Module – 1: Introduction to Data structures and Algorithms

Dr. P. Gayathri
Associate Professor
SCOPE, VIT University

Overview and importance of algorithms and data structures

- Data structures
  - Definition
  - Importance of DS – Easy and fast access
  - Types
  - Examples

Algorithm

- An Algorithm is a set of rules for carrying out calculation either by hand or on a machine
- **An Algorithm** is a well defined computational procedure that takes input and produces output
- An Algorithm is a finite sequence of instructions or steps to achieve some particular output

Algorithm

- Any Algorithm must satisfy the following criteria
  - **Input:** It generally requires finite no. of inputs
  - **Output:** It must produce at least one output
  - **Uniqueness:** Each instruction should be clear and unambiguous
  - **Finiteness:** It must terminate after a finite no. of steps

Algorithm

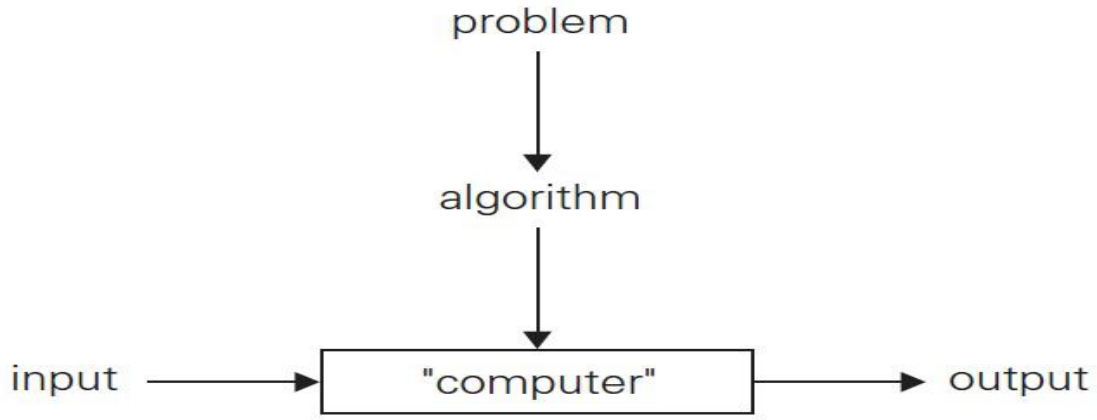
- Analysis issues of an algorithm
  - What data structures to use !!!
  - Is it correct ???
  - How efficient is it ???
  - Is there any other efficient algorithm ???

Algorithm

- An algorithm can be viewed as tool for solving a well-specified **computational problem**
- An algorithm is said to be **correct** if, for every input instance, it
  - halts with the correct output
- An incorrect algorithm
  - may not halt at all for some input instances, or
  - may halt with a wrong output

Algorithm

- A computer program is written in a programming language whereas an algorithm is written in pseudo code.



Stages of Algorithm Development

- Describe the problem – define the problem stmt.
- Identify a suitable technique
- Design an algorithm
- Check the correctness of the algorithm – works fine for all selected inputs.
- Measure the time complexity of the algorithm (see later)

Ex: Swapping of the values of two variables

- Problem Description:
  - Given two variables, a and b, exchange the values assigned to them.
- Suitable technique: Usage of a temporary variable
- Design an algorithm
- Check the correctness by sample input and output
- Measure the time complexity of the algorithm

Ex: Swapping of the values of two variables

Sample Input: Before exchange

a	b
135	245

Desired Output: After exchange

a	b
245	135

- Applications
  - Sorting Algorithms

Algorithm Design

Algorithm: EXCHANGE VALUES OF TWO VARIABLES

1. Save the original value of a in t.
2. Assign to a the original value of b.
3. Assign to b the original value of a that is stored in t.

Algorithm (pseudo code format): EXCHANGE VALUES OF TWO VARIABLES

1. t ← a
2. a ← b
3. b ← t

Proof of Correctness of the algorithm

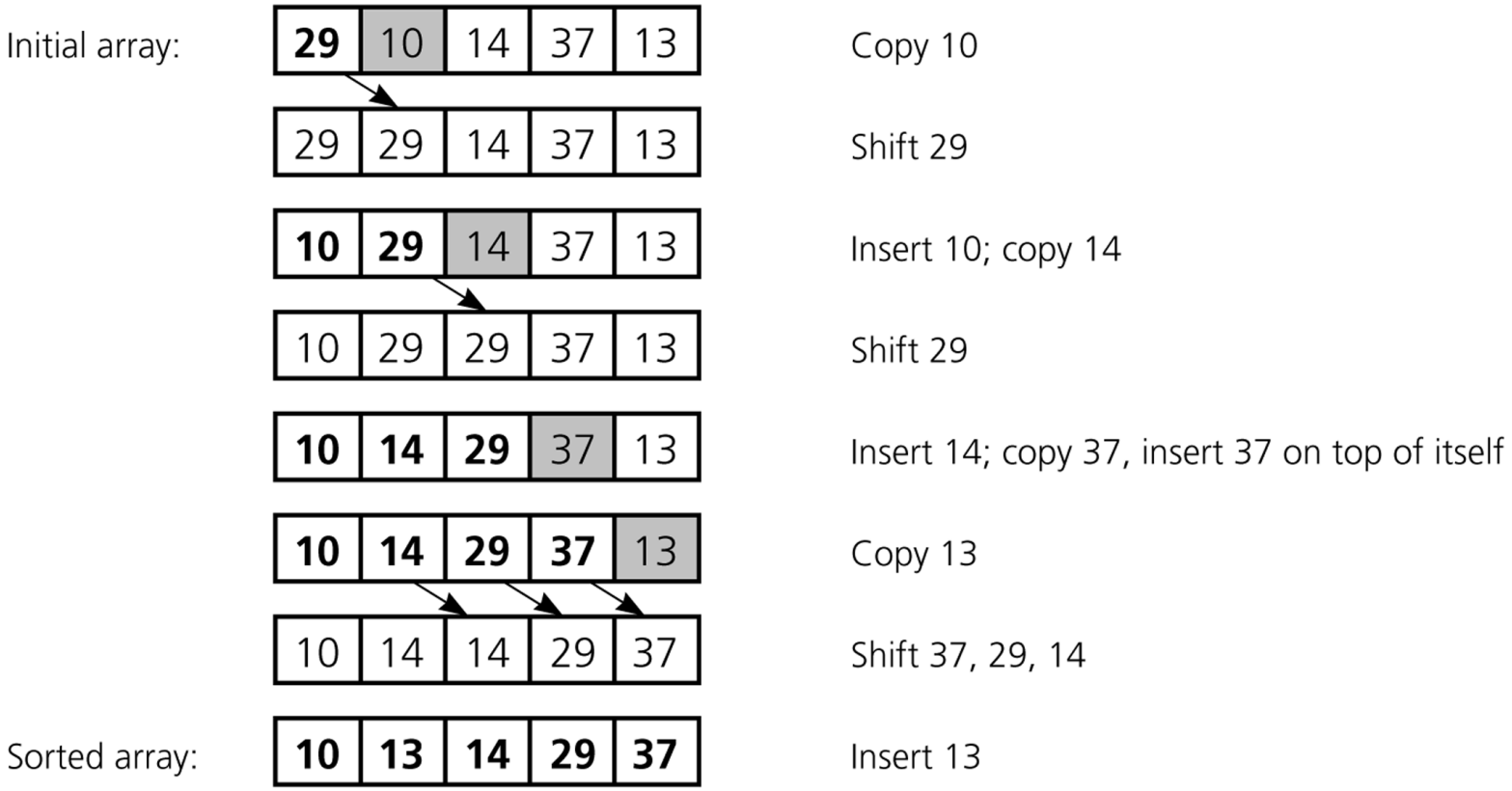
- Once an algorithm is designed, we need to prove its correctness i.e. the algorithm yields the desired output for the specific input in a finite amount of time.
- Proof of correctness depends on the algorithm.
- For some algorithms, the proof of correctness is easy and for some algorithms, it is complex
- It depends on the type of instructions in the algorithm
- In order to show that an algorithm is incorrect, one instance of input is sufficient for which the algorithm fails to give the desired output
- For the swapping of the values of two variables, proof of correctness is straightforward i.e., we can check the pre-condition and post-condition of the values of the variables

Ex. Sorting problem

- Problem Description:
  - Given the sequence of ‘n’ nos a1, a2,…an.
  - Reorder the input sequence such that

    $a1 \leq a2 \leq a3 \leq \ldots. \leq an.$
- Suitable technique: Insertion sort
  - Efficient algorithm for sorting small no. of elements
  - Works the way many people sort hand of playing cards

Sample Input and Output



Ex. Sorting problem

- Algorithm Design

```
InsertionSort (A[1..n])
{
   for j = 2 to n
   {
      key = A[j]
      i = j - 1;
      while (i > 0) and (A[i] > key)
      {
         A[i+1] = A[i]
         i = i - 1
      }
      A[i+1] = key
   }
}
```

Proof of Correctness

- We saw how this algorithm works for the given set of elements
- j indicates current card in hand. A[1 … j-1] – sorted cards in hand. A[j+1 …. n] – pile of cards still on the table
- The elements A[1 … j-1] are the elements originally in positions 1 through j-1, but now in sorted order.
- This property of sub-array A[1 … j-1] is loop invariant.
- At the beginning of each iteration of for loop, the sub-array A[1 … j-1] consists of elements originally in A[1 … j-1], but in sorted order.
- **Loops in an algorithm/program can be proven correct using mathematical induction. In general it involves something called "loop invariant"**

Loop Invariants

- Initialization
  - Check whether loop invariant holds before the first iteration of the for loop (i.e) when j = 2
  - The sub-array A[1 … j-1] has only one element A[1].
  - Sub-array with only one element is by default sorted. So property gets satisfied.

Loop Invariants

- Maintenance
  - Showing that each iteration maintains the loop invariant property
  - It is clear from our example that the property is maintained

Loop Invariants

- Termination
  - Examine whether the loop invariant is maintained when the loop terminates
  - Loop terminates when j = n+1.
  - In this case, sub-array is A[1 … n], which is in sorted order.
  - So property gets satisfied.

Loop Invariants

- Hence we can say that the algorithm is correct.

General Meaning of Loop invariants

- means checking the status of the control variables in the loop
- 3 things related to loop invariant
  - Initialization: Check whether the invariants used in the loop are properly initialized such that the condition(s) are satisfied to enter the loop i.e., it is true prior to the first iteration of the loop
  - Maintenance: Check whether the values of the invariants are updated and progress towards the termination of the loop i.e., if it is true before an iteration of the loop, it remains true before the next iteration
  - Termination: When the loop terminates, the invariant gives a useful property that helps to show that the algorithm is correct

Proof Correctness - Summation of a set of n numbers

Problem: Given a set of n numbers, design an algorithm that adds these numbers and returns the resultant sum.

Algorithm: SUMMATION
Input: array of n numbers
Output: sum of the n numbers
1. sum← 0
2. k ← 0
3. While(k< n)
4. do sum ← sum + a[k]
5. k ← k + 1
6. return sum

Proof Correctness - Summation of a set of n numbers

- The loop invariant is the variable k
- Initialization: k = 0 (pre-condition checking) before the iteration and sum = 0 implies that the sum of first zero numbers is 0
- Maintenance: k is incremented by 1 implies that "for a particular k, the value of sum gives the sum of the k numbers
- Termination: When k = n (post-condition checking), the loop terminates and at that time, sum contains the sum of n numbers

Measure Time Complexity

- Loops have impact over running time
- 3 cases
  - Best
  - Worst
  - Average

Example

Algorithm: fun(n)

s=0

For i = 1 to n

    s = s + i * i

Analysis

- Sum of squares of n non-negative integers
- Loop gets executed for n times
- O(n)
- Improvement:
  - (n(n+1)(2n+1))/6