

中华人民共和国

国家标准

信息交换用汉字编码字符集 基本集

GB 2312—80

1. 适用范围

本标准规定了汉字信息交换用的基本图形字符及其二进制编码表示。它适用于一般汉字处理、汉字通信等系统之间的信息交换。

2. 主要有关标准

- 2.1 GB 1988—80《信息处理交换用的七位编码字符集》
- 2.2 GB 2311—80《信息处理交换用七位编码字符集的扩充方法》

3. 术语含意

- 3.1 位组：表示一个字符的二进制位的有序集合。
- 3.2 代码；编码字符集：用一组明确的规则确立字符集；并确立该字符集中的字符与其位组之间一一对应关系。
- 3.3 字节：作为一个单位来处理的二进制位串。
- 3.4 图形字符：它是一种不同于控制字符的字符，通常这种字符具有可视图形，它可以用手写，印刷或显示等方法直观地表示出来。
- 3.5 转义序列：用于扩充代码的位组串。它由两个或两个以上的位组构成，其中第一个位组是转义字符（ESC）。
- 3.6 指明：在预先制订的一些字符集中，用规定的方法指出其某一特定的字符集。
- 3.7 调用：使指明的编码字符集成为能使用的状态。
- 3.8 表示：
 - (a) 按在指明和调用的字符集中字符的意义，使用规定的位组。
 - (b) 按增补的控制功能意义使用转义序列。

4. 图形字符

4.1 图形字符代码表

图形字符代码表见表1。

4.2 空白位置

代码表中，凡未填入图形字符的空白位置，均作为进一步标准化区域。

4.3 图形字符的种类及数量

本标准收一般符号、序号、数字、拉丁字母、日文假名、希腊字母、俄文字母、汉语拼音符号、汉语注音字母、汉字等，共7445个图形字符。

4.3.1 一般符号 一般符号202个。其中包括间隔、标点符号、运算符、单位符号以及制表符

国家标准总局发布
中华人民共和国第四机械工业部 提出

1981年5月1日 实施
华北计算技术研究所 起草

号等。其名称见附录 1。

4.3.2 序号 序号 60 个。其中 1~20, 共 20 个, (1)~(20) 共 20 个, ①~⑩ 共 10 个, (一)~(十) 共 10 个。

4.3.3 数字 数字 22 个。其中 0~9 共 10 个, I~XII 共 12 个。

4.3.4 拉丁字母 拉丁字母 52 个。其中大写字母 A……Z 26 个, 小写字母 a……z 26 个。

4.3.5 日文假名 日文假名 169 个。其中平假名 (包括浊音、半浊音及拗音、促音等小写假名) 83 个, 片假名 (包括浊音、半浊音及拗音、促音等小写假名) 86 个。

4.3.6 希腊字母 希腊字母 48 个。其中大写字母 A……Ω 24 个, 小写字母 α……ω 24 个。

4.3.7 俄文字母 俄文字母 66 个。其中大写字母 A……Я 33 个, 小写字母 а……я 33 个。

4.3.8 汉语拼音符号 汉语拼音符号 26 个 (包括带声调符号和其它符号的字母)。

4.3.9 汉语注音字母 汉语注音字母 ㄅ……ㄗ 37 个。

4.3.10 汉字 汉字 6763 个, 分成两级。第一级汉字 3755 个, 第二级汉字 3008 个。汉字的分级与排列见附录 2。

5. 编 码

5.1 编码表示 本标准对任意一个图形字符都采用两个字节表示, 每个字节均采用 GB 1988—80 及 GB 2311—80 中的七位编码表示。两个字节中前面的字节为第一字节, 后面的字节为第二字节。

5.2 在七位编码字符集中的指明与调用, 用三字符转义序列 ESC 2/4 4/1 指明, 通过移入字符 (SI) 调用。

5.3 图形字符在代码表中位置的表示方法 代码表分成 94 个区, 每个区有 94 位。区的编号从 1~94, 由第一字节标识; 位的编号也从 1~94, 由第二字节标识。代码表中的任何一个图形字符的位置用它所在的区号与位号标识。区号与位号之间用连字符相连。

例: 汉字“啊”用 16—01 表示。

注: 区号和位号与 GB 1988—80 中的列号与行号有如下对应关系:

列号/行号	2/1~7/14
区号	1~94
位号	1~94