Solutions: Homework Set # 3

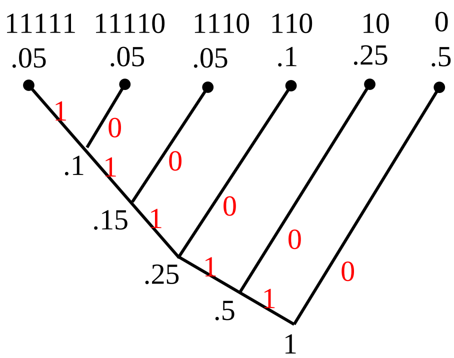
Problem 1

The Kraft's inequality implies $\sum_{i=1}^{6} D^{-l_i} \leq 1$

We define $f(D) = \sum_{i=1}^{6} D^{-l_i}$. It is clear that $f(\cdot)$ is a decreasing function.
Since $f(2) > 1 \Rightarrow D$ can not be 2.
Since $f(3) < 1 \Rightarrow D \geq 3$

Problem 2

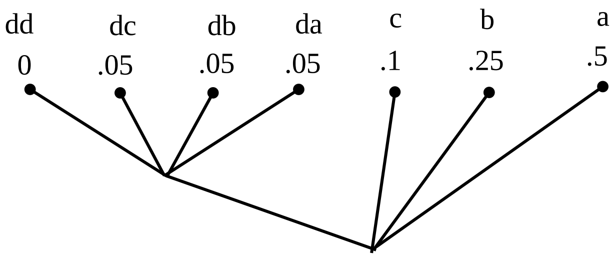
(a) The binary Huffman code for random variable X is



The average length of this code is L_H:

$$L_H = 0.5 \times 1 + 0.25 \times 2 + 0.1 \times 3 + 0.05 \times 4 + 0.1 \times 5 = 2$$

(b) Note that in order to obtain an optimal D-ary Huffman code, we have to add dummy symbols (symbols of zero probability) such that the total number of symbols becomes $k(D-1)+1$, for some integer k. It can be done by adding only one dummy symbol in this example. The quaternary Huffman code for random variable X is



The average length of this code is L_Q:

$$L_Q = 0.15 \times 2 + 0.85 \times 1 = 1.15$$

(c) By performing the mapping, we are doubling the length of each quaternary codeword. Hence

$$L_{QB} = 2L_Q = 2.30$$

(d) Since the binary Huffman code is an optimal binary code, then its average length cannot be less than that of the other binary code. So

$$L_H \leq L_{QH}$$

For the upper bound we first notice that

$$H_2(X) \leq L_H \tag{1}$$

where $H_2(X) = \sum_i p_i * \log_2 \frac{1}{p_i}$ and

$$L_Q \leq H_4(X) + 1 \tag{2}$$

where $H_4(X) = \sum_i p_i \log_4 \frac{1}{p_i} = \frac{1}{2} H_2(X)$. So

$$L_Q \leq \frac{1}{2} H_2(X) + 1 \tag{3}$$

Hence, from part (c) and using (3) we obtain

$$L_{HQ} = 2L_Q < 2(\frac{1}{2} H_2(X) + 1) = H_2(X) + 2 \leq L_H + 2$$

where the last inequality follows from (1).

(e) Suppose that X takes values from A, B, C, D with equal probabilities, $(0.25, 0.25, 0.25, 0.25)$. The binary Huffman code is then

$$\begin{aligned} A &\rightarrow 00 \\ B &\rightarrow 01 \\ C &\rightarrow 10 \\ D &\rightarrow 11 \end{aligned}$$

and the quaternary Huffman code is just a, b, c, d. Hence, the binary code we get after applying the mapping from (c) is the same as the binary Huffman code.

(f) Suppose $\mathcal{C}_2$ be the binary Huffman code with codeword lengths $(l_1, \dots, l_m)$ for a given source. Modify the code by adding a single "0" at the end of any codeword with odd length. Then perform the inverse of mapping given in part (e) to obtain a uniquely decodable quarternary code, $\mathcal{C}_4$ of codeword lengths $(l'_1, \dots, l'_m)$, where

$$l'_i = \begin{cases} \frac{1}{2}l_i & \text{if } l_i \text{ is even,} \\ \frac{1}{2}(l_i+1) & \text{if } l_i \text{ is odd.} \end{cases} = \lceil \frac{l_i}{2} \rceil. \tag{4}$$

It is clear that the average length of the new code satisfies

$$L(\mathcal{C}_4) = \sum_{i=1}^{m} p_i l'_i \leq \sum_{i=1}^{m} p_i \frac{l_i + 1}{2} = \frac{1}{2}(L_H + 1). \tag{5}$$

Since the quaternary Huffman code is optimal, we have

$$L_Q \leq L(\mathcal{C}_4) \leq \frac{1}{2}(L_H + 1), \tag{6}$$

and therefore, using the same argument as in part (d),

$$L_{QB} = 2L_Q \leq 2L(\mathcal{C}_4) \leq L_H + 1. \tag{7}$$

It can be shown that the bound is tight for the random variable X which takes value in A, B with equal probabilities.

Problem 3

(a) We will be using questions to determine the sequence $X_1, X_2, \dots, X_n$, where X_i is 1 or 0 according to whether the i^{th} object is good or defective. Thus the most likely sequence is all 1's with probability of $\prod_{i=1}^{n} p_i$, and the least likely sequence is all 0's sequence with probability $\prod_{i=1}^{n}(1-p_i)$. Since the optimal set of questions correspond to the Huffman code for the source, a good lower bound on the average number of questions is the entropy of the sequence $X_1, X_2, \dots, X_n$. But since the X_i's are independent Bernoulli random variables, we have

$$EQ \geq H(X_1, X_2, \dots, X_n) = \sum H(X_i) = \sum H(p_i).$$

(b) The last bit in the Huffman code distinguishes between the least likely source symbols. (By the condition of the problem all the probabilities are different, and thus the two least likely sequences are uniquely defined.) in this case, the two least likely sequences are $000 \cdots 00$ and $000 \cdots 01$, which have probabilities $(1-p_1)(1-p_2)\cdots(1-p_n)$ and $(1-p_1)(1-p_2)\cdots(1-p_{n-1})p_n$ respectively. Thus the last question will ask "Is $X_n = 1$?" *i.e.,* "Is the last item defective?"

(c) By the same argument as in part (a), an upper bound on the minimum average number of questions is an upper bound in the average length of a Huffman code, namely $H(X_1, X_2, \dots, X_n) + 1 = \sum H(p_i) + 1$

Problem 4

(a) The distribution of X is $\lambda p_1 + (1-\lambda)p_2$

(b) $I(X;Y|Z) = H(Y|Z) - H(Y|X,Z)$

$I(X;Y) = H(Y) - H(Y|X)$

As $p(y|x)$ is fixed,
$$\begin{aligned} H(Y|X=x) &= H(Y|X=x, Z=0) \\ &= H(Y|X=x, Z=1) \end{aligned}$$

Thus,
$$\begin{aligned} H(Y|X) &= \sum_x p(x) H(Y|X=x) \\ &= \sum_x \sum_z p(x,z) H(Y|X=x, Z=z) \\ &= H(Y|X,Z) \end{aligned}$$

Furthermore, $H(Y|Z) < H(Y)$

So

$$I(X;Y|Z) < I(X;Y)$$

Remark: Note that this inequality can be also obtained using the Markov chain $Z \leftrightarrow X \leftrightarrow Y$, as

$$\begin{aligned} I(X,Z;Y) &= I(X;Y) + I(Z;Y|X) \\ &= I(Z;Y) + I(X;Y|Z). \end{aligned}$$

The Markov chain implies $I(Z;Y|X) = 0$, and since $I(Z;Y) \geq 0$, we can conclude $I(X;Y) \geq I(X;Y|Z)$.

(c) By expansion of $I(X;Y|Z)$ with respect to Z we have

$$\begin{aligned} I(X;Y|Z) &= \sum_{z=0,1} p(Z=z) I(X;Y|Z=z) \\ &= \lambda I(X;Y|Z=0) + (1-\lambda) I(X;Y|Z=1) \\ &= \lambda I(X_1;Y) + (1-\lambda) I(X_2;Y). \end{aligned}$$

(d) Replacing $I(X;Y|Z)$ from part (c) in the inequality of part (b), we obtain $\lambda I(X_1;Y) + (1-\lambda) I(X_2;Y_2) < I(X;Y)$ which finishes the proof to the concavity of $I(X;Y)$ in $p(x)$ (when $p(y|x)$ is fixed).

Problem 5

(a) We have seen in the course that $H(x)$ is a concave function of $p(x)$. We consider the random variable X_1 with distribution $P_1 = (p_1, ..., p_i, ...p_j, ...p_m)$ and the random variable X_2 with the distribution $P_2 = (p_1, ..., p_j, ...p_i, ...p_m)$. Define the random variable X as the random variable X_1 with probability $\frac{1}{2}$ and random variable X_2 with probability $\frac{1}{2}$.

$H(X)$ is concave in $p(x)$. So

$$H(X) > \frac{1}{2} H(X_1) + \frac{1}{2} H(X_2)$$

but it is clear that $H(X_1) = H(X_2)$. Therefore, $H(X) > H(X_1)$.

(b) The same argument.

Problem 6

(a) $\lim_{n\to\infty} \frac{-1}{n} \log q(X_1, X_2, ..., X_n) = \lim_{n\to\infty} \dfrac{-1}{n} \log \prod_{i=1}^{n} q(X_i)$

$$= \lim_{n\to\infty} \frac{-1}{n} \sum_{i=1}^{n} \log q(X_i)$$

According to L.L.N (Law of Large Numbers):

$$= -E(\log q(X_1))$$

$$= \sum_{x=1}^{m} p(x) \log \frac{1}{q(x)} \frac{p(x)}{p(x)}$$

$$= D(p\|q) + H(p).$$

(b) $\displaystyle \lim_{n\to\infty} \frac{1}{n} \log \frac{q(X_1, X_2, ..., X_n)}{p(X_1, X_2, ..., X_n)} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log \frac{q(X_i)}{p(X_i)}$

$$= E[\log \tfrac{q(X_1)}{p(X_1)}]$$

$$= -D(p\|q)$$

Thus,

$$\frac{q(x_1, X_2, ..., X_n)}{p(x_1, X_2, ..., X_n)} = 2^{-nD(p\|q)}.$$