* Simplification of context free Grammars:-

Context free grammars can be generally represented in two different normal forms. To make it happen it is initially necessary to simplify or reduce the context-free grammars. This process is also called "Reduction of Context Free Grammars."

This is achieved in 3 different ways

1) Removal of Unit productions
2) Removal of Useless Symbols
3) Removal of ϵ-productions.

1) **Removal of Unit productions:-** Unit productions are defined as the rules which are in the form of $V \to T$ (or) $V \to V$

Eg:- $S \to Aa/B/C$

$B \to b$

$C \to aA/b$

$A \to b$

Here
$$\left.\begin{array}{l} S \to B \\ S \to C \\ B \to b \\ C \to b \\ A \to b \end{array}\right\} \text{are Unit productions}$$

The grammar must be rewritten by removing unit productions

$S \to Aa/b/b$

$C \to aA$

$A \to b$

(Here $A \to b$ is not removed as its is removed derivation will not be possible)

$\Rightarrow S \to Aa/b$

$C \to aA$

$A \to b$

Eg: $S \to Aa/B$
$A \to b/B$
$B \to A/a$

Sol: $S \to \cancel{B}a$
$A \to b$
$A \to \cancel{B}a$
$B \to A$
$B \to a$
} unit productions.

$$\boxed{\begin{array}{l} S \to Aa/a \\ A \to b/a \end{array}}$$

Removal of ϵ productions:- In a given context free grammar if there is any production rule in the form $X \to \epsilon$ then it should be removed by proper substitution in the production rules.

Eg: $S \to ABCd$
$A \to BC$
$B \to bB/\epsilon$
$C \to cC/\epsilon$

Sol: $B \to \epsilon$
$C \to \epsilon$

$S \to ABCd/ACd/ABd/Ad$
$A \to BC/C/B$
$B \to bB/b$
$C \to cC/c$

Eg: $A \to ABSxyabc/*/-/\epsilon$
$B \to dB/\epsilon$
$S \to \epsilon$
$X \to xy/\epsilon$
$Y \to \epsilon$

Sol. $A \to \epsilon$

$B \to \epsilon$

$S \to \epsilon$

$X \to \epsilon$

$Y \to \epsilon$

$A \to ABSXYabc \mid * \mid - \mid BSXYabc \mid ASXYabc \mid ABXYabc \mid SXYabc \mid BXYabc \mid ABSYabc \mid ABSXabc \mid BSYabc \mid BSXabc \mid AXYabc \mid ASYabc \mid ASXabc \mid AByabc \mid ABXabc \mid ABSabc \mid XYabc \mid S \ldots$

$B \to dB \mid d$

~~$S \to \epsilon$~~

$X \to xy$

~~$Y \to \epsilon$~~

* Normal forms of context free Grammar :-

A grammar is said to be context free if it exists in any one of the following normal forms.

1) Chomsky Normal form
2) Griebach Normal form

1) Chomsky Normal form:-

A Grammar $G = (V, T, S, P)$ is in chomsky norfmal form if all the production rules exists in any one of the 2 forms $A \to a$ (or) $A \to BC$. Where ~~A is~~ A, B, C are non-terminals & a is a terminal.

To convert a given grammar into CNF the following steps are to be followed:-

1) Simplify the given Context free grammar by removing NULL productions, removing unit productions & removing useless symbols.

2) The resultant context free grammar is to be checked for the CNF rule (or) lemma and if it matches then ~~it is~~ the CFG is in CNF form.

3) If the production rules doesnot match the CNF lemma then Convert the production rules to the appropriate lemma rules by transforming to new non-terminals.

Eg: Convert the given Context Free Grammar (CFG) into Chomsky Normal form (CNF)

$S \rightarrow bA / aB$

$A \rightarrow bAA / aS / a$

$B \rightarrow aBB / bS / b$

Sol: The Given CFG

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

$V: \{S, A, B\}$

$T: \{a, b\}$

$S: \{S\}$

The Grammar 'G' is free from ϵ-productions, Unit productions, useless Symbols.

If we check with CNF lemma rules.

$A \rightarrow a$ (or) $A \rightarrow BC$

then only the following production rules of G matches with lemma rules

$A \rightarrow a$

$B \rightarrow b$

$\underline{S \rightarrow bA}$:

$X \rightarrow b$ where $X \in V$.

$\underline{S \rightarrow aB}$

$Y \rightarrow a$ where $Y \in V$

$\underline{A \rightarrow bAA}$

$Z \rightarrow bA$ where $Z \in V$

$Z \rightarrow XA$, $X \in V$.

$\underline{A \rightarrow aS}$

$A \rightarrow YS$

$\underline{B \rightarrow aBB}$

$W \rightarrow aB$

$W \rightarrow YB$ where $W \in V$

$\underline{B \rightarrow bS}$

$B \rightarrow XS$

The CNF of CFG is

$S \to XA / YB$

$A \to ZA / YS / a$

$B \to WB / XS / b$

$Z \to XA$

$W \to YB$

$X \to b$

$Y \to a$

The Grammar $G : (V, T, P, S)$

$V : \{S, A, B, Z, W, X, Y\}$

$T : \{a, b\}$

$S : \{S\}$

Eg: Convert the given CFG into CNF

$S \to aAD$

$A \to aB / bAB$

$B \to b$

$D \to d$

Sol: $G : (V, T, P, S)$

$V : \{S, A, B, D\}$

$T : \{a, b, d\}$

$S : \{S\}$

The production rules of G matches with lemma rules

$B \to b$

$D \to d$

$S \to aAD$:

$X \to aA$

$X \to YA$ & $Y \to a$

$A \to aB$

$A \to YB$

$A \to bAB$

$Z \to bA$

$Z \to WA$ & $W \to b$.

The CNF of CFG is

$S \to XD$

$A \to YB / ZB$

$B \to b$

$D \to d$

$Y \to a$

$W \to b$

$Z \to WA$

$X \to YA$

$G : (V, T, P, S)$

$V : \{S, A, B, D, Y, W, Z, X\}$

$T : \{a, b\}$

$S : \{S\}$

Griebach Normal Form (GNF):-

A Grammar $G:(V,T,P,S)$ is said to be in Griebach Normal form if all the productions in P are of the form $A \in V$ & $X \in V^*$ & $a \in T$.

$A \rightarrow aX$ (or) $A \rightarrow a$ where

If the production rules are not in the GNF form then any one of the following lemma is applicable.

Lemma ①:- Let G is a context free grammar & $A \rightarrow B\alpha$ is in 'P' & if $B \rightarrow \beta_1/\beta_2/\beta_3, \ldots /\beta_n$ then A-production rule can be defined as $\boxed{A \rightarrow \beta_i \alpha}$ where $i>0$

Lemma ②:- If G is a context free grammar & let the production rule $A \rightarrow A\alpha_1/A\alpha_2/\ldots/A\alpha_n/\beta_1/\beta_2/\ldots/\beta_n$ then a new non-terminal Z is included & the grammar G is defined as $G:\{(V \cup Z), T, P', S\}$. where production P1 is given as

$P_1: A \rightarrow \beta_1/\beta_2/\ldots/\beta_n.$

$A \rightarrow \beta_1 Z/\beta_2 Z/\ldots/\beta_n Z$

~~In GNF~~

$Z \rightarrow \alpha_1/\alpha_2/\ldots/\alpha_n$

$Z \rightarrow \alpha_1 Z/\alpha_2 Z/\ldots/\alpha_n Z.$

Eg: Convert the given CFG into Griebach Normal form (GNF)

$S \rightarrow AA/a$

$A \rightarrow SS/b$

Sol:

Note:- In GNF rename the non-terminals as $A_1, A_2, \ldots A_n$ & for production rule $A_i \rightarrow A_j\alpha$. If $i \leq j$ then apply the lemma rules else convert the production rules so that it satisfies the condition $i \leq j$.

Sol: Renaming the non-terminals as A_1, A_2

$S = A_1, \; A = A_2$

$\therefore P: A_1 \rightarrow A_2A_2/a \rightarrow ①$

$A_2 \rightarrow A_1A_1/b \rightarrow ②$

~~from ② lemma ①~~ from ① $A_1 \rightarrow A_2A_2$

$A_1 \rightarrow a$

from lemma ① Substitute A_2 in above step

$A_1 \to A_1 A_1 A_2 / bA_2 / a \to$ ③

$A \to A\alpha / \beta$

from lemma ② ie $A \to A\alpha/\beta$ then

$A = A_1$, $\alpha = A_1 A_2$, $\beta = bA_2/a$

from ③ by applying lemma ②,

$A_1 \to bA_2 / a \to$ ④

$A_1 \to bA_2 Z / aZ \to$ ⑤

$Z \to A_1 A_2 \to$ ⑥

$Z \to A_1 A_2 Z \to$ ⑦

from ⑥ & ⑦, ④ & ⑤

$Z \to aA_2 / aA_2 Z$

$A_1 \to bA_2 / a / bA_2 Z / aZ$ $\to$ GNF of A_1.

from ②,

$A_2 \to A_1 A_1 / b$

from GNF form of A_1

$A_2 \to (bA_2 / a / bA_2 Z / aZ)\, A_1 / b$

$\boxed{A_2 \to bA_2 A_1 / aA_1 / bA_2 Z A_1 / aZ A_1 / b}$

∴ The GNF of G is given as

$G : \{(V \cup Z), T, P_1, S\}$

$P_1 : A_1 \to bA_2 / a / bA_2 Z / aZ$

$A_2 \to bA_2 A_1 / aA_1 / bA_2 Z A_1 / aZ A_1 / b$.

$Z \to aA_2 / aA_2 Z$.

Eg: Convert CFG into GNF.

$S \to AB$
$A \to BS/b$
$B \to SA/a$

* **Pumping lemma for Context free language:-**

Pumping lemma is used for proving that a language 'L' is not Context free.

Consider, a language 'A' as context free, then A has a pumping length 'p' such that any String 'S' exists where $|S| \geq p$ may be divided into 5 pieces UVXYZ Such that the following conditions must be true:-

1) uv^ixy^iz is in A for $i \geq 0$.
2) $|vy| > 0$
3) $|vxy| \leq p$

Example:- prove that the language $L = \{a^nb^nc^n / n>0\}$ is not context free.

Sol: Given $L = \{a^nb^nc^n / n>0\}$

1) Consider L is Context free.
2) let pumping length = s.
   Such that $|S| \geq p$ ie, $L = \{a^pb^pc^p\}$
3) let $S = a^5b^5c^5$ for $p=5$.
   $|S|$ = aaaaabbbbbccccc = 15
   $|S| = 15 \geq 5$

4) Now consider $\boxed{P=7}$
   check $|S| \geq P$
   $15 \geq 7$

4) Divide S into UVXYZ
   S = aaaaabbbbbccccc

V - Same symbol (a,b,c)
V - Different Symbols (ab,bc,abc)
Y - Same symbol (a,b,c)
Y - Different Symbol (ab, bc, abc)

Rule① $S = uv^i xy^i z$ for $i=2$
$S = aaaaabbbbbccccc$

$u = aaa$, $v = aa$, $x = bbbb$, $y = b$, $z = ccccc$

$S = uv^i xy^i z$ is in A

for $i=2$, $S = uv^i xy^i z$ is $uv^2 xy^2 z$

$= aaa\,aaaa\,bbbbbb\,ccccc$

$= a^7 b^6 c^5 \notin L$.

Given $L = \{a^n b^n c^n \mid n \geq 0\}$ is not CFL.

1) $L = \{ww \mid w \in \{0,1\}^*\}$
2) $L = \{(ab)^m c^n \mid m > n;\ m,n \geq 0\}$
3) $L = \{wcw \mid w \in \{0,1\}^*\}$
4) $L = \{wcw^R \mid w \in \{0,1\}^*\}$

Properties of Context free language:

CFL is closed under 1) union ($L_1 \cup L_2$)
2) concatenation ($L_1 \cap L_2$)
3) Kleen closure ($L_1{}^*$)

consider L_1 & L_2 be two context free languages then $L_1 \cup L_2$ is a CFL, $L_1 \cap L_2$ is a CFL & L_1^* (or) L_2^* is also a CFL.

Applications of CFL's :-

1) CFL's are applicable in the design of compilers especially gcc.
2) parser Statements
3) Debuggers
4) High level Structured programs.