

2.2 字符表示方法

2.2.1 ASCII码

❖ ASCII (American Standard Code for Information Interchange)。

- 字符总数： 128

包含数字0~9，大小写英文字母，运算符，标点符号，标示符，格式控制符等等。

- 代码宽度： 7b

- 存储宽度： 7b(有效位)+1b(奇偶校验位)
=8b

❖ ASCII码示例

① 数字型

16进制	30	31	32	33	34	35	36	37	38	39
数字	0	1	2	3	4	5	6	7	8	9

② 字符型

16进制	41	42	43	44	45	46	47	48	49	4A
字符	A	B	C	D	E	F	G	H	I	J

③ 其它型

16进制	25	28	29	2A	3C	3D	3E	3F	40	7F
字符	%	()	*	<	=	>	?	@	DEL

2.2.2 汉字编码

1. 汉字的输入码

包括：数字码、拼音码、字形码

① 数字输入码

- ✓ 常用的是国标区位码，用数字串代表汉字；
- ✓ 将6763个二级汉字分为94个区，每区94位，把汉字表示成二维数组，数组下标就是**区位码**。

※ 数字编码输入的优缺点：

无重码，输入码与内部码的转换方便，代码难以记忆。

[例] 平 ↔ 3829 （位于码表38区、29位）

②拼音输入码

以汉字拼音为基础的输入方法（如微软拼音）

※拼音输入码的优缺点：

- ✓ 简单方便；
- ✓ 同音字多，重码率高，影响输入速度。

③字形输入码

根据汉字的书写形状来进行编码（如五笔字型）。

✓ 把汉字的笔划部件用字母或数字进行编码，按笔划顺序依次输入，就能对应一个汉字。

※字形输入码的优缺点：

- ✓ 简单方便，符合书写习惯；
- ✓ 重码率高，需要记住字形结构。

2. 汉字的内码

用于汉字信息的存储、交换、检索等操作的计算机内代码, 一般用**两个字节**表示。

✓ 机内码的基础是国标码

国标码: 10进制区位码转成16进制, 再加2020H

机内码: 国标码+8080H (与ASCII区别), 或者区位码转成16进制+A0A0H

[例] “啊” 的区位码 $1601_{10}=1001H$

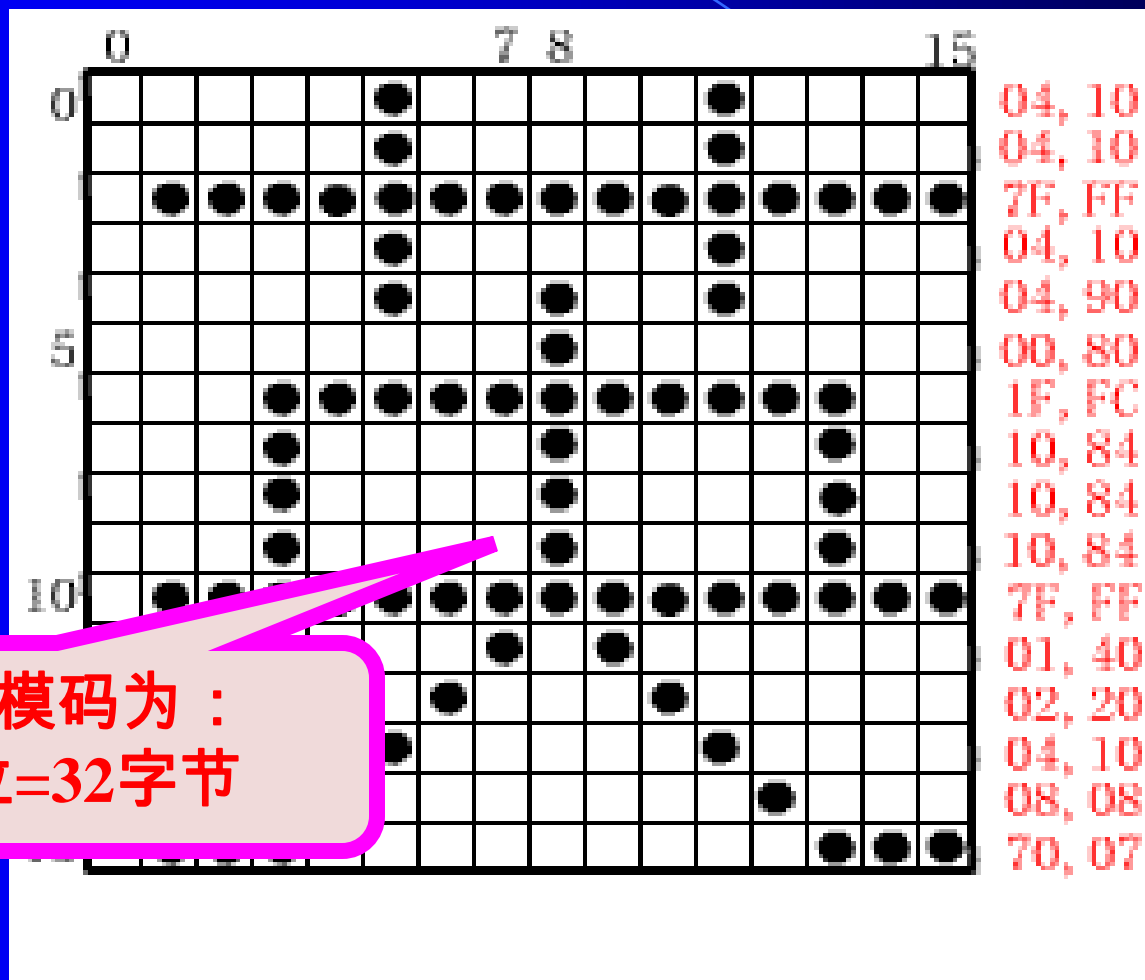
国标码= $1001H+2020H=3021H$

机内码= $1001H+A0A0H=B0A1H$

3. 汉字字模码

字模码用点阵表示的汉字字形代码, 是汉字的输出形式。

例如：



字模码

※特别说明：

- ①字模点阵用来构成汉字库，并非而机内存储。
- ②字库中存储每个汉字的点阵代码，用于汉字的显示输出或打印输出。
- ③当显示输出或打印输出时才检索字库，输出字模点阵，得到字形。

✓汉字的输入码、内码、字模码分别是用于计算机输入、内部处理、输出三种不同用途的编码；

✓各有用途，不能混为一谈。