Source Coding

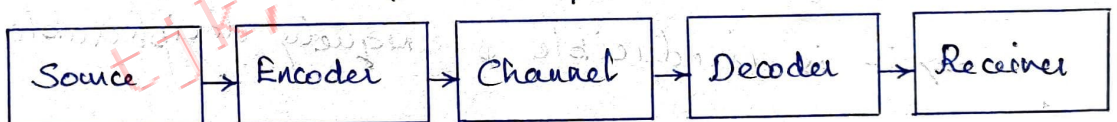**Coding:** Main purpose of coding is to improve the efficiency of the communication system.

Coding is a procedure for mapping a given set of messages $[m_1, m_2 \ldots m_N]$ into a new set of encoded messages $[c_1, c_2 \ldots c_N]$ using a code alphabet set x, in such a way that the transformation is one-to-one, ie for each message, there is only one encoded message. This is called Source Coding.

Advantages:

① It will improve the efficiency of transmission.

② Reduce the probability of error.

| Source | → | Encoder | → | Channel | → | Decoder | → | Receiver |

The msgs are first encoded by the encoder and then transmitted via channel. At the receiver end, the received msgs are first decoded in the decoder and then the original msgs are recovered.

Terminology associated with Coding:

1. **Letter, symbol or code:** Any individual number of the alphabet set.

2. **Message or word:** A finite set of letters.

3. **Length of the word:** The number of letters in a message.

4. **Encoding:** A procedure for associating words from a language to another language in one-to-one manner.

5. **Decoding:** Inverse operation of encoding. Assigning words of the second language to first language.

6. **Uniquely decipherable or uniquely decodable or seperable encoding and decoding:** Able to decode words, when there is no space between the words.

7. **Irreducibility or prefix property:** When no encoded words can be obtained from each other by the addition of more letters.

When a code is irreducible, it is also uniquely decipherable; but reverse is not true.

eg: Let $c_1 = 0$; $c_2 = 10$; $c_3 = 110$.

$$R = 0|110|10|10|000|10|110|110|10$$

$$c_1 \, c_3 \, c_2 \, c_2 \, c_1 c_1 c_2 \, c_3 \, c_3 \, c_2 . \quad \longleftarrow \text{Decoded words.}$$

⟹ it is irreducible & uniquely decipherable.

eg: $c_1 = 0$; $c_2 = 01$; $c_3 = 011$

$$R = 0|0|011|011|01|0|01|01|01|011$$

$$c_1 c_1 \, c_3 \quad c_3 \quad c_2 \, c_1 \, c_2 \, c_2 \, c_2 \, c_3 \longleftarrow \text{Decoded words}$$

⟹ Not irreducible; but uniquely decipherable.

Examples of Codes:

1. **Block Code:**

A block code is a code which maps each of the symbols of the source alphabet 's' onto some "finite sequences" of code symbols from the code alphabet 'x' and each of these finite sequence is called a "code-word".

eg: Source 's' emitting four symbols which are encoded with binary coding

$$S = \{ s_1, s_2, s_3, s_4 \} \quad \text{and} \quad x = \{ 0, 1 \}$$

<u>Block code</u> :

Source symbol	Code-A
s_1	00
s_2	01
s_3	10
s_4	11

The sequences 00, 01, 10 & 11 are called the "code-words" for source symbols s_1, s_2, s_3, s_4 respectively.

2. <u>Non - Singular Code</u> :

A block code is said to be "non-singular" iff all the code words are "distinct" and easily "distinguishable" from one another.

Consider code A ; it is non-singular.

eg: Consider $S = \{ s_1, s_2, s_3, s_4 \}$ with $x = \{ 0, 1 \}$. Let assign the code-words:

Source symbol	Code B
s_1	0
s_2	00
s_3	01
s_4	11

Code B appears non-singular, but not so. Second extension of these code words.

Source symbol	code B'	Source symbol	code B'
$s_1 s_1$	00	$s_3 s_1$	010
$s_1 s_2$	[000]	$s_3 s_2$	0100
$s_1 s_3$	001	$s_3 s_3$	0101
$s_1 s_4$	011	$s_3 s_4$	0111
$s_2 s_1$	[000]	$s_4 s_1$	110
$s_2 s_2$	0000	$s_4 s_2$	1100
$s_2 s_3$	0001	$s_4 s_3$	1101
$s_2 s_4$	0011	$s_4 s_4$	1111