Regular Expressions (RE) and Languages

Introduction:

1. Regular Language:

- A set of strings accepted by finite automata is known as regular language.

-

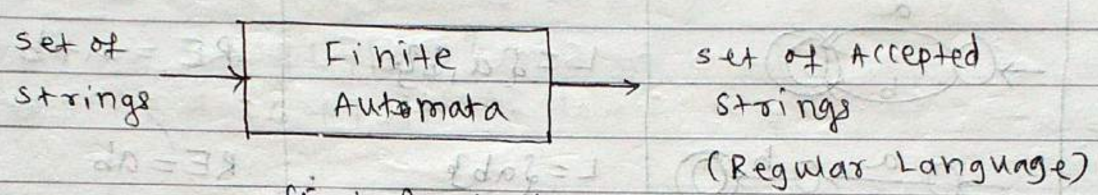


fig. 1: Regular Language

- Regular language can also be described in a compact form using a set of operators.

- The operators used to represent regular expression.
  i. Union operator (+)
  ii. cancatenation (·)
  iii. closure operation (*)

Regular Expression:

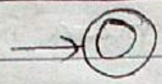
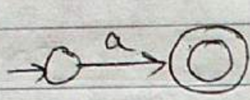
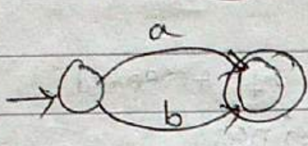
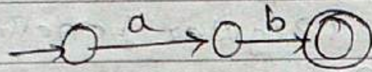
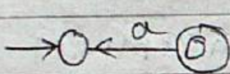
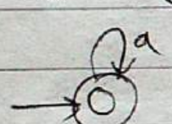
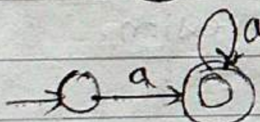
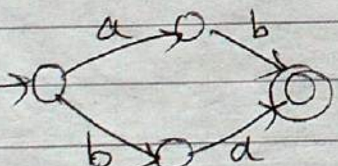
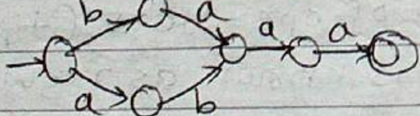
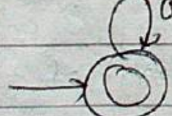
- An expression written using the set of operators (+, ·, *) and describing a regular language is known as regular expression.

- Regular expression is nothing but description of the language in algebraic form.

- The language represented by regular expression is called as regular language.

- Regular expression provides a declarative way to express the strings that machine can accept, which DFA/NFA doesn't provide.

Regular Expressions for basic Automata :

Sr.	Automata	Language	Regular Expression
1.		$L = \{\varepsilon\}$	$RE = \varepsilon$
2.		$L = \{a\}$	$RE = a$
3.		$L = \{a, b\}$	$RE = a+b$
4.		$L = \{ab\}$	$RE = ab$
5.		$L = \phi$	$RE = \phi$
6.		$L = \{\varepsilon, a, aa, aaa, \dots\}$	$RE = a^*$
7.		$L = \{a, aa, aaa, \dots\}$	$RE = a^+$ or aa^*
8.		$L = \{ab, ba\}$	$RE = ab + ba$
9.		$L = \{abaa, baaa\}$	$RE = abaa + baaa$ $\Rightarrow (ab+ba)aa$
10.		$L = \{\varepsilon, a, b, aa, ab, \dots\}$	$RE = (a+b)^*$

If R_1 & R_2 are regular expressions then:

1. $R_1 + R_2$
2. $R_1 \cdot R_2$
3. R_1^*

} are also regular.

4. R_1^+
5. R_2^*

} are also regular.

Languages to Regular Expressions (Languages to RE):

Write regular expressions for the following languages.

Example 1: set {1010}

Sol^n:

L = {1010}

RE = 1010

Example 2: L = {10, 1010}

Sol^n: ⇒ RE = 10 + 1010

Example 3: L = {ε, 10, 01}

Sol^n: ⇒ RE = ε + 10 + 01

Example 4: L = {ε, 0, 00, 000, ...}

Sol^n: ⇒ RE = 0^*

Example 5: L = {0, 00, 000, ...}

Sol^n: ⇒ RE = 0^+

Example 6: The set of strings over Σ{0,1} starting with 0

Sol^n: ⇒

L = {0, 01, 00, 010, 000, 011, ...}

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

Example 7: The set of strings over Σ = {0, 1} ending in 1.

Sol^n: ⇒

L = {1, 01, 11, 111, 001, 011, 101, ...}

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

Example 8: The set of strings over Σ = {a, b} starting with a & ending in b.

Sol^n: ⇒ RE = $a(a+b)^* b$

Exmple: The set of strings recognized by $(a+b)^3$

Soln :->

$L = (a+b)^3$

$RE = (a+b)(a+b)(a+b)$

$RE = aaa + aab + aba + abb + baa + bab + bba + bbb$

Regular set :

A word 'w' is said to be accepted by a finite automata M, if $\delta(q_0, w) = p$ for some $p \in F$. Here, q_0 is start state, F is set of final states and δ is a transition function.

A language is a regular set if it is the set accepted by some finite automata.

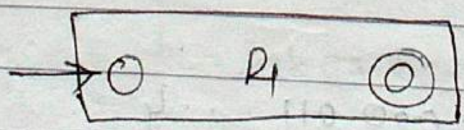
Example: Regular sets are,

1. $L = \{\varepsilon, 1, 11, ..\}$

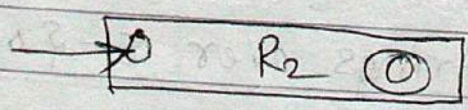
2. $L = \{\varepsilon, 00, 0000, 000000, ..\}$

Composite Finite Automata :->

Let FA for Regular Expression R_1

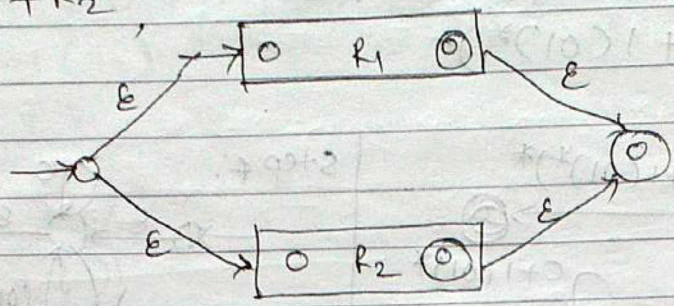


Let FA for RE R_2

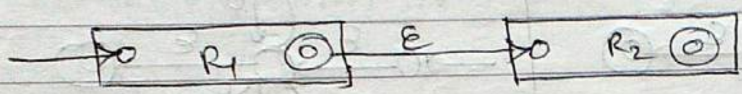


Each finite automata is assumed to have one final state. Then finite automata for following regular expressions are as,

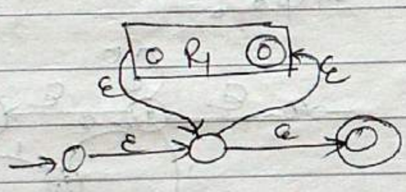
1. FA for $R_1 + R_2$



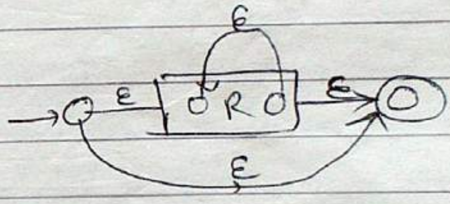
2. FA for $R_1 \cdot R_2$



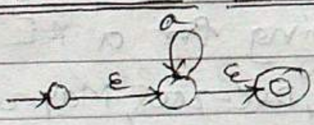
3. FA for R_1^*



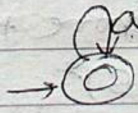
OR



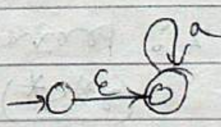
Finite automata for a^* :→



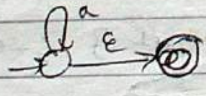
or



OR



OR



Basic properties of Regular Expression :

1. $\phi + R = R$
2. $\phi \cdot R = R \cdot \phi = \phi$
3. $\varepsilon \cdot R = R \cdot \varepsilon = R$
4. $\varepsilon^* = \varepsilon$
5. $\phi^* = \varepsilon$
6. $R + R = R$
7. $\cancel{R}\ PQ + PR = P(Q+R)$
8. $QP + RP = (Q+R)P$
9. $R^* R^* = R^*$
10. $RR^* = R^* R$
11. $(R^*)^* = R^*$
12. $\varepsilon + RR^* = R^*$
13. $(PQ)^* P = P(QP)^*$
14. $(P+Q)^* = (P^* Q^*)^* = (P^* + Q^*)^*$
15. $(P^* Q^*)^* = \varepsilon + (P+Q)^* Q$

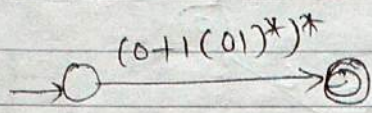
Page No.
Date :

RE to FA :

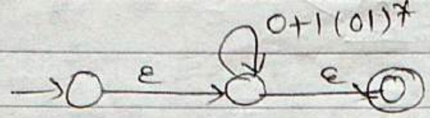
Example 1: construct FA for the RE
$R = (0 + 1(01)^*)^*$

Sol^n :⇒

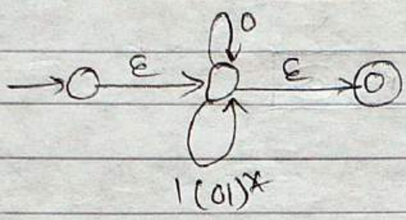
Step 1 :



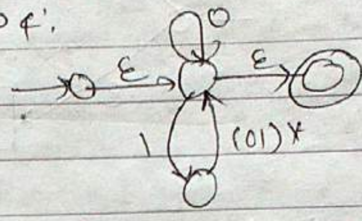
Step 2 :



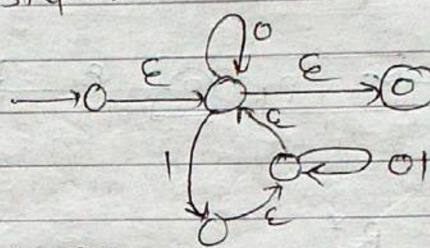
Step 3 :



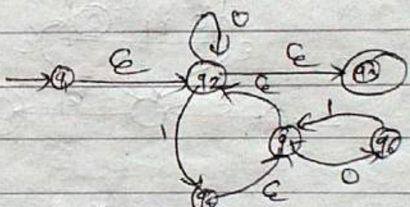
Step 4 :



Step 5 :



Step 6 :

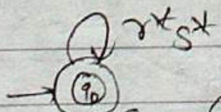


Example 2: prove or disprove the following for a RE.
$(r^*s^*)^* = (r+s)^*$ [May-2008, Dec-2008, May-2010, Dec-2010]

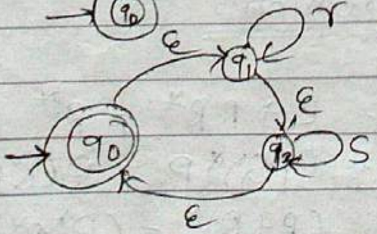
Sol^n :⇒

R.H.S $= (r^*s^*)^*$

Step 1 :



Step 2 :



Step 3 : convert ε-NFA to DFA.

Find ε-closure of every state,

$\varepsilon\text{-closure}(q_0) = \{q_0, q_1, q_2\}$
$\varepsilon\text{-closure}(q_1) = \{q_0, q_1, q_2\}$
$\varepsilon\text{-closure}(q_2) = \{q_0, q_1, q_2\}$

First subset is ε-closure of state q_0,

$\rightarrow \{q_0, q_1, q_2\}$

Find r & s successor on the set,

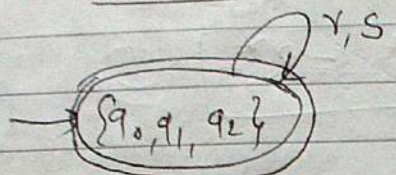
$\delta(\{q_0, q_1, q_2\}, r) = \varepsilon\text{-closure}(\delta(\{q_0, q_1, q_2\}, r))$
$\Rightarrow \varepsilon\text{-closure}(\delta(q_0, r) \cup \delta(q_1, r) \cup \delta(q_2, r))$
$\Rightarrow \varepsilon\text{-closure}(\phi \cup \{q_1\} \cup \{q_2\})$
$\Rightarrow \varepsilon\text{-closure}(q_1) \cup \varepsilon\text{-cl}(q_2)$
$\Rightarrow \{q_0, q_1, q_2\} \cup \{q_0, q_1, q_2\}$
$\Rightarrow \{q_0, q_1, q_2\}$

$\delta(\{q_0, q_1, q_2\}, s) = \varepsilon\text{-closure}(\delta(\{q_0, q_1, q_2\}, s))$
$\Rightarrow \varepsilon\text{-closure}(\delta(q_0, s) \cup \delta(q_1, s) \cup \delta(q_2, s))$

$= \epsilon\text{-closure}(\phi \cup \phi \cup \{q_2\})$

$\Rightarrow \epsilon\text{-closure}(q_2)$

$\Rightarrow \underline{\{q_0, q_1, q_2\}}$



No new subset,

Regular expression from FA is,

$\boxed{RE = (r+s)^*}$

$\boxed{\therefore (r^*s^*)^* = (r+s)^*}$

$\underline{L.H.S = R.H.S}$

Hence, proved,

Example 2: prove that,

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

Sol^n: ⇒

$LHS = \epsilon + RR^*$

$\Rightarrow \epsilon + R\{\epsilon + R + RR + \cdots\}$

$\Rightarrow \epsilon + \{R + RR + RRR + \cdots\}$

$\Rightarrow \cancel{\epsilon + R^*}\ R^*$

$= \underline{R.H.S}$

Example 3: Prove that,

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

Sol^n: ⇒

$L.H.S = (PQ)^*P$

$\boxed{(PQ)^* = \{\epsilon, PQ, PQPQ, PQPQPQ, \cdots\}}$ (1)

⇒ put eq^n (1)

$\Rightarrow \{\epsilon + PQ + PQPQ + PQPQPQ + \cdots\}P$

$\Rightarrow \{P + PQP + PQPQP + \cdots\}$

$\Rightarrow P\{\epsilon + QP + QPQP + \cdots\}$

$\Rightarrow P(QP)^*$

$\Rightarrow R.H.S$

Example 4 : prove that,

$$\varepsilon + 1^*(011)^*(1^*(011)^*)^* = (1+011)^*$$

sol^n :→

L.H.S $= \varepsilon + 1^*(011)^*(1^*(011)^*)^*$

= from identity, $\boxed{\varepsilon + RR^* = R^*}$

$\therefore R = 1^*(011)^*$

$\Rightarrow R^*$

$\Rightarrow (1^*(011)^*)^*$

$\therefore$ from identity, $(P+Q)^* = (P^*Q^*)^* = (P^*+Q^*)^*$

$\Rightarrow (1+011)^*$

⇒ R.H.S.

Example 5: Find all possible RE $L \subseteq \{a,b\}^*$ [Dec-2007, May-2008]

a) The set of strings ending in b

sol^n :⇒ $RE = (a+b)^* b$

b) The set of strings ending in ba

sol^n :⇒ $RE = (a+b)^* ba$

aa
bb ×
ab ×
ba ×

c) The set of all strings ending neither in b nor in ba

sol^n :⇒

$RE = \varepsilon + a + (a+b)^* aa$

d) The set of all strings ending in ab.

sol^n :⇒ $RE = (a+b)^* ab$

e) The set of all strings ending neither in ab nor ba.

sol^n :⇒ $RE = \varepsilon + a + b + (a+b)^*(aa+bb)$

Example 6: Find RE for subset of $\{0,1\}^*$

a) The language of all strings containing exactly two 0's

b) The language of all strings containing at least two 0's

c) The language of strings that do not end with 01.

[May-2009, Dec-2010]

Soln ⇒

a) $RE = 1^* 0 1^* 0 1^*$

b) $RE = (0+1)^* 0 \cdot (0+1)^* \cdot 0 (0+1)^*$

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

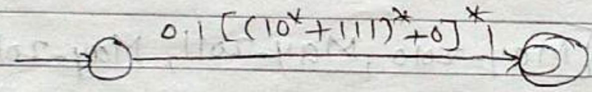
c) $RE = \epsilon + 0 + 1 + (0+1)^* (11 + 10 + 00)$

Example 7: Convert NFA with ε-moves to DFA for ~~RE = 0.1[(10~~

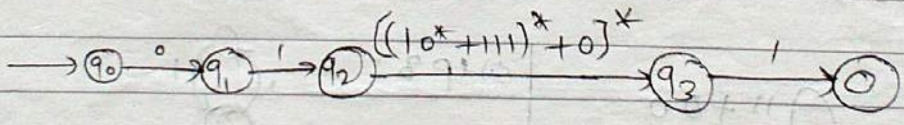
$RE = 0.1[(10^* + 111)^* + 0]^* 1.$

Soln ⇒

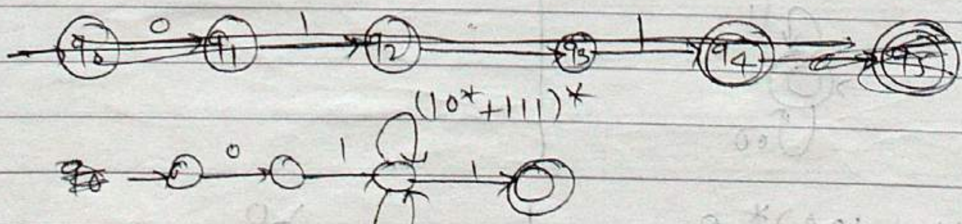
Step1:

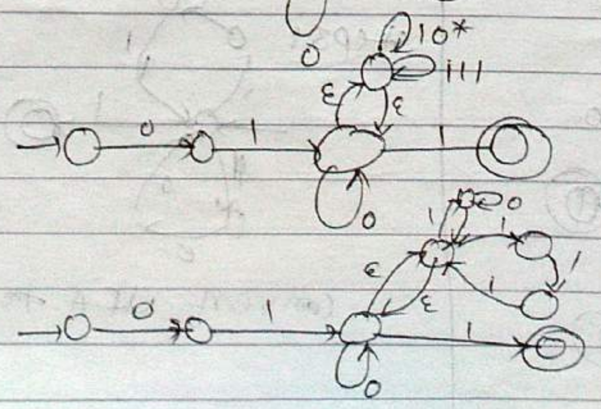


Step2:



Step3:





convert to DFA

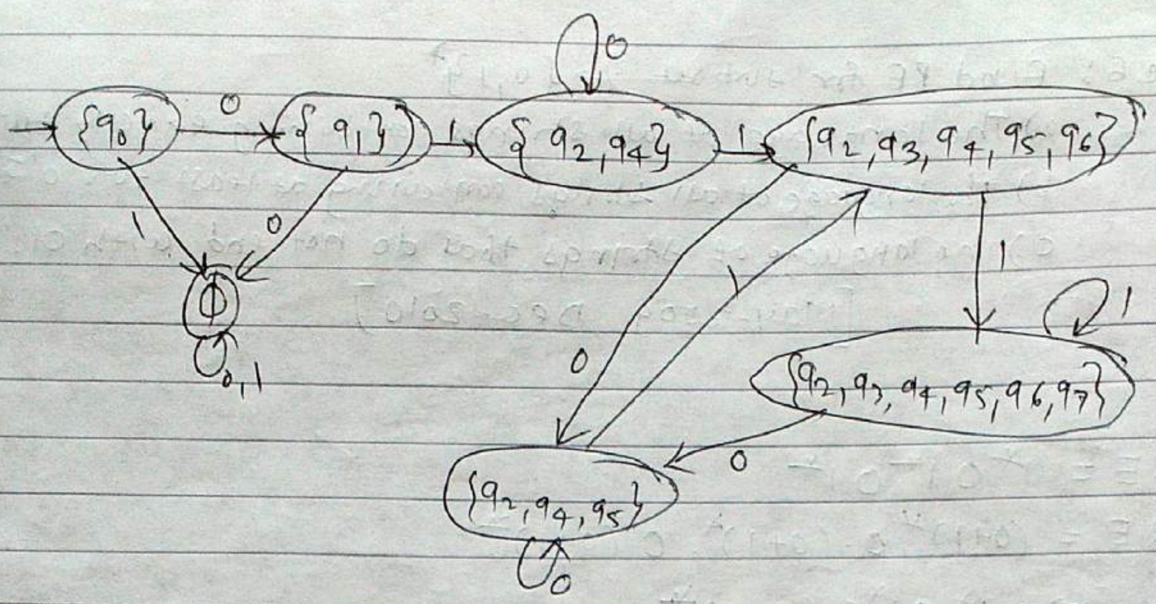


fig. DFA

Example 7: RE to DFA.

(a) $(11+00)^*$

(b) $(111+100)^* 0$

(c) $0+10^*+01^*0$

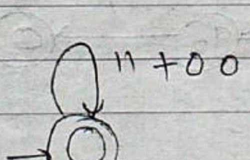
[May-2010, May-2011, May-2014]

solⁿ :⇒

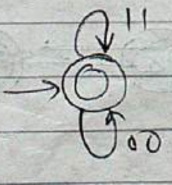
(a) $(11+00)^*$

solⁿ :⇒

step1 :



step2 :

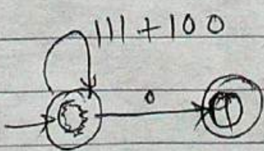


step3 :

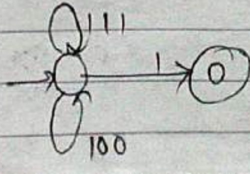
(b) $(111+100)^* 0$

solⁿ :⇒

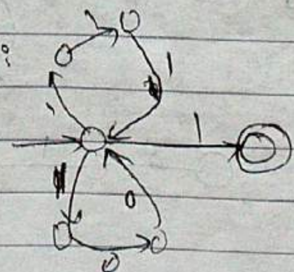
step1 :



step2 :



step3 :



convert NFA to DFA

Page No.

Date :