CS6202 - PROGRAMMING & DATA STRUCTURES I

Unit V
Part – A

1. Define sorting

Sorting arranges the numerical and alphabetical data present in a list in a specific order or sequence. There are a number of sorting techniques available. The algorithms can be chosen based on the following factors
   - Size of the data structure
   - Algorithm efficiency
   - Programmer’s knowledge of the technique.

2. Mention the types of sorting
   - Internal sorting
   - External sorting

3. What do you mean by internal and external sorting?

An internal sort is any data sorting process that takes place entirely within the main memory of a computer. This is possible whenever the data to be sorted is small enough to all be held in the main memory.

External sorting is a term for a class of sorting algorithms that can handle massive amounts of data. External sorting is required when the data being sorted do not fit into the main memory of a computing device (usually RAM) and instead they must reside in the slower external memory (usually a hard drive).

4. Define bubble sort

Bubble sort is a simple sorting algorithm that works by repeatedly stepping through the list to be sorted, comparing each pair of adjacent items and swapping them if they are in the wrong order. The pass through the list is repeated until no swaps are needed, which indicates that the list is sorted. The algorithm gets its name from the way smaller elements "bubble" to the top of the list.

5. How the insertion sort is done with the array?

It sorts a list of elements by inserting each successive element in the previously sorted sublist.

Consider an array to be sorted A[1], A[2],…., A[n]
6. **What are the steps for selection sort?**
   - The algorithm divides the input list into two parts: the sublist of items already sorted, which is built up from left to right at the front (left) of the list, and the sublist of items remaining to be sorted that occupy the rest of the list.
   - Initially, the sorted sublist is empty and the unsorted sublist is the entire input list.
   - The algorithm proceeds by finding the smallest (or largest, depending on sorting order) element in the unsorted sublist, exchanging it with the leftmost unsorted element (putting it in sorted order), and moving the sublist boundaries one element to the right.

7. **What is meant by shell sort?**
   *Shell sort*, also known as *Shell sort* or *Shell's method*, is an in-place comparison sort. It can either be seen as a generalization of sorting by exchange (bubble sort) or sorting by insertion (insertion sort). The method starts by sorting elements far apart from each other and progressively reducing the gap between them. Starting with far apart elements can move some out-of-place elements into position faster than a simple nearest neighbor exchange. Donald Shell published the first version of this sort in 1959. The running time of Shell sort is heavily dependent on the gap sequence it uses.

8. **What are the steps in quick sort?**
   The steps are:
   a. Pick an element, called a **pivot**, from the list.
   b. Reorder the list so that all elements with values less than the pivot come before the pivot, while all elements with values greater than the pivot come after it (equal values can go either way). After this partitioning, the pivot is in its final position. This is called the *partition* operation.
   c. **Recursively** apply the above steps to the sub-list of elements with smaller values and separately to the sub-list of elements with greater values.

9. **Define radix sort**
   Radix Sort is a clever and intuitive little sorting algorithm. **Radix sort** is a non-comparative integer sorting algorithm that sorts data with integer keys by grouping keys by the individual digits which share the same significant position and value. Radix Sort puts the elements in order by comparing the **digits of the numbers**.

10. **What are the advantages of insertion sort**
    Advantages
    a. Simplest sorting technique and easy to implement
    b. It performs well in the case of smaller lists.
    c. It leverages the presence of any existing sort pattern in the list
Disadvantages

- Efficiency of O(n) is not well suited for large sized lists
- It requires large number of elements to be shifted

11. Define searching
Searching refers to determining whether an element is present in a given list of elements or not. If the element is present, the search is considered as successful, otherwise it is considered as an unsuccessful search. The choice of a searching technique is based on the following factors
  a. Order of elements in the list i.e., random or sorted
  b. Size of the list

12. Mention the types of searching
The types are

- Linear search
- Binary search

13. What is meant by linear search?
*Linear search* or *sequential search* is a method for finding a particular value in a list that consists of checking every one of its elements, one at a time and in sequence, until the desired one is found.

14. What is binary search?
For binary search, the array should be arranged in ascending or descending order. In each step, the algorithm compares the search key value with the middle element of the array. If the key match, then a matching element has been found and its index, or position, is returned. Otherwise, if the search key is less than the middle element, then the algorithm repeats its action on the sub-array to the left of the middle element or, if the search key is greater, on the sub-array to the right.

15. Define hashing function
A hashing function is a key-to-transformation, which acts upon a given key to compute the relative position of the key in an array.
A simple hash function

\[ \text{HASH}(\text{KEY Value}) = (\text{KEY Value}) \mod (\text{Table size}) \]

16. What is open addressing?
Open addressing is also called closed hashing, which is an alternative to resolve the collisions with linked lists. In this hashing system, if a collision occurs, alternative cells are tired until an empty cell is found.
There are three strategies in open addressing:
- Linear probing
- Quadratic probing
- Double hashing

17. What are the collision resolution methods?
The following are the collision resolution methods
- Separate chaining
- Open addressing
- Multiple hashing

18. Define separate chaining
It is an open hashing technique. A pointer field is added to each record location, when an overflow occurs, this pointer is set to point to overflow blocks making a linked list. In this method, the table can never overflow, since the linked lists are only extended upon the arrival of new keys.

Part – B
1. Explain the sorting algorithms
2. Explain the searching algorithms
3. Explain hashing
4. Explain open addressing
5. Write a C program to sort the elements using bubble sort.
6. Write a C program to perform searching operations using linear and binary search.
7. Explain in detail about separate chaining.