This is a open book open notes exam. It is scheduled to last for 1.5 hours. All students must do their own work and are bound by the SEAS honor code. If you do not understand a question, clearly state the assumptions used in trying to find a solution.
1 (20 points)

Suppose that a Markov chain $X_1 \rightarrow X_2 \rightarrow X_3 \ldots \rightarrow X_n$ is given. This is not necessarily a stationary or even a time-invariant Markov chain.

a. Show that $I(X_1; X_n) \leq I(X_k; X_{k+1})$ for all $0 \leq k \leq n - 1$. Hint: It suffices to consider the Markov chain $X_1 \rightarrow X_k \rightarrow X_{k+1} \rightarrow X_n$.

b. Interpret part a as an “information bottleneck” limit on the end-to-end communication rate. That is, if information is passed through a sequence of systems, then $I(X_k; X_{k+1})$ is a measure of the information that goes through the $k$th system.

c. Strengthen part b in the following way. Select the probability distribution on $X_1$ to maximize $I(X_1; X_n)$. Show that the capacity of the end-to-end system is upper bounded by the capacity of any system that the information passes through along the way.
2  (20 points)

A different version of Fano’s inequality. Suppose that a random variable $X$ takes on $m$ possible values, with probabilities $p_1 \geq p_2 \geq p_3 \ldots \geq p_m$. The problem is to guess (estimate) $X$ before an experiment with the lowest probability of error and to determine a bound on the probability of error in terms of the entropy.

a. Argue that the best guess for a value of $X$ is $p_1$ and this results in a probability of error $P_e = 1 - p_1$.

b. Among all probability distributions that result in probability of error $P_e = 1 - p_1$, find the one with maximum entropy.

c. Express the entropy of the maximum entropy distribution from part b in terms of the probability of error $P_e$. 
3  (10 points)

a. You are confronted by an individual who claims to have a really good code for an source with 9 possible values. This individual announces that the lengths of the binary codewords are 2, 2, 3, 3, 4, 4, 4, 4, and 4 bits. What is your response?

b. Suppose that this individual has a second code for the same source that he claims is also good and has codeword lengths 2, 3, 3, 3, 4, 4, 4, 4, and 5 bits. What is your response?

c. (Extra Credit) What course at what institution would you suggest this individual take?
4 (35 points)

Often channels are modeled as having a state. The state determines the probability distribution between the input and output. In this problem, the states are modeled as an i.i.d. sequence of random variables $S_i$. The channel then effectively has two inputs, $X_i$ and $S_i$. The probability distribution of $S_i$ is fixed and denoted $p(s)$. The inputs $X_i$ are selected, as usual, to communicate information.

If the state is known to both the encoder and the decoder, then the capacity is $C_1 = \max_{p(x|s)} I(X;Y|S)$; this is achieved (in general) by an input probability distribution $p(x|s)$ that depends on the state $s$. If the state is not known to either the encoder or the decoder, then the capacity is $C_2 = \max_{p(x)} I(X;Y)$, as usual.

a. Prove that knowledge of the state helps. That is, show that $C_1 \geq C_2$.

b. Consider a binary channel with a state that controls which of two Z-channels is selected. If $S = 0$, then the Z-channel is

$$Q_0 = \begin{bmatrix} 1-q & q \\ 0 & 1 \end{bmatrix}. \quad (1)$$

If $S = 1$, then the Z-channel is

$$Q_0 = \begin{bmatrix} 1 & 0 \\ q & 1-q \end{bmatrix}. \quad (2)$$

For these channels, the top row are the transition probabilities for $X = 0$ and the bottom row are for $X = 1$. Each has the same value of $q$, $0 < q < 1$. Assume that $P(S = 0) = P(S = 1) = 0.5$. Compute $C_1$ and $C_2$.

Hint: The computation of $C_2$ is easy. The computation of $C_1$ depends on $q$ and is a generalization of the computation for a Z-channel. Do not derive the final expression for $C_1$, simply describe the equation that must be solved.

c. Comment about the special case of $q = 0.5$ in part b.
We proved that feedback does not increase capacity of a discrete memoryless channel. This does not mean that feedback does not help or simplify the design of codes.

Consider the binary erasure channel. With probability $1 - \epsilon$, the output equals the input. With probability $\epsilon$, the symbol is erased. That is, the output alphabet has three symbols, 0, 1, and the erasure symbol. If 0 or 1 is received at the output, then they were received correctly.

Design a simple coding strategy that achieves capacity in the presence of feedback. Recall that with feedback, the encoder has access to all previous outputs.

Hints: Recall that the capacity of the binary erasure channel is known to be $C = 1 - \epsilon$. This is equivalent to saying that the capacity equals the probability that a bit is successfully transmitted. Your design should recognize the importance of looking at the last output. Do not suggest a random coding strategy.