



全国高职高专规划教材·计算机系列

SHUJUKU JISHU JI YINGYONG XIANGMU JIAOCHENG

数据库技术及应用项目教程

(SQL Server 2008版)

李超燕◎主编



北京大学出版社
PEKING UNIVERSITY PRESS



全国高职高专规划教材·计算机系列

数据库技术及应用项目教程

(SQL Server 2008 版)

主 编 李超燕

副主编 周建良 王先花 高伟聪



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本教材获浙江省宁波市服务型重点专业群——软件与服务外包特色教材建设项目资金支持，是该项目建设的研究成果，本教材以项目驱动教学的方式来组织，符合目前高校组织教学项目驱动及“做中学”的原则。在本书编写过程中主要涉及的项目有：职工信息数据库，医疗垃圾处理数据库，学生课程数据库。这3个项目分别作为教材的课程教学、实训、课后练习编写的主线项目，其中的医疗垃圾处理数据库是与企业合作过程中由企业提供的实际开发的项目中的数据库。

本教材的第1章是数据库系统基础，让用户来了解数据库的一些基本知识。从第2章至第9章以项目驱动方式分别训练用户创建数据库、查询数据库、更新数据库、创建视图、创建存储过程和触发器、对安全性和完整性进行管理、开发数据库系统、按照数据库设计理论来完成数据库的设计这8大能力。

本教材非常适合高等院校本专科计算机及相关专业用于数据库课程的教学，也可用于计算机软件、数据库维护管理工作工程技术人员作为参考。

图书在版编目(CIP)数据

数据库技术及应用项目教程：SQL Server 2008 版/李超燕主编. —北京：北京大学出版社，2013.1

(全国高职高专规划教材·计算机系列)

ISBN 978-7-301-21901-0

I. ①数… II. ①李… III. ①关系数据库系统—高等职业教育—教材
IV. ①TP311.138

中国版本图书馆 CIP 数据核字 (2013) 第 002511 号

书 名：数据库技术及应用项目教程 (SQL Server 2008 版)

著作责任者：李超燕 主编

策 划 编 辑：胡伟晔

责 任 编 辑：胡伟晔 王慧馨

标 准 书 号：ISBN 978-7-301-21901-0/TP·1269

出 版 者：北京大学出版社

地 址：北京市海淀区成府路 205 号 100871

网 址：<http://www.pup.cn> 新浪官方微博：@北京大学出版社

电 话：邮购部 62752015 发行部 62750672 编辑部 62765126 出版部 62754962

电 子 信 箱：zyjy@pup.cn

印 刷 者：三河市博文印刷厂

发 行 者：北京大学出版社

经 销 者：新华书店

787 毫米×1092 毫米 16 开本 14.75 印张 344 千字

2013 年 1 月第 1 版 2013 年 1 月第 1 次印刷

定 价：30.00 元

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究

举报电话：(010) 62752024 电子信箱：fd@pup.pku.edu.cn

前言

Microsoft SQL Server 2008 是 Microsoft 公司推出的新一代数据管理与数据分析软件,在许多功能上比以前的版本有了很大的改进,提供了一个广泛的功能集合,扩展了可靠性、可用性、可编程性和易用性。因此,本教材所使用的就是基于 SQL Server 2008 的数据库管理系统。

本教材的教学内容经过了严格的筛选,以项目为主线来组织教材内容,符合项目驱动教学和“做中学”的原则。在本书编写过程中主要涉及的项目有:职工信息数据库,医疗垃圾处理数据库,学生课程数据库。这 3 个项目分别作为教材的课程教学、实训、课后练习编写的主线项目,其中的医疗垃圾处理数据库是与企业合作过程中由企业提供的实际开发的项目中的数据库。

本教材图文并茂,共分为 9 章,第 1 章介绍了数据库系统基础,第 2 章介绍了如何创建职工信息数据库,第 3 章介绍了如何查询职工信息数据库,第 4 章介绍了如何更新职工信息数据库,第 5 章介绍了如何为职工信息数据库创建视图,第 6 章介绍了如何为职工信息数据库创建存储过程和触发器,第 7 章介绍了如何对职工信息数据库的安全性和完整性进行管理,第 8 章介绍了如何结合 VB 完成职工信息管理系统的开发。在第 2 章至第 8 章每章都有实训,是对医疗垃圾处理数据库进行相关的操作,强化训练用户的相关技能。第 9 章介绍了如何按照数据库设计理论来完成光盘出租管理数据库的设计。

“数据库技术及应用”课程目前被列为浙江省宁波职业技术学院的院级精品课程。进入宁波职业技术学院的主页 (www.nbptweb.net),在首页的校内站点导航中进入网络课堂,在“精品课程中心”中搜索“数据库技术及应用”可以进入本课程。在本课程的精品课程中心可以下载到与课程相关的电子资源,主要包括课程的整体设计、单元设计、教材配套课件、实训项目指导、教材所涉及的数据库文件、视频教学等。

宁波职业技术学院是首批国家级示范高职院校,本校的课程教学设计也是一直走在全国高职院校的前列,本教材配套的课程整体设计、单元设计符合项目化教学的相关要求,可以作为全国高职院校计算机相关专业课程教学设计的参考文档。

本教材适合高等院校本专科计算机及相关专业用于数据库课程的教学,也可供计算机软件、数据库维护管理工作工程技术人员参考之用。

由于时间仓促,水平有限,书中不足之处在所难免,恳请广大用户提出宝贵建议与意见。若教师在上课过程中需要与本教材相关的实训和课后练习的参考答案,请与本教材的主编联系,在索要相关资料时需提供教师的相关信息,主编的 QQ 邮箱为 84405099@qq.com,并请抄送 huweiy73@sina.com。

李超燕

2012 年 5 月 9 日于宁波

目 录

第 1 章 数据库系统基础	(1)
1.1 数据库技术的产生与发展	(1)
1.1.1 数据管理的 3 个阶段	(1)
1.1.2 数据库技术的特点	(4)
1.1.3 数据库中的基本概念	(5)
1.2 数据库模型	(6)
1.2.1 概念模型	(6)
1.2.2 数据模型	(9)
1.2.3 常用数据模型	(10)
1.3 关系数据库	(12)
1.3.1 关系的数学定义	(12)
1.3.2 关系代数	(14)
1.4 数据库系统的组成	(22)
1.5 本章习题	(23)
第 2 章 创建职工信息数据库	(25)
2.1 安装 SQL Server 2008	(25)
2.2 SQL Server 2008 的常用工具	(30)
2.2.1 SQL Server Management Studio	(30)
2.2.2 SQL Server 配置管理器	(33)
2.2.3 联机丛书	(34)
2.3 创建职工信息数据库	(35)
2.3.1 创建数据库和表	(35)
2.3.2 创建关系	(41)
2.3.3 创建索引	(42)
2.4 分离和附加职工信息数据库	(45)
2.5 本章实训:创建医疗垃圾处理数据库	(46)
2.6 本章习题	(49)
第 3 章 查询职工信息数据库	(50)
3.1 基本查询	(50)
3.1.1 最简单的查询	(50)
3.1.2 条件查询	(51)



3.1.3	聚集函数	(55)
3.1.4	top 和 distinct 关键字	(57)
3.2	附加子句	(58)
3.2.1	order by 子句	(58)
3.2.2	group by 子句	(59)
3.2.3	having 子句	(60)
3.2.4	into 子句	(62)
3.3	多表连接查询	(63)
3.3.1	内部连接	(63)
3.3.2	外部连接	(66)
3.3.3	交叉连接	(69)
3.4	操作结果集	(71)
3.4.1	使用 union 组合结果集	(71)
3.4.2	使用公用表表达式	(73)
3.4.3	将结果集保存到表中	(73)
3.5	子查询	(74)
3.5.1	含 in 谓词的子查询	(74)
3.5.2	带有比较运算符的子查询	(76)
3.5.3	使用子查询代替表达式	(79)
3.5.4	带有 exists 谓词的子查询	(80)
3.6	本章实训:查询医疗垃圾处理数据库	(81)
3.7	本章习题	(87)
第 4 章	更新职工信息数据库	(89)
4.1	插入数据	(89)
4.1.1	insert... values 语句	(89)
4.1.2	insert... select 语句	(91)
4.2	修改数据	(92)
4.3	删除数据	(94)
4.4	本章实训:更新医疗垃圾处理数据库	(96)
4.5	本章习题	(98)
第 5 章	为职工信息数据库创建视图	(100)
5.1	创建视图	(100)
5.1.1	利用鼠标方式创建视图	(100)
5.1.2	利用 SQL 语句创建视图	(102)
5.2	管理视图中的数据	(104)
5.2.1	修改视图中的数据	(104)
5.2.2	插入视图中的数据	(107)
5.2.3	删除视图中的数据	(108)
5.3	删除视图	(109)

5.4 本章实训:为医疗垃圾处理数据库创建视图	(110)
5.5 本章习题	(113)
第6章 为职工信息数据库创建存储过程和触发器	(115)
6.1 程序设计基础	(115)
6.1.1 程序设计元素	(115)
6.1.2 流程控制语句	(121)
6.1.3 函数	(123)
6.1.4 游标	(129)
6.1.5 事务	(132)
6.2 存储过程	(134)
6.2.1 存储过程和种类	(135)
6.2.2 创建并调用用户存储过程	(135)
6.3 触发器	(139)
6.3.1 触发器的工作原理	(139)
6.3.2 创建并应用触发器	(139)
6.4 本章实训:为医疗垃圾处理数据库创建存储过程和触发器	(143)
6.5 本章习题	(146)
第7章 职工信息数据库的安全性和完整性管理	(149)
7.1 权限管理	(149)
7.1.1 创建用户	(149)
7.1.2 权限管理	(154)
7.1.3 数据库角色管理	(159)
7.1.4 架构安全管理	(160)
7.2 备份和还原数据库	(162)
7.2.1 创建备份设备	(163)
7.2.2 备份数据库	(164)
7.2.3 还原数据库	(168)
7.2.4 备份数据库的自动执行	(172)
7.3 完整性管理	(174)
7.3.1 约束	(175)
7.3.2 规则	(177)
7.4 本章实训:为医疗垃圾处理数据库设置安全性和完整性	(178)
7.5 本章习题	(179)
第8章 结合 VB 完成职工信息管理系统开发	(182)
8.1 系统功能分析	(182)
8.2 职工信息管理系统开发	(183)
8.2.1 口令验证模块开发	(183)
8.2.2 主界面模块开发	(185)
8.2.3 查询职工信息模块	(186)



8.2.4 插入职工信息模块	(189)
8.2.5 更新部门信息模块	(191)
8.2.6 删除职工信息模块	(193)
8.2.7 按职工统计工资信息	(195)
8.2.8 利用 VB 调用存储过程	(197)
8.3 本章实训:结合 VB 完成医疗垃圾处理系统开发	(200)
第9章 按照数据库设计理论来完成光盘出租管理数据库的设计	(201)
9.1 对光盘出租管理进行需求分析	(201)
9.2 对光盘出租管理数据库进行设计	(205)
9.2.1 概念结构设计	(205)
9.2.2 逻辑结构设计	(206)
9.3 数据异常	(207)
9.4 函数依赖的基本概念	(208)
9.5 关系的规范化	(210)
9.5.1 第一范式	(210)
9.5.2 第二范式	(210)
9.5.3 第三范式	(211)
9.5.4 BC 范式	(212)
9.6 创建光盘出租管理数据库	(212)
9.7 本章实训:按照数据库设计理论完成学生管理系统数据库的设计	(214)
9.8 本章习题	(214)
附录 A 职工信息数据库(factory)下的表结构和数据	(217)
附录 B 医疗垃圾处理数据库(medical)下的表结构和数据	(220)
附录 C 学生课程数据库(stu)下的表结构和数据	(223)
参考文献	(225)



第1章 数据库系统基础

数据库技术产生于 20 世纪 60 年代中后期,经历了近 50 年的发展,使它在理论上不断得到创新的同时也在应用上渗透到了计算机应用的各个方面,例如工农业生产、商业、行政、科学研究、工程技术和国防军事的各个部门。管理信息系统、办公自动化系统、决策支持系统等都是使用了数据库管理系统或数据库技术的计算机应用系统。现在,数据库技术已成为计算机领域中最重要技术之一,它是软件学科中一个独立的分支。

本章主要介绍与数据库技术有关的概念和术语。通过学习,用户可以初步掌握数据库的基本概念、数据库中常用的术语、数据库系统的组成等,使用户对数据库系统有一个总体上的认识,进而对数据库有宏观的理解和把握。通过对数据库的概念、术语及系统组成的学习,有助于用户准确定位自己在将来工作中应承担的角色,有利于开展有目的的自主学习。

1.1 数据库技术的产生与发展

计算机的数据处理应用,首先要把大量的信息以数据形式存放在存储器中。存储器的容量、存储的速率直接影响到数据管理技术的发展。1956 年生产的第一张磁盘容量仅为 5M 字节,而现在已达百吉(G)字节。目前光盘已广泛使用,并且容量将越来越大。磁盘是一种直接访问的存储设备,为数据库技术提供了良好的物质基础。

使用计算机以后,数据处理的速度和规模无论是相对于手工方式还是机械方式都是无可比拟的,随着数据处理量的增长,数据管理技术产生了。数据管理技术的发展与硬件(主要是外部存储器)、软件、计算机应用的范围有密切联系。数据管理技术的发展大致经过以下 3 个阶段:人工管理阶段文件系统阶段和数据库阶段。

1.1.1 数据管理的 3 个阶段

1. 人工管理阶段

人工管理阶段一般指 20 世纪 50 年代,当时计算机主要用于科学计算,这个阶段在软件上既没有操作系统也没有管理软件;在硬件上,外存只有纸带,没有磁盘等直接存取的存储设备;数据处理的方式是批处理。该阶段数据处理的主要特点是:

(1) 数据不保存。因为计算机主要用于科学计算,对于数据保存的需求尚不迫切。一般的做法是,在进行某一计算时将原始数据随程序一起输入主存,运算处理后将结果数据输出。随着计算任务的完成,数据空间随同程序空间一起被释放。



(2) 没有专用的软件对数据进行管理。每个应用程序都要包括存储结构、存取方法、输入/输出方式等, 这些完全由程序员编程实现, 从而要求程序员必须对各种数据存储结构的性能、实现和维护有较深的了解。程序中的存取子程序随着存储结构的改变而改变, 因而数据与程序不具有独立性。一旦存储结构改变, 就必须修改程序。

(3) 数据是面向应用的。即一组数据对应于一个程序, 即使两个不同应用程序涉及相同数据, 也必须各自定义, 无法互相利用、互相参照。

数据管理由各程序员在程序中进行, 程序员必须考虑数据的逻辑定义、物理存储方式及地址分配, 通过物理地址来存取数据。程序和被处理的数据紧密结合成一个整体, 数据和程序间无独立性, 如图 1-1 所示。

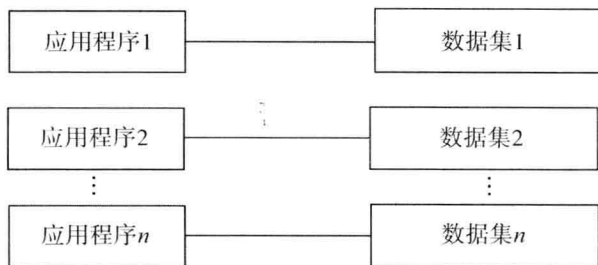


图 1-1 人工管理阶段

2. 文件系统阶段

文件系统阶段一般指 20 世纪 60 年代, 当时计算机不仅用于科学计算, 而且大量用于管理, 在软件上出现了操作系统, 并且数据管理属于操作系统的一部分, 出现了文件管理系统, 由操作系统负责对数据按照文件管理方式进行管理。在硬件上出现了直接存储的存储器, 数据可重复使用。

在文件系统中, 数据按其内容、结构和用途组成若干命名的文件, 文件一般为某用户或用户组所有, 但只可供指定的用户共享, 如图 1-2 所示。

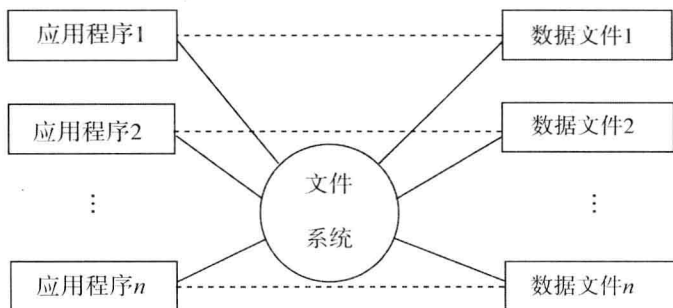


图 1-2 文件系统阶段

例如某学校的学生处、教务处、宿管中心分别要对学生的信息进行管理, 由于它们各自处理的信息不同, 因此若用文件系统实现, 则各部门的数据组织如下。

学生处要处理的信息有: 学号, 姓名, 性别, 系名, 年级, 专业, 年龄, 政治面貌, 生源地, 家庭住址等。

教务处要处理的信息有: 学号, 姓名, 性别, 系名, 年级, 专业, 课程名, 成绩,

学分等。

宿管中心要处理的信息有：学号，姓名，性别，系名，年级，专业，宿舍号，床位号，宿舍电话，寝室长等。

在该阶段，各个部门的应用程序分别要拥有自己的数据，即一个文件，但在这三个文件中，重复的数据项非常多，因此造成了比较严重的数据冗余，从该例中我们可以看出文件系统阶段存在着以下的问题：

(1) 数据冗余浪费了大量的存储空间，并且带来了潜在的不一致性。该阶段数据基本上还是面向应用的，当不同的应用使用部分相同的数据时，需要分别存储，而不能共享，因此造成了很大的冗余，并且也给数据的修改带来了潜在的不一致性。比如说，有一个学生从计算机系转到了英语系，由于某种原因可能没有通知到宿管中心，则会造成该学生在宿管中心的记录中还是计算机系，在数据处理中必然出现了错误。再比如，某个学生由于某种原因而退学了，如果没有通知到宿管中心，则他的床位将会被一直占着直到该生毕业，必然造成了学校资源的浪费。

(2) 应用程序与数据结构过分依赖，系统很难扩充。在文件系统阶段，文件是为某一个应用程序服务的，因此一旦应用程序要改变，就造成了文件数据结构的改变；反之若数据结构要改变，也造成了应用程序要作相应的变动。

比如，宿管中心在实际使用中发现对学生还要增加一个信息，如学生的身体状况，则需要在第三个文件中增加一项数据，在修改文件数据结构的同时，也要改变相应的应用程序。

(3) 缺乏对数据的统一控制能力。因为文件系统对数据进行管理的能力还是比较弱，因此对数据的安全性、正确性等方面的管理能力比较差，若用户要提高安全性，则只能在应用程序中实现，这样使应用程序的编写非常麻烦。

3. 数据库系统阶段

随着数据管理规模的不断扩大，数据量急剧增加，为了提高效率，人们开始对文件系统加以扩充，研制成倒排文件系统等。在 20 世纪 60 年代末，磁盘技术取得了重要进展，大容量和快速存取的磁盘陆续进入市场，同时成本也有了很大程度的下降，这就为数据库技术的实现提供了良好的物质条件。

60 年代中期，出现的大多数系统（如 Data Base 或 Data Bank）还不能真正称为数据库系统。数据管理技术进入数据库阶段的标志是 60 年代后期的三大事件：

- (1) 1968 年，IBM 公司成功研制了层次数据管理系统。
- (2) 美国 1971 年公布的 DBTG 报告提出网络数据库系统。
- (3) IBM 公司 1970 年发表了一系列论文，奠定了关系数据库系统的理论基础。

进入 20 世纪 70 年代后数据库技术又有了很大的发展。其发展主要表现在以下 3 个方面：

(1) 出现了许多商品化的数据库管理系统。这些计算机软件大都是基于网状模型和层次模型的数据库方法，DBTG 方法及思想对各种数据库系统影响很大。

(2) 数据库技术成为实现和优化信息系统的基本技术。商用的数据库管理系统的推出和运行使数据库技术日益广泛地应用到企业管理、交通运输、情报检索、军事指挥、政府管理和辅助决策等各个方面，并深入到人类生产和生活的各个领域。

(3) 关系方法的理论研究和软件系统的研制取得了很大成果。1974—1979 年间, IBM 公司 San Jose 研究实验室在 IBM 370 系列机上研究关系数据库实验系统 System R 获得了成功, 1981 年 IBM 公司又宣布了具有 System R 特征的新型数据库软件产品 SQL/DS 问世。与此同时, 美国加州大学伯克利分校也研制了 INGRES 关系数据库实验系统, 并紧跟着推出了商用 INGRES 系统。这些成果, 使关系方法从实验室走向了社会。

在计算机领域中, 有人把 20 世纪 70—80 年代称为数据库时代。20 世纪 80 年代, 几乎所有新开发的系统均是关系系统。同时, 微型机的关系数据库管理系统也越来越丰富, 性能越来越好, 功能越来越强, 它的应用遍及各个领域。

1.1.2 数据库技术的特点

1. 数据的共享性

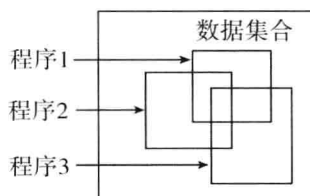


图 1-3 数据的共享

数据库系统从整体角度来看待和描述数据, 数据不再面向某个应用而是面向整个系统。数据共享是数据库系统的目的, 也是它的重要特点。一个数据库内的数据, 不仅可以为同一企业或组织内部的各个部门所共享, 也可以为不同组织、不同地区甚至不同国家的用户所共享。各个用户可以在相同时间使用同一数据库, 每个用户使用其中的一部分数据可以互相交叉和重叠, 如图 1-3 和图 1-4 所示。

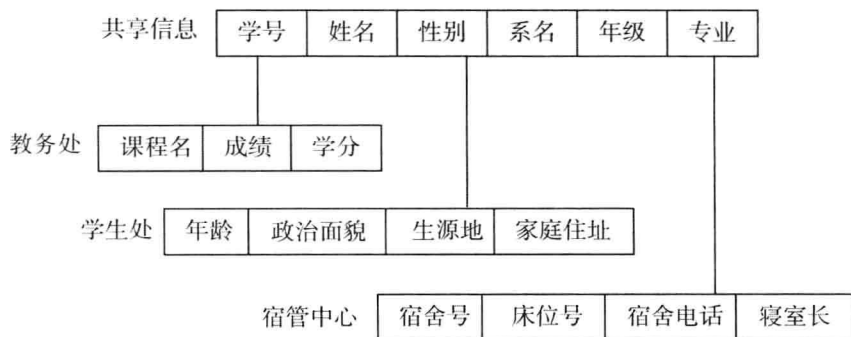


图 1-4 数据共享范例

2. 数据和程序的独立性

在数据库系统中, 系统提供映像的功能。确保应用程序对数据结构和存取方法有较高的独立性。一个数据库管理系统存在的理由就是为了在数据组织和用户的应用之间提供某种程度的独立性。数据的独立性可以分为两级:

(1) 物理独立性。即数据的物理结构, 如外存设备、存储结构、存取方法等的改变不会影响到逻辑结构, 因此也不需要改变应用程序。

(2) 逻辑独立性。即数据库的逻辑结构发生, 如定义的修改、新数据类型的加入、逻辑联系的改变等不影响原有程序数据库的存取。迄今为止, 这一独立性还不能彻底实现。

3. 数据由 DBMS 集中管理

由于数据库中的数据被不同的用户所共享，因此会引起并发操作带来的很多问题，这需要由 DBMS 集中统一控制，如图 1-5 所示。DBMS 主要提供数据完整性控制、数据安全性控制和并发控制，具体内容我们将在后面的章节中介绍。

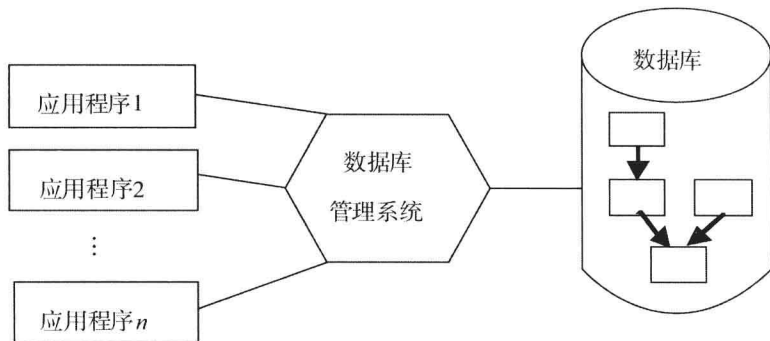


图 1-5 数据的集中管理

1.1.3 数据库中的基本概念

1. 数据 (Data)

数据实际上是描述事物的符号记录。在日常生活中人们直接用自然语言（如汉语）描述事物。在计算机中，为了存储和处理这些事物，就要抽象出这些事物的特征组成一个记录来描述。例如，在学生档案中，如果人们最感兴趣的是学生的姓名、性别、出生年月、籍贯、系名、入学时间，那么可以这样描述：

（李光，男，1984，浙江，英语系，2004）

数据与语义是不可分的。对于上面这条学生记录，了解其语义的人会得到如下信息——李光是个大学生，男，1984 年出生，浙江人，2004 年考入英语系；而不了解语义的人则无法理解其含义。可见，数据的形式本身并不能完全表达其内容，需要经过语义解释。

2. 数据库 (Database, DB)

用简明的话描述数据库时可以把数据库定义为“存放数据的仓库”，但仓库中的数据应该是有联系的，且是按照规定数据结构来组织并有专人管理的。因此对于数据库的较为全面的定义为：

数据库是长期存储在计算机中的有组织、可共享的数据集合。这些数据是现实世界中的一些相关信息，它们为特定的应用服务。

3. 数据库管理系统 (Database Management System, DBMS)

数据库管理系统是一个以统一的方式管理、维护数据库中数据的一系列软件的集合。用户一般不能直接加工或使用数据库中的数据，而必须通过数据库管理系统。DBMS 的主要功能是维持数据库系统的正常活动，接受并响应用户对数据库的一切访问要求，包括建立及删除数据文件，检索、统计、修改和组织数据库中的数据及为用户提供对数据库的维护手段等。



4. 数据库系统

数据库系统是指在计算机系统中引入数据库后的系统构成。一般由数据库、数据库管理系统、计算机软、硬件以及数据库管理员和用户等组成。

1.2 数据库模型

数据模型是数据库系统的核心和基础，在各种型号的计算机上实现的计算机管理系统都是基于某种数据模型的。

在现实生活中，模型的例子随处可见，一张地图、一座楼的设计图都是具体的模型。这些模型都能很容易使人联想到现实生活中的事物。人们对数据库的理论和实践进行研究的基础上提出了各种模型。由于计算机不能直接处理现实世界中的具体事物，所以人们必须把具体事物转换成计算机能够处理的数据。

数据库系统的主要功能是处理和表示对象和对象之间的联系。这种联系用模型表示就是数据库模型，它是人们对现实世界的认识和理解，也是对客观现实的近似描述。数据库模型更多地强调数据库的框架和数据结构形式，而不关心具体数据。不同的数据库模型实际上是提供模型化数据和信息的不同工具，根据模型应用的不同目的，可以将这些数据库模型分为两类，它们分别属于不同的层次。第一类模型是概念模型，它是按用户的观点来对数据和信息建模，主要用于数据库设计。第二类模型的数据模型，主要包括网状模型、层次模型、关系模型等。它是按计算机系统的观点对数据建模，主要用于数据库管理系统的实现。

1.2.1 概念模型

为了把现实世界中的具体事物抽象、组织为某一个数据库管理系统所能识别的数据模型，人们首先需要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，首先要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，而是概念级的模型。因此，概念模型实际上是现实世界到机器世界的一个中间层次，设计时要与用户密切合作，力求建立一个正确反映客观事实的概念模型。

概念模型有很多种，其中最为流行的一种是实体-联系模型（Entity-Relationship Model, E-R 模型），由美籍华人陈平山于 1976 年提出。人们一般用 E-R 方法建立概念模型。它的基本语义单位是实体与联系，下面介绍它的一些主要概念。

1. 概念模型中的基本概念

(1) 实体。客观存在并可相互区别的事物称为实体，是任何一种我们所关心的“事物”，可以指人，也可以指物，可以是实际的东西，也可以是抽象的、概念性的东西。例如：学生、桌子、系等。实体分为两级，一级为“个体”，如“张三”、“李四”等；另一级为“总体”，泛指某一类个体组成的集合，如“人”。

(2) 属性。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描述，例如，学生实体可以由学号、姓名、性别、出生年月等属性组成。这些属性组合起来



可以表示某一个学生。

(3) 联系。联系一般是指实体相互之间关系的抽象表示,亦即现实世界中事物之间的语义关系。例如:“教师”教导“学生”,“系”属于“学校”,“工人”生产“产品”,“学生”选修“课程”等。实体之间的联系可以分为3类。

- 一对一联系 (1:1)。两个实体集 A、B 中的每一个实体最多只能和另一个实体集中的一个实体有联系,则 A、B 之间存在着一对一的联系。现实生活中存在很多一对一联系,例如:观众与座位、乘客与车票、病人与床位等。
- 一对多联系 (1:n)。实体集 A 中至少有一个实体对应于实体集 B 中多于一个的实体,则称 A 对 B 是一对多联系。在现实生活中一对多的例子也很多,例如:部门与职员、班级和学生、国家与城市等。
- 多对多联系 (m:n)。两个实体集 A、B 中的每一个实体都和另一个实体集中任意多个实体有联系,则称这两个实体集是多对多的联系。在现实生活中的多对多的例子有:学生和课程,一个学生可以选修多门课程,一门课程可以被多个学生选修;商店与商品,一个商店可以出售多种商品,一种商品可以在多个商店出售;图书与用户,等等。

实体间的联系可以归纳为以上3类,但两个实体之间的联系到底是属于哪一种,不能光从名字上来判断,需要从实际情况出发来判断。比如,班主任和班级,如果有学校规定一个班主任只能带一个班,那么是一对一的联系,若一个班主任可以带多个班,则这两个实体之间的联系就成为一对多的联系了。

(4) 码。码又称为键或关键字,是唯一能使实体集中的每一个实体的属性或属性组。例如:学生的学号,工人的职工号,选课表的学号、课程号等。在此,大家要注意,码并不一定只有一个属性,很可能是由多个属性组组成的。

(5) 域。域为属性所取的值的范围。即一个实体集中各实体的同一属性具有的值在一定范畴之内,这一范畴称为该属性的值域,简称为域。一个属性的值域可以是整数、实数、字符串等,如“人”这个实体集的“姓名”属性的值域是字符串,“年龄”的值域是整数,“性别”的值域为“男、女”等。通常属性是个变量,属性值是变量所取得的值,而域是变量取值的集合。

2. 概念模型表示方法

E-R 图中主要用3种符号来表示实体之间的联系,下面进行介绍。

- 矩形。表示一个实体型,在矩形框内写上实体名。
- 椭圆形。表示实体的属性,在椭圆形内写上属性的名称,用无向的线条将它与相对应的实体连接起来。
- 菱形。表示联系,在菱形框内写上联系名,用无向边分别与有关实体联系起来,并且在边上标明该实体之间的联系类型 (1:1, 1:n 或 m:n)。

在此要注意,联系本身也是一种实体,也可以有属性。如果一个联系具有属性,也应该在椭圆形内写上属性名,然后用无向边与联系连接起来。

再例如用 E-R 图来表示以下的概念模型。

观众与座位在特定的影院和时间,它们之间的联系是一对一的,如图 1-6 (a) 所示。

国家与城市之间的联系是一对多的，如图 1-6 (b) 所示。

学生与课程之间的联系是多对多的，并且该联系有一个属性，即成绩，如图 1-6 (c) 所示。

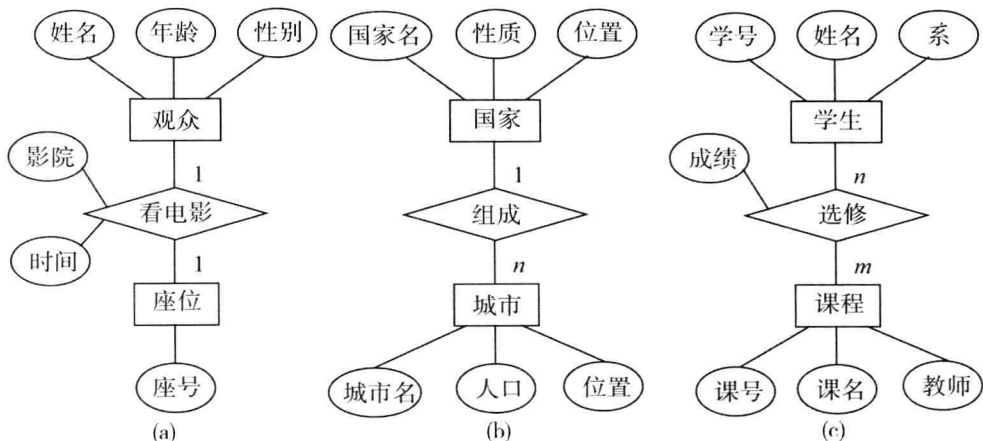


图 1-6 3 种类型联系的 E-R 图

再例如用 E-R 图来表示商店销售管理的概念模型。

图 1-7 是商店销售管理的 E-R 图，这是一个比以上几个例子都复杂的 E-R 图。但从该图中可以非常清楚地看到一个商店销售过程实体与实体之间的联系。如果为了使图不至于太复杂，可以将实体的属性另外表示，在此只给出实体、实体间的联系，以及联系的属性。

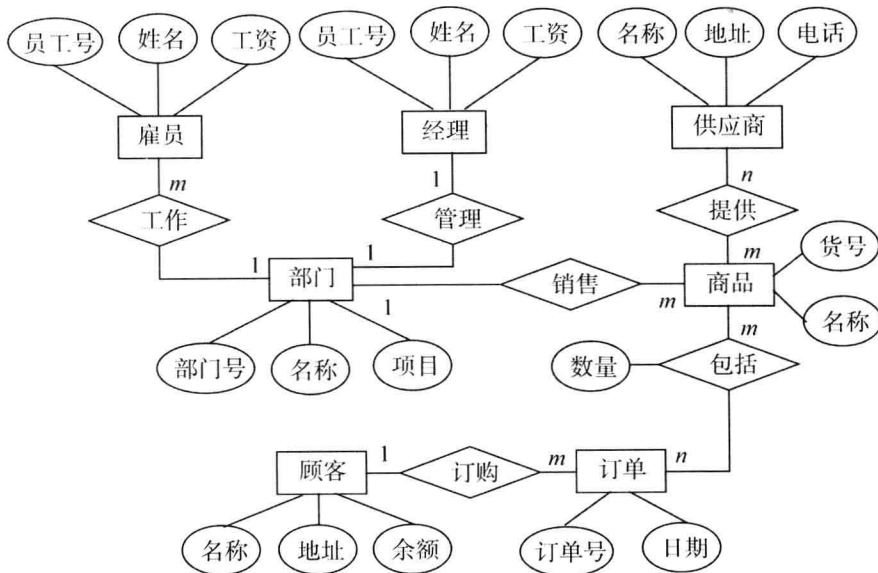


图 1-7 销售的 E-R 图

商品销售 E-R 图涉及的实体有以下几种。

- 雇员，属性有：员工号，姓名，工资。
- 经理，属性的组成同雇员，因为经理本身也是员工的一员。
- 部门，属性有：部门号，名称，项目。
- 供应商，属性有：名称，地址，电话。
- 商品，属性有：货号，名称。
- 订单，属性有：订单号，日期。
- 顾客，属性有：名称，地址，余额。

这些实体之间的联系如下。

- (1) 一个供应商可以供应多个商品，一种商品可以从多个供应商中购买；各供应商的价格不同，因此供应商和商品是多对多的关系；
- (2) 每个部门有一个经理和若干雇员，每个雇员只属于一个部门，因此部门和经理是一对一的关系，部门和雇员是一对多的关系；
- (3) 每个部门负责销售某些商品，每种商品规定只由一个部门销售，因此部门和商品是一对多的关系；
- (4) 一个订单中可以有多个商品，当然一个商品可以被包含在多个订单中，因此，商品和订单是多对多的关系；
- (5) 顾客去买东西开订单买商品，由商店送货上门，每个顾客可以开多个订单，因此，顾客和订单是一对多的关系。

1.2.2 数据模型

数据模型是一种形式化描述数据、数据之间联系以及有关语义约束的方法，是数据库系统中用于提供信息表示和操作手段的形式框架。一般来讲，任何一种数据模型都是严格定义的概念的集合。这些概念必须能够精确地描述系统的静态特性、动态特性和完整性约束条件。因此数据模型通常都是由数据结构、数据操作和完整性约束 3 个要素组成。

1. 数据结构

数据结构用于描述系统的静态特性。

它是数据模型的最基本成分，它规定如何把基本数据项组织成大的数据单位，并通过这种结构来表达数据项之间的联系。由于数据模型是向用户提供的，因此所规定的基本数据结构类型应该是简单的、基本的、易于为用户理解的；另外，这种基本数据结构类型还必须要有足够强的表达能力，即由这些基本数据结构类型可以有效地表达数据之间各种复杂的语义关系。

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，层次结构、网状结构和关系结构的数据模型分别命名为层次模型、网状模型和关系模型。

2. 数据操作

数据操作用于描述系统的动态特性。