ourdiu.com | Shared by Md. Ferdous Foysal Hridoy

Daffodil International University

Faculty of Science & Information Technology

Department of Computer Science & Engineering

Final Examination, Spring 2026

Course Code: CSE115, Course Title: Introduction to Biology & Chemistry for Computation

Level: 1 Term: 1 Batch: 71

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.		A research team is comparing genes involved in oxygen transport across different species. They find that a hemoglobin gene in humans is very similar to a gene found in chimpanzees. However, within humans, they also observe two similar genes: one responsible for adult hemoglobin and another for fetal hemoglobin, which arose due to gene duplication. The team also isolates a specific DNA fragment related to oxygen transport and inserts it into a fast-growing microorganism to amplify the gene. During sequencing, they observed that the process uses modified nucleotides that terminate chain elongation, helping determine the DNA sequence.		CO2
	a)	Identify which genes are orthologs and which are paralogs in this scenario. Justify your answer.	[4]	
	b)	List the four types of dideoxynucleotides (ddNTPs) used in the Sanger sequencing method. For each ddNTP, briefly explain its role/importance in DNA sequencing.	[6]	
2.		A bioinformatics team receives a transformed DNA sequence from a compressed genome database. However, the original sequence is missing, and only the Burrows–Wheeler Transform BWT(T) output is available: BWT(T): ACTGA$TA The team must reconstruct the original DNA sequence and analyze its structure for further processing.		CO3
	a)	Using the given BWT(T) string, apply the LF-mapping technique step-by-step to reconstruct the original DNA sequence.	[5]	
	b)	After reconstructing the sequence, generate the Burrows Wheeler Matrix (BWM) and identify the Suffix Array (SA) corresponding to the original sequence.	[5]	
3.		During a flu outbreak, 50 students were tested with a rapid flu test. Later, PCR tests showed that 20 students were actually infected.		CO3

ourdiu.com | Shared by Md. Ferdous Foysal Hridoy

		Among these 20 infected students, 15 were correctly identified as positive by the rapid test. At the same time, a part of a viral protein was sequenced: • Query Sequence: FLABBERGASTED • Target Sequence: LASTFLAG		
	a)	Calculate the **sensitivity** of the rapid test. What does the value of sensitivity mean?	[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		Draw a phylogenetic tree showing the following relationships: • A and B are closely related. • C and D share a common ancestor. • E diverged earlier than the others. • The common ancestor of A & B is more closely related to C & D than to E.		CO3
	a)	Write the corresponding Newick format for your phylogenetic tree.	[5]	
	b)	If the branch lengths are as follows, include them in the Newick notation and draw the tree. • Branch between A and B = 0.1 • Branch between C and D = 0.2 • Branch connecting (A,B) to (C,D) = 0.3 • Branch connecting E to the rest = 0.5	[5]	