EEE 551 Information Theory (Spring 2022)

Homework 3

Due: March 2, 2022

---

1. Problem 7.5 from Cover-Thomas: *Using two channels at once.* Consider two discrete memoryless channels $(\mathcal{X}_1, p(y_1|x_1), \mathcal{Y}_1)$ 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.

2. *Channel capacity.* Calculate the capacity of the following channels with the given probability transition matrices. Recall that each row of the matrix corresponds to an input symbol, and each column corresponds to an output symbol.

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

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

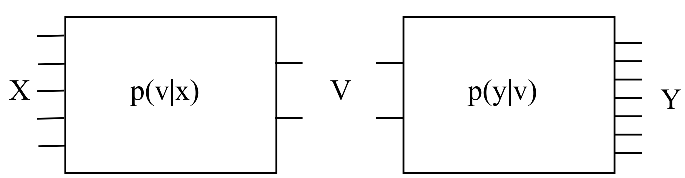
   (b) $\mathcal{X} = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $\mathcal{Y} = \{0, 1\}$

   $$p(y|x) = \begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \\ 0.3 & 0.7 \\ 0.4 & 0.6 \\ 0.5 & 0.5 \\ 0.6 & 0.4 \\ 0.7 & 0.3 \\ 0.8 & 0.2 \\ 0.9 & 0.1 \end{bmatrix}$$

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

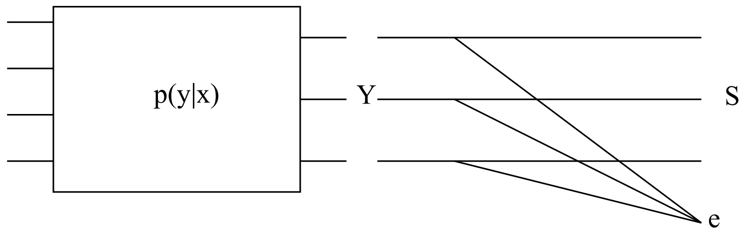
   $$p(y|x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

3. Problem 7.25 from Cover-Thomas: *Bottleneck channel.* Suppose a signal $X \in \mathcal{X} = \{1, 2, \ldots, m\}$ goes through an intervening transition $X \longrightarrow V \longrightarrow Y$:



where $\mathcal{X} = \{1, 2, \ldots, m\}$, $\mathcal{Y} = \{1, 2, \ldots, m\}$, and $\mathcal{V} = \{1, 2, \ldots, k\}$. Here $p(v|x)$ and $p(y|v)$ are arbitrary and the channel has transition probability $p(y|x) = \sum_v p(v|x)p(y|v)$. Show that $C \leq \log k$.

4. Problem 7.27 from Cover-Thomas: *Erasure channel.* Let $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$ be a discrete memoryless channel with capacity C. Suppose that this channel is cascaded immediately with an erasure channel $\{\mathcal{Y}, p(s|y), \mathcal{S}\}$ that erases α of its symbols.



Specifically, $\mathcal{S} = \{y_1, y_2, \ldots, y_m, e\}$ and

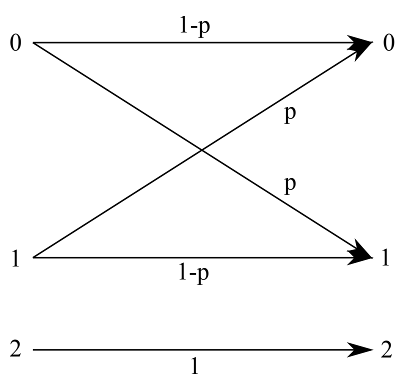
$$\Pr\{S = y|X = x\} = (1 - \alpha)p(y|x), \qquad y \in \mathcal{Y},$$
$$\Pr\{S = e|X = x\} = \alpha.$$

Determine the capacity of this channel. [You may write your answer in terms of the capacity C of the original channel $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$.]

5. Problem 7.28 from Cover-Thomas, parts (a) and (c): *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 $(\mathcal{X}_2, p_2(y_2|x_2), \mathcal{Y}_2)$ [with capacities C_1 and C_2 respectively], 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.

   (a) Show that $2^C = 2^{C_1} + 2^{C_2}$.

   (c) Use the above result to calculate the capacity of the following channel.



6. *Polar codes for the BSC.* Consider using a polar code for the binary symmetric channel (rather than the binary erasure channel, which we covered in class).

   (a) Consider a BSC(p). For the basic (i.e., single generation) polar transform, calculate the channel capacities of W^- and W^+. That is, calculate

$$I(W^-) = I(U_1; Y_1, Y_2)$$
$$I(W^+) = I(U_2; Y_1, Y_2, U_1).$$

Confirm that $I(W^-) + I(W^+) = 2I(W)$, and $I(W^-) < I(W) < I(W^+)$, where $I(W)$ is the capacity of the original channel.

(b) The following table lists the features of three polar encoders. Given is the generation number (a polar code of genertaion t has blocklength $n = 2^t$), and which U bits are frozen to 0 and which are used as message bits.

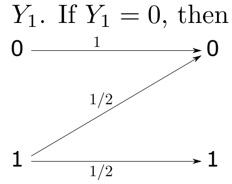
Generation	Frozen Bits	Message Bits
2	U_1, U_2, U_3	U_4
3	U_1, U_2, U_3, U_4, U_6	U_5, U_7, U_8
3	U_1, U_2, U_3, U_4, U_5	U_6, U_7, U_8

For each of these polar codes:

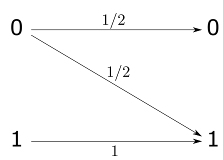
i. List all the codewords associated with the code.

ii. In the following, a "single bit flip" means an error pattern in which exactly 1 bit is changed. For a blocklength n code, there are n different single bit flip error patterns. For each code, determine which of the following is true (and explain why):
(A) The code cannot correct any single bit flip
(B) The code can correct some (but not all) single bit flips
(C) The code can correct any single bit flip

7. *Channel with memory and feedback.* Consider the following binary-input binary-output channel *with memory*; meaning that the output at a given time depends on not only the input at that time but also previous inputs. At time 1, the channel acts like a BSC(1/2); that is, the output Y_1 has a uniform distribution on $\{0, 1\}$ independently of the input. At time 2, the behavior of the channel depends on Y_1. If $Y_1 = 0$, then the channel at time 2 behaves like a Z-channel, as shown below:



Recall that the optimal input distribution for a Z-channel sends 1 with probability 2/5. If $Y_1 = 1$, then the channel at time 2 behaves like an inverted Z-channel, as shown below:



For every two subsequent transmissions, the channel replicates these two steps independently (i.e. the odd transmissions are BSC(1/2), the even transmissions are Z-channel or inverted Z-channel depending on the previous output).

(a) Find the capacity of this channel. *Hint:* Considering two transmissions at a time, this is a standard discrete memoryless channel.

(b) Find the capacity of this channel with feedback. In particular, before the second transmission, the encoder knows the output of the first transmission.

8. *Jointly typical sequences.* Let X and Y be jointly binary random variables. Let $X \sim \text{Bern}(1/2)$ be the input to a binary symmetric channel with crossover probability p and Y the output. That is, the joint distribution $p(x, y)$ is given by

$X \setminus Y$	0	1
0	$\frac{1}{2}(1-p)$	$\frac{1}{2}p$
1	$\frac{1}{2}p$	$\frac{1}{2}(1-p)$

Note that the marginal distribution of Y is also $\text{Bern}(1/2)$. Recall that the jointly typical set $A_\epsilon^{(n)}$ is defined as the set of pairs (x^n, y^n) that satisfy the three conditions

$$\left| -\frac{1}{n} \log p(x^n) - H(X) \right| \le \epsilon,$$

$$\left| -\frac{1}{n} \log p(y^n) - H(Y) \right| \le \epsilon,$$

$$\left| -\frac{1}{n} \log p(x^n, y^n) - H(X, Y) \right| \le \epsilon.$$

For a given pair of sequences (x^n, y^n), let k be the number of places in which the sequence x^n differs from y^n (i.e. the number of times a bit flip occurred). Prove that there exist numbers k_{min} and k_{max} where $(x^n, y^n) \in A_\epsilon^{(n)}$ if and only if

$$k_{\text{min}} \le k \le k_{\text{max}}.$$

What are these numbers k_{min} and k_{max}?