Aspects algorithmiques de la combinatoire

Examen du 7 mars 2013 – durée : 3 heures

Notes de cours (imprimées ou manuscrites) autorisées

Problème 1 : Génération aléatoire d'arbres de Motzkin

Un *arbre de Motzkin* de taille n est un arbre plan à n sommets, tous d'arité au plus 2 : un tel arbre a donc potentiellement des *nœuds binaires*, des *nœuds unaires* et des *feuilles*. Pour tout $n \geqslant 1$, on note $\mathcal{M}_n$ l'ensemble des arbres de Motzkin de taille n, et m_n son cardinal ; on note par ailleurs $\mathcal{M}$ la classe combinatoire formée de l'union des $\mathcal{M}_n$.

Le but de ce problème est de décrire plusieurs méthodes de tirage aléatoire dans $\mathcal{M}$. Il est découpé en quatre questions indépendantes qui peuvent être traitées dans l'ordre souhaité (mais il est sans doute préférable de terminer par la question 4 dont la fin peut prendre beaucoup de temps).

Question 1

(a). Dessiner les arbres de Motzkin de taille 1 à 5 (il y en a respectivement 1, 1, 2, 4, 9).

(b). Quelle relation existe-t-il entre la taille n d'un arbre de Motzkin, son nombre de feuilles f et son nombre de nœuds unaires u ?

(c). Proposer un codage des arbres de Motzkin par les mots d'un langage sur un alphabet de 3 lettres, analogue au codage usuel des arbres binaires par des mots de Dyck.

(d). À l'aide du lemme cyclique, en déduire la formule suivante :

$$\forall n \geqslant 1, \quad m_n \;=\; \frac{1}{n} \sum_{k \geqslant 0} \binom{n}{k} \binom{n-k}{k+1}.$$

Question 2

(a). Écrire une spécification, quadratique, décrivant la classe $\mathcal{M}$.

(b). En déduire que la série génératrice des arbres de Motzkin $M(z) = \sum_{n=1}^{\infty} m_n z^n$ est égale à :

$$M(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}. \tag{1}$$

(c). Décrire l'algorithme de génération uniforme obtenu à l'aide de la spécification de la question 2(a) par la méthode récursive. Quelle est sa complexité (en temps et en espace) ?

(d). Décrire le générateur de Boltzmann induit par la spécification de la question 2(a). Pour quelles valeurs du paramètre peut-il être instancié ? Comment choisir cette valeur en fonction de la taille visée ?
(On pourra admettre que $xM'(x) = \frac{M(x)}{\sqrt{(1+x)(1-3x)}}$)

Question 3 On va utiliser ici une autre spécification pour la classe $\mathcal{M}$, utilisant la notion de *substitution*. Soit deux classes combinatoires $\mathcal{A}$ et $\mathcal{B}$, non étiquetées, telles que :

– les atomes constituant chaque élément de $\mathcal{A}$ sont *distingables*,

– $\mathcal{B}$ ne contient pas d'objet de taille nulle.

Pour tout α dans $\mathcal{A}$ de taille k et tous $\beta_1, \ldots, \beta_k$ dans $\mathcal{B}$, on note $\alpha[\beta_1, \ldots, \beta_k]$ l'objet obtenu en substituant aux k atomes de α (distingables, donc pouvant être canoniquement numérotés de 1 à k) les k objets β_i. La classe $\mathcal{C} = \mathcal{A}[\mathcal{B}]$ est la classe formée de tous les objets obtenus. Inversement, tout objet de $\mathcal{C}$ se décompose de manière unique en $\alpha[\beta_1, \ldots, \beta_k]$. Les séries génératrices de $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ vérifient alors la relation :

$$C(z) \; = \; A(B(z)).$$

(a). Soit $\Gamma\mathcal{A}(x)$ et $\Gamma\mathcal{B}(x)$ des générateurs de Boltzmann pour $\mathcal{A}$ et $\mathcal{B}$. On définit le générateur $\Gamma\mathcal{C}(x)$ de la manière suivante :

- $\alpha \longleftarrow \Gamma\mathcal{A}(B(x)), \quad k \longleftarrow |\alpha|$
- $\forall i \in [\![1, k]\!], \quad \beta_i \longleftarrow \Gamma\mathcal{B}(x)$ (k appels indépendants)
- retourner $\alpha[\beta_1, \ldots, \beta_k]$.

Vérifier que $\Gamma\mathcal{C}(x)$ tire un élément de $\mathcal{C} = \mathcal{A}[\mathcal{B}]$ selon la distribution de Boltzmann de paramètre x.

(b). La classe $\mathcal{M}$ peut s'exprimer à l'aide d'une substitution dans la classe des arbres binaires. Donner la spécification complète correspondante, et vérifier qu'on peut ainsi retrouver l'expression (1) de la question 2(b).

(c). Décrire un nouveau générateur de Boltzmann pour $\mathcal{M}$.

Question 4 Cette fois, on cherche une spécification *holonome* de $\mathcal{M}$– autrement dit, n'utilisant que les constructions *somme*, *pointage*, et *produit par une séquence d'atomes*. On note respectivement $\mathcal{M}^\bullet$, $\mathcal{M}^\circ$ et $\mathcal{M}^\diamond$ les classes obtenues à partir de $\mathcal{M}$ par pointage d'un sommet quelconque, d'une feuille ou d'un nœud unaire.

(a). Montrer que $\mathcal{M} + \mathcal{M}^\bullet = 2\mathcal{M}^\circ + \mathcal{M}^\diamond$.

(b). Comment transformer une générateur de Boltzmann pour $\mathcal{M}+\mathcal{M}^\bullet$ en un (pseudo-)générateur de Boltzmann pour $\mathcal{M}^\bullet$ ("pseudo" signifiant ici que sa terminaison n'est pas assurée – mais sa complexité en moyenne reste du même ordre de grandeur que celle du générateur pour $\mathcal{M} + \mathcal{M}^\bullet$) ?

On peut montrer la relation de récurrence suivante :

$$(n+1)m_n = (2n-1)m_{n-1} + 3(n-2)m_{n-2},$$

qui "doit" être le reflet d'une relation entre les classes $\mathcal{M}$ et $\mathcal{M}^\bullet$:

$$(\mathcal{M} + \mathcal{M}^\bullet) = 2\mathcal{Z} + \mathcal{Z}(\mathcal{M} + 2\mathcal{M}^\bullet) + 3\mathcal{Z}^2\mathcal{M}^\bullet \tag{2}$$

(où le terme $2\mathcal{Z}$ est juste une initialisation).

(c). En admettant que l'équation (2) soit justifiée de manière constructive, elle relie les classes $\mathcal{M}^\bullet$ et $\mathcal{M}+\mathcal{M}^\bullet$. En utilisant l'astuce du (b), décrire des (pseudo-)générateurs de Boltzmann pour ces deux classes.

Pour démontrer l'équation (2), on s'appuie sur la relation montrée en (a), ce qui donne :

$$2\mathcal{M}^\circ + \mathcal{M}^\diamond = 2\mathcal{Z} + \mathcal{Z}(\mathcal{M}^\bullet + 2\mathcal{M}^\circ + \mathcal{M}^\diamond) + 3\mathcal{Z}^2\mathcal{M}^\bullet. \tag{3}$$

(d). Donner une construction simple reliant le terme $\mathcal{M}^\diamond$ de gauche et le terme $\mathcal{Z}\mathcal{M}^\bullet$ de droite.

(e). Donner une construction simple reliant les éléments de $2\mathcal{M}^\circ$ dont la feuille pointée est portée par un nœud unaire et le terme $2\mathcal{Z}\mathcal{M}^\circ$ de droite.

(f). Expliquer les derniers termes par une opération de même type que celle qui permet la croissance des arbres binaires dans l'algorithme de Rémy.

(g). En vous inspirant du principe des générateurs de Boltzmann singuliers pour les séquences supercritiques, proposer un algorithme de génération uniforme dans $\mathcal{M}_n$ (donc en taille exacte) de complexité linéaire en moyenne.