Automates à réels

Définitions & notations.

Soit A un ensemble fini. On note $A^\star$ l'ensemble des mots finis (ou suites finies) sur A et ε le mot vide. Soient $u, v \in A^*$: on note uv le mot obtenu en concaténant u et v. Étant donné un mot u et un entier n on note u^n le mot obtenu en concaténant n copies de u (avec la convention que $u^0 = \varepsilon$).

Un automate à réels est un automate comme un autre, mais disposant d'un compteur, au départ initialisé à 0. Chaque transition de l'automate peut décider de changer la valeur du compteur, en lui ajoutant une valeur λ. L'automate ne connaît à aucun moment la valeur du compteur mais peut seulement la modifier.

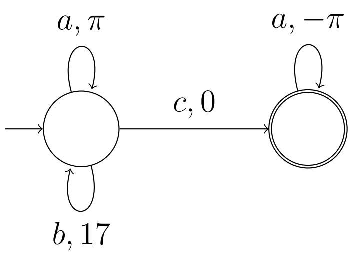
Formellement, un automate à réels est un quintuplet $\mathcal{A} = (Q, \Sigma, q_0, Q_F, T)$ où

— Q est un ensemble fini, dit ensemble d'états ; $q_0 \in Q$ est appelé état initial ; $Q_F \subseteq Q$ est l'ensemble d'états finaux ;

— Σ est l'alphabet fini ;

— $T \subseteq Q \times (\Sigma \cup \{\varepsilon\}) \times Q \times \mathbb{R}$ est un ensemble fini appelé ensemble des transitions (notez que l'on s'autorise des transitions silencieuses étiquetées par ε).

Un premier exemple d'automate à réels est représenté ci-dessous. L'état initial q_0 est repéré par une flèche entrante et les états finaux sont doublement entourés.



Une *configuration* est un élément de $Q \times \mathbb{R}$. La configuration initiale de l'automate $\mathcal{A}$ est $(q_0, 0)$. Les configurations finales sont celles de la forme $(q, 0)$ avec $q \in Q_F$.

On dit que $(q, x) \xrightarrow{a} (q', x')$ s'il existe dans T une transition de la forme (q, a, q', λ) avec $x' = x + \lambda$.

On dit que $(q, x) \xrightarrow{w} (q', x')$ si l'on peut écrire $w = w_1 \dots w_k$ avec $w_i \in \Sigma \cup \{\varepsilon\}$ et s'il existe $q_1 \dots q_{k+1}$ et $x_1 \dots x_{k+1}$ tels que $q_1 = q, x_1 = x, q_{k+1} = q', x_{k+1} = x'$ et pour tout i, $(q_i, x_i) \xrightarrow{w_i} (q_{i+1}, x_{i+1})$.

Enfin, le langage reconnu par $\mathcal{A}$ est défini par

$$L(\mathcal{A}) = \{w \mid \exists q \in Q_F, (q_0, 0) \xrightarrow{w} (q, 0)\}$$

On dit que L est $\mathbb{R}$-reconnaissable s'il existe un automate à réels tel que $L(\mathcal{A}) = L$.

Question 1.

(a) Vérifiez sur l'exemple que $aacaa \in L(\mathcal{A})$ mais que $aacab \notin L(\mathcal{A})$ et $abcaa \notin L(\mathcal{A})$.

(b) Décrire $L(\mathcal{A})$.

Si w est un mot on note $|w|_a$ le nombre d'occurrences de a dans w.

Question 2. Construire pour chacun des langages suivants un automate qui le reconnaît :

$$L_1 = \{w \in \{a,b\}^\star : |w|_a = |w|_b\}$$

$$L_2 = \{w \in \{a,b,c\}^\star : |w|_a = |w|_b = |w|_c\}$$

$$L_3 = \{a,b\}^\star \setminus L_1 = \{w \in \{a,b\}^\star : |w|_a \neq |w|_b\}$$

Question 3. Montrer que les langages $\mathbb{R}$-reconnaissables sont stables par union.

Un automate $\mathcal{A}$ est dit déterministe s'il ne contient aucune transition étiquetée par ε et que pour toute lettre a et tout état q, il existe au plus un q' et un λ tel que $(q, a, q', \lambda) \in T$.

Question 4. Montrer que le langage L_4 ci-dessous est $\mathbb{R}$-reconnaissable, mais pas par un automate déterministe. On pourra raisonner par l'absurde et raisonner sur le mot $a^n b^n c^n$ où n est plus grand que le nombre d'états d'un automate déterministe reconnaissant L_4.

$$L_4 = \{w \in \{a,b,c\}^\star : |w|_a = |w|_b \text{ ou } |w|_b = |w|_c\}$$

Question 5. Montrer que le complémentaire de L_1 est $\mathbb{R}$-reconnaissable mais qu'il n'est pas reconnaissable par un automate déterministe.

Si $\mathcal{A}$ est un automate, on appelle $\dim \mathcal{A}$ la dimension du $\mathbb{Q}$-espace vectoriel engendré par tous les coefficients λ apparaissant dans T. Dans l'exemple, on a donc $\dim \mathcal{A} = 2$.

Question 6. Montrer que tout automate *déterministe* $\mathcal{A}$ reconnaissant L_2 et qui a *un seul* état vérifie $\dim \mathcal{A} \geq 2$.

Question 7. Montrer que les langages $\mathbb{R}$-reconnaissables sont stables par intersection.

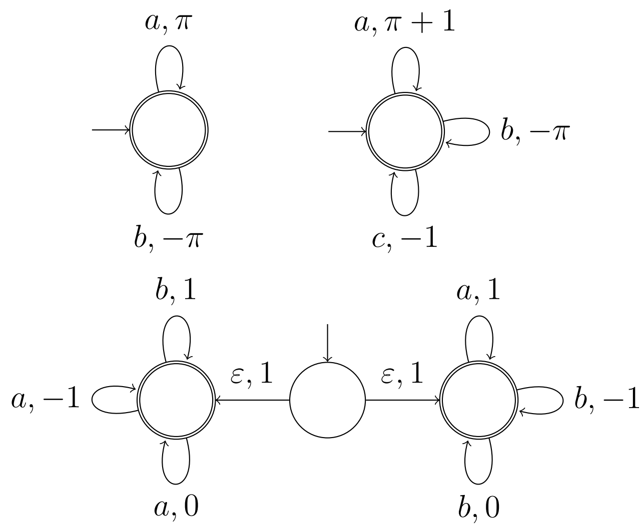
Question 8. Montrer que si L et son complémentaire sont reconnus par un automate déterministe, alors L est rationnel.

Correction

Question 1. $L(\mathcal{A}) = \{a^n c a^n | n \geq 0\}$.

Question 2. Le deuxième automate utilise le fait que π et 1 sont linéairement indépendants. Dans le troisième automate, l'état du milieu est initial.

Indication. *Pour L_3, commencer par faire le langage des mots qui ont plus de a que de b*



Question 3. Utiliser la construction classique avec ε-transitions

Question 4. Il est $\mathbb{R}$ reconnaissable par la question précédente. Prenons un automate déterministe qui le reconnaît.

Indication. *Partir du mot $a^n b^n c^n$ où n est plus grand que le nombre d'états de l'automate*

Ce mot est accepté. Comme n est grand, on boucle dans les b, il existe donc i et j différents (avec $i < j < n$) tels que $a^n b^i$ et $a^n b^j$ conduisent au même état. On ajoute pour passer de $a^n b^i$ à $a^n b^j$ une certaine quantité λ_b. Comme les mot $a^n b^i c^n$ et $a^n b^j c^n$ sont tous les deux acceptés on a nécessairement $\lambda_b = 0$ (c'est là qu'on utilise vraiment le déterminisme). De même, lors de l'exécution, on boucle aussi dans les c, donc il existe i' et j' différents (avec $i' < j' < n$) tels que la lecture de $a^n b^n c^{i'}$ et $a^n b^n c^{j'}$ nous conduisent dans le même état. Pour la même raison on a $\lambda_c = 0$. Mais alors le mot $a^n b^{n-(j-i)} c^{n+(j'-i')}$ est aussi accepté ce qui conduit à une contradiction.

Question 5. Le complémentaire de L_1 est L_3.

On raisonne par l'absurde et on considère le mot $a^n b^{2n}$ où $n = (k+5)!$ (par exemple) où k dénote le nombre d'états d'un automate déterministe reconnaissant le complémentaire de L_1. Et on utilise le même raisonnement qu'à la question précédente on montre l'existence de deux entiers $p, q < k$ tel que $a^{n+pi} b^{2n+qj} \in L_3$ pour tous couples (i, j).

Mais en prenant i et j tels que $pi = n + qj$, on a une contradiction. (prendre $j = p$ et voir que le terme de droite est bien divisible par p, donc qu'on peut définir i).

Question 6. Comme il n'y a qu'un seul état, la preuve est simple. Quand on lit $a^n b^k c^l$, le réel obtenu à la fin est $n\lambda_a + k\lambda_b + c\lambda_c$ et il doit vérifier :

— Pour tout n , $n\lambda_a + n\lambda_b + n\lambda_c = 0$

— Pour tout triplet i, j, k non simultanément égaux, $i\lambda_a + j\lambda_b + k\lambda_c \neq 0$

En particulier, $\lambda_a + \lambda_b + \lambda_c = 0$ et aucun des trois λ_i n'est nul. Sans perte de généralités disons que $\lambda_a > 0$ et $\lambda_b < 0$. Mais comme $\dim \mathcal{A} < 2$, λ_a et λ_b sont liés, par exemple par $p\lambda_a = -q\lambda_b$ avec p, q positifs distincts. Mais alors $p\lambda_a + q\lambda_b + 0\lambda_c = 0$ donc $a^p b^q$ accepté, une contradiction.

Question 7. Faire le produit cartésien des deux automates. La difficulté est qu'il faut commencer par supposer qu'ils travaillent sur des ensembles de coefficients linéairement indépendants, donc

— Se rendre compte qu'en remplaçant les coefficients λ dans un automate par d'autres coefficients β mais qui vérifient les mêmes relations sur $\mathbb{Q}$, alors on ne change pas le fonctionnement de l'automate.

— Utiliser ensuite le fait que $\mathbb{R}$ est un $\mathbb{Q}$ espace vectoriel de dimension infinie.

Question 8. Soient $\mathcal{A}_1$ et $\mathcal{A}_2$ deux automates déterministes qui reconnaissent L et son complémentaire, sur des ensembles d'états respectifs Q_1 et Q_2.

On construit un automate sur $Q_1 \times Q_2 \times \{1, 2\}$. L'idée est que l'on est dans l'état (q_1, q_2, i) après avoir lu un mot u si et seulement si l'automate $\mathcal{A}_1$ est dans l'état q_1 après avoir lu u, idem pour $\mathcal{A}_2$, et que le mot u est accepté par l'automate $\mathcal{A}_i$

L'état initial est $(q_0^{\mathcal{A}_1}, q_0^{\mathcal{A}_2}, i)$ où $i = 1$ si $q_0^{\mathcal{A}_1}$ est final dans $\mathcal{A}_1$, et $i = 2$ sinon.

Ensuite, il y a une transition de (q_1, q_2, i) à (q_1', q_2', j) si :

— Il y a une transition de q_p à q_p' dans l'automate $\mathcal{A}_p$:

— Si $i = 1$: Si q_1' est acceptant et que la transition de q_1 à q_1' porte la valeur $\lambda = 0$, alors $j = 1$. Sinon $j = 2$

— Même raisonnement si $i = 2$.

En prenant comme états acceptants les états où $i = 1$ on obtient un automate pour L.