22/8/23

U2

Grammar - Introduction

* Context Free Grammar and denoted by
$G = (V, T, P, S)$

* The main objective is to frame the rules for certain languages.

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

V ⇒ Variables or non terminal

T ⇒ Terminals or small letters or special characters

P ⇒ Production rule

S ⇒ Starting symbol

* The language generated by CFG is CFL (context Free language).

1) Let Grammar $G = \{ \{S\}, \{a, b\}, P, S \}$

$$P = (S \to aSa$$
$$S \to bSb$$
$$S \to \epsilon)$$

Find the lang. generated by this grammar.

$W = aa$.

I. $S \rightarrow aSa$

$S \rightarrow a\epsilon a$ $[S \rightarrow \epsilon]$

$\rightarrow aa$

II.

$S \rightarrow aSa$

$S \rightarrow abSba$

$S \rightarrow ab\epsilon ba$

$\rightarrow \underbrace{ab}_{w}\underbrace{ba}_{w^R}$

$L = \{ww^R / w \in \{a,b\}^*\}$

Types of grammar :-

Type ③

it generates reg. lang. & it is accepted by FA. & type ③ grammar must have a single non-terminal on LHS.

T-3 $X \rightarrow a$

$X \rightarrow aY$

$X, Y \in N(V)$

$a \in T$

$(T-1)$ (accepted by LBA)

$(T-2)$

$A \to \gamma$
$A \in N$
$\gamma \in (T \cup N)^*$

$$\boxed{\alpha A \beta \to \alpha \gamma \beta}$$

$A \in N,$
$\alpha, \beta, \gamma \in (T \cup N)^*$

α & β $\to$ maybe empty

γ $\to$ must not be empty

$(T-0)$

$$\boxed{\alpha \to \beta}$$

-> It is a phase structured grammar or recursively enumerable languages.

$\alpha \to$ terminals and non-terminals with atleast one non-terminal and 'α' can't be null.

$\beta \to$ Both terminals and non-terminals.

$T3 \leq T2 \leq T1 \leq T1$

Derivations and languages :-

1. recursive inference
2. Derivations.

Derivation types:

1. Left most derivations (LMD)
2. right most derivations (RMD)

* (LMD)

→ ~~It is~~ each step in derivation, the left most variable is replaced by it's production.

→ represented by :

$$\underset{lm}{\Longrightarrow} \quad (or) \quad \overset{*}{\underset{lm}{\Longrightarrow}}$$

Ex:

$S \to aAS/a$

$A \to ab$

$w = aaba$

& $S \to aAS$

$\underset{lm}{\Rightarrow} aabS$ [$A \to ab$]

$\underset{lm}{\Rightarrow} aaba$ [$S \to a$]

* RMD

$S \to aAS$

$S \to aAa$ [$S \to a$]

$S \to aaba$ [$A \to ab$]

"aabbabba"

$S \to aB / bA$

$A \to a / aS / bAA$

$B \to b / bS / aBB$

LMD

$S \to aB$ ~~bA~~

$S \underset{lm}{\Rightarrow} aaBB$ [$B \to aBB$]

$S \underset{lm}{\Rightarrow} aabB$ [$B \to b$]

$S \underset{lm}{\Rightarrow} aabbS$ [$B \to bS$]

$S \underset{lm}{\Rightarrow} aabbS$ [$B \to bS$]

$S \underset{lm}{\Rightarrow} aabbaB$ [$S \to aB$]

$S \underset{lm}{\Rightarrow} aabbabS$ [$B \to bS$]

$S \underset{lm}{\Rightarrow} aabbabbA$ [$S \to bA$]

$S \underset{lm}{\Rightarrow} aabbabba$ [$A \to a$]

$\Rightarrow$

RMD

aabbabba

$S \to a\underline{B}$

$S \underset{rmd}{\Rightarrow} aaB\underline{B}$

$S \underset{rmd}{\Rightarrow} aaBbb\underline{A}$

$S \underset{rm}{\Rightarrow} aa\underline{B}bba$

$S \underset{rm}{\Rightarrow} aab\underline{S}bba$

$S \underset{rm}{\Rightarrow} aabb\underline{A}bba$

$S \underset{rm}{\Rightarrow} aabbabba$

Derive the string 1000111 for left & rightmost derivation using CFG

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

$V = \{S, T\}$

$T = \{0, 1\}$

$S, P = \{ S \to T00T$

$T \to 0T / \; 1T$

Ambiguity

Ambiguous Grammar

$E \to E+E \mid E*E \mid id.$

Construct ambiguous grammar for the following

$E \to I$ $\quad E \to E*E$

$E \to E+E$ $\quad E \to (E)$

$I \to 0, 1, 2, \ldots 9$

$w = 3*2+5$

LMD:

$E \to E * E$

$E \to I * E \quad (E \to I)$

$E \to I * E+E \quad (E \to E+E)$

$E \to I * I + E \quad (E \to I)$

$E \to I * I + I \quad (E \to I)$

$E \to E + E$

$E \to E * E + E \quad (E \to E * E)$

$E \to I * E + E \quad (E \to I)$

$E \to I * I + E \quad (E \to I)$

$E \to I * I + I \quad (E \to I)$

$E \to 3 * I + I \quad (I \to 3)$

$E \to 3 * 2 + I \quad (I \to 2)$

$E \to 3 * 2 + 5 \quad (I \to 5)$

Simplication of CFG

(i) ↳ Eliminating useless production

(ii) ↳ Removing / eliminating the epsilon production

(iii) ↳ Removing / eliminating the unit production.

The productions from a grammar that can never take part in any derivations.

Ex: $S \to aSbS \,/\, C \,/\, \epsilon$

$C \to aC$

Solutions) $S \to aSbS \,/\, \epsilon$

Q) eliminate the useless symbols and useless production are from $G = (V, T, P, S)$ where

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

$T = \{a, b\}$

P :

$S \rightarrow aS / A / C$

$A \rightarrow a$

$B \rightarrow aa$

$C \rightarrow aC$

Solution:

$S \rightarrow aS / A / a / aa$

(or)

$S \rightarrow aS / A / a / A / B / aa$

(ii) Any production of CFG in the form of

$A \rightarrow e$ is called epsilon productions

Q) $\cancel{S \rightarrow aSbS / e}$

1.) $S \rightarrow abB$

$B \rightarrow Bb / e$

I $V_{nullable} = \{B\}$

II $S \rightarrow abB / ab$

$B \rightarrow Bb$

Q) $S \rightarrow BabC$

$C \rightarrow b / e$

$B \rightarrow a / e$

I $V_{nullable} = \{B, C\}$

II $S \to BabC \mid abC \mid Bab \mid ab$

$C \to b$

$B \to a$

(iii) Unit Productions: (Eliminating / Removing)

A production in a CFG in the form of

$A \to B$

where

$A, B \to$ variable (V)

Q) $S \to AaB \mid C$

$A \to a \mid bC \mid B$

$C \to a$

$B \to A \mid bb$

I. $S \to C$

$A \to B$

$B \to A$

II.

$S \to AaB \mid a$

$A \to a \mid bC \mid bb$

$C \to a$

$B \to a \mid bC \mid bb$

Q) unit production [remove all the useless production/and all ϵ)

$S \to aA / aBB$

$A \to aaA / \epsilon$

$B \to bB / bbC$

$C \to B$

I. eliminating ϵ production
$V_{nullable} = \{A\}$

$S \to a / aBB / a$

$A \to aaA / aa$

$B \to bB / bbC$

$C \to B$

II. $S \to aA / aBB / a$
$A \to aaA / aa$
$B \to bB / bbC$
$C \to bB / bbC$

iii removing productions:

$S \to aA / a$

$A \to aaA / aa$

CNF : Chomsky Normal Form.

N.T → NT. NT

N.T → T

Reducing CFG into CNF :-

To reduce a CFG into CNF, there are some following steps.

step ① i) eliminate ϵ (iii) " useless production.
(ii) " unit

Q)

S → AAC

A → aAb/e

C → aC/a

convert into CNF.

I V-nullable = {A}

S → AAC / AC / C

A → aAb/ab

C → aC/a

II removing unit production S → C

S → AAC / AC / aC / a

A → aAb/ab

C → aC/a

$\boxed{C \to aC/a}$

III There is no useless production.

IV

P1 : S → AAC is not a CNF

S → AD,

D1 → AC

$\boxed{\begin{array}{l} A \to BC \\ A \to a \end{array}}$

P2 : S → AC

P_3 :
$S \rightarrow aC$

$S \rightarrow D_2 C$

$D_2 \rightarrow a$

P_4 : $A \rightarrow aAb$

$A \rightarrow D_2 Ab$ $\qquad A \rightarrow D_2 D_3 \mid D_2 D_4$

$D_3 \rightarrow A D_4$ $\qquad D_3 \rightarrow Ab$

$\qquad D_4 \rightarrow b$

$D_4 \rightarrow b$

P_5 : $C \rightarrow aC \mid a$

$C \rightarrow D_2 C \mid a$

Q) consider a CFG

CFG $G = (\{S, A\}, \{a, b\}, P, S\}$

$S \rightarrow aAS \mid a$

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

Sol:

$S \to aAS$

$\cancel{(i)\ R_1}$

(i) $M_1 \to a$

$S \to M_1AS$

$S \to M_1M_2$

$M_2 \to AS$

(ii) $S \to a$

(iii) $A \to SbA$

$M_3 \to b$

$A \to SM_3A$

$A \to SM_4$

$M_4 \to M_3A$

(iv) $A \to SS$

(v) $A \to ba$

$A \to M_3M_1$.

CFG to CNF.

Q) $S \to ABA$

$A \to aA / \epsilon$

$B \to bB / \epsilon$

Step ①

(i) eliminate ϵ

$S \to ABA \mid BA / AA / AB / A / B$

$A \to aA / a$

$B \to bB / b$

(ii) eliminate unit production!

$\boxed{\begin{array}{l} S \to A \\ S \to B \end{array}}$

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

$A \to aA \mid a$

$B \to bB \mid b$

(iii) No useless production.

step ② CNF (for S)

(i) $S \to ABA$

$M_1 \to BA$

$\boxed{S \to AM_1}$

(ii) $S \to BA$

(iii) $S \to AA$

(iv) $S \to AB$

(v) $S \to aA$

$M_2 \to a$

$\boxed{S \to M_2A}$

(vi) $S \to a$

(vii) $S \to bB$

$M_3 \to b$

$\boxed{S \to M_3B}$

(viii) $S \to b$

(for B)

(i) $B \to bB$

$\boxed{B \to M_3B}$

(ii) $B \to b$

(for A)

(i) $A \to aA$

$\boxed{A \to M_2A}$

(ii) $A \to a$

CNF all

$S \rightarrow AM_1 \mid AB \mid BA \mid AA \mid M_2 A \mid M_3 B \mid a \mid b$

$A \rightarrow M_2 A \mid a$

$B \rightarrow M_3 B \mid b$

$M_1 \rightarrow BA$

$M_2 \rightarrow a$

$M_3 \rightarrow b.$

GNF ⇒ Greibach Normal Forms

a grammar g is (V, T, P, S) is said to be a GNF, if every production rule is in the form of $\boxed{A \rightarrow a\alpha}$

where $a \rightarrow$ terminal

$\alpha \in (V \cup T)^*$

Lemma ① Substitution Rule.

Lemma ② Elimination of Left Recursion.

Lemma ①

$$\boxed{A \rightarrow a_1 B_1 \mid a_2 B_2 \mid \dots \dots a_n \beta_n}$$

Lemma ②

(i) $A \to \beta_i$ $\left\{ \begin{array}{l} A \to \beta_1 / \beta_2 / \dots \beta_n \\ A \to \beta_1 X / \beta_2 X / \dots \beta_n X \end{array} \right.$

$A \to \beta_i X$

$(i \leq 1 \leq n)$

(ii) $X \to \alpha_i$ $\left\{ \begin{array}{l} X \to \alpha_1 / \alpha_2 / \dots \alpha_n \\ X \to \alpha_1 X \end{array} \right.$

$X \to \alpha_i X$

$(i \leq i \leq m)$

GNF Rules :-

(1) Check wether it is CNF

(1) {if not eliminate the ϵ, unit, useless}

(2) Rename the variables

Ex: $A_1, A_2 \dots$ with $S = A_1$

(3) $A_i \to A_j \alpha$ apply the following

(i) $i < j \to$ leave the prdn's as it is

(ii) $i > j \to$ substitution rule

(iii) $i = j \to$ eliminating of L.R.

(4) $A_i \to A_j \alpha$

where

$i < j$ {apply Sub. rule if A_j is in GNF}

Q) GNF for the "G"

$S \to AA \mid a$

$A \to SS \mid b$

Sol: (i) Given Gr is CNF

(ii) Rename the variable

$S \to A_1$ & $A \to A_2$

$A_1 \to A_2 A_2 \mid a$

$A_2 \to A_1 A_1 \mid b$

(iii) check i & j

$A_1 \to A_2 A_2 \mid a$ $\{ i < j$ – same as it is

$\left[\begin{matrix} i = 1 \\ j = 2 \end{matrix} \right]$

$A_2 \to A_1 A_1 \mid b$ $\{ i = 2, j = 1$; $i > j$

applying Substitution rule ;

$A_2 \to A_1 A_1 \mid b$

$A_2 \to \underbrace{A_2}_{A} \underbrace{A_2 A_1}_{\alpha} \mid \underbrace{a A_1}_{B_1} \mid \underbrace{b}_{B_2}$ } not in GNF

$i = 2, j = 2 \to$ Apply lemma 2.

(iv) $A_2 \rightarrow aA_1 / b$

$A_2 \rightarrow aA_1 Z_2 / bZ_2$

$Z_2 \rightarrow A_2 A_1$ { α terms }

$Z_2 \rightarrow A_2 A_1 Z_2$

$\boxed{A_2 \rightarrow aA_1 \mid b \mid aA_1 Z_2 \mid bZ_2}$ GNF

(v) $Z_2 \rightarrow A_2 A_1$

$Z_2 \rightarrow \boxed{aA_1A_1 \mid bA_1 \mid aA_1Z_2A_1 \mid bZ_2A_1}$ GNF

(vi) $Z_2 \rightarrow A_2 A_1 Z_2$

$\boxed{Z_2 \rightarrow aA_1A_1Z_2 \mid bA_1Z_2 \mid aA_1Z_2A_1Z_2 \mid bZ_2A_1Z_2}$ GNF

(vii) $A_1 \rightarrow A_2 A_2 \mid a$

$\boxed{A_1 \rightarrow aA_1A_2 \mid bA_2 \mid aA_1Z_2A_2 \mid bZ_2A_2 \mid a}$

GNF.

GNF depends on two lemmas:-

Lemma ① : Substitution Rule!

Let $G = (V, T, P, S)$ be a CFG. Then $A \to a_1 B$ is a production in G. Where $B \to B_1 | B_2 \ldots$

Thus we can obtain $G_1 = (V_1, T_1, P_1, S)$

where P_1 consists of

$$\boxed{A \to a_1 \beta_1 | a_2 \beta_2 | \ldots a_n \beta_n}$$

Lemma ② : Elimination of left Recursion!

$G = (V, T, P, S)$ be a CFG, is p consists of

$A \to A\alpha_1 | A\alpha_2 | \ldots | A\alpha_m | \beta_1 | \beta_2 | \beta_n$

Introduce a new variable say X,

Let $G_1 = (V_1, T_1, P_1, S)$ where $V_1 = V \cup \{X\}$

P_1 can be formed by replacing the A → production by

(i) $A \to \beta_i$
$A \to \beta_i X$
$(i \le 1 \le n)$

$A \to \beta_1 | \beta_2 | \ldots \beta_n$

$A \to \beta_1 X | \beta_2 X - \beta_n X$

(ii) $X \to \alpha_i$ $\quad$ $X \to \alpha_1 | \alpha_2 | \ldots \alpha_m$

$X \to \alpha_i X$

$(i < i \leq m)$ $\quad$ $X \to \alpha_1 X.$