

Análisis Léxico: Implementación de un DFA en Java

Yonathan Berith Jaramillo Ramírez

18 de septiembre de 2024

1 Introducción

Este documento describe en detalle el funcionamiento del autómata finito determinista (DFA) implementado en el lexer desarrollado en Java.

2 Definición de Tokens

Los tokens reconocidos por el lexer y sus expresiones regulares asociadas son:

- **Identificadores (ID)**: letras minúsculas, definidos por la expresión regular `[a-z]`.
- **Números Enteros (NUM_ENT)**: dígitos sin punto decimal, definidos por `[0-9]`.
- **Números Reales (NUM_REAL)**: dígitos con punto decimal, definidos por `[0-9]\.[0-9]`.
- **Espacios en Blanco (ESPACIO)**: espacios, tabulaciones y saltos de línea, definidos por `espacio`, `tab`, `newline`,
- **Operadores (OPERADOR)**: el operador de suma `+` y el operador de asignación `:=`.

2.1 Estados y Transiciones

El DFA consta de los siguientes estados:

- q_0 : Estado inicial.
- q_1 : Reconoce identificadores.
- q_2 : Reconoce números enteros.
- q_3 : Reconoce números reales.
- q_4 : Reconoce espacios en blanco.
- q_5 : Reconoce el operador `+`.
- q_6 : Estado intermedio para el operador `:=`.
- q_7 : Reconoce el operador `:=`.

3 Diagrama del Autómata

A continuación se presenta el diagrama del autómata utilizando TikZ.

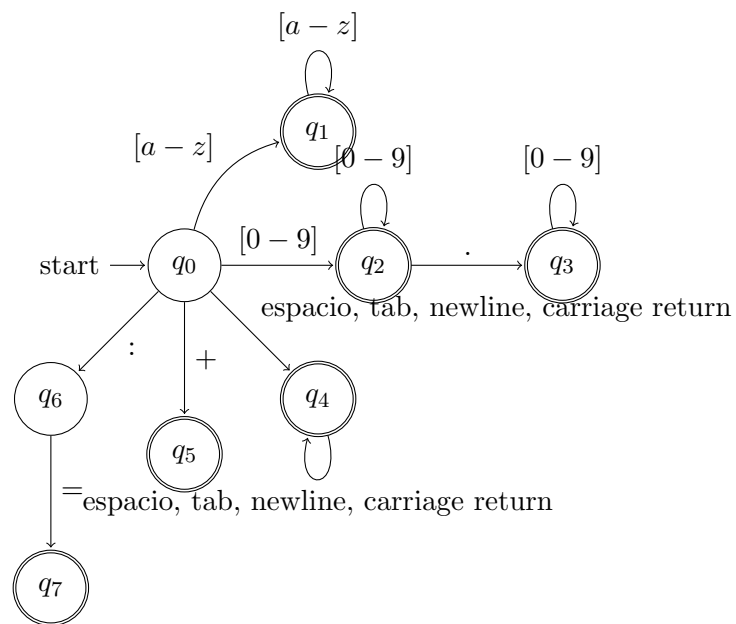


Figure 1: Diagrama del Autómata Finito Determinista