Marches, permutations et arbres binaires aléatoires

Épreuve pratique d'algorithmique et de programmation
Concours commun des Écoles Normales Supérieures

Durée de l'épreuve: 4 heures – Cœfficient: 4

Juillet 2003

N'oubliez en aucun cas de recopier votre $\mathbf{u_0}$ clairement encadré au début de votre copie.

Notes. Lorsque la description d'un algorithme est demandée, celle-ci doit être courte et précise. Vous ne devez en aucun cas recopier le code de vos procédures sur votre copie !
Quand on demande le temps de calcul d'un algorithme en fonction d'un paramètre n, on demande son ordre de grandeur en fonction du paramètre, par exemple : $O(n^2)$, $O(n \log n)$,...
Il est recommandé de commencer par lancer vos programmes sur des petites valeurs des paramètres.

Les questions sont assez largement indépendantes entre elles, et ne dépendent principalement que des algorithmes de génération des structures étudiées. Les sections 2, 3 et 5 sont totalement indépendantes des précédentes.

1 Préambule

La suite $(u_n)_{n\geqslant 0}$ suivante est dite « pseudo-aléatoire » car ses valeurs sont raisonnablement uniformes et indépendantes. Cette suite sera utilisée dans les sections suivantes pour construire nos structures « aléatoires ».

- u_0 est donné sur votre table. **À reporter en haut de votre copie !!!**
- $u_n = (16\,383 \times u_{n-1}) \text{ modulo } 59\,047, \quad \text{pour } n \geqslant 1.$

Question 1 *Que valent u_{996} et $u_{9\,547}$? Combien il y a-t-il d'indices i, $0 \leqslant i < 100\,000$, tels que $u_i = 18 \text{ modulo } 27$? Combien il y a-t-il d'indices i, $1 \leqslant i < 100\,000$, tels que $u_i = 18 \text{ modulo } 27$ et $u_{i-1} = 15 \text{ modulo } 27$?*

2 Marches aléatoires

On appelle *marche aléatoire* y de longueur n, une suite de $n+1$ entiers $(y(i))_{0 \leqslant i \leqslant n}$, telle que :

$$y(0) = 0 \quad \text{et} \quad |y(i+1) - y(i)| = 1, \quad \text{pour tout } 0 \leqslant i < n.$$

Nous considérons des marches aléatoires de longueur $n \leqslant 10\,000$. À toute marche aléatoire y, on associe la marche positive $|y|$, telle que $|y|(i) = |y(i)|$. On appelle *record* d'une marche aléatoire y tout indice $i^* > 0$, tel que $y(i) < y(i^*)$ pour tout $i < i^*$.
On se donne un paramètre entier $b \geqslant 1$. On définit la marche $(y_b(i))_{0 \leqslant i \leqslant 10\,000}$ suivante :

$$\begin{aligned} &y_b(0) = 0 \\ \text{et pour } 0 \leqslant i < 10\,000 : \quad &y_b(i+1) = y_b(i) - 1 + 2 \times (u_{b \times i + 10} \text{ modulo } 2). \end{aligned}$$

Question 2 *Combien la marche $|y_1|$ a-t-elle de records ? Quel est l'avant-dernier record de $|y_1|$? Quel est la distance maximale entre deux records consécutifs ? Mêmes questions pour $|y_3|$, et $|y_5|$. Décrivez succinctement (rapidement) vos algorithmes et donnez leurs temps de calcul en fonction de n.*

On appelle *sous-suite strictement croissante* de longueur ℓ d'une marche $(y(i))_{0 \leqslant i \leqslant n}$ toute suite d'indices $0 \leqslant i_1 < i_2 < \ldots < i_\ell \leqslant n$ telle que $y(i_1) < \ldots < y(i_\ell)$.

Question 3 *Quelles sont les longueurs des plus longues sous-suites strictement croissantes de $|y_1|$, de $|y_3|$ et de $|y_5|$? Expliquez rapidement la méthode que vous avez utilisée et donnez son temps de calcul en fonction de n.*

Question 4 *Quelles sont les longueurs des plus longues sous-suites strictement croissantes de y_1, de y_3 et de y_5 ? Détaillez et justifiez l'algorithme que vous avez utilisé et donnez son temps de calcul en fonction de n.*

3 Permutations aléatoires

On se propose de générer des permutations aléatoires sur $\{0, \ldots, n-1\}$, où $1 \leqslant n \leqslant 10\,000$.

3.1 Premier algorithme de génération

On suppose ici que $n \leqslant 100$. On se donne un paramètre entier $b \geqslant 1$. Voici l'algorithme proposé : soit $x_i = u_{b \times i + 10}$ modulo n^2 pour $0 \leqslant i < n$; s'il existe $i \neq j$ tels que $x_i = x_j$, alors renvoyer un message d'erreur ; sinon, on note $\pi_{n,b}$ l'unique permutation telle que $x_{\pi_{n,b}(0)} < \ldots < x_{\pi_{n,b}(n-1)}$.

Question 5 *Quel est le plus petit* $b \geqslant 1$*, noté* b^*_{50}*, tel que la procédure renvoie une permutation pour* $n = 50$ *? Que vaut* $\pi_{50,b^*_{50}}(37)$ *? Quel est le nombre de points fixes de* $\pi_{50,b^*_{50}} \circ \pi_{50,b^*_{50}}$ *? Mêmes questions pour* $n = 100$*.*

Question 6 *Décrivez rapidement les algorithmes et les structures de données que vous avez utilisés. Donnez pour chaque algorithme son temps de calcul en fonction de* n*.*

Nous rappelons que l'orbite d'un entier i sous l'action d'une permutation π est l'ensemble de ses itérés $\{i, \pi(i), \pi \circ \pi(i), \pi \circ \pi \circ \pi(i), \ldots\}$. L'orbite d'un élément sous l'action d'une permutation est un cycle.

Question 7 *Quelle est la longueur de la plus longue orbite de* $\pi_{50,b^*_{50}}$ *? de* $\pi_{100,b^*_{100}}$ *? Donnez le plus petit élément de ces deux orbites. Décrivez rapidement les algorithmes que vous avez utilisés, et donnez le nombre d'opérations élémentaires effectuées en fonction de* n*.*

3.2 Second algorithme de génération

On suppose maintenant $1 \leqslant n \leqslant 10\,000$. On se donne un paramètre entier $b \geqslant 0$. Nous construisons la permutation $\sigma_{n,b}$ sur $\{0, \ldots, n-1\}$, pour $1 \leqslant n \leqslant 10\,000$, ainsi. Partant de la permutation identité, on itère le processus suivant pour i allant de 1 à $n-1$, dans cet ordre : échanger dans la permutation courante les images des entiers i et (u_{i+b} modulo $(i+1)$). Puis, renvoyer la permutation obtenue.

Question 8 *Quelle est la permutation* $\sigma_{5,2}$ *?*

Question 9 *Quelle est la longueur de la plus grande orbite de* $\sigma_{9\,537,0}$ *? Donnez un élément de cette orbite. Mêmes questions pour* $\sigma_{9\,564,1}$ *et* $\sigma_{9\,337,12}$*.*

Question 10 *Quelle est la longueur de la plus longue sous-suite strictement croissante de la suite* $(\sigma_{100,0}(i))_{0 \leqslant i \leqslant 99}$ *? Donnez la liste des éléments d'une telle sous-suite maximum.*
Détaillez précisément et justifiez de façon concise l'algorithme que vous avez utilisé. Donnez son temps de calcul en fonction de la longueur n *de la suite.*

Question 11 *Quelles sont les longueurs des plus longues sous-suites strictement croissantes des suites* $(\sigma_{200,0}(i))_{0 \leqslant i \leqslant 199}$*,* $(\sigma_{500,0}(i))_{i=0 \leqslant i \leqslant 499}$ *et* $(\sigma_{1\,000,0}(i))_{0 \leqslant i \leqslant 999}$ *? Donnez les 5 premiers éléments d'une telle sous-suite maximum pour chacune.*

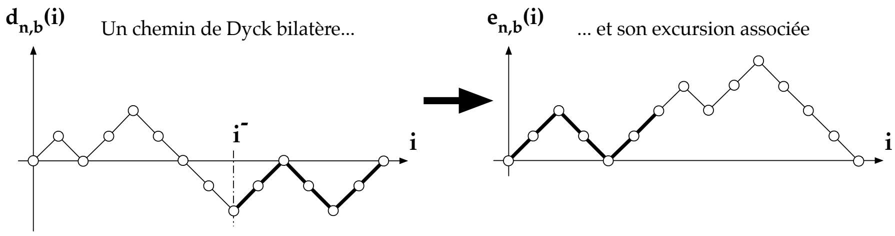
4 Excursions

On appelle *chemin de Dyck bilatère*, une marche aléatoire d de longueur $2n$ telle que $d(2n) = 0$, c'est-à-dire qui compte autant de pas montants que de pas descendants.
On se donne un paramètre entier $b \geqslant 0$. On propose de générer un chemin de Dyck bilatère aléatoire $d_{n,b}$ de longueur $2n$ ainsi : générer la permutation $\sigma_{2n,b}$; les positions des pas montants de $d_{n,b}$ sont $\sigma_{2n,b}(0), \ldots, \sigma_{2n,b}(n-1)$; et celles des pas descendants de $d_{n,b}$ sont $\sigma_{2n,b}(n), \ldots, \sigma_{2n,b}(2n-1)$.

Question 12 *Donnez* $d_{5,2}$.

Question 13 *Quelle est la valeur de* $d_{3\,270,0}(1\,557)$ *? de* $d_{4\,351,3}(1\,346)$ *? Quelle est la distance maximale entre deux records consécutifs de* $d_{3\,270,0}$ *? de* $d_{4\,351,3}$ *?*

On appelle *excursion*, un chemin de Dyck bilatère positif, i.e., une marche aléatoire e, telle que : $e(2n) = 0$, et $e(i) \geqslant 0$ pour tout $0 \leqslant i \leqslant 2n$.
On se donne un paramètre entier $b \geqslant 0$. Nous définissons l'excursion $e_{n,b}$ de la façon suivante : soit $0 \leqslant i^- < 2n$, le plus petit indice qui correspond au minimum de $d_{n,b}$ (i.e., $d_{n,b}(i) > d_{n,b}(i^-)$ pour $i < i^-$, et $d_{n,b}(i) \geqslant d_{n,b}(i^-)$ pour $i \geqslant i^-$) ; l'excursion $e_{n,b}$ est obtenue en « coupant » le chemin de Dyck bilatère $d_{n,b}$ au point i^-, et en échangeant les deux moitiés de chemins obtenus. La construction de $e_{n,b}$ est illustrée sur la figure ci-dessous.



Question 14 *Donnez* $e_{5,2}$.

Question 15 *Quelles sont les valeurs de* $e_{2\,051,0}(1\,342)$*,* $e_{2\,057,23}(1\,546)$ *? Quelle est la distance maximale entre deux records consécutifs pour ces deux excursions ?*

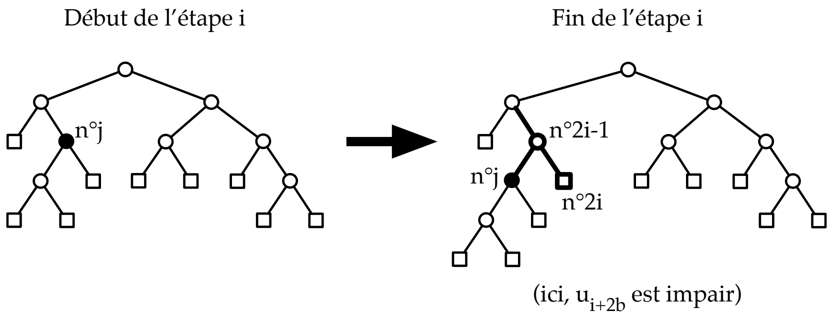
5 Génération incrémentale d'arbres binaires

Nous supposons $n \leqslant 5\,000$. Nous proposons dans cette partie de générer des arbres binaires. Un *arbre binaire* est un arbre dont les sommets ont 0 ou 2 fils. Les sommets ayant zéro fils sont appelés *feuilles* et les sommets ayant 2 fils, *sommets internes*. On démontre, par récurrence immédiate, que tout arbre binaire avec n sommets internes, a exactement $n+1$ feuilles, et compte donc $2n+1$ sommets exactement.

À défaut de liste chaînée, un arbre binaire pourra, par exemple, être codé par trois tableaux d'entiers `pere`, `fils_g`, et `fils_d`, et une variable `racine`. On se donne un paramètre entier $b \geqslant 0$. On se propose de construire l'arbre binaire $A_{n,b}$ à $2n+1$ sommets, de la façon suivante. Les sommets sont numérotés à partir de 0, par ordre de création. On commence avec un arbre réduit à un sommet (numéroté 0). Puis, pour i allant de 1 à n :

– on sélectionne le sommet n°j, avec $j = u_{i+b}$ modulo $(2i-1)$;
– on crée deux nouveaux sommets numérotés $2i-1$ et $2i$;
– le sommet n°$2i-1$ prend la place du sommet n°j,
– et le sommet n°j devient le fils droit (resp. gauche) du sommet n°$2i-1$, si u_{i+2b} est pair (resp. impair) ;
– le sommet n°$2i$ est une feuille et c'est l'autre fils du sommet n°$2i-1$;
– les relations entre les autres sommets restent inchangées.

La figure ci-dessous illustre une étape de la génération. Les sommets internes y sont représentés par des cercles, et les feuilles par des carrés. Les nouveaux sommets et nouvelles arêtes sont représentés en gras.



Question 16 *Donnez l'arbre $A_{4,1}$ (qui compte 9 sommets).*

La *hauteur* $h(x)$ d'un sommet x est définie récursivement par : $h(x) = 1$, si x est une feuille ; et $h(x) = 1 + \max(h(fg), h(fd))$ sinon, où *fg* et *fd* désignent respectivement les fils gauche et droit de x. La *profondeur* $p(x)$ d'un sommet x est sa distance à la racine (en nombre d'arêtes).

Question 17 *Quels sont le père, et les fils gauche et droit du sommet numéroté 57 dans l'arbre $A_{130,2}$? Quelles sont sa hauteur et sa profondeur ? Même question pour le sommet numéroté 1 430 dans l'arbre $A_{2\,450,3}$.*

Question 18 *Calculez pour l'arbre $A_{250,3}$:*

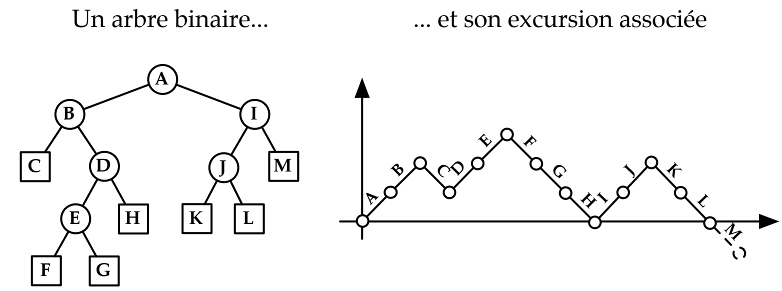
$$H_{250} = \sum_{i=0}^{500} h(i) \quad et \quad P_{250} = \sum_{i=0}^{500} p(i).$$

Calculez pour l'arbre $A_{2\,435,1}$:

$$H_{2\,435} = \sum_{i=0}^{4\,870} h(i) \quad et \quad P_{2\,435} = \sum_{i=0}^{4\,870} p(i).$$

6 Excursions et arbres binaires

Il existe une bijection entre les excursions de longueur $2n$ et les arbres binaires à n sommets internes (et $n+1$ feuilles). Soit A un arbre binaire à n sommets internes. Associons à chaque sommet i un entier $\mathrm{val}(i) \in \{-1, 1\}$. Si le sommet i est interne alors $\mathrm{val}(i) = 1$; sinon, c'est une feuille et $\mathrm{val}(i) = -1$. Soit $\sigma(0), \ldots, \sigma(2n)$ les numéros des sommets de A parcourus en ordre préfixe (moi-puis-gauche-puis-droite). On démontre, par récurrence immédiate, que la suite $(\epsilon_A(i))_{i=0..2n}$ définie par $\epsilon_A(i) = \sum_{j=0}^{i-1} \mathrm{val}(\sigma(j))$, est une excursion. Notons $\mathcal{E} : A \mapsto \epsilon_A$ la fonction qui à un arbre binaire associe son excursion. Remarquez que si on prolonge ϵ_A en $2n+1$, on a $\epsilon_A(2n+1) = -1$, indépendamment de A, car il y a toujours une feuille de plus que de sommets internes. La figure suivante donne un exemple de l'excursion associée à un arbre binaire.



Les sommets sont étiquetés alphabétiquement dans l'ordre du parcours préfixe de l'arbre

Question 19 *Donnez l'excursion associée à l'arbre $A_{4,1}$ de la question 16.*

Question 20 *Soit e, l'excursion associée à l'arbre $A_{130,2}$. Quelles sont les valeurs de $e(121)$ et $e(200)$? Quel est la distance maximale entre deux records consécutifs de e ? Mêmes questions pour les excursions associées aux arbres $A_{250,1}$ et $A_{2435,1}$.*

La fonction $\mathcal{E} : A \mapsto \epsilon_A$ est en fait une bijection des arbres binaires à n sommets internes sur les excursions de longueur $2n$ (on admet ce résultat). On s'intéresse maintenant à la fonction réciproque $\mathcal{E}^{-1}$.

Question 21 *Déterminez les arbres binaires correspondants aux excursions* e *et* e' *définies par :*
– $e(i) = i$ *pour* $0 \leqslant i \leqslant n$, *et* $e(i) = 2n - i$ *pour* $n \leqslant i \leqslant 2n$.
– $e'(i) = 0$ *si* i *est pair, et* $e'(i) = 1$ *si* i *est impair.*

Question 22 *Proposez un algorithme implémentant la fonction réciproque* $\mathcal{E}^{-1}$ *de* $\mathcal{E}$*, qui prend une excursion de longueur* $2n$ *et renvoie l'arbre binaire à* n *sommets internes associé. Justifiez votre algorithme et évaluez son nombre d'opérations en fonction de* n.

Question 23 *Soit* α *l'arbre associé à l'excursion* $e_{25,1}$ *(définition section 4). Quelle est la hauteur de* α *? Donnez la valeur de la somme cumulée des hauteurs des sommets de* α*, et la somme cumulée des profondeurs des sommets de* α.
Mêmes questions pour les arbres associés à $e_{250,1}$, $e_{2051,0}$ *et* $e_{2057,23}$.

— FIN —