

Windows95

GBK字符集汉字输入法

卢纹岱 刘书泽 宋峰 编著

99	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
4	槩 6A5C	槩 6A5D	槩 6A5E	槩 6A5F	槩 6A60	槩 6A62	槩 6A63	槩 6A64	槩 6A66	槩 6A67	槩 6A68	槩 6A69	槩 6A6A	槩 6A6B	槩 6A6C	槩 6A6D
5	槩 6A6E	槩 6A6F	槩 6A70	槩 6A72	槩 6A73	槩 6A74	槩 6A75	槩 6A76	槩 6A77	槩 6A78	槩 6A7A	槩 6A7B	槩 6A7D	槩 6A7E	槩 6A7F	槩 6A81
6	槩 6A82	槩 6A83	槩 6A85	槩 6A86	槩 6A87	槩 6A88	槩 6A89	槩 6A8A	槩 6A8B	槩 6A8C	槩 6A8D	槩 6A8F	槩 6A92	槩 6A93	槩 6A94	槩 6A95
7	槩 6A96	槩 6A98	槩 6A99	槩 6A9A	槩 6A9B	槩 6A9C	槩 6A9D	槩 6A9E	槩 6A9F	槩 6AA1	槩 6AA2	槩 6AA3	槩 6AA4	槩 6AA5	槩 6AA6	
8	槩 6AA7	槩 6AA8	槩 6AA9	槩 6AAD	槩 6AAE	槩 6AAF	槩 6AB0	槩 6AB1	槩 6AB2	槩 6AB3	槩 6AB4	槩 6AB5	槩 6AB6	槩 6AB7	槩 6AB8	槩 6AB9
9	槩 6ABA	槩 6ABB	槩 6ABC	槩 6ABD	槩 6ABE	槩 6ABF	槩 6AC0	槩 6AC1	槩 6AC2	槩 6AC3	槩 6AC4	槩 6AC5	槩 6AC6	槩 6AC7	槩 6AC8	槩 6AC9
A	槩 6ACA	槩 6ACB	槩 6ACC	槩 6ACD	槩 6ACE	槩 6ADF	槩 6AD0	槩 6AD1	槩 6AD2	槩 6AD3	槩 6AD4	槩 6AD5	槩 6AD6	槩 6AD7	槩 6AD8	槩 6AD9
B	槩 6ADA	槩 6ADB	槩 6ADC	槩 6ADD	槩 6ADE	槩 6ADF	槩 6AE0	槩 6AE1	槩 6AE2	槩 6AE3	槩 6AE4	槩 6AE5	槩 6AE6	槩 6AE7	槩 6AE8	槩 6AE9
C	槩 6AFB	槩 6AFC	槩 6AFD	槩 6AFE	槩 6AFB	槩 6AFC	槩 6AFD	槩 6AFE	槩 6AFB	槩 6AFC	槩 6AFD	槩 6AFE	槩 6AFB	槩 6AFC	槩 6AFD	槩 6AFE
	槩 6B08	槩 6B0C	槩 6B0D	槩 6B0E	槩 6B08	槩 6B0C	槩 6B0D	槩 6B0E	槩 6B08	槩 6B0C	槩 6B0D	槩 6B0E	槩 6B08	槩 6B0C	槩 6B0D	槩 6B0E
	槩 6B18				槩 6B18				槩 6B18				槩 6B18			

GBK

Windows 95 GBK 字符集 汉字输入法

卢纹岱 刘书泽 宋峥 编著



国防工业出版社

· 北京 ·

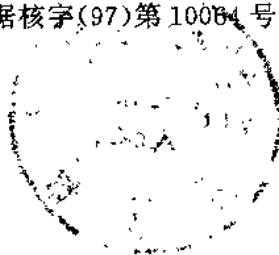
JS177/15

图书在版编目(CIP)数据

Windows 95 GBK 字符集汉字输入法/卢纹岱等编著.
—北京:国防工业出版社,1997.10
ISBN 7-118-01756-6

I. W… II. 卢… III. 窗口软件, Windows95-汉字信息
处理 IV. TP391

中国版本图书馆 CIP 数据核字(97)第 10084 号



国防工业出版社 出版发行

(北京市海淀区紫竹院南路 23 号)。

(邮政编码 100044)

北京怀柔新华印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 13¼ 313 千字

1997 年 10 月第 1 版 1997 年 10 月北京第 1 次印刷

印数:1—5000 册 定价:35.00 元

(本书如有印装错误,我社负责调换)

前 言

中国自古流传下来的文字经过数千年的发展,经过各种变异、简化和扩充,古今各种书籍里面所用的汉字浩如烟海。把具有五千年光辉灿烂的中华文化,以最新的技术、载体/媒体呈现和存储下来,用电脑来处理,是从事中文电脑事业的人们长期以来的梦想。为使这一梦想变成现实,我国有关方面的专家学者开发了各种汉字操作系统、汉字输入法,大大推动了我国计算机在各个领域中的应用。

随着改革开放的不断深入,祖国大陆、香港、台湾等海内外华人之间政治经济文化交流的不断扩大,计算机应用在中文信息处理领域中的不断普及,作为信息交流之基本载体——汉字的重要性日益增加。然而现行的各种汉字字符集标准,各自定义了不同的字汇和编码。在海峡两岸中、日、韩等使用汉字的国家、地区的电脑中文代码系存在着巨大差异。“同字不同码”的现象,给跨地区、跨国界的电子信息交流带来极大的不便。随着中文信息处理领域的扩大,GB 2312 存在所收汉字不够用、简繁不能共存、中文应用软件移植困难、代码体系间转换复杂、代价高昂等许多问题。这些问题使计算机的应用受到限制。

为了解决上述汉字字符集的扩充和汉字代码系的统一的问题,向新的国际标准 ISO 10646.1 靠拢,1995 年 10 月 5 日,全国信息技术标准化技术委员会召开特别会议,专家们通过反复研究讨论,最终确定了新的汉字扩充内码标准,并命名为“GBK”(国际扩展的汉语拼音缩写)。它突破了 GB 2312 的 6763 个汉字的限制,包括了 21003 个汉字。

微软公司在 Windows 95 中文版中,采用了 GBK 汉字内码体系,并且在内核、输入法、字型以及应用程序等方面给以全面的支持,率先在操作系统级支持大字符集,实现了简、繁、异、中、日、韩汉字的共存,并很好地兼容了原有的 GB 2312 的代码系,成为一个全新的中文平台。这对于中文信息处理、大型辞书编纂、古籍整理、电子出版、两岸及中外信息交流等都会有很大的帮助,并为将来转向 ISO10646 起到重要的作用。

本书共分八章。全面介绍了 Windows95 上可用的 GBK 字符集的中文输入法,包括 GBK 内码、国际码、流形码、郑码、表形码、全拼、双拼输入法。系统介绍了每种输入法的特点、编码规则、键位分布以及安装和使用方法等。在第八章专门介绍了 Windows95 所提供的中文输入法生成器的使用方法,使读者可以开发自己的中文输入法,并附有 GBK 字符集的字形、内码、国际码对照表。为方便广大用户,由北京翔舟新技术有限责任公司提供的流行码输入法、国际码输入法和 GBK 字符集打印程序的软盘将与本书一起出版发行。本书对使用计算机处理汉字(包括非常用汉字)信息的用户,以及想创建自己的汉字输入法的电脑爱好者将是有利的工具和学习参考用书。

由于作者水平有限,对各种输入法的理解难免有不足或不当之处,恳请有关方面的专家学者和广大读者批评指正。

本书在编写过程中,受到北京翔舟股份有限公司的大力支持。宋楚强、张福煜、王晓明、陈鲁敏、邓洪涌等同志在资料的提供、收集,文字录入、编辑等方面作了很多工作,在此一并表示感谢。

作者 1996 年 9 月

随书附带软盘使用说明

一、安装

将随书附带的软盘插入 3"软盘驱动器(假定为 A:)。在硬盘(假定为 C:)上建立一个子目录(假定为 C:\GBK),并将 A:盘内的 GBK.EXE 拷贝到该子目录中,在该目录内执行 GBK.EXE。具体操作如下:

```
C:\>MD GBK
C:\>CD GBK
C:\GBK>COPY A:GBK.EXE
C:\GBK>GBK
```

这时会在 GBK 的目录内产生流形码输入法、统一码输入法及 GBK 字符集打印程序所需的文件。

二、安装流行码输入法

起动 Windows95 后,执行 GBK 子目录中的 SETUP.EXE。详见本书的 4.4.1 节“流形码软件的安装”。

三、安装统一码输入法

起动 Windows95 后,执行 GBK 子目录中的 INSTALL.EXE。详见本书的 3.3.2 节“Unicode 输入法软件的安装”。

四、安装 GBK 字符集打印程序

将 GBK 目录内 GBKCHART 子目录内的 GBKCHART.DAT 和 VB40032.DLL 拷贝到 Windows95 的 SYSTEM 子目录(假定为 C:\WINDOWS\SYSTEM)。具体操作如下:

```
C:\GBK>CD GBKCHART
C:\GBK\GBKCHART>COPY GBKCHART.DAT \WINDOWS\SYSTEM
C:\GBK\GBKCHART>COPY VB40032.DLL \WINDOWS\SYSTEM
```

然后按附录一第 2 节的说明使用该 GBK 字符集打印程序。



内 容 简 介

GBK(国标扩展)字符集突破了 GB 2312 的 6763 个汉字的限制,包括了 21003 个汉字。

微软公司在 Windows 95 中文版中,采用了 GBK 汉字内码体系,并且在内核、输入法、字型以及应用程序等方面给以全面的支持,率先在操作系统级支持大字符集,在电脑中文内码所收汉字字汇上初步满足对汉字扩充的需求,实现了简、繁、异、中、日、韩汉字的共存,并很好地兼容了原有的 GB 2312 的代码系,成为一个全新的中文平台。这对于中文信息处理、大型辞书编纂、古籍整理、电子出版、两岸及中外信息交流等都会有很大的帮助。特别是随着电脑国际互联网(Internet)的发展,促使各国开始寻求各语言文种在共同的客户平台上进行交流的方法,使各国的语言文字占有同等的一席之地。这个共同平台的编码就是 ISO 10646。GBK 在当前充当这个平台,并为将来转向 ISO10646 将起到重要的作用。

本书共分八章,全面介绍了 Windows95 所提供的及当前市场上所有支持 GBK 字符集的中文输入法,包括 GBK 内码输入法、国际码输入法、流形码输入法、郑码输入法、表形码输入法、全拼、双拼输入法。系统介绍了每种输入法的特点、编码规则、键位分布以及安装和使用方法等。为使读者顺利、迅速掌握 GBK 字符集的输入法,本书在第二章介绍了 Windows95 中与输入法有关的操作。在第八章专门介绍了 Windows95 所提供的中文输入法生成器的使用方法,使读者可以开发自己的中文输入法。本书使用大量的实例、图示,叙述简单明了,并附有 GBK 字符集的字形、内码、国际码对照表;同时书后还附有由北京翔舟新技术有限责任公司提供的流行码输入法、国际码输入法和 GBK 字符集打印程序的软盘。

本书对所有需要使用计算机处理汉字(包括非常用汉字)信息的计算机用户,以及想创建自己的汉字输入法的电脑爱好者将是有利的工具和学习参考用书。

目 录

第一章 Windows 95 的 GBK 内码系和 GBK 内码输入法	1
1.1 GBK 内码系的产生和意义	1
1.2 GBK 字符集和 GBK 内码系	2
1.2.1 GBK 字符集	2
1.2.2 GBK 内码系	3
1.3 Windows 95 上的各种 GBK 输入法	5
1.4 GBK 内码输入法	6
第二章 Windows 95 中各种 GBK 输入法的安装与使用	7
2.1 在 Windows 95 中安装各种 GBK 输入法	7
2.2 GBK 输入法在 Windows 95 中的调用与使用	9
2.3 输入法功能和界面的设置	14
2.3.1 编码查询	15
2.3.2 词语联想	15
2.3.3 词语输入	16
2.3.4 逐渐提示	17
2.3.5 外码提示	18
2.3.6 光标跟随	18
2.4 帮助(Help)功能的使用	19
2.5 用户造词	19
2.5.1 手工造词	20
2.5.2 在线造词	21
2.6 编辑功能键	22
2.7 中文标点	22
第三章 Unicode 和 Unicode 输入法	24
3.1 Unicode 与 ISO 10646	24
3.1.1 ISO 10646 的产生	24
3.1.2 Unicode 的结构及内容	25
3.1.3 Unicode 与其他标准及代码系的关系	27
3.1.4 Unicode 的发展	28
3.2 Unicode 输入法	28
3.2.1 概述	28
3.2.2 Unicode 输入法软件的安装	29
3.2.3 Unicode 输入法的使用	30
第四章 流行码输入法	32

4.1 概述	32
4.2 输入规则	33
4.2.1 单字输入规则	33
4.2.2 词语及标点输入规则	35
4.3 键位分布	35
4.4 流行码的安装和使用	44
4.4.1 流行码软件的安装	44
4.4.2 流行码的使用	45
4.4.3 流行码帮助 (Help) 功能的使用	46
第五章 GBK 表形码输入法	52
5.1 概述	52
5.2 编码规则和取码方式	52
5.2.1 表形码的主要部件及分区	53
5.2.2 表形码的取码方式	59
5.3 键位分布与疑难字的编码	61
5.3.1 键位分布	61
5.3.2 难字编码示例	63
5.4 GBK 表形码的使用	64
5.4.1 GBK 表形码输入法的使用方法	64
5.4.2 表形码帮助 (Help) 功能的使用	65
第六章 GBK 郑码输入法	74
6.1 概述	74
6.2 编码规则	74
6.2.1 术语解释	74
6.2.2 汉字的笔画和字根	75
6.2.3 基本字根详解	79
6.3 编码规则和取码方式	91
6.3.1 汉字的分解	91
6.3.2 编码规则	93
6.3.3 取码方式	93
6.3.4 简码	96
6.3.5 繁体基根的编码	101
6.4 键位图	102
6.5 GBK 郑码输入法的使用	104
6.5.1 GBK 郑码的安装与使用	104
6.5.2 GBK 郑码输入法帮助 (Help) 功能的使用	105
第七章 GBK 全拼和双拼输入法	109
7.1 GBK 全拼输入法	109
7.1.1 GBK 全拼输入法概述	109

7.1.2 GBK 全拼输入法的编码规则	109
7.1.3 GBK 全拼输入法的安装与使用	110
7.1.4 GBK 全拼输入法帮助 (Help) 功能的使用	112
7.2 GBK 双拼输入法	113
7.2.1 概述	113
7.2.2 编码规则	114
7.2.3 GBK 双拼输入法的安装与使用	114
7.2.4 GBK 双拼输入法帮助 (Help) 功能的使用	116
7.2.5 GBK 双拼输入法键位图	118
第八章 在 Windows 95 上开发自己的输入法	119
8.1 输入法生成器 IME	119
8.1.1 码表文件的格式	119
8.1.2 调用输入法生成器 IME	121
8.1.3 输入法生成器的使用	122
8.2 帮助功能的开发	126
8.2.1 开发环境及工具	126
8.2.2 开发帮助功能的过程	127
8.2.3 工程文件	127
8.2.4 主题文件	127
 附录	
附录一 GBK 字符集打印程序使用说明	128
附录二 GBK 字符集字形、内码、Unicode 码对照表	131
 参考文献	204

第一章 Windows 95 的 GBK 内码系 和 GBK 内码输入法

微软公司的 Windows 95 因其使用简单、功能强大，特别是网络功能的加强，将个人电脑及其应用推向了一个新阶段；同时 Windows 95 又很好地保持了与现有软、硬件的兼容性，所以它一问世就受到全世界广大用户的青睐。创造了微软公司有史以来最佳的销售业绩。Windows 95 中文版（PRC），在英文版上市半年之后于 1996 年 5 月份推出。在这个全新的中文平台上，采用了被称之为 GBK（国标扩充）的电脑汉字内码。它突破了 GB 2312 的 6763 个汉字的限制，包括了 21003 个汉字，扩大了对汉字字符集的支持，以满足不断增长的中文信息处理应用的需要。

1.1 GBK 内码系的产生和意义

随着我们祖国大陆、香港、台湾等海内外华人之间政治经济文化交流的不断扩大，计算机应用在中文信息处理领域中的不断普及，作为信息交流之基本载体——汉字的重要性日益增加。然而现行的各种汉字字符集标准，如祖国大陆的 GB 2312-80、GB 12345-90，祖国台湾的 BIG-5 等，却各自定义了不同的字汇和编码方案。这就造成了汉字信息交换的巨大瓶颈，也带来了诸如汉字收字不足、简繁不能共存、中文应用软件移植困难、多代码体系间转换复杂、代价高昂等许多问题。这些问题已存在于现有的中文操作系统中，包括微软公司的中文 MS-DOS 6.22，中文 Windows 3.1/3.2 等。

我国为适应电脑中文文化的需要曾于 1980 年制定了《信息交换用汉字编码字符集-基本集》GB 2312-80，即大家常说的“国标汉字”。GB 2312-80 是使用最普遍的中文代码系。在这个标准中对 6763 个汉字定义了交换码，解决了绝大部分中文信息电脑处理问题。据某报社对一年的报纸进行的电脑汉字使用频度的统计，发现常用的汉字不过在 1200~1500 个之间，应该说 GB 2312 所收的 6763 个汉字一般来讲是够用了。

但是中国自古流传下来的文字经过数千年的发展，经过各种变异、简化和扩充，古今各种书籍里面所用的汉字浩如烟海。把具有五千年光辉灿烂的中华文化，以最新的技术、载体/媒体呈现和存储下来，用电脑来处理，是从事中文电脑事业的人们长期以来的梦想。随着中文信息处理领域的扩大，GB 2312 所收汉字不够用的困扰变得越来越严重，许多古籍书无法用电脑处理或以电子形式出版。后来虽陆续颁布了 GB 7589、GB 7590（第二、四辅助集）以及基本集对应的繁体字集 GB 12345 和第二、四辅助集对应的繁体字集等国家标准，进行汉字字符集的扩充，但它们所用的代码系与 GB 2312 是重叠的，由于不能同时使用多个字符集，因此在使用一个中文系统时可用字的总数并没有增加。

另外，现实中存在的海峡两岸在汉字简繁使用上的差异、在电脑中文内码系上的巨大差异（我国台湾和香港通行的汉字代码系为 BIG-5 码，包括 13000 多字），以及日本采用日本工业标准汉字 JIS、韩国也有韩国的汉字代码系，造成“同字不同码”的现象，给

跨地区、跨国界的电子信息交流带来极大的不便。

为解决汉字字符集的扩充和汉字代码系的统一的问题,国际标准化组织于1993年发布了《信息技术——通用多八位编码字符集》,即ISO 10646.1,国内也有人简称为“大字符集”(详见第三章)。该项国际标准共收入中、日、韩汉字共20902个,GB 2312及BIG-5的汉字均包括在内。

为了解决上述汉字字符集的扩充和汉字代码系的统一的问题,满足需求,并向新的国际标准ISO 10646.1靠拢,1995年10月5日,全国信息技术标准化技术委员会(简称“信标委”)的多八位特别工作组/汉字组在电子部计算机与微电子发展研究中心(CCID)召开特别会议,来自国内知名计算机公司、大学和研究室的专家们通过反复研究讨论,最终确定了新的汉字扩充内码标准,并命名为“GBK”(“国标扩展”的汉语拼音缩写)。

鉴于ISO 10646代码系在两字节内码系统上实现的困难,并考虑到与国内现有的中文软件的兼容性,信标委的多八位特别工作组/汉字组确定的新的汉字扩充内码标准GBK中包含了ISO 10646的大字符集,并在大字符集的基础上又扩充了101字;但并没有使用ISO 10646的代码系,而是将GB 2312的代码系加以扩充,这就是GBK代码系(详见1.2节)。

微软公司在Windows 95中文版中,采用了该汉字内码体系,并且在内核、输入法、字型以及应用程序等方面给以全面的支持。这样,Windows 95中文版率先在操作系统级支持大字符集,在电脑中文内码所收汉字字汇上初步满足对汉字扩充的需求,实现了简、繁、异、中、日、韩汉字的共存,并很好地兼容了原有的GB 2312的代码系,成为一个全新的中文平台。这对于中文信息处理、大型辞书编纂、古籍整理、电子出版、两岸及中外信息交流等都会有很大的帮助。特别是随着电脑国际互联网(Internet)的发展,促使各国开始寻求各语言文种在共同的客户平台上进行交流的方法,打破目前Internet上英文唱独角戏局面,使各国的语言文字占有同等的一席之地。这个共同平台的编码就是ISO 10646。GBK在当前充当这个平台,并为将来转向ISO 10646将起到重要的作用。另外,从避免开发同一软件的多种语言版本和减少电脑信息的重复量的角度来看,其意义也是非常深远的。

1.2 GBK 字符集和 GBK 内码系

1.2.1 GBK 字符集

GBK的目的是在保持与已有应用软件的兼容性的前提下,解决汉字收字不足、简繁共存、简化代码体系间转换等汉字信息交换的瓶颈问题,并向最终的国际统一四字节字符集标准ISO 10646.1迈进。

GBK字符集是在ISO 10646/CJK V2.0的基础上又扩充了101个汉字的字符集,它由 $20902 + 101 = 21003$ 个汉字和852个非汉字的符号组成。其汉字字汇包括:

- ① GB 2312-80中的全部汉字6763个。
- ② ISO 10646中除去GB 2312的6763个汉字后的中、日、韩(CJK)汉字共14139个。

- ③ 《简化汉字总表》中的尚未收入 ISO 10646 的 52 个简化字。
- ④ 《康熙字典》中的尚未收入 ISO 10646 的 28 个汉字部首及构件。
- ⑤ 兼容汉字 21 个。

GBK 字符集的非汉字符号包括:

- ① GB 2312-80 中的全部非汉字符号 684 个。
- ② BIG-5 中的已收入 ISO 10646 的非汉字符号 152 个。
- ③ 汉字结构符 13 个。
- ④ 追加的汉语拼音字母 2 个。

1.2.2 GBK 内码系

有了字集之后,还要给每个字符一个特定的代码,以和其他字符相区别(如:用 B0A1 代表“啊”,用 B0A2 代表“阿”等等)。这就涉及到这些代码值的范围如何、字符按代码排列的顺序如何、中间有没有空位等问题。解决了这些问题,每个字符都有了其特定的代码,就构成了一个代码系。

如 GB 2312 定义代码值的范围为 A1A1~FEFE (共 8836 个码位),非汉字符号安排在最前边,空 564 个码位后为 3755 个常用汉字,以拼音为序,再空 5 个码位之后为 3008 非常用汉字,以部首和笔划为序,最后还有 658 个空位。这就是 GB 2312 代码系(简称为 GB 代码系或国标码)。

若采用 ISO 10646 的有 20902 个汉字的字符集,用 4 个字节代表一个汉字,其前两个字节为 0000,后两个字节为 4E00~9FA5,并按汉字在《康熙字典》、《大汉和字典》、《汉语大字典》和《大字典》这四大字典中出现的页码的先后顺序排列,这就是 ISO 10646 的代码系,或称为 Unicode 代码系。

在某一中文操作系统内部,代表某一汉字的代码是唯一的,称之为内码;但这些内码很难记忆,造成了汉字输入的困难,所以又有了根据汉字字音、字形的特征进行输入的各种中文输入法,在这些输入法中也给每个汉字分配一个代码,称之为外码,外码不是唯一的。在操作系统实现时所采用的代码系又称之为该操作系统的内码系。

给 GBK 字符集中的 21844 个汉字和符号各安排一个码位,就构成了 Windows 95 的 GBK 内码系(简称为 GBK 内码系)。在该代码系中汉字的字序并没有采用 ISO 10646 中的字序,而是按照如下的规则排列的:

- ① GB 2312 的汉字仍按原 GB 2312 的顺序,即 I 级字按拼音、II 级字按部首/笔划排列。
- ② 除 GB 2312 外的其他 CJK 汉字按其 ISO 10646 中代码的大小排列。
- ③ 新追加的 101 个汉字(简化字、部首构件、兼容字)与上述两类字汇分开,单独排列。

该代码系的结构如图 1-1 所示。

从图中可以看到,原来的 GB 2312 代码系的第一字节和第二字节都向前扩充了,原来的 GB 2312 部分没有变化,这样保证了 GBK 对 GB 2312 的完全兼容。GBK 内码系总共有 23940 个码位,分为汉字区、符号区和用户区。

汉字区有三个：

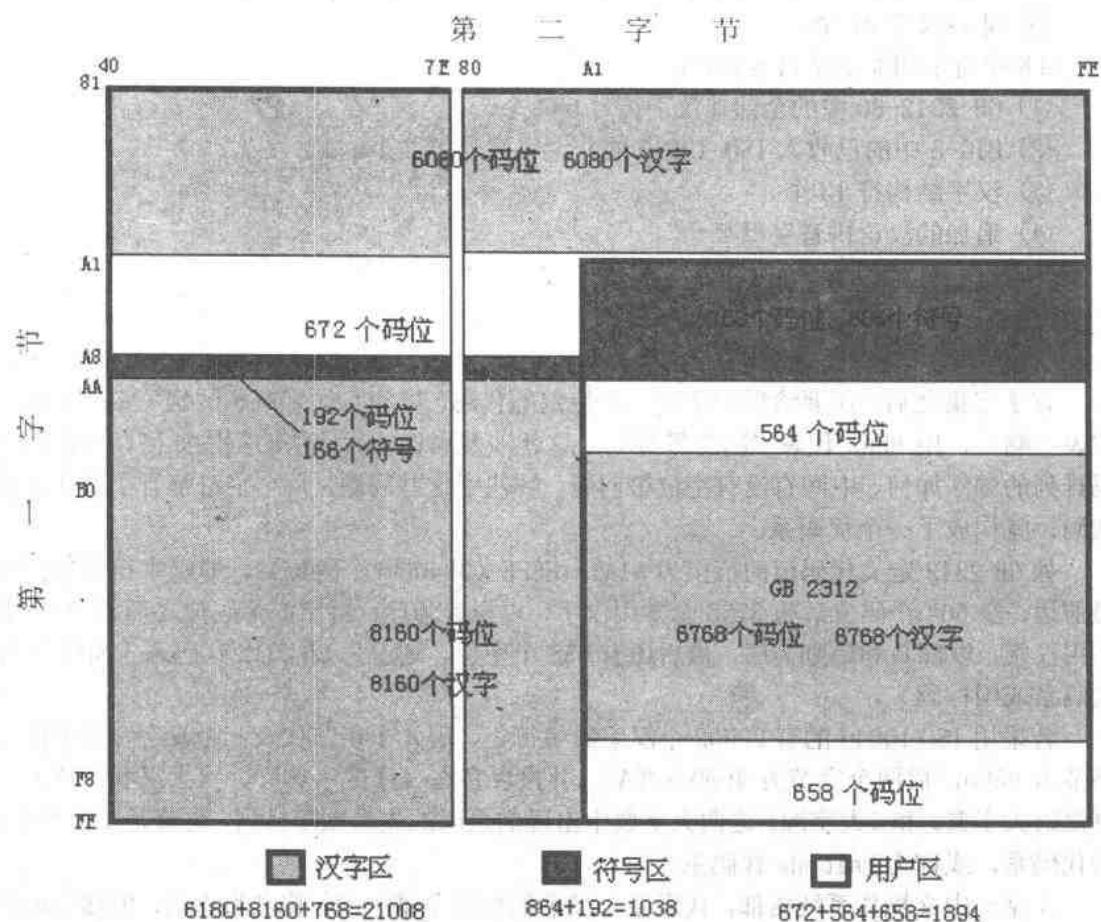


图 1-1 GBK 内码系码位图

B0A1~F7FE 共 6768 个码位，6763 个国标汉字；

8140~A0FE 共 6080 个码位，6080 个扩充汉字；

AA40~FEA0 共 8160 个码位，8160 个扩充汉字；

三区共有 21008 个码位，21003 个汉字。

符号区有两个：

A1A1~A9FE 共 846 个码位，684 个国标符号及两个追加的汉语拼音字母；

A840~A9A0 共 192 个码位，153 个扩充符号及 13 个汉字结构符；

两区共有 1038 个码位，841 个符号。

用户区有三个：

AAA1~AFFE 共 564 个码位；

F8A1~FEFE 共 658 个码位；

A140~A7A0 共 672 个码位；

三区共 1894 个码位，供用户造字用。

1.3 Windows 95 上的各种 GBK 输入法

从另外一个角度看，GBK 给各种中文输入法的设计者也提出了一个棘手问题。字多了，编码的难度也就大了，这有从量变到质变的问题。对于“声码”输入法，需要解决的是如何减少重码，提高智能化的问题。而对于“形码”输入法似乎就不那么简单了，搞不好可能会推翻原来的输入设计方案，重起炉灶了。如此说来 GBK 多少是对输入法可扩展/延伸性的一种考验，就是说原有的各种支持 GB-2312 的输入法能否继续支持 GBK？事实上也的确如此，只有少数的输入法可以继续支持 GBK。

我们知道，30 多年以来，计算机专家们一直都致力于研究用通用的电脑键盘实现中文录入，先后已出台了几百种输入法方案，而真正有实际价值的不过几十种。在众多的输入方案中，基本上可以分为两大类，即字音类和字形类。人们（无论开发者与使用者）最关心的问题在于两点——易学性和录入速度。

拼音作为我国中小学识字教育中的基本工具，普及率很高，被人们熟练掌握。这使得字音类输入法有着不可替代的易学优势。它的缺点在于：由于汉字的同音字太多，因此，重码率高成为字音类输入法速度进一步提高的主要障碍；另外随着计算机用的字符集的扩大，字音类输入法的另一个缺点——不能见字识码就显得更为突出，因为人们认识的字是有限的，一般人只认识 3000~5000 字，大量不认识的字无法输入。

字形类输入法通常运用拆字编码的方法，其重码率低的优点使熟练使用者可以达到很高的输入速度，也可以见字识码，更适合于 GBK 这样的人字符集使用。它的缺点是学会使用需要记忆很多的字根键位、繁杂的拆字方法和取码规则，这使非专业使用者在学习上产生了一定的困难。所以字形类输入法关键是如何安排其键位使之便于记忆，以及如何简化取码规则并使之符合使用者对汉字识别和拆分的习惯。

目前 Windows 95 中文版上的 GBK 输入法有：由微软公司提供的 GBK 内码输入法、GBK 全拼输入法、GBK 双拼输入法、GBK 表形码输入法和 GBK 郑码输入法；由北京翔舟新技术公司提供的 Unicode 输入法和 GBK 流行码输入法。全拼输入法和双拼输入法属于字音类输入法，郑码、表形码和流行码输入法属于字形类输入法。

本书试图用尽可能简单明了的文字，把 Windows 95 中文版的各种 GBK 输入法介绍给读者，起个抛砖引玉的作用，而不去评论各输入法的优劣，把它留给用户去评判。

1.4 GBK 内码输入法

GBK 内码输入法是 Windows 95 自带的输入法之一。其编码规则是以汉字的 GBK 内码作为输入码，击四键输入一个汉字，没有重码；GBK 内码的编码范围是：8140~FE7E 和 8180~FEFE。

使用时用鼠标左键单击屏幕右下角的输入法选择图标（或称语言指示器），并在弹出的会话框内选择“GBK 内码输入法”（如图 1-2）后，在屏幕的右下方出现输入法状态框（如图 1-3），这时就可以用 GBK 内码输入法输入汉字了。

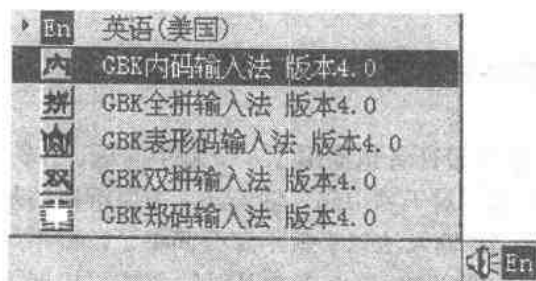


图 1-2 选择输入法



图 1-3 使用 GBK 内码输入法

如击 E785 四键，即可输入“鐳”字，并在屏幕光标位置显示“鐳”字；输入过程中输入的代码（在未达到四键时）显示在代码框内，并可用 Backspace 键进行删改，如图 1-3。从图中可见输入法状态框由五部分组成，除了指示当前使用的是哪种输入法外，还可以进行中 / 英文、全角 / 半角、中文标点 / 英文标点的切换，及调用软键盘和帮助信息等（详见第二章）。

在 GBK 内码输入法中可以使用查询键（或称通配键）功能，查询键为“？”键或空格键，查询键只能在第三键时使用。查询键的操作过程是：在输入两位合法的内码后，键入“？”键或空格键，系统会在重码选择框内显示其内码以这两个字符开头的（每页 10 个）汉字或符号（如图 1-4）。这时可用鼠标点该重码选择框上的箭头进行翻页，或用数字键选择你要输入的汉字。可用任意键退出查询状态。



图 1-4 查询键的使用

各字的 GBK 内码详见本书附录《GBK 字符集字形、内码、Unicode 码对照表》。

第二章 Windows 95 中各种 GBK 输入法的安装与使用

本章介绍 Windows 95 中各种 GBK 输入法使用的共同的部分和一些基本概念，具体每一种输入法的使用，除 GBK 内码输入法已在上一章介绍了外，均将在后续的各章中详细说明。

2.1 在 Windows 95 中安装各种 GBK 输入法

Windows 95 本身提供的各种 GBK 输入法（包括内码、全拼、双拼、郑码和表形码）在安装 Windows 95 时均已自动装入系统，并不需要专门的安装过程；由北京翔舟新技术公司提供的流行码和 Unicode 输入法的安装过程将在第三、四章中介绍。本节所说的安装并不是一般意义的软件安装。因为这些输入法虽已存在于你的 Windows 95 系统，但未必是可以随时使用的，要想使之成为可随时使用的还要经过本节所说的“安装”；如果某些输入法你不太可能用到，也可以暂时把它“删除”，以后想用时再“安装”。这种“安装”和“删除”对所有输入法都是一样的，都是用 Windows 95 的“控制面板”来完成的。

Windows 95 是独立完整的操作系统，它不再是基于其他操作系统的二级操作平台，因此，打开计算机电源，稍候，就会出现 Windows 95 的屏幕画面。在屏幕的左下角有一个“开始”按钮，用鼠标左键单击（以后在未加说明时均指左键）这个按钮，会弹出一个“开始”菜单。向上移动鼠标，当移动到“设置”的菜单项时，会弹出“设置”项的子菜单，在该子菜单上找到“控制面板”并单击它，就启动了控制面板。如图 2-1。

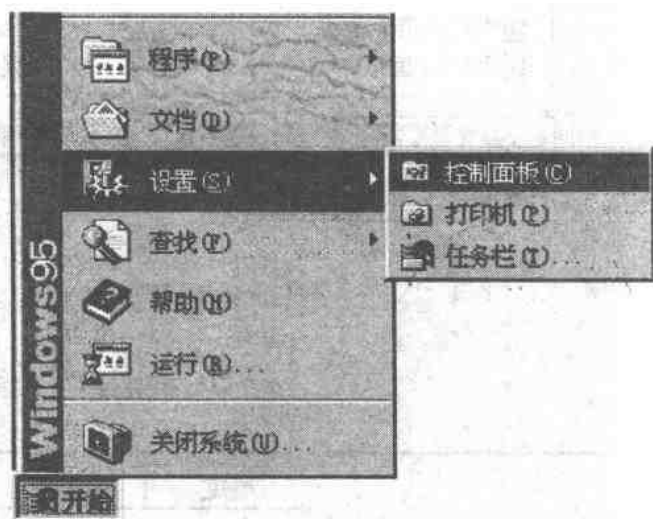


图 2-1 启动控制面板

在控制面板中,如图 2-2,用鼠标双击“输入法”图标,屏幕上就会出现“输入法”的对话框,如图 2-3。



图 2-2 控制面板

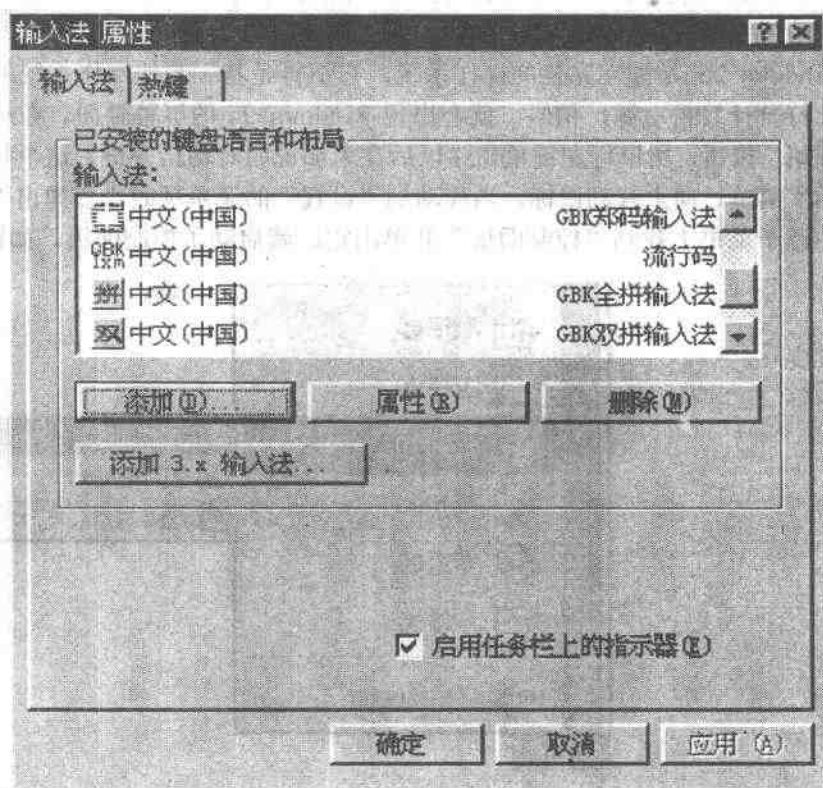


图 2-3 输入法对话框