POORNIMA
COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

Entropy (Average information Content of Symbol)

Consider that there are $M = \{m_1, m_2, m_3 \cdots\}$
diffrent message with probabilities $P = \{P_1, P_2, P_3 \cdots\}$
Suppose that a sequence of L is transmitted

$p_1 L$ message of m_1 are transmitted
$p_2 L$ " of m_2 are transmitted
⋮
$p_m L$ " of m_m are transmitted

$$I(m_1) = \log_2 \frac{1}{p_1}$$

if $(p_1 L)$ message of m_1 are transmitted

$$I_1(\text{total}) = P_1 L \log_2\left(\frac{1}{p_1}\right)$$

$$I(\text{total}) = P_1 L \log_2\left(\frac{1}{p_1}\right) + P_2 L \log_2\left(\frac{1}{p_2}\right) + \cdots + p_m L \log_2\left(\frac{1}{p_m}\right)$$

$$\text{Average Information} = \frac{\text{Total information}}{\text{No of message}}$$

$$= \frac{I(\text{total})}{L}$$

$$p_1 \log_2\left(\frac{1}{p_1}\right) + p_2 \log_2\left(\frac{1}{p_2}\right) + \cdots + p_m \log_2\left(\frac{1}{p_m}\right)$$

$$\text{Entropy}(H) = \sum_{k=1}^{m} P_k \log_2\left(\frac{1}{P_k}\right)$$

Properties of Entropy

1. Entropy is zero if the event is sure if $P_k = 1$, $P_k = 0$
2. When $p_k = 1/m$ for ~~at~~ m symbols than symbols are equally likely $\Rightarrow H = \log_2 m$
3. Upper bound on Entropy is given as

$$\boxed{H_{max} = \log_2 m}$$

Proof of Property

1. Proof 1 When $P_k = 1$

$$H = \sum_{k=1}^{m} P_k \log_2\left(\frac{1}{P_k}\right)$$

$$= \sum_{k=1}^{m} = \log_2 1$$

$$= \sum_{k=1}^{m} = \frac{\cancel{\log_{10}(1)}\ 0}{\log_{10}(2)}$$

$$\boxed{H = 0}$$

When $P_k = 0$ $\quad H = \sum_{k=1}^{m} \lim_{P_k \to 0} P_k \log_2\left(\frac{1}{P_k}\right)$

$$H = 0 \qquad \boxed{H = 0}$$

2. Proof 2 $P_k = \frac{1}{m}$; $H = \sum_{k=1}^{m} P_k \log_2\left(\frac{1}{P_k}\right)$

$$H = \sum_{k=1}^{m} \frac{1}{m} \log_2\left(1/1/m\right)$$

$$H = \sum_{k=1}^{m} \frac{1}{m} \log_2 m$$

$$H = \frac{\cancel{m}}{\cancel{m}} \log_2 m$$

$$\boxed{H = \log_2 m}$$

Source Efficiency and Redundancy:

Efficiency of the source (η)

$$\boxed{\eta_{source} = \frac{H}{H_{max}}}$$

H → entropy of source

H_{max} → max. entropy

Redundancy of the source (γ)

$$\boxed{\gamma_{source} = 1 - \eta_{source}}$$

η_{source} → efficiency

Information Rate:

$$\boxed{R = rH}$$

H → Entropy

r → rate at which message gen^n

Example for a discrete memoryless source there are three symbol with $p_1 = \alpha$ and $p_2 = p_3$ find the entropy of the source

Ans $p_1 = \alpha$

$$p_1 + p_2 + p_3 = 1$$

$$p_1 + p_2 + p_2 = 1$$

$$\alpha + 2p_2 = 1$$

$$p_2 = \left(\frac{1-\alpha}{2}\right) = p_3$$

$$P_2 = P_3 = \frac{1-\alpha}{2}$$

$$H = \sum_{K=1}^{3} P_K \log_2 \frac{1}{P_K}$$

$$= \alpha \log_2\left(\frac{1}{\alpha}\right) + \left(\frac{1-\alpha}{2}\right)\log_2\left(\frac{2}{1-\alpha}\right) + \left(\frac{1-\alpha}{2}\right)\log_2\left(\frac{2}{1-\alpha}\right)$$

$$= \alpha \log_2\left(\frac{1}{\alpha}\right) + (1-\alpha)\log_2\left(\frac{2}{1-\alpha}\right)$$

Ques 2 The source emits three message with probability $p_1 = 0.7$, $p_2 = 0.2$ and $p_3 = 0.1$. Calculate (i) Source Entropy (ii) maximum Entropy (iii) Source Efficiency (iv) Redundancy

Ans: (i) Source Entropy $(H) = \sum_{K=1}^{3} p_K \log_2\left(\frac{1}{p_K}\right)$

$$H = 0.7\log_2\left(\frac{1}{0.7}\right) + 0.2\log_2\left(\frac{1}{0.2}\right) + 0.1\log_2\left(\frac{1}{0.1}\right)$$

$$\boxed{H = 1.1568 \text{ b/m}}$$

(ii) max Entropy $(H_{max}) = \log_2 m = \log_2 3 =$

$$\frac{\log_{10} 3}{\log_{10} 2} = \boxed{1.5 \text{ b/m}}$$

(iii) Source Efficiency $\eta_{source} = \frac{H}{H_{max}} = \frac{1.1568}{1.5875}$

$$\boxed{\eta_{source} = 0.73}$$

(iv) Redundancy $\gamma_{source} = 1 - \eta_{source} = 1 - 0.73$

$$= \boxed{0.27}$$

Example

Q1 Find average self information

$p_1 = 1/2 \quad p_2 = \frac{1}{4} \quad P_3 = \frac{1}{8} \quad P_4 = \frac{1}{8}$

$$H(x_i) = \sum_{i=1}^{4} P(x_i) \log \frac{1}{P(x_i)}$$

$$= \frac{1}{2}\log\left(\frac{1}{1/2}\right) + \frac{1}{4}\log\left(\frac{1}{1/4}\right) + \frac{1}{8}\log\left(\frac{1}{1/8}\right) + \frac{1}{8}\log\left(\frac{1}{1/8}\right)$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 2^2 + \frac{\cancel{2}^{1}}{\cancel{8}_{4}}\log 2^3$$

$$= \frac{1}{2} + \frac{2}{4}(1) + \frac{3}{4}(1)$$

$$= \frac{2+2+3}{4} = \frac{7}{4} \text{ b/m}$$

$\boxed{H(x) = \frac{7}{4}}$ Average self information

Q. An event has six possible ~~outcomes~~ with the probablity $p_1 = 1/2$, $p_2 = 1/4$, $P_3 = \frac{1}{8}$, $p_4 = \frac{1}{16}$, $P_5 = \frac{1}{32}$, $P_6 = \frac{1}{32}$. find the ~~entropy~~ ~~of the~~ system. Also find the rate ~~of~~ information ~~if~~ there are 16 ~~outcomes~~ per second.

Ans

$$H = \sum_{K=1}^{6} P_K \log \frac{1}{P_K}$$

$$= \frac{1}{2}\log(2) + \frac{1}{4}\log(4) + \frac{1}{8}\log(8) + \frac{1}{16}\log(16) + \frac{2}{32}\log(32)$$

$$H = \frac{31}{16} \text{ bit/ms}$$

$\boxed{R = rH}$ R = rate ~~of~~ information
r = no ~~of~~ ~~outcome~~ generated in one second

$r = 16$ ~~outcomes~~

$$R = 16 \times \frac{31}{16} = 31 \text{ b/s}$$

$$\boxed{R = 31 \text{ b/s}}$$

Mutual Information

It is defined as the amount of information transmitted where x_i is transmitted and y_j is received

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

Conditional Probability [x_i is transmitted & y_j is received]

Average Mutual Information

Represented by $I(X;Y)$ and is calculated as bit/symbol

It is defined as amount of source information gained per received symbol

$$\boxed{I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\, I(x_i, y_j)}$$

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i|y_j)}{P(x_i)}$$

Properties of Mutual Information:

(i) Mutual information is symmetric

$$I(X;Y) = I(Y;X)$$

(iii) Mutual Information may be expressed as Entropy

$$I(X;Y) = H(X) - H(X|Y)$$
$$= H(Y) - H(Y/X)$$ Conditional Entropy.

(iv) mutual information is related to Joint Entropy $H(X,Y)$

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

Example

~~A transmitter~~ has an alphabet consisting of 5 letter $\{a_1, a_2, a_3, a_4, a_5\}$ and the receiver has an alphabet ~~of~~ four letter $\{b_1, b_2, b_3, b_4, b_5\}$

The joint probabilities ~~of~~ the system are shown as

$$P(A,B) = \begin{array}{c|cccc} & b_1 & b_2 & b_3 & b_4 \\ \hline a_1 & 0.25 & 0 & 0 & 0 \\ a_2 & 0.10 & 0.30 & 0 & 0 \\ a_3 & 0 & 0.05 & 0.10 & 0 \\ a_4 & 0 & 0 & 0.05 & 0.10 \\ a_5 & 0 & 0 & 0.05 & 0 \end{array}$$

Ans $I(A,B) = H(A) - H(A|B)$

$$H(A) \quad \sum_{i=1}^{n} P(a_i) \log_2\left(\frac{1}{P(a_i)}\right)$$

$$\boxed{H(A|B) \quad H(A,B) - H(B)}$$

$P(a_1) = 0.25$
$P(a_2) = 0.40$
$P(a_3) = 0.15$
$P(a_4) = 0.15$
$P(a_5) = 0.05$

$H(A) =$

$$0.25 \log\left(\frac{1}{0.25}\right) + 0.40 \log_2\left(\frac{1}{0.40}\right) + 2 \times 0.15 \log_2\left(\frac{1}{0.15}\right) + 0.05 \log_2\left(\frac{1}{0.05}\right)$$

$$\boxed{H(A) = 2.066 \text{ bits/symbol}}$$

for $H(B)$

$P(b_1) = 0.35$
$P(b_2) = 0.35$
$P(b_3) = 0.20$
$P(b_4) = 0.10$

$H(B) =$

$$2 \times 0.35 \log_2\left(\frac{1}{0.35}\right) + 0.20 \log_2\left(\frac{1}{0.20}\right) + 0.10 \log_2\left(\frac{1}{0.10}\right)$$

$$\boxed{H(B) = 1.85 \text{ b/sym}}$$

$H(A|B) \longrightarrow$ Get B_1 when A_1 is there.

$$H(A,B) = 0.25\log_2 \frac{1}{0.25} + 3\times 0.10 \log_2\left(\frac{1}{0.10}\right) + 3\times 0.05 \log_2\left(\frac{1}{0.05}\right)$$

$$+ 0.30 \log_2\left(\frac{1}{0.30}\right)$$

$H(A,B) = 2.666$ bits/symbol

$H(A|B) = H(A,B) - H(B)$

$H(A|B) = 2.666 - 1.857$

$= 0.809$

$I(A,B) = H(A) - H(A|B)$

$2.066 - 0.809$

$= 1.257$ b/s

Proof of mutual information Properties:

Proof 1: (i) $I(X;Y) = I(Y;X)$

~~from~~ probablity ~~theory~~

$$P(x_i, y_j) = P(x_i|y_j)\, P(y_j) \quad \text{--- (i)}$$

$$P(x_i, y_j) = P(y_j|x_i)\, P(x_i) \quad \text{--- (ii)}$$

So both equations are same

$$P(x_i|y_j)\, P(y_j) = P(y_j|x_i)\, P(x_i)$$

$$\frac{P(x_i|y_j)}{P(x_i)} = \frac{P(y_j|x_i)}{P(y_j)} \quad \text{--- (iii)}$$

$$I(X;Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i|y_j)}{P(x_i)} \quad \text{--- (iv)}$$

$$I(Y;X) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(y_j|x_i)}{P(y_j)} \quad \text{--- (v)}$$

By putting ~~the~~ value ~~of~~ (iii) into v

$$I(Y;X) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i|y_j)}{P(x_i)} \quad \text{--- (vi)}$$

So by (iv) and (vi)

$$\boxed{I(X;Y) = I(Y;X)}$$ Proved

Proof 2 $I(X;Y) = H(X) - H(X|Y)$

$H(Y) - H(Y|X)$

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\log_2 \frac{P(x_i|y_j)}{P(x_i)}$$

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\log_2 \frac{1}{P(x_i)} - \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\log_2\left(\frac{1}{P(x_i|y_j)}\right)$$

~~As~~ knows ~~that~~

$$H(X|Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\log_2\left(\frac{1}{P(x_i|y_j)}\right)$$

So by rewriting ~~the~~ equation.

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\log_2\left(\frac{1}{P(x_i)}\right) - H(X|Y)$$

~~if~~ ~~the~~ dipendency ~~of~~ $y_j = 0$

$$\sum_{i=1}^{m} P(x_i, y_j) = P(x_i)$$

$$I(X;Y) = \sum_{i=1}^{n} P(x_i)\log_2\left(\frac{1}{P(x_i)}\right) - H(X|Y)$$

$$\boxed{I(X;Y) = H(X) - H(X|Y)}$$

Scanned by CamScanner

Proof 3 mutual information is always +ve

$$I(X;Y) \geq 0$$

$$I(X;Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i|y_j)}{P(x_i)}$$

As we know that

$$P(x_i|y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

$$I(X;Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i, y_j)}{P(x_i) P(y_j)}$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i) P(y_j)}{P(x_i, y_j)}$$

$$-I(X;Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log_2 \frac{P(x_i) P(y_j)}{P(x_i, y_j)}$$

if $\boxed{\sum_{k=1}^{m} P_k \log_2 \left(\frac{q_k}{P_k}\right) \leq 0}$

$$-I(X;Y) \leq 0$$

or $\boxed{I(X;Y) \geq 0}$ Proved

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

$$H(X,Y) = H(X|Y) + H(Y) \quad \ldots (i)$$

and

$$I(X;Y) = H(X) - H(X|Y) \quad \ldots (ii)$$

~~from~~(i)

$$H(X|Y) = H(X,Y) - H(Y) \quad \ldots (iii)$$

put the value ~~of~~ (iii) in (ii)

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

Proved