

# Regular Expression

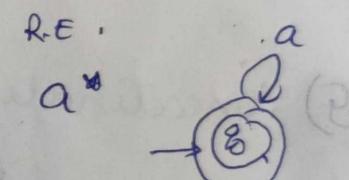
Part 1

↳ It is a method to represent a language.

means - If we want to represent a language, then we use a regular language.

→ Regular expression represent language (Regular language) All those languages which are accepted by FA (DFA | NFA, <sup>NFA with E</sup> mealy (moore)) All strings that are accepted that are regular language  
And To represent Regular language we use Regular Expression.

for ex  $L = \{ \text{a}, \text{aa}, \text{aaa} \dots \} \Rightarrow a^*$



Let 'R' be a Regular Expression over Alphabet

$\Sigma$  if R is :

(walling) collecting

- 1)  $\epsilon$  is Regular Expression denoting the set  $\{\epsilon\}$ .
- 2)  $\phi$  is " " " " empty set
- 3) for each symbol ~~a~~  $a \in \Sigma, a$  is regular

These three are primitive

$$\begin{cases}
 R = \epsilon \quad (R) = \{\epsilon\} \\
 R = \phi \quad (R) = \{\} \\
 R = a \quad (R) = \{a\}
 \end{cases}$$

4) Union of two Regular Expressions is a ~~not~~ Regular.

$$R_1 \cup R_2 = \text{Regular}$$

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

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

$\rightarrow x + y$  is RE corresponding to  $L(x) \cup L(y)$   
where  $L(x \cup y) = L(x) \cup L(y)$

$\rightarrow x \cdot y$  is RE corresponding to  $L(x) \cdot L(y)$   
where  $L(x \cdot y) = L(x) \circ L(y)$

5) Concatenation of two RE is also Regular

$$R_1 R_2 = \text{Regular}$$

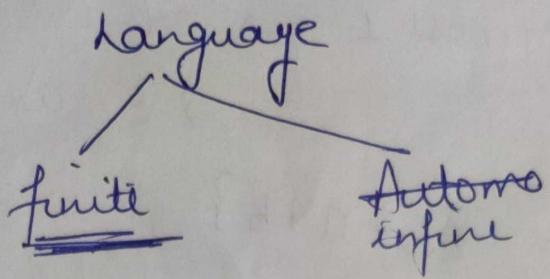
6).  $\rightarrow R^*$  is RE corresponding to  $L(R^*) = (L(R))^*$

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

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

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

eeg



- ↳ If lang. finite it must be Regular ✓
  - ↳ " " " then we create FA ✓
  - ↳ " " " then we write RE ✓
- $\Sigma = \{a, b\}$

1) No string → {}  
 $RE = \emptyset$

1) length 0 = { $\epsilon$ }

$$RE = \epsilon$$

2) length 1 = {a, b}

$$RE = (a+b)$$

Finite lang. on RE  
 $\Rightarrow$  union one RE

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

$$\begin{aligned} RE &= (ab + ba + aa + bb) \text{ or } (a+b)(a+b) \\ &= a(a+b) + b(a+b). \end{aligned}$$

5) length 3 = {aaa, aab, aba, abb, bbb, bab, baab, bbaa}

$$\therefore = (a+b)(a+b)(a+b)$$

There are  
string

6) length of string Atmost 1  $\Rightarrow (0,1)$   
 $\{ \epsilon, a, b \}$

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

7) At most 2  $\Rightarrow (\epsilon + a + b)(\epsilon + a + b)$

8) no. of b not more than 2 b's & 1 a  
 $\{ \epsilon, a, b, ab, ba, bba, bab, abb \}$

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

RE for infinite language

Language can be regular & not

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

1). All the strings having a single 'b'  
 $= \{ \epsilon, a, aa, aaa, aaaa \dots \} b. \{ \epsilon, a, aa \dots \}$

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

2). All the strings having at least one 'b'  
 $L = (a+b)^* b (a+b)^*$

3). All the strings having any no. of a &  
 and any no. of b's  
 $(a+b)^*$   
 $L = \{ \epsilon, a, b, aa, bb, \dots \}$

use mostly  
 $(*)$   $(+)$   
 Powr.

$$\begin{cases} aa^* = a^+ \\ = aa^0 \\ = a^* \epsilon \\ = a \end{cases}$$

• 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'

$$RE = a^* b b a^* \quad \cancel{abba} \quad \cancel{aabba}$$

✓  $RE = a^* b a^* b a^*$

~~abba~~  
~~aabb~~  
~~aaab~~

July Re / Rule 1 -  
Describe the following set by RE :

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

2)  $\{abbab\}$   $RE = \{abbab\}$

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

4)  $\{\lambda, ab\}$   $RE = \lambda + ab$

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

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

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

---

### ARDEN'S Theorem.

↳ To check the equivalence two RE

↳ Conversion of DFA to RE then use Arden's  
Before Applying the theorem we need to follow  
2 conditions.

{ \* FA should not contain  $\epsilon$  transition

{ \* FA should have only one initial state.

If finite automaton satisfy these two condition then  
we directly use Arden's theorem.

## Mutliplication Rule

$$\text{Rule 1} - \phi + r = R$$

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

$$3 = \varepsilon \cdot r = r \cdot \varepsilon = r \quad | \quad \wedge \cdot r = r \cdot \wedge = r$$

$$4 = \sigma + r = r$$

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

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

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

$$\wedge^* = \wedge$$

$$\phi^* = \wedge$$

$$8 = \varepsilon + \sigma^* \cdot r = \varepsilon + r \cdot r^* = \sigma^*$$

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

$$10 = (P+Q)^* = (P^* Q^*)^* = \underline{(P^* + Q^*)^*}$$

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

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

Arden's Theorem : If  $P$  &  $Q$  are two RE &  $P$  doesn't have  $\varepsilon$ -transition then equation  $R = QP^*$  will have unique solution  $\frac{R = QP^*}{\downarrow \text{eq } ①}$

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 neither but  $\varepsilon$ .

Application RE is.

↳ Validation of date

Ex compiler.

compiler some set of ~~rules~~ rules we need to write a program that observes that particular rule if any rule is violated 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 noo of 'b's over  $\Sigma = \{a, b\}$

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

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

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

$$R = Q + RP$$

— ①.

$$R = Q + RP$$

Apply eq ①

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

Apply eq ①

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

Apply eq ①

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

$$R = Q(1 + P + P^2 + P^3 - \dots) - QP^4 + RP^5$$

$$R = QP^*$$

15.

16.

$(a+b)$

$a, a^m$

$b, b^m$

$ab, ab^m$

$ab^2, ab^3$

$ab^4, ab^5$

$ab^6, ab^7$

$ab^8, ab^9$

$ab^{10}, ab^{11}$

$ab^{12}, ab^{13}$

$ab^{14}, ab^{15}$

$ab^{16}, ab^{17}$

$ab^{18}, ab^{19}$

$ab^{20}, ab^{21}$

$ab^{22}, ab^{23}$

$ab^{24}, ab^{25}$

$ab^{26}, ab^{27}$

$ab^{28}, ab^{29}$

$ab^{30}, ab^{31}$

$ab^{32}, ab^{33}$

$ab^{34}, ab^{35}$

$ab^{36}, ab^{37}$

$ab^{38}, ab^{39}$

$ab^{40}, ab^{41}$

$ab^{42}, ab^{43}$

$ab^{44}, ab^{45}$

$ab^{46}, ab^{47}$

$ab^{48}, ab^{49}$

$ab^{50}, ab^{51}$

$ab^{52}, ab^{53}$

$ab^{54}, ab^{55}$

$ab^{56}, ab^{57}$

$ab^{58}, ab^{59}$

$ab^{60}, ab^{61}$

$ab^{62}, ab^{63}$

$ab^{64}, ab^{65}$

$ab^{66}, ab^{67}$

$ab^{68}, ab^{69}$

$ab^{70}, ab^{71}$

$ab^{72}, ab^{73}$

$ab^{74}, ab^{75}$

$ab^{76}, ab^{77}$

$ab^{78}, ab^{79}$

$ab^{80}, ab^{81}$

$ab^{82}, ab^{83}$

$ab^{84}, ab^{85}$

$ab^{86}, ab^{87}$

$ab^{88}, ab^{89}$

$ab^{90}, ab^{91}$

$ab^{92}, ab^{93}$

$ab^{94}, ab^{95}$

$ab^{96}, ab^{97}$

$ab^{98}, ab^{99}$

$ab^{100}, ab^{101}$

$ab^{102}, ab^{103}$

$ab^{104}, ab^{105}$

$ab^{106}, ab^{107}$

$ab^{108}, ab^{109}$

$ab^{110}, ab^{111}$

$ab^{112}, ab^{113}$

$ab^{114}, ab^{115}$

$ab^{116}, ab^{117}$

$ab^{118}, ab^{119}$

$ab^{120}, ab^{121}$

$ab^{122}, ab^{123}$

$ab^{124}, ab^{125}$

$ab^{126}, ab^{127}$

$ab^{128}, ab^{129}$

$ab^{130}, ab^{131}$

$ab^{132}, ab^{133}$

$ab^{134}, ab^{135}$

$ab^{136}, ab^{137}$

$ab^{138}, ab^{139}$

$ab^{140}, ab^{141}$

$ab^{142}, ab^{143}$

$ab^{144}, ab^{145}$

$ab^{146}, ab^{147}$

$ab^{148}, ab^{149}$

$ab^{150}, ab^{151}$

$ab^{152}, ab^{153}$

$ab^{154}, ab^{155}$

$ab^{156}, ab^{157}$

$ab^{158}, ab^{159}$

$ab^{160}, ab^{161}$

$ab^{162}, ab^{163}$

$ab^{164}, ab^{165}$

$ab^{166}, ab^{167}$

$ab^{168}, ab^{169}$

$ab^{170}, ab^{171}$

$ab^{172}, ab^{173}$

$ab^{174}, ab^{175}$

$ab^{176}, ab^{177}$

$ab^{178}, ab^{179}$

$ab^{180}, ab^{181}$

$ab^{182}, ab^{183}$

$ab^{184}, ab^{185}$

$ab^{186}, ab^{187}$

$ab^{188}, ab^{189}$

$ab^{190}, ab^{191}$

$ab^{192}, ab^{193}$

$ab^{194}, ab^{195}$

$ab^{196}, ab^{197}$

$ab^{198}, ab^{199}$

$ab^{200}, ab^{201}$

$ab^{202}, ab^{203}$

$ab^{204}, ab^{205}$

$ab^{206}, ab^{207}$

$ab^{208}, ab^{209}$

$ab^{210}, ab^{211}$

$ab^{212}, ab^{213}$

$ab^{214}, ab^{215}$

$ab^{216}, ab^{217}$

$ab^{218}, ab^{219}$

$ab^{220}, ab^{221}$

$ab^{222}, ab^{223}$

$ab^{224}, ab^{225}$

$ab^{226}, ab^{227}$

$ab^{228}, ab^{229}$

$ab^{230}, ab^{231}$

$ab^{232}, ab^{233}$

$ab^{234}, ab^{235}$

$ab^{236}, ab^{237}$

$ab^{238}, ab^{239}$

$ab^{240}, ab^{241}$

$ab^{242}, ab^{243}$

$ab^{244}, ab^{245}$

$ab^{246}, ab^{247}$

$ab^{248}, ab^{249}$

$ab^{250}, ab^{251}$

$ab^{252}, ab^{253}$

$ab^{254}, ab^{255}$

$ab^{256}, ab^{257}$

$ab^{258}, ab^{259}$

$ab^{260}, ab^{261}$

$ab^{262}, ab^{263}$

$ab^{264}, ab^{265}$

$ab^{266}, ab^{267}$

$ab^{268}, ab^{269}$

$ab^{270}, ab^{271}$

$ab^{272}, ab^{273}$

$ab^{274}, ab^{275}$

$ab^{276}, ab^{277}$

$ab^{278}, ab^{279}$

$ab^{280}, ab^{281}$

$ab^{282}, ab^{283}$

$ab^{284}, ab^{285}$

$ab^{286}, ab^{287}$

$ab^{288}, ab^{289}$

$ab^{290}, ab^{291}$

$ab^{292}, ab^{293}$

$ab^{294}, ab^{295}$

$ab^{296}, ab^{297}$

$ab^{298}, ab^{299}$

$ab^{300}, ab^{301}$

$ab^{302}, ab^{303}$

$ab^{304}, ab^{305}$

$ab^{306}, ab^{307}$

$ab^{308}, ab^{309}$

$ab^{30}, ab^{31}$

$ab^{32}, ab^{33}$

$ab^{34}, ab^{35}$

$ab^{36}, ab^{37}$

$ab^{38}, ab^{39}$

$ab^{40}, ab^{41}$

$ab^{42}, ab^{43}$

$ab^{44}, ab^{45}$

$ab^{46}, ab^{47}$

$ab^{48}, ab^{49}$

$ab^{50}, ab^{51}$

$ab^{52}, ab^{53}$

$ab^{54}, ab^{55}$

$ab^{56}, ab^{57}$

$ab^{58}, ab^{59}$

$ab^{60}, ab^{61}$

$ab^{62}, ab^{63}$

$ab^{64}, ab^{65}$

$ab^{66}, ab^{67}$

$ab^{68}, ab^{69}$

$ab^{70}, ab^{71}$

$ab^{72}, ab^{73}$

$ab^{74}, ab^{75}$

$ab^{76}, ab^{77}$

$ab^{78}, ab^{79}$

$ab^{80}, ab^{81}$

$ab^{82}, ab^{83}$

$ab^{84}, ab^{85}$

$ab^{86}, ab^{87}$

$ab^{88}, ab^{89}$

$ab^{90}, ab^{91}$

$ab^{92}, ab^{93}$

$ab^{94}, ab^{95}$

$ab^{96}, ab^{97}$

$ab^{98}, ab^{99}$

$ab^{100}, ab^{101}$

$ab^{102}, ab^{103}$

$ab^{104}, ab^{105}$

$ab^{106}, ab^{107}$

$ab^{108}, ab^{109}$

$ab^{10}, ab^{11}$

$ab^{12}, ab^{13}$

$ab^{14}, ab^{15}$

$ab^{16}, ab^{17}$

$ab^{18}, ab^{19}$

$ab^{20}, ab^{21}$

$ab^{22}, ab^{23}$

$ab^{24}, ab^{25}$

$ab^{26}, ab^{27}$

$ab^{28}, ab^{29}$

$ab^{30}, ab^{31}$

$ab^{32}, ab^{33}$

$ab^{34}, ab^{35}$

$ab^{36}, ab^{37}$

$ab^{38}, ab^{39}$

$ab^{40}, ab^{41}$

$ab^{42}, ab^{43}$

$ab^{44}, ab^{45}$

$ab^{46}, ab^{47}$

$ab^{48}, ab^{49}$

$ab^{50}, ab^{51}$

$ab^{52}, ab^{53}$

$ab^{54}, ab^{55}$

$ab^{56}, ab^{57}$

$ab^{58}, ab^{59}$

$ab^{60}, ab^{61}$

$ab^{62}, ab^{63}$

$ab^{64}, ab^{65}$

$ab^{66}, ab^{67}$

$ab^{68}, ab^{69}$

$ab^{70}, ab^{71}$

$ab^{72}, ab^{73}$

$ab^{74}, ab^{75}</math$

$$r = \emptyset \quad L(r) = \{\emptyset\}$$

$$r = \epsilon \quad L(r) = \{\epsilon\}$$

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

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

$$5. \quad r = a \cdot b \quad L(r) = \{ab\}$$

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

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

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

$$9. \quad r = a^* \quad L(r) = \{\epsilon, a, aa, aaa, \dots\}$$

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

$$11. \quad r = (a+\epsilon)(b+\epsilon) \quad L(r) = \{ab, b\}$$

$$\text{q.e.d.} \quad \begin{cases} a+\epsilon = a, \epsilon \\ a+\emptyset = a \end{cases}$$

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

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

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

$$15. \quad r = a^* a^* = a^* \quad L(r) = \{\epsilon, a, aa, aaa, \dots\}$$

$$\{\epsilon, a, aa, \dots\} \quad \{\epsilon, a, aa, aaa, \dots\}$$

$$16. \quad r = (ab)^* \quad L(r) = \{\epsilon, ab, abab, ababab, \dots\}$$

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

$\{\epsilon, a, aa, \dots\}$

(While  
2a, b, c, d contain  
last one.)

- a)  $L(r_1) \subseteq L(r_2)$
- b)  $L(r_1) \supseteq L(r_2)$
- c)  $L(r_1) = L(r_2)$
- d)  $L(r_1) \neq L(r_2)$

two diff RE generate  
same language

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

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

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

$$1^* (011)^* = P_1$$

$$R = \lambda + \frac{P_1}{P_1} = \lambda + P_1^+ = P_1^+$$

$$= (1^* (011)^*)^* = (P_2^* P_3^*)^*$$

$$= (P_2 + P_3)^* = (1+011)^* \rightarrow \underline{\text{RMS}}$$

$$= \#$$

Write the regular expression over alphabet  $\{a, b, c\}$  containing at least one a & at least one b

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

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

$$L = \{01, 101, 011, 011, 011, 011, 011, 011, 011, 011, \dots\}$$

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

Q Write the regular expression for the set of strings 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  $L = \{w \in \{0, 1\}^*\mid w \text{ has no consecutive } 0s\}$

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

$$\text{even + even} = \text{even}$$

$$\text{odd + odd} = \text{even}$$

$$\Rightarrow (aa)^* (bb)^*$$

$$\Rightarrow a(aa)^* b(bb)^*$$

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

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

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

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

Q Write the RE for language

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

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

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

Q Write the regular Expression for the language

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

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

Q Write the RE for the language

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

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

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

$$(1+011)^*$$