Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.450 Principles of Digital Communication

Problem Set 4

Fall, 2006

Issued: Oct. 4 Due: Oct. 18

Note: Quiz 1 given in class on Wed. Oct. 11

Problem 2.27, 2.28, 3.1, 3.2, 3.3, 3.4

Extra Problem: (I almost used this as a quiz problem)

Let's start by defining some notations. Consider a prefix free code code for a DMS with distribution P_X, and consider the tree representation of the code.

For a node k on the coding tree, (excluding the leaf nodes), let p_k be the probability that the node is reached. That is, p_k is the sum of the probabilities of all the source letters that are assigned to node k and/or all its descendants. Let m and n be the two children of node k. Clearly $p_k = p_m + p_n$.

Now write $h_k = H_b(p_m/p_k)$, where H_b is the binary entropy, $H_b(a) = -a \log a - (1 - a) \log(1 - a)$.

(a) Show that the source entropy

$$H(X) = \sum_k p_k h_k$$

(b) Show that the average codeword length (bit/source symbol) is

$$E[L] = \sum_k p_k$$

(c) Now define local redundancy at node k as $r_k := 1 - h_k$, we have

$$E[L] - H(X) = \sum_k p_k \cdot r_k$$

Give an intuitive interpretation of this.

(d) Now consider the old game of weighing balls. Let there be 12 balls, all identical except one of them being either heavier or lighter. You are given a pan balance and 3 chances to use it to figure out the bad ball. Now suppose I start to use the balance the first time as follows. I uniformly pick 4 balls out of the 12, then divide into two groups and compare them. What is the average amount of information I would get when I see the result?

Argue using the above parts that this way I can not always figure out the bad ball in with 3 uses of the balance. (hint: the constraint that I have to always use the balance for no more than 3 time is much more stringent than an average number of uses constraint. The above statement is true even if I have a constraint on the average number of uses.)