

ust
中等专业学校
电子信息类

规划教材

TP39-43
L826

计算机操作应用基础

卢传友 编著



A0933032

西安电子科技大学出版社

1999

内 容 简 介

本教材内容分三部分,第一部分为计算机基础知识,包括基础知识、计算机系统概述;第二部分为计算机的基本操作,包括计算机键盘和录入操作、操作系统(DOS 6.22、Windows 95、UCDOS 5.0)、汉字输入方法、文字编辑软件 WPS 97;第三部分为计算机新技术,包括计算机病毒及其防治、计算机网络基础和 Internet 的基本操作。

本教材可以作为中等专业学校、中等职业学校计算机基础课程的教学用书,也可作为计算机短期培训班的教材或计算机工作者的工具书。

图书在版编目(CIP)数据

计算机操作应用基础/卢传友编著.

—西安:西安电子科技大学出版社,1999.11

中等专业学校电子信息类规划教材

ISBN 7-5606-0776-4

I. 计… II. 卢… III. 计算机应用-基础知识-专业学校-教材 N. TP39

中国版本图书馆 CIP 数据核字(1999)第 40444 号

责任编辑 杨宗周

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)8227828 邮 编 710071

http://www.xduph.com E-mail: xdupfb@pub.xaonline.com

经 销 新华书店

印 刷 渭南市邮电印刷厂

版 次 1999年11月第1版 1999年11月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 14.75

字 数 338千字

印 数 1~4000册

定 价 14.50元

ISBN 7-5606-0776-4/TP·0399

*** 如有印制问题可调换 ***

前 言

本教材系按电子工业部工科电子类专业教材 1996~2000 年编审出版规划,由中等专业学校计算机专业教学指导委员会推荐出版。责编委周岳山。

本教材由山东省电子工业学校卢传友担任主编,山东省信息工程学校刘同顺担任主审。

本课程的参考教学时数为 80 学时,其中理论教学 52 学时,实验教学 28 学时。另外还需安排 2 周实习。本教材主要内容为计算机操作应用。分为三部分,第一部分为计算机基础知识,包括计算机基础知识、计算机系统概述;第二部分为计算机的基本操作,包括计算机键盘和录入操作、操作系统(DOS 6.22、Windows 95、UCDOS 5.0)、汉字输入方法、文字编辑软件 WPS 97;第三部分为计算机新技术,包括计算机病毒及其防治、计算机网络基础和 Internet 的基本操作。

本课程是计算机及应用专业的专业基础课,既要把必须掌握的理论知识传授给学生,也要把计算机中常用的一些基本术语和知识介绍给学生,使学生感到计算机不难学,又使学生感到有更多的知识需要学习,激发学习计算机的兴趣。本教材本着由简到繁、由易到难、循序渐进、面向应用的原则介绍计算机的基本知识和基本操作。在讲述各种命令时,只介绍命令的常用参数,一般不介绍命令的完整格式,力求通俗易懂,并附有大量实例帮助读者理解。每章后附有上机指导和习题。

本课程是实践性较强的计算机专业基础课,在学习过程中必须注意理论知识与实践技能相结合,在熟悉计算机基础知识的同时,重点掌握计算机的基本操作。

本教材由哇碧霞编写第 4、5、7 章,卢传友编写 1、2、3、6、8、9 章并统稿。参加审阅工作的还有戚琦、郑三、刘益红等同志,他们都对本书提出了许多宝贵意见,在此表示诚挚的感谢。由于编者水平有限,书中难免还存在一些缺点和错误,殷切希望广大读者批评指正。

编 者

1999 年 9 月

第 1 章 计算机基础知识

计算机的出现是近代重大科学成就之一。它的出现,有力地推动了其它学科的发展。计算机已经成为我们工作和生活中必不可少的工具。我们学习计算机的使用,首先应当了解计算机的基础知识。本章简单介绍计算机的发展、特点及用途,计算机中数的表示和编码。

1.1 计算机的发展和分类

1.1.1 计算机的发展

自从 1946 年第一台计算机问世以来,计算机技术在短短的 50 多年里连续进行了几次重大的技术革命,每一次都留下了鲜明的标志,因此人们自然地用第一代、第二代、……来区分计算机的发展阶段。

在划分年代的依据上,有许多不同的划分原则。如按照计算机采用的电子器件来划分,结合具有里程碑意义的典型计算机来划分,考虑计算机系统的全面技术来划分等。通常人们简单地按照第一种划分原则把计算机的发展划分为四个时期,即四代。

第一个发展时期是 1946~1957 年,主要是以电子管计算机为特征;第二个发展时期是 1958~1964 年,主要是以晶体管计算机为特征;第三个发展时期是 1965~1972 年,主要是以小规模集成电路计算机为特征;第四个发展时期是 1972 年至今,主要是以大规模集成电路和超大规模集成电路计算机为特征。当前超大规模集成电路(10 万个晶体管/片)比早期(2000 个晶体管/片)的集成度高若干个数量级。

目前计算机正朝着微型、网络、多媒体、智能和巨型等方向发展。

1.1.2 计算机的分类

从原理上可分为电子模拟计算机和电子数字计算机。

从用途上可分为专用计算机和通用计算机。专用计算机是在专门场合使用的计算机,如仪器计算机,安装在固定的仪器上,作为仪器的一部分。通用计算机是人们常见的计算机,可以用于各种目的。通常所说的计算机是指电子通用数字计算机,简称计算机。

根据计算机规模大小和功能强弱,又可分为巨型计算机、大型计算机、中型计算机、小型计算机和微型计算机五种。人们常用的计算机是微型计算机,微型计算机简称为微型机或微机。

根据外型大小可把微机分为两类:台式机和便携机。台式机又称桌上型微机,从名字

可以看出它是摆在桌面上工作的，我们通常使用的微机大多是这种微机。便携机又称膝上型微机，它可以装在公文包里，外出旅行可随身携带，使用时摆在膝上即可工作。更小的则称为笔记本电脑。

根据微机规模大小和功能强弱，又可分为多板机(普通微机)、单板机和单片机。

1.2 计算机的特点及用途

1.2.1 计算机的主要特点

1. 运算速度快

随着半导体技术的发展和计算机系统的改进，计算机的运算速度已从最初的几千次每秒发展到今天的几十万次、几百万次，甚至几亿次、几十亿次每秒。

2. 精确度高

一般计算机有十几位有效数字(从理论上说还可以更高，但这使机器太复杂，或使运算速度降低，因此不必要无限制地增加有效位数)。

3. 具有记忆特性

计算机能把大量数据、程序存入存储器，也能把经过加工处理或运算的结果保存在存储器中，而且在需要使用这些信息时，可以准确、快速地把它们取出来。

4. 有逻辑判断能力

计算机可以进行各种逻辑判断，并根据判断的结果自动决定以后执行的指令。有了这种能力，再加上存储器可以存储各种数据和程序，使计算机能够快速地完成各种过程的自动控制和各种数据处理工作。

5. 可靠性高

随着大规模集成电路和超大规模集成电路技术的发展，计算机的可靠性也大大地提高了，计算机连续无故障运行时间可达几个月甚至几年。

6. 通用性强

一台计算机能适应多种用途，各行各业都能通过使用计算机来达到自己的目的。

概括起来说，计算机是一种以高速进行操作、具有存储能力、由程序控制操作过程的电子装置。

从第一代计算机到第四代计算机，计算机的体系结构没有发生根本的变化，即这些计算机都是由运算器、控制器、存储器、输入和输出(I/O)设备组成的冯·诺依曼体系结构。冯·诺依曼除了指出计算机硬件的基本结构之外，还指出了计算机必须使用二进制数。在程序运行之前，要先将指令和数据存入存储器中，然后，机器自动到存储器中取出指令和数据执行，即程序存储控制系统。

正在研制的第五代计算机将是一种非冯·诺依曼体系结构的计算机。它采取全新的工作原理和体系结构，更接近于人们思考问题的方式，即“推理”方式。第五代计算机不仅采用的技术与以前不同，而且在概念和功能方面也不同于前四代计算机。这种新型的计算机

称之为“知识信息处理系统”。其功能从目前单纯的数据处理发展到知识的智能处理，这种新型计算机具有人工智能的功能。因此，未来的第五代计算机将对计算机科学技术的发展做出突破性贡献。

1.2.2 计算机的用途

计算机以其卓越的性能和旺盛的生命力，在科学技术、国民经济及生产、生活等各个方面都得到了广泛的应用，并取得了明显的社会效益和经济效益。计算机已成为信息社会的强大支柱。计算机的应用水平已成为衡量一个国家科学技术水平的重要依据。

根据计算机的应用特点，可以归纳为科学计算、数据处理、实时控制、计算机辅助系统、办公自动化和人工智能等几大类。

1. 科学计算

用计算机来解决科学研究和工程设计等方面的数学计算问题，称为科学计算，或称数值计算。随着科学技术的不断发展，需要解决的数学问题越来越复杂，计算的量也越来越大，速度和精度也要求不断提高，如果仍用手工计算或传统的计算工具计算，已经不能满足科学技术发展的需要。

例如，天气预报工作中，有大量的气象数据需要计算，如果用传统的计算工具，大约要几个星期甚至几个月才能计算出一个近似值，显然，计算出的结果也不能称为“预报”了。如果用计算机来解决这一问题，只要几分钟就能得到准确的结果，既及时又准确。

2. 数据处理

数据处理泛指不是以单纯地求解数学问题为目的的所有计算任务和各种形式的数据资料的处理。其特点是要处理的原始数据量大，算术运算比较简单，有大量的逻辑运算与判断，结果要求以表格文件的形式存储、输出等。例如数据报表、资料统计和分析、工农业产品的合理分配、工业企业的各种计划编制、企业成本核算、人事管理、财务管理、仓库管理等。

在学校的学籍档案管理中，需要对大量的数据(如学生的学号、姓名、性别、出生年月、每学期开设的科目及各科成绩、个人总分与平均分、排列名次、奖学金等)进行收集、汇总、存档、输出等操作。这一操作可由计算机来帮助完成，使学校管理部门和教师、学生能及时掌握教学和学习情况。

我国的银行由于采用计算机管理，大大提高了工作效率，还可以在同一城市内实现通存通兑，大大方便了用户。

不少国家已使图书检索自动化。查书目、借书、查阅资料全部由计算机完成，为科学工作者提供了极大方便。

数据处理的另一个重要领域是图像处理，如卫星图像分析(通过计算机处理从卫星发回的大量数据，分析出地面上哪些是山脉、哪些是海洋、哪些是军事目标、哪些是城市等等)，再如医院中使用的CT扫描等。

3. 实时控制

实时控制是指用计算机及时地收集检测被控制对象运行情况的数据，再通过计算机来分析处理后，按照某种最佳的控制规律发出控制信号，以控制过程的进展。

应用计算机进行实时控制可以提高生产自动化水平、提高劳动效率与产品质量、降低生产成本、缩短生产周期等。应用计算机还可实现一台机器、多台机器甚至整个车间或整个工厂的自动控制。导弹的发射等国防尖端科学技术更是离不开计算机的实时控制。

4. 计算机辅助系统

所谓计算机辅助系统就是用计算机来帮助我们完成各种工作。如计算机辅助设计(CAD), 计算机辅助制造(CAM), 计算机辅助测试(CAT), 计算机辅助教学(CAI), 计算机辅助病员管理(CAPM)等。

计算机辅助设计已广泛用于船舶、飞机、建筑工程及大规模集成电路等的设计工作中。这一技术使设计工作实现了自动化或半自动化, 既缩短了设计周期、提高了设计质量, 又降低了设计成本、提高了效率。例如, 大规模集成电路版图设计要求在几平方毫米的硅片上制造上万甚至几十万个电子元件, 线条只有几微米宽, 人工根本无法设计, 这就需要借助计算机来完成。如果把CAD、CAM、CAT技术有效地结合起来, 就可以使设计、制造、测试全部由计算机来完成, 大大减轻科技人员和工人的劳动强度。

CAI技术是用计算机来代替教师去实施教学计划。它可以模拟某一物理过程, 使教学过程形象化。也可以把课程内容编制成计算机软件(称为“课件”), 对不同学生可以选择不同的内容和进度, 改变了教学的单一模式, 有利于因材施教。还可以利用计算机来辅导学生, 解答问题、批改作业、编制考题等。这种教学方式直观、形象, 能激发学生的学习兴趣, 因而能提高教学质量。

5. 办公自动化

办公自动化(OA)是指以计算机或数据处理系统来处理日常例行的事务工作。它应具有完善的文字处理功能, 较强的资料、图像处理能力和网络通讯能力。例如, 起草文稿, 收集、加工、输出各种信息等。办公自动化系统除用计算机作为信息处理工具外, 还应包括复印机、传真机、通讯设备等其它设备。

6. 人工智能

人工智能是探索计算机模拟人的感觉和思维规律的科学。它是控制论、计算机科学、仿真技术、心理学等多学科的产物。人工智能的研究和应用领域包括: 模式识别、自然语言理解、专家系统、自动程序设计等。

“自然语言理解”是人工智能的一个分支。要使计算机能理解人类用的自然语言, 就需要根据上下文和人们已有的知识才能, 分析判断某一句或某一段话的确切含义。我国已经研制成功的英汉翻译系统, 是计算机应用的一个重大突破。

“专家系统”是人工智能的又一个重要分支。它的作用是使计算机具有某一方面专家的知识, 利用这些知识来处理所遇到的问题。目前比较成熟的是计算机辅助诊断系统, 它能模拟医生分析病情, 开出处方和假条。

“机器人”是人工智能最前沿的领域, 可分为“工业机器人”和“智能机器人”, 前者可代替人进行危险作业(如海底作业、高空作业等), 后者具有某些智能, 能根据不同情况进行不同的动作(如给病人送药、代替门卫值班等)。

计算机还可以用来对弈、作曲、绘画等。目前有很多动画片就是用被称为“三维动画”的软件“3DS”制作的。

由以上可见,计算机的应用领域十分广泛。“计算机”这个名词只是由于它初期主要用于数值计算而得名。现在计算机在非数值计算中的应用已远远超过在数值方面的应用。实际上,称它为“信息处理机”更为确切。我们在沿用“计算机”这个名词时,应当对它有个全面的理解。

1.3 计算机中数的表示

前面已经提到,现在的计算机都是采用冯·诺依曼体系结构。这种结构的计算机之所以使用二进制,而不使用十进制,是由于具有两个稳定状态的电子元件容易找到,如晶体管的导通与截止、开关的开与闭、电位的高与低,都可以表示二进制中“1”和“0”两个数码。而要做到具有十个稳定状态的电子元件来表示十进制中的0、1、2、……、9十个数码是很困难的。此外,二进制的运算规则也很简单、容易实现。但是二进制数由“0”和“1”组成,不便于书写和记忆,因而在使用计算机时,还经常使用八进制和十六进制。

1.3.1 进位计数制

按照进位的原则进行计数的方法称为进位计数制(简称进制)。例如,十进制就是进位计数制的一种,它是根据“逢十进一”的原则进行计数的。

在日常生活中,我们会遇到各种进位计数制的问题。例如,数学中使用的是十进制;时间使用的是六十进制,1小时等于60分,1分等于60秒。

进位计数制中的几个基本概念定义如下:

- ① 数字:是数的表示形式。
- ② 数码:构成数字的基本符号,如0、1、……、9等。
- ③ 数值:数的大小,是一个量的概念。
- ④ 进制:按照进位的原则进行计数的方法称为进制。

每一种进制都有一个固定的基数 J ,它的每一位都可以取 J 个不同的数码。例如,十进制的基数是10,它的每一位可以是0~9十个数码中的任意一个。二进制的基数是2,它的每一位可以是0或1。十六进制的基数是16,它的每一位可以是0~9, A、B、C、D、E、F十六个数码中的任意一个。每一种进制都是“逢 J 进一”或称“借一当 J ”,小数点向左移一位或向右移一位,则该数缩小或扩大 J 倍。第 i 位数,对应一个固定的值 J^i , J^i 就称为该位的“权”。同一个数码在不同的位置代表的数值是不同的,某一位数码代表的数值的大小是该位数码与该位“权”的乘积,就是该位的位值。数字的值就是该数字所有位的位值之和,简称数值。这种计算数值的方法称为按权展开相加法。

例如,十进制数的第 i 位的权是 10^i ,十进制数101.111的百位($i=2$)的权是 $10^2=100$,百位的1代表 $1 \times 10^2=100$;个位($i=0$)的权是 $10^0=1$,个位的1代表 $1 \times 10^0=1$;它的小数点后的第一位的权是 10^{-1} ,该位的1代表 $1 \times 10^{-1}=0.1$ 。

十进制数101.111按权展开相加为:

$$1 \times 10^2 + 0 \times 10^1 + 1 \times 10^0 + 1 \times 10^{-1} + 1 \times 10^{-2} + 1 \times 10^{-3} = 101.111$$

类似地,二进制的每一位的权是 2^i 。

二进制数 101.111 按权展开相加为：

$$1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} = 5.875$$

可见，同样的数码序列在不同的进制中所表示的数值是不相等的。为了区别所使用的进制，在几种进制混合使用的场合，常用下标或代号标注。各种进制的代号为：二进制用 B、八进制用 Q、十六进制用 H。考虑到使用十进制的习惯，在其它进制已有标注的情况下，十进制也可不加标注或用“D”标注。例如 $(3123)_{10}$ 或 3123D 表示十进制数 3123。

1.3.2 不同进制数之间的转换

不同进制数之间的转换是根据“如两个有理数相等，则两数的整数部分和小数部分一定分别相等”的原则进行的。

1. 十进制数和二进制数之间的转换

(1) 二进制数转换为十进制数——按权展开相加

把二进制数转换为相应的十进制数，不管是整数部分，还是小数部分，只要将它的每一位按权展开相加，就得到相应的十进制数。

$$\begin{aligned} \text{例如，} (1011001.101)_2 &= 1 \times 2^6 + 0 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 \\ &\quad + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} \\ &= 64 + 16 + 8 + 1 + 0.5 + 0.125 = (89.625)_{10} \end{aligned}$$

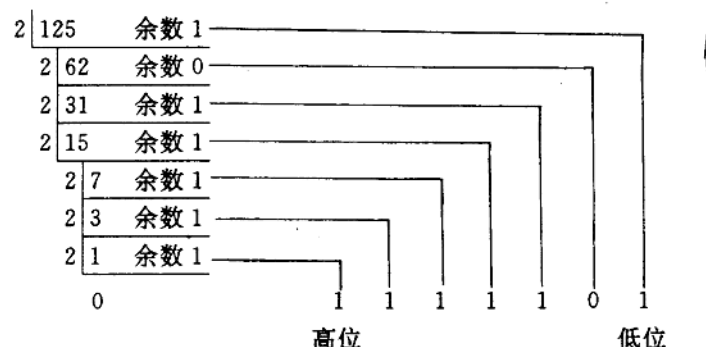
从中看到，只需把二进制数对应“1”的位的权相加即可。

$$\text{例如，} (11011001.101)_2 = 128 + 64 + 16 + 8 + 1 + 0.5 + 0.125 = (217.625)_{10}$$

(2) 十进制数转换为二进制数

① 十进制整数转换为二进制整数——除 2 取余。十进制整数转换为二进制数，采用除 2 取余法。即把十进制数反复除 2，若某次相除后余数为 1，则对应的二进制数相应位为 1；若余数为 0，则相应位为 0，直到商为 0 为止。先得到的余数是二进制的低位，后得到的余数是二进制的高位。

例如，将十进制 125 转换成二进制数：



所以， $125D = 1111101B$

② 十进制小数转换为二进制小数——乘 2 取整。十进制小数转换为二进制小数采用乘 2 取整法。即把十进制小数反复乘 2，若某次乘 2 后所得乘积的整数部分为 1，则对应的二进制数相应位为 1；若整数部分为 0，则相应位为 0。直到满足精度要求或乘积的小数部分

为0为止。先得到的整数是二进制小数的高位，后得到的整数是二进制小数的低位(远离小数点的位)。

例如，把十进制小数 0.196 转换为二进制数：

$$0.196 \times 2 = 0.392$$

$$0.392 \times 2 = 0.784$$

$$0.784 \times 2 = 1.568$$

$$0.568 \times 2 = 1.136$$

$$0.136 \times 2 = 0.272$$

.....

所以， $0.196D = 0.00110 \dots\dots B$

注意：一个二进制小数可准确地转换成一个十进制小数，因为它是按权展开的和。但一个十进制小数却不一定能完全准确地转换成二进制小数，因为某一个十进制小数不一定能恰好用二进制小数各位的权之和来标识，因此，根据需要选取适当的有效位数。

③ 十进制数转换为二进制数。将十进制数转换为二进制数，先要分成整数部分和小数部分，然后把它们分别转换成二进制，最后把这两部分用小数点连接起来。

例如，把十进制数 196.375D 转换为二进制数。

因为， $196.375D = 196D + 0.375D$

$$196D = 11000100B$$

$$0.375D = 0.011B$$

所以， $196.375D = 11000100.011B$

十进制数和八进制数、十六进制数之间的转换可采用相似的方法实现，不同的只是分别采用除 8 取余、乘 8 取整和除 16 取余、乘 16 取整，这里不再重复介绍。

2. 十六进制数和二进制数之间的转换

因为四位二进制可以表示 0~15 的数值，所以四位二进制数可用一位十六进制数来标识。例如，

$$1011B = 8 + 2 + 1 = 11D = BH$$

由于这种关系，十六进制数与二进制数之间的转换既简单又明了。

(1) 二进制数转换为十六进制数

二进制数转换为十六进制数以小数点为界，对整数部分从低位向高位每四位一组，不足四位时，在左边用 0 补足四位，然后计算出每组对应的十六进制数的值；对于小数部分从高位向低位，每四位为一组，不足四位时，在右边用 0 补足四位(注意：不补 0 的值是错误的)，然后计算对应的十六进制数的值。

例如，将二进制数 101001010111110.101B 转换成十六进制数。

$$0101 \quad 0010 \quad 1011 \quad 1110 \quad . \quad 1010$$

$$5 \quad 2 \quad B \quad E \quad A$$

所以， $101001010111110.101B = 52BE.AH$

(2) 十六进制数转换为二进制数

十六进制数转换为二进制数是上述转换方法的逆过程，只要把一位十六进制数用相应的四位二进制数表示，然后合起来即可。

例如，将十六进制数 3C8D.7EH 转换成二进制数。

3	C	8	D	.	7	E	H
0101	1100	1000	1101	.	0111	1110	B

所以，3C8D.7EH=101110010001101.0111111B

最前面的 0 和最后的 0 是没有意义的，一般可在最后结果中省去。

八进制数与二进制数之间的转换可采用相似的方法实现，区别只是每三位一组，这里不再重复介绍。

上述三种数制，十进制是人们最熟悉的，但在计算机内部目前还不能直接使用。二进制是计算机内部信息存储、处理、传送的基本数制，但人们不易使用和记忆。十六进制是介于十进制和二进制之间的一种计数方法，它把二进制缩写成人们比较容易记忆和使用的形式，解决了计算机和人之间在数制应用上的矛盾。

1.4 计算机中的编码

在计算机内部使用的都是二进制数，因为计算机只能识别“0”和“1”两个状态。为了使计算机使用的数据能够共享和传递，必须对计算机中的数、字母、符号和汉字等进行统一的编码——即用特定的代码表示。

计算机中编码的几个基本概念定义如下：

① 信息：是客观事物的特征。
② 数据：是客观事物属性的描述与记载，是信息的符号，是信息在计算机中的表示形式。

③ 数码：表达数据的符号，也称为代码；用码表达数字的时候就称为数码。

④ 编码：用码表达数据的过程或方法称为编码，有时人们也把编码的结果称为编码。

⑤ 字节：8 位二进制数(后面会详细介绍)。

目前国际上使用的编码种类很多，下面介绍计算机中常用的几种编码。

1.4.1 数的编码

计算机内部是以二进制进行运算的。但是实际应用中一般问题的原始数据大多是十进制数，这就要求输入计算机时，将十进制数转换为二进制数；输出时将二进制数转换为十进制数。这项工作是由计算机自动完成的。因此要求所采用的编码便于计算机识别和转换，通常采用若干位二进制数码来表示一位十进制数的方法，统称为十进制的二进制编码，简称二—十进制编码(BCD 码)。

由于十进制数的每一位有 0~9 十个不同的状态，因此用二进制数表示十进制数时，每一位十进制数需要由四位二进制数表示，而四位二进制数能表示十六种状态，其中六种是多余的，而这种多余便产生了多种不同的 BCD 码。这里只介绍使用最广泛的 8421BCD 码。

四位二进制数的权分别为 8、4、2、1 的 BCD 码叫 8421 码，见表 1-1。它所表示的数规律与二进制数相同，所以是最简单的。

表 1-1 8421 BCD 编码表

十进制数	8421 BCD 码	十进制数	8421 BCD 码
0	0000	8	1000
1	0001	9	1001
2	0010	10	0001 0000
3	0011	11	0001 0001
4	0100	12	0001 0010
5	0101	13	0001 0011
6	0110	14	0001 0100
7	0111	15	0001 0101

1.4.2 ASCII 码

ASCII 码是目前计算机中使用的最广泛的一种编码。它的全称是 American Standard Code for Information Interchange, 即美国标准信息交换码。

ASCII 码共有 128 个字符, 其中包括 32 个通用控制符、10 个十进制数码、52 个英文字母(大小写各 26 个)及 34 个专用符号。因为一共有 128 个字符, 所以用二进制编码共需 7 位。通常采用 8 位二进制数(一个字节)表示一个字符编码, ASCII 码使用其中的低 7 位, 一般情况下, 可认为最高位为 0。

在 ASCII 码表中, 数码的编码小于字母, 大写字母的编码小于小写字母的编码。在数码编码中, 0 的编码最小, 9 的编码最大, 字母的编码按字典顺序递增。由于每个字符对应一个编码, 而编码是可以比较大小的。因此, 字符就可以比较大小, 字符的比较在计算机内部就是对它们的 ASCII 码进行比较。

西文编码系统比较简单。在键盘上输入的就是英文词语的字母及有关字符, 它们立即转换为 ASCII 码作为计算机的内部代码进行存储或处理。进行显示时, 通过字符发生器即可获得用来显示或打印的点阵信息。

1.4.3 汉字信息编码

我国是使用汉字的国家, 要想普及计算机的使用, 首先要解决汉字的输入问题, 必须使用某一规则将数万个汉字编码, 其次还要解决汉字的输出(显示和打印)问题。

汉字信息处理的关键在于计算机对汉字代码的数据处理, 使得人们在计算机上使用汉字和使用西文一样方便和容易。在计算机内部, 西文和中文只不过被人为地定义成不同的符号, 用不同的信息表示而已。从信息处理角度来看, 汉字信息的处理和西文信息的处理没有本质的区别, 两者都是非数值信息。

计算机的汉字系统对每个汉字规定了输入计算机的代码——即汉字的外码, 从键盘输入汉字是输入汉字的外码。计算机为了识别汉字, 要把汉字的外码转换成汉字的机内码, 以处理和存储汉字信息。为了将汉字以点阵的形式输出, 还要将汉字的内码转换为汉字的

地址码,再根据地址码得到字形码,以确定一个汉字的点阵。当计算机和其它系统或设备进行信息交流时,所交流的数据还要用交换码(国标码)。下面分别介绍几个码的概念。

1. 外码

汉字主要是从键盘输入,每个汉字都对应一个汉字的外码,外码是计算机输入的代码,是输入某一个汉字时的一组键盘符号。外码也叫输入码。例如,在用拼音输入汉字时,“计”字的外码是“ji”。

一般来说,汉字输入码应具备如下特点:

- ① 易于记忆,甚至不需要记忆。
- ② 编码长度尽可能短,以加快输入速度。
- ③ 编码与汉字的对应性好,尽量减少重码,并有简便地区分重码的方法。
- ④ 编码规律性强,易于学习、掌握和使用。

在这四个特点中,有两个基本概念,第一个是码长,即汉字输入码的编码长度,人们希望编码长度越短越好,因为编码长度短,可以加快输入速度。第二个是重码,即同一个输入码对应多个汉字。

汉字的输入码有多种编码方法,衡量一个编码方法好坏的重要标准之一是重码多少,第5章将对常用的汉字输入方法作详细的介绍。

2. 交换码

国家规定了标准汉字交换码——GB2312-80(国标码),这是计算机与其它设备之间交换汉字信息的统一标准,它为计算机处理的汉字与其它设备处理的汉字之间的交换建立了桥梁。

国标码共容纳了7445个汉字及符号,其中汉字6763个,一般符号如数字、外文字母、图形符号等682个,分布在94个区中。由于给每个汉字分配了标准代码,以供交换汉字信息使用,所以,国标码又称为交换码。国标中的汉字分为两级,将其中使用频度高的常用汉字作为一级汉字,共3755个。不常用的汉字作为二级汉字,共3008个。一级汉字按汉语拼音字母顺序排列,多音字取常用的发音,同音字以起笔笔形(横竖撇点折)顺序排列,第一笔相同时,再按第二笔的顺序,以此类推。二级汉字字库按部首排列。

区 位	1~94 位
1~15	
16~55	第一级汉字(3755个)
56~94	第二级汉字(3008个)

图 1-1 GB2312 代码分布示意图

图 1-1 是 GB2312 代码分布图。

从图中可以看出,纵向分成94个区,由汉字交换码的第一个字节代表。横向把每个区分成94个位,由汉字交换码的第二个字节代表。汉字从16区开始,16~55区为一级汉字,56~94区为二级汉字。

汉字的国标码用十六进制表示。

例如,“粗”字的国标码为:

00110100	01010110
第一字节	第二字节

用十六进制表示则为 3456。

3. 区位码

区位码是根据汉字在 GB2312 代码分布的位置(区号和位号)进行编码。例如“粗”字是在 20 区 54 位, 所以区位码为 2054。

区位码也可作为一种汉字输入码。

4. 机内码

机内码亦称内部码、内码, 是指在计算机内部进行存储和处理而使用的代码。对于西文系统, 交换码和机内码是一回事, 都是 ASCII 码。但是对于汉字系统, 交换码和机内码是两回事。汉字内码由两个字节存储, 这两个字节的高位全是“1”。

例如, “粗”字的内码为:

1 0110100	1 1010110
第一字节	第二字节

用十六进制表示则为 B4D6。

每个汉字的国标码与其内码是不同的, 但二者都是用两个字节表示, 只是内码这两个字节的高位全是“1”, 而国标码这两个字节的高位全是“0”。也就是说同一个汉字的国标码与其内码之间有一一对应关系, 只要将汉字的内码的两个字节的高位的“1”全变为“0”则是该汉字的国标码。反之, 若将汉字的国标码的两个字节的高位加 1 便是该汉字的内码。

根据国标码可以计算出内码, 即汉字国标码与内码的对应关系如下:

十六进制内码 = 十六进制国标码 + 8080H。

这种转换方法称为“高位加 1 法”。

例如, “粗”字的国标码与内码的关系为:

$B4D6H = 3456H + 8080H$ 。

根据区位码也可以计算出机内码。方法是将区号和位号分别加 160, 然后用十六进制表示就是机内码。

例如, “粗”在区位码是 2054, 即区号是 20, 位号是 54。

$20 + 160 = 180$ 变为十六进制为 B4H。

$54 + 160 = 214$ 变为十六进制为 D6H。

所以, “粗”字的机内码为 B4D6H。

这样定义的内码, 有下列优点:

- ① 汉字的内码与国标码、区位码存在简单的转换关系。
- ② 汉字的内码与西文的内码容易区分。凡是最高位为“1”的两个字符就当作一个汉字看待; 凡是最高位为“0”的就当作一个西文字符。
- ③ 2 字节内码可表示汉字的个数是 $2^8 - 2 = 16\,384$, 足够覆盖常用的近 8000 个汉字。
- ④ 一个汉字用 2 个字节内码表示, 正好和汉字在显示和打印时的特征相一致, 显示和打印一个汉字正好占据 2 个西文字符的位置, 从而使中西文混合使用变得容易实现。

我们在存储汉字时, 存储的是汉字的内码。因此, 一篇有 10 000 个汉字的文章, 只需要 20 000 个字节即可保存在计算机中。

国标码、区位码、内码都是将汉字和字符按照某种顺序排列, 再按一定方式编号, 用

此编号作为汉字编码，都是采用数字编码且有序，也称为数字编码或“流水码”。

5. 汉字输出码

汉字输出主要指显示输出和打印输出汉字字形，不同字体、不同点阵的汉字字形是不同的。

(1) 字体

同一种文字可以有不同的字体，它与汉字的尺寸无关。例如，汉字的基本字体有：宋体、仿宋体、楷体、黑体等，还有由基本字体变化而来的美术体，如长体、扁体、粗体、细体、斜体、立体、中空体、阴影体等。另外按笔画不同又分为简体、繁体等。ASCII 码的字体也可分为半角字符和全角字符两种字体。

(2) 汉字字形数字化

尽管汉字字形有多种变化，但由于汉字都是方块字，每个汉字都同样大小，无论汉字的笔画多少，都可以写在同样大小的方块内。于是可以将该方块看成是 m 行 n 列的矩阵，简称点阵。一个 m 行 n 列的点阵中共有 $m \times n$ 个点，每个点可以是黑点或白点，黑点组成汉字的笔画，描绘出汉字的字形，称为点阵字形。

在计算机中可以用一组二进制数表示一个汉字的点阵，用数字“1”表示点阵中的黑点，其余用“0”表示。例如一个 16×16 点阵的汉字要用 $16 \times 16 = 256$ 个“1”或“0”表示这 256 个点所描绘的汉字，共用 32 个字节，点阵的每一列用两个字节存储，存储时是将点阵按从左到右的顺序各列存储，同列中先存储上边 8 个点，再存储下面 8 个点。这些二进制数称为该汉字的输出码。

目前常用的汉字点阵是 16×16 、 24×24 、 32×32 、 40×40 、 48×48 、 64×64 、 72×72 、 96×96 、 108×108 等。在每个汉字的点阵中行数和列数越多，描绘的汉字越细致，但每个汉字占用的存储空间越大。通常 16×16 点阵供屏幕显示汉字使用， 24×24 点阵供打印机打印汉字使用。汉字字形数字化后可得到汉字的字形码，也称字模。

(3) 汉字库

汉字字形码以二进制文件的形式存储在外存储器上构成了汉字库。汉字库也称为汉字字形库，简称字库。对于不同的汉字系统，汉字库中所包含的字数也不同，但一般汉字库应包含全部国标 GB2312 中规定的一级和二级汉字和字符。汉字库通常放在外存上，开机后调入内存，或由 ROM（只读存储器）作汉字库。微机上使用的汉卡就是具有汉字输入功能和 ROM 汉字库的插件板。

(4) 汉字地址码

用来指明汉字字形码在汉字库中存放位置的代码称为汉字地址码。需要向输出设备输出汉字时，必须通过地址码，才能在汉字库中取到所需要的字形码，最终在输出设备上形成可见的汉字字形。地址码和内部码要有简明的对应关系（不同的汉字系统，其对应关系也不同），以便输出时有效地从字库中找到相应的字形。

由于汉字信息处理的飞速发展和汉字处理的复杂性，现在汉字编码还没有完全统一，也没有完整的标准。

习 题 1

1. 计算机从原理上可分为_____和_____,从用途上可分为_____和_____,根据计算机规模大小和功能强弱可分为_____,_____,_____,_____。
2. 将十进制数 1234、5678 分别转换为二进制、八进制、十六进制。
3. 从键盘输入汉字是输入汉字的_____,为了识别汉字要把汉字的_____转换成汉字的_____,为了将汉字以点阵的形式输出,还要将汉字的_____转换成汉字的_____,再根据_____得到_____,以确定一个汉字的点阵。
4. 简述汉字国标码与内码、区位码与内码的关系;简述汉字内码表示方法的优点。

第2章 计算机系统概述

一个完整的计算机系统由硬件系统和软件系统两大部分组成。如果把硬件系统比作计算机的“躯体”，则软件系统就是计算机的“灵魂”。一个计算机系统如果只有硬件系统，没有软件系统，则是一具“僵尸”。同样，一个计算机系统如果只有软件系统，而没有硬件系统，则是一个“幽灵”。这两大部分相互依存、相互支持，缺一不可。

本章简单介绍计算机的硬件系统和软件系统及计算机系统的主要性能指标。

2.1 计算机的基本结构

2.1.1 计算机硬件的组成

计算机的硬件系统主要由运算器、控制器、存储器、输入设备和输出设备五大部分组成。存储器分为主存储器和辅助存储器。通常把运算器和控制器称为中央处理器(CPU)，CPU和主存储器称为主机，输入设备、输出设备和辅助存储器称为外部设备(外设)。它们是看得见摸得着的物理实体，是硬设备，所以称为硬件，它是计算机的物质基础。硬件的基本功能是在软件系统控制下，完成数据的输入、运算、输出等一系列操作。通常，将一个仅由硬件组成的计算机称为“裸机”。

微机的基本结构也不外乎上面所说的五大部分。微机用大规模集成电路技术把运算器和控制器集成在一个称为微处理器的芯片上，再配以大规模集成电路的主存储器芯片。通过接口电路联结输入、输出设备就构成了微机的硬件系统。

2.1.2 计算机的外部结构

我们以典型的微机来看计算机的外部结构。从外部结构上看，微机通常由显示器、机箱、键盘三大部分组成。微机的外部结构如图2-1所示。其中键盘和显示器属于输入和输出设备，程序和原始数据通过键盘输入计算机，经过加工后，再通过显示器显示运算结果(必要时可通过打印机打印出来)。机箱内装有主板、电源、显示卡、多功能卡、硬盘、软盘驱动器、光盘驱动器等很多硬件，机箱面板上有电源开关、复位键、指示灯等。

2.1.3 系统总线

总线是一组物理导线，它是各部件之间传输信息的公共通道，用以完成各部件之间的信息交换。系统总线(BUS)是连接中央处理器(CPU)、存储器、输入/输出(I/O)接口电路的一组信号线。它由数据总线、地址总线、控制总线三种总线组成。