University of Toronto
F. R. Kschischang

Department of Electrical
& Computer Engineering

ECE1502F — Information Theory

Final Examination

December 13, 2000

Instructions

You have approximately 2 hours of "in-class" time, followed by five days of "take-home" time to complete this test. Complete as much as possible during the in-class time; your grade will be computed as a weighted average of your "in-class" grade and your "take-home" grade. (Weights to be determined later.) Answer **all** five [5] questions. All questions have equal value. Show all steps and present all results clearly. Take-home due date: **Monday, December 18, 2000, 1:00 p.m.**, or earlier. Please hand in to the instructor (SY505) or to his assistant Doreen Lowe (SY 5th floor reception). All work is to be done independently. Consultation with others is **not** permitted. Good luck!

---

1. **Short Snappers**—the parts of this question all have short answers that require a minimum of computation. In all cases, justify your answer briefly.

   (a) What is $I(X;X)$?

   (b) True or false?: The capacity achieving input distribution for a discrete memoryless channel makes the output distribution uniform.

   (c) Find the capacity of the channel with channel transition matrix

   $$M = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}.$$

   (d) In the binary erasure channel, let the probability of erasure be ϵ. For which values of ϵ (if any) is this channel weakly symmetric?

   (e) The useful bandwidth W in telephone line modems is roughly 3000 Hz; the signal-to-noise ratio $P/(N_0W)$ is on the order of 20 to 30 dB. Estimate the channel capacity of a telephone line channel, assuming a telephone line can be modeled as an ideal band limited additive white Gaussian noise channel.

   (f) True or false?: Binary block codes that are constrained so that each codeword has an even number of ones can achieve the capacity of any binary symmetric channel.

   (g) True or false?: Binary block codes that are constrained so that each codeword has no more than three consecutive ones can achieve the capacity of any binary symmetric channel.

2. **A Phase-Noise Channel**—A discrete time channel channel has the set of phase angles, $\phi \in [-\pi, \pi)$, as input alphabet and output alphabet. The channel is subject to additive phase noise z, where z is independent of the input x and has a probability density which is nonzero only for $z \in [-\pi, \pi)$. The channel output y is the sum $x+z$, reduced modulo 2π to fall within the range $[-\pi, \pi)$. For example, if $x = 3\pi/4$ and $z = 3\pi/4$, then $y = 3\pi/4+3\pi/4-2\pi = -\pi/2$.

   (a) Show that the channel capacity C is achieved by a uniform input probability density $p_X(x) = \frac{1}{2\pi}$, $x \in [-\pi, \pi)$.

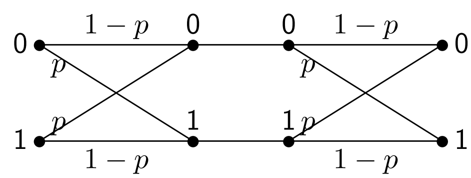
   (b) Find the channel capacity when the phase noise z is uniform in $[-\pi/M, \pi/M)$, where M is an integer greater than one, and suggest an efficient coding scheme for this channel.

   (c) Find the channel capacity when the phase noise z has probability density function

$$f_Z(z) = \frac{1}{2}\left(\frac{\alpha}{1-e^{-\alpha\pi}}\right) e^{-\alpha|z|}, \quad z \in [-\pi, \pi).$$

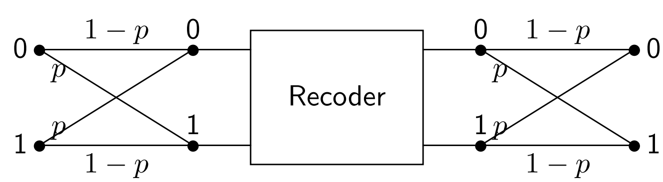
3. **Composite Channels**—Suppose that two channels are to be connected end to end to form a composite channel. Each of the original channels is a binary symmetric channel with crossover probability p.

   (a) Suppose that the output of the first channel is connected directly to the input of the second, with no processing between:



   What is the capacity, C_{cascade}, of the composite channel?

   (b) Suppose that a decoder/encoder (a recoder) is allowed between the channels:



   Find the capacity, C_{recode}, under this arrangement, and **prove** that all rates less than C_{recode} are achievable. (The starting point in your proof could be the fact that all rates less than the capacity of the subchannels is achievable in each subchannel.)

   (c) Suppose, instead of having crossover probability p, the channel following the recoder had crossover probability q. What would the capacity, C_{recode}, be now?

4. **Entropy of Functions of Random Variables**—Let X be a discrete random variable and let $Y = f(X)$ be a deterministic function of X.

(a) Show that $H(X) = H(Y) + H(X|Y)$. Under what condition is $H(X) = H(Y)$?

(b) Let X have probability mass function with values $(p_1, p_2, \ldots, p_m)$. Devise an appropriate function f to prove the grouping property of the entropy function:

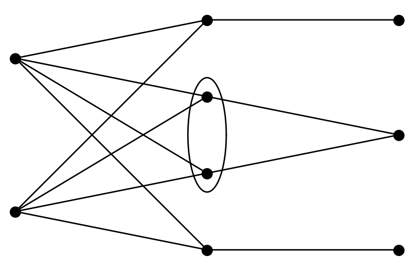
$$\begin{aligned} H(p_1, p_2, \ldots, p_m) = H(A, B) &+ AH(p_1/A, p_2/A, \ldots, p_r/A) \\ &+ BH(p_{r+1}/B, p_{r+2}/B, \ldots, p_m/B) \end{aligned}$$

where $1 \le r < m$, and

$$A = \sum_{i=1}^{r} p_i, \qquad B = \sum_{i=r+1}^{m} p_i.$$

(c) Let $Y = X_1 + X_2$, where X_1 and X_2 are discrete random variables. Under what conditions is $H(Y) = H(X_1) + H(X_2)$? (Give an example where this condition holds when X_1 and X_2 each take on two possible values.)

5. **Channel Reduction**—When output letters of a discrete memoryless channel are combined into a single letter, a *reduced channel* is said to result. Reduction is equivalent to processing the channel output by a "deterministic channel," as illustrated below, where a 2-input, 4-output channel is reduced to a 2-input 3-output channel by combining the circled outputs into a single output letter.



(a) By the data processing inequality, reduction clearly cannot increase channel capacity, and may lessen it. Prove that, if the columns of the channel matrix that correspond to the letters to be combined into a single letter are *proportional* to one another, then capacity is not reduced.

(b) Use the result of (a) to find the capacity of the discrete memoryless channel with transition matrix

$$\begin{bmatrix} 1/4 & 1/8 & 1/8 & 1/8 & 1/16 & 1/8 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 & 1/16 & 1/16 & 1/4 & 1/16 & 1/16 \\ 1/8 & 1/8 & 1/4 & 1/16 & 1/8 & 1/8 & 1/16 & 1/8 \end{bmatrix}.$$