



高等学校计算机类“十二五”规划教材

数据库原理及应用

◎主 编 唐 友 文雪巍
◎副主编 张 鑫 郝国良
司震宇 耿 姝
◎主 审 张金柱

SHUJUKU YUANLI JI
YINGYONG



西安电子科技大学出版社
<http://www.xduph.com>

高等学校计算机类“十二五”规划教材

数据库原理及应用

主 编 唐 友 文雪巍

副主编 张 鑫 郝国良 司震宇 耿 姝

主 审 张金柱

西安电子科技大学出版社

内 容 简 介

本书从数据库的基本理论、基本知识出发,通过丰富的实例介绍数据库的分析、设计过程以及开发应用等。本书主要介绍数据库的基本原理与基础知识,讲解 SQL Server 的安装、使用等各方面的知识,包括数据库的创建,表、视图和索引的使用,SQL 查询语言,Transact-SQL,存储过程和触发器,备份和恢复以及数据库的安全管理等内容,最后一章为开发与编程的案例。每章还配有习题,可帮助读者加深理解。

本书内容丰富,讲解由浅入深,例子翔实丰富。本书可作为高等院校计算机专业的教材,也可作为其他专业的“数据库应用”、“电子商务数据库”等课程的教材,还可供 SQL Server 数据库初、中级用户阅读和参考。

图书在版编目(CIP)数据

数据库原理及应用/唐友,文雪巍主编. —西安:西安电子科技大学出版社,2013.8

高等学校计算机类“十二五”规划教材

ISBN 978-7-5606-3167-7

I. ① 数… II. ① 唐… ② 文… III. ① 数据库系统—高等学校—教材 IV. ① TP311.13

中国版本图书馆 CIP 数据核字(2013)第 192658 号

策 划 邵汉平

责任编辑 许青青 王晓燕

出版发行 西安电子科技大学出版社(西安市太白南路 2 号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfxb001@163.com

经 销 新华书店

印刷单位 渭南邮电印刷厂

版 次 2013 年 8 月第 1 版 2013 年 8 月第 1 次印刷

开 本 787 毫米×1092 毫米 1/16 印 张 16.5

字 数 389 千字

印 数 1~3000 册

定 价 28.00 元

ISBN 978-7-5606-3167-7/TP

XDUP 3459001-1

如有印装问题可调换

本社图书封面为激光防伪覆膜,谨防盗版。

前 言

本书以微软公司推出的 SQL Server 大型网络关系数据库为平台,结合普通高校计算机专业数据库课程的具体要求,深入浅出地介绍了 SQL Server 的使用方法和开发技巧。本书共 10 章,分别介绍了数据库基础概述,SQL Server 2000 简介,SQL Server 数据库,表、视图和索引的基本操作,结构化查询语言 SQL, Transact-SQL,存储过程和触发器,备份与恢复,数据库管理以及人力资源管理系统。通过学习本书,读者可以了解和掌握 SQL Server 非常强大的关系数据库创建、设计以及数据库系统的开发和管理等操作。

本书编者多年来一直从事计算机数据库的教学工作,本书正是编者总结多年的教学实践编写而成的。针对初学者和自学者的特点,本书力求通俗易懂,用大量具体的操作、各种不同的实例让读者进入 SQL Server 的可视化编程环境。本书全面介绍了网络数据库技术的基本理论、实现方法、设计过程以及开发应用等。在内容编排上,本书从理论到实践、从技术基础到综合实例,循序渐进、由浅入深。由于 SQL Server 的核心是 SQL 语句,因此本书对 SQL 语句的讲解简明扼要,资料翔实,并且配有非常丰富的例子加以说明,这些例子都是针对同一个数据库里的数据来进行的,这样可以让读者对 SQL 语句有整体把握。

本书突出案例教学并配有《数据库原理及应用实践指导》。在理论讲解过程中配有大量实例,通过一个个实例的分析和操作,使读者在理解所学知识的基础上,掌握数据库应用系统的开发方法。各章(除第 10 章外)后均附有本章小结和习题,供读者练习。

本书是黑龙江省高等教育教学改革项目《基于应用型本科院校创新型卓越软件人才培养模式的研究》(JG2012010106)阶段性研究成果之一,还是物流教改教研课题《基于卓越计划的创新型物联网人才培养模式研究》(JZW2013036)阶段性研究成果之一。

唐友、文雪巍任本书主编,张鑫、郝国良、司震宇、耿姝任副主编,郭鑫、张旭、张爱军、陈秀玲、车玉生、王永辉、张靓、赵鑫任参编。编写分工如下:第 1、6 章及附录 A~E 由唐友编写,第 5、7 章及第 8 章 1、2 节由文雪巍编写,第 8 章 3、4 节及第 10 章由张鑫编写,第 9 章由郝国良编写,第 4 章由司震宇编写,第 2 章由耿姝编写,第 3 章由郭鑫、张旭、张爱军、陈秀玲、车玉生、王永辉、张靓、赵鑫合作编写,最后由唐友、文雪巍统稿。

张金柱主审了本书并提出了许多宝贵意见,在此表示衷心感谢。本书在编写过程中得到了编者单位领导的大力支持和帮助,在此一并表示感谢。

由于时间仓促,加之编者水平和经验有限,书中疏漏之处在所难免,希望广大读者提出宝贵的意见和建议,以便在以后的工作中继续提高。

编 者

2013 年 5 月

目 录

第 1 章 数据库基础概述	1	2.2.2 SQL Server 2000 的安装步骤	22
1.1 数据、信息与数据处理	1	2.3 SQL Server 2000 服务器组件	27
1.1.1 数据与信息	1	2.3.1 SQL Server 2000 服务管理器	27
1.1.2 数据处理	2	2.3.2 SQL Server 2000 主要的 服务器组件简介	28
1.2 数据库系统的组成	2	2.4 SQL Server 2000 的通信组件	29
1.2.1 数据库	2	2.4.1 服务器端网络实用工具	29
1.2.2 数据库管理系统	2	2.4.2 客户端网络实用工具	30
1.2.3 数据库系统	3	2.5 SQL Server 2000 主要的管理工具	30
1.2.4 数据库系统的用户	4	2.5.1 企业管理器	30
1.2.5 数据库系统的结构	4	2.5.2 查询分析器	31
1.3 数据管理的发展	5	2.6 注册服务器	32
1.3.1 人工管理阶段	5	2.7 本章小结	35
1.3.2 文件系统阶段	5	习题 2	35
1.3.3 数据库系统阶段	5		
1.4 数据模型	7	第 3 章 SQL Server 数据库	36
1.4.1 数据模型概述	7	3.1 SQL Server 数据库简介	36
1.4.2 概念数据模型	7	3.1.1 数据库对象	36
1.4.3 逻辑数据模型	9	3.1.2 数据库类型	37
1.5 关系型数据库及其设计	11	3.1.3 文件	37
1.5.1 关系型数据库定义	11	3.1.4 文件组	37
1.5.2 关系型数据库与表	11	3.2 创建数据库	38
1.5.3 主键与外键	11	3.2.1 使用 Transact-SQL 创建数据库	38
1.5.4 字段约束	12	3.2.2 使用企业管理器创建数据库	41
1.5.5 数据完整性	13	3.2.3 使用向导创建数据库	42
1.5.6 表的关联	14	3.3 修改数据库	45
1.5.7 数据库设计过程	15	3.3.1 使用 Transact-SQL 修改数据库	45
1.5.8 关系型数据库规范化分析	16	3.3.2 使用企业管理器修改数据库	47
1.6 本章小结	19	3.4 删除数据库	48
习题 1	19	3.5 本章小结	49
		习题 3	49
第 2 章 SQL Server 2000 简介	20		
2.1 SQL Server 2000 的特点	20	第 4 章 表、视图及索引的基本操作	50
2.2 SQL Server 2000 的安装	21	4.1 SQL Server 中的数据类型	50
2.2.1 SQL Server 2000 的硬件和 软件安装要求	21	4.1.1 数据类型	50

4.1.2 空置的含义	55	5.3.8 使用 ORDER BY 子句对 查询结果排序	94
4.2 表的基本操作	55	5.3.9 SQL 的聚合函数	95
4.2.1 创建表	55	5.3.10 使用 GROUP BY 子句对 查询结果进行分组	95
4.2.2 修改表结构	57	5.3.11 使用 HAVING 子句筛选结果表	96
4.2.3 建立表间的联系	58	5.4 基于多表的联接查询	96
4.2.4 插入表中的记录	60	5.5 子查询	101
4.2.5 修改表中的记录	61	5.6 在查询的基础上创建新表	105
4.2.6 删除表中的记录	61	5.7 本章小结	105
4.2.7 删除表	62	习题 5	105
4.3 视图的基本操作	63		
4.3.1 视图的概念	63	第 6 章 Transact-SQL	107
4.3.2 创建视图	64	6.1 常量	107
4.3.3 修改视图	68	6.2 变量	108
4.3.4 使用视图查询数据	69	6.3 流程控制语句	109
4.3.5 使用视图管理数据表中的数据	70	6.4 函数	112
4.4 索引的基本操作	74	6.5 用户定义函数	116
4.4.1 索引的概念	74	6.6 游标	119
4.4.2 索引的类型	75	6.7 本章小结	121
4.4.3 创建索引	76	习题 6	121
4.4.4 创建复合索引	84		
4.4.5 查看、修改和删除索引	84	第 7 章 存储过程和触发器	122
4.5 本章小结	86	7.1 存储过程	122
习题 4	86	7.1.1 存储过程概述	122
		7.1.2 创建存储过程	122
第 5 章 结构化查询语言 SQL	88	7.1.3 执行存储过程	128
5.1 SQL 简介	88	7.1.4 存储过程的参数	129
5.1.1 SQL 概述	88	7.1.5 存储过程的查看、修改和删除	132
5.1.2 SQL 分类	88	7.2 触发器	134
5.2 SELECT 查询语句	89	7.2.1 触发器概述	134
5.3 基于单表查询	89	7.2.2 创建触发器	134
5.3.1 查询表中指定的字段	90	7.2.3 inserted 表和 deleted 表	139
5.3.2 通配符“*”的使用	90	7.2.4 使用触发器	140
5.3.3 使用单引号加入字符串	90	7.2.5 修改触发器	142
5.3.4 使用别名	91	7.2.6 删除触发器	143
5.3.5 显示表达式的值	91	7.2.7 嵌套触发器	143
5.3.6 使用 DISTINCT 短语 消除重复的记录	91	7.3 本章小结	144
5.3.7 使用 WHERE 子句 查询特定的记录	91	习题 7	144

第 8 章 备份与恢复	145	9.3.4 两段锁协议	177
8.1 备份和恢复概述	145	9.3.5 SQL Server 的并发控制.....	178
8.1.1 数据库备份的基本概念	145	9.4 安全性	179
8.1.2 数据库恢复的基本概念	145	9.4.1 安全性概述	179
8.2 备份操作和备份命令	146	9.4.2 安全性控制	180
8.2.1 创建备份设备	146	9.4.3 统计数据库安全性	183
8.2.2 备份命令	147	9.4.4 SQL Server 的安全性管理.....	184
8.2.3 使用企业管理器进行备份	148	9.5 完整性	192
8.2.4 使用备份向导进行备份	149	9.5.1 完整性约束条件	193
8.3 恢复操作和恢复命令	154	9.5.2 完整性控制	194
8.3.1 检查点	154	9.5.3 SQL Server 的完整性策略.....	197
8.3.2 数据库的恢复命令	155	9.6 本章小结	199
8.3.3 使用企业管理器恢复数据库	156	习题 9	199
8.4 本章小结	158		
习题 8	158	第 10 章 人力资源管理系统	201
第 9 章 数据库管理	159	10.1 系统分析	201
9.1 数据库管理概述	159	10.2 系统设计	204
9.2 数据库恢复技术	159	10.3 系统实现	215
9.2.1 数据库恢复基本概念	162	10.4 本章小结	234
9.2.2 恢复实现技术	163	附录 A 常用语句	235
9.2.3 恢复策略	165	附录 B 常用函数	238
9.2.4 数据库镜像	166	附录 C @@类函数	241
9.2.5 SQL Server 数据恢复技术.....	167	附录 D 系统存储过程	243
9.3 并发控制	169	附录 E 扩展存储过程	253
9.3.1 并发控制概述	169	参考文献	256
9.3.2 封锁协议	171		
9.3.3 并发调度的可串行性	176		

第1章 数据库基础概述

数据库技术是计算机科学中的一个重要分支，它的应用非常广泛，几乎涉及所有的应用领域。要想掌握好数据库系统技术，首先必须弄清数据、数据管理、数据库、数据模型和概念模型等专业术语的内涵，了解数据库的发展过程和数据库系统的特点，弄清数据库、数据库管理系统和信息管理系统三者之间的关系。本章介绍了这些数据库系统的基本概念和基础知识。

1.1 数据、信息与数据处理

1.1.1 数据与信息

对我们每个人来说，“信息”和“数据”是非常重要的东西。“信息”可以告诉我们有用的事实和知识，“数据”可以更有效地表示、存储和抽取信息。

1. 数据

数据是用于承载信息的物理符号。这就是说，数据是信息的一种表现形式，数据通过能书写的信息编码表示信息。尽管信息有多种表现形式，它可以通过手势、眼神、声音或图形等方式表达，但数据是信息的最佳表现形式。由于数据能够书写，因而它能被记录、存储和处理，并从中挖掘出更深层的信息。必须指出的是，在不严格的情况下，往往把“数据”和“信息”两个概念混为一谈，称为数据信息。其实数据不等于信息，数据只是信息表达方式中的一种；正确的数据可表达信息，而虚假、错误的信息所表达的是谬误，不是信息。

2. 信息

在日常生活中，我们经常可以听到“信息”这个名词。什么是信息呢？简单地说，信息就是新的、有用的事实和知识，它具有实效性、有用性和知识性，是客观世界的反映。信息具有四个基本特征：一是信息的内容是关于客观事物或思想方面的知识，即信息的内容能反映已存在的客观事实，能预测未发生事物的状态，能用于指挥和控制事物的发展；二是信息是有用的，它是人们活动的必需知识，利用信息能够克服工作中的盲目性，增加主动性和科学性，可以把事情办得更好；三是信息能够在空间和时间上被传递，在空间上传递信息称为信息通信，在时间上传递信息称为信息存储；四是信息需要一定的形式表示，信息与其表现符号不可分离。

信息对于人类社会的发展有重要意义。它可以提高人们对事物的认识，减少人们活动

的盲目性；信息是社会机体进行活动的纽带，社会的各个组织通过信息网相互了解并协同工作，使整个社会协调发展，社会越发展，信息的作用就越突出；信息又是管理活动的核心，要想把事物管理好，就需要掌握更多的信息，并利用信息进行工作。

1.1.2 数据处理

围绕着数据所做的工作称为数据处理。数据处理是指对数据的收集、组织、整理、加工、存储和传播等工作。我们可以把数据处理分为 3 类：一类为数据管理，它的主要任务是收集信息，将信息用数据表示并按类别组织保存，其目的是在需要的时候，为各种应用和处理提供数据；另一类是数据加工，它的主要任务是对数据进行变换、抽取和运算，通过数据加工会得到更有用的数据，以指导或控制行为或事物的变化趋势；最后一类是数据传播，它在空间或时间上以各种形式传播信息，传播过程中不改变数据的结构、性质和内容，数据传播会使更多的人得到并理解信息，从而使信息的作用充分发挥出来。

1.2 数据库系统的组成

数据库(Data Base, DB)、数据库管理系统(Data Base Management System, DBMS)和数据库系统(Data Base System, DBS)是数据库技术中常用的术语，三者之间既有区别又有联系。

1.2.1 数据库

数据库是以一定的组织方式将相关数据组织在一起，并存储在外部存储介质上所形成的、能为多个用户共享的、与应用程序相互独立的相关数据集合。

数据库中的数据具有集中性和共享性。所谓集中性，是指把数据库看成性质不同的数据文件的集合，其中的数据冗余很小。所谓共享性，是指多个不同用户使用不同语言，为了不同应用目的可同时存取数据库中的数据。

1.2.2 数据库管理系统

数据库管理系统是以统一的方式管理、维护数据库中数据的一系列软件的集合。

数据库管理系统在操作系统的支持与控制下运行。用户一般不能直接加工和使用数据库中的数据，而必须通过数据库管理系统。数据库管理系统的主要功能是维护数据库系统的正常活动，接受并响应用户对数据库的一切访问要求，包括建立及删除数据库文件，检索、统计、修改和组织数据库中的数据以及为用户提供对数据库的维护手段等。通过使用数据库管理系统，用户可以逻辑地、抽象地处理数据，而不必关心这些数据在计算机中的存放方式以及计算机处理数据的过程细节，把一切处理数据的具体而繁杂的工作交给数据库管理系统去完成。因此，在信息素养已经成为现代人的基本素质之一的信息社会里，学习并掌握一种数据库管理系统不但重要而且必要。

数据库管理系统的功能归结起来，主要有以下几点：

1. 数据库定义(描述)功能

数据库管理系统提供数据描述语言(Data Definition Language, DDL)实现对数据库逻辑

结构的定义以及数据之间联系的描述。

2. 数据库操纵功能

数据库管理系统提供数据操纵语言(Data Manipulation Language, DML)实现对数据库检索、插入、修改、删除等基本操作。DML 通常分为两类：一类是嵌入在某种语言中的，如嵌入 C、VC++等高级语言中，这类 DML 一般不能独立使用，称之为宿主型语言；另一类是交互命令语言，它的语法简单，可独立使用，称之为自含型语言。目前，数据库管理系统广泛采用可独立使用的自含型语言，为用户和应用程序员提供操纵使用数据库的语言工具。

3. 数据库管理功能

数据库管理系统提供了对数据库的建立、更新、结构维护以及恢复等管理功能。它是数据库管理的核心部分，所有数据库的操作都要在其统一管理下进行，以保证操作的正确执行，从而保证数据库的正确有效。

4. 通信功能

数据库管理系统提供数据库与操作系统的联机处理接口以及用户与数据库的接口。作为用户与数据库的接口，用户可以通过交互式 and 应用程序方式使用数据库。交互式直接明了，使用简单，通常借助 DML 对数据库中的数据进行操作；应用程序方式则是用户或应用程序员通过文本编辑器编写应用程序，实现对数据库中数据的各种操作。

1.2.3 数据库系统

数据库系统是指在计算机系统中引进数据库后构成的系统。

数据库系统一般由四部分组成：数据库、数据库管理系统、计算机系统和人(数据库管理人员、用户)。

数据库系统的特点主要有：

(1) 数据共享。数据共享是数据库系统的目的，也是它的重要特点。数据共享是指多个用户可以同时存取数据而不相互影响。它包含三个方面的含义：所有用户可以同时存取数据；数据库不仅可以为当前用户服务，也可以为将来的新用户服务；可以使用多种语言完成与数据库的接口。

(2) 数据的独立性。数据独立是指数据与应用程序之间彼此独立，不存在着相互依赖的关系。应用程序不必随数据存储结构的改变而改变，这是数据库的一个最基本的优点。

(3) 可控冗余度。数据冗余就是数据重复。数据冗余既浪费存储空间，又容易产生数据的不一致。在数据库系统中，由于数据集中使用，从理论上说可以消除冗余，但实际上出于提高检索速度等方面的考虑，常常允许部分冗余存在。这种冗余是可以由设计者控制的，故称为“可控冗余”。

(4) 数据的一致性。数据的一致性是指数据的矛盾性。例如，员工培训管理系统中，某员工的职称信息在员工基本信息中为“讲师”，而在员工培训需求信息中为“助讲”，这就称为数据不一致。如果数据有冗余，就容易引起数据的不一致性。由于数据库能减少数据的冗余，同时提供对数据的各种检查和控制，保证在更新数据时能同时更新所有副本，因而维护了数据的一致性。

(5) 数据的安全性与完整性。数据库中加入安全保密机制,可以防止对数据的非法存取。由于实行集中控制,因此有利于控制数据的完整性。数据库系统采取了并发访问控制,保证了数据的正确性。

把反映现实世界中的事物本质的信息(如员工培训管理系统中所涉及的信息)有效组织起来,形成数据库,构建数据库系统,是一个复杂的过程。

1.2.4 数据库系统的用户

1. 系统程序员

系统程序员负责整个数据库系统的设计工作,依据用户的需求安装数据库管理系统,建立维护数据库管理系统及相关软件的工具,设计合适的数据库及表文件,并对整个数据库的存取权限作出规划。

2. 数据库管理员

数据库管理员(Database Administrator, DBA)是支持数据库系统的专业技术人员。数据库管理员的任务主要是决定数据库的内容,对数据库中的数据进行修改、维护,对数据库的运行状况进行监督,管理账号,备份和还原数据,并负责提高数据库的运行效率。

3. 应用程序员

应用程序员负责编写访问数据库的面向终端用户的应用程序,使用户可以很友好地使用数据库。可以使用 Visual Basic、Delphi、PHP、ASP 和 JSP 等来开发数据库应用程序。

4. 操作员

操作员(普通用户)只需操作应用程序软件来访问数据库,利用数据库系统完成日常的工作,不关心数据库的具体格式及其维护和管理等问题。

1.2.5 数据库系统的结构

1. 大型数据库

大型数据库由一台性能很强的计算机(称为主机或者数据库服务器)负责处理庞大的数据,用户通过终端机与大型主机相连,以存取数据。

2. 本地小型数据库

在用户较少、数据量不大的情况下,可使用本地小型数据库。小型数据库一般是由个人建立的个人数据库。常用的 Access 和 FoxPro 等即属于小型数据库。

3. 分布式数据库

分布式数据库是为了解决大型数据库反应缓慢的问题而提出的,它是由多台数据库服务器组成的,数据可来自不同的服务器。

4. 客户机/服务器数据库

在客户机/服务器数据库的网络结构中,数据库的处理可分为两个系统,即客户机(Client)和数据库服务器(Database Server),前者运行数据库应用程序,后者运行全部或者部分数据库管理系统。在客户机上的数据应用程序将该请求通过网络发送给服务器,数据库服务器

进行搜索，并将用户查询所需的数据返回到客户机。

1.3 数据管理的发展

数据处理的内容首先是数据的管理。发明计算机以后，人们一直在努力寻求如何用计算机更有效地管理数据。随着计算机硬件和软件技术的发展，计算机数据管理技术经历了从低级阶段发展到高级阶段的过程，技术上也越来越成熟。按照一般的文献划分，计算机数据管理的发展有以下几个阶段。

1.3.1 人工管理阶段

20 世纪 50 年代是第一代计算机应用阶段。当时，计算机没有磁盘这样能长期保存数据的存储设备，这个时期的数据管理是用人工方式把数据保存在卡片、纸带这类介质上，所以称为人工管理阶段。这个阶段数据管理的最大特征是数据由计算数据的程序携带，二者混合在一起，因此具有以下特点：

(1) 数据不能独立。由于数据和程序混合在一起，因此就不能处理大量的数据，更谈不上数据的独立与共享，一组数据只能被一个程序专用。此外，当程序中的数据类型、格式发生变化时，相应程序也必须进行修改。

(2) 数据不能长期保存。这个阶段计算机的主要任务是科学计算。计算机运行时，程序和数据在计算机中，程序运行结束后，数据即从计算机中释放出来。

(3) 数据没有专门的管理软件。由于计算机系统没有数据管理软件来管理数据，也就没有数据的统一存取规则。数据的存取、输入/输出方式由编写程序的程序员自己确定，这就增加了编写程序的负担。

1.3.2 文件系统阶段

随着计算机对数据处理要求的不断提高，人们对数据处理的重要性越来越重视。20 世纪 50 年代末至 60 年代，计算机操作系统中专门设置了文件系统来管理数据，计算机的数据管理进入了文件系统阶段。这个阶段的主要特征是数据文件和处理数据的程序文件分离，数据文件由文件系统管理，它确立数据文件和程序文件的接口，保证文件被正确地调用。与人工阶段相比，文件系统阶段有所进步，但还是存在以下缺点：

(1) 数据独立性差，不能共享数据。虽然从程序文件中分离了出来，但文件系统管理的数据文件只能简单地存放数据，且一个数据文件一般只能被相应的程序文件专用，相同的数据要被另外的程序使用，必须再产生数据文件，这样就出现了数据的重复存储问题，即数据冗余。

(2) 数据文件不能集中管理。由于这个阶段的数据文件没有合理和规范的结构，数据文件之间不能建立联系，使得数据文件不能集中管理，数据使用的安全性和完整性都不能保证。

1.3.3 数据库系统阶段

20 世纪 60 年代末，计算机的数据管理进入数据库系统阶段。这时，由于计算机的数

据处理量迅速增长,数据管理得到了人们的高度重视,随后在美国产生了技术成熟、具有商业价值的数据库管理系统。数据库系统不仅有效地实现了程序和数据的分离,而且把大量的数据组织在一种特定结构的数据库文件中,多个不同程序都可以调用数据库中相同的数据,从而实现了数据的统一管理及数据共享。与文件系统相比,数据库系统具有以下特点。

(1) 实现了数据共享,减少了数据冗余度。由于数据库不仅与程序文件相互独立,而且具有合理、规范的结构,使得不同的程序可以同时使用数据库中相同的数据,这样就大大节省了存储资源,减少了数据冗余度。

(2) 实现了数据独立。数据独立包括物理数据独立和逻辑数据独立。物理数据是指数据在硬件上的存储形式,其独立性是指当数据的存储结构发生变化时,不会影响数据的逻辑结构,也就不会影响程序的运行。逻辑数据是指数据在用户面前的表现形式,当逻辑数据结构发生变化时也不会影响应用程序,这就是逻辑数据的独立性。这两种数据的独立性有效地保证了数据库运行的稳定性。

(3) 采用合理的数据结构加强了数据的联系。数据库采用了合理的结构来安排其中的数据,不仅同一数据文件中的数据之间存在特定的联系,各数据文件之间也可以建立联系,这是文件系统不能做到的。

(4) 加强了数据保护。与文件系统相比,数据库系统增加了数据的多种控制功能。例如,并发控制能保证多个用户同时使用数据时不发生冲突;安全性控制能保证数据安全,不被非法用户使用和破坏;数据的完整性控制能保证数据使用过程中的正确性和有效性。

值得一提的是,有的文献把数据库系统阶段又分为集中式数据库系统阶段和分布式数据库系统阶段。早期的数据库系统是集中式的,其特点是把所有的数据无论在物理上还是在逻辑上都集中摆放在一起。这样虽然设计简单,但影响数据的流通速度。

随着计算机网络技术的高速发展,现在更多的数据库系统采用分布式的数据库系统,通过网络技术把分布在各处的计算机连接起来,数据库中的数据在物理上分布于网络中不同计算机结点上。但对用户使用来说,他不知道也不用关心数据存放在哪个地方,逻辑上看起来好像是在集中使用。分布式数据库系统提高了数据的使用效率,加快了数据的流通速度,更加符合今天人们对数据处理的需要。

关于分布式数据库系统的网络工作模式,现在使用较多的是客户机/服务器模式。在这种模式中,数据及数据处理程序放在数据服务器上,业务处理程序 and 用户界面放在客户机上。客户机/服务器模式数据库系统的结构如图 1-1 所示。数据库管理系统支持这种模式,并为开发功能强大的客户机/服务器模式的应用程序提供了专门的工具。

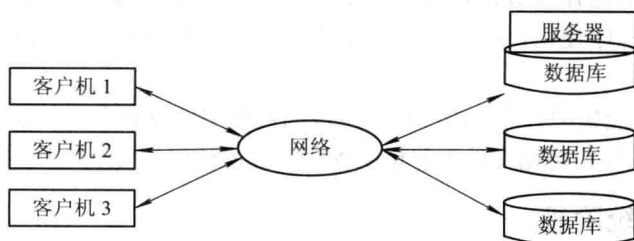


图 1-1 客户机/服务器模式数据库系统

1.4 数据模型

1.4.1 数据模型概述

数据模型是对现实世界数据特征的抽象，是用来描述数据的结构和联系的一组概念和定义，是数据库的核心内容。

由于计算机不能直接处理现实世界中的具体事物，所以必须把具体事物转换成计算机能够处理的数据。在数据库系统中，实现转换的过程通常是先把现实世界中的客观事物抽象为概念数据模型(简称概念模型)，然后再把概念数据模型转换为某一数据库管理系统所支持的逻辑数据模型(简称逻辑模型)。

概念数据模型和逻辑数据模型是数据模型的不同应用层次。概念数据模型是从现实世界到数据世界的一个中间层次，是一种面向客观世界、面向用户的模型，是数据库设计人员进行数据库设计的重要工具，也是数据库设计人员和用户之间进行交流的语言，E-R模型、扩充的E-R模型等是常用的概念模型。逻辑数据模型是一种面向数据库系统的模型，即依赖于某种具体的数据库管理系统(DBMS)，主要用于DBMS的实现，常用的逻辑数据模型包括层次模型、网状模型和关系模型等。

1.4.2 概念数据模型

数据库的内容是经过抽象、收集产生的，用于反映现实世界中事物及联系。现实世界中的事物反映到人们头脑中，产生想法、概念是一个抽象过程，在此抽象过程中所用的方法有丰富的含义，它使人们从现实世界进入信息世界(概念世界)，由信息世界再经过加工，并用一定的方法来表示，使得数据、信息能进入计算机世界(数据世界)。数据处理领域的三个世界可用图1-2来表示。

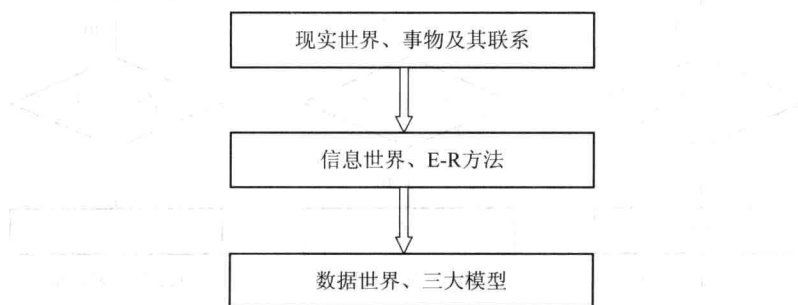


图 1-2 数据处理的三个世界

信息世界是现实世界在人们头脑中的反映。客观事物在信息世界中称为实体，信息世界的主要对象是实体以及实体间的相互联系，描述实体、实体属性以及实体之间相互联系的方法称为实体-联系模型(E-R模型、E-R图)，也叫概念模型。

1. 实体与属性

客观事物在信息世界中称为实体，它是现实世界中客观存在并可以相互区别的事物。

实体可以是具体的人或物，也可以是抽象概念，如学生成绩管理系统中的实体就有学生实体、课程实体和成绩实体。

属性是实体的特征。一个实体总是通过其属性来描述的。如学生成绩管理系统中学生实体的属性有学号、姓名、性别和出生日期等。

实体集是指同类实体的集合，即具有同一类属性的客观存在的事物的集合。在对管理对象进行分析时，不是针对个别实体，而是对实体集进行的。

2. 实体间联系

因为现实世界中的客观事物之间是彼此联系的，因此在信息世界中实体之间也是相互联系的。实体间的联系方式通常有三种：一对一联系(1:1)、一对多联系(1:n)和多对多联系(m:n)。

1) 一对一联系(1:1)

在两个不同型的实体集中，任意一方的一个实体只与另一方的一个实体相对应，则称这种联系为一对一联系。如班长与班级的联系，一个班级只有一个班长，一个班长对应一个班级。如图 1-3(a)所示。

2) 一对多联系(1:n)

在两个不同型的实体集中，一方的一个实体对应另一方的若干个实体，而另一方中的一个实体只对应本方的一个实体，则称这种联系为一对多联系。如班长与同学的联系，一个班长对应多个同学，而本班的每个同学只对应一个班长，如图 1-3(b)所示。

3) 多对多联系(m:n)

在两个不同型的实体集中，任意一方的一个实体均对应另一方的若干个实体，则称这种联系为多对多联系。如教师与学生的联系，一位教师为多个学生授课，而每个学生也有多位任课老师，如图 1-3(c)所示。

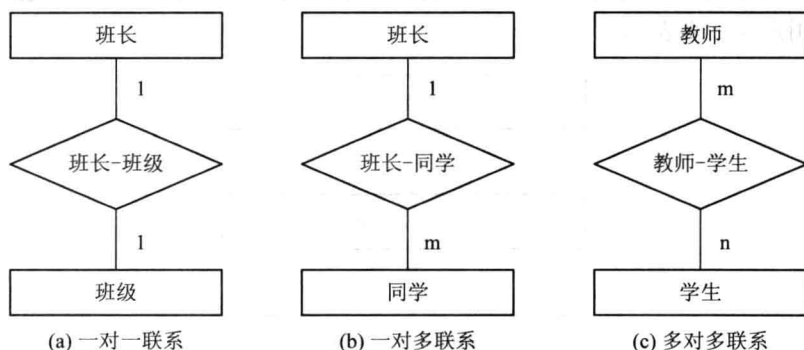


图 1-3 实体间的联系模型

3. E-R 模型(E-R 图)

使用 E-R 模型的核心是划分实体和属性，并确定实体间的联系。其表示方法为：① 实体集用矩形框表示，框内写上实体名；② 属性用椭圆形框表示，框内写上属性，并用一条无指向线标出实体与属性的联系；③ 实体间的联系用菱形框表示，框内写上实体间的联系名，并用无指向线将菱形框分别与有关的实体相连接。

学生成绩管理系统中的各实体 E-R 图如图 1-4~图 1-6 所示。各实体之间的关系 E-R 图如图 1-7 所示。

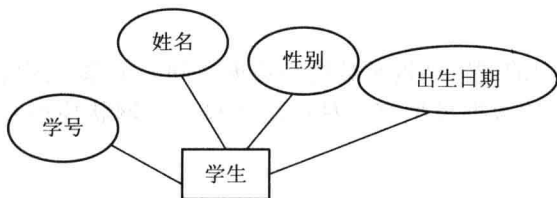


图 1-4 学生实体 E-R 图

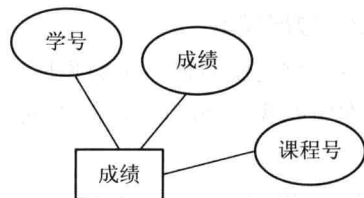


图 1-5 成绩实体 E-R 图

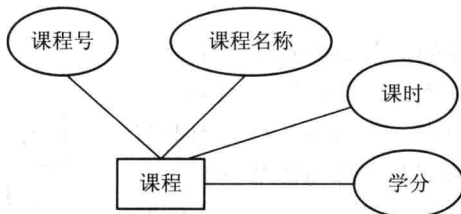


图 1-6 课程实体 E-R 图

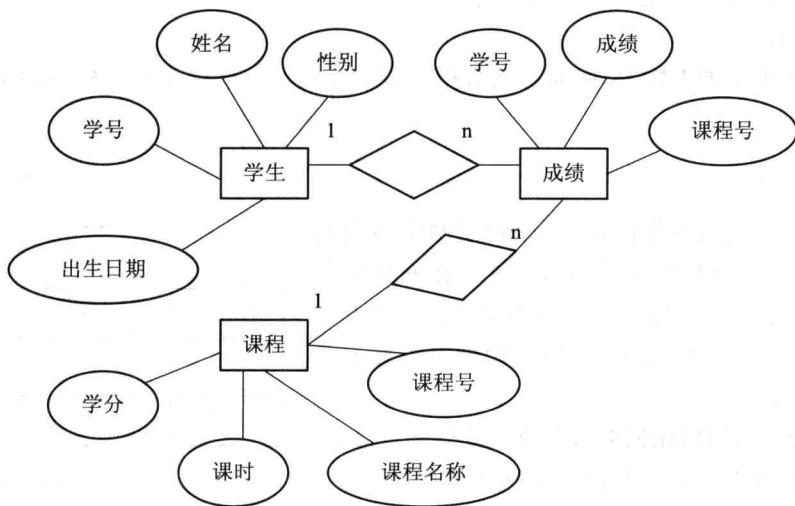


图 1-7 各实体之间的关系 E-R 图

1.4.3 逻辑数据模型

数据世界是在信息世界基础上的进一步抽象。数据模型是由信息模型转换而来的，数据模型是客观事物及其联系在数据世界中的描述，是对数据库中的数据进行逻辑组织的方法或数据库中数据间的逻辑结构。数据库设计的核心问题就是要设计一个好的数据模型。因此，应该了解与数据模型设计有关的问题。

1. 记录与数据项

在数据模型中，用数据描述的实体有对象(客观世界中的任何事物)与属性之分。描述

对象的数据称为记录，而描述属性的数据称为数据项。由于一个对象具有若干属性，因此记录就由若干数据项组成。任何数据项都包含属性名、数据类型和数据长度三个特征。

2. 数据模型

由于数据模型描述了数据库中数据间的逻辑结构及数据与数据之间的关系，因此根据数据库中数据之间的关系的不同，常用的数据模型有三种：层次模型、网状模型和关系模型。

1) 层次模型

用树形结构表示实体之间联系的模型称为层次模型，如图 1-8 所示。

树是由节点和连线组成的，节点表示实体集合，连线表示相连两实体间的联系，但只能是 1:1 或 1:m 的联系。树的最高位置只有一个节点，称之为根节点，任何一个节点的上层紧邻节点称为该节点的父节点，下层紧邻节点称为该节点的子节点。在层次模型中，有且只有根节点而无父节点，除根节点外，任何节点必须有且只有一个父节点，同时可有一个或多个子节点与其相连接。

支持层次模型的 DBMS 称为层次数据库管理系统，在这种系统中建立的数据库是层次数据库。

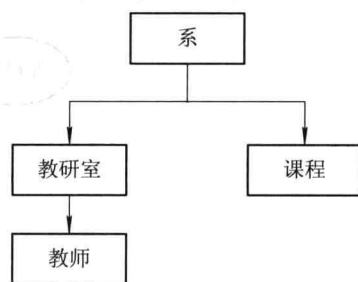


图 1-8 层次模型

2) 网状模型

网状模型是层次模型的拓展，是使用网络结构表示实体之间联系的模型，如图 1-9 所示。在网状模型中，允许有一个或一个以上节点没有父节点，至少有一个节点有多于一个的父节点。网状模型可以表示多对多联系(m:n)。

支持网状模型的 DBMS 称为网状数据库管理系统，在这种系统中建立的数据库是网状数据库。

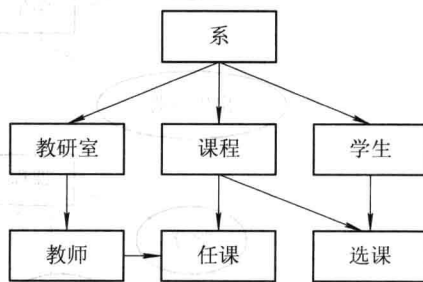


图 1-9 网状模型

3) 关系模型

用二维表格来表示一组相关的数据，既简单直观，又符合人们的习惯。用二维表格表示实体及其相互联系的模型，称之为关系模型。关系模型与前两种模型的主要差别在于它们表示实体联系的方法不同。在关系模型中，用于表示实体及其相互联系的二维表称为关系。关系不但可以表示实体间一对一联系(1:1)、一对多联系(1:m)，而且通过建立关系间的关联，还可以表示多对多联系(m:n)。

关系模型是建立在关系代数基础上的，因而具有坚实的理论基础，与层次模型和网状模型相比，具有数据结构单一、理论严密、使用方便、易学易用的特点。因此，目前绝大多数数据库系统的数据模型都采用关系模型。支持关系模型的 DBMS 称为关系型数据库管理系统。20 世纪 80 年代开始，几乎所有新开发的数据库系统都是关系数据库系统，随着