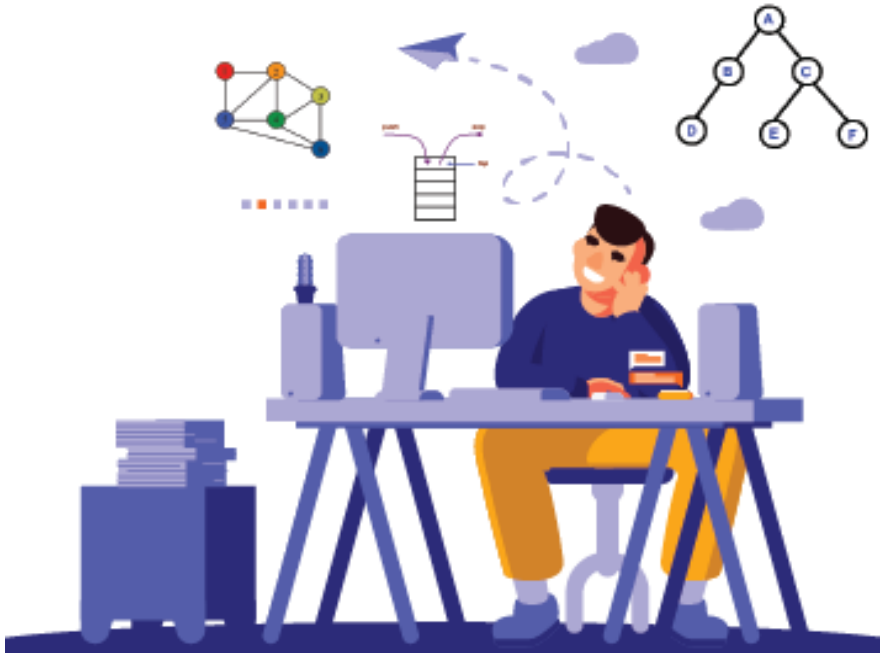
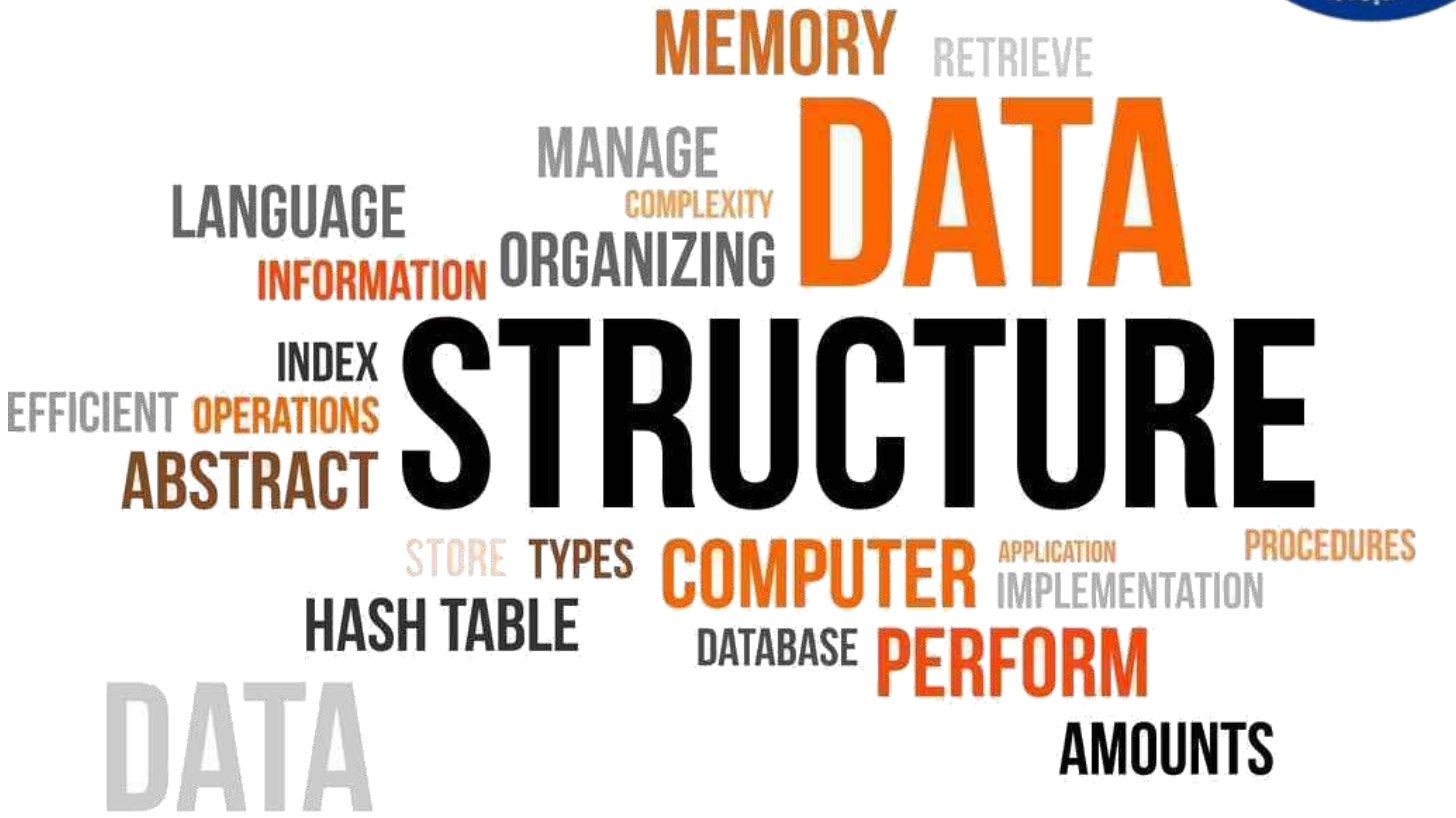


Data Structures

Course code: IT623



Dr. Rahul Mishra
Assistant Professor
DA-IICT, Gandhinagar

Lectures 9

String Processing…

String Constants:

Constants by placing the string in either single or double quotation marks.

'The End' and "To be or not to be"

↳ Mostly used by the programming languages.

String Variables:

i) Static character : whose length is defined before the program is executed and cannot change throughout the program.

ii) Semi static character : whose length may vary during the execution of the program as long as the length does not exceed a maximum value determined by the program

iii) Dynamic :

↳ Whose length can change during the execution of the program.

* These ~~characteristics~~ categories correspond, respectively, to the ways the strings are stored in the memory of the computer.

STRING OPERATIONS

> Group of consecutive elements in string (such as words, phrases and sentence), called <u>substrings</u>, may be units into themselves.

> The basic units of access in a string are usually these substrings, not individual character.

(TO) BE OR NOT TO BE

↳ substring have their own meaning

Array of Integer → 6 7 14 19 . . . are individually have meaning

* Thus, different string operations are developed which are not normally used with other kinds of arrays.

* Note: Unless otherwise stated or implied, we assume our character-type variables are <u>dynamic and have variable length.</u>

1) Substring:

Access to the substring from a given string requires three piece of information.

(a) - the name of the string or the string itself.

(b) - the position of the first character of the substring in the given string

(c) - the length of substring or the position of the last character of the string.

Simply

SUBSTRIN (string, initial, length)

String → S

initial → K position

Length → L

Example:

SUBSTRING ('TO BE OR NOT TO BE', 4, 7) = 'BE OR N'

SUBSTRING ('DSA MNC', 2, 5) = ? _____

SUBSTRING ('THE END', 4, 4) = 'END'.

2> Indexing:

Indexing, also called pattern matching, refers to finding the position where a string pattern P first appears in a given string text T.

INDEX (text, pattern)

Notably: If the pattern P does not appear in the text T then INDEX is assigned the value 0. The arguments "text" and "pattern" can be either string constant or string variable.

Example:

Let T = 'HIS FATHER IS THE PROFESSOR'

INDEX(T, 'THE') = 7

INDEX(T, 'THEN') = 0

INDEX(T, '␣THE␣') = 14

T = 'BE CONSISTENT BE EXTRAORDINARY'

INDEX(T, 'ENT') = ? INDEX(T, 'ORD') = ?

INDEX(T, 'BE␣') = ?

INDEX(T, '␣BE␣') = ?

3) Concatenation:

Let S_1 and S_2 be strings, denoted by $S_1 // S_2$

Discussed earlier

* $S_1 // S_2$ is the string consisting of S_1 followed by the characters of S_2.

Example — (a) $S_1 // S_2$ = 'DSA' // 'MNC' = 'DSAMNC'.

(b) $S_1 // ' \square ' // S_2$ = 'DSA MNC'.

4) Length:

The number of characters in a string is called its length.

LENGTH (string)

LENGTH ('DSAMNC') = 6

Another important function:

TRIM ('DSAMNC', 4) = 'DSAM'

10 minutes Problem :

For string (S) and (T)

S = 'WE THE PEOPLE' and T = 'OF THE UNITED INDIA'

(a) Find the Length of S and T ?

(b) Find (a) SUBSTRING (S, 4, 8) and (b) SUBSTING (T, 10, 5) ?

(c) Find (i) INDEX (S, 'P')

(ii) INDEX (S, 'E')

(iii) INDEX (S, 'THE')

(iv) INDEX (T, 'THE')

(v) INDEX (T, 'THEN')

(vi) INDEX (T, 'TE')