MCAC 201: Design and Analysis of Algorithms (UC)
Unique Paper Code: 223421210
Semester II
Year of Admission: 2024

Time: Three Hours **Max. Marks: 70**

***All questions are compulsory.**

1. (10) **Mathematical Review**
   a. (4) Asymptotic Notations: Fill in the blank with O, o, Θ, Ω, ω. If more than one answer is applicable, mention all. All logs are in base 2.
      i. $n^{\sqrt{n}} =$ ____ $(\sqrt{n})^n$
      ii. $2^{\log n} =$ ____ $2^{\log (n\hat{}2)}$ where $(n\hat{}2) = n^2$
      iii. $n^{\log 4} =$ _____ n^2
      iv. $(\log n)^{10} =$ ____ $\log (n^{10})$

   b. (6) Use induction to prove that that $a + a.r + a.r^2 + \ldots +$ upto nth term $= a(r^n - 1) / (r - 1)$

      1. The n^{th} term is _____
      2. Base Case: For $n =$ _____, LHS = _____, RHS = _____
      3. Induction Hypothesis: Suppose the state is true for $n = k$, i.e.

         _____

         (give the above statement mathematically)
      4. We will prove that the statement is true for $n = k + 1$. i.e.

         _____

         (give the above statement mathematically)
      5. Proof: (give the proof mathematically)

         $LHSofS(k+1) =$

         _____

         $= LHSofS(k) +$ _____

         $=$ _____ $+$ _____ (apply induction)

         $=$ _____ $+$ _____ (simplify the expression)

         $= RHSofS(k+1)$. Hence Proved.

2. (24) **Searching, Sorting , Divide and Conquer**

   a. (5) Given a sorted array A[1, ..., n] such that elements may not be distinct and a given 'key', give an O(log n) time algorithm that returns the index of the fifth occurrence of 'key' if it exists, else returns -1.
   b. (4) In India, car registration numbers follow the pattern ZZXXAYYYY, where ZZ is the state code (Eg. DL), XX are the district RTO codes (Eg 04) , A represents the registration year series (A, B, C, etc.), and YYYY is a unique four-digit number (Eg. 7212). Give a linear time algorithm to sort cars on their registration number deriving its time complexity.
   c. (6) Consider Quick sort where the last element is always chosen as the pivot? In each of the following case: count the number of comparisons, state whether it is worst case giving justification for your claim. Also, if it is not the worst case, modify it by changing the positions of at most two values - that means you can only swap one pair of numbers.
      i. 64, 32, 30, 15, 27, 3, 2, 1
      ii. 18, 5, 13, 10, 12, 7, 15, 3
   d. (5) Consider a variant IS2 of Insertion Sort in which we perform binary search in A[1 ... i - 1] instead of sequential search to find the location 'j' to insert $A[i]$. Derive the time complexity of IS2 in the worst case.
   e. (4) Consider the following algorithm Find(): Let m_1 and m_2 are two arbitrary elements of two sorted lists l_1 and l_2 respectively. If $m_1 < m_2$, then recurse on (right half of l_1, left half of l_2) else recurse on (right half of l_2, left half of l_1). The base case is reached when there is only one element in each list. Write a recurrence relation for the worst case running time of the algorithm and justifying it very briefly. Also, solve the recurrence to give the worst case running time.

3. **Greedy** (6) Given a graph G = (V, E), a subset V' of V is called an independent set if there is no edge in E between any pair of nodes in V'. We want to find a largest independent set. Consider the following greedy algorithm:
   a. Include the smallest degree vertex u in V' and
   b. Remove u and its immediate neighbours along with the edges incident on them from G.
   c. Repeat on the remaining graph

   Give an example to show that the above greedy algorithm will not give the optimal.

4. (14) **Graphs (BFS/DFS) (SP/MST)**

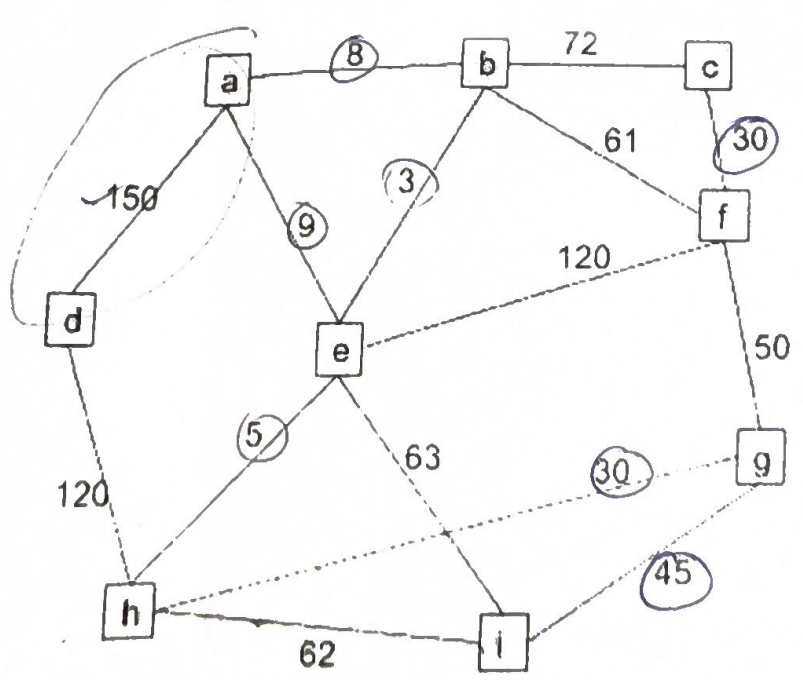
a. (7) A game has n rooms and m tunnels between them. Each room has a certain number of coins. Give an algorithm to compute the maximum number of coins you can collect while moving through the tunnels when you can freely choose your starting and ending room? Each tunnel is a one-way tunnel.

b. (7) Each department of Delhi University has its own LAN to which all the computers in the department are connected. Each of these LANs is further connected to one or more nodes in a hub of computers maintained by Delhi University Computer Center. The hub provides Internet connectivity to the LANs via one router in the hub. There is a cost of laying down cables between every pair of computers in the hub. DUCC needs to connect the computers in the hub so that each LAN is connected to the Internet. There is an important link 'e' which must be used.

i. Model the problem using graphs?

ii. What plan the company can use to efficiently connect all the nodes in the hub at minimum cost ensuring that every LAN gets connected to the Internet subject to inclusion of 'e'?

iii. Apply the above plan in part (b) on the following example with 'e = (a, d)'. Show all the steps. What is the total cost paid by the company?



(2+ 2 + 4)

5. (11) DP

a. (5) Run the DP algorithm done in the class to determine the maximum profit earned from the following set of intervals. The intervals are specified as (s_i, f_i, p_i) where s_i, f_i, p_i specify the start time, finish time and the profit earned from the i^{th} interval respectively.

(0, 3, 2), (2, 5, 4), (4, 7, 4), (2, 10, 7), (8, 11, 2), (9, 12, 1)

b. (6) A subsequence of a string is any ordered subset of characters of the string. A palindrome is a string that reads the same backward as it does forward. Fill in the blanks to give a DP solution to find the longest subsequence that forms a palindrome in a given string s. Your solution should take $O(n^2)$ time where $n = |s|$.

Suppose s = "txqybqtc". The longest palindromic subsequence is "tqbqt".

i. Let opt(i, j) represents the _____ of _____ that forms a palindrome.
ii. If s[i] == s[j], then opt(i, j) = _____.
iii. If s[i] != s[j], then opt(i, j) = _____.
iv. Combine.
v. Base Case?

6. (5) Compute the KMP failure links for the following pattern

abcwabcwtabctr