Regular Expression

- Set of strings – Algebraic Fashion
- Set of strings – language
  - Example1
    - L={ab,abb,abbb,abbbb,…..} → Regular language
    - (ab^+) → Regular Expression
  - Example2
    - R.L = {0,1,00,11}
    - R.E = (0+1+00+11)
  - Example 3
    - R.L = {0,1,00,11,000,111,0000,1111,….}
    - R.E =$(0^+ + 1^+)$

Regular Expression

Regular Expression

- Any terminals (a,b,c,….,^)
- Union of two R.E → R.E (R1,R2 →R1+R2)
- Concatenation of two R.E → R.E (R1,R2 → R1.R2)
- Iteration of R.E is an R.E (R →R*) (a*→^,a,aa,aaa,aaaa,…)

- Examples

Sets	Regular Language
{0,1,2}	R=0+1+2
{^,ab}	R=^+ab
{abb,a,b,bba,……}	R=a*b*
{^,0,00,000,0000,..}	R = 0*
{1,11,111,1111,….}	R = 1^+

Regular Expression - Examples

- Set of strings over {0,1} that end in 3 consecutives 1's
  - R.L = {111,0111,1111,00111,10111,….}
  - R.E = (0+1)*111
- Set of strings over {a,b} that has atleast 1 a → (>=1a)
  - R.L = {a,ba,ab,aa,aaa,aba,bbbba,bba,bbbaa,bbaaa,…..}
  - R.E = (a+b)* a (a+b)*
- Set of strings over {a,b} that has atmost 1 a → (<=1a)
  - R.L = {a,ba,ab,bba,abb,bbba,abbb,…..}
  - b* a b*
- R.L ={c,cc,ccc,cccc,…..}→ R.E = (c^+)

Regular Expression - Examples

- Set of string over {a,b} which has atleast 1 a →(a+b)*a(a+b)*
- Set of string over {a,b} which has atmost 1 a →b*ab*
- Set of strings over {0,1} which starts with 0 and ends with 1
  - R.L = {01,001,011,0111,01101,…..}
  - R.E = 0(0+1)*1
- Set of strings over {0,1} which has consecutives 11 in it
  - R.L = {11,011,0110,1110,01101,….}
  - R.E = (0+1)*11(0+1)*
- Set of Strings over {0,1} which doesn’t contain a substring 110
  - R.L = {0,1,001,0111,0011,01010,….}
  - R.E = (0+10)*1*

Identities of a Regular Expression

- Identities of Regular Expression:-

I_1 $\emptyset + \mathbf{R} = \mathbf{R}$

I_2 $\emptyset\mathbf{R} = \mathbf{R}\emptyset = \emptyset$

I_3 $\Lambda\mathbf{R} = \mathbf{R}\Lambda = \mathbf{R}$

I_4 $\Lambda^* = \Lambda$ and $\emptyset^* = \Lambda$

I_5 $\mathbf{R} + \mathbf{R} = \mathbf{R}$

I_6 $\mathbf{R}^*\mathbf{R}^* = \mathbf{R}^*$

I_7 $\mathbf{R}\mathbf{R}^* = \mathbf{R}^*\mathbf{R}$

I_8 $(\mathbf{R}^*)^* = \mathbf{R}^*$

I_9 $\Lambda + \mathbf{R}\mathbf{R}^* = \mathbf{R}^* = \Lambda + \mathbf{R}^*\mathbf{R}$

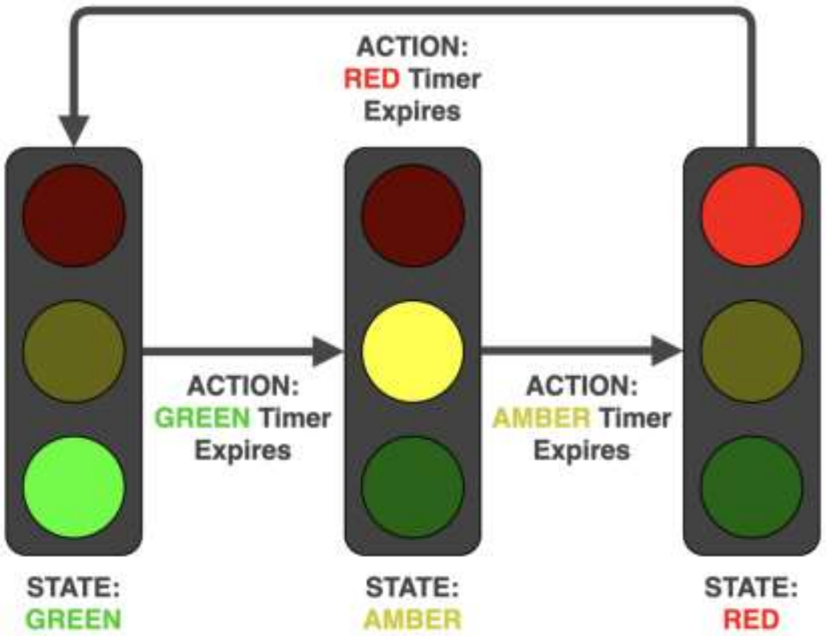
I_{10} $(\mathbf{PQ})^*\mathbf{P} = \mathbf{P}(\mathbf{QP})^*$

I_{11} $(\mathbf{P} + \mathbf{Q})^* = (\mathbf{P}^*\mathbf{Q}^*)^* = (\mathbf{P}^* + \mathbf{Q}^*)^*$

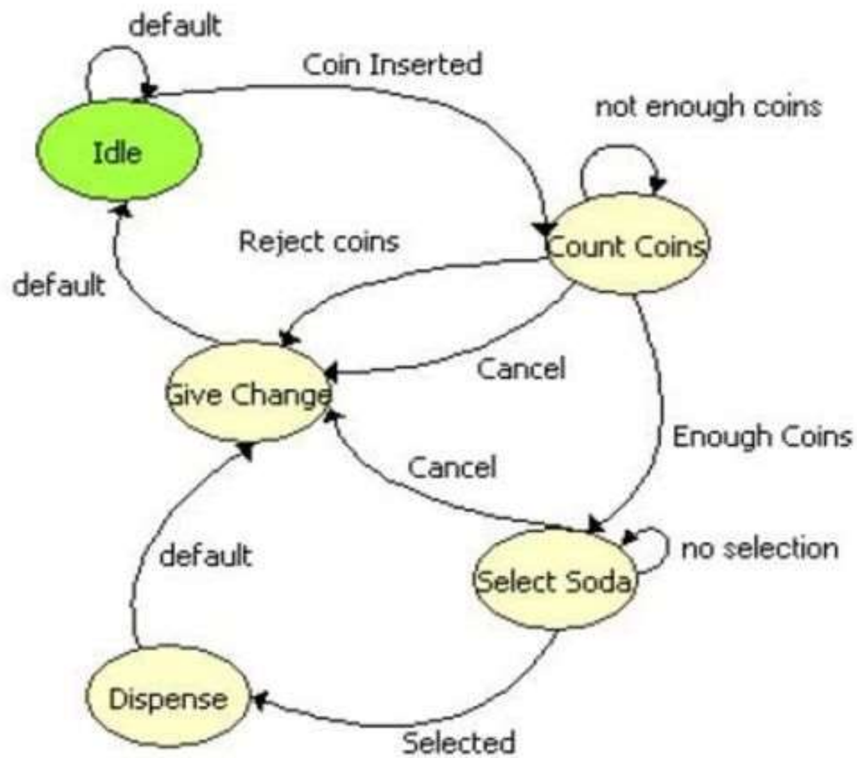
I_{12} $(\mathbf{P} + \mathbf{Q})\mathbf{R} = \mathbf{PR} + \mathbf{QR}$ and $\mathbf{R}(\mathbf{P} + \mathbf{Q}) = \mathbf{RP} + \mathbf{RQ}$

Finite State Automata

- Finite Automata – set of states and rules – transition – input
- 1State → 1 state
- Examples :Vending machine, Turnstile
  - Traffic Light



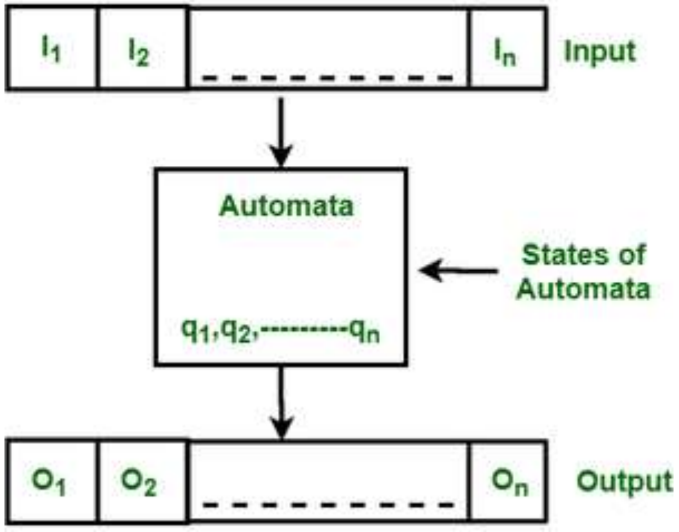
State Diagram of a Simple Soda Vending Machine



Finite State Automata

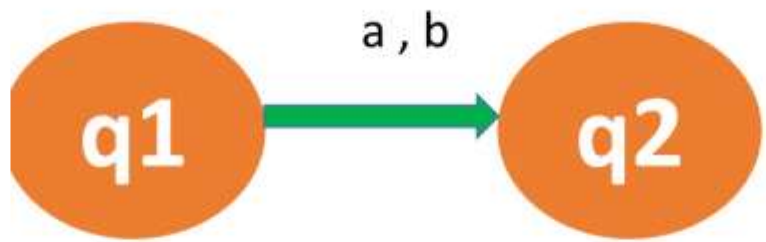
- ***FSA – Lexical Analysis of Compiler***

- FA – Tuple – $\{Q, \Sigma, q, F, \delta\}$
- Q – set of states
- Σ - set of input symbols
- q – initial state
- F – set of final states
- δ - Transitions



Regular Expression to Finite Automata

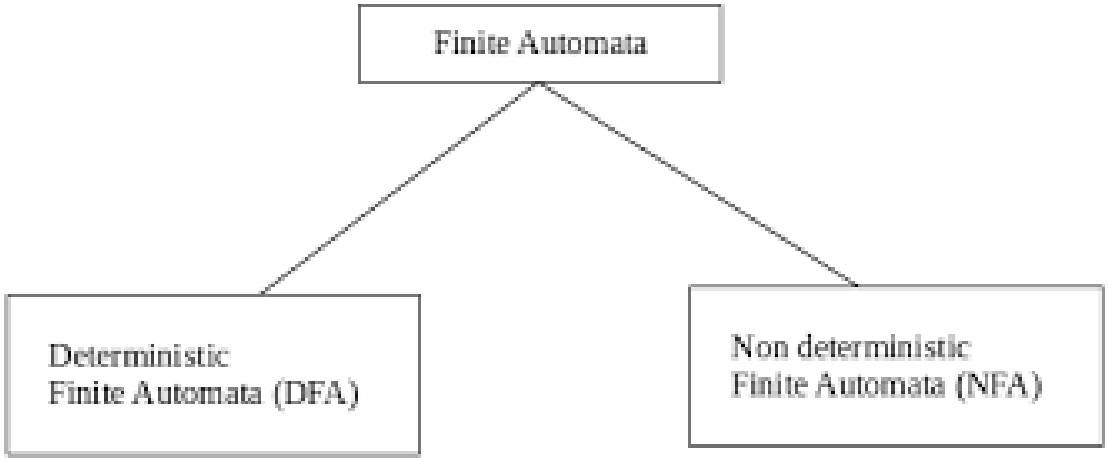
- a+b



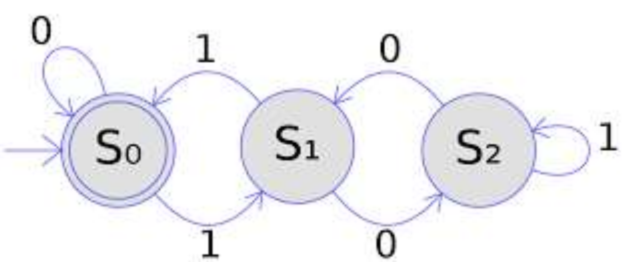
- ab

- a*

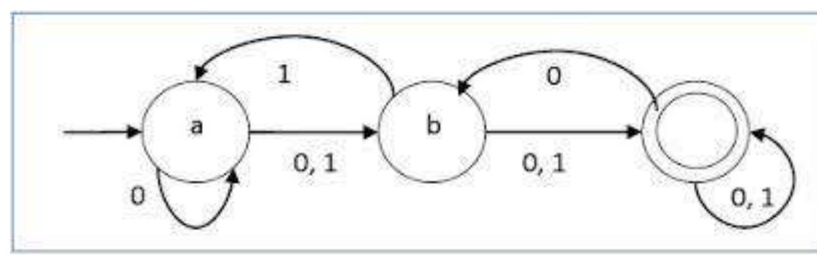
Types of Finite State Automata



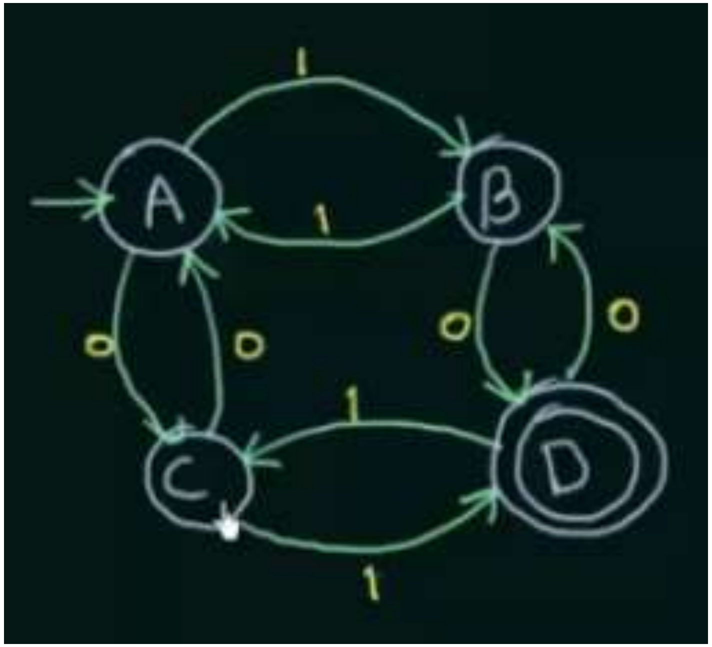
Transition → 1 state to single next state for each input symbol



Transition → 1 state to multiple next state for each input symbol



Deterministic Finite Automata (DFA)

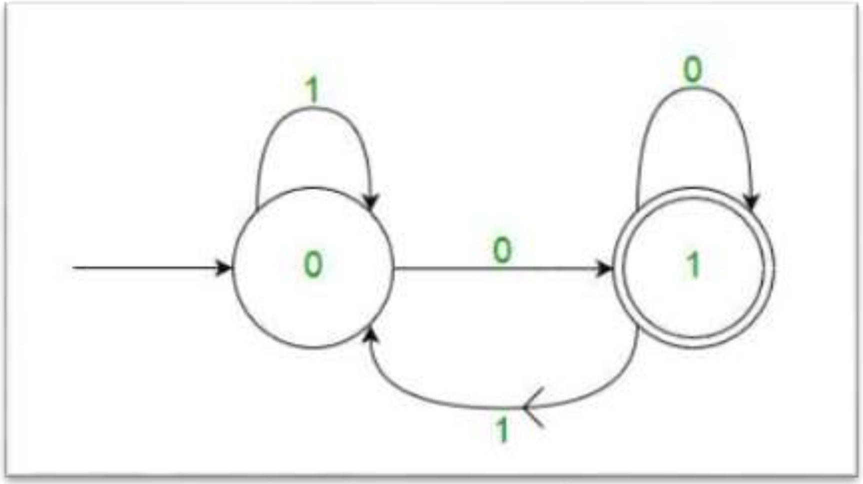


- {Q,∑, q, F, δ}
- Q = {A,B,C,D}
- ∑ = {0,1}
- Q_0 = A
- F = D
- δ → Transition function

	0	1
A	C	B
B	D	A
C	A	D
D	B	C

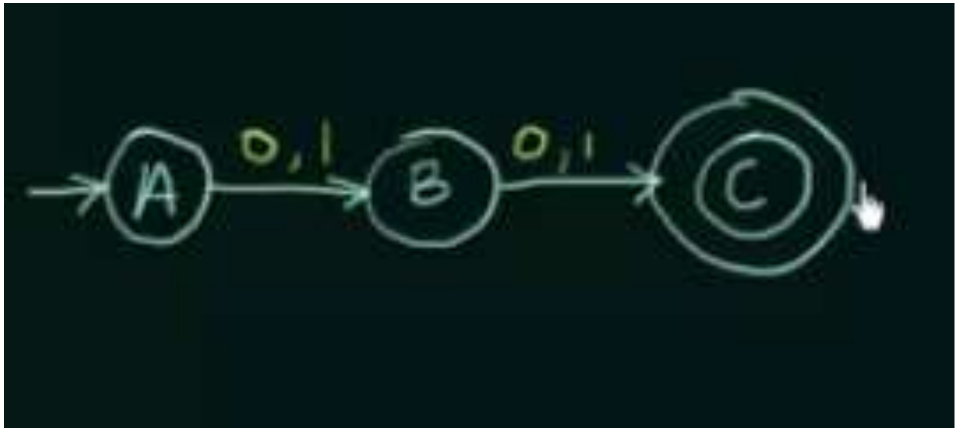
Deterministic Finite Automata (DFA)

- <u>Example 1</u>
- L1 = Set of all strings that end with ‘0’
- L1 = {000,000,010,0110,0100,01110,….}



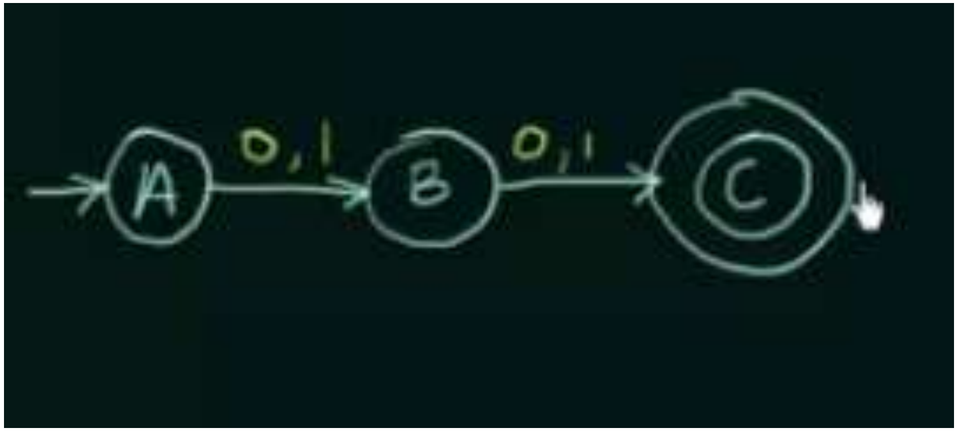
Deterministic Finite Automata (DFA)

- <u>Example 2</u>
- L1 = Set of strings over {0,1} of length 2
- L1 = {00,11,01,10}



Deterministic Finite Automata (DFA)

- <u>Example 3</u>
- L1 = Set of strings over {0,1} of length 2
- L1 = {00,11,01,10}



DFA - Applications

- Lexical Analysis – compiler
- Spelling Checker
- Search Command

DFA -Examples

- Set of strings over {0,1} that start with 0 and end with 1
- Set of strings over {a,b} that ends with bb