

Module Langages Formels

TD 8 : Langages Algébriques et Automates à Pile

Exercice 1 Lemme de l'étoile

Lemme (de l'étoile) :

Si L est un langage algébrique, alors il existe $N \in \mathbb{N}$ tel que $\forall u \in L, |u| \geq N \Rightarrow \exists x, v, y, w, z \in \Sigma^*, u = xvywz$ et

- $vw \neq \varepsilon$;
- $|vyw| \leq N$;
- $\forall i \geq 0, xv^i y w^i z \in L$

1.1. Montrer que les langages suivants ne sont pas algébriques.

1. $L_1 = \{a^i b^j c^k, i < j < k\}$;
2. $L_2 = \{a^n b^m c^m, n \leq m \leq 2n\}$;
3. $L_3 = \{a^{2^n}, n \geq 0\}$;
4. $L_4 = \{a^{n^2}, n \geq 0\}$;

Lemme (Ogden) :

Soit L un langage algébrique. Il existe un entier N tel que pour tout mot $z \in L$ dans lequel on marque au moins N positions distinctes, il est possible de décomposer z sous la forme $z = uxv y w$ avec

- x ou y contient au moins une position marquée,
- $xv y$ contient au plus N positions marquées,
- pour tout $i \geq 0, ux^i v y^i w \in L$.

1.2. On s'intéresse au langage $L_5 = \{a^i b^j c^k d^l \mid i = 0 \text{ ou } j = k = l\}$.

1.2. 1. Montrer que pour tout $N \in \mathbb{N}^+$ et tout mot $z \in L_5$ avec $|z| \geq N$, il existe une décomposition $z = uxv y w$ telle que

- $|xy| \geq 1$
- $|xv y| \leq N$
- pour tout $i \geq 0, ux^i v y^i w \in L_5$.

1.2. 2. Montrer que L_5 n'est pas algébrique.

Exercice 2

Pour tout langage L sur Σ , on définit

$$\frac{1}{2}L = \{x \in \Sigma^*, \exists y \in \Sigma^* \text{ tel que } |x| = |y| \text{ et } xy \in L\}$$

2.1. Montrer que le langage $L_6 = \{a^n b^n c^m d^{3m}, n, m \geq 1\}$ est algébrique.

2.2. Calculer $\frac{1}{2}L_6$.

2.3. Montrer que $\frac{1}{2}L_6$ n'est pas algébrique.

Exercice 3 Centre d'un langage algébrique

On note $<$ la relation « est facteur gauche » (ou préfixe) sur Σ^* et on note $FG(u) = \{v \in \Sigma^* \mid v < u\}$ et $FG(L) = \cup_{u \in L} FG(u)$.

Soit L un langage de Σ^* . On appelle centre de L , et l'on note $c(L)$, l'ensemble des mots infiniment complétables dans L , c'est-à-dire :

$$c(L) = \{u \in \Sigma^* \mid \forall n \in \mathbb{N}, \exists v_n \in L \text{ tel que } |v_n| \geq n \text{ et } u < v_n\}.$$

3.1. Calculer $c(a^*)$, $c(a^*b)$, $c(\{a^n b^n \mid n > 0\})$. Montrer que pour tous langages L , L_1 et L_2 , $c(L) \subseteq FG(L)$, $c(L_1 \cup L_2) = c(L_1) \cup c(L_2)$ et que $FG(c(L)) = c(L)$.

3.2. Montrer que si $u \in c(L)$, alors il existe une lettre $a \in \Sigma$ telle que $ua \in c(L)$. En déduire que L est infini si et seulement si $c(L)$ est infini, et qu'alors $c(c(L)) = c(L)$.

Soit L un langage algébrique sur Σ , engendré par une grammaire algébrique $G = (\Sigma, V, P, S)$ supposée réduite. On note respectivement $L_G(T)$ et $\hat{L}_G(T)$ les langages sur respectivement Σ^* et $(\Sigma \cup V)^*$ des mots générés par la grammaire G à partir du non-terminal T .

3.3. Montrer que l'on peut calculer l'ensemble $U \subseteq V$ des non-terminaux qui engendrent un langage infini. Montrer que si le mot $uvTw \in \hat{L}_G(T)$ avec $u \in \Sigma^*$ et $T \in U$, alors $u \in c(L)$.

Soit $\bar{U} = \{\bar{T} \mid T \in U\}$ un nouvel ensemble de non-terminaux en bijection avec U . On définit successivement les ensemble de règles :

$$P_1 = \{\bar{T} \rightarrow m_1 \bar{R} m_2 \mid R, T \in U, m_1 \in (\Sigma \cup V)^*, m_2 \in (\Sigma \cup V \setminus U)^*, T \rightarrow m_1 R m_2 \in P\},$$

$$\bar{P} = \{\bar{T} \rightarrow m_1 \bar{R} \mid \exists m_2 : \bar{T} \rightarrow m_1 \bar{R} m_2 \in P_1\} \cup \{\bar{T} \rightarrow \varepsilon \mid T \in U\}.$$

On pose alors $\bar{G} = (\Sigma, V \cup \bar{U}, P \cup \bar{P}, \bar{S})$.

3.4. Montrer que $L_{\bar{G}}(\bar{S}) \subseteq c(L)$.

3.5. Soit $\Sigma = \{a, b, c, d\}$, $V = \{S, R, T, W, Y, Z\}$ et

$$P : \begin{aligned} S &\rightarrow aRcTYW \\ R &\rightarrow b \\ T &\rightarrow aR \\ W &\rightarrow cTb \\ Y &\rightarrow aS + dZ \\ Z &\rightarrow cdZ + aR. \end{aligned}$$

3.5. 1. Calculer U, \bar{G} . Donner un arbre de dérivation de S en $abcabaabcabdcddcabbcabb$ et indiquer sur cet arbre comment on obtient dans \bar{G} le mot $abcabaabcabdcddcd$.

3.5. 2. Donner un exemple de mot de $c(L) \setminus L_{\bar{G}}(\bar{S})$.

3.6. Montrer que si G est sous forme normale presque-Greibach, c'est-à-dire, si tous ses membres droits commencent par une lettre terminale, $c(L) = FG(L_G(S))$.

3.7. Montrer que si L est un langage algébrique, alors $c(L)$ est également un langage algébrique, et qu'à partir d'une grammaire algébrique engendrant L , on peut effectivement construire une grammaire algébrique engendrant $c(L)$.

3.8. Montrer qu'il existe un théorème analogue pour les langages linéaires (qui ont au plus un non-terminal dans chaque membre droit de leurs règles de production).