Unit 3 - Context Free languages

* Introduction
* Formal defn of context free grammar
* Notational Conventions
* Derivation / Parsing
  - Types
  - Examples.

$x = (a+b) * c$

$A \rightarrow id = E$
$E \rightarrow E+E \mid E*E \mid (E) \mid id$

$A \Rightarrow id = E$
$\Rightarrow id = E*E$
$\Rightarrow id = (E)*E$
$\Rightarrow id = (E+E)*E$
$\Rightarrow id = (id+E)*E$
$\Rightarrow id = (id+id)*E$
$\Rightarrow id = (id+id)*id$.

Eg 1:

$S \rightarrow AS \mid \epsilon$
$A \rightarrow aa \mid ab \mid ba \mid bb$

Give LMD, RMD and Parse tree for a sentence
(i) aabbaa
(ii) baabab.

$G = (V, T, S, P)$
$V = \{S, A\}$
$T = \{a, b\} \cup \epsilon$.

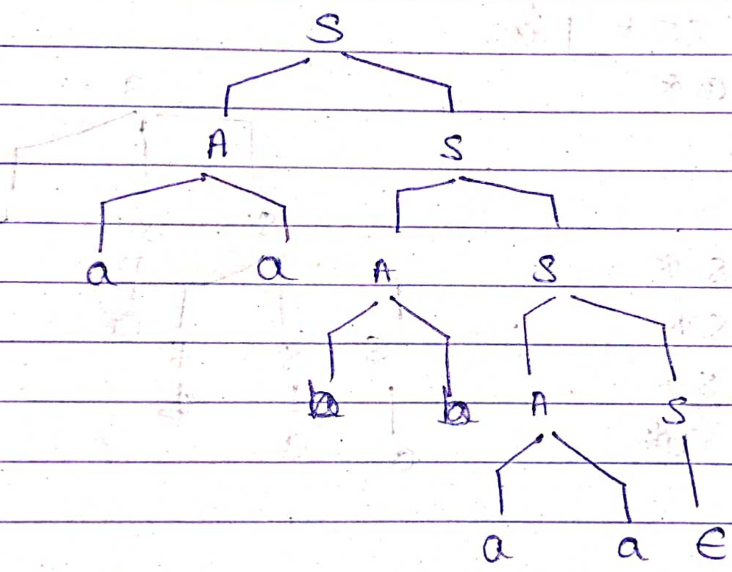
Parsing → Identification of sequence of production Application

$S \rightarrow \{S\}$

P = 6 productions.

$$\begin{aligned} S &\overset{LMD}{\Rightarrow} \underline{A}\,S \\ &\Rightarrow aa\,\underline{S} \\ &\Rightarrow aa\,\underline{A}\,S \\ &\Rightarrow aa\,bb\,\underline{S} \\ &\Rightarrow aa\,bb\,\underline{A}\,S \\ &\Rightarrow aa\,bb\,aa\,S \\ &\Rightarrow aa\,bb\,aa \end{aligned}$$

Left sentential form. (Every sentence in derivation form)



$$\begin{aligned} S &\overset{RMD}{\Rightarrow} A\,\underline{S} \\ &\Rightarrow A\,A\,\underline{S} \\ &\Rightarrow A\,A\,A\,\underline{S} \\ &\Rightarrow A\,A\,A \\ &\Rightarrow A\,A\,aa \\ &\Rightarrow A\,bb\,aa \\ &\Rightarrow aa\,bb\,aa. \end{aligned}$$

Right sentential form

$S \to AS$
$A \to aa$
$S \to AS$
$A \to bb$
$S \to AS$
$A \to aa$
$S \to \epsilon$

← Parsing
The program that performs the task of parsing is parser. (i.e derivations).

(ii) baabaa

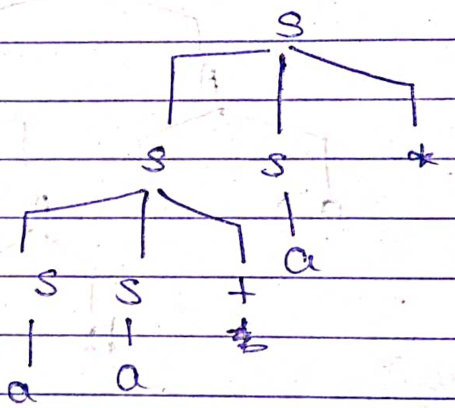
Write LMD, RMD and Parse tree for the sentence and thereby validate the sentence

(i) $S \to SS+ \mid SS* \mid a$

$w = aa+a*$

LMD

$S \Rightarrow \underline{S}S*$
$\Rightarrow \underline{S}S+S*$
$\Rightarrow a\underline{S}+S*$
$\Rightarrow aa+\underline{S}*$
$\Rightarrow aa+a*$



RMD

$S \Rightarrow S\underline{S}*$
$\Rightarrow \underline{S}a*$
$\Rightarrow S\underline{S}+a*$
$\Rightarrow \underline{S}a+a*$
$\Rightarrow aa+a*$

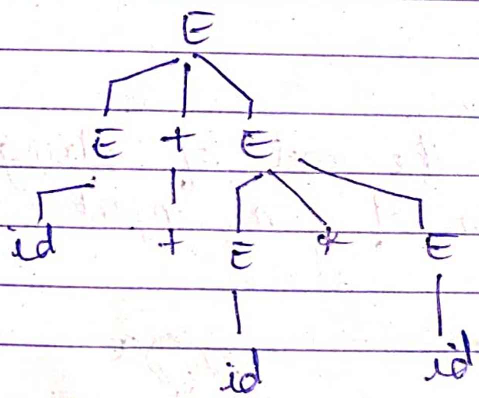
(ii) $E \to E+E \mid E*E \mid (E) \mid id$

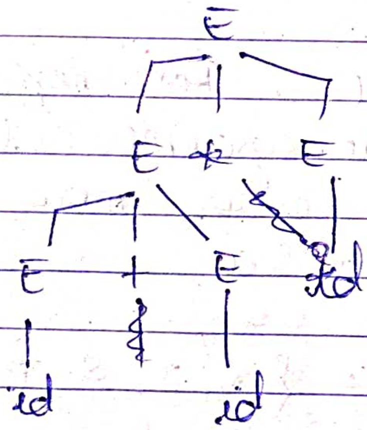
$w = id + id * id$

$$E \overset{LM}{\Rightarrow} E * E$$
$$\Rightarrow E + E * E$$
$$\Rightarrow id + E * E$$
$$\Rightarrow id + id * E$$
$$\Rightarrow id + id * id.$$

or

$$E \overset{LM}{\Rightarrow} E + E$$
$$\Rightarrow id + E$$
$$\Rightarrow id + E * E$$
$$\Rightarrow id + id * E$$
$$\Rightarrow id + id * id.$$





① $S \to AaAb \mid BbBa$

$A \to aAb \mid bAB \mid d$

$B \to aB \mid bB \mid a$

(i) aabbaa   ii) badbabaadb

~~(iii) id id id * + id +~~   ~~(iv) id id id ↑ id * id * +~~

② $E \to ET+ \mid T$

$T \to TF* \mid F$

$F \to FP\uparrow \mid P$

$P \to E \mid id$

(i) id id id * + id +

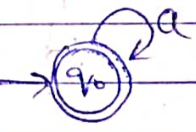
(ii) id id id ↑ id * id * +

Example : ~~For regular grammar~~.

① write grammer to accept the set of strings with 0 or more combination of an input symbol a

→

$G = \{V, T, S, P\}$



$V = \{q_0\}$ $T = \{a\}$ $S = \{q_0\}$

$q_0 \to aq_0 \mid \epsilon$ or $S \to aS \mid \epsilon$

② 0 or more combination of a's and b's.

$G = \{V, T, S, P\}$

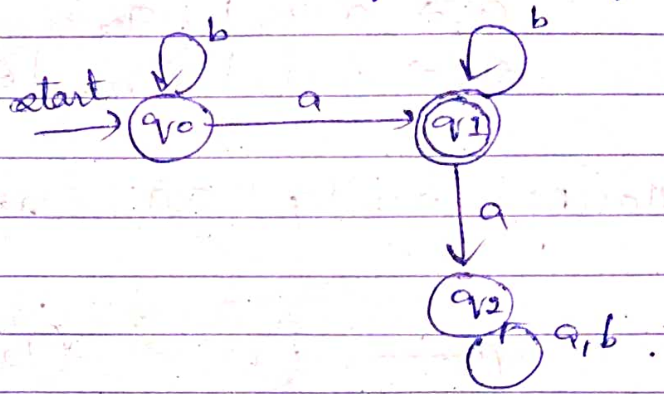
$V = \{q_0\}$ $S = \{q_0\}$

$T = \{a, b\}$ $q_0 \to aq_0 \mid bq_0 \mid \epsilon$.

ε-productions } should be removed
unit - "
useless - "

③ Write a DFA for the language accepting set of strings of a's and b's having exactly one a.

→

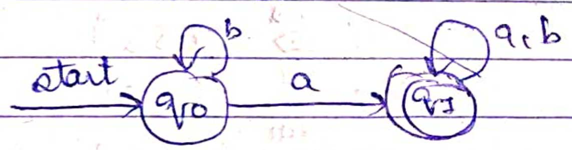


$G = \{V, T, S, P\}$

$V = \{q_0, q_1, q_2\}$
$T = \{a, b\}$
$S = \{q_0\}$

$q_0 \rightarrow aq_1 \mid bq_0$
$q_1 \rightarrow bq_1 \mid aq_2 \mid \epsilon$
$q_2 \rightarrow aq_2 \mid bq_2$.

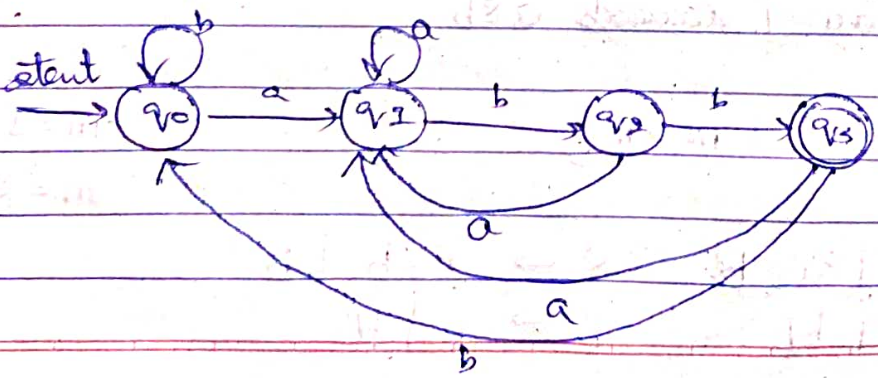
④ Atleast one a.



$V = \{q_0, q_1\}$ $T = \{a, b\}$ $S = \{q_0\}$.

$q_0 \rightarrow aq_1 \mid bq_0$
$q_1 \rightarrow aq_1 \mid bq_1 \mid \epsilon$

⑤ Set of strings of a's and b's ending with abb.



$q_0 \to a q_1 \mid b q_0$
$q_1 \to a q_1 \mid b q_2$
$q_2 \to a q_1 \mid b q_3$
$q_3 \to a q_1 \mid b q_0 \mid \epsilon$

Having substr aa regular expressio.

$(a+b)^* aa (a+b)^*$
A B A

$S \to A\,B\,A$
$A \to aA \mid bA \mid \epsilon$
$B \to aa$

① $L = \{a^n b^n \mid n >= 0\}$.

for $n = 3$

$G = \{V, T, S, P)$
$V = S$
$T = \{a, b\}$
$S \to S$
$P = S \to aSb \mid \epsilon$

$S \overset{lmd}{\Rightarrow} aSb$
$S \overset{lmd}{\Rightarrow} aaSbb$
$S \overset{lm}{\Rightarrow} aaaSbbb$
$S \overset{lm}{\Rightarrow} aaabbb$

② $L = \{a^n b^{n-3} \mid n >= 3\}$.

aaab

$S \to aaa \mid \cancel{aaaaSb}\ aSb$

③ $L = \{a^n b^m \mid m > n\}$

$m = 1 \quad n = 0$
$m = 2 \quad n = 1.$

$\cancel{S \to b \mid aSB}$
$\cancel{B \to b \mid bB}$

$S \to aSb \mid B$
$B \to b \mid bB$

(4) $L = \{a^n b^m \mid m < n\}$.

$S \rightarrow aSb \mid A$
$A \rightarrow a \mid aA$.

(5) $L = \{a^n b^m \mid m <> n\}$

$S \rightarrow aSb \mid A \mid B$
$A \rightarrow a \mid aA$
$B \rightarrow b \mid bB$

P

$V \rightarrow \{S, A, B\}$
$T \rightarrow \{a, b\}$
$S \rightarrow$ start symbol $= S$.

(6) $L = \{\omega\omega^R \mid \omega \in (a+b)^*\}$

$\omega \rightarrow a^n b^n$ $\qquad\qquad$ $\omega\omega^R \rightarrow a^n b^n b^n a^n$
$\omega^R \rightarrow b^n a^n$. $\qquad\qquad \rightarrow$ ~~$a^n b^n$~~

$S \rightarrow$ ~~$a \in \mid aa \mid S \mid bb$~~ $absb$~~a~~
$S \rightarrow aSa \mid bSb \mid \epsilon \mid a \mid b$. $\rightarrow$ for all lgths.

$S \rightarrow aSa \mid bSb \mid \epsilon$ $\rightarrow$ Ans

(7)

$[+,-]?\ [0-9]^+$

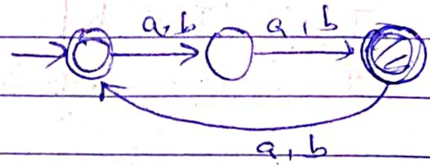
8 $\Rightarrow I \rightarrow N \mid S_I N$
$N \rightarrow D \mid ND$
$D \rightarrow 0 \mid 1 \mid 2 \mid \cdots \mid 9$.
$S_1 \rightarrow + \mid - \mid \epsilon$.

+123

$I \overset{lm}{\Rightarrow} S_1N$
$\Rightarrow +N$
$\Rightarrow +ND$
$\Rightarrow +NDD$
$\Rightarrow +DDD$
$\Rightarrow +1DD$
$\Rightarrow +12D$
$\Rightarrow +123$

→

(8) $((a+b)(a+b)(a+b))^*$



$S \rightarrow \epsilon \mid aS_1 \mid bS_1$
$S_1 \rightarrow aS_2 \mid bS_2$
$S_2 \rightarrow aS \mid bS$

$q_0 \rightarrow S$
$q_1 \rightarrow S_1$
$q_2 \rightarrow S_2$.

* $w = abaabb$

$S \overset{LMD}{\Rightarrow} aS_1$
$S \overset{LMD}{\Rightarrow} abS_2$
$S \overset{LMD}{\Rightarrow} abaS$
$S \overset{LMD}{\Rightarrow} abaaS_1$
$S \overset{LMD}{\Rightarrow} abaabS_2$
$S \overset{LMD}{\Rightarrow} abaabbS$
$S \overset{LMD}{\Rightarrow} abaabb$.

* $w = abaa$

$S \overset{LMD}{\Rightarrow} aS_1$
$S \overset{LMD}{\Rightarrow} abS_2$
$S \overset{LMD}{\Rightarrow} abaS$
$S \overset{LMD}{\Rightarrow} abaaS_1$
$S \Rightarrow$

(9)

$L = \{a^i b^j c^k, \ i = j+k, \text{ where } j, k \geq 0\}$.

$j = 1 \quad k =$
$i = 2$.

$L = a^i b^j c^k$
$= a^{j+k} b^j c^k$
$= a^j a^k b^j c^k$

$$= \underbrace{a^k \underbrace{a^j b^j}_{S_1} c^k}_{S_2}.$$

$S \rightarrow aSc \mid aS_1 \cancel{b} \mid \varepsilon$ $\qquad\qquad$ $S \rightarrow aSc \mid S_1 \mid \varepsilon$

$S_1 \rightarrow aS_1b \mid \varepsilon$. $\qquad\qquad$ $S_1 \rightarrow aS_1b \mid \varepsilon$.

$w = a^5 b^2 c^3$

$w = aaaaa\,bb\,ccc$

$S \xRightarrow{LMD} a\underline{S}c$

$S \xRightarrow{LMD} aa\underline{S}cc$

$S \xRightarrow{LMD} aaa\underline{S}ccc$

$S \xRightarrow{LMD} aaaaS_1bccc$

$S \xRightarrow{LMD} aaaaaS_1bbccc$

$S \xRightarrow{LMD} aaaaabbccc$.

(10) $L = \{a^i b^j c^k,\ k = i+j,\ \text{where } i, j \geq 0\}$.

$$L = a^i b^j c^i c^j$$
$$= \underbrace{a^i \underbrace{b^j c^j}_{S_1} c^i}_{S}$$

$S \rightarrow aSc \mid S_1 \mid \varepsilon$

$S_1 \rightarrow bS_1c \mid \varepsilon$.

(11) $L = \{a^i b^j c^k,\ j = i+k$

$L = \underbrace{a^i b^i}_{S_1}\ \underbrace{b^k c^k}_{S_2}$. $\quad a^i$

~~$S \to aSc \mid S1 \mid e$~~

~~$S1 \to aS_1b \mid b \mid e$~~

$S \to S1\,S2$

$S1 \to aS1b \mid e$

$S2 \to bS2c \mid e$

$\to$ $i = 2$ $k = 3$

$\Rightarrow w = aa\ bbbbb\ ccc$

w $S \xRightarrow{LMD} S1\ S2$

$\xRightarrow{LMD} aS1b\ S2$

$\xRightarrow{LMD} aaS1bb\ S2$

$\Rightarrow aabb\ S2$

$\Rightarrow aabb\ bS_2c$

$\Rightarrow aabbbb\ S_2cc$

$\Rightarrow aabbbbbS_2ccc$

$\Rightarrow aabbbbbccc$

(12) $L = \{ a^n w w^R b^n \mid w \in (a+b)^*, n >= 0 \}$.

$w = a^n b^n$

$w^R = b^n a^n$

$L = \underbrace{a^n \underbrace{a^n b^n b^n a^n}_{S1} b^n}_{S}$ $\quad a^n a^{2n} \cancel{b^{2n}} \cancel{b^n}$

$S \to aSb \mid e \mid S1$

$S1 \to aS_1a \mid bS_1b \mid e$.

$n = 3 \qquad w = aba.$

$L = \{ aaa\ aba\ aba\ bbb \}.$

$S \overset{LMD}{\Rightarrow} aSb$
$S \overset{LMD}{\Rightarrow} aaSbb$
$S \overset{LMD}{\Rightarrow} aaa\ Sbbb$
$S \overset{LMD}{\Rightarrow} aaa\ SI bbb$
$S \overset{LMD}{\Rightarrow} aaa\ asa\ bbb$
$S \overset{LMD}{\Rightarrow} aaa\ absba\ bbb$
$S \overset{LMD}{\Rightarrow} aaaa\ basa\ babbb$
$S \overset{LMD}{\Rightarrow} aaaa\ ba\ aba\ bbb$

* (13) $L = \{ a^n b^m c^k \mid n+2m = k \} \ m \,\&\, n \geq 0 \}.$

• Language of grammer $L(G)$

If $G = \{V, T, S, P\}$ be a context free grammar the language of G denoted as $L(G)$ defined to be the set of terminal strings that have a derivation from the start symbol

$$L(G) = \{ w \text{ in } T^* \mid S \overset{*}{\underset{G}{\Rightarrow}} w$$

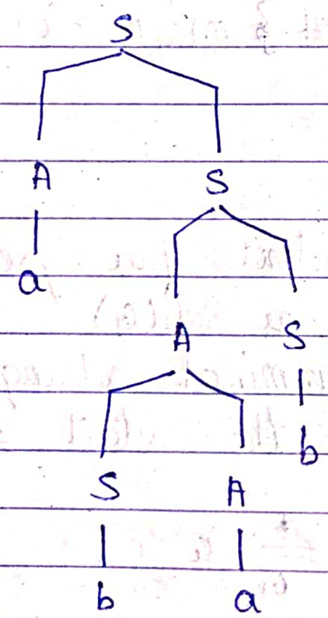
1) $S \rightarrow AS \mid b$

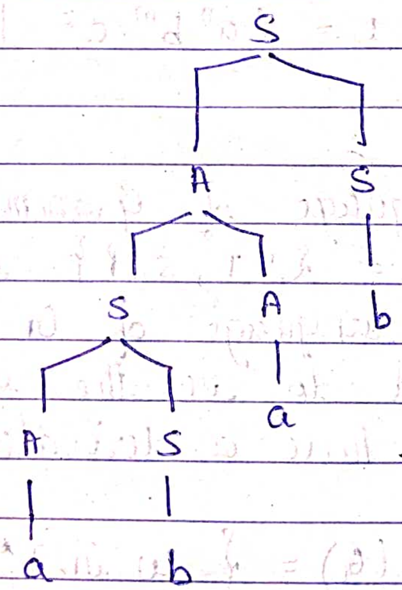
$A \rightarrow SA \mid a$

Prove that the grammar is ambigious on $w = abab$.

$\rightarrow$

$S \overset{LMD}{\Rightarrow} AS$	$S \overset{LMD}{\Rightarrow} AS$
$S \overset{LMD}{\Rightarrow} aS$	$S \overset{LMD}{\Rightarrow} SAS$
$S \overset{LMD}{\Rightarrow} aAS$	$S \overset{LMD}{\Rightarrow} ASAS$
$S \overset{LMD}{\Rightarrow} aSAS$	$S \overset{LMD}{\Rightarrow} aSAS$
$S \overset{LMD}{\Rightarrow} abAS$	$S \overset{LMD}{\Rightarrow} abAS$
$S \overset{LMD}{\Rightarrow} abas$	$S \overset{LMD}{\Rightarrow} abaS$
$S \overset{LMD}{\Rightarrow} abab$	$S \overset{LMD}{\Rightarrow} abab$.





There are 2 LMD for sentence abab, therefore 2 parse trees ∴ grammar is ambigious.

2)

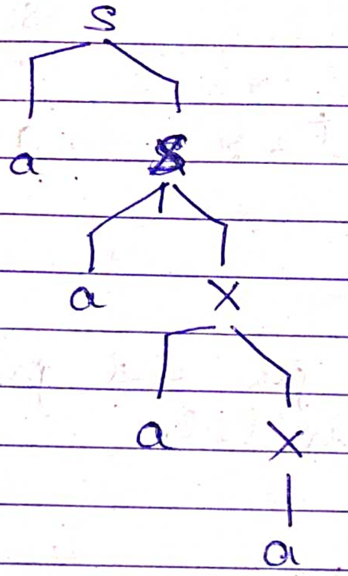
$S \rightarrow aS \mid X$

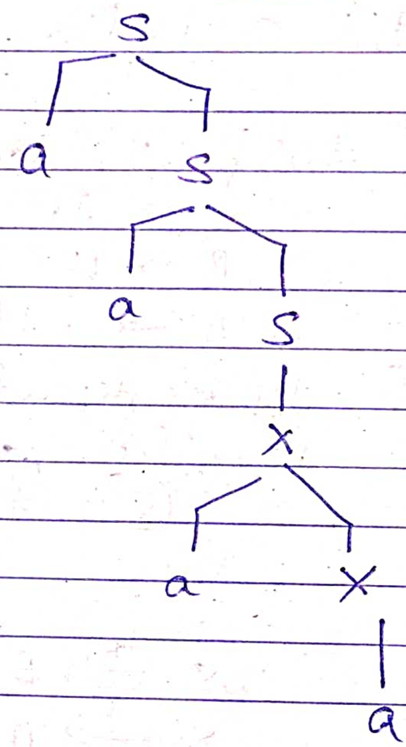
$X \rightarrow aX \mid a$.

aaaa.

$$S \overset{LMD}{\Rightarrow} a\underline{S} \overset{LMD}{\Rightarrow} aX \overset{LMD}{\Rightarrow} aaX \overset{LMD}{\Rightarrow} aaaX \Rightarrow aaaa.$$

$$S \overset{LMD}{\Rightarrow} a\underline{S} \Rightarrow aaS \Rightarrow aaX \Rightarrow aaaX \Rightarrow aaaa$$





A

$S \rightarrow ABac$

$A \rightarrow BC$

$B \rightarrow b \mid \epsilon$

$C \rightarrow D \mid \epsilon$

$D \rightarrow d$

① finding nullable variables V_N

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

② exclude ϵ productions

$S \rightarrow ABac$

$A \rightarrow BC$

$B \rightarrow b$

$C \rightarrow D$

$D \rightarrow d$

• Simplified grammar after removal of ϵ.

$S \rightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa$

$A \rightarrow BC \mid B \mid C$

$B \rightarrow b$

$C \rightarrow D$

$D \rightarrow d$

* Unit Production

$S \rightarrow Aa \mid B$

$B \rightarrow A \mid bb$

$A \rightarrow a \mid bc$

(1) Listing all Non-unit productions :

$S \rightarrow Aa$

$B \rightarrow bb$

$A \rightarrow a \mid bc$

(2) ~~depeng~~ depending graph for unit production
Variables part of unit production

$S \stackrel{*}{\Rightarrow} B$

$S \stackrel{*}{\Rightarrow} A$

$B \stackrel{*}{\Rightarrow} A$

$A \stackrel{*}{\Rightarrow} B$

$S \rightarrow Aa \mid bb \mid a \mid bc$

$B \rightarrow bb \mid a \mid bc$ → useless pro

$A \rightarrow a \mid bc \mid bb.$

* Useless Production

• Useful Variable → Let $G = \{V, T, S, P\}$ be a context free grammar, any variable $A \in V$ is said to be useful, if and only if there is atleast one $w \in L(G)$ such that

$$S \stackrel{*}{\Rightarrow} xAy \stackrel{*}{\Rightarrow} w$$

① $S \rightarrow aSb \mid \epsilon \mid A$

$A \rightarrow aA$

A is useless, becoz it doesn't lead to any terminal state.

B

② $S \rightarrow A$

$A \rightarrow aA \mid \epsilon$

$B \rightarrow bA$

B is useless becoz it is not taking part in the derivation step.

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

$A \rightarrow a$

$B \rightarrow aa$

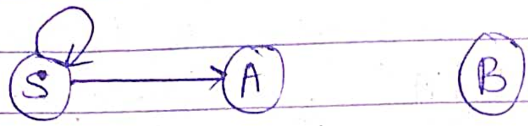
$C \rightarrow aCb$.

① Remove the variable that doesn't derive any terminal string

OV	NV	Productions
ϕ	A, B	$A \rightarrow a$ $B \rightarrow aa$
A, B	A, B, S	$S \rightarrow A$
A, B, S	A, B, S	$S \rightarrow aS$

C is removed.

② Remove the variable that is not taking part in derivation process



$S \rightarrow B$ not reachable hence it is useless

$S \rightarrow A \mid aS$

$A \rightarrow a$

① $S \rightarrow a \mid aA \mid B \mid C$

$A \rightarrow aB \mid \epsilon$

$B \rightarrow Aa$

$C \rightarrow cCD$

$D \rightarrow ddd$

(1) Removal of ϵ- productions

(1) Nullable variable $V_N = \{A\}$

$S \rightarrow a \mid aA \mid B \mid C$

$A \rightarrow aB$

$B \rightarrow Aa$

$C \rightarrow cCD$

$D \rightarrow ddd$

$S \rightarrow a \mid aA \mid B \mid C$

$A \rightarrow aB$

$B \rightarrow Aa \mid a$

$C \rightarrow cCD$

$D \rightarrow ddd.$

(ii) Removal of Unit Productions

$\cancel{S \to B}$ $S \to a \mid aA$

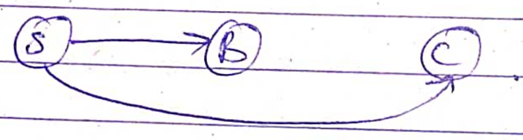
$\cancel{S \to C}$ $A \to aB$

$B \to Aa \mid a$

$C \to cCD$

$D \to ddd.$

Dependency graph.

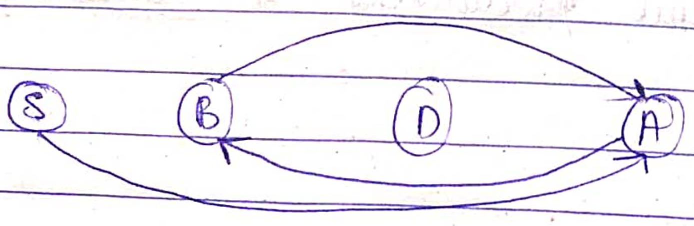


$S \overset{*}{\Rightarrow} B$

$S \overset{*}{\Rightarrow} C.$

$S \to a \mid aA \mid Aa \mid \cancel{cC} \; cCD$

OV	NV	Productions
ϕ	S, B, D	$S \to a$ $B \to a$ $D \to ddd.$
S, B, D	S, B, D, A	$A \to aB$
S, B, D, A	S, B, D, A, ~~B~~	$S \to Aa$ $B \to Aa$ $S \to aA$



$S \to a \mid Aa \mid aA \mid \cancel{cCD}$ $A \to aB$

$B \to a \mid Aa$

$D \to \cancel{ddd}.$

② $S \rightarrow aAa \mid bBb \mid \epsilon$
$A \rightarrow c \mid a$
$B \rightarrow c \mid b$
$C \rightarrow CDE \mid \epsilon$
$D \rightarrow A \mid B \mid ab$

① Removal of ϵ productions

$S \rightarrow aAa \mid bBb$ $\qquad V_n = \{S, C\}, A, B, D$
$A \rightarrow c \mid a$
$B \rightarrow c \mid b$
$C \rightarrow CDE$
$D \rightarrow A \mid B \mid ab$

$S \rightarrow aAa \mid bBb \mid aa \mid bb$
$A \rightarrow C \mid a$
$B \rightarrow C \mid b$
$C \rightarrow CDE \mid DE \mid CE \mid E$
$D \rightarrow A \mid B \mid ab$

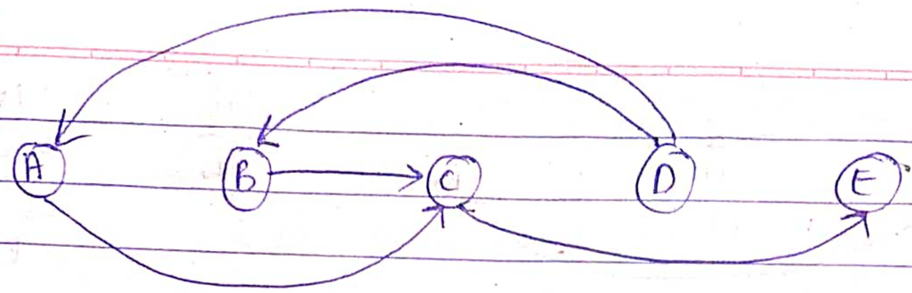
② Removal of unit productions

* Listing non-unit productions

$S \rightarrow aAa \mid bBb \mid aa \mid bb$
$A \rightarrow a$
$B \rightarrow b$
$C \rightarrow CDE \mid DE \mid CE$
$D \rightarrow ab$

(a) * Identify unit productions & write dependency.

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



$A \overset{*}{\Rightarrow} C$

$A \overset{*}{\Rightarrow} E$

$B \overset{*}{\Rightarrow} C$

$B \overset{*}{\Rightarrow} E$

$C \overset{*}{\Rightarrow} E$

$D \overset{*}{\Rightarrow} B$

$D \overset{*}{\Rightarrow} A$

$D \overset{*}{\Rightarrow} C$

$D \overset{*}{\Rightarrow} E$

(b) non-unit productions

$S \rightarrow aAa \mid bBb \mid aa \mid bb$

$A \rightarrow a \mid CDE \mid DE \mid CE$

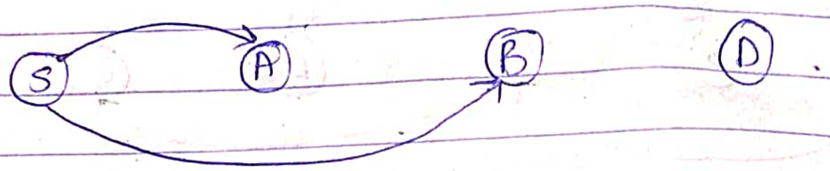
$B \rightarrow b \mid CDE \mid DE \mid CE$

$C \rightarrow CDE \mid DE \mid CE$

$D \rightarrow ab \mid a \mid b \mid CDE \mid DE \mid CE$

③ Removal of useless productions

OV	NV	productions
ϕ	S, A, B, D	$S \rightarrow aa$ $S \rightarrow bb$ $A \rightarrow a$ $B \rightarrow b$ $D \rightarrow ab$ $D \rightarrow a$ $D \rightarrow b$
S, A, B, D	S, A, B, D	$S \rightarrow aAa \mid bBb$



D is useless as it is not reached from S.

$S \rightarrow aa \mid bb \mid aAa \mid bBb$

$A \rightarrow a$

$B \rightarrow b$

* $S \rightarrow aa \mid bb \mid aaa \mid bbb$.

* CNF → Chomsky Normal form

→ If G is a grammer representing non-empty CFL without ε and unit productions in which all productions are in one of 2 simple forms either

(i) $A \rightarrow BC$ where A, B, C are all variabl

(ii) $A \rightarrow a$ where a is variable & termi

such a grammer we say that it is in chomsky normal form.

(1) Arrange that all the bodies of length 2 or more consists only of variables.

(II) Break bodies of length 3 or more into a cascade of productions each with a body consisting of 2 variables.

Ep: (1)

$S \rightarrow ABa$

$A \rightarrow aab$

$B \rightarrow AC$

(I)

$S \to AB B_a$

$B_a \to a$

$A \to B_a B_a B_b$

$B_b \to b$

$B \to A B_c$

$B_c \to c$

(II)

~~$S \to AB B_a$~~

$S \to A D_1$

$D_1 \to B B_a$

$B_a \to a$

$A \to B_a D_2$

$D_2 \to B_a B_b$

$B_b \to b$

$B \to A B_c$

$B_c \to c$

(2)

$E \to E+T \mid T$

$T \to T * F \mid F$

$F \to (E) \mid id$

~~$E \to E B_+ T$~~

(i) Remove unit productions.

- Identify unit Productions

$E \to T$

$T \to F$

~~$F \to (E)$~~

$E \to T \to F$

$E \overset{*}{\Rightarrow} T \qquad T \overset{*}{\Rightarrow} F$

$E \overset{*}{\Rightarrow} F$

- List non - unit Productions

$E \rightarrow E+T \mid T*F \mid (E) \mid id$

$T \rightarrow T*F \mid (E) \mid id$.

$F \rightarrow (E)$

$F \rightarrow id$

(I) $E \rightarrow EB_{+}T \mid TB_{*}F \mid B_{(}EB_{)} \mid id$

$B_{+} \rightarrow +$

$B_{*} \rightarrow *$

$B_{(} \rightarrow ($

$B_{)} \rightarrow)$

$T \rightarrow TB_{*}F \mid B_{(}EB_{)} \mid id$

$F \rightarrow B_{(}EB_{)}$

$F \rightarrow id$

(II) $E \rightarrow ED_1 \mid TD_2 \mid B_{(}D_3 \mid id$

$D_1 \rightarrow B_{+}T$

$D_2 \rightarrow B_{*}F$

$D_3 \rightarrow EB_{)}$

$B_{+} \rightarrow +$

$B_{*} \rightarrow *$

$B_{(} \rightarrow ($

$B_{)} \rightarrow)$

$T \rightarrow TD_2 \mid B_{(}D_3 \mid id$

$F \rightarrow B_{(}D_3$

$F \rightarrow id$

GNF (greiback Normal Form)

A context free grammer is said to be in graybach normal form if all productions are of the form

$A \rightarrow ax$

where,

$a \in T$

& $x \in V^*$

Eg: ① $S \rightarrow AB$

$A \rightarrow aA \mid bB \mid b$

$B \rightarrow b$

$S \rightarrow aAB \mid bBB \mid bB$

$A \rightarrow aA \mid bB \mid b$

$B \rightarrow b$

② $S \rightarrow AbA$

$A \rightarrow BcB$

$B \rightarrow c$

$S \rightarrow AbA$

$A \rightarrow ccB$

$B \rightarrow c$

$S \rightarrow ccBbA$

$A \rightarrow ccB$

$B \rightarrow c$

$S \rightarrow cBBB_b A$

$B_b \rightarrow b$

$A \rightarrow cBB$

$B \rightarrow c$.

4/1/22
5/1/22

Using yacc tool write a parser to validate the syntax of arithmetic expression

⇒

s → e

e → e+e | e-e | e*e | e/e | e^e | ID | D

```
%{
    #include <stdio.h>
%}
%token ID D
%%
    s : e '\n'        { printf ("%d", $1);
                        printf ("valid"); }
      ;
    e : e'+'e         { $$ = $1 + $2 ; }
      | e'-'e         { $$ = $1 - $2 ; }
      | e'*'e         { $$ = $1 * $2 ; }
      | e'/'e         { $$ = $1 / $2 ; }
      | e'^'e         { $$ = $1 ^ $2 ; }
      | ID            { $$ = $1 ; }
      | D             { $$ = $1 ; }
      ;               { ; }
%%
main ()
{
    printf (" Enter the arithmetic expression :");
    yyparse () ;
    return 0;
}
yyerror ()
{
    printf (" it is not valid");
}
```

```
%{
#include "y.tab.h"
#include <stdio.h>
extern int yylval;
%}
%%
[0-9]+                      {return D;}
[a-zA-Z]                    {return ID;}
[\+\-\*\/\(\)]              {return yytext[0];}
.|
\n                          {return yytext[0];}
%%
```

```
yacc -d exp-yacc.y          →  y.tab.c
lex exp-lex.l                  y.tab.h
cc y.tab.c lex.yy.c -ll     →  lex.yy.c
./a.out
```

6/1/22

shift
Reduce
Accept
error

write a yacc program to validate arithmatic expression and evaluate ...

yacc program

```
%{
 #include <stdio.h>
%}
%token D
%left '+' '-'
%left '*' '/'
%right '^'
%%
s : e '\n'   { printf ("Valid \n");
                printf ("value of exp %d", $1
             }
  ;
e : e'+'e    { $$ = $1 + $3; }
  | e'-'e    { $$ = $1 - $3; }
  | e'*'e    { $$ = $1 * $3; }
  | e'/'e    { if ($3 ==0)
                 { printf ("Divide of zero
                   exit(0);
                 }
               else
                   $$ = $1 / $3;
             }
  | e^e      { $$ = $1 ^ $3; }
  | D        { $$ = $1; }
```

exp valid.l

```
%{
  #include "y.tab.h"
  #include <stdio.h>
  extern int yylval;
%}
%%
  [0-9]+            { yylval = atoi (yytext); return D; }
  [\+\-\*\|\^\(\)]
  \n
%%
```

* $a^n b^n$

1) $L = \{a^n b^n \mid n >= 0\}$

2) $L = \{a^n b^{2n} \mid n >= 0\}$.

$\rightarrow$ $\delta(q_0, \epsilon, Z) = (q_f, \epsilon, Z)$ $\rightarrow n = 0$.

~~$\delta(q_0, a, Z) = (q_0, AZ)$~~

~~$\delta(q_0, a, A) = (q_0, AA)$~~

$\delta(q_0, a, Z) = (q_0, AAZ)$

$\delta(q_0, a, A) = (q_0, AAA)$

$\delta(q_0, b, A) = (q_1, \epsilon)$

$\delta(q_1, b, A) = (q_1, \epsilon)$

for $n = 2$

$(q_0, aabbbb, Z) \vdash (q_0, abbbb, AAZ)$

$\vdash (q_0, bbbb, AAAAZ)$

$\vdash (q_1, bbb, AAAZ)$

$\vdash (q_1, bb, AAZ)$

$\vdash (q_1, b, AZ)$

$\vdash (q_1, \epsilon, Z)$

$\vdash (q_f, \epsilon, Z)$

3) $L = \{w \mid w \in (a+b)^*$ and $n_a(w) = n_b(w)\}$.

$\rightarrow$ $\delta(q_0, \epsilon, Z) = \delta(q_f, Z)$.

$\delta(q_0, a, Z) = \delta(q_0, AZ)$

$\delta(q_0, b, Z) = \delta(q_0, BZ)$

$\delta(q_0, a, B) = \delta(q_0, \epsilon)$

$\delta(q_0, a, A) = \delta(q_0, AA)$

$\delta(q_0, b, A) = \delta(q_0, \epsilon)$

$\delta(q_0, b, B) = \delta(q_0, BB)$

$\delta(q_0, \epsilon, Z) = (q_f, Z)$.

$(q_0, bbabaa, Z) \vdash (q_0$

4) $L = \{ w \mid w \in (a+b)^* \text{ and } n_a(w) > n_b(w) \}$.

$\delta(q_0, \epsilon, z) = (q_f, z)$
$\delta(q_0, a, z) = (q_0, AZ)$
$\delta(q_0, b, z) = (q_0, BZ)$
$\delta(q_0, a, B) = (q_0, \epsilon)$
$\delta(q_0, a, A) = (q_0, AA)$
$\delta(q_0, b, A) = (q_0, \epsilon)$
$\delta(q_0, b, B) = (q_0, BB)$
$\delta(q_0, a, z) = (q_f, Z)$.

$L = \{ a^n b^m c^{(n+m)} \mid n >= 0 \}$.
$a^n b^m c^m c^n$

→

$\delta(q_0, \epsilon, z) = (q_f, z)$
$\delta(q_0, a, z) = (q_0, AZ)$
$\delta(q_0, a, A) = (q_0, AA)$
$\delta(q_0, b, A) = (q_1, BA)$
$\delta(q_1, b, B) = (q_1, BB)$
$\delta(q_1, c, B) = (q_2, \epsilon)$
$\delta(q_2, c, A) = (q_2, \epsilon)$
$\delta(q_2, \epsilon, z) = (q_f, z)$

$L = \{ a^{2n} b^n \mid n >= 0 \}$

$\delta(q$