COMP2610/COMP6261 – Information Theory

Tutorial 10

Manish Kumar (manish.kumar@anu.edu.au)

Question 1.

Channel capacity. Consider the discrete memoryless channel $Y = X + Z \pmod{11}$, where

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

mod 11 operation

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

for $X = 10$, $z = 1$, $z = 2$, $z = 3$

$y = 10 + 1 \pmod{11} = \frac{11}{11} = 0$

$y = 10 + 2 \pmod{11} = \frac{12}{11} = 1$

$y = \frac{10+3}{11} = 2$

So, $y \in \{0, 1, \dots, 10\}$

(a) Find the capacity.

(b) What is the maximizing $p^*(x)$?

Solution: *Channel capacity.*

$$Y = X + Z (\text{mod } 11)$$

where

$$Z = \begin{cases} 1 & \text{with probability} 1/3 \\ 2 & \text{with probability} 1/3 \\ 3 & \text{with probability} 1/3 \end{cases}$$

In this case,

$$H(Y|X) = H(Z|X) = H(Z) = \log 3,$$

$y = x + z$

independent of the distribution of X, and hence the capacity of the channel is

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} H(Y) - H(Y|X) \\ &= \max_{p(x)} H(Y) - \log 3 \\ &= \log 11 - \log 3, \end{aligned}$$

$H(Y/X) = H(Z/X)$, because, for a given X, measuring the degree of Randomness (Entropy) in Y is similar to entropy of $H(Z/X)$, for $y = x + z$ (x known)

★ In order to maximise entropy, always choose uniform distribution.

which is attained when Y has a uniform distribution, which occurs (by symmetry) when X has a uniform distribution.

(a) The capacity of the channel is $\log \frac{11}{3}$ bits/transmission.

(b) The capacity is achieved by an uniform distribution on the inputs. $p(X = i) = \frac{1}{11}$ for $i = 0, 1, \dots, 10$.

$$\max_{p(x)} H(y) = 11 \times \left(\frac{1}{11}\right) \log_2 11 = \log_2 11$$

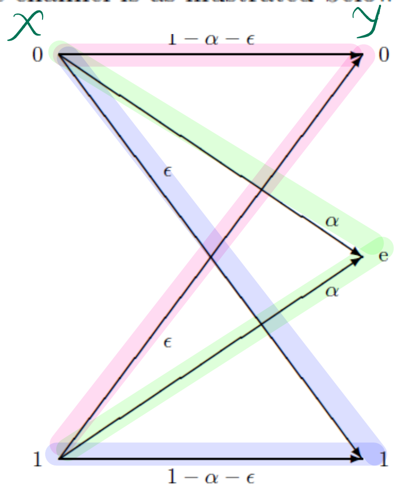
Provided $P(y=0) = P(y=1) = \dots = P(y=10) = \frac{1}{11}$

<u>Question 2.</u>

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 the channel is as illustrated below:

let $P(x=0) = \pi$

$\Rightarrow P(x=1) = 1-\pi$



$P(Y=0) = \pi(1-\alpha-\epsilon) + (1-\pi)\epsilon$

$P(Y=e) = \pi\alpha + (1-\pi)\alpha = \alpha$

$P(Y=1) = (1-\pi)(1-\alpha-\epsilon) + \pi\epsilon$

(a) Find the capacity of this channel.

(b) Specialize to the case of the binary symmetric channel ($\alpha = 0$).

(c) Specialize to the case of the binary erasure channel ($\epsilon = 0$).

Solution:

(a) As with the examples in the text, we set the input distribution for the two inputs to be π and $1-\pi$. Then

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} (H(Y) - H(Y|X)) \\ &= \max_{p(x)} H(Y) - H(1-\epsilon-\alpha, \alpha, \epsilon). \quad \text{-----} \; i \end{aligned}$$

As in the case of the erasure channel, the maximum value for $H(Y)$ cannot be $\log 3$, since the probability of the erasure symbol is α independent of the input distribution. Thus,

$$\begin{aligned} H(Y) &= H(\underbrace{\pi(1-\epsilon-\alpha)+(1-\pi)\epsilon}_{P_0}, \alpha, \underbrace{(1-\pi)(1-\epsilon-\alpha)+\pi\epsilon}_{P_0}) \quad \text{----} \; ii \\ &\leq H(\alpha) + (1-\alpha) \end{aligned}$$

with equality iff $\frac{\pi+\epsilon-\pi\alpha-2\pi\epsilon}{1-\alpha} = \frac{1}{2}$, which can be achieved by setting $\pi = \frac{1}{2}$. (The fact that $\pi = 1-\pi = \frac{1}{2}$ is the optimal distribution should be obvious from the symmetry of the problem, even though the channel is not weakly symmetric.)

$H(Y) = H(\alpha) + H(\pi(1-\epsilon-\alpha) + (1-\pi), (1-\pi)(1-\epsilon-\alpha) + \pi\epsilon)$

$$= H(\alpha) + (1-\alpha) H\left(\underbrace{\frac{\pi(1-\epsilon-\alpha)+(1-\pi)}{1-\alpha}, \frac{(1-\pi)(1-\epsilon-\alpha)+\pi\epsilon}{1-\alpha}}\right)$$

Should be 1

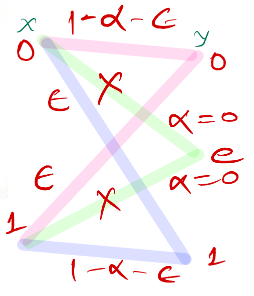
for $\max_{P(x)} H(Y)$

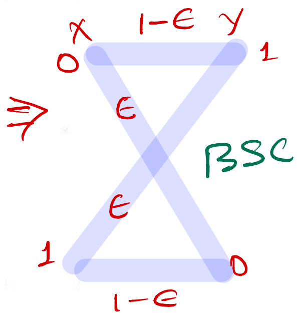
$\Rightarrow \max_{P(x)} H(Y) \leq H(\alpha) + (1-\alpha)$ ---- iii

Therefore the capacity of this channel is

from eqn. (iii)

$$
\begin{aligned}
C &= H(\alpha)+1-\alpha-H(1-\alpha-\epsilon,\alpha,\epsilon) \\
&= H(\alpha)+1-\alpha-H(\alpha)-(1-\alpha)H\left(\frac{1-\alpha-\epsilon}{1-\alpha},\frac{\epsilon}{1-\alpha}\right) \\
&= (1-\alpha)\left(1-H\left(\frac{1-\alpha-\epsilon}{1-\alpha},\frac{\epsilon}{1-\alpha}\right)\right)
\end{aligned}
$$





(b) Setting $\alpha = 0$, we get

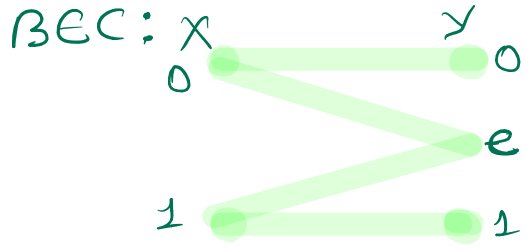
$$C = 1 - H(\epsilon),$$

which is the capacity of the binary symmetric channel.

(c) Setting $\epsilon = 0$, we get

$$C = 1 - \alpha$$

which is the capacity of the binary erasure channel.



Question 3.

Binary multiplier channel

(a) Consider the channel $Y = XZ$ where X and Z are independent binary random variables that take on values 0 and 1. Z is Bernoulli(α), i.e. $P(Z=1)=\alpha$. Find the capacity of this channel and the maximizing distribution on X.

(b) Now suppose the receiver can observe Z as well as Y. What is the capacity?

Receiver has information of Z and Y.

Prob.

$$X = \begin{cases} 0, & 1-p \\ 1, & p \end{cases}$$

$$Z = \begin{cases} 0, & 1-\alpha \\ 1, & \alpha \end{cases}$$

$$Y = \begin{cases} 0, & 1-p\alpha \\ 1, & p\alpha \end{cases}$$

Solution: *Binary Multiplier Channel*

$H(Y/X)$ if $x=0$ then y will always be 0, for any Z

$= H(Y/X=0) + H(Y/X=1)$

$= H(Y/X=1) = H(Z/X=1)$

$= P(x=1)\,H(Z) = pH(\alpha)$

(a) Let $P(X=1)=p$. Then $P(Y=1)=P(X=1)P(Z=1)=\alpha p$.

$$
\begin{aligned}
I(X;Y) &= H(Y)-H(Y|X) \\
&= H(Y)-P(X=1)H(Z) \\
&= H(\alpha p)-pH(\alpha)
\end{aligned}
$$

Now, calculate $\max_{p(x)} I(X;Y)$

We find that $p^* = \frac{1}{\alpha(2^{\frac{H(\alpha)}{\alpha}}+1)}$ maximizes $I(X;Y)$. The capacity is calculated to be $\log(2^{\frac{H(\alpha)}{\alpha}}+1)-\frac{H(\alpha)}{\alpha}$.

Can also be written as,

$$\boxed{I(X;Y) = H(\alpha p, 1-\alpha p) - pH(\alpha, 1-\alpha)}$$

(b) Let $P(X=1)=p$. Then

$$
\begin{aligned}
I(X;Y,Z) &= I(X;Z)+I(X;Y|Z) \\
&= H(Y|Z)-H(Y|X,Z) \\
&= H(Y|Z) \\
&= \alpha H(p)
\end{aligned}
$$

not ~~I(Y;X,Z)~~

because receiver can observe Z and Y (known, will act as i/p). So, joint mutual information should be calculated for $I(X;Y,Z)$ and not $I(Y;X,Z)$.

↳ if X and Z are known so, there will be no randomness in Y ⇒ $\boxed{H(Y/X,Z)=0}$

The expression is maximized for $p=1/2$, resulting in $C=\alpha$. Intuitively, we can only get X through when Z is 1, which happens α of the time.

Question 4.

Can signal alternatives lower capacity? Show that adding a row to a channel transition matrix does not decrease capacity.

Solution: *Can signal alternatives lower capacity?*
Adding a row to the channel transition matrix is equivalent to adding a symbol to the input alphabet $\mathcal{X}$. Suppose there were m symbols and we add an $(m+1)$-st. We can always choose not to use this extra symbol.

Specifically, let C_m and C_{m+1} denote the capacity of the original channel and the new channel, respectively. Then

$$\begin{aligned} C_{m+1} &= \max_{p(x_1,\ldots,x_{m+1})} I(X;Y) \\ &\geq \max_{p(x_1,\ldots,x_m,0)} I(X;Y) \\ &= C_m. \end{aligned}$$