ECE 745 - Advanced Communication Theory, Spring 2007

Final Exam

Take-Home (24 Hours)

Procedure

1. Download the exam from the course website on Thursday or Friday (anytime). **Send me an e-mail when you download the exam telling me that you have done such.**
2. **There are seven problems for 140 points.**
3. For the exam, you may use the course textbook *Elements of Information Theory,* by Cover and Thomas, and your course notes. For anything from the text, you must establish any missing steps not done in the course notes. For example, if you use the result of a theorem or homework problem, you must provide all of the steps between the course notes and the result you want to use. **No other references are allowed, including the WWW. You may not consult with anybody else for any reason - even for an issue of clarification. Instead, send me e-mail.**
4. Within 24 hours of picking up the exam, return it to one of the two following places: (1) If the Marcus 215 complex is open, slide it under the door of my office (Marcus 215H). (2) If the Marcus 215 complex is not open, place it in an envelope with "Goeckel" written prominently on the outside. Slide the envelope under the door of the "ECE Mailroom" next to Marcus 210.

Testmanship

- **Full credit will be given only to fully justified answers.**
- Giving the steps along the way to the answer will not only earn full credit but also maximize the partial credit should you stumble or get stuck. If you get stuck, attempt to neatly define your approach to the problem and why you are stuck.
- If part of a problem depends on a previous part that you are unable to solve, explain the method for doing the current part, and, if possible, give the answer in terms of the quantities of the previous part that you are unable to obtain.
- Start each problem on a new page. Not only will this facilitate grading but also make it easier for you to jump back and forth between problems.
- If you get to the end of the problem and realize that your answer must be wrong (e.g. the Fourier Transform of a real signal that is not conjugate symmetric), be sure to write "this must be wrong because . . . " so that I will know you recognized such a fact.
- Academic dishonesty **will** be dealt with **harshly -** the *minimum penalty* will be an "F" for the course.

1. [20] Consider an independent and identically distributed (IID) sequence (X_i), where each X_i is drawn from the alphabet $\mathcal{X} = \{A, B, C\}$ and the probability mass function of each of the X_i is given by:

$$p_{X_i}(x) = \begin{cases} 0.6, & x = A \\ 0.3, & x = B \\ 0.1, & x = C \end{cases}.$$

(a) Design a Huffman code that takes blocks of length 2 characters (i.e. $N = 2$) and find its rate (in output bits per input character). Show that the rate falls between easily obtained upper and lower bounds based on the source entropy $H(X)$.

(b) Without finding the Huffman code (this would take far too much time!), find lower and upper bounds to the rate (in output bits per input character) of a Huffman code that takes symbols eight at a time (i.e. $N = 8$). Full credit goes to the tightest bounds.

2. [20] We know from class that the rate R_{HUFF} of a Huffman code that takes input blocks of $N = 1$ characters can be upper bounded as $R_{HUFF} \leq H(X) + 1$, where $H(X)$ is the first-order entropy of the source. Recall that this bound was obtained by considering Shannon-Fano codes. Note that such a bound does not require precise knowledge of the source distribution $\{p_1, p_2, \ldots, p_M\}$, where p_i is the probability of the i^{th} character, and M is the number of characters in the source alphabet. However, suppose we know the probability p_1 of the first character, and define a new lossless code with codeword lengths $\{l_1, l_2, \ldots, l_M\}$ as follows:

$$l_j = \begin{cases} \lceil -\log_2\, p_1 \rceil, & j = 1 \\ \left\lceil -\log_2\, \left(p_j \left(\frac{1-2^{-\lceil -\log_2\, p_1 \rceil}}{1-p_1}\right)\right)\right\rceil, & j > 1 \end{cases}$$

Note: Perhaps this is obvious, but the numerator of the innermost term of the second line is $1 - 2^{-\lceil -\log_2\, p_1 \rceil}$ (i.e. the $-\lceil -\log_2\, p_1 \rceil$ is all in the exponent of 2).

(a) Show that $1 \leq l_j \leq \lceil -\log_2\, p_j \rceil$ for all j.

(b) Show that $\{l_1, l_2, \ldots, l_M\}$ satisfies the Kraft inequality.

(c) Obtain an upper bound on R_{HUFF} in terms of $H(X)$ and p_1 that is tighter than $H(X) + 1$.

(d) Making the obvious extension to $N = 2$, find this upper bound for an $N = 2$ Huffman code on the source of Problem 1 (note that you can pick any block to be your "p_1" block).

3. [20] **(Note: You can do parts (b), (c), and (d) without doing part (a).)** On your first job after graduate school, you encounter a stationary binary source. Your boss tells you that, because of the source physics, it never outputs two 0's in a row. She also gives you the following sample output:

$$(1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,$$
$$1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,$$
$$1,1,1,0,1,1,1,1,1,1,1,1,0,1,1,1,1,1,0)$$

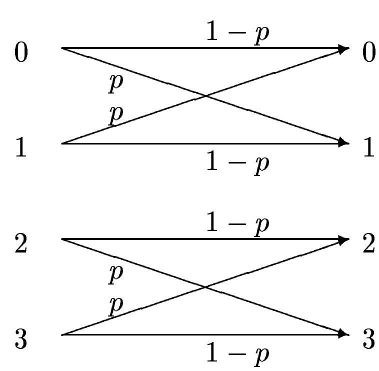
(a) Estimate the entropy rate of this source. *Note: You are going to need to make some assumptions to get an answer. Be clear about your assumptions.*

(b) Consider a lossless coding technique that simply waits for the next "0" and, when it sees the "0", sends some binary representation for the number of "1's" before that zero. For the above, it would send some binary representation of 7, 11, 4, 14, etc. Noting that the number of "1's" between successive zeroes could be arbitrarily large (i.e. it is not bounded), give an explicit lossless encoding and decoding method for an efficient version of this technique; in essence, how do you encode 7, 11, 4, 14, etc into binary?

(c) Use your scheme from (b) to encode the string of 0's and 1's above; in other words, provide the string of 0's and 1's that represent the encoder output for the string above.

(d) Will your scheme approach the entropy rate of this source for large block lengths?

4. [20] Consider the following channel:



(a) Find the capacity of this channel.

(b) Obviously, from the channel coding theorem, for $R < C$ we can build a good codebook for the channel above with 2^{NR} codewords of length N using random coding and the optimal input distribution $p_X(x)$ over $\{0,1,2,3\}$. But, suppose that I would like to use a smaller codebook, if possible. For $R < C$, give a codebook, encoding rule, and decoding rule such that a codebook of size at most $2^{N\frac{R}{2}}$ can be used on the channel above to achieve arbitrarily small error probability at rate R. *Hints: (1) The codebook entries need not necessarily be from* $\{0,1,2,3\}$ *- you may want to put some mapping from the codebook output to the channel input. (2) Think about how the capacity of this channel compares to that of the binary symmetric channel (BSC). Where does the difference come from?*

5. [20] You are communicating on a channel with additive Gaussian noise:

$$Y_i = X_i + N_i$$

where X_i is the transmitted signal, Y_i is the received signal and N_i is the noise. The interesting thing about this channel is that it can be in one of two states: "Good" or "Bad" (e.g. a wireless channel). When the channel is in a "Good" state, $N_i \sim N(0,1)$ and, when the channel is in a "Bad" state, $N_i \sim N(0,6)$. In any case, the noise is independent across i.

The state of the channel at time i is given by a stationary random process S_i, where the state at time i depends only on the state at time $i-1$. Its transitions are guided by:

$$\begin{aligned} P(S_i = \text{"Good"} | S_{i-1} = \text{"Bad"}) &= 0.45 \\ P(S_i = \text{"Bad"} | S_{i-1} = \text{"Bad"}) &= 0.55 \\ P(S_i = \text{"Good"} | S_{i-1} = \text{"Good"}) &= 0.7 \\ P(S_i = \text{"Bad"} | S_{i-1} = \text{"Good"}) &= 0.3 \end{aligned}$$

Both the transmitter and the receiver know the current channel state S_i, so the generation of a given symbol and its processing at the receiver can be based on the current state of the channel. For example (and this is intentionally a communications systems example rather than an information theory example, so do not let it confuse you), the transmitter could send a bit with +1 (bit=0) or -1 (bit=1) when the channel is "Good", and send a bit with +4 (bit=0) or -1 (bit=1) when the channel is "Bad". Because it knows the channel state, the receiver would know the corresponding mapping for any given symbol and be able to perform optimal decoding of each bit.

Now, turn your mind back to information theory (e.g. Gaussian codebooks, etc.). Given an average transmit power constraint $E[X_i^2] < 10$ (meaning that the power averaged over all i must be less than 10), provide a scheme that achieves as large an average rate as possible on this channel. Be sure to provide the details of all codebook constructions, powers employed, etc.

6. [20] Two physically-separated (i.e. joint encoding is not allowed) binary sources X and Y with joint statistics:

$$p_{X,Y}(x,y) = \begin{cases} 0.6, & x=1, y=1 \\ 0.2, & x=0, y=0 \\ 0.1, & x=0, y=1 \\ 0.1, & x=1, y=0 \end{cases}$$

generate a new pair of observations (X_i, Y_i) once per second. Each pair of observations is independent of any other pair (i.e. the sequence is independent in i). These pairs are to be transmitted over a discrete-time Gaussian multiple access channel (see sketch below) once per second; that is, the received signal for the i^{th} second is given by:

$$R_i = U_i + V_i + N_i$$

where U_i is the transmitted signal of average power $P_X = E[U_i^2]$ from source X, V_i is the transmitted signal of average power $P_Y = E[V_i^2]$ from source Y, and $N_i \sim N(0,1)$. Find the minimum total average power $P = P_X + P_Y$ for which the sources can be transmitted with arbitrarily small probability of error.