Chomsky Normal Form (CNF)

A CFG is said to be in CNF if every production of the grammar is in the form

a) Elements in R.H.S should either be two variables

or

a Terminal

→ A CFG is in CNF if the productions are in the following forms

$A \rightarrow a$

$A \rightarrow BC$

where A, B and C are non terminals & a is a terminal

CNF

A CFG is said to be in CNF if every production of the grammar is of the form

$A \to a$ or

$A \to BC$

where A, B, C are non-terminal

$\boxed{A, B, C \in V_N}$

and

$\boxed{a \in \Sigma}$

The R.H.S of the CFG into CNF will contain either (i) single terminal or (ii) Two variables.

Steps to Convert CFG ~~~~ to CNF

1) If the start symbol S occurs on some right side, create a new start symbol S' and a new production $S' \rightarrow S$

2) Remove null productions

3) Remove unit productions

4) Replace each productions $A \rightarrow B_1 --- B_n$, where $n > 2$, with $A \rightarrow B_1 X_1$ where $X_1 \rightarrow B_2 -- B_n$

Repeat this step for all productions having two or more symbols on the right side.

5) If the R.H.S of any production is in the form

$A \rightarrow aB$

where a is a terminal & A & B are non-term., then the production is replaced by

$A \rightarrow X_aB$ & $X \rightarrow a$. Repeat this step for every production which is of the form $A \rightarrow aB$.

Scanned with CamScanner

eg. Convert the CFG into CNF, whose production is given below:

$S \rightarrow ABC \mid aC$

$A \rightarrow a$

$B \rightarrow b$

$C \rightarrow c$

Soln.

CFG (V_N, Σ, P, S)

$V_N = \{S, A, B, C\}$

$\Sigma = \{a, b, c\}$

Step 1: Remove null production & unit production.

In the given CFG, no null production and unit production is present.

Step 2: ~~Cons~~ Now find out the productions that has more than two variables in R.H.S

$S \rightarrow ABC$

After removing these, we get

$S \rightarrow ADE$

$E \rightarrow BC$

$\boxed{\begin{array}{l} S \rightarrow AX_1 \\ X_1 \rightarrow BC \end{array}}$

3) Elimination of Terminals on R.H.S

Now change the prod.

consider,

$S \rightarrow aC$

Gives

$S \rightarrow DC$

$D \rightarrow a$

$\boxed{\begin{array}{l} S \rightarrow X_a C \\ X_a \rightarrow a \end{array}}$

Now, the productions

$S \rightarrow AE \mid DC$

$A \rightarrow a$

$B \rightarrow b$

$C \rightarrow c$

$D \rightarrow a$

$E \rightarrow BC$

② Convert the CFG into CNF, whose productions are given below:

$S \to aAD$

$A \to aB \mid bAB$

$B \to b$

$D \to d$

Sol.

CFG $G = \{V_N, \Sigma, P, S\}$

$V_N = \{S, A, B, D\}$

$\Sigma = \{a, b, d\}$

$S = \{S\}$

Step 1) Elimination of null production & unit production

In the given CFG, no null and unit production is present

Step 2. Elimination of terminals on R.H.S of

Consider the production, which are not in CNF.

$S \rightarrow aAD$ P_1 is constructed as follows.

$A \rightarrow aB \mid bAB$

$$
\left.
\begin{array}{l}
S \rightarrow aAD \\
\rightarrow S \rightarrow X_a AD \\
\rightarrow A \rightarrow X_a B \\
A \rightarrow X_b AB \\
X_a \rightarrow a \\
X_b \rightarrow b
\end{array}
\right]
$$

P_1

1) $S \rightarrow aAD$ gives

$S \rightarrow X_a AD$ and $X_a \rightarrow a$

2) $A \rightarrow aB$ gives rise to

$A \rightarrow X_a B$ and $X_a \rightarrow a$

3) $A \rightarrow bAB$ gives

$A \rightarrow X_b AB$ and $X_b \rightarrow b$

$V_N' = \{ S, A, B, D X_a, X_b \}$

Scanned with CamScanner

③ Restricting the no. of variables on R.H.S

$S \rightarrow X_a A D$

$A \rightarrow X_a B \mid X_b A B$

$B \rightarrow b$

$D \rightarrow d$

$C_a \rightarrow a$

$C_b \rightarrow b$

$A \rightarrow X_a B$, $B \rightarrow b$, $D \rightarrow d$, $C_a \rightarrow a$, $C_b \rightarrow b$ are added to P_2.

$S \rightarrow X_a A D$ is replaced by

$S \rightarrow X_a X_1$, $X_1 \rightarrow AD$

$A \rightarrow X_b AB$ is replaced by

$A \rightarrow X_b X_2$, $X_2 \rightarrow AB$

Let

$V_N = \{ S, A, B, D, C_a, C_b, X_1, X_2 \}$

$\{a, b, d\}$, P_2, $\{S\}$

(3). Find a Grammar in CFG.

$S \to aAbB$

$A \to aA \mid a$

$B \to bB \mid b$

Solution CFG $G(V_N, \Sigma, P, S)$

$V_N = \{S, A, B\}$

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

$S = \{S\}$

Step 1. Elimination of ~~constant~~ null & ~~variable~~ unit production

In the given CFG, no null & unit productions

Elimination of terminals on R.H.S.

Step 2: Let $G_1 = \{V'_N, \{a, b\}, P_1, S\}$

where P_1 & V'_N are constructed as follows:

(i) $A \to a$, $B \to b$ are added to P_1

(ii) $S \to aAbB$ yields

$S \to X_a A X_b B$

→ $A \to aA$ — $A \to X_a A$, $X_a \to a$

→ $B \to bB$ — $B \to X_b B$, $X_b \to b$

Scanned with CamScanner

$V'_N = \{ S, A, B, X_a, X_b \}$

Step 3: Now find out the production that has more than two variables in R.H.S.

P_1 consist of

$S \rightarrow X_a A X_b B$

$A \rightarrow X_a A$

$B \rightarrow X_b B$

$X_a \rightarrow a$

$X_b \rightarrow b$

$A \rightarrow a$

$B \rightarrow b$

$S \rightarrow \underline{X_a} \, \underline{A \, X_b B}$ is replaced by.

$X_a \underbrace{A X_b B}_{X_1}$

$S \rightarrow X_a X_1$

$A \underbrace{X_b B}_{X_2}$

$X_1 \rightarrow A X_2$

$X_2 \rightarrow X_b B$

The remaining production in P_1 are added to P_2. Let (

$G_2 = (\{ S, A, B, X_a, X_b, X_1, X_2 \} \{a, b\}, P_2, S)$

where P_2 consists of

$S \to X_a X_1$, $X_1 \to A X_2$, $X_2 \to X_b B$, $A \to X_a A$,
$B \to X_b B$, $X_a \to a$, $X_b \to b$, $A \to a$, $B \to b$.

$\therefore G_2$ is in CNF and equivalent to the given grammar.

4 eg.

Convert it into CNF :

$S \to \sim S \mid [S \supset S] \mid p \mid q$

(S being the only variable)

Solution

CFG $G(V_N, \Sigma, P, S)$

$V_N = \{S\}$

$\Sigma = \{ [, \sim, \supset,], p, q$

Step 1. Elimination of null & unit production

In the given CFG, no null & unit production

Step 2: Elimination of Terminals, which are not in the form, on R.H.s.

Let $G' = \{V_N', \Sigma, P', S)$ where

P' & V_N' are constructed as follows.

(i) $S \to p \mid q$ are added to P_1

(ii) $S \to \sim S$ induces

$S \to X_a S$, $X_a \to \sim$

(iii) $S \to [S \supset S]$ induces to

$S \to X_b\, S\, X_c\, S\, X_d$

$X_b \to [$

$X_c \to \supset$

$X_d \to]$

$V'_N = \{ S, X_a, X_b, X_c, X_d \}$

Step 3. Now find out the production that has more than two variables in R.H.S.

P₁ consist of

$S \to p \mid q$

$S \to X_a S$

$X_a \to \sim$

$S \to X_b S X_c S X_d$

$X_b \to [$

$X_c \to \supset$

$X_d \to]$

$S \rightarrow X_b\, S\, X_c\, S\, X_d$ is replaced by

$S \rightarrow X_b X_1$

$X_1 \rightarrow S X_2$

$X_2 \rightarrow X_c X_3$

$X_3 \rightarrow S X_d$

Let

$G_2 = (\{S, X_a, X_b, X_c, X_d, X_1, X_2, X_3\}, \Sigma,$

where P_2 consists of

$S \rightarrow p \mid q \mid X_a S \mid X_b X_1$

$X_a \rightarrow \sim$

$X_b \rightarrow [$

$X_c \rightarrow \supset$

$X_d \rightarrow]$

$X_1 \rightarrow S X_2$

$X_2 \rightarrow X_c X_3$

$X_3 \rightarrow S X_d$

G_2 is in CNF & equivalent to the given grammar.