REGULAR EXPRESSION

MODULE 2

Regular Expression

- Regular Language:
  - A language is regular if there exists a finite acceptor for it
  - Hence, every regular language can be described by using some NFA/DFA
- Regular Expression:
  - One of the way of describing regular languages which consists of strings, symbols and operators
  - This notation involves:
    - Combination of strings of symbols from some alphabet Σ
    - Parentheses
    - Operators $+$, $\cdot$ and $*$

Regular Expression

Examples of Regular Expression:

1. For language L1 $= \{ a \}$

   Regular Expression $R = a$

2. Language L2 $= \{ a, b, c \}$

   R. E. $R = a + b + c$

   Here + represents union operation

   Similarly . represents concatenation

   * represents star – closure

Regular Expression

Examples of Regular Expression:

3. Suppose R = (a + (b . c))*

   It stands for star closure of { a } U { bc }

   Then language will be

   L = { €, a, bc, bca, abc, aa, aaa, bcbc, aabc, … }

Regular Expression

Formal definition:

Let Σ be a given alphabet then

1. Φ, €, and a € Σ are all regular expressions.

   These are called Primitive Regular Expression

2. If r1 and r2 are regular expressions

   then r1 + r2, r1 . r2, r1* and (r1) are all regular expressions

3. A string is a regular expression if and only if it can be derived from the primitive regular expression by a finite number of applications of the rules defined in statement 2

Regular Expression

Example:

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

Then the string $(a + b . c)^* . (c + a)$ is a regular expression

$(a + b +)$ is not a regular expression

Language associated with Regular Expression

- Regular expressions can be used to describe some simple languages.
- If r is a regular expression then L (r) denote the language associated with r
- The language L (r) denoted by any regular expression r is defined by following rules:
  1. Φ is a regular expression denoting the empty set
  2. € is a regular expression denoting { € }
  3. For every a € Σ, a is a regular expression denoting { a }

Language associated with Regular Expression

If r1 and r2 are regular expression then

1. $L(r1 + r2) = L(r1) \cup L(r2)$
2. $L(r1 . r2) = L(r1) . L(r2)$
3. $L(r1^*) = (L(r1))^*$
4. $L((r1)) = L(r1)$

Language associated with Regular Expression

- r* is known as kleen closure which indicate occurrences of r for ∞ number of times
- Examples:

  1. If Σ = { a }   and   R = a*

     Then   R = { €, a, aa, aaa, aaaa, … }

  2. If Σ = { a }   and   R = a+

     Then R = { a, aa, aaa, aaaa, … }

     This is known as Positive Closure

Examples On Regular Expression

- Ex. 1 Exhibit the language in set notation

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

Solution:

$$\begin{aligned} L(a^* . (a + b)) &= L(a^*) . L(a + b) \\ &= (L(a))^* . (L(a) \cup L(b)) \\ &= \{\epsilon, a, aa, aaa, \ldots\} . \{a, b\} \\ &= \{a, aa, aaa, \ldots, b, ab, aab, aaab, \ldots\} \end{aligned}$$

Examples On Regular Expression

Ex. 2 Represent the language for given regular expression

$r = (a + b)^* . (a + bb)$ where $\Sigma = \{ a, b \}$

Solution:

$$
\begin{aligned}
L (r) &= L ((a + b)^* . (a + bb)) \\
&= L ((a + b)^*) . L (a + bb) \\
&= (L (a + b))^* . (L (a) \cup L (bb)) \\
&= (L (a) \cup L (b))^* . (\{ a, bb \}) \\
&= (\{ a, b \})^* . \{ a, bb \} \\
&= \{ \epsilon, a, b, aa, ab, ba, \ldots \} . \{ a, bb \} \\
&= \{ a, bb, aa, abb, ba, bbb, aaa, aabb, aba, abbb, \ldots \\
\}
\end{aligned}
$$

L (r) is a language consists of all the strings terminated by either a or bb

Examples On Regular Expression

Ex. 3 Exhibit the language for given regular expression

$r = (aa)^* . (bb)^* . b$

Solution:

$L (r) = \{ b, aab, bbb, aabbb, aaaabbb, \ldots \}$

$L (r) = \{ a^{2n} . b^{2m+1} \mid n \geq 0, m \geq 0 \}$

It denotes set of all strings with an even number of a's followed by odd number of b's

Examples On Regular Expression

Ex. 4 Write the regular expression for the language accepting all combinations of a's over the set $\Sigma = \{ a \}$

Solution:

$\Sigma = \{ a \}$

Possible set of language $L = \{ \epsilon, a, aa, aaa, aaaa, \ldots \}$

$R = a^*$

This is known as Kleen Closure of a

Examples On Regular Expression

Ex. 5 Write the regular expression for the language accepting all combinations of a's except the null string over the set $\Sigma = \{ a \}$

Solution:

$\Sigma = \{ a \}$

Possible set of language L = { a, aa, aaa, aaaa, … }

R = a+

This is known as Positive Closure of a

Examples On Regular Expression

Ex. 6 Design the regular expression for the language containing all the strings having any number of a's and b's over the set $\Sigma = \{ a, b \}$

Solution:

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

Possible set of language $L = \{ \epsilon, a, b, aa, ab, ba, bb, aaa, \dots \}$

$R = (a + b)^*$

Examples On Regular Expression

Ex. 7 Construct the regular expression for the language containing all the strings having any number of a's and b's except the null string over the set $\Sigma = \{ a, b \}$

Solution:

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

Possible set of language $L = \{ a, b, aa, ab, ba, bb, aaa, \ldots \}$

$R = (a + b)^+$

Examples On Regular Expression

Ex. 8 Construct the regular expression for the language containing all the strings which are ended with 00 over the set $\Sigma = \{ 0, 1 \}$

Solution:

$\Sigma = \{ 0, 1 \}$

R. E. = (Any combination of 0's and 1's) . 00

= $(0 + 1)^* 00$

Possible set of language L = { 00, 000, 100, 0100, 1100, … }

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

Examples On Regular Expression

Ex. 9 Write regular expression for the language accepting the string which are starting with 1 and ending with 0 over the set $\Sigma = \{ 0, 1 \}$

Solution:

$\Sigma = \{ 0, 1 \}$

R. E. = 1 . (Any combination of 0's and 1's) . 0

$= 1 . (0 + 1)^* . 0$

Possible set of language $L = \{ 10, 100, 110, 1000, 1010, \dots \}$

$R = 1 (0 + 1)^* 0$

Examples On Regular Expression

Ex. 10 What is the regular expression for the language starting and ending with a and having any combination of b in between over the set $\Sigma = \{ a, b \}$

Solution:

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

R. E. = a . (Any combination of b) . a

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

Possible set of language L = { aa, aba, abba, abbba, … }

$R = ab^*a$

Examples On Regular Expression

Ex. 11 For $\Sigma = \{ 0, 1 \}$ Give regular expression such that

$L(r) = \{ w \in \Sigma^* \mid w \text{ has at least one pair of consecutive zeros} \}$

Solution:

$\Sigma = \{ 0, 1 \}$

R. E. $= (0 + 1)^* . 00 . (0 + 1)^*$

Possible set of language L = { 00, 000, 100, 001, 1001, 0001, … }

$R = (0 + 1)^* \, 00 \, (0 + 1)^*$

$= (0 + 1)^* \, (00)^+ \, (0 + 1)^*$

Examples On Regular Expression

Ex. 12 Describe in simple English language

Given regular expression is $r = (a + ab)^*$

Solution:

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

R. E. $= (a + ab)^*$

Possible set of language $L = \{ \epsilon, a, ab, aba, aab, abab, aaa, \ldots \}$

The language is beginning with zero or any number of a's but not having consecutive b's

Examples On Regular Expression

Ex. 13 Write regular expression to denote the language L over Σ* where

Σ = { a, b, c } in which every string will be such that any number of a's followed by any number of b's followed by any number of c's

Solution:

Σ = { a, b, c }

R. E. = (any number of a's). (any number of b's). (any number of c's)

R. E. = a*b*c*

Possible set of language L = { ϵ, a, ab, abc, aabbc, abbccc, aaa, … }

Examples On Regular Expression

Ex. 14 Write regular expression to denote the language L over Σ^* where

$\Sigma = \{ a, b, c \}$ in which every string will be such that at least one a followed by at least one b followed by at least one c

Solution:

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

R. E. = $a^+b^+c^+$

Possible set of language L = { abc, aabc, abbc, aabbc, abbccc, aabbcc, … }

REGULAR EXPRESSION

MODULE 2

Examples On Regular Expression

Ex. 15 Write regular expression to denote the language L over Σ* where

Σ = { 0, 1 } in which every string will be such that at all strings which begins or end with 00 or 11

Solution:

R = L1 + L2

L1 – The strings which begin with 00 or 11

L2 – The strings which end with 00 or 11

L1 = (00 + 11) . (Any combination of 0's and 1's)

= (00 + 11) . (0 + 1)*

L2 = (Any combination of 0's and 1's) . (00 + 11)

= (0 + 1)* . (00 + 11)

R = (00 + 11) (0 + 1)* + (0 + 1)* (00 + 11)

Examples On Regular Expression

Ex. 16 Write regular expression to denote the language L over Σ^* where

$\Sigma = \{ a, b \}$ such that the third character from right end of the string is always a

Solution:

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

R. E. = (Any number of character a's and b's) . a . (a or b) . (a or b)

$R = (a + b)^* . a . (a + b) . (a + b)$

Possible set of language L = { aaa , aba, abb, aaba, aabb, bababb, … }

Examples On Regular Expression

Ex. 17 Write regular expression to denote the language L over Σ^* where

$\Sigma = \{ a, b \}$ which accepts all the strings with at least two b's

Solution:

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

The two b's can be surrounded with any number of a's or b's in between

R. E. = (Any combination of a's and b's) . b . (Any combination of a's and b's) . b . (Any combination of a's and b's)

$R = (a + b)^* . b . (a + b)^* . b . (a + b)^*$

Possible set of language L = { bb, bba , abb, bbb, aabb, abba, aabbaa, … }

Examples On Regular Expression

Ex. 18 Write regular expression to denote the language L having strings which should have at least one 0 and at least 1 where $\Sigma = \{ 0, 1 \}$

Solution:

R = L1 + L2

L1 – The strings in which 0 precedes 1

L2 – The strings in which 1 precedes 0

$L1 = (0 + 1)^* . 0 . (0 + 1)^* . 1 . (0 + 1)^*$

$L2 = (0 + 1)^* . 1 . (0 + 1)^* . 0 . (0 + 1)^*$

We can write resultant expression as

$R = (0 + 1)^* 0 (0 + 1)^* 1 (0 + 1)^* + (0 + 1)^* 1 (0 + 1)^* 0 (0 + 1)^*$

Examples On Regular Expression

Ex. 19 Construct regular expression which contains even length of string over the set $\Sigma = \{ 0 \}$

Solution:

Given $\Sigma = \{ 0 \}$

The language with even number of 0's can be written as

$L = \{ \epsilon, 00, 0000, 000000, \dots \}$

We cann say that 00 always comes in pair.

$R = (00)^*$

Examples On Regular Expression

Ex. 20 Write regular expression which denotes a language over the set $\Sigma = \{ 1 \}$ having odd length of strings

Solution:

Given $\Sigma = \{ 1 \}$

The language with odd number of 1's can be written as

$L = \{ 1, 111, 11111, \ldots \}$

The regular expression can be written as

$R = 1 . (11)^*$

Examples On Regular Expression

Ex. 21 Find a regular expression for the language

$L = \{ w \in \{ 0, 1 \}^* \mid w$ has no pair of consecutive zeros $\}$

Solution:

The regular expression involves repetition of the strings of the form 1 … 101 … 1

We can represent this as $r1 = (1^*011^*)^*$

If the string can be terminated by 0 then we can modify the expression as

$r1 = (1^*011^*)^* . (0 + \epsilon)$

The expression does not represent the strings with all 1's and strings with all 1's ending with 0

Examples On Regular Expression

Ex. 21 Find a regular expression for the language

$L = \{ w \in \{ 0, 1 \}^* \mid w \text{ has no pair of consecutive zeros} \}$

Solution:

It can be represented as

$r2 = 1^* . (0 + \epsilon)$

We can give resultant regular expression as

$R = r1 + r2$

$R = (1^*011^*)^* . (0 + \epsilon) + 1^* . (0 + \epsilon)$

Alternate Solution: $r = (1 + 01)^* . (0 + \epsilon)$

$r = (1^* (01)^* 1^*)^*. (0 + \epsilon)$

Examples On Regular Expression

Ex. 22 Obtain regular expression such that

$L(R) = \{ w \mid w \in \{0, 1\}^*$ and w has at least single occurrence of three consecutive zeros $\}$

Solution:

$\Sigma = \{ 0 , 1 \}$

R = (Any combination of 0's and 1's) . 000 . (Any combination of 0's and 1's)

$R = (0 + 1)^* . 000 . (0 + 1)^*$

Possible set of language L = { 000, 1000, 0001, 10001, 00000011, … }

Examples On Regular Expression

Ex. 23 Obtain regular expression such that

The set of all strings over $\Sigma = \{ 0 , 1 \}$ without length two

Solution:

$\Sigma = \{ 0 , 1 \}$

The regular expression can be splitted into three parts as

$R = r1 + r2 + r3$

Where r1 – Strings with length three or more

r2 – String with length one

r3 – String with length zero

Examples On Regular Expression

Ex. 23 Obtain regular expression such that

The set of all strings over $\Sigma = \{ 0 , 1 \}$ without length two

Solution:

The expressions can be written as

$R1 = (0 + 1)(0 + 1)(0 + 1)^+$

$R2 = (0 + 1)$

$R3 = \epsilon$

Final Expression:

$R = (0 + 1)(0 + 1)(0 + 1)^+ + (0 + 1) + \epsilon$

Examples On Regular Expression

Ex. 24 Obtain regular expression over Σ = { 0 , 1 } such that

the set of all strings with number of zeros are odd

Solution:

We need at least one zero to make count of zeros to odd. This zero can be surrounded by any number of 1's. Therefore we can write it as R1 = 1*.0.1*

Now this expression can be followed with pair of zeros. Those pair of zeros can be surrounded by any number of 1's. It can be written as R2 = (1*.0.1*.0.1*).

This R2 can be repeated any number of times.

R3 = (1*.0.1*. 0.1*.)*

Examples On Regular Expression

Ex. 24 Obtain regular expression over $\Sigma = \{ 0 , 1 \}$ such that

the set of all strings with number of zeros are odd

Solution (Contd.) :

When we concatenate R1 and R2 then no of 1's will remain odd

$R = R1.R2$

$R = 1^*.0.1^*.(1^*.0.1^*.\ 0.1^*)^*$

This expression can be further reduced as $R = 1^*.0.1^*.(0.1^*0.1^*)^*$

Final Answer: $R = 1^*.0.1^*.(0.1^*.0.1^*)^*$

Examples On Regular Expression

Ex. 25 Obtain regular expression over $\Sigma = \{ 0 , 1 \}$ such that

the set of all strings containing both 11 and 010 as substring

Solution:

$\Sigma = \{ 0 , 1 \}$

We split regular expression in two parts:

R1 = The strings in which 11 precedes 010

$= (0 + 1)^* . 11 . (0 + 1)^* . 010 . (0 + 1)^*$

R2 = The strings in which 010 precedes 11

$= (0 + 1)^* . 010 . (0 + 1)^* . 11 . (0 + 1)^*$

Final Expression: R = R1 + R2

$R = (0 + 1)^* 11 (0 + 1)^* 010 (0 + 1)^* + (0 + 1)^* 010 (0 + 1)^* 11 (0 + 1)^*$

Examples On Regular Expression

Ex. 26 Obtain regular expression over $\Sigma = \{ 0 , 1 \}$ such that

the set of all strings begin or end with 00 or 11

Solution:

$\Sigma = \{ 0 , 1 \}$

Consider six different cases:

R1 = Strings beginning with 00 and ending with 11

R2 = Strings beginning with 11 and ending with 00

R3 = Strings beginning with 00

R4 = Strings beginning with 11

R5 = Strings ending with 00

R6 = Strings ending with 11

Examples On Regular Expression

Ex. 26 Obtain regular expression over $\Sigma = \{ 0 , 1 \}$ such that

the set of all strings begin or end with 00 or 11

Solution:

Regular expression can be written as

$R1 = 00 (0+1)^*11$

$R2 = 11 (0+1)^*00$

$R3 = 00 (0+1)^*$

$R4 = 11 (0+1)^*$

$R5 = (0+1)^*00$

$R6 = (0+1)^*11$

Examples On Regular Expression

Ex. 26 Obtain regular expression over $\Sigma = \{ 0 , 1 \}$ such that

the set of all strings begin or end with 00 or 11

Solution:

Regular expression can be written as

$R = R1 + R2 + R3 + R4 + R5 + R6$

$$R = 00\,(0+1)^*11 + 11(0+1)^*00 + 00\,(0+1)^* + 11\,(0+1)^* + (0+1)^*00 + (0+1)^*11$$