ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 10 Information Theory and Coding
Homework 5 Oct. 15, 2019

PROBLEM 1. Assume $\{X_n\}_{-\infty}^{\infty}$ and $\{Y_n\}_{-\infty}^{\infty}$ are two i.i.d. processes (individually) with the same alphabet, with the same entropy rate $H(X_0) = H(Y_0) = 1$ and independent from each other. We construct two processes Z and W as follows:

- To construct the process Z, we flip a fair coin and depending on the result $\Theta \in \{0, 1\}$ we select one of the processes. In other words, $Z_n = \Theta X_n + (1 - \Theta)Y_n$.
- To construct the process W, we do the coin flip at every time n. In other words, at every time n we flip a coin and depending on the result $\Theta_n \in \{0, 1\}$ we select X_n or Y_n as follows $W_n = \Theta_n X_n + (1 - \Theta_n)Y_n$.

(a) Are Z and W stationary processes? Are they i.i.d. processes?

(b) Find the entropy rate of Z and W. How do they compare? When are they equal?
Recall that the entropy rate of the process U (if exists) is $\lim_{n\to\infty} \frac{1}{n} H(U_1, \cdots, U_n)$.

PROBLEM 2. Let X be the channel input. Assume that the channel output Y is passed through a data 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. A "K-ary erasure channel with erasure probability p" is described as follows: the input U belongs to the alphabet $\{1, \dots, K\}$, the output V belongs to the alphabet $\{1, \dots, K\} \cup \{?\}$, and if u is the input, the output V equals u with probability $1 - p$, and equals ? with probability p. Note that $\Pr(V = ?) = p$ regardless of the input distribution.

(a) Show that $\Pr(U = u|V = ?) = p_U(u)$.

(b) Show that $I(U; V) = (1 - p)H(U)$.

(c) Find the capacity of this channel and the input distribution that maximizes the mutual information.

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

(a) Find the capacity.

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

Problem 5. Suppose there are two discrete memoryless channels which are characterized by $(\mathcal{X}_1, p(x_1|y_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p(x_2|y_2), \mathcal{Y}_2)$ respectively. Assume further that $\mathcal{Y}_1$,$\mathcal{Y}_2$ and $\mathcal{X}_1$,$\mathcal{X}_2$ are disjoint (i.e. $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$ and $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$). Find the capacity C of the union of these two channels in terms of individual capacities C_1 and C_2. A union of these two channels means that the user can send one bit at a time using only one of these channels.

(Hint: You can flip a coin with optimal probability distribution to determine which channel to use.)

Problem 6. 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?