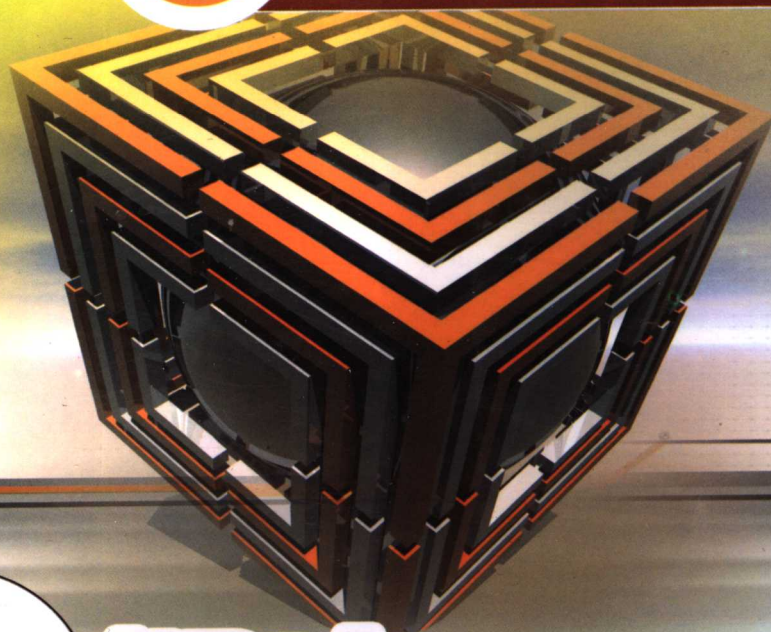




普通高等教育“十一五”规划教材

高等职业教育计算机技术系列教材



数据库

原理与应用

(SQL Server)

曹新谱 李 强 曹 蕾 编著

HIGHER TECHNICAL
AND
VOCATIONAL
EDUCATION

冶金工业出版社



数据阵

原理与应用

作者：[作者姓名]

出版社：[出版社名称]

ISBN: [ISBN号]

定价：[定价]
[其他信息]

普通高等教育“十一五”规划教材
高等职业教育计算机技术系列教材

数据库原理与应用

(SQL Server)

曹新谱 李 强 曹 蕾 编著

北 京

冶金工业出版社

内 容 简 介

本书是根据普通高等教育“十一五”国家级规划教材的指导精神而编写的。

数据库技术是目前 IT 行业中发展最快的领域之一,已经被广泛应用于各种类型的数据处理系统之中,了解并掌握数据库知识已经成为各类科技人员和管理人员的基本要求。

本书基于 SQL Server 2000 系统讨论数据库的原理和应用方法。全书分为三部分:第 1~6 章介绍数据库的一般原理及 SQL Server 2000 的基本应用;第 7 章和第 8 章介绍 SQL Server 2000 数据库的运行与管理;第 9 章~第 11 章介绍 Transaction-SQL 程序设计基础。附录 A、附录 B 和附录 C 分别列举了 SQL Server 函数、数据库模型案例集和 SQL Server 的全局变量。每章都配有练习题,书中还安排了多个上机实验题供选做。

本书内容由浅入深,循序渐进,通俗易懂,适合自学,可作为高职高专院校相关专业及其他培训班的“数据库原理与应用”课程的教学用书,也可作为计算机应用人员和计算机爱好者的自学参考书。

图书在版编目(CIP)数据

数据库原理与应用:SQL Server /曹新谱,李强,曹蕾编著. —北京:冶金工业出版社,2007.8

普通高等教育“十一五”规划教材

ISBN 978-7-5024-4367-2

I. 数… II. ①曹…②李…③曹… III. 关系数据库—数据库管理系统, SQL Server—高等学校—教材
IV. TP311.138

中国版本图书馆 CIP 数据核字(2007)第 109129 号

出版人 曹胜利(北京沙滩嵩祝院北巷 39 号,邮编 100009)

责任编辑 戈兰

ISBN 978-7-5024-4367-2

广州锦昌印务有限公司印刷;冶金工业出版社发行;各地新华书店经销

2007 年 8 月第 1 版第 1 次印刷

787mm × 1092mm 1/16; 14.5 印张; 332 千字; 224 页

28.00 元

冶金工业出版社发行部 电话:(010) 64044283 传真:(010) 64027893

冶金书店 地址:北京东四西大街 46 号(100711) 电话:(010) 65289081

(本社图书如有印装质量问题,本社发行部负责退换)

前 言

一、关于本书

本书是根据普通高等教育“十一五”国家级规划教材的指导精神而编写的。

《数据库原理与应用》作为高职高专学校某些专业的必修课程，或作为很多相关专业的选修课程，是很普遍的。这方面的书和教材也不少。本人和同事们都有多年从事这门课程教学的经历。然而，感到很难找到合适的教材。现在勉强使用的教材，普遍存在几个问题。第一是过分强调理论；第二是顾及所谓系统性，内容面面俱到，没有特色，重点不突出；第三是只讲知识结构的第一层面，很少讲系统的执行情况，出现问题的分析处理方法；第四是可帮助进行实际训练的、有应用价值的案例太少。对高职高专的培养目标而言，以上几个不足恰好是最致命的病根。因此，决心自己动手，编写更适合院校的培养目标，更贴近我们的大众的教材。

二、本书结构

本书共 11 章和 3 个附录，全部内容分成三篇：

第 1 篇 数据库系统基础（第 1~6 章）。

第 1 章：数据库系统概述。本章给读者一些关于数据库的大致轮廓和基本概念的描述，基本名词、术语的解释，如数据库系统的组成，三级模式的技术构架，数据模型的种类，数据完整性概念和关系数据模型等知识。

第 2 章：SQL Server 2000 基础。介绍 SQL Server 2000 的安装与简单的使用方法，让读者安装一个实际的、可运行的数据库服务器，并学会一些简单的操作。本章为数据库的实现作准备。

第 3 章：SQL 语言。介绍 SQL 语言的特点和 SQL Server 2000 中适用的数据类型、内部函数等，是为数据库的实现作铺垫的。也是对第 1 章中数据模型三要素的部分内容的具体展示。

第 4 章：基表。讨论数据库实现的重点内容——创建基本表及对基本表载入数据。

第 5 章：数据查询，介绍 SQL 中各类查询语句的结构，使用规则等。有大量的、可运行的查询语句实例。透过这些实例的讲解和执行情况的分析，使读者能逐步了解 SQL 语言的查询功能全貌。

第 6 章：用户视图。介绍视图的创建和使用，这是三级模式结构中外模式在应用中的具体体现。读者掌握了这些内容，对三级模式结构的科学性与合理性，对使用外模式带来的优点，自然会有切实的体会和深刻的认识。

第 2 篇 数据库运行与管理（第 7~8 章）。

第 7 章：数据库安全管理。主要内容包括数据库的安全控制、用户账户管理、用户权限管理、角色管理等。让读者对数据库的安全管理有一定程度的认识。

第 8 章：数据库的日常管理与数据传输。同第 7 章一样，了解这些知识，小型数据库

用户自己可以兼顾系统管理员的工作，保证数据库系统的正常运行。

第3篇 SQL Server 程序设计（第9~11章）。

第9章：SQL Server 程序设计。介绍 SQL Server 程序设计的基本知识，如批处理、脚本、事务等基本概念，全局变量与局部变量、自定义数据类型、规则、自定义函数等。还有编程的多种语句，如程序流程控制语句及使用规则。

第10章：存储过程与游标。通过存储过程与游标的定义与使用，使读者掌握更深入的编程方法与技巧，提高代码的质量和解决实际问题能力。

第11章：触发器。通过触发器的创建和应用，进一步提高编程的能力和水平。

附录A、附录B和附录C分别列举了SQL Server 函数、数据库模型案例集和SQL Server 的全局变量。

三、本书特点

本教材适合60~80学时左右的教学计划。课堂讲授大体占三分之二，读者实践应不少于三分之一。如果读者要达到熟练操作，还要增加实际训练的时间。

各章之后的练习题，读者一定要动手操作、完成，分析执行结果。不能满足于理性的“懂”。满足于懂，不一定真懂。即使真懂，不动手操作，就不能变成“动手能力”。

四、本书适用对象

本书内容由浅入深，循序渐进，通俗易懂，适合自学，可作为高职高专院校相关专业及其他培训班的“数据库原理与应用”课程的教学用书，对于计算机应用人员和计算机爱好者本书也是一本实用的自学参考书。

本书由于通俗易懂，深入浅出，给读者分析、解剖技术细节，非常适合自学。

由于水平、能力、经验所限，难免有诸多不妥，请各位同仁、读者批评指正。联系方式如下：

电子邮箱：service@cnbook.net

网址：www.cnbook.net

本书电子教案和习题参考答案可在该网站下载，此外，该网站还有一些其他相关书籍的介绍，可以方便读者选购参考。

编 者

2007年6月

目 录

第1章 数据库系统概述.....1

1.1 什么是数据库系统.....1

1.1.1 数据库系统的组成.....1

1.1.2 数据库系统与人.....2

1.2 数据和数据模型.....2

1.2.1 数据、数据标准和它的动态特性 ..2

1.2.2 数据模型.....4

1.2.3 概念层数据模型.....5

1.2.4 组织层数据模型.....6

1.3 实体-联系模型.....8

1.3.1 E-R 模型中常用的名词与实体 联系图.....8

1.3.2 一个实际的 E-R 模型案例.....11

1.4 关系数据模型.....13

1.4.1 关系模型的数据结构.....14

1.4.2 关系模型的数据操作.....17

1.4.3 关系模型的数据完整性约束.....17

1.5 从 E-R 模型到关系模型的转换.....20

1.5.1 从 E-R 模型转换为关系模型.....20

1.5.2 按 E-R 图转换来的关系模型的 实例.....21

1.6 数据库系统的内部结构.....22

1.6.1 三级模式.....22

1.6.2 两级映像与数据独立性.....23

1.6.3 数据字典和数据.....24

1.6.4 数据库管理系统功能及其处理 流程简述.....25

1.7 现代数据库系统的优点.....26

1.7.1 数据库系统的发展简史.....26

1.7.2 数据库系统的优点.....27

小结.....29

综合练习一.....29

一、选择题.....29

二、填空题.....30

三、问答题.....31

第2章 SQL Server 2000 基础.....32

2.1 常见的关系型数据库简介.....32

2.1.1 SQL Server 2000 概述.....32

2.1.2 其他的数据库系统.....33

2.2 SQL Server 2000 安装.....33

2.2.1 安装前的准备.....34

2.2.2 安装 SQL Server 2000 个人版.....34

2.3 系统测试和服务管理器的启动.....39

2.4 SQL Server 2000 中的两个常用工具 ..40

2.4.1 企业管理器.....40

2.4.2 查询分析器.....42

小结.....45

综合练习二.....45

一、选择题.....45

二、填空题.....46

三、问答题.....46

四、上机操作题.....46

第3章 SQL 语言.....49

3.1 SQL 语言概述.....49

3.1.1 SQL 语言的发展.....49

3.1.2 SQL 语言特点.....49

3.2 SQL 语言的功能概述.....50

3.3 SQL 允许使用的数据类型.....51

3.3.1 数值型.....51

3.3.2 近似型.....52

3.3.3 字符串型.....52

3.3.4 日期时间类型.....53

3.3.5 货币类型.....54

3.4 部分常用函数.....54

3.4.1 聚合函数.....54

3.4.2 数学函数.....55

3.4.3 字符串函数.....55

3.4.4 类型转换函数.....55

3.4.5 日期函数.....55

3.4.6 系统函数.....56

小结.....56

综合练习三.....56

一、选择题	56	5.2.6 消除取值完全相同的记录	88
二、填空题	57	5.3 条件查询	88
三、问答题	57	5.3.1 以比较运算做条件	89
四、上机操作题	57	5.3.2 用确定范围为条件	90
第4章 基表	58	5.3.3 以确定集合当查询条件	91
4.1 概述	58	5.3.4 以字符匹配为条件	92
4.2 基表的定义	58	5.3.5 涉及空值的查询	94
4.2.1 定义基本表的 SQL 语句格式和语法	58	5.3.6 多重条件(或复合)查询	94
4.2.2 关于数据的完整性约束	59	5.4 对查询结果排序和使用计算函数	95
4.2.3 实例	61	5.4.1 对查询结果排序	95
4.3 案例:创建“高校学生学籍与成绩管理数据库”	63	5.4.2 在查询语句中使用计算函数	96
4.4 修改表的结构和表名	68	5.4.3 对查询结果进行分组统计(或计算)	98
4.4.1 修改表的结构	68	5.5 多表连接查询	99
4.4.2 修改表约束	69	5.5.1 内连接	100
4.4.3 修改表名	70	5.5.2 自连接	103
4.5 删除基本表	71	5.5.3 外连接	104
4.6 图形界面的数据录入	72	5.6 子查询	105
4.7 使用 SQL 语句进行数据插入	74	5.6.1 使用子查询的结果作比较	105
4.8 数据修改与删除	75	5.6.2 使用子查询的结果作集合	106
4.8.1 数据修改	75	5.6.3 使用子查询进行逻辑测试	108
4.8.2 数据删除	76	5.7 基于外表条件的数据修改和删除	110
小结	77	5.7.1 基于外表条件的数据修改	110
综合练习四	78	5.7.2 基于外表条件的数据删除	110
一、选择题	78	小结	110
二、填空题	79	综合练习五	111
三、问答题	80	一、选择题	111
四、上机操作题	81	二、填空题	111
第5章 数据查询	83	三、问答题	112
5.1 查询语句的基本结构	83	四、上机操作题	112
5.1.1 查询语句的语句结构和语法	83	第6章 用户视图	114
5.1.2 查询约定	84	6.1 用户视图概述	114
5.2 简单查询	84	6.1.1 什么是用户视图	114
5.2.1 查询表中指定的列	84	6.1.2 视图与基本表的区别	114
5.2.2 查询表中的全部列	85	6.2 创建用户视图	115
5.2.3 结果列的值是经过计算的值	86	6.2.1 定义单源表视图	116
5.2.4 在查询结果中插入常量值的列	86	6.2.2 定义多源表视图	116
5.2.5 改变查询结果集中的列标题	87	6.2.3 在已有视图上定义新视图	117
		6.2.4 定义带表达式的视图	117
		6.2.5 含分组统计信息的视图	117

6.3 用户视图查询与删除	118	一、选择题	142
6.3.1 用户视图查询	118	二、填空题	143
6.3.2 删除用户视图	119	三、问答题	143
6.4 用户视图的作用	119	四、上机操作题	144
小结	120	第8章 数据库的日常管理与数据传输	145
综合练习六	121	8.1 备份数据库	145
一、选择题	121	8.1.1 数据库故障的种类	145
二、填空题	121	8.1.2 备份数据库	145
三、问答题	122	8.1.3 备份的方法	146
四、上机操作题	122	8.2 恢复数据库	151
第7章 数据库安全管理	123	8.2.1 自动恢复	151
7.1 SQL Server 的服务器	123	8.2.2 手工恢复	151
7.1.1 注册服务器	123	8.2.3 用企业管理器实现恢复	152
7.1.2 删除服务器	124	8.2.4 使用 Transact-SQL 语句恢复 数据库	154
7.2 安全控制	124	8.3 数据传输	155
7.2.1 一般安全控制模式	125	8.3.1 DTS 功能概述	155
7.2.2 数据库权限种类及用户的分类 ..	126	8.3.2 利用 DTS 向导实现数据的 导入和导出	156
7.2.3 SQL Server 的三级身份验证	126	小结	159
7.2.4 混合身份验证模式	127	综合练习八	160
7.3 用户登录帐户管理	129	一、选择题	160
7.3.1 建立用户登录帐户	130	二、填空题	161
7.3.2 修改登录帐户的属性	131	三、问答题	161
7.3.3 删除登录帐户	132	四、上机操作题	161
7.4 管理数据库用户	132	第9章 SQL Server 程序设计	163
7.4.1 建立数据库用户	133	9.1 批处理、脚本与事务	163
7.4.2 删除数据库用户	134	9.1.1 批处理	163
7.5 用户权限管理	134	9.1.2 脚本与注释	164
7.5.1 SQL Server 的权限管理	134	9.1.3 事务的特征	165
7.5.2 使用企业管理器管理用户权限 ..	135	9.1.4 事务处理类型	165
7.5.3 使用 Transact-SQL 语句 管理权限	137	9.2 变量与临时表	166
7.6 角色管理	138	9.2.1 全局变量	166
7.6.1 建立用户自定义角色	139	9.2.2 局部变量	167
7.6.2 为用户自定义角色授权或 回收权限	140	9.2.3 表变量	168
7.6.3 添加或删除角色中的成员	140	9.2.4 临时表	169
7.6.4 用户自定义角色的权限 修改及其他	142	9.3 程序控制语句	170
小结	142	9.3.1 BEGIN END 结构	170
综合练习七	142	9.3.2 IF ELSE 语句	170

9.3.3 WHILE 循环和 BREAK、 CONTINUE 语句	171	综合练习十	194
9.3.4 CASE 语句	172	一、选择题	194
9.3.5 GOTO 语句	173	二、填空题	195
9.4 自定义数据类型、默认值与规则	173	三、问答题	195
9.4.1 自定义数据类型	173	四、上机操作题	195
9.4.2 默认值	174	第 11 章 触发器	196
9.4.3 规则	176	11.1 触发器的创建	196
9.5 自定义函数	177	11.1.1 触发器的优点	196
9.5.1 标量函数	177	11.1.2 触发器的创建	196
9.5.2 内嵌表值函数	179	11.2 触发器启动	198
9.5.3 自定义函数的删除	181	11.3 触发器实例	198
小结	181	11.3.1 delete 触发器	198
综合练习九	181	11.3.2 insert 触发器	199
一、选择题	181	11.4 触发器综合应用	199
二、填空题	182	11.5 禁用/启用/删除触发器	202
三、问答题	182	11.5.1 禁止使用	202
四、上机操作题	183	11.5.2 启用	202
第 10 章 存储过程与游标	184	11.5.3 删除触发器	203
10.1 存储过程的基本概念	184	小结	203
10.1.1 什么是存储过程	184	综合练习十一	203
10.1.2 存储过程的种类	184	一、选择题	203
10.1.3 存储过程的特点	184	二、填空题	203
10.2 创建存储过程	185	三、问答题	203
10.2.1 用企业管理器创建存储过程	185	四、上机操作题	204
10.2.2 用 SQL 语句创建存储过程	186	附录 A SQL Server 函数	205
10.3 存储过程的执行与删除	186	A.1 聚合函数	205
10.3.1 执行存储过程	187	A.2 日期和时间函数	207
10.3.2 删除存储过程	188	A.3 数学函数	211
10.4 游标的声明	188	A.4 字符串函数	212
10.4.1 游标的功能	189	附录 B 数据库模型案例集	215
10.4.2 声明游标	189	B.1 银行储蓄业务系统	215
10.5 使用游标	191	B.2 超市业务管理系统	216
10.5.1 打开游标	191	B.3 航空售票系统	218
10.5.2 从游标中提取数据	191	B.4 产品加工信息管理系统	220
10.5.3 关闭游标	193	附录 C SQL Server 2000 的全局变量	222
10.5.4 释放游标	193	参考文献	224
小结	194		

第 1 章 数据库系统概述

随着社会的进步和科学技术的高速发展，信息已经成为社会的重要财富和资源。为适应人们对信息管理的需要，数据库技术得到了极大的发展，其应用领域也越来越广泛。数据库技术已经普遍应用于工业、农业、商业企业、政府机构管理等各领域，几乎无所不在，无所不包。有没有自动化的信息管理系统（其核心是数据库系统），已经成为衡量一个企业管理水平和信息化程度的重要标志。但是，真正要正确、全面地理解数据库系统的概念，必须深入了解数据库系统的组成和系统内部的技术构架，设计和实现数据库系统的数据模型，系统的运行机制等技术细节。本章将对数据库系统及其相关概念作一个简明地介绍。

1.1 什么是数据库系统

什么是数据库系统？一种较为通俗的说法是：数据库系统是储存数据文件的仓库。这种表述，实际上是将“数据库系统”的概念简化成了“数据库”，把它看成是静态的容器，是不全面的、较肤浅的；另一种说法是：数据库系统是储存有大量数据文件的计算机系统，它可以为人们提供多方面的数据服务功能。这一表述虽然不一定十分准确，但至少它描述了数据库系统的静态特性与动态特性，并强调了它是一个可以运行的计算机系统，能为用户提供许多服务功能，将数据库系统的本质表述较为全面。

1.1.1 数据库系统的组成

一个数据库系统（DBS）应由五部分组成，如图 1-1 所示描述了它们之间的关系。

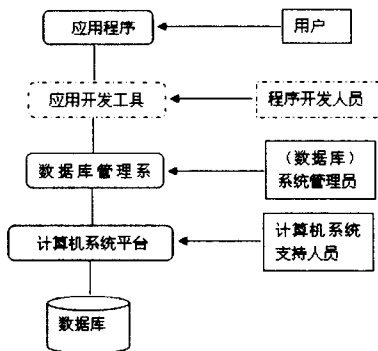


图 1-1 数据库系统构成及相关人员

（1）计算机系统平台。即计算机硬件系统及其相适应的操作系统（OS）等软件组成的计算机平台。

（2）数据库管理系统（DBMS）。它是一个大型的专用的软件包，或称数据库管理（软件）系统。这里，强调它是一个大型软件，是它对机器平台，特别对机器的操作系统，有较多的要求。也就是说，系统平台的性能、操作系统的类型及版本与数据库管理软件相适合，相匹配。

（3）数据库（DB）。简单通俗地说，存放于机器的海量存储器中的全部数据与信息组

成数据库。一般又分为数据（文件）和数据字典，将在后续内容中再作详细解释。

（4）应用程序。它是开发人员为解决用户的问题而设计、开发的程序。当然，开发用户应用程序的开发工具和环境，一般就是本系统的组成部分。

（5）应用开发工具。它是数据库实现和用户开发程序所必需的环境及软件工具。它无疑应与数据库管理系统（软件）相匹配。这一部分只是在数据库开发阶段起明显作用。当一个数据库系统建立起来，并正式运行后，应用开发工具一般不再是数据库系统必需的部分，其应用开发人员也可以撤离，而不会影响数据库系统的日常运行。因此，在图 1-1 中使用了虚线框，以示区别。

1.1.2 数据库系统与人

一个数据库系统，如果抛开其设计过程，实施过程，试运行阶段不讲，假定前几个阶段的工作已经完成，通过了验收，交付用户使用，即进入日常运行阶段。那么，数据库设计人员，应用程序开发人员，调试、测试及验收人员都已经退出。参与日常运行的只有用户、数据库系统管理员。他们各自有自己的职责。

（1）用户。它是数据库的最终使用者。一般情况下可将其分为一般用户、特殊（高级）用户。他们对各自的业务处理流程很熟悉。一般地说，用户使用数据库，都是为了实现信息管理和信息处理两大功能。而且，这两项功能几乎是经常交互式进行的。当用户遇到各种问题或困难时，需要数据库管理员的帮助或支持。

（2）数据库管理员（DBA）。有时也称为系统管理员。他是一个重要而关键的角色。他具有多方面的知识（如对数据库管理系统、操作系统的功能、工作原理与机制，对数据库的文件结构等，有深入的了解）和较丰富的数据库管理经验，能最大限度地帮助用户解决使用中可能出现的错误、故障或其他问题。他除了负责数据库用户管理外，他还负责数据库的定期或不定期的备份、恢复、日常维护、性能测试与评估等工作，甚至对数据库进行重组等。

1.2 数据和数据模型

一个大型数据库的设计、开发过程，是一个系统工程或一项软件工程。数据库用户和系统管理员，不一定要参加设计或开发工作。但他们作为数据库的使用者，有必要对数据库系统中的许多基本概念、技术有更多的了解。

1.2.1 数据、数据标准和它的动态特性

1. 数据（Data）

为了有效地描述各种事物的特点、属性，人们通常是使用“抽象”的方法从现实世界中将有用的信息抽取出来，进行集中管理或处理。例如，一个企业招聘了一批新员工，都要求新来的员工填写《员工情况登记表》。表中可能包含：员工编号、姓名、性别、出生年月、学历、专业技术特长等。

数据库系统运行过程中，主要是使用和处理数据。什么是数据？数据是人们用于描述（记录、刻画、表述）事物的特征或属性的可数值化的信息。当今社会环境下，数据库系统使用的数据，大体包括如下几大类：

数字、数值——这是数据早期的含义，是较狭义的解释。

字母、文字、符号——伴随科学技术的进步，计算机应用中的数据已经超出了数值的范围，需要保存和处理字母、文字信息，是对数据的第一次概念扩充。

图形、图像——从简单图形到地图、机械图，到人像、影像等。

语音、声音、音乐、视频——这些数据能进入计算机系统，是对数据的概念的进一步扩充。

到目前为止，这四大种类的数据，都是适合于计算机处理的可数值化信息。随着社会的发展和科学技术的进步，今后还可能出现其他新型的数据种类。

2. 数据标准和实际规范

上面只是列出了现在的数据库系统中出现的4大数据种类。对于手工管理方式，粗一点好像问题不大；对于实际使用的数据库系统而言，只知道这四大类是远远不够的。例如，在一个企业的信息管理系统中，通常有“员工基本情况表”（如表1-1所示）。此表中包含了很多员工的重要信息：表名、表中的多个栏目及其名称，是一类信息。每一个栏目，是数字，还是文字，或声音、图像，都要具体说明。如果是数字，还要说明是整数、还是实数，它的精度、取值范围，有何限制条件等等。就是说要有数据标准或数据规范。用技术名词来表述，就要求每个数据项，不仅有便于识别它的名称，而且要有具体的数据类型和取值范围（值域）等，才能建立起实用的数据库系统。这一点，无论是数据库用户，还是数据库设计、开发技术人员，一开始就必须有明确而清醒的认识。

表 1-1 员工基本情况

职工编号	姓名	性别	出生年月	最高学历	技术职称	部门号	职务

实际情况是，相当多的数据库用户，对于数据的规范化和标准化往往不重视，认为无关紧要。这一点，从许多粗糙的“用户需求说明书”中不难看出。而从软件工程或技术性角度，则要求数据必须规范，有统一的数据标准。因此有人建议：凡是有国家标准的，首选国家标准；其次，如果没有国家颁布的数据标准，也可以参考国标上通行的行业标准；即使没有以上两类标准，也应建立企业的数据标准或数据规范。这样，对数据库系统的建立和今后的运行管理是十分必要的。

关于数据标准和数据规范的问题，可参考软件工程学和其他相关资料，这里不再多述。

3. 数据的静态与动态特性

数据库系统中的数据，普遍具有两大特性：即数据的静态特性和数据的动态特性。人们容易看到其静态特性而忽略其动态特性。

1) 数据的静态特性

数据的静态特性包括三方面：数据的基本结构、数据类型及取值范围（约束条件），数据之间的联系与制约关系（约束）。

例如，一个学校的学生学籍及成绩管理数据库，一般会在系统中设计三张表，如下所示：

（1）学生基本信息表：一般应包括学号、学生姓名、性别、出生年月、所在系、专

业、班级等等。

(2) 课程表。一般应包括课程号、课程名、学分等等。

(3) 学生选课信息表。一般应包括学号、课程号、考试成绩等等。

以上给出的只是学生基本信息和学生选课信息的基本结构, 这还不够。

例如, 学号, 怎么表示? 是数字, 还是字母? 还是字母数字混合表示? 即使明确了是数字, 还要规定是多少位等等。又如性别, 是用“男”、“女”表示, 还是用英文 male、female 表示, 还是中英文都可以? 考试成绩, 用 0~100 之间的正整数, 还是可以出现带小数点的 85.6 分表示? 这些就是对数据类型及取值范围的要求。

再比如, 在学生选课信息表中, “学号”一项, 只允许出现学生信息表中实际存在的学号。只有这样, 表中的数据才有实际意义。人们只允许数据库中存储正确的、有价值的、有实际意义的数据, 而应避免一切错误和不符合规范的数据(垃圾数据)进入系统。这样一来, 就能最大限度地保证数据库的正确性与较高的运行效率。这就是数据与数据之间的联系与制约关系。

建立起这些基本概念, 搞清数据的这些静态特性, 是学好、用好数据库的基础。

2) 数据的动态特性

数据的动态特性是指数据的可操作性。任何数据库中的数据, 无疑都是给人们操作使用的。如果一个数据库中的数据, 失去了可操作性, 只是固定不变的一大堆数据, 那就完全失去了数据库存在的意义。数据库用户对数据库的操作有查询、更新数据等操作, 更新数据又可分为插入、删除、修改, 简称为增、删、改。

一般来说, 人们较容易注意到数据的静态特性, 而容易忽略其动态特性。数据的动态特性是与用户对数据库的使用密切相关的, 在数据库的概念设计阶段, 表现得并不明显, 但是, 并非不必考虑这一特性。

有时, 人们将数据的静态特性和数据的动态特性概括为数据模型的三要素, 即数据的基本结构、数据的约束条件和定义在数据集合上的数据操作。

1.2.2 数据模型

现实生活中, 人们经常使用各类模型。如: 一所大学的建筑模型, 一个用于军事的战争态势沙盘模型、飞机模型等。人体解剖图、地图都是模型。借助这些模型, 有利于人们对现实世界中某一事物的结构、组织形态、内部特征、整体与局部的关系, 及它的运动与变化等多元信息的把握和了解。而数据模型是对现实世界中各类数据的抽象和模拟。

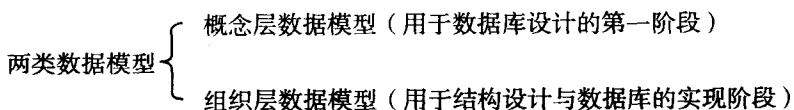
任何数据库系统的建立, 都要依赖某种数据模型, 来描述和表示信息系统。因此, 数据模型一般应满足三个要求:

(1) 需要尽可能真实地模拟或反映现实世界的数值(信息)特征。

(2) 便于人们理解 and 交流。

(3) 便于在计算机系统上实现存储和处理。

由于数据库的方案设计和数据库的实现各有特点, 因此, 不同阶段使用不同数据模型。在数据库的概念结构设计阶段, 使用概念层数据模型(也称信息模型)。在数据库的结构设计和实施阶段, 使用组织层数据模型(或简称组织模型)。



为了将现实世界的具体事物抽象、组织成一个为某一 DBMS 支持的数据模型，人们通常分两步完成：

（1）将现实世界的信息（属性、特征等）抽象为信息世界的各类数据，并使用概念模型来描述各数据的名称、数据类型、数据的精度与取值范围，不同数据之间的关联等；

（2）将概念模型转换为组织层模型，将数值化信息转换成机器数据。

这一转换过程如图 1-2 所示。

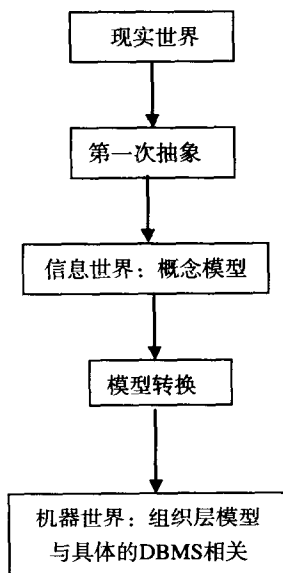


图 1-2 信息数据的转换过程

1.2.3 概念层数据模型

信息或数据从现实世界到信息世界的转换，即概念模型的设计，应该说是最难，也是最有挑战性的。因为它是一种创造性的劳动，需要设计者对课题内容深入、透彻的理解，丰富的实践经验，而且具有很高的综合素质和良好的抽象、综合能力。这种综合素质和抽象能力，当然主要不是教科书上学来的，设计者必须进行深入的用户调查等等。

相比之下，从信息世界到机器世界的转换，是比较直接和容易的。只要具有相关的专业技术知识，遵循基本规则和一定的技术规范，就不难实现。数据库应用人员的着眼点和重点应当放在数据库的实现和应用上。

概念层数据模型是用于描述、表达现实世界的多元化信息和数据的工具。同时也是数据库设计人员进行数据库设计的工具，它还是数据库设计人员与用户进行交流的工具。因此，它应具有较强的语义表达能力，也就是说，它能有效地、精确地、完整地表达人们从现实世界抽象出来的各类数据与信息特征、特性、属性，同时还包含各类数据与信息之间的关联、联系与约束。概念层数据模型应尽可能简单明了、清晰，便于人们理解与相互

交流。

概念模型是面向用户的,其主要目标是如何将现实世界的的数据表达清楚,方便用户理解和使用。在数据库概念设计阶段,设计人员应把主要精力放在了解现实世界,将用户需要的各类数据及数据之间的相互联系上,尽量准确地表述清楚。而将涉及 DBMS 的许多技术性问题暂时放下。也就是说,概念模型的设计与 DBMS 无关。

常用的概念模型有实体-联系模型 (Entity-Relationship model, E-R 模型)、语义对象模型等。在第 1.5 节将介绍使用较多的实体-联系模型。

1.2.4 组织层数据模型

所谓组织层数据模型是以数据在计算机中的组织方式为出发点,如何能更有效、更方便,更有利于设计人员与用户进行沟通、理解,且更方便用户使用。数据库技术经过近 50 年的研究与发展,最常用的组织层数据模型有 4 种:

1. 层次数据模型

它是用树形结构表示各类实体及实体之间的联系。例如,在高校的管理体系中,一个学校有许多系或其他部门,一个系又有多个教研室和多个学生队,一个教研室有许多教师的信息,学生队有许多学生的信息等等,这是一种较典型的层次结构。可以用图 1-3 来表示这些信息的组成和联系。

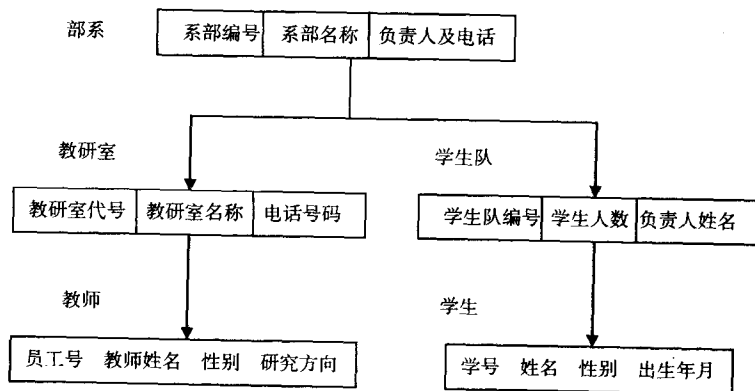


图 1-3 高校管理的数据模型

IBM 公司的 IMS 数据库管理系统 (1968 年) 是该公司推出的第一个大型商用数据库管理系统,曾经得到广泛应用,流行于 20 世纪 70 年代。

由于现实世界中很多联系可能很复杂,而并非是简单的层次关系,用层次模型来表示这些联系不仅不方便,而且增加了数据操作的难度和复杂性。经过几十年的历史变迁,使用层次模型的数据库系统在当今世界已不多见。

2. 网络数据模型

如前面所述,由于现实世界中很多联系并非层次关系而是呈更为复杂的网状关系,网状数据模型也是一种曾经较为流行的数据模型。

在如图 1-4 所示的模型中,信息之间的联系更为复杂。例如,从信息 A 出发,到信息 C2,有两条路径可走。一条是 A-A-02-C2,另一条路径是 A-B-D-C2。网络数据模型的典型代表是 DBTG 系统 (也称 CODASYL 系统),它是 20 世纪 70 年代推出的。确切地说,

它不是一个实际的软件系统，而是一个以网络数据模型为基础的数据库系统方案。它提出的基本概念、基本方法和技术处理途径，具有较普遍的意义，对网络数据库系统的研究和开发有很大的影响。但由于网络数据模型的结构复杂，数据操作语言和软件设计技术难度较大，用户不容易使用，最终也不可能成为数据模型的主流方向。

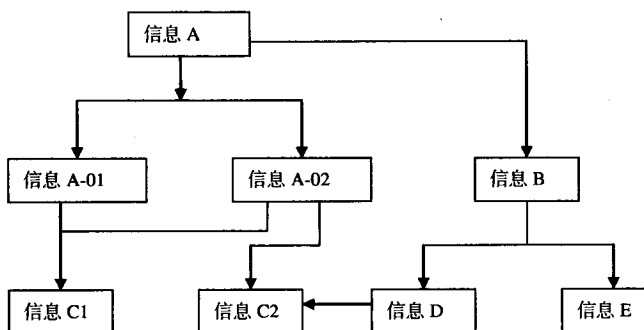


图 1-4 网状的数据模型

3. 关系数据模型

关系模型是数据库发展历史上产生较早，研究最多，使用时间最长，实际软件系统最丰富的一种数据模型。世界上许多有影响的研究、开发数据库系统的厂商所推出的数据库系统产品（软件），大多是使用关系数据模型来组织数据的。因此，关系模型是数据模型中最具影响力的，同时也是最重要的数据模型。

关系数据模型是 1970 年由美国 IBM 公司的研究员 E.F.Codd 提出的，它为关系数据理论和数据库关系方法奠定了坚实的基础。正是由于 E.F.Codd 的杰出工作和贡献，他于 1981 年获得了 ACM 图灵奖。

关系数据库系统为何经历几十年的发展，成为了数据库系统组织层数据模型的主流技术？这是由于它具有其他数据模型所没有的优势。

（1）在关系模型中，实体及实体之间的联系都用关系来表示，概念单一，数据结构简单、清晰，用户容易理解。

（2）建立在关系模型上的关系数据库系统，它的数据操作主要是关系运算（操作）。关系操作（运算）的特点是，操作对象是关系（二维表），操作的结果也是关系（二维表），特殊情形下，或许是空关系，即一张空表。

（3）关系操作在执行时，数据的存取路径对用户是透明的。用户只要知道他“干什么”或“查找什么数据”，而不必考虑“怎么干”或“如何找”，从而使用户感到工作轻松。同时，也极大地方便了用户应用程序的开发和数据库结构的修改，提高了数据库的运行效率。

（4）关系模型和关系运算是建立在关系代数理论上的，有严格的数学理论基础，一般不会出现意想不到的错误，系统具有较高的数据独立性、数据完整性和数据安全性。

关系模型是本书的重点。关于关系模型的详细内容将在本章 1.4 节中具体论述。

4. 面向对象数据模型

随着数据库应用的迅速发展，由于传统的数据模型（层次模型、网络状型、关系模型等）的数据对象简单、无复杂数据或嵌套数据；实体之间的联系简单，不能表示聚合、继