(12)

[This question paper contains 12 printed pages.]

Your Roll No..............

I

Sr. No. of Question Paper : 1057

Unique Paper Code : 2342013501

Name of the Paper : Algorithms and Advanced Data Structures

Name of the Course : B.Sc. (H) Computer Science

Semester : V – DSC (NEP-UGCF-2022)

Duration : 3 Hours   Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.

2. This question paper has **two** sections **A** and **B**.

3. Question **1** in **Section A** is compulsory.

4. Attempt any **4** questions from **Section B**.

5. Parts of a question must be attempted together.

6. **Section A** carries **30** marks and each question in **Section B** carries **15** marks.

P.T.O.

Section A

1. (a) Analyze the performance of the following operations of the List data structure as specified by the Standard Template Library of C++: push_front(e), push_back(e), pop_front( ), and pop_back( ).

Under what circumstances would you prefer a std::list over a std::vector? (3+2)

(b) Explain Kruskal's Algorithm for finding the Minimum Spanning Tree (MST) in an undirected weighted graph.

Discuss how the Union-Find data structure is used by Kruskal's algorithm to efficiently detect cycles during edge addition. (3+2)

(c) Consider the following array:

A = [6, 12, 23, 29, 31, 36, 42, 48, 59, 60]

Compare the performance of Quicksort and Randomized Quicksort for sorting the given array, A. Explain the impact of pivot selection on the partitioning process for both the algorithms and how it influences the overall time complexity.

(3+2)

(d) Given a residual graph, G with a flow, f and v(f) be the value of flow. The function bottleneck(p,f) for an augmenting path P is the minimum residual capacity with respect to the flow f. Argue that the bottleneck(P,f) always greater than zero.

If the v(f') is equal to the v(f) + bottleneck(P,f).

Prove that v(f') is greater than v(f) and the capacity condition holds for f'. (2+3)

(e) Given the sequence of probes generated by quadratic probing for keys, K

$$p(i) = h(K) + (-1)^{i-1}((i + 1)/2)^2$$

for i = 1, 2, ..., TSize – 1 i.e. h(K), h(K)+1, h(K)–1, h(K)+4, h(K)–4,...

What happens if the table size, TSize is not prime?

Can quadratic probing guarantee the use of all positions in the table for a non-prime table size? (2+2)

(f) Given that in a vh-tree, the shortest path to a leaf consists only of vertical links, and the longest path to another leaf may use both vertical and horizontal links alternately, derive the relationship between the longest path, $path_{longest}$ and the shortest path, $path_{shortest}$. Given a vh-tree with height h, derive the minimum number of nodes n_{min} that the tree can have. (2+2)

(g) What is inverted index (or inverted file) with respect to search engines. (2)

Section B

2. (a) Given a text string T="ABABDABACDABABCA BAB" and a pattern string P="ABABCABAB", perform the Knuth-Morris-Pratt (KMP) string matching algorithm to find first occurrence of P in T. Show the steps, including the construction of the failure table.

   Compare the Brute Force and KMP algorithms in terms of time complexity. (4+3)

   (b) Consider the following weighted directed graph with 4 nodes (A, B, C, D) with the following edges and weights :

   A → B (weight 3)

   A → C (weight -2)

   B → D (weight 2)

$C \rightarrow B$ (weight -1)

$D \rightarrow A$ (weight 1)

Use the Bellman-Ford algorithm to find the shortest path from vertex A to D.

Can we find the shortest path in a graph, if it contains a negative cycle? Support your answer with proper arguments. (5+3)

3. (a) Consider a hash table of size 8, with the following hash function:

$$H(k) = (k \bmod 8).$$

Insert the following set of keys : {14, 30, 22, 10, 50, 60, 90} using the given hash function and for collision resolution use linear probing.

Discuss the advantages and disadvantages of linear probing and quadratic probing in terms of collision handling and performance. (4+3)

(b) Does the splitting of a node at the time of insertion always increase the height of the tree?

Construct a B-tree of order 5 which is initially empty. Insert the following keys into the B-tree: 7,15,11,25,12,10,22,30,20,5,27,30,18. After each insertion, show the structure of the B-tree. What is the height of the final tree? (2+6)

4. (a) Given the following set of words: ["cat", "bat", "rat", "at", "batman", "sat", "saturday", "rattle", "cattle", "ate", "battle", "satin"], Construct a standard trie to store these words.

Also show the Compressed trie for the standard trie created above. What is the advantage of using a compressed trie? (3+4)

(b) Consider the array implementation of the Union-Find data structure for set S of size n, where unions keep the name of the larger set. Prove the following :

(i) Find operation takes $O(1)$ time,

(ii) MakeUnionFind(S) takes $O(n)$ time, and

(iii) Sequence of k Union operations takes at most $O(k \log k)$ time.

Explain the role of the "name of the larger set" optimization in the Union-Find data structure. How does this optimization reduce the height of the tree created during the union process? (4+4)

5. (a) Give pseudo-code for the following operations of Sequence Abstract Data Type using an array implementation :

(i) Insert from the back,

(ii) Remove from the end,

(iii) To access an element at any index

What is the running time for each of these operations? (7)

(b) Do we need to restrict the number of levels in a skip list? Justify your answer.

Consider a skip list S with n elements and the insertion policy that restricts the top-level h to a fixed value h=max{10, 2[logn]}, where n is the current number of entries in the skip list.

How does the given policy impact the structure of the skip list during insertion?How does the condition [logn] < [log(n+1)] help allow an insertion to increase the height of the skip list by atmost one level? (2+3+3)

6. (a) Consider the following unsorted array of integers:

A = [39, 12, 57, 23, 89, 11, 4, 70, 18, 42, 35, 8, 61, 26]

Find the i-th smallest element where i = 7 from the array A using the Randomized Select algorithm, where the first element of the resulting array is selected as the pivot.

When do we say a partition is helpful and show that the probability of helpful partition is atleast ½. (4+3)

(b) Consider the following directed graph, G with A, B, C, D, S, T as the nodes representing a flow network :

S $\rightarrow$ A (capacity 10)

S $\rightarrow$ C (capacity 10)

A $\rightarrow$ B (capacity 4)

A $\rightarrow$ C (capacity 2)

A $\rightarrow$ D (capacity 8)

B $\rightarrow$ T (capacity 10)

C $\rightarrow$ D (capacity 9)

D $\rightarrow$ B (capacity 6)

D $\rightarrow$ T (capacity 10)

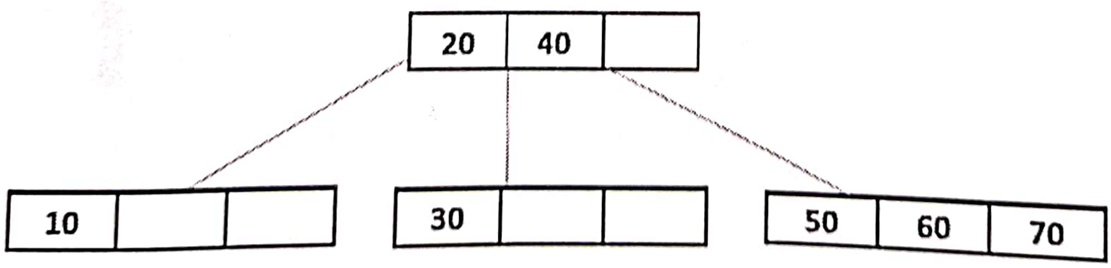
Apply the Ford-Fulkerson algorithm to find the maximum flow from source S to sink T. How does the Ford Fulkerson algorithm ensure that the flow is maximized? (8)

7. (a) Let V be a vector implemented by means of an extendable array A which is extended with [N/4] additional cells. Capacity of V increases from N to N + [N/4]. Initially V is empty and A has size N = 1, Show that the total time to perform a series of n push operations in V is O(n).

If we extend the array from N to N + k where k is a constant, can the sequence of N push operations in V still be done in O(n) time.

(4+3)

(b) Explain the deletion process in B-trees.Consider a B-tree of order 4 (maximum 3 keys per node) with the following initial structure:



Delete the keys 30 and 50 from the B-tree. Show all the necessary steps involved in the deletion process. (3+5)