Algebraic properties of regular expressions:

Kleene closure is an unary operator and Union(+) and concatenation operator(.) are binary operators.

1. Closure:

$a + b = \underline{\quad}$

NN NN NN

If r1 and r2 are regular expressions(RE), then

r1* is a RE
r1+r2 is a RE
r1.r2 is a RE

2. Closure laws –

(r*)* = r*, closing an expression that is already closed does not change the language.
∅* = ∈, a string formed by concatenating any number of copies of an empty string is empty itself.
r+ = r.r* = r*r, as r* = ∈ + r + rr+ rrr …. and r.r* = r+ rr + rrr ……
r* = r*+ ∈ $r^* = r^+ + \varepsilon$

$\phi^* = \{\varepsilon, \underbrace{\phi, \phi^2, \phi^3 \ldots}_{\text{empty}}\}$

$= \{\varepsilon\}$

3. Associativity –
If r1, r2, r3 are RE, then

i.) r1+ (r2+r3) = (r1+r2) +r3
For example : r1 = a , r2 = b , r3 = c, then
The resultant regular expression in LHS becomes a+(b+ c) and the regular set for the corresponding RE is {a, b, c}.
for the RE in RHS becomes (a+ b) + c and the regular set for this RE is {a, b, c}, which is same in both cases. Therefore, the associativity property holds for union operator.

ii.) r1.(r2.r3) = (r1.r2).r3
For example – r1 = a , r2 = b , r3 = c
Then the string accepted by RE a.(b.c) is only abc.
The string accepted by RE in RHS is (a.b).c is only abc ,which is same in both cases. Therefore, the associativity property holds for concatenation operator.
Associativity property does not hold for Kleene closure(*) because it is unary operator.

$r^+ + \varepsilon = r^*$

$(r^1, r^2, r^3 \ldots, \varepsilon) = r^*$

4. Identity –
In the case of union operators

$r + x = r$

if r+ x = r ⇒ x= ∅ as r ∪ ∅= r, therefore ∅ is the identity for +.
Therefore, ∅ is the identity element for a union operator.

set of all even length strings

even strings not divisible by 2

In the case of concatenation operator –
if r.x = r , for x= ∈
r.∈ = r ⇒ ∈ is the identity element for concatenation operator(.) .

aa, ab, ba . . .

ϕ

{aa, ab, ba . . . . ϕ} = {aa, ab, ba...}

5. Annihilator –

If r+ x = x ⇒ r ∪ x= x , there is no annihilator for +

In the case of a concatenation operator, r.x = x, when x = ∅, then r.∅ = ∅, therefore ∅ is the annihilator for the (.)operator. For example {a, aa, ab}.{ } = { }

6. Commutative property –

If r1, r2 are RE, then

r1+r2 = r2+r1. For example, for r1 =a and r2 =b, then RE a+ b and b+ a are equal.
r1.r2 ≠ r2.r1. For example, for r1 = a and r2 = b, then RE a.b is not equal to b.a.

7. Distributed property –

If r1, r2, r3 are regular expressions, then

(r1+r2).r3 = r1.r3 + r2.r3 i.e. Right distribution
r1.(r2+ r3) = r1.r2 + r1.r3 i.e. left distribution
(r1.r2) +r3 ≠ (r1+r3)(r2+r3)

8. Idempotent law –

r1 + r1 = r1 ⇒ r1 ∪ r1 = r1 , therefore the union operator satisfies idempotent property.

r.r ≠ r ⇒ concatenation operator does not satisfy idempotent property.

9. Identities for regular expression –

There are many identities for the regular expression. Let p, q and r are regular expressions.

Reference Link : https://www.geeksforgeeks.org/properties-of-regular-expressions/

Q: RE for strings of length exactly 2.

$\Sigma = \{a, b\}$

$L = \{aa, ab, ba, bb\}$

$aa + ab + ba + bb$

$a(a+b) + b(a+b)$

$(a+b)(a+b)$

either a or b

either a or b.

Q: length exactly 3

$(a+b)(a+b)(a+b)$

Q: length atleast 2

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

Q: atmost 2   0 length + 1 length + 2 length

$L = \{\varepsilon, a, b, aa, ab, ba, bb\}$

1 way: $\varepsilon + a + b + aa + ab + ba + bb$

2 way: $(\varepsilon + a + b)(\varepsilon + a + b)$

$\varepsilon \quad a \quad = \quad a$

$a \quad b \quad = \quad ab$

Q: Even length string:

0 2 4 6 . . . . .

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

Q: odd length string.

$\underbrace{\underbrace{((a+b)(a+b))^*}_{\text{Even}}(a+b)}_{\text{odd}} \quad \Big| \quad \underbrace{(a+b)\underbrace{((a+b)(a+b))^*}_{\text{Even}}}_{\text{odd}}$

Q: Starts with a $a(a+b)^*$

Q: Ends with a $(a+b)^*a$

Q: Containing a $(a+b)^* a (a+b)^*$

Q: Starts & Ends with different symbols

$\Sigma = \{a, b\}$ $a \longrightarrow b$ or $b \longrightarrow a$

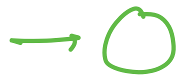
$a(a+b)^*b \quad + \quad b(a+b)^*a$

Q: Starts & ends with same symbols

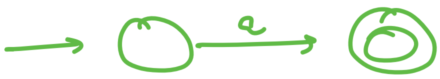
$a(a+b)^*a + b(a+b)^*b + \underline{\underline{a+b}}$

Conversion of RE to FA:

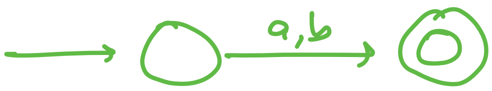
ϕ Don't accept any language



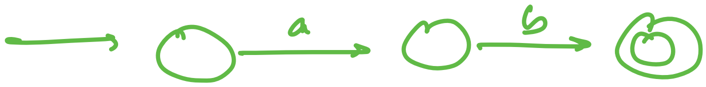
a



$a+b$

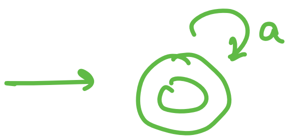


$a \cdot b$

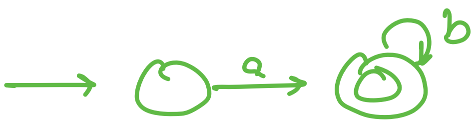


a^*

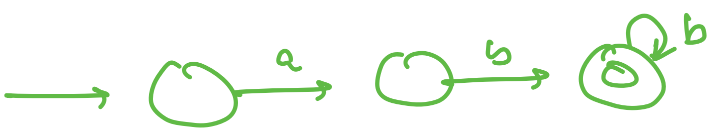
$\epsilon, a, aa, aaa \ldots$



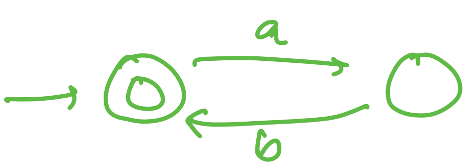
ab^*



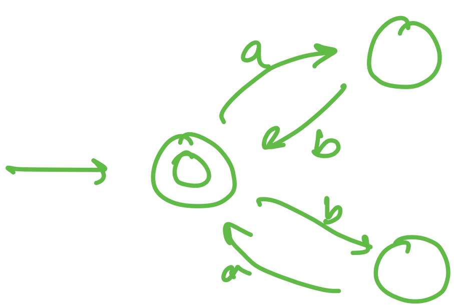
ab^+



$(ab)^*$



$(ab+ba)^*$



Conversion of FA to RE:

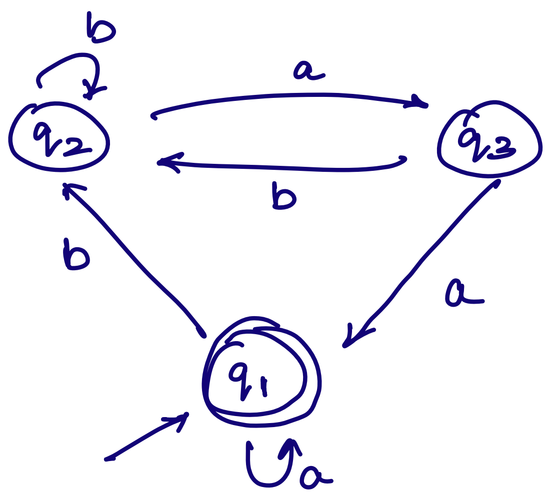
ARDEN'S THEOREM

$R = Q + RP$ $\qquad\qquad$ $R = QP^*$

$$\begin{aligned}
R &= Q + RP \\
&= Q + (Q + RP)P \\
&= Q + QP + RP^2 \\
&= Q + QP + (Q + RP)P^2 \\
&= Q + QP + QP^2 + RP^3 \\
&= Q + QP + QP^2 + QP^3 + RP^4
\end{aligned}$$

$$\begin{aligned}
&= Q + QP + QP^2 + QP^3 + QP^4 \ldots\ldots \quad QP^* \\
&= Q(\varepsilon + P + P^2 + P^3 + P^4 \ldots\ldots P^*) \\
&= QP^*
\end{aligned}$$

Q:



$q_1 = q_1 a + q_3 a + \varepsilon$ ← Initial State (1)

$q_2 = q_1 b + q_2 b + q_3 b$ (2)

$q_3 = q_2 a$ (3)

Put (3) in (1)

$q_1 = q_1 a + q_2 aa + \varepsilon$ — (4)

$R = Q + RP$

$R = QP^*$

Put (3) in (2)

$q_2 = q_1 b + q_2 b + q_2 ab$ — (5)

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

$q_2 = q_1 b (b+ab)^*$ — (6)

Put (6) in (4)

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

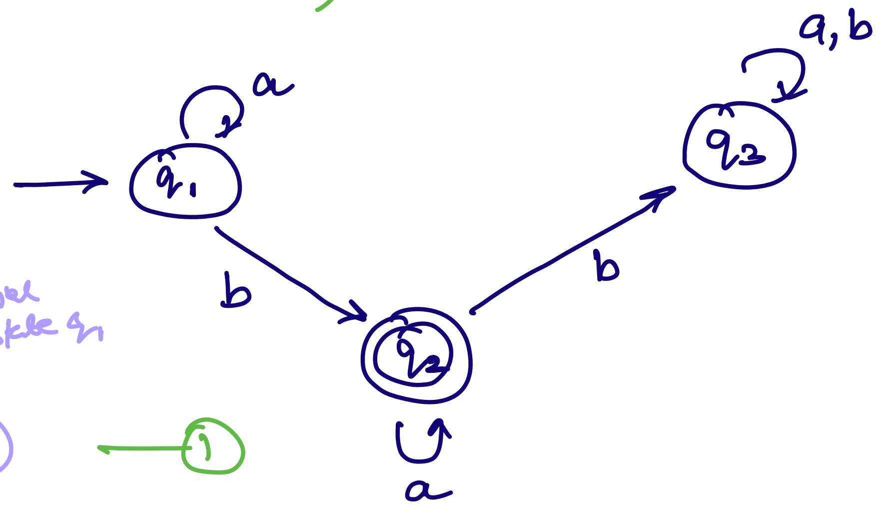
$\underbrace{q_1}_{R} = \underbrace{\varepsilon}_{Q} + \underbrace{q_1}_{R} \underbrace{(a + b(b+ab)^* aa)}_{P}$

$R = QP^*$

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

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

Q:



Unreachable

Dead:

q_3

never reach final

initial state q_1

$q_1 = q_1 a + \varepsilon$ — (1)

$q_2 = q_1 b + q_2 a$ — (2)

$q_3 = q_2 b + q_3 a + q_3 b$ — (3)

in eq^n (1)

$$\underline{\underline{q_1}} = \underline{\varepsilon} + \underline{\underline{q_1}}\,\underline{\underline{a}}$$
$$R \quad Q \quad R \quad P$$

$q_1 = \varepsilon a^* \Rightarrow q_1 = a^*$

$$q_2 = q_1 b + q_2 a$$

$$\underline{\underline{q_2}} = \underline{\underline{a^* b}} + \underline{\underline{q_2}}\,\underline{\underline{a}}$$
$$R \quad Q \quad R \quad P$$

$$\boxed{q_2 = a^* b a^*}$$