# Department of Computing

**CS 354: Compiler Construction**

**Class:** BSCS-2BC

# Lab [04]: Lexical Analysis with JFlex

**Date:** 10th,12th Oct, 2017

**Time:** [2:00pm – 4:50pm]/[2:00pm – 4:50pm]

# 

# Instructor: Saqib Nazir

**Lab Engineer**: Khurram Altaf

# Lab [04]: Lexical Analysis with JFlex

**Introduction**

The lexical analyzer is the part of the compiler that reads the source text, it may also perform certain secondary tasks at the user interface. One such task is stripping out comments and white space in the form of blanks, tabs and new line characters, from the source program. Another is correlating error messages from the compiler with the source program i.e. keeping a correspondence between errors and source line numbers.

**Objectives**

1. Successful understanding/implementation of basic Lexical Analysis using Jflex

**Tools/Software Requirement**

1. Jflex on Linux or Windows platform

**Description**

Lexical analysis is the process of converting a sequence of characters into a sequence of [tokens](http://en.wikipedia.org/wiki/Lexical_analysis#Token). A program or function which performs lexical analysis is called a lexical analyzer, lexer or scanner. A lexer often exists as a single function which is called by a [parser](http://en.wikipedia.org/wiki/Parser) or another function.

**Lab Tasks**

* **JFlex in a Nutshell (tutorial):** Go through the Jflex tutorial provided on LMS and try-out the basic examples in Jflex.
* Write a Jflex program to process TINY language with the following specifications:
  + Match integers constants.
  + Match Identifiers.
  + Operators: \*
  + Keywords: write
  + Print unrecognized characters

Use the following example code to test your lexical analyzer.

|  |
| --- |
| length := 4  breadth := 3  area := length \* breadth  write area |

Your output should resemble:

|  |
| --- |
| An identifier: length  Unrecognized character: :  Unrecognized character: =  Integer: 4  An identifier: breadth  Unrecognized character: :  Unrecognized character: =  Integer: 3  An identifier: area  Unrecognized character: :  Unrecognized character: =  An identifier: length  Operator: MULTIPLY  An identifier: breadth  Keyword: write  An identifier: area |

* **Postfix formula evaluation:** Given an input text containing non-negative integers and three operator i.e. +, - and \*, evaluate the given postfix formula using Jflex based lexical analyzer. For example given the following input:

**44 33 22 \* + 1 -**

Your output should resemble:

|  |
| --- |
| 44 0 0 0  33 44 0 0  22 33 44 0  726 44 0 0  770 0 0 0  1 770 0 0  769 0 0 0  result = 769 |

**Deliverables**

You are required to upload your task (Sources & PDF document) using the link created on LMS followed by a viva.