CID: 01183830

Electrical and Electronic Engineering

Year 4

Module: EE 4-40 Information Theory (ELEC97048)

I confirm that the answers presented in the submitted pdf file are my own work and that I have not been in contact with others during the exam period.

01183830_EE 4-40.pdf

#1, a)

i) $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
$= H(X) + H(Y) - H(X,Y)$

Mutual Information is the average amount of information obtained of X, after knowing the value of y.

OR: How much the knowledge of y reduces the amount of uncertainty in X.

ii) Minimum : 0, $I(X;Y) \geq 0$

$I(X;Y) = I(Y;X) = E\left\{\log\left(\frac{P(X,Y)}{P(X)(PY)}\right)\right\} = D(p(x,y) \| p(x)p(y)) \geq 0$

Maximum : $H(X)$ (or $H(Y)$)

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

$[= H(Y) - H(Y|X) \leq H(Y)]$

· The knowledge of y can't bring more knowledge of x, than the uncertainty of x.

The maximum of $H(X)$ is when X follows a uniform distribution.

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

iii) $D(\underline{p} \| \underline{q}) = \frac{1}{6}\log\left(\frac{1}{6}/\frac{1}{10}\right) \times 5 + \frac{1}{6}\log\left(\frac{1}{6}/\frac{1}{2}\right) = 0.34998 \simeq 0.35$

$D(\underline{q} \| \underline{p}) = \frac{1}{10}\log\left(\frac{1}{10}/\frac{1}{6}\right) \times 5 + \frac{1}{2}\log\left(\frac{1}{2}/\frac{1}{6}\right) = 0.423998 \simeq 0.42$

iv) Relative Entropy is a measure of how different two probability mass vectors $\underline{p}$ and $\underline{q}$ are.

Relative Entropy is not a distance because:

① it is asymmetric, from iii, it is clear that $D(q\|p) \neq D(p\|q)$

② it will not satisfy triangle inequality.

v) $I(X,Y) = H(X) + H(Y) - H(X,Y) = E -\log(P(X)) + E -\log(P(Y)) + E\log(P(X,Y))$

$= E\left\{\log\left\{\frac{P(X,Y)}{P(X)P(Y)}\right\}\right\} = D(P(X,Y) \| P(X)P(Y))$

#1, b) Let $[\pi, 1-\pi]$ be the 2-state-stationary distribution

i)

$$[\pi, 1-\pi]\, T = [\pi, 1-\pi]$$

$$[\pi, 1-\pi] \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} = [\pi, 1-\pi]$$

$\pi p + (1-\pi)(1-p) = 1-\pi$

$\pi p + \cancel{1-\pi} - p + \pi p - \cancel{1-\pi}$

$p(2\pi - 1) = 0$

$$\pi \cdot (1-p) + p(1-\pi) = \pi$$

$$\pi - \pi p + p - \pi p = \pi$$

$$p(2\pi - 1) = 0 \quad , \quad \pi = \tfrac{1}{2}.$$

$[\frac{1}{2}, \frac{1}{2}]$ is the stationary distribution

This makes sense as state 0 has p probability to transit to state 1, and vice versa. $[\frac{1}{2}, \frac{1}{2}]$ stationary distribution is the obvious equilibrium.

Entropy Rate: $H(x) = \lim_{n \to \infty} H(X_n | X_{n-1})$

$$= \pi \cdot H(p) + (1-\pi) H(1-p)$$

$$= \tfrac{1}{2} \cdot H(p) + \tfrac{1}{2} H(p) = H(p)$$

ii) given there is only two state.

$$H(x) = H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$$

$p = \frac{1}{2}$ maximizes $H(p)$ so maximizes $H(x)$

$$H'(p) = \log(1-p) - \log p = 0$$

$$\log\left(\frac{1-p}{p}\right) = 0, \quad \frac{1-p}{p} = 1, \quad p = \tfrac{1}{2}.$$

\#2, a) 1, 1 0, 0, 0 0, 1 1, 0 1, 0 1 0, 0 0 0, 0 1 1, 0 1 0 1,

Dictionary		Encoding
0000	ϕ	
0001	1	1
0010	10	00010
0011	0	00000
0100	00	00110
0101	11	00011
0110	01	00111
0111	010	01100
1000	000	01000
1001	011	01101
1010	0101	01111

b). $P(X=0) = 0.4 + 0.1 = 0.5 = P(Y=1) = P(Y=0) = P(Y=1) = \frac{1}{2}$.

i) $H(X) = H(\frac{1}{2}) = 1$ assume $N_\varepsilon < \infty$

$(1-\varepsilon)\, 2^{nH(x) - n\varepsilon} < |T_\varepsilon^{(n)}| \le 2^{nH(x) + n\varepsilon}$, for $n > N_\varepsilon$

$(1-0.1) \cdot 2^{10 \cdot 1 - 10 \cdot 0.1} < |T_\varepsilon^{(n)}| \le 2^{10 \cdot 1 + 10 \times 0.1}$

$460.8 < |T_\varepsilon^{(n)}| \le 2048$

$|-\frac{1}{10}\log_2 P_X(\underline{X}) - 1| < 0.1$

$0.9 < -\frac{1}{10}\log_2 P_X(\underline{X}) < 1.1$

since $P(X=0) = P(X=1) = \frac{1}{2}$. (assume X_i and X_j in sequence are i.i.d)

$P_X(\underline{X}) = \prod^{10} \frac{1}{2} = (\frac{1}{2})^{10}$

$-\frac{1}{10}\log_2[(\frac{1}{2})^{10}] = (-1)(-1)(\frac{1}{10}) \times 10 \log_2(2) = 1$, always.

$-n^{-1}\log_2 P_X(\underline{X}) - H(X) = 0$ for all sequences possible.

$2^{10} = 1024$, $\Rightarrow$ 1024 of them are in the typical set.
(or all of them)

ii) observing $P_{XY}(x,y)$ table, x and y are symmetric,
so typical set $A_\varepsilon^{(n)}(\underline{Y})$ is of same nature as $A_\varepsilon^{(n)}(\underline{X})$
All of the 1024 possible sequence of $\underline{Y}$ are in the typical set.

#2 b) iii) $H(X,Y) = -0.4\log(0.4)\times 2 - 0.1\log(0.1)\times 2 = 1.7219$

$H(X,Y) - \varepsilon < -\frac{1}{10}\log_2 P_{XY}(\underline{X},\underline{Y}) < H(X,Y) + \varepsilon$ $\qquad \varepsilon = 0.1$

$10\,H(X,Y) = -0.4\log(0.4^{10})\times 2 - 0.1\log(0.1^{10})\times 2.$

$\cancel{|J_\varepsilon^{(n)}| = 2^{-nI(X;Y)}}$

$1.6219 < -\frac{1}{10}\log_2 P_{XY}(\underline{X},\underline{Y}) < 1.8219$

$16.219 < -\log_2 P_{XY}(\underline{X},\underline{Y}) < 18.219.$

$-16.219 > \log_2 P_{XY}(\underline{X},\underline{Y}) > 18.219$

$1.311\times 10^{-5} > P_{XY}(\underline{X},\underline{Y}) > 3.277\times 10^{-6}.$

$P(x,y)$ takes either 0.4 or 0.1

let m be the number of paires of (x,y) in sequence, for which $P(x,y) = 0.4$.

$P_{XY}(\underline{x},\underline{y}) = \binom{10}{m} 0.4^m \cdot 0.1^{10-m}.$

$10C5 \times 0.4^5 \times 0.1^5 = 2.58\times 10^{-5}. \quad > 1.311\times 10^{-5}$

$10C4 \times 0.4^4 \times 0.1^6 = 5.376\times 10^{-6}$ ✓

$10C3 \times 0.4^3 \times 0.1^7 = 7.68\times 10^{-7} < 3.277\times 10^{-6}.$

so $m = 4$: in the joint typical set.

$P_{XY}(\underline{X},\underline{Y}) = \binom{10}{4} 0.4^4 \cdot 0.1^6. \quad = 5.376\times 10^{-6}$ (probability in the joint typical set)

number of $(x=0, y=0)$ or $(x=1, y=1)$ is 4.

number of x and y pairs different is 6.

~~size = $\binom{10}{4}\cdot 2^4 \cdot \binom{10}{6} 2^6 = \binom{10}{4}\binom{10}{6}\cdot 2^{10} = 45158400 > 2^{20}$~~ X

size $(1-\varepsilon)\, 2^{nH(X,Y) - n\varepsilon} < |J_\varepsilon^{(n)}| \le 2^{nH(X,Y) + n\varepsilon}$ $\qquad n=0$

$0.9 \cdot 2^{10\cdot 1.7219 - 1} < |J_\varepsilon^{(n)}| \le 2^{10\times 1.7219 + 1}$ $\qquad \varepsilon = 0.1$

$68651.2 < |J_\varepsilon^{(n)}| \le 274604.7.$

$\left.\begin{array}{l} x = y = 0, \text{ or } x = y = 1 \\ \quad + \\ 0 = x \ne y = 1,\ 1 = x \ne y = 0 \end{array}\right\}$ symmetric sequence.

size $= \binom{10}{6}\cdot\binom{10}{4}\times 2 = 88200$

01183830

#3. a) (1) definition of mutual information

(2) expand conditional mutual information $H(X|Y)$ as weighted average row entropy.

$$H(X|Y) = H(X|Y=0)\cdot P(Y=0) + H(X|Y=?)\cdot P(Y=?) + H(X|Y=1)\cdot P(Y=1)$$

due to the nature of BEC, if $y=0$ or 1, x is known for certain:

$$H(X|Y=0) = H(X|Y=1) = 0.$$

(3) from Fig 3.1

$$p(y=?) = P(Y=?|X=0)\cdot P(X=0) + P(Y=?|X=1)P(X=1)$$
$$= f\cdot p + f(1-p) = f \qquad p: Pr(X=0)$$

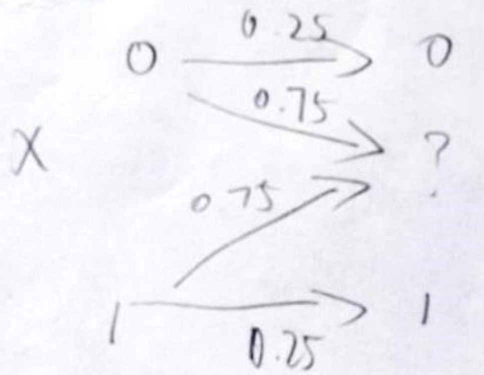
use algebra to simplify (2)

(4) $H(X) = H(p) \leq 1$, $H(X)$ is bounded by maximum $H(\frac{1}{2}) = 1$ (X is binary)

(5) Capacity: $C = \max_{x} I(x;y)$

and $I(x;y) \leq 1-f$

$\therefore C = 1-f$



b) W^- is a BEC

$W^- \sim BEC(2p-p^2)$

W^+ is a $BEC(p^2)$

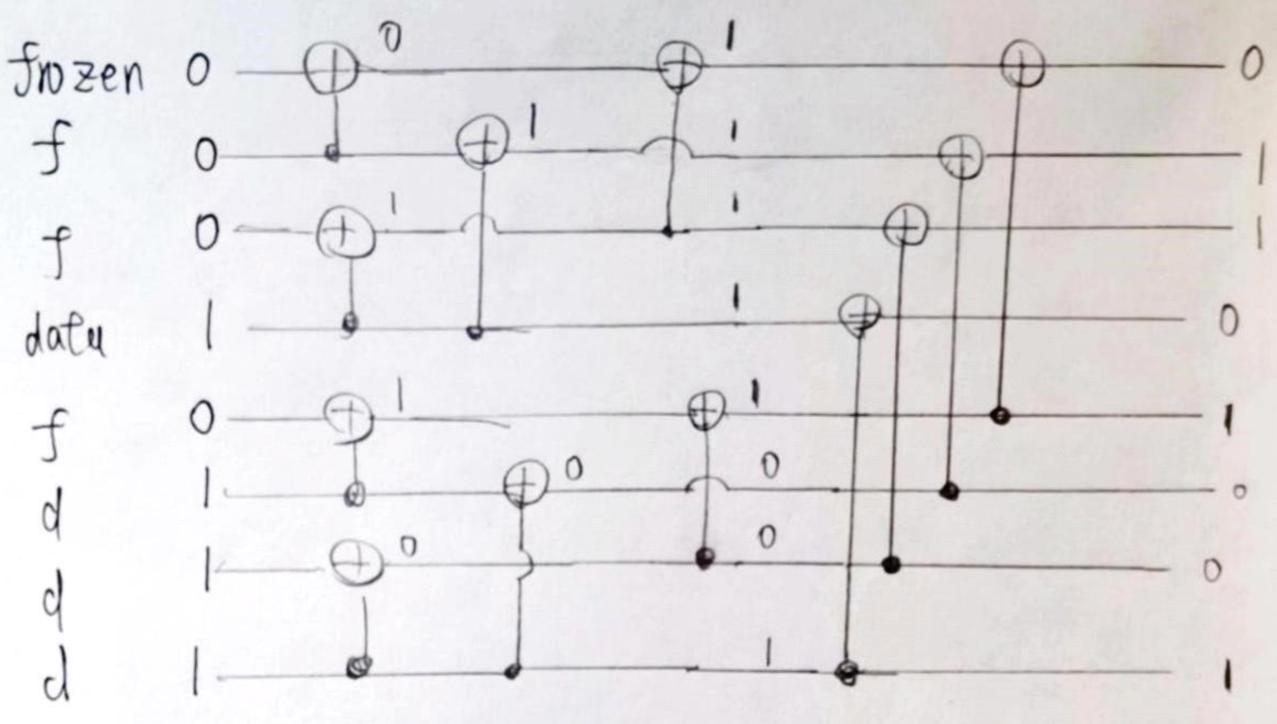
capacity of BEC: $C(W) = 1-p$ from (3 a)

$$C(W^-) = 1-(2p-p^2) = 1-(2\cdot\tfrac{3}{4}-(\tfrac{3}{4})^2) = \tfrac{1}{16}$$

$$C(W^+) = 1-p^2 = 1-(\tfrac{3}{4})^2 = \tfrac{7}{16}$$

$$C(0.75) = 2\cdot(1-0.75) = 0.5 = C(W^-)+C(W^+)$$

#3. (C)



output codeword is 01101001

b) more explanation.

W^{-} : input: U_1

output: $\underbrace{(U_1+U_2, U_2), (?, U_2), (U_1+U_2, ?), (?, ?)}$ erasure

$Pr(\text{erasure}) = 1 - Pr(\text{not erasure})$

$= 1 - (1-p)^2 = 2p - p^2.$

W^{+} : input: U_2.

output (U_1+U_2, U_2, U_1)

$(?, U_2, U_1)$

$(U_1+U_2, ?, U_1)$

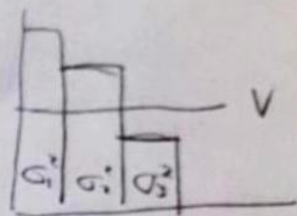
$(?, ?, U_1)$

$P_r(\text{erasure}) = p^2.$

Capacity of BEC: ~~BEC is~~

shown in 3, a)

#4. a) i) single channel.



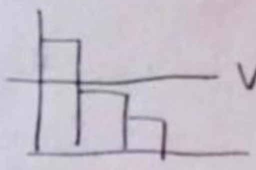
$V = 3P + \sigma_3^2$, $V - \sigma_3^2 = 3P$

$\sigma_3^2 + 3P \leq \sigma_2^2$

$3P \leq \sigma_2^2 - \sigma_3^2$, $P < \frac{1}{3}(\sigma_2^2 - \sigma_3^2)$

$C = \frac{1}{2}\log(1+\frac{P_3}{\sigma_3^2}) = \frac{1}{2}\log(1+\frac{3P}{\sigma_3^2})$

ii.) a pair of channels



$\sigma_1^2 > V > \sigma_2^2$.

$V - \sigma_2^2 + V - \sigma_3^2 = 3P$

$2V = 3P + \sigma_2^2 + \sigma_3^2$, $V = \frac{3}{2}P + \frac{1}{2}\sigma_2^2 + \frac{1}{2}\sigma_3^2$

$\frac{1}{3}(\sigma_2^2 - \sigma_3^2) < P < \frac{1}{3}(2\sigma_1^2 - \sigma_2^2 - \sigma_3^2)$

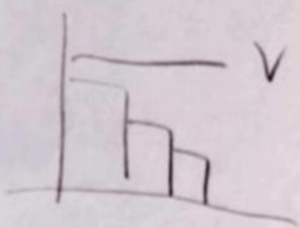
$P_2 = V - \sigma_2^2 = \frac{3P - \sigma_2^2 + \sigma_3^2}{2}$

$P_3 = V - \sigma_3^2 = \frac{1}{2}(3P + \sigma_2^2 - \sigma_3^2)$

$C = \frac{1}{2}\log(1+\frac{P_2}{\sigma_2^2}) + \frac{1}{2}\log(1+\frac{P_3}{\sigma_3^2})$

$= \frac{1}{2}\log(1+\frac{3P - \sigma_2^2 + \sigma_3^2}{2\sigma_2^2}) + \frac{1}{2}\log(1+\frac{3P+\sigma_2^2-\sigma_3^2}{2\sigma_3^2})$

iii) 3 channels



$V \geq \sigma_1^2$

$3V - \sigma_1^2 - \sigma_2^2 - \sigma_3^2 = 3P$.

$V = P + \frac{1}{3}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) \geq \sigma_1^2$

$P \geq \frac{2}{3}\sigma_1^2 - \frac{1}{3}\sigma_2^2 - \frac{1}{3}\sigma_3^2$

$P_1 = V - \sigma_1^2 = P - \frac{2}{3}\sigma_1^2 + \frac{1}{3}\sigma_2^2 + \frac{1}{3}\sigma_3^2$

$P_2 = P + \frac{1}{3}\sigma_1^2 - \frac{2}{3}\sigma_2^2 + \frac{1}{3}\sigma_3^2$

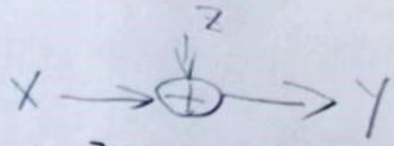
$P_3 = P + \frac{1}{3}\sigma_1^2 + \frac{1}{3}\sigma_2^2 - \frac{2}{3}\sigma_3^2$

$C = \frac{1}{2}\log(1+\frac{P_1}{\sigma_1^2}) + \frac{1}{2}\log(1+\frac{P_2}{\sigma_2^2}) + \frac{1}{2}\log(1+\frac{P_3}{\sigma_3^2})$

#4, b)

$X \in \{0, 1, -1\}$

$Y = X + Z, \quad Z \sim U[-1, 1]$



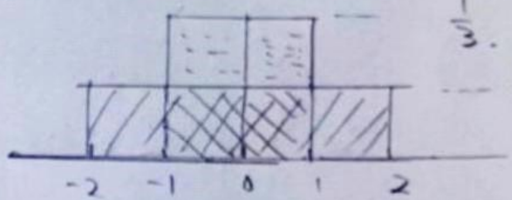
$X=0\ [-1, 1]$
$X=-1\ [-2, 0]$
$X=1\ [0, 2]$

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

given $x=0$, y will be uniformly distributed in $[-1, 1]$

similar for give $x=-1$, and 1 so.

Probability of Y: (pdf)



$\frac{1}{3}$.

$\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$.

$\frac{1}{6} \times 4 + \frac{1}{3} \times 2 = 1$.

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

X & Z are independent: $H(Z|X) = \cancel{H(Z)} = h(Z)$

$h(Z) = \log(b-a) = \log(1+1) = 1 = H(Y|X)$

$$h(y) = -\int_{-\infty}^{\infty} f_Y(y) \log f_Y(y)\, dy$$

$$= -\int_{-2}^{-1} \frac{1}{6} \log\left(\frac{1}{6}\right) dy - \int_{-1}^{1} \frac{1}{3} \log\left(\frac{1}{3}\right) dy - \int_{1}^{2} \frac{1}{6} \log\left(\frac{1}{6}\right) dy$$

$$= -\left[\frac{1}{6} \log\left(\frac{1}{6}\right) y\right]_{-2}^{-1} - \left[\frac{1}{3} \log\left(\frac{1}{3}\right) y\right]_{-1}^{1} - \left[\frac{1}{6} \log\left(\frac{1}{6}\right) y\right]_{1}^{2}$$

$$= -\frac{1}{6} \log\left(\frac{1}{6}\right) - \frac{2}{3} \log\left(\frac{1}{3}\right) - \frac{1}{6} \log\left(\frac{1}{6}\right) = 1.918$$

$C = I(X;Y) = h(y) - h(Y|X) = 1.918 - 1 = 0.918$ (bits per channel use)