HCMC University of Technology
Faculty of Computer Science & Engineering



Assignment 1

Recognizer

cuu duong than cong . com

Author PhD. Nguyen Hua Phung

cuu duong than cong . com

September 6, 2015

CuuDuongThanCong.com https://fb.com/tailieudientucntt

Contents

1 Specification **2**
1.1 Phase 1: Lexer 2
1.2 Phase 2: Recognizer 3
1.3 Phase 3: AST generation 3

2 Requirements **3**

cuu duong than cong . com

cuu duong than cong . com

CuuDuongThanCong.com https://fb.com/tailieudientucntt

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.
- use Scala to implement a visitor to generate an AST for a program.

1 Specification

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

cuu duong than cong . com

- read carefully the specification of BKOOL language
- Download and unzip file assignment1.zip, which contains Main.scala, BKOOL.g4 and some other script for your convience.
- Modify BKOOL.g4. **Please fill in your id in the headers of these files**.

This assignment is divided three phases: lexer phase, recognizer phase and AST generation 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:

cuu duong than cong . com

- Modify BKOOL.g4 to detect tokens in BKOOL language.
- 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.

CuuDuongThanCong.com https://fb.com/tailieudientucntt

  - "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 BKOOL. To complete this phase, you need to:

- Modify BKOOL.g4.
- You can assume that there is at most one error in each test case.

1.3 Phase 3: AST generation

After the deadline of the phase 2, the AST files together with some other files will be uploaded so you should download then. To complete this phase, you need to:

- Modify ASTGeneration.scala.
- You can assume that there is no error in each test case.

cuu duong than cong . com

2 Requirements

The operating system when cheking your submission is Windows. Make sure that your program can be compiled and run in the environment.
Note that you must NOT compress your files when submit them. You must submit the following files:

- Phase 1:
  - BKOOL.g4: the lexer
  - The deadline of phase 1 is 16:00 18th Sep. 2015.
- Phase 2:
  - BKOOL.g4: the recogniser
  - The deadline of phase 2 is 16:00 25th Sep. 2015.
  - Lexer is also checked in this phase.
- Phase 3:

CuuDuongThanCong.com https://fb.com/tailieudientucntt

  - BKOOL.g4 and ASTGeneration.scala
  - The deadline of phase 3 is 16:00 2nd Oct. 2015.
  - Recogniser and Lexer are also checked in this phase.
- Result = min(Lexer1,Lexer2,Lexer3) + min(Recogniser2,Recogniser3) + ASTGeneration

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.

cuu duong than cong . com

cuu duong than cong . com

CuuDuongThanCong.com https://fb.com/tailieudientucntt