

数据库原理及 实例教程

shujuku yuanli ji shili jiaocheng

computer

主编 马 乐 主审 李仲麟



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

内 容 简 介

本书全面、系统地介绍了数据库的基础理论知识,包括数据模型、数据库系统结构、关系数据库、结构化查询语言 SQL、关系数据库设计理论和数据库保护。全书每章后均附有习题,最后还附录了 2 个实例。

本书体系完整,内容翔实,实例丰富,图文并茂,适合作为高等院校本科生数据库课程的教学用书,也可供数据库领域的技术人员及数据库技术爱好者参考。

图书在版编目 (CIP) 数据

数据库原理及实例教程 / 马乐主编. —广州: 华南理工大学出版社, 2012. 1
ISBN 978 - 7 - 5623 - 3571 - 9

I. ①数… II. ①马… III. ①数据库系统 - 高等学校 - 教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2011) 第 273479 号

总 发 行: 华南理工大学出版社 (广州五山华南理工大学 17 号楼, 邮编 510640)

营销部电话: 020 - 87113487 87110964 87111048 (传真)

E-mail: scutcl3@scut.edu.cn <http://www.scutpress.com.cn>

责任编辑: 胡 元

印 刷 者: 佛山市浩文彩色印刷有限公司

开 本: 787mm × 1092mm 1/16 印张: 13.25 字数: 340 千

版 次: 2012 年 1 月第 1 版 2012 年 1 月第 1 次印刷

印 数: 1 ~ 1000 册

定 价: 28.00 元

版权所有 盗版必究

本书编委会

主 编 马 乐

副主编 陈 放 姜思羽 郝阳阳 杨 艺

参 编 袁俐萍 翁智生 谭兆麟 李 妍 王方丽
张伟娜 余小华 郑馥丹 秦映波 王素丽

前 言

数据库技术是计算机科学中一个重要的组成部分,是计算机软件领域的一个重要分支。数据库技术诞生于20世纪60年代,经过50年的积累,数据库技术已经成为计算机科学技术发展速度最快、应用领域最广的技术之一。

本书主要面向计算机专业学生编写,全书分7章。

第1章为绪论,概括地介绍了数据库的一些基本概念、数据模型、数据库系统结构。

第2章介绍了关系数据库,包括关系数据库结构、关系的完整性及关系代数。关系数据库是目前各类数据库中最重要、最流行的数据库,全面地掌握好关系数据库的基本概念、理论、方法,能为进一步学好各种数据库实例打下坚实的基础。

第3章介绍了关系数据库标准语言SQL,包括SQL的基本语法、操作和应用。SQL是一个功能强大的数据库语言,其主要功能就是同各种数据库建立联系,进行沟通。SQL语言通常用于完成一些数据库的操作任务,比如在数据库中更新数据,或者从数据库中检索数据,为用户灵活地完成数据库任务提供了很大的便利。

第4章介绍了关系数据库设计理论。关系数据库的设计归根结底是如何构造关系。本章从有待改进的数据库模式实例出发,阐明关系模式规范化理论研究的实际背景,然后介绍规范化理论的有关概念和方法,包括可能存在的异常问题和解决方法、函数依赖定义及其推理规则、各种范式及其相互关系、关系模式的分解特性等内容。

第5章介绍了数据库的保护,包括数据库的安全性、完整性、并发控制及数据库恢复。安全性阐述了如何防止因用户非法使用数据库而造成数据泄露、更改或破坏;完整性阐述了如何控制数据库以保证数据的正确;并发控制阐述了如何充分地共享和利用数据库资源;数据库恢复阐述了如何从各种故障中将数据库从错误状态恢复到某种逻辑一致的状态。本章还以SQL Server 2005为实例,进一步介绍了以上内容的实际应用。

第6章介绍了数据库设计,包括需求分析、概念结构设计、逻辑结构设计及数据库物理设计。数据库设计是建立数据库及其应用系统的核心和基础,本章按照规范化的设计方法,将数据库设计分为以上四个阶段进行阐述。

第7章介绍了数据库技术新进展,包括数据模型及数据库系统的发展历程、与其他计算机技术结合后诞生的新一代数据库技术、在特定应用中出现的新型数据库等。

本书最后还附录了网上书店数据库设计及在线考试系统数据库设计的案例,以对全书的重要知识做一个总体的应用示例,也可作为贯穿全书的教学案例供教师使用。

由于编者水平有限,书中难免会存在一些不足之处,殷切希望广大读者朋友能不吝赐教,批评指正。联系方式:hitmy@163.com,谢谢。

编 者
2011年9月

目 录

第1章 绪论	(1)
1.1 引言	(1)
1.1.1 基本概念	(1)
1.1.2 数据库技术的产生与发展	(3)
1.1.3 数据库技术的研究领域	(5)
1.2 数据模型	(5)
1.2.1 数据模型三要素	(6)
1.2.2 概念模型	(6)
1.2.3 数据模型	(9)
1.3 数据库系统结构	(11)
1.3.1 数据库系统模式结构	(12)
1.3.2 数据库系统体系结构	(13)
习题1	(16)
第2章 关系数据库	(17)
2.1 概述	(17)
2.2 关系数据结构	(18)
2.2.1 关系	(18)
2.2.2 关系模式	(20)
2.2.3 关系数据库	(21)
2.3 关系的完整性	(21)
2.4 关系代数	(22)
2.4.1 传统的集合运算	(23)
2.4.2 专门的关系运算	(25)
习题2	(28)
第3章 关系数据库标准语言 SQL	(29)
3.1 SQL 概述	(29)
3.1.1 SQL 的特点	(29)
3.1.2 基本概念	(31)
3.2 数据定义	(31)

3.2.1 基本数据类型	(32)
3.2.2 定义、删除与修改基本表	(32)
3.2.3 创建与删除索引	(34)
3.3 查询	(35)
3.3.1 单表查询	(37)
3.3.2 连接查询	(48)
3.3.3 嵌套查询	(52)
3.3.4 集合查询	(58)
3.4 数据更新	(60)
3.4.1 插入数据	(60)
3.4.2 修改数据	(61)
3.4.3 删除数据	(62)
3.5 视图	(63)
3.5.1 定义视图	(64)
3.5.2 查询视图	(67)
3.5.3 更新视图	(68)
3.5.4 视图的优点	(70)
3.6 数据控制	(71)
3.6.1 权限与角色	(71)
3.6.2 授权与收回权限	(72)
3.7 嵌入式 SQL	(75)
3.7.1 嵌入式 SQL 的说明部分	(75)
3.7.2 嵌入式 SQL 的可执行语句	(76)
3.7.3 动态 SQL 简介	(78)
3.8 存储过程	(79)
习题 3	(81)

第4章 关系数据库设计理论

4.1 关系模型设计中的一些问题	(83)
4.2 数据依赖	(85)
4.3 范式	(87)
4.3.1 第一范式 (1NF)	(87)
4.3.2 第二范式 (2NF)	(88)
4.3.3 第三范式 (3NF)	(90)
4.3.4 BC 范式 (BCNF)	(91)
4.3.5 多值依赖与第四范式 (4NF)	(93)
4.4 关系模式的规范化	(96)
习题 4	(97)

第 5 章 数据库保护	(99)
5.1 安全性	(99)
5.1.1 一般安全措施	(99)
5.1.2 SQL Server 的安全性	(101)
5.2 完整性	(110)
5.2.1 完整性规则	(111)
5.2.2 SQL Server 的完整性	(113)
5.3 并发控制	(117)
5.3.1 概述	(117)
5.3.2 并发操作的可串行化	(120)
5.3.3 封锁	(121)
5.3.4 死锁与活锁	(126)
5.3.5 SQL Server 的并发控制	(128)
5.4 恢复	(132)
5.4.1 故障种类	(133)
5.4.2 恢复实现技术	(134)
5.4.3 恢复策略	(135)
5.4.4 SQL Server 的恢复技术	(136)
习题 5	(140)
第 6 章 数据库设计	(141)
6.1 概述	(141)
6.1.1 数据库设计的任务、内容及特点	(141)
6.1.2 设计步骤	(142)
6.2 需求分析	(143)
6.2.1 需求分析的任务	(143)
6.2.2 需求分析的主要内容	(144)
6.2.3 需求分析的主要步骤	(144)
6.2.4 需求分析说明书	(146)
6.3 概念结构设计	(146)
6.3.1 概念结构设计概述	(147)
6.3.2 数据抽象与局部视图设计	(148)
6.3.3 视图集成	(150)
6.4 逻辑结构设计	(155)
6.4.1 图向数据模型的转换	(155)
6.4.2 关系模型向 DBMS 支持的数据模型转换	(158)
6.4.3 数据模型的优化	(158)
6.4.4 设计外模式	(159)

6.5 数据库物理设计	(160)
6.5.1 聚簇设计	(160)
6.5.2 索引设计	(161)
6.5.3 分区设计	(161)
6.5.4 评价物理设计	(162)
6.6 数据库实施	(163)
6.7 数据库的运行与维护	(165)
习题6	(166)
第7章 数据库技术新进展	(167)
7.1 数据库技术发展概述	(167)
7.2 数据模型及数据库系统的发展	(167)
7.2.1 第一代数据库系统	(168)
7.2.2 第二代数据库系统——关系数据库系统	(169)
7.2.3 新一代数据库技术的研究和发展	(170)
7.3 数据库新技术	(174)
7.3.1 数据模型的发展	(174)
7.3.2 数据库技术与其他技术的结合	(176)
7.3.3 面向应用领域的数据库新技术	(183)
习题7	(186)
附录1 网上书店数据库设计	(187)
1 系统需求分析	(187)
2 E-R 模型设计	(187)
3 数据库模型	(190)
3.1 数据库需求分析	(190)
3.2 数据库表	(190)
3.3 创建数据库脚本	(191)
附录2 在线考试系统数据库设计	(193)
1 系统概述	(193)
2 需求分析	(194)
3 系统建模	(196)
4 数据库分析与设计	(197)
参考文献	(202)

第1章 绪 论

数据库是数据管理的最新技术，是计算机科学的重要分支。当今，信息资源已成为各个部门的重要财富和资源，建立一个满足各级部门信息处理要求的信息系统也成为企业或组织生存和发展的重要条件。因此，作为信息系统核心和基础的数据库技术得到越来越广泛的应用。

从20世纪50年代中期开始，计算机的应用由科学研究逐渐扩展到企业、行政部门等社会领域，数据处理已成为计算机的主要应用。60年代末，数据库技术就是作为数据处理中的一门新技术发展起来的。经过50多年的发展，现已形成了较为完整的理论体系和实用技术。

1.1 引 言

1.1.1 基本概念

数据、数据库、数据库系统和数据库管理系统是与数据库技术密切相关的4个基本概念。

1. 数据 (data)

数据是数据库中存储的对象。从狭义的角度看，数据指数字；从广义的角度看，数据的种类很多，如文字、图形、图像、声音、各种格式的记录等都是数据。数据是描述事物的符号记录，是信息存在的一种形式，只有经过语义的解释或者处理才能成为有用的信息。

人们研究各种事物，可以通过自然语言，也可以使用形式化的方式。数据通常是对事物特征或者属性的描述。例如，下面一组数据：

(朱莉，女，1985，广东，外语系，200431102118)

了解这组数据含义的人，可以从中得出以下信息：朱莉是位外语系的女学生，1985年出生，广东人，学号是200431102118。

2. 数据库 (database，简称DB)

人们经常需要从现实世界的一个事物中收集并抽取出一系列有用的数据，然后再将其保存起来，供进一步加工处理，以获取有用信息。在科学技术飞速发展的今天，人们的视野越来越广阔，数据量急剧增加。过去人们把数据存放在文件柜里，现在人们借助计算机

和数据库技术科学地保存和管理大量的复杂数据，以便能方便而充分地利用这些宝贵的数据资源。

数据库是按照数据结构来组织、存储和管理数据的仓库，其将数据按一定格式存储在计算机的存储设备上，是有组织、可共享的数据集合。数据库中的数据按一定的数据模型进行组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。

3. 数据库管理系统 (database management system, 简称 DBMS)

收集并抽取出应用所需的大量数据之后，为了能高效地处理这些数据，需要由一个软件系统——数据库管理系统进行管理。

数据库管理系统是位于用户和数据库之间的数据管理软件，是数据库系统的重要组成部分。

数据库管理系统负责数据库中的数据组织、数据操纵、数据维护、数据控制及保护等，其目的是尽可能地降低用户程序在数据管理、维护和使用上的复杂程度，提高数据库应用的简明性和方便性。

4. 数据库系统 (database system, 简称 DBS)

数据库系统是指在计算机系统中引入数据库后的系统，它由硬件系统、软件系统和人员三部分组成，其中软件系统包括操作系统、数据库、数据库管理系统、编程工具、应用软件等。通常把数据库系统简称为数据库。

数据库管理员 (database administrator, 简称 DBA) 是对数据库进行规划、设计、协调、维护 and 管理的 IT 技术人员。数据库应用系统是使用数据库语言开发的能够满足数据处理需求的应用程序。用户可以通过数据库管理系统直接操纵数据库，或者通过数据库应用程序操纵数据库。

数据库系统可以用图 1-1 表示。

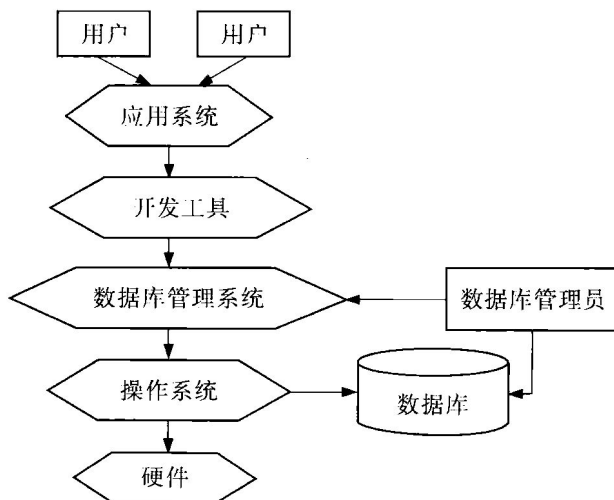


图 1-1 数据库系统

1.1.2 数据库技术的产生与发展

数据管理技术的发展在应用需求的推动下,经历了人工管理、文件系统和数据库系统三个发展阶段。

1. 人工管理阶段

人工管理阶段在20世纪50年代中期以前,计算机主要用于科学计算,并没有管理数据的软件,数据处理方式是批处理。这个时期的数据管理有如下特点:

(1) 数据不保存。

由于当时计算机没有磁盘存储设备,数据无法在计算机中长期保存,只是在需计算某一结果时将数据输入,使用完后就撤走。该阶段不仅对数据如此处理,对程序也是如此。

(2) 没有数据管理软件。

程序员不仅要规定数据的逻辑结构,而且还要在程序中设计物理结构,包括存储结构、存取方法、输入/输出方式等,编程工作量大。

(3) 数据不能共享。

一个程序中的数据无法被其他程序利用,因此程序与程序之间存在大量的重复数据,数据的冗余度大。

(4) 程序和数据不具有独立性。

一组数据只对应一组程序,如果数据的类型、格式等改变了,程序必须做出相应的修改。

2. 文件系统阶段

20世纪50年代后期到60年代中期,随着计算机的发展,操作系统中的文件系统成为专门管理外存的数据管理软件。其处理方式不仅有文件批处理,而且有联机实时处理。这一阶段的数据管理有以下特点:

(1) 数据可以长期保存。

磁盘等直接存储设备的出现,使得数据可以长期保留在计算机内反复进行查询、修改、插入和删除等操作。

(2) 由文件系统管理数据。

文件系统把数据组织成相互独立的数据文件,利用“按文件名访问,按记录进行存取”的管理技术,可以对文件进行修改、插入和删除的操作。但文件系统的功能无法和数据库管理系统相比。

(3) 数据共享性差。

在文件系统中,一个(或一组)文件基本上对应于一个应用程序,即文件仍然是面向应用的。当不同的应用程序具有部分相同的数据时,必须建立各自的文件,而不能共享相同的数据。同时,由于相同数据的重复存储、各自管理,容易造成数据的不一致性,给数据的修改和维护造成了困难。

(4) 数据独立性差。

文件系统上的文件是为某一特定应用服务的,文件的逻辑结构对该应用程序来说是优化的,一旦数据的逻辑结构改变,必须修改应用程序,修改文件结构的定义。因此,数据

与程序之间仍缺乏独立性。

3. 数据库系统阶段

数据库系统的开发应用始于 20 世纪 60 年代末。数据库系统克服了文件系统的缺陷,提供了对数据更高级、更有效的管理。

数据库技术的主要目的是有效地管理和存取大量的数据资源,提高数据的共享性,使多个用户能并发存取数据库中的数据,减小数据的冗余度,以提高数据的一致性和完整性,实现数据和应用程序的独立性,减少应用程序的维护代价。数据库系统管理数据具有如下特点:

(1) 数据结构化。

数据的结构化是数据库的主要特征之一,是与文件系统的根本区别。数据库中的数据是按照一定的数据结构组织起来的,能表示出数据之间所存在的有机关联,从而反映现实世界事物之间的联系。

(2) 数据共享,减少冗余。

在数据库系统中,对数据的定义和描述已经从应用程序中分离出来,通过数据库管理系统来统一管理。在建立数据库时,应当以面向整个系统的观点组织数据,在数据库中存放整个系统的综合性数据,数据可以被多个应用程序共享使用,这样可以大大减少数据冗余,节省存储空间。

(3) 数据具有较高的独立性。

数据独立具有两个含义:物理数据独立性和逻辑数据独立性。物理数据独立性是指数据库物理结构(包括数据的组织、存储、存取方法等)发生改变时,不会影响到逻辑结构,而程序使用的是逻辑结构,所以不必改动程序;逻辑数据独立性是指数据库逻辑结构发生改变时,程序有时也可以不变,因为总体逻辑结构变化并没有影响该程序所使用的局部逻辑结构。

(4) 数据的统一管理和控制。

数据库为多个用户和应用程序所共享,对数据的存取往往是并发的,数据库管理系统可提供以下 4 个方面的数据控制功能:数据的安全性控制、数据的完整性控制、数据的并发控制以及故障恢复控制,从而保障多用户对数据库中数据的共享。

综上所述,数据库是长期存储在计算机内有组织的大量的共享的数据集合。它可以供各种用户共享,具有较小冗余度和较高的数据独立性。DBMS 在数据库建立、运用和维护时对数据库进行统一控制,以保证数据的完整性、安全性,并在多用户同时使用数据库时进行并发控制,在发生故障后对系统进行恢复。

数据库系统的出现使信息系统的研制从以围绕加工数据的程序为中心转变为围绕共享的数据库来进行。这既便于数据的集中管理,也有利于应用程序的研制和维护,提高了数据的利用率和相容性,从而提高了做出决策的可靠性。因此,大型复杂的信息系统大多以数据库为核心。

数据库技术从 20 世纪 60 年代中期产生至今仅有 50 年的历史,但其发展速度之快、使用范围之广是其他技术所不及的。60 年代末出现了第一代数据库——网状数据库、层次数据库,70 年代出现了第二代数据库——关系数据库。目前关系数据库系统已逐渐淘汰了网状数据库和层次数据库,成为当今最为流行的商用数据库系统。而 80 年代出现的以面

向对象模型为主要特征的数据库系统又在向关系数据库系统挑战。数据库技术与网络通信技术、人工智能技术、面向对象程序设计技术、并行计算技术等互相渗透、互相结合,成为当前数据库技术发展的主要特征。

1.1.3 数据库技术的研究领域

目前虽然已有了一些比较成熟的数据库技术,但随着计算机硬件的发展和应用范围的扩大,数据库技术也需要不断向前发展。概括地讲,当前数据库技术主要的研究范围有以下三个领域。

1. 数据库管理系统软件的研制

数据库管理系统(DBMS)是数据库系统的基础。DBMS的研制包括研制DBMS本身以及以DBMS为核心的一组相互联系的软件系统。研制的目标是扩大功能、提高性能和提高用户的生产效率。

随着数据库应用领域的不断扩大,许多新的应用领域如自动控制、计算机辅助设计等要求数据库能够处理与传统数据类型不同的新的数据类型,例如,声音、图像等非格式化数据,面向对象的数据库系统、扩展的数据库系统、多媒体数据库系统等兴起就是应这些新的需求和应用背景而产生的。

2. 数据库应用系统的设计与开发

在DBMS的支持下,数据库应用系统设计与开发的主要任务是按照应用的要求,为某一行业或企业设计一个结构合理、使用方便、效率较高的数据库及其应用系统。其中主要的研究方向是数据库设计方法学和设计工具,包括数据库设计方法、开发工具和设计理论的研究,数据模型和数据建模的研究,计算机辅助数据库设计方法及其软件系统的研究,数据库设计规范和标准的研究等。

3. 数据库理论

很长一段时间,数据库理论的研究主要集中于关系的规范化理论、关系数据理论等方面,近年来,也开始了面向对象数据库、分布式数据库、多媒体数据库的理论研究。随着人工智能与数据库理论的结合及并行计算机的发展,出现数据库逻辑演绎和知识推理、并行算法等理论研究,演绎数据库系统、知识库系统和数据仓库的研制都已成为新的研究方向。

1.2 数据模型

数据模型(data model)是现实世界数据特征的抽象。由于计算机不可能直接处理现实世界中的具体事物,所以人们必须事先把具体事物转换成计算机能够处理的数据。也就是首先要数字化,把现实世界中具体的人、物、活动、概念用数据模型来抽象、表示和处理。简言之,数据模型就是对现实世界的模拟。

数据模型是数据库系统的核心和基础。现有的数据库系统均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基础。

根据数据模型应用的不同目的,可以将数据模型划分为三种,分别为概念模型、逻辑模型和物理模型。

- 概念模型,即按用户的观点来对数据和信息建模,主要用于数据库设计。
- 逻辑模型,它是按计算机系统的观点对概念模型的表示,主要用于 DBMS 的实现。
- 物理模型,描述数据在磁盘上的存储方式和存取方法,是面向计算机系统的。物理模型的具体实现是 DBMS 的任务,一般用户不必考虑物理的细节。

1.2.1 数据模型三要素

数据模型所描述的内容分三个部分,分别是数据结构、数据操作和数据约束。

1. 数据结构

数据结构用于描述系统的静态特征,包括数据的类型、内容、性质及数据之间的联系等。它是数据模型的基础,也是刻画一个数据模型性质最重要的方面。在数据库系统中,人们通常按照其数据结构的类型来命名数据模型。例如,层次模型和关系模型的数据结构就分别是层次结构和关系结构。

2. 数据操作

数据操作用于描述系统的动态特征,包括数据的插入、修改、删除和查询等。数据模型必须定义这些操作的确切含义、操作符号、操作规则及实现操作的语言。

3. 数据约束

数据的约束条件实际上是一组完整性规则的集合。完整性规则是指给定数据模型中的数据及其联系所具有的制约和存储规则,用以限定符合数据模型的数据库及其状态的变化,以保证数据的正确性、有效性和相容性。例如,限制一个表中学号不能重复,或者年龄的取值不能为负,都属于完整性规则。

1.2.2 概念模型

数据模型是数据库系统的核心和基础。各种机器上实现的 DBMS 软件都是基于某种数据模型的。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机器世界。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,这种信息结构并不依赖于具体的计算机系统,也不是某一个 DBMS 支持的数据模型,而是概念级的模型;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 1-2 所示。不难看出,概念模型实际上是现实世界到机器世界的一个中间层次。

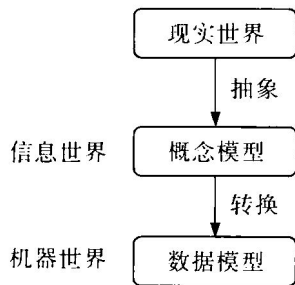


图 1-2 数据模型的抽象过程

概念模型用于信息世界的建模,是现实世界到信息世界的第一层抽象,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, 概念模型一方面应该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; 另一方面它还应该简单、清晰, 易于用户理解。

1. 基本概念

(1) 实体 (entity)

客观存在并可相互区别的事物称为实体。实体可以是具体的人、事、物, 也可以是抽象的概念或联系, 例如, 一个职工、一个学生、一个部门、一门课、学生的一次选课、仓库的一次进货。

(2) 属性 (attribute)

实体所具有的某一特性称为属性。一个实体可以由若干个属性来刻画。例如, 学生实体可以由学号、姓名、性别、出生年份、系、入学时间等属性组成 (200831198766, 李琳, 女, 1990, 计算机系), 这些属性组合起来描述了一个学生。

(3) 域 (domain)

属性的取值范围称为该属性的域。例如, 学号的域为 12 位整数, 姓名的域为字符串集合, 年龄的域为大于 0 的整数, 系别的域为学校实际存在的系名称。

(4) 码 (key)

唯一标识实体的属性集称为码。例如, 学号可以唯一标识一名学生, 则学号是学生实体的码。

(5) 实体型 (entity type)

具有相同属性的实体必然具有共同的特征和联系。用实体名及其属性名集合来抽象和刻画同类实体, 称为实体型。例如, 学生 (学号, 姓名, 性别, 出生年份, 系) 就是一个实体型。

(6) 实体集 (entity set)

同型实体的集合称为实体集。例如, 所有学生就是一个实体集。

(7) 联系 (relationship)

现实世界中事物内部及事物之间的关联, 可以反映为实体内部和实体之间的联系。实体型之间的联系可以分三类, 如图 1-3 所示。

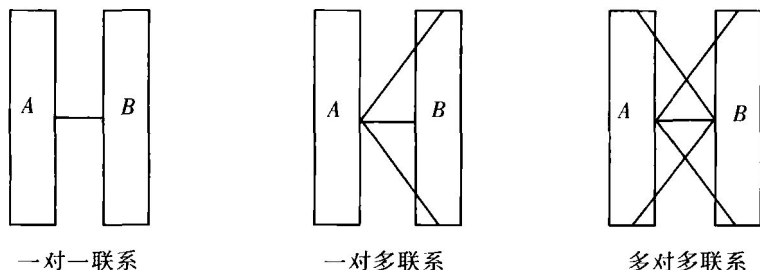


图 1-3 实体间的三种联系

• 一对一联系 (1:1)

如果对于实体集 A 中的每一个实体, 实体集 B 中至多只有一个实体与之对应, 同样对于实体集 B 中的每一个实体, 实体集 A 中至多只有一个实体与之对应, 则称实体集 A 与实体集 B 有一对一联系, 记为 1:1。

例如, 国家与首都之间具有一对一联系。

• 一对多联系 ($1:n$)

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

例如, 国家与省份之间具有一对多联系。

• 多对多联系 ($m:n$)

如果对于实体集 A 中的每一个实体, 实体集 B 中有 n 个实体 ($n \geq 0$) 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 m 个实体 ($m \geq 0$) 与之联系, 则称实体集 A 与实体集 B 具有多对多联系, 记为 $m:n$ 。

例如, 一门课程同时有若干个学生选修, 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。

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。

实体型之间的这种一对一、一对多、多对多联系不仅存在于两个实体型之间, 也存在于两个以上的实体型之间。

多个实体集间也存在以上三种联系。例如, 学生、老师、课室三个实体就存在着学生在课室听老师授课的联系。

2. E-R 模型与 E-R 图

概念模型最为常用的表示方法是实体 - 联系方法 (entity-relationship approach)。该方法用 E-R 图来描述概念模型, E-R 图方法也称为 E-R 模型。E-R 图提供了表示实体、属性和联系的方法。

• 实体型: 用矩形表示, 矩形框内写明实体名。

• 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。

• 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 $m:n$)。

需要注意的是, 联系本身也是一种实体型, 也可以有属性。如果一个联系具有属性, 则这些属性也要用无向边与该联系连接起来。

图 1-4 是一个学生成绩管理系统的 E-R 图。

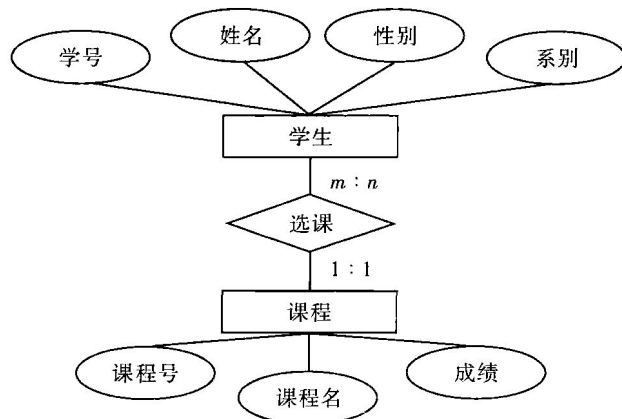


图 1-4 E-R 图实例

1.2.3 数据模型

数据库的类型是根据数据模型来划分的，而任何一个 DBMS 也是根据数据模型有针对性地设计出来的，这就意味着必须把数据库组织成符合 DBMS 规定的数据库模型。目前成熟地应用在数据库系统中的数据库模型有：层次模型、网状模型和关系模型。它们之间的根本区别在于数据之间联系的表示方式不同（即记录型之间的联系的方式不同）。层次模型以“树结构”表示数据之间的联系，网状模型是以“图结构”表示数据之间的联系，而关系模型则是用“二维表”（或称为关系）来表示数据之间的联系的。

在非关系模型中，实体用记录表示，实体之间的联系转换成记录之间的两两联系。非关系模型数据库结构的基本单位是基本层次联系。所谓基本层次联系，是指两个记录以及它们之间的一对多（包括一对一）的联系。

1. 层次模型 (hierarchical model)

层次模型是数据库系统最早使用的一种模型，它的数据库结构是一棵“有向树”，如图 1-5 所示。根结点在最上端，层次最高，子结点在下，逐层排列。层次模型的特征是：

- 有且仅有一个结点，没有父结点，其本身就是根结点；
- 其他结点有且仅有一个父结点。

层次数据库系统只能处理一对多的实体关系。

最有影响的层次模型 DBMS 是 20 世纪 60 年代末，由 IBM 公司推出的 IMS 层次模型数据库系统。

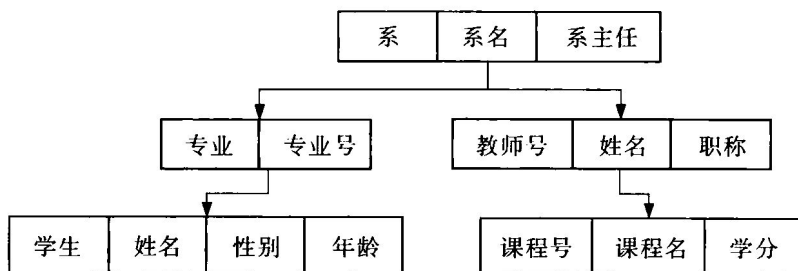


图 1-5 层次模型

2. 网状模型 (network model)

网状模型以网状结构表示实体与实体之间的联系，使用链接指针来实现。网状模型可以表示多个从属关系的联系，也可以表示数据间的交叉关系，即数据间的横向关系与纵向关系，它是层次模型的扩展，如图 1-6 所示。网状模型可以方便地表示各种类型的联系，但结构复杂，实现的算法难以规范化。其特征是：

- 允许结点有多于一个父结点；
- 可以有一个以上的结点没有父结点。