

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: B.TECH
BRANCH: CS/IT/ECE/CE

SEMESTER : VII
SESSION : MO/2022

SUBJECT: BE316 BIOINFORMATICS ALGORITHMS (OE III)

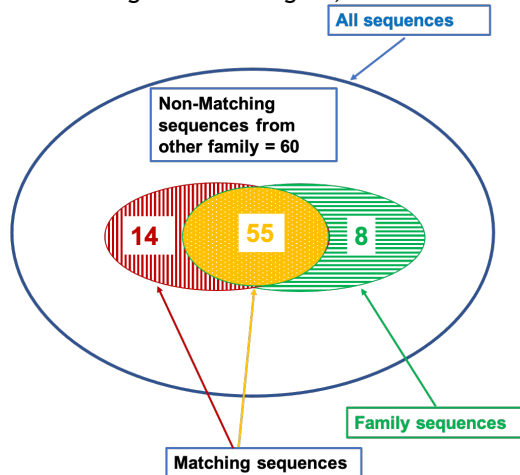
TIME: 3:00 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Before attempting the question paper, be sure that you have got the correct question paper.
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

- Q.1(a) What are the tuples and attributes in the relational database? Explain with an example. [2]
Q.1(b) Discuss the content and file formats of NCBI, PDB, and KEGG databases. [3]
Q.1(c) Using the information from various databases and bioinformatics tools, how can you contribute to the fight against COVID-19 diseases? [5]
- Q.2(a) How does binary search work? Explain with a hypothetical example of at least ten numbers. [2]
Q.2(b) Heap sort is not a stable algorithm. Demonstrate with an example. [3]
Q.2(c) Why do we need sorting and searching algorithms for bioinformatics research? Explain with suitable examples. [5]
- Q.3(a) Select the correct option of the regular expressions that would fit for the sequence EWILKDF, also explain your choice. [2]
- i) E-M-x-[ILV]-x{2}-F
ii) [EN]-W-x-[ILV]-[RKH]-x-F
iii) [EN]-W-x{2}-[ILV]-F
iv) E-W-I-[ILMV]-x-K-[FA]
- Q.3(b) Considering the below figure, calculate the sensitivity and selectivity. [3]



- Q.3(c) Considering the following sequences, calculate the scores of all the possible motifs of length 4 using Gibbs's sampler algorithm for the outlier sequence. [5]

Seq-1: ACCGTGGTGT

Seq-2: TGGCACAAGC

Seq-3: GCCGATAGTC

Seq-4: AGTGGCGAAC

Seq-5: CCTGTGGTCA

Note:

1. Consider sequence-2 as an outlier sequence and only perform the calculation for the first iteration.
2. Select the random motif of length 4 in all the sequences.

PTO

