LECTURE 7

Context Free Grammar (CFG) Chomsky Normal Form (CNF) .....

Introduction

There are many CFG's for any given CFL. When reasoning about CFL's, it often helps to assume that a grammar for it has some particularly simple form.

1. Chomsky Normal Form (CNF) :

A context-free grammar is said to be in Chomsky normal form if every production is of one of these two types:

A → BC (where B and C are nonterminals).

A → σ (where σ is a single terminal).

For languages that include the empty string ε, the rule S →ε may also be allowed, where S is the start symbol, as long as S does not occur on the right -hand side of any rule

2. Converting a CFG to CNF:

The conversion of CFG to CNF can be achieved by the following steps :

1) Add new start symbol.
2) Eliminating ε-productions (A → ε, where A is any nonterminal)
   i. Identify nullable nonterminal (nonterminals from which ε can be derived).
   ii. Add new production rules, if necessary, to remove ε.
3) Eliminating unit productions (A → B, where A and B are any nonterminals)
   i. Identify A-derivable nonterminals, If A →B is a production and B ≠ A, then B is A-derivable.
   ii. Remove unit productions by adding the production rules of any A-derivable to A.
4) Converting remaining rules to proper form.

Example 1:

Convert the following Context-free grammar production rules to Chomsky Normal Form:

S → TU | V
T → aTb | ε
U → cU | ε
V → aVc | W
W → bW | ε

Solution:

1) Add new start symbol
   S_0→S
   S → TU | V
   T → aTb | ε
   U → cU | ε
   V → aVc | W
   W → bW | ε
2) Eliminating ε-productions:
   i. nullable nonterminal: {S, T, U, V, W}.
   ii. Remove ε by adding new production rules, if necessary.
   S_0→S
   S → TU | V | T | U
   T → aTb | ab
   U → cU | c
   V → aVc | W | ac
   W → bW | b
3) Eliminating unit productions:
   i. S_0-derivable: { S }.
   S-derivable: { T, U, V, W }.
   T-derivable: { }.
   U-derivable: { }.
   V-derivable: { W }.
   W-derivable: { }.
   ii. Add the production rules of S_0-derivable set to S_0, S-derivable set to S, and the production rules of V-derivable set to V.
   S_0 → TU | aTb | ab | cU | c | aVc | bW | b | ac
   S → TU | aTb | ab | cU | c | aVc | bW | b | ac
   T → aTb | ab
   U → cU | c
   V → aVc | bW | b | ac
   W → bW | b

4) Converting to Chomsky normal form.

$S_0 \rightarrow TU \mid X_1Y_1 \mid X_1X_2 \mid X_3U \mid c \mid X_1Y_2 \mid X_2W \mid b \mid X_1X_3$

$S \rightarrow TU \mid X_1Y_1 \mid X_1X_2 \mid X_3U \mid c \mid X_1Y_2 \mid X_2W \mid b \mid X_1X_3$

$X_1 \rightarrow a$

$Y_1 \rightarrow TX_2$

$X_2 \rightarrow b$

$X_3 \rightarrow c$

$Y_2 \rightarrow VX_3$

$T \rightarrow X_1Y_1 \mid X_1X_2$

$U \rightarrow X_3U \mid c$

$V \rightarrow X_1Y_2 \mid X_2W \mid b \mid X_1X_3$

$W \rightarrow X_2W \mid b$

Example 2:

Convert the following CFG production rules to Chomsky Normal Form (CNF):

S→ AA

A→ B | BB

B→ abB | b | bb

Solution:

1) Add new start symbol

$S_0 \rightarrow S$

S→ AA

A→ B | BB

B→ abB | b | bb

2) Eliminating ε-productions:

   i. nullable nonterminal: {}.

$S_0 \rightarrow S$

S→ AA

A→ B | BB

B→ abB | b | bb

3) Eliminating unit productions:

   i. S_0-derivable: { S }.
      A-derivable: { B }.

   ii. Add the production rules of S_0-derivable set to S_0, and the production rules of A-derivable set to A.

$S_0 \rightarrow AA$
$S \rightarrow AA$
$A \rightarrow abB \mid b \mid bb \mid BB$
$B \rightarrow abB \mid b \mid bb$

4) Converting to Chomsky normal form.

$S_0 \rightarrow AA$
$S \rightarrow AA$
$A \rightarrow XB \mid b \mid ZZ \mid BB$
$X \rightarrow YZ$
$Y \rightarrow a$
$Z \rightarrow b$
$B \rightarrow XB \mid b \mid ZZ$

Example 3:

Convert the following CFG production rules to Chomsky Normal Form (CNF):

$S \rightarrow AB$

$A \rightarrow aAA \mid \varepsilon$

$B \rightarrow bBB \mid \varepsilon$

Solution:

1) Add new start symbol

$S_0 \rightarrow S$
$S \rightarrow AB$
$A \rightarrow aAA \mid \varepsilon$
$B \rightarrow bBB \mid \varepsilon$

2) Eliminating ε-productions:
   i. nullable nonterminal: $\{S_0, S, A, B\}$.
   ii. Remove ε by adding new production rules, if necessary.

$S_0 \rightarrow S$
$S \rightarrow AB \mid A \mid B$
$A \rightarrow aAA \mid aA \mid a$
$B \rightarrow bBB \mid bB \mid b$

3) Eliminating unit productions:
   i. S_0-derivable: { S }.
      S-derivable: { A, B }.

ii. Add the production rules of S_0-derivable set to S_0, and the production rules of S-derivable set to S.

$S_0 \rightarrow AB \mid aAA \mid aA \mid a \mid bBB \mid bB \mid b$
$S \rightarrow AB \mid aAA \mid aA \mid a \mid bBB \mid bB \mid b$
$A \rightarrow aAA \mid aA \mid a$
$B \rightarrow bBB \mid bB \mid b$

4) Converting to Chomsky normal form.

$S_0 \rightarrow AB \mid PA \mid QA \mid a \mid RB \mid SB \mid b$
$S \rightarrow AB \mid PA \mid QA \mid a \mid RB \mid SB \mid b$
$P \rightarrow QA$
$Q \rightarrow a$
$R \rightarrow SB$
$S \rightarrow b$
$A \rightarrow PA \mid QA \mid a$
$B \rightarrow RB \mid SB \mid b$

Example 4:

Convert the following CFG production rules to Chomsky Normal Form (CNF):
$S \rightarrow ASA \mid aB$
$A \rightarrow B \mid S$
$B \rightarrow b \mid \varepsilon$

Solution:

1) Add new start symbol

$S_0 \rightarrow S$
$S \rightarrow ASA \mid aB$
$A \rightarrow B \mid S$
$B \rightarrow b \mid \varepsilon$

2) Eliminating ε-productions:
   i. nullable nonterminal: { A, B}.
   ii. Remove ε by adding new production rules, if necessary.

$S_0 \rightarrow S$
$S \rightarrow ASA \mid aB \mid AS \mid SA \mid \mathbf{S} \mid a$
$A \rightarrow B \mid S$
$B \rightarrow b$

Note: The rule S→S is useless rule and it must be remove.

3) Eliminating unit productions:
   i. S_0-derivable: { S }.
      A-derivable: { B, S }.
   ii. Add the production rules of S_0-derivable set to S_0, and the production rules of A-derivable set to A.

$S_0 \rightarrow ASA \mid aB \mid AS \mid SA \mid a$
$S \rightarrow ASA \mid aB \mid AS \mid SA \mid a$
$A \rightarrow b \mid ASA \mid aB \mid AS \mid SA \mid a$
$B \rightarrow b$

4) Converting to Chomsky normal form.

$S_0 \rightarrow XA \mid YB \mid AS \mid SA \mid a$
$X \rightarrow AS$
$Y \rightarrow a$
$S \rightarrow XA \mid YB \mid AS \mid SA \mid a$
$A \rightarrow b \mid XA \mid YB \mid AS \mid SA \mid a$
$B \rightarrow b$

3. Homework:

HW 1:

Convert the following CFG to Chomsky Normal Form (CNF):

$S \rightarrow aY \mid Ybb \mid Y$
$X \rightarrow \varepsilon \mid a$
$Y \rightarrow aXY \mid bb \mid XXa$