Regular Grammar

$\alpha \to \beta$

$\alpha \equiv$ single Non Terminal

$\beta \in (V+T)^*$ $\quad |\beta| \leq 2$

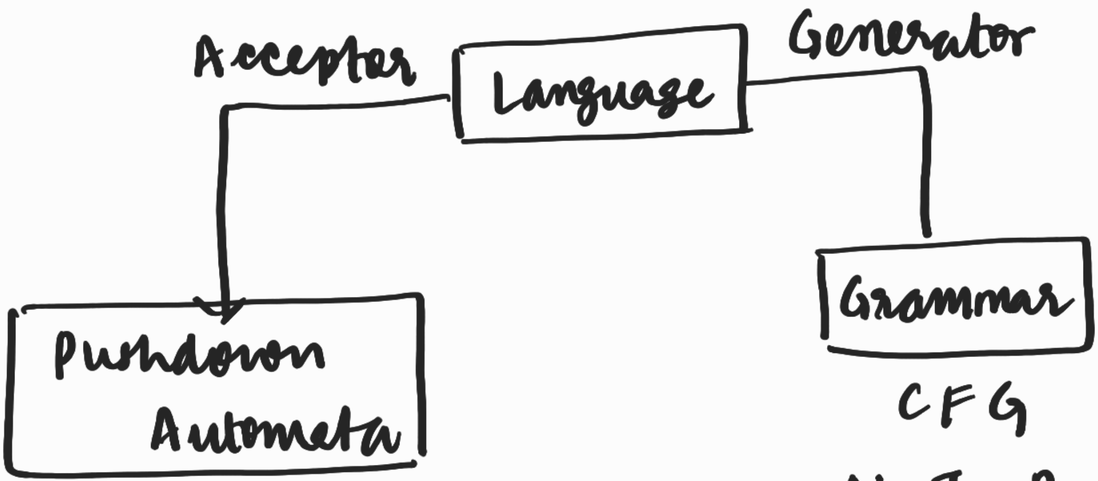
combination of terminus & non terminus

i.e $A \to \beta x$ or $A \to x\beta$

Type 2 Context Free Grammar

$\alpha \to \beta$ $\quad \alpha \Rightarrow$ single NT

$B \to (V+T)^*$



CFG

N, T, P, S

Eg $\to$ $\{ S \to aSB$

$S \to aB$

$B \to b \}$

N: $\{S, B\}$

T: $\{a, b\}$

$S \to S$

$S \to aSB \to aaBB \to aabB \to aabb$

Sentential / Sequential form

Q $L = \{aa, ab, ba, bb\}$

$S \to aa \mid ab \mid ba \mid bb$

Q $(a|b) = \{a, b\}$

$(a|b)(a|b)$

$S \to AA$

$A \to a|b$

Q $L = \{a^n \mid n \geq 0\}$

$\epsilon, a, aa, \ldots\ldots$

$A \to \epsilon \mid aA$ (or) $A \to \epsilon \mid Aa$

$A \to aA \to aaA \to aaaA \to aaaaA$

Q $(a|b)^*$ ϵ, a, b, aaa, bba

$S \to aS \mid bS \mid \epsilon$

let string be abab

$S \to aS \to abS \to abaS \to ababS \to abab$

Q set of all strings of length atleast 2

mod(w) = 2

$(a|b)(a|b)(a|b)^*$

$S \to AAB$

$A \to a|b$

$B \to aB|bB|\varepsilon$

Q Strings of length atmost 2.

$\varepsilon|a|b|aa|ab|ba|bb$

$\to (a|b|\varepsilon)(a|b|\varepsilon)$

$S \to AA$

$A \to a|b|\varepsilon$

Q starting with 'a' and ending with 'b'

$a(a|b)^* b$

$S \to aAb$

$A \to aA|bA|\epsilon$

Q Starts with 'a' and ends with 'b' or starts with 'b' and ends with 'a'

$$(a(a|b)^* b) \mid (b(a|b)^* a)$$

$S \to aAb \mid bAa$

$A \to aA \mid bA \mid \epsilon$

Q starts and ends with the same symbol

$a \mid b \mid a(a|b)^* a \mid b(a|b)^* b$

$S \to aAa \mid bAb \mid a \mid b$

$A \to aA \mid bA \mid \epsilon$

Q $a^n b^n \mid n \geq 1$

$S \to aSb \mid ab$

Q Palindrome word (word ^{reverse})

$\Sigma = \{a, b\}$

$S \to aSa \mid bSb \mid \epsilon \mid a \mid b$

Q generating even length string

$S \to BS \mid \epsilon$ $\qquad ((a|b)(a|b))^*$

$A \to a \mid b$ A should be repeated twice

$B \to AA$ B should be repeated n times

Q $a^n b^m \mid n, m \geq 1$

$S \to AB$

$A \to aA \mid a$

$B \to bB \mid b$

Q $a^n b^n c^m \mid n, m \geq 1$

$S \to AB$

$A \to aAb \mid ab$

$B \to cB \mid c$

Q $a^n b^m c^n \mid n, m \geq 1$

$S \to aSc \mid aBc$

$B \to bB \mid b$

Q $a^n b^{2n} \mid n \geq 1$

$S \to aSbb \mid abb$

Q $a^n b^m c^m d^n \mid n \geq 1$

$S \to aSd \mid aBd$

$B \to bBc \mid bc$

Q $a^n b^n c^m d^m \mid n, m \geq 1$

$S \to AB$

$A \to aAb \mid ab$

$B \to cBd \mid cd$

Q $a^n b^{n+m} c^m \mid n, m \geq 1$

$a^n b^n b^m c^m \mid n, m \geq 1$

$S \to AB$

$A \to aAb \mid ab$

$B \to bBc \mid bc$

Q $a^n b^n c^n \mid n \geq 1$

Not a context free language.

Q $a^{m+n} b^m c^n \mid m, n \geq 1$

Q $a^m b^m c^{n+m} \mid m, n \geq 1$

Q $a^n b^m c^n d^m \mid m, n \geq 1$

Q $a^n a^m b^m c^n \mid n, m \geq 1$

$S \to aSc \mid aAc$

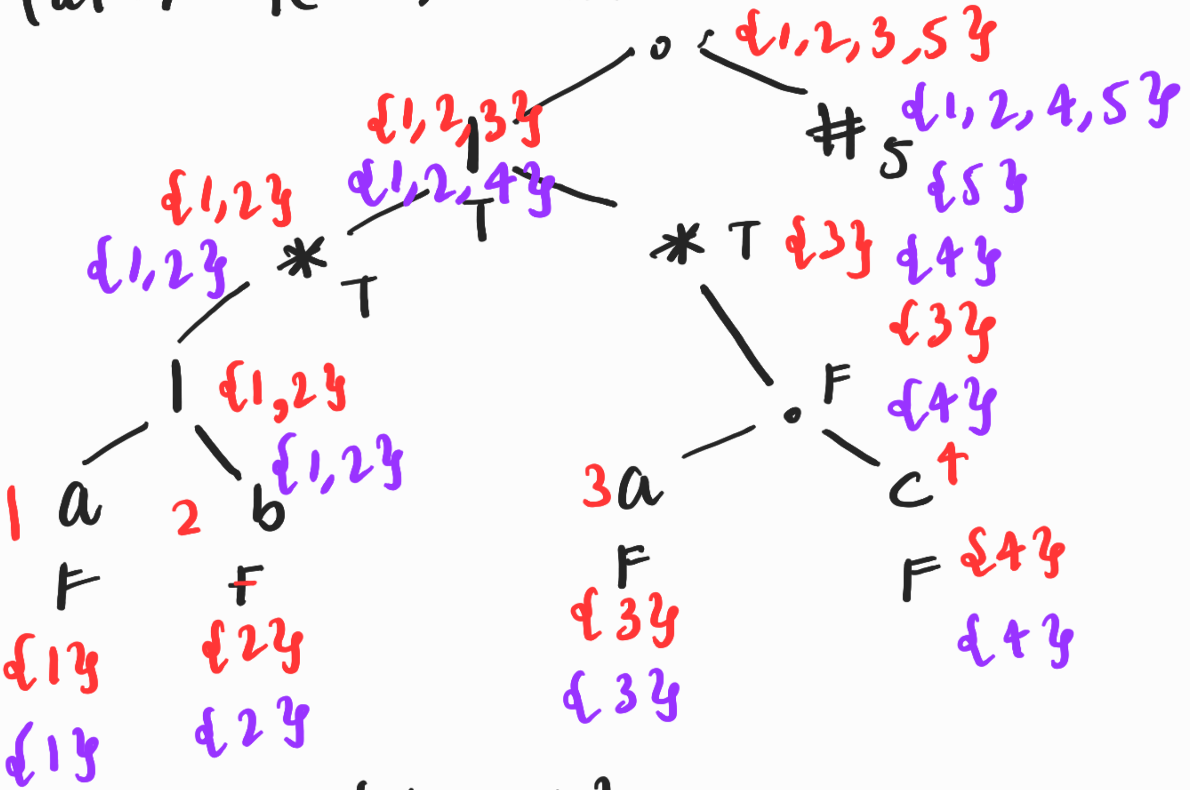
$A \to aAb \mid ab$ $(a^m b^m)$

Q $a^m b^m c^m c^n$

$S \to aSc \mid aAc$

$A \to bAc \mid bc$

Q $(a|b)^* \mid (a \cdot c)^*$ Make DFA



followpos (1) = {1, 2, 5}

followpos (2) = {1, 2, 5}

followpos (3) = {4}

followpos (4) = {3, 5}

followpos (5) = {ϕ}

$\{1, 2, 3, 5\} \equiv A$ start state

$$Dtran[A, a] = followpos(1) \cup followpos(3) = \{1, 2, 4, 5\} \quad B$$

$$Dtran[A, b] = followpos(2) = \{1, 2, 5\} \quad C$$

$$Dtran[A, c] = \phi$$

$$Dtran[B, a] = followpos(1) = \{1, 2, 5\} = C$$

$$Dtran[B, b] = followpos(2) = \{1, 2, 5\} = C$$

$$Dtran[B, c] = followpos(4) = \{3, 5\} \quad D$$

$$Dtran[C, a] = followpos(1) = \{1, 2, 5\}$$

$$Dtran[C, b] = followpos(2) = \{1, 2, 5\}$$

$$Dtran[C, c] = \phi$$

$$Dtran[D, a] = followpos(3) = \{4\} = E$$

$$Dtran[D, b] = \phi$$

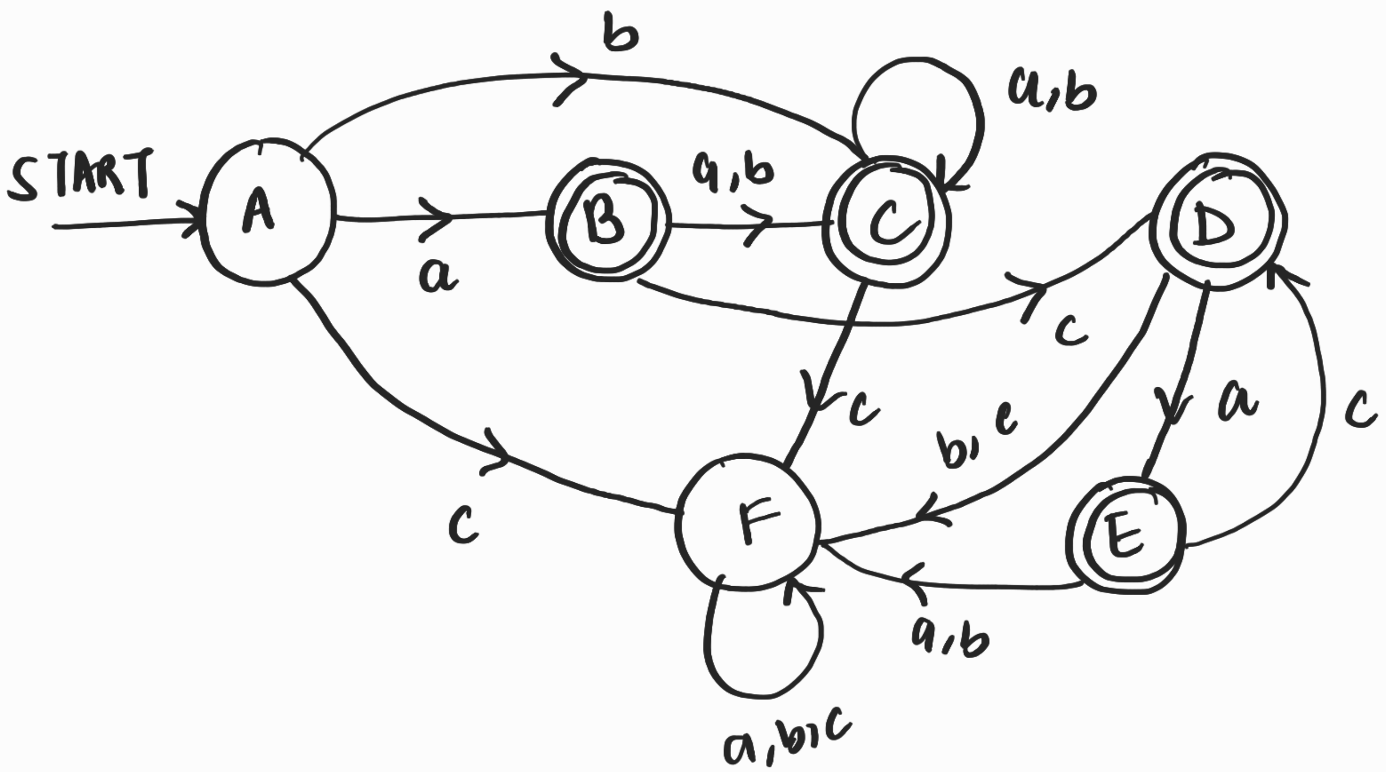
$$Dtran[D, c] = \phi$$

$$Dtran[E, a] = \phi$$

$$Dtran[E, b] = \phi$$

$$Dtran[E, c] = followpos(4) = \{3, 5\}$$

	a	b	c
→ A	B	C	F
* B	C	C	D
* C	C	C	F
* D	E	F	F
* E	F	F	D
F	F	F	F



Chomsky Normal Form

$E \to E + T \mid T$

$T \to T * F \mid F$

$F \to (E) \mid id$

(remove ambiguity in grammar)

$A \to BC$

$A \to a$

id + id * id

Left most derivation → reduce leftmost Non terminals first.

E → E + T → T + T

→ F + T

→ id + T

→ id + T * F

→ id + id * F

→ id + id * id

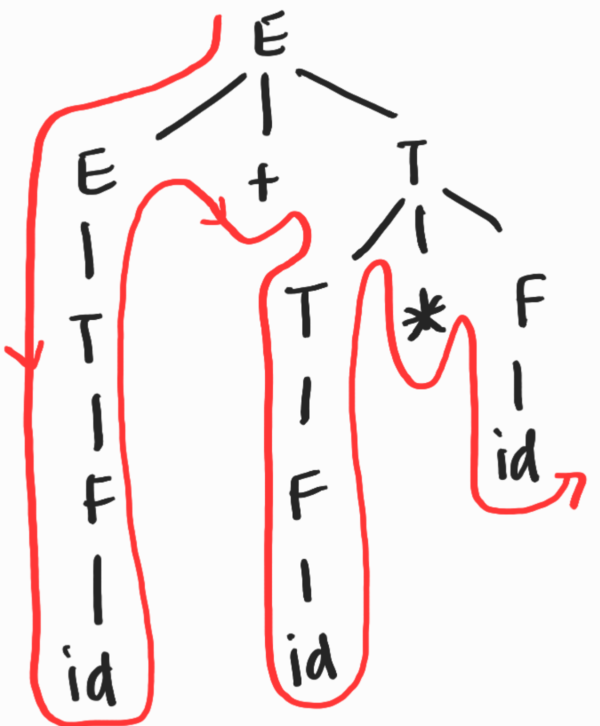
RMD E → E + T

→ E + T * F → E + T * id → E + F * id

→ E + id * id → T + id * id → F + id * id

→ id + id * id

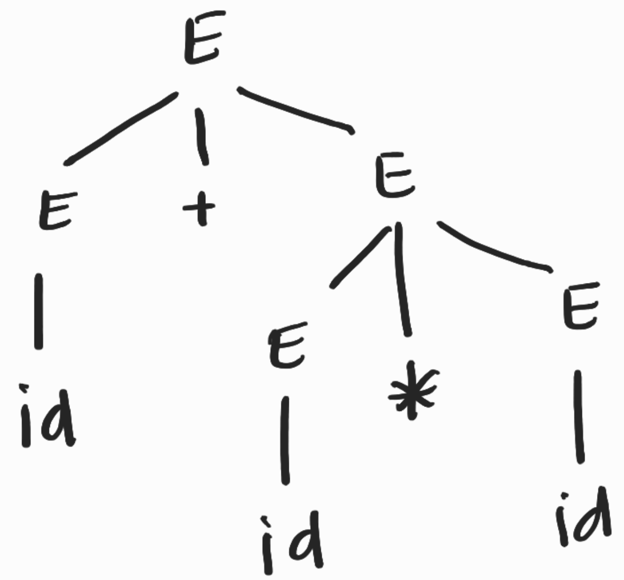
Parse Tree

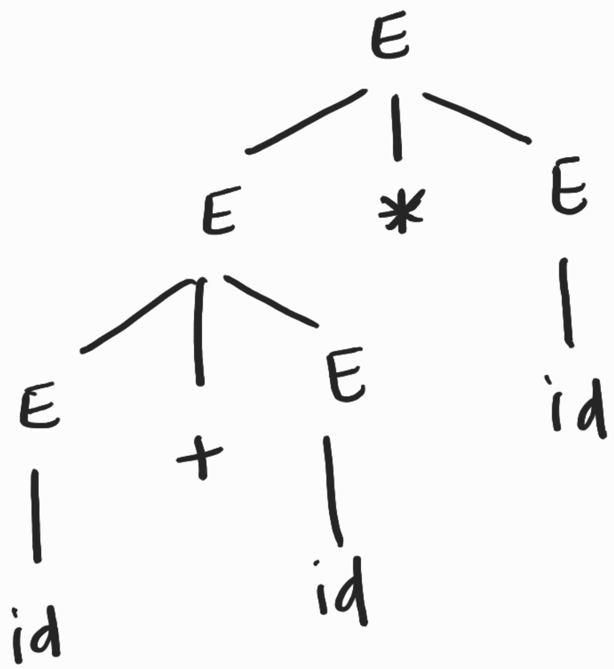


Ambiguity- For a single grammar if we are able to generate 2 or more parse trees then it is called ambiguity.

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

string → $id+id*id$





Chomsky Normal Form (CNF)

$A \rightarrow BC$ or $A \rightarrow a$

* 2 non terminals or 1 terminal at RHS.

Simplification of CFG:-

1. eliminate the useless symbols.
   - Non Generating Symbol
   - Non Reachable symbol
2. eliminate ϵ production
3. eliminate unit production

1. Eg → $S \rightarrow AB \mid a$

   $A \rightarrow b$

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

   $T = \{a, b\}$

   Start symbol = $\{a, b\}$

S & A are generating

B has no production → non generating

→ useless.

$S \rightarrow a$

$A \rightarrow b$

Since A is non reachable from start state

Thus A is useless

→ $S \rightarrow a$

2 Eliminating ϵ productions

$S \to AB$

$A \to aAA \mid \epsilon$

$B \to bBB \mid \epsilon$

A & B are having empty productions

If $A \to \epsilon$ $B \to \epsilon$ $\Rightarrow$ $S \to \epsilon\epsilon \to \epsilon$

$S \to AB \mid B \mid A$

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

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

3 Eliminating unit production : $A \to B$

where A & B are Non Terminals.

$S \to AB$

$A \to a$

$B \to C \mid b$

$C \to D$

$D \to E$

$E \to a$

unit productions

$B \to C$

$C \to D$

$D \to E$

$B \to a \mid b$

non reachable states $\left\{ \begin{array}{l} C \to a \\ D \to a \\ E \to a \end{array} \right\}$

→ $S \to AB$

$A \to a$

$B \to a|b$

Q $S \to a|b|Sa|Sb|S0|S1$

$F \to S|(E)$

$T \to F|T*F$ → E is the start symbol

$E \to T|E*T$

! unit Productions

$E \to T$

$T \to F$

$F \to S \to a|b|Sa|Sb|S0|S1$

$S \to a|b|Sa|Sb|S0|S1$

$F \to a|b|Sa|Sb|S0|S1|(E)$

$T \to a|b|Sa|Sb|S0|S1|(E)|T*F$

$E \to a|b|Sa|Sb|S0|S1|(E)|T*F|E*T$

Safe Order to Simplify CFG:-

(1) eliminate ∈ productions

(2) eliminate unit production

(3) eliminate useless productions

Conversion to CNF

Q $S \to TU \mid V$

$T \to aTb \mid \epsilon$

$U \to cU \mid \epsilon$

$V \to aVc \mid W$

$W \to bW \mid \epsilon$

1 eliminate ϵ productions

T, U, W are nullable

$V \to W \Rightarrow V$ is nullable

$S \to TU/V \Rightarrow S$ is nullable

$S \to TU \mid U \mid T \mid V$

$T \to aTb \mid ab$

$U \to cU \mid c$

$V \to aVc \mid ac \mid W$

$W \to bW \mid b$

2 Unit Productions

$S \to T \mid U \mid V$

$V \to W$

$S \to TU \mid aTb \mid ab \mid cU \mid c \mid aVc \mid ac \mid bW \mid b$

$T \to aTb \mid ab$

$U \to cU \mid c$

$V \to aVc \mid ac \mid bW \mid b$

$W \to bW \mid b$

Introduce 3 NTs $X_a\,(a)$

$X_b\,(b)$

$X_c\,(c)$

3 Useless Productions

No useless productions.

— x —

$S \to TU \quad X_a T X_b \mid X_a X_b \mid X_c U \mid c \mid X_a V X_c \mid X_a X_c \mid X_b W \mid b$

$T \to X_a T X_b \mid X_a X_b$

$U \to X_c U \mid c$

$V \to X_a V X_c \mid X_a X_c \mid X_b W \mid b$

$W \to X_b W \mid b$

$X_a \to a$

$X_b \to b$

$X_c \to c$

— x —

$S \to TU \mid X_a Y_1 \mid X_a X_b \mid X_c U \mid c \mid X_a Y_2 \mid X_a X_c \mid X_b W \mid b$

$Y_1 \to T X_b$

$Y_2 \to V X_c$

$T \to X_a Y_1 \mid X_a X_b$

$U \to X_c U \mid c$

$W \to X_a Y_2 \mid X_a X_c \mid X_b W \mid b$

$W \to X_b W \mid b$

$X_a \to a$

$X_b \to b$

$X_c \to c$

Q

$S \to aXbX$

$X \to aY \mid bY \mid \epsilon$

$Y \to X \mid c$

1 eliminate null productions

X is nullable

Y becomes nullable as $Y \to X$

$S \to aXbX \mid aXb \mid abX \mid ab$

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

$Y \to X \mid c$

2 unit productions

$Y \to X$

$Y \to aY \mid bY \mid a \mid b \mid c$

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

$S \to aXbX \mid aXb \mid abX \mid ab$

3 No useless productions

$S \to E_1E_2 \mid E_1D_b \mid D_aE_2 \mid D_aD_b$

$X \to D_aY \mid D_bY \mid a \mid b$

$Y \to D_aY \mid D_bY \mid a \mid b \mid c$

$D_a \to a$

$D_b \to b$

$E_1 \to D_aX$

$E_2 \to D_bX$

Q $S \to AbA$

$A \to Aa \mid \epsilon$

1 ϵ productions

A is nullable

$S \to AbA \mid Ab \mid bA \mid b$

$A \to Aa \mid a$

2. No unit productions

3. No useless productions

$S \to AP_1 \mid AX_b \mid X_bA \mid b$

$A \to AX_a \mid a$

$X_a \to a$

$X_b \to b$

$P_1 \to X_bA$

Greibach's Normal Form

$A \to b$ or $A \to bC_1C_2 \ldots C_n$

b is terminial

$A, C_1, C_2 \ldots C_n$ or non terminials

1. Grammar Should be simplified
2. Grammar should be in CNF.

Q. $S \to CA \mid BB$

$B \to b \mid SB$

$C \to b$

$A \to a$

Already in CNF.

S1 Rename all non terminals as A_i

$S \Rightarrow A_1$

$C \Rightarrow A_2$

$A \Rightarrow A_3$

$B \Rightarrow A_4$

$A_1 \rightarrow A_2 A_3 \mid A_4 A_4$

$A_4 \rightarrow b \mid A_1 A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

If we have a production of the form

$A_i \rightarrow A_j x \Rightarrow i < j$

$A_1 \rightarrow A_2 A_3 \mid A_4 A_4$ satisfies

$A_4 \rightarrow b \mid A_1 A_4$ does not satisfy

place productions of A_1

only check this

$A_4 \rightarrow b \mid A_2 A_3 A_4 \mid A_4 A_4 A_4$

$A_4 \rightarrow b \mid b A_3 A_4 \mid A_4 A_4 A_4$

left recursion $(i = j)$

$A \rightarrow A\alpha \mid \beta$

To reduce left recursion :- Introduce additional Non Terminals

$A' \rightarrow \alpha A' \mid \alpha$

$A \rightarrow \beta A' \mid \beta$

$$\underbrace{A_4}_{A} \rightarrow \underbrace{b \mid bA_3A_4}_{\beta} \mid \underbrace{A_4}_{A}\ \underbrace{A_4A_4}_{\alpha}$$

Q left recursion eg

$$\underbrace{E}_{A} \rightarrow \underbrace{E}_{A}\underbrace{+T}_{\alpha} \mid \underbrace{T}_{\beta}$$

$E' \rightarrow +TE' \mid +T$

$E \rightarrow TE' \mid T$

$A_4' \rightarrow A_4A_4A_4' \mid A_4A_4$

$A_4 \rightarrow (b \mid bA_3A_4)A_4' \mid b \mid bA_3A_4$

$A_4 \rightarrow bA_4' \mid bA_3A_4A_4' \mid b \mid bA_3A_4$

$A_1 \rightarrow A_2A_3 \mid A_4A_4$

$A_4 \rightarrow bA_4' \mid bA_3A_4A_4' \mid b \mid bA_3A_4$

$A_4' \rightarrow A_4A_4A_4' \mid A_4A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

$A_1 \rightarrow bA_3 \mid bA_4'A_4 \mid bA_3A_4A_4'A_4 \mid bA_4 \mid bA_3A_4A_4$

$A_4 \rightarrow bA_4' \mid bA_3A_4A_4' \mid b \mid bA_3A_4$

$A_4' \rightarrow bA_4'A_4A_4' \mid bA_3A_4A_4'A_4A_4' \mid bA_4A_4' \mid bA_3A_4A_4A_4' \mid bA_4'A_4 \mid bA_3A_4A_4'A_4 \mid bA_4 \mid bA_3A_4A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

Q $S \to AY \mid ZA \mid AZ \mid b$

$Y \to ZA$

$A \to AX \mid a$

$X \to a$

$Z \to b$

already in CNF

1 Renaming Variables

$S \Rightarrow P_1$

$A \Rightarrow P_2$

$Y \Rightarrow P_3$

$Z \Rightarrow P_4$

$X \Rightarrow P_5$

$P_1 \to P_2 P_3 \mid P_4 P_2 \mid P_2 P_4 \mid b$

$P_3 \to P_4 P_2$

$P_2 \to P_2 P_5 \mid a$

$P_5 \to a$

$P_4 \to b$

2 $A_i \to A_j x \to i < j$

$P_1 \to P_2 P_3 \mid P_4 P_2 \mid P_2 P_4 \mid b$ ✓✓

$P_3 \to P_4 \; P_2$ ✓✓

$P_2 \to P_2 P_5 \mid a$ left recursion

$A \quad A \; \alpha \; \beta$

$A' \to \alpha A' \mid \alpha$

$A \to \beta A' \mid \beta$

$P_2' \to P_5 P_2' \mid P_5$

$P_2 \to a \; P_2' \mid a$

//

$P_1 \to P_2 P_3 \mid P_4 P_2 \mid P_2 P_4 \mid b$

$P_3 \to P_4 P_2$

$P_2' \to P_5 P_2' \mid P_5$

$P_2 \to a \; P_2' \mid a$

$P_5 \to a$

$P_4 \to b$

//

$P_1 \to a P_2' P_3 \mid a P_3 \mid b P_2 \mid a P_2' P_4 \mid a P_4 \mid b$

$P_3 \to b P_2$

$P_2' \to a P_2' \mid a$

$P_2 \to a \; P_2' \mid a$

$P_5 \to a$

$P_4 \to b$