Chapter III Problems Solutions

1.
   - Find the entropy of a source that emits one of three symbols A, B and C in a statistically independent sequence with probabilities $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{4}$ respectively.
   - Entropy of a sequence is:

$$H(s) = \sum_{i=1}^{3} p_i \log_2 \frac{1}{p_i}$$

$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{4}\log_2(4) = 1.5 \text{ bits/symbol}$$

2. • A discrete source emits one of five symbols once every millisecond. The symbol probabilities are $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$ and $\frac{1}{16}$, respectively. Find the source entropy and information rate.

•

$$H(s) = \sum_{i=1}^{5} p_i \log_2 \frac{1}{p_i}$$

$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{16}\log_2(16) + \frac{1}{16}\log_2(16)$$

$$= 0.5 + 0.5 + 0.375 + 0.25 + 0.25 = 1.875 \text{ bits/symbol}$$

Information rate, R

$$R = r_s H(s) \text{ bits/sec} = 1000 \times 1.875 \text{ bits/sec}$$

3. • For a zero-memory binary source, the source alphabet S is just $\{0,1\}$. Plot the entopy of such a source as a function of $P(0)$. What is the entropy of this source when $P(0) = 0.3$?
Prove the inequality

$$\sum_{i=1}^{N} P_i \log(\frac{1}{P_i}) \leq \sum_{i=1}^{N} P_i \log(\frac{1}{Q_i})$$

with the constraint

$$\sum_{i=1}^{N} P_i = \sum_{i=1}^{N} Q_i = 1$$

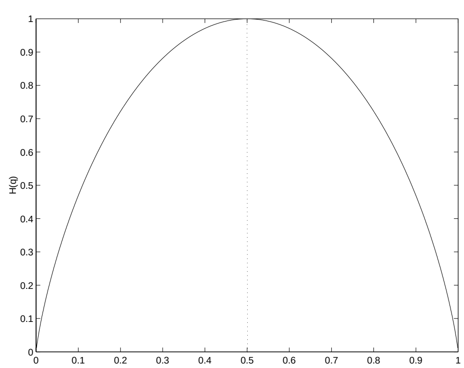
• Let the probability of having a 0 be q, thus the probability of having a 1 is $\bar{q} = 1 - q$. Hence, the entropy of the source is

$$H(s) = q \log(\frac{1}{q}) + \bar{q} \log(\frac{1}{\bar{q}}) \ \text{ bits/symbol}$$

We usually refer to the last function as $H(q)$ which is to be distinguished from $H(S)$, the latter being the entopy of a source S, while the former is a function of the variable q defined over $[0,1]$. Now,

$$\lim_{q\to 0} q \log(\frac{1}{q}) = \lim_{q\to 0} \frac{\log(\frac{1}{q})}{q^{-1}} = \lim_{q\to 0} \frac{\frac{d}{dq}\log(\frac{1}{q})}{\frac{dq^{-1}}{dq}} = \lim_{q\to 0} \frac{q}{\frac{-1}{q^2}} = 0$$

A plot of the entropy function is given below.
At $q = 0.3$, $H(q) = -0.3\log(0.3) - 0.7\log(0.7)$



To find $\log_2 r$, let

$$\log_2 r = x \implies 2^x = r \implies x\ln(2) = \ln(r) \implies x = \frac{\ln(r)}{\ln(2)}$$

Therefore $H(q)|_{q=0.3} = 0.8813$

To prove that
$$\sum_{i=1}^{N} P_i \log(\frac{1}{P_i}) \leq \sum_{i=1}^{N} P_i \log(\frac{1}{Q_i})$$
we first note that this equivalent to proving that

$$\sum_{i=1}^{N} P_i \ln(\frac{1}{P_i}) \leq \sum_{i=1}^{N} P_i \ln(\frac{1}{Q_i})$$

(because $\log_2(x) = \ln(x)/\ln(2)$)
From the text we know that $\ln(x) \leq (x-1)$

$$\sum_{i=1}^{N} P_i \log(\frac{Q_i}{P_i}) \leq \sum_{i=1}^{N} P_i(\frac{Q_i}{P_i} - 1) = \sum_{i=1}^{N}(Q_i - P_i) = 0$$

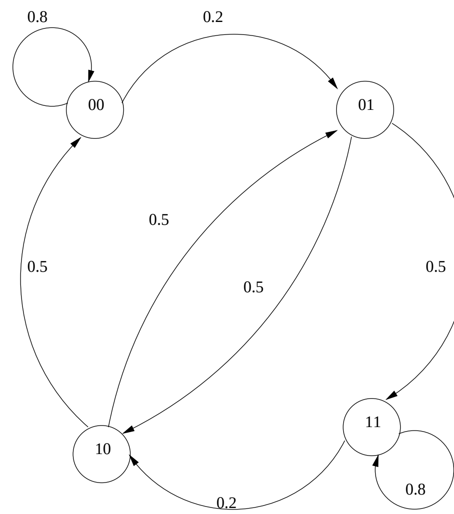
$$\implies \sum_{i=1}^{N} P_i \log(\frac{Q_i}{P_i}) \leq 0 \implies \sum_{i=1}^{N} P_i \log(\frac{1}{P_i}) \leq \sum_{i=1}^{N} P_i \log(\frac{1}{Q_i})$$

4. - A second order Markov source with the binary alphabet $S = \{0, 1\}$ has the conditional symbol probabilities

$$\begin{aligned} P(0|00) = P(1|11) &= 0.8 \\ P(1|00) = P(0|11) &= 0.2 \\ P(0|01) = P(0|10) = P(1|10) = P(1|01) &= 0.5 \end{aligned}$$

   Draw the state diagram of this source.

   - Let N be the number of symbols the source can emit, i.e. $N = 2$, m be the order of the Markov process, i.e. $m = 2$. Thus the total number of states $= N^m = 4$
   The state diagram is shown below: Note that $\sum_{j=1}^{4} p_{ij} = 1$

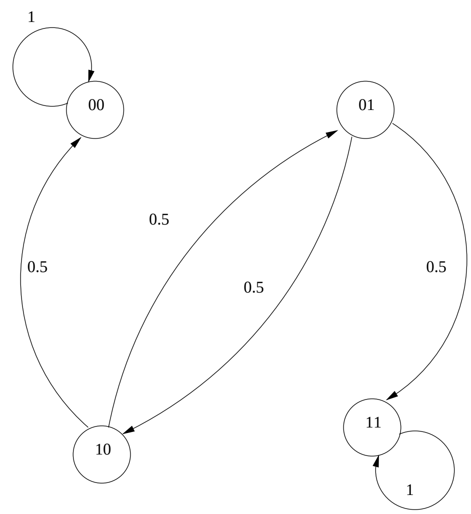


5. • A second-order Markov source with the binary alphabet $S = \{0, 1\}$ has the conditional symbol probabilities

$$P(0|00) = P(1|11) = 1$$
$$P(0|01) = P(0|10) = P(1|10) = P(1|01) = 0.5$$

Draw the state diagram of this source. Is this source ergodic?

• The state diagram is shown below:
Note that if we arrive at either state 00 or 11, we stay in that state



forever. Thus this is not an ergodic source.

6. • The transition matrix of a first-order Markov source having M states i sdenoted by Φ. Let

$$\Phi^n = \begin{bmatrix} q_{11} & \cdots & q_{1M} \\ \vdots & & \vdots \\ q_{M1} & \cdots & q_{MM} \end{bmatrix}$$

Show that $\sum_{j=1}^{M} q_{ij} = 1$.

• Let Φ and Θ be two matrices of the M^{th} order. That is

$$\Phi = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1M} \\ \vdots & & \vdots \\ \phi_{M1} & \cdots & \phi_{MM} \end{bmatrix} \quad \text{and} \quad \Theta = \begin{bmatrix} \theta_{11} & \cdots & \theta_{1M} \\ \vdots & & \vdots \\ \theta_{M1} & \cdots & \theta_{MM} \end{bmatrix}$$

such that
$\sum_{j=1}^{M} \phi_{ij} = \sum_{j=1}^{M} \theta_{ij} = 1$

Let

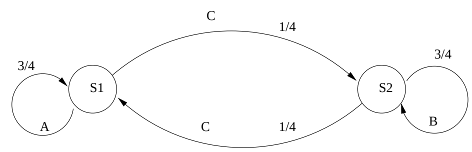
$$\Psi = \Phi.\Theta = \begin{bmatrix} \psi_{11} & \cdots & \psi_{1M} \\ \vdots & & \vdots \\ \psi_{M1} & \cdots & \psi_{MM} \end{bmatrix}$$

Then

$$\sum_{j=1}^{M} \psi_{ij} = \sum_{j=1}^{M} \sum_{k=1}^{M} \phi_{ik}\theta_{kj} = \sum_{k=1}^{M} \phi_{ik}[\sum_{j=1}^{M} \theta_{kj}] = \sum_{k=1}^{M} \phi_{ik} = 1$$

Hence conclusion follows

7. • Consider an information source modeled bya first-order Markov process. The source has two states σ_1 and σ_2 and can emit three symbols A, B and C. The probability of emitting any of the thress symbols from each state is indicated in the figure below. The probability of



the states, $P(\sigma_1) = P(\sigma_2) = \frac{1}{2}$. Find the source entropy H. Consider a sequence m consisting of n symbols emitted from the source. We define the average information content per symbol in messages containing n symbols

$$G_n = \frac{1}{n} \sum_i P(m_i) \log_2 \frac{1}{P(m_i)}$$

where the sum is over all sequences m_i consisting of n symbols. Find G_1, G_2 and G_3.

• Let H_i be the entropy of the source at state σ_i, $i = 1, 2$.

$$H_i = \sum_{j=1}^{2} p_{ij} \log \frac{1}{p_{ij}} \implies H_1 = \frac{1}{4} \log(4) + \frac{3}{4} \log \frac{4}{3} = 0.8113$$

Similarly, $H_2 = \frac{1}{4} \log(4) + \frac{3}{4} \log \frac{4}{3} = 0.8113$
Thus the source entropy

$$H(s) = \sum_{i=1}^{2} P_{\sigma_i} H_i = 0.8113 \text{ bits/symbol}$$

– Let us consider the messages containing one ($n = 1$) symbol. There are 3 messages and the probability of each can be calculated as follows.

$$P(A) = P_{\sigma_1}.P(A|\sigma_1) = \frac{1}{2}.\frac{3}{4} = \frac{3}{8}$$
$$P(B) = P_{\sigma_2}.P(B|\sigma_2) = \frac{1}{2}.\frac{3}{4} = \frac{3}{8}$$
$$P(C) = P_{\sigma_1}.P(C|\sigma_1) = P_{\sigma_2}.P(C|\sigma_2) = \frac{1}{2}.\frac{1}{4} + \frac{1}{2}.\frac{1}{4} = \frac{1}{4}$$

Thus the average information content per symbol in messages containing one symbol is

$$G_1 = \sum_{i=1}^{3} P(m_i) \log_2 \frac{1}{P(m_i)} = 1.5612 \text{ bits/symbol}$$

– Consider the messages containing two ($n = 2$) symbols. There are 9 message ($= N^2$). The probability of the messages is obtained as follows:

$$P(AA) = P_{\sigma_1}.P(A|\sigma_1).P(A|\sigma_1) = \frac{1}{2}.\frac{3}{4}.\frac{3}{4} = \frac{9}{32}$$
$$P(AB) = P_{\sigma_1}.P(A|\sigma_1).P(B|\sigma_1) = 0$$
$$P(AC) = P_{\sigma_1}.P(A|\sigma_1).P(C|\sigma_1) = \frac{3}{32}$$
$$P(BA) = 0$$
$$P(BB) = P_{\sigma_2}.P(B|\sigma_2).P(B|\sigma_2) = \frac{9}{32}$$
$$P(BC) = P_{\sigma_2}.P(B|\sigma_2).P(C|\sigma_2) = \frac{3}{32}$$
$$P(CA) = P_{\sigma_2}.P(C|\sigma_2).P(A|\sigma_1) = \frac{3}{32}$$
$$P(CB) = P_{\sigma_1}.P(C|\sigma_1).P(B|\sigma_2) = \frac{3}{32}$$
$$P(CC) = P_{\sigma_1}.P(C|\sigma_1).P(C|\sigma_2) + P_{\sigma_2}.P(C|\sigma_2).P(C|\sigma_1) = \frac{2}{32}$$

$$\Longrightarrow \quad G_2 = \frac{1}{2}\sum_{i=1}^{9} P(m_i)\log_2\frac{1}{P(m_i)} = 1.2799 \text{ bits/symbol}$$

– In a similar way, the probabilities of the 27 messages of 3 symbols each are calculated and are found to be

$$\begin{array}{rrr}
P(AAA) = \frac{27}{128} & P(AAB) = 0 & P(AAC) = \frac{9}{128} \\
P(ABA) = 0 & P(ABB) = 0 & P(ABC) = 0 \\
P(ACA) = 0 & P(ACB) = \frac{9}{128} & P(ACC) = \frac{3}{128} \\
P(BAA) = 0 & P(BAB) = 0 & P(BAC) = 0 \\
P(BBA) = 0 & P(BBB) = \frac{27}{128} & P(BBC) = \frac{9}{128} \\
P(BCA) = \frac{9}{128} & P(BCB) = 0 & P(BCC) = \frac{3}{128} \\
P(CAA) = \frac{9}{128} & P(CAB) = 0 & P(CAC) = \frac{3}{128} \\
P(CBA) = 0 & P(CBB) = \frac{9}{128} & P(CBC) = \frac{3}{128} \\
P(CCA) = \frac{3}{128} & P(CCB) = \frac{3}{128} & P(CCC) = \frac{2}{128}
\end{array}$$

$$\Longrightarrow \quad G_3 = \frac{1}{3}\sum_{i=1}^{27} P(m_i)\log_2\frac{1}{P(m_i)} = 1.0970 \text{ bits/symbol}$$

Note that $G_1 > G_2 > G_3$

8. • Use the Shannon-Fano coding procedure to design a source encoder for the information source given in Problem 7. For $n = 1, 2$ and 3, calculate the average number of bits per symbol (i.e., L/n), G_n, and the rate efficiency $\eta = \frac{H}{L/n}$

• From Problem 7, $H = 0.8113$. For $n = 1$, the Shannon-Fano codewords and corresponding wordlengths are given below:

Message	Prob's P_i	$\log \frac{1}{P_i}$	l_i	F_i	Code
$n = 1$					
A	$\frac{3}{8}$	1.415	2	0	00
B	$\frac{3}{8}$	1.415	2	$\frac{3}{8}$	01
C	$\frac{1}{4}$	2	2	$\frac{6}{8}$	11
$n = 2$					
AA	$\frac{9}{32}$	1.83	2	0	00
BB	$\frac{9}{32}$	1.83	2	$\frac{9}{32}$	01
AC	$\frac{3}{32}$	3.415	4	$\frac{9}{16}$	1001
BC	$\frac{3}{32}$	3.415	4	$\frac{21}{32}$	1010
CA	$\frac{3}{32}$	3.415	4	$\frac{24}{32}$	1100
CB	$\frac{3}{32}$	3.415	4	$\frac{27}{32}$	1101
CC	$\frac{2}{32}$	4.00	4	$\frac{30}{32}$	1111
$n = 3$					
AAA	$\frac{27}{128}$	2.245	3	0	000
BBB	$\frac{27}{128}$	2.245	3	$\frac{27}{128}$	001
AAC	$\frac{9}{128}$	3.83	4	$\frac{54}{128}$	0110
ACB	$\frac{9}{128}$	3.83	4	$\frac{63}{128}$	0111
BBC	$\frac{9}{128}$	3.83	4	$\frac{72}{128}$	1001
BCA	$\frac{9}{128}$	3.83	4	$\frac{81}{128}$	1010
CAA	$\frac{9}{128}$	3.83	4	$\frac{63}{128}$	1011
CBB	$\frac{9}{128}$	3.83	4	$\frac{99}{128}$	1100
ACC	$\frac{3}{128}$	5.14	6	$\frac{108}{128}$	110110
BCC	$\frac{3}{128}$	5.14	6	$\frac{111}{128}$	110111
CAC	$\frac{3}{128}$	5.14	6	$\frac{114}{128}$	111001
CBC	$\frac{3}{128}$	5.14	6	$\frac{117}{128}$	111010
CCA	$\frac{3}{128}$	5.14	6	$\frac{120}{128}$	111100
CCB	$\frac{3}{123}$	5.14	6	$\frac{123}{128}$	111101
CCC	$\frac{2}{128}$	6	6	$\frac{126}{128}$	111111

For $n = 1$ Average wordlength, $L = 2$.

Average wordlength per symbol, $L/n = 2$ bits/symbol

Average information content per symbol ,$G_1 = 1.5612$ bits/symbol

Efficiency, $\eta = \frac{H}{L/n} \frac{0.8113}{2} = 40.56\%$

For $n = 2$ Average wordlength, $L = 2.88$.

Average wordlength per symbol, $L/n = 1.44$ bits/symbol

Average information content per symbol ,$G_2 = 1.2799$ bits/symbol

Efficiency, $\eta = \frac{0.8113}{2} = 56.34\%$

For $n = 3$ Average wordlength, $L = 3.89$.

Average wordlength per symbol, $L/n = 1.30$ bits/symbol

Average information content per symbol ,$G_3 = 1.097$ bits/symbol

Efficiency, $\eta = \frac{0.8113}{1.30} = 62.40\%$

9. • Let S be a zero-memory information source with source alphabet $\{s_1, s_2, \ldots, s_N\}$ and the probability of s_i equal to P_i. Let the n^{th} extension of S be S^n. Show that the entropy of S^n is given by

$$H(S^n) = nH(S)$$

• Let x_i be the symbol of the n^{th} extension, S^n, of the zero-memory source S.
$x_i \equiv \{s_{i_1}, s_{i_2}, \ldots, s_{i_n}\}$. Now the entropy of S^n is given by:

$$H(S^n) = \sum_{i_1, i_2, \ldots, i_n} P(s_{i_1}, s_{i_2}, \ldots, s_{i_n}) \log \frac{1}{P(s_{i_1}, s_{i_2}, \ldots, s_{i_n})}$$

But, from independence, we have

$$\sum_{i_1, i_2, \ldots, i_n} P(s_{i_1}, s_{i_2}, \ldots, s_{i_n}) = \sum_{i_1, i_2, \ldots, i_n} P(s_{i_1}) P(s_{i_2}) \cdots P(s_{i_n})$$

Therefore,

$$\log \frac{1}{P(s_{i_1}, s_{i_2}, \ldots, s_{i_n})} = \log \left(\frac{1}{P(s_{i_1})} \right)^n = n \log \frac{1}{P(s_{i_1})}$$

$$\Longrightarrow \quad H(S^n) = \sum_{i_1} P(s_{i_1}) \{ n \log \frac{1}{P(s_{i_1})} \} \sum_{i_2} P(s_{i_2}) \sum_{i_3} P(s_{i_3}) \cdots \sum_{i_n} P(s_{i_n})$$

$$= n \sum_{i_1} P(s_{i_1}) \log \frac{1}{P(s_{i_1})} = nH(S)$$

10. • Consider the source $S = \{s_1, s_2, s_3\}$ with $P(s_1) = \frac{1}{2},, P(s_2) = P(s_3) = \frac{1}{4}$. Calculate the entropy of the second extension of S.

• Second extension contains $3^2 = 9$ symbols and their probabilities are:

Symbols of $S \times S$	Probability $P(x_i)$
$x_1 = s_1 s_1$	$\frac{1}{4}$
$x_2 = s_1 s_2$	$\frac{1}{8}$
$x_3 = s_1 s_3$	$\frac{1}{8}$
$x_4 = s_2 s_1$	$\frac{1}{8}$
$x_5 = s_2 s_2$	$\frac{1}{16}$
$x_6 = s_2 s_3$	$\frac{1}{16}$
$x_7 = s_3 s_1$	$\frac{1}{18}$
$x_8 = s_3 s_2$	$\frac{1}{16}$
$x_9 = s_3 s_3$	$\frac{1}{16}$

$$H(S^2) = \sum_{i=1}^{9} P(x_i) \log \frac{1}{P(x_i)} = 3 \text{ bits/symbol} = 2H(S)$$

11. • Construct two sets of compact codes for a source S consisting of six symbols, $s_1, s_2, s_3, s_4, s_5, s_6$ with corresponding probabilities $0.4, 0.3, 0.1, 0.1, 0.06, 0.04$

• First Set of Codes

Original Source			Reduced Sources							
Symbols	Prob's	Code	S_1		S_2		S_3		S_4	
s_1	0.4	1	0.4	1	0.4	1	0.4	1	→0.6	0
s_2	0.3	00	0.3	00	0.3	00	0.3	00→	0.4	1
s_3	0.1	011	0.1	011	→0.2	010→	→0.3	01→		
s_4	0.1	0100	0.1	0100→	0.1	011→				
s_5	0.06	01010→	→0.1	0101→						
s_6	0.04	01011→								

Second Set of Codes

Original Source			Reduced Sources							
Symbols	Prob's	Code	S_1		S_2		S_3		S_4	
s_1	0.4	1	0.4	1	0.4	1	0.4	1	→0.6	0
s_2	0.3	00	0.3	00	0.3	00	0.3	00→	0.4	1
s_3	0.1	0100	→0.1	011	→0.2	010→	→0.3	01→		
s_4	0.1	0101	0.1	0100→	0.1	011→				
s_5	0.06	0110→	0.1	0101→						
s_6	0.04	0111→								

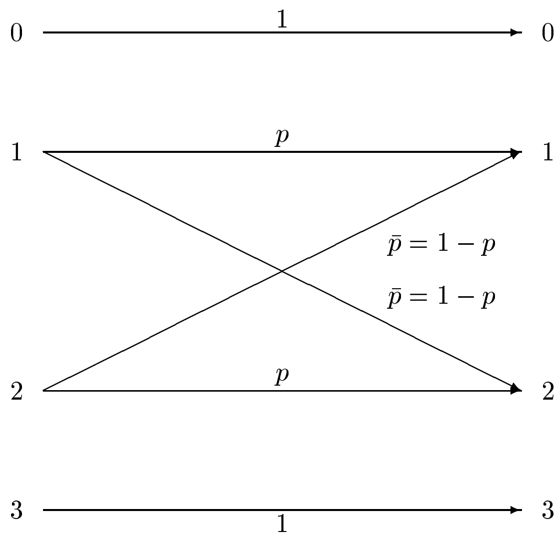
12. • A source S consisting of six symbols, $s_1, s_2, s_3, s_4, s_5, s_6$ with corresponding probabilities $0.30, 0.25, 0.15, 0.12, 0.10, 0.08$ respectively, is coded into 4-ary digits. Find a set of compact codes for the source and calculate the code efficiency.

• We find a compact r-ary code in a way similar to that for a binary code. For an r-ary code, we shall have exactly r messages left in the last reduced set if and only if the total number of original messages is equal to $r + k(r-1)$, where k is an integer. This is obvious since each reduction decreases the number of messages by $(r-1)$. Thus if $r = 4$, in order to have 4 messages remaining in the last reduced source, we must start with $4 + k.3$ messages.
Now, original number of messages is 6. Thus we have to add a dummy message of probability 0 to make 7 messages. Hence we have the following compact code:

Original Source			Reduced Sources	
Messages	Probabilities	Code		
m_1	0.30	0	0.30	0
m_2	0.25	2	→0.30	1
m_3	0.15	3	0.25	2
m_4	0.12	10→	0.15	3
m_5	0.10	11→		
m_6	0.08	12→		
m_7	0.00	13→		

13. • Calculate the capacity of the discrete channel shown below. Assume $r_s = 1$ symbol/sec.

$$P(x=0) = P$$
$$P(x=1) = Q$$
$$P(x=2) = Q$$
$$P(x=3) = P$$



• Let $\alpha = -(p \log p + \bar{p} \log \bar{p})$, then from the definition of channel capacity,

$$C = \max[H(x) - H(x|y)]r_s = \max[I(x;y)r_s]$$

subject to the constraint

$$2P + 2Q = 1 \quad \text{or} \quad Q = \frac{1}{2} - P$$

$$H(x) = -(2P \log P + 2Q \log Q)$$

and

$$H(x|y) = -2Q(p \log p + \bar{p} \log \bar{p}) = 2\alpha Q$$

Hence

$$I(x;y) = -2P \log P - 2Q \log Q - 2\alpha Q$$

We want to maximize $I(x;y)$ with respect to P and Q subject to $Q = \frac{1}{2} - P$. Thus,

$$I(x;y) = -2P \log P - 2(\frac{1}{2} - P) \log(\frac{1}{2} - P) - 2\alpha(\frac{1}{2} - P)$$

$$\Longrightarrow \quad \frac{dI(x;y)}{dP} = 2[-\log_2 e - \log_2 P + \log_2 e + \log_2(\frac{1}{2} - P) + \alpha] = 0$$

$$\Longrightarrow \qquad -\log P + \log(\frac{1}{2} - P) = \alpha = 0 \quad \Longrightarrow \quad \log \frac{\beta}{P}(\frac{1}{2} - P) = 0$$

where $\beta = 2^{\alpha}$ Therefore,

$$P = \frac{\beta}{2(1+\beta)} = \frac{2^{\alpha}}{2(1+2^{\alpha})} \quad \text{and} \quad Q = \frac{1}{2(1+2^{\alpha})}$$

The channel capacity is

$$-2(P\log P + Q\log Q + \alpha Q)r_s = \log \frac{2(\beta+1)}{\beta} \quad \text{bits/sec}$$

14. • A binary channel matrix is given by

$$\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}$$

The probabilities of the two symbols being transmitted are $\frac{1}{3}$ and $\frac{2}{3}$, respectively.

(a) Determine the probabilities of the two symbols received at the destination.

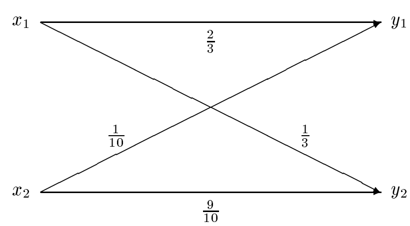
(b) Determine $H(x)$, $H(x|y)$ and $I(x;y)$.

• The channel can be represented as shown.

(a)

$$P(y_1) = P(y_1|x_1)P(x_1) + P(y_1|x_2)P(x_2) = \frac{13}{45}$$

$$P(y_2) = P(y_2|x_1)P(x_1) + P(y_2|x_2)P(x_2) = \frac{32}{45}$$



(b)

$$H(x) = P(x_1)\log\frac{1}{P(x_1)} + P(x_2)\log\frac{1}{P(x_2)} = 0.918 \text{ bits/symbol}$$

To compute $H(x|y)$, we find

$$P(x_1|y_1) = \frac{P(y_1|x_1)P(x_1)}{P(y_1)} = \frac{10}{13}$$

$$P(x_1|y_2) = \frac{P(y_2|x_1)P(x_1)}{P(y_2)} = \frac{5}{32}$$

$$P(x_2|y_1) = \frac{P(y_1|x_2)P(x_2)}{P(y_1)} = \frac{3}{13}$$

$$P(x_2|y_2) = \frac{P(y_2|x_2)P(x_2)}{P(y_2)} = \frac{54}{64}$$

$$H(x|y_1) = P(x_1|y_1)\log\frac{1}{P(x_1|y_1)} + P(x_2|y_1)\log\frac{1}{P(x_2|y_1)} = 0.776$$

$$H(x|y_2) = P(x_1|y_2)\log\frac{1}{P(x_1|y_2)} + P(x_2|y_2)\log\frac{1}{P(x_2|y_2)} = 0.624$$

Hence

$$H(x|y) = H(x|y_1)P(y_1) + [H(x|y_2) = P(y_2) = 0.668$$

$$I(x; y) = H(x) - H(x|y) = 0.25 \ \ \text{bits/symbol}$$

15. • Consider a channel defined by the channel matrix

$$\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mn} \end{bmatrix}$$

where $P_{ij} = P(y_j|x_i)$. The channel is said to be uniform if the terms in every row and every column of the channel matrix consist of an arbitrary permutation of the terms in the first row.
Show that the capacity of a uniform channel is given by:

$$C = \log n - \sum_{j=1}^{n} P(y_j|x_i) \log \frac{1}{P(y_j|x_i)}$$

• To find the capacity of a uniform channel, we first find the mutual information $I(x;y)$

$$I(x;y) = I(y;x) = H(y) - H(y|x) = H(y) - \sum_{i=1}^{m} P(x_i) \sum_{j=1}^{n} P(y_j|x_i) \log \frac{1}{P(y_j|x_i)}$$

Now $P(y_j|x_i)$ for a given i are all the terms in the i^{th} row of the channel matrix, i.e. $P(y_j|x_i) = P_{ij}$.
For a uniform channle, all rows contain the same elements in various permutations, thus the sum $(\sum_{j=1}^{n} P_{ij} \log \frac{1}{P_{ij}})$ is independent of i.

$$\Longrightarrow \quad \sum_{i=1}^{m} P(x_i) \sum_{j=1}^{n} P_{ij} \log \frac{1}{P_{ij}}) = \sum_{j=1}^{n} P_{ij} \log \frac{1}{P_{ij}})$$

and

$$I(x;y) = H(y) - \sum_{j=1}^{n} P_{ij} \log \frac{1}{P_{ij}})$$

(Note that the last term in the above equation is independent of $P(x_i)$; a characteristic of uniform channels.)
But,

$$C = \max_{P(x_i)} I(x;y) \max_{P(x_i)} H(y) - \sum_{j=1}^{n} P_{ij} \log \frac{1}{P_{ij}})$$

Since the last term in the above equation is independent of $P(x_i)$, then maximizing $I(x;y)$ comes down to maximizing $H(y)$. But,

$$H(y) \leq \log n \quad \text{(since there are } n \text{ received symbols)}$$

$$\Longrightarrow \quad C = \log n - \sum_{j=1}^{n} P(y_j|x_i) \log \frac{1}{P(y_j|x_i)} \quad \text{bits/symbol}$$

16. • An r-ary symmetric channel (rSC) is a particular case of a uniform channel with r input and r output symbols such that the channel matrix is given by

$$\begin{bmatrix} \bar{p} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \frac{p}{r-1} & \bar{p} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \bar{p} \end{bmatrix}$$

Calculate the capacity of the rSC,

• Using the result from problem 15, we have for a r-ary symmetric channel.

$$C = \log r - [\bar{p} \log \frac{1}{\bar{p}} + (r-1)(\frac{p}{r-1} \log \frac{1}{p/(r-1)})]$$

$$= \log r - p\log(r-1) - [p\log\frac{1}{p} + \bar{p}\log\frac{1}{\bar{p}}] = \log r - p\log(r-1) - \Omega(p)$$

17. • Consider a random variable x with uniform probability distribution over the interval (-1,1). THis random variable is amplified to form another random variable y such taht $y = 2x$. Using Eq.(III-80), calculate the entropies of x and y, and explain why the two entropies so calculated differ in value.

•

$$p_X(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

The entropy $H(x)$ is given by

$$H(x) = \int_{-1}^{1} \frac{1}{2} \log 2dx = 1 \text{ bit}$$

$$y = 2x \quad \Longrightarrow \quad p_Y(y) \begin{cases} \frac{1}{4} & -2 \leq x \leq 2 \\ 0 & \text{Otherwise} \end{cases}$$

and

$$H(y) = \int_{-2}^{2} \frac{1}{4} \log 4dx = 2 \text{ bits}$$

Superficially, the entropy of y is twice that of x. But amplification can neither add nor subtract information. Therefore the results must be wrong!!!

We have to remember that $H(x)$ and $H(y)$ are relative entropies and they will be equal only if their reference entropies are equal. The reference entropy R_1 for x is given by

$$R_1 = \lim_{\Delta x \to 0} -\log \Delta x$$

while that for y is R_2 and is given by

$$R_2 = \lim_{\Delta y \to 0} -\log \Delta y$$

$$R_1 - R_2 = \lim_{\Delta x, \Delta y \to 0} \log \frac{\Delta y}{\Delta x} = \log \frac{dy}{dx} = \log 2 = 1 \text{ bit}$$

Thus R_1, the reference entropy for x is hier that the reference entropy R_2 for y. Thus the absolute entropies for both x and y will be equal.

18. • For a continuous random variable x constrained to a peak magnitude M, i.e. $(-M < x < M)$, show that the entropy is a maximum when x is uniformly distributed in the range $(-M, M)$ and has zero probability density function outside this range. Find the maximum entropy for this random variable.

•

$$H(x) = \int_{-\infty}^{\infty} p(x) \log \frac{1}{p(x)} dx = \int_{-M}^{M} p(x) \log \frac{1}{p(x)} dx$$

Also,

$$\int_{-M}^{M} p(x) dx = 1$$

Thus, comparing this with Eqs (III-84,85), we have

$$F(x,p) = -p \log p \implies \frac{\partial F}{\partial p} = -(1 + \log p)$$

$$\Phi_1(x,p) = p \implies \frac{\partial \Phi_1}{\partial p} = 1$$

Substituting these quantities in equation (III-86), we have

$$\frac{\partial F}{\partial p} + \alpha_1 \frac{\partial \Phi_1}{\partial p} = 0 \implies p = 2^{\alpha_1 - 1}$$

and ,

$$\int_{-M}^{M} p(x) dx = \int_{-M}^{M} 2^{\alpha_1 - 1} dx = 2M(2^{\alpha_1 - 1}) = 1$$

Hence,

$$2^{\alpha_1 - 1} = \frac{1}{2M} \implies p(x) = \frac{1}{2M}$$

Thus, maximum entropy occurs when x is evenly distributed. Maximum entropy is given by

$$H(x) = \int_{-M}^{M} p(x) \log \frac{1}{p(x)} dx = \int_{-M}^{M} \frac{1}{2M} \log 2M dx = \log 2M$$

19. • Show that for a continuous channel

$$I(x;y) = I(y;x)$$

• From Eqs.(III-94,95)

$$I(x;y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{XY}(x,y) i(x;y) dx dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{XY}(x,y) \log \frac{p(x|y)}{p(x)} dx dy$$

Using conditional probability rule, $p(x|y) = \frac{p_{XY}(x,y)}{p(y)}$ Thus

$$I(x;y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{XY}(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy$$

which is symmetric with respect to x and y. Thus

$$I(x;y) = I(y;x)$$