Solutions to Homework # 6

ECE 6605 Information Theory — 11/05/03

Prof. Steven W McLaughlin

1. a) First since Y_1 and Y_2 are conditionally independent and conditionally identically distributed given X which implies $H(Y_1,Y_2|X) = H(Y_1|X) + H(Y_2|X)$ and $I(X;Y_1) = I(X;Y_2)$
   Now consider $I(X;Y_1,Y_2)$ and expand as follows

$$\begin{aligned}
I(X;Y_1,Y_2) &= H(Y_1,Y_2) - H(Y_1,Y_2|X) \\
&= H(Y_1) + H(Y_2) - I(Y_1;Y_2) - H(Y_1,Y_2|X) \\
&= H(Y_1) + H(Y_2) - I(Y_1;Y_2) - (H(Y_1|X) + H(Y_2|X)) \\
&= (H(Y_1) - H(Y_1|X)) + (H(Y_2) - H(Y_2|X)) - I(Y_1;Y_2) \\
&= I(X;Y_1) + I(X;Y_2) - I(Y_1;Y_2) \\
&= 2I(X;Y_1) - I(Y_1;Y_2)
\end{aligned}$$

   b) Since $I(Y_1;Y_2) \geq 0$ we have that $I(X;Y_1,Y_2) \leq 2I(X;Y_1)$ and therefore $\max I(X;Y_1,Y_2) \leq \max 2I(X;Y_1)$ so that $C_t \leq C_1$ where C_t is the capacity of the channel with two observations and C_1 is the capacity of the channel with one observation.

2. First let us analyze the original transition matrix T, we will call its capacity C and $\mathbf{p}$ will be the input distribution that achieves capacity, that is $\mathbf{p}$ is the input distribution for which we get $\max I(X;Y)$. Now adding a row to the transition matrix is equivalent to adding a new input to the channel (since each row in T represents a channel input). Then let T_e be the new channel matrix such that $T_e = [T;\ \mathbf{r}_{n+1}]$ and C_e be the capacity of the expanded channel.

   Now we want to show that $C_e \geq C$. To do this we need to find an input distribution $\mathbf{p_e}$ which can always be found that guarantees that C_e is at least equal to C. It is simple to realize that if we let $\mathbf{p_e} = [\mathbf{p}\ \ 0]$, that is we keep the old input distribution and just zero out the new element then we can guarantee that $C_e = C$ with this input distribution. Any other input distributions can only increase the value of C_e since if they don't we have that $C_e = C$ and therefore we have proved that adding a row to the channel transition matrix can't decrease capacity.

3. Let $\tilde{Y}$ denote the output of the cascaded channel, Y the output of the BSC and X the input to the BSC, also α is the erasure probability for the Erasure channel and p is the crossover probability of the BSC. Then we can show $I(X;\tilde{Y}) = (1-\alpha)I(X;Y)$, this is actually true for the cascade of any arbitrary DMC and an erasure channel. By taking the maximum on both sides we find $C_t = (1-\alpha)C$ where C is the capacity of the DMC. For the BSC $C = 1 - H(p)$. Also notice $1-\alpha$ is the capacity of the erasure channel, so basically we will show $C_t = C_e C$. Let

$$E = \begin{cases} 1, & \tilde{Y} = e \\ 0, & \tilde{Y} = Y \end{cases}$$

Then since E is a function of Y,

$$\begin{aligned} H(\tilde{Y}) &= H(\tilde{Y},E) \\ &= H(E)+H(\tilde{Y}|E) \\ &= H(\alpha)+\alpha H(\tilde{Y}|E=1)+(1-\alpha)H(\tilde{Y}|E=0) \\ &= H(\alpha)+(1-\alpha)H(Y) \end{aligned}$$

the last equality comes from the construction of E. Also,

$$\begin{aligned} H(\tilde{Y}|X) &= H(\tilde{Y},E|X) \\ &= H(E|X)+H(\tilde{Y}|X,E) \\ &= H(E)+\alpha H(\tilde{Y}|X,E=1)+(1-\alpha)H(\tilde{Y}|X,E=0) \\ &= H(\alpha)+(1-\alpha)H(Y|X) \end{aligned}$$

Therefore

$$I(X;\tilde{Y}) = H(\tilde{Y} - H(\tilde{Y}|X) = (1-\alpha)I(X;Y)$$

Therefore $C = (1-\alpha)(1-H(p))$.

4. We find the channel transition matrix by finding the number of ways of going from state i to state j in one jump, since we have two states this is a two by two matrix given by:

$$T = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}$$

We find the eigenvalues using $\det(T-\lambda I)=0$ and we get the characteristic equation $\lambda^2-2\lambda-4=0$ which has solutions $\lambda = 1\pm\sqrt{5}$, then $C = \log_2(1+\sqrt{5}) = 1.69424$ bits/ch.use

To obtain a code we will colapse state 1 into state 2 by using 2 channel uses. Let L be a letter space, W be a word space and - dash and . dot, we have a one state transition diagrams with transitions: -, ., L-, L., W-, W. we can design a code for these assuming they are equally likely and using a Huffman code, therefore the code assignemnt will be:

Symbol	$C(X)$
-	11
.	10
L-	011
L.	010
W-	001
W.	000

Then the average word length can be calculated using the fact that the first two symbols send two bits per channel use, whereas the last four symbols send three bits for every two channel uses. Then $L_x = \frac{1}{6}(2+2+4\left(\frac{3}{2}\right)) = 1.666$ bits/ch.use

5. For a (1,7) constrained channel the transition matrix can be obtained as

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then using Matlab we find that $\lambda = 1.6013$ is the largest eigenvalue and therefore $C = \log_2 \lambda = 0.6792$ bits