創想館　14-201

2008-10

Tuesday　SPRING

13:00－14:30

2008年度

「情報工学」

第10回講義

Spring 2008　Prof.M.Nakajima　KEIO Univ. Yagami

前回学んだこと

復 習

- 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

✓ ‘ハフマンの符号化法’について学んだ.

☞ シャノンの符号化法は,‘効率’という観点から見て,誤りが生じない通信路においては,ベストな符号化法である！

- 符号化の手順を,必ず覚えておくこと！
- 他の符号化法（シャノンの符号化法や不作為に行った符号化の結果と‘符号化の効率’や‘符号化の冗長度’で比較できるようにしておくこと.

✓ ‘誤りが生じる通信路用の符号化法’についてを学んだ.

① パリティー検査符号

・ 奇数パリティ, 偶数パリティ, がある.

② 巡回符号

・ 冗長であるが, エラーに極めて強い. また実用的である.

ANIMATION

前回学んだこと

復 習

- 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

✓ ‘ハミング距離 $d(u_i, u_j)$’について学んだ.

- ハミング距離を用いることの意味を理解しておくこと.
- ハミング距離を用いて誤り検出や誤り訂正を行う方法を覚えておくこと.
- ハミング距離とエラー検出(ビット数)及びエラー訂正の関係(ビット数)を知っておくこと.

☞ ハミング距離一定の符号の作り方：

⇨ 専門書に譲ることにする.

ANIMATION

前回の講義の補足

- 符号化の結果は，必ずしも一様とは限らない！
- 前回の例題について，考えてみよう・・・

§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．

情報	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

前回の講義の補足

- 符号化の結果は，必ずしも一様とは限らない！
- 前回の例題について，考えてみよう・・・

[STEP.1] 生起確率大の順に並べる ⇒ 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．

A	B	C	D	E
0.3	0.25	0.20	0.15	0.10
			0	1

D U E = 0.25

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

A	B	D U E	C
0.3	0.25	0.25	0.20
		0	1

D U E U C = 0.45

前回の講義の補足

- 符号化の結果は，必ずしも一様とは限らない！
- 前回の例題について，考えてみよう・・・

［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.

DUEUC	A	B
0.45	0.30	0.25
	0	1

AUB = 0.55

後は同じ・・・

AUB	DUEUC
0.55	0.45
0	1

AUBUDUEUC = 1.0 ⇐ *最後まで行った！*

‖

AUBUCUDUE

前回の講義の補足

- 符号化の結果は，必ずしも一様とは限らない！
- 前回の例題について，考えてみよう・・・

［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

ANIMATION

前回の講義の補足

ちょっと，やり方を変えてみます！

[STEP.1]　生起確率大の順に並べる

[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．

A	B	C	D	E
0.3	0.25	0.20	0.15	0.10
			0	1

D U E ＝ 0.25

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

A	D U E	B	C
0.3	0.25	0.25	0.20
		0	1

B U C ＝ 0.45

ANIMATION

前回の講義の補足

ちょっと，やり方を変えてみます！

[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．

B U C	A	D U E
0.45	0.30	0.25
	0	1

A U D U E ＝ 0.55

後は同じ・・・

A U D U E	B U C
0.55	0.45
0	1

A U D U E U B U C ＝ 1.0

‖

A U B U C U D U E

前回の講義の補足

ちょっと，やり方を変えてみます！

[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	前回求めた最終符号
E	AUDUE=0.55(0),　DUE=0.25(1) E=0.1(1)	011	101
D	AUDUE=0.55(0),　DUE=0.25(1) D=0.15(0)	010	100
C	BUC=0.45(1),　C=0.2(1)	11	11
B	BUC=0.45(1),　B=0.25(0)	10	01
A	AUDUE=0.55(0),　A=0.3(0)	00	00

ANIMATION

前回の講義の補足

ちょっと，やり方を変えてみます！

[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	前回求めた最終符号
E	AUDUE=0.55(0),　DUE=0.25(1) E=0.1(1)	011	101
D	AUDUE ... 25(1)	010	100
C		11	11
B	BUC=0.45(1),　B=0.25(0)	10	01
A	AUDUE=0.55(0),　A=0.3(0)	00	00

どちらも‘符号化の効率’は同じ！

先週出題のクイズ(クイズ.3)について

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!

Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami

クイズ.3の略解

① A B C D
0.5 0.25 0.125 0.125
0 1
CUD = 0.25

② A B CUD
0.5 0.25 0.25
0 1
BUCUD = 0.5

③ A BUCUD
0.5 0.5
0 1
AUBUCUD = 1.0

ANIMATION

クイズ.3の略解　　・・・つづき

	作られた和事象のうち、その文字を含むもの	符号	符号の長さ	シャノンの符号化の結果
D	B∪C∪D=0.5 (1) C∪D=0.25 (1) D=0.125 (1)	111	3	111
C	B∪C∪D=0.5 (1) C∪D=0.25 (1) C=0.25 (0)	110	3	110
B	B∪C∪D=0.5(1) B=0.25(0)	10	2	10
A	A=0.5(0)	0	1	0

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った！！！

すなわち,

$$C = 1(bit/sec)\ ,\quad \eta = 1$$ （符号化の効率 100%）

ANIMATION

§6 情報の伝送　（その2）

§6.1 マルコフ過程の符号容量

§6.2 誤りがある場合の符号容量

かなり難しい'話'だ！

§6.1 マルコフ過程の符号容量

これまでの話では、文字(通報)の生起に関して、相互に何の制約もなかった.

この先は、文字間に制約がある場合の符号容量について考える.

$p(j/i)$ を考える必要がある場合・・・

memo

エントロピーの計算では・・・

単なる確率過程の場合：　$H = -\sum p(i)\log_2 p(i)$

単純マルコフ過程の場合：　$H_{M1} = -\sum p(i)p(j/i)\log_2 p(j/i)$

まず、確率的性質が、下図のようなシャノン線図で表される A, B の文字列を考える.

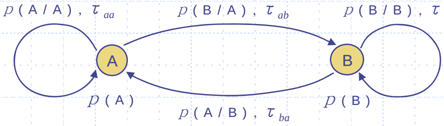

Fig.6.1

ANIMATION

各文字 A , B の 01符号を, それぞれ「0」,「10」とし,

文字 A , B によって作られた文字列(文章)を,

A B A A B A B B A B A A A B A B B A・・・とすると,

0 10 0 0 10 0 10 10 0 10 0 0 0 10 0 10 10 0・・・

2 1 1 2 1 2 2 1 2 1 1 1 2 1 2 2 1

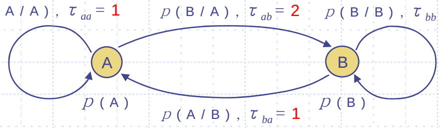

Fig.6.1

☞（注）実際には‘文字間スペース’,‘単語間スペース’が必要となるが,
これまでと同様, ここでもそれらを無視して話を進める.

T 秒間に送れる「A」の個数は,　$N_a(T) = N_a(T-\tau_{aa}) + N_b(T-\tau_{ba})$

T 秒間に送れる「B」の個数は,　$N_b(T) = N_a(T-\tau_{ab}) + N_b(T-\tau_{bb})$

ANIMATION

A B AA B A B B A B AAA B A B B A・・・

0 10 0 0 10 0 10 10 0 10 0 0 0 10 0 10 10 0・・・

2 1 1 2 1 2 2 1 2 1 1 1 2 1 2 2 1

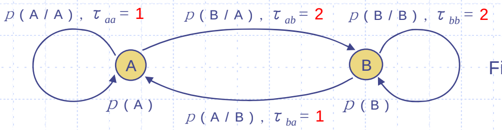

Fig.6.1

T 秒間に送れる「A」の個数は, $N_a(T) = N_a(T-\tau_{aa}) + N_b(T-\tau_{ba})$

T 秒間に送れる「B」の個数は, $N_b(T) = N_a(T-\tau_{ab}) + N_b(T-\tau_{bb})$

よって, $N_a(T) = N_a(T-1) + N_b(T-1)$　（１）

$N_b(T) = N_a(T-2) + N_b(T-2)$　（２）

ここで、ある情報を文字 A,B の組み合わせによって送る場合を考える. 1秒（正しくは, 単位時間）間以内に送れ「A」で終わる組合せは、「A」のみで、「B」で終わる組合せはないが、2秒あれば、「A」で終わる組合せは、「AA」,「B」で終わる組合せは「B」で共に1個となる. さらに3秒あれば、「A」で終わる組合せは、「AAA」,「BA」で2個, 「B」で終わる組合せは, 「AB」で1個ということになる.

ここで、ある情報を文字 A,B の組み合わせによって送る場合を考える. 1秒（正しくは, 単位時間）間以内に送れ「A」で終わる組合せは、「A」のみで、「B」で終わる組合せはないが、2秒あれば、「A」で終わる組合せは、「AA」,「B」で終わる組合せは「B」で共に1個となる. さらに3秒あれば、「A」で終わる組合せは、「AAA」,「BA」で2個, 「B」で終わる組合せは, 「AB」で1個ということになる. これを表にすると・・

T (sec)	「A」で終わる文章	N_0 (T)	「B」で終わる文章	N_1 (T)
1	A	1		0
2	AA	1	B	1
3	AAA, BA	2	AB	1
4				
5				

Table-6.1

ここで、ある情報を文字 A,B の組み合わせによって送る場合を考える． 1秒（正しくは，単位時間）間以内に送れ「A」で終わる組合せは、「A」のみで、「B」で終わる組合せはないが、2秒あれば、「A」で終わる組合せは、「AA」,「B」で終わる組合せは「B」で共に1個となる. さらに3秒あれば、「A」で終わる組合せは、「AAA」,「BA」で2個,「B」で終わる組合せは,「AB」で1個ということになる． これを表にすると・・

T (sec)	「A」で終わる文章	N_0 (T)	「B」で終わる文章	N_1 (T)
1	A	1		0
2	AA	1	B	1
3	AAA, BA	2	AB	1
4	AAAA, ABA, BAA	3	AAB, BB	2
5	AAAAA, AABA, ABAA, BAAA, BBA	5	AAAB, ABB, BAB	3

Table-6.1

ANIMATION

§6.1 マルコフ過程の符号容量

T (sec)	「A」で終わる文章	N_0 (T)	「B」で終わる文章	N_1 (T)
1	A	1		0
2	AA	1	B	1
3	AAA, BA	2	AB	1
4	AAAA, ABA, BAA	3	AAB, BB	2
5	AAAAA, AABA, ABAA, BAAA, BBA	5	AAAB, ABB, BAB	3

Table-6.1

Table-6.1１ が式（1）、（2）を満たしているかどうか、チェックしてみよう！

T=3 時

$$\begin{cases} \underset{2}{N_a(3)} = N_a(3-1) + N_b(3-1) = \underset{1}{N_a(2)} + \underset{1}{N_b(2)} \\ \underset{1}{N_b(3)} = N_a(3-2) + N_b(3-2) = \underset{1}{N_a(1)} + \underset{0}{N_b(1)} \end{cases}$$

OK!

ANIMATION

T=4 時

$$\begin{cases} \underset{3}{N_a(4)} = N_a(4-1)+N_b(4-1) = \underset{2}{N_a(3)}+\underset{1}{N_b(3)} \\ \underset{2}{N_b(4)} = N_a(4-2)+N_b(4-2) = \underset{1}{N_a(2)}+\underset{1}{N_b(2)} \end{cases}$$

OK!

T=5 時

$$\begin{cases} \underset{5}{N_a(5)} = N_a(5-1)+N_b(5-1) = \underset{3}{N_a(4)}+\underset{2}{N_b(4)} \\ \underset{3}{N_b(5)} = N_a(5-2)+N_b(5-2) = \underset{2}{N_a(3)}+\underset{1}{N_b(3)} \end{cases}$$

OK!

さて、それでは（１）、（２）式を連立方程式として解いてみよう.

解の形として、$N_a(T) = A_a W^T$、$N_b(T) = A_b W^T$ を考え、

（１）、（２）式に代入すると、

$$A_a W^T = A_a W^{T-1} + A_b W^{T-1}$$
$$A_b W^T = A_a W^{T-2} + A_b W^{T-2}$$

両辺を W^T で割って整理すると、

$$A_a (W^{-1} - 1) + A_b W^{-1} = 0$$
$$A_a W^{-2} + A_b (W^{-2} - 1) = 0$$

すなわち、$\begin{vmatrix} W^{-1}-1 & W^{-1} \\ W^{-2} & W^{-2}-1 \end{vmatrix} = 0$

よって、 $W^{-2} + W^{-1} - 1 = 0$

$W^{-1} = X$ とすると、 $X^2 + X - 1 = 0$

この方程式の最大根は、 $X = \dfrac{-1+\sqrt{5}}{2}$

よって、 $W = X^{-1} = \dfrac{\sqrt{5}+1}{2}$

よって、$C = log_2 \dfrac{\sqrt{5}+1}{2} \cong 0.69 \quad bit/sec$

Example：6.1

Table-6.2 に示されるように、3つの符号「A」、「B」、「C」があり、それらを送るのに必要な時間は、其々2、1、1秒である.

❖ 各符号として, 例えば A:01, B:0, C:1 等をイメージ

（1） 符合間の遷移には、特段の制約がないものとして、通信速度の最大値（符号容量 C ）を求めよ.

（2） 符合間の遷移が、左図 Fig.6.2 のようになるとして符号容量 C を求めよ.

符号	A	B	C
τ_{ij} (sec)	2	1	1
p_i	p_a	p_b	p_c

Table-6.2

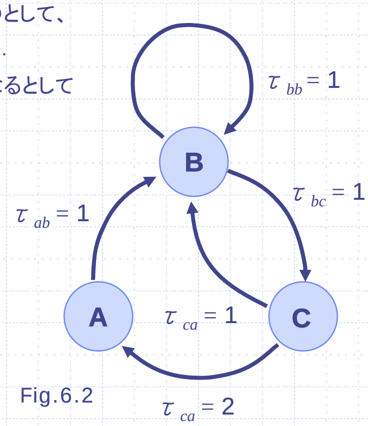

Fig.6.2

【 略 解 】

（1）単なる確率過程（生起確率のみ）を考えるのであれば ・・・

$$W^{-2}+W^{-1}+W^{-1}=1$$

$$X=W^{-1} \quad \text{として} \quad X^2+2X-1=0$$

$$X=-1\pm\sqrt{2}\Rightarrow -1+\sqrt{2}$$

よって、

$$C=log_2 W=log_2\frac{1}{X}=log_2\frac{1}{-1+\sqrt{2}}=log_2\sqrt{2}+1$$

$$\cong 1.27 \quad bit/sec$$

また、この Cを与える各符号の生起確率は、

$$p_a=2^{-C\tau_a}=2^{-1.27\times 2}.\ .\ .\ .\ .=0.17$$

$$p_b=2^{-C\tau_b}=2^{-1.27\times 1}.\ .\ .\ .\ .=0.41$$

$$p_c=2^{-C\tau_c}=2^{-1.27\times 1}.\ .\ .\ .\ .=0.41$$

（2）次に遷移確率まで考えると（マルコフ過程）・・・

$$N_a(T) = N_c(T-2)$$
$$N_b(T) = N_a(T-1) + N_b(T-1) + N_c(T-1)$$
$$N_c(T) = N_b(T-1)$$

この連立方程式を解くことになる ・・・

まず、解の形として　$N_a(T) = A_a W^T$、$N_b(T) = A_b W^T$、$N_c(T) = A_c W^T$　を考え、

上記方程式に代入

$$\begin{cases} A_a W^T = A_c W^{T-2} \\ A_b W^T = A_a W^{T-1} + A_b W^{T-1} + A_c W^{T-1} \\ A_c W^T = A_b W^{T-1} \end{cases}$$

$\div W^T$

$$\begin{cases} -A_a + A_c W^{-2} = 0 \\ -A_b + A_a W^{-1} + A_b W^{-1} + A_c W^{-1} = 0 \\ -A_c + A_b W^{-1} = 0 \end{cases}$$

§6.1 マルコフ過程の符号容量

よって、

$$\begin{vmatrix} -1 & 0 & W^{-2} \\ W^{-1} & W^{-1}-1 & W^{-1} \\ 0 & W^{-1} & -1 \end{vmatrix} = 0$$

$$\therefore\ W^{-4} + W^{-2} + W^{-1} - 1 = 0$$

$X = W^{-1}$　として、　$X^4 + X^2 + X - 1 = 0$

この最大根は、　$X \cong 0.6$　　電卓で探す！

よって、

$$C = \log_2 W = \log_2 \frac{1}{X} = \log_2 \frac{1}{0.6} \cong 0.74 bit/\sec$$

Spring 2008　Prof.M.Nakajima　KEIO Univ. Yagami
ANIMATION

クイズ.4

- ✓ 学籍番号末尾の数字が偶数の人のみ，そのまま教室に残って下さい．奇数の人は，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．

ANIMATION

クイズ.4

〔問題〕3つの符号 A,B,C があり，各符号の長さは 1,2,2（単位時間）である．ある情報をこれらの符号の組み合わせによって送る場合について，次に問に答えなさい．

（1）８単位時間内に送れる符号 A, B, C の組み合わせ総数を求めなさい．

（2）次に，この系における通信速度の最大値を求めなさい．

（3）また，その場合の各符号の生起確率を求めなさい．

今回も，ノートを見て良い！

2008年度

「情報工学」

第10回講義　おわり