Syntax Analysis :-

* Introduction.
* Role of Syntax Analysis.

Introduction :-

→ It is the Second Phase of the compilation.

- It checks for the Syntax of the language.
- Syntax analyzer takes the tokens from the lexical analyzer and groups them in some programming structure called "Syntax tree (or) Parse tree".
- If any syntax cannot be recognised then the syntax error will be generated.

Definition :-

- A Parsing (or) Syntax analysis is a process which takes input string "w" and produce either a parse tree (or) generate the syntactic errors.
- It is also called "Parser".

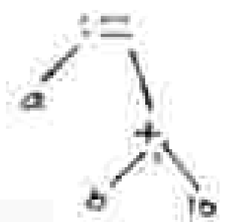
Ex:- consider the source program statement $A := B + 10$

then the lexical analyzer reads the above statement and breaks it into the set of tokens like ~~'A' is identifier~~

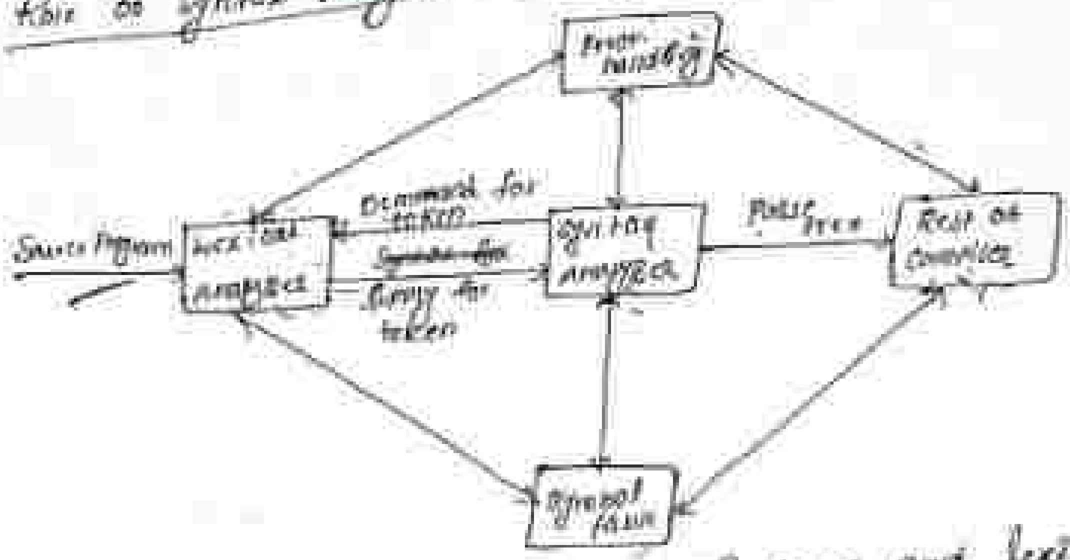
A — is identifier
:= — Assignment operator
B — identifier
+ — Assignment operator
10 — number (or) constant.

- Now the syntax analyzer collect the above tokens from lexical analyzer and arrange them into a structure is called "Parse tree (or) Syntax tree".

⇒ a := b + 10.



Role of Syntax Analyzer :- (or) Parser :-



In the Process of compilation the Parser and lexical analyzer work together that means.

1. When the Parser requires String of tokens it invokes lexical analyzer.

2. In turn the lexical Analyzer Supply tokens to Syntax Analyzer.

→ The Parser Collects Sufficient number of tokens and build a parse tree.

→ It finds Syntactical Errors at the time of construction Parse tree.

→ This Errors are ~~recorded~~ recovery by Error handler.

Context-free Grammar :-

* Introduction * Derivation and Parse tree * Ambiguous Grammar

→ Introduction

A context free Grammar "G" is a four tuples like

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

V ⟶ set of non terminal Symbols.

T ⟶ set of Terminal Symbols.

P ⟶ set of Production ~~form~~ rules of the form α, β where α ∈ V, β ...

S → start symbol.

Ex:- Let the language $L = a^n b^n, n \geq 1$

$L = a^n b^n$

minimum string = ab

$S \rightarrow a^n b^n$

$\rightarrow a\, a^{n-1} b^{n-1} b$

$\rightarrow aa\, a^{n-2} b^{n-2} bb$

P: $S \rightarrow aSb$

$S \rightarrow aSb$

$S \rightarrow ab$

From the above $G = (V, T, P, S)$ where is a Context free Grammar

where $V \rightarrow$ set of $\{S\}$

$T \rightarrow$ set of $\{a, b\}$

$P \rightarrow \{aSb, ab\}$

$S \rightarrow$ start symbol $\{S\}$

→ Derivation and Parse tree :-

Derivation from 'S' means generation of a string 'W' from 'S'. For constructing derivation two things are important.

1. Choice of non-terminal form several others.
2. choice of rule from production rules for corresponding non terminal.

→ Definition of Derivation tree :-

Let $G = (V, T, P, S)$ be a context free grammar. The Derivation tree is a tree which can be constructed by following properties.

1. The root node has label 'S'.
2. Every vertex can be derived from $\{V \cup T \cup \epsilon\}$

3. If there exist a vertex 'A' with children $x_1, x_2, x_3 \ldots x_n$ then there should be a production rule

$$A \rightarrow x_1 x_2 x_3 \ldots x_n$$

4. The leaf nodes are from set T and interior nodes are from set {V}.

→ 1. Left Most Derivation (LMD)

In left most derivation the left most non terminal is replaced by a terminal (or) non-terminal in each step begining with a start symbol.

→ Right Most Derivation (RMD):-

In Right most derivation the Right most non terminal in each step is replaced by a terminal (or) non terminal begining with a start symbol.

Ex:-> Consider the grammar given below

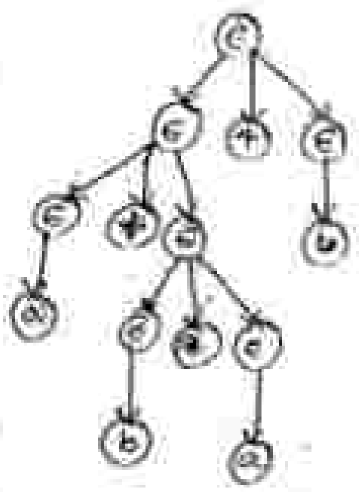
$$E \rightarrow E+E$$
$$E \rightarrow E-E$$
$$E \rightarrow E*E$$
$$E \rightarrow E/E$$
$$E \rightarrow a/b.$$

Obtain 1. Left Most Derivation 2. Right Most Derivation 3. Parse tree for the input string $a+b*a+b$.

(a) Left Most Derivation:-

$$E \rightarrow E+E$$
$$\rightarrow E+E+E$$
$$\rightarrow a+E+E$$
$$\rightarrow a+E*E+E$$
$$\rightarrow a+b*E+E$$
$$\rightarrow a+b*a+E$$
$$\rightarrow a+b*a+b$$

Parse tree:-



2) Right Most Derivation:- a+b*a+b

$E \rightarrow E+E$

$\rightarrow E+E+E$

$\rightarrow E+E+b$

$\rightarrow \cancel{E+a+b}$

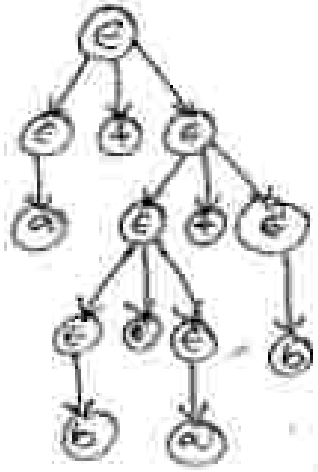
$\rightarrow E+E*E+b$

$\rightarrow E+E*a+b$

$\rightarrow E+b*a+b$

$\rightarrow a+b*a+b$

Parse tree:-



8. consider the grammar given below $S \rightarrow (L) \mid a$
$L \rightarrow S, S \mid S$.

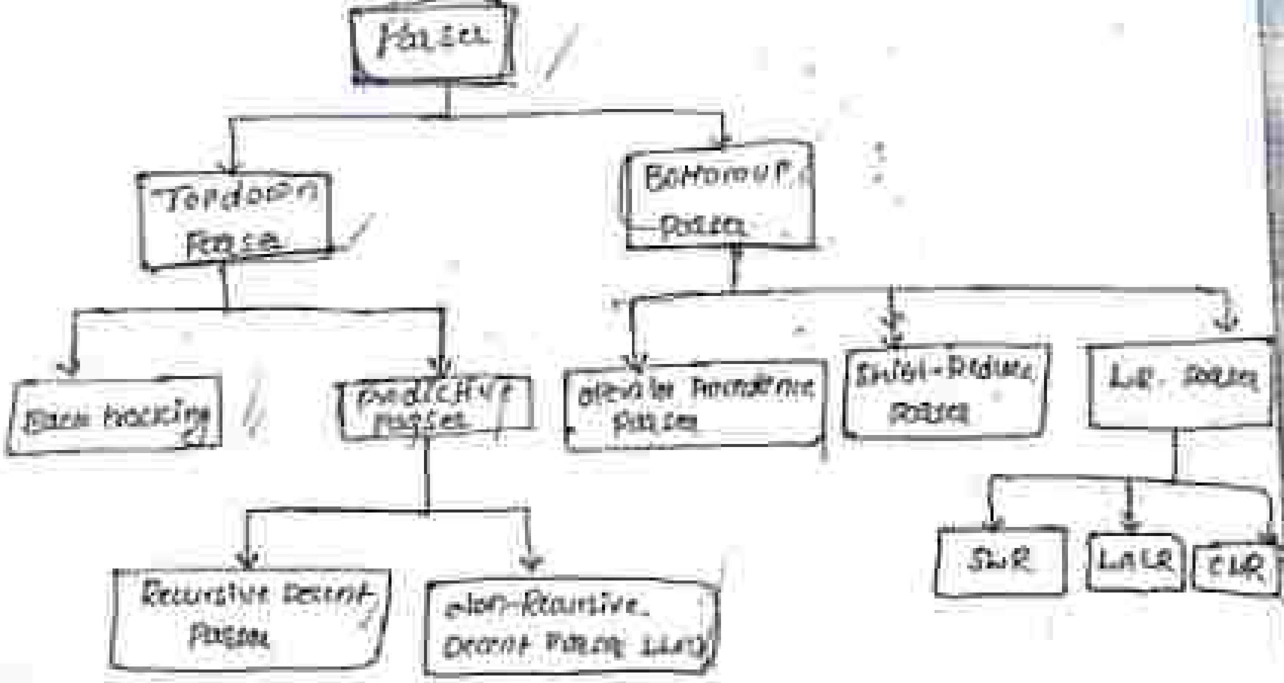
input string is (a, (a,a)).

→ Writing a Context-free Grammar:

* Lexical Analyser Vs Syntax Analyser
* Classification of Parsing techniques.
* Problems with Parsing Techniques.

1. Back tracking.
2. Left Recursion.
3. Left factoring.
4. Ambiguity.

→ Classification of Parsing techniques:-



Top down Parser:-

. The Process of constructing a syntax tree (or) Parse tree from root node to leaf node is called "Topdown Parser".

. Top down Parser classified into two types.

1. Back tracking 2. Predictive Parser.

→ Back tracking :-

Back tracking is an technique in which for ~~[illegible]~~ Expansion of non terminal symbol we choose one alternative and if some mismatch occurs then we try another alternative if any.

ex:- consider the grammar $S \rightarrow xPz$
$P \rightarrow yw|y$

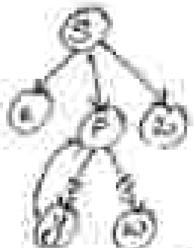
Now we obtain an input string xyz from the above grammar.

$\rightarrow$ 1

$\rightarrow$ we can take the first production $S \rightarrow xPz$ then the corresponding parse tree is $S \rightarrow xPz$

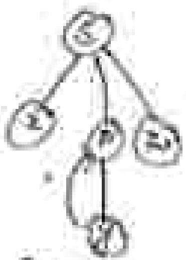
$\rightarrow$ 2. Now non-terminal symbol 'p' is replace by the first alternative of it that is $P \rightarrow yw$ then the parse tree is

$S \rightarrow xPz$
$P \rightarrow yw|y$



This string doesn't derived the given i/p string. So we move backward to 'P' and remove the corresponding branch from it and apply another alternative of it. then the corresponding tree is

$S \rightarrow xPz$
$P \rightarrow xyz$



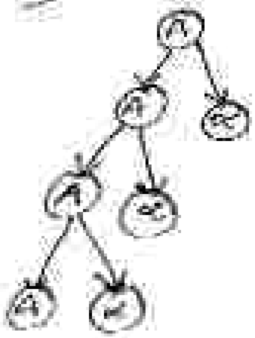
This parsing technique increases lot of over head in implementation of parse tree. So we need to eliminate the backtracking by modifying the grammar.

$\rightarrow$ **Left Recursion:-**

A left recursive grammar is a grammar which contains the production rule is like $A \rightarrow A\alpha$ where $A \in V$ and $\alpha \in (V \cup T)^*$.

· If left-Recursion is is present in the grammar the top down parser can enter into infinity loop like

$A \to A\alpha$



→ Elimination of left Recursion:-

To Eliminate left Recursion we need to modified the grammar.

Let 'G' = (V, T, P, S) be a CFG with the productions rule having left recursion $A \to A\alpha$, $A \to \beta$. then we Eliminate the left Recursion by rewriting the production rule as

1. $A \to \beta A'$
2. $A' \to \alpha A'$
3. $A' \to \epsilon$

$A \to A\alpha$
$A \to \beta$

Ex:- consider the grammar

$E \to E + T$ } $A \to A\alpha$ $A \to A\alpha$
$E \to T$ } $A \to \beta$ $A = E$, $\alpha = +T$,
$T \to T * F$ } $A \to A\alpha$
$T \to F$ } $A \to \beta$
$F \to (E)$
$F \to id$

Eliminate left recursion from the given grammar.

A: Step 1:- $E \to E + T$ $T \to T * F$ $F \to (E)$
$E \to T$ $T \to F$ $F \to id$.
we can map this grammar production rule with the rule $A \to A\alpha$, $A \to \beta$. where $A = E$
$\alpha = +T$
$\beta = T$

Know we can eliminate left Recursion

1. $A \to \beta A'$
   ↳ $E \to TE'$
2. $A' \to \alpha A'$
   ↳ $E' \to +TE'$
3. $A' \to \epsilon$
   ↳ $E' \to \epsilon$

∴ After eliminating left recursion the new Production rules are.

$E \rightarrow TE'$

$E' \rightarrow +TE'$

$E' \rightarrow \epsilon$

~~T → *FT'~~ similarly for the rules

$T \rightarrow T*F$

$T \rightarrow F$

we can map the grammar rule with the rule

$A \rightarrow A\alpha$

$A \rightarrow \beta$

where $A = T$

$\alpha = *F$

$\beta = F$

know we can eliminate left recursion.

1. $A \rightarrow \beta A'$

 $\hookrightarrow T \rightarrow FT'$

2. $A' \rightarrow \alpha A'$

 $\hookrightarrow T' \rightarrow *FT'$

3. $A' \rightarrow \epsilon$

 $\hookrightarrow T' \rightarrow \epsilon$

Therefore the grammar with out left recursion is

$E \rightarrow TE'$

$E' \rightarrow +TE'$

$E' \rightarrow \epsilon$

$T \rightarrow FT'$

$T' \rightarrow *FT'$

$T \rightarrow \epsilon$

$F \rightarrow (E)$

$F \rightarrow id$.

Q. Consider the Grammar
$A \rightarrow ABd$
$A \rightarrow Aa$
$A \rightarrow a$
$B \rightarrow Be$
$B \rightarrow b$
remove left recursion.

Sol:- $A \to ABd$ $\quad$ $A \to Aa$ $\quad$ $B \to Bc$

$A \to a$ $\quad$ $A \to a$ $\quad$ $B \to b$

1. $A \to aA'$ $\quad$ 1. $A \to aA'$ $\quad$ 1. $B \to bB'$
2. $A' \to BdA'$ $\quad$ 2. $A' \to aA'$ $\quad$ 2. $B' \to cB'$
3. $A' \to \epsilon$ $\quad$ 3. $A' \to \epsilon$ $\quad$ 3. $B' \to \epsilon$

→ Left factoring :-

1. A Grammar may not be suitable for recursive decent Parsing even if there is no left recursion.
2. for example consider the grammar $S \to iEtS \mid iEtSeS \mid a$, $E \to b$
3. A useful method for manipulating the grammar into a form suitable for recursive decent Parsing is left factoring.

→ Left - factoring :-

The Process of factoring out the common prefix of alternative. let $A \to \alpha\beta_1 \mid \alpha\beta_2 \mid \alpha\beta_3 \mid \ldots \mid \alpha\beta_n$ are 'n' number of 'A' Production rule and 'α' is not equal to null. after left factoring the grammar will become

1. $A \to \alpha A'$
2. $A' \to \beta_1 \mid \beta_2 \mid \beta_3 \mid \ldots \mid \beta_n$

Ex:- consider the Grammar $S \to iEtS \mid iEtSeS \mid a$, $E \to b$. do the left factoring on above Grammar.

Sol:- consider the Production rule with common Prefix Part. $S \to iEtS \mid iEtSeS$ to now Perform left-factoring then after the two Productions are

We can map this Grammar rule with the Rules

$A \to \alpha\beta_1 \mid \alpha\beta_2$

where $A = S$, $\alpha = iEtS$, $\beta_1 = \epsilon$, $\beta_2 = eS$. After left factoring the new Productions rule are

1. $A \to \alpha A'$
   $S \to iEtS\,S'$
2. $A' \to \beta_1 \mid \beta_2$
   $S' \to \epsilon \mid eS$
3. $S \to a$
   $E \to b$

∴ The grammar after left factoring is

$S \rightarrow iEtS S'$

$S' \rightarrow \epsilon \mid eS$

$S \rightarrow a$

$E \rightarrow b$

2, To the left factoring in the following Grammar.

$A \rightarrow aAB \mid aA \mid a$

$B \rightarrow bB \mid b$

Sol: $A \rightarrow aAB \mid aA \mid a$

where $A = A$

$\alpha = a$

$\beta_1 = AB$

$\beta_2 = A$

$\beta_3 = \epsilon$

After left factoring the new production rule is

$A \rightarrow \alpha A'$ $\quad$ & $A' \rightarrow \beta_1 \mid \beta_2 \mid \beta_3$ $\quad$ $B \rightarrow bB \mid b$

$A \rightarrow aA'$ $\quad$ $A' \rightarrow AB \mid A \mid \epsilon$ $\quad$ $A = B$, $\alpha = b$, $\beta_1 = B$, $\beta_2 = \epsilon$

∴ the grammar after left factoring is

$A \rightarrow aA'$

$A' \rightarrow AB \mid A \mid \epsilon$

$A \rightarrow a$

$B \rightarrow bB'$

$B' \rightarrow B \mid \epsilon$

→ Ambiguity:-

A Grammar which has more than one left most derivation (or) Right most Derivation (or) parse tree for the same input string is called "Ambiguous grammar."

Ex:- Consider the grammar which has more than one left most derivation for the input string $a+a+a$

$S \rightarrow S+S$

$S \rightarrow a$

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

$A \rightarrow \beta$

1, LMD:-

$S \rightarrow S+S$

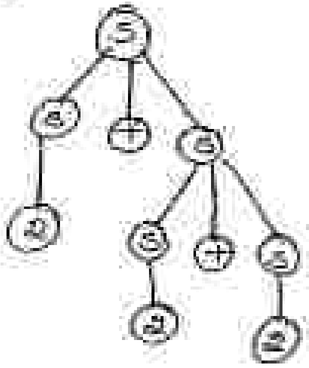
$\rightarrow a+S$

$\rightarrow a+S+S$

$\rightarrow a + a + S$ $\quad A \rightarrow \alpha A$

$\rightarrow a + a + a.$ $\quad A \rightarrow \beta$

Parse tree :-



✗ (doesn't current one)

*R, LMD

Draw back :-

The computer may be confused at the time of computing mathematical expressions due to the grammar is Ambiguous.

2 LMD :-

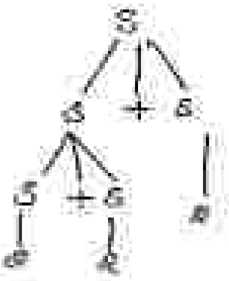
$S \rightarrow S+S$

$\rightarrow S+S+S$

$\rightarrow a+S+S$

$\rightarrow a+a+S$

$\rightarrow a+a+a$



✓

For Removing ambiguity :

1. If the grammar as left associative operators $(+, -, *, /, \pm)$ then induce left Recursion $\quad A \rightarrow A\alpha$, $A \rightarrow \beta$
2. If the grammar as Right associative operator $(=, \uparrow,)$ then induce the Right Recursion. $\quad A \rightarrow \alpha A$, $A \rightarrow \beta$

Ex:- Consider the Grammar

$S \rightarrow S+S$

$S \rightarrow S*S$

$S \rightarrow a$. And the i/p string is $a+a*a$.

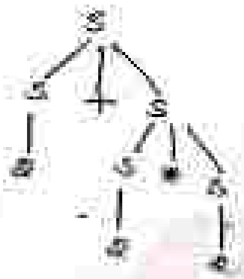
LMD :- $S \rightarrow S+S$

$\rightarrow a+S$

$\rightarrow a+S*S$

$\rightarrow a+a*S$

$\rightarrow a+a*a.$

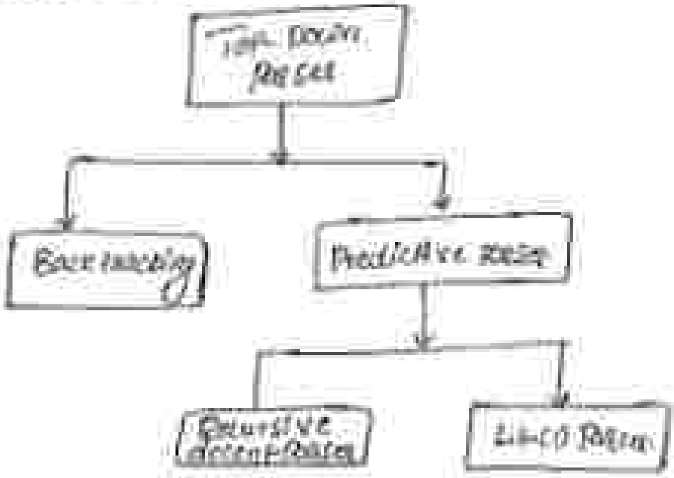


$\rightarrow$ S+S*S
$\rightarrow$ 2+S*S
$\rightarrow$ 2+2*S
$\rightarrow$ 2+2*2

$\rightarrow$ 6
X.

$S \rightarrow S+S$ $\Rightarrow$ $S \rightarrow S+T$
$S \rightarrow T$
$S \rightarrow S*S$ $\Rightarrow$ $T \rightarrow T*F$
$S \rightarrow 2$ $\Rightarrow$ $T \rightarrow F$
$F \rightarrow 2$

$\rightarrow$ Top-down Parser:-



Recursive decent Parser:-

- A Parser that uses collection of recursive Procedure for Parsing the given input string is called "Recursive Decent Parser".
- In this type of Parser the CFG is used to build the recursive Procedures.
- The RHS of the Production rule is directly converted to a Program.
- For each non-terminal a separate Procedure is where written and body of the Procedure is RHS of the corresponding non-terminal.

STEPS for Construction of Recursive decent Parser:-

The RHS of the Production rule is directly converted into Program code symbol by symbol.

Step 1:- If the i/p symbol is non-terminal then a call to the procedure corresponding to that non-terminal symbol.

Step 2:- If the i/p symbol is terminal then it is matched with the lookahead symbol from the input.

Step 3:- If the production rule has many alternatives then all the alternatives has to be combined into a single body of procedure.

Step 4:- The parser should be activated by a procedure corresponding to start symbol.

Ex:- construct a recursive decent parser for the following grammar

$E \to E+T$
$E \to T$
$T \to TF$
$T \to F$
$F \to F* | a | b$

Note:-
the recursive decent parser is works on a CFG without left recursion.

Sol:- The Given grammar is

$E \to E+T$
$E \to T$
$T \to TF$
$T \to F$
$F \to F* | a | b$.

The above grammar contains left recursion. we should eliminate left recursion from the given grammar before constructing RDP.

Elimination of Left Recursion:-

$A \to A\alpha$, $A \to \beta$ ⇒ $A \to \beta A'$, $A' \to \alpha A'$, $A' \to \epsilon$

1. $E \to E+T$
$E \to T$
⇓
$E \to TE'$
$E' \to +TE'$
$E' \to \epsilon$

2. $T \to TF$
$T \to F$
⇓
$T \to FT'$
$T' \to FT'$
$T' \to \epsilon$

3. $F \to F*$
$F \to a$
⇓
$F \to aF'$
$F' \to *F'$
$F' \to \epsilon$

4. $F \to F*$
$F \to b$
⇓
$F \to bF'$
$F' \to *F'$
$F' \to \epsilon$

∴ The resultant Grammar without left Recursion is

$E \rightarrow TE'$

$E' \rightarrow +TE'$

$E' \rightarrow \epsilon$

$T \rightarrow FT'$

$T' \rightarrow FT'$

$T' \rightarrow \epsilon$

$F \rightarrow aF'$

$F' \rightarrow *F'$

$F' \rightarrow \epsilon$

$F \rightarrow bF'$

→ Construction of Recursive decent Parser:-

```
Procedure E()
{
    T();
    E'();
    if (lookahead == $)
    printf("\n string accepted");
    else
    printf("\n string rejected");
}
Procedure E'()
{   if (lookahead == '+')
    {
        match('+');
        T();
        E'();
    }
    else
    {
        null;
    }
}
Procedure T()
{
    F();
    T'();
}

procedure T'()
{  F(); { if (true)
}T'();
```

```
Else
{ null; }
  }
}
procedure F ()
 {
     if (lookahead =='a')
     { match (a);
       F' ();
     }
     Else
     {
       if (lookahead = 'b')
       {
         match (b);
         F' ();
       }
     }
   }
Procedure F' ()
 {
     if (lookahead ='*')
     {
       match (*);
       F' ();
     }
     Else
     {
        null;
     }
 }
procedure match (char c)
{  if (lookahead ='c')
     { lookahead ++;
}
```

6

LL(1) Parser:

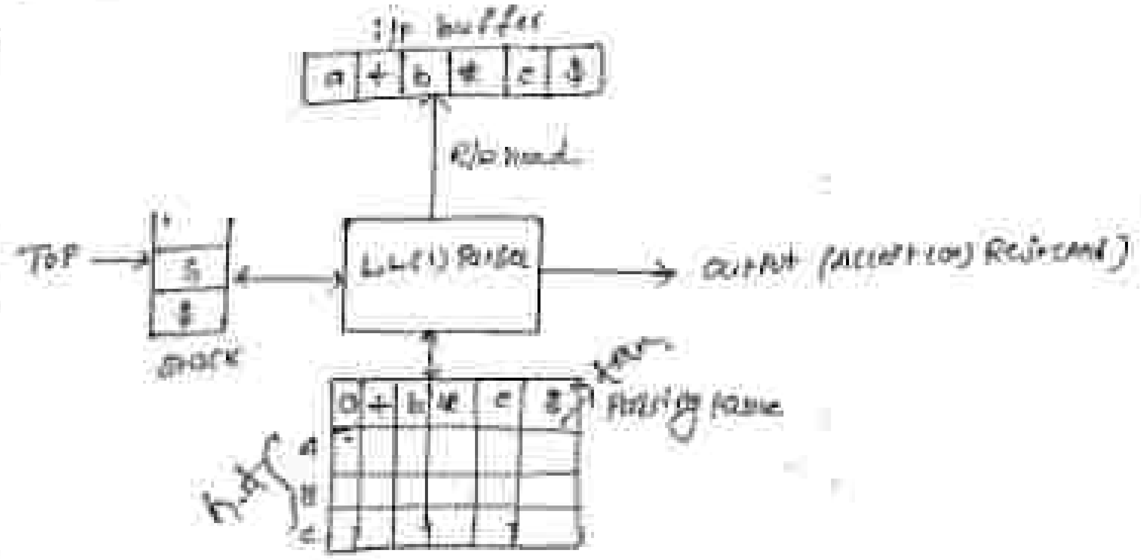
* Introduction.
* Model of LL(1) Parser.
* Construction of LL(1) Parser.

(a) Introduction:

Note:- LL Parser is works on only the CFG without left Recursion.

* Top down parser.
* Non-Recursive parser.
* Predictive parser.
* LL(1) means
  - L → Reads the given i/p string from left side to right side.
  - L → Parse the given i/p string by using "left most derivation".
  - 1 → Reads only one lookahead i/p symbol at a time.
* It represents a "LL(1) Predictive Parsing table".

(b) Model of LL(1) Parser:-



It contains three data structures like
1. Input buffer. 2. Stack 3. Parsing table.

• Input buffer:- LL Parser uses input buffer to store the i/p tokens.

• Stack:- LL Parser uses stack to hold the left sentential form i.e., the symbols in the RHS of rule are placed (Pushed) into the stack in reverse order that is from right to left.

• Parsing table:- It is a two dimensional array contains rows and columns. rows represents non-terminals, columns represent terminals. The table can be represented by M[A,a]

where A is a non terminal.
a is a current i/p symbol.

→ Construction of LL(1) Parser :-

Steps:-

1. computation of FIRST and Follow functions.
2. construction of LL(1) parse table using FIRST and Follow.
3. construction of LL(1) parsing algorithm using parse table.

1. computation of FIRST and Follow functions :-

1. FIRST function:- FIRST is a set of Terminal symbols that are FIRST symbols appearing at RHS of Production rule.

Rule for computing first function :-

1. If the terminal symbol 'a' then FIRST (a) = {a}
2. If there is a rule 'X' derives $X \rightarrow \epsilon$ then FIRST (X) = {ε}
3. for the rule $A \rightarrow x_1 x_2 x_3 \ldots x_n$ then FIRST (A) = {FIRST (x_1) ∪ FIRST (x_2) ∪ FIRST(x_3) ∪ ...... ∪ FIRST (x_n)}.

Ex: compute FIRST function on the following grammar.

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

The grammar without left recursion is

$E \rightarrow TE'$
$E' \rightarrow +TE'$
$E' \rightarrow \epsilon.$
$T \rightarrow FT'$
$T' \rightarrow *FT'$
$T' \rightarrow \epsilon$
$F \rightarrow (E)$
$F \rightarrow id.$

Production rule	FIRST
$E \rightarrow TE'$	FIRST (T) = {(, id}
$E' \rightarrow +TE' / \epsilon$	{+, ε}
$T \rightarrow FT'$	FIRST (F) = {(, id}
$T' \rightarrow *FT' / \epsilon$	{*, ε}
$F \rightarrow (E) / id$	{(, id}

2. compute FIRST function on the following grammar $S \to (L) | a$
$L \to L,S | a$

the Grammar without left recursion is

$S \to (L) | a$
$L \to L,S | S$
⇓
$L \to SL'$
$L' \to ,SL'$
$L' \to \epsilon$

$S \to (L) | a$
$L \to SL'$
$L' \to ,SL'$
$L' \to \epsilon$

production rule	FIRST
$S \to (L) \| a$	{ (, a }
$L \to SL'$	FIRST(S) = { (, a }
$L' \to ,SL' \| \epsilon$	{ , , ϵ }

3. compute FIRST function on the following grammar.

$S \to aAAb | BbBa$
$A \to \epsilon$
$B \to \epsilon$

the given grammar doesn't contain left recursion.

production rule	FIRST
$S \to aAAb \| BbBa$	{a, b}
$A \to \epsilon$	{ϵ}
$B \to \epsilon$	{ϵ}

4. compute FIRST grammar.

A $S \to aAB$ (or) bA (or) ϵ.
$A \to aAb$ (or) ϵ
$B \to bB$ (or) ϵ

production rule	FIRST
$S \to aAB \| bA \| \epsilon$	{a, b, ϵ}
$A \to aAb \| \epsilon$	{a, ϵ}
$B \to bB \| \epsilon$	{b, ϵ}

FOLLOW function:-

FOLLOW(A) is a set of terminal symbols that appear immediately to the right of (A). i.e;

$FOLLOW(A) = \{a \mid S \rightarrow \alpha A a \beta\}$

where α, β is grammar symbols.
a. is terminal symbols

Rules for Computing follow:-

1. FOLLOW (S) = {\$} where S is the start symbol.
2. If there is a production rule $A \rightarrow \alpha B \beta$ then if
   FOLLOW (B) = FIRST (β) accept 'ϵ' in FIRST (β)
3. If there is a production rule ~~$A \rightarrow B\beta$ FOLLOW (B) = FIRST (β)~~ if FIRST (β) contains ϵ.
   $A \rightarrow B\beta$ FOLLOW (B) = FIRST (β) $\cup$ FOLLOW (A) if FIRST (β) contains ϵ.
4. If there is a production rule of the form $A \rightarrow \alpha B$ then
   FOLLOW (B) = FOLLOW(A).

Ex:- find follow function on the following grammar.
$S \rightarrow Bb \mid Cd$
$B \rightarrow aB \mid \epsilon$
$C \rightarrow cC \mid \epsilon$

Sol:- FOLLOW (S) = {\$}
FOLLOW (B) = {b}
FOLLOW (C) = {d}

Ex:- compute ~~first~~ FIRST and follow functions on the following grammar.

$S \rightarrow ABCDE$
$A \rightarrow a \mid \epsilon$
$B \rightarrow b \mid \epsilon$
$C \rightarrow c$
$D \rightarrow d \mid \epsilon$
$E \rightarrow e \mid \epsilon$

Sol:- $S \rightarrow ABCDE$
FIRST (S) = FIRST (A)
= {a, ϵ}

$A \rightarrow \underline{a}$
$A \rightarrow \underline{\epsilon}$
FIRST (A) = {a} $\cup$ {ϵ}
= {a, ϵ}

$B \rightarrow \underline{b}$
$B \rightarrow \underline{\epsilon}$
FIRST (B) = {b} $\cup$ {ϵ}
= {b, ϵ}

$C \to c$
$C \to \epsilon$

$FIRST(C) = \{c\} \cup \{\epsilon\}$
$= \{c, \epsilon\}$

$D \to d$
$D \to \epsilon$

$FIRST(D) = \{d\} \cup \{\epsilon\}$
$= \{d, \epsilon\}$

$E \to e$
$E \to \epsilon$

$FIRST(E) = \{e\} \cup \{\epsilon\}$
$= \{e, \epsilon\}$

$S \to ABCDE$
$S \to aBCDE$

$FIRST(S) = \{a\}$

$S \to \epsilon BCDE$
$S \to BCDE$

$FIRST(S) = FIRST(B)$
$= \{b, \epsilon\}$

$S \to BCDE$
$S \to bCDE$

$FIRST(S) = \{b\}$

$S \to \epsilon CDE$
$S \to CDE$

$FIRST(S) = FIRST(C)$
$= \{c\}$

$\therefore FIRST(S) = \{a\} \cup \{b\} \cup \{c\}$
$= \{a, b, c\}$

$\to$ Computation of Follow:-

$S \to ABCDE$
$A \to a \mid \epsilon$
$B \to b \mid \epsilon$
$C \to c$
$D \to d \mid \epsilon$
$E \to e \mid \epsilon$

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

$FOLLOW(A) = \{b, c\}$

$FOLLOW(B) = \{c\}$

$FOLLOW(C) =$

S → ABCDE

FOLLOW(A) = FIRST(B)
= {b, c}

1. S → ABCDE
S → AbCDE
FOLLOW(A) = {b}

2. S → ABCDE
S → A$\underline{C}$CDE
S → ACDE
FOLLOW(A) = FIRST(C)
= {c}

S → ABCDE
FOLLOW(B) = FIRST(C)
= {c}

S → AB$\underline{C}$DE
FOLLOW(C) = FIRST(D)
= {d, ε}

① S → ABCDE
S → AB$\underline{C}$dE
FOLLOW(C) = {d}

② S → ABCDE
S → ABCεE
S → ABCE

→ Construction of LL(1) Parse table :

Algorithm :-

For the rule A → α of grammar 'G'.

Step1:- For each 'a' in FIRST(α) create entry 'M[A, a] = A → α
where 'a' is a terminal symbol.

Step2: For 'ε' in FIRST(α) create entry M[A, b] = A → α.
where 'b' is a terminal symbol. In FOLLOW(A).

Step3:- If ε in FIRST(α) and $ is in FOLLOW(A) then create entry
in the table M[A, $] = A → α.

Step4:- All the remaining entries in the table M are marked as "SYNTAX ERRORS".

~~Implementation~~ :- ~~Ex~~ :

Ex:- construct

$E \rightarrow TE'$

$E' \rightarrow +TE' | \epsilon$

$T \rightarrow FT'$

$T' \rightarrow *FT' | \epsilon$

$F \rightarrow (E) | id$

$\rightarrow$ FIRST (E) = FIRST (T) = { (, id }

FIRST (T) = FIRST (F) = { (, id }

FIRST (E') = { +, ε }

FIRST (T') = { *, ε }

FIRST (F) = { (, id }

FOLLOW (E) = { $,) }

FOLLOW (E') = FOLLOW (E) = { $,) }

FOLLOW (T) = { +, $,) }

FOLLOW (T') = { +, $,) }

$E \rightarrow TE'$

FOLLOW(T) = FIRST(E')

= { +, ε }

$E \rightarrow TE$

$E \rightarrow T$

FOLLOW(T) = FOLLOW(E)

= { $,) }

∴ FOLLOW(T) = { +, $,) }

$E' \rightarrow TE$

$E' \rightarrow +T$

FOLLOW(T) = FOLLOW(E')

= { $,) }

$T' \rightarrow *FT'$

FOLLOW(F) = FIRST (T')

= { *, ε }

$T' \rightarrow *F\epsilon$

$T' \rightarrow *F$

FOLLOW(F) = FOLLOW (T') = { +, $,) }

$T \rightarrow FT'$

FOLLOW(F) = FIRST(T')

= { *, ε }

$T \rightarrow FF$

$T \rightarrow F$

FOLLOW (F) = FOLLOW(T)

= { +, $,) }

$T \rightarrow FT'$

FOLLOW (T') = FOLLOW(T)

= { +, $,) }

$\boxed{\therefore \text{ FOLLOW (F) = \{ +, *, \$,) \}}}$

LL(1) Parser table :-

	+	*	(	)	id	$
E			E→TE'		E→TE'	
E'	E'→+TE'			E'→ε		E'→ε
T			T→FT'		T→FT'	
T'	T'→ε	T'→*FT'		T'→ε		T'→ε
F			F→(E)		F→id	

$\rightarrow E \rightarrow TE'$

FIRST (TE') = FIRST(T)

= { (, id }

$E' \rightarrow +TE'$

FIRST (+TE') = FIRST(+)

= { + }

4. $T \rightarrow FT'$
FIRST(FT') = FIRST(F) = {(, id}

3. $E' \rightarrow \epsilon$
FIRST(ϵ) = {ϵ}
FOLLOW(E') = {$,)}

5. $T' \rightarrow *FT'$
FIRST(*FT') = FIRST(*) = {*}

6. $T' \rightarrow \epsilon$
FIRST(ϵ) = {ϵ}
FOLLOW(T') = {+, $,)}

7. $F \rightarrow (E)$
FIRST((E)) = {(}

$F \rightarrow id$
FIRST(id) = {id}.

LL(1) Parsing algorithm:-

consider the i/p string ~~id+id*id~~. ((a),a)

STACK	Input string	ACTION
$S	((a),a)$	$S \rightarrow (L)$
$)L(	((a),a)$	pop
$)L	(a),a)$	$L \rightarrow SL'$
$)L'S	(a),a)$	$S \rightarrow (L)$
$)L')L(	(a),a)$	pop
$)L')L	a),a)$	$L \rightarrow SL'$
$)L')L'S	a),a)$	$S \rightarrow a$
$)L')L'a	a),a)$	pop
$)L')L'	),a)$	$L' \rightarrow \epsilon$
$)L')$\epsilon$	),a)$	–
$)L')	),a)$	pop
$)L'	,a)$	$L' \rightarrow SL'$
$)L'S,	,a)$	pop
$)L'S	a)$	$S \rightarrow a$
$)L'a	a)$	pop
$)L'	)$	$L' \rightarrow \epsilon$
$)$\epsilon$	)$	–
$)	)$	pop
		accepted.

(1) construct LL(1) parse table for the following grammar

$S \rightarrow (L) | a$

$L \rightarrow L, S | S$ and check the input string ((a), a).

is replaced or not by the LL(1) parser.

(2) construct LL(1) parser table for the following grammar.

$S \rightarrow iEtSS'$

$S \rightarrow a$

$S' \rightarrow eS | \epsilon$

$E \rightarrow b$

Error Recovery in predictive Parser:-

→ An Error is detected during predictive parsing when the terminal on the top of the stack does not match the next input symbol (or).

→ when non-terminal A on the top of the stack a is the next input symbol and parsing table entry & M [A,a] is empty

→ the process of reducing number of Errors in the parser table is called Error recovery.

→ LL(1) parser uses "panic mode" Error recovery technique

"Panic mode Error recovery":

It is based on the idea of skipping symbols on the i/p until a synchronizing token is selected.

synchronizing token:- It is a set of terminals obtained from follow of non-terminal in the given grammar

Ex: FOLLOW(E) = {$}

FOLLOW(E') = {$,)}

FOLLOW(T) = {+, $,)}

FOLLOW(T') = {+, $,)}

FOLLOW(F) = {*, +, $,)}

After applying panic mode error recovery technique modified LL(1) parser table is.

	+	*	(	)	id	$
E			E→TE'	sync	E→TE'	sync
E'	E'→+TE'			E'→ε		E'→ε
T	sync		T→FT'	sync	T→FT'	sync
T'	T'→ε	T'→*FT'		T'→ε		T'→ε
F	sync	sync	F→(E)	sync	F→id	sync

Parsing Algorithm:-

* If the parser look up the entry M[A,a] as a blank then the i/p symbol 'a' is skipped.
* If the entry is 'sync' then the non-terminal the top of the stack is poped.
* If the token on the top of the stack doesn't match the i/p symbol then we pop off the token from the stack.

Ex:- consider the i/p string +id * * id.

Stack	i/p string	Action
$E	+id**id$	skipped +
$E	id **id$	E→TE'
$E'T	id ** id$	T→FT'
$E'T'F	id **id $	F→id
$E'T'id	id** id$	pop
$E'T'	** id$	T'→*FT'
$E'T'F*	** id$	pop
$E'T'F	* id$	sync, pop
$E'T'	* id$	T'→*FT'
$E'T'F*	* id$	pop
$E'T'F	id$	F→id
$E'T'id	id$	pop
$E'T'	$	T'→ε
$E'		