Al-Mustaqbal University College
Department of Computer
Engineering Techniques

Information Theory and coding

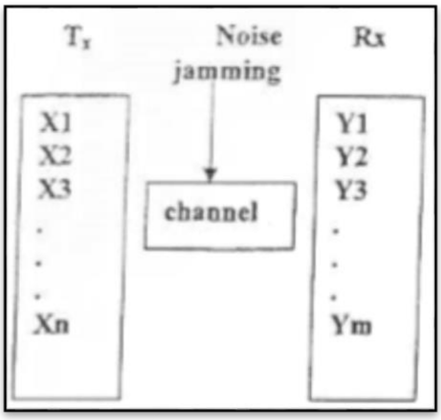
Fourth stage

Lecture 6

Mutual information for noisy channel

By:
MSC. Ridhab Sami

Mutual information for noisy channel:



Consider the set of symbols $x1$, $x2$,….,xn, the transmitter Tx my produce. The receiver Rx may receive $y1$, $y2$ ………ym. Theoretically, if the noise and jamming is neglected, then the set X=set Y. However and due to noise and jamming, there will be a conditional probability $P(yj \mid xi)$:

1- $P(xi)$ to be what is so called the a priori probability of the symbol xi , which is the prob of selecting xi for transmission.

2- 2- $P(yj \mid xi)$ to be what is called the aposteriori probability of the symbol xi after the reception of yj . The amount of information that yj provides about xi is called ***the mutual information*** between xi and yi . This is given by:

$$I(x_i, y_j) = \log_2\left(\frac{aposterori\ prob}{apriori\ prob}\right) = \log_2\left(\frac{P(y_j \mid x_i)}{P(x_i)}\right)$$

Properties of $I(xi , yj)$:

1- It is symmetric, $I(xi , yj) = I(yj , xi)$.

2- $I(xi , yj) > 0$ if aposteriori probability > a priori probability, yj provides +ve information about xi .

3- $I(xi , yj) = 0$ if aposteriori probability = a priori probability, which is the case of statistical independence when yj provides no information about xi .

4- $I(xi , yj) < 0$ if aposteriori probability < a priori probability, yj provides -ve information about xi , or yj adds ambiguity.

Example: Show that I(X, Y) is zero for extremely noisy channel.

Solution: For extremely noisy channel, then yjgives no information about xi the receiver can't decide anything about xi as if we transmit a deterministic signal xi but the receiver receives noise like signal yj that is completely has no correlation with xi . Then xi and yj are statistically independent so that $P(xi \mid\mid yj) = P(xi)and\ P(yj \mid\mid xi) = P(xi)\ for\ all\ i\ and\ j,then$: $I(xi , yj) = \log 21 = 0\ for\ all\ i\ \&\ j,then\ I(X, Y) = 0$

1. **Joint entropy:**

   In information theory, joint entropy is a measure of the uncertainty associated with a set of variables.

$$H(X,Y) = H(XY) = -\sum_{j=1}^{m}\sum_{i=1}^{n} P(xi,yj)log_2P(xi,yj) \quad bits/symbol$$

2. **Conditional entropy:**

   In information theory, the conditional entropy quantifies the amount of information needed to describe the outcome of a random variable Y given that the value of another random variable X is known.

$$H(Y \mid X) = -\sum_{j=1}^{m}\sum_{i=1}^{n} P(xi,yj)log_2P(yj \mid xi) \quad bits/symbol$$

3. **Marginal Entropies:**

   Marginal entropies is a term usually used to denote both source entropy H(X) defined as before and the receiver entropy H(Y) given by:

$$H(y) = -\sum_{j=1}^{m} P(yj)log_2P(yj) \qquad bit/symbol$$

4. **Relationship between joint, conditional and transinformation:**

<u>Noise entropy:</u> $H(Y \mid X) = H(X,Y) - H(X)$

<u>Loss entropy:</u> $H(X \mid Y) = H(X,Y) - H(Y)$

Also we have transinformation (average mutual information):

$$I(X,Y) = H(X) - H(X \mid Y)$$

$$I(X,Y) = H(Y) - H(Y \mid X)$$

Example: The joint probability of a system is given by:

$$P(X,Y) = \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \begin{bmatrix} 0.5 & 0.25 \\ 0 & 0.125 \\ 0.0625 & 0.0625 \end{bmatrix}$$

Find:

1- Marginal entropies.

2- Joint entropy.

3- Conditional entropies.

4- The transinformation.

Solution:

	x1	x2	x3		y1	y2
P(x)=	[0.75	0.125	0.125],	P(y)=	[0.5625	0.4375]

1-

$$H(x) = -\sum_{i=1}^{n} P(xi)log_2P(xi) = -\left[\frac{0.75ln0.75 + 2 * 0.125ln0.125}{ln2}\right]$$

$$= 1.06127 \ \ bit/symbol$$

$$H(y) = -\sum_{j=1}^{m} P(yj)log_2P(yj) = -\left[\frac{0.5625ln0.5625 + 0.4375ln0.4375}{ln2}\right]$$

$$= 0.9887 \ bit/symbol$$

2-

$$\mathrm{H(x,y)} = -\sum_{j=1}^{m}\sum_{i=1}^{n} \mathrm{P(xi,yj)}\log_2 \mathrm{P(xi,yj)}$$

$$= -\left[\frac{0.5ln0.5 + 0.25ln0.25 + 0.125ln0.125 + 2 * 0.0625ln0.0625}{ln2}\right]$$

$$= \quad 1.875\ bit/symbol$$

3-

$$H(y \mid x) = H(x,y) - H(x) = 1.875 - 1.06127 = 0.813 \quad bit/symbol$$

$$H(x \mid y) = H(x,y) - H(y) = 1.875 - 0.9887 = 0.886 \quad bit/symbol$$

4-

$$I(x,y) = H(x) - H(x \mid y) = 1.06127 - 0.886 = 0.175\ \ bit/symbol$$