CSC310 - Fall 2007
Assignment 2

due on Tuesday, November 13th.

For all questions below, show your work. Do not just write down numbers. Do not show every single addition, but show enough work to convince the reader you know what you are doing.

Problem 1 [10p]

Consider the following encoding scheme for bit strings of length $M + 2$. Given such a string x, send the first two bits without modification. Compute the counts of 0-s and 1-s within these two bits, and use the Laplace model for estimating probabilities for the entire remaining portion of the string (i.e. these probabilities are computed after the first two bits and they do not change any more). Encode the rest of the string using arithmetic coding with those estimated probabilities.

Assume the initial string is generated by flipping bits independently of each other with probability .6 of getting a 1. Assume $M = 100$.

What is the entropy of the source string? What is the average encoding length achieved by this scheme?

Problem 2 [15p]

Consider the following Binary Symmetric/Erasure Channel (BSEC), generalizing both BSC and BEC: it has parameters p and f, the input alphabet is $\{0, 1\}$, the output alphabet is $\{0, 1, ?\}$, and each time a bit (either 0 or 1) is sent, the probability of being transmitted correctly is p, the probability of being flipped is $f(1-p)$, and the probability of being erased is $(1-f)(1-p)$. Let X denote the input bit and Y denote the output bit. Assume $p_0 = \Pr[X = 0] = .6$, $p = .8$, and $f = .4$.

Compute the channel transmission probabilities $Q_{j|i}$. Compute the receiver probabilities $q_j = \Pr[Y = j]$. Compute $H(X), H(Y), H(X, Y), I(X; Y), H(X|Y), H(Y|X)$.

Compute the capacity of this channel (for this, you might want to consider other p_0-s).

Problem 3 [10p]

Consider the BSC channel with $f = .3$ and the simple repetition code $\{00000, 11111\}$ for its 5-th extension. For what values of $p_0 = \Pr[w = 00000]$ should 01101 be decoded as 00000, and 10111 be decoded as 11111? What is the smallest p_0 for which 11111 should be decoded as 00000?

Problem 4 [20p]

Consider the Z channel with $f = .2$ and the code $C = \{0101, 0110, 1001, 1010\}$ for its 4th extension. Since it has 4 codewords, it can be used to transmit blocks of 2 message bits. Assume the probabilities of sending each codeword are equal. We saw that under this assumption, the optimal decoder is by maximum likelihood.

Devise such a maximum likelihood decoder for this code: for every possible received block, you must decide what codeword this decoder should output. Note that, in certain cases, even the optimal decoder has to make a choice.

Compute the probability that the optimal decoder you devised above is wrong, i.e., it produces a codeword different than the one sent. Note that in those cases where the optimal decoder has to make a choice, sometimes its choice is right.

Compute the mutual information between the string sent and the first bit of the string received.