

21世纪高等学校计算机类课程创新规划教材 · 微课版



数据库原理及应用

— SQL Server 2012

◎ 刘金岭 冯万利 张有东 编著

8小时
微课视频

236个
经典实例

100道
习题

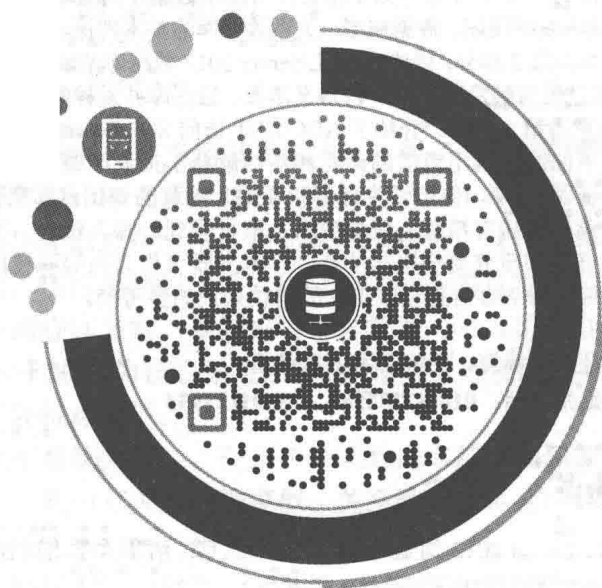
15个
上机实验

2个
课程设计

清华大学出版社



21世纪高等学校计算机类课程创新规划教材 · 微课版



数据库原理及应用

—— SQL Server 2012

◎ 刘金岭 冯万利 张有东 编著



内 容 简 介

本书是为高等院校应用型本科计算机专业或相关专业精心编写的一本数据库课程教学用书,它以 SQL Server 2012 为核心系统,较完整地论述了数据库系统的基本概念、基本原理和 SQL Server 的应用技术。

本书分为 3 个部分,共 12 章。第 1~3 章为第 1 部分,讲述了数据库的基本理论知识及数据库设计的相关技术,其内容主要包括数据库系统概述、数据模型、关系数据库的基本理论、关系模式的规范化以及数据库设计思想和方法;第 4~11 章为第 2 部分,讲述了 SQL Server 2012 数据库基础、数据库与表管理、SELECT 数据查询、视图与索引、存储过程与触发器及用户自定义函数、数据库并发控制、数据库安全管理及数据库的备份与恢复等内容。第 12 章为第 3 部分,介绍了 ADO.NET 访问 SQL Server 数据库的简单应用。

本书结合应用型本科学生的特点,用通俗的语言和实例解释了抽象的概念,将抽象概念融合到具体的数据库管理系统 SQL Server 2012 中,便于学生理解和掌握。每章的知识点都配有视频讲解。

本书在编写过程中,力求做到语言精练、概念清晰、取材合理、深入浅出、突出应用,为读者进一步从事数据库系统的应用、开发和研究奠定坚实的基础。本书既可作为高等院校应用型本科有关专业的数据库原理及应用教材,也可作为从事信息领域工作的科技人员的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理及应用:SQL Server 2012/刘金岭等编著. —北京:清华大学出版社,2017
(21 世纪高等学校计算机类课程创新规划教材·微课版)

ISBN 978-7-302-46651-2

I. ①数… II. ①刘… III. ①关系数据库系统—高等学校—教材 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2017)第 032997 号

责任编辑:魏江江 李 晔

封面设计:刘 键

责任校对:时翠兰

责任印制:宋 林

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:20.25

字 数:498 千字

版 次:2017 年 7 月第 1 版

印 次:2017 年 7 月第 1 次印刷

印 数:1~2000

定 价:39.50 元

前 言

2009年,结合当时的教学和应用开发需求,出版了《数据库原理及应用》教材,受到了广大应用型本科师生和计算机爱好者的欢迎,并于2011年被评为江苏省高等院校精品教材。同时开展了“数据库原理及应用”精品课程的建设,获江苏省成教精品课程。随着数据库技术的不断升级,应用越来越广泛,结合广大师生的反馈意见以及新的教学和应用开发经验,制订了全新的修订方案,重新编写了本教材,并入选江苏省重点规划教材项目。

本书以SQL Server 2012为核心系统,在《数据库原理及应用》基础上进行了全面的调整、修改和优化,主要特点如下:

(1)以关系数据库系统为核心。在系统论述数据库基本知识的基础上,着重讨论了关系数据库的基本理论,其中对关系数据模型、关系数据库体系结构、关系规范化理论等进行了简要的讲解。

(2)本书对传统数据库的内容进行了精简,如对层次数据库、网状数据库仅对其模型做了简要介绍,删除了一些与操作系统联系较密切的存储理论等。

(3)为了反映当前数据库领域的新技术、新水平和新趋势,本书简要介绍了处理大规模数据的数据库系统,如分布式数据库系统、并行数据库、NoSQL数据库、云数据库和XML数据库的相关概念。

(4)强化了数据库设计技术。第3章在数据库设计理论的基础上,结合作者多年的数据库开发经验,介绍了一些行之有效的数据库设计与开发方法和技巧。

(5)将抽象理论融合到具体模型中。结合具体的SQL Server 2012数据库管理系统,讲解了数据库的一些管理技术和应用,如数据库的完整性约束、存储过程与触发器、数据库并发控制、数据库的安全与保护、备份与恢复等。使读者在学习理论的同时了解了具体的应用,也为读者维护管理大中型数据库系统打下坚实的基础。

(6)在内容选取、章节安排、难易程度、例子选取等方面充分考虑到理论教学和实践教学的需要。力求使教材概念准确、清晰、重点明确,内容精练,便于取舍。每章均配有习题,便于教学。

为了方便课程的学习及数据库技术的应用,作者还组织编写了辅助教材《数据库原理及应用实验与课程设计指导——SQL Server 2012》,作为读者学习本课程时的实践用书。另外清华大学出版社的网站上还有本书的习题参考答案和教学课件等教学材料供教师教学参考。

清华大学出版社魏江江老师对本书的编写给出了指导性的意见,在此表示衷心的感谢。

由于编者水平有限,书中疏漏谬误之处在所难免,殷切希望广大读者批评指正。

作者简介:

刘金岭:教授,男,主持完成了多项教改项目,获得多项教学成果。主要进行数据库

原理及应用、Oracle 数据库、Web 数据库应用技术等课程的教学。

冯万利：副教授，男，主要进行数据库原理及应用、网站开发技术、ASP.NET 程序设计、Object Oriented Programming Using C#等课程的教学。

张有东：男，博士，教授，长期进行数据库的教学和研究。

编者

2017年3月

目 录

第 1 章 数据库系统概述	1
1.1 数据管理技术的发展.....	1
1.1.1 数据和数据管理.....	1
1.1.2 数据管理发展的三个阶段.....	2
1.2 数据库系统.....	5
1.2.1 数据库系统的组成.....	5
1.2.2 数据库系统结构.....	8
1.3 数据模型.....	9
1.3.1 数据处理的三个阶段.....	9
1.3.2 常见的数据模型.....	13
1.4 处理大规模数据的数据库.....	17
1.4.1 分布式数据库.....	17
1.4.2 并行数据库.....	20
1.4.3 NoSQL 数据库.....	23
1.4.4 云数据库.....	25
1.4.5 XML 数据库.....	28
习题 1.....	30
第 2 章 关系型数据库基本理论	31
2.1 关系数据模型.....	31
2.1.1 关系数据结构.....	31
2.1.2 关系运算.....	35
2.1.3 关系的完整性约束.....	36
2.2 关系代数基本理论.....	38
2.2.1 传统的集合运算.....	39
2.2.2 专门的关系运算.....	42
2.2.3 关系代数表达式及其应用实例.....	46
2.3 关系数据库的规范化理论.....	47
2.3.1 关系模式规范化的必要性.....	47
2.3.2 函数依赖.....	49
2.3.3 关系的范式及规范化.....	51

2.3.4	关系模式的分解	53
习题 2		56
第 3 章	数据库设计	59
3.1	数据库设计概述	59
3.1.1	数据库设计目标和方法	59
3.1.2	数据库设计的基本步骤	61
3.2	需求分析	62
3.2.1	需求分析的任务和目标	63
3.2.2	需求分析的步骤	63
3.2.3	数据流图	64
3.2.4	数据字典	67
3.3	概念结构设计	68
3.3.1	概念结构设计任务和 E-R 模型的特点	68
3.3.2	概念结构设计的基本方法	68
3.3.3	概念结构设计的主要步骤	69
3.3.4	局部 E-R 模型的设计	70
3.3.5	全局 E-R 模型的设计	77
3.3.6	概念结构设计实例	81
3.4	逻辑结构设计	86
3.4.1	E-R 模型向关系模式的转换	86
3.4.2	关系模式的优化	90
3.5	物理结构设计	93
3.5.1	设计物理结构	93
3.5.2	评价物理结构	94
3.6	数据库的实施	95
3.7	数据库运行和维护	95
习题 3		96
第 4 章	SQL Server 系统概述	98
4.1	SQL Server 系统简介	98
4.1.1	SQL Server 的版本	98
4.1.2	SQL Server 系统数据库	99
4.1.3	SQL Server 三个关键系统表	99
4.2	Transact-SQL 简介	102
4.2.1	SQL 语言的发展与特点	102
4.2.2	Transact-SQL 语法基础	103
4.3	Transact-SQL 流程控制语句	112
4.3.1	BEGIN ... END 语句	112

4.3.2	分支语句	113
4.3.3	循环语句	115
4.3.4	RETURN 语句	116
4.3.5	WAITFOR 语句	117
4.3.6	TRY...CATCH 语句	117
4.4	SQL Server 存储机制	118
4.4.1	SQL Server 数据页	118
4.4.2	SQL Server 数据页类型	119
4.4.3	DBCC 命令	120
4.4.4	实例分析	121
习题 4	125
第 5 章	数据库和数据表管理	126
5.1	SQL Server 数据库概述	126
5.1.1	数据库文件类型	126
5.1.2	数据库文件组	127
5.2	SQL Server 数据库基本管理	127
5.2.1	创建用户数据库	127
5.2.2	数据库结构的修改	132
5.2.3	数据库文件的更名、删除	136
5.3	SQL Server 数据表管理	137
5.3.1	表的创建与维护	137
5.3.2	表中数据的维护	146
5.3.3	数据表的行列互换输出	148
习题 5	149
第 6 章	数据查询与游标机制	150
6.1	基本查询	150
6.1.1	SELECT 查询语句的结构	150
6.1.2	简单查询	151
6.1.3	带有 WHERE 子句的查询	154
6.1.4	带有 ORDER BY 子句的查询	156
6.1.5	带有 GROUP BY 子句的查询	157
6.1.6	输出结果选项	158
6.1.7	联合查询	159
6.2	多表查询	160
6.2.1	连接查询	160
6.2.2	子查询	164
6.3	游标机制	170

6.3.1	游标概述	170
6.3.2	游标的管理	171
6.3.3	利用游标修改和删除表数据	176
习题 6		176
第 7 章	视图与索引	178
7.1	视图	178
7.1.1	视图的基本概念	178
7.1.2	创建视图	179
7.1.3	修改视图	184
7.1.4	删除视图	185
7.1.5	使用视图	185
7.2	索引	187
7.2.1	索引的基本概念	187
7.2.2	创建索引	189
7.2.3	管理索引	192
习题 7		196
第 8 章	存储过程、触发器和用户定义函数	198
8.1	存储过程	198
8.1.1	存储过程概述	198
8.1.2	创建存储过程	200
8.1.3	调用存储过程	206
8.1.4	管理存储过程	207
8.2	触发器	210
8.2.1	触发器概述	210
8.2.2	创建触发器	211
8.2.3	管理触发器	217
8.3	用户定义函数	219
8.3.1	用户定义函数概述	219
8.3.2	创建用户定义函数	220
8.3.3	管理用户定义函数	223
习题 8		224
第 9 章	数据库并发控制	225
9.1	事务	225
9.1.1	事务概述	225
9.1.2	管理事务	227
9.2	并发数据访问管理	233

9.2.1	并发数据操作引起的问题	233
9.2.2	封锁机制	234
9.2.3	事务隔离级	238
习题 9		242
第 10 章	数据库安全管理	243
10.1	身份验证	243
10.1.1	Windows 验证模式	243
10.1.2	混合验证模式	243
10.2	身份验证模式的设置	244
10.2.1	使用“编辑服务器注册属性”	244
10.2.2	使用“对象资源管理器”	245
10.3	登录账户管理	245
10.3.1	创建登录账户	245
10.3.2	管理登录账户	249
10.4	数据库用户管理	250
10.4.1	创建数据库用户	251
10.4.2	删除数据库的用户	252
10.5	角色管理	252
10.5.1	SQL Server 角色类型	252
10.5.2	固定服务器角色管理	254
10.5.3	固定数据库角色管理	256
10.5.4	用户定义数据库角色	258
10.6	权限管理	260
10.6.1	语句权限	260
10.6.2	对象权限	262
10.6.3	隐含权限	263
10.6.4	授予用户或角色权限	263
10.6.5	拒绝用户或角色权限	265
10.6.6	撤销用户或角色权限	265
10.6.7	使用系统存储过程查看权限	266
习题 10		267
第 11 章	数据库备份与还原	269
11.1	备份与还原概述	269
11.1.1	备份方式	269
11.1.2	备份与还原策略	271
11.2	分离和附加数据库	272
11.2.1	分离数据库	273

11.2.2	附加数据库	275
11.3	数据库备份	276
11.3.1	创建和删除备份设备	276
11.3.2	备份数据库方法	279
11.4	数据库还原	287
11.4.1	数据库还原的技术	287
11.4.2	数据库还原的方法	291
习题 11	295
第 12 章	ADO.NET 访问 SQL Server 数据库	296
12.1	数据库访问技术 ADO.NET	296
12.1.1	ADO.NET 概述	296
12.1.2	数据库访问模式	299
12.2	数据库的连接	301
12.2.1	数据库的连接	301
12.2.2	ASP.NET 连接数据库测试	302
12.3	数据库的基本操作	302
12.3.1	用户登录界面	302
12.3.2	向数据库添加数据	303
12.3.3	记录数据管理	305
12.4	存储过程调用	306
12.4.1	无参数存储过程调用	306
12.4.2	带参数存储过程调用	308
12.4.3	用户自定义函数调用	308
12.5	执行 SQL 事务处理	310
习题 12	311
参考文献	312

数据库技术是现代信息科学与技术的重要组成部分，是计算机数据处理与信息管理系统核心。数据库技术研究和解决了计算机信息处理过程中大量数据有效地组织和存储的问题，在数据库系统中减少数据存储冗余、实现数据共享、保障数据安全以及高效地检索数据和处理数据。数据库技术的根本目标是要解决数据的共享问题。

数据库系统涉及许多基本概念，主要包括信息、数据、数据处理、数据库、数据库管理系统以及数据库系统等。

1.1 数据管理技术的发展

数据管理技术的发展是和计算机技术及其应用的发展联系在一起的，经历了由低级到高级的发展过程。

1.1.1 数据和数据管理

数据库系统的核心任务是数据管理。数据库技术是一门研究如何存储、使用和管理数据的技术，是计算机数据管理技术的最新发展阶段。数据库应用涉及数据（data）、信息（information）、数据处理（data processing）和数据管理（data management）等基本概念。

1. 数据和信息

现代社会是信息的社会，信息以惊人的速度增长，因此，如何有效地组织和利用它们成为急需解决的问题。数据库系统的目的是高效地管理及共享大量的信息，而信息与数据是分不开的。

数据和信息是数据处理中的两个基本概念，有时可以混用，但有时必须分清。数据是信息的一种表现形式，数据通过能书写的信息编码表示信息。信息有多种表现形式，它通过手势、眼神、声音或图形等方式表达，但是数据是信息的最佳表现形式。由于数据能够书写，因而它能够被记录、存储和处理，从中挖掘出更深层的信息。因此，也可以说，数据是原材料，信息是产品，信息是数据的含义。例如数据 1、3、5、7、9、11、13、15，它是一组原始数据，对它进行分析便可以看出：它是一组首项是 1，公差为 2 的等差数列，我们可以比较容易地知道它的任意项的值和前 n 项的和，这便是一条信息。而从数据 1、3、-23、4、5、-1、41 中不能提炼出任何有用的内容，故它不是信息。

数据和信息可以混用，有时一些数据对某些人来说可能是信息，而对另外一些人而言则可能只是数据。例如，在运输管理中，运输单对司机来说是信息，这是因为司机可以从该运输单上知道什么时候要为什么客户运输什么物品。而对负责经营的管理者来说，运输

单只是数据，因为从单张运输单中，无法知道本月经营情况，也不能掌握现有可用的司机、运输工具等情况。

2. 数据处理与数据管理

数据处理是对数据的采集、存储、检索、加工、变换和传输，基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出对于某些特定的人们来说是有价值、有意义的数据。例如，某省全体高考学生各门课程成绩的总分按从高到低的顺序进行排序、统计各个分数段的人数等，进而可以根据招生人数确定录取分数线。数据处理是系统工程和自动控制的基本环节。数据处理贯穿于社会生产和社会生活的各个领域。数据处理技术的发展及其应用的广度和深度，极大地影响着人类社会发展的进程。

数据管理是利用计算机硬件和软件技术对数据进行有效的收集、存储、检查、维护并实现对数据的各种运算和操作。其目的在于充分有效地发挥数据的作用。实现数据有效管理的关键是数据组织。随着计算机技术的发展，数据管理经历了人工管理、文件系统管理、数据库系统管理三个阶段。

1.1.2 数据管理发展的三个阶段

计算机硬件、系统软件的发展和计算机应用范围不断扩大是促使数据管理技术发展的主要因素。随着数据管理技术的不断发展，标志着数据存储冗余不断减小、数据独立性不断增强以及数据操作更加方便和简单。

1. 人工管理阶段

在这一阶段（20世纪50年代中期之前），计算机主要用于科学计算，其他工作还没有展开。外部存储器只有磁带、卡片和纸带等，还没有磁盘等直接存取存储设备。软件也处于初级阶段，只有汇编语言，没有操作系统和数据管理方面的软件。数据处理方式基本是批处理。这个阶段有如下几个特点：

（1）数据不保存，数据也无须长期保存。

（2）计算机系统不提供对用户数据的管理功能。用户编制程序时，必须全面考虑好相关的数据，包括数据的定义、存储结构以及存取方法等。程序和数据是一个不可分割的整体。数据脱离了程序就无任何存在的价值，数据无独立性。

（3）只有程序的概念，没有文件的概念。数据的组织形式必须由程序员自行设计。

（4）数据不能共享。不同的程序均有各自的数据，这些数据对不同的程序通常是不相同的，也不可共享。即使不同的程序使用了相同的一组数据，这些数据也不能共享，程序中仍然需要各自加入这组数据，不能省略。基于这种数据的不可共享性，必然导致程序与程序之间存在大量的重复数据，浪费了存储空间。

（5）数据面向程序，不单独保存数据。基于数据与程序是一个整体，数据只为本程序所使用，数据只有与相应的程序一起保存才有价值，否则就毫无用处。所以，所有程序的数据均不单独保存。

例如，学校管理系统有人事管理、学生管理和课程管理三部分，在人事管理阶段应用程序与数据之间的依赖关系如图 1.1 所示。

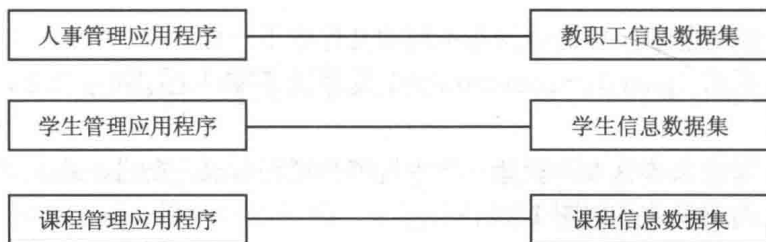


图 1.1 应用程序和数据的依赖关系

2. 文件管理阶段

在这一阶段（20 世纪 50 年代后期至 60 年代中期），计算机不仅用于科学计算，还用于信息管理。随着数据量的增加，数据的存储、检索和维护问题成为紧迫的需要，数据结构和数据管理技术迅速发展起来。此时，外部存储器已有磁盘、磁鼓等直接存取存储设备。软件领域出现了高级语言（如 Fortran、Algol60 等）和操作系统（如 GM-NAA I/O、IBSYS 等）。操作系统中的文件系统是专门管理外存的数据管理软件。数据处理的方式有批处理，也有联机实时处理。

这一阶段的数据管理有以下特点：

（1）数据以“文件”形式可长期保存在外部存储器的磁盘上。由于计算机的应用转向了信息管理，因此对文件要进行大量的查询、修改和插入等操作。

（2）数据的逻辑结构与物理结构有了区别，但比较简单。程序与数据之间具有“设备独立性”，即程序只需用文件名就可以进行数据操作，不必关心数据的物理位置。由操作系统的文件系统提供存取方法（读/写）。

（3）文件组织已多样化。有索引文件、链接文件和直接存取文件等。但文件之间相互独立、缺乏联系等问题。数据之间的联系要通过程序去构造。

（4）数据不再属于某个特定的程序，可以重复使用，即数据面向应用。但是文件结构的设计仍然是基于特定的用途，程序基于特定的物理结构和存取方法。因此程序与数据结构之间的依赖关系并未根本改变。

在文件系统阶段，由于具有设备独立性，因此当改变存储设备时，不必改变应用程序。但这只是初级的数据管理，在数据的物理结构修改时，仍然需要修改用户的应用程序，即应用程序具有“程序-数据依赖”性。有关物理表示的知识和访问技术将直接体现在应用程序的代码中。

（5）对数据的操作以记录为单位。这是由于文件中只存储数据，不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除及修改等所有操作，都要用程序来实现。

文件系统阶段是数据管理技术发展中的一个重要阶段。在这一阶段，得到充分发展的数据结构和算法丰富了计算机科学，为数据管理技术的进一步发展打下了基础。

随着数据管理规模的扩大，数据量急剧增加，文件系统显露出了三个明显的缺陷：

（1）数据冗余（redundancy）。由于文件之间缺乏联系，造成每个应用程序都有对应的文件，有可能同样的数据在多个文件中重复存储。

（2）数据不一致（inconsistency）。这往往是由数据冗余造成的，在进行数据更新操作

时,稍不谨慎,就可能使同样的数据在不同的文件中不一样。

(3) 数据联系弱 (poor data relationship)。这是由于文件之间相互独立,相互之间又缺乏联系造成的。

例如,学校管理系统有人事管理、学生管理和课程管理三部分,在文件管理阶段应用程序与数据之间的依赖关系如图 1.2 所示。



图 1.2 应用程序和数据的依赖关系

3. 数据库管理阶段

在 20 世纪 60 年代末,磁盘技术取得了重要进展,具有数百兆容量和快速存取的磁盘陆续进入市场,成本也不高。同时,计算机在管理中应用规模更加庞大、数据量急剧增加,为数据库技术的产生提供了良好的物质条件。数据库系统克服了文件系统的缺陷,提供了对数据更高级、更有效的管理。概括起来,数据库阶段的管理方式具有以下特点:

(1) 数据结构化。数据结构化是数据库系统与文件系统之间的根本区别。数据库中包含许多单独的数据文件,这些文件的数据具有特定数据结构(数据项不可再分、类型相同)。文件之间也存在相互的联系,在整体上也服从一定的组织形式,从而满足管理大量数据的需求。

(2) 较高的数据共享性。数据共享是指数据不再面向某个应用而是面向整个系统,使得多个用户同时存取数据而互不影响。如同一企业中的不同部门,甚至不同企业、不同地区的用户都可以使用同一个数据库中的数据,减少了数据冗余。

数据共享性可以大大减少数据冗余,节约存储空间,还能够避免数据之间的不相容性与不一致性。

(3) 统一管理和控制数据。利用专门的数据库管理系统实现对数据定义、操作、统一管理和控制。在应用程序和数据库之间保持高度的独立性,数据具有完整性、一致性和安全性,并具有充分的共享性,有效地减少了数据冗余。

例如,学校管理系统有人事管理、学生管理和课程管理三部分,在数据库管理阶段应用程序与数据之间的依赖关系如图 1.3 所示。

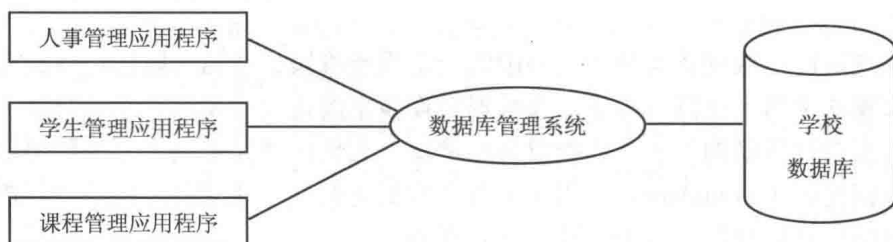


图 1.3 应用程序和数据的依赖关系

1.2 数据库系统



数据库系统 (DataBase Systems, DBS) 是为适应数据处理需要而发展起来的一种较为理想数据处理的核机构。它是一个实际可运行的, 用于存储、维护数据的软件系统并且可以向应用系统提供数据, 是存储介质、处理对象和管理系统的集合体。

1.2.1 数据库系统的组成

数据库系统是指在计算机系统中引入数据库后的系统, 一般由计算机硬件、数据库、数据库管理系统、数据库开发工具、数据库应用系统、数据库管理员和用户构成。数据库系统各元素的层次结构如图 1.4 所示。

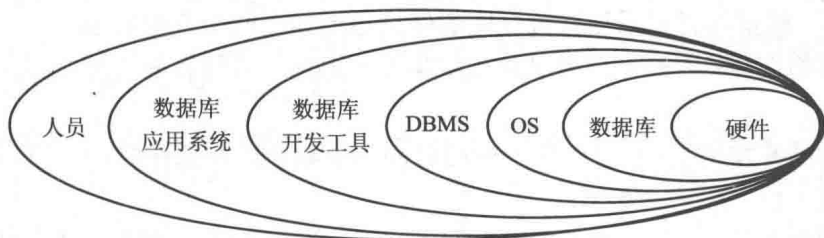


图 1.4 数据库系统层次图

1. 计算机硬件

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、输入输出设备以及计算机网络环境。

2. 数据库

数据库 (DataBase, DB) 是指长期存储在计算机内、有组织的、统一管理的相关数据的集合。数据库能为各种用户共享, 具有较小的数据冗余度, 数据间联系紧密而又有较高的数据独立性等。

特别需要指出的是, 数据库中存储数据具有“集成的”和“共享的”的特点。

所谓“集成”, 是指把某个特定应用环境中的与各种应用相关的数据及其数据之间的联系 (联系也是一种数据) 全部集中并按照一定的结构形式进行存储, 或者说, 把数据库看成是若干个性质的数据文件的联合和统一的数据整体, 并且在文件之间局部或全部消除了冗余, 使数据库系统具有整体数据结构化和数据冗余小的特点。

所谓“共享”, 是指数据库中的一块块数据可为多个不同的用户所共享, 即多个不同的用户使用多种不同的语言, 为了不同的应用目的同时存取数据库, 甚至同时存取同一块数据。共享实际上是基于数据库是“集成的”这一事实的结果。

3. 计算机软件

数据库系统中的软件包括操作系统 (Operating System, OS)、数据库管理系统、数据库开发工具及数据库应用系统。操作系统给用户提供良好的应用接口, 数据库系统必须在操作系统的支持下才能正常使用。

数据库管理系统 (DataBase Management System, DBMS) 为用户或应用程序提供访问数据库的方法, 包括数据库的建立、查询、更新及各种数据控制。数据库系统各类用户对数据库的各种操作请求, 都是由数据库管理系统来完成的, 它是数据库系统的核心软件。

数据库管理系统有大小之分, 常见的大中型数据库管理系统有 Oracle、DB2、SQL Server、Sybase 等, 小型数据库管理系统有 Foxpro、Access、MySQL 等。

数据库管理系统的主要功能有以下五个方面。

(1) 数据库的定义功能: 数据库管理系统提供数据定义语言 (Data Definition Language, DDL) 定义数据库的三级结构、两级映像, 定义数据的完整性约束、保密限制约束等。

(2) 数据库的操纵功能: 数据库管理系统提供数据操纵语言 (Data Manipulation Language, DML) 实现对数据的操作。基本的数据操作有两类: 检索 (查询) 和更新 (包括插入、删除、更新)。

(3) 数据库的保护功能: 数据库中的数据是信息社会的战略资源, 对数据的保护是至关重要的大事。数据库管理系统对数据库的保护通过如下四个方面实现。

- 数据库的恢复: 在数据库被破坏或数据不正确时, 系统有能力把数据库恢复到正确的状态。
- 数据库的并发控制: 在多个用户同时对同一个数据进行操作时, 系统应能加以控制, 防止破坏数据库中的数据。
- 数据完整性控制: 保证数据库中数据及语义的正确性和有效性, 防止任何对数据造成错误的操作。
- 数据安全性控制: 防止未经授权的用户存取数据库中的数据, 以避免数据的泄露、更改或破坏。

(4) 数据库的维护功能: 这一部分包括数据库的数据载入、转换、转储, 数据库的改组以及性能监控等功能。

(5) 数据字典: 数据库系统中存放三级结构定义的数据库称为数据字典 (Data Dictionary, DD)。对数据库的操作都要通过访问数据字典才能实现。数据字典中还存放数据库运行时的统计信息, 例如记录个数、访问次数等。管理数据字典的实用程序称为“数据字典系统”, 访问数据字典中的数据是由数据字典系统实现的。在现有的大型系统中, 把数据字典系统单独抽出来成为一个软件工具。

有两点需要说明: 一是数据库管理系统功能强弱随系统而异, 大系统功能较强、较全, 小系统较弱、较少; 二是应用程序并不属于数据库管理系统范围, 应用程序是用主语言和 DML 编写的, 程序中 DML 语句由数据库管理系统执行, 而其余部分仍由主语言编译程序完成。

4. 人员

开发、管理和使用数据库系统的人员主要是数据库管理员 (DataBase Administrator, DBA)、系统分析员 (System Analyst, SA) 和数据库设计人员、应用程序员和终端用户。

1) 数据库管理员

在数据库环境下, 有两类共享资源: 一类是数据库, 另一类是数据库管理系统软件。因此需要有专门的管理机构来监督和管理数据库系统。DBA 则是这个机构的一个 (组) 人员, 负责全面管理和控制数据库系统。具体职责包括: