TP: Cartographie d'arbres

MP2I Lycée Pierre de Fermat

Ce TP reprend le format des TP d'informatique du concours des ENS tel qu'ils existaient avant l'apparition des classes MP2I/MPI. Les sujets comportent des questions de programmation qui demandent de coder une fonction et de donner sa valeur sur des entrées bien précises. Pour que chaque candidat.e ait des réponses différentes, et empêcher la triche, on leur attribue au début de l'épreuve une graine qui servira à "générer de l'aléatoire". Ainsi, la première question demande presque toujours de calculer les termes d'une suite dont le premier terme est la graine donnée. De plus, pour pouvoir vérifier vos réponses au fur et à mesure, le sujet donne une fiche réponse pour une valeur de graine spécifique. Pour ce TP, utilisez cette valeur (qui se trouve en fin de sujet sur la fiche réponse) de graine.

Aux questions de programmation s'ajoutent des questions écrites, demandant d'expliquer un algorithme, d'étudier une complexité, etc... Aux vraies épreuves, ces questions sont ensuites développées pendant un entretien oral après la partie programmation. Pour ce TP, rédigez-les sur papier, pour vous entraîner.

Je vous conseille de lire l'intégralité du sujet avant de vous lancer dedans, pour voir de quelles structures vous aurez besoin. Pensez aussi à tester vos fonctions sur des petits exemples que vous pouvez résoudre à la main pour vous aider à débugger.

1 Introduction

Dans ce sujet, on va chercher des moyens de se déplacer sur des arbres binaires, et d'optimiser leur stockage sur des supports de taille limitée.
On considère la suite d'entiers (u_k) définie pour $k \geqslant 0$ par :

$$u_k = \begin{cases} \textit{votre } \mathsf{u}_0 \textit{ (à reporter sur votre fiche)} & \text{si } k = 0 \\ 15\,091 \times u_{k-1} \mod 64\,007 & \text{si } k \geqslant 1 \end{cases}$$

Question 1 *Que valent :* **a)** u_{10} **b)** u_{100} **c)** u_{1000}

On définit l'arbre binaire A_n de la façon suivante :
- A_0 est l'arbre réduit à un seul nœud, appelé <u>feuille</u>.
- Pour $n > 0$, A_n est un nœud muni de deux sous-arbres :
  - son fils gauche est l'arbre $A_{u_{2n} \mod n}$,
  - son fils droit est l'arbre $A_{u_{2n+1} \mod n}$.

On appelle <u>père</u> d'un nœud n_i dans l'arbre le nœud n_j tel que n_i est un fils de n_j, et <u>racine</u> de l'arbre le seul nœud qui ne possède pas de père.

Question 2 *Donnez le nombre de nœuds et de feuilles des arbres suivants :*
a) A_{10} **b)** A_{100} **c)** A_{1000}

2 Schéma d'adressage des arbres

On veut pouvoir trouver son chemin facilement dans les arbres, c'est-à-dire savoir se déplacer d'un nœud n_i vers un nœud n_j. Pour cela, on munit les nœuds d'une adresse qui permet de savoir facilement comment rejoindre un nœud, ou que l'on soit dans l'arbre.
L'adresse d'un nœud n_i est un couple d'entiers (l_i, m_i) tels que :
- l_i est le numéro de n_i dans un parcours en profondeur préfixe, en supposant que la racine porte le numéro 0.
- m_i est la plus grande valeur de l_j pour les nœuds n_j du sous-arbre enraciné en n_i.

Dans la suite du sujet, on repérera un nœud n_i de l'arbre par son indice l_i.

Question 3 *Pour les arbres et les valeurs de l_i suivantes, on donnera les valeurs de m_i telles que (l_i, m_i) est l'adresse d'un nœud n_i de l'arbre :*
a) A_{10}, $l_i = 3$ **b)** A_{100}, $l_i = 9$ **c)** A_{1000}, $l_i = 30$.

Question à développer pendant l'oral : Expliquez comment trouver son chemin d'un nœud n_i à un nœud n_j dans l'arbre grâce à ces adresses.

Question 4 *Donnez la distance parcourue en nombre de nœuds traversés lorsqu'on se déplace sur les arbres suivants du nœud n_i au nœud n_j donnés par leurs valeurs l_i et l_j (si $n_i = n_j$, on considèrera que cette distance vaut zéro) :*
a) A_{10}, $l_i = 3$ *et* $l_j = 9$ **b)** A_{100}, $l_i = 5$ *et* $l_j = 30$ **c)** A_{1000}, $l_i = 30$ *et* $l_j = 90$

3 Cartographie d'arbres

On veut créer une représentation cartographique de l'arbre sur lequel on se déplace. La taille d'une carte est limitée, il faut donc plusieurs cartes pour couvrir tout l'arbre. On suppose qu'une carte peut contenir jusqu'à p nœuds de l'arbre. Lorsqu'on traverse cet arbre, de la racine jusqu'à une feuille, on a besoin d'emporter toutes les cartes nécessaires, de telle sorte que le chemin de la racine jusqu'à la feuille soit couvert. On appelle cartographie de taille k une partition des nœuds de l'arbre en k sous-ensembles disjoints contenant chacun au plus p nœuds.

On appelle coût d'un nœud pour une cartographie donnée le nombre des cartes qui couvrent le chemin depuis la racine de l'arbre jusqu'à ce nœud. On note $c(n_i)$ le coût d'un nœud n_i. On va surtout s'intéresser ici au coût des feuilles de l'abre. On note F_A l'ensemble des feuilles de l'arbre A. Pour une cartographie donnée, on considère le coût moyen $C^{\text{moy}}(A)$ des feuilles de l'arbre, ainsi que le coût maximum $C^{\max}(A)$ des feuilles de l'arbre :

$$C^{\max}(A) = \max_{f \in F_A} c(A), \qquad C^{\text{moy}}(A) = \frac{1}{|F_A|} \sum_{f \in F_A} c(A).$$

Question 5 *On propose une méthode de cartographie simple : pour $k = 0..\lceil n/p \rceil$, la carte k contient les nœuds d'indice l_j avec $p \times k \leqslant l_j < p \times (k+1)$. Pour les arbres et les valeurs de p suivantes, donnez le coût moyen des feuilles C^{moy} et le coût maximum des feuilles $C^{\max}$:*

a) A_{10}, $p = 3$ **b)** A_{100}, $p = 6$ **c)** A_{1000}, $p = 9$

Question à développer pendant l'oral : Quelle est la complexité de votre algorithme ?

3.1 Minimisation du coût maximum

Pour un arbre A donné, on cherche maintenant à obtenir une cartographie telle que le coût maximum des feuilles $C^{\max}(A)$ est minimal. Dans une telle cartographie, on note $p(A)$ la taille minimale d'une carte contenant la racine de A.

Question à développer pendant l'oral : Donnez une expression de $C^{\max}(A)$ et $p(A)$ en fonction de $C^{\max}(A_g)$, $C^{\max}(A_d)$, $p(A_g)$ et $p(A_d)$ où A_g et A_d sont les sous-arbres gauche et droite de l'arbre.

Question 6 *Pour les arbres et valeurs de p suivants, donnez le coût maximum des feuilles $C^{\max}(A)$ ainsi que la taille minimale de la carte contenant la racine dans une cartographie minimisant $C^{\max}(A)$:*

a) A_{10}, $p = 3$ **b)** A_{100}, $p = 6$ **c)** A_{1000}, $p = 9$

Question à développer pendant l'oral : Comment limiter le nombre de cartes tout en gardant une cartographie qui minimise le coût maximum des feuilles ? On pourra donner une borne par rapport au nombre minimal de cartes possible.

3.2 Minimisation du coût moyen

On cherche ici à minimiser le coût moyen $C^{\mathrm{moy}}(A)$ des feuilles de A dans une cartographie. On suppose ici que dans une cartographie, les cartes sont des parties convexes de l'arbre. Ainsi, si deux nœuds n_i et n_j sont dans une même carte, tous les nœuds sur le chemin de n_i à n_j sont aussi dans cette carte.
Question à développer pendant l'oral : Montrez qu'il existe une cartographie minimisant le coût moyen des feuilles telle que la carte contenant la racine possède exactement p nœuds dès que l'arbre contient plus de p nœuds.
Soit $C(i, k)$ la valeur minimale de la somme des coûts des feuilles du sous-arbre enraciné en n_i lorsque la carte contenant n_i possède au plus k nœuds. Donnez une expression de $C(i, k)$ en fonction des $C(j, k')$ pour les fils n_j de n_i, et $k' = 1, \ldots, p$. On pourra noter $f(n_i)$ le nombre de feuilles du sous-arbre enraciné en n_i

Question 7 *Quel est le coût moyen des feuilles $C^{\mathrm{moy}}(A)$ minimal pour les arbres et valeurs de p suivants ?*
a) A_{10}, $p = 3$ **b)** A_{100}, $p = 6$ **c)** A_{1000}, $p = 9$

Question à développer pendant l'oral : Donnez la complexité en temps et en espace de votre algorithme. Peut-on les améliorer ?

4 Minimisation du coût maximum et du coût moyen

On cherche désormais à optimiser à la fois le coût maximum des feuilles et leur coût moyen.
Question à développer pendant l'oral : Expliquez comment calculer, pour un arbre et une valeur de p donnés, le plus petit coût moyen des feuilles dans une cartographie quand on impose que le coût maximum des feuilles est inférieur ou égal à une valeur V donnée. Quelle est la complexité en temps et en espace de votre algorithme ?

Pour un arbre A est une valeur de p donnés, on note $C^1(A, p)$ le plus petit coût moyen des feuilles dans une cartographie, et $C^2(A, p)$ le plus petit coût moyen des feuilles dans une cartographie où le coût maximum des feuilles est minimal.

Question 8 *Pour les arbres suivants, donnez la plus petite valeur de $p \geqslant 3$ telle que $C^1(A, p) \neq C^2(A, p)$, ainsi que les valeurs de $C^1(A, p)$ et $C^2(A, p)$ correspondantes.*
a) A_{500} **b)** A_{1000} **c)** A_{3000}

Fiche réponse type: Cartographie d'arbres

$\widetilde{u_0}$: 7

Question 1

a) 13238

b) 61807

c) 28695

Question 2

a) 21, 11

b) 107, 54

c) 2459, 1230

Question 3

a) 5

b) 11

c) 40

Question 4

a) 3

b) 4

c) 8

Question 5

a) 2.727, 4

b) 3.963, 6

c) 7.293, 15

Question 6

a) 3, 1

b) 3, 6

c) 5, 3

Question 7

a) 2.545

b) 2.815

c) 3.853

Question 8

a) 8, 3.511, 3.517

b) 24, 2.819, 2.820

c) 22, 2.895, 2.897