IFT436 – Algorithmes et structures de données

Université de Sherbrooke

Devoir 6

Enseignant:	Michael Blondin
Date de remise:	jeudi 5 novembre 2019 à 10h30
À réaliser:	en équipe de deux ou individuellement
Modalités:	à remettre en classe, en copie imprimée ou copie manuscrite lisible
Bonus:	les questions bonus sont indiquées par ★
Pointage:	max. 20 points + 1 points bonus

Nous savons que 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, et soit s' la séquence obtenue en triant s en ordre croissant. Le *rang* d'un élément $x \in s$, dénoté $\text{rang}(x)$, correspond à la position de x dans s' (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 $\text{rang}(89) = 8$. De plus, 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 impaire:

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. Autrement dit, $\texttt{pseudomed}_1$, $\texttt{pseudomed}_3, \texttt{pseudomed}_5, \texttt{pseudomed}_7, \ldots$ sont tous des algorithmes différents;

— nous supposons que le choix d'un nombre aléatoire est 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. 3 pts

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. 7 pts

(d) Donnez un algorithme (déterministe) qui détermine en temps $\mathcal{O}(|s|)$ si un élément $x \in s$ est raisonnable. 4 pts

(e) 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 le *temps espéré* de cette procédure? Justifiez. 4 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 l'algo. de la sous-question (d)`
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 pt