Solutions to Homework # 3

ECE 6605 Information Theory — 9/24/03

Prof. Steven W McLaughlin

1. a) The state transition diagram looks like:

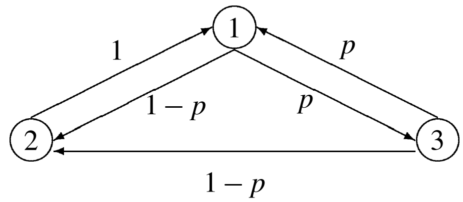

Figure 1: Optical Recording Channel model

Then we can write the following set of equations for $\mu P = \mu$:

$$\begin{aligned} \mu_2 + \mu_3 p &= \mu_1 \\ (1-p)\mu_1 + (1-p)\mu_3 &= \mu_2 \\ \mu_1 p &= \mu_3 \\ \mu_1 + \mu_2 + \mu_3 &= 1 \end{aligned}$$

Solving them is a simple process, and we get the following:

$$\begin{aligned} \mu_1 &= \frac{1}{2+p-p^2} \\ \mu_2 &= \frac{1-p^2}{2+p-p^2} \\ \mu_3 &= \frac{p}{2+p-p^2} \end{aligned}$$

b) Use $H(\mathscr{X}) = \sum_i \mu_i \sum_j (-p_{ij} \log_2 p_{ij})$ then

$$\begin{aligned} H(\mathscr{X}) &= \mu_1[-(1-p)\log_2(1-p) - p\log_2 p] + \mu_2[-1\log_2 1] \\ &\quad + \mu_3[-(1-p)\log_2(1-p) - p\log_2 p] \\ &= \mu_1 H(p) + \mu_3 H(p) \\ &= \frac{1}{2+p-p^2}H(p) + \frac{p}{2+p-p^2}H(p) \\ &= \frac{1+p}{2+p-p^2}H(p) \end{aligned}$$

which evaluated por $p = 1/2$ gives $H(\mathscr{X}) = 2/3$ bits, since $H(p) = 1$ and $(1+p)/(2+p-p^2)|_{p=1/2} = 2/3$

2. For this problem we will use the following definitions: $H(\mathbf{X}_{n-1}) = H(X_1,\dots,X_{n-1})$ if X_1 is not included we will use the starting point, and the ending point, for example $H(\mathbf{X}_{2,n-1}) = H(X_2,\dots,X_{n-1})$.

   a) Begin with the left hand side and use the chain rule for entropy:

$$\begin{aligned}\frac{H(X_1,X_2,\dots,X_n)}{n} &= \frac{1}{n}\sum_{i=1}^{n} H(X_i|\mathbf{X}_{i-1})\\ &= \frac{H(X_n|\mathbf{X}_{n-1})}{n} + \frac{1}{n}\sum_{i=1}^{n-1} H(X_i|\mathbf{X}_{i-1})\\ &= \frac{H(X_n|\mathbf{X}_{n-1})}{n} + \frac{H(X_1,\dots,X_{n-1})}{n}\end{aligned}$$

   Now we will use the stationarity of the process toghether with the fact that conditioning can't increase entropy to establish $H(X_i|\mathbf{X}_{i-1}) = H(X_n|X_{n-1},\dots,X_{n-i+1}) \geq H(X_n|\mathbf{X}_{n-1})$ now we can sum on both sides from $i = 1$ to $n-1$ and then divide by $n-1$ to get:

$$\begin{aligned}\sum_{i=1}^{n-1} H(X_i|\mathbf{X}_{i-1}) &\geq \sum_{i=1}^{n-1} H(X_n|\mathbf{X}_{n-1}) = (n-1)H(X_n|\mathbf{X}_{n-1})\\ \frac{H(X_1,\dots,X_{n-1})}{n-1} &\geq H(X_n|\mathbf{X}_{n-1})\end{aligned}$$

   Combining this result from the one in the previous set of equations we get

$$\begin{aligned}\frac{H(X_1,X_2,\dots,X_n)}{n} &= \frac{H(X_n|\mathbf{X}_{n-1})}{n} + \frac{H(X_1,\dots,X_{n-1})}{n}\\ &\leq \frac{1}{n}\left[\frac{H(X_1,\dots,X_{n-1})}{n-1} + H(X_1,\dots,X_{n-1})\right]\\ &\leq \frac{H(X_1,\dots,X_{n-1})}{n-1}\end{aligned}$$

   b) For this part use the stationarity and conditioning rule from a)

$$\begin{aligned}H(X_n|\mathbf{X}_{n-1}) &\leq H(X_i|\mathbf{X}_{i-1})\\ \sum_{i=1}^{n} H(X_n|\mathbf{X}_{n-1}) &\leq \sum_{i=1}^{n} H(X_i|\mathbf{X}_{i-1})\\ nH(X_n|\mathbf{X}_{n-1}) &\leq H(X_1,\dots,X_n)\\ H(X_n|\mathbf{X}_{n-1}) &\leq \frac{H(X_1,\dots,X_n)}{n}\end{aligned}$$

3. a) This statement is true, we can write

$$\begin{aligned}H(X_n|X_0) &= H(X_n,X_0) - H(X_0)\\ H(X_{-n}|X_0) &= H(X_{-n},X_0) - H(X_0)\end{aligned}$$

   But $H(X_n|X_0) = H(X_{-n}|X_0)$ because of stationarity.

b) This statement is not true in general, although it is true if the process is a Markov Chain. However if it is not we can imagine a periodic process with period n, such that $X_0, X_1, \ldots, X_{n-1}$ are i.i.d. uniformly distributed binary random variables. Let $X_k = X_{k-n}$ for $k \geq n$. Then for this process $H(X_n|X_0) = 0$ and $H(X_{n-1}|X_0) = 1$, which contradicts the statement $H(X_n|X_0) \geq H(X_{n-1}|X_0)$.

c) This statement is true. Use stationarity and then conditioning as follows: $H(X_n|\mathbf{X}_{n-1}X_{n+1}) = H(X_{n+1}|\mathbf{X}_{2,n}X_{n+2} \geq H(X_{n+1}|\mathbf{X}_n X_{n+2})$. Therefore the statement is satisfied.

d) Again the statement is true and we prove it in pretty much the same manner as with c). $H(X_{n+1}|\mathbf{X}_n\mathbf{X}_{n+2,2n+1}) = H(X_{n+2}|\mathbf{X}_{2,n+1}\mathbf{X}_{n+3,2n+2}) \geq H(X_{n+2}|\mathbf{X}_{n+1}\mathbf{X}n+3, 2n+2)$. Therefore the statement is satisfied.

4. We need to calculate the limit as $n \to \infty$ of the following quantity

$$\frac{1}{n}\log_2 \frac{p(X_1, X_2, \ldots, X_n)p(Y_1, Y_2, \ldots, Y_n)}{p(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n)} \tag{1}$$

First notice that since $Z_i = (X_i, Y_i)$ are i.i.d. that implies that (X_i) and (Y_i) are also i.i.d. This then allows us to use our greatly simplified evaluation:

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\log_2 \frac{p(\mathbf{X}_n)p(\mathbf{Y}_n)}{p(\mathbf{X}_n, \mathbf{Y}_n)} &= \lim_{n\to\infty} \frac{1}{n}\log_2 \prod_{i=1}^{n} \frac{p(X_i)p(Y_i)}{p(X_i, Y_i)} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{p(X_i)p(Y_i)}{p(X_i, Y_i)} \\
&\longrightarrow E_{p(x,y)}\left(\log_2 \frac{p(X)p(Y)}{p(X,Y)}\right) \\
&= -I(X;Y)
\end{aligned}$$

5. a) We begin by writing $I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$ now we know from conditioning that $H(X|Z) \leq H(X)$ so we can know write $I(X;Y|Z) \leq H(X) - H(X|Y,Z)$ now we write $I(X,Y) = H(X) - H(X|Y)$, comparing this and the right hand side we realize that we could prove $I(X;Y|Z) \leq I(X;Y)$ if and only if we can say that $H(X|Y,Z) \geq H(X|Y)$ but since we know that conditioning reduces entropy it can't be greater so we would have to show that $H(X|Y,Z) = H(X|Y)$, which we can't guarantee in general.

   As a counterexample, let X, Y be independent fair binary random variables and let $Z = X + Y$. In this case $I(X;Y) = 0$ from independence, however $I(X;Y|Z) = H(X|Z) = 1/2$. Therefore in this case $I(X;Y) < I(X;Y|Z)$.

   b) We've already laid the groundwork in part a), all that is missing is what we can say about $H(X|Y,Z)$ when we have a markov chain. Mainly it is not hard to show that in this case $p(x|y) = p(x,y|z)/p(y|z)$ which basically implies that $H(X|Y,Z) = H(X,Y|Z) - H(Y|Z) = H(X|Y)$, therefore we have from above that $I(X;Y|Z) \leq H(X) - H(X|Y,Z) = H(X) - H(X|Y) = I(X;Y)$ or $I(X;Y|Z) \leq I(X;Y)$.

6. We start by using the chain rule:

$$\begin{aligned} H(X_1,\ldots,X_n) &= \sum_{i=1}^{n} H(X_i|\mathbf{X}_{i-1}) \\ &= H(X_1)+H(X_2|X_1)+\sum_{i=3}^{n} H(X_i|\mathbf{X}_{i-1}) \\ &= H(X_1)+H(X_2|X_1)+\sum_{i=3}^{n} H(X_i|X_{i-1}X_{i-2}) \end{aligned}$$

This last step comes from the fact that all we need is the last position and the direction of the movement to specify the next state, and we can get both by knowing the last two states. Now we know that $H(X_1) = 1$ bit and $H(X_2|X_1) = H(p)$. We can now restate the problem for the rest of the states into one that says the following. If we are moving in a positive direction the probability that we make a negative jump is $p = 0.1$ and the probability that we make a positive jump is $1 - p$. It is likewise if we are in the negative case. We see therefore that by considering the direction of jumping (as opposed to the position and direction) we have reduced the problem to that of a simple two state markov chain. Then $H(X_i|X_{i-1}X_{i-2}) = H(p)$. Therefore

$$\begin{aligned} H(X_1,\ldots,X_n) &= 1+H(p)+\sum_{i=3}^{n} H(p) \\ &= 1+H(p)+(n-2)H(p) \\ &= 1+(n-1)H(p) \end{aligned}$$

Now to find the entropy rate $H(\mathscr{X})$ we find the limit as $n \longrightarrow \infty$ of $H(X_1,\ldots,X_n)/n$. Then

$$\begin{aligned} \lim_{n\longrightarrow\infty} \frac{H(X_1,\ldots,X_n)}{n} &= \lim_{n\longrightarrow\infty} \frac{1}{n} + \frac{(n-1)H(p)}{n} \\ &= H(p) \lim_{n\longrightarrow\infty} (1-1/n) \\ H(\mathscr{X}) &= H(p) \end{aligned}$$