CS 5319
Advanced Discrete Structure

Lecture 14:
Introduction to Group Theory IV

Outline

- Introduction
- Groups and Subgroups
- Generators
- Cosets (and Lagrange’s Theorem)
- Permutation Group (and Burnside’s Theorem)
- Group Codes

Codes

- A sequence of symbols is called a word
- The coding problem is to represent distinct message using distinct words
  - In particular, in digital communication, the words are formed by 0 or 1
- A code is the set of words used for the distinct messages in a certain scheme
  - These words are called codewords

Codes

- There are a few reasons for designing codes

1. To save storage or communication time

   Ex : Huffman, run-length, arithmetic, …

2. To detect error in transmission

   Ex : Repetition, parity, checksum, …

3. To correct error in transmission

   Ex : Golay, Reed-Solomon, Reed-Muller, group, convolution, …

Block Codes

- A block code is a code whose codewords all have the same length
  - One reason for having block codes is for error detection or error correction
- Let A = the set of all length-n binary words

  $\oplus$ = binary XOR operation on A

  Ex : $10111 \oplus 00101 = 10010$
- Note that $(A, \oplus)$ is a group

Block Codes

- For any x in A, we define the weight of x, denoted by $w(x)$, to be # of 1's in x

  Ex : $w(10111) = 4$, $w(00101) = 2$

- The distance between any x and y in A, denoted by $d(x, y)$, is the weight of $x \oplus y$
  - That is, $d(x, y) = w(x \oplus y)$

    ➔ $d(x, y) = d(y, x)$
  - Also, $d(x, y)$ = # bits that x and y are different

Block Codes

Lemma 1 (Triangle Inequality) :

> For any x, y, and z in A,
>
> $$d(x, z) \le d(x, y) + d(y, z)$$

Proof :

First, $w(u \oplus v) \le w(u) + w(v)$ [why?]

$$\Rightarrow \quad w(x \oplus z) = w(x \oplus y \oplus y \oplus z)$$
$$\le w(x \oplus y) + w(y \oplus z)$$

Thus, $d(x, z) \le d(x, y) + d(y, z)$

Block Codes

- We now define the distance of a code, which is closely related to its error-correction power

> Definition :
>
> The distance of a code G is the minimum distance between any two words in G

- Now, suppose a word y is received from sender
  - If y is in G, we assume y is the word sent
  - If y is not in G, we assume the word x in G, with $d(x, y)$ minimized, is the word sent

Block Codes

- The previous decoding method is called the minimum-distance decoding criterion

Theorem 1 :

> A code of distance $2t + 1$ can correct t or fewer transmission errors if we use minimum-distance decoding criterion

Block Codes

Proof : Suppose there are at most t errors

Let x = codeword sent, y = word received

z = codeword not equal to x

Then we have: $d(x, y) \leq t$

Also, we have: $d(x, z) \geq 2t + 1$

so that $d(x, y) + d(y, z) \geq 2t + 1$

➔ For any $z \neq x$, $d(y, z) \geq t + 1$

➔ y is decoded correctly without ambiguity

Group Codes

- We now study a class of block codes called group codes
- A subset G of A is called a group code if $(G, \oplus)$ is a subgroup of $(A, \oplus)$
  - A = set of all length-n binary words

Ex :

{ 0000, 0011, 1101, 1110 } is a group code

{ 00000, 01110 } is also a group code

Group Codes

- The distance of a group code can be easily determined based on the following theorem

Theorem 2 :

> Let G be a group code, and x be the minimum-weight non-zero codeword in G. Then,
>
> distance of G is equal to $w(x)$

Group Codes

Proof :

First, since **0** (the word with all 0's) is in G,

$$w(x) = d(x, \mathbf{0}) \geq \text{distance of } G$$

However, for any y and z in G,

$$d(y, z) = w(y \oplus z) \geq w(x) \qquad \text{[why?]}$$

➔ distance of $G = \min_{y,z} d(y, z) \geq w(x)$

Thus, distance of $G = w(x)$

Group Codes

- For group codes, if we use minimum-distance decoding criterion, there is an easy way to find the codeword corresponding to the received word
- Firstly, suppose word y is received
  - Corresponding codeword x in G should be one that minimizes $w(x \oplus y)$
  - Equivalently, if e is the word with smallest weight in the coset $G \oplus y$

  ➔ $e = y \oplus x$, or $x = y \oplus e$

Group Codes

- Now, suppose that
  - for each coset C of G, we remember its smallest weight word e_C
  - For each y, we remember which coset $C(y)$ that y belongs to

➔ Then upon receiving y, we can decode it back to the transmitted codeword x by :

$$x = y \oplus e_{C(y)}$$

Ex : Suppose we use the group code

$$G = \{\ 00000,\ 00111,\ 11010,\ 11101\ \}$$

whose distance = 3 (thus can correct 1 error)

- Cosets of G :

$$\left.\begin{aligned}
G \oplus 00000 &= \{\ 00000,\ 00111,\ 11010,\ 11101\ \} \\
G \oplus 00001 &= \{\ 00001,\ 00110,\ 11011,\ 11100\ \} \\
G \oplus 00010 &= \{\ 00010,\ 00101,\ 11000,\ 11111\ \} \\
G \oplus 00100 &= \{\ 00100,\ 00011,\ 11110,\ 11001\ \} \\
G \oplus 01000 &= \{\ 01000,\ 01111,\ 10010,\ 10101\ \} \\
G \oplus 10000 &= \{\ 10000,\ 10111,\ 01010,\ 01101\ \}
\end{aligned}\right\}\ \boxed{\text{Decoding is unique}}$$

$$\left.\begin{aligned}
G \oplus 10001 &= \{\ 10001,\ 10110,\ 01011,\ 01100\ \} \\
G \oplus 10100 &= \{\ 10100,\ 10011,\ 01110,\ 01001\ \}
\end{aligned}\right\}\ \boxed{\text{Decoding not unique}}$$