Algorithmique - Méthode probabiliste

1 Méthode probabiliste

La méthode probabiliste est un moyen de montrer l'existence d'objets par un simple résonnement de dénombrement. Si la probabilité d'obtenir un objet ayant certaine propriété est strictement positive, alors un tel objet existe. Par exemple, si la probabilité de tiré au sort un ticket gagnant est non nul, alors il y a au moins un ticket gagnant.

Cette méthode est non-constructiviste en générale, on montre juste l'existence sans vraiment donner l'objet. Mais le plus souvent, une preuve utilisant la méthode probabiliste est équivalente à un algorithme probabiliste permettant de construire cet objet.

2 Argument d'espérance

On utilisera le lemme suivant :

Lemme 1. *On considère un univers Ω et X une variable aléatoire sur Ω. Si $\mathbb{E}(X) = \mu$, alors $P(X \geqslant \mu) \geqslant 0$ et $P(X \leqslant \mu) > 0$.*

Ainsi, il existe au moins une réalisation de X supérieure ou égal à μ, et au moins une réalisation de X inférieure ou égal à μ.

2.1 Application : MAX-SAT

Théorème 1. *Étant donné un ensemble de m clauses, on note k_i le nombre de littéraux dans la $i$ème clause pour $1 \leqslant i \leqslant m$. Soit $k = \min_{1\leqslant i\leqslant m} k_i$. Il existe une distribution de valeurs de vérités qui satisfait au moins*

$$\sum_{i=1}^{m}(1-2^{-k_i}) \geqslant m(1-2^{-k})$$

clauses.

Démonstration. Pour chaque variable, on tire au sort sa valeur de vérité. En notant X_i pour $1 \leqslant i \leqslant m$ la variable aléatoire qui vaut 1 la $i$ème clause est satisfaite, 0 sinon, on a $P(X_i = 0) = 2^{-k_i}$, par indépendance des choix. En notant N le nombre de clause satisfaite, on a $N = \sum_{i=1}^{n} X_i$, et par linéarité de l'espérance :

$$\mathbb{E}(N) = \sum_{i=1}^{n}\mathbb{E}(X_i) = \sum_{i=1}^{m}(1-2^{-k_i})$$

Par le lemme ci-dessus, on peut conclure. □

Ainsi, la preuve ci-dessus peut être transformer en un algorithme qui assigne à chaque variable une valeur de vérité au hasard. Son temps d'exécution est un $\mathcal{O}(m)$, et il renvoie une distribution qui satisfait en espérance $m(1-2^{-k})$ clauses. Cet algorithme est donc une $1-2^{-k}$-approximation randomisé, ce qui généralise l'algorithme ci-dessus.

2.2 Application : Ensemble indépendant

Théorème 2. *Soit $G = (S, A)$ un graphe connexe à n sommets et $m \geqslant n/2$ arêtes. Alors G contient un ensemble indépendant ayant au moins $n^2/4m$ sommets.*

Démonstration. Soit $d = 2m/n \geqslant 1$ le degré moyen d'un sommet dans G. On considère l'algorithme probabiliste suivant :

Algorithm 1: EnsembleIndépendantRandomisé(G)

Supprimer chaque sommet de G (avec ses arêtes adjacentes) de manière indépendante avec une probabilité $1 - 1/d$;

Pour chaque arête restante, la supprimer avec un des sommets adjacent pris au hasard.

L'algorithme renvoie un ensemble indépendant puisque chaque arête a été supprimée.

Remarque : l'algorithme se passe en deux phases, une phase d'échantillonnage, et une phase de modification. On parle de méthode *Sample and Modify.*

Soit X le nombre de sommet qui survive à la première étape de l'algorithme. X suit une loi binomiale de paramètre n et $1/d$. Ainsi, $\mathbb{E}(X) = n/d$.

Soit Y le nombre d'arêtes qui survivent à la première étape. Chaque arête survit ssi ses deux sommets adjacents survivent, donc Y est la somme de m variables de Bernoulli de paramètre $(1/d)^2$, et donc

$$\mathbb{E}(Y) = m\left(\frac{1}{d}\right) = \frac{nd}{2d^2} = \frac{n}{2d}$$

La seconde étape retire toutes les arêtes restantes et au plus Y sommets. Ainsi, lorsque l'algorithme termine, le nombre de sommets restants est $X - Y$, et

$$\mathbb{E}(X - Y) = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d}$$

Ce qui conclut notre preuve. □