# Automates et langages

Corrigé de l'examen — RICM1— 8 janvier 2003

## Exercice 1 : Un automate et son langage

1. Voici les productions de grammaire obtenues directement à partir de l'automate:
   $P \to aP \quad P \to aQ$
   $Q \to bP \quad Q \to R$
   $R \to bR \quad R \to cQ \quad R \to bP \quad R \to \epsilon$
   Les non-terminaux de la grammaire sont $\{P,Q,R\}$, le symbole initial est $P$.
2. En dénotant avec $X_p$, $X_q$, $X_r$ les langages acceptés à partir des états $p$, $q$ et $r$ respectivement, le système d'équations pour ces langages est:

$$\left\{\begin{array}{lcl} X_p & = & aX_p + aX_q \\ X_q & = & bX_p + X_r \\ X_r & = & bX_r + cX_q + bX_p + \epsilon \end{array}\right.$$

   En remplaçant $X_q$ dans les équations 1 et 3 on obtient:

$$\left\{\begin{array}{lcl} X_p & = & aX_p + abX_p + aX_r \\ X_r & = & bX_r + cbX_p + cX_r + bX_p + \epsilon \end{array}\right.$$

   La deuxième équation implique:

$$X_r = (b+c)^*(cbX_p + bX_p + \epsilon)$$

   d'où, en remplaçant dans la première équation:

$$X_p = (a+ab)X_p + a(b+c)^*(cb+b)X_p + a(b+c)^*$$

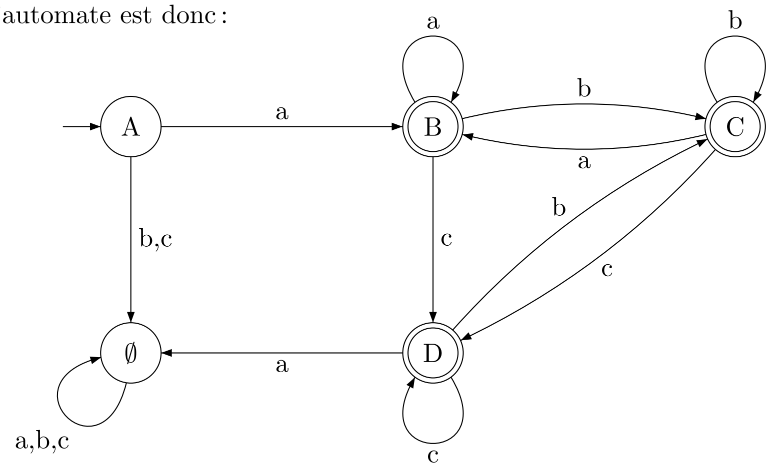
   Finalement, on obtient l'expression régulière équivalente à l'automate :

$$X_p = (a + ab + a(b+c)^*(cb+b))^* \, a(b+c)^*$$

3. Le tableau de successeurs obtenu par déterminisation est :

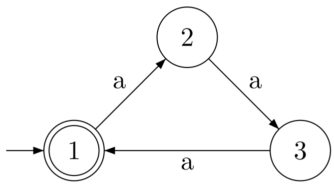
| | $A$ $\{p\}$ | $B$ $\{p,q,r\}$ | $C$ $\{p,r\}$ | $D$ $\{q,r\}$ |
|---|---|---|---|---|
| $a$ | $\{p,q,r\}$ | $\{p,q,r\}$ | $\{p,q,r\}$ | $\emptyset$ |
| $b$ | $\emptyset$ | $\{p,r\}$ | $\{p,r\}$ | $\{p,r\}$ |
| $c$ | $\emptyset$ | $\{q,r\}$ | $\{q,r\}$ | $\{q,r\}$ |

L'automate est donc :

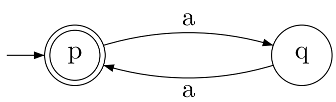



## Exercice 2: Un langage régulier plus intéressant

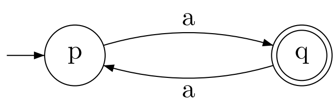
1. Le langage contient les mots de longueur multiple de 3 et impaire, formés avec le symbole $a$.
2. Soit $K = (aaa)^*$ et $L = (aa)^*$. L'automate acceptant $K$ est :



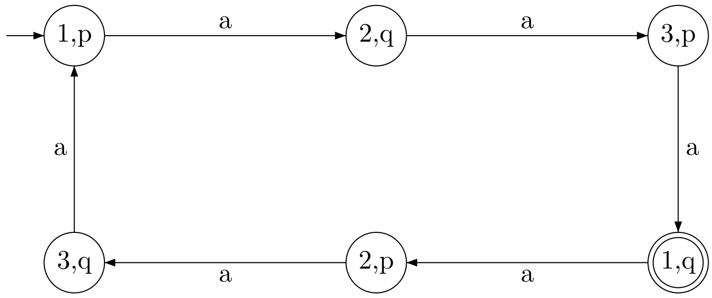

L'automate acceptant $L$ est :




L'automate acceptant $\overline{L}$ est :



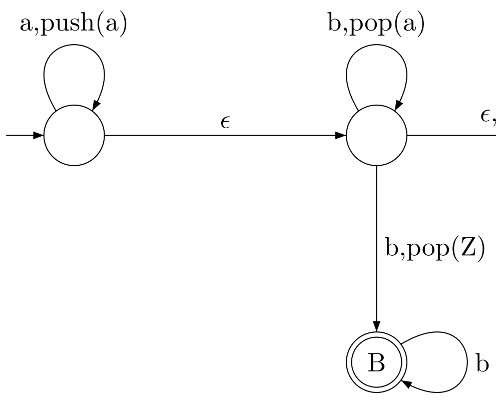
L'automate acceptant $M = K - L = K \cap \overline{L}$ est:




3. $M$ n'est pas vide, car l'automate contient un chemin $1p \to 2q \to 3p \to 1q$ de l'état initial à l'état final.
   $M$ contient une infinité de mots, car l'automate contient un cycle $\sigma = 1p \to 2q \to 3p \to 1q \to 2p \to 3q \to 1p$ accessible de l'état initial, tel que l'état final est accessible à partir de $\sigma$, étiqueté par un mot non vide $a^6$.
4. $a^3(a^6)^*$
5. $3 + 6k \mid k \in \mathbb{N}$

# Exercice 3: Un langage non-régulier

1. La grammaire qui engendre $\{a^i b^j \mid i < j\}$ peut être donnée avec les productions suivantes :
   $$I \to aIb \quad | \quad Ib \quad | \quad b$$
   La grammaire qui engendre $\{a^i b^j \mid i > j\}$ peut être donnée avec les productions suivantes :
   $$J \to aJb \quad | \quad aJ \quad | \quad a$$
   La grammaire qui engendre T est donnée par les productions antérieures, plus les productions suivantes :
   $$S \to I \quad | \quad J$$
   Les non-terminaux de la grammaire sont $\{S,I,J\}$. Le symbole initial est $S$.
2. Voici un automate à pile pour $T$, en supposant que l'automate commence avec une pile contenant le symbole $Z$ (l'état A sert a accepter les mots qui ont plus de $a$, l'état B ceux avec plus de $b$.)



3. $a^*b^* = \{a^n b^m \mid n,m \in \mathbb{N}\}$
   $a^*b^* - T = \{a^n b^m \mid n,m \in \mathbb{N} \wedge \neg(n \neq m)\} = \{a^i b^i \mid i \in \mathbb{N}\}$
4. Supposons que $T$ est régulier. Alors, par les propriétés de fermeture des langages réguliers, $\overline{T}$ est aussi régulier. Comme $a^*b^*$ est régulier, il en résulte que $a^*b^* \cap \overline{T} = a^*b^* - T$ est aussi régulier.
   Mais $a^*b^* - T = \{a^i b^i \mid i \in \mathbb{N}\}$ n'est pas régulier (vu en cours).
   Par conséquent, on a une contradiction de l'hypothèse de régularité de $T$.

# Exercice 4: Une intersection

1. On a vu ces langages en TD.

$$\begin{array}{rcllcl} P & = & \{a^n b^m a^m b^n \mid n,m \in \mathbb{N}\} & & & \\ Q & = & ((a+b)^3)^* & & = & \{w \mid w \in \{a,b\}^* \text{ et } |w| \text{ est multiple de } 3\} \end{array}$$

2. Nous allons utiliser comme non-terminaux pour l'intersection $S_0,S_1,S_2,R_0,R_1,R_2$. La signification du symbole $S_i$ (resp. $R_i$) est la même que celle de $S$ (resp. $R$), et encore que le nombre de terminaux (c-à-d $a$ et $b$) déjà générés est $i(\bmod 3)$.
   Chaque transition de la grammaire de départ donne plusieurs transitions de la nouvelle grammaire, ainsi pour $S \to aSb$ on obtient trois règles: $S_0 \to aS_2b; S_1 \to aS_0b; S_2 \to aS_1b$ (si avant la production on avait $i(\bmod 3)$ terminaux, alors après on a 2 terminaux de plus, c-à-d $i+2(\bmod 3)$).
   On peut terminer la dérivation si est seulement si notre symbole est R, et le nombre de terminaux est multiple de 3, ce qui correspond à $R_0$ dans la nouvelle grammaire. Pour cette raison la seule production qui donne $\epsilon$ est $R_0 \to \epsilon$.
   Finalement on obtient la grammaire suivante:

$$\begin{array}{rcl|l} S_0 & \to & a\ S_2 b & R_0 \\ S_1 & \to & a\ S_0 b & R_1 \\ S_2 & \to & a\ S_1 b & R_2 \\ R_0 & \to & b\ R_2 a & \epsilon \\ R_1 & \to & b\ R_0 a & \\ R_2 & \to & b\ R_1 a & \end{array}$$

   avec $S_0$ comme symbole initial.