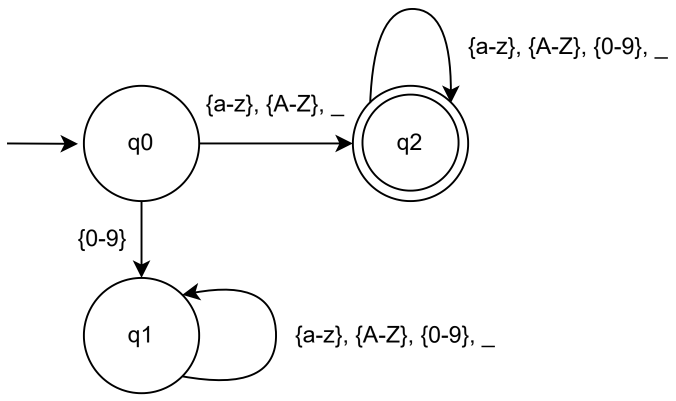
CPSC 323 Compilers and Languages Project 1 FSA Diagrams

DFA for Identifiers

RE: [a-zA-Z_][a-zA-Z0-9_]*

DFA for identifiers of any character length. In many programming languages, identifiers can only contain alphabetic letters, numbers, and underscores. Identifiers can also not start with a number. These rules are reflected by the DFA below. q1 is a dead state.

Graph



Table

	{a-z}	{A-Z}	{0-9}	_
→q0	{q2}	{q2}	{q1}	{q2}
q1	{q1}	{q1}	{q1}	{q1}
<u>q2</u>	{q2}	{q2}	{q2}	{q2}

Example Inputs

variable1_ : This input will be accepted as it satisfies the rules of an identifier. Starting from state q0, the first character 'v' will move to the accepting state q2. All subsequent characters will loop back to q2, therefore this input is accepted.

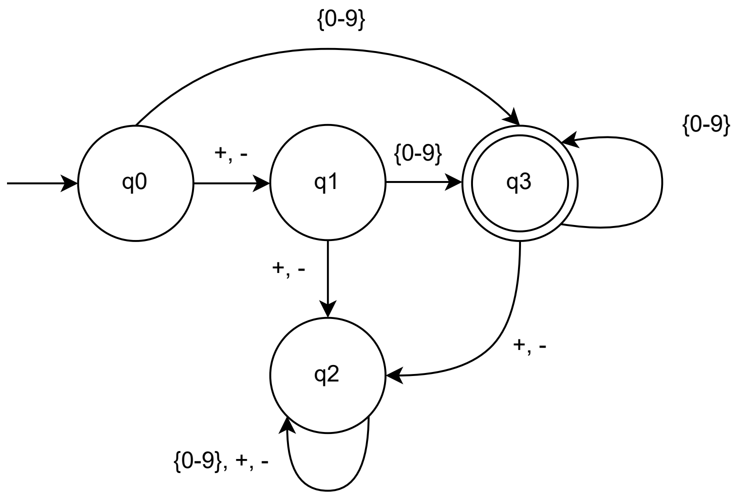
1variable_: This input will not be accepted as it breaks the rules of an identifier by starting with a number. Starting from state q0, the first character '1' will move to state q1. q1 is a dead state, so any remaining inputs will loop back to this same state. q1 is not an accepting state, therefore, this input is not accepted.

DFA for Integers

RE: [0-9+-][0-9]*

DFA that accepts integers of any length. A plus or minus sign may also be optionally added at the beginning of the input. q2 is a dead state.

Graph



Table

	{0-9}	+	-
→q0	{q3}	{q1}	{q1}
q1	{q3}	{q2}	{q2}
q2	{q2}	{q2}	{q2}
q3	{q3}	{q2}	{q2}

Example Inputs

-9: This input will be accepted. Starting from state q0, the character '-' moves to state q1. The next character '9' will move to the accepting state q3, therefore, this input is accepted..

+1+: This input will not be accepted as there is an extra plus sign after the integer. Starting from q0, the character '+' will move to state q1. The character '1' then moves to the accepting state q3. Lastly, the '+' character moves to state q2. q2 is not an accepting state, therefore this input is not accepted.