University of Toronto
F. R. Kschischang

Department of Electrical
& Computer Engineering

ECE1502F — Information Theory
Final Examination
December 11, 2007

Instructions

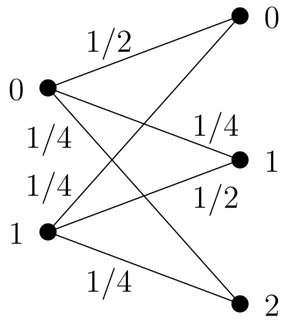
You have approximately 2 hours of "in-class" time, followed by three days of "take-home" time to complete this test. Complete as much as possible during the in-class time; your grade will be computed as a weighted average of your "in-class" grade and your "take-home" grade. (Weights to be determined later.) Answer **all** five [5] questions. All questions have equal value. Show all steps and present all results clearly. Take-home due date: **Friday, December 14, 2007, 10:30 a.m.**, or earlier. Please hand in to the instructor (BA4132) or to the teaching assistant Benjamin Smith (BA4165). All work is to be done independently. Consultation with others is **not** permitted. Good luck!

---

1. **Short Snappers**—the parts of this question are independent. In all cases, justify your answer briefly.
   (a) If $I(X;Y) = 0$ and $I(X;Z) = 0$, is $I(X;Y,Z) = 0$?
   (b) Prove or find a counterexample: $I(X;Y) = 0$ implies $I(X;Y|Z) = 0$ for any random variable Z.
   (c) True or false? Among all densities over the non-negative reals having mean $m > 0$, the exponential density $f(x) = (1/m)e^{-x/m}$ has maximum differential entropy.
   (d) Find a quaternary Huffman code (i.e., a Huffman code over the alphabet $\{0, 1, 2, 3\}$) for the random variable X that takes values uniformly in the set $\{0, 1, 2, 3, 4, 5, 6, 7\}$.
   (e) Find the capacity of the ternary-input ternary-output discrete memoryless channel with transition matrix

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

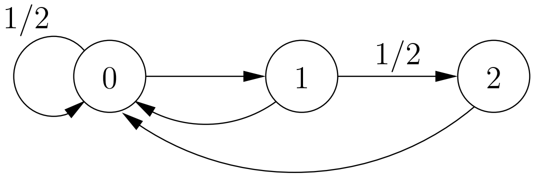
2. Consider the binary-input, ternary-output discrete memoryless channel with transition probabilities as shown in the diagram below.



(a) Is this channel symmetric? Weakly symmetric?

(b) Find the capacity of this channel and a capacity-achieving input distribution.

3. Consider the three-state Markov chain with transitions between states and transition probabilities as shown in the diagram below. (Transition probabilities not shown can be inferred from those shown.)



(a) The Markov chain operates in steady state, i.e., forming a stationary process. Find the entropy rate of this process.

(b) Let $X_0, X_1, X_2, X_3, \ldots$ denote the sequence of values taken on by this stationary Markov process starting at time 0, where X_i denotes the value of the process at time i. Find an efficient binary code to describe this sequence. [Hint: consider a code that describes X_0, and then efficiently describes the sequence of state-transitions.] Why is your code efficient?

(c) Use your code to encode the state sequence 101200.

4. Consider the following transmission strategy for a discrete-time additive white Gaussian noise channel with power constraint P and noise variance N.

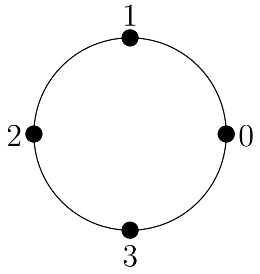
Split the power into two parts $P_1 = \alpha P$ and $P_2 = (1 - \alpha)P$, where $0 < \alpha < 1$. Design two independent Gaussian codebooks Ω_1 and Ω_2, where Ω_1 is a $(2^{nR_1}, n)$ codebook with power constraint P_1 and Ω_2 is a $(2^{nR_2}, n)$ codebook with power constraint P_2. The transmitter selects a codeword $X^{(1)}$ from Ω_1 and a codeword $X^{(2)}$ from Ω_2 and transmits their vector sum $X = X^{(1)} + X^{(2)}$. The receiver decodes the received word $Y = X + Z$ in two stages (here Z is noise). First it processes Y to find a codeword $\hat{X}^{(1)} \in \Omega_1$ (treating the contribution of $X^{(2)}$ as additional noise). It then forms $Y^{(2)} = Y - \hat{X}^{(1)}$ and processes $Y^{(2)}$ to find a codeword $\hat{X}^{(2)} \in \Omega_2$.

(a) How should R_1 and R_2 be chosen so that both decodings are successful with high probability in the limit as $n \to \infty$?

(b) Suppose the receiver reverses the order of processing (i.e., first decoding a codeword of Ω_2 and then decoding a codeword of Ω_1). How does this change your answer to part (a)?

(c) For fixed α, what is the maximum rate sum $R_1 + R_2$ that is achievable by this scheme? (A rate sum is achievable if there exists a sequence of $((2^{nR_1}, n), (2^{nR_2}, n))$ code-pairs for which the maximum probability of error approaches zero as $n \to \infty$.) Compare your maximum achievable rate sum with the capacity of the channel.

5. Let $Q = \{0, 1, 2, 3\}$ be a quaternary alphabet. For $x, y \in Q$, the "Lee distance" $L(x, y)$ between x and y is defined as the shortest "circular" distance between x and y on the circle shown in the figure and as tabulated below.



$L(x, y)$

$y =$	0	1	2	3
$x = 0$	0	1	2	1
1	1	0	1	2
2	2	1	0	1
3	1	2	1	0

Similarly, for $x \in Q$, the "Lee weight" $\rho(x)$ is defined as $\rho(x) = L(x, 0)$. Observe that $d(x, y) = \rho(x - y)$, when $x - y$ is computed modulo 4.

(a) Let X be a random variable with sample space Q. Let $w = E[\rho(X)]$ be the average Lee weight associated with X. For fixed $w \le 1$, show that the distribution that achieves maximum entropy with average Lee weight no greater than w is given by

$$\begin{aligned} p(0) &= \frac{(2-w)^2}{4} \\ p(1) = p(3) &= \frac{w(2-w)}{4} \\ p(2) &= \frac{w^2}{4} \end{aligned}$$

(b) A quaternary memoryless source produces symbols from Q with uniform probability. The output of the source is to be reproduced with letters from the same alphabet, and distortion is measured using the Lee distance as a distortion measure.

i. Determine $D_{\max}$, the distortion value associated with $R = 0$. Determine $R_{\max}$, the rate associated with $D = 0$.

ii. For $0 \le D \le D_{\max}$, determine the rate-distortion function $R(D)$.

iii. Verify that your rate-distortion function satisfies $R(D_{\max}) = 0$ and $R(0) = R_{\max}$.

(c) Explain how a rate-distortion code designed for a binary source with Hamming distortion might be adapted for use as a rate-distortion code for the quaternary source of (b) with Lee distortion. (*Hint:* provide an appropriate binary labeling for the circle shown in the figure.)