

第2章 コンピュータと情報

第0節 位取記数法

第1節 コンピュータで扱う情報

第2節 コンピュータがはたらくしくみ(後日)

第3節 さまざまな情報の表現

データのデジタル化の方法

コンピュータはデータをどんな規則でデジタル化するか？ その方法を個別に学習する。

基本は
1対1対応

数値のデジタル表現

固定小数点表示(整数), 浮動小数点表示(実数)

文字のデジタル表現

文字コード, 1バイト文字, 2バイト文字, (フォント)

音声のデジタル表現

標本化, 量子化, 符号化, D-A変換, A-D変換

画像のデジタル表現

標本化, 量子化, 符号化, 解像度, 3原色, 静止画, 動画, ペイント系, ドロー系

データの容量と圧縮

圧縮, 可逆性, 非可逆性, 解凍(伸張)

ファイル形式と拡張子

品質, 特徴

古くから使われている文字集合

関連: pp.42-43

ASCII

JIS X 0201

	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0000	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F
0001	10	11	12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	1F
0010	20	21	22	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F
0011	30	31	32	33	34	35	36	37	38	39	3A	3B	3C	3D	3E	3F
0100	40	41	42	43	44	45	46	47	48	49	4A	4B	4C	4D	4E	4F
0101	50	51	52	53	54	55	56	57	58	59	5A	5B	5C	5D	5E	5F
0110	60	61	62	63	64	65	66	67	68	69	6A	6B	6C	6D	6E	6F
0111	70	71	72	73	74	75	76	77	78	79	7A	7B	7C	7D	7E	7F

これだけなら
8bit(1Byte)で
足りる。
(7bitで十分)

問い: 表を用いて **Kobe** を, 16進数と2進数で表そう(ノートに)。

問い: 表を用いて
Kobe を,
16進数と
2進数で表そう。
答:

K 4B 0100 1011
o 6F 0110 1111
b 62 0110 0010
e 65 0110 0101

<http://www.7b.biglobe.ne.jp/~akane/other/ascii.htm>より

文字のデジタル化:用語の定義

関連: pp.42-43

よく使われる専門用語の定義

- ① **文字集合** (character set):
文字の一覧。**符号化文字集合** (coded character set) 等ともいう。
- ② **文字符号化方式** (character encoding scheme):
文字集合①に含まれる各文字に1対1で対応づけた記号を,
データ列(**ビット列**)に変換する方法(規則)。
- ③ **文字コード** (character code):
文字符号化方式②で文字に対応したビット列。
- ④ **文字コード体系** (character encoding system):
文字コードの一覧。単に**文字コード**と呼ぶことも多い(紛らわしい)。
- ⑤ **エンコーディング** (符号化処理):
文字等のデータをビット列に置き換える処理。
逆に元に戻すことを, **デコーディング**という。

補足: 紛らわしい専門用語の解説

関連: pp.42-43

- ①文字集合(符号化文字集合)
- ②文字符号化方式
- ③文字コード
- ④文字コード体系, 等

A: 人間用の符号	B: 文字例	C: コンピュータ用符号 (=数値)
1-4-73	ら	1000 0010 1110 0111 (82E7)
1-4-74	り	1000 0010 1110 1000 (82E8)
1-4-75	る	1000 0010 1110 1001 (82E9)

- JIS規格等, 「A」とBの対応表」が複数種類存在する。
「①文字集合」とは, これら「A」とBの対応表」のことである。
- コンピュータは数値しか扱えないから「B」とCの対応表」も必要である。
「④文字コード体系」とは, この「B」とCの対応表」のことである。
- ④を作る際の(すなわち数値化するための)規則が②であり, ②の規則に沿ってきた文字ごとの数値が③である。

文字コード体系の例

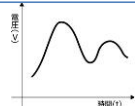
関連: pp.42-43

(符号化)文字集合 (Coded) character set	文字符号化方式 character encoding scheme	文字コード, 文字コード(体)系 character code, character encoding system
ASCII	なし	ASCII 7bit
JIS X 0201 & JIS X 0208	ISO-2022-JP Shift_JIS	JIS(漢字)コード 1,2Byte(7bit表現) Shift_JISコード 1,2Byte
ISO-2022-JP: 半角の片仮名含まず	EUC	EUC-JPコード 1,2Byte
Unicode	UTF-8, UTF-8N	UTF-8 ASCII部分: 1Byte 他: 2-6Byte Webで使用
	UTF-16BE, UTF-16LE	UTF-16 2Byte or 4Byte
	UTF-32BE, UTF-32LE	UTF-32 4Byte固定長

音声のA-D変換(実習とまとめ)

関連: p46-47

実習: デジタル化 ⇒ データ量を計算



1秒に何回標本化?
何段階?



標本化定理 もとの音の最高周波数の2倍以上で標本化。
人の可聴帯域 約20Hz～20,000Hz。

音楽CD サンプル周波数: 44,100Hz (1秒間に44100回)
量子化ビット数: 16bit = $2^{16} = 65536$ 段階 (0～65535)
チャンネル数: 2ch (マイク2つ)

階調(まとめと)

関連: p44-45

● 階調 濃淡の段階

世の中でよく使われるのは...

- 赤 緑 青: 各8bit (2^8 通り)
すなわち 256階調 (0～255)
- $256 \times 256 \times 256 =$ 約1678万色

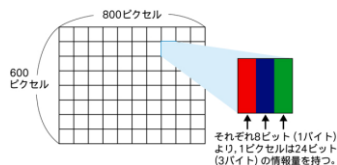
これを
フルカラー(24ビットカラー)
という。



画像と文字のデータ量比較

関連: p44-45

例: 800×600 ピクセルの
フルカラー画像の
データ量について
考えてみよう。



(計算を簡単にするために、接続辞は1024ではなく1000を使って)

- ① 容量1.44MBのフロッピーディスク(FD)に、 800×600 ピクセルの(上例の)画像は、約何枚くらい保存できると考えられるか。
- ② FDに、1ページあたり40字×18行=720字の日本語(2バイト文字)が印刷された文庫本の文字情報は、約何ページ保存できると考えられるか。
- ③ コンピュータにとって、どちらの処理がしんどそう??

3次元コンピュータグラフィックス(3DCG)

3DCG: Three-Dimensional Computer Graphics

モデリング(modeling)

形状をつくること。
(面をポリゴンで表現)

レンダリング(Rendering)

画像を生成すること。
(質感や陰影等を計算)



データの容量と圧縮・解凍

関連: p44-47

アナログに近い状態でデジタル化したい ⇒ データ量が非常に増加!

データの**圧縮**: ファイルのデータ量(ファイルサイズ)を小さくすること。

可逆圧縮

非可逆圧縮

- 元データに戻せる
- 例: ランレングス圧縮
run length encoding
- 元データに(完全には)戻せない
- 例: 音声... 聞こえない周波数の音をカット
大きな音に隠れた小さな音をカット
- 例: 画像... 色の違いを間引く
- 例: 動画... 変化した部分(差分)だけを記録

データの**解凍(展開・伸張)**: 使用するためにもとの状態に戻すこと。

ファイル形式の違いによるデータ量を調べる(実習)

実習: BMPファイル(写真とイラスト)を様々な形式で保存して、
結果(データ量や見た目の変化)を考察する。

- ① [レポート等提出] ⇒ [2019xx-x_画像圧縮x組]
⇒ [2019xx-x_ファイル形式(xxx)]フォルダを開く。
- ② 名前を付けて保存を6回。
- ③ ファイルを観察(保存後の変化を記録)。

		BMP(元)	JPEG	GIF	PNG
写真	ファイルサイズ				
	特徴等				
イラスト	ファイルサイズ				
	特徴等				