Automates finis

On introduit les automates fini qui permettant de représenter d'une manière fini certaines ensemble (les ensembles rationnels) du monoïde libre Σ^*. Un automate est formé par l'ensemble des différents états possibles du système, reliés entre eux par des conditions: le systéme passe d'un état dans un autre quand une condition donnée est vérifiée.

Définition 1: Un automate est un 5-uplet $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ où,

- Σ est un ensemble fini dit l'alphabet d'entrée.
- Q est un ensemble dit l'ensemble des états.
- $I \subseteq Q$ est l'ensemble des états initiaux.
- $F \subseteq Q$ est l'ensemble des états finals.

$\delta \subseteq Q \times \Sigma \times Q$ est l'ensemble des transitions.

Un automate fini est un automate $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ dont l'ensemble d'états Q est fini.

Un automate fini $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ est automate fini déterministe et on le note AFD si $|I| = 1$.

Nous représentons un automate fini $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ de la manière suivante:

les états de $\mathcal{A}$ sont les sommets d'un graphe orienté et sont représentés par des cercles.

Si $\delta(q, \alpha) = q', q, q' \in Q, \alpha \in \Sigma$, alors on trace un arc orienté de q vers q' et d'étiquette $\alpha, q \underset{\longrightarrow}{\alpha} q'$. Les états finals sont représentés à un double cercle et l'état initial est désigné par une flèche entrante sans étiquette. Enfin si deux lettres α, β sont telles que $\delta(q, \alpha) = q'$ et $\delta(q, \beta) = q'$, on s'autorise à dessiner un unique arc portant deux étiquettes séparés par une virgule $q \underset{\longrightarrow}{\alpha, \beta} q'$. Cette convention s'adapte à plus de deux lettres.

Exemple 2: Soit l'automate $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ où $\Sigma = \{\alpha, \beta\}, Q = \{1, 2\}, I = \{1\}, F = \{2\},$

$\delta = \{(1, \alpha, 1), (1, \beta, 2), (2, \beta, 2)\}$.

Remarque 3: Soit l'automate $\mathcal{A} = (Q, I, F, \Sigma, \delta)$. On étend naturellement la fonction de transition δ à $Q \times \Sigma^*$ de la manière suivante: $\delta(q, \epsilon) = q$ et $\delta(q, \alpha w) = \delta(\delta(q, \alpha), w), \alpha \in \Sigma, w \in \Sigma^*, q \in Q$.

Définition 4: Soit $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ un automate fini. Un mot c de δ^* est un chemin dans $\mathcal{A}$ ssi il s'écrit,

$c = (q_1, x_1, q_1')(q_2, x_2, q_2') \ldots (q_i, x_i, q_i') \ldots (q_n, x_n, q_n')$ et il vérifie la condition $c = \epsilon$ ou $\forall i \in [1, n-1], q_i' = q_{i+1}$.

- n est la longueur du chemin.
- Ce chemin mêne de l'état q_1 à l'état q_n.
- On convient que ϵ mêne de q à q, pour tout $q \in Q$.
- On appelle trace le morphisme $h : \delta^* \longrightarrow \Sigma^*$ tel que: $\forall (q, x, q') \in \delta, h((q, x, q')) = x$.
- Un mot $w \in \Sigma^*$ est accepté (ou reconnu) par $\mathcal{A}$ ssi il existe un chemin c dans $\mathcal{A}$ menant d'un état $q_i \in I$ à un état $q_f \in F$ tel que $h(c) = w$.
- On appelle langage accepté (ou reconnu) par $\mathcal{A}$, noté $L_{\mathcal{A}}$, l'ensemble des mots acceptés par $\mathcal{A}$ .i. e,

$$L_{\mathcal{A}} = \{w \in \Sigma^*, \exists q_i \in I, q_f \in F : \delta^*(q_i, w) = q_f\}.$$

Définition 5: Soit l'automate $\mathcal{A} = (Q, I, F, \Sigma, \delta)$. La fonction de transitions δ définit un morphisme Ψ de Σ^* dans le monoïde Q^Q de toutes les applications de Q dans Q: pour tout $w \in \Sigma^*, \Psi(w)$ est l'application qui à tout $q \in Q$ associe l'élément $q' \in Q$ ssi $\delta^*(q, w) = q'$. Le sous monoïde $\Psi(\Sigma^*)$ de Q^Q est appelé le monoïde de transitions de $\mathcal{A}$.

Exercice 6: Construit l'automate fini déterministe qui reconnaît le langage L suivant:

$$L = \{w \in \{0, 1\}^* : |w|_1 \equiv 0\,[4]\}.$$

Proposition 6: Soit $\mathcal{A} = (Q, q_o, F, \Sigma, \delta)$ un automate fini déterministe.

1. La relation $\mathcal{R}_{\mathcal{A}}$ définie sur Σ^* par $w_1 \mathcal{R}_{\mathcal{A}} w_2 \iff \delta^*(q_0, w_1) = \delta^*(q_0, w_2)$, est une relation d'équivalence.

2. Pour tout langage L de Σ^*, la relation $\mathcal{R}_L$ définie sur Σ^* par $w_1 \mathcal{R}_L w_2 \iff (\forall x \in \Sigma^* : w_1 x \in L \iff w_2 x \in L)$ est une relation d'équivalence.

3. Les relations $\mathcal{R}_{\mathcal{A}}$ et $\mathcal{R}_L$ sont invariantes à droite.

4. La relation $\mathcal{R}_{\mathcal{A}}$ est d'indice fini.

Théorème 7:

Soit L un langage sur un alphabet Σ. Les assertions suivantes sont équivalentes:

1. L est accepté par un automate fini déterministe.

2. L est la réunion de classes d'équivalence d'une relation invariante à droite d'indice fini.

3. La relation $\mathcal{R}_L$ est d'indice fini.

Exercice 8: On considére l'automate $\mathcal{A} = (Q, I, F, \Sigma, \delta)$ où $\Sigma = \{\alpha, \beta\}, Q = \{q_0, q_1, q_2\}, I = \{q_0\}, F = \{q_2\},$

$$\delta = \{(q_0, \alpha, q_1), (q_1, \alpha, q_1), (q_1, \beta, q_2), (q_2, \beta, q_2), (q_2, \alpha, q_1)\}.$$

Déterminer le langage $L_{\mathcal{A}}$.

Définition 9: Soit Σ un alphabet. Supposons que $0, \epsilon, +, ., (,), *$ sont des lettres n'appartenant pas à Σ. L'ensemble ER_Σ des expressions régulières sur Σ est défini récursivement par:

- 0 appartient à ER_Σ,
- pour tout $\sigma \in \Sigma$, σ appartient à ER_Σ,
- si φ et ψ appartient à ER_Σ, alors $\varphi + \psi, \varphi \cdot \psi$ et φ^* sont aussi appartient à ER_Σ.

A une expression régulière, on associe un langage grâce à l'application $\mathcal{L} : ER_\Sigma \longrightarrow 2^{\Sigma^*}$ définie par:

$\mathcal{L}(0) = \emptyset, \mathcal{L}(\epsilon) = \{\epsilon\}, \mathcal{L}(\sigma) = \{\sigma\}$ pour toute σ de Σ, $\mathcal{L}(\varphi + \psi) = \mathcal{L}(\varphi) \cup \mathcal{L}(\psi), \mathcal{L}(\varphi \cdot \psi) = \mathcal{L}(\varphi) \cdot \mathcal{L}(\psi), \mathcal{L}(\varphi^*) = (\mathcal{L}(\varphi))^*$.

Exemple 10: Soit $\Sigma = \{\alpha, \beta\}$, voici quelques exemples d'expressions régulières:

$e_1 = (\varepsilon + (\alpha\beta)), e_2 = (((\alpha\beta)\alpha) + \beta^*)^*, e_3 = ((\alpha + \beta)^* (\alpha\beta))$ et par conséquent,

$\mathcal{L}(e_1) = \{\varepsilon, \alpha\beta\}, \mathcal{L}(e_2) = \{\{\alpha\beta\alpha\} \cup \{\beta\}^*\}^*, \mathcal{L}(e_3) = \{\alpha, \beta\}^* \{\alpha\beta\}$.

Définition 11: 1. Un langage L sur Σ est régulier (ou rationnel) s'il existe une expression $e \in ER_\Sigma$ telle que $L = \mathcal{L}(e)$. Si φ et ψ deux expressions régulières telles que $\mathcal{L}(\varphi) = \mathcal{L}(\psi)$, alors on dit que φ et ψ sont équivalentes.

2. Les langages rationnels sur un alphabet Σ est la plus petite famille de langage R telle que :

$\emptyset \in R$ et $\{\sigma\} \in R$ pour toute lettre σ de Σ.

R est close pour les opérations rationnelles (l'union, la concaténation et l'étoile).

Théorème 12 (Kleene): Un langage sur un alphabet Σ est régulier si et seulement si il reconnu par un automate fini.