ECE 563 Midterm Exam Fall 2009 Name:

1. A quaternary channel has transition matrix Q

$$Q = \begin{bmatrix} 1-\beta & \beta & 0 & 0 \\ \beta & 1-\beta & 0 & 0 \\ 0 & 0 & 1-\gamma & \gamma \\ 0 & 0 & \gamma & 1-\gamma \end{bmatrix}$$

a) Suppose that $\gamma = 0$, $\beta = 1$. What is the capacity?

b) Suppose that $\gamma = \beta$. What is the capacity?

c) Suppose that $\gamma \neq \beta$. What is the capacity?

a) With $\gamma = 0, \beta = 1$, $H(X|Y) = 0$, so

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(X) = \log 4 = 2 \text{ bits/symbol}$$

b) If $\gamma = \beta$, the channel is symmetric, so optimal input distribution is $p(x) = \frac{1}{4}$, $x = 1,2,3,4 \Rightarrow p(y) = \frac{1}{4}$, $y = 1,2,3,4$ and $H(Y|X) = H(\beta)$, So

$$C = I(X;Y) = H(Y) - H(Y|X) = 2 - H(\beta)$$

c) When $\gamma \neq \beta$, the channel is no longer symmetric, but the two component BSCs are, so we expect the optimal input to be $\left(\frac{p}{2}, \frac{p}{2}, \frac{1-p}{2}, \frac{1-p}{2}\right)$ for some p. Then the output distribution is the same, so

$$H(Y) = H\left(\tfrac{p}{2}, \tfrac{p}{2}, \tfrac{1-p}{2}, \tfrac{1-p}{2}\right) = -p \log \tfrac{p}{2} - (1-p) \log \tfrac{1-p}{2} \quad 1 + H(p)$$

and $H(Y|X) = \frac{p}{2} H(Y|X=1) + \frac{p}{2} H(Y|X=2) + \frac{1-p}{2} H(Y|X=3) + \frac{1-p}{2} H(Y|X=4) = pH(\beta) + (1-p)H(\gamma)$

So $C = \max_p I = 1 + H(p^*) - p^* H(\beta) - (1-p^*) H(\gamma)$

Where p^* is obtained from

$$I' = -(\log p + \cancel{1} - \log(1-p) - \cancel{1}) - H(\beta) + H(\gamma) = 0 \quad \left(\text{check to make sure } p^* \text{ is a max}\right)$$

$$\Rightarrow p^* = \frac{1}{2^{H(\beta) - H(\gamma)} + 1}$$

2. A binary source has distribution $p = (0.89, 0.11)$. Determine if it is possible to encode (combined source code and channel code) so that the bit error rate is smaller than any given $\epsilon > 0$ for the ternary channel with transition matrix

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

$H(p) = H(0.11) \cong 0.5$ bits

$$C = \max_{p(x)} I(Y;X) = \max_{p(x)} H(Y) - H(Y|X)$$

We can lower-bound this by choosing any $p(x)$

Since the channel is symmetric, we expect $p(x) = \frac{1}{3}$ to result in high mutual information.

Notice that $p(x) = \frac{1}{3} \Rightarrow p(y) = \frac{1}{3}$.

So $H(Y) = H(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = \log 3 \cong 1.58$

and $H(Y|X) = H(\frac{1}{2}) = \log 2 = 1$

So $C \geq 1.58 - 1 = 0.58 > H(p)$

So Shannon's Coding Theorem implies that it is possible to encode so that the bit error rate is arbitrarily small.

3. A Huffman code is used to encode a dictionary of 4,096 English words into sequences of bits. Prove that every codeword is at least two bits long if the most probable symbol has probability less than $\frac{1}{3}$. Does this conclusion depend on the size of the dictionary?

Proof by contradiction:

Assume there is a codeword $c(x)$ for word x that is only one bit long. Since there are more than 2 words, it is clear that there can only be one such word, so x must be the most probable word, or we could modify the code to get one with shorter expected length.

Now consider the encoding process. If x is to recieve a codeword of length 1, its probability must remain at least as large as the sums of probabilities of the groupings of codewords at each stage, until the last stage.

Consider the second to last stage of the encoding process. Let $a_1, \dots, a_{n_1}$ be the words in one group and $b_1, \dots, b_{n_2}$ be the words in the other group.

Let $P_A = P(a_1) + \dots + P(a_{n_1})$, $P_B = P(b_1) + \dots + P(b_{n_2})$, $P_x = P(x)$.

Assume wlog that $P_A \le P_B \le P_x < \frac{1}{3}$.

But then $P_A + P_B + P_x \le 3P_x < 1$, thus we arrive at a contradiction. So it must be that every codeword is at least 2 bits long.

This conclusion does not depend on the size of the dictionary. (Note: $P_{max} < \frac{1}{3} \Rightarrow |\mathcal{X}| > 3$).

4. Use Stirling's approximation ($n! \approx \sqrt{2n\pi}\left(\frac{n}{e}\right)^n$) to show the following statements for large n.

a) Show that $\binom{n}{k} \approx e^{nH_b\left(\frac{k}{n}\right)}$.

b) Show that $\frac{n!}{\prod_k (n_k!)} \approx e^{nH(\mathbf{p})}$ where $p_k = n_k/n$.

c) Consider a data compaction code that processes a source block of length n over a K-ary alphabet by first counting the frequency of the kth source symbol (n_k) for the actual block, then describing the sourceword in two parts, first stating the composition $(n_0, n_1, \ldots, n_{K-1})$ of the block in a compact binary representation, then stating the specific sourceword by giving its position in an ordered list of all sourcewords with this composition. Is this a universal code for large blocklength?

a) $$\binom{n}{k} = \frac{n!}{k!(n-k)!} \approx \frac{\sqrt{2n\pi}\left(\frac{n}{e}\right)^n}{\sqrt{2k\pi}\left(\frac{k}{e}\right)^k \sqrt{2(n-k)\pi}\left(\frac{n-k}{e}\right)^{n-k}} = \left(\frac{n}{k(n-k)2\pi}\right)^{1/2} \frac{n^n}{k^k (n-k)^{n-k}} \frac{\cancel{e^k e^{n-k}}}{\cancel{e^n}}$$

$$\Rightarrow \frac{1}{n}\ln\binom{n}{k} \approx \frac{1}{n}\left[\ln\left(\frac{n}{k(n-k)2\pi}\right)^{1/2} + n\ln n - k\ln k - (n-k)\ln(n-k)\right]$$

$$\approx \ln n - \frac{k}{n}\ln k - \frac{n-k}{n}\ln(n-k)$$

$$= \frac{k}{n}\ln n + \frac{n-k}{n}\ln n - \frac{k}{n}\ln k - \frac{n-k}{n}\ln(n-k)$$

$$= \frac{k}{n}\ln\frac{n}{k} + \frac{n-k}{n}\ln\frac{n}{n-k} = H_b\left(\frac{k}{n}\right)$$

So $\binom{n}{k} \approx e^{nH_b\left(\frac{k}{n}\right)}$

b) Similarly, $$\frac{n!}{\prod_k n_k!} \approx \left(\frac{2\pi n}{\prod_k 2\pi n_k}\right)^{1/2} \frac{n^n}{\prod_k n_k^{n_k}}$$

So $$\frac{1}{n}\ln\frac{n!}{\prod_k n_k!} \approx \frac{1}{n}\left[n\ln n - \sum_k n_k \ln n_k\right]$$

$$= \sum_k \frac{n_k}{n}\ln n - \sum_k \frac{n_k}{n}\ln n_k$$

$$= \sum_k \frac{n_k}{n}\ln\frac{n}{n_k} = H_b\left(\frac{n_1}{n}, \frac{n_2}{n}, \ldots, \frac{n_K}{n}\right) = H_b(p)$$

So $\frac{n!}{\prod_k n_k!} \approx e^{nH(p)}$

4. Use Stirling's approximation ($n! \approx \sqrt{2n\pi}\left(\frac{n}{e}\right)^n$) to show the following statements for large n.

a) Show that $\binom{n}{k} \approx e^{nH_b\left(\frac{k}{n}\right)}$.

b) Show that $\frac{n!}{\prod_k (n_k!)} \approx e^{nH(\mathbf{p})}$ where $p_k = n_k/n$.

c) Consider a data compaction code that processes a source block of length n over a K-ary alphabet by first counting the frequency of the kth source symbol (n_k) for the actual block, then describing the sourceword in two parts, first stating the composition $(n_0, n_1, \ldots, n_{K-1})$ of the block in a compact binary representation, then stating the specific sourceword by giving its position in an ordered list of all sourcewords with this composition. Is this a universal code for large blocklength?

c) Each block cannot contain more than n of any symbol, so we need no more than $K \log n$ bits to state the composition.

Given a composition, there are $\frac{n!}{\prod_k n_k!}$ words with that composition, so we need another $\log \frac{n!}{\prod_k n_k!}$ bits to specify the specific sourceword, so, for large block length, the average number of bits per symbol is

$$\frac{1}{n}\left[K\log n + \log \frac{n!}{\prod_k n_k}\right] \approx \frac{1}{n} \log e^{nH(p)} \cong H(p)$$

So this code is universal for large block length.