Unit-1 : Basic Foundations

DATE

Review of set Theory

* Set : A set is any well defined un-ordered collection of distinct objects, called as elements of the set. Ex : All vowel alphabets, odd number.

* Cardinality of set : Represents the number of elements within a set.

* Representation of a set :

1) Descriptive Method.
=> Set is specified by a verbal description. Ex: the set of numbers 1, 2, 3 is : S = the set of positive integers less than 4.

2) Tabulation Method.
=> Set is specified by listing all the elements in the set.
Ex : $S = \{1, 2, 3\}$

3) Rule method or Set-builder Method.
=> Set is specified by stating a characteristic property common to all elements in the set.
Ex : $S = \{x : x$ is an integer $1 \leq x \leq 3\}$

* Subset : A set A is a subset of a set B if each element of A is also element of B and is denoted by

* Finite set : A set consisting of finite number of elements. Ex : Set of days in a week.

DATE

* Infinite set : Set consisting of infinite number of elements. Ex: set of all odd numbers.

* Empty set:- (void set) : Set without any element. ϕ or $\{\}$

* Unit set : Set consisting of only one element. is called a unit set or singleton set.
  Ex: $N = \{2\}$ is a unit set with single element 2

* Universal set : A set of all entities in the current context.

* Power set : Set of all possible subsets from any set. A set with n elements has 2^n subsets.

* Set operations:-

  1) Union : Union of two sets A & B, denoted by $A \cup B$ is the set of only those elements which belongs to either A or B or both A and B.
  $$A \cup B = \{ x \mid x \in A \text{ or } x \in B \}$$

  2) Intersection:- Intersection of two sets A & B, denoted by $A \cap B$ is the set of only those elemets which belongs to both A and B.
  $$A \cap B = \{ x \mid x \in A \text{ and } x \in B \}$$

  3) Complement: If A is a subset of universal set U. Then the complement of A is the set of all those elements U which do not belong to A. denoted by $\bar{A}$.
  $$\bar{A} = \{ x \mid x \in U \text{ and } x \notin A \}$$

  4) Difference:- Diff. of 2 sets A & B denoted by A-B is a set of all those elemens which belongs A but not B.
  $$A - B = \{ x \mid x \in A \text{ and } x \notin B \}$$

DATE

Review of Logic

Logic is the study of truth and the form of valid inference. A valid inference is one where there is a specific relation of logical support between the assumptions of the inference & its conclusion.

* Propositional logic:-

Proposition is a declarative sentence that declares a fact. It is either true or false but not both.
Propositional logic is a branch of mathematical logic which studies the logical relationships between propositions.

Ex: $2+2=5$ (False) is a proposition.
"Open the door" is not a proposition.

* Predicate logic:

Any declarative statements involving mathematical variables often found in mathematical assertion and in computer programs, which are neither true nor false when the values of variables are not specified is called predicate.
The logic involving predicates is called predicate logic.

Ex: $x > 4$? Is this a proposition? No, until the value of x is specified.

DATE

* Quantifiers.

Quantifiers, are the tools that change the propositional function into a proposition. These are the words that refers to the quantities such as "Some" or "all" and indicates how frequenty a certain statement is true. Construction of propositions from the predicates using quantifiers is quantification.

→ Universal quantifier: The phrase "for all" denoted by $\forall$, is called universal quantifier.

Ex: All students of such college are hardworking
Let, $S(x)$: x students of such college
$H(x)$: x students are hardworking

$$\forall x\,(S(x) \rightarrow H(x))$$

→ Existential quantifier: The phrase "there exist", denoted by $\exists$, is called existential quantifie

Ex: Only few students of such college are hardworking

$$\exists x\,(S(x) \wedge H(x)).$$

DATE

Review of Functions.

* Function.

If A and B are two sets, a function f from A to B is a rule that assigns each elements of A to the elements of B.

For a function f from A to B, the elements of A are domain, elements of B are co-domain, and the elements of B to which elements of A are assigned is range.

* Relations:

A binary relation r on two sets A and B is a subset of $A \times B$. Ex: $A = \{1, 2, 3\}$, $B = \{a, b, c\}$, then $\{(1, a), (2, b), (3, c)\}$ is a binary relation on two set.

* Equivalence relation

A binary relation r is an equivalence relation if r satisfies following:

i) r is reflexive, ie. $\forall x$, $(x, x) \in R$.

ii) r is symmetric, ie. $\forall x \,\&\, y$, $(x, y) \in R$ implies $(y, x) \in R$

iii) r is transitive, $\forall x, y, z$. $(x, y) \,\&\, (y, z) \in r$ then $(x, z) \in r$

Review of Proofs

1) Direct proof: To prove $p \rightarrow q$, we prove $p \rightarrow q$

2) Indirect →
   - Contraposition: To prove $p \rightarrow q$, we prove $\neg q \rightarrow \neg p$
   - Contradiction: To prove $p \rightarrow q$, we prove $\neg q \rightarrow \neg p$

3) Induction proof

(see in discrete structure notes)

PAGE

classmate

DATE

* Theory of Computation (TOC):-

It is a study of power and limits of computing having three interacting components:- Automata Theory, Computability Theory and Complexity Theory.

Computability Theory
- What can be computed?
- Are there problems that no program can solve?

Complexity Theory
- What can be computed efficiently?
- Are there problems that no program can solve in a limited amount of time or space?

Automata Theory.
- Study of abstract machine & their properties, providing a mathematical notation of " Computer"
- Automata are abstract mathematical models of machines that perform computations on an input by moving through a series of states or configurations. If the computation of an automata reaches an accepting configuration it accepts that input.

Study of Automata
- For software designing & checking behaviour of digital circuits
- For designing software for checking large body of text as a collection of web pages, patterns etc. like pattern recognizer
- Designing "lexical analyzer" of a compiler, that breaks input text into logical units called "tokens"

Abstract Model

An abstract model of computer system (considered either as hardware or software) is constructed to allow a detailed & precise analysis of how the computer system works.

→ Such a model usually consists of input, output & operations that can be performed and so can be thought of as a processor. eg: an abstract machine that models a banking system can have operations like "deposit", "withdraw", "transfer", etc.

* Basic Concepts of Automata Theory :

- **Symbol:** It is the basic building block of TOC.
  Ex: 0, 1, a, b, ℓ
- **Alphabet:** It is a finite, non-empty set of symbols.
  It is denoted by Σ (sigma).
  Ex: $\Sigma = \{0, 1\}$ → binary alph. $\Sigma = \{+, -, *\}$ → special symbols
  $\Sigma = \{a, b, c\}$
- **Strings:** String is a finite sequence of symbols taken from some alphabet. Ex: 0110 is a string from {0,1}.
  Ex: Given, $\Sigma = \{0, 1\}$
  Some possible strings over Σ are:- "00", "01", "01000", etc.
  If $\Sigma = \{a, b, c, \ldots, z\}$, then "Computation" is a string from given alphabet Σ.

DATE

Note: An empty string exists in every language. Such an empty string is denoted by ϵ (epsilon)

$|\epsilon| = 0$ ie. length of Empty string is 0

$|"abc"| = 3$ ie. length is 3

Q. Given any alphabet Σ such that $|\Sigma| = n$, $n \geq 1$, how many strings of length m, $m \geq 1$ are possible over Σ?

$\Rightarrow$ $\boxed{n^m}$ Ex: $\Sigma = \{a, b\}$

Then, strings of length 2 possible over Σ ie

$n=2, m=2$ are $\{aa, ab, ba, bb\}$. ie. 4. ie. $2^2 = 4$

- Length of String: The length of string w, denoted by $|w|$ is the number of symbols in w. denoted by $|w|$.

- Empty String: It is a string with zero symbols. Denoted by 'ϵ' (Epsilon). The length of an Empty String is zero. $|\epsilon| = 0$.

- Power of Alphabet: The set of all strings of certain length k from an alphabet is the kth power of that alphabet ie. $\Sigma^k = \{w \mid |w| = k\}$.

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

$\Sigma^0 = \{\epsilon\}$ ie. set of all strings with length 0.

$\Sigma^1 = \{0, 1\}$ ie. set of all strings with length 2

$\Sigma^2 = \{00, 01, 10, 11\}$ ie. " " " " lengths

$\Sigma^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$

DATE

- Kleene Closure (Σ^*)

The set of all the strings of any length over an alphabet Σ is called Kleene closure of Σ and it is denoted by Σ^*. Thus, Kleene Closure is the set of all the strings over alphabet Σ with length 0 or more. It is also called mother of all languages & it has infinite length.

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3$ ----

Eg: $A = \{0\}$ then $A^* = \{0^n$ $n = 0, 1, 2, --\}$

- Positive closure (Σ^+)

The set of all strings over an alphabet Σ except the empty string is called positive closure & is denoted by Σ^+.

$\therefore \Sigma^+ = \Sigma^* - \Sigma^0$

- Concatenation of strings:-

If x and y are two strings over an alphabet, the concatenation of x and y is written by xy and consists of the symbols of x followed by those of y.

Ex: $x = aaa$ & $y = bbb$

Then, $xy = aaabbb$, $yx = bbbaaa$

Note: Concatening the empty string 'ε' with another string, the result is just the other string.

- Suffix of a string:

A string (s) is called a suffix of another string 'w' if it is obtained by removing zero or more leading symbols in w. Ex: $w = abcd$

$s = bcd$ is suffix of w

s is proper suffix if $s \neq w$

classmate PAGE

DATE

- Prefix of a String:
  ~~At~~ A String 's' is called a prefix of a string 'w' if it is obtained by removing zero or more trailing symbols of w. Ex: $w = abcd$
  $S = abc$ is a prefix of w
  S is a proper prefix if $s \neq w$

- Substring: A string 'S' is called a substring of string w if it is obtained by removing zero or more leading or trailing symbols in w. It is proper substring if $s \neq w$.

- Language:
  A language L over an alphabet Σ is subset of all the strings that can be formed out of Σ. ie. A language is a subset of kleen closure over an alphabet. ie; $L \subseteq \Sigma^*$.
  A language can be finite or infinite.

  Ex: $\Sigma = \{0, 1\}$
  L_1: set of all strings over Σ such that strings have length 2.
  $\therefore L_1 = \{00, 01, 10, 11\}$. It is a finite language
  L_2: Set of all strings over Σ starting with 0.
  $\therefore L_2 = \{0, 01, 00, 001, \ldots\ldots\}$
  It is an infinite language.

DATE

- Empty language:

A language is empty if it does not have any strings within it. ϕ is an empty language. It does not contain any string.

- Membership in a Language.

Membership in a language defines association of some string with the language. A membership problem is the question of deciding whether a given string is a member of some particular language or not. ie. whether the string belongs to the given language or not.

Ex: Given, L = {DBMS, TOC, GRAPHICS} then for w = "DBMS", the membership of w is true in L.
For, w = "HELLO", the membership is false in L.

(final ex.)

Questions asked from this Chapter

Q. Define the term alphabet, prefix and suffix of string concatenation & kleen closure with example. (2078 - 5 marks)

Q. Give the regular expressions for the following language over alphabet {a, b} (2078 - 5 marks)

a. Set of all strings with substring bab or abb
⇒ {

b. Set of all strings whose 3rd symbol is 'a' and 5th symbol is 'b'.
⇒

Q. Define the term : alphabet, substring/prefix/suffix of a string with example. (2076 - 5 marks)

Q. Define the term: Kleene closure, union, concatenation and power of an alphabet with example. (2072 - 5 marks)