Parsing

Sujit Kumar Chakrabarti

IIITB

Recursive Descent Parsing

Recursive Descent Parsing

- Appropriate for manual implementation
- Top-down parsing
- Starts with the root
- Prepares the parse tree in pre-order depth first sequence
- Finds the *leftmost derivation* for a string
- LL(k) Grammars

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation	Rightmost Derivation
E	

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation	**Rightmost Derivation**
$E \quad \Rightarrow \quad E + E$	

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation
E	$\Rightarrow$	$E + E$	
	$\Rightarrow$	$\mathbf{id}_1 + E$	

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation
E	$\Rightarrow$	$E + E$	
	$\Rightarrow$	$\mathbf{id}_1 + E$	
	$\Rightarrow$	$\mathbf{id}_1 + (E)$	

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation
E	$\Rightarrow$	$E + E$	
	$\Rightarrow$	$\mathbf{id}_1 + E$	
	$\Rightarrow$	$\mathbf{id}_1 + (E)$	
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$	

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation

$$\begin{aligned} E &\Rightarrow E + E \\ &\Rightarrow \mathbf{id}_1 + E \\ &\Rightarrow \mathbf{id}_1 + (E) \\ &\Rightarrow \mathbf{id}_1 + (E + E) \\ &\Rightarrow \mathbf{id}_1 + (\mathbf{id}_2 + E) \end{aligned}$$

Rightmost Derivation

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation
E	$\Rightarrow$	$E + E$	E
	$\Rightarrow$	$\mathbf{id}_1 + E$	
	$\Rightarrow$	$\mathbf{id}_1 + (E)$	
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$	
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$	
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$	

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation		
E	$\Rightarrow$	$E + E$	E	$\Rightarrow$	$E + E$
	$\Rightarrow$	$\mathbf{id}_1 + E$			
	$\Rightarrow$	$\mathbf{id}_1 + (E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$			

Leftmost and Rightmost Derivation

Grammar:

$$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation

$$\begin{aligned} E &\Rightarrow E + E \\ &\Rightarrow \mathbf{id}_1 + E \\ &\Rightarrow \mathbf{id}_1 + (E) \\ &\Rightarrow \mathbf{id}_1 + (E + E) \\ &\Rightarrow \mathbf{id}_1 + (\mathbf{id}_2 + E) \\ &\Rightarrow \mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3) \end{aligned}$$

Rightmost Derivation

$$\begin{aligned} E &\Rightarrow E + E \\ &\Rightarrow E + (E) \end{aligned}$$

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation		
E	$\Rightarrow$	$E + E$	E	$\Rightarrow$	$E + E$
	$\Rightarrow$	$\mathbf{id}_1 + E$		$\Rightarrow$	$E + (E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E)$		$\Rightarrow$	$E + (E + E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$			

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation		
E	$\Rightarrow$	$E + E$	E	$\Rightarrow$	$E + E$
	$\Rightarrow$	$\mathbf{id}_1 + E$		$\Rightarrow$	$E + (E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E)$		$\Rightarrow$	$E + (E + E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$		$\Rightarrow$	$E + (E + \mathbf{id}_3)$
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$			
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$			

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation		
E	$\Rightarrow$	$E + E$	E	$\Rightarrow$	$E + E$
	$\Rightarrow$	$\mathbf{id}_1 + E$		$\Rightarrow$	$E + (E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E)$		$\Rightarrow$	$E + (E + E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$		$\Rightarrow$	$E + (E + \mathbf{id}_3)$
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$		$\Rightarrow$	$E + (\mathbf{id}_2 + \mathbf{id}_3)$
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$			

Leftmost and Rightmost Derivation

Grammar:

$E \quad \rightarrow \quad E + E \mid (E) \mid \mathbf{id}$

Example:

$1 + (2 + 3) \rightarrow \mathbf{id}_1 + \mathbf{id}_2 + \mathbf{Id}_3$

Leftmost Derivation			Rightmost Derivation		
E	$\Rightarrow$	$E + E$	E	$\Rightarrow$	$E + E$
	$\Rightarrow$	$\mathbf{id}_1 + E$		$\Rightarrow$	$E + (E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E)$		$\Rightarrow$	$E + (E + E)$
	$\Rightarrow$	$\mathbf{id}_1 + (E + E)$		$\Rightarrow$	$E + (E + \mathbf{id}_3)$
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + E)$		$\Rightarrow$	$E + (\mathbf{id}_2 + \mathbf{id}_3)$
	$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$		$\Rightarrow$	$\mathbf{id}_1 + (\mathbf{id}_2 + \mathbf{id}_3)$

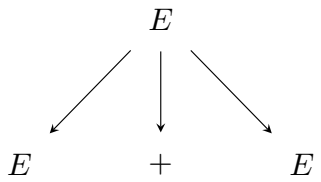
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation

E

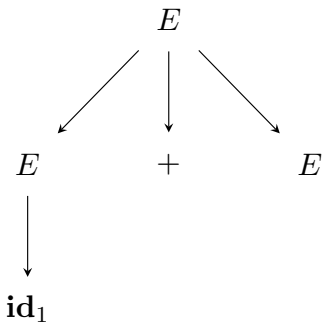
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



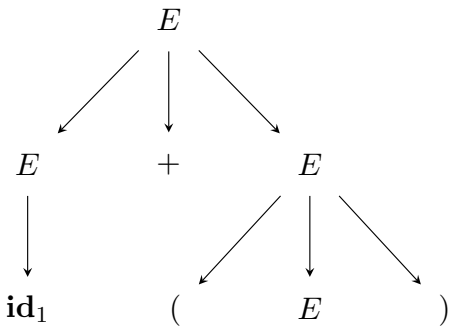
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



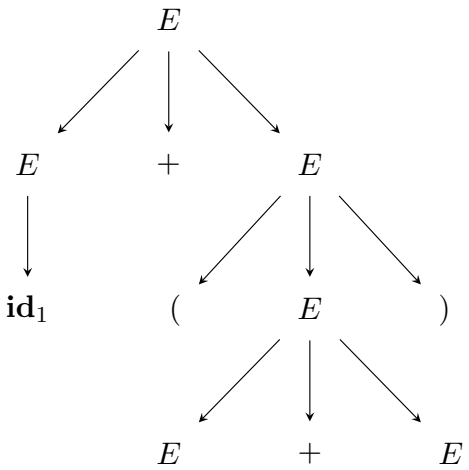
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



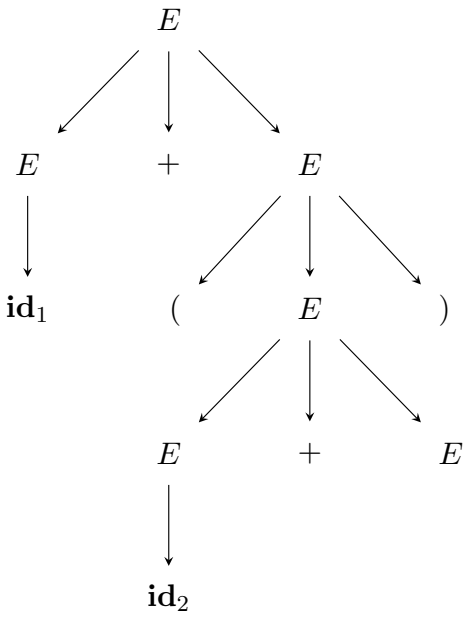
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



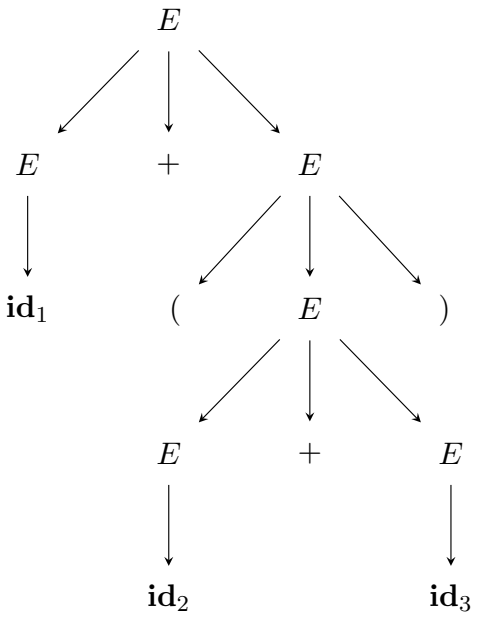
Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



Leftmost and Rightmost Derivation

Formation of Parse Tree with Leftmost Derivation



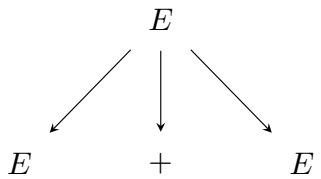
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation

E

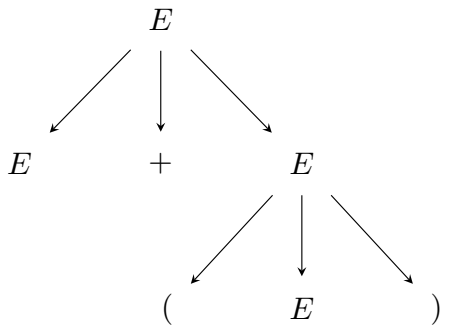
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



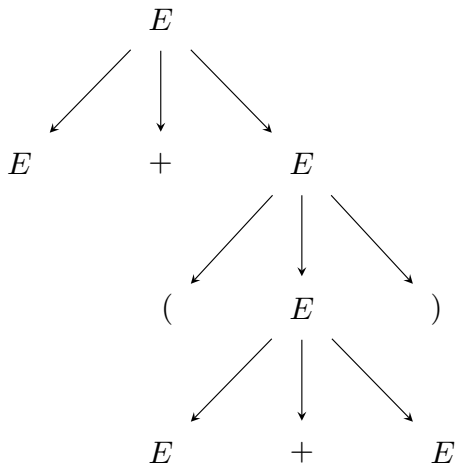
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



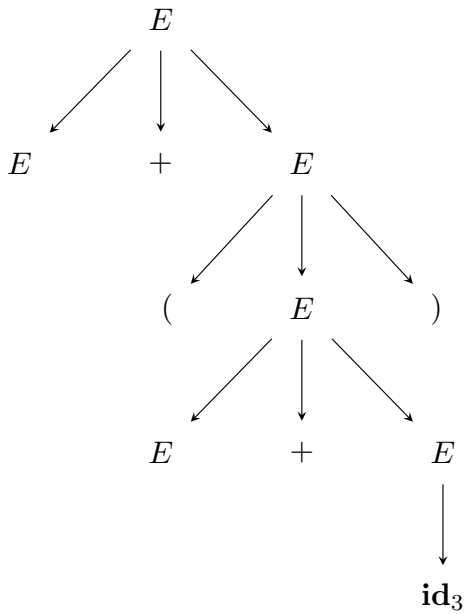
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



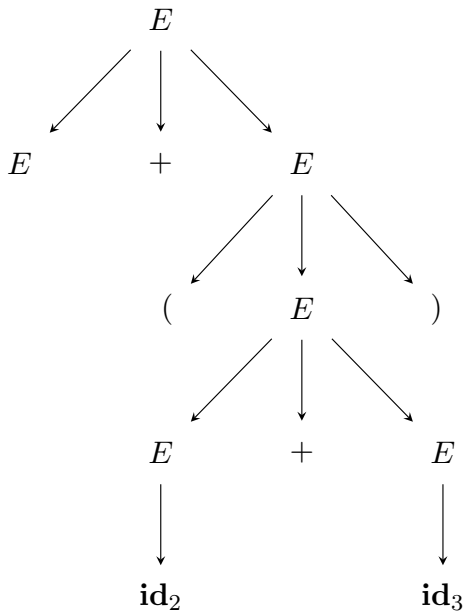
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



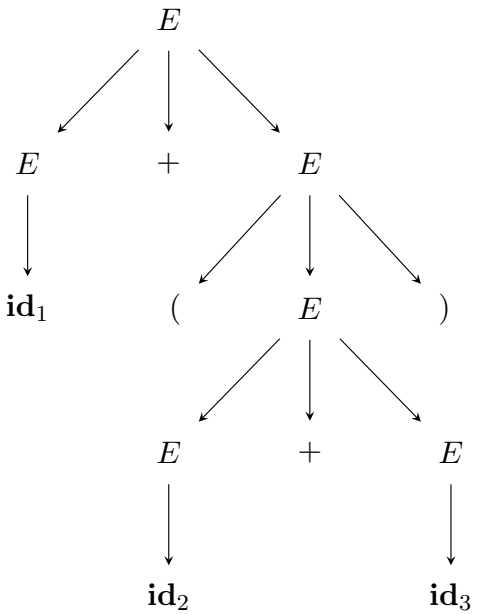
Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



Leftmost and Rightmost Derivation

Formation of Parse Tree with Rightmost Derivation



Recursive Descent Parsing

With backtracking – Example

Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ `"cad"`

Recursive Descent Parsing

With backtracking – Example

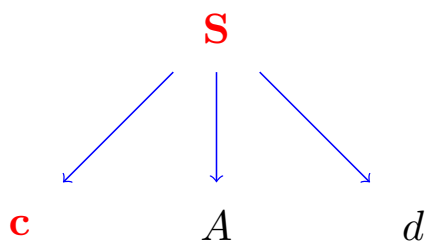
Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ "cad"

Step 1:



Recursive Descent Parsing

With backtracking – Example

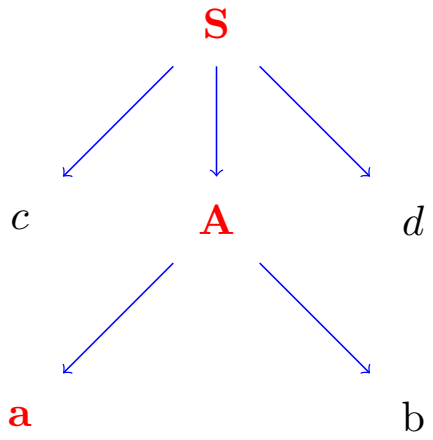
Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ "cad"

Step 2:



Recursive Descent Parsing

With backtracking – Example

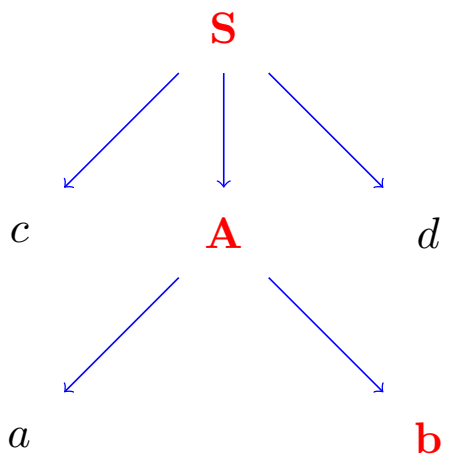
Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ "cad"

Step 3:



Failure

Recursive Descent Parsing

With backtracking – Example

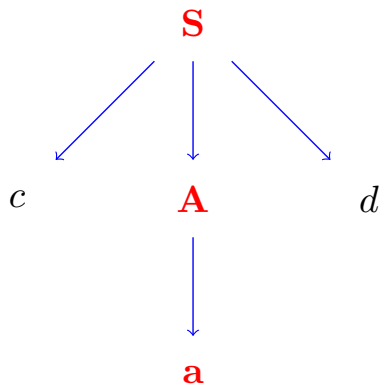
Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ "cad"

Step 4:



Backtracking

Recursive Descent Parsing

With backtracking – Example

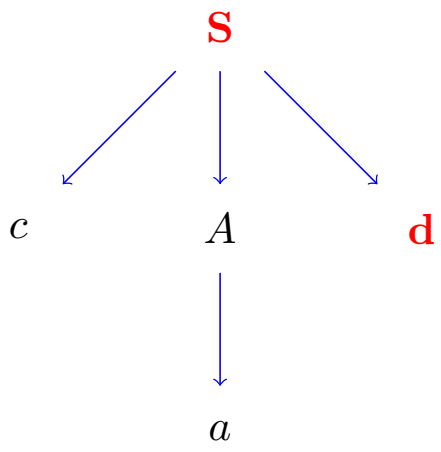
Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & c\ A\ d \\
A & \rightarrow & a\ b \\
A & \rightarrow & a
\end{array}
$$

Input string:

$w =$ "cad"

Step 5:



Success

Recursive Descent Parsing

With backtracking

```
procedure S(pos)
    if MATCH(pos,'c') then
        if A1(pos + 1) then
            if match(pos + 3, 'd') then
                return true
            else
                return false
        else if (A2(pos + 1)) then
            if MATCH((pos + 2, 'd')) then
                return true
    else
        return false
```

Recursive Descent Parsing

With backtracking

```
procedure A1(pos)
    if MATCH(pos, 'a') and MATCH(pos + 1, 'b') then
        return true
    else
        return false

procedure A2(pos)
    if MATCH(pos, 'a') then
        return true
    else
        return false
```

Recursive Descent Parsing

With backtracking

- Powerful algorithm
- Backtracking
- Back and forth movement of input pointer
- May lead to inefficiency and complexity

Recursive Descent Parsing

Predictive Parsing – Example

Grammar:

$$
\begin{aligned}
E &\rightarrow \mathbf{num}\ E' \\
E' &\rightarrow +\ E\ E' \mid \epsilon
\end{aligned}
$$

Example:

$1 + 2 + 3 \rightarrow \mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$

Recursive Descent Parsing

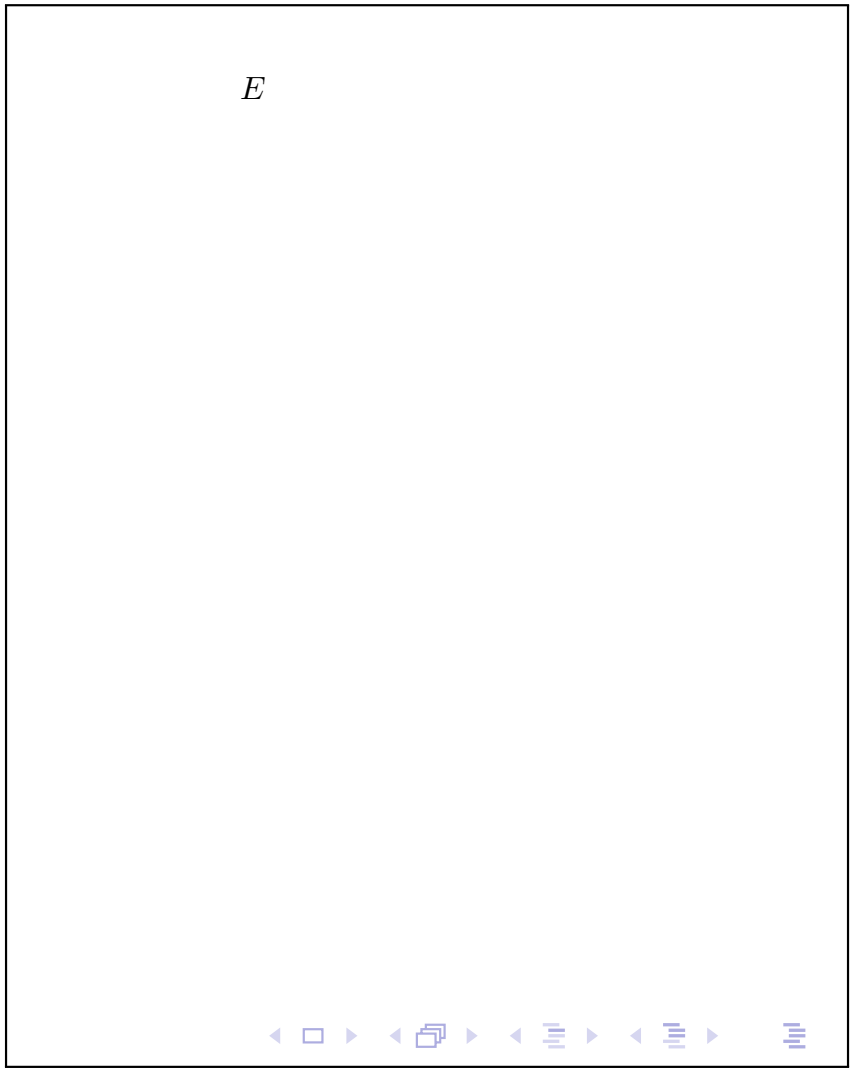
Predictive Parsing – Example

Grammar:

$$
\begin{array}{lcl}
E & \rightarrow & \textbf{num}\ E' \\
E' & \rightarrow & +\ E\ E' \mid \epsilon
\end{array}
$$

Example:

$\textbf{num}_1 + \textbf{num}_2 + \textbf{num}_3$



Recursive Descent Parsing

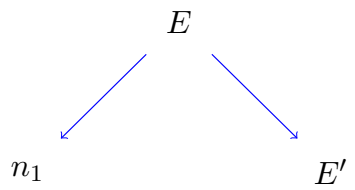
Predictive Parsing – Example

Grammar:

$$E \rightarrow \mathbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$



Recursive Descent Parsing

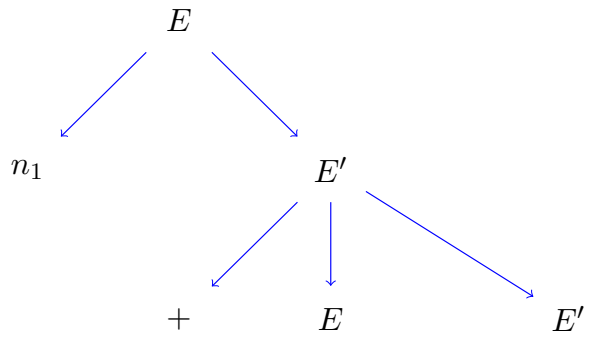
Predictive Parsing – Example

Grammar:

$$E \rightarrow \mathbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$



Recursive Descent Parsing

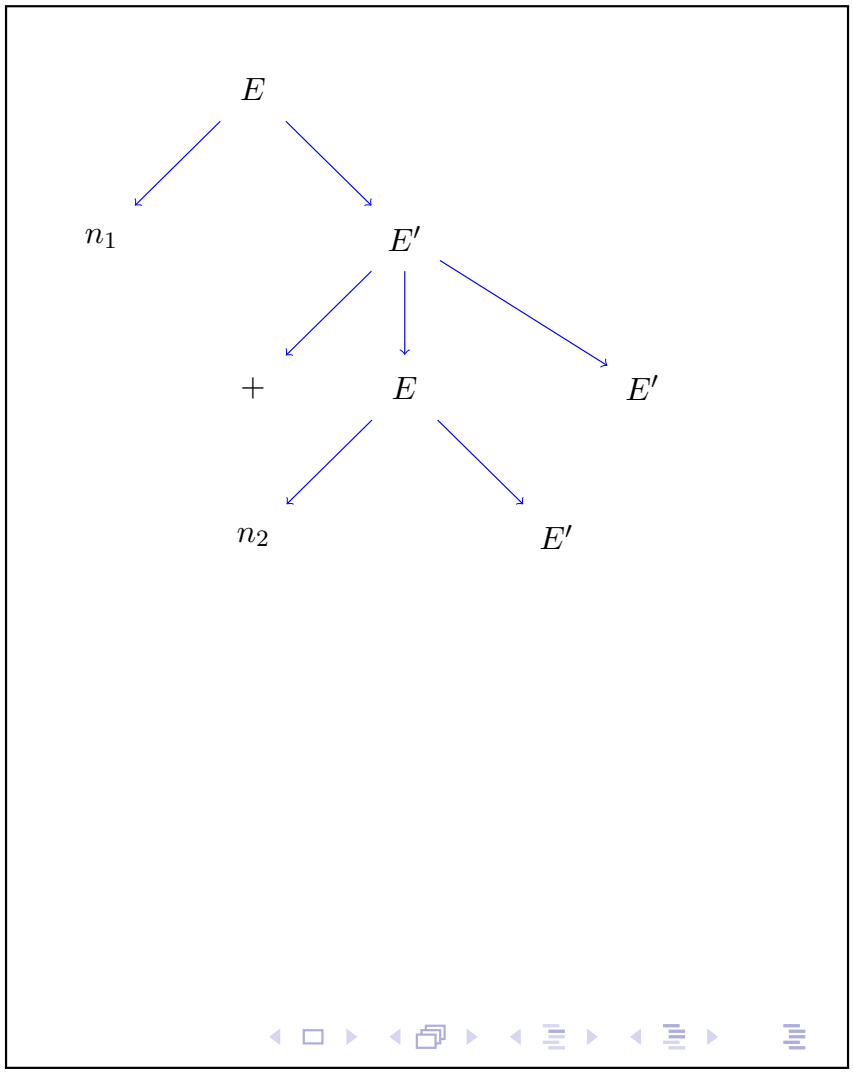
Predictive Parsing – Example

Grammar:

$$E \rightarrow \textbf{num}\ E'$$
$$E' \rightarrow + E\ E' \mid \epsilon$$

Example:

$\textbf{num}_1 + \textbf{num}_2 + \textbf{num}_3$



Recursive Descent Parsing

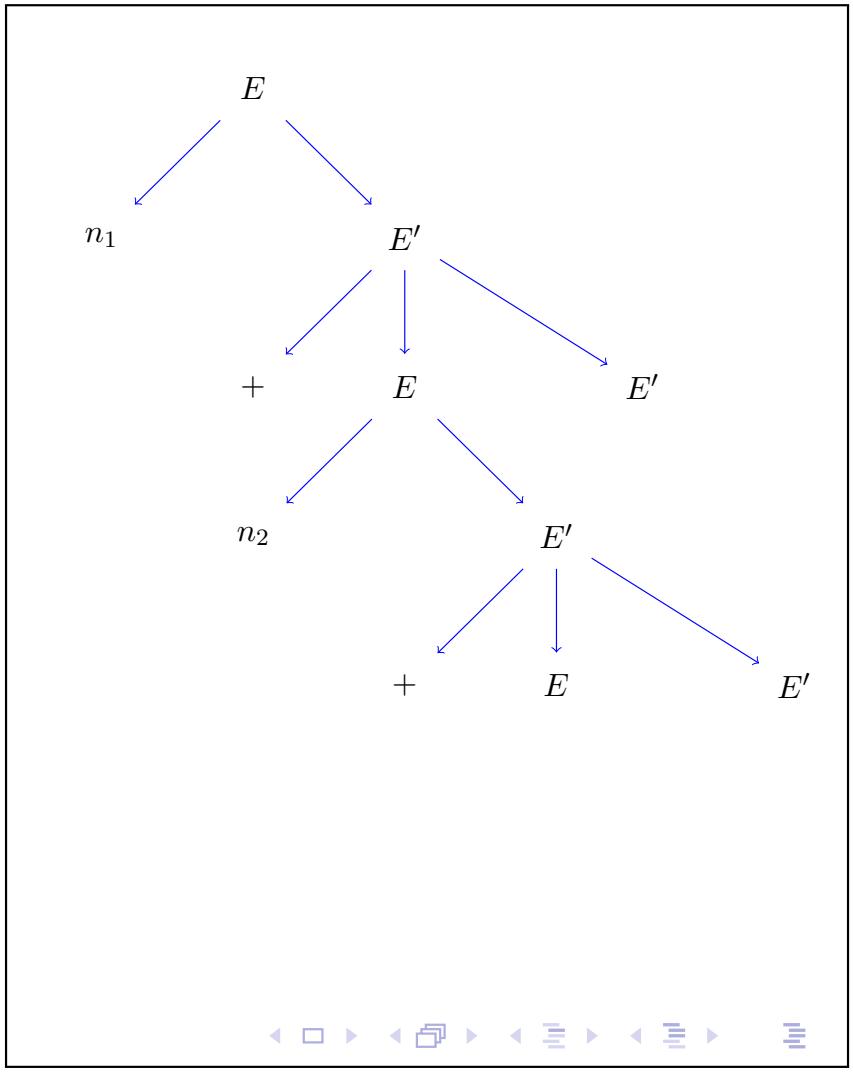
Predictive Parsing – Example

Grammar:

$$E \rightarrow \mathbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$



Recursive Descent Parsing

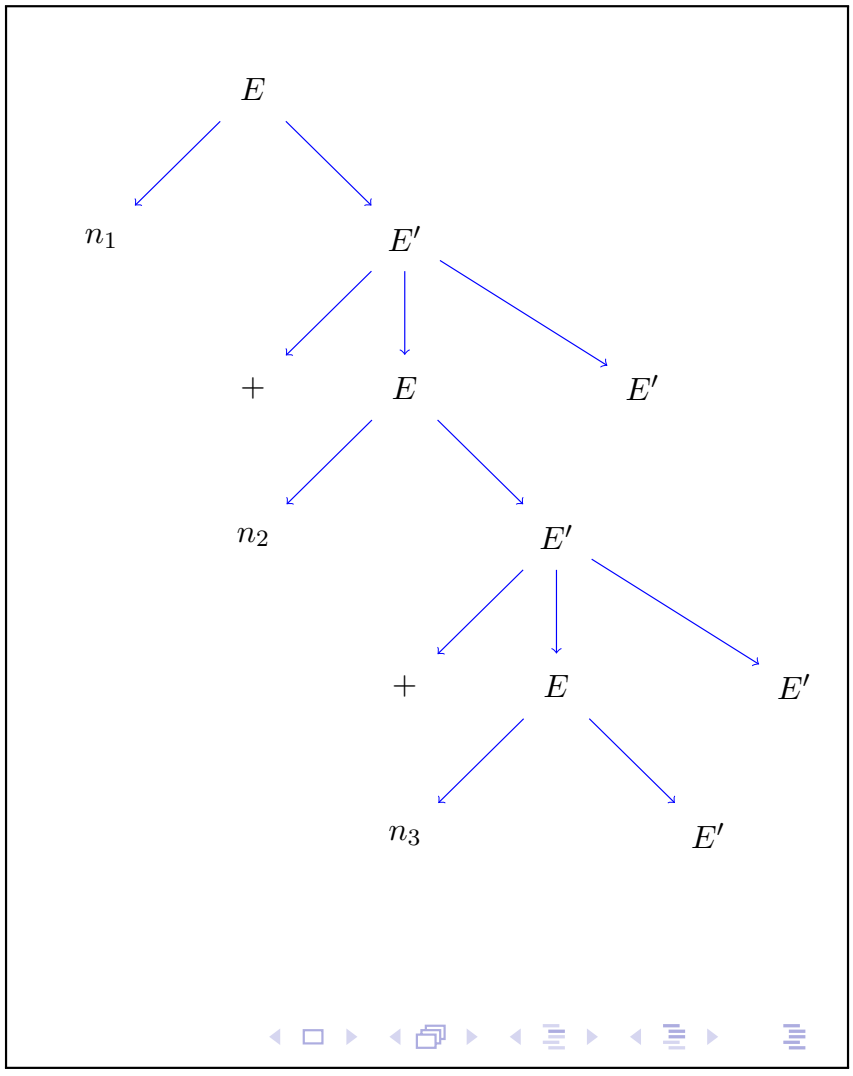
Predictive Parsing – Example

Grammar:

$$E \rightarrow \mathbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$



Recursive Descent Parsing

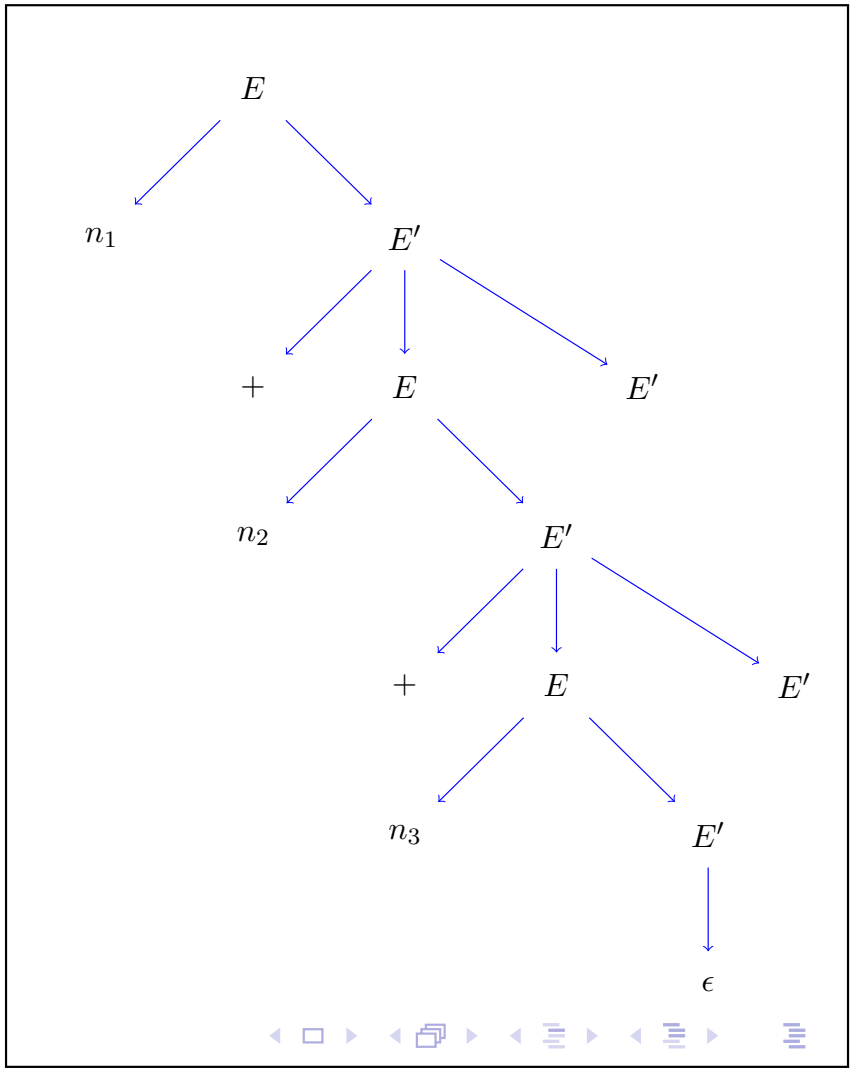
Predictive Parsing – Example

Grammar:

$$E \rightarrow \textbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\textbf{num}_1 + \textbf{num}_2 + \textbf{num}_3$



Recursive Descent Parsing

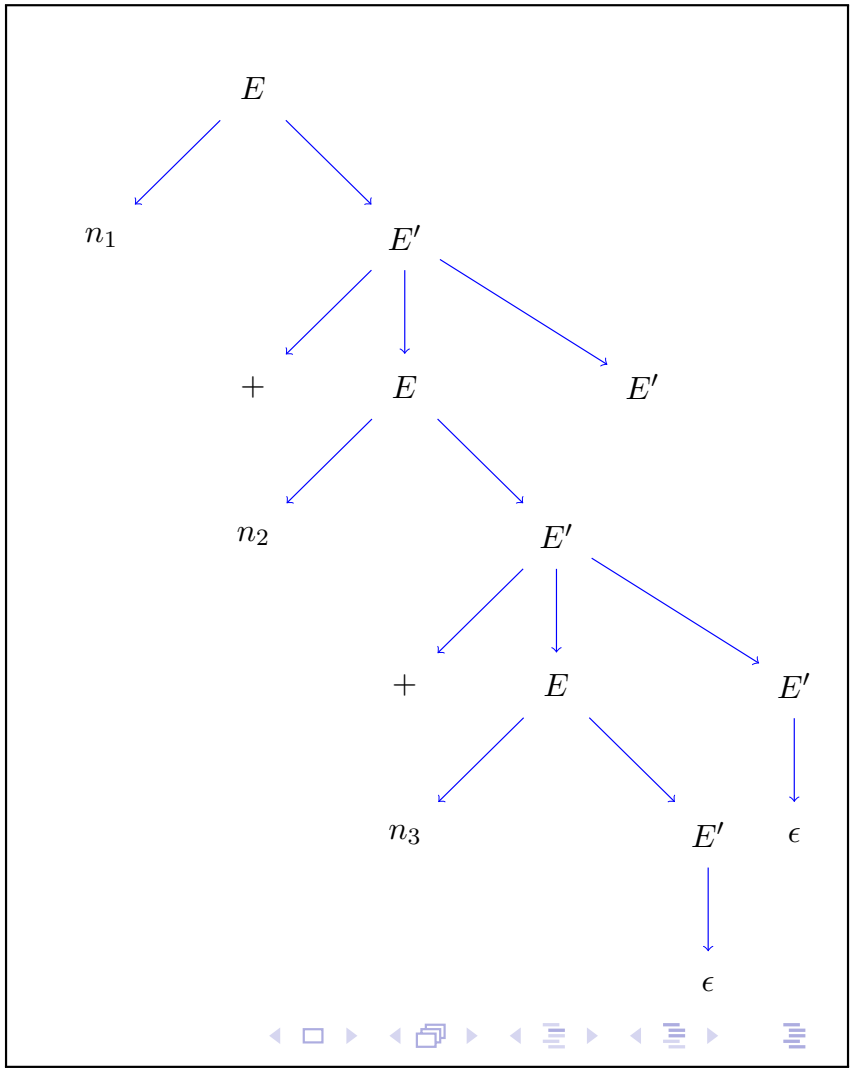
Predictive Parsing – Example

Grammar:

$$E \rightarrow \mathbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\mathbf{num}_1 + \mathbf{num}_2 + \mathbf{num}_3$



Recursive Descent Parsing

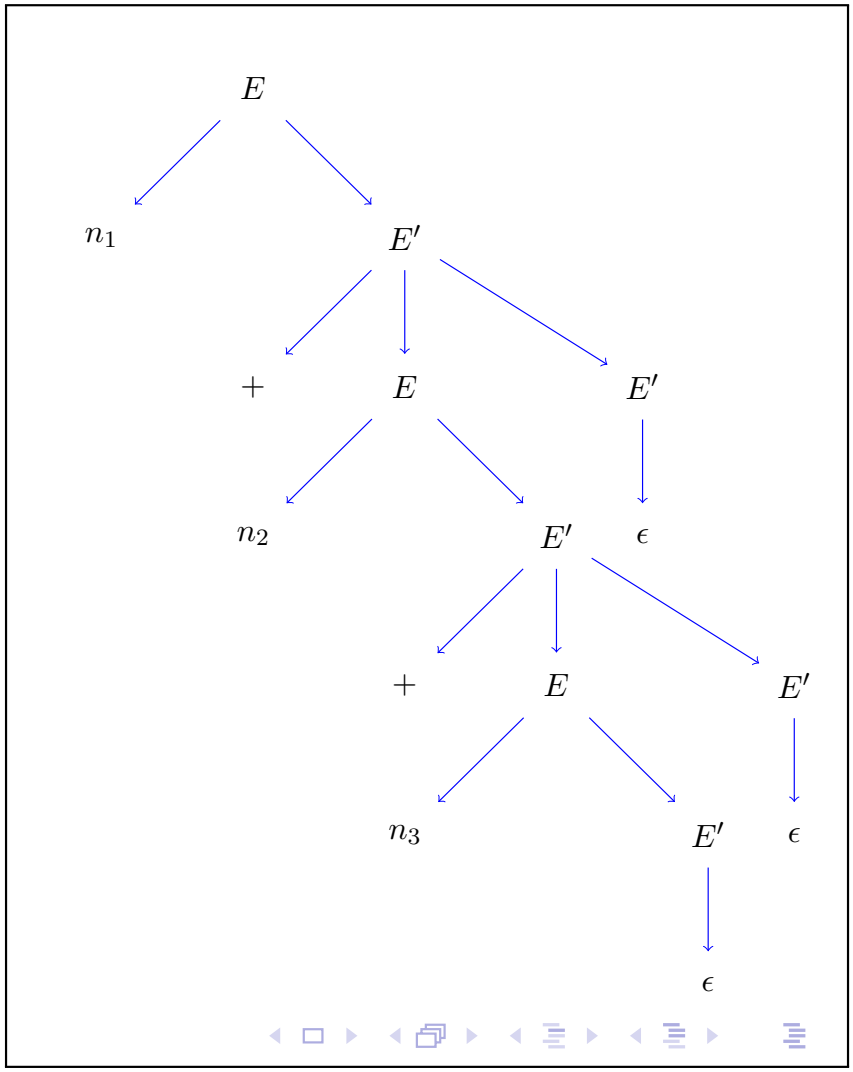
Predictive Parsing – Example

Grammar:

$$E \rightarrow \textbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

Example:

$\textbf{num}_1 + \textbf{num}_2 + \textbf{num}_3$



Recursive Descent Parsing

Algorithm

Grammar:

$$
\begin{array}{lcl}
E & \rightarrow & \textbf{num}\ E' \\
E' & \rightarrow & +\ E\ E' \mid \epsilon
\end{array}
$$

Recursive Descent Parsing

Algorithm

Grammar:

$$E \rightarrow \textbf{num}\ E'$$
$$E' \rightarrow +\ E\ E' \mid \epsilon$$

```
procedure E
    return MATCH(num) and E'

procedure E'
    return E'_1 or E'_2

procedure E'_1
    return MATCH(+) and E and E'

procedure E'_2
    return true
```

Recursive Descent Parsing

Algorithm – Activity

Grammar:

$$
\begin{array}{lcl}
expr & \rightarrow & \textbf{num} \mid term \\
term & \rightarrow & factor \mid factor + term \\
factor & \rightarrow & \textbf{num} \mid \textbf{num} * factor
\end{array}
$$

Recursive Descent Parsing

Left Recursion

Grammar:

$$E \quad \rightarrow \quad E + T \mid \textbf{num}$$

Recursive Descent Parsing

Left Recursion

Grammar:

$$E \quad \rightarrow \quad E + T \mid \textbf{num}$$

```
procedure E
    return E and MATCH(+) and MATCH(num)

procedure T
    return E'_1 or E'_2
```

Recursive Descent Parsing

Left Recursion

Grammar:

$$E \quad \rightarrow \quad E + T \mid \mathbf{num}$$

Recursive Descent Parsing

Left Recursion

Grammar:

$$E \rightarrow E + T \mid \textbf{num}$$

$\Rightarrow$

Modified Grammar:

$$\begin{aligned} E &\rightarrow \textbf{num}\ E' \\ E' &\rightarrow +\ E\ E' \mid \epsilon \end{aligned}$$

Recursive Descent Parsing

Left Recursion

Grammar:

$$E \rightarrow E + T \mid \textbf{num}$$

$\Rightarrow$

Modified Grammar:

$$\begin{aligned} E &\rightarrow \textbf{num}\ E' \\ E' &\rightarrow +\ E\ E' \mid \epsilon \end{aligned}$$

- Algorithm available for removing left recursion
- Self-study