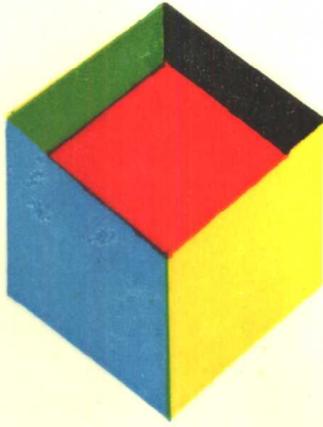


计算机应用培训教程

JISUANJI YINGYONG
PEIXUN JIAOCHENG



五笔字型输入法

WBX

- 上海市中等职业教育中心组编
- 戴卿 编著
- 华东理工大学出版社

计算机应用培训教程

五笔字型输入法

上海市中等职业教育中心 组编

戴 卿 编著

华东理工大学出版社

内 容 简 介

本书较为详尽地介绍了五笔字型输入法的编码规则和输入方法技巧,包括键盘操作、五笔字型输入法、五区字根介绍及练习、五笔字型编码方案综述、五键五笔划输入法等,并且循序渐进地配有较多的练习题和上机材料,供读者上机操作和检查学习情况。附录提供五笔字型词汇集。本书深入浅出、内容详尽、讲练结合,适合初学者作为学习五笔字型,也可作为学习计算机应用教程(初级)的配套参考书和上机指导书。

(沪)新登字 208 号

计算机应用培训教程

五笔字型输入法

上海市中等职业教育中心 组编

戴 卿 编著

华东理工大学出版社出版发行

上海市梅陇路 130 号

邮政编码 200237

新华书店上海发行所发行经销

上海展望印刷厂 印刷

开本 787×1092 1/16 印张 10.75 字数 270 千字

1994 年 12 月第 1 版 1994 年 12 月第 1 次印刷

印数 1-17000 册

ISBN 7-5628-0550-4/TP·73 定价 9.40 元

序

90年代上海紧缺人才培训工程,在上海市人民政府统一规划和部署下,已经拉开序幕,各级各类从业人员和广大市民,积极参加计算机应用能力的培训和考核;不少学校配合课程教学改革也把让学生掌握计算机基础知识和学会操作作为重要课程;不少学习计算机的培训班、短期班也应运而生,乃至计算机已经开始逐渐进入家庭。一个普及计算机知识,推广计算机应用的高潮正在上海兴起。

我校于80年代中期在国内率先创办的办公室自动化专业,就是以培养会操作应用计算机的人员为主的专业。不少教师在近十年的专业教学中,积累了许多经验。为了配合上海市市民计算机应用能力的培训和考核,更好地推动计算机知识的普及和应用,同时也促进我校办公室自动化专业更好地与市计算机人才培训工程接轨,我校专业组老师在参加计算机应用能力培训教学的基础上,邀请了职教界的有关同仁编写了这套《计算机应用培训教程》系列丛书。

这套丛书暂为四册:《磁盘操作系统DOS》、《五笔字型输入法》、《文字处理系统WPS》、《FOXBASE+数据库原理与应用》。这是一套通俗易懂及讲练结合的初学者入门教材,同时配备详尽的上机操作指导以及丰富的练习题,便于初学者自学或自己上机操作。它既可以作为职业技术学校的教学用书,又是上海计算机应用教程(初级、中级)很好的配套参考书和上机指导书。

由于计算机技术的迅速发展,近年来早已问世的有关计算机方面的书籍和刊物为我们编写这套系列丛书提供了不少的借鉴,但是限于我们的水平有限,错误和疏漏之处在所难免,恳请广大读者和专家同仁批评指正。

上海市中等职业教育中心

校长:张荫生

前 言

五笔字型输入法是河南省王永民研制的,已先后推出了四代版本。目前使用较为广泛的是1986年第四版,本书内容主要适合于这一版本。它的字库中收录有《中华人民共和国国家标准(GB2312—80)通讯用汉字字符集(基本集)》中的全部汉字和常用的符号。

汉字编码,就是采用一种科学可行的方法,为每一个方块汉字编取一个或一组唯一的代码,以便计算机辨认、接收和处理。众多的汉字编码方案,是利用汉字的形、音、义三属性进行编码。常用的编码方案大致分为拼音编码、整字编码和字形分解编码三大类。五笔字型编码方案是采用字形分解编码的。

五笔字型编码方案之所以在众多的编码方案中脱颖而出,是由于符合计算机对汉字输入码的基本要求,即击键数少(平均每字2.6键)、输入效率高、重码少(低于万分之二);又由于其规则简明、好学易记、记忆量少、规律性强、操作直观、击键与写字相仿、词组输入不用换挡等受到操作人员的欢迎。输入人员只要熟悉汉字的书写笔顺,花上不长的时间,就可以学会这种输入方法。

本书编写力求适合操作者自学的需要,精选典型的例子,较为详尽地介绍五笔字型输入法的编码规则和输入方法技巧,并循序渐进地配有较多的练习题和上机材料,供读者练习和检查学习情况之用。在附录内还提供了五笔字型词汇集。本书编撰过程中,得到本校OA教研组有关同志及王宪、马烈、钱丽萍等先生的帮助和指正,由张鹏、吕骏、黄允哲同志造字、排版,特此致谢。由于笔者水平有限,编写时间较为仓促,书中不足和错误之处在所难免,请读者不吝批评指正。

目 录

第一章 键盘操作	1
§ 1.1 键盘的构成	1
§ 1.2 键盘操作简介	3
第二章 五笔字型输入法	5
§ 2.1 汉字的字根	5
§ 2.2 汉字的拆分原则	8
§ 2.3 汉字的末笔字型交叉识别码.....	10
§ 2.4 五笔字型的编码与输入.....	15
第三章 五区字根介绍及练习	25
§ 3.1 横起类——第一区字根练习.....	26
§ 3.2 竖起类——第二区字根练习.....	29
§ 3.3 撇起类——第三区字根练习.....	31
§ 3.4 捺起类——第四区字根练习.....	34
§ 3.5 折起类——第五区字根练习.....	37
第四章 五笔字型编码方案综述	41
§ 4.1 五笔字型输入法总结.....	41
§ 4.2 重码和容错码.....	43
§ 4.3 选择式易学输入法.....	46
第五章 五键五笔划输入法	48
§ 5.1 五键五笔划键盘.....	48
§ 5.2 五键五笔划单字的编码输入.....	49
§ 5.3 词汇的编码和输入.....	50
第六章 上机实验指导	52
实验一 英文字母指法练习	52
实验二 常用汉字拆分练习	57
实验三 末笔字形交叉识别码输入练习	60
实验四 第一区字根练习	68
实验五 第二区字根练习	75
实验六 第三区字根练习	80
实验七 第四区字根练习	87
实验八 第五区字根练习	96
实验九 文章综合练习.....	102
附录一 五笔字型词汇集	107
附录二 常见非基本字根拆法	162

第一章 键盘操作

计算机汉字输入是通过键盘来完成的，键盘是计算机的一个重要组成部分，是计算机使用者向计算机输入数据和控制计算机的工具。键盘的按键标记大多数是相同的，其键位排列都采用标准(科蒂)键盘排列法。标准键盘中有 26 个英文字母键、10 个数字键和一些常用的符号键。这些按键的分布与标准的英文打字机上的分布是相同的。用五笔字型输入法向计算机输入汉字时，主要使用三排字母键和字母键下方长条状的空格键，有时还要用到个别数字键。

§ 1.1 键盘的构成

计算机键盘上有一条电缆引出线，用来同主机后面的 DIN 插座相接，这条电缆线中包括了四条线：+5V 电源、地线和两条双向信号线。电缆大约有 6 英尺长，并绕成了螺旋形，就像电话机听筒线一样。

键盘上有 101 个键，排列成四个主要部分：打字机键盘、功能键、光标控制键、数字键。

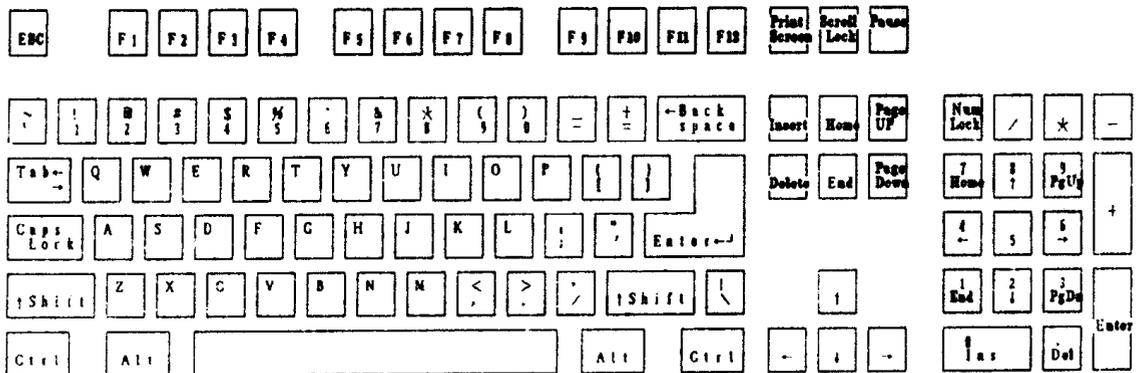


图 1-1 101 键加强型键盘

一、功能键

根据用户的需要，键盘上还特意设置了几个功能键 F1~F12，再加上 Esc 键，共 13 个键，它们的具体功能由操作系统或应用程序来定义。

二、打字键盘

它是键盘的主要部分,各类计算机的打字键盘部分设置基本一致,与通常的英文打字机相同,但计算机上只有一些附加功能。

1. 空格键

字母键下方长条状的空格键,按下此键时,会把一个空白格输入计算机,光标向右移动一个字符。

2. 大小写转换键 Shift

当处于非大写锁定状态时,按下转换键并同时按其他某个键,便可转换字母键、符号键和数字键,并使小写状态转为大写状态。

3. 控制键 Ctrl

这个键,总是与其他键同时使用以实现各种功能,这些功能是在操作系统中或其他应用程序中进行定义的。

4. Alt 键

它总是与其他键同时使用,以切换各种输入法。

5. 大写锁定键 CapsLock

这个键可将字母 A 到 Z 锁定为大写状态,而对其他键无影响。当处于大写锁定状态时,Shift 键会将大写转换为小写。再按下大写锁定键就退出大写锁定状态。

6. 退格键 BackSpace

用它可以删除当前光标位置左边的一个字符,并将光标左移一个位置。

7. 跳格键 Tab

这个键用来将光标右移到下一个跳格位置。同时按下“Shift”键和“Tab”键时,将把光标左移到前一个跳格位置。跳格位置总是被设为 8 个字符间隔,除非另作改变。

三、光标控制键

1. 光标移动键

按此键,光标将按箭头方向移动一格或一行。

2. Ctrl+Break 键

同时按下 Ctrl 和 Break 键被作为强行中止,常用来中止程序的执行。

3. 插入键 Ins

此键用来切换插入状态和替换状态,一个字符被插入后,光标右侧的所有字符向右移动一位。再次按 Ins 键则返回到替换方式。

4. 删除键 Del

它用来删除光标当前位置上的字符,当一个字符被删除后,光标右侧的所有字符将向左移动一个位置。

5. Home 键

按此键时光标移到屏幕的左上角。

6. End 键

按此键时光标移到本行中最后一个字符的右侧。

7. PgUp 键和 PgDn 键

常用来实现光标的快速移动。其具体的用法取决于操作系统或应用程序。

四、数字键

按下 NumLock 键,机器处于数字锁定状态。这时键的功能为输入数字和运算符号。

§ 1.2 键盘操作简介

初学者在练习五笔字型输入法之前,应适当化些时间进行指法训练,这不仅有利于提高输入速度、减少差错率,而且为尽早形成条件反射、实现盲打创造良好条件。

一、正确的姿势

初学键盘输入时,首先必须注意的是击键姿势。如果初学时姿势不当,就不能做到准确快速地输入,也容易疲劳。

1. 身体应保持笔直,稍偏于键盘右方。
2. 应将全身重量置于椅子上,桌椅要放置到便于手指操作的高度,两脚平放。
3. 两肘轻轻贴于腋边,手指轻放于规定的字母键上,手腕平直。人与键盘的距离可移动椅子或键盘的位置来调节,以调节到人能保持正确的击键姿势为好。
4. 显示器宜放在键盘的正后方,放输入原稿前,先将键盘右移 5CM,再将原稿紧靠键盘左侧放置,以便阅读。

二、正确的键入指法

1. 基准键及其手指的对应关系

(1) 基准键位,位于键盘的第二行,共有八个字键,分别为“A、S、D、F、J、K、L、;”如图 1-2 中所示。左、右手的各手指必须放在图中所规定的字键上。

(2) 图 1-2 两组基准键之间的键,即“G、H”键不属于基准键。

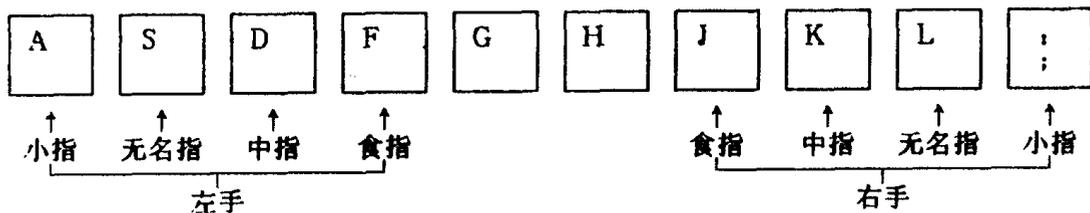


图 1-2 基准键位示意图

2. 字键的击法

(1) 手腕要平直,手臂要保持静止,全部动作仅限于手指部分(上身其他部位不得接触工作台或键盘)。

(2) 手指要保持弯曲,稍微拱起,指尖后的第一关节微成弧形,分别轻轻地放在字键的

中央。

(3) 输入时,手抬起,只有要击键的手指才可伸出击键。击毕要立即缩回,不可用摩擦手法,也不可停留在已击的字键上(除8个基准键外)。

(4) 输入过程中,要用相同的节拍轻轻地击字键,不可用力过猛。

3. 空格的击法

右手从基准键上迅速垂直上抬1-2CM,大拇指横着向下一击并立即回归,每击一次输入一个空格。

4. 换行的击法

需要换行时,起右手伸小指击一次 RETURN 键。击后,右手立即退回原基准键位,在手回归的过程中,小指提前弯曲,以免把“;”键带入。

三、键盘的指法分区

在基准键位的基础上,对于其他字母、数字、符号都采用与八个基准键的键位相对应的位置来记忆。例如,用原击 D 键的左手中指击 E 键,用原击 L 键的右手无名指击 O 键等。

键盘的指法区如图 1-3 所示,凡两斜线范围内的字键,都必须由规定的手的同一手指管理,这样,既便于操作,又便于记忆。

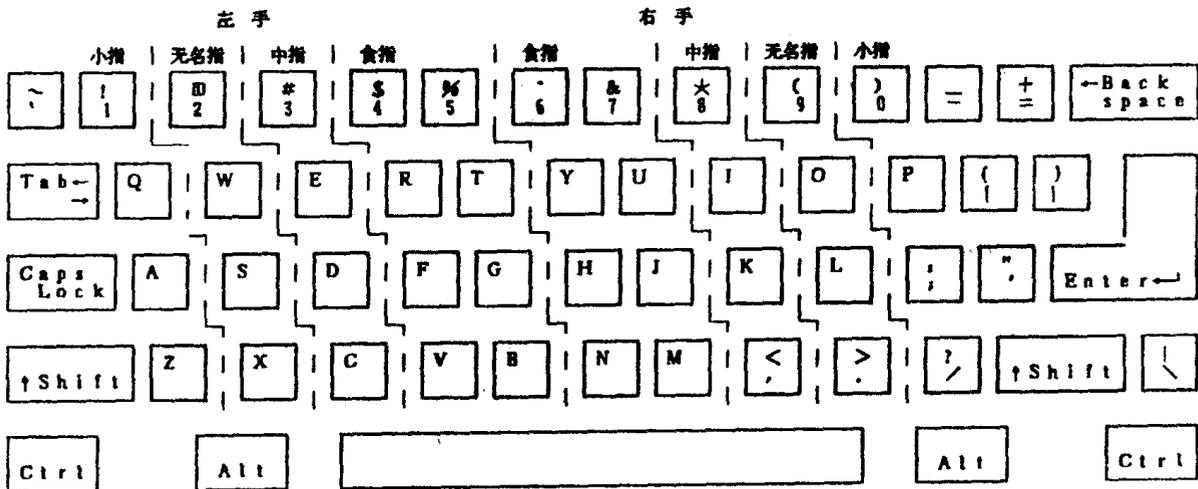


图 1-3 指法分区图

第二章 五笔字型输入法

标准的英文键盘上一般都有 26 个英文字母键。这些键分上、中、下三排,双手十指各司其职,上下移动不用跨排。英文输入速度之所以很高,不能不说与这种小巧的键盘设计有关。然其根源在于英文的基本单位是单词,其组成是 26 个字母,真所谓万变不离其宗。作为方块字的汉字,曾因其笔划形态多变、字数冗多、字形复杂,而现有计算机的字母键又非常少,使汉字输入成举世关注的难事。

于是,国内外许多研究者煞费十数年之苦心,终于找出了结构各自独立的汉字中存在着类似英文单词中的“字母”——字根。它们按照一定规律分布于英文字母键上,这样键盘就成了“字根键盘”,输入汉字时,只要按一下相关字根的字母键,字根就会自动组合成你要的汉字。例如:输入“胡”字,分别键入字根“古”、“月”的字母键“D”、“E”,再按一下条形空格键,字根“古”和“月”就会自动组合成汉字“胡”,并显示在屏幕上,操作人员在输入汉字时,仿佛在玩儿时的拼字游戏,这种编码方案就是所谓的拼形编码(或称字形分解编码)。

五笔字型编码方案就是遵从人们的书写汉字的习惯顺序,将字根按一定位置关系拼合起来构成汉字的。

§ 2.1 汉字的字根

五笔字型输入法认为,汉字由字根组成,字根由笔划构成,笔划、字根、整字(单字)是构成汉字的三个层次。即:由若干笔划复合、连接、交叉,形成相对不变的结构组成字根;再将字根按一定的位置关系拼合起来就构成了汉字(整字)。

一、汉字的五种笔划

在书写汉字时,不间断地一次连续写成的一个线条叫做汉字的笔划。笔划是构成汉字的第一层次。据此,对成千上万的汉字及其字根加以分析,就只考虑笔划的运笔方向,而不计其轻重长短,从而得出一个正确的结论:汉字的笔划归于横、竖、撇、捺、折五种。为了便于记忆和应用,根据其使用频率的高低,依次用 1、2、3、4、5 作为代号,如表 2-1 所示。

上述对笔划的分类及其结果也与人们的习惯看法是基本一致的。对此,要注意如下三点:

1. 上述分类只考虑笔划的走向而不论其长短轻重。
2. 笔划“/”(可称之为“提”)与“一”同属横笔;笔划“丨”(常称为“竖左钩”)与“|”同属竖笔;撇笔只有“丿”一种;点笔“丶”与“丶”同属捺笔。除上述笔划外,汉字的其他笔划均有转折,统统属于折笔,用“乙”来表示。

3. 熟记五种笔划,对学习五笔字型输入法是很重要的。既要记住“1、2、3、4、5”分别代表“横、竖、撇、捺、折”,更要熟记提、竖右钩、点三种笔划的代号。

表 2-1 汉字的基本笔划

代 号	笔划名称	笔划走向	笔划及其变形
1	横	左→右	— /
2	竖	上→下	
3	撇	右上→左下	丿
4	捺	左上→右下	丶
5	折	带转折	乙 ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏

二、字根及单体结构

由若干笔划交叉连接而形成的相对不变的结构,就叫做字根。字根是有形有义的,在多数情况下还有称谓的构字基本单位,这些基本单位,经过拼形组合,就产生出为数众多的汉字。

但是,字根并不像汉字那样,有公认的标准和一定的数量。哪些结构算字根,哪些结构不算字根,历来没有严格的界限。不同的研究者,不同的应用目的,其筛选的标准和选定的数量差异很大。

前面说过,我们可以把字根看成是组成汉字(整字)的“字母”,一定数量的字根,就可以像 26 个字母构成任何一个英文单词那样,经过一定的组合,构成字库中 6763 个汉字中的任何一个。这就要求我们所选出的“字母”(字根)必须具有尽可能多的代表性,即同一个字根可以在较多的汉字中找得到,是这些汉字的相同部分。

例如,汉字“夺、达、天、央、因、奇、庆……”有相同的部分“大”;汉字“睡、睛、咱、厢、盲……”有相同的部分“目”;汉字“认、谁、信、誉”有相同的部分“言”。这些相同的部分“大”、“目”、“言”,就是字根。

根据字根必须具备代表性的原则,优选出了 130 种字根(也有称之为基本字根),如计各字根变形在内,则有 200 个字根,只有这 130 种字根才有“权力”组字或词。

在汉字拼形编码的研究中,所选字根多数是一些传统的汉字部首。但由于某种需要,有时也选用一些不是部首的笔划结构作为字根,甚至于硬造出一些“字根”来。相反,有些虽然是传统部首,但鉴于它们组字很少或者其本身便于拆分成几个部分,在制定编码方案时也未必选取。例如,在五笔字型方案中,“足、鱼、礻、礻、羊、皮、麻、户”都不选作字根。这些落选的汉字结构(也有称之为非基本字根)称之为单体结构,可以通过“单体结构拆分原则”,拆分成彼此交连套迭的几个基本字根。如:“足”拆成字根“口”与“乚”,“鱼”拆成字根“𩺰”与“一”,“礻”拆成字根“礻”与“丶”等等。从中可以看出,一切汉字都是由“字根”组成的。字根是构成汉字的最基本的单位,是构字的灵魂,笔划只是起辅助作用。

三、字根在键盘上的分布

我们尽管把字根比喻为构成汉字的“字母”,但由于汉字结构多种多样,汉字数量成千上万,字根不可能“精简”到跟英文字母一样多,而众多的字根如何更科学、更合理地分布在英

文字母键上,便于操作人员尽快掌握和快速输入是很有学问的。

五笔字型方案选用了25个(A~Y)英文字母键,Z键在选择易学输入法时用。我们把这25个字母键分成5个区,每区又分成5个位,这样共得到:11~15,21~25,……,51~55共25个键位。结合英文打字键盘,字母排列如图2-1所示。

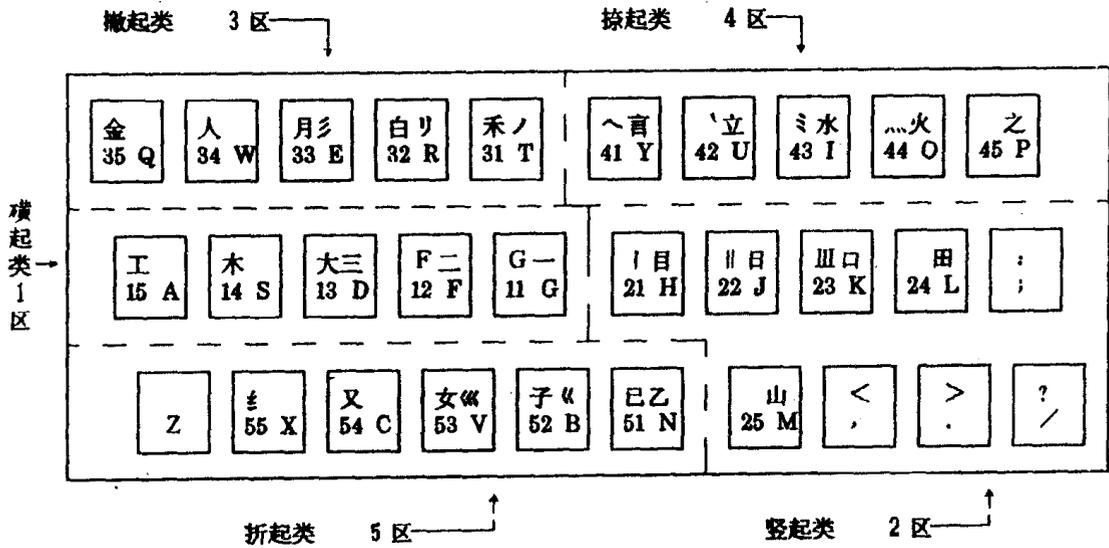


图 2-1 键盘上的字母排列

从图 2-1 中可以看出,任何一个英文字母键都可以用一个 2 位数字来表示,如“G”为“11”,表示第 1 区第 1 位;“J”为“22”,表示第 2 区第 2 位;“V”为“53”,表示第 5 区第 3 位。这个 2 位数(区位号)称为数字码,对应的英文字母称为字母码,二者统称为键位代码(或代码)。

接着,我们把 130 种字根根据首笔笔划代号也分成五类,分别分布于五区英文字母键上。如首笔为横的字根位于第 1 区,首笔为竖的字根位于第 2 区,其余类推。位号由字根的次笔决定,如字根“王”,首笔为横(1),次笔也为横(1),则位于字母键 G(11)上;字根“言”,首笔为点(4),次笔为横(1),代码为 Y(41)。于是每个字根都可以“对号入座”了,这称为字根分布的“区位法则”。

每个英文字母键上有一组字根,从中选出一个有代表性的字根,作为该键的键名。这样,可以编成一首“键名谱”:

- 1 区 横起类:王土大木工
- 2 区 竖起类:目日口田山
- 3 区 撇起类:禾白月人金
- 4 区 捺起类:言立水火之
- 5 区 折起类:巳子女又纟

尽快掌握每个字根的区位,对初学者来说可以起到事半功倍的作用。设计者力求将 130 种字根分布有规律可循,不杂乱,尽量使同一键上的字根在形、音、义方面能产生所需的联想,便于记忆和迅速熟练掌握。见图 3-1 五笔字型键盘字根总图。

1. 区位法则

以字根的首笔代号确定区号；字根的次笔代号确定位号。

2. 复合散笔字根法则

同一种笔划分散结合而形成的字根称为复合散笔字根，如表 2-2 所示，共有 17 个。这些字根都由同种笔划组成，笔划的代号就是该字根的区号，笔划数等于该字根的位号。如：点的代号为 4，那么数字码 41 代表的字根是一个点“丶”；数字码 42 代表的字根是两个点“冫”；数字码 43 代表字根“彡”；44 则代表“灬”。

3. 形似联想法则

与键名或其他字根形态相近，构成联想。

如：“王”字键上有“王”、“五”；“土”字键上有“土”、“干”等字根与之形似。

五笔字型键盘字根总图还给出每个字母键上一组字根的助记词，利用这些助记特性和口诀，不难记忆各字根的键位。记忆终究是要靠自己用脑，别人无法代替，自己不妨仔细观察分析一下字根总表和键盘图，根据上述三条规律寻找、记住可助记的各类字根的特征。

表 2-2 字根键盘的区位号、键名和笔划字根

位号 区号	1	2	3	4	5
横 1	G 王 一	F 土 二	D 大 三	S 木	A 工
竖 2	H 目 丨	J 日 丨	K 口 川	L 田 川	M 山
撇 3	T 禾 丿	R 白 丿	E 月 夕	W 人	Q 金
捺 4	Y 言 丶	U 立 丶	I 水 彡	O 火 灬	P 之
折 5	N 乙 乙	B 子 丨	V 女 川	C 又	X 纟

§ 2.2 汉字的拆分原则

汉字是由基本字根拼合而成的，包括许多没有资格入选为字根的单体结构，这些单体结构并不一定都是汉字，如“巾、夫”，但也全部是由字根与字根或单笔划与字根，按照一定关系组成的。在输入汉字时，必须将整字（包括单体结构）拆成字根，按笔顺先后输入，即按下这些字根的字母键来完成，因此，将整字拆分成字根是输入汉字的一个很重要的环节。

一、汉字的结构分析

字根在组成汉字时，按照它们之间的位置关系可以分成四种类型：

1. 单根结构

字根本身就单独成为一个汉字，如汉字“口、木、王、山、田、马、寸”等，五笔字型键盘字根总表上所有的汉字，都属于这一结构，称之为单根字。包括键名汉字和成字字根汉字两部分。

2. 连笔结构

这类结构的汉字由一个单笔划(一、丨、丿、丶、乙)和另一个字根连接组成的。例如,汉字“自”由单笔划“丿”下连字根“目”所组成;汉字“亏”由字根“二”下连单笔划“乙”组成;汉字“且”由字根“月”下连单笔划“一”组成;汉字“主”由单笔划“丶”下连字根“王”组成。在这种结构中,其中的单笔划与字根相连,可连前(如“主、自”字),也可连后(如“亏、且”字)。因此,一切单笔划与字根相连构成汉字时,不管书写时单笔划是否与字根相连,一律视为相连(如“主”字)。

3. 散根结构

指构成汉字的字根之间可以保持一定的距离。如汉字“吕、汉、足、表、因、识、照、这”等。这里要引起注意的是,“连笔结构”与“散根结构”的区别;两个字根相连,如果有一个是单笔划字根(其实单笔划也应该属于字根,是一种特殊的字根),或者两个都是单笔划字根(如“入”字),则属连笔结构;如果两个都是非单笔划字根,则属散根结构(如“足、表”字)。

4. 交叉结构

是指由两个或几个字根交叉套迭之后构成的汉字。如“农”字是由字根“一”、“艹”交叉构成,汉字“申”由字根“日”、“丨”交叉构成,汉字“里”由字根“日”、“土”交叉构成,汉字“夷”由字根“一”、“弓”、“人”交叉构成。

二、汉字的拆分原则

在上述四种汉字结构中,单根结构只由一个字根组成,无需拆分,其输入法有专门公式。前面的一些实例中,汉字(包括一些单体结构)的拆分较直观,如汉字“胡”拆成字根“古”和“月”,汉字“足”拆成字根“口”和“止”,汉字“鱼”拆成字根“鱼”和“一”,单体结构“夨”拆成字根“冫”和“人”,“禛”拆成字根“禛”和“丶”。而绝大多数汉字可能有几种拆分形式,如汉字“天”可拆成字根“二”和字根“人”相交,也可拆成字根“一”和字根“大”相连。当然,正确的拆分方式只有一个,因此,我们要遵循下列的汉字拆分原则进行拆分(也有的称之为单体结构拆分原则)。

1. 按笔顺拆分

在汉字拆分中,应按习惯书写顺序拆分成几个部分。如“片”字拆分成的字根依次是“丿”、“丨”、“一”、“乙”,汉字“长”应拆成字根“丿”、“七”、“丶”,汉字“义”应拆成字根“丶”和“乂”。

2. 能散不连

在保证按书写顺序拆分的前提下,如果一个汉字可以按照“散根结构”来拆分,就不要按照“连笔结构”拆分。例如,汉字“午”应拆成字根“勹”与字根“十”上下“散”的关系,而不是单笔划“丿”与字根“干”之间“连”的关系。

3. 能连不交

同样按照书写顺序拆分汉字时,如果一个汉字可以按照“连笔结构”来拆分,就不要按照“交叉结构”去拆分。例如,汉字“于”可以按照“连”的关系拆成单笔划“一”和字根“十”,又可按照“交”的关系拆成字根“二”和“丨”,按照“能连不交”的原则,前一种拆法是正确的。同样,汉字“生”可以按“连”的关系拆成字根“丿”和“土”,就不要按“交”的关系拆成字根“勹”和“土”。

4. 取大优先

在保证满足上述三条原则的前提下,假如有几种可能的拆分方法可提供选择的话,优先考虑每次都拆出最大的字根的那种拆分方法。这包含两层意思:一是在拆分过程中,保证前面的字根尽量大(笔划数多)。例如,汉字“尺”有两种拆法,即拆成字根“尸”和“丶”及拆成字根“㇇”和“人”,比较二种拆法的第一字根的笔划数哪个多,便可得出结论,前一种拆分方法是正确的。同样,汉字“夹”应拆成字根“一”、“亠”和“人”而不能拆成字根“一”、“丶”和“大”。二是在几种可能的拆法中,以拆分出的字根数最少的那种拆法优先。例如,汉字“京”可拆成两个字根“亠”和“小”,也可拆成三个字根“亠”、“口”和“小”,根据取大优先原则,字根数少的那种优先,自然拆成两个字根的拆法是正确的。单体结构“缶”也应拆分成字根“亠”和“山”,而不能拆分成字根“一”、“十”和“凵”。

总之,拆分应当兼顾上述几条原则的要求,在遵循按笔顺拆分的前提下,一般先满足“取大优先”中的字根数最少的原则。而拆分成的字根数相同的几种可能的拆法中,则应先考虑“散”比“连”优先,“连”比“交”优先,再满足“取大优先”中的前面的字根尽量大的原则。例如,汉字“矢”有二种拆分可能性,即拆分成字根“一”和“大”及拆分成字根“亠”和“人”,因为都拆成二个字根,就应考虑汉字结构上的“散”、“连”、“交”之间的先后关系,前一种拆法结构上是“散”,比后一种“交”优先,则前一种拆法正确,而不应该考虑“取大优先”中的前面字根尽量大的原则,否则会错误地选择后一种拆法。常见非基本字根拆分见附录2。

§ 2.3 汉字的末笔字型交叉识别码

我们知道,用五笔字型编码方案输入汉字时,除了字根总图上的所有汉字,即所谓的单根字外,所有的汉字都是以输入字根代码来完成的。每个汉字最多击键四次,即由四个字根组成的四根字依次输入,超过四个字根的汉字称为多根字,依次击前三个字根和最后一个字根的代码,即1、2、3、末。对少于四个字根的二根字、三根字,其多出来的一键、二键可以用空格键来代替完成。由于每个字母键上的一组字根有2~6种之多,并且这些二根、三根汉字字根数较少,即“信息量不足”,在输入时极易出现重码。如输入二根字“见”时,由于字根“儿”与“乂”同处于字母键“Q”上,当击二键“MQ”(字根“冂”在“M”键上)后,用空格来结束,屏幕上只出现汉字“风”。若击“MQZ”(“Z”为万能学习键,在后面有关章节会专门介绍),再用空格键来结束,此时屏幕上提示行出现五个汉字供选择,击数字键“5”即可输入汉字“见”。此法虽简单,但由于输入时要从提示行中选字,大大影响输入速度,因而在输入二根字、三根字时有必要补充“信息”,以便于分别输入这些出现重码的汉字,即所谓“离散重码”。

一、汉字的三种字型

在前一节我们分析了汉字结构,根据字根的位置关系,把汉字分成单根结构、散根结构、连笔结构和交叉结构。除了单根结构外,从整体上来看,散根结构中的绝大多数汉字明显由几个部分组成,称之为“合体字”,两个部分合并在一起的汉字称为双合字,三个部分合并在一起的称为三合字。当然也有由四个部分或超过四个部分组成的汉字,其信息已够丰富,就不属于考虑范围。根据合体字组成部分之间的位置关系,左右排列的叫左右型汉字,上下排

列的叫上下型汉字。而连笔结构和交叉结构的汉字,无法分成几个组成部分,称为“独体字”,加上散根结构中一部分无法从结构上分成左右型或上下型的汉字,统称为杂合型汉字。这样,根据构成汉字的各字根之间的位置关系,我们又把成千上万的方块汉字分为三种类型(字型):左右型、上下型和杂合型。同样,也按照它们拥有汉字的字数多少以1至3编号,如表2-3所示。

表 2-3 汉字的三种字型

字型代号	字型	图示	字例
1	左右	口 口 田 田	叭 渐 经 新
2	上下	日 日 田 田	只 莫 花 碧
3	杂合	囧 囧 田 田 田 田 田 田	固 凶 冈 斗 这 式 应 乘 天

1. 一型——左右型汉字

在左右型汉字中,包括两种情况:

(1) 在双合字中,两个部分分列左右,整个汉字中有着明显的界线。如:叭、胡、枫、汉、咽、拥等。

汉字枫和咽的右边也由两个字根构成,虽然这两个字根之间是外内型关系(杂合型),但整个汉字却属于左右型。

(2) 在三合字中,整字的三个部分从左到右并列;或者单独占据一边的部分与另外两个部分呈左右排列。如:渐、经、新,都应属于左右型。

2. 二型——上下型汉字

上下型汉字也包括两种情况:

(1) 双合字中,两个部分分列上下,其间有一定距离。如:只、字、节、看等。

(2) 三合字中,三个部分上下排列,或者单占一层的部分与另外两部分作上下排列。如:莫、花、碧、想等。

3. 三型——杂合型(外内型和单体型汉字)

三型是指组成整字的各部分之间没有简单明确的左右或上下型关系者。如:固、凶、冈、这、斗、头、飞、乘、本、重、天、且、册、成、主等。

汉字的图型特征,是每一个有文化的中国人从小学起就熟知的。这里可以用来作为识别汉字的一个重要依据。如字根“口”、“八”上下排列为汉字“只”,左右排列即为汉字“叭”等。因此,我们还可以把三种字型叫做字根的三种排列方式。在我们向计算机输入汉字时,除了按正确的笔划顺序键入组成汉字的字根外,有时还有必要告诉机器那些键入的字根是按什么方式排列的,也就是补充键入一个字型信息,即所谓的“末笔字型交叉识别”代码,简称识别码。

在学习识别码之前,应注意以下几点:

字根单独成字,即单根字无需判断字型;

属散根结构的汉字,才有可能分成左右型或上下型等;

属连笔结构或交叉结构的汉字,其字型一律为杂合型。

此外,其它不分左右、上下的汉字(汉字结构为散、连、交等的混合类型或外内型、单体