

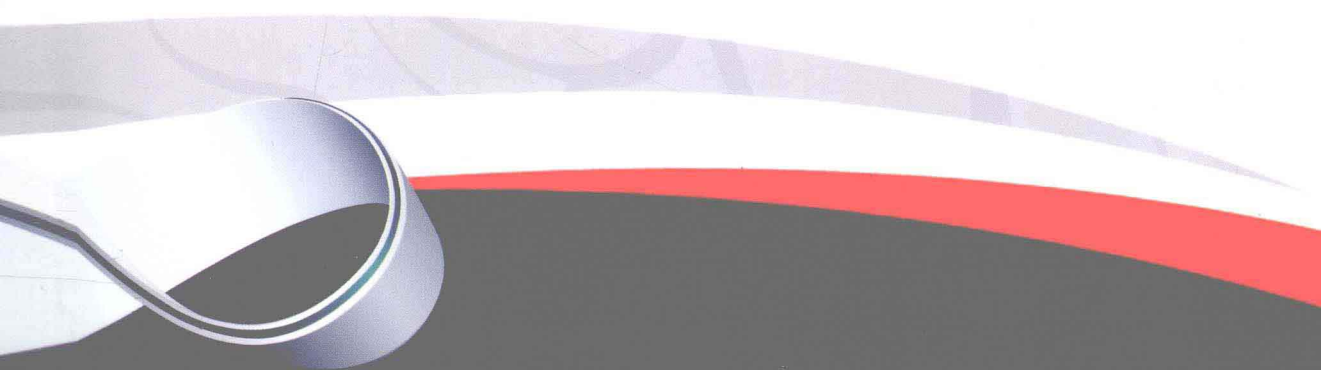


计算机专业职业教育实训系列教材

文字录入与编辑 实训教程

WENZI LURU YU BIANJI
SHIXUN JIAOCHENG

宋福英 主编



机械工业出版社
CHINA MACHINE PRESS



配 电 子 教 案

计算机专业职业教育实训系列教材

文字录入与编辑实训教程

主 编 宋福英

副主编 王永斌

参 编 石立峰 吴启翔 郭栩栩

主 审 王小平



机械工业出版社

本书是各职业学校计算机应用、电子商务、电算会计等专业的办公自动化方向的实践课教材,实践性和应用性很强。它以实用为出发点,帮助学生熟练掌握中英文盲打技术和技能,熟练掌握五笔字型输入法,了解其他常用汉字输入法,为学生学习其他的计算机课程及以后的工作打好基础,培养学生的上机操作能力和吃苦耐劳的精神。

本书共分6章,分别讲述了文字录入基础知识、汉字处理基础知识、五笔字型输入法、其他常用汉字输入法、文字编辑基础知识和方正书刊排版系统。学好这门课程,学生将能尽快了解计算机的基本知识,掌握计算机的基本操作技术,成为具有计算机文化知识的人才。本书还配有电子课件,读者可在机械工业出版社 www.cmpedu.com 上以教师身份免费注册下载,或联系编辑(010-88379194)索取。

本书适合各职业学校计算机应用等相关专业学生,同时也适合于各类培训班学员及计算机录入者。

图书在版编目(CIP)数据

文字录入与编辑实训教程/宋福英主编. —北京:机械工业出版社, 2011

计算机专业职业教育实训系列教材

ISBN 978-7-111-33065-3

I. ①文… II. ①宋… III. ①文字处理—职业教育—教材 IV. ①TP391.1

中国版本图书馆CIP数据核字(2011)第008102号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:梁伟 责任编辑:蔡岩

封面设计:鞠杨 责任印制:乔宇

三河市宏达印刷有限公司印刷

2011年3月第1版第1次印刷

184mm×260mm·7.5印张·176千字

0001—3000册

标准书号:ISBN 978-7-111-33065-3

定价:19.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010) 88361066

销售一部:(010) 68326294

销售二部:(010) 88379649

读者服务部:(010) 68993821

门户网:<http://www.cmpbook.com>

教材网:<http://www.cmpedu.com>

封面无防伪标均为盗版

前言

本书是各职业学校计算机应用、电子商务、电算会计等专业的办公自动化方向的实践课教材,实践性和应用性很强。它以实用为出发点,帮助学生熟练掌握中英文盲打技术和技能,熟练掌握五笔字形输入法,了解其他常用汉字输入法,为学生学习其他的计算机课程及以后的工作打好基础,培养学生的上机操作能力和吃苦耐劳的精神。通过学习本教材学生能懂得文字录入与编辑方面的基本知识,对文字录入与编辑有较完整、较系统的了解,培养学生熟练地进行文字录入和处理工作的能力,通过训练达到中级文字录入处理员的水平。使用本书时,注意以下几个方面:

(一) 本书特色

本书基于“快速掌握、即查即用、学以致用”的原则,根据日常工作和学习的需要取材谋篇,以应用为目的,并配以大量实例。具有以下特点:

1. 内容上注重“实用为先”

本书在内容上注重“实用为先”,精选最需要的知识,介绍最实用的操作技巧和最典型的应用案例。真正将计算机使用者的技巧和心得完完全全地传授给读者,教会您生活和工作当中真正能用到的东西。

2. 方法上注重“活学活用”

本书在方法上注重“活学活用”,根据用户实际使用的需要,以应用为目的,将软件的功能完全发掘给读者,教会读者更多、更好的应用方法。

3. 讲解上注重“丰富有趣”

本书在文字上注重“丰富有趣”,风趣幽默的语言搭配生动有趣的实例,采用全程图解的方式,细致地进行分步讲解,读者翻看时会感到兴趣盎然,回味无穷。同时还提供了大量“提示”、“注意”、“技巧”的精彩点滴,让读者在学习过程中随时认真思考,对初、中级用户在使用计算机过程中随时进行贴心的技术指导,迅速将“新手”打造成为“高手”。

(二) 本课程与其他课程的联系

文字录入作为计算机专业的基础课程,应在学生学习了《计算机文化基础》、《计算机常用工具软件》课程后开设,不需要太多其他预备知识。学好本课程,可以为学生再进一步学习其他专业课程打下良好的基础。

(三) 本课程的实训实践环节

1. 实践课的总体要求

(1) 课前预习:学生在进行上机前必须认真准备,明确上机目的,掌握上机操作的步骤和方法。

(2) 认真操作:上机操作能力是在上机过程中逐步培养和提高的,学生应在教师指导下,严格按照上机作业和要求,独立完成作业。力求在每一次上机课中都能正确掌握每节课的内容。

(3) 写好上机报告：学生上机课结束以后应写上机报告，它是对实践过程和结果的分析 and 总结。要求学生能正确地分析并总结过程，从而加深学生对实践内容所涉及原理的理解。

2. 考核办法

本课程为考试课，考核成绩=试卷（20%）+上机（60%）+平时成绩（20%）。其中试卷部分主要考核学生对课程中基本概念和重要理论的掌握情况；上机部分主要考核学生经过一学期的学习之后的动手实践能力；平时成绩主要是根据平时的提问及出勤等来确定的。

3. 作业、操作、测试的重点内容和方法

(1) 作业：本课程是一门实践性极强的课程，作业主要是上机作业和上机报告两种。上机报告主要是学生在上机过程中的操作步骤和注意事项，以统一印发的报告纸形式上交；上机作业主要是学生在上机过程中制作的实例，以电子作品的形式上交或展示。

(2) 操作：本课程操作内容比较多，基本上是在机房以上机练习的形式进行。一般操作课堂上要布置一个具体的任务留给学生进行练习以便当堂检查考核。

(3) 测试：和作业相对应，本课程测试分为上机考试和卷面考试两部分。其中，上机考试在机房进行，占总成绩的 60%；卷面考试以试卷的形式进行，占总成绩的 20%。

(四) 本课程教学建议

训练由易到难，逐步展开。练习时要循序渐进、持之以恒、反复训练，宜采用集中练习法，如集中时间反复练习输入同一段文字。教学中应注重培养学生盲打能力。建议使用金山打字作为学生的学习软件。

(1) 本课程涉及的知识点较多，是实践性很强的一门课程，除了在课堂上讲授一定的基本知识和原理外，要充分利用多媒体技术或计算机室进行直观教学，讲授操作方法，演示操作步骤和操作效果。采用启发、讲解、现场指导等多种方法进行课堂教学，以提高教学效果。

(2) 根据该课程实践性、操作性强的特点，应尽量采取讲练结合的教学方式，突出上机操作训练。

(3) 适应计算机软、硬件技术迅速发展的需要，各任课教师应及时总结课程教学经验，补充并更新教学内容，改进教学方法，以提高课程教学质量。

(4) 课时分配如下：

章节内容	讲授学时	上机学时
导学 第1章 文字录入基础知识	2	2
第2章 汉字处理基础知识	2	2
第3章 五笔字型输入法	14	14
第4章 其他常用汉字输入法	2	2
第5章 文字编辑基础知识	2	2
第6章 方正书刊排版系统	14	14
合 计	36	36

本书由宋福英主编，王小平任主审。参加编写的还有王永斌、石立峰、吴启祥、郭栩栩。由于作者水平有限，书中难免有不妥和错误之处，恳请各位专家、读者批评指正。

编 者

目 录

前言	
导学	1
第1章 文字录入基础知识	3
1.1 键盘键位及其功能	3
1.2 键盘指法	5
第2章 汉字处理基础知识	9
2.1 字符与汉字编码	9
2.2 汉字录入方法介绍	12
2.3 计算机汉字处理流程	14
2.4 总结	14
第3章 五笔字型输入法	15
3.1 汉字的基本结构	15
3.2 汉字的构成	16
3.3 五笔字型的字根键盘	17
3.4 汉字的拆分	20
3.5 五笔字型的汉字编码	22
3.6 简码、重码和容错码	25
3.7 词语的输入	26
3.8 王码	27
3.9 智能五笔	29
3.10 万能五笔	32
3.11 掌握五笔字型输入法的技巧	33
3.12 总结	34
第4章 其他常用汉字输入法	35
4.1 区位法汉字录入	35
4.2 智能ABC输入法	37
4.3 微软拼音输入法	45
4.4 搜狗拼音输入法	46
4.5 二笔输入法	56
第5章 文字编辑基础知识	60
5.1 基本常识	60

5.2 排版工艺常识.....	67
5.3 校对知识.....	70
5.4 电子排版工艺.....	75
第6章 方正书刊排版系统.....	78
6.1 初步认识方正书刊 10.0 排版系统.....	78
6.2 字符效果.....	80
6.3 段落效果.....	87
6.4 整篇排版效果.....	91
6.5 表格制作.....	98
6.6 数学排版.....	102
6.7 化学排版效果.....	106
参考文献.....	112

导 学

本书将教会大家两大技能：文字录入和文字排版。

文字录入

在计算机被广泛应用于社会各行各业的今天，大多数计算机所做的工作是进行信息处理。要进行信息处理，首先要做的工作是把收集的数据录入到计算机中，然后按照企业的要求对信息进行编辑和排版。因此，文字录入和编辑排版的重要性和必要性表现得越来越突出。提高文字录入和排版速度，已成为很多人的追求。近几年操作能力强、能迅速进入职业角色的中等职业学校的学生越来越受到企业的青睐，还出现了“供不应求”的情况。

学习文字录入，也就是要选择一种适合自己的输入法，通过科学规范的方法学习并练习后达到：

初级：掌握至少一种汉字输入方法，能实现盲打，速度为 30 字/分；

中级：熟练掌握至少一种汉字输入方法，能实现盲打，速度 60 字/分；

高级：熟练掌握至少一种汉字输入方法，能实现盲打，速度为 90 字/分；

能手级：熟练掌握至少一种汉字输入方法，能实现盲打，速度为 120 字/分；

速录员：对语音信息的采集速度是每分钟不低于 140 字；

速录师：每分钟不低于 180 字；

高级速录师：每分钟不低于 220 字。

速录师：让文字“音速”飞行

速录师是指运用速录机设备，从事语音信息实时采集并生成电子文本的人员。通俗地说，就是运用专业的速录机，边听边打字，做到“语音落、记录完、文稿成”，说话者讲完即可把讲话稿打印出来。

速录师就好像在弹钢琴一样，手指在键盘上不断地跳跃。

陈红学习速录始于 2003 年，那年她中专毕业，在学校接受了速录的培训，培训后她的速录速度达到 160 字/分钟，在同一批培训的学生当中，名列三甲。大部分同学毕业后去了法院当速记员。而陈红坚持自己的选择，最终走进企业，并于 2007 年创办了自己的公司。

速录师平均工资 3000 元以上。

据了解，目前广州只有几家速录公司。近来，陈红希望招聘一名能听懂、打字速度在 200 字/分钟以上的速录师。“但到现在还是找不到符合要求的人才。”

其实速录师的待遇一点儿也不低，一般以 200 元/小时或 1200 元/天计价。但速录师的经历至少几个月的磨砺，单独完成合格的速录后，收入才会慢慢增加。在广州，一般的速录师平均月工资可以达到 4000 元以上。

如今，速录已开始广泛应用于行业或企业举行的各种重要会议、峰会、法庭审讯、记者采访、直播等领域。

文字排版

在古代，要印刷书本，需要人工将字按文件原稿排好，所以有“排版”一说。而到了现代，排版实际上就是指掌握一些文件编辑软件的法，并按照常见的行文规则和规范在计算机中对原稿进行排版。

“三分长相，七分打扮”

俗话说：“三分长相，七分打扮”，这是用来形容人的外表要靠精心装饰来改善。节假日中的人们，经过梳妆打扮以后，穿上漂亮合体的衣服，总能给人焕然一新的感觉。

排版对于出版物，就是那件漂亮衣服，就是那些化妆品。在人们的日常生活中，都会接触到排版物，不论是每天必看的报纸、阅读的书、查看的资料、浏览的杂志和刊物，还是街头巷尾铺天盖地的广告，都或多或少涉及排版。

而评价出版物的好坏，不能只看内容是否精彩，还要看排版是否优美。只有经过专业排版润色的出版物，才能给人赏心悦目、心旷神怡的感觉，并油然而生一种美感，才是一件真正的好作品。

不管是排报纸、杂志、书还是平面广告，都要处理文字、图形和图像等素材，并把这些素材安排在一个页面内，这个版面制作过程主要由排版软件来完成。

在印前制作领域中，本书将要讲解的排版软件——北大方正的飞腾被誉为“排版天才”。它大量应用于报社和出版社，受到了很多好评。这款排版软件以其强大的功能占据了广泛的市场。

学习完本书后，学生可具有使用计算机从事秘书、文书、信息资料与档案管理、文字处理、图书出版、报刊、印刷广告等专门工作所必需的中英文文字输入技能、制作中英文文稿版式的基本知识和基本技能，制作出赏心悦目的文章，为学生走向企业、事业等文秘工作岗位打下坚实的基础。

第1章 文字录入基础知识

掌握规范的录入指法，已经成为操作计算机必备的前提，如果您现在对文字录入还一窍不通或不甚了解，那么请跟随我们，让我们一起与它来一次亲密接触吧！

本章将教给你

- 对键盘键位的认识
- 对键盘键位功能的了解
- 规范科学的指法

学完本章后你应该

- 能熟记键盘各键位及功能
- 了解录入操作的基本原则
- 掌握正确的录入操作姿势
- 精通科学规范的指法

键盘是实现人与计算机之间对话的主要途径，各种数据信息都可以通过键盘输入到计算机中。为提高录入速度，必须进行专门的指法训练。在进行指法训练前，首先要了解键盘上的键位分布情况，并按照正确的指法进行操作，以便逐步实现盲打，提高数据录入效率。在这一章里，我们将介绍键盘键位，对键盘的操作有一个深入的认识。通过学习，了解键盘上各键的功能，认清各键位置，进行指法练习。

1.1 键盘键位及其功能

要想学好打字就必须熟练操作键盘，要想熟练操作键盘就必须掌握键盘的布局和指法分区。

1.1.1 键盘键位

键盘大体上可分为标准键盘、非标准键盘和专用键盘三种。其中非标准键盘和专用键盘主要用于专用设备和特殊设备中，很少与计算机配套使用。

早期的计算机标准键盘为 83 键，分为功能键区、标准字符键盘区和数字键盘区三个区域。随着计算机的不断发展，键盘也向多功能型和方便型发展：出现了标准的 101 键键盘；或为配合 Windows 95、Windows 98 使用的专用键盘；为减少打字疲劳而设计的人体工学键盘等。这些键盘基本上都可分为五个区域，即功能键区、标准字符键盘区、控制键区、数字键盘区和状态指示区。101 键或 104 键的通用扩展键盘键位总图，如图 1-1 所示。

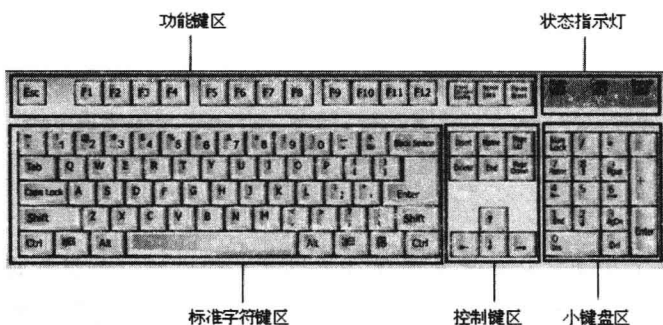


图 1-1 键盘示意图

在键盘操作中，不同的键区和不同的键有着不同的使用特点。

1. 标准字符区

标准字符区位于键盘中央偏左的大片区域，是使用键盘的主要区域，如图 1-2 所示。



图 1-2 标准字符区

2. 功能键区

功能键区位于键盘的最上面一行，如图 1-3 所示。

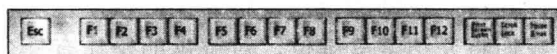


图 1-3 功能键区

3. 编辑键区

编辑键主要是指在整个屏幕范围内，进行光标的移动操作和有关的编辑操作等，如图 1-4 所示。

4. 小键盘区（数字/全屏幕操作键区）

小键盘区位于键盘的右侧，又叫数字键区，主要用于快速输入数字，如图 1-5 所示。

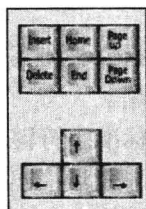


图 1-4 编辑键区

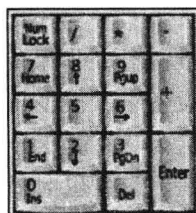


图 1-5 小键盘区

1.1.2 键盘操作

速录成功的关键是对计算机键盘的熟练使用。

操作时计算机应放置在专用的桌子上，高度为适中。座位最好是可以调节高度的转椅。在初学键盘操作时，必须十分注意打字姿势。正确的姿势可以有效、高速地向计算机输入有关信息，提高打字速度，正确姿势的具体要求是：

1) 打字者正坐在椅子上，全身重心平稳，腰杆挺直，背部与椅子成直角，两腿自然平放在桌子下。

2) 椅子高度要适当，一般都使用可以方便调节座位高低的转椅，眼睛距显示器的距离为 30 cm 左右。

3) 两肩放松，两肘悬空，手腕平直，手指自然弯曲，轻放于规定的键位上，手指要放在键位的中央。身体与电脑桌保持一定的距离。

4) 原稿应放在键盘左侧，故可将键盘稍稍右移。

力求做到“盲打”，即视线要投注在显示器上，不可常看键盘，以免视线往返，增加眼睛的疲劳。同时要经常注意用眼卫生，如眼睛与屏幕要保持适中的距离，眼睛要平视、放松，屏幕显示要调节到恰到好处的亮度，连续操作时间不宜过长，要注意眼睛休息。

键盘输入时的正确坐姿如图 1-6 所示。

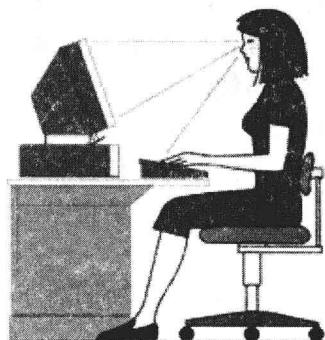


图 1-6 键盘输入时的正确坐姿

1.2 键盘指法

“工欲善其事，必先利其器”。在学习任何一种录入法之前都必须掌握且熟练科学规范的键盘指法。如图 1-7 所示，十个手指应各司其职、各尽其责、齐心协力、通力合作。

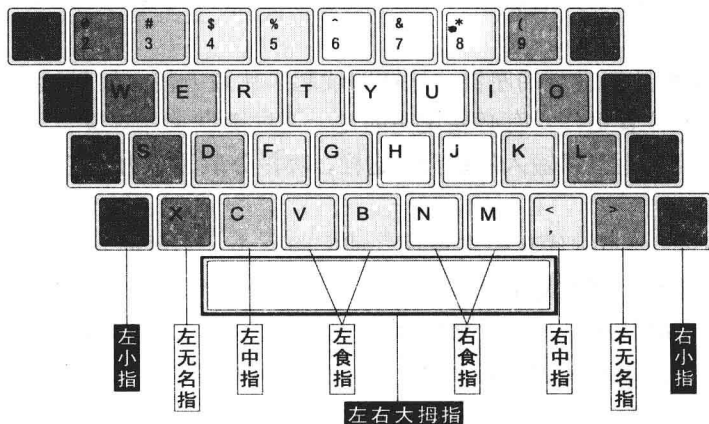


图 1-7 指法示意图

1.2.1 指法训练

操作键盘时,应将手指放在键盘的8个基准键位上。基准键位是指主键盘区第二排字母键中的【A】、【S】、【D】、【F】、【J】、【K】、【L】和【;】。

要想掌握规范的指法,在练习中就要不断强化下面的打字技巧和要领。

1) 手指的键位分工是指把键盘上的键位合理地分配给10个手指。除拇指外,其余8个手指各有一定的活动范围,每个手指负责各自区域内字符的输入。

2) 当使用键盘输入字符或数字时,每个手指都只能在自己的区域内活动,不能敲击其他键。在击键时应注意:不能长时间按住一个键不放,当击键结束后,各个手指应立即退回相应的基准键位上。

3) 打字者在操作时必须集中精力。击键要果断、迅速,击键后要立即弹起,手指退回原位。击键的力量也要均匀。

4) 准备打字时,双手拇指放在空格键,其余8个手指垂放在各自的基准键上。

5) 击键时手指自然弯曲,手指向上略微拱起,手指的第一关节呈微弧型,手指放在按键中央。

6) 非击键的手仍自然地停留在基准键上,两手同时击键时除外。

7) 击键完毕,手指应立即回到基准键上。

1.2.2 指法训练的难点

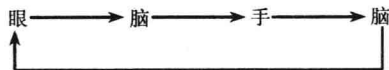
在进行文字录入训练时,除了强调正确的姿势外,还必须强调技术训练和心理训练相结合,这是指法训练的关键也是难点。

1) 训练时,在做好准备工作时,必须专心练习,不能受其他事情的干扰。阅读原稿的速度,以手能跟得上为宜。击键过程中,要注意体会击打不同键位上的按键时手指动作的差别和手指的键感,尽力记住准确的击键动作。

2) 在基础练习阶段,要把准确性放在第一位。基础训练阶段特别强调姿势的正确;逐个字符地记忆键位;训练手指的动作;练习眼、脑、手的协调等也很重要,这是以后提高速度的基础。

3) 一般可将原稿放在键盘的左侧,这样阅读起来比较方便。如果配置有专用的原稿架,放在键盘后面的中间位置更佳。有必要说明的是,通过键盘录入时最忌讳边看原稿边看计算机键盘或看显示器屏幕上已经输入的信息,这样,容易分散注意力,造成多打、漏打或串组、串行等差错。

4) 阅读原稿时,要将视线集中在单词(或字组)上;击键时,视线要集中到第一个单字,击完第一个单字后,视线移到下一个单字(空格并归到前一单词或字组)。在眼看与手击之间,脑是桥梁。眼睛所看见的反映到脑子里,脑指挥手的动作以完成击键;手之键感返回通知大脑动作完成,眼睛又去收集信息,这一过程可归纳为:



直到输入结束,该循环才结束。为了集中精力、加深记忆和保证动作协调,初学者还可以采用默念的办法,即把眼睛看到的内容不出声地边念边击键,使键位印象清晰,落指无误。

5) 随着所要练习的字符的增加,训练难度也不断加大。学习者一定要按照键盘输入的操作要领,认真练习,多练多记才会熟能生巧,逐步达到得心应手的效果。而心不在焉、马马虎虎地练习,是决不会达到预期效果的。

1.2.3 基本键的训练

1. 基本键 A、S、D、F、J、K、L、; 的训练

在键盘中,A、S、D、F和J、K、L、;这8个键称为基准键位,如图1-8所示。

网上有许多录入训练小软件如:金山打字,圆圆打字高手、键盘精灵——指法练习系统、打字训练之指法练习等,可下载使用。

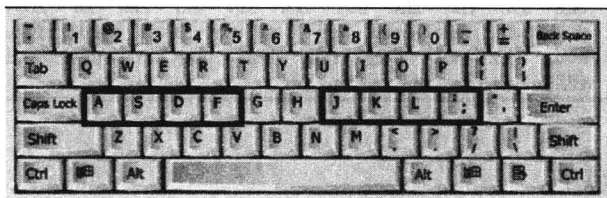


图 1-8 基本键位

反复练习输入下面的内容:

```
fff fff fff jjj jjj jjj ddd ddd ddd kkk kkk kkk sss sss sss lll
lll lll aaa aaa aaa ;;; ;;; ;;; sdf sad dsa dsa jlk jlk klj lk ass
add all all dad das sdf ask ask fall sak dlk lad lad lss las sls
ls! sls ad sad fla fasd kjlk lkds asda sdfd klj; kljj aksd al;sf
aasd lj;s asfd ;lkj jkl; asfd ;lkj skdl a;l afdk lasd j!a asd;k ;sdf
```

2. G、H 键的训练

反复练习输入下面的内容:

```
hghg hggj fgfg fgfg fghg gjgj gjgj jhjh jh fgf fgf fgf jhj jhj
jhj fgh fgh fgh jhg jhg had had glad glad sdf ga hjks lkjha
lkjhg kjhgf daslk hjk sa hj kla kjlas lkjgh kjhfg jghds fhjg hgj fjh
fhjg fhjg jfhj gfhg fhg hgjg gjj gff ffg ggj fj fjjj jf hg fj
```

3. E、R、T、Y、U、I 键的训练

反复练习输入下面的内容:

```
au eu ue el iaa aai ill ill eiei aiai fiei adfis afis aidi her it
it is fed fed fed ill ill ill lid lid lid ask ask ask sail sail
sail; kill kill kill; jail jail jail file file file; jade jade jade; desk desk
desk desk at a future date; the judge is just; at least a year use
the regular rate rest a little after that date a safe ride free the
```

4. Q、W、O、P 键的训练

反复练习输入下面的内容:

wwp ppw oow opq icop ow qo pq qc ytp qp qqp tpp tqq
pqp qqp qwp qqw wwp ppw oow opq oww poow opw owo
owe owe we op wee qwp ppe qc ytp qp qqp

5. V、B、N、M 键的训练

反复练习输入下面的内容：

bab fvb fvv fbb vbv bvb mmb mmb vbb bvv fmm jmm bnd mnu
inn bvn bni vnp vuvn bnn nbn ffb fin jnn mmb nub mnn vbn
nvb mnn base need best nine bear able rain mine mean abut
maid turn dind land bile train under balck until bring brush gabit

6. Z、X、C 键的训练

反复练习输入下面的内容：

zzx cex dkz aaz eze xke zks zek zee zkl zka xll zuu xss xxz
cde xft abed efg hijk lmn opq rst uvw xyz abed efg hijk lmn
opq rst uvw xyz abc defg hijk lmn opq rst uvw xyz in the
end we will conserve only zxx cex

1.2.4 范围键的训练

1. 符号键的训练

键盘上还有一些字符，如+、-、*、/、(、)、#、\$、!、%、&等，这些字符的输入也必须按照它们各自的分区，用相应的手指按规则击键输入。只要熟悉了字母键和符号的击键方法，字符的输入也就很简单了。

反复练习输入下面的内容

@@ @ @ ## ## ## !!! \$\$\$ % % % & & &
*** ((())) = = + + + + , , , ? ? ? ?
< < < > > > ^ !!! ~~~~~ ||| + _ _ ++ HAU ABCD
EFGH IJL MN ZXD afafH ioqQ WM lss kZUW abedEFH

2. 数字键的训练

数字位于键盘上标准键盘区的最上端，如图 1-9 所示。



图 1-9 数字键区

反复练习输入下面的内容：

1111 1111 2222 2222 3333 3333 4444 4444 5555 5555 6666 6666
7777 7777 8888 8888 9999 9999 0000 0000 0123456789 1238 244
1475 1831 1712 1514 1049 8750 6654 3479 0098 1257 8701 391
921; 238 1475 1831 1712 1514 1049 8750 alal s2s2 d3d3 f4f4
g5g5 h6h6 j7j7 k8k8 l9l9 ;0;0 111a 222s 333d 444f 555g 6h k8

第2章 汉字处理基础知识

了解和掌握汉字处理基础知识，将对提高汉字的录入和处理能力以很大的帮助。在这一章，我们将学习汉字在计算机中不同于其他字符的一些处理方法和基本知识！

🔗 本章将教给你

- 👉 汉字的特点及汉字编码
- 👉 汉字输入法简介
- 👉 汉字处理的基本知识和流程

🔗 学完本章后你应该

- 👉 了解不同字符的编码
- 👉 知道常用的不同类型的汉字输入法
- 👉 掌握几种不同的汉字编码
- 👉 掌握汉字的处理流程

2.1 字符与汉字编码

2.1.1 字符编码

大家知道，对于计算机来说，它所能识别和处理的内容只能是二进制数。所以，我们需要把处理的所有数据都进行数字化，也就是把汉字、英文的大小写字母、数字、各种符号、图像和声音等都用二进制数来表示。这样，计算机才能识别和处理这些对象。

为了在应用中有一个统一的标准，所有的英文字母、数字和各种符号都按一定规则用二进制编码来表示。汉字也不例外，同样有汉字编码字符集。

目前计算机中用得最广泛的字符集及其编码，是由美国国家标准局(ANSI)制定的 ASCII 码 (American Standard Code for Information Interchange, 美国标准信息交换码)，它已被国际标准化组织 (ISO) 定为国际标准，称为 ISO 646 标准。适用于所有拉丁文字字母，ASCII 码有 7 位码和 8 位码两种形式。

因为 1 位二进制数可以表示两种状态：0、1；而 2 位二进制数可以表示 4 种状态：00、01、10、11；依此类推，7 位二进制数可以表示 128 种状态，每种状态都唯一地编为一个 7 位的二进制码，对应一个字符（或控制码），这些码可以排列成 1 个十进制序号 0~127。所以，7 位 ASCII 码是用 7 位二进制数进行编码的，可以表示 128 个字符，详细编码对照见表 2-1。

表 2-1 常见字符的 ASCII 码对照表

$b_6b_5b_4$ $b_3b_2b_1b_0$	010	011	100	101	110	111
0000	SP	0	@	P	.	p
0001	!	1	A	Q	a	q
0010	"	2	B	R	b	r
0011	#	3	C	S	c	s
0100	\$	4	D	T	d	t
0101	%	5	E	U	e	u
0110	&	6	F	V	f	v
0111	,	7	G	W	g	w
1000	(	8	H	X	h	x
1001	)	9	I	Y	i	y
1010	*	:	J	Z	j	z
1011	+	;	K	[	k	{
1100	,	<	L	\	l	
1101	-	=	M	]	m	}
1110	.	>	N	^	n	~
1111	/	?	O	_	o	Del

当计算机使用 7 位 ASCII 码时, 每个字节只占用了 7 位, 还有一位没用, 也就是 8 位中的最高位, 其值恒为 0, 因此我们称 7 位 ASCII 码为基本 ASCII 码。当最高位的值为 1 时, ASCII 码被扩充了, 又可以多表示 128 种字符, 称为扩充 ASCII 码。通常扩充 ASCII 码被各个国家用来作为自己国家语言文字的代码。

2.1.2 汉字编码字符集 (GB 2312—80)

同 ASCII 码提供了字母和数字及其他符号的编码一样, 我们的汉字也有自己的编码。

GB 2312 或 GB 2312—80 是一个简体中文字符集的中国国家标准, 全称为《信息交换用汉字编码字符集·基本集》, 又称为 GB 0, 由中国国家标准总局发布, 它是一个简化字的编码规范。1981 年 5 月 1 日实施。GB 2312 编码通行于中国大陆; 新加坡等地也采用此编码。中国大陆几乎所有的中文系统和国际化的软件都支持 GB 2312。

GB 2312 标准共收录 6763 个汉字, 其中一级汉字 3755 个, 二级汉字 3008 个; 同时, GB 2312 收录了包括拉丁字母、希腊字母、日文平假名及片假名字母、俄语西里尔字母在内的 682 个全角字符。在这个字符集中每个字符用两个字节表示。整个字符集分成 94 个区, 每区有 94 个位。每个区位上只有一个字符, 因此可用所在的区和位来对汉字进行编码, 称为区位码。这个码是唯一的, 不会有重码字。把换算成十六进制的区位码加上 2020H, 就得到国标码。国标码加上 8080H, 就得到常用的计算机机内码。

GB 2312 的出现, 基本满足了汉字的计算机处理需要, 它所收录的汉字已经覆盖中国内