

COMPUTER SCIENCE TRIPOS Part II – 2017 – Paper 7

4 Bioinformatics (PL)

- (a) Discuss the time and space complexity of dynamic programming algorithms in the context of sequence alignment problems. [6 marks]
- (b) Describe the differences between Soft and Hard Clustering. [6 marks]
- (c) Construct the de Bruijn graph of TACCTTCAGCGCCTTC by splitting it into k -mers for a suitable value of k . Comment on how the choice of k affects the construction. Discuss the use of de Bruijn graphs in the context of genome sequencing. [8 marks]