Option Informatique en Spé MP et MP*

Devoir surveillé du mardi 11 décembre 2001

Autour du mot de Kolakoski

Résumé

Le mot de KOLAKOSKI est l'unique mot infini sur l'alphabet $\Sigma = \{1,2\}$ commençant par un 2 et égal à la suite des longueurs de ses blocs. On ne sait pas grand'chose de ce mot, à part le fait qu'il n'est pas ultimement périodique. Des conjectures sont ouvertes sur la répartition des 1 et des 2, sur la complexité ...

Bien évidemment, l'objet de ce texte n'est pas de vous proposer l'attaque de telles conjectures ! La première partie présente l'opérateur Φ qui, à un mot u sur l'alphabet $\{1, 2\}$ ne contenant pas trois lettres consécutives identiques, associe la suite des longueurs de ses blocs, considérée comme un mot sur ce même alphabet.

La deuxième partie s'intéresse de près à l'itération de cet opérateur.

La troisième partie définit le mot de KOLAKOSKI et en donne deux méthodes de calcul.

Chaque partie comporte une ou plusieurs questions difficiles ; elles sont explicitement signalées par $\star\star\star$ (trois étoiles). Les étudiant(e)s qui visent une E.N.S. en option informatique ou math/info sont invités à attaquer ces questions, sans pour autant négliger le reste.

Ce texte vous donnera à plusieurs reprises l'occasion de montrer votre habileté en Caml. Rappel : l'emploi de références ou de structures de données mutables n'est pas autorisé. On attend des programmes courts, et accompagnés de quelques explications. Vous devez respecter rigoureusement les spécifications de l'énoncé : noms des fonctions, types des arguments et du résultat.

$\boxed{\textit{Veuillez rédiger chaque partie sur une copie séparée.}}$

Table des matières

1 Un morphisme alternant

► Dans ce problème, $\Sigma = \{1,2\}$. Σ^+ désigne l'ensemble $\Sigma^* \setminus \{\varepsilon\}$ des mots sur Σ autres que le mot vide.

► Nous noterons $\overline{1} = 2$ et $\overline{2} = 1$. Soit $u \in \Sigma^+$; u admet une et une seule décomposition en blocs de la forme $u = x^{e_1}\overline{x}^{e_2}\ldots$ où $x \in \Sigma$ et les e_k sont dans $\mathbb{N}^*$. Par exemple, le mot 1121112212222 admet la décomposition $1^2\,2^1\,1^3\,2^2\,1^1\,2^4$.

► Nous dirons que $u \in \Sigma^+$ est *stérile* si l'un au moins des deux mots 111 et 222 est facteur de u. Nous noterons K_1 l'ensemble des mots non stériles ; ce sont donc les mots dans la décomposition en blocs desquels n'interviennent que les exposants 1 et 2. Notez bien que $\varepsilon \notin K_1$.

Question 1 • Donnez une expression rationnelle de K_1 ; vous justifierez rigoureusement votre proposition.

Question 2 • Décrivez un automate fini déterministe complet reconnaissant K_1 ; vous justifierez rigoureusement votre proposition.

Question 3 • À la réflexion, quelle est la preuve la plus courte de la rationalité de K_1 ?

► Nous définissons le type Caml suivant :

```
type sigma = U | D ;;
```

Un mot sur Σ sera représenté par une `sigma list` ; par exemple, le mot 12112122 sera représenté par `[U;D;U;U;D;U;D;D]`.

Question 4 • Rédigez en Caml une fonction de signature

```
sterile : sigma list -> bool
```

spécifiée comme suit : `sterile u` indique si le mot u est stérile ; vous lèverez une exception si $u = \varepsilon$.

Question 5 • Exhibez un langage L non rationnel et contenu dans K_1. Bien entendu, vous devez prouver le caractère non rationnel de L.

► Soit $u \in K_1$; nous appelerons *fils* de u, et nous noterons $\Phi(u)$ le mot sur Σ formé par la suite des exposants dans la décomposition en blocs de u. Par exemple, $\Phi(2112212211212) = 122122111$. Vous pourrez noter $u \rightarrow v$ lorsque $v = \Phi(u)$.

Question 6 • Énoncez une condition nécessaire et suffisante simple portant sur u et v pour que l'on ait $\Phi(uv) = \Phi(u)\Phi(v)$.

Question 7 • Rédigez en Caml une fonction de signature

```
fils : sigma list -> sigma list
```

spécifiée comme suit : `fils u` construit le fils d'un mot $u \in \Sigma^+$ non stérile ; vous lèverez une exception si $u = \varepsilon$ ou si u est stérile.

Question 8 • Soit $u \in \Sigma^+$; montrez qu'il existe exactement deux éléments $v \in 1\Sigma^*$ et $w \in 2\Sigma^*$ de K_1 qui ont pour fils u.

► Nous dirons que v est le *père* et w la *mère* de u, et nous noterons $\pi(u) = v$ et $\mu(u) = w$.

Question 9 • Rédigez en Caml une fonction de signature

```
parents : sigma list -> sigma list * sigma list
```

spécifiée comme suit : `parents u` construit le couple $\big(\pi(u), \mu(u)\big)$ formé des parents de u ; vous lèverez une exception si $u = \varepsilon$.

Question 10 • Quel est l'intérêt de la définition du type `sigma` ?

► La suite de Fibonacci est définie par $F_0 = 0$, $F_1 = 1$ et la relation de récurrence $F_n = F_{n-1} + F_{n-2}$ pour $n \geqslant 2$.

Question 11 • Donnez une expression simple du nombre Q_n de mots de K_1 de longueur n.

► Nous noterons K_2 l'ensemble des mots u de K_1 tels que $\Phi(u) \in K_1$: ce sont donc les mots qui ont un «petit-fils». Plus généralement, nous noterons K_{n+1} l'ensemble des mots u de K_1 tels que $\Phi(u) \in K_n$.

Question 12 • Soit $u \in K_1$. Montrez que $|\Phi(u)| \leqslant |u|$. Dans quel(s) cas a-t-on l'égalité ?

Question 13 • Soit $u \in K_2$, de longueur au moins égale à 2. Montrez que $|\Phi^2(u)| < |u|$.

▶ Pour chacune des deux questions suivantes, vous décrirez soigneusement votre construction ; vous pouvez vous contenter d'une justification rapide.

Question 14 ⋆⋆⋆ • Soit L un langage rationnel sur l'alphabet Σ. Montrez que $\Phi^{-1}(L)$ est rationnel.

Question 15 ⋆⋆⋆ • Soit L un langage rationnel sur l'alphabet Σ. Montrez que si L est contenu dans K_1, alors $\Phi(L)$ est rationnel.

2 Itération de Φ

▶ La descendance d'un mot $u \in \Sigma^+$ est l'ensemble vide si u est stérile, la réunion de $\Phi(u)$ et de la descendance de ce dernier dans le cas contraire.

Question 16 • Montrez que la descendance d'un mot $u \in \Sigma^+$ est finie.

▶ Soit M un ensemble de mots non vide. Nous noterons $\widehat{M}$ l'ensemble des mots qui n'ont aucun facteur dans M. Ainsi, $K_1 = \widehat{M_1}$ avec $M_1 = \{111, 222\}$.

Question 17 • Montrez que si M est rationnel, alors $\widehat{M}$ est rationnel.

Question 18 ⋆⋆ • Soit $n \geqslant 1$. Montrez qu'il existe un langage M_n de cardinal $2^{n+1}-2$, tel que $K_n = \widehat{M_n}$.

▶ Nous venons de montrer que K_n est rationnel. Notons $\mathbb{K} = \bigcap_{n\geqslant 1} K_n$.

Question 19 • Soit $u \in \Sigma^+$. Montrez que $u \in \mathbb{K}$ ssi il existe un exposant d tel que $\Phi^d(u) = 1$.

Question 20 • Montrez que $\mathbb{K}$ est infini.

▶ Soient u et v deux mots sur Σ. Nous noterons $u \sim v$ s'il existe des exposants p et q tels que $\Phi^p(u) = \Phi^q(v)$.

Question 21 • Montrez que $\sim$ est une relation d'équivalence sur Σ^+.

Question 22 • Que représente $\mathbb{K}$ pour $\sim$?

Question 23 • Montrez que chaque classe modulo $\sim$ est infinie.

▶ Soient L un langage et $u \in L$. Soit y un facteur de u autre que le mot vide : $u = xyz$. Nous dirons que y est un *facteur itérant* de u vis-à-vis du langage L lorsque $xy^kz \in L$ pour tout $k \in \mathbb{N}$. Ainsi, le lemme de l'étoile affirme simplement que dans un langage rationnel, tout mot suffisamment long possède un facteur itérant vis-à-vis de ce langage. Nous noterons $\mathsf{FI}_L(u)$ l'ensemble (éventuellement vide) des facteurs itérants de u vis-à-vis du langage L

Question 24 ⋆⋆⋆ • Montrez que si u possède un facteur itérant vis-à-vis de $\mathbb{K}$, alors $\Phi(u)$ possède lui aussi un facteur itérant vis-à-vis de $\mathbb{K}$.

Question 25 • $\mathbb{K}$ est-il rationnel ?

▶ Soit $u \in \mathbb{K}$. Appelons *durée de vie* du mot u le plus petit exposant $d \in \mathbb{N}$ tel que $\Phi^d(u) = 1$. Nous le noterons $d(u)$. Ainsi $d(1) = 0$, $d(2) = 1$ et $d(12) = 3$.

Question 26 • Calculez la durée de vie du mot $u = 2211212212211211221211$.

Question 27 • Rédigez en Caml une fonction de signature

```
duree_de_vie : sigma list -> int
```

spécifiée comme suit : `duree_de_vie u` calcule la durée de vie du mot $u \in \mathbb{K}$; vous lèverez une exception si $u \notin \mathbb{K}$.

3 Le mot de Kolakoski

► Un *mot infini* sur l'alphabet Σ est une application $u : \mathbb{N}^* \mapsto \Sigma$. Nous noterons $u = u_1u_2u_3 \dots u_n \dots$ Un *facteur* de u est un mot fini de la forme $u[p..q] = u_pu_{p+1} \dots u_{q-1}u_q$, avec $1 \leqslant p \leqslant q$.

► Un mot infini u est ultimement périodique s'il existe $n_0 \geqslant 1$ et $p \geqslant 1$ tels que $u_{n+p} = u_n$ pour tout $n \geqslant n_0$. Lorsque $p = 1$, nous dirons que u est *stationnaire.*

► Soient $x \in \Sigma^*$ et $y \in \Sigma^+$; nous noterons xy^ω le mot infini $u = xyyy\dots$ obtenu en concaténant une infinité d'exemplaires du mot y derrière le mot x. Ce mot infini est ultimement périodique, avec $n_0 = |x| + 1$ et $p = |y|$. Remarquez que $y^\omega = yy^\omega$.

► Soit u u mot infini non stationnaire ; nous pouvons définir sa décomposition en blocs : $x^{e_1}\overline{x}^{e_2}\dots$ où $x \in \Sigma$ et les e_k sont dans $\mathbb{N}^*$. Nous dirons que u est *stérile* s'il contient un facteur 111 ou facteur 222. Sinon, nous pouvons définir le *fils* de u : c'est le mot infini $\Phi(u)$ sur Σ formé par la suite des longueurs des blocs dans la décomposition de u. Par exemple, $\Phi\big(21221(221121)^\omega\big) = 1121(2211)^\omega$; vous noterez (sans avoir besoin de le prouver) que le résultat établi à la question 6 pour les mots finis s'étend aux mots infinis. La descendance d'un mot infini u se définit comme celle d'un mot fini ; nous dirons que u est *ultimement stérile* s'il existe $n \in \mathbb{N}$ tel que $\Phi^n(u)$ soit stérile.

Question 28 ⋆⋆⋆ • Montrez que la descendance d'un mot infini ultimement périodique est finie.

Question 29 • Montrez que Φ possède exactement deux points fixes $\kappa \in 2\Sigma^*$ et $\kappa' \in 1\Sigma^*$. Vous utiliserez la remarque suivante : le j-ième caractère de l'un de ces points fixes est aussi la longueur de son j-ième bloc.

Question 30 • Quelle relation simple existe-t-il entre κ et κ' ?

Question 31 • κ est-il ultimement périodique ?

Question 32 • Rédigez en Caml une fonction de signature

```
kappa : int -> sigma
```

spécifiée comme suit : `kappa n` calcule la n-ième lettre du mot infini κ.

Question 33 • Montrez que $\kappa = \mu(\kappa_1)\pi(\kappa_2)\mu(\kappa_3)\pi(\kappa_4)\dots$

Question 34 • Proposez alors une autre version de la fonction `kappa`.

FIN