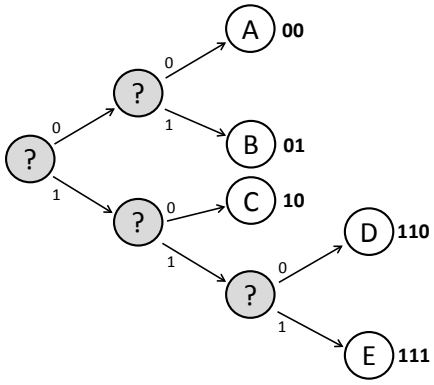
COMPUTER SCIENCE TRIPOS Part II – 2015 – Paper 9

6 Information Theory and Coding (JGD)

(a) Huffman trees enable construction of uniquely decodable prefix codes with optimal codeword lengths. The five codewords shown here for the alphabet {A,B,C,D,E} form an instantaneous prefix code.



(i) Give a probability distribution for the five letters that would result in such a tree.
(ii) Calculate the entropy of that distribution.
(iii) Compute the average codeword length for encoding this alphabet, and relate your results to the Source Coding Theorem.

[3×2 marks]

(b) What does it mean for a function to be "self-Fourier"? Show that the Gaussian function is self-Fourier. Name two other functions of importance in information theory that are self-Fourier, and in both cases mention a topic or theorem exploiting this property. [6 marks]

(c) (i) In the FFT algorithm, if a discrete data sequence consists of N sample values (nominally N is some power of 2), what complex number is the *primitive N^{th} root of unity* which, raised to various powers, generates all the complex numbers needed to perform a discrete Fourier transform? [2 marks]

(ii) If all the N^{th} roots of unity are known, by what mechanism are sequences of them selected that are needed for the k^{th} frequency component? [2 marks]

(d) Define the Kolmogorov algorithmic complexity K of a string of data. What approximate relationship is expected between K and the Shannon entropy H for the same source? Give a reasonable estimate of the Kolmogorov complexity K of a fractal, and explain why it is reasonable. [4 marks]