Chapitre 9 : Grammaires non contextuelles

January 22, 2025

1 Grammaire non contextuelle

1.1 Vocabulaire

Définition 9.1 - *grammaire au sens général, grammaire de type 0*

Une *grammaire* est défini par un quadruplet (Σ, V, P, S) où :

- Σ est un alphabet fini de *symboles terminaux*, dit aussi *alphabet terminal*
- V est un alphabet fini de *symboles non terminaux* (ou *variables*), dit aussi *alphabet non terminal*
- $P \subset (\Sigma \cup V)^* \times (\Sigma \cup V)^*$ est un ensemble de *règles de production*.
  Une règle de production $(w_1, w_2) \in P$, notée $w_1 \to w_2$ est un couple de mots écrits avec des symboles terminaux et non terminaux.
- $S \in V$ est un symbole non terminal avec un statut particulier de *symbole initial* (ou *axiome*, *variable initiale*)

Une grammaire sans propriété particulière est dite *type 0*.

Remarque 9.2 - *grammaires*

On note usuellement par des majuscules les symboles non terminaux, et en minuscule les terminaux.

Exemple 9.3 - *de grammaire de type 0*

Pour $\Sigma = \{a\}$ et $V = \{S, D, F, X, Y, Z\}$ et $P = \{\mathtt{S} \to \mathtt{DXaD}, \mathtt{Xa} \to \mathtt{aaX}, \mathtt{XF} \to \mathtt{YF}, \mathtt{aY} \to \mathtt{Ya}, \mathtt{DY} \to \mathtt{DX}, \mathtt{XZ} \to \mathtt{Z}, \mathtt{aZ} \to \mathtt{Za}, \mathtt{DZ} \to \epsilon\}$, $G = (\Sigma, V, P, S)$ est une grammaire de type 0.

Définition 9.4 - *dérivation immédiate*

Soit $G = (\Sigma, V, P, S)$ une grammaire. On dit que $\alpha \in (\Sigma \cup V)^*$ se dérive immédiatement en $\beta \in (\Sigma \cup V)^*$ lorsqu'il existe $(\alpha_2, \beta_2) \in P$ tel que :

$$\exists(\alpha_1, \alpha_3) \in \Big((\Sigma \cup V)^*\Big)^2, \begin{cases} \alpha = \alpha_1\alpha_2\alpha_3 \\ \beta = \alpha_1\beta_2\alpha_3 \end{cases}$$

Le cas échéant, on note $\alpha \Rightarrow \beta$. On parle de *dérivation immédiate*. Moralement, le facteur α_2 est remplacé par le facteur β_2.

Définition 9.5 - *clôture reflexive et transitive*

On note $\Rightarrow^*$ la *clôture reflexive et transitive* de la relation $\Rightarrow$ de dérivabilité immédiate.

$\Rightarrow^*$ est définie comme la plus petite relation au sens de l'inclusion tel que :

- $\forall \alpha \in (\Sigma \cup V)^*,\ \alpha \Rightarrow^* \alpha$
- $\forall (\alpha, \beta) \in \Big((\Sigma \cup V)^*\Big)^2, (\alpha \Rightarrow \beta) \implies (\alpha \Rightarrow^* \beta)$
- $\forall (\alpha, \beta, \gamma) \in \Big((\Sigma \cup V)^*\Big)^3,\ (\alpha \Rightarrow^* \beta \text{ et } \beta \Rightarrow^* \gamma) \implies (\alpha \Rightarrow^* \gamma)$

Autrement dit, $\alpha \Rightarrow^* \beta$ lorsqu'il existe $(\alpha = \alpha_0,\ \ldots,\ \alpha_k = \beta)$ une suite de mots dans $(\Sigma \cup V)^*$ telle que :

$$\forall i \in [\![0,\ k-1]\!], \alpha_i \Rightarrow \alpha_{i+1}$$

Exemple 9.6 - *de dérivation*

Dans la grammaire précédemment introduite :
Pour $\Sigma = \{a\}$ et $V = \{S, D, F, X, Y, Z\}$ et $P = \{\mathtt{S} \to \mathtt{DXaD}, \mathtt{Xa} \to \mathtt{aaX}, \mathtt{XF} \to \mathtt{YF}, \mathtt{aY} \to \mathtt{Ya}, \mathtt{DY} \to \mathtt{DX}, \mathtt{XZ} \to \mathtt{Z}, \mathtt{aZ} \to \mathtt{Za}, \mathtt{DZ} \to \epsilon\}$, $G = (\Sigma, V, P, S)$ est une grammaire de type 0.

$$
\begin{aligned}
S &\Rightarrow \mathtt{DXaF} \\
&\Rightarrow \mathtt{DaaXF} \\
&\Rightarrow \mathtt{DaaYF} \\
&\Rightarrow \mathtt{DaYaF} \\
&\Rightarrow \mathtt{DYaaF} \\
&\Rightarrow \mathtt{DXaaF} \\
&\Rightarrow \mathtt{DaaXaF} \\
&\Rightarrow \mathtt{DaaaaXF} \\
&\Rightarrow \mathtt{DaaaaZ} \\
&\Rightarrow \mathtt{DaaaZa} \\
&\Rightarrow \mathtt{DaaZaa} \\
&\Rightarrow \mathtt{DaZaaa} \\
&\Rightarrow \mathtt{DZaaaa} \\
&\Rightarrow \mathtt{aaaa}
\end{aligned}
$$

D'où $S \Rightarrow^* \mathtt{aaaa}$

Définition 9.7 - *langage engendré par une grammaire depuis un mot*

Soit $G = (\Sigma, V, , S)$ une grammaire et $\alpha \in (\Sigma \cup V)^*$.
On définit le *langage engendré par G depuis α* comme l'ensemble des mots de Σ^* que l'on peut obtenir par dérivation depuis α en utilisant les règles de production de G.

$$\mathcal{L}_G = \{u \in \Sigma^*,\, \alpha \Rightarrow^* u\}$$

Le *langage élargi engendré par G depuis α* est :

$$\widehat{\mathcal{L}_G(\alpha)} = \{\beta \in (\Sigma \cup V)^*,\, \alpha \Rightarrow^* \beta\}$$

Le *langage engendré par G* désigne $\mathcal{L}_G(S)$ le langage engendré par G depuis le symbole initial.

Exemple 9.8

Pour la grammaire de l'exemple, on pourrait montrer que $\mathcal{L}_G(S) = \{a^{2^n},\, n \in \mathbb{N}^*\}$ On montre que ce langage n'est pas régulier (absurde + lemme de l'étoile tmtc)

Définition 9.9 - *langage de type 0*

On dit qu'un *langage est de type 0* s'il peut être engendré par une grammaire de type 0.

Théorème 9.10 - *de Chomsky (HP)*

Les langages de type 0 sont exactement les langages récursivement énumérables, c'est-à-dire les langages reconnaissables par une machine de Turing.