



全国计算机技术与软件专业技术资格（水平）考试指定用书

程序员教程

张淑平 覃桂敏 主编

全国计算机专业技术资格考试办公室 组编



清华大学出版社

第5版

技术与软件专业技术资格（水平）考试指定用书

程序员教程

（第5版）

张淑平 覃桂敏 主编

清华大学出版社
北京

内 容 简 介

本书作为全国计算机技术与软件专业技术资格(水平)考试(简称“软考”)的初级职称指定教材,具有比较权威的指导意义。本书根据《程序员考试大纲》(2018 年审定通过)的重点内容,组织了共 11 章的内容,考生在学习教材内容的同时,还须对照考试大纲,认真学习和复习大纲要求的知识点。

本书是在《程序员考试大纲》的指导下,对《程序员教程(第 4 版)》进行再编后完成的。

本书适合参加相关考试的考生和在校大学生作为教材使用。

本书扉页为防伪页,封面贴有清华大学出版社防伪标签,无上述标识者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

程序员教程/张淑平,覃桂敏主编.—5 版.—北京:清华大学出版社,2018

(全国计算机技术与软件专业技术资格(水平)考试指定用书)

ISBN 978-7-302-49123-1

I. ①程… II. ①张…②覃… III. ①程序设计-资格考试-自学参考资料 IV. ①TP311.1

中国版本图书馆 CIP 数据核字(2017)第 313212 号

责任编辑:杨如林 柴文强

封面设计:常雪影

责任校对:白 蕾

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:清华大学印刷厂

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×230mm 印 张:31.75 防伪页:1 字 数:672 千字

版 次:2004 年 7 月第 1 版 2018 年 2 月第 5 版 印 次:2018 年 2 月第 1 次印刷

印 数:1~5000

定 价:89.00 元

产品编号:075519-01

第 5 版前言

全国计算机技术与软件专业技术资格（水平）考试从实施至今已有二十余年，在社会上产生了很大的影响，对我国软件产业的形成和发展做出了重要的贡献。为了适应我国计算机信息技术发展的需求，人力资源和社会保障部、工业和信息化部决定将考试的级别拓展到计算机信息技术行业的各个方面，以满足社会上对计算机信息技术人才的需要。

编者受全国计算机专业技术资格考试办公室委托，对《程序员教程（第 4 版）》一书进行再编，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，每个章节的内容都能构成相关领域的一门课程，因此编写本书的难度很高。考虑到参加考试的人员已有一定的基础，所以本书中只对考试大纲中所涉及的知识领域的要点加以阐述，但限于篇幅所限，不能详细地展开，请读者谅解。

全书共分 11 章，各章的内容安排如下。

第 1 章 计算机系统基础知识：主要介绍计算机系统硬件组成、数据在计算机中的表示和运算、校验码基础知识、指令系统和多媒体系统基础知识。

第 2 章 操作系统基础知识：主要介绍操作系统的类型和功能等基本概念，进程管理、存储管理、设备管理、文件管理和作业管理等基础知识。

第 3 章 程序设计语言基础知识：主要介绍程序设计语言的类型和特点、程序设计语言的基本成分以及编译、解释等基本语言翻译基础知识。

第 4 章 数据结构与算法：主要介绍线性表和链表、栈、队列、数组、树、图等基本数据结构以及查找、排序等常用算法。

第 5 章 软件工程基础知识：主要介绍软件工程和项目管理基础、面向对象分析与设计方法、软件需求分析、软件设计、编码和测试、软件系统运行与维护、软件质量管理等基础知识。

第 6 章 数据库基础知识：主要介绍数据库管理系统的主要功能和特征、数据库模式、数据模型和 ER 图、关系运算和 SQL 等基础知识。

第 7 章 网络与信息安全基础知识：主要介绍网络的功能、分类、组成和拓扑结构，基本的网络协议与标准，常用网络设备与网络通信设备的作用和特点、局域网（LAN）和互联网（Internet）基础知识，以及信息安全、网络安全基础知识。

第 8 章 标准化和知识产权基础知识：主要介绍标准化的基本概念和知识产权的概念与特点、计算机软件著作权和商业秘密权基础知识。

第9章 C程序设计：主要介绍C程序基础、语句、函数、指针与简单C程序中常见错误。

第10章 C++程序设计：主要介绍C++程序基础、类与对象、继承与多态、输入与输出流库、异常处理和常用STL模板库。

第11章 Java程序设计：主要介绍Java程序语言基础和特点、类与接口、异常、文件和输入/输出流以及Java类库等基础知识。

本书第1章由张淑平、马志欣编写，第2章由王亚平编写，第3章和第4章由张淑平编写，第5章由褚华、霍秋艳编写，第6章由王亚平编写，第7章由严体华编写，第8章由刘强编写，第9章由张淑平、覃桂敏编写，第10章由张淑平、宋胜利编写，第11章由霍秋艳编写，全书由张淑平、覃桂敏统稿。

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，书中难免存在欠妥之处，望读者指正，以利改进和提高。

编 者

2018年1月

目 录

第 1 章 计算机系统基础知识1	
1.1 计算机系统的基本组成.....1	
1.2 数据的表示及运算.....3	
1.2.1 计算机中数据的表示.....3	
1.2.2 校验码.....13	
1.2.3 逻辑代数及逻辑运算.....17	
1.2.4 机器数的运算.....19	
1.3 计算机的基本组成及工作原理.....23	
1.3.1 总线的基本概念.....23	
1.3.2 中央处理单元.....25	
1.3.3 存储系统.....30	
1.3.4 输入/输出技术.....36	
1.4 指令系统简介.....40	
1.5 多媒体系统简介.....44	
1.5.1 数字声音.....45	
1.5.2 图形与图像.....48	
1.5.3 动画和视频.....55	
第 2 章 操作系统基础知识62	
2.1 操作系统概述.....62	
2.2 进程管理.....65	
2.2.1 基本概念.....65	
2.2.2 进程控制.....67	
2.2.3 进程通信.....68	
2.2.4 进程调度.....72	
2.2.5 死锁.....74	
2.2.6 线程.....76	
2.3 存储管理.....76	
2.3.1 基本概念.....76	
2.3.2 存储管理方案.....77	
2.3.3 分页存储管理.....79	
2.3.4 分段存储管理.....80	
2.3.5 虚拟存储管理.....82	
2.4 设备管理.....85	
2.4.1 设备管理概述.....85	
2.4.2 设备管理技术.....86	
2.4.3 磁盘调度.....88	
2.5 文件管理.....89	
2.5.1 基本概念.....89	
2.5.2 文件的结构和组织.....90	
2.5.3 文件目录.....92	
2.5.4 存取方法、存取控制.....93	
2.5.5 文件的使用.....95	
2.5.6 文件的共享和保护.....95	
2.5.7 系统的安全与可靠性.....96	
2.6 作业管理.....98	
2.6.1 作业管理.....98	
2.6.2 作业调度.....99	
2.6.3 人机界面.....102	
第 3 章 程序设计语言基础知识104	
3.1 程序设计语言概述.....104	
3.1.1 程序设计语言的基本概念.....104	
3.1.2 程序设计语言的分类和特点.....105	
3.1.3 程序设计语言的基本成分.....109	
3.2 语言处理程序基础.....116	

3.2.1 汇编程序基础	116	5.2 软件需求分析	212
3.2.2 编译程序基础	118	5.2.1 软件需求的定义	212
3.2.3 解释程序基础	126	5.2.2 软件需求分析的基本任务	213
第4章 数据结构与算法	129	5.2.3 需求建模	213
4.1 线性结构	129	5.3 软件设计	214
4.1.1 线性表	129	5.3.1 软件设计的基本任务	214
4.1.2 栈和队列	137	5.3.2 软件设计原则	216
4.1.3 串	143	5.4 结构化分析与设计方法	219
4.2 数组	144	5.4.1 结构化分析方法	219
4.3 树与二叉树	147	5.4.2 结构化设计方法	221
4.3.1 树的基本概念	148	5.4.3 结构化程序设计方法	223
4.3.2 二叉树	149	5.5 面向对象分析与设计方法	224
4.3.3 树和森林	153	5.5.1 面向对象的基本概念	224
4.3.4 最优二叉树	156	5.5.2 面向对象分析与设计	227
4.3.5 二叉查找树	157	5.5.3 UML 概述	228
4.4 图	158	5.5.4 设计模式	232
4.5 查找	163	5.6 软件测试与运行	235
4.5.1 顺序查找与折半查找	164	5.6.1 软件测试的目的及原则	235
4.5.2 树表查找	167	5.6.2 软件测试方法	238
4.5.3 哈希表及哈希查找	170	5.6.3 软件测试过程	241
4.6 算法	173	5.6.4 软件测试设计和管理	242
4.6.1 算法概述	173	5.6.5 软件调试	243
4.6.2 排序算法	177	5.6.6 软件运行与维护	243
4.6.3 递归算法	187	5.7 软件项目管理	245
4.6.4 字符串运算	188	5.7.1 管理范围	245
4.6.5 图的相关算法	191	5.7.2 成本估算	246
第5章 软件工程基础知识	199	5.7.3 风险分析	247
5.1 软件工程概述	199	5.7.4 进度管理	248
5.1.1 软件生存周期	199	5.8 软件质量与软件质量保证	250
5.1.2 软件生存周期模型	201	5.8.1 软件质量特性	250
5.1.3 软件过程评估	207	5.8.2 软件质量保证	252
5.1.4 软件工具	209	第6章 数据库基础知识	254
5.1.5 软件开发环境	211	6.1 基本概念	254

6.1.1 数据库系统	254	7.2.1 计算机网络互连设备	305
6.1.2 数据库管理技术的发展	255	7.2.2 计算机网络传输媒体	311
6.1.3 大数据	257	7.3 TCP/IP	314
6.2 数据模型	260	7.3.1 TCP/IP 模型	314
6.2.1 数据模型的基本概念	260	7.3.2 TCP/IP 协议	315
6.2.2 数据模型的三要素	261	7.3.3 IP 地址	318
6.2.3 E-R 模型	261	7.4 Internet 基础知识	322
6.2.4 基本的数据模型	265	7.4.1 Internet 服务	322
6.3 DBMS 的功能和特征	267	7.4.2 因特网接入方式	327
6.3.1 DBMS 的功能	267	7.4.3 TCP/IP 的配置	328
6.3.2 DBMS 的特征与分类	268	7.4.4 浏览器的设置与使用	329
6.4 数据库模式	270	7.5 局域网基础知识	332
6.4.1 模式	270	7.6 信息安全基础知识	337
6.4.2 三级模式两级映像	272	7.7 网络安全概述	341
6.5 关系数据库与关系运算	272	第 8 章 标准化和知识产权基础知识	346
6.5.1 关系数据库的基本概念	272	8.1 标准化基础知识	346
6.5.2 关系数据库模式	275	8.1.1 软件工程标准化	346
6.5.3 完整性约束	276	8.1.2 能力成熟度模型简介	348
6.5.4 关系代数运算	276	8.2 知识产权基础知识	349
6.6 关系数据库 SQL 语言简介	280	8.2.1 基本概念	349
6.6.1 SQL 概述	280	8.2.2 计算机软件著作权	351
6.6.2 SQL 数据定义	281	8.2.3 计算机软件的商业秘密权	363
6.6.3 SQL 数据查询	285	第 9 章 C 程序设计	366
6.6.4 SQL 数据更新	293	9.1 C 语言基础	366
6.6.5 SQL 的访问控制	295	9.1.1 数据类型	366
6.6.6 嵌入式 SQL	296	9.1.2 运算符与表达式	372
6.7 数据库设计	297	9.1.3 输入/输出	375
第 7 章 网络与信息安全基础知识	300	9.2 控制语句	377
7.1 计算机网络概述	300	9.3 函数	381
7.1.1 计算机网络的组成	300	9.4 指针	385
7.1.2 计算机网络的分类	301	9.4.1 指针的定义	385
7.1.3 ISO/OSI 参考模型	304	9.4.2 指针与数组	387
7.2 计算机网络硬件	305	9.4.3 指针与函数	389

9.4.4 指针与链表	393	11.2.1 Java 基本数据类型	449
9.5 常见的 C 程序错误	395	11.2.2 控制结构	455
第 10 章 C++ 程序设计	407	11.2.3 Java 核心类	458
10.1 C++ 程序基础	407	11.3 类与接口	462
10.1.1 数据类型	407	11.3.1 类的定义与使用	462
10.1.2 运算符、表达式和语句	409	11.3.2 对象的初始化	464
10.1.3 基本输入/输出	410	11.3.3 包	466
10.1.4 函数	412	11.3.4 继承	468
10.1.5 类与对象	415	11.3.5 抽象类与接口	471
10.2 继承与多态	423	11.4 异常	478
10.3 异常处理	428	11.4.1 异常的处理	478
10.4 标准库	430	11.4.2 自定义异常	480
10.4.1 I/O 流库	431	11.5 输入、输出和流	483
10.4.2 string	434	11.5.1 字节流	483
10.4.3 STL	438	11.5.2 字符流	486
10.4.4 vector	443	11.5.3 标准输入/输出流	487
第 11 章 Java 程序设计	447	11.6 Java 类库的使用	489
11.1 Java 语言概述	447	11.6.1 基本类型的包装类型	489
11.2 Java 语言基础	448	11.6.2 泛型和集合类	493

第 1 章 计算机系统基础知识

本章主要介绍计算机系统的基本组成、计算机中数据的表示和运算、计算机系统硬件基础组成、指令系统以及多媒体系统等基础知识。

1.1 计算机系统的基本组成

计算机系统是由硬件系统和软件系统组成的，通过运行程序来协同工作。计算机硬件是物理装置，计算机软件是程序、数据和相关文档的集合。计算机系统的基本组成如图 1-1 所示。

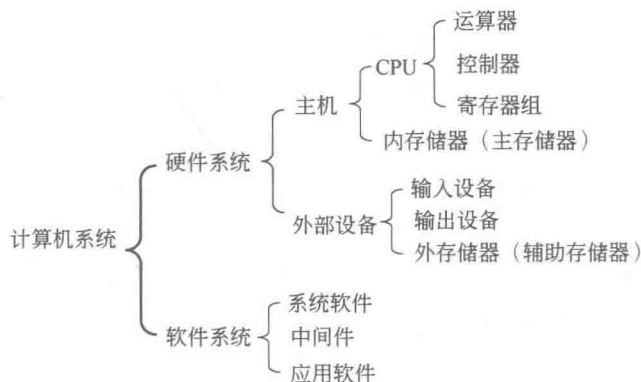


图 1-1 计算机系统的基本组成

1. 计算机硬件

基本的计算机硬件系统由运算器、控制器、存储器、输入设备和输出设备五大部件组成，随着网络技术的发展和运用，通信部件也成为计算机系统的基本组件。运算器和控制器及其相关部件已被集成在一起，统称为中央处理单元（Central Processing Unit, CPU）。CPU 是硬件系统的核心，用于数据的加工处理，能完成各种算术、逻辑运算及控制功能。

运算器是对数据进行加工处理的部件，它主要完成算术和逻辑运算。控制器的主要功能则是从主存中取出指令并进行分析，控制计算机的各个部件有条不紊地完成指令的

功能。

存储器是计算机系统中的记忆设备，分为内部存储器（Main Memory，MM，简称内存、主存）和外部存储器（简称外存，辅存）。相对来说，内存速度快、容量小，一般用来临时存储计算机运行时所需的程序、数据及运算结果。外存容量大、速度慢，可用于长期保存信息。寄存器是 CPU 中的存储器件，用来临时存放少量的数据、运算结果和正在执行的指令。与内存存储器相比，寄存器的速度要快得多。

习惯上将 CPU 和主存储器的有机组合称为主机。输入/输出（I/O）设备位于主机之外，是计算机系统与外界交换信息的装置。所谓输入和输出，都是相对于主机而言的。输入设备的作用是将信息输入计算机的存储器中，输出设备的作用是把运算结果按照人们所要求的形式输出到外部设备或存储介质上。

2. 计算机软件

计算机软件是指为管理、运行、维护及应用计算机系统所开发的程序和相关文档的集合。如果计算机系统中仅有硬件系统，则只具备了计算的基础，并不能真正计算，只有将解决问题的步骤编制成机器可识别的程序并加载到计算机内存开始运行，才能完成计算。

软件是计算机系统中的重要组成部分，通常可将软件分为系统软件、中间件和应用软件。系统软件的主要功能是管理系统的硬件和软件资源，应用软件则用于解决应用领域的具体问题，中间件是一类独立的系统软件或服务程序，常用来管理计算资源和网络通信，提供通信处理、数据存取、事务处理、Web 服务、安全、跨平台等服务。

3. 计算机分类

计算机技术的发展异常迅速，将更多的元件集成到单一的半导体芯片上，使得计算机变得更小，功耗更低，速度更快。

（1）个人移动设备（Personal Mobile Device，PMD）。指一类带有多媒体用户界面的无线设备，如智能手机、平板电脑等。

（2）桌面计算机。桌面计算机的产品范围非常广泛，包括低端的上网本、台式计算机、笔记本电脑以及高配置的工作站，核心部件是基于超大规模集成电路技术的 CPU。台式计算机和笔记本电脑属于微型计算机，常用于一般性的办公事务处理等，工作站则是一种高档的微型计算机，通常配有高分辨率的大屏幕显示器及容量很大的内存储器和外部存储器，具备强大的数据运算与图形、图像处理能力，主要面向工程设计、动画制作、科学研究、软件开发、金

融管理、信息服务、模拟仿真等专业应用领域。

(3) 服务器。不同于桌面计算机,服务器代替了传统的大型机,主要提供大规模和可靠的文件及计算服务,强调可用性、可扩展性和很高的吞吐率。

(4) 集群/仓库级计算机。集群机是将一组桌面计算机或服务器用网络连接在一起,运行方式类似于一个大型的计算机。将数万个服务器连接在一起形成的大规模集群称为仓库级计算机。

(5) 超级计算机。超级计算机的基本组成在概念上与个人计算机无太大差异,但规格高,性能要强大许多,具有很强的计算能力,但是能耗巨大。我国的超级计算机主要有银河、天河、曙光、神威四个系列。例如,神威·太湖之光由40个运算机柜和8个网络机柜组成,共有40960块处理器,每一块处理器相当于20多台常用笔记本电脑的计算能力。

(6) 嵌入式计算机。嵌入式计算机是专用的,是针对某个特定的应用,如针对网络、通信、音频、视频或针对工业控制,对功能、可靠性、成本、体积、功耗有严格要求的计算机系统。日常生活中常见的微波炉、洗衣机、数码产品、网络交换机和汽车中都采用嵌入式计算机技术。

1.2 数据的表示及运算

1.2.1 计算机中数据的表示

在计算机内部,数值、文字、声音、图形图像等各种信息都必须经过数字化编码后才能被传送、存储和处理。所谓编码,就是采用少量的基本符号,选用一定的组合原则,来表示大量复杂多样的信息。基本符号的种类和这些符号的组合规则是一切信息编码的两大要素。例如,用10个阿拉伯数码表示数字,用26个英文字母表示英文词汇等,都是编码的典型例子。

1. 进位计数制及其转换

在采用进位计数的数字系统中,如果只用 r 个基本符号表示数值,则称其为 r 进制(Radix- r Number System), r 称为该数制的基数(Radix)。不同数制的共同特点如下。

(1) 每一种数制都有固定的符号集。例如,十进制数制的基本符号有十个:0, 1, 2, ..., 9。二进制数制的基本符号有两个:0和1。

(2) 每一种数制都使用位置表示法。即处于不同位置的数符所代表的值不同,与它所在位置的权值有关。例如,十进制数1234.55可表示为

$$1234.55 = 1 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 4 \times 10^0 + 5 \times 10^{-1} + 5 \times 10^{-2}$$

计算机中常用的进位数制有二进制、八进制、十进制和十六进制，如表 1-1 所示。

表 1-1 常用计数制

进 位 制	二 进 制	八 进 制	十 进 制	十 六 进 制
规则	逢二进一	逢八进一	逢十进一	逢十六进一
基数	$r=2$	$r=8$	$r=10$	$r=16$
数符	0, 1	0, 1, 2, ..., 7	0, 1, 2, ..., 9	0, 1, 2, ..., 9, A, B, ..., F
权	2^i	8^i	10^i	16^i
形式表示符	B	O	D	H

可以看出，十进制计数制中权的值恰好是基数 10 的某次幂，其他计数制同理。因此，对任何一种进位计数制，其表示的数都可以写成按权展开的多项式，在此基础上实现不同计数制的相互转换。

1) 十进制计数法与二进制计数法的相互转换

在十进制计数制中， $r=10$ ，基本符号为 0, 1, 2, ..., 9。

在二进制计数制中， $r=2$ ，基本符号为 0 和 1。二进制数中的一个 0 或 1 称为 1 位 (bit)。

将十进制数转换成二进制数时，整数部分和小数部分分别转换，然后再合并。十进制整数转换为二进制整数的方法是“除 2 取余”；十进制小数转换为二进制小数的方法是“乘 2 取整”。

【例 1-1】 把十进制数 175.71875 转换为相应的二进制数。

算式	商	余数	算式	乘积
175 / 2	87	1	0.71875 * 2	1.43750
87 / 2	43	1	0.4375 * 2	0.8750
43 / 2	21	1	0.875 * 2	1.750
21 / 2	10	1	0.75 * 2	1.50
10 / 2	5	0	0.5 * 2	1.0
5 / 2	2	1	0.71875 ₁₀ = 0.10111 ₂	
2 / 2	1	0		
1 / 2	0	1		

$$175_{10} = 10101111_2$$

因此，175.71875₁₀ = 10101111.10111₂

在熟悉 2 的整幂次情况下，可将十进制数写成按二进制数权的大小展开的多项式，按权值从高到低依次取各项的系数就可得到相应的二进制数。

$$\begin{aligned}
 (175.71875)_{10} &= 2^7 + 2^5 + 2^3 + 2^2 + 2^1 + 2^0 + 2^{-1} + 2^{-3} + 2^{-4} + 2^{-5} \\
 &= 10101111.10111_2
 \end{aligned}$$

二进制数转换成十进制数的方法是：将二进制数的每一位数乘以它的权，然后相加，即可求得对应的十进制数值。

【例 1-2】 把二进制数 100110.101 转换成相应的十进制数。

$$\begin{aligned}
 (100110.101)_2 &= 1 \times 2^5 + 0 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} \\
 &= 32 + 0 + 0 + 4 + 2 + 0 + 0.5 + 0 + 0.125 \\
 &= 38.625
 \end{aligned}$$

2) 八进制计数法与十进制、二进制计数法的相互转换

八进制计数制的基本符号为 0, 1, 2, ..., 7。

十进制数转换为八进制数的方法是：对于十进制整数采用“除 8 取余”的方法转换为八进制整数；对于十进制小数则采用“乘 8 取整”的方法转换为八进制小数。

二进制数转换成八进制数的方法是：从小数点起，每三位二进制位分成一组（不足 3 位时，在小数点左边时左边补 0，在小数点右边时右边补 0），然后写出每一组的等值八进制数，顺序排列起来就得到所要求的八进制数。

【例 1-3】 将二进制数 10101111.10111 转换为相应的八进制数。

$$10101111.10111_2 = 010\ 101\ 111.101\ 110_2 = 257.56_8$$

依照同样的思想，将一位八进制数用三位二进制数表示，就可以直接将八进制数转换成二进制数。

二进制、八进制数和十六进制数之间的对应关系如表 1-2 所示。

表 1-2 二进制、八进制和十六进制数之间的对应关系

二进制	八进制	二进制	十六进制	二进制	十六进制
000	0	0000	0	1000	8
001	1	0001	1	1001	9
010	2	0010	2	1010	A
011	3	0011	3	1011	B
100	4	0100	4	1100	C
101	5	0101	5	1101	D
110	6	0110	6	1110	E
111	7	0111	7	1111	F

3) 十六进制计数法与十进制、二进制计数法的相互转换

在十六进制计数制中， $r=16$ ，基本符号为 0, 1, 2, ..., 9, A, B, ..., F。

十进制数可以转换为十六进制数的方法是：十进制数的整数部分“除 16 取余”，十进制

数的小数部分“乘 16 取整”。

由于一位十六进制数可以用 4 位二进制数来表示,因此二进制数与十六进制数的相互转换就比较容易。二进制数转换成十六进制数的方法是:从小数点开始,每 4 位二进制数为一组(不足 4 位时,在小数点左边时左边补 0,在小数点右边时右边补 0),将每一组用相应的十六进制数符来表示,即可得到正确的十六进制数。

【例 1-4】 将二进制数 10101111.10111 转换为相应的十六进制数。

$$(1010\ 1111.1011\ 1)_2 = \text{AF.B8}_{16}$$

2. 二进制运算规则

(1) 加法:二进制加法的进位规则是“逢二进一”。

$$0+0=0 \quad 1+0=1 \quad 0+1=1 \quad 1+1=0 \text{ (有进位)}$$

(2) 减法:二进制减法的借位规则是“借一当二”。

$$0-0=0 \quad 1-0=1 \quad 1-1=0 \quad 0-1=1 \text{ (有借位)}$$

(3) 乘法:

$$0 \times 0=0 \quad 1 \times 0=0 \quad 0 \times 1=0 \quad 1 \times 1=1$$

3. 机器数和码制

各种数据在计算机中表示的形式称为机器数,其特点是采用二进制计数制,数的符号用 0、1 表示,小数点隐含表示而不占位置。机器数对应的实际数值称为数的真值。

对于带符号数,机器数的最高位是表示正、负的符号位,其余位则表示数值。若约定小数点的位置在机器数的最低数值位之后,则是纯整数;若约定小数点的位置在机器数的最高数值位之前(符号位之后),则是纯小数。无符号数是指全部二进制位均代表数值,没有符号位。

为了便于运算,带符号的机器数可采用原码、反码和补码、移码等不同的编码方法。

1) 原码表示

数值 X 的原码记为 $[X]_{\text{原}}$,如果机器字长为 n (即采用 n 个二进制位表示数据),则最高位是符号位,0 表示正号,1 表示负号,其余的 $n-1$ 位表示数值的绝对值。数值零的原码表示有两种形式: $[+0]_{\text{原}}=00000000$, $[-0]_{\text{原}}=10000000$ 。

【例 1-5】 若机器字长 n 等于 8,则

$$[+1]_{\text{原}}=00000001 \quad [-1]_{\text{原}}=10000001$$

$$[+127]_{\text{原}}=01111111 \quad [-127]_{\text{原}}=11111111$$

$$[+45]_{\text{原}}=00101101 \quad [-45]_{\text{原}}=10101101$$

$$[+0.5]_{\text{原}}=0 \diamond 1000000 \quad [-0.5]_{\text{原}}=1 \diamond 1000000 \text{ (其中 } \diamond \text{ 是小数点的位置)}$$

2) 反码表示

数值 X 的反码记作 $[X]_{\text{反}}$, 如果机器字长为 n , 则最高位是符号位, 0 表示正号, 1 表示负号, 其余的 $n-1$ 位表示数值。正数的反码与原码相同, 负数的反码则是其绝对值按位求反。数值 0 的反码表示有两种形式: $[+0]_{\text{反}}=00000000$, $[-0]_{\text{反}}=11111111$ 。

【例 1-6】 若机器字长 n 等于 8, 则

$[+1]_{\text{反}}=00000001$	$[-1]_{\text{反}}=11111110$
$[+127]_{\text{反}}=01111111$	$[-127]_{\text{反}}=10000000$
$[+45]_{\text{反}}=00101101$	$[-45]_{\text{反}}=11010010$
$[+0.5]_{\text{反}}=0 \diamond 1000000$	$[-0.5]_{\text{反}}=1 \diamond 01111111$ (其中 $\diamond$ 是小数点的位置)

3) 补码表示

数值 X 的补码记作 $[X]_{\text{补}}$, 如果机器字长为 n , 则最高位为符号位, 0 表示正号, 1 表示负号, 其余的 $n-1$ 位表示数值。正数的补码与其原码和反码相同, 负数的补码则等于其反码的末尾加 1。在补码表示中, 0 有唯一的编码: $[+0]_{\text{补}}=00000000$, $[-0]_{\text{补}}=00000000$ 。

【例 1-7】 若机器字长 n 等于 8, 则

$[+1]_{\text{补}}=00000001$	$[-1]_{\text{补}}=11111111$
$[+127]_{\text{补}}=01111111$	$[-127]_{\text{补}}=10000001$
$[+45]_{\text{补}}=00101101$	$[-45]_{\text{补}}=11010011$
$[+0.5]_{\text{补}}=0 \diamond 1000000$	$[-0.5]_{\text{补}}=1 \diamond 1000000$ (其中 $\diamond$ 是小数点的位置)

相对于原码和反码表示, n 位补码表示法有一个例外, 当符号位为 1 而数值位全部为 0 时, 它表示整数 2^{n-1} , 即此时符号位的 1 既表示负数又表示数值。

设计补码时, 有意识地引用了模运算在数理上对符号位的处理, 利用模的自动丢弃实现了符号位的自然处理。

用补码表示数时, 由于符号位和数值部分一起编码, 很难从码值形式直接判断真值的大小。例如, $45 > -45$, 而其补码 00101101 在形式上小于 11010011。

4) 移码表示

移码表示法是在数 X 上增加一个偏移量来定义的, 常用于表示浮点数中的阶码。如果机器字长为 n , 在偏移量为 2^{n-1} 时, 只要将补码的符号位取反便可获得相应的移码表示。偏移量也可以是其他值。采用移码表示时, 码值大者对应的真值就大。

【例 1-8】 若机器字长 n 等于 8, 则

$[+1]_{\text{移}}=10000001$	$[-1]_{\text{移}}=01111111$
$[+127]_{\text{移}}=11111111$	$[-127]_{\text{移}}=00000001$
$[+45]_{\text{移}}=10101101$	$[-45]_{\text{移}}=01010011$
$[+0]_{\text{移}}=10000000$	$[-0]_{\text{移}}=10000000$

4. 定点数和浮点数

1) 定点数

所谓定点数,就是表示数据时小数点的位置固定不变。小数点的位置通常有两种约定方式:定点整数(纯整数,小数点在最低有效数值位之后)和定点小数(纯小数,小数点在最高有效数值位之前)。

设机器字长为 n , 各种码制表示下的带符号数的范围如表 1-3 所示。当机器字长为 n 时, 定点数的补码和移码可表示 2^n 个数, 而其原码和反码只能表示 $2^n - 1$ 个数(0 表示占用了两个编码), 因此, 定点数所能表示的数值范围比较小, 运算中很容易因结果超出范围而溢出。

表 1-3 机器字长为 n 时各种码制表示的带符号数的范围

码 制	定 点 整 数	定 点 小 数
原码	$-(2^{n-1}-1) \sim +(2^{n-1}-1)$	$-(1-2^{-(n-1)}) \sim +(1-2^{-(n-1)})$
反码	$-(2^{n-1}-1) \sim +(2^{n-1}-1)$	$-(1-2^{-(n-1)}) \sim +(1-2^{-(n-1)})$
补码	$-2^{n-1} \sim +(2^{n-1}-1)$	$-1 \sim +(1-2^{-(n-1)})$
移码	$-2^{n-1} \sim +(2^{n-1}-1)$	$-1 \sim +(1-2^{-(n-1)})$

2) 浮点数

浮点数是小数点位置不固定的数, 浮点表示法能表示更大范围的数。

在十进制中, 一个实数可以写成多种表示形式。例如, 83.125 可写成 $10^3 \times 0.083125$ 或 $10^4 \times 0.0083125$ 等。同理, 一个二进制数也可以写成多种表示形式。例如, 二进制数 1011.10101 可以写成 $2^4 \times 0.101110101$ 、 $2^5 \times 0.0101110101$ 或 $2^6 \times 0.00101110101$ 等。

一个含小数点的二进制数 N 可以表示为更一般的形式:

$$N = 2^E \times F$$

其中, E 称为阶码, F 为尾数, 这种表示数的方法称为浮点表示法。

在浮点表示法中, 阶码通常为带符号的纯整数, 尾数为带符号的纯小数。浮点数的表示格式一般如下:

阶符	阶码	数符	尾数
----	----	----	----

很明显, 一个数的浮点表示不是唯一的。当小数点的位置改变时, 阶码也相应改变, 因此可以用多种浮点形式表示同一个数。