Langages Formels

Licence Informatique – ENS Cachan

Partiel du 26 mars 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

Soient A et B deux alphabets et $M = A^* \times B^*$. On définit un produit binaire sur M par $(u_1, u_2) \cdot (v_1, v_2) = (u_1v_1, u_2v_2)$ pour $(u_1, u_2), (v_1, v_2) \in M$.

[1] **a)** Montrer que $(M, \cdot, 1_M)$ est un monoïde avec un élément neutre 1_M que l'on précisera. Montrer que M est engendré par $X = (A \times \{\varepsilon\}) \cup (\{\varepsilon\} \times B)$.

Un M-automate est un quadruplet $\mathcal{A} = (Q, T, I, F)$ avec Q l'ensemble fini d'états, $I, F \subseteq Q$, et $T \subseteq Q \times X \times Q$ l'ensemble des transitions. Un couple $(u, v) \in M$ est accepté par $\mathcal{A}$ s'il existe un calcul $q_0 \xrightarrow{x_1} q_1 \xrightarrow{x_2} q_2 \cdots \xrightarrow{x_n} q_n$ avec $q_0 \in I$, $q_n \in F$ et $(u, v) = x_1 \cdot x_2 \cdots x_n$. On note $\mathcal{R}(\mathcal{A}) \subseteq M$ l'ensemble des couples acceptés par $\mathcal{A}$.

Une relation $R \subseteq M$ est rationnelle s'il existe un M-automate $\mathcal{A}$ tel que $R = \mathcal{R}(\mathcal{A})$. On note $\mathsf{Rat}(M)$ l'ensemble des relations rationnelles sur M.

[3] **b)** On suppose $a, c \in A$ et $b, d \in B$.
Montrer que la relation $R_1 = \{(a^n, b^n) \mid n \geq 0\}$ est rationnelle.
Montrer que la relation $R_2 = \{(a^n c^m, b^m d^n) \mid n, m \geq 0\}$ n'est pas rationnelle.

Le domaine d'une relation $R \subseteq M$ est $\mathrm{dom}(R) = \{u \in A^* \mid \exists v \in B^*,\ (u, v) \in R\} \subseteq A^*$.

[2] **c)** Montrer que le domaine d'une relation rationnelle est un langage rationnel : si $R \in \mathsf{Rat}(M)$ alors $\mathrm{dom}(R) \in \mathsf{Rat}(A^*)$.

L'image d'un langage $L \subseteq A^*$ par une relation $R \subseteq M$ est $R(L) = \{v \in B^* \mid \exists u \in L,\ (u, v) \in R\}$.

[3] **d)** Montrer que l'image d'un langage rationnel par une relation rationnelle est un langage rationnel : si $R \in \mathsf{Rat}(M)$ et $L \in \mathsf{Rat}(A^*)$ alors $R(L) \in \mathsf{Rat}(B^*)$.

[5] **e)** Soient C un nouvel alphabet, $\varphi \colon C^* \to A^*$ et $\psi \colon C^* \to B^*$ deux morphismes. Pour $K \subseteq C^*$ on définit $(\varphi, \psi)(K) = \{(\varphi(w), \psi(w)) \mid w \in K\} \subseteq M$.
Montrer que si $K \in \mathsf{Rat}(C^*)$ alors $(\varphi, \psi)(K) \in \mathsf{Rat}(M)$.
Soit $R \in \mathsf{Rat}(M)$. Montrer qu'il existe un alphabet C, deux morphismes $\varphi \colon C^* \to A^*$ et $\psi \colon C^* \to B^*$, et un langage $K \in \mathsf{Rat}(C^*)$ tels que $R = (\varphi, \psi)(K)$.

[5] **f)** Soit $R \subseteq M$ une relation. On définit le langage $\theta(R) = \{u\tilde{v} \mid (u,v) \in R\} \subseteq A^*B^*$ (où $\tilde{v}$ dénote le miroir du mot v).
Soit $R \in \mathsf{Rat}(M)$. Montrer que $\theta(R)$ peut être engendré par une grammaire linaire.
On suppose que $A = B$. Soit $L \subseteq A^*$ un langage linéaire. Montrer qu'il existe une relation rationnelle $R \subseteq M$ telle que $L = \theta(R)$.
On suppose que $A \cap B = \emptyset$. Montrer qu'il existe un langage linéaire $L \subseteq A^*B^*$ tel que pour toute relation rationnelle $R \subseteq M$, on a $L \neq \theta(R)$.

2 Arbres

Soit Σ un alphabet, $\# \notin \Sigma$ et $\Sigma_\# = \Sigma \cup \{\#\}$.

Pour un arbre t' sur l'alphabet $\Sigma_\#$ on définit les propriétés suivantes :

(P-1) les nœuds internes de t' sont étiquetés $\#$.

(P-2) t' n'a pas de nœud interne unaire.

(P-3) tous les nœuds internes de t' sont binaires.

Un langage d'arbres L satisfait une propriété si tous les arbres de L satisfont cette propriété.

Pour un langage d'arbres L, on note $\mathsf{Fr}(L)$ l'ensemble des frontières des arbres de L.

[6] **a)** Soit $L \subseteq T_p(\Sigma)$ un langage d'arbres reconnaissable.
Montrer qu'il existe un langage d'arbres $L' \subseteq T_p(\Sigma_\#)$ reconnaissable tel que $\mathsf{Fr}(L') = \mathsf{Fr}(L)$ et vérifiant la propriété (P-1).
Montrer qu'il existe un langage d'arbres $L' \subseteq T_p(\Sigma_\#)$ reconnaissable tel que $\mathsf{Fr}(L') = \mathsf{Fr}(L)$ et vérifiant les propriétés (P-1) et (P-2).
Montrer qu'il existe un langage d'arbres $L' \subseteq T_2(\Sigma_\#)$ reconnaissable tel que $\mathsf{Fr}(L') = \mathsf{Fr}(L)$ et vérifiant les propriétés (P-1) et (P-3).