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

► On rappelle l'énoncé du lemme de l'étoile :

> Soit L un langage reconnaissable ; il existe un naturel N tel que tout mot u de L, de longueur au moins égale à N, se décompose en $u = xyz$ avec $y \neq \varepsilon$, $|xy| \leqslant N$ et $xy^n z \in L$ pour tout $n \in \mathbb{N}$.

Question 1 • Donnez une démonstration du lemme de l'étoile.

Question 2 • Utilisez le lemme de l'étoile pour montrer que le langage $L_1 = \{a^{n!} \mid n \in \mathbb{N}\}$ n'est pas reconnaissable.

► On rappelle que tout réel x possède un et un seul développement décimal propre (c'est-à-dire ne se finissant pas par une infinité de 9) ; et que ce développement est périodique à partir d'un certain rang si et seulement si x est rationnel.

► Soit α un réel. Notons $d(\alpha) = d_1 d_2 d_3 \ldots$ le développement décimal propre de $\alpha - \lfloor \alpha \rfloor$ (le zéro et la virgule étant omis), considéré comme un mot infini sur l'alphabet formé des chiffres 0 à 9. Par exemple, $d(1/3) = 333\ldots$ Notons $\mathrm{Lang}(\alpha)$ le langage formé par les préfixes (finis) de ce mot infini ; par exemple, $\mathrm{Lang}(e) = \{\varepsilon, 7, 71, 718, 7182, 71828, \ldots\}$ puisque $e = 2{,}71828\ldots$

Question 3 • Montrez que $\mathrm{Lang}(1/7)$ est reconnaissable ; pour ce faire, vous construirez un automate fini déterministe complet reconnaissant ce langage.

Question 4 • Utilisez le lemme de l'étoile pour montrer que $\mathrm{Lang}(\sqrt{2})$ n'est pas reconnaissable.

► Il est clair que, si $\mathrm{Lang}(\alpha)$ contient le mot u, il contient aussi tous les préfixes de u. Un langage qui possède cette propriété est dit *préfixiel*.

Question 5 • Montrez que l'on peut décider (au moyen d'un algorithme que vous décrirez brièvement) si un langage reconnaissable L donné est préfixiel.

► Dans la littérature, le lemme de l'étoile est aussi désigné sous le nom de *lemme de pompage* (en anglais : *pumping lemma*) puisque son application permet, une fois que l'on a «amorcé» avec le facteur y, de «pomper» indéfiniment celui-ci en n'obtenant que des mots du langage.

► Guo-Qiang Zhang et E. Rodney Canfield ont présenté un *lemme de non-pompage*, dont voici l'énoncé :

> Soit L un langage reconnaissable ; à tout mot y, on peut associer des naturels m et n vérifiant $m > n > 0$ et tels que $y^m z \in L \iff y^n z \in L$ pour tout mot z.

Nous allons commencer par démontrer ce lemme. Notons $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe reconnaissant L. La fonction de transition étendue est notée δ^* ; elle est définie par $\delta^*(q, \varepsilon) = q$ et $\delta^*(q, ux) = \delta\big(\delta^*(q, u), x\big)$ pour $q \in Q$, $x \in A$ et $u \in A^*$.

Question 6 • Expliquez l'affirmation suivante : dans la preuve du lemme de non-pompage, il n'est pas nécessaire de considérer le cas $y = \varepsilon$.

Question 7 • Soit $y \neq \varepsilon$. En observant la suite de terme général $q_k = \delta^*(i, y^k)$, terminez la preuve du lemme de non-pompage.

Question 8 • Utilisez le lemme de non-pompage pour montrer que le langage L_2 formé des mots sur l'alphabet A dont la longueur est un carré parfait, n'est pas reconnaissable.

Question 9 • On note u^R le *miroir* du mot u, défini par les relations $\varepsilon^R = \varepsilon$ et :

$$(u_1 u_2 \ldots u_n)^R = u_n u_{n-1} \ldots u_2 u_1$$

Utilisez le lemme de non-pompage pour montrer que le langage suivant n'est pas reconnaissable :

$$L_3 = \{v v^R w \mid v, w \in A^+\}$$

Question 10 • Peut-on utiliser le lemme de l'étoile pour montrer que le langage L_3 n'est pas reconnaissable ?

Question 11 ⋆⋆⋆ • Peut-on utiliser le lemme de non-pompage pour montrer que le langage $\mathrm{Lang}(\sqrt{2})$ n'est pas reconnaissable ?

FIN