Option Informatique en Spé MP et MP*
Miroir d'un facteur : le corrigé

Question 1 • La relation $\rightarrow$ n'est pas transitive : prenons $\Sigma = \{a,b\}$, $u = ababb$, $v = abbab$ et $w = babba$. Clairement, $u = ab(ab)b \rightarrow vab(\widetilde{ab})b$ et $v = abbab \rightarrow w = \widetilde{v}$; pourtant, on n'a pas $u \rightarrow w$: aucune des trois transformations amenant un b en tête ne convient.

Question 2 • Soient x, y et z tel que $u = xyz$ et $v = x\widetilde{y}z$. Notant $x' = \widetilde{x}$, $y' = \widetilde{y}$ et $z' = \widetilde{z}$ nous aurons $\widetilde{u} = z'y'x'$ et $\widetilde{v} = z'yx' = z'\widetilde{y'}x'$ donc $\widetilde{u} \rightarrow \widetilde{v}$.

Question 3 • Le sens direct est clair, et il suffit d'établir la réciproque lorsque $t \in \Sigma$. Soient donc x, y et z tels que $tu = xyz$ et $tv = x\widetilde{y}z$. Si $x \neq \varepsilon$, alors $x = tx'$ avec $|x'| = |x| - 1$, puis $tu = tx'yz$ et $tv = tx'\widetilde{y}z$ donc $u = x'yz$ et $v = x'\widetilde{y}z$, si bien que $u \rightarrow v$. Si $x = \varepsilon$, alors $tu = yz$ et $tv = \widetilde{y}z$. Ainsi, t est la première lettre de y et de $\widetilde{y}$; c'est donc aussi la dernière lettre de y. Si $|y| = 1$, alors $y = t$, puis $u = v = z$ donc $u \rightarrow v$. Sinon, $y = ty't$ puis $tu = ty'tz$ et $tv = \widetilde{ty't}z = t\widetilde{y'}tz$; nous en déduisons $u = y'tz$ et $v = \widetilde{y'}tz$ ce qui montre que $u \rightarrow v$.

Question 4 • En appliquant répétitivement le résultat précédent, nous pouvons nous ramener au cas où u et v n'ont ni préfixe commun, ni suffixe commun. Dans ces conditions, $u \rightarrow v$ ssi $u = \widetilde{v}$. Ceci peut être effectué pour un coût $\mathcal{O}(n)$: notant p (resp. q) la longueur du plus long préfixe (resp. suffixe) commun à u et v, on effectue $p + q + 2$ comparaisons de caractères pour les éliminer ; puis $n - p - q$ comparaisons pour décider si $u[p + 1..n - q]$ et $v[p + 1..n - q]$ sont miroirs l'un de l'autre.

Question 5 • Comme nous travaillons sur des listes, il nous faut les retourner pour éliminer leur plus long suffixe commun. Le coût total sera donc de $2n$ opérations ≪Cons≫ au lieu de n.

```
let rec efface_prefixe_commun = function
 | (t::x,u::y) when t=u -> efface_prefixe_commun (x,y)
 | c -> c ;;

let equivalents u v =
 let (u1,v1) = efface_prefixe_commun (u,v) in
 let (u2,v2) = (rev u1,rev v1) in
 let (u3,v3) = efface_prefixe_commun (u2,v2) in
  rev(u3) = v3 ;;
```

Question 6 • La réponse est positive. Soit $\mathcal{A} = (Q, \delta, I, F)$ un automate fini (pas forcément déterministe) reconnaissant L. Soient q et q' deux états de $\mathcal{A}$; nous noterons $L_{q,q'}$ l'ensemble des mots u tels que $q \xrightarrow{u} q'$: il est clair que ce langage est rationnel, et son miroir $\widetilde{L_{q,q'}}$ l'est donc également.

• Nous allons montrer que $\Phi(L)$ est la réunion des langages $L_{i,q} \cdot \widetilde{L_{q,q'}} \cdot L_{q',f}$ où $i \in I$, $f \in F$ et $q, q' \in Q$. Ces langages sont en nombre fini, plus précisément $|Q|^4$; et chacun d'eux est rationnel en tant que produit de rationnels. Donc $\Phi(L)$ sera rationnel.

• Soit $v \in \Phi(L)$: il existe $x, y, z \in \Sigma^*$ tels que $xyz \in L$ et $v = x\widetilde{y}z$. Observons un calcul réussi de $\mathcal{A}$ d'étiquette xyz ; il est de la forme $i \xrightarrow{x} q \xrightarrow{y} q' \xrightarrow{z} f$ avec $i \in I$ et $f \in F$. Alors $v = x\widetilde{y}z$ appartient au produit $L_{i,q} \cdot \widetilde{L_{q,q'}} \cdot L_{q',f}$.

• Réciproquement, considérons un mot v appartenant à la réunion décrite plus haut. Il existe donc $i \in I$, $f \in F$ et $q, q' \in Q$ tels que $v \in L_{i,q} \cdot \widetilde{L_{q,q'}} \cdot L_{q',f}$. Donc $v = xyz$ avec $x \in L_{i,q}$, $y \in \widetilde{L_{q,q'}}$ et $z \in L_{q',f}$. Notons $u = x\widetilde{y}z$: ce mot est l'étiquette d'un calcul réussi de $\mathcal{A}$, à savoir $i \xrightarrow{x} q \xrightarrow{\widetilde{y}} q' \xrightarrow{z} f$. Donc u appartient à L, et finalement v appartient à $\Phi(L)$.

Question 7 • Prenons comme alphabet $\Sigma = \{a,b\}$. Alors $L = \{a^n b a^n \mid n \in \mathbb{N}\}$ répond à la question. Il est clair qu'il n'est pas rationnel ; et $\Phi(L) = \{a^p b a^q \mid p + q \text{ est pair}\}$ est rationnel.

Question 8 • Il suffit de choisir un langage non rationnel L tel que $\Phi(L) = L$. Ainsi, avec $\Sigma = \{a,b\}$ le langage $L = \{u \in \Sigma^* \mid |u|_a = |u|_b\}$ convient.

Question 9 • Il est clair que, si $u \rightarrow v$, alors $|u|_x = |v|_x$ pour toute lettre $x \in \Sigma$; il en est donc de même lorsque $u \xrightarrow{*} v$. Montrons que cette condition est suffisante : il suffit de prouver que, si $\Sigma = \{x_1, \ldots, x_p\}$, alors le mot $u = u_1 \ldots u_n$ est équivalent au mot ${x_1}^{|u|_{x_1}} \ldots {x_p}^{|u|_{x_p}}$. Prouvons-le par récurrence sur n ; le résultat est clair si $n = 0$ ou $n = 1$. Supposons-le acquis au rang n, et considérons un mot u de longueur $n+1$: $u = x_j v$ avec $x_j \in \Sigma$ et $|v| = n$. L'hypothèse de récurrence s'applique à v : $v \xrightarrow{*} {x_1}^{|v|_{x_1}} \ldots {x_p}^{|v|_{x_p}}$. En utilisant les relations

$|u|_{x_i} = |v|_{x_i} + \delta_{i,j}$ et $xy^p \to y^p x$, nous aurons clairement :

$$\begin{aligned} u &= x_j {x_1}^{|v|_{x_1}} \dots {x_p}^{|v|_{x_p}} \to {x_1}^{|v|_{x_1}} x_j {x_2}^{|v|_{x_2}} \dots {x_p}^{|v|_{x_p}} \xrightarrow{*} {x_1}^{|v|_{x_1}} x_j {x_{j-1}}^{|v|_{x_{j-1}}} {x_j}^{|v|_{x_j}} \dots {x_p}^{|v|_{x_p}} \\ &\to {x_1}^{|v|_{x_1}} {x_{j-1}}^{|v|_{x_{j-1}}} x_j {x_j}^{|v|_{x_j}} \dots {x_p}^{|v|_{x_p}} = {x_1}^{|v|_{x_1}} {x_{j-1}}^{|v|_{x_{j-1}}} {x_j}^{1+|v|_{x_j}} \dots {x_p}^{|v|_{x_p}} \\ &= {x_1}^{|u|_{x_1}} {x_{j-1}}^{|u|_{x_{j-1}}} {x_j}^{|u|_{x_j}} \dots {x_p}^{|u|_{x_p}} \end{aligned}$$

Question 10 • D'après l'analyse effectuée à la question précédente, il existe autant de classes de mots de longueur n modulo $\xrightarrow{*}$ que de triplets (i, j, k) de naturels vérifiant $i + j + k = n$; ce nombre est :

$$\sum_{0\leqslant i\leqslant n} \sum_{0\leqslant j\leqslant n-i} 1 = \sum_{0\leqslant i\leqslant n} (n - i + 1) = \sum_{1\leqslant i\leqslant n+1} i = \frac{(n+1)(n+2)}{2}$$

Question 11 • Montrons que le langage $L = \{a^n ba^n ba^n ba^n \mid n \in \mathbb{N}\}$ répond à la question.

• L n'est pas rationnel : s'il l'était, le lemme de l'étoile assurerait l'existence d'une constante N telle que tout mot u de L de longueur au moins égale à N possède une factorisation $u = xyz$ vérifiant $|xy| \leqslant N$, $y \neq \varepsilon$ et $xy^k z \in L$ pour tout $k \in \mathbb{N}$. Choisissons alors $u = a^N ba^N ba^N ba^N$; xy est préfixe de a^N, donc $x = a^i$ et $y = a^j$ avec $j \neq 0$, donc $z = a^{N-i-j} ba^N ba^N ba^N$. Mais alors le mot $xz = a^{N-j} ba^N ba^N ba^N$ n'appartient pas à L car $N - j < N$.

• Montrons que $\Phi(L)$ n'est pas rationnel. L'idée est que tout mot de ce langage est de la forme $a^\alpha ba^\beta ba^\gamma ba^\delta$, deux des exposants α, β, γ et δ devant être égaux. Supposons $\Phi(L)$ rationnel, et notons N la constante fournie par le lemme de l'étoile. Soit $v = a^N ba^N ba^N ba^N$, qui appartient à L. La factorisation donnée par le lemme de l'étoile est de la forme xyz avec $x = a^i$, $y = a^j$ et $z = a^{N-i-j} ba^N ba^N ba^N$, avec $j \neq 0$. Notons $v = xz = a^{N-j} ba^N ba^N ba^N$ et prouvons que ce mot n'appartient pas à $\Phi(L)$. C'est clair si j n'est pas multiple de 4, puisque Φ conserve le nombre d'occurrences de a dans un mot. Soient donc $p = N - j/4$, $w = a^p ba^p ba^p ba^p$ et w' un mot tel que $w \to w'$: il existe des mots r, s et t tels que $w = rst$ et $w' = r\widetilde{s}t$. Si $|s|_b = 0$, ou si $s = a^k ba^k$ ou $s = a^k bs' ba^k$, alors $w' = w$ donc $w' \neq v$. Il reste six cas à envisager selon les b qui apparaissent dans s. Examinons un de ces cas : si s contient le premier b de $'$, alors $r = a^{p-j}$, $s = a^j ba^k$ et $t = a^{p-k} ba^p ba^p$ avec $j \neq k$; par suite $w' = a^{p-j+k} na^{p-k+j} ba^p ba^p$ donc $w' \neq v$. On se convaincra aisément que la situation est la même dans les cinq autres cas : w' est de la forme $a^p ba^q ba^r ba^s$

• Il résulte de la question 9 que $\Phi^*(L)$ est l'ensemble des mots $u \in \Sigma^*$ tels que $|u|_b = 3$ et $|u|_a$ est multiple de 4 ; il est clair que ce langage est rationnel.

FIN