332:542 **Information Theory and Coding** **Examination 1**

March 8, 2002

This is an 80 minute exam. You may have an additional 100 minutes to answer the following five questions in the notebooks provided. You are permitted 2 double-sided sheet of notes. Make sure that you have included your name, ID number (last 4 digits only) and signature in each book used (*5 points*). Read each question carefully. All statements must be justified. Computations should be simplified as much as possible.

1. *20 points* Let $Y = g(X)$ be deterministic function of discrete random variable X.
   (a) *5 points* Give an example of a random variable X and a function $Y = g(X)$ such that $H(Y) < H(X)$.
   (b) *5 points* Give an example of a random variable X and a function $Y = g(X)$ such that $H(Y) = H(X)$
   (c) *10 points* Either give an example of a random variable X and a function $Y = g(X)$ such that $H(Y) \geq H(X)$ or prove that $H(Y) \leq H(X)$. Hint: look at $H(X, Y)$.
2. *20 points* Let be random variables such that
$$X \to Y \to Z, \qquad Y \to Z \to X, \qquad Z \to X \to Y$$
   If $I(X; Y) = 3$, find $I(X; Z)$ and $I(Y; Z)$. Or, if these quantities cannot be known, find tight upper and lower bounds. Make sure to justify your answers.
3. *15 points* Consider the code $\{0, 01\}$. Justify your answers to the following questions:
   (a) *5 points* Is the code instantaneous?
   (b) *5 points* Is the code nonsingular?
   (c) *5 points* Is the code uniquely decodable?
4. *10 points* The source coding theorem shows that the optimal source code for random variable X has expected length $L \leq H(X) + 1$. Find an example of a random variable X for which the optimal code has $L > H(X) + 1 - \epsilon$ for any small $\epsilon > 0$. Make sure you justify your answer.
5. *30 points* A source has an alphabet of 4 letters, a_1, a_2, a_3, a_4 with probabilities $p_1 \geq p_2 \geq p_3 \geq p_4$.
   (a) Suppose $p_1 > p_2 = p_3 = p_4$. Find the smallest number q such that $p_1 > q$ implies $n_1 = 1$, where n_1 is the length of the code word for a_1 in a binary Huffman code for the source.
   (b) Show by example that if $p_1 = q$ (your answer in part (a)), then a Huffman code exists with $n_1 > 1$.
   (c) Now assume the more general condition $p_1 > p_2 \geq p_3 \geq p_4$. Does $p_1 > q$ still imply that $n_1 = 1$? Why or why not?
   (d) Now assume that the source has an arbitrary number K of letters, with $p_1 > p_2 \geq \cdots \geq p_K$. Does $p_1 > q$ now imply that $n_1 = 1$? Explain.
   (e) *20 points* Now assume the source has K letters $a_1, \ldots, a_K$, with $p_1 \geq p_2 \geq \cdots \geq p_K$. Find the largest number q' such that $p_1 < q'$ implies that $n_1 > 1$.

Announcement: Class on Thursday March 13 will be canceled. Instead, we will have class on Thursday March 20 (during Spring break) at the usual hour.