TD5 : Grammaires algébriques

On note $G = (\Sigma, V, P, S)$, où Σ est un alphabet, V l'ensemble des variables, $S \in V$ est l'axiome, et $P \subset (\Sigma \cup V)^\star \times (\Sigma \cup V)^\star$ désigne un ensemble fini de règles.

Exercice 1 (Langages de Dyck)**.** Soit $\Sigma = \{a_1, \ldots, a_n\} \cup \{\bar{a}_1, \ldots, \bar{a}_n\}$ l'alphabet formé de n paires de parenthèses. Soit $G_n = (\Sigma_n, V, P_n, S)$ la grammaire définie par $S \to a_1 S \bar{a}_1 S + \ldots + a_n S \bar{a}_n S + \epsilon$. Le langage $D_n^* = \mathcal{L}_{G_n}(S)$ est appelé langage de Dyck sur n paires de parenthèses.

1. Montrer que $D_1^* = \{w \in \Sigma_1^* \mid |w|_{a_1} = |w|_{\bar{a}_1} \text{ et } |v|_{a_1} \geq |v|_{\bar{a}_1} \text{ pour tous } v \leq w\}$.
2. On considère le système de réécriture (de type 0) $R_n = (\Sigma_n, P'_n)$ dont les règles sont $P'_n = \{(a_i \bar{a}_i, \epsilon) \mid 1 \leq i \leq n\}$.
   Montrer que $D_n^* = \{w \in \Sigma_n^* \mid w \xrightarrow{*} \epsilon \text{ dans } R_n\}$.
3. Soit Γ un alphabet disjoint de Σ_n, $\Sigma = \Sigma_n \cup \Gamma$ et $L \subset \Sigma^*$ un langage.
   On définit la clôture $\mathsf{clot}(L) = \{v \in \Sigma^* \mid \exists w \in L, w \xrightarrow{*} v \text{ dans } R_n\}$.
   Montrer que si L est reconnaissable, alors $\mathsf{clot}(L)$ l'est aussi.
   On définit la réduction $\mathsf{red}(L) = \{v \in \mathsf{clot}(L) \mid v \not\rightarrow \text{ dans } R_n\}$.
   Montrer que si L est reconnaissable, alors $\mathsf{red}(L)$ l'est aussi.

Exercice 2.
1. Quels sont les langages engendrés par les grammaires $S \to \epsilon \mid aaaS$, $S \to ab \mid aSb$
2. Donner des grammaires algébriques qui engendrent les langages : $L_1 = \{a^n b^{2n} \mid n \leq 0\}$, $L_2 = ab^*$, $L_3 = (ab^*)^*$, $L_4 = \{a^n b^p \mid n \leq p\}$, $L_5 = \{a^n b^p c^q \mid n = p \text{ ou } p = q\}$

Exercice 3. Donner la forme normale de Chomsky de la grammaire $(\{S, A, B\}, \{a, b\}, P, S)$ définie par les productions :

$$\begin{aligned} S &\to bA|aB \\ A &\to bAA|aS|a \\ B &\to aBB|bS|b \end{aligned}$$

Exercice 4. On considère le langage $L = L_G(S)$ où la grammaire G est définie par $S \to aSSb + c$. Montrer que tout langage rationnel inclus dans L est fini.

Exercice 5. Soit K un langage rationnel et L un langage algébrique. Montrer que

1. l'inclusion $K \subset L$ est indécidable ;

2. l'inclusion $L \subset K$ est décidable.

Exercice 6. Montrer que tout langage algébrique sans ε est engendré par une grammaire algébrique dont toutes les productions sont de la forme $A \to a$, $A \to aB$ et $A \to aBC$.

Exercice 7. Une grammaire *opérateur* est une grammaire algébrique sans ε-productions et telle qu'il n'existe pas deux symboles consécutifs dans les parties droites des productions qui soient des variables. Montrer que tout langage algébrique sans ε a une grammaire opérateur.