Daffodil International University
Faculty of Science & Information Technology
Department of Computer Science & Engineering
Final Semester Examination, Fall 2024
Course Code: CSE115 Course Title: Introduction to Biology and Chemistry for Computation
Level: 1 Term: 1 Batch: 67

Time: 2:00 Hrs **Marks: 40**

Answer ALL Questions

[The figures in the right margin indicate the full marks and corresponding course outcomes. All portions of each question must be answered sequentially.]

<table>
<tr><td>1.</td><td>Computational chemistry uses computer simulations to solve chemical problems, modeling molecular structures and reactions with theoretical methods and algorithms

Compare and Contrast the principles and applications of Molecular Mechanics (MM) and Quantum Mechanics (QM) in the study of molecular systems.

Provide examples of when each method is most suitable for studying molecular properties and interactions, highlighting their strengths and limitations.</td><td>[5]</td><td>CO2</td></tr>
<tr><td>2.</td><td>Dr. Smith, a geneticist, was studying the DNA of two patients to identify mutations linked to a rare genetic disorder using advanced computational tools, she compared their DNA sequences to a reference genome. After running the analysis, she identified two sequences of interest:

Patient 1 DNA sequence: TTATGCAATT
Patient 2 DNA sequence: TGCCTTCA

Apply an efficient algorithm to analyze the sequences, identify similarities, and detect the gene responsible for the disorder. Also, Describe how the alignment scores are justified.

Consider the following scoreboard for calculation.<table><tr><td>Match: +2</td><td>Mismatch: -2</td><td>Gap Penalty: -1</td></tr></table></td><td>[10]</td><td>CO3</td></tr>
<tr><td>3.</td><td>The Burrows-Wheeler Transform (BWT) is a significant algorithm in bioinformatics to efficiently compress and index genomic data. It facilitates rapid sequence alignment, variant detection, and genome assembly, making it crucial for analyzing large-scale genomic datasets. Suppose a bioinformatician has given you the following BWT(T) sequence: AIONCRM$OTHDIA

I. Discover the original gene using LF Mapping from the given BWT(T).
II. Construct Suffix Array, Suffix tree and Keyword Tree from the original sequence you retrieved in (I).</td><td>[10]</td><td>CO3</td></tr>
</table>

4. Dr. Towmid, an astrophysicist, receives a mysterious transmission from outer space. The message is a long string of symbols (the query sequence) that seems to hold a hidden pattern. To decode the message, Dr. Towmid compares it to a massive database of known alien signal fragments to find matching patterns that might provide clues to its meaning. To speed up the process, his teammate Dr. Surovi wants to use the FASTA algorithm and creates a hash table. **[10]** **CO3**

The alien message (query sequence) is: **ZXQWERTYUIO**

A known alien signal fragment (target sequence) is: **ERTYZXQWE**

Both the doctors have decided to take the K values as 1 so they can get more possible matches.

I. **Apply** Fasta Algorithm using hash table to show the match between the query sequence and the target sequence, and at what positions the match starts?

II. **Analyze** the change in the Final Table if the value of K were 5.

5. Sajid is working with a system of DNA sequence and a dataset of nitrogen bases that is given below: **[5]** **CO3**

T	C	A	G	C	A	T	G
A	G	C	A	G	T	T	T
C	T	G	C	T	A	G	G

Sajid is fond of Thymine nitrogen base. So he was searching for T. Each time he found a T, he printed it. So the expected outcome should have been T T T T T but he found A T G T T.

Compute the Selectivity and Sensitivity of the given scenario and **Discuss** the steps.