



职业教育精品实用教材

ZHIYE JIAOYU JINGPIN SHIYONG JIAOCAI

汉字录入与编辑技术

主 编 杨化文 琚 赞

西北工业大学出版社

职业教育精品实用教材

汉字录入与编辑技术

主 编 杨化文 琚 赟

西北工业大学出版社

【内容简介】本书是针对职业教育的基础性、先进性、实用性和操作性等特点,参照计算机相关行业的职业技能鉴定规范,以及中级技术工人等级标准进行编写的。全书共分为10章,分别为计算机汉字处理概述、常用汉字输入方法介绍、五笔输入法、其他常用汉字输入法、Word 2003 简介、文本的输入与编辑、文档的排版、表格、图形处理和综合操作。

本书内容新颖、结构严谨,在内容安排上循序渐进、由浅入深、重点突出,在讲解方法上采用理论引导、操作指导的方式教学,并通过图文并茂的方式形象地展现给读者。

图书在版编目(CIP)数据

汉字录入与编辑技术/杨化文, 琚贇主编. —西安: 西北工业大学出版社, 2008.6
职业教育精品实用教材
ISBN 978-7-5612-2383-3

I. 汉… II. ①杨…②琚… III. 文字处理—职业教育—教材 IV. TP391.1

中国版本图书馆 CIP 数据核字 (2008) 第 073527 号

出版发行: 西北工业大学出版社

通信地址: 西安市友谊西路 127 号 邮编: 710072

电 话: (029) 88493844 88491757

网 址: www.nwpup.com

印 刷 者: 陕西向阳印务有限公司

开 本: 787 mm×1 092 mm 1 / 16

印 张: 14

字 数: 334 千字

版 次: 2008 年 6 月第 1 版 2008 年 6 月第 1 次印刷

定 价: 22.90 元

出版说明

为了更好地贯彻《中共中央国务院关于深化教育改革全面推进素质教育的决定》精神，全面落实《面向 21 世纪教育振兴行动计划》中提出的职业教育课程改革和教材建设规划，职业教育精品实用教材编写组组织相关力量对实现职业教育培养目标、保障重点专业建设的主干课程的教材进行了规划和编写。

职业教育精品实用教材是面向职业教育的规范性教材，严格按照国家最新颁发的教学大纲编写，并通过了专家的审定。本套教材深入贯彻素质教育的理念，突出职业教育的特点，注重对学生的创新能力和实践能力的培养，在内容编排、例题设置和图示说明等方面努力创新，在满足不同学制、不同专业以及不同办学条件教学需求的同时，实现教学效果的最优化。

我们希望各地、各校在使用本套教材的过程中，及时提出改进意见和建议，使之不断地得到完善和提高。

职业教育精品实用教材编写组

前 言

随着计算机技术的日益普及和飞速发展,计算机已经成为人们日常工作和生活中必不可少的工具,也是知识经济时代的代表。无论是从事计算机软件、硬件工作的专业人员,还是普通的办公人员,在工作、学习和生活中都会接触到计算机,因此,汉字录入和编辑成为人们使用计算机必备的一项基本技能。

本书针对职业教育的基础性、先进性、实用性、操作性等特点,注重培养学生的实际操作能力和创新能力,书中内容以讲解实际应用中的案例为主,引出理论知识,再结合实验部分强化练习,将理论与实践紧密联系在一起。同时各章都配有相应的练习,以提升学生对本章内容实际应用的能力。

本书内容新颖,结构严谨,在内容安排上循序渐进,由浅入深,重点突出,在讲解方法上采用理论引导、操作指导的方式教学,并通过图文并茂的方式形象地展现给读者。

全书分为 10 章,第 1 章介绍了计算机汉字处理的基础知识,全面系统地概括了计算机汉字处理的有关概念及应用环境;第 2 章介绍了常用的汉字输入方法,内容主要包括汉字输入的简介,键盘汉字输入方法的分类,区位法汉字的输入及智能 ABC 汉字输入法;第 3、4 章介绍了各种常用输入法的知识及应用,主要包括五笔输入法、王码输入法、二笔输入法、郑码输入法和紫光输入法;第 5 章简单地介绍了 Word 2003 的概念及界面;第 6~9 章讲解了 Word 2003 的基本操作和编辑排版等知识,主要包括文本的输入与编辑、文档的排版、表格的应用、图形处理等;第 10 章通过 3 个综合操作对所学知识进行实际练习。

鉴于本书的内容全面及实例丰富,评审专家充分肯定了本书的编写思想、内容选择和编写体例,专家认为,“教材在内容选择和篇幅上符合学习需求,力求立足于基本操作,较好地采用了‘任务驱动’的编写方法。在教学内容的选择和编排上作了一些合理的调整和探索,比较适合职业教育的特点和实际教学需求。”

由于编者的编写经验有限,书中难免存在一些不足,恳请广大师生及读者批评指正,在此深表谢意。

编 者

目 录

第 1 章 计算机汉字处理概述	1	2.1.1 汉字输入的现状	29
1.1 编码及汉字处理	1	2.1.2 语音输入与 OCR 识别	29
1.1.1 英文字符与 ASCII 码	1	2.2 汉字键盘输入法分类	30
1.1.2 汉字及其特点	2	2.2.1 概述	30
1.1.3 计算机汉字处理	3	2.2.2 汉字输入的编码方法	31
1.2 计算机的汉字编码	3	2.3 区位码汉字输入	32
1.2.1 汉字内部码和汉字输入码	3	2.3.1 分区与划位	32
1.2.2 汉字交换码	4	2.3.2 区位汉字输入法	33
1.2.3 汉字输出码	6	2.4 智能 ABC 汉字输入法	33
1.2.4 汉字地址码	6	2.4.1 智能 ABC 输入法简介	33
1.2.5 各种汉字编码之间的关系	6	2.4.2 基本输入法	34
1.3 键盘及基本指法	7	2.4.3 汉字输入的基本过程	37
1.3.1 键盘的种类	7	2.4.4 智能 ABC 输入法的特殊 功能	37
1.3.2 键的分布和键区功能	7	2.5 上机训练	38
1.3.3 计算机打字的基本要求	9	2.5.1 五笔汉字输入法训练	38
1.3.4 打字基本指法	10	2.5.2 区位汉字输入法训练	39
1.4 Windows XP 操作系统简介	10	2.5.3 智能 ABC 输入法训练	40
1.4.1 Windows XP 简介	10	2.6 小结·练习	41
1.4.2 Windows XP 的基本操作	16	2.6.1 本章小结	41
1.4.3 Windows XP 的窗口	18	2.6.2 练习	42
1.4.4 Windows XP 窗口的操作	22	第 3 章 五笔输入法	43
1.4.5 系统设置窗口	24	3.1 五笔字型输入法简介	43
1.4.6 对话框	26	3.2 五笔字型输入法	44
1.5 英文打字训练	27	3.2.1 五笔字型字根及排列	44
1.6 小结·练习	28	3.2.2 五笔字型输入法的编码 规则	45
1.6.1 本章小结	28	3.2.3 五笔编码输入技巧	47
1.6.2 练习	28	3.2.4 提高输入速度的方法	48
第 2 章 常用汉字输入方法	29		
2.1 汉字输入简介	29		

3.3 常用五笔输入法	50	4.5.4 郑码输入法训练	85
3.3.1 智能五笔输入法	50	4.6 小结·练习	86
3.3.2 万能五笔输入法	56	4.6.1 本章小结	86
3.4 上机训练	60	4.6.2 练习	86
3.4.1 五笔字型输入法训练	60	第5章 Word 2003 简介	88
3.4.2 智能五笔输入法训练	61	5.1 Word 2003 概述	88
3.4.3 万能五笔输入法训练	61	5.2 Word 2003 的启动与退出	90
3.5 小结·练习	62	5.2.1 Word 2003 的启动	90
3.5.1 本章小结	62	5.2.2 Word 2003 的退出	91
3.5.2 练习	63	5.3 Word 2003 的工作界面	91
第4章 其他常用汉字输入法	64	5.3.1 标题栏	91
4.1 王码	64	5.3.2 菜单栏	92
4.1.1 王码简介	64	5.3.3 工具栏	93
4.1.2 汉字的5种笔画	65	5.3.4 标尺	95
4.1.3 汉字的3种字型	65	5.3.5 分割框	95
4.1.4 王码的键盘分布	65	5.3.6 文本编辑区	96
4.1.5 关于简码	71	5.3.7 滚动条	96
4.2 二笔输入法	72	5.3.8 状态栏	96
4.2.1 二笔输入法简介	72	5.4 Word 2003 基本操作训练	96
4.2.2 二笔输入法的特点	74	5.5 小结·练习	98
4.3 紫光拼音输入法	74	5.5.1 本章小结	98
4.3.1 安装及属性设置	75	5.5.2 练习	99
4.3.2 状态条介绍	76	第6章 文本的输入与编辑	100
4.3.3 输入风格	77	6.1 创建新文档	100
4.3.4 常用的输入技巧	77	6.2 输入文本	101
4.4 郑码输入法	81	6.2.1 输入文字	101
4.4.1 郑码简介	81	6.2.2 段落的断开、合并	102
4.4.2 郑码的基本原理	81	6.3 插入点的移动与字符的修改	102
4.4.3 助记词	82	6.3.1 移动插入点	102
4.5 上机训练	83	6.3.2 字符的插入、删除和 修改	103
4.5.1 王码输入法训练	83	6.4 选定文本	104
4.5.2 二笔输入法训练	83	6.4.1 用鼠标选定文本	104
4.5.3 紫光拼音输入法训练	84		

6.4.2 用键盘选定文本	104	7.2.3 分栏排版	124
6.5 文本的基本编辑	105	7.3 页面设置	125
6.5.1 移动文本	105	7.3.1 设置页眉/页脚	125
6.5.2 复制文本	105	7.3.2 页码	125
6.5.3 不同文件之间文本的复制或 移动	106	7.3.3 设置预留装订线区域	128
6.5.4 删除文本	106	7.3.4 设置节	129
6.6 查找和替换	106	7.3.5 设置脚注和尾注	130
6.6.1 查找并替换文字	106	7.4 边框和底纹	131
6.6.2 查找并替换特定格式	107	7.5 打印预览及打印输出	131
6.6.3 查找并替换段落标记、分页 符及其他项目	108	7.5.1 打印预览	131
6.6.4 使用通配符调整搜索条件	108	7.5.2 打印输出	132
6.7 撤销和恢复	108	7.6 样式	133
6.7.1 撤销操作	108	7.6.1 创建样式	134
6.7.2 恢复操作	109	7.6.2 应用样式	137
6.8 保存文档	109	7.6.3 修改样式	137
6.8.1 手动保存文档	109	7.6.4 删除样式	138
6.8.2 自动保存文档	110	7.7 模板	139
6.9 打开和关闭文档	111	7.7.1 使用模板创建文档	139
6.9.1 打开文档	111	7.7.2 创建模板	142
6.9.2 关闭文档	112	7.7.3 应用模板	143
6.10 输入和编辑文本训练	113	7.8 文档排版训练	143
6.11 小结·练习	116	7.9 小结·练习	144
6.11.1 本章小结	116	7.9.1 本章小结	144
6.11.2 练习	116	7.9.2 练习	144
第7章 文档的排版	118	第8章 表格	146
7.1 字符格式化	118	8.1 创建表格	146
7.1.1 字符格式化的方法	118	8.1.1 创建规则表格	146
7.1.2 设置字符格式	118	8.1.2 绘制自由表格	148
7.2 段落格式化	120	8.1.3 将现有文本转换成表格	149
7.2.1 设置段落格式	120	8.2 修改表格	150
7.2.2 分隔符	122	8.2.1 选定表格中的内容	150
		8.2.2 移动或复制表格中 的内容	151

8.2.3 调整整个表格或部分 表格的尺寸	152	9.1.1 绘图画布和绘图工具栏	167
8.3 插入和删除表格、行或列	153	9.1.2 绘制基本图形	168
8.3.1 删除表格或表格中的 内容	153	9.1.3 编辑图形	173
8.3.2 插入或删除行或列	153	9.1.4 图形的其他操作	176
8.4 合并、拆分单元格和表格	154	9.2 使用文本框	180
8.4.1 合并单元格	154	9.2.1 插入、编辑文本框	181
8.4.2 拆分单元格	154	9.2.2 文本框的链接	182
8.4.3 拆分、合并表格	154	9.2.3 插入图文框	183
8.5 在表格中嵌套表格	155	9.3 插入图片对象	184
8.6 表格格式化和表格数据 的简单计算	155	9.3.1 插入剪贴画	184
8.6.1 设置表格中文字的 格式和方向	155	9.3.2 从文件中插入图片	185
8.6.2 设置单元格中文本 的垂直对齐方式	156	9.3.3 设置图片格式	186
8.6.3 添加表格边框和底纹	157	9.4 艺术字	191
8.6.4 改变表格位置	159	9.4.1 插入艺术字	191
8.6.5 文字环绕表格	159	9.4.2 插入艺术字实例	192
8.6.6 跨页操作表格	160	9.5 图形、图片处理训练	195
8.6.7 表格计算和排序	160	9.6 小结·练习	196
8.7 表格制作训练	163	9.6.1 本章小结	196
8.8 小结·练习	165	9.6.2 练习	196
8.8.1 本章小结	165	第10章 综合训练	198
8.8.2 练习	165	10.1 综合操作一	198
第9章 图形、图片处理	167	10.2 综合操作二	200
9.1 绘制与编辑图形	167	10.3 综合操作三	201
		附录	206
		附录一 Windows 快捷键大全	206
		附录二 常用字体与字号	212
		附录三 十进制、十六进制与 字符对应表	213

第1章 计算机汉字处理概述

1.1 编码及汉字处理

计算机内部采用的是二进制数的处理方式，因此数值、文字、图形或其他各种字符，在计算机内部都是通过编码标准把它们转换为二进制数来进行处理。

1.1.1 英文字符与 ASCII 码

计算机在处理信息的过程中，只能采用数字编码的形式传输、存储和处理。例如，数值、文字、符号、图形、图像等都要先转化成数字编码形式，计算机才能进行处理。计算机广泛采用二进制编码处理信息，也就是采用 0、1 两个符号组成的基本码。因此，数字、英文字母和其他各种符号在计算机中都必须按编码标准统一用二进制编码表示。

美国标准信息交换代码，缩写为 ASCII，是目前在计算机中使用最广泛的编码。

ASCII 码由 8 位二进制数组成，其中最高位为校验位，用于在传输过程中检验数据的正确性；其余 7 位二进制数表示一个字符，共有 128 种组合。这些组合对应着十进制值的 0~127，其中有 95 个计算机终端能够输入并且可以显示的字符（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）；另外 33 个字符是通用控制字符，这些字符不能显示或打印，它们用于控制计算机某些外围设备的工作和计算机的一些运行情况。ASCII 字符编码如表 1.1 所示。

表 1.1 ASCII 字符编码表

高位 低位	000	001	010	011	100	101	110	111
0000	NUL	DLE	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
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	

续 表

高位 低位	000	001	010	011	100	101	110	111
1101	CR	GS	-	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	—	o	DEL

7 位版本的 ASCII 码使用 7 位二进制数进行编码，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。一般来说各国都把扩充的 ASCII 码作为自己国家语言文字的代码。



小贴士

字符：它是抽象的最小文本单位。字符就是对某种意义的图画表示，或者说形状表示。“A”是一个字符，“¥”也是一个字符。

字符集：字符的集合。例如汉字字符集，拉丁字符集等。

编码字符集：也就是一个字符集的编码形式，它为每一个字符分配一个唯一的数字。在没有包含全人类的字符编码集出来之前，各个国家都是自编自用，也就出现了西欧语言（又称 Latin - 1）、编码字符集（ISO 8859 - 1）、中文编码字符集（GB2312/BIG5/GBK/GB18030）、日文编码字符集，CJK 中日韩统一编码字符集等。

1.1.2 汉字及其特点

世界上所有的文字，都是记录语言的符号系统，是人类的书面交际工具，这是文字的共性。但是，各种文字又有它们各自不同的特点。汉字经过几千年的发展，也形成了汉字独有的特点，主要有以下几个方向：

（1）汉字的字形非常复杂。汉字是一种二维的方块字，每个汉字由笔画、字根或整字组成，在计算机中没有直接输入汉字的键盘。

（2）汉字的数量巨大。仅常用的汉字就有几千个，最新的字典中收录有 60 000 多个汉字，这是其他外国语言不能媲美的。也正是因为这个特点，使得汉字在计算机处理时比较麻烦。

（3）汉字的“字义”非常复杂。在不同的环境下，相同的汉字代表着不同的含义，这就限制使用“字义”作为汉字的输入方法了。

（4）汉字的字体多样化。由于汉字的字体繁多，在建立汉字字库时为了满足不同需求，需要配备多种不同字体的汉字库，这样就无形中增加了字库的存储量。

（5）汉字的发音多样化。汉字中的同音字和多音字现象很普遍，因此出现了重码问题。

（6）汉字的排序非常困难。由于汉字数量巨大，结构复杂，存在同音字与多音字等现象，使得对汉字的排序非常困难。

1.1.3 计算机汉字处理

计算机汉字信息处理,即计算机直接对汉字信息进行输入/输出和加工处理的技术。汉字信息处理就是把汉字信息数据化,即用一个固定的编码代表一个汉字,这些编码就称为“代码”,计算机在处理汉字信息时,只需要处理这些“代码”即可。汉字信息处理的内容包括以下3个方面:

- ✧ 汉字信息的输入(指汉字的各种输入方法);
- ✧ 汉字信息的处理(指汉字的各种编辑方法);
- ✧ 文字信息的输出(指显示和打印汉字的方法)。

汉字的字种繁多,字型复杂,汉字的信息处理与通用的字母数字类信息处理存在着很大的差异,突出表现在汉字输入/输出技术和汉字处理系统软件方面。但是,汉字信息在信息结构、交换、信息加工等方面与其他外国语言信息的加工又存在共性。因此,在汉字信息处理中多采用与外国语言信息处理兼容的途径,以便充分利用已取得的计算机信息处理技术资源。同时,汉字信息处理还包括研究适合汉字特点的操作系统和汉字计算机语言。

1.2 计算机的汉字编码

为使计算机可以处理汉字,同样需要对汉字进行编码。计算机进行汉字处理的过程实际上是各种汉字编码间的转换过程。这些汉字编码有汉字信息交换码、汉字输入码、汉字内码、汉字字形码和汉字地址码等。下面分别进行介绍。

1.2.1 汉字内部码和汉字输入码

1. 汉字内部码

汉字内部码又称为内码或机内码,是汉字在设备和信息处理系统内部最基本的表达形式,即计算机设备和信息处理系统内部存储、处理和传输汉字使用的代码。

汉字内部码是为了能在计算机内部对汉字进行处理、存储和传输而编制的汉字编码。因此,不论用何种形式输入何种格式的字符信息,在机器内部都要转换成统一的汉字内部码,然后才能在机器内部传输、存储和处理。

目前,一般以 ASCII 码作为汉字内部码来设计计算机系统。在这种机制中一般用两个字节来存放汉字的内码,两个字节共有 16 位,可以表示 65 536 个可区别码。如果两个字节各有 7 位,则可以表示 16 384 个可区别码,这样对于汉字就已经够用了。



小贴士

汉字字符必须和英文字符能相互区别开,以免造成混淆。英文字符的机内代码是 7 位 ASCII 码,最高位为“0”,汉字机内代码中两个字节的最高位均为“1”。不同的计算机系统所采用的汉字内部码有可能不同。

2. 汉字输入码

汉字输入码是为了使用户能够使用西文键盘输入汉字而编制的编码,又称为外码。汉字输入码有许多种不同的编码方案,一般可分为以下几类:

- ◇ 音码:以汉语拼音字母和数字进行编码,例如,全拼输入法和双拼输入法;
- ◇ 形码:根据汉字的字形结构对汉字进行编码,例如,五笔字型输入法;
- ◇ 音形码:以拼音为主,字形字义为辅进行编码,例如,自然码输入法;
- ◇ 数字码:用固定位数的数字进行编码,例如,区位输入法。

同一个汉字在不同的输入码编码方案中的编码一般也不同,例如,使用全拼输入法输入“嵌”字,就要键入编码“qian”(然后选字);而用五笔字型输入法,其输入码则是“mafW”。

1.2.2 汉字交换码

汉字交换码是系统间或计算机间进行信息交换时所用的代码。汉字交换码是中文信息处理技术的基础,因此,各系统或计算机所用的汉字交换码必须完全一样。

目前,我国使用的交换码有 GB1988(信息处理—信息交换用 ISO 编码字符集)和 GB2312—1980(信息交换用汉字编码字符集——基本集)两个标准。

1. GB1988

GB1988 与国际通用的基本代码集相同,主要是字母、数字、符号等字符。

2. GB2312—1980

汉字信息交换码是用于汉字信息处理系统之间或汉字信息处理系统与通信系统之间进行信息交换的汉字代码。

GB2312—1980 是我国规定的标准汉字交换码,又称国标码。在 GB2312—1980 中每个符号用两个字节表示,每个字节用低 7 位二进制(最高位为 0)。

通常使用“高位加 1 法”进行国标码和内码之间的转换。“高位加 1 法”是原两字节高位均为 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,反之也可以通过某汉字的内码求出它的国标码。

对应于国标码,一个汉字的内码也用两个字节存储。因为 ASCII 码是英文的机内码,要使汉字内码与 ASCII 码不发生混淆,应把国标码每个字节的最高位设为“1”,作为汉字内码。

- ◇ 汉字国标码与内码之间的关系是:汉字内码=汉字的国标码+(8080)₁₆。

例如:“大”字的国标码是 3473H,将国标码转换为内码的方法是,只须将国标码加上 (8080)₁₆ 即可。换算关系如下式:

$$\begin{array}{rcl}
 3473 & 16 & \text{国标码} \\
 + & 8080 & 16 \\
 \hline
 B4F3 & 16 & \text{内码}
 \end{array}$$

3. 汉字字形码

汉字字形码是存放汉字字形信息的编码,它与汉字内部码一一对应。每个汉字的字形码是

预先存放在计算机内的,通常称为汉字库。当输出汉字时,计算机根据内码在字库中查到该汉字对应的字形码,从而得知该汉字的字形信息,然后显示或打印输出。

目前,描述汉字字形的方法主要有“点阵字形”和“轮廓字形”两种方法。由于“轮廓字形”不常用,这里我们只介绍“点阵字形”。

由于汉字是由笔画组成的“方块字”,所以对于同一规格的汉字,无论其笔画是多少,都可以写在相同大小的方框里。如果使用 m 行 n 列的小圆点组成这个方块(这个方块称为点阵),那么每一个汉字都可以由点阵中的某些点组成。如图 1.1 所示的是汉字“工”字的 16×16 点阵字形。

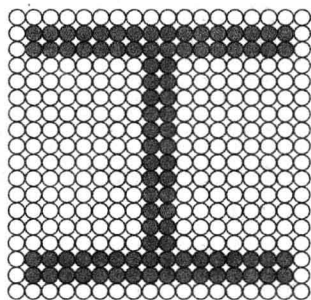


图 1.1 汉字“工”字的 16×16 点阵字形

在计算机中,用一组二进制数表示一个点阵。当某一点的二进制数为“1”时,表示该点为黑点,为“0”时表示该点为白点。一个 16×16 点阵有 256 个点,需要用 32 ($16 \times 16 \div 8$) 个字节来表示。同理, 24×24 点阵的汉字输出码需要 72 ($24 \times 24 \div 8$) 个字节存储空间, 32×32 点阵的汉字输出码需要 128 ($32 \times 32 \div 8$) 个字节存储空间。

显然,点阵中的行、列数越大,字形的质量越好,锯齿现象也就越少,但存储汉字输出码所占用的存储容量也就越大。汉字字形通常分为通用型和精密型两类。

◇ 通用型汉字字形点阵分为简易型 16×16 点阵、普通型 24×24 点阵和提高型 32×32 点阵 3 种;

◇ 精密型汉字字形点阵用于常规的印刷排版,字形一般在 96×96 点阵以上,占用的字节量较大,通常都采用信息压缩存储技术处理。

我国于 1981 年颁布了国家标准的汉字编码集,即《信息交换用汉字编码字符集——基本集》,国家标准代号是“GB 2312—1980”,简称交换码或国标码。下面介绍国标码的有关知识。

(1) 国标码的字符集,共收录了 7 445 个图形符号和两级常用汉字等。其中,有 682 个非汉字图形符号和 6 763 个汉字的代码。汉字代码中有一级常用汉字 3 755 个,二级常用汉字 3 008 个。

(2) 国标码的存储,国标码可以说是扩展了的 ASCII 码。两个字节存储一个国标码;国标码的编码范围为 2121H~7E7E。

(3) 区位码,也称为国标区位码,是国标码的一种变形。它把全部一级、二级汉字和图形符号排列在一个 94 行 $\times$ 94 列的矩阵中,构成一个二维表格,类似于 ASCII 码表。

◇ “区”:阵中的每一行,用区号表示,区号范围是 1~94;

◇ “位”:阵中的每一列,用位号表示,位号范围也是 1~94;

◇ “区位码”:汉字的区号与位号的组合(高两位是区号,低两位是位号);

◇ 实际上,区位码也是一种汉字输入码,其最大优点是一字一码,无重码,最大的缺点是难以记忆。

(4) 区位码与国标码之间的关系, $\text{国标码} = \text{区码} + (20)_{16} \text{ 位码} + (20)_{16}$ 。

例如,“中”字的输入区位码是“5448”。首先分别将其区号、位号转换为十六进制,得 $(3630)_{16}$;然后把区号和位号分别加上 20H,得“中”字的国标码是 $(3630)_{16} + (2020)_{16} = (5650)_{16}$ 。

1.2.3 汉字输出码

汉字输出码又称汉字字形码。汉字输出码是汉字字库中存储的汉字字形的数字化信息,用于汉字的显示和打印。字形码也称字模码,是用点阵来表示的汉字字形代码,它是汉字的输出形式,根据输出汉字的要求不同,点阵的多少也不同。简易型汉字为 16×16 点阵,普通型汉字为 24×24 点阵,提高型汉字有 32×32 点阵、 48×48 点阵等。

字形码的点阵信息量很大,所占存储空间也很大,以 16×16 点阵为例,每个汉字就要占用 32 个字节,两级汉字大约占用 256KB。

每个完整的汉字信息处理都离不开从输入码到机内码,再由机内码到字形码的转换。虽然汉字输入码、机内码、字形码目前并不统一,但是只要在信息交换时,使用统一的国家标准,就可以达到信息交换的目的。

我国国家标准局于 2000 年 3 月颁布的国家标准 GB8030—2000《信息技术和信息交换用汉字编码字符集·基本集的扩充》,收录了 27 万多个汉字。它彻底解决了邮政、户政、金融、地理信息系统等迫切需要的人名、地名所用汉字,同时也为汉字研究、古籍整理等领域打下了统一的信息平台基础。

1.2.4 汉字地址码

汉字地址码是指汉字库中存储汉字字形信息的逻辑地址码。在汉字库中,字形信息都是按一定顺序(大多数按标准汉字国标码中汉字的排列顺序)连续存放在存储介质中的,所以汉字地址码也大多是连续有序的,而且与汉字机内码间有着简单的对应关系,从而简化了汉字内码到地址码的转换。

1.2.5 各种汉字编码之间的关系

汉字的输入、输出和处理的过程,实际上是汉字的各种编码之间的转换过程。

首先通过汉字输入码将汉字输入计算机内,然后由输入字典将输入码转换为内码,再以内码的形式进行存储和处理。在汉字通信过程中,处理机先将汉字内码转换为适合于通信用的交换码,然后进行通信处理。

在汉字的显示和打印输出过程中,处理机首先根据汉字的机内码计算出地址码,然后根据地址码从字库中提取输出码,再显示或打印输出汉字。汉字的各种编码在汉字信息处理系统中的地位及它们之间的关系如图 1.2 所示。

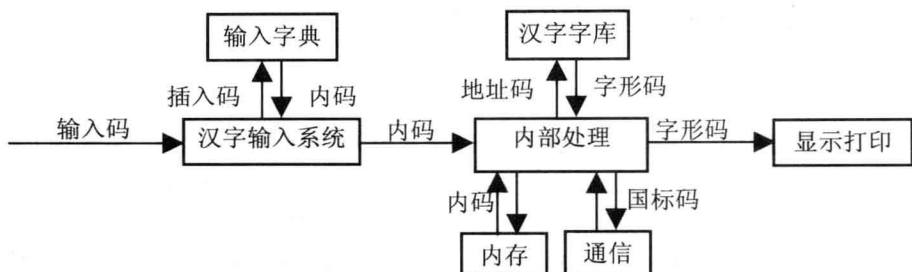


图 1.2 各种汉字编码之间的关系

1.3 键盘及基本指法

1.3.1 键盘的种类

根据键盘的作用可将键盘分为标准键盘、非标准键盘和专用键盘 3 类。其中非标准键盘和专用键盘大部分用于专用设备或特殊设备，一般很少用于电脑，这里我们只介绍标准键盘。

虽然标准键盘有很多种，如美、英、法、德、意大利、西班牙等国家都有各自的标准键盘，但这些标准键盘的键数都相同，而且键的大小、位置等也基本相同，只是键上标注的字符因各个国家的语言不同而不同。

不同国家的键盘都有自己的键盘程序，以保证能够正确地将键盘上的字符读入计算机中。例如，美国键盘程序固化在 ROM 中，电脑启动后默认所带键盘为美式键盘。我国计算机中一般都使用此类方法。

1.3.2 键的分布和键区功能

在键盘的使用中，有时也根据键数来区分键盘，其中常用的有 83 键键盘和 108 键键盘。

✧ 83 键键盘分为功能键区、打字键区和数字键区 3 个区域；

✧ 108 键键盘分为功能键区、主键盘区、编辑键区、小键盘区和指示灯，如图 1.3 所示。

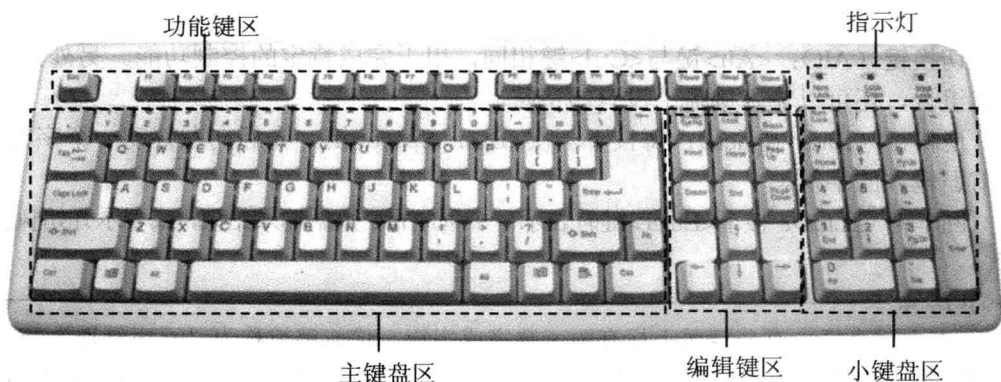


图 1.3 标准 108 键键盘

这里就以 108 键键盘为例介绍一下各键区的功能。

1. 功能键区

功能键区位于键盘的最上面,包括撤消键(Esc)、F 功能键(F1~F12)、唤醒键(Wake Up)、休眠键(Sleep)和电源键(Power)。功能键主要用于完成某些特殊的功能操作,具体功能由操作系统或应用程序定义,也可以由使用者自定义。使用功能键可以简化操作和节省时间。

✧ 撤消键(Esc):该键用于中止某个命令,在一些软件中,按下 Esc 键将弹出系统菜单。

✧ F 功能键(F1~F12):F 功能键在不同的情况下功能也不尽相同,主要根据软件定义的功能完成特定的操作。

✧ 休眠键(Sleep):按下 Sleep 键,可以使计算机处于休眠状态。

✧ 唤醒键(Wake Up):在计算机处于休眠状态时按下 Wake Up 键,可以唤醒计算机,使计算机处于工作状态。

✧ 电源键(Power):在开机状态下按下 Power 键,可以直接关闭计算机。

2. 主键盘区

主键盘区由基本字符键和辅助功能键组成。

(1) 基本字符键,主要包括英文字母键、数字键、标点符号键、特殊字符键和空格键。

(2) 辅助功能键,主要包括跳格键(Tab)、大写字母锁定键(Caps Lock)、Windows 键、换档键(Shift)、控制键(Ctrl)、切换键(Alt)、退格键(Back Space)和回车键(Enter)。

✧ 跳格键(Tab):<Tab>键用于移动光标,每按一次键,光标向右移动 n 个字节(n 可以由使用者自定义,默认为 8 个字节)的空格位置。

✧ 大写字母锁定键(Caps Lock):系统默认为小写字母输入状态,当按下<Caps Lock>键时,系统即处于大写字母输入状态,再次按下<Caps Lock>键时,退出大写输入状态,恢复小写字母输入状态。

✧ Windows 键:操作系统键,按下<Windows>键时可以快速打开开始菜单。

✧ 换档键(Shift):当需要键入大写字母或双字符键中的上档字符时,可以在按下<Shift>键的同时,按下相应字符键。

✧ 控制键(Ctrl):<Ctrl>键与其他键同时使用,可以实现操作系统中或其他应用程序中定义的功能。

✧ 切换键(Alt):<Alt>键与<Ctrl>键相同,用于完成特定的控制功能。另外,<Alt>键与数字键组合,还可以输入键盘上没有的字符。

✧ 退格键(Back Space):每按下一次<Back Space>键,光标向左移动一个字符,即删除一个前面输入的字符。

✧ 回车键(Enter):按下<Enter>键时表示一个程序行、数据输入结束,或表示开始执行命令。



小贴士

在主键盘区,有些键面上标有两个数字或字符,这些键被称为“双字符键”。双字符键上面的字符称为“上档字符”,下面的字符相应地称为“下档字符”。<Shift>键的作用就是切换上下档字符。