

 高等教育规划教材

数据库原理与 SQL Server 2012 应用教程

第 2 版

程云志 张 勇 赵艳忠 等编著



提供电子教案和习题解答

下载网址 <http://www.cmpedu.com>



机械工业出版社
CHINA MACHINE PRESS

高等教育规划教材

数据库原理与 SQL Server 2012 应用教程

第2版

程云志 张 勇 赵艳忠 等编著



机械工业出版社

本书全面讲述了数据库的原理与 SQL Server 2012 的应用。本书首先介绍了数据库系统的基本概念和理论, 以及数据库的设计方法等。然后以 SQL Server 2012 数据库管理系统为教学开发平台, 详细介绍了 SQL Server 2012 的基本知识, 数据库操作, 表和表数据操作, T-SQL 语言, 数据查询, 数据完整性、规则和索引, 视图和用户定义函数, 存储过程、触发器和游标, 系统安全管理, 事务、批、锁和作业, 数据库的备份还原与导入/导出, 最后讲述了 VB 2010、C# 2010、ASP.NET 数据库应用程序的开发基础。

本书可作为大学本科、专科、高等职业院校和软件职业技术学院的数据库以及数据库应用课程教材, 也可作为初学者学习数据库的入门教材, 以及数据库应用系统开发人员的技术参考书。

本书配有电子教案, 需要的教师可登录 www.cmpedu.com 免费注册, 审核通过后下载, 或联系编辑索取 (QQ: 2966938356, 电话: 010-88379739)。

图书在版编目 (CIP) 数据

数据库原理与 SQL Server 2012 应用教程/程云志等编著. —2 版. —北京: 机械工业出版社, 2015. 4

高等教育规划教材

ISBN 978-7-111-50122-0

I. ①数… II. ①程… III. ①关系数据库系统 IV. ①TP311.138

中国版本图书馆 CIP 数据核字 (2015) 第 090373 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 和庆娣 责任校对: 张艳霞

责任印制: 李 洋

北京振兴源印务有限公司印刷

2015 年 6 月第 2 版 · 第 1 次印刷

184mm×260mm · 20 印张 · 495 千字

0001—3000 册

标准书号: ISBN 978-7-111-50122-0

定价: 39.90 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: (010) 88379833

机工官网: www.cmpbook.com

读者购书热线: (010) 88379649

机工官博: weibo.com/cmp1952

教育服务网: www.cmpedu.com

封面无防伪标均为盗版

金书网: www.golden-book.com

出版说明

当前,我国正处在加快转变经济发展方式、推动产业转型升级的关键时期。为经济转型升级提供高层次人才,是高等院校最重要的历史使命和战略任务之一。高等教育要培养基础性、学术型人才,但更重要的是加大力度培养多规格、多样化的应用型、复合型人才。

为顺应高等教育迅猛发展的趋势,配合高等院校的教学改革,满足高质量高校教材的迫切需求,机械工业出版社邀请了全国多所高等院校的专家、一线教师及教务部门,通过充分的调研和讨论,针对相关课程的特点,总结教学中的实践经验,组织出版了这套“高等教育规划教材”。

本套教材具有以下特点:

1) 符合高等院校各专业人才的培养目标及课程体系的设置,注重培养学生的应用能力,加大案例篇幅或实训内容,强调知识、能力与素质的综合训练。

2) 针对多数学生的学习特点,采用通俗易懂的方法讲解知识,逻辑性强、层次分明、叙述准确而精炼、图文并茂,使学生可以快速掌握,学以致用。

3) 凝结一线骨干教师的课程改革和教学研究成果,融合先进的教学理念,在教学内容和方法上做出创新。

4) 为了体现建设“立体化”精品教材的宗旨,本套教材为主干课程配备了电子教案、学习与上机指导、习题解答、源代码或源程序、教学大纲、课程设计和毕业设计指导等资源。

5) 注重教材的实用性、通用性,适合各类高等院校、高等职业学校及相关院校的教学,也可作为各类培训班教材和自学用书。

欢迎教育界的专家和老师提出宝贵的意见和建议。衷心感谢广大教育工作者和读者的支持与帮助!

机械工业出版社

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，也是应用最为广泛的技术之一，它已经成为计算机信息系统与应用系统的核心技术和重要基础，已经广泛应用于各种领域，小到工资管理、人事管理、学籍管理，大到企业级的信息管理、银行系统管理等。

本书以关系数据库系统为核心，全面、系统地阐述了数据库系统的基本概念、基本原理和 SQL Server 2012 数据库管理系统的应用技术。主要内容包括：数据库系统基本概念、理论和数据库设计，并详细地介绍了现今最流行的关系型数据库管理系统——SQL Server 2012。通过大量的实例，全面、深入地介绍了 SQL Server 2012 数据库管理系统软件的安装、配置、操作，以及 SQL Server 2012 数据库操作，表和表数据操作，T-SQL 语言，数据查询，数据完整性、规则和索引，视图和用户定义函数，存储过程、触发器和游标，系统安全管理，事务、批、锁和作业，数据库的备份还原与导入/导出，最后讲述了 VB 2010、C# 2010、ASP.NET 数据库应用程序的开发基础。本书有下列特点：

- 1) 以 SQL Server 2012 数据库管理系统中文版为教学和开发平台。
- 2) 体系完整，内容丰富，符合大学计算机专业和非计算机专业对数据库知识的要求。
- 3) 本书首先介绍了数据库的基本概念，为以后的学习奠定了较好的理论基础。
- 4) 书中引用了大量的实例，更加突出实用性，并配以详细的操作步骤和抓图。

本书作者从事大学本科计算机专业教学，不仅具有丰富的教学经验，同时还具有多年的数据库开发经验。作者依据长期的教学经验，深知数据库原理的主要知识点、重点与难点，以及读者对数据库应用中最感兴趣的方面，逐渐形成了本书严谨的、适合于学习的结构体系。

本书内容丰富、结构新颖、系统性与实用性强，注重理论教学和实践教学相结合，叙述准确而精炼，图文并茂，具体而且直观。既可作为高等学校计算机专业、信息管理与信息系统专业及非计算机专业本科数据库应用课程的教学用书，也可作为从事信息领域工作的科技人员的自学参考书。对于计算机应用人员和计算机爱好者，本书也是一本实用的工具书。

本书由程云志、张勇、赵艳忠等编著。第 1、2 章由程云志编写，第 3、5 由张勇编写，第 4、6 章由郑珂编写，第 7、10 章由赵艳忠编写，第 8、15 章由曹志伟编写，第 9 章由李晗编写，第 11 由郑磊编写，第 12、16 章由臧国轻编写，第 14、17 章由王晶姝编写，第 13 章、教学资源的制作等由刘瑞新、吴蕾、史洪智、梁宏伟、韩顺友、党兰学、段金卯、蔡军、张敬来、乔家君、马同森、程遂营、刘克纯、徐维维、谢紫安、缪丽丽、徐云林、骆秋容、田金雨、王如雪、曹媚珠、张曼完成。全书由刘瑞新教授审阅统稿。

因编者水平有限，书中疏漏之处在所难免，敬请读者批评指正。

编 者

目 录

出版说明

前言

第 1 章 数据库系统概述	1	3.1.1 函数依赖	30
1.1 数据库系统简介	1	3.1.2 范式	31
1.1.1 数据库技术的发展历史	1	3.2 数据库设计概述	35
1.1.2 数据库系统的基本概念	3	3.2.1 数据库设计的特点	35
1.2 数据库系统结构	5	3.2.2 数据库设计的步骤	36
1.2.1 数据库系统的三级模式结构	5	3.3 需求分析阶段	36
1.2.2 数据库系统的二级映像	6	3.4 概念结构设计阶段	36
1.3 习题	6	3.4.1 概念结构设计的任务	37
第 2 章 数据模型	7	3.4.2 概念结构设计的步骤	37
2.1 信息的三种世界	7	3.5 逻辑结构设计阶段	38
2.1.1 现实世界	7	3.5.1 逻辑结构设计的任务	38
2.1.2 信息世界	8	3.5.2 逻辑结构设计的步骤	38
2.1.3 计算机世界	8	3.6 物理结构设计阶段	39
2.1.4 三种世界的转换	9	3.6.1 物理结构设计的任务	39
2.2 概念模型	9	3.6.2 物理结构设计的步骤	39
2.2.1 概念模型的基本概念	9	3.7 数据库实施阶段	40
2.2.2 概念模型表示	11	3.8 数据库运行和维护	40
2.3 数据模型	13	3.9 数据库设计实例	41
2.3.1 数据模型的基本概念	14	3.9.1 学生成绩管理数据库设计	41
2.3.2 常用的数据模型	14	3.9.2 图书出版管理数据库设计	42
2.4 关系数据库	19	3.10 关系数据库管理系统	43
2.4.1 关系模型的组成	19	3.11 习题	45
2.4.2 关系的数学定义	20	第 4 章 SQL Server 2012 基本知识	46
2.4.3 关系代数	21	4.1 SQL Server 2012 发展简介	46
2.4.4 传统的集合运算	22	4.2 SQL Server 2012 的版本	46
2.4.5 专门的关系运算	24	4.2.1 SQL Server 2012 版本的	
2.5 关系查询优化	27	分类	46
2.6 习题	27	4.2.2 SQL Server 2012 Standard 功能	
第 3 章 数据库设计	30	简介	47
3.1 规范化	30	4.3 SQL Server 2012 Standard 的	

安装与卸载	48	5.6 数据库的收缩	80
4.3.1 安装 SQL Server 2012 Standard 的系统需求	48	5.6.1 手动收缩	81
4.3.2 SQL Server 2012 的安装	48	5.6.2 自动收缩	83
4.3.3 SQL Server 2012 的卸载	53	5.7 移动数据库	84
4.4 SQL Server 2012 组件和管理 工具	54	5.8 数据库快照	84
4.4.1 服务器组件	54	5.8.1 数据库快照的优点	84
4.4.2 管理工具	54	5.8.2 数据库快照的操作	85
4.4.3 文档	55	5.9 习题	86
4.5 SQL Server 2012 服务器的 管理	56	第 6 章 表和表数据操作	87
4.5.1 启动/停止服务器	56	6.1 表概念	87
4.5.2 服务器的注册	58	6.1.1 表结构	87
4.6 习题	59	6.1.2 表类型	88
第 5 章 数据库操作	60	6.1.3 数据类型	88
5.1 数据库基本概念	60	6.2 创建表	92
5.1.1 物理数据库	60	6.2.1 管理工具界面方式创建表	92
5.1.2 逻辑数据库	61	6.2.2 命令行方式创建表	95
5.1.3 SQL Server 2012 的系统数据库 和用户数据库	63	6.3 修改表	98
5.1.4 报表服务器和报表数据库	64	6.3.1 管理工具界面方式修改表	98
5.2 创建数据库	64	6.3.2 命令行方式修改表	98
5.2.1 管理工具界面方式创建 数据库	64	6.4 删除表	100
5.2.2 命令行方式创建数据库	67	6.5 表数据操作	100
5.3 修改数据库	72	6.5.1 管理工具界面方式操作表 数据	101
5.3.1 管理工具界面方式修改 数据库	72	6.5.2 命令行方式操作表数据	102
5.3.2 命令行方式修改数据库	73	6.6 习题	106
5.4 删除数据库	76	第 7 章 T-SQL 语言	108
5.4.1 管理工具界面方式删除 数据库	77	7.1 SQL 语言基本概念	108
5.4.2 命令行方式删除数据库	78	7.1.1 T-SQL 语言简介	108
5.5 数据库的分离和附加	78	7.1.2 T-SQL 语言的语约定	109
5.5.1 分离数据库	78	7.1.3 标识符	110
5.5.2 附加数据库	79	7.1.4 常量和变量	110
		7.1.5 注释	114
		7.1.6 运算符	115
		7.1.7 函数	115
		7.1.8 表达式	122
		7.2 流程控制语句	126
		7.2.1 SET 语句	126

7.2.2	BEGIN...END 语句	126	9.2.4	绑定规则	177
7.2.3	IF...ELSE 语句	127	9.2.5	解除规则	177
7.2.4	WHILE、BREAK、 CONTINUE 语句	128	9.2.6	删除规则	178
7.2.5	RETURN 语句	129	9.3	索引	178
7.2.6	WAITFOR 语句	129	9.3.1	索引的分类	179
7.2.7	GOTO 语句	129	9.3.2	创建索引	181
7.2.8	TRY...CATCH 语句	130	9.3.3	查看索引	183
7.2.9	GO 语句	132	9.3.4	修改索引	184
7.2.10	EXECUTE 语句	132	9.3.5	删除索引	185
7.2.11	T-SQL 语句的解析、编译 和执行	132	9.3.6	其他类型索引	186
7.3	数据定义语句	132	9.3.7	优化索引	190
7.4	习题	133	9.4	数据库关系图	192
第 8 章	数据查询	134	9.5	习题	196
8.1	数据查询语句	134	第 10 章	视图和用户定义函数	197
8.1.1	投影列	135	10.1	视图	197
8.1.2	选择行	140	10.1.1	视图概述	197
8.1.3	连接	146	10.1.2	视图的类型	198
8.2	数据汇总	150	10.1.3	创建视图准则	199
8.3	排序	151	10.1.4	创建视图	199
8.4	分组	152	10.1.5	查询视图	202
8.5	子查询	156	10.1.6	可更新视图	202
8.5.1	无关系查询	156	10.1.7	修改视图定义	204
8.5.2	相关子查询	159	10.1.8	删除视图	205
8.6	集合操作	161	10.2	用户定义函数	205
8.7	存储查询结果	164	10.2.1	标量值函数	206
8.8	习题	165	10.2.2	内嵌表值函数	209
第 9 章	数据完整性、规则和索引	166	10.2.3	多语句表值函数	210
9.1	数据完整性	166	10.2.4	修改和重命名用户定义 函数	211
9.1.1	实体完整性	166	10.2.5	删除用户定义函数	212
9.1.2	域完整性	170	10.3	习题	212
9.1.3	引用完整性	173	第 11 章	存储过程、触发器和游标	213
9.2	规则	176	11.1	存储过程	213
9.2.1	规则的概念	176	11.1.1	存储过程概述	213
9.2.2	创建规则	176	11.1.2	存储过程的类型	214
9.2.3	查看规则	177	11.1.3	创建存储过程	215
			11.1.4	调用存储过程	218

11.1.5	获取存储过程信息	219	13.1.4	事务和批	265
11.1.6	修改和重命名存储过程	221	13.1.5	事务隔离级	266
11.1.7	重新编译存储过程	222	13.2	锁	267
11.1.8	删除存储过程	222	13.2.1	锁概述	267
11.2	触发器	223	13.2.2	锁的模式	268
11.2.1	触发器概述	223	13.2.3	锁的信息	269
11.2.2	触发器的类型	223	13.2.4	死锁及处理	271
11.2.3	触发器的设计规则	225	13.3	数据库引擎优化顾问	272
11.2.4	使用触发器	226	13.3.1	数据库引擎优化顾问 概述	272
11.2.5	启用、禁用和删除 触发器	228	13.3.2	数据库引擎优化顾问的 使用	272
11.2.6	嵌套触发器和递归 触发器	228	13.4	作业	274
11.3	游标	229	13.5	习题	275
11.3.1	游标概述	230	第 14 章 数据库的备份还原与 导入/导出		
11.3.2	游标的类型	231	14.1	数据库的备份还原	276
11.3.3	游标的使用	232	14.1.1	备份还原概述	276
11.4	习题	238	14.1.2	恢复模式	277
第 12 章 系统安全管理			14.1.3	数据库备份	278
12.1	身份验证模式	239	14.1.4	数据库还原	281
12.1.1	身份验证概述	239	14.2	数据库的导入/导出	284
12.1.2	身份验证方式设置	241	14.2.1	数据库表数据导出	284
12.2	账号和角色	242	14.2.2	数据库表数据导入	287
12.2.1	账号	242	14.3	习题	289
12.2.2	角色	247	第 15 章 VB 2010/SQL Server 2012 开发		
12.3	授权的主体	251	15.1	ADO.NET 技术概述	290
12.4	授权的安全对象	251	15.1.1	ADO.NET 模型	290
12.5	权限操作	253	15.1.2	ADO.NET 结构	291
12.5.1	在 SQL Server Management Studio 中设置权限	253	15.1.3	数据控件	292
12.5.2	T-SQL 语句授权	254	15.1.4	数据访问步骤控件	292
12.6	习题	258	15.2	数据库应用程序实例	294
第 13 章 事务、批、锁和作业			15.2.1	数据库设计	294
13.1	事务	259	15.2.2	系统设计	294
13.1.1	事务概述	259	15.2.3	程序设计	295
13.1.2	事务的类型	260	15.3	习题	298
13.1.3	事务处理语句	262			

第 16 章 C# 2010/SQL Server 2012

开发	299
16.1 C#语言简介	299
16.2 数据库应用程序实例	299
16.3 语言集成查询 (LINQ)	302
16.3.1 LINQ 简介	302
16.3.2 LINQ 的使用	303
16.3.3 LINQ 在 C# 中的使用	303

16.3.4 对象关系设计器

(O/R 设计器)	304
-----------------	-----

16.4 习题	305
---------------	-----

第 17 章 ASP.NET/SQL Server 2012

开发	306
17.1 ASP.NET 简介	306
17.2 数据库应用程序实例	307
17.3 习题	310

第 1 章 数据库系统概述

数据库技术是计算机科学的重要分支，是信息系统的一个核心技术，是使用计算机对各种信息、数据进行收集、管理的必备知识。数据库技术所研究的问题就是如何科学地组织和存储数据，如何高效地获取和处理数据。

本章主要介绍数据库的基本知识，包括数据库的发展历史、概念描述、体系结构等。

1.1 数据库系统简介

数据库技术是随着信息社会对数据处理任务的需要而产生的。随着社会对数据处理任务的要求不断提高，数据库也随之产生并不断发展。数据库的诞生和发展给计算机信息管理带来了一场巨大的革命。数据库技术从问世到现在，形成了坚实的理论基础、成熟的商业产品和广泛的应用领域，吸引了越来越多的研究者加入。对于一个国家来说，数据库的建设规模、数据库信息量的大小和使用频度已经成为衡量这个国家信息化程度的重要标志。

1.1.1 数据库技术的发展历史

在 20 世纪 60 年代，随着数据处理自动化的发展，数据库技术应运而生。在计算机应用领域中，数据处理越来越占主导地位，数据库技术的应用也越来越广泛。数据库是数据管理的产物。数据管理是数据库的核心任务，内容包括对数据的分类、组织、编码、储存、检索和维护。从数据管理的角度看，数据库技术到目前共经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

人工管理阶段是指计算机诞生的初期，即 20 世纪 50 年代后期之前。这个时期的计算机主要用于科学计算。从硬件看，起初没有磁盘等直接存取的存储设备，后来可以存储在磁带上；从软件看，没有操作系统和管理数据的软件，数据处理方式是批处理。那时的数据管理非常简单，通过大量的分类、比较和表格绘制的机器运行数百万穿孔卡片或读写磁带来进行数据的处理和存储。

这个时期的数据管理具有以下四个特点。

1) 数据基本不保存。该时期的计算机主要应用于科学计算，由于技术限制，一般不需要将数据长期保存，只是在计算某一课题时将数据输入，用完后不保存原始数据，也不保存计算结果。

2) 没有对数据进行管理的软件系统。程序员不仅要规定数据的逻辑结构，而且还要在程序中设计物理结构，包括存储结构、存取方法、输入/输出方式等。因此，程序中存取数据的子程序随着存储结构的改变而改变，数据与程序不具有一致性。卡片或磁带都只能顺序读取。

3) 没有文件的概念。数据的组织方式必须由程序员自行设计。数据是面向应用的，一组数据只能对应一个程序。即使两个程序用到相同的数据，也必须各自定义、各自组织，数据无法共享、无法相互利用和互相参照，程序和程序之间有大量重复的数据。因此，造成程序与程

序之间有大量的冗余数据。

4) 数据不具有独立性。当数据的逻辑结构或物理结构发生变化后, 必须对应用程序做相应的修改, 这就加重了程序员的负担。

人工管理阶段应用程序与数据之间的对应关系如图 1-1 所示。

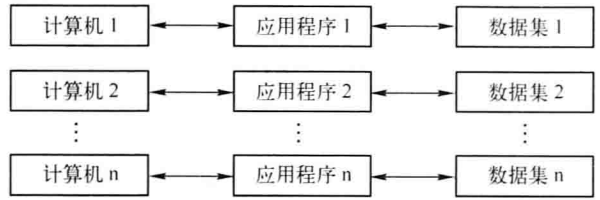


图 1-1 人工管理阶段应用程序与数据之间的对应关系

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中期, 这时在硬件方面, 外存储器有了磁盘、磁鼓等直接存取的存储设备; 在软件方面, 操作系统中已经有了专门用于管理数据的软件, 称为文件系统。处理方式上不仅有了批处理, 还有能够联机的实时处理。

这个时期的数据管理具有以下四个特点。

1) 数据可以长期保存。由于计算机大量用于数据处理, 经常对文件进行查询、修改、插入和删除等操作, 所以, 数据需要长期保留在外存储器中, 便于反复操作。

2) 由文件系统管理数据。操作系统提供了文件管理功能和访问文件的存取方法, 程序和数据之间有了数据存取的接口, 程序可以通过文件名和数据打交道, 不必再寻找数据的物理存放位置。至此, 数据有了物理结构和逻辑结构的区别, 但此时程序和数据之间的独立性尚不充分。

3) 文件的形式已经多样化。由于已经有了直接存取的存储设备, 文件也就不再局限于顺序文件, 还有了索引文件、链表文件等。因而, 对文件的访问可以是顺序访问, 也可以是直接访问。

4) 数据具有一定的独立性。文件系统文件是为某一特定应用服务的, 因此, 系统不容易扩充, 仍旧有大量的冗余数据, 数据共享性差, 独立性差。

文件系统阶段应用程序与数据之间的对应关系如图 1-2 所示。

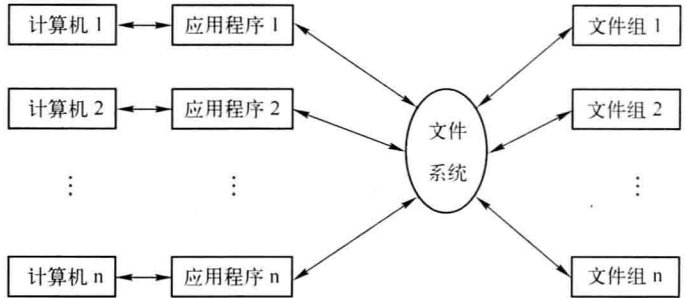


图 1-2 文件系统阶段应用程序与数据之间的对应关系

3. 数据库系统阶段

数据库系统的萌芽出现于 20 世纪 60 年代。当时的计算机开始广泛地应用于数据管理, 对数据的共享提出了越来越高的要求。传统的文件系统已经不能满足人们的需要。能够统一管理和共享数据的数据库管理系统 (DataBase Management System, DBMS) 便应运而生。

在这个阶段中, 数据库中的数据不再是面向某个应用或某个程序, 而是面向整个企业或整个应用的, 处理的数据量急剧增长。这时在硬件方面, 磁盘容量越来越大, 读写速度越来越快; 在软件方面, 编制越来越复杂, 功能越来越强大。处理方式上, 联机处理要求更多。

这个时期的数据管理具有以下四个特点。

1) 采用复杂的结构化的数据模型。数据库系统不仅要描述数据本身, 还要描述数据之间的联系。这种联系是通过存取路径来实现的。数据结构化是数据库与文件系统的根本区别。

2) 较高的数据独立性。数据和程序彼此独立, 数据存储结构的变化尽量不影响用户程序的使用。数据与程序的独立把数据的定义从程序中分离出去, 加上数据由数据库管理系统管

理，从而简化了应用程序的编制和程序员的负担。

3) 最低的冗余度。数据库系统中的重复数据被减少到最低程度，这样，在有限的存储空间内可以存放更多的数据，并减少存取时间。数据冗余度低，共享性高，易于扩充。

4) 数据控制功能。数据库系统具有数据的安全性，以防止数据的丢失和被非法使用；具有数据的完整性，以保护数据的正确、有效和相容；具有数据的并发控制，避免并发程序之间的相互干扰；具有数据的恢复功能，在数据库被破坏或数据不可靠时，系统有能力把数据库恢复到最近某个时刻的正确状态。

数据库系统阶段应用程序与数据之间的对应关系如图 1-3 所示。

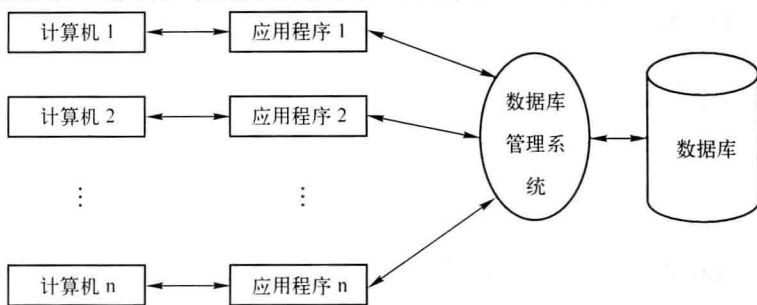


图 1-3 数据库系统阶段应用程序与数据之间的对应关系

事实上，只要有大量的信息需要处理，需要大量数据支持工作，都可以使用数据库技术。目前数据库技术应用到了社会生活中几乎所有的领域，在金融业、航空业、学校、人力资源和军事等领域，数据库已经成为不可或缺的组成部分了。

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，现有数据库已经不能完全满足需求，于是新一代数据库孕育而生。新一代的数据库支持多种数据模型，并和诸多新技术相结合，广泛应用于更多领域。

总之，随着科学技术的发展，计算机技术不断应用到各行各业，数据存储不断膨胀，对未来的数据库技术将会有更高的要求。

1.1.2 数据库系统的基本概念

数据库系统作为信息系统的核心和基础，涉及一些常用的术语和基本概念。

1. 数据

数据 (Data) 是数据库中存储的基本对象。数据不仅仅指的是具体的数字和文字，还包括图形、声音、人的身体状况记录、计算机运行情况等，这些形式的数据经过数字化后都可以存储到计算机中。因此，数据是人们用各种物理符号，把信息按一定格式记载下来的有意义的符号组合。

例如，学生李飞扬的计算机网络考试成绩为 91 分，91 就是数据。学生李飞扬还有学号、姓名、性别、出生日期、学院名称等信息。所以可这样来描述学生李飞扬：

(2012130045, 李飞扬, 男, 1994-12-5, 计算机学院)

这些描述也是数据，它们都可以经过数据化处理后被计算机识别。通过这些数据可以掌握该学生的基本信息，但却无法正确理解数据的含义，例如，数据“2012130045”的含义是学号还是身份证号，从数据中无法知道。可见，数据的形式还不能完全表达其内容，需要解释。因此，数据和关于数据的解释是不可分的。数据的解释是指对数据含义的说明，称为数据的语义，数据与其语义是不可分的。

2. 数据库

数据库 (DataBase, DB) 即存放数据的仓库, 这个仓库就是计算机存储设备。数据库中的数据并不是简单的堆积, 数据之间相互关联。严格地讲, 数据库是长期存储在计算机内、有组织的、可共享的大量数据的集合。数据库中的数据按照一定的数据模型组织、描述和存储, 具有较小的冗余度、较高的数据独立性和易扩展性, 并可为各种用户共享。

3. 数据库管理系统

数据库管理系统 (DataBase Management System, DBMS) 是专门用于管理数据库的计算机系统软件, 介于应用程序与操作系统之间, 是数据管理软件。数据库管理系统能够为数据库提供数据的定义、建立、维护、查询和统计等操作功能, 并具有对数据完整性、安全性进行控制的功能。

现今广泛使用的数据库管理系统有微软公司的 SQL Server、Access, 甲骨文公司的 Oracle、MySQL, IBM 公司的 DB2、Informix 等。

数据库管理系统的主要功能包括数据库定义功能、数据库存取功能、数据库管理功能、数据库建立维护功能。

4. 数据库系统

数据库系统 (DataBase System, DBS) 是指在计算机系统中引入了数据库后的系统, 由计算机硬件、数据库、数据库管理系统、应用程序 (Application) 和用户 (User) 构成, 即由计算机硬件、软件和使用人员构成。数据库系统是一个计算机应用系统。

计算机硬件是数据库系统的物质基础, 是存储数据库及运行数据库管理系统的硬件资源, 主要包括主机、存储设备、I/O 通道以及计算机网络环境等。

软件主要包括操作系统以及数据库管理系统本身。此外, 为了开发应用程序, 还需要各种高级语言及其编译系统, 以及各种以数据库管理系统为核心的应用开发工具软件。

数据库管理系统是负责数据库存取、维护和管理系统软件, 是数据库系统的核心, 其功能的强弱是衡量数据库系统性能优劣的主要指标。

数据库中的数据由数据库管理系统进行统一管理和控制, 用户对数据库进行的各种操作都是由数据库管理系统实现的。

应用程序是在数据库管理系统的基础上, 由用户根据应用的实际需要开发的、处理特定业务的应用程序。

用户是指管理、开发、使用数据库系统的所有人员, 通常包括数据库管理员、应用程序员和终端用户。

综上所述, 在数据库系统中, 数据库中包含的数据是存储在存储介质上的数据文件的集合; 每个用户均可使用其中的部分数据, 不同用户使用的数据可以重叠, 同一组数据可以为多个用户共享; 数据库管理系统为用户提供对数据的存储组织、操作管理功能; 用户通过数据库管理系统和应用程序实现数据库系统的操作与应用。数据库强调的是数据, 数据库管理系统强调的是系统软件, 数据库系统强调的是系统。在一般不发生混淆的情况下, 常常把数据库系统也称为数据库。

数据库系统在整个计算机系统中的地位如图 1-4 所示。

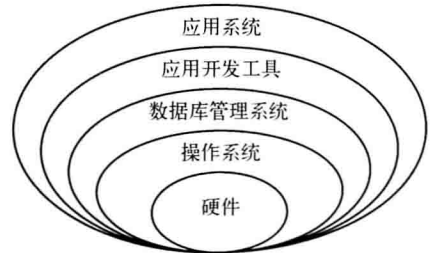


图 1-4 数据库系统在整个计算机系统中的地位

1.2 数据库系统结构

数据库系统虽然是一个庞大、复杂的系统，但它都有一个总的框架。虽然数据库系统软件产品众多，且建立在不同的操作系统之上，但从数据库系统管理角度看，数据库系统通常采用三级模式结构。

1.2.1 数据库系统的三级模式结构

数据库系统通常采用三级模式结构：外模式、模式和内模式。三级模式结构如图 1-5 所示。

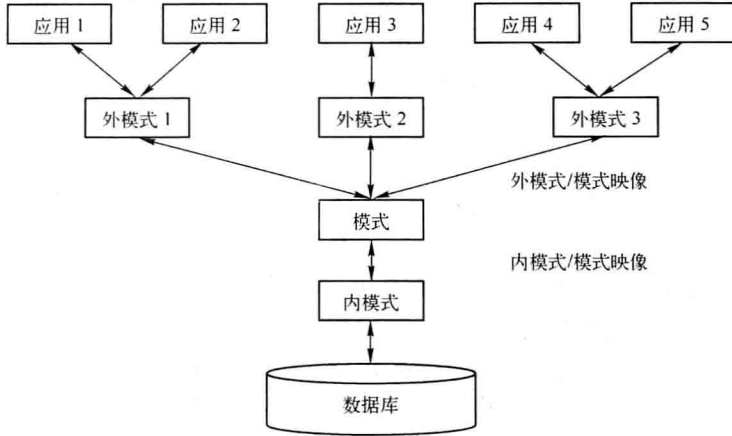


图 1-5 数据库系统的三级模式结构

1. 模式

模式也称为逻辑模式，是对数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图。

它是数据库系统模式结构的中间层，既不涉及数据的物理存储细节和硬件环境，也与具体的应用程序和开发工具无关。模式实际上是数据库数据在逻辑级上的视图，一个数据库只能有一个模式。数据库模式以某一种数据模型为基础，综合考虑了所有用户的需求，并将这些需求有机地整合成一个逻辑整体。模式只是对数据库结构的一种描述，而不是数据库本身，是装配数据的一个框架。例如，数据记录由哪些数据项构成，数据项的名字、类型、取值范围等，而且要定义数据之间的联系，定义与数据有关的安全性、完整性要求。

数据库系统提供数据模式描述语言（Data Description Language, DDL）来严格地表示这些内容。

2. 外模式

外模式也称为子模式或用户模式，是数据库用户看到的数据视图，是与某一应用有关的数据的逻辑表示。

外模式是模式的子集，它是各个用户的数据视图。由于不同的用户其需求不同，看待数据的方式不同，对数据的要求不同，使用的程序设计语言也不同，因此不同用户的外模式描述是不同的。即使对模式中同一数据，在外模式中的结构、类型、保密级别等都可以不同。一个数据库可以有多个外模式。

数据库系统提供外模式描述语言来描述用户数据视图。

3. 内模式

内模式也称为存储模式，是数据在数据库系统内部的表示或底层描述，即对数据库物理结构和存储方式的描述。

一个数据库只能有一个内模式。例如，记录的存储方式是顺序存储、链式存储还是按哈希（Hash）方式存储；索引按照什么方式组织；数据是否压缩，是否加密等。

数据库系统提供内模式描述语言来描述数据库的物理存储。

1.2.2 数据库系统的二级映像

数据库系统的三级模式是对数据的三个抽象级别，它把数据的具体组织留给 DBMS 管理，使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的表示和存储。为了实现这三个层次上的联系和转换，数据库系统在这三级模式中提供了两层映像：外模式/模式的映像和模式/内模式的映像。

1. 外模式/模式的映像

模式描述的是数据的全局逻辑结构，外模式描述的是数据的局部逻辑结构。对于每一个外模式，数据库都有一个外模式/模式的映像，它定义并保证了外模式与数据模式之间的对应关系。这些映像定义通常包含在各自的外模式中。

当模式改变时（例如，增加新的关系、新的属性、改变属性数据类型等），外模式/模式的映像要做相应的改变，由数据库管理员（Data Base Administrator, DBA）负责，以保证外模式保持不变。应用程序是根据数据的外模式编写的，因此应用程序不必修改，保证了数据与程序的逻辑独立性，即数据的逻辑独立性。

2. 模式/内模式的映像

数据库的内模式依赖于它的全局逻辑结构，即模式。由于一个数据库只有一个模式，也只有一个内模式，所以模式/内模式的映像也是唯一的。它定义并保证了数据的逻辑模式与内模式之间的对应关系。

当数据库的存储结构改变了，模式/内模式的映像也必须作相应的修改（仍由 DBA 负责），使得模式保持不变，保证了数据与程序的物理独立性，即数据的物理独立性。

正是由于上述二级映像功能，才使得数据库系统中的数据具有较高的逻辑独立性和物理独立性。二级映像保证了数据库外模式的稳定性，从而从底层保证了应用程序的稳定性。数据与程序之间的独立性使得数据的存取由 DBMS 管理，用户不必考虑存取路径等细节，从而简化了应用程序的编制，大大减少了应用程序的维护和修改。

1.3 习题

1. 数据库的发展历史分为哪几个阶段？各有什么特点？
2. 简述数据、数据库、数据库管理系统、数据库系统的概念。
3. 使用数据库系统有什么好处？
4. 试述数据库系统的三级模式结构和二级映像的特点。
5. 什么是数据与程序的逻辑独立性？什么是数据与程序的物理独立性？
6. 举例说明，在实际工作生活中，有哪些单位部门使用数据库，这些数据库所起的作用如何。

第2章 数据模型

数据库不仅反映数据本身所表达的内容，而且还反映数据之间的联系。由于计算机不能直接处理现实世界中的具体事物，所以人们必须事先将具体事物转换成计算机能够处理的数据。在数据库系统的形式化结构中如何抽象、表示、处理现实世界中的信息和数据呢？这就是数据库的数据模型。通过数据模型这个工具来抽象、表示和处理现实世界中的信息和数据。

本章主要介绍信息的三种世界的概念，概念模型（E-R 图）和数据模型的组成，三种常用的数据模型。

2.1 信息的三种世界

在信息社会中，信息成为重要的资源，在国民经济中占据主导地位，并构成社会信息化的物质基础。以计算机、微电子和通信技术为主的信息技术革命是社会信息化的动力源泉，从根本上改变了人们的生活方式、行为方式和价值观念。

信息（Information）就是通过各种方式传播的能被感受的声音、文字、图像、符号等。简单地说，信息就是新的、有用的事实和知识。

信息需要载体才能表示，例如，今天的温度信息用数字“30”表示，“30”就是数据。对每个人来说，“信息”和“数据”是两种非常重要的东西。“信息”可以告诉人们有用的事实和知识，“数据”可以更有效地表示、存储和抽取信息。信息和数据是数据库管理的基本内容和对象。信息是现实世界事物状况的反映，通过加工，它可以用一系列数据来表示。

不同的领域，数据的描述也有所不同。人们在研究和处理数据的过程中，常常把数据的转换分为三个领域——现实世界、信息世界、计算机世界。这三个世界间的转换过程就是将客观现实的信息反映到计算机数据库中的过程。

2.1.1 现实世界

现实世界（Real World）就是人们所能看到的、接触到的世界。信息的现实世界是指人们要管理的客观存在的各种事物、事物之间的相互联系及事物的发生、变化过程。客观存在的世界就是现实世界，它不依赖于人们的思想。现实世界存在无数事物，每一个客观存在的事物可以看作是一个个体，每个个体都有属于自己的特征。比如，计算机有价格、品牌、型号等特征。而不同的人，只会关心其中的一部分特征，并且一定领域内的个体有着相同的特征。用户为了某种需要，必须将现实世界中的部分需求用数据库实现。此时，它设定了需求及边界条件，这为整个转换提供了客观基础与初始启动环境。人们所见到的客观世界中的划定边界的一个部分环境就是现实世界。现实世界主要涉及以下三个概念。

1. 实体（Entity）

现实世界中存在的可以相互区分的客观事物或概念称为实体。例如，计算机、电视、大象、人。