

Introduction

Coding theory is the branch of mathematics concerned with transmitting data across noisy channels and recovering the message.

It is about making messages easy to read: don't confuse it with *cryptology* which is the art of making messages hard to read!

Introduction

- We assume that our message is in the form of binary digits or bits, strings of 0 or 1.
- We have to transmit these bits along a channel (such as a telephone line) in which errors occur randomly, but at a predictable overall rate.
- Occasionally, noise on the channel, perhaps in the form of atmospheric disturbances or hardware malfunctions, causes the 0 to be received as a 1.

Error-detecting and error-correcting codes

Yotsanan Meemark

Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok, Thailand

<http://pioneer.netserv.chula.ac.th/~myotsana/>

Abstract Algebra I

Introduction

When photographs are transmitted to Earth from deep space, error-detecting and error-correcting codes are used to guard against the noise caused by lightning and other atmospheric interruptions.

Compact discs (CDs) use error-detecting and error-correcting codes so that a CD player can read data from a CD even if it has been corrupted by noise in the form of imperfections on the CD.

Introduction

- We would like to develop ways to combat these errors that can occur during data transmission.
- To compensate for the errors we need to transmit **more bits** than there are in the original message.

Shannon's Communication Channel

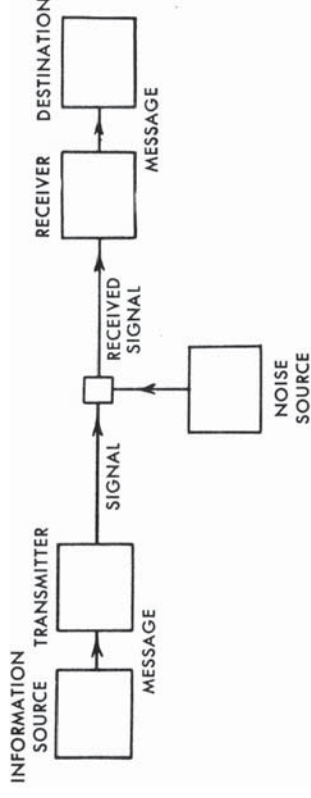


Fig. 1. — Schematic diagram of a general communication system.

Example (ISBN-10)

The International Standard Book Number (ISBN-10) Code is used throughout the world by publishers to identify properties of each book. The first nine digits (bounded between 0 and 9, inclusive) of each ISBN represent information about the book including its language, publisher, and title. In order to guard against errors, the nine-digit "message" is encoded as a ten-digit codeword. The appended tenth digit is a 'check digit' chosen so that the whole ten-digit string $x_1x_2 \dots x_{10}$ satisfies

$$10x_1 + 9x_2 + 8x_3 + 7x_4 + 6x_5 + 5x_6 + 4x_7 + 3x_8 + 2x_9 + x_{10} \equiv 0 \pmod{11}.$$

If x_{10} should be equal to 10, an 'X' is used.

Example (ISBN-10)

E.g., (a) 3-540-26596- x_{10}

$$\begin{aligned} 10(3) + 9(5) + 8(4) + 7(0) + 6(2) + 5(6) + 4(5) + 3(9) + 2(6) + x_{10} \\ \equiv 208 + x_{10} \equiv 0 \pmod{11}. \end{aligned}$$

$$\text{so } x_{10} = 1.$$

(b) Check if 0-495-83253-0 is a correct ISBN-10.

Example (ISBN-10)

- The two most common errors in handling an ISBN-10 (e.g., typing or writing it) are a single altered digit or the transposition of adjacent digits.
- The ISBN check digit method therefore ensures that it will always be possible to detect these two most common types of error.
- Since there is no way of telling where the error is, the code is not error-correcting.
- In contrast, it is possible for other types of error, such as two altered non-transposed digits, or three altered digits, to result in a valid ISBN number (although it is still unlikely).

Binary codes

- We shall consider from now on information which is stored or transmitted in binary form.
- English texts may be converted by replacing each letter, numeral, space or punctuation mark by a suitable binary-based code (such as ASCII) for it.
- We shall think of the set $\{0, 1\}$ as coming equipped with the operations: addition and multiplication modulo 2.
- It is customary in this context to write \mathbb{B} instead of \mathbb{Z}_2 .

Binary codes

- A **word of length n** is a string of n binary digits.
- We shall think of words of length n as members of \mathbb{B}^n , the Cartesian product of n copies of the binary set \mathbb{B} regarded as an abelian group under addition.
- That is, $\mathbb{B}^n = \{x_1 x_2 \dots x_n : x_1, x_2, \dots, x_n \in \{0, 1\}\}$ for $n \in \mathbb{N}$.
- E.g., 0001, 1110 and 0000 are words of length 4.

Codewords

- Suppose that our original messages are composed of words of length m .
- We choose a *code function* $f : \mathbb{B}^m \rightarrow \mathbb{B}^n$ and, instead of sending a word w , we send the word $f(w)$.
- Any word of length n in the image of f is called a **codeword**. Thus,
$$\mathcal{C} = \text{im } f = \{f(w) : w \in \mathbb{B}^m\}$$
is the set of all codewords.

Parity-check codes

Define $f : \mathbb{B}^m \rightarrow \mathbb{B}^{m+1}$ by

$$f(w) = wx,$$

where the **parity-check digit** of w is

$$x = \begin{cases} 0 & \text{if the number of 1's in } w \text{ is even,} \\ 1 & \text{if the number of 1's in } w \text{ is odd.} \end{cases}$$

E.g., take $m = 3$: we have 8 words in \mathbb{B}^3 and under each is its image under f .

000	001	010	011	100	101	110	111
0000	0011	0101	0110	1001	1010	1100	1111

Parity-check codes

- This code enables us to any single error in the transmission of a codeword since, if a single digit is changed, the word received will have an odd number 1 and so not be a codeword. E.g., 1101 is not a codeword.
- In fact any odd number of errors will be detected, but an even number of errors will fail to be detected.
- However, this code does not allow one to correct an error without re-transmission of the word.

Repetition codes

Define $f : \mathbb{B}^m \rightarrow \mathbb{B}^{3m}$ by $f(w) = www$, i.e., the word repeated three times.

E.g. if $m = 6$ and $w = 101111$, then $f(w) = 10111110111111011111$

- This code will detect any single error of any two errors.
- If the above $f(w)$ is received as 1001111010111101111 (two errors) then we can detect some error has occurred in transmission since the received message is not a six-letter word three times repeated.

Repetition codes

- However, this code does not necessarily detect three errors. If $f(w)$ above were received as 001111001111001111 (three errors) then it looks as if the original word was 001111, where as w was 101111.
- Although, two errors can be detected, this code can correct only one error.
- If $f(w)$ were received as 101011101111101111 then we would consider it *most likely* that the original word was 101111, not 101011.

Repetition codes

Let us be more specific.

- Suppose the www is the word sent and m is the word received.
- Breaking m into three block abc where a, b and c are word of length 6.
- If no errors have been made in transmission then $a = b = c$.
- If one error occurs, then two of a, b and c are equal to each other (so, necessarily, to w), so we correct the message and conclude (correctly) that the original word is w .
- If two errors have been made then it could happen that $abc = w'ww'$ (say), so we would conclude (incorrectly) that the original word was w' .

Repetition codes

- The disadvantage of the repetition scheme is that it multiplies the number of bits transmitted by a factor which may prove unacceptably high.
- In 1948, Claude Shannon, working at Bell Laboratories in the USA, inaugurated the whole subject of coding theory by showing that it was possible to encode messages in such a way that the number of extra bits transmitted was as small as possible. Unfortunately his proof in 'A Mathematical Theory of Communication' did not give any explicit recipes for these optimal codes.
- Two years later, Hamming, also at Bell Labs, published details of his work on explicit error-correcting codes with information transmission rates more efficient than simple repetition.

Weight and distance

- The **weight** of a binary word w , $\text{wt}(w)$, is defined to be the number of 1s in its binary expression.
- The **distance** between binary words v, w of the same length is defined to be the weight of their difference:

$$d(v, w) = \text{wt}(v - w) = \text{wt}(v + w) \text{ (because } 1 = -1 \text{ in } \mathbb{B}).$$

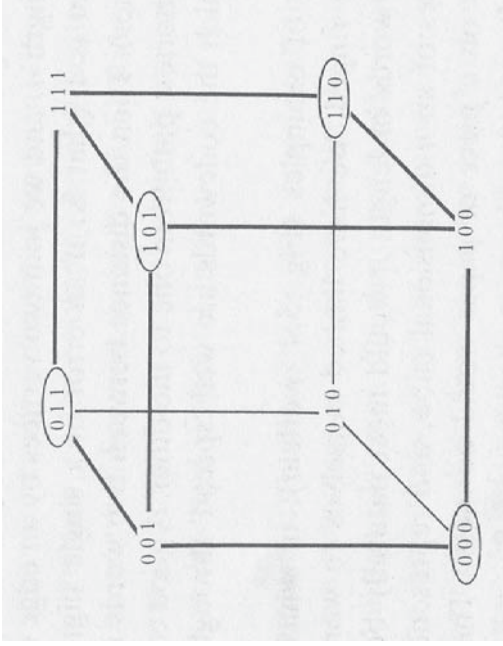
- E.g., $\text{wt}(001101) = 3$, $\text{wt}(000000) = 0$ and $d(010110, 100000) = \text{wt}(110110) = 4$.
- Note that the distance between v and w is simply the number of places at which they differ.
- $d(v, w) = d(w, v)$, $d(u, v) \geq 0$ and $(d(u, v) = 0 \Leftrightarrow u = v)$.

Weight and distance

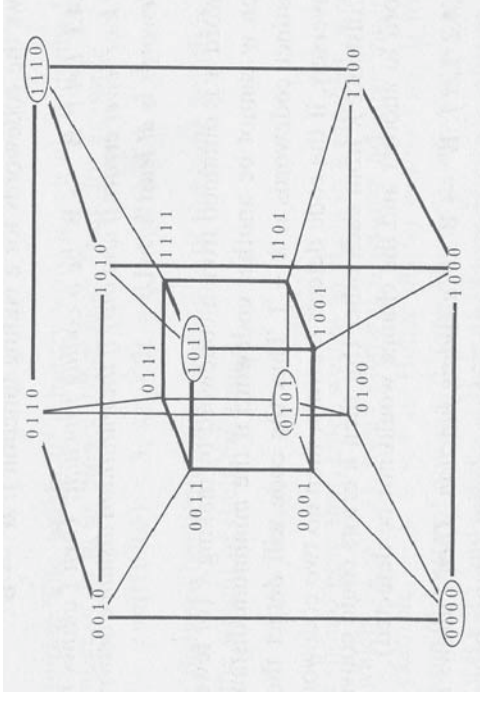
- $d(010101, 111100) =$
- $d(1111, 0101) =$
- $d(w, w) =$
- $d(w, \vec{0}) =$

If w is a word that is transmitted and is received as v then the net number of errors that have occurred in transmission is the distance $d(v, w)$. 'Therefore, a good coding function $f : \mathbb{B}^m \rightarrow \mathbb{B}^n$ will be one that *maximizes* the distance between codewords'.

Parity-check codewords in \mathbb{B}^3 and distance



Parity-check codewords in \mathbb{B}^4 and distance



Let $f : \mathbb{B}^m \rightarrow \mathbb{B}^n$ be a coding function and $\mathcal{C} = \text{im } f$.

The **minimum distance between distinct codewords** is

$$d(\mathcal{C}) = \min\{d(u, v) : u, v \in \mathcal{C} \text{ and } u \neq v\}.$$

Theorem

Let $f : \mathbb{B}^m \rightarrow \mathbb{B}^n$ be a coding function.

Then f allows the detection of $\leq k$ errors \iff the minimum distance between distinct codewords is $\geq k + 1$.

Theorem

Let $f : \mathbb{B}^m \rightarrow \mathbb{B}^n$ be a coding function.

Then f allows the correction of $\leq k$ errors \iff the minimum distance between distinct codewords is $\geq 2k + 1$.

Example

Define the coding function $f : \mathbb{B}^4 \rightarrow \mathbb{B}^9$ by $f(w) = ww\bar{x}$ where x is the parity-check digit of w . We can list all w and $f(w)$ as follows.

w	$f(w)$
0000	000000000
0010	001000101
0100	0100001001
0110	0110001100
1000	100010001
1010	101010100
1100	110011000
1110	111011101

We may check, by computing $d(u, v)$ for all $u \neq v$, that the minimum distance between codewords is 3 and so the code can detect up to two errors and can correct any single error.