Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 09-10-2025 **Time:** 09.30am to 11.00am

School: SOCSE	**Program:** COMPUTER SCIENCE AND ENGINEERING	
Course Code: CSE3507	**Course Name:** Algorithms in Computational Biology	
Semester: V	**Max Marks**: 50	**Weightage**: 25%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**14**	**36**	**---**	**---**	**---**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **5Q x 2M=10M**

1	Describe the fundamental components that define an algorithm, including its structure, clarity, and effectiveness in solving a problem.	**2 Marks**	**L2**	**CO1**
2	Interpret how to write a pandas DataFrame to a CSV file in Python while excluding the index column. Provide the relevant syntax and reasoning.	**2 Marks**	**L2**	**CO1**
3	Distinguish between Comparative ImmunoProfiling and the Consecutive Integer Problem. List four differences in their domain, purpose, data type, and analytical method.	**2 Marks**	**L2**	**CO2**
4	Summarize two practical strategies used to address NP-complete problems when exact polynomial-time solutions are not feasible. Include one example for each.	**2 Marks**	**L2**	**CO2**
5	Interpret the pattern of four consecutive odd integers using algebraic expressions, and determine a set whose sum equals 136.	**2 Marks**	**L2**	**CO2**

Part B

Answer the Questions. **Total Marks 40M**

6.	**a.**	Summarize how the merge sort algorithm is implemented in Python. Interpret how its divide-and-conquer strategy organizes data during sorting.	**10 Marks**	**L2**	**CO1**
		Or			
7.	**a.**	Describe how binary search operates on a sorted list. Interpret why its time complexity is more efficient than linear search, using Python to illustrate the process.	**10 Marks**	**L2**	**CO1**

8.	**a.**	Interpret the concept of asymptotic analysis by describing its key ideas, standard notations, and core properties. Illustrate how it is performed using a relevant example, and explain its role in evaluating algorithmic efficiency.	**10 Marks**	**L2**	**CO 2**
	b.	Explain the process of identifying and proving that a problem is NP-complete. What are the key steps involved, and why is this classification important.	**5 marks**	**L2**	**CO 2**
		Or			
9.	**a.**	Explain how a custom Fibonacci-like sequence can be generated starting from two user-defined numbers. How does this differ from the traditional Fibonacci sequence? Illustrate the output for the first 10 terms when the starting numbers are 5 and 8.	**10 Marks**	**L2**	**CO 2**
	b.	Explain the concept of Polynomial-Time Reducibility. Why is it important in computational complexity theory, and how can one reduce a problem A to another problem B in polynomial time.	**5 marks**	**L2**	**CO 2**

10.	**a.**	Explain how DNA fragments are used in Comparative ImmunoProfiling (CIP) to reconstruct biological sequences through overlap analysis. Then, apply the overlap rule (≥ 5 bases) to the fragments below to identify valid links, build the longest chain, reconstruct the full sequence by trimming overlaps, and justify any exclusions. F_1: GATCGTACGA F_2: TACGATCGGA F_3: ATCGGATACC	**10Marks**	**L2**	**CO 2**

		F_4: GGATACCGTT F_5: ACCGTTAGCA			
	b.	Outline the main procedures of Heap Sort using pseudocode. Interpret how the heap structure supports efficient sorting operations.	**5 Marks**	**L2**	**CO 2**
		Or			
11.	**a.**	Interpret the TSP-based DNA assembly approach by writing an algorithm that uses suffix-prefix overlap as edge cost. Apply this algorithm to the fragments below to determine an efficient traversal order and reconstruct the final DNA sequence. F_1: GATCGTACGA F_2: TACGATCGGA F_3: ATCGGATACC F_4: GGATACCGTT F_5: ACCGTTAGCA	**10 Marks**	**L2**	**CO 2**
	b.	Demonstrate your understanding of Radix Sort by writing its pseudocode. Interpret how the algorithm processes digits to achieve sorting.	**5 marks**	**L2**	**CO 2**