

Syntactic Analysis

Ambiguity and Error Recovery

Copyright 2009, Pedro C. Diniz, all rights reserved.

Students enrolled in the Compilers class at Instituto Superior Técnico (IST/UTL) have explicit permission to make copies of these materials for their personal use.



Ambiguous Languages

- A Context-Free Language (CFL) is inherently ambiguous if every Context-Free Grammar that generates the language is ambiguous
- However most ambiguous grammars found in practice are for unambiguous languages
 - Would like to make it unambiguous



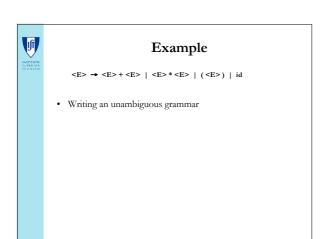
Ambiguous Grammars

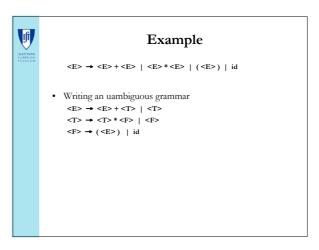
- If you have an ambiguous grammar for an unambiguous language, you can:
 - Write an unambiguous grammar
 - Use precedence and associativity to resolve parsing action conflicts

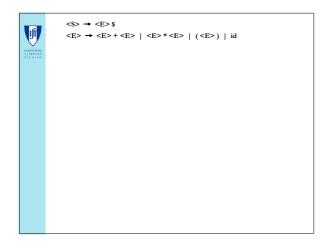


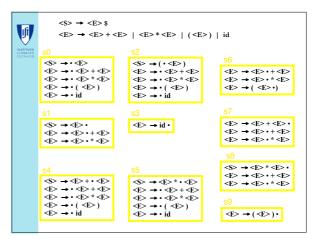
Example

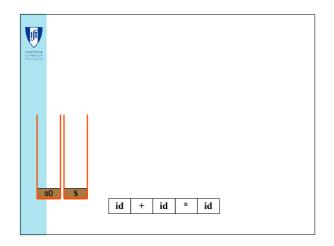
<E> → <E> + <E> | <E> * <E> | (<E>) | id

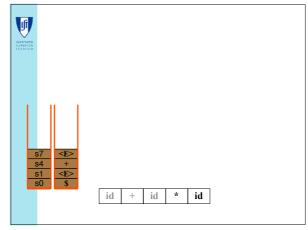


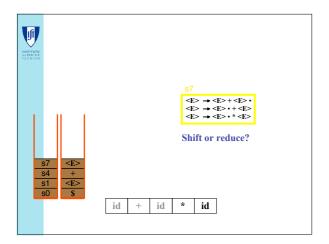


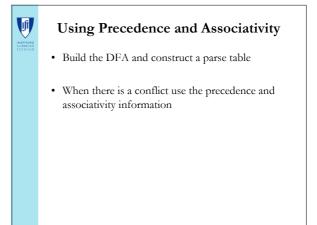


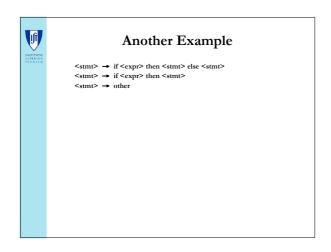


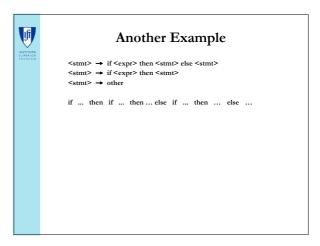


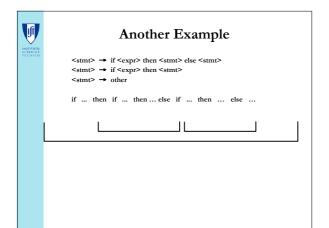


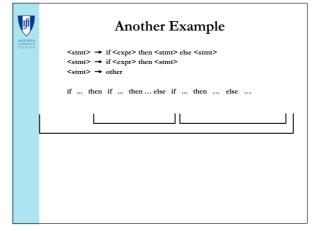


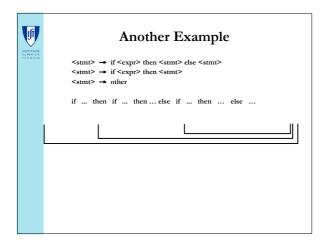


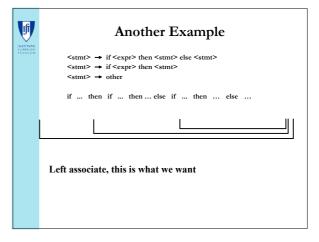


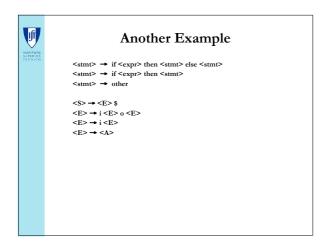


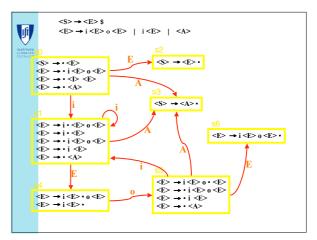


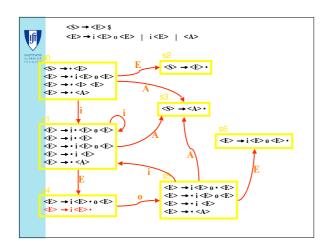


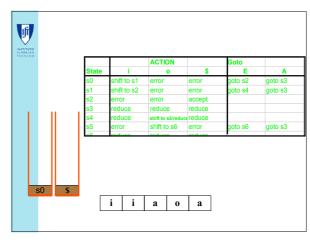


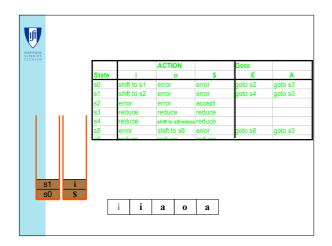


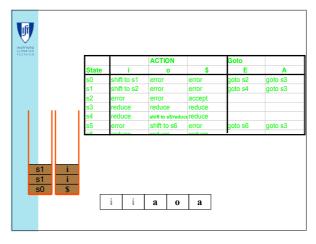


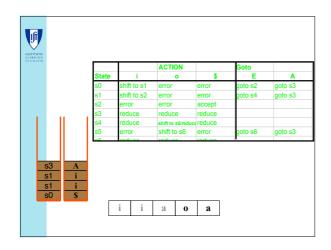


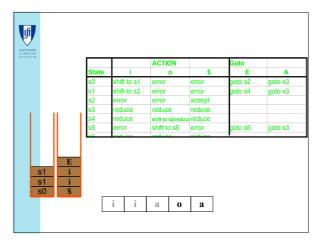


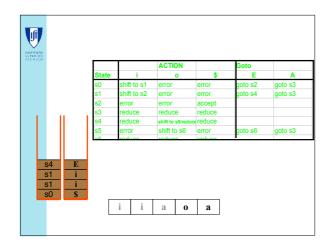


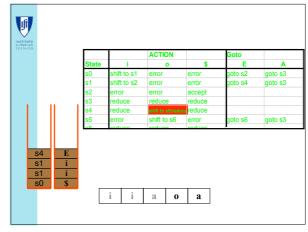


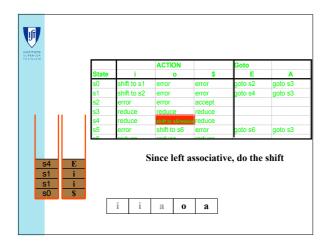


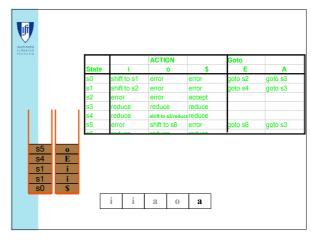


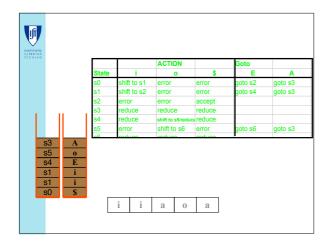


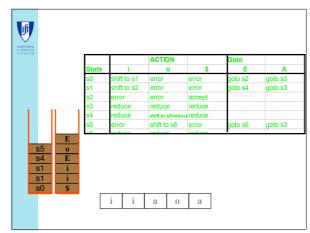


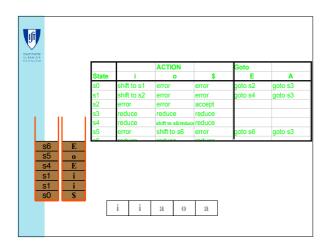


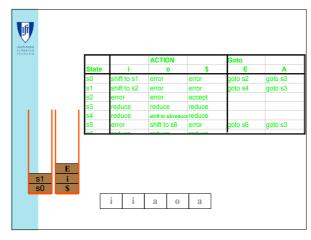


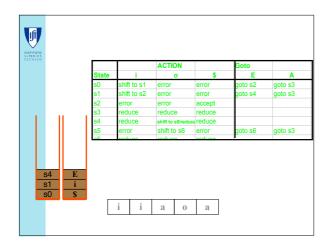


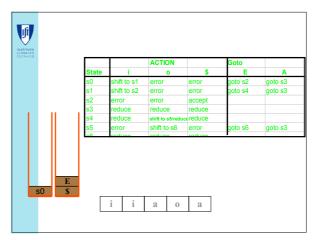


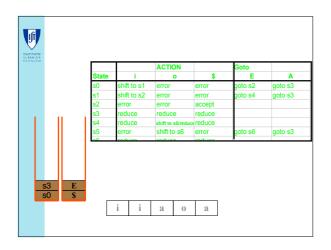


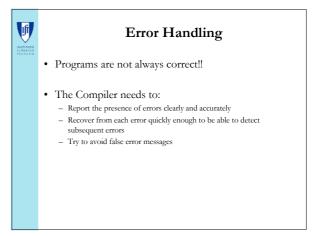


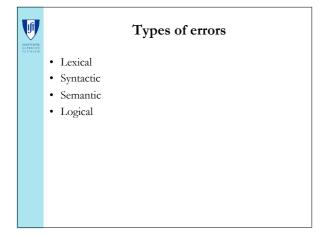


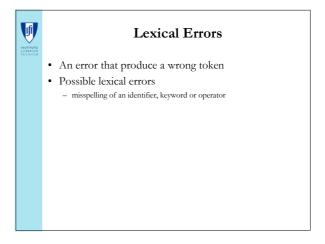














Syntactic Errors

- A program that does not satisfy the CFG of the language
- Examples
 - arithmetic expression with unbalanced parentheses
 missing semicolon



Semantic Errors

- An error that needs context sensitive information to identify
- Examples
 - Operator applied to an incompatible operand
 Accessing an undeclared variable



Logical Errors

- Errors in the Execution Model
- Examples
 - Infinitely recursive call
 - Accessing an array out of bounds
 - Dereferencing a null pointer



Types of Syntax Error Recovery

- · Panic mode recovery
- · Parse level recovery
- Error productions
- · Global correction



Panic Mode Recovery

- On Discovering an Error
 - Pop zero or more states/symbols off the stack
 - Discard zero or more input symbols
 - Until we reach a point where parsing can continue
- User Defined Panic Non-terminals
 - Example: right bracket, semicolon



Panic Mode Recovery

- On Error
 - Pops stack till it reaches a state X that can perform a goto transition on one or more panic non-terminals
 - · parsing is terminated if none found
 - Discards tokens from the input buffer till it finds a synchronizing
 - A token that belongs to follow(Λ) of the panic non-terminal Λ encountered in the previous step
 Pushes A and the state goto(S,A) on the stacks and resume parsing
- Same as pretending having seen A
- Issue: Which A and terminals to Choose?
 - Typical choices; statements and ';' or "end" tokens to synchronize

