EE 121 - Introduction to Digital Communications
Homework 4 Solutions

March 5, 2008

1 (a) To show that $l_i \leq l_j$ for all $i < j$, assume the contrary, that there exists an optimum prefix-free code with $p_i > p_j$ and $l_i > l_j$ for some i and j. Then if we interchange these codewords we have a new code with average length

$$\begin{aligned}\overline{L}' &= \overline{L} - (p_i l_i + p_j l_j) + (p_i l_j + p_j l_i) \\ &= -p_i(l_i - l_j) - p_j(l_j - l_i) \\ &= (p_i - p_j)(l_j - l_i) \\ &< \overline{L},\end{aligned}$$

which contradicts the optimality of the code.

Using this result, we see that l_k is minimized when all preceding constraints of the form $l_i \leq l_j$ for $i < j$ are met with equality, i.e. when we have $l_1 = l_2 = \cdots = l_k$. In this case we have a tree where all the leaves are at the same depth, i.e. $l_k = \log_2 k$. Thus in general $l_k \geq \log_2 k \geq \lfloor \log_2 k \rfloor$.

(b) The unary-binary code uses $l_j = 2\lfloor \log j \rfloor + 1$ bits to encode the length j of the match, therefore

$$\lim_{j \to \infty} \frac{l_j}{\log j} = 2.$$

2 (a)

$$\begin{aligned}\int_{\infty}^{\infty} f_U(u) du &= 2 \int_0^1 c(1-u) du \\ &= c \\ &\Rightarrow c = 1.\end{aligned}$$

(b) For a uniform quantizer we need $\lceil \log M \rceil$ bits.

(c) The minimum achievable average number of bits per source symbol required, is the entropy. Let the r.v. X denote the quantizer output.

$$H(X) = -\sum_k \Pr(X = k) \log_2 \Pr(X = k)$$

We need to find the pmf of X. There are two cases, M even and M odd. We provide the solution for M even. For $k \in \{-M/2, \ldots, \ldots, M/2\}\backslash\{0\}$

$$\begin{aligned}\Pr(X = k) &= \Pr(2(|k|-1)/M < U < 2|k|/M) \\ &= \int_{2(|k|-1)/M}^{2|k|/M} f_U(u)du \\ &= \int_{2(|k|-1)/M}^{2|k|/M} 1 - udu \\ &= \frac{2}{M}\left(1 + \frac{1-2|k|}{M}\right).\end{aligned}$$

Thus

$$\begin{aligned}H(X) &= -\sum_{k \in \{-M/2,\ldots,\ldots,M/2\}\backslash\{0\}} \frac{2}{M}\left(1+\frac{1-2|k|}{M}\right)\log_2 \frac{2}{M}\left(1+\frac{1-2|k|}{M}\right) \\ &= -\sum_{k=1}^{M/2} \frac{4}{M}\left(1+\frac{1-2k}{M}\right)\log_2 \frac{2}{M}\left(1+\frac{1-2k}{M}\right)\end{aligned}$$

(d) When M is large we can approximate this sum by an integral

$$\begin{aligned}H(X) &\to -\int_1^{M/2} \frac{4}{M}\left(1+\frac{1-2x}{M}\right)\log_2 \frac{2}{M}\left(1+\frac{1-2x}{M}\right)dx \\ &= \frac{1}{M^2}\left((M-1)^2\log_2\left(\frac{M-1}{M^2}\right) + \frac{M(1-M/2)}{\ln 2} + 2M(M-2) - \log_2\frac{1}{M^2}\right) - 1 \\ &\to \log_2\frac{1}{M} + 1 - \frac{1}{2\ln 2}\end{aligned}$$

as $M \to \infty$. Thus gap is bounded and converges to $1 - 1/2\ln 2 \approx 0.28$ bits.

3 (a) The number of transmissions is a *geometric* random variable, i.e. $P(N = n) = p^{n-1}(1-p)$

(b) The mean of a geometric r.v. is $\mathbb{E}N = 1/(1-p)$, therefore the system is stable for arrival rates less than $\lambda < 1/\mathbb{E}N = 1-p$ packets per time slot.

(c) No. The best rate of an FEC on the erasure channel is $1-p$.

(d) Answer to part (b) is unchanged as the steady state behavior remains the same.

4 (a) $\Pr(\text{unable to decode}) = p^n$.

(b) Detection rule is to decode the symbol that appears the most times, picking randomly if there is a tie. Suppose a 0 is sent. A 1 will be decoded if there are more than $\lceil n/2 \rceil$ bit flips. Thus

$$\Pr(\text{error}) = \sum_{k=\lceil n/2 \rceil}^{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

(c) When n is large we can use the Gaussian approximation to the Binomial distribution. The mean and variance of the relevant Gaussian are np and $np(1-p)$, respectively. Let N denote the number of bit flips. The pdf of the Gaussian is

$$p_N(x) = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-(x-np)^2/2np(1-p)}.$$

Thus

$$\begin{aligned}
\Pr(\text{error}) &= \Pr(N > n/2) \\
&\to \int_{n/2}^{\infty} \frac{1}{\sqrt{2\pi np(1-p)}} e^{-(x-np)^2/2np(1-p)} dx.
\end{aligned}$$

We can't evaluate this integral explicitly, but we can bound it tail behavior.

$$\Pr(\text{error}) \approx e^{-(n/2-np)^2/2np(1-p)}$$

More precisely the tail behavior can be captured by the following statement concerning the exponent.

$$\lim_{n\to\infty} -\frac{\log \Pr(\text{error})}{n} = \frac{(1/2-p)^2}{2p(1-p)}.$$

The right hand side of the above equation tells us *how fast* the probability of error decays with n. We can compare this decay rate to the erasure channel decay rate, which is

$$\begin{aligned}
\lim_{n\to\infty} -\frac{\log \Pr(\text{unable to decode})}{n} &= \lim_{n\to\infty} -\frac{\log p^n}{n} \\
&= \lim_{n\to\infty} -\frac{\log e^{\log p^n}}{n} \\
&= \lim_{n\to\infty} -\frac{n \log p}{n}. \\
&= \log p
\end{aligned}$$

In simple language, this means that when n is large, the error probability for the erasure channel looks like $\approx e^{-n \log p}$ whereas the error probability for the binary symmetric channel looks like $\approx e^{-n \frac{(1/2-p)^2}{2p(1-p)}}$. If we plot the two functions $\log p$ and $\frac{(1/2-p)^2}{2p(1-p)}$ we see that $\log p$ is much larger for all values of p in the desired range $[0, 1/2]$. Thus the probability of error decays much faster when communicating over the erasure channel, for the same value of p. This is one sense in which the erasure channel is "easier" to communicate over, though it is a minor one. The dominant reason is that we can use simple feedback schemes on the erasure channel (such as in problem 3), because the receiver knows which bits are lost.

5 (a) An encoding matrix for the $(7,4)$ Hamming code is

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

We can decode if and only if after removing the rows of G corresponding to the erasures, the remaining matrix $\overline{G}$ is full rank over the field of integers mod 2, $\mathcal{GF}(2)$. If there are fewer than 3 erasures $\overline{G}$ is always full rank over $\mathcal{GF}(2)$. If there are more than 3 erasures it is never full rank. $\overline{G}$ is full rank for certain 3-erasure combinations, but not all. By enumeration there are 7, 3-erasure combinations that result in $\overline{G}$ having reduced rank. Thus

$$\begin{aligned} \Pr(\text{error}) &= \binom{7}{0}p^7 + \binom{7}{1}p^6(1-p) + \binom{7}{2}p^5(1-p)^2 + \binom{7}{3}p^4(1-p)^3 + 7p^3(1-p)^4 \\ &= p^7 + 7p^6(1-p) + 21p^5(1-p)^2 + 35p^4(1-p)^3 + 7p^3(1-p)^4. \end{aligned}$$

This probability does not depend on the transmitted codeword, only on the rank of $\overline{G}$. This is an artifact of the linearity of the code. When p is small

$$\Pr(\text{error}) \approx 7p^3.$$

This is the probability that one of the 7, 3-erasure combinations that results in $\overline{G}$ having reduced rank, occurs.