INFO 010 – Théorie des langages – Partie pratique

Introduction aux grammaires

Sébastien COLLETTE et Gilles GEERAERTS

Définition

Une grammaire est un quadruplet $G = \langle V, T, P, S\rangle$ où

- V est l'ensemble des variables ;
- T est l'ensemble des terminaux ;
- P est l'ensemble des règles de production. $P \subseteq (V \cup T)^* V (V \cup T)^* \times (V \cup T)^*$;
- $S \in V$ est le symbole de départ.

Exemple

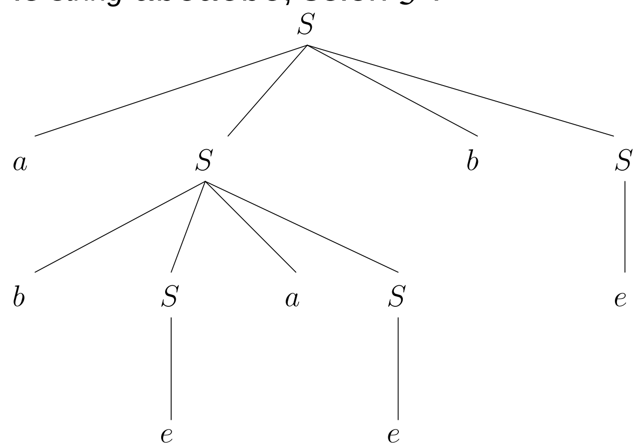
Soit une grammaire $\mathcal{G}$ dont les règles de P sont :

$$\begin{array}{lcl} S & \rightarrow & aSbS \\ & & bSaS \\ & & e \end{array}$$

S est la seule variable (ou symbole non-terminal) ; il est aussi le symbole de départ. On a $T = \{a, b, e\}$.

Arbre de dérivation

Pour le *string* abeaebe, selon $\mathcal{G}$:



Hiérarchie de Chomsky

Classe 0 : Grammaires non restreintes
Pas de restrictions sur la forme des règles

Hiérarchie de Chomsky – suite

Classe 1 : Grammaires *context sensitive* Toute règle dans P est de la forme :

$$\alpha \rightarrow \beta$$

avec $|\alpha| \leq |\beta|$. Il peut en outre y avoir une règle de la forme :

$$S \rightarrow \varepsilon$$

où S est le symbole de départ, et à condition que S n'apparaisse pas dans une partie droite.

Hiérarchie de Chomsky – suite

Définition alternative des grammaires CS
Toute règle dans P est de la forme :

$$\alpha A \gamma \rightarrow \alpha \beta \gamma$$

avec $\beta \in (V \cup T)^+$. Ou bien de la forme :

$$S \rightarrow \varepsilon$$

et S n'apparaît dans aucun membre de droite.

Hiérarchie de Chomsky – suite

Classe 2 : Grammaires *context free* Toute règle est de la forme :

$$A \rightarrow \alpha$$

Classe 3 : Grammaires régulières 2 classes :

Grammaire linéaires droites règles du type :

$$A \rightarrow wB \text{ ou } A \rightarrow w$$

Grammaires linéaires gauches règles du type :

$$A \rightarrow Bw \text{ ou } A \rightarrow w$$

Exercice 1

Décrivez, en français, les langages générés par les grammaires suivantes, et donnez leur type :

(a)
$$\begin{array}{rcl} S & \rightarrow & abcA \\ & & Aabc \\ A & \rightarrow & \varepsilon \\ Aa & \rightarrow & Sa \\ cA & \rightarrow & cS \end{array}$$

(b)
$$\begin{array}{rcl} S & \rightarrow & 0 \\ & & 1 \\ & & 1S \end{array}$$

(c)
$$\begin{array}{rcl} S & \rightarrow & a \\ & & *SS \\ & & +SS \end{array}$$

Exercice 2

Soit la grammaire $\mathcal{G}$ suivante :

$$\begin{array}{rcl} S & \rightarrow & AB \\ A & \rightarrow & Aa \\ & & bB \\ B & \rightarrow & a \\ & & Sb \end{array}$$

- Cette grammaire est-elle régulière ?
- Donnez :
  - l'arbre de dérivation pour $baabaab$, $bBABb$ et $baSb$
  - les dérivations gauche et droite de $baabaab$.

Exercices 3 & 4

- Écrivez une grammaire *context-free* qui génère toutes les chaînes de a et de b (dans n'importe quel ordre), tel qu'il y a plus de a que de b. Testez votre grammaire sur $baaba$ en en donnant une dérivation.
- Écrivez une grammaire *context-sensitive* qui génère toutes les chaînes de a, de b et de c (quel que soit l'ordre), ayant le même nombre de a, de b et de c. Donnez la dérivation de $cacbab$ selon votre grammaire.

Solution de l'exercice 1

1. Grammaire non restreinte donnant toutes les suites de abc.
2. Grammaire linéaire droite donnant toutes les suites de 1 éventuellement terminées par un 0, ou bien le *string* 0.
3. Grammaire context free donnant toutes les expressions arithmétiques utilisant la somme et le produit sur la variable a, et ce, en notation polonaise.

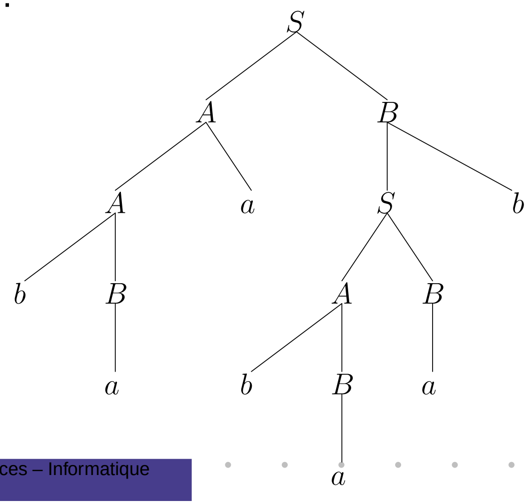
Solution de l'exercice 2

Il s'agit d'une grammaire context free mais pas régulière.

Justification : il ne peut y avoir à la fois une règle de la forme $A \rightarrow \alpha B$ et une règle de la forme $A \rightarrow B\alpha$ dans une grammaire régulière.

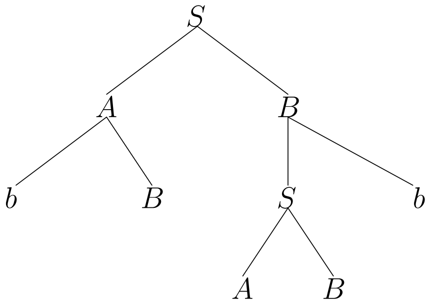
Solution de l'exercice 2 – suite

$baabaab$:



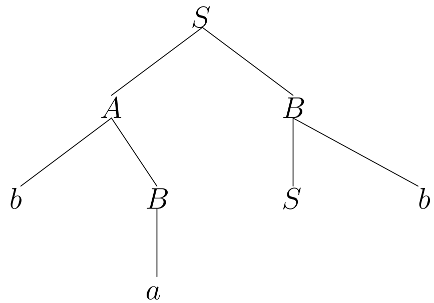
Solution de l'exercice 2 – suite

$bBABb$:



Solution de l'exercice 2 – suite

$baSb$:



Solution de l'exercice 2 – suite

La dérivation gauche de $baabaab$ est :
$S \Rightarrow AB \Rightarrow AaB \Rightarrow bBaB \Rightarrow baaB \Rightarrow baaSb \Rightarrow baaABb \Rightarrow baabBBb \Rightarrow baabaBb \Rightarrow baabaab$

La dérivation droite est : $S \Rightarrow AB \Rightarrow ASb \Rightarrow AABb \Rightarrow AAab \Rightarrow AbBab \Rightarrow Abaab \Rightarrow Aabaab \Rightarrow bBabaab \Rightarrow baabaab$

Solution de l'exercice 3

$$S \rightarrow bSa|aSb|abS|baS|Sab|Sba|Sa|aS|a$$

On dérive $baaba$ de la façon suivante :

$$S \Rightarrow baS \Rightarrow baabS \Rightarrow baaba$$

Solution de l'exercice 4

$$\begin{array}{rcl|rcl}
S & \rightarrow & S' \mid \varepsilon & & & \\
S' & \rightarrow & ABCS \mid ABC & CA & \rightarrow & AC \\
AB & \rightarrow & BA & CB & \rightarrow & BC \\
AC & \rightarrow & CA & A & \rightarrow & a \\
BA & \rightarrow & AB & B & \rightarrow & b \\
BC & \rightarrow & CB & C & \rightarrow & c
\end{array}$$

$S \Rightarrow ABCS \Rightarrow ABCABCS \Rightarrow ABCABC \Rightarrow ACBABC \Rightarrow CABABC \Rightarrow CABACB \Rightarrow CABCAB \Rightarrow CACBAB \Rightarrow cACBAB \Rightarrow caCBAB \Rightarrow cacBAB \Rightarrow cacbAB \Rightarrow cacbaB \Rightarrow cacbab$