UNIT-2

Regular Expressions

The language of FA is represented by an expression called regular expression (RE).

Kleen closure $(L^*) = \bigcup_{i=0}^{\infty} L^i = \{L^0, L^1, L^2, \ldots\}$ → including ϵ

Positive closure $(L^+) = \bigcup_{i=1}^{\infty} L^i = \{L^1, L^2, \ldots\}$ → excluding ϵ

$$L^+ = L^* - \{\epsilon\}$$

Eg: $\left.\begin{array}{l} 0^* = \{\epsilon, 0, 00, 000, \ldots\} \\ 0^+ = \{0, 00, 000, \ldots\} \end{array}\right\}$ for $\Sigma = \{0\}$

Consider $\Sigma = \{0,1\} \Rightarrow L^0 = \{\epsilon\}, L^1 = \{0,1\}, L^2 = \{00, 01, 10, 11\},$
$L^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}, \ldots$

$\Rightarrow L^* = \bigcup_{i=0}^{\infty} L^i = \{\epsilon, 0, 1, 00, 01, 10, 11, \ldots\}$

↳ L^* can be rep as $(0+1)^*$ or $(0|1)^*$

$\Rightarrow L^+ = \bigcup_{i=1}^{\infty} L^i = \{0, 1, 00, 01, 10, 11, \ldots\}$

Same with L^+

Formal defn of RE

Let Σ be an alphabet or a symbol, then the RE over Σ and the sets that are represented can be defined recursively.

→ ϕ is a reg exp which denotes empty set

→ ϵ is a reg exp which denotes $\{\epsilon\}$

→ $\forall a \in \Sigma$ where a is a reg exp which denotes $\{a\}$

→ if r, s are REs denoting the languages R, S respectively, then $(r+s), (rs), (r^*)$ are reg exps which denote $(R \cup S)$, (RS), (R^*) respectively.

Q: Represent the given set as a RE using closure.

$L = \{0, 00, 000, \ldots\} \Rightarrow RE = 00^*$ $\quad 0^+$

Q: The set of all strings over $\Sigma = \{0,1\}$ which has at most two zeroes consecutively, find its equivalent reg exp.

(no zeroes + one zero + two zeroes)

$\downarrow$ 1^* $\downarrow$ 1^*01^* $\downarrow$ 1^*001^*

$$RE = 1^* + 1^*01^* + 1^*001^*$$

Q: $L = a^n$ such that n is divisible by 2 or 3 or $n=5$ }

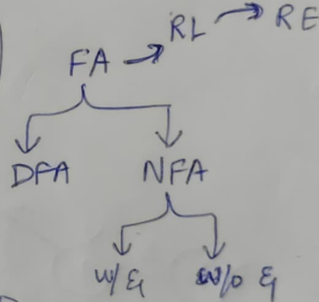
find equivalent regular expressions

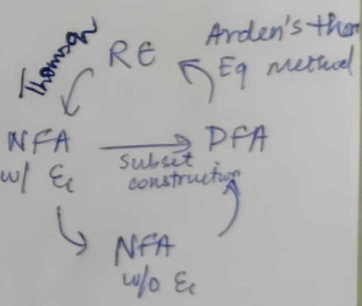
Q: $L = a^2, a^5, a^8$

↳ $RE = aa(aaa)^*$

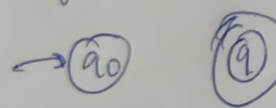
24.01.2025

Thomson's construction → convert RE into ϵ NFA representation

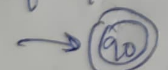




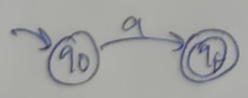
Rule ①: If ϕ is a RE it is empty set



Rule ②: If ϵ is a RE it is $\{\epsilon\}$



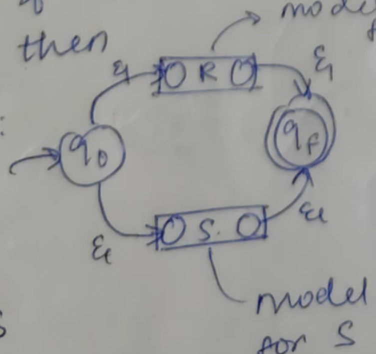
Rule ③: If $a \in \Sigma$ is a RE then rep is $\{a\}$



basis steps

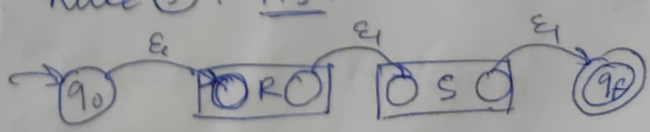
Rule ④: If r, s are REs rep by sets R, S, then

r+s:

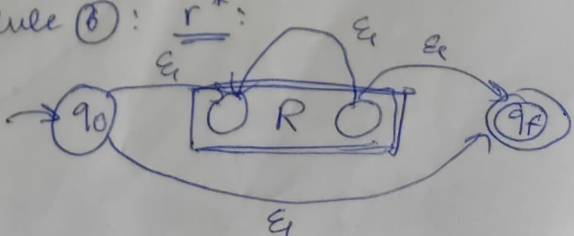


R ∪ S

Rule ⑤: r.s:

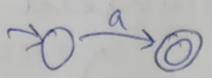


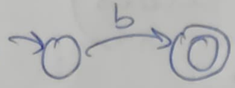
Rule ⑥: r*:

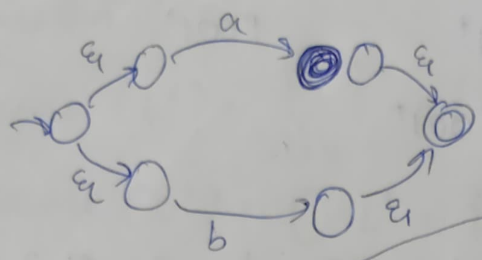


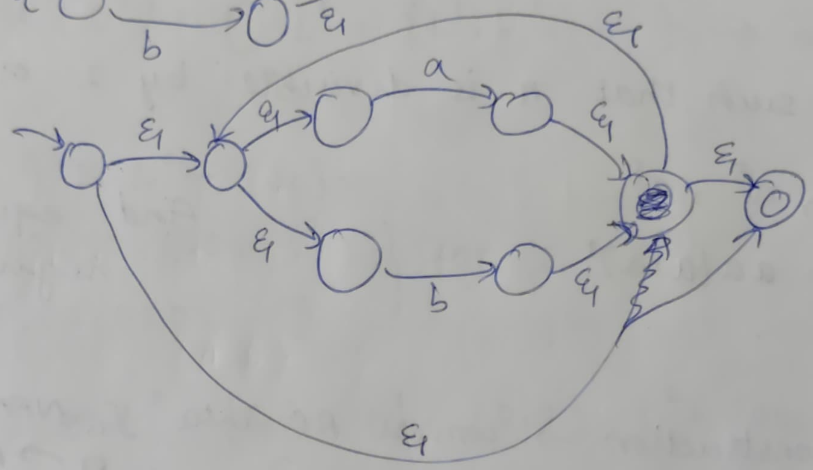
ind^n steps

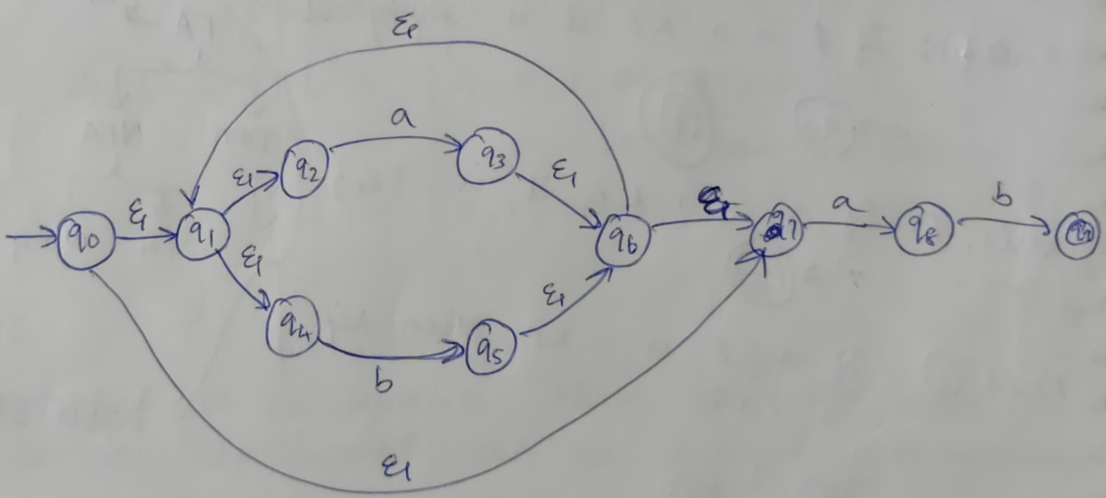
Q: Construct the ε NFA for the given RE = $(a+b)^* ab$

$a \Rightarrow$ 

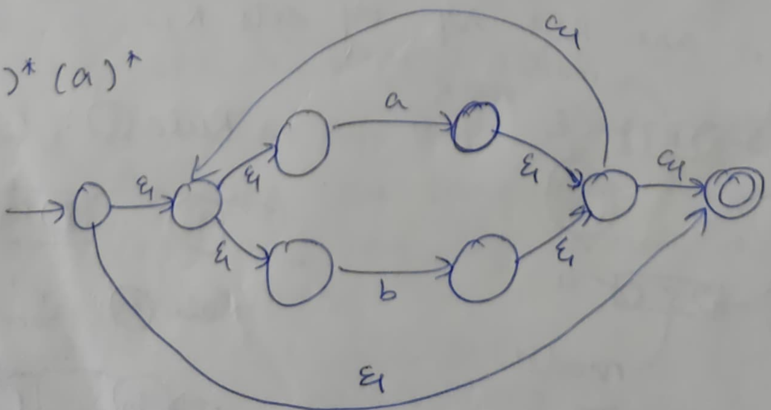
$b \Rightarrow$ 

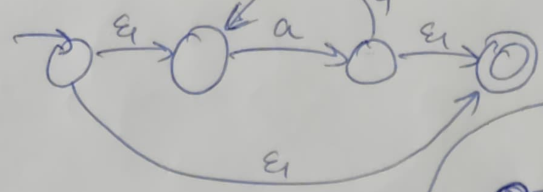
$a+b \Rightarrow$ 

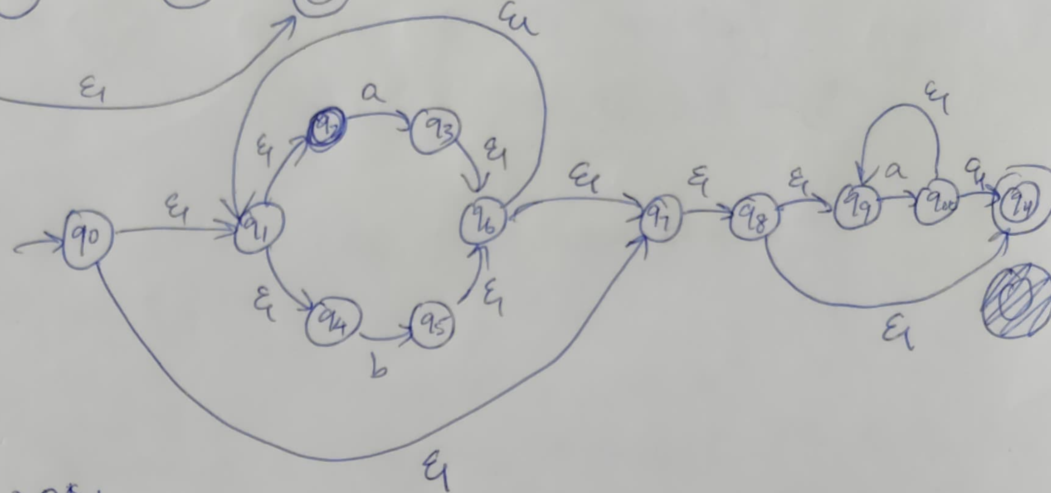
$(a+b)^* \Rightarrow$ 

$(a+b)^* ab \Rightarrow$ 

Q: RE = $(a+b)^* (a)^*$

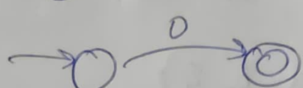
$(a+b)^* \Rightarrow$ 

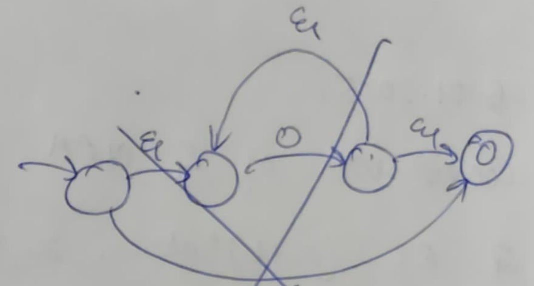
$a^* \Rightarrow$ 

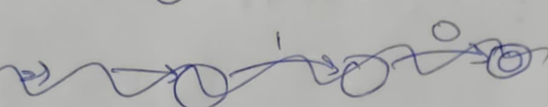
$(a+b)^*(a)^* \Rightarrow$ 

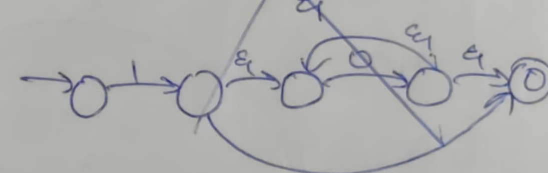
Q: RE = $\boxed{10^+}(0+11)0^*1$

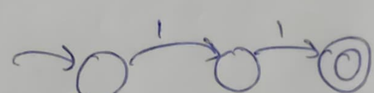
$1 \Rightarrow$ 

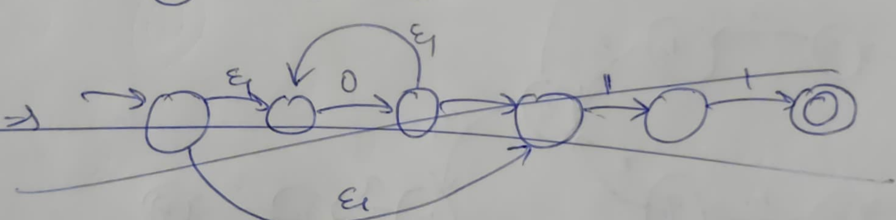
$0 \Rightarrow$ 

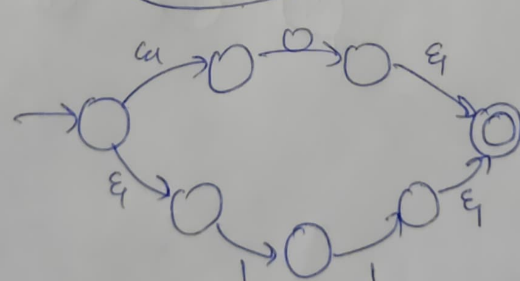
$0^* \Rightarrow$ 

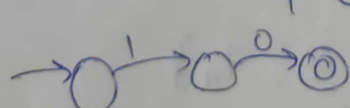
~~10 =>~~ 

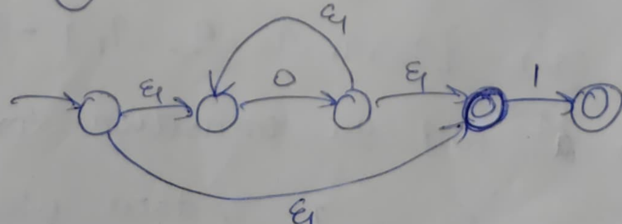
$10^+ \Rightarrow$ 

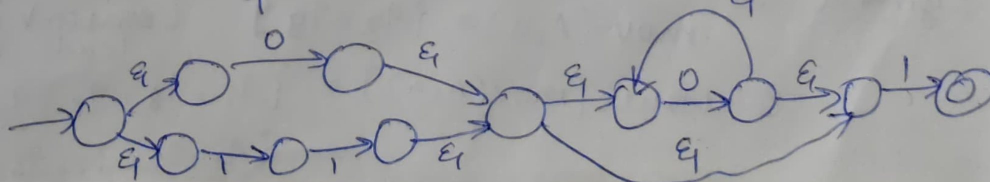
$11 \Rightarrow$ 

~~(0+11) =>~~ 

$(0+11) \Rightarrow$ 

$10 \Rightarrow$ 

$0^*1 \Rightarrow$ 

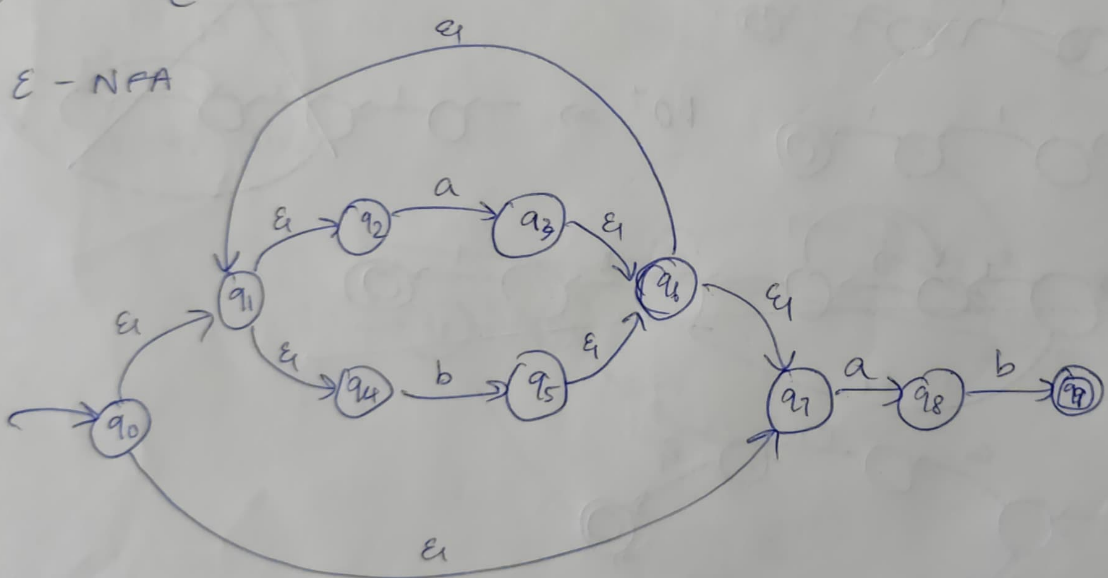
$(0+11)0^+1 \Rightarrow$ 

25.01.2025

Conversion of ϵ NFA into DFA using subset construction

Q: $RE = (a+b)^* ab$, $\Sigma = \{a, b\}$, $\Sigma^* = \{a, b, \epsilon\}$

ϵ - NFA



① find the ϵ closure (q_0)

ϵ closure $(q_0) = \{q_0, q_1, q_7, q_2, q_4\}$

ϵ closure $(q_0) = \{q_0, q_1, q_2, q_4, q_7\} \rightarrow$ A

② for the new state A, find the transition for each symbol Σ.

$mov(A, a) = \{q_3, q_8\}$ (states where 'a' transition ~~exists~~ in ϵ closure (q_0) leads to from states)

ϵ closure $[mov(A, a)] = \{q_3, q_8, q_6, q_7, q_1, q_2, q_4\}$

$= \{q_1, q_2, q_3, q_4, q_6, q_7, q_8\}$ — B

$mov(A, b) = \{q_5\}$ $\quad (q_4 \in A,\ q_4 \xrightarrow{b} q_5)$

$\epsilon\ closure\ [mov(A, b)] = \{q_5, q_6, q_7, q_1, q_2, q_4\}$

$= \{q_1, q_2, q_4, q_5, q_6, q_7\}$ — C

(3) Find the mov and ε closure for the new states B and C for each input symbol Σ.

$mov(B, a) = \{q_3, q_8\}$

$\epsilon\ closure\ [mov(B, a)] = B$

$mov(B, b) = \{q_5, q_9\}$

$\epsilon\ closure\ [mov(B, b)] = \{q_5, q_9, q_6, q_7, q_1, q_2, q_4\}$

$= \{q_1, q_2, q_4, q_5, q_6, q_7, q_9\}$ — D

$mov(C, a) = \{q_3, q_8\}$

$\epsilon\ closure\ [mov(C, a)] = B$

$mov(C, b) = \{q_5\}$

$\epsilon\ closure\ [mov(C, b)] = C$

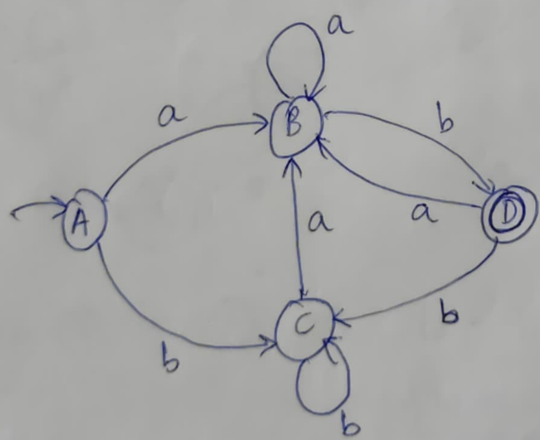
$mov(D, a) = \{q_3, q_8\}$

$\epsilon\ closure\ [mov(D, a)] = B$

$mov(D, b) = \{q_5\}$

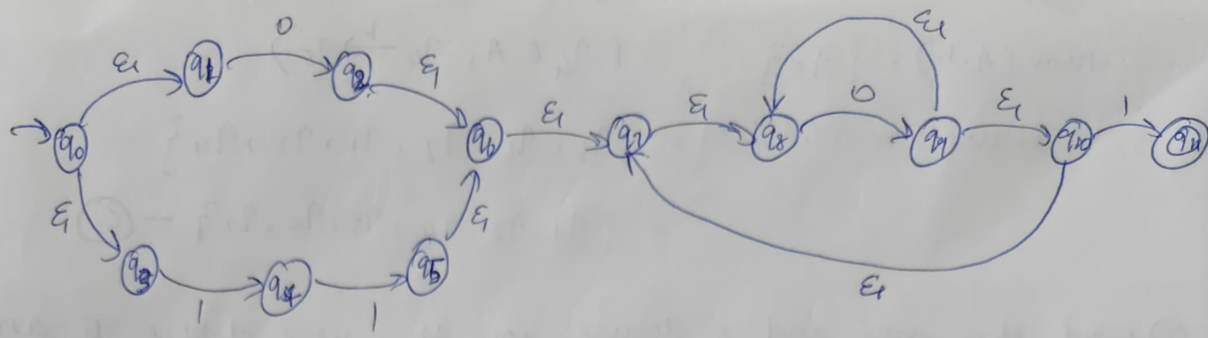
$\epsilon\ closure\ [mov(D, b)] = C$

DFA:



States	a	b
→ A	B	C
B	B	D
C	B	C
* D	B	C

Q:



	0	1
A		
B		
C		
D		B
E	ϕ	ϕ
F		

28.01.2025

Conversion of DFA to RE using Arden's theorem

Conditions: (1) The FA should not have ϵ transition

(2) The FA should have only one start state

(3) α_{ij} denotes the set of labels of edges from Q_i to Q_j, if there is no edge then $\alpha_{ij} = \phi$

Assume the states are $q_1, q_2, \ldots, q_n$, then

$$q_1 = q_1 \alpha_{11} + q_2 \alpha_{21} + \cdots + q_n \alpha_{n1} + \epsilon$$

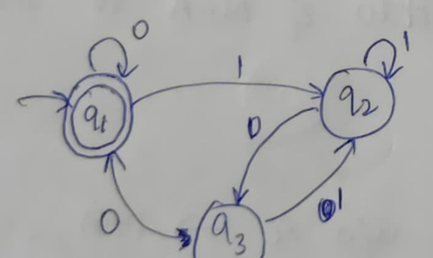
$$q_2 = q_1 \alpha_{12} + q_2 \alpha_{22} + \cdots + q_n \alpha_{n2}$$

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

(4) Apply repeated substitution to express RE in terms of α_{ij}.

(5) If the eqn is of the form $R = Q + RP$, then the solution is $R = QP^*$

Q: Construct a RE to the given FA using Arden's theorem.



$q_1 = q_1 0 + q_3 0$ — (1)

$q_2 = q_1 1 + q_2 1 + q_3 1$ — (2)

$q_3 = q_2 0$ — (3)

Substitute (3) in (2) $\Rightarrow$ $q_2 = q_1 1 + q_2 1 + q_2 01$

$$\boxed{q_2 = q_1 1 + q_2 (1 + 01)}$$

$\downarrow$

$R = Q + RP \rightarrow$ solⁿ is $R = QP^*$

$\Rightarrow \boxed{q_2 = q_1 1 (1 + 01)^*}$

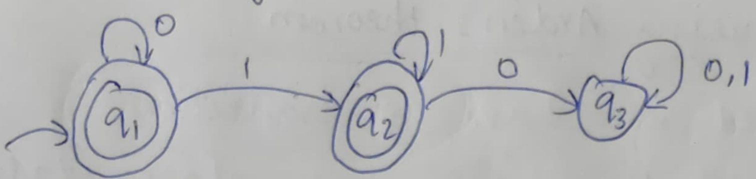
Substitute (3) in (1) $\Rightarrow$ $q_1 = q_1 0 + q_2 00 + \epsilon$

$q_1 = q_1 0 + q_1 1 (1 + 01)^* 00 + \epsilon$

can be ignored

$\boxed{q_1 = \epsilon (0 + 1(1+01)^* 00)^*} \Leftarrow \boxed{q_1 = q_1 (0 + 1(1+01)^* 00) + \epsilon}$

Q: convert the given DFA into RE using Arden's theorem.



$q_1 = q_1 0 + \epsilon \rightarrow R = Q + RP \Rightarrow R = QP^* = \epsilon 0^* \Rightarrow q_1 = \epsilon 0^*$

$\Rightarrow \boxed{q_1 = 0^*}$

$q_2 = q_1 1 + q_2 1$

$q_3 = q_2 0 + q_3 0 + q_3 1$

$q_2 = q_2 1 + q_1 1 \rightarrow R = Q + RP \Rightarrow R = QP^* \Rightarrow \boxed{q_2 = q_1 1 1^*}$

$q_2 = 0^* 1 1^*$

$q_3 = q_2 0 + q_3 (0+1)$

$q_3 = q_1 1 1^* 0 + q_3 (0+1)$

$\boxed{q_3 = q_1 1 1^* 0 (0+1)^*}$ $q_3 = 0^* 1 1^* 0 (0+1)^*$

$\Rightarrow \boxed{q_1 + q_2 = 0^* + 0^* 1 1^*}$ ($\because q_1, q_2$ are both final states)

30.01.2025

HW Q: $RE = (0+1)^* 1^* (00+11)$: a) convert RE into ϵ NFA to DFA to RE

RE $\rightarrow$ ϵ NFA $\rightarrow$ DFA $\rightarrow$ RE

b) convert RE $\rightarrow$ NFA ϵ $\rightarrow$ NFA w/o ϵ $\rightarrow$ DFA

Theorem: For every RE 'r', there exists a NFA with ϵ transition that accepts L(R) (Thomson construction)

Proof: This thm statement can be proved by induction on the no. of operators in the RE that there is a NFA w/ ϵ transition having 1 final state and there is no transition out of this final state such that $L(M) = L(R)$

Basis step: Assume basis step is equivalent to zero operators.

$\Rightarrow$ the cases will be $r = \epsilon$, $r = \phi$, $r = a$

Induction step: Assume one or more operators. Let the given RE have i operators ($i \geq 1$).

case (i): union $\Rightarrow$ let $r = r_1 + r_2$ $\Rightarrow$ both r_1 and r_2 must have less than i operators

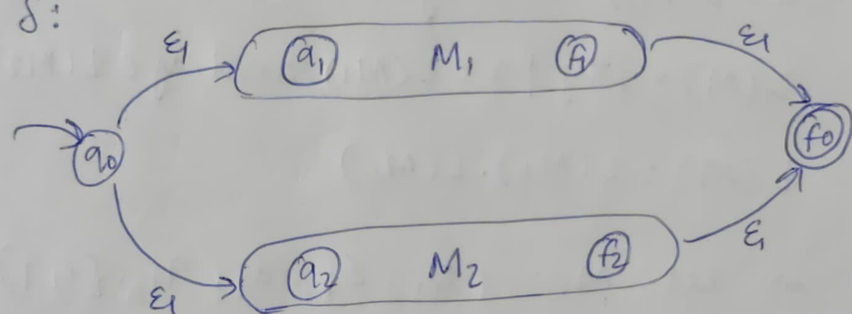
$M_1 = (Q_1, \Sigma_1, \delta_1, q_1, \{f_1\})$ $\quad L(M_1) = L(r_1)$

$M_2 = (Q_2, \Sigma_2, \delta_2, q_2, \{f_2\})$ $\quad L(M_2) = L(r_2)$

Assume Q_1 and Q_2 are disjoint sets. Let q_0, f_0 be the new start and final states.

$\Rightarrow M = (Q_1 \cup Q_2 \cup \{q_0, f_0\}, \Sigma_1 \cup \Sigma_2, \delta, q_0, \{f_0\})$

δ:



① $\delta(q_0, \epsilon) = \{q_1, q_2\}$

② $\delta(q, a) = \delta_1(q, a)$ $\quad q \in Q_1 - f_1,\ a \in \Sigma_1 \cup \{\epsilon\}$

③ $\delta(q, a) = \delta_2(q, a)$ $\quad q \in Q_2 - f_2,\ a \in \Sigma_2 \cup \{\epsilon\}$

④ $\delta(f_1, \epsilon) = \delta(f_2, \epsilon) = f_0$

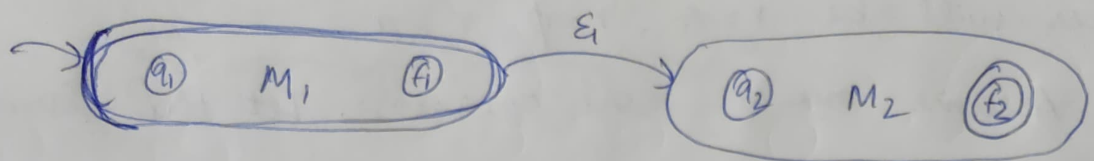
Conclusion: Let x be a string, then $x \in M$ and the path from q_0 to f_0 iff there is a path labeled x in M_1 from q_1 to f_1 or in M_2 from q_2 to f_2.

$\therefore$ $L(M) = L(M_1) \cup L(M_2)$

Case (ii): concatenation $\Rightarrow$ let $r = r_1 . r_2$

$M_1, M_2, L(M_1), L(M_2)$ $\rightarrow$ same as in the case of union

Assume Q_1 and Q_2 are disjoint, q_1 and f_2 are new start and final states



$M = (Q_1 \cup Q_2, \Sigma_1 \cup \Sigma_2, \delta, q_1, \{f_2\})$

$\delta(q, a) = \delta_1(q, a) \qquad q \in Q_1 - f_1, \ a \in \Sigma_1 \cup \{\varepsilon\}$

$\delta(q, a) = \delta_2(q, a) \qquad q \in Q_2, \ a \in \Sigma_2 \cup \{\varepsilon\}$

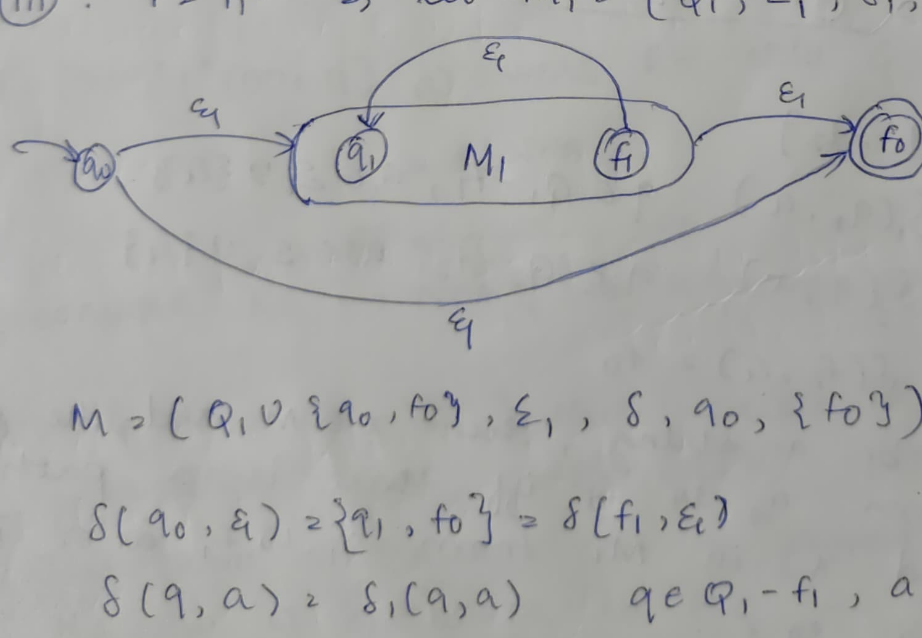
$\delta(f_1, \varepsilon) = q_2$

conclusion: Assume xy is a string, then $L(M) = xy$ such that $x \in L(M_1)$ and $y \in L(M_2)$

$L(M) = \{xy \mid x \in L(M_1) \text{ and } y \in L(M_2)\}$

$\therefore \ L(M) = L(M_1) \cdot L(M_2)$

case (iii) : $r = r_1^*$ $\Rightarrow$ let $M_1 = (Q_1, \Sigma_1, \delta_1, q_1, \{f_1\})$



$M = (Q_1 \cup \{q_0, f_0\}, \Sigma_1, \delta, q_0, \{f_0\})$

$\delta(q_0, \varepsilon) = \{q_1, f_0\} = \delta(f_1, \varepsilon)$

$\delta(q, a) = \delta_1(q, a) \qquad q \in Q_1 - f_1, \ a \in \Sigma_1 \cup \{\varepsilon\}$

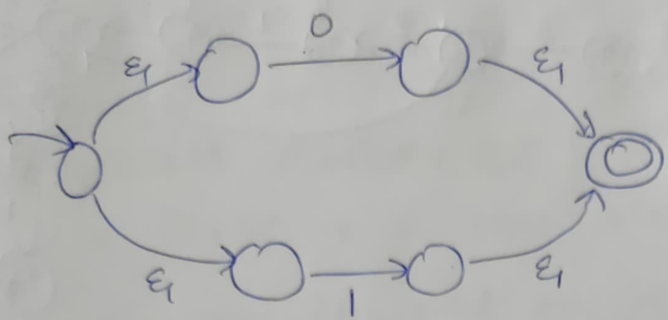
conclusion: There is a path in M from q_0 to f_0 labeled 'x' iff $x = x_1 x_2 x_3 \cdots x_j$ where $j \geq 0$ such that each $x_i \in L(M_1)$.

$\therefore \ L(M) = L(M_1)^*$ is proved.

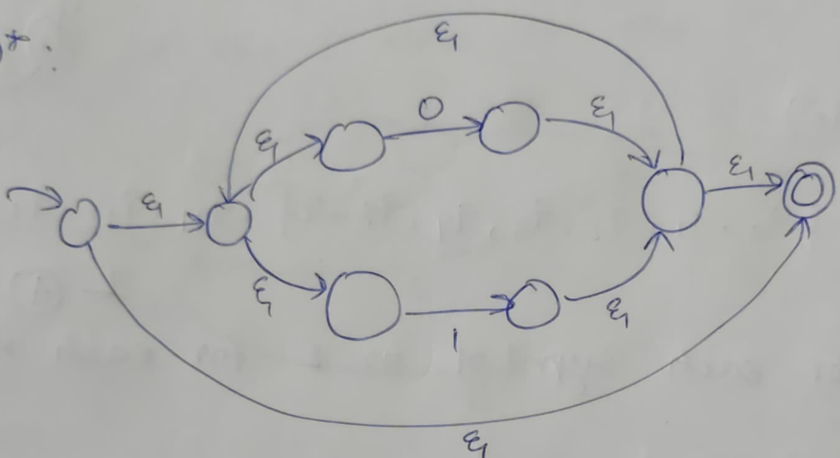
HW

Q: RE = $(0+1)^* 1^* (00+11)$ RE → ε NFA → DFA

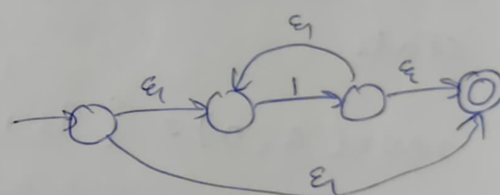
0+1:



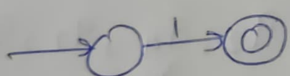
$(0+1)^*$:



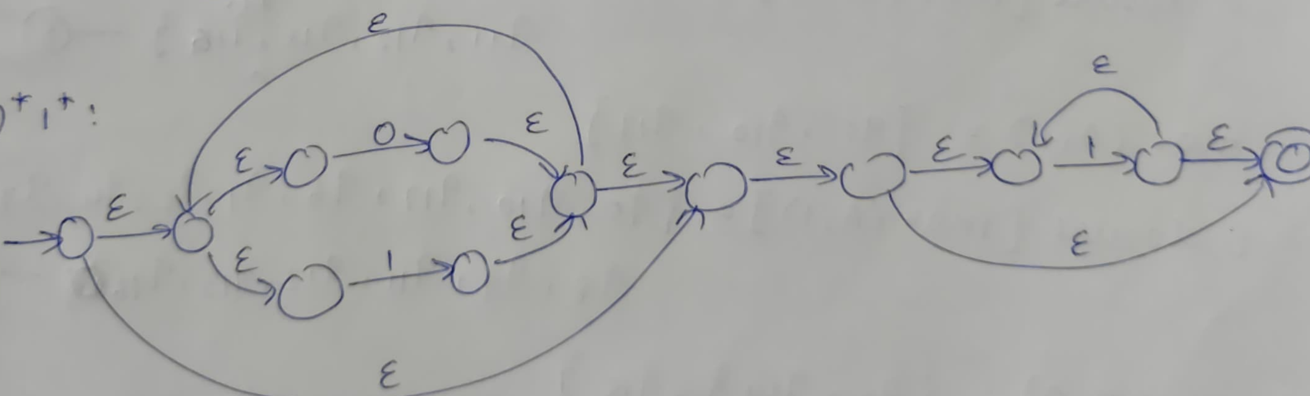
1^*:



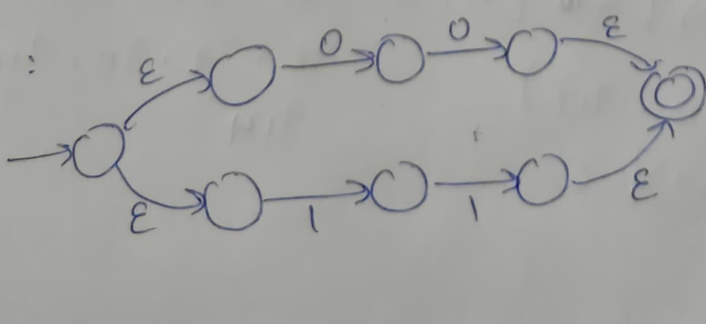
1^*:



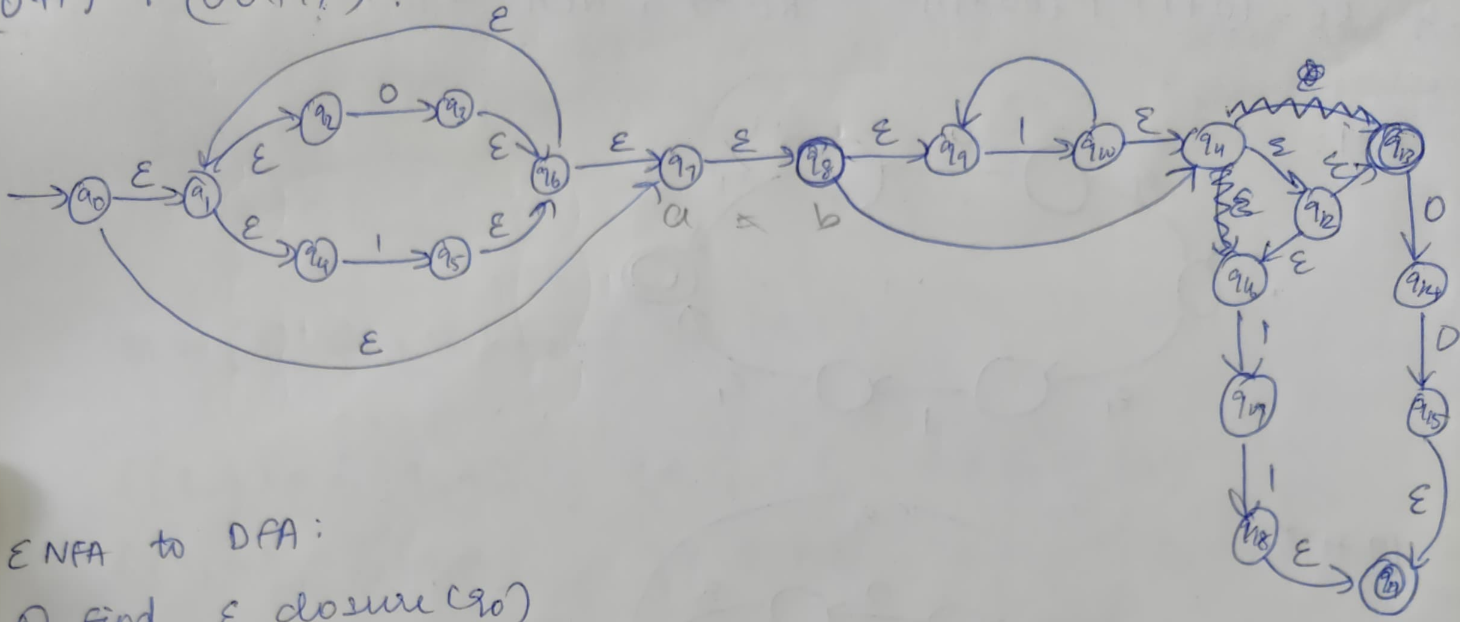
$(0+1)^* 1^*$:



$(00+11)$:



$(0+1)^* 1^*(00+11)$:



ε NFA to DFA:

① Find ε closure(q_0)

ε closure(q_0) = $\{q_0, q_1, q_2, q_3, q_7, q_8, q_9, q_{11}, q_{12}, q_{13}, q_{16}\}$ — (A)

② Find transition for each symbol in Σ for each new state.

mov(A, 0) = $\{q_3, q_{14}\}$

ε closure [mov(A, 0)] = $\{q_3, q_6, q_1, q_2, q_4, q_7, q_8, q_9, q_{11}, q_{12}, q_{13}, q_{14}, q_{16}\}$ — (B)

mov(A, 1) = $\{q_5, q_{10}, q_{17}\}$

ε closure [mov(A, 1)] = $\{q_5, q_{10}, q_{17}, q_6, q_1, q_2, q_4, q_7, q_8, q_9, q_{11}, q_{12}, q_{13}, q_{16}\}$ — (C)

mov(B, 0) = $\{q_3, q_{14}, q_{15}\}$

ε closure [mov(B, 0)] = $\{q_3, q_{14}, q_{15}, q_6, q_1, q_2, q_4, q_7, q_8, q_9, q_{11}, q_{12}, q_{13}, q_{14}$

mov(B, 1) = $\{q_5, q_{10}, q_{17}\}$

ε closure [mov(B, 1)] = C

Q: Find the RE for the given language: set of all strings of 0's and 1's beginning with 0 ending with 1. $0(0+1)^*1$

Q: $\Sigma = (a,b)$ where no. of a's is divisible by 3. $(b^*ab^*ab^*a)^*$

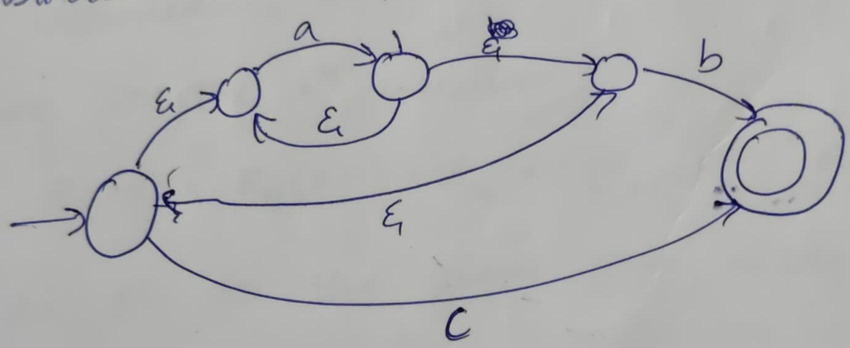
Q: ~~[illegible]~~ $w \in \{0,1^*$ | w does not contain 00 and is not empty }

~~[illegible]~~ $1^* + 1^*0 + 1^*01^*$

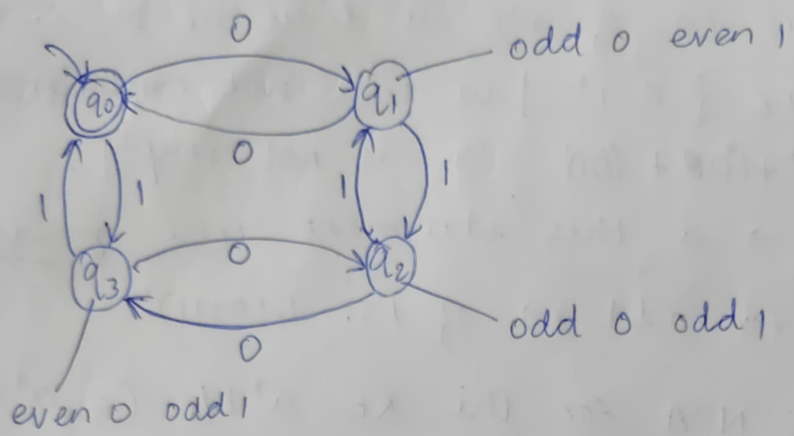
Q: $(r+s)^* \equiv r^* + s^*$ → is this statement true or false?

Q: Write the RE for odd no. of 1's. $1.(0+11)^*$

Q: Construct the ε NFA for the RE a^*b/c ($\equiv a^*b+c$)



Q: Even no. of 0's and even no. of 1's.



04.02.2025

RE equation method $R_{ij}(k)$

For every DFA $A = (Q, \Sigma, \delta, S, \{F\})$ there is a RE 'R' such that $L(R) = L(A)$

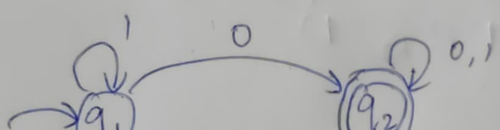
for n states (applying inductive method)

$$\underbrace{R_{ij}(K)}_{\text{RE from } i \text{ to } j \text{ passing through } k \text{ intermediate states}} = L\left(\bigcup_{j \in F} \underbrace{R_{1j}(n)}_{\text{from start to reach each final state}}\right) = L(A)$$

Expanding for the k^{th} transition,

$$R_{ij}(K) = R_{ij}(K-1) + R_{ik}(K-1)(R_{KK}(K-1))^* R_{Kj}(K-1)$$

Q: Find the RE for the given DFA using eq^n method.



Basis step: $k = 0 \Rightarrow R_{ij}^{(0)} = \boxed{R_{12}^{(0)} = a}$ → moving from one state to another ↳ input symbol is must

$\boxed{R_{11}^{(0)} = a + \varepsilon}$ ↳ staying in same state, input symbol is not compulsory

Induction step: $k \geq 1$

$$\Rightarrow \boxed{R_{ij}^{(K)} = R_{ij}^{(K-1)} + R_{ik}^{(K-1)} \cdot (R_{KK}^{(K-1)})^* \cdot R_{Kj}^{(K-1)}}$$

$R_{12}^{(2)}$ → no. of states

$$R_{12}^{(2)} = R_{12}^{(1)} + R_{11}(1) \cdot (R_{11}^{(1)})^* \cdot R_{12}^{(1)}$$

R_{12}: 1 → start state, 2 → final state

$\cancel{R_{12}^{(2)}}$

(k=0)

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

$R_{12}^{(0)} = 0$

$R_{21}(0) = \phi$

$R_{22}(0) = 0 + 1 + \varepsilon$

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

(K=1)

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

$R_{11}^{(1)} = (1+\varepsilon)\left[\varepsilon + \underbrace{(1+\varepsilon)^* (1+\varepsilon)}_{\lambda + R^* R = R^*}\right]$

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

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

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

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

$\quad\ \lambda \qquad R \qquad R^*$

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

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

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

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

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

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

k=2 → for k=n, do only for required expression

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

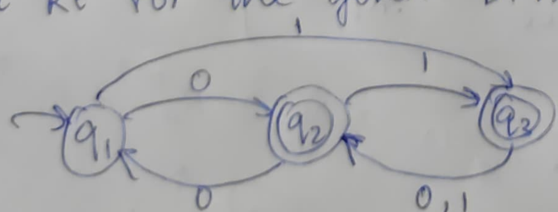
$= \cancel{0+(1+\varepsilon)(1+\varepsilon)^* 0}$

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

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

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

Q: Find the RE for the given DFA using eqn method $R_{ij}^{(n)}$



$\Rightarrow RE = R_{12}^{(3)} + R_{13}^{(3)}$

k=0 →

$R_{11}^{(0)} = \varepsilon$	$R_{21}^{(0)} = 0$	$R_{31}^{(0)} = \varepsilon$
$R_{12}^{(0)} = 0$	$R_{22}^{(0)} = \varepsilon$	$R_{32}^{(0)} = 0+1$
$R_{13}^{(0)} = 1$	$R_{23}^{(0)} = 1$	$R_{33}^{(0)} = \varepsilon$

$k=1 \Rightarrow R_{11}(1) = R_{11}^{(0)} + R_{11}^{(0)} ($