2009 Paper 9 Question 9

Information Theory and Coding

(*a*) Calculate the entropy in bits for each of the following random variables:

(*i*) Pixel values in an image whose possible grey values are all the integers from 0 to 255 with uniform probability. [1 mark]

(*ii*) Humans grouped by whether they are, or are not, mammals. [1 mark]

(*iii*) Gender in a tri-sexual insect population whose three genders occur with probabilities 1/4, 1/4, and 1/2. [2 marks]

(*iv*) A population of persons classified by whether they are older, or not older, than the population's median age. [1 mark]

(*b*) Let $p(x)$ and $q(x)$ be two discrete probability distributions.

(*i*) What is the *Kullback–Leibler distance* (KL) between these distributions? [2 marks]

(*ii*) If we have devised an optimally compact code for the random variable described by $q(x)$, what does the KL tell us about the effectiveness of our code if the probability distribution is $p(x)$ instead of $q(x)$? [1 mark]

(*iii*) Which axiom of distance metrics is violated by this distance? [1 mark]

(*iv*) What happens to this metric if there are some forbidden values of x for which $p(x) = 0$, and other values of x for which $q(x) = 0$? [1 mark]

(*c*) Consider a continuous real signal whose bandwidth extends from 0 to 1 KHz. We wish to represent a 1.0 second interval of it exactly, using just a finite list of numbers obtained by sampling the signal at discrete, periodic, points in time.

(*i*) What is the length of the shortest list of such discrete samples needed to reconstruct this interval of the signal from them completely? [2 marks]

(*ii*) Name, define, and sketch a plot of the function used to reconstruct the continuous signal from its samples, by superimposing shifted copies of this function weighted by the discrete samples. [3 marks]

(*d*) Explain why data can be compressed by encoding it into transforms (such as the DCT, Fourier or Gabor) that result in coefficients that have a more narrow, peaked, distribution than the original data. Without going into details about particular transforms, explain why the coefficients obtained have distributions with less entropy than the original signal or image, and why this enables compression. [5 marks]