

## 「情報の表現 2」

### 1. ビットとバイト

#### ビット (bit)

情報の最小単位

2進数字 (binary digit) の略

2進数の1桁が1ビット

1ビットで2通りの状態 (0か1か) が表現できる

#### バイト (byte)

8ビットで1バイト (B)

1000バイトが1キロバイト (KB)

1000キロバイトが1メガバイト (MB)

1000メガバイトが1ギガバイト (GB)

1000ギガバイトが1テラバイト (TB)

※1024B=1KB, 1024KB=1MB…という換算もある

### 2. 文字コード

文字をコンピュータで符号 (ビット列) として扱えるようにしたもの。

英数字は1バイト (最大 256 文字) で表現する

ASCII コードが世界標準。

文字種の多い文字 (漢字など) は2バイト (最大 65,536 文字) で表現する。

日本語には JIS コード, EUC コード, シフト JIS コード, Unicode などが使用される。

### 3. ASCII コード

1 バイト (8ビット) だが, 最上位の1ビットを除き, 7 ビット (128 文字) でコード化する。

最上位の1ビットは通信時の誤り検出 (パリティ) に用いられる。

通常の文字 (数字, アルファベット, 記号), 制御文字に対応している。

大文字と小文字は区別される。

数字も文字としてコード化されている。

制御文字とは, たとえば, ESC, BS, DEL など。

7ビットを上位3ビットと下位4ビットに分けている。

例：

Ryukoku を 16 進数の ASCII コードで表すと  
52 79 75 6B 5F 6B 75

### ASCII コード表

上位 3 ビット

	0	1	2	3	4	5	6	7
0	NUL	DLE	SP	0	@	P	~	p
1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3	#	3	C	S	c	s
4	EOT	DC4	\$	4	D	T	d	t
5	ENO	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
B	VT	ESC	+	;	K	[	k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M	]	m	}
E	SO	RS	.	>	N	^	n	~
F	SI	US	/	?	O	_	o	DEL

下位 4 ビット

#### 4. 日本語の文字コード

**ISO-2022-JP (通称 JIS コード)**

電子メールで利用されている。

**EUC**

Unix というシステムで利用される。

**SHIFT\_JIS**

パソコンで利用される。

**Unicode**

UTF-8, UTF-16 など。多言語化に対応。

※このようなコードの対応の誤りなどにより、いわゆる「文字化け」が起こる。

## 5. アナログとデジタルの意味

### アナログ (analog) の本来の意味

analogy は類似・相似を意味する.

例: 温度計 (温度を「長さ」や「角度」で表す)

### デジタル (digital) の本来の意味

digit は本来「指」という意味であり, 指で数字を数えることから「数字で表す」ことを示すようになった.

※表示 (表現) と内部の処理とは分けて考える必要がある.

### アナログ表示とデジタル表示

ある量 (例: 温度) が長さや角度などで表示されていればアナログ表示であり, 数字で表示されていればデジタル表示.

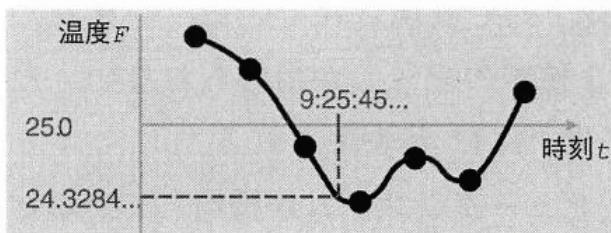
### アナログ処理とデジタル処理

ある量が電圧・電流などの物理量で表され, そのまま処理されるのがアナログ処理であり, 数値に変換されて処理されるのがデジタル処理.

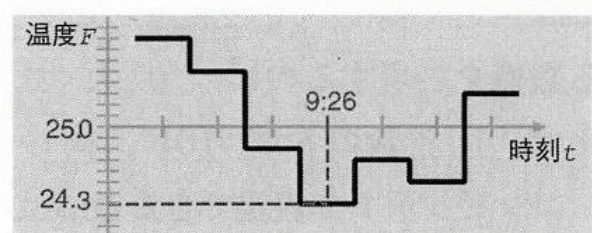
## 6. アナログ量とデジタル量

アナログ信号は, 時間的にも数値的にも連続的に変化する.

デジタル信号は, 時間的にも数値的にも離散的 (とびとびの値) に変化する.



(a)



(b)

アナログ量 (a) とデジタル量 (b) の対比

## 7. アナログからデジタルへの変換方法

### 標本化 (サンプリング)

元の連続信号を一定間隔ごとに抽出する.

サンプリングする時間間隔をサンプリング周期, 周波数をサンプリング周波数という.

### 量子化

サンプリングした信号の値を一定間隔ごとに表現 (離散的な値に近似) する.

量子化をどのぐらいの段階 (範囲) で行うかをビット深度という

(例: 8 ビット, 12 ビット, 16 ビットなど)

## 8. どのぐらいの時間間隔でサンプリングすればよいか？

### 標本化定理

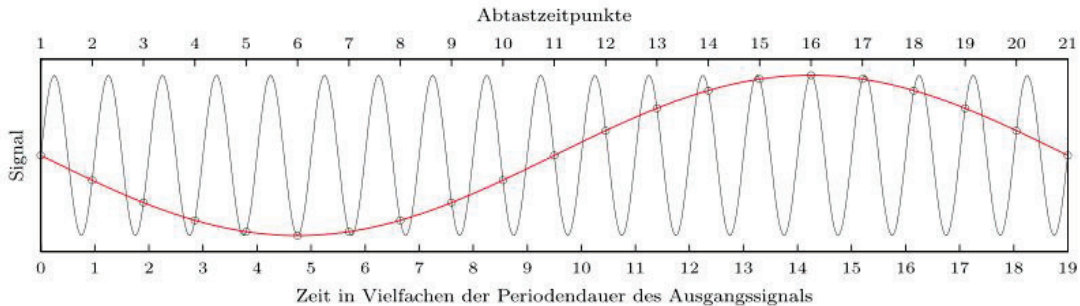
元の信号に含まれる最高周波数の2倍よりも高い周波数（周期でいえば半分より短い周期）でサンプリングすれば、元の信号は必ず再現できる。

もし元の信号に、サンプリング周波数の1/2よりも高い周波数成分が含まれていると、本来は存在しない信号（エイリアス）が再現されてしまう。

そのような成分が含まれないように、あらかじめフィルタにより除去しておく必要がある。

### エイリアス信号の例：

元の信号の最高周波数の2倍よりも低い周波数でサンプリングしているため、元の正弦波とは違う波形（赤線）が再現されている。



## 9. CDの規格

### サンプリング周波数：44.1kHz

標本化定理により、元の音源の最高周波数は22.05kHzよりも低いことになる。

これは人間の聴覚の特性（20kHzぐらまでしか聞こえない）に合っている。

### ビット深度：16bit

16ビットは65,536段階であり、最も小さな音と最も大きな音の違いが65,536倍ということになる。

これは人間の聴覚の特性（ $10^6$ 倍程度の範囲が聞き取れる）よりもやや狭い範囲である。

※CDでは1秒分の音楽データを44,100個に分割し、1個あたり16ビットで記録しているということ。

### 74分間のCD音源のデータ量は？

44.1kHz、16ビット、ステレオなので、1秒分のデータ量は  $44,100 \times 16 \times 2 \div 8 = 176,400$  バイトである。

これが50秒間だと  $176,400 \times 50 = 8,820,000$  バイトとなり、約8.8MBとなる。

74分間（CDの規格上の容量）だと約783MBである。

（注：1024B = 1KB、1024KB = 1MBという換算では、約747MBとなる。）

## 10. AD変換とDA変換

アナログからデジタルへの変換をアナログ→デジタル変換（AD変換）という。

デジタルからアナログへの変換をデジタル→アナログ変換（DA変換）という。

人間が直接認識できるのはアナログ信号なので、デジタル→アナログ変換が必要になる（たとえば、CDやDVDの再生）。

身の回りのデジタル機器（パソコン、iPod、携帯電話、デジカメ…）には、DA変換（場合によってはAD変換も）の回路が含まれている。

## 11. デジタル処理の特徴

### デジタル処理の利点

ノイズ（雑音）に強い（ノイズが入りにくい，入ったノイズを検出・修正する方法がある）．  
完全なコピーを作成できる．

### デジタル処理の欠点

標本化の分解能，量子化の精度に依存する．

著作権の問題が生じる．

記録メディアによっては意外と寿命が短く，一気にデータが失われる可能性がある．