

UNIT-II GRAMMAR.

①

Grammar Introduction:

A grammar of a language (G_1) is defined as,

$$G_1 = (V, T, P, S)$$

V - finite set of objects called Variables (Non-terminal)

T - finite set of objects called Terminals

$S \in V$ - start symbol

P - finite set of productions.

Types of Grammar:

- Type 0 grammars / unrestricted grammars (Recursively Enumerable Language)
- Type 1 grammars / content sensitive grammars
- Type 2 grammars / content free grammars
- Type 3 grammars / Regular grammar.

1) Type 0 grammars

- No restrictions on the production rules
- Production rule is of the form

$$\alpha \rightarrow \beta \mid \alpha \neq \beta$$

Where $\alpha, \beta \rightarrow$ can be strings composed by terminals and non-terminals.

- This grammar can be modeled using Turing Machine.

2) Type 1 grammars

- Content sensitive grammar (CSG) :

- Production rule is of the form

$$\alpha A \beta \Rightarrow \alpha \gamma \beta$$

Here $A \rightarrow$ Non-terminal symbol

$\alpha, \beta, \gamma \rightarrow$ combination of terminals and non-terminals.

- This grammar can be modeled using linear bounded automata.

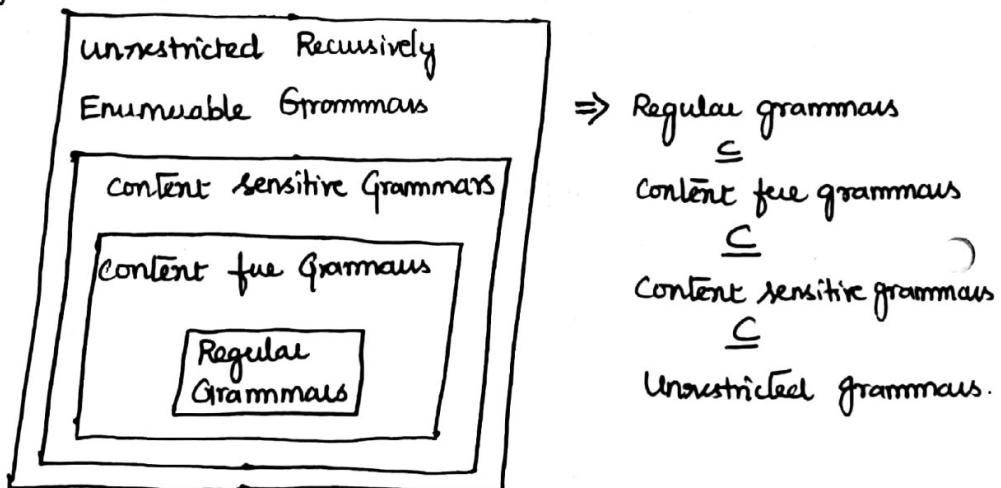
3) Type 2 grammar:

- Content Free Grammar (CFG).
- production rule is of the form $A \rightarrow \alpha$
 - A - Non terminal symbol
 - α - Terminal or non-terminal symbol.
- This grammar can be modeled using push down automata.

4) Type 3 grammar:

- Regular grammar that describe regular / formal languages.
- production rule consists of,
 - only one terminal at the left hand side.
 - Right hand side having a single terminal and may or may not be followed by non terminals.
- $A \rightarrow a, A \rightarrow aB$
- This grammar can be modeled using finite automata.

Chomsky hierarchy.



Content free language and Grammars.

Definition: The content free grammar can be formally defined as a set denoted by $G_1 = (V, T, P, S)$ where V and T are sets of non-terminals and terminals respectively.

P is set of production rules. $NT \rightarrow NT$

$NT \rightarrow T$

S is a start symbol.

Example:

$$P = \{ \begin{array}{l} S \rightarrow S+S \\ S \rightarrow S*S \\ S \rightarrow (S) \\ S \rightarrow \epsilon \end{array} \}$$

Syntax of any English statement is,

SENTENCE \rightarrow Noun VERB

NOUN \rightarrow Rama / Seeta / Gopal

VERB \rightarrow goes / Writes / sings.

Derived strings: "Rama sings"

(2)

Problems:

1) Construct the CFG for the regular expression $(0+1)^*$.

Solution: $CFG_1 = (V, T, P, S)$

$$= (\{S\}, \{0, 1\}, P, S)$$

Example:

$$(0+1)^* = \{\epsilon, 0, 1, 01, 10, 00, 11, \dots\}$$

$$\text{where } P = \{ \begin{array}{l} S \rightarrow 0S | 1S \\ S \rightarrow \epsilon \end{array} \}$$

2) Construct CFG for the language L which has all the strings which are all Palindrome over $\Sigma = \{a, b\}$

Solution: $G_1 = (\{S\}, \{a, b\}, P, S)$

$$S \rightarrow aSa | bSb | a | b, \epsilon$$

Example: abaaba

$$\rightarrow a \underline{s} a$$

$$\rightarrow a b \underline{s} b a$$

$$\rightarrow a b a \underline{s} a b a$$

$$\rightarrow a b a \epsilon a b a$$

$$\rightarrow a b a a b a$$

3) Construct CFG for $\{0^m 1^n \mid 1 \leq m \leq n\}$

Solution: $G_1 = (V, T, P, S)$

$$V = \{S, A\}, T = \{0, 1\}$$

Example: 00111

$$P = \{ \begin{array}{l} S \rightarrow 0S1 | 0A | 01 \\ A \rightarrow 1A | 1 \end{array} \}$$

$$\begin{aligned} S &\rightarrow 0 \underline{S} 1 \\ &\rightarrow 0 0 \underline{S} 1 \\ &\rightarrow 0 0 1 \underline{A} 1 \\ &\rightarrow 0 0 1 1 1 \end{aligned}$$

4) Construct CFG for $L = \{a^m b^n c^p \mid m+n=p \text{ and } p \geq 1\}$

Solution: $G_1 = (V, T, P, S)$

$$V = \{S, A\}, T = \{a, b, c\}$$

$$P = \{ \begin{array}{l} S \rightarrow aSc | bAc | ac | bc \\ A \rightarrow bc | bc \end{array} \}$$

5) Consider the alphabet $\Sigma = \{a, b, (,), +, *, \cdot, /, \dots, \epsilon\}$. Construct a CFG that generates all strings in E^* that are RE over $\{a, b\}$.

Solution:

$$\begin{aligned} E &\rightarrow E+E \\ E &\rightarrow E \cdot E \\ E &\rightarrow E/E \\ E &\rightarrow a/b/\epsilon. \end{aligned}$$

Problems: Grammar to Language

1) If $S \rightarrow aSb/aAb$, $A \rightarrow bAa$, $A \rightarrow ba$ is a CFG then determine CFL.

Solution:

$$\begin{aligned} S &\rightarrow a\underline{S}b \\ &\rightarrow aa\underline{s}bb \\ &\rightarrow aa\underline{a}A\underline{b}bb \\ &\rightarrow \underbrace{aa}_{a^n} \underbrace{ababb}_{b^n} \end{aligned}$$

$$L(G) = \{a^n b^m a^m b^n \mid m, n \geq 1\}$$

2) If $S \rightarrow aSa/bSb/\epsilon$ is CFG. Find $L(G)$.

Solution:

$$\begin{aligned} S &\rightarrow a\underline{S}a \\ &\rightarrow aa\underline{S}aa \\ &\rightarrow aa\underline{b}Sbaa \\ &\rightarrow \underbrace{aa}_{w} \underbrace{bbbaa}_{w^R} \end{aligned}$$

$$L(G) = \{ww^R \mid w \in (a,b)^*\}$$

3) Find the context free language for the following grammar.

(i) $S \rightarrow aSbS/bSaS/\epsilon$

(ii) $S \rightarrow aSb/ab$.

Solution

(i) $S \rightarrow a\underline{S}bS \quad \therefore L$ containing equal number of 'a's and 'b's

$$\begin{aligned} &\rightarrow ab\underline{S}aSbS \\ &\rightarrow ab\underline{a}\underline{S}bS \\ &\rightarrow abab\underline{S} \\ &\rightarrow abab \end{aligned}$$

$$\begin{aligned} S &\rightarrow b\underline{S}aS \\ &\rightarrow ba\underline{S} \\ &\rightarrow ba \end{aligned}$$

(ii) $S \rightarrow aSb/ab$

$$\begin{aligned} S &\rightarrow a\underline{S}b \\ &\rightarrow aa\underline{S}bb \\ &\rightarrow aaabb \end{aligned}$$

L containing equal number of 'a's followed by equal number of 'b's

$$L = \{a^n b^n \mid n \geq 1\}$$

Derivations, Ambiguity, Derivation tree

Derivations: Use the productions from head to body (i.e) from start symbol expanding till reaches the given string.

Two types of derivations are,

(i) Left Most Derivation(LMD)

(ii) Right Most Derivation(RMD)

⇒ LMD is a derivation in which the leftmost non-terminal is replaced first from the sentential form.

(i) $S \xrightarrow[\text{Lm}]{*} \alpha$, then α is left sentential form.

⇒ RMD is a derivation in which rightmost non-terminal is replaced first from the sentential form.

(ii) $S \xrightarrow[\text{Rm}]{*} \alpha$, then α is right sentential form.

Derivation tree (parse tree)

- It is a graphical representation for the derivation of the given production rules for a given CFG.

Properties

i) Root node is always a node indicating start symbol.

ii) Derivation is read from left to right.

iii) Leaf nodes are always terminal nodes.

iv) Interior nodes are always non-terminal nodes.

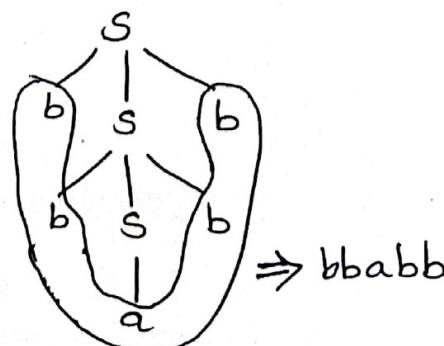
Example:

Consider the grammar G has the production

$S \rightarrow bSb \mid a \mid b$ and string "bbabb"

Derivation:

$$\begin{aligned} S &\rightarrow bSb \\ &\rightarrow b \underline{b} \underline{s} b b \\ &\rightarrow bbabb \end{aligned}$$



Problems: Construct the derivation tree for the string "aaabbabbba" using LMD and RMD. using $S \rightarrow aB/bA$, $A \rightarrow a/as/bAA$, $B \rightarrow b/bS/aBB$

Solution:

LMD:

$$S \Rightarrow aB$$

$$\Rightarrow aaBB \quad (B \rightarrow aBB)$$

$$\Rightarrow aaaBBB \quad (B \rightarrow aBB)$$

$$\Rightarrow aaabBB \quad (B \rightarrow b)$$

$$\Rightarrow aaabbB \quad (B \rightarrow b)$$

$$\Rightarrow aaabbAB \quad (B \rightarrow aBB)$$

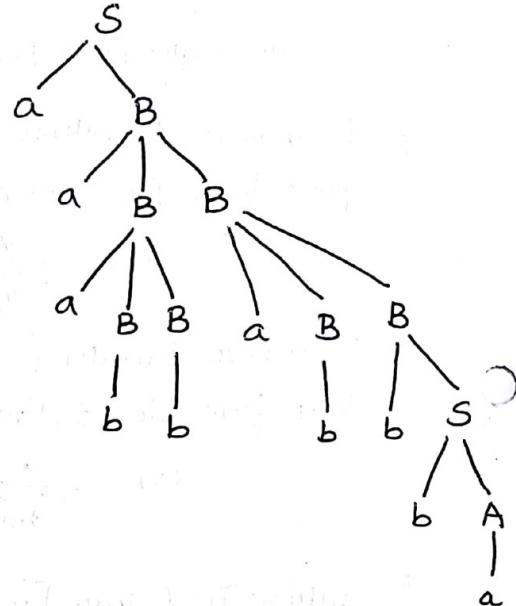
$$\Rightarrow aaabbabB \quad (B \rightarrow b)$$

$$\Rightarrow aaabbabbS \quad (B \rightarrow bs)$$

$$\Rightarrow aaabbabbA \quad (S \rightarrow bA)$$

$$\Rightarrow aaabbabbba \quad (A \rightarrow a)$$

Parse tree:



RMD:

$$S \Rightarrow aB$$

$$\Rightarrow aaBB \quad (B \rightarrow aBB)$$

$$\Rightarrow aabbs \quad (B \rightarrow bs)$$

$$\Rightarrow aaBbbA \quad (S \rightarrow bA)$$

$$\Rightarrow aaBbba \quad (A \rightarrow a)$$

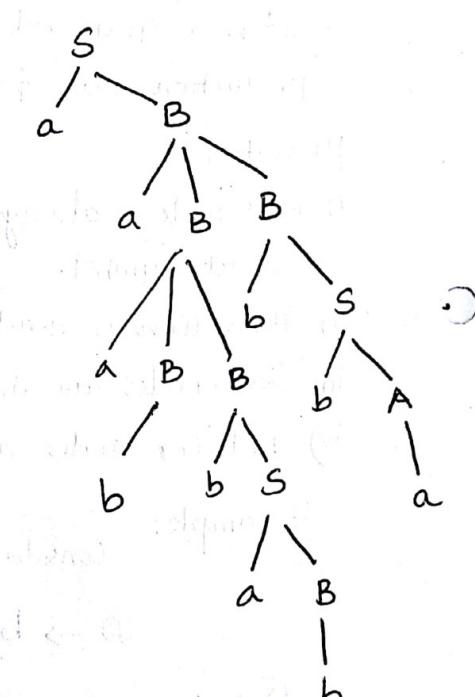
$$\Rightarrow aaabbBbba \quad (B \rightarrow aBB)$$

$$\Rightarrow aaaBbsbba \quad (B \rightarrow bs)$$

$$\Rightarrow aaaBbabBbba \quad (S \rightarrow aB)$$

$$\Rightarrow aaaBbabbbba \quad (B \rightarrow b)$$

$$\Rightarrow aaabbabbba \quad (B \rightarrow b)$$



Pbm 2: Write a grammar G_1 to recognize all prefix expressions involving all binary arithmetic operators. Construct the parse tree for the sentence ' $- * + abcde$ '.

Solution:

$$G_1 = (V, T, P, S) \text{ where}$$

$$V = \{S, A, B, C, D, E\}$$

$$T = \{+, -, *, /, a, b, c, d, e\}$$

S is a start symbol.

Productions are,

$$S \rightarrow A \mid B \mid C \mid D \mid E$$

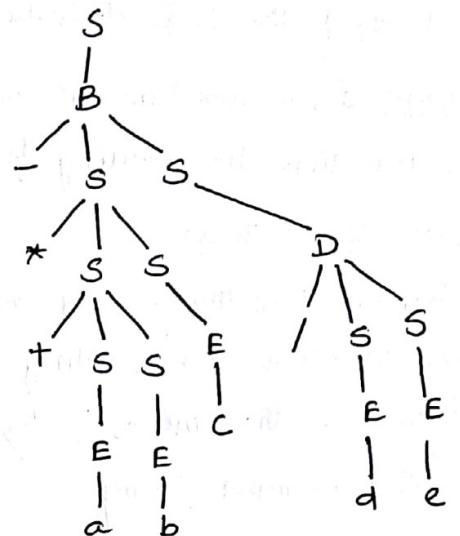
$$A \rightarrow +SS$$

$$B \rightarrow -SS$$

$$C \rightarrow *SS$$

$$D \rightarrow /SS$$

$$E \rightarrow ab \mid b \mid c \mid d \mid e$$



Ambiguity:

If there exists more than one parse trees for a given grammar, that means there could be more than one leftmost or rightmost derivation possible and then that grammar is said to be ambiguous grammar.

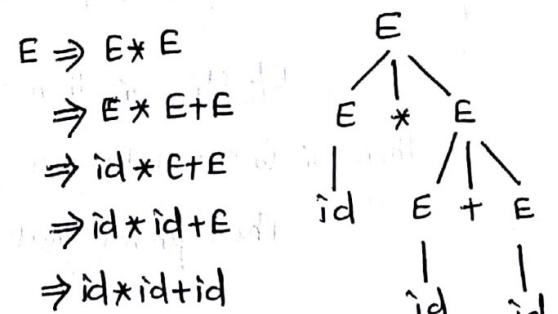
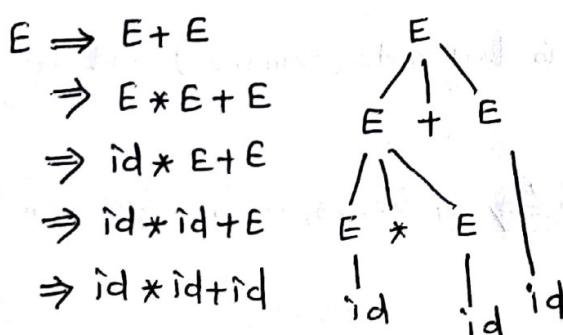
Problems-

1) The CFG is given by $G_1 = (V, T, P, S)$ where $V = \{E\}$, $T = \{id\}$, $S = \{E\}$

$P = \{E \rightarrow E+E, E \rightarrow E*E, E \rightarrow id\}$. Is the grammar ambiguous?

Solution:

Consider the string $id * id + id$.



Here, we obtain two different parse tree for the string $id * id + id$.

\therefore The given grammar is ambiguous.

Relationship between derivation and derivation trees.

Theorem: Let $G_1 = (V, T, P, S)$ be a context free grammar. Then $S \xrightarrow{*} a$ if and only if there is a derivation tree in grammar G_1 which gives the string a .

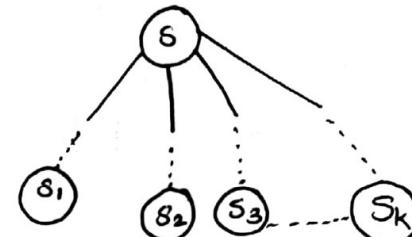
Proof: For a non-terminal S there exists $S \xrightarrow{*} w$ if and only if there is a derivation tree starting from root S and yielding w .

Basis of induction:

Assume that there is only one interior node S .

The derivation tree yielding $S_1, S_2, S_3, \dots, S_n$.

From S is that means $S \xrightarrow{*} S_1, S_2, \dots, S_n$
 $\xrightarrow{*} a$ is input string.



Induction hypothesis:

⇒ We assume that for $k-1$ nodes the derivation tree can be drawn. We then prove that for k vertices also we can have a derivation tree.

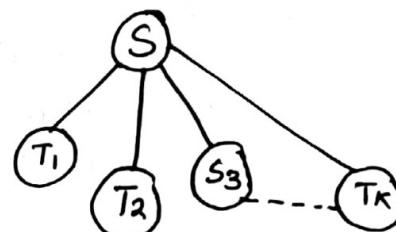
⇒ That means the input string can be derived as $S \rightarrow S_1 S_2 S_3 \dots S_k$.

⇒ There are two cases,

(i) S_i may be a leaf variable

(ii) S_i may be an interior node yielding a .

⇒ The S derives a by fewer number of k steps
 then $a \in S_1 S_2 S_3 S_4 \dots S_k$.



If $a_i = S_i$ then S_i is leaf node (terminal) and if $S_i \xrightarrow{*} a_i$; then S_i is an interior node.

This proves that $S \xrightarrow{*} S_1, S_2, S_3, \dots, S_n \xrightarrow{*} a$ can be obtained.

Simplification of CFG1.

Simplification of grammar means reduction of grammar by removing useless symbols.

- Elimination of useless symbols.
- Elimination of unit productions.
- Elimination of Null production(ϵ).

Elimination of useless symbols.

Any symbol is useful when it appears on the right hand side, in the production rule and generates some terminal string. If no such derivation exists then it is supposed to be an useless symbol.

Example: Eliminate the useless symbol from the following grammar.

$$S \rightarrow aS \mid A \mid C$$

$$A \rightarrow a$$

$$B \rightarrow aa$$

$$C \rightarrow aCb$$

Solution: Production with terminal symbols are

$$A \rightarrow a$$

$$B \rightarrow aa$$

$$\text{Start symbol } S \rightarrow aS \mid A \mid C$$

Here, there is no production for B.

$\therefore B$ is useless symbol.

Similarly

$$S \rightarrow C \rightarrow aCb \rightarrow aacbb \rightarrow aaaCbbb \rightarrow \dots$$

Here, no terminating symbol for C.

$\therefore C$ is useless symbol.

Eliminate B and C, we get

$S \rightarrow aS \mid A$
$A \rightarrow a$

Elimination of ϵ production:

(5)

If there is ϵ production, remove it, without changing the meaning of the grammar.

Example: Eliminate ϵ -productions from the CFG.

$$A \rightarrow OB|IB|I$$

$$B \rightarrow OB|IB|\epsilon$$

Solution:

$$A \rightarrow OB|IB|I$$

$$B \rightarrow \epsilon, A \rightarrow OI|II$$

$$B \rightarrow OB|IB$$

$$B \rightarrow \epsilon, B \rightarrow OI$$

After Elimination,

$$A \rightarrow OB|IB|OI|II$$

$$B \rightarrow OB|IB|OI$$

Removing Unit productions

The unit productions are the productions in which one non-terminal gives another non-terminal.

$$X \rightarrow Y$$

Example: Eliminate the unit production from following grammar.

$$S \rightarrow AB$$

$$A \rightarrow a$$

$$B \rightarrow C|b$$

$$C \rightarrow D$$

$$D \rightarrow E|bc$$

$$E \rightarrow d|Ab$$

Here,

$$B \rightarrow C$$

$$C \rightarrow D$$

$D \rightarrow E$ are unit productions.

$D \rightarrow E|bc$ can be written as $D \rightarrow d|Ab|bc$

by $C \rightarrow E|bc$, B becomes $B \rightarrow d|Ab|bc|b$.

After removing unit productions.

$$S \rightarrow AB$$

$$A \rightarrow a$$

$$B \rightarrow d|Ab|bc|b$$

$$C \rightarrow d|Ab|bc$$

$$D \rightarrow d|Ab|bc \quad X$$

$$E \rightarrow d|Ab \quad X$$

Chomsky normal form (CNF)

(6)

A context free grammar $G_1 = (V, T, P, S)$ is said to be in CNF if each production in G_1 is of the form

$$X \rightarrow YZ$$

$$\begin{bmatrix} NT \rightarrow NT \cdot NT \\ NT \rightarrow T \end{bmatrix}$$

$$X \rightarrow \alpha, \text{ where } X, Y, Z \in V, \text{ and } \alpha \in T^*$$

Problems:

1) Convert the given CFG₁ to CNF $S \rightarrow aSa | bSb | a | b$

Solution:

Productions are

$$S \rightarrow aSa$$

$$S \rightarrow bSb$$

$$S \rightarrow a$$

$$S \rightarrow b$$

Here, the productions which are already in CNF is

$$S \rightarrow a$$

$$S \rightarrow b$$

Apply CNF rule to other productions,

$$S \rightarrow aSa \quad | C_a \rightarrow a$$

$$S \rightarrow CaSCa$$

$$S \rightarrow CaA$$

$$A \rightarrow SCa$$

$$S \rightarrow bSb$$

$$S \rightarrow C_b SC_b$$

$$| C_b \rightarrow b$$

$$S \rightarrow C_b B$$

$$B \rightarrow SC_b$$

The resultant productions are,

$$S \rightarrow CaA | C_b B | a | b$$

$$A \rightarrow SCa$$

$$B \rightarrow SC_b$$

$$C_a \rightarrow a$$

$$C_b \rightarrow b$$

2) Reduce the following grammar to Chomsky normal form.

$$\begin{aligned} S &\rightarrow a \mid AAB \\ A &\rightarrow ab \mid aB \mid \epsilon \\ B &\rightarrow aba \mid \epsilon. \end{aligned}$$

Solution:

Productions are,

$$\begin{aligned} S &\rightarrow a \\ S &\rightarrow AAB \\ A &\rightarrow ab \\ A &\rightarrow aB \\ A &\rightarrow \epsilon \\ B &\rightarrow aba \\ B &\rightarrow \epsilon \end{aligned}$$

Productions already in CNF is

$$\begin{aligned} S &\rightarrow a \\ S &\rightarrow AA \\ S &\rightarrow AB \\ A &\rightarrow a \end{aligned}$$

Apply CNF rules to other productions,

$$S \rightarrow \underbrace{AAB}_S$$

$$\boxed{S \rightarrow SB}$$

$$\begin{aligned} S &\rightarrow aba \\ S &\rightarrow \underbrace{CaC_bC_a}_{C_1} \\ &\quad | \\ &\quad C_a \rightarrow a \\ &\quad C_b \rightarrow b \end{aligned}$$

$$\boxed{\begin{aligned} S &\rightarrow CaC_1 \\ C_1 &\rightarrow C_b C_a \end{aligned}}$$

$$\boxed{\begin{aligned} A &\rightarrow CaC_b \\ A &\rightarrow CaB \end{aligned}}$$

$$\begin{aligned} B &\rightarrow CaC_bC_a \\ &\quad | \\ &\quad C_1 \end{aligned}$$

$$\boxed{B \rightarrow CaC_1}$$

Here, Eliminate ϵ production.

$$A \rightarrow \epsilon \text{ and } B \rightarrow \epsilon$$

\Rightarrow After elimination, productions are

$$\begin{aligned} S &\rightarrow a \mid AAB \mid AA \mid AB \mid B \\ A &\rightarrow ab \mid aB \mid a \\ B &\rightarrow aba \end{aligned}$$

\Rightarrow Eliminate unit production. $\Rightarrow E$

$$\begin{aligned} S &\rightarrow a \mid AAB \mid AA \mid AB \mid aba \\ A &\rightarrow ab \mid aB \mid a \\ B &\rightarrow aba \end{aligned}$$

\therefore The resultant productions, are,

$$S \rightarrow SB \mid C_a C_1 \mid AA \mid AB \mid a$$

$$C_1 \rightarrow C_b C_a$$

$$A \rightarrow CaC_b \mid a \mid CaB \cancel{| C_a B |}$$

$$B \rightarrow CaC_1$$

3) Convert the given CFG1 to CNF.

$$\begin{array}{ll}
 S \rightarrow aB & A \rightarrow bAA \\
 S \rightarrow bA & B \rightarrow b \\
 A \rightarrow a & B \rightarrow bS \\
 A \rightarrow aS & B \rightarrow aBB
 \end{array}$$

Solution:

Productions already in CNF is,

$$\begin{array}{l}
 A \rightarrow a \\
 B \rightarrow b
 \end{array}$$

Apply CNF rules to other productions.

$$C_a \rightarrow a, C_b \rightarrow b$$

$$S \rightarrow C_a B$$

$$S \rightarrow C_b A$$

$$A \rightarrow C_a S$$

$$A \rightarrow C_b \underbrace{AA}_{C_1}$$

$$A \rightarrow C_b C_1$$

$$C_1 \rightarrow AA$$

$$B \rightarrow C_b S$$

$$B \rightarrow C_a \underbrace{BB}_{C_2}$$

$$B \rightarrow C_a C_2$$

$$C_2 \rightarrow BB$$

The resultant productions are,

$$S \rightarrow C_a B$$

$$S \rightarrow C_b A$$

$$A \rightarrow C_a S$$

$$A \rightarrow a$$

$$A \rightarrow C_b C_1$$

$$C_1 \rightarrow AA$$

$$B \rightarrow b$$

$$B \rightarrow C_b S$$

$$B \rightarrow C_a C_2$$

$$C_2 \rightarrow BB$$

4) Convert the grammar with productions into CNF

$$A \rightarrow bAB \mid \lambda, B \rightarrow BAa \mid \lambda$$

5) Convert the grammar $S \rightarrow AB \mid aB, A \rightarrow aab \mid \epsilon, B \rightarrow bba$ into CNF.

6) Convert to Chomsky Normal Form.

$$\begin{array}{lll}
 S \rightarrow A \mid CB & B \rightarrow 1B \mid 1 & D \rightarrow 2D \mid 2 \\
 A \rightarrow C \mid D & C \rightarrow 0C \mid 0 &
 \end{array}$$

Greibach Normal Form (GNF)

A grammar $G = (V, T, P, S)$ is said to be in GNF if every production rule is of the form.

$$X \rightarrow a\alpha$$

where $a \in T$, $X \in V$

$$\alpha \in V^*$$

Right hand side of every production starts with a terminal, followed by a string of variables of zero/more length.

Problems:

- Convert the given CFG to GNF

$$S \rightarrow ABA$$

$$A \rightarrow aA | \epsilon$$

$$B \rightarrow bB | \epsilon.$$

Solution:

Simplify the CFG, Eliminate ϵ production $A \rightarrow \epsilon, B \rightarrow \epsilon$.

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

$$A \rightarrow aA | a$$

$$B \rightarrow bB | b$$

Eliminate unit productions,

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

$$A \rightarrow aA | a$$

$$B \rightarrow bB | b.$$

Apply GNF rules,

$$S \rightarrow \underline{ABA}$$

$$S \rightarrow aABA | aBA$$

$$S \rightarrow \underline{AB}$$

$$S \rightarrow aAB | aB$$

$$S \rightarrow BA$$

$$S \rightarrow bBA | bA$$

$$S \rightarrow AA$$

$$S \rightarrow aAA | aA$$

∴ The resultant productions are,

$$S \rightarrow aABA | aBA | aAB | aB | bBA | bA$$

$$S \rightarrow aAA | aA | a | b$$

$$A \rightarrow aA | a$$

$$B \rightarrow bB | b$$

2) Convert given CFG₁ to G_{NF} where $V = \{S, A\}$, $T = \{0, 1\}$ and P is,

$$S \rightarrow AA | 0$$

$$A \rightarrow SS | 1$$

Solution:

Replace S as A₁ and A as A₂

CFG₁ becomes, $A_1 \rightarrow A_2 A_2 | 0$

$$A_2 \rightarrow A_1 A_1 | 1$$

start with $A_2 \rightarrow A_1 A_1 | 1$

$$\underbrace{A_2 \rightarrow A_2 A_2 A_1}_{\text{left recursion.}} | 0 A_1 | 1$$

Introduce B₂,
to eliminate left recursion } $\Rightarrow A_2 \rightarrow A_2 A_2 A_1, A_2 \rightarrow 0 A_1 | 1$
 $B_2 \rightarrow A_2 A_1 | A_2 A_1 B_2, A_2 \rightarrow 0 A_1 B_2 | 1 B_2$

The productions are

$$\begin{aligned} A_2 &\rightarrow 0 A_1 | 1 \\ A_2 &\rightarrow 0 A_1 B_2 | 1 B_2 \end{aligned} \Rightarrow A_1 \rightarrow A_2 A_2 | 0$$

$$A_1 \rightarrow 0 A_1 A_2 | 1 A_2 | 0 A_1 B_2 A_2 | 1 B_2 A_2 \quad (\text{G}_N\text{F})$$

$$B_2 \rightarrow A_2 A_1$$

$$B_2 \rightarrow 0 A_1 | 1 A_1 | 0 A_1 B_2 A_1 | 1 B_2 A_1 \quad (\text{G}_N\text{F})$$

$$B_2 \rightarrow A_2 A_1 B_2$$

$$B_2 \rightarrow 0 A_1 A_1 B_2 | 1 A_1 B_2 | 0 A_1 B_2 A_1 B_2 | 1 B_2 A_1 B_2$$

Converting Back $A_1 = S$ and $A_2 = A$ (G_NF)

$$S \rightarrow 0 S A | 1 A | 0 S B_2 A | 1 B_2 A | 0$$

$$A \rightarrow 0 S | 1 | 0 S B_2 | 1 B_2$$

$$B_2 \rightarrow 0 S S | 1 S | 0 S B_2 S | 1 B_2 S$$

$$B_2 \rightarrow 0 S S B_2 | 1 S B_2 | 0 S B_2 S B_2 | 1 B_2 S B_2$$

3) Convert the grammar $S \rightarrow AB$, $A \rightarrow BS|b$, $B \rightarrow SA|a$ into Greibach Normal form. ⑨

Solution:

Consider the grammar,

$$S \rightarrow AB, A \rightarrow BS|b, B \rightarrow SA|a$$

Assume $S = A_1$, $A = A_2$, and $B = A_3$.

$$A_1 \rightarrow A_2 A_3$$

$$A_2 \rightarrow A_3 A_1 | b$$

$$A_3 \rightarrow A_1 A_2 | a$$

Consider,

$$A_3 \rightarrow A_1 A_2 | a$$

$$A_3 \rightarrow A_2 A_3 A_2 | a$$

$$A_3 \rightarrow A_3 A_1 A_3 A_2 | b A_3 A_2 | a$$

Now,

$$\underbrace{A_3 \rightarrow A_3}_{\text{Introduce } B_3 \text{ to eliminate left recursion}} A_1 A_3 A_2, A_3 \rightarrow b A_3 A_2 | a$$

$$\left. \begin{array}{l} \text{Introduce } B_3 \text{ to} \\ \text{eliminate left recursion} \end{array} \right\} \Rightarrow B_3 \rightarrow A_1 A_3 A_2 | A_1 A_3 A_2 B_3$$
$$A_3 \rightarrow b A_3 A_2 | a | b A_3 A_2 B_3 | a B_3. (G_{NF})$$

$$A_2 \rightarrow A_3 A_1 | b$$

$$A_2 \rightarrow b A_3 A_2 A_1 | a A_1 | b A_3 A_2 B_3 A_1 | a B_3 A_1 | b (G_{NF})$$

$$A_1 \rightarrow A_2 A_3$$

$$A_1 \rightarrow b A_3 A_2 A_1 A_3 | a A_1 A_3 | b A_3 A_2 B_3 A_1 A_3 | a B_3 A_1 A_3 | b A_2$$

$$B_3 \rightarrow A_1 A_3 A_2$$

$$\rightarrow b A_3 A_2 A_1 A_3 A_3 A_2 | a A_1 A_3 A_3 A_2 | b A_3 A_2 B_3 A_1 A_3 A_3 A_2 | a B_3 A_1 A_3 A_3 A_2 | b A_3 A_2 A_1 A_3 A_2.$$

$$B_3 \rightarrow A_1 A_3 A_2 B_3$$

$$\rightarrow b A_3 A_2 A_1 A_3 A_3 A_2 B_3 | a A_1 A_3 A_3 A_2 B_3 | b A_3 A_2 B_3 A_1 A_3 A_3 \\ A_2 B_3 | \\ a B_3 A_1 A_3 A_2 B_3 | b A_3 A_3 A_2 B_3 .$$

∴ The resultant productions are,

$$A_1 \rightarrow b A_3 A_2 A_1 A_3 | a A_1 A_3 | b A_3 A_2 B_3 A_1 A_3 | b A_3$$

$$A_2 \rightarrow b A_3 A_2 A_1 | a A_1 | b A_3 A_2 B_3 A_1 | a B_3 A_1 | b$$

$$A_3 \rightarrow b A_3 A_2 | a | b A_3 A_2 B_3 | a B_3$$

$$B_3 \rightarrow b A_3 A_3 A_2 b A_3 A_2 A_1 A_3 A_2 | a A_1 A_3 A_3 A_2 |$$

$$b A_3 A_2 B_3 A_1 A_3 A_3 A_2 | a B_3 A_1 A_3 A_3 A_2 |$$

$$b A_3 A_2 A_1 A_3 A_3 A_2 B_3 | a A_1 A_3 A_3 A_2 B_3 |$$

$$b A_3 A_2 B_3 A_1 A_3 A_3 A_2 B_3 | a B_3 A_1 A_3 A_2 B_3 | b A_3 A_3 A_2 B_3 .$$