

Introduction aux langages formels

Grammaires

Frédéric Guyomarch et Cédric Lhoussaine

24 février 2025



Université
de Lille

Définition : Grammaire

Une **grammaire** est un quadruplet $G = (V, \Sigma, \mathcal{P}, S)$ où

- V est un alphabet de symboles **non-terminaux** (ou **variables**) (vocabulaire) ;
- Σ est un alphabet de symboles **terminaux** ;
- $\mathcal{P} \subseteq (V \cup \Sigma)^+ \times (V \cup \Sigma)^*$ est un ensemble de **règles** ou **productions** de la forme $u \rightarrow v$.
- $S \in V$ est l'**axiome** (à partir duquel les mots sont générés).

Définition : Dérivation de mots

Soit $G = (V, \Sigma, \mathcal{P}, S)$, $u \in (V \cup \Sigma)^+$ et $v \in (V \cup \Sigma)^*$, alors v peut être **dérivé** de u à partir de G , dénoté $u \xRightarrow{G} v$, ssi $\exists x, y \in V^*$ tels que :

- $u = xu'y$
- $v = xv'y$
- $u' \rightarrow v' \in \mathcal{P}$.

Exemple : Dérivation de $1+(2*3)$

$$\begin{aligned} S &\Rightarrow \underline{S}+S \Rightarrow \underline{C}+S \Rightarrow 1+\underline{S} \Rightarrow 1+(\underline{S}) \Rightarrow 1+(\underline{S}*S) \\ &\Rightarrow 1+(\underline{C}*S) \Rightarrow 1+(2*\underline{S}) \Rightarrow 1+(2*\underline{C}) \Rightarrow 1+(2*3) \end{aligned}$$

Définition : Langage engendré

Le **langage engendré** (ou généré) par une grammaire $G = (V, \Sigma, \mathcal{P}, S)$ est

$$\mathcal{L}(G) = \{w \mid S \xRightarrow[G]{*} w\}$$

Définition : Grammaire régulière

Une grammaire est dite **régulière** si toutes ses règles sont de la forme $A \rightarrow wB$ ou $A \rightarrow w$ où $A, B \in V$ et $w \in \Sigma^*$.