CUET PG 2026 Bioinformatics Question Paper with Solutions(Memory Based)

Time Allowed :1 Hour 30 Mins	**Maximum Marks :**300	**Total Questions :**75

General Instructions

Read the following instructions very carefully and strictly follow them:

- The exam lasts 90 minutes (1 hour 30 minutes).
- There are 75 Multiple Choice Questions (MCQs) to be answered.
- +4 marks for every correct answer. -1 mark (negative marking) for every incorrect answer. 0 marks for unanswered or un-attempted questions.
- For any discrepancy in questions, the English version is considered final (except for language-specific papers).
- Click one of the four options to choose an answer.
- You must click "Save & Next" to confirm your response. Only saved answers are considered for evaluation.
- Use "Mark for Review & Next" to flag a question for later. You can unselect or change your answer using the "Clear Response" button.
- All calculations must be done on the Rough Sheets provided at the centre. These must be returned to the invigilator after the exam.

1. Which algorithm is used for performing Global Sequence Alignment?
(A) Smith–Waterman Algorithm
(B) Needleman–Wunsch Algorithm
(C) BLAST Algorithm
(D) FASTA Algorithm

Correct Answer: (B) Needleman–Wunsch Algorithm

Solution:

Concept:
Sequence alignment is a fundamental technique in **bioinformatics** used to identify similarities between DNA, RNA, or protein sequences. There are two main types of sequence alignment:

- **Global Alignment:** Aligns two sequences from beginning to end, considering their entire length.
- **Local Alignment:** Finds the best matching subsequences within larger sequences.

The **Needleman–Wunsch algorithm** is a dynamic programming method specifically designed for **global sequence alignment**.

Step 1: Understanding Global Sequence Alignment.
Global sequence alignment compares two sequences along their entire length to determine the optimal alignment. This approach is useful when the sequences are of similar length and expected to be similar overall.
Step 2: Identifying the algorithm used.
The **Needleman–Wunsch algorithm** performs global alignment using dynamic programming by constructing a scoring matrix and tracing back the optimal path to produce the best alignment.
Step 3: Eliminating other options.

- **Smith–Waterman Algorithm** → Used for **local sequence alignment**.
- **BLAST** → Used for fast database similarity searches.
- **FASTA** → Another heuristic search algorithm for sequence comparison.

Conclusion:
Therefore, the algorithm used for **Global Sequence Alignment** is the **Needleman–Wunsch Algorithm**.

Quick Tip

Global Sequence Alignment → **Needleman–Wunsch Algorithm**
Local Sequence Alignment → **Smith–Waterman Algorithm**

---

2. What is the primary difference between PAM and BLOSUM scoring matrices?
(A) PAM matrices are based on global alignments, while BLOSUM matrices are based on local alignments
(B) PAM matrices are used only for DNA sequences, while BLOSUM matrices are used for proteins
(C) PAM matrices are derived from highly divergent sequences, while BLOSUM matrices use closely related sequences
(D) PAM matrices are heuristic methods, while BLOSUM matrices use dynamic programming

Correct Answer: (A) PAM matrices are based on global alignments, while BLOSUM matrices are based on local alignments

Solution:

Concept:
In bioinformatics, **scoring matrices** are used in sequence alignment to assign scores for amino acid substitutions. Two widely used matrices are:

- **PAM (Point Accepted Mutation)**
- **BLOSUM (BLOcks SUbstitution Matrix)**

Both matrices are used for protein sequence alignment, but they differ in how they are constructed.

Step 1: Understanding PAM matrices.
PAM matrices are derived from **global alignments of closely related protein sequences**. They estimate mutation probabilities over evolutionary time by extrapolating observed mutations.
Step 2: Understanding BLOSUM matrices.
BLOSUM matrices are created from **local alignments of conserved protein regions (blocks)** found in related proteins. They group similar sequences and calculate substitution frequencies from these conserved blocks.
Step 3: Identifying the main difference.
The key distinction is the **alignment strategy used to build the matrices**:

- PAM → Derived from **global sequence alignments**
- BLOSUM → Derived from **local conserved blocks**

Conclusion:
Thus, the primary difference is that **PAM matrices are based on global alignments, while BLOSUM matrices are based on local alignments**.

Quick Tip

PAM → Global alignment based matrix
BLOSUM → Local conserved block based matrix

---

3. Which biological database is considered a curated, secondary database for protein sequences?
(A) GenBank
(B) UniProtKB/Swiss-Prot
(C) EMBL
(D) DDBJ

Correct Answer: (B) UniProtKB/Swiss-Prot

Solution:

Concept:
Biological databases store information about DNA, RNA, and protein sequences. These databases are generally classified into two categories:

- **Primary databases** – Contain raw experimental sequence data submitted directly by researchers.
- **Secondary (curated) databases** – Contain processed, annotated, and reviewed data derived from primary databases.

Secondary databases provide more reliable information because experts carefully analyze and annotate the sequences.

Step 1: Understanding primary databases.
Primary databases such as **GenBank**, **EMBL**, and **DDBJ** store nucleotide sequences submitted by scientists. These entries are not always manually reviewed and may contain redundant or incomplete annotations.
Step 2: Understanding curated secondary databases.
UniProtKB/Swiss-Prot is a curated protein sequence database. Each entry is manually reviewed by experts who add detailed information such as:

- Protein function
- Domain structure
- Post-translational modifications
- Biological pathways
- Literature references

Step 3: Selecting the correct answer.
Among the options, only **UniProtKB/Swiss-Prot** is a curated secondary database specifically designed for protein sequence annotation.
Conclusion:
Therefore, the curated secondary database for protein sequences is **UniProtKB/Swiss-Prot**.

Quick Tip

Primary databases → GenBank, EMBL, DDBJ
Curated secondary protein database → UniProtKB/Swiss-Prot

---

4. In the Michaelis-Menten equation, what does the constant K_m represent?
(A) Maximum reaction velocity
(B) Substrate concentration at half of V_{max}
(C) Enzyme concentration
(D) Rate of enzyme degradation

Correct Answer: (B) Substrate concentration at half of V_{max}

Solution:

Concept:
The **Michaelis–Menten equation** describes the relationship between the rate of an enzyme-catalyzed reaction and the concentration of substrate. It is expressed as:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

where:

- v = reaction velocity

- V_{max} = maximum reaction velocity
- $[S]$ = substrate concentration
- K_m = Michaelis constant

Step 1: Understanding the meaning of K_m.
The constant K_m represents the substrate concentration at which the reaction rate reaches **half of its maximum velocity** ($V_{max}/2$).
Step 2: Interpreting enzyme affinity.
The value of K_m also provides information about the **affinity of an enzyme for its substrate**:

- **Low K_m value** → High affinity (enzyme binds substrate easily)
- **High K_m value** → Low affinity (enzyme binds substrate less efficiently)

Step 3: Selecting the correct option.
Since K_m indicates the substrate concentration at which the reaction rate is half of V_{max}, option (B) is correct.
Conclusion:
Thus, the Michaelis constant K_m represents the **substrate concentration required to reach half of the maximum reaction velocity**.

Quick Tip

K_m = Substrate concentration when reaction rate = $\frac{V_{max}}{2}$

---

5. Which amino acid is responsible for forming disulfide bridges in a protein structure?
(A) Glycine
(B) Cysteine
(C) Lysine
(D) Alanine

Correct Answer: (B) Cysteine

Solution:

Concept:
Proteins are stabilized by several types of chemical interactions such as hydrogen bonds, ionic interactions, hydrophobic interactions, and **disulfide bonds**. These disulfide bonds are important for maintaining the three-dimensional structure of proteins.
Step 1: Understanding disulfide bridges.
A disulfide bridge (–S–S–) is a covalent bond formed between two sulfur atoms of amino acids.
Step 2: Identifying the amino acid involved.
The amino acid **cysteine** contains a sulfhydryl group (–SH). When two cysteine residues come close together during protein folding, their sulfhydryl groups oxidize and form a **disulfide bond**.

Step 3: Role in protein structure.
Disulfide bonds help stabilize:

- Tertiary protein structure
- Quaternary protein structure
- Extracellular proteins such as antibodies and hormones

Conclusion:
Therefore, the amino acid responsible for forming disulfide bridges is **cysteine**.

Quick Tip

Two cysteine residues → Oxidation → Disulfide bond (S–S)

---

6. What is the specific function of the BLASTn program?
(A) Compare protein sequences
(B) Compare nucleotide sequences
(C) Predict protein structure
(D) Align protein structures

Correct Answer: (B) Compare nucleotide sequences

Solution:

Concept:
BLAST (Basic Local Alignment Search Tool) is a widely used bioinformatics program that compares biological sequences with sequences stored in databases to identify similarities.
Step 1: Understanding BLAST programs.
Different versions of BLAST are used for different types of sequences:

- **BLASTn** – Nucleotide vs nucleotide comparison
- **BLASTp** – Protein vs protein comparison
- **BLASTx** – Translated nucleotide vs protein comparison
- **tBLASTn** – Protein vs translated nucleotide database

Step 2: Function of BLASTn.
BLASTn specifically compares a nucleotide query sequence (DNA or RNA) against a nucleotide database to identify regions of similarity.
Step 3: Applications of BLASTn.
BLASTn is commonly used for:

- Identifying homologous genes
- Detecting species similarity
- Finding conserved DNA regions

- Genome annotation

Conclusion:
Therefore, the primary function of **BLASTn** is to **compare nucleotide sequences with nucleotide databases to find similar regions**.

Quick Tip

BLASTn → Nucleotide vs Nucleotide comparison

---

7. Which tool is most commonly used for Multiple Sequence Alignment (MSA)?
(A) BLAST
(B) ClustalW
(C) FASTA
(D) Needleman–Wunsch

Correct Answer: (B) ClustalW

Solution:

Concept:
Multiple Sequence Alignment (MSA) is a computational technique used in bioinformatics to align three or more biological sequences (DNA, RNA, or proteins). The goal of MSA is to identify conserved regions, functional domains, and evolutionary relationships among sequences. MSA plays an important role in:

- Studying evolutionary relationships
- Identifying conserved protein domains
- Predicting protein structure and function
- Phylogenetic analysis

Step 1: Understanding sequence alignment tools.
Different tools are used for sequence comparison:

- **BLAST** – Used for fast database searching and local sequence alignment.
- **Needleman–Wunsch** – Used for pairwise global sequence alignment.
- **FASTA** – A heuristic algorithm used for sequence similarity searching.

Step 2: Identifying the MSA tool.
ClustalW is one of the most widely used tools for performing multiple sequence alignment. It aligns multiple sequences by progressively aligning the most similar sequences first and building a guide tree based on sequence similarity.
Step 3: Selecting the correct option.
Among the given choices, the tool specifically designed for Multiple Sequence Alignment is **ClustalW**.

Conclusion:
Therefore, the most commonly used tool for Multiple Sequence Alignment is **ClustalW**.

Quick Tip

Multiple Sequence Alignment (MSA) tool → **ClustalW**

---

8. The p53 protein is primarily known for which cellular role?
(A) DNA replication
(B) Tumor suppression
(C) Protein synthesis
(D) Cell membrane transport

Correct Answer: (B) Tumor suppression

Solution:

Concept:
The **p53 protein** is one of the most important regulatory proteins in cells and is often referred to as the **"guardian of the genome."** It plays a crucial role in maintaining genomic stability by preventing the proliferation of cells with damaged DNA.
Step 1: Role of p53 in DNA damage response.
When DNA damage occurs due to radiation, chemicals, or replication errors, the p53 protein becomes activated. It can stop the cell cycle and allow time for DNA repair mechanisms to correct the damage.
Step 2: Cell cycle regulation and apoptosis.
If the DNA damage cannot be repaired, p53 triggers **programmed cell death (apoptosis)** to eliminate the damaged cell. This prevents the accumulation of mutations that could lead to cancer.
Step 3: Connection to cancer prevention.
Because of its ability to control cell growth and eliminate damaged cells, p53 functions as a **tumor suppressor protein**. Mutations in the p53 gene are found in many types of human cancers.
Conclusion:
Therefore, the primary cellular role of the p53 protein is **tumor suppression**.

Quick Tip

p53 protein → "Guardian of the genome" → Prevents cancer by stopping damaged cells from dividing

---

9. Which molecular technique is used to detect specific mRNA molecules in a sample?

(A) Southern Blot
(B) Northern Blot
(C) Western Blot
(D) PCR

Correct Answer: (B) Northern Blot

Solution:

Concept:
Blotting techniques are important molecular biology methods used to detect specific biomolecules such as DNA, RNA, or proteins. Each blotting method is designed to identify a particular type of molecule.
Step 1: Understanding different blotting techniques.
The major blotting techniques include:

- **Southern blot** – Used to detect specific DNA sequences.
- **Northern blot** – Used to detect RNA molecules, especially mRNA.
- **Western blot** – Used to detect specific proteins.

Step 2: Role of Northern blot.
The **Northern blot** technique separates RNA molecules using gel electrophoresis, transfers them onto a membrane, and detects specific RNA sequences using labeled probes.
Step 3: Application in gene expression analysis.
Northern blot is commonly used to:

- Detect specific mRNA molecules
- Measure gene expression levels
- Study transcription patterns

Conclusion:
Therefore, the molecular technique used to detect specific **mRNA molecules** in a sample is the **Northern blot** technique.

Quick Tip

Southern blot → DNA
Northern blot → RNA (mRNA detection)
Western blot → Protein

---

10. In computer science, what does the Markov Property assume about future states?
(A) Future states depend on all previous states
(B) Future states depend only on the current state
(C) Future states depend only on the initial state
(D) Future states are completely random

Correct Answer: (B) Future states depend only on the current state

Solution:

Concept:
The **Markov Property** is a fundamental concept in probability theory and computer science, especially in models such as **Markov Chains**, **Hidden Markov Models (HMM)**, and **Markov Decision Processes (MDP)**.
The Markov Property states that the probability of transitioning to the next state depends only on the **current state**, and not on the sequence of events that occurred before it.
Step 1: Understanding state transitions.
In many computational models, systems move between different states over time. These transitions can be represented probabilistically.
Step 2: Memoryless assumption.
The Markov Property introduces a **memoryless assumption**. This means the future state depends solely on the present state and ignores past states.
Step 3: Applications of the Markov Property.
This property is widely used in:

- Hidden Markov Models used in bioinformatics
- Speech recognition systems
- Natural language processing
- Reinforcement learning algorithms

Conclusion:
Thus, the Markov Property assumes that **future states depend only on the current state and not on previous states**.

Quick Tip

Markov Property → Future state depends only on the **current state** (memoryless system)

---

11. Which database serves as the main archive for 3D macromolecular structures?
(A) GenBank
(B) Protein Data Bank (PDB)
(C) UniProt
(D) EMBL

Correct Answer: (B) Protein Data Bank (PDB)

Solution:

Concept:
Structural biology focuses on determining the three-dimensional structures of biological macromolecules such as proteins, DNA, and RNA. These structures are essential for understanding molecular function and interactions.
Step 1: Understanding structural databases.
Specialized databases store experimentally determined 3D structures of biological molecules.
Step 2: Protein Data Bank (PDB).
The **Protein Data Bank (PDB)** is the primary global archive that stores three-dimensional structures of biological macromolecules determined using techniques such as:

- X-ray crystallography
- Nuclear Magnetic Resonance (NMR) spectroscopy
- Cryo-electron microscopy (Cryo-EM)

Step 3: Importance of PDB.
Researchers use PDB data for:

- Drug design
- Protein structure prediction
- Studying molecular interactions
- Structural bioinformatics research

Conclusion:
Therefore, the main archive for **3D macromolecular structures** is the **Protein Data Bank (PDB)**.

Quick Tip

3D structures of proteins, DNA, and RNA → stored in **Protein Data Bank (PDB)**

---

12. What is the purpose of Emulsion PCR (ePCR) in Next-Generation Sequencing?
(A) To sequence DNA directly without amplification
(B) To amplify individual DNA fragments in isolated microreactors
(C) To detect proteins in sequencing samples
(D) To separate RNA molecules from DNA

Correct Answer: (B) To amplify individual DNA fragments in isolated microreactors

Solution:

Concept:
Next-Generation Sequencing (NGS) technologies require millions of copies of DNA fragments to generate detectable sequencing signals. One important amplification technique used in some NGS platforms (such as 454 pyrosequencing and Ion Torrent sequencing) is **Emulsion PCR (ePCR)**.

Step 1: Understanding emulsion formation.
In ePCR, DNA fragments are mixed with beads, primers, and PCR reagents in an oil-water emulsion. This mixture forms millions of tiny droplets.
Step 2: Microreactor concept.
Each droplet acts as an independent **microreactor** containing:

- A single DNA fragment
- PCR reagents
- A bead coated with primers

Step 3: DNA amplification.
PCR amplification occurs inside each droplet, producing many copies of the same DNA fragment attached to the bead. This process ensures that each bead carries identical DNA sequences.
Step 4: Role in sequencing.
After amplification, these beads are placed onto a sequencing chip where the amplified DNA fragments can be read simultaneously during the sequencing process.
Conclusion:
Therefore, the purpose of **Emulsion PCR (ePCR)** in Next-Generation Sequencing is to **amplify individual DNA fragments in isolated microreactors**.

Quick Tip

ePCR → Amplifies single DNA fragments inside microscopic droplets for NGS sequencing

---

13. Which component is strictly part of a Microprocessor System Unit, excluding peripherals?
(A) Keyboard
(B) Monitor
(C) CPU
(D) Printer

Correct Answer: (C) CPU

Solution:

Concept:
A **Microprocessor System Unit** refers to the core internal components of a computer responsible for processing data and executing instructions. It excludes external input and output devices, which are called **peripherals**.
Typical components of a system unit include:

- Central Processing Unit (CPU)
- Memory (RAM)
- Motherboard

- Internal buses

Peripherals such as keyboards, monitors, and printers are external devices connected to the system unit.
Step 1: Understanding the role of the CPU.
The **Central Processing Unit (CPU)** is the main processing component of the computer. It performs arithmetic calculations, logical operations, and controls the execution of instructions.
Step 2: Distinguishing peripherals from system unit components.
Devices like the keyboard, monitor, and printer are input/output peripherals used for interaction with the computer system.
Step 3: Selecting the correct answer.
Since the CPU is an internal component of the system unit and not a peripheral device, it is the correct answer.
Conclusion:
Therefore, the component that is strictly part of a microprocessor system unit is the **CPU**.

Quick Tip

CPU, RAM, and motherboard → System unit components
Keyboard, monitor, printer → Peripheral devices

---

14. How does a 2D array differ from a 1D array in a programming language like C++?
(A) A 2D array stores characters only
(B) A 2D array has rows and columns
(C) A 2D array stores only integers
(D) A 2D array cannot store multiple values

Correct Answer: (B) A 2D array has rows and columns

Solution:

Concept:
Arrays are data structures used to store multiple values of the same data type in contiguous memory locations.
Step 1: Understanding a 1D array.
A **one-dimensional array (1D array)** stores elements in a single linear sequence. Each element is accessed using one index.
Example in C++:

$$intarr[5] = \{1, 2, 3, 4, 5\};$$

Here each element is accessed using a single index such as arr[0], arr[1], etc.
Step 2: Understanding a 2D array.
A **two-dimensional array (2D array)** stores data in a matrix-like structure consisting of rows and columns.
Example:

$$intmatrix[3][3];$$

Elements are accessed using two indices such as:

$$matrix[row][column]$$

Step 3: Key difference.
The primary difference is the number of dimensions used to organize the data:

- 1D array → single dimension
- 2D array → two dimensions (rows and columns)

Conclusion:
Thus, a 2D array differs from a 1D array because it stores data in **rows and columns**.

Quick Tip

1D array → Linear list
2D array → Matrix structure (rows × columns)

---

15. Which type of bond characterizes the primary structure of DNA?
(A) Hydrogen bond
(B) Ionic bond
(C) Phosphodiester bond
(D) Disulfide bond

Correct Answer: (C) Phosphodiester bond

Solution:

Concept:
DNA structure consists of several hierarchical levels. The **primary structure** refers to the linear sequence of nucleotides in the DNA strand.
Step 1: Understanding DNA components.
Each DNA nucleotide contains three components:

- A phosphate group
- A deoxyribose sugar
- A nitrogenous base

Step 2: Bond connecting nucleotides.
Adjacent nucleotides in the DNA strand are linked by **phosphodiester bonds**. These bonds connect the 3' hydroxyl group of one sugar molecule to the 5' phosphate group of the next nucleotide.
Step 3: Distinguishing from other bonds.

- **Hydrogen bonds** hold complementary bases together between two DNA strands.
- **Disulfide bonds** occur in proteins.

- **Ionic bonds** are not responsible for linking DNA nucleotides.

Conclusion:
Therefore, the bond that characterizes the **primary structure of DNA** is the **phosphodiester bond**.

Quick Tip

DNA backbone = Sugar + Phosphate
Nucleotides linked by → **Phosphodiester bonds**
