MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

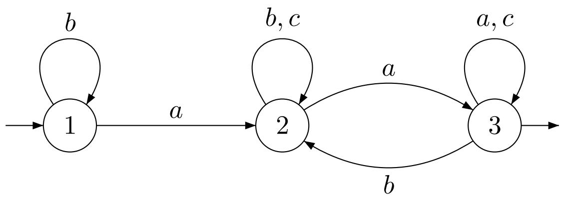
Partiel du 1er décembre 2010. Durée: 2h, notes de cours autorisées

⋆⋆⋆

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

Partie 1. Calcul d'un monoïde syntactique.

Soit $A = \{a, b, c\}$. On considère l'automate $\mathcal{A}$ représenté ci-dessous:



Question 1. Donner une expression rationnelle pour le langage L reconnu par $\mathcal{A}$.

Question 2. Calculer le monoïde syntactique M de L. On donnera la liste de ses éléments (vous devriez trouver 8 éléments, en comptant l'élément neutre) et les relations permettant de le définir.

Question 3. Calculer l'image P de L dans M par le morphisme syntactique.

Question 4. Déterminer l'idéal minimal et les idempotents de M.

Question 5. Déterminer la structure en $\mathcal{D}$-classes de M (on dessinera les diagrammes boîtes à œufs). Le monoïde M est-il commutatif? $\mathcal{R}$-trivial? $\mathcal{L}$-trivial? apériodique?

Question 6. Montrez que L est sans-étoile et donnez une expression sans étoile pour L.

Question 7. Donner une formule φ de la logique $\mathbf{FO}[<]$ telle que $L = L(\varphi)$. Même question avec une formule utilisant seulement deux variables. Est-ce possible avec seulement une variable? Justifiez votre réponse.

Partie 2. Constantes

Soit M un monoïde et P une partie de M. On dit qu'un élément s de M est une *constante pour* P si pour tout $s_1, s_2, s_3, s_4 \in M$,

$$s_1 s s_2, s_3 s s_4 \in P \text{ entraîne } s_1 s s_4 \in P$$

On note $C(P)$ l'ensemble des constantes pour P. Dans la suite du problème, on fixe un langage reconnaissable L de A^*, on note $\eta : A^* \to M$ son morphisme syntactique et on pose $P = \eta(L)$.

Question 8. Montrer que si $s \in P$ et si s est une constante pour P, alors la condition $usv \in P$ entraîne $us, sv \in P$.

Question 9. Montrer qu'un mot u de A^* est une constante pour L si et seulement si $\eta(u)$ est une constante pour P.

Question 10. Soit $\mathcal{A} = (Q, A, \cdot, q_0, F)$ l'automate minimal (déterministe, émondé mais pas nécessairement complet) de L. Le *rang* d'un mot u est l'entier

$$\mathrm{rg}(u) = |\{q \cdot u \mid q \in Q \text{ et } q \cdot u \text{ est défini}\}|.$$

Montrer qu'un mot u est une constante pour L si et seulement si il est de rang $\leqslant 1$.

Question 11. Déterminer les constantes pour P dans le cas où L est le langage considéré dans la première partie du problème.

On dit qu'un langage L de A^* est *dense* si on peut prolonger n'importe quel mot de A^* en un mot de L, autrement dit, si pour tout mot u de A^*, on a $A^*uA^* \cap L \neq \emptyset$.

Question 12. Montrez que $C(P)$ est un idéal de M. Montrer que cet idéal est apériodique.

Question 13. Montrer que si L n'est pas dense, alors M a un zéro. Montrer que si s est une constante pour P, alors $sMs \subseteq \{s, 0\}$.

Question 14. Montrer que si L est dense et si s est une constante pour P, alors $sMs = \{s\}$.

MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

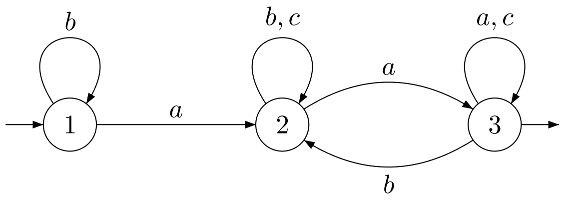
December 1st, 2010. Duration: 2h.

⋆ ⋆ ⋆

Warning : Clearness, accuracy and concision of the writing will be rewarded.

Part 1. Computation of a syntactic monoid.

Let $A = \{a, b, c\}$. Let $\mathcal{A}$ be the automaton represented in the picture below:



Question 1. Give a rational (= regular) expression for the language L accepted by $\mathcal{A}$.

Question 2. Compute the syntactic monoid M of L. Give the list of its elements (you should find 8 elements, including the identity) and of its defining relations.

Question 3. Compute the image of L in M under the syntactic morphism.

Question 4. Compute the minimal ideal and the idempotents of M.

Question 5. Compute the $\mathcal{D}$-class structure of M (draw the egg-box pictures). Is the monoid M comutative? $\mathcal{R}$-trivial? $\mathcal{L}$-trivial? aperiodic?

Question 6. Show that L star-free and give a star-free expression for L.

Question 7. Give a formula φ of the logic $\mathbf{FO}[<]$ such that $L = L(\varphi)$. Same question with a formula with only two variables. Is it possible with only one variable? Justify your answer.

Part 2. Constants

Let M be a monoid and let P be a subset of M. An element s of M is said to be a *constant for P* if for all $s_1, s_2, s_3, s_4 \in M$,

$$s_1ss_2, s_3ss_4 \in P \text{ implies } s_1ss_4 \in P$$

We denote by $C(P)$ the set of all constants for P. In the remainder of this problem, we fix a recognisable language L of A^*, we denote by $\eta : A^* \to M$ its syntactic morphism and we set $P = \eta(L)$.

Question 8. Show that if $s \in P$ and if s is a constant for P, then the condition $usv \in P$ implies $us, sv \in P$.

Question 9. Show that a word u of A^* is a constant for L if and only if $\eta(u)$ is a constant for P.

Question 10. Let $\mathcal{A} = (Q, A, \cdot, q_0, F)$ be the minimal automaton (deterministic, trim but not necessarily complete) of L. The *rank* of a word u is the nonnegative integer

$$\mathrm{rg}(u) = |\{q \cdot u \mid q \in Q \text{ et } q \cdot u \text{ is defined}\}|.$$

Show that a word u is a constant for L if and only if its rank is $\leqslant 1$.

Question 11. Give the constants for P when L is the language of the first part of the problem.

A language L of A^* is said to be *dense* if one can extend any word of A^* into a word of L, that is, if for all word u of A^*, one has $A^*uA^* \cap L \neq \emptyset$.

Question 12. Show that $C(P)$ is an idéal of M. Show that this ideal is aperiodic.

Question 13. Show that if L is nondense, then M has a zero. Show that if s is a constant for P, then $sMs \subseteq \{s, 0\}$.

Question 14. Show that if L is dense and if s is a constant for P, then $sMs = \{s\}$.

Corrigé

Partie 1. Calcul d'un monoïde syntactique.

Question 1. Une expression rationnelle pour L is $b^*a(b+c)^*a(a+c+b(b+c)^*a)^*$. Une expression beaucoup plus simple est $b^*aA^*ac^*$.

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

	1	2	3
$*\ 1$	1	2	3
a	2	3	3
$*\ b$	1	2	2
$*\ c$	0	2	3
$*\ a^2$	3	3	3
$*\ ab$	2	2	2
$*\ bc$	0	2	2
$*\ ca$	0	3	3

dans lequel les idempotents sont indiqués par une étoile. Les relations définissant M sont

$$ac = a \qquad ba = a \qquad bb = b \qquad cb = bc \qquad cc = c \qquad a^3 = a^2 \qquad a^2b = ab$$
$$abc = ab \qquad bca = ca \qquad caa = ca \qquad cab = bc$$

Question 3. L'image de L est obtenu en sélectionnant les éléments x de M tels que $1 \cdot x = 3$. Ici $P = \{a^2\}$.

Question 4. L'idéal minimal est $\{a^2, ab, bc, ca\}$ et les idempotents sont

$$E(M) = \{1, b, c, a^2, ab, bc, ca\}$$

Question 5. La structure en $\mathcal{D}$-classes est la suivante:

$$\boxed{{}^*\,1}$$

$$\boxed{{}^*\,b} \qquad \boxed{a} \qquad \boxed{{}^*\,c}$$

$$\begin{array}{|c|c|}\hline {}^*\,a^2 & {}^*\,ab \\ \hline {}^*\,ca & {}^*\,bc \\ \hline \end{array}$$

Ce monoïde n'est ni commutatif, ni $\mathcal{R}$-trivial, ni $\mathcal{L}$-trivial, mais il est apériodique.

Question 6. Comme M est apériodique, le langage L est sans-étoile. Une expression sans étoile pour L s'obtient en observant d'abord que $A^* = \emptyset^c$, puis que si B est un sous-ensemble de B, $B^* = A^* - \sum_{b \in A-B} A^*bA^*$. Si on part de l'expression $b^*aA^*ac^*$ donnée plus haut, on obtient ainsi une expression sans étoile.

Question 7. Il suffit de traduire l'expression $b^*aA^*ac^*$ en formule. Il suffit de dire qu'il existe deux positions $x < y$ qui portent la lettre a, qu'il n'y a que des b avant la position x et que des c après la position y. Ce qui donne

$$\varphi = \exists x\ \exists y\ (x < y) \wedge \mathbf{a}x \wedge \mathbf{a}y \wedge \forall z\ (z < x \rightarrow \mathbf{b}z) \wedge (z > y \rightarrow \mathbf{c}z)$$

On peut économiser une variable en prenant

$$\psi = (\exists x\ \exists y\ (x < y) \wedge \mathbf{a}x \wedge \mathbf{a}y) \wedge \forall y\ (y < x \rightarrow \mathbf{b}z) \wedge \forall x\ (x > y \rightarrow \mathbf{c}z)$$

Si on ne dispose que d'une seule variable, on ne peut définir que des langages commutatifs. Or le langage L n'est pas commutatif.

Partie 2. Constantes

Question 8. Soit $s \in P$ tel que s soit une constante pour P et supposons que $usv \in P$ entraîne $us, sv \in P$. En appliquant la définition d'une constante avec $s_1 = u$, $s_2 = v$ et $s_3 = s_4 = 1$, on obtient $us \in P$ et en l'appliquant avec $s_1 = s_2 = 1$, $s_3 = u$ et $s_4 = v$, on trouve $sv \in P$.

Question 9. Soit u un mot de A^* et soit $s = \eta(u)$. Supposons que u sont une constante pour L et soient $s_1, s_2, s_3, s_4 \in M$. Comme η est surjective, on peut trouver des mots u_1, u_2, u_3, u_4 tels que $\eta(u_i) = s_i$ pour $1 \leqslant i \leqslant 4$. Supposons que $s_1ss_2, s_3ss_4 \in P$. On a $L = \eta^{-1}(P)$, $\eta(u_1uu_2) = s_1ss_2$ et $\eta(u_3uu_4) = s_3ss_4$ et donc $u_1uu_2, u_3uu_4 \in L$ et $u_1uu_4 \in L$ puisque u est une constante de L. Donc s_1ss_3, qui est égal à $\eta(u_1uu_4)$, appartient à P. Par conséquent, s est une constante pour P.

Réciproquement, si s est une constante pour P et si u_1, u_2, u_3, u_4 sont des mots de A^* tels que $u_1uu_2, u_3uu_4 \in L$, on a, en posant $s_i = \eta(u_i)$ $(1 \leqslant i \leqslant 4)$, $s_1ss_2, s_3ss_4 \in \eta(L) = P$ et donc $s_1ss_4 \in P$, d'où finalement $u_1uu_4 \in L$.

Question 10. Rappelons que les états de $\mathcal{A}$ s'identifient aux langages non vides de la forme $u^{-1}L$ $(u \in A^*)$. Soit u une constante pour L et soient $q_1 = u_1^{-1}L$ et $q_3 = u_3^{-1}L$ deux états de $\mathcal{A}$. Supposons que les états $q_2 = q_1 \cdot u$ et $q_4 = q_3 \cdot u$ soient définis et soient $u_2 \in q_2$ et $u_4 \in q_4$. Comme $q_2 = q_1 \cdot u = (u_1u)^{-1}L$, on a $u_1uu_2 \in L$ et de même $u_3uu_4 \in L$. Comme u est une constante pour L, il en résulte $u_1uu_4 \in L$ et donc $u_4 \in q_2$. On montrerait de la même façon que $u_4 \in q_2$ entraine $u_4 \in q_4$ et donc $q_2 = q_4$.

Supposons maintenant que u soit de rang $\leqslant 1$. Soient u_1, u_2, u_3, u_4 des mots de A^* tels que $u_1uu_2, u_3uu_4 \in L$. Posons $q_1 = u_1^{-1}L$ et $q_3 = u_3^{-1}L$. On a alors $u_2 \in q_1 \cdot u$ et $u_4 \in q_3 \cdot u$ et donc $q_1 \cdot u = q_3 \cdot u$. Il en résulte que $u_1uu_4 \in L$ et donc u est une constante pour L.

Question 11. Si on reprend l'exemple de la première partie, l'ensemble des constantes pour P s'obtient en prenant les mots de rang $\leqslant 1$ dans l'automate: ce sont donc exactement les éléments de l'idéal minimal de M.

Question 12. Revenons au cas général. La formule $\mathrm{rg}(uv) \leqslant \max\{\mathrm{rg}(u), \mathrm{rg}(v)\}$ montre que $C(L)$ est un idéal. Dans le cas d'un monoïde quelconque M, c'est également évident. Prenons une constante s et des éléments x, y de M. Si $s_1(xsy)s_2, s_3(xsy)s_4 \in P$, on a $(s_1x)s(ys_4) \in P$ puisque s est une constante et donc xsy est une constante. Par conséquent, $C(P)$ est un idéal de M. Cet idéal est apériodique car les constantes sont des éléments de rang $\leqslant 1$ dans M.

Question 13. Si L n'est pas dense, il existe un mot u incomplétable dans L, i.e. tel que $A^*uA^* \cap L = \emptyset$. L'élément $\eta(u)$ est alors un zéro de M. En effet, si $s, t \in A^*$, le mot sut est incomplétable dans L et donc $sut \sim_L u$ et $\eta(sut) = \eta(u)$. Par conséquent, $\eta(u)$ est un zéro de M.

Soit s une constante pour P et soit $u \in M$. Supposons que $sus \neq 0$. Il existe alors des éléments $x, y \in M$ tels que $xsusy \in P$. En prenant $s_1 = x$, $s_2 = usy$, $s_3 = xsu$ et $s_4 = y$, on en déduit

$xsy \in P$. Réciproquement, supposons que $xsy \in P$. Comme $sus \neq 0$, il existe $r, t \in M$ tels que $rsust \in P$. On en déduit $xsust \in P$ puis $xsusy \in P$. On en déduit que $s \sim_P sus$ et donc $sus = s$ puisque M est le monoïde syntactique de P. Par conséquent $sMs \subseteq \{s, 0\}$.

Question 14. Même raisonnement, mais il n'y a pas de zéro.