DFA TO RE

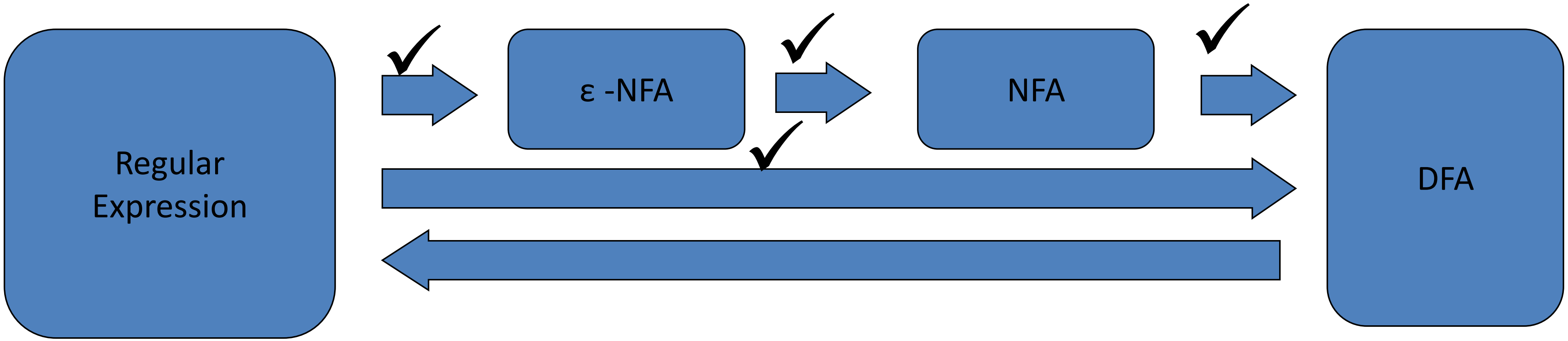
Dr. A. Beulah

AP/CSE

LEARNING OBJECTIVE

- To construct finite automata for any given pattern and find its equivalent regular expressions
  - To understand the equivalence of FA and RE

ROAD MAP



CONVERSION OF DFA TO RE

1. Regular Expression equation method - $R_{ij}^{(k)}$
2. Arden's Theorem.
3. State elimination technique.

RE EQUATION METHOD - $R_{IJ}^{(K)}$

- **Theorem**

  If L=L(M) for some DFA M= (Q, ∑, δ, S, F), then there is a regular expression r such that L= L(r).

- **Proof**

  Let L be the set accepted by the DFA

  Given a DFA M = (Q, ∑, δ, S, F), where $Q=\{q_1, q_2, \ldots, q_n\}$, i.e., $|Q| = n$.

RE EQUATION METHOD - $R_{IJ}^{(K)}$

$R_{ij}^{(K)}$ → RE describing the set of all strings x such that $\delta(q_i, x) = q_j$ going through intermediate states $\{q_1, q_2, \ldots\ldots q_K\}$ only.

- **Basis**

  K = 0 → no intermediate states.

  $R_{ij}^{(0)}$ → a set of strings which is either ε(or) single symbol.

RE EQUATION METHOD - $R_{IJ}^{(K)}$

- Case i

$R_{ij}^{0} = \{a \in \Sigma \mid \delta(q_i, a) = q_j\}$ if $i \neq j$

- Case ii

$R_{ij}^{0} = \{a \in \Sigma \mid \delta(q_i, a) = q_j\} \cup \{\varepsilon\}$ if $i = j$

R_{ij}^{k}

'n'

$k=n$

$k=0$

R_{ij}^{0} 2

$i \to j$

R_{ij}^{n}

RE EQUATION METHOD - $R_{IJ}^{(K)}$

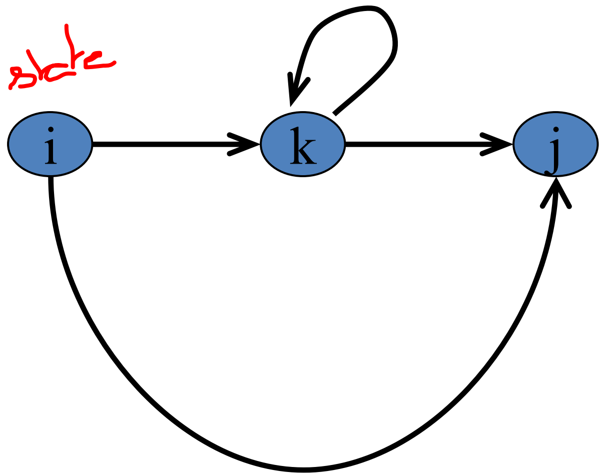
- Induction
  - It involves regular expression operations : union, concatenation and closure.
  - $R_{ij}^{k} = \underline{R_{ik}^{k-1}} (\underline{R_{kk}^{k-1}})^* \underline{R_{kj}^{k-1}} + \underline{R_{ij}^{k-1}}$

$\underline{k=1 \ldots n}$ $- k$

$\underline{k=0}$
no intermediate state

1

$k=1$ R_{ik}^{0}



RE EQUATION METHOD - $R_{IJ}^{(K)}$

- The observation of this proof is that regular expression

$L(M) = \{w \in \Sigma^* \mid \delta(q_1, w) = q_j \in F\}$

$= \cup_{qj \in F} R_{1j}^{n}$

where $R_{1j}^{(n)}$ denotes the labels of all paths from q_1 to q_j

where $F = \{q_{j1}, q_{j2}, \ldots\ldots q_{jp}\}$,

so $L(M) = R_{1j1}^{(n)} + R_{1j2}^{(n)} + \ldots\ldots R_{1jp}(n)$

IDENTITIES FOR REGULAR EXPRESSIONS

I1 $\phi + R = R$

I2 $\phi R = R\phi = \phi$

I3 $\lambda R = R\lambda = R$

I4 $\lambda^* = \lambda$

I5 $R + R = R$

I6 $R^*R^* = R^*$

I7 $RR^* = R^*R$

I8 $(R^*)^* = R^*$

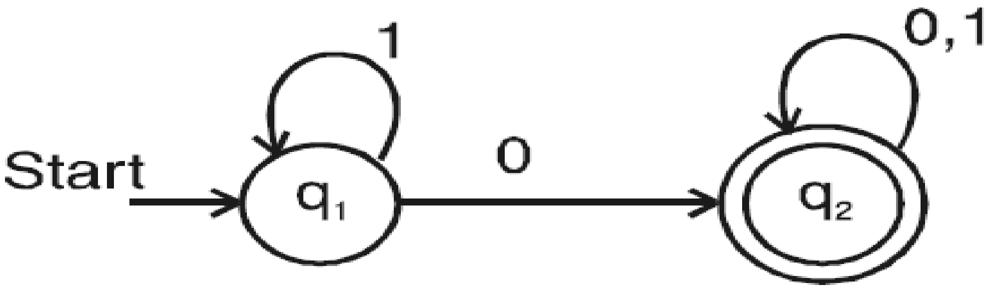
I9 $\lambda + RR^* = R^* = \lambda + R^*R$

I10 $(PQ)^*P = P(QP)^*$

I11 $(P + Q)^* = (P^*Q^*)^* = (P^* + Q^*)^*$

I12 $(P + Q)R = PR + QR$ and $R(P + Q) = RP + RQ$

EXAMPLE



K=0

$R_{ij} = a \quad i=j \to \varepsilon + a$

$R_{11}^0 = 1+\varepsilon$

$R_{21}^0 = \phi$

$R_{12}^0 = 0$

$R_{22}^0 = 0+1+\varepsilon$

K=1 $\quad R_{11}^1, R_{12}^1, R_{21}^1, R_{22}^1$

$R_{ij}^k = R_{ij}^{k-1} + R_{ik}^{k-1}\left(R_{kk}^{k-1}\right)^* R_{kj}^{k-1}$

$R_{11}^1 = R_{11}^0 + R_{11}^0\left(R_{11}^0\right)^* R_{11}^0$

$q_0\, q_1, q_2$

1, 2, 3 . . .

1 2

i j

11

12

21

22

i=1

j=1

k=1

$= (1+\varepsilon) + (1+\varepsilon)(1+\varepsilon)^*(1+\varepsilon)$

$= (1+\varepsilon) + \underline{1^*(1+\varepsilon)}$

1^*

$R_{11}^1 = (1+\varepsilon) + 1^*$

$R_{12}^1 = R_{12}^0 + R_{11}^0\left(R_{11}^0\right)^* R_{12}^0$

$= 0 + (1+\varepsilon)(1+\varepsilon)^* 0$

$= \underset{R\ P}{\varepsilon 0} + \underset{Q\ P}{1^* 0}$

$(R+Q)P$

$= (\varepsilon + 1^*) 0$

$R_{12}^1 = 1^* 0$

$(1+\varepsilon)(1+\varepsilon)^*$

$(P+Q)^*$

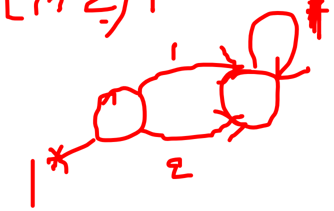
$(1^*\varepsilon^*)^*$

$(P^*Q^*)^*$

$(1^*\varepsilon)^*$

$(1^*)^*$

$(1+\varepsilon)1^*$



EXAMPLE

$R_{21}^{1} = R_{21}^{0} + R_{21}^{0} (R_{11}^{0})^{*} R_{11}^{0}$

$= \phi + \phi \cdots\cdots$

$R_{21}^{1} = \phi$

$R_{22}^{1} = R_{22}^{0} + R_{21}^{0} (R_{11}^{0})^{*} R_{12}^{0}$

$= (0+1+\varepsilon) + \phi \cdots\cdots$

$R_{22}^{1} = 0+1+\varepsilon$

$\underline{K=2}$

R_{if}^{n}

$\boxed{R_{12}^{2}}$

$R_{12}^{2} = R_{12}^{1} + R_{12}^{1} (R_{22}^{1})^{*} R_{22}^{1}$

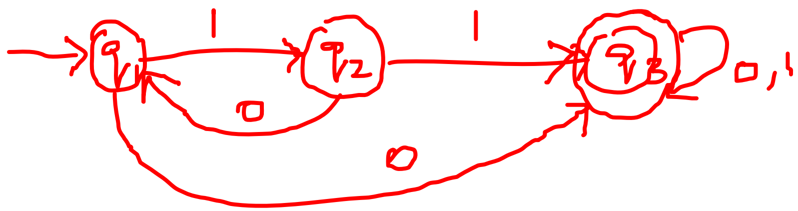
$= 1^{*}0 + 1^{*}0 \cdot \underline{(\underbrace{0+1+\varepsilon}_{R})^{*} \cdot (\underbrace{0+1+\varepsilon}_{R})}$

$= \underbrace{1^{*}0}_{R} \underbrace{+}_{P} \underbrace{1^{*}0 \cdot}_{R} \underbrace{(0+1)^{*}}_{Q}$

$= 1^{*}0 (\varepsilon + (0+1)^{*})$

$\underline{= 1^{*}0 \cdot (0+1)^{*}}$ RE

EXAMPLE



1, 2, 3,

k=0

$R_{11}^0 = \varepsilon$ $R_{21}^0 = 0$ $R_{31}^0 = \phi$ $\cancel{\phi + \varepsilon}$

$R_{12}^0 = 1$ $R_{22}^0 = \varepsilon$ $R_{32}^0 = \phi$

$R_{13}^0 = 0$ $R_{23}^0 = 1$ $R_{33}^0 = 0 + 1 + \varepsilon$

k=1

89

k=0

k=1

k=2

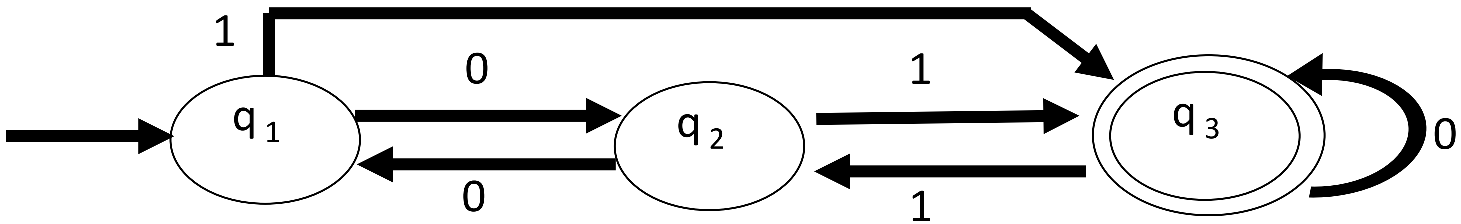
$\boxed{k=3}$

$\boxed{R_{13}^3}$ →

→ $R_{13}^3 = R_{13}^2 + R_{12}^2 (R_{33}^2)^* R_{33}^2$

EXAMPLE

- Find a regular expression representing the set L over an alphabet $\sum=\{0, 1\}$ accepted by the following DFA M.



EXAMPLE

	k=0	k=1	k=2
r(1,1,k)	e	e	0(00)*0+ e
r(1,2,k)	0	0	0(00)*
r(1,3,k)	1	1	0(00)*(1+01)+1
r(2,1,k)	0	0	(00)*0
r(2,2,k)	e	00+e	(00)*
r(2,3,k)	1	1+01	(00)*(1+01)
r(3,1,k)	∅	∅	1(00)*0
r(3,2,k)	1	1	1(00)*
r(3,3,k)	0+ e	0+ e	1(00)*(1+01)+0+ e

EXAMPLE

$r_{1,3}^{3} = r_{1,3}^{2} + r_{1,3}^{2} (r_{3,3}^{2})^* r_{3,3}^{2}$

= (0(00)*(1+01)+1)+(0(00)*(1+01)+1)(1(00)*(1+01)+0+ε)*(1(00)*(1+01)+0+ε)

= (0(00)*(1+01)+1) (1(00)* (1+01)+0)*

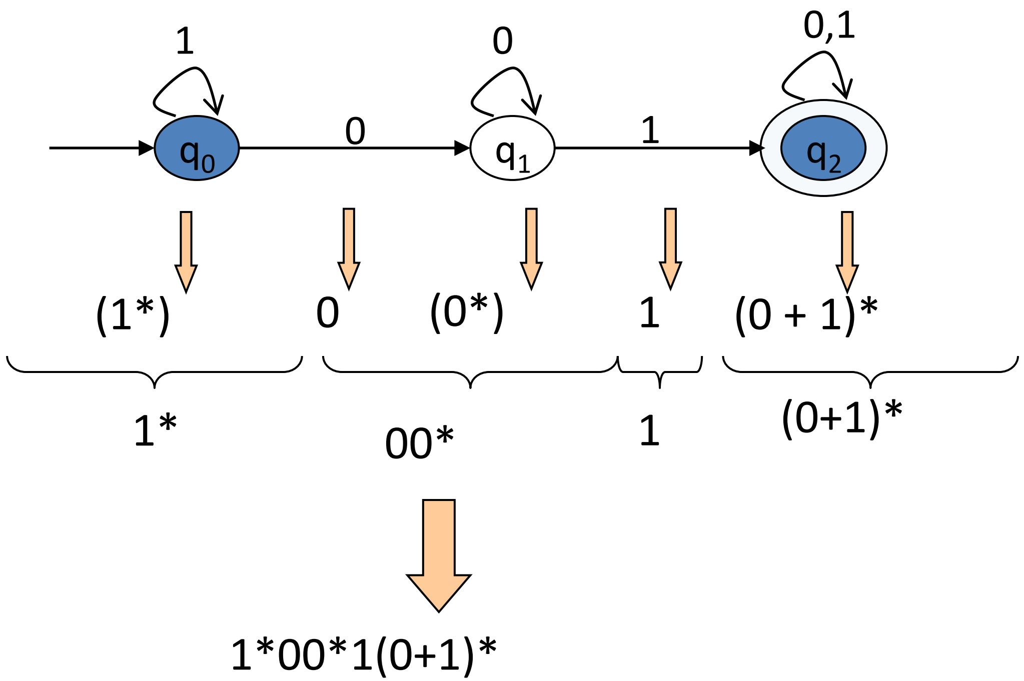
= (0*1) (1(00)* (1+01)+0)*

= (0*1) (10*1+0)*

DFA TO RE CONSTRUCTION

- Informally, trace all distinct paths (traversing cycles only once)
  from the start state to *each of the* final states
  and enumerate all the expressions along the way

Example:



ARDEN’S THEOREM

- In order to find out a RE of a FA, use Arden’s Theorem along with the properties of regular expressions.
- **Lemma:**

Let **P** and **Q** be two regular expressions. If **P** does not contain null string, then **R = Q + RP** has a unique solution that is **R = QP***

- **Proof** $R = Q + RP$

$R = Q + (Q + RP)P$ I12 [After putting the value R = Q + RP]

$= Q + QP + RPP$ $= Q + QP + (Q + RP)PP = Q + QP + QPP + RPPP$

When we put the value of **R** recursively again and again, we get the following equation –

$R = Q + QP + QP^2 + QP^3.....$ I12

$R = Q\ (\varepsilon + P + P^2 + P^3 + ...)$

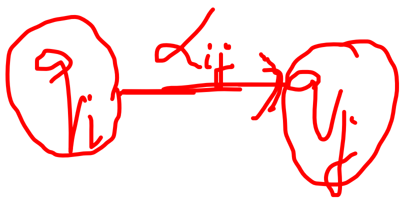
$R = QP^*$ [As P* represents $(\varepsilon + P + P^2 + P^3 + ...)$]

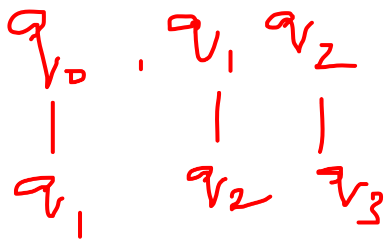
Hence, proved.

ASSUMPTIONS FOR APPLYING ARDEN'S THEOREM

- The transition diagram must not have NULL transitions
  ε
- It must have only one initial state
- Its states are $q_1, q_2 \ldots\ldots q_n$
- α_{ij} denotes the set of labels of edges from q_i to q_j.

α_{ij}





METHOD

- **Step 1**
  - Create equations as the following form for all the states of the DFA having n states with initial state q_1.

If there is no edge $\boldsymbol{\alpha_{ij} = \phi}$.

$$q_1 = q_1\alpha_{11} + q_2\alpha_{21} + \ldots\ldots + q_n\,\alpha_{n1} + \varepsilon$$

$$q_2 = q_1\,\alpha_{12} + q_2\,\alpha_{22} + \ldots\ldots + q_n\alpha_{n2}$$

……..

$$q_n = q_1\,\alpha_{1n} + q_2\alpha_{2n} + \ldots\ldots + q_n\alpha_{nn}$$

n

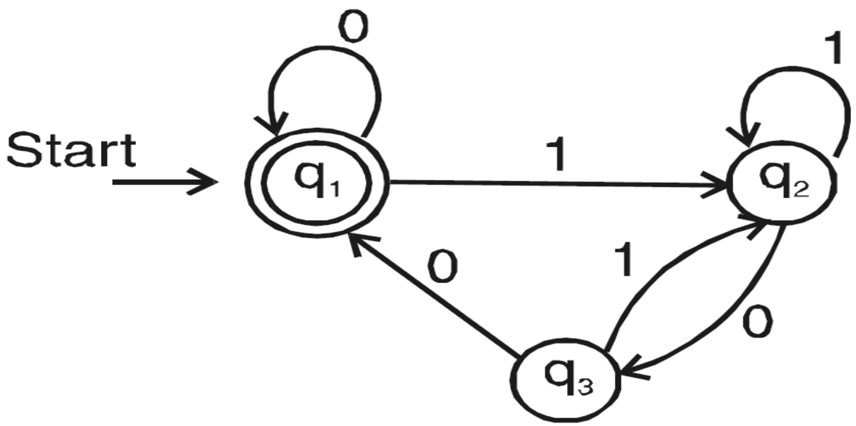
α_{ij}

q_1

$q_i \rightarrow q_1$ RE $\rightarrow$ FA

- **Step 2**
  - Solve these equations to get the equation for the final state in terms of $\boldsymbol{\alpha_{ij}}$

EXAMPLE



$q_1 = q_1\alpha_{11} + q_2\alpha_{21} + \ldots\ldots + q_n\,\alpha_{n1} + \varepsilon$

$q_2 = q_1\,\alpha_{12} + q_2\,\alpha_{22} + \ldots\ldots + q_n\alpha_{n2}$

……..

$q_n = q_1\,\alpha_{1n} + q_2\alpha_{2n} + \ldots\ldots + q_n\alpha_{nn}$

3

$q_1 = q_1\alpha_{11} + q_2\alpha_{21} + q_3\alpha_{31} + \varepsilon$

$q_2 = q_1\alpha_{12} + q_2\,\alpha_{22} + q_3\alpha_{32}$

$q_3 = q_1\alpha_{13} + q_2\,\alpha_{23} + q\,\alpha_{33}$

3 states

α_{ij}

$q_1 = q_1 0 + q_2 \phi + q_3 0 + \varepsilon$ —— (1)

$q_2 = q_1 1 + q_2 1 + q_3 1$ —— (2)

$q_3 = q_1 \phi + q_2 0 + q_3 \phi$ —— (3)

$q_1 \to F \to RE$

$q_3 = q_2 0 \longrightarrow (4)$

$q_2 = q_1 1 + q_2 1 + q_3 1$

$q_2 = q_1 1 + q_2 1 + q_2 01$

$RP + RQ = R(P+Q)$ 12

$\to \underbrace{q_2}_{R} = \underbrace{q_1 1}_{Q} + \underbrace{q_2}_{R}\underbrace{(1+01)}_{P}$

$R = Q + RP$

$R = QP^*$

$q_2 = q_1 1 (1+01)^*$ —— (5)

EXAMPLE

Substitute (5) in (4)

$$q_3 = q_2 0 \longrightarrow (4)$$

$$q_2 = q_1 \cdot 1(\cancel{0+}\, 1+01)^* \text{ —— } (5)$$

$$\underline{q_3} = q_1 \cdot 1(1+01)^* 0 \text{ —— } (6)$$

Substitute (6) in (1)

$$(1) \rightarrow q_1 = q_1 0 + \underline{q_3} 0 + \varepsilon$$

$$q_1 = \underbrace{q_1}_{R} \underbrace{0}_{P} + \underbrace{q_1}_{R} \cdot \underbrace{1(1+01)^* 0}_{Q} + \varepsilon$$

$RP + RQ$

$$\underbrace{q_1}_{R} = q_1 \underbrace{(0 + 1(1+01)^* 0)}_{P} + \underbrace{\varepsilon}_{Q}$$

$R(P+Q)$

Ardens lemma

$R = RP + Q$

$R = QP^*$

$$q_1 = \varepsilon(0+1(1+01)^* 0)^*$$

$$= (0+1(1+01)^* 0)^*$$

RE

EXAMPLE

- 2 final states

$q_1 = q_1 0 + q_2 \phi + q_3 \phi + \varepsilon$

$q_2 = q_1 1 + q_2 1 + q_3 \phi$

$q_3 = q_1 \phi + q_2 0 + q_3 (0+1)$

q_1

q_2

$q_1 = q_1 \cdot 0 + q_2 \phi + q_3 \phi + \varepsilon = \varepsilon + q_1 \cdot 0 : q_1 = 0^* \quad (\because R = Q + RP)$

$q_2 = q_1 \cdot 1 + q_2 \cdot 1 + q_3 \phi \cancel{+\varepsilon} \qquad q_2 = 0^* \cdot 1 + q_2 \cdot 1 \quad (R = Q + RP)$

$\downarrow$

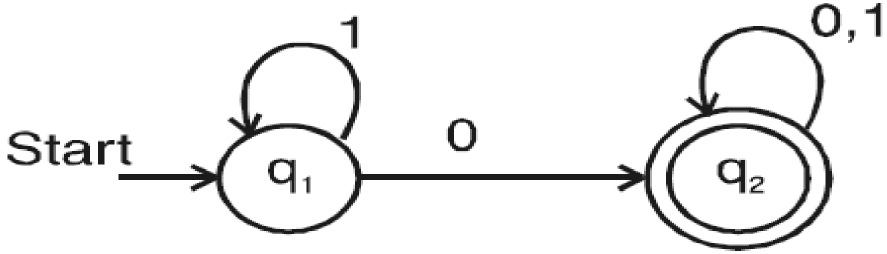
$q_3 = q_1 \cdot \phi + q_2 \cdot 0 + q_3 (0+1) \qquad q_2 = 0^* \cdot 1 (1^*)$

$q_3 = q_2 \, 0 \cdot (0+1)^*$

Final Regex:

$q_1 + q_2 = (0^* + 0^* \cdot 1 \cdot (1^*))$

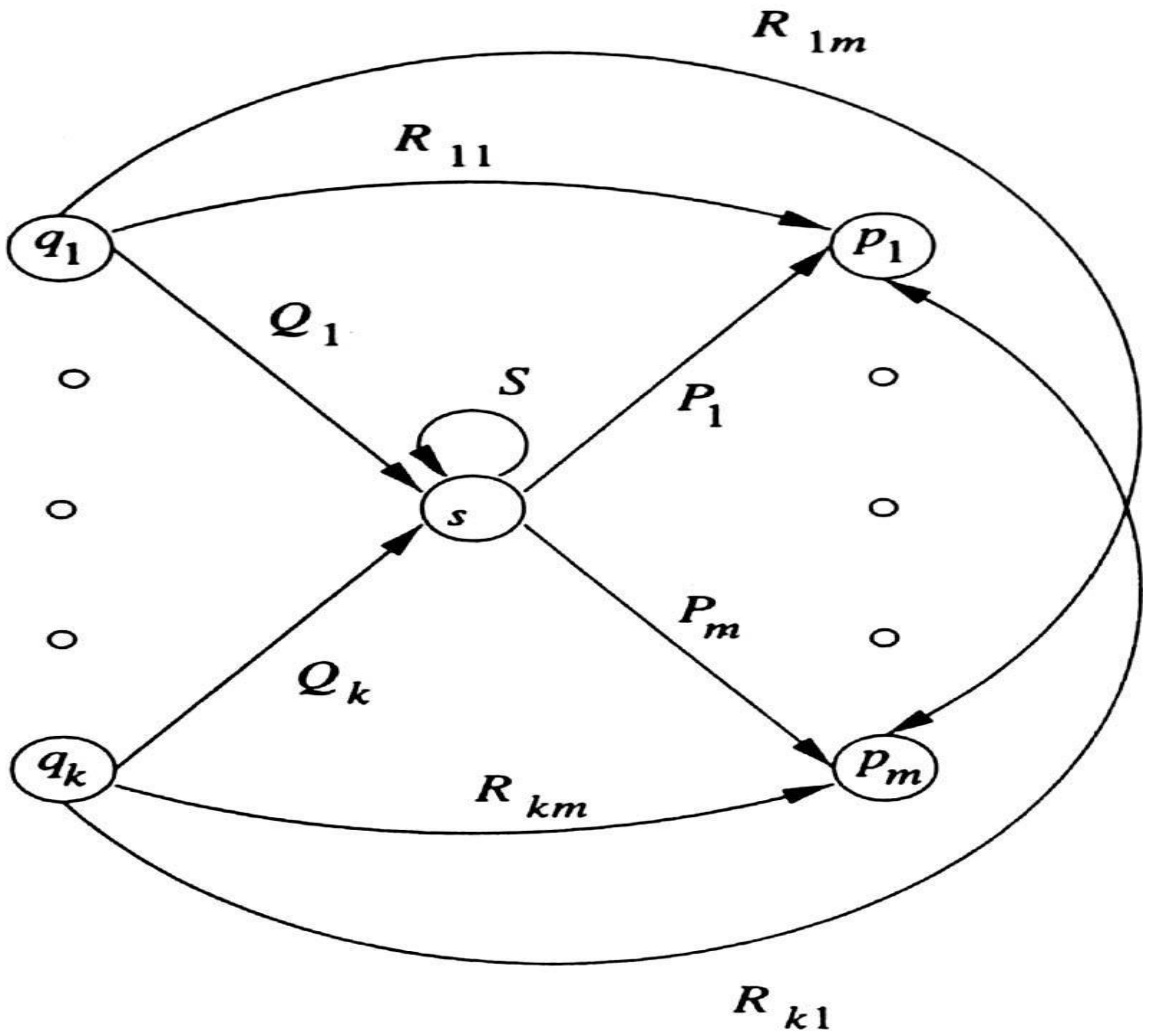
EXAMPLE

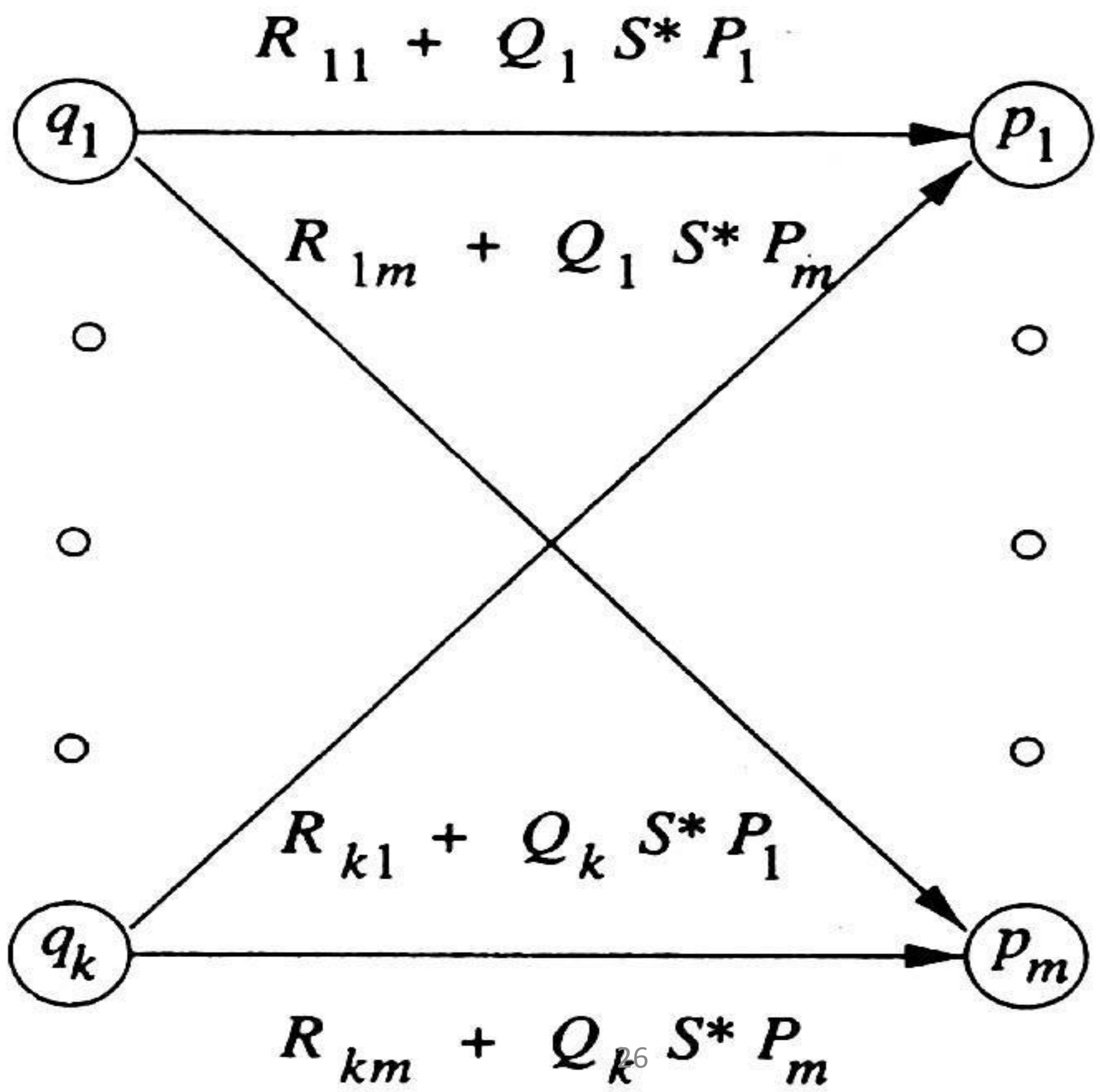


STATE ELIMINATION METHOD

- S → intermediate state
- Predecessor of S → q1, q2 … qk
- Successor of S → p1, p2, … pm

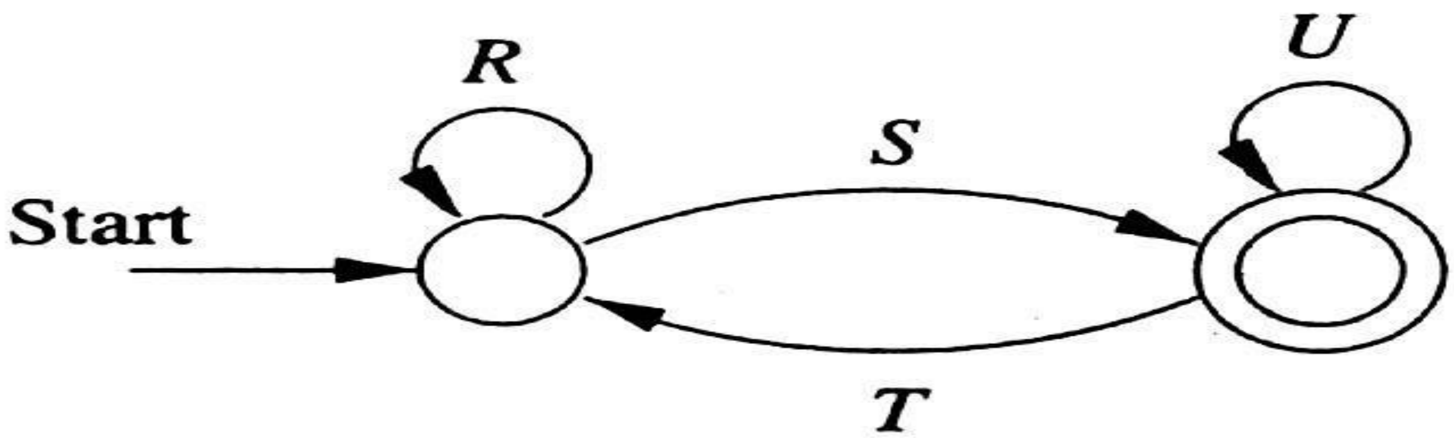
S BEFORE & AFTER ELIMINATION





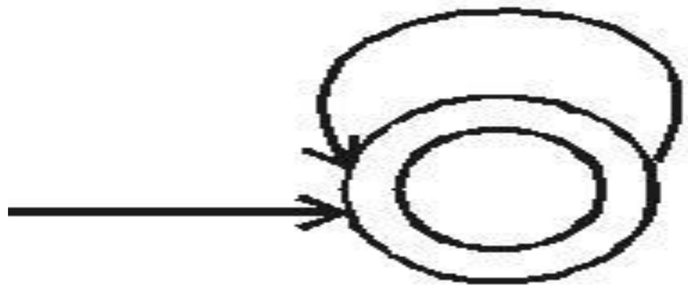
STEPS

1. Eliminate all states except q and the start state q_0
2. $q \neq q0$
   – $(R + SU^*T)^*SU^*$



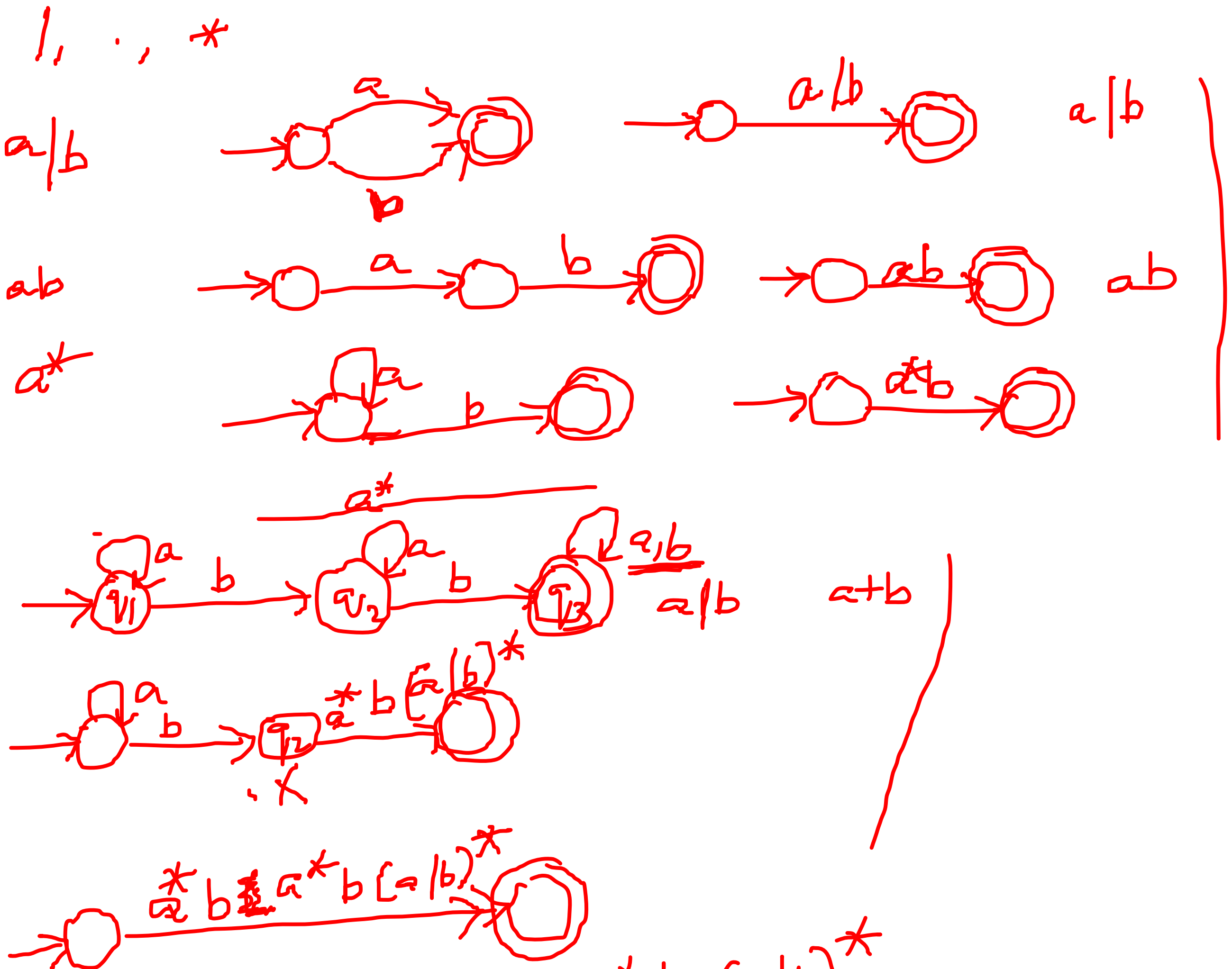
STEPS

3. Start state = final state



4. Union of all expressions derived from 2 and 3

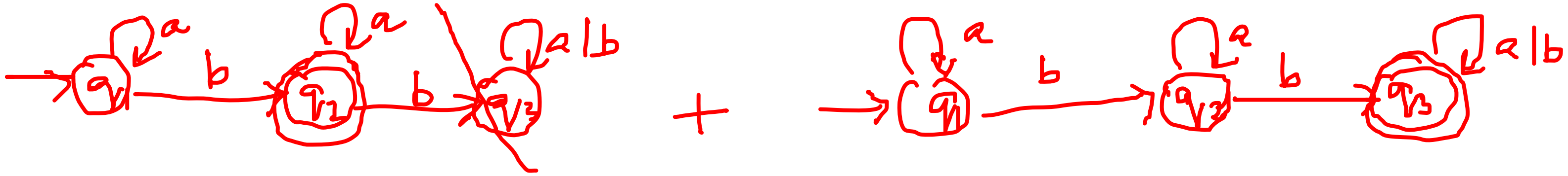
EXAMPLE



RE $\rightarrow$ $a^*ba^*b(a|b)^*$

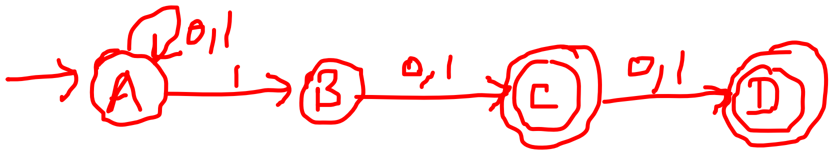
EXAMPLE

- 2 final states



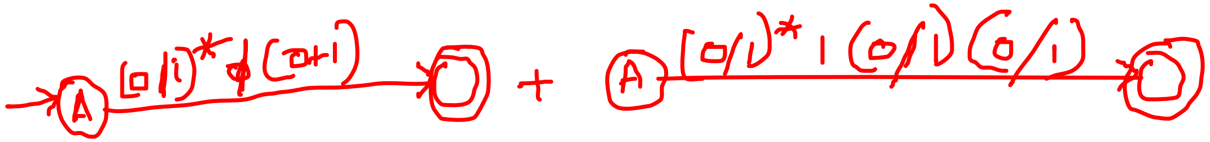
$$a^*ba^* + a^*ba^*b(a|b)^*$$

EXAMPLE



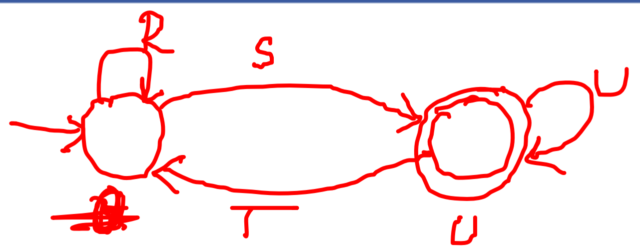
C → RE D → RE

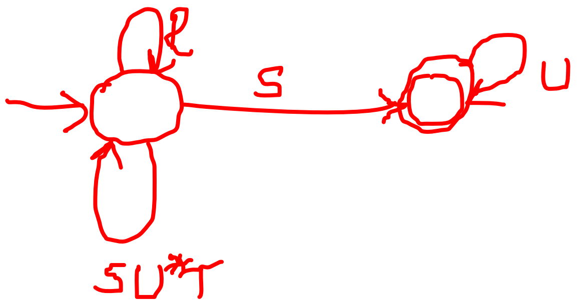
+

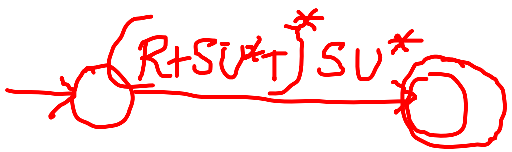


$(0/1)^* \phi (0/1) + (0/1)^* 1 (0/1)(0/1)$

EXAMPLE

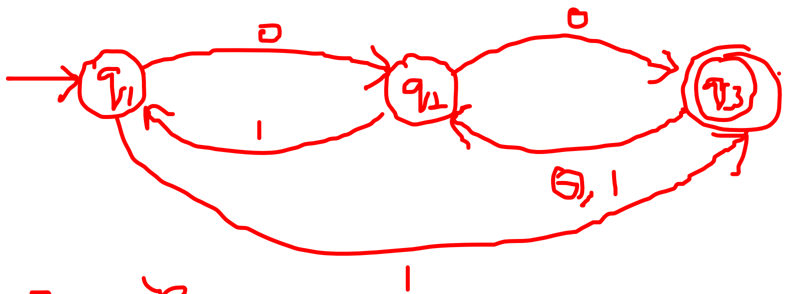






RE → $(R+SU^*T)^*SU^*$

EXAMPLE



$q_1 \rightarrow q_1$ $q_1 \rightarrow q_1 + q_1 \rightarrow q_2 \rightarrow q_1$

01 $\phi + 01$

$q_1 \rightarrow q_3$ $q_1 \rightarrow q_3 + q_1 \rightarrow q_2 \rightarrow q_3$

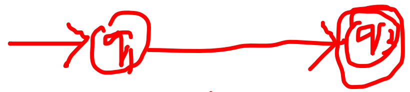
$1+00$ $1 + 00$

$q_3 \rightarrow q_1$ $q_3 \rightarrow q_1 + q_3 \rightarrow q_2 \rightarrow q_1$

$(0+1)1$ $\phi + (0+1)1$

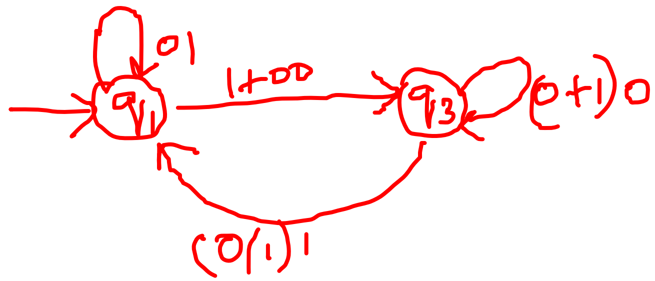
$q_3 \rightarrow q_3$ $q_3 \rightarrow q_3 + q_3 \rightarrow q_2 \rightarrow q_3$

$(0+1)0$ $\phi + (0/1)0$



13

11
13
31
33



$(R + SU^*T)^* SU^*$

$[01 / 1+00((0+1)0)^* (0/1)1]^* (1+00)((0+1)0)^*$

EXAMPLE

$q_1 \to q_1$:

$\Rightarrow q_1 \to q_1 + q_1 \to q_2 \to q_1$

$\Rightarrow \phi + 1\phi$

$\Rightarrow \phi$

$q_1 \to q_3$:

$\Rightarrow q_1 \to q_3 + q_1 \to q_2 \to q_3$

$\Rightarrow 0 + 11^*0$

$q_3 \to q_1$:

$\Rightarrow q_3 \to q_1 + q_3 \to q_2 \to q_1$

$\Rightarrow \phi + 1\phi$

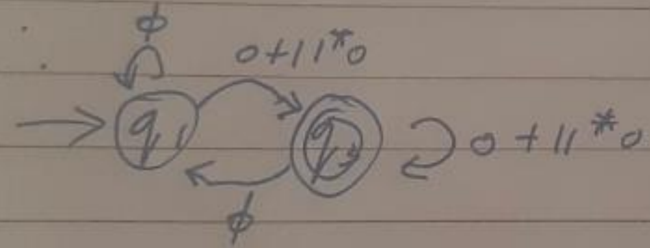
$\Rightarrow \phi$

$q_3 \to q_3$:

$\Rightarrow q_3 \to q_3 + q_3 \to q_2 \to q_3$

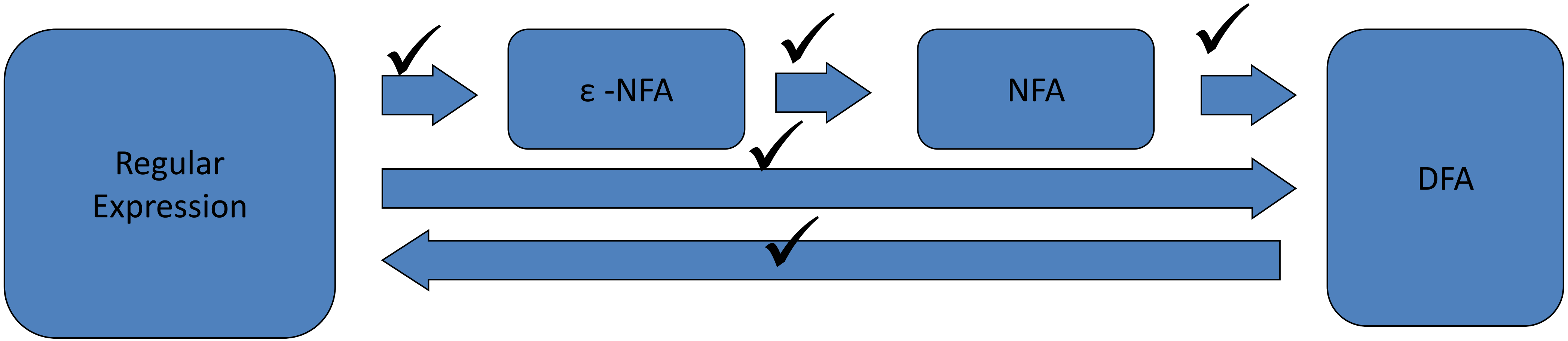
$= 0 + 11^*0$

$\therefore$



$R.E = (0 + 11^*0)(0 + 11^*0)^*$

ROAD MAP



A language is regular iff it is accepted by a DFA, NFA, eNFA, or regular expression

SUMMARY

- Definition of RE
- Precedence, identities, properties of RE.
- Thomson’s construction to convert RE to NFA and then to DFA

TEST YOUR KNOWLEDGE

- Which of the following does not represents the given language?

  Language: {0,01}

  a) 0+01

  b) {0} U {01}

  c) {0} U {0}{1}

  d) {0} ^ {01}

TEST YOUR KNOWLEDGE

- Regular Expression R and the language it describes can be represented as:
  a) R, R(L)
  b) L(R), R(L)
  c) R, L(R)
  d) All of the mentioned

LEARNING OUTCOME

On successful completion of this topic, the student will be able to:

- Understand equivalence of FA and RE(K3)

REFERENCE

- Hopcroft J.E., Motwani R. and Ullman J.D, “Introduction to Automata Theory, Languages and Computations”, Second Edition, Pearson Education, 2008