COM2067/ COM267

Chapter 2: Introduction to Data Structures and Algorithms

Data Structures Using C, Second Edition
Reema Thareja

- BASIC TERMINOLOGY
- CLASSIFICATION OF DATA STRUCTURES
- OPERATIONS ON DATA STRUCTURES
- ABSTRACT DATA TYPE
- ALGORITHMS
- DIFFERENT APPROACHES TO DESIGNING AN ALGORITHM
- CONTROL STRUCTURES USED IN ALGORITHMS
- TIME AND SPACE COMPLEXITY

BASIC TERMINOLOGY

- We know how to write, debug, and run simple programs in C language.
- Our aim has been to design good programs, where a good program is defined as a program that
  - runs correctly
  - is easy to read and understand
  - is easy to debug
  - is easy to modify

BASIC TERMINOLOGY

- A program should give correct results, but along with that it should also run efficiently.
- A program is said to be efficient when it executes in minimum time and with minimum memory space.
- In order to write efficient programs we need to apply certain data management concepts.
- Data structure is a crucial part of data management.
- A data structure is basically a group of data elements that are put together under one name, and which defines a particular way of storing and organizing data in a computer so that it can be used efficiently.

BASIC TERMINOLOGY

- Data structures are used in almost every program or software system.
- Some common examples of data structures are arrays, linked lists, queues, stacks, binary trees, and hash tables.
- Representing information is fundamental to computer science.
- The primary goal of a program or software is not to perform calculations or operations but to store and retrieve information as fast as possible.

BASIC TERMINOLOGY

- The application of an appropriate data structure provides the most efficient solution.
- A solution is said to be efficient if it solves the problem within the required resource constraints like the total space available to store the data and the time allowed to perform each subtask.
- And the best solution is the one that requires fewer resources than known alternatives.
- Moreover, the cost of a solution is the amount of resources it consumes.
- The cost of a solution is basically measured in terms of one key resource such as time, with the implied assumption that the solution meets the other resource constraints.

BASIC TERMINOLOGY

- Today computer programmers do not write programs just to solve a problem but to write an efficient program.
- For this, they first analyze the problem to determine the performance goals that must be achieved and then think of the most appropriate data structure for that job.
- However, program designers with a poor understanding of data structure concepts ignore this analysis step and apply a data structure with which they can work comfortably.
- The applied data structure may not be appropriate for the problem at hand and therefore may result in poor performance (like slow speed of operations).

BASIC TERMINOLOGY

- If a program meets its performance goals with a data structure that is simple to use, then it makes no sense to apply another complex data structure just to exhibit the programmer's skill.
- When selecting a data structure to solve a problem, the following steps must be performed.
  - Analysis of the problem to determine the basic operations that must be supported. For example, basic operation may include inserting/deleting/searching a data item from the data structure.
  - Quantify the resource constraints for each operation.
  - Select the data structure that best meets these requirements.
- This three-step approach to select an appropriate data structure for the problem at hand supports a data-centred view of the design process.

BASIC TERMINOLOGY

- In the approach, the first concern is the data and the operations that are to be performed on them.
- The second concern is the representation of the data, and the final concern is the implementation of that representation.
- There are different types of data structures that the C language supports.
- While one type of data structure may permit adding of new data items only at the beginning, the other may allow it to be added at any position.
- While one data structure may allow accessing data items sequentially, the other may allow random access of data.
- So, selection of an appropriate data structure for the problem is a crucial decision and may have a major impact on the performance of the program.

BASIC TERMINOLOGY

Elementary Data Structure Organization

- Data structures are building blocks of a program.
- A program built using improper data structures may not work as expected.
- So as a programmer it is mandatory to choose most appropriate data structures for a program.
- The term data means a value or set of values.
- It specifies either the value of a variable or a constant (e.g., marks of students, name of an employee, address of a customer, value of pi, etc.).