Université de Paris – UFR d'Informatique
L2 Informatique et Math-Info

Automates et analyse lexicale
Année 2020-2021

TD n°2

Expressions rationnelles et automates

Les exercices marqués (∗) sont optionnels

Expressions rationnelles

Exercice 1 *Soit* $\Sigma = \{a, b, c\}$. *Donner des expressions rationnelles décrivant les langages ci-dessous :*

$L_1 = \{u \in \Sigma^* :$ *toute occurrence de b est immédiatement suivie de deux occurrences de a*$\}$,
$L_2 = \{u \in \Sigma^* :$ *le nombre d'occurrences de a dans u est multiple de 3*$\}$,
(∗) $L_3 = \{u \in \Sigma^* :$ *le nombre d'occurrences de a dans u est congru à* 3 *modulo 4*$\}$,
(∗) $L_4 = \{u \in \Sigma^* :$ *les blocs de a dans u sont alternativement de longueur paire et impaire et le premier bloc de a, s'il existe, est de longueur paire, le dernier aussi*$\}$.

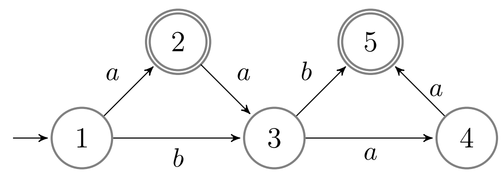
Exercice 2 *Dans la suite, a et b désignent des lettres.*

1. *Simplifier les expressions rationnelles suivantes :*
   (a) $(aa)^*a + (aa)^*$;
   (b) $(a + \varepsilon)a^*b$;
   (c) $(a + \varepsilon)(\varepsilon + aa)^+a$.
2. (∗) *Montrer les égalités suivantes :*
   (a) $(a^2 + a^3)^* = (a^2a^*)^*$;
   (b) $a^*(a + b)^* = (a + ba^*)^*$;
   (c) $(ba)^+(a^*b^* + a^*) = (ba)^*ba^+b^*$.

Initiation informelle aux automates

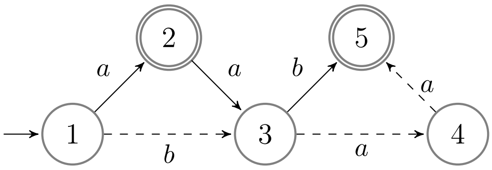
Une autre façon de décrire un langage est de dessiner un automate. Nous introduisons cette notion informellement, elle sera reprise en cours ensuite. Voici à quoi ressemble un automate.

$\mathcal{A}$:



Vocabulaire : Les cercles seront appelés *états*, les flèches *transitions* et les lettres associées aux flèches sont appelées *étiquettes*. Les états avec une petite flèche entrante sans étiquette sont les *états initiaux*, ici l'état 1 est initial. Ceux avec un double cercle sont les *états terminaux* (ou *finaux*), ici 2 et 5.

Pour savoir si un mot est *accepté* par un automate, on lit ce mot en commençant par un état initial et en prenant à chaque fois une transition qui permet de lire la prochaine lettre, pour voir si on aboutit à un état final. Par exemple, *baa* se lit sur le chemin donné par les transitions en pointillé ci-dessous.

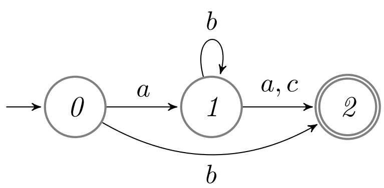


On pourra noter le chemin $1 \xrightarrow{b} 3 \xrightarrow{a} 4 \xrightarrow{a} 5$, et 5 est bien un état final donc le mot est reconnu.
Si un tel chemin n'existe pas, le mot n'est pas accepté. Par exemple le mot *aa* peut se lire, mais on aboutit sur l'état 3 qui n'est pas final. Le mot *bbb* quant à lui ne peut même pas se lire sur l'automate.
Le *langage reconnu* par un automate donné est l'ensemble des mots acceptés par cet automate.

Exercice 3 *1. Quels sont les mots acceptés par l'automate $\mathcal{A}$ ci-dessus ? Déduisez-en le langage reconnu.*

2. Faites de même avec l'automate $\mathcal{B}$ ci-dessous (donnez une expression rationnelle correspondante).

$\mathcal{B}$:



Vocabulaire (suite) : les automates $\mathcal{A}$ et $\mathcal{B}$ sont dit *déterministes* car il y a un seul état initial et, depuis un état donné, il y a au plus une transition possible pour une même lettre. Dans le cas contraire, ils sont dit *non déterministes* (voir plus bas).

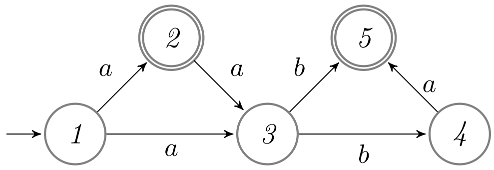
Exercice 4 (Dessinez des automates déterministes) *1. Dessinez un automate déterministe qui reconnaît le langage $\{car, bar, or\}$. Faites en sorte qu'il ait le moins d'états possibles.*

2. Dessinez un automate déterministe qui reconnaît le langage donné par l'expression rationnelle $a^*(a+b)c^*$.

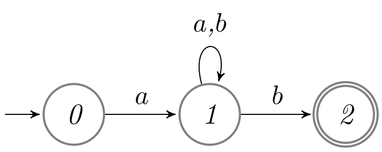
3. $(*)$ Dessinez un automate déterministe qui reconnaît le langage des mots de longueur paire sur l'alphabet $\{a\}$.

Exercice 5 ($(*)$ Automates non déterministes) *Quand on a des automates non déterministes, un mot peut être lu de plusieurs manières différentes dans le même automate. Il est accepté dès qu'un des chemins (au moins) finit par un état terminal, même si d'autres chemins aboutissent à d'autres états.*
On considère les deux automates suivants :

$\mathcal{C}_1$:



$\mathcal{C}_2$:



1. Les mots aab, abbb, abab et aba sont-ils acceptés par l'automate $\mathcal{C}_1$? Sont-ils acceptés par l'automate $\mathcal{C}_2$?

2. Décrire par une expression rationnelle les langages reconnus par ces automates.