5

Properties of Context Free Languages

Objectives :

- Concept of normal form
  1. Chomsky's normal form
  2. Greibach normal form
- Applications of context free grammar
- Properties of context free languages.

5.1 Introduction

As we have discussed in earlier chapters context free languages can be defined by Context Free Grammar (CFG). The context free grammar is most important factor in the compilation of program, when we check the syntax of any programming construct we need CFG to verify it. Hence it is necessary to bring the given grammar to some normal form. We will discuss two commonly used normal forms in this chapter. Then we will discuss various properties of context free languages.

5.2 Normal Forms

As we have seen the grammar can be simplified by reducing the ε production removing useless symbols, unit production. There is also a need to have grammar in some specific form. As you have seen in CFG at the right hand of the production there are any number of terminal or non terminal symbols in any combination. We need to normalize such a grammar. That means we want the grammar in some specific format. That means there should be fixed number of terminals and non terminals, in the context free grammar.

There are two important normal forms - Chomsky's Normal Form and Greibach Normal Form. We will study these with the help of examples.

5.2.1 Chomsky's Normal Form (CNF)

The Chomsky's Normal Form can be defined as

Non terminal → Non terminal. Non terminal
Non terminal → terminal

The given CFG should be converted in the above format then we can say that the grammar is in CNF. Before converting the grammar into CNF it should be in reduced form. That means remove all the useless symbols, ε productions and unit productions from it. Thus this reduced grammar can be then converted to CNF.

Let us solve some examples for Chomsky's Normal Form.

Problems on CNF :

Example 5.1 : *Convert the following CFG into CNF*

$S \rightarrow aaaaS$

$S \rightarrow aaaa$

Solution : As we know the rule for Chomsky's Normal Form is

Non terminal → Non terminal. Non terminal

Non terminal → terminal

But the CFG given is

$S \rightarrow aaaaS$ rule 1

$S \rightarrow aaaa$ rule 2

If we add a rule

$A \rightarrow a$ which is in CNF.

Then out rule 1 and 2 becomes,

$A \rightarrow AAAAS$

$S \rightarrow AAAA$

Let us take

$S \rightarrow A\ \boxed{AAAS}$ can be replaced by P_1

If we define

$P_1 \rightarrow AAAS$ then the rule becomes

$S \rightarrow A\ P_1$ which is in CNF

But the new rule P_1 is not CNF, so let us convert it

$P_1 \rightarrow A\ \boxed{AAS}$ can be replaced by P_2

then $P_1 \rightarrow A\ P_2$ which is in CNF

and $P_2 \rightarrow AAS$ which is not in CNF so let us convert it to CNF

$P_2 \rightarrow A\ \boxed{AS}$ can be replaced by P_3

$P_2 \rightarrow A\ P_3$ where $P_3 \rightarrow AS$

Now both P_2 and P_3 are in CNF.

Collectively rewrite these rules (these all indicate rule 1 of the given CFG).

$$\boxed{\begin{array}{l} S \rightarrow A\ P_1 \\ P_1 \rightarrow A\ P_2 \\ P_2 \rightarrow A\ P_3 \\ P_3 \rightarrow A\ S \end{array}}$$

Now consider rule 2

$S \rightarrow AAAA$

If we break the rule 2 as

$$S \rightarrow \underbrace{\boxed{AA}}_{P_4}\ \underbrace{\boxed{AA}}_{P_5} \quad \text{indicated by}$$

The rule becomes

$S \rightarrow P_4\ P_5$ which is in CNF.

But P_4 and P_5 indicate the same rule. So we can eliminate either of them.

Hence rule 2 becomes

$S \rightarrow P_4\ P_4$

Finally we can collectively show the CFG converted to CNF as

$S \rightarrow A\ P_1$

$P_1 \rightarrow A\ P_2$

$P_2 \rightarrow A\ P_3$

$$P_3 \rightarrow A\,S$$

$$S \rightarrow P_4\,P_4$$

$$P_4 \rightarrow AA$$

$$A \rightarrow a$$

Example 5.2 : *Convert the given CFG to CNF* $S \rightarrow aSa \mid bSb \mid a \mid b$.

Solution : Let us start by adding new symbols for the terminals.

$$S \rightarrow ASA$$

$$A \rightarrow a$$

$$S \rightarrow BSB$$

$$B \rightarrow b$$

Note that although $s \rightarrow a$ and $A \rightarrow a$ are similar still we are not replacing any S. This is because, S is not simply giving a. It has other productions also. We want such a non terminal having only one production and that is a. Same is with B. Hence we have added the rules $A \rightarrow a$ and $B \rightarrow b$.

Let us take $S \rightarrow ASA$ for converting to CNF.

$$S \rightarrow A\,\boxed{SA}$$

replace it by A_1

$$\left.\begin{array}{l} S \rightarrow A\,A_1 \\ A_1 \rightarrow SA \end{array}\right\} \text{both are in CNF}$$

Take, $S \rightarrow B\,\boxed{SB}$

replace it by A_2

$S \rightarrow B\,A_2$ both are in CNF

$$A_2 \rightarrow SB$$

Hence we can write the rule in CNF as

$$S \rightarrow A\,A_1$$

$$A_1 \rightarrow SA$$

$$A \rightarrow a$$

$$S \rightarrow B\,A_2$$

$A_2 \rightarrow SB$

$B \rightarrow b$

$S \rightarrow a$

$S \rightarrow b$

➡ **Example 5.3 :** *Consider* $a=(\{S,A\},\{a,b\},P,S)$

Where P consists of

$S \rightarrow aAS \mid a$

$A \rightarrow SbA \mid SS \mid ba$

Convert it to its equivalent CNF.

ution : Let us consider,

$S \rightarrow aAS$

If we put $R_1 \rightarrow a$

then $S \rightarrow R_1 AS$

and if $R_2 \rightarrow AS$

then S becomes

$S \rightarrow R_1 R_2$ is now in CNF

$S \rightarrow a$ is already in CNF

$A \rightarrow SbA$

Let $R_3 \rightarrow b$

Hence $A \rightarrow S R_3 A$

Add $R_4 \rightarrow R_3 A$

then $A \rightarrow S R_4$

$A \rightarrow SS$ is already in CNF

Now as we have defined $R_1 \rightarrow a$ and $R_3 \rightarrow b$.

$A \rightarrow R_1 R_3$ will be a conversion in CNF of $A \rightarrow ab$.

Finally, we can write

$S \rightarrow R_1 R_2$ $A \rightarrow S R_4$

$S \rightarrow a$ $R_4 \rightarrow R_3 A$

$R_1 \to a$ $\quad R_3 \to b$

$R_2 \to AS$ $\quad A \to SS$

$A \to R_1 R_3$

Example 5.4 : *Convert the given CFG to CNF.*

Consider $G = (V, T, P, S)$

Where $V = \{S, A, B\}$

$T = \{a, b\}$

P consists of

$S \to aB$ $\quad A \to bAA$

$S \to bA$ $\quad B \to b$

$A \to a$ $\quad B \to aS$

$A \to aS$ $\quad B \to aBB$

Solution : Let us start with first rule.

$R_1 \to a$

$S \to R_1 B$ is in CNF for $S \to aB$

$R_2 \to b$

$S \to R_2 A$ is in CNF for $S \to bA$

$A \to a$

$A \to a S$ can be written as

$A \to R_1 S$

Now $A \to bAA$ can be written as

$A \to R_2 \boxed{AA}$ replace it by R_3

i.e. $R_3 \to AA$

then $A \to R_2 R_3$

$B \to a$

$B \to bS$ can be written as

$B \to R_2 S$ is in CNF

Now, $B \to aBB$

i.e. $B \to R_1 BB$ $\quad \because R_1 \to a$

Let $R_4 \rightarrow BB$

then $B \rightarrow R_1 R_4$

Finally we can write,

$$S \rightarrow R_1 B$$
$$S \rightarrow R_2 A$$
$$A \rightarrow R_1 S$$
$$A \rightarrow R_2 R_3$$
$$B \rightarrow b$$
$$B \rightarrow R_2 S$$
$$B \rightarrow R_1 R_4$$
$$R_1 \rightarrow a$$
$$R_2 \rightarrow b$$
$$R_3 \rightarrow AA$$
$$R_4 \rightarrow BB$$

is a Chomsky's Normal Form.

Example 5.5 : *Convert the following CFG to CNF.*

$$S \rightarrow ABA$$
$$A \rightarrow aA \mid \varepsilon$$
$$B \rightarrow bB \mid \varepsilon$$

Solution : In the Chomsky's Normal Form the ε production is not allowed. So first we will eliminate ε productions.

$$A \rightarrow \varepsilon \quad \text{and} \quad B \rightarrow \varepsilon$$

If put ε instead of A and B

We can get

$$S \rightarrow ABA \mid AB \mid BA \mid AA \mid A \mid B$$

Similarly $A \rightarrow aA \mid a$

$$B \rightarrow bB \mid b$$

Now S → A and S → B is unit production getting introduced in the grammar after removal of ε production. So we will remove unit production also

$$S \rightarrow ABA \mid AB \mid BA \mid AA \mid aA \mid a \mid bB \mid b$$

$$A \rightarrow aA \mid a$$

$$B \rightarrow bB \mid b$$

Now let us convert this grammar to Chomsky's Normal Form.

Let

$$S \rightarrow ABA$$

$$\boxed{S \rightarrow ABA}$$

$$\boxed{R_1 \rightarrow BA}$$

then $S \rightarrow AR_1$

Now $\boxed{S \rightarrow AB}$ $\boxed{S \rightarrow BA}$ $\boxed{S \rightarrow AA}$ already in CNF

Let $S \rightarrow aA$

We will define $R_2 \rightarrow a$

∴ $\boxed{\begin{array}{l} S \rightarrow R_2A \\ R_2 \rightarrow a \end{array}}$

$\boxed{S \rightarrow a}$ already in CNF

$S \rightarrow bB$

Let $\boxed{\begin{array}{l} R_3 \rightarrow b \\ S \rightarrow R_3B \end{array}}$ then

$S \rightarrow b$ already in CNF

Now consider the rules,

$$\boxed{A \rightarrow aA}$$

We can rewrite it as

$$\boxed{\begin{array}{l} A \rightarrow R_2A \\ A \rightarrow a \end{array}}$$

Also, $B \rightarrow bB$ can be rewritten as

$$B \rightarrow R_3\ B$$

$$B \rightarrow b$$

Collect all the CNF s to integrate and we will get

$$S \rightarrow A\ R_1\ |\ AB\ |\ BA\ |\ AA\ |\ R_2 A\ |\ R_3 B\ |\ a\ |\ b$$

$$A \rightarrow R_2 A\ |\ a$$

$$B \rightarrow R_3\ B\ |\ b$$

$$R_1 \rightarrow BA$$

$$R_2 \rightarrow a$$

$$R_3 \rightarrow b$$

is done.

Example 5.6 : *Convert the following grammar to Chomsky's Normal Form* $S \rightarrow a\ |\ b\ |\ a\ S\ S.$

Solution : The Chomsky's Normal Form is

$$NT \rightarrow NT.\ NT$$

$$NT \rightarrow \text{terminal}$$

We will convert the given grammar in this form

$S \rightarrow a$ already in CNF

$S \rightarrow b$ already in CNF

$S \rightarrow a\ S\ S$

Now we will convert

$S \rightarrow a\ S\ S$ into equivalent CNF

$$S \rightarrow R_1\ R_2$$

$$R_1 \rightarrow a$$

$$R_2 \rightarrow S\ S$$

∴ The equivalent CNF of given grammar is -

$$S \rightarrow a$$

$$S \rightarrow b$$

$$S \rightarrow R_1 R_2$$
$$R_1 \rightarrow a$$
$$R_2 \rightarrow S\,S.$$

Example 5.7 : *Convert the following grammar to Chomsky's Normal form* $S \rightarrow \sim S|[S \supset S]|p|q.$

Solution : Chomsky's Normal Form is -

Non terminal → Non terminal.Non terminal

Non terminal → terminal

Consider given grammar rule by rule and let us convert it into CNF

$$S \rightarrow \sim S$$

Can be written as

$$S \rightarrow AS$$
$$A \rightarrow \sim$$

Similarly $S \rightarrow [S \supset S]$

Can be written as

$$S \rightarrow BC$$
$$B \rightarrow DE$$
$$E \rightarrow SF$$
$$D \rightarrow [$$
$$F \rightarrow \supset$$
$$C \rightarrow SG$$
$$G \rightarrow]$$

And, $S \rightarrow p$

$$S \rightarrow q$$

is already in CNF

Hence the complete grammar written in CNF is -

$$S \rightarrow AS$$
$$A \rightarrow \sim$$
$$S \rightarrow BC$$

$B \rightarrow DE$

$E \rightarrow SF$

$D \rightarrow [$

$F \rightarrow \supset$

$C \rightarrow SG$

$G \rightarrow]$

$S \rightarrow p$

$S \rightarrow q$

5.2.2 Greibach Normal Form (GNF)

Now we will discuss one more interesting normal form called **Greibach Normal Form.**

The rule for GNF is

Non- terminal → One terminal. Any number of non-terminals

In short,

$$\boxed{NT \rightarrow t.\ NT}$$

For example :

$$\left.\begin{array}{l} S \rightarrow aA \\ S \rightarrow a \end{array}\right\} \text{is in GNF}$$

But

$S \rightarrow AA$ or

$S \rightarrow Aa$ is not in GNF

To convert given CFG into GNF very interesting procedure is followed. We can use two important lemmas based on which it is easy to convert given CFG to GNF.

Lemma 1 : Let,

G = (V, T, P, S) be a given CFG and if there is a production $A \rightarrow Ba$ and $B \rightarrow \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$.

Then we can convert A rule to GNF as

$$A \rightarrow \beta_1 a \mid \beta_2 a \mid \beta_3 a \mid \dots \mid \beta_n a$$