Name ________________________ ID3 _ _ _

ECS 452: Digital Communication Systems **2018/2**

HW 5 — Due: Not Due

Lecturer: Prapun Suksompong, Ph.D.

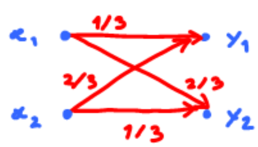
Problem 1. For each of the DMCs whose corresponding transition probability matrices $\mathbf{Q}$ are specified below, (i) draw the channel diagram and (ii) compute its capacity C and the corresponding $\underline{\mathbf{p}}$ that achieves it.

(a)

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

(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 $\mathcal{Y} = \{y_1, y_2\}$.



(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,

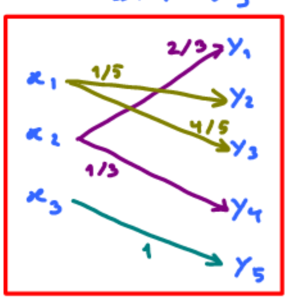
$C = \log_2 |S_Y| - H(\underline{r}) = \log_2 2 - H([\frac{1}{3} \ \frac{2}{3}])$

$= 1 - 0.9183 \approx 0.0817$

(b)

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

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



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

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

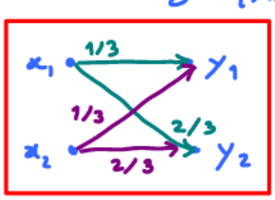
So, $C = \log_2 |S_X|$

$= \log_2 3 \approx 1.5850$ bpcu.

(c)

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

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



(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

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.

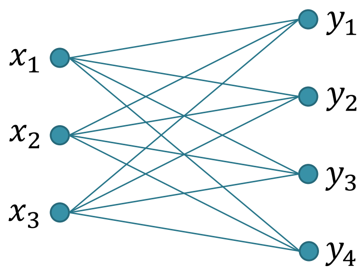


Figure 5.1: Channel diagram for Problem 2.

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

(a) What are the dimensions of its matrix **Q**?

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

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

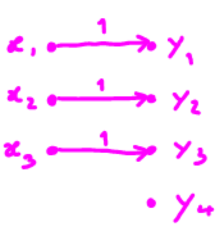
We know that $C \le \min\{\underbrace{\log_2|\mathcal{X}|}_{=\log_2 3}, \underbrace{\log_2|\mathcal{Y}|}_{=\log_2 4}\} = \log_2 3 \approx 1.5850$

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

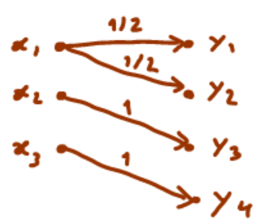
(c) Is it possible to find appropriate values for the **Q** matrix of this DMC such that $C = \log_2 3$ bpcu? If so, give an example of such **Q**.

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:



or



The correspond Q are

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

, respectively.

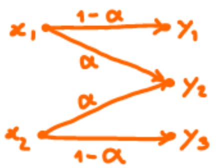
Problem 3 (Calculus*)**.** On many occasions, we have to work with a DMC whose $\mathbf{Q}$ matrix ***does not match*** any of our special cases discussed in class. As discussed in class, capacity calculation for these more general cases can be done via treating them as optimization problems in calculus. 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.

$p_0 \leftarrow$

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 Q matrix are



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

(a)

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

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,

$$Q = \begin{array}{c|ccc} x \backslash y & 0 & e & 1 \\ \hline 0 & 1-\alpha & \alpha & 0 \\ 1 & 0 & \alpha & 1-\alpha \end{array} \quad \begin{array}{l} \to H(Y|X=0) = H(\alpha) \\ \to H(Y|X=1) = H(\alpha) \end{array} \Bigg\} \Rightarrow H(Y|X) = H(\alpha)$$

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

Write these as $\bar{\alpha}$

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

Write this as $\bar{p}_0$

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

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

and

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

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

$$\frac{d}{dx} f(x)\ln f(x) = f'(x)\ln f(x) + f(x)\frac{1}{f(x)}f'(x)$$
$$= f'(x)(1+\ln f(x)) = f'(x)\ln(e f(x))$$
$$\Big\downarrow \times \frac{1}{\ln 2}$$
$$\frac{d}{dx} f(x)\log_2 f(x) = f'(x)\log_2 e f(x)$$

$$\frac{d}{dp_0} I(X;Y) = -\cancel{\bar{\alpha}}\log_2(\cancel{e}p_0\cancel{\bar{\alpha}}) - (-\cancel{\bar{\alpha}})\log_2(\cancel{e}\bar{p}_0\cancel{\bar{\alpha}}) = 0$$

$p_0 = \bar{p}_0$

$p_0 = 1 - p_0$

$p_0 = \frac{1}{2}$ ← 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

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

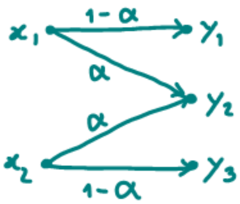
$$= -\bar{\alpha}\log_2\frac{\bar{\alpha}}{2} + \bar{\alpha}\log_2\bar{\alpha} = \bar{\alpha}\log_2\frac{\bar{\alpha}}{\bar{\alpha}/2} = \bar{\alpha} = 1-\alpha$$

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

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

Alternatively, we can solve this part by trying to apply properties of entropy (the same way that we solve the examples in class) and avoid using calculus.

Let $P[X=x_1]=p_0$



$$Q = \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & \alpha & 1-\alpha \end{bmatrix} \begin{matrix} \xrightarrow{\times p_0} \\ \xrightarrow[\times(1-p_0)]{} \end{matrix} \begin{bmatrix} (1-\alpha)p_0 & \alpha p_0 & 0 \\ 0 & \alpha(1-p_0) & (1-\alpha)(1-p_0) \end{bmatrix}$$

$\Sigma \downarrow$ α

Observe that $H(X|Y=y_1)=0$ (When $Y=y_1$, we know that X must be x_1.)

Similarly, $H(X|Y=y_3)=0$ (" $Y=y_3$ " " x_2.)

Now for $H(X|Y=y_2)$,

$$P[X=x_1|Y=y_2] = \frac{P[Y=y_2|X=x_1]P[X=x_1]}{P[Y=y_2]} = \frac{\alpha p_0}{\alpha} = p_0$$

Bayes' theorem

This gives $P[X=x_2|Y=y_2] = 1 - p_0 = 1-p_0$.

Hence, $H(X|Y=y_2) = H([p_0 \;\; 1-p_0]) = H(X)$

Therefore, $H(X|Y) \equiv \sum_y q(y) H(X|Y=y) = 0 + \alpha H(X)$

Now, $I(X;Y) = H(X) - H(X|Y) = H(X) - \alpha H(X) = (1-\alpha)H(X)$.

So, maximizing $I(X;Y)$ is the same as maximizing $H(X)$.

Because X is binary, we already know (from Def. 2.47 and Figure3 in Ch. 2) that

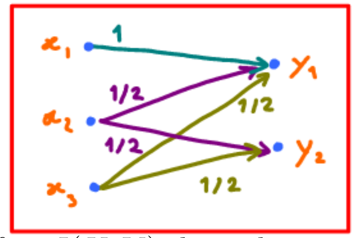
$H(X)$ is maximized when X is uniform ($\underline{p} = [1/2 \;\; 1/2]$) with max value of 1.

Therefore, the capacity of this channel is $C = (1-\alpha)\times 1 = 1-\alpha$ [bpcu] and the capacity-achieving input probability vector is $\underline{p} = [1/2 \;\; 1/2]$.

(b)

$$\begin{array}{c|cc} x \backslash y & y_1 & y_2 \\ \hline x_1 & & \\ x_2 & & \\ x_3 & & \end{array} \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \Rightarrow$$

(i)



Hint: Let $\underline{\mathbf{p}} = [p_1, p_2, p_3]$. Note that the values of p_2 and p_3 only affect $I(X,Y)$ through $p_2 + p_3$, which is $1 - p_1$. So, we are back to optimizing over only one variable.

(ii) As hinted, we let $\underline{p} = [p_1\ p_2\ p_3]$.

Then,

$$\underline{q} = \underline{p}Q = [p_1\ p_2\ p_3]\begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \left[p_1 + \tfrac{1}{2}p_2 + \tfrac{1}{2}p_3 \quad 0 + \tfrac{1}{2}p_2 + \tfrac{1}{2}p_3\right] = \left[p_1 + \frac{p_2+p_3}{2} \quad \frac{p_2+p_3}{2}\right]$$

Next, from each row of the Q matrix, we can find

$$\left.\begin{array}{l} H(Y|X=x_1) = H([1\ 0]) = 0, \\ H(Y|X=x_2) = H([\tfrac{1}{2}\ \tfrac{1}{2}]) = \log_2 2 = 1, \text{ and} \\ H(Y|X=x_3) = \quad '' \quad\quad '' \quad = 1. \end{array}\right\} \Rightarrow \text{Therefore, } H(Y|X) = \sum_x p(x)H(Y|x) = p_1 \times 0 + p_2 + p_3 = p_2 + p_3.$$

Note that $\underline{q}$ and $H(Y|X)$ do not depend directly on the values of p_2 and p_3. They use p_2 and p_3 via the sum p_2+p_3.

Hence, $H(Y)$ (which is calculated from $\underline{q}$) and $I(X;Y) = H(Y) - H(Y|X)$ also do not depend directly on p_2 and p_3. Only the sum p_2+p_3 is involved.

Note that $p_1+p_2+p_3 = 1$. So, $p_2+p_3 = 1-p_1$.

Hence, $\underline{q} = \left[p_1 + \frac{p_2+p_3}{2} \quad \frac{p_2+p_3}{2}\right] = \left[p_1 + \frac{1-p_1}{2} \quad \frac{1-p_1}{2}\right] = \left[\frac{1+p_1}{2} \quad \frac{1-p_1}{2}\right]$

$H(Y|X) = p_2+p_3 = 1-p_1$

So, $I(X;Y) = H(Y) - H(Y|X) = -\frac{1+p_1}{2}\log_2\frac{1+p_1}{2} - \frac{1-p_1}{2}\log_2\frac{1-p_1}{2} - (1-p_1)$

$$\frac{d}{dp_1}I(X;Y) = -\frac{1}{2}\log_2 e\,\frac{1+p_1}{2} - \left(-\frac{1}{2}\right)\log_2 e\,\frac{1-p_1}{2} + 1 = 0$$

$$\frac{1}{2}\log_2\frac{1-p_1}{1+p_1} + 1 = 0 \Rightarrow \frac{1+p_1}{1-p_1} = 2^2 = 4 \Rightarrow 1+p_1 = 4-4p_1 \Rightarrow 5p_1 = 3 \Rightarrow p_1 = \frac{3}{5}$$

With $p_1 = \frac{3}{5}$,

$$\underline{q} = \left[\frac{1+\frac{3}{5}}{2} \quad \frac{1-\frac{3}{5}}{2}\right] = \left[\tfrac{4}{5}\ \tfrac{1}{5}\right]$$

So, $C = H([\tfrac{4}{5}\ \tfrac{1}{5}]) - (1-\tfrac{3}{5}) \approx 0.7219 - \tfrac{2}{5} = 0.3219$ bpcu

The pmf $\underline{p}$ of X that achieves this capacity value is

$\underline{p} = [\tfrac{3}{5}\quad p_2\quad p_3]$ where $p_2+p_3 = \frac{2}{5}$

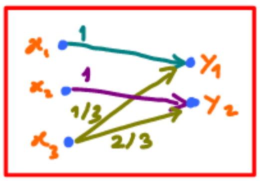
$\uparrow$

Ex. $p_2 = p_3 = \frac{1}{5}$

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

$$\mathbf{Q} = \begin{array}{c} x \backslash y \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{cc} y_1 & y_2 \end{array} \\ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 1/3 & 2/3 \end{array}\right] \end{array}.$$

(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 also that without the last row in the matrix $\mathbf{Q}$, the channel is noiseless. One may eliminate a "useless" channel input by not using it at all.

Note that $|\mathcal{Y}| = 2$. So, $I(X;Y) \le H(Y) \le \log_2 |\mathcal{Y}| = \log_2 2 = 1$.

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

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

$$\underline{q} = [1/2\ \ 1/2\ \ 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/3 & 2/3 \end{bmatrix} = [\tfrac{1}{2}\ \ \tfrac{1}{2}]$$

So, $H(Y) = H([1/2\ \ 1/2]) = 1$ and

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

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 $\underline{p} = [1/2\ \ 1/2\ \ 0]$ achieved the maximum $I(X;Y)$.

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

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

%% 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)

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

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

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

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

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

```
>> Capacity_HW_blahut
ps =
 [0.5000   0.5000] = p
C =
   0.0817  bpcu
ps =
 [0.3333   0.3333   0.3333] = p
C =
   1.5850  bpcu
ps =
 [0.5000   0.5000] = p
C =
   0  bpcu
ps =
 [0.3333   0.3333   0.3333] = p
C =
   1.5850  bpcu
ps =
 [0.5000   0.5000] = p
C =
   0.3333
                    one of the answers
ps =
 [0.6000   0.2000   0.2000] = p
C =
   0.3219  bpcu
ps =
 [0.5000   0.5000   0.0000] = p
C =
   1.0000  bpcu
```