Quantum Information Theory —Homework Lecture 4

Gavin King

October 31, 2023

Question 1

We begin by observing that:

$$\begin{aligned}\sum_{i=1}^{n}\left[I(X_{i+1}^n;Y_1^i)-I(X_i^n;Y_1^{i-1})\right] &= \sum_{i=1}^{n} I(X_{i+1}^n;Y_1^i)-\sum_{i=0}^{n-1} I(X_{i+1}^n;Y_1^i)\\ &= I(X_{n+1}^n;Y_1^i)-I(X_1^n;Y_1^0)\\ &= I(\emptyset;Y_1^i)-I(X_1^n;\emptyset)\\ &= 0\end{aligned}$$

Then, by application of the chain rule for mutual information:

$$\begin{aligned}\sum_{i=1}^{n} I(X_{i+1}^n;Y_i|Y_1^{i-1}) &= \sum_{i=1}^{n}\left[I(X_{i+1}^n;Y_1^i)-I(X_{i+1}^n;Y_1^{i-1})\right]\\ &= \sum_{i=1}^{n}\left[I(X_{i+1}^n;Y_1^i)-I(X_i^n;Y_1^{i-1})+I(X_i;Y_1^{i-1}|X_{i+1}^n)\right]\\ &= \sum_{i=1}^{n} I(X_i;Y_1^{i-1}|X_{i+1}^n)\end{aligned}$$

Question 2

Part (a)

For any X, Y, Z:

$$H(X|Z)\le H(X|Z)+H(Y|X,Z)=H(X,Y|Z)=H(Y|Z)+H(X|Y,Z)\le H(Y|Z)+H(X|Y)$$

Part (b)

For independent X, Y:

$$H(X+Y)\ge H(X+Y|Y)=H(X)$$

To see that the inequality need not be saturated, consider a sum of Bernoulli variables.

Part (c)

We're given that:

$$p_{Y_1,Y_2|X_1,X_2}(y_1,y_2|x_1,x_2)=p_{Y_1|X_1}(y_1|x_1)\,p_{Y_2|X_2}(y_2|x_2)$$

We may marginalize over Y_1 to see that $X_1 \leftrightarrow X_2 \leftrightarrow Y_2$ or Y_2 over Y_2 to see that $Y_1 \leftrightarrow X_1 \leftrightarrow X_2$. Therefore:

$$\begin{aligned}
I(X_1, X_2; Y_1, Y_2) &= I(X_1, X_2; Y_1) + I(X_1, X_2; Y_2|Y_1) \\
&= I(X_1; Y_1) + I(X_2; Y_1|X_1) + I(X_2; Y_2|Y_1) + I(X_1; Y_2|Y_1, X_2) \\
&\leq I(X_1; Y_1) + I(X_2; Y_1|X_1) + I(X_2; Y_2) + I(X_1; Y_2|X_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2)
\end{aligned}$$

At the inequality, and also at the last equality we used that $Y_1 \leftrightarrow X_1 \leftrightarrow X_2 \leftrightarrow Y_2$ form a Markov chain.

Part (d)

We're given that $p_{X_1,X_2}(x_1, x_2) = p_{X_1}(x_1)\, p_{X_2}(x_2)$, that is, that X_1 and X_2 are independent.

$$\begin{aligned}
I(X_1, X_2; Y_1, Y_2) &= I(X_1; Y_1, Y_2) + I(X_2; Y_1, Y_2|X_1) \\
&= I(X_1; Y_1) + I(X_1; Y_2|Y_1) + I(X_2; Y_2|X_1) + I(X_2; Y_1|X_1, Y_2) \\
&\geq I(X_1; Y_1) + I(X_1; Y_2|Y_1) + I(X_2; Y_2) + I(X_2; Y_1|Y_2) \\
&\geq I(X_1; Y_1) + I(X_2; Y_2)
\end{aligned}$$

At the first inequality we used the independence of X_1 and X_2.

Question 3

The Z channel is defined by the conditional PMF $p_{Y|X}$:

$$p_{Y|X}(0|0) = 1 \qquad p_{Y|X}(1|1) = \tfrac{1}{2} \qquad p_{Y|X}(0|1) = \tfrac{1}{2}$$

Thus, taking $Z \sim \text{Bern}(p)$, the joint distribution $p_{Y,X}$ is:

$$p_{Y,X}(0,0) = 1 - p \qquad p_{Y,X}(1,0) = 0 \qquad p_{Y,X}(1,1) = \tfrac{p}{2} \qquad p_{Y,X}(0,1) = \tfrac{p}{2}$$

Using $i(y|x) = -p_{Y,X}(y,x) \log \frac{p_{Y,X}(y,x)}{p_X(x)}$ we obtain:

$$i(0|0) = i(1|0) = 0 \qquad i(0|1) = i(1|1) = \tfrac{p}{2}$$

and so $H(Y|X) = \sum i(y|x) = p$.

On the other hand, the marginal distribution $p_Y(y)$ is:

$$p_Y(0) = 1 - p + \tfrac{p}{2} = 1 - \tfrac{p}{2} \qquad p_Y(1) = \tfrac{p}{2}$$

and so $Y \sim \text{Bern}(\frac{p}{2})$ and thus $H(Y) = H(\frac{p}{2})$.

We may now compute the mutual information:

$$I(X;Y) = H(Y) - H(Y|X) = H(\tfrac{p}{2}) - p = -p - \tfrac{p}{2}\log\tfrac{p}{2} - (1 - \tfrac{p}{2})\log\left(1 - \tfrac{p}{2}\right)$$

This quantity is maximized when:

$$0 = \frac{dI(X;Y)}{dp} = \tfrac{1}{2}\log\left(\tfrac{1}{2p} - \tfrac{1}{4}\right) \qquad \Rightarrow \qquad p = \tfrac{2}{5}$$

and so the channel capacity is:

$$C = \max_{p \in [0,1]} I(X;Y) = I(X;Y)\Big|_{p=\frac{2}{5}} \simeq 0.322 \text{ bits}$$

Question 4

The Noisy Typewriter channel is defined by the conditional PDF:

$$p_{Y|X}(y|x) = \tfrac{1}{2}\delta_{x,y} + \tfrac{1}{2}\delta_{x-1,y} \qquad\qquad x, y = 1, 2, \ldots, N$$

And we will write $p_x \equiv p_X(x)$ for the distribution of X.
Then the joint distribution is:

$$p_{Y,X}(y,x) = p_{Y|X}(y|x)\, p_X(x) = \tfrac{1}{2}\delta_{x,y}p_x + \tfrac{1}{2}\delta_{x-1,y}p_{x-1}$$

which we may marginalize to obtain:

$$p_Y(y) = \sum_x p_{Y,X}(y,x) = \sum_x \left(\tfrac{1}{2}\delta_{x,y}p_x + \tfrac{1}{2}\delta_{x-1,y}p_{x-1}\right) = \tfrac{1}{2}(p_y + p_{y-1})$$

Now, the entropy of Y is:

$$\begin{aligned}
H(Y) &= -\sum_y \tfrac{1}{2}(p_y + p_{y-1}) \log\left(\tfrac{1}{2}(p_y + p_{y-1})\right) \\
&= -\tfrac{1}{2}\sum_y (p_y + p_{y-1})[\log(p_y + p_{y-1}) - 1] \\
&= -\tfrac{1}{2}\sum_y (p_y + p_{y-1}) \log(p_y + p_{y-1}) + \tfrac{1}{2}\sum_y p_y + \tfrac{1}{2}\sum_y p_{y-1} \\
&= 1 - \tfrac{1}{2}\sum_y (p_y + p_{y-1}) \log(p_y + p_{y-1})
\end{aligned}$$

For $X \sim \text{Unif}(1, N)$, that is, for $p_X = 1/N$, we may directly evaluate this expression to obtain $H(Y) = \log N$, which is the maximum entropy for a random variable with N outcomes.
On the other hand, if we set:

$$p_x = \begin{cases} \frac{2}{N} & x \text{ odd} \\ 0 & x \text{ even} \end{cases}$$

then we also find $H(Y) = \log N$.
We may also evaluate the conditional entropy:

$$\begin{aligned}
H(Y|X) &= -\sum_{x,y} p_{Y,X}(y,x) \log p_{p_{Y|X}(y|x)} \\
&= -\sum_y \sum_x \left(\tfrac{1}{2}\delta_{x,y}p_x + \tfrac{1}{2}\delta_{x-1,y}p_{x-1}\right) \log\left(\tfrac{1}{2}\delta_{x,y} + \tfrac{1}{2}\delta_{x-1,y}\right) \\
&= -\sum_y \left(\tfrac{1}{2}p_y + p_{y-1}\right) \log \tfrac{1}{2} = \sum_y p_y = 1
\end{aligned}$$

and so the mutual information of X and Y is:

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - 1$$

But we've already noticed that $H(Y)$ is maximized for the two distributions mentioned above, and so the channel capacity is:

$$C = \max_{p_X(x)} I(X;Y) = \log N - 1$$

which, taking an alphabet of size $N = 26$, yields $C = \log_2 13$ bits.