Regular Expressions

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

1) Any terminal symbol i.e. symbols $\in \Sigma$ including $\wedge$ (Empty) and ϕ (null) are regular expression.

   example → $a, b, c \ldots \wedge, \phi$

2) The union of two regular expressions is also a regular expression.

   example → $R_1, R_2 \Rightarrow (R_1 + R_2)$

3) The concatenation of two regular expressions is also a regular expression.

   example → $R_1, R_2 \Rightarrow (R_1 . R_2)$

4) The iteration (or closure) of a regular expression is also a regular expression.

   example → $R \rightarrow R^*$ $\quad a^* = \phi, a, aa, aaa \ldots$

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

Examples of Regular Expression

1. Starting with 1 and end with 0.

First symbol is 1 and ends with 0 in middle it can be anything.

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

2. Starting and ending with a and any no. of b's in between.

$a(b, \varepsilon, bb, bbb, bbbbbbb)a$

$RE = a(b)^*a$

3. Starting with a and not having consecutive b's.

$a(a, \varepsilon, ab, aba, abab, ababaa, aaab, aaabaaba)$

$RE = (a + ab)^*$

4. Any number of a's, any number of b's, any number of c's followed all together.

$L = \{abc, a, ab, ac, bc, bcc\}$

$RE = a^* b^* c^*$

5. RE. for even length of string over $\Sigma = \{0\}$

$RE = (00)^*$

6. R.E. for atleast one 0 and one 1

$RE = \cancel{(0^{*}}$

$[(0+1)^* 0 (0+1)^* 1 (0+1)^*] + [(0+1)^* 1 (0+1)^* 0 (0+1)^*]$

Identities of Regular Expression

1) $\phi + R = R$
2) $\phi R + R\phi = \phi$
3) $\varepsilon R = \varepsilon R = R$
4) $\varepsilon^* = \varepsilon$ and $\phi^* = \varepsilon$
5) $R + R = R$
6) $R^* R^* = R^*$
7) $RR^* = R^*R$
8) $(R^*)^* = R^*$
9) $\varepsilon + RR^* = \varepsilon + R^*R = R^*$
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

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. $R = QP^*$

$R = Q + RP$ — (1) Replace R with $R = QP^*$

$= Q + QP^*P$

$= Q(\varepsilon + P^*P)$ $\varepsilon + R^*R = R^*$ Identity

$= QP^*$

Proof for unique solution.

$R = Q + RP$ Replace R with $R = Q + RP$

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

$R = Q + QP + RP^2$ Replace R with $R = Q + RP$

$R = Q + QP + (Q + RP)P^2$

$R = Q + QP + QP^2 + RP^3$

⋮ n time

$R = Q + QP + QP^2 \ldots QP^n + RP^{n+1}$ Replace R with QP^*

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

$R = Q(\underbrace{\varepsilon + P + P^2 \ldots P^n + P^*P^{n+1}}_{})$

↳ closure of P i.e P^*

$R = QP^*$

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

LHS:- $(1+00^*1)+(1+00^*1)(0+10^*1)^*(0+10^*1)$

$\Rightarrow (1+00^*1)[\varepsilon+(0+10^*1)^*(0+10+1)]$ $\quad \varepsilon+R^*R=R$

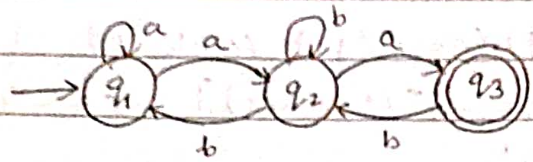
$\Rightarrow (1+00^*1)(0+10^*1)^*$

$\Rightarrow (\varepsilon\cdot1+00^*1)(0+10^*1)^*$ $\quad \varepsilon\cdot R=R$

$\Rightarrow (\varepsilon+00^*)1(0+10^*1)^*$ $\quad \varepsilon+R^*R=R$

$\Rightarrow 0^*1(0+10^*1)^*$

NFA to regular expression Conversion :-



$q_3 \rightarrow q_2 a$ — ①

$q_2 \rightarrow q_1 a + q_2 b + q_3 b$ — ②

$q_1 \rightarrow \underset{\text{start state}}{\varepsilon} + q_1 a + q_2 b$ — ③

$q_3 \rightarrow q_2 a$ receive which input from which state

Solve all three expression above :-

① →

$q_3 = q_2 a$

$= (q_1 a + q_2 b + q_3 b) a$ putting value of q_2 from ②

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

② →

$q_2 \Rightarrow q_1 a + q_2 b + q_3 b$

$\Rightarrow q_1 a + q_2 b + q_2 ab$ putting q_3 from ①

$\underbrace{q_2}_{R} \Rightarrow \underbrace{q_1 a}_{Q} + \underbrace{q_2}_{R} \underbrace{(b + ab)}_{P}$ $R = Q + RP = QP^*$

$q_2 \Rightarrow (q_1 a)(b + ab)^* \longrightarrow$ ⑤

③ →

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

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

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

$q_1 = \cancel{(q_1} \varepsilon ((a + a(b+ab)^*) b)^*$ $\varepsilon \cdot R = P$

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

Final State $q_3 \rightarrow$

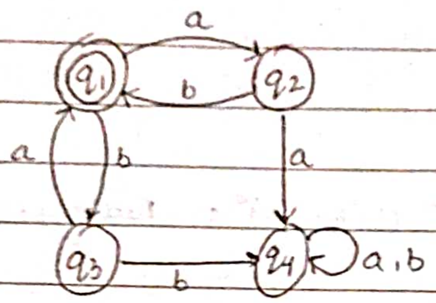
$q_3 = q_2 a$ from eqn (5)

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

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

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

DFA to Regular Expression Conversion



$q_1 = \varepsilon + q_2 b + q_3 a$ — (I)

$q_2 = q_1 a \rightarrow$ (II)

$q_3 = q_1 b \rightarrow$ (III)

$q_4 = q_2 a + q_3 b + q_4 a + q_4 b \rightarrow$ (4)

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

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

$q_1 = \varepsilon + q_1 ab + q_1 ba$

$\underbrace{q_1}_{R} = \underbrace{\varepsilon}_{Q} + \underbrace{q_1}_{R}\underbrace{(ab + ba)}_{P}$ $\qquad R = Q + RP = QP^*$

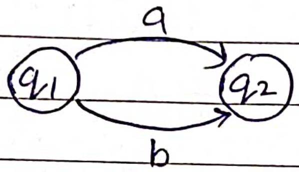
$q_1 = \varepsilon \cdot (ab + ab)^*$ $\qquad \varepsilon \cdot R = R$

$q_1 = (ab + ba)^*$

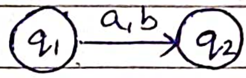
Conversion of Regular Expression to Finite Automata

Rules to Remember :-

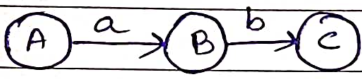
1. (a+b)



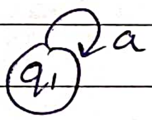
or



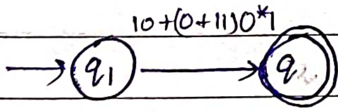
2. (a·b) or (ab)

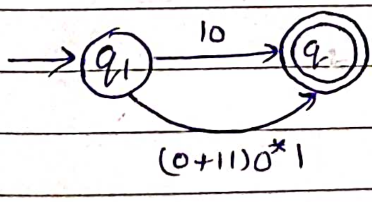


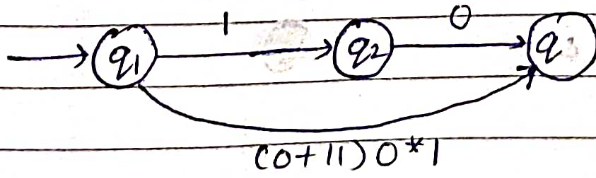
3. a*

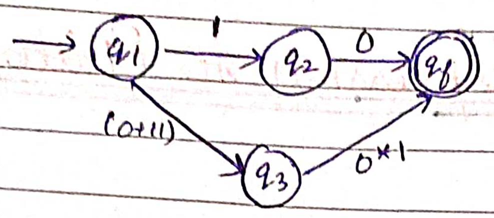


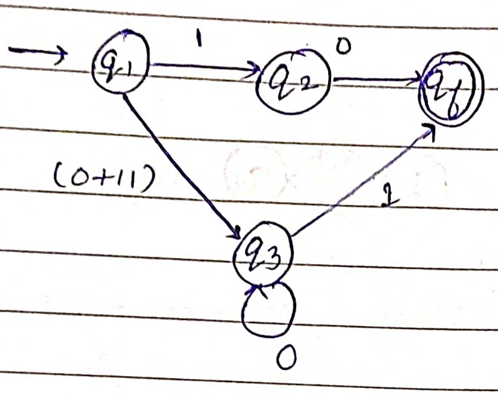
Example RE to FA (10 + (0+11)0* 1)

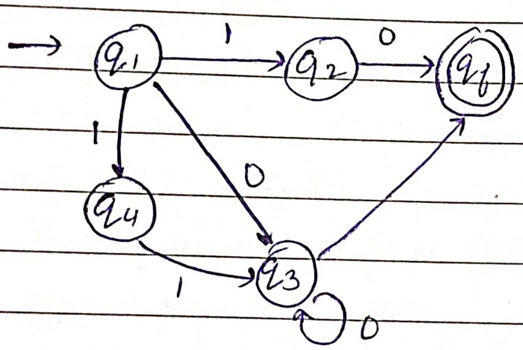












* First convert to NFA with Epsilon that is ε-NFA and then to NFA without epsilon and then to DFA.