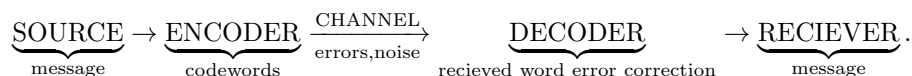


Introduction

We model communication:



Examples: optical signals, electrical telegraph, SMS (compression), postcodes, CDs (error correction), zip/gz files (compression).

Given a source and a channel, modelled probabilistically, the basic problem is to design an encoder and decoder to transmit messages economically (noiseless coding; compression) and reliably (noisy coding).

Examples:

- Noiseless coding: Morse code: common letters are assigned shorter code-words, e.g $A \mapsto \bullet-$, $E \mapsto \bullet$, $Q \mapsto --\bullet-$, $S \mapsto \bullet\bullet\bullet$, $O \mapsto --$, $Z \mapsto --\bullet\bullet$. Noiseless coding is adapted to source.
- Noisy coding: Every book has an ISBN $a_1, a_2, \dots, a_9, a_{10}$, $a_i \in \{0, 1, \dots, 9\}$ for $1 \leq i \leq 9$ and $a_{10} \in \{0, 1, \dots, 9, X\}$ with $\sum_{j=1}^{10} ja_j \equiv 0 \pmod{11}$. This detects common errors - e.g one incorrect digit, transposition of two digits. Noisy coding is adapted to the channel.

Plan:

- (I) Noiseless coding - entropy
- (II) Error correcting codes - noisy channels
- (III) Information theory - Shannon's theorems
- (IV) Examples of codes
- (V) Cryptography

Books: [GP], [W], [CT], [TW], Buchmann, Körner. Online notes: Carne, Körner.

Basic Definitions

Definition (Communication channel). A *communication channel* accepts symbols from a alphabet $\mathcal{A} = \{a_1, \dots, a_r\}$ and it outputs symbols from alphabet $\mathcal{B} = \{b_1, \dots, b_s\}$. Channel modelled by the probabilities $\mathbb{P}(y_1 \dots y_n \text{ recieved} | x_1 \dots x_n \text{ sent})$. A *discrete memoryless channel* (DMC) is a channel with

$$p_{ij} = \mathbb{P}(b_j \text{ recieved} | a_i \text{ sent})$$

the same for each channel use and independent of all past and future uses. The channel matrix is $P = (b_{ij})$, a $r \times s$ stochastic matrix.

Definition (Binary symmetric channel). The *binary symmetric channel* (BSC) with error probability $p \in [0, 1)$ from $\mathcal{A} = \mathcal{B} = \{0, 1\}$. The channel matrix is

$$\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}.$$

A symbol is transmitted correctly with probability $1 - p$. Usually assume $p < 1/2$.

The *binary erasure channel* (BEC) has $\mathcal{A} = \{0, 1\}$, $\mathcal{B} = \{0, 1, *\}$. The channel matrix is

$$\begin{pmatrix} 1-p & 0 & p \\ 0 & 1-p & p \end{pmatrix}.$$

So $p = \mathbb{P}(\text{symbol can't be read})$.

Definition. We model n uses of a channel by the n th extension, with input alphabet \mathcal{A}^n and output alphabet \mathcal{B}^n . A *code* C of length n is a function $\mathcal{M} \rightarrow \mathcal{A}^n$ where \mathcal{M} is the set of possible messages. Implicitly we also have a decoding rule $\mathcal{B}^n \rightarrow \mathcal{M}$. The *size* of C is $m = |\mathcal{M}|$. The *information rate* is $\rho(C) = \frac{1}{n} \log_2 m$. The *error rate* is $\hat{e}(C) = \max_{x \in \mathcal{M}} \mathbb{P}(\text{error} | x \text{ sent})$.

Remark. For the remainder of the course we write \log instead of \log_2 .

Definition. A channel can *transmit reliably at rate* R if there exists $(C_n)_{n=1}^\infty$ with each C_n a code of length n such that

$$\lim_{n \rightarrow \infty} \rho(C_n) = R \text{ \& } \lim_{n \rightarrow \infty} \hat{e}(C_n) = 0.$$

The *capacity* is the supremum of all reliable transmission rates. We'll see in Chapter 9 that a BSC with error probability $p < 1/2$ has non-zero capacity.

1 Noiseless coding

1.1 Prefix-free codes

For an alphabet \mathcal{A} , $|\mathcal{A}| < \infty$, let $\mathcal{A}^* = \bigcup_{n \geq 0} \mathcal{A}^n$, the set of all finite strings from \mathcal{A} . The *concatenation* of strings $x = x_1 \dots x_r$ and $y = y_1 \dots y_s$ is $xy = x_1 \dots x_r y_1 \dots y_s$.

Definition. Let \mathcal{A}, \mathcal{B} be alphabets. A code is a function $c : \mathcal{A} \rightarrow \mathcal{B}^*$. The strings $c(a)$ for $a \in \mathcal{A}$ are called *codewords* or *words* (CWS).