a prime

Tuesday, 8 December, 2020 10:51 AM

P.T. $L = \{a^p \mid p \text{ is a prime}\}$ is not regular.

$= \{a^2, a^3, a^5, a^7, a^{11}, - - \}$

$= \{aa, aaa, aaaaa, - - - \}$

Step I : Let L be a RL and $\underline{n}$ be the states in FA M accepting L.

Step II : Let $w = a^p \in L$, $p \geq n$

$|w| = p \geq n$

aaaaa $\in L$

$w = xyz$, $\underbrace{|xy| \leq n, \ |y| > 0}$

$y = a^m \Rightarrow |y| = m$

$1 \leq m \leq n$

$\underbrace{|xyz| = p}$

$\underline{1} \leq |y| \leq \underline{n}$

Step III : $\underbrace{xy^iz \notin L \text{ for any suitable } i}$

y^i $\quad$ $a^3 = aaa$

yyy $\quad$ $- i$

$|xy^iz| = |xyz| + (i-1)|y|$

$= \underbrace{p + (i-1)m}$

$\underbrace{i = p+1}$

$xy^iz \notin L$ $= p + (p+\cancel{1}-\cancel{1})m$

$= p + pm$

$\underbrace{|xy^iz|} = p(1+m)$ for $\underbrace{i = p+1}$

not a prime

$0 \leq i = p+1 \leq \infty$

$xy^iz \notin L$ for $i = p+1$

$\therefore L$ is not a regular language

Q: P.T. $L = \{a^k b^k \mid k \ge 1\}$ is not regular.

$= \{ab, aabb, aaabbb, aaaabbbb, ---\}$

Step I — — —

Step II $w = a^n b^n \in L \Rightarrow w = xyz$, $\underbrace{|xy| \le n, \ |y| > 0}$

$|w| = 2n > n$ $\qquad 1 \le |y| \le n$ $\qquad |xyz| \ge n$

$|xyz| = 2n \Rightarrow |x| + |y| + |z| = 2n$ $\qquad \underbrace{aaaabbbb}$ $\qquad a^n b^n$ $\quad |xy| \le n$

$y = a^l \mid 1 \le l \le n$

Step III $\underbrace{xy^0z} = \frac{xyz}{y} = \frac{a^n b^n}{a^l} = a^{n-l} b^n$

$n - l \ne n$

$xy^0z = xz \notin L$

$xy^iz \notin L$ for $i = 0$

∴ L is not a regular language

$xyz \in L$ as $xy^iz \in L$ $\forall i \ge 0$

Input Buffering

To Do ✓ $\qquad$ greq, ×

else

area, x

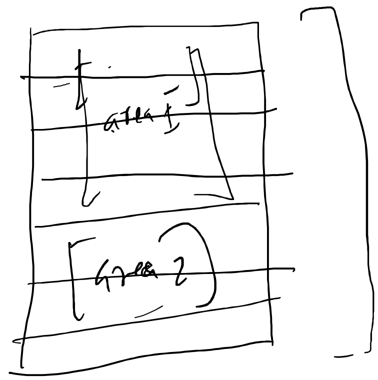
Lexeme Beginning →

Forward →

int area1, area2;

Symbol table

name	type	size	starting location
		128	0
area1	int	4	500
area2	int		



2^n bytes

Error recovery actions

1) deleting an extra character

foro(i=1;i<=10;i++) ⇒ for(i=1;i<=10;i++)

extra character → delete extra character

2) fr(" " ") ⇒ for(i=1; " ")

missing → Insert a missing character

3) fir(" " ") ⇒ for(" " ")

incorrect character ⇒ replaced with correct charact

4) fro(i=1;i<=10;i++) ⇒ for(" " ")

↳ transposing adjacent characters

Panic mode recovery

for(i=1 ; i<=10 ; i++)

Regular definitions of tokens

keyword → else | if | — —

if → if
else → else
letter → a|b|---|z|A|B|---|Z → letter [a-zA-Z]
digit → 0|1|---|9 → digit [0-9]
relop → <|<=|>|>=|<>|= → relop [<|<=|>|>=|<>|=]
id → letter (letter/digit)* → id [letter(letter+digit)*]

$R^+ = RR^*$

a^+

a^*

num → digit$^+$ (. digit$^+$)? (E(+/-)? digit$^+$)?

130.0

only → 105

107.25

Int

~~130.47.21~~

107E10
+10

107E-20 = 107^{-20}
Integer and Exp

108.20E-10 = 108.20^{-10}
Int dec Exp digit+

delim → blank/ tab/ newline
ws → delim$^+$

int else;

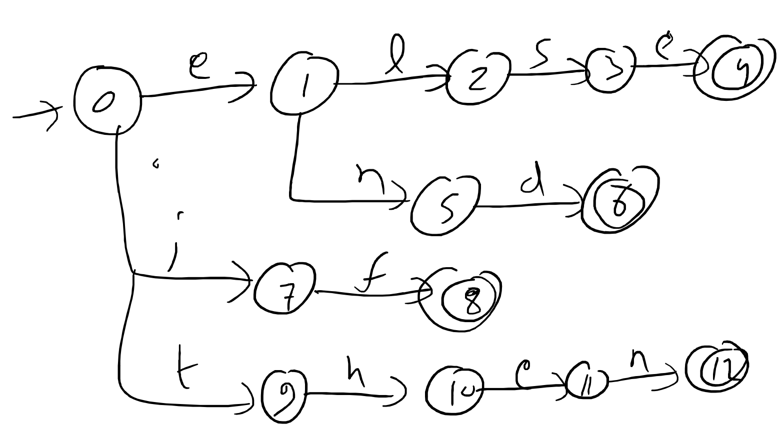
end , then
if

Input Buffering

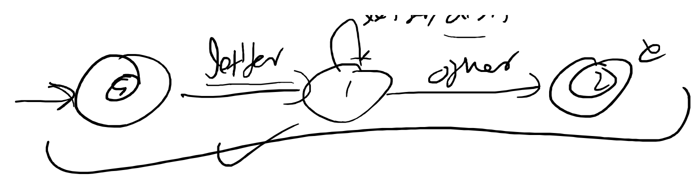
Fixed e l s e 1

Forward

int else1



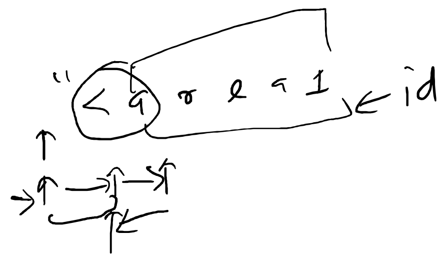
< > → HE



Fixed Forward →

LE

→ HE



Fix forward →

LEX: Auxiliary Definitions

$D_1 = R_1$

$D_2 = R_2$

Each D_i is a distinct name and each R_i is a RE formed using $\Sigma \cup \{D_1, D_2, \ldots D_{i-1}\}$

letter (D_1)	[a-zA-Z] (R_1)
digit (D_2)	[0-9] (R_2)
id (D_3)	[letter(letter+digit)*] (R_3)

Translation Rules

P_1 {action1}

P_2 {action2}

P_3 {action3}

|

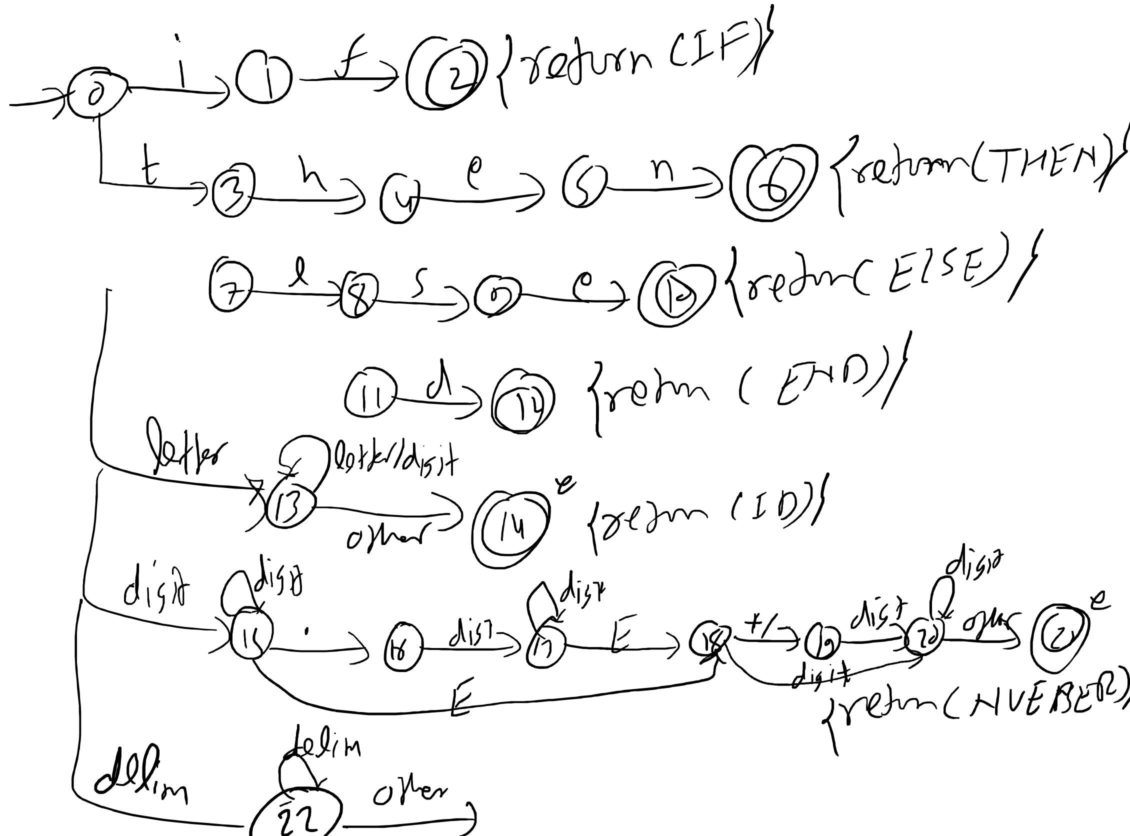
P_n {actionn}

if, then, else, end, id, number, relop

I/P: Lex specifications
- Manifest constants
- Auxiliary defns (Regular definitions)
- Translation Rules

else, end

O/P: Lexical Analyzer



Grammar (or Unstructured or Phrase Structured) (or type 0)

A grammar G is defined as 4-tuple (V_N, Σ, P, S)

$S \in V_N$:- Start symbol (start variable)

V_N – Non empty finite set of nonterminals (variables)

Σ – Nonempty finite set of terminals (terminal symbols)

P – Set of Productions (or production rules or rewriting rules)

$S \in V_N$: start symbol

→ P consists the productions of the form

$\alpha \to \beta$ $\quad \alpha, \beta \in (V_N \cup \Sigma)^*$ $\quad \alpha$ must contain at least one symbol from V_N

$\phi A \psi \to \phi \alpha \psi$

$\phi, \psi \in (V_N \cup \Sigma)^*$, $A \in V_N$

$$\boxed{V_N \cap \Sigma = \phi}$$

Types of Grammars

- Type 0 (UG or PSG)
- Type 1 (CSG – Context Sensitive Grammar)
- Type 2 (CFG – Context-free Grammar)
- Type 3 (RG – Regular grammar or Linear grammar)

Type 1 (CSG):

$G = (V_N, \Sigma, P, S)$

Context sensitive production

$\alpha \to \beta$

Left context → $\underbrace{\phi A \psi}_{L.H.S} \to \underbrace{\phi \alpha \psi}_{R.H.S}$ ← Right context

$\phi, \psi \in (V_N \cup \Sigma)^*$, $A \in V_N$

$\alpha \in (V_N \cup \Sigma)^*$

1) Contexts preservation condition

Lines

1) Contexts preservation condition

2) Length condition

$|\phi A \psi| \le |\phi \alpha \psi|$

if

<u>exception case</u> : if $\alpha = \varepsilon$

$|\phi A \psi| > |\phi \alpha \psi|$

We allow this case provided A should not appear in RHS of any other production.

A grammar G is Type 1 or CSG if every production of this grammar is Type 1.

<u>Context-free grammar</u> (Type-2)

$G = (V_N, \Sigma, P, S)$

P consists productions of the form $A \to \alpha$

$A \in V_N$, $\alpha \in (V_N \cup \Sigma)^*$

$S \in V_N$: Start symbol

$V_N \cap \Sigma = \phi$

<u>Regular Grammar</u> (Type3 grammar) <u>exception case</u> : $\alpha = \varepsilon$

$G = (V_N, \Sigma, P, S)$

production should be of the form $A \to \Sigma^* B \mid \Sigma^*$ (Right

or

$A \to B \Sigma^* \mid \Sigma^*$ (Left Linear)

Linear)

$A, B \in V_N$

$y = mx$ (Linear)

$y = 4x^2$ (quadratic)

$y^2 = 4ax$ (,,)

$G : (V, \Sigma, P, S)$

$S \to aCa$, $C \to aCa$
$C \to b$

$G : (V, \Sigma, P, S)$

$P: \{ S \to aCa,\ C \to aCa \mid b \} \Rightarrow C \to aCa,\ C \to b$ $\quad y^2 = \ldots$

$V = \{S, C\}, \quad \Sigma = \{a, b\}$

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

$L(G) = ?$

$S \Rightarrow aCa$ Left Right $\quad [\because S \to aCa]$

$\Rightarrow aba \quad w \quad [\because C \to b]$

$aba \in L(G)$

$S \Rightarrow aCa$

$\Rightarrow aaCaa$

$\Rightarrow aabaa = w$

$aabaa \in L(G)$

$L(G) = \{ aba, aabaa, aaabaaa, \ldots$

$= \{ a^n b a^n \mid n \geq 1 \}$

$S \Rightarrow aCa$

$\Rightarrow \Rightarrow$

$L(G) = \{ aba, aabaa, aaabaaa, aaaabaaaa, aaaaabaaaaa, \ldots \}$

$= \{ a^n b a^n \mid n \geq 1 \}$

Ex

Ex

$(\{S\}, \{0,1\}, \{S \to 0S1 \mid 01\}, S)$
$\quad V \quad \Sigma$

$01 \in L(G)$

$S \Rightarrow 01$

$0011 \in L(G)$

$S \Rightarrow 0S1 \Rightarrow 0011$

$000111 \in L(G)$

$S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 000111$

$L(G) = \{01, 0011, 000111, 00001111, \ldots\}$
$= \{0^n 1^n \mid n \geq 1\}$

$S \to 0S1 \mid \varepsilon$

$L(G) = \{\varepsilon, 01, 0011, 000111, \ldots\}$
$= \{0^n 1^n \mid n \geq 0\}$

L = set of all palindrome strings over $\{a, b\}$

$\{\varepsilon, a, b, aa, bb, aaa, aba, bab, bbb, \ldots\}$

$\to S \to aSa \mid bSb \mid aa \mid bb \mid \varepsilon$

$S \Rightarrow aa$

aba

bab

$\boxed{S \to aSa \mid bSb \mid a \mid b \mid \varepsilon}$

to generate all length palindromes

abba

baab

$S \Rightarrow aSa$

odd length P: $\{ S \to aSa \mid bSb \mid a \mid b$

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

aabaa

LMD
Left most Derivation

$\to S \Rightarrow bSb$ $[S \to bSb]$

$\Rightarrow baSab$ $[\because S \to aSa]$

$\Rightarrow babab = w$ $[\because S \to b]$

$babab \in L(G)$

$L(G) = \{a, b, aaa, aba, bab, bbb, aaaaa, ababa, \ldots\}$
= set of all ~~odd length~~ palindromes over $\{a, b\}$

Even length palindrome: $G = (\{S\}, \{a, b\}, P, S)$

P: $\{ S \to aSa \mid bSb \mid \varepsilon \}$

$S \Rightarrow abba$

Even length palindrome

$P: \{ S \to aSa \mid bSb \mid \varepsilon \}$

$L(G) = \{ \varepsilon, aa, bb, abba, baab, bbbb, \ldots \}$

aSa [$\because S \to aSa$]

$\boxed{S \overset{*}{\underset{G}{\Rightarrow}} abba}$

$w = abba$

$abba \in L(G)$

$S \Rightarrow \varepsilon$ (Direct or one-step derivation)

$S \overset{*}{\underset{G}{\Rightarrow}} aa \leftarrow w$

derivation

$S \Rightarrow aSa$ [$\because S \to aSa$]

$\Rightarrow abSba$ [$\because S \to bSb$]

$\Rightarrow ab\varepsilon ba = abba$ [$\because S \to \varepsilon$]

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

Example Consider the grammar given below as

$S \to S+S$

$S \to S*S$

$S \to b$

$P: \{ S \to \overset{\alpha_1}{S+S} \mid \overset{\alpha_2}{S*S} \mid \overset{\alpha_3}{a} \mid \overset{\alpha_4}{b} \}$

$A \in V_H$

$G = (\{S\}, \underbrace{\{a, b, +, *\}}_{\Sigma}, P, S)$

V_H

$S \Rightarrow S+S$ ~~[$\because S \to S+S \in P$]~~

~~$\Rightarrow a+S*S$~~ [$\because S \to S*S \in P$]

LMD

$\Rightarrow a+a*S$ [$\because S \to a \in P$]

$\Rightarrow a+a*b = w$ (yield) [$\because S \to b \in P$]

b-) Derivation tree (Parse

Left forms

$S \Rightarrow S+S$ [$\because S \to S+S$]

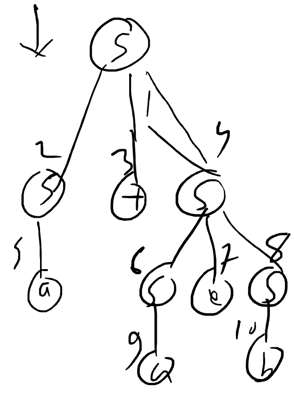
$\Rightarrow S+S*S$ [$\because S \to S*S$]

$S+S*b$ [$\because S \to b$]

$\Rightarrow S+a*b$ [$\because S \to a$]

Right forms

$\Rightarrow a+a*b$ [$\because S \to a$]



$\underbrace{a+a*b}$

$\underbrace{S \to S+S}$

(Parse tree)
Derivation tree: It is a tree with following characteristic:

root node will be labled with start variable. ... terminals).

2. Internal nodes will be labled with Variables (non ter...)

3. External " (leaf nodes) are labled with terminals or ε.

4. If a internal node X has children $\underbrace{X_1, X_2 \dots X_k}$ then

$$X \rightarrow X_1 X_2 \dots X_k \in P \qquad \underbrace{S \rightarrow S+S} \in P$$

$$S \rightarrow S*S \in P$$

5. A leaf node with label ε will not have any sibling.

G: $\boxed{S \rightarrow aSa \mid bSb \mid \varepsilon}$ ← even length palindromes

$S \Rightarrow aSa$ [$\because S \rightarrow aSa$]

$\Rightarrow abSba$ [$\because S \rightarrow bSb$]

$\Rightarrow ab\varepsilon ba = abba$ [$\because S \rightarrow \varepsilon$]

$\boxed{abba \in L(G)}$

a b b a

$L(G)$?

$\boxed{L(G) = \{ \varepsilon, aa, bb, aaaa, abba, baab, bbbb, \dots \}}$

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

$S \Rightarrow \varepsilon \qquad [S \rightarrow \varepsilon]$

$\underbrace{S \Rightarrow aSa}_{\gamma_0 \quad \gamma_1} \Rightarrow a\varepsilon a = aa = w$

$\gamma_2 = \gamma_n = w$

~~$\gamma_0 \quad \gamma_1$~~ $\quad S \overset{*}{\Rightarrow} aa$

LMD or RMD

$\underbrace{\overbrace{S \Rightarrow \gamma_1 \Rightarrow \gamma_2 \Rightarrow \gamma_3 \Rightarrow \dots \Rightarrow \gamma_n}^{*} = w}$ (yield)

(γ_0)

Sentential form

Left Right

Ambiguity in CFG:

Ambiguity in CFG:

$S \to S+S \mid S*S \mid a \mid b$

$w = a+a*b$

LMD 1

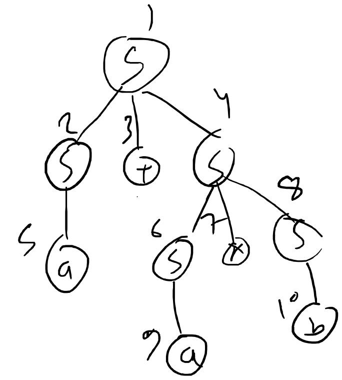
$S \Rightarrow S+S$ $[\because S \to S+S \in P]$

$\Rightarrow a+S$ $[S \to a \in P]$

$\Rightarrow a+S*S$ $[S \to S*S \in P]$

$\Rightarrow a+a*S$

$\Rightarrow a+a*b$ $[S \to b \in P]$



a+a*b

Bhopal

$\to S \Rightarrow S*S$

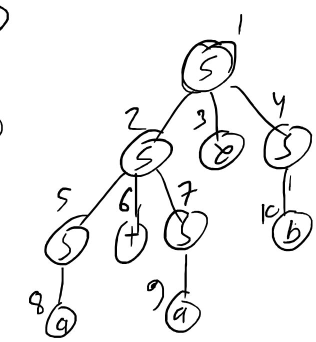
$\Rightarrow S+S*S$

$\Rightarrow a+S*S$

$\Rightarrow a+a*S$

$\Rightarrow a+a*b$

a+a*b



a+a*b

PT. $S \to SbS \mid a$ is ambiguous.

Derivation tree 1

LMD 1

$S \Rightarrow SbS$ $[\because S \to SbS \in P]$

$\Rightarrow SbSbS$ $[\because S \to SbS \in P]$

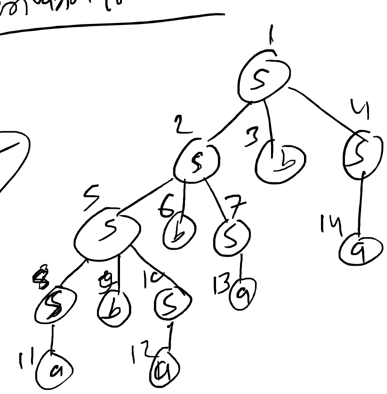
$\Rightarrow SbSbSbS$ ["]

$\Rightarrow abSbSbS$ $[S \to a \in P]$

$\Rightarrow ababSbS$ $[S \to a \in P]$

$\Rightarrow abababS$ $[S \to a \in P]$

$\Rightarrow abababa$ $w =$ abababa $|w| = 7$

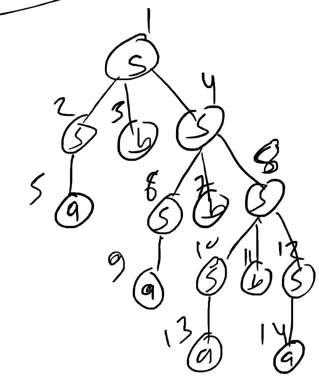


abababa

Derivation tree 2

$\Rightarrow ababa\ldots$

Derivation tree 2



LMD2

$S \Rightarrow SbS$ [$\because S \to SbS \in P$]

$\Rightarrow abS$ [$\because S \to a \in P$]

$\Rightarrow abSbS$ [$\because S \to SbS \in P$]

$\Rightarrow ababS$ [$\because S \to a \in P$]

$\Rightarrow ababSbS$ [$\because S \to SbS \in P$]

$\Rightarrow abababS$ [$\because S \to a \in P$]

$\Rightarrow abababa$ [$\because S \to a \in P$]

$\underbrace{abababa}$

$S \Rightarrow SbS \Rightarrow abS \Rightarrow aba$

1) $S \to 0S1 \mid 01$

2) $S \to 0S11 \mid 011$

3) $S \to 0S1 \mid \varepsilon$

4) $S \to 0S11 \mid \varepsilon$

5) $S \to aS \mid bS \mid a \mid b$. $L(G) = ?$

$S \Rightarrow a$ $\quad a \in L(G)$

$S \Rightarrow b$ $\quad b \in L(G)$

$aa \in L(G)$ $\quad ab \in L(G)$

$ba \in L(G)$

$bb \in L(G)$

$S \Rightarrow aS \Rightarrow aa$ ✓

$S \Rightarrow aS \Rightarrow ab$

$S \Rightarrow bS \Rightarrow ba$

$S \Rightarrow bS \Rightarrow bb$

$S \Rightarrow aS \Rightarrow aaS \Rightarrow aaa$

$L(G) = \{a, b, aa, ab, ba, bb, aaa, aab, aba, abb, baa, bab, bba, bbb, \ldots\}$

$= \{a, b\}^+$

$= \Sigma^+$

$L = \{0^n 1^n \mid n \geq 1\} = \{01, 0011, 000111, 00001111, \ldots\}$

$L = \{0^n 1^n \mid n \geq 1\} = \{01, 0011, 000111, 00001111, \ldots\}$

$\Downarrow$

$G = ?$

$S \to 01$

$S \to 0S1$

$G = (\{S\}, \{0,1\}, \{S \to 0S1, S \to 01\}, S)$

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

$= \{abb, aabbbb, aaabbbbbb, \ldots\}$

$S \to abb$

$S \to aSbb$

$S \Rightarrow 0S1 \Rightarrow 0011$

$S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow \underline{000111}$ ✓

$L = \{0^n 1^n \mid n \geq 0\}$

$G = (\{S\}, \{0,1\}, \{S \to 0S1 \mid \varepsilon\}, S)$

$S \Rightarrow 0S1 \Rightarrow 0\varepsilon 1 = 01$

$S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 00\varepsilon 11 = 0011$

Simplification of CFGs

③) Elimination of Useless symbols and production

①) Elimination of null productions $(A \to \varepsilon)$

②) Elimination of Unit productions $(A \to B)$

Elimination of null production:

$S \to AB \mid aS$

$A \to \varepsilon \mid b$

$B \to \varepsilon \mid c$

$D \to b$

$S \Rightarrow aS \Rightarrow a\varepsilon = a$

$\varepsilon \in L(G)$ if $S \overset{*}{\underset{G}{\Rightarrow}} \varepsilon$

$G' = G$

$L(G') = L(G)$

Step I. Construction of set of nullable variables

Step I · Construction of set of nullable variables

$L(G^{''}) = L(G)$

$+ \ w \in \Sigma^*$

$W_1 = \{A, B\}$ as $A \to \varepsilon$, $B \to \varepsilon$

$W_2 = \{A, B\} \cup \{S\}$ as $S \to AB$

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

$W_3 = \{S, A, B\} \cup \phi = \{S, A, B\} = W_2$

G' such that

1) $\boxed{L(G') = L(G) - \{\varepsilon\}}$

2) $\boxed{L(G') = L(G)}$

$\boxed{W = \{S, A, B\} = W_2}$

Case I

Step II : Elimination of null productions

$S \to aAB$

$S \to a\varepsilon\varepsilon = a$

i) $D \to b$ will included in P' ✓

ii) $S \to AB$ gives rise to $S \to AB$, $S \to B$, $S \to A$

iii) $S \to aS$ gives rise to $S \to aS$, $S \to a$

iv) $A \to b$ included in P' ✓

v) $D \to c$,, ,, ✓

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

$G' = (\{S, A, B, D\}, \{a, b, c\}, P', S)$

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

$B \to c$

$D \to b$

Case II $\} \ L(G') = L(G)$

$$P' = \left.\begin{array}{l} S_1 \to S \mid \varepsilon \\ S \to AB \mid A \mid B \mid aS \mid a \\ A \to b \\ B \to c \\ D \to b \end{array}\right\}$$

$G' = (\{S_1, S, A, B, D\}, \{a, b, c\}, P', S_1)$

Elimination of Unit Productions :

A production of the form $A \to B$ is called unit production.

$S \to AB,\ A \to a,\ B \to C|b,\ C \to D|c,\ D \to E|d,\ E \to a$

Step I ~~Construction~~ of set of reachable variables in unit form

$W_0(S) = \{S\}$

$W_1(S) = \{S\} \cup \phi$

$\cancel{W_1(B) = \{B\} \cup \{C\} = \{B, C\}}$

$W_2(B) = \{B, C\} \cup \{D\} = \{B, C, D\}$

$W_3(B) = \{B, C, D\} \cup \{E\} = \{B, C, D, E\}$

$W_4(B) = \{B, C, D, E\} \cup \phi$

$\boxed{W(S) = \{S\}}$

$\cancel{W_0(A) = \{A\}}$

$W_1(A) = \{A\} \cup \phi$

$W(A) = \{A\}$

$\boxed{W(A) = \{A\}}$

$\boxed{W(B) = \{B, C, D, E\}}$

$W_0(C) = \{C\}$

$W_1(C) = \{C\} \cup \{D\}$

$= \{C, D\}$

$W_2(C) = \{C, D\} \cup \{E\} = \{C, D, E\}$

$W_3(C) = \{C, D, E\} \cup \phi = \{C, D, E\}$

$\boxed{W(C) = \{C, D, E\}}$

$W_0(D) = \{D\}$

$W_1(D) = \{D\} \cup \{E\} = \{D, E\}$

$W_2(D) = \{D, E\} \cup \phi = \{D, E\}$

$\boxed{W(D) = \{D, E\}}$

$W_0(E) = \{E\}$

$W_1(E) = \{E\} \cup \phi = \{E\}$

$\boxed{W(E) = \{E\}}$

$C \to c,\ D \to d$ ✓

Step II

i) $S \to AB,\ A \to a,\ B \to b,\ E \to a$ are included in P'

ii) $B \to C$ gives rise to $B \to a|d|c$

iii) $C \to D$ gives rise to $C \to a|d$

iv) $D \to E$,, ,, $D \to a$ ✓

$G' = (\{S, A, B, C, D, E\}, \{a, b, c, d\}, P', S)$
$\quad\quad\quad V_N \quad\quad\quad \Sigma$

$P' = \{$
$S \to AB$
$A \to a$
$B \to a|b|c|d$
$C \to a|d|c$
$D \to a|d$
$E \to a$
$\}$

Elimination of Useless Symbols and Productions

Construct a reduced grammar equivalent to the grammar
S → aAa, A → Sb | bCC | DaA. C → abb | DD, E → aC, D → aDA

Step I : Construction of set of terminating variables

$W_1 = \{C\}$ as $C \to abb$

$W_2 = \{C\} \cup \{A, E\}$ as $A \to bCC$, $E \to aC$
$= \{A, C, E\}$

$W_3 = \{A, C, E\} \cup \{S\}$ as $S \to aAa$
$= \{S, A, C, E\}$

$W_4 = \{S, A, C, E\} \cup \{A\}$ as $A \to Sb$
$= \{S, A, C, E\}$

$\boxed{W = \{S, A, C, E\} = W_3}$

$G_1 = (\{S, A, C, E\}, \{a, b\}, P_1, S)$

$P_1 := \{ S \to aAa$
$A \to Sb \mid bCC$
$C \to abb,$

$\} \; E \to aC$

Step II: Find symbols reachable from start symbol

$W_0 = \{S\}$

$W_1 = \{S\} \cup \{A, a\}$ as $S \to aAa$

$= \{S, A, a\}$

$W_2 = \{S, A, a\} \cup \{S, C, b\}$ as $A \to Sb \mid bCC$

$= \{S, A, C, a, b\}$

$W_3 = \{S, A, C, a, b\} \cup \{a, b\}$ as $C \to abb$

$= \{S, A, C, a, b\}$

$G_2 = (\underbrace{\{S, A, C\}}_{V_H}, \underbrace{\{a, b\}}_{\Sigma}, P_2, S)$

$P_2 := \{ S \to aAa$

$A \to Sb \mid bCC$

$C \to abb \}$

Build the simplified Context-free Grammars equivalent to the given grammars.

(i) S→AB
A→a | B
B→b | C
C→aC
D→b

(ii) S→0A0 | 1B1 | BB
A→C
B→A
C→ε

Step I Elimination of null production

Step II ,, ,, Unit ,,

$S \to AB$

$A \to a \mid \cancel{B} \mid aC$

$B \to b \mid aC$

$C \to aC$

$D \to b$

Step III Elimination of useless symbols & productions

(i) $S \to AB$

S

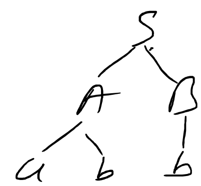
Step III Example

(i) $S \to AB$

$A \to a/b$

$B \to b$

$D \to b \Leftarrow$ useless



(ii) $\left.\begin{array}{l} S \to AB \\ A \to a/b \\ B \to b \end{array}\right\}$

$$\boxed{\begin{array}{l} G' = (\{S, A, B\}, \{a, b\}, P', S) \\ P' = \{S \to AB, A \to a/b, B \to b\} \end{array}}$$

Normal Forms (Module-3)

- CNF (Chomsky Normal Form)
- GNF (Greibach Normal Form)

CNF: Every production is of the form

$A \to BC \mid a$

Convert the following grammar to Chomsky Normal Form

$S \to bA \mid aB$, $A \to bAA \mid aS \mid a$, $B \to aBB \mid bS \mid b$

Step I: Eliminate null and unit productions

$G_1 = G$

Step II: (i) $S \to bA$ gives rise to $S \to C_b A$, $C_b \to b$

(ii) $S \to aB$ " " " $S \to C_a B$, $C_a \to a$

(iii) $A \to bAA$ " " " $A \to C_b AA$

(iv) $A \to aS$ " " " $A \to C_a S$

(v) $A \to a$ ✓ is in CNF ✓

(vi) $B \to aBB$ " " : $B \to C_aBB$ ✗

(vii) $B \to bS$ " " . $B \to C_bS$ ✓

(viii) $B \to b$ is in CNF

Step III (i) $A \to \underbrace{C_bA}_{C_1}A$ gives rise to $A \to C_1A$ ✓, $C_1 \to C_bA$ ✓

(ii) $B \to \underbrace{C_aB}_{C_2}B$ " " $B \to C_2B$ ✓, $C_2 \to C_aB$ ✓

$G' = (\{S, A, B, C_a, C_b, C_1, C_2\}, \{a, b\}, P', S)$

$P' = \{$

$S \to C_bA \mid C_aB$

$A \to C_aS \mid C_1A \mid a$

$B \to C_bS \mid b \mid C_2B$

$C_1 \to C_bA$

$C_2 \to C_aB$

$C_a \to a$

$C_b \to b$

$\}$

Chomsky Normal Form, where P is given as

$E \to E + T \mid T, \quad T \to T \star F \mid F, \quad F \to (E) \mid id$

Step I : Eliminate null and unit production

$E \to E+T \mid T\star F \mid (E) \mid id$

$T \to T\star F \mid (E) \mid id$

$F \to (E) \mid id$

Step II (i) $E \to E+T$ gives rise to $E \to EAT$ ✗, $A \to +$ ✓

(ii) $E \to T\star F$ " " $E \to TBF$ ✗, $B \to \star$ ✓

(iii) $E \to (E)$ " " $E \to CED$ ✗, $C \to ($ ✓, $D \to)$ ✓

(iv) $E \to id$ is in CNF ✓

(v) $T \to T \circ F$ " " $T \to TBF$

(vi) $T \to (E)$ " " $T \to CED$

(vii) $T \to id$ is in CNF

(viii) $F \to (E)$ " " $F \to CED$

(ix) $F \to id$ is in CNF

Step III (i) $E \to EAT$ gives rise to $E \to GT$, $G \to EA$

(ii) $E \to TBF$ " " $E \to HF$, $H \to TB$

(iii) $E \to CED$ " " $E \to ID$, $I \to CE$

$$A \to \underbrace{BC}_{C_1}\underbrace{DE}_{C_2}$$

$$A \to C_1C_2 \quad , \; C_1 \to BC, \quad C_2 \to DE$$

GNF (Greibach Normal Form)

$$A \to a\alpha^{\epsilon} \quad , \quad \alpha \in V_N$$

or

$$A \to a\alpha \quad , \; \alpha \in V_N^{\epsilon}$$

Find the GNF equivalent to the following:

S→AA | a, A→SS | b

Step (A): Eliminate null, unit and useless symbols

$G_1 = G$

Step I : Rename the variables

$S \to A_1$, $A \to A_2$

$A_1 \to A_2A_2 \mid a$

$A_2 \to A_1A_1 \mid b$

$A_i \to A_j \gamma$

Step II : $A_1 \to a$, $A_2 \to b$ are in GNF

$\boxed{}$ → next step

$>$ i (Convert into $i = j$ case)

$i < j$ (Step IV)

$A_2 \to A_2A_2A_1 \mid aA_1 \mid b$

$A_2 \to A_2A_2A_1 \mid aA_1 \mid b$ $\quad i < j$ (step IV)

$A_2 \to aA_1 \mid b$

Step III: $A_2 \to A_2A_2A_1$

$A_n \to A_n\gamma$ $\quad A_n = A_2$, $\gamma = A_2A_1$

$A_n \to A_nZ_n$

$Z_n \to \gamma \mid \gamma Z_n$

$G' = (\{A_1, A_2, Z_2\}, \{a, b\}, P', A_1)$

The productions in Eq (1) to (3) are in P' that are in CNF.

$A_2 \to A_2Z_2 \Rightarrow A_2 \to aA_1Z_2 \mid bZ_2$

$Z_2 \to A_2A_1 \mid A_2A_1Z_2$ (take into step V)

Step V $Z_2 \to A_2A_1 \mid A_2A_1Z_2$ gives rise to

$Z_2 \to aA_1Z_2A_1 \mid bZ_2A_1 \mid aA_1A_1 \mid bA_1 \mid$

$aA_1Z_2A_1Z_2 \mid bZ_2A_1Z_2 \mid$

$aA_1A_1Z_2 \mid bA_1Z_2$ --- (3)

Step IV $A_2 \to aA_1Z_2 \mid bZ_2 \mid aA_1 \mid b$ --- (1)

$A_1 \to aA_1Z_2A_2 \mid bZ_2A_2 \mid aA_1A_2 \mid bA_2 \mid a$ --- (2)

Convert the following Context Free Grammar to Chomsky Normal Form

S→AaB | aaB, A→ ε, B→bbA | ε

Step I: Eliminate null and unit productions

$S \to AaB \mid aB \mid Aa \mid a \mid aaB \mid aa$

$B \to bbA \mid bb$

$S \to aB \mid aaB \mid aa \mid a$

$B \to bb$

Step II: (i) $S \to aB$ gives rise to $S \to C_aB$, $C_a \to a$

(ii) $S \to aaB$ gives " " $S \to C_aC_aB$

(iii) $S \to aa$ " " , $S \to C_aC_a$

(iv) $S \to a$ is in CNF

(v) $B \to bb$ gives rise to $B \to C_bC_b$, $C_b \to b$

Step III (i) $S \to C_aC_aB$ gives rise to $S \to C_aC_1$, $C_1 \to C_aB$

The resultant grammar in CNF can be written as

$G' = (\{S, B, C_1, C_a, C_b\}, \{a, b\}, P', S)$

$P' = \{ S \to C_aB \mid C_aC_1 \mid C_aC_a \mid a,\ B \to C_bC_b,\ C_1 \to C_aB,\ C_a \to a,\ C_b \to b \}$

Greibach Normal Form (GNF)

$G = (\{S, A, B\}, \{0,1\}, P, S)$ $\quad A \to \alpha$ (CFG)

$P\ \{$ S --> A0, A --> 0B. B --> A0, B -->1 $\}$ $\quad \alpha \in V_N$

$A \to a\alpha,\ \alpha \in V_N^*$

Step (a) Simplification of G

Step (b) $S \to A0$ gives rise to $S \to AC,\ C \to 0$

$B \to A0$ " " " $B \to AC$

$G_1 = (\{S, A, B, C\}, \{0,1\}, P_1, S)$

$P_1 = \{ S \to AC,\ A \to 0B,\ B \to AC \mid 1,\ C \to 0 \}$

Step 1 Rename the variables

$S \to A_1,\quad A \to A_2,\ B \to A_3,\ C \to A_4$

$A_1 \to A_2A_4$

$A_2 \to 0A_3$ ✓

$A_3 \to A_2A_4 \mid 1$ ✓

$A_4 \to 0$ ✓

$A_i \to A_j\gamma$

Step II (i) $A_2 \to 0A_3,\ A_3 \to 1,\ A_4 \to 0$ are in GNF

(ii) $A_3 \to A_2A_4$ gives rise to $A_3 \to 0A_3A_4$ ✓

$(i > j)$

(iii) $A_1 \to A_2A_4$ " " " $A_1 \to 0A_3A_4$ ✓

$G' = (\{A_1, A_2, A_3, A_4\}, \{0,1\}, P', A_1)$

$P' = \{ A_1 \to 0A_3A_4,\ A_2 \to 0A_3,\ A_3 \to 0A_3A_4 \mid 1,\ A_4 \to 0 \}$

Ex: $E \to E+T \mid T$

$T \to T*F \mid F$

$F \rightarrow (E) \mid a$

Step (a): Elimination of null & unit production

$$P \left\{ \begin{array}{l} E \rightarrow E+T \mid T*F \mid (E) \mid a \\ T \rightarrow T*F \mid (E) \mid a \\ F \rightarrow (E) \mid a \end{array} \right.$$

$G = (\{E, T, F\}, \{+, *, (,), a\}, P, E)$

Step (b):

$E \rightarrow E+T$ gives rise to $E \rightarrow EAT$, $A \rightarrow +$

$E \rightarrow T*F$ " " " $E \rightarrow TBF$, $B \rightarrow *$

$E \rightarrow (E)$ " " " $E \rightarrow (EC$, $C \rightarrow)$

$T \rightarrow T*F$ " " " $T \rightarrow TBF$

$T \rightarrow (E)$ " " " $T \rightarrow (EC$

$F \rightarrow (E)$ " " " $F \rightarrow (EC$

$$\begin{array}{l} E \rightarrow EAT \mid TBF \mid (EC \mid a \\ T \rightarrow TBF \mid (EC \mid a \\ F \rightarrow (EC \mid a \\ A \rightarrow + \\ B \rightarrow * \\ C \rightarrow) \end{array}$$

Step I: Rename the variables

$E \rightarrow A_1$, $A \rightarrow A_2$, $T \rightarrow A_3$, $B \rightarrow A_4$, $F \rightarrow A_5$, $C \rightarrow A_6$

$$\begin{array}{l} A_1 \rightarrow A_1 A_2 A_3 \mid A_3 A_4 A_5 \mid (A_1 A_6 \mid a \\ A_3 \rightarrow A_3 A_4 A_5 \mid (A_1 A_6 \mid a \\ A_5 \rightarrow (A_1 A_6 \mid a \\ A_2 \rightarrow + \\ A_4 \rightarrow * \\ A_6 \rightarrow) \end{array}$$

Step II $A_1 \rightarrow (A_1 A_6 \mid a$

Step II

$$\left.\begin{array}{l} A_1 \to (A_1A_6 \mid a \\ A_3 \to (A_1A_6 \mid a \\ A_5 \to (A_1A_6 \mid a \\ A_2 \to + \\ A_4 \to e \\ A_6 \to\ ) \end{array}\right\}$$ are in GNF

$A_1 \to A_1 \overset{e}{A_2} A_3 \mid (A_1A_6 \mid a$

~~$A_3 \to A_3 \overset{e}{A_4} A_5 \mid (A_1A_6 \mid a$~~

Step III

$A_1 \to A_1 \underbrace{A_2A_3}_{r}$ $\qquad$ $A_1 \to (A_1A_6 \mid a$

$A_3 \to A_3 \underbrace{A_4A_3}_{r}$ $\qquad$ $A_3 \to (A_1A_6 \mid a$

$A_n \to A_n r \Rightarrow A_n \to A_n Z_n$

$Z_n \to r \mid r Z_n$

$A_1 \to A_1 Z_1 \longrightarrow A_1 \to (A_1A_6Z_1 \mid aZ_1$

$Z_1 \to A_2A_3 \mid A_2A_3Z_1$

$A_3 \to A_3Z_3 \Longrightarrow A_3 \to (A_1A_6Z_3 \mid aZ_3$

$Z_3 \to A_4A_5 \mid A_4A_5Z_3$

Step IV

$A_3 \to (A_1A_6Z_3 \mid aZ_3 \mid (A_1A_6 \mid a$ — (1)

$A_1 \Rightarrow A_3A_4A_5$ gives rise to $A_1 \to (A_1A_6Z_3A_4A_5 \mid aZ_3A_4A_5 \mid$
$(A_1A_6A_4A_5 \mid aA_4A_5$

$A_1 \to (A_1A_6Z_3A_4A_5 \mid aZ_3A_4A_5 \mid (A_1A_6A_4A_5 \mid aA_4A_5 \mid (A_1A_6Z_1 \mid aZ_1 \mid$
$(A_1A_6 \mid a$ — (2)

$A_2 \to +$ — (3)

$A_4 \to e$ — (4)

$A_5 \to (A_1A_6 \mid a$ — (5)

$A_6 \to\ )$ — (6)

Step V

$Z_1 \to A_2A_3 \mid A_2A_3Z_1$ gives rise to

$\Downarrow$

$Z_1 \to +A_3 \mid +A_3Z_1$ — (7)

$Z_3 \to A_4A_5 \mid A_4A_5Z_3$ gives rise to

$\Downarrow$

$Z_3 \to \varepsilon A_3 \mid \varepsilon A_3Z_3$ — — (8)

$G' = (\{A_1, A_2, A_3, A_4, A_5, A_6, Z_1, Z_3\}, \{+, \varepsilon, (,), a\}, P', A_1)$

The productions from ~~(1)~~ (1) to (8) are in P' that are in GNF.