Birla Institute of Technology & Science - Pilani, Hyderabad Campus
Second Semester 2017-2018
CS F211: Data Structures and Algorithms
Comprehensive Examination

Type: Closed **Time: 180 mins** **Max Marks: 105** **Date: 03.05.2018**

All parts of the same question should be answered together.

1.a. In a group of n persons (numbered from 1 to n), a *star* is someone who does not know anybody else but who is known by all other persons. Our goal is to identify a star, if one exists, in the group. The only actin that can be taken is to ask a question to any person i: "Do you know person j?" We assume that everybody tells the truth. [2 + 6 Marks]

(i). How many stars can exist in the group?

(ii). Design an algorithm to find the star (if any) that requires O(n) questions.

1.b. The goal is to compute simultaneously the maximum and the minimum of n integers, and we study the complexity of the algorithm in terms of number of comparisons in the worst case.

(i). Design a naïve algorithm and give its complexity.

(ii). One idea to improve the algorithm is to group elements by pairs, in order to decrease the number of comparisons that must be done. Design an algorithm based on this idea, and analyze its complexity and show that number of comparisons is Ciel(3n/2) – 2. [2 + 6 Marks]

2.a. Given a sequential file that contains at most 4 billion 32-bit integers in random order (one billion = 109). Find a 32-bit integer that is not present in the file. [6 Marks]

Note: It is obvious to observe that multiple numbers will be missing. Assume you could create temporary files but you have only a few hundred bytes of main memory to use.

2.b. An editor of *The History of the 20th Century World Science* wants to efficiently find out the year of 20th century when the most number of prominent scientists were alive. The prominent scientists are, by definition, the people mentioned in the book with the dates of their birth and death. (No living scientists are included in the book.) Devise an algorithm for this task. The entries in the index are names of scientists sorted alphabetically and each entry has the scientist's birth and death years. If person A died the same year person B was born, assume that the death event happened before the birth. Assume that you are given the book's index as input and that you can only traverse the list of scientists only once. [6 Marks]

3.a. You are given a dictionary of all English words. Using your knowledge of data structures and algorithms, store the words in such a way that all anagrams of a word can be found efficiently. [6 Marks]

Note: An anagram is a word formed by rearranging the letters of a different word, typically using all the original letters exactly once. For example, the word anagram can be rearranged into "nagaram".

3.b. Two elements of a binary search tree (BST) are swapped by mistake. You are given the pointer to root of tree. You can swap the values of any two nodes. Without modifying the tree structure and using no extra space (you can use at most 3 pointers), write steps to recover the original tree. [6 Marks]

3.c. Prove that the height of a Red-Black Tree is O(log n). [6 Marks]

4.a. Derive best, worst and average time complexities of searching for an element in hash tale where in collision is resolved by chaining. [8 Marks]

4.b. What is the running time of HEAPSORT on an array A of length n that is already sorted in increasing order? What about decreasing order? [6 Marks]

4.c. Let T be a table of n relative integers. We want to find the maximum sum of contiguous elements, namely, two indices i and j ($1 \leq i \leq j \leq n$) that maximizes $\Sigma_{k=i}^{j} T[k]$.

(i) If the values in the table are T[1] = 2. T[2] = 18, T[3] = -22, T[4] = 20, T[5] = 8, T[6] = -6, T[7] = 10, T[8] = -24, T[9] = 13, and T[10] = 3, can you return the two indices and the corresponding optimal sum?

(ii) Design an algorithm that returns the maximum sum of contiguous elements with a divide and conquer algorithm. [2 + 7 Marks]

5.a. A Toeplitz matrix, or diagonal constant matrix, named after Otto Toeplitz, is an nXn matrix with ($a_{i,j}$) coefficients ($1 \leq i, j \leq n$) and such that $a_{i,j} = a_{i-1,j-1}$ for $2 \leq i, j \leq n$. [2 + 4 Marks]

(i) Let A and B be two Toeplitz matrices. Is A + B a Toeplitz matrix? And the product A X B?

(ii) Give an algorithm to add two Toeplitz matrices in O(n).

5.b. Design an efficient algorithm (time complexity – $O(n^3)$ and space complexity $O(n^2)$)for the matrix chain multiplication problem that outputs a fully parenthesized expression for how to multiply the matrices in the chain using the minimum number of operation. [6 Marks]

5.c. Given a matrix A of size n X m with coefficients in {0, 1}, we want to find the maximum width K of a square of ones in A, as well as the coordinates (I, J) of the top left corner of such a square. In other words, for all i, j such that $I \leq i \leq I + K - 1$ and $J \leq j \leq J + K - 1$, we have A[i, j] = 1. [6 + 2 Marks]

(i) Design a dynamic programming algorithm to solve this problem.

(ii) What is the complexity of this problem?

Hint: Consider t[i,j], the width of the biggest square of ones whose top left corner is (i, j).

6.a. Woody the woodcutter will cut a given log of wood, at any place you choose, for a price equal to the length of the given log. Suppose you have a log of length L, marked to be cut in n different locations labeled 1, 2, . . ., n. For simplicity, let indices 0 and n + 1 denote the left and right endpoints of the original log of length L. Let di denote the distance of mark i from the left end of the log, and assume that $0 = d0 < d1 < d2 < \ldots < dn < dn+1 = L$. The wood-cutting problem is the problem of determining the sequence of cuts to the log that will cut the log at all the marked places and minimize your total payment. Give an efficient algorithm to solve this problem. [8 Marks]

6.b. Find the shortest distance between vertex ‘s’ and all other vertices ‘a’,’b’,’c’,’d’ by applying Dijkstra’s algorithm. [8 Marks]