Redundancy

COMMUNICATION NETWORK.
NOISE CHARACTERISTICS OF A CHANNEL

Communication Network

- Consider a source of communication with a given alphabet. The source is linked to the receiver via a channel.
- The system may be described by a joint probability matrix: by giving the probability of the joint occurrence of two symbols, one at the **input** and another at the **output**.

Communication Network

- x_k – a symbol, which was sent; y_j - a symbol, which was received
- The joint probability matrix:

$$\left[P\{X,Y\}\right]=\begin{pmatrix} P\{x_1,y_1\} & P\{x_1,y_2\} & \dots & P\{x_1,y_m\} \\ P\{x_2,y_1\} & P\{x_2,y_2\} & \dots & P\{x_2,y_m\} \\ \dots & \dots & \dots & \dots \\ P\{x_n,y_1\} & P\{x_n,y_2\} & \dots & P\{x_n,y_m\} \end{pmatrix}$$

Communication Network: Probability Schemes

- There are following five probability schemes of interest in a product space of the random variables X and Y:
- $[P\{X,Y\}]$ – joint probability matrix
- $[P\{X\}]$ – marginal probability matrix of X
- $[P\{Y\}]$ – marginal probability matrix of Y
- $[P\{X|Y\}]$ – conditional probability matrix of $X|Y$
- $[P\{Y|X\}]$ – conditional probability matrix of $Y|X$

Communication Network: Entropies

- There is the following interpretation of the five entropies corresponding to the mentioned five probability schemes:

- $H(X,Y)$ – average information per pairs of transmitted and received characters (the entropy of the system as a whole);
- $H(X)$ – average information per character of the source (the entropy of the source)
- $H(Y)$ – average information per character at the destination (the entropy at the receiver)
- $H(Y|X)$ – a specific character x_k being transmitted and one of the permissible y_j may be received (a measure of information about the receiver, where it is known what was transmitted)
- $H(X|Y)$ – a specific character y_j being received ; this may be a result of transmission of one of the x_k with a given probability (a measure of information about the source, where it is known what was received)

Communication Network: Entropies' Meaning

- $H(X)$ and $H(Y)$ give indications of the probabilistic nature of the transmitter and receiver, respectively.
- $H(X,Y)$ gives the probabilistic nature of the communication channel as a whole (the entropy of the union of X and Y).
- $H(Y|X)$ gives an indication of the noise (errors) in the channel
- $H(X|Y)$ gives a measure of equivocation (how well one can recover the input content from the output)

Communication Network: Derivation of the Noise Characteristics

- In general, the joint probability matrix is not given for the communication system.
- It is customary to specify the noise characteristics of a channel and the source alphabet probabilities.
- From these data the joint and the output probability matrices can be derived.

Communication Network: Derivation of the Noise Characteristics

- Let us suppose that we have derived the joint probability matrix:

$$\left[P\{X,Y\}\right]=\begin{pmatrix} p\{x_1\}\,p\{y_1|x_1\} & p\{x_1\}\,p\{y_2|x_1\} & \dots & p\{x_1\}\,p\{y_m|x_1\} \\ p\{x_2\}\,p\{y_1|x_2\} & p\{x_2\}\,p\{y_2|x_2\} & \dots & p\{x_2\}\,p\{y_m|x_2\} \\ \dots & \dots & \dots & \dots \\ p\{x_n\}\,p\{y_1|x_n\} & p\{x_n\}\,p\{y_2|x_n\} & \dots & p\{x_n\}\,p\{y_m|x_n\} \end{pmatrix}$$

Communication Network: Derivation of the Noise Characteristics

- In other words :

$$\left[P\{X,Y\}\right]=\left[P\{X\}\right]\left[P\{Y|X\}\right]$$

- where:

$$\left[P\{X\}\right]=\begin{pmatrix} p\{x_1\} & 0 & 0 & \dots & 0 \\ 0 & p\{x_2\} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p\{x_{n-1}\} & 0 \\ 0 & 0 & \dots. & 0 & p\{x_n\} \end{pmatrix};$$

Communication Network: Derivation of the Noise Characteristics

- If $[P\{X\}]$ is not diagonal, but a row matrix (n-dimensional vector) then

$$\left[P\left\{Y\right\}\right]=\left[P\left\{X\right\}\right]\left[P\left\{Y\mid X\right\}\right]$$

- where $[P\{Y\}]$ is also a row matrix (m-dimensional vector) designating the probabilities of the output alphabet.

Communication Network: Derivation of the Noise Characteristics

- Two discrete channels of our particular interest:
- Discrete noise-free channel (an ideal channel)
- Discrete channel with independent input-output (errors in the channel occur, thus noise is presented)

Noise-Free Channel

- In such channels, every letter of the input alphabet is in a one-to-one correspondence with a letter of the output alphabet. Hence the joint probability matrix is of diagonal form:

$$\left[P\{X,Y\}\right]=\begin{pmatrix} p\{x_1,y_1\} & 0 & 0 & \dots & 0 \\ 0 & p\{x_2,y_2\} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p\{x_{n-1},y_{n-1}\} & 0 \\ 0 & 0 & \dots. & 0 & p\{x_n,y_n\} \end{pmatrix};$$

Noise-Free Channel

- The channel probability matrix is also of diagonal form:

$$\left[P\{X\,|\,Y\}\right]=\left[P\{Y\,|\,X\}\right]=\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots. & 0 & 1 \end{pmatrix};$$

- Hence the entropies

$$H\left(Y\,|\,X\right)=H\left(X\,|\,Y\right)=0$$

Noise-Free Channel

- The entropies H(X,Y), H(X), and H(Y):

$$H\left(X,Y\right)=H(X)=H(Y)=$$

$$=-\sum_{i=1}^{n} p\left\{x_i,y_i\right\}\log p\left\{x_i,y_i\right\}$$

Noise-Free Channel

- Each transmitted symbol is in a one-to-one correspondence with one, and only one, received symbol.
- The entropy at the receiving end is exactly the same as at the sending end.
- The individual conditional entropies are all equal to zero because any received symbol is completely determined by the transmitted symbol and vise versa.

Discrete Channel with Independent Input-Output

- In this channel, there is no correlation between input and output symbols: any transmitted symbol x_i can be received as any symbol y_j of the receiving alphabet with equal probability:

$$\left[P\{X,Y\}\right] = \underbrace{\begin{pmatrix} p_1 & p_1 & \cdots & p_1 \\ p_2 & p_2 & \cdots & p_2 \\ \cdots & \cdots & \cdots & \cdots \\ p_m & p_m & \cdots & p_m \end{pmatrix}}_{n \text{ identical columns}} ; \sum_{i=1}^{m} p_i = \frac{1}{n} = p\{y_j\}; \; p\{x_i\} = np_i$$

Discrete Channel with Independent Input-Output

- Since the input and output symbol probabilities are statistically independent, then

$$p\left\{x_i, y_j\right\} = \underbrace{p\left\{x_i\right\}}_{np_i} \underbrace{p\left\{y_j\right\}}_{1/n} = np_i \frac{1}{n} = p_i$$

$$p\left\{x_i \mid y_j\right\} = p_1\left\{x_i\right\} = np_i$$

$$p\left\{y_j \mid x_i\right\} = p_1\left\{y_j\right\} = \frac{1}{n}$$

Discrete Channel with Independent Input-Output

$$H(X,Y) = -n\left(\sum_{i=1}^{m} p_i \log p_i\right)$$

$$H(X) = -\sum_{i=1}^{m} np_i \log np_i = -n\left(\sum_{i=1}^{m} p_i \log p_i\right) - \log n$$

$$H(Y) = -n\left(\frac{1}{n}\right)\log\frac{1}{n} = \log n$$

$$H(X|Y) = -\sum_{i=1}^{n} np_i \log np_i = H(X); \; H(Y|X) = -\sum_{i=1}^{m} np_i \log\frac{1}{n} = \log n = H(Y)$$

- The last two equations show that this channel conveys no information: a symbol that is received does not depend on a symbol that was sent

Noise-Free Channel vs Channel with Independent Input-Output

- Noise-free channel is a loss-less channel, but it carries no information.
- Channel with independent input/output is a completely lossy channel, but the information transmitted over it is a pure noise.
- Thus these two channels are two “extreme“ channels. In the real world, real communication channels are in the middle, between these two channels.

Basic Relationships among Different Entropies in a Two-Port Communication Channel

- We have to take into account that

$$p\left\{x_k, y_k\right\} = p\left\{x_k \mid y_j\right\} p\left\{y_j\right\} = p\left\{y_j \mid x_k\right\} p\left\{x_k\right\}$$

$$\log p\left\{x_k, y_k\right\} = \underbrace{\log p\left\{x_k \mid y_j\right\} p\left\{y_j\right\}}_{\log p\left\{x_k \mid y_j\right\} + \log p\left\{y_j\right\}} = \underbrace{\log p\left\{y_j \mid x_k\right\} p\left\{x_k\right\}}_{\log p\left\{y_j \mid x_k\right\} + \log p\left\{x_k\right\}}$$

- Hence

$$H(X,Y) = H(X \mid Y) + H(Y) = H(Y \mid X) + H(X)$$

Basic Relationships among Different Entropies in a Two-Port Communication Channel

- Fundamental Shannon's inequalities:

$$H(X) \geq H(Y|X) \quad H(Y) \geq H(Y|X)$$

- The conditional entropies never exceed the marginal ones.
- The equality sigh hold if, and only if X and Y are statistically independent and therefore

$$\frac{p\{x_k\}}{p\{x_k | y_j\}} = \frac{p\{y_j\}}{p\{y_j | x_k\}} = 1$$

Mutual Information

- What is a mutual information between x_i , which was transmitted and y_j, which was received, that is, the information conveyed by a pair of symbols (x_i, y_j)?

$$I\left(x_i; y_j\right) = \log \frac{\overbrace{p\left\{x_i \mid y_j\right\}}^{\frac{p\left\{x_i, y_j\right\}}{p\left\{y_j\right\}}}}{p\left\{x_i\right\}} = \log \frac{p\left\{x_i, y_j\right\}}{p\left\{x_i\right\} p\left\{y_j\right\}}$$

Mutual Information

- This probability determines the a posteriori knowledge of what was transmitted

$$I\left(x_i; y_j\right) = \log \frac{p\left\{x_i \mid y_j\right\}}{p\left\{x_i\right\}}$$

- This probability determines the a priori knowledge of what was transmitted
- The ratio of these two probabilities (more exactly – its logarithm) determines the gain of information

Mutual and Self-Information

- The function $I(x_i, x_i)$ is the self-information of a symbol x_i (it shows a priori knowledge that x_i was transmitted with the probability $p(x_i)$ and a posteriori knowledge is that x_i has definitely been transmitted).
- The function $I(y_j, y_j)$ is the self-information of a symbol y_i (it shows a priori knowledge that y_i was received with the probability $p(y_i)$ and a posteriori knowledge is that y_i has definitely been received).

Mutual and Self-Information

- For the self-information:

$$I\left(x_i\right) = I\left(x_i, x_i\right) = \log \frac{p\left\{x_i \mid x_i\right\}}{p\left\{x_i\right\}} = \log \frac{1}{p\left\{x_i\right\}}$$

- The mutual information does not exceed the self-information:

$$I\left(x_i; y_j\right) \le I\left(x_i; x_i\right) = I\left(x_i\right)$$

$$I\left(x_i; y_j\right) \le I\left(y_j; y_j\right) = I\left(y_j\right)$$

Mutual Information

- The mutual information of all the pairs of symbols can be obtained by averaging the mutual information per symbol pairs:

$$I(X;Y) = \overline{I(x_i, y_j)} = \sum_j \sum_i p\{x_i, y_j\} I(x_i, y_j) =$$

$$= \sum_j \sum_i p\{x_i, y_j\} \underbrace{\log \frac{p\{x_i \mid y_j\}}{p\{x_i\}}}_{I(x_i, y_j)} =$$

$$= \sum_j \sum_i p\{x_i, y_j\} \left(\log p\{x_i \mid y_j\} - \log p\{x_i\} \right)$$

Mutual Information

- The mutual information of all the pairs of symbols $I(X;Y)$ shows the amount of information containing in average in one received message with respect to the one transmitted message
- $I(X;Y)$ is also referred to as transinformation (information transmitted through the channel)

Mutual Information

- Just to recall:

$$H(X) = -\sum_{k=1}^{n} p\{x_k\} \log p\{x_k\} \qquad H(Y) = -\sum_{j=1}^{m} p\{y_j\} \log p\{y_j\}$$

$$H(X|Y) = -\sum_{j=1}^{m}\sum_{k=1}^{n} p\{y_j\} p\{x_k|y_j\} \log p\{x_k|y_j\}$$

$$H(Y|X) = -\sum_{k=1}^{n}\sum_{j=1}^{m} p\{x_k\} p\{y_j|x_k\} \log p\{y_j|x_k\}$$

$$I(X;Y) = \sum_j \sum_i p\{x_i, y_j\}\left(\log p\{x_i|y_j\} - \log p\{x_i\}\right) =$$

$$= \underbrace{\sum_j \sum_i \underbrace{p\{x_i, y_j\}}_{p\{y_j\}p\{x_k|y_j\}} \log p\{x_i|y_j\}}_{-H(X|Y)} - \underbrace{\sum_i \underbrace{\sum_j p\{x_i, y_j\}}_{p\{x_i\}} \log p\{x_i\}}_{H(X)}$$

Mutual Information

- It follows from the equations from the previous slide that:

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

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

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

- $H(X|Y)$ shows an average loss of information for a transmitted message with respect to the received one
- $H(Y|X)$ shows a loss of information for a received message with respect to the transmitted one

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

Mutual Information

- For a noise-free channel, $I(X;Y)=H(X)=H(Y)=H(X,Y)$,which means that the information transmitted through this channel does not depend on what was sent/received. It is always completely predetermined by the transmitted content.

Mutual Information

- For a channel with independent input/output , $I(X;Y)=H(X)\text{-}H(X|Y)= H(X)\text{-}H(X)=0$,which means that no information is transmitted through this channel.

Channel Capacity

- The channel capacity (bits per symbol) is the maximum of transinformation with respect to all possible sets of probabilities that could be assigned to the source alphabet (C. Shannon):

$$C = \max I\left(X;Y\right) = \max\left[H(X) - H(X\,|\,Y)\right] =$$

$$= \max\left[H(Y) - H(Y\,|\,X)\right]$$

- The channel capacity determines the upper bound of the information that can be transmitted through the channel

Rate of Transmission of Information through the Channel

- If all the transmitted symbols have a common duration of t seconds then the rate of transmission of information through the channel (bits per second or capacity per second) is

$$C_t = \frac{1}{t} C$$

Absolute Redundancy

- Absolute redundancy of the communication system is the difference between the maximum amount of information, which can be transmitted through the channel and its actual amount:

$$R_a = C - I\left(X;Y\right)$$

Relative Redundancy

- Relative redundancy of the communication system is the ratio of absolute redundancy to channel capacity:

$$R_r = \frac{R_a}{C} = \frac{C - I\left(X;Y\right)}{C} = 1 - \frac{I\left(X;Y\right)}{C}$$