

可下载教学资料

<http://www.tup.tsinghua.edu.cn>



高等学校教材
信息管理与信息系统

网络数据库原理 与应用

刘 翔 编著

清华大学出版社



高等学校教材
信息管理与信息系统

网络数据库原理与应用

刘翔 编著

清华大学出版社
北京

内 容 简 介

本书系统地介绍了网络数据库的概念和应用。通过本书的学习,可使学生掌握互联网数据库技术的概念和运作机制,能在工作和生活的各个方面灵活地应用互联网数据库技术,并具有设计互联网数据库的能力。本书内容包括网络数据库应用系统的基本概念、当前流行的关系数据库管理系统 SQL Server 网络数据库功能特征、SQL Server 数据库基本数据对象管理、SQL Server 存储过程与触发器、SQL Server 网络数据库安全管理、.NET 平台及其网络数据库系统开发工具环境、ASP.NET 网络数据库应用程序设计、信息管理网络数据库应用实例、Web 服务网络数据库应用系统分析与设计,以及网络数据库数据仓库等。

本书适用于高等学校经济类、管理类专业作为教材,也可作为网络数据库应用系统开发人员的技术参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

网络数据库原理与应用/刘翔编著. —北京:清华大学出版社,2008.10

(高等学校教材·信息管理与信息系统)

ISBN 978-7-302-18258-0

I. 网… II. 刘… III. 关系数据库—数据库管理系统—高等学校—教材

IV. TP311.138

中国版本图书馆 CIP 数据核字(2008)第 113655 号

责任编辑:索 梅

责任校对:时翠兰

责任印制:孟凡玉

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185×260 印 张:13.5 字 数:336 千字

版 次:2008 年 10 月第 1 版 印 次:2008 年 10 月第 1 次印刷

印 数:1~4000

定 价:21.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:026649-01

改改革开放以来,特别是党的十五大以来,我国教育事业取得了举世瞩目的辉煌成就,高等教育实现了历史性的跨越,已由精英教育阶段进入国际公认的大众化教育阶段。在质量不断提高的基础上,高等教育规模取得如此快速的发展,创造了世界教育发展史上的奇迹。当前,教育工作既面临着千载难逢的良好机遇,同时也面临着前所未有的严峻挑战。社会不断增长的高等教育需求同教育供给特别是优质教育供给不足的矛盾,是现阶段教育发展面临的基本矛盾。

教育部一直十分重视高等教育质量工作。2001年8月,教育部下发了《关于加强高等学校本科教学工作,提高教学质量的若干意见》,提出了十二条加强本科教学工作提高教学质量的措施和意见。2003年6月和2004年2月,教育部分别下发了《关于启动高等学校教学质量与教学改革工程精品课程建设工作的通知》和《教育部实施精品课程建设提高高校教学质量和人才培养质量》文件,指出“高等学校教学质量和教学改革工程”是教育部正在制定的《2003—2007年教育振兴行动计划》的重要组成部分,精品课程建设是“质量工程”的重要内容之一。教育部计划用五年时间(2003—2007年)建设1500门国家级精品课程,利用现代化的教育信息技术手段将精品课程的相关内容上网并免费开放,以实现优质教学资源共享,提高高等学校教学质量和人才培养质量。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上;精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展、顺应并符合新世纪教学发展的规律、代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻

性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。首批推出的特色精品教材包括:

(1) 高等学校教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 高等学校教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 高等学校教材·电子信息——高等学校电子信息相关专业的教材。

(4) 高等学校教材·软件工程——高等学校软件工程相关专业的教材。

(5) 高等学校教材·信息管理与信息系统。

(6) 高等学校教材·财经管理与计算机应用。

清华大学出版社经过 20 多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

E-mail: dingl@tup.tsinghua.edu.cn

规范化的数据是管理的重要资源,正如人们常说的“三分技术、七分管理、十二分数据”,数据管理是现代管理的基石。互联网时代,组织所获得的信息正爆炸性地增长,数据成为一个组织重要资产的同时,面对海量的管理业务数据,管理者如何管理好数据、充分利用数据资源是现代管理者面临的重要问题。为了管理海量而复杂的各种业务数据,管理者必须借助数据管理工具对数据进行生产与管理,以简化数据管理和数据抽取工作,否则,管理数据将变成巨大负担,以至于获取和管理数据的代价将远远超过从数据发掘出的价值。

数据库技术是信息技术最主要的核心。大容量的数据存储、快速的数据查询、方便的向导和工具、友好与亲切的用户界面的数据库技术在管理业务中已经广泛应用,如企业电子商务、企业资源规划、客户关系管理等都是以数据库为基础的。数据库是管理规范化数据的核心工具。从现代管理中的数据管理角度出发,了解与学习数据库原理与应用对当代经济类、管理类专业大学生是十分必要的。通过本书的学习,可使学生基本掌握数据库的原理,了解如何将经济和管理中的业务数据转化为数据库的数据范式,并利用数据库管理系统对业务数据进行有效的管理,为经济类、管理类大学生今后从事管理工作打下基础。

在数据库的发展中,最重要的是数据库已经进入了互联网时代。使用数据库系统存储数据并通过 Web 浏览器浏览数据,已是目前实现管理的主流。通过 Web 可存取的表单界面来产生查询请求,并使用诸如 HTML 的标记语言将查询结果格式化,从而便于在浏览器中显示。所有数据库提供商都在增加相关数据库系统的互联网功能,使之更适于在互联网上部署。随着越来越多在线数据的产生,并且这些数据通过计算机网络越来越容易获得,数据库变得更加重要了。目前,管理与经济等众多领域的管理系统(例如企业资源规划系统、客户关系管理系统、财务系统、经济预测系统、银行管理与交易系统、股票交易系统等)都在广泛地使用基于互联网环境的数据库。基于互联网环境的数据库管理系统正代表着最大和最具活力的数据管理的商业市场,基于互联网环境的数据库管理系统的原理与应用是目前数据库原理与应用的代表和主流。本书以基于互联网环境的数据库管理系统为主线,介绍数据库原理与应用。对于经济类、管理类本科生而言,本书试图用最简单的方法,直观、系统、快速地学习最新和实用的数据库原理与应用内容,以提高学习效果。

本书以 SQL Server 2005 为平台,从管理者了解与应用数据库的角度,以非常简明的方式,介绍数据库的基本原理与应用方法。全书基本主线可描述如下:首先,对数据库基本概念、系统结构进行较全面的介绍;其次,介绍数据库系统工具、数据对象的管理方法、安全技术等;再次,介绍网络数据库前端开发平台 .NET 与开发语言 C#.NET;最后,介绍网络数据库应用系统的分析、设计与实现,包括 ASP.NET 网络数据库、Web 服务网络数据库与网络商务智能及数据仓库。本书实例丰富,应用性强,主要特色是以实例为导向,简明扼要地介绍数据库的数据管理原理与应用,系统地介绍网络数据库应用系统面向服务的分析与设计方法,较全面地介绍网络数据仓库的规划、分析、设计与实现方法。

本书中的源程序与课件可从清华大学出版社网站(<http://www.tup.com.cn>)下载。

本书适合于经济类、管理类专业本科生作为数据库教材。

本书是由作者使用多年的教学讲义编写的,由于作者水平有限,难免有疏漏之处,诚请读者批评指正。

编 者
于上海大学
2008 年 8 月

第 1 章 数据库基础	1
1.1 数据库概述	1
1.2 关系数据库的主要概念	3
1.2.1 数据描述、存储、查询与数据库设计过程	3
1.2.2 ER 模型	4
1.2.3 关系模型	5
1.2.4 数据规范化	6
1.2.5 主流数据库产品简介	8
1.3 网络数据库系统概念、特征及其组成	9
1.3.1 网络数据库系统概念	9
1.3.2 网络数据库系统的特点	10
1.4 网络数据库系统体系结构	11
1.4.1 主机-终端体系结构	12
1.4.2 分层体系结构	12
1.4.3 基于 Web 服务的体系结构	12
1.5 思考题	13
第 2 章 SQL Server 网络数据库系统工具	14
2.1 SQL Server 版本简介	14
2.2 SQL Server 网络数据库功能特征	15
2.3 部署 SQL Server 网络数据库功能	17
2.4 Transact-SQL 语言基础	23
2.4.1 Transact-SQL 组成部分	23
2.4.2 SQL Server Transact-SQL 新语法	25
2.5 思考题	26
第 3 章 SQL Server 数据库基本对象操作和管理	27
3.1 数据库	27

3.2 表	35
3.2.1 表设计与创建	35
3.2.2 向表中插入数据	55
3.2.3 使用 UPDATE 修改表中数据	57
3.2.4 使用 DELETE 删除数据	58
3.2.5 使用 Transact-SQL 修改表结构	59
3.3 索引	60
3.3.1 可视化创建索引	60
3.3.2 使用 Transact-SQL 创建索引	62
3.4 SQL Server 数据查询与表现	64
3.4.1 检索表中列技术	64
3.4.2 操纵数据技术	64
3.4.3 选择行技术	65
3.5 思考题	66
第 4 章 SQL Server 存储过程与触发器	68
4.1 SQL Server 存储过程	68
4.1.1 存储过程概述	68
4.1.2 存储过程的创建与执行	69
4.1.3 修改存储过程	71
4.1.4 用 C# 创建存储过程	71
4.2 SQL Server 触发器	73
4.2.1 触发器的作用	73
4.2.2 触发器的分类	73
4.2.3 创建与执行触发器	74
4.2.4 查看触发器基本信息	76
4.2.5 查看触发器代码	76
4.2.6 修改触发器	77
4.2.7 删除触发器	78
4.3 思考题	78
第 5 章 SQL Server 数据库系统安全管理	79
5.1 网络传输安全性	79
5.2 SQL Server 服务器安全性	79
5.2.1 Windows 身份验证	80
5.2.2 混合身份认证	80
5.2.3 设置 SQL Server 的安全认证模式	80
5.3 SQL Server 数据库安全性	81
5.4 SQL Server 数据对象安全性	83

5.5 思考题	83
第 6 章 .NET 平台及其网络数据库系统开发环境	84
6.1 Microsoft .NET	84
6.1.1 .NET 平台概述	84
6.1.2 .NET Framework	85
6.2 C# 语言	86
6.2.1 C# 语言概述	87
6.2.2 C# 语言的优点	88
6.2.3 C# 语言和其他语言的比较	89
6.2.4 理解命名空间	90
6.2.5 编写和编译 C# 程序	91
6.3 .NET 提供的数据库开发环境	92
6.3.1 .NET 常用数据访问技术	92
6.3.2 选择合适的 .NET 数据访问技术	93
6.3.3 .NET 中用于数据处理的命名空间	94
6.3.4 .NET 中可视化数据库工具	94
6.4 SQL Server 与 XML	96
6.4.1 XML 定义	96
6.4.2 SQL Server XML 数据类型	98
6.5 思考题	99
第 7 章 ASP.NET 网络数据库应用程序设计	100
7.1 ASP.NET 概述	100
7.2 ASP.NET 应用程序	102
7.2.1 ASP.NET 应用程序建立	102
7.2.2 服务器控件	106
7.2.3 ASP.NET 页面值	107
7.3 数据库数据源控件与数据显示控件	110
7.3.1 数据源控件	110
7.3.2 数据显示控件	111
7.4 ASP.NET 网络数据库应用系统	111
7.4.1 数据库设计	112
7.4.2 数据库业务逻辑处理程序设计	113
7.4.3 数据库 Web 页面设计	118
7.4.4 在 ASP.NET 中执行数据库操作命令	125
7.4.5 在 ASP.NET 中配置 Web. config 数据库接口	127
7.5 思考题	128

第 8 章 信息管理网络数据库应用系统实例	129
8.1 开发环境与系统配置	129
8.2 功能描述	132
8.3 系统设计与实现	133
第 9 章 Web 服务网络数据库应用系统分析与设计	146
9.1 Web 服务概念	146
9.2 Web 服务技术	149
9.3 Web 服务创建与调用	151
9.4 Web 服务网络数据库系统分析	157
9.4.1 Web 服务网络数据库系统分析方法	157
9.4.2 Web 服务网络数据库系统分析原则	161
9.4.3 Web 服务网络数据库系统分析内容	162
9.5 Web 服务网络数据库系统设计	164
9.5.1 Web 服务网络数据库系统体系结构	164
9.5.2 Web 服务网络数据库系统设计原则	165
9.5.3 Web 服务网络数据库系统设计内容	166
9.6 Web 服务网络数据库系统的服务管理与实现	169
9.7 Web 服务网络数据库实例	170
9.7.1 Web 服务 ERP 系统	170
9.7.2 Web 服务虚拟组织	171
9.8 思考题	174
第 10 章 网络数据库数据仓库	175
10.1 数据仓库概念	175
10.2 SQL Server 数据仓库	177
10.2.1 概览	177
10.2.2 分析服务	178
10.2.3 报表服务	180
10.3 数据仓库开发过程	182
10.3.1 规划及可行性报告	182
10.3.2 系统需求	185
10.3.3 系统分析	189
10.3.4 系统设计	193
10.3.5 系统实现	200
10.3.6 试用	200
10.3.7 扩充	201
10.4 思考题	202
参考文献	203

数据库基础

1.1 数据库概述

数据库广泛应用于管理业务中的数据管理,例如仓库管理、档案管理、财务管理、生产管理等数据处理问题。这类数据的特点是数据量比较大、数据检索频繁,数据处理的主要内容是数据的存储、查询、修改、分类排序等,以及支持决策功能。

数据库是指相关数据的集合,这些数据按照一定的数据模型在计算机系统中组织存储和使用。数据库管理系统(DataBase Management System,DBMS)是辅助用户管理和利用大数据集的软件工具。DBMS具有以下特征。

(1) 数据独立性。应用程序可独立于数据表达和存储细节。DBMS提供数据的抽象视图,从而把应用代码与数据细节分开。

(2) 持久性。数据库中数据可长久地保持,如银行的业务数据、单位的人事数据等,这些数据需要保存几年、几十年甚至更长的时间。

(3) 可靠性。系统发生软硬件故障时,DBMS可以有办法迅速可靠地把系统恢复到发生故障前的状态,例如可以对数据进行经常性的备份,以便当系统遭到破坏、数据丢失后可以把数据恢复到系统破坏前的状态。

(4) 共享性。指许多用户能够按照一定、有序的方式存取数据库中的数据,避免同步存取时可能会造成的并发性错误。例如在铁路部门或航空部门的售票系统中,经常会出现多个售票点对同一数据源中数据的操作。

(5) 有效的数据存取。DBMS采用各种复杂技术有效地存储和检索数据。这个特征非常重要,当数据存储在外部存储设备上时尤其如此。

(6) 数据完整性和安全性。如果数据总是通过DBMS存取,则DBMS能增强数据完整性约束。例如,在插入一个雇员的工资信息之前,DBMS可以检查是否超过部门预算。另外,DBMS也可以通过实行存取控制,确保数据对不同用户的可见性是不同的。

(7) 并发存储。DBMS并发调度数据的存取,所有用户在存取数据时感觉就像只有一个用户在操作一样。

数据库技术产生于20世纪60年代中期,根据数据模型的发展,可以划分为3个阶段:第一代网状、层次数据库系统;第二代关系数据库系统;第三代面向对象数据库系统。

1. 第一代数据库

第一代数据库的代表有：

(1) 1969 年 IBM 公司研制的层次模型的数据库管理系统 IMS。

(2) 20 世纪 70 年代美国数据库系统语言研究会(CODASYL)下属数据库任务组(DBTG)提议的网状模型。层次数据库的数据模型是有根的定向有序树,网状模型对应的是有向图。这两种数据库具有如下共同点。

① 支持三级模式(外模式、模式、内模式)。保证数据库系统具有数据与程序的物理独立性和一定的逻辑独立性。

② 用存取路径来表示数据之间的联系。

③ 有独立的数据定义语言。

④ 导航式的数据操纵语言。

2. 第二代关系数据库

其主要特征是支持关系数据模型。关系模型具有以下特点。

(1) 实体和实体之间的联系用关系来表示。

(2) 以关系代数为基础。

(3) 关系数据库语言是非过程化的。

3. 第三代面向对象数据库

第三代面向对象的数据库产生于 20 世纪 80 年代,主要有以下特征。

(1) 支持数据管理、对象管理和知识管理。

(2) 保持和继承了第二代数据库系统的技术。

(3) 对其他系统开放,支持数据库语言标准,支持标准网络协议,有良好的可移植性、可连接性、可扩展性和互操作性等。第三代数据库支持分布处理技术、并行计算技术。

网络数据库技术是目前数据库技术代表性内容。它允许用户开发的应用程序把多个物理分开的、通过网络互联的数据库当作一个完整的数据库看待。例如,Oracle 10g 数据库是第一个为网格计算设计的数据库。它利用了基于网格的硬件革新,并使用户更容易地在这些标准化、模块化的硬件组件上安装和配置 Oracle 数据库。作为网格资源(服务器和存储器)的消费者以及数据的提供者,Oracle 10g 数据库利用了网格计算的 3 个基本属性,即每一层计算栈之间的虚拟化、基于企业策略和动态需求的工作和资源供应、集中资源,以提高利用率,使 Oracle 数据库使用硬件组件(存储器和服务器)的方式虚拟化。它自动地将集群化的存储器和服务器供应给运行在网格中的不同数据库。作为数据的提供者,Oracle 10g 数据库提供了各种技术,Oracle 管理员可以利用这些技术来为网格用户和应用程序集中、虚拟化和供应数据。此外,运行在网格上的企业在安全性、高可用性、自依赖性和可管理性方面还具有严格的操作要求。

1.2 关系数据库的主要概念

1.2.1 数据描述、存储、查询与数据库设计过程

如图 1.1 所示, DBMS 中的数据描述的三级抽象为概念模式、物理模式和外模式。

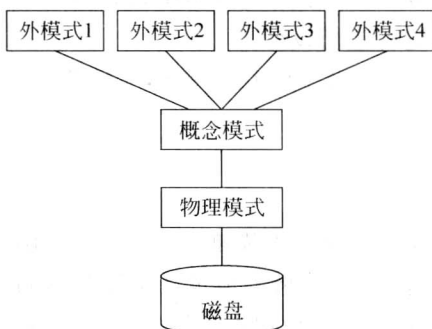


图 1.1 DBMS 中抽象的级别

概念模式(也称为逻辑模式)以 DBMS 数据模型的形式描述存储的数据。在关系型 DBMS 中,概念模式描述存储在数据库中的所有关系,这些关系包括实体信息(如企业管理者)和联系信息(如企业管理者的登记信息)。例如,假设所有企业管理者实体都是用 Managers 关系中的记录来描述,每一个实体集合和联系集合都描述为一个关系,于是产生下面的概念模式:

```
Managers(mid: string, name: string, age: integer)
Business(mid: string, bid: string, content: string)
```

物理模式描述概念模式中关系在磁盘和磁带等二级存储设备上实际是如何存储的。物理模式描述存储细节。通常要依据数据存取特点选择物理模式。例如,一个企业数据库的物理模式如下:

- (1) 使用未排序的记录文件存储所有关系。
- (2) 在 Managers 关系的第 1 列上创建索引。

外模式(也称为 DBMS 的数据模型)允许为单个用户和用户组定制数据存取。任何给定的数据库只有一组存储关系,只有一个概念模式和一个物理模式,但它可以有多个外模式,每一个外模式由一个或多个视图和来自概念模式的关系组成。

例如,如果允许员工查找管理者的名字和业务范围,则可通过定义下面的视图来实现:

```
Managersfo(mid: string, name: string, content: string)
```

针对 DBMS 中存储的数据的提问称为查询,用户利用 SQL 语言可以创建、修改和查询数据库中数据。

为了应用数据库管理业务数据,合理的数据库基本设计过程是很重要的,其基本过程如下:

- (1) 需求分析。了解用户业务需求,调查清楚用户面临的数据管理问题。
- (2) 概念数据库设计。使用高级数据模型来对需求分析信息进行描述。
- (3) 逻辑数据库设计。选择一个 DBMS 来实现数据库设计,并且将概念数据库设计转换为以所选 DBMS 的数据模型表示的数据库模式。
- (4) 模式的细化。分析关系数据库模式中的关系集合,找出潜在的问题并且改进它。用规范化关系理论对这些关系进行重组以确保所需要的特征。
- (5) 物理数据库设计包括在一些表上建立索引,以及对一些表进行聚簇,或者也可能对前面所得的部分数据库模式进行重新设计。
- (6) 应用与安全设计。任何涉及 DBMS 的软件项目都必然包含数据库之外的内容。

1.2.2 ER 模型

语义数据模型是一种抽象的高级数据模型,一个广泛使用的语义数据模型是实体-联系(Entity Relationship,ER)模型,它允许形象地定义实体及它们之间的联系。DBMS 不可能直接支持 ER 模型,但可以将它转变为数据库支持的数据模型。

实体是客观世界的一个对象。例如,汽车、汽车公司、汽车公司生产部、财务部、销售部等。相似实体的集合称为实体集。实体集之间可以是相交的;所有汽车公司销售部雇员的集合与所有汽车公司生产部雇员的集合可能都包含李明(他恰巧在两个部门工作)。可以定义实体集 Managers,它包含了汽车公司的所有管理人员。

属性指实体具有的某种特性。属性用来描述一个实体。对于属性的选择反映了希望描述实体信息的细节。例如,Managers 实体可以用姓名(name)、工作内容作为属性,就可以在数据库中为每个管理人员存储其姓名、工作内容等信息,但没有存储管理人员的住址、年龄等信息。对于每个与实体集相关的属性,须确定该属性的值域。例如,属性 name 的值域可以为一个 30 个字符的字符串集合。对于每个实体集,都选取一个码,码的值能够唯一地确定实体集中的每个实体,有可能存在多个候选码,这时可在其中指定一个作为主码。

图 1.2 表示实体集 Managers 有 3 个属性 mid、name 及 age。其中,实体用矩形表示,属性用椭圆表示,主码用下划线标出。

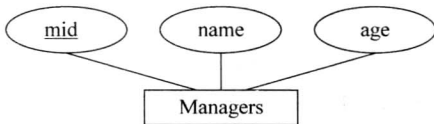


图 1.2 实体集 Managers

联系就是两个或多个实体之间的一种关联。与实体一样,希望把相似的联系放到一起,组成一个联系集。在信息世界中,实体之间的联系可分为两类:实体内部的联系,如组成实体的各属性之间的关系;实体之间的联系。设 A、B 为两个包含若干个体的总体,其间建立了某种联系,其联系方式可分为以下 3 类。

- 一对一联系 如果对于 A 中的一个实体,B 中至多有一个实体与其发生联系,反之,B 中的每一实体至多对应 A 中一个实体,则称 A 与 B 是一对一联系。

- 一对多联系 如果对于 A 中的每一实体, 实体 B 中有一个以上实体与之发生联系, 反之 B 中的每一实体至多只能对应于 A 中的一个实体, 则称 A 与 B 是一对多联系。
- 多对多联系 如果 A 中至少有一实体对应于 B 中一个以上实体, 反之, B 中也至少有一个实体对应于 A 中一个以上实体, 则称 A 与 B 为多对多联系。

联系集可以具有描述性属性。这些属性用来记录有关联系自身的一些信息, 而不是记录包含在该联系集中实体集的信息。

图 1.3 描述了汽车公司管理人员与部门的信息模型。联系集 Works_In 用菱形表示, 其中每个联系都表明管理人员在哪个部门工作。

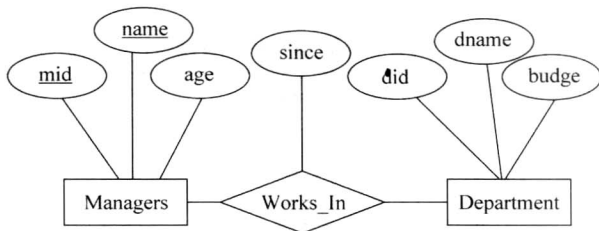


图 1.3 汽车公司管理人员与部门的信息模型

在 Works-In 中加了一个 since 属性用来记录管理人员何时进入某部门工作。在 Works_In 联系集中, 每个 Works_In 联系都能用管理人员的 mid 和部门的 did 来唯一标识。

ER 模型的不精确性可能会导致数据冗余, 从而引起很多问题。将在后续讨论数据冗余及相关问题, 并提出一种称为规范化的方法来从数据库表中消除冗余。

1.2.3 关系模型

1970 年, IBM 公司 San Jose 研究所的 E. F. Code 发表了题为“大型共享数据库的数据关系模型”的论文, 开创了数据库的关系方法和关系规范化的理论研究。关系模型是建立在数学概念的基础上, 应用关系代数和关系演算等数学理论处理数据库系统的方法。关系模型中用于描述数据的主要结构是关系。关系模型是数据库领域的一次革命。关系模型简单、精巧。数据库由一个或多个关系共同组成, 每个关系是行和列组成的表。

一个关系由关系模式和实例共同组成。实例就是一个表中记录行, 而关系模式则对表的每个列进行描述。关系模式需要给出关系名, 每个字段(称之为列或属性)的名称, 以及每个字段的域。关系模式中的域由一个域名表示, 它具有与之相关的具体值。在关系模型下, 数据的逻辑结构是一张二维表, 其中, 表名为关系名, 表头中各列名称为字段名。关系模型中的主要术语如下。

- (1) 关系。一个关系对应于一张二维表。
- (2) 记录(元组)。表中一行称为一个记录。
- (3) 字段。表中一列称为一个字段, 给每列起一个名即为字段名。
- (4) 主码(主关键字)。表中的某个属性组, 它的值唯一地标识一个元组。
- (5) 域。字段的取值范围。

- (6) 分量。记录中的一个字段值。
- (7) 关系模式。对关系的描述,用关系名(属性 1,属性 2,...,属性 n)来表示。

例 1.1 汽车公司数据库中管理人员信息的关系模式。

```
Managers(mid: string, name: string, age: integer)
```

一个关系的实例可以看成是记录的一个集合,每条记录具有的字段数目与在关系模式中定义的相符。一个关系的元组可以看成是一个表,其中每条记录就是表中的一行,所有行具有相同数目的字段。

例 1.2 关系 Managers 的实例,如表 1.1 所示。表中有 3 条记录,并与关系模式一样具有 3 个字段。任何 2 行数据都不完全相同。

表 1.1 关系 Managers 的实例

mid	name	age
1001	王新	29
1008	李高	38
1012	张明	30

对于实体和实体之间的联系均以关系来表示。

例 1.3 实体 Managers 和 Department 之间的联系。

```
Managers(mid: string, name: string, age: integer)
Works_In(mid, did, since)
Department(did: string, dname: string, budge: numeric)
```

其中,Works_In 关系引用了 Managers 关系中的 mid 字段,Department 关系引用了 Works_In 关系的 did 字段。

对于关系之间的联系可通过包含的属性表示,如例 1.3 中 Works_In 的“mid, did”。

1.2.4 数据规范化

数据库中数据的冗余是指在数据库中存储了同一信息的多个副本。它会引起以下一些问题:

- 冗余存储 信息被重复存储。
- 更新异常 当重复信息的一个副本被修改,所有的副本都必须进行同样的修改,否则就会造成不一致性。
- 插入异常 只有当一些信息事先已经存储在数据库中时,另外一些信息才能存入到数据库中。
- 删除异常 在删除某些信息时可能会丢失其他的信息。

关系型数据库的规范化理论是数据库设计的理论基础,其目的是研究关系模式中各个属性之间的依赖关系及其对关系模式的影响。规范化理论不仅能够作为数据库设计优劣的判断依据,而且还可以预测数据库系统可能出现的问题。给定一个关系模式,需要确定它是否是一个好的设计,或者是否需要将其分解为一些小的关系。要做出这些决定,就必须全面