ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
Communication Systems Department

Handout 28 Information Theory and Coding
Final Exam February 20, 2001

Problem 1. (30 points.) We design a Huffman code for a source U, where $\Pr(u)$ for each source letter is an integer power of 1/2.

(a) (6 points) Show that the Huffman code will have an expected codeword length that is identical to the entropy of the source. [Hint: there is no need to run the Huffman algorithm, find a set of codeword lengths that meets the Entropy bound and satisfies the Kraft inequality.]

Consider a discrete stationary source, $\ldots, X_{-1}, X_0, X_1, X_2, \ldots$. Assume for simplicity that all the probabilities of interest are integer powers of 1/2. We consider two alternative source coding methods:

METHOD 1: We take consecutive pairs $\ldots$, (X_1, X_2), (X_3, X_4),$\ldots$, (X_{2k-1}, X_{2k}),$\ldots$, and encode each pair using the Huffman code for the pair.

METHOD 2: We encode each X_k using a Huffman code based on the conditional probability of X_k given the value of the preceding source letter X_{k-1}. (Note that one needs, in general, $|\mathcal{X}|$ different Huffman codes, one per each possible preceding letter.)

(b) (8 points) Find the expected codeword length per letter for method 1.

(c) (8 points) Find the expected codeword length per letter for method 2.

(d) (8 points) Which of the two methods is better?

Problem 2. (35 points) Consider an additive noise channel $Y = X + Z$ where X is the input of the channel, Y is the output of the channel and Z is noise. The set of inputs to the channel are *non-negative* real numbers. Furthermore, the channel input is contrained in its average value: a codeword $\mathbf{x} = (x_1, \ldots, x_n)$ has to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leq P.$$

This translates to a constraint $E[X] \leq P$ when we are trying to maximize the mutual information $I(X;Y)$. The noise Z is independent of the input X, and has the exponential distribution with $E[Z] = 1$, i.e.,

$$f_Z(z) = \begin{cases} \exp(-z) & z \geq 0 \\ 0 & \text{else.} \end{cases}$$

(a) (5 points) The capacity of this channel is given by

$$C = \max_{\substack{X:E[X]\leq P \\ X \text{ is non-negative}}} I(X;Y).$$

Express the mutual information in terms of the differential entropy of Y and the differential entropy of Z.

(b) (5 points) What is the differential entropy of Z?

(c) (4 points) For a random variable X that satisfies the input constraints, what are the constraints on the range and the expectation of Y?

(d) (8 points) Find the maximum possible differential entropy of Y subject to the constraints you found in part (c). Show that the capacity is upper bounded by

$$C \leq \log(1 + P).$$

(e) (8 points) Show that if X has density

$$f_X(x) = \begin{cases} \mu\delta(x) + (1-\mu)\mu e^{-x\mu} & x \geq 0 \\ 0 & \text{else,} \end{cases}$$

then $Y = X + Z$ is a non-negative random variable with density $f_Y(y) = \mu e^{-y\mu}$, $y \geq 0$. [Recall if A and B are two independent random variables with densities f_A and f_B, the density of their a sum C is given by $f_C(c) = \int f_A(a) f_B(c-a)\, da$.]

(f) (5 points) Conclude that the upper bound of part (d) is actually an equality, i.e.,

$$C = \log(1 + P).$$

Problem 3. (35 points) Consider a discrete memoryless channel with input alphabet $\mathcal{X} = \{0,1\}$, output alphabet $\mathcal{Y} = \{0,1\}$ and $Y = X + Z \bmod 2$, where Z is independent of X and

$$\Pr(Z=0) = 1-\epsilon, \quad \Pr(Z=1) = \epsilon.$$

Assume that $0 \le \epsilon < 1/2$.

(a) (5 points) What is the capacity of this channel and what is the input distribution with which this capacity is achieved.

(b) (8 points) Consider a code with only two codewords, $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{w} = (v_1, \dots, v_n)$, and let d be the number of places $\mathbf{u}$ and $\mathbf{w}$ differ. Suppose that the decoder is the minimum distance decoder: if $\mathbf{y}$ is the received sequence, it finds out which of $\mathbf{u}$ and $\mathbf{w}$ is closest to $\mathbf{y}$ in the Hamming distance, and declares that one. If $\mathbf{y}$ is equally close to $\mathbf{u}$ and $\mathbf{w}$, we assume that the decoder makes an error. Recall that the Hamming distance between two sequences is the number or places they differ.

Suppose that $\mathbf{u}$ and $\mathbf{w}$ differ in their *first* d positions. Show that

$$\Pr(\text{Error}) = \Pr\left(\sum_{i=1}^{d} Z_i \ge d/2\right).$$

where $\mathbf{Z} = (Z_1, \dots, Z_d)$ is the first d elements of the channel noise sequence.

(c) (10 points) The Chernoff bound says that for a sequence of i.i.d. random variables $Z_1, Z_2, \dots,$

$$\Pr\left(\sum_{i=1}^{d} Z_i \ge d\alpha\right) \le \left(e^{-s\alpha} E[e^{sZ_1}]\right)^d,$$

for any $s > 0$. Use this bound with $e^s = (1-\epsilon)/\epsilon$ to show that

$$\Pr(\text{Error}) \le \left[4\epsilon(1-\epsilon)\right]^{d/2}.$$

(d) (8 points) Assuming that d is odd, show that

$$\Pr(\text{Error}) \ge \binom{d}{(d+1)/2} \epsilon^{(d+1)/2} (1-\epsilon)^{(d-1)/2}.$$

(e) (4 points) Assuming that the codewords are drawn randomly with the codewords chosen independently, and each entry of the codewords chosen independently with the distribution found in (a), what is the expected value of d?