

マルチメディア符号化と信号処理(1)

- 情報理論
 - ◆ 情報源符号化、エントロピー
 - ◆ 情報源符号化定理:「情報源符号化の効率の限界と、その限界に任意に近い効率を持つ情報源符号の存在を保証する。」
 - ◆ ハフマン符号(コンパクト符号)
 - ◆ 情報量とエントロピー
 - ◆ 通信路符号化、通信路容量、通信路符号化定理
 - ◆ 符号理論、誤り検出符号、誤り訂正符号

1

IT(情報技術)世界の標準語は“0”と“1”

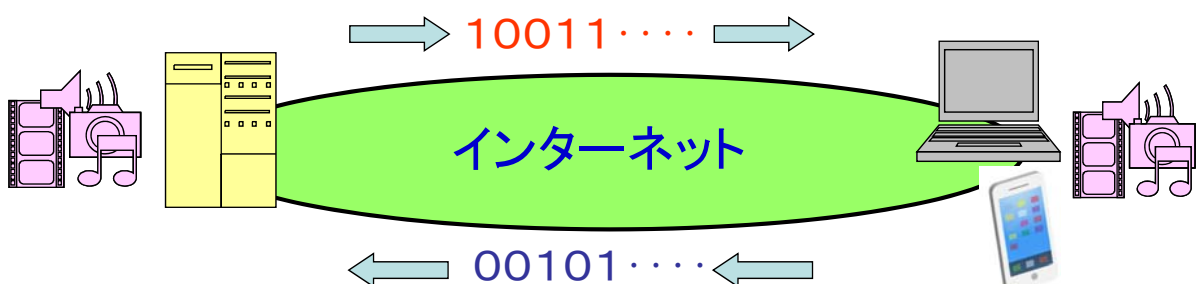
コンピュータやネットワークの中では、すべての情報が“0”と“1”で記憶され、計算され、伝えられる。

例えば、A ⇔ 01000001

% ⇔ 00100101

.....

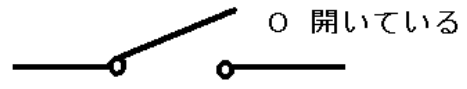
単位は
ビット



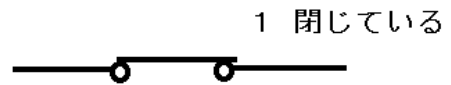
2

コンピュータの内部では、

“1” ↔ 素子が活性化した状態
(電圧や電流や磁気がある状態)



“0” ↔ 素子が鎮静化した状態
(電圧や電流や磁気がない状態)

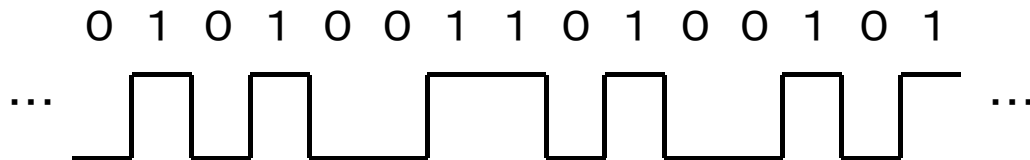


スイッチ回路

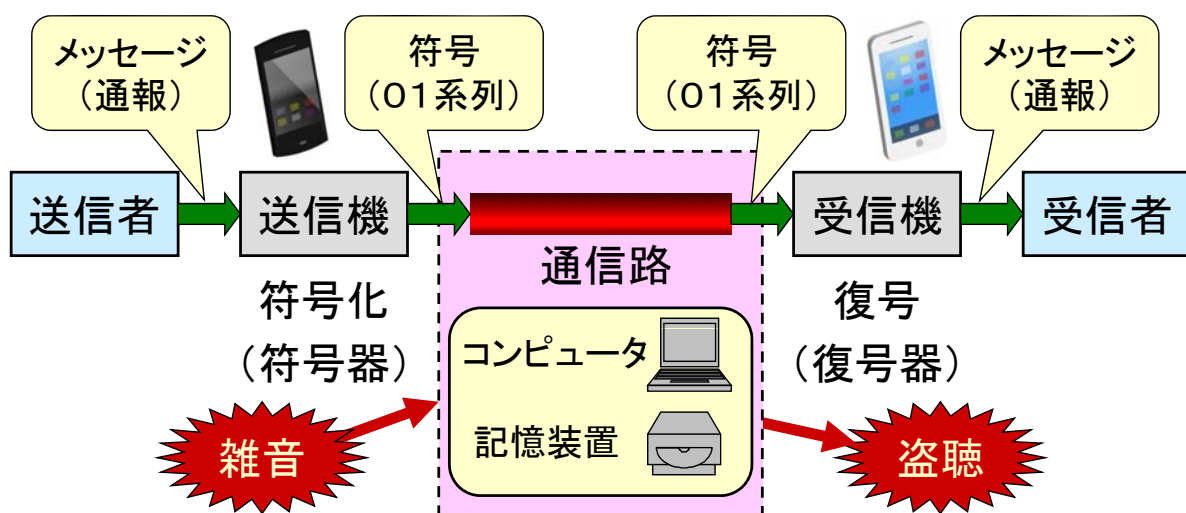
情報を伝える(伝送する)場合は、

“1” ↔ パルスを送る、 正のパルスを送る、...

“0” ↔ パルスを送らない、負のパルスを送る、...



情報の符号化 — 情報伝送のモデル —



私たちがITとうまく付き合うためには、人間の言葉を“0”と“1”だけで綴られるITの言葉にじょうずに翻訳しなければならない。

じょうずな符号化とは？

【条件1】 ～符号化の効率の追求～

なるべく“0”と“1”の系列の長さが短くなるように符号化する。

➡ 一定の時間内に送れる符号の数が一定なら、ある情報を表す“0”と“1”の系列が短いほど、短時間にたくさんの情報を伝えることができる。

【条件2】 ～一意的に復号可能～

メッセージが私たちの言葉(文字や数字)で表されている場合には、その境目が分かるようにくふうする。


➡ “0”と“1”の系列から一意的にもとのメッセージを復元できる(もとのメッセージが一つに確定される)。

5

(例) 成績 A, B, C, D を符号化する。

記号	符号 C_1	符号 C_2	符号 C_3
A	0	00	0
B	1	01	10
C	10	10	110
D	11	11	111

符号 C_1 の場合

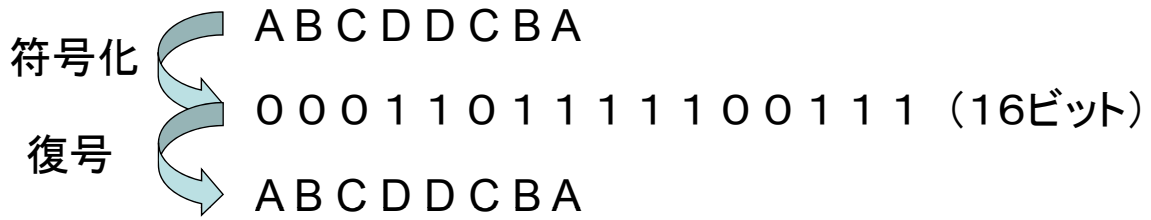
符号化  ABCDDCBA
01101111010 (12ビット)

復号  A D A B D B C C
A B B A B B D B A C
.....

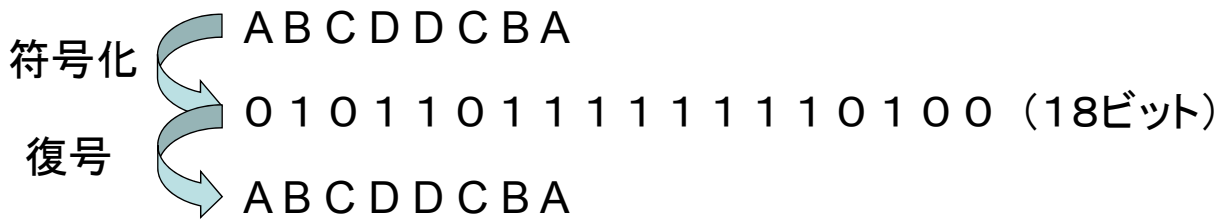
条件2を満たさない
符号としては失格

6

符号C₂の場合

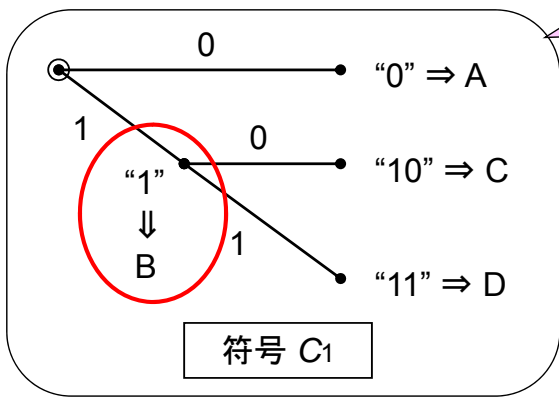


符号C₃の場合



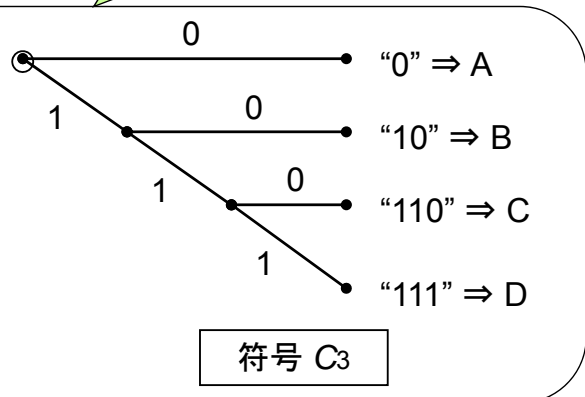
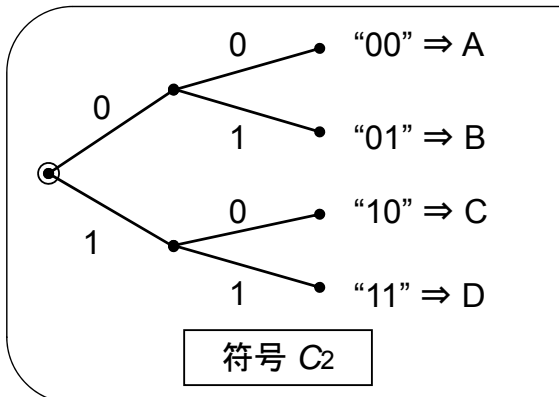
条件2を満たす(一意的に復号可能)

符号の木



符号の木の葉以外の節点にも符号(記号)が割り当てられている
⇒ 一意的に復号可能でない

符号の木のすべての葉に符号(記号)が割り当てられている
⇒ 一意的に復号可能である



条件1の符号化の効率について考えよう（その1）

➡ 平均符号長 L (ビット/記号)・・・符号の長さの期待値

記号	出現確率	符号 C_2	符号 C_3
A	0.25	00	0
B	0.25	01	10
C	0.25	10	110
D	0.25	11	111

符号 C_2 の場合

$$L = 0.25 \times 2 + 0.25 \times 2 + 0.25 \times 2 + 0.25 \times 2 = 2.0$$

符号 C_3 の場合

$$L = 0.25 \times 1 + 0.25 \times 2 + 0.25 \times 3 + 0.25 \times 3 = 2.25$$

9

条件1の符号化の効率について考えよう（その2）

➡ 平均符号長 L (ビット/記号)・・・符号の長さの期待値

記号	出現確率	符号 C_2	符号 C_3
A	0.5	00	0
B	0.25	01	10
C	0.125	10	110
D	0.125	11	111

符号 C_2 の場合

$$L = 0.5 \times 2 + 0.25 \times 2 + 0.125 \times 2 + 0.125 \times 2 = 2.0$$

符号 C_3 の場合

$$L = 0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3 + 0.125 \times 3 = 1.75$$

10

(その1)

記号	A	B	C	D
出現確率	0.25	0.25	0.25	0.25

メッセージ ABCDDCBA

符号C₂ 0001101111100111 (16ビット)

符号C₃ 0101101111111110100 (18ビット)

(その2)

記号	A	B	C	D
出現確率	0.5	0.25	0.125	0.125

メッセージ AABCDDBAA

符号C₂ 0000011011010000 (16ビット)

符号C₃ 00101101111000 (14ビット)

11

じょうずな符号化とは？

【条件1】 ～符号化の効率の追求～

なるべく“0”と“1”の系列の長さが短くなるように符号化する。



記号(文字)の出現確率に応じて、長さの異なる符号を割り当て、全体の符号の長さを短縮する(情報圧縮)。
⇒ 可変長符号化

出現確率が高い記号(文字)には短い符号を、出現確率が小さい記号(文字)には長い符号を割り当てる！

最も効率的な符号(平均符号長が一番短い符号: **ハフマン符号**)の作り方

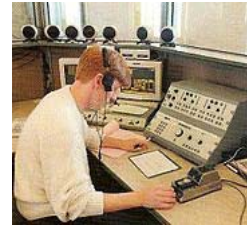
「情報理論・確率論」(2年前期)の講義の中で学びました。

12

モールス符号

電流や光や音などを断続させて情報を伝えるための符号

原始的だけど、しかし原始的だからこそ、最後の通信手段として役に立つ！



SOS(助けてくれ)

--- — — — --- (トントントン ツーツーツ トントントン)

1つのモールス符号は、だいたい1~6個の短点(点)と長点(線)の組み合わせから構成されている。

(例) S: 3個の短点(点)、O: 3個の長点(線)

このモールス符号の構成については、法令(無線局運用規則)で定められている。

欧文モールス符号表(無線局運用規則別表第1号)

1 文字	2 数字	3 記号
·- A	·- - - -	
- - - - B	· · · - -	
- - - - C	· · · - -	
- - - D	· · · - -	
· E	· · · · ·	
· · · F	· · · · ·	
- - - G	· · · · ·	
· · · H	· · · · ·	
· · I	· · · · ·	
· - - - J	· · · · ·	
- - - K	· · · · ·	
· · · L	· · · · ·	
- - M	· · · · ·	
· · N	· · · · ·	
- - - O	· · · · ·	
· - - - P	· · · · ·	
- - - - Q	· · · · ·	
· · · R	· · · · ·	
· · · S	· · · · ·	
- T	· · · · ·	
· · · U	· · · · ·	
· · · V	· · · · ·	
· - - W	· · · · ·	
· · · X	· · · · ·	
- - - Y	· · · · ·	
- - - Z	· · · · ·	

[モールス電信の符号の線及び間隔]

- ① 1線の長さは、3点に等しい。
- ② 1符号を作る各線又は点の間隔は、1点に等しい。
- ③ 2符号の間隔は、3点に等しい。
- ④ 2語の間隔は、7点に等しい。

モールス符号の構成

短点の構成	·	E
長点の構成	-	T
1符号の構成	· - - -	C
2符号の構成	· - - - · - - -	CQ
2語の間隔	· - - - · - - - · - - -	I am

[参考文献] 電信教則本(CQ出版社)

モールス符号の性能

～効率的な符号化という見地から～

英文におけるアルファベットの出現確率は下表のようになっている。モールス符号は効率的な符号になっているでしょうか？

英文におけるアルファベットの出現確率(言語のくせ)

文字 (出現確率%)

a (8.2)	b (1.5)	c (2.8)	d (4.3)	e (12.7)	f (2.2)	g (2.0)	h (6.1)
i (7.0)	j (0.2)	k (0.8)	l (4.0)	m (2.4)	n (6.7)	o (7.5)	p (1.9)
q (0.1)	r (6.0)	s (6.3)	t (9.1)	u (2.8)	v (1.0)	w (2.4)	x (0.2)
y (2.0)	z (0.1)						

デジタル画像の符号化



モノクロ画像

デジタル画像
平面上の各点(画素)に濃淡(濃度値)あるいは色情報(RGB)を数値で指定することによって表現される

23	29
27	45

← 黒 白 →



0 127 255

(8ビット表現)



カラー画像

40	38
44	46

R(赤)

12	24
18	22

G(緑)

88	73
93	77

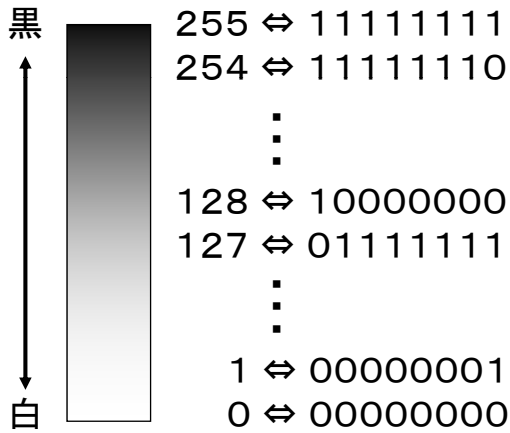
B(青)

デジタル画像(モノクロの場合)

1バイト(B)

1 KB = 1024 B
1 MB = 1024 KB

1画素の濃淡値(輝度値)を8ビットで表現する(256階調)



1024 × 1024画素(約100万画素)のデジタル画像の容量は?

1024 × 1024 × 8 (ビット、bit)
= 1024 × 1024 (バイト、B(Byte))
= 1024 × 1 (キロバイト、KB)
= 1 (メガバイト、MB)



フロッピーディスク
約1枚分の容量

256MBのSDメモリでは何枚?



画素の濃淡値の出現回数(出現確率)に偏りがあれば、全体の符号の長さを短縮できる(情報圧縮できる)。

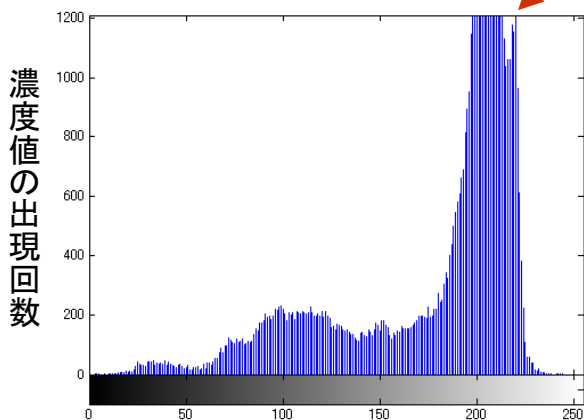
画像 Airplane



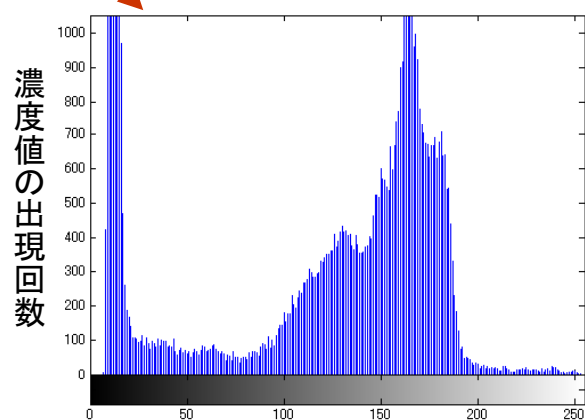
画像 Cameraman



濃淡値をどのように符号化すれば、画像の容量を小さくできますか?

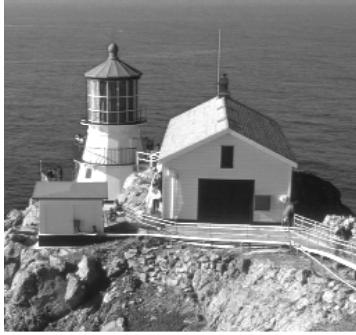


画素の濃淡値(輝度値)



画素の濃淡値(輝度値)

画像 Lighthouse



画像 Woman



濃淡値をどのように符号化すれば、画像の容量を小さくできますか？

