

数据库技术及应用

谭 龙 屈广宁 宋成杰 编著

哈尔滨地图出版社

数据库技术及应用

SHUJUKU JISHU JI YINGYONG

谭龙 屈广宁 宋成杰 编著

哈尔滨地图出版社

·哈尔滨·

内 容 提 要

本书将数据库的基础理论、数据库管理系统、数据库应用产品和数据库应用技术融为一体,覆盖了数据库的基本理论和知识、数据库管理系统的实现原理,以及数据库设计开发的方法,有助于读者全面理解和掌握数据库的应用技术。

全书包括数据库基础、数据库系统、数据库设计、数据库产品介绍和数据库应用等内容,共五篇,具体内容主要包括:数据库理论的基础知识、数据库设计、SQL Server 数据库技术及具体的应用系统实例等。

本书既可以作为高等院校计算机、信息管理与信息系统等相关专业数据库课程的教材,也可供从事计算机软件以及数据库应用、管理和开发的科技人员、工程技术人员及其他有关人员阅读参考。

图书在版编目(CIP)数据

数据库技术及应用/谭龙,屈广宁,宋成杰编著
哈尔滨:哈尔滨地图出版社,2006.1

ISBN 7-80717-332-7

I. 数... II. ①谭...②屈...③宋... III. 数据库
系统 IV. TP311.13

中国版本图书馆 CIP 数据核字(2006)第 044244 号

哈尔滨地图出版社出版发行

(地址:哈尔滨市南岗区测绘路 2 号 邮政编码:150086)

哈尔滨市民强印刷厂印刷

开本:787mm×1092mm 1/16 印张:17.875 字数:458 千字

2006 年 1 月第 1 版 2006 年 1 月第 1 次印刷

印数:1~1000 定价:29.80 元

前 言

数据库技术是计算机科学技术中发展最快的领域之一。数据库应用理论和技术是计算机及其相关专业大学生的必修内容，也是计算机工程技术人员必备的知识和技能。

本书覆盖了数据库的基本理论和知识、数据库管理系统的实现原理，以及数据库设计开发的方法，有助于读者全面理解和掌握数据库的应用技术。本书在内容选择上注重理论的基础性和技术的实用性；在技术实践的相关章节中，本书结合当前应用广泛的 SQL Server、Oracle 等具体数据库产品进行介绍。本书侧重介绍了数据库设计技巧和方法，因此无论是对在校的大学生的课程实践或者毕业设计，还是计算机工程技术人员的应用开发，均具有广泛的指导作用。

全书共分为数据库理论知识、数据库系统设计、SQL Server 技术、数据库技术应用四部分，主要内容分布为：

- 第1章 介绍了数据描述和数据模型，数据管理技术的发展等内容；
- 第2章 介绍关系数据库理论，包括关系代数和关系演算。
- 第3章 介绍关系数据库的规范化理论。
- 第4章 介绍了数据库设计基本过程和方法。
- 第5章 通过实例对三大范式的设计进行分析，并介绍了常用的数据库设计技巧。
- 第6章 介绍了数据库与应用系统之间的关系。
- 第7章 介绍了 SQL Server 2000 的体系结构，安装，以及主要工具。
- 第8章 介绍了如何创建数据库和表。
- 第9章 介绍了 SQL Server 2000 简单查询、高级查询以及视图的建立和使用。
- 第10章 介绍了 SQL Server 存储过程和触发器的特点、创建方法和使用管理方法。
- 第11章 介绍了如何通过设置特定的服务器配置，更改网络连接或者设置服务器配置选项来提高 SQL Server 性能。
- 第12章 介绍了 SQL Server 2000 的数据安全性及完整性操作。
- 第13章 介绍了 SQL Server 2000 的备份和恢复操作。
- 第14章 介绍了 SQL Server 2005 的主要特性，特点，新增功能以及升级工作等。
- 第15章 介绍了一个实际的管理系统设计与分析过程。

本书内容详尽、示例丰富、通俗易懂，结构合理，语言简洁，能帮助读者在掌握理论知识的同时，提高应用能力。

在本书的编写过程中，参考了很多同行的著作，在此表示深深的谢意。

限于时间和水平，书中错误之处在所难免，敬请广大读者和专家批评指正。

编 者
2005.8

目 录

第 1 章 数据库技术概述	1
1.1 数据与信息	1
1.2 数据管理技术的发展	2
1.3 数据及数据模型	7
1.4 数据库系统的组成	14
1.5 数据库的保护	16
1.6 数据库技术的发展	18
第 2 章 关系型数据库理论	27
2.1 关系及关系代数	27
2.2 关系型数据库标准语言 SQL	39
第 3 章 关系型数据库的规范化理论	58
3.1 函数依赖	60
3.2 规范化的关系模式	62
3.3 关系模式的分解	65
第 4 章 数据库系统设计	71
4.1 概述	71
4.2 需求分析	73
4.3 概念设计	76
4.4 逻辑设计	81
4.5 物理设计	82
第 5 章 范式设计和数据库设计技巧	84
5.1 数据库的范式设计实例	84
5.2 数据库设计技巧	87
5.3 大型数据库的设计原则	91
第 6 章 数据库与应用系统	94
6.1 应用系统的层次结构	94
6.2 数据库与应用系统	98
第 7 章 SQL Server 概述	104
7.1 SQL Server 2000 的体系结构	104
7.2 SQL Server 2000 的新特性	105
7.3 SQL Server 2000 的安装	107
7.4 SQL Server 2000 的两个主要工具	114
第 8 章 数据库和表的创建	120
8.1 数据库的创建	120
8.2 数据表的创建	132
第 9 章 数据查询	152
9.1 Transact-SQL 简单查询	152

9.2 Transact-SQL 高级查询.....	1 7 1
9.3 视图.....	1 7 9
第 10 章 存储过程和触发器.....	1 9 0
10.1 存储过程.....	1 9 0
10.2 触发器.....	1 9 5
第 11 章 SQL Server 服务器的管理.....	2 0 3
11.1 注册 SQL Server 服务器.....	2 0 3
11.2 配置 SQL Server 服务器.....	2 0 6
第 12 章 数据安全性及完整性.....	2 1 0
12.1 SQL Server 的安全性机制.....	2 1 0
12.2 管理服务器的安全性.....	2 1 1
12.3 SQL Server 数据库的安全性.....	2 1 4
12.4 表和列级的安全性.....	2 1 9
12.5 完整性的概念.....	2 2 3
12.6 使用约束实施数据完整性.....	2 2 3
12.7 使用规则.....	2 2 4
12.8 使用默认值.....	2 2 5
12.9 使用 IDENTITY 列.....	2 3 0
第 13 章 备份和恢复.....	2 3 3
13.1 备份概述.....	2 3 3
13.2 备份数据库.....	2 3 5
13.3 恢复数据库.....	2 3 9
第 14 章 SQL Server 2005.....	2 4 2
14.1 SQL Server 2005 概述.....	2 4 2
14.2 SQL Server 2005 的主要特性.....	2 4 3
14.3 SQL Server 2005 的特点.....	2 4 4
14.4 SQL Server 2005 的新增功能.....	2 4 6
14.5 升级到 SQL Server 2005.....	2 5 8
第 15 章 企业人事库存管理系统设计与分析.....	2 5 9
15.1 应用需求分析.....	2 5 9
15.2 系统功能模块划分.....	2 6 0
15.3 系统数据库设计.....	2 6 1
15.4 系统应用程序设计.....	2 6 9
15.5 人事管理模块的实现.....	2 7 6
15.6 库存管理模块的实现.....	2 7 9
参考文献.....	2 8 3

第1章 数据库技术概述

随着信息社会发展,要求计算机不但能进行科学计算,而且能进行大量数据的简单处理(如数据的查询和更新),使计算机的应用从科学研究部门逐步扩展到企业、行政部门,因而产生了数据库技术。数据库技术是计算机应用领域的重要分支,其核心任务是进行数据管理。数据库技术产生于20世纪60年代末,现已形成相当规模的理论体系和使用技术。本章主要讲解以下几个问题,使读者对数据库技术有一个完整的了解:

- 数据管理技术的发展
- 数据描述和数据模型
- 数据库系统综述

1.1 数据与信息

数据库技术产生于20世纪60年代末,70年代初期,其主要目的是有效地管理和存取大量的数据资源。在计算机应用中,数据处理和以数据处理为基础的信息系统所占的比重最大。人类的一切活动都离不开数据,离不开信息。

1.1.1 数据、信息与数据处理

1. 数据:是指存储在某一媒体上可加以鉴别的符号资料。

数据的概念包括两个方面,其一,数据内容是事物特性的反映或描述;其二,数据是存储在某一媒体上符号的集合。这里的符号,不仅指数字、字母、文字和其他特殊字符,而且还包括图形、图像、动画、影像、声音等多媒体数据。存储不仅是在纸上,包括记录在磁介质、光介质上、半导体存储器里。

2. 信息:是关于现实世界事物的存在方式或运动形态的综合反映,是人们进行各种活动所需要的知识。

数据与信息的关系:数据是载荷信息的物理符号或称为载体。信息是人们消化理解了的数据,是对客观世界的认识,即知识。

3. 数据处理:是指将数据转换成信息的过程。广义地讲,处理包括对数据的收集、存储、加工、分类、检索、传播等一系列活动。狭义地讲,处理是指对所输入的数据进行加工处理。可用以下式简单表示:

$$\text{信息} = \text{数据} + \text{处理}$$

1.1.2 信息系统

1. 信息系统

计算机信息系统是指为了某些明确的目的而建立的由人员、设备、程序和数据集合构成的统一整体。根据系统实际应用目的,信息系统可分为面向外部,实现对外信息服务的开放式信息系统和面向内部业务和管理的管理信息系统两大类。

2. 管理信息系统分类

面向不同管理层次,管理信息系统可分为三类:

数据处理系统 (EDP Electronic Data Processing): 操作型的数据库系统, 为提高工作效率和质量。

管理信息系统 (MIS Management Information System): 有一定的综合能力, 为各部门提供一定的信息服务。

决策支持系统 (DSS Decision Support System): 为管理者提供决策依据。

1.2 数据管理技术的发展

数据管理技术是对数据的分类、组织、存储、操作和维护的技术。简单地说计算机是数据处理机, 输入原始数据, 经过计算机的处理, 获得我们所需要的信息。在计算机处理中, 数据的管理显示了其更加重要的作用。数据管理技术经历了人工管理阶段、文件系统阶段、数据库系统阶段。

1.2.1 人工管理阶段 (20 世纪 50 年代中期以前)

计算机发展的初期主要应用于科学计算, 这一阶段计算机的软、硬件的发展也处于初级阶段, 计算机的硬件上只有磁带、卡片、纸带, 没有磁盘等直接存取的存储设备; 软件上没有操作系统实现对计算机数据的统一管理和调度, 数据是由程序员设计应用程序时设计, 交给应用程序进行管理。其特点为:

1. 数据不能存储。由于没有直接存取的存储设备, 加上计算机主要用于科学计算, 一般不需要将数据长期保存, 在程序的运行过程中进行数据输入, 程序运行结束后, 程序和数据所占有的存储空间被释放或被其他的程序或数据覆盖。
2. 没有软件对数据实施统一管理。数据需要程序员设计应用程序进行管理, 包括数据存取方式、输入/输出方式、数据的操作方式。在程序中不但要设计数据的逻辑结构, 而且还要设计数据的物理结构, 对编写应用程序的程序员, 编制这种类型的程序有相当的难度。
3. 数据不能共享。由于数据是面向应用程序的 (写在程序中或在程序运行时由键盘输入), 一组数据对应一个程序, 即使两个应用程序使用的是完全相同的数据, 这些数据仍然需要分别定义、分别输入, 无法相互利用。数据存在冗余数据。
4. 数据的独立性差。由于数据和应用程序组织在一起, 当数据的逻辑结构或物理结构改变时, 要通过改变应用程序来实现; 当应用程序改变时, 数据的逻辑结构和物理结构也要发生相应的改变。数据和应用程序之间缺乏独立性。

1.2.2 文件系统阶段 (20 世纪 50 年代后期至 60 年代中期)

随着计算机软、硬件技术的发展, 计算机的应用范围从最初的科学计算扩展到各个行业, 要求计算机不但能用来进行高精度的科学计算, 而且能用来处理大批量的数据。计算机硬件技术的发展使得数据的长期存储成为可能; 操作系统的出现使得能够将数据单独组织成为数据文件存储在外存中, 由操作系统统一管理, 应用程序能够方便的访问这些数据。这一阶段数据管理的特点:

1. 数据可以长期保存。数据以独立数据文件的形式长期存储在外存储器上, 可以被应用程序随时访问。
2. 应用程序和数据之间具有“设备独立”性。数据和应用程序以文件方式分别存储在外存中, 程序与数据具有了一定的独立性。

3. 数据可以重复利用。由于数据是独立存放的, 如果应用程序使用的数据与数据文件中存储的数据在数据的内容和数据的格式上是完全一致的, 那么这个数据文件可以被这类应用程序使用, 在一定程度上避免了数据的冗余。

1.2.3 数据库系统阶段 (20 世纪 60 年代末期以后)

文件系统的产生是数据管理技术发展中的一个重要阶段, 使数据具有了“设备独立性”, 即数据的存储设备发生改变时不影响应用程序。但文件系统阶段的数据管理技术在进行大量数据的管理时仍然存在着一定的局限性: 数据不具备“物理结构独立性”和“逻辑结构独立性”, 即数据的存储结构或逻辑结构发生改变时, 应用程序也要做相应的改变; 数据文件基本上是为特定的应用程序设计, 数据的格式没有统一的规定, 只有当一组应用程序使用的数据在内容和其格式完全相同时, 数据可以达到共享。随着数据管理要求的提高, 处理的数据量不断增加, 数据的共享性的要求越来越强, 为提高数据处理的效率, 加上计算机硬件技术的提高, 数据库技术应运而生。数据库阶段数据管理的特点:

1. 有专门的数据管理软件—数据库管理系统 (DBMS), 对数据实施统一的管理和控制。数据库管理系统是计算机厂商提供的数据库管理软件, 数据库系统中的数据由 DBMS 统一管理和控制, 这些控制包括数据的完整性控制、数据的恢复、并发控制、事务支持。数据库系统的数据对系统中的用户是共享资源, 由 DBMS 统一管理, 所有的应用程序可以同时存取数据库中的数据, DBMS 还提供方便的用户接口, 利用命令 (数据库语言) 对数据库中的数据实施查询和更新操作, 从某种程度上说, DBMS 系统软件的发展水平标志着数据库技术的发展水平。

2. 数据结构化。数据库系统的数据都是统一设计, 用复杂的数据模型表示数据。数据模型不仅描述数据本身的特点, 还描述数据之间的联系。数据不再面向特定的一个或多个应用程序, 而是面向整个应用系统。

3. 数据共享性高、冗余度小。数据库系统中数据是独立于应用系统的, 不为某个或某类应用系统而设计, 多个应用程序可以同时访问数据库中的数据, 甚至可以同时访问同一个数据, 达到数据的共享。同时数据库中的数据是专门设计的, 有效的控制了数据的冗余, 保持了数据的一致性。

4. 数据独立性高。数据库系统中的数据具有较高的数据独立性, 这主要是指数据与应用程序之间的独立。数据库系统中数据结构采用三级结构: 用户逻辑结构; 数据库整体逻辑结构; 数据库的物理结构。每级结构之间具有相应的独立性, 这种独立性使得在数据库的整体逻辑结构改变时尽量不影响用户的逻辑结构; 当数据库的物理结构发生改变时尽量不影响数据的整体逻辑结构。数据和应用程序具有较好的独立性。

5. 数据操作灵活。可以整体操作数据库中的数据、可以方便地操作其中的某些数据或数据项, 进行数据的查询和更新操作, 也可以方便地进行数据的扩展。

以上五个方面构成了数据的主要特征。综上所述, 数据库系统中的数据是长期存储在计算机中大量的、有组织的、可以共享的, 冗余度小、独立性强, 可以统一管理和控制的数据的集合。

1.2.4 数据的传统管理方式与数据库管理方式的比较

数据管理技术的发展经历了人工管理、文件系统、数据库系统三个阶段, 这三个阶段中

数据与程序的联系、系统设计各有自己的特点,从总体上来说数据管理方式可分为传统管理方式和数据库管理方式,其中人工管理方式和文件系统为传统管理方式,数据库系统阶段为数据库方式。

1. 三个阶段应用程序和数据之间的关系

数据管理技术发展的三个阶段最明显的特征是数据与应用程序的联系程度不同,数据管理技术发展的目标是增加数据与应用程序的独立性,减少数据和应用程序的联系,从而增强数据共享性、减少数据的冗余;增强数据之间的联系,保证数据的一致性和安全性。程序和数据之间的联系特点如图 1-1 所示。

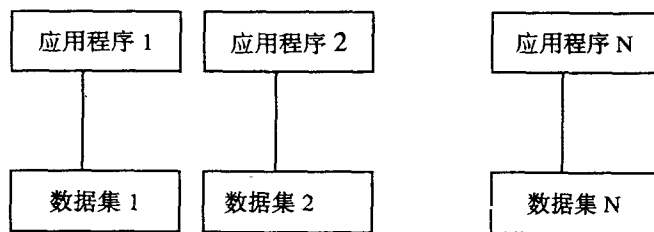


图 1-1 人工管理阶段数据和应用程序之间关系图

图 1-1 反映了数据人工管理阶段的特点:应用程序和所处理的数据之间的关系是一一对应的,而且数据之间没有联系,数据是由程序员设计、应用程序管理的。

图 1-2 中反映了文件系统阶段数据管理的特点:应用程序和处理的数据之间的关系是多对多的关系,即一个应用程序可以操作多个数据文件中的数据、一个数据文件可以被多个应用程序操作;数据文件、应用程序由操作系统统一管理,数据对应用程序具有一定程度的共享性。但是,在文件系统阶段,数据文件之间没有任何的关系,使得数据的共享性、一致性、冗余度受到一定的限制。

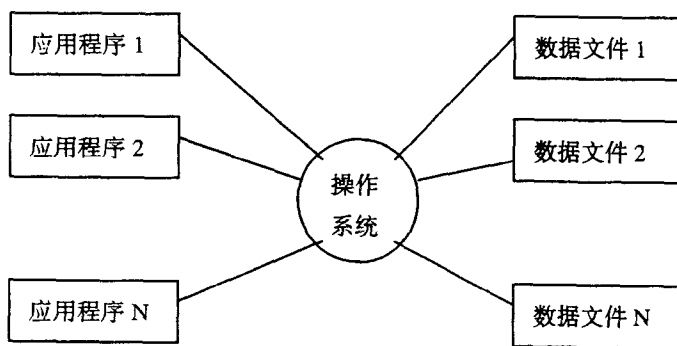


图 1-2 文件系统阶段数据和应用程序之间关系图

图 1-3 中反映了数据库系统阶段数据管理的特点:应用程序和所处理的数据之间的关系是多对多的;应用程序和数据由数据库管理系统统一管理;用数据库中的数据既可以描述数据也可以描述数据之间的联系;数据库系统中的数据由 DBMS 提供简单的数据接口,可以使应用程序方便地对数据库中的数据进行操作。数据库系统中,数据的共享性、一致性提高,数据的冗余度降低,系统的可扩展性增强。

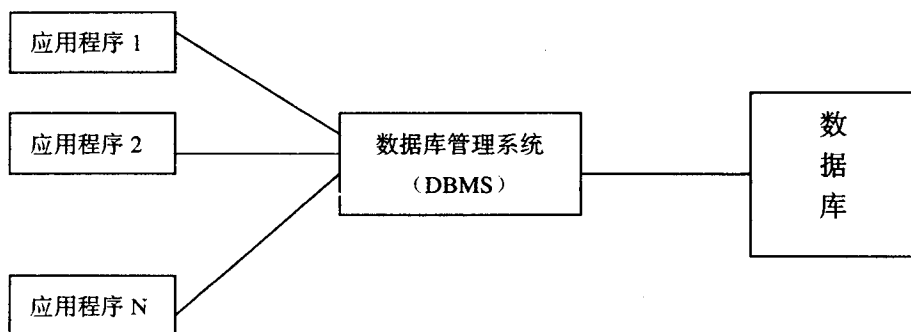


图 1-3 数据库系统阶段数据和应用程序之间的关系图

2. 传统方式和数据库方式系统设计的特点

数据管理技术发展的三个阶段,从计算机系统设计的角度看可以分为传统方式和数据库方式,其中人工管理阶段和文件系统阶段属于传统方式,数据库管理阶段属于数据库方式。在传统方式下,系统设计是以算法为中心,系统设计的中心任务是算法的设计;数据库系统阶段,系统设计的是以数据为中心,数据的建立、操作、维护都由数据库管理系统完成,用户的中心任务就是设计出能反映现实实际的数据,并将这些数据组织成数据库的形式存储在计算机中,所以数据库系统的设计实际上是数据的设计。图 1-4 反映了数据管理系统设计方法的演变。

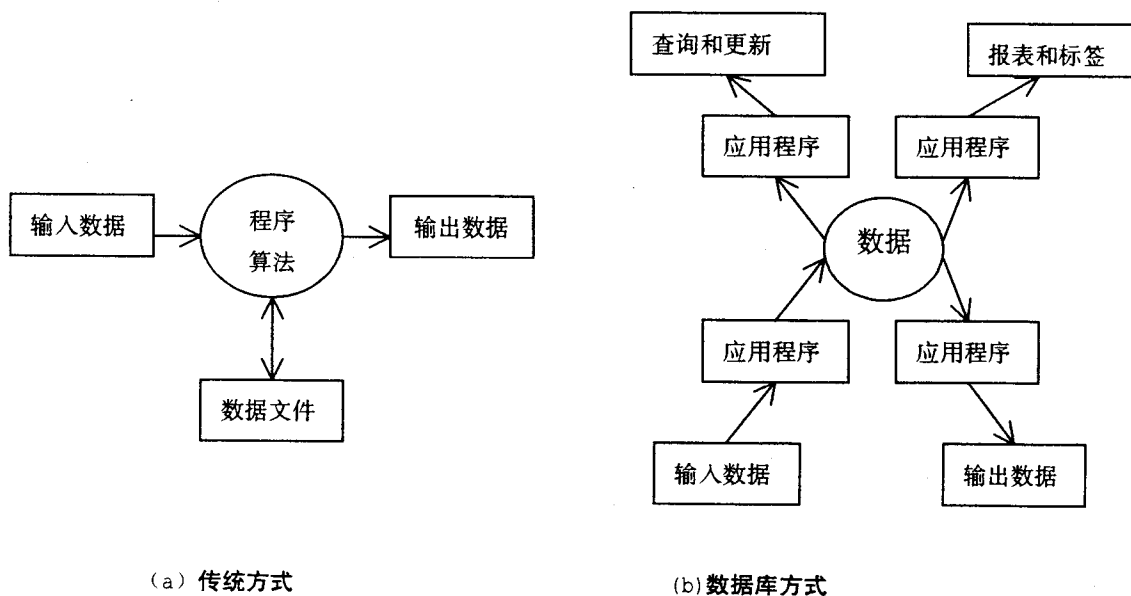


图 1-4 数据管理系统设计方法的演变

数据管理技术从传统方式到数据库方式的演变是一个重大的变化,传统方式系统的设计以程序、算法为中心,数据的逻辑结构、存储结构在应用程序中定义,应用程序的设计困难,

数据库方式系统的设计以数据为中心，数据的逻辑结构、存储结构由数据库管理系统定义，数据的定义、操作、维护也由数据库管理系统提供的用户接口（命令）实现，系统设计的中心任务是数据的设计。

数据库技术的产生使计算机的应用深入到人类社会生活的各个领域，如企业管理、银行业务、商业系统、邮电通讯系统、行政管理等。数据库技术还在不断地发展，不断地与其他计算机技术相渗透。数据库技术与网络通讯技术相结合，产生了分布式数据库系统；数据库技术与面向对象技术相结合，产生了面向对象数据库系统等。

1.3 数据及数据模型

上一节我们讨论了数据管理技术的发展阶段,数据库系统是以数据为中心的、专门进行数据管理的计算机系统,因此数据库系统设计的主要任务是数据的设计,是研究如何将客观世界描述成数据库技术所研究的数据。

1.3.1 数据描述的三个领域

1. 现实世界

现实世界是存在于人们头脑之外的客观世界,是客观事物及其相互联系。例如,学校教学中涉及的学生管理、教师管理、课程管理。管理者要求:每个学期开学时制作学生选修课程情况表,内容包括学号、姓名、课程名、选修课程类别(类别分为必修、选修);每个学期结束时制作学生选修课程成绩表,内容包括学号、姓名、课程名、选修课程的类别、总评成绩;制作教师授课安排表,内容包括教师号、教师名、课程名、授课类别(授课类别分为主讲、辅导、实验)、学时数、班级数等。这就是现实世界,是数据库设计者接触到的最原始的数据,数据库设计者对这些原始数据进行综合、抽象成为数据库技术所研究的数据,现实世界描述数据的形式,就成为信息世界。

2. 信息世界

信息世界是现实世界的符号描述,即将客观世界用数据来描述。例如,学生是客观世界中的个体,可以用一组数据(学号、姓名、性别、年龄、班级、成绩)来描述,有这样一组数据不见其人便可了解该学生的基本情况。因此可以说信息世界就是我们所说的数据世界。信息世界中的术语:

实体:客观世界存在的、可以区别的事物称为实体。实体可以是具体的事物,例如,学生李、教师张、数学课;也可以是抽象的事件,例如:本学期学生李选修了哪些课程,教师张教授了哪门课程,读者的一次借阅活动等。

属性:实体有很多特性,每个特性称为实体的一个属性,每个属性有一个类型。例如,学生实体的属性有:学号、姓名、性别、年龄、班级、成绩,其中学号、姓名、班级的类型为字符型,性别的类型为逻辑型,年龄的类型为整型。

实体集:性质相同的实体的集合。例如,全体学生的集合,全体教师的集合等。

实体标识符:能够惟一标识实体的属性或属性的集合。如学生实体集的属性:学号,能够唯一确定一个学生,因此可以作为学生实体集的标识符。

3. 机器世界

信息世界中的数据在机器世界中的存储,成为计算机的数据。机器世界中对数据的描述采用数据库技术的专业术语,对应于信息世界中的术语有以下四个专业术语:

记录:对应于每一实体的数据。例如,学生这一实体的一组数据(20010310,王兰,女,19,筑42,87)就是一条记录。

字段:对应于信息世界中的属性,在数据库技术中称为字段,学生实体中学号、姓名、性别、班级、成绩都是字段,每个字段都有类型、取值范围。字段的取值范围称为字段的域。

数据文件:对应于信息世界的实体集。由若干条记录组成的数据集合,在计算机中以文件的形式存放。

关键字:能够惟一标识记录的字段或字段表达式,与信息世界中的实体标识符相对应,

例如, 学生实体中的学号, 在没有重名的前提下姓名也可以作为关键字。

从客观世界到信息世界不是简单的数据描述, 而是从客观世界中抽象出适合数据库技术研究的数据, 同时要求这些数据能够很好地反映客观世界的事物; 从信息世界到机器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。所谓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; 所谓数据的物理结构是实际存储在存储设备上的数据。

在数据库系统中, 数据的逻辑结构与数据的物理结构之间可以差别很大, 数据的逻辑结构面向程序员, 数据的物理结构面向机器。数据库管理软件的功能之一, 就是要能够把数据的逻辑结构映象为数据的物理结构、把数据的物理结构映象为数据的逻辑结构。

1.3.2 数据模型

模型是现实世界特征的模拟和抽象。在数据库技术中, 用数据模型的概念描述数据库的结构和语义, 是对现实世界的数据库抽象, 数据模型是研究数据库技术的核心和基础。

数据库技术中研究的数据模型分为两个层面, 一层是面向用户的, 我们称为概念模型, 另一个层面是面向计算机系统的, 我们称为结构模型。下面我们分别做介绍。

1.3.2.1 概念数据模型

概念数据模型是独立于计算机系统的数据模型, 用来描述某个特定组织关系的信息结构, 属于信息世界的建模, 所以概念模型应该能够方便、准确地表示信息世界中常用的概念。另外概念数据模型也是用户和应用系统设计员互相交流的桥梁, 以保证数据模型能够正确地描述客观世界。

概念模型的表示方法很多, 其中最为常用的是 P.P.Chen 于 1976 年提出的实体—联系图方法 (Entity — Relationship Approach), 简称 E-R 模型。E-R 实体联系图是直观表示概念模型的工具, 其中包含了实体、联系、属性三个成分, 联系的方法分为一对一 (1:1)、一对多 (1:N)、多对多 (M:N) 三种方式, 联系是哪种方式取决于客观实际本身。

实体: 用矩形框表示, 矩形框内填上实体的名称。

联系: 用菱形框表示, 菱形框内填上联系的名称。

属性: 用椭圆形框表示, 椭圆形框内填上属性的名称。

实体和实体之间的联系用无向线段连接, 在线段上标注联系的类型, 实体和联系都有各自的属性。例如, 在学生选课管理系统中涉及学生和课程两个实体, 同时这两个实体之间是有联系的, 可以用图 1-5 的 E-R 模型图表示。

图 1-5 的 E-R 图表示的是学生与所选课程之间的联系。图中反映现实世界中的两个实体学生和课程, 其中学生由学号、姓名、性别、年龄四个特性描述, 即通过这四个特性可以了解学生的基本情况, 同样, 通过课程号、课程名、计划学时数这三个属性描述了课程实体的特性, 图中还表示了实体学生和实体课程之间的联系, 它们的联系方式是多 (M) 对多 (N) 的, 通常表示为 (M:N), 即一个学生可以选多门课程, 一门课程可以有多个学生选修, 联系也有名称, 如联系名: 学生—课程, 图示为联系 (S-C), 联系也有属性, 属性名为成绩, 描述的现实世界的意义为: 某个学生选修某门课程应有一个选修该课程的成绩。

从上面的 E-R 实体—联系图中可以看出: E-R 图, 既表示实体, 也表示实体之间的联系, 是现实世界的抽象, 与计算机系统没有关系, 是可以被用户理解的数据描述方式。通过 E-R 图可以使用户了解系统设计者对现实世界的抽象是否符合实际情况, 从某种程度上说 E-R

图也是用户与系统设计者进行交流的工具, E-R 模型已成为概念模型设计的一个重要设计方法。另外, 学生实体的实际特性除了 E-R 图中表示的几种外还有身高、体重等没有在 E-R 图中表示出来的特性, 那么 E-R 图中应表示哪些特性, 不表示哪些特性呢? 这要视实际管理工作中所涉及的问题而定, 与工作有关的特性就表示, 相反与管理工作无关的特性就不表示。

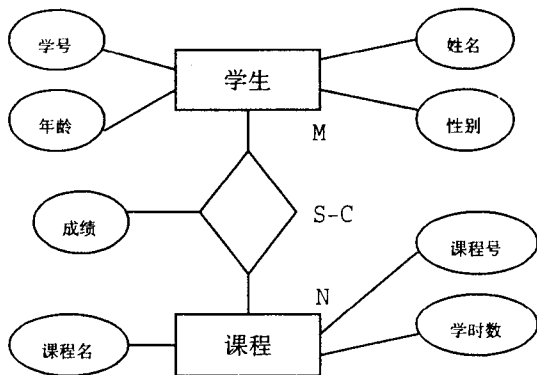


图 1-5 学生、课程实体联系图 (E-R 图)

1.3.2.2 结构数据模型

概念数据模型是对现实世界的描述, 这种数据模型最终要转换成计算机能实现的数据模型。现实世界的第二层抽象是直接面向数据库的逻辑结构, 称为结构数据模型, 这类模型涉及到计算机系统和数据库管理系统。结构数据模型的三个组成部分是

数据结构: 实体和实体间联系的表示和实现。

数据操作: 数据库的查询和更新操作的实现。

数据完整性约束: 数据及其联系应具有制约和依赖规则。

不同的结构数据模型有不同的数据结构形式, 常用的数据结构模型有层次模型、网状模型、关系模型和面向对象模型。其中层次模型和网状模型属于非关系模型, 在数据库技术发展的初期 (20 世纪 70 年代) 非常流行, 在数据库系统产品中占主导地位, 到了 20 世纪 80 年代, 逐渐被关系型的数据库系统所取代。也正是关系型数据库系统的出现, 使得数据库技术得到快速的发展。下面介绍层次、网状、关系、面向对象数据模型。

1. 层次模型

层次模型是数据库系统中最早出现的数据模型, 这种数据模型用树型结构表示实体及实体之间的联系, 树的结点是记录类型, 表示的是实体及联系, 有一个根结点, 每个非根结点有且只有一个父结点, 上一层结点和下一层结点之间的联系方式为一对多。我们看到层次模型实际上描述了实体及实体与实体之间的联系, 是概念模型的计算机存储结构, 因此在数据库系统设计时, 需要将概念模型转换为层次模型。图 1-5 所表示的 E-R 模型转换为层次模型为图 1-6 所示。

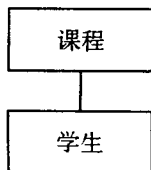


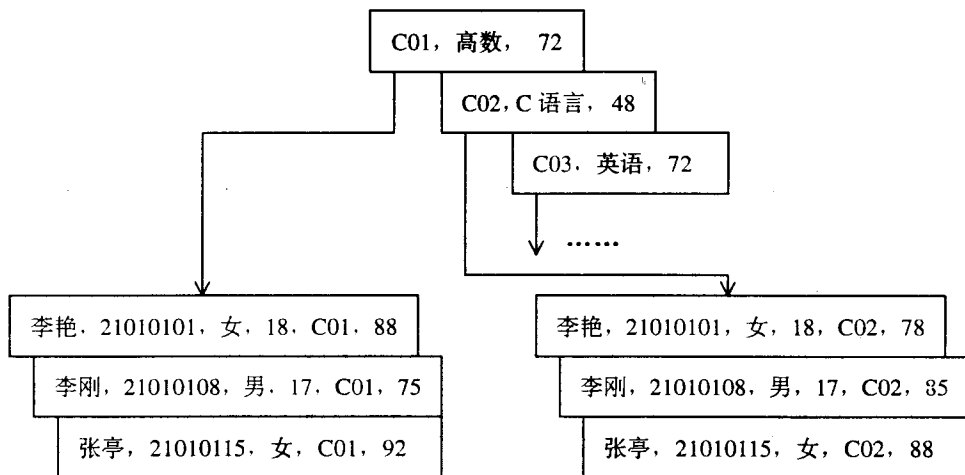
图 1-6 层次模型

在图 1-5 的 E-R 模型中学生和课程之间的关系是多对多，即一个学生可选修多门课程，一门课程可以有多个学生选修，在层次模型中把学生与课程之间的关系转换成一对多，联系类型 S-C 合并到记录类型学生中。可以用两类记录型数据描述此概念模型。

课程（课程号，课程名，学时数）

学生（学号，姓名，性别，年龄，课程号，成绩）

该模型的一个具体实例如下：



从上面的实例中我们看到，在层次模型中，记录之间的联系通过指针作为向导的，操作困难，查询效率低。

2. 网状模型

在网状模型中用有向图表示实体及实体之间的联系。在有向图中，结点是记录型的，有向边表示从箭尾一端的记录类型到箭头一端的记录类型间的联系是一对多，网状模型中实体及实体之间的联系没有层次模型中严格的规定，因此将 E-R 概念模型转换成网状模型时很容易，只是将 E-R 模型中的实体和联系均转换成记录型，每个多对多联系用两个一对多联系实现，图 1-5 的 E-R 模型转换成网状模型如图 1-7 所示。

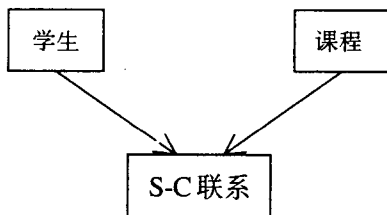


图 1-7 网状模型

在网状模型中，实体及实体间的多对多联系容易实现，查询效率较高。记录之间的联系依然通过指针作向导来实现，操作困难，编写应用程序复杂，程序员必须熟悉数据库的逻辑结构。由于层次模型和网状模型的操作都是通过指针来实现的，所以应用程序的编制比较复杂，从 20 世纪 80 年代中期起，其市场已被关系模型的数据库系统所取代。

3. 关系模型

关系模型的理论基础是数学理论,数据的操作通过关系运算实现。在关系模型中用二维表表示实体及实体之间的联系,关系模型的实例称为关系,而从数学的观点上看,关系是集合,其他元素是元组(记录)。同时可以遵循一定的规则,将E-R模型转换成关系模型。

将E-R模型转换成关系模型的规则:E-R模型中的主要成分是实体及实体间的联系,对于实体的转换方式是:

①将一个实体转换成为一个关系模式。实体的属性为关系模式的属性,实体的标识符为关系模式的关键字,如图1-5的E-R模型中有两个实体:学生、课程,可以分别转换成学生模式和课程模式:学生模式(学号,姓名,性别,年龄),学号是学生模式的关键字。

课程模式(课程号,课程名,学时数),课程号是课程模式的关键字。

②联系转换成为关系模式。联系转换成为关系模式时,要根据联系方式的不同采用不同的转换方式:

●若联系的方式是一对一的(1:1),可以在两个实体关系模式中的任意一个关系模式中加入另一个关系模式的关键字和联系类型的属性。

●若联系方式是一对多的(1:N),则在N端实体的关系模式中加入1端实体关系模式的关键字和联系类型的属性。

●若联系方式是多对多的(M:N),则将联系也转换成关系模式,其属性为互为联系的两个实体的关键字和联系的属性,图1-5中的E-R模型中学生实体和课程实体之间的联系是多对多的,则将联系也转换为关系模型,其属性为学生实体的关键字(学号)和课程实体的关键字(课程号)以及联系的属性(成绩),联系转换成的关系模式应为:

S-C模式(学号,课程号,成绩),学号+课程号,是该关系模式的关键字。

例1 图1-8是学校职工管理系统的E-R模型图,将其转换成关系模型。

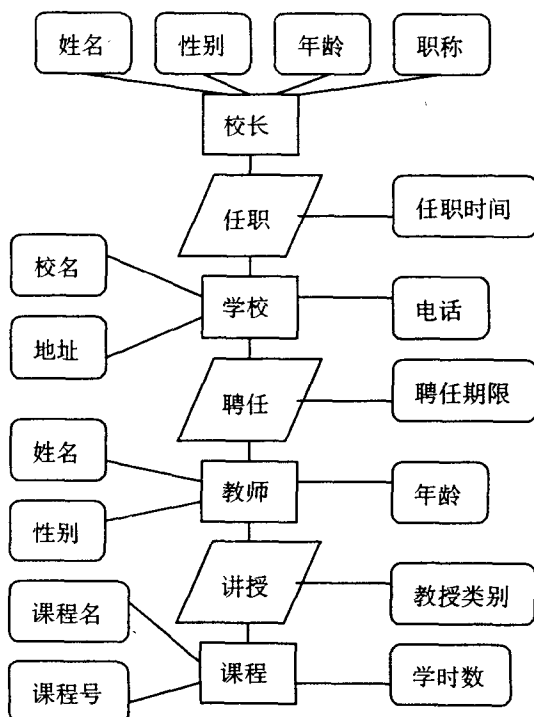


图 1-8 职工管理系统的 E-R 模型图