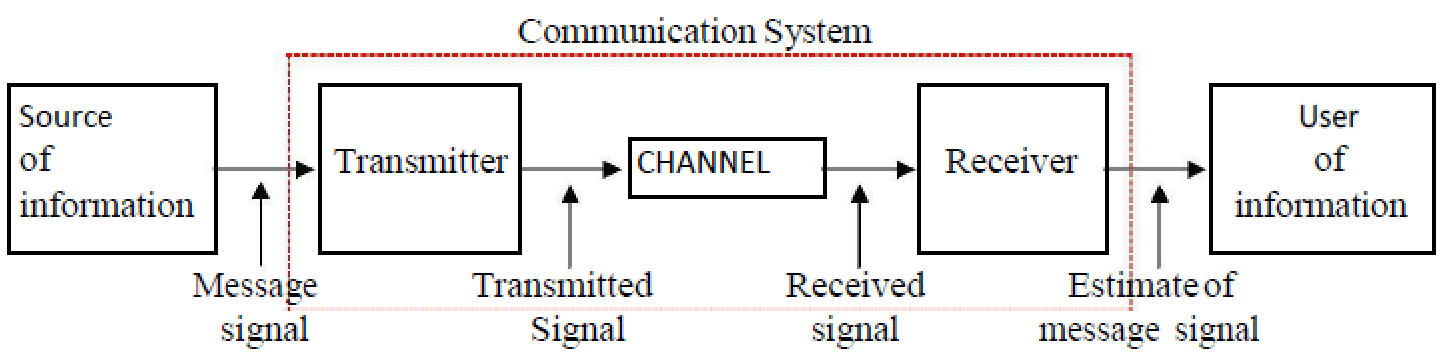
Information Theory & Coding

- Communication Model
- Introduction to ITC (Shannon's Theory)
- Amount of Information
- Average Information (Entropy)
- Information rate

Communication System



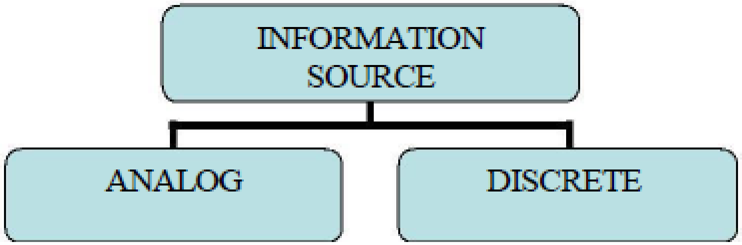
Communication Model

Transmitter	Channel	Receiver
Source	**Physical (wire)**	Antenna
Transducer	Coaxial Cable	Filter
Amplifier	Optical Cable	Demodulator
Modulator	**Wireless**	Decoder
Encoder	Air	Destination
Antenna	Space	

Shannon's Theory (Basic Concept of ITC)

1) Measure of Source Information
2) Information Capacity of Channel
3) Coding

Information Source



Source definition

➢**Analog** : Emit a continuous – amplitude, continuous – time electrical wave from.

➢**Discrete** : Emit a sequence of letters of symbols.

The output of a discrete information source is a string or sequence of symbols (dice , coin)

Amount of Information

- Consider the following examples:

1. Today is Monday
2. It is a cloudy day
3. Possible snow fall today

Continuous ….

Amount of information received is obviously different for these messages.

- Message (1) Contains very little information since the everyone know that today is Monday

- Message (2) The forecast of "cloudy day" contains more information , since it is not an event that occurs often.

- Message (3) In contrast, the forecast of "snow fall" convey even more information, since the occurrence of snow in jodhpur is a rare/impossible event.

Continuous ….

1) It is related to the probability of occurrence of the event.
2) Message associated with an event "least likely to occur" contains most information
3) Information is proportional to uncertainty of an event
4) Information is inversely proportional to probability

Information

an information source emits one of "m" possible messages $x_1, x_2 \ldots\ldots x_m$ with $p_1, p_2 \ldots\ldots p_m$ as their probabilities of occurrence.

The information content of the i^{th} message, can be written as

$$I \quad \alpha \quad 1/P_i$$

$$I(x_i) = \log_b \frac{1}{P(x_i)} = -\log_b P(x_i)$$

Base of log decide unit of information

b=2 , bits (Binary digit)

b=10 , decit (Decimal digit)

b=e , nats (Natural digit)

Properties

Basic

$$I(x_i) = 0 \quad \text{for} \quad P(x_i) = 1$$

$$I(x_i) \geqslant 0$$

$$I(x_i) > I(x_j) \quad \text{if } P(x_i) < P(x_j)$$

$$I(x_i x_j) = I(x_i) + I(x_j) \quad \text{if } x_i \text{ and } x_j \text{ are independent}$$

Question

- Information content in a universally true event is………

  $p=1$

  $I = \log 1/p = \log 1 = 0$

- A discrete source emits five symbols with the probability $P_1=1/2$, $P_2=1/4$, $P_3=1/8$, $P_4=1/16$ and $P_5=1/16$. Find information content of each symbols.

  $I = \log 1/p = \text{Log } 1/(1/2) = \log 2 = 1$ bits

  (Ans 1,2,3, 4 and 4 bits)

ENTROPY

Average Information per symbol

$x_1\, x_2\, x_1\, x_3 \,---$

length of Message = L

Set $\{x_1, x_2, x_3 ---- x_m\}$,

probability $\{P_1, P_2, P_3, ---- P_m\}$

Symbol	Respective probability	Information content each symbol	Nos. of symbol	Total Info. by a symbol
x_1	P_1	$I_1 = \log_2 \frac{1}{P_1}$	$P_1 L$	$P_1 L \log_2 \frac{1}{P_1}$
x_2	P_2	$I_2 = \log_2 \frac{1}{P_2}$	$P_2 L$	$P_2 L \log_2 \frac{1}{P_2}$
x_3	P_3	$I_3 = \log_2 \frac{1}{P_3}$	$P_3 L$	$P_3 L \log_2 \frac{1}{P_3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	P_m	$P_m = \log_2 \frac{1}{P_m}$	$P_m L$	$P_m L \log_2 \frac{1}{P_m}$

Total Information

$$I_{total} = P_1 L \log_2 \frac{1}{P_1} + P_2 L \log_2 \frac{1}{P_2} + P_3 L \log_2 \frac{1}{P_3} + --- + P_m L \log_2 \frac{1}{P_m}$$

$$= L\left(P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2} + P_3 \log_2 \frac{1}{P_3} + --- + P_m \log_2 \frac{1}{P_m}\right)$$

$$I_{total} = L \sum_{i=1}^{m} P_i \log_2 \frac{1}{P_i}$$

$$\frac{I_{total}}{L} = \sum_{i=1}^{m} P_i \log_2 \frac{1}{P_i}$$

$$H(x) = \sum_{i=1}^{m} P_i \log_2 \frac{1}{P_i} = -\sum_{i=1}^{m} P_i \log_2 P_i = \frac{-1}{\log_{10} 2} \sum_{i=1}^{m} P_i \log_{10} P_i$$

ENTROPY

Formula

$$
\begin{aligned}
H(X) = E[I(x_i)] &= \sum_{i=1}^{m} P(x_i) I(x_i) \\
&= -\sum_{i=1}^{m} P(x_i) \log_2 P(x_i) \quad \text{b/symbol}
\end{aligned}
$$

Base Conversion formula

$$\log_2 a = \frac{\ln a}{\ln 2} = \frac{\log a}{\log 2}$$

Question

- A discrete source emits five symbols with the probability $P_1=1/2$, $P_2=1/4$, $P_3=1/8$, $P_4=1/16$ and $P_5=1/16$. Find Entropy of source .

$$H(X) = \sum_{j=1}^{m} P_i \log_2 \frac{1}{P_i}$$

$$= \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{16}\log_2 16 + \frac{1}{16}\log_2 16$$

$$= \frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(3) + \frac{1}{16}(4) + \frac{1}{16}(4)$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{1}{4} + \frac{1}{4}$$

$$= \frac{4+4+3+2+2}{8}$$

$$= \frac{15}{8} \text{ bits/symbol}$$

Information Rate

If a discrete source emits symbol at rate **"r"** symbol per sec with average information per symbol (entropy) **"H"** bits per symbols then information rate **"R"** of source will be given by

$$R = r\,H \text{ bits per second}$$

Question

- An event has five possible outcomes with the probability $P_1=1/2$,$P_2=1/4$,$P_3=1/8$,$P_4=1/16$ and $P_5=1/16$. find the rate of information if there are 1000 outcomes per sec.

Answer : H = 15/8 bits/symbol , r = 1000 outcomes per sec so rate of information

$$R = r\,H$$

$$=1000*15/8=1875 \text{ bits per second}$$

Objective for learning

If the rate of information from a source does not exceed the capacity of a given communication channel ,then we can opt a coding technique to get maximum utilization of channel such that the information can be transmitted over the channel with small error ,despite the presence of channel noise.

Summary

- Basic of ITC
- Amount of Information (I)
- Average Information (Entropy) H
- Information rate R