

HW 5 — Due: Not Due

Lecturer: Asst. Prof. Dr. Prapun Suksompong

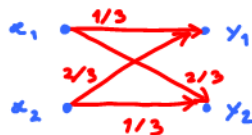
Problem 1. For each of the DMCs whose corresponding transition probability matrices Q are specified below, (i) draw the channel diagram and (ii) compute its capacity C and the corresponding \mathbf{p} that achieves it.

(a)

(i) The problem does not explicitly specify \mathcal{X} and \mathcal{Y} . However, from the size of the Q matrix, we know that $|\mathcal{X}| = |\mathcal{Y}| = 2$. So, we will set $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$.

$$Q = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}$$

(ii) Note that this is a BSC with $p = \frac{2}{3}$. We know that the capacity of BSC is $C = 1 - H(p) = 1 - H(\frac{2}{3}) = 1 - 0.9183 \approx 0.0817$ bpcu.



Alternatively, this channel is weakly symmetric. Therefore, $C = \log_2 |S_Y| - H(\kappa) = \log_2 2 - H([\frac{1}{3} \frac{2}{3}]) = 1 - 0.9183 \approx 0.0817$

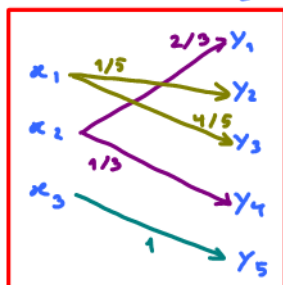
(b)

(i) Let $\mathcal{X} = \{x_1, x_2, x_3\}$ and $\mathcal{Y} = \{y_1, y_2, y_3, y_4, y_5\}$.

$$Q = \begin{bmatrix} 0 & 1/5 & 4/5 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(ii) Note that this is a noisy channel with non-overlapping outputs.

(Only one non-zero element in each column of Q .)

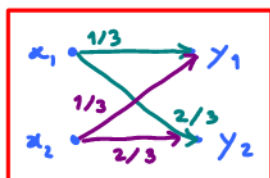


So, $C = \log_2 |S_X| = \log_2 3 \approx 1.5850$ bpcu.

(c)

(i) As in (a), we let $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$

$$Q = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix}$$



(ii) The rows of Q are the same.

So, $Q(y|x)$ does not depend on x .

can "pull" $Q(y|x)$ out of the summation

$$\text{Therefore, } q(y) = \sum_x p(x) Q(y|x)$$

$$= Q(y|x) \sum_x p(x) = Q(y|x)$$

which implies $X \perp\!\!\!\perp Y$.

Therefore, $I(X;Y) = 0$ for any input distribution. Hence, $C = 0$ bpcu.

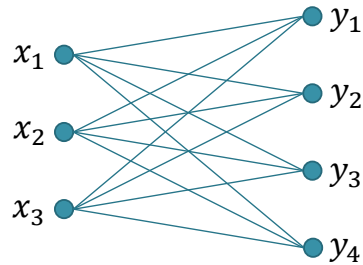


Figure 5.1: Channel diagram for Problem 2.

Problem 2. The channel diagram for a DMC is shown in Figure 4.1.

- (a) What are the dimensions of its matrix \mathbf{Q} ?

There are $3 = |\mathcal{X}|$ possible channel inputs and $4 = |\mathcal{Y}|$ possible channel outputs.
Therefore, the \mathbf{Q} matrix must be 3×4 .

- (b) Is it possible to find appropriate values for the \mathbf{Q} matrix of this DMC such that $C = 1.6$ bpcu? If so, give an example of such \mathbf{Q} .

We know that $C \leq \min \left\{ \log_2 |\mathcal{X}|, \log_2 |\mathcal{Y}| \right\} = \log_2 3 \approx 1.5850$

However $1.6 > 1.5850$. Therefore $C = 1.6$ bpcu is impossible.

- (c) Is it possible to find appropriate values for the \mathbf{Q} matrix of this DMC such that $C = \log_2 3$ bpcu? If so, give an example of such \mathbf{Q} .

Note that $\log_2 3 = \log_2 |\mathcal{X}|$. Therefore, let's try to construct a noiseless channel or a noisy channel with non-overlapping output

Examples of answers:



The corresponding \mathbf{Q} are

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ or } \mathbf{Q} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

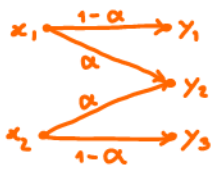
, respectively.

Problem 3 (Calculus*). On many occasions, we have to work with a DMC whose \mathbf{Q} matrix *does not match* any of our special cases discussed in class. As discussed in class, capacity calculation for these more general cases can be done via treating them as optimization problems in calculus. When $|\mathcal{X}| = 2$, $\underline{\mathbf{p}} = [p_0 \ 1 - p_0]$. So, the optimization is over one variable. In general, when $|\mathcal{X}| = n$, the optimization is over $n - 1$ variables.

For each of the DMCs whose transition probability matrices \mathbf{Q} are specified below, (i) draw the channel diagram and (ii) compute its capacity C and the corresponding $\underline{\mathbf{p}}$ that achieves it by first calculating $I(X; Y)$ for an arbitrary $\underline{\mathbf{p}}$ (with appropriate dimension) and then set appropriate (partial) derivative(s) to 0 to solve for $\underline{\mathbf{p}}$.

(i)

The given DMC belongs to a family of channels called **Binary Erasure Channel (BEC)**. The general form of the channel diagram and the corresponding \mathbf{Q} matrix are



$$\mathbf{Q} = \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & \alpha & 1-\alpha \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/3 \end{bmatrix}$$

Usually, $x_1 = y_1 = 0, x_2 = y_3 = 1, y_2 = e$

the case where the bit is "erased"



Here, we have $\alpha = 2/3$. So, the channel diagram is

(ii)

For general α ,

$$\mathbf{Q} = \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & \alpha & 1-\alpha \end{bmatrix} \rightarrow \begin{cases} H(Y|X=0) = H(\alpha) \\ H(Y|X=1) = H(\alpha) \end{cases} \Rightarrow H(Y|X) = H(\alpha)$$

Let $\underline{\mathbf{p}} = [p_0 \ 1-p_0]$. Then,

$$\mathbf{q}_Y = \underline{\mathbf{p}} \mathbf{Q} = [p_0 \ 1-p_0] \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & \alpha & 1-\alpha \end{bmatrix} = \begin{bmatrix} p_0(1-\alpha) & \alpha & (1-p_0)(1-\alpha) \end{bmatrix}$$

Write these as $\bar{\alpha}$

$$= [p_0 \bar{\alpha} \ \alpha \ \bar{p}_0 \bar{\alpha}]$$

Write this as \bar{p}_0

Therefore, $H(Y) = -p_0 \bar{\alpha} \log_2 p_0 \bar{\alpha} - \alpha \log_2 \alpha - \bar{p}_0 \bar{\alpha} \log_2 \bar{p}_0 \bar{\alpha}$

and

$$I(X; Y) = H(Y) - H(Y|X) = H(Y) - H(\alpha)$$

Some useful facts: For any positive function $f(x)$,

$$\frac{d}{dx} f(x) \ln f(x) = f'(x) \ln f(x) + f(x) \frac{1}{f(x)}$$

$$= f'(x) (1 + \ln f(x)) = f'(x) \ln (e f(x))$$

$$\downarrow \times \frac{1}{\ln 2}$$

$$\frac{d}{dx} f(x) \log_2 f(x) = f'(x) \log_2 e f(x)$$

$$\frac{d}{dp_0} I(X; Y) = -\bar{\alpha} \log_2 (p_0 \bar{\alpha}) - (-\bar{\alpha}) \log_2 (p_0 \bar{\alpha}) = 0$$

$p_0 = \bar{p}_0$
 $p_0 = 1 - p_0$
 $p_0 = \frac{1}{2}$

One may say that the uniform pmf is "obvious" from the "symmetry" in the transition probabilities from the inputs.

with $p_0 = \frac{1}{2}$, we have

$$C = I(X; Y) = -\frac{\bar{\alpha}}{2} \log_2 \frac{\bar{\alpha}}{2} - \alpha \log_2 \alpha - \frac{\bar{\alpha}}{2} \log_2 \frac{\bar{\alpha}}{2} - (-\alpha \log_2 \alpha - \bar{\alpha} \log_2 \bar{\alpha})$$

$$= -\bar{\alpha} \log_2 \frac{\bar{\alpha}}{2} + \bar{\alpha} \log_2 \bar{\alpha} = \bar{\alpha} \log_2 \frac{\bar{\alpha}}{\bar{\alpha}/2} = \bar{\alpha} = 1 - \alpha$$

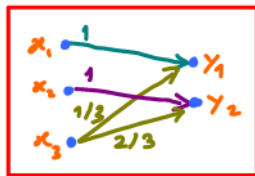
When $\alpha = \frac{2}{3}$, $C = 1 - \frac{2}{3} = \frac{1}{3}$ bpcu.

The capacity-achieving $\bar{\mathbf{p}}$ is $\bar{\mathbf{p}} = [\frac{1}{2} \ \frac{1}{2}]$.

Problem 4. The selected examples of the \mathbf{Q} matrices presented in class are not the only examples in which capacity values can be found without solving the full-blown optimization problems (from which you have suffered in Problem 3). Here is one more example. Consider a DMC whose transition probability matrix \mathbf{Q} is

$$\mathbf{Q} = \begin{array}{c} \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \backslash \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/3 & 2/3 \end{bmatrix} \end{array}.$$

(a) Draw the channel diagram.



(b) Compute its capacity C and the corresponding $\underline{\mathbf{p}}$ that achieves it.

Hint: Bound the capacity from the dimensions of \mathbf{Q} then try to find a $\underline{\mathbf{p}}$ that achieves the bound. Note that without the last row in the matrix \mathbf{Q} , the channel is noiseless. One may also eliminate a “useless” channel input by not using it at all.

Note that $|Y| = 2$. So, $I(X; Y) \leq H(Y) \leq \log_2 |Y| = \log_2 2 = 1$.

Note also that without the last row, this \mathbf{Q} corresponds to a noiseless channel with capacity $\log_2 |X| = \log_2 2 = 1$ which is achieved by uniform pmf on the input. ($\underline{\mathbf{p}} = [\frac{1}{2} \frac{1}{2}]$.)

Now, when the last row of \mathbf{Q} is included, we may choose not to use it by using $\underline{\mathbf{p}} = [\frac{1}{2} \frac{1}{2} 0]$. This gives

$$\underline{\mathbf{p}} = [\frac{1}{2} \frac{1}{2} 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/3 & 2/3 \end{bmatrix} = [\frac{1}{2} \frac{1}{2}]$$

so, $H(Y) = H([\frac{1}{2} \frac{1}{2}]) = 1$ and

$$H(Y|X) = \frac{1}{2} \times H([1 \ 0]) + \frac{1}{2} \times H([0 \ 1]) + 0 \times H([\frac{1}{3} \ \frac{2}{3}]) = 0.$$

Therefore, $I(X; Y) = H(Y) - H(Y|X) = 1 - 0 = 1$ which is the same as the bound above.

Because $I(X; Y)$ can not exceed the bound, we know that this $\underline{\mathbf{p}} = [\frac{1}{2} \frac{1}{2} 0]$ achieved the maximum $I(X; Y)$.

Therefore, $C = 1$ bpcu and corresponding $\underline{\mathbf{p}} = [\frac{1}{2} \frac{1}{2} 0]$

Problem 5 (Blahut-Arimoto algorithm). A MATLAB function `capacity_blahut` is provided on the course web site. It calculates the capacity $C = \max_{\underline{p}} I(\underline{p}, \mathbf{Q})$ and the corresponding capacity-achieving input pmf \underline{p} using Blahut-Arimoto algorithm.

Write a MATLAB script to check your answers in all earlier problems with the help of the provided MATLAB function.

`close all; clear all;`

```

%% 1.a
Q = [1/3 2/3; 2/3 1/3];
[ps C] = capacity_blahut(Q)

```

→

```

>> Capacity_HW_blahut
ps =
 [0.5000 0.5000] = p
C =
 0.0817 bpcu

```

```

%% 1.b
Q = [0 1/5 4/5 0 0; 2/3 0 0 1/3 0; 0 0 0 0 1];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.3333 0.3333 0.3333] = p
C =
 1.5850 bpcu

```

```

%% 1.c
Q = [1/3 2/3; 1/3 2/3];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.5000 0.5000] = p
C =
 0 bpcu

```

```

%% 2.c
Q = [1 0 0 0; 0 1 0 0; 0 0 1 0];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.3333 0.3333 0.3333] = p
C =
 1.5850 bpcu

```

```

%% 3.a
Q = [1/3 2/3 0; 0 2/3 1/3];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.5000 0.5000] = p
C =
 0.3333

```

one of the answers

```

%% 3.b
Q = [1 0; 1/2 1/2; 1/2 1/2];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.6000 0.2000 0.2000] = p
C =
 0.3219 bpcu

```

```

%% 4
Q = [1 0; 0 1; 1/3 2/3];
[ps C] = capacity_blahut(Q)

```

→

```

ps =
 [0.5000 0.5000 0.0000] = p
C =
 1.0000 bpcu

```