PRINCIPLES OF COMMUNICATIONS

UNIT-5

LECTURE-1

Source Coding

1. Source symbols encoded in binary
2. The average codelength must be reduced
3. Remove redundancy $\Rightarrow$ reduces bit-rate

Consider a discrete memoryless source on the alphabet

$$S = \{s_0, s_1, \cdots, s_k\}$$

Let the corresponding probabilities be

$$\{p_0, p_1, \cdots, p_k\}$$

and codelengths be

$$\{l_0, l_1, \cdots, l_k\}.$$

Then, the average codelength(average number of bits per symbol) of the source is defined as

$$L = \sum_{k=0} p_k l_l$$

If L_{min} is the minimum possible value of L, then the coding efficiency of the source is given by η.

$$\eta = \frac{L_{min}}{L}$$

For an efficient code η approaches unity.

The question: What is smallest average codelength that is possible?

The Answer: Shannon's source coding theorem

Given a discrete memoryless source of entropy H(s), the average codeword length L for any distortionless source encoding scheme is bounded by

$$L \geq H(s)$$

Since, H(s) is the fundamental limit on the average number of bits/symbol, we can say

$$L_{min} \geq H(s)$$

$$\Longrightarrow \eta = \frac{H(s)}{L}$$

Data Compaction:

1. Removal of redundant information prior to transmission.
2. Lossless data compaction - no information is lost.
3. A source code which represents the output of a discrete memoryless source should be uniquely decodable.

Source Coding Schemes for Data Compaction

Prefix Coding

1. The *Prefix Code* is variable length source coding scheme where no code is the prefix of any other code.
2. The prefix code is a uniquely decodable code.
3. But, the converse is not true i.e., all uniquely decodable codes may not be prefix codes.

Table 1: Illustrating the definition of prefix code

Symbol	Prob.of Occurrence	Code I	Code II	Code III
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111

Table 2: Table is reproduced from S.Haykin's book on Communication Systems

From 1 we see that Code I is not a prefix code. Code II is a prefix code. Code III is also uniquely decodable but not a prefix code.

Prefix codes also satisfies Kraft-McMillan inequality which is given by