



普通高等教育“十一五”国家级规划教材

计算机应用基础

(第5版)

· 高等院校非计算机专业信息技术基础教材 ·

郭永青 李祥生 胡加立/主编



北京大学医学出版社

普通高等教育“十一五”国家级规划教材

高等院校非计算机专业信息技术基础教材

计算机应用基础

(第5版)

主 编

郭永青 李祥生 胡加立

编 委

(以姓氏笔画为序)

王 静	北京大学
王爱环	山西省心血管病医院
王路漫	北京大学
兰顺碧	华中科技大学
齐惠颖	北京大学
李祥生	山西医科大学
郑 凤	北京大学
胡 彬	华中科技大学
胡加立	北京大学
秦立轩	华中科技大学
郭永青	北京大学
郭建光	北京大学
温厚津	北京大学

北京大学医学出版社

图书在版编目(CIP)数据

计算机应用基础 / 郭永青, 李祥生, 胡加立主编. — 5 版.
— 北京: 北京大学医学出版社, 2009.2

ISBN 978-7-81116-713-9

I. 计… II. ①郭… ②李… ③胡… III. 电子计算机—医学院校—教材
IV.TP3

中国版本图书馆 CIP 数据核字 (2009) 第 009667 号

主 编

立加胡 主祥李 青永郭

委 编

(书号图第册数划)

学大京北

编 王

学大京北

编 王

学大京北

编 王

学大京北

编 王

计算机应用基础 (第 5 版)

主 编: 郭永青 李祥生 胡加立

出版发行: 北京大学医学出版社 (电话: 010-82802230)

地 址: (100191) 北京市海淀区学院路 38 号 北京大学医学部院内

网 址: <http://www.pumppress.com.cn>

E-mail: booksale@bjmu.edu.cn

印 刷: 北京瑞达方舟印务有限公司

经 销: 新华书店

责任编辑: 罗德刚 张其鹏

责任校对: 齐欣

责任印制: 郭桂兰

开 本: 787mm × 1092mm 1/16 印张: 19 字数: 475 千字

版 次: 2009 年 2 月第 5 版 2009 年 2 月第 1 次印刷 印数: 1 - 3000 册

书 号: ISBN 978-7-81116-713-9

定 价: 34.50 元

版权所有, 违者必究

(凡属质量问题请与本社发行部联系退换)

前 言

《计算机应用基础》是普通高等教育“十一五”国家级规划教材，面向医药类院校等非计算机专业的学生。《计算机应用基础》从 1994 年第 1 版到这次的第 5 版的出版发行，每次再版都凝聚了致力于计算机基础教育教师们的辛勤汗水。在这里首先要感谢北京大学医学部及北京大学医学出版社为计算机基础教学的教材建设所提供的大力支持，还要感谢参加教材编写的华中科技大学网络与计算中心教研室、山西医科大学计算中心、北京大学医学部公共教学部计算机教研室的各位老师。

随着社会的信息化、数字化的进程快速推进，为贯彻科学发展观，切实落实教育部《关于进一步深化本科教学改革全面提高教学质量的若干意见》（教高[2007] 2 号）的精神，结合教育部高等学校计算机基础课程教学指导委员会提出的 1+X 模式，我们重新调整了教学内容，将第 4 版《微机应用基础》的上下两册合并成为一册，更改名为《计算机应用基础》，将它作为医药类学生进入大学的第一门计算机基础课程教材，以适应医药类学生的医学课程任务重而学习信息技术课时少的特点，以及近年来入校学生计算机知识和素养有所提高的现实情况。

我们总结了医药学校多年来的教学经验，将网络基础应用这部分内容并入第一章介绍给学生；将数据库与程序设计（如数据库采用 Access 和程序设计采用 VBA）作为最后一章编写，便于授课的教师根据课时来进行内容扩展或缩减。

本书共分为七章，第一章计算机与网络（郭永青、胡彬、李祥生编写），第二章计算机硬件系统（郭建光编写），第三章软件系统（王静、郭永青编写），第四章常用应用软件（郑凤、温厚津、李祥生、王爱环、齐惠颖编写），第五章多媒体技术应用（兰顺碧、秦立轩、王路漫编写），第六章网页制作软件（齐惠颖编写），第七章数据库基础与应用（胡加立、齐惠颖编写）。在每一章节后面，我们都精心编写了习题和上机练习题，便于读者自测练习和有目的的上机操作练习。全书由郭永青、胡加立统稿，郑凤负责排版。

由于水平有限，书中如有错误之处，敬请读者批评指正。

编者

2008 年 12 月 31 日

目 录

第一章 计算机与网络	1
1.1 电子计算机的发展、特点和应用	1
1.2 计算机中采用的计数制	4
1.3 计算机中的数据信息的表示	7
1.4 计算机的组成	10
1.5 信息技术与信息数字化	13
1.6 网络应用基础	14
1.7 计算机与网络安全	33
第二章 计算机硬件系统	46
2.1 计算机硬件系统结构	46
2.2 微型计算机的核心组件	46
2.3 微型计算机总线	55
2.4 输入输出设备接口	57
2.5 输入输出设备	59
2.6 多媒体计算机技术	60
第三章 软件系统	62
3.1 计算机软件概述	62
3.2 操作系统概述	63
3.3 Windows XP 操作系统	69
3.4 软件开发基础知识	84
第四章 常用应用软件	96
4.1 文字处理软件 Word	96
4.2 电子表格处理软件 Excel	124
4.3 幻灯片制作软件 PowerPoint	145
4.4 PDF 文档阅读与制作	155
第五章 多媒体技术应用	162
5.1 多媒体概述	162
5.2 多媒体环境的建立	166
5.3 声音处理软件	171
5.4 Flash 多媒体动画素材的编辑与制作	175
5.5 Photoshop 图像处理软件	188
第六章 网页制作软件	202
6.1 网页制作基础	202

6.2	使用 Dreamweaver 8 建立站点	205
6.3	制作网页	206
6.4	表格的应用	209
6.5	创建超链接	211
6.6	使用 CSS 样式表	213
6.7	层和行为	216
6.8	库和模板	219
6.9	框架	222
6.10	表单	223
6.11	制作多媒体网页	225
6.12	站点发布	226
第七	数据库基础与应用	229
7.1	数据库技术概述	229
7.2	数据库的建立与操作	237
7.3	数据查询	247
7.4	窗体	256
7.5	宏	266
7.6	报表概述	270
7.7	VBA 模块	279
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		

第一章 计算机与网络

电子计算机 (Electronic Computer) 是一种用电来进行各种信息加工的机器, 它可以按照预先编好的程序自动执行各种操作, 以完成信息的输入、存取、加工处理及输出。在当今信息化时代, 计算机是信息自动化处理的最基本、最有效的工具。计算机技术与通信技术的结合促使计算机网络产生, 计算机网络迅速发展, 给世界带来了很大的变化, 已成为人们生活和工作不可或缺的部分。在以网络为基础的信息社会里, 人们的行为方式、思想方式甚至社会形态都发生了显著的变化。

1.1 电子计算机的发展、特点和应用

回顾计算机发展历史, 自 1946 年世界第一台电子计算机问世, 经历了电子管、半导体、集成电路、大规模集成电路、超大规模集成电路等几代的发展, 其性能提高程度以指数形式增长。超级计算机、光计算机、生物计算机等的研制开发取得了显著成果。计算机技术的发展, 使人类利用“0”和“1”编码技术, 来实现对一切声音、文字、图像和数据的编码和解码, 使各类信息的采集、处理、储存和传输实现了标准化和高速处理。

1.1.1 计算机发展简史

人类的计算技术有着悠久的历史, 我国的祖先发明的算盘至今在某些领域还在使用。在 19 世纪, 由于西方国家生产力的发展, 使普通的计算工具难以完成计算的需要, 因此, 人类一直在寻求新的计算技术。19 世纪 50 年代, 英国数学家乔治·布尔 (George Boole) 创立了逻辑代数, 用二进制进行运算, 是当前电子计算机的数学基础。1936 年英国科学家图灵 (Alan Mathison Turing) 首次提出逻辑机的通用模型, 即“图灵机”, 建立了算法理论, 为计算机的出现提供了重要的理论依据, 被称为计算机之父。1946 年 2 月世界上第一台电子计算机在美国宾夕法尼亚大学诞生, 被命名为 ENIAC (Electronic Numerical Integrator and Calculator) 即“电子数字积分计算器”。它共使用了 18 000 只电子管, 耗电 150 千瓦, 重约 30 吨, 占地约 170 平方米, 每秒能进行 5000 次加法运算。ENIAC 的问世表明了电子计算机时代的到来, 它的出现具有划时代的意义。鉴于 ENIAC 的缺点, 美籍匈牙利数学家冯·诺依曼 (John Von Neumann) 于 1946 年 6 月发表了“电子计算机装置逻辑结构初探”的论文, 提出三个要点: 其一是计算机所有数据和程序都采用二进制; 其二是将程序和指令顺序存放在内存存储器, 且能自动依次执行指令; 其三是计算机由输入设备、输出设备、内存存储器、运算器和控制器五大部分组成。采用以上结构的计算机, 被称为冯·诺依曼结构计算机。

英国剑桥大学威尔克斯 (M.V. Wilkes) 教授在 1946 年接受了冯·诺依曼的存储程序计算机结构原理后, 在剑桥大学设计了 EDSAC (The Electronic Delay Storage Automatic Computer) 计算机, 于 1949 年 5 月研制成功并投入运行。它是世界上首台“存储程序”

电子计算机。

1951年6月14日,第一台计算机作为商品交付使用,从此计算机从实验室走向社会,标志着人类进入计算机时代。

至今,所有的计算机仍然没有摆脱冯·诺依曼的理论,目前使用的仍是冯·诺依曼结构计算机。

1.1.2 计算机的分类

按用途分类,计算机分为专用计算机和通用计算机。专用计算机一般用于对其他设备的控制,比如工业自动生产流水线、医疗设备的控制分析等。通用计算机就是我们日常所用的可以对各种数据进行加工处理的计算机。

按其规模分类,计算机分为巨型机、大型机、小巨型机、小型机、工作站以及微型计算机(微机)。巨型机和大型机一般用于尖端科学,它的功能是最强的,速度和精度也最高。小型机一般用于大中型企业以及比较大的科研单位,功能仅次于巨型机和大型机。工作站一般用于计算机辅助设计。在这些机种中,微机的功能是最弱的,但应用领域最为广泛,大到企、事业单位,小到普通家庭,渗透于各领域。微机又分为台式机、便携机(笔记本电脑),近年来又出现了手持计算机(掌上电脑),总之,这一类计算机越来越小,而功能也越来越多。

由于微型计算机的发展,使过去由一台大型计算机带若干终端的集中化使用模式向人手一机、独立使用的分散化模式转变。近年计算机网络快速发展,就是由于微机的出现才使得家庭上网得以实现,另外它也促进了多媒体技术的发展。

我国的计算机发展起步较晚,但发展极为迅速,1956年国家制定12年科学规划时,把发展计算机、半导体等技术定为重点学科。1958年我国组装调试成第一台电子管计算机(103机),1959年研制成大型通用电子管计算机(104机),其运算速度为10 000次/秒。1964年我国推出了第一批晶体管计算机,其运算速度为10万~20万次/秒。1971年我国研制成功第三代集成电路计算机。1982年采用大、中规模集成电路研制成功16位计算机DJS-150。

1983年长沙国防科技大学推出向量运算速度达1亿次的银河I巨型计算机。目前世界上只有很少几个国家能生产巨型机,我国是其中之一。目前,我国已形成了相当规模的计算机产业。

当前,计算机正在向巨型化、微型化、网络化和智能化的方向发展,许多人正在探索新一代计算机的研究。对新一代计算机有各种各样的设想方案:有的研究非冯·诺依曼结构的计算机;有的研究具有高度智能的计算机,可以模仿人的大脑处理各种事务。除此之外,神经网络计算机、生物计算机、光子计算机也是许多人研究的热门课题。相信不久的将来各种类型的新一代计算机会出现在我们的工作、学习和生活当中。

1.1.3 计算机的应用

按应用领域进行划分,主要包括以下几个方面:

1. 科学计算 科学计算也称为数值计算,是指用于完成科学研究和工程技术中提出的数学问题的计算,如航天、气象分析等会产生大量的数值数据,都需要用计算机进行运

算处理。

2. 数据处理 数据处理也称为非数值计算,它可以对大量的数据进行加工处理。与科学计算不同,数据处理涉及的数据范围非常广泛,文字、图形图像、声音等都是计算机可处理的非数值数据。典型的应用是办公自动化和多媒体。

3. 过程控制 过程控制也称实时控制,又称自动控制,利用计算机及时采集数据,将数据处理后按最佳值迅速准确地对控制对象进行控制,代替人的手进行操作。典型的应用是工业生产流水线的自动化生产。

4. 计算机辅助系统 计算机辅助系统也是计算机应用方面的一大分支,它所包含的内容主要有 CAD、CAM 和 CBE。

计算机辅助设计 CAD (Computer Aided Design) 就是利用计算机帮助各类设计人员进行设计,例如飞机设计、船舶设计、建筑设计、机械设计、大规模集成电路设计等。

计算机辅助制造 CAM (Computer Aided Manufacturing) 是指计算机进行生产设备的管理、控制和操作的技术。在产品的制造工程,利用计算机对生产的各个环节进行监测,始终让生产工艺处于最佳状态,可以提高产品质量,缩短生产周期,减轻工人的劳动强度。

计算机辅助教育 CBE (Computer Based Education): 包括计算机辅助教学 CAI (Computer Assisted Instruction)、计算机辅助测试 CAT (Computer Aided Test) 和计算机管理教学 CMI (Computer Management Instruction)。近年由于多媒体和网络技术的发展,推动了 CBE 的发展,打破了过去传统教学手段,某些学科抽象的理论可以用计算机直观地模拟出来,另外远程教育也得到很快发展。

1.1.4 计算机技术在医药卫生领域的应用

在医药卫生领域,计算机的应用已渗透到医药学各学科,其主要应用包括以下几个方面:

1. 远程诊断 由于近年来网络日益成熟,许多医院都开通了远程诊断,使得许多专家不用到现场就可以通过网络的传输看到病人的所有病历资料,其中包括病人的即时动态医学影像(比如彩超动态图像),据此可以对疑难病例进行会诊,迅速作出诊断,及时采取措施,为抢救病人赢得宝贵的时间。

2. 医学图像处理 现代医学离不开影像信息的支持,如病理切片图像、X 射线透视图像、CT 和 MRI 扫描图像、核医学图像、超声影像、红外线热成像图像及窥镜图像等。功能各异的医学影像可分为结构影像技术与功能影像技术两大类,前者主要用于获取人体各器官解剖结构图像,借助此类结构透视图像,不经解剖检查,医务人员就可以诊断出人体器官的器质性病变,如 CT 和 MRI 扫描图像。然而在人体器官发生早期病变,但器官的外形结构仍表现为正常时,器官的某些生理功能,如新陈代谢等却已开始发生异常变化,此时采用结构解剖基于 SPECT 及 PET 的功能影像技术。功能影像能够检测到人体器官的生化活动状况,并将其以功能影像的方式呈现出来。

3. 人工脏器方面 以人工脏器替换病变或损伤的器官已经是很成熟的技术,其实是计算机技术才使人工脏器得以实现。以人工肾即血液透析机为例,血液透析机的体外循环系统包括血泵、肝素泵、血流量表、动静脉压表和空气探测器由计算机控制;透析液系统包括比例泵、透析流量计,超滤系统等也是由计算机控制;透析机的监测控制装置还要由计

算机统一管理。没有计算机,人工肾是不可能正常工作的。

4. 流行病学数据处理 流行病学涉及的范围极其广泛,大量调查资料的分析,利用计算机进行处理不仅能提高运算效率,也可以提高运算的准确性,大样本的数据资料都是用计算机处理,目前,SAS、SPSS 等都是很优秀的统计软件,尤其在科研上,它是医药卫生人员进行科学研究强有力的助手。

5. 医院信息管理 数字化医院是我国现代医疗发展的新趋势,数字化医院系统是由医院业务软件、数字化医疗设备、计算机网络平台所组成的三位一体的综合信息系统,数字化医院工程有助于医院实现资源整合、流程优化,降低运行成本,提高服务质量、工作效率和管理水平。

数字化医院的业务软件通常由以下几部分组成:

- (1) HIS (Hospital Information System) 医院信息系统
- (2) PACS (Picture Archiving and Communication Systems) 医学图像档案管理和通信系统
- (3) LIS (Laboratory Information System) 检验信息系统
- (4) CIS (Clinic Information System) 临床管理信息系统
- (5) RIS (Radiology Information System) 放射科信息系统
- (6) GMIS (Globe Medical Information Service) 区域医疗卫生服务

此外还有:CAE (计算机辅助教学系统)、CAD (计算机辅助诊断系统)、CAT (计算机辅助治疗系统)、CAS (计算机辅助外科系统)、RTIS (放射治疗系统)等。

6. 生物信息学 (Bioinformatics) 生物信息学是在生命科学的研究中采用计算机技术和信息论方法对蛋白质及其核酸序列等多种生物信息采集、加工、储存、传递、检索、分析和解读的一门科学,是现代生命科学与信息科学、计算机科学、数学、统计学、物理学和化学等学科相互渗透而形成的交叉学科。它是当今生命科学和自然科学的重大前沿领域之一,同时也将是 21 世纪自然科学的核心领域之一。生物信息学研究的内容包括了序列和结构比对、蛋白质结构预测、基因识别、分子进化、比较基因组学、序列重叠群、药物设计、基因芯片、基因表达谱等方面。

1.2 计算机中采用的计数制

在日常生活中我们习惯使用十进制,对于十进制可以用 0, 1, 2, …, 9 这十个数码表示,把这些数码的个数称为基数,即十进制的基数为十。采用逢基数进一的规则,则称为进位计数制。

除了十进制,人们也使用其他进制,有时还采用十六进制、六十进制,比如在很早以前使用过十六两为一斤的秤。

由于计算机内部使用的是数字电路,即用脉冲表示信号,而脉冲信号只有两种状态,电压的有无(即高电平和低电平)、灯光的亮与暗(灯泡加没加电),两种状态都可以用数码 0、1 来表示。两种状态的电路最容易实现,而且稳定、可靠,用开关的开启和关闭即可实现,只不过开关是由电子开关完成,如果出现三种以上的稳定状态,电路上实现起来

就复杂了。所以计算机中处理的各种信息都是用二进制代码来表示的, 现在对这些计数制作一简单介绍。

1.2.1 十进制数制 (Decimal Number System)

十进制使用 0~9 这十个数码表示, 它的基数是十, 它的计数规律为逢十进一。

十进制数的书写规则是将该数后面加 D 或在括号外加数字下标, 例如 237.68D 或 $(237.68)_{10}$ 都是表示十进制的 237.68。通常都省略。

在十进制中, 数值的大小不仅和其所用的代码有关, 还与其所在的位置有关, 比如 262.84 这个数, 六个代码中出现了两个 2, 但它的大小是不一样的, 小数点左边的 2 代表 2, 最左位上的 2 代表 200, 同样是数码 2, 但在不同的位置它具有不同的值, 我们称之为位权, 也称权重 (Weight)。为便于观察, 我们可以把该数展开, 即:

$$262.84 = 2 \times 10^2 + 6 \times 10^1 + 2 \times 10^0 + 8 \times 10^{-1} + 4 \times 10^{-2}$$

10^n 、 10^{n-1} …… 10^0 …… 10^{-m+1} 、 10^{-m} 就是我们所说的位权。可见, 位权是数码在该位置所具有的值。

1.2.2 二进制数制 (Binary Number System)

二进制使用 0、1 两个数码, 基数为 2, 计数规则为逢二进一。

二进制数的书写规则是将该数后面加 B 或在括号外加数字下标, 例如 $(1011.101)_2$ 或 1011.101B 都表示该数为二进制数。

与十进制相仿, 它的位权为 2 的整数幂。所以一个二进制数也可以将它展开, 展开后各项值的和是十进制表示的值。例如:

$$101101.11B = 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2}$$

二进制相加, 遵照逢二进一的规则, 如: $1011B + 101B = 10000B$

二进制数书写长, 不好读, 不好记。计算机中常用十六进制来对二进制进行“缩写”。

1.2.3 十六进制 (Hexadecimal Number System)

十六进制数基数为十六, 使用 0, 1, 2, 3……9, A, B, C, D, E, F 这十六个数码, 其规则为逢十六进一。

十六进制的书写方法为将该数后面加 H 或在括号外加数字下标, 例如 13D2H 或 $(13D2)_{16}$ 都表示该数为十六进制。

与十进制相仿, 它的位权为 16 的整数幂。一个十六进制数也可以将它展开, 展开后各项值之和是十进制表示的值。例如: $(13D8)_{16} = 1 \times 16^3 + 3 \times 16^2 + 13 \times 16^1 + 8 \times 16^0$

1.2.4 各种进制数之间的转换

1. 将二进制数转换为十进制数 将二进制数展开, 然后用十进制运算规则计算每一项的值再相加, 即可转换为十进制数, 例如:

$$\begin{aligned} (1011.101)_2 &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} \\ &= 8 + 0 + 2 + 1 + 0.5 + 0 + 0.125 = (11.625)_{10} \end{aligned}$$

2. 将十进制数转换为二进制数 可以将整数部分和小数部分分开, 整数部分采用除 2 取余逆排法; 小数部分采用乘 2 取整顺排法。例如: 将 $(13.375)_{10}$ 转换为二进制, 方法为:

整数部分:

小数部分:

2 | 13 余 ...1

2 | 6 余 ...0

2 | 3 余 ...1

2 | 1 余 ...1

0

2⁰

2¹

2²

2³

0.375

× 2

0.750

× 2

1.500

× 2

1.000

取整数 ...0

取整数 ...1

取整数 ...1

即: $(13.375)_{10} = (1101.011)_2$

3. 将二进制数转换为十六进制数 通过表 1.1 可以找出规律:一位十六进制数可以用四位二进制数表示,因此二进制的整数部分转换为十六进制时,只需从二进制数的小数点往左,每四位为一组,与一位十六进制数相对应,最后若不够四位,可以在其左端用 0 补齐。二进制小数部分转换为十六进制数时,则以小数点开始往右,每四位一组,最后若不够四位,在其右端用 0 补足。

例如:将二进制数 $(10110101101.10101)_2$ 转换为十六进制数,方法为:

0101 1010 1101 1010 1000

5 A D A 8

因此, $(10110101101.10101)_2 = (5AD.A8)_{16}$

表 1.1 十进制、二进制和十六进制关系对照表

十进制	二进制	十六进制
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F

4. 将十六进制数转换为二进制 将十六进制数的每位用四位二进制表示,转换方法是:对整数部分,小数点以左,每一位十六进制数用相应的四位二进制数表示,不足四位时,在其左端添“0”补足。小数部分则是从小数点开始往右,用四位二进制数表示一个十六进制数。例如:将 $(5A3B.AF)_{16}$ 转换为二进制数,其方法如下:

5 A 3 B • A F
0101 1010 0011 1011 • 1010 1111

因此: $(5A3B.AF)_{16} = (101101000111011.10101111)_2$

5. 十六进制和十进制数之间的转换：整数部分可以采用除十六取余法，小数部分采用乘 16 取整法，但较为复杂，建议通过二进制作为一个桥梁进行转换。

1.3 计算机中的数据信息的表示

1.3.1 符号数据的表示

计算机不仅可以处理数值数据，还可以处理字符、图形符号、汉字等，它们都是非数值数据。在计算机中，它们都有自己的代码，只不过用二进制代码表示，比如英文字母“A”的代码为 01000001，“B”的代码为 01000010，就像学生有学号一样。通常用 8 位二进制代码表示一个英文字符或控制符，我们把 8 位二进制代码称为一个字节，用 Byte 表示，也就是说，一个英文字符占一个字节，它就是我们以后要讲到的计算机存储器容量的单位，8 位二进制代码其中的某一位，不管它是 0 或是 1，我们把它记作 1 个信息单位，称为 1 个比特，用英文 Bit 表示。

在 8 位二进制代码中，最高位为 0 作为校验位，后 7 位用于字符编码，即 $b_7b_6b_5b_4b_3b_2b_1b_0$ ，其中 $b_7=0$ ，这样一共有 $2^7=128$ 种组合，用这 128 种不同状态的组合分别表示英文的 128 个字符，美国标准信息交换码就是采用的这种 7 位编码方案，简称 ASCII 码 (American Standard Code for Information Interchange)，表示每个字符的二进制代码称为该字符的 ASCII 值。英文字符的 ASCII 值见表 1.2。

表 1.2 英文字符 ASCII 表

$b_6 \ b_5 \ b_4$				000	001	010	011	100	101	110	111
b_3	b_2	b_1	b_0								
0	0	0	0	NUL	DLE	SP	0	@	P	,	p
0	0	0	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	STX	DC2	"	2	B	R	b	r
0	0	1	1	ETX	DC3	#	3	C	S	c	s
0	1	0	0	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	ACK	SYN	&	6	F	V	f	v
0	1	1	1	BEL	ETB	'	7	G	W	g	w
1	0	0	0	BS	CAN	(	8	H	X	h	x
1	0	0	1	HT	EM	)	9	I	Y	i	y
1	0	1	0	LF	SUB	*	:	J	Z	j	z
1	0	1	1	VT	ESC	+	;	K	[	k	{
1	1	0	0	FF	FS	<	<	L	\	l	
1	1	0	1	CR	GS	=	=	M	]	m	}
1	1	1	0	SO	RS	>	>	N	^	n	~
1	1	1	1	SI	US	/	?	O	_	o	DEL

在表 1.2 中上横栏为 ASCII 码的高四位，由于 b_7 为校验位，所以未标出，左面竖栏为低四位，ASCII 码可以用十或十六进制表示。比如：数字 0，从表中可以得到其 $b_6b_5b_4$ 为 011，而 $b_3b_2b_1b_0$ 为 0000，所以它的 ASCII 码为 0110000，它的机内码是在 ASCII 码最高位前加校验位 0 构成一个字节，即为 00110000，其十进制表示为 48。

1.3.2 汉字编码

英文字符数量相对较少而且它们本身有序,所以对它们进行数字化编码是较为容易的。而对汉字字符进行数字化编码难度要大得多,因为汉字既多又复杂,几千个汉字要对应几千个编码,除了这些,汉字的字形比其他国家的文字要复杂得多,不同的汉字有不同的形状,即便是同一汉字,又有宋体、楷体等多种字体,每个汉字之间又缺乏联在性,所以对汉字进行编码要考虑很多因素,比如汉字的排列顺序、汉字如何输入以及汉字字形如何在计算机中表示等。

1. 常用的汉字信息编码标准 在汉字信息编码标准中,常用的是简体中文 GB2312、GB18030 等,繁体中文 Big5 码等。

GB2312 码是中华人民共和国国家汉字信息交换用编码,全称《信息交换用汉字编码字符集基本集》,由国家标准总局发布,1981 年 5 月 1 日实施,简称国标码。通行于中国内地,新加坡等地也使用此编码。GB2312 收录简化汉字及符号、字母、日文假名等共 7445 个图形字符,其中汉字占 6763 个。GB2312 规定“对任意一个图形字符都采用两个字节表示,每个字节均采用七位编码表示”,习惯上称第一个字节为“高字节”,第二个字节为“低字节”。该字符集是几乎所有的中文系统和国际化的软件都支持的中文字符集,这也是最基本的中文字符集。

由于 GB2312-80 仅收汉字 6763 个,这大大少于现有汉字。随着计算机的广泛应用,国标 GB2312-80 已不能适应发展需要,为了解决这些问题,以及配合电脑业界组织的 UNICODE 的实施,全国信息化技术委员会于 1995 年 12 月 1 日发布了《汉字内码扩展规范》,之后信息产业部和国家质量技术监督局于 2000 年 3 月 17 日发布了两项新的国家标准:GB18030-2000 和 GB18031-2000。GB18030-2000《信息技术信息交换用汉字编码字符集基本集的扩充》(简称 GBK),共收录了 27 484 个汉字,具体规定了图形字符的单字节编码和双字节编码,并对四字节编码体系结构做出了规定。该标准是一个强制性标准。与现有的绝大多数操作系统、中文平台在计算机内码一级兼容,能够支持现有的应用系统。

Big5 是台湾地区的 IIIT1984 年发明的,CNS 11643-1992(Chinese National Standard)是其扩展版本。

Hong Kong GCCS (Government Chinese Character Set) 是香港政府在 Big5 基础上增加了 3049 个字符;之后又制定了 HKSCS 标准,它包括了 Big5 和 ISO10646 的编码,也可以说 HKSCS 是 GCCS 的增强版。

一般来讲,计算机内汉字编码中包括机内码、输入码和汉字输出码。

2. 机内码 以国标码 GB2312 码为例,说明汉字信息编码与实际存储在机器中的机内码之间的关系。

在计算机内,如果直接采用国标码,势必会造成与 ASCII 码混淆,例如:

汉字“大”的国标码为 00110100 01110011,而数字“4”和“s”的 ASCII 码分别为 00110100、01110011,如果不加以指定,计算机会把 0011010001110011 当成两个英文字符 4s 来处理。鉴于以上情况,必须把国标码变成机内码才可以让计算机处理,方法是将每个字节的最高位置 1,只要每个字节的最高位为 1 即为汉字,这样就构成了汉字机内

码。无论是国标码还是机内码，书写时都可用 16 进制。还是以汉字的“大”为例，它的国标码、机内码和 ASCII 码有如下对应关系：

名称	编码（十六进制）	编码（二进制）	
国标码	3473	00110100 01110011	
机内码	B4F3	10110100 11110011	
ASCII 码	3473	00110100 01110011	代表英文“4s”

由此可见，如果用十六进制表示，将国标码转为机内码的方法只需将国标码加 8080 即可。即：

$$\begin{array}{r}
 3473 \quad \text{国标码} \\
 +8080 \\
 \hline
 B4F3 \quad \text{机内码}
 \end{array}$$

3. 输入码 由于汉字的独立性，使汉字的输入变得较为复杂，常用的汉字输入法基本分为两大类：

（1）编码汉字输入：编码汉字输入现基本分三类，以音为主的拼音输入和以形为主的笔形输入以及音形结合的输入方法。它们各有特色，拼音易学好记，但相同读音的汉字太多，即重码率高，检字困难，典型的拼音输入法有全拼输入法和微软拼音输入法。由于相同形状的汉字很少，所以笔形输入重码率低，但掌握困难，典型的笔形输入是五笔输入编码。无论采用哪一种，它们都称为输入码，也称外码。当向计算机输入外码时，一般都要转换成机内码后才能进行存储和处理，当然这是各种汉字操作系统所要解决的问题，使用者只需输入汉字的外码，剩下的由计算机自己处理。

（2）非编码汉字输入：近年人们发明了不少用于汉字输入的设备，如手写板输入和语音输入。手写板输入需要一特定的硬件（手写板），用户可以将汉字手工写在手写板上，计算机自动将手写体识别成可编辑的文本，这种方法只要会写汉字就可以向计算机输入汉字，如果所写汉字不是很草，识别率还是很高的。语音输入是指用麦克风按正常的说话速度朗读，计算机通过声卡和识别软件，将语音自动识别成可编辑的文本，但这种方法存在个体差异，即由于每个人的声调、发音均不同，会造成识别错误，因此要对计算机进行训练，让计算机逐渐能够适应你的发音，才能有较高的识别率。除上述两种非编码输入之外，用扫描仪将书报上的文稿以图像的形式扫到计算机中，再通过识别软件（OCR）进行识别，还原成可编辑的文本，如果原稿比较清楚，其识别率可达 90% 以上。

4. 汉字输出码 汉字输出码是地址码、字形存储码和字形码的统称。

地址码：是指汉字字形信息在汉字字模库中存放的首地址。每个汉字在字库中占有一个固定大小的连续区域，其中首地址即是该汉字的地址码。

字形存储码：是指存放在字库中的汉字字形点阵码。不同的字体有不同的字库，如黑体、仿宋体等，点阵的点数越多字的质量越高，越美观。

由于汉字都是方块字，每个汉字看作是一个有 M 行 N 列点组成的矩阵，称为汉字的点阵字模，简称点阵。如果用二进制数 1 代表点阵中的黑点，用 0 表示无黑点。一个汉字若用 16×16 点阵表示，则共有 256 个点，如图 1.1 所示：

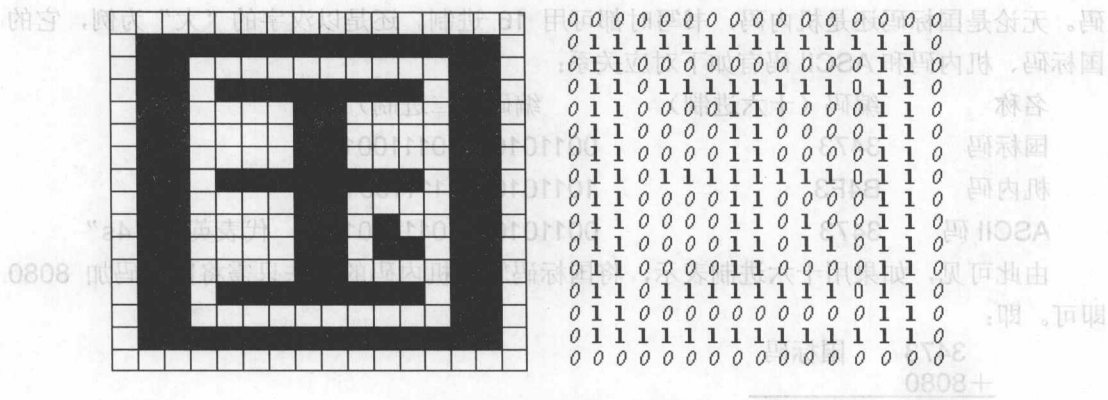


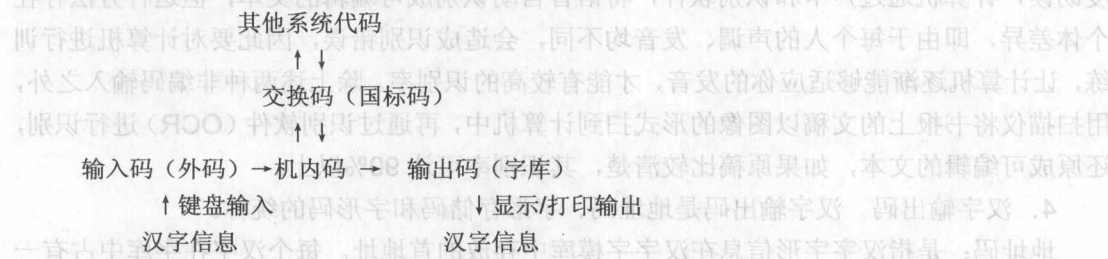
图 1.1 汉字字形码示意图

由图 1.1 可以看出，一个汉字可以用 16 行二进制代码表示，一行为 16 位，正好为 2 个字节，所以一个 16 点阵字库要占 $16 \times 2 = 32$ 个字节。对于 16 点阵字库要用于打印，其质量可见而知，所以 16 点阵字库主要用于显示，真正用于打印，应采用 24 点阵以上的字库，而 24 点阵字库每个汉字要占 $24 \times 3 = 72$ 个字节。

汉字字形库直接存储点阵码时占用的存储空间大，为了减少字库所占的容量，采用了数据压缩技术。使用较多的字库压缩方法有哈夫曼树法、矢量法和字根压缩法。近年来开发的新的汉字操作系统中常使用矢量汉字。所谓矢量汉字就是经过矢量法把基本点阵字模进行压缩后得到的汉字。这些汉字信息存在矢量字库中，显示和打印时要经过相应的转换程序进行还原和变换，得到不同的字体。

字形码：指在输出设备上输出汉字时所送出的汉字字形点阵码。点阵数据的组织是按照输出设备的特性及输出字体的一些特点（如倾斜角度，放大倍数）进行的，是对基本字库中数据进行变换得到的。

以上所介绍的各种汉字编码之间的关系为：



1.4 计算机的组成

一个完整的计算机系统是由硬件系统和软件系统两大部分组成。所谓硬件就是构成计算机实体的所有器件，而软件是指那些看不到、摸不着却又实实在在存在的那些计算机中存储的数据、程序等等。软件系统是计算机的灵魂，以硬件系统为依托，对硬件设备进行控制和管理。只有硬件、软件系统相互结合，才能发挥计算机系统的强大功能。

1.4.1 计算机硬件

从原理上讲,计算机是由输入设备、输出设备、存储器、控制器和运算器五大部分组成,如图 1.2 所示。下面简单地讲述各部分的功能。

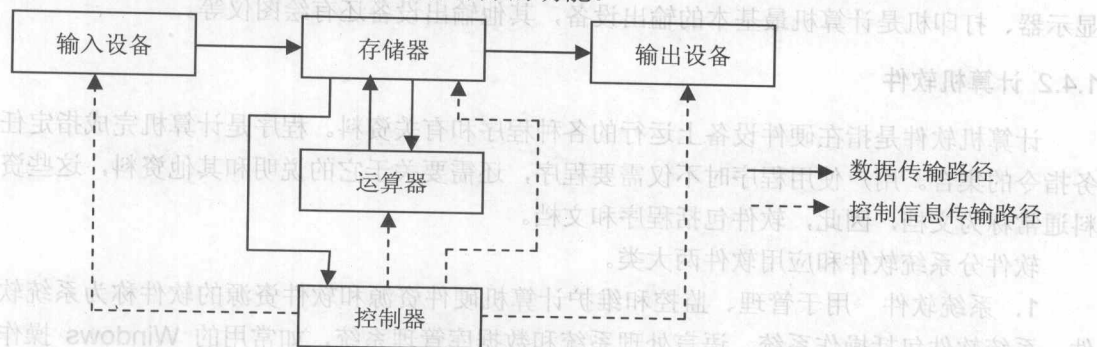


图 1.2 计算机硬件的基本组成

1. 运算器 (Calculator) 运算器也称为算术逻辑运算单元 (Arithmetic and Logic Unit, ALU), 计算机中所有的算术运算、逻辑运算和信息传送都在这里进行, 它由加法器、移位电路、逻辑部件、信息传送部件以及寄存器等电路组成。由于任何数学运算最终可以用加法和移位这两种基本操作来完成, 因而加法器是 ALU 的核心部件。寄存器则用来暂时存放参与运算的操作数和运算结果。

2. 控制器 (Controller) 控制器也称控制电路 (Control Circuit), 它就像一个部队的指挥部, 是整个计算机的控制中心, 其任务是按预定的顺序不断取出指令进行分析, 然后根据指令的要求向运算器、存储器等各部分发出控制信号让其完成指令所规定的操作, 指令是一条命令, 就是让计算机做什么, 是由一串二进制代码组成, 不同的代码表示不同的命令, 是给计算机约定好的。控制电路由指令计数器、指令寄存器、指令译码器和操作控制部件等组成。指令计数器用于提供指令的存放地址。指令寄存器可以把从存储器取出的指令暂存起来。指令译码器是把取到的指令译成操作控制部分所能识别的信号, 使其完成指令所规定的操作。

3. 存储器 (Memory) 存储器是计算机存放数据的地方, 它由一片片连续的存储单元组成, 每个单元都赋有编号, 称为地址, 就像楼房的房间号一样, 每个单元都可以存放一组二进制代码, 就像房间里住的是谁。信息存入内存的过程称为写入, 取出的过程称为读出。存储器的基本指标是容量和读写速度。存储器分内部存储器和外部存储器, 内部存储器可以由 CPU 直接访问, 内存的读写速度快但其存储空间是有限的, 外部存储器作为内部存储器的扩展存储, 存储容量大但读写速度相对较慢。存储器的计量单位是字节 (Byte), 由于计算机存储器容量很大, 所以用字节很不方便, 一般用千字节 (KB)、兆字节 (MB) 等表示, 其换算如下:

$$1\text{KB}=1024\text{Byte} \left(2^{10}\right)$$

$$1\text{MB}=1024\text{KB} \left(2^{20}\right)$$

$$1\text{GB}=1024\text{MB} \left(2^{30}\right)$$

$$1\text{TB}=1024\text{GB} \left(2^{40}\right)$$