ECE 695C, Midterm #1
6:30-7:30pm Tuesday, September 29, 2009, ME 312,

1. Enter your name, student ID number, e-mail address, and signature in the space provided on this page, **NOW!**
2. This is a closed book exam.
3. Some time-consuming and bonus questions are assigned a small number of points. Make sure you manage your time well during the exam.
4. There are 4 questions and 12 pages in the exam booklet. Use the back of each page for rough work.
5. Neither calculators nor help sheets are allowed.

Name:

Student ID:

E-mail:

Signature:

Question 1: [20%]

Consider a hypothesis testing problem with observation Y:

H_0: Y is of distribution $P_Y \sim e^{-2|y-1|}$, and

H_1: Y is of distribution $Q_Y \sim e^{-\pi(y+1)^2}$.

1. [6%] Write down the Maximum Likelihood (ML) detector in the form of a log-likelihood-ratio test.
2. [6%] Find out the decisions for different observations $Y = 0.5$, $Y = -0.5$, and $Y = -3$, respectively.
3. [8%] Prove that the ML detector minimizes the summation of the false-alarm and the misdection probabilities.

Question 2: [30%] Consider a binary code of the following 1 by N parity-check matrix

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \end{pmatrix}. \tag{1}$$

Suppose we pass this code through an i.i.d. binary-input/ternary-output channel with a parameter $\epsilon \in [0, 1]$. Namely, for each $X = x$, $x \in \{0, 1\}$, we have

$$P(Y = y|X = x) = \begin{cases} 1 - \epsilon & \text{if } y = (-1)^x \\ \epsilon & \text{if } y = 0 \\ 0 & \text{if } y = (-1)^{x+1} \end{cases} \tag{2}$$

1. [3%] What is the code rate of this code.

2. [8%] Suppose the received output is $\overrightarrow{y} = y_1 y_2 y_3 \cdots y_N$ where each y_i takes value in $\{-1, 0, 1\}$. Plot the factor graph representation of the joint probability $P_{\overrightarrow{X}\overrightarrow{Y}}(\overrightarrow{x}, \overrightarrow{y})$, assuming each valid codeword is equally likely to be chosen. Please explicitly describe the factor functions used in your factor graph representation.

3. [8%] Use the BCJR decoder on this factor graph to find the Maximum A posteriori Probability (MAP) decision of the third bit x_3, assuming the observation is $\overrightarrow{y} = -1, 1, 0, -1, 1, 1, 1, -1$ with $N = 8$. Please write down all the forward (and feedback) metrics $a_{X_i}(x_i)$ (and $b_{X_j}(x_j)$) in the following form

$$a_{X_i}(x_i) = \begin{cases} \cdots & \text{if } x = \cdots \\ \cdots & \text{if } x = \cdots \end{cases}. \tag{3}$$

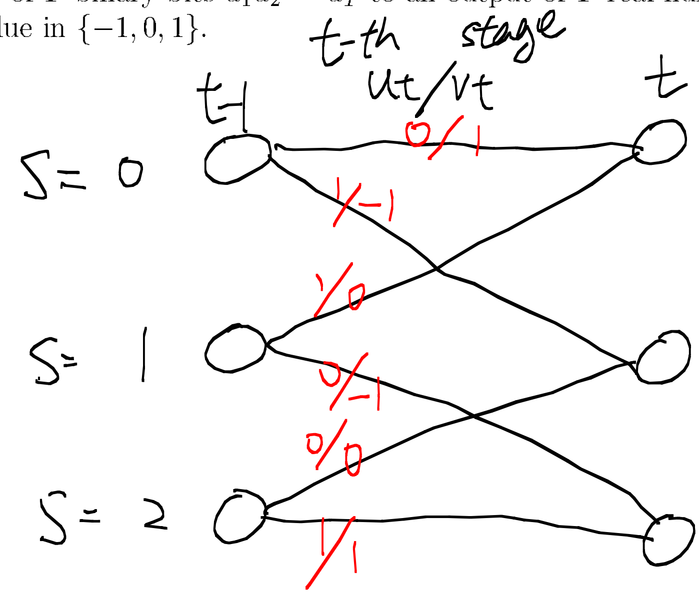
   And write down the decision rule in the final stage.

4. [5%] Repeat the above question for general $\overrightarrow{y} = y_1 y_2 y_3 \cdots y_N$. For this sub-question, there is no need to write down the detailed derivation of the state metrics. A final result in terms of the observation $y_1 y_2 y_3 \cdots y_N$ would be sufficient.

   Note: when there is a tie between the computed posterior probabilities, please declare x_3 as "uncertain" rather than choose it to be 0 or 1.

5. [6%] Suppose the all-zero codeword $\overrightarrow{x} = 0, 0, 0, 0, \cdots, 0$ was transmitted. Find the closed-form expression of the probability that the third bit x_3 is declared as "uncertain". Your answer should be in terms of ϵ and N.

Question 3: [35%] Consider the following trellis structure with three states, which takes an input of T binary bits $u_1 u_2 \cdots u_T$ to an output of T real numbers $v_1 v_2 \cdots v_T$. Each v_i takes value in $\{-1, 0, 1\}$.



We thus have 2^T different valid codewords. A popular application of trellis codes is "trellis quantization." That is, a trellis quantizer takes as input any T-dimensional real vector $\overrightarrow{x} = (x_1, x_2, \cdots, x_T)$ and outputs a binary string $\hat{\overrightarrow{u}} = (\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_T)$ such that among all 2^T valid trellis codewords, the $\hat{\overrightarrow{v}}$ real vector corresponding to $\hat{\overrightarrow{u}}$ is the one that is the closest to $\overrightarrow{x}$ in terms of the Euclidean distance $\sqrt{\sum_{i=1}^{T}(x_i - v_i)^2}$.

Describe how to implement such trellis quantization by describing the following three steps of your iterative trellis quantizers:

- For the t-th stage, write down the path-metric assignment for the six branches in the trellis when the input is x_t.
- Describe the update rules of your trellis quantizer.
- Describe carefully the final decision rule of your trellis quantizer.

Question 4: [15% + bonus 10%] **The description of this problem is very long but the questions are actually much simpler.**

Consider the following communication scheme over a binary symmetric channel with cross-over probability p. Namely, for any time slot t, the transmitter can remain silent $x_t = 0$ or transmit a bit $x_t = 1$. When transmitting a zero (remaining silent), the transmitter uses no power (cost = 0). When transmitting a one, the transmitter uses power 1 (cost = 1). There are totally T slots available.

Suppose we use $H_2(p)$ to denote the binary entropy function

$$H_2(p) \triangleq -p\log_2(p) - (1-p)\log_2(1-p). \tag{4}$$

The maximum number of bits one can transmit over this channel is $T(I(X;Y)) = T(1 - H_2(p))$, which requires $P(X=0) = P(X=1) = 1/2$. The total cost for the transmission is thus $T/2$. The data rate per cost is thus $\frac{T(1-H_2(p))}{T/2} = 2(1-H_2(p))$. On the other hand, we can choose non-uniform prior distribution $P(X=0) = 1-w$ and $P(X=1) = w$ instead. One can show that with the non-uniform prior distribution the achievable rate per unit cost is

$$\frac{H_2(w+p-2pw) - H_2(p)}{w}.$$

By some simple calculus, one can show that

$$\max_{w\in[0,1]} \frac{H_2(w+p-2pw) - H_2(p)}{w} = D(P_1||P_0), \tag{5}$$

where the two distributions P_1 and P_0 are

$$P_1(y) = P_{Y|X}(y|x=1) = \begin{cases} p & \text{if } y=0 \\ (1-p) & \text{if } y=1 \end{cases},$$
$$P_0(y) = P_{Y|X}(y|x=0) = \begin{cases} (1-p) & \text{if } y=0 \\ p & \text{if } y=1 \end{cases} \tag{6}$$

Namely, the maximum achievable capacity per unit cost (ACPUC) is $D(P_1||P_0)$ **with** $p^* = 0$**.**

Question: Consider the following code construction that achieves this optimal rate per unit cost and answer the following subquestions.

Our construction has $M+1$ valid codewords. We first divide the T bits into M different sub-periods, each period has T/M bits. The zero-th codeword $\vec{x}_0 = 00\cdots0$ contains T zeros. The second codeword is $\vec{x}_0 = \underbrace{11\cdots1}_{T/M \text{ bits}}\ \underbrace{00000\cdots0}_{(M-1)T/M \text{ bits}}$. Namely, the first sub-period is all one and the remaining bits are all zero. Similarly, the i-th codeword $\overrightarrow{x}_i$, $i = 2, \cdots, M$ has the i-th subperiod being all one while the remaining subperiods being all zero.

For example, if $T = 6$ and $M = 3$, the four valid codewords are: $\overrightarrow{x}_0 = 000000$, $\overrightarrow{x}_1 = 110000$, $\overrightarrow{x}_2 = 001100$, and $\overrightarrow{x}_3 = 000011$.

We choose the M value satisfying $M + 1 = 2^{\frac{T}{M}\cdot(D(P_1||P_0)-\delta)}$ for some small $\delta > 0$

1. [3%] With $M + 1 = 2^{\frac{T}{M}\cdot(D(P_1||P_0)-\delta)}$, how many bits can we transmit using this scheme?

2. [3%] With the above construction, in average, how much cost is it to transmit a codeword?

3. [2%] What is the rate per unit cost of the above scheme.

4. [7% + bonus 10%=17%] The remaining question is whether the receiver can successfully distinguish among the $M + 1$ different codewords $\overrightarrow{x}_i$, $i = 0, \cdots, M$ based on the noisy observation over the binary symmetric channel. Explain in words (or prove that) why using the above scheme the receiver can indeed successfully distinguish among $M + 1$ different codewords $\overrightarrow{x}_i$, $i = 0, \cdots, M$ when T is sufficiently large.

   Hint 1: Quantify the probability that the $\overrightarrow{x}_0$ being misdetected as $\overrightarrow{x}_i$ for some $i \neq 0$.

   Hint 2: Design a decoder that looks at each of the M subperiods individually first and then make a decision by jointly considering all M subperiods together.