2006 Paper 13 Question 10

Bioinformatics

(*a*) Hidden Markov models (HMM) are widely used in Bioinformatics.

    (*i*) In a HMM when would you use the Baum–Welch algorithm, and when the Viterbi algorithm, and why? Give biologically motivated examples.
        [8 marks]

    (*ii*) Any machine learning model (such as a HMM) for protein secondary structure determination or gene finding relies on discovering characteristic statistical properties of protein sequences. Name a property (and justify your answer) that helps to localise (and distinguish) transmembrane segments and coils in a protein sequence, or exon/intron boundaries in a genomic region.
        [2 marks]

(*b*) Discuss the complexity of an algorithm to reconstruct a genetic network from microarray perturbation data.
    [7 marks]

(*c*) What is the difference in terms of connectivity between a scale-free network and a random network? Give biological examples of scale-free networks.
    [3 marks]