

Analyse de problèmes d'optimisation par les outils de mécanique statistique

Vitor Sessak

6 juin 2007

Résumé

Ce rapport a comme but faire une introduction à l'utilisation des outils de physique statistique pour résoudre des problèmes d'informatique théorique. En particulier, il se concentre en deux exemples : d'abord, la connexité des graphes aléatoires, qui se prête à une résolution simple mais intéressante et le K-SAT, qui est notable par son importance en informatique.

1 Introduction

Dans la plupart des problèmes de mécanique statistique, la détermination de l'état fondamental est une partie importante de la compréhension du problème. En particulier, cet état détermine complètement le système quand il est en température zéro et en équilibre. Comme l'état fondamental est par définition l'état qui minimise l'énergie, on est rapporté directement à un problème de optimisation.

Malgré ce lien clair entre ces deux domaines d'étude, on a dû attendre les outils de la théorie des systèmes désordonnés pour que la mécanique statistique puisse réellement agrandir la compréhension de certains problèmes de optimisation. En plus, il a fallu chercher des relations moins triviales entre les deux domaines que la simple détermination de l'état fondamental pour finalement avoir des résultats intéressants.

Dans ce rapport, je me concentrerai en deux problèmes classiques d'informatique théorique : d'abord la connectivité de graphes aléatoires, où j'irais retrouver quelques résultats déjà connus et deuxièmement le problème de K-SAT.

2 La connectivité des graphes aléatoires

Dans cette section on s'intéresse aux graphes aléatoires. Pour les définir, on fixe une fois pour toutes le nombre d'arrêts N et pour chaque un des $N(N-1)/2$ arrêts possibles on décide avec une probabilité $p = \gamma/N$ (resp. $1-p = 1-\gamma/N$) de les avoir liées (resp. pas liées) dans le graphe.

Maintenant, on se pose la question de combien de composants connexes le graphe possède dans la limite de N grand. Plus précisément, on s'intéresse à la valeur $c = C/N$ où C est le nombre de composants connexes du graphe. Évidemment, pour $\gamma = N$ le graphe est complet ($c = 0$) et pour $\gamma = 0$ le graphe est complètement disconnexe ($c = 1$). Un théorème par Erdős et Renyi dit qu'il existe un certain γ_c tel que si $\gamma > \gamma_c$ (resp. $\gamma < \gamma_c$), $c = 0$ (resp. $c = 1$) avec probabilité 1. Dans cette section on s'intéresse à retrouver ces résultats par des méthodes de mécanique statistique.

Pour faire une analogie avec une fonction de partition on définit la densité de probabilité des valeurs de c par

$$\rho(c) = \sum_G P(G) \delta(c - c(G))$$

où $P(G)$ est la probabilité de tirer au sort un certain graphe et est donnée par

$$P(G) = p^{N_L(G)}(1-p)^{N(N-1)/2-L(G)}$$

où $L(G)$ est le nombre de liens du graphe G .

On définit maintenant la fonction génératrice de la variable aléatoire c

$$\begin{aligned} Y(q) &= \int_0^1 dc \rho(c) q^{Nc} \\ &= \int_0^1 dc \sum_G P(G) \delta(c - c(G)) \\ &= \sum_G p^{L(G)} (1-p)^{N(N-1)/2-L(G)} q^{C(G)} \end{aligned} \quad (1)$$

où la somme s'effectue sur tous les graphes possibles à N arrêts.

2.1 Le modèle de Potts

Maintenant, on s'intéresse à relier l'éq. 1 à un modèle physique connu, le modèle de Potts. Ce modèle est analogue au modèle d'Ising avec la différence que chaque spin porte une valeur $\sigma_i = \{0, 1, \dots, n\}$ au lieu de juste ± 1 . On dit que deux spins voisins interagissent si et seulement si ils portent la même valeur. Cela nous donne le hamiltonien

$$H_{\text{Potts}} = - \sum_{i < j} \delta_{\sigma_i, \sigma_j}$$

où on a assumé une itération ferromagnétique et nous nous avons restreint au cas de dimension infinie. Il est utile de noter qu'on retrouve le modèle d'Ising pour le cas $n = 2$.

Si on introduit la variable $v = e^\beta - 1$, on a

$$\begin{aligned} Z_{\text{Potts}} &= \sum_{\{\sigma_i\}} \prod_{i < j} [1 + v \delta_{\sigma_i, \sigma_j}] = \\ &= \sum_{\{\sigma_i\}} \left[1 + v \sum_{i < j} \delta_{\sigma_i, \sigma_j} + v^2 \sum_{i < j, k < l, (i,j) \neq (k,l)} \delta_{\sigma_i, \sigma_j} \delta_{\sigma_k, \sigma_l} + \dots \right] \end{aligned} \quad (2)$$

Chaque somme de cette expression peut être vu comme une somme sur tous les graphes possibles de k liens, où k est la puissance du terme en v qui accompagne la sommation. Cet interprétation nous permet de réécrire l'éq. 2 comme

$$Z_{\text{Potts}} = \sum_{\{\sigma_i\}} \sum_G \left[v^{L(G)} \prod_{(i,j) \text{ lien de } G} \delta_{\sigma_i, \sigma_j} \right]$$

où on somme sur tous les graphes G possibles à N arrêts. Si on échange l'ordre de sommation, on voit que, pour que le produit ne s'annule pas, il faut que tous les spins dans un même composant du graphe aient la même valeur. On a alors

$$Z_{\text{Potts}} = \sum_G v^{L(G)} n^{C(G)} = \sum_G \left(\frac{p}{1-p} \right)^{L(G)} n^{C(G)} = e^{N\gamma/2} Y(n)$$

où on retrouve la fonction génératrice de la distribution des c .

Le calcul de la énergie libre du modèle de Potts est analogue à la résolution du modèle d'Ising en champs moyen. La principale différence est qu'au lieu de écrire l'énergie en fonction de la magnétisation, on l'écrit en fonction de la fraction de spins en chaque une des n configurations possibles. Ce calcul donne comme résultat

$$Z_{\text{Potts}} = \int_0^1 \prod_{i=0}^{n-1} dx_i \delta \left(\sum_i x_i - 1 \right) e^{-Nf(x_1, \dots, x_n)}$$

où x_i est la fraction de spins dans l'état i et

$$f(x_1, \dots, x_n) = \sum_{i=0}^{n-1} \left[-\frac{\gamma}{2} x_i^2 + x_i \ln x_i \right]$$

on peut alors dire, par la méthode du col, que $Z_{\text{Potts}} = e^{-Nf_{\text{Potts}}}$, où $f_{\text{Potts}} = \min_{\{x_k\}} f(x_1, \dots, x_n)$, en imposant la restriction $\sum_i x_i = 1$.

On pourrait imaginer que comme il n'a pas de direction privilégié, il suffit de chercher une solution de la forme $x_1 = x_2 = \dots = x_n = x$. Par contre, pour certains valeurs de γ , cela ne correspond même pas à un minima local de f (il suffit d'en calculer l'Hessien pour en rendre compte). Cet effet de brisure de symétrie est très semblable à celui du modèle de Sherrington-Kirkpatrick. Heureusement, dans ce cas on peut retrouver le minimum d'une façon beaucoup plus simple que dans la solution de Parisi : il suffit de supposer que tous les x_i sont identiques sauf un.

Finalement, avec un peu de calcul [1], on retrouve le résultat déjà connu $\gamma_c = 1$.

3 Le problème de K-Satisfaisabilité aléatoire

Un des problèmes centraux de l'informatique théorique est de savoir si un système donné d'équations booléennes possède une solution. En particulier, quand chaque équation du système porte sur K variables, on appelle ce problème de K-Satisfaisabilité ou K-SAT.

Il est bien connu en informatique qu'une expression booléenne peut être écrite en n'utilisant que les opérateurs OR et NOT. On peut alors mettre n'importe quel système boolean sur la même forme que l'exemple suivant :

$$\begin{aligned} & (x_1) \text{ OR } (\text{NOT } x_5) \text{ OR } \dots \text{ (K-3 termes) } \dots \text{ OR } (x_{p_1}) = \text{TRUE}, \\ & (\text{NOT } x_3) \text{ OR } (x_9) \text{ OR } \dots \text{ (K-3 termes) } \dots \text{ OR } (x_{p_2}) = \text{TRUE}, \\ & \quad \quad \quad \vdots \quad \quad \quad \dots \quad \quad \quad \vdots \end{aligned}$$

On décide maintenant de raisonner en termes de spin, en définissant $S_i = 1$ (resp. $S_i = -1$) si x_i est TRUE (resp. FALSE). On définit C_{li} par

$$\begin{cases} C_{li} = 1 & \text{si } x_i \text{ apparaît dans la clause } l \text{ du système} \\ C_{li} = -1 & \text{si } x_i \text{ apparaît dans la clause } l \text{ du système précédé par un NOT} \\ C_{li} = 0 & \text{si } x_i \text{ n'apparaît pas dans la clause } l \text{ du système} \end{cases} \quad (3)$$

On peut dire alors que la solution du problème est donné par les $\{S_i\}$ qui annulent l'énergie

$$E[\mathbf{C}, \mathbf{S}] = \sum_{l=1}^M \delta \left(\sum_{i=1}^N C_{li} S_i + K \right) \quad (4)$$

où on dénoté le delta de Kronecker par $\delta(x) = \delta_{x,0}$ pour alléger les notations.

On s'intéresse dans la suite au modèle du K-SAT aléatoire, où les C_{i_l} sont choisis aléatoirement. On voudrais prouver que, comme pour les graphes aléatoires, si l'on pose $\alpha = \frac{M}{N}$, il existe un α_c critique qui sépare une phase où le système a presque sûrement une solution et une phase où il n'a presque sûrement pas de solution. Pour en faire, on calcule $\overline{\log Z}$ par la méthode des répliques. Alors,

$$\begin{aligned} \overline{Z(\mathbf{C})^n} &= \overline{\sum_{\mathbf{S}^1, \dots, \mathbf{S}^n} \exp\left(-\frac{1}{T} \sum_{a=1}^n E[\mathbf{C}, \mathbf{S}^a]\right)} = \sum_{\mathbf{S}^1, \dots, \mathbf{S}^n} \left[\exp\left(-\frac{1}{T} \sum_{a=1}^n \delta\left(\sum_{i=1}^N C_i S_i^a + K\right)\right) \right]^M \\ &= \sum_{\mathbf{S}^1, \dots, \mathbf{S}^n} \left\{ \frac{1}{2^K} \sum_{C_1, \dots, C_K = \pm 1} \frac{1}{N^K} \sum_{i_1, \dots, i_K = 1}^N \exp\left[-\frac{1}{T} \sum_{a=1}^n \prod_{l=1}^K \delta(S_{i_l}^a + C_l)\right] \right\}^M \end{aligned} \quad (5)$$

où pour écrire l'éq. 5 on a supposé que chaque $\{C_i\}$ est homogènement distribuées parmi les vecteurs qui ont K valeurs ± 1 et $N - K$ zéros. On peut voir que, dans la même façon que dans le modèle de Potts, l'énergie ne dépend que du numéro de sites dans chaque une des configurations $\{S^1, \dots, S^n\}$ possibles des répliques sur chaque site. On introduit alors les variables $x(\sigma)$ qui correspond à ce numéro divisé par N . Le vecteur σ est un vecteur à n composants qui détermine une certaine configuration.

Après un calcul analogue à ce des graphes aléatoires, on retrouve que $\overline{Z}^n = \exp(-N f_{\text{opt}}/T)$ où f_{opt} est le minimum par rapport aux 2^n possibles valeurs de $\{x(\sigma)\}$ du fonctionnel

$$f[x] = \alpha \ln \left[\sum_{\sigma_1, \dots, \sigma_K} x(\sigma_1) \dots x(\sigma_K) \exp\left(-\frac{1}{T} \sum_{a=1}^n \prod_{l=1}^K \delta(\sigma_l^a - 1)\right) \right] + \frac{1}{T} \sum_{\{\sigma\}} x(\sigma) \ln x(\sigma)$$

Si on fait l'hypothèse de symétrie des répliques, qui dans ce cas correspond à $x(\sigma) = x_0$, on ne retrouve un minimum de f que si α est plus petit qu'un certain $\alpha_c(K)$. Cela est analogue au calcul réalisé pour les graphe aléatoires, mais aussi au cas du modèle de Sherrington-Kirkpatrick, où à partir d'une certaine température, il y a une brisure spontané de cette symétrie. Alors, pour étudier le comportement du système quand $\alpha > \alpha_c$, il faut faire une hypothèse sur la forme des $x(\sigma)$ de et vérifier la cohérence du modèle résultant. Cela a été fait pour $K = 2$, où on a un accord avec la solution rigoureuse déjà connu de $\alpha_c = 1$ et pour $K = 3$, où il n'existe pas de preuve rigoureuse, mais les valeurs des simulations numériques montrent un bon accord avec les résultats trouvés.

4 Conclusion

Dans ce rapport, nous avons étudié deux problèmes d'informatique théorique : la connectivité des graphes aléatoires et le problème de K-SAT pour des clauses aléatoires. Dans les deux cas, nous avons pu observer la présence d'une transition de phase et calculer l'énergie libre associé au système.

Les cas étudiés montrent que la physique statistique peut être un outil intéressant pour analyser des problèmes d'informatique théorique, surtout quand le problème qu'on s'intéresse fait intervenir des grandeurs aléatoires et on s'intéresse à analyser son comportement typique.

Les deux exemples montrent aussi que pour ce type de problème on retombe souvent sur des équations semblables à celles retrouvés pour les systèmes désordonnés, ce que nous invite à utiliser des outils comme la méthode des répliques et le prolongement analytique.

Références

- [1] Olivier C. Martin, Rémi Monasson, and Riccardo Zecchina. Statistical mechanics methods and phase transitions in optimization problems. *Theor. Comput. Sci.*, 265(1-2) :3–67, 2001.