06 Anikate Agrawal 201500088

Automata Assignment II

1. Explain why the given grammar is ambiguous.

$S \to 0A \mid 1B$ $\quad A \to 0AA \mid 1S \mid 1$ $\quad B \to 1BB \mid 0S \mid 0$

LMD 1	LMD 2
$S \to 1B$	$S \to 1B$
$\to 11BB$	$\to 11BB$
$\to 110SB$	$\to 110B$
$\to 1100AB$	$\to 1100S$
$\to 11001B$	$\to 11001B$
$\to 110010$	$\to 110010$

We have 2 left most derivative for same string 110010. So the given grammar is ambiguous.

2. Given the following ambiguous context free grammar

$S \to Ab \mid aaB$ $\quad A \to a \mid Aa$. $\quad B \to b$

(a) Find the string generated by the grammar that has two leftmost derivations. Show the derivations.

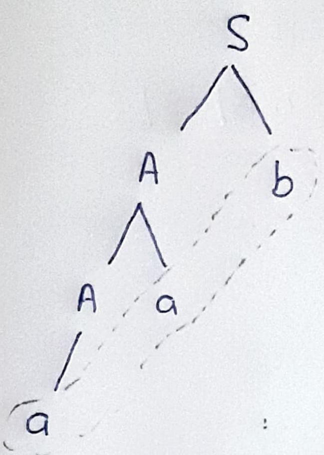
LMD 1	LMD 2
$S \to Ab$	$S \to aaB$
$\to Aab$	$\to aab$
$\to aab$	

So the string aab has two LMD.

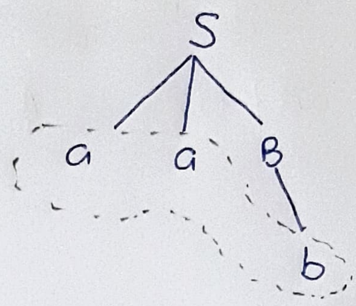
06 Anikate Agrawal 201500088

(b) Show the two derivation trees for the string.

String - aab



LMD 1



LMD 2

3. State and prove pumping lemma for CFG.

Let L be a CFL. There exists some integer n such that for all w in L with $|w| \geq n$, $w = uvxyz$ with $|vxy| \leq n$ and $|vy| > 0$ such that $uv^i xy^i z \in L$ for all $i = 0, 1, 2, 3 \ldots$

Proof :-

- Let G be a grammar in CNF generating the language $L - \{\varepsilon\}$
- Let the grammar have m variables.
- Pick $n = 2^m$.
- Let $w \in L(G)$ & $|w| \geq n$
- Any derivation tree for w has height at least $m+1$.
- If any derivation tree for w has height at least $m+1$, then the number of nodes in the path is atleast $m+2$.

06 Anikate Agrawal 201500088

- One node is the leaf node labeled with a terminal.
- At least m+1 nodes are internal nodes labeled with non-terminals.
- Since there are only m non-terminals in the grammar, and since m+1 appear on this long path, it follows that some non-terminal (and perhaps many) appears at least twice on this path.

Prove that the given language is not context free.

i) $L = \{ a^i b^j c^i d^j \mid i, j \geq 0 \}$

Let L be a CFL

Let n be length of pumping lemma

$z = a^n b^m c^n d^m$, $|z| \geq n$

Split $z = uvwxy$

$u = a^n b^m$, $vwx = c^n$. $y = d^m$

$|vwx| \leq n$, $vx = c^{n-p}$, $p < n$, $|vx| \geq 1$

$uv^k wx^k y = u\,v\,v^{k-1}\,w\,x\,x^{k-1}\,y$

$= uvwx\,(vx)^{k-1}\,y$

$= a^n b^m c^n (c^{n-p})^{k-1} d^m$

$= a^n b^m c^n (c^{nk-n-pk+p}) d^m$

$= a^n b^m c^{\cancel{n}+nk\cancel{-n}-pk+p} d^m$

$= a^n b^m c^l d^m \notin L$

So our assumption is wrong

L is not context free.

06 Anikate Agrawal 201500088

ii) $L = \{ a^i b^i c^i \mid i \geq 1 \}$

Let L be a CFL. Since L is infinite, the pumping lemma can be applied. Let n be the constant of the lemma.

Let $w \in L$ & $w = a^k b^k c^k$ for some $k > n$.

$w = uvxyz$, $|vy| > 0$ $|vxy| \leq n$

Since $|vxy| \leq n$, therefore it cannot contain all 3 symbols - a, b & c.

- Case I : Suppose vxy contains one of the symbol in Σ.

  Suppose vxy contains at least one a.

  Then uv^2xy^2z will more a's than b or c.

  Therefore uv^2xy^2z does not belong to L.

  This is a contradiction of the pumping Lemma.

- Case II : Suppose vxy contains two of the symbol in Σ.

  Suppose vxy contains at least one a & at least one b.

  Then uv^2xy^2z will have more a's & b's than c.

  Therefore uv^2xy^2z does not belong to L.

  This is contradiction of the pumping lemma.

In all the cases we get a contradiction.

Therefore our assumption that L is context free is wrong.

06 Anikate Agrawal 201500088

4. Eliminate null productions

i) $S \to aSb \mid aAb \mid ab \mid a$ $\quad A \to \Lambda$

$V_1 = \{A\}$

$V_2 = \{A\} = V_1$

P' : $S \to aSb \mid aAb \mid ab \mid a$

(ii) $S \to aXbX$ $\quad X \to aY \mid bY \mid \Lambda$ $\quad Y \to X \mid d$

$V_1 = \{X\}$

$V_2 = \{X, Y\}$

$V_3 = \{X, Y\} = V_2$

P' : $S \to aXbX \mid abX \mid aXb \mid ab$

$X \to aY \mid a \mid bY \mid b$

$Y \to X \mid d$

5. Eliminate null productions

$S \to AB$

$A \to aAA \mid \Lambda$

$B \to bBB \mid \Lambda$

$V_1 = \{A, B\}$

$V_2 = \{S, A, B\}$

$V_3 = \{S, A, B\} = V_2$

P' ; $S \to AB \mid A \mid B$

$A \to aAA \mid aA \mid a$

$B \to bBB \mid bB \mid b$

06 Anikate Agrawal 201500088

5. Find if grammar is ambiguous or not.

$S \to 0S \mid 1AA$ $\quad$ $B \to 1 \mid 0BB$

$A \to 0 \mid 1A \mid 0B$ $\quad$ for string 0100110

$S \to 0S$
$\to 01AA$
$\to 010BA$
$\to 0100BBA$
$\to 01001BA$
$\to 010011A$
$\to 0100110$

LMD 1

$S \to 0S$
$\to 01AA$
$\to 010A$
$\to 0100B$
$\to 01001$

LMD 2 cannot derive string

So, the grammar is not ambiguous as not 2 left most derivative tree are possible.

7. For the grammar $S \to aAS \mid a$ $\quad$ $A \to SbA \mid SS \mid ba$.

Generate the string aabaaabbaaa. find :-

a) LMD

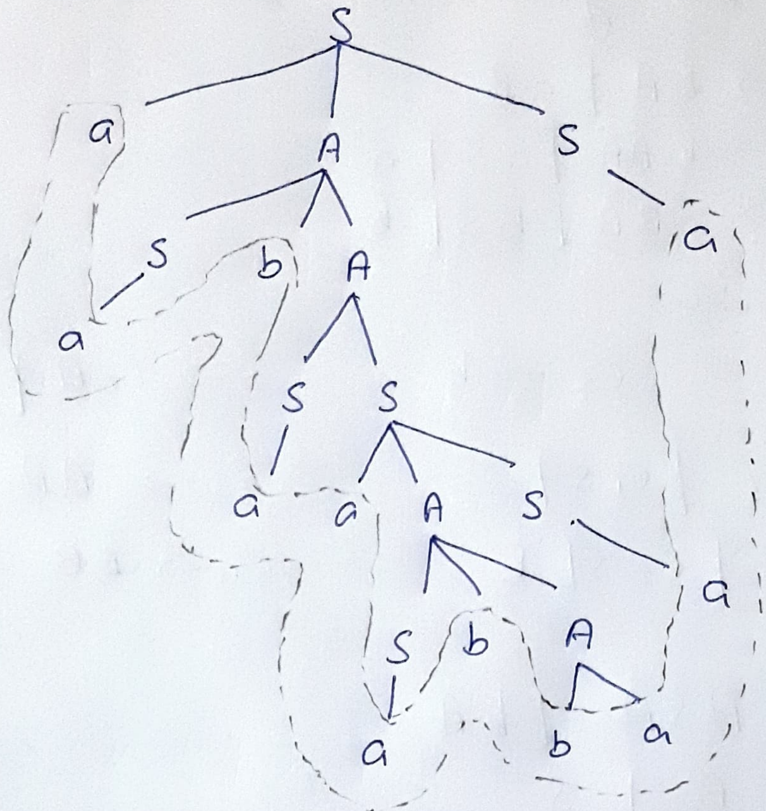
$S \to aAS$
$\to aSbAS$
$\to aabAS$
$\to aabSSS$
$\to aabaSS$
$\to aabaaASS$
$\to aabaaSbASS$
$\to aabaaabASS$
$\to aabaaabbaSS$
$\to aabaaabbaaS$
$\to aabaaabbaaa$

b) RMD

$S \to aAS$
$\to aAa$
$\to aSbAa$
$\to aSbSSa$
$\to aSbSaASa$
$\to aSbSaAaa$
$\to aSbSaSbAaa$
$\to aSbSaSbbaaa$
$\to aSbSaabbaaa$
$\to aSbaaabbaaa$
$\to aabaaabbaaa$

06 Anikate Agrawal 201500088

c. Parse Tree



8. Reduce the given CFG into CNF.

$S \to \sim S \mid [SX] \mid p \mid q$ (S being only variable)

$S \to p \mid q$ (Already in CNF)

$S \to C_1 S$, $C_1 \to \sim$

$S \to C_2 C_3$, $C_2 \to [$

$C_3 \to S C_4$, $C_4 \to C_5 C_6$, $C_5 \to X$, $C_6 \to]$

Final Productions $P' \Rightarrow$

$S \to p \mid q \mid C_1 S \mid C_2 C_3$ $C_1 \to \sim$, $C_2 \to [$

$C_3 \to S C_4$ $C_4 \to C_5 C_6$, $C_5 \to X$, $C_6 \to]$

06 Anikate Agrawal 201500088

9. Reduce the given CFG into CNF

i) $S \to bA \mid aB$

$A \to bAA \mid aS \mid a$

$B \to aBB \mid bS \mid b$

Final productions

$C_1 \to a$ $\quad C_2 \to b$ $\quad S \to C_2A \mid C_1B$

$A \to C_2C_3 \mid C_1S \mid a$ $\quad C_3 \to AA$

$B \to C_1C_4 \mid C_2S \mid b$ $\quad C_4 \to BB$

ii) $S \to ASA \mid bA$

$A \to B \mid S$

$B \to a$

Eliminating unit productions

$S \to ASA \mid bA$

$A \to a \mid ASA \mid bA$

$B \to a$

Converting to CNF

$S \to AC_1 \mid C_2A$ $\quad C_1 \to SA$ $\quad C_2 \to b$

$A \to a \mid AC_1 \mid C_2A$

$B \to a$

06 Anikate Agrawal 20150088

10. Convert the grammar into GNF.

i) $S \to AB$ $A \to BS \mid a$ $B \to SA \mid b$

Renaming S, A, B as A_1, A_2, A_3

$A_1 \to A_2 A_3$
$A_2 \to A_3 A_1 \mid a$
$A_3 \to A_1 A_2 \mid b$

i) A_1 - production $A_1 \to A_2 A_3$ is in required form.

ii) A_2 - production $A_2 \to A_3 A_1 \mid a$ are in required form.

iii) $A_3 \to b$ is in required form

Applying lemma 1 to $A_3 \to A_1 A_2$

$A_3 \to A_2 A_3 A_2$
$\to A_3 A_1 A_3 A_2 \mid a A_3 A_2$

lemma 2 to $A_3 \to A_3 A_1 A_3 A_2$

$\boxed{A_3 \to b \mid a A_3 A_2 \mid b Z_3 \mid a A_3 A_2 Z_3}$

$Z_3 \to A_1 A_3 A_2 \mid A_1 A_3 A_2 Z_3$

$\boxed{A_2 \to a}$

$A_2 \to A_3 A_1$ Applying lemma 1

$\boxed{A_2 \to b A_1 \mid a A_3 A_2 A_1 \mid b Z_3 A_1 \mid a A_3 A_2 Z_3 A_1}$

Applying Lemma 1 to $A_1 \to A_2 A_3$

$\boxed{A_1 \to b A_1 A_3 \mid a A_3 \mid a A_3 A_2 A_1 A_3 \mid b Z_3 A_1 A_3 \mid a A_3 A_2 Z_3 A_1 A_3}$

06 Anikate Agrawal 201500088

$Z_3 \rightarrow A_1 A_3 A_2 \mid A_1 A_3 A_2 Z_3$ Applying lemma 1

$$\boxed{\begin{array}{l} Z_3 \rightarrow aA_3A_3A_2 \mid aA_3A_3A_2Z_3 \\ Z_3 \rightarrow bA_1A_3A_3A_2 \mid bA_1A_3A_3A_2Z_3 \\ Z_3 \rightarrow aA_3A_2A_1A_3A_3A_2 \mid aA_3A_2A_1A_3A_3A_2Z_3 \\ Z_3 \rightarrow bZ_3A_1A_3A_3A_2 \mid bZ_3A_1A_3A_3A_2Z_3 \\ Z_3 \rightarrow aA_3A_2Z_3A_1A_3A_3A_2 \mid aA_3A_2Z_3A_1A_3A_3A_2Z_3 \end{array}}$$

ii) $S \rightarrow ABb \mid a$

$A \rightarrow aaA$

$B \rightarrow bAb$

$C_1 \rightarrow a$, $C_2 \rightarrow b$

$S \rightarrow ABC_2 \mid a$

$A \rightarrow aC_1A$

$B \rightarrow bAC_2$

Applying lemma 1 to $S \rightarrow ABC_2$

$$S \rightarrow aC_1ABC_2$$

Final productions in GNF are:-

$S \rightarrow aC_1ABC_2 \mid a$

$A \rightarrow aC_1A$

$B \rightarrow bAC_2$

$C_1 \rightarrow a$

$C_2 \rightarrow b$

06 Anikate Agrawal 20150008

11. Eliminate unit production from the given grammar and convert it into GNF.

$E \to E+T \mid T$

$T \to T*F \mid F$

$F \to (E) \mid a$

Eliminating unit productions

$E \to E+T \mid T*F \mid (E) \mid a$

$T \to T*F \mid (E) \mid a$

$F \to (E) \mid a$

To convert to CNF we introduce new variables A, B, C corresponding to +, *,)

The modified productions are

$E \to EAT \mid TBF \mid (EC \mid a$

$T \to TBF \mid (EC \mid a$

$F \to (EC \mid a$

$A \to +$, $B \to *$, $C \to)$

The variables A, B, C, F, T and E are renamed as $A_1, A_2, A_3, A_4, A_5, A_6$. Then the productions become

$\boxed{A_1 \to + \quad , \quad A_2 \to * , \quad A_3 \to) , \quad A_4 \to (A_6A_3 \mid a}$

$A_5 \to A_5A_2A_4 \mid (A_6A_3 \mid a$

$A_6 \to A_6A_1A_5 \mid A_5A_2A_4 \mid (A_6A_3 \mid a$

06 Anikate Agrawal 201500088

We have to modify only A_5 and A_6 productions

$A_5 \rightarrow A_5 A_2 A_4$ Applying lemma 2

$$\boxed{A_5 \rightarrow (A_6 A_3 \mid a \mid (A_6 A_3 Z_5 \mid a Z_5}$$

$Z_5 \rightarrow A_2 A_4 \mid A_2 A_4 Z_5$

$A_6 \rightarrow A_5 A_2 A_4$ Applying Lemma 1

$A_6 \rightarrow (A_6 A_3 A_2 A_4 \mid a A_2 A_4 \mid (A_6 A_3 Z_5 A_2 A_4 \mid a Z_5 A_2 A_4$

$A_6 \rightarrow (A_6 A_3 \mid a$ are in proper form

$A_6 \rightarrow A_6 A_1 A_5$ applying lemma 1

$$\boxed{\begin{array}{l} A_6 \rightarrow (A_6 A_3 A_2 A_4 \mid a A_2 A_4 \mid (A_6 A_3 Z_5 A_2 A_4 \mid a Z_5 A_2 A_4 \mid (A_6 A_3 \mid a \\ A_6 \rightarrow (A_6 A_3 A_2 A_4 Z_6 \mid a A_2 A_4 Z_6 \mid (A_6 A_3 Z_5 A_2 A_4 Z_6 \mid a Z_5 A_2 A_4 Z_6 \mid (A_6 A_3 Z_6 \mid a Z_6 \end{array}}$$

$Z_6 \rightarrow A_1 A_5 \mid A_1 A_5 Z_6$ Lemma 1

$$\boxed{Z_6 \rightarrow + A_5 \mid + A_5 Z_6}$$

$Z_5 \rightarrow A_2 A_4 \mid A_2 A_4 Z_5$ lemma 1

$$\boxed{Z_5 \rightarrow * A_4 \mid * A_4 Z_5}$$

The productions in box are the required GNF form.

06 Anikate Agrawal 201500088

12. Eliminate unit production from the given grammar.

$S \rightarrow AB$ $\quad A \rightarrow a$ $\quad B \rightarrow C \mid b$ $\quad C \rightarrow D$ $\quad D \rightarrow E$ $\quad E \rightarrow a$

Final productions are :-

$S \rightarrow AB$ $\quad A \rightarrow a$ $\quad B \rightarrow a \mid b$ $\quad C \rightarrow a$ $\quad D \rightarrow a$ $\quad E \rightarrow a$

13. Remove the useless symbol from the given context free grammar.

$S \rightarrow aB \mid bX$

$A \rightarrow BAd \mid bSX \mid a$

$B \rightarrow aSB \mid bBX$

$X \rightarrow SBD \mid aBAX \mid ad$

Phase 1 $\quad V_1 = \{A, X\}$

$V_2 = \{S, A, X\}$

$V_3 = \{S, A, X\} = V_2$

$P' \Rightarrow$ $\quad S \rightarrow \cancel{aB \mid}\ bX$

$A \rightarrow \cancel{BAd \mid}\ bSX \mid a$

$X \rightarrow ad$

Phase 2 : $\quad V_1 = \{S\}$

$V_2 = \{S, X\}$

$V_3 = \{S, X\} = V_2$

$P'' \Rightarrow$ $\quad S \rightarrow bX$

$X \rightarrow ad$

06 Anikate Agrawal 201500088

14. Consider a context free grammar G with the following productions :

$S \rightarrow 1S1 \mid T$

$T \rightarrow 1X1 \mid X$

$X \rightarrow 0X0 \mid 1$

i) Write four strings of L(G)

$S \rightarrow 1S1$

$\rightarrow 1T1$

$\rightarrow 1X1$

$\rightarrow \boxed{111}$

$S \rightarrow T$

$\rightarrow X$

$\rightarrow \boxed{1}$

$S \rightarrow 1S1$

$\rightarrow 1T1$

$\rightarrow 11X11$

$\rightarrow \boxed{11111}$

$S \rightarrow T$

$\rightarrow X$

$\rightarrow 0X0$

$\rightarrow \boxed{010}$

(ii) Give an example of a string $w \in \{0,1\}^*$ such that $|w| > 7$ and $w \notin L(G)$

0000000 is the required string

15. Give a CFG for the language $L = \{x \in \{a,b\}^* \mid$ x starts and ends with different symbols

Ans $S \rightarrow aAb \mid bAa$

$A \rightarrow \wedge \mid aA \mid bA \mid a \mid b$