



高等教育“十三五”规划教材

数据库原理教程

(第二版)

范明 叶阳东 邱保志 职为梅 编著

融合数字化资源的
新形态教材



科学出版社

普通高等教育“十三五”规划教材

数据库原理教程

(第二版)

范 明 叶阳东 邱保志 职为梅 编著

科学出版社

北 京

内 容 简 介

本书第1版于2008年出版,被列为普通高等教育“十一五”规划教材。本书在第1版的基础之上,结合现代高等教育对人才培养提出的新要求(从“知识传授”到“能力培养”)修订而来。

本书全面阐述数据库系统的基本概念、理论、方法和数据库新技术。全书共13章,包括数据库系统概述、实体-联系模型、关系数据模型、关系数据库标准语言SQL、数据库的完整性、数据库的安全性、关系数据库的设计理论、数据库设计、查询处理与优化、事务与并发控制、数据库的恢复技术、数据仓库和大数据管理。本书涵盖数据库系统原理入门课程的基本内容,同时包含数据仓库和大数据管理等实用的新知识点。

本书可作为高等学校计算机、信息科学及其他相关专业本科生的数据库课程基础教材,也可供数据库工程技术人员学习参考。

图书在版编目(CIP)数据

数据库原理教程/范明等编著. —2版. —北京:科学出版社,2018.3

普通高等教育“十三五”规划教材

ISBN 978-7-03-056814-4

I. ①数… II. ①范… 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(2018)第048352号

责任编辑:于海云/责任校对:郭瑞芝

责任印制:吴兆东/封面设计:迷底书装

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

北京教图印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2008年4月第 一 版 开本:787×1092 1/16

2018年3月第 二 版 印张:18 1/2

2018年3月第七次印刷 字数:420 000

定价:56.00元

(如有印装质量问题,我社负责调换)

前 言

数据库技术发展非常迅速,为了适应数据库技术的新进展,我们对《数据库原理教程》第1版进行了修订。为了更好地对本书进行修订,我们征求了多名使用过该教材的老师意见,并对使用该教材的学生做了问卷调查,调查表明第1版教材存在以下问题。

(1)第1版教材中案例较少。有些深奥难懂的理论知识点没有足够的例子帮助学生理解。

(2)第1版教材不能反映数据库学科的新技术和新发展。近年来,数据库学科发展迅速,自2008年第1版教材出版以来,数据库出现了很多新技术,这要求我们对其进行升级,以跟随新技术的发展。

(3)第1版教材不能适应新时代教育理念的发展。现代高等教育理念已经从传统的“知识传授”为主导转化为以“能力培养”为主导的教学模式,这是大学教育的一场深刻变革,教材必须体现新的教育理念,以培养学生综合应用能力和解决复杂工程问题的能力。

《数据库原理教程(第二版)》是为本科生编写的教程,作为数据库的入门教材,我们保持原书的宗旨和风格不变,将重点放在基本概念、原理和技术的讲解,包括数据库系统的三级模式和两级映像、实体-联系模型、关系数据模型、SQL、数据的完整性和安全性、关系数据库规范化理论、数据库设计、查询处理和优化、事务处理、并发控制和数据库的恢复等。

在第1版的基础上,第2版主要修改的内容包括以下几方面。

(1)增加一些案例,帮助学生理解复杂的理论知识,如关系模式、数据字典等概念。增加直观明了的图表,便于学生理解。

(2)数据库学科发展飞快,增加了一些反映数据库新技术的内容,如数据仓库和数据挖掘技术、大数据管理技术等。限于篇幅删去了XML和ODBC编程。

(3)增加视频讲解,并以二维码的方式呈现,增强学生的自主学习能力。同时,为了进一步加强实验教学环节,根据教材章节的内容安排了9个提高学生动手能力的实验项目,培养学生的综合应用能力和动手能力。

本书由职为梅主持修订。职为梅编写第12章和第13章。范明、叶阳东、邱保志对各章进行通读并做了修改。

全书共13章,主要内容如下:

第1章是数据库系统的一般综述,所涉及的内容将在之后各章进一步展开讨论。

第2章介绍实体-联系模型(E-R模型)。E-R模型概念简单,并具有很强的语义表达能力,广泛用于对现实世界建模。

第3章介绍数据库系统的关系数据模型,包括与之相关的抽象语言关系代数和关系演算。这些抽象语言是实际数据库系统语言(如SQL)的基础。

第4章介绍广泛使用的关系语言SQL。SQL是事实上的关系数据库标准语言,几乎所有的RDBMS都支持SQL。SQL的介绍主要基于SQL-92,同时包含SQL-99的部分特色。

第5章介绍数据库的完整性,包括SQL对完整性的支持。

第6章介绍数据库的安全性,包括SQL对安全性的支持。

第7章介绍关系数据库的设计理论,包括数据依赖和规范化。为了适应不同读者的需要,

理论结果的证明从正文分离,在该章附录给出。

第 8 章介绍数据库应用系统的设计,涵盖从需求分析到数据库的建立和维护整个过程。

第 9 章介绍查询处理与优化,包括基本运算算法与基于保持等价性的查询变换和查询优化方法。

第 10 章、第 11 章分别介绍事务处理和数据库恢复。其中,第 10 章介绍事务的一般概念,讨论并发控制,第 11 章介绍在出现故障的情况下保证事务正确执行的数据库恢复技术。

第 12 章、第 13 章分别介绍数据仓库和大数据管理。为了反映数据库学科的新成果和应用的最新方向,本书中加入目前数据仓库和大数据的内容。随着大数据时代的来临,数据仓库对企业决策的支持作用越来越大,数据仓库也成为各大厂商看重并着力发展的业务领域。而大数据是目前全世界关注的热点。

在 20 多年的数据库教学中,我们使用和参考过许多国内外优秀教材。这些教材不仅为我们过去的教学提供了方便和支持,而且深深影响本书的编写。我们向这些教材(见参考文献)的所有作者和译者表示敬意与感谢。

感谢选修数据库系统原理课程的历届学生。在 20 多年的教学过程中,他们的求知欲推动我们不断思考如何选择和组织教学内容,如何用简洁而又不失严谨的表述方式讲述数据库的基本概念和技术。感谢所有使用过该教材的老师,他们对再版教材提出了很多良好的建议。感谢参与问卷调查的全体学生,他们对再版教材提出了很好的建议。本书的修订受国家自然科学基金(基金号为 61772475)的资金支持,也受到郑州大学 2016 年度教材建设的立项支持,他们的支持使得本书能够顺利改版,在此一并向他们表示感谢。

在本书的修订过程中,我们努力将教学过程中的经验体会和研究成果纳入书中内容,但由于学科发展太快,我们水平有限,书中难免有不足之处,希望各位学术同仁不吝赐教。意见和建议请发送至 iewmzhi@zzu.edu.cn,我们将不胜感激。

编者

2017 年 12 月于郑州大学

目 录

前言

第 1 章 数据库系统概述	1
1.1 基本概念	1
1.1.1 数据	1
1.1.2 什么是数据库	2
1.1.3 数据库管理系统	4
1.1.4 数据库系统	4
1.2 数据模型	5
1.2.1 概念模型	6
1.2.2 数据模型的三要素	6
1.2.3 关系模型	7
1.2.4 其他数据模型	8
1.3 数据库系统的结构	9
1.3.1 数据库系统的外部结构	9
1.3.2 数据库系统的三级模式结构	10
1.3.3 二级映像与数据独立性	11
1.4 数据库语言	12
1.4.1 数据定义语言	12
1.4.2 数据操纵语言	13
1.4.3 数据控制语言	13
1.5 数据库管理员与数据库用户	13
1.5.1 数据库管理员	13
1.5.2 数据库用户	14
1.6 数据库技术的发展	15
1.7 小结	15
习题	16
第 2 章 实体-联系模型	17
2.1 基本概念	17
2.1.1 实体	17
2.1.2 联系	19
2.2 实体-联系图	22
2.3 一个例子	24
2.4 弱实体集	25
*2.5 扩展的 E-R 图	28
2.5.1 特殊化和一般化	28

2.5.2	属性继承	29
2.5.3	一般化约束	30
2.5.4	聚集	31
2.6	小结	32
	习题	33
第3章	关系数据模型	35
3.1	关系数据库的结构	35
3.1.1	关系的相关概念	35
3.1.2	关系的码	37
3.1.3	关系模式	38
3.1.4	关系数据库	38
3.2	从 E-R 模型到关系模型	39
3.2.1	基本 E-R 图的转换	39
*3.2.2	扩展 E-R 图到关系模型的转换	42
3.3	关系操作	44
3.3.1	关系语言的分类	44
3.3.2	关系操作的特点	44
3.4	关系的完整性约束	44
3.4.1	空值	45
3.4.2	实体完整性	45
3.4.3	参照完整性	45
3.4.4	用户定义的完整性	47
3.5	关系代数	47
3.5.1	基本运算	47
3.5.2	附加的关系运算	49
3.5.3	用关系代数表达式表示查询	51
3.5.4	扩展的关系运算	56
3.6	关系演算	58
3.6.1	元组关系演算	59
3.6.2	域关系演算	62
3.6.3	关系语言的表达能力	63
*3.7	关系数据库的更新	63
3.7.1	插入	63
3.7.2	删除	64
3.7.3	修改	64
3.8	小结	65
	习题	65
	实验一：认识数据库管理系统	67
第4章	关系数据库标准语言 SQL	68
4.1	SQL 概述	68

4.1.1	发展历程	68
4.1.2	SQL 功能	69
4.1.3	SQL 的特点	69
4.2	学生选课数据库	70
4.3	数据定义	72
4.3.1	SQL 的数据类型	72
4.3.2	模式的定义和删除	73
4.3.3	定义、修改和删除基本表	74
4.3.4	建立和删除索引	78
4.4	数据查询	79
4.4.1	SELECT 语句的一般形式	79
4.4.2	不带 WHERE 的简单查询	80
4.4.3	带 WHERE 子句的查询	81
4.4.4	排序和分组	83
4.4.5	连接查询	85
4.4.6	嵌套查询	88
*4.4.7	子查询导出的表	92
4.4.8	集合运算	92
4.5	数据更新	94
4.5.1	插入	94
4.5.2	删除	95
4.5.3	修改	96
4.6	视图	96
4.6.1	定义视图和删除视图	97
4.6.2	基于视图的查询	99
4.6.3	基于视图的更新	99
4.6.4	视图的作用	100
4.7	嵌入式 SQL	101
4.7.1	概述	101
4.7.2	不使用游标的 SQL 语句	104
4.7.3	使用游标的 SQL 语句	105
4.7.4	动态 SQL	110
4.8	小结	111
	习题	112
	实验二: 交互式 SQL 语句	113
	实验三: 嵌入式 SQL 语句	113
第 5 章	数据库的完整性	114
5.1	完整性概述	114
5.2	实体完整性	115
5.3	参照完整性	115

5.3.1	参照完整性与 E-R 模型	116
5.3.2	违反参照完整性的更新	116
5.3.3	SQL 中的参照完整性	117
5.4	用户定义的完整性	117
5.4.1	域约束	117
5.4.2	属性约束	120
5.4.3	关系约束	120
5.4.4	断言与数据库约束	121
5.5	触发器	122
5.5.1	为什么需要触发器	122
5.5.2	SQL 中的触发器	122
5.6	小结	124
	习题	125
	实验四：完整性控制	125
第 6 章	数据库的安全性	126
6.1	安全性概述	126
6.2	用户标识与鉴别	127
6.3	存取控制	128
6.4	自主存取控制	129
6.4.1	数据对象和存取权限	129
6.4.2	存取控制的任务	129
6.4.3	权限的授予和回收	129
6.4.4	角色	131
6.4.5	其他特性	133
6.4.6	SQL 授权的局限	133
6.5	强制存取控制	134
6.6	视图机制	135
6.7	其他安全措施	136
6.7.1	审计技术	136
6.7.2	数据加密	136
6.8	小结	137
	习题	137
	实验五：安全性控制	138
第 7 章	关系数据库的设计理论	139
7.1	问题提出	139
7.1.1	不好的设计可能导致的问题	139
7.1.2	数据依赖与冗余	141
7.2	函数依赖	141
7.2.1	函数依赖的定义	141
7.2.2	函数依赖的意义	142

7.2.3	逻辑蕴涵和依赖集的闭包	143
7.3	函数依赖的推导	144
7.3.1	Armstrong 公理	144
7.3.2	属性集的闭包	145
7.3.3	函数依赖集的等价和极小覆盖	147
7.4	关系模式的范式	148
7.4.1	范式与规范化概述	149
7.4.2	2NF、3NF 和 BCNF	149
7.4.3	函数依赖与范式	151
7.5	关系模式的分解	151
7.5.1	无损连接分解	151
7.5.2	保持函数依赖的分解	153
7.6	将关系模式分解成高级范式	156
7.6.1	具有无损连接性的 BCNF 分解	156
7.6.2	具有无损连接性和保持函数依赖 3NF 分解	157
7.7	多值依赖与 4NF	158
7.7.1	多值依赖	159
7.7.2	函数依赖和多值依赖的公理	160
7.7.3	函数依赖和多值依赖的闭包	161
7.7.4	4NF	162
7.7.5	具有无损连接性的 4NF 分解	162
7.7.6	嵌入型多值依赖	163
7.8	在设计中使用规范化理论	164
7.8.1	关于规范化的附注	164
7.8.2	E-R 模型与规范化	165
7.8.3	泛关系设计方法	166
7.8.4	逆规范化	167
7.9	小结	167
	习题	168
	附录: 本章引理和定理证明	169
第 8 章	数据库设计	175
8.1	数据库设计概述	175
8.1.1	什么是数据库设计	175
8.1.2	数据库设计的特点	176
8.1.3	数据库设计的步骤	177
8.2	需求分析	178
8.2.1	需求分析的步骤	178
8.2.2	数据字典	180
8.3	数据库概念设计	182
8.3.1	什么是概念结构设计	182

8.3.2	概念结构设计的方法	182
8.3.3	数据抽象	183
8.3.4	分 E-R 图设计	184
8.3.5	分 E-R 图集成	185
8.4	逻辑结构设计	187
8.4.1	关系模式的规范化和优化	188
8.4.2	外模式的设计	189
8.5	数据库物理设计	189
8.5.1	概述	189
8.5.2	存取方法的选择	190
8.5.3	存储结构	192
8.6	数据库的实施和维护	192
8.6.1	数据库建立与试运行	192
8.6.2	数据库的运行与维护	193
8.7	小结	194
	习题	195
	实验六: 数据库设计	196
第 9 章	查询处理与优化	197
9.1	查询处理概述	197
9.2	选择运算的实现	199
9.2.1	基本算法	199
9.2.2	使用索引的选择	200
9.2.3	复杂选择的实现	201
9.3	连接运算的实现	202
9.3.1	基本算法	202
9.3.2	索引嵌套循环连接	203
9.3.3	排序-归并连接	203
9.3.4	散列连接	204
9.3.5	复杂的连接	205
9.4	查询优化概述	205
9.5	代数优化	207
9.5.1	关系代数表达式的等价变换规则	207
9.5.2	代数优化的启发式方法	210
9.6	物理优化	214
9.6.1	基于代价的优化	214
9.6.2	物理优化的启发式方法	216
9.6.3	一个物理优化的例子	217
9.7	小结	219
	习题	219
	实验七: 查询优化	220

第 10 章 事务与并发控制	221
10.1 事务的概念	221
10.1.1 事务的特性	221
10.1.2 事务状态	222
10.1.3 SQL 对事务的支持	223
10.2 并发控制概述	224
10.2.1 事务的并发执行	224
10.2.2 并发执行可能导致的问题	225
10.3 封锁	226
10.3.1 共享锁与排它锁	227
10.3.2 封锁协议	227
10.3.3 活锁与死锁	228
10.3.4 死锁的预防	229
10.3.5 死锁的检测与解除	230
10.4 并发调度的可串行化	231
10.4.1 串行调度与并发调度	231
10.4.2 冲突可串行化	233
10.4.3 视图可串行化	234
10.5 两阶段锁协议	235
10.6 多粒度封锁	237
10.6.1 粒度的层次结构	237
10.6.2 意向锁	238
10.7 小结	239
习题	239
实验八：事务与并发控制	240
第 11 章 数据库的恢复技术	241
11.1 数据库恢复概述	241
11.2 存储器结构	242
11.2.1 存储器类型	242
11.2.2 稳定存储器的实现	243
11.2.3 数据访问	243
11.3 基于日志的恢复技术	244
11.3.1 日志	245
11.3.2 延迟更新技术	245
11.3.3 即时更新技术	247
11.4 基于检查点的恢复技术	248
11.5 缓冲技术	250
11.5.1 日志缓冲	250
11.5.2 数据库缓冲	250
11.6 介质故障恢复技术	251

11.6.1	转储	251
11.6.2	介质故障恢复	252
11.7	其他恢复技术	252
11.7.1	影子分页技术	252
11.7.2	数据库镜像	253
11.8	小结	254
	习题	254
	实验九：数据库的备份与恢复	255
第 12 章	数据仓库	256
12.1	数据仓库概述	256
12.1.1	什么是数据仓库	256
12.1.2	数据仓库的构建	257
12.1.3	数据仓库的多维数据模型	257
12.1.4	联机分析处理技术	259
12.2	数据挖掘	260
12.2.1	什么是数据挖掘	261
12.2.2	数据挖掘的主要任务	262
12.3	大数据下的新型数据仓库	264
12.3.1	大数据环境下企业面临的挑战	264
12.3.2	大数据环境下的新型数据仓库	265
12.4	小结	266
	习题	266
第 13 章	大数据管理	267
13.1	大数据概述	267
13.1.1	什么是大数据	267
13.1.2	大数据的特征	267
13.2	大数据关键技术	269
13.2.1	大数据采集	269
13.2.2	大数据预处理	269
13.2.3	大数据存储及管理	269
13.2.4	大数据分析与管理	270
13.2.5	大数据展现与应用	271
13.3	大数据的典型应用	273
13.4	大数据处理系统	274
13.4.1	Hadoop	274
13.4.2	NoSQL	277
13.4.3	NewSQL	278
13.5	小结	278
	习题	279
	参考文献	280

第 1 章 数据库系统概述

数据库技术产生于 20 世纪 60 年代中期，经过 50 余年的发展，已经成为计算机学科的重要分支。数据库技术是信息系统的核心和基础，它的出现极大地促进了计算机应用向各行各业的渗透。现在，数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

随着网络技术和信息技术的发展，越来越多的用户通过计算机或者移动终端订购商品、机票、火车票、预订酒店；通过网上银行存取款、检索和管理账户；数据库技术和 Internet 技术的结合使得数据库已经成为每个人生活中不可缺少的部分。因此，“数据库原理”课程不仅是计算机相关专业的核心课程，而且是许多非计算机专业的选修课程。

本章是数据库系统的概述，简要介绍数据库的一些基本概念，在后续的章节中会对这些内容进一步讨论和分析。

1.1 基本概念

今天，数据库已经无处不在。本节我们简略介绍什么是数据、数据库、数据库管理系统和数据库系统的相关概念。

1.1.1 数据



数据是数据库中存储的基本对象，是描述现实世界中各种具体事物或抽象事物的符号，这些符号可以是数字，也可以是文字、声音、图形、图像、视频等。它们经过处理后存储在数据库中。

单纯的数据不能完全表达其内容，需要和它的语义背景结合起来理解，换句话说数据和其语义环境密切相关。例如，(李明, 2013)是一个数据记录，该记录放到本科生毕业管理系统中，表示李明是一个研究生，他 2013 年毕业；放到幼儿园学生管理系统中，表示他是一个幼儿园的孩子，他 2013 年出生。同一个数据放到不同的语义环境中数据的含义就发生了变化，因此数据和它的语义背景是分不开的。

数据通常有型和值之分。数据的型指的是数据的结构，数据的值指的是数据的具体取值。例如，表 1.1 中，学生数据由“学号”“姓名”“性别”“出生年月”“专业”数据项构成，这几个数据项构成了学生数据的型，学生数据的值是表 1.1 中第二行起，每一行表示一个学生数据在这个数据型下的具体取值。

表 1.1 学生数据

学号	姓名	性别	出生年月	专业
201705001	王万里	男	1998.12	计算机科学与技术
201706001	张晓霞	女	1999.01	软件工程
...	...	...	...	...

数据库中常见的数据类型有数值型、字符型、日期型等。数据是受数据类型和取值范围约束的，例如，“学号”，该数据既可以定义为整数类型也可以定义为字符类型，如果定义为整数类型，则不能出现以 0 开头的学号，如果定义为字符类型则没有此限制。当一个数据的数据类型确定后，还可以约定该数据的取值范围，仍以“学号”为例，假设定义它为整数类型，约定它取值范围是 2016001~2016999。

现实生活中的数据往往不是孤立存在的，它们之间是有联系的，如在本科生管理系统中的“学生”数据和“课程”数据之间就存在联系，“学生”选修“课程”，“课程”被“学生”选修，它们之间存在“选修”联系，如图 1.1 所示。我们在数据库中保存数据的时候不仅要保存数据还要保存数据和数据之间的联系。

学生 (Students)				
学号	姓名	性别	出生年月	专业
201705001	王万里	男	1998.12	计算机科学与技术
201706001	张晓霞	女	1999.01	软件工程
...	...	...	...	...

课程 (Courses)				选修 (SC)		
课程号	课程名	学时	学分	学号	课程号	成绩
CS101	计算科学导论	32	2	201705001	CS101	89
CS102	数据库系统	64	4	201705001	CS102	92
CS103	操作系统	64	4	201705002	CS103	90
...	...	...	...	...	...	...

图 1.1 数据和数据之间的联系

1.1.2 什么是数据库

数据库技术本质上是一种数据管理技术，数据管理顾名思义就是对数据的管理，具体来说是指使用电子计算机对数据进行分类、组织、存储、检索和维护。几乎所有的行业和部门都存在并且不断产生的大量数据。为了维持正常运作，这些行业和部门都需要持久地存储、维护和管理它们的数据。

1. 数据管理的例子

几乎所有的行业和部门都存在并且不断产生的大量数据。为了维持正常运作，这些行业和部门都需要持久地存储、维护和管理它们的数据。下面是数据管理的一些典型例子。

- (1) 零售业：管理产品、客户和购买信息。
- (2) 银行业：管理客户、账户和存贷款信息。
- (3) 制造业：管理供应、订单、库存、销售等信息。
- (4) 交通：例如，航空公司管理航班和订票信息，铁路部门管理客车和火车票销售信息，公路交通部门管理班车和客车票销售信息。
- (5) 电信业：管理通信网络信息、存储通话记录、维护电话卡余额。
- (6) 图书馆：管理图书资料、读者和借阅信息。
- (7) 政府部门：例如，税务部门需要管理纳税人信息和纳税信息，交管部门需要管理车

辆信息和驾驶员信息。

(8)学校：管理学生、教师、课程信息和学生成绩。

要将这些数据存储在计算机系统中，所面临的共同问题是如何合理地组织数据、如何有效地访问数据。数据库技术与数据库管理系统的产生和发展的直接动力正是源于数据管理的需要。

数据库(Database, DB)是持久存储在计算机中有组织的、可共享的大量数据的集合。数据库中的数据按一定的数据模型组织、描述和存储，可以被各种用户共享，具有较小的冗余度、较高的数据独立性，并且易于扩展。

2. 数据管理的基本操作

不同的应用涉及不同的数据。许多应用表面看来很不相同，似乎没有什么共同点。例如，超市存储的数据与银行存储的数据内容很不相同，用法也很不相同。然而，对于数据管理，所有的应用都需要如下基本操作。

(1)**数据查询**：从数据库系统中(我们暂且称它为数据库)查找用户需要的信息。例如，查找图书、查阅学生成绩、查找商品价格等。

(2)**数据插入**：将新的数据输入到数据库中。用数据库的术语，这种操作称作“插入”。例如，新产品的信息需要输入数据库，新的银行账户信息需要输入数据库等。

(3)**数据删除**：从数据库删除不再需要的数据。例如，车辆报废，其相关信息要从数据库中删除；账户注销，其信息要从相关数据库中删除等。

(4)**数据修改**：修改数据库中某些数据。例如，某种商品降价，修改它的价格。本质上，修改可以用删除+插入实现。然而，修改作为一种单独的操作是方便的。

数据的插入、删除和修改统称**数据更新**。几乎所有的更新操作都涉及隐含的查询。

例如，订购8月1日从郑州到广州CZ3971航班的机票本质上是导致该航班的机票数减1。但是，我们必须首先查询是否有8月1日CZ3971航班；如果有，是否有剩余机票。

3. 实际应用对数据操作的要求

实际应用是在并发、充满故障和错误的环境下运行的。这对数据操作的实现提出了很高的要求。这些要求包括以下几方面。

(1)**并发访问**：允许多个用户同时对数据库中的数据进行访问。

(2)**面临故障**：各种各样的故障都可能发生，必须确保数据在任何情况下都不被破坏。对于许多应用，这一点至关重要。例如，银行存款数据不能因突然停电而丢失或破坏。

(3)**数据的安全性**：防止用户对数据进行未经授权的访问。例如，我们可以允许银行客户查看自己的存款，但不允许他们修改。

(4)**数据的完整性**：防止不符合语义的数据进入数据库。例如，我们不能将负数作为学生的成绩输入到数据库中。

(5)**数据的一致性**：防止数据库进入不一致状态。有些操作必须作为一个整体(原子性)。例如，从账户A转1000元到账户B涉及将账户A的存款额减去1000元，将账户B的存款增加1000元。这两个操作要么都做，要么都不做。否则，数据库将进入不一致状态。故障可能导致两个操作中的一个完成，而另一个未完成。必须保证即使发生故障也不会影响数据库的一致性。

这些要求，加上查询条件的多样性和复杂性，使得数据查询和更新的实现很复杂。共同

的需要值得开发专门的软件程序实现，不必每个应用都写类似的程序。开发专门的软件系统管理数据，这种软件系统就是数据库管理系统，即 DBMS。

1.1.3 数据库管理系统



数据库管理系统 (Database Management System, DBMS) 是位于用户和计算机操作系统之间的数据管理软件，专门用来管理数据，是计算机系统的基础软件。DBMS 提供如下功能。

- (1) 数据定义：提供数据定义语言 (Data Definition Language, DDL)，用于定义数据库中的数据对象和它们的结构。
- (2) 数据操作：提供数据操作语言 (Data Manipulation Language, DML)，用于操作数据，实现对数据库的基本操作 (查询、插入、删除和修改)。
- (3) 事务管理和运行管理：统一管理数据、控制对数据的并发访问，保证数据的安全性、完整性，确保故障时数据库中数据不被破坏，并且能够恢复到一致状态。
- (4) 数据存储和查询处理：确定数据的物理组织和存取方式，提供数据的持久存储和有效访问；确定查询处理方法，优化查询处理过程。
- (5) 数据库的建立和维护：提供实用程序，完成数据库数据批量装载、数据库转储、介质故障恢复、数据库的重组和性能监测等。
- (6) 其他功能：包括 DBMS 与其他软件通信、异构数据库之间数据转换和互操作等。

1.1.4 数据库系统

数据库系统由数据库、DBMS (及其开发工具)、应用系统和数据库管理员组成，如图 1.2 所示。

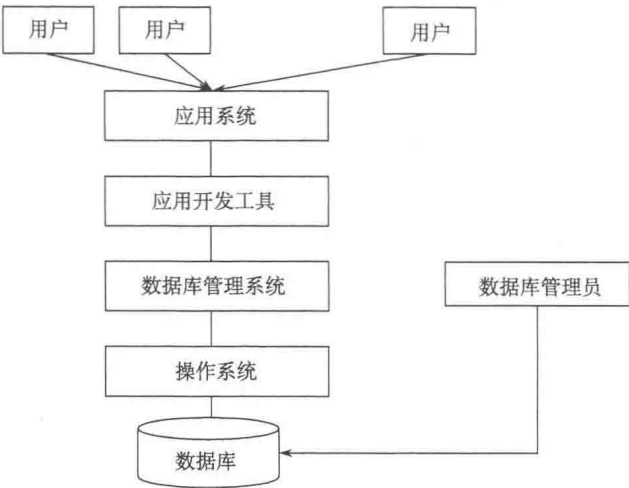


图 1.2 数据库系统的构成

- 使用数据库进行信息管理具有如下优点。
- (1) 数据整体结构化：在数据库中，数据的组织面向整个机构、面向所有可能的应用，而不是某个具体部门或某个特定的应用。数据结构化是整体结构化，数据结构不仅描述现实世界的对象，而且描述对象之间的联系。
 - (2) 数据共享：数据库中的数据面向整个机构组织使得它能够更好地被多个用户、多个