Indian Institute of Technology Roorkee

Department of Computer Science and Engineering

CSN-362: Compiler Laboratory (Spring 2024-2025)

Lab Assignment-2 (L2) **Date: 28 Jan 2025** **Duration: 3 hrs**

Problem Statement 1:

Write a C/C++ program to recognize strings as per following regular expressions (REs): 'a*', 'a*b+', 'abb', but not anything else.

Test Cases.

a) Enter a String: aaaabbbbb //2 Marks output: aaaabbbbb is accepted under rule 'a*b+'
b) Enter a string: cdgs // 2 Marks output: cdgs is not recognized
c) Enter a string: aaaaa // 2 Marks output: aaaaa is accepted under rule ‘a’
d) Enter a string: aabaaa // 2 Marks output: aabaaa is not recognized
e) Enter a string: aaaaac // 2 Marks output: aaaaac is not recognized

Submission folder P1 should contains:

1. Source code file
2. Snapshot image files for at least 4 testcases after running the code on a given set of REs by the TA.

Problem Statement 2:

Lexical Analysis is the first phase of compiler also known as scanner. It converts the input program into a sequence of Tokens. A C program consists of various tokens and a token is either a keyword, an identifier, a constant, a string literal, or a symbol. Write a C/C++ program to print all the keywords, literals, valid identifiers, invalid identifiers, integer number, real number in a given C program.

Testcase:

Any simple C code containing the following:
(1) Keywords: Examples- for, while, if, etc.
(2) Identifier: Examples- Variable name, function name, etc.
(3) Operators: Examples- '+', '++', '-', etc.
(4) Separators: Examples- ', ' ';' etc.

For one C program statement:

```
char str[100] = "int a = b + 1c; ";
```

Output:

'int' IS A KEYWORD
'a' IS A VALID IDENTIFIER
'=' IS AN OPERATOR
'b' IS A VALID IDENTIFIER
'+' IS AN OPERATOR
'1c' IS NOT A VALID IDENTIFIER

Submission folder P2 should contains:

1. Source code file
2. Input C program file having all the types mentioned in the question (.c file)
3. Snapshot image file showing the outputs on the console after running the code on your C program file.