Automata Design Document

Student Name: Umar Javed & Sheharyar Ahmed

**Roll No:** 22i-1050 & 22i-1171

Assignment: 01

---

1. Regular Expressions

The following regular expressions define the 7 token types implemented in the scanner.

| Token Type	Regular Expression	Description
IDENTIFIER	`[A-Z][a-zA-Z0-9_]*`	Starts with Uppercase, followed by alphanumeric/underscore. Max length 31 (enforced in code).
INTEGER	`[+-]?[0-9]+`	Optional sign, followed by one or more digits.
FLOAT	`[+-]?[0-9]+\.[0-9]+`	Integer part, decimal point, fractional part.
STRING	`` `"([^"\] ``	`` \.)*"` ``
BOOLEAN	`true|false`	Exact literal matches.
COMMENT	`##[^\r\n]*`	Starts with `##`, consumes characters until newline.
WHITESPACE	`[ \t\n\r]+`	Spaces, tabs, and newlines (skipped by scanner).

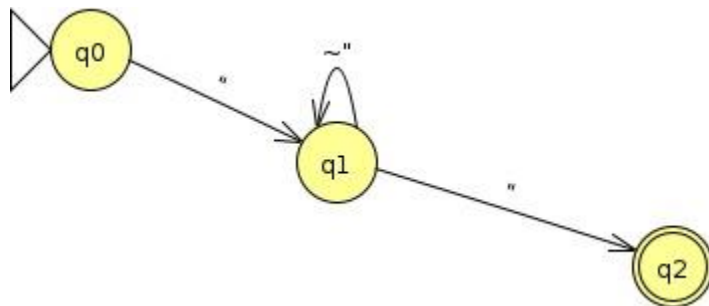# 2. NFA Diagrams

2.1 Identifier NFA

q0

A-Z

a-z, 0-9, _

q1
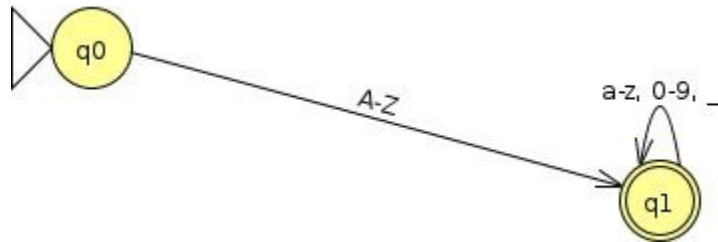
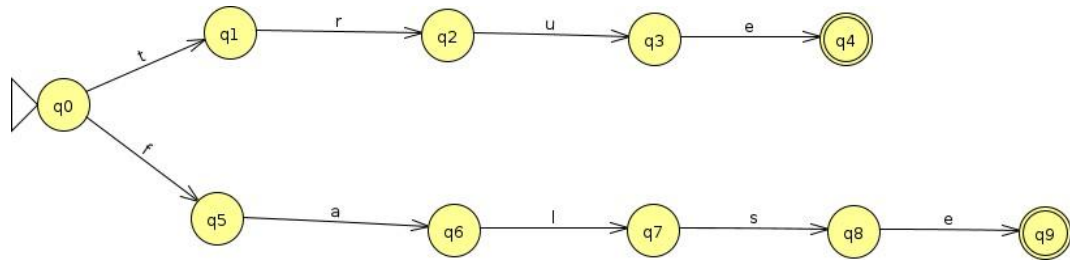## 2.2 Integer literal NFA

q0

0-9

+, -

0-9

0-9

q1

q2
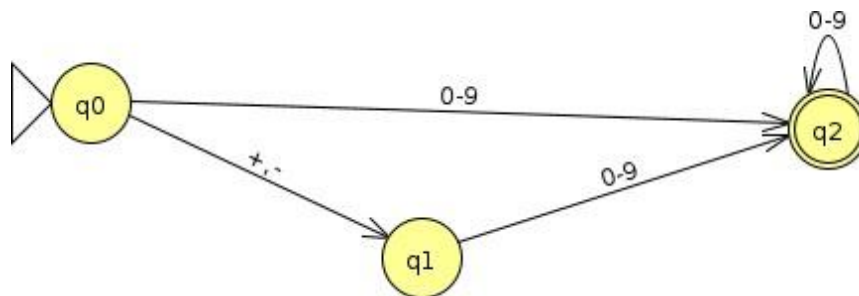
2.3 Float Literal NFA
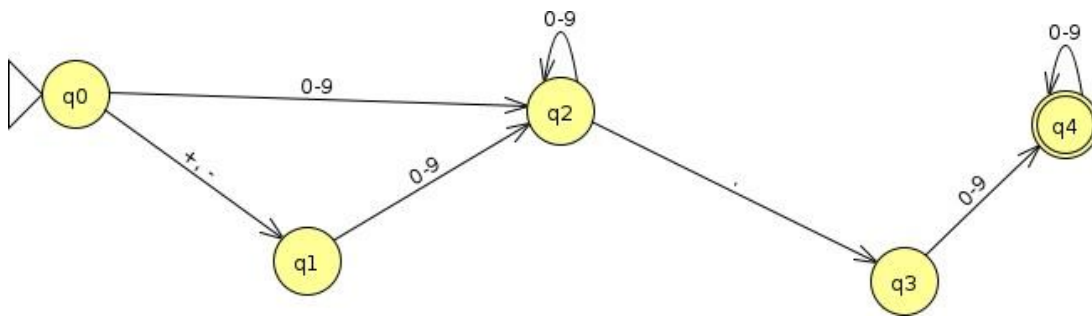
2.4 String Literal NFA

2.5 Boolean NFA

2.6 Comment NFA
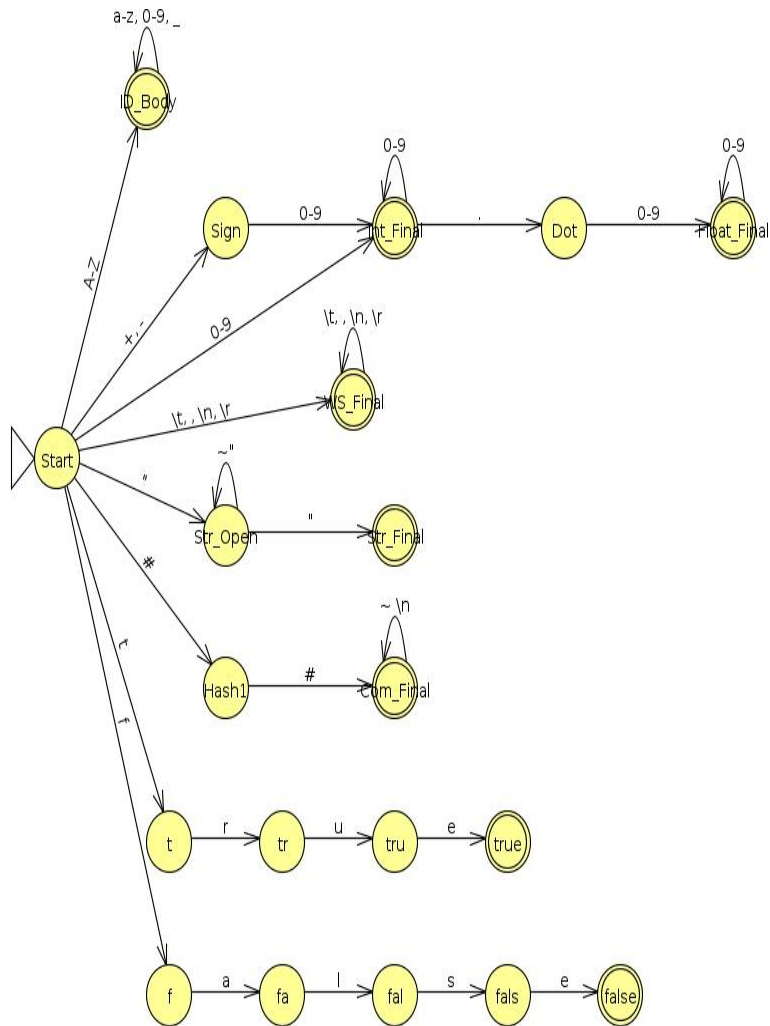
2.7 Whitespace NFA

q0

\t, ,\n,\r

\t, ,\n,\r

q1
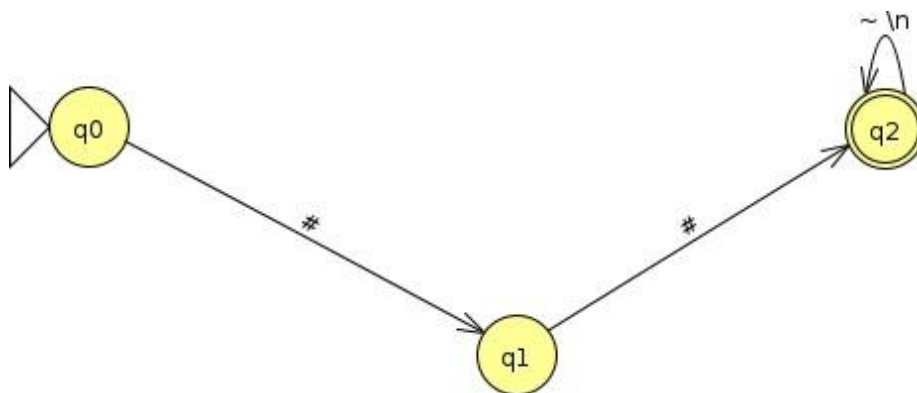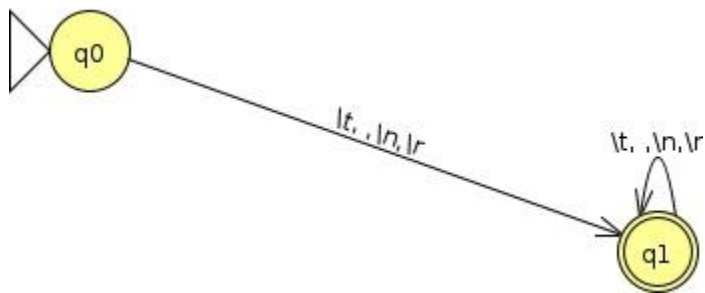
3. Minimized DFA

3.1 DFA Diagram

a-z, 0-9, _

ID_Body

0-9

0-9

Sign — 0-9 → Int_Final — . → Dot — 0-9 → Float_Final

A-z

+,-

0-9

\t, , \n, \r

WS_Final

\t, , \n, \r

Start

~"

"

Str_Open — " → Str_Final

#

~ \n

Hash1 — # → Com_Final

t

t — r → tr — u → tru — e → true

f

f — a → fa — l → fal — s → fals — e → false

3.2 Transition State Table

Current State	Input A-Z	Input a-z, 0-9	Input .	Input "	Input #	Input \n, \t	Next State / Action
0 (Start)	1	-	-	7	9	6	Determine Token Type
0 (Start)	-	2 (digit)	-	-	-	-	Start Number (Int)
0 (Start)	-	11 (t), 15 (f)	-	-	-	-	Start Boolean
1 (ID Body)	1	1	-	-	-	-	**Accept: IDENTIFIER**
2 (Int)	-	2	4	-	-	-	**Accept: INTEGER**
4 (Dot)	-	5	-	-	-	-	Transition to Float
5 (Float)	-	5	-	-	-	-	**Accept: FLOAT**
6 (WS)	-	-	-	-	-	6	**Accept: WHITESPACE**
7 (Str Open)	-	(any char)	-	8	-	-	Read String Content
8 (Str End)	-	-	-	-	-	-	**Accept: STRING**
9 (Hash 1)	-	-	-	-	10	-	Check 2nd Hash
10 (Comm)	-	(any char)	-	-	-	0	**Accept: COMMENT**
14 (true)	-	-	-	-	-	-	**Accept: BOOLEAN**
19 (false)	-	-	-	-	-	-	**Accept: BOOLEAN**