UFR d'Informatique

Paris 7 – Paris Diderot

Année 2015–2016

Notes de cours d'algorithmique – M1

François Laroussinie

francois.laroussinie@liafa.univ-paris-diderot.fr

Page web du cours : http://www.liafa.univ-paris-diderot.fr/~francoisl/m1algo.html

1 Recherche du k-ème élément - Analyse en moyenne de quickselect

Pour la présentation des différents algorithmes, voir les transparents (fichier pdf en ligne).

Ici on ne traite que de la complexité en moyenne de l'algorithme *quickselect*. L'objectif est de trouver une justification « théorique » de ce bon comportement en pratique...

On note $\mathcal{C}(n,k)$ le coût moyen (en nombre de comparaisons) de la recherche du k-ème élément dans un tableau de taille n (ou plus exactement une zone de taille n) par le quickselect. On suppose que le *rang* du pivot est équiprobable, il peut prendre les valeurs 1, 2,...,n. Bien sûr si le rang du pivot est k, alors le pivot est l'élément recherché (et la recherche est alors terminée) ! On peut écrire :

$$\mathcal{C}(n,k) \;=\; \underbrace{\frac{1}{n}\Big(\sum_{i=1}^{k-1}\mathcal{C}(n-i,k-i) \;+\; \sum_{i=k+1}^{n}\mathcal{C}(i-1,k)\Big)}_{i \text{ correspond à la pos. du pivot.}} \;+\; \underbrace{n-1}_{\text{pivotage}}$$

La première somme correspond à la poursuite de la recherche dans la partie droite du tableau et la seconde à la recherche dans la partie gauche... La moyenne pour tous les k est alors : $\mathcal{C}(n) = \frac{1}{n}\sum_{k=1}^{n}\mathcal{C}(n,k)$.

D'où :

$$\mathcal{C}(n) \;=\; \frac{1}{n^2}\Big(\underbrace{\sum_{k=1}^{n}\sum_{i=1}^{k-1}\mathcal{C}(n-i,k-i)}_{A} \;+\; \underbrace{\sum_{k=1}^{n}\sum_{i=k+1}^{n}\mathcal{C}(i-1,k)}_{B}\Big) \;+\; n-1$$

Et on détaille (voir les figures "calcul de A" et "calcul de B") :

— A : $\sum_{k=1}^{n}\sum_{i=1}^{k-1}\mathcal{C}(n-i,k-i) = \sum_{i=1}^{n-1}\sum_{k=1}^{i}\mathcal{C}(i,k)$.

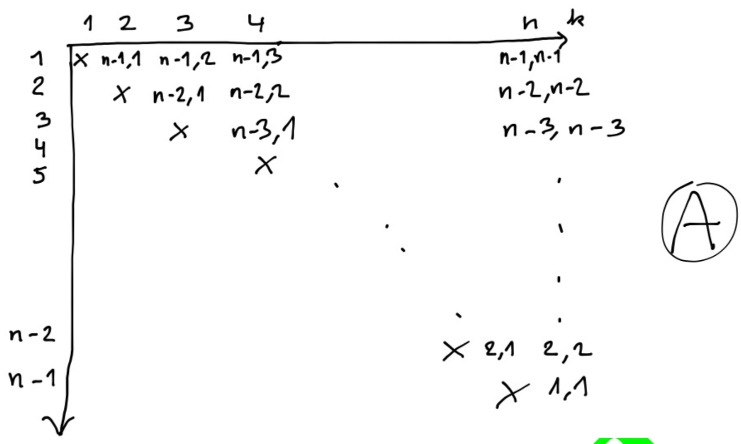

Figure 1 – Calcul de A

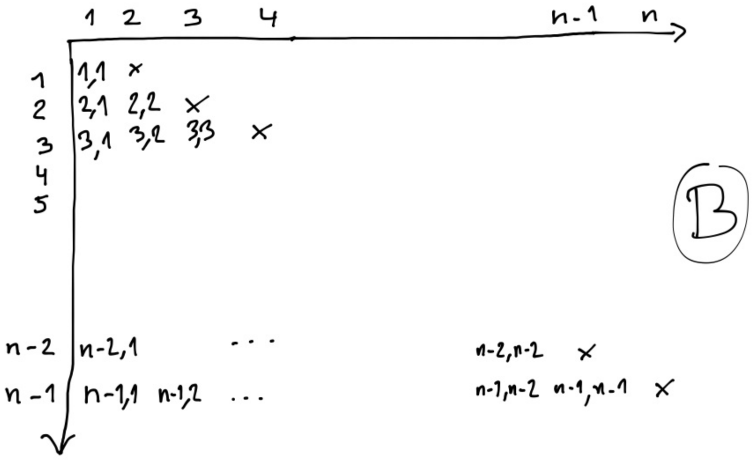

Figure 2 – Calcul de B

— $B : \sum_{k=1}^{n} \sum_{i=k+1}^{n} \mathcal{C}(i-1,k) = \sum_{k=1}^{n} \sum_{i=k}^{n-1} \mathcal{C}(i,k) = \sum_{i=1}^{n-1} \sum_{k=1}^{i} \mathcal{C}(i,k).$

On en déduit :

$$\begin{aligned}
\mathcal{C}(n) &= \frac{2}{n^2} \sum_{i=1}^{n-1} \Big(\sum_{k=1}^{i} \mathcal{C}(i,k) \Big) \quad + \quad n-1 \\
&= \frac{2}{n^2} \sum_{i=1}^{n-1} \Big(i \cdot \mathcal{C}(i) \Big) \quad + \quad n-1
\end{aligned} \tag{1}$$

On résout…

$$
\begin{array}{rcl}
n^2 \cdot \mathcal{C}(n) & = & 2\sum_{i=1}^{n-1} i \cdot \mathcal{C}(i) \quad + \; n^2(n-1) \\
(n-1)^2 \cdot \mathcal{C}(n-1) & = & 2\sum_{i=1}^{n-2} i \cdot \mathcal{C}(i) \quad + \; (n-1)^2(n-2) \\
\hline
n^2 \cdot \mathcal{C}(n) - (n-1)^2 \cdot \mathcal{C}(n-1) & = & 2(n-1)\mathcal{C}(n-1) + n^2(n-1) - (n-1)^2(n-2) \\
n^2 \cdot \mathcal{C}(n) & = & ((n-1)^2 + 2n - 2)\mathcal{C}(n-1) + n^3 - n^2 + 2n^2 - 4n + 2 - n^3 + 2n^2 - n \\
n^2 \cdot \mathcal{C}(n) & = & (n^2 - 2n + 1 + 2n - 2)\mathcal{C}(n-1) + 3n^2 - 5n + 2 \\
n^2 \cdot \mathcal{C}(n) & = & (n^2 - 1)\mathcal{C}(n-1) + 3n^2 - 5n + 2
\end{array}
$$

D'où on déduit :

$$\frac{n^2 \cdot \mathcal{C}(n)}{n(n+1)} = \frac{(n+1)(n-1)}{n(n+1)}\mathcal{C}(n-1) + \frac{3n^2 - 5n + 2}{n(n+1)}$$

Et donc :

$$\frac{n}{n+1} \cdot \mathcal{C}(n) = \frac{n-1}{n}\mathcal{C}(n-1) + \frac{3n^2 - 5n + 2}{n(n+1)}$$

On pose : $s_n = \frac{n}{n+1} \cdot \mathcal{C}(n)$, et on a : $s_n = s_{n-1} + \frac{3n^2-5n+2}{n(n+1)}$, d'où :

$$s_n = s_{n-1} + \frac{3n-5}{n+1} + \frac{2}{n} - \frac{2}{n+1}$$

Avec $s_1 = \frac{\mathcal{C}(1)}{2} = 0$. Mais on peut aussi prendre de manière équivalente (mais plus « régulière » pour des indices allant de 1 à n pour le calcul ci-dessous) la convention : $s_0 = 0$ (car il donne aussi $s_1 = 0$). Et finalement, pour $n \geq 1$, on a :

$$
\begin{aligned}
s_n &= \sum_{i=1}^{n} \underbrace{\frac{3i-5}{i+1}}_{3-\frac{8}{i+1}} + \frac{2}{i} - \frac{2}{i+1} \\
&= 3n - \underbrace{8\sum_{i=1}^{n} \frac{1}{i+1}}_{O(\ln n+1)} \quad + \quad \underbrace{2\sum_{i=1}^{n}(\frac{1}{i} - \frac{1}{i+1})}_{=2(1-\frac{1}{n+1})}
\end{aligned} \tag{2}
$$

Et on a donc : $\mathcal{C}(n) = O(n)$.