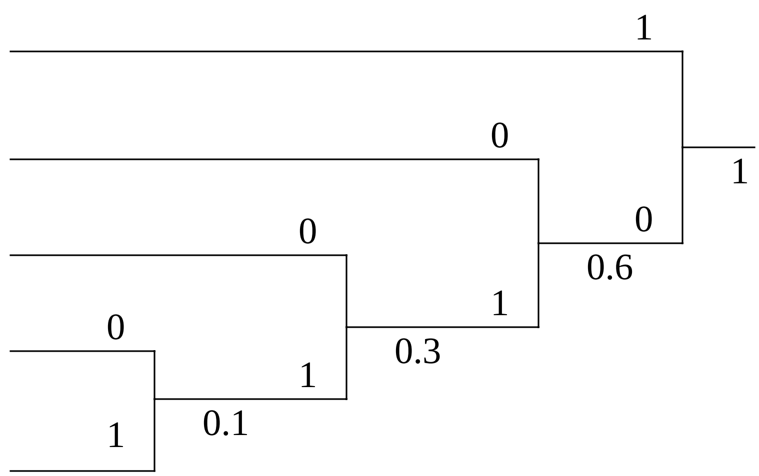
Solutions to Test 2

Problem No. 1

(a)

{A}	P(A)	Huffman Code
4	0.4	1
3	0.3	00
2	0.2	010
1	0.05	0110
0	0.05	0111



$$E[l(x)] = 1(0.4) + 2(0.3) + 3(0.2) + 4(0.5) + 4(0.5) = 2$$

(b)

Yes, the Huffman code above is an example of an optimal prefix code. We can construct a non-prefix code that has a shorter expected codeword length by modifying the longest two codewords as follows:

{A}	Codewords
4	1
3	00
2	010
1	011
0	101

$$E[l(x)] = 1(0.4) + 2(0.3) + 3(0.2) + 3(0.5) + 3(0.5) = 1.9$$

(c)

$$Cost = (2)^{l_i} = C_i$$

l_i = length of ith codeword

The average cost of X is

$$C = \sum_{i=1}^{n} p_i c_i l_i$$

which is what we would like to minimize.

We let

$$r_i = 2^{l_i}$$

$$Q = \sum_i p_i c_i$$

$$q_i = p_i c_i / Q$$

q_i forms a probability distribution.

$$\begin{aligned}
C &= \sum p_i c_i l_i \\
&= -\sum p_i c_i \log l_i \\
&= -\sum Q q_i \log r_i \\
&= Q \sum q_i \log \frac{1}{r_i} \\
&= Q[\sum (q_i \log 1 - q_i \log r_i)] \\
&= Q\left(\sum q_i \log \frac{q_i}{r_i} - \sum q_i \log q_i\right) \\
&= Q(D(q||r) + H(q))
\end{aligned}$$

Since the only freedom is in the choice of r_i, we can minimize C by choosing

$\bar{r} = \bar{q}$

$$2^{-l_i} = \frac{p_i c_i}{\sum p_i c_i}$$

$$l_i^* = -\log \frac{p_i c_i}{\sum p_i c_i}$$

Therefore, the minimum cost C* for the assignment of codewords is

C* = Q H(q)

We con construct a Huffman code with minimum expected cost by using distribution q instead of distribution p.

Problem No. 2

(a)

$$C = max_p I(X;Y) = max_p [H(Y) - H(Y|X)] \le max_p [1 - H(Y|X)]$$

equally holds when the input distribution is uniform. The value of p that minimizes the capacity of the channel is p = 1/2.

C = 1 - 1 = 0

(b)

The transistion probability matrix for a single binary sysmetric channel is

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

???????

Problem No. 3

(a)

$$h(f) = -\int_0^{\infty} f(x)\log f(x)dx$$

$$= -\int_0^{\infty} e^{-x}\ln e^{-x}dx = \int_0^{\infty} e^{-x}(-x)dx = \int_0^{\infty} xe^{-x}dx = \langle xe^{-x} - e^{-x}| = 1$$

(b)

$$h(f) = \int_0^{\infty} 2e^{-2x}\ln(2e^{-2x})dx$$

$$= \int_0^{\infty} e^{-2x}(\ln(2) + \ln(e^{-2x}))dx$$

$$= -2\int_0^{\infty} \ln 2e^{-2x}dx - 2\int_0^{\infty} -2xe^{-2x}dx$$

$$= -2\ln 2\langle\frac{1}{-2}e^{-2x}| - \langle(-x)\frac{1}{-2}e^{-2x}| - \int_0^{\infty}\frac{-1}{2}e^{-2x}1dx = 1.693$$

This answer agrees with the scaling theorem.

Scaling Theorem: h(aX) = h(X) + log|a|
= 1 + log2 = 1.693

(c)

$$h(x) = -\int_3^\infty xe^{-x}\ln e^{-x}dx = -e^3\int_3^\infty e^{-x}\ln e^{-x}e^3dx$$

$$= -e^3\left[\int_3^\infty e^{-x}(\ln e^{-x} + \ln e^{e})dx\right] = -e^3\int_3^\infty(-x)e^{-x}dx - e^3\int_3^\infty 3e^{-x}dx$$

$$= -e^3[-\langle xe^{-x} - e^{-x}| - \langle(-3)e^{-x}|] = 1$$

The answer agrees with the translation theorem. Part c is a translation of part a.

Translation Theorem: h(X+c) = h(X)
Translation does not change the differential entropy.

(d)

$$D(f||g) = \int_0^\infty f\ln\frac{f}{g}dx = \int_0^\infty f(\ln f - \ln g)dx = \int_0^\infty f\ln f dx - \int_0^\infty f\ln g dx$$

$$= \int_0^\infty e^{-x}\ln e^{-x}dx - \int_0^\infty e^{-s}\ln(2e^{-2x})dx = \int_0^\infty e^{-x}\ln e^{-x}dx - \int_0^\infty e^{-x}(\ln 2 + \ln e^{-2x})dx$$

$$= \int_0^{\infty} xe^{-x}dx - \int_0^{\infty} e^{-x}\ln 2dx - \int_0^{\infty} (-e^{-x}\ln e^{-2x})dx$$

$$= \langle xe^{-x} - e^{-x}| - \langle -\ln 2 \cdot e^{-x}| - \langle -(-2x)e^{-x}| = 1 - \ln 2 = 0.3069$$