CSC 3310 Concepts of Programming Languages

Fall 2017

Assignment Name	Python Programming Assignment
Due Date	November 17th, 2017
Delivery Method	Through Canvas Only
Points	10% of final grade

Motivation

Python has become a popular language, several universities use it as a first entry programming language, and Python is the programming language used in many popular open source software. In addition to the above-mentioned uses, bioinformatics and data science projects use Python as the preferred tool.

This project consists in the development of the front end of the front end of a compiler. By programming the Lexical Analyzer (Lexer) for the Mini – Power grammar you will gain further understanding of the lexical analysis and the production of tokens needed for the Syntax Analyzer (Parser).

Description

Write a program in Python that takes a program written in Mini – Power, and outputs the tokens and lexemes into a new file.

The program should run like this:

```
prompt> python lexer.py input.txt
Processing input file input.txt
14 tokens produced
Result in file input.out
prompt>
```

The tokens in the grammar are:

- SEMICOLON
- PRINT
- ID
- PLUS
- MINUS
- TIMES
- DIV
- POWER
- ASSIGN
- INT_CONST
- REAL_CONST
- STRING

Given the following program written in this language:

```
a% = 5.5;
b# = 3;
c% = a% + b#;
PRINT “The result is”;
PRINT c%
```

The output file should contain the following information:

```
ID    a     REAL
ASSIGN
REAL_CONST  5.5
SEMICOLON
ID    b     INTEGER
ASSIGN
INT_CONST   3
SEMICOLON
ID    c     REAL
ASSIGN
ID    a     REAL
PLUS
ID    b     INTEGER
SEMICOLON
PRINT
STRING      The result is
SEMICOLON
PRINT
ID    c     REAL
```

Notice that the ID tokens have the partial lexeme associated, along with the type (from the suffix of the ID); the constant values have their value associated too.

Assignment Requirements

- Good programming practices
    - Indentation (required by python)
    - Meaningful identifier naming
    - Consistent variable naming convention
- This is an strictly individual assignment

Delivery Method

- Files to be uploaded
    - lexer.py [You MUST name your program this, failure will result in zero grade]
    - test01.txt [Test Program written in Mini – Power]
    - test02.txt [Test Program written in Mini – Power]
- Uploaded in Canvas

Assessment and Grading

Assessment will consider the following factors in the grading of the project:

- Good programming practices
- Adherence to instructions
- Correct function of the program
- No runtime errors
- Late deliveries will have a zero mark
- Plagiarism will have a double zero mark (in addition to losing 15% of your final grade, the group that plagiarizes will lose an additional 15% of their final grade), besides there will be a report filed in the students' academic record.

Extra Challenge (14 points, 4 bonus points)

Instead of using the Mini – Power grammar use the Mini – Pascal grammar. You will need to identify the tokens to construct the lexer.