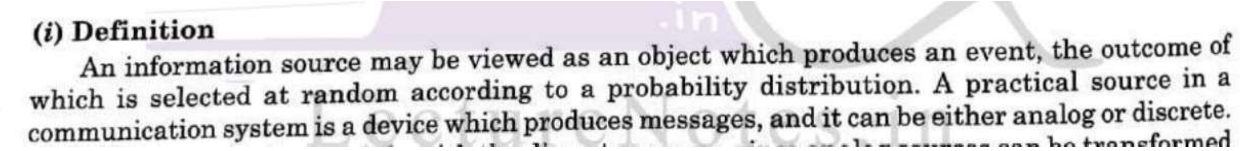
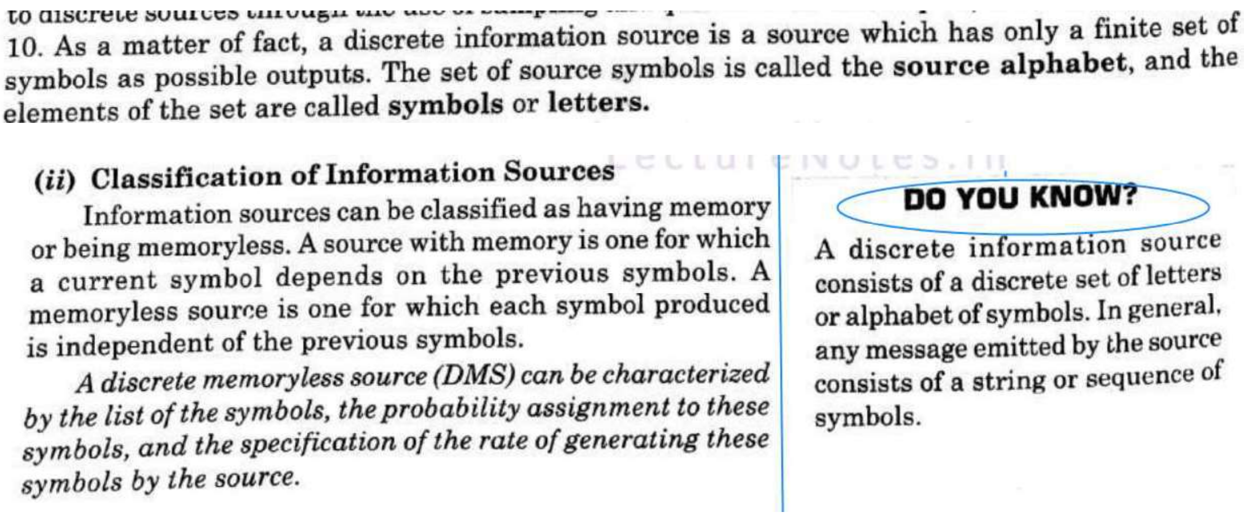
INFORMATION THEORY:– Entropy, Information rate, classification of codes, Kraft McMillan inequality, Source coding theorem, Shannon-Fano coding, Huffman coding, Extended Huffman coding - Joint and conditional entropies, Mutual information - Discrete memory less channels – BSC, BEC – Channel capacity, Shannon limit



(*i*) Definition

An information source may be viewed as an object which produces an event, the outcome of which is selected at random according to a probability distribution. A practical source in a communication system is a device which produces messages, and it can be either analog or discrete.

can be transformed



to discrete sources through the use of sampling

10. As a matter of fact, a discrete information source is a source which has only a finite set of symbols as possible outputs. The set of source symbols is called the **source alphabet**, and the elements of the set are called **symbols** or **letters.**

(*ii*) Classification of Information Sources

Information sources can be classified as having memory or being memoryless. A source with memory is one for which a current symbol depends on the previous symbols. A memoryless source is one for which each symbol produced is independent of the previous symbols.

A discrete memoryless source (DMS) can be characterized by the list of the symbols, the probability assignment to these symbols, and the specification of the rate of generating these symbols by the source.

DO YOU KNOW?

A discrete information source consists of a discrete set of letters or alphabet of symbols. In general, any message emitted by the source consists of a string or sequence of symbols.

****

Information is the source of a communication system, whether it is analog or digital.

Information theory is a mathematical approach to the study of coding of information along with the quantification, storage, and communication of information.

Conditions of Occurrence of Events

If we consider an event, there are three conditions of occurrence.

- If the event has not occurred, there is a condition of **uncertainty**.
- If the event has just occurred, there is a condition of **surprise**.
- If the event has occurred, a time back, there is a condition of having some **information**.

These three events occur at different times. The differences in these conditions help us gain knowledge on the probabilities of the occurrence of events.

The performance of the communication system is measured in terms of its error probability. An errorless transmission is possible when probability of error at the receiver approaches zero.

The information theory is related to the concepts of statistical properties of messages/sources, channels, noise interference etc. The information theory is used for mathematical modeling and analysis of the communication systems.

1.1.1 Uncertainty

Consider the source which emits the discrete symbols randomly from the set of fixed alphabet i.e.

$$X = \{x_0, x_1, x_2, \ldots\ldots x_{K-1}\} \quad \ldots (1.1.1)$$

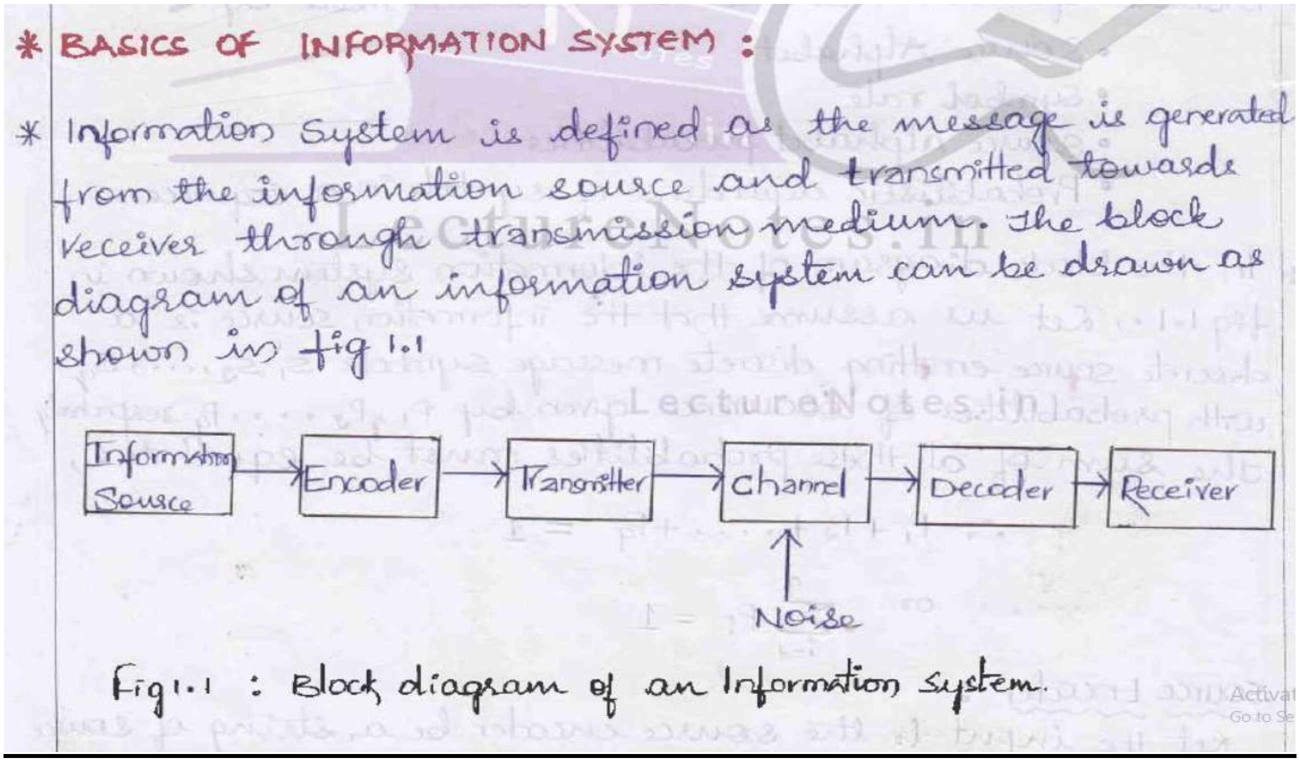
The various symbols in 'X ' have probabilities of $p_0, p_1, p_2, \ldots\ldots\ldots$ etc., which can be written as,

$$P(X=x_k) = p_k \qquad k = 0, 1, 2, \ldots\ldots K-1 \quad \ldots (1.1.2)$$

This set of probabilities satisfy the following condition,

$$\sum_{k=0}^{K-1} p_k = 1 \quad \ldots (1.1.3)$$

Such information source is called discrete information source. The concept of 'Information' produced by the source is discussed in the next section. This idea of Information is related to 'Uncertainty' or 'Surprise'. Consider the emission of symbol

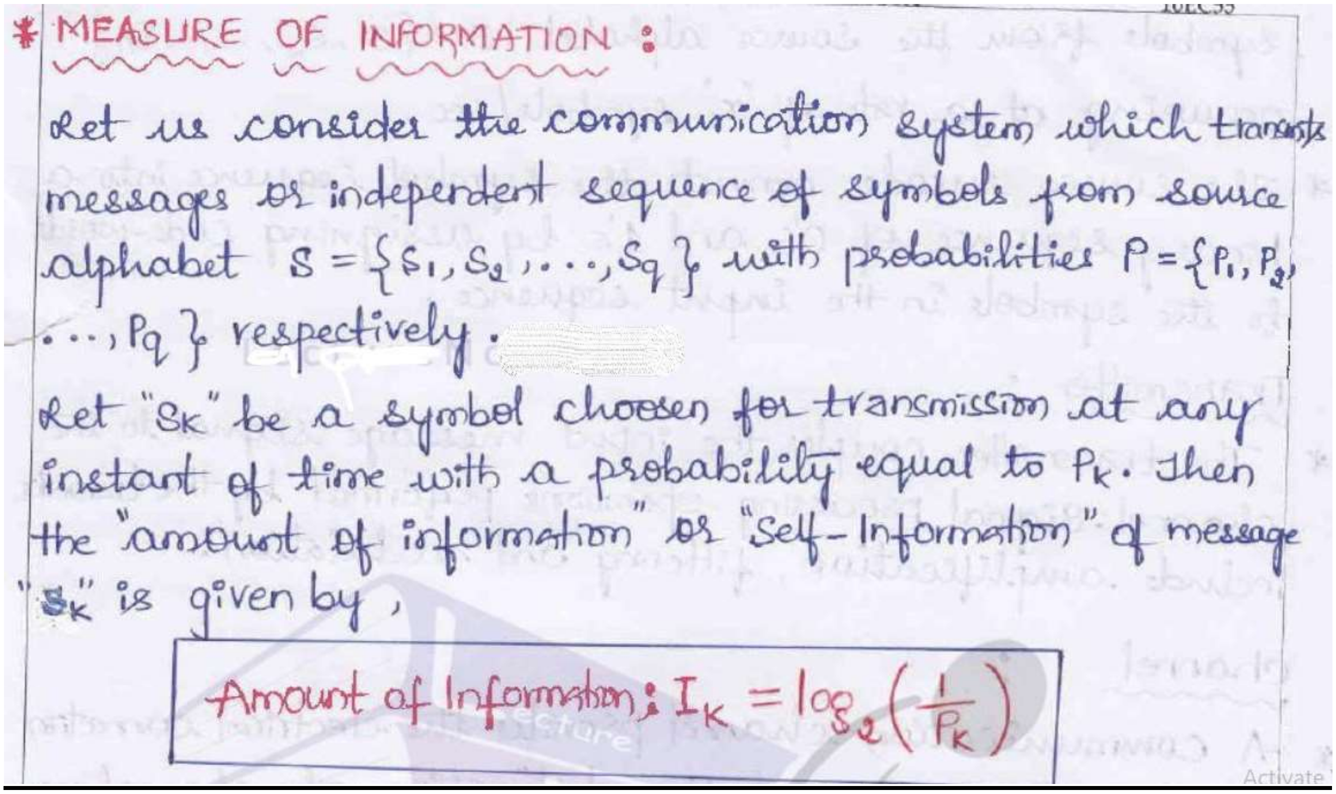


*** BASICS OF INFORMATION SYSTEM :**

* Information System is defined as the message is generated from the information source and transmitted towards receiver through transmission medium. The block diagram of an information system can be drawn as shown in fig 1.1

Fig1.1 : Block diagram of an Information system.

Definition of Information:



*** MEASURE OF INFORMATION :**

Let us consider the communication system which transmits messages or independent sequence of symbols from source alphabet $S = \{s_1, s_2, \dots, s_q\}$ with probabilities $P = \{P_1, P_2, \dots, P_q\}$ respectively.

Let "S_K" be a symbol choosen for transmission at any instant of time with a probability equal to P_K. Then the "amount of information" or "Self-Information" of message "S_K" is given by,

$$\boxed{\text{Amount of Information: } I_K = \log_2\left(\frac{1}{P_K}\right)}$$

* Unit of Information :

In the above equation $\log_2\left(\frac{1}{P_k}\right) = \frac{\log_{10}(1/P_k)}{\log_{10} 2}$, if the base of the logarithm is 2, then the units are called "BITS", which is the short form of Binary Units". If the base is "10", the units are "HARTLEYS" or "DECITS". If the base is "e", the units are "NATS" and if the base, in general, is "r", the units are called "r-ary units".

* The most widely used unit of information is "BITS", where the base of the logarithm is 2.

INFORMATION RATE :

Let us suppose that, the symbols are emitted by the source at a fixed time rate "r_s" symbols/sec. The average source information rate R_s" in bits/sec is defined as the product of the average information content per symbol and the message symbol rate "r_s".

$$\therefore \boxed{R_s = r_s H(S)} \text{ bits/sec or BPS}$$

1.2.1 Properties of Information

Following properties can be written for information.

i) If there is more uncertainty about the message, information carried is also more.

ii) If receiver knows the message being transmitted, the amount of information carried is zero.

iii) If I_1 is the information carried by message m_1, and I_2 is the information carried by m_2, then amount of information carried combinely due to m_1 and m_2 is $I_1 + I_2$.

iv) If there are $M = 2^N$ equally likely messages, then amount of information carried by each message will be N bits.

Prove the following statement, "If receiver knows the message being transmitted, the amount of information carried is "Zero".

Here it is stated that receiver "knows" the message. This means only one message is transmitted. Hence probability of occurrence of this message will be $P_k = 1$.

Therefore, the amount of information carried by this type of message is,

$$I_k = \log_2\left(\frac{1}{P_k}\right)$$

$$= \frac{\log_{10} 1}{\log_{10} 2}$$

$$\therefore \boxed{I_k = 0 \text{ bits}}$$

Example 1.2.1 : *Calculate the amount of information if* $p_k = \frac{1}{4}$.

Solution : From equation 1.2.1 we know that amount of information is given as,

$$I_k = \log_2\left(\frac{1}{p_k}\right) = \frac{\log_{10}\left(\frac{1}{p_k}\right)}{\log_{10} 2}$$

$$= \frac{\log_{10} 4}{\log_{10} 2}$$

$$= 2 \text{ bits}$$

➡ **Example 1.2.3 :** *In binary PCM if '0' occur with probability $\frac{1}{4}$ and '1' occur with probability $\frac{3}{4}$, then calculate amount of information conveyed by each binit.*

Solution : Here binit '0' has $p_1 = \frac{1}{4}$

and binit '1' has $p_2 = \frac{3}{4}$

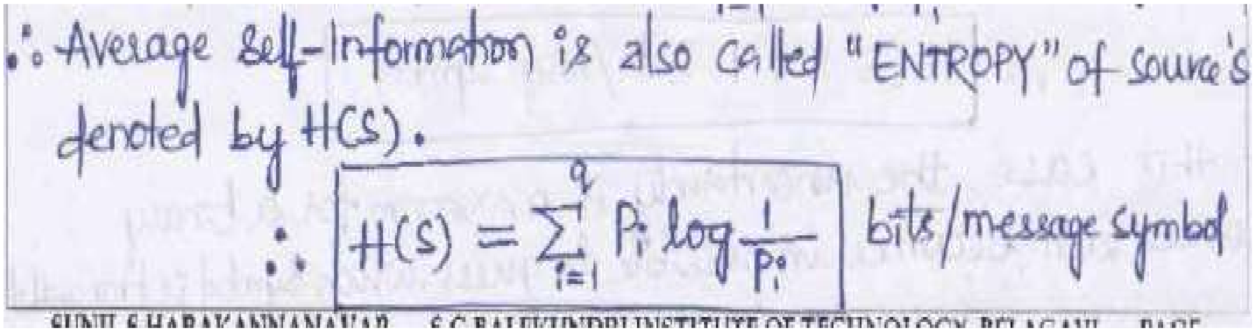
Then amount of information is given by equation 1.2.1 as,

$$I_k = \log_2\left(\frac{1}{p_k}\right)$$

$$\left.\begin{array}{lll} \text{with} & p_1 = \frac{1}{4}, & I_1 = \log_2 4 = \frac{\log_{10} 4}{\log_{10} 2} = 2 \text{ bits} \\ \text{and with} & p_2 = \frac{3}{4}, & I_2 = \log_2\left(\frac{4}{3}\right) = \frac{\log_{10}(4/3)}{\log_{10} 2} = 0.415 \text{ bits} \end{array}\right\} \quad \dots (1.2.4)$$

Here observe that binit '0' has probability $\frac{1}{4}$ and it carries 2 bits of information. Whereas binit '1' has probability $\frac{3}{4}$ and it carries 0.415 bits of information. This shows that if probability of occurrence is less, information carried is more, and vice versa.

ENTROPY:



∴ Average self-Information is also called "ENTROPY" of source's denoted by H(S).

$$\therefore \boxed{H(S) = \sum_{i=1}^{q} P_i \log \frac{1}{P_i}} \text{ bits/message symbol}$$

Illustration 1 :

Let us consider a binary source with source alphabet $S = \{s_1, s_2\}$ with probabilities $P = \left\{\frac{1}{256}, \frac{255}{256}\right\}$

Then, Entropy $H(S) = \sum_{i=1}^{2} P_i \log \frac{1}{P_i}$

$$= \frac{1}{256}\log_2 256 + \frac{255}{256}\log_2 \frac{256}{255}$$

$$\therefore \boxed{H(S) = 0.037 \text{ bits/msg symbol.}}$$

∴ The average uncertainty is very very small and is relatively very very easy to guess whether s_1 or s_2 will occur.

Illustration 2 :

Let $S' = \{S_3, S_4\}$ with $P' = \left\{\frac{7}{16}, \frac{9}{16}\right\}$

Then, Entropy $H(S') = \frac{7}{16}\log_2 \frac{16}{7} + \frac{9}{16}\log_2 \frac{16}{9}$

$$\therefore \boxed{H(S') = 0.989 \text{ bits/msg symbol}}$$

In this case, it is hard to guess whether S_3 or S_4 is transmitted

Illustration 3 :

Let $S'' = \{S_5, S_6\}$ with $P'' = \{1/2, 1/2\}$

Then, Entropy $H(S'') = \frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2$

$$\therefore \boxed{H(S'') = 1 \text{ bits/msg symbol}}$$

In this case, the uncertainty is maximum for a binary source and becomes impossible to guess which symbol is transmitted

Consider a source $S = \{s_1, s_2, s_3\}$ with $P = \{1/2, 1/4, 1/4\}$

Find (a) self Information of each message

(b) Entropy of source 'S'.

(a) Self Information of $s_1 = I_1 = \log_2 \frac{1}{P_1} = \log_2 2 = 1$ bits

self Information of $S_2 = I_2 = \log_2 \frac{1}{P_2} = \log_2 4 = 2$ bits

self Information of $S_3 = I_3 = \log_2 \frac{1}{P_3} = \log_2 4 = 2$ bits

b). Average Information Content or Entropy is given by,

$$H(S) = \sum_{i=1}^{3} P_i I_i = P_1 I_1 + P_2 I_2 + P_3 I_3$$

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

$$\therefore \quad \boxed{H(S) = 1.5 \text{ bits/msg symbol}}$$