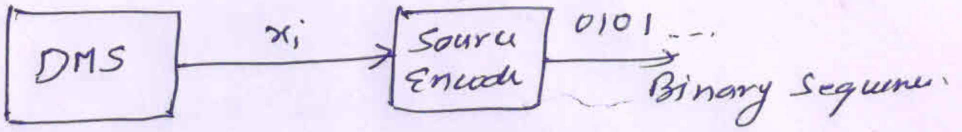
Source Coding

A conversion of the output of a DMS into a sequence of binary code word is called source coding.



$X = \{x_1, x_2 ---x_m\}$

AIM of source coding: Minimize the average bit rate required for representation of the source by reducing the redundancy of the information source.

Key Word

(i) Codeword length: Let X be a DMS with finite entropy H(X) and an alphabet $\{x_1, ---- x_m\}$ with corresponding prob. of occurance $P(x_i)$ ($i = 1, 2, ---m$). Let the binary code word assigned to symbol x_i by the encoder have length n_i measured in bits. The length of a codeword is the number of binary digit in the code word.

(ii) Average Code word length

$$L = \sum_{i=1}^{m} P(x_i)\, n_i$$

(iii) Code Efficiency

$$\eta = \left(\frac{H}{L} \times 100\right) \%$$

(iv) Code Redundancy

$$\gamma = (100 - \eta)\%$$

Classification of Code

x_i	code 1	Code 2	Code 3	code 4	Code 5	code 6
x_1	00	00	0	0	0	1
x_2	01	01	1	10	01	01
x_3	00	10	00	110	011	001
x_4	11	11	11	111	0111	0001

(1) Fixed length code

all symbol have fixed length block binary digit

code 1 and code 2.

(2) Variable length code

code word length is not fixed as all symbols donot have equiprobable.

code 3, code 4, code 5 and code 6.

(3) Prefix Code

No code word is prefix of other code word.

code 4, code 6. code 2

(4) Uniquely Decodeable

if the origional source sequence can be reconstructed perfectly from the encoded binary sequence

* Note all prefix code are uniquely decodeable but vice versa is not true.

Example code 5 is not prefix free but it is uniquely decodable.

Code 3 is not uniquely decodable as 1001 sequence corresponds to source sequence $x_2 x_3 x_2$ or $x_2 x_1 x_1 x_2$

5. Instantaneous Code

if the end of any codeword is recognizable without examing subsequent code symbols.

all prefix free code are instantaneous code.

6. Optimal code.

if $\bar{H} = L$

Kraft Inequality.

let X be a DMS with alphabet $\{x_1, x_2, --- x_m\}$ assigned binary codeword length $\{n_1, n_2, ---- n_m\}$

The necessary and sufficient conditions for the existance of an instantaneous binary code is

$$K = \sum_{i=1}^{m} 2^{-n_i} \leq 1$$

Which is known inequality