## The Solutions

B—bookwork, A—application, E—new example, T—new theory

1.

a)

i) $H(Y) = 0.81$ [3E]

ii) $I(X_1 ; Y) = 0.31$ [3E]

iii) $I(X_{1:2} ; Y) = 0.81$ [3E]

b)

$D(\mathbf{p}||\mathbf{q}) = \sum p_i \log\frac{p_i}{q_i} = \frac{1}{3}\log\frac{2}{3} + \frac{2}{3}\log\frac{4}{3} = 0.0817$ [3E]

$D(\mathbf{q}||\mathbf{p}) = \sum q_i \log\frac{q_i}{p_i} = \frac{1}{2}\log\frac{3}{2} + \frac{1}{2}\log\frac{3}{4} = 0.085$ [3E]

c) i) Firstly note that the output distributions are given by

$$p'_y = \sum_{x\in X} p_x W(y|x)$$

$$q'_y = \sum_{x\in X} q_x W(y|x)$$

[2T]

Thus,

$$D(\mathbf{p}'||\mathbf{q}') = \sum_{y\in Y} p'_y \log\frac{p'_y}{q'_y} = \sum_{y\in Y}\sum_{x\in X} p_x W(y|x)\cdot\log\frac{\sum_{x\in X} p_x W(y|x)}{\sum_{x\in X} q_x W(y|x)}$$

[2T]

Now, using the log-sum inequality, the inner sum

$$\sum_{x\in X} p_x W(y|x)\cdot\log\frac{\sum_{x\in X} p_x W(y|x)}{\sum_{x\in X} q_x W(y|x)}$$

$$\le \sum_{x\in X} p_x W(y|x)\log\frac{p_x W(y|x)}{q_x W(y|x)} = \sum_{x\in X} p_x W(y|x)\log\frac{p_x}{q_x}$$

[3T]

Substituting this back to $D(\mathbf{p}'||\mathbf{q}')$, we obtain

$$D(\mathbf{p}'||\mathbf{q}') \le \sum_{y\in Y}\sum_{x\in X} p_x W(y|x)\log\frac{p_x}{q_x}$$

$$= \sum_{x\in X} p_x \sum_{y\in Y} W(y|x)\log\frac{p_x}{q_x} = \sum_{x\in X} p_x \log\frac{p_x}{q_x} = D(\mathbf{p}||\mathbf{q})$$

where we use the fact that

$$\sum_{y\in Y} W(y|x) = 1$$

[3T]

2.

a) i) (1) chain rule [1B]

(2) chain rule in another way [1B]

(3) $H(e|x,y) \geq 0$ entropy is non-negative [1B]

$H(e|y) \leq H(e)$ conditioning reduces entropy [1B]

(4) total probability theorem

(5) $H(e) = H(p_e)$ [2B]

Given $y$ and $e=0$, $x=y$, so entropy $=0$.

Given $y$ and $e=1$, $x \neq y$ but can take any of the $|X|-1$ values, so entropy $\leq \log(|X|-1)$

(6) algebra [1B]

(7) $H(p_e) \leq 1$ [1B]

b) The optimum estimator is given by [3E]

$$\hat{x} = \begin{cases} 1 & y = a \\ 2 & y = b \end{cases}$$

Then the error probability is given by [2E]

$$P_e = \frac{1}{12} + \frac{1}{12} + \frac{1}{12} + \frac{1}{12} = \frac{1}{3}$$

To apply Fano's inequality, we need to calculate conditional entropy [3E]

$$H(Y|X) = \frac{1}{2}H\left(\frac{1}{3},\frac{1}{12},\frac{1}{12}\right) + \frac{1}{2}H\left(\frac{1}{12},\frac{1}{3},\frac{1}{12}\right) = H\left(\frac{1}{3},\frac{1}{12},\frac{1}{12}\right) = 1.252$$

Then Fano's inequality gives [2E]

$$\frac{H(Y|X) - 1}{\log 2} = 0.252$$

which is smaller than 1/3.

c) If |X| is infinity, then Fano's inequality is vacuous, namely, it only says Pe >= 0. [2T]

To get a meaningful bound, one may modify step (5) to get [3T]

$$\overset{(4)}{=} H(e) + H(x \mid y, e=0)(1-p_e) + H(x \mid y, e=1)p_e$$

$$\overset{(5)}{\leq} H(p_e) + 0 \times (1-p_e) + H(x)p_e$$

$$\overset{(6)}{\Rightarrow} p_e \geq \frac{\left(H(x \mid y) - H(p_e)\right)}{H(x)} \overset{(7)}{\geq} \frac{\left(H(x \mid y) - 1\right)}{H(x)}$$

where we use conditioning reduces entropy: [1T]

$$H(x \mid y, e=1) \leq H(x)$$

Then, the bound is meaningful as long as $x$ has finite entropy. [1T]

(Solution is not unique; student will get credit as long as it makes sense.)

3.

a)

| | |
|---|---|
| (1) definition of Gaussian pdf | [1B] |
| (2) definition of differential entropy | [1B] |
| (3) algebra, definition of mean | [1B] |
| (4) tr(AB) = tr(BA) | [1B] |
| (5) since trace is linear, E and tr can exchange order | [1B] |
| (6) definition of covariance matrix K | [1B] |
| (7) tr(KK^-1) = tr(I) = n | [1B] |
| (8) algebra | [1B] |
| (9) identity \|cA\| = c^n\|A\| | [1B] |
| (10) same | [1B] |

b) Using the chain rule of mutual information, [3A]

$$I(X;Y,V) = I(X;V) + I(X;Y|V) = I(X;Y) + I(X;V|Y).$$

But $I(X;V) = 0$ due to independence, and $I(X;V|Y) \geq 0$ trivially, so [2A]

$$I(X;Y|V) = I(X;Y) + I(X;V|Y) \geq I(X;Y).$$ [3A]

c)

i) This is well known $C = \frac{1}{2}\log(1 + \frac{P}{N})$ [2B]

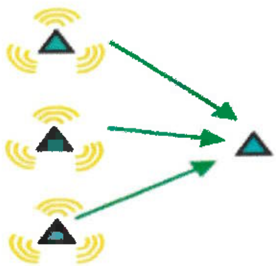
ii) This looks tricky but really is simple: you can just think of two noise terms, where one has variance N while the other has variance N1. [3A]

So

$$C = \frac{1}{2}\log(1 + \frac{P}{N+N_1})$$ [2A]

4.

a) i) Multi-access channel is a many-to-one channel where there are many senders but only one receiver. A typical example is the uplink of a cellular communication system. [3B]



[2B]

ii)

$$R_1 \le W \log\left(1+\frac{P_1}{N W}\right)$$

$$R_2 \le W \log\left(1+\frac{P_2}{N W}\right) \qquad [3B]$$

$$R_1 + R_2 \le W \log\left(1+\frac{P_1}{N W}\right) + W \log\left(1+\frac{P_2}{N W}\right)$$

The capacity region is a pentagon. To achieve corner points, one may use onion peeling. [2B]

iii) If W goes to infinity, then

$$R_1 \to \frac{P_1}{N}$$

$$R_2 \to \frac{P_2}{N} \qquad [3T]$$

$$R_1 + R_2 \to \frac{P_1}{N} + \frac{P_2}{N}$$

where the unit is nat (there is an extra factor log(e) if the unit is bit). Therefore, senders can transmit as if there were no interference. [2T]

b)

i) Let $Y = (Y_1, Y_2)$. Denote its covariance matrix by $K_Y$, which is given by [2A]

$$K_Y = \begin{bmatrix} \alpha^2 P + N & \alpha(1-\alpha)P \\ \alpha(1-\alpha)P & (1-\alpha)^2 P + N \end{bmatrix}$$

Its determinant [1A]

$$\det(K_Y) = \alpha^2 PN + (1-\alpha)^2 PN + N^2$$

Then, since $h(Y_1, Y_2)$ is maximized when X is Gaussian,

$$C = I(X;Y) = h(Y_1, Y_2) - h(Z_1, Z_2) = \frac{1}{2}\log\left(\frac{\det K_Y}{N^2}\right)$$
$$= \frac{1}{2}\log\left(1 + \frac{\alpha^2 P}{N} + \frac{(1-\alpha)^2 P}{N}\right)$$

[2A]

ii) The problem is equivalent to a degraded broadcast channel

$$y_1' = x + z_1/\alpha$$
$$y_2' = x + z_2/(1-\alpha)$$

The noise terms have variance $N_1 = N/\alpha^2$ and $N_2 = N/(1-\alpha)^2$, respectively. [2A]

Assume $N_1 < N_2$, i.e., $\alpha < \frac{1}{2}$, its capacity region is given by

$$R_1 \le C\left(\frac{\beta P}{N_1}\right)$$
$$R_2 \le C\left(\frac{(1-\beta)P}{\beta P + N_2}\right)$$

where $\beta \in [0,1]$. [2A]

If $\alpha > \frac{1}{2}$, the rates are reversed

$$R_1 \le C\left(\frac{(1-\beta)P}{\beta P + N_1}\right)$$
$$R_2 \le C\left(\frac{\beta P}{N_2}\right)$$

[1A]