

高等院校计算机基础教育规划教材

数据库原理与应用

主编 尤 峥 副主编 刘 楠 咎 超



WUHAN UNIVERSITY PRESS

武汉大学出版社

内 容 简 介

本书旨在较全面地介绍 SQL Server 2005 的基本功能、相关命令和数据库系统开发的基本方法。通过相关理论知识的学习和实践操作训练,使读者能够选择适当的开发平台,正确地、合理地进行数据库系统的开发设计,从而提高软件开发的整体质量。

本书针对职业技术教育的特点,突出计算机科学与技术学科的特点,强调理论与实践相结合的教学方式,用示例讲解的方式,引导读者学习、掌握实际应用技术。

本书适合作为高等专科学校、高等职业学校、成人高等学校以及高等院校主办的二级职业技术学院计算机及相关专业学生使用的教材。目的在于培养学生的实际动手的能力,使得学生更加适合用人单位的技能要求。

图书在版编目(CIP)数据

数据库原理与应用/尤峥主编. —武汉:武汉大学出版社,2007.5
高等院校计算机基础教育规划教材
ISBN 978-7-307-05567-4

I. 数… II. 尤… III. 数据库系统—高等学校—教材
IV. TP311.13

中国版本图书馆 CIP 数据核字(2007)第 057663 号

出版发行:武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件:wdp4@whu.edu.cn 网址:www.wdp.com.cn)

印刷:北京市昌平百善印刷厂

开本:787×1092 1/16 印张:18.75

字数:468千字 印数:1~3000册

版次:2007年5月第1版 2007年5月第1次印刷

ISBN 978-7-307-05567-4/TP·249 定价:28.00元

版权所有,不得翻印;凡购买我社的图书,如有缺页、倒页、脱页等质量问题,请与当地图书销售部门联系调换。

高等院校计算机基础教育规划教材

编 委 会

主 任：赵云冲

副主任：涂兰敬 王勰媛 高锐 韦爱荣

编 委：（按姓氏笔画排序）

丁青云	凡大林	王 伟	王艳梅
王海梅	王 娟	王朝晖	王 辉
方美秀	白海波	邢宇飞	向 蕾
刘少华	刘年超	祁昌平	孙怀东
孙贤龙	花庆毅	杨希鹏	杨 涛
李运生	李 琴	沈 丹	张国良
陈 曦	金守兵	赵天广	赵红芬
胡 俊	胡胜利	咎 超	贾云娇
钱 勇	徐 楠	殷洪菊	黄宝龙
黄 磊	梁 辉	韩丽彦	程灵枝
程灵波	程宗米	解 平	熊化武

前言

PREFACE

随着计算机技术与网络通信技术的发展,数据库技术已成为信息社会中对大量数据进行组织与管理的重要手段及软件技术,是网络信息化管理系统的基础。数据库原理及应用是计算机类专业的核心课程,也是许多专业中涉及信息处理的基础课程。

SQL Server 2005 是 Microsoft 公司推出的一个全面的数据库平台,使用集成的商业智能 (BI) 工具提供了企业级的数据管理。SQL Server 2005 数据库引擎为关系型数据和结构化数据提供了更安全可靠的存储功能,使用户可以构建和管理用于业务的高可用和高性能的数据应用程序。

本教材针对职业技术教育的特点,突出计算机科学与技术学科的特点,强调理论与实践相结合的教学方式,用示例讲解的方式,引导读者学习、掌握实际应用技术。

本教材旨在较全面地介绍 SQL Server 2005 的基本功能、相关命令和数据库系统开发的基本方法。通过相关理论知识的学习和实践操作训练,使学生能够选择适当的开发平台,正确地、合理地进行数据库系统的开发设计,从而提高软件开发的整体质量。

全书共分 13 章,主要内容包括:数据库系统的基础知识,SQL Server 2005 的安装与配置,关系型数据库通用语言 (Transact-SQL) 的基本应用,数据库系统的设计基础,SQL Server 数据库的创建和管理,数据表的创建和使用,数据的基本查询方法,数据库对象视图、索引、存储过程、触发器及游标的创建与管理,数据库的安全管理机制,数据的备份与恢复,数据的导入与导出,数据库系统的开发与发布等。

本书适合作为高等专科学校、高等职业学校、成人高等学校以及高等院校主

办的二级职业技术学院计算机及相关专业学生使用的教材。

由于编者水平有限,加之编写时间仓促,在教材中难免会有疏漏和不妥之处,欢迎读者批评指正。

编 者

2007 年 3 月

目 录

CONTENTS

第 1 章 数据库系统概述	1
1.1 引言	1
1.2 数据库技术的发展	2
1.2.1 信息、数据及数据处理	2
1.2.2 数据管理	2
1.3 数据模型	4
1.3.1 信息的三个世界	4
1.3.2 实体联系模型	7
1.3.3 逻辑模型	7
1.4 数据库系统的组成与结构	10
1.4.1 数据库系统的组成	10
1.4.2 数据库系统的结构	10
1.4.3 数据库管理系统	12
1.5 关系数据库简介	15
1.5.1 关系数据库的定义	15
1.5.2 关系模型的组成	15
1.5.3 数据操作及操作语言	15
1.5.4 关键字	16
1.5.5 完整性约束	16
1.5.6 关系数据库的特点	17
1.5.7 常见的关系数据库	17
本章小结	17
习题 1	18
第 2 章 SQL Server 2005 介绍	20
2.1 SQL Server 2005 简介	20
2.1.1 SQL Server 的发展历史	20
2.1.2 SQL Server 2005 的版本及功能比较	21

2.1.3 SQL Server 2005 的新特性	22
2.2 SQL Server 2005 的安装与卸载	24
2.2.1 系统要求	24
2.2.2 安全准备	27
2.2.3 安装与卸载 SQL Server 2005	28
2.3 SQL Server 2005 工具简介	37
2.3.1 SQL Server Management Studio	37
2.3.2 SQL 命令提示实用工具	38
2.3.3 数据库引擎优化顾问	39
2.3.4 配置管理器	39
2.3.5 外围应用配置器	39
2.3.6 SQL Server Profiler	40
2.3.7 Business Intelligence Development Studio	40
本章小结	40
习题 2	40

第 3 章 关系数据库语言

3.1 SQL 语言	42
3.2 Transact-SQL 语言程序设计基础	44
3.2.1 数据类型	45
3.2.2 运算符	50
3.2.3 注释符与通配符	51
3.2.4 T-SQL 语言流程控制命令	51
3.2.5 函数	55
3.2.6 用户自定义数据类型	60
3.3 SQL Server 2005 中 T-SQL 语言的扩展	61
本章小结	64
习题 3	65

第 4 章 数据库系统设计方法

4.1 概述	67
4.2 数据库系统的设计	68
4.2.1 数据库系统设计的目标与方法	68
4.2.2 数据库系统设计的步骤	68
4.3 数据库系统设计实例	75
4.3.1 系统需求分析	75
4.3.2 概念结构设计及数据分析	76
本章小结	81
习题 4	81

第 5 章 创建和管理数据库	83
5.1 SQL Server 数据库基础知识	83
5.2 数据库的创建	90
5.2.1 设计数据库	90
5.2.2 创建数据库	91
5.3 管理数据库	97
5.3.1 显示数据库信息	97
5.3.2 修改数据库	98
5.3.3 删除数据库	99
5.4 设置数据库的选项	100
5.5 实例解析	101
本章小结	102
习题 5	103
第 6 章 创建和管理数据表	104
6.1 数据表概述	104
6.2 创建数据表	105
6.2.1 设计数据表	105
6.2.2 创建数据表	107
6.3 数据表的管理	110
6.3.1 数据表的更名与删除	110
6.3.2 修改数据表结构	110
6.3.3 数据的添加、更新、删除	114
6.4 数据完整性概述	116
6.4.1 数据完整性的概念	116
6.4.2 约束	117
6.5 实例解析	121
本章小结	123
习题 6	123
第 7 章 数据查询	125
7.1 数据查询概述	125
7.2 SELECT 子句	127
7.2.1 SELECT 子句的语法	127
7.2.2 INTO 子句	128
7.2.3 SELECT 语句应用	128
7.3 FROM 子句	129
7.3.1 FROM 子句的语法	129
7.3.2 FROM 子句的应用说明	131
7.4 WHERE 子句	132

7.4.1 WHERE 子句的语法	132
7.4.2 数据的过滤	132
7.5 GROUP BY 和 HAVING 子句	134
7.5.1 GROUP BY 子句	134
7.5.2 多次分组	135
7.5.3 HAVING 子句	136
7.6 ORDER BY 子句	136
7.6.1 ORDER BY 子句的语法	136
7.6.2 ORDER BY 子句应用说明	137
7.7 实例解析	137
本章小结	144
习题 7	144

第 8 章 视图与索引 147

8.1 视图概述	147
8.1.1 视图的类型	147
8.1.2 视图的优点和缺点	148
8.2 创建视图	148
8.2.1 创建视图的基本原则	149
8.2.2 使用 SQL Server Management Studio 创建视图	149
8.2.3 使用 T-SQL 语句创建视图	151
8.3 管理视图	153
8.3.1 视图的修改与更名	153
8.3.2 视图的删除	155
8.3.3 视图数据的管理	156
8.4 索引概述	156
8.4.1 索引的类型	157
8.4.2 索引设计基础	158
8.5 索引的创建	159
8.5.1 使用 SQL Server Management Studio 创建索引	159
8.5.2 使用 T-SQL 语言创建索引	162
8.6 索引的管理	164
8.6.1 使用 SQL Server Management Studio 管理索引	164
8.6.2 使用 T-SQL 语句管理索引	165
本章小结	166
习题 8	166

第 9 章 存储过程与触发器 169

9.1 存储过程概述	169
9.2 存储过程的创建与管理	170

9.2.1 存储过程的设计	170
9.2.2 创建存储过程	172
9.2.3 存储过程的管理	177
9.3 触发器概述	178
9.4 触发器的创建与管理	180
9.4.1 DML 触发器计划	180
9.4.2 使用 T-SQL 语言创建触发器	182
9.4.3 使用 SQL Server Management Studio 创建触发器	183
9.4.4 触发器的管理	184
9.5 实例解析	189
本章小结	190
习题 9	190
第 10 章 游标的应用	193
10.1 游标概述	193
10.2 使用游标	194
10.2.1 声明游标	194
10.2.2 打开游标	196
10.2.3 检索游标返回的数据	196
10.2.4 关闭游标	197
10.2.5 释放游标	197
10.3 定位修改删除游标	199
本章小结	200
习题 10	200
第 11 章 数据库的安全管理	202
11.1 SQL Server 2005 的验证模式	202
11.2 登录账户管理	204
11.3 用户管理	211
11.3.1 数据库用户概述	211
11.3.2 数据库用户管理	211
11.4 角色管理	214
11.4.1 角色简介	214
11.4.2 角色的管理	215
11.5 权限管理	219
本章小结	222
习题 11	222
第 12 章 数据备份与导入/导出	225
12.1 数据库的备份	225

12.1.1 概述	225
12.1.2 备份设备	226
12.1.3 备份类型	228
12.1.4 备份操作	229
12.2 数据库的恢复	234
12.2.1 使用 SQL Server Management Studio 恢复数据库备份	235
12.2.2 使用 T-SQL 语句进行数据库的恢复	236
12.3 数据导入与导出	238
12.3.1 概述	239
12.3.2 使用 SQL Server Management Studio 导出与导入数据	241
12.3.3 SSIS 设计器简介	249
本章小结	251
习题 12	251
第 13 章 数据库系统开发	254
13.1 数据库系统开发简介	254
13.2 数据访问技术概述	255
13.3 ADO 数据访问技术	259
13.3.1 ADO 编程模型	259
13.3.2 ADO 数据库访问控件	260
13.4 Web 数据库访问技术简介	262
13.5 数据的复制与发布	263
13.5.1 基本概念	263
13.5.2 执行复制	266
13.5.3 发布的方法	268
本章小结	273
习题 13	273
第 14 章 实训	275
实训 1 数据库的创建与管理	275
实训 2 数据查询	276
实训 3 数据库对象的创建与管理	277
实训 4 数据库的安全管理	278
实训 5 数据库的备份/恢复与导入/导出	279
附录 I	281
附录 II	282
参考文献	284

第 1 章 数据库系统概述

本章要点

随着计算机技术与网络通信技术的发展,数据库技术已成为信息社会中对大量数据进行组织与管理的重要手段及软件技术,是网络信息化管理系统的基础。本章主要介绍数据库系统以及所涉及的一些基本概念及相关基础知识,包括:

- ◆ 数据库技术的产生和发展
- ◆ 数据模型
- ◆ 数据库的系统结构
- ◆ 数据库系统的组成、结构以及工作流程
- ◆ 关系数据库的基本概念

1.1 引言

随着人类社会的进步、科学技术的发展,呈现在人们面前的信息越来越多,通过对大量的信息进行识别、处理、传递、存储,建立高效的信息管理系统已经成为越来越多用户的迫切要求。为此,在 20 世纪 60 年代,数据库技术作为软件学科的一个分支应运而生了。

“数据库(Database)”这个名词起源于 20 世纪 50 年代初,当时美国为了战争的需要,把各种情报集中在一起,存储在计算机里,称为 Information Base 或 Database。

1963 年,美国 Honeywell 公司的 IDS(Integrated Data Store)系统投入运行,揭开了数据库技术的序幕。1965 年,该系统帮助设计了阿波罗登月火箭,推动了数据库技术的产生。当时美国社会上产生了许多形形色色的 Database 或 Databank,但是,它们基本上都是文件系统的扩充或是倒排文件系统。1968 年,美国 IBM 公司推出层次模型的 IMS 数据库系统(1969 年形成产品);1969 年,美国 CODASYL(数据库系统语言协会)组织的数据库任务组(DBTG)发表关于网状模型的 DBTG 报告(1971 年正式通过);1970 年,IBM 公司的 E. F. Codd 连续发表论文提出关系模型。这三件事奠定了现代数据库技术的基础。

20 世纪 70 年代是数据库蓬勃发展的年代,网状系统和层次系统占据了整个数据库商用市场,而关系系统仅处于实验阶段。许多计算机厂商研制了各种数据库管理系统,许多学者发表了大量研究性论文,使数据库技术在实践中和理论上得到飞速发展,数据库技术也日趋成熟。

20 世纪 80 年代,随着计算机硬件性能的改善,关系系统由于使用简便,逐步代替网状系统和层次系统占领了市场。同时关系数据库的理论日趋完善,走向更高级的阶段。数据库的理论以及新系统的研究继续进行,不断深化。此时关系数据库语言 SQL 成了国际标准语言。

20 世纪 90 年代,关系数据库已成为数据库技术的主流。由于受到计算机应用领域以及

其他分支学科的影响,数据库技术与面向对象技术、网络技术相互渗透,人们在对象数据库技术和网络数据库技术方面进行了深入的研究。

进入 21 世纪以后,无论是市场的需求还是技术条件的成熟,对象数据库技术、网络数据库技术的推广和普及已成定局。数据库技术使得计算机应用迅速渗透到工农业生产、商业、行政、科学研究、工程技术和国防军事各个部门,渗透到社会的每一角落,并改变着人们的工作方式和生活方式。管理信息系统(MIS)、办公信息系统(OIS)、计算机集成制造系统(CIMS)、地理信息系统(GIS)、因特网(Internet)技术等都使用了数据库技术的计算机应用系统。

1.2 数据库技术的发展

1.2.1 信息、数据及数据处理

信息是人类社会发展中维持生产、经济和社会活动必不可少的重要资源。表示信息的最准确方法是数据(Data)。如若需要描述一所学校的规模,最直观的方法就是列出学校的教职工人数、在校学生人数、占地面积、建筑面积等一系列具体的数据。

数据是对客观事物属性的一种描述,是信息的载体。数据的表现形式可以是数字、文字、图像、声音等,各种形式的数据经过数字化后存储在计算机中。数据是数据库中存储的基本对象。

数据处理是现代计算机应用中的一个重要组成部分。数据处理是指对各种形式的数据进行分类、组织、编码、存储、检索、更新和维护等一系列活动的总和。其目的是从大量的、杂乱无章的原始数据中推导出对人们有价值的信息,用来作为行动和决策的依据。

1.2.2 数据管理

数据处理的中心问题是数据管理。数据管理是随着计算机硬件和软件技术的发展而不断发展的。数据管理的发展经历了以下几个阶段:

1. 人工管理阶段

20 世纪 50 年代中期以前是应用计算机管理数据的初级阶段。该阶段的特点是计算机系统仅提供基本的输入/输出,尚无统一的数据管理软件。在此阶段,计算机的程序和数据合二为一,数据与程序不具有独立性,各个程序使用自己的数据。由于数据是面向程序的,程序间的数据不能共享,从而造成数据的冗余和难以管理。

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,随着计算机硬件性能的改进和软件的发展,人们开始应用文件系统来管理数据。程序和数据之间部分实现了相互的独立性,一些数据不再属于某个特定程序,可以重复使用。这个阶段尽管比人工管理阶段有了很大的改进,但一些根本性的问题依然存在,如数据冗余度大、易造成数据的不一致性、程序与数据相互依赖、应用程序设计困难等。

3. 数据库系统阶段

20 世纪 60 年代后期发生的对数据库技术有着奠基作用的三件大事,标志着以数据库系统为基本方式的数据管理阶段的开始。在这个阶段,计算机硬件技术和软件研究水平的快速提高使得数据处理这一领域取得了长足的进步。伴随着大容量、高速度、低价格的存储设

备的出现,用来存储和管理大量信息的“数据库管理系统”应运而生,成为当代数据管理的主要方法。

数据库系统的特点是采用数据模型表示复杂的数据结构、有较高的数据独立性、为用户提供了方便的接口和统一的数据控制功能(数据并发控制、数据的备份与恢复、数据安全性控制、数据完整性控制等)、增强了系统的灵活性。在数据库系统阶段对数据的操作既可以记录为单位,也可以数据项为单位。

4. 高级数据库技术阶段

20世纪70年代中期以来,随着计算机技术以及其他相关技术的不断发展,数据库技术已经和网络通信技术、面向对象技术、人工智能技术等交叉融合在一起,出现了分布式数据库系统、面向对象数据库系统、智能型知识数据库系统、多媒体数据库系统以及数据仓库等。

数据库技术成为当今计算机领域发展最快的技术之一,它的实用性很强,应用范围很广,其应用领域几乎覆盖社会生活的各个方面。建设一个以数据库管理系统为核心的信息管理系统对于提高管理效率、增强竞争力具有极其重要的意义。

(1) 分布式数据库系统

分布式数据库系统(Distributed Database Systems, DDBS)是分布在计算机网络不同结点上的数据集合。分布式数据库系统在物理上是分散的,在逻辑上是统一的。其主要特点是网络上各结点上的数据库都具有独立的处理能力,一般数据处理本地完成,不能处理的才交给其他处理机处理;计算机之间通过通信网络连接,每个结点上的计算机可以访问本地数据库,也可以访问其他结点上的数据库;系统中数据冗余相对增多;个别结点的失效不会引起系统的瘫痪;多结点并行工作,数据处理效率提高。

分布式数据库系统已经成为信息处理学科的重要领域,正在迅速发展之中。

(2) 面向对象数据库系统

面向对象数据库系统(Object Oriented Database System, OODBS)是由面向对象数据模型定义的对象集合,该集合中的对象反映了面向对象程序设计所支持的对象语义。面向对象数据库系统是一个面向对象数据库的永久、共享存储器和程序。

面向对象数据库系统的基本设计思想是:一方面把面向对象语言向数据库方向扩展,使应用程序能够存取并处理对象;另一方面扩展数据库系统,使其具有面向对象的特征,提供一种综合的语义数据建模概念集,以便对现实世界中复杂应用的实体和联系建模。因此,面向对象数据库系统首先是一个数据库系统,具备数据库系统的基本功能,其次是一个面向对象的系统,是针对面向对象的程序设计语言的永久性对象存储管理而设计的,充分支持完整的面向对象概念和机制。

(3) 智能型知识数据库系统

智能型知识数据库系统是一门新的学科,该系统将人工智能的知识获取技术与机器学习的理论引入到数据库系统中,通过抽取隐含在数据库实体间的逻辑蕴含关系和隐含在应用中的数据操纵之间的因果关系,形式化地描述数据库中的实体联系。在智能型知识数据库中可以把语义知识自动提供给推理机,从已有的事实知识推出新的事实知识。它主要研究知识的表示、结构、存储、获取等技术,是专家系统、知识处理系统的重要组成部分。

(4) 多媒体数据库

多媒体数据库系统(Multi-media Database System, MDBS)是数据库技术与多媒体技术相结合的产物。在许多数据库应用领域中,都涉及大量的多媒体数据。多媒体数据比传统

的数字、字符等格式化数据复杂,因此,多媒体数据库系统处理的数据量大,数据需要实时处理和传输,数据类型复杂。从实际应用的角度考虑,多媒体数据库管理系统(MDBMS)应具有如下基本功能:

①有效地表示多种媒体数据,对于不同媒体的数据,如文本、图形、图像、声音等能够按应用的不同,采用不同的表示方法。

②能够处理各种媒体数据,正确识别和表示各种媒体数据的特征、各种媒体间的空间或时间关联。

③能够像其他格式化数据一样对多媒体数据进行操作,包括对多媒体数据的浏览、查询、检索,对不同的媒体提供不同的操纵,如声音的合成、图像的缩放等。

④具有开放功能,提供多媒体数据库的应用程序接口等。

(5)数据仓库

随着信息技术的高速发展,数据和数据库在急剧增长,数据库应用的规模、范围和深度不断扩大,一般的事务处理已不能满足应用的需要,企业界需要基于大量信息数据的决策支持(Decision Support,DS),数据仓库(Data Warehouse,DW)技术的兴起满足了这一需求。数据仓库作为决策支持系统(Decision Support System,DSS)的有效解决方案,涉及三方面的技术内容:数据仓库技术、联机分析处理(On-Line Analysis Processing,OLAP)技术和数据挖掘(Data Mining,DM)技术。

1.3 数据模型

在数据库系统中存储和管理的数据,均是现实世界客观事物的属性描述。若要将现实世界客观事物的属性抽象为可由计算机存储和处理的数据,一般需要经过三个阶段的转化过程:现实世界→信息世界→机器世界。所有的数据库系统均是基于某种数据模型的,数据模型直接影响数据库的性能。因此,数据模型的选择是设计数据库的一项首要任务。

1.3.1 信息的三个世界

1. 现实世界

现实世界就是客观事物及由事物本身的性质决定的事物之间的联系。现实世界中的事物之间既有“共性”,又有“个性”。要解决现实问题,就要从中找出反映实际问题的对象,研究它们的性质及其内在规律,从而寻求解决的方法。现实世界是数据库系统设计者接触到的最原始的数据。

如学校的收费管理系统中涉及基本信息管理(院系信息管理、专业信息管理、学生信息管理)、学生收费管理(收费项目维护、学生收费管理)、查询统计管理(日收费统计管理、月收费统计管理)以及系统用户管理(用户权限管理、用户密码设置)等。

2. 信息世界

现实世界中的事物及其相互间的联系经过人们的感知、分析、归纳、抽象形成相应的信息。对这些信息的记录、整理、分类和格式化,就构成了信息世界。因此,信息世界也称为概念世界,是现实世界在人们头脑中的反映,是用数据对客观事物及其联系的一种抽象描述。其中,经过抽象描述的现实世界中的客观事物在信息世界中称为“实体”。

如基本信息管理中的专业信息表,可以用专业编号、专业名称、专业描述以及院系编号等一组数据来描述,通过这样一组数据就可以了解相关部门的基本情况。

信息世界仅涉及对某个特定组织所关心的特定信息结构的描述,而不涉及信息在系统中的表示。信息世界中的主要概念包括:

(1)实体(Entity):现实世界中客观存在并且可以相互区别的对象。如各院系、各专业、一名学生等。

(2)实体集(Entity Set):同类型实体的集合称为实体集。如全部院系、全部专业、全体学生等。

(3)属性(Attribute):对实体或实体间联系的描述称为属性。一个实体可以由若干个属性来描述,实体间的联系也有自己特有的属性。例如,学生的学号、姓名、性别、出生年份、院系、专业、入学时间等均为学生实体的属性;一名学生应交学费的金额即是“收费”联系的属性。

(4)域(Domain):属性的取值范围。每个属性有一个值域,其类型可以是整数型、实数型、字符串型或其他类型。例如,性别的域为(男,女)。

(5)联系(Relation):实体与实体之间的关系。通过联系,可以用一个实体的信息来查找另一个实体的信息。实体之间的联系经抽象后,可以分为三类,如图 1-1 所示。

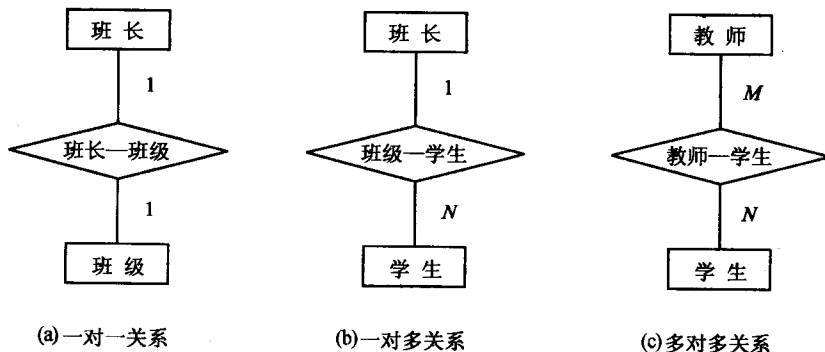


图 1-1 实体间的三种联系

①一对一关系(1:1)。对于实体集 A 中的每一个实体,实体集 B 中有并且至多有一个实体和它有联系,反之亦然。如一个班级有一个班长,每个班长对应一个班级,班长和班级之间的关系是一对一关系。

②一对多关系(1:N)。对于实体集 A 中的每一个实体,实体集 B 中有多于一个的实体和它有联系,反之,对于实体集 B 中的每一个实体,实体集 A 中有并且至多有一个实体和它有联系。如一个班级有多个学生,而每个学生只能属于一个班级,班级和学生之间的关系是一对多的关系。

③多对多关系(M:N)。对于实体集 A 中的每一个实体,实体集 B 中有多于一个的实体和它有联系,反之对于实体集 B 中的每一个实体,实体集 A 中也有多于一个的实体和它有联系。如每个教师为多位学生授课,而每位学生又有多位任课教师,教师和学生之间的关系就是多对多的关系。

3. 机器世界

机器世界是数据化了的信息世界,故也称为数据世界。数据世界中的数据既能反映信息世界,同时又使数据的计算机处理成为可能。机器世界中的主要概念如下:

(1)数据项(Item):标记实体属性的命名单位,也称为元素或字段。它是可以命名的最

小信息单位。数据项的取值范围称为域。域以外的任何值对该数据项都是无意义的。例如,表示月份的数据项的域是1~12,则13就是无意义的值。数据项的值可以是数值、字母、汉字等形式。数据项的物理特点是它具有确定的物理长度,一般用字节数表示。

(2)记录(Record):由若干相关联的数据项的集合构成。记录是应用程序输入/输出的逻辑单位。对于大多数据库系统来讲,记录是处理和存储信息的基本单位。通常用一条记录描述一个实体。构成该记录的数据项表示实体的若干属性。如实体“学生”的一组数据(200740201,黎明,男,计算机)就是一条记录。其数据项包括学号、姓名、性别及专业。

(3)文件(File):由同一给定类型记录的集合构成。文件对应于信息世界的实体集,因此,文件又可定义为描述一个实体集的所有记录集。如所有的学生记录组成一个学生文件。文件用文件名称标识。文件根据记录的组织方式和存取方法可以分为:顺序文件、索引文件、直接存取文件等。

(4)数据库(Database):比文件更大的数据组织形式。数据库是具有特定联系的数据的集合,也可以看成是具有特定联系的多种类型的记录的集合。

(5)关键字(Key):能够惟一标识文件中的每一条记录的字段或字段集,如学生的“学号”可以作为学生记录的关键字,企事业单位职工的“编号”可以作为职工记录的关键字。

从现实世界到信息世界不是简单的数据描述,而是从现实世界中抽象出适合数据库技术处理的数据即实体。同时要求这些实体能够很好地反映客观世界中的事物;从信息世界到数据世界也不是简单的数据对应存储,而是要设计数据的逻辑结构和物理存储结构。数据的逻辑结构,是指程序员或用户用以操作的数据形式,是抽象的概念化数据;数据的物理存储结构,是指数据实际存储在存储设备上的形式。

4. 三个世界之间的关系

现实世界、信息世界和数据世界三者之间的关系如图 1-2 所示。



图 1-2 三个世界的关系

从图 1-2 可知,将现实世界中的事物及其联系抽象、数据化为数据世界,需要使用两个不同层次的数据模型。

(1)概念数据模型:一种独立于计算机系统,完全不涉及信息在计算机中的表示,仅用于描述某个特定组织所关心的信息结构的数据模型,也称信息模型。概念模型是按用户的观点来对数据和信息进行建模的,是用户和数据库系统设计者之间进行交流的工具,其既独立于计算机系统,又与具体的数据库管理系统无关。这一类模型中最著名的是“实体联系模型”。

概念模型的特征包括:具有较强的语义表达能力,能够方便、直接地表达应用中的各种语义知识;简单、清晰、易于理解;独立于计算机系统的模型,完全不涉及信息在计算机中的表示;独立于任何 DBMS,可以方便地转换为计算机上任一数据库管理系统所支持的特定数