ECE 154C

Homework #1
Due: Wednesday, April 8, 2009

1. Consider a source which produces an i.i.d. sequence of symbols from the alphabet {A,B,C} with probabilities {0.4, 0.3, 0.3} respectively. For n=1, 2, and 3, find binary Huffman codes for taking n source symbols at a time. In each case compute the average number of binary code symbols per source symbol and compare it to the entropy (base 2).

2. Repeat problem 1, except find Shannon Fano codes instead of Huffman codes.

3. The purpose of this problem is to see what happens when you design a code for the wrong set of probabilities. Assume that you design a Huffman code for the second extension of a source where the probabilities of the source letters are {0.6, 0.2, 0.1, 0.1} but that the actual source probabilities are {0.1, 0.1 0.2, 0.6}. Find the average number of binary code symbols per source symbol and compare it with the entropy of the source.

4. Denote the n-th extension of the i.i.d. source S as S^n. S^n is a new an i.i.d. source that produces n letters from the original source S at a time. Prove that the entropy of S^n, $H(S^n)$, is equal to n H(S).

5. Simulate the system described in problem 1 with n=2. That is first simulate a source that produces source letters with probabilities {0.4, 0.3, 0.3}, then encode them two at a time using the Huffman code found in problem 1 for n=2. Encode 1000 such pairs and count the number of binary code digits required to encode. Then compute the ratio of the number of binary code digits to the number of source letters (which should be 2000. Compare your answer with the theoretical answer found in problem 1.