(CS1.301) Algorithm Analysis and Design (Monsoon 2022)

End of Semester Examination

Alloted time: 180 minutes

Total marks: 90

Please read the instructions VERY carefully.

- There are a total of 12 questions with varying credit, printed on pages 2 till 5.
- Discussions amongst the students are not allowed. No electronic devices (including smart watches) nor notes/books of any kind are allowed.
- Any dishonesty shall be penalized heavily.
- Any theorem/lemma/claim/fact that was proved in the class can be used without proof in the exam, only by explicitly writing its statement, and a clear remark that it was chosen from the class notes.
- Questions have been framed to be disambiguous, and queries may not be answered during the examination. In case you find any ambuguity, please mention that in your answer scripts and work with it. Answers got by misreading of questions may not be given credit.
- Be clear in your arguments. Partial marking is available for every question but vague arguments shall not be given any credit.
- You do not have to write a pseudocode unless explicitly asked.
- Analysis of running times, and proofs of correctness need to be done unless explicitly asked not to.
- Questions start on the next page.
- There are hints for a couple of problems on page 5.

Question 1

[3+3 marks]

(i) For a very large $n \geqslant n_0$, write the following asymptotic terms in an increasing order of Big-O notation.

(a) $n^{\log(\log n)}$

(b) $n^{(\log n)^2}$

(c) 2^n

(d) $n^{\log n}$

(e) $2^{\sqrt{n}}$

(f) n^{100}

2^{64} $16 \times 16 = 256$ 2560

(ii) For n that is large enough, and constants a, b, c, compute the asymptotic growth of the following function.

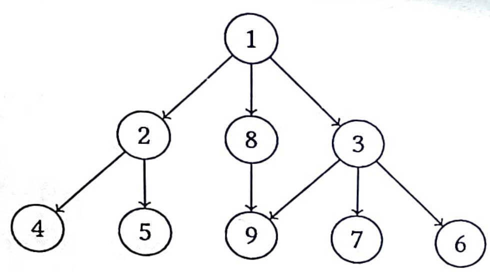
$$T(n) = a \cdot T\left(\frac{n}{b}\right) + n^c.$$

That is, what is the value of N in terms of n, a, b, c such that $T(n)$ can be written as $O(N)$.

Question 2

[3+3 marks]

Let $G = (V, E)$ be a graph as drawn below. Suppose we were to run graph search algorithms from



the node 1 to discover the node 7, what is the sequence of nodes discovered before discovering the node 7 if the search algorithm used is

(a) Breadth-first search.

(b) Depth-first search.

Assume that each of these search algorithms prioritises a neighbour of least value.

Question 3

[8 marks]

A vertex u in a connected (undirected) graph $G = (V, E)$ is called an articulation point if removal of u from the vertex set, and removal of edges incident on the vertex u disconnects the graph. Give an algorithm to find articulation point(s) in a given graph.

Question 4 [4+2 marks]

Suppose the symbols a, b, c, d, e occur with frequencies $0.5, 0.25, 0.125, 0.0625, 0.0625$, respectively.

(a) What is the Huffman encoding of this alphabet?

(b) If this encoding is applied to a file consisting of 1,000,000 characters from a, b, c, d, e, with the given frequencies, what is the length of the encoded file in bits.

Question 5 [10 marks]

A path is monotonically increasing if the weight of every edge (considered in sequence) on that path is strictly increasing. Given an edge-weighted directed graph, find a monotonically increasing shortest path from s to every other vertex.

Question 6 [8 marks]

Given a sorted array of distinct integers $A[1, n]$, we want to find out whether there is an index i for which $A[i] = i$. Give a divide-and-conquer algorithm that runs in time $O(\log n)$.

Question 7 [8 marks]

A k-way merge operation is as follows. Suppose you have k sorted arrays, each with n elements, and you want to combine them into a single sorted array of kn elements (for example, in each iteration of merge sort, we do a 2-way merge operation). Give an efficient solution to this problem, using divide-and-conquer.

Question 8 [6 marks]

We say that a $2n$-length string over the letters L and R is balanced if it is properly nested. L could be thought of as an opening bracket and R could be thought of as a closing bracket.

For example,

- for $n = 1$, <u>LR</u> (which corresponds to [ ]) is balanced but <u>RL</u> (which corresponds to] [ ) is not balanced.
- For $n = 2$, <u>LRLR</u> (which corresponds to [ ] [ ]) and <u>LLRR</u> (which corresponds to [[ ]]) are balanced but <u>RLRL</u> (which corresponds to] [ ] [ ) and <u>RLLR</u> (which corresponds to] [[ ]) are not balanced.

Write a recursive function (either mathematically or using pseudocode) to enumerate (or print out) all the strings of length $2n$ over the letters L, R, that are balanced. In other words, using recursion print out all the balanced $2n$-length strings over letters L, R. (No proof of correctness is required.) Here, the input to the recursive function is n.

$\frac{10}{8}$ 0.125 200 ~~400~~ 16 0.06

$T(n) : 2^k . T(\frac{n}{2^k})$

$n = 2^k$ $n \log n$ $T(1)$ n

[[ ]

$T(n) = 2T(n/2)$

$T(n/2) = 2T(n/4)$

$T(n) = 2(2T(\frac{n}{4})$

$n = 2^k \Rightarrow$ $2^k . T(\frac{n}{2^k})$

~~160000~~ 4 125000 × 3 ~~250000~~

1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8

Question 9 [10 marks]

A polygon is convex if every line that does not contain any edge of the polygon intersects the polygon in at most two points.
You are given a convex polygon P on n fixed vertices in the plane (specified by their x and y coordinates). A triangulation of P is a collection of $n-3$ diagonals of P such that no two diagonals intersect (except possibly at their endpoints). Notice that a triangulation splits the polygon's interior into $n-2$ disjoint triangles (see Figure 1 for a visualization).
The cost of a triangulation is the sum of the perimeters of the triangles in it. Give an efficient algorithm for finding a triangulation of minimum cost.

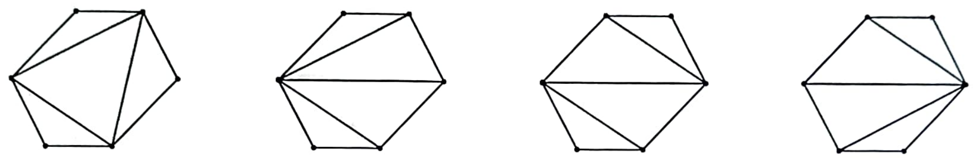


Figure 1: Some of the many different ways to triangulate a convex polygon on 6 vertices. Note that the second and the last figures are distinct.

Question 10 [8 marks]

For the following network, with edge capacities as shown in Figure 2, find the maximum flow from s to t along with the associated minimum cut.

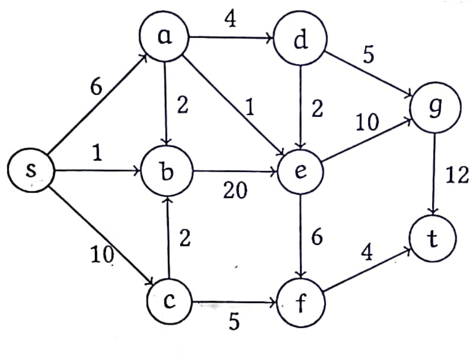


Figure 2: Figure for question 10

Please work this problem out by first starting with a feasible flow and then listing out all the intermediate steps of computing the residual graphs and augmentations until the maximum flow is found.

Question 11 [8 marks]

An edge of a flow network is called critical if decreasing the capacity of this edge (by any positive integer-valued quantity) results in a decrease in the maximum flow. Give an efficient algorithm that finds a critical edge in a network.

Question 12 [3+3 marks]

Argue why the following problems are in NP.

1. **Minimum Spanning Tree**: Given a graph $G = (V, E)$ along with edge weights $\{w_e \mid e \in E\}$, find the minimum spanning tree of G.

2. **Integer factoring**: Given integer A and k, check if there is an integer b such that $1 < b \leqslant k$, and b divides A.

Hints

- Hint for question 3: One of the search algorithms discussed in the class could be of use.

- Hint for question 9: Label the vertices of P by $\{1, 2, \ldots, n\}$, starting from an arbitrary vertex and walking clockwise. For $1 \leqslant i < j \leqslant n$, let the subproblem $A(i, j)$ denote the minimum cost triangulation of the polygon spanned by vertices $i, i + 1, \ldots, j$.