

# Top-down Parsing

Dinesh Gopalani  
dgopalani.cse@mnit.ac.in

# Top-down Parsing

- It is a process of finding a leftmost derivation for an input string  $w$  starting with the start symbol  $S$ .
- At each step an appropriate production is applied to a non-terminal (say  $A$ ).
- After selecting  $A$ -production, next the process consists of matching the terminal symbols in the production body with the input string.

# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Input String w : c a d



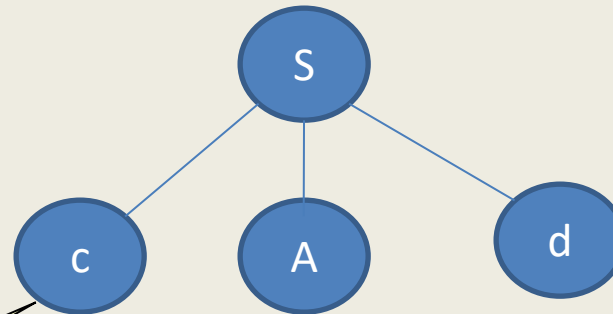
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is  
successful

Input String  $w$  : **c** a d  
                          ↑

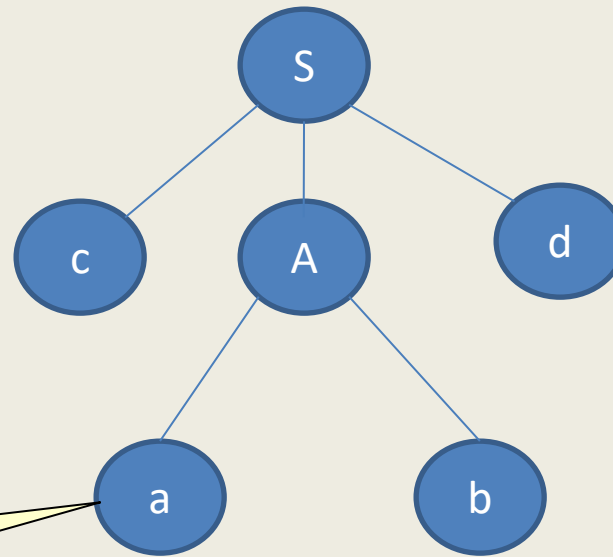
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is  
successful

Input String  $w$  : c a d  
                                  ↑

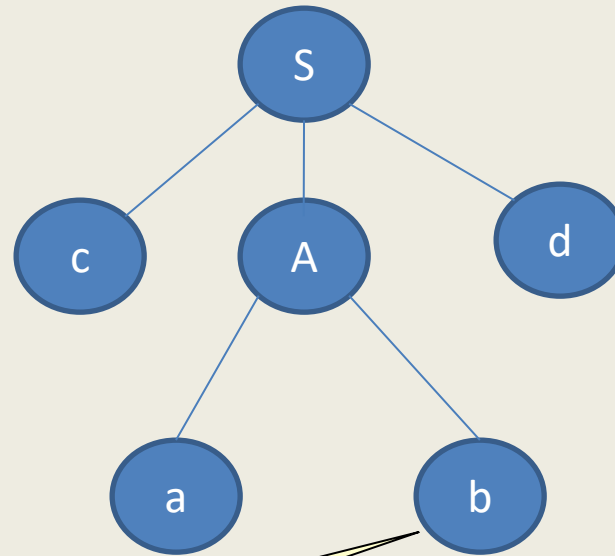
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is failed  
- Backtracking

Input String  $w$  : c a d



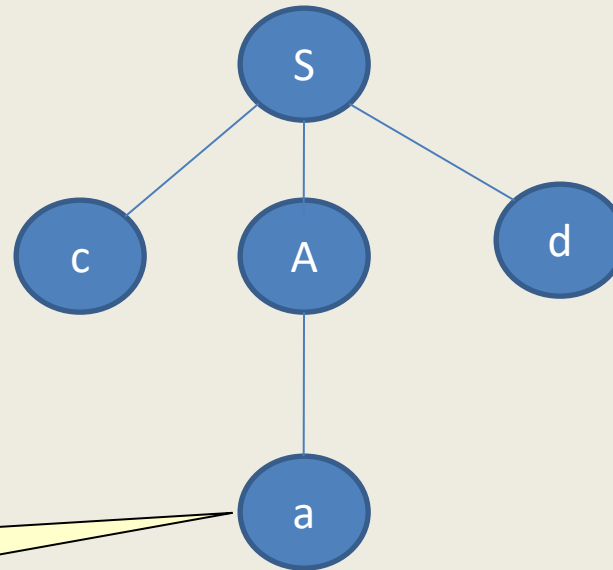
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is  
successful

Input String  $w$  : c a d



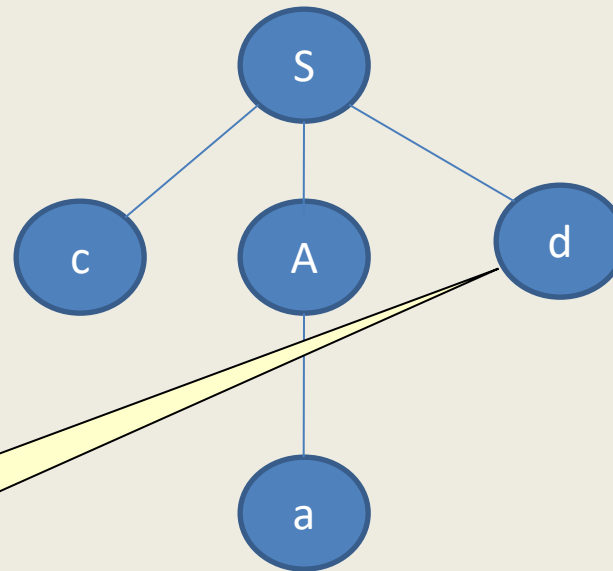
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is  
successful

Input String  $w$  : c a d  
↑



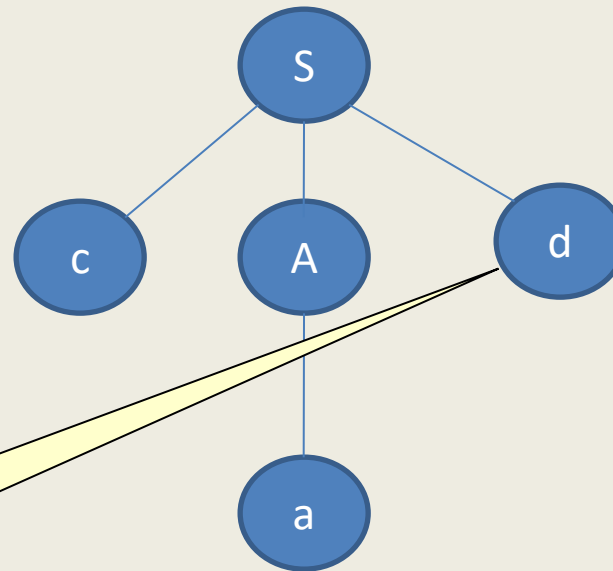
# Top-down Parsing

Approach:

Grammar G:

$S \rightarrow c A d$

$A \rightarrow a b \mid a$



Match is  
successful

Input String  $w$  : c a d

Leftmost Derivation :  $S \Rightarrow c A d \Rightarrow c a d$



# Recursive-Descent Parser

- It consists of set of procedures (functions) for each non-terminal.
- The process starts with the function for the start symbol.
- It may require backtracking – may require repeated scans over the input.
- The backtracking shall be avoided as it will make parsing inefficient.

# Recursive-Descent Parser

Example:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$

# Recursive-Descent Parser

Example:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$

```
int D()  
{  
    T();  
    L();  
    return 1;  
}
```

# Recursive-Descent Parser

Example:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$

```
int D()
{
    T();
    L();
    return 1;
}
```

```
int T()
{
    if (input=="int"){
        Advance();
        return 1; }
    else if (input=="char"){
        Advance();
        return 1; }
    else
        return 0;
}
```

# Recursive-Descent Parser

Example:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$

```
int D()
{
    T();
    L();
    return 1;
}
```

```
int T()
{
    if (input=="int"){
        Advance();
        return 1; }
    else if (input=="char"){
        Advance();
        return 1; }
    else
        return 0;
}
```

```
int L()
{
    isave=ip;
    if (input=="id"){
        Advance();
        if (input==","){
            Advance();
            L();
            return 1; }
        ip=isave;
    }
    if (input=="id"){
        Advance();
        if (input==";"){
            Advance();
            return 1; }
    }
    else
        return 0;
}
```

# Problems with Top-down Parsing

## Backtracking:

- If a sequence of wrong productions are used and subsequently a mismatch is discovered then the parser has to backtrack and try other set of productions.
- One major reason is common prefixes present in the input grammar.
- Example:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id} , L \mid \text{id} ;$

# Problems with Top-down Parsing

## Backtracking:

- To fix the common prefixes problem, the process of left factoring may be used.
- Left Factoring :

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

$$A \rightarrow \alpha A'$$

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



# Problems with Top-down Parsing

## Backtracking:

- To fix the common prefixes problem, the process of left factoring may be used.
- Left Factoring :

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$

$L \rightarrow \text{id} L'$

$L' \rightarrow , L \mid ;$

# Problems with Top-down Parsing

## Left-Recursion:

- A grammar  $G$  is said to be left-recursive if it has a non-terminal  $A$  such that there is a derivation  $A \xRightarrow{+} A \alpha$  for some  $\alpha$ .
- A left-recursive grammar can cause a top-down parser to go into an infinite loop – while trying to expand  $A$ , it may eventually find again trying to expand  $A$  without consuming any input.
- Therefore in order to use top-down parsing, all left-recursions must be eliminated from the input grammar.

# Elimination of Left-Recursion

## Immediate Left-Recursion:

- If there are left-recursive pair of productions

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

(here  $\beta$  does not begin with A)

- Then left-recursion can be eliminated by replacing this pair of productions with the following

$$A \rightarrow \beta A'$$

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

# Elimination of Left-Recursion

## Immediate Left-Recursion: (General Case)

- If there are left-recursion among A-productions as

$$A \rightarrow A \alpha_1 \mid A \alpha_2 \mid \dots \mid A \alpha_m \mid \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$$

(here no  $\beta_i$  begins with A)

- Then left-recursion can be eliminated by replacing as

$$A \rightarrow \beta_1 A' \mid \beta_2 A' \mid \dots \mid \beta_n A'$$

$$A' \rightarrow \alpha_1 A' \mid \alpha_2 A' \mid \dots \mid \alpha_m A' \mid \wedge$$

# Elimination of Left-Recursion

## Left-Recursion:

1. Arrange all non-terminals in some order  $A_1, A_2, \dots, A_n$ .
2. Eliminate immediate left-recursion among  $A_1$ -productions.
3. for  $i = 2$  to  $n$  do  
begin  
for  $j = 1$  to  $i-1$   
replace each production of the form  $A_i \rightarrow A_j \gamma$  by the  
productions  $A_i \rightarrow \delta_1 \gamma \mid \delta_2 \gamma \mid \dots \mid \delta_k \gamma$ , where  
 $A_j \rightarrow \delta_1 \mid \delta_2 \mid \dots \mid \delta_k$  are current  $A_j$ -productions.  
Eliminate immediate left-recursion among  $A_i$ -productions.  
end

# Elimination of Left-Recursion

## Left-Recursion:

1. Arrange all non-terminals in some order  $A_1, A_2, \dots, A_n$ .
2. Eliminate immediate left-recursion among  $A_1$ -productions.
3. for  $i = 2$  to  $n$  do

begin

for  $j = 1$  to  $i-1$

replace each production of the form  $A_i \rightarrow A_j \gamma$  by the productions  $A_i \rightarrow \delta_1 \gamma \mid \delta_2 \gamma \mid \dots \mid \delta_k \gamma$ , where

$A_j \rightarrow \delta_1 \mid \delta_2 \mid \dots \mid \delta_k$  are current  $A_j$ -productions.

Eliminate immediate left-recursion among  $A_i$ -productions.

end

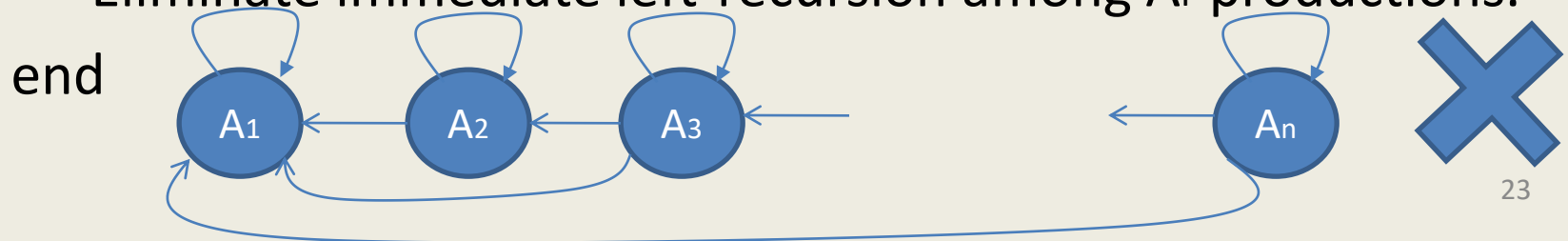


# Elimination of Left-Recursion

## Left-Recursion:

1. Arrange all non-terminals in some order  $A_1, A_2, \dots, A_n$ .
2. Eliminate immediate left-recursion among  $A_1$ -productions.
3. for  $i = 2$  to  $n$  do  
begin  
for  $j = 1$  to  $i-1$   
replace each production of the form  $A_i \rightarrow A_j \gamma$  by the  
productions  $A_i \rightarrow \delta_1 \gamma \mid \delta_2 \gamma \mid \dots \mid \delta_k \gamma$ , where  
 $A_j \rightarrow \delta_1 \mid \delta_2 \mid \dots \mid \delta_k$  are current  $A_j$ -productions.

Eliminate immediate left-recursion among  $A_i$ -productions.



# Elimination of Left-Recursion

## Example:

$$S \rightarrow A a \mid b$$

$$A \rightarrow A c \mid S d \mid e$$

1. Rename  $S$  as  $A_1$  and  $A$  as  $A_2$ .

$$A_1 \rightarrow A_2 a \mid b$$

$$A_2 \rightarrow A_2 c \mid A_1 d \mid e$$

2. Eliminate immediate left-recursion among  $A_1$ -productions –  
Not Applicable here.

3. for ( $i=2, j=1$ )

Replace  $A_2 \rightarrow A_1 d$  by

$$A_2 \rightarrow A_2 a d \mid b d$$

Now  $A_2$  productions are  $A_2 \rightarrow A_2 a d \mid b d \mid A_2 c \mid e$



# Elimination of Left-Recursion

Example:

3. (Contd.) for (i=2)

Eliminate immediate left recursions among  $A_i$  productions:

$A_2 \rightarrow A_2 a d \mid b d \mid A_2 c \mid e$

$A_2 \rightarrow b d A_2' \mid e A_2'$

$A_2' \rightarrow a d A_2' \mid c A_2' \mid \wedge$

So, final grammar is

$A_1 \rightarrow A_2 a \mid b$

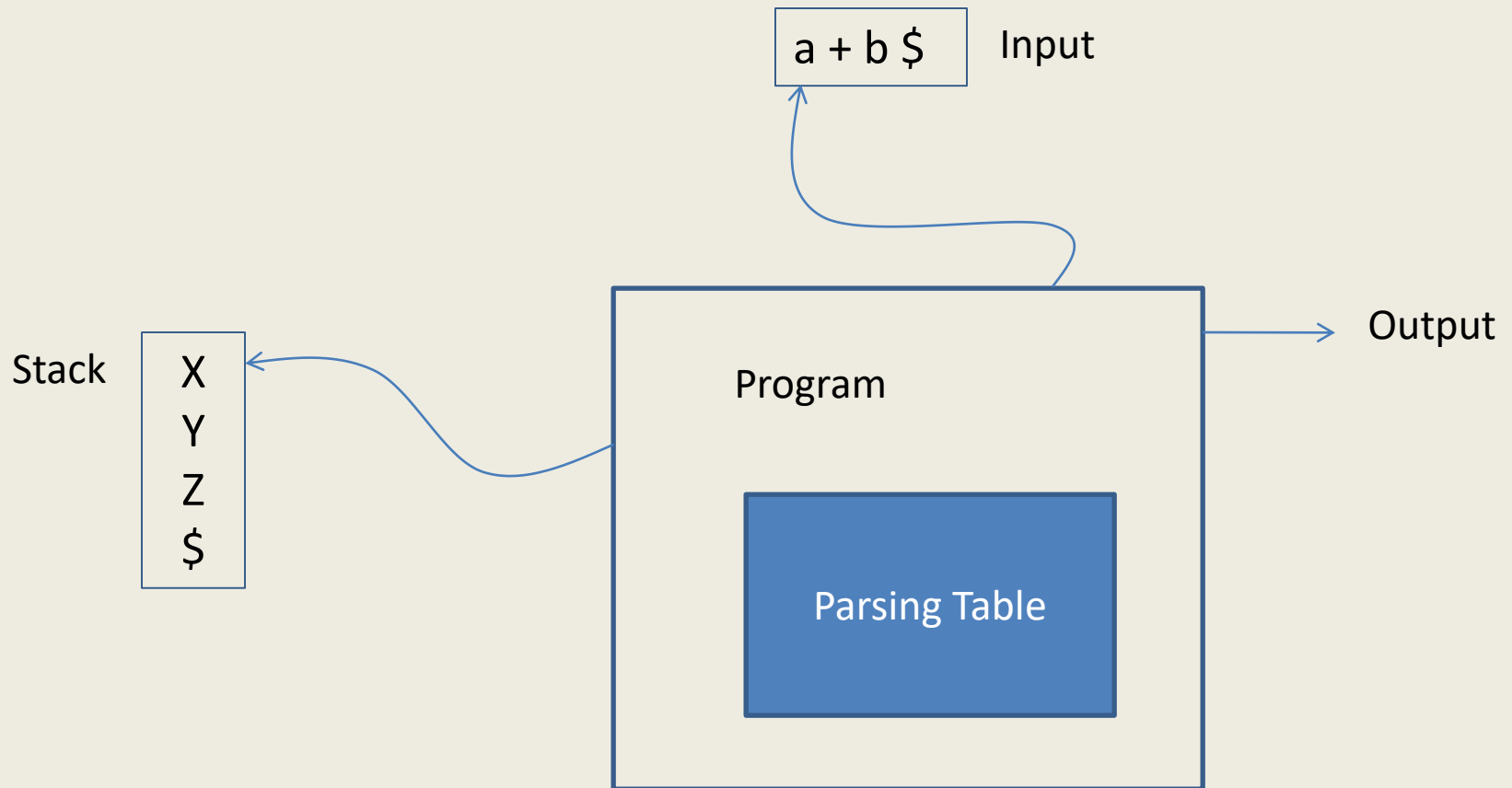
$A_2 \rightarrow b d A_2' \mid e A_2'$

$A_2' \rightarrow a d A_2' \mid c A_2' \mid \wedge$

# LL(1) Parser

- Based on Top-down approach.
- Significance of the name:
  - L - Left-to-right scanning of input
  - L - Leftmost derivation
  - (1) - Number of input symbol(s) to make parsing decision
- It avoids backtracking.
- It uses Parsing Table to decide what production is to be applied at each step.

# LL(1) Parser



# Working of LL(1) Parser

The parser determines  $X$ , the symbol on TOS and  $a$ , the current input symbol :

1. If  $X = a = \$$ , the parser halts and announces successful completion of parsing.
2. If  $X = a \neq \$$ , the parser pops  $X$  off the Stack and input pointer is advanced.
3. If  $X$  is a non-terminal then the parser checks parsing table  $M[X, a]$  which is either an  $X$ -production or an error. If entry is a production ( $X \rightarrow UVW$ ), the parser replaces  $X$  on TOS by  $WVU$  (with  $U$  on top).

# Construction of LL(1) Parsing Table

First and Follow functions:

$\text{First}(\alpha)$  – It is the set of terminals that appear in the beginning of the strings derived from  $\alpha$ .

If  $\alpha \xRightarrow{*} a \beta$ , then  $a$  is in  $\text{First}(\alpha)$ .

If  $\alpha \xRightarrow{*} \wedge$ , then  $\wedge$  is also in  $\text{First}(\alpha)$ .

$\text{Follow}(A)$  – It is the set of terminals that can appear immediately on the right of  $A$  in some sentential form.

If  $S \xRightarrow{*} \alpha A a \beta$ , for some  $\alpha$  and  $\beta$ , then  $a$  is in  $\text{Follow}(A)$ .

If  $S \xRightarrow{*} \alpha A$ , then  $\$$  is in  $\text{Follow}(A)$ .

# First Function

First(X):

1. If X is a terminal (say a)

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

2. If X is a non-terminal (say A)

(i) If  $A \rightarrow a \alpha$  is a production then add a to First(A).

For  $A \rightarrow \wedge$ , add  $\wedge$  to First(A).

(ii) If  $A \rightarrow Y_1 Y_2 \dots Y_k$  is a production, then

- Add every non- $\wedge$  symbol in First( $Y_1$ ) to First(A)
- If  $\wedge$  is in First( $Y_1$ ), then add every non- $\wedge$  symbol in First( $Y_2$ ) to First(A)
- If  $\wedge$  is in both First( $Y_1$ ) and First( $Y_2$ ), then add every non- $\wedge$  symbol in First( $Y_3$ ) to First(A)
- ... (The process goes on)
- Finally, if  $\wedge$  is in First( $Y_i$ ) for all  $i=1,2, \dots, k$  then add  $\wedge$  to First(A).

# First Function

$\text{First}(\alpha) : \alpha = X_1 X_2 \dots X_k$

- Add every non- $\wedge$  symbol in  $\text{First}(X_1)$  to  $\text{First}(\alpha)$
- If  $\wedge$  is in  $\text{First}(X_1)$ , then add every non- $\wedge$  symbol in  $\text{First}(X_2)$  to  $\text{First}(\alpha)$
- If  $\wedge$  is in both  $\text{First}(X_1)$  and  $\text{First}(X_2)$ , then add every non- $\wedge$  symbol in  $\text{First}(X_3)$  to  $\text{First}(\alpha)$
- ... (The process goes on)
- Finally, if  $\wedge$  is in  $\text{First}(X_i)$  for all  $i=1,2, \dots, k$  then add  $\wedge$  to  $\text{First}(\alpha)$ .

# Follow Function

Follow(A) :

- If A is a start symbol then add \$ in Follow(A)
- If there is a production  $B \rightarrow \alpha A \beta$  ( $\beta \neq \wedge$ ), then add every non- $\wedge$  symbol in  $\text{First}(\beta)$  to Follow(A)
- If there is a production  $B \rightarrow \alpha A$ , or a production  $B \rightarrow \alpha A \beta$  where  $\text{First}(\beta)$  contains  $\wedge$ , then add every symbol in Follow(B) to Follow(A)



# Follow Function

Follow(A) :

- If A is a start symbol then add \$ in Follow(A)
- If there is a production  $B \rightarrow \alpha A \beta$  ( $\beta \neq \wedge$ ), then add every non- $\wedge$  symbol in  $\text{First}(\beta)$  to Follow(A)
- If there is a production  $B \rightarrow \alpha A$ , or a production  $B \rightarrow \alpha A \beta$  where  $\text{First}(\beta)$  contains  $\wedge$ , then add every symbol in Follow(B) to Follow(A)

If  $S \xRightarrow{*} \gamma B \alpha \delta \Rightarrow \gamma \alpha A \alpha \delta$

then  $\alpha$  which is following B, will also follow A as well.

# First and Follow Functions

Example:

$$S \rightarrow A B \mid d S$$
$$A \rightarrow a A b \mid ^$$
$$B \rightarrow b B A \mid c$$
$$\text{First}(S) = \{ \}$$
$$\text{First}(A) = \{a, ^\}$$
$$\text{First}(B) = \{b, c\}$$

# First and Follow Functions

Example:

$$S \rightarrow A B \mid d S$$
$$A \rightarrow a A b \mid ^$$
$$B \rightarrow b B A \mid c$$
$$\text{First}(S) = \{a, b, c, d\}$$
$$\text{First}(A) = \{a, ^\}$$
$$\text{First}(B) = \{b, c\}$$

# First and Follow Functions

Example:

$$S \rightarrow A B \mid d S$$
$$A \rightarrow a A b \mid ^$$
$$B \rightarrow b B A \mid c$$
$$\text{First}(S) = \{a, b, c, d\}$$
$$\text{First}(A) = \{a, ^\}$$
$$\text{First}(B) = \{b, c\}$$
$$\text{Follow}(S) = \{\$ \}$$
$$\text{Follow}(A) = \{b, c, a, \$ \}$$
$$\text{Follow}(B) = \{\$, a\}$$

# Construction of LL(1) Parsing Table

1. For each production  $A \rightarrow \alpha$  of the grammar, do steps 2 and 3.
2. For each terminal  $a$  in  $\text{First}(\alpha)$ , add  $A \rightarrow \alpha$  to  $M[A,a]$ .
3. If  $\wedge$  is in  $\text{First}(\alpha)$ , add  $A \rightarrow \alpha$  to  $M[A,b]$  for each terminal  $b$  (including  $\$$ ) in  $\text{Follow}(A)$ .
4. Make all remaining entries of  $M$  as Errors.

# Construction of LL(1) Parsing Table

Example:

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

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

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

Eliminate left-recursion:

$$E \rightarrow T E'$$

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

$$T \rightarrow F T'$$

$$T' \rightarrow * F T' \mid \wedge$$

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

# Construction of LL(1) Parsing Table

Example:

$$E \rightarrow T E'$$

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

$$T \rightarrow F T'$$

$$T' \rightarrow * F T' \mid \wedge$$

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

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

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

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

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

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

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

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

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

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

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

# Construction of LL(1) Parsing Table

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

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

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow \text{id}$			$F \rightarrow ( E )$		

First (E) = { (, id }

First (E') = { +,  $\wedge$  }

First (T) = { (, id }

First (T') = { \*,  $\wedge$  }

First (F) = { (, id }

Follow(E) = { \$, ) }

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

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

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

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



# Working of LL(1) Parser

The parser determines  $X$ , the symbol on TOS and  $a$ , the current input symbol :

1. If  $X = a = \$$ , the parser halts and announces successful completion of parsing.
2. If  $X = a \neq \$$ , the parser pops  $X$  off the Stack and input pointer is advanced.
3. If  $X$  is a non-terminal then the parser checks parsing table  $M[X, a]$  which is either an  $X$ -production or an error. If entry is a production ( $X \rightarrow UVW$ ), the parser replaces  $X$  on TOS by  $WVU$  (with  $U$  on top).

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

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

Stack

\$E

Input

id+id\*id\$

Action

$E \rightarrow T E'$

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow \text{id}$			$F \rightarrow ( E )$		

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

$F \rightarrow ( E ) \mid id$

Stack

Input

\$E

id+id\*id\$

\$E'T

id+id\*id\$

Action

$E \rightarrow T E'$

$T \rightarrow F T'$

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow id$			$F \rightarrow ( E )$		

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

$F \rightarrow ( E ) \mid id$

Stack

Input

\$E      id+id\*id\$

\$E'T      id+id\*id\$

\$E'T'F      id+id\*id\$

Action

$E \rightarrow T E'$

$T \rightarrow F T'$

$F \rightarrow id$

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow id$			$F \rightarrow ( E )$		

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

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

Stack

Input

\$E id+id\*id\$

\$E'T id+id\*id\$

\$E'T'F id+id\*id\$

\$E'T'id id+id\*id\$

Action

$E \rightarrow T E'$

$T \rightarrow F T'$

$F \rightarrow \text{id}$

Match

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow \text{id}$			$F \rightarrow ( E )$		

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

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

Stack

Input

\$E id+id\*id\$

\$E'T id+id\*id\$

\$E'T'F id+id\*id\$

\$E'T'id id+id\*id\$

\$E'T' +id\*id\$

Action

$E \rightarrow T E'$

$T \rightarrow F T'$

$F \rightarrow \text{id}$

Match

$T' \rightarrow \wedge$

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow \text{id}$			$F \rightarrow ( E )$		

# Working of LL(1) Parser

Example:

$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

$F \rightarrow ( E ) \mid id$

Stack

Input

\$E id+id\*id\$

\$E'T id+id\*id\$

\$E'T'F id+id\*id\$

\$E'T'id id+id\*id\$

\$E'T' +id\*id\$

\$E' +id\*id\$

\$E'T+ +id\*id\$

\$E'T id\*id\$

\$E'T'F id\*id\$

\$E'T'id id\*id\$

\$E'T' \*id\$

\$E'T'F\* \*id\$

\$E'T'F id\$

\$E'T'id id\$

\$E'T' \$

\$E' \$

\$ \$

Action

$E \rightarrow T E'$

$T \rightarrow F T'$

$F \rightarrow id$

Match

$T' \rightarrow \wedge$

$E' \rightarrow + T E'$

Match

$T \rightarrow F T'$

$F \rightarrow id$

Match

$T' \rightarrow * F T'$

Match

$F \rightarrow id$

Match

$T' \rightarrow \wedge$

$E' \rightarrow \wedge$

Accepted

	id	+	*	(	)	\$
E	$E \rightarrow T E'$			$E \rightarrow T E'$		
E'		$E' \rightarrow + T E'$			$E' \rightarrow \wedge$	$E' \rightarrow \wedge$
T	$T \rightarrow F T'$			$T \rightarrow F T'$		
T'		$T' \rightarrow \wedge$	$T' \rightarrow * F T'$		$T' \rightarrow \wedge$	$T' \rightarrow \wedge$
F	$F \rightarrow id$			$F \rightarrow ( E )$		

# LL(1) Grammar

- A grammar whose LL(1) Parsing table has no multiply-defined entries is said to be LL(1).
- A grammar is NOT LL(1), if
  - Ambiguous
  - Left-recursive
  - Common Prefixes



# LL(1) Grammar

A grammar  $G$  is LL(1) if and only if whenever  $A \rightarrow \alpha \mid \beta$  are two distinct productions of  $G$ , the following conditions must hold:

1. For no terminal  $a$  do both  $\alpha$  and  $\beta$  derive strings beginning with  $a$ .
2. At most one of  $\alpha$  and  $\beta$  can derive the empty string.
3. If  $\beta \xRightarrow{*} \wedge$ , then  $\alpha$  must not derive any string beginning with a terminal in  $\text{Follow}(A)$ .

# LL(1) Grammar

A grammar  $G$  is LL(1) if and only if whenever  $A \rightarrow \alpha \mid \beta$  are two distinct productions of  $G$ , the following conditions must hold:

1. For no terminal  $a$  do both  $\alpha$  and  $\beta$  derive strings beginning with  $a$ .  
 $\text{First}(\alpha) \cap \text{First}(\beta) = \phi$
2. At most one of  $\alpha$  and  $\beta$  can derive the empty string.  
Either  $\text{First}(\alpha)$  or  $\text{First}(\beta)$  contains  $\wedge$ , and not both
3. If  $\beta \xRightarrow{*} \wedge$ , then  $\alpha$  must not derive any string beginning with a terminal in  $\text{Follow}(A)$ .  
 $\text{First}(\alpha) \cap \text{Follow}(A) = \phi$

# LL(1) Grammar

- Examples:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$



# LL(1) Grammar

- Examples:

$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id}, L \mid \text{id};$



$D \rightarrow T L$

$T \rightarrow \text{int} \mid \text{char}$

$L \rightarrow \text{id } L'$

$L' \rightarrow , L \mid ;$



# LL(1) Grammar

- Examples:

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


# LL(1) Grammar

- Examples:

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

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



$E \rightarrow T E'$

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

$T \rightarrow F T'$

$T' \rightarrow * F T' \mid \wedge$

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



# LL(1) Grammar

- Examples:

$S \rightarrow B b \mid C c$

$B \rightarrow a$

$C \rightarrow a$



# LL(1) Grammar

- Examples:

$S \rightarrow B b \mid C c$

$B \rightarrow a$

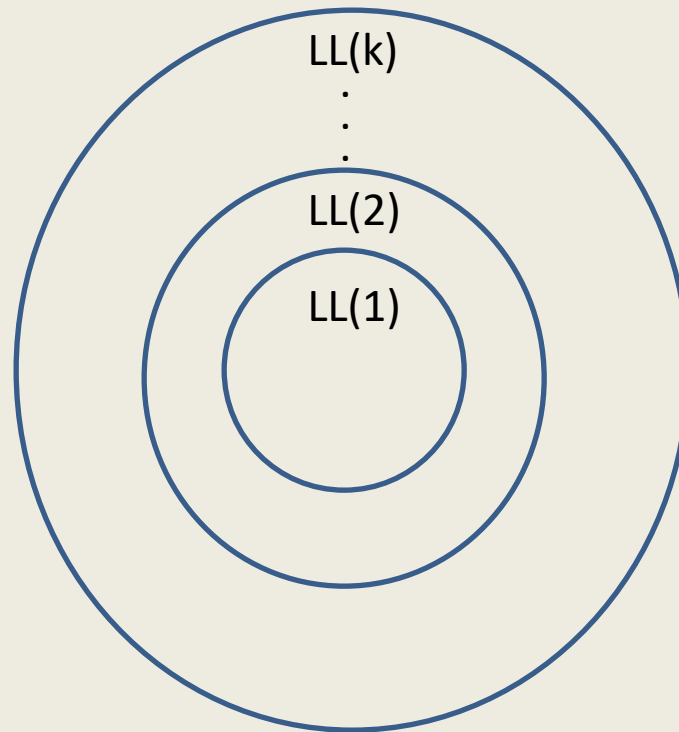
$C \rightarrow a$





# LL(k) Parser and LL(k) Grammar

LL(k) Grammar for  $k \geq 2$  is superset of LL(1) Grammar.



# LL(2) Grammar and LL(2) Parser

Example:

$S \rightarrow B b \mid C c$

$B \rightarrow a$

$C \rightarrow a$

	a\$	b\$	c\$	aa	ab	ac	ba	bb	bc	ca	cb	cc
S					$S \rightarrow B b$	$S \rightarrow C c$						
B					$B \rightarrow a$							
C						$C \rightarrow a$						

# LL(k) Parser and LL(k) Grammar

- LL(k) Parsers for  $k \geq 2$  are complex and their Parsing tables have many entries.
- Also,
$$L(\text{LL}(k) \text{ Grammar}) = L(\text{LL}(1) \text{ Grammar})$$
- Therefore, LL(k) grammar can be converted into LL(1) grammar.
- For all these reasons LL(1) is preferred.