MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

Examen du 2 février 2010. Durée: 2h 30, notes de cours autorisées

⋆ ⋆ ⋆

Avertissement : On attachera une grande importance à la clarté, à la précision et à la concision de la rédaction. Les parties 2, 3 et 4 sont indépendantes.

Dans ce problème, on s'intéresse à la classe de langages $\mathcal{V}$ ainsi définie: pour chaque alphabet A, $\mathcal{V}(A^*)$ est l'ensemble des langages qui sont unions finies de langages de la forme

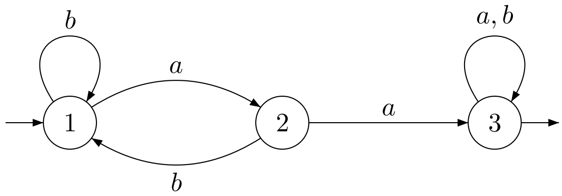
$$(*) \qquad A_0^* a_1 A_1^* \cdots a_k A_k^*$$

où $A_0, \ldots, A_k$ sont des parties de A (éventuellement vides) et $a_1, \ldots, a_k \in A$.

On rappelle qu'un langage est *sans étoile non ambigu* si et seulement si son monoïde syntactique appartient à la variété $\mathbb{D}\mathbf{A}$.

1. Quelques exemples

On considère sur l'alphabet $A = \{a, b\}$ le langage $L = A^*aaA^*$, dont voici l'automate minimal:



Question 1. Calculer le monoïde syntactique M de L (on trouvera 6 éléments). Donner la liste des idempotents de M (on trouvera 5 idempotents) et sa structure en $\mathcal{D}$-classes (on trouvera 3 $\mathcal{D}$-classes).

Question 2. Est-ce que M est commutatif? apériodique? Appartient-il à la variété $\mathbb{D}\mathbf{A}$? Justifier chacune de vos réponses.

Question 3. On rappelle la définition de l'ordre syntactique $\leqslant_L$ sur A^*: on a $u \leqslant_L v$ si et seulement si, pour tout $x, y \in A^*$, $xvy \in L \implies xuy \in L$. Montrer que $a^2 <_L a$, $1 <_L b$, $ab <_L 1$ et $ba <_L 1$. En déduire une description complète de l'ordre syntactique sur M.

On considère un langage K de la forme $(*)$ qui vérifie les conditions supplémentaires suivantes: $a_1 \notin A_0 \cup A_1$, $a_2 \notin A_1 \cup A_2$, ..., $a_k \notin A_{k-1} \cup A_k$. L'automate minimal de K se trouve à [Chapitre 9, section 1.3] dans le support de cours.

Question 4. Montrer que le monoïde syntactique de K est $\mathcal{J}$-trivial (on pourra utiliser le Théorème 1.10 p.173). Montrer que K est sans étoile non ambigu.

2. Projections

Un morphisme φ de A^* dans B^* est une *projection* si, pour chaque lettre $a \in A$, $\varphi(a)$ est une lettre de B.

Question 5. Montrer que la classe $\mathcal{V}$ contient les langages sans étoile non ambigus et est fermée par projection. Plus formellement, montrer que si φ est une projection de A^* dans B^* et si $L \in \mathcal{V}(A^*)$, alors $\varphi(L) \in \mathcal{V}(B^*)$.

Question 6. Réciproquement, montrer que tout langage de $\mathcal{V}$ est projection d'un langage sans étoile non ambigu. Plus formellement: si $L \in \mathcal{V}(A^*)$, il existe un alphabet C, une projection $\varphi : C^* \to A^*$ et un langage sans étoile non ambigu K de C^* tels que $L = \varphi(K)$.

Question 7. Soit $\varphi : A^* \to B^*$ une projection et soit L un langage de A^*. Montrer que si L est reconnu par un monoïde M, alors $\varphi(L)$ est reconnu par le monoïde $\mathcal{P}(M)$ des parties de M (voir p.15 du support de cours).

3. Un peu de logique...

On rappelle que $\Sigma_2[<]$ est l'ensemble des formules de $\mathbf{FO}[<]$ de la forme

$$\exists x_1 \exists x_2 \ \cdots \ \exists x_p \ \forall y_1 \forall y_2 \ \cdots \ \forall y_q \ \psi(x_1, \ldots, x_p, y_1, \ldots, y_q)$$

où ψ est une formule sans quantificateurs.

Question 8. Donner explicitement une formule de $\Sigma_2[<]$ définissant le langage L de la question 1.

Question 9. Plus généralement, montrez que tout langage de $\mathcal{V}$ peut être défini par une formule de $\Sigma_2[<]$.

3. Retour sur $\mathcal{V}$

Question 10. Montrer que $\mathcal{V}$ est une variété positive.

Question 11. Soient x et y deux mots contenant les mêmes lettres (comme *acbaaacb* et *bac*). Montrer que tout langage de $\mathcal{V}$ vérifie l'équation $x^\omega y x^\omega \leqslant x^\omega$.

MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

February 3, 2009. Duration: 2h 30.

⋆⋆⋆

Warning : Clearness, accuracy and concision of the writing will be rewarded. Parts 2, 3 and 4 are independent.

In this problem, we are interested in the class of languages $\mathcal{V}$ defined as follows: for each alphabet A, $\mathcal{V}(A^*)$ is the set of languages that are finite unions of languages of the form

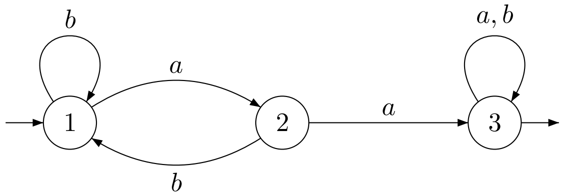
$$(*) \qquad A_0^* a_1 A_1^* \cdots a_k A_k^*$$

where $A_0, \ldots, A_k$ are subsets of A (possibly empty) and $a_1, \ldots, a_k \in A$.

Recall that a language is *unambiguous star-free* if and only if its si syntactic monoid belongs to the variety $\mathbb{D}\mathbf{A}$.

1. A few examples

Consider on the alphabet $A = \{a, b\}$ the language $L = A^*aaA^*$, whose minimal automaton is:



Question 1. Compute the syntactic monoid M of L (one will find 6 elments). Give the list of idempotents of M (one will find 5 idempotents) and its $\mathcal{D}$-class structure (one will find 3 $\mathcal{D}$-classes).

Question 2. Is M commutative? aperiodic? Does it belong to the variety $\mathbb{D}\mathbf{A}$? Justify each of yours answers.

Question 3. Let us recall the definition of the syntactic order $\leqslant_L$ on A^*: one has $u \leqslant_L v$ if and only if, for all $x, y \in A^*$, $xvy \in L \implies xuy \in L$. Show that $a^2 <_L a$, $1 <_L b$, $ab <_L 1$ et $ba <_L 1$. Derive from these relations a complete description of the syntactic order on M.

Consider a language K of the form $(*)$ which also satisfies the following conditions: $a_1 \notin A_0 \cup A_1$, $a_2 \notin A_1 \cup A_2$, ..., $a_k \notin A_{k-1} \cup A_k$. The minimal automaton of K can be found at [Chapitre 9, section 1.3] in the yellow book.

Question 4. Show that the syntactic monoid of K is $\mathcal{J}$-trivial (one could use for instance Theorem 1.10 p.173). Show that K is unambiguous star-free.

2. Projections

A morphism φ from A^* to B^* is a *projection* if, for each letter $a \in A$, $\varphi(a)$ is a letter of B.

Question 5. Show that the class $\mathcal{V}$ conatins the unambiguous star-free languages and is closed under projection. More formally, show that if φ is a projection from A^* to B^* and if $L \in \mathcal{V}(A^*)$, then $\varphi(L) \in \mathcal{V}(B^*)$.

Question 6. Conversely, show that every language of $\mathcal{V}$ is the projection of an unambiguous star-free language. More formally: if $L \in \mathcal{V}(A^*)$, there exist an alphabet C, a projection $\varphi : C^* \to A^*$ and an unambiguous star-free language K of C^* such that $L = \varphi(K)$.

Question 7. Let $\varphi : A^* \to B^*$ be a projection and let L be a language of A^*. Show that if L is recognized by a monoid M, then $\varphi(L)$ is recognized by the monoid $\mathcal{P}(M)$ of subsets of M (see p.15 of the yellow book).

3. Some logic...

Recall that $\Sigma_2[<]$ is the set of formulas of $\mathbf{FO}[<]$ of the form

$$\exists x_1 \exists x_2 \ \cdots \ \exists x_p \ \forall y_1 \forall y_2 \ \cdots \ \forall y_q \ \psi(x_1, \ldots, x_p, y_1, \ldots, y_q)$$

where ψ is a quantifier-free formula.

Question 8. Give explicitly a formula of $\Sigma_2[<]$ defining the language L of Question 1.

Question 9. More generally, prove that every language of $\mathcal{V}$ can be defined by a $\Sigma_2[<]$-formula.

4. Back to $\mathcal{V}$

Question 10. Show that $\mathcal{V}$ is a positive variety.

Question 11. Let x and y be two words containing the same letters (like *acbaaacb* and *bac*). Prove that every language of $\mathcal{V}$ satisfies the equation $x^\omega y x^\omega \leqslant x^\omega$.

Solution

An example

Question 1. The syntactic monoid of L is generated by the following generators:

	1	2	3
a	2	3	3
b	1	1	3

Elements:

	1	2	3
$* \ 1$	1	2	3
a	2	3	3
$* \ b$	1	1	3
$* \ a^2$	3	3	3
$* \ ab$	1	3	3
$* \ ba$	2	2	3

Note that a^2 is a zero of M. Thus we set $a^2 = 0$. The other relations defining M are:

$$b^2 = b \qquad aba = a \qquad bab = b$$

Idempotents:

$$E(S) = \{1, b, a^2, ab, ba\}$$

$\mathcal{D}$-classes:

^{*1}

a	$^*ab$
$^*ba$	$^*b$

$^*a^2$

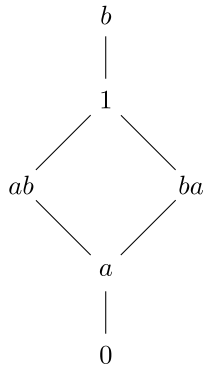
Question 2. This monoid is not commutative, since $ab \neq ba$. This monoid is aperiodic. This monoid is not in $\mathbb{D}\mathbf{A}$ since the $\mathcal{D}$-class of b is regular, but contains a, which is not idempotent. Other proof: the identity $(xy)^\omega(yx)^\omega(xy)^\omega = (xy)^\omega$ is not satisfied for $x = a$ and $y = b$.

Question 3. One has $u \leqslant_L v$ if and only if, for all $x, y \in A^*$, $xvy \in L$ implies $xuy \in L$. Since L is the set of words containing the factor aa, one has $aa \leqslant_L v$ for all words v.

Next observe that if $xby \in L$, then $xy \in L$ and if $xy \in L$, then both $xaby$ and $xbay$ are in L. It follows that $1 \leqslant_L b$, $ab \leqslant_L 1$ et $ba \leqslant_L 1$. Since these words have distinct syntactic images in M, all these relations are actually strict. It follows that in M, $0 < a$, $ab < 1$, $ba < 1$ and $1 < b^2$. Since the order is stable, one also gets $a < ab$ and $a < ba$. Finally, the syntactic order is

$$0 < a \qquad a < ab \qquad a < ba \qquad ab < 1 \qquad ba < 1 \qquad 1 < b$$

and it can be represented in the following diagram



Question 4. Let L be a language is of the form $A_0^*a_1A_1^*a_2 \cdots a_kA_k^*$, where $k \geqslant 0$, $a_1, \ldots, a_k \in A$, $A_k \subseteq A$ and, for $0 \leqslant i \leqslant k-1$, A_i is a subset of $A \setminus \{a_{i+1}\}$. It follows from Theorem 1.10 (p.173) that its syntactic monoid is $\mathcal{R}$-trivial.

Since K and K^r are both languages of this form, their syntactic monoids are both $\mathcal{R}$-trivial. It follows that the syntactic monoid of K is both $\mathcal{R}$-trivial and $\mathcal{L}$-trivial and hence is $\mathcal{J}$-trivial. In particular, it is in $\mathbb{D}\mathbf{A}$ and thus K is unambiguous star-free.

2. Projections

Question 5. Every unambiguous star-free language is a finite union of unambiguous products of the form $(*)$. Therefore, $\mathcal{V}$ contains all the unambiguous star-free languages.

Since projections commute with union, it suffices to prove that the projection of a language of the form $(*)$ is of the same form, which is trivial.

Question 6. Again, it suffices to consider a language L of the form $(*)$. For $1 \leqslant i \leqslant k$, let B_i be a copy of A_i and let $b_1, \ldots, b_k$ be new distinct letters. Let C be the disjoint union of the sets B_i $(0 \leqslant i \leqslant k)$ and $\{b_1, \ldots, b_k\}$. Let $\varphi : C^* \to A^*$ be the projection sending each B_i onto its copy A_i and each b_i onto a_i. By Question 4, the language $K = B_0^*b_1B_1^* \cdots b_kB_k^*$ is unambiguous star-free and satisfies $\varphi(K) = L$.

Question 7. Let $\varphi : A^* \to B^*$ be a projection and let $\eta : A^* \to M$ be a morphism recognizing L. Let

$$P = \eta(L) \text{ and } Q = \{X \in \mathcal{P}(M) \mid X \cap P \neq \emptyset\}$$

Let $\gamma : B^* \to \mathcal{P}(M)$ be the morphism defined, for each $b \in B$, by $\gamma(b) = \eta(\varphi^{-1}(b))$. One gets

$$\begin{aligned}
\gamma^{-1}(Q) &= \{u \in B^* \mid \gamma(u) \in Q\} = \{u \in B^* \mid \gamma(u) \cap P \neq \emptyset\} \\
&= \{u \in B^* \mid \eta(\varphi^{-1}(u)) \cap P \neq \emptyset\} = \{u \in B^* \mid \varphi^{-1}(u) \cap \eta^{-1}(P) \neq \emptyset\} \\
&= \{u \in B^* \mid \varphi^{-1}(u) \cap L \neq \emptyset\} = \varphi(L)
\end{aligned}$$

Therefore $\varphi(L)$ is recognized by $\mathcal{P}(M)$.

3. Some logic...

Question 8. One has $L = L(\varphi)$, with

$$\varphi = \exists x_1 \, \exists x_2 \, \forall y \; (x_1 < x_2) \wedge \mathbf{a}x_1 \wedge \mathbf{a}x_2 \wedge \big((x_1 < y) \rightarrow (x_2 \leqslant y)\big)$$

Question 9. It suffices to prove the result for $L = A_0^* a_1 A_1^* \cdots a_k A_k^*$. then one has $L = L(\varphi)$ with $\varphi = \exists x_1 \, \exists x_2 \, \cdots \, \exists x_k \; \forall y \; \psi$, where

$$\begin{aligned}\psi = (x_1 < x_2 < \, \cdots \, < x_k) \wedge \Bigg((y < x_1) \rightarrow \Big(\bigvee_{a \in A_0} \mathbf{a}y\Big)\Bigg) \\ \wedge \Bigg((x_0 < y < x_1) \rightarrow \Big(\bigvee_{a \in A_1} \mathbf{a}y\Big)\Bigg) \wedge \cdots \wedge \Bigg((x_{k-1} < y < x_k) \rightarrow \Big(\bigvee_{a \in A_{k-1}} \mathbf{a}y\Big)\Bigg) \\ \wedge \Bigg((x_k < y) \rightarrow \Big(\bigvee_{a \in A_k} \mathbf{a}y\Big)\Bigg)\end{aligned}$$

Therefore, every language of $\mathcal{V}$ can be defined by a $\Sigma_2[<]$-formula.

4. Back to $\mathcal{V}$

Question 10. Let us show that $\mathcal{V}$ is closed under quotients. By induction, it suffices to prove that if $L \in \mathcal{V}(A^*)$ and $a \in a$, then $a^{-1}L$ and La^{-1} are both in $\mathcal{V}(A^*)$. Since quotients commute with union, we may suppose that $L = A_0^* a_1 A_1^* \cdots a_k A_k^*$. If $a_1 \neq a$, one has

$$a^{-1}L = \begin{cases} L & \text{if } a \in A_0 \\ \emptyset & \text{if } a \notin A_0 \end{cases}$$

and if $a_1 = a$, one gets

$$a^{-1}L = \begin{cases} L \cup A_1^* a_2 \cdots a_k A_k^* & \text{if } a \in A_0 \\ A_1^* a_2 \cdots a_k A_k^* & \text{if } a \notin A_0 \end{cases}$$

The proof for La^{-1} is symmetrical.

Let $\varphi : B^* \rightarrow A^*$ be a morphism. Let us show that if $L \in \mathcal{V}(A^*)$, then $\varphi^{-1}(L) \in \mathcal{V}(B^*)$.

Question 11. Let $L = A_0^* a_1 A_1^* \cdots a_k A_k^*$. Since x and y have the same content, their images in the syntactic monoid of $A_0^*, \ldots, A_k^*$ are equal and idempotent. Therefore, by [Chap.XIII, Theorem 5.1], L satisfies the equation $x^\omega y x^\omega \leqslant x^\omega$. Now, the languages satisfying this equation form a lattice and are in particular closed under finite union. It follows that every language of $\mathcal{V}$ satisfies this equation.