CS258 Information Theory Homework 6

Zhou Litao 518030910407 F1803016

April 5, 2020

Exercise 1 (BSC) Calculate the channel capacity of BSC.

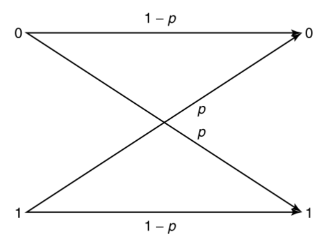

Figure 1: Binary Symmetric Channel

Solution.

$$
\begin{aligned}
C &= \max_{p(x)} I(X;Y) \\
&= \max_{p(x)} (H(Y) - H(Y|X)) \\
&= \max_{p(x)} (H(Y) - \sum p(x) H(Y|X=x)) \\
&= \max_{p(x)} (H(Y) - \sum p(x) H(p)) \\
&= \max_{p(x)} (H(Y) - H(p)) = \log 2 - H(p)
\end{aligned} \tag{1}
$$

The maximal value can be obtained when X and Y are uniformly distributed. □

Exercise 2 (BSC) Calculate the channel capacity of BEC.

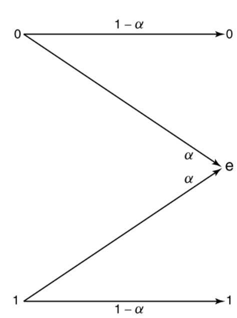

Figure 2: Binary Erasure Channel

Solution.

$$\begin{aligned}
C &= \max_{p(x)} I(X;Y) \\
&= \max_{p(x)} (H(Y) - H(Y|X)) \\
&= \max_{p(x)} (H(Y) - H(\alpha))
\end{aligned} \tag{2}$$

Let $\Pr(X=1) = \pi$, then

$$\begin{aligned}
H(Y) &= H((1-\pi)(1-\alpha), \alpha, \pi(1-\alpha)) \\
&= H(\alpha) + (1-\alpha)H(\pi) \\
C &= \max_{p(x)} (H(Y) - H(\alpha)) \\
&= \max_{\pi} ((1-\alpha)H(\pi) + H(\alpha) - H(\alpha)) \\
&= \max_{\pi} (1-\alpha)H(\pi) = 1-\alpha
\end{aligned} \tag{3}$$

The maximal value can be obtained when X is uniformly distributed. □

Exercise 3 (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 $\mathcal{Y}_1, \mathcal{Y}_2$. Find the capacity of this channel.

Solution. By condition we know that $p(y_1, y_2|x_1, x_2) = p(y_1|x_1) \times p(y_2|x_2)$. It follows by definition that

$$H(Y_1, Y_2|X_1, X_2) = H(Y_1|X_1) + H(Y_2|X_2). \tag{4}$$

Therefore,

$$\begin{aligned}
I(X_1, X_2; Y_1, Y_2) &= H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2) \\
&= H(Y_1, Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \\
&\le H(Y_1) + H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \\
&= I(Y_1; X_1) + I(Y_2; X_2) \\
&\le C_1 + C_2
\end{aligned} \tag{5}$$

Hence $C = C_1 + C_2$. The equality holds when $p(x_1, x_2) = p^*(x_1)p^*(x_2)$, where $p^*(x_1)$ and $p^*(x_2)$ are the optimal distribution corresponding to the original channel capacities. □

Exercise 4 (Z-channel) The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

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

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

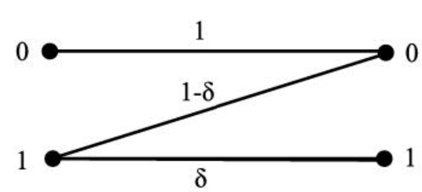

Figure 3: Z-Channel

Solution. Assume that $\Pr(X = 1) = \pi$, given the transition matrix Q, we have that $\Pr(Y = 1) = \frac{\pi}{2}$, $\Pr(Y = 1) = 1 - \frac{\pi}{2}$.

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(\frac{\pi}{2}) - \pi H(\frac{1}{2}) - (1-\pi)H(1) \\ &= -\frac{\pi}{2}\log(\frac{\pi}{2}) - \left(1-\frac{\pi}{2}\right)\log\left(1-\frac{\pi}{2}\right) - \pi \end{aligned} \tag{6}$$

Since the function is a concave function, we find its maximal value by taking the derivative.

$$\begin{aligned} \frac{dI(X;Y)}{d\pi} &= -\frac{1}{2}\log(\frac{\pi}{2}) - \frac{1}{2\ln 2} + \frac{1}{2\ln 2} + \left(1-\frac{1}{2}\right)\log\left(1-\frac{\pi}{2}\right) - 1 \\ &= \frac{1}{2}\log\frac{2-\pi}{\pi} - 1 := 0 \end{aligned} \tag{7}$$

It follows that the mutual information is at its maximum when $\pi = \frac{2}{5}$. The channel capacity is $C \approx 0.32193$. □

Exercise 5 (Erasures and errors in a binary channel) Consider a channel with binary inputs that has both erasures and errors. Let the probability of error be ϵ and the probability of erasure be α, so the channel is follows:

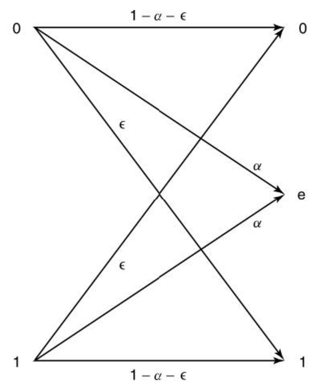

Figure 4: Erasures and Errors

Find the capacity of this channel.

Solution. The transition matrix is as follows.

X \ Y	0	e	2
0	$1-\alpha-\epsilon$	α	ϵ
1	ϵ	α	$1-\alpha-\epsilon$

The matrix is row symmetric, we have that

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(\alpha, \epsilon, 1-\alpha-\epsilon) \tag{8}$$

Furthermore, note that the distribution of Y in terms of the distribution of X is symmetric, i.e.

$$H(Y)|_{\Pr(X=0)=\pi} = H(Y)|_{\Pr(X=0)=1-\pi} \tag{9}$$

and that the entropy function is a concave function. It follows that the maximal value must be obtained at $p(x) = \frac{1}{2}$ for $x = 0, 1$.

$$C = H(\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}) - H(\alpha, \epsilon, 1-\alpha-\epsilon) \tag{10}$$

□

Exercise 6 (Additive noise channel) Find the channel capacity of the following discrete memoryless channel:

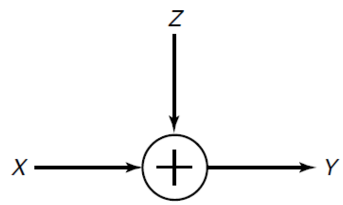

Figure 5: Discrete Memoryless Channel

where $\Pr\{Z=0\} = \Pr\{Z=a\} = \frac{1}{2}$. The alphabet for x is $\mathbf{X} = \{0,1\}$. Assume that Z is independent of X. Observe that the channel capacity depends on the value of a.

Solution. The condition implies that $a \neq 0$, but after the plus operation the distribution of Y may vary. Hence we discuss the value of a in several cases.

1. $a = \pm 1$, the result of $X + a$ and $X + 0$ may overlap. We take the case of $a = 1$ as example. The transition matrix is as follows.

X \ Y	0	1	2
0	$\frac{1}{2}$	$\frac{1}{2}$	0
1	0	$\frac{1}{2}$	$\frac{1}{2}$

   The matrix is row symmetric, assume that $\Pr(X=0) = \pi$we have that

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(\frac{1}{2}\pi, \frac{1}{2}, \frac{1}{2}\pi) - H(\frac{1}{2}) \\ &\leq H(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\pi) - H(\frac{1}{2}) = \frac{1}{2} \end{aligned} \tag{11}$$

   $C = \frac{1}{2}$. The maximal value is obtained at $\pi = \frac{1}{2}$.

2. $a \neq \pm 1$. Then the output Y will not overlap for different X. Hence Y is a function of X. We have

$$C = \max I(X;Y) = \max H(X) = 1 \tag{12}$$

   The maximal value is obtained when $p(x) = \frac{1}{2}$ for $x = 0, 1$.

□

Exercise 7 (Channel capacity) Consider the discrete memoryless channel $Y = X + Z(\text{mod}11)$, where

$$Z = \begin{pmatrix} 1, & 2, & 3 \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{pmatrix}$$

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

1. Find the capacity.
2. What is the maximizing $p^*(x)$?

Solution. Note that the transition matrix of this channel is

$$
\begin{array}{c|ccccccc}
X \backslash Y & 0 & 1 & 2 & \dots & 9 & 10 \\
\hline
0 & 0 & \frac{1}{3} & \frac{1}{3} & \dots & 0 & 0 \\
1 & 0 & 0 & \frac{1}{3} & \dots & 0 & 0 \\
2 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
9 & \frac{1}{3} & \frac{1}{3} & 0 & \dots & 0 & \frac{1}{3} \\
10 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \dots & 0 & 0
\end{array}
$$

Note that the matrix is both column and row symmetric. Therefore we have

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - H(\mathbf{r}) \\
&\leq \log|\mathcal{Y}| - H(\mathbf{r}) = \log 11 - \log 3 = \log\frac{11}{3}
\end{aligned} \tag{13}
$$

$C = \log\frac{11}{3}$. The maximal value is obtained when $p^*(x) = \frac{1}{11}$ for every $x \in \{0, 1, \dots, 10\}$ □

Exercise 8 (Zero-error capacity) A channel with alphabet {0,1,2,3,4} has transition probabilities of the form

$$
p(y|x) = \begin{cases} 1/2 & \text{if } y = x \pm 1 \bmod 5 \\ 0 & \text{otherwise} \end{cases}
$$

(a) Compute the capacity of this channel in bits.

(b) The zero-error capacity of a channel is the number of bits per channel use that can be transmitted with zero probability of error. Clearly, the zero-error capacity of this pentagonal channel is at least 1 bit (transmit 0 or 1 with probability 1/2). Find a block code that shows that the zero-error capacity is greater than 1 bit. Can you estimate the exact value of the zero-error capacity? (Hint: Consider codes of length 2 for this channel.)

Solution.

1. Note that the transition matrix is row and column symmetric. It follows that

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - H(\frac{1}{2}) \\
&\leq \log|\mathcal{Y}| - H(\frac{1}{2}) = \log 5 - 1 \approx 1.322
\end{aligned} \tag{14}
$$

   $C = 1.322$ bits. The maximal value is obtained when X and Y are uniformly distributed.

2. According to the hint, we try to build a 2-tuple code with the alphabet {0,1,2,3,4}. Note that for every input codeword (X_1, X_2), there are four possibilities of output, namely $((X_1 \pm 1) \bmod 5, (X_1 \pm 1) \bmod 5)$, but the good news is that if we can ensure that there are no other input codewords (and their possible outputs) occupying the 4 entries, we can infer the input determintly.

   Figure 6 gives an example of zero-error codes, where cells identified by different colors represent a codeword. The input codewords are (0,0),(1,2),(2,4),(3,1) and (4,3). Given an output codeword, its corresponding input codeword can be determined directly from the table.

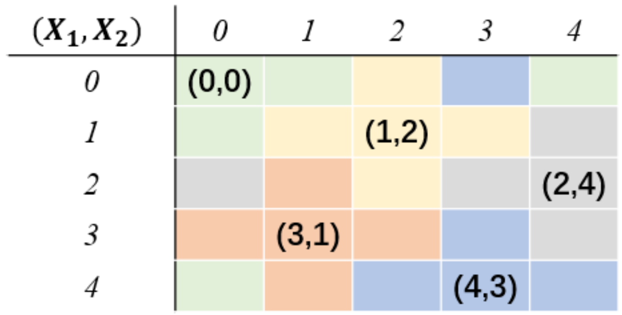

Figure 6: A Construction of Zero-Error Codes

With the codes above, we can find that the number of bits per channel use is $\frac{1}{2}\log 5 > 1$.

$\square$