Information Theory and Coding
Prof. Suhas Diggavi

EPFL Winter Semester 2008/2009
Handout #19, Tuesday, 25 November 2008

Homework Set #8

Due 4 December 2008, before 12:00 noon, INR 031/032/038

Problem 1 (Feedback Capacity of Erasure Channels with Memory)

Consider a binary memoryless erasure channel with $\alpha = .2$ fraction of bits erased as shown in Figure 1. Assume that this channel is provided with feedback; i.e., all the received symbols are sent back

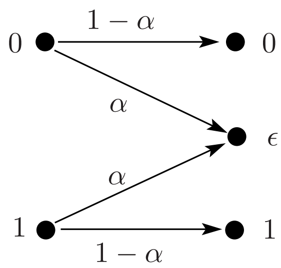


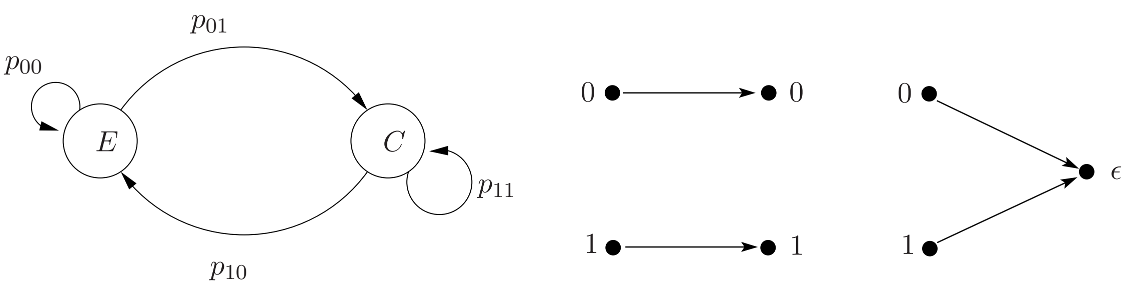
Figure 1: Binary Memoryless Erasure Channel

immediately and noiselessly to the transmitter, which can then use them and decide which symbol to send next.

(a) Assume that the transmitter retransmits every bit until it gets through. What is the rate of transmission?

(b) Compute the capacity of this feedback channel.

Now assume that the above binary erasure channel is not memoryless and erasure occurs based on a 2-state Markov process as shown in Figure 2(a). State E is when an erasure occurs and state C is when no erasure occurs. The equivalent channel at each state is shown in Figure 2(b) and 2(c).



(a) Markov chain of the channel state (b) Channel in state C (c) Channel in state E

Figure 2: Binary Erasure Channel with Memory

Assume that the transition matrix is $P = \begin{pmatrix} .8 & .2 \\ .5 & .5 \end{pmatrix}$.

(c) What is the capacity of this channel if the feedback is not present?

(d) Now consider the channel described above and assume that feedback is present. Compare the capacity of this feedback channel with the results of part (c). *Hint:* Verify the following chain of equalities and inequalities.

$$\begin{aligned}
I(W;Y^n) &\stackrel{(I)}{=} I(W;Y^n,Q^n) \\
&\stackrel{(II)}{=} I(W;Y^n|Q^n) \\
&= \sum_{i=1}^{n} H(Y_i|Y^{i-1},Q^n) - \sum_{i=1}^{n} H(Y_i|Y^{i-1},W,Q^n) \\
&\stackrel{(III)}{=} \sum_{i=1}^{n} H(Y_i|Y^{i-1},Q^n) - \sum_{i=1}^{n} H(Y_i|Y^{i-1},W,Q^n,X_i) \\
&\stackrel{(IV)}{=} \sum_{i=1}^{n} H(Y_i|Y^{i-1},Q^n) - \sum_{i=1}^{n} H(Y_i|Y^{i-1},Q^n,X_i) \\
&\stackrel{(V)}{\leq} \sum_{i=1}^{n} H(Y_i|Q_i) - \sum_{i=1}^{n} H(Y_i|Y^{i-1},Q^n,X_i) \\
&\stackrel{(VI)}{=} \sum_{i=1}^{n} H(Y_i|Q_i) - \sum_{i=1}^{n} H(Y_i|Q_i,X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i|Q_i)
\end{aligned}$$

Problem 2 (Arimoto-Blahut algorithm)

In this problem we study the Arimoto-Blahut algorithm which is an iterative algorithm to find the optimal input distribution of a discrete memoryless channel, and therefore its capacity. Note that the capacity of an MDC with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined as

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x)q(y|x)\log\frac{p(x)q(y|x)}{p(x)r(y)},$$

where $q(y|x)$ is the channel transition probability, and $r(y) = \sum_x p(x)q(y|x)$ is the probability of observing the symbol $y \in \mathcal{Y}$ at the receiver. Define the conditional probability of an input symbol given an output as

$$w^*(x|y) = \frac{p(x)q(y|x)}{r(y)} \qquad x \in \mathcal{X}, y \in \mathcal{Y}$$

and

$$f(p,w) = \sum_{x,y} p(x)q(y|x)\log\frac{w(x|y)}{p(x)}.$$

Thus, one can rewrite the capacity expression as

$$C = \max_{p,w} f(p,w).$$

(a) Assume that $p(x)$ is fixed. Show that $w^*(x|y)$ maximizes $f(p, w)$, *i.e.,* $f(p, w^*) \geq f(p, w)$ for arbitrary w.

(b) Now, assume that $w(x|y)$ is fixed. Find the optimal $p(x)$ that maximizes $f(p, q)$.
Hint: Note that $f(p, q)$ is a concave function with respect to $p(x)$ (you do not need to prove this). Use the Kuhn-Tucker conditions to find the maximizing $p(x)$.

The Arimoto-Blahut algorithms works as follows.

1. Set $n = 0$ and starts with an arbitrary initial distribution $p^{(0)}(x)$ where $p^{(0)}(x) > 0$ for $\forall x \in \mathcal{X}$.
2. Find the corresponding conditional distribution $w^{(n)}(x|y)$.
3. Find the $p^{(n+1)}$ which maximizes $f(p, w^{(n)})$ over p.
4. Increment n and goto step (2).

It can be shown that

$$C - f(p^{(n+1)}, w^{(n)}) \leq \sum_{x \in \mathcal{X}} p^*(x) \log \frac{p^{(n+1)}(x)}{p^{(n)}(x)}, \tag{1}$$

where $p^*(x)$ is the optimal input distribution.

(c) Show that $f(p^{(n+1)}, w^{(n)})$ tends to the capacity as $n \to \infty$.
Hint: Find the sum of the terms in LHS of (1) for $n = 0, 1, \ldots, N$ for some large enough N and show that $|C - f(p^{(n+1)}, w^{(n)})|$ converges to zero.

Problem 3 (Symmetric discrete memoryless channels)

In this problem we determine the capacity of general (not necessarily binary) symmetric discrete memoryless channels.

Definition 1 (Symmetric channels) *Let us define a matrix W such that the entry in the x^{th} row and y^{th} column corresponds to $w(y|x)$. We say that a channel is symmetric if the set of columns of W can be partitioned into subsets so that each subset has a property that the rows of the channel transition probabilities $w(y|x)$ are permutations of each other and the columns are permutations of each other.*

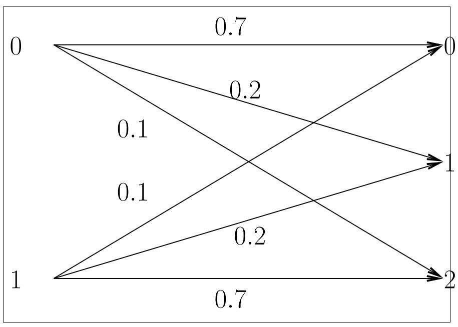


Figure 3: Symmetric channel with 2 inputs and 3 outputs

For example, the channel in Figure 3 is a symmetric channel. The matrix W for this channel is given by

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}$$

which can be partitioned into

$$\begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.7 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$$

You can verify that these partitions fulfill the conditions of Definition 1.

(a) Prove that, for a symmetric discrete memoryless channel, capacity is achieved by using the inputs with equal probability.

(b) If α is a probability vector, then define

$$H_k(\alpha) = \alpha_1 \log \frac{1}{\alpha_1} + \ldots + \alpha_k \log \frac{1}{\alpha_k}.$$

Using the result from part (a) and the KKT conditions, prove that the capacity of a symmetric channel is

$$C_S = \log |\mathcal{X}| - H_{|\mathcal{Y}|}(r) - \sum_{y \in \mathcal{Y}} p(y|x_t) \log(\sum_{x \in \mathcal{X}} p(y|x))$$

where $\mathcal{X}$ and $\mathcal{Y}$ are the input and output alphabet, respectively, and r is the row of the transition matrix W. Notice that the above expression does not depend on the choice of x_t, i.e. it holds for any $x_t \in \mathcal{X}$.

Problem 4 (Concavity of determinants)

Let θ be a binary random variable distributed according to Bernoulli(λ) for $0 \le \lambda \le 1$, and

$$Z = \begin{cases} X_1 & \text{if } \theta = 0 \\ X_2 & \text{if } \theta = 1, \end{cases} \tag{2}$$

where $X_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_1)$ and $X_2 \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_2)$, where $\mathcal{N}(\mathbf{0}, \mathbf{K})$ denotes a Gaussian distribution with $\mathbf{0}$ mean and covairance matrix $\mathbf{K}$.

(a) What is the distribution of Z?

(b) Find $h(Z)$.

(c) Find $h(Z|\theta)$.

(d) Compare the results of parts (b) and (c). Conclude an inequality which involves the $|\mathbf{K}_1|$ and $|\mathbf{K}_2|$, the determinants of $\mathbf{K}_1$ and $\mathbf{K}_2$.