Option Informatique en Spé MP et MP*

Devoir surveillé du samedi 9 février 2008

Langages rationnels
Ordre lexicographique
Numération exotique

Résumé

Le premier problème propose l'étude d'une relation de réécriture, ainsi que des classes modulo cette relation ; on montre que l'ensemble des représentants de ces classes est un langage rationnel ; on demande la programmation de deux fonctions ; enfin, on établit un résultat plus général, lié à l'ordre lexicographique.

À côté de la numération de position en base 10 ou en base 2, il existe bien d'autres systèmes de numération ≪exotiques≫ : numération redondante (avec les chifres 0, 1 et -1) ; numération en base de FIBONACCI (système de ZECKENDORF) ; numération en base complexe. Le deuxième problème présente un système utilisant une base fractionnaire ; vous serez à nouveau amené à écrire des fonctions en Caml.

Veuillez rédiger chaque partie sur une copie séparée.

Table des matières

Consignes de programmation

▶ L'usage des mots `ref`, `for`, `if`, `then`, `else` est interdit. Vous ne devez utiliser ni vecteurs, ni chaînes de caractères. Comptent pour une ligne : l'en-tête d'une fonction ; chaque motif ; une fonction définie par un seul motif (`let f x = x + 1;;`) ne compte que pour une ligne.

1 Problème 1 : YAOPRL

1.1 Une relation de réécriture

▶ Dans cet exercice, l'alphabet utilisé est $\Sigma = \{a,b,c\}$. Soient u un mot et $k \in [\![0, |u|]\!]$; $u[1..k]$ désigne le préfixe de u de longueur k. Soit E un ensemble fini ; le nombre d'éléments de E noté $|E|$.

▶ Nous définissons sur Σ^* une relation notée $\leftrightarrow$, définie comme suit : $u \leftrightarrow v$ s'il existe des mots x et y tels que $\{u,v\} = \{xaby,xbay\}$. Cette relation est clairement symétrique. Nous noterons $u \overset{*}{\leftrightarrow} v$ si $u = v$, ou s'il existe une suite $(m_0, m_1, \dots, m_n)$ de mots tels que $m_0 = u$, $m_n = v$ et $m_i \leftrightarrow m_{i+1}$ pour $0 \leqslant i < n$. Il est clair que $\overset{*}{\leftrightarrow}$ est une relation d'équivalence.

▶ Soit $u \in \Sigma^*$; la classe de u modulo $\overset{*}{\leftrightarrow}$, notée $\overline{u}$, est l'ensemble des mots v vérifiant $v \overset{*}{\leftrightarrow} u$; dans cette classe, tous les mots ont la même longueur. Par exemple, $\{abcabb, abcbab, abcbba, bacabb, bacbab, bacbba\}$ est une classe modulo $\overset{*}{\leftrightarrow}$.

Question 1 Montrez que chaque classe modulo $\overset{*}{\leftrightarrow}$ est un langage rationnel.

Question 2 Combien existe-t-il de classes de mots de longueur 3 ?

Question 3 Le nombre de classes modulo $\overset{*}{\leftrightarrow}$ est-il fini ou infini ?

Question 4 Donnez un encadrement de $|\overline{u}|$, en fonction de $|u|$. Le majorant proposé doit être un $o(3^n)$.

Question 5 Dans cette question, nous fixons $n \in \mathbb{N}$. Soient u et v deux mots de longueur n, vérifiant $u \overset{*}{\leftrightarrow} v$; la *distance* de u à v est la longueur minimale d'un chemin menant de u à v ; nous la noterons $d(u,v)$. Déterminez, en fonction de n, la valeur maximale de $d(u,v)$; vous préciserez un couple (u,v) pour lequel ce maximum est atteint.

1.2 Représentants minimaux pour l'ordre lexicographique

▶ Nous imposons sur notre alphabet l'ordre $a < b < c$. Rappelons la définition de l'ordre lexicographique sur Σ^* : si u et v sont deux mots sur Σ, nous dirons que u précède v, et nous noterons $u \preceq v$, si l'une des deux conditions suivantes est satisfaite :

- u est préfixe de v ;
- il existe un indice i tel que $u[1..i-1] = v[1..i-1]$ et $u_i < v_i$.

Nous savons que la relation $\preceq$ est un ordre total sur Σ^*.

Question 6 Donnez une preuve *très simple* de la transitivité de la relation $\preceq$. Indication : ajoutez une lettre bien choisie à l'alphabet.

▶ Soit $\mathcal{R}$ une classe modulo $\overset{*}{\leftrightarrow}$. Le *représentant* de $\mathcal{R}$ est le plus petit élément de $\mathcal{R}$ pour l'ordre lexicographique. Par exemple, le représentant de la classe de *bacbab* est *abcabb*.

▶ Notons M l'ensemble des représentants des classes modulo $\overset{*}{\leftrightarrow}$.

Question 7 Donnez une expression rationnelle du langage M.

Question 8 Dessinez un automate fini déterministe complet reconnaissant M.

▶ Pour les questions de programmation, définissons le type Caml `lettre` comme suit :

```
type lettre = A | B | C ;;
```

Chaque mot de Σ^* sera représenté par une `lettre list`. Par exemple, le mot *babccaacb* sera représenté par la liste `[B;A;B;C;C;A;A;C;B|`. La définition donnée plus haut du type `lettre` a une conséquence agréable : Caml considère que `A` $<$ `B` $<$ `C`.

Question 9 • Rédigez en Caml une fonction :

```
compare_lex : lettre list * lettre list -> int
```

spécifiée comme suit : `compare_lex(u,v)` rend le résultat 0 (si $u = v$), 1 (si $u \preceq v$) ou 2 (si $v \preceq u$). Le coût de cette fonction doit être un $\mathcal{O}\big(\min(|u|,|v|)\big)$. Vous ne devez pas utiliser plus de sept lignes.

Question 10 Décrivez un algorithme qui, à un mot $u \in \Sigma^*$, associe le représentant de la classe de u modulo $\overset{*}{\leftrightarrow}$. Le coût de cet algorithme doit être un $\mathcal{O}\big(|u|\big)$.

Question 11 $\star\star\star$ • Rédigez en Caml une fonction :

```
représentant : lettre list -> lettre list
```

spécifiée comme suit : `représentant u` calcule le représentant de u modulo $\mathcal{R}$. Le coût de cette fonction doit être un $\mathcal{O}\big(|u|\big)$. Vous ne devez pas utiliser plus de dix lignes.

1.3 Encore une transformation conservant le caractère rationnel

► Dans cette partie, l'alphabet utilisé est $\Sigma = \{a,b\}$; nous munissons Σ^* de l'ordre lexicographique induit par la relation $a < b$. Soit L un langage sur Σ. Nous dirons que l'entier n est *actif* si $L \cap \Sigma^n$ n'est pas vide ; notons J l'ensemble des entiers n actifs. À chaque $n \in J$, nous associons le mot $\mu(n)$, qui est le plus petit élément de $L \cap \Sigma^n$ pour l'ordre lexicographique. Enfin, notons $\Phi(L) = \{\mu(n) \mid n \in J\}$. Par exemple, si L est l'ensemble des mots de longueur paire, alors $\Phi(L)$ est décrit par l'expression rationnelle $(aa)^*$.

► Nous nous proposons d'établir le résultat suivant : si L est rationnel, alors $\Phi(L)$ est rationnel.

Question 12 Explicitez $\Phi(L)$ lorsque L est l'ensemble des mots contenant au moins une occurrence de la lettre b.

Question 13 Explicitez $\Phi(L)$ lorsque L est l'ensemble des mots contenant deux occurrences non consécutives de la lettre b.

► Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe reconnaissant L. Començons par définir deux fonctions $\mathbf{D} : \wp(Q) \mapsto \wp(Q)$ et $\mathbf{T} : \wp(Q) \times \Sigma \mapsto \wp(Q)$ comme suit :

- $q' \in \mathbf{D}(S)$ ssi il existe $q \in S$ et $x \in \Sigma$ tels que $\delta(q,x) = q'$;
- $q' \in \mathbf{T}(S,x)$ ssi il existe $q \in S$ et $y \in \Sigma$ tels que $y < x$ et $\delta(q,y) = q'$.

► Définissons maintenant un nouvel automate $\mathcal{B}$:

- ses états sont les couples (q,S) où $q \in Q$ et $S \subset Q$;
- son état initial est le couple $(i, \emptyset)$;
- un état (q,S) de $\mathcal{B}$ est final ssi $q \in F$ et $S \cap F = \emptyset$;
- sa fonction de transition Δ est définie comme suit : $\Delta\big((q,S),x\big) = \big(\delta(q,x),S'\big)$ où S' est la réunion de $\mathbf{D}(S)$ et de $\mathbf{T}(S,x)$.

► Soit u un mot de L, de longueur $n \geqslant 1$. Pour $k \in [\![0,n]\!]$, notons (q_k, S_k) le k-ième état par lequel passe l'automate $\mathcal{B}$, au cours de la lecture du mot u.

Question 14 Montrez que $q \in S_k$ ssi il existe un mot v de longueur k, qui est strictement inférieur à $u[1..k]$ pour l'ordre lexicographique et vérifie $\delta^*(i,v) = q$.

Question 15 Montrez que l'automate $\mathcal{B}$ reconnait le langage $\Phi(L)$.

Question 16 $\star\star\star$ Soit L un langage rationnel non vide sur Σ, reconnu par un automate fini $\mathcal{A}$. Notons w le plus petit mot de L, pour l'ordre lexicographique. Construisez un automate $\mathcal{M}$ reconnaissant $\{w\}$.

► Les résultats que nous venons d'établir se généralisent aux langages sur un alphabet à plus de 2 lettres.

2 Problème 2 : un système de numération exotique

▶ Dans tout l'exercice, p et q sont deux naturels premiers entre eux, vérifiant $p > q$. Σ désigne l'ensemble $[\![0, p-1]\!]$.

▶ Soit $n \in \mathbb{N}^*$. notons $N_0 = n$; puis, pour i commençant à 0, définissons N_{i+1} et a_i par les deux conditions $qN_i = pN_{i+1} + a_i$ et $0 \leqslant a_i < p$. En clair : N_{i+1} est le quotient et a_i le reste dans la division euclidienne de qN_i par p.

Question 1 • Montrez que la suite $(N_i)_{i\in\mathbb{N}}$ est décroissante.

Question 2 • Montrez qu'il existe un naturel k tel que $N_k > 0$ et $N_i = 0$ pour tout $i > k$.

Question 3 • Montrez que $a_k > 0$.

▶ La représentation de n en base p/q est le mot $a_k a_{k-1} \dots a_1 a_0$ sur l'alphabet Σ.

Question 4 • Vérifiez que la représentation de 5 en base 3/2 est 2101 ; vous déroulerez les étapes de l'algorithme exposé plus haut.

Question 5 • Qu'obtient-on dans le cas particulier $q = 1$?

Question 6 • Quelle est la décomposition de 17 en base 5/3 ?

Question 7 • Rédigez en Caml une fonction de signature :

```
representation : int -> int -> int -> int list
```

spécifiée comme suit : `representation p q n` donne la représentation de n en base p/q. Vous ne devez pas utiliser plus de cinq lignes.

Question 8 • Montrez qu'il existe une suite $(u_i)_{i\in\mathbb{N}}$ de rationnels (qui dépendent de p et q) telle que, pour tout $n \geqslant 1$, la suite $(a_i)_{0\leqslant i\leqslant k}$ qui représente n en base p/q vérifie $\sum_{i\geqslant 0} a_i u_i = n$.

Question 9 • Rédigez en Caml une fonction de signature :

```
conversion : int -> int -> int list -> int
```

spécifiée comme suit : `conversion p q l` donne le naturel n dont ℓ est la représentation en base p/q. Attention : si vous effectuez des divisions, elle doivent «tomber juste». Vous ne devez pas utiliser plus de cinq lignes.

▶ Nous nous intéressons au langage $L_{p/q}$, ensemble des représentations des naturels non nuls en base p/q.

Question 10 • Montrez que si w appartient à $L_{p/q}$, alors tout préfixe (non vide) de w appartient à $L_{p/q}$.

Question 11 • Montrez que tout mot w appartenant à $L_{p/q}$ est préfixe d'au moins un autre mot de $L_{p/q}$.

Question 12 • Exhibez un langage rationnel infini disjoint de $L_{3/2}$.

FIN