Grammaires algébriques

FOND1 - ENS Rennes

11 Exemples de grammaires algébriques

Exemple 61 (grammaire qui engendre un langage de phrases en français).

$$\begin{array}{ll} \mathit{PHRASE} & \rightarrow \mathit{SUJET\ VERBE\ COMPLÉMENT} \\ \mathit{SUJET} & \rightarrow \mathit{DETERMINANT\ NOM} \\ \mathit{COMPLÉMENT} & \rightarrow \mathit{DETERMINANT\ NOM} \\ \mathit{DETERMINANT} & \rightarrow \mathit{la} \mid \mathit{une} \\ \mathit{NOM} & \rightarrow \mathit{fleur} \mid \mathit{fille} \\ \mathit{VERBE} & \rightarrow \mathit{mange} \\ \mathit{VERBE} & \rightarrow \mathit{dévore} \end{array}$$

Exemple 62 (une grammaire qui engendre le langage $\{a^n b^n \mid n \in \mathbb{N}\}$).

$$S \quad \rightarrow aSb \mid \epsilon$$

Exemple 63 (grammaire qui engendre le langage de Dyck).

$$\begin{array}{ll} S & \rightarrow ST \mid \epsilon \\ T & \rightarrow a_1 S \overline{a_1} \mid \cdots \mid a_n S \overline{a_n} \end{array}$$

Exemple 64 (grammaire qui engendre le langage de Łukasiewicz, i.e. expressions préfixes).

$$S \quad \rightarrow +SS \mid a$$

Exemple 65 (grammaire qui engendre des expressions postfixes).

$$S \quad \rightarrow SS+ \mid a$$

Exemple 66 (grammaire qui engendre des expressions arithmétiques complétement parenthésées).

$$E \quad \rightarrow (E+E) \mid (E \times E) \mid a$$

Exemple 67 (grammaire ambiguë qui engendre des expressions arithmétiques).

$$E \quad \rightarrow E+E \mid E \times E \mid (E) \mid a$$

Exemple 68 (grammaire non ambiguë qui engendre des expressions arithmétiques).

$$\begin{array}{ll} E & \rightarrow E+T \mid T \\ T & \rightarrow T \times F \mid F \\ F & \rightarrow (E) \mid a \end{array}$$

12 Définitions formelles

Définition 69 (grammaire algébrique)**.** *Une* grammaire algébrique *est un 4-uplet* $G = (V, \Sigma, R, S)$ *où*
- V *ensemble fini de variables (ou non-terminaux) ;*
- Σ *ensemble fini de terminaux tel que* $\Sigma \cap V = \emptyset$ *;*
- $R \subseteq V \times (V \cup \Sigma)^*$ *ensemble fini de règles de production ;*
- $S \in V$ *variable initiale (ou axiome).*

Pour $(X, \alpha) \in R$, on note $X \to \alpha$.

Définition 70 (partie gauche, partie droite)**.** *Etant donné une règle* $X \to \alpha$*,* X *s'appelle la partie gauche et* α *s'appelle la partie droite.*

Définition 71 (dérivation)**.** *Soit* $G = (V, \Sigma, R, S)$ *une grammaire algébrique et soient* $\alpha, \beta \in (V \cup \Sigma)^*$*. On dit que* α *se dérive en* β*, noté* $\alpha \to \beta$ *s'il existe* $u, v \in (V \cup \Sigma)^*$ *et* $X \in V$ *tels que :*

$$\alpha = uXv, \quad \beta = u\gamma v, \quad \text{si } X \to \gamma \text{ est dans } R.$$

Définition 72 (dérivation gauche, dérivation droite)**.** *La dérivation est* gauche *(resp. droite) si* $u \in \Sigma^*$ *(resp.* $v \in \Sigma^*$*).*

Notation 73 (fermeture réflexive transitive)**.** *On note* $\to^*$ *pour la fermeture réflexive transitive de* $\to$*.*

Définition 74 (langage engendré par une variable)**.** *Soit* $G = (V, \Sigma, R, S)$ *une grammaire algébrique. Le langage engendré par une variable* X *est défini par :*

$$\hat{L}_X(G) = \{\alpha \in (V \cup \Sigma)^* \mid X \to^* \alpha\}.$$

$$L_X(G) = \{w \in \Sigma^* \mid X \to^* w\}.$$

Représentation graphique : arbre de dérivation

Définition 75 (langage engendré)**.** *Le langage engendré par une grammaire algébrique* G *est défini par :*

$$L(G) = L_G(S).$$

Notation 76. *Si* w *est un mot et* a *une lettre, on note* $|w|_a$ *le nombre d'occurrences de la lettre* a *dans* w*.*

Proposition 77. *Le langage de Łukasievicz est le langage :*

$$K = \left\{w \in \{a, +\}^* \mid |w|_a = |w|_+ + 1 \text{ et pour tout préfixe propre de } u \text{ de } w,\ |u|_a \leq |u|_+\right\}.$$

Définition 78 (langage algébrique)**.** $L \subseteq \Sigma^*$ *est* algébrique *s'il existe une grammaire algébrique* G *tel que* $L = L(G)$*. On note* $Alg(\Sigma)$ *la classe des langage algébrique sur* Σ*.*

13 Ambiguïté

Définition 79 (grammaire ambiguë)**.** *Une grammaire* G *est ambiguë s'il existe un mot* $u \in L(G)$ *ayant au moins 2 suites différentes de dérivations gauches.*

Définition 80 (langage intrinsèquement ambigu)**.** *Un langage algébrique est intrinsèquement ambigu si pour toute grammaire* G*,* $L = L(G)$ *implique que* G *est ambiguë.*

14 Stabilité

14.1 Par opérations rationnelles

Théorème 81. $Alg(\Sigma)$ *est stable par concaténation, union et étoile de Kleene.*

Corollaire 82. $Rat(\Sigma) \subseteq Alg(\Sigma)$*.*

14.2 Lemme de pompage

Lemme 83 (Bar-Hillel, Perles et Shamir)**.** *Soit* L *un langage algébrique. Il existe un entier* K *tel que pour tout mot* $u \in L$ *avec* $|u| \geq K$*, il existe une décomposition* $u = xvywz$ *telle que*
1. $vw \neq \epsilon$ *;*
2. $|vyw| \leq K$ *;*
3. *Pour tout* $i \in \mathbb{N}$*,* $xv^iyw^iz \in L$*.*

Corollaire 84. $\{a^nb^nc^n \mid n \geq 0\}$ *n'est pas algébrique.*

14.3 Non-stable par intersection et complémentaire

Corollaire 85. $Alg(\Sigma)$ *n'est pas close par intersection et par complémentaire.*