

Chapitre 1:

Les grammaires

Un langage L étant fixé, comment savoir si une chaîne W donnée appartient à L ?

Solution possible :

Decrire une grammaire adéquate engendrant L qui pourrait permettre de :

- Produire toute les chaînes de L
- Tester une chaîne en renvoyant un message d'error si elle n'est pas dans L

1) Définition formelle d'une grammaire :

Une grammaire est un quadruplet

$G: (X, N, S, P)$

- X : Est un ensemble fini de symboles terminaux
- N : est un ensemble de symboles non terminaux (variables appelé alphabet non terminal).

On note par V l'ensemble $X \cup N$

• **S** : $(S \in N)$: c'est une variable particulière (start).

• **P** : est un ensemble finie de regle de production. Les regles doivent etre de la forme :
ou $A \in V^+$ et $v \in V^*$.

L'ensemble P est noté comme suit :
$$P = \{ (u, v) \mid u \in V^+ \text{ et } v \in V^* \}$$

Les elements de P seront noté $u \rightarrow v$ au lieu de (u, v) : un element de P est une regle de production.

N/B : Les regles de production seront ordonné en respectant l'ordre des variables auxquels s'applique.

La 1^{ere} regle concerne tjrs l'axiome, le ~~types de~~

Les types de regles :

* **Regle vide :**

C'est une regle de la forme $A \rightarrow \epsilon$.
ce type de regle est tres utile car il nous permet d'autoriser certaine possibilites sans obligation.

* Règle directement réursive :

C'est une règle de la forme ~~$A \rightarrow A$~~
 $A \rightarrow v A w$ avec v et $w \in V^*$

autrement dit la variable A figure encore dans la chaîne de production de a .

De telles règles permettent de multiplier des occurrences de symboles et de rallonger indéfiniment des expressions.

En effet une telle règle permet ~~d'obtenir~~ d'obtenir à partir de A des chaînes, $v A w$, $vv A ww$, $vvv A xxxw$, ...

Exp :

Soit la grammaire G suivante définie par les règles :

- 1 $S \rightarrow a b$
- 2 $S \rightarrow a A S b$
- 3 $A \rightarrow b S b$
- 4 $A S \rightarrow b$
- 5 $A \rightarrow \epsilon$

• Donner la def formelle de G .

$$G = (X, N, S, P) /$$

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

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

$$S = \{S\}$$

$$P = \left\{ (S, ab), (S, aASb), (A, bSb), (AS, b), (A, \epsilon) \right\}$$

Prog:

$$G: 1: S \rightarrow ab$$

$$2: S \rightarrow \epsilon$$

$$3: A \rightarrow Aa$$

$$4: A \rightarrow \epsilon$$

Afin de donner une représentation plus synthétique de la grammaire. On peut se servir d'une barre vertical comme séparateur.

entre les règles de n variable,
de la manière suivante

$$G: S \rightarrow abSA \mid E \quad (2)$$

$$A \rightarrow Aa \mid E \quad (3) \quad (4)$$

2) Dérivations du langage engendré par une grammaire:

1) Définitions:

On appelle langage par une grammaire l'ensemble des chaînes terminales produites par cette grammaire.

Une chaîne terminale, composé uniquement des symboles terminaux est dite engendré par une grammaire ~~est~~ si elle peut être dérivée la production de la grammaire en partant de l'axiome

Definition 1:

Soit A une variable et $x, y, u \in V^*$

Soit A une variable en appliquant une

$A \leftarrow u$
 règle $A \leftarrow u$ à la variable qui figure
 dans une chaîne V qui est égale $x \cancel{A} y$, $x A y$
 en produit la chaîne $W = x u y$, cette opération
 est dite dérivation et elle est notée par
 $\cancel{x A y} \Rightarrow x u y$

On peut répéter ceci autant de fois que l'on
 veut en appliquant de nouveaux à la chaîne
 ainsi produite

Definition 2:

Une chaîne W dérive d'une chaîne V quelconque
 si à partir de V on peut arriver à W par un
 nombre finie d'application de la grammaire,
 règle

Rq: On peut préciser le numéro de la règle appliquée
 au-dessus de la ~~grande~~ double flèche si cela
 s'avère utile.

$$V \xRightarrow{n} W$$

Exemple:

Soit la grammaire $G = (X, N, S, P) / \Lambda = \{a, b\}$
 $N = \{S, A\}$ $S = \{S\}$

$$P = \{ (S, ab), (S, aASb), (A, bSb), (AS, b), (A, \epsilon) \}$$

1. $S \rightarrow ab$	$S \rightarrow ab aASb$
2. $S \rightarrow aASb$	$A \rightarrow bSb \epsilon$
3. $A \rightarrow bSb$	$AS \rightarrow b$
4. $AS \rightarrow b$	
5. $A \rightarrow \epsilon$	

$\textcircled{1} \swarrow$
 $S \rightarrow ab \in L(G)$

$\textcircled{2} \searrow$
 $aASb \xRightarrow{\textcircled{4}} abb \in L(G)$

$\textcircled{5} \searrow$
 $abSbSb \Rightarrow$
 $\textcircled{1} \searrow$
 $aSb \Rightarrow$
 $\textcircled{2} \searrow$
 $aAabb \Rightarrow$
 $aAaASbb \Rightarrow$

NB: à chaque dérivation on applique
 une seule règle

Definition 3: On appelle squelette de la dérivation la suite des numéros de règle utilisés pour générer un mot donné

Exp:

$$f_0 \Rightarrow f_1 \Rightarrow f_2 \Rightarrow f_3 \Rightarrow f_4$$

$f_4 \Rightarrow x^*$

axiome ϵ

le mot est

On appelle dérivation la suite 1 - 3 - 2 - 4.

Definition 4:

Le langage formel engendré par G est l'ensemble $L(G) = \{ w \in X^* \mid \exists S \Rightarrow w \}$ → dérivation

C'est l'ensemble de tous les mots qui appartiennent à X^* qui dérive de l'axiome S .

~~$$A \Rightarrow a \mid \epsilon \in L(G)$$~~

Exemple:

Soit $G : S \rightarrow abSA \mid \epsilon$

$A \rightarrow Aa \mid \epsilon$

Prouver que la chaîne ub aba est dans $L(G)$.

$abab^u$

$S \xrightarrow{1} abSA \xrightarrow{2} ababSAA$
 $\xrightarrow{3} ababAA \Rightarrow ababAaA \Rightarrow$

Exercice :

Soit $G : S \rightarrow aaaSbb$

1. $S \rightarrow a$
2. $S \rightarrow \epsilon$

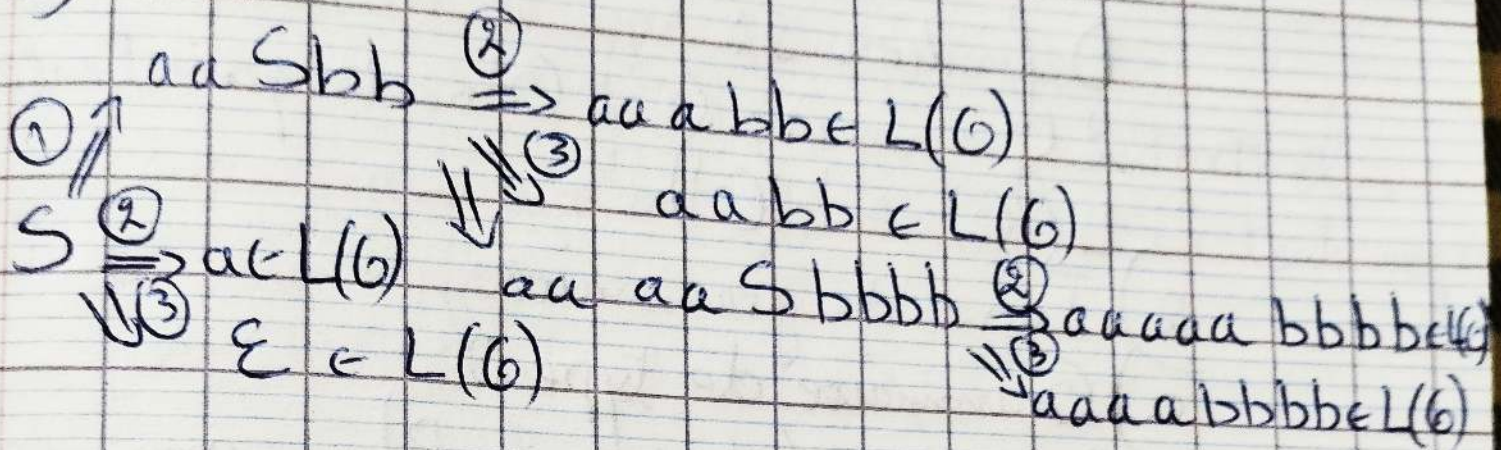
- 1) Donner une définition formelle de G .
- 2) Trouver des mots générés par la grammaire G qui nécessitent un nombre de dérivation ≤ 3 .

$$1) G = (X, N, S, P) \quad X = \{a, b\}$$

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

$$P = \{ \cancel{a}, \cancel{b}, \cancel{\epsilon} (S, aaaSbb), (S, a), (S, \epsilon) \}$$

2)



$$L(G) = \{ aabbb, aabb, aaaaa bbbb, \\ aaaa bbbb, a, \epsilon \}$$

Definition 5:

2 Grammaire G et G' sont équivalentes si
 $L(G) = L(G')$

2) Types de grammaire:

a) Grammaire de type 0:

Soit $G = (X, N, S, P)$ une grammaire formelle.

Une grammaire formelle G de type 0 est une grammaire où on introduit aucune contrainte sur la forme des règles

b) Grammaire de type 1:

Se sont des grammaire contextuelle.
Une grammaire est dite de type 1 si toute les règles de production sont de la forme:

$u \cdot A \cdot v \rightarrow u \cdot w \cdot v$

avec $A \in N$

$u, v, w \in V^*$ ($V = N \cup X$)

On dit que A se transforme en w dans le contexte $u \cdot v$ où u et v sont des chaînes quelconque.

c) Grammaire de type 2:

Se sont des grammaire non contextuel.

Une grammaire est dite de type 2 si toute règle est de la forme : $A \rightarrow W$ avec $A \in N$ et $W \in V^*$

d) Grammaire de type 3:

Grammaire régulière

Une grammaire formelle G de type 3 est dite régulière si toutes les règles sont de la forme.

Grammaire régulière à gauche

Grammaire régulière à droite

$A \rightarrow a.B$

$A \rightarrow a$

$A \rightarrow \epsilon$

$A \rightarrow B.a$

$A \rightarrow a$

$A \rightarrow \epsilon$

avec $A \in N$ et $a \in X$

$B \in N$

→ Tout grammaire de type 3 est de type 2 et 1

- Les grammaires de type 2 sont utilisées pour analyser la syntaxe des langages de programmation. Les grammaires de type 3 sont utiles pour effectuer l'analyse lexicale au niveau des compilateurs.

Exp:

$$G: (X, N, S, P)$$

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

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

$$S = \{A\}$$

$$P = \{A \rightarrow A \mid bB \mid BB$$

$$B \rightarrow A$$

$$aA \mid bB \rightarrow aC \mid bD$$

$$aC \rightarrow bA$$

$$bD \rightarrow aB$$

$$\begin{aligned} C &\rightarrow \epsilon \\ D &\rightarrow \epsilon \end{aligned}$$

Le grammaire G est une grammaire de type 0
 puisque il n'y a aucune contrainte de forme
 de la grammaire

Note:

Si les règles de la grammaire contiennent
 une seule variable à gauche et si elle
 n'est pas de type 3 elle est forcément
 de type 2.

Exp: $S \rightarrow aSb ; S \rightarrow \epsilon$

G est de type 2 car toutes les règles sont
 de la forme $A \rightarrow W$

Exercice:

$$G = \begin{cases} S \rightarrow aSBC \\ S \rightarrow aBC \\ aB \rightarrow ab \\ bB \rightarrow bb \\ bC \rightarrow bc \\ cC \rightarrow cc \end{cases}$$

$G = (X, N, S, P)$ $X = \{a, b, c, S, B, C\}$
 $S = \{S\}$ $N = \{S, A, B, C\}$
 Donner Définition formelle:
 1) $P = \{(S, ASBC), (S, aBC), (aB, ab), (bB, bb), (BC, bc), (C, cc)\}$

- 2) G est elle de type 3?
 3) G est elle " " " ?

2) La grammaire G ni de type 3 ni de type 2 car les règles 3 à 6 contiennent plus qu'un symboles à leurs partie gche.

3) Type 1:

$\epsilon . S . \epsilon \rightarrow \epsilon . a S B C . \epsilon$
 $\epsilon . S . \epsilon \rightarrow \epsilon a B c . \epsilon$
 $a B \epsilon \rightarrow a B \epsilon$
 $b B \epsilon \rightarrow b B \epsilon$
 $b C \epsilon \rightarrow B C \epsilon$
 $C \epsilon \rightarrow C C \epsilon$