IFT436 – Algorithmes et structures de données

Université de Sherbrooke

Devoir 5

Enseignant:	Michael Blondin
Date de remise:	lundi 2 décembre 2024 à 23h59
À réaliser:	en équipe de trois ou moins
Modalités:	remettre en ligne sur Turnin dans un fichier PDF
Bonus:	les questions bonus sont indiquées par ★
Pointage:	max. 15 points + 1,5 points bonus

Le pivot idéal pour le tri rapide est la médiane. En théorie, on peut calculer la médiane d'une séquence en temps linéaire, mais cela s'avère peu efficace en pratique. Une autre approche consiste à choisir un élément à une distance raisonnable de la médiane, comme un élément compris entre le premier et troisième quartile.

Soit s une séquence de n éléments comparables distincts. Le *rang* d'un élément $x \in s$, dénoté $\text{rang}(x)$, correspond à la position de x dans s ordonnée en ordre croissant (en comptant à partir de 1). Nous disons qu'un élément $x \in s$ est *raisonnable* si

$$\lceil n/4 \rceil < \text{rang}(x) \leq \lfloor 3n/4 \rfloor.$$

Par exemple, considérons la séquence $s = [50, 13, 20, 67, 41, 89, 70, 30]$. Nous avons $\text{rang}(13) = 1$, $\text{rang}(67) = 6$ et $\text{rang}(89) = 8$. Les pivots raisonnables de s sont 30, 41, 50 et 67.

Afin d'identifier un pivot raisonnable, une approche probabiliste consiste à choisir k éléments aléatoirement (de façon uniforme), puis de retourner la médiane de ces k éléments, où $k \in \mathbb{N}$ est un paramètre *impair*:

Entrées : séquence s de $n \in \mathbb{N}_{\geq 3}$ éléments comparables distincts
Résultat : un élément raisonnable $x \in s$

`pseudomed`$_k(s)$ *:*
  $t \leftarrow [\,]$
  faire k **fois**
    choisir $i \in [1..n]$ aléatoirement de façon uniforme
    ajouter $s[i]$ **à** t
  trier t
  retourner $t[\lceil k/2 \rceil]$

Remarques:

— la constante k ne fait *pas* partie de l'entrée, il s'agit d'un paramètre fixe et toujours impair. Autrement dit, `pseudomed`$_1$, `pseudomed`$_3$, `pseudomed`$_5$, `pseudomed`$_7$, ... sont tous des algorithmes différents;

— le choix d'un nombre aléatoire est ici une opération élémentaire qui s'effectue *toujours* en temps constant.

Question 1.

(a) L'algorithme $\texttt{pseudomed}_k$ est-il de Las Vegas ou de Monte Carlo? Justifiez. 2 pts

(b) Quelle est la probabilité que $\texttt{pseudomed}_3$ retourne la médiane de $s = [42, 9000, 0]$? Justifiez. 4 pts

Indice: combien y a-t-il de séquences avec 2 éléments identiques, $3\ldots$

Pour les sous-questions suivantes, supposez que n est divisible par 4 (cela simplifie les calculs).

(c) Donnez la probabilité d'erreur de $\texttt{pseudomed}_1$ et de $\texttt{pseudomed}_3$, puis généralisez en donnant une expression symbolique qui décrit la probabilité d'erreur de $\texttt{pseudomed}_k$. Justifiez. 6 pts

Indice: considérez trois types d'éléments: raisonnables, déraisonnables de gauche et déraisonnables de droite; oubliez les nombres, pensez à des chaînes formées avec les caractères $\{\texttt{r}, \texttt{g}, \texttt{d}\}$, puis au coefficient binomial.

(d) La procédure ci-dessous exécute $\texttt{pseudomed}_3$ jusqu'à ce qu'un élément raisonnable soit identifié. Ainsi, sa valeur de retour est toujours correcte. Quel est son *temps espéré* sous notation $\mathcal{O}$? Justifiez. 3 pts

Entrées : séquence s de $n \in \mathbb{N}_{\geq 3}$ éléments comparables distincts
Résultat : un élément raisonnable $x \in s$

faire
  $x \leftarrow \texttt{pseudomed}_3(s)$
tant que x *n'est pas raisonnable* // déterminé avec algo. déterministe en temps $\mathcal{O}(n)$
retourner x

★ Montrez qu'il existe une valeur de k pour laquelle la probabilité d'erreur de $\texttt{pseudomed}_k$ est $\leq 1/2^{275}$. ★ 1,5 pts