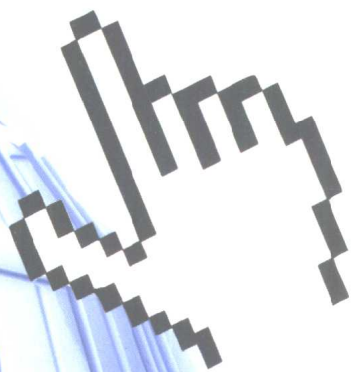


中国计算机软件专业技术资格和水平考试指定用书

信息产业部计算机软件专业技术资格和水平考试办公室 组编

王春森 主编

程序员教程



清华大学出版社

<http://www.tup.tsinghua.edu.cn>

中国计算机软件专业技术资格和水平考试指定用书
信息产业部计算机软件专业技术资格和水平考试办公室 组编

程序员教程

王春森 主编

NO. 5100

清华大学出版社

(京)新登字 158 号

内 容 简 介

本书是信息产业部计算机软件专业技术资格和水平考试办公室组织编写的考试指定用书。本书遵循考试大纲的要求,对程序员考试应必备的知识和技能的重点作了阐述。

本书内容包括:计算机硬件、程序语言、操作系统、软件工程、数据库、多媒体和网络等方面的基础知识以及与程序编制有关的数据结构知识和常用算法设计方法。

本书是程序员考试应试者的必读教材,也可作为各类计算机信息技术培训和辅导的教材,还可作为大专院校师生和广大工程技术人员学习计算机信息技术基础知识的参考书。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

书 名: 程序员教程

作 者: 王春森 主编

出版者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

[http:// www. tup. tsinghua. edu. cn](http://www.tup.tsinghua.edu.cn)

印刷者: 北京市密云胶印厂

发行者: 新华书店总店北京发行所

开 本: 787×960 1/16 **印张:** 27.75 **字数:** 556 千字

版 次: 2001 年 5 月第 1 版 2001 年 6 月第 2 次印刷

书 号: ISBN 7-302-04371-X/TP·2567

印 数: 30001~60000

定 价: 38.00 元

序

人类已跨入新世纪,正在进入信息时代。我国国民经济和社会发展第十个五年计划将国民经济和社会信息化作为覆盖现代化建设全局的战略举措,要求加速发展信息产业,大力推进信息化。现在,信息技术(IT)的应用越来越普及,不但促进了社会的高速发展,也改变着人们的工作、学习、生活 and 娱乐的方式以及思想观念。各行各业对信息技术专业人才的需求也迅速增长。尤其是计算机软件和计算机网络人才,出现了严重的短缺。国务院颁布了鼓励软件产业发展的若干政策,要求进一步扩大软件人才的培养规模。为此,国家人事部和信息产业部决定进一步发展中国计算机软件专业技术资格和水平考试。

过去十年来,中国计算机软件专业技术资格和水平考试培养和选拔了十多万合格的计算机软件专业人才,在国内外产生了很大的影响,对我国软件产业的形成和发展做出了重要的贡献。根据形势发展的需要,2001年考试的级别拓展为:初级程序员、程序员、系统设计师(高级程序员)、系统分析员、网络程序员和网络设计师,以后还将逐步拓展到信息技术领域的其他方面。

为了规范培训和考试工作,我们组织有关专家编写了中国计算机软件专业技术资格和水平考试的指定教材和辅导用书。这套丛书将遵循考试大纲的要求,全面介绍有关的知识和技能,帮助考生学习和备考。

我们相信,经过大家的努力,中国计算机软件专业技术资格和水平考试将会成为我国信息技术领域专业水平的重要考试,将对培养大批信息技术专业人才,推进国民经济和社会信息化做出更大的贡献。

信息产业部计算机软件专业技术资格和水平考试办公室

2001年3月1日

前 言

中国计算机软件专业技术资格和水平考试实施至今已经历了十余年,在社会上产生了很大影响,对我国软件产业的形成和发展做出了重要的贡献。为了适应我国计算机信息技术发展的需求,国家人事部和信息产业部决定将考试的级别逐步拓展到计算机信息技术行业的各个方面,以满足社会上对各种计算机信息技术人才的需要。

编者受信息产业部计算机软件专业技术资格和水平考试办公室委托,在《程序设计(程序员级)》一书的基础上编写成本书,以适应新的考试大纲要求。在考试大纲中,要求考生掌握的知识面很广,一个条目在大学里可能是一学期的课程,因此编写的难度很高。考虑到参加考试的人员已有一定的基础,所以书中只对考试大纲中所涉及到的知识领域的要点加以阐述,限于篇幅不能详细地展开。再考虑到参加考试的人员一般都已熟悉一种以上的程序设计语言和已具有基本的编程能力,因此本书以数据结构和基本算法为主线,使读者在原有的程序设计知识和编程能力方面得以总结、拓宽和提高,并增强程序的编制和阅读能力,掌握常用数据结构和算法的原理和要点。

全书共分9章。第1章计算机硬件基础知识由丛锁、高传善编写;第2章程序语言知识由张守志编写;第3章操作系统基础知识由于玉编写;第4章软件工程基础知识由钱乐秋编写;第5章数据库基础知识由丁宝康编写;第6章多媒体基础知识由白雪峰编写;第7章网络基础知识由曹邦伟编写;第8章数据结构基础和第9章常用算法设计方法由夏宽理编写。最后,由王春森统稿。

在本书的编写过程中,参考了许多相关的书籍和资料,编者在此对这些参考文献的作者表示感谢。同时感谢清华大学出版社在本书出版过程中所给予的支持和帮助。

因水平有限,书中难免存在错漏和不妥之处,望请读者指正,以利改进和提高。

编 者

2001年3月于复旦大学

目 录

第1章 计算机硬件基础知识	1	2.3 解释程序基础知识	79
1.1 计算机系统概述	1	2.4 编译程序基础知识	81
1.1.1 计算机史简介	1	2.4.1 编译过程概述	81
1.1.2 计算机系统的结构	2	2.4.2 文法和语言的形式描述	84
1.2 数据的计算机表示	3	2.4.3 词法分析	88
1.2.1 数制	3	2.4.4 语法分析	93
1.2.2 数值数据的表示方法	7	2.4.5 语法制导翻译和中间代码 生成	100
1.2.3 符号数据的表示方法	13	2.4.6 代码优化	105
1.2.4 汉字的处理	14	2.4.7 代码生成	109
1.2.5 数据的校验方法	15	第3章 操作系统基础知识	112
1.3 计算机算术运算和逻辑运算	22	3.1 概述	112
1.3.1 二进制数的算术运算	22	3.1.1 引言	112
1.3.2 逻辑代数及逻辑运算	25	3.1.2 什么是操作系统	113
1.4 计算机基本结构	28	3.1.3 操作系统的功能	115
1.4.1 总线结构	28	3.1.4 操作系统的工作原理	116
1.4.2 中央处理器	30	3.1.5 操作系统的硬件基础	117
1.4.3 存储器系统	42	3.2 并发进程	118
1.4.4 I/O 系统	47	3.2.1 程序、进程和处理器	118
1.5 指令系统	53	3.2.2 进程间的通信	118
1.5.1 机器指令的格式	53	3.2.3 信号量与 P,V 操作	119
1.5.2 指令的种类和功能	56	3.2.4 高级通信原语	120
1.5.3 指令的寻址方式	59	3.3 系统核心	121
1.5.4 指令的执行过程	60	3.3.1 基本硬件机构	121
第2章 程序语言知识	61	3.3.2 进程	121
2.1 程序语言基础知识	61	3.3.3 进程管理	122
2.1.1 概述	61	3.4 存储管理	124
2.1.2 程序语言的数据类型	64	3.4.1 引言	124
2.1.3 程序语言的控制结构	71	3.4.2 实存管理	125
2.2 汇编程序基础知识	76	3.4.3 虚存组织	128
2.2.1 汇编语言	76	3.4.4 虚存管理	130
2.2.2 汇编程序	77		

3.5	设备管理	131	4.4.2	结构化设计方法简介 ...	160
3.5.1	引言	131	4.4.3	详细设计	161
3.5.2	设备的分配	133	4.5	编码	166
3.5.3	磁盘调度算法	133	4.5.1	程序设计语言	166
3.5.4	RAID	134	4.5.2	对源程序的质量要求 ...	168
3.6	文件管理	134	4.5.3	编码风格	169
3.6.1	引言	134	4.6	软件测试	171
3.6.2	文件的结构和组织	135	4.6.1	测试的基本概念	171
3.6.3	文件访问方法	137	4.6.2	测试步骤	173
3.6.4	文件存储设备管理	138	4.6.3	白盒测试的测试用例 设计	174
3.6.5	文件控制块和文件目录 ...	139	4.6.4	黑盒测试的测试用例 设计简介	177
3.6.6	文件的使用	140	4.7	面向对象方法的基本概念	178
3.7	作业管理和用户界面	141	4.8	软件维护	181
3.7.1	引言	141	4.8.1	软件维护的分类	181
3.7.2	用户作业管理	141	4.8.2	与软件维护有关的问题 ...	182
3.7.3	作业调度	142	4.8.3	可维护性	183
3.7.4	用户界面	143	4.8.4	软件维护活动流程	183
3.8	其他管理	143	4.9	软件管理	184
3.8.1	死锁问题	143	4.10	软件质量保证	187
3.8.2	多处理器系统与线程 ...	144	4.10.1	软件质量特性	187
3.9	实例	144	4.10.2	软件质量保证概述 ...	190
3.9.1	UNIX 系统	144			
3.9.2	Windows NT 系统	148	第 5 章	数据库基础知识	192
第 4 章	软件工程基础知识	151	5.1	数据管理技术的发展	192
4.1	软件工程概述	151	5.1.1	数据管理技术的发展 阶段	192
4.1.1	软件生存周期	151	5.1.2	数据库技术的基本概念 ...	194
4.1.2	软件开发模型	153	5.1.3	数据库系统的进一步 发展	195
4.2	系统分析和软件项目计划	154	5.2	数据描述	195
4.2.1	系统分析	154	5.2.1	数据的描述	195
4.2.2	可行性分析	154	5.2.2	数据联系的描述	197
4.2.3	编制项目开发计划	155	5.3	数据模型	198
4.3	需求分析	155	5.3.1	数据模型的定义	198
4.3.1	需求分析的任务	155	5.3.2	概念数据模型	199
4.3.2	结构化分析方法简介 ...	156			
4.4	软件设计	157			
4.4.1	软件设计的概念和原则 ...	157			

5.3.3 基本数据模型	200	6.2.2 图形、图像格式	249
5.4 数据库系统的结构	203	6.3 声音(音频)	251
5.4.1 数据库的数据体系结构...	203	6.3.1 基本概念	252
5.4.2 数据库系统	205	6.3.2 语音合成与语言合成 ...	253
5.4.3 数据库管理系统	207	6.3.3 音乐合成和 MIDI	254
5.4.4 用户访问数据的过程 ...	208	6.4 视频(动画)	257
5.4.5 数据库系统的全局结构...	210	6.4.1 基本概念	258
5.5 关系模型和关系运算	212	6.4.2 视频格式	259
5.5.1 关系模型的基本概念 ...	212	6.5 多媒体创作工具的发展与方向...	259
5.5.2 关系运算	213	6.5.1 多媒体创作工具发展	
5.5.3 关系数据库 DML 的		概述	259
特点	219	6.5.2 多媒体创作工具技术	
5.6 关系数据库 SQL 语言简介	219	发展	260
5.6.1 SQL 数据库的数据体系		第 7 章 网络基础知识	264
结构	220	7.1 网络发展简史	264
5.6.2 SQL 语言的组成	220	7.2 网络的主要应用与网络的分类 ...	266
5.6.3 SQL DDL	221	7.2.1 计算机网络的功能	266
5.6.4 SQL DML 的数据查询		7.2.2 若干典型的计算机网络	
语句	223	应用系统	267
5.6.5 SQL DML 的数据更新		7.2.3 计算机网络的分类	268
语句	230	7.3 网络协议和标准	270
5.6.6 SQL 的访问控制	231	7.3.1 网络的标准化	271
5.6.7 嵌入式 SQL	232	7.3.2 网络协议模型	272
5.7 数据库设计过程	232	7.3.3 ISO/OSI 网络协议层次	
5.7.1 数据库工程	233	模型	274
5.7.2 数据库系统生存期	233	7.3.4 TCP/IP 协议组	274
5.7.3 数据库设计的具体		7.4 常用的网络操作系统	275
过程	233	7.4.1 网络操作系统的基本	
第 6 章 多媒体基础知识	240	功能	275
6.1 多媒体的概念和特征	240	7.4.2 Novell Netware	276
6.1.1 什么是多媒体	240	7.4.3 Microsoft Windows NT	
6.1.2 多媒体的关键技术	240	Server	277
6.1.3 多媒体系统的构成	242	7.4.4 UNIX	278
6.1.4 声卡和显卡	244	7.5 构建 LAN 网络	279
6.2 图像和图形	247	7.5.1 LAN 涉及的主要网络	
6.2.1 基本概念	247	专用设备	279

7.5.2	网络采用的主要传输媒体	280	7.9.3	防火墙的基本概念	300
7.5.3	LAN 的协议标准和主要的媒体访问控制方式 ..	281	第 8 章	数据结构基础	301
7.5.4	LAN 的硬件构建示例 ..	283	8.1	线性表	301
7.6	构建 WAN 网络	284	8.2	栈	311
7.6.1	通信服务和网络接入技术	285	8.3	队列	322
7.6.2	两类重要的网络互联设备	286	8.4	数组	328
7.6.3	WAN 的几种典型连接方式	288	8.4.1	数组的基本概念	328
7.7	Internet 的应用	290	8.4.2	稀疏矩阵	338
7.7.1	Internet 的服务	290	8.5	字符串	343
7.7.2	IP 地址和域名	292	8.6	树和二叉树	351
7.7.3	Intranet 初步	293	8.6.1	树	352
7.8	网络应用的主要方式	294	8.6.2	二叉树	356
7.8.1	文件服务器模型与域模型	294	8.6.3	二叉查找树	359
7.8.2	客户机/服务器使用方式	295	8.6.4	穿线树	363
7.8.3	网络计算初步	296	8.6.5	树形数据结构程序设计实例	367
7.8.4	电子商务介绍	297	8.7	排序	371
7.9	网络的安全性初步	298	8.8	查找	386
7.9.1	网络安全性的基本概念 ..	298	第 9 章	常用算法设计方法	397
7.9.2	网络的信息安全技术初步	298	9.1	迭代法	397
			9.2	穷举搜索法	398
			9.3	递推法	401
			9.4	递归	403
			9.5	回溯法	414
			9.6	贪婪法	424
			9.7	分治法	430
			9.8	动态规划法	432

第 1 章 计算机硬件基础知识

1.1 计算机系统概述

1.1.1 计算机史简介

自从 1946 年 2 月现代电子计算机的鼻祖 ENIAC(electronic numerical integrator and computer)在美国宾夕法尼亚大学问世以后,短短 50 年里,计算机技术经历了巨大的变革。习惯上根据计算机系统所采用的硬件技术来划分计算机的发展阶段。

从 1946 年到 50 年代后期(1946—1957)为电子管计算机时期。计算机的元器件主要由电子管(vacuum tube)组成。其特点是体积庞大、功耗高、运算速度较低。如 ENIAC 占地 170m²,重达 30t,功耗为 140kW,有 18000 多个电子管,每秒钟能进行 5000 次加法计算。这一阶段,计算机主要用于军事、国防等尖端技术领域。除了 ENIAC 以外,1945 年左右,冯·诺依曼等人在研制 EDVAC(electronic discrete variable computer)时,提出了存储程序(stored-program)概念,奠定了以后计算机发展的基石。IBM 公司 1954 年 12 月推出的 IBM650 是第一代计算机的代表。

从 20 世纪 50 年代后期到 60 年代中期(1958—1964)为晶体管计算机时期。自从 1947 年晶体管(transistor)在贝尔实验室诞生后,引发了一场影响深远的电子革命。体积小、功耗低、价格便宜的晶体管取代了电子管,不仅提高了计算机的性能,也使计算机在科研、商业等领域内广泛地应用。第二代计算机不仅采用了晶体管器件,而且存储器改用速度更快的磁芯存储器;与此同时高级编程语言和系统软件的出现,也大大提高了计算机的性能和拓宽了其应用领域。这一时期计算机的代表主要有 DEC 公司 1957 年推出的 PDP-I,IBM 公司于 1962 年推出的 7094 以及 CDC 公司 1964 年研制成功的 CDC6600。1969 年 CDC 公司研制的 CDC7600 平均速度达到每秒千万次浮点运算。

从 20 世纪 60 年代中期到 70 年代初期(1965—1971)为集成电路计算机时代。第一代和第二代计算机均采用分离器件(discrete component)组成。集成电路(integrated circuit)的出现,宣告了第三代计算机的来临。由于采用了集成电路,使得计算机的制造成本迅速下降;同时因为逻辑和存储器件集成化的封装,大大提高了运行速度,功耗也随之下降;集成电路的使用,使得计算机内各部分的互联更加简单和可靠,计算机的体积也进一步缩小。这一时期的代表为 IBM 的 system/360 和 DEC 的 PDP-8。

从 20 世纪 70 年代初期到 70 年代后期(1972—1978)为大规模集成电路(LSI)计算机时代。20 世纪 70 年代初半导体存储器的出现,迅速取代了磁芯存储器,计算机的存储器向大容量、高速度的方向飞速发展。存储器芯片从 1kbit,4kbit,16kbit,64kbit,256kbit,1Mbit,4Mbit 发展到 16Mbit(1992 年)。

接着就进入了超大规模集成电路(VLSI)计算机时代。随着技术的日新月异,软件和通信的重要性也逐步上升,成为和硬件一样举足轻重的因素。同时系统结构的特点对计算机的性能也有巨大的影响(中断系统、Cache 存储器、流水线技术等等)。实际上在第三代计算机以后,就很难找到一个统一的标准进行划分。

目前更流行的是认为计算机的发展经过了三次浪潮(wave)。第一次浪潮为单个主机(Mainframe)的时期,IBM360、370 是其典型代表。第二次浪潮为客户机/服务器(Client/Server)的时期,这时期出现了小型机、微型机和局域网。现在正处于第三次浪潮,网络计算的时期,即以网络为中心或以网络为基础的计算时期。

1.1.2 计算机系统的结构

计算机的工作过程一般是由用户(程序员)使用各种编程语言把所需要完成的任务以程序的形式提交给计算机,然后翻译成计算机能直接执行的机器语言程序,在计算机上运行以实现用户需要的功能。

计算机系统可以看作是一个多层的系统,不同的用户(或者应用)与不同的层次打交道。通常把计算机系统分作 4 层:物理机器(机器语言机器)、操作系统语言机器、汇编语言机器和高级语言机器,如图 1.1 所示。第 1 层实际上就是由硬件组成的计算机本身,它所能理解和执行的是机器语言。第 2~4 层其实都是“虚拟”机器,通过软件组成的层次结构,隐藏了机器的硬件实现细节。用户在其上见到的“机器”实际上是在物理机器上运行的软件表示的虚拟机器。

1945 年冯·诺依曼在设计 EDVAC 时提出了存储程序的概念,后来他和同事合作设计的存储程序计算机 IAS 的结构被其后几乎所有的通用计算机采用。这种结构也就称

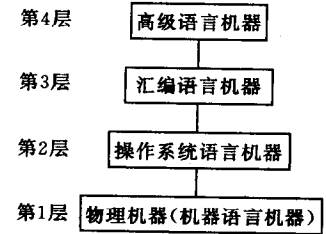


图 1.1 计算机系统的层次结构

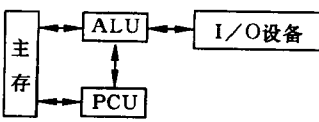


图 1.2 IAS 计算机结构

为冯·诺依曼机,如图 1.2 所示。其中主存(main memory)用以存放数据和程序;算术逻辑运算单元 ALU(arithmetic logical unit)完成二进制算术和逻辑运算;程序控制单元 PCU(program control unit)负责解释和执行指令;I/O(input/output)设备由 PCU 控制,实现和外部环境的数据交换。

1.2 数据的计算机表示

1.2.1 数制

计算机的存储器(memory)和寄存器(register)是两态器件,因此,各种信息是以二进制的方式在计算机里存储和运算。但是为了方便起见,有时也需采用不同的数制(number system)来表示数据,例如十进制、十六进制、八进制等。

1. 数制介绍

一种数制由基数 r (radix) 和 r 个不同的数码(symbol)组成。任意一个数用 r 进制可以表示为

$$\begin{aligned} N &= (R_n R_{n-1} \cdots R_1 R_0 R_{-1} \cdots R_{-m})_r \\ &= R_n \times r^n + R_{n-1} \times r^{n-1} + \cdots + R_1 \times r + R_0 + R_{-1} \times r^{-1} + \cdots + R_{-m} \times r^{-m} \end{aligned}$$

对于十进制数而言共有 10 个数码:0,1,2,3,4,5,6,7,8,9,十进制数 538.4 即表示 $5 \times 10^2 + 3 \times 10^1 + 8 \times 10^0 + 4 \times 10^{-1}$ 。对于八进制数而言 8 个数码为 0~7;而十六进制的 16 个数码分别为 0~9 和 A~F。二进制的基数为 2,数码为 0 和 1。

上式称为数 N 在 R 进制中的按权(power)展开式。

【例 1.1】 把八进制数 1704.25 转化为十进制数。

解:

$$\begin{aligned} (1704.25)_8 &= 1 \times 8^3 + 7 \times 8^2 + 0 \times 8^1 + 4 \times 8^0 + 2 \times 8^{-1} + 5 \times 8^{-2} \\ &= 1 \times 512 + 7 \times 64 + 4 \times 1 + 2 \times 0.125 + 5 \times 0.015625 \\ &= (964.328125)_{10} \end{aligned}$$

【例 1.2】 把十六进制数 2EC.F 转化为十进制数。

解:

$$\begin{aligned} (2EC.F)_{16} &= 2 \times 16^2 + 14 \times 16^1 + 12 \times 16^0 + 15 \times 16^{-1} \\ &= 2 \times 256 + 14 \times 16 + 12 \times 1 + 15 \times 0.625 \\ &= (748.9375)_{10} \end{aligned}$$

表 1.1 列出了常用的十、十六、八和二进制数的对照表。

表 1.1 常用进制对照表

十进制	十六进制	八进制	二进制	十进制	十六进制	八进制	二进制
0	0	0	0	9	9	11	1001
1	1	1	1	10	A	12	1010
2	2	2	10	11	B	13	1011
3	3	3	11	12	C	14	1100
4	4	4	100	13	D	15	1101
5	5	5	101	14	E	16	1110
6	6	6	110	15	F	17	1111
7	7	7	111	16	10	20	10000
8	8	10	1000				

2. 数制间的转换

(1) 从十进制向 r 进制转换

把一个十进制数转换为等价的 r 进制数应分成两步进行: 首先把该数的整数部分和小数部分分别转换为对应的 r 进制数; 然后再把这两部分合起来即可。

算法 1: 除 r 取余(把十进制整数 I 转换为对应的 r 进制数)

- ① $k := 1$
- ② 把 I 除以 r 得到商 Q 和余数 R_k
- ③ 若 $Q=0$ 则转到④, 否则 $I := Q; k := k+1$, 并转到②
- ④ 把 R_k 按 k 从大到小排列起来就是所求结果

【例 1.3】 把 $(307)_{10}$ 转换为对应的二进制数。

解:

商(Q).....	余数(R_k)
307	
153	1
76	1
38	0
19	0
9	1
4	1
2	0
1	0
0	1

$$\begin{array}{r} \times 2 \\ 1 \quad .0 \end{array}$$

$$(0.8125)_{10} = (0.1101)_2$$

【例 1.7】 把 $(0.8125)_{10}$ 转换为十六进制数和八进制数。

解：

$$\begin{array}{r} 0.8125 \\ \times 16 \\ D \quad .0 \end{array}$$

$$\begin{array}{r} 0.8125 \\ \times 8 \\ 6 \quad .50 \\ \times 8 \\ 4 \quad .0 \end{array}$$

$$(0.8125)_{10} = (0.D)_{16} = (0.64)_8$$

(2) 二进制数 / 八进制数 / 十六进制数之间的相互转换

因为 $2^3=8$, $2^4=16$, 所以每一位八进制数码都对应着一个 3 位二进制数, 每一位十六进制数码则对应着一个 4 位二进制数, 如表 1.2 和表 1.3 所示。

表 1.2 八进制数和二进制数对应表

0	1	2	3	4	5	6	7
000	001	010	011	100	101	110	111

表 1.3 十六进制数和二进制数对应表

0	1	2	3	4	5	6	7
0000	0001	0010	0011	0100	0101	0110	0111
8	9	A	B	C	D	E	F
1000	1001	1010	1011	1100	1101	1110	1111

二进制向八进制转换的方法为:从小数点开始,分别向左、右每 3 位二进制数编成一组,若不够 3 位,则小数点左侧的最高位和右侧的最低位用 0 补充;每一组用对应的八进制数码表示即可。八进制向二进制转换的方法为:从小数点开始,把每一位八进制数码转换成对应的 3 位二进制数即可。其小数点左侧的最高位或右侧的最低位的 0(如果有的话)可以省去。二进制和十六进制数之间的转换与此类似,只不过每组的二进制位数是 4 而不是 3。

【例 1.8】 把 $(273.64)_8$ 转换为二进制数。

解:

$$\begin{array}{cccccc} 2 & 7 & 3 & . & 6 & 4 \\ 010 & 111 & 011 & . & 110 & 100 \\ (273.64)_{10} = (10111011.1101)_2 \end{array}$$

【例 1.9】把(7CE2.9A)₁₆转换为二进制数。

解:

$$\begin{array}{cccccc} 7 & C & E & 2 & . & 9 & A \\ 0111 & 1100 & 1110 & 0010 & . & 1001 & 1010 \\ (7CE2.9A)_{16} = (111110011100010.1001101)_2 \end{array}$$

在计算机中八进制数、十六进制数通常被认为是二进制数的另一种表示方法。

(3) 特殊方法

① 由于八进制与二进制之间的转换十分简单,因此有时十进制与二进制之间的转换可以以八进制为中介。

$$(273.71875)_{10} = (421.56)_8 = (100010001.10111)_2$$

② 如果一个分数的分母为 2 的整数次幂则可以用以下方法转换。

$$(3/32)_{10} = 3 \times 2^{-5} = (11)_2 \times (0.00001)_2 = (0.00011)_2$$

③ 从十进制向八进制/十六进制转换可以先把十进制转换为二进制,然后再从二进制转换为八进制/十六进制。

$$(27.625)_{10} = (11011.101)_2 = (33.5)_8 = (1B.A)_{16}$$

1.2.2 数值数据的表示方法

因为当前的计算机所采用的存储器件都是两态器件,所以适合于存放二进制数据。尽管为了方便起见,也使用其他进制,如十进制、十六进制,但是在计算机内部数据还是以二进制的形式存放和处理。二进制只使用了两个不同的数字符号,易于用物理器件来实现:在物理世界中具有两个稳定状态的物理器件很多,如晶体管的“截止”与“导通”、电容的“充电”与“放电”、电压信号的“高”与“低”、脉冲的“有”与“无”,电磁单元的“正向磁化”与“反向磁化”等等,只要规定其中一种稳定状态表示“1”,另一种稳定状态表示“0”,就可以用来表示二进制数位了。其次,二进制的运算规则非常简单,易于用电子器件来实现。

计算机里用到的数据类型主要可分为两类:表示数量的数值数据和非数值性的符号数据。所有的数据都是以二进制的形式在计算机里处理和存储。本节主要讲述数值数据在计算机里的表示方法。

1. 符号位的处理

正数和 0 都可以表示为无符号数,而负数通常是用“—”号来表示。但在计算机里,任

何符号都是用“1”和“0”来表示的。所以,习惯上计算机里都用最高位(最左位)来表示数的符号:0——正数,1——负数。

例如, $X_1 = (+0010010)_2$ 和 $X_2 = (-0010010)_2$ 在计算机里分别表示为 00010010 和 10010010。

2. 小数点的处理

计算机在处理数值数据时,对小数点的处理有两种不同的方法:定点法和浮点法,也就对应了两种不同形式的数据表示方法。

(1) 定点数据表示方法

在这种表示方法中,小数点的位置固定不变。根据其所在位置的不同,分别用来表示整数和纯小数:

- 定点整数格式: $N = b_s b_m b_{m-1} \cdots b_1 b_0$

小数点位置在最低数据位 b_0 的右边, b_s 为符号位, b_m 为其最高数据位,其取值范围为 $|N| \leq 2^m - 1$, 即 $-(2^m - 1) \leq N \leq (2^m - 1)$ 。

- 定点小数格式: $N = b_s . b_{-1} b_{-2} \cdots b_{-n}$

小数点位置固定在符号位 b_s 与最高数据位 b_{-1} 之间。其取值范围 $|N| \leq 1 - 2^{-n}$ 。

需要注意的是,无论是在定点整数还是定点小数表示法中,小数点作为一个符号并不明确地表示出来,而是由其表示方法所隐含。例如:

$N = 10111010$, 若表示定点小数则其值为 -0.011101 ; 若表示定点整数,则其值为 -111010 。

(2) 浮点数据表示方法

这种表示方法把数值数据表示成

$$N = M \cdot R^E$$

其中, M 被称为 N 的尾数, E 是 N 的指数或称阶码,而 R 则是该阶码的基数。在这种表示方法下,同一个数可以有不同的浮点表示形式,阶码的大小可以用来调节数值中小数点的位置。例如: 0.081 可以表示为 $81E-3$, 也可以表示为 $8.1E-2$ 或者 $0.81E-1$ 等等,这里 E 只是尾数和阶码之间的一个分隔符,而隐含基数为 10。在这种意义上,其小数点位置是可以“浮动”的,所以把该方法称为浮点数据表示法。

需要指出的是在我们上面的表示方法中,其基数 10 并未出现在表示形式中,而是隐含规定的。在计算机中,基数 R 通常由硬件部件(浮点运算部件)隐含规定,也不明确地出现在表达式中。在计算机里浮点数通常以如下形式表示:

M_s	E	M
-------	-----	-----