BCSE 3rd Year 2nd Semester - 2011

Design and Analysis of Algorithms

All questions carry equal marks

Answer **any five** questions

Sub parts of a question must be answered contiguously.

Answers must be precise.

Time : 3 Hours | Full Marks: 100

1 (a) State and explain the fundamental characteristics of a valid algorithm.

(b) Define Big-O notation as applied to represent the complexity of running time of an algorithm.

Running time of T(n) of an algorithm with input size n is expressed as

$$T(n) = pn^2 + qn + r$$

State with justification in each of the following cases, whether T(n) can be expressed as (i) $O(n^2)$ (ii) $O(n^3)$

(c) Suppose you have algorithms with following running times. How much slower does each of these algorithms get when the input size is doubled ?

(i) n^2 (ii) 2^n

(d) Suppose you have algorithms with following running times. Also, suppose you have a computer that can perform 10^{10} operations per second, and you need to compute a result in at most an hour of computation. For each of the following algorithms, what is the largest input size n for which you would be able to get the result within an hour?

(i) n^2 (ii) 2^n

(e) Derive an expression for computation time required to solve "Tower of Hanoi" problem. 6 + (2 + 2 X 2) + 2 X 1 + 2 X 1 + 4

2. (a) Consider the following problem:

An Interval Scheduler receives n requests labelled 1......n with each request i specifying a start time s_i and a finish time f_i ; $s_i < f_i$ for all i. Two requests i and j are compatible if the requested intervals do not overlap, i.e. if $f_j <= s_i$ or $f_i <= s_j$. A subset A of requests is compatible if all pairs of requests in A are compatible.

Design a Greedy algorithms to select a compatible subset of requests of maximum possible size and prove informally that this algorithm yields correct results.

(b) Now, suppose each request (as mentioned in 2(a)) interval i has an associated value $v_i > 0$ and the problem is to find a compatible subset of intervals of maximum total value.

Explain a Greedy algorithm will not work in this case.

Design an algorithm based on the technique of dynamic programming, to solve this problem of weighted interval scheduling.

(c) What is the computation time complexity in the average case, for sorting a sequence of n integers in non-decreasing order using

(i) Merge sort (ii) Quicksort

If you do not have any a-priori knowledge about the ordering of integers in input sequence, which one of these two algorithms will you prefer and why?

(5 + 2) + (2 + 5) + (2 X 1 + 4)

3. (a) Define, with the help of appropriate examples, the terms P, NP, NP Complete and NP Hard as used in the context of expressing time complexity of algorithms.

(b) In a social networking web-site environment, every member has a set of friends. A set of members in this environment is said to form a group if any two members in the set are friends. Show that the problem to find a group with maximum number of members is NP Complete.

(c) Suggest a problem which would be computationally even harder to solve as compared with the problem stated in 3(b). Justify your answer critically.

(4 X 3) + 4 + 4

4. (a) Given a rod of length n inches and a table of prices p_i for i = 1,2,3,....,n, the problem is to find the maximum revenue r_n obtainable by cutting the rod and selling the pieces. Note that if the price p_n for a rod of length n is large enough, an optimal solution may require no cutting at all. Design an efficient algorithm to solve this problem.

Explain the methodology suggested by you with the help of an example when the length of the rod is 7 inches and price-wise prices are as mentioned below :

Length i	1	2	3	4	5	6	7	
Piece p_i		1	5	8	9	10	17	17

(b) You have an n X n square cell. A path in the cell goes along the horizontal and the vertical edges of the grid. A monotonic path from the lower left corner to the upper right corner is a path along the horizontal and vertical edges of the grid that do not cross the diagonal from the lower left corner to the upper right corner i.e. it stays below the diagonal. Find out the number of such monotonic paths in the n X n cell.

(c) A list x_1 , x_2 ,......, x_n is said to be cyclically sorted if the smallest number is the list is x_i for some unknown i and the list x_i , x_{i+1} ,......, x_n , x_1 , x_2 ,......., x_{i-1} are sorted in non-decreasing order. As an example, the list 4,5,1,2,3 is cyclically sorted. Given such a cyclically sorted list of n numbers design an efficient algorithm to find the presence (or not) of a particular number in this list. Also, comment on the time complexity of the algorithm suggested by you.

(5 + 3) + 7 + (3 + 2)

5 (a) Solve the following 0-1 Knapsack problem using branch and bound technique :

The input consists of 4 objects and
Values of objects (V_i) : 20, 3, 7, 5
Weights of objects (W_i) : 5, 2, 4, 1
Capacity of Knapsack (C) : 8

Comment critically on the complexity of the algorithm you have deployed for this purpose and indicate the change in complexity when the problem does not have any restriction on "0-1" condition.

(b) Suppose you are managing the construction of advertisements boards on an Expressway that runs west-east for M miles. The possible sites of the advertisement boards are given by numbers x_1, x_2, ..., x_n each in the interval [0, M] specifying their position along the Expressway, measured in miles from the western end. If you place an advertisement board at location x_i, you receive a revenue of $r_i > 0$.

Regulations imposed by the Expressway Authority of the country require that no two advertisement boards be within less than or equal to 5 miles of each other. The problem you need to solve is to place the advertisement boards at a subset of the sites so as to maximise your total avenue, subject to this restriction.

Design an algorithm using the technique of dynamic programming to solve this problem. The running time of this algorithm should be polynomial in n

Give a trace of the algorithm suggested by you for the following example :

$$M = 20, \qquad n = 4$$

$$[\, x_1, x_2, x_3, x_4 \,] = [\, 6, 7, 12, 14 \,]$$

and

$$[\, r_1 , r_2 , r_3 , r_4 \,] = [\, 5, 6, 5, 1 \,]$$

(c) Discuss Kruskal's algorithm for finding Minimum Spanning Tree of a connected graph and

Critically comment on the optimality of the solution provided by this algorithm.

(5 + 2) + (5 + 2) + (4 + 2)

6. (a) We are provided with a sequence of n numbers a_1, a_2, ..., a_n which we assume to be all distinct – and, we define an inversion to be pair $i < j$ such that $a_i > a_j$. Now, let us call a pair as significant inversion if $i < j$ and $a_i > 2a_j$.

Design an efficient algorithm to count the number of significant inversions in such a sequence of numbers. Comment critically on the complexity of the algorithm proposed by you.

(b) Given two strings X = [x_1, x_2, ..., x_n] and Y = [y_1, y_2, ..., y_m], the Shortest Common Supersequence [SCS] is a minimum length string Z such that both X and Y are subsequences of Z.

As for example, if X = [abcbdab] and Y = [bdcaba], one SCS is Z = [abdcabdab]

Design an efficient algorithm to find the SCS Z and its length, given two input strings X and Y of length n and m respectively. Comment critically on the complexity of algorithm proposed by you.

(c) Discuss how the methodology of Backtracking can be used to solve 8-Queen problem and comment critically on the complexity of the algorithm you suggest for this purpose.

Illustrate your approach with the help of 4-Queen problem.

(4 + 2) + (6 + 2) + (4 + 2)

7. (a) Let A and B be two matrices with dimensions m X n and n X p respectively. The conventional methodology requires m.n.p number of multiplications to compute A X B. Now, let M_1 X M_2 X ....X M_r be a chain of matrix products. The chain may be evaluated in several different ways. The costs of any computation is the number of multiplications used and let us assume that the conventional methodology as indicated above, is followed for multiplication of any two matrices.

(i) Consider the case r = 4 and matrices M_1 through M_4 with dimensions 100 X 1, 1 X 100, 100 X 1 and 1 X 100 respectively. What is the best case and the worst case values of the cost to compute M_1 X M_2 X M_3 X M_4.

(ii) Derive an expression for the number of ways in which the final value of the product of the chain of matrices can be arrived at.

(iii) Design an efficient algorithm for finding the final value of the product of the chain of matrices, using the conventional methodology for deriving the product of any two multip0lication-compatible matrices. Discuss the complexity of the algorithm proposed by you for this purpose.

(b) Consider a file containing characters [a,b,c,d,e,f] with frequency [45, 13, 12, 16, 9, 5]. Discuss the methodology of "Huffman Code" for generating binary codes of these characters using this example. Indicate the fundamental characteristics of this methodology as a technique for generating unique binary codes for characters.

(c) What do you mean by the term "Randomized algorithms"? Explain the domains of applicability of such algorithms. Discuss how Randomized algorithms may be deployed for deriving efficient solution of "Hiring problem".

(2 + 4 + 4) + 4 + (2 + 2 + 2)