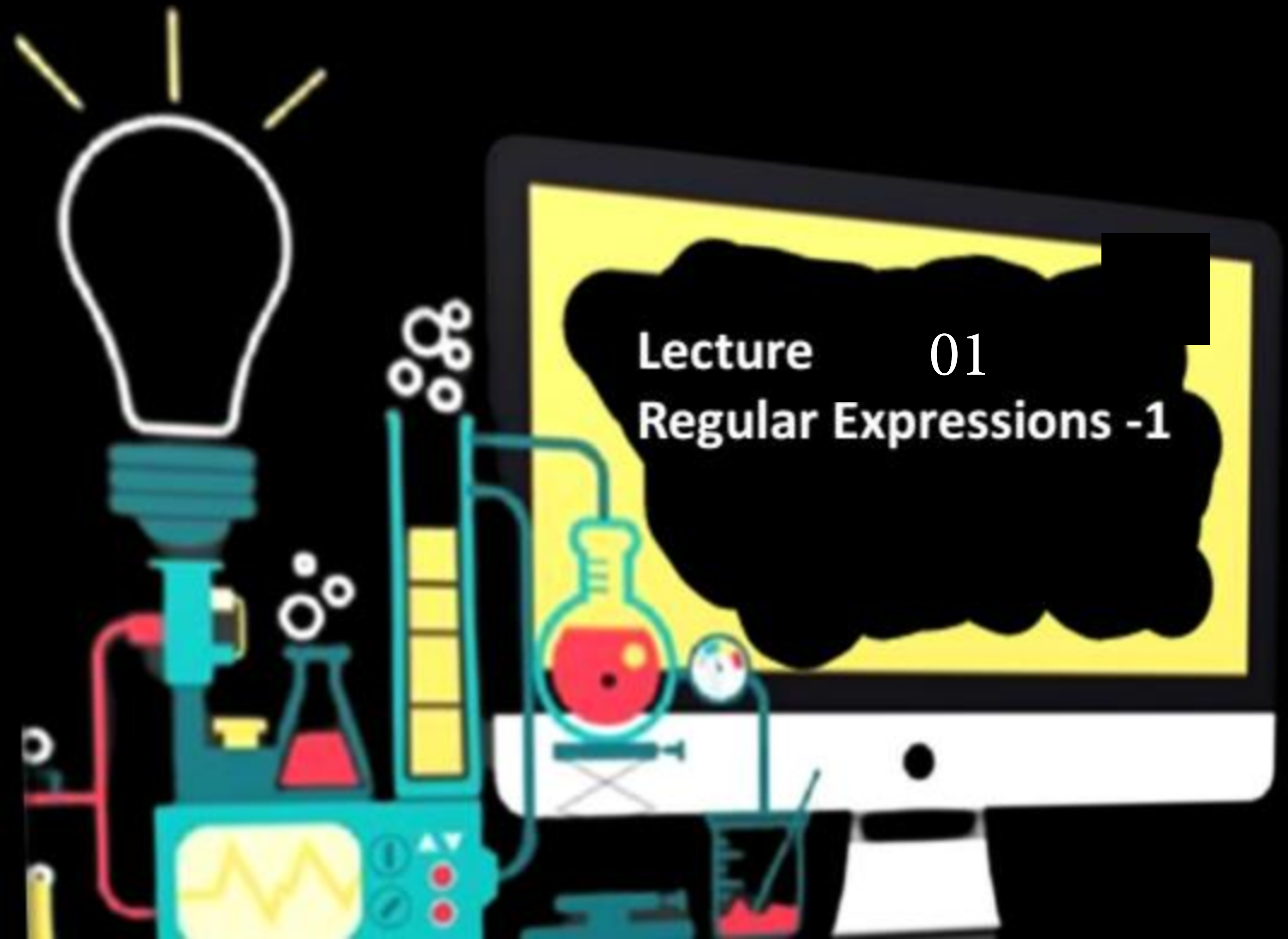
Theory of Computation



Lecture 01
Regular Expressions -1

Deva sir

Topics to be covered

- Operations on Strings
- Regular Expresions

Topic: Basics

Symbol : Smallest

(Σ) Alphabet : set of symbols [$|\Sigma| =$ finite]

→ String (word) : Sequence of symbols

(Set) Language : collection of strings

Automata : It is a machine that represents a Set

Grammar : It is collection of productions (rules)

Note:

Topic: Strings

Prefix (Init) Beginning Sequence	Suffix ending Sequence	Substring (Subword) part of string
Begin	end	
$w = abc$ 3 length	$w=abc$	$w=abc$
Prefixes of w:	suffixes of w:	Substrings of w:
ε, a, ab, abc — 4 prefixes	ε, c, bc, abc — 4 suffixes	ε, a, b, c, ab, bc, abc
$\text{Prefix}(w) = \{u \mid \underbrace{u}_{\text{beginning}}v = w\}$	$\text{Suffix}(w) = \{v \mid u\underbrace{v}_{\text{ending}} = w\}$	$\text{Substring}(w) = \{y \mid xyz = w\}$

ac is subsequence

not substring

Note:

Topic: String

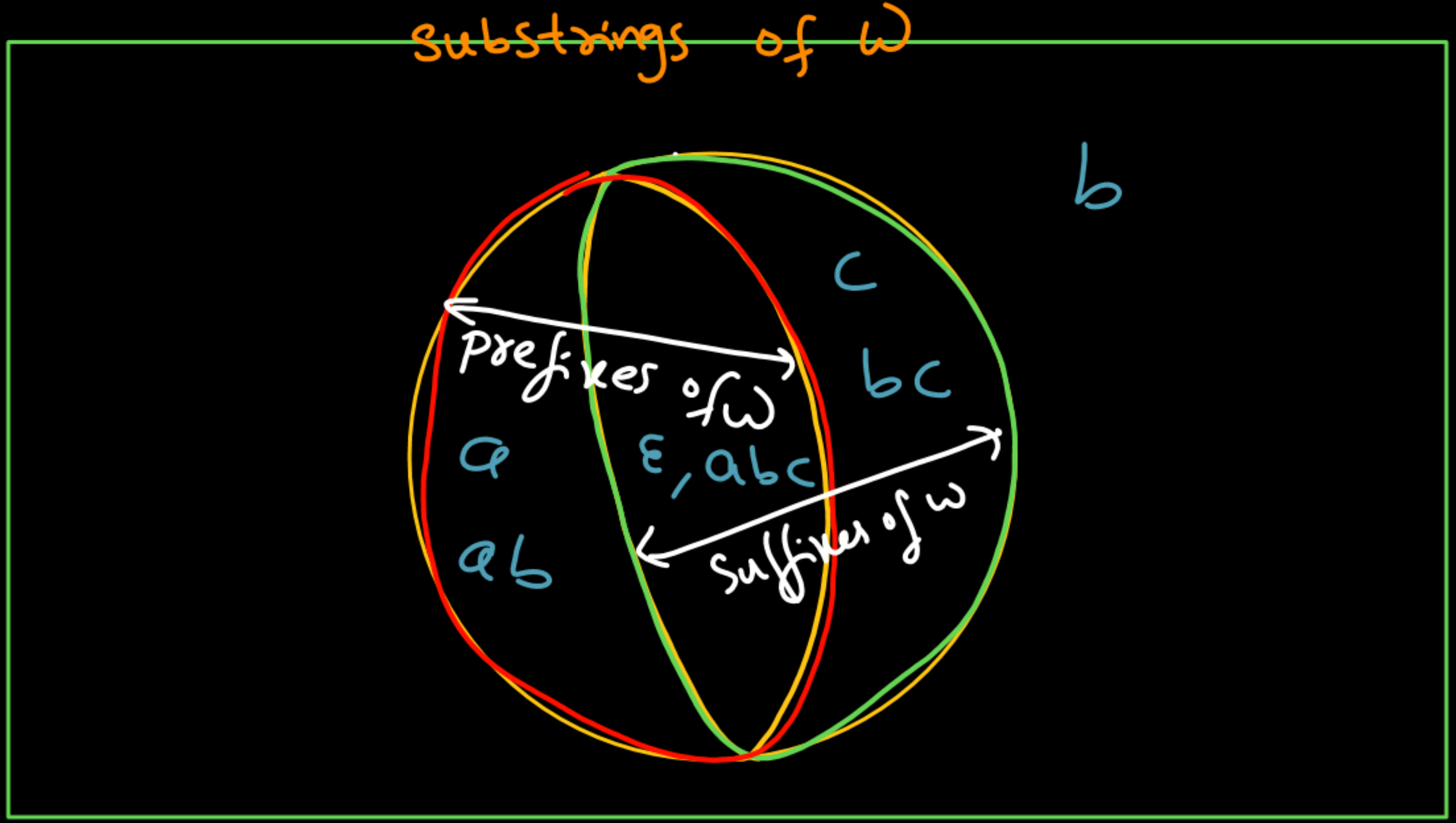
$w = abc$

Prefixes:

ε
a
ab
abc

Suffixes:

ε
c
bc
abc



Every prefix is substring
Every suffix is substring

Substrings

ε
a
b
c
ab
bc
abc

Note:

Topic: How many strings?

Let $|w| = n$. Then
(w → string)

Q1) How many prefixes of w? $= n+1$

Q2) " " Suffixes of w? $= n+1$

Q3) " " Substrings of w?

- $\boxed{\text{Min } n+1}$ when all n symbols are same
- $\boxed{\text{Max } \frac{n(n+1)}{2}+1}$ when all n symbols are distinct in w

Note:

Topic: Substrings

$w = aaaa$ (n)

ε

a

aa

aaa

aaaa

5 substrings

$(n+1)$

$w = abcd$ (n)

$\varepsilon \rightarrow 1$ zero len

n one length: a, b, c, d

n-1 2 length: ab, bc, cd

n-2 3 length: abc, bcd

n-3 4 length: abcd

$\Rightarrow$ Sum of n numbers $= \frac{n(n+1)}{2}$

Note:

Topic:

$\star\star\star$

Q4) How many different length prefixes ? $= n+1$

Q5) " " " " Suffixes ? $= n+1$

Q6) " " " " Substrings ? $= n+1$

$W = aab$	prefixes	Suffixes	Substrings
Strings	ε, a, aa, aab	ε, b, ab, aab	ε, a, b, aa, ab, aab
different lengths	0, 1, 2, 3	0, 1, 2, 3	0, 1, 2, 3
No. of different lengths	4	4	4

Note:

Topic: Language (L)

- → It is set of strings over Σ
- → It represents a problem
- → It is a subset of Σ^*

$$\boxed{L \subseteq \Sigma^*}$$

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \ldots$

Σ^0 → Zero length strings

Σ^1 → One length strings

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

strings : $\varepsilon, a, b, aa, ab, ba, bb, \ldots$

$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$ = set of all strings [universal set]

Note:

Topic:

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

$\boxed{L \subseteq \Sigma^*}$

$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$

$L_1 = \{\ \}$	$A_1 = \{a, \varepsilon\}$	$B_1 = \{\varepsilon, a, b\}$	- - - -
$L_2 = \{\varepsilon\}$	$A_2 = \{a, b\}$	$B_2 = \{\varepsilon, a, aa\}$	
$L_3 = \{a\}$	$A_3 = \{aa, \varepsilon\}$	$\vdots$	
$L_4 = \{b\}$	$\vdots$		
$L_5 = \{aa\}$			
$L_6 = \{ab\}$			
$\vdots$			

Note:

Topic: Automata

FA

DFA $\cong$ NFA

PDA

DPDA $<$ PDA (NPDA)

LBA

DLBA ? NLBA

TM

DTM $\cong$ NTM

Note:

Topic: Regular Expressions

What is Regular Expression?

- → It represents a regular set.
  (denotes)
  (generates)
- → It uses 4 operators to represent regular set.

Regular Language

$\cong$

Regular Expression

$\Downarrow \;\cong\; \Uparrow$

Finite Automata

$\Downarrow \;\cong\; \Uparrow$

Regular Grammar

Note:

C expression

C++ expression

python expression

DBMS

$\pi_{\ldots}(A \times B)$

Digital logic

$x + y \cdot z$

propositional logic

$x \vee y \wedge z$

mathematics

$x + y \times z$

$$[X] + [Y] \cdot [Z]$$

Expression

└→ $\boxed{\text{operators}}$ & operands

- → unary
- → Binary
- ⋮
- → n-ary

Topic:

what is + ?

1) which subject ?

2) which topic ?

3) which context ?

In digital:

$1 \cdot 1 = 1$

In TOC :

$1 \cdot 1 = 11$

$1 \cdot 1 \neq 1$

$$\underbrace{a + b \cdot c}$$

Digital

$+ \rightarrow$ OR

$\cdot \rightarrow$ AND

$a + bc = (a+b) \cdot (a+c)$ ✓

TOC

$+ \rightarrow$ OR

$\cdot \rightarrow$ Concatenation

$a + bc \neq (a+b) \cdot (a+c)$

Note:

Topic: Regular Expression $\Rightarrow$ 4 operators

OR	Concatenation	Kleene Star	Kleene plus
$(+)$	$(\cdot)$	$(*)$	$(+)$
Binary	Binary	Unary	Unary
$a+b$	$a.b$	a^*	a^+

Note:

Topic: OR $(+)$ $(\cup)$ $(\mid)$

Expression

$R_1 + R_2$

$R_1 \cup R_2$

$R_1 \mid R_2$

$\boxed{\text{Either } R_1 \text{ or } R_2}$

Set

$L(R_1) \cup L(R_2)$

Note:

Topic:

$\underline{a + b} \Longrightarrow \{a, b\}$

- $\rightarrow a$ ✓
- $\rightarrow b$ ✓
- $\rightarrow ab$ ✗
- $\rightarrow ba$ ✗
- $\rightarrow \varepsilon$ ✗
- $\rightarrow aa$ ✗
- $\rightarrow bb$ ✗

Note:

Topic: OR simplification

① $a + a = a$

② $a + b = b + a$

③ $a + \phi = a$

④ $a + \varepsilon = a + \varepsilon = \varepsilon + a$

⑤ $\phi + \phi = \phi$

⑥ $\phi + \varepsilon = \varepsilon$

⑦ $\varepsilon + \varepsilon = \varepsilon$

any regular exp

⑧ $\boxed{R + \phi = R}$

⑨ $\boxed{R + \varepsilon = \varepsilon + R}$

If R generates ε then $R+\varepsilon=R$

Otherwise $R+\varepsilon \neq R$

$\phi \rightarrow$ empty expression

$\cong$

empty set

Note:

Topic:

Regular Exp	Regular set
ϕ	$\{\ \} = \phi$
ε	$\{\varepsilon\}$
a	$\{a\}$
$a+b$	$\{a, b\}$
$a+\varepsilon$	$\{a, \varepsilon\}$
$a+\phi = a$	$\{a\} \cup \{\ \} = \{a\}$
$\phi + \varepsilon = \varepsilon$	$\{\ \} \cup \{\varepsilon\} = \{\varepsilon\}$
$\varepsilon + \varepsilon = \varepsilon$	$\{\varepsilon\} \cup \{\varepsilon\} = \{\varepsilon\}$

Note:

Topic: Concatenation $(\cdot)$

$R_1 \cdot R_2$ $\quad$ R_1 followed by R_2

$\boxed{a \cdot b}$

- $\rightarrow$ a ×
- $\rightarrow$ b ×
- $\rightarrow$ ab ✓
- $\rightarrow$ ba ×
- $\rightarrow$ ε ×
- $\vdots$

correct $R_1 + R_2 = R_2 + R_1$

wrong $R_1 \cdot R_2 = R_2 \cdot R_1$

Note:

Topic:

$R+R=R$

$R+\phi=R$

$$L(R_1 \cdot R_2) = L(R_1) \cdot L(R_2)$$

$$L(R_1+R_2) = L(R_1) \cup L(R_2)$$

$R \cdot \phi = \phi \cdot R = \phi$

$R \cdot \varepsilon = \varepsilon \cdot R = R$

1. $\phi \cdot \phi = \phi$
2. $\phi + \phi = \phi$
3. $\varepsilon \cdot \varepsilon = \varepsilon$
4. $\varepsilon + \varepsilon = \varepsilon$
5. $\phi \cdot \varepsilon = \varepsilon \cdot \phi = \phi$
6. $\phi + \varepsilon = \varepsilon$
7. $\phi \cdot a = \phi$
8. $\phi + a = a$
9. $\varepsilon \cdot a = a$
10. $\varepsilon + a = \varepsilon + a$

Note:

Topic:

I

$$\phi \cdot a = \phi$$

$$\{\ \} \cdot \{a\} = \{\ \}$$

II

$$\varepsilon \cdot a = a$$

$$\{\varepsilon\} \cdot \{a\} = \{\varepsilon a\} = \{a\}$$

Note:

Topic:

properties	OR	Concatenation
① Identity	$R+\phi=\phi+R=R$ $\boxed{\phi \text{ is identity}}$	$R.\varepsilon=\varepsilon.R=R$ $\boxed{\varepsilon \text{ is identity}}$
② Commutative	$R_1+R_2=R_2+R_1$ $\boxed{\text{Holds}}$	$R_1R_2=R_2R_1$ ✗ $\boxed{\text{not holds}}$
③ Associative	$(R_1+R_2)+R_3=R_1+(R_2+R_3)$ $\boxed{\text{Holds}}$	$(R_1.R_2).R_3=R_1.(R_2.R_3)$ $\boxed{\text{Holds}}$
④ Annihilator (Dominator)	$R+\Sigma^*=\Sigma^*$ Σ^* is dominator for +	$R.\phi=\phi.R=\phi$ ϕ is dominator for .

Note:

Topic:

(5) Distribution

I) OR over concatenation

Not holds

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

$$\{a, bc\} \neq \{aa, ac, ba, bc\}$$

II) Concatenation over OR

Holds

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

$$\{ab, ac\} \qquad \{ab, ac\}$$

Note:

Topic: OR , concatenation

1. $\boxed{a}+\boxed{aa} = a+aa$
2. $\boxed{a.a}+\boxed{a} = aa+a$
3. $a+\underbrace{a\phi} = a+\phi = a$
4. $a+\underbrace{a\varepsilon} = a+a = a$
5. $\phi+aa = aa$
6. $\phi+a\varepsilon = a\varepsilon = a$

7. $\phi.aa = \phi$
8. $\varepsilon.aa = aa$
9. $\underbrace{\phi.a}_{\phi}+a = a$
10. $\phi+\underbrace{\varepsilon.a}_{a} = a$
11. $\varepsilon+\underbrace{\phi.\varepsilon}_{\phi} = \varepsilon$
12. $\phi+\underbrace{\varepsilon.\phi}_{\phi} = \phi$

Higher .

$+ \downarrow$

Note:

Topic: R^* R^+

$$\boxed{R^0 = \phi^0 = \varepsilon^0 = a^0 = b^0 = \varepsilon}$$

Kleene star of R

Kleene closure of R

$R^* = R^{\geq 0}$

$= R^0 + R^1 + R^2 + \dots$

R^*

$L(a^*) = \{a^0, a^1, a^2, \dots\}$

$= \{\varepsilon, a, aa, \dots\}$

$L(\varepsilon^*) = \{\varepsilon^0, \varepsilon^1, \varepsilon^2, \dots\}$

$= \{\varepsilon\}$

Kleene plus of R

Positive closure of R

R^+

$R^+ = R^{\geq 1}$

$= R^1 + R^2 + R^3 + \dots$

$L(a^+) = \{a^1, a^2, a^3, \dots\}$

$= \{a, aa, aaa, \dots\}$

$L(\varepsilon^+) = \{\varepsilon^1, \varepsilon^2, \dots\}$

$= \{\varepsilon\}$

Note:

Topic:

$$L((ab)^*) = \{(ab)^0, (ab)^1, (ab)^2, \ldots\}$$

$$= \{\varepsilon, ab, abab, \ldots\}$$

$$\boxed{(ab)^* = \varepsilon + ab + abab + \ldots}$$

$$(ab)^+ = ab + abab + ababab + \ldots$$

$$R^2 = R.R$$

$$R^3 = R.R.R$$

$$\boxed{2^3 = 222}$$

Note:

Topic: $\phi^0 = \varepsilon^0 = a^0 = R^0 = \varepsilon$ $\quad \boxed{aa^* = a^*a = a^+}$

① $\phi^* = \underbrace{\boxed{\phi^0}}_{\varepsilon} + \underbrace{\phi^1 + \phi^2 + \cdots}_{\phi} = \varepsilon$

② $\phi^+ = \phi$

equivalent
③ $\varepsilon^* = \varepsilon^0 + \varepsilon^1 + \varepsilon^2 + \cdots = \varepsilon$

④ $\varepsilon^+ = \varepsilon$

Same
⑤ $\underset{\text{must}}{a} \cdot \underset{\text{must}}{a^*} = a \cdot (a^0 + a^1 + a^2 + \cdots)$

$= a(\varepsilon + a + aa + \cdots) = a + a^2 + a^3 + \cdots$

$= a^+$

⑥ $\underset{\text{must}}{a^*} \cdot \underset{\text{must}}{a} = (\varepsilon + a + a^2 + \cdots) \cdot a = a + a^2 + a^3 + \cdots = a^+$

⑦ $\boxed{a} + a^* = a^*$

⑧ $a^+ + \boxed{a} = a^+$

H.W.
* ⑨ $(a+\varepsilon)^* = a^*$

H.W.
* ⑩ $(a+\varepsilon)^+ = a^*$

$a + \varepsilon \neq a$

Note:

Topic:

$(a+\varepsilon)^*$

- $\varepsilon \leftarrow (a+\varepsilon)^0$ ✓
- a ✓
- aa ✓
- aaa ✓
- $\vdots$

$= a^*$

$(a+\varepsilon)^2 = (a+\varepsilon).(a+\varepsilon)$

- aa ✓
- a ✓
- ε ✓

$$\boxed{\begin{gathered} a+\varepsilon \neq a \\ (a+\varepsilon)^* = a^* \end{gathered}}$$

Note:

Thank You