Algorithms and Applications (15-499)
Assignment #1a Due Date: 28 Jan 03

Problem 1: (10pt) Given the following conditional probabilities for a two state Markov Chain what factor would one save by using the conditional entropy instead of the unconditional entropy?

$$p(w|w) = .9 \quad p(b|w) = .1$$
$$p(w|b) = .2 \quad p(b|b) = .8$$

Problem 2: (10pt) Devise a uniquely decipherable code that is not a prefix code.

Problem 3: (10pt) Given the following probability model:

Letter	$p(a_i)$	$f(a_i)$
a	.1	0
b	.2	.1
c	.7	.3

Decode the 4 letter message given by `01001110110` assuming it was coded using arithmetic coding. Why is this message longer than if we simply had used a fixed-length code of 2 bits per letter, even though the entropy of the set $\{.1, .2, .7\}$ is just a little more than 1 bit per letter. Note: once you figure out how to do the decoding, it should not take more than five minutes on a calculator or scripting language.

Problem 4: (20pt)

A. Prove the first part of the Kraft-McMillan inequality for Prefix Codes. In particular show that for any prefix code C,

$$\sum_{(s,w)\in C} 2^{-l(w)} \leq 1.$$

Hint: think of the prefix code as a binary tree (as discussed in class) and try some form of induction. You will be graded on the clarity as well as the correctness of your proof. Remember that making the proof longer does not necessarily increase the information content.

B. Prove that if you have $n = 2^k$ codewords in a prefix code and that if one of them is shorter than k bits, then at least two must be longer than k bits.