Université de Strasbourg **UFR de Mathématiques et Informatique**
L3S5 d'Informatique **Session de mai 2011**

Théorie des Langages

Contrôle Terminal

Durée : 3 heures
Responsable : Prof. Christian RONSE
Tous documents en papier autorisés mais non partagés
Calculettes inutiles
Téléphones et appareils électroniques éteints et rangés dans un sac fermé

Justifiez soigneusement vos réponses

(1) Langages algébriques ou non.
Considérons l'alphabet $\Sigma = \{a, b, c, d\}$. Parmi les langages suivants, déterminer ceux qui sont algébriques (hors-contexte). Dans l'affirmative, décrire la grammaire qui engendre le langage et donner un exemple de dérivation ; dans la négative, expliquer pourquoi (en utilisant une propriété vue en cours).

(i) $\{a^n b^{(n^2)} \mid n \in \mathbb{N}\}$.
(ii) $\{a^m b^n c^p d^{m+n+p} \mid m, n, p \in \mathbb{N}\}$.
(iii) $\{a^m b^n \mid m, n \in \mathbb{N},\ m \neq n\}$.
(iv) $\{a^m b^n c^p \mid m, n, p \in \mathbb{N},\ n > m + p\}$.
(v) $\{a^{(n^3)} \mid n \in \mathbb{N}\}$.

(2) Un langage algébrique.
Considérons l'alphabet $\Sigma = \{a, b\}$. Le langage $L = \{a^m b^n \mid m, n \in \mathbb{N},\ m \leq n \leq 3m\}$.

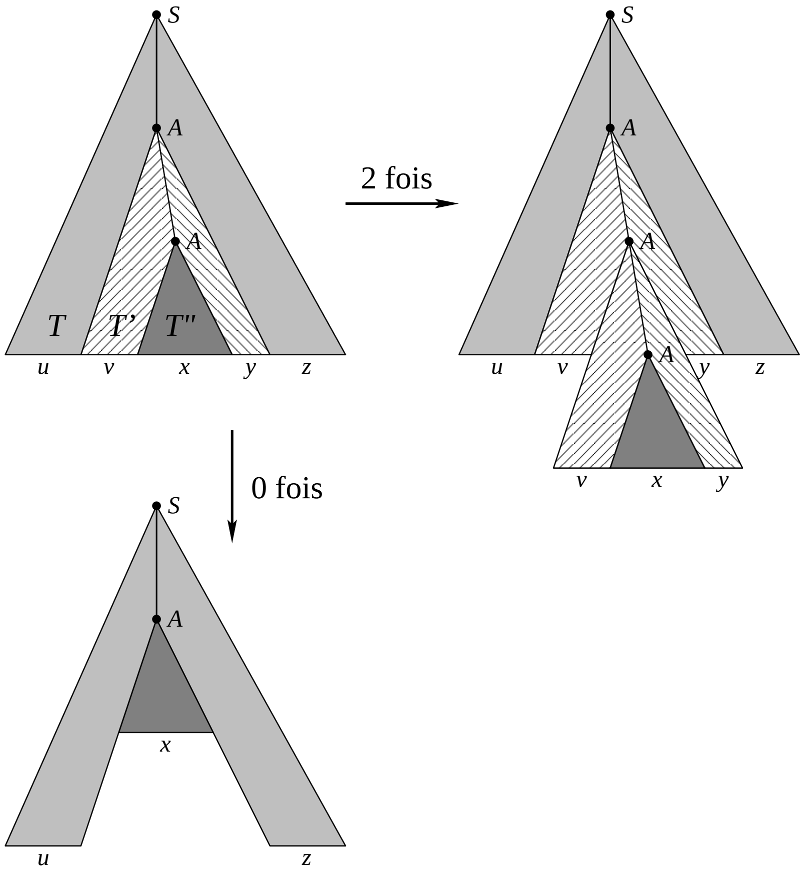
(i) Donner une grammaire algébrique qui engendre L et donner une dérivation du mot $a^3 b^6$.
(ii) Décrire un automate à pile acceptant L.

(3) Amélioration du lemme de la double étoile.
Soit $G = (V, \Sigma, R, S)$ un grammaire algébrique. Posons $\Phi(G)$ comme le maximum des longueurs $|w|$ des mots w à droite dans les règles $A \to w$ de R. Soit $|V - \Sigma|$ le nombre de symboles non terminaux. Dans la preuve du lemme de la double étoile, étant donné un mot $w \in L(G)$ de longeur $|w| > \Phi(G)^{|V-\Sigma|}$, on prend l'arborescence T de dérivation de w la plus petite possible ; ayant plus que $\Phi(G)^{|V-\Sigma|}$ feuilles, celle-ci est de hauteur supérieure à $|V - \Sigma|$, donc elle a un chemin C de la racine S vers une feuille, de longeur au moins $|V - \Sigma| + 1$; la feuille est étiquetée dans Σ, les autres sommets de C sont étiquetés dans $V - \Sigma$. Donc sur ce chemin C deux sommets ont la même étiquette $A \in V - \Sigma$. On obtient alors la décomposition de l'arborescence T illustrée dans la page suivante, où T' et T'' sont les sous-arborescences dont les racines sont ces deux sommets étiquetés A, et on a la décomposition $w = uvxyz$ avec $vy \neq \varepsilon$, telle que pour tout $n \in \mathbb{N}$, $uv^n xy^n z \in L(G)$; nous illustrons la construction des arborescences de dérivation de uxz ($n = 0$) et de uv^2xy^2z ($n = 2$).

On va choisir sur ce chemin C de la racine vers une feuille une répétition d'une même étiquette dans $V - \Sigma$, telle que la première occurrence de cette étiquette soit la plus basse possible. On écrit $s_m, \ldots, s_0$ pour la suite des sommets de C, où $c_m = S$ est la racine et c_0 est la feuille. Soit h le plus petit $s \in \{1, \ldots, m\}$ tel qu'il existe $t \in \{1, \ldots, s-1\}$ pour lequel c_s et c_t ont la même étiquette dans $V - \Sigma$ (donc c_t est sous c_s, il est dans la sous-arborescence de racine c_s).

(i) Donner une borne supérieure pour h.

(ii) En déduire une borne supérieure pour $|vxy|$.



(4) Nombre pair ou impair d'occurrences.

Considérons l'alphabet $\Sigma = \{a_1, \ldots, a_n\}$. Pour un mot $m \in \Sigma^*$ et pour $i = 1, \ldots, n$, soit $p_i(m)$ la parité du nombre d'occurrences de la lettre a_i dans m ; donc $p_i(m) = 0$ si la lettre a_i apparaît un nombre pair de fois dans le mot m, et $p_i(m) = 1$ si elle y apparaît un nombre impair de fois. Le vecteur $\big(p_1(m), \ldots, p_n(m)\big)$ sera appelé *vecteur de parité de m*.

(i) Pour $m_1, m_2 \in \Sigma^*$, décrire comment on obtient le vecteur $\big(p_1(m_1m_2), \ldots, p_n(m_1m_2)\big)$ à partir des deux vecteurs $\big(p_1(m_1), \ldots, p_n(m_1)\big)$ et $\big(p_1(m_2), \ldots, p_n(m_2)\big)$.

Soit L le langage formé de tous les mots $m \in \Sigma^*$ tels que chacune des lettres de Σ apparaît un nombre pair de fois dans m : $p_1(m) = \ldots = p_n(m) = 0$.

(ii) Donner la relation entre le vecteur de parité et la congruence $\approx_L$ selon le langage L (rappel : $m_1 \approx_L m_2$ ssi pour tout $w \in \Sigma^*$, $m_1 w \in L \Leftrightarrow m_2 w \in L$).

(iii) Décrire l'automate standard (minimal) acceptant L.

(5) Commutation de mots.

Soit Σ un alphabet, et soient $u, v \in \Sigma^*$ deux mots. Montrer que ces deux mots commutent, $u \cdot v = v \cdot u$, si et seulement si u et v sont puissances d'un même mot : $\exists\, w \in \Sigma^*$, $\exists\, m, n \in \mathbb{N}$, $u = w^m$ et $v = w^n$. (Indication : la preuve se fait par récurrence sur la longueur de uv, et vous pouvez utiliser le Lemme de Lévy).

(6) Union et intersection infinie de langages rationnels.

On sait que l'union et l'intersection de deux langages rationnels est un langage rationnel ; donc par récurrence, cela reste vrai pour l'union et l'intersection d'un nombre fini quelconque de langages rationnels. Que peut-on dire de l'union et de l'intersection d'une *infinité dénombrable* de langages rationnels ? Quel type de langage peut-on obtenir de cette manière ?