高等院校计算机专业系列教材

龙守谌 编著

数据库技术与应用

671

科学出版社

高等院校计算机专业系列教材

数据库技术与应用

龙守谌 编著

科学出版社

1995

（京）新登字 092 号

内 容 简 介

本书是在多年数据库课程教学和数据库技术科研的基础上编写而成的，系统、完整地阐述了数据库系统的有关概念、原理与方法。全书分为 10 章。第一章概要介绍数据库系统的基本概念与特点。第二章重点说明数据物理组织中常用的索引顺序方法。第三章至第五章介绍当前占主导地位的关系数据库系统，包括关系系统的一般性问题，如关系数据结构、关系系统的完整性以及关系代数和元组关系演算；关系系统实例——ORACLE 以及标准数据库语言 SQL，关系系统的分类与查询优化。第六章较深入地讲述了关系数据理论。第七章为非关系系统，简要介绍了网状模型(DBTG)系统和层次模型系统 IMS。第八章介绍数据库系统的数据保护功能。第九章讲述数据库设计的基本技术，主要是概念设计和逻辑设计。第十章介绍数据库技术的进展，包括语义数据模型和分布式数据库。

本书内容全面，概念清晰，简明扼要，可作为高等院校有关专业数据库课程的教材，也可供广大计算机工作者、工程技术人员以及其他有关人员参考。

高等院校计算机专业系列教材

数据库技术与应用

龙守谌 编著

责任编辑 王春晖

科学出版社出版

北京东黄城根北街 16 号

邮政编码：100717

北京云浩印制厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1995 年 5 月第一版 开本：787×1092 1/16

1995 年 5 月第一次印刷 印张：11 3/8

印数：1—4 000 字数：258 000

ISBN 7-03-004505-4/TP・411

定价：18.00 元

前　言

当代新技术革命的蓬勃发展，带来社会生产力新的飞跃，引起整个社会的巨大变革，电子计算机技术是新技术革命中最活跃的核心技术，在工农业生产、流通领域、国防建设和科学研究方面得到越来越广泛的应用。

党的十一届三中会全以来，我国计算机应用事业的发展相当迅速。到目前为止，全国装机量已突破数十万台，16 位和 32 位以下微型计算机已经形成产业和市场规模，举国上下在计算机科研、开发、生产、应用、经营等方面取得了优异的成绩，创造了显著的经济效益和社会效益，并在商业、城建、金融、科技、文教、卫生、公安等广阔的领域中积极开发应用计算机技术，为开拓计算机应用的新局面作出了重要贡献。实践证明，人才是计算机开发应用的中心环节，我们必须把计算机专业人才的开发与培养放在计算机应用事业的首位，要坚持不懈地抓住人才培养这个关键。

北京工业大学计算机学院（原北京计算机学院）建院 15 年来，拥有一批以计算机专家、教授为骨干的教学、科研师资队伍，培养了数千名计算机专业技术人才，在多年的教学科研工作中，积累了许多经验。承蒙科学出版社的大力支持，教师们决心将多年总结的教学经验编写出来，出版一套计算机专业的系列教材。本套教材在内容上注重科学性，工程性和实用性，具有简明清晰，通俗易懂，方便教学，易于自学等特点。可供大专院校的师生学习计算机专业知识使用，或作为计算机专业的技术人员的参考资料，也可以提供给有志于学习和使用计算机的人员入门与提高使用。出版本套教材是人才培养和开发方面的一件很有意义的工作。

这套系列教材包括《计算机导论》、《计算机组成原理》、《数字系统逻辑设计》、《微型计算机接口技术》、《数据库技术与应用》、《宏汇编语言程序设计》、《计算机操作系统原理》、《编译原理》、《计算机控制》和《数据安全与软件加密》等 10 本。自 1994 年 9 月开始上述各教材将陆续出版。

数据库技术是数据管理的最新技术，是计算机科学的一个重要分支，广泛地应用于国民经济、文化教育、科学研究、科技情报、人工智能以及计算机辅助设计等领域。目前，大型主机和中小型计算机及众多的微型计算机都配备了数据库管理系统（DBMS）。此外，由于数据库具有数据结构化、最低冗余度、较高数据独立性、易于扩充和编制应用程序等特点，所以利用数据库技术做数据管理已经十分普遍。

鉴于上述情况，近年来我国高等院校多已把数据库课程作为计算机专业的一门主课；同时，数据库理论及应用研究也十分活跃，特别是关系数据库。由于其概念清晰、模型结构简单、语言一体化、用户接口方便、可做集合处理并有坚实的数学基础，已在数据库研究与应用中占主导地位。因此，本书以介绍关系数据库为主。

全书共分为十章，前两章为绪论和数据库数据的物理组织，概述数据库系统和数据模

型的基本内容，介绍在 DBMS 中用得较多的索引顺序文件组织方式，作为学习后续章节的导引。第三章至第六章全面介绍关系方法的理论与实现，是全书的重点，也是数据库的发展方向。第七章讲述非关系系统，包括网状模型系统和层次模型系统。第八章讲解数据库系统的数据保护功能。第九章是数据库设计，以介绍概念设计和逻辑设计为主。

计算机技术和应用的发展十分迅速，从而对数据管理的要求日益提高。例如，数据库系统不仅要提供数据，而且要提供知识；计算机应用组态由单机趋于联网；随着办公自动化、辅助设计、地理信息系统等的广泛应用，出现了除文本和数值以外的多种媒体的数据等。这样，就要求模拟更为复杂的数据结构以及相应的处理。因此，我们列入了第十章中的数据库技术的新进展，介绍语义数据模型和分布式数据库。语义数据模型是面向对象的数据库和多媒体数据库的基础。

本书内容精练、深入浅出，适合于大专院校的有关专业师生教学之用，也可供从事计算机应用工作的广大技术人员参考。

由于作者水平有限，书中难免有错误和不当之处，请予批评、指正。

需要全本请访问 www.ertongbook.com

目　录

第一章　绪　论 …… 1
1.1　数据库技术与数据管理 …… 1
1.1.1　数据库系统简介 …… 1
1.1.2　数据管理的进展 …… 2
1.2　数据模型化 …… 5
1.2.1　数据模型的三个要素 …… 5
1.2.2　数据抽象 …… 5
1.2.3　概念数据模型 …… 7
1.2.4　概念模型的建模方法 …… 7
1.2.5　主要数据模型 …… 9
1.3　数据库系统的结构 …… 11
1.3.1　数据库的三级结构与三级模式 …… 12
1.3.2　两级映象与数据独立性 …… 12
1.3.3　数据库管理系统 …… 13
习题 …… 14
第二章　数据的物理组织 …… 16
2.1　文件组织概述 …… 16
2.2　顺序文件组织 …… 16
2.3　直接文件组织 …… 17
2.3.1　码到地址的转换算法 …… 17
2.3.2　冲突处理 …… 18
2.4　索引文件组织 …… 20
2.4.1　索引无序文件 …… 20
2.4.2　索引顺序文件 …… 20
2.5　B-树 …… 25
2.5.1　B-树 …… 26
2.5.2　B+树 …… 27
习题 …… 30
第三章　关系数据库 …… 32
3.1　关系模型的基本概念 …… 32
3.1.1　关系数据库的实例 …… 32
3.1.2　关系的数学定义 …… 33
3.1.3　关系的性质 …… 34

3.2 关系的完整性规则 …… 34
3.2.1 主码与外来码 …… 35
3.2.2 两个完整性规则 …… 35
3.3 关系的运算 …… 36
3.3.1 关系代数 …… 36
3.3.2 关系演算 …… 44
习题 …… 47
第四章 ORACLE 系统与 SQL …… 48
4.1 ORACLE 系统 …… 48
4.1.1 ORACLE 系统的产品 …… 48
4.1.2 ORACLE 系统的特点 …… 49
4.2 结构查询语言 SQL …… 49
4.2.1 SQL 的数据定义功能 …… 50
4.2.2 SQL 的数据操纵功能 …… 51
4.2.3 视图 …… 63
4.2.4 嵌入式 SQL …… 65
习题 …… 68
第五章 关系系统与查询优化 …… 70
5.1 关系系统 …… 70
5.1.1 关系数据子语言 …… 70
5.1.2 关系系统的定义和分类 …… 70
5.2 关系系统的查询优化 …… 71
5.2.1 查询优化的方法学 …… 72
5.2.2 查询优化的算法 …… 75
5.2.3 查询处理代价的估算 …… 76
5.2.4 索引和聚簇对查询优化的影响 …… 78
习题 …… 78
第六章 关系数据理论 …… 79
6.1 函数依赖的基本概念 …… 79
6.1.1 关系模式的存储异常与数据依赖 …… 79
6.1.2 函数依赖诸定义 …… 81
6.1.3 数据依赖的逻辑蕴涵 …… 82
6.2 函数依赖的推导公理——Armstrong 公理 …… 82
6.2.1 Armstrong 函数依赖推导公理 …… 82
6.2.2 Armstrong 公理的完备性 …… 83
6.2.3 闭包的计算 …… 84
6.3 函数依赖的等价和覆盖 …… 85
6.3.1 等价与覆盖 …… 85
6.3.2 函数依赖集的最小集 …… 85

6.4 关系模式的规范化 …… 87
6.4.1 第一范式（1NF）到第三范式（3NF） …… 87
6.4.2 Boyce-Codd 范式（BCNF） …… 89
6.4.3 多值依赖与第四范式 …… 91
6.5 关系模式的分解 …… 93
6.5.1 分解的连接无损性 …… 94
6.5.2 函数依赖集的保持 …… 97
6.5.3 模式分解的算法 …… 98
习题 …… 100
第七章 非关系系统 …… 102
7.1 网状模型系统 …… 102
7.1.1 DBTG 系统概述 …… 102
7.1.2 DBTG 数据结构 …… 102
7.1.3 网状数据模型的描述 …… 104
7.1.4 DBTG 的数据完整性 …… 110
7.1.5 DBTG 的数据操纵 …… 110
7.2 层次模型系统 …… 115
7.2.1 层次数据结构 …… 115
7.2.2 IMS 数据库的描述 …… 117
7.2.3 逻辑数据库 …… 119
7.2.4 IMS 的数据操纵 …… 121
习题 …… 123
第八章 数据库的数据保护功能 …… 125
8.1 安全性 …… 125
8.1.1 保证数据安全的一般方法 …… 125
8.1.2 ORACLE 系统中的安全措施 …… 126
8.2 完整性 …… 127
8.2.1 完整性规则 …… 128
8.2.2 完整性规则的表述 …… 128
8.2.3 ORACLE 系统的完整性措施 …… 129
8.3 并行性 …… 131
8.3.1 并发控制 …… 131
8.3.2 死锁及避免死锁的办法 …… 132
8.3.3 可串行性 …… 133
8.3.4 两段锁协议 …… 133
8.4 恢复 …… 134
8.4.1 转储 …… 134
8.4.2 日志文件 …… 134
习题 …… 135

第九章　数据库的概念设计与逻辑设计 …… 136
9.1　数据库设计的方法与阶段 …… 136
9.2　概念设计 …… 138
9.2.1　基数 …… 138
9.2.2　局部视图设计与集成 …… 139
9.3　数据库设计的功能分析 …… 147
9.3.1　用数据流模型进行功能分析 …… 147
9.3.2　概念设计与功能分析结合 …… 148
9.4　逻辑设计 …… 152
9.4.1　关系模型的逻辑设计 …… 152
9.4.2　网状模型的逻辑设计 …… 156
习题 …… 158
第十章　数据库技术进展 …… 160
10.1　语义数据模型 …… 160
10.1.1　实体-联系（E-R）模型 …… 161
10.1.2　嵌套的关系模型 …… 162
10.2　分布式数据库 …… 166
10.2.1　分布式数据库概述 …… 167
10.2.2　分布式数据库系统的特点 …… 167
10.2.3　分布式数据库系统的基本结构 …… 168
10.2.4　数据的分片 …… 169
10.2.5　分布数据的查询分解及优化 …… 171
10.2.6　分布事务管理 …… 172
习题 …… 174
参考文献 …… 174

第一章 绪 论

早期的计算机主要应用于科学计算，随着生产的发展，社会的进步，计算机应用进入数据处理时代。这是一个很大的转折，其重要标志就是出现了数据库技术，使得计算机获得更广泛的应用，成为广大科技工作者和管理人员的得力助手和工具。

数据库技术是在文件技术基础上发展起来的，它是数据管理的最新技术。这一章将从数据管理的发展过程简要地介绍数据库系统所涉及的基本概念，包括文件技术与数据库技术的特点、数据模型、数据库系统的体系结构以及数据库管理系统（DBMS）的主要功能和组成成分等，使读者对数据库技术有一个概要的了解，为后续各章的学习打下基础。

1.1 数据库技术与数据管理

1.1.1 数据库系统简介

数据库系统是计算机化的信息系统，由一种称做数据库管理系统(DBMS)的通用软件包管理数据。数据库系统由三个主要部分组成，它们是数据库、数据库管理系统(软件)和计算机硬件。下面简要介绍这每一部分。

(1)数据库(database)

数据库是由 DBMS 建立的相关数据文件的集合。数据库的内容是把来自企业内各个部门的数据组织起来形成的，所有用户都可以使用，冗余数据减至最少。

数据库中的数据通常由两大部分组成：一部分是有关应用所需的工作数据的集合，即所谓的用户数据，称做物理数据库，它是数据库的主体；另一部分是关于各级数据结构的描述数据，即所谓的系统数据，称作描述数据库，通常由一个数据字典系统管理。

(2)数据库管理系统

数据库管理系统(DBMS)是数据库系统的主要软件部件。DBMS 是大量相关软件例行程序的集合，其中每个例行程序负责一个特定的任务。DBMS 的主要功能是：

1)建立和组织数据库。

2)创立并维护数据库的存取路径，因此，可以快速存取数据库任何部分的数据。

3)应用户请求，操纵数据库中的数据。

4)维护数据的完整性和安全性。

5)为用户使用数据库实行联机。

DBMS 解释和处理用户请求，从数据库中检索信息。图 1.1 示出了 DBMS 在用户请求与数据库之间的界面作用。

在大多数情况下，查询请求在存取物理数据库之前要穿过几层的 DBMS 软件以及操作系统。DBMS 引用适当的子程序来解释查询，找到数据库中所希望的数据，并以一定顺序提取出来。

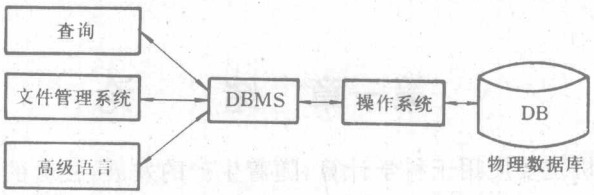


图 1.1 DBMS 的界面作用

(3)数据库系统硬件

数据库系统的硬件部分包括直接存取存储设备(DASD)、主机及其附属设备,如主存,I/O 处理器以及设备控制器。

计算机系统的两种基本存储设备是主存储器(又叫内存储器)和辅存储器(又叫外存储器)。主存的存取速度在微秒以下,主要用作程序与数据的暂时存储。辅存的存取速度要慢得多,主要因为有机电移动,大约在毫秒级,主要用作永久性数据存储。

(4)数据库管理员

在大多数企业组织中,终端用户并不直接控制数据库中的数据,而是由一个或一个小组的专家来控制与管理数据库,这个个人或小组便称作数据库管理员(DBA)。DBA 按企业的需要与全部终端用户打交道,总结他们的需求。DBA 的成员大多是 DBMS、数据库设计、主机操作系统、数据通信、计算机硬件和应用程序设计等某方面的专家。

1.1.2 数据管理的进展

数据库技术是做数据管理用的,数据管理是数据处理的核心。数据管理并非一开始就应用数据库技术,这里有一个发展过程。

数据是以一定方式记录下来的、可以加以鉴别的符号。信息是现实世界中事物的存在方式和运算状态的反映。

数据与信息的关系是非常密切的。数据是信息的符号表示或载体。信息则是数据的内涵,是对数据的语义解释。

人们在了解、掌据事物的运动规律和事物之间固有联系的基础上,从一些已知的信息出发,运用演绎推理,导出新的数据或信息,为人类社会生活的各种需要服务,就是数据或信息处理。数据处理包括数据的收集、加工、存储、检索、维护和传递等工作。

随着计算机硬件和软件的不断发展,数据管理经历了人工管理、文件系统、数据库技术三个阶段。图 1.2 示出了数据管理的发展情况。

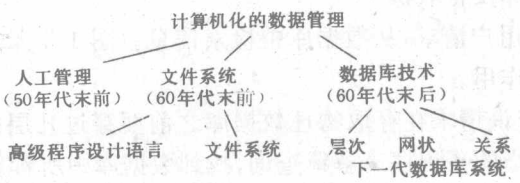


图 1.2 数据管理进展情况

需要全本请访问 www.ertongbook.com

1. 人工管理阶段

50年代末以前，计算机主要用于科学计算，少量用于数据处理。这时，对于数据保存的需求并不迫切，一般是在算题时，将原始数据与程序一起输入主存，运算处理后将结果数据输出。随着计算机任务的完成，用户作业退出计算机系统，数据空间随同程序空间一起释放。数据由人工保管，这便是人工管理阶段。

2. 文件系统阶段

50年代末以后，计算机应用普及，数据处理逐渐上升为主要应用领域。这时由于大量数据存储、检索和维护成为迫切需要，因此，数据结构设计与数据管理技术研究迅速发展，在硬件方面提供了可能性，于是便进入“文件技术”阶段。这一阶段的特点是：

1)硬件方面，外存储器成了计算机系统的重要组成部分。

2)软件方面，出现了专门用于数据管理的系统软件，对驻留于外存的数据文件实施统一管理，它构成操作系统的一个重要组成部分，称为文件系统。

这时，应用程序不再需要了解数据在存储介质上的实际存储情况，只把存储文件看成是按某种方式组织的一组数据记录的集合。从记录到存储信息的映象是通过系统提供的存取方法(access methods)进行的。

3)文件组织形式已经多样化，由于已有了直接存取存储设备，出现了索引文件、直接存取文件，既可满足批处理应用需要，又能有效地随机存取。

总的说，这时的数据定义主要是面向应用的，即数据文件为应用所私有(private)。数据文件间联系很少，数据冗余大，有潜在的不一致性。

例1.1 我们考虑一个小型商业系统，它基于下述活动组织数据：

1)销售(transaction)，建应收帐(accounts receivable)程序。

2)订货(inventory order)，建应付帐(accounts payable)程序。

3)存货(inventory)，含库存程序。

用文件技术实现这个系统的方案是，为应收帐程序建立销售文件，程序生成发票(invoices)。该文件存储顾客事务以及货物销售情况。类似地，为应付帐程序建立订货文件，生成给供应商的支票(cheques)。这个文件包含关于供应商和订货方面的信息，库存子系统中许多程序应用库存文件，生成存货报表、采购需求和修改库存。如图1.3所示。

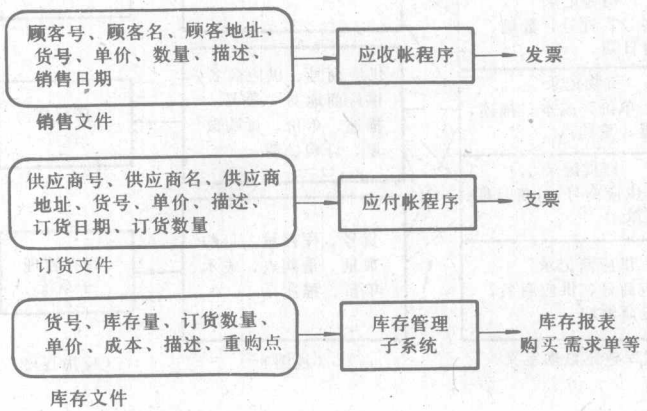


图1.3 文件系统实现小型商业系统

由此可以看出，数据文件是依特定应用建立的。应用程序是基于文件的指定物理结构和存取方法编制的。因此，除了数据冗余大而外，数据的物理结构与应用程序之间依赖性很大，独立性差。

在60年代，文件技术得到充分发展，计算机的应用被推向高潮。在许多组织中，在个别部门的应用逐渐发展到多个部门的应用；由简单应用发展为彼此相关的复杂应用。这带来了数据管理上的一些新问题，主要有：

数据冗余大。文件系统中的文件基本上对应于某个应用程序。就是说，数据是面向应用的。当不同的应用程序所需要的数据有部分相同时，也必须建立各自的文件，而不能共享相同的数据。数据冗余大，浪费存储空间。由于相同的数据的重复存储，各自管理，给数据的修改和维护带来了困难，容易造成数据的不一致性。

数据和应用程序独立性差。由于应用程序和数据结构互相依赖，一旦因某种原因导致数据存储结构改变，必须修改应用程序，修改文件结构的定义。这就使得编制应用程序，特别是维护应用程序的成本大大增加。

此外，文件系统缺乏对数据操作的控制方法，这里安全性、完整性方面的控制完全由用户应用程序负责，这使得程序的编制更为繁琐。

由于传统文件系统的这些缺点，促使人们寻求新的数据管理技术，目标包括：

1)实现不同应用对数据的共享，减少数据的重复存储，消除潜在的不一致性。

2)实现数据独立性，即应用程序独立于数据的存储结构，降低维护成本。

3)由系统软件提供数据安全性、完整性等数据控制功能。

结果，于60年代末出现了数据库技术，进入数据库技术时代。

3. 数据库技术阶段

60年代末，大容量和快速存取磁盘设备取得重大进展，为数据库技术的实现提供了良好的物质条件。数据库技术的特点如下：

(1)数据冗余小

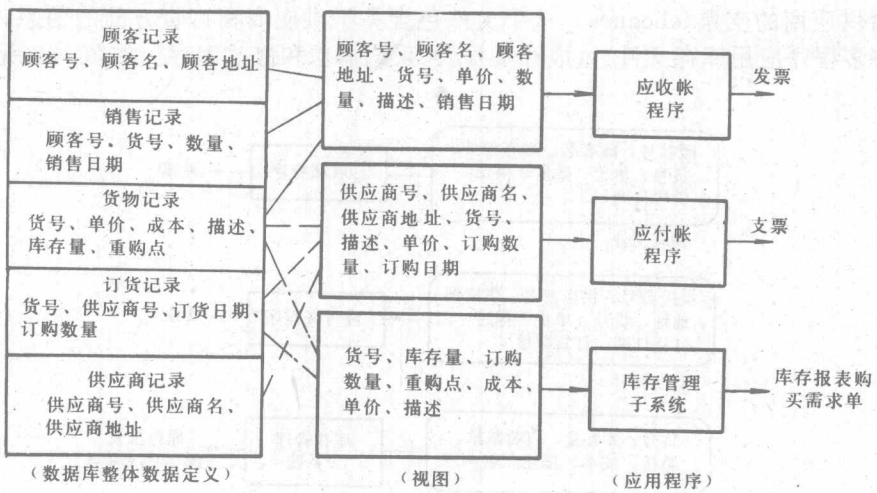


图 1.4　小型商业系统的数据库实现

我们仍用例1.1的小型商业系统说明这个问题。这里，将各个应用程序涉及到的数据项统一加以组织，然后分组成适当的记录类型。图1.4示出了小型商业系统的数据库实现风格。图的最左面是信息系统的数据模式，包括5个记录类型。这里已经去掉了一切不必要的重复存储，使数据冗余减至最小，这5个记录类型的联系如图1.5所示。

(2)数据独立性好

数据库系统提供了两方面的映象功能，一个是数据的存储结构与逻辑结构之间的映象；另一个是数据的总体逻辑结构与某类应用的局部逻辑结构之间的映象，这两级映象功能确保了应用程序对数据存储结构和存取方法有较高的数据独立性。

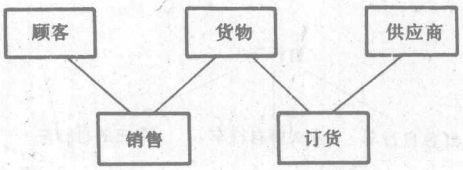


图1.5　小型商业系统的整体数据结构

(3)对数据实行统一的控制

数据库中的数据是各用户的共享资源，这个共享一般是并发的。因此，确保数据的安全与完整十分重要。数据库管理系统提供统一的数据控制功能，除确保系统数据安全而外，还大大减轻了编制应用程序的负担。

从文件系统发展到数据库系统是信息处理领域的一个重大变革。

1.2　数据模型化

通过前面一节的讨论，我们对数据库和数据库系统有了一个概要的了解，本节将进一步对数据库的数据结构作简要介绍。从用户的观点看，数据库是模拟现实世界中某应用环境活动的信息集合，它不仅反映客观事物本身的内容，而且反映它们之间的联系。这种模拟是通过数据模型来进行的。在实际的数据处理中，首先是将现实世界的事物及其联系抽象为信息世界的信息模型；然后再将信息模型抽象为计算机世界的数据模型。前者是按用户观点的抽象，后者是按计算机系统观点的抽象。

1.2.1　数据模型的三个要素

一般讲，数据模型描述系统的静态特性、动态特性和完整性约束条件。因此，数据模型通常由数据结构、数据操纵和完整性约束三部分组成：

1)数据结构是所研究的对象类型的集合。包括对事物本身的描述，以及对联系的描述。在基于记录的程序设计范畴，主要概念是记录、数据项和类型等。

2)数据操纵指对数据库中各种对象的实例允许执行的操作的集合，包括操作及有关的操作规则。数据库的操作包括检索和更新两大类。

3)完整性约束是完整性规则的集合。完整性规则用以限定符合数据模型的数据库状态及状态的变化，以保证数据的正确、有效。

1.2.2　数据抽象

数据模型是对现实世界的抽象。一般地讲，所谓抽象是对实际的人、物、事和概念的人为处理，它抽取人们关心的共同特性，忽略非本质的细节，并把这些特性用各种概念精确

地描述出来，构成某种模型。一般有三种抽象。

(1)分类(classification)

定义某一概念作为现实世界中一组对象的类型，这些对象具有某些共同的特性和行为，它抽象对象实例与类之间的“is member of”语义。例如，‘自行车’这个概念便是一个类，它的成员是全部自行车，包括红色自行车、永久牌自行车和张三的自行车等。分类抽象如图 1.6 所示。

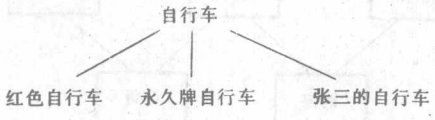


图 1.6　分类抽象



图 1.7　聚集抽象

(2)聚集(aggregation)

由表示成分的一组类定义一个新类。它抽象了对象的内部类型和成分之间的“is part of”语义。例如，由类 NAME，SEX 和 AGE 抽象为类 PERSON，是由若干属性聚集组成实体型。如图 1.7 所示。

分类和聚集是建造数据库和许多程序设计语言的数据结构的两种基本抽象。由对象实例抽象成为类，一般称作分类；而由对象类抽象为更一般的类称为聚集。例如，图 1.8 所示的情况。

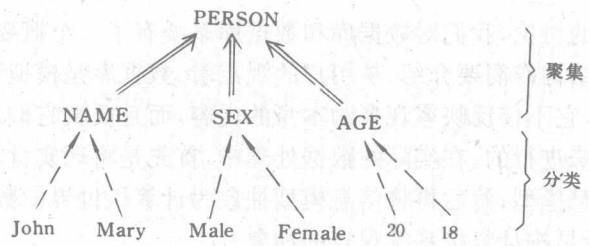


图 1.8　分类与聚集

(3)概括(generalization)

定义两个(或多个)类之间的子集联系。它抽象了类之间的“is －a”联系。例如，学生是一个类，本科生和研究生是学生的子集(类)，我们把学生称为超类(superclass)，本科生、研究生为学生的子类(subclass)。如图 1.9 所示。



图 1.9　概括抽象

概括抽象有一个重要的性质，即继承性(inheritance)。子类可以继承超类上定义的所有抽象。如本科生、研究生可继承学生类的属性。当然，子类可以有自己的某些特殊属性。

1.2.3 概念数据模型

现实世界的事物经过一次抽象进入信息世界，这是按用户观点的抽象，所得到的概念数据模型不依赖于具体的计算机系统，具有相当的适应性与稳定性。

信息世界涉及的主要概念有5种。

(1)实体(entity)

可区分的客观事物均可称为实体。实体可以是人，也可以是物；可以指实际对象，也可指某些概念；可以指事物本身，也可以指事物与事物之间的联系等。

(2)属性(attribute)

事物某方面的性质。一个实体可以由若干个属性来刻画。例如，库存管理中的零件(part)可以由零件号(pno)、零件名(pname)、重量(weight)和颜色(color)来描述。

(3)域(domain)

属性的取值范围，它是一个值集。域的形式定义为

$$D=(vi|pi)$$

其中，vi是满足谓词pi的值。如：

$$AGE=(x|x\geqslant 0 \text{ and } x\leqslant 150)$$

(4)实体的型和值

型指类型(type)，用实体名和属性名来表示。具有相同属性的实体有共同的特性，称为实体型。例如，学生(学号、姓名、性别、年龄)是一个实体型。同型实体的集合称为实体集，是一个群体的概念。值指具体的事物或特征。它是个体的概念。

(5)联系

现实世界的事物之间是相互关联的。这种联系必然反映在信息世界中。一般存在两类联系：一是实体之间的联系，二是实体内部的联系，如组成实体的属性之间的联系。在介绍实体-联系(E-R)方法一节中，介绍实体之间的联系。

1.2.4 概念模型的建模方法

建立概念模型的方法有多种。最常用的是实体-联系(entity-relationship)方法，它是美籍华人P. P. S. Chen于1976年提出的。这个方法用E-R图来描述某一组织的概念模型，具有简单直观的特点。

1. E-R图图例

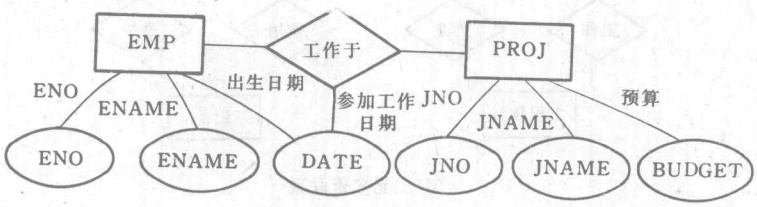


图1.10 E-R图的例子

长方形框表示实体，实体名写在其中。菱形表示联系。椭圆表示属性域。弧线表示属

性。如果属性域与属性同名，可以用属性域代属性名，以简化图形。图 1.10 给出了 E-R 图的例子。

2. 联系(relationships)

两个实体集之间的联系可按参与实体型的数目分为二元联系、多元联系与环形(一元)联系。按实体集间实体映射可分为一对一、一对多和多对多的联系。

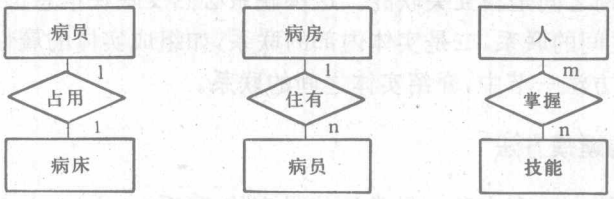
(1)二元联系

1)一对一(1:1)联系。若对于实体集 A 中的每一个实体，实体集 B 中至多有一个实体与之联系；反之亦然，则称实体集 A 与实体集 B 具有一对一联系。例如，医院中病员(patient)与病床(bed)的联系是一对一的。即一个病员只占用一张病床，一张病床只被一个病员占用。

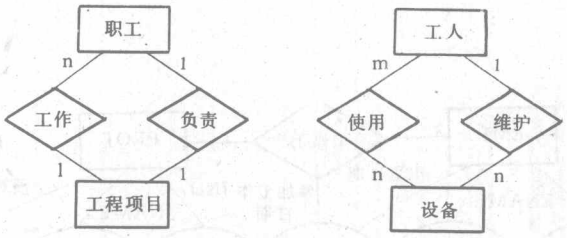
2)一对多(1:n)联系。若对于实体集 A 中的每一个实体，实体集 B 中有 n 个实体($n \geq 0$)与之联系；反之，对于实体集 B 中每一个实体，实体集 A 中最多只有一个实体与之联系，称实体集 A 与实体集 B 有一对多的联系。

例如，医院的病房(room)与病员的联系是一对多的。一个病房可住多个病员，每个病员只住一个病房。

3)多对多(m:n)联系。若对于实体集 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($m \geq 0$)实体与之联系，则称实体集 A 与实体集 B 具有多对多的联系。例如，人(person)与技能(skill)之间的联系是多对多的。一个人可以有多种技能，一种技能可为多人掌握。图 1.11 示出这三类二元联系。



(a)1:1 的联系　(b)1:n 的联系　(c)m:n 的联系



(d)二元多重联系

图 1.11　二元联系

4)两个实体集间可有多重联系，如图 1.11(d)所示。职工与工程项目之间可有两种联系。工人与设备之间有工人使用设备和工人维修设备的联系。

需要全本请访问 www.ertongbook.com