Information Theory Exercise Sheet #6

(Channel Capacities)

University of Amsterdam, Master of Logic, Fall 2015
Lecturer: Christian Schaffner
TA: Mathias Madsen

Out: Wednesday, 2 December 2015
(due: Wednesday, 9 December 2015, 9:00)

1. **Capacity of Multiple Channel Uses** Prove Lemma 1 below stating that the capacity per transmission is not increased if we use a discrete memoryless channel many times. For inspiration, look again at the proof of the converse of Shannon's noisy-channel coding theorem.

   Lemma 1 (Lemma 7.9.2 in [CT]) *Let Y^n be the result of passing X^n through a discrete memoryless channel of capacity C. Then, $I(X^n;Y^n) \leq nC$ for all P_{X^n}.*

   Does your proof also work for the feedback case (i.e. where X_{i+1} is allowed to depend on X^iY^i)? If not, point out the steps in your proof where you use that there is no feedback.

2. **Symmetric Channels.** Consider the channel with transition matrix

$$P_{Y|X} = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{bmatrix} .$$

   In a transition matrix, the entry in the xth row and yth column denotes the conditional probability $P_{Y|X}(y|x)$ that y is received when x has been sent.

   Definition 1 *A channel is said to be* symmetric *if the rows of the channel transition matrix $P_{Y|X}$ are permutations of each other and the columns are permutations of each other. A channel is said to be* weakly symmetric *if every row of the transition matrix is a permutation of every other row and all the column sums $\sum_x P_{Y|X}(y|x)$ are equal.*

   For instance, the channel $P_{Y|X}$ above is symmetric and the channel

$$Q_{Y|X} = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{bmatrix}$$

   is weakly symmetric but not symmetric.

(a) Find the optimal input distribution and channel capacity of $Q_{Y|X}$.

(b) Give a general strategy how to compute the capacity for weakly symmetric channels. What is the optimal input distribution?

3. **Zero-error vs non-zero-error Shannon capacity:** Let $P_{Y|X}$ be a discrete memoryless channel with confusability graph G and capacity $C = \max_{P_X} I(X;Y)$.

   (a) Show that $\log(\alpha(G)) \leq C$.

   (b) Show that for any $n \geq 1$, $\log(\alpha(G^{\boxtimes n})) \leq \max_{P_{X^n}} I(X^n;Y^n)$, where the Y^n are obtained by using the channel n times, i.e. $P_{Y^n|X^n}(y^n|x^n) = \Pi_{i=1}^n P_{Y|X}(y_i|x_i)$ for all x^n, y^n.

   (c) Conclude that the zero-error Shannon capacity of G is at most the channel capacity C.

4. **Additive noise channel.** Find the channel capacity of the following discrete memoryless channel. On input X from $\mathcal{X} = \{0,1\}$, the output Y is obtained by adding (over the reals) another real random variable Z, i.e. $Y = X + Z$ with distribution $P_Z(0) = P_Z(a) = \frac{1}{2}$ independent of X. Compute the channel capacity for all possible values of $a \in \mathbb{R}$.

5. **Tall, fat people.** Suppose that the average height of people in a room is 1.5m. Suppose that the average weight is 50kg.

   (a) Argue that no more than one third of the population is 4.5m tall.

   (b) Find an upper bound on the fraction of people who are simultanously tall (say, at least 3m) and fat (say, at least 150kg).

6. **Erasures and errors in a binary channel** Consider a channel with binary inputs that has both erasures and errors. Let the probability of error be ε and the probability of erasure be α, so the channel is as described in Figure 1.

   (a) Find the channel capacity of this channel.

   (b) Specialize to the case of the binary symmetric channel ($\alpha = 0$).

   (c) Specialize to the case of the binary erasure channel ($\varepsilon = 0$).

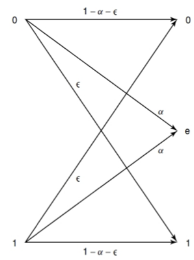

Figure 1: Erasures and errors in a binary channel.

7. **Encoder and decoder as part of the channel.** Consider a binary symmetric channel (BSC) $P_{Y|X}$ with crossover probability $\varepsilon = 0.1$. A possible coding scheme for this channel with two codewords of length 3 is to encode message w_1 as 000 and w_2 as 111. The decoder uses majority vote. With this coding scheme, we can consider the combination of encoder, channel, and decoder as forming a new BSC $Q_{Y|X}$, with two inputs w_1 and w_2 and two outputs w_1 and w_2.

   (a) Calculate the crossover probability of this new channel $Q_{Y|X}$.

   (b) What is the capacity of this new channel in bits per transmission of the original channel $P_{Y|X}$?

   (c) What is the capacity of the original BSC $P_{Y|X}$ with crossover probability $\varepsilon = 0.1$. Compare the two capacities.

   (d) Prove the following general result: For any channel, considering the encoder, channel, and decoder together (as a new channel from message W to estimated messages $\hat{W}$) will not increase the capacity in bits per transmission of the original channel.

Homework is exercises 2,3,4,6.