

情報と科学：情報の誤り訂正について

平成8年10月16日(水)

担当：理学部 情報科学科 塩田

1. デジタル信号

0と1の組み合わせで表現される信号をデジタル信号と呼びます。例えば音声の信号は、音の波形を数値に変換し、さらにその数値を2進数に変換することによってデジタル化することができます。

0または1のひとつをビットと呼ぶことを覚えておいてください。

2. デジタルは美しい

コンパクトディスク(CD)・レーザーディスク(LD)・デジタルビデオ(DV)などデジタルで情報を記録・再生する機械を用いると、雑音の少ない美しい音声・画像を楽しむことができます。

しかし、単に「デジタルだから雑音が少ない」という訳ではありません。CDやLDの表面にも埃やキズは付いていますが、そのことによる情報の誤りを自動的に検出・訂正する仕組みが使われているのです。

3. 誤り訂正とは(日常言語の場合)

たとえば下宿中の息子から「カメオクレ ケンタ」という電報が届いたとしましょう。生物学科の学生ならいざ知らず、普通は「亀を送れ」ということは考えられません。常識的に考えてこれは「カネオクレ ケンタ」の間違いだと判断するでしょう。この様に、日常言語では常識・状況などから多少の誤りを訂正することができます。

4. 誤り検出とは(日常言語の場合)

高知県には鏡村と香我美町があります(どちらも読みは「かがみ」です)。

「かがみむらのちょうちょうさん」

という文章は、村には町長さんは居ないので、間違っていることはわかります(誤り検出)。しかし、これが「鏡村の村長さん」の間違いなのか、「香我美町の町長さん」の間違いなのかまではわかりません。

5. デジタル信号の誤り訂正

今、1%の確率でビットを誤る通信路を使って、YesかNoかを連絡する必要があるという状況を考えてください。

先ずYesなら0、Noなら1を送信するとすると、100回に1回の割合でYes、Noを間違えてしまいます。

ところがYesなら000、Noなら111を送信するとするとどうでしょう。もし受信した信号が001ならば、000のひとつのビットが化けて001になる確率の方が、111の2つのビットが化けて001になる確率よりも高いので、もとの信号は000、すなわちYesを送信したのだと判断することにします。この方法だと1万回に3回の割合でしかYes、Noを間違えずに済みます。

6. 誤り訂正符号

この様に必要最小限のビット数（情報ビット）以外に検査ビットという「おまけのビット」を付けてデジタル信号を作り、そのおまけの効果によって情報の誤りを訂正するテクニックを誤り訂正符号と言います。

7. デジタル信号の誤り検出

5. の例でも、Yes を 00、No を 11 として送信した場合、受信した信号が 01 ならば、それが 00 の後ろのビットが化けたものである確率と 11 の前のビットが化けたものである確率とが等しいので、誤り検出はできますが誤り訂正はできません。

8. ハミング距離

長さの等しいふたつのビット列を比べたとき、異なっているビットの個数をその「ハミング距離」と言います。長さ 3 のビット列では

$$\begin{aligned}d(000, 000) &= 0, & d(000, 100) &= d(000, 010) = d(000, 001) = 1, \\d(000, 110) &= d(000, 101) = d(000, 011) = 2, & d(000, 111) &= 3\end{aligned}$$

のように数えます。

ハミング距離の概念を使うと 5. の誤り訂正の例は、『受信した信号と比べて 000 と 111 のどちらがハミング距離が近いかを計算し、その近い方が送信された信号だと判断する』と言い表すことができます。

9. 例

A,B,C,D,E,F,G の 7 文字の情報を連絡するとき、それぞれ次のビット列を送信するとしましょう。

$$\begin{aligned}A &\rightarrow 0001011 & B &\rightarrow 1000101 & C &\rightarrow 1100010 & D &\rightarrow 0110001 \\E &\rightarrow 1011000 & F &\rightarrow 0101100 & G &\rightarrow 0010110\end{aligned}$$

この 7 つのビット列はどの 2 つもハミング距離で 4 離れていますので、7 ビット中の 1 ビットの誤りを正しく訂正することができ、7 ビット中の 3 ビットの誤りを検出することができます。

10. 法則

上の例のように、送信するビット列のどの 2 つもハミング距離で距離 D 以上離れているときには、 $\frac{(D-1)}{2}$ 個までの誤りを正しく訂正することができ、 D 個未満の誤りを検出することができます。（詳しくは講義で説明します。）

11. 補足

良い誤り訂正符号を作るためには（例えば誤り訂正能力の高いなど）、様々な数学的構造が使われています。例えば 9. の例は有限射影幾何学という数学を用いて作ったものです。座標が 0 と 1 しかない世界で幾何学図形を考えるという代物ですが、興味があったら勉強してみてください。

参考文献：岩垂好裕 著、符号理論入門（昭晃社）、1992.