IIIT Hyderabad **Scribed By: D. Shritishma Reddy(20161165)**
Introduction to Coding Theory (EC5.205) **Lecture #1**
Instructor: Dr. Prasad Krishnan **03.01.2020**

Keywords. *Channel Coding. Source Coding. Information Theory. Types of Channel Models.*

1 Introduction

We assume that the data we deal with is primarily in digital form. Data is majorly processed in two ways

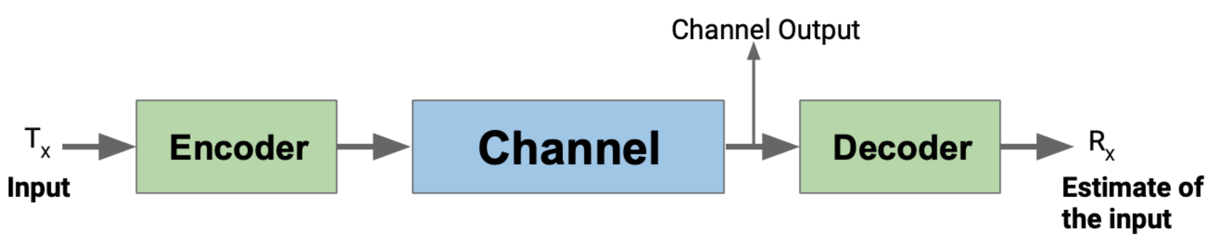
1. Data Storage which includes Data Retrieval
2. Data Communication which includes Data Reception

2 Block Diagram of Channel Coding

Basic Block Diagram of Communication

Communicating data from one location to another requires some form of pathway or medium. These pathways, called communication channels. They are often not perfect. We modify the characteristics of the channel to suit our needs according to our problem requirements.

The Block Diagram of Channel Coding is shown below. It assumes coding for "reliable" communication.



Basic Block Diagram of Channel Coding

- T_x is input data transmitted
- R_x is the estimate of the input data

The following are the desired design characteristics of the above block diagram

- Low decoding complexity
- High rate of Communication of the channel
- Low probability of error of R_x
- An additional power or area/size constraints might be present depending on the application such as satellite communications etc.

3 Source Coding

The use of variable-length codes for the minimum lossless representation of the information, for a given size of alphabet is called source coding. Source Coding exploits the inherent redundancy in information.

Source Coding is done by understanding the statistical properties of data which helps to reducing redundancy of pictures.

Huffman Coding Algorithm is an example for a Source Coding Algorithm. It requires the probability distribution of the input alphabet.

According to Shannon, the entropy value of a piece of information provides an absolute limit on the shortest possible average length of a message, or how much it can be compressed, without losing information as it is transmitted. Thus, the absolute limits of communication is given by Entropy.

The measure of information entropy associated with each possible data value is the negative logarithm of the probability mass function for the value:

$$S = -\sum_i P_i \log P_i = -\mathrm{E}_P[\log P], where$$

$$\mathrm{E}_P[X] = \sum_i P_i X_i$$

is the expectation defined by the probability P.

4 Types of Channel Models

4.1 Discrete to Discrete Channel

Let X be some discrete (typically finite) set which is the input alphabet. Let Y be some discrete (typically finite) set which is the output alphabet. Characteristics of a deterministic channel:

- No randomness in the channel
- Determinable and Precise [y=f(x)]
- No coding problem for this channel

Discrete-Discrete Channel

For a non-deterministic channel:

- we use **Channel Conditional Probability Distribution** given by P(Y| X).
- Is is essentially a Probability Mass Function [PMF].
- For every x ϵ X, we have a distribution on Y, P(output=y|input=x).

Example: Storage Channels such as Hard Disk for which both the input and output data are bits (discrete). Bits are stored into the Hard Disk while writing(input) into it and bits are read(output) from it when required.

4.2 Discrete to Continuous Channel

Let X be some discrete (typically finite) set which is the input alphabet. Let Y be some continuous set which is the output alphabet. For a non-deterministic channel:

- We use **Channel Conditional Probability Distribution** that is essentially a Probability Density Function[PDF].
- It is given by p(y|input=x), for every x ϵ X.

Discrete-Continuous Channel

Example: For any digital space communication system

- Input is On/Off or High/Low voltage signal. (Square Waves)
- Output is most commonly a distorted irregular wave. If this is randomly sampled, the output will belong to R and not specifically belong to {High,Low} or {On,Off}

4.3 Continuous to Continuous Channel

Let X be some continuous set which is the input alphabet.Let Y be some continuous set which is the output alphabet.

Continuous-Continuous Channel

For a non-deterministic channel:

- We use **Channel Conditional Probability Distribution** that is essentially a Probability Density Function[PDF].
- It is given by p(y|x).

In real life we run experiments to obtain the values of PMF's and PDF's.

References

[1] Lecture Notes