



UNIVERSITY WITH A PURPOSE



Content of this lecture

- Introduction to Parsing Technique
- Top down parser

Parsing Techniques

- Parsing technique is used to check tokens generated are satisfying the grammatical rules of the language or not, if they satisfy then corresponding Parse tree is generated otherwise error is reported to error handler.
- Based on the methodology adopted for the generation of parse tree, parsing techniques can be categorized in to two categories:
 - Top down parser
 - Bottom parser

Parsing Techniques

- Top Down Parser
 - Input is scanned from left to right one symbol at a time
 - Build the parse tree from top(root) to bottom(leaves)
 - We have to decide which rule $A \rightarrow \beta$ should be applied to a node labelled A.
 - Expanding A results in new nodes (children of A) labelled by symbols in β
 - It can be also viewed as finding a leftmost derivation for an input string
- Bottom up Parser
 - Input is scanned from left to right one symbol at a time
 - Build the parse tree from bottom(leaves) to top(root).
 - Input string will be reduced to starting symbol.
 - Reducing rule $A \rightarrow \beta$ results in adding a node A to the tree. A's children are the nodes labelled by the symbols in β

Parsing Techniques

- Consider the grammar $G1 = (\{E, F, T\}, \{a, (,), *, +\}, P, E)$ where the productions are:

$$E \rightarrow E + T$$

$$E \rightarrow T$$

$$T \rightarrow T * F$$

$$T \rightarrow F$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

and the string $x = a + a * a$

Parsing Techniques

- Top-down method: rules are considered in the same order as a leftmost derivation

$$E \Rightarrow \underline{E} + T$$

$$E \Rightarrow \underline{I} + T$$

$$E \Rightarrow \underline{F} + T$$

$$E \Rightarrow a + \underline{I}$$

$$E \Rightarrow a + \underline{I} * F$$

$$E \Rightarrow a + \underline{F} * F$$

$$E \Rightarrow a + a * \underline{F}$$

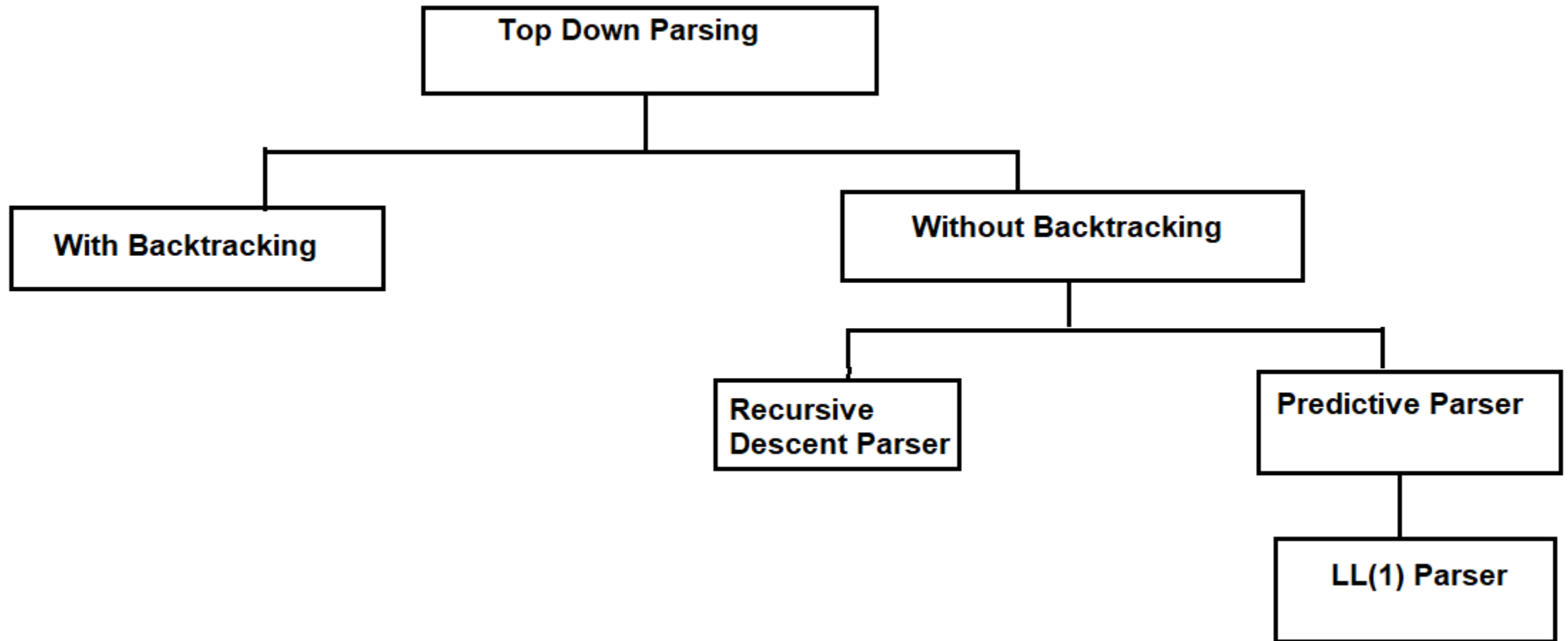
$$E \Rightarrow a + a * a$$

Parsing Techniques

- Bottom-up method: rules are considered in the same order as a reverse rightmost derivation

$a + a * a$
 $\Leftarrow \underline{F} + a * a$
 $\Leftarrow \underline{T} + a * a$
 $\Leftarrow \underline{E} + a * a$
 $\Leftarrow E + \underline{F} * a$
 $\Leftarrow E + \underline{T} * a$
 $\Leftarrow E + T * \underline{F}$
 $\Leftarrow E + \underline{T}$
 $\Leftarrow \underline{E}$

Top Down Parsing Techniques



Top Down Parsing with Backtracking

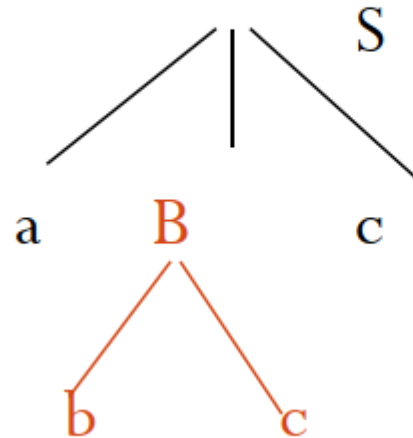
- In this category, parser can make repeated scans of input.
- If required input string is not achieved by applying one production rule then another production rule can be applied at each step to get the required string.

Top Down Parsing with Backtracking

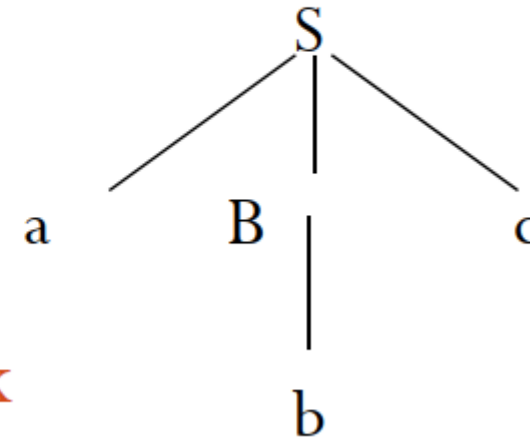
$S \rightarrow aBc$

$B \rightarrow bc \mid b$

input: abc



fails,
backtrack



Top Down Parsing with Backtracking Algorithm

- $S \rightarrow aAd$
- $A \rightarrow bc \mid b$

Top Down Parsing with Backtracking Algorithm

- $S \rightarrow aAd$
- $A \rightarrow bc \mid b$

```
Procedure S ()
{
    if (input symbol = 'a') then
    {
        advance();
        if A() then
        {
            if (input symbol='d') then
            {
                advance();
                declare success;
            }
            else error;
        }
        else error;
    }
    else error
}
```

Top Down Parsing with Backtracking Algorithm

- $S \rightarrow aAd$
- $A \rightarrow bc|b$

```
Procedure S ()
{
    if (input symbol = 'a') then
    {
        advance();
        if A() then
        {
            if (input symbol='d') then
            {
                advance();
                declare success;
            }
            else error;
        }
        else error;
    }
    else error
}
```

```
Procedure A ()
{
    if (input symbol = 'b') then
    {
        advance();
        if (input symbol='c') then
        {
            advance();
            declare success;
        }
        else error;
    }
    if (input symbol = 'b') then
    {
        advance();
        declare success;
    }
    else error;
}
```

Top Down Parsing with Backtracking- Limitations

- Its not efficient as undoing semantic action requires lot of overhead.
- Entries made in the symbol table during parsing have to be removed while backtracking.
- The order in which the alternative production rules are applied can also affect the efficiency.
- Existence of left recursion will cause parser to enter into infinite loop. This must me removed.
- Its difficult to locate the position of error on occurrence of failure.

Top Down Parsing without Backtracking

- In this category, once the production rule is applied, it can't be undone. The parser which follows top down parsing without backtracking are:
 - Recursive Descent Parser
 - Predictive Parser
 - LL(1) parser
- Need to remove Left recursion and Left factoring from grammar before doing top down parsing without backtracking.

Recursive Descent Parser

- It is defined as a parser that uses various recursive procedures to process the input string with no backtracking.
- It can be easily implemented using a language which is recursive in nature.
- The first symbol of the string on R.H. S. of a production will uniquely determines the correct alternative to choose.
- In this technique, there is no need for the procedures to return an indication of success or failure since the calling procedure has no scope of trying another alternate. Rather on failure, an error correcting routine is called.

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \mathbf{id}$

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \mathbf{id}$

Procedure E()

{

 T();

 E'();

}

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \mathbf{id}$

Procedure E()

{

 T();

 E'();

}

Procedure E' ()

{

 if (input symbol = +) then

 {

 advance();

 T();

 E'();

 }

}

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \text{id}$

Procedure E()

{

 T();

 E'();

}

Procedure E' ()

{

 if (input symbol = +) then

 {

 advance();

 T();

 E'();

 }

}

Procedure T()

{

 F();

 T'();

}

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \mathbf{id}$

Procedure E()

{

 T();

 E'();

}

Procedure E' ()

{

 if (input symbol = +) then

 {

 advance();

 T();

 E'();

 }

}

Procedure T()

{

 F();

 T'();

}

Procedure T' ()

{

 if (input symbol = *) then

 {

 advance();

 F();

 T'();

 }

}

Recursive Descent Parser

- Write down the algorithm using recursive procedure to implement following grammar:

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid \text{id}$

Procedure E()

```
{
    T();
    E'();
}
```

Procedure E' ()

```
{
    if (input symbol = '+') then
    {
        advance();
        T();
        E'();
    }
}
```

Procedure T()

```
{
    F();
    T'();
}
```

Procedure T' ()

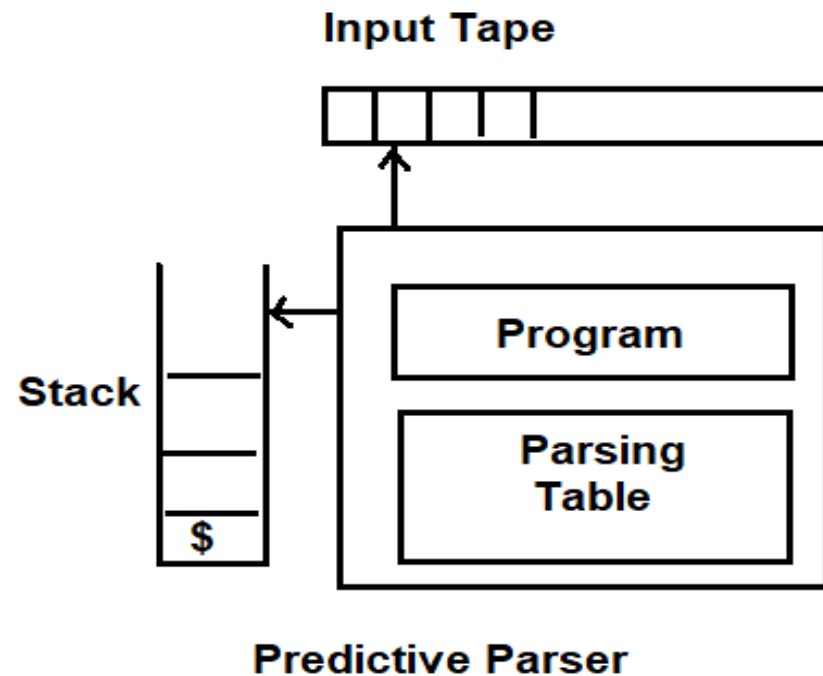
```
{
    if (input symbol = '*') then
    {
        advance();
        F();
        T'();
    }
}
```

Procedure F ()

```
{
    if (input symbol = 'id')
    {
        advance();
    }
    else if (input symbol = '(')
    {
        advance();
        E();
        if (input symbol = ')')
            advance();
        else error();
    }
    else error();
}
```

Predictive Parser

- It is a top down backtracking less parser which generates the parse tree in LL(1) manner.
- There should be no left recursion and left factoring in the grammar.



Components of Predictive Parser

- Input tape: which contains the string to be parsed, only terminals are there in the input tape. \$ symbol is there at the end of the string which indicates that string has finished.
- Stack: which contains sequence of grammar symbols (terminals and non terminals both) preceded by \$ symbol. \$ indicates the bottom of the stack.
- Parsing Table: It is a two dimensional array $M[A, a]$ where A is representing the non terminal and a is representing the terminal or \$ symbol.
- Program: Parser is controlled by a program.

How Program controls Parser

- If current input symbol and top stack symbol are \$ then parser halts and announce successful completion of parsing.
- If current input symbol and top stack symbol matches but both are not \$ then parser pop off the top stack symbol and advances the input pointer to next input symbol.
- If top of the stack is Non terminal (X) and present input symbol is 'a' then program consults entry M [X,a] in the parsing table and take the action accordingly.
 - If this entry is an error entry then parser calls the error recovery routine.
 - If $M[X,a] = \{X \rightarrow UvW\}$ then parser replaces the X on top of the stack by WvU where U will be top.
- Initial configuration of parser is Stack: \$S Input: w\$
- To fill the entries of parsing table, we use First() and Follow() functions.

Steps for finding First

- $\text{First}(\alpha)$ represents the set of the terminal symbols which occurs as first symbol in the strings derived from α where α is any string of grammar symbols.
 - If $X \rightarrow a$ then $\text{First}(X) = a$
 - If $X \rightarrow \lambda$ then $\text{First}(X) = \lambda$
 - First of terminal will be that terminal itself
 - If $X \rightarrow a\alpha$ then $\text{First}(X) = a$
 - If $X \rightarrow Y_1 Y_2 Y_3 \dots Y_k$ then
 - $\text{First}(X) = \text{First}(Y_1)$;if $\text{First}(Y_1)$ does not contain λ
 - $\text{First}(X) = \text{First}(Y_1) - \{\lambda\} \cup \text{First}(Y_2)$; if $\text{First}(Y_1)$ contain λ and $\text{First}(Y_2)$ does not contain λ
 - $\text{First}(X) = \text{First}(Y_1) - \{\lambda\} \cup \text{First}(Y_2) - \{\lambda\} \cup \text{First}(Y_3)$; if $\text{First}(Y_1)$ and $\text{First}(Y_2)$ contain λ and $\text{First}(Y_3)$ does not contain λ

Steps for finding Follow

- Follow(A) is the set of the terminals which occur immediately after (follow) the non-terminal A in the strings derived from the starting symbol
 - \$ is in Follow(X) if X is start symbol
 - If $A \rightarrow \alpha B \beta$ and $\beta \neq \lambda$ then
 - If first(β) does not contain λ then Follow(B)={First(β)}
 - If first(β) contains λ then Follow(B)={First(β)}- $\{\lambda\}$ U Follow(A)
 - If $A \rightarrow \alpha B$ then Follow(B)={Follow(A)}

Numerical

- Find first and follow for the following grammar
 - $S \rightarrow AaA \mid BbB$
 - $A \rightarrow bB$
 - $B \rightarrow \epsilon$

Numerical

- Find first and follow for the following grammar
 - $S \rightarrow AaA \mid BbB$
 - $A \rightarrow bB$
 - $B \rightarrow \epsilon$

 - $\text{First}(B) = \epsilon$
 - $\text{First}(A) = b$
 - $\text{First}(S) = b$

Numerical

- Find first and follow for the following grammar

- $S \rightarrow AaA \mid BbB$

- $A \rightarrow bB$

- $B \rightarrow \epsilon$

- $\text{First}(B) = \epsilon$ $\text{Follow}(B) = \{\$, a, b\}$

- $\text{First}(A) = b$ $\text{Follow}(A) = \{\$, a\}$

- $\text{First}(S) = b$ $\text{Follow}(S) = \{\$\}$

Numerical

- Find first and follow for the following grammar
 - $E \rightarrow E+T \mid T$
 - $T \rightarrow T * F \mid F$
 - $F \rightarrow (E) \mid id$

Numerical

- Find first and follow for the following grammar
 - $E \rightarrow E+T \mid T$
 - $T \rightarrow T * F \mid F$
 - $F \rightarrow (E) \mid id$

Numerical

- Find first and follow for the following grammar
 - $E \rightarrow E+T \mid T$
 - $T \rightarrow T * F \mid F$
 - $F \rightarrow (E) \mid id$

- $\text{First}(E) = \{ (, id \}$
- $\text{First}(T) = \{ (, id \}$
- $\text{First}(F) = \{ (, id \}$

Numerical

- Find first and follow for the following grammar

- $E \rightarrow E+T \mid T$

- $T \rightarrow T * F \mid F$

- $F \rightarrow (E) \mid id$

- $First(E) = \{ (, id \}$ $Follow(E) = \{ \$, +,) \}$

- $First(T) = \{ (, id \}$ $Follow(T) = \{ \$, +,), * \}$

- $First(F) = \{ (, id \}$ $Follow(F) = \{ \$, +,), * \}$

Numerical

- Find first and follow for the following grammar
 - $S \rightarrow L=R$
 - $S \rightarrow R$
 - $L \rightarrow *R$
 - $L \rightarrow \text{id}$
 - $R \rightarrow L$

Numerical

- Find first and follow for the following grammar
 - $S \rightarrow L = R$
 - $S \rightarrow R$
 - $L \rightarrow *R$
 - $L \rightarrow \text{id}$
 - $R \rightarrow L$

- $\text{First}(S) = \{*, \text{id}\}$
- $\text{First}(L) = \{*, \text{id}\}$
- $\text{First}(R) = \{*, \text{id}\}$

Numerical

- Find first and follow for the following grammar

- $S \rightarrow L=R$

- $S \rightarrow R$

- $L \rightarrow *R$

- $L \rightarrow \text{id}$

- $R \rightarrow L$

- $\text{First}(S) = \{*, \text{id}\}$

- $\text{Follow}(S) = \{\$ \}$

- $\text{First}(L) = \{*, \text{id}\}$

- $\text{Follow}(L) = \{\$, =\}$

- $\text{First}(R) = \{*, \text{id}\}$

- $\text{Follow}(R) = \{\$, =\}$

Numerical

- Find first and follow for the following grammar
 - $E \rightarrow TE'$
 - $E' \rightarrow +TE' \mid \varepsilon$
 - $T \rightarrow FT'$
 - $T' \rightarrow *FT' \mid \varepsilon$
 - $F \rightarrow (E) \mid id$

Numerical

- Find first and follow for the following grammar
 - $E \rightarrow TE'$
 - $E' \rightarrow +TE' \mid \varepsilon$
 - $T \rightarrow FT'$
 - $T' \rightarrow *FT' \mid \varepsilon$
 - $F \rightarrow (E) \mid id$
 - $First(E) = \{ (, id \}$
 - $First(T) = \{ (, id \}$
 - $First(F) = \{ (, id \}$
 - $First(E') = \{ +, \varepsilon \}$
 - $First(T') = \{ *, \varepsilon \}$

Numerical

- Find first and follow for the following grammar

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

– $\text{First}(E) = \{ (, id \}$

$\text{Follow}(E) = \{), \$ \}$

– $\text{First}(T) = \{ (, id \}$

$\text{Follow}(T) = \{ +,), \$ \}$

– $\text{First}(F) = \{ (, id \}$

$\text{Follow}(F) = \{ +, *,), \$ \}$

– $\text{First}(E') = \{ +, \epsilon \}$

$\text{Follow}(E') = \{), \$ \}$

– $\text{First}(T') = \{ *, \epsilon \}$

$\text{Follow}(T') = \{ +,), \$ \}$

Rules for filling the entries in Parsing Table

- For each production rule $A \rightarrow \alpha$ in the grammar do the following:
 - For each terminal a in $\text{First}(\alpha)$ add $A \rightarrow \alpha$ in $M[A, a]$
 - If ϵ is in $\text{First}(\alpha)$, and b is in $\text{Follow}(A)$ then add $A \rightarrow \alpha$ to $M[A, b]$
 - If ϵ is in $\text{First}(\alpha)$, and $\$$ is in $\text{Follow}(A)$ then add $A \rightarrow \alpha$ to $M[A, \$]$
 - All the remaining entries in the table are error.

Algorithm-Predictive Parser

- Input: String to be parsed and parsing table M for Grammar
- Output: Either stack will become empty or it will give an error
- Steps:
 - While(Stack is not empty)do
 - {
 - Let X be top stack symbol and let a be next input symbol
 - If (X is terminal or \$) then
 - If(X=a) then
 - » Pop X from stack and remove a from the input
 - Else
 - » Error()
 - Else
 - If $M[X,a]=A \rightarrow Y_1Y_2...Y_n$ then
 - » Pop X and Push $Y_nY_{n-1}...Y_1$ onto stack
 - Else
 - » Error()
 - }

Numerical

- Construct predictive parsing table for the following grammar and check whether string $id+id*id$ is accepted or not.
 - $E \rightarrow TE'$
 - $E' \rightarrow +TE' \mid \varepsilon$
 - $T \rightarrow FT'$
 - $T' \rightarrow *FT' \mid \varepsilon$
 - $F \rightarrow (E) \mid id$

Numerical

- $E \rightarrow TE'$
- $E' \rightarrow +TE' \mid \varepsilon$
- $T \rightarrow FT'$
- $T' \rightarrow *FT' \mid \varepsilon$
- $F \rightarrow (E) \mid id$

- Step 1: Eliminate left recursion and left factoring: Already removed

Numerical

- $E \rightarrow TE'$
- $E' \rightarrow +TE' \mid \varepsilon$
- $T \rightarrow FT'$
- $T' \rightarrow *FT' \mid \varepsilon$
- $F \rightarrow (E) \mid id$

- Step 1: Eliminate left recursion and left factoring: Already removed
- Step 2: Compute First of all non terminals: $First(E) = First(T) = First(F) \{ (, id \}$ $First(E') = \{ +, \varepsilon \}$ $First(T') = \{ *, \varepsilon \}$

Numerical

- $E \rightarrow TE'$
- $E' \rightarrow +TE' \mid \varepsilon$
- $T \rightarrow FT'$
- $T' \rightarrow *FT' \mid \varepsilon$
- $F \rightarrow (E) \mid id$

- Step 1: Eliminate left recursion and left factoring: Already removed
- Step 2: Compute First of all non terminals: $First(E) = First(T) = First(F) \{ (, id \}$ $First(E') = \{ +, \varepsilon \}$ $First(T') = \{ *, \varepsilon \}$
- Step 3: Compute Follow of all non terminals: $Follow(E) = Follow(E') \{), \$ \}$ $Follow(T) = Follow(T') \{ +,), \$ \}$
 $Follow(F) = \{ +, *,), \$ \}$

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

Non Terminal			INPUT SYMBOLS			
	id	+	*	()	\$
E						
E'						
T						
T'						
F						

- Step 1: Eliminate left recursion and left factoring: Already removed
- Step 2: Compute First of all non terminals: $\text{First}(E) = \text{First}(T) = \text{First}(F) \{ (, id \}$ $\text{First}(E') = \{ +, \epsilon \}$ $\text{First}(T') = \{ *, \epsilon \}$
- Step 3: Compute Follow of all non terminals: $\text{Follow}(E) = \text{Follow}(E') \{), \$ \}$ $\text{Follow}(T) = \text{Follow}(T') \{ +,), \$ \}$
 $\text{Follow}(F) = \{ +, *,), \$ \}$
- Step 4: Construct the predictive parsing table

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

Non Terminal			INPUT SYMBOLS			
	id	+	*	()	\$
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E \rightarrow +TE'$			$E' \rightarrow \epsilon$	$E' \rightarrow \epsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \epsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$
F	$F \rightarrow id$			$F \rightarrow (E)$		

- Step 1: Eliminate left recursion and left factoring: Already removed
- Step 2: Compute First of all non terminals: $\text{First}(E) = \text{First}(T) = \text{First}(F) \{ (, id \}$ $\text{First}(E') = \{ +, \epsilon \}$ $\text{First}(T') = \{ *, \epsilon \}$
- Step 3: Compute Follow of all non terminals: $\text{Follow}(E) = \text{Follow}(E') \{), \$ \}$ $\text{Follow}(T) = \text{Follow}(T') \{ +,), \$ \}$
 $\text{Follow}(F) = \{ +, *,), \$ \}$
- Step 4: Construct the predictive parsing table

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

Non Terminal	INPUT SYMBOLS					
	id	+	*	()	\$
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \epsilon$	$E' \rightarrow \epsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \epsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$
F	$F \rightarrow id$			$F \rightarrow (E)$		

- Step 1: Eliminate left recursion and left factoring: Already removed
- Step 2: Compute First of all non terminals: $\text{First}(E) = \text{First}(T) = \text{First}(F) \{ (, id \}$ $\text{First}(E') = \{ +, \epsilon \}$ $\text{First}(T') = \{ *, \epsilon \}$
- Step 3: Compute Follow of all non terminals: $\text{Follow}(E) = \text{Follow}(E') \{), \$ \}$ $\text{Follow}(T) = \text{Follow}(T') \{ +,), \$ \}$
 $\text{Follow}(F) = \{ +, *,), \$ \}$
- Step 4: Construct the predictive parsing table.
- Step 5: Check the acceptance of the string $id+id*id$ using predictive parsing program.

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

Non Termin al	INPUT SYMBOLS					
	id	+	*	()	\$
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E \rightarrow +TE'$			$E' \rightarrow \epsilon$	$E' \rightarrow \epsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \epsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$
F	$F \rightarrow id$			$F \rightarrow (E)$		

MATCHED	STACK	INPUT	ACTION
	E\$	id+id * id\$	
	TE'\$	id+id * id\$	E->TE'
	FT'E'\$	id+id * id\$	T->FT'
	id T'E'\$	id+id * id\$	F->id
id	T'E'\$	+id * id\$	Match id
id	E'\$	+id * id\$	T'->€
id	+TE'\$	+id * id\$	E'-> +TE'
id+	TE'\$	id * id\$	Match +
id+	FT'E'\$	id * id\$	T-> FT'
id+	idT'E'\$	id * id\$	F-> id
id+id	T'E'\$	* id\$	Match id
id+id	* FT'E'\$	* id\$	T'-> *FT'
id+id *	FT'E'\$	id\$	Match *
id+id *	idT'E'\$	id\$	F-> id
id+id * id	T'E'\$	\$	Match id
id+id * id	E'\$	\$	T'-> €
id+id * id	\$	\$	E'-> €

LL(1) Grammar

- The algorithm for constructing a predictive parsing table can be applied to any grammar G to produce a parsing table M .
- For some grammars however, M may have some entries that are multiply-defined.
- If G is left recursive or ambiguous, then M will have at least one multiply-defined entry.
- A grammar whose parsing table has no multiply-defined entries is said to be LL(1).

LL(1) Grammar

- LL(1) grammars have several distinctive properties.
 - Parsing table has no multiply-defined entries is said to be LL(1).
 - No ambiguous or left recursive grammar can be LL(1).
 - A grammar G is LL(1) iff whenever $A \rightarrow \alpha \mid \beta$ are two distinct productions of G , the following conditions hold :-
 - $FIRST(\alpha)$, $FIRST(\beta)$ must be disjoint.
 - At most one of the strings α or β can derive ϵ i.e. $FIRST(\alpha) \cap FIRST(\beta) \neq \epsilon$.
 - If $\beta \Rightarrow^* \epsilon$, then $FIRST(\alpha) \cap FOLLOW(A) = \phi$ i.e. α does not derive any string beginning with the terminal in $FOLLOW(A)$
- In predictive parser when we restrict grammar to be only LL(1) grammar then it is called LL(1) parser.

LL(1) Grammar

$G1 = (\{E, F, T\}, \{a, (,), *, +\}, P, E)$ where the productions are:

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid a$

Is this LL(1) grammar

LL(1) Grammar

$G1 = (\{E, F, T\}, \{a, (,), *, +\}, P, E)$ where the productions are:

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid a$

It is not LL(1) grammar

LL(1) Grammar

$G2 = (\{S, A\}, \{a, b, c, d\}, P, S)$ where the productions are:

$S \rightarrow cAd$

$A \rightarrow ab|a$

Is this LL(1) grammar?

LL(1) Grammar

$G_2 = (\{S, A\}, \{a, b, c, d\}, P, S)$ where the productions are:

$S \rightarrow cAd$

$A \rightarrow ab|a$

It is not LL(1) grammar?

LL(1) Grammar

$G_3 = (\{E, E', T, T', F\}, \{a, *, +, (,)\}, P, E)$ with the set of productions P :

$E \rightarrow TE'$ $T \rightarrow FT'$ $F \rightarrow (E) \mid a$

$E' \rightarrow +TE' \mid \varepsilon$ $T' \rightarrow *FT' \mid \varepsilon$

Is this LL(1) grammar?

LL(1) Grammar

$G_3 = (\{E, E', T, T', F\}, \{a, *, +, (,)\}, P, E)$ with the set of productions P :

$E \rightarrow TE'$ $T \rightarrow FT'$ $F \rightarrow (E) \mid a$

$E' \rightarrow +TE' \mid \varepsilon$ $T' \rightarrow *FT' \mid \varepsilon$

It is LL(1) grammar

LL(1) Grammar

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

Is this LL(1) grammar

LL(1) Grammar

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string abba

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string 'abba'

$\text{First}(S) = \{a\}$

$\text{First}(B) = \{b, \epsilon\}$

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string abba

$\text{First}(S) = \{a\}$ $\text{Follow}(S) = \{\$ \}$

$\text{First}(B) = \{b, \epsilon\}$ $\text{Follow}(B) = \{a\}$

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string abba

$\text{First}(S) = \{a\}$

$\text{Follow}(S) = \{\$ \}$

$\text{First}(B) = \{b, \epsilon\}$

$\text{Follow}(B) = \{a\}$

	a	b	\$
S			
B			

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string abba

First (S)={a} Follow(S)={\$}

First (B)={b, ϵ } Follow(B)={a}

	a	b	\$
S	$S \rightarrow aBa$		
B	$B \rightarrow \epsilon$	$B \rightarrow bB$	

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar, now do the parsing of the string abba

First (S)={a}

Follow(S)={\$}

First (B)={b, ϵ }

Follow(B)={a}

<u>stack</u>	<u>input</u>	<u>output</u>
\$S	abba\$	$S \rightarrow aBa$
\$aBa	abba\$	
\$aB	bba\$	$B \rightarrow bB$
\$aBb	bba\$	
\$aB	ba\$	$B \rightarrow bB$
\$aBb	ba\$	
\$aB	a\$	$B \rightarrow \epsilon$
\$a	a\$	
\$	\$	accept, successful completion

	a	b	\$
S	$S \rightarrow aBa$		
B	$B \rightarrow \epsilon$	$B \rightarrow bB$	

LL(1) Parser

$G_4 = (\{S, B\}, \{a, b\}, P, S)$ with the set of productions P :

$S \rightarrow aBa$

$B \rightarrow bB \mid \epsilon$

It is LL(1) grammar,

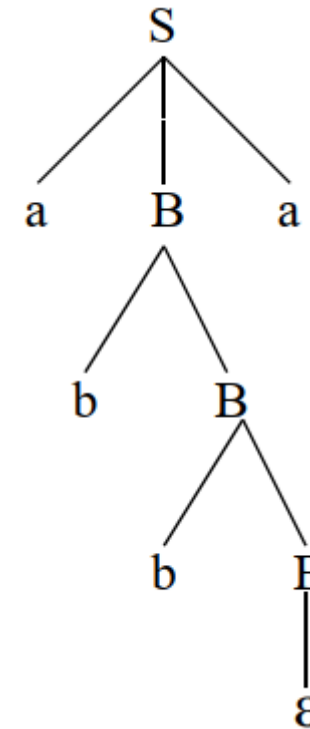
now do the parsing of the string abba

$\text{First}(S) = \{a\}$

$\text{Follow}(S) = \{\$ \}$

$\text{First}(B) = \{b, \epsilon\}$

$\text{Follow}(B) = \{a\}$



<u>stack</u>	<u>input</u>	<u>output</u>
$\$S$	abba\$	$S \rightarrow aBa$
$\$aBa$	abba\$	
$\$aB$	bba\$	$B \rightarrow bB$
$\$aBb$	bba\$	
$\$aB$	ba\$	$B \rightarrow bB$
$\$aBb$	ba\$	
$\$aB$	a\$	$B \rightarrow \epsilon$
$\$a$	a\$	
$\$$	\$	accept, successful completion

	a	b	\$
S	$S \rightarrow aBa$		
B	$B \rightarrow \epsilon$	$B \rightarrow bB$	

LL(1) Grammar

$G_1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$

Construct parsing table for this and check if this grammar is LL(1)

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$ Construct parsing table for this and check if this grammar is LL(1)

Step 1: Remove Left recursion

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$ Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$ Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

Step 3: Calculate First and Follow.

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$

Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

Step 3: Calculate First and Follow.

$\text{First}(S) = \{i, a\}$

$\text{First}(S') = \{e, \epsilon\}$

$\text{First}(C) = \{b\}$

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$ Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

Step 3: Calculate First and Follow.

$\text{First}(S) = \{i, a\}$ $\text{Follow}(S) = \{\$, e\}$

$\text{First}(S') = \{e, \epsilon\}$ $\text{Follow}(S') = \{\$, e\}$

$\text{First}(C) = \{b\}$ $\text{Follow}(C) = \{t\}$

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$

Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

Step 3: Calculate First and Follow.

$\text{First}(S) = \{i, a\}$ $\text{Follow}(S) = \{\$, e\}$

$\text{First}(S') = \{e, \epsilon\}$ $\text{Follow}(S') = \{\$, e\}$

$\text{First}(C) = \{b\}$ $\text{Follow}(C) = \{t\}$

Step 4: Construct parsing table.

	a	b	e	i	t	\$
S						
S'						
C						

LL(1) Grammar

$G1 = (\{S, S', C\}, \{i, t, a, e, b\}, P, S)$ where the productions are:

$S \rightarrow iCtSS' \mid a$

$S' \rightarrow eS \mid \epsilon$

$C \rightarrow b$

Construct parsing table for this and check is this grammar LL(1)

Step 1: Remove Left recursion

Step 2: Remove Left Factoring

Step 3: Calculate First and Follow.

$\text{First}(S) = \{i, a\}$ $\text{Follow}(S) = \{\$, e\}$

$\text{First}(S') = \{e, \epsilon\}$ $\text{Follow}(S') = \{\$, e\}$

$\text{First}(C) = \{b\}$ $\text{Follow}(C) = \{t\}$

Step 4: Construct parsing table.

	a	b	e	i	t	\$
S	$S \rightarrow a$			$S \rightarrow iCtSS'$		
S'			$S' \rightarrow eS$ $S' \rightarrow \epsilon$			$S' \rightarrow \epsilon$
C		$C \rightarrow b$				

LL(1) Grammar

$G1 = (\{S, L\}, \{ (,), a, , \}, P, S)$ where the productions are:

$S \rightarrow (L) \mid a$

$L \rightarrow L, S \mid S$

Construct parsing table for this, check is this grammar LL(1), also check string '(a,a)' is accepted:

THANK YOU



Dr Anurag Jain

Assistant Professor (SG), Virtualization Department, School of Computer Science,
University of Petroleum & Energy Studies, Dehradun - 248 007 (Uttarakhand)

Email: anurag.jain@ddn.upes.ac.in , dr.anuragjain14@gmail.com

Mobile: (+91) -9729371188