

COMPUTER SCIENCE TRIPOS Part II – 2013 – Paper 7

3 Bioinformatics (PL)

Given the two DNA sequences: GCACTT and CCCAAT

- (a) Compute the alignment (using the edit graph) and the final score with the following rules: match score = +1, mismatch = -1, gap penalty = -1. [4 marks]
- (b) Discuss how the alignment score and the quality of the result depend on the match score, mismatch, and gap penalty. [6 marks]
- (c) Generate four, short DNA sequences (a,b,c,d) such that their relations as a tree are approximately the following: ((a,b),(c,d)). [5 marks]
- (d) How is the score matrix used in phylogenetic tree building techniques? [5 marks]