

ビットとバイト

パソコンやスマートフォンは無論、ハードディスクや DVD レコーダなどのあらゆるデジタル機器では、データの取り扱い容量を示す時などに、バイト(Byte)という単位が使われています。

また、光回線や無線 LAN などのデータ通信においてデータの転送速度を示す時には、バイトだけでなく、ビット(bit)という単位も使われます。

2 進数とビット(bit)

そもそもコンピュータでは、0 と 1 の組み合わせによる 2 進法で計算をします。なぜ 2 進数が使われているかというと、電気ので動くコンピュータは、電気が流れる(オン)、流れない(オフ)で判断をする方法がもっとも適しているからです。

電気のオンオフの状態を数字に置き換えると、「オフの状態を0」、「オンの状態を1」とすることができ、使える数字は 0 と 1 の二つだけとなります。この 0 と 1 だけの組み合わせの最小単位をビット(bit)と呼びます。bit は、binary digit(2 進数)から生まれた造語です。

bit を 0 と 1 を収めるマス、つまり桁数と考えると良いでしょう。1bit なら 1 マス(1 桁)なので、0 か 1 を収めることができ、2 通りの組み合わせが構成できます。bit が 2 つあると、00、01、10、11 の 4 通りの組み合わせができます。3 つあると、000、001、010、011、100、101、110、111 の 8 通りの組み合わせができます。

つまり、bit の数が増えれば、2、4、8、16、32、64、128、256 といった「2 のべき乗」の形で、コンピュータで扱える数を増やしていけることとなります。

バイト(byte)

バイト(byte)は、複数のビットという意味の用語で、データ量あるいは情報量の単位で、8ビットが1バイトです。

【便利知識】

現在では、1バイト=8ビットで定着していますが、何ビットを1組の単位としてデータ処理を行うかは、コンピュータの歴史の中で、様々でした。

黎明期の中央処理装置(CPU)は 4 ビット単位でしかデータ処理ができないものがありましたし、データ処理(メモリのアドレス付け)の単位をバイトではなく、バイトに比べ

て長い固定長の「ワード(word)」で処理するワードマシンや、バイトに比べて短い文字(「キャラクタ(character)」)単位で可変長で処理するキャラクタマシンというものもありました。

データ処理の単位をバイトという単位で行うようになったのは IBM 製の System/360 が誕生して以降です。ただバイトは必ずしも 8 ビットではなく、メーカーや機種によっては 6 ビットや 7 ビットのものなどもありました。

バイトが 6 ビット以上なのは、小文字と大文字の英字と数字を表すのに $26+26+10=52$ 以上の組み合わせが必要だからでした。6 ビットあれば 64 通りの組み合わせが表現できます。扱うことのできる記号の数が増えると 6 ビットでは不足 7 ビットを単位とするものが生まれ、やがて通信での誤り検出のためのパリティビットを加えた 8 ビットで処理することが標準的になったのです。

ISO(国際標準化機構)や IEC(国際電気標準会議)で「1 バイト=8 ビット」と制定されたのは 2008 年で、つい最近のことです。

【便利知識】

バイト(byte)を省略して大文字の B で表わすことが少なくありません。KB(キロバイト)、MB(メガバイト)、GB(ギガバイト)、TB(テラバイト)といった使い方がされます。これに対してビット(bit)を省略して小文字の b で表わすこともないわけではありません。典型的な例がデータ通信での伝送速度を示す、bps(bits per second)です。B と b を単独で使うことは、避けた方が無難です。

【便利知識】

一般的に K(キロ)は 1,000 を意味しますが、コンピュータの世界では、2 の 10 乗の 1,024 として扱われることがあり、KB が 1,000 バイトなのか、1,024 バイトなのかははっきりしないことがあります。そのため、コンピュータの容量や記憶装置の大きさを示す時などには、1,024 バイトであることを示すために KiB(キビバイトと読みます)という表記がされることがあります。

16 進数

社会生活の上では、0 から 9 までの数字で表現される、10 進数が使われています。10 進数が定着したのは、人間の指が左右で 10 本あるからだ、とも言われています。慣れているということもありますが、10 進数は人間にとって理解しやすく、共通の概念として有効なものであることは言うまでもありません。

10進数では9を超えると10、99を超えると100という形で、桁上がりします。

コンピュータの世界では2進数が使われ、0と1の数字で表現されます。1を超えると10に、11を超えると100に桁上がりします。2進数では、2のべき乗で桁上がりしますので、どうしても桁数が大きくなり、読みにくいなど、人間にとって取り扱いが面倒になります。

そこで、コンピュータに関する仕様書やプログラム内では、4ビット毎にまとめて、0から9までの数字とAからFまでの英字で表現する16進数が用いられるようになりました。Aは10進数の10、Fは10進数の15に当たります。

1バイト=8ビットを4ビットずつ区切って16進数で表現すると、8桁の2進数が2桁の16進数で表現できます。つまり、16進数はコンピュータで複雑な処理をすることなく、すぐに2進数に変換できるのです。

16進数は、10進数よりも桁数が小さくなり、使う文字も紛らわしいものが無いので、人間とコンピュータの双方にとって扱いやすいのです。

10進数・2進数・16進数【比較表】

10進数	2進数	16進数
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	10000	10
17	10001	11
(中略)	(中略)	(中略)
255	11111111	FF