NF93 – Grammaires formelles

Chabot Simon

20 avril 2011

1 Définition d'une grammaire et d'un langage

Définition 1.1 (Grammaire formelle). *Une* grammaire formelle G *est un ensemble constitué de quatre objets :*

- *Un ensemble fini* Σ *de symboles terminaux, conventionnellement notés en minuscules. Ce sont les* lettres *du langage.*
- *Un ensemble fini* $\mathcal{V}$ *de symboles non-terminaux, conventionnellement notés en majuscules. Ce sont les* variables[1].
- *Un élément* $S \in \mathcal{V}$*, appelé* axiomes.
- *Un ensemble* P *des* règles de production. *De la forme* $\alpha \to \beta$*, avec* $\alpha \in (\mathcal{V} \cup \Sigma)^* \, \Sigma \, (\mathcal{V} \cup \Sigma)^*, \alpha \neq \varepsilon, \beta \in (\mathcal{V} \cup \Sigma)^*$

On note $G = (\Sigma, \mathcal{V}, S, P)$.

Définition 1.2 (Puissance d'un alphabet). *La* k^e *puissance d'un alphabet* Σ*, noté* Σ^k *est définie par :*

$$\begin{cases} \Sigma^0 & = \{\varepsilon\} \\ \Sigma^k & = \Sigma^{k-1} \times \Sigma \end{cases}$$

Σ^* *est l'ensemble de tous les mots que l'on peut former à partir de* Σ*. On définit* $\Sigma^+ = \Sigma^* \backslash \{\varepsilon\}$.

Définition 1.3 (Langage). *On appelle* langage *défini sur un alphabet* Σ*, un sous ensemble de* Σ^*.

Définition 1.4 (Puissance d'un langage). *La* k^e *puissance d'un langage* L*, noté* L^k *est définie par :*

$$\begin{cases} L^0 & = \{\varepsilon\} \\ L^k & = L^{k-1} \times L \end{cases}$$

1. On a $\Sigma \cap \mathcal{V} = \emptyset$

Exemple 1.1. *Si* $L = \{ab, cd\}$*, alors* $L^0 = \{\varepsilon\}$*, et* $L^1 = L$*,* $L^2 = \{abab, abcd, cdabc, cdcd\}$*.*

L^* *est l'ensemble de tous les mots que l'on peut obtenir en concaténant un nombre quelconque de mots de* L*. On défini* $L^+ = L^* \backslash \{\varepsilon\}$*.*

Définition 1.5 (Dérivation)**.** *Une dérivation est l'application d'une des règles de* P*. On note* $\alpha \Rightarrow \beta$*.*
- *La clôture transitive de* $\Rightarrow$ *sur* P *est noté* $\Rightarrow^+$
- *La clôture transitive et réflexive de* $\Rightarrow$ *sur* P *est noté* $\Rightarrow^*$

Définition 1.6 (Langage engendré par une grammaire)**.** *Le langage* $L(G)$ *engendré par la grammaire* G *est défini par*

$$L(G) = \{u \in \Sigma^* | S \Rightarrow^* u\}$$

Définition 1.7 (Arbre de dérivation)**.** *On appelle* arbre de dérivation *un arbre tel que :*
- *La racine est l'axiome*
- *Les feuilles sont des symboles terminaux, ou* ε*.*
- *Les nœuds sont des symboles non-terminaux,*
- *Les fils d'un nœud* α *sont* $\beta_0, \ldots, \beta_n$ *ssi* $\alpha \Rightarrow \beta_0 \ldots \beta_n$ *est une production.*

Définition 1.8 (Grammaire ambigüe)**.** *On dit qu'une grammaire est* ambigüe *s'il existe* $u \in L(G)$ *tel que* u *est au moins deux arbres de dérivation distincts.*

2 Hiérarchie de Chomsky

Type 0 Aucune restriction

Type 1 Grammaire contextuelle, c'est à dire que le règles de production sont du type :

$$\alpha A \beta \to \alpha \gamma \beta$$

avec $A \in \mathcal{V}, \alpha, \beta, \gamma \in (\mathcal{V} \cup \Sigma)^*, \gamma \neq \varepsilon$.

Type 2 Grammaire hors-contextuelle ou algébrique, c'est à dire que les règles de production sont du type :

$$X \to \alpha$$

avec $\alpha \in (\mathcal{V} \cup \Sigma)^*, X \in \mathcal{V}$.

Type 3 Grammaires :
- régulière à *gauche*,

$$X \to Ya \qquad \text{ou}$$
$$X \to a$$

avec $X, Y \in \mathcal{V}$ et $a \in \Sigma$.
- régulière à *droite*,

$$X \to aY \qquad \text{ou}$$
$$X \to a$$

avec $X, Y \in \mathcal{V}$ et $a \in \Sigma$.

Définition 2.1 (Type d'un langage)**.** *On dit qu'un langage est de type n s'il existe une grammaire de type n qui puisse l'engendrer et qu'il n'existe pas de grammaire de type $m, m > n$ qui le puisse.*

Index