COMPUTER SCIENCE TRIPOS Part II − 2023 − Paper 8

2 Bioinformatics (pl219)

(a) Calculate the dynamic programming matrix and one or more optimal alignment(s) for the sequences `GAATTC` and `GATTA`, scoring $+2$ for a match, -1 for a mismatch and with a linear gap penalty of $d = 2$. [5 marks]

(b) Determine whether the RNA string `GGACCACCAGG` should be folded into two substructures. [7 marks]

(c) Discuss how to carry out the multiple sequence alignment of the following three sequences: `TTTTAAAA`, `AAAACCCC`, `CCCCTTTT`. [4 marks]

(d) Discuss which steps of the 1994 Adleman's DNA computation approach would particularly affect the scalability of the number of the visited cities. [4 marks]