DEA d'ALGORITHMIQUE 1998-1999

Analyse d'algorithmes: Modèles combinatoires
Philippe.Flajolet@inria.fr
http://algo.inria.fr/flajolet/Teach/

Règle du jeu. Rédiger votre copie en au plus deux heures et trente minutes au total pour les exercices I et II, et en au plus 30 minutes pour le QCM. Ceux qui le souhaitent pourront ensuite compléter leurs réponses aux exercices sur feuillets séparés clairement marqués comme tels.

Postez vos copie le **jeudi 11 mars 1999** au plus tard à:
Philippe Flajolet,
(DEA ALGO)
INRIA Rocquencourt
F-78153 Le Chesnay

Pensez a garder une copie pour vous au cas où $\cdots$

Partie I. Prédictions et scrutins

Le problème des prédictions de résultats de scrutin consiste à observer de manière partielle un dépouillement de sorte à tenter de prédire avec bonne probabilité ce qui va en être le résultat final. Ce problème statistique est le "problème inverse" du problème de probabilités discrètes suivant: "On considère 2 candidates en lice appellés Arlette et Bernadette. Qu'observe-t-on au cours du dépouillement?"

Soit $\mathcal{A} = \{a, b\}$ un alphabet à deux lettres. la taille d'un mot est sa longueur. On utilise les mêmes lettres pour un langage, sa série génératrice, et la suite des coefficients correspondants. On considère les langages suivants:

— $\mathcal{W} = \{a, b\}^\star$ est le langage de tous les mots;

— $\mathcal{E}$ est le langage des mots contenant autant de a que de b: $\mathcal{E} = \{\epsilon, ab, ba, aabb, bbaa, abab, \ldots\}$;

— $\mathcal{P}$ est le sous ensemble des mots w de E tels que, de plus, dans tout prefixe de w le nombre de a soit toujours supérieur ou égal au au nombre de b: $\mathcal{P} = \{\epsilon, ab, aabb, abab, \ldots\}$;

— $\mathcal{Q}$ est le langage obtenu à partir de $\mathcal{P}$ en echangeant le rôle des lettres a et b.

On pourra s'aider d'un dessin ou un mot est représenté par une ligne polygonale partant de $(0, 0)$, la lettre a étant codée par l'ajout du vecteur $(1, 1)$ et la lettre b par l'ajout de $(1, -1)$. Il n'est pas nécessaire de détailler les calculs intermédiaires et on peut s'appuyer sur un système de calcul formel pour l'obtention des résulats.

Q1. Donner la série génératrice $W(z)$. Observer que $\mathcal{P}$ vérifie l'equation

$$\mathcal{P} = \mathtt{Seq}(a \cdot \mathcal{P} \cdot b),$$

et en déduire $P(z)$. Montrer une fois pour toutes que pour toute série $f(z)$, on a: $[z^{2n}]f(z^2) = [z^n]f(z)$. Expliciter le coefficient $P_{2n} = [z^{2n}]P(z)$.

Q2. Montrer que

$$\mathcal{E} = \mathtt{Seq}\left((a \cdot \mathcal{P} \cdot b) \cup (b \cdot \mathcal{Q} \cdot a)\right).$$

En déduire la série génératrice $E(z)$ ainsi que E_{2n}.

Évaluer exactement et estimer asymptotiquement le quotient

$$\frac{P_{2n}}{E_{2n}},$$

lequel donne la probabilité qu'Arlette ne soit jamais dominée par Bernadette au cours du dépouillement, sachant qu'Arlette et Bernadette sont à égalité de votes.

Q3. Soit $E_{n,k}$ le nombre de mots de $\mathcal{E}_n$ ayant exactement k préfixes (de longueur entre 1 et n) qui sont dans $\mathcal{E}$, et $E(z,u) = \sum_{n,k} E_{n,k} z^n u^k$. Graphiquement, il s'agit des lignes polygonales touchant k fois l'axe horizontal. Montrer que

$$E(z,u) = \frac{1}{1 - uz^2(P(z) + Q(z))}.$$

On pose

$$\mu_{2n} := \frac{1}{E_{2n}}[z^{2n}]\left(\frac{\partial}{\partial u}E(z,u)\right)_{u=1}.$$

Justifier d'abord que μ_{2n} est l'espérance du nombre de fois où Arlette et Bernadette sont à égalité au cours du dépouillement, sachant qu'elles ont recueilli un nombre égal de voix.

Utiliser les expressions exactes ou l'analyse de singularité pour obtenir un equivalent asymptotique de μ_{2n}.

Q4. Pour un scrutin particulier $w = w_1 \cdots w_{2n} \in \mathcal{E}$, soit $\chi[w]$ la valeur du premier indice j avec $0 \le j < n$ tel que $w_1 \cdots w_{j+1}$ comporte plus de b que de a; on pose $\chi[w] = n$ si un tel indice n'existe pas. (On mesure donc la durée de la période initiale où Arlette domine Bernadette.) Par exemple, $\chi[baab] = 0$, $\chi[aabbbba] = 4$, $\chi[aaababbb] = 8$. Soit

$$F(z,v) := \sum_{w \in \mathcal{E}} z^{|w|} v^{\chi(|w|)}.$$

Montrer que

$$F(z,v) = 1/(1 - (z^2 v^2 P(zv) + z^2 Q(z)))$$

En déduire la valeur exacte et asymptotique de l'espérance

$$\nu_{2n} = \frac{1}{E_{2n}}[z^{2n}]\left(\frac{\partial}{\partial v}F(z,v)\right)_{v=1}.$$

Partie II. Modèles d'urnes

On donne ici une approche combinatoire à un modèle de transfert de particules et de chaleur étudié par Pólya et Ehrenfest.

Q1. Le nombre de manières de placer les n éléments $\{1, 2, \ldots, n\}$ dans k urnes distinguées, avec chaque urne contenant un nombre pair de boules (possiblement égal à 0), vaut

$$G_{n,k} = n!\,[z^n](\cosh z)^k,$$

où $\cosh(z) = (e^z + e^{-z})/2$.

Montrer que $(\cosh z)^k$ se développe en une combinaison linéaire finie d'exponentielles de la forme $e^{\alpha z}$. Exprimer $G_{n,k}$.

Q2. Soit $f(z) = \sum_n f_n z^n/n!$ une série génératrice exponentielle, $\widetilde{f}(z) = \sum_n f_n z^n$ la série ordinaire correspondante. Déterminer $\widetilde{f}(z)$ lorsque

$$f(z) = e^{\alpha z}, \quad f(z) = \cosh(z), \quad f(z) = ze^z.$$

Montrer que la transformation $f \mapsto \widetilde{f}$ est linéaire.

Déterminer la série génératrice ordinaire $\widetilde{f}(z)$ correspondant à $f(z) = (\cosh(z))^k$ sous forme de décomposition en éléments simples.

Obtenir un équivalent asymptotique de $G_{2n,k}$ lorsque k est fixé et n tend vers l'infini.

Q3. Deux cavités appelées A et B contenant au total k particules (distinctes) sont en communication par un trou très fin. A chaque seconde, une particule passe de A à B ou inversement. Au temps 0, la cavité A est vide et la cavité B est pleine.

Donner la série génératrice des fonctions de $\{1, \ldots, n\}$ dans $\{1, \ldots, k\}$ telles que l'image inverse de tout $j \in [1, k]$ soit de taille paire.

On postule que toutes les évolutions possibles en n étapes sont équiprobables. Relier le problème aux questions précédentes.

Justifier par la méthode de votre choix que la série génératrice exponentielle de toutes ces évolutions vaut e^{kz}.

Déterminer un équivalent asymptotique en grand temps de la probabilité que la cavité A soit vide au bout de n étapes.

Q4. La série gńératrice exponentielle des évolutions qui aboutissent à avoir r particules dans la cavité A vaut

$$\binom{k}{r}(\cosh z)^{k-r}(\sinh z)^r.$$

Justifier ce résultat.

Déterminer les deux premiers termes du développement asymptotique ($n \to +\infty$) de l'espérance E_n de la variable aléatoire représentant le nombre d'atomes contenus dans l'urne A au temps n; déterminer ainsi a quelle vitesse on se "rapproche de l'équilibre". (Suggestion: former une série bivariée où u marque les différentes valeurs possible de r.)

III. QUESTIONNAIRE

Cochez une case et justifiez en au plus deux lignes votre choix.

Q1. Soit $A(z)$ la série génératrice ordinaire d'une classe $\mathcal{A}$. Soit $\mathcal{C}$ la classe des paires ordonnées (α, β) d'éléments de $\mathcal{A}$ telles que $|\alpha| = |\beta|$. Alors la série génératrice ordinaire de $\mathcal{C}$ vérifie:

☐ $C(z) = \sum_n (A_n z^n)^2$

☐ $C(z) = \left(\sum_n A_n z^n \right)^2$

....

....

Q2. Le nombre X_n d'arbres unaires-binaires (dont les sommets ot degré 0, 1, ou 2) comportant deux fois plus de sommets binaires que de feuilles (sommets de degré 0) vérifie

☐ $\dfrac{X_n}{2^n} \to +\infty$

☐ $\dfrac{X_n}{2^n} \to +0$

....

....

Q3. Le nombre de partitions de n en sommants premiers, c-à-d le nombre des $(\nu_1, \nu_2, \ldots, \nu_k)$ avec $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_k$, $\nu_1 + \cdots + \nu_k = n$, chaque $\nu_j \in \{2, 3, 5, \ldots\}$, a pour série génératrice ordinaire (OGF)

☐ $\displaystyle\prod_{p \text{ premier}} \frac{1}{1 - z^p}$

☐ $\dfrac{1}{1 - \sum_{p \text{ premier}} z^p}$

....

....

Q4. Le nombre de compositions de l'entier n en sommants qui sont des puissances de 2, c-à-d le nombre des $(\nu_1, \nu_2, \ldots, \nu_k)$ avec $\nu_1 + \cdots + \nu_k = n$, chaque $\nu_j \in \{1, 2, 4, 8, \ldots\}$ a pour série génératrice ordinaire (OGF)

☐ $\displaystyle\sum_{n=0}^{\infty} \frac{z^{2^n}}{1 - z^{2^n}}$

☐ $\dfrac{1}{1 - \sum_{n=0}^{\infty} z^{2^n}}$

....

....

Q5. Une "quasi-permutation" est un objet étiqueté (EGF) defini comme un ensemble de cycles non-orientés (de taille ≥ 1). La série génératrice exponentielle des quasipermutations est:

- ☐ $\dfrac{e^{z/2}}{\sqrt{1-z}}$
- ☐ $\dfrac{e^{-z/2}}{1-z}$

....

....

Q6. On considère $f_n = [z^n]e^{-1+e^{-1+e^z}} = [z^n]\exp(-1+\exp(-1+e^z))$. Alors:

- ☐ $f_n \to 0$
- ☐ $f_n \to \infty$

....

....

Q7. Soit $h_n = [z^n]\log\dfrac{1}{1-z/2}e^{z^2+z^3}$. Alors:

- ☐ $h_n \sim e^{12}2^{-n}/n$
- ☐ $h_n \sim e^{-12}2^{n}/n$

....

....

Q8. On rappelle que $[z^n]\dfrac{1}{\sqrt{1-z}} \sim \dfrac{1}{\sqrt{\pi n}}$ et $[z^n]\log\dfrac{1}{1-z} = \dfrac{1}{n}$. Soit $k_n = [z^n]\dfrac{1}{\sqrt{1-3z}}\log\dfrac{1}{1-2z}$. Alors:

- ☐ $k_n \sim \log 3\dfrac{3^n}{\sqrt{\pi n}}$
- ☐ $k_n \sim \sqrt{2}\dfrac{2^n}{n}$

....

....

Q9. La probabilité p_n q'un mot binaire de longueur n contienne au moins une fois le motif abb vérifie:

- ☐ $p_n \to 1$
- ☐ $p_n \to \dfrac{1}{8}$

....

....

Q10. Étant donné un graphe connexe dont un sommet source a et un sommet destination b ont été distingués. Soit L_n le nombre de chemins de a à b dans le graphe ayant longueur n. Alors:

- ☐ $\sum_n L_n z^n$ est une fraction rationnelle
- ☐ $\sum_n L_n \frac{z^n}{n!}$ est une fraction rationnelle

....

....