4. Regular Languages and Regular Grammars

→ Introduction:

- The languages accepted by finite automata are described by regular expressions. So to prove a language is accepted by finite automata it is sufficient to prove the regular expression of that language is accepted by finite automata.

- Regular expressions are used in many programming languages and language tools like lex, vi editor, PHP, PERL etc. They are used as powerful tools in search engines.

→ Regular Expressions:

- Regular Expressions are shorthand notations to describe a language.
- A Regular expression is a notation to represent certain sets of strings in an algebric fashion.
- Regular Expression describes the language accepted by FA.

Definition

*) Any terminal symbol/element of Σ is R.E.

Eg: ϕ, ϵ, a in Σ

ϕ is a regular expression & denotes the empty set.

ϵ is a regular expression and denotes the set $\{\epsilon\}$.

a is a regular expression and denotes the set $\{a\}$.

Scanned by CamScanner

*) Union of two regular expressions R1 and R2 is regular expression R ($R = R1 + R2$).

Eg: Let "a" be regular expression R1

"b" be regular expression R2

(a+b) is also a regular expression R having the elements {a, b}.

*) Concatenation of two regular expressions R1 and R2 written as R1.R2 is also regular expression R ($R = R1 \cdot R2$)

Eg: Let "a" be regular expression R1

"b" be regular expression R2

(a.b) also a regular expression R having the element {ab}.

*) Iteration (closure) of a regular expression R written as R^* is also a regular expression.

Eg: Let "a" be regular expression

then ϵ, a, aa, aaa .... are also regular expression.

If L is a language represented by the regular expression R then Kleen closure of L is denoted as L^* and given as

$$L^* = \bigcup_{i=0}^{\infty} L^i$$

The positive closure of L, denoted L^+, is the set

$$L^+ = \bigcup_{i=1}^{\infty} L^i$$

*) If R is a r.e. then R^* is also a regular expression.

*) R.E over Σ is precisely those obtained recursively by the applicatio of the above rules once or several times.

Scanned by CamScanner

- The basic operations performed on regular expressions are

  Union

  Concatenation

  Kleene's closure.

Among the three closure has highest priority (precedence) next highest is for concatenation and least is for union.

→ **REGULAR SET :**

✓ Any set represented by a regular expression is called a regular set.

Eg: $\Sigma = \{0,1\}$ then 0 denotes a set $\{0\}$.

✓ If a, b are the elements of Σ then regular expressions

- ϕ is a regular expression which denotes an emptyset.
- ϵ is a regular expression which denotes a set $\{\epsilon\}$.
- a denotes the set $\{a\}$
- a+b denotes the set $\{a,b\}$
- ab denotes the set $\{ab\}$
- a^* denotes the set $\{\epsilon, a, aa, aaa \dots\}$
- $(a+b)^*$ denotes the set $\{\epsilon, a, b, aa, ab, ba, bb \dots\}$

✓ Representing the regular set by regular expression

Regular set	Regular expression
$\{101\}$	101
$\{\epsilon, a\}$	$\epsilon+a$
$\{\epsilon, a, aa, ab, ba, bb \dots\}$	$(a+b)^*$
$\{ab, ba\}$	$(ab+ba)$.

Scanned by CamScanner

→ Describing the following sets by regular expression.

1) All strings of 0's and 1's.

$(0+1)^*$

2) Set of all strings of 0's and 1's ending with 00

$(0+1)^*00$

3) Set of all strings ending with 011

$(0+1)^*011.$

4) Set of all strings of 0's and 1's beginning with 00.

$00(0+1)^*$

5) Set of all strings of 0's and 1's beginning with 0 and ending with 1.

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

6) Set of all strings of 0's and 1's beginning with 1 and ending with 00.

$1(0+1)^*00.$

7) Strings of 0's and 1's with atleast two consecutive 0's

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

8) Regular expression that denotes all the words with atleast two 0's.

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

9) Regular expression that denotes exactly two 0's

$1^*01^*01^*.$

10) Set of all strings having even no. of 1's

$(11)^*$

11) Set of all strings having odd number of 1's

$1(11)^*$ or $(11)^*1.$

12) Regular expression to denote any number of 0's followed by any number of 1's followed by any number of 2's

$0^*1^*2^*.$

Scanned by CamScanner

(13) Regular expression to denote any number of 0's followed by any number of 1's followed by any number of 2's with atleast one of each symbol.

$$00^*11^*22^* = 0^+1^+2^+$$

(14) Regular expression for the set of strings, that consists of alternating 0's and 1's

$$(\epsilon+1)\ (01)^*\ (\epsilon+0)$$

(or).

$$(\epsilon+0)\ (10^*)\ (\epsilon+1)$$

(15) Set of all strings of 0's and 1's whose last two symbols are the same.

$$(0+1)^*\ (00+11).$$

(16) Set of strings in which every 0 is immediately followed by atleast two 1's

To find R.E consider two possibilities

(a) String with only 1's

(b) Every 0 preceded by 11. Ie, 011

Hence, $\underline{\underline{(1+011)^*}}$

(17) All strings of 0's and 1's beginning with 1 or 0 and not having two consecutive 0's.

$$(0+\epsilon)\ (1+10)^*.$$

Scanned by CamScanner