Information Theory & Coding (ITC)

Source coding | channel

Information Theory:

- Branch of applied mathematics and and developed by claude.

E. Shanon to find the fundamental limit on signal processing operation such as compressing data and Reliable stary and communicating data.

Applications:

- cryptography : widely used to hide information and applications include, cash ~~mach~~ machines, computer passwords, etc.
- Natural language processing
- Network.

Information: Gives some message.

The key measure of information is entropy.

- Average no. of bits needed for storage or communication.

Self Information:

Information about the individual event.

Self information can be $I_K = \frac{1}{\log_2 P_K}$ bits.

P_K = probability of events

~~$I_K = \log$~~ $\boxed{I_K = -\log_2 P_K}$

Base of $\log_2 \rightarrow$ unit is bits.
" " $\log_{10} \rightarrow$ " " Hartley.

Natural logarithm $\rightarrow$ unit is nat.

$$\boxed{\text{Information} \propto \frac{1}{\text{probability}}}$$

Properties of self information

① $I(S_K) = 0$ if $P_K = 1$.
where S_K is an event.

In the case, if the outcome of the event certain even before it occurs, there is no information gained.

② $I(S_K) \geq 0$ if $0 \leq P_K \leq 1$

The occurrence of the events provide some or no information but never brings about loss of information.

③ $I(S_K) > I(S_i)$ for $P_K < P_i$

The less probable events gives more information.

21/06/2021.

ENTROPY: (measure of information)

It is the mean of self information.

Mathematically,

$$H = -\sum_{K=0}^{K-1} P_K \log_2 P_K$$

$$\Big[\; E[I(S_K)]$$

$$\Rightarrow \sum_{K=0}^{K-1} P_K \, I(S_K)$$

$$\sum_{K=0}^{K-1} P_K \log_2 \left(\frac{1}{P_K} \right) \Big] \quad \text{— (1)}$$

$\frac{1}{p}$

Source entropy

$$\underset{\text{(Source)}}{X} \longrightarrow \boxed{C} \longrightarrow y \; \text{(Receiver)}$$

$$H(X) = -\sum_{i=1}^{M} P(x_i) \log_2 P(x_i)$$

Destination entropy

$$H(Y) = -\sum_{j=1}^{N} P(y_j) \log_2 P(y_j)$$

ation.

Joint entropy.

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log_2 P(x_i, y_j)$$

Conditional Entropy.

$H(Y/X)$

- Measure of information (entropy) about the same receiving port, that it is known that X- transmitter. It gives an indication of noise of either in channel

Equivocation: $H(X/Y)$

- Measure of information about source where it is known the Y- receiver.
- It gives how well one can recover the i/p content from output.

V.V.Imp. properties of entropy

consider discrete continuous source whose mathematical modal is defined by

$$\mathcal{Y} = \{s_0, s_1, \ldots s_{K-1}\}$$

with prob

$$P(S = S_K) = P_K, \quad K = 0, 1, \ldots K-1$$

The entropy of the source is bounded as follows :

~~H~~ $H(\mathcal{X})$.

$$\boxed{0 \leq H(\mathcal{X}) \leq \log_2 K}$$

$\log_2 K$ $\log p_k$

$(\log(1)) \to 0$ $p_k = 1$

$K \to$ No. of symbols. (data).

$\boxed{\text{Case - I}}$

Entropy $\boxed{H(\mathcal{X}) = 0}$

If and only $p_k = 1$ & remaining probability in the set are zero (0).

- This lower bound in entropy corresponds to no uncertainity

$\boxed{\text{Case - II}}$

Entropy $\boxed{H(\mathcal{X}) = \log_2 K}$ tossing a coin $\log_2 2 = 1$

If and only if $p_K = \frac{1}{K}$ for all K.

(All the symbols in the alphabet are ~~equi~~ equiprobable)

- Upper bound on entropy corresponds to maximum uncertainity.

Proof:

~~probality~~ prob $p_k \le 1$ $H(\mathcal{S}) \rightarrow$ Non-negative

$\Rightarrow H(\mathcal{S}) \geq 0$

<u>proof for upper bound:</u>

$[\ln x \le x-1, \ x \geq 0]$

Two probability distributions

$\{p_0, p_1, \ldots\ldots p_{K-1}\}$ and

$\{q_0, q_1, \ldots q_{K-1}\}$ on the

alphabet $\mathcal{S} = \{s_0, s_1, \ldots s_{K-1}\}$

of a discrete memoryless source.

$$\sum_{k=0}^{K-1} p_k \log_2 \frac{q_k}{p_k} = \frac{1}{\ln 2} \sum_{k=0}^{K-1} p_k \log_{10} \frac{q_k}{p_k} \quad \left[\because \log_2 x = \frac{\log_{10} x}{\log 2}\right]$$

$$\sum_{k=0}^{K-1} p_k \log_2 \frac{q_k}{p_k} \le \frac{1}{\ln 2} \sum_{k=0}^{K-1} p_k \left(\frac{q_k}{p_k} - 1\right)$$

$$\le \frac{1}{\ln 2} \sum_{k=0}^{K-1} q_k - p_k$$

$$\le \frac{1}{\ln 2}\left(\sum_{k=0}^{K-1} q_k - \sum_{k=0}^{K-1} p_k\right) = 0.$$

$$\therefore \sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) \le 0 \quad \text{If } p_k = q_k \ \forall k$$

$q_k = \frac{1}{K}$

$K = \frac{1}{q_k}$

If $q_k = \frac{1}{K} \rightarrow$ equiprobable

$$H(\mathcal{S}) = \sum_{k=0}^{K-1} q_k \log_2 \frac{1}{q_k} = \log_2 K$$

$$\sum_{k=0}^{K-1} p_k \log \frac{1}{p_k} \le \log_2 K$$

$$\boxed{H(y) \le \log_2 K}$$

22/08/11.

Relationship between joint and Conditional entropy

Joint E, $H(X,Y)$

CE, $H(X/y)$ or $H(Y/X)$

Case - I :

If x and y are statically independent

$P(x_i, y_j) = p(x_i)\, p(y_j)$, $\forall\ i, j$

$$H(X,Y) = -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log p(x_i, y_j)$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \cdot \log \left[p(x_i) \cdot p(y_j) \right]$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{N} \log p(x_i)\, P(x_i, y_j) + $$

$$- \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log P(y_j)$$

$$= -\sum_{i=1}^{M} \log P(x_i) \sum_{j=1}^{N} P(x_i, y_j)$$

$$- \sum_{j=1}^{N} \log P(y_j) \sum_{j=1}^{N} p(x_i, y_j)$$

$$= -\sum_{j=1}^{M} p(x_i) \log p(x_i) - \sum_{j=1}^{N} p(y_j) \log p(y_j)$$

$$\therefore \boxed{H(X,Y) = H(X) + H(Y)}$$

$$\left(\because \sum_{j=1}^{N} p(x_i, y_j) = p(x_i) \right)$$

al

<u>Case-II</u>: If x and y are dependent then,

$$p(x_i, y_j) = p(x_i)\, p(y_j/x_i) \quad (\text{or})$$

$$= p(y_j) \cdot p(x_i/y_j)$$

$$H(X,Y) = -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log \left[p(x_i) p(y_j/x_i) \right]$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log p(x_i) - \sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log p(y_j/x_i)$$

$$= -\sum_{i=1}^{M} \log p(x_i) \sum_{j=1}^{N} p(x_i, y_j) - \sum_{j=1}^{N} \log p(y_j/x_i) \sum_{j=1}^{M} p(x_i, y_j)$$

$$\boxed{H(X,Y) = H(X) + H(Y/X)}$$

$\text{''}ly$,

$$H(X,Y) = -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log \left[p(y_j) \cdot p(x_i/y_j) \right]$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log \left[p(y_j) + p(x_i/y_j) \right]$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log p(y_j) - \sum_{i=1}^{M} \sum_{j=1}^{N} p(x_i, y_j) \log p(x_i/y_j)$$

$$= -\sum_{j=1}^{N} \log p(y_j) \sum_{i=1}^{M} p(x_i, y_j) - \sum_{i=1}^{M} \log p(x_i/y_j) \sum_{j=1}^{N} p(x_i, y_j)$$

$$= -\sum_{j=1}^{N} p(y_j) \log p(y_j) - \sum_{i=1}^{M} p(x_i) \log p(x_i/y_j)$$

$$= \cancel{H(Y) + H(X)} \quad \boxed{H(Y) + H(X/Y)}$$

→ The conditional entropy is just defined as each satisfying an important inequality.

$$0 \leq H(Y/X) \leq H(Y)$$

$$0 \leq H(X/Y) \leq H(X)$$

Usual Information

- It measures the amount of information that can be obtained about one random variable by observing another.

- It is important in communication where it can be used to maximize the amount of information shared between sent and received signal.

$$\boxed{I(x_i, y_i) = \log\left[\frac{P(x_i/y_i)}{P(x_i)}\right]} \quad -① $$

mutual information.

Properties:

i) If ~~x and y~~ X and Y are independent,

$$P(x_i/y_j) = P(x_i) \quad \text{--- (2)}$$

$\therefore$ from (1) & (2)

i) $\boxed{I(x_i, y_j) = 0}$

ii) when the occurrence of the event $Y = y_j$, unity determines the occurrence of $X = x_i$, the conditional probability is unity.

$$I(x_i, y_j) = \log\left(\frac{1}{P(x_i)}\right)$$

$$\boxed{\therefore I(x_i) = -\log P(x_i)}$$

Average Mutual Information:

$$I(X, Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j)\, I(x_i, y_j)$$

contd..

23/06/11

Average Mutual Information

properties.

i) when X and Y are statistically independent,

$I(X,Y) = 0$

$\therefore \boxed{I(X,Y) \geq 0}$

ie. Mutual information is always non-negative.

ii) symmetric properties.

$$I(X,Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j)\, I(x_i, y_j)$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log \left[\frac{P(x_i/y_j)}{P(x_i)} \right]$$

$[\because P(x_i, y_j) = P(x_i/y_j) \cdot P(y_j)$

or

$= P(y_j/x_i) \cdot P(x_i)$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(x_i) \cdot P(y_j)}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log \frac{P(y_j/x_i) \cdot \cancel{P(x_i)}}{\cancel{P(x_i)} \cdot P(y_j)}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log \frac{P(y_j/x_i)}{P(y_j)}$$

$$= I(Y,X)$$

$\therefore \boxed{I(X,Y) = I(Y,X)}$

//

iii)

$$I(X,Y) = \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log\left[\frac{P(x_i/y_j)}{P(x_i)}\right]$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log\left[\frac{P(x_i, y_j)}{P(x_i)\cdot P(y_j)}\right]$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j)\left[\log P(x_i, y_j) - \log(x_i) - \log(y_j)\right]$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j)\log P(x_i, y_j) - \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j)\log(x_i) - \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j)\log P(y_j)$$

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

$$\therefore \boxed{I(X,Y) = H(X) + H(Y) - H(X,Y)} \quad - (1)$$

iii)
iv) $I(X,Y) \Rightarrow H(Y) - H(Y/X) \geq 0$

$H(X) - H(X/Y) \geq 0$

Proof:

$$\left[\because H(X,Y) = H(X) + H(Y)\right.$$
$$H(X,Y) = H(Y) + H(X/Y) \text{ or}$$
$$\left. H(X) + H(Y/X)\right] - (2)$$

from (1) & (2)

$$I(X,Y) = H(X) + H(\cancel{Y}) - H(\cancel{Y}) - H(X/Y)$$

$$\boxed{I(X,Y) = H(X) - H(X/Y)}$$

$\|^{ly}$. $I(X,Y) = H(\cancel{X}) + H(Y) - H(\cancel{X}) - H(Y/X)$

$$\boxed{I(X,Y) = H(Y) - H(Y/X)}$$

Joint entropy of two ensembles X and Y are maximum when the ensembles are independent.

Q: Discuss about entropy of binary memoryless source.

Consider a binary source which has two symbols

0 → prob P_0

1 → prob $P_1 = 1 - P_0$

∵ source is memoryless, successive symbols emitted by the source are statistically independent.

$$H(X) = -\sum_{k=0}^{K-1} P_k \log_2 P_k$$

$$= -\sum_{k=0}^{1} \cancel{P_0 \log_2 P_0} \; P_k \log P_k$$

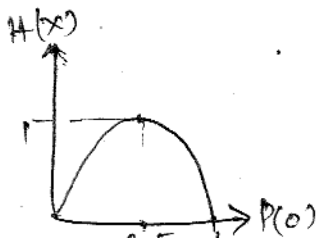
$$= -P_0 \log P_0 - P_1 \log P_1$$

$$H(X) = -P_0 \log P_0 - (1-P_0)\log(1-P_0)$$

when $\boxed{\begin{array}{l} P_0 = 0, \; H(X) = 0 \\ P_1 = 1, \; H(X) = 0 \end{array}}$

when, $\boxed{P_0 = 1/2 = P_1 \text{ then, } H(X) = \text{Maximum.}}$

symbols 1 and 0 are equiprobable.



$H(X) = \text{Max}^m$
$= 1$ bit

3/6/01

Unique decipherability & Instantaneous Codes

No code word should be prefix for the another code word.

The aim of noiseless coding is to produce codes with the following two properties;

i) unique decipherability

ii) Minimum average length for a given source.

⇒ Codes which have both the above properties are said to be optimal codes.

unique but not in sequence.

x_1	0
x_2	1
x_3	00
x_4	11

00111 ...

$x_3 x_4 x_2$ ...

$x_3 x_2 x_4$ ...

$x_1 x_1 x_2 x_2 x_3$ ...

$x_1 x_1 x_4 x_2$ ....

ANT

OD.

Eg.

such code is not uniquely decipherable.

A	B	C	D
0	0	0	0
1	10	01	10
00	110	011	110
11	111	0111	1110

B and D are instantaneous codes.

C and D are comma codes.

i.e. any sequence of code words can be decoder by sub-~~diving~~ dividing 0's & 1's to the left of every '0' for both 'C' & 'D' and to rewrite every zero to code D.

The character '0' is the 1st or last character of every code ~~word~~ as Comma.

Q. Find ~~the~~ ~~cont~~ which is not uniquely decipherable.

- A is not uniquely decipherable. Instantaneous is sufficient but not necessary Condition for code to be uniquely decipherable.

Q.1) Determine wheather Codes are uniquely decipherable.

S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
	d	ba	ce	ac	c	eac	ac	c	eac	ac
abc	abd			ab	cd	eab	ab	cd	eab	ab
abcd							ϕ		bace	d
e										
dba										
bace										
ceac										
ceab										
eabd										

$S_7 = S_{10}$

Non of the sets S_0' ~~are~~ Contains Code word

So code is uniquely decipherable.

Q.2)

S_0

010
0001
0110
1100
00011
00110
11110
10101

Q.2)

S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
010	1	100	11	00	01	0	10	1011	1011
0001		1110	110		011	10	001	10	10
0110		01011			110	001	110	0	001
1100					0	110	0011		110
00011						0011	0110		0011
00110						0110			0110
11110									
101011									

$$\boxed{S_8 = S_{10}}$$

eous on)

S_8 S_9 ... each ... each ... bace

... of the sets S_0 contains

20/06/11.

Q. Consider a discrete memoryless source with source alphabet $S = S_0, S_1, S_2$ with probability $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$ respectively. find its entropy.

Soln:

$$P(S_0) = 1/4$$
$$P(S_1) = 1/4$$
$$P(S_2) = 1/2$$

$$H = -\sum_{k=0}^{2} P_k \log_2 P_k$$

$$= -\left[P_0 \log_2 P_0 + P_1 \log_2 P_1 + P_2 \log_2 P_2\right]$$

$$= -\left[\tfrac{1}{4}\log_2 \tfrac{1}{4} + \tfrac{1}{4}\log_2 \tfrac{1}{4} + \tfrac{1}{2}\log_2 \tfrac{1}{2}\right]$$

$$= -\left[\frac{1}{4}\ \frac{\log 1/4}{\log 2} + \frac{1/4 \log_{10} 1/4}{\log_{10} 2} + \frac{1/2 \log_2 1/2}{\log 2}\right]$$

$= 3/2$ bits.

Q: In binary PCM, if zero occurs probability 1/4 and 1 occurs with probability 3/4 then calculate the amount of information conveyed by each unit.

Soln:

Binary

0	1
$\frac{1}{4}$	$3/4$

$I_K = -\log P_K$

$I_0 = -\log 1/4 = \underline{\underline{2 \text{ bits}}}$

$I_1 = -\log 3/4 = -\dfrac{\log 3/4}{\log_{10} 2} = \underline{\underline{0.4 \text{ bits}}}$.

Q: The joint probability of matrix of a channel with binary input is given as $P(x,y) = \begin{pmatrix} 1/4 & 1/4 \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}$ and

marginal probability is given as

$P(x_1) = P(x_2) = 1/2$

$P(y_1) = P(y_2) = 1/2$

Find its average mutual information.

Soln:

$$p(x,y) = \begin{array}{c|cc} & y_1 & y_2 \\ \hline x_1 & 1/4 & 1/4 \\ x_2 & 1/4 & 1/4 \end{array}$$

only from joint probability, we can find marginal probability

$p(x_1) = 1/4 + \frac{1}{4} = 1/2$

$p(x_2) = 1/2$

$p(y_1) = 1/2$

$p(y_2) = 1/2$

$I(X,Y) = H(X) + H(Y) - H(X,Y)$

$H(X) = -(1/2 \log 1/2 + 1/2 \log 1/2)$

$H(Y) = -(1/2 \log 1/2 + 1/2 \log 1/2)$

$H(X,Y) = -[1/4 \log 1/4 + 1/4 \log 1/4 + \frac{1}{4} \log 1/4 + \frac{1}{4} \log \frac{1}{4}]$

$\therefore I(X,Y) = -[1/2 \log 1/2 + 1/2 \log 1/2 + 1/2 \log 1/2 + 1/2 \log 1/2]$

$- [1/4 \log 1/4 + \frac{1}{4} \log 1/4 + \frac{1}{4} \log 1/4 + \frac{1}{4} \log 1/4]$

$= -2 \log 1/2 - \log 1/4$

$= 2 - 2$

$= 0$

25/06/11.

Given the noise matrix of a channel ~~R-yix~~ $P(Y/X)$. Find $I(X,Y)$

$$P(Y/X) = \begin{array}{c c} & Y \\ X & \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Marginal probf.

$$P(X) = \left[\frac{1}{4}, 2/5, 3/20, \frac{3}{20}, \frac{1}{20}\right]$$

Steps:

$\Rightarrow P(X,Y) = P(X) \cdot P(Y/X)$

$P(Y) =$ Adding the columns of $P(X,Y)$.

$$H(Y) = -\sum_{j=0}^{K-1} P(y_j) \log P(y_j)$$

$$H(X) = -\sum_{i=0}^{K-1} P(x_i) \log P(x_i)$$

$$H(X,Y) = -\sum_{i=0}^{K-1} \sum_{j=0}^{K-1} P(x_i, y_j) \log P(x_i, y_j)$$

$$P(X/Y) = \frac{P(X,Y)}{P(Y)}$$

$$H(X/Y) = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} P(x_i, y_j) \log P(x_i/y_j)$$

$$H(Y/X) = \sum \sum P(x_i, y_j) \log P(y_j/x_i)$$

$I(X,Y) = H(Y) - H(Y/X)$ (or)

$= H(X) - H(X/Y)$

$$P(X,Y) = \begin{array}{c} \begin{array}{cccc} y_0 & y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{10} & 3/10 & 0 & 0 \\ 0 & 1/20 & 1/10 & 0 \\ 0 & 0 & 1/20 & 1/10 \\ 0 & 0 & 1/20 & 0 \end{bmatrix} \end{array}$$

10 ~~4, 10, 10, 20, 10, 20, 20, 10~~

2 ~~4, 1, 1, 2, 1, 2, 2, 1~~

2, 1, 1, 1, 1, 1, 1..

$P\cancel{(Y) = P(y_0) + P(y_1) + P(y_2) + P(y_3)}$

$\cancel{= \left(\frac{1}{4} + \frac{1}{10}\right) + \left(\frac{3}{10} + \frac{1}{20}\right) + \left(\frac{1}{10} + \frac{1}{20} + \frac{1}{20}\right) + \frac{1}{10}}$

$= \cancel{40}$

$P(Y) \Rightarrow$

$P(y_0) = \frac{1}{4} + \frac{1}{10} = 7/20$

$P(y_1) = \frac{3}{10} + \frac{1}{20} = 7/20$

$P(y_2) = 1/10 + \frac{1}{20} + \frac{1}{20} = 1/5$

$P(y_3) = 1/10$

$$P(X/Y) = \begin{bmatrix} 5/7 & 0 & 0 & 0 \\ \frac{2}{7} & 6/7 & 0 & 0 \\ 0 & 1/7 & 1/2 & 0 \\ 0 & 0 & 1/4 & 1 \\ 0 & 0 & 1/4 & 0 \end{bmatrix}$$

$$H(X) \Rightarrow \frac{-1}{\log_{10} 2}\left[\frac{1}{4}\log_{10} 1/4 + 2/5 \log_{10} 2/5 + 3/20 \log_{10} 3/20 + 3/20 \log 3/20 + 1/20 \log_{10} 1/20\right]$$

$$= \cancel{-0.3 \left[-0.4 - 1 + \right.} -$$

$$= \cancel{\frac{1}{-0.3}\left(-1.505 - 1.59 - 1.23 - 1.23 - 0.065\right)}$$

$$= -\frac{1}{0.3}\left[-0.15 - 0.159 - 0.123 - 0.123 - 0.065\right]$$

$$= \frac{0.611}{0.3} = 2.0367$$

$$\therefore \boxed{H(X) = 2.0367}$$

$$H(Y) = -\frac{1}{\log_{10} 2}\left[\frac{7}{20}\log_{10} 7/20 + \frac{7}{20}\log_{10} 7/20 + \frac{3}{20}\log_{10} 3/20 + \frac{1}{10}\log_{10} 1/10\right]$$

$$= \frac{-1}{0.3}\left[-0.159 - 0.159 - 0.123 - 0.1\right]$$

$$= \frac{0.541}{0.3} = 1.8$$

$$\therefore \boxed{H(Y) = 1.8}$$

$H(X, Y) = 2.665$

For $H(X, Y)$,

$$p(X, Y) = \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{10} & \frac{3}{10} & 0 & 0 \\ 0 & \frac{1}{20} & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{20} & \frac{1}{10} \\ 0 & 0 & \frac{1}{20} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{10} & \frac{3}{10} & 0 & 0 \\ 0 & 1/20 & 1/10 & 0 \\ 0 & 0 & 1/20 & 1/10 \\ 0 & 0 & 1/20 & 0 \end{bmatrix}$$

$$\therefore H(X, Y) = -\frac{1}{\log_{10} 2}\left[\frac{1}{4}\log\frac{1}{4} + \frac{1}{10}\log\frac{1}{10} + \frac{3}{10}\log\frac{3}{10} + \frac{1}{20}\log 1/20 + \frac{1}{10}\log 1/10 + \frac{1}{20}\log\frac{1}{20} + \frac{1}{20}\log\frac{1}{20} + \frac{1}{10}\log\frac{1}{10}\right]$$

$$= -\frac{1}{0.3}\left[-0.15 - 0.1 - 0.15 - 0.065 - 0.1 - 0.065 - 0.065 - 0.1\right]$$

$$= -\frac{1}{0.3}(-0.795)$$

$$\boxed{H(X, Y) = 2.65}$$

$$P(Y/X)$$

$$P(X,Y) = \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{4} & 3/10 & 0 & 0 \\ 0 & 1/20 & 1/10 & 0 \\ 0 & 0 & 1/20 & 1/10 \\ 0 & 0 & 1/20 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{4} & 3/4 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$H(Y/X) = \sum\sum P(x,y) \log p(y/x)$$

$$= -\frac{1}{\log 2}\Big[1/4 \log 1 + \frac{1}{4}\log 1/4 + 3/10 \log 3/4 + 1/20 \log 1/3$$
$$1/10 \log 2/3 + 1/20 \log 1/3 + 1/20 \log 1 +$$
$$+ 1/10 \log 2/3\Big]$$

$$= -\frac{1}{0.3}\Big[0 - 0.15 - 0.0374 - 0.0238 - 0.0176$$
$$- 0.0238 \cancel{-0.0238} - 0.0176\Big]$$

$$= \cancel{0.75}\ \frac{0.2402}{0.3} = \boxed{0.8}$$

$$\therefore I(X,Y) \Rightarrow H(Y) - H(Y/X)$$

$$= 1.85 - 0.6$$

$$= 1.25\ \text{bits/Symbol}.$$

Q. ~~The single unbiased~~

A single dice is tossed once. If the face of the die is 1, 2, 3, 4 and a coin is tossed once. If the face of the die is 5 or 6. The coin is tossed twice. Find the information conveyed ~~by~~ about the face of the die by the no. of H's obtained.

Sol^n.