

# Bottom-Up Parsing

Copyright 2010, Keith D. Cooper & Linda Torczon, all rights reserved.



- Top-down parsers build syntax tree from root to leaves
- Left-recursion causes non-termination in top-down parsers
  - -Transformation to eliminate left recursion
  - Transformation to eliminate common prefixes in right recursion

ELAWARE 1743 %

Recap of Top-down Parsing (cont'd)

- FIRST, FIRST<sup>+</sup>, & FOLLOW sets + LL(1) condition
  - —LL(1) uses <u>left-to-right scan</u> of the input, <u>leftmost derivation</u> of the sentence, and <u>1</u> word lookahead
  - —LL(1) condition means grammar works for predictive parsing
- Given an LL(1) grammar, we can
   Build a table-driven LL(1) parser
- LL(1) parser keeps lower fringe of partially complete tree on the stack

## Parsing Techniques



Top-down parsers (LL(1), recursive descent)

- Start at root of the parse tree and grow toward leaves
- Pick a production & try to match the input
- Bad "pick"  $\Rightarrow$  may need to backtrack
- Some grammars are backtrack-free (predictive parsing)

#### Bottom-up parsers (LR(1), operator precedence)

- Start at the leaves and grow toward root
- As input consumed, encode possibilities in internal state
- Start in a state valid for legal first tokens
- Bottom-up parsers handle a large class of grammars



The point of parsing is to construct a <u>derivation</u>

A derivation consists of a series of rewrite steps

 $\boldsymbol{\mathcal{S}} \Rightarrow \boldsymbol{\gamma}_{0} \ \Rightarrow \boldsymbol{\gamma}_{1} \ \Rightarrow \boldsymbol{\gamma}_{2} \ \Rightarrow ... \ \Rightarrow \boldsymbol{\gamma}_{n-1} \Rightarrow \boldsymbol{\gamma}_{n} \Rightarrow \boldsymbol{\textit{sentence}}$ 

- Each  $\gamma_i$  is a sentential form
  - -If  $\gamma$  contains only terminal symbols,  $\gamma$  is a sentence in L(G)
  - —If  $\gamma$  contains 1 or more non-terminals,  $\gamma$  is a sentential form



 $\boldsymbol{\mathcal{S}} \Rightarrow \gamma_{0} \ \Rightarrow \gamma_{1} \ \Rightarrow \gamma_{2} \ \Rightarrow ... \ \Rightarrow \gamma_{n-1} \Rightarrow \gamma_{n} \Rightarrow \textit{sentence}$ 

- To get  $\gamma_i$  from  $\gamma_{i-1}$ , expand some NT  $A \in \gamma_{i-1}$  by using  $A \rightarrow \beta$ —Replace the occurrence of  $A \in \gamma_{i-1}$  with  $\beta$  to get  $\gamma_i$ 
  - -In a leftmost derivation, it would be first NT  $A \in \gamma_{i-1}$
- A *left-sentential form* occurs in a *leftmost* derivation
- A *right-sentential form* occurs in a *rightmost* derivation

Bottom-up parsers build rightmost derivation in reverse



A bottom-up parser builds derivation by working from input sentence <u>back</u> toward the start symbol S



**Bottom-up** Parsing

(definitions)



In terms of parse tree, it works from leaves to root

- Nodes with no parent in partial tree form upper fringe
- Each replacement of  $\beta$  with  $\boldsymbol{A}$  shrinks the upper fringe, we call this a *reduction*.
- "Rightmost derivation in reverse" processes words *left to* right



**Bottom-up** Parsing

(definitions)



In terms of parse tree, it works from leaves to root

- Nodes with no parent in partial tree form upper fringe
- Each replacement of  $\beta$  with  $\boldsymbol{A}$  shrinks the upper fringe, we call this a *reduction*.
- "Rightmost derivation in reverse" processes words *left to* right





#### Finding Reductions

Sentential	Next Reduction		
Form	Prod'n	Pos'n 🔪	
abbcde	2	2	
<u>a</u> A <u>bcde</u>			

And the input string <u>abbcde</u>

"Position" specifies where the right end of  $\beta$  occurs in the current sentential form. We call this position k.



#### Finding Reductions

Can	aida	n +ha a					
Con	side	r The g	ran	mar.	Sentential	Next R	eduction
	0	Goal	$\rightarrow$	<u>a</u> A B <u>e</u>	Form	Prod'n	Pos'n 🔪
	1	A	$\rightarrow$	A <u>b c</u>	<u>abbcde</u>	2	2
	2			<u>b</u>	<u>a</u> A <u>bcde</u>	1	4
	3	В	$\rightarrow$	<u>d</u>	<u>a</u> A <u>de</u>	3	3
		<u>a</u> A B <u>e</u>	0	4			
Anc	i the	input	stri	ng <u>abbcde</u>	Goal	_	_

"Position" specifies where the right end of  $\beta$  occurs in the current sentential form. We call this position k. Finding Reductions





Parser must find substring  $\beta$  at parse tree's frontier that matches some production  $A \rightarrow \beta$ 

 $(\Rightarrow \beta \rightarrow A \text{ is in Reverse Rightmost Derivation})$ 

We call substring  $\beta$  a handle





Formally,

- $A \rightarrow \beta \in P$  and k is the position in  $\gamma$  of  $\beta$ 's rightmost symbol.
- If  $\langle A \rightarrow \beta, k \rangle$  is a handle, then replacing  $\beta$  at k with A produces the right sentential form from which  $\gamma$  is derived in the rightmost derivation.



0	Goal	$\rightarrow$	Expr
1	Expr	$\rightarrow$	Expr + Term
2			Expr - Term
3			Term
4	Term	$\rightarrow$	Term * Factor
5			Term / Factor
6			Factor
7	Factor	$\rightarrow$	number
8			id
9			(Expr)

Bottom up parsers can handle either left-recursive or right-recursive grammars.

A simple left-recursive form of the classic expression grammar



parse

#### On ChalkBoard Example

				Prod'	Sentential Form	Handle
0	Goal	$\rightarrow$	Expr	n		
1	Expr	$\rightarrow$	Expr + Term	8	<id,<u>x&gt; - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u></id,<u>	8,1
2		Ι	Expr - Term		<i>Factor</i> - <num<u>,2&gt; * <id,y></id,y></num<u>	
3			Term			
4	Term	$\rightarrow$	Term * Factor			
5			Term / Factor			
6			Factor			
7	Factor	$\rightarrow$	number			
8			id			
9			<u>( Expr )</u>			

A simple left-recursive form of the classic expression grammar

Handles for rightmost derivation of  $\underline{x} = \underline{2} \stackrel{*}{\underline{}} \underline{y}$ 



parse

#### On ChalkBoard Example

				Prod'	Sentential Form	Handle
0	Goal	$\rightarrow$	Expr	n		
1	Expr	$\rightarrow$	Expr + Term	8	<id<u>,x&gt; - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u></id<u>	8,1
2			Expr - Term	6	<i>Factor</i> - <num<u>,2&gt; * <id,y></id,y></num<u>	6,1
3			Term	3	<i>Term</i> - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u>	3,1
4	Term	$\rightarrow$	Term * Factor	7	<i>Expr</i> - <num<u>,2&gt; * <id,<u>y&gt;</id,<u></num<u>	7,3
5			Term / Factor	6	Expr - Factor * <id,¥></id,¥>	6,3
6			Factor	8	Expr - Term * <id,y></id,y>	8,5
7	Factor	$\rightarrow$	number	4	Expr - Term * Factor	4,5
8		Ι	id	2	Expr - Term	2,3
9			<u>(</u> Expr <u>)</u>	0	Expr	0,1
				-	Goal	-

A simple left-recursive form of the classic expression grammar

Handles for rightmost derivation of  $\underline{x} - \underline{2} + \underline{y}$ 

(Abstract View)



A bottom-up parser repeatedly finds a handle  $A \rightarrow \beta$  in current right-sentential form and replaces  $\beta$  with A.

To construct a rightmost derivation

 $\boldsymbol{\mathcal{S}} \Rightarrow \boldsymbol{\gamma}_0 \ \Rightarrow \boldsymbol{\gamma}_1 \ \Rightarrow \boldsymbol{\gamma}_2 \ \Rightarrow ... \ \Rightarrow \boldsymbol{\gamma}_{n-1} \Rightarrow \boldsymbol{\gamma}_n \Rightarrow \boldsymbol{\mathcal{W}}$ 

Apply the following conceptual algorithm

for  $i \leftarrow n$  to 1 by -1 Find the handle  $\langle A_i \rightarrow \beta_i, k_i \rangle$  in  $\gamma_i$ Replace  $\beta_i$  with  $A_i$  to generate  $\gamma_{i-1}$ 

of course, *n* is unknown until the derivation is built

This takes 2n steps

Bottom-up Parsing



Bottom-up parsers finds rightmost derivation

- Process input left to right
- Handle always appears at upper fringe of partially completed parse tree

### LR parsing



- Keep upper fringe of the partially completed parse tree on a <u>stack</u>
  - -Stack makes position information irrelevant
  - -Handles appear at top of the stack (TOS)

If G is unambiguous, then every right-sentential form has a unique handle.

## ELAWARE 1743

#### More on Handles

Prod'n	Sentential Form	Handle	*
8	<id,<u>x&gt; - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u></id,<u>	8,1	
6	Factor - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u>	6,1	* <id,<u>y&gt;</id,<u>
3	<i>Term</i> - <num,<mark>2&gt; * <id,y></id,y></num,<mark>	3,1	Rest of input from scanner
7	<i>Expr</i> - <num,<u>2&gt; * <id,<u>y&gt;</id,<u></num,<u>	<i>→</i> 7,3	
Factor -	→ <u>number</u>	K=:	<num,2> &lt;</num,2>
			stack

VIVERSITY OF ELAWARE

Shift-Reduce Parsing

To implement a bottom-up parser, we adopt the shiftreduce paradigm

- A shift-reduce parser is a stack automaton with <u>four</u> actions
- *Shift* next word is shifted onto the stack
- Reduce right end of handle is at top of stack
   Located handle (rhs) on top of stack
   Pop handle off stack & push appropriate *lhs*

<u>Shift</u> is just a push and a call to the scanner <u>Reduce</u> means found a handle, takes |*rhs*| pops & 1 push

But how does parser know when to shift and when to reduce? It shifts until it has a handle at the top of the stack.



- Accept stop parsing & report success
- *Error* call an error reporting/recovery routine

Accept if no input and Goal symbol on top of stack (TOS) Error otherwise

But how does parser know when to shift and when to reduce? It shifts until it has a handle at the top of the stack.

A simple *shift-reduce parser:* 

```
push INVALID
token ← next_token( )
repeat until (top of stack = Goal and token = EOF)
   if the top of the stack is a handle A \rightarrow \beta
      then // reduce \beta to A
         pop |\beta| symbols off the stack
         push A onto the stack
      else if (token ≠ EOF)
          then // shift
             push token
              token ← next_token( )
       else // need to shift, but out of input
       report an error
```



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input

A simple *shift-reduce parser*:

push INVALID token ← next\_token( ) repeat until (top of stack = Goal and token = EOF) if the top of the stack is a handle  $A \rightarrow \beta$ then *// reduce*  $\beta$  to A pop  $|\beta|$  symbols off the stack push A onto the stack else if (token ≠ EOF) then // shift push token token ← next\_token( ) else // need to shift, but out of input report an error



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input

A simple *shift-reduce parser*:

push INVALID repeat until (top of stack = Goal and token = EOF) if the top of the stack is a handle  $A \rightarrow \beta$ then *// reduce*  $\beta$  to A pop  $|\beta|$  symbols off the stack push A onto the stack else if (token ≠ EOF) then // shift push token token ← next\_token( ) else // need to shift, but out of input report an error



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input

A simple *shift-reduce parser*:

push INVALID token ← next\_token( ) repeat until (top of stack = Goal and token = EOF) if the top of the stack is a handle  $A \rightarrow \beta$ then *// reduce*  $\beta$  to A pop  $|\beta|$  symbols off the stack push A onto the stack else if (token ≠ EOF) then // shift push token token ← next\_token( ) else // need to shift, but out of input report an error



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input

A simple *shift-reduce parser:* 

push INVALID token ← next\_token( ) repeat until (top of stack = Goal and token = EOF) if the top of the stack is a handle  $A \rightarrow \beta$ then *// reduce*  $\beta$  to A pop  $|\beta|$  symbols off the stack push A onto the stack else if (token ≠ EOF) then // shift push token token ← next\_token() else // need to shift, but out of input report an error



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input

A simple *shift-reduce parser:* 

```
push INVALID
token ← next_token( )
repeat until (top of stack = Goal and token = EOF)
   if the top of the stack is a handle A \rightarrow \beta
      then // reduce \beta to A
         pop |\beta| symbols off the stack
         push A onto the stack
      else if (token ≠ EOF)
          then // shift
             push token
              token ← next_token( )
       else // need to shift, but out of input
       report an error
```



What happens on an error?

- It fails to find a handle
- Thus, it keeps shifting
- Eventually, it consumes all input



#### Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>			0	Goal	$\rightarrow$	Expr
				1	Expr	$\rightarrow$	Expr + Term
				2		Ι	Expr - Term
				3			Term
				4	Term	$\rightarrow$	Term* Factor
				5		I	Term / Factor
				6			Factor
				7	Factor	$\rightarrow$	<u>number</u>
				8		I	<u>ıd</u>
				9			<u>( Expr )</u>

1. Shift until the top of the stack is the right end of a handle



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action	-			-
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>			1	Expr	$\rightarrow$	Expr + Term
				2		Ι	Expr - Term
				3			Term
				4	Term	$\rightarrow$	Term * Factor
				5		Ι	Term / Factor
				6		I	Factor
				7	Factor	$\rightarrow$	<u>number</u>
				8		I	<u>id</u>
				9			<u>( Expr )</u>



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action		<b>a</b> 1		-
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
				2		Ι	Expr - Term
				3			Term
				4	Term	$\rightarrow$	Term * Factor
				5		Ι	Term / Factor
				6		Ι	Factor
				7	Factor	$\rightarrow$	number
				8		Ι	<u>id</u>
				9			<u>( Expr )</u>



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action	-			_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>				- /		-,
				2			Expr - Term
				3			Term
				4	Term	$\rightarrow$	Term * Factor
				5		I	Term / Factor
				6		Ι	Factor
				7	Factor	$\rightarrow$	number
				8		I	id
				9			<u>( Expr )</u>



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>			2		I	Expr - Term
				3			Term
				4	Term	$\rightarrow$	Term * Factor
				5		Ι	Term / Factor
				6		Ι	Factor
				7	Factor	$\rightarrow$	<u>number</u>
				8		Ι	id
				9			<u>( Expr )</u>

- 1. Shift until the top of the stack is the right end of a handle
- 2. Find the left end of the handle and reduce



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action	-			_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2		I	Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>			3		I	Term
				4	Term	$\rightarrow$	Term * Factor
				5		Ι	Term / Factor
				6		Ι	Factor
				7	Factor	$\rightarrow$	<u>number</u>
				8			id
				9			<u>(</u> Expr <u>)</u>



#### Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action	-			_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6				,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>			3			Term
				4	Term	$\rightarrow$	Term * Factor
<i>Expr</i> is not a handle a	it this point bec	ause redu	ucing now	5		Ι	Term / Factor
will cause backtrackir	ng.		-	6		Ι	Factor
While that statement sounds like oracular, we will see that the decision can be automated efficiently.					Factor	$\rightarrow$	<u>number</u>
		4	,	8		Ι	id
				9			<u>(</u> Expr <u>)</u>

1. Shift until the top of the stack is the right end of a handle



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ <i>Expr</i> -	<u>num</u> * <u>id</u>			4	Term	$\rightarrow$	Term* Factor
				5		Ι	Term / Factor
				6		Ι	Factor
				7	Factor	$\rightarrow$	number
				8			id
				9		Ι	<u>(</u> Expr <u>)</u>

- 1. Shift until the top of the stack is the right end of a handle
- 2. Find the left end of the handle and reduce



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action	-			-
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ Expr -	<u>num</u> * <u>id</u>	none	shift	4	Term	$\rightarrow$	Term* Factor
\$ Expr - <u>num</u>	* <u>id</u>			5		I	Term / Factor
				6		I	Factor
				7	Factor	$\rightarrow$	number
				8			<u>id</u>
				9		I	(Expr)

- 1. Shift until the top of the stack is the right end of a handle
- 2. Find the left end of the handle and reduce



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ Expr -	<u>num</u> * <u>id</u>	none	shift	4	Term	$\rightarrow$	Term * Factor
\$ Expr - <u>num</u>	* <u>id</u>	7,3	reduce 7	5		I	Term / Factor
\$ Expr - Factor	* <u>id</u>			5		I	
				6			Factor
				7	Factor	$\rightarrow$	<u>number</u>
				8			id
				9			<u>( Expr )</u>



Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ Expr -	<u>num</u> * <u>id</u>	none	shift	4	Term	$\rightarrow$	Term* Factor
\$	* <u>id</u>	7,3	reduce 7	5		Т	Term / Factor
\$ Expr - Factor	* <u>id</u>	6,3	reduce 6	0		I	
\$ Expr - Term	* <u>id</u>			6		Ι	Factor
				7	Factor	$\rightarrow$	number
				8		I	id
				9			<u>( Expr )</u>



#### Back to <u>x - 2 \* y</u>

Stack	Input	Handle	Action				-
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ Expr -	<u>num</u> * <u>id</u>	none	shift	4	Term	$\rightarrow$	Term * Factor
\$ Expr - <u>num</u>	* <u>id</u>	7,3	reduce 7	5		I	Term / Factor
\$ Expr - Factor	* <u>id</u>	6,3	reduce 6	Ũ		I	
\$ Expr - Term	* <u>id</u>	none	shift	6		I	Factor
\$ Expr - Term *	<u>id</u>	none	shift	7	Factor	$\rightarrow$	<u>number</u>
\$ Expr - Term * <u>id</u>							
				8		I	<u>Id</u>
				9			<u>( Expr )</u>

1. Shift until the top of the stack is the right end of a handle



1 accept

#### Back to <u>x</u> <u>-</u> <u>2</u> <u>\*</u> <u>y</u>

Stack	Input	Handle	Action				_
\$	<u>id</u> - <u>num</u> * <u>id</u>	none	shift	0	Goal	$\rightarrow$	Expr
\$ <u>id</u>	- <u>num</u> * <u>id</u>	8,1	reduce 8	1	Expr	$\rightarrow$	Expr + Term
\$ Factor	- <u>num</u> * <u>id</u>	6,1	reduce 6		,		,
\$ Term	- <u>num</u> * <u>id</u>	3,1	reduce 3	2			Expr - Term
\$ Expr	- <u>num</u> * <u>id</u>	none	shift	3			Term
\$ Expr -	<u>num</u> * <u>id</u>	none	shift	4	Term	$\rightarrow$	Term * Factor
\$ Expr - <u>num</u>	* <u>id</u>	7,3	reduce 7	5		I	Term / Factor
\$ Expr - Factor	* <u>id</u>	6,3	reduce 6	5		I	
\$ Expr - Term	* <u>id</u>	none	shift	6			Factor
\$ Expr - Term *	id	none	shift	7	Factor	' →	number
\$ Expr - Term * <u>id</u>		8,5	reduce 8				
\$ Expr - Term * Factor		4,5	reduce 4	8			
\$ Expr - Term		2,3	reduce 2	9		I	<u>( Expr )</u>
\$ Expr		0,1	reduce 0		5 shifts + 9 reduces +		±
\$ Goal		none	accept				TS + UCES +

1. Shift until the top of the stack is the right end of a handle



#### Back to <u>x</u> <u>-</u> <u>2</u> <u>\*</u> <u>y</u>

Stack	Input	Action
\$	<u>id</u> - <u>num</u> * <u>id</u>	shift
\$ <u>id</u>	- <u>num</u> * <u>id</u>	reduce 8
\$ Factor	- <u>num</u> * <u>id</u>	reduce 6
\$ Term	- <u>num</u> * <u>id</u>	reduce 3
\$ Expr	- <u>num</u> * <u>id</u>	shift
\$	<u>num</u> * <u>id</u>	shift
\$	* <u>id</u>	reduce 7
\$ Expr - Factor	* <u>id</u>	reduce 6
\$ Expr - Term	* <u>id</u>	shift
\$ Expr - Term *	<u>id</u>	shift
\$ Expr - Term * <u>id</u>		reduce 8
\$ Expr - Term * Factor		reduce 4
\$ Expr - Term		reduce 2
\$ Expr		reduce 0
\$ Goal		accept



Corresponding Parse Tree



# An Important Lesson about Handles

- A handle must be a substring of a sentential form  $\gamma$  such that :
  - —Must match rhs  $\beta$  of some rule  $A \rightarrow \beta$ ; and
- Simply looking for right hand sides that match strings is not good enough

An Important Lesson about Handles



 Critical Question: How can we know when we have found a handle without generating lots of different derivations?



- Critical Question: How can we know when we have found a handle without generating lots of different derivations?
  - Answer: We use left context, encoded in the sentential form, left context encoded in a "parser state", and a lookahead at the next word in the input. (Formally, 1 word beyond the handle.)
  - —We build all of this knowledge into a handlerecognizing DFA

## LR(1) Parsers



- LR(1) parsers are table-driven, shift-reduce parsers that use a limited right context (1 token) for handle recognition
- The class of grammars that these parsers recognize is called the set of LR(1) grammars

LR(1) means left-to-right scan of the input, rightmost derivation (in reverse), and 1 word of lookahead.

## Informal definition:

A grammar is LR(1) if, given a rightmost derivation

$$S \Rightarrow \gamma_0 \Rightarrow \gamma_1 \Rightarrow \gamma_2 \Rightarrow ... \Rightarrow \gamma_{n-1} \Rightarrow \gamma_n \Rightarrow sentence$$

### We can

- 1. isolate the handle of each right-sentential form  $\gamma_{i\prime}$  and
- 2. determine the production by which to reduce,
- by scanning  $\gamma_i$  from *left-to-right*, going at most 1 symbol beyond the right end of the handle of  $\gamma_i$