ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

**Handout 17** Information Theory and Coding

Homework 7 November 5, 2013

---

Problem 1. Consider two discrete memoryless channels. The first channel has input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$; the second channel has input alphabet $\mathcal{Y}$ and output alphabet $\mathcal{Z}$. The first channel is described by the conditional probabilities $P_1(y|x)$ and the second channel by $P_2(z|y)$. Let the capacities of these channels be $C_1$ and $C_2$. Consider a third memoryless channel described by probabilities

$$P_3(z|x) = \sum_{y \in \mathcal{Y}} P_2(z|y) P_1(y|x), \quad x \in \mathcal{X},\, z \in \mathcal{Z}.$$

(a) Show that the capacity $C_3$ of this third channel satisfies

$$C_3 \leq \min\{C_1, C_2\}.$$

(b) A helpful statistician preprocesses the output of the first channel by forming $\tilde{Y} = g(Y)$. He claims that this will strictly improve the capacity.

(b1) Show that he is wrong.

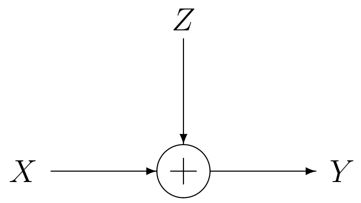
(b2) Under what conditions does he not strictly decrease the capacity?

Problem 2. Let $X$ be the channel input. Assume that the channel output $Y$ is passed through a date processor in such a way that no information is lost. That is,

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

where $Z$ is the processor output. Find an example where $H(Y) > H(Z)$ and find an example where $H(Y) < H(Z)$. Hint: The data processor does not have to be deterministic.

Problem 3. Find the channel capacity of the following discrete memoryless channel:




where $\Pr\{Z = 0\} = \Pr\{Z = a\} = \frac{1}{2}$ and $a \neq 0$. The alphabet for $x$ is $\mathcal{X} = \{0, 1\}$. Assume that $Z$ is independent of $X$.

Observe that the channel capacity depends on the value of $a$.

Problem 4. Consider the discrete memoryless channel $Y = X + Z \pmod{11}$, where

$$\Pr(Z = 1) = \Pr(Z = 2) = \Pr(Z = 3) = 1/3$$

and $X \in \{0, 1, \ldots, 10\}$. Assume that $Z$ is independent of $X$.

(a) Find the capacity.

(b) What is the maximizing $p^*(x)$?

PROBLEM 5. We are given a memoryless stationary binary symmetric channel BSC($\epsilon$). I.e, if $X_1, \ldots, X_n \in \{0,1\}$ are the input of this channel and $Y_1, \ldots, Y_n \in \{0,1\}$ are the output, we have:

$$P(Y_i|X_i, X^{i-1}, Y^{i-1}) = P(Y_i|X_i) = \begin{cases} 1-\epsilon & \text{if } Y_i = X_i, \\ \epsilon & \text{otherwise.} \end{cases}$$

Let $W$ be a random variable that is uniform in $\{0,1\}$ and consider a communication system with feedback which transmits the value of $W$ to the receiver as follows:

- At time $t = 1$, the transmitter sends $X_1 = W$ through the channel.
- At time $t = i+1 \leq n$, the transmitter gets the value of $Y_i$ from the feedback and sends $X_{i+1} = Y_i$ through the channel.

(a) Give the capacity $C$ of the channel in terms of $\epsilon$, and show that $C = 0$ when $\epsilon = \frac{1}{2}$.

(b) Show that if $\epsilon = \frac{1}{2}$, $I(X^n; Y^n) = n - 1$. This means that $I(X^n; Y^n) \leq nC$ does not hold for this system.

(c) Show that although $I(X^n; Y^n) > nC$ when $\epsilon = \frac{1}{2}$, we still have $I(W; Y^n) \leq nC$.

Note that since $W$ is the useful information that is being transmitted, it is the value of $I(W; Y^n)$ that we are interested in when we want to compute the amount of information that is shared with the receiver.

PROBLEM 6. Consider a random source $\mathcal{S}$ of information, and let $W$ be a random variable which represents the first $L$ symbols $U_1, \ldots, U_L$ of this source, i.e., $W = U_1^L$. We want to transmit the value of $W$ using a memoryless stationary channel as follows:

- At time $t = 1$, we send $X_1 = f_1(W)$ through the channel.
- At time $t = i+1 \leq n$, we send $X_{i+1} = f_i(W, Y^i)$ through the channel. $Y_1, \ldots, Y_i$ are the output of the channel at times $t = 1, \ldots, i$ respectively,

$f_1, \ldots, f_n$ are $n$ mappings that constitute the encoder. Clearly, this is a communication system with feedback as we are using the value of $Y^i$ in the computation of $X_{i+1}$.

In the previous problem, we gave an example which satisfies $I(X^n; Y^n) > nC$ and $I(W; Y^n) \leq nC$. Show that the inequality $I(W; Y^n) \leq nC$ always holds by justifying each of the following equalities and inequalities:

$$\begin{aligned} I(W; Y^n) &\overset{(a)}{=} \sum_{i=1}^{n} I(W; Y_i|Y^{i-1}) \overset{(b)}{\leq} \sum_{i=1}^{n} I(W, Y^{i-1}; Y_i) \overset{(c)}{\leq} \sum_{i=1}^{n} I(W, X_i, X^{i-1}, Y^{i-1}; Y_i) \\ &\overset{(d)}{=} \sum_{i=1}^{n} I(X_i, X^{i-1}, Y^{i-1}; Y_i) \overset{(e)}{=} \sum_{i=1}^{n} I(X_i; Y_i) \overset{(f)}{\leq} nC. \end{aligned}$$

Since $I(W; Y^n)$ represents the amount of information that is shared with the receiver, the inequality $I(W; Y^n) \leq nC$ shows that feedback does not increase the capacity.