C. Roul

COURSE CODE: 10B11BI311   MAX. MARKS: 15
COURSE NAME: BIOLOGICAL COMPUTATION
COURSE CREDITS: 04   MAX. TIME: 1Hour

Note: All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.

Q1. Each question carries 2 marks. Answer any three.   (2x3=6)

i. What were the important factors that led to emergence of Bioinformatics (BI) as separate branch? What are the functional aspects of BI?

ii. Unrelated sequences are also considered as random sequences, why? BLAST software also calculates P & E values. How does it plot extreme value distribution and determine P values from that plot? (0.5+1.5)

iii. For one protein, structure, function and all other information are available. Suppose a new protein sequence is matching significantly with above protein. What kind of information you can provide about the new sequence and how?

iv. Local alignment method with two protein sequences of length 200 each gave the alignment:

F W L E V E G N S M T A P
F W L D V Q G D SM T A G

If raw score is 65, K = 0.1 and λ=0.2
Determine the alignment is significant or not?

Q2. Each question carries 3 marks.   (3x3=9)

i. How do you implement different steps of BLAST to identify high scoring pair (HSP)? How are HSPs connected to determine larger matching regions? (2+1)

ii. What is affine gap penalty and discuss the basis of its use? Different scoring scheme may provide different alignments, why and how do you determine which alignment correct? How do you manually curate (rectify) the alignment? (1.5+1+0.5)

iii. Two protein sequences are given below:
Seq1: G A W H E G E
Seq2: P A W H A E
Use BLOSUM 62 scoring matrix and Gap penalty -8 to determine the local alignment (Smith Waterman method)?