

Option Informatique en Spé MP et MP*

Langages reconnaissables sur un alphabet à une seule lettre

Le corrigé

Question 1 • L est reconnaissable, car $L = \{a^r\} \cdot \{a^s\}^*$. Soient $Q = \llbracket 0, r + s - 1 \rrbracket$, $i = 0$, $F = \{r\}$ et δ définie par $\delta(q, a) = q + 1$ si $q < r + s - 1$ et $\delta(r + s - a, a) = r$. Il est clair que l'automate (Q, δ, i, F) reconnaît L .

Question 2 • Rappelons qu'un relatif r est un *résidu quadratique modulo q* s'il existe $n \in \mathbb{N}$ tel que $r \equiv n^2 \pmod{q}$. Il est clair que, si l'on connaît les résidus quadratiques modulo q appartenant à $\llbracket 0, q - 1 \rrbracket$, on connaît aussi tous les autres.

• Considérons alors l'ensemble $F = \{r_1, r_2, \dots, r_p\}$ des éléments de $\llbracket 0, 1988 \rrbracket$ qui sont des résidus quadratiques modulo 1999 : c'est une partie finie et non vide de $\llbracket 0, 1988 \rrbracket$. Pour $i \in \llbracket 1, p \rrbracket$, notons $L(i) = \{a^{r_i + 1999k} \mid k \in \mathbb{N}\}$; nous savons que $L(i)$ est rationnel, d'après la question précédente. Alors L est reconnaissable en tant que réunion de la famille finie de langages $(L(i))_{1 \leq i \leq p}$.

Voici un script Maple dressant la liste des résidus quadratiques modulo n :

```

liste_residus_quadratiques := proc(n)
local E, k;
  E := {};
  for k from 0 to n-1 do E := E union {irem(k^2, n)} od;
  E;
end

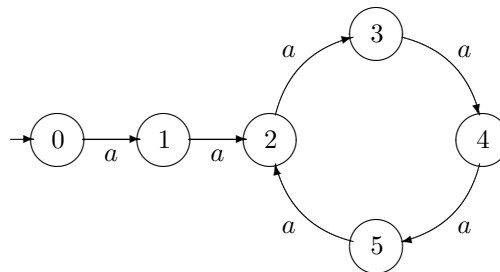
```

On constate qu'il existe 1000 résidus quadratiques modulo 1999. Ceci est une propriété générale (que vous démontrerez sans peine) : si $p > 2$ est premier, il existe exactement $\frac{p+1}{2}$ résidus quadratiques modulo p .

Question 3 • Notons $R = \{i \in P \mid i < n_0\}$ et $S = \{j \in P \mid n_0 \leq j < n_0 + p\}$; notons ensuite $\mathcal{R} = \{a^i \mid i \in R\}$ et $\mathcal{S}_j = \{a^{j+kp} \mid k \in \mathbb{N}\}$, pour $j \in S$. Alors $\mathcal{L}(P)$ est reconnaissable en tant que réunion du langage \mathcal{R} (qui est fini, donc reconnaissable) et de la famille finie $(\mathcal{S}_j)_{j \in S}$ de langages (tous reconnaissables d'après la question 1).

Question 4 • Si L est fini, il est ultimement périodique : en effet, notant m la longueur maximale d'un mot de L , il suffit de prendre $n_0 = m + 1$ et $p = 1$. Alors $a^{n+p} \in L \iff a^n \in L$ pour tout $n \geq n_0$, puisque ces deux assertions sont simultanément fausses ! L sera reconnu par l'automate (déterministe, mais non complet) (Q, δ, i, F) où $Q = \llbracket 0, m \rrbracket$, $i = 0$, $F = \{|u|, u \in L\}$ et $\delta : q \in \llbracket 0, m - 1 \rrbracket \mapsto q + 1$.

• Supposons maintenant L infini et considérons un automate fini déterministe $\mathcal{A} = (Q, \delta, i, F)$ reconnaissant \mathcal{L} . La suite de terme général $\delta^*(i, a^n)$ prend ses valeurs dans l'ensemble fini Q ; il existe donc des exposants n et m distincts tels que $\delta^*(i, a^n) = \delta^*(i, a^m)$. Notons n_0 le plus petit naturel pour lequel il existe $m > n_0$ tel que $\delta^*(i, a^n) = \delta^*(i, a^m)$, puis $p = m - n_0$. Le graphe de \mathcal{A} aura l'allure ci-dessous («poêle à frire»). Le dessin correspond au cas $n_0 = 2$, $p = 4$; on n'a pas spécifié quels états étaient finals.



Il est clair que, si $n \geq n_0$, alors $\delta^*(i, a^{n+p}) = \delta^*(i, a^n)$ si bien que $\lambda(L)$ est ultimement périodique.

Question 5 • Si $R = \{\varepsilon\}$, alors $R^* = \{\varepsilon\}$ et le PGCD cherché est 0. Sinon, l'ensemble D des diviseurs des longueurs des éléments de R est une partie de \mathbb{N}^* non vide (car contenant 1) et majorée (par la plus petite longueur d'un élément de R^*); elle possède donc un plus grand élément d qui est le PGCD cherché.

Question 6 • Par définition de d , tout élément de R appartient à $(a^d)^*$; il en est donc de même de tout élément de R^* . Il reste à prouver que $\{a^{kd} \mid k \in \mathbb{N}\} \setminus R^*$ est un ensemble fini; pour ce faire, nous allons exhiber N tel que $k \geq N$ implique $a^{kd} \in R^*$. Il est clair que l'on peut supposer $d = 1$ (quitte à découper les mots considérés en tranches de longueur d).

• On sait qu'il existe une famille $(x_i)_{1 \leq i \leq n}$ de naturels non nuls et une famille $(\alpha_i)_{1 \leq i \leq n}$ de relatifs non nuls vérifiant $\sum_{1 \leq i \leq n} \alpha_i x_i = 1$ et tels que $a^{x_i} \in R$ pour tout $i \in \llbracket 1, n \rrbracket$. En ne gardant dans le membre de gauche que

les $\alpha_i > 0$, on peut écrire ceci $\sum_{1 \leq j \leq p} \beta_j y_j = 1 + \sum_{1 \leq j \leq q} \gamma_j z_j$ où les β_j, y_j, γ_j et z_j sont des naturels non nuls, avec

$a^{y_j} \in R$ pour $1 \leq j \leq p$ et $a^{z_j} \in R$ pour $1 \leq j \leq q$. Soit alors $N = y_1 \sum_{1 \leq j \leq q} \gamma_j z_j$. Comme $a^{N+y_1} = a^N a^{y_1}$ est le

produit de deux éléments de \mathbb{R}^* , il suffit de montrer que $a^{N+k} \in R^*$ pour $k \in \llbracket 1, y_1 - 1 \rrbracket$. Or on peut écrire :

$$N + k = y_1 \sum_{1 \leq j \leq q} \gamma_j z_j + k \left(\sum_{1 \leq j \leq p} \beta_j y_j - \sum_{1 \leq j \leq q} \gamma_j z_j \right) = \sum_{1 \leq j \leq q} (y_1 - k) \gamma_j z_j + k \sum_{1 \leq j \leq p} \beta_j y_j$$

Comme $y_1 - k \geq 0$, il est clair que $a^{N+k} \in R^*$, ce qui termine la preuve.

Question 7 • En tant que langage fini, G est reconnaissable; donc $A^* \setminus G$ l'est aussi. $\{a^{kd} \mid k \in \mathbb{N}\} = (a^d)^*$ est reconnaissable, donc $\{a^{kd} \mid k \in \mathbb{N}\} \setminus G$ est reconnaissable en tant qu'intersection de deux langages reconnaissables.

Question 8 • Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe reconnaissant L . Nous allons construire un automate $\mathcal{A}' = (Q', \delta', i, F)$ qui reconnaît $\varphi(L)$. Pour construire δ' , nous distinguons pour chaque lettre $x \in X$ trois cas de figure selon la longueur de $\varphi(x)$:

- si $|\varphi(x)| = 1$, on remplace chaque transition de la forme (q, x, q') qui apparaissait dans δ par une transition $(q, \varphi(x), q')$;
- si $|\varphi(x)| = 0$, on remplace chaque transition de la forme (q, x, q') qui apparaissait dans δ par une transition instantanée (q, ε, q') ;
- si $|\varphi(x)| > 1$, alors, notant $n = |\varphi(x)|$ et $\varphi(x) = y_1 y_2 \dots y_n$, on va, pour chaque transition (q, x, q') qui apparaissait dans δ introduire de nouveaux états q_1, q_2, \dots, q_{n-1} ainsi que les transitions (q, u_1, q_1) , (q_1, u_2, q_2) et ainsi de suite jusqu'à (q_{n-1}, u_n, q') .

Q' est la réunion de Q et de l'ensemble des nouveaux états ainsi introduits. On vérifie sans peine que \mathcal{A}' reconnaît $\varphi(L)$.

• On peut aussi donner une preuve en travaillant sur les expressions rationnelles. Définissons par induction structurelle une application $\widehat{\varphi}$ de l'ensemble des expressions rationnelles sur X , sur lui-même, au moyen des règles suivantes :

- $\widehat{\varphi}(\emptyset) = \emptyset$
- $\widehat{\varphi}(\varepsilon) = \varepsilon$
- $\widehat{\varphi}(x) = \varphi(x)$ pour tout $x \in X$

- $\widehat{\varphi}(e + e') = \widehat{\varphi}(e) + \widehat{\varphi}(e')$
- $\widehat{\varphi}(e \cdot e') = \widehat{\varphi}(e) \cdot \widehat{\varphi}(e')$
- $\widehat{\varphi}(e^*) = (\widehat{\varphi}(e))^*$

On vérifie aisément que, à une expression rationnelle e décrivant un langage L , $\widehat{\varphi}$ associe une expression rationnelle $\widehat{\varphi}(e)$ décrivant le langage $\varphi(L)$.

Question 9 • Considérons le morphisme φ qui envoie toutes les lettres de X sur l'unique lettre a d'un alphabet Y . Si $L \subset X^*$ est reconnaissable, alors $\varphi(L)$ l'est aussi, donc $\lambda(\varphi(L))$ est ultimement périodique. Mais, comme le morphisme φ considéré conserve la longueur (on dit qu'il est *strictement alphabétique*), on a clairement $\lambda(\varphi(L)) = \lambda(L)$ si bien que $\lambda(L)$ est ultimement périodique.

Question 10 • Notons L ce langage, et supposons-le reconnaissable. $\lambda(L)$ serait ultimement périodique. Soient donc $n_0 \geq 0$ et $p > 0$ tels que $n \in \lambda(L) \iff n + p \in \lambda(L)$ pour tout $n \geq n_0$. Notons $\mu(n)$ la longueur de l'écriture décimale de n ; par exemple, $\mu(7!) = \mu(5040) = 4$. Choisissons n suffisamment grand pour que $\mu(n!) \geq n_0$ et $n + 1 \geq 10^{p+1}$. Nous mettons en évidence une contradiction : $(n + 1)! = (n + 1)n! \geq 10^{p+1}n!$, donc $\mu((n + 1)!) \geq p + 1 + \mu(n!) > \mu(n!) + p$; or il n'existe aucun élément de $\lambda(L)$ strictement compris entre $\mu(n!)$ et $\mu((n + 1)!)$.

Question 11 • Il est clair que ceci n'est pas possible avec un alphabet à une seule lettre ! Avec un alphabet à deux lettres a et b , il suffit de prendre le langage $\{a^n b^n \mid n \in \mathbb{N}\}$ dont il est bien connu qu'il n'est pas rationnel : $\lambda(L) = 2\mathbb{N}$ est ultimement périodique.

Question 12 • Les coefficients de cette série sont égaux à 0 ou à 1, donc son rayon de convergence est au moins égal à 1. Si L est fini, ces coefficients sont nuls APCR, donc S est un polynôme, et le rayon de convergence est infini. Sinon, comme il existe une infinité de coefficients égaux à 1, le rayon de convergence ne peut être supérieur à 1, et vaut donc exactement 1.

Question 13 • Comme L est rationnel, $\lambda(L)$ est une partie de \mathbb{N} ultimement périodique. Soient $n_0 \geq 0$ et $p > 0$ tels que $n \in \lambda(L) \iff n + p \in \lambda(L)$ pour tout $n \geq n_0$. Notons $T = \sum_{n < n_0, n \in \lambda(L)} X^n$ et $U = \sum_{n_0 \leq n < p, n \in \lambda(L)} X^n$: ces deux polynômes sont à coefficients dans \mathbb{N} (et même dans $\{0,1\}$). Et l'on a :

$$\begin{aligned}
S(x) &= \sum_{n \in \lambda(L)} x^n = \sum_{n < n_0, n \in \lambda(L)} x^n + \sum_{n \geq n_0, n \in \lambda(L)} x^n = T(x) + \sum_{n_0 \leq n < n_0 + p, n \in \lambda(L)} \sum_{k \in \mathbb{N}} x^{n+kp} \\
&= T(x) + \sum_{n_0 \leq n < n_0 + p, n \in \lambda(L)} x^n \sum_{k \in \mathbb{N}} (x^p)^k = T(x) + \frac{1}{1 - x^p} \sum_{n_0 \leq n < n_0 + p, n \in \lambda(L)} x^n \\
&= T(x) + \frac{U(x)}{1 - x^p}
\end{aligned}$$

Il est clair que $\deg(T) < n_0$ et que $U = X^{n_0}V$, avec $\deg(V) < p$.

FIN