

MSc and EEE/ISE PART IV: MEng and ACGI

Q4(c)

Time allowed: 3:00 hours

Answer ALL questions.

All questions carry equal marks

Examiners responsible First Marker(s) : C. Ling
Second Marker(s) : A. Manikas

Information for students

Notation:

- (a) Random variables are shown in Tahoma font. x , \mathbf{x} , \mathbf{X} denote a random scalar, vector and matrix respectively.
- (b) The size of a set A is denoted by $|A|$.
- (c) \oplus denotes the exclusive-or operation, or modulo-2 addition.
- (d) “i.i.d.” means “independent identically distributed”.

The Questions

1.

- a) Let the joint distribution of two random variables X and Y be given by

$p(x, y)$	$y=0$	$y=1$	$y=2$
$x=0$	1/4	0	0
$x=1$	0	1/4	0
$x=2$	0	1/4	1/4

Compute:

- i) The entropy $H(X)$, $H(Y)$
- ii) The conditional entropy $H(X|Y)$, $H(Y|X)$
- iii) The joint entropy $H(X, Y)$
- iv) The mutual information $I(X, Y)$
- v) Draw a Venn diagram for the above quantities.

[10]

- b) Let X and Y be two independent discrete random variables taking integer values. X is uniformly distributed over $\{1, 2, 3, 4\}$ and $P(Y = k) = 2^{-k}$, $k = 1, 2, 3, \dots$.

- i) Find $H(X)$.
- ii) Find $H(Y)$.
- iii) Find $H(X + Y, X - Y)$.

[7]

- c) Let x_1 and x_2 be discrete random variables drawn according to probability mass function $p_1(\cdot)$ and $p_2(\cdot)$ over the respective alphabets $X_1 = \{1, 2, \dots, m\}$ and $X_2 = \{m+1, \dots, n\}$. Let

$$X = \begin{cases} X_1 & \text{with probability } \alpha \\ X_2 & \text{with probability } 1 - \alpha \end{cases}$$

Find $H(X)$ in terms of $H(X_1)$, $H(X_2)$ and α .

[8]

2.

- a) Consider the rate-distortion function $R(D) = \min I(X; \hat{X}), E_{X, \hat{X}} d(X, \hat{X}) \leq D$, where $E_{X, \hat{X}}$ denotes the expectation with respect to X, \hat{X} . Justify each step in the following derivation of the rate-distortion function for a Bernoulli source $X = \{0, 1\}$, $p_X = \{1-p, p\}$ ($p \leq 1/2$), and $d(x, \hat{x}) = x \oplus \hat{x}$. In the following, (1), (2), ..., (6) are the step numbers.

$$\text{If } D \geq p \Rightarrow R(D) \stackrel{(1)}{=} 0;$$

$$\text{If } D < p \leq 1/2,$$

$$I(X; \hat{X}) \stackrel{(2)}{=} H(X) - H(X | \hat{X}) \stackrel{(3)}{=} H(p) - H(X \oplus \hat{X} | \hat{X})$$

$$\stackrel{(4)}{\geq} H(p) - H(X \oplus \hat{X}) \stackrel{(5)}{\geq} H(p) - H(D) \Rightarrow R(D) \stackrel{(6)}{\geq} H(p) - H(D).$$

[9]

- b) Huffman coding. Consider the probability distribution of a random variable X :

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 0.05 & 0.05 & 0.25 & 0.2 & 0.15 & 0.3 \end{pmatrix}$$

- i) Find a binary Huffman code for X .
- ii) Find the expected code length for this code.

[6]

- c) Lempel-Ziv coding. Consider the following all-zero sequence of length n :
 $x^n = 000000000000\dots$

- i) Give the LZ78 parsing and encoding. You may simply use numbers 1, 2, 3, ... to represent the locations.
- ii) Show that the number of encoding bits per symbol for this sequence goes to zero as $n \rightarrow \infty$.

[10]

3.

- a) Calculate the capacity of the following channels with probability transition matrices

i)
$$\mathbf{Q} = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} \quad x, y \in \{0, 1, 2\}$$

ii)
$$\mathbf{Q} = \begin{bmatrix} 1/3 & 1/6 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{bmatrix} \quad x \in \{0, 1\}, y \in \{0, 1, 2\}$$

[8]

- b) Compute the capacity of the concatenated binary symmetric channel shown in Fig. 3.1, where the cross-over probabilities are p and q , respectively.

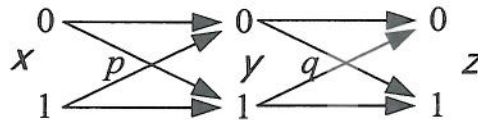


Fig. 3.1. Concatenated binary symmetric channel.

[7]

- c) Fano's inequality. Consider the Markov chain shown in Fig. 3.2, where x is the channel input, y is the channel output, and the estimate of x is simply $\hat{x} = y$.

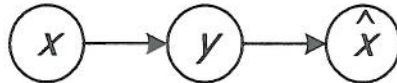


Fig. 3.2. Markov chain.

The input alphabet $X = \{1, 2, 3, 4, 5\}$, probability mass vector $\mathbf{p}_x = [0.35, 0.35, 0.1, 0.1, 0.1]$. The output alphabet $Y = \{1, 2\}$; if $x \leq 2$, then $y = x$ with probability $6/7$, while if $x > 2$, then $y = 1$ or 2 with equal probability.

- Compute the actual error probability.
- Compute the conditional entropy $H(x|y)$.
- Using Fano's inequality $H(x|y) \leq 1 + p_e \log(|X| - 1)$ where P_e is the error probability, compute the Fano bound on the error probability, and compare with i).

[10]

4.

- a) With reference to Fig. 4.1, justify each step in the following proof of the converse of the Gaussian channel coding theorem. That is, if the error probability $P_e^{(n)} \rightarrow 0$ and $n^{-1} \mathbf{x}^T \mathbf{x} < P$ for each $\mathbf{x}(w) = \mathbf{x}_{1:n}$, then the rate $R \leq \frac{1}{2} \log(1 + PN^{-1})$. Here P is the signal power, while N is the noise power. (1), (2), ..., (9) are step numbers.

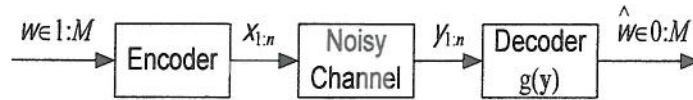


Fig. 4.1. Communication over a noisy channel. $M = 2^{nR}$.

$$\begin{aligned}
 nR &\stackrel{(1)}{=} H(W) \stackrel{(2)}{=} I(W; \mathcal{Y}_{1:n}) + H(W | \mathcal{Y}_{1:n}) \stackrel{(3)}{\leq} I(X_{1:n}; \mathcal{Y}_{1:n}) + H(W | \mathcal{Y}_{1:n}) \\
 &\stackrel{(4)}{=} h(\mathcal{Y}_{1:n}) - h(\mathcal{Y}_{1:n} | X_{1:n}) + H(W | \mathcal{Y}_{1:n}) \stackrel{(5)}{\leq} \sum_{i=1}^n h(Y_i) - h(Z_{1:n}) + H(W | \mathcal{Y}_{1:n}) \\
 &\stackrel{(6)}{\leq} \sum_{i=1}^n I(X_i; Y_i) + 1 + nRP_e^{(n)} \stackrel{(7)}{\leq} \sum_{i=1}^n \frac{1}{2} \log(1 + PN^{-1}) + 1 + nRP_e^{(n)} \\
 &\stackrel{(8)}{\Rightarrow} R \leq \frac{1}{2} \log(1 + PN^{-1}) + n^{-1} + RP_e^{(n)} \stackrel{(9)}{\Rightarrow} R \leq \frac{1}{2} \log(1 + PN^{-1}) \text{ as } n \rightarrow \infty
 \end{aligned}$$

[10]

- b) Calculate the differential entropy of a zero-mean Gaussian random vector with correlation matrix

$$\mathbf{K} = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}.$$

[5]

- c) Slepian-Wolf coding. Let (X, Y) have the joint probability mass function

$p(x,y)$	1	2	3
1	α	β	β
2	β	α	β
3	β	β	α

where $4\beta + 3\alpha = 1$. (Note: this is a joint, not a conditional, probability mass function.)

- i) Find the Slepian-Wolf rate region for this source.
 ii) What is the rate region if $\alpha = 1/3$?

[10]

B - Bookwork
E - new example
A - new application

1. a) The marginal distributions

$p(x)$	
0	$\frac{1}{4}$
1	$\frac{1}{4}$
2	$\frac{1}{2}$

$p(y)$	
0	$\frac{1}{4}$
1	$\frac{1}{2}$
2	$\frac{1}{4}$

$\frac{1}{8}$

i) $H(x) = \frac{1}{4} \times 2 + \frac{1}{4} \times 2 + \frac{1}{2} \times 1 = 1.5$ bits

[2E]

$H(y) = 1.5$ bits

ii) $H(y|x) = \sum_{k=0}^2 H(y|x=k) p(x=k)$
 $= 0 + 0 + \frac{1}{2} = \frac{1}{2}$

[2E]

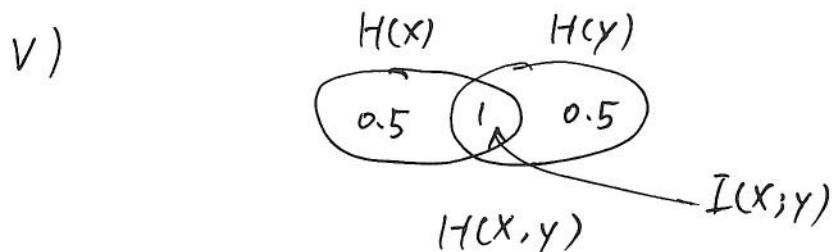
$H(x|y) = 0 + \frac{1}{2} + 0 = \frac{1}{2}$

iii) $H(x, y) = \log 4 = 2$

[2E]

iv) $I(x; y) = H(x) - H(x|y) = 1$

[2E]



[2E]

b) i) $H(x) = \log 4 = 2$

[2B]

ii) $H(y) = - \sum_{k=1}^{\infty} p(y=k) \log p(y=k)$
 $= - \sum 2^{-k} \cdot k$
 $= \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = 2$

[2E]

iii) The mapping $(x, y) \rightarrow (x+y, x-y)$ is one-to-one.

$\therefore H(x+y, x-y) = H(x, y) = H(x) + H(y) = 4$

[3A]

c) The probability mass vector for X is

$$[2A] \quad \alpha p_1(1) \quad \alpha p_1(2) \quad \dots \quad \alpha p_1(m), (1-\alpha)p_2(m+1), \dots, (1-\alpha)p_2(n)$$

$$[2A] \quad H(X) = - \sum_{k=1}^m \alpha p_1(k) \log [\alpha p_1(k)] - \sum_{k=m+1}^n (1-\alpha) p_2(k) \log [(1-\alpha) p_2(k)]$$

$$= - \sum_{k=1}^m \alpha p_1(k) [\log \alpha + \log p_1(k)]$$

$$[2A] \quad - \sum_{k=m+1}^n (1-\alpha) p_2(k) [\log (1-\alpha) + \log p_2(k)]$$

$$= - \alpha \log \alpha - (1-\alpha) \log (1-\alpha) + \alpha \sum_{k=1}^m p_1(k) \log p_1(k) + (1-\alpha) \sum_{k=m+1}^n p_2(k) \log p_2(k)$$

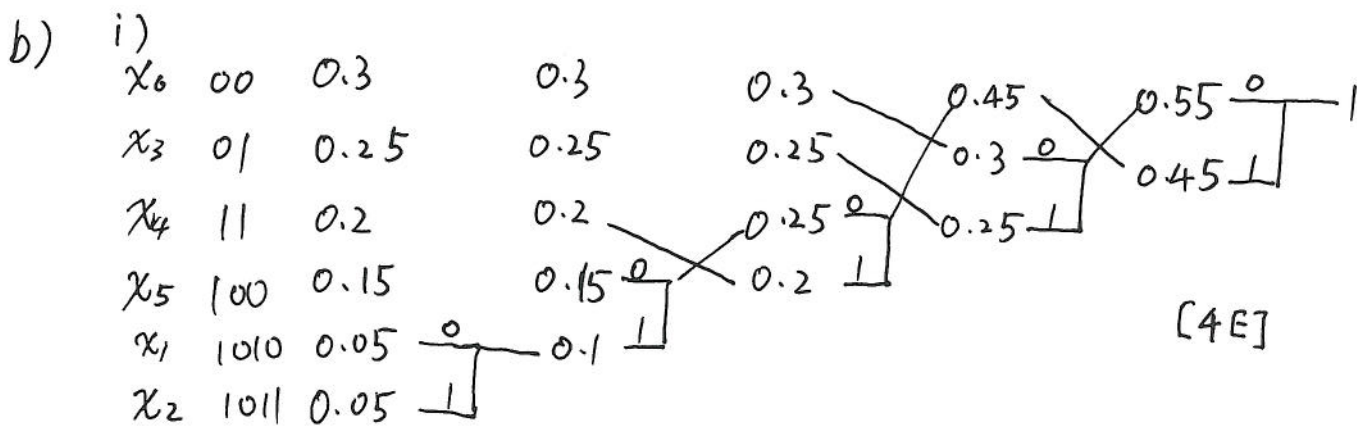
$$[2A] \quad = H(\alpha) + \alpha H(X_1) + (1-\alpha) H(X_2)$$

2. a)
- (1) send nothing when $D \geq p$, as the distortion is p . [1B]
 - (2) definition of mutual information [1B]
 - (3) $H(x) = H(p)$
 $H(x|\hat{x}) = H(x \oplus \hat{x} | \hat{x})$ [1B]
 - (4) conditioning reduces entropy [1B]
 - (5) $H(x \oplus \hat{x}) \leq H(D)$ [3B]
 This is because $E_{x, \hat{x}} d(x, \hat{x}) \leq D$

$$= 1 \cdot \Pr(x \oplus \hat{x} = 1) + 0 \cdot (\Pr(x \oplus \hat{x} = 0))$$

$$= \Pr(x \oplus \hat{x} = 1) \leq D$$

 $H(x \oplus \hat{x}) \leq H(D)$ as $H(p)$ is monotonic when $p \leq \frac{1}{2}$
 - (6) $I(x; \hat{x}) \geq H(p) - H(D)$ [2B]
 $\Rightarrow \min I(x; \hat{x}) \geq H(p) - H(D)$
 $R(D) \geq H(p) - H(D)$



ii) expected length

$$\bar{L} = 2 \times 0.75 + 3 \times 0.15 + 4 \times 0.1$$

$$= 2.35$$

c)

i) location \rightarrow 1 2 3 4 5 [4E]
 0, 00, 000, 0000, 00000, ...

encoding 0, 10, 20, 30, 40, ...

ii) for a length- n sequence, the number k of phrases is given by the equation

$$1 + 2 + 3 + \dots + k = n \quad [2A]$$

$$k(k+1) = n$$

$$k < \sqrt{n}$$

The number of encoded bits

$$< k \cdot (\log_2 k + 1) \quad [2A]$$

The number of bits / symbol

$$< \frac{k \cdot (\log_2 k + 1)}{n}$$

$$< \frac{\sqrt{n} \cdot (\log_2 \sqrt{n} + 1)}{n}$$

[2A]

$$< \frac{\log_2 \sqrt{n} + 1}{\sqrt{n}}$$

$$\rightarrow 0 \text{ as } n \rightarrow \infty$$

3. a) Both are weakly symmetric channels

$$\begin{aligned} \text{i) } C &= \log |Y| - H(Q_{1,:}) \\ &= \log 3 - \left(\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 2 \right) \quad [4E] \\ &= \log 3 - \frac{3}{2} \approx 0.08 \end{aligned}$$

$$\begin{aligned} \text{ii) } C &= \log 3 - \left(\frac{1}{3} \times \log 3 + \frac{1}{6} \times \log 6 + \frac{1}{2} \times 1 \right) \\ &\approx 0.12 \quad [4E] \end{aligned}$$

b) The transitional matrix of the concatenated channel is

$$\begin{aligned} &\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix} \quad [2A] \\ &= \begin{pmatrix} 1-p-q+2pq & p+q-2pq \\ p+q-2pq & 1-p-q+2pq \end{pmatrix} \quad [2A] \end{aligned}$$

This is still a BSC

$$C = 1 - H(p+q-2pq) \quad [3A]$$

c) i) Joint distribution

$y \equiv \hat{x} \backslash x$	1	2	3	4	5
1	0.3	0.05	0.05	0.05	0.05
2	0.05	0.3	0.05	0.05	0.05

[2 E]

error probability is when $\hat{x} \neq x$

$$P_e = 0.4$$

[2 E]

ii) $H(x|y)$ = average row entropy

$$= \cancel{-0.3 \log 0.3 - 4 \times 0.05 \log 0.05}$$

$$= -0.6 \log 0.6 - 4 \times 0.1 \log 0.1$$

[2 E]

$$= 1.771$$

iii)

$$P_e \geq \frac{H(x|y) - 1}{\log(|X| - 1)}$$

[2 E]

$$= \frac{1.771 - 1}{\log 4}$$

$$= \frac{0.771}{2}$$

$$= 0.3855$$

This is a valid lower bound of the actual error probability.

[2 E]

4.

7.

a) (1) uniformly distributed [1 B]

(2) by definition [1 B]

(3) $I(W; Y_{1:n}) \leq I(X_{1:n}; Y_{1:n})$ Markov chain [1 B] $W \rightarrow X \rightarrow Y \rightarrow \hat{W}$

(4) definition [1 B]

(5) indep. bound [1 B]

(6) Fano's inequality [1 B]

(7) $I(X_i; Y_i) \leq \frac{1}{2} \log(1 + \frac{P}{N})$ Capacity is the maximum mutual information.

(8) algebra [1 B]

(9) $P_e^{(n)} \rightarrow 0, \quad n^{-1} \rightarrow 0 \quad \text{as } n \rightarrow \infty$ [2 B]b) $h(x) = \frac{1}{2} \log((2\pi e)^2 |K|)$ [2 E]

$$= \frac{1}{2} \log((2\pi e)^2 \cdot 2)$$

$$= 4.57$$

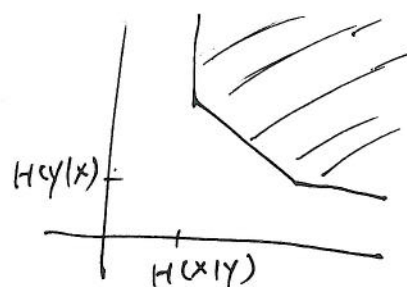
[3 E]

c) i) Shannon - Wolf region

$$R_x \geq H(X|Y)$$

$$R_y \geq H(Y|X)$$

$$R_x + R_y \geq H(X, Y)$$



$$H(X|Y) = H(3\alpha, 3\beta, 3\beta)$$

$$H(Y|X) = H(3\alpha, 3\beta, 3\beta)$$

$$H(X, Y) = H(X|Y) + H(Y)$$

$$= H(3\alpha, 3\beta, 3\beta) + \log 3$$

Both X and Y are uniformly distributed

$$\Rightarrow H(X) = H(Y) = \log 3$$

ii) If $\alpha = \frac{1}{3}$, $\beta = 0$

$$H(X|Y) = H(Y|X) = 0$$

$$H(X, Y) = \log 3$$

$$R_x + R_y \geq \log 3$$

$$R_x \geq 0 \quad R_y \geq 0$$

