HCMC University of Technology
Faculty of Computer Science & Engineering

Assignment 1

Lexer & Recognizer

Author Dr. Nguyen Hua Phung

June 22, 2018

Contents

Assignment 1

version 1.0

After completing this assignment, you will be able to

- define formally lexicon of a programming language.
- use ANTLR to implement a lexer for a programming language.
- define formally grammar of a programming language.
- use ANTLR to implement a recognizer for a programming language.

1 Specification

In this assignment, you are required to write a lexer and a recognizer for a program written in MC. To complete this assignment, you need to:

- read carefully the specification of MC language
- Modify MC.g4. in the initial code to describe formally MC language.**Please fill in your id in the header of this file**.
- Add more test in LexerSuite and ParserSuite in the initial code.

This assignment is divided two phases: lexer phase and recognizer phase. **These phases are assessed independently.**

1.1 Phase 1: Lexer

In this phase, you are required to write a lexer for a program written in ANTLR. To complete this phase, you need to:

- Modify MC.g4 to detect tokens in MC language.
- Make 100 testcases for LexerSuite to test your code.
- For lexical errors, please print out as follows:
  - "ErrorToken "+ <char>: when the lexer detects an unrecognized character
  - "Unclosed string: "+<unclosed string>: when the lexer detects an unterminated string.

  - "Illegal escape in string: "+<wrong string>: when the lexer detects an illegal escape in string. The wrong string is from the beginning of the string to the illegal escape.

- You can assume that there is only one error in each test case.

1.2 Phase 2: Recognizer

In this phase, you are required to write a recognizer for a program written in MC. To complete this phase, you need to:

- Modify MC.g4.
- Make 100 testcases for ParserSuite to test your code.
- You can assume that there is at most one error in each test case.

2 Requirements

Note that you must NOT compress your files when submit them. You may test your code at the website:
http://www.cse.hcmut.edu.vn/onlinejudge, but you MUST submit three files MC.g4, LexerSuite.py and ParserSuite.py in BKeL.

The deadline of both phases of assignment 1 is announced in the class website.

You must complete the assignment by yourself and do not let your work seen by someone else, otherwise, you will be punished by the university rule for plagiarism.