# 2001 Paper 7 Question 12

## Information Theory and Coding

(*a*) Consider $n$ different discrete random variables, named $X_1, X_2, \ldots, X_n$, each of which has entropy $H(X_i)$.

Suppose that random variable $X_j$ has the smallest entropy, and that random variable $X_k$ has the largest entropy.

(*i*) What is the upper bound on the joint entropy $H(X_1, X_2, \ldots, X_n)$ of all these random variables? [2 marks]

(*ii*) Under what condition will this upper bound be reached? [2 marks]

(*iii*) What is the lower bound on the joint entropy $H(X_1, X_2, \ldots, X_n)$ of all these random variables? [2 marks]

(*iv*) Under what condition will this lower bound be reached? [2 marks]

(*b*) (*i*) Define the Kolmogorov algorithmic complexity $K$ of a string of data. [2 marks]

(*ii*) What relationship is to be expected between the Kolmogorov complexity $K$ and the Shannon entropy $H$ for a given set of data? [2 marks]

(*iii*) Give a reasonable estimate of the Kolmogorov complexity $K$ of a fractal, and explain why it is reasonable. [2 marks]

(*c*) The signal-to-noise ratio $SNR$ of a continuous communication channel might be different in different parts of its frequency range. For example, the noise might be predominantly high frequency hiss, or low frequency rumble. Explain how the information capacity $C$ of a noisy continuous communication channel, whose available bandwidth spans from frequency $\omega_1$ to $\omega_2$, may be defined in terms of its signal-to-noise ratio as a function of frequency, $SNR(\omega)$. Define the bit rate for such a channel's information capacity, $C$, in bits/second, in terms of the $SNR(\omega)$ function of frequency.

[Note: This question asks you to generalise beyond the material lectured.]

[6 marks]