



全国高等职业教育规划教材

数据库技术与应用

—— SQL Server 2008

胡国胜 易著梁 主编

李曼 张红娟 张辉 等编著

电子教案下载网址 www.cmpedu.com



机械工业出版社
CHINA MACHINE PRESS

全国高等职业教育规划教材

数据库技术与应用

——SQL Server 2008

胡国胜 易著梁 主编
李曼 张红娟 张辉 等编著



机械工业出版社

本书系统全面地阐述了数据库的基本原理及应用。全书内容围绕图书馆管理数据库系统展开,系统地介绍了 SQL Server 2008 数据库的应用;书中实训环节使用宾馆管理信息系统,充分体现了“项目驱动、案例教学、理论与实践相结合”的教学理念。通过最后的综合应用实例的学习,读者可以进一步掌握实际而又全面的数据库应用知识。本书为教师配有电子教案。

本书以培养学生的数据库开发能力为目标,内容新颖,通俗易懂,实用性强,适合作为高职高专院校计算机相关专业的教材,也可供广大技术人员及自学者参考。

图书在版编目(CIP)数据

数据库技术与应用——SQL Server 2008 / 胡国胜, 易著梁主编. —北京: 机械工业出版社, 2010.1
(全国高等职业教育规划教材)
ISBN 978-7-111-29463-4

I. 数… II. ①胡…②易… III. 关系数据库—数据库管理系统, SQL Server 2008—高等学校: 技术学校—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字(2010)第 001647 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 鹿 征

责任印制: 李 妍

北京振兴源印务有限公司印刷

2010 年 1 月第 1 版·第 1 次印刷

184mm×260mm·16.5 印张·407 千字

0001—4000 册

标准书号: ISBN 978-7-111-29463-4

定价: 28.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010) 88361066

门户网:<http://www.cmpbook.com>

销售一部:(010) 68326294

教材网:<http://www.cmpedu.com>

销售二部:(010) 88379649

读者服务部:(010) 68993821

封面无防伪标均为盗版

前 言

在计算机科学技术中，数据库是发展最快的技术之一，近年来，各种应用领域对数据管理的需求越来越多，数据库技术的重要性也越来越被大家所认识。尤其是 Internet 的发展以及多种信息技术的交叉与发展，给数据库应用提供了更多的机遇，同时也推动了数据库技术的发展和完善。

Microsoft 公司的新一代服务器产品 SQL Server 2008 新增了许多新的特性和关键的改进，是至今为止最强大和最全面的 SQL Server 版本。本书全面介绍了 SQL Server 2008 的主要功能、相关命令和开发应用系统的一般技术，力求最有效地帮助读者快速而全面地掌握数据库技术的基本原理和应用。

本书共有 15 章，可分为 3 个部分。第 1 部分是第 1、2 章，介绍数据库技术和关系数据库的基本概念；第 2 部分是第 3~14 章，介绍 SQL Server 2008 的基本数据管理功能，主要包括数据库设计、SQL Server 2008 系统概述、创建与使用数据库、创建与使用数据表、SQL 查询、T-SQL 编程基础、视图与索引、存储过程、触发器、事务和锁、数据库的安全保护，以及数据库的备份与还原；第 3 部分是第 15 章，介绍以 SQL Server 2008 为数据库平台的基于 C/S 模式的数据库应用系统设计开发实例。附录部分介绍以 ERwin 为工具实现数据库模型的设计。每章后面都配实训和习题。

本书作者精心设计了两个具体的数据库管理系统实例，在教学环节中使用图书馆管理系统，在实训过程中使用宾馆管理信息系统，充分体现了“项目驱动、案例教学、理论与实践相结合”的教学理念。通过这两个实例，读者可以对数据库技术有一个比较全面的了解，进而掌握数据库理论与数据库应用的基本知识，了解数据库应用系统的开发模式，并具有初步的数据库应用开发能力，实现编程技能的逐步提高。本书是具有丰富教学经验的教师与经验丰富的企业级应用程序开发工程师相结合的成果，一线教师与企业人员的优势互补保障了教材的质量，是校企结合的范例。

本书由胡国胜、易著梁主编，刘瑞新主审，李曼、张红娟、张辉等编著，参加编写的作者还有陈克坚、张建华、贾燕玲、李晓娟、魏蔚、张玉胜、郭银、雷成茂、翟春雷、彭春艳、刘克纯、岳爱英、岳香菊、崔瑛瑛。由于编者水平有限，书中错误与疏漏之处在所难免，敬请读者批评指正。

为方便教师授课，本书提供了电子教案、源代码等，读者可到机械工业出版社教材服务网 www.cmpedu.com 下载。

编 者

目 录

出版说明

前言

第 1 章 数据库技术基础	1	2.6 习题	23
1.1 关于数据库的基本术语	1	第 3 章 数据库设计	25
1.1.1 信息与数据	1	3.1 数据库设计概述	25
1.1.2 数据库与数据库系统	1	3.1.1 数据库设计的内容	25
1.1.3 数据库管理系统	2	3.1.2 数据库设计的步骤	26
1.2 数据管理技术的发展	3	3.2 图书馆管理系统数据库的 设计	30
1.2.1 人工管理阶段	3	3.2.1 需求分析	30
1.2.2 文件系统阶段	3	3.2.2 数据库结构设计	30
1.2.3 数据库系统阶段	4	3.3 实训	33
1.3 数据模型	5	3.4 习题	34
1.3.1 数据模型的概念及要素	6	第 4 章 SQL Server 2008 系统 概述	35
1.3.2 概念模型及表示	7	4.1 SQL Server 2008 简介	35
1.3.3 常用的数据模型	9	4.1.1 SQL Server 的发展历史	35
1.3.4 数据库系统的三级数据模式 结构	12	4.1.2 SQL Server 2008 的版本	35
1.4 实训	13	4.2 SQL Server 2008 的安装	36
1.5 习题	14	4.2.1 环境需求	36
第 2 章 关系数据库	15	4.2.2 SQL Server 2008 的安装 过程	37
2.1 关系	15	4.3 SQL Server 2008 的组件	47
2.2 关系运算	16	4.4 SQL Server 2008 管理工具	48
2.2.1 选择	16	4.4.1 SQL Server Management Studio	49
2.2.2 投影	16	4.4.2 SQL Server Business Intelligence Development Studio	49
2.2.3 连接	16	4.4.3 SQL Server Configuration Manager	50
2.3 关系的完整性及约束	18	4.4.4 SQL Server Profiler	50
2.3.1 关系的完整性	18	4.4.5 数据库引擎优化顾问	51
2.3.2 约束	19	4.5 实训	51
2.4 关系设计的规范化	20	4.6 习题	52
2.4.1 第一范式 (1NF)	20		
2.4.2 第二范式 (2NF)	21		
2.4.3 第三范式 (3NF)	21		
2.4.4 BC 范式 (BCNF)	22		
2.5 实训	23		

第 5 章 创建与使用数据库	53		
5.1 SQL Server 2008 中的数据库			
基础知识	53		
5.1.1 数据库常用对象	53		
5.1.2 文件和文件组	54		
5.1.3 系统数据库	55		
5.2 数据库的创建	55		
5.2.1 使用 SQL Server Management			
Studio 创建用户数据库	55		
5.2.2 使用 CREATE DATABASE 语句			
创建用户数据库	58		
5.3 查看数据库信息	62		
5.3.1 使用 SQL Server Management			
Studio 查看数据库信息	63		
5.3.2 使用 T-SQL 语句查看数据库			
信息	64		
5.4 修改数据库	64		
5.4.1 更改数据库的所有者	64		
5.4.2 添加和删除数据文件、日志			
文件	65		
5.4.3 重命名数据库	69		
5.5 删除数据库	70		
5.5.1 使用 SQL Server Management			
Studio 删除用户数据库	70		
5.5.2 使用 DROP DATABASE 语句			
删除用户数据库	71		
5.6 实训	71		
5.7 习题	71		
第 6 章 创建与使用数据表	73		
6.1 数据类型	73		
6.1.1 基本数据类型	73		
6.1.2 用户定义数据类型	75		
6.2 创建数据表	76		
6.2.1 数据表的概念及内容	76		
6.2.2 使用 SQL Server Management			
Studio 工具创建数据表	77		
6.2.3 使用 Transact-SQL 创建			
数据表	79		
6.3 修改表结构	81		
		6.3.1 使用 SQL Server Management	
		Studio 修改数据表结构	81
		6.3.2 使用 Transact-SQL 修改数据表	
		结构	82
		6.4 数据表中插入、修改与删除	
		数据	84
		6.4.1 向数据表中插入数据	84
		6.4.2 修改数据表中数据	87
		6.4.3 删除数据表中数据	87
		6.5 删除数据表	89
		6.6 实训	89
		6.7 习题	90
		第 7 章 SQL 查询	92
		7.1 查询的基本结构	92
		7.2 简单查询	93
		7.2.1 简单查询语句	93
		7.2.2 对查询结果排序	94
		7.2.3 将查询结果分组统计	95
		7.2.4 使用聚合函数进行查询	96
		7.3 条件查询	97
		7.3.1 比较查询条件	97
		7.3.2 范围查询条件	98
		7.3.3 列表查询条件	99
		7.3.4 模糊 LIKE 查询	100
		7.4 连接查询	102
		7.4.1 内连接	102
		7.4.2 外连接	103
		7.4.3 交叉连接	104
		7.5 子查询	104
		7.5.1 使用比较运算符的子	
		查询	105
		7.5.2 使用 IN 的子查询	105
		7.5.3 使用 SOME 和 ANY 的子	
		查询	106
		7.5.4 使用 ALL 的子查询	107
		7.5.5 使用 EXISTS 的子查询	107
		7.6 实训	108
		7.7 习题	109
		第 8 章 T-SQL 编程基础	111

8.1	Transact-SQL 基础知识	111	9.6.1	什么是索引	141
8.1.1	SQL 与 Transact-SQL	111	9.6.2	索引类型	141
8.1.2	T-SQL 语法格式	111	9.7	创建索引	143
8.2	批处理	112	9.7.1	使用 SQL Server Management Studio 工具创建索引	143
8.3	常量和变量	112	9.7.2	使用 CREATE INDEX 语句创建索引	144
8.3.1	常量	113	9.8	管理索引	146
8.3.2	全局变量	113	9.8.1	修改索引	146
8.3.3	局部变量	114	9.8.2	删除索引	147
8.4	运算符与表达式	115	9.8.3	查看索引	147
8.5	Transact-SQL 函数	116	9.9	实训	148
8.6	流程控制语句	119	9.10	习题	148
8.6.1	BEGIN...END 语句	119	第 10 章	存储过程	151
8.6.2	IF...ELSE 语句	119	10.1	存储过程概述	151
8.6.3	CASE...END 语句	120	10.1.1	存储过程的概念	151
8.6.4	WHILE 语句	121	10.1.2	存储过程的优点	151
8.6.5	WAITFOR 语句	122	10.2	创建和执行存储过程	152
8.6.6	注释	123	10.2.1	创建存储过程	152
8.7	游标的使用	123	10.2.2	执行存储过程	156
8.7.1	游标概述	123	10.3	管理存储过程	159
8.7.2	游标的基本操作	124	10.3.1	查看存储过程	159
8.8	实训	130	10.3.2	修改存储过程	161
8.9	习题	130	10.3.3	重命名存储过程	162
第 9 章	视图与索引	132	10.3.4	删除存储过程	163
9.1	视图的基础知识	132	10.4	实训	165
9.1.1	视图的概念	132	10.5	习题	165
9.1.2	视图的作用	132	第 11 章	触发器	167
9.1.3	视图的类型	133	11.1	触发器的概念	167
9.2	创建视图	133	11.1.1	触发器的作用	167
9.2.1	使用 SQL Server Management Studio 工具创建视图	133	11.1.2	触发器的分类	167
9.2.2	使用 Transact-SQL 语句创建视图	135	11.1.3	比较触发器与约束	169
9.3	修改、删除和重命名视图	136	11.2	创建触发器	169
9.3.1	修改视图	136	11.3	修改、删除、重命名和查看触发器	171
9.3.2	删除视图	137	11.3.1	修改触发器	171
9.3.3	重命名视图	138	11.3.2	删除触发器	171
9.4	使用视图操作数据表	138	11.3.3	重命名触发器	171
9.5	查看视图信息	139	11.3.4	查看触发器信息	172
9.6	索引概述	141			

11.4	触发器的使用	172	14.3	还原	213
11.4.1	INSERTED 表和 DELETED 表	172	14.3.1	还原概述	213
11.4.2	触发器示例	173	14.3.2	还原数据库	213
11.5	实训	176	14.4	分离和附加数据库	215
11.6	习题	177	14.4.1	分离数据库	216
第 12 章	事务和锁	178	14.4.2	附加数据库	217
12.1	事务	178	14.5	数据的导入与导出	219
12.1.1	事务的特性及管理	178	14.5.1	数据的导出	219
12.1.2	事务控制语句	179	14.5.2	数据的导入	223
12.2	锁	182	14.6	实训	225
12.2.1	锁的模式	183	14.7	习题	226
12.2.2	死锁	184	第 15 章	SQL Server 2008 综合应用	
12.3	实训	184		实例——Windows 应用程序	
12.4	习题	185		开发	227
第 13 章	数据库的安全保护	187	15.1	Windows 应用概述	227
13.1	数据库安全性控制	187	15.1.1	C/S 模式的工作原理	227
13.1.1	SQL Server 的身份验证 模式	187	15.1.2	.NET 框架	228
13.1.2	设置身份验证模式	188	15.1.3	Visual Studio 编程环境	228
13.2	用户和角色管理	189	15.2	基于 Windows 环境的图书馆 管理系统的设计与实现	229
13.2.1	用户管理	189	15.2.1	系统设计	229
13.2.2	角色管理	197	15.2.2	系统实现	229
13.3	权限管理	201	15.3	实训	241
13.3.1	登录账号权限管理	201	15.4	习题	242
13.3.2	用户账号权限管理	202	附录 A	使用 ERwin 设计数据库	
13.4	实训	204		模型	243
13.5	习题	204	A.1	ERwin 简介	243
第 14 章	数据库的备份与还原	205	A.2	使用 ERwin 设计数据库 模型	243
14.1	故障的种类	205	A.2.1	启动 ERwin 并创建逻辑模型 文件	244
14.2	备份	206	A.2.2	实现图书馆管理系统的数据库 模型	247
14.2.1	备份类型	206	A.3	实训 使用 ERwin 进行数据 建模	254
14.2.2	备份设备的类型	207	A.4	习题	254
14.2.3	创建备份设备	207			
14.2.4	删除备份设备	209			
14.2.5	备份数据库	210			

第 1 章 数据库技术基础

当今社会是一个信息化的社会，数据库技术的发展，已经成为先进信息技术的重要组成部分。数据是信息的载体，数据库是互相关联的数据的集合。数据库技术的发展，归根结底是由实际应用需求推动的。目前，绝大多数的计算机应用系统都离不开数据库的支撑。大到一个国家，小到一个集团的内部，数据库的建设规模、数据库信息量的大小和使用频度已经成为衡量其信息化程度的重要标志。

数据库领域有其自身显著的特点，涉及相当多的理论及概念。本章将逐步引出这些概念，使读者从知晓概念到加深对概念的理解。

1.1 关于数据库的基本术语

1.1.1 信息与数据

数据库是计算机信息管理的基础，它的研究对象是数据，一提到数据，人们往往就会想到信息。但是数据并不是信息的本身。

一般认为，信息是对现实世界中不同事物的存在特征、运动形态以及不同事物间的相互联系等多种属性的描述，通过抽象形成概念。信息是可以被认识、理解、表达、加工、推理和传播的诸如数字、文字、图像和声音等符号所表示的某一事物的消息和知识。

信息的表达必须借助于符号，数据是对事实和概念的描述，是表达信息的符号记录。在现实生活中，数据无处不在，文字、图像、声音、员工的档案记录等。例如，在员工的档案中，对于员工的基本信息，人们最感兴趣的是员工的工号、姓名、性别、出生日期、入职时间、所属部门，可以这样描述：

(1001, 张洪, 男, 1983-1-5, 2008-3-1, 销售部)

上面的这条员工记录就是数据。数据本身并不能完全表达内容，一定要通过语义解释。了解语义的人会从上面的记录中得到：张洪是该公司销售部门的一名男职员，工号是 1001，1983 年 1 月 5 日出生，2008 年 3 月 1 日入职。

可见，数据与信息是两个既有联系又有区别的概念，数据是信息的符号表示，信息则是数据的内涵，是对数据的语义解释。但是，在计算机领域，两者并不严格区分，一般统称“数据”。

1.1.2 数据库与数据库系统

数据库 (Database, DB)，顾名思义，就是存放数据的仓库，是一个长期存储在计算机内，相互联系的数据集合，数据库中的数据按照一定的数据模型组织、描述和存储，具有较少冗余和较高的数据独立性，允许多个用户共享使用，并且提供数据的安全性维护和完整性检查措施。

这里需要注意的是，数据库与数据仓库的概念不同；数据库是具有逻辑关系的数据集合，逻辑上无关的数据集合不能称做数据库；数据库是对现实世界的描述，可以是一个单位或组织，其内部的某些改变应及时反映到数据库中。

数据库系统（DataBase System, DBS）是指在计算机系统中引入数据库后的系统，一般由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员（DataBase Administrator, DBA）和用户构成。数据库系统可以用图 1-1 来表示。

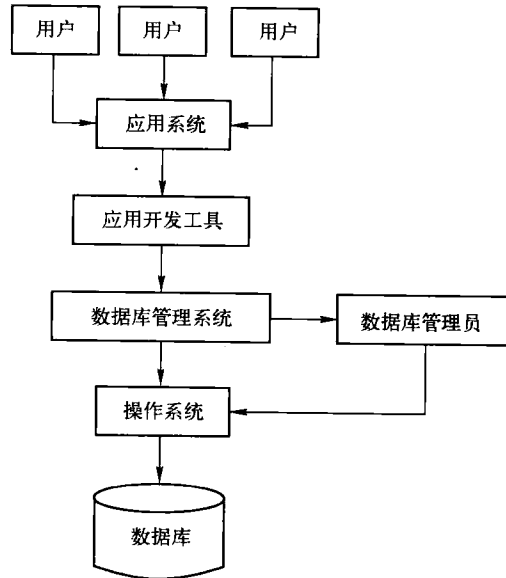


图 1-1 数据库系统

1.1.3 数据库管理系统

数据库管理系统（DataBase Management System, DBMS）是位于用户与操作系统之间的一层数据管理软件，它是为了建立、使用和维护数据库而配置的系统软件。如 Access, Microsoft SQL Server, Oracle 等。它建立在操作系统的基础上，对数据库进行统一的管理和控制。主要功能如下。

1. 数据定义功能

DBMS 提供数据定义语言（Data Definition Language, DDL）对数据库中的数据对象进行定义，如对表、视图、索引、存储过程等进行的定义。

2. 数据操纵功能

DBMS 提供数据操纵语言（Data Manipulation Language, DML），用户可以使用 DML 操纵数据，实现对数据库的基本操作，如查询、插入、删除和修改等。

3. 数据库的运行管理功能

数据库在建立、运用和维护时由 DBMS 统一管理、统一控制，以保证数据的安全性、完整性和多用户对数据库使用的并发控制及发生故障后的系统恢复等。数据库的运行管理是 DBMS 的核心部分。

4. 数据库的建立和维护功能

数据库的建立和维护功能包括初始数据的输入、转换，数据库的转储、恢复，数据库的重组和性能检测分析等功能。

1.2 数据管理技术的发展

数据是一个单位或组织的重要资源，为了组织的长远发展，必须对组织的各种数据施行有效的管理。所谓数据管理，是指对数据进行收集、整理、存储、检索、加工和传递等一系列活动的总和。数据管理的最终目的是从数据中获取有用的信息，以服务于组织的管理工作。数据处理是数据管理的中心工作，将原始数据转换成信息的过程称作数据处理。

数据管理技术是随着计算机硬件技术和软件技术的发展而不断发展的，它经历了人工管理阶段、文件系统阶段和数据库系统三个阶段。

1.2.1 人工管理阶段

20 世纪 50 年代中期以前，计算机数据管理的能力很差，这一阶段称为人工管理阶段。此时计算机的硬件存储设备主要有磁带、卡片、纸带等；没有操作系统和专门管理数据的软件；数据处理方式是批处理；计算机主要用于科学计算。

在人工管理阶段，数据与程序之间是一一对应的关系，如图 1-2 所示。

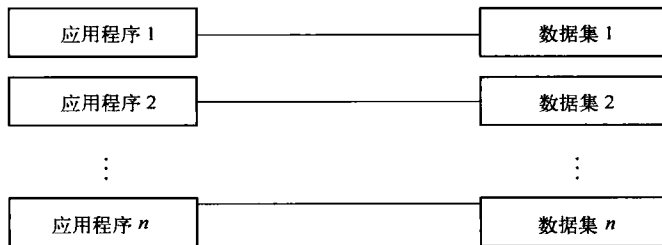


图 1-2 人工管理阶段应用程序与数据之间的对应关系

人工管理阶段的缺点如下。

1. 数据不保存

只是在计算某一具体题目时将数据输入，运行结束后得到输出结果，输入、输出和中间结果均不保存。

2. 数据不共享，冗余度大

一组数据只对应一个应用程序，即使多个应用程序使用相同的数据，也要各自定义，不能共享，导致冗余度大。

3. 数据缺乏独立性

数据与程序是紧密结合在一起的，数据的逻辑结构、物理结构和存取方式都有程序规定。没有文件的概念，数据的组织方式完全由程序员决定。

1.2.2 文件系统阶段

20 世纪 50 年代后期到 60 年代中期，计算机数据管理技术进入到文件系统阶段。此时

的计算机已经有了磁盘、磁鼓等直接存储的设备；出现了操作系统和专门的数据管理软件，称为文件系统；处理方式上不仅有批处理，还能够实现联机实时处理；计算机不仅用于科学计算，还广泛用于数据处理。

在文件系统阶段，文件系统把数据组织成文件形式存储在磁盘上，这些数据文件相互独立，长期保存在存储设备上。文件可以命名，应用程序利用“按文件名访问，按记录进行存取”的方式，对文件中的数据进行修改、插入和删除操作。

这一阶段的数据还是面向应用程序的，数据文件基本上与各自的应用程序相对应，如图 1-3 所示。

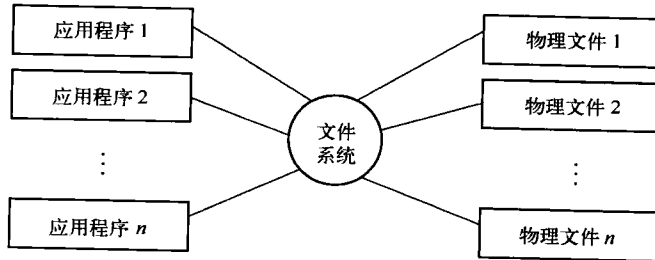


图 1-3 文件系统阶段应用程序与数据之间的对应关系

文件系统阶段对数据的管理有了长足的进步，但它还是从应用程序的角度来组织、看待和处理数据的。这一阶段的缺点如下。

1. 数据共享性差，数据的冗余度较大

文件系统提供了数据的物理独立性，实现了一定程度的数据共享，但只能实现文件级共享。文件的设计很难满足多种应用程序的不同要求，数据冗余不可避免。

2. 数据存在不一致性

在文件系统中，没有维护数据一致性的监控机制，数据的一致性由用户自己维护。同一数据在多个地方同时存放，尤其在大型信息系统中，很容易造成不一致现象的发生。

3. 数据的独立性差

文件系统只实现了数据的物理独立，而没有实现数据的逻辑独立。文件结构的每一处修改都将导致应用程序的修改。因此，文件系统的程序与数据之间缺乏逻辑独立性。

1.2.3 数据库系统阶段

20 世纪 60 年代中后期开始，计算机数据管理技术进入到数据库系统阶段。此时的计算机在硬件方面有了大容量的磁盘，软件方面出现了大量的系统软件；硬件的价格在下降，软件的价格在上升；在处理方式上，联机实时处理要求增多，并开始考虑和提出分布式处理。

为了解决多用户共享数据，使数据为尽可能多的应用服务，一种新的数据管理技术——数据库技术应运而生。数据库技术从 20 世纪 60 年代中期开始萌芽，至 60 年代末 70 年代初，数据库技术已经发展到成熟阶段。

与文件系统不同的是，数据库系统是面向数据的不是面向程序的，各个处理功能通过数据管理软件从数据库中获取所需要的数据和存储处理结果。它克服了文件系统的弱点，为用户提供了一种方便、功能强大的数据管理手段。数据库系统阶段的数据处理过

程如图 1-4 所示。

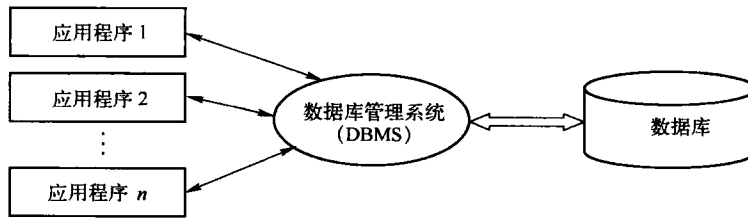


图 1-4 数据库系统阶段应用程序与数据之间的对应关系

数据库系统是在文件系统的基础上发展起来的新技术，与文件系统相比具有如下主要特点。

1. 数据库系统以数据模型为基础

数据库设计的基础是数据模型。在进行数据库设计时，要站在全局需要的角度抽象和组织数据；要完整地、准确地描述数据自身和数据之间联系的情况；要建立适合整体需要的数据模型。

2. 数据库系统的数据冗余度小，数据共享性高

由于数据库系统是从整体角度上看待和描述数据的，数据不再是面向某个应用，而是面向整个系统，所以数据库中同样的数据不会多次重复出现。这就使得数据库中的数据冗余度小，从而避免了由于数据冗余度大带来的数据冲突问题，也避免了由此产生的数据维护和数据统计错误等问题。

数据库系统通过数据模型和数据控制机制提高数据的共享性。数据共享度高会提高数据的利用率，它使得数据更有价值和更容易、方便地被使用。

3. 数据库系统的数据和程序之间具有较高的独立性

由于数据库中的数据定义功能和数据管理功能是由 DBMS 提供的，所以数据对应用程序的依赖度大大降低，数据和程序之间具有较高的独立性。

4. 数据库系统通过 DBMS 进行数据安全性和完整性的控制

数据的安全性控制是指保护数据库，以防止不合法的使用造成的数据泄露、破坏和更改。

数据的完整性控制是指为了保证数据的正确性、有效性和相容性，防止不符合语义的数据输入或输出所采用的控制机制。

5. 数据库中数据的最小存取单位是数据项

在文件系统中，由于数据的最小存取单元是记录，结果给使用及数据操作带来许多不便。数据库系统改善了其不足之处，它的最小数据存取单位是数据项，即使用时可以按数据项或数据项组存取数据，也可以按记录或记录组存取数据。

1.3 数据模型

数据模型是数据库系统的核心，要为一个数据库建立数据模型，首先要深入到信息的现实世界中进行系统需求分析，用概念模型真实、全面地描述现实世界中的管理对象及联系，

然后再将概念模型转换成数据模型。

1.3.1 数据模型的概念及要素

模型是对现实世界特征的模拟和抽象，而数据模型是对现实世界数据特征的抽象，是一组描述数据、数据之间的联系、数据的语义和完整性约束的概念工具的集合。

现实世界的物质要在计算机中得以表示和处理，一般要经过两个阶段的抽象，从现实世界到信息世界的抽象，再从信息世界到计算机世界的抽象。下面先介绍这三个世界（领域）。

1. 现实世界

现实世界泛指存在于人脑之外的客观世界。信息的现实世界是指人们要管理的客观存在的各种事物、事物之间的相互联系及事物的发生、变化过程。通过对现实世界的了解和认识，使得人们对要管理的对象、管理的过程和方法有个概念模型。认识信息的现实世界并用概念模型加以描述的过程称为系统分析。

2. 信息世界

现实世界中的事物反映到人们的头脑里，经过认识、选择、命名、分类等综合分析而形成了印象和概念，从而得到了信息。当事物用信息来描述时，即进入了信息世界。信息世界最主要的特征是可以反映数据之间的联系。

3. 计算机世界

信息世界中的信息，经过数字化处理形成计算机能够处理的数据，就进入了计算机世界。计算机世界也叫做机器世界或者数据世界。计算机世界是数据在计算机上的存储和处理，这些数据必须具有自己特定的数据结构，能够反映信息世界中数据之间的联系。

现实世界、信息世界和计算机世界这三个领域是由客观到认识、由认识到使用管理的三个不同层次，后一领域是前一领域的抽象描述。三者的转换关系如图 1-5 所示。

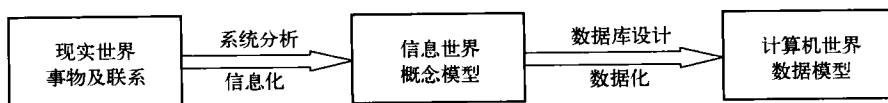


图 1-5 三种世界的联系和转换过程

从图中可以看出，现实世界的事物及联系，通过系统分析成为信息世界的概念模型，而概念模型经过数据化处理转换为数据模型。

数据模型的三要素是数据结构、数据操作和数据的约束条件。

1. 数据结构

数据结构描述的是数据库中数据的组成、特性及其相互间联系。在数据库系统中通常按数据结构的类型命名数据模型，如层次模型、网状模型、关系模型。数据结构是对系统静态特性的描述，是数据模型三要素中的首要内容。

2. 数据操作

数据操作是对数据库中各种对象的实例允许执行的操作的集合，包括操作及操作规则。数据库的操作主要有检索、插入、删除、修改，操作规则有优先级别等。数据操作是对系统动态特性的描述。

3. 数据的约束条件

数据的约束条件是一组完整性规则的集合，用于限定符合数据模型的数据库状态及变化，保证数据的完整性。

1.3.2 概念模型及表示

概念模型是对信息世界的管理对象、属性及联系等信息的描述形式。概念模型不依赖计算机及数据库管理系统，它是对现实世界的真实、全面反映。

1. 信息世界的基本概念

(1) 实体

现实世界中可以相互区分的能被人们识别的事物和概念称为实体 (Entity)。实体可以是实实在在的物体，也可以是抽象的概念或联系。例如，一个学生、一台机器、一部汽车等是事物实体，一门课程、一个班级等称为概念实体。

(2) 实体集

具有相同特征或能用同样特征描述的实体的集合称为实体集 (Entity Set)，如学生、汽车等都是实体集。实体集不是孤立存在的，实体集之间有着各种各样的联系，例如，学生和课程之间有“选课”联系。

(3) 属性

属性 (Attribute) 是实体的某一方面特征的抽象表示。例如，学生可以通过其“学号”、“姓名”、“性别”、“年龄”、“政治面貌”等特征来描述。此时，“学号”、“姓名”、“性别”、“年龄”、“政治面貌”等就是学生的属性。

(4) 属性值

属性值是属性的具体取值。例如，某一学生的学号为“09001”，姓名为“王刚”，性别为“男”，年龄为“20”，政治面貌为“党员”，这些具体的描述就称为属性值。

(5) 域

属性的取值范围称为属性的域 (Domain)。例如，学生的年龄为 16~45 之间的正整数，其数据域为 (16~45)。

(6) 码

码 (Key) 能唯一标识实体的属性或属性集，也可称为关键字。例如，学生的学号可以作为学生实体的码，学生的姓名则不一定可以作为学生实体的码，因为姓名可能重复。学生的选课情况实体集则要把学号和课程号的组合作为码。

2. 概念模型的表示方法

概念模型是对信息世界的建模，概念模型应该能够全面、准确地描述出信息世界中的基本概念。概念模型的表示方法很多，其中最为著名和使用最为广泛的是 P.P.Chen 于 1976 年提出的实体-联系方法 (Entity-Relationship Approach)，简称 E-R 图法。该方法用 E-R 图来描述现实世界的概念模型，提供了表示实体集、属性和联系的方法。E-R 图也称为 E-R 模型。在 E-R 图中：

1) 用长方形表示实体集，长方形内写实体名。

2) 用椭圆形表示实体集的属性，并用线段将其与相应的实体集连接起来。例如，学生具有学号、姓名、性别、年龄和所在系 5 个属性，用 E-R 图表示如图 1-6 所示。

由于实体集的属性比较多，有些实体可具有多达上百个属性，所以在 E-R 图中，实体集

的属性可不直接画出。

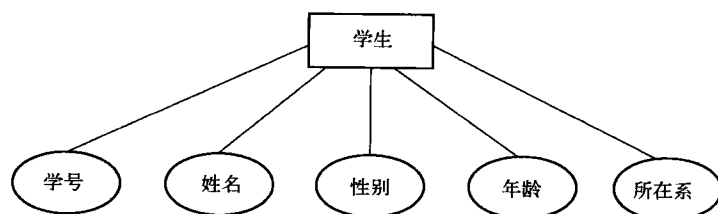


图 1-6 学生及属性的 E-R 图

3) 用菱形表示实体集间的联系, 菱形内写上联系名, 多用动词描述。并用线段分别与有关实体集连接起来, 同时在线段旁标出联系的类型。如果联系具有属性, 则该属性仍用椭圆框表示, 仍需要用线段将属性与其联系连接起来。例如, 供应商、项目和零件之间存在有供应联系, 该联系有供应量属性, 如图 1-7 所示。

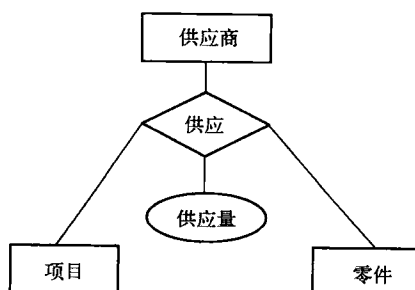


图 1-7 实体间联系的属性及其表示

3. 实体联系的类型

实体集之间的联系可概括为如下 3 种。

1) 一对一联系 (1:1)。设有两个实体集 A 和 B, 如果实体集 A 与实体集 B 之间具有一对一联系, 则对于实体集 A 中的每一个实体, 在实体集 B 中至多有一个 (也可以没有) 实体与之联系; 反之, 对于实体集 B 中的每一个实体, 实体集 A 也至多有一个实体与之联系。两实体集间的一对一联系记做 1:1。例如, 在一个工厂里面只有一个厂长, 而一个厂长只能在一个工厂里任职, 则工厂与厂长之间具有一对一联系。

2) 一对多联系 (1:n)。设有两个实体集 A 和 B, 如果实体集 A 与实体集 B 之间具有一对多联系, 则对于实体集 A 的每一个实体, 实体集 B 中有一个或多个实体与之联系; 而对于实体集 B 的每一个实体, 实体集 A 中至多有一个实体与之联系。实体集 A 与实体集 B 之间的一对多联系记做 1:n。例如, 一个学校里有多名教师, 而每个教师只能在一个学校里教学, 则学校与教师之间具有一对多联系。

3) 多对多联系 (m:n)。设有两个实体集 A 和 B, 如果实体集 A 与实体集 B 之间具有多对多联系, 则对于实体集 A 的每一个实体, 实体集 B 中有一个或多个实体与之联系; 反之, 对于实体集 B 中的每一个实体, 实体集 A 中也有一个或多个实体与之联系。实体集 A 与实体集 B 之间的多对多联系记做 m:n。例如, 工厂里的一个职工可以参加多种体育组织, 而一个体育组织也可以有多名职工, 体育组织与职工之间具有多对多联系。

实际上，一对一联系是一对多联系的特例，而一对多联系又是多对多联系的特例。图 1-8 是用 E-R 图表示两个实体集之间的 1:1、1:n 或 m:n 联系的例子。

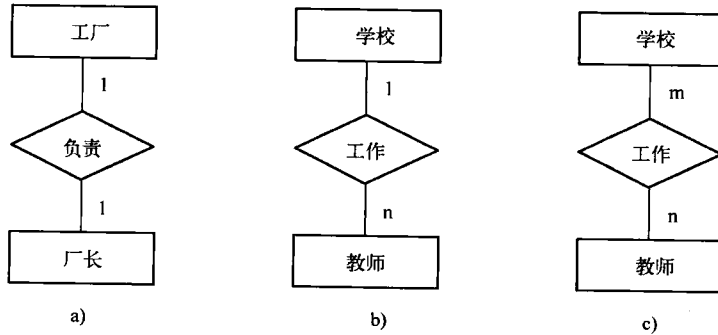


图 1-8 两个实体集联系例子
a) 1:1 联系 b) 1:n 联系 c) m:n 联系

以上三种实体间的联系都是发生在两个实体集之间的。实际上，三个或三个以上实体集之间也可以同时发生联系。例如，图 1-9a 所示的教师、课程、参考书之间的联系，一门课程可以有若干教师讲授，一个教师只讲授一门课程；一门课程使用若干本参考书，每一本参考书只供一门课程使用。因此，课程与教师、参考书之间的联系是一对多的。又如，供应商、项目和零件之间的联系，一个供应商可以供给多个项目多种零件；每个项目可以使用多个供应商供应的零件；每种零件可由不同供应商供给。因此，供应商、项目、零件三个实体集之间是多对多的联系，如图 1-9b 所示。

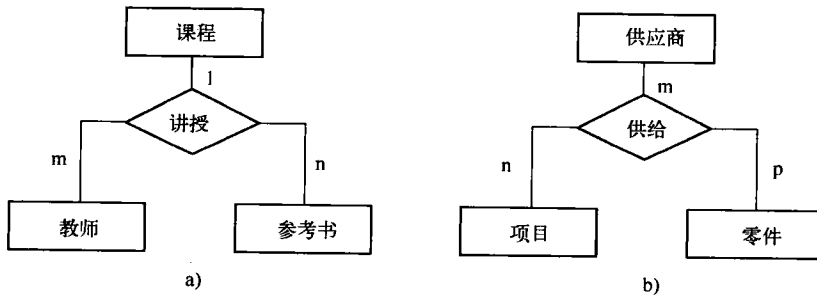


图 1-9 三个实体集联系例子
a) 一对多的联系 b) 多对多的联系

另外，在一个实体集的实体之间也存在一对一、一对多或多对多的联系。例如，职工是一个实体集，职工中有领导，而领导自身也是职工。职工实体集内部具有领导与被领导的联系，即某一个职工领导若干名职工，而一个职工仅被一个领导所管，这种联系是一对多的联系，如图 1-10 所示。

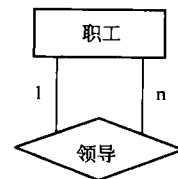


图 1-10 同一实体集内的一对多联系实例

1.3.3 常用的数据模型

数据模型是按计算机系统的观点对数据建模，是现实世界数据特征的抽象，用于 DBMS 的实现。数据库领域最常用的数据模型主要有 4 种，它们