



全国计算机技术与软件专业技术资格（水平）考试辅导用书

新编软件设计师 考试辅导

全国计算机专业技术资格考试办公室组编
谢树煜 主编

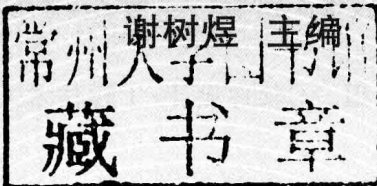
清华大学出版社



全国计算机技术与软件专业技术资格（水平）考试辅导用书

新编软件设计师 考试辅导

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
北京

内 容 简 介

本书根据全国计算机技术及软件专业技术资格（水平）考试《软件设计师考试大纲》的要求，参照《软件设计师教程》的结构及历年软件设计师专业资格考试试题编写，内容紧扣考试大纲。全书共分 13 章，每章都由内容提要、例题分析、思考练习题组成。内容提要是对有关章节知识的提炼，给出考试要点和学习难点。例题分析是全书重点，着重解析考试大纲要求的基本知识及其综合应用方法。思考练习题供读者检验自己对有关内容掌握的程度。为了帮助读者提高理解程序、编制程序及软件设计的能力，本书专门增加了 C/C++ 语言程序设计、面向对象技术、软件系统分析与设计各章，并在有关章节中加大了软件工程、数据结构和常用算法设计方法的比重。

本书供参加软件设计师考试的读者应试复习时使用，也可供大专院校师生及计算机技术人员参考。

本书扉页为防伪页，封面贴有清华大学出版社防伪标签，无标签者不得销售。

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

新编软件设计师考试辅导 / 谢树煜主编. —北京：清华大学出版社，2010.10

（全国计算机技术与软件专业技术资格（水平）考试辅导用书）

ISBN 978-7-302-22593-5

I. ①新… II. ①谢… III. ①软件设计-工程技术人员-资格考核-自学参考资料
IV. ①TP311.5

中国版本图书馆 CIP 数据核字（2010）第 079732 号

责任编辑：柴文强 赵晓宁

责任校对：徐俊伟

责任印制：王秀菊

出版发行：清华大学出版社

地 址：北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编：100084

社 总 机：010-62770175

邮 购：010-62786544

投稿与读者服务：010-62795954, jsjic@tup.tsinghua.edu.cn

质 量 反 馈：010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者：清华大学印刷厂

经 销：全国新华书店

开 本：185×230 印 张：46.75 防 伪 页：1 字 数：1080 千字

版 次：2010 年 10 月第 1 版 印 次：2010 年 10 月第 1 次印刷

印 数：1~5000

定 价：79.00 元

产品编号：034389-01

前 言

计算机技术及软件专业技术资格（水平）考试是国家人力资源与社会保障部与工业和信息化部主办的国家级考试，20多年来为国家选拔和培养了十多万名合格的软件技术人才，在国内外产生了很大的影响，得到了社会各界的广泛认同。

2002年1月，为了推动中日两国间信息技术的交流与合作，信息产业部电子教育中心与日本信息处理技术人员考试中心就IT考试标准相互认证达成协议，中国计算机技术与软件专业技术资格（水平）考试成为中日双方互相承认的软件技术资格考试，中国软件考试逐步走上与国际接轨之路。后来中韩两国也达成考试标准互认协议。

2003年10月，国家人事部与信息产业部发布的第39号文件规定，把计算机技术与软件专业技术资格考试纳入全国专业技术人员职业资格证书制度的统一规划中。通过考试取得技术资格证书的人员，表明已具备相应的专业岗位工作水平和能力，用人单位可择优聘任担任相应的专业技术职务。同时决定今后不再进行相应专业技术职务任职资格的评审工作，因此这种考试既是职业资格考试，又是技术资格考试。我们相信这种以考代评的重大改革，对软件专业技术人才培养将起到巨大的推动作用。

为了规范计算机专业技术资格考试，全国计算机专业技术资格考试办公室公布了统一的考试大纲，2009年公布了新版考试大纲与培训指南。为了帮助广大考生深入理解考试大纲的要求，掌握有关课程的基本概念、基本内容和基本方法，进一步提高应试能力，在全国计算机专业技术资格考试办公室的领导下，在清华大学出版社的支持下，我们编写了《新编软件设计师考试辅导》。该书是根据考试大纲的要求，配合学员考试自学复习的需要编写的。书中介绍了有关专业要求的基本知识和技能，内容涵盖软件专业的核心课程。全书包括计算机系统知识、操作系统、数据库、多媒体、计算机网络、程序设计语言的基础知识，标准化和知识产权的基础知识。强化了软件工程、数据结构、常用算法设计方法的内容，为了帮助学员应对下午试卷，学习C和C++程序设计语言，本书增加了C/C++语言程序设计、面向对象技术、软件系统分析与设计3章内容。

考试大纲要求学员掌握的知识面很宽，考虑到学员复习的时间有限，书中针对考试大纲及教材的内容要点和学习难点作了剖析，并把重点放在例题分析上，这些题目都是作者从自己切身教学经验和历届考题中精选出来的，例题分析中对有关解题思路、解题方法、应用的基本知识和基本原理，做了详尽介绍，一定会对参加考试的学员有所启发和帮助。每章还附有思考练习题及答案，供学员自我检查练习时使用。

本书由清华大学谢树煜教授主编，全书共分13章。第1章计算机系统知识由谢树煜编写；第2章程序设计语言基础由北京大学丁文魁教授编写；第3章操作系统知识、第5章

网络基础知识由中国农业大学孙瑞志教授编写，北京大学方裕教授在操作系统方面的工作对他也很有帮助应该表示感谢；第 6 章多媒体基础知识由清华大学林福宗教授编写；第 7 章数据库技术基础由北京科技大学王道平教授编写，第 4 章系统开发与运行、第 8 章数据结构、第 10 章面向对象技术由清华大学殷人昆教授编写；第 9 章常用算法设计方法、第 11 章标准化基础知识与第 12 章 C/C++ 语言程序设计由清华大学孙甲松副教授完成，第 11 章中知识产权基础知识由国家软件保护中心李维高级工程师编写；最后一章软件系统分析与设计由殷人昆、王道平共同完成。他们都是有关学科的专家和教授，有丰富的教学经验，特别是很多老师多年来担任过软件专业技术资格和水平考试的考前培训工作，积累了很多经验。对在本书编写过程中清华大学出版社柴文强编审的大力支持深表感谢。

由于水平和时间有限，书中不妥之处敬请指正。

编者

2010 年 1 月于清华园

目 录

第 1 章 计算机系统知识1	
1.1 内容提要.....1	
1.1.1 计算机系统的组成和特性.....1	
1.1.2 数据表示.....3	
1.1.3 计算机算术运算和逻辑运算.....8	
1.1.4 CPU 基本构造及功能.....9	
1.1.5 指令系统与 RISC 技术.....11	
1.1.6 存储系统.....15	
1.1.7 外围设备与 I/O 技术.....20	
1.1.8 并行处理技术.....27	
1.1.9 计算机安全性、可靠性及 系统性能评价.....30	
1.2 例题分析.....34	
1.3 思考练习题及答案.....78	
1.3.1 思考练习题.....78	
1.3.2 思考练习题答案.....85	
第 2 章 程序设计语言基础86	
2.1 内容提要.....86	
2.1.1 程序语言的知识.....86	
2.1.2 语言处理程序概述.....87	
2.1.3 构造编译程序的基本 知识和基本原理.....88	
2.2 例题分析.....88	
2.3 思考练习题及答案.....114	
2.3.1 思考练习题.....114	
2.3.2 思考练习题答案.....118	
第 3 章 操作系统知识120	
3.1 内容提要.....120	
3.1.1 操作系统基础知识.....120	
3.1.2 处理机管理.....121	
3.1.3 存储管理.....122	
3.1.4 设备管理.....123	
3.1.5 文件管理.....124	
3.1.6 作业管理.....125	
3.1.7 网络与嵌入式操作系统 基础知识.....126	
3.1.8 UNIX 操作系统实例.....126	
3.2 例题分析.....126	
3.3 思考练习题及答案.....153	
3.3.1 思考练习题.....153	
3.3.2 思考练习题答案.....158	
第 4 章 系统开发与运行160	
4.1 内容提要.....160	
4.1.1 软件工程概述.....160	
4.1.2 系统分析.....161	
4.1.3 需求分析.....162	
4.1.4 软件设计.....163	
4.1.5 软件实现.....165	
4.1.6 程序设计.....166	
4.1.7 软件测试.....167	
4.1.8 系统转换.....169	
4.1.9 软件维护.....170	
4.1.10 软件工程管理.....171	
4.1.11 软件质量保证.....175	
4.1.12 软件过程及过程能力 评估模型.....177	
4.1.13 软件开发工具与环境.....180	
4.2 例题分析.....180	

4.3	思考练习题及答案	229	6.3	思考练习题及答案	333
4.3.1	思考练习题	229	6.3.1	思考练习题	333
4.3.2	思考练习题答案	244	6.3.2	思考练习题答案	338
第 5 章	网络基础知识	247	第 7 章	数据库技术基础	340
5.1	内容提要	247	7.1	内容提要	340
5.1.1	网络概述	247	7.1.1	数据库管理系统的 功能和特征	340
5.1.2	计算机网络的体系结构	248	7.1.2	数据库模型	341
5.1.3	网络的传输控制	248	7.1.3	数据模型和 E-R 图	342
5.1.4	常用网络设备及特点	249	7.1.4	关系数据库规范化理论	345
5.1.5	客户机/服务器模式与 网络计算	249	7.1.5	数据操作	347
5.1.6	局域网 (LAN) 技术	250	7.1.6	关系数据库 SQL	349
5.1.7	Internet 基础及应用	251	7.1.7	数据库的保护	353
5.1.8	网络软件	254	7.1.8	数据仓库和分布式数据库 基础知识	355
5.1.9	网络安全	254	7.2	例题分析	357
5.1.10	网络管理	254	7.3	思考练习题及答案	371
5.2	例题分析	255	7.3.1	思考练习题	371
5.3	思考练习题及答案	287	7.3.2	思考练习题答案	379
5.3.1	思考练习题	287	第 8 章	数据结构	382
5.3.2	思考练习题答案	295	8.1	内容提要	382
第 6 章	多媒体基础知识	296	8.1.1	线性表	382
6.1	内容提要	296	8.1.2	栈	384
6.1.1	多媒体的概念	296	8.1.3	队列	387
6.1.2	多媒体计算技术	297	8.1.4	数组	389
6.1.3	多媒体存储技术	299	8.1.5	字符串	393
6.1.4	多媒体网络应用	301	8.1.6	树与二叉树	395
6.1.5	多媒体内容编辑语言	304	8.1.7	图	401
6.2	例题分析	305	8.1.8	排序	409
6.2.1	多媒体的概念	305	8.1.9	查找	414
6.2.2	多媒体计算技术	306	8.2	例题分析	419
6.2.3	多媒体存储技术	327	8.3	思考练习题及答案	443
6.2.4	多媒体网络应用	329			
6.2.5	多媒体内容编辑语言	331			

第 9 章 常用算法设计方法	477	11.1.4 标准的代号和编号	576
9.1 内容提要	477	11.1.5 国际标准和国外先进标准	577
9.1.1 迭代法	477	11.1.6 信息技术标准化	579
9.1.2 穷举搜索法	479	11.1.7 标准化组织	580
9.1.3 递推法	479	11.1.8 ISO 9000 标准简介	582
9.1.4 递归法	480	11.1.9 ISO/IEC 15504 过程	
9.1.5 回溯法	481	评估标准简介	582
9.1.6 贪婪法	482	11.2 知识产权内容提要	582
9.1.7 分治法	482	11.2.1 知识产权的概念与特点	583
9.1.8 动态规划法	483	11.2.2 我国保护软件知识	
9.2 例题分析	483	产权的法律法规	584
9.3 思考练习题及答案	504	11.2.3 计算机软件著作权保护	586
9.3.1 思考练习题	504	11.2.4 计算机软件商业秘密	
9.3.2 思考练习题答案	508	法律保护	594
第 10 章 面向对象技术	512	11.3 例题分析	596
10.1 内容提要	512	11.4 思考练习题及答案	601
10.1.1 面向对象方法的概述	512	11.4.1 思考练习题	601
10.1.2 面向对象软件的开发过程	513	11.4.2 思考练习题答案	602
10.1.3 面向对象的分析和设计建模	515	第 12 章 C/C++ 语言程序设计	604
10.1.4 面向对象的实现和测试	523	12.1 内容提要	604
10.1.5 设计模式	528	12.1.1 C 程序的构成	604
10.2 例题分析	530	12.1.2 变量的定义	606
10.3 思考练习题及答案	561	12.1.3 数据类型	606
10.3.1 思考练习题	561	12.1.4 算术表达式	608
10.3.2 思考练习题答案	571	12.1.5 赋值表达式	608
第 11 章 标准化与知识产权		12.1.6 ++、-- 和逗号运算符	608
基础知识	572	12.1.7 三目运算符	609
11.1 标准化内容提要	572	12.1.8 输入输出	609
11.1.1 标准化的基本概念	572	12.1.9 选择结构 if	612
11.1.2 标准化过程模式	573	12.1.10 switch 语句	613
11.1.3 标准的分类	574	12.1.11 标号语句和 goto 语句	614
		12.1.12 while 语句	614
		12.1.13 do-while 语句	614
		12.1.14 for 语句	615

12.1.15	continue 和 break 语句	615	12.3.1	思考练习题	662
12.1.16	字符型数据	616	12.3.2	思考练习题答案	671
12.1.17	文件引用	616			
12.1.18	宏定义	617	第 13 章	软件系统分析与设计	676
12.1.19	函数	618	13.1	结构化分析与设计	676
12.1.20	数组	620	13.1.1	软件需求规格说明	676
12.1.21	指针	622	13.1.2	结构化分析	679
12.1.22	字符串	623	13.1.3	结构化设计	689
12.1.23	函数的进一步讨论	624	13.1.4	详细设计	691
12.1.24	作用域和存储类型	625	13.2	面向对象分析与设计	693
12.1.25	结构与联合	626	13.2.1	面向对象分析与设计	693
12.1.26	位运算	628	13.2.2	用例建模	695
12.1.27	文件操作	628	13.2.3	类与对象建模	700
12.1.28	C++简介	629	13.2.4	交互建模	704
12.1.29	关于 C++ 的几个基本 问题	630	13.2.5	行为建模	711
12.1.30	类	632	13.3	数据库分析与设计	719
12.1.31	函数重载	633	13.3.1	数据库设计的步骤	719
12.1.32	操作符重载	634	13.3.2	需求分析	720
12.1.33	类的继承和派生	634	13.3.3	概念结构设计	721
12.1.34	模板	640	13.3.4	逻辑结构设计	722
12.1.35	异常处理	643	13.3.5	数据库的物理设计	724
12.2	例题分析	644	13.3.6	数据库实施与维护	726
12.3	思考练习题及答案	662	13.3.7	案例分析	727
			13.4	思考练习题及答案	731

第 1 章 计算机系统知识

计算机系统包括硬件系统和软件系统。

计算机系统结构是程序设计人员看到的计算机系统的外部属性，即系统的功能和结构特性。计算机组成主要研究计算机的内部属性，计算机各部件的功能和联系，以及组成工作原理。计算机软件系统知识，在以后各章中叙述。

本章内容根据考试大纲的要求和历届考试命题的情况制定。主要包括数据表示、指令系统、中央处理器、存储系统、I/O 技术、并行处理、计算机安全性、可靠性和系统性能评价。

1.1 内容提要

- ① 计算机系统的组成和特性，层次结构及分类方法。
- ② 计算机中数据表示方法，数制、码制、浮点数表示法及数据校验码。
- ③ CPU 基本构造及功能。
- ④ 指令系统及 RISC 技术。
- ⑤ 存储系统，包括 Cache 及虚拟存储器工作原理。
- ⑥ 外围设备，包括接口、总线及 I/O 控制方式。
- ⑦ 计算机安全性、可靠性及系统性能评价。

1.1.1 计算机系统的组成和特性

1. 计算机系统构成

根据冯·诺依曼提出的经典的计算机结构模型，计算机由运算器、控制器、存储器、输入装置和输出装置五大部件构成，运算器为中心，被运算的数据用二进制数表示，机器指令包括操作码和地址码，也用二进制数编码表示。计算程序和数据都放在存储器中，存储单元地址采用线性编址，使用程序计数器 PC 决定程序中指令执行的顺序。冯·诺依曼结构计算机是基于控制驱动的存储程序计算机。

虚拟计算机：人根据逻辑设计使用各种电子器件研制成功的计算机称为物理机器，是实实在在的硬件计算机。人们使用机器语言（二进制指令）与物理机器打交道。计算机系统包含硬件系统和软件系统。不同的用户使用不同层次的软件语言与计算机交往，可以与不同层次语言的虚拟计算机交往。这种机器实际上是不存在的，都必须通过编译程序等翻译成机器语言，才能在物理机器上运行。

2. 计算机系统的层次结构

用户在不同层次使用不同语言与计算机打交道，均可实现程序要求，故可以把计算机看作一个多层次的系统。

第一层（核心层）是物理机器。人们使用二进制机器指令与机器交往。

第二层是操作系统级机器。操作系统用于管理计算机的软件和硬件资源。人们通过系统调用，方便有效地使用和管理计算机资源，把这个层次叫做操作系统级机器，也可叫做操作系统虚拟机。

第三层是汇编语言机器。人们使用容易记忆的符号表示的指令（汇编语言）与机器打交道。当然，最终执行这些汇编语言时还要通过汇编器把汇编语言翻译成机器语言才能在物理机器上执行。可把这层看作汇编语言虚拟机。

第四层是高级语言机器。人们使用高级语言与机器打交道。运行程序时，首先通过编译程序把高级语言翻译成机器语言才能执行，这层也可叫做高级语言虚拟机。

3. 计算机的特性

50年来，计算机在各个领域得到了广泛应用，其主要原因和技术关键在于以下几点。

- ① 由于采用高速电子开关电路，决定了计算机高速度运算的功能。
- ② 由于采用二进制数表示被运算的数据，决定了机器具有非常高的运算精度。
- ③ 逻辑运算能力决定了计算机的控制功能和人工智能领域的应用。
- ④ 数字化信息编码技术，决定了计算机处理文字、语音、图像的通用性。
- ⑤ 由于存储程序技术的采用，决定了自动执行程序的能力和计算的快速性。

因此通常讲的计算机，准确地说应该称为电子数字计算机。

4. 计算机的分类

① 按照计算机逻辑电路采用的器件划分，计算机可分为电子管、晶体管、集成电路和大规模集成电路（一、二、三、四代）计算机。

② 按照机器的功能和规模可将计算机划分为巨、大、中、小、微各种类型。

③ 按照计算机的应用领域，可将计算机分成通用机和专用机。例如工作站、服务器，嵌入式计算机，以及图像处理计算机、控制计算机等。

④ 由于并行处理技术的发展，按照计算机系统中处理机的数目和作用可分为单处理机、多处理机、并行处理机、超标量处理机、超流水线处理机、MPP 大规模并行处理机和机群系统等。

⑤ 1966年 M.J.Flynn 提出了按照指令流和数据流的不同组织方式的 Flynn 分类法，计算机系统可以分为以下几类。

SISD：单指令流单数据流计算机，即普通的计算机。

SIMD：单指令流多数据流的计算机，如阵列处理计算机。

MISD：多指令流单数据流计算机，有人把流水线机器划归这一类。

MIMD：多指令流多数据流计算机，即多处理机系统。

⑥ 1972年冯泽云按照最大并行度的概念，提出了冯氏分类法，他把计算机的字长和位片宽度（每次处理的字数）作为计算机分类的主要指标，最大并行度指计算机系统单位时间内能够同时处理的二进制位数。

字串位串 WSBS 为第一代串行计算机，计算机中对各个字的处理是串行进行的，每个字的各位二进制数又是串行进行的，速度很慢。

字并位串 WPBS 指传统计算机，计算机一个字的各位二进制数是并行运算的，但各个字是串行进行的，每次只处理一个字。

字串位并 WSBP 处理机的字宽是1位，但同时可处理多个字，如 STARAN、MPP 计算机等。

字并位并 WPBP 处理机字长各位二进制数同时运算，同时机器可对多个字并行处理，如 ILLIACIV、Cmmp 等。

在技术和应用的推动下，计算机的系统结构取得了很大的发展，例如采用数据驱动的数据流计算机，面向函数式语言的归约计算机等，它们都是非诺依曼计算机，但计算机的主流和基础仍是冯·诺依曼结构。

1.1.2 数据表示

1. 内容要点

1) 数据分类

计算机中处理的数据有两类：数值数据和非数值数据。

数值数据指表示数量的数据，有正负和大小之分，在计算机中的数据表示为二进制的形式进行运算和存储。

非数值数据包括字符、汉字、声音和图像等，在计算机中处理前必须用某种编码形式表示并转换成二进制数。

2) 数制

常用的十进制数的计数法则是：表示一位数有10个不同的符号（0, 1, 2, 3, 4, 5, 6, 7, 8, 9），相邻数位之间的关系是“逢十进一”（各位数的位权是 10^n ），它所表示的数值是各位数按权展开的和，十进制数中的10称为该计数制的基数。

二进制数的基数为2，表示一位数有两个符号：0和1，相邻数位间之进位关系为“逢二进一”，表示的十进制数值为各位数按权展开的和。

为了方便，计算机中二进制数常用八进制数、十六进制数表示，应熟记八进制数、十六进制数与二进制数的对应关系。

BCD 码：每位十进制数用4位二进制数表示，又可分为有权码和无权码两种方案。

3) 二进制数与其他记数制数的转换方法

二进制数转换为十进制数是把二进制数各位按权展开求和。

十进制数转换为二进制数时分整数和小数两部分，分开进行转换然后相加。十进制整

数部分采用“除 2 取余”法，直到商数为 0，最后得到的余数是二进制数的最高位。十进制小数部分采用“乘 2 取整”的方法，首先得到的整数部分为转换成二进制小数的最高位，直到要求的精度。

以此类推，十进制数转换成任意进制数（如 r 进制）的方法，整数部分为“除 r 取余”，小数部分为“乘 r 取整”。

4) 机器数

各种数据在计算机中表示的形式称为机器数，其特点是数的符号用 0、1 表示，如 0 表正数，1 表负数。小数点隐含表示不占位置。机器数对应的实际数值称为该数的真值。

机器数有两种：无符号数和带符号数。

无符号数表示正数，没有负数，机器数中无符号位。

无符号整数的小数点，固定在该数最低位之后，是纯整数。

无符号小数的小数点固定在该数最高位之前，是纯小数。

带符号机器数的最高位是表示正、负数的符号位，其余为数值位。带符号整数的小数点固定在该数最低位之后，是纯整数。带符号小数的小数点固定在该数最高位（符号位）之后，最高数值位之前，是纯小数。这种表示数的方式称为定点数。

浮点数小数点的位置不是固定的，如 $N = MR^E$ 。小数点位置用阶码 E 表示， E 为定点整数。

浮点数的数值部分用尾数 M 表示， M 为定点小数， R 为阶 E 的底数，在一个指定的机器中 R 是固定的，在浮点数表示中不再出现。

5) 定点数据编码方法

为了运算方便，带符号的机器数有不同的编码方法，称为码制。

① 原码：又称符号绝对值码。

该数最高位为符号位，正数用 0 表示，负数用 1 表示。其他位为数据位，用二进制数绝对值表示。原码与真值转换方便，但作加减运算不便，且零有 +0 和 -0 两种表示方法。

② 反码：正数的反码表示与原码相同。

负数的反码，符号位用 1 表示，数值位由其绝对值各位取反得到。

反码零也有 +0 和 -0 两种表示方法。因运算不便使用较少。

③ 补码：为了加减运算方便引入补码概念。关键思想是用加法代替减法。

正数的补码与原码表示相同。

负数的补码，符号位用 1 表示，数值位用其绝对值的补数表示（即原码各位求反，末位加 1）。

补码最大的优点是作加减运算方便，如 $(x + y)_{补} = (x)_{补} + (y)_{补}$ ， $(x - y)_{补} = (x)_{补} + (-y)_{补}$ 。

补码符号位参加运算，不单独处理，现代计算机中大都采用补码系统。补码的另一优点是零的表示形式是唯一的，即 $(+0000)_{补} = (-0000)_{补} = 00000$ 。

补码的表数范围比原码、反码略宽。在定点小数中，补码可以表示 -1。

$(-1)_{\text{补}} = 1.0000$ ，而原码、反码不能表示绝对值为1的数。

④ 移码：为了比较两个整数的大小，引入移码的概念。

移码与补码有类似的地方（数值部分），但符号位与补码相反，即正数的移码符号位为1，负数的移码符号位为0。或者说求一个数的移码，先求其补码再将其符号位变反即可得到。移码表数范围与补码整数的表数范围相同。

6) 非数值数据（符号数据）

英文字符编码的国际标准是ASCII码。用7位二进制数表示，可表示128个符号。扩展的二/十进制交换码EBCDIC，采用8bits表示一个字符，可表示 $2^8 = 256$ 个不同符号。

汉字编码有很多种方法。常用的数字编码方式是区位码，将常用汉字分成94个区，每个区又分94位，每个汉字的区位编号用两个字节十进制数表示。

拼音码和字形码也是汉字常用输入编码方法。

汉字国标码也是数字编码，它是汉字信息交换码国家标准，与区位码一一对应，但区位号用十六进制数表示，且第一个汉字放在 $(21)_{16}$ 区 $(21)_{16}$ 位。

计算机内存储汉字的编码方法与输入编码不同，通常用两个字节汉字国标码表示，为与ASCII码区分，将每个字节最高位置1表示汉字字符，而ASCII码最高位为0，低7位表示其编码，这种汉字编码称为汉字机内编码，简称内码。

汉字输出时通过内码找到其对应字模码（点阵字型）逐点输出点阵字形。如果一个汉字用 16×16 点阵表示，则每个汉字要占2字节 $\times 16 = 32$ 字节，两级汉字共6763个字模，占用大量存储空间。

7) 数据校验方法

计算机中的数据在传送、存储过程中可能出错，为了及时发现和纠正错误，编码中引入差错检查机制。

常用的校验编码有奇偶校验码、海明校验码和CRC循环冗余校验码。

奇偶校验是最常用的校验方法，可以发现一位错或奇个数位同时出错。

2. 学习难点

1) 码制

码制是为了运算方便而提出的数值数据的编码方法。

① 4种码制中，正数的原码、补码、反码的表示都是一样的，即符号位为0，数值位为其绝对值。而正数的移码表示中，虽然其数值也是其绝对值，但符号位相反，即正数的符号位为1，负数的符号位为0。

② 负数的原码、反码、补码的符号位都为1，但数值位表示方法是不同的。

原码的数值位为该数的绝对值。

反码的数值位为该数的绝对值每位求反。

补码的数值位为该数绝对值的补数，即其绝对值每位求反，末位加1。

负数的移码，其数值位与补码相同，即该数绝对值的补数，但符号位与补码不同（与

原码反码也不同), 即负数移码的符号位为 0。

③ 4 种编码中零的表示不同。

$$(+0)_{\text{原}} = 0000 \quad (-0)_{\text{原}} = 1000 \quad (+0)_{\text{反}} = 0000 \quad (-0)_{\text{反}} = 1111$$

$$(+0)_{\text{补}} = 0000 \quad (-0)_{\text{补}} = 0000 \quad (+0)_{\text{移}} = 1000 \quad (-0)_{\text{移}} = 1000$$

④ 4 种编码表示数的范围不同, 以 8 位二进制带符号整数 x 为例表示如下。

$$11111111 \leq (x)_{\text{原}} \leq 01111111, \quad -(2^7-1) \leq (x)_{\text{原}} \leq (2^7-1),$$

$$\text{即 } -127 \leq (x)_{\text{原}} \leq +127;$$

$$10000000 \leq (x)_{\text{反}} \leq 01111111, \quad -(2^7-1) \leq (x)_{\text{反}} \leq (2^7-1),$$

$$\text{即 } -127 \leq (x)_{\text{反}} \leq +127;$$

$$10000000 \leq (x)_{\text{补}} \leq 01111111, \quad -2^7 \leq (x)_{\text{补}} \leq (2^7-1),$$

$$\text{即 } -128 \leq (x)_{\text{补}} \leq +127;$$

注意: 此时 $(x)_{\text{补}}$ 最小值是 $-2^7 = -128$, 而不是 $-(2^7-1) = -127$ 。

$$00000000 \leq (x)_{\text{移}} \leq 11111111, \quad -2^7 \leq (x)_{\text{移}} \leq (2^7-1),$$

$$\text{即 } -128 \leq (x)_{\text{移}} \leq +127。$$

2) 浮点数

n 位定点数的补码或移码可表示 2^n 个数, 而其原码、反码只能表示 $2^n - 1$ 个数, 表数范围小。

浮点数是小数点位置不固定的数, 同时为了扩大定点数的表示范围, 引入浮点数概念。任意浮点数 N 可用阶码 E 和尾数 M 两个部分来表示: $N = MR^E$ 。

M 称为浮点数尾数, 表示其数值的有效数字, 是定点小数。

E 称为浮点数的阶码, 表示小数点的位置, 是定点整数。

R 是阶码的底。在浮点数表示中 R 是固定不变的, 隐含表示, 通常取 $R = 2$, 则 $N = M \times 2^E$ 。机器中只需给出 E 和 M , 即可知该数之数值。 E 和 M 都是定点数, 也可分别指定其为原码、反码、补码、移码, 同样可求出其表数范围。需要注意, 对于浮点数 N , 有以下几种情况。

① 最大正数: M 是最大正数, E 是最大正数。

② 最小正数: M 是最小正数, E 是最小负数。

③ 最大负数: M 是最大负数, E 是最小负数。

④ 最小负数: M 是最小负数, E 是最大正数。

注意补码、移码的最小负数与原码、反码是不同的。计算机中一个数的浮点数有许多种表示方法。为了规范, 规定在运算结束将运算结果存到机器中时, 必须是规格化的浮点数, 即浮点数尾数的最高数值位是有效数字, 即 $1/2 \leq |M| < 1$ 。

3) 海明校验码

常用于发现纠正 1 位数据出错。其编码规则是在 n 位被校验数据位间, 插入 K 个校验位, 其校验位之个数满足关系 $2^k - 1 \geq n + k$; 校验位在海明码中的位置是固定的, 即海明码的

1, 2, 4, 8...位; 一个校验位可校验多个数据位, 每个校验位的取值等于其被校验数据位之和。其中被校数据位海明位号等于各校验位海明位号之和。

当某个数据位出错时, 则引起有关的校验位改变; 当所有海明位均正确时, 则有关的校验值全为 0。

当某个校验位出错时, 则有关的校验值只有一位不为 0, 且其编码为该出错校验位之海明位号; 当某个数据位出错时, 则有关的校验值有两位或 3 位不为 0, 且其编码为该出错数据位之海明位号。

纠正错误时, 只要将出错位变反即可, 因此可以自动纠正 1 位错。发现多位数据出错或纠正多位出错的情况要复杂得多。

4) CRC 循环冗余校验码

用于发现和纠正信息传送过程中连续出现的多位错误。CRC 码是指在 k 位被校验数据之后拼接 r 位校验码, 得到 $k+r$ 位编码。需设计一种算法, 使得发送方根据 k 位数据算出 r 位校验位之值, 一起传给对方; 接收方根据同一算法对 $k+r$ 位数据进行校验, 即可判断传送是否出错。关键是找出这种算法。

我们把一个二进制代码看成一个多项式 $M(x)$ 按 x 的降幂排列的多项式的系数。例如, 可将 1101 看成 $1x^3+1x^2+0x^1+1x^0$ 的各个系数, 一个 4 位二进制代码看作一个三次多项式, 一个 n 位代码可看作 $n-1$ 次多项式。代码左移 3 位相当于多项式乘以 x^3 , 如 $(x^3+x^2+1)x^3 = x^6+x^5+x^3$ 。

一个多项式除以另一个多项式, 其商和余数也都是多项式, 余数多项式次数比除数多项式的次数少 1。对多项式系数的运算按模二运算进行, 运用模二运算时各位数据间没有进位关系。

为了产生 CRC 码的 r 位校验码, 选择一个生成多项式 $G(x)$, $G(x)$ 是一个 r 次的多项式。

如果被校验的代码为 k 位二进制数据, 可用 $M(x)=M_{k-1}M_{k-2}\cdots M_1M_0$ 来表示; k 位数据后增加 r 位校验位, 则将 $M(x)$ 乘以 x^r , 左移 r 位, 再除以 $G(x)$, 所得余数多项式 $R(x)$, 为 $r-1$ 次多项式, 余数相应代码为 r 位; 将该 r 位代码接在左移 r 位的 $M(x)$ 后面得到 $M(x)x^r+R(x)$, 它是 $G(x)$ 的倍式, 可以被 $G(x)$ 整除。其各位系数即构成 $k+r$ 位 CRC 循环冗余校验码。接收方收到这 $k+r$ 位信息后用同样的生成多项式 $G(x)$ 去除, 如果传送信息完全正确, 则应能除尽, 余数为 0。如果余数不为 0, 则说明传送出错, 并根据不同的余数判断哪一位出错。将该位变反, 即可纠正错误。显然生成的多项式 $G(x)$ 是经过严格挑选的, 它具有以下性质。

- 任何一位出错, CRC 码除以 $G(x)$ 的余数不为 0。
- 不同的码位出错余数不能相同。
- 给定生成多项式后, 余数与出错码位之间对应关系不变, 与被校验数据无关。
- 对余数继续作模二除法运算, 应使余数循环。

1.1.3 计算机算术运算和逻辑运算

1. 内容要点

(1) 定点加减法运算。

一般采用补码加减法实现，对于定点小数，

加法： $(x + y)_{\text{补}} = (x)_{\text{补}} + (y)_{\text{补}} \bmod 2$ ；

减法： $(x - y)_{\text{补}} = (x)_{\text{补}} + (-y)_{\text{补}} \bmod 2$ 。

定点数的模数可看作最高位（符号位）之进位，定点小数之模为 2。

定点数相加减，可能出现溢出，它不是最高位之进位。

溢出判断方法有两种：双符号位法和进位判别法。双符号位法，数的符号位用两位二进制数表示：00 为正数，11 为负数，两个符号不相同同时溢出。

进位判别法判断结果溢出，要求符号位进位 C_n 与次高位（最高数据位）进位 C_{n-1} 之中有一个有进位，但不能同时有进位，即 $C_n \oplus C_{n-1} = 1$ 。

(2) 定点数乘法用原码方便，结果的符号为运算两数符号之异或，积（或商）为两数绝对值之积（或商）。

(3) 浮点数加减法运算步骤。

① 对阶，参加运算两数阶码必须相同。

② 尾数加减。

③ 尾数规格化。

④ 舍入。

⑤ 溢出处理，判断阶码是否溢出。

(4) 浮点数的乘除运算。

浮点数的乘除运算：浮点数相乘，其积的阶码为两数阶码之和，积的尾数为两尾数之积；浮点数相除，其商的阶码为两数阶码之差，商的尾数为两尾数之商。其结果均需规格化。

(5) 逻辑代数与基本逻辑运算。

① 必须熟记基本逻辑运算常用公式，对逻辑表达式化简很有用。

② 逻辑运算（与、或、非、异或）都是对应位之间的运算，相邻数位没有进位关系。

2. 学习难点

1) 浮点加减法运算

① 首先必须完成对阶操作。即当两数阶码不同时，要把两数的阶码都变成大阶，此时阶码小的数变大，为了与原数保持相等，必须把该数之尾数右移两数阶差那么多位。

② 浮点数两尾数相加减后必须对结果进行规格化，尾数为定点小数，所谓规格化即要求尾数的绝对值要大于等于 0.5，且小于 1。

在补码表示方法中，要求尾数的符号位与数值最高位不同，即 $0.1 \times \dots \times$ 或 $1.0 \times \dots \times$ 。