

EECS 769 Exam Redo

Elise McElhiney

December 1, 2016

0.1 Problem 1

Consider M discrete memoryless channels, channel k has input alphabet \mathcal{X}_k , output alphabet \mathcal{Y}_k , and is described by probabilities $P_k(y|x)$. Let C_k denote the capacity of channel k (computed in log base e).

Consider a new channel described as follows: the inputs to the channel is a pair (k, x) where $k \in \{1, \dots, M\}$, and $x \in \mathcal{X}_k$. When the input is (k, x) , the output of the channel is (k, y) where y is chosen according to $P_k(y|x)$. [In other words, the user of the new channel chooses which of the M original channels he wants to use with which input, and the channel output indicates which channel was used, and the output of that channel.]

(a) Show that

$$\begin{aligned} I(K, X; K, Y) &= H(K) + I(X; Y|K) \\ I(K, X; K, Y) &= I(K; K, Y|X) + I(X; K, Y|X) \quad \text{by chain rule} \\ &= H(K) + I(X; Y|K) \end{aligned} \tag{1}$$

(b) Show that

$$I(K, X; K, Y) \leq H(K) + E[C_K]$$

where $E[C_K] = \sum_{k=1}^M Pr(K = k)C_k$

Indicate the conditions under which the inequality will become equality.

$$\begin{aligned} I(K, X; K, Y) &\leq H(K) + E[C_K] \\ I(X; Y|K) + H(K) &\leq H(K) + E[C_K] \\ I(X; Y|K) &\leq E[C_K] \end{aligned}$$

The mutual information (or what we can send on the channel) given the channel K will be less than or equal to the expected value of the capacity of the channel. This is true because of the source-channel coding theorem that states that we can't exceed capacity. We get equality iff we maximize this for $p_x(y|x)$, that is:

$$\max_{p_x(y|x)} I(X; Y) = I(X; Y|K)$$

(c) Let $p_k = Pr(K = k)$. Show that $H(K) + E[C_K]$ is maximized if $p_k = \exp(C_k) / \sum_{i=1}^M \exp(C_i)$.

$H(K)$ is maximized when X is a uniform distribution over the alphabet. Since we have $k = 1, 2, \dots, M$ channels we know that the alphabet size of each channel is $\exp C_k$. We also know that the effective alphabet size for the combination of all the parallel channels is $\sum_{i=1}^M \exp C_i$.

This equation is the extended form of $\exp C_1 + \exp C_2 = \exp C_{1,2}$. If there is a third channel we then have $\exp C_{1,2,3} = \exp C_{1,2} + \exp C_3$. From the above equation we can now write:

$$\exp C_{1,2,\dots,M} = \sum_{i=1}^M \exp C_i$$

If X is to be uniformly distributed over the entire effective alphabet, the probability of the M channels must be distributed proportional to their effective alphabet size. Or (with p_k being the probability that the channel is used):

$$p_k = \frac{\exp C_i}{\sum_{i=1}^M \exp C_i}$$

(d) Show that the capacity of the channel is

$$\ln \sum_{i=1}^M \exp(C_i)$$

As stated above, we know that:

$$\begin{aligned} \exp C_K &= \sum_{i=1}^M \exp C_i \\ \therefore C_K &= \ln \sum_{i=1}^M \exp C_i \end{aligned}$$

0.2 Problem 2

(a) Let X_1, X_2 , and Y be 3 random variables such that X_1 and X_2 are independent. Is:

i) $I(X_1; Y|X_2) \leq I(X_1; Y)$; or

ii) $I(X_1; Y|X_2) \geq I(X_1; Y)$; or

iii) a definitive relationship can not be established

Give proof for your assertion

$$I(X_1; Y, X_2) = I(X_1; Y, X_2)$$

$$I(X_1; Y) + I(X_1; X_2|Y) = I(X_1; X_2) + I(X_1; Y|X_2)$$

Since X_1 and X_2 are independent we know that $I(X_1; X_2) = 0$ this becomes:

$$I(X_1; Y|X_2) - I(X_1; Y) = I(X_1; X_2|Y) \geq 0 \quad \text{by property of non-negativity}$$

Therefore:

$$I(X_1; Y|X_2) \geq I(X_1; Y)$$

(b) We can think of X_1 and X_2 as the transmitted symbols of two non-cooperating users in a multiple access channel whereby two users are trying to communicate with a single receiver simultaneously:

Although we have not formally discussed multiple-access channels, can you venture and interpretation of your result in (a) in terms of what is says about the performance limits of communication in such channels? (Hint: X_1 and X_2 are interfering with each other)

We will be limited in performance because of the interference. My conclusion in part a suggests that the interference is strictly destructive and that our upper bound for capacity for either channel is the capacity of the channel without interference (i.e. if we know the other value x_2)

0.3 Problem 3

(a) A source has an alphabet of 4 letters a_1, a_2, a_3, a_4 , and we have the condition $p(a_1) > p(a_2) = p(a_3) = p(a_4)$. Find the smallest number q such that $p(a_1) > q$ implies $n_1 = 1$, where n_1 throughout is the length of the code word for a_1 in a binary Huffman code for the source.

$$p(a_1) \geq p(a_2) = p(a_3) = p(a_4)$$

In order for $n=1$ (i.e. $p(a_1)$ is never required to be removed via Huffman reduction) we must have:

$$p(a_1) \geq p(a_2) \quad \text{known to be true from the initial equation}$$

$$p(a_1) \geq p(a_3) + p(a_4)$$

$$p(a_1) \geq 2p(a_2)$$

We also know:

$$p(a_1) + p(a_2) + p(a_3) + p(a_4) = 1$$

$$p(a_1) = 1 - 3p(a_2) \quad \text{since } p(a_2) = p(a_3) = p(a_4)$$

Combining above equations we get:

$$1 - 3p(a_2) \geq 2p(a_2)$$

$$1 \geq 5p(a_2)$$

$$\frac{1}{5} \geq p(a_2)$$

$$\therefore p(a_1) \geq \frac{2}{5}$$

(e) Assume $p(a_1) \geq p(a_2) \geq \dots \geq p(a_K)$. Find the largest number q' such that $p(a_1) < q'$ implies that $n_1 > 1$.

$q' = 0.4$. If $p(a_1) = 0.4 - \delta$ then the final four values could easily achieve the values $0.4 - \delta, 0.2, 0.2, 0.2 + \delta$. Therefore the final three values are $0.4 - \delta, 0.4, 0.2 + \delta$ as you can see, our next reduction will include $p(a_1) = 0.4 - \delta$ and we will have $n_1 > 1$ as a result. The explanation of the converse is written out in part (d) where I explain why for any value $p(a_1) > 0.4$ will result in $n_1 = 1$

0.4 Problem 4

(a) Consider the Z-channel below with input distribution $q = Pr(X = 0), 1 - q = Pr(X = 1)$. Denote the capacity of the Z-channel as C_Z . Find the value of q (as a function of ϵ) that achieves capacity.

$$\begin{aligned}
 H(Y|X) &= P(X=0)H(1-\epsilon, \epsilon) + P(X=1)H(1, 0) = P(X=0)H(1-\epsilon, \epsilon) = qH(1-\epsilon, \epsilon) \\
 H(Y) &= H(\epsilon) = H(Pr(Y=0)) = H(q(1-\epsilon)) \\
 I(X;Y) &= H(Y) - H(Y|X) \\
 &= H(q(1-\epsilon)) - qH(1-\epsilon, \epsilon) \\
 &= -q(1-\epsilon) \log q(1-\epsilon) - q(-(1-\epsilon) \log(1-\epsilon) + (\epsilon) \log \epsilon) \\
 &= -q((1-\epsilon) \log q(1-\epsilon) - ((1-\epsilon) \log(1-\epsilon) + (\epsilon) \log \epsilon)) \\
 &= -q((1-\epsilon) \log q(1-\epsilon) - (1-\epsilon) \log(1-\epsilon) - (\epsilon) \log \epsilon) \\
 &= q(1-\epsilon) \log \frac{1}{q(1-\epsilon)} - q(1-\epsilon) \log \frac{1}{(1-\epsilon)} - q(\epsilon) \log \frac{1}{\epsilon} \\
 \frac{d}{dq} &= \frac{(\epsilon-1)(-\log \frac{1}{q-\epsilon q}) + \log \frac{1}{1-\epsilon} + 1}{\log 2}
 \end{aligned} \tag{2}$$

I used wolfram alpha to find the root for the variable q:

$$q = -\frac{1-\epsilon}{\exp(\epsilon-1)}$$

Which I believe will maximize the mutual information $I(X;Y)$ and as such achieve capacity.

(b) Consider a cascade of two channels, as shown below. The output of the first channel directly input to the second. Give the transition probabilities for the effective XZ channel.

	z = 0	z = 1
x = 0	$1 - \epsilon + \epsilon\delta$	$\epsilon(1 - \delta)$
x = 1	δ	$(1 - \delta)$

(c) Find the value of δ (in terms of ϵ) that makes the effective XZ channel symmetric, and find the corresponding capacity of the XZ channel.

In order for the XZ channel to be symmetric the following two equations must be true.

$$\delta = \epsilon(1 - \delta)$$

$$1 - \epsilon + \delta\epsilon = 1 - \delta$$

By solving this system of equations, we get:

$$\delta = \frac{\epsilon}{\epsilon + 1}$$

We know that for a symmetric channel we have:

$$C = \log |Z| - H(\text{row of transition matrix})$$

$$C = \log 2 - H\left(\frac{\epsilon}{\epsilon + 1}, 1 - \frac{\epsilon}{\epsilon + 1}\right)$$

Which is a Bernoulli function where $\alpha = \frac{\epsilon}{\epsilon + 1}$:

$$C = 1 - \frac{\epsilon}{\epsilon + 1} \log \frac{\epsilon}{\epsilon + 1} - \left(1 - \frac{\epsilon}{\epsilon + 1}\right) \log 1 - \frac{\epsilon}{\epsilon + 1}$$

(d) What is the relationship between the capacity computed in (c) and C_Z ? That is, explain the conditions where the capacity computed in (c) is less than, equal to, or greater than C_Z .

We know from the data processing inequality, we know that we can never gain information through processing. Therefore $\max I(X; Y) \geq \max I(X; Z)$ Equality occurs iff we can fully retrieve the value of X given the value of Y. In this particular example, we could only accomplish this if $\epsilon = 0$ since for all other cases, we can't fully determine any X given Y due to the transition probability when $X = 0$. For all other cases, the capacity computed in (c) will be less than C_Z . The capacity for (c) will never be greater than C_Z since the additional channel will increase the uncertainty.

(e) Suppose now that the output sequence from the first channel can be arbitrarily processed and the processed result can be sent on from Y to Z. Find the resulting capacity of the combined system from X to Z, C_{XZ} (Hint: try to express C_{XZ} in terms of C_Z and C_{YZ}).

$X, Y, \text{ and } Z$ clearly form a Markov chain: $X \rightarrow Y \rightarrow Z$ We know from the data processing inequality that $I(X; Y) \geq I(X; Z)$ We know that the maximum effective capacity two channels in series is the the product of the two channels capacities. That is:

$$C_{XZ} \leq C_Z * C_{YZ}$$

Since we process the the output sequence at Y we know that we either retain or lose mutual information, but we can never gain mutual information. This means that the total capacity $C = \max I(X; Z)$ will decrease due to processing unless the processing results in a Y where $X = g(Y)$ in which case it will be equal.

0.5 Problem 5

Let the input random variable X to a channel be uniformly distributed over the interval $-1/2 \leq x \leq +1/2$. Let the output of the channel be $Y = X + Z$, where the noise random variable is uniformly distributed over the interval $-a/2 \leq z \leq +a/2$.

The density of Y is as follows:

For $a < 1$:

$$p_Y(y) = \begin{cases} (\frac{1}{2a})(y + \frac{1+a}{2}) & -\frac{1+a}{2} \leq y \leq -\frac{1-a}{2} \\ 1 & -\frac{1-a}{2} \leq y \leq +\frac{1-a}{2} \\ (\frac{1}{2a})(-y - \frac{1+a}{2}) & +\frac{1-a}{2} \leq y \leq +\frac{1+a}{2} \end{cases}$$

For $a = 1$: Y is triangular over $[-1, +1]$

For $a > 1$:

$$p_Y(y) = \begin{cases} y + \frac{1+a}{2} & -\frac{1+a}{2} \leq y \leq -\frac{1-a}{2} \\ \frac{1}{a} & -\frac{1-a}{2} \leq y \leq +\frac{1-a}{2} \\ -y - \frac{1+a}{2} & +\frac{1-a}{2} \leq y \leq +\frac{1+a}{2} \end{cases}$$

Find $I(X; Y)$ as a function of a .

If $a < 1$ the output Y will be conditionally uniformly distributed with probability $1 - a$ over the interval $[-\frac{1-a}{2}, +\frac{1-a}{2}]$. Y will have a triangular density with the base of the triangle having width a with probability a .

$$h(Y) = H(a) + (1 - a) \ln(1 - a) + a(\frac{1}{2} + \ln a)$$

$$h(Y) = -a \ln a - (1 - a) \ln(1 - a) + (1 - a) \ln(1 - a) + \frac{a}{2} + a \ln a$$

$$h(Y) = \frac{a}{2} \text{ nats}$$

If $a > 1$ Y 's trapezoidal density be scaled by a factor a :

$$h(Y) = \ln a + \frac{1}{2a}$$

Therefore:

$$I(X; Y) = \begin{cases} \frac{a}{2} - \ln a & a \leq 1 \\ \frac{1}{2a} & a \geq 0 \end{cases}$$

For $a = 1$ find the capacity of the channel when the input X is peak-limited; that is, the range of X is limited to $-1/2 \leq x \leq +1/2$. What probability distribution on X maximizes the mutual information $I(X; Y)$?

$$I(X; Y) = h(Y) - h(Y|X) = h(Y) - h(Z)$$

Both X and Y are restricted to the interval $[-\frac{1}{2}, +\frac{1}{2}]$. This means that since $Y = X + Z$, it is limited to the interval $[-1, +1]$. The differential entropy of Y is \leq the entropy of a uniformly distributed random variable on the interval $[-1, +1]$. This means, $h(Y) \leq 1$. This equation for maximum entropy achieves equality iff $x = \pm 1$ with probabilities of $\frac{1}{2}$ each.

$$I(X; Y) = h(Y) - h(Z) = 1 - 0 = 1$$

$$\hat{X} = \begin{cases} -\frac{1}{2} & y < 0 \\ +\frac{1}{2} & y \geq 0 \end{cases}$$