Langages Formels

Licence Informatique – ENS Cachan

Examen du 26 mai 2015

durée 3 heures

Document autorisé : polycopié du cours.
Toutes les réponses devront être correctement justifiées.
Le chiffre en regard d'une question est une indication sur sa difficulté ou sa longueur.

1 Mots et arbres

Soit Σ un alphabet et $L \subseteq \Sigma^*$ un langage. La *racine* de L est le langage $\sqrt{L} = \{v \in \Sigma^* \mid vv \in L\}$.

[1] **a)** Montrer que si L est un langage rationnel alors $\sqrt{L}$ est aussi rationnel.

[3] **b)** Montrer que $L = \{a^n b^p a^{2p} b^q a^n \mid n, p, q \geq 1\}$ est un langage linéaire.
Montrer que $\sqrt{L}$ n'est pas algébrique.

[4] **c)** Soit Σ un alphabet et $L \subseteq T_p(\Sigma)$ un langage d'arbres. La *racine* de L est le langage

$$\sqrt{L} = \{t \in T_p(\Sigma) \mid \exists s \in T_{p,\square}(\Sigma), \exists a \in \Sigma \text{ tels que } t = s \cdot a \text{ et } s \cdot s \cdot a \in L\} .$$

Montrer que si L est reconnaissable alors $\sqrt{L}$ est aussi reconnaissable.

[3] **d)** Soit $w = a_1 \cdots a_n \in \Sigma^*$ un mot avec $a_k \in \Sigma$ pour $1 \leq k \leq n$. On note $w[i,j] = a_i \cdots a_j$ le facteur de w défini par $1 \leq i \leq j \leq n$. Dans la suite, $u \in \Sigma^*$ et $L \subseteq \Sigma^*$ est un langage *fini*.
Définir une formule du premier ordre $\varphi_u(x,y) \in \mathrm{FO}(\Sigma,<)$ ayant deux variables libres x et y, telle que pour tout $w \in \Sigma^*$ et $1 \leq i \leq j \leq |w|$ on a

$$w[i,j] = u \qquad \Longleftrightarrow \qquad w, x \mapsto i, y \mapsto j \models \varphi_u(x,y) .$$

Définir une formule du premier ordre $\varphi_L(x,y) \in \mathrm{FO}(\Sigma,<)$ ayant deux variables libres x et y, telle que pour tout $w \in \Sigma^*$ et $1 \leq i \leq j \leq |w|$ on a

$$w[i,j] \in L \qquad \Longleftrightarrow \qquad w, x \mapsto i, y \mapsto j \models \varphi_L(x,y) .$$

Définir une formule close $\varphi_{L^*} \in \mathrm{MSO}(\Sigma,<)$ telle que pour tout $w \in \Sigma^*$ on a

$$w \in L^* \qquad \Longleftrightarrow \qquad w \models \varphi_{L^*} .$$

2 Langages linéaires et automates à un pic

Un *automate à un pic* est un automate à pile tel que dans tout calcul valide, la taille de la pile n'augmente plus une fois qu'elle a diminué. La taille de la pile peut donc augmenter (au sens large) pendant une première partie du calcul, puis elle ne fait que diminuer (au sens large).

Un *langage est à un pic* s'il peut être accepté *par pile vide* par un automate à un pic.

[3] **a)** Montrer que le langage $L = \{a^n b^n \mid n > 0\} \cup \{a^n b^{2n} \mid n > 0\}$ est un langage linéaire.
Montrer que L est un langage à un pic.
Montrer que le langage $K = \{ba^{i_1}ba^{i_2}b \cdots ba^{i_n}b \mid n > 0 \text{ et } \exists j, i_j \neq j\}$ est un langage à un pic.

[3] **b)** Montrer que tout langage linéaire est un langage à un pic.

[3] **c)** Soit $G = (\Sigma, V_1 \cup V_2, P_1 \cup P_2)$ une grammaire vérifiant :

$$\begin{aligned} & V_1 \cap V_2 = \emptyset \\ & P_1 \subseteq V_1 \times (V_1\Sigma^* \cup \Sigma^*) \\ & P_2 \subseteq V_2 \times (\Sigma^* V_2 V_1^* \cup \Sigma^*) \end{aligned}$$

Montrer que pour tout $x \in V_1 \cup V_2$, $\mathcal{L}_G(x)$ est un langage linéaire.

Soit $\mathcal{A} = (Q, \Sigma, Z, T, (q_0, z_0))$ un automate à pile arbitraire.
Pour $p, q \in Q$ et $z \in Z$ on définit les langages

$$\begin{aligned} K_{zp,q} &= \{w \in \Sigma^* \mid \exists\ zp \xrightarrow{w} q \text{ calcul dans } \mathcal{A} \text{ avec une pile de taille toujours } \leq 1\}, \\ L_{zp,q} &= \{w \in \Sigma^* \mid \exists\ zp \xrightarrow{w} q \text{ calcul à un pic dans } \mathcal{A}\}. \end{aligned}$$

[2] **d)** Montrer que le langage $K_{zp,q}$ peut être engendré par une grammaire linéaire *droite*.

[3] **e)** Construire une grammaire vérifiant les conditions de la question **(c)** et qui engendre les langages $L_{zp,q}$ avec les variables de V_2 et les langages $K_{zp,q}$ avec les variables de V_1.
Montrer que tout langage à un pic est un langage linéaire.