定義6.2. $\{0,1\}^n$ 上の距離 d が $x=x_1\cdots x_n$, $y=y_1\cdots y_n$ に対し.

$d(x,y)=|\{i \mid x_i \neq y_i,\ 1\leq i\leq n\}|$ で与えられるとき.

この距離をハミング距離という.

定義6.3.

符号 σ は等長符号とする. $(\sigma:\Sigma\to\{0,1\}^n)$

σ が d-誤り検出(resp. 訂正)可能を

$\overset{\text{def}}{\iff}$ 含まれている誤りが d 個以下なら誤りを検出(resp. 訂正)できることをいう.

定義6.4. σ の最小距離を.

$m(\sigma):=\min\{d(\sigma(\alpha),\sigma(\beta)) \mid \alpha,\beta\in\Sigma,\alpha\neq\beta\}$ で定義する.

定理6.5. 次が成立.

(1) σ が d-誤り検出可能 $\iff m(\sigma)\geq d+1$

(2) σ が d-誤り訂正可能 $\iff m(\sigma)\geq 2d+1$.

(証)(1)($\Leftarrow$) $\sigma(\alpha)$ に誤りが d 個以下であるとき. これは他のどの $\sigma(\beta)$ とも一致しない.

∴誤りがあることがわかる.

($\Rightarrow$) 対偶を示す. $m(\sigma)\leq d$ のとき. $\exists\alpha,\beta\in\Sigma$ s.t. $d(\sigma(\alpha),\sigma(\beta))\leq d$

とできる. ここで $\sigma(\alpha)$ の中の $d(\sigma(\alpha),\sigma(\beta))$ 個に誤りがあり, その結果 $\sigma(\beta)$ になっている場合, 誤りを検出できない. //

(2)($\Leftarrow$). $\sigma(\alpha)$ の高々 d 個を変えて x となったとする. ここで. $\alpha\neq\beta\in\Sigma$ に対し.

$d(x,\beta)\geq d(\alpha,\beta)-d(\alpha,x)\geq d+1$ である.

∴ xから距離 d 以下の符号語は1つしかないので.

xを$\sigma(\alpha)$にすればよい

($\Rightarrow$) 逆に. $m(\sigma) \leq 2d$ のとき. $d(\sigma(\alpha), \sigma(\beta)) = 2d$ となる α, β がとれる.

ここで x を $d(\sigma(\alpha), x) = d$, $d(\sigma(\beta), x) = d$ となるようとれば.

xは. $\sigma(\alpha), \sigma(\beta)$ のどちらが誤ったものかわからない //

例 (1). $a, b, \cdots, h$ をそれぞれ 000, 001, ⋯, 111 に対応させる符号 σ_1 は

$m(\sigma_1) = 1$ より 誤り検出.訂正ではない.

(2). (1) を3回繰り返す符号 σ_2.

例 $\sigma_2(a) = 000000000$, $\sigma_2(b) = 001001001$ は

$m(\sigma_2) = 3$ より. 2-誤り検出, 1-誤り訂正である.

(3). (1) の符号の和を最後につけた符号 σ_3

$a : 0000$ $b : 0011$, $c : 0101$, $d : 0110$,

$e : 1001$ $f : 1010$, $g : 1100$, $h : 1111$ は

$m(\sigma_3) = 2$ より 1-誤り検出可能である.

$\{0.1\}^n$ の中である符号語 x と距離1にある符号語の数は.

${}_nC_1 = n$ 個である.

距離2にあるのは ${}_nC_2 = \frac{n(n-1)}{2}$ である.

なので $m(\sigma) = 3$ となるためには. $|\Sigma| = k$ として.

$k(1+n) \leq 2^n$ であることが必要である.

なお. (3) の符号はパリティ符号とよばれる.

(4) これをさらに拡張する. $\sigma_1(x) = x_1x_2x_3$ に対して.

$$y_1 = x_2 + x_3$$
$$y_2 = x_3 + x_1 \quad — ☆$$
$$y_3 = x_1 + x_2.$$
$$y_4 = x_1 + x_2 + x_3 \quad (\mathrm{mod}\ 2)$$ とし.

$\sigma_4(x) = x_1x_2x_3y_1y_2y_3y_4$ とすると.

a : 0000000　b : 0011101　c : 0101010 … となるが.

もし. $x_1x_2x_3$ と $x_1'x_2'x_3'$ の1つだけちがうとき ($x_1 \neq x_1'$) のときは.

y_2, y_3, y_4 の3つが違う.

もし. 2つが違うとき. ($x_1 \neq x_1'$, $x_2 \neq x_2'$) は.

y_1, y_2 が違う.

もし. 3つ全部違うときは

y_4 が違う.

∴ $m(\sigma_4) = 4$ となり. 3-誤り検出, 1-誤り訂正となる.

これをハミング符号という.

ここで 1-誤り訂正であるが. これも次を使って. 簡単にできる.

$$\begin{cases} S_1 = x_2 + x_3 + y_1 \\ S_2 = x_3 + x_1 + y_2 \\ S_3 = x_1 + x_2 + y_3 \\ S_4 = x_1 + x_2 + x_3 + y_4 \end{cases}$$

とする. ☆との関連に注意.

このSをシンドロームという.

誤りがなければ全て 0 になるはずである。

誤りが 1 であるとき。

誤り	S_1	S_2	S_3	S_4
x_1	0	1	1	1
x_2	1	0	1	1
x_3	1	1	0	1
y_1	1	0	0	0
y_2	0	1	0	0
y_3	0	0	1	0
y_4	0	0	0	1

となり、どれが誤りかはシンドロームを計算すればよい。