REGULAR EXPRESSIONS

- Formal recursive def.^{n} of Regular Expression (RE) over Σ as follows:

1). Any Terminal Symbol ($a \in \Sigma$), $\wedge$ & ϕ are RE.

2). Union of R_1 & R_2 is written as $R_1 + R_2$ is also a RE.

3). concatenation of R_1 & R_2, written as $R_1 R_2$ is also RE.

4). Iteration or closure of R, written as R^* is also RE.

5). If R is RE then (R) is also a RE.

→ RE over Σ can be obtained by applying rule 1-5 any number of times.

Def^n: Any set represented by RE is called a Regular set.

Ex.: 1) L_1 : set of all strings of 0's & 1's ending in 00.

R_1 : $(0+1)^* 00$

2) L_2 : Set of all strings of 0's & 1's beginning with 0 and ending with 1.

R_2 : $0(0+1)^* 1$.

3). L_3 : $\{\wedge, 11, 1111, 111111, \ldots\}$

R_3 : $(11)^*$

Identities for REs:

1. $\phi + R = R$
2. $\phi R = R\phi = \phi$
3. $\Lambda R = R\Lambda = R$
4. $\Lambda^* = \Lambda$ & $\phi^* = \Lambda$
5. $R + R = R$
6. $R^* R^* = R^*$
7. $R R^* = R^* R$
8. $(R^*)^* = R^*$
9. $\Lambda + R R^* = R^* = \Lambda + R^* R$
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: Let P & Q be RE over Σ. If P does not contains Λ, then the following equation in R, viz.

$$R = Q + RP \text{ has a unique sol}^n$$

given by,

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

Prove that

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

LHS

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

$$= (1+00^*1)\left[\Lambda + (0+10^*1)^*(0+10^*1)\right] \quad \{\text{Using } 12\}$$

$$= (1+00^*1)(0+10^*1)^* \quad \{\text{Using } 9\}$$

$$= (\Lambda+00^*)1(0+10^*1)^* \quad \{\text{using } 12 \text{ for } 1+00^*1\}$$

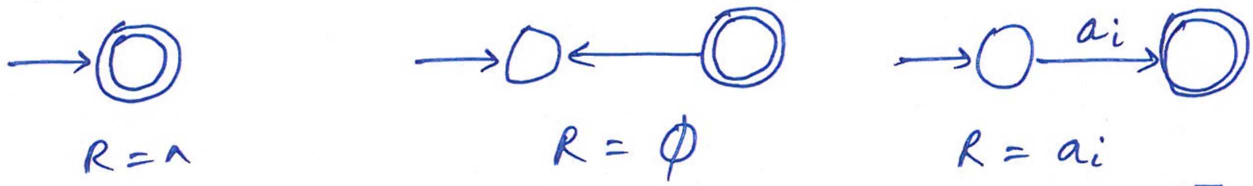
$$= 0^*1(0+10^*1)^* \quad \{\text{using } 9\}$$

$= RHS.$

———— x ————

FINITE AUTOMATA & REGULAR EXPRESSIONS

Theorem 1: Every RE R can be recognized by a transition system, i.e. for every string $w \in R$, there exists a path from initial state to final state with path value w.



$R = \Lambda$ $\qquad$ $R = \phi$ $\qquad$ $R = a_i$, $a_i \in \Sigma$

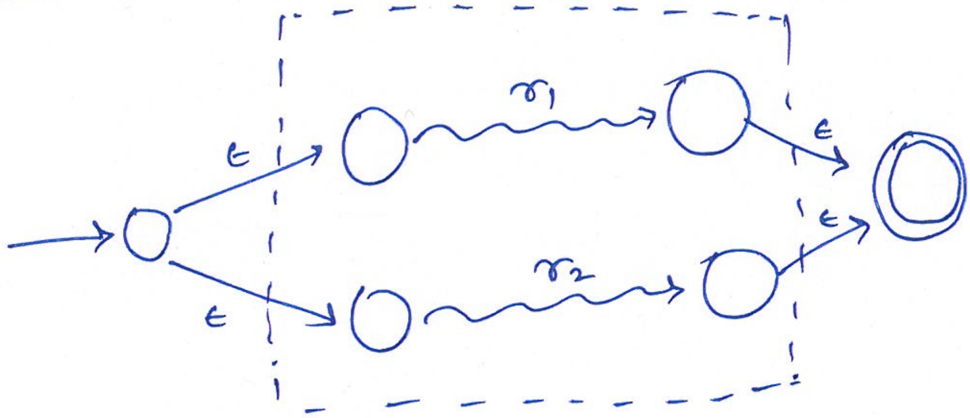
Theorem 2: Any set L accepted by a FA is represented by a RE.

(Proof for both theorem left to work out)

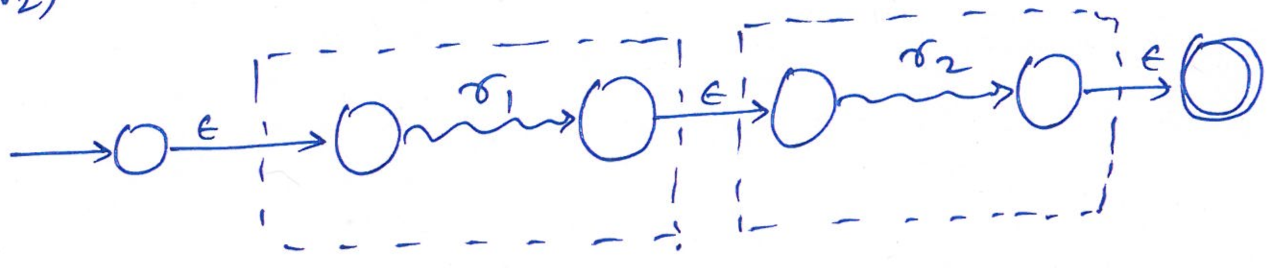
—x—

Automata for REs

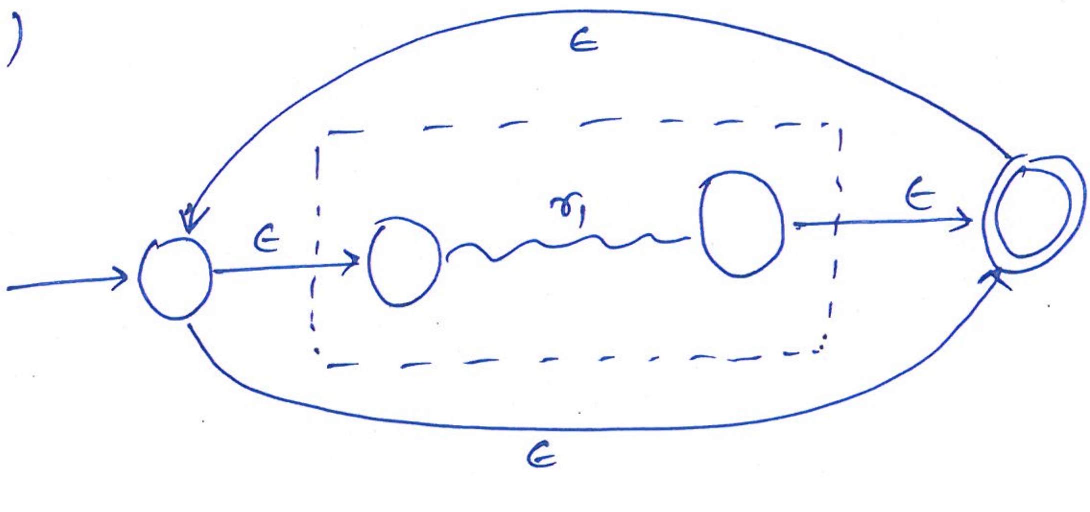
$L(r_1 + r_2)$



$L(r_1 r_2)$

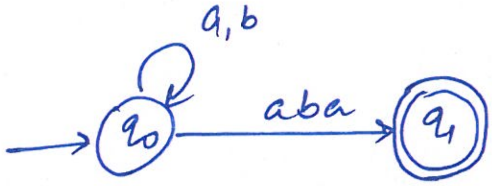


$L(r_1^*)$

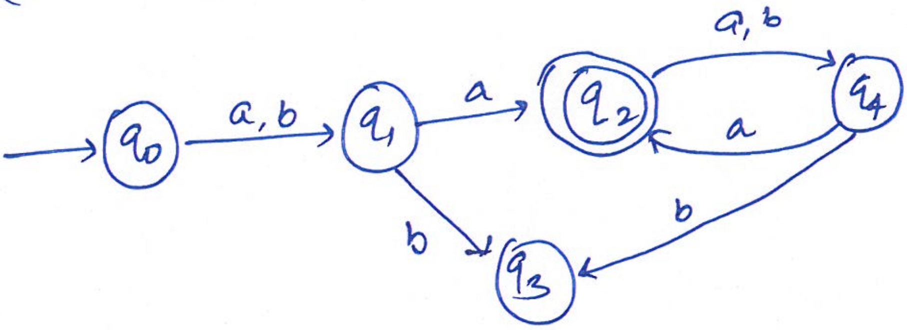


Design NDFA for

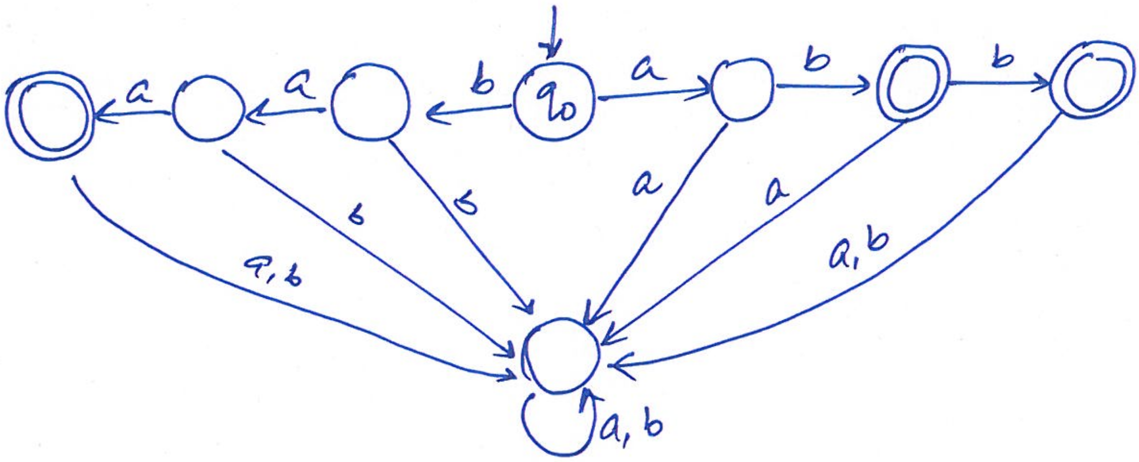
1) $(a+b)^* aba$



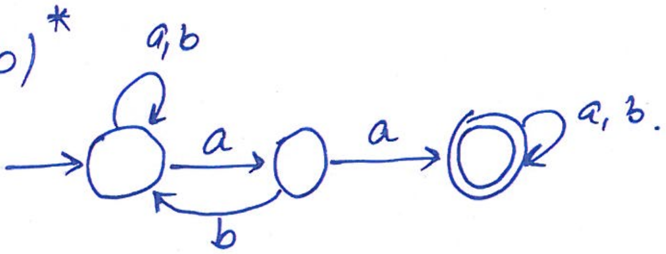
2) $(a+b)a((a+b)a)^*$



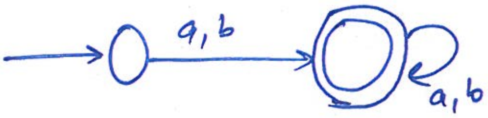
3) $baa + ab + abb$



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



5) $(a+b)^+$



Algebric Method using Arden's Theorem

Assumptions:

1. Transition graph does not have ε-moves.
2. It has only one initial state, say q_1.
3. Its vertices are $q_1, q_2, \ldots, q_n$.
4. v_i is the RE representing the set of strings accepted by the system even though q_i is the final state.
5. α_{ij} denotes RE representing set of labels of edges from q_i to q_j. When there is no edge, $\alpha_{ij} = \phi$.

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

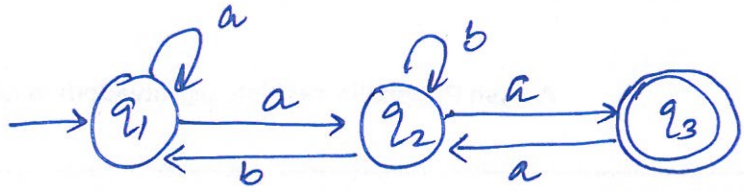
$$q_2 = q_1\alpha_{12} + q_2\alpha_{22} + \cdots + q_n\alpha_{n2}$$

$$\vdots$$

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

Apply substitutes & Arden's theorem, express q_i in terms of α_{ij}'s.

Q.



string Accepted

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

- FA does not have ϵ-moves.
- FA has unique starting state q_1.

Now,

$$q_1 = q_1 a + q_2 b + \epsilon \quad \text{---(1)}$$

$$q_2 = q_1 a + q_2 b + q_3 a \quad \text{---(2)}$$

$$q_3 = q_2 a \quad \text{---(3)}$$

Put q_3 in eq.^{n} (2),

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

$$= q_1 a (b+aa)^* \quad \left[\text{Arden's theorem: } R = Q + RP \Rightarrow R = QP^*\right]$$

Now, put q_2 value in eq^n (1)

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

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

$$q_1 = \epsilon (a + a(b+aa)^* b)^* \quad [\text{Arden's theorem}]$$

Now,

$$q_2 = (a + a(b+aa)^* b)^* a (b+aa)^*$$

$$q_3 = (a + a(b+aa)^* b)^* a (b+aa)^* a. \quad \text{---(4)}$$

Since q_3 is the final state in the system, hence eq.^{n} (4) will be RE accepted by the given transition system.