Regular Expression

Regular Expressions are used for representing certain sets of strings in an algebraic fashion.

1) Any terminal symbol i.e. symbols $\in \Sigma$ including $\wedge$ and Φ are regular expressions. — $a, b, c, \ldots \wedge, \phi$

   ϕ = Null

   $\wedge$ → empty

2) The Union of two regular expressions is also a regular expression. — $R_1, R_2 \quad (R_1 + R_2)$

3) The Concatenation of two regular expressions is also a regular expression. — $R_1, R_2 \rightarrow (R_1 . R_2)$

4) The iteration (or Closure) of a regular expression is also a regular expression. — $R \rightarrow R^*$ $\quad a^* = \wedge, a, aa, aaa, \ldots$

5) The regular expression over Σ are precisely those obtained recursively by the application of the above rules once or several times.

Regular Expression - Examples

Describe the following sets as Regular Expressions

1) $\{0,1,2\}$ — 0 or 1 or 2

   $R = 0 + 1 + 2$

2) $\{\wedge, ab\}$

   $R = \wedge\ ab$

3) $\{abb, a, b, bba\}$ — abb or a or b or bba

   $R = abb + a + b + bba$

4) $\{\wedge, 0, 00, 000, \ldots\}$ — closure of 0

   $R = 0^*$

   $\wedge = \epsilon$

5) $\{1, 11, 111, 1111, \ldots\}$

   $R = 1^+$

Identities of Regular Expression

1) $\varnothing + R = R$
2) $\varnothing R + R\varnothing = \varnothing$
3) $\epsilon R = R\epsilon = R$
4) $\epsilon^* = \epsilon$ and $\varnothing^* = \epsilon$
5) $R + R = R$
6) $R^*R^* = R^*$

$R^nR^* = R^*R^n = R^+$

$R + R^2 + R^3 + \ldots + R^n + R^* = R^+$

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

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

R^+

$R^*R = R^+$

ARDEN'S THEOREM

If P and Q are two Regular Expressions over Σ, and if P does not contain ϵ, then the following equation in R given by $R = Q + RP$ has a unique solution i.e. $R = QP^*$

$$R = Q + RP \longrightarrow (1)$$

$$= Q + QP^*P$$

$$= Q(\epsilon + P^*P)$$

$$= QP^*$$

Proved

$R = QP^*$

$[\epsilon + R^*R = R^*]$

$$R = Q + RP$$
$$= Q + [Q + RP]P$$
$$= Q + QP + RP^2$$
$$= Q + QP + [Q + RP]P^2$$
$$= Q + QP + QP^2 + RP^3$$
$$\vdots$$
$$= Q + QP + QP^2 + \ldots QP^n + RP^{n+1}$$

$$\boxed{R = QP^*}$$

$$= Q + QP + QP^2 + \ldots QP^n + QP^* P^{n+1}$$
$$= Q[\epsilon + P + P^2 + \ldots P^n + P^* P^{n+1}]$$

$$= Q + [Q + RP]P$$
$$= Q + QP + RP^2$$
$$= Q + QP + [Q + RP]P^2$$
$$= Q + QP + QP^2 + RP^3$$
$$\vdots$$
$$= Q + QP + QP^2 + \ldots QP^n + RP^{n+1}$$

$$\boxed{R = QP^*}$$

$$= Q + QP + QP^2 + \ldots QP^n + QP^* P^{n+1}$$

$$= Q\underbrace{[\epsilon + P + P^2 + \ldots P^n + P^* P^{n+1}]}_{P^*}$$

$$R = QP^*$$

An Example Proof using Identities of Regular Expressions

Prove that (1+00*1) + (1+00*1) (0+10*1)* (0+10*1) is equal to 0*1 (0+10*1)*

$$LHS = (1+00^*1) + (1+00^*1)(0+10^*1)^*(0+10^*1)$$

$$= (1+00^*1)\left[\epsilon + (0+10^*1)^*(0+10^*1)\right]$$

$$= (1+00^*1)(0+10^*1)^*$$

$$= (\epsilon . 1+00^*1)(0+10^*1)^*$$

$$= (\epsilon+00^*)1(0+10^*1)^*$$

$$= 0^*1(0+10^*1)^* = RHS$$

$$\epsilon + R^*R = R^*$$

$$\epsilon . R = R$$

Designing Regular Expressions - Examples (Part-1)

Design Regular Expression for the following languages over {a,b}

1) Language accepting strings of length exactly 2
2) Language accepting strings of length atleast 2
3) Language accepting strings of length atmost 2

Soln

1) $L_1 = \{aa, ab, ba, bb\}$

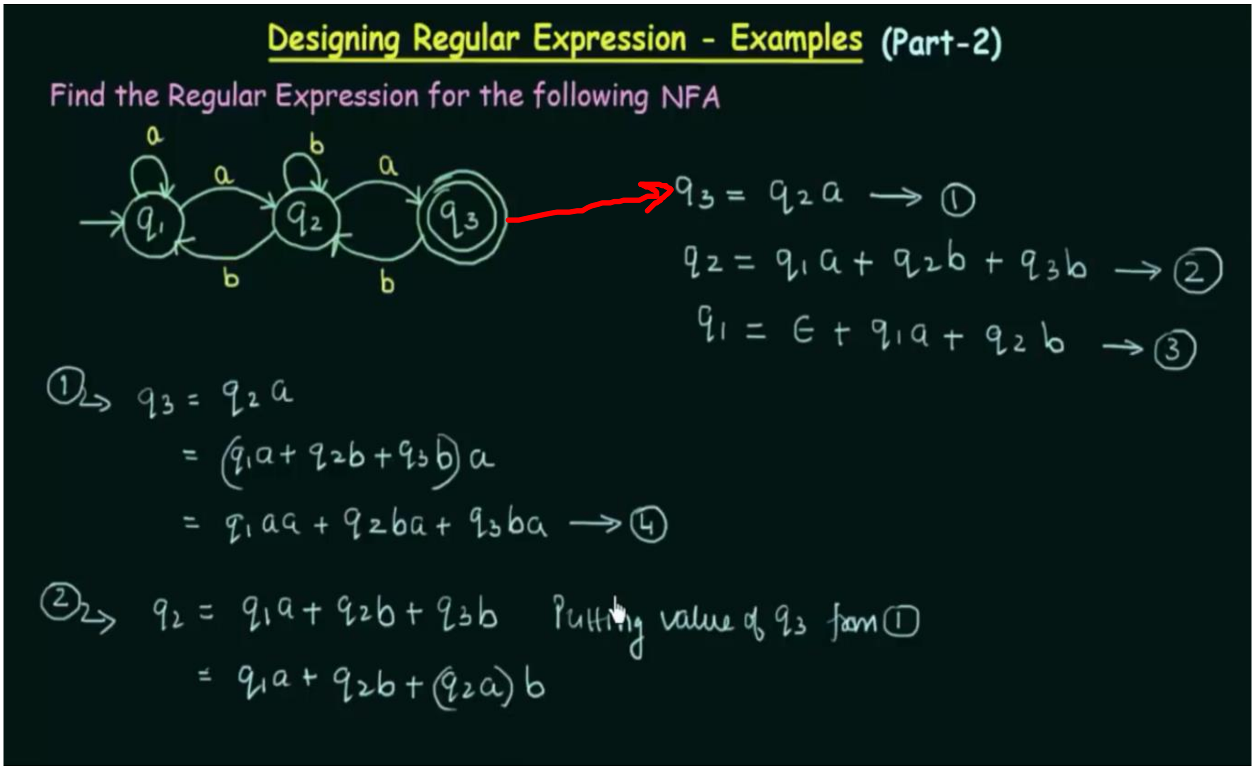
$$R = aa+ab+ba+bb$$
$$= a(a+b)+b(a+b)$$
$$= (a+b)(a+b)$$

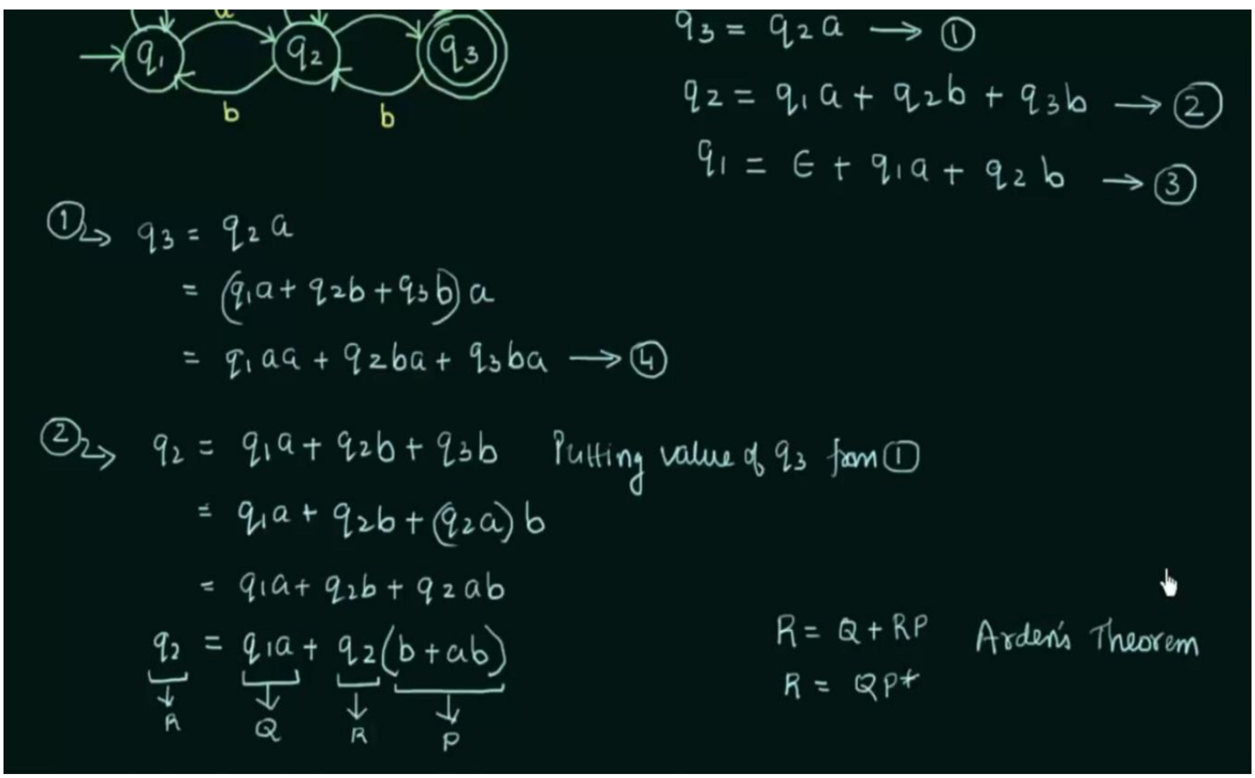
2) $L_1 = \{aa, ab, ba, bb, aaa, \ldots\}$

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

3) $L_1 = \{\epsilon, a, b, aa, ab, ba, bb\}$

$$R = \epsilon+a+b+aa+ab+ba+bb$$
$$= (\epsilon+a+b)(\epsilon+a+b)$$


Designing Regular Expression - Examples (Part-2)
Find the Regular Expression for the following NFA
q_1 q_2 q_3 a a b a b b
$q_3 = q_2 a \rightarrow$ (1)
$q_2 = q_1 a + q_2 b + q_3 b \rightarrow$ (2)
$q_1 = \epsilon + q_1 a + q_2 b \rightarrow$ (3)
(1) $\Rightarrow q_3 = q_2 a$
$= (q_1 a + q_2 b + q_3 b) a$
$= q_1 aa + q_2 ba + q_3 ba \rightarrow$ (4)
(2) $\Rightarrow q_2 = q_1 a + q_2 b + q_3 b$
Putting value of q_3 from (1)
$= q_1 a + q_2 b + (q_2 a) b$


q_1 q_2 q_3 b b
$q_3 = q_2 a \rightarrow$ (1)
$q_2 = q_1 a + q_2 b + q_3 b \rightarrow$ (2)
$q_1 = \epsilon + q_1 a + q_2 b \rightarrow$ (3)
(1) $\Rightarrow q_3 = q_2 a$
$= (q_1 a + q_2 b + q_3 b) a$
$= q_1 aa + q_2 ba + q_3 ba \rightarrow$ (4)
(2) $\Rightarrow q_2 = q_1 a + q_2 b + q_3 b$
Putting value of q_3 from (1)
$= q_1 a + q_2 b + (q_2 a) b$
$= q_1 a + q_2 b + q_2 ab$
$\underbrace{q_2}_{R} = \underbrace{q_1 a}_{Q} + \underbrace{q_2}_{R} \underbrace{(b + ab)}_{P}$
$R = Q + RP$ Arden's Theorem
$R = QP^*$

$$q_1 = \epsilon + q_1 a + q_2 b \rightarrow 3$$

1 $\Rightarrow$ $q_3 = q_2 a$

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

$$= q_1 aa + q_2 ba + q_3 ba \rightarrow 4$$

2 $\Rightarrow$ $q_2 = q_1 a + q_2 b + q_3 b$ Putting value of q_3 from 1

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

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

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

$R = Q + RP$ Arden's Theorem

$R = QP^*$

$$q_2 = (q_1 a)(b+ab)^* \rightarrow 5$$

$$q_2 = (q_1 a)(b+ab)^* \rightarrow 5$$

3 $\Rightarrow$ $q_1 = \epsilon + q_1 a + q_2 b$

Putting value of q_2 from 5

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

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

$R = Q + RP$

$R = QP^*$

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

$\epsilon \cdot R = R$

$$q_1 = (a + a(b+ab)^* b)^* \rightarrow 6$$

$R \quad Q \quad R \quad P$

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

$q_1 = (a+a(b+ab)^* b)^* \rightarrow (6)$

Final state (q_3)

$q_3 = q_2 a$

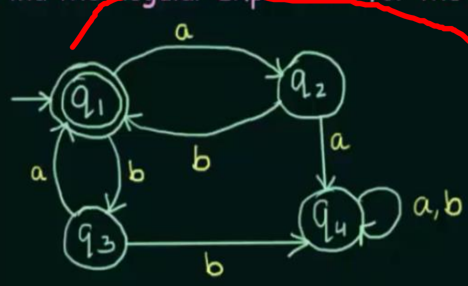
$= q_1 a(b+ab)^* a$ Putting value of q_2 from (5)

$q_3 = (a+a(b+ab)^* b)^* a(b+ab)^* a$ Putting value of q_1 from (6)

= Required Regular Expression for the given NFA

Designing Regular Expression - Examples (Part-3)

Find the Regular Expression for the following DFA



$q_1 = \epsilon + q_2 b + q_3 a \rightarrow (I)$

$q_2 = q_1 a \rightarrow (II)$

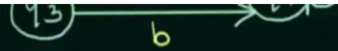
$q_3 = q_1 b \rightarrow (III)$

$q_4 = q_2 a + q_3 b + q_4 a + q_4 b \rightarrow (IV)$

(I) $\rightarrow$ $q_1 = \epsilon + q_2 b + q_3 a$

Putting values of q_2 and q_3 from (II) and (III)

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



$q_4 = q_2 a + q_3 b + q_4 a + q_4 b \rightarrow$ (IV)

(I) $\Rightarrow$ $q_1 = \epsilon + q_2 b + q_3 a$

Putting values of q_2 and q_3 from (II) and (III)

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

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

$R = Q + RP$

$R = QP^*$ Arden's theorem

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

$\epsilon . R = R$

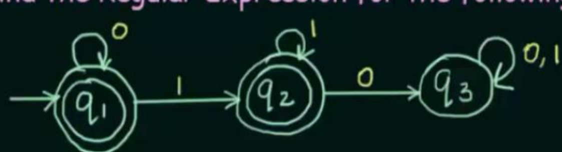
$q_1 = (ab + ba)^*$

$\Rightarrow$ Regular Expression .

Designing Regular Expression - Examples (Part-4)

(When there are Multiple Final States)

Find the Regular Expression for the following DFA



$q_1 = \epsilon + q_1 0 \rightarrow$ (I)

$q_2 = q_1 1 + q_2 1 \rightarrow$ (II)

$q_3 = q_2 0 + q_3 0 + q_3 1 \rightarrow$ (III)

Final state (q1)

(I) $\Rightarrow$ $$\underbrace{q_1}_{R} = \underbrace{\epsilon}_{Q} + \underbrace{q_1}_{R}\underbrace{0}_{P}$$

$R = Q + RP$

$R = QP^*$ Arden's theorem

$q_1 = \epsilon . 0^*$

$\epsilon . R = R$

$q_1 = 0^* \rightarrow$ (4)

$q_1 = 0^* \rightarrow$ (4)

Final state (q_2)

(11) $\hookrightarrow$ $q_2 = q_1 1 + q_2 1$

$\underbrace{q_2}_{R} = \underbrace{0^* 1}_{Q} + \underbrace{q_2}_{R}\underbrace{1}_{P}$ Putting value of q_1 from (4)

$R = Q + RP$

$R = QP^*$

$q_2 = 0^* 1 (1)^*$

R = union of both Final states

$= 0^* + 0^* 1 1^*$

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

Final State (q_2)

(11) $\hookrightarrow$ $q_2 = q_1 1 + q_2 1$

$\underbrace{q_2}_{R} = \underbrace{0^* 1}_{Q} + \underbrace{q_2}_{R}\underbrace{1}_{P}$ Putting value of q_1 from (4)

$R = Q + RP$

$R = QP^*$

$q_2 = 0^* 1 (1)^*$

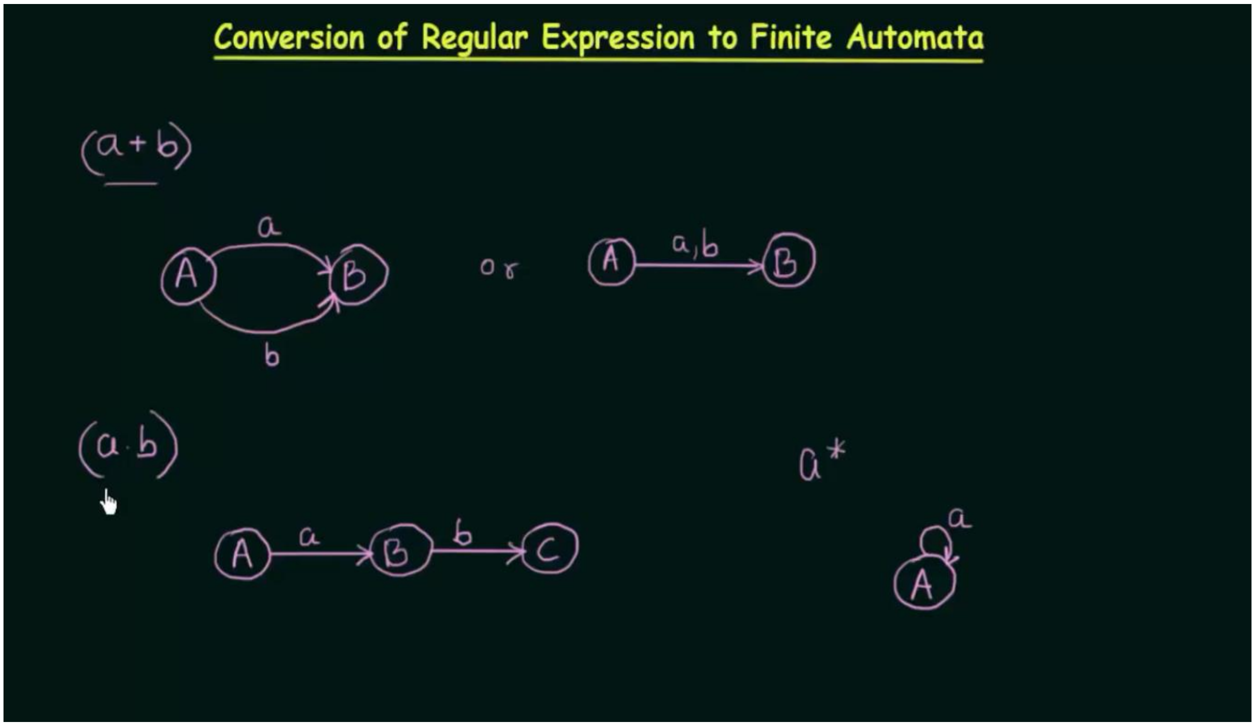
R = union of both Final states

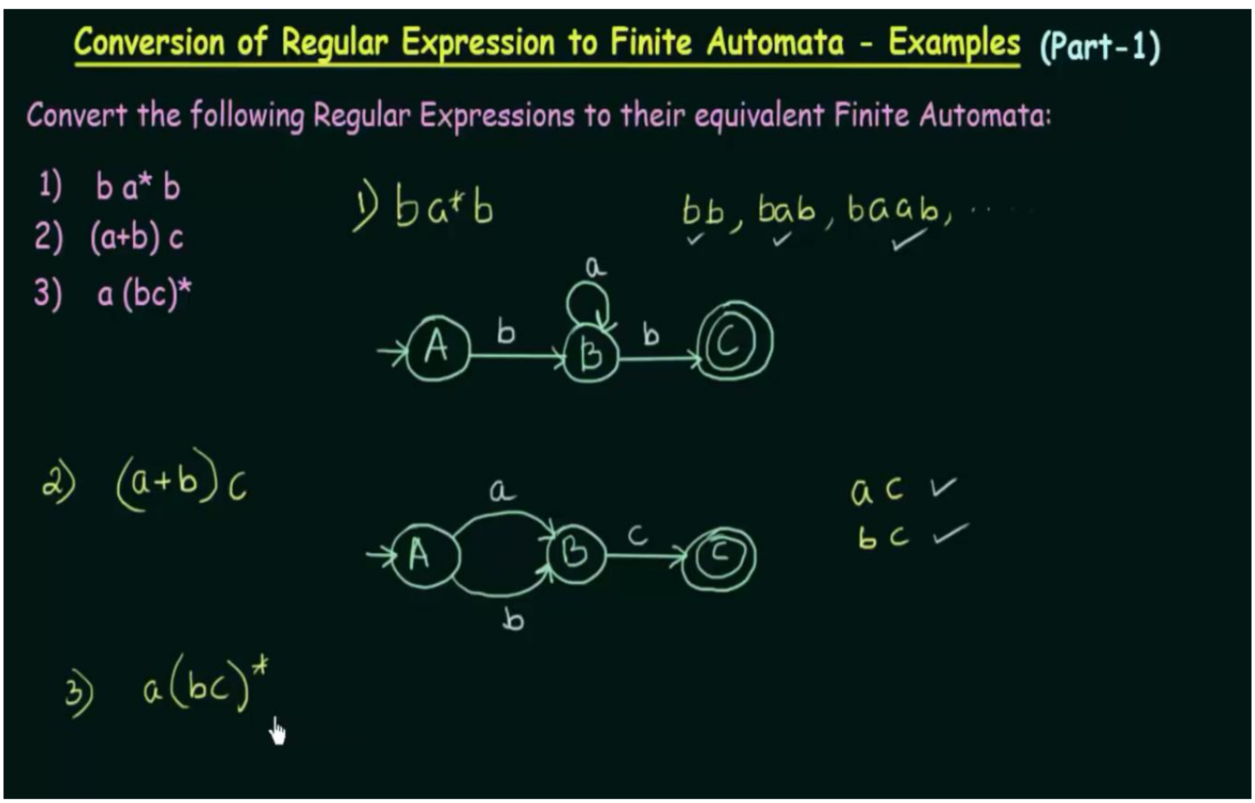
$= 0^* + 0^* 1 1^*$

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

$= 0^* 1^*$

$\hookrightarrow$ Regular Expression

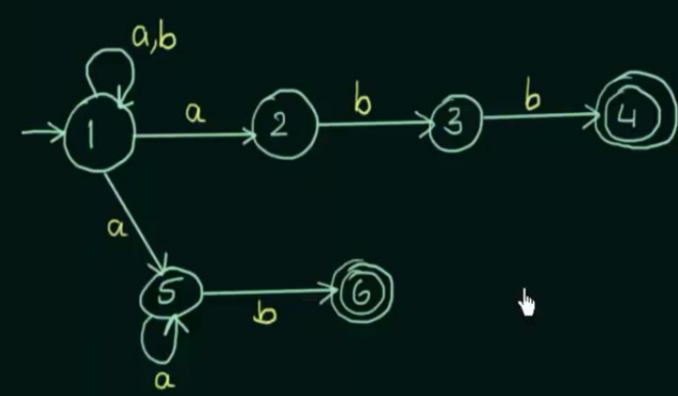

Conversion of Regular Expression to Finite Automata
(a+b)
a
A
B
b
or
A
a,b
B
(a·b)
A
a
B
b
C
a*
a
A


Conversion of Regular Expression to Finite Automata - Examples (Part-1)
Convert the following Regular Expressions to their equivalent Finite Automata:
1) b a* b
2) (a+b) c
3) a (bc)*
1) ba*b
bb, bab, baab, ...
a
A
b
B
b
C
2) (a+b)c
a
A
B
c
C
b
ac ✓
bc ✓
3) a(bc)*

Conversion of Regular Expression to Finite Automata - Examples (Part-2)

Convert the following Regular Expression to its equivalent Finite Automata:

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



$a^+ = \{a, aa, aaa, \ldots\}$

$a^* = \{\epsilon, a, aa, \ldots\}$

Conversion of Regular Expression to Finite Automata - Examples (Part-3)

Convert the following Regular Expression to its equivalent Finite Automata:

10 + (0 + 11) 0*1