

Advanced Programming Languages. Practice 2.

EPIC Institute of Technology

September 29, 2023

1 Parsing a language

Usually a two-stage process:

- Lexer
 - Transforms the text into a stream of tokens
 - Uses regular grammar
 - Example tokens:
 - * Keywords
 - * Operators
 - * Integers
 - * String literals
 - * Comments
 - * ...
- Parser
 - Builds an AST from the stream of tokens
 - Uses (almost) context-free grammar

2 Tools

- Parse the grammar manually
- Use a parser generator
 - Generates source code from the grammar
- Available lexer generators:
 - Flex
 - Ragel
- Available parser generators:
 - Yacc
 - Bison
 - ANTLR (lexer + parser)
 - tree-sitter (lexer + parser)

Now, let's try ANTLR in action.