

## 2005 Paper 3 Question 2

### Data Structures and Algorithms

- (a) Briefly outline how a sequence of symbols can be encoded as a sequence of Huffman codes, and explain under what assumptions Huffman encoding generates optimally compact code. [6 marks]
- (b) Estimate the number of bits needed to Huffman encode a random permutation of As, Bs and Cs, with each letter occurring one million times. [3 marks]
- (c) Estimate the number of bits needed to Huffman encode a random permutation of As, Bs and Cs, where A occurs two million times and B and C each occur one million times. [3 marks]
- (d) Estimate how many bits would be needed to encode the sequence in part (b) above using arithmetic coding. You may assume that  $\log_2 3$  is about 1.6. [4 marks]
- (e) Estimate, with justification, how many bits would be needed to encode the sequence in part (c) above using arithmetic coding. [4 marks]