Problem Solving Session 4

5.2

How many fingers has a Martian? Let

$$S = \begin{pmatrix} S_1, \ldots, S_m \\ p_1, \ldots, p_m \end{pmatrix}.$$

The S_i's are encoded into strings from a D-symbol output alphabet in a uniquely decodable manner. If $m = 6$ and the codeword lengths are $(l_1, l_2, \ldots, l_6) = (1, 1, 2, 3, 2, 3)$, find a good lower bound on D. You may wish to explain the title of the problem.

Solution

Uniquely decodable codes satisfy Kraft's inequality. Therefore

$$f(D) = D^{-1} + D^{-1} + D^{-2} + D^{-3} + D^{-2} + D^{-3} \leq 1.$$

We have $f(2) = 7/4 > 1$, hence $D > 2$. We have $f(3) = 26/27 < 1$. So a possible value of D is 3. Our counting system is base 10, probably because we have 10 fingers. Perhaps the Martians were using a base 3 representation because they have 3 fingers. (Maybe they are like Maine lobsters ?)

5.3

Slackness in the Kraft inequality. An instantaneous code has word lengths $l_1, l_2, \ldots, l_m$ which satisfy the strict inequality

$$\sum_{i=1}^{m} D^{-l_i} < 1.$$

The code alphabet is $\mathcal{D} = \{0, 1, 2, \ldots, D-1\}$. Show that there exist arbitrarily long sequences of code symbols in $\mathcal{D}^*$ which cannot be decoded into sequences of codewords.

Solution

Slackness in the Kraft inequality. Instantaneous codes are prefix free codes, i.e., no codeword is a prefix of any other codeword. Let $n_{max} = \max\{n_1, n_2, ..., n_q\}$. There are $D^{n_{max}}$ sequences of length n_{max}. Of these sequences, $D^{n_{max}-n_i}$ start with the i-th codeword. Because of the prefix condition no two sequences can start with the same codeword. Hence the total number of sequences which start with some codeword is $\sum_{i=1}^{q} D^{n_{max}-n_i} = D^{n_{max}} \sum_{i=1}^{q} D^{-n_i} < D^{n_{max}}$. Hence there are sequences which do not start with any codeword. These and all longer sequences with these length n_{max} sequences as prefixes cannot be decoded. (This situation can be visualized with the aid of a tree.)
Alternatively, we can map codewords onto dyadic intervals on the real line corresponding to real numbers whose decimal expansions start with that codeword. Since the length of the interval for a codeword of length n_i is D^{-n_i}, and $\sum D^{-n_i} < 1$, there exists some interval(s) not used by any codeword. The binary sequences in these intervals do not begin with any codeword and hence cannot be decoded.

5.5

More Huffman codes. Find the binary Huffman code for the source with probabilities $(1/3, 1/5, 1/5, 2/15, 2/15)$. Argue that this code is also optimal for the source with probabilities $(1/5, 1/5, 1/5, 1/5, 1/5)$.

Solution

More Huffman codes. The Huffman code for the source with probabilities $(\frac{1}{3}, \frac{1}{5}, \frac{1}{5}, \frac{2}{15}, \frac{2}{15})$ has codewords $\{00,10,11,010,011\}$.
To show that this code (*) is also optimal for (1/5, 1/5, 1/5, 1/5, 1/5) we have to show that it has minimum expected length, that is, no shorter code can be constructed without violating $H(X) \leq EL$.

$$H(X) = \log 5 = 2.32 \text{ bits.} \tag{380}$$

$$E(L(*)) = 2 \times \frac{3}{5} + 3 \times \frac{2}{5} = \frac{12}{5} \text{ bits.} \tag{381}$$

Since

$$E(L(\text{any code})) = \sum_{i=1}^{5} \frac{l_i}{5} = \frac{k}{5} \text{ bits} \tag{382}$$

for some integer k, the next lowest possible value of $E(L)$ is 11/5 = 2.2 bits ¡ 2.32 bits. Hence (*) is optimal. Note that one could also prove the optimality of (*) by showing that the Huffman code for the (1/5, 1/5, 1/5, 1/5, 1/5) source has average length 12/5 bits. (Since each Huffman code produced by the Huffman encoding algorithm is optimal, they all have the same average length.)

5.9

Optimal code lengths that require one bit above entropy. The source coding theorem shows that the optimal code for a random variable X has an expected length less than $H(X) + 1$. Give an example of a random variable for which the expected length of the optimal code is close to $H(X) + 1$, i.e., for any $\epsilon > 0$, construct a distribution for which the optimal code has $L > H(X) + 1 - \epsilon$.

Solution

Optimal code lengths that require one bit above entropy. There is a trivial example that requires almost 1 bit above its entropy. Let X be a binary random variable with probability of $X = 1$ close to 1. Then entropy of X is close to 0, but the length of its optimal code is 1 bit, which is almost 1 bit above its entropy.