Problem Set 4

These problems have been selected from the Course Textbook by Cover and Thomas.

7.1 *Preprocessing the output.* One is given a communication channel with transition probabilities $p(y \mid x)$ and channel capacity $C = \max_{p(x)} I(X;Y)$. A helpful statistician preprocesses the output by forming $\tilde{Y} = g(Y)$, claiming that this will strictly improve the capacity.

(a) Show that the statistician is wrong.

(b) Under what conditions does the statistician not strictly decrease the capacity?

7.3 *Channels with memory have higher capacity.* Consider a binary symmetric channel with $Y_i = X_i \oplus Z_i$, where $X_i, Y_i \in \{0,1\}$ and $\oplus$ is modulo 2 addition. Suppose that $Z_1, Z_2, \ldots$ have constant marginal probabilities $\Pr(Z_i = 1) = p = 1 - \Pr(Z_i = 0)$, but that $Z_1, Z_2, \ldots$ are not necessarily independent. For any positive integer n, assume that $(Z_1, \ldots, Z_n)$ is independent of the channel input $(X_1, \ldots, X_n)$. Show that

$$\max_{p(x_1,\ldots,x_n)} I(X_1,\ldots,X_n;Y_1,\ldots,Y_n) \geq n(1 - H(p, 1-p)).$$

7.5 *Using two channels at once.* Consider two discrete memoryless channels $(\mathcal{X}_1, p(y_1 \mid x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p(y_2 \mid x_2), \mathcal{Y}_2)$, with capacities C_1 and C_2, respectively. A new channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1 \mid x_1) \cdot p(y_2 \mid x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is formed in which the pair $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ is sent, resulting in the pair $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ being received. Find the capacity of this channel.

7.8 *Z-channel.* The Z-channel has binary input and output alphabets $\mathcal{X} = \mathcal{Y} = \{0,1\}$, and transition probabilities $p(y \mid x)$ given by

$$p(0 \mid 0) = 1, \quad p(1 \mid 0) = 0, \quad p(0 \mid 1) = \frac{1}{2}, \quad p(1 \mid 1) = \frac{1}{2}.$$

Find the capacity of the Z-channel and the maximizing input probability distribution.

7.28 *Choice of channels.* Find the capacity C of the union of two channels $(\mathcal{X}_1, p(y_1 \mid x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p(y_2 \mid x_2), \mathcal{Y}_2)$ with respective capacities C_1 and C_2, where at each time one can send a symbol either over channel 1 or over channel 2, but not both. Assume that the output alphabets are distinct and do not intersect (i.e., the receiver can unambiguously determine which channel was selected).

(a) Show that $2^C = 2^{C_1} + 2^{C_2}$. Thus 2^C is the effective alphabet size of a channel with capacity C.

(b) Compare with problem 2.10 (Problem Set 1) where $2^H = 2^{H_1} + 2^{H_2}$, and interpret part (a) in terms of the effective number of noise-free symbols.

(c) Use the above result to calculate the capacity of the channel where $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$ and $p(y \mid x)$ is defined via

$$\begin{array}{lll} p(0 \mid 0) = 1 - p, & p(1 \mid 0) = p, & p(2 \mid 0) = 0, \\ p(0 \mid 1) = p, & p(1 \mid 1) = 1 - p, & p(2 \mid 1) = 0, \\ p(0 \mid 2) = 0, & p(1 \mid 2) = 0, & p(2 \mid 2) = 1. \end{array}$$

7.9 *Suboptimal codes.* For the Z-channel of Problem 7.8, assume that we choose a $(2^{nR}, n)$ code at random, where each codeword is a sequence of *fair* coin tosses. This will not achieve capacity. Find the maximum rate R such that the probability of error $P_e^{(n)}$, averaged over the randomly generated codes, tends to zero as the block length n tends to infinity.

7.14 *Channels with dependence between the letters.* Consider the channel with $\mathcal{X} = \mathcal{Y} = \{00, 01, 10, 11\}$ where the channel output is determined as a deterministic function of the input according to $00 \mapsto 01$, $01 \mapsto 10$, $10 \mapsto 11$, and $11 \mapsto 00$. Thus if the 2-bit sequence 01 is the input to the channel, the output is 10 with probability 1. Let (X_1, X_2) denote the input symbols and let (Y_1, Y_2) denote the output symbols, where $X_1, X_2, Y_1, Y_2 \in \{0, 1\}$.

(a) Calculate the mutual information $I(X_1, X_2; Y_1, Y_2)$ as a function of the input distribution on the four possible pairs of inputs.

(b) Show that the capacity of the channel is 2 bit/channel use.

(c) Show that under the maximizing input distribution, $I(X_1; Y_1) = 0$. Thus, the distribution on the input sequences that achieves capacity does not necessarily maximize the mutual information between individual symbols and their corresponding outputs.

7.19 *Capacity of the carrier pigeon channel.* Consider a commander of an army besieged in a fort for whom the only means of communication to his allies is a set of carrier pigeons. Assume that each carrier pigeon can carry one letter (8 bits), that pigeons are released once every 5 minutes, and that each pigeon takes exactly 3 minutes to reach its destination.

(a) Assuming that all the pigeons reach their destination safely, what is the capacity of this link in bit/hour?

(b) Now assume that the enemies try to shoot down the pigeons and that they manage to hit a fraction α of them. Since the pigeons are sent a constant rate, the receiver knows when the pigeons are missing. What is the capacity of this link in bit/hour?

(c) Now assume that the enemy is more cunning and that every time they shoot down a pigeon, the send out a dummy pigeon carrying a random letter (chosen uniformly from all 8-bit letters). What is the capacity of this link in bit/hour?

Set up an appropriate model for the channel in each of the above cases, and indicate how to go about finding the capacity.

7.27 *Erasure channel.* Let $(\mathcal{X}, p(y \mid x), \mathcal{Y})$ be a discrete memoryless channel with capacity C. Suppose that this channel is cascaded immediately with an erasure channel $(\mathcal{Y}, p(s \mid y), \mathcal{Y} \cup \{?\})$ that erases its symbols with probability α. Denote the input to the first channel as X, the output of the first channel (which is the input to the erasure channel) as Y, and the output of the erasure channel as S. We then have

$$\begin{aligned} \Pr(S = y \mid X = x) &= (1-\alpha)p(y \mid x), \\ \Pr(S = ? \mid X = x) &= \alpha. \end{aligned}$$

Determine the capacity of this channel.

7.35 *Capacity.* Suppose the channel described by $\mathcal{P}$ has capacity C, where $\mathcal{P}$ is an $m \times n$ channel matrix.

(a) What is the capacity of

$$\tilde{P} = \begin{bmatrix} \mathcal{P} & 0 \\ 0 & 1 \end{bmatrix}?$$

(b) What is the capacity of

$$\hat{P} = \begin{bmatrix} \mathcal{P} & 0 \\ 0 & I_k \end{bmatrix},$$

where I_k is the $k \times k$ identity matrix. *Hint: use the result of Problem 7.28.*