北海道大学
Hokkaido University

情報エレクトロニクス学科共通科目・2年次・夏ターム〔必修科目〕

講義「情報理論」

第12回

第8章 通信路符号化法(2)

8.2 ハミング符号(ハミング距離と誤り訂正能力)

[復習]単一パリティ検査符号

$x_1x_2\cdots x_k$ →通信路符号化→ w→2元通信路→y→誤り検出

w: 符号語、 y: 受信語

[単一パリティ検査符号の場合]

$w = x_1x_2\cdots x_k c$ [または$w=(x_1, x_2, \cdots, x_k, c)$ と表す]

ただし

$x_1x_2\cdots x_k$: 長さkの0,1の系列 ---------------情報記号(列)

$c = x_1 + x_2 + \cdots + x_k \pmod 2$ -----------検査記号

単一パリティ検査符号は、一記号の誤りを検出できる誤り検出符号

（error-detecting code）

[復習]組織符号

- k 個の情報記号から、$n-k$ 個の検査記号を一定の方法で求め、付加することにより符号化される符号長 n の符号を組織符号（systematic code）と呼ぶ。

$$w = \underbrace{x_1 x_2 \cdots x_k\, c_1 c_2 \cdots c_{n-k}}_{n}$$

- 符号長 n、情報記号数 k の組織符号を (n, k) 符号と書く。
  (n, k) 2元符号の効率は、

$$\eta = R / R_{\max} = (\log_2 M)/n = (\log_2 2^k)/n = k / n$$

  である。
- 単一パリティ検査符号は $(k+1, k)$ 組織符号であり、効率は$\eta = k / (k+1)$ である。

k を大きくとれば効率は上がるが、
冗長度が低くなり信頼性は小さくなる

[復習]線形符号とパリティ検査方程式

- **線形符号（linear code）**

  $c = x_1 + x_2 + \cdots + x_k$ のように、検査記号が情報記号の線形な式で与えられる符号

  線形符号の最も基本的な性質

  「任意の二つの符号語について、その成分ごとの和をとると、それがまた符号語になる」（線形符号となるための必要十分条件）

- **パリティ検査方程式**

  ＝0 という形で線形符号の符号語となるための必要十分条件を与える式（または式の組）

  [例]単一パリティ検査符号のパリティ検査方程式

  $$w_1 + w_2 + \cdots + w_{n-1} + w_n = 0$$

  （符号語に含まれる1の個数が偶数）

- **シンドローム（syndrome）**

  受信語 y をパリティ検査方程式の左辺に代入した結果 s 。すなわち、

  $$s = y_1 + y_2 + \cdots + y_n$$

  誤りがない⇒$s = 0$,　　1個の誤り ⇒$s = 1$

[復習]水平垂直パリティ検査符号

- 水平垂直パリティ検査符号

以下のように生成された符号語が$(x_1,x_2,x_3,x_4,c_1,c_2,c_3,c_4,c_5)$の(9,4)組織符号

$$\begin{array}{cc|c} x_1 & x_2 & c_1 \\ x_3 & x_4 & c_2 \\ \hline c_3 & c_4 & c_5 \end{array}$$

$c_1 = x_1 + x_2$

$c_2 = x_3 + x_4$

$c_5 = c_1 + c_2 = x_1 + x_2 + x_3 + x_4 = c_3 + c_4$

$c_3 = x_1 + x_3 \quad c_4 = x_2 + x_4$

1個の誤りが訂正と、2個の誤りを検出が可能。このような符号を、誤り訂正検出符号、あるいは簡単に誤り訂正符号(error-correcting code)と呼ぶ。

パリティ検査方程式

$$\begin{aligned} w_1 + w_2 + w_5 &= 0 \\ w_3 + w_4 + w_6 &= 0 \\ w_1 + w_3 + w_7 &= 0 \\ w_2 + w_4 + w_8 &= 0 \\ w_1 + w_2 + w_3 + w_4 + w_9 &= 0 \end{aligned}$$

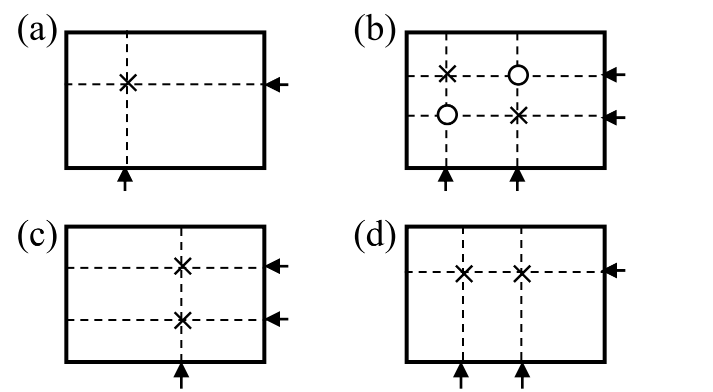


図: 単一誤りの訂正と2重誤りの検出

[復習](7,4)ハミング符号(1)

- (7,4)ハミング符号

4個の情報ビット x_1, x_2, x_3, x_4 に対し、

$$c_1 = x_1 + x_2 + x_3$$
$$c_2 = \quad x_2 + x_3 + x_4$$
$$c_3 = x_1 + x_2 \quad + x_4$$

により、検査ビット c_1, c_2, c_3 を作り、

$$\mathrm{w} = (x_1, x_2, x_3, x_4, c_1, c_2, c_3)$$

という符号語に符号化。

情報ビットが4個、符号語は $2^4 = 16$個

表:(7,4)ハミング符号

x_1	x_2	x_3	x_4	c_1	c_2	c_3
0	0	0	0	0	0	0
1	0	0	0	1	0	1
0	1	0	0	1	1	1
1	1	0	0	0	1	0
0	0	1	0	1	1	0
1	0	1	0	0	1	1
0	1	1	0	0	0	1
1	1	1	0	1	0	0
0	0	0	1	0	1	1
1	0	0	1	1	1	0
0	1	0	1	1	0	0
1	1	0	1	0	0	1
0	0	1	1	1	0	1
1	0	1	1	0	0	0
0	1	1	1	0	1	0
1	1	1	1	1	1	1

[復習](7,4)ハミング符号(2)

- 符号語を $\boldsymbol{w}=(w_1, w_2, \cdots, w_7)$ する。

<u>パリティ検査方程式</u>

$$w_1+w_2+w_3+w_5=0$$
$$w_2+w_3+w_4+w_6=0$$
$$w_1+w_2+w_4+w_7=0$$

<u>受信語 $\boldsymbol{y}=(y_1, y_2, \cdots, y_7)$ に対するシンドローム</u>

$$\boldsymbol{s}_1=y_1+y_2+y_3+y_5$$
$$\boldsymbol{s}_2=y_2+y_3+y_4+y_6$$
$$\boldsymbol{s}_3=y_1+y_2+y_4+y_7$$

- （パリティ）検査行列（parity check matrix）

(7,4)ハミング符号のパリティ検査方程式の係数行列 $\boldsymbol{H}$

$$\begin{array}{c} \begin{array}{ccccccc} x_1 & x_2 & x_3 & x_4 & c_1 & c_2 & c_3 \end{array} \\ \boldsymbol{H}=\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \boldsymbol{H}^{\mathrm{T}}=\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

パリティ検査方程式　と　シンドロームの計算式は

$$\boldsymbol{w}\,\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{0} \qquad \boldsymbol{s}=\boldsymbol{y}\,\boldsymbol{H}^{\mathrm{T}}$$

と書ける。

[復習](7,4)ハミング符号(3)

ハミング符号は1個の誤りが訂正が可能な誤り訂正符号

誤りパターン							シンドローム		
e_1	e_2	e_3	e_4	e_5	e_6	e_7	s_1	s_2	s_3
1	0	0	0	0	0	0	1	0	1
0	1	0	0	0	0	0	1	1	1
0	0	1	0	0	0	0	1	1	0
0	0	0	1	0	0	0	0	1	1
0	0	0	0	1	0	0	1	0	0
0	0	0	0	0	1	0	0	1	0
0	0	0	0	0	0	1	0	0	1
0	0	0	0	0	0	0	0	0	0

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

単一誤りに対するシンドロームの行 ⟷ 検査行列の列

ハミング距離とハミング重み(1)

２つのn次元ベクトル$\boldsymbol{u}=(u_1,u_2,\cdots,u_n)$, $\boldsymbol{v}=(v_1,v_2,\cdots,v_n)$の間の**ハミング距離**$d_H(\boldsymbol{u},\boldsymbol{v})$

$$\overset{def}{\Leftrightarrow}\quad d_H(u,v)=\sum_{i=1}^{n}\delta(u_i,v_i) \qquad \text{ただし} \quad \delta(u,v)=\begin{cases}0 & \text{if } u=v\\ 1 & \text{if } u\neq v\end{cases}$$

$d_H(u,v)$はuとvの対応する位置にある成分の対のうち、互いに異なるものの数

ハミング距離は**距離の３公理**を満たす。

距離の３公理

任意のn次元ベクトル$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$に対して以下のことが成り立つ。

(i) $d_H(\boldsymbol{v}_1,\boldsymbol{v}_2)\geqq 0$であり、等号が成立するのは$\boldsymbol{v}_1=\boldsymbol{v}_2$のときに限る。

(ii) $d_H(\boldsymbol{v}_1,\boldsymbol{v}_2)=d_H(\boldsymbol{v}_2,\boldsymbol{v}_1)$

(iii) $d_H(\boldsymbol{v}_1,\boldsymbol{v}_2)+d_H(\boldsymbol{v}_2,\boldsymbol{v}_3)\geqq d_H(\boldsymbol{v}_1,\boldsymbol{v}_3)$ (三角不等式)

ハミング距離とハミング重み(2)

n次元ベクトル$\boldsymbol{v}$の**ハミング重み**または**重み**$w_H(\boldsymbol{v})$ $\overset{\text{def}}{\Leftrightarrow}$ $w_H(\boldsymbol{v})=d_H(\boldsymbol{v},\boldsymbol{0})$

$w_H(\boldsymbol{v})$は$\boldsymbol{v}$の0でない成分の数

ハミング距離はハミング重みを用いて次のように表せる。

$$d_H(\boldsymbol{u},\boldsymbol{v})=w_H(\boldsymbol{u}-\boldsymbol{v})$$

（例） 符号語$\boldsymbol{w}$を送りt個の誤りが生じて$\boldsymbol{y}=\boldsymbol{w}+\boldsymbol{e}$が受信された場合

$$d_H(\boldsymbol{w},\boldsymbol{y})=w_H(\boldsymbol{e})=t$$

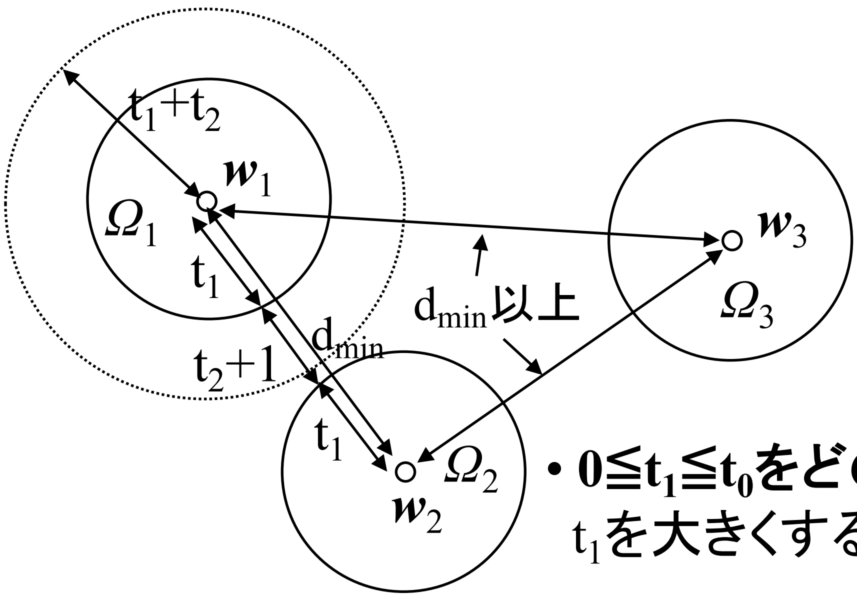
最小距離と誤り訂正能力(1)

符号Cの**最小ハミング距離**または**最小距離**(minimum distance) d_{min}

$$\overset{\text{def}}{\Leftrightarrow} \quad d_{min} = \min_{\boldsymbol{u}\neq\boldsymbol{v},\ \boldsymbol{u},\boldsymbol{v}\in C} d_H(\boldsymbol{u},\boldsymbol{v})$$

限界距離復号法

式 $d_{min} \geqq 2t_1+1$ を満たす整数t_1を定め、t_1以下の誤り訂正を行う復号法



- t_1の最大値$t_0 = \lfloor (d_{min}-1)/2 \rfloor$を**誤り訂正能力**という。
- $t_2 = d_{min} - 2t_1 - 1$とおけば $t_1+1 \leqq t \leqq t_1+t_2$個の誤りは訂正はできないが検出は可能
- **$0 \leqq t_1 \leqq t_0$をどのように選ぶかは重要な問題**
  t_1を大きくする→正しく復号される確率は増大するが
  誤って復号される確率も増大
  検出さえできれば、再送要求などの救済措置が可能

最小距離と誤り訂正能力(2)

【例】$d_{min}=5$の符号による誤りの訂正と検出

t_1	訂正可能な誤り	訂正できないが検出可能な誤り
0	ー	1～4個
1	1個	2～3個
2	2個	ー

線形符号の最小距離＝0でない符号語のハミング重みの最小値

最小ハミング重みまたは**重み**

何故ならば $d_{min} = \min_{u\neq v,\ u,v\in C} d_H(\boldsymbol{u},\boldsymbol{v}) = \min_{u\neq v,\ u,v\in C} w_H(\boldsymbol{u}-\boldsymbol{v}) = \min_{w\in C, w\neq 0} w_H(\boldsymbol{w})$

[ハミング符号]　最小距離d_{min}＝3、 誤り訂正能力$t_0=1$

（例）(7,4)ハミング符号の場合　最小距離d_{min}＝最小ハミング重み=3

[水平垂直パリティ検査符号]　最小距離$d_{min}=4$、 誤り訂正能力$t_0=1$

単一誤り訂正・2重誤り検出符号

(single-error-correcting/double-error-detecting code;SEC/DED符号)