

Compte-rendu minimal du mini-projet SDA : LCA et TH

Auteur : PEYREMORTE Romain

Groupe de TP : E

Contents

1 Exercice 1	1
1.1 Question 1.4	1
2 Évaluation expérimentale.	2
2.1 Performance comparée de LCA et TH	2
2.2 Qualité du générateur aléatoire	3
3 Principales difficultés rencontrées	3
4 Informations complémentaires	4
5 Bilan personnel	4

Consigne : Vous devez écrire vos réponse à la place des ... en laissant une ligne vide avant et deux après votre réponse.

Remarque : Ce document utilise le langage Markdown. On peut en engendrer une version PDF en faisant par exemple :

```
pandoc --toc -N -o LISEZ-MOI.pdf LISEZ-MOI.txt
```

1 Exercice 1

1.1 Question 1.4

Indiquer les inconvénients/avantages d'une implantation par listes chaînées d'une SDA.

Un inconvénient est qu'il faut parcourir la liste pour trouver un élément (et donc soit pour savoir s'il est présent, soit pour le modifier ou soit pour le supprimer) ce qui fait que certaines tâches atteignent une complexité égale à la taille de la liste, ce qui est gênant pour des grandes listes.

Les avantages sont la forme linéaire qui permet de mieux visualiser et mesurer la liste, ainsi que le fait que les clés et données peuvent être de tout types (pas besoin de comparer comme dans un arbre binaire).

2 Évaluation expérimentale.

2.1 Performance comparée de LCA et TH

Indiquer ici les résultats obtenus.

Borne	Taille	Temps LCA	Temps TH	Min	Max	Facteur
10	10	1ms	1ms	1	2	1e-1
10	100	1ms	2ms	4	19	1,5e-1
10	1000	7ms	16ms	88	119	3,1e-2
10	10000	51ms	135ms	949	1026	7,7e-3
10	100000	172ms	1s 16ms	9874	10129	2,52e-3
10	1000000	1s 722s	9s 796ms	99630	100461	8,31e-4
10	10000000	17s 100ms	1m 37s 764ms	998812	1002177	3,365e-4
100	10	1ms	>1ms	1	2	1e-1
100	100	9ms	6ms	1	5	4e-2
100	1000	131ms	15ms	4	20	1,6e-2
100	10000	1s 462ms	145ms	77	134	5,7e-3
100	100000	14s 166ms	1s 408ms	923	1089	1,66e-3
100	1000000	2m 21s 783ms	14s 90ms	9664	10260	5,96e-4
100	10000000	X	2m 20s 17ms	99093	100934	1,841e-4
1000	10	2ms	3ms	1	1	0
1000	100	37ms	7ms	1	2	1e-2
1000	1000	3s 139ms	43ms	1	6	5e-3
1000	10000	2m 2s 865ms	589ms	2	24	2,2e-3
1000	100000	24m 47 495ms	2s 157ms	919	1089	1,7e-3
1000	1000000	X	19s 869ms	9753	10348	5,95e-4
1000	10000000	X	13m 13s 846ms	9688	10277	5,89e-5
10000	10	1ms	>1ms	1	1	0
10000	100	17ms	3ms	1	1	0
10000	1000	19s 713ms	54ms	1	3	2e-2
10000	10000	X	2s 668ms	1	6	5e-4
10000	100000	X	53s 804ms	1	24	2,3e-4
10000	1000000	X	9m 32s 946ms	65	139	7,4e-5
10000	10000000	X	96m 18s 515ms	884	1135	2,51e-5
100000	10	1ms	1ms	1	1	0
100000	100	20ms	>1ms	1	1	0
100000	1000	21s 593ms	58ms	1	2	1e-3
100000	10000	X	3s 468ms	1	5	4e-4
100000	100000	X	8m 5s 129ms	1	8	7e-5
100000	1000000	X	X	X	X	X

Borne	Taille	Temps LCA	Temps TH	Min	Max	Facteur
100000	10000000	X	X	X	X	X

Les tests marqués d'un X sont des tests non réalisés par manque de temps car ces derniers prenaient beaucoup de temps (plus d'une heure). On peut remarquer que le LCA est plus rapide que le TH que pour la borne à 10, ce qui s'explique par le hachage qui est à 1000 pour le TH.

Les tests n'ont pas été réalisés pour des tailles supérieures ou des bornes supérieures car l'exécution était longue (à distance, plus de 45 min sans réponse donc fenêtre fermée) et par manque de temps. Le Sda LCA semble plus adapté à des faibles bornes et des faibles tailles, surtout lorsque la borne est inférieure au facteur de hachage de TH. Cependant, le Sda TH est plus adapté et fonctionne mieux pour des bornes importantes et des tailles importantes. Plus ce facteur est élevé, plus la différence entre le min et le max est importante par rapport à l'échantillon, ce qui est un mauvais indicateur du générateur. En effet, en divisant la différence par la taille de l'échantillon, on peut mieux juger cette dernière (une différence de 1000 est moins gênante pour un échantillon de 10000000 qu'une différence de 1 pour un échantillon de 10). On en déduit donc de ce facteur que le générateur aléatoire est bon pour de grands échantillons ainsi que pour de grandes bornes, mais moins par rapport à des échantillons faibles dans des bornes faibles (borne = 10 et taille = 10).

2.2 Qualité du générateur aléatoire

Indiquer les conclusions quant à la qualité du générateur aléatoire.

On n'obtient pas de fréquence nulle avec ce programme car on calcule la fréquence des nombres de l'échantillon aléatoire (et pas de l'ensemble des nombres de l'intervalle). On observe que pour des très faibles tailles d'échantillons par rapport à la borne (taille/borne ≤ 0.1), on obtient une différence soit nulle soit importante (un rapport de 2 ou plus entre min et max), cela est dû au fait que les fréquences sont faibles, de l'ordre de l'unité, souvent égales à 1. Cependant, pour mieux observer la qualité du générateur, j'ai calculé un facteur qui correspond à la fréquence maximale moins la fréquence minimale, le tout divisé par le nombre d'échantillon.

3 Principales difficultés rencontrées

Indiquer ici les principales difficultés rencontrées lors de la réalisation de ce projet et comment elles ont été surmontées ou contournées.

Les principales difficultés rencontrées sont sur la genericité et sa définition (tous les détails à ne pas oublier), sur l'utilisation d'un paramètre générique dans l'interface d'un programme générique (T_LCA dans th.ads). Ces difficultés ont

été surmontées en demandant de l'aide et des explications à des camarades ayant mieux compris (étant déjà passer par là, ils ont pu plus rapidement détecter l'erreur, le détail qui manquait).

4 Informations complémentaires

Indiquer ici les informations qui pourraient aider à la compréhension du travail réalisé.

Pour l'exercice 3, de ce que j'ai discuté avec des amis, j'ai eu une interprétation différente de la phrase "Le principe du programme sera de mesurer la fréquence d'apparition des nombres d'un échantillon de nombres tirés aléatoirement dans un intervalle 1..Borne et d'afficher la plus petite fréquence et la plus grande fréquence". J'ai interprété qu'il fallait calculer la fréquence d'apparition des nombres de l'échantillon, donc je n'ai aucune fréquence nulle (car tous les nombres de l'échantillon sont dans l'échantillon), de ce fait, $\min > 0$.

5 Bilan personnel

Quel bilan personnel tirez vous de ce mini-projet ?

Le mini-projet était intéressant mais beaucoup trop gros : nous n'avions qu'une seule séance de TP pour bien communiquer avec les professeurs (les mails n'étant pas optimum pour expliquer le problème et en trouver la source). Je suis d'accord que pour le projet final, le nombre de séance de TP est insuffisant pour que travailler durant ces dernières, cependant, un minimum de 2 séances permet de commencer à la première, de bloquer entre les deux, et de pouvoir poser sa question en physique à la deuxième. De plus, le mini-projet abordait certains détail (généricité dans généralité) seulement aborder rapidement en cours pour ce qui est de la forme ada à employer, et non en TP, ce qui est à l'origine des principales difficultés que j'ai rencontrées.