Chapitre II : *Les grammaires*

I- *Introduction :*

pour définir un langage contenant un nombre fini de mots, il suffit de citer ces mots (définition par extension). Mais lorsque le langage est infini, nous avons recours à un système qui nous permet de générer les mots, ce système est appelé « grammaire ».

Exemple : soit Q le langage des nombres décimaux :

L'alphabet : {0,1,2,…,9,• }
ND →E•F (entier, fractionnaire)
E →CN / C
C → 0/1/2/../9
N→ C / CN
F →CN / C

II- *Définition formelle d'une grammaire :*

1- **Définition :** une grammaire **G** est un 4-uplet **G=(V_T,V_N,S,R)** tel que :

- V_T : Vocabulaire Terminal $|V_T| = n$
- V_N, : Vocabulaire Non Terminal $|V_N| = p$

  tel que $V_T \cap V_N = \emptyset$ et $V_T \cup V_N = V$

- S $\in V_N$ un symbole particulier appelé Axiome (start symbol)
- R est un ensemble fini de règles tel que :
  $(u \rightarrow v) \in R$ et $u \in V^+$ et $v \in V^*$, « → » signifie que u se réécrit en v.

2- La relation dérivation directe « $\underset{G}{\Rightarrow}$ » :

Soit **G=(V_T,V_N,S,R)** et x,y de mots de V* , on dit que y dérive directement de x

$(x \underset{G}{\Rightarrow} y)$ **ssi** $\exists (u \rightarrow v) \in R$ et $x=\alpha u\beta$ et $y=\alpha v\beta$ avec $\alpha, \beta \in V^*$

3- Relation « $\overset{*}{\Rightarrow}_G$» fermeture transitive de « $\underset{G}{\Rightarrow}$ » :

On dit que y dérive de x $(x \overset{*}{\Rightarrow} y)$ s'il existe une suite finie $w_0,\ldots,w_p$ ($w_i \in V^*$) tel que :

$$\mathbf{x= w_0 \quad y=w_p \ \text{ et } w_i \underset{G}{\Rightarrow} w_{i+1}}$$

$$\mathbf{x \underset{G}{\Rightarrow} w_1 \underset{G}{\Rightarrow} w_2 \ldots w_{i-1} \underset{G}{\Rightarrow} y}$$

4- Langage généré par une grammaire :

Le langage engendré par une grammaire G est définie par :

$$\mathbf{L(G) = \{ x \,/\, (x \in V_T^*) \text{ et } (S \overset{*}{\Rightarrow} x) \}}$$

Remarque : Soient G1 et G2 deux grammaires :

[L(G1) = L(G2)] ⇔ [G1 est équivalente à G2]

Exemples :

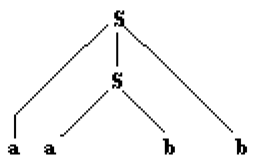
1- Soit G1= ({a,b},{S},S,R) avec R : S → aSb / ab
   L(G1) = $\{a^n b^n \ / \ n \geq 1\}$
2- Soit G2= ({a,b},{S,A,B},S,R) avec R : S → aA   A→ aA / aB   B → bB / b
   L(G2) = $\{a^p b^n \ / \ p \geq 2 \ , \ n \geq 1\}$

5- Arbres de dérivation :

Soit la grammaire **G=(V_T,V_N,S,R)** , arbre étiqueté est un " **arbre de dérivation"** dans G <u>**ssi**</u> :

- L'alphabet des étiquettes est inclus dans $V_N \cup V_T$.
- Les nœuds sont étiquetés par des éléments de V_N.
- Les feuilles sont étiquetées par des éléments de V_T.
- L'étiquette de la racine est l'axiome
- Pour tout nœud étiqueté par **A et ses fils étiquetés par** $f_1,f_2,...,f_n$ est associé une règle R de la forme $\mathbf{A} \longrightarrow f_1 f_2 \ldots f_n$ dans G

Exemple : mot a^2b^2 dans G_1



Règles de G_1 appliquées : $S \rightarrow^1 aSb \rightarrow^2 aabb$

III- *Classification de Chomsky :* il existe 4 types de grammaires. Cette classification est faite suivant la forme des règles de réécriture.

1- Grammaire de type 0 :(grammaire sans restriction)

Une grammaire est de type 0 si ses règles ne sont soumises à aucune condition **: u→v avec u∈V^+ et v∈V***

2- Grammaire de type 1 :(grammaire sous contexte, à contexte lié)

Une grammaire est de type 1 si toutes les règles sont de la forme **:**

u→v avec |u|≤| v|

Exemple :

Soit G= ({a,b,c},{S,B,C},S,R) avec

R : S → aSBC / aBC   CB→ BC   aB →ab   bB→bb
bC→bc   cC→cc

3- **Grammaire de type 2 :(grammaire algébrique, à contexte libre)** G est de type 2 si toutes ses règles sont de la forme :

A→ α , *avec* α ∈ V* et A ∈ V_N

4- **Grammaire de type 3 :(grammaire régulière)** G est de type 3 si toutes ses règles sont de la forme :

$A \rightarrow aB$, $A \rightarrow a$ $avec$ $a \in V_T$ et $A,B \in V_N$ **Linéaire à Droite**

Ou exclusif

$A \rightarrow Ba$, $A \rightarrow a$ $avec$ $a \in V_T$ et $A,B \in V_N$ **linéaire à Gauche**

Remarque : toute grammaire de type i est aussi de type i-1.

IV- Types de langages :

Exemple : soit G1= ({a,b},{S,A,B},S,R) avec

R : $S \rightarrow aS$ /ABb $A \rightarrow aA$ / a $B \rightarrow b$
L(G1) = $\{a^p b^2 / p \geq 1\}$ **G1 est de type 2**

Soit G2= ({a,b},{S,A,B},S,R) avec

R : $S \rightarrow aS$ /aA $A \rightarrow aA$ / bB $B \rightarrow b$
L(G2) = $\{a^p b^2 / p \geq 1\}$ **G1 est de type 3**

L(G1) = L(G2) donc L est de type 3

Définition : le type d'un langage est le type maximum des grammaires qui l'engendrent.

Remarques :

- Le type d'une grammaire augmente quand on augmente les conditions sur la forme des règles.
- Le type d'un langage diminue quand on augmente les conditions sur la forme des mots appartenant à ce langage.
- Une grammaire est dite ambiguë si un mot a au moins deux arbres de dérivations différents ou deux dérivations les plus à gauche différentes dans G.