

6. 文字データの世界 (日本語文字)

- (1) コンピュータによる漢字の入力
- ・漢字キーボードの時代 (60 字×36 字=2160 字)
 - ・日本語ワープロの誕生 (東芝 JW-10)
→かな漢字変換/ローマ字漢字変換

(2) 漢字の種類 (数)

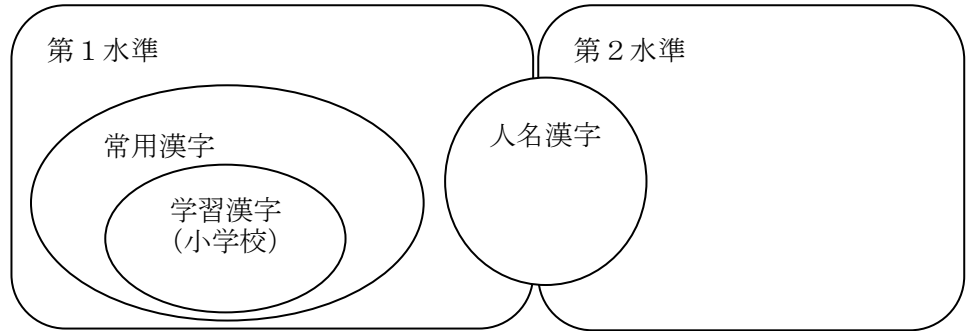
<現実の世界>

- ・学習漢字
- ・常用漢字
- ・人名漢字

- ・大漢和辞典の収録漢字

<コンピュータの世界>

- ・第1水準
- ・第2水準



(3) 全角文字のコード (JIS 漢字コード)

1 文字を 16 ビット (2 バイト) で表現 → 通りの表現が可能

文 →	 	→ (4 A 3 8)
教 →	 	→ (3 6 3 5)

4桁のコード番号で識別

- ・欠字 … コードが存在しない漢字 (森鷗外、中国の地名など)
- ・外字 … 登録されていない欠字や特殊記号などに任意にコードを付けたもの
 - ・メーカー外字 → 高 崎 青 薊 ① ② ③ I II III 絵文字など
(日本語入力バー「IMEパッド」-「文字一覧」)
 - ・ユーザー外字

(4) いろいろな文字コード

	JIS 漢字	シフト JIS	EUC	Unicode	種類
文	4A38	95B6	CAB8	6587	第1水準
教	3635	8BB3	B6B5	6559	第1水準
高	(967E)	(FBFC)	(外字)	9AD9	外字 (IBM 文字)
盆	4B5F	967E	CBDF	76C6	第1水準
①	(2D21)	(8740)	(外字)	2460	外字 (NEC 文字)
I	(2D35)	(8754)	(外字)	2160	外字 (NEC 文字)

- ①JIS 漢字コード
- ②シフト JIS コード
- ③EUC コード (拡張 UNIX コード Extended Unix Code)
- ④Unicode

(5) 文字コードの課題

- ①欠字・外字の問題
- ②略字体の代用?
- ③Unicode の課題