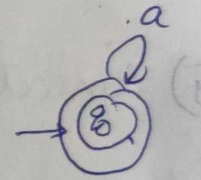
Regular Expression (Part 1)

↳ It is a method to Represent a language.

means – If we want to represent a language(L) then we use a regular language

→ Regular Expression represent language (Regular Language) All those languages which are accept FA (DFA | NFA, NFA with ε, mealy | moore) All strings that are accepted that are regular language And To represent Regular language we use Regular Expression.

for ex $L = \{\lambda, a, aa, aaa --\}$ ⇒ R.E. a^*



Let 'R' be a Regular Expression over Alphabet Σ if R is :

1) ε is Regular Expression denoting the set {ε} (null string) (collection of string)
2) φ is " " " { } (empty set)
3) for each symbol $a \in \Sigma$, a is regular Expression denoting set {a}.

$R = \varepsilon \quad L(R) = \{\varepsilon\}$
$R = \phi \quad L(R) = \{\}$
$R = a \quad L(R) = \{a\}$

These three are premitive.

4) Union of two Regular Expression is also Regular.

$R_1 \cup R_2$ = Regular $(R_1 + R_2)$.

$a \cup b = \{a, b\}$.

If x is RE over language $L(x)$

Y " " over language $L(Y)$

→ $X+Y$ is RE corresponding to $L(X) \cup L(Y)$

where $L(X \cup Y) = L(X) \cup L(Y)$.

→ $X \cdot Y$ is RE corresponding to $L(X) \cdot L(Y)$

where $L(X \cdot Y) = L(x) \cdot L(Y)$

5) Concatination of two RE is ~~also~~ also Regular

$R_1 R_2$ = Regular.

6). → R^* is RE corresponding to $L(R^*) = (L(R))^*$.

$a^* = L\{\lambda, a, aa, aaa \ldots\}$

* RE is said to be valid iff it can be derived from the primitive RE by finite no. of application of the rules

$r^*, r^+, r_1 \cdot r_2, r_1 + r_2$

eg

Language

finite ~~Automo~~ infinite

↳ If lang. finite it must be Regular ✓

↳ " " " then we create FA ✓

↳ " " " then we write RE ✓

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

Lang with

1) No string → $\{ \}$

RE = ϕ

1) length 0 = $\{\varepsilon\}$

RE = ε

2) length 1 = $\{a, b\}$

RE = $(a+b)$

(finite lang. को RE मे union कर देना है

4) length 2 = $\{ab, ba, aa, bb\}$

RE = $(ab + ba + aa + bb)$ or $(a+b)(a+b)$

= $a(a+b) + b(a+b)$.

5) length 3 = $\{aaa, aab, aba, abb, bbb, bab, baa, bba\}$

= $(a+b)(a+b).(a+b)$

6) length of string Atmost 1 ⇒ (0,1)

$\{\varepsilon, a, b\}$

$RE = \{\varepsilon + a + b\}$

7) At most 2 ⇒ $(\varepsilon+a+b)(\varepsilon+a+b)$

8) No. of b not more than 2 b's & 1 a

$\{\varepsilon, a, b\ ab, ba\ bba, bab, abb\}$

$RE = \varepsilon, a+b+ab+ba+bba+bab+abb$

use rightly $a \cdot a^*$

(*) (+)

Powers.

RE for infinite language

Language can be Regular & not

$aa^* = a^+$

$= aa^0$

$= a^*\varepsilon$

$= a$

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

1). All the string having a single 'b'

$= \{\varepsilon, a, aa, aaa, aaaa \dots\}\ b.\{\varepsilon, a, aa \dots\}$

$RE = (a^*)\ b\ (a^*)$

2). All the string having at least one 'b'

$L = (a+b)^*\ b\ (a+b)^*$

3 All the string having any no. of a & and any no. of b's

$(a+b)^*$

$L = \{\varepsilon, a, b, ab \dots\}$

strings having bbbb as substring.

* There are multiple way to write RE

$RE = (a+b)^* bbbb (a+b)^*$.

5) All strings end with ab.

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

6) All string start with 'ba'

$ba(a+b)^*$

7) beginning ~~with~~ & end with 'a'

$a(a+b)^* a$

8) All the string containing 'a'

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

9) All the string starting & end with different symbol

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

10) All the string having double 'b'

$RE = a^* bb a^*$ abba. aabba

✓ $RE = a^* b a^* b a^*$ ab aba

Describe the following set by RE.

1) $\{101\}$ RE = ~~(~~ (101)

2) $\{abba\}$ RE = (abba)

3) $\{01, 10\}$ RE = (01 + 10)

4) $\{\Lambda, ab\}$. RE = $\Lambda + ab$

5) $\{abb, a, b, bba\}$ ⇒ RE = $abb + a + b + bba$

6) $\{\Lambda, 0, 00, 000\}$ RE = 0^*.

7) $\{1, 11, 111 \ldots\}$ RE = $(1)^+$
= $1(1)^*$

ARDEN's Theorem.

↳ To check the equivalence two RE

↳ Conversion of DFA to RE then use Arden's

Before Applying this theorm we need to follow 2 conditions.

- FA should not contain ε transition
- FA should have only one initial state.

If Finite Automata satisfy these two condition then we directly this Arden's theorm.

...ntity Rule

Rule 1 - $\phi + r = R$

2 - $\phi \cdot r = r \cdot \phi = \phi$

3 = $\varepsilon \cdot r = r \cdot \varepsilon = r$ / $\Lambda \cdot r = r \Lambda = r$

4 = $r + r = r$ $\Lambda^* = \Lambda$

5 = $r^* \cdot r^* = r^*$ $\phi^* = \Lambda$

6 = $r^* r = r \cdot r^* = r^+$

7 = $(r^*)^* = r^*$.

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

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

10 = $(P+Q)^* = (P^* \cdot Q^*)^* = (P^* + Q^*)^*$

11 = $(P+Q)r = Pr + Qr$ &

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

Arden's Theorem : If P & Q are two RE & P doesn't have ε-transition then equation $\boxed{R = Q + RP}$ will have unique solution $R = QP^*$ → eq ①.

Proof $R = Q + RP$

Apply eq ①

$R = Q + QP^* P$

$R = Q + QP^* P$

$= Q(1 + P^* P)$

$\Rightarrow Q(\varepsilon + P^* P)$

$\boxed{R = QP^*}$

1 is nothing but ε.

Application RE is.

↳ Validation of data

Ex Compiler.

compiler some set of ~~rules~~. rules we need to write a program that obsses that particular rule

If any rule is volitele then we get syntax error

→ RE denoting language with strings having any number of a's over $\Sigma = \{a\}$

$$RE = a^*$$

→ RE denoting language with string having any number of a's & any no. of b's over $\Sigma = \{a, b\}$

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

→ RE denoting language with string start with 'a' & end with 'b'

$RE =$ ~~$a((a+b)^*)b$~~ $a(a+b)^* b$

Of

$$R = Q + RP \quad \text{—(1).}$$

$$R = Q + RP$$

Apply Eq (1)

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

9

10 Apply eq (1)

11 $$R = Q + QP + (Q + RP)P^2$$
$$= Q + QP + QP^2 + RP^3$$

Apply eq (1)

$$R = Q + QP + QP^2 + (Q + RP)P^3$$
$$= Q + QP + QP^2 + QP^3 + RP^4$$

12.

$$\text{---} \quad QP^4 + RP^5$$

$$R = Q(\Lambda + P + P^2 + P^3 \text{ ---})$$

13.

(14) $$R = QP^*$$

15.

16

$(a+b)^*$

a, aa

b, bb

ab, abab ab

ϵ $(1+011)^*$

{1, 11, 1111 ...

011, 011011, 110

1011, 11011, 1101

111011011

~~0111011111~~

$(a+b)$

1101

$(1+011)^*$

ϵ

$(110+1)^*$

110, 1, 011

1, 011, 111 11 1011011011011011

1, 011

ϕ $L(r) = \{\ \}$

$r = \varepsilon$ $L(r) = \{\varepsilon\}$

3 $r = a$ $L(r) = \{a\}$

4. $r = a+b$ $L(r) = \{a + b\}$

5. $r = a.b$ $Lr = \{ab\}$

6. $r = a+b+c$ $L(r) = \{a, b, c\}$

7. $r = (ab+a)b$ $\cancel{L(r)}$. $L(r) = \{abb + ab\}$

8 $r = a^+$ $L(r) = \{a, aa, aaa \ldots\}$

9 $r = a^*$ $L(r) = \{\varepsilon, a, aa, aaa \ldots\}$

10. $r = (a+ba)(b+a)$ $L(r) = \{ab, aa, bab, baa$

11 $r = (a+\varepsilon)(b\cancel{+\phi})$ $L(r) = \{ab, b\}$

imp $\begin{cases} a+\varepsilon = a, \varepsilon \\ a+\phi = a \end{cases}$

12. $r = (a+b)^2$. $L(r) = \{aa, ab, ba, bb\}$

$\Downarrow$

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

13. $r = (a+b)^*$ $L(r) = \{\wedge, a, b, aa, ab, ba \ldots\}$

(14) $r = (a+b)^*(a+b)$ $\cancel{L(r)} =$ $L(r) = \{a, b, aa, ab, ba, bb \ldots\}$

$\Downarrow$

$(a+b)^+$

15. $r = a^* a^* = a^*$ $L(r) = \{\varepsilon, a, aa, aaa \ldots\}$

$\{\varepsilon, a, aa \ldots\}\{\varepsilon, a, aa, aaa \ldots\}$

16 $r = (ab)^*$ = $L(r) = (\varepsilon, ab, abab, ababab \ldots\}$

Q $r_1 = a^*$ $\qquad r_2 = a^* + (aa)^*$

$\{\varepsilon, a, aa \ldots\}$ $\qquad \{\varepsilon, a, aa, aaa \ldots$ $\qquad \varepsilon, a^2, a^4 \ldots$

a) $L(r_1) \subseteq L(r_2)$ $\qquad$ two diff RE generate same language

b) $L(r_1) \supseteq L(r_2)$

c) $L(r_1) = L(r_2)$

d) $L(r_1) \neq L(r_2)$

If it is possible that for any language there are more than one RE.

Q. $R = \lambda + 1^*(011)^*(1^*(011)^*)^*$

$= (1 + 011)^*$

$1^*(011)^* = P1$

$R = \lambda + \underline{P1\ P1^*} \quad = \lambda + P1^+ = P1^*$

$= (1^*(011)^*)^* = (P2^* P3^*)^*$

$= (P2 + P3)^* = (1 + 011)^* \to \underline{RHS}$

$= \cancel{P^*}$

Q. Write the regular Expression over alphbet {a,b,c} containing at least one a & at least one b

→ $RE = (a+b+c)^* \, a \, (a+b+c)^* \, b \, (a+b+c)^* + (a+b+c)^* \, b \, (a+b+c)^* \, a \, (a+b+c)^*$

Q Write the regular Expression for the set of string of 0's & 1's whose tenth symbol from right end is 1

$L = \{0/1 \;\; 1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \;\; 0/1 \ldots\}$

$RE = (0+1)^* \, 1 \, (0+1)^9$

Q Write the regular expression for the set of string of an equal number of 0's & 1's such that in every prefix the no. of 0's differ from number of 1's by at most 1

$RE = (01 + 10)^*$

Q Write the regular Expression for the language $L = \{w \in \{0,1\}^* : w$ has no consio

Q Write the RE for the language $L = \{a^n b^m : (n+m)$ is even$\}$.

even + even = even ⇒ $(aa)^* (bb)^*$

odd + odd = even ⇒ $a(aa)^* \, b(bb)^*$

$(aa)^* (bb)^* + a(aa)^* \, b(bb)^*$

Q Write the RE over alphabet {a, b} for the set of strings with even number of a's followed by odd no. of b's thus is for a language

$L = \{a^{2n} b^{2m+1} : n \geq 0 \ \& \ m \geq 0.$

$(aa)^* b (bb)^*$

Q Write the RE for language

$L = \{a^n b^m \mid n \geq 4, m \leq 3\}.$

$a = 4, 5, 6, 7$ $\quad$ $b = \{0, 1, 2, 3\}.$

$a^4 (a)^* (\epsilon + b + b^2 + b^3)$

Q Write the regular Expression for the language

$L = \{ab^n w : n \geq 3 \ \ w \in (a, b)^+\}$

$RE = ab^3 b^* (a+b)^*$

Q Write the RE for the language

$L = \{w : |w| \bmod 3 = 0 \quad w \in \{a, b\}^*\}$

$0, 3, 6, 9$

$RE = ((a+b)^3)^*$

Q. Write RE in which every 0 is immediately followed by at least 2 1's.

$(1 + 011)^*$