Université Mouloud MAMMERI de Tizi-Ouzou
Faculté de génie électrique et informatique
Département d'informatique
Module : **Théorie des Langages** – L2 – année : 2017/2018

Chapitre II

Classification des grammaires et des langages

Par : Mr HABET Md-Said
(reprise des notes de cours de Mr KHEMLICHE Salem)

Le but de ce chapitre est de présenter la classification des grammaires, puis des langages, introduite par **Chomsky**. La classification, ou hiérarchie, des grammaires induit une classification des langages définis par ces grammaires.
Cette classification comporte quatre classes de grammaires, nommées de type 0 à type 3, comme c'est présenté dans la définition suivante.

Définition II.1 :

Soit g = <π, N, S, P> une grammaire formelle.
La grammaire g est dite de type i (i = 0, 1, 2, 3) si elle vérifie les conditions du type i. Ces conditions portent sur la forme des règles de production de g.

0) Type 0 :
Les grammaires de ce type (type 0) n'ont aucune restriction sur leurs règles de production.

1) Type 1 :
Toute règle de production de P est de la forme : $\alpha \rightarrow \beta$ avec $|\alpha| \leq |\beta|$; sauf pour la règle $S \rightarrow \varepsilon$, qui, si elle est présente dans P, impose que S n'apparaisse dans aucune partie droite de règle de production de P. Pour α et β, ils sont des éléments quelconques de $(\pi \cup N)^*.N.(\pi \cup N)^*$ et $(\pi \cup N)^+$.
Une grammaire vérifiant ces conditions est dite *grammaire monotone* ou *grammaire de type 1*.

2) Type 2 :
Toute règle de production de P est de la forme : $A \rightarrow w$ avec $A \in N$ et $w \in (\pi \cup N)^*$.
Une grammaire vérifiant ces conditions est dite *grammaire à contexte libre* ou *grammaire de type 2*.

3) Type 3 :
Toute règle de production de P est de la forme : $A \rightarrow wB$ ou $A \rightarrow w$ avec $A, B \in N$ et $w \in \pi^*$.
Une grammaire vérifiant ces conditions est dite *grammaire régulière* ou *grammaire de type 3*.

Exemple II.2 :

Soit g = <π, N, S, P> la grammaire définie par :
π = {a, b},
N = {S, A, B},
P = { $S \rightarrow AB$;
$A \rightarrow aA \mid \varepsilon$;
$B \rightarrow bB \mid \varepsilon$ }

Cette grammaire est de type 2, car toutes ses règles de production vérifient les conditions du type 2, à savoir, que toutes les parties gauches de règles de production sont réduites à un seul non terminal, les parties droites étant quelconques.
On peut remarquer que les règles de production (2) à (5), de P, vérifient les conditions du type 3 ; mais cela n'est pas suffisant pour dire que g est de type 3 : pour qu'une grammaire soit de type i (i= 0, 1, 2, 3) il faut que toutes ses règles de production vérifient les conditions du type i.

Définition II.3 :

Une grammaire régulière g = <π, N, S, P>, telle que présentée à la définition II.1, est dite aussi *grammaire régulière à droite*. Il existe aussi la définition d'une *grammaire régulière à gauche* :
les règles de production d'une telle grammaire vérifient :
$A \to Bw$ ou $A \to w$ avec $A, B \in N$ et $w \in \pi^*$. Si g vérifie ces conditions, elle est également de type 3.

Remarque II.4 :

Pour qu'une grammaire soit régulière à droite il faudrait que toutes ses règles vérifient la clause 3) de la définition II.1. De même, pour qu'une grammaire soit régulière à gauche il faudrait que toutes ses règles vérifient la définition II.3.
Une grammaire, dont certaines règles sont de la forme $A \to wB$ et d'autres de la forme $A \to Bw$ avec $A, B \in N$ et $w \in \pi^+$, n'est pas régulière. C'est le cas, par exemple, de la grammaire ayant les règles :
$\{S \to aS \mid Sb \mid \varepsilon\}$: elle n'est ni régulière à droite ni régulière à gauche. Mais il existe quelques grammaires qui sont à la fois régulières à droite et à gauche comme celle ayant les règles :
$\{S \to ab \mid bab \mid A \,;\; A \to bb \mid a\}$.

Définition II.5 :

Soit g = <π, N, S, P> une grammaire.
Lorsque toute règle de production de P est de la forme : $\alpha_1 A \alpha_2 \to \alpha_1 w \alpha_2$; où :
$A \in N$ et $\alpha 1, \alpha 2 \in (\pi \cup N)^*$ et aussi $w \in (\pi \cup N)^+$; on dit que la grammaire g est à *contexte lié*.

Théorème II.6 :

Toute grammaire monotone est équivalente à une grammaire à contexte lié.
Démonstration :
La démonstration de ce théorème sera présentée dans le chapitre sur les langages de type 1.

Remarque II.7 :

En général, toute règle de production de type i (i = 1, 2, 3) est admise dans le type i-1 ; à une exception : les règles de productions, de type 2, de la forme $A \to \varepsilon$ (avec $A \neq S$) ne sont pas admises dans le type 1.
Cette constatation révèle des relations d'inclusion entre les classes de grammaires :
- toute grammaire de type 3 est aussi de type 2 ;
- toute grammaire de type 1 est de type 0.
- les grammaires de type 2 qui sont de type 1 sont celles qui ne contiennent pas de règle : $A \to \varepsilon$ (avec $A \neq S$; ou si $A=S$, alors S apparait dans une partie droite de règle de production). Cependant une grammaire, qui contient de telles règles, peut être transformée en une autre grammaire équivalente qui soit de type 1 : la méthode sera donnée dans le chapitre sur les langages de type 2.

Lorsqu'on parle de type i d'une grammaire g, on entend le plus petit type satisfaisant la grammaire. C'est ainsi que lorsqu'on dit qu'une grammaire est de type 0, cela sous-entend qu'elle n'est pas de type 1 (ni de type 2, ni 3). De même si on dit qu'une grammaire est de type 2, on sous-entend qu'elle n'est pas de type 3 …

Exemple II.8 :

Soit g = <π, N, S, P> la grammaire définie par :

$\pi = \{a, b\}$,

$N = \{S, A\}$,

P : $S \rightarrow aSb \mid A$;

$A \rightarrow cA \mid \varepsilon$

Cette grammaire est de type 2. Elle génère le langage $\{ a^n.c^m.b^n \ / \ n, m \geq 0 \}$.

Elle n'est pas de type 1, mais elle peut être transformée en une grammaire, équivalente, de type 2 et qui est aussi de type 1. Une telle grammaire est la suivante :

g = <π, N', S', P'> la grammaire définie par :

$\pi = \{a, b\}$,

$N = \{S', S, A\}$,

P : $S' \rightarrow S \mid \varepsilon$;

$S \rightarrow aSb \mid ab \mid A$;

$A \rightarrow cA \mid c$

Exemple II.9 :

Soit g = <π, N, S, P> la grammaire définie par :

$\pi = \{a, b\}$,

$N = \{S, A, B, C\}$,

P : $S \rightarrow AC$;

$A \rightarrow aAb \mid B$

$Bb \rightarrow bbB$

$BC \rightarrow \varepsilon$

Cette grammaire est de type 0. Elle engendre le langage : $L = \{ a^n.b^{2n} \ / \ n \geq 0\}$.

On peut trouver une grammaire de type 2 équivalente g' = <π, N', S, P'> définie par :

$N' = \{S\}$,

P' : $S \rightarrow aSbb \mid \varepsilon$

Mais on ne peut pas, pour cette grammaire, trouver une grammaire de type 3 qui lui soit équivalente (cela sera démontré dans le chapitre sur les langages réguliers). Il en découle que le langage L n'est pas régulier.

Définition II.10 :

Un langage engendré par une grammaire de type i (i=0, 1, 2, 3) est dit langage de type i.

Il existe, donc, quatre classes de langages :

- langages de type 0 ;
- langages de type 1 ou langages à contexte lié ;
- langages de type 2 ou langages à contexte libre, ou encore langages algébriques ;
- langages de type 3 ou langages réguliers.

Remarque II.11 :

En notant T_i = classe des langages de type i (i=0, 1, 2, 3) ; on a :

$T_3 \subset T_2 \subset T_1 \subset T_0$; comme cela est représenté dans la figure suivante :