Option Informatique en Spé MP et MP*
DS du jeudi 17 février 2005 : le corrigé

Question 1 • Soit $u \in L$; nous savons que u admet au moins une décomposition $u = vabw$. Considérons celle pour laquelle $|v|$ est minimale : le mot v ne contient pas le facteur ab, donc cette décomposition répond à la question. Notons que le mot v s'écrit $b^i a^j$, avec $i \geqslant 0$ et $j \geqslant 0$.

▶ Remarquons que L est stable pour la concaténation, donc la question 2 a un sens ; mieux : nous avons clairement $L = \Sigma^* L \Sigma^*$, donc L est un *idéal* de Σ^*.

Question 2 • φ n'est pas un morphisme, car $\varphi(abab) = baab$, tandis que $\varphi(ab)\varphi(ab) = baba$.

Question 3 • φ n'est pas injectif, car $\varphi(abba) = \varphi(baab) = baba$.

Question 4 • Montrons que $\varphi(L) = \{vbaw \mid v, w \in \Sigma^*\}$. Le sens $\subset$ est clair. Pour la réciproque, considérons un mot de la forme $vbaw$; nous pouvons supposer que $|v|$ est minimal ; alors v ne contient pas le facteur ba, si bien que $vbaw = \varphi(vabw)$.

▶ $\boldsymbol{i}$ est l'état initial de l'automate $\mathcal{A}$.

Question 5 • Notons $Q' = Q \times [\![1, 4]\!]$, $i' = (\boldsymbol{i}, 1)$ et $F' = F \times \{4\}$. Définissons δ' comme suit :

- pour chaque transition $p \xrightarrow{a} q$ de $\mathcal{A}$, δ' contient les transitions $(p,1) \xrightarrow{a} (q,2)$, $(p,1) \xrightarrow{b} (q,3)$, $(p,2) \xrightarrow{a} (q,2)$, $(p,2) \xrightarrow{b} (q,3)$, $(p,4) \xrightarrow{a} (q,4)$;
- pour chaque transition $p \xrightarrow{b} q$ de $\mathcal{A}$, δ' contient les transitions $(p,1) \xrightarrow{b} (q,1)$, $(p,3) \xrightarrow{a} (q,4)$, $(p,4) \xrightarrow{b} (q,4)$.

Je dis que $\mathcal{B} = (Q', \delta', i', F')$ reconnait $\varphi(P)$.

• Soit $u \in \varphi(P)$; d'après la question 1, il existe $i \geqslant 0$, $j \geqslant 0$ et $w \in \Sigma^*$ tels que $b^i a^j abw \in P$ et $u = \varphi(b^i a^j abw) = b^i a^j baw$. Notons $r = |w|$. Nous allons distinguer plusieurs cas de figure.

- $i = 0$ et $j = 0$: abw est reconnu par un calcul de $\mathcal{A}$ de la forme $q_0 = \boldsymbol{i} \xrightarrow{a} q_1 \xrightarrow{b} q_2 \xrightarrow{w_1} q_3 \cdots \xrightarrow{w_r} q_{r+2}$. Alors baw est reconnu par le calcul suivant de $\mathcal{B}$: $(q_0,1) \xrightarrow{b} (q_1,3) \xrightarrow{a} (q_2,4) \xrightarrow{w_1} (q_3,4) \cdots \xrightarrow{w_r} (q_{r+2},4)$. Ceci inclut le cas où $u = ba$.
- $i = 0$ et $j > 0$: $a^{j+1}bw$ est reconnu par un calcul de $\mathcal{A}$ de la forme $q_0 = \boldsymbol{i} \xrightarrow{a} q_1 \cdots \xrightarrow{a} q_{j+1} \xrightarrow{b} q_{j+2} \xrightarrow{w_1} q_{j+3} \cdots \xrightarrow{w_r} q_{j+r+2}$. Alors $a^j baw$ est reconnu par le calcul suivant de $\mathcal{B}$: $(q_0,1) \xrightarrow{a} (q_1,2) \cdots \xrightarrow{a} (q_j,2) \xrightarrow{b} (q_{j+1},3) \xrightarrow{a} (q_{j+2},4) \xrightarrow{w_1} (q_{j+3},4) \cdots \xrightarrow{w_r} (q_{j+r+2},4)$. Ceci inclut le cas où $u = a^j ba$.
- $i > 0$ et $j = 0$: $b^i abw$ est reconnu par un calcul de $\mathcal{A}$ de la forme $q_0 = \boldsymbol{i} \xrightarrow{b} q_1 \cdots \xrightarrow{b} q_i \xrightarrow{a} q_{i+1} \xrightarrow{b} q_{i+2} \xrightarrow{w_1} q_{i+3} \cdots \xrightarrow{w_r} q_{i+r+2}$. Alors $b^{i+1}aw$ est reconnu par le calcul suivant de $\mathcal{B}$: $(q_0,1) \xrightarrow{b} (q_1,1) \cdots \xrightarrow{b} (q_i,1) \xrightarrow{b} (q_{i+1},3) \xrightarrow{a} (q_{i+2},4) \xrightarrow{w_1} (q_{i+3},4) \cdots \xrightarrow{w_r} (q_{i+r+2},4)$. Ceci inclut le cas où $u = b^{i+1}a$.
- $i > 0$ et $j > 0$: $b^i a^{j+1} bw$ est reconnu par un calcul de $\mathcal{A}$ de la forme $q_0 = \boldsymbol{i} \xrightarrow{b} q_1 \cdots \xrightarrow{b} q_i \xrightarrow{a} q_{i+1} \cdots \xrightarrow{a} q_{i+j} \xrightarrow{a} q_{i+j+1} \xrightarrow{b} q_{i+j+2} \xrightarrow{w_1} q_{i+j+3} \cdots \xrightarrow{w_r} q_{i+j+r+1}$. Alors $b^i a^j baw$ est reconnu par le calcul suivant de $\mathcal{B}$: $(q_0,1) \xrightarrow{b} (q_1,1) \cdots \xrightarrow{b} (q_i,1) \xrightarrow{a} (q_{i+1},2) \xrightarrow{a} (q_{i+2},2) \cdots \xrightarrow{a} (q_{i+j},2) \xrightarrow{b} (q_{i+j+1},3) \xrightarrow{a} (q_{i+j+2},4) \xrightarrow{w_1} (q_{i+j+3},4) \cdots \xrightarrow{w_r} (q_{i+j+r+1},4)$. Ceci inclut le cas où $u = b^i a^{j+1} a$.

Concluons : dans tous les cas, u est reconnu par $\mathcal{B}$.

• Soit u reconnu par $\mathcal{B}$. Observons un calcul réussi de $\mathcal{B}$, d'étiquette u : ce calcul commence dans l'état $(\boldsymbol{i}, 1)$ et se termine dans un état $(f, 4)$ avec $f \in F$; il *doit* donc emprunter une transition $(s,3) \xrightarrow{a} (t,4)$. Une fois atteint l'état $(t,4)$, le calcul ne peut passer que par des états appartenant à $Q \times \{4\}$. Pour atteindre l'état $(s,3)$, il y a deux possibilités :

- passer par au moins un état $(p,2)$; le calcul emprunte $i \geqslant 0$ transitions $(x,1) \xrightarrow{b} (y,1)$; puis une transition $(x,1) \xrightarrow{a} (y,2)$; puis $j \geqslant 0$ transitions $(x,2) \xrightarrow{a} (y,2)$; puis une transition $(x,2) \xrightarrow{b} (s,3)$; la fin du calcul a été décrite plus haut. Nous avons $u = b^i a^{j+1} baw$, où w est l'étiquette de la partie du calcul qui commence dans l'état $(t,4)$. De façon détaillée, le calcul de $\mathcal{B}$ a l'allure suivante : $(q_0,1) \xrightarrow{b} (q_1,1) \cdots \xrightarrow{b} (q_i,1) \xrightarrow{a} (q_{i+1},2) \xrightarrow{a} (q_{i+2},2) \cdots \xrightarrow{a} (q_{i+j+1},2) \xrightarrow{b} (q_{i+j+2},3) \xrightarrow{a} (q_{i+j+3},4) \xrightarrow{w_1} (q_{i+j+4},4) \cdots \xrightarrow{w_r} (q_{i+j+r+3},4)$ avec $q_0 = \boldsymbol{i}$, $q_{i+j+2} = s$, $q_{i+j+3} = t$ et $q_{i+j+r+3} \in F$. Dans ce cas, $u = \varphi(z)$, où $z = b^i a^{j+2} bw$ est reconnu par le calcul suivant de $\mathcal{A}$: $q_0 \xrightarrow{b} q_1 \cdots \xrightarrow{b} q_i \xrightarrow{a} q_{i+1} \cdots \xrightarrow{a} q_{i+j+1} \xrightarrow{a} q_{i+j+2} \xrightarrow{b} q_{i+j+3} \xrightarrow{w_1} q_{i+j+4} \cdots \xrightarrow{w_r} q_{i+j+r+3}$; justifions cette affirmation :

  - $q_0 = \boldsymbol{i}$ et $q_{i+j+r+3} \in F$;
  - pour $0 \leqslant k < i$, $q_k \xrightarrow{b} q_{k+1}$ est une transition de $\mathcal{A}$ parce que $(q_k,1) \xrightarrow{b} (q_{k+1},1)$ est une transition de $\mathcal{B}$;
  - $q_i \xrightarrow{b} q_{i+1}$ est une transition de $\mathcal{A}$ parce que $(q_i,1) \xrightarrow{a} (q_{i+1},2)$ est une transition de $\mathcal{B}$;
  - pour $1 \leqslant k \leqslant j$, $q_{i+k} \xrightarrow{a} q_{i+k+1}$ est une transition de $\mathcal{A}$ parce que $(q_{i+k},2) \xrightarrow{a} (q_{i+k+1},2)$ est une transition de $\mathcal{B}$;
  - $q_{i+j+1} \xrightarrow{a} q_{i+j+2}$ est une transition de $\mathcal{A}$ parce que $(q_{i+j+1},2) \xrightarrow{b} (q_{i+j+2},3)$ est une transition de $\mathcal{B}$;
  - $q_{i+j+2} \xrightarrow{b} q_{i+j+3}$ est une transition de $\mathcal{A}$ parce que $(q_{i+j+2},3) \xrightarrow{a} (q_{i+j+3},4)$ est une transition de $\mathcal{B}$;
  - pour $1 \leqslant k \leqslant r$, $q_{i+j+k+2} \xrightarrow{w_k} q_{i+j+k+3}$ est une transition de $\mathcal{A}$ parce que $(q_{i+j+k+2},4) \xrightarrow{w_k} (q_{i+j+k+3},4)$ est une transition de $\mathcal{B}$.

- ne passer par aucun état $(p,2)$; le calcul emprunte $i \geqslant 0$ transitions $(x,1) \xrightarrow{b} (y,1)$; puis une transition $(x,1) \xrightarrow{b} (s,3)$; la fin du calcul a été décrite plus haut. Nous avons cette fois $u = b^{i+1} aw$, où w est défini comme dans le cas précédent. De façon détaillée, le calcul de $\mathcal{B}$ a l'allure suivante : $(q_0,1) \xrightarrow{b} (q_1,1) \cdots \xrightarrow{b} (q_i,1) \xrightarrow{b} (q_{i+1},3) \xrightarrow{a} (q_{i+2},3) \xrightarrow{w_1} (q_{i+3},4) \cdots \xrightarrow{w_r} (q_{i+r+2},4)$ avec $q_0 = \boldsymbol{i}$, $q_{i+1} = s$, $q_{i+2} = t$ et $q_{i+r+2} \in F$. Dans ce cas, $u = \varphi(z)$, où $z = b^i abw$ est reconnu par le calcul suivant de $\mathcal{A}$: $q_0 \xrightarrow{b} q_1 \cdots \xrightarrow{b} q_i \xrightarrow{a} q_{i+1} \xrightarrow{b} q_{i+2} \xrightarrow{w_1} q_{i+3} \cdots \xrightarrow{w_r} q_{i+r+2}$; la justification détaillée de cette dernière affirmation est semblable à celle donnée, plus haut, pour le premier cas, et n'est donc pas détaillée ici.

▶ Nous noterons $\deg(u)$ le degré du mot u.

Question 6 • Notons $u \to v$ pour indiquer que $\varphi(u) = v$. Alors :

$$ba\mathbf{ab}abb \to b\mathbf{ab}aabb \to bbaa\mathbf{ab}b \to bba\mathbf{ab}ab \to bb\mathbf{ab}aab \to bbbaa\mathbf{ab} \to bbba\mathbf{ab}a \to bbb\mathbf{ab}aa \to bbbbaaa$$

Ce dernier mot n'appartient pas à L, donc $\deg(baababb) = 8$. Remarque : à chaque étape, le facteur ab concerné a été indiqué en **gras**.

Question 7 • Notons J l'ensemble des indices j compris entre 1 et $|u|$ tels que $u_j = b$; puis, pour $j \in J$, notons d_j le nombre d'occurrences de a dans le préfixe $u[1..j-1]$; définissons alors $\delta(u) = \sum_{j \in J} d_j$. Remarquons que cette somme est nulle ssi $u \notin L$. Nous allons établir $\deg(u) = \delta(u)$. Ceci traduit l'idée intuitive suivante : appliquer répétitivement φ au mot u revient à faire «remonter» les b vers le début du mot et, corrélativement, à faire redescendre les a vers la fin du mot.

• La preuve est simple : si un mot n'appartient pas à L, il est de la forme $b^p a^q$; on a donc $\deg(u) = \delta(u) = 0$. Si $u \in L$, considérons la décomposition $u = vabw$ dont l'existence a été établie à la question 1. Alors $\varphi(u) = vbaw$; il est clair que $\delta(vbaw) = \delta(vabw) - 1$, et $\deg(vbaw) = \deg(vabw) - 1$. Ceci montre que la quantité $\deg(u) - \delta(u)$ est conservée à chaque application de φ. Comme sa valeur finale est nulle, il en est de même de sa valeur initiale, donc $\deg(u) = \delta(u)$ pour tout mot u.

• Soit u de longueur $n = 2p$. Si $u = a^p b^p$, alors $\deg(u) = p^2$ puisque chacune des p occurrences de b est précédée par les p occurrences de a. Si $u = a^{p-j} b^{p+j}$ avec $0 < j \leqslant p$, alors $\deg(u) = (p-j)(p+j) = p^2 - j^2 < p^2$; même conclusion si $u = a^{p+j} b^{p-j}$ avec $0 \leqslant j < p$. Enfin, supposons que u contient le facteur ba : $u = b^{k+1} aw$ avec

$k \geqslant 0$; le mot $v = b^k abw$ vérifie $\varphi(v) = u$, donc $\deg(v) = \deg(u) + 1$. Or v a même longueur que u. Concluons : la valeur maximale du degré d'un mot de longueur $2p$ est p^2, et cette valeur est atteinte par le mot $a^p b^p$ et lui seul.

• Un raisonnement très semblable montre que la valeur maximale du degré d'un mot de longueur $2p+1$ est $p(p+1)$, et cette valeur est atteinte par les mots $a^p b^{p+1}$ et $a^{p+1} b^p$ et eux seuls.

Question 8 • L'arrêt se produit lorsque l'on rencontre le facteur ab ou, au plus tard, en fin du mot u, si ce dernier ne contient pas le facteur ab.

```
let rec phi = function
 | [] | [_] -> failwith "pas dans L"
 | A::B::q -> B::A::q
 | t::q -> t::(phi q) ;;
```

Question 9 • Nous mettons en oeuvre les idées développées à la question 7 : après lecture du préfixe de longueur ℓ de u, `d` donne le nombre d'occurrences de a dans ce préfixe ; et la valeur de `accu` est $\sum_{j \in J, j \leqslant \ell} d_j$, soit le degré du préfixe $u[1..\ell]$.

```
let degré u =
 let rec aux d accu = function
  | A::q -> aux (d+1) accu q
  | B::q -> aux d (accu+d) q
  | _ -> accu
 in aux 0 0 u ;;
```

Question 10 • Prenons $M = \{a^n b a^n \mid n \geqslant 0\}$. Il est clair que M est une partie de L. Si M est un langage rationnel, alors, d'après le lemme de l'étoile, il existe un entier N tel que tout mot u de M, de longueur au moins égale à N, se décompose en $u = xyz$ avec $y \neq \varepsilon$, $|xy| \leqslant N$ et $xy^k z \in M$ pour tout $k \in \mathbb{N}$. Appliquons ceci au mot $a^N b a^N$, dont la longueur $2N+1$ est supérieure à N ; comme $|xy| \leqslant N$, xy est préfixe de a^N, donc $x = a^i$, $y = a^j$ et $z = a^{N-i-j} b a^N$, avec $i \geqslant 0$ et $j > 0$. Alors le mot $xz = a^{N-j} b a^N$ devrait appartenir à M, ce qui n'est pas le cas. Donc M n'est pas un langage rationnel.

• Par ailleurs, le langage $\psi(M) = \{ba^{2n} | n \geqslant 0\}$ est rationnel, puisqu'il est décrit par l'expression rationnelle $b(aa)^*$.

FIN