ECE 745 - Advanced Communication Theory, Spring 2009
Homework #1 (INCOMPLETE)

1. A Huffman code finds the optimal codeword to assign to a given block of N source symbols.

   (a) Show that $\{01, 100, 101, 1110, 1111, 0011, 0001\}$ cannot be a Huffman code for any N for any source distribution where every string to be coded has non-zero probability.

   (b) For a source producing an **IID sequence** of discrete random variables, each drawn from source alphabet $\mathcal{X}$, it has been found that a Huffman code on blocks of length 2 (i.e. $N = 2$ source symbols are taken at a time) has rate 2 bits/symbol, and a Huffman code on blocks of length 3 (i.e. $N = 3$ source symbols are taken at a time) has rate 1.6 bits/symbol.

   - Find upper and lower bounds to the **first-order entropy** of the source.
   - Find upper and lower bounds to the **entropy rate** of the source.
   - What can be deduced about the size of the source alphabet?

   (c) Does there exist a source producing an **IID sequence** of discrete random variables, and integers $N \in \{1, 2, 3, 4, \ldots\}$ and $M \in \{2, 3, 4, \ldots\}$, such that a Huffman code on blocks of length N has (strictly) **smaller** rate (in bits/symbol) than a Huffman code on blocks of length MN?

2. An independent and identically distributed (IID) binary source has $P(X_k = 0) = 0.8$, $P(X_k = 1) = 0.2$, $\forall k$. Suppose that I take blocks of $N = 3$ bits at a time from this source as the input to my source coder.

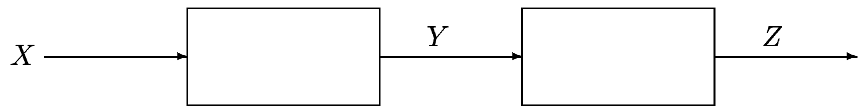
   - Find a Huffman code for this situation (i.e. a Huffman code that assigns to each 3-bit input sequence a variable-length output bit sequence.)
   - Find the rate of your code from (a) in average output bits per input bit. **Also, (do not forget this part)**, use the entropy of the source to show that the rate of your Huffman code falls between some easily obtained (but non-trivial) bounds.
   - Your boss has read a little information theory and has fallen in love with Shannon-Fano coding. Find a Shannon-Fano code for the source (i.e. a Shannon-Fano code that assign to each 3-bit input sequence a variable-length output bit sequence) and verify that its rate is consistent with your answers from (a) and (b). *Hint: The Shannon-Fano definition is non-constructive, but you should be able to find a code that meets the definition.*

   (b) Now, your boss comes to you with two pieces of good news: (i) There are no complexity limits (i.e. you can let your N be as big as you want), and (ii) the source, which is stationary, is correlated. In fact, he tells you that, for any n, the conditional probability of the n^{th} symbol only depends on the previous symbol; that is,

$$p_{X_n|X_1,X_2,X_3,\ldots,X_{n-1}}(x_n|x_1,x_2,x_3,\ldots,x_{n-1}) = p_{X_n|X_{n-1}}(x_n|x_{n-1}) = \begin{cases} 0.9 & x_n = 0, x_{n-1} = 0 \\ 0.1 & x_n = 1, x_{n-1} = 0 \\ 0.4 & x_n = 0, x_{n-1} = 1 \\ 0.6 & x_n = 1, x_{n-1} = 1 \end{cases}$$

   Using stationarity and this conditional probability, you can show that the marginal probabilities are still $P(X_n = 0) = 0.8$, $P(X_n = 1) = 0.2$. Find the rate that I tell my boss to expect out of the source coder **and compare it to your answer from (a).**

3. Consider discrete random variables X, Y, and Z. Suppose that X and Z are conditionally independent given Y as implied by the diagram below:



Are each of the following statements true or false? ("True" means for **all** X, Y, and Z such that X and Z are conditionally independent given Y. "False" means **there exists** X, Y, and Z, where X and Z are conditionally independent given Y, such that the statement does not hold). Be sure to give justification for your answers.

(a) $I(X; Y|Z) = 0$

(b) $H(X|Z) = H(X)$

(c) $H(X|Y, Z) = H(X|Y)$

(d) $I(X; Z) \leq H(Y)$

(e) $H(Y|Z) \leq H(X|Z)$

(f) $I(X; (Y, Z)) = I(X; Y)$

4. A discrete-valued source outputs an independent and identically distributed (IID) sequence of random variables (X_i), each drawn from the alphabet $\mathcal{X} = \{a_1, a_2, \ldots, a_Q\}$. I have not been able to find the probability mass function $p_{X_i}(x)$; however, I have been able to show that for any $\epsilon > 0$, there are sets $B_1, B_2, B_3, \ldots$ that have the following three properties:

(i) B_N is a subset of $\mathcal{X}^N$.

(ii) $P((X_1, X_2, \ldots, X_N)^T \in B_N) \to 1$ as $N \to \infty$.

(iii) $|B_N| = 4^{3\epsilon N + 5N}$.

What can be said about the entropy of the source?

5. The following is a formal restatement of our theorem for almost lossless fixed length source codes:

Let (X_i) be an IID sequence of discrete random variables with common distribution $p_X(x)$ and first-order entropy $H(X)$.

(1) Then, for any $R > H(X)$, any $\delta > 0$, there exists a fixed length source code with rate less than R and probability of decoding error $P_e < \delta$.

(2) Conversely, for any $R < H(X)$ and any $\epsilon > 0$, there is a (large) integer N_0 such that for all $N \geq N_0$, a fixed length source code that operates on N symbols at a time has probability of decoding error bounded as:

$$P_e > 1 - \epsilon - 2^{N(R - H(X) + \epsilon)}$$

By applying the Typical Sequences Theorem, show:

(a) The positive statement (statement (1)).

(b) The converse (statement (2)).

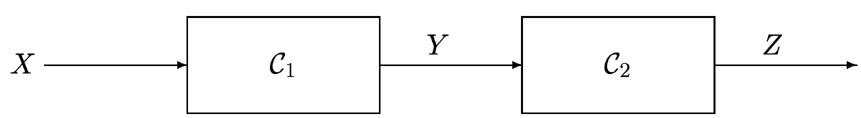
(c) Show that for $R < H(X)$, $\lim_{N\to\infty} P_e = 1$.

6. A source produces an **IID sequence** of discrete random variables, each drawn from source alphabet $\mathcal{X} = \{a, b, c, d\}$, where the elements have respective probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}$. We want to encode this source with a **fixed length almost lossless** source code. Assume that a (very large number) N source symbols are taken at a time and mapped into K binary digits.

   (a) Assuming N can be chosen arbitrarily large, find the infimum of the rates $\frac{K}{N}$ for which the decoding error probability can be made arbitrarily small.

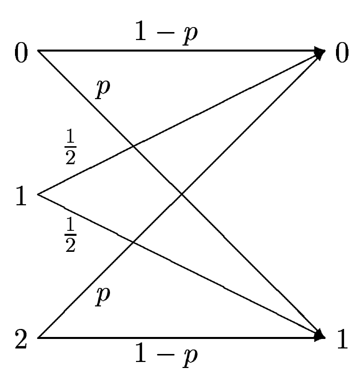
   (b) Now suppose that I restrict each of the **output** sequences of K binary digits to have **at most** $\frac{K}{4}$ 1's. Assuming N can be chosen arbitrarily large, **estimate** the infimum of the rates $\frac{K}{N}$ for which the decoding error probability can be made arbitrarily small (Yes, there is a nice closed form solution).

7. Consider two discrete memoryless channels $\mathcal{C}_1$ and $\mathcal{C}_2$ with capacities C_1 and C_2, respectively. Assume that the output alphabet of $\mathcal{C}_1$ is equal to the input alphabet of $\mathcal{C}_2$ and that they are connected as:



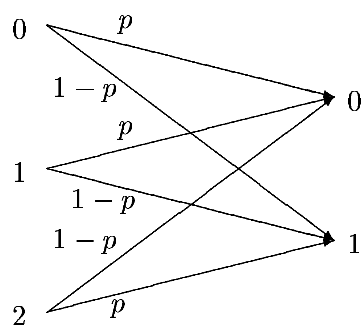
   Show that the capacity of the cascade of channels (i.e. the channel with input X and output Z) is less than or equal to the minimum of C_1 and C_2.

8. Find the **capacity** of the channel shown below:

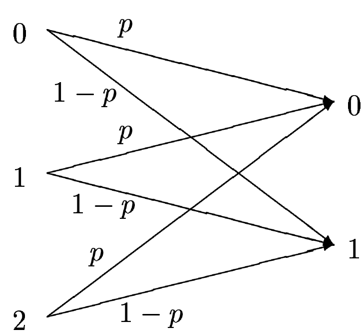


9. *This is an old exam problem:*

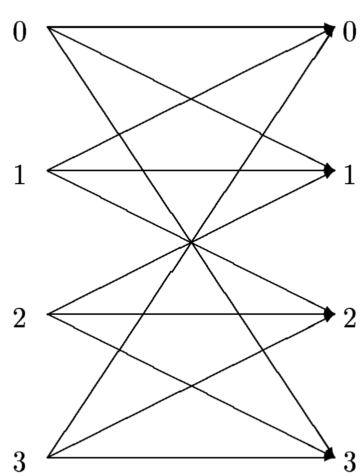
(a) Find the capacity of the discrete memoryless channel (DMC) denoted by:



(b) Find the capacity of the discrete memoryless channel (DMC) denoted by:



(c) Find the capacity of the discrete memoryless channel (DMC) denoted by:



where each of the transitions shown has probability $\frac{1}{3}$.

10. Suppose that we are trying to ”guess” X from an observation of Y with the decision rule $\hat{X}(Y)$. Let $P_e = P(\hat{X}(Y) \neq X)$.

(a) Show that any optimum decision rule can never have P_e greater than $\frac{|\mathcal{X}|-1}{|\mathcal{X}|}$.

(b) Show that Fano’s function $\mathcal{F}_{|\mathcal{X}|}(p)$ is maximized at $p = \frac{|\mathcal{X}|-1}{|\mathcal{X}|}$.

(c) Suppose $I(X;Y) = H(X)$. Show that an optimal decision rule has $P_e = 0$.