

Cartouche du document

Année : ING 1

Matière : Théorie des langages

Activité : Travail dirigé

Objectifs

- Notions de langage hors contexte ou algébrique.
- Normalisation de Chomsky
- Algorithme CKY
- Application au traitement des langues

Définition :

Un langage hors contexte est aussi appelé langage algébrique

Une grammaire hors contexte (ou algébrique) est un quadruplet T, N, S, P où :

- T : ensemble des éléments terminaux.
- N : ensemble des éléments non terminaux.
- S : élément non terminal initial
- R : ensemble de règles où les règles peuvent être de la forme :
 - $X \rightarrow a$ où $a \in T$ et $X \in N$
 - $X \rightarrow Y$ où $Y \in (N \cup T)^*$ et $X \in N$

Définition :

Une grammaire hors contexte (sans ϵ) est dite sous forme normale de Chomsky si et seulement si toutes les règles sont de la forme :

$A \rightarrow a$ où $a \in T$

$A \rightarrow BC$ où $B, C \in N$

Sommaire des exercices

- 1 - Normalisation de Chomsky
- 2 - Des propositions très relatives

Corps des exercices

1 - Normalisation de Chomsky

Énoncé :

Il s'agit dans cet exercice de normaliser une grammaire puis de se servir de cette normalisation pour définir un algorithme (en n^3) qui teste si un mot appartient ou n'appartient pas au langage engendré par la grammaire.

Question 1)**Énoncé de la question**

Soit la grammaire G basée sur l'alphabet $\{a, b\}$.

- (1) $S \rightarrow b A$
- (2) $S \rightarrow a B$
- (3) $A \rightarrow b A A$
- (4) $A \rightarrow a S$
- (5) $A \rightarrow a$
- (6) $B \rightarrow a B B$
- (7) $B \rightarrow b S$
- (8) $B \rightarrow b$

Mettre cette grammaire sous forme normalisée de Chomsky.

Procédure de transformation sous forme normale de Chomsky:

- 1. remplacer tous les terminaux x en partie droite des règles par des non terminaux en ajoutant des règles de la forme $X \rightarrow x$
- 2. transformer les parties droites des règles comme suit : $X \rightarrow Y Z W$ par deux règles $X \rightarrow Y V$ et $V \rightarrow Z W$
- 3. transformer les parties droites des règles comme suit :
 $X \rightarrow Y$ par $X \rightarrow W Z$ si $Y \rightarrow W Z$

Solution de la question**Etape 1**

Pour chaque terminal, on crée une règle avec un nouveau terminal comme membre gauche et le terminal comme membre droit.

- (9') $A1 \rightarrow a$
- (10') $B1 \rightarrow b$

Etape 2

Pour chaque règle comportant plus de deux membres à droite ((3) et (6)), on scinde la partie droite en créant de nouveaux non terminaux et de nouvelles règles.

- (11') $A2 \rightarrow A A$
- (12') $B2 \rightarrow B B$

Étape 3

On normalise toutes les règles et on obtient la grammaire suivante :

(1') $S \rightarrow B1 A$

(2') $S \rightarrow A1 B$

(3') $A \rightarrow B A2$

(4') $A \rightarrow A1 S$

(5') $A \rightarrow a$

(6') $B \rightarrow A1 B2$

(7') $B \rightarrow B1 S$

(8') $B \rightarrow b$

(9') $A1 \rightarrow a$

(10') $B1 \rightarrow b$

(11') $A2 \rightarrow A A$

(12') $B2 \rightarrow B B$

Question 2)**Énoncé de la question**

Montrer que le mot **aabbab** appartient au langage engendré par la grammaire. On utilisera l'algorithme de Chomsky.

Algorithme de Cocke, Younger et Kasami qui permet de tester si un mot w est reconnu par une grammaire sous forme normale de Chomsky.

- On note n la longueur du mot w : $n = |w|$. On définit v une matrice de dimensions (n,n)

•

Pour $i = 1$ à n

DEBUT

$v[i,1] = \{A \text{ tel que } A \text{ est le membre gauche d'une règle } A \rightarrow a \text{ et } a \text{ est le } i\text{ème} \text{ symbole du mot } w\}$

FIN

Pour $j = 2$ à n

DEBUT

Pour $i = 1$ à $n - j + 1$

DEBUT

$v[i,j] =$

Pour $k = 1$ à $j - 1$

```

DEBUT
v[i,j] = v[i,j]  {A tel que A est le membre gauche
d'une règle A → B C avec B = v[i,k] et
C = v[i+k,j-k]}
FIN
FIN
FIN
FIN

```

- Le mot w est reconnu par la grammaire $\Leftrightarrow S \in v[1,n]$

Solution de la question

Le tableau v avec le mot **aabbab** est le suivant :

a	a	b	b	a	b
A,A1	A,A1	B,B1	B,B1	A,A1	B,B1
A2	S	B2	S	S	
A	B	B	B		
S	S	B2			
A	B				
S					

Nous constatons que $S \in v[1,6]$, donc "aabbab" est un élément du langage.

2 - Des propositions très relatives

Énoncé :

On s'intéresse aux constructions de phrases avec des propositions relatives. On définit la grammaire suivante : On considère la grammaire hors contexte suivante:

```

s  sn sv
sn  det n reln | det n rela | np reln | np rela
reln pron sv
rela proa sn vt
sn  det n
sn  np
sv  vi | vt sn
proa que
pron qui
vt  regarde | regardent | mange | mangent

```

vi dort | dorment | tombe | tombent
det une | un | la | le | des | les
n pommes | pomme | femme | femmes
np pierre | marie

Légende :

- reln <==> proposition relative nominative
- rela <==> proposition relative accusative
- proa <==> pronom relatif accusatif
- pron <==> pronom relatif nominatif
- vi <==> verbe intransitif
- vt <==> verbe transitif
- det <==> déterminant
- n <==> nom commun
- np <==> nom propre

Question 1)

Énoncé de la question

Vérifier l'appartenance de la phrase **Une pomme que Pierre regarde tombe** au langage reconnu par la grammaire en utilisant un arbre syntaxique.

Solution de la question

Question 2)

Énoncé de la question

Appliquer l'algorithme CKY pour vérifier :

- l'appartenance de la phrase **Une pomme que Pierre regarde tombe** au langage engendré par la grammaire
- le rejet de la phrase **Une pomme qui Pierre regarde tombe** du langage engendré par la grammaire

Solution de la question

Se conférer à la réponse globale de l'exercice