Homework 4 Solutions

Chapter 7

6) *Noisy typewriter.*

a) If the typewriter prints out whatever key is struck, then the output, Y, is the same as the input, X, and

$$C = \max I(X;Y) = \max H(X) = \log 26, \tag{611}$$

attained by a uniform distribution over the letters.

b) In this case, the output is either equal to the input (with probability $\frac{1}{2}$) or equal to the next letter (with probability $\frac{1}{2}$). Hence $H(Y|X) = \log 2$ independent of the distribution of X, and hence

$$C = \max I(X;Y) = \max H(Y) - \log 2 = \log 26 - \log 2 = \log 13, \tag{612}$$

attained for a uniform distribution over the output, which in turn is attained by a uniform distribution on the input.

c) A simple zero error block length one code is the one that uses every alternate letter, say A,C,E,. . . ,W,Y. In this case, none of the codewords will be confused, since A will produce either A or B, C will produce C or D, etc. The rate of this code,

$$R = \frac{\log(\#\text{ codewords})}{\text{Block length}} = \frac{\log 13}{1} = \log 13. \tag{613}$$

In this case, we can achieve capacity with a simple code with zero error.

8) *The Z channel.* First we express $I(X;Y)$, the mutual information between the input and output of the Z-channel, as a function of $x = \Pr(X = 1)$:

$$\begin{aligned} H(Y|X) &= \Pr(X=0)\cdot 0 + \Pr(X=1)\cdot 1 = x \\ H(Y) &= H(\Pr(Y=1)) = H(x/2) \\ I(X;Y) &= H(Y) - H(Y|X) = H(x/2) - x \end{aligned}$$

Since $I(X;Y) = 0$ when $x = 0$ and $x = 1$, the maximum mutual information is obtained for some value of x such that $0 < x < 1$.
Using elementary calculus, we determine that

$$\frac{d}{dx} I(X;Y) = \frac{1}{2}\log_2 \frac{1 - x/2}{x/2} - 1\,,$$

which is equal to zero for $x = 2/5$. (It is reasonable that $\Pr(X = 1) < 1/2$ because $X = 1$ is the noisy input to the channel.) So the capacity of the Z-channel in bits is $H(1/5) - 2/5 = 0.722 - 0.4 = 0.322$.

7) *Cascade of binary symmetric channels.* There are many ways to solve this problem. One way is to use the singular value decomposition of the transition probability matrix for a single BSC.
Let,

$$A = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

be the transition probability matrix for our BSC. Then the transition probability matrix for the cascade of n of these BSC's is given by,

$$A_n = A^n.$$

Now check that,

$$A = T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1-2p \end{bmatrix} T$$

where,

$$T = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Using this we have,

$$\begin{aligned} A_n &= A^n \\ &= T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & (1-2p)^n \end{bmatrix} T \\ &= \begin{bmatrix} \frac{1}{2}(1+(1-2p)^n) & \frac{1}{2}(1-(1-2p)^n) \\ \frac{1}{2}(1-(1-2p)^n) & \frac{1}{2}(1+(1-2p)^n) \end{bmatrix}. \end{aligned}$$

From this we see that the cascade of n BSC's is also a BSC with probablility of error,

$$p_n = \frac{1}{2}(1-(1-2p)^n).$$

The matrix, T, is simply the matrix of eigenvectors of A.
This problem can also be solved by induction on n.
Probably the simplest way to solve the problem is to note that the probability of error for the cascade channel is simply the sum of the odd terms of the binomial expansion of $(x+y)^n$ with $x = p$ and $y = 1-p$. But this can simply be written as $\frac{1}{2}(x+y)^n - \frac{1}{2}(y-x)^n = \frac{1}{2}(1-(1-2p)^n$.