There are four problems below. Please choose three to solve. If you choose to solve all four, only the lowest three scores will count. Show all your work and write in a readable way.

Question 1

In computational biology, DNA can be represented as a sequence of characters drawn from an alphabet of four letters, A, C, T, and G, representing the four nucleotides. Given two sequences S_1 and S_2 of n and m characters, respectively, describe what is meant by a local alignment. Given a similarity score of +2, a mismatch penalty of -1, and a gap score of 0, give an efficient sequential algorithm to compute the score of the best local alignment between S_1 and S_2. What is the asymptotic complexity of your algorithm? What are the space requirements? Suppose now that you are given a distributed memory parallel computer with p nodes (with $1 < p < \min(n, m)$), design and analyze a parallel algorithm for sequence similarity problem using local alignments that scales with the number of nodes.

Question 2

Let S be a subset of the first n integers. Design a data structure supporting the following three primitive operations:

(a) INSERT(i): Insert an assigned integer in S (void if it is already there).

(b) DELETE(i): Delete an assigned integer from S (void if it is not there).

(c) BETWEEN(i, j): Output the number of elements that fall between i and j. For instance, if S includes 2,4,6,8 and 11, and the two items in the query are 3 and 7, then the answer is 2.

Each primitive must be executed in time $O(\log n)$.

Question 3

Let G be a weighed, undirected and connected graph with positive edge weights. Prove or confute the following claims:

A) If all edge weights in G are distinct, then G has a unique minimum spanning tree.

B) If there are two edges in G with the same weight, then G has more than one minimum spanning tree.

Question 4

The classical Traveling Salesperson Problem requires to input a weighted undirected graph and find a minimum-weight cycle that includes all vertices of the graph. We consider the modified version of the problem, which requires to find a minimum-weight *path* through all vertices; that is, the salesperson has to visit all vertices without returning to the initial vertex. The classical Traveling Salesperson Problem is known to be NP-complete. Use this fact to prove that the modified problem is also NP-complete.