EE276: Homework #4

Due on Friday Feb 9, 5pm

1. **Channel capacity.**
   Find the capacity of the following channels with given probability transition matrices, where the element p_{ij} of the matrix represents $p(y = j|x = i)$:

   (a) $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$

   $$p(y|x) = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

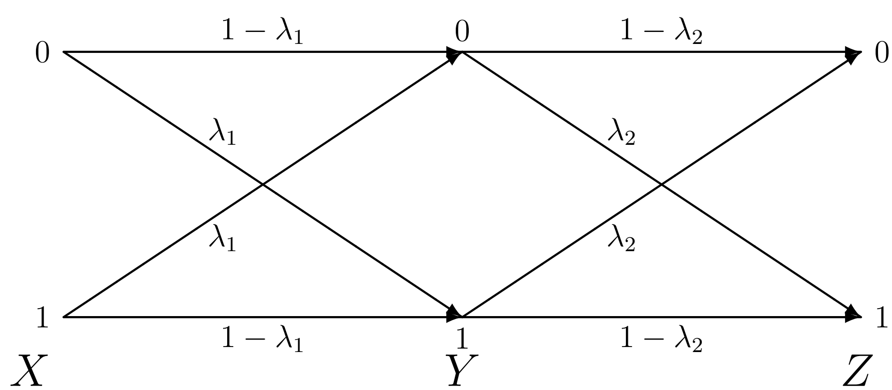
   (b) $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$

   $$p(y|x) = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

   (c) $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ (the Z-channel)

   $$p(y|x) = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix}$$

2. **Cascaded BSCs.**
   Let $p_1(y|x)$ and $p_2(z|y)$ be binary symmetric channels with crossover probabilities λ_1 and λ_2 respectively.



   (a) What are the capacities C_1 and C_2 of $p_1(y|x)$ and $p_2(z|y)$, respectively?

   (b) We now cascade these channels as shown in the figure. Thus $p_3(z|x) = \sum_y p_1(y|x)p_2(z|y)$. What is the capacity C_3 of $p_3(z|x)$?

   (c) What is the capacity of the cascade in part b) if the receiver can view *both* Y and Z, i.e., the channel is defined by the transition $p((y, z)|x) = p_1(y|x)p_2(z|y)$?

   (d) Now let us actively intervene between channels 1 and 2, rather than passively transmit y^n. What is the capacity of channel 1 followed by channel 2 if you are allowed to decode the output y^n of channel 1 and then reencode it as $\tilde{y}^n$ for transmission over channel 2? (Think $W \longrightarrow x^n(W) \longrightarrow y^n \longrightarrow \tilde{y}^n(y^n) \longrightarrow z^n \longrightarrow \hat{W}$

where W represents the information bits.)
Note: In part (d), think of capacity as the maximum achievable rate for reliable communication, not as the maximum mutual information, and argue about both the achievability and converse.

3. **Choice of channels.**
Let $\mathcal{C}_1 \equiv (\mathcal{X}_1, p_1(y_1|x_1), \mathcal{Y}_1)$ and $\mathcal{C}_2 \equiv (\mathcal{X}_2, p_2(y_2|x_2), \mathcal{Y}_2)$ be two channels with capacities C_1, C_2 respectively. Assume the input and output alphabets for the two channels are disjoint, i.e., $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$ and $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$. Consider a channel $\mathcal{C}$, which is a union of the two channels $\mathcal{C}_1$, $\mathcal{C}_2$, where at each time, one can send a symbol over $\mathcal{C}_1$ or over $\mathcal{C}_2$ but not both. In this problem, we calculate the capacity of $\mathcal{C}$ in terms of the C_1 and C_2.

Define θ as an indicator random variable denoting whether channel $\mathcal{C}_1$ or $\mathcal{C}_2$ is used in a particular transmission. Let X and Y denote the input and output for the channel $\mathcal{C}$. Note that X follows a mixture distribution over the (disjoint) alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$:

$$X = \begin{cases} X_1 & \text{with probability } \alpha \\ X_2 & \text{with probability } 1-\alpha \end{cases}$$

where X_1 and X_2 are random variables taking values in $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively.

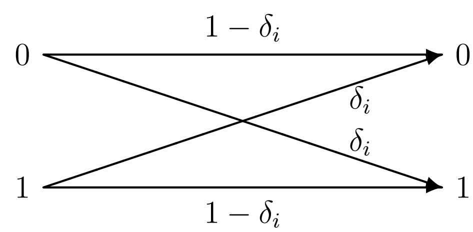
(a) Argue that $H(\theta|X) = H(\theta|Y) = 0$.

(b) Show that
$$I(X;Y) = h_2(\alpha) + \alpha I(X_1;Y_1) + (1-\alpha) I(X_2;Y_2)$$
where Y_1 and Y_2 are the channel outputs when X_1 and X_2 are transmitted through $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively.
Hint: Start with $I(Y;X,\theta) = I(Y;\theta) + I(Y;X|\theta) = I(Y;X) + I(Y;\theta|X)$.

(c) Let C be the capacity of the channel $\mathcal{C}$. Maximize the expression in (b) over α, P_{X_1}, P_{X_2} to show that $2^C = 2^{C_1} + 2^{C_2}$.

(d) Let $C_1 = C_2$. Then show that $C = C_1 + 1$ and give an intuitive explanation.

4. **Time-varying channels.**
Consider a time-varying discrete *memoryless* channel. Let $Y_1, Y_2, \ldots, Y_n$ be conditionally independent given $X_1, X_2, \ldots, X_n$, with conditional distribution given by $p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n p_i(y_i|x_i)$ (where $p_i(y_i|x_i)$ is a $BSC(\delta_i)$ as shown in figure). Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$.



In this problem, we show that

$$\max_{P_{\mathbf{X}}} I(\mathbf{X};\mathbf{Y}) = \sum_{i=1}^{n}(1 - h(\delta_i))$$

(a) Show that $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n}(1 - h(\delta_i))$ for any $P_{\mathbf{X}}$ (It may be helpful to wait until Tuesday's lecture, and to use a chain of inequalities similar to the channel coding converse proof).

(b) Find a distribution over $\mathbf{X}$ for which $I(\mathbf{X};\mathbf{Y}) = \sum_{i=1}^{n}(1 - h(\delta_i))$.

5. **Maximum Differential Entropy** (It may also be helpful to wait until Tuesday's lecture to solve this problem, even though the necessary concepts have been defined last Thursday.)

   (a) Show that among all distributions supported in an interval [a,b], the uniform distribution maximizes differential entropy.

   (b) Let X be a continuous random variable with $\mathcal{E}[X^4] \leq \sigma^4$ and Y be a continuous random variable with a probability density function $g(y) = c\exp\left(-\frac{y^4}{4\sigma^4}\right)$ where $c = \frac{1}{\int_{-\infty}^{\infty}\exp\left(-\frac{y^4}{4\sigma^4}\right)dy}$. Show that

$$h(X) \leq h(Y)$$

   with equality if and only if X is distributed as Y.
   [Hint: you can use the fact that $\mathcal{E}[Y^4] = \sigma^4$.]