**Birla Institute of Technology & Science, Pilani**

**Work-Integrated Learning Programmes Division**

**First Semester 2020-2021**

**M.Tech (Data Science and Engineering)**

**End-Semester Test (EC-3 Regular)**

Course No. : DSECF ZG519

|  |
| --- |
| No. of Pages = 3  No. of Questions = 11 |

Course Title : DATA STRUCTURE ALGORITHMS AND DESIGN

Nature of Exam : Closed Book/open Book

Weightage : 40%

Duration : 150 Min

Date of Exam :

Note:

1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

*Answer All the Questions (Only in the pages mentioned against questions. If you need more pages. Continue remaining answers from page 16 onwards, we have provided extra space at each question)*

1. Suppose we’re using quick sort to process data that we’re receiving from a connection in a networked system. We want to cover our system from the possibility of being “sabotaged” by hostile connections — we could receive data that is specifically crafted to cause quick sort to have its worst- case performance and thus make our system consume excessive resources and time (rendering it unable to efficiently respond to other connections).
   1. Assuming that quick sort simply chooses the first element as the pivot (instead of the median of first, last, and middle), what is the arrangement of data that produces the worst-case performance in quick sort? **[2M]**
   2. Suggest a simple strategy (hopefully requiring no more than linear time) to avoid the problem. That is, a strategy to guarantee that quick sort will run in O(n log n) most of the time, regardless of input data, even if this input data is maliciously created. **[2M]**
2. Solve the following recurrence relation : *f(n) = 4f(n/2)+* Hint: replace the with a generic representative  **[4M]**
3. We want to store values between 0 to 9999 in a hash table of size 10.the hash function operates as follows: given a value x, add the four digits of x and take the last (right -most) digit.
   1. Insert, in the given order the values 3836,7209,2373,9412,6950,471,5569,9703 handle collusion using linear probing. **[3M]**
   2. Then, remove in the given order 3836,9412 and 5569. After every value deletion you have to rehash the existing values **[1M]**
4. Give the following directed graph

*![Diagram

Description automatically generated]()*

Use Dijkstra’s algorithm to find the shortest path from 1 to 8.  **[4M]**

1. Insert the following sequence of elements into an AVL tree, starting with an empty tree: 100,200,300,400,500,600,700,450,475,550,800, 750,725,10,5,150,350. Show all steps. **[4M]**
2. Perform a breadth-first search of the following graph, where E is the starting node. In-order words, show the output if we issue the call BFS (E). provide two cases: (a) use a counter clockwise ordering from the top (12oo clock position) (b) use a clockwise ordering from the top. **[1+1=2M]**

![A picture containing transport

Description automatically generated]()

1. The following table shows the current undergraduate and M.E enrolments for the college of engineering (COE)

|  |  |  |
| --- | --- | --- |
| Course(Dept) | No of students | Prob |
| I (Civil & Ev) | 121 | .07 |
| II ( Mech.Eng) | 389 | .23 |
| III (Mat.Sci) | 127 | .07 |
| VI (EECS) | 645 | .38 |
| X (Chem.eng) | 237 | .13 |
| XVI(Aero & Astro) | 198 | .12 |
| **Total** | **1717** | **1.0** |

* 1. Design a variable length Huffman code the minimizes the average number of bits in messages encoding the departments of randomly chosen groups of student. Show your Huffman tree and give the code each course.  **[3M]**
  2. If you code is used to send messages containing only the encodings of the departments for each student in groups of 100 randomly chosen students, what’s the average length of such messages? write an expression. **[1M]**

1. The matrix chain order P ={P0,P1, P2 ,P3 ,P4, P5} = {5,10,3,12,5,50} the objective is to find minimum number of scalar multiplications required to multiply the 5 matrices and also find the optimal sequence of multiplications. **[5M]**
2. Show the tree constructed by kruskal’s minimum spanning tree algorithm for the graph in the following fig. label each edge of your tree by a number that indicates the order in which the edges were added to the tree ( so the first edge added will be labelled “1”, the second edge added will be labelled “2” etc..,) **[3M]**

![Diagram

Description automatically generated]()

1. Use floyd’s algorithm to determine the lengths of the shortest paths between all vertices in the following undirected weighted graph: ![Shape, arrow

   Description automatically generated]()

Explain what you have computed at each stage **[4M]**

1. You must find 1000 most expensive items from an unsorted price list contraining 107 different items. Two schemes of solution are as follows.

**Scheme A**: repeat 1000 times the sequential search (with the linear complexity O(n)).

**Scheme B**: convert the list into an array (the same complexity O(n) as for search), then sort the array (complexity O(n log n)) and fetch 1000 top items.

Which scheme would you prefer assuming that searching for 100 items in the unsorted list takes 0.1 millisecond (ms) and sorting of 100 items also takes 0.1 ms? Time for fetching data should not be taken into account. **[2M]**

**\*\*\*End of Test\*\*\***