Unit-2 Regular Expression

• Regular expression is nothing but a way of representing Regular language.

Ex:- $a^2+b^2+2ab = (a+b)^2$

~~Q) L = {a, aa, aaa, aa~~

R.E.

Kleen Clouser = [*] = A^*

Positive Clouser = [+] = A^+

Concatenation = [.] = $A.B$

Union = [+] = $A+B$

Language to Expression

Q) $L = \{a, b, aa, bb, aaa, bbb \ldots\}$

R.E = $(a,b)^+$

Q) $L = \{\epsilon, a, b, aa, bb, aaa, bbb \ldots\}$

R.E = $(a,b)^*$

Q) $L = \{a\}$

R.E = a

Q) $L = \{a, b\}$

R.E = $(a+b)$

Q) $L = \{a, b, c\}$

R.E = $a+b+c$

Expression to Language

Q) $r(\phi) = L(r) = \{\}$

$r(\epsilon) = L(r) = \{\epsilon\}$

$r(a) = L(r) = \{a\}$

$r(a+b) = L(r) = \{a, b\}$

$r(a+b+c) = L(r) = \{a, b, c\}$

$R(a.b) = L(r) = \{ab\}$

$R(a^+) = L(r) = \{a, aa, aaa, \ldots\}$

$R(a^*) = L(r) = \{\epsilon, a, aa, aaa, \ldots\}$

$R(a^+)^+ = L(r) = R(a^+)$

$R(a^*)^* = R(a^*)$

$r(a^*.a^*) = L(r) = a^*$

$r(ab+a).b$:

$L(r) = \{abb, ab\}$

$r(a+b) = L(r) = L(a, b)$

$r(a+b) = L(a, b)$

$r(a+b)^2 = L(a,b)(a,b) = \{aa, ab, ba, bb\}$

$r(a+b)^* = \{\epsilon, a, b, aa, bb, aaa, bbb, \ldots\} \Rightarrow (a^*+b^*)^*$

$r(a+b)^*(a,b) = L(r) = \{a, b\} = (a+b)^+$

$r(ab)^* = L(r) = \{\epsilon, ab, abab, \ldots\ldots\}$

Precedence Rule of R.E.

BODMAS

( ) → Braces

$*$ → Klnee clousre

$+$ → Positive clousure

$\cdot$ → Concatenation.

$+$ → Union

Identities of R.E

1) $\phi + R = R$
2) $\phi R = R\phi = \phi$
3) $\epsilon R = R\epsilon = R$
4) $\epsilon^* = \epsilon$ & $\phi^* = \epsilon$
5) $R + R = R$
6) $R^* R^* = R^*$
7) $RR^* = R^* R$
8) $(R^*)^* = R^*$
9) $\epsilon + RR^* = \epsilon + R^*.R = R^*$
10) $(PQ)^*.P = P.(QP)^*$
11) $(P+Q)^* = (P^* Q^*)^* = (P^* + Q^*)^*$
12) $(P+Q).R = PR + QR$ and.
    $R(P+Q) = RP + RQ$.

Arden's Theorem

Design R.E.

① Starting with ab.
R.E $ab \cdot (a+b)^*$

Starting with abb
R.E $abb(a+b)^*$

② Ending with abb
R.E $(a+b)^* \cdot abb$

③ Contain substring abb
R.E $(a+b)^* abb (a+b)^*$

④ Starting & ending with a
$a \cdot (a+b)^* \cdot a$

⑤ Starting & Ending symbol same.
$\Sigma = \{a, b\}$
R.E $(a(a+b)^* a) + (b(a+b)^* \cdot b)$

Starting & Ending with different symbol.
$\Sigma = \{a, b\}$
R.E $(a(a+b)^* b) + (b(a+b)^* a)$

⑥ $|W| = 3$ $\Sigma = \{a, b\}$
R.E
$(a+b)(a+b)(a+b)$
(or)
$(a+b)^3$

$|W| \le 3$ $\Sigma = \{a, b\}$
R.E $(\epsilon+a+b)(\epsilon+a+b)(\epsilon+a+b)$
(or)
$(\epsilon+a+b)^3$

$|W| \geq 3$ $\Sigma = \{a, b\}$
R.E $(a+b)^3 (a+b)^*$

⑦ $|W|_a = 2$
This means there must be Two a's
R.E $b^* a b^* a b^*$

$|W|_a \le 2$
R.E $b^*(a+\epsilon) b^* (a+\epsilon) b^*$

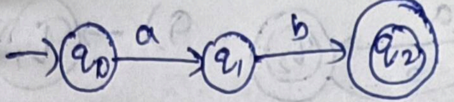
$|W|_a \geq 2$
R.E
$(a+b)^* a (a+b)^* a (a+b)^*$

⑧ 3rd symbol from left end is b
R.E
$(a+b)(a+b) b (a+b)^*$

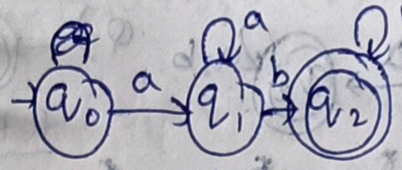
⑨ 28th symbol from right is a
$(a+b)^* a (a+b)^{27}$

$$\boxed{\begin{array}{c}\text{Arden's Theorem}\\ R = Q + RP \\ R = QP^*\end{array}}$$

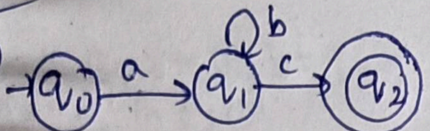
Finite Automata to R·E

Q) 

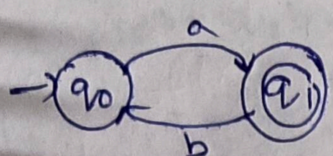
R.E ab.

Q) 

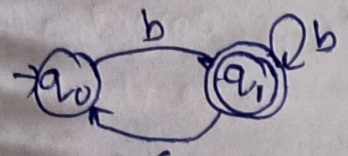
R·E a^*b^*

Q) 

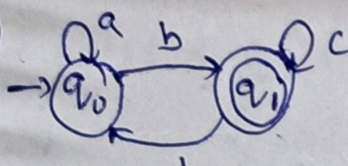
R·E ab^*c

Q) 

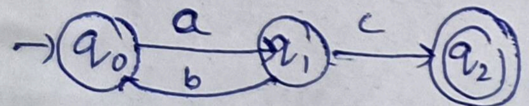
R·E $a(ba)^*$

Q) 

$b(b^*+(cb)^*)^*$

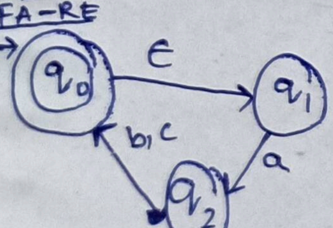
Q) 

R·E $a^*b(c^*+da^*b)^*$

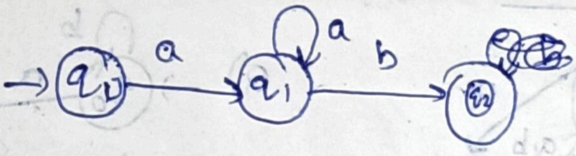
Q) 

R·E $a(ba)^*c$

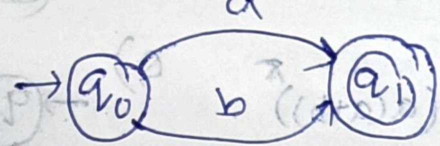
ENFA-RE

Q) 

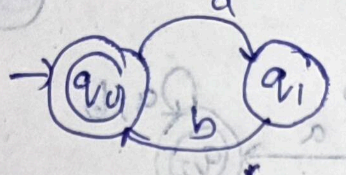
RE $(\epsilon\, a(b+c))^*$

Q) 

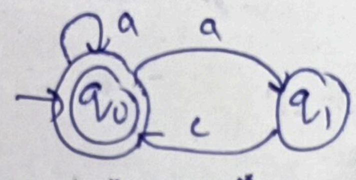
R·E aa^*b

Q) 

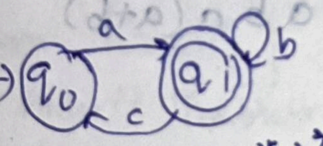
R·E $(a+b)$

Q) 

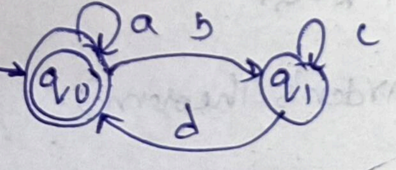
R·E $(a\cdot b)^*$

Q) 

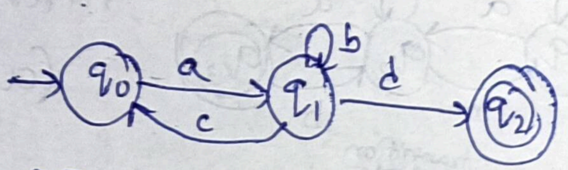
R·E $a^*+(ac)^* \Rightarrow (a+(ac))^*$

Q) 

$a(b^*+(ca)^*)^*$

Q) 

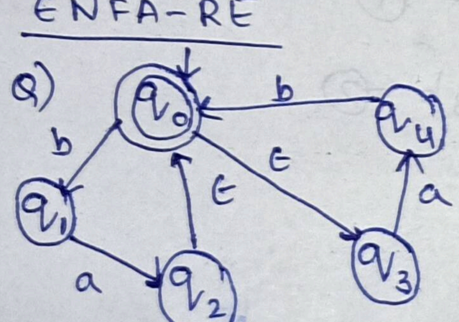
R·E $\{a^*+(bc^*d)^*\}^*$ (or) $(a+bc^*d)^*$

Q) 

R·E $a(b^*+(ca)^*)d$ (or)

$a(b+ca)^*d$.

ENFA-RE

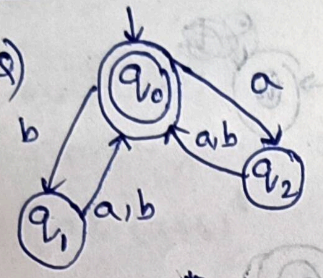
Q) 

$(ba\epsilon)^*+(\epsilon ab)^*$

(or)

$(ba\epsilon+\epsilon ab)^*$

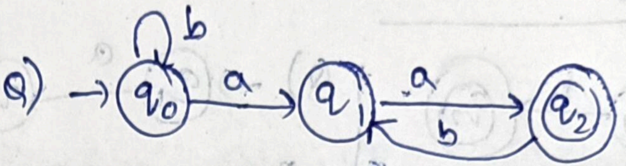
NFA to RE

q)


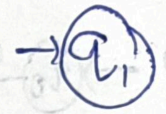
$(b(a+b))^* + (a(a+b))^*$

(or)

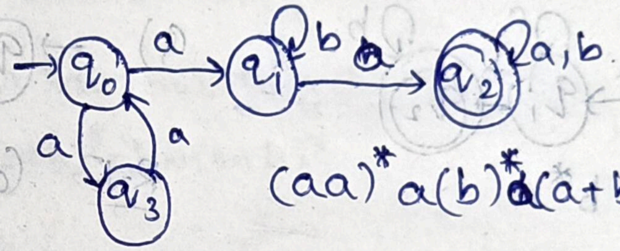
$(b(a+b) + a(a+b))^*$

q)


$b^* aa(ba)^*$

q)


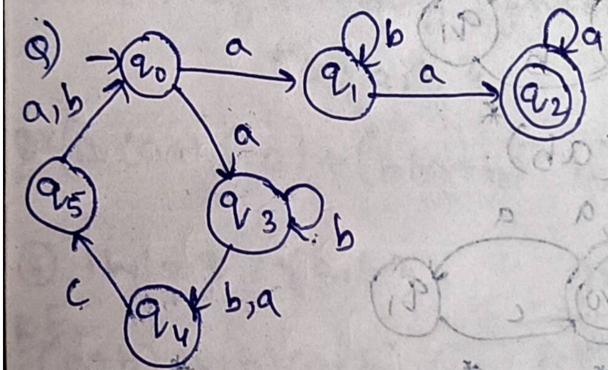
R.E ϕ

q)


$(aa)^* a(b)^* a(a+b)^*$

q)


R.E ϵ

q)


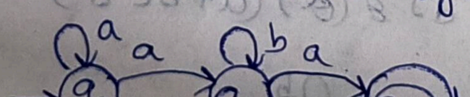
$(ab^*(b+a)c(a+b)^*)^* ab^* a(a+b)^*$

Arden's Theorem

$R = Q + RP$ then $R = QP^*$

FA to R.E (using Arden's Theorem)

q) Consider transition system



After proving we get

$(a + a(b+ab)^* b)^* a(b+ab)^* a$.

Incoming transition states

$q_1 = \epsilon + q_1 a + q_2 b$ — ①

$q_2 = q_1 a + q_2 b + q_3 b$ — ②

$q_3 = q_2 a$ — ③

~~Eq using ③④②~~

$q_3 = q_2 a$

Put value of q_2 in eq ③

$q_3 = (q_1 a + q_2 b + (q_2 a) b) a$

$q_3 = (q_1 a + q_2 (b + ab)) a$

$q_3 = q_1 aa + q_2 (ba + aab)$ —— ④

Eq ②

$q_2 = q_1 a + q_2 b + q_3 b$

put q_3 in Eq ②

$q_2 = q_1 a + q_2 b + (q_2 a) b$

$q_2 = q_1 a + q_2 b + q_2 ab$

$$\underbrace{q_2}_{R} = \underbrace{q_1 a}_{Q} + \underbrace{q_2}_{R} \underbrace{(b + ab)}_{P}$$

$q_2 = q_1 a (b + ab)^*$ —— ⑤

[From $R = Q + RP$
then $R = QP^*$]
By Arden's Theorem.

~~Eq①~~
Choose q_1

$q_1 = \epsilon + q_1 a + q_2 b$

Put value of q_2 in above Eq.

$q_1 = \epsilon + q_1 a + (q_1 a (b + ab)^*) b$

$$\underbrace{q_1}_{R} = \underbrace{\epsilon}_{Q} + \underbrace{q_1}_{R} \underbrace{(a + (a(b + ab)^*) b)}_{P}$$

$q_1 = \epsilon (a + \{a (b + ab)^* b\})^*$

[$R = Q + RP = QP^*$
by Arden's Theorem]

$q_1 = (a + a(b + ab)^* b)^*$ —— ⑥

$q_3 = q_2 a$

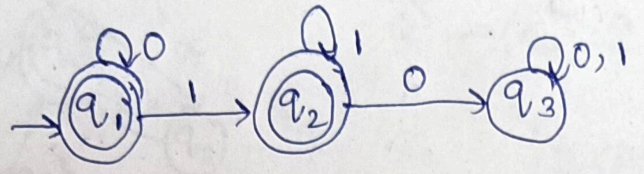
Put value of q_2

$q_3 = (q_1 a (b + ab)^*) a$

Put value of q_1 in above eq.

$q_3 = (a + a(b+ab)^* b)^* (a(b+ab)^*)\, a$.

Q)



$q_1 = \epsilon + q_0$ — ①

$q_2 = q_1 1 + q_2 1$ — ②

$q_3 = q_2 0 + q_3 0 + q_3 1$ — ③

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

Now taking

$\underset{R}{q_1} = \underset{Q}{\epsilon} + \underset{R}{q_1}\underset{P}{0}$

$q_1 = \epsilon 0^*$ [By Arden's Theorem $R = Q + RP = QP^*$]

$\boxed{q_1 = 0^*}$

Now

$q_2 = q_1 1 + q_2 1$

Put ~~q~~ value of q_1 in q_2

$\underset{R}{q_2} = \underset{Q}{(0^*)1} + \underset{R}{q_2}\underset{P}{1}$ (By Arden's Theorem)

$\boxed{q_2 = (0^* 1 1^*}$

Because q_3 is ~~not~~ a nonproductive state so we are not solving q_3.

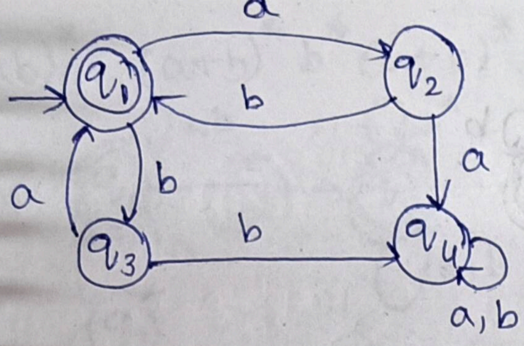
q_1 & q_2 are final so we make union $q_1 + q_2$

$q_1 + q_2 = 0^* + 0^* 1 1^*$ [$\epsilon + RR^* = R^*$]

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

$\boxed{q_1 + q_2 = 0^* 1^*}$

8)



$\epsilon + ((ab)^* + (ba)^*)$

$= (ab+ba)^*$

$q_1 = \epsilon + q_2 b + q_3 a$ — ①

$q_2 = q_1 a$ — ②

$q_3 = q_1 b$ — ③

Eq ② & Eq ③ in Eq ①

$q_1 = \epsilon + q_1 ab + q_1 ba$

$q_1 = \epsilon + q_1 (ab + ba)$

$\underbrace{q_1}_{R} = \underbrace{\epsilon}_{Q} + \underbrace{q_1}_{R} \underbrace{(ab+ba)}_{P}$

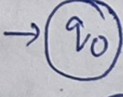
(By Arden's $R = Q + RP = QP^*$)

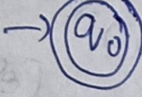
$q_1 = \epsilon (ab+ba)^*$

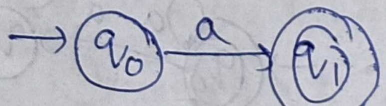
$q_1 = \epsilon (ab+ba)^*$

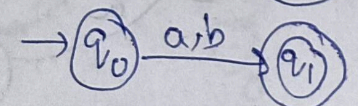
$\boxed{q_1 = (ab+ba)^*}$

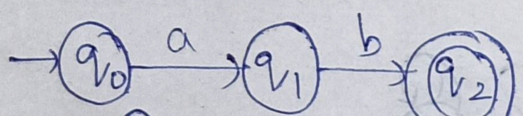
R.E to FA

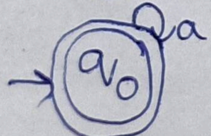
$\phi =$ 

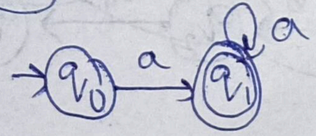
$\epsilon =$ 

$a =$ 

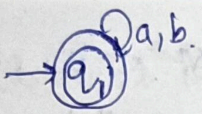
$a+b =$ 

$a \cdot b =$ 

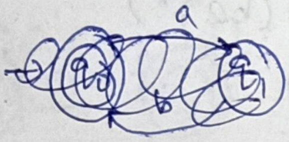
$a^* =$ 

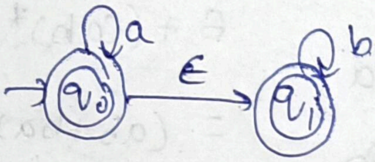
$a^+ =$ 

$(a+b)^*$

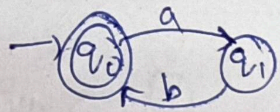


a^*b^*

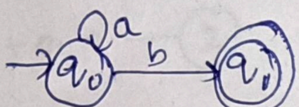




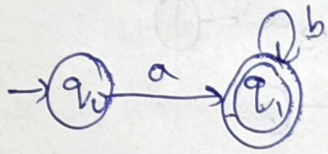
$(ab)^*$



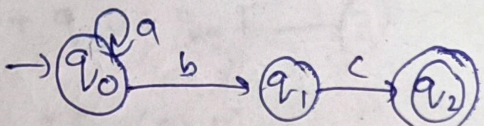
a^*b



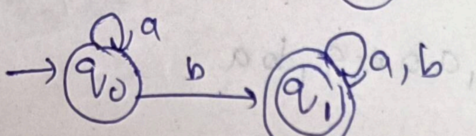
ab^*



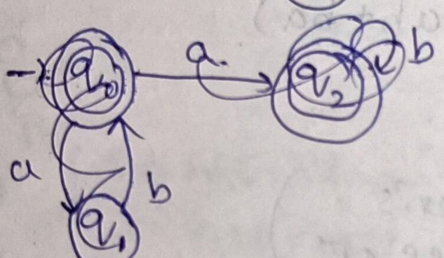
a^*bc

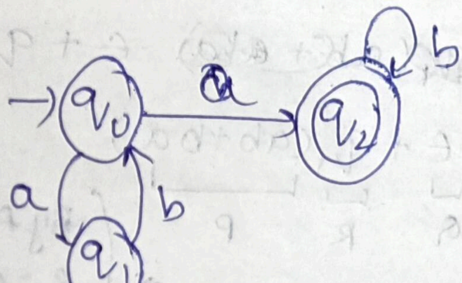


$a^*b(a+b)^*$

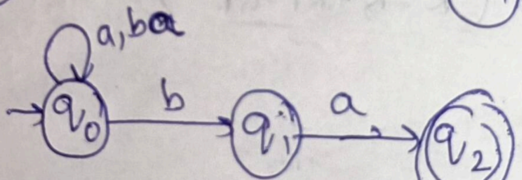


$(ab)^*ab^*$

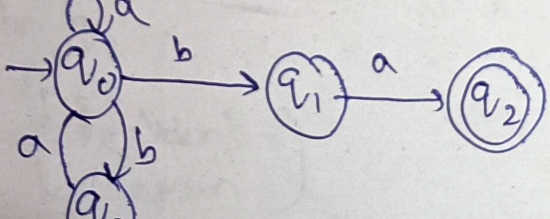




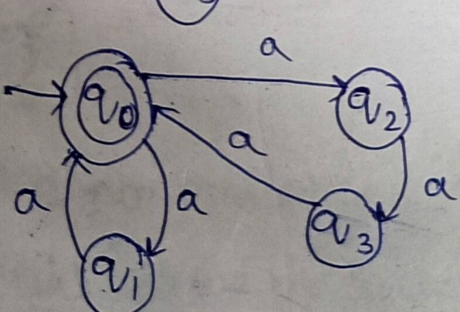
$(a+ba)^*ba$



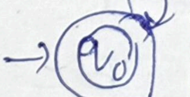
⇓

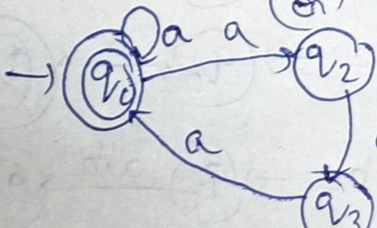


$(aa+aaa)^*$

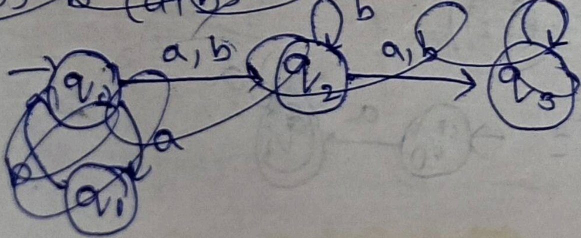


$(a+aaa)^*$

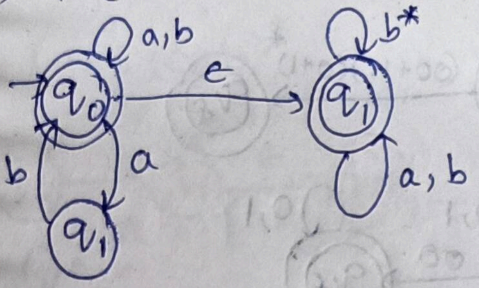




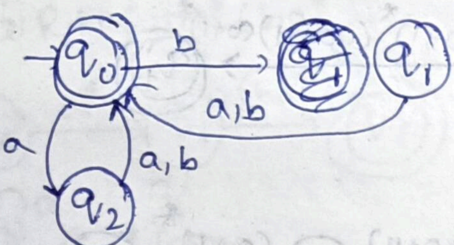
~~(ab)* + (a+b)* b* (a+b)*~~



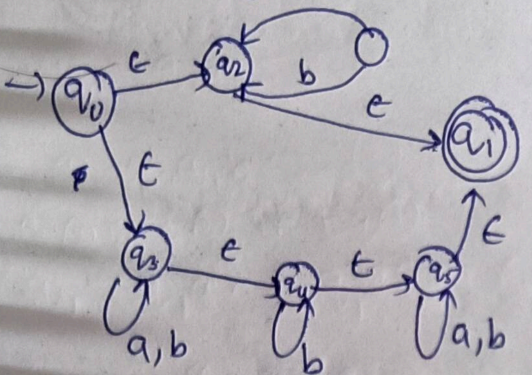
$(ab)^* + (a+b)^* b^* (a+b)^*$



$[b(a+b) + a(a+b)]^*$



(Q)

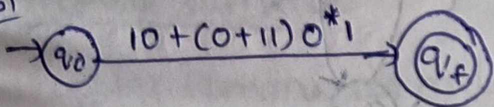


R.E to FA

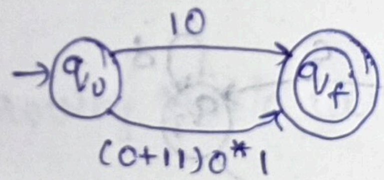
Elimination Method

Q) $10 + (0+11)0^*1$

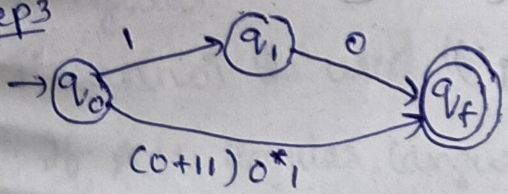
Step1



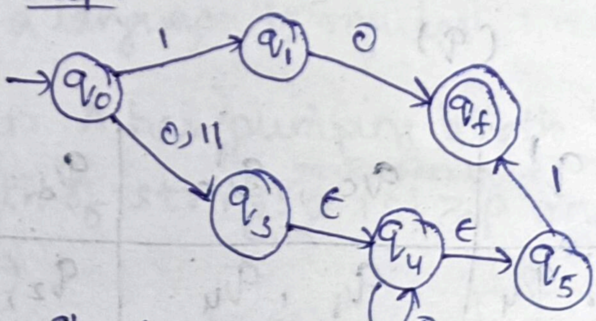
Step2



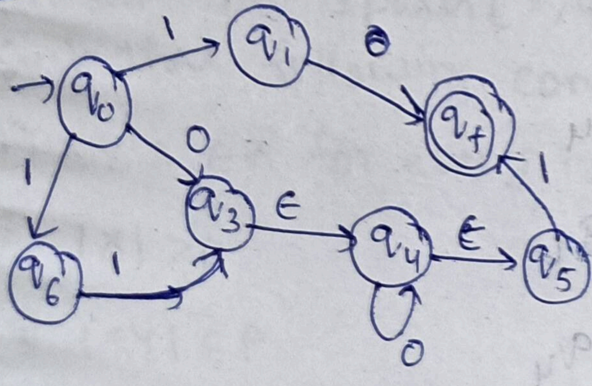
Step3



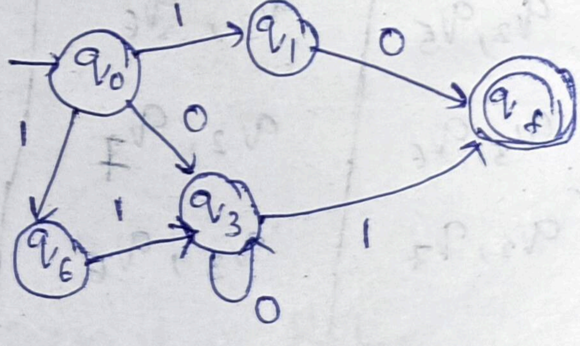
Step4



Step5

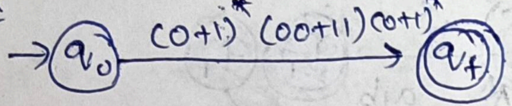


Step6

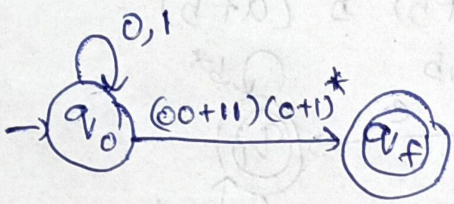


Q) $(0+1)^*(00+11)(0+1)^*$

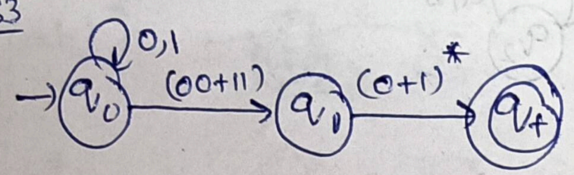
S1



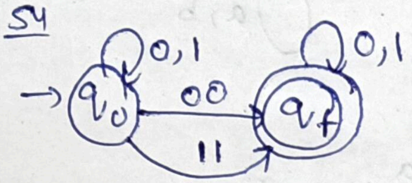
S2



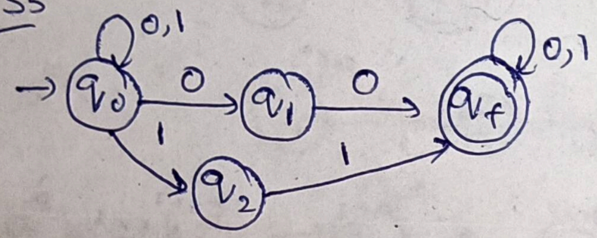
S3



S4

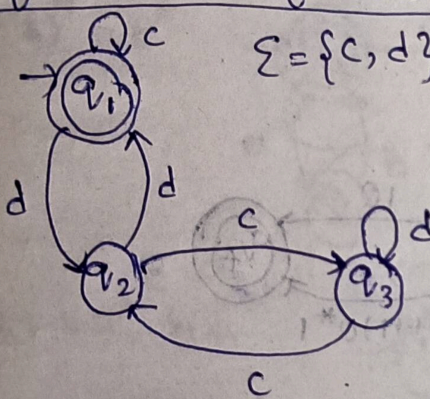


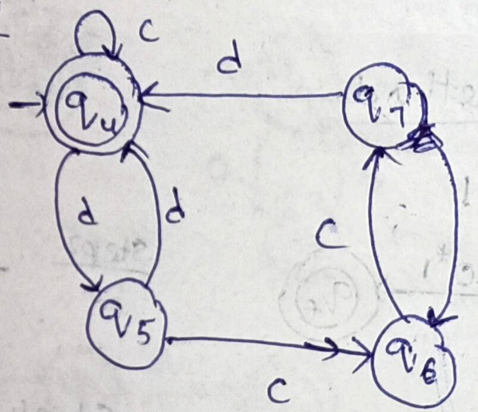
S5



Equivalence of Two DFA

Q) $\Sigma = \{c, d\}$





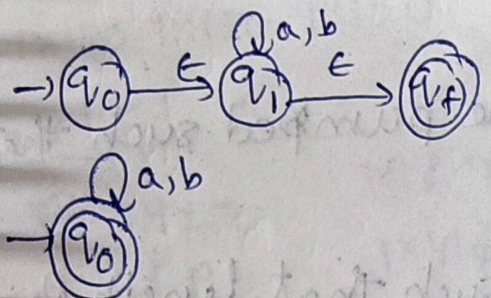
(q) (q')

q, q'	q_c, q'_c	q_d, q'_d
q_1, q_4	q_1, q_4	q_2, q_5
q_2, q_5	q_3, q_6	q_1, q_4
q_3, q_6	q_2, q_7	q_3, q_6
q_2, q_7	q_3, q_6	q_1, q_4

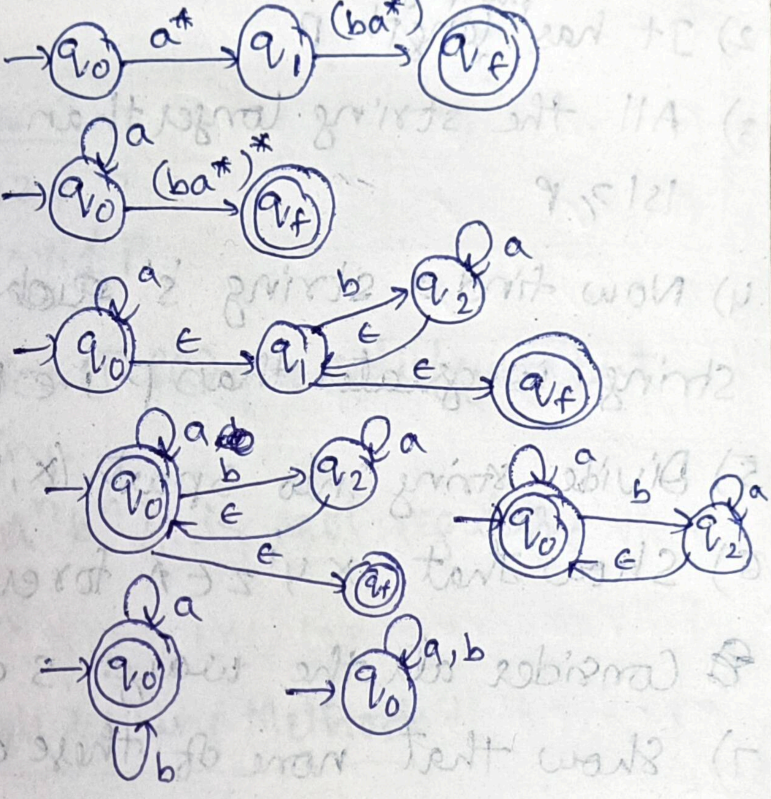
Here we started with q_1, q_4 and also we are getting the final states same as end and there is ~~not~~ no new states coming. So they are equivalent.

Q) $(a+b)^* = a^*(ba^*)^*$ (Check this)

Take L.H.S $(a+b)^*$



Take R.H.S $a^*(ba^*)^*$



$\therefore$ LHS = RHS

$\therefore (a+b)^* = a^*(ba^*)^*$

Pumping Lemma

* Pumping lemma is only use for proving that a language is not regular.
* It cannot be used to prove a language is regular.

Def:- If A is Regular language then A has pumping length 'p' such that any string 's' where length of strings greater than p i.e $|s| \geq p$ may be divided into 3 parts (x, y, z).
such that following conditions are true

$s \rightarrow x, y, z$

① $xy^iz \in A$ for every $i \geq 0$

② $|x| > 0$

③ $|xy| \leq p$

If these 3 conditions are true then it is R.L.

Steps for solving pumping lema

1) Assume A is regular ~~it has~~
2) It has pumping length P
3) All the string longer than p can not be pumped such that $|s| \geq P$
4) Now find a string 's' ~~such~~ in A such that length of string is greater than p i.e $|s| \geq P$
5) Divide string into 3 parts (x, y, z) ~~there show that~~
6) Show that $xy^iz \in A$ for every $i \geq 0$

Consider all the ways s can be divided into x y z

7) Show that none of these can satisfy all the 3 pumping conditions at same time.

Q) $A = \{a^n b^n \mid n \geq 0\}$

Proof

Assume A is regular

$P = a^P b^P$ Let $P = 7$

~~Case~~ $a^7 b^7$

s → x y z

$S = \underbrace{a}_{x}\underbrace{aaaaa}_{y}\underbrace{abbbbbbb}_{z}$

Case 1 Y is in b part

$\underbrace{aaaaaaa\,b}_{x}\underbrace{bbbbb}_{y}\underbrace{b}_{z}$

Case 2

Y is in a part

$\underbrace{aa}_{x}\underbrace{aaa}_{y}\underbrace{aabbbbbbb}_{z}$

Case 3

Y is in ab part

$\underbrace{aaaaa}_{x}\underbrace{aabb}_{y}\underbrace{bbbbb}_{z}$

$xy^iz \in A$

Assume value of $i = 2$

Case 1 xy^2z

$\underbrace{aaaaaaab}_{x}\underbrace{bbbbbbbb}_{y}\underbrace{bb}_{z}$

Case 2 xy^2z

$\underbrace{aa}_{x}\underbrace{aaaaaaa}_{y}\underbrace{abbbbbbb}_{z}$

Case 3 xy^2z

$\underbrace{aaaaa}_{x}\underbrace{aabbaabb}_{y}\underbrace{bbbbb}_{z}$

~~1x1~~>0

Case 1 Case 2 Case 3

$|x|>0$ $|x|>0$ $|x|>0$

Case 1

~~...~~ $9 \neq 7$

$|xy| \notin P$

Case 2

$10 \neq 7$

$|xy| \notin P$

Case 3

$13 \neq 7$

$|xy| \notin P$

Here in this question condition ① & ② are satisfying but condition ③ is not satisfying.

so Given expression $A = \{a^n b^n\}$ is not regular.

And it doesn't follow pattern

Minimization of DFA by Table Filling Method:

(Myhill Neorode theorem)

steps

1) Draw A table for all pairs of P, Q

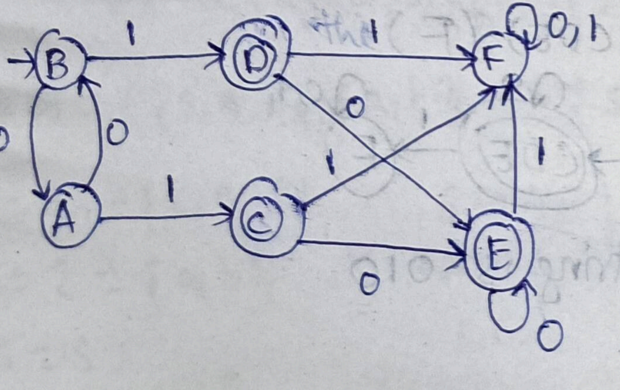
2) Mark all pairs where $P \in F$ and $Q \notin F$

3) If there are any unmarked pairs in (P,Q) such that $[\delta(P,x), \delta(Q,x)]$ is marked then mark (P,Q)

(x is any input symbol) Repeat until no more marking can be made.

4) Combine all the unmarked ~~pairs~~ pairs and make them a single state in the minimized DFA.

Q)



~~A~~	A	B	C	D	E	F
A						
B						
C	✓	✓				
D	✓	✓				
E	✓	✓				
F	✓	✓	✓	✓	✓	

Both are final → no mark

Both are non-final → no mark

One final & one non final → mark

S-2

(B, A) = $\delta(B,0) = A$ } No mark
$\delta(A,0) = B$

$\delta(B,1) = D$ } No mark
$\delta(A,1) = C$

(D, C) = $\delta(D,0) = E$ } No mark
$\delta(C,0) = E$

$\delta(D,1) = F$ } No Mark
$\delta(C,1) = F$

(E, C) = $\delta(E,0) = E$ } No
$\delta(E,0) = E$

$\delta(E,1) = F$ } No
$\delta(C,1) = F$

(E, D) = $\delta(E,0) = E$ } No
$\delta(D,0) = E$

$\delta(E,1) = F$ } No
$\delta(D,1) = F$

(F, A)
$\delta(F,0) = F$
$\delta(A,0) = B$

$\delta(F,1) = F$ } Mark
$\delta(A,1) = C$

(F, B)

$\delta(F,0) = F$

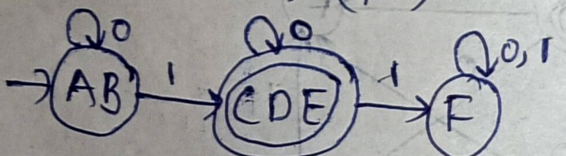
$\delta(B,0) = A$

$\delta(F,1) = F$ } mark
$\delta(B,1) = D$

S-3

(B, A) (D, C) (E, C) (E, D) Make then one pair

(AB) (CDE) (F)



Accepted string is 010