CSL302: Compiler Design

Hands on Session on Lexical Analyzer Tool

Date: 7-Aug-2025

Instructions:

1. Prepare a zip file and upload your solutions on canvas
2. Include a readme file
3. Submit the solutions by the end of tomorrow

Task-0: Lex tool set up

If you are using ubuntu, you can install lex using the below command

sudo apt install flex

Task-1: Execute the demo files

Download the demo lex files discussed in the class from canvas (corresponding to lex tutorial). Try to compile the lex files and check if you are getting the expected output after executing them.

You can use the following steps for compiling and running

1. lex <filename>.l
2. gcc lex.yy.c
3. ./a.out

Task-3: Writing Custom Lexical Analyzer

Write a lexical analyzer that reads the character stream from the user and it prints the tokens as per the following.

Type of text	**Token**
Name	NAME
Phone number	PHONE_NUMBER
Email Address	EMAIL
Credit Card Number	CREDIT_CARD

The following are the constraints for the tokens to be recognized.

Name: Contains First name followed by Last name (with space separated). First and Last names start with capital letters followed by any number of lower letter alphabets.

Phone number: It has 10 digits, in the format as (123) 456–8911 i.e (3 digits)(space)(3 digits)(hyphen)(4 digits).

Email Address: It has the format username@Domain.
The username should start with an alphabet (lower/upper) followed by any number of alphabet or digits. Domain should use <string>.<string>.<string>, where strings contain only lower Alphabets.

Credit Card Number: It has 16 digits in the below format <4 Digits><space><4 Digits><space><4 Digits><space><4 Digits>

Task-4: Implementing wc with Lex:

The linux command "**wc"** takes a text file as input and outputs the number of lines, words, and characters in the file, in the following manner.

$wc myfile.txt

39 85 570 myfile.txt

Write a lexical analyzer (in lex) that implements the above wc command. Your program should print the output as explained above.