MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

Partiel du 20 novembre 2007. 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.

Exercice 1. Automate minimal

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

Question 2. Calculer le monoïde syntactique M de L; on donnera la liste des éléments et les relations permettant de définir M.

Question 3. Déterminer les idempotents de M.

Question 4. Déterminer l'idéal minimal de M et sa structure en $\mathcal{D}$-classes (on dessinera les diagrammes boîtes à œufs).

Question 5. Selon vous, L est-il sans-étoile? Testable par morceaux? Justifier votre réponse.

Question 6. On considère la signature logique $\{\mathbf{a} \mid a \in A\} \cup \{<, S\}$, où $<$ est interprétée comme la relation d'ordre habituelle sur les entiers et S comme la relation successeur. Trouvez une formule φ, aussi simple que possible, telle que $L \cap A^+ = L(\varphi)$.

Question 7. On prend maintenant la signature restreinte $\{\mathbf{a} \mid a \in A\} \cup \{S\}$. Est-il possible de décrire $L \cap A^+$ à l'aide d'une formule du premier ordre dans cette signature?

Exercice 2. Langages de la forme F^* avec F fini.

Soit L un langage de A^*. On dit qu'un mot u de A^* est *complétable* dans L s'il existe des mots $x, y \in A^*$ tels que $xuy \in L$, *incomplétable* dans le cas contraire. Un langage est *dense* si tous les mots de A^* sont complétables dans L.

Question 8. Déterminer, pour chacun des langages ci-dessous, s'ils sont denses ou pas, en justifiant votre réponse (l'alphabet est $A = \{a, b\}$).
(1) $\{aa, aba, b\}^*$
(2) $\{aa, aba, abb, ba, bb\}^*$
(3) $\{a, aab, bab, bb\}^*$

Question 9. Montrer que les mots incomplétables de L sont tous équivalents pour $\sim_L$, la congruence syntactique de L.

Question 10. En déduire que si L n'est pas dense, le monoïde syntactique de L possède un zéro.

Soit F un langage *fini* de A^*. On s'intéresse aux propriétés syntactiques du langage $L = F^*$. On note M le monoïde syntactique, $\eta : A^* \to M$ le morphisme syntactique de L et on pose $P = \eta(L)$.

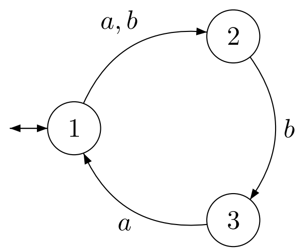
Question 11. Soit u un mot non vide et complétable dans L. On suppose que $\eta(u)$ est un idempotent de M. Montrer qu'il existe une factorisation $u = xy$ et un entier $n > 0$ tels que $(yx)^n \in L$.

Question 12. En déduire que P rencontre chaque $\mathcal{D}$-classe régulière de M autre que celle de 1 et différente du zéro dans le cas où L n'est pas dense.

Corrigé

Exercice 1. Automate minimal

Question 1. L'automate minimal du langage $L = \{aba, bba\}^*$ 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	1
b	2	3	0
a^2	0	0	2
ab	3	0	2
ba	0	1	0
b^2	3	0	0
$*\ a^3$	0	0	0
$*\ a^2b$	0	0	3
$*\ aba$	1	0	0
$*\ ba^2$	0	2	0
a^2ba	0	0	1
aba^2	2	0	0
ba^2b	0	3	0

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

$$abb = a^2b \qquad bab = ba^2 \qquad b^2a = aba \qquad b^2b = 0 \qquad a^3 = 0$$
$$ba^2a = 0 \qquad a^2baa = a^2 \qquad aba^2b = b^2 \qquad ba^2ba = ba$$

Question 3. Les idempotents sont:

$$E(M) = \{1, a^3, a^2b, aba, ba^2\}$$

Question 4. L'idéal minimal est:

$$I = \{a^3\}$$

et la structure en $\mathcal{D}$-classes est la suivante:

^{*1}

a	b

ab

$^*ba^2$	ba	ba^2b
aba^2	$^*aba$	b^2
a^2	a^2ba	$^*a^2b$

$^*a^3$

Question 5. Comme M est apériodique, le langage L est sans-étoile. En revanche, M n'est pas $\mathcal{J}$-trivial (puisque par exemple $aa \mathrel{\mathcal{R}} a^2b$) et donc L n'est pas testable par morceaux.

Question 6. Voir question suivante.

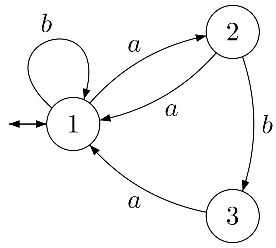
Question 7. Montrons que $L \cap A^+$ est égal au langage K des mots commençant par *aba* ou *bba*, finissant par *ba* et ne contenant pas les facteurs *aaa*, *bbb*, *baba* et *aabb*. Ces contraintes s'expriment simplement à l'aide du successeur.

Partie 2. Langages de la forme F^* avec F fini.

Question 8. Pour ces questions, on peut par exemple calculer l'automate minimal des langages considérés et vérifier que cet automate est complet et fortement connexe, car alors, on peut, partant d'un état accessible quelconque, atteindre un état final.

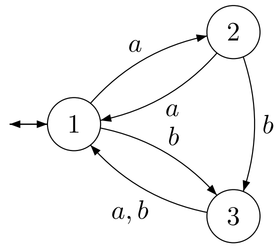
En fait, comme ces langages sont tous de la forme F^* où F est un code fini, il existe une façon beaucoup plus expéditive (mais hors programme...) de résoudre cet exercice: il suffit de vérifier que, en prenant $\pi(a) = \pi(b) = 1/2$, la probabilité de F est égale à 1.

(1) L'automate minimal du langage $\{aa, aba, b\}^*$ est



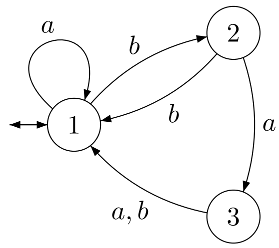
Le mot *abbabb* est incomplétable dans le langage $\{aa, aba, b\}^*$ car aucun chemin de cet automate n'est étiqueté par ce mot. Ici $\pi(F) = 1/4 + 1/8 + 1/2 < 1$.

(2) L'automate minimal du langage $\{aa, aba, abb, ba, bb\}^*$ est donné ci-dessous:



Cet automate est complet et fortement connexe. On a aussi $\pi(F) = 1/4 + 1/8 + 1/8 + 1/4 + 1/4 + 1/4 = 1$.

(3) Une petite astuce pour simplifier les calculs. Il est clair que L est dense si et seulement si le langage L^r, obtenu en retournant tous les mots du L, est dense. On a ici $L^r = \{a, baa, bab, bb\}^*$. L'automate minimal de ce langage est:



On a aussi $\pi(F) = 1/2 + 1/8 + 1/8 + 1/4 = 1$.

Question 9. Soient u et v incompétables. Alors par définition, on a, pour tout $x, y \in A^*$, $xuy, xvy \notin L$. Donc $u \sim_L v$.

Question 10. Si u est un mot incomplétable, tous les mots de A^*uA^* sont incomplétables et sont donc syntactiquement équivalents à u. Soit $\eta : A^* \to M$ le morphisme syntactique et soit $z = \eta(u)$. Il résulte de ce qui précède que $\eta(A^*uA^*) = \eta(u)$ et donc $MzM = z$. Autrement dit, z est un zéro de M.

Question 11. Puisque u est complétable, il existe des mots $x, y \in A^*$ tels que $xuy \in L$. Et comme $\eta(u)$ est idempotent, on a, pour tout $n > 0$, $u \sim_L u^n$ et donc $xu^ny \in L$. Soit k la longueur maximale des mots de F et prenons $r = 1+|u^{k+1}|$ et $n = (k+1)r$. Par construction, chaque facteur u^{k+1} est plus long que tous les mots de F. On peut donc trouver des mots $f_0, g_0, f_1, g_1, \ldots, f_r, g_r$ tels que $|f_0|, \ldots, |f_r| \leqslant k$, $f_0g_0 = f_1g_1 = \cdots = f_rg_r = u^{k+1}$ et $xf_0, g_0f_1, \ldots, g_{r-1}f_r, g_ry \in F^*$. D'après le choix de r, il existe deux entiers $i < j$ tels que $f_i = f_j$. On a alors $g_if_{i+1} \cdots g_{j-1}f_j = (g_if_i)^{j-i} \in F^*$.

Question 12. Soit D une $\mathcal{D}$-classe régulière autre que celle de 1 et différente du zéro dans le cas où L n'est pas dense. Ces hypothèses assurent que tous les mots u tels que $\eta(u) \in D$ sont non vides et complétables. Comme D est régulière, on peut choisir un tel mot tel que $\eta(u) = e$ soit idempotent. D'après la question précédente, il existe une factorisation $u = xy$ et un entier $n > 0$ tels que avec $(yx)^n \in L$. On a alors $\eta(u^n) = e^n = e$. Comme $(yx)^n$ est un conjugué de $(xy)^n$, $\eta((yx)^n)$ est une idempotent conjugué de e. Il est donc $\mathcal{D}$-équivalent à D et il appartient à P puisque $(yx)^n \in L$.