Information Theory

Exercise Sheet 6

Exercise 1 (Preprocessing the output)**.** *One is given a communication channel with transition probabilities* $p(y|x)$ *and channel capacity* $C = \max_{p(x)} I(\mathbf{X}, \mathbf{Y})$. *A helpful statistician preprocesses the output by forming* $\widetilde{\mathbf{Y}} = g(\mathbf{Y})$. *He claims that this will strictly improve the capacity.*

1. *Show that he is wrong.*
2. *Under what conditions does he not strictly decrease the capacity?*

Exercise 2 (Some capacities)**.** *Compute the capacity of the following noisy channels,*

1. *The typewriter channel.*
2. *The Binary Symmetric Channel (BSC) with probability of errors* f.
3. *The Binary Erasure Channel (BEC) with probability of errors* f.
4. *The Z channel with probability of errors* f. *What happens to the distribution reaching the capacity if the noise level* f *is very close to* 1?

Hint: *remember the useful formula* $H(\mathbf{Y} \mid \mathbf{X}) = \sum_i p(\mathbf{Y} = i) H(\mathbf{X} \mid \mathbf{Y} = i)$

Exercise 3 (Simulating channel)**.** *A binary erasure channel with input* x *and output* y *has transition probability matrix:*

$$\mathbf{Q} = \begin{pmatrix} 1-q & 0 \\ q & q \\ 0 & 1-q \end{pmatrix} \qquad \begin{tikzcd}[row sep=tiny] 0 \arrow[r] \arrow[dr] & 0 \\ & \perp \\ 1 \arrow[r] \arrow[ur] & 1 \end{tikzcd}$$

Find the mutual information $I(\mathbf{X}, \mathbf{Y})$ *between the input and output for general input distribution* (p_0, p_1), *and show that the capacity of this channel is* $C = 1 - q$ *bits.*

A Z channel has transition probability matrix:

$$\mathbf{Q} = \begin{pmatrix} 1 & q \\ 0 & 1-q \end{pmatrix} \qquad \begin{tikzcd} 0 \arrow[r] & 0 \\ 1 \arrow[r] \arrow[ur] & 1 \end{tikzcd}$$

Show that, using a $(2,1)$-block code, two uses of a Z channel can be made to emulate one use of an erasure channel, and state the erasure probability of that erasure channel. Hence show that the capacity of the Z channel, C_Z, satisfies $C_Z \geq \frac{1}{2}(1-q)$ bits.

Explain why the result $C_Z \geq \frac{1}{2}(1-q)$ is an inequality rather than an equality.

Exercise 4 (Time-varying channels.)**.** *Consider a time-varying discrete memory-less channel. Let $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$ be conditionally independent given $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, with conditional distribution given by*

$$p(y \mid x) = \prod_{i=1}^{n} p_i(y_i \mid x_i).$$

Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$, $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)$. Find $\max_{p(x)} I(\mathbf{X}, \mathbf{Y})$.

Exercise 5 (Channels with memory have higher capacity)**.** *Consider a binary symmetric channel with $\mathbf{Y}_i = \mathbf{X}_i + \mathbf{Z}_i \bmod 2$ where $\mathbf{X}_i, \mathbf{Y}_i \in \{0,1\}$. Suppose that $\{\mathbf{Z}_i\}_i$ has constant marginal probabilities $p(\mathbf{Z}_i = 1) = p = 1 - p(\mathbf{Z}_i = 0)$, but that $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$ are not necessarily independent. Assume that $\mathbf{Z}_n$ is independent of the input $\mathbf{X}_n$. Let $C = 1 - h(p)$. Notice that nC corresponds to the capacity of this channel when the $\mathbf{Z}_i$'s are independent. Show that*

$$\max_{p(x_1, x_2, \dots, x_n)} I\Big((\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n) \Big) \geq nC.$$

Exercise 6 (Encoder and decoder as part of the channel)**.** *Consider a Binary Symmetric Channel (BSC) with crossover probability* 0.1. *A possible coding scheme for this channel with two codewords of length* 3 *is to encode message a_1 as* 000 *and a_2 as* 111. *With this coding scheme, we can consider the combination of encoder, channel, and decoder (where the decoder, given $b_1b_2b_3$ chooses a_b where b is the majority bit in $b_1b_2b_3$) as forming a new BSC, with two inputs a_1 and a_2 and two outputs a_1 and a_2.*

1. *Calculate the crossover probability of this channel.*
2. *What is the capacity of this channel in bits per transmission of the original channel?*

3. *What is the capacity of the original BSC with crossover probability* 0.1*?*

4. *Prove a general result that for any channel, considering the encoder, channel, and decoder together as a new channel from messages to estimated messages will not increase the capacity in bits per transmission of the original channel.*

Exercise 7 (Transatlantic cables)**.** *A transatlantic cable contains* $N = 20$ *indistinguishable electrical wires. You have the job of figuring out which wire is which, that is, to create a consistent labelling of the wires at each end. Your only tools are the ability to connect wires to each other in groups of two or more, and to test for connectedness with a continuity tester. What is the smallest number of transatlantic trips you need to make, and how do you do it?*

How would you solve the problem for larger N *such as* $N = 1000$*?*

As an illustration, if N *were* 3 *then the task can be solved in two steps by labelling one wire at one end* a*, connecting the other two together, crossing the Atlantic, measuring which two wires are connected, labelling them* b *and* c *and the unconnected one* a*, then connecting* b *to* a *and returning across the Atlantic, whereupon on disconnecting* b *from* c*, the identities of* b *and* c *can be deduced.*

This problem can be solved by persistent search, but the reason it is posed is that it can also be solved by a greedy approach based on maximizing the acquired information. Let the unknown permutation of wires be x*. Having chosen a set of connections of wires* $\mathcal{C}$ *at one end, you can then make measurements at the other end, and these measurements* y *convey information about* x*. How much? And for what set of connections is the information* y *conveys about* x *maximized?*