Théorie des Langages - Feuille nº 8

AUTOMATES À PILE ET LANGAGES

CORRECTION

Exercice 1 - Donnez un automate à pile avec acceptation par pile vide correspondant à chacune des grammaires suivantes mises sous forme normale de Greibach :

$G_1 = \langle V_1, \Sigma_1, P_1, S_1 \rangle$	$G_2 = \langle V_2, \Sigma_2, P_2, S_2 \rangle$
$V_1 = \{a, b, S_1, B, C\}$	$V_2 = \{a, b, c, S_2, D, E\}$
$\Sigma_1 = \{a, b\}$	$\Sigma_2 = \{a, b, c\}$
$S_1 \to a \mid aS_1 \mid aBC$	$S_2 \to cD$
$B \to aB \mid bC$	$D \to aS_2 \mid bE$
$C \to b$	$E \to aD \mid cS_2 \mid b$

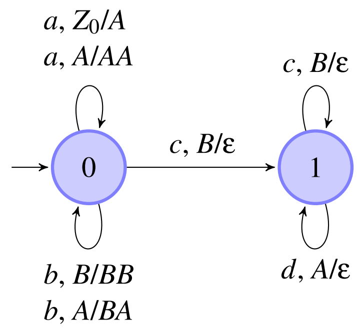
Pour la grammaire G_1 :
Automate $M_1 = (\Sigma, \Gamma, Z_0, Q, q_0, F, \delta)$, avec

- $\Sigma = \{a, b\}$
- $\Gamma = \{S_1, B, C\}$
- $Q = \{q_0\}$
- $F = \emptyset$
- δ :
  - $(q_0, a, S_1) \to (q_0, \varepsilon)$
  - $(q_0, a, S_1) \to (q_0, S_1)$
  - $(q_0, a, S_1) \to (q_0, BC)$
  - $(q_0, a, B) \to (q_0, B)$
  - $(q_0, b, B) \to (q_0, C)$
  - $(q_0, b, C) \to (q_0, \varepsilon)$

Pour la grammaire G_2 :
Automate $M_2 = (\Sigma, \Gamma, Z_0, Q, q_0, F, \delta)$, avec

- $\Sigma = \{a, b, c\}$
- $\Gamma = \{S_2, D, E\}$
- $Q = \{q_0\}$
- $F = \emptyset$
- δ :
  - $(q_0, c, S_2) \to (q_0, D)$
  - $(q_0, a, D) \to (q_0, S_2)$
  - $(q_0, b, D) \to (q_0, E)$
  - $(q_0, a, E) \to (q_0, D)$
  - $(q_0, c, E) \to (q_0, S)$
  - $(q_0, b, E) \to (q_0, \varepsilon)$

Exercice 2 - Soient les deux automates à pile (M_1 à haut, M_2 à bas), avec acceptation par pile vide, suivants. Donnez les grammaires algébriques correspondantes.



La correction de M_2 est à envoyer en devoir maison. Pour M_1 :

- Première étape de l'algorithme : on met l'axiome S et tous les triplets possibles $< etat, pile, etat >$.

  $V \setminus \Sigma = \{S, \langle 0,Z_0,0\rangle, \langle 0,Z_0,1\rangle, \langle 1,Z_0,0\rangle, \langle 1,Z_0,1\rangle, \langle 0,A,0\rangle, \langle 0,A,1\rangle, \langle 1,A,0\rangle, \langle 1,A,1\rangle, \langle 0,B,0\rangle, \langle 0,B,1\rangle, \langle 1,B,0\rangle, \langle 1,B,1\rangle\}$

  NB : il y a ici plus de triplets qu'il n'y a de transitions dans l'automate !

- Deuxième étape : $P \leftarrow \emptyset$. On construit P itérativement :

- Troisième étape : $S \rightarrow \langle 0,Z_0,0\rangle | \langle 0,Z_0,1\rangle$

- Quatrième étape : il faut traiter les transitions $(c, B/\varepsilon)$ et $(d, A/\varepsilon)$

  $$\langle 0,B,1\rangle \rightarrow c \qquad \langle 1,B,1\rangle \rightarrow c \qquad \langle 1,A,1\rangle \rightarrow d$$

- Cinquième étape : on va traiter les transitions restantes une à une.

  - Pour $(a, Z_0/A)$: $\langle 0,Z_0,0\rangle \rightarrow a\langle 0,A,0\rangle \qquad \langle 0,Z_0,1\rangle \rightarrow a\langle 0,A,1\rangle$
  - Pour $(a, A/AA)$:

    $$\begin{array}{lllllll} \langle 0,A,0\rangle & \rightarrow & a\langle 0,A,0\rangle\langle 0,A,0\rangle & \qquad & \langle 0,A,0\rangle & \rightarrow & a\langle 0,A,1\rangle\langle 1,A,0\rangle \\ \langle 0,A,1\rangle & \rightarrow & a\langle 0,A,0\rangle\langle 0,A,1\rangle & & \langle 0,A,1\rangle & \rightarrow & a\langle 0,A,1\rangle\langle 1,A,1\rangle \end{array}$$

  - Pour $(b, A/BA)$:

    $$\begin{array}{lllllll} \langle 0,A,0\rangle & \rightarrow & b\langle 0,B,0\rangle\langle 0,A,0\rangle & \qquad & \langle 0,A,0\rangle & \rightarrow & b\langle 0,B,1\rangle\langle 1,A,0\rangle \\ \langle 0,A,1\rangle & \rightarrow & b\langle 0,B,0\rangle\langle 0,A,1\rangle & & \langle 0,A,1\rangle & \rightarrow & b\langle 0,B,1\rangle\langle 1,A,1\rangle \end{array}$$

  - Pour $(b, B/BB)$:

    $$\begin{array}{lllllll} \langle 0,B,0\rangle & \rightarrow & b\langle 0,B,0\rangle\langle 0,B,0\rangle & \qquad & \langle 0,B,0\rangle & \rightarrow & b\langle 0,B,1\rangle\langle 1,B,0\rangle \\ \langle 0,B,1\rangle & \rightarrow & b\langle 0,B,0\rangle\langle 0,B,1\rangle & & \langle 0,B,1\rangle & \rightarrow & b\langle 0,B,1\rangle\langle 1,B,1\rangle \end{array}$$

- On renomme, et on obtient :

$$\begin{array}{lllllllll} S & \to & K|L & H & \to & c|bGH|bHJ & J & \to & c \\ F & \to & d & K & \to & aC & L & \to & aD \\ C & \to & aCC|aDE|bGC|bHE & D & \to & aCD|aDF|bGD|bHF & G & \to & bGG|bHI \end{array}$$

- Finalement, on réduit et on nettoie :
  - Symboles productifs : $\{H,J,F,D,L,S\}$.
    $\Rightarrow$ On supprime les règles qui contiennent C,E,G,I et K.

$$\begin{array}{lllllllll} S & \to & L & H & \to & c|bHJ & J & \to & c \\ F & \to & d & & & & L & \to & aD \\ & & & D & \to & aDF|bHF & & & \end{array}$$

  - Symboles accessibles : $\{S,L,D,H,F,J\}$
    $\Rightarrow$ Pas de suppression de règles de production.
  - Traitement des règles unitaires (ici $S \to L$)

$$\begin{array}{lllllllll} S & \to & aD & H & \to & c|bHJ & J & \to & c \\ F & \to & d & & & & L & \to & aD \\ & & & D & \to & aDF|bHF & & & \end{array}$$

  - Symboles accessibles : $\{S,D,F,H,J\}$

On obient donc $G = \langle V,\Sigma,P,S\rangle$, avec

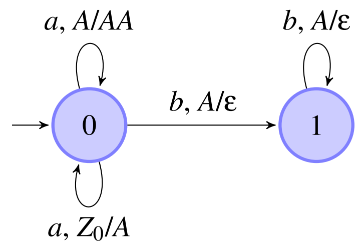
- $\Sigma = \{a,b,c,d\}$; $V \setminus \Sigma = \{S,D,F,H,J\}$
- P :

$$\begin{array}{lll} S & \to & aD \\ H & \to & c|bHJ \\ F & \to & d \\ J & \to & c \\ D & \to & aDF|bHF \end{array}$$

On peut même aller un peu plus loin car on remarque que F (resp. J) peut être remplacer directement par d (resp. c). Ce qui nous donne la grammaire équivalente $G' = \langle V,\Sigma,P,S\rangle$, avec

- $\Sigma = \{a,b,c,d\}$; $V \setminus \Sigma = \{S,D,H\}$
- P :

$$\begin{array}{lll} S & \to & aD \\ H & \to & c|bHc \\ D & \to & aDd|bHd \end{array}$$



Pour M_2.

- $V \setminus \Sigma = \{S, \langle 0,Z_0,0\rangle, \langle 0,Z_0,1\rangle, \langle 1,Z_0,0\rangle, \langle 1,Z_0,1\rangle, \langle 0,A,0\rangle, \langle 0,A,1\rangle, \langle 1,A,0\rangle, \langle 1,A,1\rangle\}$

- $P \leftarrow \emptyset$. On construit P itérativement :

- $S \to \langle 0,Z_0,0\rangle | \langle 0,Z_0,1\rangle$

- Pour les transitions $(b, A/\varepsilon)$

$$\begin{array}{lcl} \langle 0,A,1\rangle & \to & b \\ \langle 1,A,1\rangle & \to & b \end{array}$$

- Pour $(a, Z_0/A)$

$$\begin{array}{lcl} \langle 0,Z_0,0\rangle & \to & a\langle 0,A,0\rangle \\ \langle 0,Z_0,1\rangle & \to & a\langle 0,A,1\rangle \end{array}$$

- Pour $(a, A/AA)$

$$\begin{array}{lcl} \langle 0,A,0\rangle & \to & a\langle 0,A,0\rangle\langle 0,A,0\rangle \\ \langle 0,A,0\rangle & \to & a\langle 0,A,1\rangle\langle 1,A,0\rangle \\ \langle 0,A,1\rangle & \to & a\langle 0,A,0\rangle\langle 0,A,1\rangle \\ \langle 0,A,1\rangle & \to & a\langle 0,A,1\rangle\langle 1,A,1\rangle \end{array}$$

- On renomme, et on obtient :

$$\begin{array}{lcl@{\qquad}lcl} S & \to & F|G & B & \to & aBB|aCD \\ C & \to & aBC|aCE|b & E & \to & b \\ F & \to & aB & G & \to & aC \end{array}$$

- On réduit et on nettoie :
  - Symboles productifs : $\{C,E,G,S\}$.
  - Symboles accessibles : $\{S,G,C,E\}$
  - Règle unitaire, on obtient :

$$\begin{array}{lcl@{\qquad}lcl} S & \to & aC & & & \\ C & \to & aCE|b & E & \to & b \\ G & \to & aC & & & \end{array}$$

  - Symboles accessibles, on a : $G = \langle V, \Sigma, P, S\rangle$, avec
    - $\Sigma = \{a,b\}$; $V \setminus \Sigma = \{S,C\}$
    - P :

$$\begin{array}{lcl} S & \to & aC \\ C & \to & aCb|b \end{array}$$

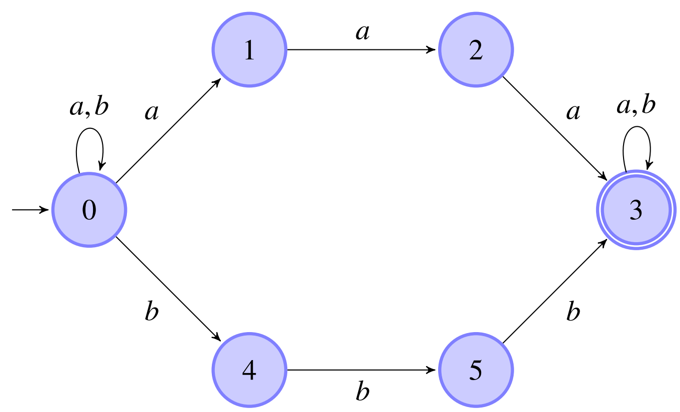
Exercice 3 - Pour chacun des langages suivantes, déterminer s'il s'agit i) d'un langage régulier, ii) d'un langage algébrique mais non-régulier ou iii) d'un langage non-algébrique. Justifier votre réponse.

1. $L_1 = \{uvw \mid u,w \in \Sigma^* \text{ et } v \in \{aaa, bbb\}\}$ avec $\Sigma = \{a,b,c\}$;

L_1 est un langage régulier. Pour le prouver, il suffit de donner au choix (a) une expression régulière, (b) un automate fini (déterministe ou non) ou (c) une grammaire régulière qui capture ce langage.

(a) $(a+b)^*(aaa+bbb)(a+b)^*$

(b)



(c) On a $G = \langle V, \Sigma, P, 0 \rangle$, avec

- $\Sigma = \{a,b\}$; $V \setminus \Sigma = \{0,1,2,3,4,5\}$
- P :

$$\begin{array}{lcl}
0 & \rightarrow & a0|b0|a1|b4 \\
1 & \rightarrow & a2 \\
2 & \rightarrow & a3 \\
3 & \rightarrow & a3|b3|\varepsilon \\
4 & \rightarrow & b4 \\
5 & \rightarrow & b3
\end{array}$$

2. $L_2 = \{0^{2^n} | n \geq 0\}$ avec $\Sigma = \{0\}$;

L_2 n'est pas algébrique. Utilisons le théorème de pompage pour le prouver.
Supposons que L_2 est algébrique.
Soit $w = 0^{2^p}$. Comme $|w| = 2^p$, on a bien $|w| \geq p$.
Si L_2 est algébrique, quelle que soit la décomposition de w en $xuyzt$ qui vérifie :

(a) $|uz| > 0$

(b) $|uyz| \leq p$

Si on trouve un $n \geq 0$ tel que $xu^n yz^n t \notin L_2$ alors L_2 n'est pas algébrique.
Soit $xuyzt$ une décomposition de w vérifiant les deux conditions ci-dessus.
Supposons $xu^2yz^2t \in L_2$ (on va montrer que $n = 2$ permet de montrer que L_2 n'est pas algébrique en raisonnant sur la longueur du mot),

$$|xu^2yz^2t| = |xuyzt| + |uz| \leq |xuyzt| + |uyz| \leq |xuyzt| + p = 2^p + p$$

Cependant, on sait que mathématiquement $2^p + p < 2^{p+1}$ donc on a $|xu^2yz^2t| < 2^{p+1}$.

Rappellons que $\forall w \in L_2$, $\exists q$ tel que $w = 0^{2^q}$ et $|w| = 2^q$.
On a donc :

- $|xu^2yz^2t| > 2^p$, puisque $|xuyzt| = 2^p$ et $|uz| > 0$
- $|xu^2yz^2t| < 2^{p+1}$

Donc il n'existe pas de q tel que $xu^2yz^2t = 0^{2^q}$, et donc $xu^2yz^2t \notin L_2$.
L_2 n'est donc pas algébrique.

3. $L_3 = \{a^n b^m \mid n < m\}$

L_3 est un langage algébrique mais non-régulier. Pour prouver que L_3 n'est pas régulier, nous utiliserons le lemme de l'étoile et, pour prouver que L_3 est algébrique, il suffit de trouver un automate à pile ou une grammaire algébrique qui capture L_3.

Prouvons que L_3 n'est pas un langage régulier en utilisant le lemme de l'étoile.

Supposons que L_3 est un langage régulier.

Il existe donc un entier positif p tel que $\forall w \in L_3$ avec $|w| \geq p$, il existe une décomposition $w = xuy$ tel que $|xu| \leq p$, $|u| > 0$ (ou $u \neq \varepsilon$) et $\forall n \geq 0, xu^n y \in L_3$.

Considérons le mot $w = a^p b^{p+1} \in L_3$. On a bien $|w| = 2p+1 \geq p$.

Tout préfixe de w de longueur inférieur à p est nécessairement composé uniquement de a. Comme xu est un préfixe de w et que $|xu| \leq p$, c'est vrai également pour xu. Autrement dit, x et u sont uniquement composés de a (mais pas forcément le même nombre). Donc il existe deux entiers i, j (avec $j > 0$) tels que $\underbrace{a^i}_{x}\underbrace{a^j}_{u}\underbrace{a^k b^{p+1}}_{y}$ avec $i + j + k = p$.

Montrons maintenant que tous les mots de la forme $a^i (a^j)^n a^k b^{p+1}$ avec $n \geq 0$ n'appartiennent pas forcément à L_3. Commençons par réecrire la formule :

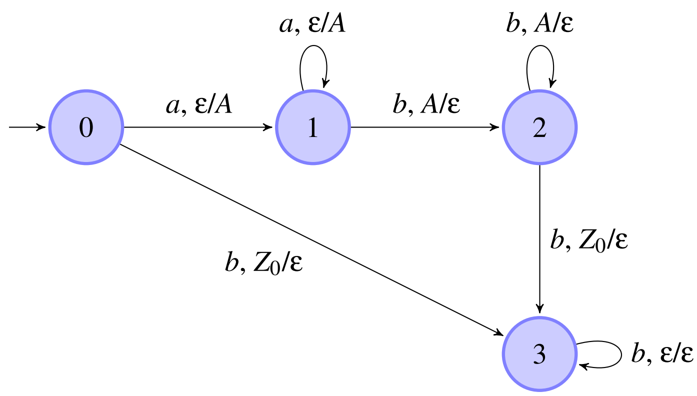
$$a^i (a_j)^n a^k b^{p+1} = a^i a^j (a^j)^{n-1} a^k b^{p+1} = a^i a^j a^k (a^j)^{n-1} b^{p+1} = a^p (a^j)^{n-1} b^{p+1}$$

On voit maintenant clairement que comme $j > 0$, pour n'importe quel $n > 1$ on a $a^p (a^j)^{n-1} b^{p+1} \notin L_3$ car on obtient au moins autant de a que de b. L'hypothèse est donc fausse. Par conséquent, L_3 n'est pas un langage régulier.

Prouvons que L_3 est un langage algébrique.

L'automate à pile M_3 utilisant l'acceptation par pile vide suivant capture L_3.

On a ainsi $M_3 = (\{a,b\}, \{Z_0, A\}, Z_0, \{0,1,2,3\}, 0, \emptyset, \delta)$ avec δ :



Vous pouvez vérifier en regardant que les mots b, bb, abb, $aaabbbbbb$ sont bien acceptés alors que les mots ε, ab, $aabb$, aab ne le sont pas (impossible d'obtenir une pile vide).

4. $L_4 = \{a^n b^m \mid n + m \leq 512\}$

L_4 est un langage régulier car il est fini. Sa taille est précisément égale à $\sum_{i=1}^{513} i$.