1 Bioinformatics (PL)

(*a*) Define the Longest Common Subsequence (LCS) problem between two strings and find a solution for the case of the two strings:

ACGT and GGTTTAAGCCGT

[4 marks]

(*b*) Discuss the limitations of the Nussinov algorithm for RNA folding prediction. [4 marks]

(*c*) Which algorithm would you choose for clustering sequence and gene expression datasets and why? [3 marks]

(*d*) Discuss the complexity of the Markov Clustering (MCL) algorithm. [4 marks]

(*e*) Discuss the assumptions for the use of the Gillespie algorithm in the study of reaction networks and comment on the algorithm's complexity in terms of the number of reactions. [5 marks]