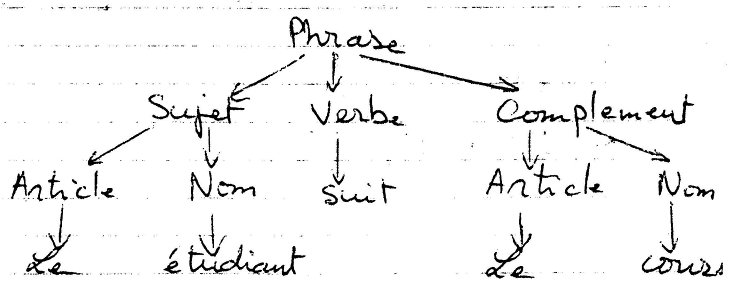
Exemple

La phrase "l'etudiant suit le cour" est construite à partir des règles décrites si deçu

Phrase → sujet verbe complement
Sujet → Article Nom / Article Adjectif Nom
Verb → suit / regarde /
Complement → Article Nom
Article → Le / La / Un / Une
Nom → Etudiant / cours
Adjectif → sérieux



On distingue deux types de symboles

a) Symboles non-terminaux :
Ce sont des symboles intermédière pour produir de nouveau objets
(ce sont les symboles qui font encore définir)

b) Symboles terminaux
Ce sont les lettre du langage qui sont les objets non productif

Grammaires :

Appelée aussi "système de substitution", intuitivement système de substitu
est un ensemble de règles qui remplacent des sous-mots par d'autres sous-mots

Une grammaire est une description générative (intensive) d'un langage

Grammaire $\xrightarrow{\text{générer}}$ Langage

Formellement :
Une grammaire est un quadruplet $G < T, N, S, P >$ où
T : un ensemble non vide de Términaux (l'alphabet sur lequel est définie le langa
N : un ensemble de non Términaux tel que $T \cap N = \emptyset$ (les symboles intermédiaire
S : L'axium $\in N$
P : ensemble de règles ou de production (ou de récriture) chaque règle est de la

Une règle de production $\alpha \rightarrow \beta$ précise que la séquence de symbole α peut être remplacée par la séquence de symbole β.

α est appelé membre gauche de la dérivation, β est appelé membre droit

Notation :

Les elements de T sont désigné par des lettres minuscules de l'alphabet latin et les elements de N sont désigné par les lettres majuscules

$T : \{a, b, c, \dots\}$

$N : \{A, B, C, \dots\}$

Remarque :

Plusieurs règles ayant même membre gauche seront notés en écrivant à droite du symbole flèche $\rightarrow$ les differentes membres droit séparées par /

$$\begin{matrix} \alpha \rightarrow \beta_1 \\ \alpha \rightarrow \beta_2 \end{matrix} \iff \alpha \rightarrow \beta_1 / \beta_2$$

Le langage définie par une grammaire est l'ensemble des mots composé de términaux qui peuvent être obtenue à partir de l'axiome par application des règles.

~~Formellement~~ : Derivation directe et indirecte :

On introduit la Notion de dérivation entre mots. D'abord dérivation directe (en une étape), en suite dérivation indirecte (en plusieurs étapes).

Une dérivation consiste en l'application d'une ou plusieurs règles à partir d'un mot.

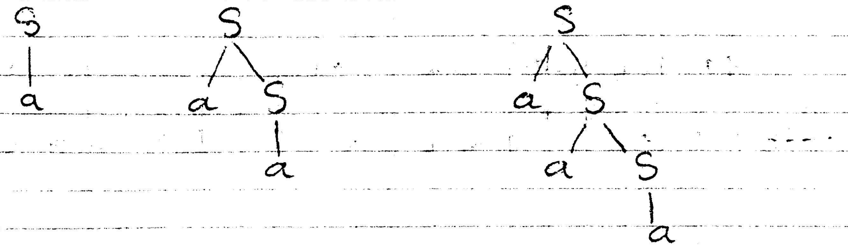
Exemple

Soit la grammaire $G < T, N, S, P >$ où

$T = \{a\}$ $N = \{S\}$

$$P = \left\{ \begin{matrix} S \rightarrow aS \\ S \rightarrow a \end{matrix} \right\} \Rightarrow S \rightarrow aS/a$$

Quel est le langage généré par la grammaire ?



a aa aaa

ou bien

$S \rightarrow a$

$S \rightarrow aS \rightarrow aa$

$S \rightarrow aS \rightarrow aaS \rightarrow aaa$

$\vdots$

$L = \{ a, aa, aaa, \cdots \}$

$L = \{ a, a^2, a^3, \dots \}$

ou bien $L = \{ a^n \ / \ n \geq 1 \}$

Une dérivation directe s'obtien en une étape par contre une dérivation indirecte s'obtien en plusieurs étapes

Une dérivation indirecte est notée $\overset{*}{\Longrightarrow}$

$S \overset{*}{\Longrightarrow} aaa$

<u>Exemple 2 :</u>

Soit $G < T, N, S, P >$ où $T = \{0, 1\}$

$N = \{S\}$

$P = \{ S \rightarrow 0S1 \ / \ \varepsilon \}$

Quel est le langage généré par la grammaire ?

$S \rightarrow \varepsilon$

$S \rightarrow 0S1 \rightarrow 01$

$S \rightarrow 0S1 \rightarrow 00S11 \rightarrow 0011 \iff 0^2 1^2$

$S \rightarrow 0S1 \rightarrow 00S11 \rightarrow 000S111 \rightarrow 000111 \iff 0^3 1^3$

$\vdots$

$L = \{ \varepsilon, 01, 0^2 1^2, 0^3 1^3, \dots \}$ donc $L = \{ 0^n 1^n \ / \ n \geq 0 \}$

0S1 est obtenue directement à partir de S

Définition :

Un langage engendré (généré) par une grammaire G notée L(G) est définie par l'ensemble des mots appartenent à T^* généré (directement ou indirectement) à partir de l'axiume

$$L(G) = \{ w / S \stackrel{*}{\Longrightarrow} w \text{ et } w \in T^* \}$$

Grammaire équivalente :

Deux grammaire G et G` sont diter équivalentes si elles engendrent (génèrent) le même langage

$$G \equiv G` \iff L(G) = L(G`)$$

Exemple :

Trouver une grammaire qui génère le langage suivant :

$$L = \{ a^n b^m / n \geqslant 0 \quad m \geqslant 0 \}$$

$$L = \{ \varepsilon, a, b, ab, aaab, bbb, b\cancel{ba}, \dots \}$$

$$G_1 \begin{cases} S \longrightarrow AB \\ A \longrightarrow aA / \varepsilon \\ B \longrightarrow bB / \varepsilon \end{cases} \qquad G_2 \begin{cases} S \longrightarrow aS / B \\ B \longrightarrow bB / \varepsilon \end{cases}$$

Remarque.

Un langage peut être généré par plusieurs grammaire et une grammaire ne génère qu'un seule langage

Classification de Grammaires et des langages

CHOMSKY a définit 4 types de grammaires formelles suivant la nature des règles de production des grammaires

type 3 : Grammaires Régulières

a) Grammaires Régulière Droites :

Si toutes les règles de production sont de la forme :

$A \rightarrow wB$ ou $A \rightarrow w$ avec $A, B \in N$ et $w \in T^*$

b) Grammaires Régulière Gauche :

Si toutes les règles de production sont de la forme :

$A \rightarrow Bw$ ou $A \rightarrow w$ avec $A, B \in N$ et $w \in T^*$

type 2 : Grammaires Algébriques (Grammaire à contexte libre)

Si toute les productions sont de la forme $A \rightarrow w$

avec $A \in N$ et $w \in (T \cup N)^*$

$A \rightarrow aABb$ ✓
$A \rightarrow aAb$ ✓
$B \rightarrow ABCD$ ✓
$A \rightarrow \varepsilon$ ✓
$A \rightarrow aa$ ✓

type 1 : Grammaire à contexte lié

Si toute les règles de production de P son de la forme

$\alpha A \beta \rightarrow \alpha w \beta$ $\alpha, \beta \in (T \cup N)^*$ $A \in N$, $w \in (T \cup N)^*$

$aAb \rightarrow accb$
$aA \rightarrow accc$
$\varepsilon BB \rightarrow cd\, B$
$\varepsilon A \varepsilon \rightarrow aAb$

Et seule l'axiome peut générer le mot vide et dans ce cas S n'apparaît pas dans un membre droit d'une production.

Les Grammaires Monotones :

Une grammaire monotone est une grammaire dont toute les règles sont de la forme $\alpha \rightarrow \beta$ avec $|\alpha| \leqslant |\beta|$ et seule l'axiome peut générer le mot vide et dans ce cas S n'apparait pas dans un membre droit d'une production

$AB \rightarrow aab$ ✓
$aAb \rightarrow aAAb$ ✓
$aAb \rightarrow acb$ ✓
$AB \rightarrow a$ ✗
$A \rightarrow ab$ ✓

Théorème :

Une grammaire monotone est de type 1

Remarque :

Toute grammaire monotone peut être transformée vers une grammaire à contexte lié et intuitivement.

type 0 : Grammaire sans réstriction

La forme des règles de production dans P n'est l'objet d'aucune restriction

Classification des Langages

A chaque type de grammaire est associé un type de langage.

Les grammaires de type 3 génèrent les langages réguliers.

Les grammaires de type 2 génèrent les langages Algébriques.

Les grammaires de type 1 génèrent les langages à contexte lié.

Les grammaires de type 0 génèrent tous les langages décidable (reconnu en un temps fini par la machine)

$A \rightarrow a\beta$ 3, 2, 1, 0

$A \rightarrow a\beta b$ $\cancel{3}$, 2, 1 0

$AB \rightarrow aaB$ $\cancel{3}$, $\cancel{2}$, 1 0

$ABC \rightarrow d$ $\cancel{3}$, $\cancel{2}$, $\cancel{1}$, 0

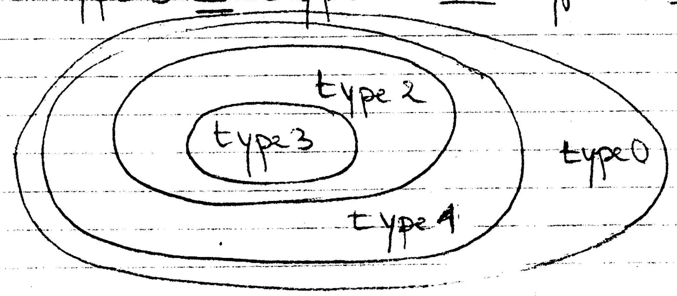
Tout langage est de type 0

Tout langage de type 1 est de type 0

Tout langage de type 2 est de type 1 et type 0

Tout langage de type 3 est de type 2, type 1 et type 0

Ainsi type 3 $\subseteq$ type 2 $\subseteq$ type 1 $\subseteq$ type 0



Remarque :

Pour un Langage donné il faut trouver la grammaire du type le plus stricte (du plus haut rond)

Si une donne une grammaire de type 1 qui génère un langage cela veut dire qu'il n'existe pas une grammaire de type $i+1$ qui le génère.

Exemple classique de grammaire

Le langage classique de type 3 est

Type 3 .

$L = \{ a^n b^m \ / \ n \geq 0 \quad m \geq 0 \}$

$G = < \{a, b\}, \{S, R\}, S,$
$\quad P = \{ S \rightarrow aS / R$
$\quad\quad R \rightarrow bR / \varepsilon \} >$

Type 2 :

$L = \{ a^n b^n \ / \ n \geq 0 \}$

$G = < \{a, b\}, \{S\}, S,$
$\quad\quad P = \{ S \rightarrow aSb / \varepsilon . \} >$

Type 1 :

Le langage de type 1 classique est :

$L = \{ a^n b^n c^n \ / n \geqslant 1 \}$

$G = < \{a, b, c\}, \{S, R, T\}, S$

$P = \{ S \rightarrow aRbc \ / abc$

$R \rightarrow aRTb \ / aTb$

$Tb \rightarrow bT$

$Tc \rightarrow cc \} >$

$S \rightarrow aRbc \rightarrow aaRTbbc \rightarrow aaaTbTbbc \rightarrow aaaTbbTbc \rightarrow$

$\rightarrow aaaTbbbTc \rightarrow aaaTbbbcc \rightarrow aaabTbbcc \rightarrow aaabbTbcc$

$\rightarrow aaabbbTcc \rightarrow aaabbbccc$