

ICS 35.040
L 71



中华人民共和国国家标准

GB 18030—2005

信息技术 中文编码字符集

Information technology—Chinese coded character set

2005-11-08 发布

2006-05-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	Ⅲ
1 范围	1
2 规范性引用文件	1
3 原则	1
4 术语和定义	1
5 字汇	1
6 总体结构	2
7 字符的排列顺序	3
8 码位分配	4
附录 A (规范性附录) 双字节字符表	7
A.1 表的内容	7
A.2 表的说明	7
附录 B (规范性附录) 表意文字描述符	89
附录 C (规范性附录) 追加的汉字及部首/构件	90
附录 D (规范性附录) 四字节字符表	91
D.1 表的内容	91
D.2 表的说明	91
附录 E (规范性附录) 部分字符代码的说明	499

前 言

本标准的单字节编码部分、双字节编码部分和四字节编码部分的 CJK 统一汉字扩充 A(即 0x8139EE39—0x82358738)部分为强制性。

本标准代替 GB 18030—2000《信息技术 信息交换用汉字编码字符集 基本集的扩充》。本标准与上一版比,增加了编码汉字的数量,并补充规定了我国部分少数民族文字在本标准中的代码位置;本标准重新规定了字符“m”的编码位置;本标准的编码体系结构保持不变。

本标准的附录 A、附录 B、附录 C、附录 D 和附录 E 是规范性附录。

本标准由中华人民共和国信息产业部提出。

本标准由中国电子技术标准化研究所归口。

本标准起草单位:信息产业部电子工业标准化研究所、北京大学计算机技术研究所、北大方正集团、北京方正新天地信息网络科技有限责任公司、四通集团公司、中国电子信息产业发展研究院、中科院软件所、长城软件公司、四通利方公司、中软总公司、金山软件公司、联想集团有限公司。

本标准主要起草人:陈堃铎、黄疆、胡万进、张建国、陈壮。

本标准于 2000 年首次发布,本次为第一次修订。

信息技术 中文编码字符集

1 范围

本标准作为 GB/T 2311 体系的编码字符标准,规定了信息技术用的中文图形字符及其二进制编码的十六进制表示。

本标准适用于图形字符信息的处理、交换、存储、传输、显现、输入和输出。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准,然而,鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本标准。

GB/T 2311—2000 信息处理 字符代码结构与扩充技术(idt ISO/IEC 2022:1994)

GB 2312—1980 信息交换用汉字编码字符集 基本集

GB/T 11383—1989 信息处理 信息交换用八位代码结构和编码规则(idt ISO 4873:1986)

GB 12345—1990 信息交换用汉字编码字符集 辅助集

GB 13000.1—1993 信息技术 通用多八位编码字符集(UCS) 第一部分:体系结构与基本多文种平面(idt ISO/IEC 10646-1:1993)

3 原则

本标准向下与国家标准 GB 2312 信息处理交换码所对应的内码兼容。

本标准在字汇上支持 GB 13000 的全部中、日、韩(CJK)统一汉字(包括 CJK 统一汉字扩充 A、CJK 统一汉字扩充 B)字符和我国部分少数民族文字的字符。

4 术语和定义

下列术语和定义适用于本标准。

4.1

字符 character

供组织、控制或表示数据用的元素集合中的一个元素。

4.2

编码字符 coded character

字符及其编码表示。

4.3

字汇 repertoire

用编码字符集表示的一个指定的字符集合。

4.4

保留区 reserved zone

本标准中留作未来国家标准规定的区域。

5 字汇

本标准收录的字符以单字节、双字节或四字节编码。