TD 1 - Langages réguliers

définitions

Longueur (nombre de symboles) d'un mot :

$$|\varepsilon| = 0 \qquad\qquad |aw| = 1 + |w|$$

Nombre d'occurences d'un symbole :

$$|\varepsilon|_a = 0 \qquad\qquad |bw|_a = \begin{cases} 1 + |w|_a & si\ a = b \\ |w|_a & si\ a \neq b. \end{cases}$$

Image miroir (inverse l'ordre des lettres) :

$$\overline{\varepsilon} = \varepsilon \qquad\qquad \overline{aw} = \overline{w}a$$

Concaténation et étoile de langages :

$$L_1 \cdot L_2 := \{u_1 \cdot u_2 \mid u_1 \in L_1,\ u_2 \in L_2\}$$

$$L^0 := \{\varepsilon\} \qquad\qquad L^{n+1} := L \cdot L^n$$

$$L^\star := \bigcup_{n\in\mathbb{N}} L^n = \{u_1 \cdots u_n \mid n \in \mathbb{N},\ \forall 1 \leqslant i \leqslant n,\ u_i \in L\}\,.$$

Exercice 1. Langages réguliers

Montrer que les langages suivants sont réguliers.

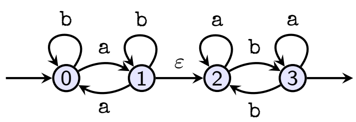
1. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ commence par $\mathtt{ab}$ et finit par $\mathtt{bb}\}$;
2. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ a (au moins) une occurence de $\mathtt{bb}\}$;
3. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ a exactement une occurence de $\mathtt{bb}\}$;
4. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ ne contient pas 3 occurences successives de la même lettre$\}$;
5. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ contient un nombre impair de $\mathtt{a}\}$;
6. $\{w \in \{\mathtt{a},\mathtt{b}\}^\star \mid w$ ne contient pas la lettre $\mathtt{b}\}$.

Exercice 2. Élimination des ε-transitions

Un automate avec ε-transitions (ou automate asynchrone) est une structure

$$\mathscr{A} := \langle Q, \Sigma, \Delta, I, F\rangle$$

où Q est un ensemble fini d'états, Σ est un alphabet fini, $I, F \subseteq Q$ sont les ensembles d'états respectivement initiaux et finaux, et $\Delta \subseteq Q \times (\Sigma \cup \{\varepsilon\}) \times Q$ est l'ensemble des transitions. Voici un exemple :



Une exécution dans un tel automate est une séquence

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} q_n$$

telle que $\forall 1 \leqslant i \leqslant n$, $q_{i-1} \xrightarrow{a_i} q_i \in \Delta$, et $\langle q_0, q_n\rangle \in I \times F$. Le mot reconnu par l'exécution est obtenu à partir de la séquence $a_1 \cdots a_n$ en effaçant les εs, et en ne gardant que les symboles de l'alphabet Σ. Le langage de l'automate est l'ensemble des mots reconnus par des exécutions de l'automate.

Le langage de l'automate donné en exemple ci-dessus est l'ensemble des mots $w \in \{\mathtt{a}, \mathtt{b}\}$ qui peuvent se décomposer en $w = uv$, où u a un nombre impair de $\mathtt{a}$, et v a un nombre impair de $\mathtt{b}$. En effet, toute exécution de cette machine se décompose comme suit : $0 \xrightarrow{u}^{*} 1 \xrightarrow{\varepsilon} 2 \xrightarrow{v}^{*} 3$. L'exécution $0 \xrightarrow{u}^{*} 1$ nous assure que le mot u contient un nombre impair de $\mathtt{a}$ (voir la question 5 de l'exercice 1), et de manière similaire la fin de l'exécution vérifie que v contient un nombre impair de $\mathtt{b}$.

Question. Montrer que les langages reconnus par les automates asynchrones sont exactement les langages réguliers.

Exercice 3. Propriétés de clôture des langages réguliers

Montrer que les langages réguliers sont clos par :

1. union
2. intersection
3. concaténation
4. étoile (répétition)
5. image miroir.

Exercice 4. Lemme de l'étoile

Montrer que les langages suivants ne sont pas réguliers :

1. $L_1 := \{\mathtt{a}^n\mathtt{b}^n \mid n \in \mathbb{N}\}$.
2. $L_2 := \left\{w \in \{\mathtt{a}, \mathtt{b}\}^\star \;\middle|\; |w|_{\mathtt{a}} = |w|_{\mathtt{b}}\right\}$.
3. $L_3 := \left\{w \in \{\mathtt{a}, \mathtt{b}\}^\star \;\middle|\; w = \overline{w}\right\}$.
4. $L_4 := \left\{ww \;\middle|\; w \in \{\mathtt{a}, \mathtt{b}\}^\star\right\}$.