Exercise 8.1 Deterministic Channel (EE5139)

Consider a memoryless channel that takes pairs of bits as input and produces two bits as output as follows: 00 → 01, 01 → 10, 10 → 11, 11 → 00 (to read: input → output). Let (X_1, X_2) denote the two input bits and (Y_1, Y_2) the two output bits.

a.) Calculate the mutual information $I(X_1, X_2; Y_1, Y_2)$ for a given joint PMF of the four pairs of input bits. You can express your answer in terms of

$$\begin{aligned} p_{00} &= \Pr(X_1 = 0, X_2 = 0) \\ p_{10} &= \Pr(X_1 = 1, X_2 = 0) \\ p_{01} &= \Pr(X_1 = 0, X_2 = 1) \\ p_{11} &= \Pr(X_1 = 1, X_2 = 1) \end{aligned}$$

Solution: Since the channel is deterministic, we have

$$I(X_1, X_2; Y_1, Y_2) = H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2) = H(Y_1, Y_2) - 0 = H(Y_1, Y_2).$$

Given input distribution $(p_{00}, p_{01}, p_{10}, p_{11})$, the output distribution can be expressed as $(p_{11}, p_{00}, p_{01}, p_{10})$. Thus,

$$I(X_1, X_2; Y_1, Y_2) = H(p_{11}, p_{00}, p_{01}, p_{10}) = -\sum_{i,j\in\{0,1\}} p_{ij} \log p_{ij}.$$

b.) Show that the channel mutual information is 2 and indicate the units.

Solution: The maximizing input distribution is clearly uniform, *i.e.*,

$$p^\star_{ij} = 1/4 \quad \forall\, (i, j) \in \{0, 1\}^2.$$

Thus,

$$C = H(p^\star_{11}, p^\star_{00}, p^\star_{01}, p^\star_{10}) = 4 \cdot (-\frac{1}{4} \log_2 \frac{1}{4}) = 2 \text{ bits per channel use.}$$

c.) Show that, surprisingly, $I(X_1; Y_1) = 0$ for the capacity-achieving distribution of the input you derived in part (b) (that is, information is only transferred by considering both bits). **Hint:** Find the joint pmf of X_1 and Y_1.

Solution: Note that

$$Y_1 = X_1 \oplus X_2.$$

We now calculate the conditional probability $p_{Y_1|X_1}$ directly. Consider

$$\begin{aligned} p_{Y_1|X_1}(y_1|0) &= \sum_{x_2} p_{Y_1|X_1X_2}(y_1|0, x_2) p_{X_2|X_1}(x_2|0) \\ &= \sum_{x_2} p_{Y_1|X_1X_2}(y_1|0, x_2) p_{X_2}(x_2) \end{aligned}$$

For $y_1 = 0$, we have

$$\begin{aligned} p_{Y_1|X_1}(0|0) &= \frac{1}{2} p_{Y_1|X_1X_2}(0|0, 0) + \frac{1}{2} p_{Y_1|X_1X_2}(0|0, 1) \\ &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0. \end{aligned}$$

Hence,

$$p_{Y_1|X_1}(y_1|0) = \begin{cases} 1/2 & y_1 = 0 \\ 1/2 & y_1 = 1 \end{cases}$$

Similarly,

$$p_{Y_1|X_1}(y_1|0) = \begin{cases} 1/2 & y_1 = 0 \\ 1/2 & y_1 = 1 \end{cases}$$

Therefore, using the capacity achieving distribution, the sub-channel is completely noisy, and $I(X_1, Y_1) = 0$.

Exercise 8.2 Symmetric Channel (all)

For two positive integers k and m, let $(k \bmod m)$ be the *remainder* when k is divided by m. Find the capacity of the m-input discrete memoryless channel in which

$$Y = (X + Z) \mod m,$$

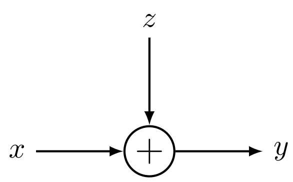
where $X \in \{0, 1, \ldots, m-1\}$, $\Pr[Z = 1] = \frac{3}{4}$, and $\Pr[Z = 0] = \frac{1}{4}$.
Solution: The capacity is $\log_2 m - h(1/4)$. Consider

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - h(1/4).$$

Note that $H(Y)$ is maximized at the value $\log_2 m$ and this is achievable using the uniform input distribution $p(x) = 1/m$ for all $x \in \{0, 1, \ldots, m-1\}$.

Exercise 8.3 Additive noise channel (EE5139)

Find the channel capacity of the following discrete memoryless channel:



where $\Pr[Z = 0] = \Pr[Z = a] = \frac{1}{2}$. The alphabet for x is $\mathcal{X} = \{0, 1\}$. Assume that Z is independent of X. Observe that the channel capacity depends on the value of a.
Solution: We can identify two different values for the capacity of the channel according to the value of a.
When $a \neq \pm 1$, the outputs of the channel are non-overlapping. In such cases the channel is a noisy channel with non-overlapping outputs. It is known that the capacity of such a channel is 1 bit since:

$$C = \max I(X;Y) = \max H(Y) - H(Y|X) = \max H(Y) = 1$$

where the third equality follows from the non-overlapping fact.
When $a = 1$ or -1, the outputs of the channel are overlapped. If $a = 1$, Y can be 1 for either inputs, but is 0 (or 2) only if $X = 0$ (or 1). If $a = -1$, Y can be 0 for either inputs, but is -1 (or 1) only if $X = 0$ (or 1). We compute the channel capacity for $a = 1$. By symmetry, the capacity is same when $a = -1$.

$$\begin{aligned} H(Y) &= -p_Y(0)\log_2 p_Y(0) - p_Y(1)\log_2 p_Y(1) - p_Y(2)\log_2 p_Y(2) \\ &= -\frac{2}{4}\log_2\frac{1}{4} - \frac{1}{2}\log_2\frac{1}{2} = 1 + \frac{1}{2} \end{aligned}$$

and

$$H(Y|X) = -p_X(0)H(Y|X=0) - p_X(1)H(Y|X=1) = \frac{1}{2} + \frac{1}{2} = 1$$

Thus,

$$C = \max I(X;Y) = \max H(Y) - H(Y|X) = 1 + \frac{1}{2} - 1 = \frac{1}{2}.$$

Exercise 8.4 Channel Mutual Information (EE5139)

Let X and Z be independent random variables taking values on $\{1,\ldots,n\}$ and $\{0,1\}$, respectively, with $p_X(i)=q_i$ (for each i) and $p_Z(1)=p$. Define the random variable $Y := X \cdot Z$.

a.) Write $H(Y)$ in terms of $H(X)$ and $H(Z)$.

Solution: Note that the pmf of Y can be written as

$$P_Y(y)=\begin{cases}1-p, & y=0\\ pq_i, & y=1,2,\ldots,n\\ 0, & \text{otherwise.}\end{cases}$$

Thus,

$$\begin{aligned}H(Y)&=(1-p)\log\frac{1}{1-p}+\sum_{i=1}^{n}pq_i\log\frac{1}{pq_i}\\ &=(1-p)\log\frac{1}{1-p}-p\log p\sum_{i=1}^{n}q_i-p\sum_{i=1}^{n}q_i\log q_i\\ &=H(Z)+pH(X).\end{aligned}$$

b.) Find p and $\boldsymbol{q}=(q_1,\ldots,q_n)$ that maximize $H(Y)$.

Solution: Without loss of generality, we measure in nats in this question. Note that $H(Y)=h(p)+p\cdot H(X)$, where h is the binary entropic function (in nats). One must have $H(Y)\le h(p)+p\log n$ since $H(X)\le\log n$, and the equality is attained when $\boldsymbol{q}=(\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n})$. Taking derivatives of $h(p)+p\log n$ with respect to p, we have

$$\frac{\mathrm{d}}{\mathrm{d}p}H(Y)=\log\frac{1-p}{p}+\log n,$$

which is 0 when $p=p^\star=\frac{n}{n+1}$, positive when $p<p^\star$, and negative when $p>p^\star$. As a result,

$$h(p)+p\log n\le h(\frac{n}{n+1})+\frac{n}{n+1}\log n.$$

Thus,

$$H(Y)\le h(\frac{n}{n+1})+\frac{n}{n+1}\log n$$

with equality when $\boldsymbol{q}=(\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n})$ and $p=\frac{n}{n+1}$.

c.) Suppose X and Y are input and output of a DMC channel. For a fixed $p\in[0,1]$, what is the channel mutual information $I(p)$?

Solution:

$$\begin{aligned}I(p):&=\max_{\boldsymbol{q}}I(X:Y)\\ &=\max_{\boldsymbol{q}}H(Y)-H(Y|X)\\ &=\max_{\boldsymbol{q}}H(Z)+pH(X)-H(XZ|X)\\ &=\max_{\boldsymbol{q}}H(Z)+pH(X)-H(Z)\\ &=\max_{\boldsymbol{q}}pH(X)\\ &=p\log n.\end{aligned}$$

Exercise 8.5 Using two channels at once (EE6139)

Consider two discrete memoryless channels $(X_1, p(y_1|x_1), Y_1)$ and $(X_2, p(y_2|x_2), Y_2)$ with capacities C_1 and C_2, respectively. A new channel $(X_1 \times X_2, p(y_1|x_1) \times p(y_2|x_2), Y_1 \times Y_2)$ is formed in which $x_1 \in X_1$ and $x_2 \in X_2$ are sent simultaneously, resulting in y_1, y_2. Find the channel mutual information of this channel.

Solution: Firstly, note that the new channel has input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$, output alphabet $\mathcal{Y}_1 \times \mathcal{Y}_2$ and channel law

$$P_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2) = P_{Y_1|X_1}(y_1|x_1)P_{Y_2|X_2}(y_2|x_2)$$

This is a *product channel.* So we must find a input distribution on $\mathcal{X}_1 \times \mathcal{X}_2$ that maximizes $I(X_1X_2; Y_1Y_2)$. Notice that since the joint distribution of (X_1, X_2, Y_1, Y_2) factorizes as

$$P_{X_1X_2}(x_1, x_2)P_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2) = P_{X_1X_2}(x_1, x_2)P_{Y_1|X_1}(y_1|x_1)P_{Y_2|X_2}(y_2|x_2)$$

we have $Y_1 - X_1 - X_2 - Y_2$ forming a Markov chain.

Let us first find an upper bound on the mutual information. We have

$$\begin{aligned}
I(X_1X_2; Y_1Y_2) &= H(Y_1Y_2) - H(Y_1Y_2|X_1X_2) \\
&= H(Y_1Y_2) - H(Y_1|X_1X_2) - H(Y_2|X_1X_2) \\
&= H(Y_1Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \\
&\leq H(Y_1) - H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2)
\end{aligned}$$

where the first and second inequalities follow from the Markov chain $Y_1 - X_1 - X_2 - Y_2$. Note that equality in the above chain of inequalities holds when X_1 and X_2 are independent. Hence,

$$\begin{aligned}
C :&= \max_{p_{X_1X_2}} I(X_1X_2; Y_1Y_2) \\
&\leq \max_{p_{X_1X_2}} I(X_1; Y_1) + I(X_2; Y_2) \\
&= \max_{p_{X_1}} I(X_1; Y_1) + \max_{p_{X_2}} I(X_2; Y_2) \\
&= C_1 + C_2.
\end{aligned}$$

with equality iff $P_{X_1X_2} = P^\star_{X_1}P^\star_{X_2}$ where $P^\star_{X_j}$ maximizes the mutual information of the j-th channel.

Thus the input distribution that maximizes $I(X_1X_2; Y_1Y_2)$ is of the product form $P^\star_{X_1X_2} = P^\star_{X_1}P^\star_{X_2}$ where the constituent distributions $\{P^\star_{X_j}\}_{j=1,2}$ are those maximizing the mutual information of the j-th channel. The channel mutual information, *i.e.*, the maximal value of $I(X_1X_2; Y_1Y_2)$, is $C_1 + C_2$.

Exercise 8.6 Concatenation of channels (EE6139)

We concatenate n binary symmetric channels as depicted below.

$$X_0 \longrightarrow \boxed{\mathsf{BSC}_1} \longrightarrow X_1 \longrightarrow \boxed{\mathsf{BSC}_2} \longrightarrow X_2 \longrightarrow \cdots \longrightarrow \boxed{\mathsf{BSC}_n} \longrightarrow X_n$$

Let the crossover probability of all of the BECs to be p. Show that the concatenated channel is equivalent to a BSC with crossover probability

$$\frac{1}{2}\left[1 - (1 - 2p)^n\right]$$

and show that $\lim_{n\to\infty} I(X_0; X_n) = 0$ regardless of the distribution of X_0.

Solution: We use mathematical induction to show the the concatenated channel to be BSC($P_e^{(n)}$), where

$$P_e^{(n)} := \frac{1}{2}\left[1-(1-2p)^n\right]$$

for each $n \in \mathbb{N}$.
When $n = 1$, the concatenated channel is BSC(p)= BSC($P_e^{(1)}$) since $P_e^{(1)} = \frac{1}{2}(1-(1-2p)) = p$.
We suppose that for $n = k$, the the concatenated channel is BSC($P_e^{(k)}$). Then, for $n = k+1$,

$$\begin{aligned}
W(0|0) &= W_{\mathrm{BSC}(p)}(0|0) \cdot W_{\mathrm{BSC}(P_e^{(k)})}(0|0) + W_{\mathrm{BSC}(p)}(0|1) \cdot W_{\mathrm{BSC}(P_e^{(k)})}(1|0) \\
&= (1-p) \cdot \frac{1}{2}\left[1+(1-2p)^k\right] + p \cdot \frac{1}{2}\left[1-(1-2p)^k\right] \\
&= \frac{1}{2}\left[1+(1-2p)^{k+1}\right] \\
W(1|0) &= 1 - W(0|0) = \frac{1}{2}\left[1-(1-2p)^{k+1}\right] = P_e^{(k+1)} \\
W(1|1) &= \cdots = \frac{1}{2}\left[1+(1-2p)^{k+1}\right] \\
W(0|1) &= 1 - W(1|1) = \frac{1}{2}\left[1-(1-2p)^{k+1}\right] = P_e^{(k+1)}
\end{aligned}$$

which, by definition, shows that the concatenated channel is BSC($P_e^{(k+1)}$).
Therefore, by axiom of choice, the concatenation of the n BSC(p) is BSC($P_e^{(n)}$) for any $n \in \mathbb{N}$.
On the other hand, $P_e^{(n)} \to 1/2$ as $n \to \infty$, namely $\Pr(X_n = 0) = \Pr(X_n = 1) \to \frac{1}{2}$. Thus

$$\lim_{n\to\infty} I(X_0; X_n) = \lim_{n\to\infty} (H(X_n) - H(X_n|X_0)) = \lim_{n\to\infty} (1-1) = 0.$$