Groups and Coding（群与编码）

Zhang Yanmei

ymzhang@bupt.edu.cn

College of Computer Science & Technology

Beijing University of Posts & Telecommunications

Content

- Concept
- Encoding function（编码函数） encoding function
  - (m,n) encoding function
  - parity(m,m+1) check code
  - (m,3m) encoding function
- Error detection（差错检测）
  - Hamming distance（海明距离） Hamming distance
  - Properties of the distance function
  - Minimal distance of an encoding function
  - Theorem 2
- Group codes（群码）
  - Definition
  - Theorem 3

Coding theory

- In today's modern world of communication, data items are constantly being transmitted from point to point.

- The basic problem in transmission of data is that of receiving the data as sent and not receiving a distorted （失真） piece of data.

Unit of information

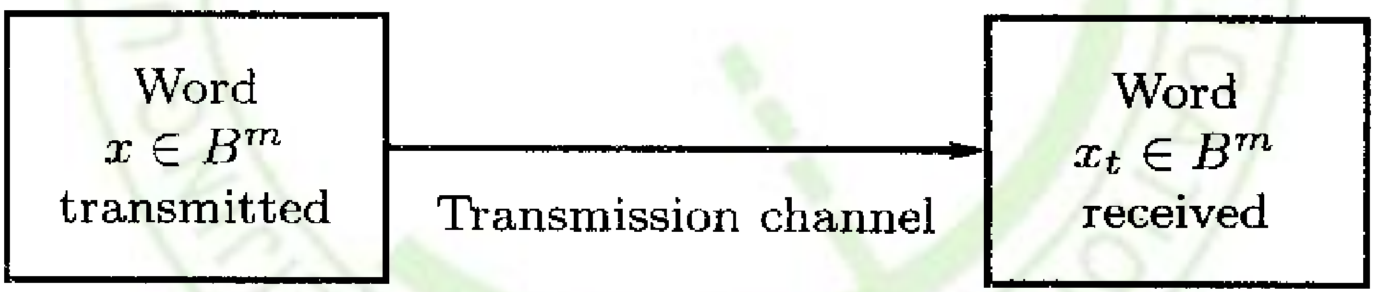
- *Message*(信息) is a finite sequence of characters from a finite alphabet $B = \{0, 1\}$
- *Word*(码字) is a sequence of m 0's and 1's.

Groups B^m

- The set B is a group under the binary operation + (mod 2 addition)
- It follows that $B^m = B \times B \times \ldots \times B$ (m factors) is a group under the operator $\oplus$ defined by
  - $(x_1, x_2, \ldots, x_m) \oplus (y_1, y_2, \ldots, y_m)$
    $= (x_1 + y_1, x_2 + y_2, \ldots, x_m + y_m)$
- An element in B^m will be written as $(b_1, b_2, \ldots, b_m)$ or more simply as $b_1 b_2 \ldots b_m$

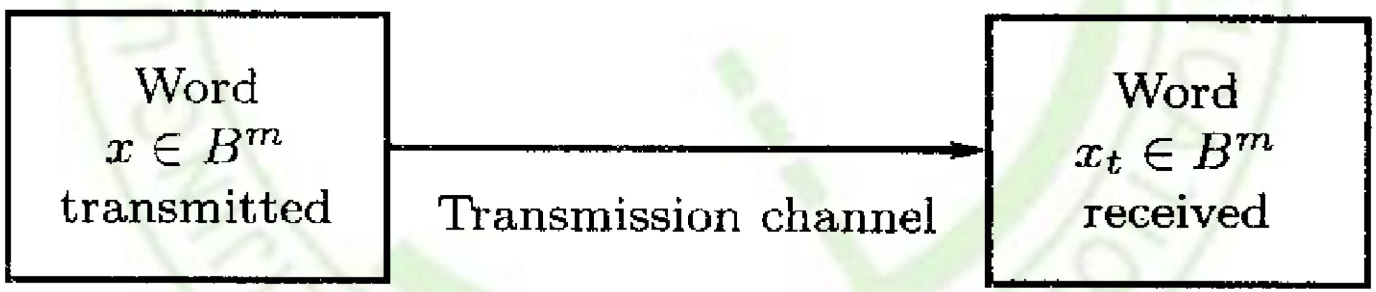
Transmission channel and Noise （传输信道与噪音）

- An element $x \in B^m$ is sent through the transmission channel and is received as an element $x_t \in B^m$.



Transmission channel and Noise

- *Noise*(噪声) in the transmission channel may cause a 0 to be received as a l, or vice versa, lead $x \neq x_t$



对 错

Coding theory

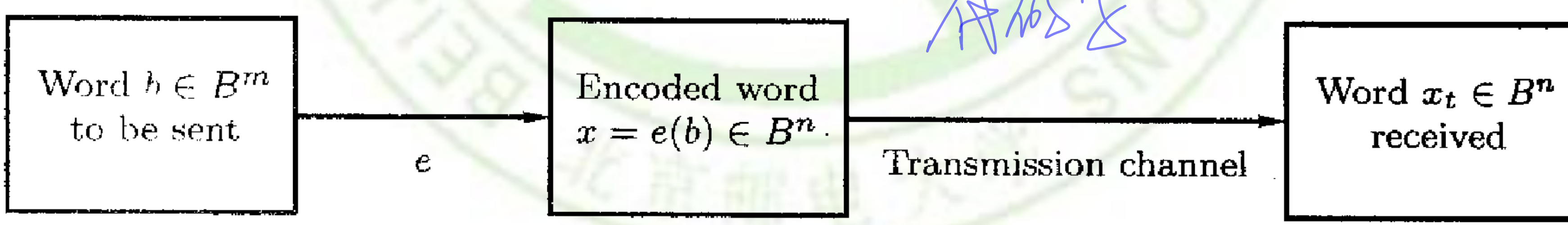
- Coding theory has developed techniques for introducing redundant information（引入冗余信息） in transmitted data that help in detecting, and sometimes in correcting, errors.
- Some of these techniques make use of group theory.

引入冗…

Encoding function - 编码函数

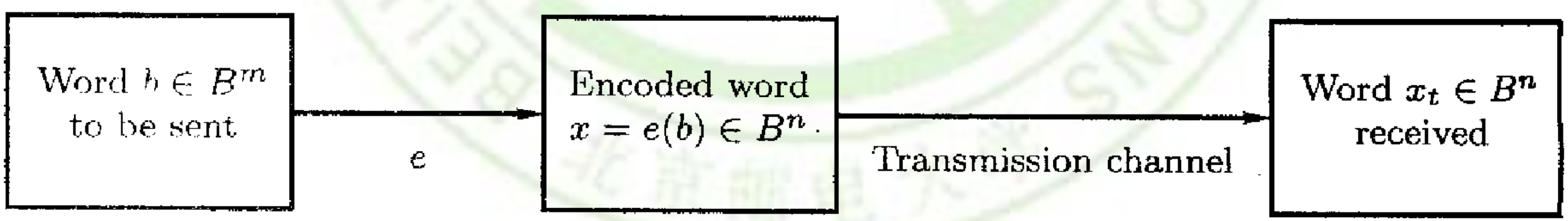
- Choose: an integer $n > m$ and a one-to-one function $e: B^m \to B^n$
  编码函数
  - e is called an (m, n) *encoding function*, representing every word in B^m as a word in B^n.
  - If $b \in B^m$, then $e(b)$ is called the *code word*(码字) representing b.

码字



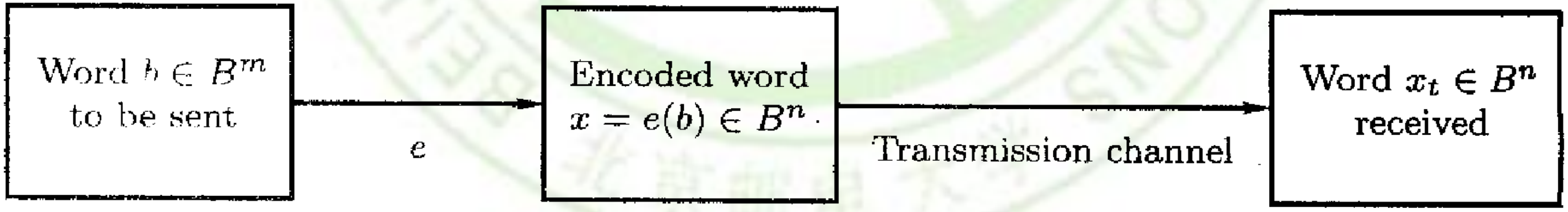
Encoding function

- If the transmission channel is noiseless, then $x_t = x$ for all x in B^n.
- In this case $x = e(b)$ is received for each $b \in B^m$, and since e is a known function, b may be identified.



Encoding function

- In general, errors in transmission do occur.
- We will say that the code word $x = e(b)$ has been transmitted with *k or fewer errors* if x and x_t differ in at least 1 but no more than k positions.



Error detect – 差错检测

- Let $e: B^m \to B^n$ be an (m, n) encoding function.
- e *detects k or fewer errors* if whenever $x = e(b)$ is transmitted with k or fewer errors, then x_t is not a code word (thus x_t could not be x and therefore could not have been correctly transmitted).

Definition (weight)

- For $x \in B^n$, the number of l's in x is called the weight(权) of x and is denoted by $|x|$.
- Find the weight of each of the fol1owing words in B^5. weight
  - x=01000 1
  - x=11100 3
  - x=00000 0
  - x=11111 5

Example (parity check code)奇偶校验码

- The following encoding function e: $B^m \to B^{m+1}$ is called the *parity (m, m+1) check code*:
- If $b = b_1b_2\ldots b_m \in B^m$, define

$$e(b) = b_1b_2\ldots b_m b_{m+1}$$

- where

$$b_{m+1} = \begin{cases} 0 & \text{if } |b| \text{ is even} \quad \text{偶} \\ 1 & \text{if } |b| \text{ is odd.} \quad \text{奇} \end{cases}$$

Example (parity check code)

- Let $m = 3$. Then

$$e(000) = 0000$$
$$e(001) = 0011$$
$$e(010) = 0101$$
$$e(011) = 0110$$
$$e(100) = 1001$$
$$e(101) = 1010$$
$$e(110) = 1100$$
$$e(111) = 1111$$

$x = e(b) = (11)$

- Suppose now that $b = 111$. Then $x = e(b) = 1111$

Example (3m encoding function)

- Consider the $(m, 3m)$ encoding function e: $B^m \to B^{3m}$.
- If $b = b_1 b_2 \ldots b_m \in B^m$, define

$$e(b) = b_1 b_2 \ldots b_m b_1 b_2 \ldots b_m b_1 b_2 \ldots b_m$$

$$\left.\begin{array}{l} e(000) = 000000000 \\ e(001) = 001001001 \\ e(010) = 010010010 \\ e(011) = 011011011 \\ \hline e(100) = 100100100 \\ e(101) = 101101101 \\ e(110) = 110110110 \\ e(111) = 111111111 \end{array}\right\} \text{code word}$$

码字

Example (3m encoding function)

- Suppose now that $b = 011$
  - Then $e(011) = 011011011$.
- Assume now we receive the word 011111011. This is not a code word, so we have detected the error.

Hamming distance(海明距离)

- Let x and y be words in B^n. The *Hamming distance* $\delta(x, y)$ between x and y is the weight, $|x \oplus y|$, of $x \oplus y$.
- The distance between $x = x_1x_2\ldots x_n$ and $y = y_1y_2\ldots y_n$ is the number of various of i such that $x_i \neq y_i$, that is, the number of positions in which x and y differ.
- Using the weight of $x \oplus y$ is a convenient way to count the number of different positions.

Example

- Find the distance between x and y:
  - (a) $x = 110110$, $y = 000101$
  - (b) $x = 001100$, $y = 010110$.
- Solution
  - (a) $x \oplus y = 110011$, so $| x \oplus y | = 4$
  - (b) $x \oplus y = 011010$, so $| x \oplus y | = 3$

Theorem(properties of distance function)

- Let x, y, and z be elements of B^n. Then
  - (a) $\delta(x, y) = \delta(y, x)$
  - (b) $\delta(x, y) \geq 0$
  - (c) $\delta(x, y) = 0$ if and only if $x = y$
  - (d) $\delta(x, y) \leq \delta(x, z) + \delta(z, y)$

$\delta(x, y) \leq \delta(x, z) + \delta(z, y)$

$\delta(x, y) = 0$, if and only if $x = y$

Theorem (properties of distance function)

- Let x, y, and z be elements of B^n. Then
  - (d) $\delta(x, y) \leq \delta(x, z) + \delta(z, y)$

- Proof of (d)
  - $|x \oplus y| \leq |x|+|y|$; $a \oplus a = \mathbf{0}$
  - $\delta(x, y) = |x \oplus y| = |x \oplus \mathbf{0} \oplus y|$

$$= |x \oplus z \oplus z \oplus y|$$

$$\leq |x \oplus z|+|z \oplus y|$$

minimum distance 最小距离

Minimum distance（最小距离）

- The *minimum distance* of an encoding function $e: B^m \to B^n$ is the minimum of the distances between all distinct pairs of code words; that is,

$$\min\{\delta(e(x), e(y)) \mid x, y \in B^m\}$$

Example 5

- Consider the following (2, 5) encoding function e:

$$\left.\begin{aligned} e(00) &= 00000 \\ e(01) &= 00111 \\ e(10) &= 01110 \\ e(11) &= 11111 \end{aligned}\right\} \text{code word}$$

- Minimum distance?

Theorem 2

- An (m, n) encoding function $e: B^m \to B^n$ can detect k or fewer errors
  - if and only if
- its minimum distance is at 1east $k + 1$.

Proof

$\Leftarrow$ the minimum distance is at least $k + 1$

- Let $b \in B^m$, and let $x = e(b) \in B^n$ be the code word representing b.
  - x is transmitted and is received as x_t. If x_t were a code word different from x, then $\delta(x, x_t) \geq k+1$, so x would be transmitted with $k + 1$ or more errors.
  - Thus, if x is transmitted with k or fewer errors, then x_t cannot be a code word.
  - This means that e can detect k or fewer errors.

Proof

e can detect k or fewer errors $\Rightarrow$

- Suppose that the minimum distance between code words is $r \leq k$
- Let x and y be code words with $\delta(x, y) = r$.
  - If $x \,{}_{t}\!= y$, that is, if x is transmitted and is mistakenly received as y, then $r \leq k$ errors have been committed and have not been detected.
- Thus it contradict with e can detect k or fewer errors.

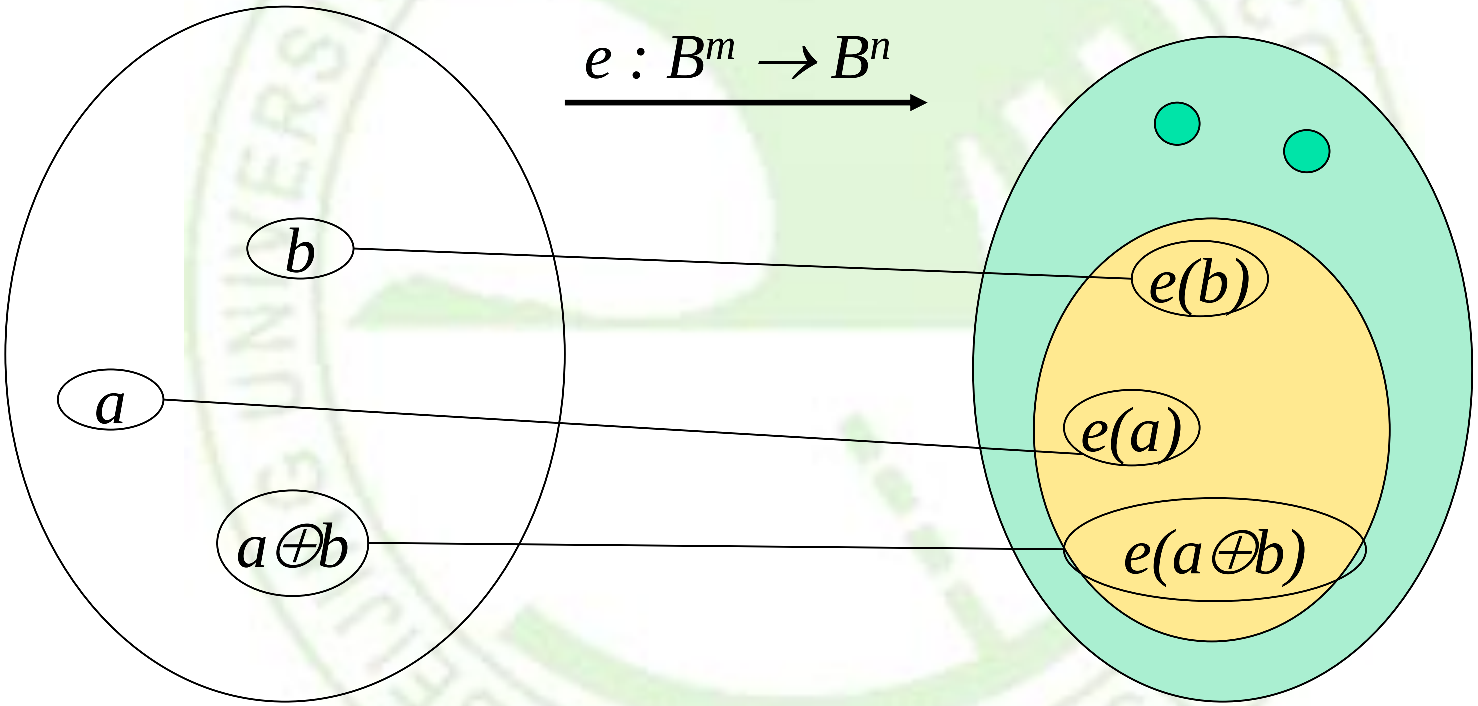
- Q.E.D.

Example 6
How many errors will *e* detect?

- Consider the (3, 8) encoding function $e: B^3 \to B^8$ defined by

$$
\left.
\begin{aligned}
e(000) &= 00000000 \\
e(001) &= 10011100 \\
e(010) &= 00101101 \\
e(011) &= 10010101 \\
e(100) &= 10100100 \\
e(101) &= 10001001 \\
e(110) &= 00011100 \\
e(111) &= 00110001
\end{aligned}
\right\} \text{code word}
$$

Encoding function



e is a one-to-one function, but may be not a homomorphism function.

Group Codes – 群码

- An (m, n) encoding function $e: B^m \to B^n$ is called a *group code* if
  - $e(B^m) = \{e(b) \mid e(b) \in B^n\} = \text{Ran}(e)$
- is a subgroup of B^n

group code.

Subgroups

- Recall from the definition of subgroup give in Section 9.4 that N is a subgroup of B^n if
  - (a) the identity of B^n is in N,
  - (b) if x and y belong to N, then $x \oplus y \in N$, and
  - (c) if x is in N, then its inverse is in N.

EXAMPLE 7(is *e* a group code?)

- Consider the (3, 6) encoding function $e: B^3 \to B^6$ defined by

$$\left.\begin{array}{l}
e(000) = 000000 \\
e(001) = 001100 \\
e(010) = 010011 \\
e(011) = 011111 \\
\hline
e(100) = 100101 \\
e(101) = 101001 \\
e(110) = 110110 \\
e(111) = 111010
\end{array}\right\} \text{code word}$$

Example 7: is e a group code?

- We must show that the set of all code words
  - N = {000000, 001100, 010011, 011111, 100101, 101001, 110110, 111010}
- is a subgroup of B^6.

Theorem 3

- Let e: $B^m \to B^n$ be a group code. The minimum distance of e is the minimum weight of a nonzero code word.

Proof(1) of Theorem 3

- Let δ be the minimum distance of the group code, and suppose that $\delta = \delta(x, y)$, where x and y are distinct code words.
- Also, let η be the minimum weight of a nonzero code word and suppose that $\eta = |z|$ for a code word z.

Proof(2) of Theorem 3

- Since e is a group code, so $x \oplus y$ is a nonzero code word. Thus
  - $\delta = \delta(x, y) = |x \oplus y| \geq \eta$.
- On the other hand, since $\mathbf{0}$ and z are distinct code words,
  - $\eta = |z| = |z \oplus \mathbf{0}\,| = \delta(z, \mathbf{0}) \geq \delta$
- Hence $\eta = \delta$.

- Q.E.D.

Example 8

- The minimum distance of the group code in

$$\left.\begin{array}{l} e(000)=000000 \\ e(001)=001100 \\ e(010)=010011 \\ e(011)=011111 \\ \hline e(100)=100101 \\ e(101)=101001 \\ e(110)=110110 \\ e(111)=111010 \end{array}\right\}\text{code word}$$

- is 2
- To check this directly would require 28 different calculations.

Constructing

group code

群

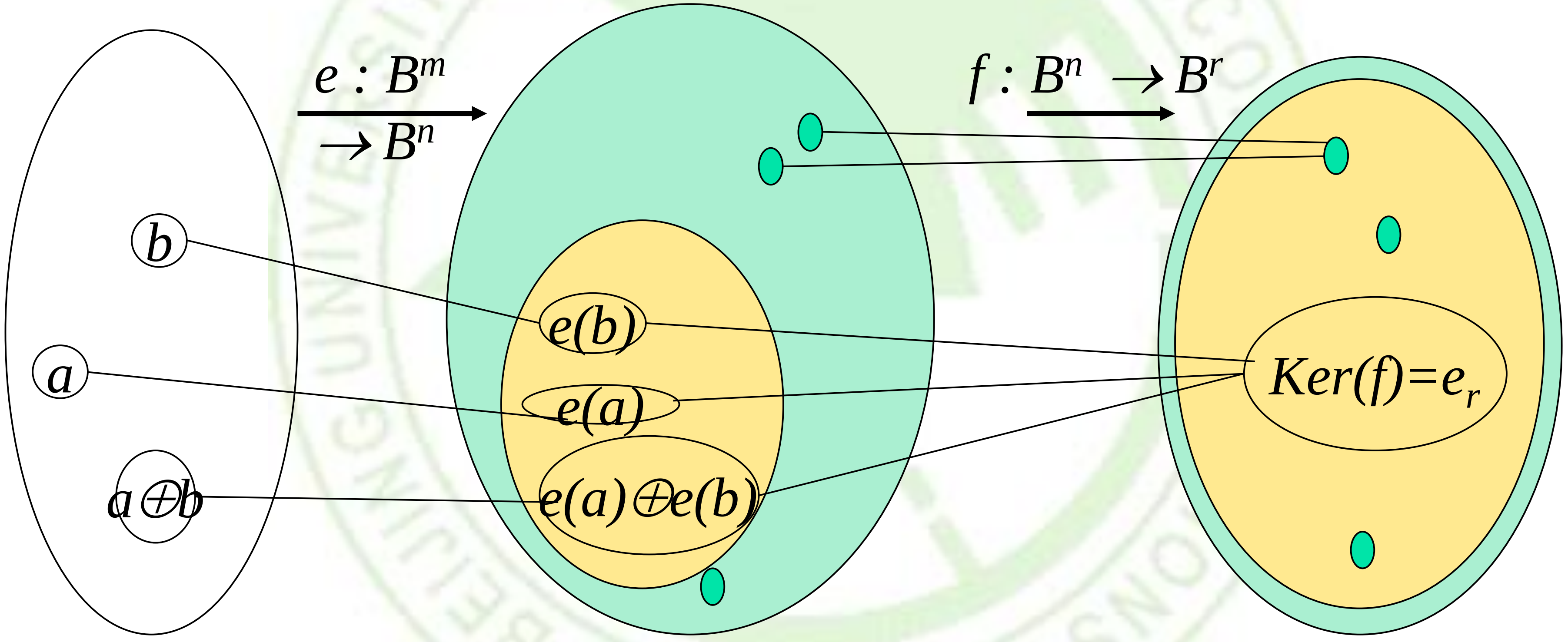
Zhang Yanmei

ymzhang@bupt.edu.cn

College of Computer Science & Technology

Beijing University of Posts & Telecommunications

Find the e and f function



e is a homomorphism and $e(B^m)$ is subgroup.

f is onto homomorphic, and $e(B^m)$ is ker(f).

Notation

element

- We shall now consider the element
  - $b = b_1 b_2 \ldots b_m \in B^m$ as the $1 \times$ m matrix
  - $[b_1\, b_2\, \ldots\, b_m\,]$ 1×m matrix
  - $x = x_1 x_2 \ldots x_n \in B^n$ as the $1 \times n$ matrix
  - $[x_1\, x_2\, \ldots\, x_n\,]$

Theorem 5

- Let m and n be nonnegative integers with $m < n$, $r = n - m$, and let $\boldsymbol{H}$ be an $n \times r$ Boolean matrix.
- Then the function $f_H: B^n \to B^r$ defined by
  - $f_H(x) = x * \boldsymbol{H},\ x \in B^n$;
- is a homomorphism from the group B^n to the group B^r.

Proof

- If x and y are elements in B^n, then
- $f_H(x \oplus y) = (x \oplus y) * \mathbf{H}$

  $= (\mathrm{x} * \mathbf{H}) \oplus (\mathrm{y} * \mathbf{H})$ by Theorem 4

  $= f_H(x) \oplus f_H(\mathrm{y})$.
- Hence f_H is a homomorphism from B^n to B^r .

A review on Boolean matrices

- **mod-2 sum $\mathbf{D} \oplus \mathbf{E}$**
- **mod-2 Boolean product $\mathbf{D} * \mathbf{E}$**

- Theorem 4
  - Let **D** and **E** be $m \times p$ Boolean matrices, and let **F** be a $p \times n$ Boolean matrix. Then
    - $(\mathbf{D} \oplus \mathbf{E}) * \mathbf{F} = (\mathbf{D} * \mathbf{F}) \oplus (\mathbf{E} * \mathbf{F})$.
  - That is, a distributive property holds for $\oplus$ and $*$.

Parity check matrix 一致性检验矩阵

- Let $m < n$ and $r = n - m$, the following $n \times r$ Boolean matrix is called a *parity check matrix*.

$$H = \begin{bmatrix} H_{m\times r} \\ I_r \end{bmatrix} = \left[\begin{array}{cccc} h_{11} & h_{12} & \cdots & h_{1r} \\ h_{21} & h_{22} & \cdots & h_{2r} \\ \vdots & \vdots & & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{mr} \\ \hline 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{array}\right]$$

奇偶校验
矩阵

Theorem 3

- Let m and n be nonnegative integers with $m < n$, $r = n - m$, and let $\boldsymbol{H}$ be an *parity check matrix*.
- And the function f_H: $B^n \rightarrow B^r$ defined by
  - $f_H(x) = x^*\mathbf{H}$, $x \in B^n$
- Then f_H is a **onto homomorphism** from the group B^n to the group B^r.

Proof $f_H(x) = x * H$ is onto

- Proof
  - Let $b = b_1 b_2 \ldots b_r$ be any element in B^r.
  - Letting $x = 0_1 \ldots 0_m b_1 b_2 \ldots b_r$
  - Then $x * \boldsymbol{H} = b$.
  - Thus $f_H(x) = b$, so f_H is onto.

- Q.E.D

Corollary 1

- Let m, n, r, $\mathbf{H}$, and f_H be as in Theorem 3. Then
  - $N = \{x \in B^n \mid x * \mathbf{H} = \mathbf{0}\}$
  - is a **normal subgroup** of B^n.
- Proof:
  - N is the kernel of the homomorphism f_H, so it is a normal subgroup of B^n.

e_H: $B^m \to B^n$ in matrix format

$$e_{\mathrm{H}}(B^m) = B^m * \begin{bmatrix} \mathrm{I}_m & \mathrm{H}_{m\times r} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{2^m 1} & b_{2^m 2} & \cdots & b_{2^m m} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & h_{11} & h_{12} & \cdots & h_{1r} \\ 0 & 1 & \cdots & 0 & h_{21} & h_{22} & \cdots & h_{2r} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & h_{m1} & h_{m2} & \cdots & h_{mr} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} & x_{11} & x_{12} & \cdots & x_{1r} \\ b_{21} & b_{22} & \cdots & b_{2m} & x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_{2^m 1} & b_{2^m 2} & \cdots & b_{2^m m} & x_{2^m 1} & x_{2^m 2} & \cdots & x_{2^m r} \end{bmatrix}$$

Theorem 6

- Let $x = y_1 y_2 \ldots y_m x_1 \ldots x_r \in B^n$. Then
  - $x * \mathbf{H} = \mathbf{0}$
- if and only if
  - $x = e_H(b)$ for some $b \in B^m$.

- Define an encoding function $e_H: B^m \rightarrow B^n$
  - If $b = b_1 b_2 \ldots b_m$,
  - let $x = e_H(b) = b_1 b_2 \ldots b_m \, x_1 x_2 \ldots x_r$, where

$$
\begin{aligned}
x_1 &= b_1 \cdot h_{11} + b_2 \cdot h_{21} + \ldots + b_m \cdot h_{m1} \\
x_2 &= b_1 \cdot h_{12} + b_2 \cdot h_{22} + \ldots + b_m \cdot h_{m2} \\
&\vdots \\
x_r &= b_1 \cdot h_{1r} + b_2 \cdot h_{2r} + \ldots + b_m \cdot h_{mr}
\end{aligned}
\qquad (1)
$$

PROOF: $x * \mathbf{H} = \mathbf{0} \Rightarrow x = e_H(b)$ for some $b \in B^m$

- Suppose that $x * \mathbf{H} = \mathbf{0}$

$$
\begin{aligned}
& y_1 \cdot h_{11} + y_2 \cdot h_{21} + \ldots + y_m \cdot h_{m1} + x_1 = 0 \\
& y_1 \cdot h_{12} + y_2 \cdot h_{22} + \ldots + y_m \cdot h_{m2} + x_2 = 0 \\
& \vdots \\
& y_1 \cdot h_{1r} + y_2 \cdot h_{2r} + \ldots + y_m \cdot h_{mr} + x_r = 0
\end{aligned}
$$

Proof: $x * \mathbf{H} = \mathbf{0} \Rightarrow x = e_H(b)$ for some $b \in B^m$

- Note that $x_i + x_i = 0$. So add x_i to i^{th} row and get

$$x_i = y_1 \cdot h_{1i} + y_2 \cdot h_{2i} + \ldots + y_m \cdot h_{mi}$$

- Letting $b_1 = y_1$, $b_2 = y_2, \ldots, b_m = y_m$, we see that $x_1, x_2, \ldots, x_r$ satisfy the equations in (l). Thus $b = b_1 b_2 \ldots b_m \in B^m$ and $x = e_H(b)$

Proof: $x * \mathbf{H} = \mathbf{0} \Leftarrow x = e_H(b)$ for some $b \in B^m$

- Conversely if $x = e_H(b)$, the equations in (l) can be rewritten by adding x_i to both sides of the i^{th} equation, i = l, 2,..., n, as

$$
\begin{aligned}
&b_1 \cdot h_{11} + b_2 \cdot h_{21} + \ldots + b_m \cdot h_{m1} + x_1 = 0 \\
&b_1 \cdot h_{12} + b_2 \cdot h_{22} + \ldots + b_m \cdot h_{m2} + x_2 = 0 \\
&\vdots \\
&b_1 \cdot h_{1r} + b_2 \cdot h_{2r} + \ldots + b_m \cdot h_{mr} + x_r = 0
\end{aligned}
$$

- which shows $x * \mathbf{H} = \mathbf{0}$

- Q.E.D.

Corollary 2

- $e_H(B^m) = \{e_H(b) \mid b \in B^m\}$ is a subgroup of B^n.

Proof :

- The result follows from the observation that
  - $e_H(B^m) = \ker(f_H)$
  - and from Corollary l.
- Thus e_H is a group code.

Example 11

- Let $m = 2$, $n = 5$, and

$$H = \begin{bmatrix} H_{2\times 3} \\ I_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathrm{H} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Determine the group code e_H: $B^2 \to B^5$.

Solution 11

$$
\begin{aligned}
e_H(B^m) &= B^m * [I_2 \; H_{2\times3}] \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

- What does H means?

Example 11

$$H = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Let $m = 2$, $n = 5$, and

$$e(B^m) = B^m * \mathrm{H}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Solution 11

- $e_H(B^m)$ *H= $\begin{bmatrix} 00000 \\ 01011 \\ 10110 \\ 11101 \end{bmatrix} * \begin{bmatrix} 110 \\ 011 \\ 100 \\ 010 \\ 001 \end{bmatrix} = \begin{bmatrix} 000 \\ 000 \\ 000 \\ 000 \end{bmatrix}$
- $e_H(B^m)$ *is group code because closure and e.*
- any x$\in B^n$ and $\notin e_H(B^m)$, *example*

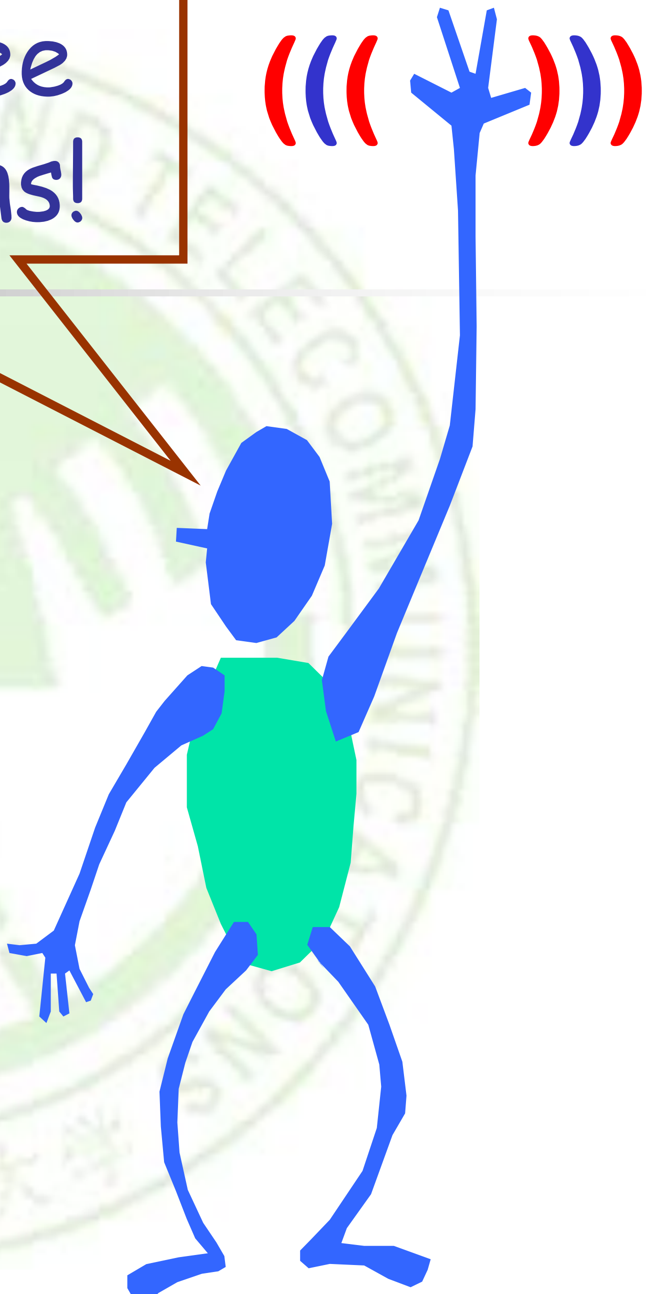
$$[00001] * \begin{bmatrix} 110 \\ 011 \\ 100 \\ 010 \\ 001 \end{bmatrix} = [001]$$

e_H: $B^m \to B^n$ is homomorphism

- The function e_H: $B^m \to B^n$ defined by
  - $e_H(b) = b * \mathbf{H}$, $b \in B^m$;
- is a homomorphism from the group B^m to the group B^n.
- proof:
  - *e_H is a one-to-one function.*
  - $e_H(a \oplus b) = (a \oplus b) * \mathbf{H} = (\boldsymbol{a}{*}\mathbf{H}) \oplus (b{*}\mathbf{H})$

  $= e_H(a) \oplus e_H(b)$

Please feel free to ask questions!



Homework

- 16,18,20,26 @412

编程作业：给定H(读取文件方式，第一行两个整数m,n，第二行$m\times(n-m)$个0或1，也就是矩阵H的上半部分，下半部单位矩阵自行生成)，计算群码编码函数e_H。

1. 计算该编码函数能检测到多少位错误
2. 交互输出字的码字

编程作业：针对$(8,12)$编码e，找出最小距离最大的群码编码函数，输出H及最小距离。