UNIVERSITÉ DE STRASBOURG

Département d'Informatique
LICENCE 3ème année
Examen de Théorie des langages
Mardi 4 janvier 2011

Les notes de cours et des travaux dirigés sont autorisées.

Dans toute la suite on considérera l'alphabet $\mathcal{A} = \{0,1\}$ et tous les langages considérés sont sur cet alphabet.

Notations : Soit $n \in \mathbb{N}$.

- $\mathcal{A}^n = \{m \in \mathcal{A}^\star \ : \ |m| = n\}$ est l'ensemble des mots de taille n sur l'alphbet $\mathcal{A}$ (i.e. $\mathcal{A}^\star = \bigcup_{n\geq 0} \mathcal{A}^n$).
- $\mathcal{A}^{\leq n} = \{m \in \mathcal{A}^\star \ : \ |m| \leq n\} = \bigcup_{0\leq i\leq n} \mathcal{A}^i$.
- Soit $m = a_1a_2...a_n$ un mot sur l'aphabet $\mathcal{A}$ (i.e. $a_1, a_2, ..., a_n \in \mathcal{A}$). $m^R = a_na_{n-1}...a_2a_1$.

I. Vrai/Faux

Dire pour chacune des affirmations suivantes si elle est vraie ou fausse en justifiant vos réponses :

1. Tout langage fini sur l'alphabet $\mathcal{A}$ est un langage régulier.
2. $card(\mathcal{A}^{\leq n}) = 2^{n+1} - 1$.
3. Si L est un langage régulier, alors $L^R = \{m^R \ : \ m \in L\}$ est un langage régulier.
4. Si L est un langage régulier, alors le langage $L_{pref} = \{m \ : \ mm' \in L \text{ et } m, m' \in \mathcal{A}^\star\}$ des préfixes de L est un langage régulier.
5. Il existe un entier postif n tel que $\mathcal{A}^\star \setminus \mathcal{A}^{\leq n}$ n'est pas un langage algébrique.
6. Un langage sur $\mathcal{A}$ est soit un langage régulier, soit un langage algébrique.

II. Langages réguliers ... algébriques

Définitions : Soit $m \in \mathcal{A}^\star$.

- Un morphise sur $\mathcal{A}^\star$ est une application $\mu : \mathcal{A}^\star \mapsto \mathcal{A}^\star$ telle que si $m = a_1a_2, ..., a_n$ avec $a_1, a_2, ..., a_n \in \mathcal{A}$, alors $\mu(m) = \mu(a_1)\mu(a_2)...\mu(a_n)$. Autrement dit, la fonction μ est complétement caractérisée par la donnée de ces valeurs sur les lettres de l'alphbet $\mathcal{A}$. Par exemple, si $\mu(0) = 0$ et $\mu(1) = 01$, alors $\mu(10010) = 0100010$.
- Le mot m est dit avec carré, s'il existe $u, v, w \in \mathcal{A}^\star$, tels que $v \neq \varepsilon$ et $m = uvvw$, et il est dit sans carré si une telle décomposition de m n'exsite pas.

- $L_2(\mathcal{A})$ (respectivement $L'_2(\mathcal{A}) = A^\star \setminus L_2(\mathcal{A})$) est l'ensemble des mots avec carré (respectivement sans carré) sur l'aphabet $\mathcal{A}$. Par exemple, le mot $m = 0\underline{1010}1$ est avec carré (i.e. $m \in L_2(\mathcal{A})$), par contre le mot $m' = 010$ est sans carré (i.e. $m' \in L'_2(\mathcal{A})$).

- Le mot m est dit avec cube, s'il existe $u, v, w \in \mathcal{A}^\star$, tels que $v \neq \varepsilon$ et $m = uvvvw$ et il est dit sans cube si une telle décomposition de m n'exsite pas.

- $L_3(\mathcal{A})$ (respectivement $L'_3(\mathcal{A}) = A^\star \setminus L_3(\mathcal{A})$) est l'ensemble des mots avec cube (respectivement sans cube) sur l'aphabet $\mathcal{A}$. Par exemple, le mot $m = 1\underline{101010}0$ est avec cube (i.e. $m \in L_3(\mathcal{A})$), par contre le mot $m' = 0110100$ est sans cube (i.e. $m' \in L'_3(\mathcal{A})$).

1. Caractériser tous les mots sans carré sur l'alphabet $\mathcal{A}$ c'est-à-dire donner explicitement tous les mots du langage $L'_2(\mathcal{A})$.

2. Déduire de 1. que le langage $L_2(\mathcal{A})$ des mots avec carré sur l'alphabet $\mathcal{A}$ est un langage régulier et donner un automate déterministe reconnaissant le langage $L_2(\mathcal{A})$.

3. Soit le morphisme θ sur $\mathcal{A}^\star$ défini par $\theta(0) = 01$ et $\theta(1) = 10$ et considérons la suite des mots $(m_i)_{i\geq 0}$ définie par : $m_0 = 0$ et $m_{i+1} = \theta(m_i)$ pour $i \geq 0$. Calculer m_i pour $1 \leq i \leq 7$.

4. Montrer que pour tout $i \geq 0$, m_i est un préfixe de m_{i+1}.

5. Montrer que pour tout $i \geq 0$, m_i est sans cube. En déduire que le langage $L'_3(\mathcal{A})$ des mots sans cube sur l'alphabet $\mathcal{A}$ est infini.

6. Montrer que le langage $L'_3(\mathcal{A})$ n'est pas un langage algébrique.

7. Le langage $L_3(\mathcal{A})$ est-il un langage régulier ? Justifier votre réponse.

8. Le langage $L_3(\mathcal{A})$ est-il un langage algébrique ? Justifier votre réponse.

III. Langages algébriques

1. Dire pour chacun des langages suivant s'il est un langage algébrique ou pas en justifiant vos réponses :

   (a) $\{0^i 1^j \ : \ i, j \in \mathbb{N} \text{ et } i \neq j\}$.

   (b) $\mathcal{A} \setminus \{0^n 1^n \ : \ n \in \mathbb{N}\}$.

   (c) $\{mmm \ : \ m \in \mathcal{A}^\star\}$.

   (d) $\{m \in \mathcal{A}^* \ : \ m = m^R\}$.

   (e) $\{0^{n^2} \ : \ n \in \mathbb{N}\}$.

   (f) $\{0^p 1^n \ : \ p, n \in \mathbb{N},\ p \text{ premier et } p > n\}$.

2. Décrire un algorithme qui étant donné une grammaire algébrique G décide si le langage $L(G)$ engendré par G est un langage fini ou infini.
   Indication : Utiliser le lemme de la double étoile.

3. Quelle est la complexité de votre algorithme ?