



WM9801

王永民 著

98

版反

五笔字型 培训教程

新旧版五笔字型的区别

王码键盘的码元分布规律

无拆分编码法

简码重码和容错码

词语输入规则

98 王码标准教材之二



電子工業出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

TP391
W462

98 王码标准教材之二

98 版五笔字型培训教程

王永民 著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

JS/36/5/

内 容 简 介

98 王码是五笔字型发明人王永民教授 10 年磨一“键”,在 86 版五笔字型科学体系的基础上推出的形码音码兼备、简体繁体相容的一整套汉字输入法及其多平台自动适配的系列软件,被国内外专家称之为世界上“汉字键盘输入的第一个全面解决方案”。

《98 王码标准教材》丛书全方位、多层次地介绍了 98 王码包含的各类输入法及其系列软件。本书是这套丛书的第二本,针对五笔字型的新、老用户,由浅入深、条理清楚地讲述了 98 版五笔字型的使用及训练方法。通过学习本书,读者将达到高级录入水平。

本书适合于各类学校、培训班作为培训教材,也可供电脑爱好者自学使用。

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,翻印必究。

丛 书 名: 98 王码标准教材

书 名: 98 版五笔字型培训教程

著 者: 王永民

责任编辑: 李 影

特约编辑: 刘维华

策 划: 郭 立

排版制作: 电子工业出版社计算机排版室

印 刷 者: 北京天宇星印刷厂

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: 15.25 彩插: 2 字数: 340 千字

版 次: 1998 年 11 月第 1 版 1998 年 11 月第 1 次印刷

书 号: ISBN 7-5053-4956-2
TP·2436

定 价: 26.00 元

凡购买电子工业出版社的图书,如有缺页、倒页、脱页者,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: (010)68279077

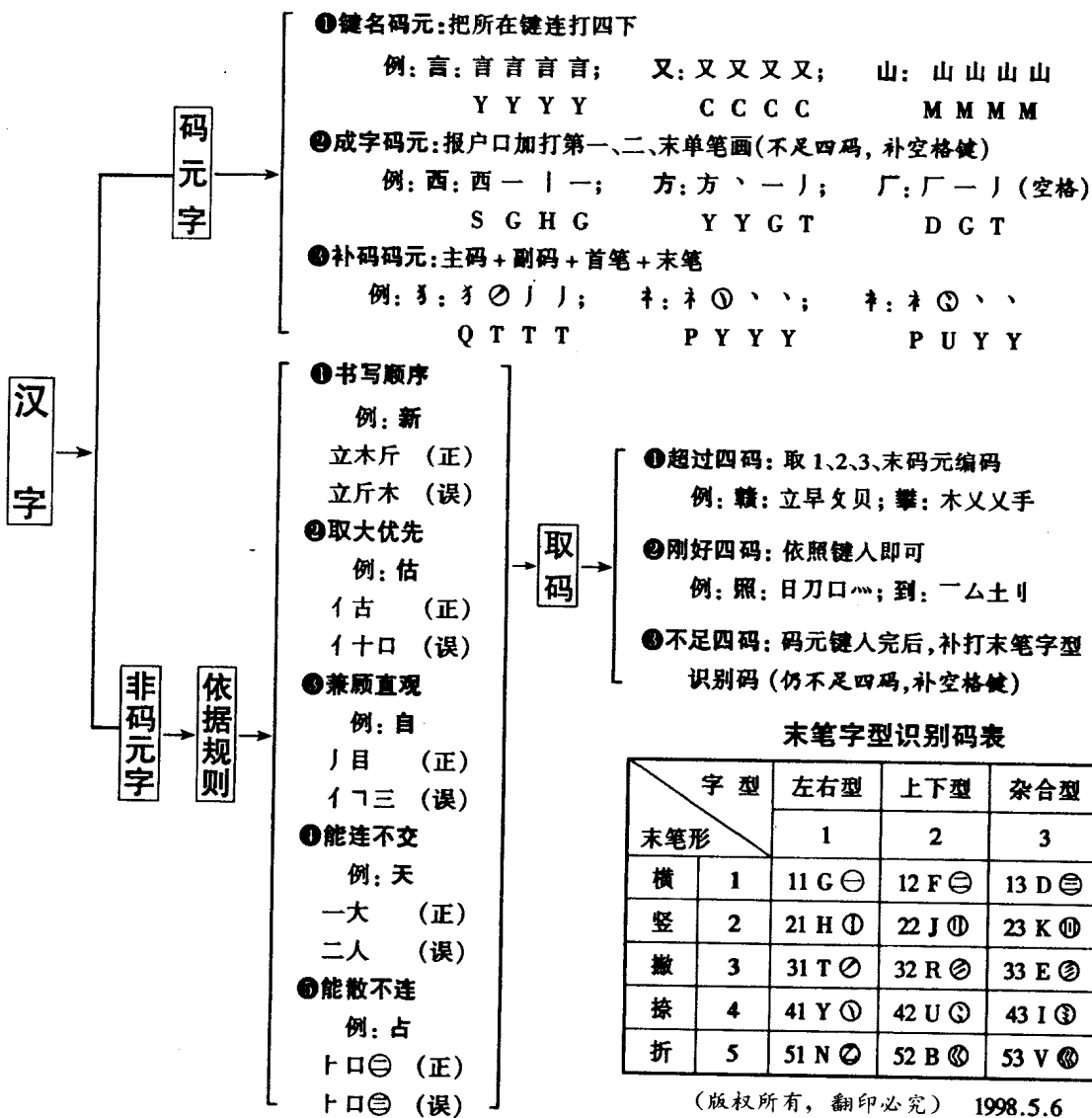


设计·监制：王永民教授



金	人	月	白	禾	言	立	水	火	之
夕	尹	力	丘	久	文	辛	永	业	之
鳥	几	毛	叉	彳	方	六	南	广	之
35 Q	34 W	33 E	32 R	31 T	41 Y	42 U	43 I	44 O	45 P
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊	十	圭	上	月	川	甲	
廿	西	艮	未	夫	止	早		四	
升	甫	厂	甘	卦	七	早		皿	
15 A	14 S	13 D	12 F	11 G	21 H	22 J	23 K	24 L	
工	木	大	土	王	目	日	口	田	
戈	丁	戊</							

98 王码编码流程图



学一阵子，用一辈子

——代前言

为了向电脑输入汉字，就得由汉字的平面图形入手，从中选取必要信息，再编制成代码，即所谓“汉字编码”。

之所以必须这么做，必须“拐个弯”，间接地输入汉字，是因为外国人发明电脑时，丝毫没有考虑汉字；而我们的祖先创造汉字时，也不可能预见到今天会有电脑出现。因此，当中国人不得不用电脑的时候，就发现，谁也没有办法把几千甚至几万个汉字，一个个地摆到只有 26 个字母键的键盘上，像英文那样直接按键输入。

接触汉字输入的人们，儿童、少年、青年、老年有之；小学、中学、大学文化水平者有之；业余、兼职、专职录入人员有之；北方人、南方人、海外华人有之。总之，凡是遇到电脑要处理汉字的场合，几乎都会遇到“汉字输入”，这本是一个文化问题，却成了一个社会问题。

在我们还没来得及将汉字输入技术大规模地纳入中小学教育，使中小生长大之后，人人“不用学习”便会输入汉字之前，大批中青年甚至老年人士，都必须“补课”，才能“愉快”地输入汉字。既然是“补课”，就常遇到没时间、不习惯、记忆力不好、手指不听使唤等问题。

不少人把打键进步慢，归罪于五笔字型，这是不科学的。君不见英美人士如不从小打键，长大了再“补课”，也必须交 200 美元学费，再练上一两个月才能熟练掌握标准指法？这跟“汉字”和“王码”有何关系呢？指法训练从小开始，一星期就能掌握，长大了再补课，一个月也不会灵活。

不信，从今天起，你换个手拿筷子吃饭，或换个手拿笔写字，体会一下就不难明白：成年人的手有多笨！由此你不难推知，要让从未训练过的十个“笨”手指“总动员”，协调灵活来打键，是多么难办！

如今，主张废除汉字的呼声确实小了，但不学就想有“术”的呼声却处处可闻。其实，不少人在“补课”中，之所以到处寻找“不用学习”、“天生会用”、“五分钟学会”、“读音输入、出口成章”等等“全新”的输入法，甚至天天抱怨打不快而“嫁祸于人”，其根本原因，是把技能技巧的训练看得太简单、太容易了。

速度和效益已成为信息时代的普遍追求。我们大可不必要求国人都像纽约

人那样,在传送带状的电梯上跑步争分夺秒,在地铁车厢中头戴耳机学习新知识而时时有紧迫感,可我们某些直到下岗才“突然发现”自己什么都不会的人士,确实也太懒了点。他有大把的时间打牌、“修长城”、唱卡拉 OK,却拿不出一两天的时间学习编码,更不要说练习打键,学习软件,还一天到晚期盼着“不学就会”、“张口输入”的奇迹出现。

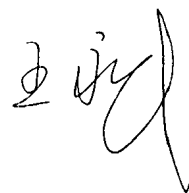
作者认为,键盘输入,学一阵子,用一辈子,是决不会被“淘汰”的!因为这是最基本的、不必增加设备、不受条件制约,因而到处有用的输入方式。正像有了汽车、自行车,人却永远需要走路一样。

使用键盘,在西方社会已有 100 多年的历史,早已成了连小孩子都具备的文化知识和技能。可中国的“键盘文化”才刚刚开始,不用不可,用了头疼!中国人要适应这种“洋玩艺”,要知道寻找“灵丹妙药”,代替打键是多么无知,要体味到不会打键就没人给你工作和饭吃的痛苦,至少还需要一代人的时间。

为满足不同层次读者学习五笔字型的第二代定型版本——98 王码的需求,作者与电子工业出版社合作,用短短的两个月时间,日夜劳作,一气呵成,推出了《98 王码标准教材》系列丛书,全方位、多层次地介绍 98 王码系列软件。

这套丛书用全面、经典、系统的语言诠释王码公司十年磨一“键”才开发完成的软件产品,并配备了根据十几年王码培训的经验精心设置的练习。《98 版五笔字型培训教程》是这套丛书的第二册,共分 10 章,面向王码新老用户,全面介绍了 98 王码系列软件。

作者希望本书对学用电脑,特别是学用王码者有所助益,并愿和出版社及广大读者一起,为推进我国文字处理信息产业的历史进程,再接再厉,百尺竿头,更上一层楼。



于一九九八年八月

目 录

第一章 汉字输入法概述	(1)
第一节 汉字的优缺点	(3)
第二节 汉字输入法的分类	(4)
第三节 86 版五笔字型	(8)
第四节 98 版五笔字型的特点	(10)
第五节 98 版五笔字型的新增功能	(12)
第六节 98 版五笔字型与 86 版五笔字型的区别	(13)
第七节 怎样学好 98 版五笔字型	(15)
第二章 98 王码基础知识	(17)
第一节 汉字的五种笔画	(19)
第二节 笔画间的三种关系	(20)
第三节 汉字结构的三个层次	(21)
第四节 汉字的三种字型	(21)
第五节 汉字编码的单位——码元	(23)
第六节 码元顺序与笔顺规范	(24)
第七节 三种编码表达式	(25)
第三章 认识王码键盘	(27)
第一节 标准键盘及其指法练习	(29)
第二节 什么是王码键盘	(36)
第三节 王码键盘的键位排列特点	(38)
第四节 王码键盘的码元分布规律	(40)
第五节 王码键盘的键面符号介绍	(41)
第六节 快速记住码元的区位号	(42)
第四章 码元分区助记及码元总表	(47)
第一节 第 1 区	(49)
第二节 第 2 区	(52)
第三节 第 3 区	(55)
第四节 第 4 区	(58)
第五节 第 5 区	(62)
第六节 码元总表	(65)

第五章 码元汉字输入法	(69)
第一节 键名汉字输入	(71)
第二节 成字码元输入	(72)
第三节 补码码元及其输入	(73)
第六章 合体字的输入	(77)
第一节 合体字的取码规则	(79)
第二节 多元字的编码规则	(82)
第三节 四元字的编码规则	(83)
第四节 二元字和三元字的编码规则	(83)
第五节 识别码的选取规则	(85)
第六节 98 王码汉字编码流程图	(87)
第七节 无拆分编码法及 1000 个常用字练习	(89)
第七章 简码重码和容错码	(119)
第一节 简码输入	(121)
第二节 重码	(124)
第三节 容错码	(125)
第四节 万能学习键 Z	(130)
第八章 词语输入	(133)
第一节 词语输入法的特点	(135)
第二节 词语输入规则	(136)
第三节 使用词库生成器	(137)
第九章 98 王码的使用	(161)
第一节 启动王码输入法	(163)
第二节 浏览输入法界面	(165)
第三节 使用王码输入法	(168)
第四节 设置高级输入特性	(173)
第十章 指法练习	(177)
第一节 键盘指法练习	(179)
第二节 范文输入练习	(181)
附录 A 国标汉字 98 王码编码表	(185)
附录 B 98 王码助记歌	(236)

第 一 章

汉字输入法概述

- 第一节 汉字的优缺点
 - 第二节 汉字输入法的分类
 - 第三节 86 版五笔字型
 - 第四节 98 版五笔字型的特点
 - 第五节 98 版五笔字型的新增功能
 - 第六节 98 版五笔字型与 86 版五笔字型的区别
 - 第七节 怎样学好 98 版五笔字型
-

第一节 汉字的优缺点

汉字是我们祖先发明的用于记录语言的一套符号体系,是中国 5000 年文化的产物和结晶。在世界众多的文字当中,汉字有以下几种优点:

(1) 汉字总数很多,但常用字很少

在汉语中,汉字大概有五六万个。总数虽多,但常用字很少。有人曾经作过统计,在报纸和杂志中,占普通文章 40% 的汉字只有 100 个,占 90% 的汉字只有 1000 个。由此可见,只要掌握 1000 个常用汉字就可以阅读一般文章。

(2) 构词灵活,适应性强

随着社会的发展和科学技术的进步,新的名称和术语不断涌现。由于汉语的词语是由单字拼成的,所以,无论公司注册、婴儿降生,还是新型化纤、药物的研制成功,都不会使汉字的数目增多。新词的出现只不过是几个汉字某种新的组合。

(3) 认字便能认词

对于汉字来说,认识汉字就可以认识词语。

(4) 汉字易于识别

汉字是平面结构。由于人类眼睛的视野总是一个面,所以汉字易于辨认。人们常说读中文书时能够“一目十行”,就是指汉字的阅读效率非常高。

(5) 汉字篇幅比较少

汉字书写简洁、表意深邃,占用的篇幅比较少,这是汉字的又一重要特点。

汉字的诸多优点,使汉字长盛不衰,沿用至今。博大精深的中国文化之所以与西方不同,主要也是与中国汉字有关。但是,随着科技的发展,特别是电脑的出现,汉字的种种优势又成了它与电脑发生矛盾的焦点。在电脑时代,人们迫切需要有一种方法能够将汉字输入到电脑中,然后利用电脑来处理中文信息。然而汉字的下列缺点使得人们很难直接找到一种好的解决方法:

(1) 汉字的字数太多,无法与 26 个键对应

英文只有 26 个字母,利用这 26 个字母就可以处理全部的书信、报刊、图书和资料,而汉字有五六万之多,其中常用的汉字也有好几千。这么多的汉字,仅靠电脑键盘上的 26 个键是无法直接输入的。

(2) 汉字的笔画太多,一个字要敲若干次键

组成汉字的笔画繁简差别太大,少的仅有一画,多的有几十画。由于大多数汉字都由多个笔画组成,因此,若采用一画敲一键的方法,汉字的输入速度将会慢

得惊人。

(3) 汉字的字型结构复杂,同形字多

在汉语中,当将相同的偏旁部首进行不同的排列时,可以组成不同的汉字。例如,“口”和“八”可组成汉字“叭”和“只”,“口”和“巴”可组成汉字“吧”和“邑”。另外,有些汉字仅有一点差别,但读音、意义各异。甚至有些汉字乍看起来很相似,但细看又不全相同。由于这些汉字的形状笔画相差无几,无疑给电脑造成了识别困难。

(4) 同音字太多,不易区分

现代汉语只有 417 个基本音节,若考虑标注 5 个声调(平声、去声、阴平、阳平和轻声),总共才有 1330 多个音节形式。面对 5 万多个汉字,必然存在大量的同音汉字。汉字的同音字如此之多,再加上方言,在采用拼音的方法输入时必然会产生各种歧义。

总之,在当今信息社会里,由于汉字字数繁多、字形复杂、难于机械地处理、与现有计算机的硬件环境不相适应等,使得汉字输入成为电脑中文化的瓶颈。汉字输入处理技术如不加以突破,我国的信息交换、情报处理、通讯系统、管理系统、军事指挥以及印刷排版等,便不能赶上时代的步伐,汉字自然也就会被淘汰。

第二节 汉字输入法的分类

为了使中华古老文化能够步入信息时代,加快我国信息产业的发展,我国广大科技人员对汉字输入法进行了广泛深入的研究。到目前为止,全世界研究发表的汉字输入方法据说已有 1000 种,其中经过上机试验的方案不下 100 种。所有的汉字输入法可划分为两大类:键盘输入法和非键盘输入法。这两类输入方法都有自己的特点,下面分别进行介绍。

一、非键盘输入法

非键盘输入法是指不通过键盘就可以输入汉字。这类方法主要包括以下几种:扫描识别输入法、手写识别输入法、语音识别输入法。

1. 扫描识别输入法

扫描识别输入法就是指用光电扫描仪把符合一定要求的汉字图形扫描到电脑,并自动与标准的字形进行比较,从而识别输入的汉字。扫描识别输入法的优点是不用按键,缺点是只对印刷体汉字有较好的效果。因为汉字字体过多过杂,

常常使识别的准确率受到影响。

扫描识别输入法对于某些用户有相当的吸引力,尤其对于图书馆和情报中心来说,使用扫描识别输入法来复制那些字体比较规范的现存文本,其效率是非常高的。因此,使用扫描识别输入法有着相当大的实用价值和广阔的应用前景。

但是,扫描识别输入法也有它自己的缺点:第一是识别设备价格高昂,如果不是制作专业资料,一般不采用这种方法;第二是不能识别书写不规范的手稿文章。对于手稿文章,一般还要采用键盘输入法;第三是扫描进入电脑总有一定的错误率,错的地方还要设法改正,因而更费时间。

2. 手写识别输入法

手写识别输入法是指在一个特制的书写板上,按照一定要求书写汉字,机器自动地将手写的字与电脑里存放的字进行比较,如果手写的字与电脑里存放的某个字匹配,就将该字作为输入的汉字。

手写识别输入法使用起来比较简单,用户只需在书写板上写字,手写输入软件就会自动将所写的字输入到计算机中。因此,一个人只要会写汉字,就可以使用该输入方法。

但是,由于技术上的原因,手写识别输入法只能识别书写比较规范的汉字,对于稍微潦草的字,它是很难识别的。因此,在使用手写识别输入法时,用户不但要把每个汉字都写得很规范,还不能写得太小。这样便带来另一个问题,那就是书写速度很慢。一般来说,每分钟能写 20~25 个规范字已经相当费劲了。所以,使用手写识别输入法输入汉字的速度不会很快,其用途范围是可想而知的。

3. 语音识别输入法

语音识别输入法是一种新的输入方法。在使用这种方法时,输入人员拿着话筒说话,计算机听人说话,并把听到的声音识别成汉字。

语音识别输入法是一种非常有前途的输入方法,研究该输入法的人和单位很多。由于语音识别输入法涉及到语言学、语音学、计算机科学和人工智能等多种学科,同时汉语同音字特别多,发音差异特别大,所以,研究语音识别输入法的难度非常大,目前该输入法的研究还只停留在试验室阶段。即使有部分产品已经推向市场,例如 IBM 公司推出的汉语语音识别系统,但由于汉语中同音同调词语非常多(例如“外交理论”和“外焦里嫩”、“南方”和“男方”等),所以错误率总是比较高。再加上受方言口音的限制以及要求周围环境不能有噪声等,致使这种输入方法难以广泛地推广使用。

二、键盘输入法

键盘输入法是指通过键盘输入汉字。由于汉字只有字根和笔画,没有字母,

因此,我们不能像输入英文那样,利用键盘上的 26 个字母直接向电脑输入汉字。

如果要通过键盘向电脑输入汉字,就必须将汉字或汉字的构成部分编成代号,也就是代码,然后将它们设计在电脑键盘的英文字母键上。这样,通过用电脑键盘输入这些代码,就可以间接地输入汉字。

为汉字编制代码的方法,叫作汉字编码。根据编码方法的不同,键盘输入法主要包括以下几种:电报码输入法、区位码输入法、拼音输入法、拼形输入法。

1. 电报码输入法

电报码是一个四位码,它是由一个丹麦人为清朝政府设计的。他从 5 万多个汉字中挑选 1 万个常用汉字,然后从 0000 开始,依次编制流水号,直到 9999 为止。一个汉字分配一个流水号,一个流水号就代表一个汉字。例如:

中:0022 国:0948

人:0086 民:3046

电报码的优点主要表现在两个方面:第一是唯一性强,由于每个码对应于一个汉字,所以无重码。第二是学习方法简单。用户只要记住每个汉字对应的码是多少,就可以输入汉字。

然而,1 万个汉字对应于 1 万个与字毫不相干的码,学习难度是可想而知的。因而在电脑时代,以电报码作为社会广泛采用的汉字电脑输入法,肯定是不可取的。

有人说,电报员不是译码速度很高吗?但是这些人必须经过较长时间的训练,而且,他们真正记熟的字也只不过一两千个,稍微冷僻的字,还必须查编码字典,不能做到见字识码。

2. 区位码输入法

区位码是指中华人民共和国国家信息交换编码。它用四位阿拉伯数字对国标汉字库中的汉字或其它字符编码。在四位阿拉伯数字中,前两位是区码,取值为 01~94,后两位是位码,取值也为 01~94,区码+位码就构成了区位码。在区位码中,1~9 区是各种符号和英语、日语、俄语、希腊语的大小写字母等,总共 202 个。16~87 区为汉字,共收入一、二级汉字 6763 个,其中 16~55 区是使用频度较高的汉字(一级汉字),按照汉语拼音的顺序排列。56~87 区是使用频度较低的汉字(二级汉字),按照汉语笔顺排列。

由于国标汉字库中,每个汉字都有一个区位码,每一个区位码都代表一个汉字,因此,通过区位码也可以输入汉字。在利用区位码输入汉字时,应先输入区码,然后输入位码。例如,

“五”的区码是 46,位码是 69

“笔”的区码是 17,位码是 42

“字”的区码是 55,位码是 54

“型”的区码是 48,位码是 45

所以,利用区位码输入法分别输入下列数字:4669 1742 5554 4845,就可以将“五笔字型”这四个汉字输入到计算机中。一般的中文系统中都配有区位码输入法。

区位码输入法的学习方法很简单,和电报码一样,只有一个字——“背”,用户只要记住每个汉字的区位号就可以将汉字输入到电脑中。但由于区位码与汉字本身毫不相关,所以记住汉字的区位号也不是一件容易的事。

3. 拼音输入法

拼音输入法(简称为音码)是根据汉语拼音输入汉字的一种输入方法。由于汉字读音简单,只用 25 个英文字母便可以拼出全部汉字的 400 多种读音,因此,只要学过汉语拼音,几乎不需要再学习,便可以在键盘上把汉字的音敲进去。拼音输入法已被广大非专业录入人员所采用。目前市场上常用的拼音输入法包括全拼输入法和智能拼音输入法。

虽然拼音输入法易学易用,但与其他输入方法相比,拼音输入法又存在下列非常明显的缺点:

(1)汉语的同音字很多。使用拼音输入汉字时,虽然键盘输入比较容易,但由于重码字很多,用户从屏幕选字非常难,不能实现盲打。

(2)中国幅员辽阔,民族众多,方言复杂,读音各异,许多省市和地区“Zh、Ch、Sh”与“Z、C、S”分不清。因此,用拼音输入常常也并不十分有效。

(3)每个人能够认识的汉字总是有限的。如果遇到不认识的字,就无法根据拼音输入了。

4. 拼形输入法

拼形输入法(简称为形码)的基本思想是:先从汉字中取出常见的字根,把它们依科学原理分配在电脑的键盘上,作为输入汉字的基本单位;当要输入汉字时,按照汉字的书写顺序依次按键盘上与字根对应的键,系统根据输入的字根组成的代码,检索出所要的字。拼形输入法是一种不考虑读音,完全依据笔画和字形特征来为汉字编制代码的方法。与其它输入法相比,拼形输入法具有下列特点:

(1)方言虽多,汉字的字形却是全国统一的,其字形和结构不受方言的影响。所以,拼形输入法比拼音输入法有更广泛的适用范围。

(2)汉字本质上是一种图形文字,所以,拼形输入法在输入汉字时,给人的印象是使用一块块字根,像搭积木那样把汉字组合起来,符合汉字的构字规律。

(3)汉字的字数多,一个字一个图形,笔画不同,形态各异。绝大多数情况下,按一组键只出来一个字,重码字少,不用挑选,用户可以实现盲打,其输入效率可以很高。

当然,与其他类型的方案相比,拼形方案也有它的缺点。其一是设计难度很大。要想设计一个优秀的拼形输入方案,就必须付出艰苦的劳动。其二是需要学习才能掌握。学员要想掌握这种方法,先要熟悉键盘上的字根,还要学会汉字的编码方法。

在拼形输入法中,应该说至今最有代表性、使用者最多的输入方法是 86 版五笔字型输入法。86 版五笔字型输入法是由作者于 1986 年发明的,它是完全依照汉字字形进行编码的一种汉字输入方法。由于作者姓王,所以,国内外用户都将 86 版五笔字型简称为“86 王码”。本书中所讲的 86 王码,也是指 86 版五笔字型。实际上,86 王码除了 86 版五笔字型以外,它还包括五笔画输入法。

第三节 86 版五笔字型

86 版五笔字型输入法是一种高效的汉字输入法,它是一种使用 25 个字母键,在键盘上用字根拼形组字的输入法。与其它输入方法相比,该输入法具有以下特点:

(1)它是适合专职和非专职人员共同使用的一种汉字输入法。五笔字型是一种按字形来设计的编码方法,编码的唯一性好,重码率非常低,平均每输入 10 000 个汉字,才有 1~2 个字需要挑选。因此,输入效率特别高。

(2)五笔字型用双手 10 个手指击键。经过标准的指法训练,每分钟向电脑输入 100 个汉字是很平常的。1992 年全国五笔字型大奖赛中,冠军的速度是生稿(错 1 罚 5)每分钟 243 个汉字。

(3)用五笔字型既能输入单字,也能输入词语。无论多么复杂的汉字,最多只按 4 下键,无论多长的词语也只按四下键。而且,字与词之间不需要任何换档或附加操作。因此,它既符合汉字构词灵活的特点,又能大幅度地提高输入速度。

(4)五笔字型集文字学、计算机科学、信息论以及人机工程学等多种学科的原理、方法于一体,字根组合的“相容性”使重码最少,键位字根安排的“规律性”使得方案相对易学,而“协调性”使得手指的击键负担趋于合理。

(5)五笔字型经过 10 年大规模社会实践的检验,已成为国内占主导地位的汉