



普通高等教育“十一五”国家级规划教材

高等学校计算机科学与技术系列教材

# 高级数据库技术与应用

汤庸 叶小平 汤娜 吴凌坤 编著



高等教育出版社  
Higher Education Press

TP311.13/342

2008

普通高等教育“

材

高等学校计算机科学与技术系列教材

# 高级数据库技术与应用

汤庸 叶小平

汤娜 吴凌坤 编著

高等教育出版社

## 内容提要

数据库技术一直是发展最快的信息技术之一。本书根据现代数据库应用的网络化、智能化、多维化和协同化等特征,选择基础的和有代表性的数据库技术组织内容,使读者对数据库新技术有一个整体的了解,为进一步从事相关研究和开发提供一些思路。

全书共6章。首先回顾数据库的发展,简要介绍数据库基础理论,然后按照基于时间、知识、网络等方面介绍时态数据库、实时数据库、时空数据库、知识库、主动数据库、数据仓库与数据挖掘、分布式数据库、Web数据库、移动数据库、XML数据库等技术,最后介绍协同数据库和综合应用实例等。附录中给出了关于数据库技术的学习资源以及ATSQL语言的描述。

本书可以作为高等学校计算机专业本科高年级选修课程或研究生课程的教材,同时也可以作为广大数据库研究与开发人员的参考书籍。

## 图书在版编目(CIP)数据

高级数据库技术与应用 / 汤庸等编著. —北京: 高等教育出版社, 2008.2

ISBN 978-7-04-023222-6

I. 高… II. 汤… III. 数据库系统-高等学校-教材  
IV. TP311.13

中国版本图书馆CIP数据核字(2008)第003712号

策划编辑 刘艳 责任编辑 郭福生 封面设计 于文燕 责任绘图 吴文信  
版式设计 张岚 责任校对 朱惠芳 责任印制 毛斯璐

出版发行 高等教育出版社  
社址 北京市西城区德外大街4号  
邮政编码 100011  
总机 010-58581000

经销 蓝色畅想图书发行有限公司  
印刷 北京未来科学技术研究所  
有限责任公司印刷厂

开本 787×1092 1/16  
印张 20.25  
字数 450 000

购书热线 010-58581118  
免费咨询 800-810-0598  
网址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.landaco.com>  
<http://www.landaco.com.cn>  
畅想教育 <http://www.widedu.com>

版次 2008年2月第1版  
印次 2008年2月第1次印刷  
定价 25.40元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23222-00

# 高等学校计算机科学与技术系列教材编审委员会

主任：李未

副主任：傅育熙 王志英 齐治昌 陈平 蒋宗礼 马殿富

委员：（按姓氏笔画为序）

王 戟（国防科学技术大学）

宁 洪（国防科学技术大学）

刘 强（清华大学）

孙吉贵（吉林大学）

庄越挺（浙江大学）

何炎祥（武汉大学）

何钦铭（浙江大学）

张晨曦（同济大学）

李宣东（南京大学）

李晓明（北京大学）

陈 钟（北京大学）

陈道蓄（南京大学）

周立柱（清华大学）

周傲英（华东师范大学）

孟祥旭（山东大学）

岳丽华（中国科学技术大学）

罗军舟（东南大学）

姚淑珍（北京航空航天大学）

胡事民（清华大学）

骆 斌（南京大学）

徐宝文（东南大学）

黄虎杰（哈尔滨工业大学）

蒋建伟（上海交通大学）

廖明宏（哈尔滨工业大学）

熊 璋（北京航空航天大学）

樊晓桢（西北工业大学）

# 序

---

计算机和通信技术的迅猛发展，不仅形成了融合度最高、潜力最大、增长最快的信息产业，而且成为推动全球经济快速增长和全面变革的关键因素。进入 21 世纪，我国的信息产业虽然已取得了长足的发展，但与发达国家相比，还有不小的差距。国家信息化的发展和信息产业国际竞争能力的提高，迫切需要高素质、创新型的计算机专业人才。

高素质计算机专业人才的培养离不开高质量的计算机教育。我们的专业虽然机会多，处于非常有利的条件，但是我们同样面临着一件事，就是从规模发展向质量提高的转变。怎么提高质量？专业素质的教育和应用素质的训练非常重要。尤其是我国高等教育进入大众化发展阶段，社会对计算机专业人才呈现出了多样化的需求。而与此同时，计算机学科的发展已极大地突破了原有的学科体系框架，形成了在“计算机科学与技术”之下向多个专业方向发展的新格局。在这种背景下，教育部高等学校计算机科学与技术教学指导委员会编制了《高等学校计算机科学与技术专业发展战略研究报告暨专业规范（试行）》（以下简称“专业规范”）。专业规范按照“培养规格分类”的指导思想，提出了三种类型、四个方向，即科学型（计算机科学方向），工程型（计算机工程方向、软件工程方向），应用型（信息技术方向）的计算机专业发展建议，体现了社会对不同人才类型的需求，对于指导我国计算机教学改革与建设，规范计算机教学工作，促进计算机教学质量的提高都具有重要的意义。

高水平的教材是一流教育质量的重要保证。为了配合专业规范的试行，便于广大高等学校教师按照新的专业规范组织实施教学，高等教育出版社在大力支持专业规范研究与起草工作的同时，还邀请规范起草小组的有关专家成立“高等学校计算机科学与技术系列教材编审委员会”，组织规划了结合计算机专业规范、面向全国高等学校计算机专业本科生的“高等学校计算机科学与技术系列教材”。令人高兴的是，一批有创新、改革精神，且有丰富教学经验的高等学校教师投身到新体系计算机专业教材的编写中来，他们用自己创造性的思维、辛勤的汗水诠释专业规范的思想，把新的课程体系和教学内容生动地传达给师生，并进行着有意义的教学实践。

“高等学校计算机科学与技术系列教材”以专业规范和 CC2001 - CC2005 有关教程为依据，以强化基础、突出实践、注重创新为原则，体现了学科课程体系和教学内容改革的新成果。此外，这一系列教材还配有丰富的教学辅助资源，并与现代教育技术手段相结合，充分发挥网络平台的作用，使教材更有利于广大教师和学生使用。目前，这一系列教材有不少选题已列入普通高等教育“十一五”国家级规划教材，希望这些教材的出版能够对新形势下我国高等学校计

II ▶▶ 序

计算机专业课程改革与建设起到积极的推动作用，使我国高等学校的计算机专业教学质量再上一个台阶。



中国科学院院士  
2006—2010年教育部高等学校计算机科学与技术教学指导委员会主任  
二〇〇七年十一月

# 前言

---

数据库技术从 20 世纪 60 年代中期产生至今已有四十多年的历史，数据库技术一直是最活跃、发展速度最快的 IT 技术之一。一般来说，数据库发展可以分为 3 个阶段：第一代数据库以网状模型和层次模型为特征；第二代数据库以关系模型为特征，关系数据库在 20 世纪 70 年代末和 80 年代得到了广泛应用；20 世纪 90 年代，随着网络和多媒体技术的迅猛发展，数据库的应用也得到进一步拓展。现代数据库技术融合多种技术，数据库技术的发展进入一个新的时期。

现代数据库应用范围非常广泛，在实际数据库应用开发中难以使用单一的数据库技术实现，往往需要融合应用许多方面的相关技术。网络化、智能化、多维化和协同化已成为现代数据库应用的主要特征。现代数据库技术所涵盖的范围很广，每种具体数据库技术都可以编写成独立教材，因此本书中不可能详细介绍每种数据库技术。本书根据网络化、智能化、多维化和协同化等特征，选取介绍其中一些基本的和有代表性的数据库技术，目的是使读者可以通过这门课程的学习对数据库新技术有一个整体的认识，为进一步从事相关研究和开发提供一些思路。

全书共分 6 章。第 1 章简要回顾数据库的发展，讨论新一代数据库技术的范畴和实现途径，给出现代数据库技术及应用展望。

第 2 章简要介绍关系数据库理论，主要包括关系数据模型、关系模式设计和事务处理等关系理论基础，然后介绍面向对象数据库、对象关系数据库的基本概念和基础知识等。本章知识点是高级数据库技术的基础。

第 3 章介绍基于时间的数据库技术。时间是自然界的重要属性，也是新一代数据库的重要特征。本章主要介绍多维数据库技术，包括时间数据演算、时态数据库技术、实时数据库技术、时空数据库技术等。

第 4 章介绍基于知识的数据库技术。主要以知识库为基础，介绍主动数据库技术、数据挖掘技术、决策支持系统等需要处理知识的数据库应用技术。

第 5 章介绍基于网络的数据库技术。主要包括分布式数据库技术、Web 数据库技术、移动数据库技术、XML 数据库、Web 搜索技术等。

第 6 章介绍数据库协同管理、数据交换与信息发布、 workflow 数据库等数据库协同技术。然后介绍时态知识数据库综合应用和两个综合应用实例，在实例中综合应用了面向对象技术、时态数据库技术、主动数据库、知识库、数据库协同工作等多种数据库技术。

最后，附录 A 给出了对于学习和研究数据库新技术有帮助的一些资源，包括重要学术会议、学术期刊和相关教材的介绍，附录 B 给出了时态数据库 TempDB 中 ATSQL 语言的 BNF 描述。

本书定位为“导论”，旨在给读者对数据库新技术的学习和研究起到一个“引导”作用。

主要特色是基础技术、发展方向与综合实践相结合。

① 本书按现代数据库应用的主要特征——多维化、智能化、网络化、协同化——这 4 个方面来组织内容,分别介绍每个方面最基础的数据库技术,然后介绍几种相关的数据库新技术,使读者既掌握相关基础,又了解相关技术领域及其发展方向。

② 书中介绍了国际数据库界重要的学术会议和期刊,点评了现有的高级数据库相关著作,目的是培养读者的研究素质,引导读者了解数据库技术的新动向和新方向。这些学习资源是本书内容的延伸,这样可以保持本书内容的新颖性。

③ 注重理论研究和实践应用的结合。书中融入了我们近期数据库研究和开发成果,介绍的应用实例,涉及面向对象、时态数据库、主动数据库、知识库、协同技术等多种数据库新技术综合应用,有较好的示范性,启发读者在实际研究工作中注重多种研究成果的综合应用。

本书与我们编写的《数据库系统原理教程》、《数据库系统实验指导教程》(清华大学出版社)一同入选普通高等教育“十一五”国家级规划教材。我们在 3 部教材编写过程中,统筹安排,形成数据库课程链。前两部教材作为计算机专业必选课程教材,本书可作为高年级选修课和研究生教材。建议计划学时 40~60 学时,教学方式采用讲授与讨论相结合的方式,数据库技术部分采用讲授(32~40 学时),根据书中讨论题组织数据库文献阅读和专题讨论(8~20 学时),有条件的可以进行综合实例分析与实践。

本书由汤庸统筹,叶小平、汤娜、吴凌坤、李建国等参加编写主要内容,参与编写的还有马慧、刘海、左亚尧、陈国华、潘炎、何庆、冀高峰、朱君、郭欢等博士生。数据库课程组的老师印鉴教授、刘玉葆、万海、范昭赋博士以及“高级数据库技术”课程讨论班的研究生给予了许多支持,中国计算机学会数据库专业委员会副主任、四川大学唐常杰教授审阅了书稿并提出了宝贵意见,本书的编写还参阅了大量的著作和相关文献,在此一并表示衷心的感谢!书中若有错漏之处,敬请指正。

本书编写工作得到国家自然科学基金项目(60673135、60373081)、教育部新世纪优秀人才支持计划(NCET-04-0805)和广东省自然科学基金项目(重点 04105503、7003721、05003348)、教育部-微软精品课程(2006)等项目支持。

编者

2007 年 7 月于中山大学



# 作者简介



汤庸先后获武汉大学计算机科学与技术学士和硕士学位、中国科技大学计算机科学与技术博士学位；现任中山大学计算机系教授、博士生导师、信息学院副院长、协同软件研究开发中心主任，中国计算机学会理事、数据库专业委员会委员，广东省计算机学会副理事长，广东开源软件促进会副会长，高校 Linux 推进联盟理事长，教育部-微软精品课程负责人等。

长期从事数据库与协同软件的研究、开发与教学工作，主持过各类项目 30 多项，出版著作 9 部，论文被 SCI、EI、ISTP 等收录 100 多篇次，取得软件版权 11 项，申请发明专利 1 项；1995 年至今指导硕士生 80 多名、博士生 21 名、博士后 2 名、访问学者 4 名。入选教育部“新世纪优秀人才支持计划”、广东省“千百十工程”（百人层次）、广东省委组织部高级专家库等。曾获“宝钢教育奖”、“丁颖科技奖”、教育部科技进步二等奖、广东省科技进步一等奖等。

# 目 录

<b>第 1 章 绪论</b> .....	1	2.3.4 对象关系数据库系统	56
1.1 数据库发展回顾	1	2.4 本章小结	58
1.1.1 数据管理的初级阶段	2	2.5 讨论题	59
1.1.2 第一代: 层次与网状数据库	2	<b>第 3 章 基于时间的数据库技术</b> .....	60
1.1.3 第二代: 关系数据库	3	3.1 时间模型及演算	60
1.1.4 新一代数据库的出现	4	3.1.1 时间模型与时间系统	60
1.2 数据库新技术概述	5	3.1.2 时间元素与时间粒度	62
1.2.1 数据模型的新特征	5	3.1.3 时间元素算术运算	65
1.2.2 数据库系统的新要求	6	3.1.4 时间元素关系演算	66
1.2.3 数据库新技术的范畴	7	3.1.5 时间数据与时态信息	69
1.2.4 现代数据库的实现途径	9	3.2 时态数据库技术	69
1.3 数据库的研究热点与展望	10	3.2.1 时态数据库的发展阶段	69
1.3.1 数据库的研究热点	10	3.2.2 时态数据库的基本概念	72
1.3.2 数据库新技术的综合应用	13	3.2.3 时态关系代数	80
1.4 本书的内容架构	14	3.2.4 时态数据模型 BCDM	86
1.5 讨论题	15	3.2.5 时态查询语言 TSQL2	89
<b>第 2 章 数据库理论基础</b> .....	16	3.2.6 时态数据库管理系统	95
2.1 关系数据理论	16	3.2.7 时态数据库的现状与展望	98
2.1.1 关系数据模型	16	3.3 时空数据库技术	99
2.1.2 关系模式设计	22	3.3.1 空间数据库简介	100
2.1.3 事务处理	30	3.3.2 时空数据建模	106
2.2 面向对象数据库	34	3.3.3 时空数据索引	107
2.2.1 面向对象概念	35	3.3.4 时空数据查询	109
2.2.2 面向对象数据模型	39	3.3.5 时空数据库应用	110
2.2.3 ODMG 中的数据操作	41	3.4 实时数据库技术	111
2.2.4 面向对象数据库系统	46	3.4.1 实时数据库概述	111
2.3 对象关系数据库	47	3.4.2 实时数据库基本特征	112
2.3.1 对象关系数据模型	48	3.4.3 实时事务模型	116
2.3.2 ORDB 中的数据定义	50	3.4.4 实时数据库管理系统	120
2.3.3 ORDB 中的数据查询	54	3.5 本章小结	122

3.6 讨论题	123	5.2.4 移动数据管理与分布数据管理	200
<b>第4章 基于知识的数据库技术</b>	124	<b>5.3 Web 数据库技术</b>	201
4.1 知识库技术	124	5.3.1 Web 数据库体系结构	201
4.1.1 知识的概念	124	5.3.2 Web 数据库访问连接	206
4.1.2 知识库与知识库系统	126	5.3.3 Web 信息检索技术	214
4.1.3 知识库语言	128	<b>5.4 XML 数据库技术</b>	222
4.1.4 几种经典知识表达与推理	129	5.4.1 XML 技术基本思想	222
4.1.5 数据库与知识库	138	5.4.2 XML 数据库概念	227
<b>4.2 主动数据库技术</b>	140	5.4.3 XML 数据库模式设计	229
4.2.1 主动数据库简介	141	5.4.4 XML 数据库管理系统	232
4.2.2 主动数据库的体系结构	141	<b>5.5 本章小结</b>	233
4.2.3 主动规则结构	143	<b>5.6 讨论题</b>	234
4.2.4 主动数据库实现途径	148	<b>第6章 数据库协同与综合应用</b>	235
<b>4.3 决策支持系统</b>	148	<b>6.1 数据库协同技术</b>	235
4.3.1 决策支持系统发展阶段	149	6.1.1 数据库协同管理	236
4.3.2 决策支持系统概述	149	6.1.2 信息集成技术	238
4.3.3 决策支持系统的体系结构	151	6.1.3 数据交换与信息发布	242
4.3.4 决策支持系统实现基础	155	6.1.4  workflow 数据库	247
<b>4.4 数据仓库与数据挖掘</b>	157	<b>6.2 时态信息处理综合应用</b>	249
4.4.1 从事务数据到决策知识	157	6.2.1 时态数据库软件 TempDB	250
4.4.2 数据仓库概述	159	6.2.2 时态知识处理软件 TempKB	257
4.4.3 数据挖掘技术	165	6.2.3 时态应用系统开发	272
<b>4.5 本章小结</b>	171	<b>6.3 数据库综合应用</b>	275
<b>4.6 讨论题</b>	172	6.3.1 工资时态知识数据库系统	275
<b>第5章 基于网络的数据库技术</b>	173	6.3.2 中小型数据中心协同应用系统	288
<b>5.1 分布式数据库系统</b>	173	<b>6.4 本章小结</b>	295
5.1.1 分布式数据库概述	173	<b>6.5 讨论题</b>	295
5.1.2 分布式数据库体系结构	177	<b>附录A 相关资源介绍</b>	296
5.1.3 分布式数据存储	180	A1.1 相关学术会议	296
5.1.4 分布式查询处理	184	A1.2 相关学术期刊	298
5.1.5 分布式事务管理	185	A1.3 相关教材	300
<b>5.2 移动数据库技术</b>	187	<b>附录B TempDB 2.0 中 ATSQL2 的 BNF 定义</b>	303
5.2.1 移动数据库概述	188	<b>参考文献</b>	307
5.2.2 移动数据库的关键技术	190		
5.2.3 移动代理技术	197		

# 第 1 章 绪 论

数据库技术从 20 世纪 60 年代中期产生至今已有四十多年的历史，数据库技术一直是最活跃、发展速度最快的 IT 技术之一。20 世纪 90 年代，随着网络和多媒体技术的迅猛发展，数据库的应用也得到了更广泛的拓展，在应用中的新需求直接推动了数据库技术的研究与发展。数据库技术的发展进入了一个新的时期。

作为本书的绪论，本章简要回顾了数据库的发展，讨论了新一代数据库的需求和特征以及新一代数据库技术的范畴和实现途径，最后介绍了本书的组织结构。

## 1.1 数据库发展回顾

数据库技术从 20 世纪 60 年代中期产生至今已有 40 多年历史。一般来说，数据库发展可以分为 3 个阶段：第一代数据库以网状模型和层次模型为特征；第二代数据库以关系模型为基础，关系数据库在 20 世纪 70 年代末和 80 年代得到了广泛应用；20 世纪 90 年代，数据库系统已经从第一代的网状、层次数据库系统，第二代的关系数据库系统，发展到第三代的以面向对象为主要特征的数据库系统。

图 1-1 是数据库技术的发展历程回顾。

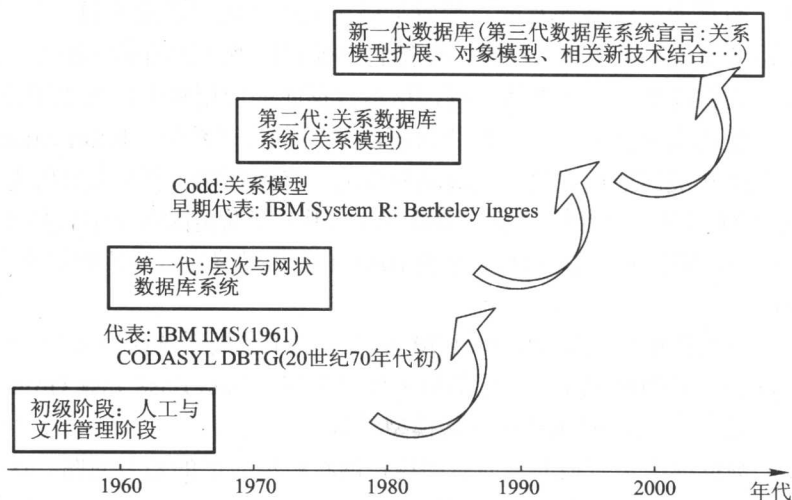


图 1-1 数据库技术的发展历程

### 1.1.1 数据管理的初级阶段

数据库技术是应数据管理任务的需要而产生的。数据库的发展随着对数据管理的要求的提高而不断发展。在数据库出现以前, 数据管理经历了人工管理阶段和文件管理阶段。

数据的处理是对各种数据进行收集、存储、加工和传播等的活动。数据的管理则是指对数据进行分类、组织、编码、存储、检索和维护, 数据管理是数据处理的中心问题。

计算机的基本作用是进行复杂的科学计算, 然而随着计算机技术的发展, 其应用已经远远超过了这个范围, 对数据进行处理成了计算机的主要用途。在计算机软件和硬件的发展基础上, 对数据的管理已从人工的管理、文件系统发展到了 20 世纪 60 年代的数据库系统。

人工对数据进行管理, 工作反复、效率低、数据不具有独立性、难以共享。但是在早期没有磁盘等可直接存取的设备, 也没有操作系统支持等条件下, 这种办法是对数据进行管理的唯一且较有成效的办法。

在文件系统阶段对数据的管理有了较大的改善, 数据可以长期保存。文件系统提供了专门的软件对数据进行管理, 为数据管理提供了有效手段。但是, 文件系统中的数据冗余度较大、数据的共享性差、数据独立性差。随着对数据管理性能要求的提高, 比如更高的共享性、更好的数据独立性, 就要求对于数据管理的方法和技术要有新的突破, 数据库技术应运而生。

#### 1.1.2 第一代: 层次与网状数据库

20 世纪 60 年代的层次和网状模型是数据库技术中的初级系统, 它们为统一管理与共享数据提供了有力的支撑, 以层次和网状模型为基础的数据库被称为第一代数据库系统。

在这一时期, 数据库系统蓬勃发展, 形成了历史上著名的“数据库时代”。但是这两种模型均脱胎于文件系统比较简单的数据结构, 使得它们受到物理结构的影响较大, 用户对数据库的使用需要对数据的物理结构有详细的了解, 这给数据库的应用和推广带来了许多麻烦。

层次数据库系统的典型代表是 IBM 公司研制的信息管理系统 (Information Management System, IMS)。1961 年美国通用电气公司成功开发的网状数据库 (也是人类历史上的第一个数据库)。1966 年, IBM 与其客户合作开发新型数据库, 用于帮助美国国家宇航局管理宏大的“阿波罗登月计划”中的烦琐资料, 并在 1968 年由 IBM 工程师最终完成。该数据库在 1969 年发布时被命名为“IMS”。

网状数据库系统的典型代表是 20 世纪 70 年代初美国数据系统语言协会 (Conference On Data System Language, CODASYL) 下属的数据库任务组 (Data Base Task Group, DBTG) 提出了 DBTG 报告, 建立了数据库系统的网状模型结构。

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。这两种数据库奠定了现代数据库发展的基础, 具有如下共同特点。

##### 1. 支持三级模式的体系结构

三级模式通常指外模式、模式、内模式。模式之间具有转换 (或称为映射) 功能。层次数

据库和网状数据库均支持三级模式结构，通过外模式与模式、模式与内模式之间的映像，保证了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. 用存取路径表示数据之间的联系

数据库不仅存储数据，而且存储数据之间的联系。数据之间的联系在层次和网状数据库系统中是用存取路径来表示和实现的。这是数据库系统和文件系统的主要区别之一。

## 3. 独立的数据定义语言

层次数据库系统和网状数据库系统都有独立的数据定义语言，用以描述数据库的外模式、模式、内模式以及相互映像。但是模式一经定义，就很难修改。修改模式必须首先把数据全部卸出，然后重新定义、重新生成诸模式，最后编写实用程序把卸出的数据按新模式的定义装入新数据库中。因此在许多实际运行的层次、网状数据库系统中，模式不是能够轻易重构的。数据库设计人员在建立数据库应用系统时，不仅应当充分考虑用户的当前需求，还要充分了解需求可能的变化和发展，因此对数据库设计水平的要求比较高。

## 4. 导航的数据操纵语言

所谓导航就是指用户不仅要了解“要干什么”，而且要指出“怎么干”。用户必须使用某种高级语言编写程序，一步一步地“引导”程序按照数据库中某一条预先定义的存取路径来访问数据库，最终达到要访问的数据目标。在访问数据库时，每次只能存取一条记录值。若该记录值不满足要求，则沿着存取路径查找下一条记录值。层次和网状数据库的数据查询和数据操纵语言是一次一条记录的导航式的过程化语言。这类语言通常嵌入某一种高级语言，如 COBOL、FORTRAN、PL/1 中。

### 1.1.3 第二代：关系数据库

关系数据库系统形成于 20 世纪 70 年代中期，并在 80 年代得到了充分的发展，它具有简单的结构方式并且极为方便，因此 80 年代它逐步取代了层次和网状数据库，成为占主导地位的数据库，并发展为第二代数据库系统。第二代数据库的主要特征是支持关系数据模型。

1970 年，IBM 公司的 San Jose 试验室的研究员 Edgar F. Codd 在其论文《大型共享数据库数据的关系模型》中开创性地提出了关系数据模型，为关系数据库技术奠定了理论基础。正是由于他在关系数据库研究方面的杰出贡献，Codd 在 1981 年被授予 ACM 图灵奖。

20 世纪 70 年代是关系数据库理论研究和原型开发的时代，大量的理论成果和实践经验终于使关系数据库从实验室走向了社会，IBM 公司的 San Jose 研究试验室开发的 System R 和加州大学伯克利分校研制的 Ingres 是典型代表。20 世纪 80 年代几乎所有新开发的系统都是关系数据库，其中涌现出许多商品化关系数据库管理系统，如 DB2、Ingres、Oracle、Informix、Sybase 等。

关系数据库以关系模型为基础，关系模型的基本组成包括数据结构、关系操作、数据完整性 3 部分。

### 1. 数据结构

数据结构包括域和域上定义的关系等。关系模型中，实体、实体与实体之间的联系都是通过关系这种单一的结构类型来表示的。

### 2. 关系操作

关系操作的能力可用关系代数（或等价的关系演算）中的并、交、差、广义笛卡儿积、选择、投影、连接、除等操作来表示。关系操作的特点是集合操作。无论是操作的对象还是操作的结果都是集合。这种操作方式被称为一次一集合（set-at-a-time）的方式，与非关系型的一次一条记录（record-at-a-time）的方式相对应。

### 3. 数据完整性

数据完整性包括实体完整性、参照完整性以及与应用有关的完整性。实体完整性和参照完整性是关系模型必须满足的完整性约束条件，应由关系数据库系统自动支持；与应用有关的完整性由用户定义，它是针对某一具体数据库的约束条件，由应用环境决定，关系数据库系统应提供定义和检验这类完整性的机制。

关系模型具有如下主要特点。

- 关系模型的概念单一，实体和实体之间的联系采用关系来表示。
- 以关系代数为基础，易于进行形式化描述。
- 数据独立性强，数据的物理存储和存取路径对用户透明。
- 关系数据库语言是非过程化的，可以将用户从编程数据库记录的导航式检索中解脱出来，大大降低了用户编程的难度。

## 1.1.4 新一代数据库的出现

20世纪80年代，数据库技术在商业领域的巨大成功刺激了其他领域对数据库技术需求的迅速增长。随着科学技术的不断进步，各个行业领域对数据库技术提出了更多的需求，关系数据库已经不能完全满足需求，以关系数据库为代表的传统数据库已经很难胜任新领域的要求。这样就必须要新的数据库技术才能满足现实需要。为了支持现代工程的应用，必须将数据库技术与其他现代信息、数据处理技术，如面向对象技术、时序和实时处理技术、人工智能技术、多媒体技术“完善”地集成，以形成“新一代数据库技术”。

1990年，美国的高级DBMS功能委员会发表了《第三代数据库系统宣言》，提出了第三代数据库管理系统应具有的特征，包括3条基本原则和13个命题。其基本思想如下。

- 第三代数据库系统应支持数据管理、对象管理和知识管理。第三代数据库系统应集数据管理、对象管理和知识管理为一体，支持丰富的对象结构和规则。即面向对象数据模型是第三代数据库系统的重要特征。
- 第三代数据库系统必须保持或继承第二代数据库系统的技术。必须保持第二代数据库系统的非过程化数据存取方式和数据独立性。即第三代数据库是关系模式的扩展。
- 第三代数据库系统必须对其他系统开放。数据库系统的开放性表现在：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。

现代数据库技术融合多种技术，面向对象成为其主要特征。但是，由于面向对象模型本身发展尚未成熟，所以第三代数据库没有像第二代关系数据库那样具有统一的数据模型，第三代数据库的数据模型是以关系模型为基础，支持多种数据模型（具有面向对象特征）的复杂数据模型。

在现代数据库研究和应用中，往往需要与诸多新技术相结合（比如分布处理技术、并行计算技术、人工智能技术、多媒体技术、模糊技术），因此也衍生出多种新的数据库技术。这些数据库新技术的研究和发展导致了众多不同于传统数据库的系统诞生，构成了当今数据库系统的大家族。由于第三代数据库系统没有统一的数据模型，人们更多地使用“新一代数据库系统”或“现代数据库”概念。在本书中，我们经常使用现代数据库技术、数据库新技术等术语。

## 1.2 数据库新技术概述

数据库技术从 20 世纪 60 年代中期发展到今天仅仅是几十年的历史，但发展速度之快、使用范围之广是其他技术远远不能及的。从 20 世纪 80 年代开始，数据库技术在商业领域的巨大成功刺激了其他领域对数据库技术需求的迅速增长。一方面，新的数据库应用领域，如计算机辅助设计/计算机辅助管理（CAD/CAM）、过程控制、雷达跟踪、办公自动化系统、地理信息系统（GIS）、计算机集成制造（CIM）系统和知识库系统等，为数据库的应用开辟了新的天地；另一方面，在应用中管理方面的新需求也直接推动了数据库技术的研究与发展。

以关系数据库为代表的传统数据库已经很难胜任新领域的要求，因为新的应用要求数据库能处理复杂性较高的数据，如处理与时间有关的属性，甚至还要求数据库有动态性和主动性。这样就必须有新的数据库技术才能满足现实需要。为了满足现代应用的需要，必须将数据库技术与其他现代信息、数据处理技术（如面向对象技术、时序和实时处理技术、人工智能技术、多媒体技术）“完善”地集成，以形成“新一代数据库技术”，也可称为“现代数据库技术”，如时态数据库技术、实时数据库技术和多媒体数据库技术等。

### 1.2.1 数据模型的新特征

现代应用所涉及的范围很广泛，所呈现出的特征行为非常复杂，对数据库的要求也十分多样，例如数据库要有主动性，数据库能表达数据的时间属性等。现代应用对数据库的要求的一个具体体现是对数据模型的要求。数据模型主要是指数据库的数据结构、对数据的操作以及数据的约束条件。现代数据库对数据模型的新要求主要体现在以下方面。

#### 1. 数据特征

现代应用中的数据表现出与传统应用数据不同的特征。

##### (1) 多维性

每个数据对象除了用值来表示外，每个值还有与其相联系的时间属性，即数据是二维的；



更进一步,如果联系到空间,其值就是三维的;如果考虑到时间的两维性(有效时间和事务时间)以及空间的三维性,数据的维就更加复杂。

### (2) 易变性

数据对象频繁地发生变化,其变化不仅表现在数据的值上,而且表现在它的定义上,也就是数据的定义可以动态改变。

### (3) 多态性

数据对象不仅是传统意义下的值,还可以是过程、规则、方法和模型,甚至是声音、影像和图形等。

## 2. 数据结构

在数据结构的表示方面,现代应用的要求更高。

### (1) 数据类型

不仅要求能表达传统的基本数据类型,如整型、实型和字符型等,还要求能表达更复杂的数据类型,如集合、向量、矩阵、时间类型和抽象数据类型等。

### (2) 数据之间的联系

数据之间有了各种复杂的联系,如 $n$ -元联系;多种类型之间的联系,如时间、空间、模态联系;非显式的联系,如对象之间隐含的关系。

### (3) 数据的表示

除了表示结构化、格式化的数据,还要表示非结构化、半结构化的数据,非格式、超格式的数据。

## 3. 数据的操作

对数据的操作在现代应用中也表现出了新的特征,主要有:

### (1) 数据操作的类型

数据操作的类型不仅包含通常意义上的插入、删除、修改和查询;还要进行各种其他类型的特殊操作,如执行、领域搜索、浏览和时态查询等;另外还要能够执行用户自定义操作。

### (2) 数据的互操作性

要求数据对象可以在不同模式下执行交互操作,数据可以在不同模式的视图下进行交互作用。

### (3) 数据操作的主动性

传统数据库中的数据操作都是被动的、单向的,即只能由应用程序控制数据操作,其作用方向只能是应用程序到数据;而现代应用要求数据的使用加强主动性和双向作用,即数据的状态和状态变迁可主动地驱动操作,除了应用作用于数据外,数据也可以作用于应用。

### 1.2.2 数据库系统的新要求

现代应用的复杂性、主动性和时态性等特性对数据库系统的要求是多方面的,它涉及数据库的很多方面,从数据建模到数据查询,从数据存储到数据库管理等多方面,大致归纳如下。