

2009 Paper 7 Question 5

Bioinformatics

- (a) Discuss, with one example, the complexity of the Nussinov algorithm for RNA folding. [5 marks]
- (b) In the context of algorithms on strings, what is the advantage of using spaced seeds in database search? [3 marks]
- (c) Hidden Markov models (HMM) are used to identify genes in genome sequencing projects.
- (i) Describe how you would build a hidden Markov model to identify genes in a genome sequence. [7 marks]
- (ii) How would you assess the sensitivity and specificity performance of the HMM? [5 marks]