# MPRI, Fondations mathématiques de la théorie des automates

Olivier Carton, Jean-Éric Pin

Partiel du 23 novembre 2015. Durée: 1h 30, 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. Étude d'un langage.

Soit $A = \{a, b, c\}$ et soit $L = c(ab)^*$.

**Question 1.** Trouver l'automate minimal de $L$.

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

**Question 3.** Donner la liste des idempotents de $M$.

**Question 4.** Déterminer la structure en $\mathcal{D}$-classes de $M$ et dessiner les diagrammes boîtes à œufs.

**Question 5.** Le monoïde $M$ est-il apériodique? $\mathcal{R}$-trivial? Justifier votre réponse.

## Partie 2. Langages de la forme $P^*$ avec $P$ fini.

Dans ce problme, $P$ est un langage de $A^+$. On dit que $P$ est un *code prfixe* si aucun mot de $P$ n'est prfixe propre d'un autre mot de $P$, autrement dit si les conditions $u, uv \in P$ entrainent $v = 1$.

**Question 6.** Déterminer si les ensembles suivants sont des codes prfixes (l'alphabet est $\{a, b, c\}$ pour les deux premiers exemples et $\{a, b\}$ pour les deux derniers):

(1) $c\{a, b\}^*$
(2) $\{a, b\}^* c$
(3) $\{a, ba, bb\}$
(4) $\{aa, aba, ba, bb\}$

**Question 7.** Montrer que si $P$ est un code prfixe, alors, pour tout $u, v \in A^*$, les conditions $u, uv \in P^*$ entrainent $v \in P^*$. En dduire que tout mot de $P^*$ s'crit de faon unique comme produit de mots de $P$.

**Question 8.** On dit qu'un code prfixe fini est *maximal* s'il n'est contenu dans aucun autre code prfixe fini autre que lui-mme. Quels sont les codes prfixes finis maximaux de la question 6?

**Question 9.** Soit $P$ un code prfixe fini. Montrer que les conditions suivantes sont quivalentes:

(1) $P$ est un code prfixe fini maximal,

(2) pour tout $u \in A^*$, il existe $v \in A^*$ tel que $uv \in PA^*$,

(3) pour tout $u \in A^*$, il existe $v \in A^*$ tel que $uv \in P^*$.

Soit $P$ un code prfixe fini maximal. On note $\eta : A^* \to M$ le morphisme syntactique du langage $P^*$.

**Question 10.** Montrer que si $u$ est un mot tel que $\eta(u)$ est idempotent, alors il existe une factorisation $u = xy$ et un entier $n > 0$ tels que $(yx)^n \in P^*$.

**Question 11.** En déduire que $\eta(P^*)$ rencontre chaque $\mathcal{D}$-classe régulière de $M$.

# Partie 3. Un peu de probabilits.

Soit $\pi : A \to ]0, 1]$ une fonction telle que $\sum_{a \in A} \pi(a) = 1$. On tend $\pi$  $A^*$ en posant $\pi(1) = 0$ et $\pi(a_1 a_2 \cdots a_n) = \pi(a_1)\pi(a_2) \cdots \pi(a_n)$. Finalement, on pose $\pi(P) = \sum_{u \in P} \pi(u)$.

**Question 12.** On prend $A = \{a, b\}$ et on pose $\pi(a) = p$ et $\pi(b) = q = 1 - p$. Calculer $\pi(P)$ pour $P = \{a, ba, bb\}$ et $P = \{aa, aba, ba, bb\}$.

**Question 13.** Montrer que si $P$ est un code prfixe fini, alors $\pi(P) \leqslant 1$.

**Question 14.** Montrer que si $P$ est un code prfixe fini maximal, alors $\pi(P) = 1$.

**Question 15.** Montrer que si $P$ est un code prfixe fini tel que $\pi(P) = 1$, alors $P$ est un code prfixe fini maximal.

# MPRI, Mathematical foundations of automata theory

Olivier Carton, Jean-Éric Pin

November 23, 2015. Duration: 1h 45.

⋆ ⋆ ⋆

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

## Part 1. Study of a language

Let $A = \{a, b, c\}$ and let $L = c(ab)^*$.

**Question 1.** Find the minimal automaton of $L$.

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

**Question 3.** Give the list of all idempotents of $M$.

**Question 4.** Give the $\mathcal{D}$-class structure of $M$ and draw the corresponding egg-box pictures.

**Question 5.** Is $M$ aperiodic? $\mathcal{R}$-trivial? Justify your answer.

## Part 2.

In this problem, $P$ is a language of $A^+$. One says that $P$ is a *prefix code* if no word of $P$ is a proper prefix of another word of $P$, that is, if the conditions $u, uv \in P$ imply $v = 1$.

**Question 6.** Find out whether the following sets are prefix codes (the alphabet is $\{a, b, c\}$ for the firsst two examples and $\{a, b\}$ for the last two ones):

(1) $c\{a, b\}^*$
(2) $\{a, b\}^*c$
(3) $\{a, ba, bb\}$
(4) $\{aa, aba, ba, bb\}$

**Question 7.** Prove that if $P$ is a prefix code, then for all $u, v \in A^*$, the conditions $u, uv \in P^*$ imply $v \in P^*$. Deduce that every word of $P^*$ can be uniquely written as a product of words of $P$.

**Question 8.** A finite prefix code is *maximal* if it is not contained in any other finite prefix code (apart from itself). Find out the maximal finite prefix codes of question 6.

**Question 9.** Let $P$ be a finite prefix code. Prove that the following conditions are equivalent:

(1) $P$ is a maximal finite prefix code,
(2) for all $u \in A^*$, there exists $v \in A^*$ such that $uv \in PA^*$,
(3) for all $u \in A^*$, there exists $v \in A^*$ such that $uv \in P^*$.

Let $P$ be a maximal finite prefix code. We denote by $\eta : A^* \to M$ the synctactic morphism of the language $P^*$.

**Question 10.** Show that if $u$ is a word such that $\eta(u)$ is idempotent, then there exists a factorisation $u = xy$ and an integer $n > 0$ such that $(yx)^n \in P^*$.

**Question 11.** Conclude that $\eta(P^*)$ meets every regular $\mathcal{D}$-class of $M$.

# Partie 3. A bit of probabilities.

Let $\pi : A \to ]0, 1]$ be a function such that $\sum_{a \in A} \pi(a) = 1$. We extend $\pi$ to $A^*$ by setting $\pi(1) = 0$ and $\pi(a_1 a_2 \cdots a_n) = \pi(a_1)\pi(a_2) \cdots \pi(a_n)$. Finally, we set $\pi(P) = \sum_{u \in P} \pi(u)$.

**Question 12.** Let $A = \{a, b\}$ and let us set $\pi(a) = p$ and $\pi(b) = q = 1 - p$. Compute $\pi(P)$ for $P = \{a, ba, bb\}$ and $P = \{aa, aba, ba, bb\}$.

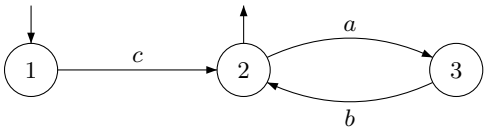
**Question 13.** Prove that if $P$ is a finite prefix code, then $\pi(P) \leqslant 1$.

**Question 14.** Prove that if $P$ is a maximal finite prefix code, then $\pi(P) = 1$.

**Question 15.** Prove that if $P$ is a finite prefix code such that $\pi(P) = 1$, then $P$ is a maximal finite prefix code.

# Corrigé

**Question 1.** L'automate minimal de $L$ est représenté ci-dessous




**Question 2.** Le monoïde syntactique de $L$ est

| | 1 | 2 | 3 |
|---|---|---|---|
| $* 1$ | 1 | 2 | 3 |
| $a$ | 0 | 3 | 0 |
| $b$ | 0 | 0 | 2 |
| $c$ | 2 | 0 | 0 |
| $* a^2$ | 0 | 0 | 0 |
| $* ab$ | 0 | 2 | 0 |
| $* ba$ | 0 | 0 | 3 |
| $ca$ | 3 | 0 | 0 |

Relations:

$$ac = 0 \qquad b^2 = 0 \qquad bc = 0 \qquad cb = 0 \qquad c^2 = 0 \qquad a^2 = 0 \qquad aba = a$$
$$bab = b \qquad cab = c$$

**Question 3.** Idempotents:

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

**Question 4.** $\mathcal{D}$-classes:

| |
|---|
| $* 1$ |

| | |
|---|---|
| $* ba$ | $b$ |
| $a$ | $* ab$ |

| | |
|---|---|
| $ca$ | $c$ |

| |
|---|
| $* a^2$ |

**Question 5.** Le monoïde syntactique de $L$ est apériodique mais n'est pas $\mathcal{R}$-trivial, puisque $b \mathcal{R} ba$.

## Partie 2. Langages de la forme $P^*$ avec $P$ fini.

**Question 6.** Le langage $c\{a,b\}^*$ n'est pas un code prfixe puisque $c, ca \in c\{a,b\}^*$, mais $a \notin c\{a,b\}^*$. En revanche, les trois autres langages $\{a,b\}^*c$, $\{a, ba, bb\}$ et $\{aa, aba, ba, bb\}$ sont des codes prfixes.

**Question 7.** Soit $P$ un code prfixe. Soit $w$ un mot de $P^*$ de longueur minimale tel que $w = uv$ avec $u \in P^*$ mais $v \notin P^*$. Ecrivons $w = w_1w'$ et $u = u_1u'$ avec $w_1, u_1 \in P$. On a alors $w = w_1w' = u_1u'v$, et comme $P$ est prfixe, on a ncessairement $w_1 = u_1$, d'o $w' = u'v$, ce qui contredit la minimalit de $w$.

Soit $w$ un mot de $P^*$ de longueur minimale ayant deux factorisations distinctes en mots de $P$: $w = x_1 \cdots x_r = y_1 \cdots y_s$. On a comme ci-dessus $x_1 = y_1$ et donc le mot $w' = x_2 \cdots x_r = y_2 \cdots y_s$ a deux factorisations distinctes, ce qui contredit la minimalit de $w$.

**Question 8.** Le code prfixe fini $\{aa, aba, ba, bb\}$ est contenu dans le code prfixe fini $\{aa, aba, abbba, bb\}$. Il n'est donc pas maximal. En revanche, le code prfixe fini $P = \{a, ba, bb\}$ est maximal, puisque tout mot de $A^*$ est soit prfixe d'un mot de $P$, soit a un prfixe dans $P$.

**Question 9.** Soit $P$ un code prfixe fini.
(1) entraine (3). Supposons qu'il existe un mot $u \in A^*$ tel que pour tout $v \in A^*$, $uv \notin P^*$. On va montrer que $P$ n'est pas maximal en montrant que $P \cup \{u\}$ est un code prfixe. En effet, si $u$ est prfixe d'un mot de $P$, il existe un mot $v$ tel que $uv \in P$ (et donc $uv \in P^*$), ce qui contredit l'hypothse. Si maintenant $u$ a un prfixe dans $P$, il eiste $v$ tel que $uv \in P$, ce qui contredit galement l'hypothse.
(3) entraine (2). Trivial.
(2) entraine (1). Supposons  prsent que pour tout $u \in A^*$, il existe $v \in A^*$ tel que $uv \in PA^*$. Si $P$ n'est pas maximal, on peut trouver un mot $u \notin P$ tel que $P \cup \{u\}$ est un code prfixe. D'aprs l'hypothse, il existe $v \in A^*$ tel que $uv \in PA^*$. On a donc $uv = pw$ avec $p \in P$ et $w \in A^*$. On en dduit qu'ou bien $p$ est prfixe de $u$, ou bien $u$ est prfixe de $p$, ce qui contredit le fait que $P \cup \{u\}$ est un code prfixe.

**Question 10.** Puisque $P$ est maximal, il existe $y \in A^*$ tels que $uy \in P^*$. Et comme $\eta(u)$ est idempotent, on a, pour tout $n > 0$, $u \sim_L u^n$ et donc $u^n y \in L$. Soit $k$ la longueur maximale des mots de $P$ 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 $P$. 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 $f_0, g_0f_1, \ldots, g_{r-1}f_r, g_ry \in P^*$. D'après le choix de $r$, il existe deux entiers $i < j$ tels que $f_i = f_j$. On a alors $g_if_{i+1} \cdots g_{j-1}f_j = (g_if_i)^{j-i} \in P^*$.

**Question 11.** Soit $D$ une $\mathcal{D}$-classe régulière de $M$. 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 $(yx)^n \in P^*$. On a alors $\eta(u^n) = e^n = e$. Comme $(yx)^n$ est un conjugué de $(xy)^n$, $\eta((yx)^n)$ est un idempotent conjugué de $e$. Il est donc $\mathcal{D}$-équivalent à $D$ et appartient à $\eta(P^*)$.

## Partie 3. Un peu de probabilits.

Soit $\pi : A \to ]0,1]$ une fonction telle que $\sum_{a \in A} = 1$. On tend $\pi$  $A^*$ en posant $\pi(1) = 0$ et $\pi(a_1a_2 \cdots a_n) = \pi(a_1)\pi(a_2) \cdots \pi(a_n)$. Finalement, on pose $\pi(P) = \sum_{u \in P} \pi(u)$.

**Question 12.** Pour $P = \{a, ba, bb\}$, on a $\pi(P) = p + qp + qq = p + q(p + q) = p + q = 1$. Pour $P = \{aa, aba, ba, bb\}$, on a $\pi(P) = pp+pqp+qp+qq = pp+pqp+q(p+q) = p^2+p^2(1-p)+(1-p) = 1 - p + 2p^2 - p^3$.