

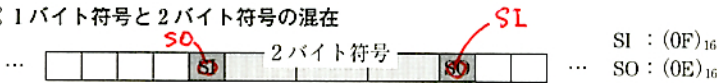
2.16 符号化

※ 1バイト符号

- ☒ ASCII (American Standard Code for Information Interchange)
- ☒ EBCDIC (Extended Binary Coded Decimal Interchange Code)
- ☒ JIS 8 単位符号 (JIS : Japan Industrial Standard)

| 16進 | 下位4ビット (16進表示) | | | | | | | | | | | | | | | |
|-----|----------------|-----|-----|-----|-----|-----|-----|-----|-----|----|-----|-----|----|----|----|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
| 0 | NUL | SOH | STX | ETX | EOT | ENQ | ACK | BEL | BS | HT | LF | VT | FF | CR | SO | SI |
| 1 | DLE | DC1 | DC2 | DC3 | DC4 | NAK | SYN | ETB | CAN | EM | SUB | ESC | FS | GS | RS | US |
| 2 | SP | ! | " | # | \$ | % | & | ' | (|) | * | + | , | - | . | / |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | : | ; | < | = | > | ? |
| 4 | @ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| 5 | P | Q | R | S | T | U | V | W | X | Y | Z | [| \ |] | ^ | _ |
| 6 | ` | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o |
| 7 | p | q | r | s | t | u | v | w | x | y | z | { | | } | ~ | DEL |

※ 1バイト符号と2バイト符号の混在



コンピュータで文字情報を扱うには、各種の文字や記号を**符号化** (coding, encode) しておく必要がある。英語で利用する文字や記号の個数は128を超えないので、ASCIIコードやEBCDICなどの1バイトの符号で表すことができる。図にJIS 8 単位符号の例を示す。

これらの符号では、0～9の数字に対して、その数を表す4ビットの2進数を下位4ビットとし、上位4ビットに $(3)_{16} = (0011)_2$ を付加した符号を割り当てている。したがって、10進数の293を記憶しておく際、この上位4ビットを除去し、2, 9, 3を表す4ビットの2進数を順に並べ、(0010 1001 0011)として記憶しておくことメモリ量の節約になる。このような形式のデータをpacked decimal (圧縮形式の10進数) という。

10進数字を4ビットの2進数で表したこのような符号は、BCD (Binary Coded Decial) 符号と呼ばれ、**2進法10進数** (BCD数) を表す際に用いられる。BCD数のビット系列は、packed decimalと同じで、293は $(0010\ 1001\ 0011)_{BCD}$ と表す。本書では、このようなビット系列がBCD数であることを示すため、BCDという添字を付加する。なお、BCD符号では、0～9のほかに、 $(C)_{16}$ および $(D)_{16}$ をそれぞれプラスおよびマイナスの記号を表すために用いている。また、BCD数を表すために、2.4節の3余り符号を用いることもできる。

漢字などの文字も符号化しようとする、1バイトではできないので2バイト符号が必要となる。そのような符号には、JIS符号、シフトJIS符号、EUC (Extended UNIX Code) などがある。また、さらに多くの文字を統一的に扱おうとするUnicodeも考えられており、符号のビット数も増加している。

1バイト符号と2バイト符号が混在している場合、1バイト符号を2バイトに変換していたのではむだなビットが多くなる。そこで、どの符号にも、2バイト符号の開始を意味する **SI (Shift In)** および **SO (Shift Out)** 符号が準備されている。

符号の場合、数の桁に対応するビットの場所を**位置**と呼ぶ。

(E識別方式、あふいは

2バイト符号

(あふいは符号系列)