TD 3. Grammaires algébriques

Exercice 1. Lemme fondamental

Démontrer le résultat suivant qui permet de décomposer des dérivations au sein d'une grammaire algébrique.

Lemme 1. *Soit* $G = (N, T, P, S)$ *une grammaire algébrique et* $m \geq 1$*. Si* $\alpha_1, \ldots, \alpha_m, \beta \in V^*$ *sont tels que, pour un certain* $n \in \mathbb{N}$*, on ait la dérivation* $\alpha_1 \cdots \alpha_m \Rightarrow^n \beta$ *dans* G*, alors il existe* $n_1, \ldots, n_m$ *dans* $\mathbb{N}$ *et* $\beta_1, \ldots, \beta_m$ *dans* V^* *tels que* $n = n_1 + \cdots + n_m$*,* $\beta = \beta_1 \cdots \beta_m$*, et pour tout* $1 \leq i \leq m$*,* $\alpha_i \Rightarrow^{n_i} \beta_i$ *dans* G*.*

Exercice 2. Langage de Dyck

On considère la grammaire algébrique $G = (N, T, P, S)$ où $N \stackrel{\text{def}}{=} \{S\}$, $T \stackrel{\text{def}}{=} \{(,)\}$ et l'ensemble de productions P est

$$S \to (S)S \mid \varepsilon$$

Pour un mot $w \in T^*$, on définit son poids $p(w) \stackrel{\text{def}}{=} |w|_(- |w|_)$ comme la différence entre son nombre de parenthèses ouvrantes et son nombre de parenthèses fermantes. Un tel mot est *bien parenthésé* si $p(w) = 0$ et pour tout préfixe u de w, $p(u) \geq 0$.

(a) Montrer que tout mot engendré par cette grammaire est bien parenthésé.

(b) Montrer que tout mot bien parenthésé est engendré par cette grammaire. Est-elle ambiguë ?

Exercice 3. Forme quadratique

Montrer que, pour toute grammaire algébrique $G = (N, T, P, S)$, on peut construire une grammaire algébrique équivalente $G' = (N', T, P', S)$ telle que, pour toute production $A \to \alpha$ dans P', $|\alpha| \leq 2$.

Exercice 4. Clôture par intersection avec un langage rationnel

(a) Soit $G = (N, \Sigma, P, S)$ une grammaire algébrique et $A = (Q, \Sigma, \delta, I, F)$ un automate fini. Construire une grammaire algébrique G' telle que $L(G') = L(G) \cap L(A)$.

Indice : considérer des non terminaux $(q, X, q') \in Q \times V \times Q$ *tels que* $L_{G'}(q, X, q') = L_G(X) \cap \{w \in \Sigma^* \mid q' \in \delta^*(q, w)\}$*.*

(b) Modifier la construction pour ne produire que des non terminaux co-accessibles dans G'.