

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

数据库 原理与开发

Database Principles and Developments

赵明砚 单世民 赵凤强 编著

- 采用MS SQL Server 2005和 .NET 2005, 平台先进
- 分为基础篇、应用篇和实战篇三部分, 结构合理
- 强调实用性、应用性和开发能力培养, 特色鲜明



高校系列

 人民邮电出版社
POSTS & TELECOM PRESS

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

数据库 原理与开发

Database Principles and Developments

赵明砚 单世民 赵凤强 编著



高校系列

人民邮电出版社
北京

人民邮电出版社
样书
专用章

图书在版编目 (CIP) 数据

数据库原理与开发 / 赵明砚, 单世民, 赵凤强编著. —北京: 人民邮电出版社, 2008.8
21 世纪高等学校计算机规划教材
ISBN 978-7-115-18106-0

I. 数… II. ①赵…②单…③赵… III. 关系数据库—数据库管理系统, SQL Server 2005—高等学校—教材
IV. TP311.138

中国版本图书馆 CIP 数据核字 (2008) 第 067397 号

内 容 提 要

本书共分 3 部分。第 1 部分是基础篇, 由第 1 章至第 6 章组成, 内容包括数据库系统概述、关系数据库理论、关系模型数据操作标准语言、数据库的物理存储、数据库设计和关系数据库管理系统; 第 2 部分是应用篇, 由第 7 章至第 10 章组成, 内容包括 SQL Server 2005 应用基础、存储过程及触发器、数据库安全性和数据恢复与数据转移; 第 3 部分是实战篇, 由第 11 章和第 12 章组成, 介绍了数据库的各种访问接口及其实现, 并通过一个实际的项目——图书馆管理系统的实现, 从需求分析、数据库设计到程序设计与实现完整地介绍数据库系统的开发方法。本书附录 A 部分给出了一个具体的应用领域——物流管理领域的实际业务过程; 附录 B 给出了系统数据库的逻辑结构, 供大家参考。

本书实用性较强, 在学习理论知识的同时结合 SQL Server 2005 的应用, 通过 .NET 2005 开发平台实现一个数据库系统的开发, 使读者能够真正独立地开发数据库系统。

本书可作为普通高等院校计算机科学与技术、软件工程、网络工程等专业数据库相关课程的教材, 也可作为相关院校非计算机理工类专业数据库课程的教材。

21 世纪高等学校计算机规划教材

数据库原理与开发

- ◆ 编 著 赵明砚 单世民 赵凤强
责任编辑 蒋 亮
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京楠萍印刷有限公司印刷
- ◆ 开本: 787×1092 1/16
印张: 14.25
字数: 371 千字
印数: 1—3 000 册

2008 年 8 月第 1 版

2008 年 8 月北京第 1 次印刷

ISBN 978-7-115-18106-0/TP

定价: 24.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

反盗版热线: (010)67171154

现今社会对人才的基本要求之一就是应用计算机的能力。在高等学校,培养学生应用计算机的能力,主要是通过计算机课程的体制改革,即计算机教学分层、分类规划与实施;密切联系实际,恰当体现与各专业其他课程配合;教学必须以市场需求为导向,目的是培养高素质创新型人才。

人民邮电出版社经过对教学改革新形势充分的调查研究,依据目前比较成熟的教学大纲,组织国内优秀的有丰富教学经验的教师编写一套体现教学改革最新形势的“高校系列计算机教材”。在本套教材的出版过程中,我社多次召开教材研讨会,广泛听取了一线教师的意见,也邀请众多专家对大纲和书稿做了认真的审读与研讨。本套教材具有以下特点。

1. 覆盖面广, 突出教改特色

本套教材主要面向普通高等学校(包括计算机专业和非计算机专业),是在经过大量充分的调研基础上开发的计算机系列教材,涉及计算机教育领域中的所有课程(包括专业核心骨干课程与选修课程),适应了目前经济、社会对计算机教育的新要求、新动向,尤其适合于各专业计算机教学改革的特点特色。

2. 注重整体性、系统性

针对各专业的特点,同一门课程规划了组织结构与内容不同的几本教材,以适应不同教学需求,即分别满足不同层次计算机专业与非计算机专业(如工、理、管、文等)的课程安排。同时本套教材注重整体性的策划,在教材内容的选择上避免重叠与交叉,内容系统完善。学校可根据教学计划从中选择教材的各种组合,使其适合本校的教学特点。

3. 掌握基础知识, 侧重培养应用能力

目前社会对人才的需要更侧重于其应用能力。培养应用能力,须具备计算机基础理论、良好的综合素质和实践能力。理论知识作为基础必须掌握,本套教材通过实践教学与实例教学培养解决实际问题的能力和知识综合运用能力。

4. 教学经验丰富的作者队伍

高等学校在计算机教学和教材改革上已经做了大量的工作,很多教师在计算机教育与科研方面积累了相当多的宝贵经验。本套教材均由有丰富教学经验的教师编写,并将这些宝贵经验渗透到教材中,使教材独具特色。

5. 配套资源完善

所有教材均配有 PPT 电子教案,部分教材配有实践教程、题库、教师手册、学习指南、习题解答、程序源代码、演示软件、素材、图书出版后要更新的内容等,以方便教与学。

我社致力于优秀教材的出版,恳请大家在使用的过程中,将发现的问题与提出的意见反馈给我们,以便再版时修改。

前言

国家批准成立软件学院已经有 10 年之久,在这 10 年间,随着 IT 行业及软件学院特色教育的发展,跟其他院系相比,软件学院更多地向 IT 行业输送应用性较强的社会人才,逐步形成了软件学院自己独特的实用性较强的教学模式和风格,这就要求软件学院应该采用实用性更强的教材,本书就是应这种需求而撰写的。

本书共分 3 部分。第 1 部分是基础篇,这部分由第 1 章~第 6 章组成,包括数据库管理的发展历程、数据库系统的基本概念,数据库结构、关系数据模型的数据库理论、关系模型的数据结构、关系模型的完整性约束、关系代数数据操作、规范化理论、标准 SQL 语言、数据库的物理组织、数据库设计过程及设计实例、数据库管理系统的原理知识等内容。这一部分的内容是数据库系统的理论基础,是进行数据库设计以及数据库系统实现的原理知识。但是这部分内容不容易理解,需要与后续部分的内容前后呼应、贯通学习。

第 2 部分是应用篇,介绍了关系数据库管理系统 MS SQL Server 2005 的基本操作与数据库的程序开发。这部分由第 7 章~第 10 章组成,包括 SQL Server 2005 的版本介绍、SQL Server 2005 的安装、SQL Server Configuration Manager 的操作使用、SQL Server Management Studio 的操作使用、SQL Server Query Analyzer 的操作使用、存储过程及触发器的操作和开发、数据库安全性管理的实施、数据库的恢复和数据转移等功能的操作使用等内容。这一部分的内容实用性较强,需要大家在上机操作过程中学习体会,并熟练掌握。最好在学习这部分内容的时候结合第 1 部分的知识内容,能够更进一步加深对数据库原理知识的理解,从而更好地指导实际应用。

第 3 部分是实战篇,介绍了数据库的各种访问接口及其实现,并通过一个实际的项目——图书馆管理系统的实现,从需求分析、数据库设计到程序设计与实现完整地介绍了数据库系统的开发方法。这部分由第 11 章和第 12 章组成,这一部分主要侧重数据库系统的实现,而并不是讲解开发工具的具体语言,因此大家在学习这一部分时,需要参照.NET 的 C#语言来同时学习。

本书的附录 A 部分给出了一个具体的应用领域——物流管理领域的实际业务过程。物流管理系统是一个典型的数据库应用系统,作者负责分析、设计、开发实现了该系统。在第 5 章中给出了该应用系统的数据库设计思路及设计过程,在设计该应用实例时并没有完整给出设计方案,希望读者能够在设计思路及设计过程的指导下独立分析设计完成其余部分。本书在附录 B 中给出了系统数据库的逻辑结构,供大家参考。

本书在多处给读者留了思考问题,提出这些问题,一方面是可以复习过去的知识,另一方面是可以将知识融汇在一起,提高读者分析问题的能力。

本书不仅使读者掌握数据库系统的原理知识,而且使读者掌握一个具体的关系数据库管理系统的基本应用,通过.NET 2005 开发平台实现一个数据库系统的开发,能够真正独立地开发数据库系统。

为了便于教师使用本书教学,本书为教师免费提供电子教案及有关程序代码,可登录人民邮电出版社的网站下载。

本书在编写过程中,胡颖涛、向楷等从不同方面分别做了一些工作,在此表示感谢!特别感谢王秀坤教授对本书所提出的改进意见,使得本书内容更为充实、合理。

由于水平所限,书中难免有不足之处,希望读者及同仁给予批评指正。

编者

2008年7月

目 录

第1篇 基础篇

第1章 数据库系统概述.....2

- 1.1 数据管理的形成和发展.....2
- 1.2 数据库系统的基本概念.....5
 - 1.2.1 信息及信息模型.....5
 - 1.2.2 数据及数据模型.....8
 - 1.2.3 数据库、数据库管理系统及数据库系统.....9
- 1.3 数据库系统结构.....10
 - 1.3.1 数据库系统物理结构.....10
 - 1.3.2 数据库系统的外部结构.....12
- 1.4 数据库开发人员及研究领域.....13
 - 1.4.1 数据库开发人员.....13
 - 1.4.2 数据库研究领域.....14
- 1.5 本章小结.....15
- 习题.....15

第2章 关系数据库理论.....16

- 2.1 关系模型数据结构.....16
 - 2.1.1 关系的定义.....16
 - 2.1.2 关系的形式化表示.....18
- 2.2 关系数据结构的规范化.....18
 - 2.2.1 规范化理论基本概念.....18
 - 2.2.2 范式.....19
- 2.3 关系模型的数据操作原理.....22
 - 2.3.1 关系操作概述.....22
 - 2.3.2 关系代数.....23
 - 2.3.3 关系演算.....28
- 2.4 关系的完整性.....33
- 2.5 本章小结.....34
- 习题.....34

第3章 关系模型数据操作标准语言.....36

- 3.1 SQL 概述.....36

- 3.2 数据定义.....37
 - 3.2.1 定义、删除与修改基本表.....37
 - 3.2.2 定义与删除视图.....39
 - 3.2.3 建立与删除索引.....40
- 3.3 数据查询.....41
 - 3.3.1 单表查询.....41
 - 3.3.2 多表查询.....47
 - 3.3.3 嵌套查询.....50
 - 3.3.4 集合查询.....54
 - 3.3.5 视图查询.....54
 - 3.3.6 查询语句小结.....55
- 3.4 数据更新.....55
 - 3.4.1 数据插入.....55
 - 3.4.2 修改数据.....56
 - 3.4.3 删除数据.....57
 - 3.4.4 视图更新.....57
- 3.5 数据权限控制.....58
 - 3.5.1 定义权限.....58
 - 3.5.2 收回权限.....59
- 3.6 本章小结.....59
- 习题.....59

第4章 数据库的物理存储.....61

- 4.1 存储设备.....61
 - 4.1.1 外部存储器.....62
 - 4.1.2 数据处理方式.....63
- 4.2 数据存储方式.....64
 - 4.2.1 数据文件.....64
 - 4.2.2 索引.....68
- 4.3 本章小结.....74
- 习题.....75

第5章 数据库设计.....76

- 5.1 数据库设计概述.....76

5.1.1 数据库设计的方法及特点	77
5.1.2 数据库设计的基本步骤	77
5.2 需求分析	79
5.2.1 需求分析的任务	79
5.2.2 需求分析的方法	79
5.3 概念结构设计	80
5.3.1 概念结构设计策略	80
5.3.2 局部 E-R 模型设计	81
5.3.3 全局 E-R 模型设计	81
5.3.4 精化概念模型	82
5.4 逻辑结构设计	82
5.4.1 概念模型向关系模型转换	83
5.4.2 精化数据模型	84
5.4.3 设计外模式	84
5.5 物理结构设计	85
5.5.1 确定存取方法	86
5.5.2 确定数据库的存储结构	86
5.5.3 评价物理结构	87
5.6 数据库的实施和维护	87
5.6.1 数据加载和应用程序的编制	87
5.6.2 数据库的试运行	88
5.6.3 数据库的运行和维护	88
5.7 数据库设计实例	89
5.7.1 需求分析	89
5.7.2 概念结构设计	90
5.7.3 逻辑结构设计	90

第 2 篇 应 用 篇

第 7 章 SQL Server 2005 应用基础

7.1 SQL Server 2005 概述	112
7.2 SQL Server 2005 版本介绍	113
7.3 SQL Server 2005 简洁版安装	114
7.4 SQL Server 2005 常用工具使用	116
7.4.1 配置管理器 (SQL Server Configuration Manager)	116
7.4.2 数据管理器 (SQL Server Management Studio Express)	116
7.4.3 查询分析器 (SQL Server Query Analyzer)	120

5.7.4 物理结构设计	91
5.8 本章小结	91
习题	91

第 6 章 关系数据库管理系统

6.1 事务	92
6.2 数据库恢复技术	93
6.2.1 故障的种类	93
6.2.2 数据库转储	94
6.2.3 日志文件	95
6.2.4 恢复策略	96
6.3 并发控制	97
6.3.1 并发控制概述	97
6.3.2 封锁	98
6.3.3 封锁协议	99
6.3.4 死锁	101
6.3.5 并发调度的可串行性	102
6.3.6 两段锁协议	103
6.3.7 多粒度封锁	104
6.3.8 意向锁	104
6.4 数据库安全性	105
6.4.1 数据库安全性概述	105
6.4.2 数据库安全性控制	106
6.5 关系系统的查询优化	108
6.6 本章小结	110
习题	110

7.5 本章小结	122
习题	122

第 8 章 存储过程及触发器

8.1 存储过程	123
8.1.1 创建存储过程	123
8.1.2 删除存储过程	125
8.1.3 存储过程的嵌套	125
8.1.4 存储过程实例	125
8.2 触发器	128
8.2.1 创建 DML 触发器	128
8.2.2 创建 DDL 触发器	129
8.2.3 获取触发器的信息	130

8.2.4 触发器实例	130	9.5 本章小结	144
8.3 本章小结	133	习题	144
习题	133	第 10 章 数据恢复与数据转移	145
第 9 章 数据库安全性管理	134	10.1 数据库的备份与还原概述	145
9.1 SQL Server 的安全性机制	134	10.2 恢复模式	145
9.2 登录和用户	135	10.3 数据库备份	147
9.2.1 登录身份验证模式及其设置	135	10.3.1 使用 SSMS 进行完整备份	147
9.2.2 创建登录	136	10.3.2 使用 SSMS 进行完整差异备份	148
9.2.3 创建用户	137	10.4 数据库还原	149
9.3 权限管理	138	10.4.1 还原完整备份示例	149
9.3.1 服务器权限	138	10.4.2 还原完整差异备份	149
9.3.2 数据库对象权限	138	10.5 数据库的导出与导入	150
9.3.3 数据库权限	141	10.5.1 数据库表数据导出	150
9.4 角色管理	141	10.5.2 数据库表数据导入	153
9.4.1 固定服务器角色	141	10.6 本章小结	155
9.4.2 数据库角色	143	习题	155
第 3 篇 实 战 篇			
第 11 章 数据库访问	158	12.3 应用系统的设计与实现	176
11.1 数据访问接口概述	158	12.3.1 界面设计	176
11.2 ODBC 数据访问	159	12.3.2 工程创建与代码编写	178
11.2.1 ODBC 体系结构	159	12.4 本章小结	188
11.2.2 建立 ODBC 数据源	160	习题	188
11.3 OLE DB 数据访问	162	附录 A 某物流公司业务介绍	189
11.4 ADO 数据库访问	163	A.1 某物流公司组织结构	189
11.5 JDBC 数据访问	165	A.2 人员岗位职责	189
11.6 ADO.NET 数据访问	167	A.3 仓库业务管理	192
11.7 本章小结	171	A.4 配送业务管理	199
习题	171	A.5 单据流转管理	202
第 12 章 .NET 2005 与 SQL		A.6 与财务往来业务	204
Server 2005 数据库系统		附录 B 某物流系统数据库逻辑	
开发实例	172	结构	207
12.1 图书馆管理的需求分析	172	参考文献	218
12.2 数据库的设计与实现	173		

第1篇

基础篇

- 第1章 数据库系统概述
- 第2章 关系数据库理论
- 第3章 关系模型数据操作标准语言
- 第4章 数据库的物理存储
- 第5章 数据库设计
- 第6章 关系数据库管理系统

第1章 数据库系统概述

数据库系统概述

数据库系统概述

数据库系统概述

数据库系统概述

数据库系统概述

第1章

数据库系统概述

当今,随着信息化进程的不断发展,对信息管理的要求越来越高,管理信息系统(MIS)、企业资源优化(ERP)系统、决策支持系统(DSS)等的应用,都是以数据库为核心建立起来的上层应用,信息资源已成为宝贵的物质财富,建立一个满足需求、性能卓越、稳定性高并且行之有效的数据库应用系统已成为一个团体生存和发展的基础。数据库技术得到前所未有的广泛应用,越来越多新的应用领域采用数据库技术管理数据。对于一个国家来说,数据库的建设规模、安全性和稳定性已成为衡量这个国家信息化程度的重要标志。因此数据库技术已经成为计算机技术的重要分支,数据库课程是计算机科学与技术专业、信息管理专业的重要课程。

计算机的应用大大提高了人类工作的效率,人类发展电子化与信息化的过程,实质上是将现实世界转化成机器世界,也就是虚拟世界的过程。人类在计算机上以数据的形式描述现实世界,数据库技术就是研究如何科学地将描述现实世界的数据库管理组织起来,为人们提供可共享的、安全的、可靠的数据。数据库系统就是应用数据库技术对现实世界进行管理的计算机系统。

1.1 数据管理的形成和发展

从20世纪30年代开始,人们借助计算机进行数据处理。研制计算机的初衷是利用它进行复杂的科学计算。随着计算机技术的发展,其应用远远超出了人们的想象。在应用需求的推动下,在计算机硬件、软件发展的基础上,数据管理技术经历了人工管理、文件系统管理、数据库管理系统管理3个阶段。

1. 人工管理阶段

20世纪50年代中期以前,计算机主要用于科学计算。当时外存只有纸带、卡片、磁带,没有磁盘等直接存取的存储设备;没有操作系统和管理数据的软件。

人工管理数据具有如下特点。

(1) 数据不能被大量保存。

当时计算机主要用于数学计算,程序需要的数据量不是很大,一般通过纸带载入数据,不需要将数据长期保存,并且当时的硬件状况也不允许大量数据的存储。

(2) 数据不具有独立性。

数据需要由计算程序自己定义,没有相应的数据管理软件系统。程序员需要通过程序定义数据的逻辑结构、物理结构,包括存储结构、存取方法、输入方式等。如果数据发生变化必须修改计算程序,数据不具有独立性。

(3) 数据无法共享。

数据是针对某一个计算程序的,一组数据只能对应一个计算程序。由于计算程序必须对数据各自定义,因此数据无法互相利用、互相参照,不能共享。

(4) 数据无结构。

由于数据不保存,针对计算程序定制,数据没有结构性可言。

2. 文件系统管理阶段

到20世纪60年代中期,出现了磁盘、磁鼓等直接存取存储的硬件设备;出现了简单的操作系统,并有了专门的数据管理软件,一般称为文件系统;处理方式上不仅有了批处理,而且能够联机实时处理。

文件系统管理数据具有如下特点。

(1) 数据可以大量长期保存。

由于有了直接存储的硬件设备,数据长期保存成为可能。

(2) 文件系统管理数据。

文件系统通过数据文件组织管理数据,把数据组织成相互独立的数据文件,能够对文件进行修改、插入和删除等操作。

(3) 程序有了一定的独立性。

程序和数据之间可以通过文件系统实现存取转换,使计算程序与数据之间有一定的独立性,程序员可以不必过多地考虑数据的物理细节,节省了开发和维护程序的工作量。

(4) 数据有了内部结构。

文件系统实现了某种类型数据的结构,但仍没有整体结构。在文件系统中,相互独立的文件的记录内部是有结构的。传统文件的最简单形式是等长同格式的记录集合。

例如:一个学生基本信息记录文件,每个记录都有如图1-1所示的记录格式。

学号	姓名	性别	专业	年龄	政治面貌	籍贯	家庭成员	奖惩情况
----	----	----	----	----	------	----	------	------

图1-1 学生认识记录文件

其中前7项数据是任何学生必须具有的而且基本上是等长的,而后两项数据的长度是可变的。

假设用文件系统来实现学生管理系统,要求对学生的基本信息和选课情况进行管理,这时我们就需要创建3个文件:F1——学生基本信息文件;F2——课程基本信息文件;F3——选课信息文件。文件中定义好了数据的简单结构,每一行代表一条记录,应用程序就是要实现对文件进行读写的操作。

思考:请大家思考如何对文件进行读写操作。

但是,文件系统仍存在以下缺点。

(1) 不支持对文件的并发访问。

难以实现多个程序对同一文件的并发访问,这主要是源于操作系统对文件访问的实现原理。

(2) 无安全控制。

难以控制用户对文件内记录的访问,不能定制可用的安全性。例如,无法定制某一用户对某一字段的修改权限等。

(3) 数据关联性差。

虽然在建立文件时是有结构的,但这种结构只是针对一个应用而言。一个学校或一个组织涉及许多应用,不仅要考虑某个应用的数据结构,还要考虑整个组织的数据结构。例如,一个学校

的信息管理系统中不仅要考虑学生的人事管理,还要考虑学籍管理、选课管理,同时还要考虑教员的人事管理、科研管理等应用,这就要求在描述数据时不仅要描述数据本身,还要描述数据之间的联系。在文件系统中,尽管其记录内部已有了某些结构,但记录之间没有联系。

(4) 数据共享性差。

在文件系统中的文件基本上是对应于一个应用程序。对于应用相同数据结构的应用程序可以实现共享,但数据结构有一点变化,就要建立新的数据文件,因此不能大规模共享数据。

(5) 数据独立性差。

数据与程序虽然有了一定的独立性,但是数据文件的数据结构一旦改变,应用程序也要跟着改变,要扩展应用程序的功能是一件较为困难的事。因此数据与程序之间仍缺乏独立性。

3. 数据库管理系统管理阶段

随着计算机应用的越来越广泛,数据规模的不断增长,应用系统功能的不断扩展,数据共享的需求不断增加,要求数据有更好的结构性、独立性、共享性。

20 世纪 60 年代后期,硬件飞速发展,已有大容量磁盘,价格开始下降;软件的规模及需求量增加,编制和维护系统软件 and 应用程序所需的成本不断增加;在处理方式上,不仅要联机实时处理,还要开始提出和考虑分布式处理。文件系统数据管理手段已经不能满足应用的需求,于是结构性、独立性、共享性更优越的数据库管理技术便应运而生,出现了统一管理数据的专门软件系统——数据库管理系统。

人们不再用单个的文件来保存数据,而是把所有的应用数据都保存在数据库中,我们在理解数据库的时候可以把它想象成存放数据的仓库,而数据库管理系统(DataBase Management System, DBMS)就是仓库保管员,数据库应用系统就是数据的加工工厂,用数据库管理系统来管理数据比文件系统管理数据具有明显的优点,从文件系统到数据库管理系统,标志着数据管理技术的飞跃。

 **思考:** 请大家思考数据库应用系统与数据库管理系统的区别。

数据库管理系统管理数据的特点如下。

(1) 数据结构化。

数据库管理系统通过数据库实现整体数据的结构化,是数据库的主要特征之一,也是数据库管理系统与文件系统的本质区别。

在数据库管理系统中,数据库是面向某一应用领域的,它是对该应用领域现实世界的描述,具有整体的结构化。

(2) 数据共享性高。

由于数据库从整体应用领域组织和描述数据,因此数据可以被多个部门、多个应用同时共享使用,并且数据库应用系统的功能易于扩展。

(3) 数据独立性高。

应用程序与存储在硬盘上的数据互不干扰,是相互独立的,一般称为数据库的物理独立性。硬盘上的数据由数据库管理系统管理,应用程序的改变不需要改变硬盘上的数据结构。

当数据库的逻辑结构改变时,一般不需要改变应用程序,从而实现了数据与应用程序的逻辑独立性。简化了应用程序的编制,使应用系统更加易于开发和维护。

(4) 数据由数据库管理系统管理。

数据由统一的数据管理软件数据库管理系统进行管理。应用程序需要与数据库管理系统进行交互实现对数据的各种操作。

(5) 能够实现并发操作。

多个用户可以同时存取数据库中的数据,数据库管理系统提供了并发控制功能来保证并发操作时数据的正确性。

(6) 数据有安全保护机制。

数据库管理系统提供了用户及登录、权限设置、存取控制等安全保护功能,使得数据能够防止由于不合法的使用造成的数据泄密和破坏,使每个用户只能按规定的权限,对某些数据以某些方式进行使用和处理。

(7) 数据可恢复。

当计算机系统出现各种故障使数据库遭到破坏时,数据库管理系统能够将数据库从错误状态恢复到某一已知的正确状态,这就是数据库的恢复功能。

综上所述,数据库是长期存储在计算机内有组织的可共享的大量的数据集合。它使应用程序的开发与数据的建立分离,具有很高的共享性、结构性、安全性,并且能够统一管理控制,使得数据库应用系统的开发进入一个新的时代。

目前,数据库管理系统技术的发展已经非常成熟,关系型数据库管理系统已经成为开发大规模数据库应用系统最为核心的应用基础。

下面我们对数据库的基本概念加以学习。

1.2 数据库系统的基本概念

在上一节已经提到,人类利用计算机的过程实际上是在计算机上模拟现实世界的过程。在这一过程中人类首先要认识世界,并能够准确地描述现实世界,这一过程实际上是人脑对现实世界的抽象,从人脑到电脑的过程又是更高层次的抽象过程。这两个抽象过程是否准确将对我们软件系统的成败起着至关重要的作用。

当今是信息社会,也有人说是信息爆炸的时代,我们几乎每天都在谈论信息,但信息究竟是什么呢?信息与数据有什么区别和联系呢?数据库系统与数据库管理系统有哪些不同呢?这都是我们必须解决的问题。

1.2.1 信息及信息模型

1. 信息

请看下面文字:

(1) “姚明身高为 2.29m。”

(2) “在对小牛的比赛中姚明得到 19 分, 8 个篮板。”

这两句话都有数据,都是通过数据告诉我们信息。看看这两句话有没有什么区别?

第一句话通过数据告诉我们姚明的身高,除此之外,在我们的大脑中也反映出姚明真高,为什么会这样呢,因为我们知道人的平均身高也就是 1.70 ~ 1.90m。

第二句话通过数据告诉我们姚明比赛的得分情况,但是我们看不出这场比赛姚明到底打得好还是不好。

通过以上例子,可以看出信息是数据在人脑中的反映。数据来源于现实世界,因此信息是人脑对现实世界的反映,当然它的物理介质是人脑,没有人脑就没有信息,数据是人脑对现实世界

的抽象。数据不一定非要是数字,比如“人类的性别有男女”,“男,女”就是文字数据。

2. 信息模型

汽车模型、建筑模型、飞机航模都是具体的模型。看到模型,就会使人联想到真实的物体,但汽车模型并不是真正的汽车,因为它不具备汽车的功能,之所以能够使人联想,是因为模型具备了真实物体的典型特征,汽车模型具备汽车的典型外表特征,这些典型特征已经能够形象地描述该真实物体,这就是抽象。模型是现实世界物体特征的模拟,是对现实世界物体的抽象。信息模型(Data Model)也是一种模型,它是现实世界信息特征的抽象。

信息模型也称为概念模型,是现实世界到计算机世界的一个中间层次,是人脑对现实世界的理解,是对描述现实世界的信息进行建模,它是向计算机世界转化的第一步,是表达现实世界构成特征的重要工具,是数据库设计人员进行数据库设计的有力工具,也是数据库设计人员和用户之间进行交流的语言,因此概念模型不仅应该具有较强的语义表达能力,能够方便、直接地表达现实世界中的各种语义知识,而且它还应该简单、清晰、易于用户理解。

在人脑中人们通过信息来描述现实世界,信息模型就是要对人脑中的这些信息进行规范,它是现实世界在人脑中的反映,是一系列图形规则的总称,通常我们用 P.P.S.Chen 于 1976 年提出的实体—联系方法(Entity-Relationship Approach, E-R 方法)来建立信息模型。该方法用 E-R 图来描述现实世界的概念模型,因此 E-R 方法也称为 E-R 模型。这种方法得到了广泛的应用,也是描述信息模型最常用的方法。下面介绍 E-R 方法。

(1) 实体(Entity)

实体是客观存在具有公共特性并可相互区别的事物。“客观存在”不仅仅是指看得见摸得着的具体物体,也可以是某一个抽象概念,如一个老师、一个学生、一个员工、一门课程、一次订货等都是实体。实体必须能够相互区别,因此必须要唯一标识一个实体。在 E-R 方法中实体用矩形表示,矩形框内写明实体名。

(2) 属性(Attribute)

人们在描述一个具体事物时,总是要描述一些该事物的典型特性,描述实体的这些特性称为属性。例如学生实体可以由学号、姓名、性别、政治面貌等属性组成。(94002268, 张山, 男, 团员)这些属性组合起来表征了一个学生。对于不同的用户,描述实体的属性可能会存在差异,这主要是不同用户所关注的实体特性有所不同,例如,学校的教务处和财务处所关注的学生属性是不同的,但是对于数据库设计人员,应该从整体上描述一个实体。唯一标识实体的属性集称为码。例如,学号是学生实体的码。属性的取值范围称为属性域。它与实数域、整数域是类似的概念。例如,性别的域为(男,女)。在 E-R 方法中属性用椭圆形表示,并用无向边将其与相应的实体连接起来。例如,学生实体用 E-R 图表示,如图 1-2 所示。

(3) 联系(Relationship)

在现实世界中,事物内部以及事物之间是有联系的,这些联系在信息世界中反映为实体内部的联系和实体之间的联系。实体内部的联系通常是指组成实体的各属性之间的联系。实体之间的联系通常是指不同属性的实体之间的联系。

两个实体型之间的联系可以分为 3 类。

(1) 一对一联系(1:1)

如果对于实体 A 的每一个实例,实体 B 至多有一个(也可以没有)实例与之联系,反之亦然,则称实体 A 与实体 B 具有一对一联系,记为 1:1。

例如,公司里面,一个部门只有一个经理,而一个经理只在一个部门任职,则部门与经理之

间具有一对一联系。

(2) 一对多联系 (1:n)

如果对于实体 A 的每一个实例, 实体 B 有 n 个实例 ($n \geq 0$) 与之联系, 反之, 对于实体 B 的每一个实例, 实体 A 至多只有一个实例与之联系, 则称实体 A 与实体 B 具有一对多联系, 记为 $1:n$ 。

例如, 一个部门中有若干名员工, 而每个员工只在一个部门中工作, 则部门与员工之间具有一对多联系。

(3) 多对多联系 ($m:n$)

如果对于实体 A 的每个实例, 实体 B 有 n 个实例 ($n \geq 0$) 与之联系, 反之, 对于实体 B 的每个实例, 实体 A 也有 m 个实例 ($m \geq 0$) 与之联系, 则称实体 A 与实体 B 具有多对多联系, 记为 $m:n$ 。

例如, 一门课程同时有若干学生选修, 而一个学生可以同时选修多门课程, 则课程与学生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。

在 E-R 方法中联系用菱形表示, 菱形框内写明联系名, 并用无向边分别与有关实体连接起来, 同时在无向边旁标上联系的类型 ($1:1$, $1:n$ 或 $m:n$)。用 E-R 图形来表示两个实体型之间的这 3 类联系, 如图 1-2 所示。

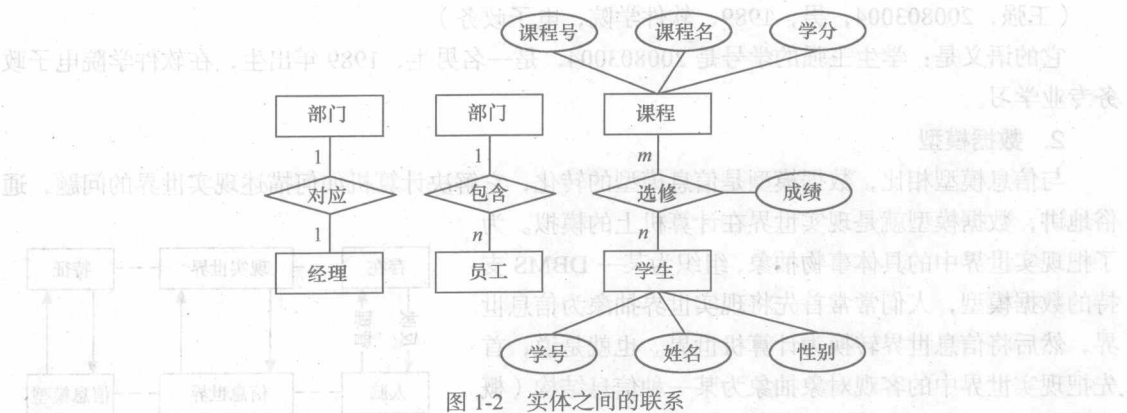


图 1-2 实体之间的联系

联系本身也可能具有属性, 则这些属性也要用无向边与该联系连接起来。如图 1-2 所示中, 课程与学生的联系。

E-R 方法也可以描述两个以上的实体之间的一对一、一对多、多对多联系。

例如, 对于课程、教师与学生 3 个实体, 如果一门课程可以有多个教师讲授, 每一个教师可以讲授多门课程, 每一个学生可以选修多门课程, 每一门课程可以有多个学生选修, 则课程与教师、学生之间的联系用 E-R 图表示, 如图 1-3 所示。

又如, 有 3 个实体型: 顾客、商品、售货员, 每个顾客可以从多个售货员那里购买多种商品; 每个售货员可以向多名顾客销售多种商品; 每种商品可由多个售货员销售给多名顾客。由此看出顾客、商品、售货员三者之间是多对多的联系。要注意, 3 个实体型之间多对多的联系和 3 个实体型两两之间的 (3 个) 多对多联系的语义是不同的。请读者给出顾客、商品、售货员 3 个实体型两两之间的多对多联系的语义, 并画出相应的 E-R 图。

同一个实体内的实例之间也可能存在一对一、一对多、多对多的联系。例如, 学生实体内部具有学生与班长的管理联系, 即某一学生 (班长) “管理” 若干名学生, 而一个学生仅被另外一个

学生直接管理，因此这是一对多的联系，如图 1-4 所示。

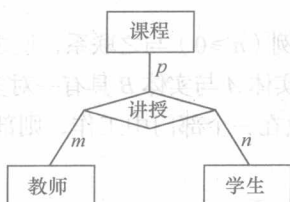


图 1-3 多个实体之间的联系

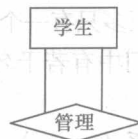


图 1-4 同一实体内部联系

1.2.2 数据及数据模型

1. 数据

由信息的定义我们可以看到，数据是描述现实世界的载体，它的形式是多种多样的。因此，数据是描述现实世界事物的符号记录。数据一定是来自现实世界的，没有现实世界也就没有数据，因此，数据必须是有语义的，所谓语义也就是描述现实世界的语言。

在计算机上是通过数据来描述现实世界的，它有特定的格式：记录，比如学生的数据记录格式：

（王强，200803004，男，1989，软件学院，电子政务）

它的语义是：学生王强的学号是 200803004，是一名男生，1989 年出生，在软件学院电子政务专业学习。

2. 数据模型

与信息模型相比，数据模型是信息模型的转化，它解决计算机如何描述现实世界的问题，通俗地讲，数据模型就是现实世界在计算机上的模拟。为了把现实世界中的具体事物抽象，组织为某一 DBMS 支持的数据模型，人们常常首先将现实世界抽象为信息世界，然后将信息世界转换为计算机世界。也就是说，首先把现实世界中的客观对象抽象为某一种信息结构（概念模型），然后再把概念模型转换为计算机上某一 DBMS 支持的数据模型，这一过程如图 1-5 所示。

数据库是反映某个应用领域的数据的综合，它不仅要反映数据本身的内容，而且要反映数据之间的联系。

数据模型是数据库的核心和基础，用数据模型这个工具

来抽象、表示和处理现实世界中的数据。各种 DBMS 软件都是基于某种数据模型的。因此，了解数据模型的基本概念是学习数据库的根源所在。

对数据模型的要求有：能比较真实地模拟现实世界；容易为人所理解；便于在计算机上实现。在数据库系统中针对不同的使用对象和应用目的，采用不同的数据模型。

（1）数据模型的类型

数据模型主要包括网状模型（Network Model）、层次模型（Hierarchical Model）、关系模型（Relational Model）、面向对象模型（Object Oriented Model）等，它是按计算机系统的观点对数据建模，主要用于 DBMS 的实现。

其中层次模型和网状模型统称为非关系模型。应用层次模型和网状模型所创建的数据库系统称

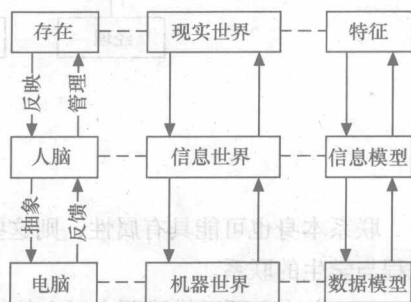


图 1-5 现实世界到机器世界的转变