TD numéro 1

Pierre-Alain Fouque

October 5, 2004

1 Mots

Exercice 1: Deux mots u et v sont *conjugués* s'il existe des mots x et y tels que

$$u = xy, \qquad v = yx$$

(a) Vérifier que la conjuguaison est une relation d'équivalence.

(b) Soient u et v deux mots non vides. Montrer que les conditions suivantes sont équivalentes:

1. $uv = vu$,
2. il existe deux entiers $n, m \geq 1$ tels que $u^n = v^m$,
3. il existe un mot w non vide et deux entiers $k, \ell \geq 1$ tels que $u = w^k$ et $v = w^\ell$.

2 Automates

Exercice 2:

(a) Construire directement, à partir d'automates reconnaissant les langages K et L, des automates reconnaissant les langages $K \setminus L = K \cap \bar{L}$ et $K \Delta L = (K \cup L) \setminus (K \cap L)$.

En conclure que si $\mathcal{A}$ et $\mathcal{B}$ sont deux automates finis quelconques, alors on sait décider de l'égalité des langages reconnaissables $L(\mathcal{A})$ et $L(\mathcal{B})$.

(b) Montrer que la famille des langages reconnaissables est close par préfixe, suffixe et facteur.

(c) Un mot $u = a_1 \dots a_n \in A^*$ est un *sous-mot* d'un mot $v \in A^*$ s'il existe des mots $u_0, \dots, u_n \in A^*$ tels que $v = u_0 a_1 u_1 \dots a_n u_n$. Si $L \subseteq A^*$ est un langage, on note $\mathrm{SM}(L)$ l'ensemble des sous-mots des mots de L. Montrer que si L est un langage reconnaissable, il en est de même du langage $\mathrm{SM}(L)$.

(d) Soient $u, v \in A^*$ deux mots. On définit l'ensemble des *mélanges* ou *shuffles* des mots u et v par :

$$\begin{aligned} u \bowtie v = \{w \in A^* | \exists u_1, v_1, \dots, u_n, v_n \in A^* \text{ tels que} \\ u = u_1 \dots u_n,\ v = v_1 \dots v_n \text{ et } w = u_1 v_1 \dots u_n v_n\} \end{aligned}$$

Par exemple,

$$aba \bowtie bc = \{ababc, abbac, babac, abbca, babca, abcba, bacba, bcaba\}$$

Si $K, L \subseteq A^*$ sont deux langages, on définit : $K \bowtie L = \{w \in A^* | \exists u \in K, \exists v \in L \text{ tels que } w \in u \bowtie v\}$

Montrer que si K et L sont des langages reconnaissables, il en est de même de $K \bowtie L$.

(e) Montrer la proposition suivante:
La classe des langages reconnus par les automates avec ϵ-transitions est la classe des langages reconnaissables. De plus, on dispose d'un algorithme permettant, étant donné un automate avec ϵ-transitions, de construire un automate "sans ϵ-transitions" reconnaissant le même langage.

Exercice 3: Déterminisation d'automate non-déterministe

(a) Construire un automate fini déterministe, sur l'alphabet $A = \{a, b\}$, reconnaissant les mots qui ne contiennent pas le facteur abb.

(b) Construire un automate fini déterministe qui reconnaît les mots qui ne contiennent ni le facteur abb, ni le facteur baa.

Exercice 4: Algorithme de recherche de motifs dans un texte

Un algorithme efficace de recherche de motifs peut être construit en utilisant des automates. L'automate a la propriété de ne reconnaître qu'un seul mot appelé le motif.

(a) Trouver un algorithme naïf de recherche de motif dans un texte et calculer sa complexité.

(b) Trouver un automate permettant de reconnaître le motif $abaabbaaa$.

(c) Évaluer dans le cas général la complexité de cet algorithme en précisant le coût de la construction de l'automate.

Exercice 5: Soit $A = \{a, b\}$

(a) Donner un automate non déterministe avec n états pour $L = A^* a A^{n-2}$

(c) Montrer que tout automate déterministe reconnaissant ce langage L a au moins 2^{n-1} états.