Chapter 7

Channel Capacity

Exercise 1

A source produces independent, equally probable symbols from an alphabet $\{a_1, a_2\}$ at a rate of one symbol every 3 seconds. These symbols are transmitted over a binary symmetric channel (with error probability $p < 1/2$) which is used once per second by an encoding the source symbol a_1 as 000 and the source symbol a_2 as 111. If in the corresponding 3 second interval of the channel output, any of the sequences $000, 001, 010, 100$ is received, a_1 is decoded. otherwise, a_2 is decoded.

1. For each possible received 3bit sequence find the probability that a_1 came out of source given that received sequence.
2. Find the probability of an incorrect decision (using question 1. is not the easiest way).
3. Show that the above decoding rule minimizes the probability of an incorrect decision.

Answer of exercise 1

1. Let us assume the sequence y was received. Let $\sharp_0 y$ denotes the number of '0' in y and $\sharp_1 y$ analogously for '1'. We want to determine the probability $\mathcal{P}(a_1 \text{ sent}|y \text{ received})$, shortly $\mathcal{P}(a_1|y)$.
   It is easy to see that

   $$\mathcal{P}(a_1|y) = \frac{\mathcal{P}(a_1, y)}{\mathcal{P}(y)} = \frac{\mathcal{P}(y|a_1)\mathcal{P}(a_1)}{\mathcal{P}(y)}.$$

   We know that $\mathcal{P}(a_1) = 1/2$, and $\mathcal{P}(y|a_1)$ is uniquely determined by the number of bits that were flipped. Since a_1 is encoded as 000, the number of bits the channel flips when transforming 000 to y is determined by the Hamming weight of y. We get

   $$\mathcal{P}(y|a_1) = p^{\sharp_1 y}(1-p)^{\sharp_0 y}.$$

It remains to determine $\mathcal{P}(y)$:

$$\begin{aligned}\mathcal{P}(y) =&\mathcal{P}(y|a_1)\mathcal{P}(a_1) + \mathcal{P}(y|a_2)\mathcal{P}(a_2)\\ =&\frac{1}{2}(p^{\sharp_1 y}(1-p)^{\sharp_0 y} + p^{\sharp_0 y}(1-p)^{\sharp_1 y})\end{aligned}$$

Finally,

$$\mathcal{P}(a_1|y) = \frac{p^{\sharp_1 y}(1-p)^{\sharp_0 y}}{p^{\sharp_1 y}(1-p)^{\sharp_0 y} + p^{\sharp_0 y}(1-p)^{\sharp_1 y}}.$$

2. We calculate the respective error probabilities provided the signal a_1 (a_2) was sent

$$\begin{aligned}\lambda_1 &= \sum_{y\in\{011,101,110,111\}} \mathcal{P}(y|a_1) = p^3 + 3p^2(1-p)\\ \lambda_2 &= \sum_{y\in\{100,010,001,000\}} \mathcal{P}(y|a_2) = p^3 + 3p^2(1-p).\end{aligned}$$

Since $\lambda_1 = \lambda_2$, the (arithmetic) average error probability $P_e^{(3)}$ and maximum error probability λ_{max} coincide.

3. We have only two codewords, so the probability that a sequence y decoded as (WLOG) a_1 was decoded incorrectly is proportional to the number of flips needed to map 111 to y and reads

$$\mathcal{P}^e_{a_1}(y) \overset{def}{=} p^{\sharp_0 y}(1-p)^{\sharp_1 y}.$$

To minimize error probability for particular y, we have to decide to decode it either as a_1 with the aforementioned error probability

$$\mathcal{P}^e_{a_1}(y) = p^{\sharp_0 y}(1-p)^{\sharp_1 y} = \mathcal{P}(y|a_2).$$

or as a_2 with the error probability

$$\mathcal{P}^e_{a_2}(y) = p^{\sharp_1 y}(1-p)^{\sharp_0 y} = \mathcal{P}(y|a_1).$$

We observe that the (arithmetic) average error probability is the sum

$$P_e^{(3)} = \frac{1}{2}\sum_{y\in\{0,1\}^3} p_y, \tag{7.1}$$

where p_y is, depending on the decoding rule, either $\mathcal{P}(y|a_1)$ or $\mathcal{P}(y|a_2)$. To minimize the expression (7.1), we have to simply minimize all summands, i.e. for each y set $p_y = \min\{\mathcal{P}(y|a_1), \mathcal{P}(y|a_2)\}$.
Recall that $p < 1/2$, what gives that the optimum has to maximize the power of p. Hence, our decoding is minimal, observing that $p^2(1-p) < p(1-p)^2$ and $p^3 < (1-p)^3$.
We have that for each fixed decoding $\lambda_{max} \geq P_e^{(3)}$. Our decoding minimizes $P_e^{(3)}$ to be $\tilde{P}_e^{(3)}$. So, the maximal error probability for our decoding is

$$\tilde{\lambda}_{max} = \tilde{P}_e^{(3)} < P_e^{(3)} \leq \lambda_{max}. \tag{7.2}$$

Exercise 2

Consider a binary symmetric channel with $Y_i = Z_i \oplus X_i$, where $X_i, Y_i, Z_i \in \{0,1\}$. Suppose that all Z_i have the same marginal probability distribution $P(Z_i = 1) = p = 1 - P(Z_i = 0)$, but that $Z_1, Z_2, \ldots, Z_n$ are not necessarily independent. Assume that Z^n is independent of X^n (random vectors are independent!). Let $C = 1 - H(p, 1-p) = 1 - H(Z_1)$. Show that

$$\max_{X_1, X_2, \ldots, X_n} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \geq nC.$$

Answer of exercise 2

We evaluate

$$\begin{aligned} &I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \\ =&H(Y_1, Y_2, \ldots, Y_n) - H(Y_1, Y_2, \ldots, Y_n | X_1, X_2, \ldots, X_n) \\ =&H(Y_1, Y_2, \ldots, Y_n) - H(Z_1, Z_2, \ldots, Z_n). \end{aligned} \tag{7.3}$$

The last expression is maximized when the joint distribution of $Y_1, Y_2, \ldots, Y_n$ is uniform, what is achieved regardless of distribution on $Z_1, Z_2, \ldots, Z_n$ via the uniform distribution on $X_1, X_2, \ldots, X_n$, since X^n and Z^n are independent. Using the maximum value $H(Y_1, Y_2, \ldots, Y_n) = n$ we get the required result

$$\begin{aligned} \max_{X_1, X_2, \ldots, X_n} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) =&n - H(Z_1, Z_2, \ldots, Z_n) \\ \geq&n - nH(Z_1) = nC. \end{aligned} \tag{7.4}$$

Exercise 3

Consider the discrete memoryless channel $Y = X + Z \mod 11$, where Z is given by

i	1	2	3
$\mathcal{P}(Z = i)$	1/3	1/3	1/3

and $X \in \{0, 1, \ldots, 10\}$. Calculate the capacity of this channel assuming that Z is independent of X.

Answer of exercise 3

We calculate

$$\max_X I(X;Y) = \max_X (H(Y) - H(Y|X)) \overset{(*)}{=} \max_X H(Y) - \log 3, \tag{7.5}$$

where $(*)$ holds because the entropy $H(Y|X = x)$ does not depend on x, giving it does not depend on the distribution of X, and equals to $\log 3$. It remains to maximize $H(Y)$, but it is easy to see that we can achieve the uniform distribution on Y.

The maximum value

$$\max_X I(X;Y) = \log 10 - \log 3 \tag{7.6}$$

is achieved for the uniform distribution of X.

Exercise 4

Consider two discrete memoryless channels (with independent noise) $(X_1, p(y_1|x_1), Y_1)$ and $(X_2, p(y_2|x_2), Y_2)$ with the respective capacities C_1 and C_2. We form a new channel $(X_1 \times X_2, p((y_1, y_2)|(x_1, x_2)), Y_1 \times Y_2)$, where we send (x_1, x_2) simultaneously through the previously defined channels to obtain (y_1, y_2). The transition probability is given by

$$p((y_1, y_2)|(x_1, x_2)) = p(y_1|x_1)p(y_2|x_2). \tag{7.7}$$

Find the capacity of the combined channel.

Answer of exercise 4

Note that the maximum here should be taken over all joint (!) distributions of X_1 and X_2. We start with

$$\max_{X_1,X_2} I(X_1, X_2; Y_1, Y_2) = \max_{X_1,X_2} (H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2)).$$

Let us evaluate

$$\begin{aligned}
H(Y_1, Y_2|X_1, X_2) &= -\sum_{x_1,x_2,y_1,y_2} p(x_1, x_2, y_1, y_2) \log p(x_1, x_2|y_1, y_2) \\
&\overset{(*)}{=} -\sum_{x_1,x_2,y_1,y_2} p(x_1, x_2, y_1, y_2) \log(p(y_1|x_1)p(y_2|x_2)) \\
&= -\sum_{x_1,x_2,y_1,y_2} p(x_1, x_2, y_1, y_2) \log p(y_1|x_1) - \sum_{x_1,x_2,y_1,y_2} p(x_1, x_2, y_1, y_2)p(y_2|x_2)) \\
&= -\sum_{x_1,y_1} p(x_1, y_1) \log p(y_1|x_1) - \sum_{x_2,y_2} p(x_2, y_2)p(y_2|x_2)) \\
&= H(Y_1|X_1) + H(Y_2|X_2)
\end{aligned}$$

using Equation (7.7) to get $(*)$. We proceed with the capacity

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

On the other hand, we may achieve the capacity C_1+C_2, since if we maximize only over marginal distributions of X_1 and X_2, we get

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

This gives the result

$$\max_{X_1,X_2} I(X_1,X_2;Y_1,Y_2)=C_1+C_2. \tag{7.8}$$

Exercise 5

Consider a 26 key typewriter.

- If pushing a key results in printing the associated letter (no errors), what is the capacity C in bits?
- Suppose that pushing a key results in printing that letter of the next letter (with equal probabilities). Thus, $A \to A$ or B, $B \to B$ or C, …, $Z \to Z$ or A. What is the capacity?
- What is the highest rate code with block length one that you can find that achieves zero probability of error for the channel in part 2?

Answer of exercise 5

??

Exercise 6

Let the channel has binary input and output alphabet and transition probabilities $p(y|x)$ given by

$X\backslash Y$	0	1
0	1	0
1	1/2	1/2

Find capacity of this channel and the input probability distribution achieving this capacity.

Answer of exercise 6

??