Université de Rouen — Année en cours
Faculté des sciences — Théorie des langages
Licence Informatique 3ème année — Résumé de cours pour le TD n° 2

Les automates et les langages reconnaissables

1 Définition

Un *automate fini* sur un alphabet Σ est un quintuplet $M = (\Sigma, Q, \delta, I, T)$ où :
— Q est un ensemble fini d' *états*,
— δ est un ensemble fini de *transitions* ou *flèches* $\delta \subseteq Q \times \Sigma \times Q$
— I est l'ensemble des *états initiaux* $I \subseteq Q$
— T est l'ensemble des *états terminaux* ou *états finaux* $T \subseteq Q$

Exemple 1 :

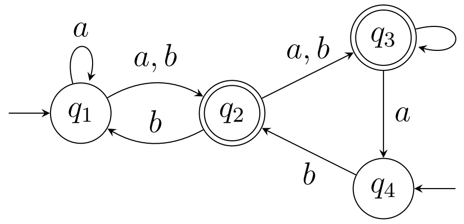

Figure 1 – Automate M_1

$\Sigma = \{a, b\}$, $Q = \{q_1, q_2, q_3, q_4\}$, $I = \{q_1, q_4\}$, $T = \{q_2, q_3\}$,
$\delta = \{(q_1, a, q_1),(q_1, a, q_2),(q_1, b, q_2),(q_2, a, q_3),(q_2, b, q_1),(q_2, b, q_3),(q_3, a, q_3),(q_3, a, q_4),(q_4, b, q_2)\}$.
Dans la transition $(1, a, 1)$, a est appelée la *trace* ou l'*étiquette* de la transition.

On peut représenter δ à l'aide d'une table appelée *table de transitions* :

	a	b
$\rightarrow q_1$	$\{q_1, q_2\}$	$\{q_2\}$
$\leftarrow q_2$	$\{q_3\}$	$\{q_1, q_3\}$
$\leftarrow q_3$	$\{q_3, q_4\}$	$\emptyset$
$\rightarrow q_4$	$\emptyset$	$\{q_2\}$

Exemple 2 :

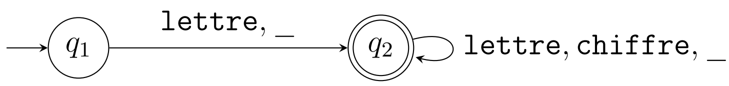

FIGURE 2 – Automate M_2

$\Sigma = \{\texttt{lettre}, \texttt{chiffre}, \texttt{\_}\}$, $Q = \{q_1, q_2\}$, $I = \{q_1\}$, $T = \{q_2\}$
$\delta = \{(q_1, \texttt{lettre}, q_2), (q_2, \texttt{lettre}, q_2), (q_2, \texttt{chiffre}, q_2)\}$

On appelle *chemin* dans l'automate une suite de flèches consécutives :
$q_2 \xrightarrow{a} q_3 \xrightarrow{a} q_3 \xrightarrow{a} q_4 \xrightarrow{b} q_2$ est un chemin de M_1, sa *trace* est la suite des traces de ses flèches soit $aaab$ et sa *longueur*, la longueur de sa trace, soit 4 ici.

$q_1 \xrightarrow{\texttt{lettre}} q_2 \xrightarrow{\texttt{lettre}} q_2 \xrightarrow{\texttt{chiffre}} q_2$ est un chemin de M_2 de trace `lettre lettre chiffre` et de longueur 3.
$p \xrightarrow{\varepsilon} p$ est un chemin de longueur 0 et de trace ε.

Un *chemin réussi* est un chemin qui , partant d'un état initial, se termine dans un état final. Le chemin $q_1 \xrightarrow{\texttt{lettre}} q_2 \xrightarrow{\texttt{lettre}} q_2 \xrightarrow{\texttt{chiffre}} q_2$ de M_2 est un chemin réussi.

On appelle *langage reconnu par un automate M* et on note $L(M)$ l'ensemble des traces des chemins réussis.

L'automate M_2 reconnait les mots qui sont formés d'une suite non vide de lettres, de chiffres et de _ commençant par une lettre ou _. Autrement dit, cet automate donne la règle de construction des identificateurs de variables pour le langage `C` et reconnaît l'ensemble des identificateurs dans ce langage de programmation.

Exemple 3 :

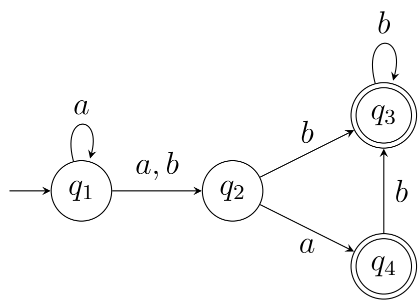

FIGURE 3 – Automate M_3

$$
\begin{aligned}
L(M_3) &= \{a\}^* \cdot \Sigma \cdot (\{b\}^+ \cup \{a\} \cup \{a\} \cdot \{b\}^+) \\
&= \{a\}^* \cdot \Sigma \cdot (\{b\}^+ \cup \{a\} \cdot (\{\varepsilon\} \cup \{b\}^+)) \\
&= \{a\}^* \cdot \Sigma \cdot (\{b\} \cdot \{b\}^* \cup \{a\} \cdot \{b\}^*) \\
&= \{a\}^* \cdot \Sigma \cdot \{a, b\} \cdot \{b\}^* \\
&= \{a\}^* \cdot \Sigma^2 \cdot \{b\}^*
\end{aligned}
$$

Exemple 4 :

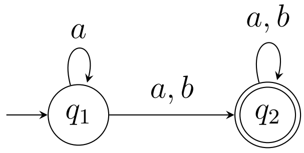

FIGURE 4 – Automate M_4

$L(M_4) = \{a\}^* \cdot \Sigma \cdot \Sigma^* = \{a\}^* \cdot \Sigma^+$

Deux automates sont dits *équivalents* s'ils reconnaissent le même langage. On dit qu'un langage est *reconnaissable* si il existe un automate qui le reconnait et on note $\mathrm{Rec}(\Sigma^*)$ l'ensemble des langages reconnaissables sur Σ.