

Note: All questions are compulsory. Marks are indicated against each question in square brackets.

- Q.1. Elucidate the analytical version of distance and similarities in biological sequence alignment and analysis. Discuss the respective role of distance and similarity used for the functional annotations of biological sequences. (CO-1,2) [3]
- Q.2. What is a spectrum? Explain with an example, how two different set of fragments can generate same spectrum? What is the significance of l -mers in spectrum analysis? (CO-2,3) [4]
- Q.3. Solve the given fragments problem through Euler's approach. Generate non-redundant fragments before solving it. $S = \{ ATG, TGC, GTG, GGC, GCA, GCG, CGT \}$ (CO-2,3) [4]
- Q.4. Calculate the coverage of a fragment assembly where 200 reads of length 550 bp were provided. The target genome (contig) length was 5K. ? For the same problem if we need 7X coverage, how many reads we will require? (CO-2-4) [3]
- Q.5. What are various complications of fragment assembly? Discuss 4 important complications with suitable examples. ? (CO-2-4) [4]
- Q.6. Elaborate the journey of graph theory based implementations to solve the problem of linearity of the genomes. ? Discuss the same through interval graphs also. (CO-2-4) [3]
- Q.7. Convert the given SSP into TSP and vice versa for a given set of fragments. Set of strings are: {000, 001, 010, 011, 100, 101, 110, 111}. Do the same for a 6 DNA based fragments set of your choice. (CO-3-4) [4]