ASSIGNMENT 4

Exercise 1 (Capacity of two channels)**.** Consider two DMCs $(\mathcal{X}_1, p(y_1|x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p(y_2|x_2), \mathcal{Y}_2)$ with capacities C_1 and C_2, respectively. A new channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1|x_1) \times p(y_2|x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is formed in which $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$ are sent simultaneously, resulting in y_1, y_2. Find the capacity of this channel.

Exercise 2 (Choice of channels)**.** Find the capacity C of the union of two channels $(\mathcal{X}_1, p_1(y_1|x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p_2(y_2|x_2), \mathcal{Y}_2)$, where at each time, one can send a symbol over channel 1 or channel 2 but not both. Assume that the output alphabets are distinct and do not intersect. Show that $2^C = 2^{C_1} + 2^{C_2}$. Thus, 2^C is the effective alphabet size of a channel with capacity C.

Exercise 3 (Z-channel)**.** The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

$$Q = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, x, y \in \{0, 1\}$$

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

b. 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.

Exercise 4 (Unused symbols)**.** Show that the capacity of the channel with transition probability matrix

$$Q = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

is achieved by a distribution that places zero probability on one of input symbols. Give an intuitive reason as to why that symbol is not used. What is the capacity of this channel?

Exercise 5 (Erasures and errors in a binary channel)**.** A binary erasure channel with erasure probability β, denoted BEC(β) has output alphabet $\{0, 1, e\}$ and transitions given by $P(0|0) = P(1|1) = 1 - \beta$, and $P(e|0) = P(e|1) = \beta$ where e is the erasure symbol.

a. Show that the capacity of BEC(β) is $1 - \beta$.

b. Consider a channel with binary inputs that has both erasures and errors. Let the probability of error be α and the probability of erasure be β, which means that when we send symbol 0, with probability $1 - \alpha - \beta$ we receive symbol 0, with probability α we receive symbol 1 and with probability β we receive an erasure symbol. Find the capacity of this channel.

Exercise 6 (Binary multiplier channel)**.** Consider the channel $Y = X \cdot Z$, where X and Z are independent binary random variables and $Z \sim Ber(\alpha)$. [i.e., $P(Z = 1) = \alpha$].

a. Find the capacity of this channel and the maximizing distribution on X.

b. Now suppose that the receiver can observe Z as well as Y. What is the capacity?

Exercise 7 (Memoryless channels without feedback)**.** Consider a channel W given by $(\mathcal{X}^n, P(y^n|x^n), \mathcal{Y}^n)$.

1. Suppose that W is the n-th extension of a *memoryless channel* $(\mathcal{X}, P(y|x), \mathcal{Y})$. That is,

$$P(y_i|x^i, y^{i-1}) = P(y_i|x_i).$$

2. Also, suppose that W is used *without feedback* (the input symbols do not depend on the past output symbols). Namely,

$$P(x_i|x^{i-1}, y^{i-1}) = P(x_i|x^{i-1}).$$

Conclude that

$$P(y^n|x^n) = \prod_{i=1}^{n} P(y_i|x_i).$$

Hint – Expand $P(x^n, y^n)$ using Bayes' rule.

Exercise 8. (Conditional KL divergence) The conditional KL divergence between P and Q, probability distributions on $(\mathcal{X} \times \mathcal{Y})$, is given by

$$D(P_{X|Y}||Q_{X|Y}|Y) = \sum_{x,y} P(x,y) \log \frac{P(x|y)}{Q(x|y)}.$$

a. Prove that like the KL divergence, the conditional KL divergence is also non-negative.

b. Prove that

$$I(X;Y) = \max_{V(x|y)} \sum_{x,y} Q(x)P(y|x) \log \frac{V(x|y)}{Q(x)},$$

and the maximum is attained by $V^*(x|y) = \frac{Q(x)P(y|x)}{\sum_x Q(x)P(y|x)}$.

Exercise 9. (Capacity-achieving output distribution is unique) Consider a discrete memoryless channel and let $P_X^{(1)}$ and $P_X^{(2)}$ by two capacity-achieving input distributions. Define

$$P_X = \theta P_X^{(1)} + (1-\theta) P_X^{(2)}.$$

a. Let $X^{(1)} \sim P_X^{(1)}$ and $X^{(2)} \sim P_X^{(2)}$. Define

$$X = ZX^{(1)} + (1-Z)X^{(2)}$$

where $Z \sim Ber(\theta)$. Convince yourself that $X \sim P_X$.

b. Does $Z - X - Y$ form a Markov chain? Namely, verify that for all x, y, z,

$$P(z, y|x) = P(z|x)P(y|x).$$

c. Show that $I(Z;Y) = 0$. *Hint* – Apply chain rule to $I(X, Z; Y)$

d. Conclude that the capacity-achieving output distribution is unique.