Analyse probabiliste d'algorithmes

Vallon

22 mars 2016

Problème : Pour un même problème P comment comparer l' "efficacité" de deux algorithmes A_1 et A_2 ?

Exemples :

1. Pour calculer l'image d'un nombre a donné par un polynôme P donné, dans quelle mesure peut on dire que la méthode de Horner est plus efficace que la méthode naïve ?
2. Pour trier une liste d'entiers donné dans l'ordre croissant, dans quelle mesure peut on dire que le tri par fusion est plus efficace que le tri par insertion ?

- On définit généralement la taille de l'objet traité par l'algorithme par un entier n. Par exemple le nombre de coefficients du polynôme , ou le nombre d'entiers de la liste
- On construit ensuite pour chaque algorithme une fonction temps d'exécution de l'algorithme $T(n)$
- Finalement on compare les fonctions temps d' exécution lorsque $n \to \infty$
- Que vaut $\lim\limits_{n \to +\infty} \dfrac{n^2}{n \ln(n)}$?
- En déduire que pour tout a et b réels strictement positifs il existe un rang N tel que pour tout $n \geqslant N$ on a $a \times n \ln(n) \leqslant b \times n^2$

- La durée d'un tri par insertion est "a peu de chose près" proportionnelle à n^2
- La durée d'un tri par fusion est "a peu de chose près" proportionnelle à $n \ln(n)$
- Il existera donc une taille N à partir de laquelle le tri par fusion sera plus rapide que le tri par insertion

- Est ce que la taille n est la seule grandeur à prendre en compte ?
- Comment prendre en compte toutes les configurations possibles pour une même taille n ?
- D'où l'idée d' apporter une analyse probabiliste à l'étude des algorithmes

Problème de l'embauche : Pour un poste une suite de candidats numérotés de 1 à n sont susceptibles d'être embauchés selon leur niveau de qualification $(r_1, r_2, ..., r_n)$ où r_i est un entier entre 1 et n, la qualification du candidat i . Plus r_i est élevé plus le candidat est qualifié. Les candidats sont testés et éventuellement embauchés dans l' ordre d'apparition selon cet algorithme

- meilleur = 0
- Pour i = 1 à n
- Tester candidat i
- Si candidat i plus qualifié que meilleur
- meilleur = i
- embaucher candidat i

Analyse

- Le coût de l'entretien est minime par rapport au coût de l'embauche
- Il s'agit donc d'évaluer le nombre de fois où l'instruction conditionnelle sera exécutée
- On ne peut le faire qu'en moyenne
- Soit N la variable aléatoire égale au nombre de candidats embauchés
- $N = N_1 + N_2 + ... + N_n$
- $N_i = 1$ si le candidat i a été embauché , 0 sinon
- $E(N) = E(N_1) + ... + E(N_n) = p_1 + p_2 + ... + p_n$
- $p_i = \frac{1}{i}$
- $E(N) = 1 + \frac{1}{2} + ... + \frac{1}{n}$ "du même ordre que " $\ln(n)$