# Programming a DFA

Ximena Silva Bárcena
A01785518

## Program Overview

This program is a lexical analyzer written in Python, which reads arithmetic expressions from a file and classifies each token into categories such as:

- Numbers
  - Integer
  - Float
  - Scientific notation
- Operator ( =, +, -, *, /, ^)
- Identifiers (variables)
- Special symbols
- Comments

It uses Deterministic Finite Automata DFA logic for classification.

## How to run the program

Requirements:
- Python 3.x installed
- The program file: final.py
- Input file: expressions.txt

Install Required Libraries:
- Python standard libraries are used

Run the Program
  Open a terminal and navigate to the folder containing both files.
- Run the script with:

  python final.py

## Input Format
  The input should be placed in a file named expressions.txt

## Output Format
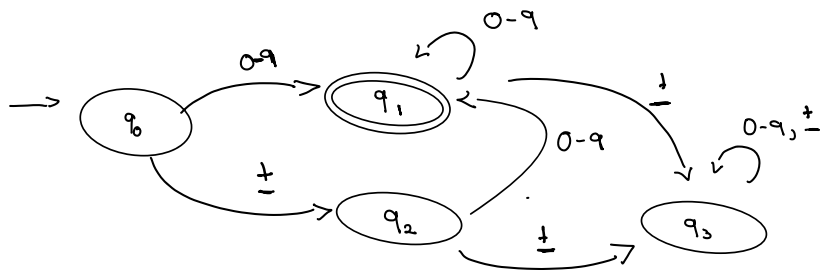  The program prints the results to the console.
  Each line of output has two columns:
     Token - the substring from the expression
     Type - the classification of the token

# Numbers

## Integer
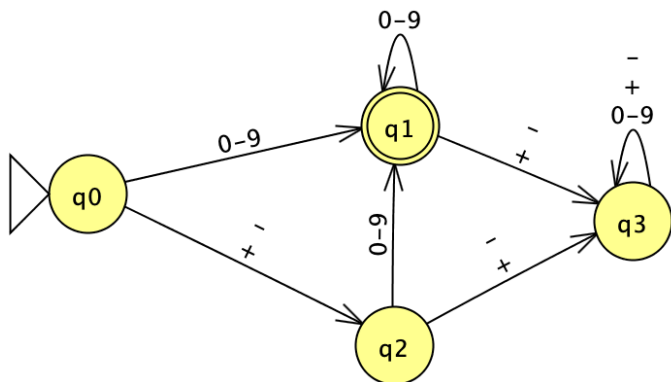


$Q = \{q_0, q_1, q_2, q_3\}$
$\Sigma = \{0-9, -, +\}$

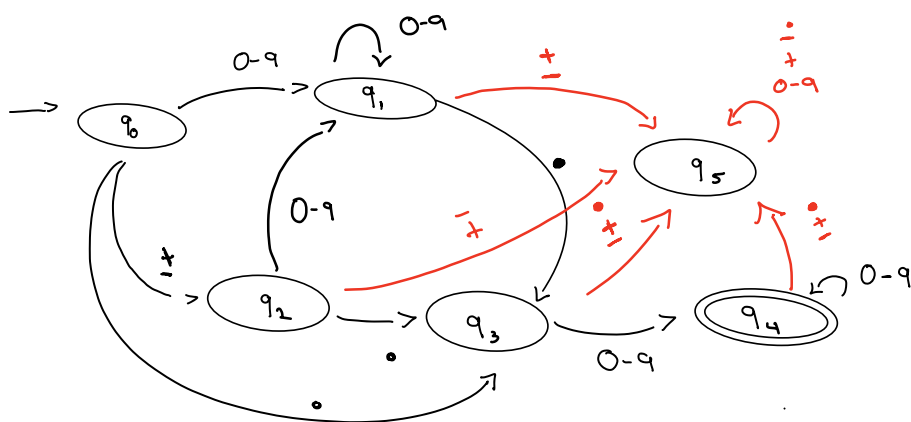$\delta =$

|  | 0-9 | +/- |
|---|---|---|
| $q_0$ | $q_1$ | $q_2$ |
| $q_1$ | $q_1$ | $q_3$ |
| $q_2$ | $q_1$ | $q_3$ |
| $q_3$ | $q_3$ | $q_3$ |

$q_0 \in Q$
$F = \{q_1\}$



## Floats



$Q = \{q_0, q_1, q_2, q_3, q_4, q_5\}$
$\Sigma = \{0-9, -, +, \cdot\}$

$\delta =$

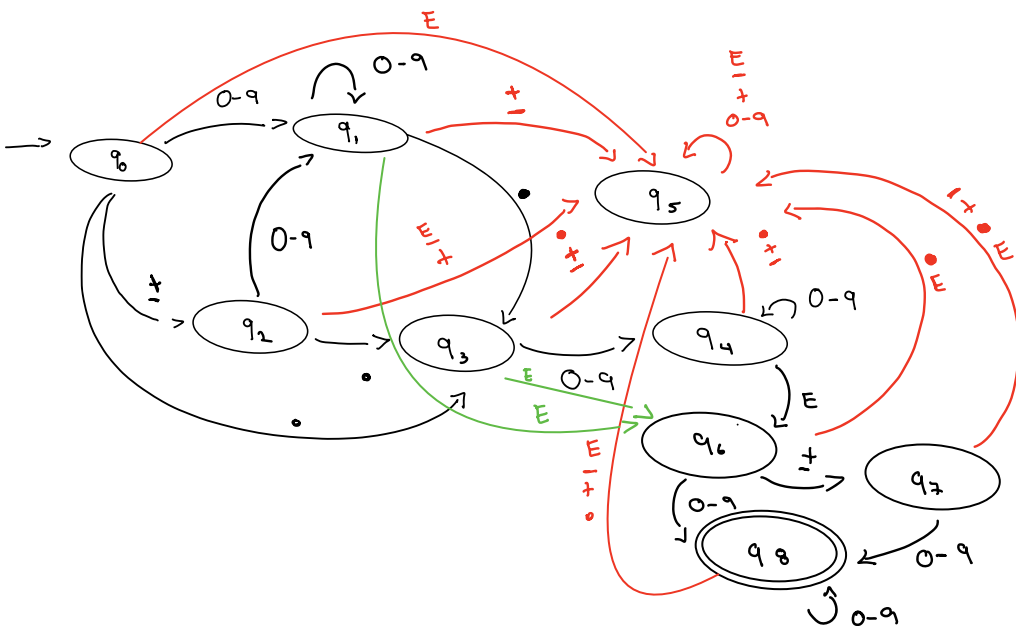|  | 0-9 | +/- | . |
|---|---|---|---|
| $q_0$ | $q_1$ | $q_2$ | $q_3$ |
| $q_1$ | $q_1$ | $q_5$ | $q_3$ |
| $q_2$ | $q_1$ | $q_5$ | $q_3$ |
| $q_3$ | $q_4$ | $q_5$ | $q_5$ |
| $q_4$ | $q_4$ | $q_5$ | $q_5$ |
| $q_5$ | $q_5$ | $q_5$ | $q_5$ |

$q_0 \in Q$
$F = \{q_4\}$
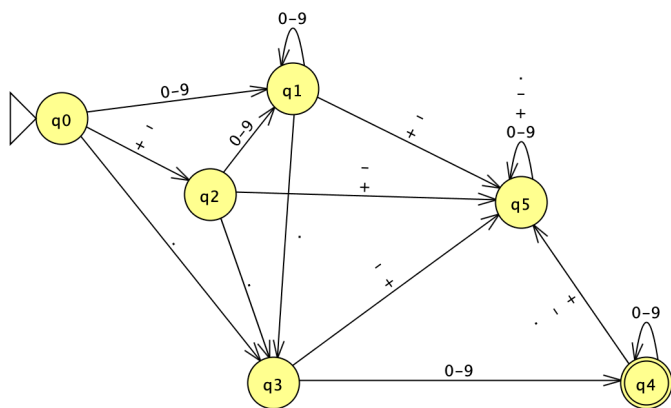
# Scientific notation



$$Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8\}$$
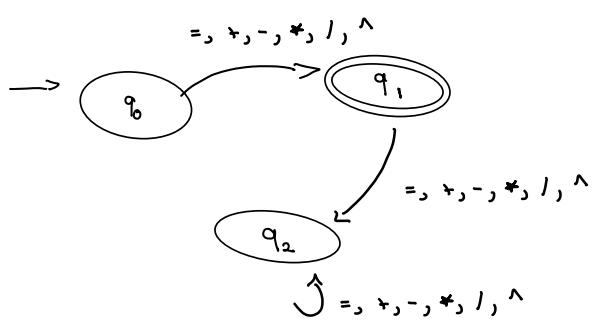$$\Sigma = \{0-9, -, +, \cdot, E\}$$

| $\delta =$ | | 0-9 | +/- | $\cdot$ | E |
|---|---|---|---|---|---|
| | $q_0$ | $q_1$ | $q_2$ | $q_3$ | $q_5$ |
| | $q_1$ | $q_1$ | $q_5$ | $q_3$ | $q_6$ |
| | $q_2$ | $q_1$ | $q_5$ | $q_3$ | $q_5$ |
| | $q_3$ | $q_4$ | $q_5$ | $q_5$ | $q_6$ |
| | $q_4$ | $q_4$ | $q_5$ | $q_3$ | $q_6$ |
| | $q_5$ | $q_5$ | $q_5$ | $q_5$ | $q_3$ |
| | $q_6$ | $q_8$ | $q_7$ | $q_5$ | $q_5$ |
| | $q_7$ | $q_8$ | $q_5$ | $q_5$ | $q_5$ |
| | $q_8$ | $q_8$ | $q_5$ | $q_5$ | $q_5$ |

$$q_0 \in Q$$
$$F = \{q_8\}$$

# Operators

$=, +, -, *, /, \wedge$

$q_0 \rightarrow q_1$

$=, +, -, *, /, \wedge$

$q_2$

$\circlearrowright =, +, -, *, /, \wedge$
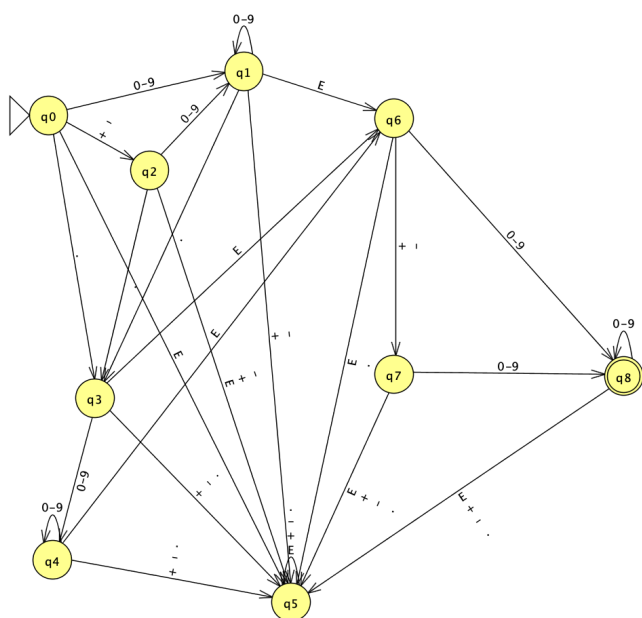
$Q = \{q_0, q_1, q_2\}$

$\Sigma = \{=, +, -, *, /, \wedge\}$

| $\delta =$ | | $=, +, -, *, /, \wedge$ |
|---|---|---|
| | $q_0$ | $q_1$ |
| | $q_1$ | $q_2$ |
| | $q_2$ | $q_2$ |

$q_0 \in Q$

$F = \{q_1\}$



# Identifiers

$a-z$

$a-z$

$q_0 \rightarrow q_1$ — $0-9$

$0-9$
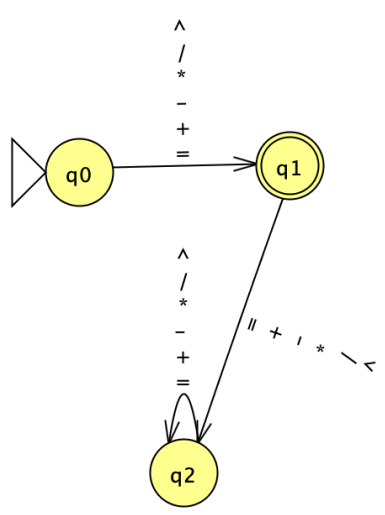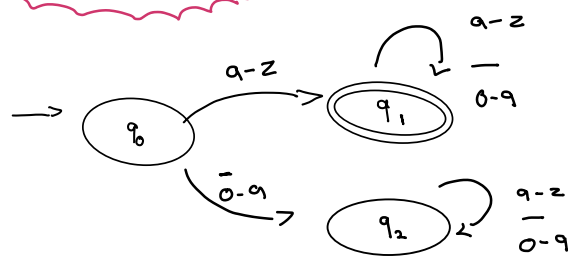
$q_2$ $a-z$ $0-9$

$Q = \{q_0, q_1, q_2\}$

$\Sigma = \{a-z, -, 0-9\}$

| $\delta =$ | | $a-z$ | $-10 \cdot 9$ |
|---|---|---|---|
| | $q_0$ | $q_1$ | $q_2$ |
| | $q_1$ | $q_1$ | $q_1$ |
| | $q_2$ | $q_2$ | $q_2$ |

$q_0 \in Q$

$F = \{q_1\}$

$Q = \{q_0, q_1\}$

$\Sigma = \{(,)\}$



$$\delta = \begin{array}{c|c} & () \\ \hline q_0 & q_0 \\ q_1 & q_1 \end{array}$$
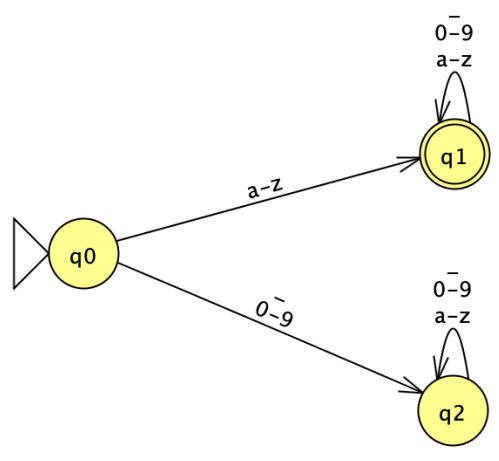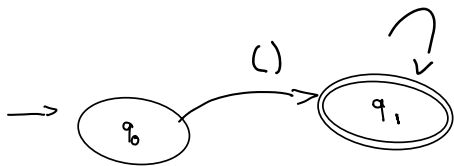
$q_0 \in Q$

$F = \{q_1\}$



Comments
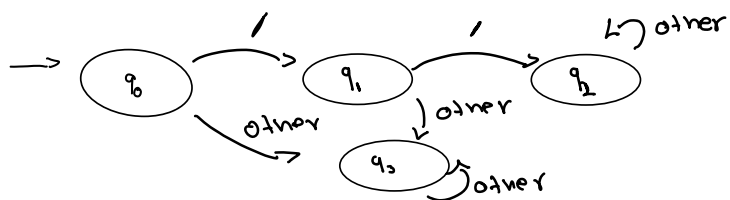
$Q = \{q_0, q_1, q_2\}$

$\Sigma = \{/, \text{other}\}$



$$\delta = \begin{array}{c|c|c} & / & \text{other} \\ \hline q_0 & q_1 & q_3 \\ q_1 & q_2 & q_3 \\ q_2 & q_2 & q_2 \end{array}$$

$q_0 \in Q$

$F = \{q_1, q_2\}$