

全国计算机技术与软件专业技术资格（水平）考试参考用书

数据库系统工程师 考试全程指导

全国计算机技术与软件专业技术资格（水平）考试办公室推荐
丁宝康 陈坚 主编 施伯乐 主审

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）考试参考用书

数据库系统工程师 考试全程指导

全国计算机技术与软件专业技术资格（水平）考试办公室推荐
丁宝康 陈坚 主编 施伯乐 主审

清华大学出版社
北京

内 容 简 介

本书是全国计算机技术与软件专业技术资格(水平)考试办公室组织编写的考试辅导用书。本书遵循考试大纲的要求,对数据库系统工程师级别考试应该必备的知识、技能和习题进行了详细的解析。

全书共分16章,内容包括:计算机系统知识、数据结构与算法、操作系统、程序设计语言、网络、多媒体、数据库基本概念、关系数据库理论、SQL语言、系统开发与运行、数据库设计、数据库运行与管理、网络与数据库、数据库新技术、信息化与知识产权、标准化知识。

本书旨在帮助考生快速把握考试的重点和难点,熟悉试题的形式,掌握解答问题的方法和技术。本书不仅是数据库系统工程师级别考试的学习用书,也可作为各类信息技术、数据库系统培训和辅导的教材,还可作为从事数据库领域工作的科技人员的参考书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

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。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

数据库系统工程师考试全程指导 / 丁宝康, 陈坚主编. —北京: 清华大学出版社, 2006.4

(全国计算机技术与软件专业技术资格(水平)考试参考用书)

ISBN 7-302-12657-7

I. 数… II. ①丁… ②陈… III. 数据库系统-工程技术人员-资格考核-自学参考资料 IV. TP311.13

中国版本图书馆 CIP 数据核字(2006)第 018886 号

出 版 者: 清华大学出版社 地 址: 北京清华大学学研大厦
http://www.tup.com.cn 邮 编: 100084
社 总 机: 010-62770175 客 户 服 务: 010-62776969

组稿编辑: 柴文强

文稿编辑: 薛 阳

印 刷 者: 清华大学印刷厂

装 订 者: 三河市春园印刷有限公司

发 行 者: 新华书店总店北京发行所

开 本: 185×230 印张: 49.75 防伪页: 1 字数: 1146 千字

版 次: 2006年4月第1版 2006年5月第2次印刷

书 号: ISBN 7-302-12657-7/TP·8091

印 数: 3001~6000

定 价: 65.00元

前 言

自 2004 年起, 全国计算机技术与软件专业技术资格(水平)考试执行人事部、信息产业部新的考试大纲。“数据库系统工程师”是为该考试的信息系统专业中级资格层次设置的一个资格名称。通过数据库系统工程师级别考试的学生, 即可被用人单位择优聘任为工程师职务。

据查, 这两年参加数据库系统工程师考试的考生为数不少。这是由于“数据库”是计算机得到广泛应用的两大基础之一(另一个是“网络”), 所有的计算机应用系统都与数据有关, 要建立数据库。

在考试大纲中, 要求考生以掌握数据库系统的知识为主, 同时适当地掌握硬件、数据结构、操作系统、程序设计语言、网络、多媒体、信息化与知识产权、标准化等基础知识。通过本考试的合格人员将具有较强的能力参与信息系统的开发, 能主持数据库结构的设计和数据库应用系统的开发, 并具有工程师的实际工作能力和业务水平。

为了满足广大考生的需要, 我们组织了参与过多年资格考试命题或辅导的教师, 以新的考试大纲为依据, 编写了《数据库系统工程师考试全程指导》这本书。由于要求考生了解、熟悉和掌握的内容很多, 而考生都已有一定的基础, 因此, 本书从总结和提高的角度出发, 对涉及的知识点加以阐述。本书的每一章由基本要求、基本内容、重点习题解析和模拟试题 4 部分组成。本书的宗旨是尽可能帮助考生快速把握考试的重点、难点和关键点, 熟悉试题的形式, 掌握解答问题的方法和技巧, 以在考试中取得良好的成绩。

本书共分 16 章, 第 1、4、6 章由乔健编写, 第 2 章由邓桂英编写, 第 3 章由于玉编写, 第 5、10、13 章由周敏子编写, 第 7、8、9、11 章由丁宝康编写, 第 12、14 章由陈坚编写, 第 15、16 章由许建军编写。还有夏根女、陈长洪、薛剑虹、王晓雯、徐美娟、曾宇昆、杨卫稼等老师参加了编写工作。全书由丁宝康、陈坚担任主编, 负责组稿、修改和统稿。

复旦大学施伯乐教授对本书的编写进行了指导, 并审阅了全稿, 提出了许多宝贵的意见, 我们在此表示衷心的感谢。

在本书的编写过程中, 我们还参考了计算机同仁的许多相关书籍和资料, 在此对这些参考文献的作者也表示由衷的感谢。

限于水平, 书中存在不足之处, 敬请广大专家和读者批评指正。对本书的意见请按电子邮件地址 dn@citiz.net 反馈给我们, 非常感谢!

丁宝康 陈 坚

2005 年 7 月

目 录

第 1 章 计算机系统知识1	
1.1 基本要求.....1	
1.2 基本内容.....1	
1.2.1 计算机系统的组成.....1	
1.2.2 计算机的基本工作原理.....4	
1.2.3 计算机体系结构.....8	
1.2.4 存储系统.....11	
1.2.5 指令系统.....20	
1.2.6 输入输出技术.....21	
1.2.7 流水线操作.....23	
1.2.8 总线结构.....25	
1.2.9 多处理机与并行处理.....27	
1.2.10 计算机安全性技术.....28	
1.2.11 计算机可靠性模型.....37	
1.2.12 计算机系统的性能评价与 故障诊断.....38	
1.2.13 小结.....39	
1.3 重点习题解析.....40	
1.3.1 填空题.....40	
1.3.2 简答题.....42	
1.3.3 选择题.....43	
1.4 模拟试题.....46	
第 2 章 数据结构与算法55	
2.1 基本要求.....55	
2.2 基本内容.....56	
2.2.1 数据结构与算法概念.....56	
2.2.2 线性表.....56	
2.2.3 栈和队列.....58	
2.2.4 串.....59	
2.2.5 数组和广义表.....60	
2.2.6 树和二叉树.....61	
2.2.7 图.....65	
2.2.8 查找.....69	
2.2.9 排序.....75	
2.2.10 常见算法设计方法.....81	
2.2.11 小结.....84	
2.3 重点习题解析.....84	
2.3.1 判断题.....84	
2.3.2 填空题.....86	
2.3.3 简答题.....88	
2.3.4 选择题.....89	
2.4 模拟试题.....90	
第 3 章 操作系统知识99	
3.1 基本要求.....99	
3.2 基本内容.....100	
3.2.1 基础知识.....100	
3.2.2 进程描述与控制.....111	
3.2.3 进程互斥与同步.....115	
3.2.4 存储器管理.....128	
3.2.5 处理机管理.....142	
3.2.6 设备管理.....147	
3.2.7 文件系统.....150	
3.2.8 死锁.....160	
3.2.9 网络操作系统和嵌入式 操作系统的基础知识.....161	
3.2.10 小结.....163	
3.3 重点习题解析.....164	
3.3.1 选择题.....164	
3.3.2 简答题.....166	

3.4 模拟试题	168	6.2.2 音频	257
第 4 章 程序设计语言基础	171	6.2.3 图形和图像	262
4.1 基本要求	171	6.2.4 动画和视频	268
4.2 基本内容	171	6.2.5 多媒体网络	271
4.2.1 程序设计语言的基础知识	171	6.2.6 多媒体计算机系统	272
4.2.2 汇编程序的基本原理	178	6.2.7 虚拟现实	274
4.2.3 编译程序的基本原理	179	6.2.8 小结	275
4.2.4 解释程序的基本原理	203	6.3 重点习题解析	275
4.2.5 小结	204	6.3.1 填空题	275
4.3 重点习题解析	205	6.3.2 简答题	277
4.3.1 填空题	205	6.3.3 选择题	278
4.3.2 简答题	206	6.4 模拟试题	281
4.3.3 选择题	208	第 7 章 数据库技术基础	284
4.4 模拟试题	210	7.1 基本要求	284
第 5 章 网络基础知识	215	7.2 基本内容	284
5.1 基本要求	215	7.2.1 数据管理技术的发展阶段	284
5.2 基本内容	215	7.2.2 数据库的基本术语	288
5.2.1 网络概述	215	7.2.3 数据描述	288
5.2.2 ISO/OSI 网络体系结构	218	7.2.4 数据模型	292
5.2.3 网络的协议与标准	222	7.2.5 数据库的体系结构	298
5.2.4 网络设备和介质	231	7.2.6 数据库管理系统	303
5.2.5 网络搭建	236	7.2.7 数据库系统	305
5.2.6 Internet 及应用	239	7.2.8 数据库的控制功能	313
5.2.7 网络安全	246	7.2.9 小结	328
5.2.8 小结	249	7.3 重点习题解析	329
5.3 重点习题解析	249	7.3.1 填空题	329
5.3.1 填空题	249	7.3.2 简答题	331
5.3.2 简答题	251	7.3.3 多项选择题	335
5.4 模拟试题	252	7.4 模拟试题	336
第 6 章 多媒体基础知识	256	第 8 章 关系数据库基本理论	342
6.1 基本要求	256	8.1 基本要求	342
6.2 基本内容	256	8.2 基本内容	343
6.2.1 多媒体的基本概念	256	8.2.1 关系模型的基本概念	343
		8.2.2 关系代数	346

8.2.3	关系演算	353	9.5.2	设计题	480
8.2.4	关系代数表达式的优化	357	第 10 章 系统开发与运行 484		
8.2.5	关系模式的设计准则	363	10.1	基本要求	484
8.2.6	函数依赖	365	10.2	基本内容	484
8.2.7	关系模式的分解特性	370	10.2.1	软件工程基础知识	484
8.2.8	范式	376	10.2.2	软件项目管理知识	489
8.2.9	其他数据依赖和范式	381	10.2.3	系统分析基础知识	495
8.2.10	小结	385	10.2.4	系统设计知识	505
8.3	重点习题解析	387	10.2.5	系统实施知识	512
8.3.1	填空题	387	10.2.6	系统测试与调试	515
8.3.2	简答题	388	10.2.7	系统运行基础知识	522
8.4	模拟试题	403	10.2.8	系统维护基础知识	523
8.4.1	单项选择题	403	10.2.9	小结	525
8.4.2	设计题	415	10.3	重点习题解析	525
第 9 章 SQL 语言 417			10.3.1	填空题	525
9.1	基本要求	417	10.3.2	选择题	527
9.2	基本内容	417	10.3.3	简答题	532
9.2.1	SQL 简介	417	10.4	模拟试题	533
9.2.2	SQL 的数据定义	419	10.4.1	选择题	533
9.2.3	SQL 的数据查询	425	10.4.2	数据流图设计	541
9.2.4	SQL 数据查询中的 限制和规定	431	第 11 章 数据库设计 547		
9.2.5	SQL 的数据更新	439	11.1	基本要求	547
9.2.6	视图	442	11.2	基本内容	547
9.2.7	嵌入式 SQL	443	11.2.1	数据库应用系统设计的 全过程	547
9.2.8	SQL 的数据控制	450	11.2.2	数据库设计工具介绍	554
9.2.9	主流数据库厂商及产品	462	11.2.3	概念设计与 ER 模型	556
9.2.10	小结	462	11.2.4	逻辑设计与转换规则	577
9.3	对 SELECT 语句的深入理解	463	11.2.5	ER 模型实例分析	581
9.4	重点习题解析	464	11.2.6	面向对象的高级概念 建模	586
9.4.1	填空题	464	11.2.7	小结	592
9.4.2	简答题	466	11.3	重点习题解析	593
9.4.3	设计题	469	11.3.1	填空题	593
9.5	模拟试题	477			
9.5.1	单项选择题	477			

11.3.2 简答题.....	595	第 14 章 数据库发展趋势与新技术	655
11.3.3 设计题.....	596	14.1 基本要求.....	655
11.4 模拟试题.....	598	14.2 基本内容.....	656
11.4.1 选择题.....	598	14.2.1 从关系数据库到对象	
11.4.2 设计题.....	600	数据库.....	656
第 12 章 数据库系统的运行与管理	603	14.2.2 面向对象的数据类型	
12.1 基本要求.....	603	系统.....	656
12.2 基本内容.....	603	14.2.3 OODB.....	658
12.2.1 DBS 的运行计划.....	603	14.2.4 ORDB.....	664
12.2.2 DBS 的运行与维护.....	606	14.2.5 ERP 的发展历程.....	674
12.2.3 数据库的管理.....	609	14.2.6 DW.....	681
12.2.4 性能调整.....	610	14.2.7 数据转移.....	691
12.2.5 用户支持.....	613	14.2.8 OLAP.....	695
12.2.6 小结.....	613	14.2.9 DM.....	703
12.3 重点习题解析.....	614	14.2.10 DSS 的建立.....	710
12.4 模拟试题.....	615	14.2.11 小结.....	714
第 13 章 网络与数据库	617	14.3 重点习题解析.....	716
13.1 基本要求.....	617	14.4 模拟试题.....	721
13.2 基本内容.....	617	第 15 章 知识产权与信息化基础	
13.2.1 分布式数据库的定义		知识.....	726
和特点.....	617	15.1 基本要求.....	726
13.2.2 分布式数据库的体系		15.2 基本内容.....	726
结构.....	620	15.2.1 知识产权的概念与特点.....	726
13.2.3 分布式数据库的查询		15.2.2 计算机软件著作权.....	728
优化.....	628	15.2.3 计算机软件著作权的	
13.2.4 分布式数据库的管理.....	631	侵权问题.....	734
13.2.5 数据库与 WWW.....	636	15.2.4 专利法概述.....	738
13.2.6 动态 Web 网页和开发.....	638	15.2.5 企业知识产权的保护.....	740
13.2.7 XML 与数据库.....	644	15.2.6 信息化基础知识.....	741
13.2.8 小结.....	647	15.2.7 远程教育、电子商务、	
13.3 重点习题解析.....	648	电子政务.....	745
13.3.1 填空题.....	648	15.2.8 小结.....	750
13.3.2 简答题.....	649	15.3 重点习题解析.....	751
13.4 模拟试题.....	652	15.3.1 填空题.....	751

15.3.2 简答题.....	752	16.2.7 标准化组织.....	771
15.4 模拟试题.....	755	16.2.8 ISO 9000 标准简介.....	773
第 16 章 标准化基础知识	761	16.2.9 能力成熟度模型 CMM 简介.....	775
16.1 基本要求.....	761	16.2.10 ISO/IEC 15504 过程评 估标准简介.....	776
16.2 基本内容.....	761	16.2.11 小结.....	778
16.2.1 标准化的基本概念.....	761	16.3 重点习题解析.....	779
16.2.2 标准化过程模式.....	763	16.3.1 填空题.....	779
16.2.3 标准的分类.....	764	16.3.2 简答题.....	781
16.2.4 标准的代号和编号.....	766	16.4 模拟试题.....	783
16.2.5 国际标准和国外先进 标准.....	767	参考文献	786
16.2.6 信息技术标准化.....	769		

第 1 章 计算机系统知识

1.1 基本要求

1. 学习目的与要求

本章的学习目的和要求是：通过对本章的学习，掌握计算机的一些基础知识，包括计算机的组成、基本工作原理、体系结构、存储系统、计算机安全、可靠性与系统性能评测基础知识等。

2. 本章重点内容

- (1) 计算机系统的组成：计算机的发展以及硬件、软件组成。
- (2) 计算机的基本工作原理：数制、汉字编码和 CPU 的结构工作流程。
- (3) 计算机体系结构的知识：体系结构的发展和分类、存储系统、指令系统、输入输出技术、流水线、总线、并行处理。
- (4) 计算机安全：安全的概述、加密和认证技术、计算机病毒。
- (5) 计算机系统的可靠性、性能等的评估。

1.2 基本内容

1.2.1 计算机系统的组成

计算机系统分为硬件和软件两大部分。硬件是计算机系统的机器部分，它是计算机工作的物质基础；软件则是为了运行、管理和维护计算机而编制的各种程序的总和，广义的软件还应该包括与程序有关的文档。本小节中首先介绍了计算机的发展历程，然后分别介绍计算机组成的两部分：硬件和软件。

1. 计算机发展概述

计算机的发明是 20 世纪人类最伟大的成就之一，它标志着信息时代的开始。半个世纪以来，计算机技术一直处于发展和变革之中。至今，计算机的发展经历了以下 5 个重要阶段。

(1) 大型机阶段：1946 年美国宾州大学研制的第一台计算机 ENIAC 被认为是大型机的鼻祖。它采用电子管作为基本逻辑部件，有体积大、耗电多、成本高等缺点。大型机的发展经历了以下几代。

- ① 第 1 代：采用电子管制作的计算机。
- ② 第 2 代：采用晶体管制作的计算机。

③ 第3代：采用中、小规模集成电路制作的计算机。

④ 第4代：采用大规模、超大规模集成电路制作的计算机，如 IBM 360 等。

(2) 小型机阶段：小型机又称小型计算机，通常用以满足部门的需要，被中小型企业使用，如 DEC 公司的 VAX 系列。

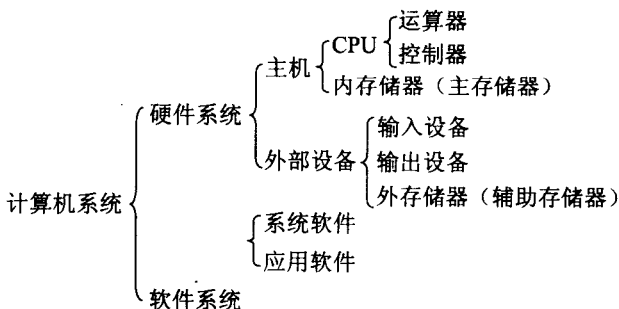
(3) 微型机阶段：微型机又称个人计算机 (PC)，它面向个人和家庭，价格便宜，应用相当普及，如 Apple II、IBM-PC 系列机。

(4) 客户机/服务器阶段：早期的服务器主要是为客户机提供资源共享的磁盘服务器和文件服务器，现在的服务器主要是数据库服务器和应用服务器等。

客户机/服务器 (Client/Server) 模式是对大型机的一次挑战。由于客户机/服务器模式的结构灵活，适用面广，成本较低，因此得到了广泛的应用。如果服务器的处理能力强，客户机的处理能力弱，则称为瘦客户机/胖服务器；否则称为胖客户机/瘦服务器。

(5) 互联网阶段：自 1969 年美国国防部 ARPANET 运行以来，计算机广域网开始发展起来。1983 年 TCP/IP 正式成为 ARPANET 的标准协议，以它为主干发展起来的 Internet 得到飞速发展。

计算机系统由硬件系统和软件系统两部分组成。计算机硬件是计算机系统中看得见、摸得着的物理装置；计算机软件是程序、数据和相关文档的集合。计算机系统的组成如下。



2. 计算机硬件系统结构

计算机的硬件分成 5 大组成部件：运算器、控制器、存储器、输入设备和输出设备。其中，运算器和控制器是计算机的核心，合称中央处理单元 (Central Processing Unit, CPU) 或处理器。CPU 的内部还有一些高速存储单元，被称为寄存器。其中运算器执行所有的算术和逻辑运算；控制器负责把指令逐条从存储器中取出，经译码后向计算机发出各种控制命令；而寄存器为处理单元提供操作所需要的数据。

存储器是计算机的记忆部分，用来存放程序以及程序中涉及的数据。它分为内部存储器和外部存储器。内部存储器用于存放正在执行的程序和使用的数据，其成本高、容量小，但速度快。外部存储器可用于长期保存大量程序和数据，其成本低、容量大，但速度较慢。

输入设备和输出设备统称为外部设备，简称外设或 I/O 设备，用来实现人机交互和机间通信。微型机中常用的输入设备有键盘、鼠标等，输出设备有显示器、打印机等。

计算机硬件的典型结构主要包括单总线结构、双总线结构和采用通道的大型系统

结构。

(1) 单总线结构：图 1-1 就是单总线的计算机系统结构，即用一组系统总线将计算机的各部件连接起来，各部件之间可以通过总线交换信息。这种结构的优点是易于扩充新的 I/O 设备，并且各种 I/O 设备的寄存器和主存储器可以统一编址，使 CPU 访问 I/O 设备更方便灵活；其缺点是同一时刻只能允许挂在总线上的一对设备之间相互传送信息，即只能分时使用总线，这限制了信息传送的吞吐量。这种结构一般用在小型和微型计算机中。

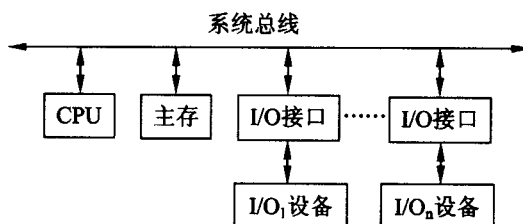


图 1-1 单总线结构

(2) 双总线结构：为了消除信息传送的瓶颈，常设置多组总线，最常见的是在内存和 CPU 之间设置一组专用的高速存储总线。以 CPU 为中心的双总线结构如图 1-2 所示，将连接 CPU 和外围设备的系统总线称为输入输出 (I/O) 总线。这种结构的优点是控制线路简单，对 I/O 总线的传送速率要求很低，缺点是 CPU 的工作效率很低，因为 I/O 设备与主存之间的信息交换要经过 CPU 进行。以存储器为中心的双总线结构如图 1-3 所示，主存储器可通过存储总线与 CPU 交换信息，同时还可以通过系统总线与 I/O 设备交换信息。这种结构的优点是信息传送速率高，缺点是需要增加新的硬件投资。

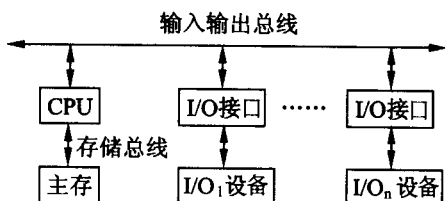


图 1-2 以 CPU 为中心的双总线结构

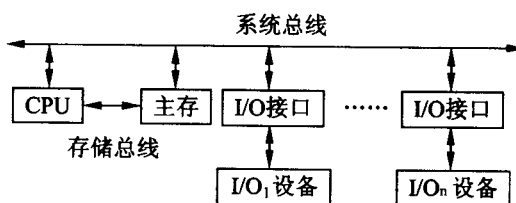


图 1-3 以存储器为中心的双总线结构

(3) 采用通道的大型系统结构：为了扩大系统的功能和提高系统的效率，在大、中型计算机系统中采用通道结构。在这种结构中，一台主机可以连接多个通道，一个通道可以连接一台或多台 I/O 设备，所以它具有较大的扩展余地。另外，由通道来管理和控制 I/O 设备，减轻了 CPU 的负担，提高了整个系统的效率。

3. 计算机软件

软件是计算机系统的重要组成部分，它可以使计算机更好地发挥作用。计算机软件是程序、数据和相关文档的集合。软件可以分为系统软件和应用软件。

系统软件是指为了方便使用、维护和管理计算机系统而编制的软件及其文档，包括操

作系统、语言翻译程序等。

应用软件是解决某一问题的程序及其文档。它覆盖了计算机应用的所有方面，每个应用都有相应的应用软件。

1.2.2 计算机的基本工作原理

1. 数制

生活中习惯用十进制计数，但在计算机中以二进制的形式表达数值，为了便于表示，还使用十六进制数。如果在数字系统中只使用 r 个基本符号表示数值，则称为 r 进制， r 称为该数制的基。如对于二进制，只有符号 0 和 1。对于十六进制，除了 0~9 十个数字外，还有 A~F 六个字母，分别代表 10~15。

不同数制都采用位置表示法，不同位上的字符代表的值与该位的权值有关。一个 n 位的 r 进制数 N 可以表示为

$$N = d_{n-1}d_{n-2} \cdots d_0$$
$$W_{n-1} W_{n-2} \cdots W_0$$

其中 d_i 是从低位数起的第 i 位数字， W_i 是该位的权值， $W_i = r^i$ ，如

$$123 = 1 \times 100 + 2 \times 10 + 3 \times 1 = 1 \times 10^2 + 2 \times 10^1 + 3 \times 10^0$$

数字 N 等值的十进制数可以用如下公式求得：

$$N = d_{n-1} \times W_{n-1} + d_{n-2} \times W_{n-2} + \cdots + d_0 \times W_0 = \sum_{i=0}^{n-1} d_i \times W_i = \sum_{i=0}^{n-1} d_i \times r^i$$

$$\begin{aligned} \text{如十六进制数 } A3B4 &= 10 \times 16^3 + 3 \times 16^2 + 11 \times 16^1 + 4 \times 16^0 \\ &= 41908 \end{aligned}$$

不同的数制的数之间可以互相转换。十进制数转换为二进制和十六进制数可以使用除法，二进制和十六进制数之间存在对应关系，每 4 个二进制位对应 1 个十六进制位，所以二进制和十六进制数之间转换非常方便直接。

例：十进制数 123 转换为二进制、八进制、十六进制表示。

首先将 123 转换为十六进制， $123 \div 16 = 7$ 余 11。

所以 123 的十六进制表示为 7B。

再将 7B 转换为二进制，每个字符转换成 4 个二进制字符 0111 1011。

再将 0111 1011 转换成八进制，每 3 个二进制字符转换成 1 个八进制字符，为 173。

2. 算术逻辑运算

由于计算机中使用二进制表示数值，所以计算机中最基本的是二进制的算术逻辑运算。

(1) 二进制加法：逢二进一

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

(2) 二进制减法：借一当二

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

(3) 二进制乘法

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

(4) 二进制除法：与乘法相反，除 0 为非法

$$1 \div 1 = 1 \quad 0 \div 1 = 0$$

(5) 二进制与运算，又称为逻辑乘

$$0 \wedge 0 = 0 \quad 0 \wedge 1 = 0 \quad 1 \wedge 0 = 0 \quad 1 \wedge 1 = 1$$

(6) 二进制或运算，又称为逻辑加

$$0 \vee 0 = 0 \quad 0 \vee 1 = 1 \quad 1 \vee 0 = 1 \quad 1 \vee 1 = 1$$

(7) 二进制异或运算

$$0 \oplus 0 = 0 \quad 0 \oplus 1 = 1 \quad 1 \oplus 0 = 1 \quad 1 \oplus 1 = 0$$

3. 机器数和码制

在计算机中，各种字符只能以二进制编码方式表示。字母和字符需要按照特定的规律编码在计算机中表示，最常用的一种编码是 ASCII 码（美国标准信息交换码：American Standard Code for Information Interchange）。ASCII 码用 7 位二进制编码，故有 128 个。

对十进制数字编码可以使用所谓的“二进制编码的十进制数（Binary Coded Decimal, BCD）”，在这种编码中 1 位十进制数用 4 位二进制编码表示，最常用的 BCD 码是 8421BCD 码，它用 4 位二进制的低 10 个表示 0~9 这 10 个数字，如表 1-1 所示。

表 1-1 不同进制间与 BCD 码的对应关系

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

例：对于 BCD 码：0100 1001 0111 1000.0001 0100 1001

可以知道它表示的数字为 4978.149

数值在计算机中表示的二进制编码通常称为机器数，它对应的实际数值称为机器数的真值。机器数分为无符号数和有符号数。无符号数只能表示正数，所有位按照二进制编码表示数字。有符号数的最高位为符号位，0 表示正数，1 表示负数。有符号数的表示有原码、反码和补码等不同编码方式，这些编码方式称为码制。

(1) 原码：最高有效位表示符号（正数用 0，负数用 1），其他位表示数值大小。

例: $X = 106 = 01101010B$ $【X】_{原} = 01101010B$

$X = -106$ $【X】_{原} = 11101010B$

(2) 反码: 正数的反码与原码相同, 最高符号位用 0 表示, 其余位为数值位。而负数的反码则为它的正数的各位 (包括符号位) 按位取反而形成的, 即 0 变成 1, 1 变成 0。

例: $X = 106 = 01101010B$ $【X】_{反} = 01101010B$

$X = -106$ $【X】_{反} = 10010101B$

负数的反码与原码有很大区别: 最高符号位仍用 1 表示, 但是数值位不同。

对于数值 0, 在原码和反码中有 +0 和 -0 两种表示法。8 位二进制原码和反码所能表示的数值范围为 $-127 \sim +127$ 。

(3) 补码: 正数的补码表示与原码相同, 即最高符号位用 0 表示, 其余位为数值位。而负数的补码则为它的反码在最低位上加 1 得到。

例: $X = 106 = 01101010B$ $【X】_{补} = 01101010B$

$X = -106$ $【X】_{补} = 10010110B$

补码不分所谓的 +0 或 -0, 只有一个表示形式。在计算机中, 有符号数默认采用补码形式, 所以补码应用最广泛。8 位二进制补码所能表示的数值为 $-128 \sim +127$ 。-128 的补码为 10000000, -1 的补码为 11111111。

补码的一个好处是不同符号数相加不需要通过减法来实现, 而可以直接按照二进制加法法则计算。

例: $37 + (-69) = -32$

$00100101 + 10111011 = 11100000$

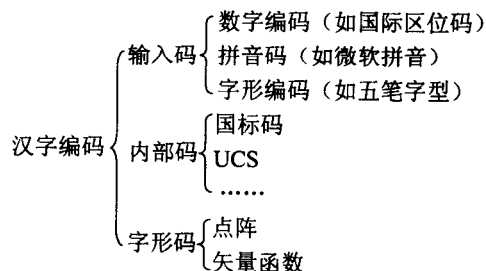
同符号数的补码相加可能产生溢出, 即结果超过了规定的数值范围。溢出会产生计算错误, 使两个正数相加变成负数, 两个负数相加变成正数。

例: $89 + 67 = 156$

$01011001 + 01000011 = 10011100 = -100$

4. 汉字编码

在计算机上处理汉字, 必须先对汉字进行编码。汉字的处理主要包括编码输入、存储和输出 3 部分, 分别对应着输入码、内部码和输出字形码。



(1) 输入码: 输入码解决的主要问题是利用西文标准的键盘将汉字输入到计算机中。汉字的输入码主要分为 3 类: 数字编码、拼音码和字形编码。

① 数字编码是用数字串代表汉字，比较常用的是国标区位码。它将国家标准局公布的 6 763 个汉字分成 94 个区，每区 94 位，区码和位码各为两位的十进制数字，它们一起确定一个汉字。如“中”字在第 54 区 48 位，区位码为 5 448。数字编码输入的优点是无重码，缺点是编码难以记忆，而且输入一个汉字要按键 4 次，不便于输入。

② 拼音码以汉字拼音为基础，敲入拼音后在同音的字中进行选择，如微软拼音、智能 ABC 等。拼音码的优点是便于记忆，缺点是拼音输入重码率太高，选择比较耗时，影响了输入速度。

③ 字形编码以汉字的形状为基础，将汉字拆成偏旁部首和笔划，对这些部件用字母或数字进行编码，通过键入这些笔划组合成汉字，如五笔字型等。字形编码的重码率小，输入速度快，记忆有规律，比较流行。

(2) 内部码：汉字内码是汉字在计算机或其他信息处理设备中存储、传输和处理的形式。国家标准局 GB2312-80 中规定了汉字的国标码，使用两个字节存放一个汉字的内码，每个字节的最高位为 1，这样两个字节各用 7 位，共可以表示 16 384 个汉字。以汉字“大”为例，国标码为 3473H，将两个字节的最高位置 1，得到机内码为 B4F3H。另外国标码等于区位码加上 2020H。

1993 年国际标准化组织公布了“通用多八位编码字符集”的国际标准 ISO/IEC10646，简称 USC，其中包括中、日、韩的文字，每个字符使用 4 个字节来表示。

(3) 输出字形码

输出的汉字字形码表示汉字字形的字模数据，是汉字的输出方式。它通常用点阵、矢量函数等方式来表示。在用点阵表示汉字时，根据汉字输出的要求不同，点阵的大小也不同，简易型汉字为 16×16 点阵，高精度的汉字为 24×24 点阵、 32×32 点阵、 48×48 点阵等。汉字的矢量表示法将汉字看作笔划组成的图形，保存每个笔划矢量的坐标值，输出时将所有笔划组合起来得到字形信息。不同的汉字矢量表示占用的空间不同，而点阵法中每个汉字占用空间都相同。

5. 中央处理器 CPU

(1) CPU 的组成

CPU 主要由运算器和控制器组成，下面分别介绍这两部分。

① 运算器

运算器是对数据进行加工处理的部件，它主要完成算术运算和逻辑运算，完成对数据的加工和处理。运算器基本都是由算术/逻辑运算单元 (ALU)、累加器 ACC、寄存器组、多路转换器和数据总线等逻辑部件组成。

② 控制器

计算机能执行的基本操作称为指令，一台计算机的所有指令组成指令系统。指令由操作码和地址码两部分组成，操作码指明操作的类型，地址码则指明操作数及运算结果存放的地址。

控制器的主要功能是从内存中取出指令，并指出下一条指令在内存中的位置。将取出

的指令经指令寄存器送往指令译码器，经过对指令的分析发出相应的控制和定时信息，控制和协调计算机的各个部件的工作，以完成指令所规定的操作。

控制器主要由程序计数器（PC）、指令寄存器（IR）、状态条件寄存器、时序产生器、微操作信号发生器组成。

- 程序计数器（PC）：当程序顺序执行时，每取出一条指令，PC 内容自动增加一个值，指向下一条要取的指令。当程序出现转移时，则将转移地址送入 PC，然后由 PC 指向新的程序地址。
- 指令寄存器（IR）：用于存放当前要执行的指令。
- 指令译码器（ID）：对现行指令进行分析，确定指令类型、指令所要完成的操作以及寻址方式。
- 状态/条件寄存器：用于保存指令执行完成后产生的条件码，例如运算是否有溢出，结果为正还是为负，是否有进位等。此外，状态/条件寄存器还保存中断和系统工作状态等信息。
- 微操作信号发生器：把指令提供的操作信号、时序产生器提供的时序信号及由控制功能部件反馈的状态信号等综合成特定的操作序列，从而完成取指令的执行控制。

执行指令一般分为取指令、指令译码、按指令操作码执行和形成下一条指令地址 4 个步骤。

（2）CPU 的功能

CPU 的基本功能如下。

- ① 程序控制：CPU 通过执行指令来控制程序的执行顺序。
- ② 操作控制：一条指令功能的实现需要若干操作信号来完成，CPU 产生每条指令的操作信号并将操作信号送往不同的部件，控制相应的部件按指令的功能要求进行操作。
- ③ 时间控制：CPU 对各种操作进行时间上的控制。
- ④ 数据处理：CPU 对数据进行算术运算及逻辑运算等方式进行加工处理。

1.2.3 计算机体系结构

1. 计算机体系结构概述

计算机体系结构（Computer Architecture）是程序员所看到的计算机的属性，即概念性结构与功能特性。按照计算机系统的多级层次结构，不同级程序员所看到的计算机具有不同的属性。一般来说，低级机器的属性对于高层机器程序员基本是透明的，我们通常所说的计算机体系结构主要指机器语言级机器的系统结构。

一般来说，计算机体系结构的属性包含以下方面。

- 机内数据表示：硬件能直接辨识和操作的数据类型和格式；
- 寻址方式：最小可寻址单位、寻址方式的种类、地址运算；
- 寄存器组织：操作寄存器、变址寄存器、控制寄存器及专用寄存器的定义、数量