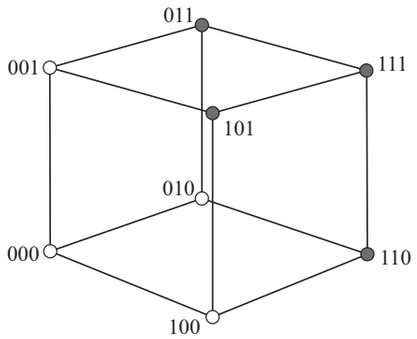
通信による情報伝達では，情報源から出された情報（文字，画像，音声など）はそのまま伝達できないので，数字や文字，記号による情報記号に変換され，「符号語」に符号化されて通信路を通って受信者へと伝達される．現在使われている，いわゆるデジタル技術では，情報記号や符号語として 000，1101110 のようなすべて 0 と 1 の記号の列が用いられる．これらは，$\mathbb{F}_2 = \{0, 1\}$ の元を成分とするベクトル $(0, 0, 0)$, $(1, 1, 0, 1, 1, 1, 0)$ とみなすことができる．

符号語とは，情報記号に伝送中に生じる誤りを検出・訂正するための検査記号を付加したものである．そして，符号語の集合を**符号**と呼ぶ．また，符号語に用いられている記号の数を符号の**長さ**と呼ぶ．例えば，長さ 3 の符号語は

$$000,\quad 001,\quad 010,\quad 011,\quad 100,\quad 101,\quad 110,\quad 111$$

の計 8 個ある．



1 長さ 4 の符号に含まれる符号語をすべてあげよ．

符号語はすべて 0 と 1 だけからなる記号列であるから，符号語において誤りが生じるということは，該当する箇所の 0 と 1 が入れ替わるということである．ここで $1 + 1 = 0$ となる $\mathbb{F}_2$ における加法を用いると，符号語のある桁（ビットと呼ぶ）に誤りが生じることはすなわち，そのビットに $\mathbb{F}_2$ の意味で 1 を加えるということに他ならない．例えば，長さ 8 の符号 11101010 の 3 ビット目と 6 ビット目に誤りが生じたとすると，11101010 をベクトル $(1, 1, 1, 0, 1, 0, 1, 0)$ とみなし，これに $(0, 0, 1, 0, 0, 1, 0, 0)$ を加え，

$$(1, 1, 1, 0, 1, 0, 1, 0) + (0, 0, 1, 0, 0, 1, 0, 0) = (1, 1, 1, 0, 0, 1, 1, 1, 0)$$

が受信側が受け取る符号語となる．一般に，送信した符号語を $\vec{u}$，誤りパターンを $\vec{e}$，受信した符号語を $\vec{r}$ とすると，

$$\vec{u} + \vec{e} = \vec{r}$$

が成り立つ．もちろん，ベクトルの成分の加法は $\mathbb{F}_2$ における加法とする．

入学年度		学部	学科		組		番号				検	フリガナ	
		D		1								氏名	

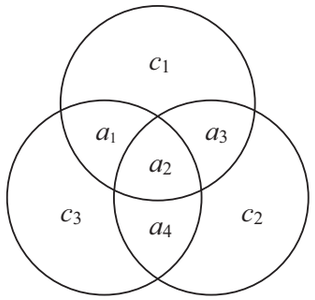
2 つの符号語 $\vec{a} = (a_1, a_2, \ldots, a_n)$ と $\vec{b} = (b_1, b_2, \ldots, b_n)$ の**ハミング距離**とは，対応するビット成分で値が異なるものの数である．例えば，上の $(1, 1, 0, 0, 1, 1, 1, 0)$ と $(1, 0, 1, 0, 1, 1, 0, 1)$ のハミング距離は 4 である．

$$\begin{array}{r} (\ 1, 1, 0, 0, 1, 1, 1, 0\ ) \\ +(\ 1, 0, 1, 0, 1, 1, 0, 1\ ) \\ \hline (\ 0, 1, 1, 0, 0, 0, 1, 1\ ) \end{array} \rightarrow 1 \text{ が } 4 \text{ 個}$$

2 01001101 と 10111100 のハミング距離を求めよ．

• ハミング符号 $H(7, 4)$

いま，長さ 4 の**情報ビット** $a_1a_2a_3a_4$（a_i は 0 または 1）をベクトルで表し $\vec{a} = (a_1, a_2, a_3, a_4)$ とし，右の図の各円内のパリティ（偶奇）が 0 になるように c_1, c_2, c_3 を決める．すなわち



$$(\star) \qquad \begin{cases} c_1 = a_1 + a_2 + a_3 \\ c_2 = a_2 + a_3 + a_4 \\ c_3 = a_1 + a_2 + a_4 \end{cases}$$

となるように決める．そして，$\vec{a}$ に**検査ビット** $c_1c_2c_3$ を付加して長さ 7 の符号 $\vec{u} = (a_1, a_2, a_3, a_4, c_1, c_2, c_3)$ を作る．

たとえば，情報ビットが 1101 であれば，

$$\begin{cases} c_1 = 1 + 1 + 0 = 0 \\ c_2 = 1 + 0 + 1 = 0 \\ c_3 = 1 + 1 + 1 = 1 \end{cases}$$

となり，$\vec{u} = (1, 1, 0, 1, 0, 0, 1)$ となる．

◦ ハミング符号の誤り訂正の原理

ハミング符号は 1 個の誤りを訂正できる．誤り訂正には**シンドローム**と呼ばれる数の組が用いられる．符号語 $\vec{u} = (a_1, a_2, a_3, a_4, c_1, c_2, c_3)$ に対して，シンドローム s_1, s_2, s_3 を次のように計算する．

$$\begin{cases} s_1 = a_1 + a_2 + a_3 + c_1 \\ s_2 = a_2 + a_3 + a_4 + c_2 \\ s_3 = a_1 + a_2 + a_4 + c_3 \end{cases}$$

各シンドロームは，表の図の 3 つの円内の 4 ビットがパリティ検査符号であると考えて，誤りがあるかを検査するものである．例えば，$(s_1, s_2, s_3) = (1, 1, 0)$ であるとする．このとき，誤りはひとつだけあると仮定すると，それは上の円と右下の円内に属し，左下の円内にはないことを意味する．したがって，誤りは a_3 であることがわかる．

【裏に続く】

3 整数 0 から 15 までを 2 進法で表し，それを長さ 4 の情報ビットとみなし，それぞれについて検査ビットを求め，以下の表を完成させよ．

10 進法	2 進法	検査ビット	符号語
0	0000	000	0000000
1	0001		
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13	1101	001	1101001
14			
15	1111	111	1111111

4 1011011 の誤りを訂正せよ．

◦ ハミング符号 $H(7,4)$ と $\mathbb{F}_8$

実は，ハミング符号は 8 元体 $\mathbb{F}_8$ と密接に関連している．いま，長さ 4 の情報ビット $a_1a_2a_3a_4$ に対し，$\mathbb{F}_2[x]$ の元

$$q(x) = a_1x^3 + a_2x^2 + a_3x + a_4$$

を対応させる．そして，$q(x)$ に x^3 をかけて 6 次式とし，$q(x)x^3$ を $g(x) = x^3 + x + 1$ で割った余りを付けくわえて，$u(x)$ を

$$u(x) = q(x)x^3 + \big(\, q(x)x^3 \text{ を } g(x) \text{ で割った余り} \,\big)$$

と定義する．$\mathbb{F}_2[x]/(g(x)) \simeq \mathbb{F}_8$ だから，$q(x)x^3$ を $g(x)$ で割った余りはすなわち $q(x)x^3$ に $\mathbb{F}_8$ の生成元 α（$\alpha^3 + \alpha + 1$ をみたす元）を代入したものを α の 2 次以下の式に直したものと解釈できる．

担当：鍬田 政人 (kuwata.04f@g.chuo-u.ac.jp)

例えば，情報ビットを 1101 としたとき，$q(x) = x^3 + x^2 + 1$ であり，$q(x)x^3$ に $x = \alpha$ を代入すると，

$$(\alpha^3 + \alpha^2 + 1)\alpha^3 = \alpha^6 + \alpha^5 + \alpha^3 = (\alpha^2 + 1) + (\alpha^2 + \alpha + 1) + \alpha + 1 = 1$$

したがって，$q(x)x^3$ を $g(x)$ で割った余りは 1 であり，

$$u(x) = x^6 + x^5 + x^3 + 1$$

となる．これを 0 と 1 の列に直すと 1101001 となり，先の問題で 1101 に検査ビット 001 を加えて 1101001 にしたこと全く同じになる．

5 情報ビット $a_1a_2a_3a_4$ が与えられたとき，上の操作で得られる $u(x)$ は

$$u(x) = a_1x^6 + a_2x^5 + a_3x^4 + a_4x^3 + c_1x^2 + c_2x + c_3$$

で与えられることを示せ．ただし，c_1, c_2, c_3 は $(\star)$ で与えられる数（$\mathbb{F}_2$ の元）である．

6 上の操作で得られる $u(x)$ は $g(x) = x^3 + x + 1$ で割りきれることを示せ．

【次回以降のお楽しみ】符号 $a_1a_2a_3a_4c_1c_2c_3$ が送信され，それを受信者が $b_1b_2b_3b_4b_5b_6c_7$ と受信したとする．誤りの個数（ビット数）が 1 以下であるとき，これを訂正するにはどのようにすればよいだろうか?