GI12/4C59 – Homework #2 (Due 12am, November 8, 2005)

Aim: To get familiarity with the basic concepts of Information Theory (entropy, mutual information, etc.) and some coding principles. Presentation, clarity, and synthesis of exposition will be taken into account in the assessment of these exercises. This document is available from
http://www.cs.ucl.ac.uk/staff/Rob.Smith/Internal/information_theory.htm

1. **[30 pts]** Let X and Y be two random variables with values in the sets $X = \{0,1,2\}$ and $Y = \{0,1\}$ respectively. Define the probability distribution p on $X \times Y$ by the table

$$p = \begin{bmatrix} & X=0 & X=1 & X=2 \\ Y=0 & \frac{1}{3} & \frac{1}{12} & \frac{1}{6} \\ Y=1 & \frac{1}{12} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

   (a) Compute the joint entropy of X and Y, $H(X,Y)$.
   (b) Find the marginal distribution of X and the conditional distribution of Y given X. Then, use these quantities to compute the entropy of X, $H(X)$ and the relative entropy of Y given X, $H(Y|X)$.
   (c) Verify the entropy results above by using the chain rule which relates $H(X,Y)$ to $H(X)$ and $H(Y|X)$.
   (d) Compute the mutual information between X and Y.

2. **[30 pts]** This one might require you to use a computer, but only in the most basic way. The frequency p_n of the n^{th} most frequent word in English (and in many other languages) is roughly approximated by:

$$p_n \approx \begin{cases} \frac{0.1}{n} & \text{for } n \in 1...12,367 \\ 0 & n > 12,367 \end{cases}$$

(this is known as Zipf's Law). Assuming English is generated by selecting words at random from this distribution, what is the information entropy of English?

3. **[10 pts]** The following strings are received in a (7,4) Hamming code. Decode, please:
   (a) **r = 1101011**
   (b) **r = 0110110**
   (c) **r = 0100111**
   (d) **r = 1111111**

4. **[30 pts]** Imagine constructing a Huffman code for symbols made up of blocks of n bits. We will call each of these codes X^n. If the probability of 0 is 0.8, and the probability of 1 is 0.2, determine optimal length Huffman codes X^2 and X^3. Calculate the expected length of the codewords, and the information entropy for the codes. Please show your work.