Telecom ParisTech SOCOM205
Prof. Aslan Tchamkerten March 2017

ASSIGNMENT 4: SOLUTION

Exercise 1 (Additive channel)**.** Consider a channel $Y = Z + X$, where X is a binary input, Y a binary output, and Z, the noise in the channel, a random variable that is independent of X and that takes values $\{0, a\}$ with probabilities $Pr(Z = 0) = 1 - Pr(z = a) = \frac{1}{3}$. Find the capacity of this channel with respect to a.

Solution. We consider two cases:

- If $a \notin \{-1, +1\}$, the possible outputs, $\{0, 1, a, 1 + a\}$, do not overlap, so $H(X|Y) = 0$ and

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(X) = 1,$$

  where the capacity is derived by uniform distribution for input.

- If $a = 1$. the possible outputs are $\{0, 1, 2\}$, with transition probability

$$Q(y|x) = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

  Here, assuming $p(X = 0) = 1 - p(x = 1) = p$, we have

$$H(Y|X) = H(Y|X = 0)p(X = 0) + H(Y|X = 1)p(X = 1) = h_b(\frac{1}{3})$$

$$C = \max_{p(x)} I(X;Y) = \max_{p} H(Y) - h_b(\frac{1}{3}),$$

  and we have

$$p(y = 0) = \frac{p}{3}$$
$$p(y = 1) = \frac{1 + p}{3}$$
$$p(y = 2) = \frac{2 - 2p}{3}$$

  By taking derivative with respect to p, it can be seen that capacity is maximized for $p \simeq 0.54$ and so $C \simeq 0.54$.

- If $a = -1$, similarly we have $C \simeq 0.54$.

□

Exercise 2 (Mod 11 channel)**.** Let $U \in \{0, 1, ..., 10\}$ be the input of a DMC. The output is determined by the formula $V = U + Z$ (mod 11), where Z is independent of U, and

$$P\{Z = 1\} = P\{Z = 2\} = P\{Z = 3\} = \frac{1}{3}.$$

Find the channel capacity.

Solution.

$$\begin{aligned} C &= \max_{p(u)} I(U;V) = \max_{p(u)} H(V) - H(V|U) \\ &\overset{(a)}{=} \max_{p(u)} H(V) - \log(3) \\ &\le \log(11) - \log(3) \end{aligned}$$

where (a) is due to the fact that $H(V|U) = H(Z|U) = H(Z) = \log(3)$.

Notice that with uniform distribution over input (which yields uniform distribution over output) equality can be achieved. So, $C = \log(\frac{11}{3})$.

□

Exercise 3. Find the capacity and optimal input distribution for the three-input, three-output channel whose transition probabilities are given by

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

Solution. Let $X \in \{1,2,3\}$ be the input and $Y \in \{1,2,3\}$ of the channel. By symmetry, in the optimal distribution we have $p(X=2) = p(X=3)$. So, let $p(X=1) = 1-p$ and $p(X=2) = p(X=3) = \frac{p}{2}$, which yields $p(Y=1) = 1-p$ and $p(Y=2) = p(Y=3) = \frac{p}{2}$:

$$\begin{aligned} H(Y|X) &= H(Y|X=1)p(X=1) + H(Y|X=2)p(X=2) + H(Y|X=3)p(X=3) \\ &= 0 + h_b(\frac{1}{3})\frac{p}{2} + h_b(\frac{1}{3})\frac{p}{2} = p \cdot h_b(\frac{1}{3}) \end{aligned}$$

$$H(Y) = -(1-p)\log(1-p) - \frac{p}{2}\log(\frac{p}{2}) - \frac{p}{2}\log(\frac{p}{2}) = -(1-p)\log(1-p) - p\log(p) + p$$

$$C = \max_p I(X;Y) = \max_p -(1-p)\log(1-p) - p\log(p) + p(1 - h_b(\frac{1}{3}))$$

By taking derivative it can be verified that it will be maximized for $0 \simeq 0.514$ which gives $C \simeq 1.04$.

□

Exercise 4 (Cascade of BSCs)**.** Consider the cascade of k binary symmetric channels, each having the same crossover probability p:

1. Show that the binary channel obtained is equivalent to a BSC with crossover probability $\frac{1}{2}(1-(1-2p)^k)$. This shows that the capacity of the cascade channel goes to zero as $k \to \infty$.

2. Suppose now that at the output of each channel arbitrary information processing is allowed. What is the capacity of the cascade channel?

Solution. a. Suppose that the input of channel is X and the output for the cascade of i BSC is Y_i. We show by induction on i that the channel is equivalent to a BSC with crossover probability $\frac{1}{2}(1-(1-2p)^k)$. For $i = 1$ trivially it holds. Suppose for $i = l-1$ holds, now for $i = l$ we have

$$
\begin{aligned}
&Pr(Y_l = 1|X = 0) \\
&= Pr(Y_l = 1|Y_{l-1} = 0, X = 0)Pr(Y_{l-1} = 0|X = 0) + Pr(Y_l = 1|Y_{l-1} = 1, X = 0)Pr(Y_{l-1} = 1|X = 0) \\
&= Pr(Y_l = 1|Y_{l-1} = 0)Pr(Y_{l-1} = 0|X = 0) + Pr(Y_l = 1|Y_{l-1} = 1)Pr(Y_{l-1} = 1|X = 0) \\
&= p(1 - \frac{1}{2}(1 - (1 - 2p)^{l-1})) + (1 - p)\frac{1}{2}(1 - (1 - 2p)^{l-1}) \\
&= \frac{1}{2}(1 - (1 - 2p)^l)
\end{aligned}
$$

b. First note that $X - Y_1 - Y_k$ forms a Markov chain, hence

$$C = \max_{p(x)} I(X; Y_k) \le \max_{p(x)} I(X; Y_1) = C_{\text{BSC}} = 1 - p.$$

Furthermore, $C = 1 - p$, since the rate $1 - p$ is achievable by recovering X at each intermediate node and retransmitting it.

□

Exercise 5 (Repetition Code)**.** The symbols 0 or 1 are to be transmitted through a binary symmetric channel with crossover probability $p < 1/2$. The channel code has two codewords (000) and (111), the code (000) is transmitted if the message is 0, and (111) is transmitted if the message is 1. The decoder is using a majority decision; if the received word has more than one zero, then the decoder output is 0, otherwise it is 1. What are the probabilities of incorrect decoding $P_{e,1}$ and $P_{e,2}$?

Solution. Error happens when more than one bit is changed.

$$P_{e,1} = P_{e,2} = 3p^2(1 - p) + p^3.$$

□

Exercise 6 (Byte channel)**.** The input to a channel Q is a word of 8 bits. The output is also a word of 8 bits. Each time it is used, the channel flips exactly one of the transmitted bits, but the receiver does not know which one. The other seven bits are received without error. All 8 bits are equally likely to be flipped. Derive the capacity of this channel.

Solution. Note that for each input there are 8 equiprobable outputs, so $H(Y|X) = \log 8 = 3$.

$$C = \max_{p(x)} I(X; Y) = \max_{p(x)} H(Y) - H(Y|X) \le 8 - 3 = 5,$$

where inequality is tight for uniform distribution for input which yields uniform distribution for output.

□

Exercise 7 (Zero-error, pentagon channel)**.** Consider the "pentagon" channel with input and output alphabets $\{0, 1, 2, 3, 4\}$ and conditional probabilities given by

$$Q(y|x) = \begin{cases} 1/2 & \text{if} \quad y = x \pm 1 \mod 5 \\ 0 & \text{otherwise.} \end{cases}$$

For instance, symbol "4" becomes "3" or "0" with probability $1/2$.

1. Compute the capacity of this channel.
2. The "zero-error" capacity is the number of information bits per channel use that can be transmitted with exactly zero error probability.
   (a) By considering coding with block length equal to 1, show that the zero-error capacity of this channel is at least 1 bit per channel use.
   (b) Exhibit a code with 5 codewords of length 2 which shows that the zero-error capacity is strictly greater than 1 bit per channel use.

Shannon showed in 1956 that $0.5 \log 5$ bits per channel use is achievable without error over the pentagon channel. But it was more than twenty years after (1979(!)) that Lovàsz showed that $0.5 \log 5$ is indeed the highest rate one can get (*i.e.*, it is the zero-error capacity), even when considering codewords of length ≥ 2.

Solution. 1. Note that for each input there are 2 equiprobable outputs, so $H(Y|X) = \log 2 = 1$.

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X) \leq \log 5 - 1,$$

where inequality is tight for uniform distribution for input which yields uniform distribution for output.

2. (a) One can send the symbols 1 and 2 with probability half. Note that if we send 1, the output is either 2 or 0, and if we send 2, the output is either 1 or 3. So, the input can be decoded always correctly. Using this coding scheme we can send one bit per channel use, so the capacity is at least one.
   (b) We choose the following codewords with equal probabiltiy:

$$\{00, 13, 21, 34, 42\}.$$

   It can be verified that all possible outputs of each codeword is different than anothers, which means $H(X^2|Y^2) = 0$. Using, this coding scheme, we can send $\frac{\log 5}{2} \simeq 1.16$ bit per channel use, so the capacity is strictly greater than one.

□

Exercise 8 (Processing the output)**.** Consider a channel with a conditional probability distribution $\{Q(y|x)\}$. Instead of using directly the output Y, we first process the output as $\tilde{Y} = g(Y)$, and then use $\tilde{Y}$ for decoding. Is this processing useful?

Solution. Due to the Markov chain $X - Y - \tilde{Y}$ and data processing inequality we have $I(X;\tilde{Y}) \leq I(X;Y)$, so

$$\begin{aligned} C_{\text{Processed}} &= \max_{p(x)} I(X;\tilde{Y}) \\ &\leq \max_{p(x)} I(X;Y) = C \end{aligned}$$

which means that processing the output will not increase the capacity of the channel.

□