Lecture 9: Introduction to Data Structures

BT 3051 – Data Structures and Algorithms for Biology

Karthik Raman
Department of Biotechnology
Indian Institute of Technology Madras

August 26, 2014

What is a data structure?

- Data structures organise (*structure*) data on a computer, for efficient use by algorithms, e.g. Array
- Abstract data types (ADTs) are theoretical models of data structures defining both the **type of data** and the **operations that can be performed** on the data, e.g. Set
- Data structures are *physical* implementations of ADTs on a computer

What is a data structure?

- We already know the basic data types: integers, characters, floats, Booleans etc.
- Abstract data types (ADTs) organise and structure collections of such data types
- Many data structures can implement the same ADT
- Data structures organise data efficiently — so that we can find, update, add and delete parts of it efficiently …

Philosophy of Data Structures

- Every data structure has costs and benefits
- Rarely is one data structure better than another in all situations
- Data structures require:
  - space for each data item it stores (data + overheads)
  - time to perform each basic operation,
  - programming effort (Some data structures/algorithms can be very complicated!)
- Each problem has constraints on available space and time
- Careful analysis of problem characteristics $\longrightarrow$ best data structure for the task

Common Goals of Data Structures [and Algorithms]

- Correctness
- Efficiency
- Robustness
- Reusability
- Adaptability

Selecting a Data Structure

- Analyse problem to determine resource constraints a solution must meet
- What are the basic operations that must be supported?
  - What are the resource constraints?
- Which data structure best meets these requirements?
  - Typically we want the “simplest” data struture that will meet requirements

Selecting a Data Structure

Important Questions

- Are all data inserted into the data structure at the beginning, or are insertions interspersed with other operations?
  - i.e. are the data static or dynamic?
- Can data be deleted?
  - This may often demand a more complex representation
- Are all data processed in some well-defined order, or is random access allowed?