

創想館 14-201

2008-9



Tuesday SPRING
13:00-14:30

 Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami



2008年度

「情報工学」

第9回講義

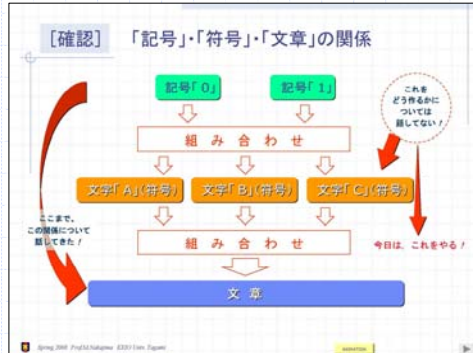
 Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami



前回学んだこと

- ◆ 覚えていなければいけないこと
- ◆ 答えられなければならないこと

✓ '01符号'によって各'文字'(A, B, C...など)をどう表すかについて学んだ.



前回学んだこと

- ◆ 覚えていなければいけないこと
- ◆ 答えられなければならないこと

✓ '01符号'によって各'文字'(A, B, C...など)をどう表すかについて学んだ.

✓ まず、「符号化とは」、「符号化の方針とは」について学んだ.

✓ 引き続き、「具体的な'01符号'を作る方法」としての「シャノンの符号化法」を学んだ.

- ☞ シャノンの符号化法は、「効率」という観点から見ると、ベストな符号化法ではない
したがって、現在は使われていないが、歴史的意義は大きい！
➤ やり方については、一応覚えておこう！

✓ 符号化の効率： $\eta = \frac{R}{C}$ ， 符号化の冗長度： $1 - \eta = 1 - \frac{R}{C}$

前回学んだこと

- ◆ 覚えていなければいけないこと
- ◆ 答えられなければならないこと

- ✓ '01符号'によって各'文字'(A, B, C...など)をどう表すかについて学んだ.
- ✓ まず,「符号化のとは」,「符号化の方針は」について学んだ.
- ✓ 引き続き,具体的な'01符号を作る方法'としての「シャノンの符号化法」を学んだ.

☞ シャノンの符号化法は, '効率'という観点から見ると, ベストな符号化法ではない
したがって, 現在は使われていないが, 歴史的意義は大きい!
➤ やり方については, 一応覚えておこう!

- ✓ 符号化の効率 **今日は, 'ベストな符号化法'を学ぶ!**

§ 5.2.2 ハフマンの符号化法

➤ 現在, '誤りのない系ではベストな符号化法'と言われている.

ハフマンの符号化法の手順

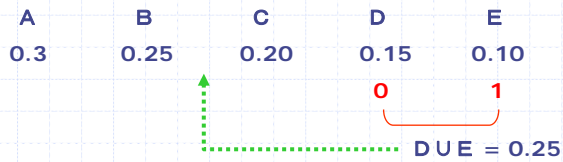
例によって, 具体的な例を用いて説明する

☞ 4つの文字(A,B,C,D)で作られた非常に長い文章を調べたら, 各文字の生起確率が, 下表の通りであった. 誤りのない系において, 各文字に対する最適な'01符号'を作ってみよう.

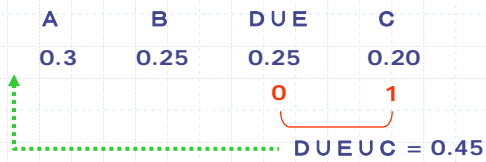
通報	生起確率
A	$p_1=0.30$
B	$p_2=0.25$
C	$p_3=0.20$
D	$p_4=0.15$
E	$p_4=0.10$

[STEP.1] 生起確率大の順に並べる \Rightarrow 既に、並んでいる！

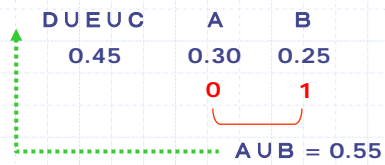
[STEP.2] 小さい方から、'和事象'をとる。
同時に、小さい方の2つの情報に、「0」と「1」をあてがう。



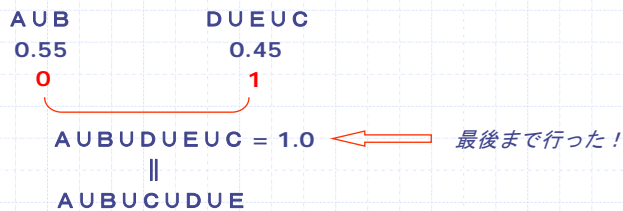
[STEP.3] 確率順に並べ直した後、STEP.2 のプロセスを繰り返す。



[STEP.4] 和事象が作れなくなるまでSTEP.2 のプロセスを繰り返す。



後は同じ・・・



[STEP.6] 作られた和事象の中から、その文字を含むも和事象に割り振られた「0」、「1」を生起確率大の順に並べ、最終符号とする。

文字	作られた和事象のうち、その文字を含むもの	最終符号
E	$DUEUC=0.45(1)$, $DUE=0.25(0)$ $E=0.1(1)$	101
D	$DUEUC=0.45(1)$, $DUE=0.25(0)$ $D=0.15(0)$	100
C	$DUEUC=0.45(1)$, $C=0.20(1)$	11
B	$AUB=0.55(0)$, $B=0.25(1)$	01
A	$AUB=0.55(0)$, $A=0.3(0)$	00

宿題

- (1) 上記例題を‘シャノンの符号化法’で符号化しなさい。
- (2) 次に、‘無作為な符号化法’により符号化しなさい。
- (3) これら3つの方法による符号化の結果を‘符号化の効率’により比較しなさい。

以上、誤りの生じない通信路での符号化法について述べてきた。

ここから、誤りの生じる通信路での符号化法について述べる。

§ 5.3 誤りがある場合の符号化法

- 誤りの発生に対する基本的な対策

⇒ 繰り返し送信

- 誤りの発生への対策として一般的な手法はない！

⇒ 個々に色々な方法が提案されている。

以下、誤りのある場合の符号化法について、
代表的なものを2つを紹介する・・・

1. パリティ検査符号
2. 巡回符号

§ 5.3.1 パリティ検査符号

各符号の末尾に最終ビットとして、「0」か「1」を付加し、各符号の「0」、
「1」の和が常に偶数または奇数のどちらかになるようにしたもの。

⇒ 奇数パリティ or 偶数パリティ

具体例：
(偶数パリティ)

A = 0 0 0 0	E = 1 0 0 1
B = 0 0 1 1	F = 1 0 1 0
C = 0 1 0 1	G = 1 1 0 0
D = 0 1 1 0	H = 1 1 1 1

➡ 「0」、「1」を入れ替えれば、奇数パリティになる。

● 特長：常に「1ビットのエラーを検出」できる。

§ 5.3.2 巡回符号



前もって知っておく必要があること

文字 u_1 と u_2 の
2つの「0,1符号」

ハミング距離 $d(u_1, u_2)$

$$u_1 = 100101111 \leftarrow x_i$$

$$u_2 = 101101101 \leftarrow y_i$$

$$d(u_1, u_2) = \sum_{i=1}^N (x_i - y_i)^2 = (1-1)^2 + (0-0)^2 + \underline{(0-1)^2} + \dots + \dots + \underline{(1-0)^2} + (1-1)^2 = (-1)^2 + 1^2 = 2$$

ハミング距離を使うことの意味？



ハミング距離を使うことの意味？

▷ ハミング距離一定の符号

- A 000
- B 011
- C 110
- D 101

d=2の3桁の符号系列

送信信号 000 101 011 110 101
 $d=2$ $d=2$ $d=2$ $d=2$

受信信号 000 101 001 110 101
 $d=2$ $d=1$ $d=1$ $d=2$
 ↑? ↑?

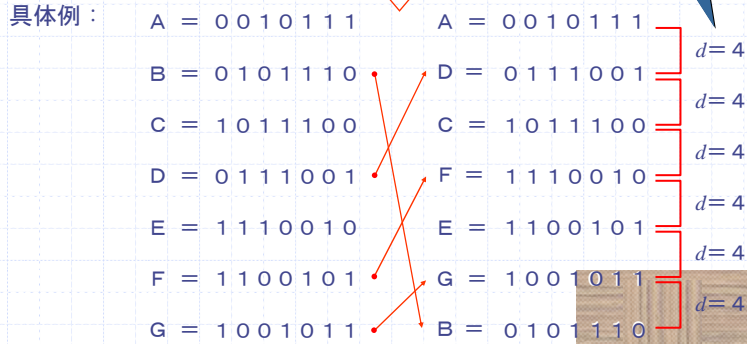
- ・ 符号001は誤りであることがわかる！
 この符号だけ、前後の符号とのハミング距離がd=1になっているから・・・
- ・ しかし、本当は何であったのかまではわからない！



§ 5.3.2 巡回符号

- ▶ 特殊ではあるが、実用的
- ▶ 作り方は、名(巡回)の示す通り

やっぱり全て、 $d=4$

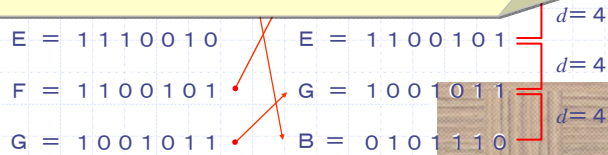


§ 5.3.2 巡回符号

- ▶ 特殊ではあるが、実用的
- ▶ 作り方は、名(巡回)の示す通り

やっぱり全て、 $d=4$

- 特徴：冗長度大 ⇨ エラーには極めて強い！
しかし、前例では、7ビットもあるのに、7事象の符号しか作れない。
- メリット：作り易い ⇨ 電子回路向き（ハードウェア構成が容易）
シフトレジスタの使用により、簡単に作れる。



§ 5.3.2 巡回符号

- ▶ 特殊ではあるが、実用的
- ▶ 作り方は、名(巡回)の示す通り

具体例：

A =	0 0 1 0 1 1 1	}	$d=4$
B =	0 1 0 1 1 1 0	}	$d=4$
C =	1 0 1 1 1 0 0	}	$d=4$
D =	0 1 1 1 0 0 1	}	$d=4$
E =	1 1 1 0 0 1 0	}	$d=4$
F =	1 1 0 0 1 0 1	}	$d=4$
G =	1 0 0 1 0 1 1	}	$d=4$

では、'種'はどう作るか？

考えてくること！
宿題！！

どの符号間も $d=4$
となっている！



ハミング距離とエラー検出、エラー訂正の関係

- 例えば、ハミング距離 d が常に「2」であるなら、符号系列において、常に1ビットの誤りを検出できる。

理由 → 1ビットの誤りがあると、 $d=1$ または $d=3$ になってしまうから。

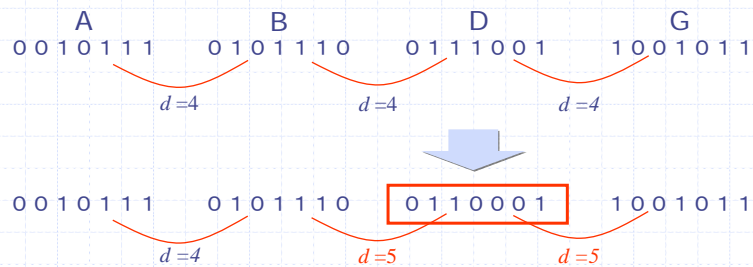
では、一般論としてはどうなるか？

d	検出可能な エラービット数	訂正可能な エラービット数
1	0	0
2	1	0
3	2	1
4	3	1
⋮	⋮	⋮
⋮	⋮	⋮
2r-1	2(r-1)	r-1
2r	2r-1	r-1



【確認】

例えば先ほどの巡回符号で、 $A \rightarrow B \rightarrow D \rightarrow G$ と送る場合について考える。



過りのある符号は、0110001. これを前出7つの符号と(ハミング距離で)比べ、その最小となる符号を選ぶことにより、正しい符号は、0111001であると判明.



クイズ.3

- ✓ 学籍番号末尾の数字が奇数の人のみ、そのまま教室に残って下さい。偶数の人は、速やかに退室して下さい。
- ✓ 教室に残った人は、隣りの人との間に一席空けて座って下さい。
- ✓ 回答時間は、15分間です。
今回は、自筆ノート等の参照を許します。
ただし、周囲の人との相談は許しません。また、無駄なことを書くと減点します。
- ✓ 終わった人は、回答用紙を教壇上の机上に提出し、他の人に迷惑をかけぬよう注意してご退出下さい。



クイズ.3

【問題】 下表のような生起確率をもつ4つの文字(A,B,C,D)に対する符号「01符号」を‘ハフマンの符号化法’によって定めなさい。そして、‘シャノンの符号化法’の結果と‘効率’を比較しなさい。

文字	生起確率
A	$P_1=1/2$
B	$P_2=1/4$
C	$P_3=1/8$
D	$P_4=1/8$

❖ シャノンの符号化法の説明で使ったものと同じ。

今回は、ノートを見て良い！



2008年度

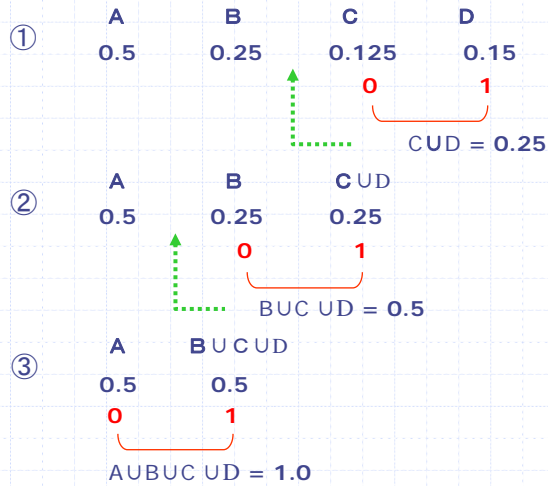
「情報工学」

第9回講義 おわり



クイズ.3

Example : 5.2.1 の場合をハフマンの符号化法で符号化してみよう.



	作られた和事象のうち、その通報を含むもの	符号	長さ
D	BUCUD=0.5(1) CUD=0.25(1) D=0.125(1)	111	3
C	BUCUD=0.5(1) CUD=0.25(1) C=0.25(0)	110	3
B	BUCUD=0.5(1) B=0.25(0)	10	2
A	A=0.5(0)	0	1

$$N(T) = N(T) + N(T-1) + 2N(T-3)$$

$$1 = W^{-1} + W^{-2} + 2W^{-3} \dots$$

偶然ではあるが、シャノンの符号化法で符号化した場合と同じになる!!!

小テスト問題 060613

議事録作成のために会議を録音する場合と音楽を高音質で録音する場合とについて比較してみよう。録音はデジタル方式とし、両者ともにサンプリング周波数 40kHz で保存された。

- (1) 保存された標本値はともに3億個であった。会議および音楽は、ともに何分間録音されたことになるか？
- (2) 一般に、会議記録程度の音声録音なら 0~5kHz の周波数情報を保存すれば良いが、高音質の音楽録音では 0~20kHz の周波数情報が完全に保存されなければならないと言われている。保存された標本値の数には、それぞれの場合に冗長があるか？ あるとすれば、3億個の中で、それぞれ何個を捨てることが可能か？
- (3) 録音結果を CD-R に保存することにした。必要な CD-R の枚数の比を、会議録音と音楽録音の場合で比較しなさい。

小テスト問題 060613

議事録作成のために会議を録音する場合と音楽を高音質で録音する場合とについて比較してみよう。録音はデジタル方式とし、両者ともにサンプリング周波数 40kHz で保存された。

- (1) 保存された標本値はともに3億個であった。会議および音楽は、ともに何分間録音されたことになるか？
- (2) 一般に、会議記録程度の音声録音なら 0~5kHz の周波数情報を保存すれば良いが、高音質の音楽録音では 0~20kHz の周波数情報が完全に保存されなければならないと言われている。保存された標本値の数には、それぞれの場合に冗長があるか？ あるとすれば、3億個の中で、それぞれ何個を捨てることが可能か？
- (3) 録音結果を CD-R に保存することにした。必要な CD-R の枚数の比を、会議録音と音楽録音の場合で比較しなさい。

頭の柔らかさのチェック

