ourdiu.com | Shared by Shaker Abdullah

Daffodil International University

Faculty of Science & Information Technology

Department of Computer Science & Engineering

Final Examination, Fall 2025

Course Code: CSE115

Course Title: Introduction to Biology and Chemistry for Computation

Level: 1 Term: 1 Batch: 70

Time: 02:00 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.]

1.	Mr. Hasan, a prospective graduate student, was exploring the genetic similarities between domestic cats and tigers and discovered that their DNA sequences are nearly 95% identical. To examine a specific gene region, he isolated a small fragment of DNA and introduced it into a simple unicellular organism so it could replicate. During analysis, he noticed that the sequencing process required primers and a special type of nucleotide that stops strand extension, allowing him to read the sequence in an ordered manner.			**CO2**
	a)	**Analyze** which ancient unicellular organism would be suitable for accepting and replicating foreign nucleotide molecules.	**[4]**	
	b)	**Examine** the scenario and find out the appropriate sequencing method and **explain** the complete procedure step-by-step.	**[6]**	
2.	A biotechnology research lab is studying the genomes of two closely related plant species to understand patterns of gene duplication and mutation. The researchers often need to store, search, and align large DNA sequence datasets efficiently. During one stage of the analysis, the lab receives a short DNA sequence fragment: **ATCATG** To improve both storage efficiency and search performance, the bioinformatics team decides to apply the Burrows–Wheeler Transform.			**CO3**
	a)	**Demonstrate** how the lab can construct a Burrows–Wheeler Matrix (BWM) for the given DNA sequence and generate the Suffix array (SA) and BWT.	**[6]**	
	b)	**Apply** the LF-mapping process to show how the original DNA sequence can be reconstructed during decoding and mention the reason behind using occurrence numbers when the same character appears multiple times.	**[4]**	

ourdiu.com | Shared by Shaker Abdullah

3.	During a campus health surveillance week, the university clinic screened 160 students using a new COVID-19 rapid test. All students were then sent for confirmatory PCR testing. The PCR results revealed that 50 students were actually infected. Of these infected students, 32 had shown symptoms but 18 were asymptomatic. After that, the rapid test results were reviewed, and it was found that the rapid test correctly identified 30 of the 32 symptomatic infected students and 10 of the 18 asymptomatic infected students as positive. During the evaluation of a new COVID-19 rapid test, a university research lab performs genetic verification on samples collected from students. For one of the samples, the lab extracts and sequences a portion of the SARS-CoV-2 spike protein, producing the following query fragment: Query sequence (Q): VYGPTLNYN To confirm whether this fragment truly matches the known viral genome, the lab aligns it with a reference viral gene segment stored in their database: Target sequence (T): QKVYGPTLYYNE			CO3
	a)	**Calculate** the sensitivity of the COVID-19 rapid test for detecting infected students, and explain how a higher sensitivity would affect the test's precision in campus-wide screenings.	[4]	
	b)	**Illustrate** an algorithm that will find the match between query sequence and target sequence where k=1 and show the regions of similarity.	[6]	
4.	A team of campus molecular epidemiologists at the University Health Lab screened 200 students with a new influenza rapid test, sequenced short viral fragments from positive samples, and sampled five local viral variants (V–Z). They produced aligned 10-base sequences (There is a gap in Z). **V:** A A A A A A A A C G **W:** T T T T A A A A C G **X:** T T A A A A T T C G **Y:** A A T T T A T A C G **Z:** G G G C C _ A A C G Now they are planning to compute pairwise **p-distances** and use UPGMA to infer a phylogenetic tree that shows how the five variants are related under a molecular-clock assumption.			CO3
	a)	**Apply** the UPGMA algorithm on the given data to build a phylogenetic tree, and label all branches with their calculated branch lengths.	[10]	