**CS323 Documentation**

1. **Problem Statement**

To design a lexical analyzer using FSMs. The program will read a file containing source code and return a token and the appropriate lexeme when it is needed.

1. **How to use your program**
   1. Download the zipped folder onto a computer running OS X. (The program runs on Computer D6 in room 200)
   2. Unzip the file into a known directory.
   3. add any extra test cases into the same directory.
   4. copy down the file path for the directory.
   5. double click on the executable labeled “Rat18F-Compiler”.
   6. When prompted enter in the file path of the directory all the files are currently in.
   7. To run the given test cases enter a 1, 2, or 3 followed by the enter key.
   8. To run a user supplied test case enter the filename including the file extension followed by the enter key.
   9. once you see “Lexicon Analyzer Complete” show up on your screen the Analyzer is complete
      1. If you see “Unable to open specified file.” there is either a problem with the file or there is an error with the file path you provided.
         1. Check to make sure you copied the file directory path correctly, and you spelled the file name correctly.
   10. In the same directory as the test files there will be a new text document called output\_”file name”.txt. This document will have the table of tokens and lexers from the supplied test case.
   11. repeat steps E through J for all test cases.
2. **Design of your program**

REs

Let D = { 1, 2, 3, 4, 5, 6, 7, 8, 9, 0 }

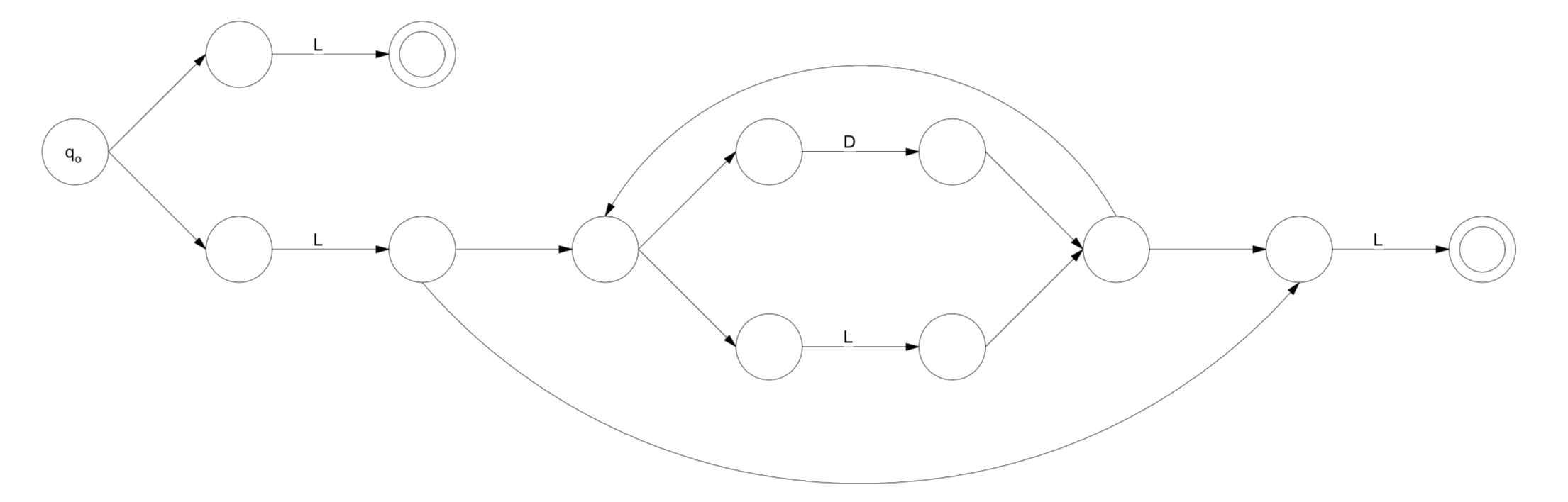
L = { a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z }

Identifier: (( L (L | D) \* L ) | L)

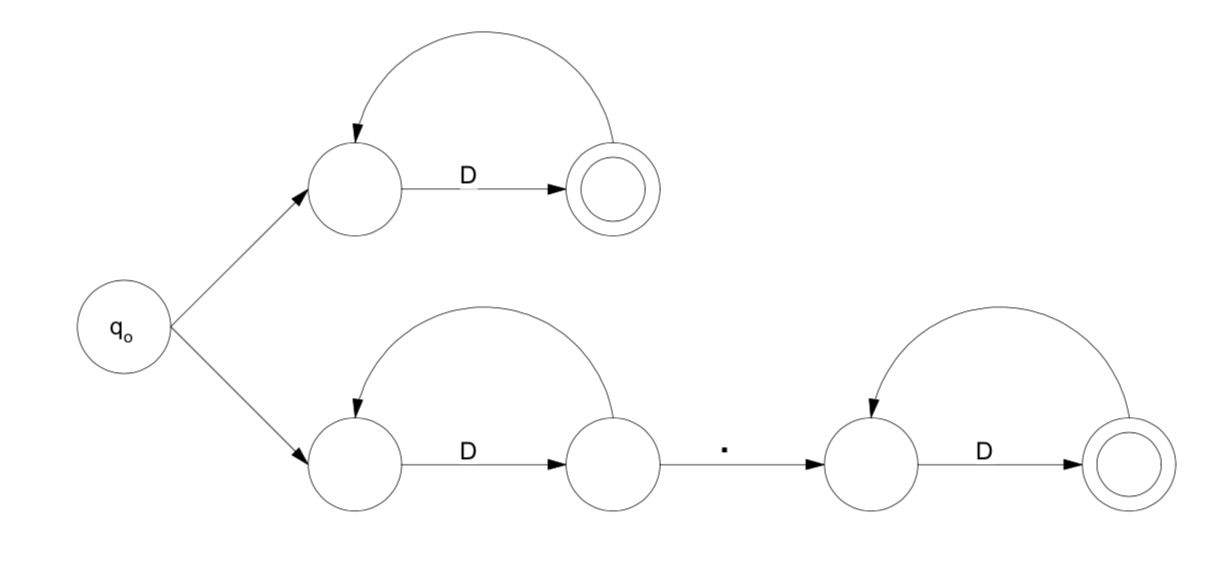
Integer: D+

Reals: D+ . D+

NFSM using Thompson method

Identifier:

Integer/Real:



In order to choose what DFSM to check we utilize If Else statements.

Data Structures:

Tuple: Used to return one variable with both the lexeme and token.

1. **Any Limitation**

The Lexeme might not fit all on one line if the length is longer than the window. It will drop down onto the next line if that is the case.

1. **Any shortcomings**

None.