

## 2009 Paper 9 Question 3

### Bioinformatics

- (a) Compute the global alignment between the two strings  $s_1 = \text{ACCGTT}$  and  $s_2 = \text{AGTTCA}$ , considering the following scoring parameters: +1 for match, -1 for mismatch, and -1 for a gap.
- (i) What is the maximum similarity score between the two sequences  $s_1$  and  $s_2$ ? [2 marks]
  - (ii) Find an alignment with this similarity score. [2 marks]
  - (iii) Is the alignment you found unique, or are there multiple alignments achieving the maximum similarity score? [1 mark]
- (b) Discuss the complexity of the Sankoff parsimony algorithm. [4 marks]
- (c) Discuss the main differences between K-means, Superparamagnetic and Markov clustering algorithms. [7 marks]
- (d) Discuss the utility of the Gillespie algorithm in system biology. [4 marks]