Option Informatique en Spé MP et MP*
Calcul de la médiane en temps linéaire

▶ Soit $\ell = (x_i)_{1 \leqslant i \leqslant n}$ une liste de $n \geqslant 1$ éléments distincts, pris dans un ensemble ordonné. Nous noterons $(\widehat{x_i})_{1 \leqslant i \leqslant n}$ la liste obtenue en triant la liste ℓ par ordre croissant : ses éléments sont donc ceux de ℓ, mais vérifient $\widehat{x_i} \leqslant \widehat{x_{i+1}}$ pour $1 \leqslant i < n$. Un indice $j \in [\![1, n]\!]$ sera dit *médian* vis-à-vis de cette liste si :

- $x_j = \widehat{x_{p+1}}$, si $n = 2p + 1$;
- $x_j = \widehat{x_p}$ ou $x_j = \widehat{x_{p+1}}$, si $n = 2p$.

Par exemple, l'indice 4 est médian pour la liste $(17, 2, 3, 8, 15)$; les indices 2 et 4 sont médians pour la liste $(17, 8, 1, 9)$.

▶ La détermination d'un indice médian pour une liste est un problème qui possède une solution de coût linéaire, et nous allons le prouver en construisant un algorithme $\mathcal{A}$ spécifié comme suit :

- $\mathcal{A}$ accepte en entrée un couple (E, k) où E est une liste d'éléments d'un ensemble ordonné, et $k \in [\![1, n]\!]$, avec $n = |E|$;
- $\mathcal{A}$ détermine le k-ième élément de $\widehat{E}$, pour un coût $C(n) = \mathcal{O}(n)$.

▶ Lorsque l'on demande de calculer un coût, l'unité est la comparaison de deux éléments d'une liste.

Question 1 • Expliquez comment trier une liste de sept éléments d'un ensemble ordonné, en effectuant au plus 21 comparaisons.

Question 2 • Soit $\mathcal{A}$ un algorithme capable de trier toute liste de sept éléments en effectuant des comparaisons. Justifiez l'affirmation suivante : il existe au moins une liste que $\mathcal{A}$ ne peut trier en moins de treize comparaisons.

▶ Vous admettrez qu'il existe un algorithme permettant de trier une liste de sept éléments en effectuant au plus 13 comparaisons.

Question 3 • Notons $14p + 7$ le plus petit multiple impair de 7, supérieur ou égal à n. Donnez une expression de p en fonction de n, en faisant intervenir la fonction $x \mapsto \lceil x \rceil$.

▶ La première étape de l'algorithme consiste à décomposer la liste de longueur n en $2p + 1$ listes de sept éléments (quitte à rajouter quelques éléments à la liste initiale), et à trier chacun des «paquets» de longueur 7.

Question 4 • Quel est le coût de cette étape ? Comment faut-il choisir les éléments à ajouter ?

▶ Dans la deuxième étape, on place dans une liste de longueur $2p+1$ la valeur médiane de chaque paquet, et on détermine la médiane m de cette liste.

Question 5 • Quel est le coût de cette étape ?

Question 6 • Avec les informations obtenues à ce stade, combien d'éléments sont certainement inférieurs à m ? Combien sont certainement supérieurs à m ? Pour vous aider à comprendre, faites un dessin, dans lequel vous placerez les paquets de sept en colonnes.

▶ Il reste un certain nombre d'éléments dont la place par rapport à m n'est pas connue. Nous voulons découper notre liste en deux listes : l'une, contenant les minorants stricts de m ; l'autre, contenant les majorants de m et m lui-même.

Question 7 • Combien de comparaisons faut-il effectuer pour réaliser cette étape ?

Question 8 • Donnez un encadrement du nombre r de minorants stricts de m, en fonction de p.

Question 9 • Expliquez comment se poursuit l'algorithme, dans chacun des cas suivants : $k \leqslant r$; $k = r + 1$; $k > r + 1$.

Question 10 • Pour $n > 1024$, donnez une formule exprimant $C(n)$, faisant intervenir des valeurs prises par C en un ou plusieurs entiers *strictement inférieurs* à n.

Question 11 ⋆⋆ • Prouvez la majoration $C(n) \leqslant 15n - 163$ pour $n > 32$. Vous admettrez (ou prouverez) que pour $32 < n \leqslant 1024$, on peut trier une liste de longueur n en effectuant au plus $10n$ comparaisons.

FIN