

第3章 電子情報ネットワークの成り立ち

3.1 情報の単位

3.1.1 ビット --- 情報の最小単位

世の中にはさまざまな情報が満ちあふれていますが、その最小単位は何でしょう。たとえば「今、雨が降っていますか?」という質問に対する答は、できるだけ単純にした場合「はい」「いいえ」のどちらかになるはずです。「はい」「いいえ」に限りませんが、このような「2つのうちどちらか」という情報は、これ以上簡単にできない情報の最小単位だと考えることができます。この最小単位をビット(bit)と呼びます。

用語

ビット(bit)とは、「2つのうちどちらか」をあらわす、情報の最小単位のことです。

演習

世の中の情報で、「2つのうちどちらか」という形で表せるようなものをもっとたくさん考えてみよう。

3.1.2 ビット列と2進数

ビットを簡潔に書き表すために、1つのビットを2進法のひとつの桁に対応させて0または1で表すことにします。こうしておけば、複数のビットの組合せも見やすく表せます。

メモ

実はビット(bit)ということばは、「2進法のひと桁」 (binary digit)から来ています。

たとえば、雨の降りかたをもう少し細かく表したいとします。そのため、たとえば2ビットを使って

振り方	表現
降っていない	00
小雨	01
本降り	10
土砂降り	11

のように表せます。ビット数が増えたことと、雨の降りかたに関する情報が増えたことが対応しているのが分かりますね。一般に、ビットが N 個あれば、 2^N 通りの場合のうちのどれであるかを表すことができます。実は2進数も、 N ビットで「 $0 \sim 2^N - 1$ のどれであるか」を表現しているものだと考えられます。

演習

世の中にある情報のうち、「 N 通りの場合のうちどれか」という形で表せるようなものをもっとたくさん考えてみよう。

3.1.3 デジタルとアナログ

ここまで挙げてきたような、ビットの並び(ビット列)で表せる情報をデジタル情報といいます。デジタル情報は「 2^N 通りのうちどれか」という形で表されますから、「どちらでもない中間的な状態」というものはありません。これを離散的である、といいます。これに対し、ものの重さや長さなどの連続的な情報のことをアナログ情報といいます。実用的には、ある決まった小さな値を単位としてそれがいくつあるか、という形に読みかえることで、アナログ情報をデジタル情報に変換して表すことができます。これを量子化と呼びます。

用語

デジタル情報とは、「 N 通りのうちどれか」という形で表される離散的な情報をいいます。これに対し、アナログ情報とは、長さや重さなどのように連続的な情報をいいます。アナログ情報を最少の細かさを定めてデジタル情報に変換することを量子化といいます。

たとえば、皆さんの身長を表すのであれば、1mm を単位として表すので十分だと言えるでしょう。身長であれば 4m という人はいませんから、 $2^{12} = 4096$ であることを考えると、この方法なら 12 ビットのデジタル情報として表せることになります。ただし、1mm より小さい部分の情報を捨てていることを忘れてはいけません。これを量子化誤差といいます。もっと細かい値が必要なら、より多くのビット数を用いることになります。

演習

世の中の情報のうち、おもにアナログ情報として扱われているものはどれか。おもにデジタル情報として扱われているのはどれか。またアナログでもデジタルでも扱われているものはどれか。3 つずつ考えてみよ。

3.2 さまざまな情報の表現

3.2.1 情報の表現と符号化

さまざまな情報をビット列に対応させて表すことを符号化といい、その対応づけの規則を符号化規則といいます。たとえば「雨が降っていない」を 0、「雨が降っている」を 1 であらわす、というのも符号化規則です。

用語

情報をビット列に対応させて表すことを符号化という。

デジタル情報を受け渡すとき、情報の送り手と受け手で同じ符号化規則を使っていないと、情報が正しく伝わりません。このため、情報の種類ごとに標準の符号化規則をさだめて、共通に使うようにするのが一般的です。以下で、代表的な情報の種類ごとにその表現方法(符号化)について見てみましょう。

3.2.2 整数の表現

0 以上の整数(非負の整数)をビット列で表すばあいには、2 進数によって表すのが一般的です。たとえば 3 ビットの 2 進数について見てみます。

数	ビット列
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

$2^3 = 8$ ですから、0 ~ 7 までの整数が表せることになります。では、負の数まで表したい場合はどうでしょう。現在コンピュータで多く使われている 2 の補数表現では、「すべてのビットが 1」を - 1 に対応させ、3 ビットの場合次のように対応させます。

数	ビット列
0	000
1	001
2	010
3	011
-4	100
-3	101
-2	110
-1	111

同じビット数で負の数まで表すため、正の数については表せる範囲が半分になっています。2 の補数表現では、ある数が負であるかどうかを調べるには最上位(一番左)のビットが 1 かどうかを調べれば分かります。

メモ

コンピュータで整数を扱うときには、32 ビットや 64 ビットの 2 の補数表示が多く使われます。32 ビットの場合、10 進数で 10 桁程度の大きさまで扱えることになります。

メモ

コンピュータで扱う情報表現のビット数は 8 の倍数が使われることが多いのに気がつきましたか。これは、現在のコンピュータは 8 ビット単位でデータを扱うのに都合がよくできていることに関係しています。なお、8 ビットのことを 1 **バイト**(byte)と呼ぶことがあります。たとえば 32 ビットは 4 バイト、64 ビットは 8 バイトですね。

メモ

なお、コンピュータでは整数だけでなく実数(少数点つきの数)もよく扱います。その表現については話が長くなるので説明しませんが、決まったビット数で表すため、整数と同じく「表せる最大の数」「表せる最少の数」があります。さらに実数では「表せる最少のきざみ幅」も問題になります。そういう意味では、コンピュータで扱う実数は数学で言う実数とは違って、「実数の近似値」とでも言うべきものです。

3.2.3 実数の表現

少数点以下の値も含んだ数値を表す場合には、先に出て来たように「小さい値を単位として」表すこともできますが、それだと(絶対値が)非常に大きい数や小さい数を表すのが不便なので、*指数方式*または*浮動少数点*と呼ばれる表現を使うのが普通です。電卓にも、指数部と仮数部(有効数字の部分)を分けて表示 / 計算できるものがありますが、それと同じだと考えればよいでしょう。注意すべきなのは、指数部も仮数部も表せる桁数が決まっている、ということです。たとえば、(10 進数で)指数 2 桁+符号、仮数 5 桁+符号とすると、

$$10000000000 = 1.0000 \times 10^{+10}$$

$$0.0000000001 = 1.0000 \times 10^{-10}$$

などの値は表せますが、これらを足そうとすると

$$10000000000.0000000001 \sim 1.0000 \times 10^{+10}$$

のように近似値に丸められてしまい、「0 でないものを足したのに値が増えない」という不思議なことが起こります。そもそも 123456 という数値でさえ

$$123456 \sim 1.2346 \times 10^{+05}$$

のように概数でしか表せません。実際のコンピュータでは指数部も仮数部ももう少し余裕がありますが、まったく同様のことが起こります。これは、有限のビット数のデジタル表現では原理的に避けられない制約なのです。

演習

上記の「仮数 5 桁+符号、指数 2 桁+符号」表現で次の計算を実行すると結果はどうなりますか。またその違いはなぜ生じるのでしょうか。

$$3.1416 - 3.1234 + 0.0000001$$

$$3.1416 + 0.0000001 - 3.1234$$

3.2.4 画像の表現

画像(image)を符号化する場合には、まず画像全体を細かく縦横のます目に分割します。ここで 1 インチ(約 2.5cm)あたりいくつのます目に分割するかを*解像度*(resolution)と呼びます。表示画面では 75dpi (dot per inch --- インチあたりのます目数)、プリンタ出力では 300dpi ~ 600dpi くらいの解像度が標準的です。1 つのます目のことを*ピクセル*と呼びます。白黒で階調のない画像の場合は、1 つのピクセルは「白」「黒」のいずれかですから、1 ビットで表すことができます。カラー画像の場合には、図 3.1 のように、1 つのピクセルの色や明るさを、「赤」「緑」「青」3 原色の強さをそれぞれ 0 ~ 255 の段階で(つまり 8 ビットで)表し、1 ピクセルについて 24 ビットを使うやり方が標準的です。しかしもっとビット数を節約する流儀もいろいろあります。

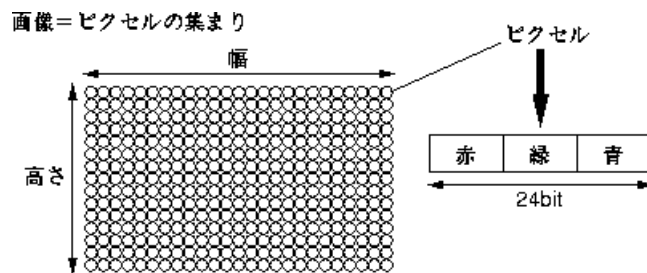


図 3.1 画像のデジタル表現

画像全体の情報は、すべてのピクセルの情報を合わせたものになります。画像のサイズとは横方向と縦方向にいくつずつ目があるかを意味し、たとえば「200×150 ピクセル」のように書きあらわします。画像を直接扱うようなお絵描きソフトを(次節で述べるドロー系と対比して) ペイント系のお絵描きソフト、と呼びます。

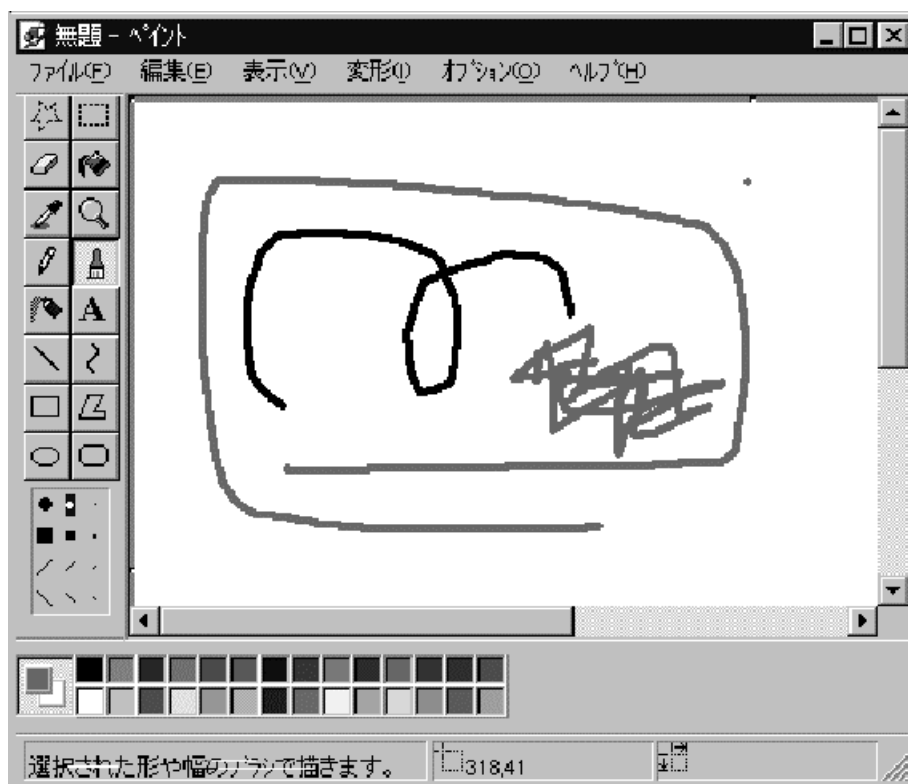


図 3.2 ペイント系のソフト

実際の画像では、ある小さい範囲のピクセルはどれも同じ色を持つことが多いので、これを利用して符号化のビット数を節約できます。これを画像の**圧縮**(compression)と呼びます。圧縮方法まで含めた画像の表現方式にもいくつかの規格があり、それぞれ名前がついています。 GIF、JPEG、TIFF など代表的な画像の表現形式です。

演習

ペイント系のお絵描きソフトで絵を描いてみよう。拡大機能を使って、実はその絵が多数の点から成っていることを確認しよう。(画像)イメージを読み込んだり、それを少し修正してみても、同じように確認しよう。

3.2.5 図形の表現

コンピュータで幾何学的な図形などから成る図を表現する場合、上で述べた画像による表現も可能ですが、それよりも「どこから、どこまで、どのような太さの、どのような色の、どのような線を引く」という情報を符号化して表す方が次の点で優っています。

- ・一般に画像よりも少ないビット数で済む
- ・自由な大きさに拡大/縮小できる
- ・後から変更することが容易である

図形の一種として文字や(小さな)画像を含めて配置することもできます。このような考え方に基づくお絵描きソフトを **ドロー系のお絵描きソフト**、と呼びます(図 3.3)

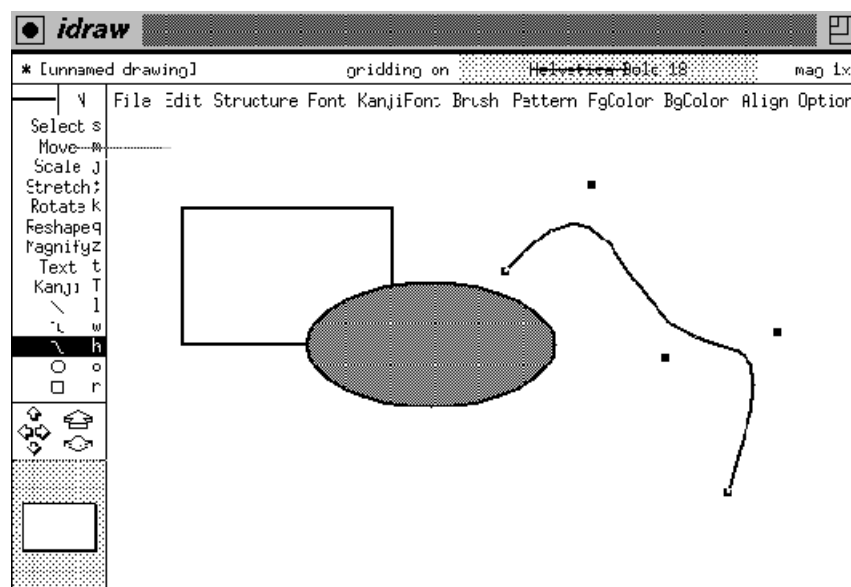


図 3.3 ドロー系のソフト

演習

ドロー系のお絵描きソフトで絵を描いてみよう。一度配置した図形等を移動したり、変形できることを確認してみよう。

3.2.6 音声の表現

音をアナログ信号の形で送る場合には、空気の振動をマイクで電圧の変化に変換し、時間とともに電圧が変化するようにそのまま伝送して、受け取った側でそれをそのままの形でスピーカに送ることで元の音が再生されます。これに対し、音をデジタル情報にする場合には、非常に細かい時間間隔ごとに電圧を調べ(サンプリング)、電圧の大小も先に説

明したような形でデジタル化して伝送します。典型的には、1 秒間に 44,000 回(44kHz)のサンプリングを行い、各サンプルの値を 16 ビットで表現します(音楽 CD もこのような形で音をデジタル記録しています)。音声の情報に対しても圧縮を行うことができます。音声の表現方式にもいくつかの規格があり、それぞれ名前がついています。μ-law、AIFF、WAV など代表的な音声の表現形式です。

3.2.7 動画の表現

動画(ムービー)の場合は、音と同様にサンプリングを行いますが、ただし各サンプルは画像情報ということになります(これをフレームと呼びます)。フレーム数は毎秒 10~20 程度が普通です。動画の情報はそのままでは非常に大きくなる(ビット数が多くなる)ので、圧縮はほとんど必須だと言えます。動画の表現形式としては、MPEG、Indeo、Cinepak、QuickTime などが代表的です。

3.3 文字とテキスト形式

3.3.1 文字の表現

文字で表された情報はコンピュータで多く扱うので、この節でやや詳しく見て行きましょう。文字を表すのも、これまでやってきたのと同様に、それぞれの文字をビット列に対応づけることで符号化します。この対応のことをコード系と呼びます。また、それぞれの文字を表すビット列、またはそれを 2 進数とみなしたときの数値を、その文字の文字コードと呼びます。コード系には、1 つの文字を 8 ビット単位で表すものと 16 ビット単位で表すものがあります。8 ビット単位では最大 $2^8 = 256$ の文字しか表せませんから、漢字が扱えるのは 16 ビット単位のコード系だけです。コンピュータどうしが通信を行なう際に、それらの間でコード系が一致していないと、ある文字を送ったはずが、受け取った側では全然違う文字だと思われてしまい、めちゃくちゃになります。このため、標準のコード系が規格として定められています。コード系の規格としては、アメリカ通商産業省の定めた文字規格である **ASCII**、日本の通産省の定める **JIS X0208** があります。ASCII は 8 ビット単位、JIS X0208 は 16 ビット単位のコード系です。さらに、国際標準化機構が定めた **ISO-2022** という、「世界中の文字コードを混在させる規格」が知られています。これを使えば、ASCII の文字と日本語の漢字とを 1 つのファイルで混ぜて使うことができるのです。異なるコード系を使っているコンピュータ同士では、ネットワークを用いた情報交換は出来ません。どうしても情報交換を行なう必要がある場合には、通信の際に文字コードの変換を必要とします。

演習

日本語の五十音のそれぞれに適当な番号を振って、対応表を書いてみましょう。さらに、句読点や他の記号にも番号を振ってみましょう。また、自分の対応表を、他の人が書いたものと比較してみましょう。

3.3.2 ASCII コード

英語のアルファベットで使用されている「ローマ文字」のコード系は、ほぼ世界中のコンピュータが ASCII 規格にしたがっています。ほぼ世界中のコンピュータがこの規格を採

用しているため、これらの文字コードの違いを意識する必要はほとんどありません。ASCII コードは、アルファベット・数字・いくつかの記号に 32～126 の文字コードを割り振っています。コンピュータの内部では、その番号を 2 進数で表記したのと同じビット列がそれぞれの文字を表します。

32	48 0	64 @	80 P	96 ‘	112 p
33 !	49 1	65 A	81 Q	97 a	113 q
34 "	50 2	66 B	82 R	98 b	114 r
35 #	51 3	67 C	83 S	99 c	115 s
36 \$	52 4	68 D	84 T	100 d	116 t
37 %	53 5	69 E	85 U	101 e	117 u
38 &	54 6	70 F	86 V	102 f	118 v
39 ’	55 7	71 G	87 W	103 g	119 w
40 (56 8	72 H	88 X	104 h	120 x
41)	57 9	73 I	89 Y	105 i	121 y
42 *	58 :	74 J	90 Z	106 j	122 z
43 +	59 ;	75 K	91 [107 k	123 {
44 ,	60 <	76 L	92 \	108 l	124
45 -	61 =	77 M	93]	109 m	125 }
46 .	62 >	78 N	94 ^	110 n	126 ~
47 /	63 ?	79 O	95 _	111 o	

図 3.4 ASCII 文字コード表

演習

次の文字列を、ASCII 文字コード表にしたがって、数の列に置き換えてみましょう。

ima no kion wa, 20 do desuka?

3.3.3 日本語の漢字コード

コンピュータはアメリカで生まれ、したがって、情報処理技術も初期の段階ではアメリカのみで開発・発展してきました。そのため、コンピュータ普及の初期の頃は、コンピュータ上で日本語の漢字を用いるために特殊な方法を必要としました。現在では、日本の通産省が定めた JIS(日本工業規格)という規格に日本語の漢字を取り扱うための規格があり、ほとんどすべてのコンピュータメーカーの、ほとんどすべてのコンピュータがこの JIS 規格にしたがって、日本語の漢字を取り扱っています。しかし、「JIS 規格による日本語の漢字コード」と ASCII コードを混在させて通信を行なう際には、コンピュータそのものの内部構造の違いやネットワークへの対応方法の違いが原因となって、

- ・ JIS コード(ISO-2022 に基づく)
- ・ シフト JIS コード
- ・ 日本語 EUC コード

の 3 通りの規格が定められており、どの方法も広く使われています。また、ISO は書類番号 10646 番で新しい文字コードを定義していて、Unicode とよばれています。異なる日本

語の漢字コードを用いているコンピュータ同士が通信をする場合や、そのコンピュータが標準的に提供している日本語の漢字コードとは異なる日本語の漢字コードを用いているファイルを読む場合には、変換を行なう必要があります。

演習

JIS コード表に載っていない漢字を漢和辞典などからさがしてみましょう。それらの漢字をインターネットを使って通信するにはどのようにするべきでしょうか？

演習

世界中の言語にはどんな言葉があるかを調べてみましょう。また、それらの言語で使われている文字を画面に表示したり、電子メールでやりとりしたりするための工夫を考えてみましょう。

3.3.4 テキストデータとバイナリーデータ

数字・記号・アルファベットなどの文字コードのみから成るデータをテキストデータといいます。それに対して、テキストデータ以外のデータをバイナリーデータといいます。バイナリーデータは、画像データなどのようなデータを表すために用います。ところで、電子メールはテキストデータしか送信・受信出来ないように設計されています。したがって、電子メールなどを通して画像データなどのバイナリーデータを送ろうとするには、適当な規則に基づいて変換を施してから送る必要があります、さらに、受信した人は、変換規則に基づいて元のデータに戻す必要があります。

演習

画像ファイルを電子メールに添付して他の人に送ってみましょう。また、そのとき送られた電子メールの大きさを調べてみましょう。さらに、送ったメールを受け取ってもらったら、どんな画像ファイルだったかを電子メールで返信してもらいましょう。

3.3.5 使ってはいけない文字

インターネットで日本語・英語の通信を行なう際に使用可能な文字は、RFC やその他の規格によって定められています。これらの規格に合わない文字番号を通信に使うと、通信相手に正確に内容が伝わらないばかりか、場合によっては、送られたメッセージの中継の途中で動作中のサーバーを止めてしまうこともあるかもしれません。この節では、インターネットでは使ってはいけない文字番号を持つ文字の中で代表的なものについて説明します。これら以外の文字を使う際も、読む人が正常に文字を表示できるかどうかに注意をして、使う文字を選ぶべきでしょう。

外字

JIS で定められた日本語の漢字コードは、6355 文字分です。しかし、日本人の人名・日本の地名に用いられている文字の総数は、この数を遥かに越えます。さらに、コンピュータを国語の研究などに用いようとすれば、古い字体や既に用いられなくなった字、中国語の

文字で日本語に影響を与えた文字などを取り扱う必要があります。これらの文字には JIS の定めた文字コードがないので、「外字」と呼ばれる OS の機能を使って、結果として、そのコンピュータを使う人が自分で管理することになります。たとえば、ある JIS 漢字コード に定義されていない日本語の漢字を、「外字コード 0001 番」などとして決めるということです。このようにしてしまうと、使っているコンピュータや使っている人同士が異なれば、同じ文字番号が異なる文字に対応し、その結果として、正常な通信が出来ないという結果をもたらします。外字は、必要上用いるものですから、利用すべきではないとはいえません。したがって、外字で表している文字をどうしても通信内容に含める必要がある時には、その文字の形を別途の方法で伝えたり、その文字の形の画像ファイルを同時に送信したりする必要があるでしょう。これらの問題点を解決すべく、JIS 漢字コードの拡張が検討されています。この拡張によって、外字の問題のほとんどが解決することが期待されています。

半角カタカナ

用いてはいけない文字の中で、誤って用いられることの多い文字が、8 ビット単位のカタカナ、俗に「半角カタカナ」といわれている文字です。これは、むかしコンピュータで漢字を扱うのが難しかったころ、8 ビット単位の文字コードでカタカナと必要最小限の記号を扱っていたことに由来しています。多くのコンピュータでは、8 ビットカタカナ文字は通常の漢字より狭い幅で表示されます。ですからその文字がカタカナの場合には、字の幅に注意をすればすぐに判別できますが、句読点を表す「、。」や、促音をあらわす「ー」、中グロを表す「・」の 8 ビットカタカナ文字は、画面上でもなかなか判別がつかみませんので注意が必要です。

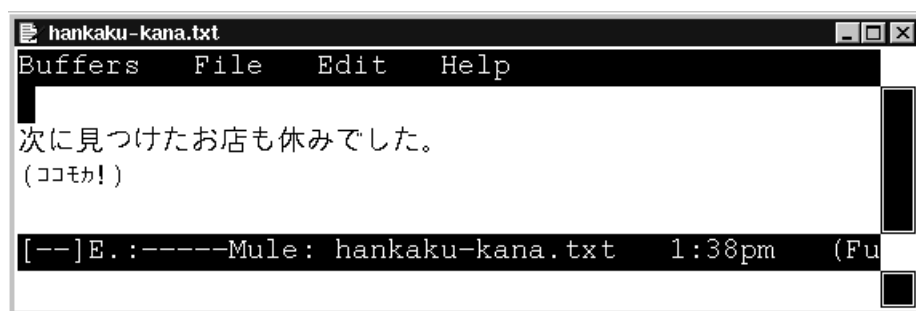


図 3.5 半角カタカナの表示が正常な例

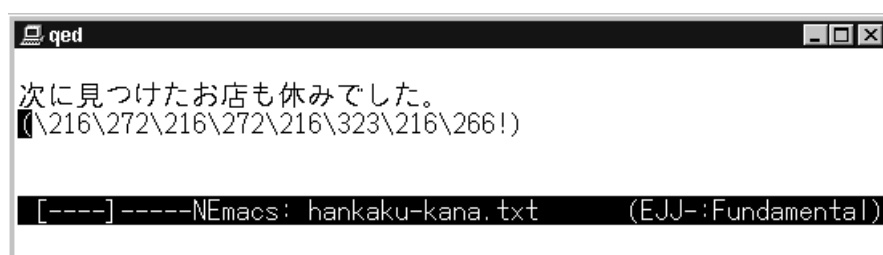


図 3.6 8 ビットカタカナの表示が正しくない例

例に上がっているのは、8 ビットカタカナを正常に表示可能なソフトウェアと、8 ビットカタカナを正常に表示不可能なソフトウェアに無理矢理に 8 ビットカタカナを含む文書を表示させた例です。「ココモカ!」の部分の文字がおかしくなっています。これらの文字に用いられる文字コードは、ASCII で規定している文字コードに使用可能な部分を逸脱しているため、RFC では 8 ビットカタカナ用の文字コードを用いることは出来ないように規定されています。そのため、多くの電子メールクライアントやネットニュースクライアントでは、半角カタカナを表示できないものが多く存在しています。さらに、インターネットにある中継ホストの中には、8 ビットカタカナをうまく中継できないものがあります。しかし、もともとインターネットでは、8 ビットカタカナを表示できるようにする必要はありませんので、受けとった側がこのような状況で困ってしまう場合には、送り手に問題があることになります。したがって、インターネットメッセージでは 8 ビットカタカナを用いてはいけません。

文字コードに関するまとめ

ここまでで取り扱った項目をまとめると、「書き手の環境・読み手の環境に気をつけよう」ということです。世界中にはさまざまなコンピュータがあり、さまざまな OS で動作しています。同じ文字番号を使っているからといって、同じ文字が表示されるとは限りません。文字の表示という基本的な問題だけを取り上げても、このようにさまざまな問題を抱えているのが現在のインターネットなのです。コード系は、8 ビット(1 バイト)単位で符号化するものと 16 ビット(2 バイト)単位で符号化するものがおもで、それぞれにさらにいく通りものコード系が存在しています。

演習

インターネットで使ってはいけない文字についてまとめてみましょう。

3.4 コンピュータの構成

3.4.1 コンピュータとは

ここまで、コンピュータが扱うさまざまな情報について見てきましたが、それではコンピュータとは何でしょう？ コンピュータ(computer、計算機)とは元はその名の通り「数の計算を高速に/大量に行なうための機械」でしたが、今ではこれまでに見てきたさまざまなデジタル情報(ビット列)を取り扱う(蓄積・加工する)ための機械であると考えerるほうが実情に合っています。

用語

コンピュータ(計算機)とは、デジタル情報を蓄積したり加工するための装置を言います。

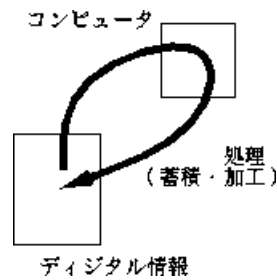


図 3.7 コンピュータの概念

3.4.2 コンピュータの働き

「デジタル情報を処理する」とは具体的にはどういうことでしょうか？ これはつまり、コンピュータにどのような機能が備わっているかということにほかなりません。コンピュータには具体的には次のような 3 つの機能が備わっています。

- ・ 入出力 --- 外部とコンピュータとの間でデジタル情報をやりとりする機能
- ・ 記憶 --- コンピュータ内部にデジタル情報を格納しておく機能
- ・ 演算と制御 --- コンピュータ内部でデジタル情報を加工する機能、および全体を制御する機能

これらの機能について順番に見てみましょう。

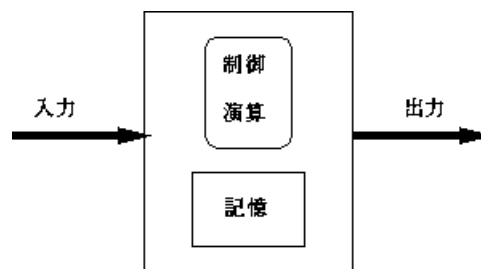


図 3.8 コンピュータの働き

3.4.3 入出力

外部からコンピュータにデジタル情報を取り込むことを入力(input)といいます。逆に、コンピュータから外部にデジタル情報を取り出すことを出力(output)といいます。入力と出力をあわせて入出力といいます。コンピュータの機器のうちで、入出力を行なうための機器を入出力装置といいます。代表的な入力装置としては、キーボード、マウス、音声入力装置、画像入力装置などがあります。代表的な出力装置としては、ディスプレイ (表示装置)、プリンタ、音声出力装置などがあります。そのほか、コンピュータに備わっている各種のインタフェース(やりとりのための接続口)は入力と出力の両方を兼ねています。たとえばあるコンピュータからネットワークインタフェースを経由して情報を出力し、それを他のコンピュータのネットワークインタフェースを経由して入力することで、デジタル情報を伝送できます。

3.4.4 記憶

コンピュータにはデジタル情報を記憶しておく機能も必要です。たとえば、入力した情報は演算を行なうまで記憶しておく必要がありますし、演算の結果も記憶しておき、利用者から指示があった時はじめて出力するのがふつうです。コンピュータの機器のうちで、記憶を行なう部分を**記憶装置(storage)**といいます。記憶装置はさらに、高速にし入れが可能だが容量が限られている**主記憶装置(main storage)**と、低速だが大量に記憶でき、コンピュータの電源を切っても内容が失われない**補助記憶装置**ないし**二次記憶装置(secondary storage)**に分類できます。主記憶装置は、メモリチップと呼ばれる数ミリ角の半導体素子の集まりで構成されています。1つのメモリチップに、約 64,000,000 ~ 256,000,000 ビットの情報が格納できます。補助記憶装置は、入出力装置の一種と考えることができます。つまり、計算機から出力された情報を格納しておき、そこから取り出した情報は計算機への入力になるのです。補助記憶装置にはさまざまな種類のものがあります。代表的なものには、ハードディスク、フロッピーディスク、MO ディスク(光磁気ディスク)、磁気テープなどがあります。

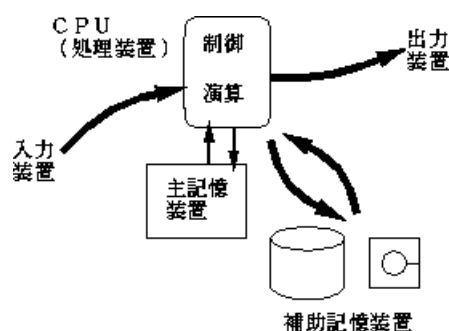


図 3.9 コンピュータの構成

3.4.5 演算と制御

コンピュータの内部でデジタル情報を加工することを**演算**といいます(数値以外の情報に対してもこういいます)。コンピュータの機器のうちで、演算を行なう部分を**処理装置**または**CPU**(central processing unit)などといいます。CPU も小さな半導体素子で構成されています。コンピュータがいつどのようにデジタル情報を入出力し、記憶し、加工するかは、すべて処理装置によって制御されています。つまり、処理装置には演算機能に加えて、**制御機能**も内蔵されているわけです。コンピュータが具体的にどのような制御を行なうか(つまりコンピュータが全体としてどのように動作するか)は、すべて**プログラム**によって定められます。プログラムもデジタル情報(ビット列)の形で表され、主記憶装置の中に格納されています。CPU の中の制御機能は、主記憶装置に格納されているプログラムを参照しながら、その指示に従ってコンピュータ全体を制御します。

3.4.6 ハードウェアとソフトウェア

ハードウェアとは、コンピュータの機器全般をいいます。これに対し**ソフトウェア**とは、プログラムやデータ(プログラムが扱う情報)など、コンピュータで有用な作業をするために必要なデジタル情報全般をいいます。コンピュータは、ハードウェアに加えて、ソフト

ウェアがなければ動作できません。ハードウェアとソフトウェアはコンピュータを動かす車の両輪だといえるでしょう。

演習

現在使っているコンピュータのハードウェアとソフトウェアを購入するのにそれぞれどれくらいの金額が必要か調べてみなさい。

3.4.7 さまざまなソフトウェア

電子メール、WWW、文書作成、ゲームなど、特定の作業をおこなうのに必要なソフトウェアを**応用ソフトウェア**または**アプリケーションソフトウェア**と呼びます。コンピュータショップなどに行くと、さまざまな応用ソフトウェアが販売されているのを見かけます。これに対し、コンピュータを動かす上で必要となる機能全般を提供するソフトウェアを**基本ソフトウェア**と呼びます。基本ソフトウェアの代表的なものに**オペレーティングシステム(OS)**があります。OSは、コンピュータ上で常時動作していて、応用ソフトウェアを起動させてくれたり、応用ソフトウェアが動作するときの下支えとなる機能を内部的に提供してくれるソフトウェアです。

3.5 マルチメディア文書

3.5.1 マルチメディアとは

コンピュータで扱う情報のうちでテキストが占める割合が高いことは説明しましたが、もちろんその他の種類の情報も扱えます。コンピュータで扱うテキスト以外の情報種別のことを総称して**マルチメディア**と呼びます。まずテキストが基本で、それに他のものを加えるから複数種類の(マルチ)媒体(メディア)、というわけです。しかし技術の進歩とともに、最近ではテキスト情報を全く含まないようなマルチメディアのコンテンツ(内容)も一般的になっています。

演習

では、情報を伝えるのにテキスト以外にどのような媒体(メディア)の種別があり得るかできるだけ多く挙げよ。続いてそれらを、(1)既にコンピュータで扱うことが実用化されているもの、(2)コンピュータで扱うことが研究されているもの、(3)まだコンピュータで扱えるかどうか分からないもの、に分類してみよ。

3.5.2 さまざまなマルチメディア素材

ではもう少し具体的に、WWWなどで使われているマルチメディアの素材(コンテンツを作る材料)について見てみましょう。それぞれの素材は通常、ファイルの形で保存しておき、必要に応じて取り出して利用します。ある特定種類のメディアの素材でも、ファイルに格納する方法(や、そのメディアの情報を0、1に対応づける方法)はひと通りとは限りません。これを**ファイルフォーマット**と呼んでいます。さらに、ファイルの内容が分からなくなないように、ファイルには格納されている情報のフォーマットに対応した名前(具体的には拡張子)をつけるのが通例です。以下では代表的なファイルフォーマットについても簡単に触れます。

- ・ 画像 --- コンピュータの画面に表示されるさまざまな絵は、基本的な素材です。画像のファイルフォーマットとしては、GIF(.gif)、JPEG(.jpeg または .jpg)などがあります。
- ・ サウンド --- 各種の音(サウンド)も素材としてよく使われます。サウンドのファイルフォーマットとしては、AU 形式(.au)、WAVE 形式(.wav)などがあります。
- ・ 動画 --- 動く画像は、大変インパクトがありますが、情報の量も多いので大きなファイルになります。動画のファイルフォーマットとしては、Quicktime 形式(.mov)、Video for Windows 形式(.avi)、MPEG 形式(.mpeg または .mpe)などがあります。
- ・ アプリケーション --- ワードプロソフトや、表計算ソフトや、アニメーションソフトなど、各種ソフト固有のファイルを媒体(メディア)として使うことや、任意のプログラムを媒体(メディア)として使うこともあり得ます。その場合、アプリケーションやプログラム独自の機能が使えるという利点がありますが、どこでもそのようなアプリケーションやプログラムが動かせるとは限らないという制約もあります。ファイルフォーマットはアプリケーションやプログラムの動く環境に応じてさまざまです。

演習

用意された素材集を見てみよう。また、ファイルの大きさと素材の内容にはどのような関係があると思うか、考えてみよう。

3.5.3 マルチメディア文書の設計

テキスト以外の情報も含んだ文書のことを、マルチメディア文書と言います。また、使用するソフトウェアにもよりますが、マルチメディア文書ではテキストの部分も、見出し、本文、引用、強調箇所など内容に応じて字体や文字の大きさや色を変えるなど、表現の工夫が可能である場合が多いです。マルチメディア文書では、テキストだけで表せないことを読み手に伝えられるという点では優れていますが、読み手、書き手の双方において、各種のメディアや表現などに目を奪われてしまい、肝心の伝えたい内容がおろそかになってしまうというおそれもあります。そのような罅に陥らないためには、文書が伝えたい内容が何であるかを明確にして、文書の構成を決め(ここまでは普通の文書の設計法と同じです!)、それに沿って各種のメディアや表現の使い方を決めて行くべきです。

メモ

各種メディアや表現の使い方については、次の点に気をつけましょう。

- ・ 画像は、必要以上に大きくせず、図などテキストで表現できない情報を効率的に伝える、人物の顔を知らせる、確かに伝えようとしているできごとがあったことを納得してもらい、など明確な目的を持って使う。
- ・ サウンドや動画は、それ自体が重要な情報である場合(しゃべり方や音曲そのものを伝えたい、動きそのものを伝えたい等)に限定して使うのがよい。
- ・ 文書に「臨場感を与える」飾りものとして画像、サウンド、動画を使う場合には、最小限にとどめる。長いサウンドや動画はくりかえし見る人には退屈であり、コンピュータに与える負荷も大きくなる。
- ・ 文書の表現を工夫するときは、「どんな項目があるかさっと見てとって欲しいから見

出しを大きくする」「この段落だけは熟読して欲しいからその色を変える」など目的を持って行い、その目的と関係のない表現は控える。

演習

先の演習で見た素材を使って、マルチメディア文書を企画しなさい。テーマは先生からもらうこと。

3.5.4 マルチメディア文書の作成

マルチメディア文書を作成する手助けをしてくれるようなソフトウェアを一般にオーサリングソフトと言います。最近ではワープロソフトにも画像などが入れられるものがありますから、両者の区別はあいまいになりつつありますが、おもにテキスト情報を使うのがワープロソフト、マルチメディアをおもに考えているのがオーサリングソフトだと言えます。オーサリングソフトでは、画面からテキストを打ち込むのに加えて、範囲指定やメニューを使って「ここは見出しである」「ここは箇条書きである」「ここにN行M列の表を入れる」「ここにこれこれの素材(画像や音)を入れる」といった指示を行うことによって、マルチメディア文書を組み立てて行くことができるようになっています。

演習

オーサリングソフトを使って、先の演習で企画したマルチメディア文書を実際に作ってみよ。

3.5.5 マルチメディア文書とバインディング

オーサリングソフトで作成し保存したファイルの中はどうなっているのでしょうか？WWW用のオーサリングソフトでは、作成したドキュメントをHTMLと呼ばれる形式で保存してくれ、HTMLは普通のテキストエディタで見たり編集したりできます。実際に眺めてみてください。そうすると、打ち込んだテキストそのものがそのまま現われている部分と、テキストにない何だかよく分からないものとが混ざりあっているのが分かるはずです。ここでテキストにない部分は、「この範囲は見出し」「ここはこういう色」「ここはこういう表」といった情報を表しています(図3.10)。

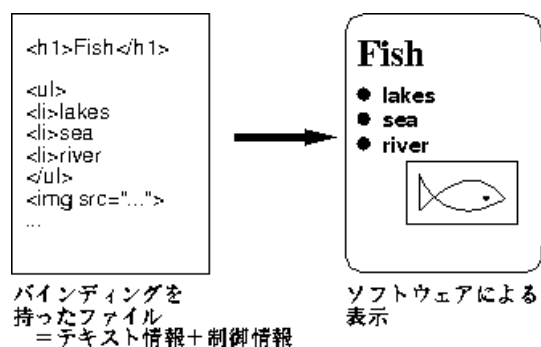


図 3.10 バインディングの概念

このように、もともとの内容(コンテンツ)に別の情報を付加して、全体としてより豊富な表現や機能を提供することをバインディングと呼びます。ここでは HTML 形式を見ましたが、ワープロソフトはそれぞれ固有の形式のバインディングを付加していますし、コンピュータで扱う多くのファイル形式にも同様の考え方が見られます。

演習

先の演習で作った文書を眺めて、内容(コンテンツ)とバインディングの関係がどうなっているか考えてみよ。