

C
计算机技术专业

汉字录入与编辑技术(第三版)

中等职业学校电子信息类教材

电子工业出版社



教育部规划教材



中等职业学校电子信息类教材 计算机技术专业

汉字录入与 编辑技术

(第三版)

王式杰 邢小茹 郝 健 编
马开颜 审



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

URL: <http://www.phei.com.cn>

教育部规划教材

中等职业学校电子信息类教材(计算机技术专业)

汉字录入与编辑技术(第三版)

王式杰 邢小茹 郝 健 编
马开颜 审

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

全书共分九章,前四章对汉字处理做了介绍,讲解了区位、拼音、智能 ABC、五笔字型及 98 王码输入方法;后四章着重讲解了 Word 97 的初步知识,主要是编辑功能和表格的制作。为了取得更好的学习效果,本书还配有上机指导与习题集供师生选用。习题多来自作者多年的教学实践,也参考了各种计算机考试的内容和要求,使读者在学习本课程的同时,也能顺利通过相应的考试。

本书适用于各级各类中等职业教育学校的学生在计算机基础课中使用,也可做为各种短训班的教材使用。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究。

图书在版编目(CIP)数据

汉字录入与编辑技术/王式杰,邢小茹,郝健编.3 版.-北京:电子工业出版社,2000.4

中等职业学校电子信息类教材

ISBN 7-5053-5850-2

I. 汉… II. ①王… ②邢… ③郝… III. 汉字信息处理-输入-方法-技术学校-教材 IV. TP391.14

中国版本图书馆 CIP 数据核字(2000)第 04206 号

丛 书 名: 中等职业学校电子信息类教材(计算机技术专业)

书 名: 汉字录入与编辑技术(第三版)

编 者: 王式杰 邢小茹 郝 健

审 者: 马开颜

责任编辑: 吕 迈

排版制作: 电子工业出版社计算机排版室

印 刷 者: 北京四季青印刷厂

出版发行: 电子工业出版社 URL: <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×1092 1/16 印张: 12 字数: 308 千字

版 次: 2000 年 4 月第 3 版 2000 年 9 月第 3 次印刷

书 号: ISBN 7-5053-5850-2
G·510

印 数: 30100 册 定价: 16.00 元

JS416/32

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换;若书店售缺,请与本社发行部联系调换。电话 68279077。

出版说明

职业教育的教育质量和办学效益,直接关系到我国 21 世纪劳动者和专门人才的素质,关系到经济发展的进程。要培养具备综合职业能力和全面素质,直接在生产、服务、技术和管理第一线工作的跨世纪应用型人才,必须进一步推动职业教育教学改革,确立以能力为本位的教学指导思想。在课程开发和教材建设上,以社会和经济需求为导向,从劳动力市场和职业岗位分析入手,努力提高教育质量。

电子工业出版社受国家教育部的委托,负责规划、组织并出版全国中等职业技术学校计算机技术与实用电子技术两个专业的教材。电子工业出版社以电子工业为背景,以本行业的科技力量为依托,与教研、教学第一线的教研人员和教师相结合,已组织编写、出版计算机技术专业 and 实用电子技术专业的教材 70 余种,受到了广大职业学校师生的好评,为促进职业教育做出了积极的努力。

随着科学技术水平日新月异,计算机和电子技术的发展更是突飞猛进,而职业教育直接面向社会、面向市场,这就要求教材内容必须密切联系实际,反映新知识、新技术、新工艺和新方法。好的教材应该既要让学生学到专业知识,又能让学生掌握实际操作技能,而重点放在学生的操作和技能训练方面。在这一思想指导下,电子工业出版社根据《职业教育法》及劳动部颁发的《职业技能鉴定规范》,在教育部等相关部门的领导下,会同电子行业的专家、教育教研部门研究人员以及广大职业学校的领导和教师,在深入调查研究的基础上,制定了两个专业的指导性教学计划。该计划强调技能培养,充分考虑各学校课程设置、师资力量、教学条件的差异,突出了“宽基础多模块、大菜单小模块”灵活办学的宗旨。

新版教材具有以下突出的特点:

1. 发挥产业优势,以本行业的科技力量为依托,充分适应职业学校推行的学业证书和职业资格证书的双证制度,突出教材的实用性、先进性、科学性和趣味性。
2. 教材密切反映电子技术、特别是计算机技术的发展,不断推陈出新。实用电子技术专业教材突出数字化、集成化技术;计算机技术专业教材内容涉及多种流行软件及实用技术。
3. 教材与职业学校开设的专业课程相配套,注意贯穿能力和技能培养于始终,精心安排例题、习题,在把握难易、深广度时,以易懂、广度优先,理论原理为操作技能服务,够用即可。
4. 教材的编写一改过去又深又厚的模式,突出“小模块”的特点,为不同学校依据自己的师资力量和办学条件灵活选择不同专业模块组合提供方便。

另外,为满足广大职业学校教师的教学需要,我们还将根据每种教材的具体情况推出配套的教师辅助参考书以及供学生使用的上机操作/练习指导书。

随着教育体制改革的进一步深化,加之科学技术的迅猛发展,编写职业技术学校教材始终是一个新课题。希望全国各地职业学校的广大师生多提宝贵意见,帮助我们紧跟职业教育和科学技术的发展,不断提高教材的编写质量,以便更好地为广大师生服务。

全国职业高中电子信息类教材工作领导小组

1998 年 12 月

全国职业高中电子信息类教材工作领导小组

组长:

姚志清(原电子工业部人事教育司副司长)

副组长:

牛梦成(教育部职成教司教材处处长)

蔡继顺(北京市教委职教处副处长)

李 群(黑龙江省教委职教处处长)

王兆明(江苏省教委职教办主任)

陈观诚(福建省职业技术教育学会副秘书长)

王 森(解放军军械工程学院计算机应用研究所教授)

吴金生(电子工业出版社副社长)

成员:

褚家蒙(四川省教委职教处副处长)

尚志平(山东省教学研究室副主任)

赵丽华(天津市教育局职教处处长)

潘效愚(安徽省教委职教处处长)

郭菊生(上海市教委职教处)

翟汝直(河南省教委研究室主任)

李洪勋(河北省教委职教处副处长)

梁玉萍(江西省教委职教处处长)

吴永发(吉林省教育学院职教分院副院长)

王家诒(上海现代职业技术学校副校长)

郭秀峰(山西省教委职教处副处长)

彭先卫(新疆教委职教处)

李启源(广西教委职教处副处长)

彭世华(湖南省职教研究中心主任)

许淑英(北京市教委职教处副处级调研员)

姜昭慧(湖北省职教研究中心副主任)

张雪冬(辽宁省教委中职处副处长)

王志伟(甘肃省教委职教处助理调研员)

李慕瑾(黑龙江教委职教教材站副编审)

何雪涛(浙江省教科院)

杜锡强(广东省教育厅职业与成人教育处副处长)

秘书长:

林 培(电子工业出版社)

全国职业高中电子信息类教材编审委员会

名誉主任委员:

杨玉民(原北京市教育局副局长)

主任委员:

马叔平(北京市教委副主任)

副主任委员:

邢 晖(北京市教科院职教所副所长)

王家诒(上海现代职业技术学校副校长)

王 森(解放军军械工程学院计算机应用研究所教授)

韩广兴(天津广播电视大学高级工程师)

[实用电子技术编审组]

组长:

刘志平(北京市职教所教研部副主任)

副组长:

陈其纯(苏州市高级工业学校特级教师)

杜德昌(山东省教学研究室教研员)

白春章(辽宁教育学院职教部副主任)

张大彪(河北师大职业技术学院电子系副主任)

王连生(黑龙江省教育学院职教部副教授)

组员:

李蕴强(天津市教育教研室教研员)

孙介福(四川省教科所职教室主任)

沈大林(北京市回民学校教师)

朱文科(甘肃省兰州职业中专)

郭子雄(长沙市电子工业学院高级教师)

金国砥(杭州中策职业高级中学教研组长)

李佩禹(山东省家电行业协会副秘书长)

邓 弘(江西省教委职教处助理调研员)

刘 杰(内蒙古呼和浩特市第一职业中专教师)

高宪宏(黑龙江省佳木斯市职教中心)

朱广乃(河南省郑州市教委职教室副主任)

黄新民(上海现代职业技术学校)

[计算机技术编审组]

组长:

吴清萍(北京市财经学校副校长)

副组长:

史建军(青岛市科协计算机普及教育中心副主任)

钟葆(上海现代职业技术学校教研组长)

周察金(四川省成都市新华职业中学教研组长)

组员:

刘逢勤(郑州市第三职业中专教研组长)

戚文正(武汉市第一职教中心教务主任)

肖金立(天津市电子计算机职业中专教师)

严振国(无锡市电子职业中学教务副主任)

魏茂林(青岛市教委职教室教研员)

陈民宇(太原市实验职业中学教研组长)

徐少军(兰州市职业技术学校教师)

白德淳(吉林省冶金工业学校高级教师)

陈文华(温州市职业技术学校教研组长)

邢玉华(齐齐哈尔市职教中心学校主任)

谭枢伟(牡丹江市职教中心学校)

谭玉平(石家庄第二职教中心副校长)

要志东(广东省教育厅职业教育研究室教研员)

[通信技术编审组]

组长:

徐治乐(广州市电子职业高级中学副校长)

副组长:

陶宏伟(北京市西城电子电器职高主任)

陈振源(厦门教育学院职业教育教研室高级教师)

组员:

赖晖煜(福建省厦门电子职业中专学校主任)

许林平(石家庄市职业技术教育中心主任)

邱宝盛(山东省邮电学校副校长)

邹开跃(重庆龙门浩职业中学主任)

前 言

本书是在《汉字录入与编辑技术》第二版的基础上修订的版本。考虑到目前计算机的发展速度和现状,书中将原来的汉字编辑软件 WPS 及 CCED 撤换为目前流行的 Word 系统,在汉字录入方面也增加了一些新的内容。读者可以不经 DOS 的学习而直接使用本书。

全书共分为九章。考虑到目前计算机实际应用现状,我们除向读者介绍了计算机汉字处理的总体情况外,在汉字输入方法中保留了原有的区位法、全拼法、五笔字型法,新增加了智能 ABC 法和 98 王码的知识内容。在编辑软件方面,去掉了第二版中的 WPS 和 CCED 软件的内容,将 Word 97 做为编辑软件来讲解。书中用了四章的篇幅介绍了 Word 97 的初步知识,着重讲解了它在汉字输入、编辑、表格等方面的内容,而没有对该软件进行全面的阐述。

对于中等职业学校的学生来说,计算机课程是一门实践性很强的课程,汉字录入与编辑技术尤为如此。它是学习计算机的基础之一,也是学生应该掌握的基本技能之一。要想使学生能有一手过硬的操作技能,适当加大上机的操作训练是必不可少的。为了适应教学的需要,获得好的上机操作效果,本书配有上机指导与练习,其内容来自作者的教学实践以及社会上的各种计算机考试,采用循序渐进、逐步深化的方法,引导读者不断提高操作水平。

建议本课程至少使用 108 课时,按每学期 18 周计算,周课时为 6 课时,上一个学期;也可以安排成周课时为 3 课时,共两个学期,但不如前者效果好。这里包括了课堂授课与上机练习的时间。其中操作练习应占大多数课时。

本书第一、二两章由北京宣武职教中心学校郝健编写;第三、四两章由同校的王式杰老师编写;第五至八章由石家庄市第二职业中专邢小茹老师编写。第九章的上机指导 1~3,12~15 以及练习 1、2、5~8 由郝健编写,其余上机指导与练习由王式杰编写。全书由北京市职教计算机教研员马开颜老师审阅。由于时间的关系,加之作者水平有限,书中错误之处在所难免,敬请广大读者在使用本书的过程中提出宝贵意见。

编 者

1999 年 8 月

目 录

第一章 计算机汉字处理概述	(1)
第一节 编码及汉字处理	(1)
一、英文字符与 ASCII 码	(1)
二、汉字及其特点	(2)
三、计算机汉字处理	(2)
第二节 汉字字库	(3)
第三节 汉字的编码	(3)
一、计算机的汉字内部码	(3)
二、计算机的汉字交换码	(4)
三、计算机的汉字输出码	(4)
第四节 键盘及基本指法	(5)
一、键盘的种类	(5)
二、键的分布和使用特点	(5)
三、特殊功能键介绍	(6)
四、计算机打字的基本要求	(7)
五、打字的基本指法	(7)
第五节 Windows 简介	(8)
一、Windows 95 的启动与退出	(8)
二、基本操作	(9)
三、中文输入法的设置	(11)
第二章 汉字录入方法介绍	(13)
第一节 汉字录入简介	(13)
一、汉字录入的现状	(13)
二、语言录入与 OCR 识别	(13)
第二节 键盘汉字录入方法分类	(13)
一、概述	(13)
二、汉字录入的编码方法	(14)
第三节 区位法汉字录入	(14)
一、分区与划位	(14)
二、区位输入法	(15)
第四节 智能 ABC 汉字输入法	(16)
一、基本输入方式	(16)
二、汉字输入基本过程	(19)
三、智能特色与设置	(20)
四、使用技巧	(25)
五、提高指南	(26)
第五节 形码输入法	(27)

第三章 五笔字型汉字输入法	(28)
第一节 五笔字型输入法简介	(28)
第二节 五笔字型输入法	(28)
一、汉字的三个层次	(28)
二、基本字根及其键位	(29)
三、汉字的拆分原则	(33)
四、单字的输入	(36)
五、词语的输入	(39)
六、简码	(40)
七、容错码	(42)
八、学习键	(42)
第四章 新五笔字型汉字输入法	(44)
第一节 98 王码简介	(44)
第二节 系统的安装	(45)
第三节 98 王码系列软件的基本思想	(47)
第四节 王码键盘与码元布局	(47)
第五节 98 王码与 86 五笔字型的异同	(49)
一、98 王码与 86 五笔字型的共性	(49)
二、98 王码与 86 五笔字型的区别	(50)
第六节 98 王码的造词功能	(51)
第五章 Word97 简介	(52)
第一节 Word97 的启动和退出	(52)
一、从“开始”菜单中启动 Word97	(52)
二、利用快捷方式启动 Word97	(52)
三、退出 Word97	(53)
第二节 Word97 的界面	(54)
一、Word97 屏幕	(54)
二、Word97 菜单	(55)
第三节 使用工具按钮与快捷菜单	(56)
一、使用工具按钮	(56)
二、快捷菜单	(57)
第四节 使用帮助	(58)
一、使用“帮助”菜单中的“目录索引”命令	(58)
二、利用“屏幕提示”获取帮助	(59)
三、激活“Office 助手”	(60)
第五节 Word97 视图方式	(61)
第六章 输入和编辑文本	(63)
第一节 建立新文档	(63)
第二节 输入文本	(64)
一、进入中文输入状态	(64)
二、输入文本	(64)
三、段落的断开、合并	(65)
四、特殊符号的输入	(66)

第三节 插入点的移动与字符的修改	(67)
一、移动插入点	(67)
二、字符的插入、删除和修改	(68)
第四节 选定文本	(69)
一、使用鼠标选定文本	(69)
二、使用键盘选定文本	(71)
第五节 删除、复制和移动文本	(72)
一、选定文本的删除	(72)
二、选定文本的复制和移动	(72)
第六节 查找和替换	(76)
一、查找和替换文字	(76)
二、查找和替换特定格式	(77)
三、回到原先的位置编辑	(77)
四、定位	(78)
第七节 撤消和重复	(78)
一、撤消操作	(78)
二、重复一项操作	(79)
第八节 文件的保存	(80)
一、保存文件	(80)
二、自动保存	(81)
第九节 打开文档和插入文件	(82)
一、打开文档	(82)
二、插入文件	(82)
第七章 文档的排版	(84)
第一节 字符格式化	(84)
一、使用“格式”工具对字符进行格式化	(84)
二、使用“格式”菜单对字符进行格式设置	(88)
第二节 段落格式化	(93)
一、段落缩进	(93)
二、行距与段间距	(97)
三、对齐方式	(99)
第三节 边框和底纹	(102)
一、为选定段落添加边框	(102)
二、为选定段落添加底纹	(102)
三、应用举例	(103)
第四节 页面设置和插入分页符	(104)
一、页面设置	(104)
二、插入分页符	(106)
第五节 页眉、页脚、页码	(107)
一、创建页眉与页脚	(107)
二、浏览、编辑或编排页眉、页脚	(108)
三、页码	(109)
第八章 表格	(111)
第一节 创建与填写表格	(111)

一、创建表格	(111)
二、表格的填写	(112)
第二节 修改行高、单元格宽度和列宽	(113)
一、表格的选定	(114)
二、修改行高、单元格宽度和列宽	(114)
第三节 插入单元格、行和列	(118)
一、插入单元格	(118)
二、插入行	(120)
三、插入列	(120)
第四节 移动、复制单元格、行或列中的内容	(122)
一、用拖动的方法移动或复制单元格、行或列中的内容	(122)
二、用命令移动或复制单元格、行或列的内容	(123)
三、用快捷键移动或复制单元格、行或列的内容	(123)
第五节 删除单元格、行或列	(123)
一、删除单元格	(124)
二、删除行	(124)
三、删除列	(124)
第六节 合并、拆分单元格和表格	(125)
一、合并单元格	(125)
二、拆分单元格	(126)
三、拆分表格	(127)
第七节 移动表格的位置	(128)
第八节 绘制自由表格	(128)
第九章 上机指导与练习	(131)
第一节 英文打字上机指导	(131)
计算机基础知识练习	(132)
第二节 区位法汉字输入上机指导	(134)
第三节 智能 ABC 汉字输入法上机指导	(135)
汉字录入基础练习	(136)
第四节 键名与一级简码上机指导	(138)
第五节 二级简码上机指导	(139)
第六节 三级简码及成字字根的输入上机指导	(140)
第七节 两字词的输入上机指导	(141)
第八节 三字词的输入上机指导	(142)
第九节 四字词及多字词的输入上机指导	(144)
第十节 综合练习上机指导	(146)
五笔字型(86 五笔)练习	(149)
第十一节 98 王码上机指导	(151)
五笔字型(98 王码)练习	(154)
第十二节 Word97 的进入上机指导	(156)
Word97 基础练习	(157)
第十三节 Word97 基本编辑上机指导	(158)
Word97 编辑方法练习	(159)
第十四节 Word97 排版上机指导	(161)

Word97 排版练习	(162)
第十五节 Word97 制表上机指导	(164)
Word97 制表练习	(165)
第十六节 综合练习(一)上机指导	(166)
第十七节 综合练习(二)上机指导	(169)
附录一 Windows 95 与 Word 97 常用快捷键列表	(174)
附录二 字体与字号	(178)

第一章 计算机汉字处理概述

这一章中我们主要讲解在汉字计算机处理中有关的一些基本概念和计算机的基本操作,以使读者对这一工作的范畴及过程有个总体的了解。

第一节 编码及汉字处理

一、英文字符与 ASCII 码

在计算机中,处理文字符号时,要对文字和符号进行数字化处理,也就是用二进制数的编码来表示文字和符号。所以英文字母、数字和各种符号都必须按约定的规则用二进制编码在机器中表示。

在计算机系统中使用得最广泛的是美国信息交换标准码(American Standart Cord for Information Interchange),缩写为 ASCII。现在已被国际标准化组织(ISO)认定为国际标准,并在世界范围内通用。

ASCII 码有两个版本:7 位版本和 8 位版本。国际上通用的是 7 位版本,即用 7 位二进制数来表示英文字母、字符。7 位版本的 ASCII 码有 128 个元素,十进制编码值为 0~127,其中通用控制字符 33 个,控制字符不对应任何可显示或打印的实际字符,主要是用于控制计算机某些外围设备的工作和某些计算机的运行情况。另外的 95 个字符则是计算机终端能输入并且可以显示的字符,其中大、小写英文字母 52 个,阿拉伯数字 10 个,各种标点符号和运算符 33 个。具体字符编码见表 1.1。

表 1.1

高位 低位	000	001	010	011	100	101	110	111
0000	NUL	DEL	SP	0	@	P	'	P
0001	SOH	DC1	!	1	A	Q	A	Q
0010	STX	DC2	"	2	B	R	B	R
0011	ETX	DC3	#	3	C	S	C	s
0100	EOT	DC4	\$	4	D	T	D	t
0101	ENQ	NAK	%	5	E	U	E	u
0110	ACK	SYN	&	6	F	V	F	v
0111	BEL	ETB	'	7	G	W	G	w
1000	BS	CAN	(	8	H	X	H	x
1001	HT	EM	)	9	I	Y	I	y
1010	LF	SUB	*	:	J	Z	J	z

续表

高位 低位	000	001	010	011	100	101	110	111
1011	VT	ESC	+	;	K	[	K	{
1100	FF	FS	,	<	L	\	L	
1101	CR	GS	-	=	M	]	M	}
1110	SO	RS	.	>	N	^	N	~
1111	SI	US	/	?	O	-	O	DEL

在计算机中使用 7 位 ASCII 码时,每个字节只占用了 7 位,最高位恒为 0。8 位版本的 ASCII 码使用 8 位二进制数进行编码,当最高位为 0 时,称为基本 ASCII 码;当最高位为 1 时,形成扩充的 ASCII 码,又可表示 128 种字符。通常各个国家都把扩充的 ASCII 码作为自己国家语言文字的代码。

二、汉字及其特点

第一,汉字的文字量很大,最新的字典中有汉字六万多个,常用字也有几千个,这是一般外文不能相比的,因此,汉字的这一特点给汉字的计算机处理带来了很大的困难。

第二,汉字的字形复杂,它是一种二维的方块字,每个汉字又可以分为笔画、字根、整字等几个层次,无法直接使用计算机的英文键盘。

第三,汉字有多种字体,为了适应不同的需求,就要备有多种不同字形的汉字库,这就造成了计算机存储量过大的情况。

第四,汉字的发音呈多样化,一字多音与多字同音的现象屡见不鲜,造成汉字输入编码不易解决重码的问题。

第五,汉字的字义复杂,同一汉字在不同情况下使用,可以代表多种含义,因此,单独使用汉字的字义作为输入方法是不可行的。

第六,汉字排序难,英文只有 26 个字母,排序很容易,而汉字数量大、结构复杂、同音字多。

三、计算机汉字处理

文字信息处理的应用范围非常广泛,编辑文稿、建立文件档案资料、排版印刷等等。可以说,凡需用文字表达信息的场所,都可以使用文字信息处理技术。文字信息处理的实质就是把文字信息数字化,就是用一个固定的数码代表一个字母或文字,这一数码就叫做“代码”。在计算机内部处理文字信息时,就可以像处理纯数字一样来进行。需要输出时,再把替代的数码还原成相应的字母或文字。也就是说,汉字信息处理就是计算机对汉字代码的数据处理。

我们所讲的计算机文字处理的内容包括以下三个方面:

文字信息的输入(主要指各种输入方法)

文字信息的处理(主要指各种编辑方法)

文字信息的输出(主要指显示和打印的方法)

汉字信息处理就是对文字信息中的汉字信息进行处理的过程。现在世界上使用汉字的人越来越多,而我国是汉字的发源地,研究和开发汉字信息处理技术,是我国计算机工作者的责任,学习计算机汉字处理技术也是我国学生的义务。

第二节 汉字字库

要想让计算机能处理汉字,就需要在计算机中存有汉字的字模(字形),这就像铅字印刷时要有各种铅字字模一样。产生汉字字形的方法有模拟式和数字式两种,前者如字模板、全息照相等,文字质量较高,价格较便宜;后者是在计算机中使用的主要方法,它输出文字一致性好、速度快。

我们把每个汉字分别写在一个划分为 M 行、N 列的网格方块内,方块内的每个小方格是一个点,有笔画的方格涂成黑点,用二进制中的数字 1 代表;没有笔画的方格内部空白,用二进制中的数字 0 代表,这样一个汉字,就可以用若干个二进制数字来表示了。这个方块就叫做一个 $M \times N$ 的点阵,如果 M 等于 N,我们就简称这个点阵为 M 点阵,比如 16 点阵、24 点阵、32 点阵等。不论多大的点阵,字形信息一般都是以字节的形式存储在存储器中的。点阵数越大,所表现的字形越精确、越逼真,所占用的存储量也越大。一个 16 点阵的汉字字形就要用 16 乘 16 等于 256 个二进制位表示,也就是说要占 32 个字节的存储量,而这仅仅是实用中最小的点阵。图 1.1 是一个 16 点阵的汉字字形。

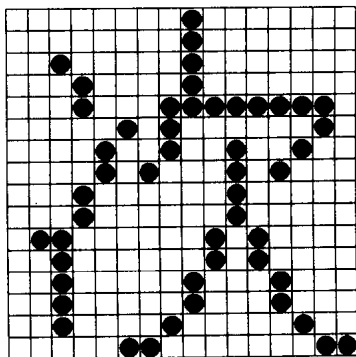


图 1.1 16 点阵的汉字字形

汉字字形按其存储方式可分为整字存储和压缩存储两大类。把汉字字形点阵信息按字节全部存入存储器中,需要时直接读出,就是整字存储,它使用方便、原理简单、速度快,但需要较多的存储量;而压缩存储不将字形信息直接存储,它先将点阵信息用压缩技术进行压缩,再将压缩后的信息存入存储器,使用时先将压缩信息还原成点阵信息再用。压缩的方法很多,应用较广泛的矢量法就是其中之一。

汉字字模除了字模尺寸和点阵大小的规格外,还要求有多种字体。存储这些信息的存储器叫做汉字库(或称汉字字模库)。一般采用磁盘或只读存储器做存储介质,前者存储成本较低,但读取速度也较慢,这种用磁盘做存储介质的汉字库叫做“软字库”。用只读存储器芯片(ROM)做存储介质的汉字库叫做“硬字库”,它读取速度快,价格稍高于软字库。

第三节 汉字的编码

一、计算机的汉字内部码

汉字内部码也称为内码或机内码。计算机输入汉字的外码以后,一般要转换为内码,才能进行存储、计算、传递等处理。内码是在设备和系统内部处理时所使用的汉字代码,它应能满足存储、处理和传输的要求,一个汉字的内码一般用两个字节来表示。中文信息处理系统有不同的代码系列,其内码亦不相同,有两字节、三字节和四字节汉字内码等。但国标基本集规定每个符号都用两字节代码表示。

计算机处理汉字实际上是处理汉字的代码,汉字内码包括存储码、运算码和传输码三种。存储码是存储汉字信息内容的代码,可为不等长码,以便节省存储空间;运算码是参与各种运算处理(如分类、合并、增、删、改等)的代码,一般为等长码;传输码是系统内部汉字传输(如送

显示器显示或送打印机打印等)用的代码,多为等长码。最理想的状况是存储码、运算码和传输码能统一起来,减少代码转换,但现在还没有做到。

内码通常是用汉字在字库中的物理位置来表示的,如用字库中的序号或存储位置来表示。两字节内码一般不与西文字符编码发生冲突,并与标准交换码有比较简明的对应关系,这样中西文兼容性较好。

二、计算机的汉字交换码

汉字交换码是在系统间或计算机间进行通讯或信息交换时用的代码,它是中文信息处理技术的基础,各系统或计算机所用的汉字交换码应该完全一致。

目前我国使用的交换码的标准主要有两个,一个是 GB1988(信息处理——信息交换用 ISO 编码字符集),它与国际通用的基本代码集相同,主要是字母、数字和符号等字符。另一个是 GB2312-80(国标基本集),它是我国规定的标准汉字交换码,在该集中每个符号用两字节表示,每个字节仅用低 7 位二进制位,最高位为 0。基本集的汉字内码与国标码(交换码)有一个明确的对应关系,即国标码加 8080(十六进制)就得到其内码。这种转换方法叫做“高位加 1 法”,即对原两字节高位均为 0 的国标码,分别在两字节的高位加 1,就得到了这个汉字的内码,反之也可从某汉字的内码求出它的国标码。

微机在英文文本工作方式时,输入码、交换码、存储、运算和传输都是用 ASCII 代码,对于汉字来说,问题就比较复杂,在中英文兼容的系统中,一字节的西文字符随时都可以出现在汉字之间,这就要求系统能正确区分英文和汉字,为此,必须对汉字交换码做出某种标识,经过标识的交换码就变成了内码,前面讲的高位加 1 法就是标识的方法之一。

三、计算机的汉字输出码

汉字输出码也称汉字字形码。对汉字字形点阵数字化后的一串二进制数,就叫做汉字的输出码。在显示和打印时,要将内码转换成相应的点阵信息,在屏幕上显示出来或在打印机上打出来。目前 16 点阵、24 点阵和 32 点阵的字型库都已经标准化,都有相应的国家标准。

汉字处理的基本过程是:对于要处理的汉字信息先由键盘输入外码,再经过软件变换为内码,根据内码算出汉字在字库中的地址并检索出该汉字的字形点阵,将字形点阵送往输出设备输出。若与其他设备有信息交换,则由内码转换为交换码再行交换。

输出汉字的字形时输出设备要有较高的分辨率,对于汉字显示规格,现在一般使用 15 × 16 点阵汉字字模,一般屏幕每行显示 40 个汉字,每屏显示 24 行字以及一、两行提示信息。显示器的分辨率为 640 × 420 以上,若要显示 24 点阵汉字,显示器的分辨率不能低于 1000 × 700。

常用的汉字的打印设备有针式打印机、喷墨打印机和激光印字机等品种。针式打印机比较耐用,但使用时噪音较大;喷墨打印机的打印效果远较针式打印机好,但消耗品价格较高;打印效果最好的是激光印字机,但价格比前两者要高得多。

汉字计算机处理中各种编码间的关系可用图 1.2 表示。