MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

Partiel du 14 novembre 2006. Durée: 1h 30, tous documents autorisés

⋆ ⋆ ⋆

Avertissement : On attachera une grande importance à la clarté, à la précision et à la concision de la rédaction.

Partie 1: Monoïde syntactique

Question 1. Soit $A = \{a, b\}$. Calculer l'automate minimal du langage b^*abA^*.

Question 2. Calculer le monoïde syntactique M de L.

Question 3. Quels sont les idempotents de M?

Question 4. Calculer l'idéal minimal de M.

Question 5. Calculer les relations de Green dans M.

Question 6. Le langage L est-il sans-étoile? Testable par morceaux?

Partie 2. Automates extensifs

Soit $\mathcal{A} = (Q, A, \cdot, i, F)$ un automate déterministe fini complet. On dit que $\mathcal{A}$ est extensif s'il existe un ordre total $\leqslant$ sur Q tel que, pour tout $q \in Q$ et $a \in A$, $q \leqslant q \cdot a$.

Question 7. Montrer que le monoïde de transition d'un automate extensif est nécessairement $\mathcal{R}$-trivial.

Question 8. Montrer que, réciproquement, tout langage reconnu par un monoïde $\mathcal{R}$-trivial est reconnu par un automate extensif.

Partie 3. Langages reconnus par des monoïdes $\mathcal{R}$-triviaux.

Un langage est dit $\mathcal{R}$-simple s'il est de la forme $A_0^*a_1A_1^*a_2A_2^* \cdots a_kA_k^*$ avec $k \geqslant 0$, $a_1, \ldots, a_k \in A$, $A_0, \ldots, A_k \subseteq A$ et $a_1 \notin A_0$, $a_2 \notin A_1$, $\ldots$, $a_k \notin A_{k-1}$.

Question 9. Montrer que le monoïde syntactique d'un langage $\mathcal{R}$-simple est $\mathcal{R}$-trivial.

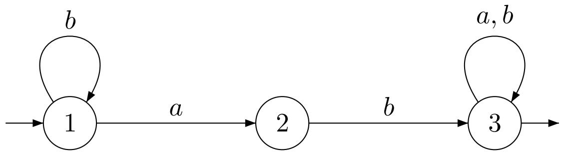
Question 10. Proposer un algorithme efficace pour tester si le monoïde de transition d'un automate déterministe fini complet donné est $\mathcal{R}$-trivial. On précisera la complexité de l'algorithme proposé en fonction du nombre d'états de l'automate et de la taille de l'alphabet.

Question 11. Démontrer que si le monoïde syntactique d'un langage est $\mathcal{R}$-trivial, ce langage est union disjointe de langages $\mathcal{R}$-simples.

Corrigé

Partie 1: Monoïde syntactique

Question 1. L'automate minimal du langage b^*abA^* est donné ci-dessous



Question 2. Le monoïde syntactique M de L est donné par le tableau suivant

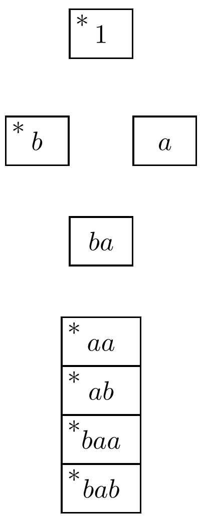
		1	2	3
$*$	1	1	2	3
	a	2	0	3
$*$	b	1	3	3
$*$	aa	0	0	3
$*$	ab	3	0	3
	ba	2	3	3
$*$	baa	0	3	3
$*$	bab	3	3	3

dans lequel les idempotents sont indiqués par une étoile. Les relations définissant M sont $bb = b$, $aaa = aa$, $aab = aa$ et $aba = ab$.

Question 3. L'ensemble des idempotents de M est $\{1, b, aa, ab, baa, bab\}$.

Question 4. L'idéal minimal de M est $\{aa, ab, baa, bab\}$.

Question 5. Les relations de Green dans M sont représentées ci-dessous:



Question 6. Comme M est apériodique, le langage L est sans-étoile. En revanche, M n'est pas $\mathcal{J}$-trivial (puisque par exemple $aa \;\mathcal{L}\; ab$) et donc L n'est pas testable par morceaux.

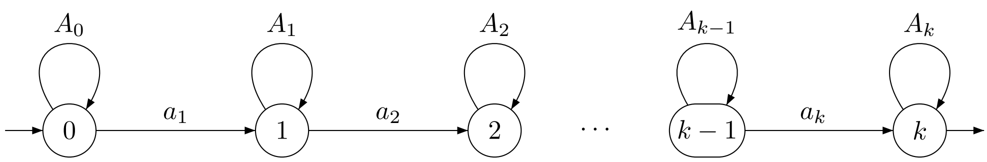
Partie 2. Automates extensifs

Question 7. Soit $\mathcal{A} = (Q, A, \cdot\,, i, F)$ un automate extensif. et soit M son monoïde de transition. Si $u \;\mathcal{R}\; v$, il existe des éléments x, y de M tels que $ux = v$ et $vy = x$. On a alors, pour tout état $q \in Q$, $q\cdot u \leqslant (q\cdot u)\cdot x = q\cdot v$ et de même $q\cdot v \leqslant (q\cdot v)\cdot y = q\cdot u$. On en déduit que pour tout $q \in Q$, $q\cdot u = q\cdot v$, et donc $u = v$. Par conséquent, M est $\mathcal{R}$-trivial.

Question 8. Soit L un langage de A^* reconnu par un monoïde $\mathcal{R}$-trivial M. Il existe donc un morphisme $\varphi : A^* \to M$ et une partie F de M tels que $L = \varphi^{-1}(F)$. On utilise la construction classique pour passer d'un monoïde à un automate déterministe $\mathcal{A} = (M, A, \cdot\,, 1, F)$, où les transitions sont données par $q\cdot a = q\varphi(a)$. On sait que $\mathcal{A}$ reconnait L. De plus, puisque M est $\mathcal{R}$-trivial, la relation $\leqslant_{\mathcal{R}}$ est un ordre partiel sur M. Prolongeons l'inverse de cet ordre partiel en un ordre total sur M, noté $\leqslant$. On suppose donc par définition que si $u \leqslant_{\mathcal{R}} v$, alors $v \leqslant u$. Puisque pour tout $q \in M$, $q\varphi(a) \leqslant_{\mathcal{R}} q$, on a $q \leqslant q\varphi(a)$ et donc $q \leqslant q\cdot a$, ce qui montre que l'automate $\mathcal{A}$ est extensif.

Partie 3. Langages reconnus par des monoïdes $\mathcal{R}$-triviaux.

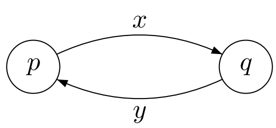
Question 9. Considérons le langage $\mathcal{R}$-simple $L = A_0^* a_1 A_1^* a_2 A_2^* \cdots a_k A_k^*$, avec $k \geqslant 0$, $a_1, \ldots, a_k \in A$, $A_0, \ldots, A_k \subseteq A$ et $a_1 \notin A_0$, $a_2 \notin A_1$, $\ldots$, $a_k \notin A_{k-1}$. L'automate déterministe représenté ci-dessous reconnait L.



Cet automate est extensif par construction est d'après la question 8, son monoïde de transition M est $\mathcal{R}$-trivial. Comme le monoïde syntactique de L divise M, il est également $\mathcal{R}$-trivial.

Question 10. Soit $\mathcal{A} = (Q, A, \cdot\,, i, F)$ un automate déterministe complet. On va montrer que le monoïde de transition M de $\mathcal{A}$ est $\mathcal{R}$-trivial si et seulement si les composantes fortement connexes de $\mathcal{A}$ (considéré comme un graphe dirigé) sont triviales. On peut alors calculer les composantes fortement connexes du graphe par l'algorithme de Tarjan et vérifier qu'elles sont triviales. La complexité de cet algorithme est en temps linéaire par rapport à la taille du graphe (nombre de sommets + nombre d'arêtes), donc ici en $O(|A||Q|)$.

Supposons que l'une des composantes fortement connexes du graphe de $\mathcal{A}$ contienne deux états distincts p et q: il existe alors deux mots x et y tels que $p\cdot x = q$ et $q\cdot x = p$.



Soit n l'exposant de M. Les éléments $(xy)^n$ et $(xy)^n x$ sont alors $\mathcal{R}$-équivalents puisque $(xy)^n x = [(xy)^n]x$ et $[(xy)^n x]y(xy)^{n-1} = (xy)^{2n} = (xy)^n$. Il sont pourtant distincts puisque $p\cdot (xy)^n = p$ et $p\cdot (xy)^n x = q$. Donc M n'est pas $\mathcal{R}$-trivial.

Réciproquement, supposons qu'il existe deux éléments distincts u et v de M $\mathcal{R}$-équivalents. Il existe des éléments x, y de M tels que $ux = v$ et $vy = x$ et il existe au moins un état r tel que $r\cdot u \neq r\cdot v$. Posant alors $p = r\cdot u$ et $q = r\cdot v$, on a $p\cdot x = r\cdot ux = r\cdot v = q$ et de même $q\cdot y = p$. Donc les composantes fortement connexes du graphe de $\mathcal{A}$ ne sont pas triviales.

Question 11. D'après la question 8, un langage L dont le monoïde syntactique est $\mathcal{R}$-trivial est reconnu par un automate extensif $\mathcal{A} = (Q, A, \cdot, i, F)$. Comme Q est totalement ordonné, on peut supposer que $Q = \{0, 1, \ldots, n\}$. Si on pose $A_i = \{a \in A \mid i\cdot a = i\}$, le langage L est égal à l'union disjointe des langages $\mathcal{R}$-simples $A_{i_0}^* a_1 A_{i_1}^* a_2 \cdots a_k A_{i_k}^*$, où l'union est étendue à l'ensemble (fini) des $(i_0, a_1, i_1, a_2, \ldots, a_k, i_k)$ tels que $0 = i_0 < i_1 < i_2 < \ldots < i_k \in F$ et $i_0\cdot a_1 = a_1$, $i_1\cdot a_2 = i_2$, $\ldots$, $i_{k-1}\cdot a_k = i_k$.