

Top-Down Parsing

Lecture 7
Section 4.4

ROKAN UDDIN FARUQUI

Associate Professor
Dept of Computer Science and Engineering
University of Chittagong, Bangladesh
Email: *rokan@cu.ac.bd*

1 Left Factoring

$E \rightarrow TE'$
 $E' \rightarrow +TE' \mid \text{epsilon}$

2 Nullability

$E \rightarrow E + T \mid T$

3 The FIRST Function

4 The FOLLOW Function

$S \rightarrow AA'$
 $A' \rightarrow b \mid c$

5 The Parse Table Entries

6 The Parse Table

7 Predictive Parsing

8 Assignment



Outline

- 1 Left Factoring
- 2 Nullability
- 3 The FIRST Function
- 4 The FOLLOW Function
- 5 The Parse Table Entries
- 6 The Parse Table
- 7 Predictive Parsing
- 8 Assignment



Left Factoring

- A problem occurs when two productions for the same nonterminal begin with the same token.
- We cannot decide which production to use.
- This is not necessarily a problem since we could process the part they have in common, then make a decision based on what follows.



Left Factoring

- Consider the grammar

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

- We use **left factorization** to transform it into the form

$$A \rightarrow \alpha A'$$

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

- Now we can apply the productions immediately and unambiguously.



Example

Example (Left Factoring)

- In the earlier example, we had the productions

$$C \rightarrow \text{id} == \text{num} \mid \text{id} != \text{num} \mid \text{id} < \text{num}$$

- To perform left factoring, introduce a nonterminal C' :

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

$$C' \rightarrow == \text{num} \mid != \text{num} \mid < \text{num}$$

Example

Example (Left Factoring)

- Consider the grammar of **if** statements.

$$\begin{aligned} S \rightarrow & \text{ if } C \text{ then } S \text{ else } S \\ | & \text{ if } C \text{ then } S \end{aligned}$$

- We rewrite it as

$$\begin{aligned} S \rightarrow & \text{ if } C \text{ then } S \ S' \\ S' \rightarrow & \text{ else } S \mid \varepsilon \end{aligned}$$

Outline

① Left Factoring

② Nullability

③ The FIRST Function

④ The FOLLOW Function

⑤ The Parse Table Entries

⑥ The Parse Table

⑦ Predictive Parsing

⑧ Assignment



Table-Driven LL Parsing

- To build the parsing table, we need three concepts:
 - Nullability
 - The FIRST function
 - The FOLLOW function



Nullability

$$\begin{array}{l} E' \rightarrow +TE' \\ E' \rightarrow e \end{array}$$

Definition (Nullable)

A nonterminal A is **nullable** if

$$A \xrightarrow{*} \varepsilon.$$



Nullability

$B \Rightarrow e$
 $C \Rightarrow B \Rightarrow e$
 $C \Rightarrow^* e$
 $A \Rightarrow BC \Rightarrow C \Rightarrow e$
 $A \Rightarrow^* e$

Example (Nullability)

- In the following grammar, A , B , and C are nullable.

$$\begin{aligned} A &\rightarrow B\ C \\ B &\rightarrow A \mid a\ C \mid \varepsilon \\ C &\rightarrow a\ B \mid B\ A \mid B \end{aligned}$$


Nullability

- In other words, A is nullable if there is a production

$$A \rightarrow \varepsilon,$$

or there is a production

$$A \rightarrow B_1 B_2 \dots B_n,$$

where B_1, B_2, \dots, B_n are nullable.



Example

Example (Nullability)

$$E \rightarrow E + E - \text{ambiguous}$$

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

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

- In the grammar

$$F \rightarrow (E) \mid \text{id} - \text{unambiguous but left recursive}$$

$$E \rightarrow T E'$$

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

$$T \rightarrow F T'$$

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

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

$$E \rightarrow T E' \mid$$

- E' and T' are nullable.
- E , T , and F are not nullable.

Example

Example (Nullability)

Nonterminal	Nullable
E	No
E'	Yes
T	No
T'	Yes
F	No



Outline

- ① Left Factoring
- ② Nullability
- ③ The FIRST Function
- ④ The FOLLOW Function
- ⑤ The Parse Table Entries
- ⑥ The Parse Table
- ⑦ Predictive Parsing
- ⑧ Assignment



The FIRST Function

id1 + (id2 * id3)



Definition (The FIRST Function)

For any string of grammar symbols α , $\text{FIRST}(\alpha)$ is the set of all tokens that may appear as the first symbol in a replacement string of α .



Computing FIRST(X)

Computing FIRST(X)

- Let X be a grammar symbol. FIRST(X) is computed as follows.
 - If X is a token, then $\text{FIRST}(X) = \{X\}$.
 - If X is a nonterminal, then for every production

$$X \rightarrow Y_1 \ Y_2 \ \dots \ Y_n,$$

- FIRST(Y_1) \subseteq FIRST(X).
- Furthermore, if Y_1, Y_2, \dots, Y_k are nullable, then

$$\text{FIRST}(Y_{k+1}) \subseteq \text{FIRST}(X).$$

FIRST

- We are concerned with $\text{FIRST}(X)$ only for the nonterminals of the grammar.
- $\text{FIRST}(X)$ for terminals is trivial.
- According to the definition, to determine $\text{FIRST}(A)$, we must inspect all productions that have A as their head.



Example

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

$$\text{FIRST}(E') = \{+, \text{e}\} - \text{nullable}$$

$$\text{FIRST}(T') = \{*, \text{e}\} - \text{nullable}$$

Example (FIRST)

- Let the grammar be

$$A \rightarrow BCD$$

$\text{FIRST}(A) = \text{FIRST}(D)$ if both B and C are nullable

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

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

$$\text{FOLLOW}(T') = \{ \}$$

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

$$\underline{E} \rightarrow \underline{T} E'$$

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

$$T \rightarrow \underline{E} T'$$

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

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



Example

Example (FIRST)

- Find $\text{FIRST}(E)$.
 - E occurs as the head in only one production

$$E \rightarrow T E'$$

- Therefore,

$$\text{FIRST}(T) \subseteq \text{FIRST}(E).$$

- Furthermore, T is not nullable.

- Therefore,

$$\text{FIRST}(E) = \text{FIRST}(T).$$

- We have yet to determine $\text{FIRST}(T)$.

Example

Example (FIRST)

- Find $\text{FIRST}(E)$.
 - E occurs as the head in only one production

$$E \rightarrow T E'$$

- Therefore,

$$\text{FIRST}(T) \subseteq \text{FIRST}(E).$$

- Furthermore, T is not nullable.

- Therefore,

$$\text{FIRST}(E) = \text{FIRST}(T).$$

- We have yet to determine $\text{FIRST}(T)$.

Example

Example (FIRST)

- Find $\text{FIRST}(E)$.
 - E occurs as the head in only one production

$$E \rightarrow T E'$$

- Therefore,

$$\text{FIRST}(T) \subseteq \text{FIRST}(E).$$

- Furthermore, T is not nullable.

- Therefore,

$$\text{FIRST}(E) = \text{FIRST}(T).$$

- We have yet to determine $\text{FIRST}(T)$.

Example

Example (FIRST)

- Find $\text{FIRST}(E)$.
 - E occurs as the head in only one production

$$E \rightarrow T E'$$

- Therefore,

$$\text{FIRST}(T) \subseteq \text{FIRST}(E).$$

- Furthermore, T is not nullable.

- Therefore,

$$\text{FIRST}(E) = \text{FIRST}(T).$$

- We have yet to determine $\text{FIRST}(T)$.

Example

Example (FIRST)

- Find $\text{FIRST}(E)$.
 - E occurs as the head in only one production

$$E \rightarrow T E'$$

- Therefore,

$$\text{FIRST}(T) \subseteq \text{FIRST}(E).$$

- Furthermore, T is not nullable.

- Therefore,

$$\text{FIRST}(E) = \text{FIRST}(T).$$

- We have yet to determine $\text{FIRST}(T)$.

Example

Example (FIRST)

- Find $\text{FIRST}(T)$.
 - T occurs as the head in only one production

$$T \rightarrow F T'$$

- Therefore,

$$\text{FIRST}(F) \subseteq \text{FIRST}(T).$$

- Furthermore, F is not nullable.

- Therefore,

$$\text{FIRST}(T) = \text{FIRST}(F).$$

- We have yet to determine $\text{FIRST}(F)$.

Example

Example (FIRST)

- Find FIRST(T).
 - T occurs as the head in only one production

$$T \rightarrow F T'$$

- Therefore,

$$\text{FIRST}(F) \subseteq \text{FIRST}(T).$$

- Furthermore, F is not nullable.

- Therefore,

$$\text{FIRST}(T) = \text{FIRST}(F).$$

- We have yet to determine $\text{FIRST}(F)$.

Example

Example (FIRST)

- Find FIRST(T).
 - T occurs as the head in only one production

$$T \rightarrow F T'$$

- Therefore,

$$\text{FIRST}(F) \subseteq \text{FIRST}(T).$$

- Furthermore, F is not nullable.

- Therefore,

$$\text{FIRST}(T) = \text{FIRST}(F).$$

- We have yet to determine FIRST(F).

Example

Example (FIRST)

- Find $\text{FIRST}(T)$.
 - T occurs as the head in only one production

$$T \rightarrow F T'$$

- Therefore,

$$\text{FIRST}(F) \subseteq \text{FIRST}(T).$$

- Furthermore, F is not nullable.

- Therefore,

$$\text{FIRST}(T) = \text{FIRST}(F).$$

- We have yet to determine $\text{FIRST}(F)$.

Example

Example (FIRST)

- Find $\text{FIRST}(T)$.
 - T occurs as the head in only one production

$$T \rightarrow F T'$$

- Therefore,

$$\text{FIRST}(F) \subseteq \text{FIRST}(T).$$

- Furthermore, F is not nullable.

- Therefore,

$$\text{FIRST}(T) = \text{FIRST}(F).$$

- We have yet to determine $\text{FIRST}(F)$.

Example

Example (FIRST)

- Find $\text{FIRST}(F)$.
 - $\text{FIRST}(F) = \{(), \text{id}\}$.
 - Therefore,
 - $\text{FIRST}(E) = \{(), \text{id}\}$.
 - $\text{FIRST}(T) = \{(), \text{id}\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(F)$.
 - $\text{FIRST}(F) = \{\langle, \text{id}\}$.
- Therefore,
 - $\text{FIRST}(E) = \{\langle, \text{id}\}$.
 - $\text{FIRST}(T) = \{\langle, \text{id}\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(F)$.
 - $\text{FIRST}(F) = \{\langle, \text{id}\}$.
- Therefore,
 - $\text{FIRST}(E) = \{\langle, \text{id}\}$.
 - $\text{FIRST}(T) = \{\langle, \text{id}\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(E')$.
 - $\text{FIRST}(E') = \{+\}$.
- Find $\text{FIRST}(T')$.
 - $\text{FIRST}(T') = \{*\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(E')$.
 - $\text{FIRST}(E') = \{+\}$.
- Find $\text{FIRST}(T')$.
 - $\text{FIRST}(T') = \{*\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(E')$.
 - $\text{FIRST}(E') = \{+\}$.
- Find $\text{FIRST}(T')$.
 - $\text{FIRST}(T') = \{*\}$.



Example

Example (FIRST)

- Find $\text{FIRST}(E')$.
 - $\text{FIRST}(E') = \{+\}$.
- Find $\text{FIRST}(T')$.
 - $\text{FIRST}(T') = \{*\}$.



Example

Example (FIRST)

Nonterminal	Nullable	FIRST
E	No	$\{ (, \text{id} \})$
E'	Yes	$\{ + \}$
T	No	$\{ (, \text{id} \})$
T'	Yes	$\{ * \}$
F	No	$\{ (, \text{id} \})$



Example

Example (FOLLOW)

- The grammar

$$R \rightarrow R \cup R | R R | R^* | (R) | \text{id}$$

generates all regular expressions over a set of identifiers. Find $\text{FIRST}(X)$ for each nonterminal X .



Outline

- 1 Left Factoring
- 2 Nullability
- 3 The FIRST Function
- 4 The FOLLOW Function
- 5 The Parse Table Entries
- 6 The Parse Table
- 7 Predictive Parsing
- 8 Assignment



The FOLLOW Function

Definition (The FOLLOW Function)

For any string of grammar symbols α , $\text{FOLLOW}(\alpha)$ is the set of all tokens that may follow α in a derivation.



Computing FOLLOW(A)

Computing FOLLOW(A)

- Let A be a nonterminal. FOLLOW(A) is computed as follows.

- If A is the start symbol, then $\$ \in \text{FOLLOW}(A)$.

- If $B \rightarrow \alpha A \beta$ is a production, then

$$\text{FIRST}(\beta) \subseteq \text{FOLLOW}(A).$$

- If $B \rightarrow \alpha A$ is a production, or $B \rightarrow \alpha A \underline{\beta}$ is a production and β is nullable, then

$$\text{FOLLOW}(B) \subseteq \text{FOLLOW}(A).$$



FOLLOW

- We are concerned about $\text{FOLLOW}(A)$ only for the nonterminals of the grammar.
- According to the definition, to determine $\text{FOLLOW}(A)$, we must inspect all productions that have A in their body.



Example

Example (FOLLOW)

- Let the grammar be

FOLLOW(T) = FIRST(E') U
FOLLOW(E)
 $= \{+, \$,)\}$

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

$\text{FOLLOW}(T) = \text{FIRST}(E') \cup$

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

$\text{FOLLOW}(E') = \text{FOLLOW}(E)$

$\text{FOLLOW}(T') = \text{FOLLOW}(T)$

$E \rightarrow T E'$

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

$T \rightarrow F T'$

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

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



Example

Example (FOLLOW)

- Find FOLLOW(E).
 - E is the start symbol, therefore

$$\$ \in \text{FOLLOW}(E).$$

- E occurs in the body of only one production.

$$F \rightarrow (E)$$

- Therefore

$$\text{FOLLOW}(E) = \{ \$,) \}.$$

Example

Example (FOLLOW)

- Find FOLLOW(E).
 - E is the start symbol, therefore

$$\$ \in \text{FOLLOW}(E).$$

- E occurs in the body of only one production.

$$F \rightarrow (E)$$

- Therefore

$$\text{FOLLOW}(E) = \{ \$,) \}.$$

Example

Example (FOLLOW)

- Find FOLLOW(E).
 - E is the start symbol, therefore

$$\$ \in \text{FOLLOW}(E).$$

- E occurs in the body of only one production.

$$F \rightarrow (E)$$

- Therefore

$$\text{FOLLOW}(E) = \{ \$,) \}.$$

Example

Example (FOLLOW)

- Find FOLLOW(E').
 - E' occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore,

$$\text{FOLLOW}(E') = \text{FOLLOW}(E) = \{\$,)\}.$$

Example

Example (FOLLOW)

- Find FOLLOW(E').
 - E' occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore,

$$\text{FOLLOW}(E') = \text{FOLLOW}(E) = \{\$\,,\,\}\,.$$



Example

Example (FOLLOW)

- Find FOLLOW(T).
 - T occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore, $\text{FOLLOW}(T) \supseteq \text{FIRST}(E') = \{+\}$.
- However, E' is nullable, therefore it also contains $\text{FOLLOW}(E) = \{\$\),)\}$ and $\text{FOLLOW}(E') = \{\$\),)\}$.
- Therefore,

$$\text{FOLLOW}(T) = \{+, \$,)\}.$$

Example

Example (FOLLOW)

- Find FOLLOW(T).
 - T occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore, $\text{FOLLOW}(T) \supseteq \text{FIRST}(E') = \{+\}$.
- However, E' is nullable, therefore it also contains $\text{FOLLOW}(E) = \{\$\),)\}$ and $\text{FOLLOW}(E') = \{\$\),)\}$.
- Therefore,

$$\text{FOLLOW}(T) = \{+, \$,)\}.$$

Example

Example (FOLLOW)

- Find FOLLOW(T).
 - T occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore, $\text{FOLLOW}(T) \supseteq \text{FIRST}(E') = \{+\}$.
- However, E' is nullable, therefore it also contains $\text{FOLLOW}(E) = \{\$\),)\}$ and $\text{FOLLOW}(E') = \{\$\),)\}$.
- Therefore,

$$\text{FOLLOW}(T) = \{+, \$,)\}.$$

Example

Example (FOLLOW)

- Find FOLLOW(T).
 - T occurs in the body of two productions.

$$\begin{aligned}E &\rightarrow T \ E' \\E' &\rightarrow + \ T \ E'\end{aligned}$$

- Therefore, $\text{FOLLOW}(T) \supseteq \text{FIRST}(E') = \{+\}$.
- However, E' is nullable, therefore it also contains $\text{FOLLOW}(E) = \{\$\),)\}$ and $\text{FOLLOW}(E') = \{\$\),)\}$.
- Therefore,

$$\text{FOLLOW}(T) = \{+, \$,)\}.$$

Example

Example (FOLLOW)

- Find FOLLOW(T').
 - T' occurs in the body of two productions.

$$\begin{aligned}T &\rightarrow F T' \\T' &\rightarrow * F T'\end{aligned}$$

- Therefore,

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



Example

Example (FOLLOW)

- Find FOLLOW(T').
 - T' occurs in the body of two productions.

$$\begin{aligned}T &\rightarrow F T' \\T' &\rightarrow * F T'\end{aligned}$$

- Therefore,

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



Example

Example (FOLLOW)

- Find FOLLOW(F).
 - F occurs in the body of two productions.

$$\begin{aligned} T &\rightarrow F \ T' \\ T' &\rightarrow * \ F \ T'. \end{aligned}$$

- Therefore,
 $\text{FOLLOW}(F) \supseteq \text{FIRST}(T') = \{*\}.$

- However, T' is nullable, therefore it also contains
 $\text{FOLLOW}(T) = \{+, \$,)\}$ and $\text{FOLLOW}(T') = \{+, \$,)\}.$
- Therefore,

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

Example

Example (FOLLOW)

- Find FOLLOW(F).
 - F occurs in the body of two productions.

$$\begin{aligned} T &\rightarrow F \ T' \\ T' &\rightarrow * \ F \ T'. \end{aligned}$$

- Therefore,

$$\text{FOLLOW}(F) \supseteq \text{FIRST}(T') = \{*\}.$$

- However, T' is nullable, therefore it also contains $\text{FOLLOW}(T) = \{+, \$,)\}$ and $\text{FOLLOW}(T') = \{+, \$,)\}$.
- Therefore,

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

Example

Example (FOLLOW)

- Find FOLLOW(F).
 - F occurs in the body of two productions.

$$\begin{aligned} T &\rightarrow F \ T' \\ T' &\rightarrow * \ F \ T'. \end{aligned}$$

- Therefore,
 $\text{FOLLOW}(F) \supseteq \text{FIRST}(T') = \{*\}.$
- However, T' is nullable, therefore it also contains
 $\text{FOLLOW}(T) = \{+, \$,)\}$ and $\text{FOLLOW}(T') = \{+, \$,)\}.$
- Therefore,
 $\text{FOLLOW}(F) = \{*, +, \$,)\}.$

Example

Example (FOLLOW)

- Find FOLLOW(F).
 - F occurs in the body of two productions.

$$\begin{aligned} T &\rightarrow F \ T' \\ T' &\rightarrow * \ F \ T'. \end{aligned}$$

- Therefore,

$$\text{FOLLOW}(F) \supseteq \text{FIRST}(T') = \{*\}.$$

- However, T' is nullable, therefore it also contains $\text{FOLLOW}(T) = \{+, \$,)\}$ and $\text{FOLLOW}(T') = \{+, \$,)\}$.
- Therefore,

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

Example

Example (FOLLOW)

Nonterminal	Nullable	FIRST	FOLLOW
E	No	$\{ (, \text{id} \})$	$\{ \$,) \}$
E'	Yes	$\{ + \}$	$\{ \$,) \}$
T	No	$\{ (, \text{id} \})$	$\{ \$,), + \}$
T'	Yes	$\{ * \}$	$\{ \$,), + \}$
F	No	$\{ (, \text{id} \})$	$\{ *, \$,), + \}$



Example

Example (FOLLOW)

- The grammar

$$\begin{aligned} S &\rightarrow (L) \mid \text{id} \\ L &\rightarrow L , S \mid S \end{aligned}$$

generates nested lists of identifiers. Find FOLLOW(X) for each nonterminal X .



Example

Example (FOLLOW)

- The grammar

$$R \rightarrow R \cup R | R R | R^* | (R) | \text{id}$$

generates all regular expressions over a set of identifiers. Find FOLLOW(X) for each nonterminal X .



Outline

- 1 Left Factoring
- 2 Nullability
- 3 The FIRST Function
- 4 The FOLLOW Function
- 5 The Parse Table Entries
- 6 The Parse Table
- 7 Predictive Parsing
- 8 Assignment



Nullability and FIRST for Strings

LL(1) Parsing / Table Driven LL Parsing

L = read input left to right

L = leftmost derivation

$M[E', +]$

5 x 6

- The parse table has
 - One row for each nonterminal, E, F, T, E', T'
 - One column for each terminal and \$. $id, +, *, (,), \$$
- \$ is the end-of-file marker.
- Each cell in the table represents a combination (A, a) of a nonterminal A and a terminal a .

$M[E, id]$
 $M[E', +]$

	id	+	*	()	\$
E						
E'						

Nullability and FIRST for Strings

$E \rightarrow TE' | E'$
 $A \rightarrow A_1A_2A_3\dots A_9$
 $\text{FIRST}(A) =$

- We need to extend the definitions of nullable and FIRST to *strings* of grammar symbols.
 - ε is nullable.
 - Let $\alpha = X_1X_2\dots X_n$ be a string of grammar symbols. Then α is nullable if each X_i is nullable.

Nullability and FIRST for Strings

- By definition, $\text{FIRST}(\epsilon) = \emptyset$.
- For the string $\alpha = X_1X_2\dots X_n$,
 - Let k be the largest integer for which X_1, X_2, \dots , and X_k are nullable.
 - Then

$$\text{FIRST}(\alpha) = \text{FIRST}(X_1) \cup \dots \cup \text{FIRST}(X_{k+1}).$$

- We will not need FOLLOW for strings.

Parse Table Construction : Algorithm

INPUT : Grammar G .

OUTPUT : Parsing table M

METHOD :

$$\begin{aligned}
 E &\rightarrow TE' \\
 E' &\rightarrow +TE' \mid \epsilon \\
 T &\rightarrow FT' \\
 T' &\rightarrow *FT' \mid \epsilon \\
 F &\rightarrow (E) \mid id
 \end{aligned}$$

$$\text{FIRST}(TE') = \text{FIRST}(T) = \{(, id\}$$

$$M[E, ()] = E \rightarrow TE'$$

$$M[E, id] = E \rightarrow TE'$$

- Each table entry is a production $A \rightarrow \alpha$.
- Rules for entering $A \rightarrow \alpha$ in the table:
 - For every rule $A \rightarrow \alpha$ and for every $a \in \text{FIRST}(\alpha)$, write $A \rightarrow \alpha$ in cell (A, a) .
 - For every rule $A \rightarrow \alpha$, where α is nullable, and for every $a \in \text{FOLLOW}(A)$, write $A \rightarrow \alpha$ in cell (A, a) .
- Write “error” in all cells that do not contain a production.

Parse Table Entries

- The interpretation of $A \rightarrow \alpha$ in cell (A, a) is that if A is on top of the stack and we encounter a in the input, then we should replace A by α on the stack.
- Push the symbols of α onto the stack in *reverse order*, from right to left, so that the first (leftmost) symbol is on the top of the stack.



Example

Input: G

Output: M

1. E --> TE'

 $A \rightarrow \text{alpha}$, FIRST(alpha) = FIRST(TE') = FIRST(T) = {(, id)}

Example (Parse Table)

M[E, ()] = E --> TE'

M[E, id] = E --> TE'

- Let the grammar be

4. T --> FT'

FIRST(FT') = FIRST(F) = {(, id)}

M[T, ()] = T --> FT'

M[T, id] = T --> FT'

2. E' --> +TE'

FIRST(+TE') = {+}

M[E', +] = E' --> +TE'

 $E \rightarrow T E'$ $E' \rightarrow + T E' \mid \epsilon$ $T \rightarrow F T'$ $T' \rightarrow * F T' \mid \epsilon$ $F \rightarrow (E) \mid \text{id}$ $F \rightarrow (E)$ $F \rightarrow \text{id}$ 3. E' --> e
FOLLOW(E') = {\$,) }

M[E', \$] = E' --> e

M[E',]) = E' --> e



Example

Example (Parse Table)

- Recall

Nonterminal	Nullable	FIRST	FOLLOW
E	No	$\{(), \text{id}\}$	$\{\$\,,)\}$
E'	Yes	$\{+\}$	$\{\$\,,)\}$
T	No	$\{(), \text{id}\}$	$\{\$\,,), +\}$
T'	Yes	$\{*\}$	$\{\$\,,), +\}$
F	No	$\{(), \text{id}\}$	$\{*, \$\,,), +\}$



Outline

- 1 Left Factoring
- 2 Nullability
- 3 The FIRST Function
- 4 The FOLLOW Function
- 5 The Parse Table Entries
- 6 The Parse Table
- 7 Predictive Parsing
- 8 Assignment



Example

Example (Parse Table)

- Consider the production $E \rightarrow T E'$.
 - $\text{FIRST}(T E') = \text{FIRST}(T) = \{\text{(), id}\}$.
 - Therefore, enter $E \rightarrow T E'$ in cells $(E, \text{()})$ and (E, id) .
 - $T E'$ is not nullable.



Example

Example (Parse Table)

- Consider the production $E' \rightarrow \varepsilon$.
- $\text{FIRST}(\varepsilon) = \emptyset$.
- ε is nullable and $\text{FOLLOW}(E') = \{\$\},)\}$.
- Therefore, enter $E' \rightarrow \varepsilon$ in cells $(E', \$)$ and $(E',))$.



Example

 $E \rightarrow E + E$

Example (Parse Table)

- Handle the other productions similarly.
- In which cells do we enter $E' \rightarrow + T E'$?
- In which cells do we enter $T \rightarrow F T'$?
- In which cells do we enter $T' \rightarrow * F T'$?
- In which cells do we enter $T' \rightarrow \epsilon$?
- In which cells do we enter $F \rightarrow (E)$?
- In which cells do we enter $F \rightarrow \text{id}$?



Example

$$\begin{array}{l} E' \rightarrow +TE' \\ \text{FIRST}(+TE') = \{ + \} \\ M[E', +] = E' \rightarrow TE' \end{array}$$

$$\begin{array}{l} E' \rightarrow +TE' \\ A \rightarrow \text{alpha} \\ \text{FIRST}(\text{alpha}) = \{ \text{alpha} \} \\ T \rightarrow FT \\ \text{FIRST}(FT) = \text{FIRST}(F) = \{ \text{id}, (\} \\ M[T, \text{id}] = T \rightarrow FT \\ M[T, (] = T \rightarrow FT' \\ \text{FIRST}(FT') = \{ * \} \end{array}$$

$$\begin{array}{l} M[E, ()] = E \rightarrow TE' \\ M[E, \text{id}] = E \rightarrow TE' \end{array}$$

Example (Parse Table)

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

$$E' \rightarrow \epsilon$$

Is it LL(1)?

$$\begin{array}{l} T' \rightarrow e \\ \text{FOLLOW}(T') = \{ +, (\} \\ \text{FOLLOW}(E') = \{ \$,) \} \\ M[E', \$] = E' \rightarrow \epsilon \\ M[E', ()] = E' \rightarrow \epsilon \end{array}$$



Outline

1 Left Factoring

2 Nullability

3 The FIRST Function

4 The FOLLOW Function

5 The Parse Table Entries

6 The Parse Table

7 Predictive Parsing

8 Assignment



Predictive Parsing

Definition (LL(1) Grammar)

An **LL(1) grammar** is a grammar whose predictive parse table does not contain any multiple entries.

- A multiple entry would indicate that we could not decide which production to apply.



Predictive Parsing

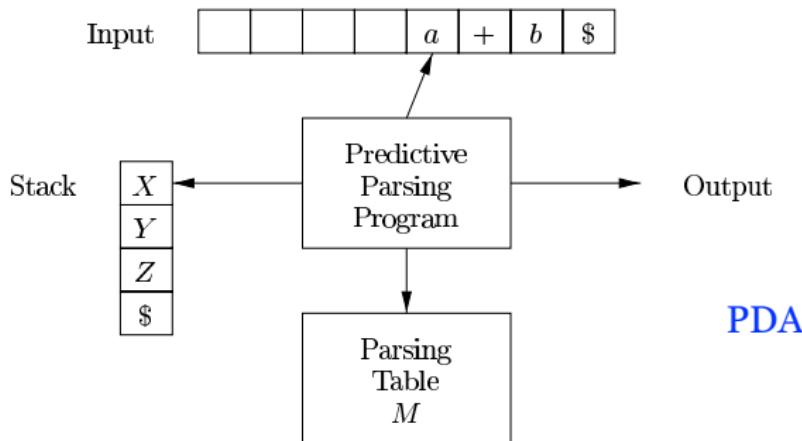
$$\begin{aligned} E &\rightarrow E + T \mid T \\ T &\rightarrow T * F \mid F \\ F &\rightarrow (E) \mid \text{id} \end{aligned}$$


Figure: Model of a table-driven predictive parsing



Predictive Parsing Algorithm

- The predictive parsing algorithm uses
 - The parse table,
 - An input buffer containing a sequence of tokens,
 - A stack of grammar symbols.



Predictive Parsing Algorithm



The Parsing Algorithm

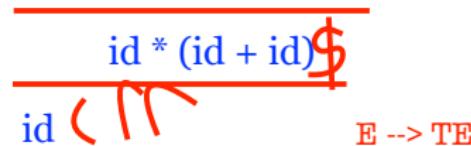
- Initialize the input buffer to the input followed by \$.
- Initialize the stack to \$ and S , with S on the top.

E\$

id+id * id\$



Predictive Parsing Algorithm

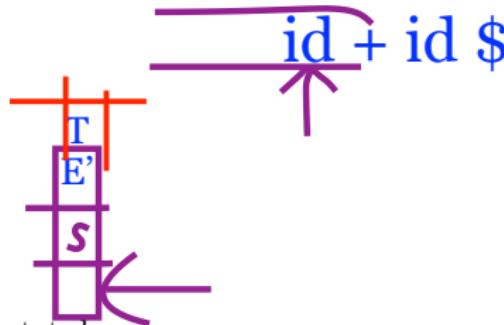


The Parsing Algorithm

- Consider the top stack symbol X .
- There are three possibilities.
 - ✓ X is a terminal.
 - ✓ X is a nonterminal.
 - ✓ X is \$.



Predictive Parsing Algorithm



The Parsing Algorithm

- If X is a terminal, then
 - If X matches the current token,
 - Pop X from the stack.
 - Advance to the next token.
 - Otherwise, it is an error.



Predictive Parsing Algorithm

The Parsing Algorithm



- If X is a nonterminal, then
 - Use X together with the current token to get the entry from the parse table.
 - If the entry is a production,
 - Pop X from the stack.
 - Push the symbols on the right-hand side of the production, from right to left, onto the stack.
 - Otherwise, it is an error.



Predictive Parsing Algorithm

The Parsing Algorithm

- If X is \$, then
 - If the current token is also \$,
 - Accept the input.
 - Otherwise, it is an error.



Example

M[E, ()]

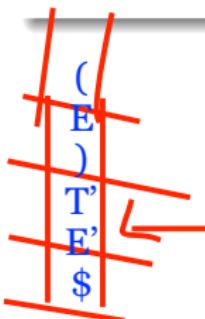
Stack

E\$
TE'\$
FT'E'\$
(id)T'E'\$
id)T'E'\$

Input
(id + id) * id \$
(id + id) * id \$
(id + id) * id \$
id + id) * id \$

Example (Predictive Parsing)

- Parse the string **(id + id)*id.**



Stack
\$E
\$E'T
\$E'T'F
\$E'T')E(
\$E'T')E

Input
(id + id) * id\$
(id+id) * id\$
(id+id) * id\$
(id+id) * id\$
id+id)*id\$

$X = ($
Input = (
 $M[T, ()] = E \rightarrow TE'$

Example M [E, C] = E → TE'

(id + id) * id \$

\$E'T**\$E**input: (
 stack: E

Example (Predictive Parsing)

Stack	Input
\$ E	(id + id) * id \$
\$ E' T	(id + id) * id \$
\$ E' T' F	(id + id) * id \$
\$ E' T') E ((id + id) * id \$
\$ E' T') E	id + id) * id \$
\$ E' T') E' T	id + id) * id \$
\$ E' T') E' T' F	id + id) * id \$
\$ E' T') E' T' id	id + id) * id \$
\$ E' T') E' T'	+ id) * id \$
\$ E' T') E'	+ id) * id \$
\$ E' T') E' T +	+ id) * id \$
\$ E' T') E' T	id) * id \$

E, C] = E → TE'

Example

Example (Predictive Parsing)

Stack	Input
\$ E	(id + id) * id \$
\$ E' T	(id + id) * id \$
\$ E' T' F	(id + id) * id \$
\$ E' T') E ((id + id) * id \$
\$ E' T') E	id + id) * id \$
\$ E' T') E' T	id + id) * id \$
\$ E' T') E' T' F	id + id) * id \$
\$ E' T') E' T' id	id + id) * id \$
\$ E' T') E' T'	+ id) * id \$
\$ E' T') E'	+ id) * id \$
\$ E' T') E' T +	+ id) * id \$
\$ E' T') E' T	id) * id \$

Example

$$M[F, () = F \rightarrow \underline{(E)}]$$

Example (Predictive Parsing)

Stack	Input
\$ E	(id + id) * id \$
\$ E' T	(id + id) * id \$
\$ E' T' F	(id + id) * id \$
\$ E' T') E ((id + id) * id \$
\$ E' T') E	id + id) * id \$
\$ E' T') E' T	id + id) * id \$
\$ E' T') E' T' F	id + id) * id \$
\$ E' T') E' T' id	id + id) * id \$
\$ E' T') E' T'	+ id) * id \$
\$ E' T') E'	+ id) * id \$
\$ E' T') E' T +	+ id) * id \$
\$ E' T') E' T	id) * id \$

Example

$$F \rightarrow (E)$$

Example (Predictive Parsing)

Stack	Input
\$ E	(id + id) * id \$
\$ E' T	(id + id) * id \$
\$ E' T' F	(id + id) * id \$
<u>\$ E' T') E (</u>	(id + id) * id \$
\$ E' T') E	id + id) * id \$
\$ E' T') E' T	id + id) * id \$
\$ E' T') E' T' F	id + id) * id \$
\$ E' T') E' T' id	id + id) * id \$
\$ E' T') E' T'	+ id) * id \$
\$ E' T') E'	+ id) * id \$
\$ E' T') E' T +	+ id) * id \$
\$ E' T') E' T	id) * id \$

Example

Example (Predictive Parsing)

Stack	Input
\$ E	(id + id) * id \$
\$ E' T	(id + id) * id \$
\$ E' T' F	(id + id) * id \$
\$ E' T') E ((id + id) * id \$
\$ E' T') E	(id + id) * id \$
\$ E' T') E T	id + id) * id \$
\$ E' T') E' T' F	id + id) * id \$
\$ E' T') E' T' id	id + id) * id \$
\$ E' T') E' T'	+ id) * id \$
\$ E' T') E'	+ id) * id \$
\$ E' T') E' T +	+ id) * id \$
\$ E' T') E' T	id) * id \$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E$	$(id + id) * id \$$
$\$ E' T$	$(id + id) * id \$$
$\$ E' T' F$	$(id + id) * id \$$
$\$ E' T') E ($	$(id + id) * id \$$
$\$ E' T') E$	$id + id) * id \$$
$\$ E' T') E' T$	$id + id) * id \$$
$\$ E' T') E' T' F$	$id + id) * id \$$
$\$ E' T') E' T' id$	$id + id) * id \$$
$\$ E' T') E' T'$	$+ id) * id \$$
$\$ E' T') E'$	$+ id) * id \$$
$\$ E' T') E' T +$	$+ id) * id \$$
$\$ E' T') E' T$	$id) * id \$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T'$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	id) * id \$
$\$ E' T') E' T' id$	id) * id \$
$\$ E' T') E' T'$) * id \$
$\$ E' T') E'$) * id \$
$\$ E' T')$) * id \$
$\$ E' T'$	* id \$
$\$ E' T' F *$	* id \$
$\$ E' T' F$	id \$
$\$ E' T' id$	id \$
$\$ E' T'$	\$
$\$ E'$	\$
$\$$	\$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	id) * id \$
$\$ E' T') E' T' id$	id) * id \$
$\$ E' T') E' T'$) * id \$
$\$ E' T') E'$) * id \$
$\$ E' T')$) * id \$
$\$ E' T'$	* id \$
$\$ E' T' F *$	* id \$
$\$ E' T' F$	id \$
$\$ E' T' id$	id \$
$\$ E' T'$	\$
$\$ E'$	\$
$\$$	\$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	id) * id \$
$\$ E' T') E' T' id$	id) * id \$
$\$ E' T') E' T'$) * id \$
$\$ E' T') E'$) * id \$
$\$ E' T')$) * id \$
$\$ E' T'$	* id \$
$\$ E' T' F *$	* id \$
$\$ E' T' F$	id \$
$\$ E' T' id$	id \$
$\$ E' T'$	\$
$\$ E'$	\$
$\$$	\$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$\text{id}) * \text{id} \$$
$\$ E' T') E' T' \text{id}$	$\text{id}) * \text{id} \$$
$\$ E' T') E' T'$	$) * \text{id} \$$
$\$ E' T') E'$	$) * \text{id} \$$
$\$ E' T')$	$) * \text{id} \$$
$\$ E' T'$	$* \text{id} \$$
$\$ E' T' F *$	$* \text{id} \$$
$\$ E' T' F$	$\text{id} \$$
$\$ E' T' \text{id}$	$\text{id} \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

Example (Predictive Parsing)

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Example

 $\$E$ $)id + id) * id$

Example (Predictive Parsing)

 $\$E$
 $M[E,)] =$
 $M[E,)]$

Stack	Input
$\$ E' T') E' T' F$	$id) * id \$$
$\$ E' T') E' T' id$	$id) * id \$$
$\$ E' T') E' T'$	$) * id \$$
$\$ E' T') E'$	$) * id \$$
$\$ E' T')$	$) * id \$$
$\$ E' T'$	$* id \$$
$\$ E' T' F *$	$* id \$$
$\$ E' T' F$	$id \$$
$\$ E' T' id$	$id \$$
$\$ E' T'$	$\$$
$\$ E'$	$\$$
$\$$	$\$$

Outline

- ① Left Factoring
- ② Nullability
- ③ The FIRST Function
- ④ The FOLLOW Function
- ⑤ The Parse Table Entries
- ⑥ The Parse Table
- ⑦ Predictive Parsing
- ⑧ Assignment



Assignment

Homework

- For the grammar

$$S \rightarrow (L) \mid \text{id}$$

$$L \rightarrow L , S \mid S$$

construct a parse table for this grammar.

- Parse the expression **(id, id)**.

