

Department of Computer Science and Engineering

Indian Institute of Technology, Kharagpur

Compiler Theory: CS31003

3rd year CSE, 5th Semester

Assignment - 2: Lexical Analysis

Assign Date: September 14, 2020

Mode: Individual Submission

Marks: 10

Submit Date: 23:55, September 22, 2020

1. Write regular expressions for the following languages: $4 \times 1 = 4$
 - (a) All strings of lowercase letters that contain the five vowels in order (they need not be together).
 - (b) All strings of a's and b's with an even number of a's and an odd number of b's.
 - (c) All strings of a's and b's that do not contain the substring 'abb'.
 - (d) All strings of a's and b's that do not contain the subsequence 'abb'.
2. Write Regular Expression Declarations and Translation rules in order to tokenize a minimal C function. You only need to generate tokens that the below example requires, that is, KEYWORD, IDENTIFIER, OPERATOR, SPECIAL_CHARACTER. Only declarations and rules are required. 4

```
float Function2Calculate(int a, double b, float c)
{
    b = b ^ c;
    a += b;
    c /= a;
    return c;
}
```

Give the list of tokens generated for this function. 2

3. Consider the below python function. Think about which features of the lexer you have learnt will be useful in creating tokens that python interpreter can use. (Bonus - No marks.)

```
def func(a, b, c):
    c = a + b
    return a + b + c
```

Note: Please submit your answer handwritten in paper. Upload your answer in .pdf format in the moodle server. File name should be named as ass2-roll.pdf, where "roll" is your roll number.