B.Sc. Engg. (CSE) 7th Semester

30 May 2018

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)
ORGANISATION OF ISLAMIC COOPERATION (OIC)
Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION **WINTER SEMESTER, 2017-2018**

DURATION: 3 Hours **FULL MARKS: 150**

CSE 4741: Bioinformatics

Programmable calculators are not allowed. Do not write anything on the question paper.
There are **8 (eight)** questions. Answer any **6 (six)** of them. Figures in the right margin indicate marks.

---

1. a) Two biological sequences of length five and six respectively are needed to be aligned both locally and globally. It has been decided that *Smith-Waterman algorithm* will be used for local alignment and *Needleman-Wunsch algorithm* will be used for global alignment. Show Initialized scoring matrix for both of the methods. Explain how the difference(s) of the initial scoring matrices impact on the final outcomes. — 10

   b) Four RNA sequences are given as follows-

   Seq 1: ACGCATTGAATGATGATAAT
   Seq 2: ACGCGTTGGGCGATGGCAAC
   Seq 3: ACACATTGAGTGATAATAAT
   Seq 4: ACGCGTTGGGCGACGGTAAT

   i. Build distance matrix for the sequences. — 4
   ii. Using Fitch-Margoliash algorithm to build a phylogenetic tree for the sequences. — 11

2. a) Discuss the steps of translation process in *Central Dogma*. — 5
   b) Explain various theories about how multi-translation takes place from a single mRNA. — 11
   c) Write short notes on the followings- — 3x3
   Poly Adenylation, mRNA Degradation, mRNA Half-Life

3. a) Along with a graphical presentation name various types of RNA secondary structures. — 7
   b) Discuss basic principles and advantages of MFE-folding algorithm for RNA structure prediction. — 6
   c) State *Projection Algorithm* for Motif search. Define various terms used in the algorithm. — 8+4

4. a) Discuss in detail how the propensity values can be calculated for Chow-Fasman algorithms. — 8
   b) Write down the Chow-Fasman algorithms to predict protein secondary structures (alpha helix, beta strand, beta turn) for an amino acid chain. — 5x3
   c) How to resolve conflicts between outcomes of Chow-Fasman algorithms? — 2

5. a) For an experiment you are needed to calculate a proximity matrix for a set of gene clusters. Each cell M(i,j) of the proximity matrix represents similarity between cluster(i) and cluster(j). Discuss various possible attributes to calculate cluster similarity in this context. — 12
   b) Mention one of the major drawback of K-means algorithm. Discuss a method to overcome this drawback. Write down the modified K-means algorithm as per discussion. — 1+5+7

6. a) Compare divisive and agglomerative clustering methods. 6
   b) What is the importance of SSE (Sum of Squared Error) in the field of clustering? How SSE can be calculated? 5+2
   c) Classify various types of RNAs and discuss their functions. 9
   d) How RNAs are different then DNAs? 3

7. a) What is RNA editing? How it is different than gene evolution? Explain. 8
   b) *Cephalopods have something really strange in their genes* - discuss on this topic. 12
   c) What is *DNA microarray*? Discuss its importance in the field of bioinformatics. 5

8. a) What do you mean by gene expression? Discuss control regions of a gene. 2+4
   b) Discuss the mechanism of gene regulation process. 9
   c) Define Boolean Network. Explain how Boolean Network can be used to model a gene regulatory network. 10