ΕΠΛ323 - Θεωρία και Πρακτική Μεταγλωττιστών

Lecture 6b

Syntax Analysis

Elias Athanasopoulos
eliasathan@cs.ucy.ac.cy

Bottom-up Parsing (shift-reduce parsing)

$S \rightarrow aABe$
$A \rightarrow Abc \mid b$
$B \rightarrow d$

abbcde	
aAbcde	A→b
aAde	A→Abc
aABe	B→d
S	S→aABe

Rightmost derivation

$$S \underset{rm}{=>} aABe \underset{rm}{=>} aAde \underset{rm}{=>} aAbcde \underset{rm}{=>} \mathrm{abbcde}$$

Handles

- A “handle” of a string is a substring that matches the right side of a production.
- The nonterminal of the left side of the production represents one step along the reverse of rightmost derivation.

abbcde	
aAbcde	*A→b*
aAde	*A→Abc*
aABe	*B→d*
S	*S→aABe*

abbcde	
aAbcde	*A→b*
aAAcde	*A→b*
	No production

Not all substrings are handles!

Example

$E \rightarrow E+E$
$E \rightarrow E*E$
$E \rightarrow (E)$
$E \rightarrow \mathbf{id}$

Rightmost derivation 1

$$
\begin{aligned}
E &\underset{rm}{\Rightarrow} \underline{E + E} \\
&\underset{rm}{\Rightarrow} E + \underline{E * E} \\
&\underset{rm}{\Rightarrow} E + E * \underline{\mathbf{id}} \\
&\underset{rm}{\Rightarrow} E + \underline{\mathbf{id}} * \mathbf{id} \\
&\underset{rm}{\Rightarrow} \underline{\mathbf{id}} + \mathbf{id} * \mathbf{id}
\end{aligned}
$$

Rightmost derivation 2

$$
\begin{aligned}
E &\underset{rm}{\Rightarrow} \underline{E * E} \\
&\underset{rm}{\Rightarrow} E * \underline{\mathbf{id}} \\
&\underset{rm}{\Rightarrow} \underline{E + E} * \underline{\mathbf{id}} \\
&\underset{rm}{\Rightarrow} E + \underline{\mathbf{id}} * \mathbf{id} \\
&\underset{rm}{\Rightarrow} \underline{\mathbf{id}} + \mathbf{id} * \mathbf{id}
\end{aligned}
$$

Ambiguous grammar, two rightmost derivations produce the same string.

Handle Pruning

Right-Sentenial Form	Handle	Reducing Production
id + id * id	id	E → id
E + id * id	id	E → id
E + E * id	id	E → id
E + E * E	E * E	E → E*E
E + E	E + E	E → E+E
E		

Stack Implementation of Shift-Reduce Parsing

STACK	INPUT	ACTION
$\$$	$\mathbf{id} + \mathbf{id} * \mathbf{id}\$$	shift
$\$\mathbf{id}$	$+ \mathbf{id} * \mathbf{id}\$$	reduce by $E \rightarrow \mathbf{id}$
$\$E$	$+ \mathbf{id} * \mathbf{id}\$$	shift
$\$E +$	$\mathbf{id} * \mathbf{id}\$$	shift
$\$E + \mathbf{id}$	$* \mathbf{id}\$$	reduce by $E \rightarrow \mathbf{id}$
$\$E + E$	$* \mathbf{id}\$$	shift
$\$E + E *$	$\mathbf{id}\$$	shift
$\$E + E * \mathbf{id}$	$\$$	reduce by $E \rightarrow \mathbf{id}$
$\$E + E * E$	$\$$	reduce by $E \rightarrow E*E$
$\$E + E$	$\$$	reduce by $E \rightarrow E+E$
$\$E$	$\$$	accept

Actions

- In a *shift* action, the next input symbol is shifted onto the top of the stack.
- In a *reduce* action, the parser knows the right end of the handle is at the top of the stack. It must then locate the left end of the handle within the stack and decide with what nonterminal to replace the handle.
- In an *accept* action, the parser announces successful completion of parsing.
- In an *error* action, the parser discovers that a syntax error has occurred and calls an error recovery routine.