**CS323 Project 1:Lexical Analysis**

**1.** **Problem Statement**

Write a Lexical Analyzer that would be written in the language of your choice. The Analyzer would read in the file written in the Rat20F language and would output the Tokens and corresponding Lexeme values to an output file.

**2.** **How to use your program**

Step1) Click on lexer.exe

Step2) Enter the input file you want to open. (Ex. If you want to open “test.txt”, type in “test”. **No .txt is needed**) The input file has to be in the same directory with lexer.exe. Program would output an output file titled with output followed by the input file that was specified.

Step3) Either type in another file that you would want to open or type in “Quit” to exit the program.

**3.** **Design of your program**

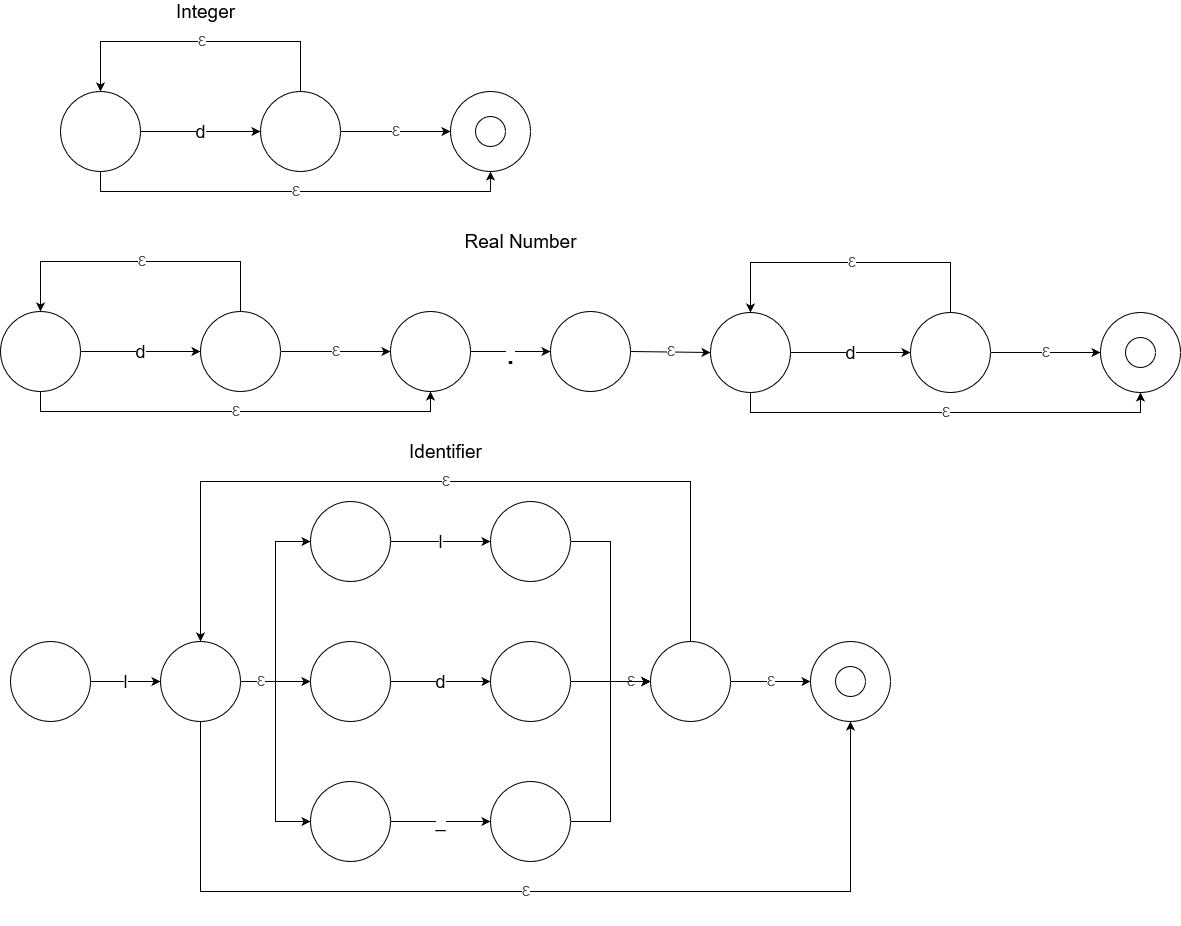
**vector<string> would be utilized to organize the tokens and Lexeme**

**Regular Expressions(RE):**

**Integer: d+**

**Real: d+.d+**

**Identifier:l(l|d|\_)\***

**Thompson Construct**

Created two FSM:

One for determining the Difference between Integer or Real Numbers when the first char is a digit.

One for identifiers when the first char is either an upper or lowercase letter. Anything that follows the first char must either be a letter, digit or \_.

Finite State Machine for Integer or Real Numbers

int FSM[5][3] = { //Accepting States is 4

0, 'i', '.', //Finite State Machine Table

1, 2, 0, //Starting State

2, 2, 3, // Integer

3, 4, 0, // Period

4, 4, 0 }; // Digit with Period Already

//Finite State Machine for identifier

int FSM[5][4] = {

0, 'i', 'l', '\_',

1, 0, 2, 0, //Starting State

2, 3, 2, 4, //Upper/Lower Case Letter

3, 2, 3, 4, //Integer

4, 2, 2, 4 }; // '\_'

**4.** **Any Limitation**

*None*

**5.** **Any shortcomings**

*None*