par variable α et pour trois tailles différentes. Le temps est estimé à partir de la grosseur de l'arbre de recherche engendré par l'algorithme de résolution (figure 5). On montre ici le temps médian pour 10 000 problèmes tirés au hasard, le temps moyen pouvant être perturbé par la présence de quelques tirages exceptionnels, peu représentatifs, et très difficiles à résoudre. La croissance avec la taille est maximale au seuil α_c . Elle est exponentielle autour du seuil et dans toute la région non satisfaisable, mais de moins en moins fortement lorsque α augmente.

La figure 2 montre le temps médian pour résoudre un problème 3-SAT aléatoire, c'est-à-dire pour en trouver une solution (en deçà du seuil α_c) ou vérifier qu'il n'y en a pas (au-delà du seuil). A première vue, la complexité computationnelle est concentrée autour du seuil et augmente exponentiellement avec la taille au voisinage de la transition. Bien évidemment, la courbe de la figure 2 dépend quantitativement de la méthode de résolution utilisée. Cependant, cette allure de la complexité en fonction de α , à savoir la succession de régions *grosso modo* facile, difficile et moins difficile, reste valable pour tous les algorithmes disponibles actuellement.

Pourquoi les physiciens devraient-ils s'intéresser au problème 3-SAT, qui est un objet de nature purement mathématique ? Tout d'abord, 3-SAT est compliqué. Les mathématiciens

Dictionnaire

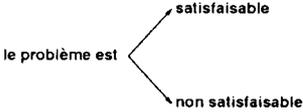
K-SAT	physique statistique
variable booléenne $x = \text{vraie, fausse}$	spin d'Ising $s_x = +1, -1$
clauses	couplages et champs agissant sur les spins
nombre de clauses violées par une configuration logique	énergie E de la configuration des spins
par exemple :	
2 SAT x ou \bar{y} $(x$ ou $\bar{y})$ et $(\bar{x}$ ou $z)$	$E = \frac{1}{4}(1-s_x)(1+s_y)$ $E = \frac{1}{4}(1-s_x)(1+s_y) + \frac{1}{4}(1+s_x)(1-s_y)$
3 SAT x ou \bar{y} ou z	$E = \frac{1}{8}(1-s_x)(1+s_y)(1-s_z)$
nombre minimal de clauses violées	énergie de l'état fondamental
	énergie de l'état fondamental = 0
le problème est 	énergie de l'état fondamental > 0

Figure 3 - Dictionnaire permettant le passage des objets définis pour SAT (colonne de gauche) à ceux habituellement rencontrés en physique statistique (colonne de droite).

sont parvenus à établir des bornes rigoureuses encadrant la valeur du seuil, $3.14 \leq \alpha_c \leq 4.51$, mais la valeur exacte du seuil semble hors d'atteinte par les méthodes actuelles. Les méthodes d'approximation développées par les physiciens pour étudier les transitions de phase permettent non seulement d'estimer le seuil, mais aussi de dévoiler la structure microscopique des solutions et le mécanisme de leur disparition. Les concepts de diagramme de phase, de flot dynamique, d'évolution markovienne, se révèlent également utiles pour comprendre le fonctionnement des algorithmes. La physique statistique est ainsi à même de fournir des intuitions précieuses, dont certaines ont déjà été utilisées par les mathématiciens pour établir des résultats rigoureux.

UN VERRE DE SPINS A TEMPÉRATURE NULLE

Pour étudier cette transition à l'aide des concepts et méthodes de la physique statistique moderne, commençons par traduire les ingrédients

du problème 3-SAT (variables booléennes, clauses, ...) en termes plus familiers en physique (spins d'Ising, interactions, ...). Un dictionnaire est établi à cet effet figure 3. L'idée principale est d'introduire une énergie, représentant le nombre de clauses non satisfaites pour chaque configuration des variables-spins, et de déterminer l'état fondamental du système de spins correspondant. La valeur nulle ou strictement positive de l'énergie fondamentale indique alors sans ambiguïté si le problème 3-SAT est satisfaisable ou non.

D'un point de vue physique, 3-SAT s'apparente par son hamiltonien à des matériaux très étudiés en matière condensée depuis 25 ans, les verres de spins. Par rapport aux systèmes magnétiques habituels, les verres de spins sont désordonnés : les ions magnétiques qu'ils renferment, du fait de leurs positions figées et quelconques dans la matrice métallique, ont des interactions d'intensités et de signes aléatoires. L'absence d'homogénéité dans les interactions rend le paysage d'énergie extrêmement complexe, ce qui provoque l'apparition de nom-

breux états métastables et bouleverse les propriétés physiques de basse température (encadré 2). A température nulle, une telle richesse de comportement peut résulter de la présence d'un fort désordre. Dans le cas de 3-SAT, ce désordre provient des clauses aléatoires et est d'autant plus fort que la valeur de α est élevée.

L'image « physique » de 3-SAT aléatoire est la suivante. Il existe un seuil α_c au-delà duquel l'énergie du fondamental cesse, avec probabilité un, d'être nulle et devient strictement positive. A cause de la difficulté des calculs analytiques, des approximations sont nécessaires et nous avons obtenu $\alpha_c \simeq 4.48$, supérieur d'environ 4 % aux valeurs numériques. Certains aspects de la transition sont surprenants. Immédiatement à gauche du seuil, le nombre de solutions est encore énorme, environ égal à $2^{0.14N}$. Le franchissement du seuil, consécutif à l'ajout d'un nombre de clauses très petit devant N , suffit à faire disparaître toutes ces solutions ! La transition s'accompagne en outre de l'apparition brutale de nombreuses variables sur-contraintes, ayant toujours les mêmes valeurs dans toutes les configurations d'énergie minimale. Ce squelette de variables joue le rôle de paramètre d'ordre. Il est nul en dessous du seuil et saute discontinûment à α_c , à une valeur de l'ordre de 15 %. La transition de 3-SAT peut être interprétée comme étant du premier ordre.

Cette transition brutale contraste avec celle advenant pour le problème 2-SAT aléatoire, moins intéressant du point de vue de la complexité que 3-SAT, car facilement soluble en temps polynômial (encadré 1). La transition de 2-SAT aléatoire à $\alpha_c = 1$ est du deuxième ordre et aucun squelette n'émerge brutalement. Pour mieux cerner d'éventuels liens entre les caractéristiques de la transition de phase et la complexité computationnelle, nous avons introduit un modèle mixte, appelé 2+p-SAT, qui juxtapose

Encadré 2

SOLUTION DE SAT ET VERRE DE SPINS A TEMPÉRATURE NULLE

Le nombre de clauses non satisfaites pour le problème du bal de l'ambassade s'écrit à l'aide du dictionnaire introduit en figure 3 :

$$E = -(s_p s_q - s_q s_r - s_p s_r - 3)/4.$$

La fonction E définit un hamiltonien pour les trois spins d'Ising s_i , d'indices $i = p, q, r$, situés aux sommets d'un triangle (figure). Ce modèle est un cas particulier de verre des spins. Tout d'abord, les couplages ne sont pas tous de même signe comme dans le cas du modèle d'Ising. En outre, ces signes sont des variables aléatoires qui dépendent du tirage particulier des clauses et définissent un désordre figé. Enfin, si dans le cas du bal de l'ambassade, les trois énergies de lien peuvent être minimisées simultanément (les produits des deux spins aux extrémités de chaque lien peuvent être rendus égaux aux signes des couplages portés par ces mêmes liens), en général cela n'est pas le cas et le système est frustré (pour cela, il suffirait d'inverser un seul couplage du triangle ci-contre). Désordre et frustration sont les deux caractéristiques fondamentales des verres de spins.

Les problèmes SAT conduisent à des modèles de verres des spins particulièrement intéressants. Chaque spin interagit avec un nombre limité de spins parmi tous les autres, ce qui donne lieu à un réseau à connectivité finie mais variable. Ce réseau est lui aussi désordonné et dépend du tirage des clauses. De surcroît, 3-SAT définit un système plutôt inhabituel en physique parce que les spins interagissent, non pas deux par deux comme dans la 2-SAT de l'exemple ci-contre, mais trois par trois (figure 3).

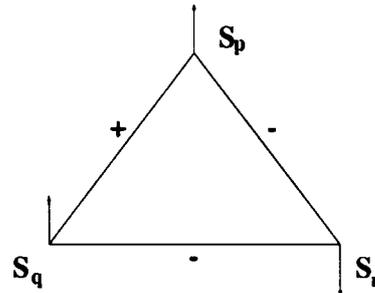


Figure - Spins et interactions pour le problème du bal.

Dans les verres de spins, les observables thermodynamiques, comme par exemple l'énergie libre, dépendent a priori de la réalisation du désordre figé. Cependant, dans la limite thermodynamique, elles sont auto-moyennantes, c'est-à-dire ne fluctuent pas de réalisation à réalisation. On peut donc obtenir la densité d'énergie libre en calculant la valeur moyenne, sur la distribution du désordre, du logarithme de la fonction de partition.

Le calcul de cette valeur moyenne est difficile et requiert l'utilisation de la méthode des répliques, propre aux systèmes désordonnés. Ainsi, nous avons pu calculer de manière approchée la valeur du seuil α_c , et caractériser certaines propriétés microscopiques des solutions (fraction de variables fixées, nombre de spins qui diffèrent d'une solution à une autre, ...).

$(1 - p)\alpha N$ clauses de longueur 2 et $p\alpha N$ clauses de longueur 3. Ce modèle interpole continûment entre la 2-SAT ($p = 0$) et la 3-SAT ($p = 1$). Il est caractérisé par un point tricritique à $p_0 \simeq 0.4$, qui sépare une transition du type 3-SAT ($p_0 < p < 1$), du premier ordre en le squelette, d'une transition du type 2-SAT ($0 < p < p_0$), sans apparition brutale de variables sur-contraintes (figure 4).

ALGORITHME DE RÉOLUTION ET TRAJECTOIRES

Pourquoi la transition influe-t-elle sur la complexité de résolution de

3-SAT ? Pour répondre à cette question, il est nécessaire d'exposer brièvement la méthode de résolution, dite de Davis et Putnam (DP). Selon une stratégie très répandue en optimisation, la procédure de DP consiste en une succession d'essais et de corrections d'erreurs, qui peuvent être représentés sous forme d'un arbre (figure 5). Brièvement, (1) un nœud de l'arbre correspond au choix d'une variable et de sa valeur. Selon que cette dernière est vraie ou fausse, on emprunte l'une des deux branches possibles. (2) Le long d'une branche, on analyse toutes les implications logiques du dernier choix effectué. (3) Si une contradiction (une clause insatisfaite) surgit,

on modifie le dernier choix (remonter dans l'arbre) et on poursuit sur une nouvelle branche en allant à l'étape 2 ; si toutes les clauses sont satisfaites, une solution est trouvée et on arrête la recherche ; sinon, on retourne à l'étape 1.

La complexité computationnelle est directement donnée par la taille de l'arbre de recherche, plus précisément par le nombre de ses nœuds. Les performances peuvent être améliorées plus ou moins empiriquement en jouant sur la règle heuristique de choix des variables (étape 1).

L'algorithme de DP définit un processus dynamique complexe, non markovien et original par rapport

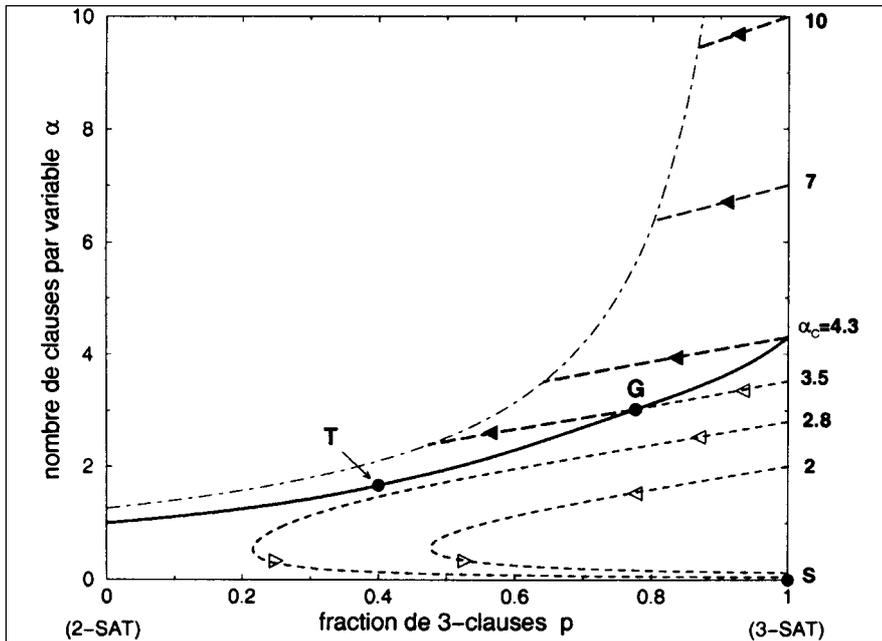


Figure 4 - Diagramme des phases du modèle $2 + p$ -SAT et flot dynamique de l'algorithme de Davis et Putnam. La ligne critique $\alpha_c(p)$ (ligne pleine foncée) sépare la région satisfaisable (en dessous de la ligne) de la phase sans solution (au-dessus). Ses extrémités ont pour ordonnées 1 ($p = 0$, 2-SAT) et 4.3 ($p = 1$, 3-SAT). T est le point tricritique séparant les transitions continues (à gauche de T) de celles discontinues (à droite de T). Les lignes en tirets représentent les trajectoires de branche unique (lignes minces, flèches évidées) et d'arbre touffu (lignes épaisses, flèches foncées). Les flèches indiquent la direction du mouvement. Les points de départ pour les problèmes 3-SAT sont situés sur l'axe vertical $p = 1$ avec les rapports α correspondant. Pour $\alpha < \alpha_L \approx 3.003$, les trajectoires de branche unique restent confinées à la phase satisfaisable et se terminent en S , de coordonnées $(1,0)$, où une solution est trouvée. Dans l'intervalle $\alpha_L < \alpha < \alpha_C$, la trajectoire de branche unique traverse la ligne critique en G , dont la position dépend de la valeur initiale de α . Un arbre touffu croît alors dans la phase sans solution, comme lorsque le rapport initial α est au-dessus du seuil $\alpha_C \approx 4.3$. Les trajectoires des branches dominantes des arbres touffus s'arrêtent sur la ligne de contradiction (trait mixte). Dès lors, DP atteint son point de remontée maximale dans l'arbre, c'est-à-dire le premier noeud si $\alpha > \alpha_C$ ou le noeud G si $\alpha_L < \alpha < \alpha_C$. Dans ce dernier cas, une solution peut être trouvée le long d'une nouvelle branche alors que, dans le premier cas, l'absence de solution est prouvée (figure 6).

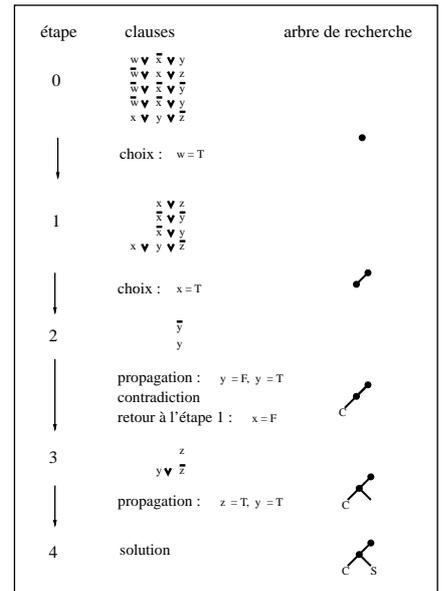


Figure 5 - L'algorithme de Davis-Putnam (DP) est une méthode exacte permettant de résoudre les problèmes SAT sans effectuer l'énumération exhaustive des 2^N configurations possibles des variables booléennes. **Étape 0** : Le problème 3-SAT initial porte sur 5 clauses et $N = 4$ variables. Chaque variable peut être vraie (V) ou fautive (F); \bar{x} désigne la négation logique de x . Les colonnes représentent le numéro de l'étape de résolution, l'ensemble des clauses à satisfaire et l'arbre de recherche indiquant l'historique des opérations déjà effectuées. **1** : DP choisit une variable selon une heuristique, c'est-à-dire une règle empiriquement connue pour être efficace. Dans le cas présent, l'heuristique consiste à choisir une variable aléatoirement (ici w) dans une clause de longueur minimale (la première) et à donner à cette variable la valeur qui satisfait la clause ($w = V$). Un nœud (point noir) et une ramification (segment) sont ajoutés à l'arbre de recherche pour représenter le choix de la variable (w) et de sa valeur (V) respectivement. **2** : les clauses contenant w sont satisfaites et éliminées, celles où figure \bar{w} sont simplifiées ; quant aux autres, elles restent inchangées. **3** : comme 1. **4** : quand des clauses unitaires (de longueur un) sont présentes, les variables qu'elles contiennent sont obligatoirement fixées pour les satisfaire. C'est la phase de propagation. **5** : une contradiction (C) survient et la branche en cours est terminée. **6** : DP remonte dans l'arbre jusqu'au choix précédent (nœud x), inverse la valeur de la variable correspondante ($V \rightarrow F$) et crée une nouvelle branche. **7** : comme 4. **8** : le processus s'arrête quand toutes les clauses sont satisfaites (une solution $S = \{V, F, V, V\}$ est trouvée) ou quand toutes les branches se terminent par C (l'existence de solution est alors prouvée).

aux évolutions habituelles des systèmes physiques. Afin de l'étudier, nous sommes partis de la constatation suivante (figure 5). La procédure récursive de DP transforme le problème 3-SAT de départ en un problème mixte $2 + p$ -SAT contenant des clauses de longueurs deux et trois, les clauses contenant une seule variable étant éliminées à l'étape 2 de l'algorithme ci-dessus. Donc, si l'on symbolise le problème 3-SAT initial comme un point de coordonnées $p = 1, \alpha$ dans le diagramme de phase de la figure 4, ce point va évoluer sous l'action dynamique de l'algorithme et décrire une trajectoire. Trois intervalles différents de rapports ini-

tiaux α doivent être distingués sur l'axe vertical $p = 1$ correspondant à 3-SAT.

Pour $\alpha < \alpha_L = 3.003$, l'arbre de recherche est essentiellement constitué d'une seule branche (figure 6A), qui débouche sur une solution. La trajectoire correspondant à cette branche dépend de l'heuristique de choix des variables utilisée par DP (figure 5), et est représentée figure 4. Elle se dirige d'abord vers la gauche, puis rebrousse chemin pour venir toucher l'axe de la 3-SAT à une valeur de α inférieure à celle de départ. La trajectoire s'arrête au point S où toutes les clauses sont éliminées ($\alpha = 0$) et une solution est trouvée. La complexité typique est

d'ordre N ; résoudre un problème 3-SAT est donc facile dans cette région, comme la figure 2 l'atteste.

Encadré 3

DYNAMIQUE DE CROISSANCE DE L'ARBRE ET TEMPS DE RÉOLUTION

L'algorithme de DP construit l'arbre de recherche de manière séquentielle, nœuds après nœuds et branches après branches, jusqu'à son achèvement. Comme décrit en figure 5, chaque nœud correspond à un problème $2+p$ -SAT, caractérisé par ses paramètres p et α . Imaginons maintenant un processus parallèle de construction, couche par couche, où les deux branches issues de chaque nœud croissent simultanément. Chaque branche suit les règles d'évolution de DP (choix des variables, propagation des clauses unitaires et arrêt en cas de contradiction), de sorte que l'arbre final est identique à l'arbre de recherche engendré séquentiellement. A chaque « moment » t , l'arbre inachevé comporte de nombreuses branches. Chacune des extrémités de ces branches correspond à un problème $2+p$ -SAT, caractérisé par un couple de paramètres p et α . Considérons l'histogramme $\mathcal{B}(p, \alpha; t)$ des valeurs de ces paramètres.

Cet histogramme croît exponentiellement avec la taille N . Il est commode d'étudier son logarithme

$\Omega(p, \alpha; t) = \frac{1}{N} \log_2 \mathcal{B}(p, \alpha; t)$. Ce dernier peut-être vu comme une surface dans l'espace tridimensionnel p, α, Ω (figure 1).

L'ajout d'une nouvelle couche à l'arbre de recherche fait croître le nombre de branches, en modifiant les paramètres des problèmes SAT qu'elles portent. Ainsi, on se ramène à un problème de croissance bidimensionnelle hors équilibre, où le « temps » t n'est autre que la profondeur de l'arbre en construction. La surface-histogramme évolue selon une équation de croissance où \mathcal{H} incorpore les détails de

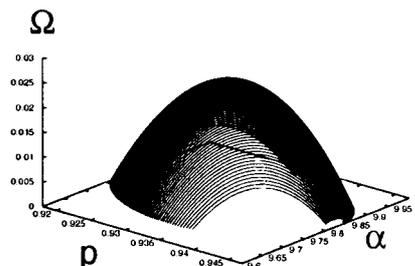


Figure 1 - Exemple d'histogramme $\Omega(p, \alpha)$, à un « instant » t donné.

$$\frac{\partial \Omega}{\partial t} = \mathcal{H} \left[\frac{\partial \Omega}{\partial p}, \frac{\partial \Omega}{\partial \alpha}, p, \alpha, t \right] \quad (1)$$

l'heuristique de choix de DP. Le nombre de branches $\mathcal{B}(t)$ à profondeur t vaut, à l'ordre dominant en N , $2^{N \Omega^*(t)}$, où $\Omega^*(t)$ est la hauteur du sommet de la surface atteint au point de coordonnées $p^*(t), \alpha^*(t)$. Les branches dominantes portent donc des problèmes $2+p$ -SAT ayant $p^*(t), \alpha^*(t)$ pour paramètres caractéristiques, et définissent les trajectoires d'arbre de la figure 5.

La croissance cesse à la profondeur t_a , lorsque des contradictions apparaissent au bout des branches dominantes (ligne d'arrêt en trait mixte de la figure 4). Le nombre final de branches, c'est-à-dire d'extrémités C (figure 6), est $\mathcal{B}(t_a) = 2^{N \Omega^*(t_a)}$. Pour n'importe quel arbre complet, comme on peut le vérifier sur la figure 6B, ce nombre est égal au nombre de nœuds plus un. Dans la limite de N grand, le calcul de $\Omega^*(t_a)$ nous donne directement ω .

A l'opposé, pour $\alpha > \alpha_c \simeq 4.3$, les problèmes 3-SAT n'ont (presque) jamais de solution, mais la procédure de DP met un temps considérable avant de s'en assurer. L'arbre de recherche est symbolisé figure 6B. Il est composé de beaucoup de branches inégales et est très touffu. Le nombre de nœuds croît exponentiellement avec N et s'exprime commodément comme $2^{N\omega}$. ω , qui dépend de α , peut être calculé analytiquement (encadré 3). L'idée principale est de suivre la croissance des branches dominantes (les plus nombreuses, voir encadré 3) de l'arbre jusqu'à leur extinction, due à l'apparition de contradictions. Les trajectoires des branches dominantes sont

dessinées figure 4 et permettent d'interpréter aisément les principales caractéristiques de la figure 2. En effet, ω est d'autant plus grand que les branches dominantes de l'arbre ont poussé en profondeur et donc que les trajectoires sont longues. Au-dessus du seuil, les trajectoires ont toujours une longueur finie : ω est positif et la complexité est exponentielle. En outre, ω est maximal au seuil et décroît rapidement aux grands rapports α (le problème devient de plus en plus sur-contraint et il est de plus en plus facile de trouver un groupe de clauses incompatibles entre elles).

Enfin, la région intermédiaire $\alpha_L < \alpha < \alpha_c$ est la plus riche car

elle conjugue les deux comportements précédents (voir l'arbre de la figure 6C). La trajectoire de branche unique, issue du point $(p = 1, \alpha)$ correspondant au problème 3-SAT initial (et ayant des solutions), coupe la ligne critique $\alpha_c(p)$ en un point G . A ce moment, l'algorithme engendre, sans le savoir, un problème $2+p$ -SAT qui n'a pas de solution. La formation d'un sous-arbre touffu commence en aval de G , arbre que l'algorithme doit parcourir entièrement. Pour trouver une solution, il n'est pas nécessaire de remonter plus haut que G , car les problèmes $2+p$ -SAT immédiatement en amont de G sont tous satisfaisables dans la limite de grande taille N (figure 6C). DP trouvera finale-

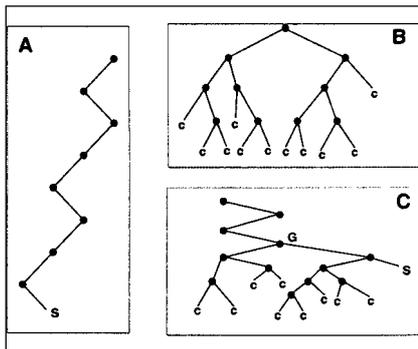


Figure 6 - Structures des arbres de recherche. Les nœuds (cercles noirs) représentent les choix de variables par l'heuristique et les traits entre deux nœuds, l'élimination de variables figurant dans des clauses unitaires. Lorsqu'une contradiction apparaît, la branche meurt et son extrémité est marquée d'un C. L'algorithme s'arrête dès qu'une solution est trouvée (S) ou si l'arbre a été entièrement parcouru en vain. **A.** branche unique (région $0 < \alpha < \alpha_L = 3.003$). **B.** arbre touffu (région $\alpha > \alpha_c$). **C.** cas mixte, branche puis arbre (région $\alpha_L < \alpha < \alpha_c$).

ment une solution au bout de la branche marquée par S sur la figure 6C. Cette branche est non typique et contribue très peu à la complexité, puisque presque toutes les branches de l'arbre sont décrites par la trajectoire de branches dominantes issue du point G (figure 4).

La complexité de résolution du problème 3-SAT initial se résume donc essentiellement à celle du problème 2+p-SAT critique situé en G, et est calculable analytiquement (encadré 3). Dans cette région intermédiaire, 3-SAT est exponentiellement difficile à résoudre.

CONCLUSION ET PERSPECTIVES

L'existence d'une transition et son influence sur la complexité de réso-

lution n'est pas l'apanage du seul problème SAT. De nombreuses autres tâches d'optimisation (couverture de sommets, coloriage de graphes, ...) se comportent de manière très similaire.

L'approche de physique statistique permet une compréhension, non rigoureuse mais précieuse, de ces phénomènes critiques. Elle s'écarte aussi de la théorie de la complexité computationnelle (encadré 1), fondée sur la notion de pire des cas, en soulignant l'importance de la notion de cas typique (qui survient avec très grande probabilité). Ainsi, 3-SAT est NP-complet, donc aucun algorithme ne peut résoudre tous les problèmes en un temps polynômial mais, pour $\alpha < \alpha_L$, la procédure de DP y parvient avec la probabilité un ! La capacité à ne pas tenir compte d'événements rares est un avantage considérable de l'approche statistique. L'espoir, pour l'instant déjà partiellement concrétisé si l'on songe par exemple à l'importance de méthodes d'optimisation issues de la physique, comme le recuit simulé, est d'utiliser cette nouvelle compréhension pour améliorer les algorithmes actuels. La figure 4 montre que la complexité ne serait pas exponentielle en dessous du seuil si l'on disposait d'une heuristique engendrant des trajectoires capables d'éviter la ligne critique.

En retour, les objets étudiés en informatique se révèlent être une source originale et féconde de problèmes théoriques, aux aspects dynamiques inhabituels (dynamique induite par les algorithmes) qui obligent à approfondir et à mieux comprendre certains domaines de la phy-

sique statistique des systèmes désordonnés hors équilibre.

POUR EN SAVOIR PLUS

Hayes (B.), *American Scientist*, **85** 108, 1997.

Cette excellente introduction au problème SAT est disponible sur le web : <http://www.amsci.org/amsci/issues/Comsci97/compsci9703.html>.

Delahaye (J-P.), *Pour la Science*, **213**, juillet 1995.

Hogg (T.), Huberman, (B.A.), Williams, (C.), *Artificial Intelligence*, **81**, I & II, 1996.

De nombreuses références spécialisées ainsi qu'un compte rendu de nos travaux sur les aspects statistiques de la transition se trouvent dans :

Monasson (R.), Zecchina (R.), Kirkpatrick (S.), Selman (B.), Troyansky (L.), *Nature*, **400**, 133, 1999.

La dynamique de recherche par l'algorithme DP est quant à elle exposée dans :

Cocco (S.), Monasson (R.), *Physical Review Letters*, **86**, 1654, 2001.

Pour une introduction à la physique des verres de spins et à la méthode des répliques :

Mézard (M.), Toulouse (G.), *La Recherche*, **232**, 616, 1991.

Article proposé par :
Giulio Biroli, Simona Cocco et Rémi Monasson, tél. 01 44 32 25 13,
remi.manasson@lpt.ens.fr.