Option Informatique en Spé MP et MP*
Lemme de pompage et lemme de non-pompage : le corrigé

Question 1 • Soient L un langage reconnaissable, et $\mathcal{A} = (Q, \delta, i, F)$ un automate fini reconnaissant L ; notons $N = |Q|$. Considérons un mot u vérifiant $|u| \geqslant N$. Au cours de la lecture de ce mot, l'automate passe par $N + 1$ états au moins. D'après le principe des tiroirs, il existe au moins un état j par lequel l'automate passe au moins deux fois. Ceci revient à dire que u peut se décomposer en xyz, avec $y \neq \varepsilon$ et $\delta^*(i, xy) = \delta^*(i, x) = j$. Une récurrence immédiate montre que $\delta^*(i, xy^n) = j$; on en déduit $\delta^*(xy^n z) = \delta^*(xyz)$ et donc $xy^n z \in L$ pour tout $n \in \mathbb{N}$.

• La condition $|xy| \leqslant N$ sera obtenue en choisissant j comme suit : pour $0 \leqslant k \leqslant |u|$, notons $q_k = \delta^*(i, u_1 u_2 \dots u_k)$ l'état dans lequel se trouve l'automate après lecture du préfixe de longueur k de u ; alors $j = \min\{k : 0 \leqslant k \leqslant |u| \text{ et } \mathrm{Card}\{q_0, q_1, \dots, q_k\} \leqslant k\}$. Autrement dit : tant que l'on n'a pas lu j lettres de u, on n'est pas passé deux fois par le même état.

Question 2 • Procédons par l'absurde. Supposons L_1 reconnaissable : le lemme de l'étoile affirme l'existence d'un naturel N tel que tout mot $u \in L_1$, de longueur au moins égale à N, se décompose en $u = xyz$ avec $y \neq \varepsilon$, $|xy| \leqslant N$ et $xy^n z \in L_1$ pour tout $n \in \mathbb{N}$. Comme $(N + 2)! \geqslant N$, nous pouvons appliquer ce lemme au mot $u = a^{(N+1)!}$; soit donc xyz la factorisation de u dont le lemme affirme l'existence ; on aura $1 \leqslant y \leqslant N$, donc $(N + 1)! + 1 \leqslant |xy^2 z| \leqslant (N + 1)! + N$. Il est clair que $(N + 1)! < (N + 1)! + 1$; par ailleurs, $N + 1 > N$ et $(N + 1)! \geqslant 1$ impliquent :

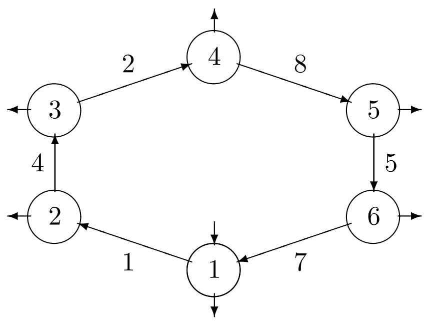
$$(N + 2)! = (N + 2)(N + 1)! = (N + 1)! + (N + 1)(N + 1)! > (N + 1)!$$

On a donc $(N + 1)! < |xy^2 z| < (N + 2)!$, si bien que $xy^2 z \notin L_2$. Ceci contredit la conclusion du lemme de l'étoile, et montre donc que L_2 n'est pas reconnaissable.

Question 3 • Le développement décimal illimité de $1/7$ est $0{,}142857142857\dots$; donc $\mathrm{Lang}(1/7)$ est décrit par l'expression rationnelle suivante :

$$(142857)^*(\varepsilon + 1 + 14 + 142 + 1428 + 14285)$$

Ceci prouve que ce langage est rationnel. En prime, voici un automate reconnaissant ce langage :



Question 4 • Procédons par l'absurde. Supposons $\mathrm{Lang}(\sqrt{2})$ reconnaissable : le lemme de l'étoile affirme l'existence d'un naturel N tel que tout mot $u \in \mathrm{Lang}(\sqrt{2})$, de longueur au moins égale à N, se décompose en $u = xyz$ avec $y \neq \varepsilon$, $|xy| \leqslant N$ et $xy^n z \in \mathrm{Lang}(\sqrt{2})$ pour tout $n \in \mathbb{N}$. Notons $p = |x|$ et $q = |y|$; alors le mot infini $d(\sqrt{2}) = d_1 d_2 d_3 \dots$ vérifie $d_i = x_i$ pour $1 \leqslant i \leqslant p$, puis $d_j = d_{j+q}$ pour tout $j > p$: pour le voir, il suffit de considérer un préfixe de la forme xy^n suffisamment long pour que $p + nq \geqslant j$. Ainsi, le développement décimal de $\sqrt{2}$ serait périodique à partir d'un certain rang, or ceci est contradictoire puisque $\sqrt{2}$ n'est pas rationnel.

Question 5 • Soient L un langage reconnaissable et $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe reconnaissant L. Dire que L est préfixiel revient à dire que tout préfixe d'un mot de L est encore dans L ; sur l'automate, ceci revient à dire que tous les états traversés lors d'un calcul réussi sont des états d'acceptation. Si l'on a pris le soin d'émonder $\mathcal{A}$, la condition se réduit à $F = Q$, puisque dans ce cas tout état participe à au moins un calcul réussi.

L'algorithme se résume donc comme suit : on émonde l'automate ; on vérifie ensuite qu'il ne reste que des états d'acceptation.

Question 6 • Si $y = \varepsilon$, alors $y^m z = z = y^n z$ quels que soient m et n. On peut donc prendre $m = 2$ et $n = 1$.

Question 7 La suite de terme général $q_k = \delta^*(i, y^k)$ $(k > 0)$ est à valeurs dans l'ensemble fini Q ; il existe donc des exposants m et n distincts, tels que $q_m = q_n$; on peut supposer $m > n > 0$ pour fixer les idées. Alors

$$\begin{aligned} y^m z \in L &\iff \delta^*(i, y^m z) \in F \iff \delta^*\big(\delta^*(i, y^m), z\big) \in F \\ &\iff \delta^*\big(\delta^*(i, y^n), z\big) \in F \iff \delta^*(i, y^n z) \in F \iff y^n z \in L \end{aligned}$$

Ce qui termine la preuve du lemme.

Question 8 • Supposons L_2 reconnaissable et considérons le mot $y = a$. Le lemme de non-pompage assure l'existence de naturels m et n tels que $m > n > 0$ et $y^m z \in L \iff y^n z \in L$ quel que soit le mot z. Prenons $z = a^{m^2+m+1}$: alors $y^m z = a^m a^{m^2+m+1} = a^{m^2+2m+1} = a^{(m+1)^2}$ appartient à L_2. En revanche, $y^n z = a^n a^{m^2+m+1} = a^{m^2+m+n+1}$ n'est pas élément de L_2 ; en effet :

$$m^2 < m^2 + m + n + 1 < m^2 + 2m + 1 = (m+1)^2$$

Par conséquent, $m^2 + m + n + 1$ n'est pas un carré parfait.

Question 9 • Supposons L_3 reconnaissable et considérons le mot $y = ab$. Le lemme de non-pompage assure l'existence de naturels m et n tels que $m > n > 0$ et $y^m z \in L \iff y^n z \in L$ quel que soit le mot z. Prenons $z = (ba)^n a$: alors $y^n z = (ab)^n (ba)^n a$ appartient à L_3, il suffit de prendre $x = (ab)^n$ et $w = a$ pour s'en convaincre. En revanche, $y^m z = (ab)^m (ba)^n a$ n'admet aucune décomposition de la forme $vv^R w$ avec v et w non vides, donc $y^m z \notin L_3$.

Question 10 • La réponse est négative. Soit $N = 4$. Considérons un mot $vv^R w$ de L_3, de longueur au moins égale à N. Comme $v \neq \varepsilon$, on peut écrire $v = au$ où a est une lettre ; deux cas de figure se présentent. Si $u \neq \varepsilon$, prenons $x = \varepsilon$, $y = a$ et $z = uu^R aw$; on a bien $|xy| = 1 \leqslant N$, $y \neq \varepsilon$ et $xy^n z \in L_3$ pour tout $n \in \mathbb{N}$ puisque $xy^0 z = uu^R aw$ (avec $u \neq \varepsilon$ et $aw \neq \varepsilon$) et $xy^n z = aa^R a^{n-2} uu^R aw$ si $n \geqslant 2$ (avec $a \neq \varepsilon$ et $a^{n-2} uu^R aw \neq \varepsilon$). Si $u = \varepsilon$, prenons $x = a^2$, $y = b$ et $z = w'$ où b est la première lettre de w et w' est tel que $w = bw'$; on a bien $|xy| = 3 \leqslant N$, $y \neq \varepsilon$, et $xy^n z \in L_3$ pour tout $n \in \mathbb{N}$ puisque ce mot s'écrit $aa^R y^n z$ avec $a \neq \varepsilon$ et $y^n z \neq \varepsilon$. Ainsi, L_3 vérifie la *conclusion* du lemme de l'étoile ; on ne peut donc pas espérer utiliser celui-ci pour montrer que ce langage n'est pas reconnaissable.

Question 11 • La réponse est négative. Considérons un mot $y \in \mathrm{Lang}(\sqrt{2})$. Comme le développement décimal de $\sqrt{2}$ n'est pas périodique, il existe un exposant $n > 0$ tel que $y^n \notin \mathrm{Lang}(\sqrt{2})$. Alors, comme ce langage est préfixiel, il ne contient pas non plus y^{n+1}. Notons $m = n + 1$; on a $m > n > 0$. Aucun des deux mots $y^m z$ et $y^n z$ n'appartient à $\mathrm{Lang}(\sqrt{2})$ puisque ce dernier est préfixiel. Ainsi, ce langage vérifie la *conclusion* du lemme de non-pompage ; on ne peut donc pas espérer utiliser celui-ci pour montre que ce langage n'est pas reconnaissable.

FIN