

深入
浅出 系列规划教材

深入浅出

数据库系统及应用基础

程学先 主编

清华大学出版社

深入
浅出 系列规划教材

深入浅出

数据库系统及应用基础

程学先 主编

清华大学出版社
北京

内 容 简 介

本书介绍了最基础的数据库系统的基本概念、基本原理和基本方法。内容包括数据库概念,SQL语言,关系代数、数据库基础理论、数据模型、关系优化、数据库系统管理。为帮助学生更好地理解数据库基本理论、基本概念与基本方法,做到理论联系实际,突出应用,提供了一个小型仓储管理信息系统实例,介绍了从需求分析、建模、建立数据库与生成应用系统的全过程。为学以致用,本书联系社会上广泛应用的SQL Server数据库组织教学。为了解决编程难题,设计了基于部件库的管理信息系统软件生产线,使得无任何编程语言知识与技能就能完成应用系统的设计与建设。本书内容分为基础篇与实践篇,在实践篇中结合有关章、节组织上机操作,其中包括应用我们开发的实验工具程序进行的实验与应用软件生产线进行系统开发的实验,通过实践环节的学习使建立对数据库及其应用的感性认识、理论联系实际以加深对数据库基本概念与基本方法的理解、学以致用使提高动手能力。

为配合教学,在清华大学出版社网站中提供了PowerPoint教学幻灯片,管理信息系统软部件库(Java版)及软件生产线程序、实践篇中例题的电子文档。

本书可作为高等院校应用类型本、专科学生新生计算机数据库基础课教材,也可作为转型学校和高职高专学校数据库应用课程的教材。

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

深入浅出数据库系统及应用基础/柴文强编. —北京:清华大学出版社,2015

深入浅出系列规划教材

ISBN 978-7-302-40824-6

I. ①深… II. ①柴… III. ①数据库系统—教材 I. ①TP311.13

中国版本图书馆CIP数据核



责任编辑:白立军 柴文强

封面设计:傅瑞学

责任校对:胡伟民

责任印制:刘海龙

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:三河市少明印务有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:15

字 数:375千字

版 次:2015年9月第1版

印 次:2015年9月第1次印刷

印 数:1~2000

定 价:29.00元

产品编号:062978-01



为什么开发深入浅出系列丛书?

目的是从读者角度写书,开发出高质量的、适合阅读的图书。

“不积跬步,无以至千里;不积小流,无以成江海。”知识的学习是一个逐渐积累的过程,只有坚持系统地学习知识,深入浅出,坚持不懈,持之以恒,才能把一类技术学习好。坚持的动力源于所学内容的趣味性和讲法的新颖性。

计算机课程的学习也有一条隐含的主线,那就是“提出问题→分析问题→建立数学模型→建立计算模型→通过各种平台和工具得到最终正确的结果”,培养计算机专业学生的核心能力是“面向问题求解的能力”。由于目前大学计算机本科生培养计划的特点,以及受教学计划和课程设置的原因,计算机科学与技术专业的本科生很难精通掌握一门程序设计语言或者相关课程。各门课程设置比较孤立,培养的学生综合运用各方面的知识能力方面有欠缺。传统的教学模式以传授知识为主要目的,能力培养没有得到充分的重视。很多教材受教学模式的影响,在编写过程中,偏重概念讲解比较多,而忽略了能力培养。为了突出内容的案例性、解惑性、可读性、自学性,本套书努力在以下方面做好工作。

1. 案例性

所举案例突出与本课程的关系,并且能恰当反映当前知识点。例如,在计算机专业中,很多高校都开设了高等数学、线性代数、概率论,不言而喻,这些课程对于计算机专业的学生来说是非常重要的,但就目前对不少高校而言,这些课程都是由数学系的老师讲授,教材也是由数学系的老师编写,由于学科背景不同和看待问题的角度不同,在这些教材中基本都是纯数学方面的案例,作为计算机系的学生来说,学习这样的教材缺少源动力并且比较乏味,究其原因,很多学生不清楚这些课程与计算机专业的关系是什么。基于此,在编写这方面的教材时,可以把计算机上的案例加入其中,例如,可以把计算机图形学中的三维空间物体图像在屏幕上的伸缩变换、平移变换和旋转变换在矩阵运算中进行举例;可以把双机热备份的案例融入到马尔科夫链的讲解;把密码学的案例融入到大数分解中等。

2. 解惑性

很多教材中的知识讲解注重定义的介绍,而忽略因果性、解释性介绍,往往造成知其然而不知其所以然。下面列举两个例子。

(1) 读者可能对 OSI 参考模型与 TCP/IP 参考模型的概念产生混淆,因为两种模型之

间有很多相似之处。其实,OSI参考模型是在其协议开发之前设计出来的,也就是说,它不是针对某个协议族设计的,因而更具有通用性。而TCP/IP模型是在TCP/IP协议栈出现后出现的,也就是说,TCP/IP模型是针对TCP/IP协议栈的,并且与TCP/IP协议栈非常吻合。但是必须注意,TCP/IP模型描述其他协议栈并不合适,因为它具有很强的针对性。说到这里读者可能更迷惑了,既然OSI参考模型没有在数据通信中占有主导地位,那为什么还花费这么大的篇幅来描述它呢?其实,虽然OSI参考模型在协议实现方面存在很多不足,但是,OSI参考模型在计算机网络的发展过程中起到了非常重要的作用,并且,它对未来计算机网络的标准化、规范化的发展有很重要的指导意义。

(2)再例如,在介绍原码、反码和补码时,往往只给出其定义和举例表示,而对最后为什么在计算机中采取补码表示数值?浮点数在计算机中是如何表示的?字节类型、短整型、整型、长整型、浮点数的范围是如何确定的?下面我们来回答这些问题(以8位数为例),原码不能直接运算,并且0的原码有+0和-0两种形式,即00000000和10000000,这样肯定是不行的,如果根据原码计算设计相应的门电路,由于要判断符号位,设计的复杂度会大大增加,不合算;为了解决原码不能直接运算的缺点,人们提出了反码的概念,但是0的反码还是有+0和-0两种形式,即00000000和11111111,这样是不行的,因为计算机在计算过程中,不能判断遇到0是+0还是一0;而补码解决了0表示的唯一性问题,即不会存在+0和-0,因为+0是00000000,它的补码是00000000,-0是10000000,它的反码是11111111,再加1就得到其补码是10000000,舍去溢出量就是00000000。知道了计算机中数用补码表示和0的唯一性问题后,就可以确定数据类型表示的取值范围了,仍以字节类型为例,一个字节共8位,有00000000~11111111共256种结果,由于1位表示符号位,7位表示数据位,正数的补码好说,其范围从00000000~01111111,即0~127;负数的补码为10000000~11111111,其中,11111111为-1的补码,10000001为-127的补码,那么到底10000000表示什么最合适呢?8位二进制数中,最小数的补码形式为10000000;它的数值绝对值应该是各位取反再加1,即为 $01111111+1=10000000=128$,又因为是负数,所以是-128,即其取值范围是-128~127。

3. 可读性

图书的内容要深入浅出,使人爱看、易懂。一本书要做到可读性好,必须做到“善用比喻,实例为王”。什么是深入浅出?就是把复杂的事物简单地描述明白。把简单事情复杂化的是哲学家,而把复杂的问题简单化的是科学家。编写教材时要以科学家的眼光去编写,把难懂的定义,要通过图形或者举例进行解释,这样能达到事半功倍的效果。例如,在数据库中,第一范式、第二范式、第三范式、BC范式的概念非常抽象,很难理解,但是,如果以一个教务系统中的学生表、课程表、教师表之间的关系为例进行讲解,从而引出范式的概念,学生会比较容易接受。再例如,在生物学中,如果纯粹地讲解各个器官的功能会比较乏味,但是如果提出一个问题,如人的体温为什么是37℃?以此为引子引出各个器官的功能效果要好得多。再例如,在讲解数据结构课程时,由于定义多,表示抽象,这样达不到很好的教学效果,可以考虑在讲解数据结构及其操作时用程序给予实现,让学生看到直接的操作结果,如压栈和出栈操作,可以把PUSH()和POP()操作实现,这样效果会好

很多,并且会激发学生的学习兴趣。

4. 自学性

一本书如果适合自学学习,对其语言要求比较高。写作风格不能枯燥无味,让人看一眼就拒人千里之外,而应该是风趣、幽默,重要知识点多举实际应用的案例,说明它们在实际生活中的应用,应该有画龙点睛的说明和知识背景介绍,对其应用需要注意哪些问题等都要有提示等。

一书在手,从第一页开始的起点到最后一页的终点,如何使读者能快乐地阅读下去并获得知识?这是非常重要的问题。在数学上,两点之间的最短距离是直线。但在知识的传播中,使读者感到“阻力最小”的书才是好书。如同自然界中没有直流的河流一样,河水在重力的作用下一定沿着阻力最小的路径向前进。知识的传播与此相同,最有效的传播方式是传播起来损耗最小,阅读起来没有阻力。

是为序。

欢迎老师投稿: bailj@tup.tsinghua.edu.cn。

2014年12月15日

前言



数据库技术是一门发展迅速,在科学与社会各个领域广泛应用的技术,是计算机科学的重要分支,是建设各类信息系统的基础,目前其应用已从信息检索、一般管理扩大到科技计算、计算机辅助设计、人工智能等各个领域。

本书力求由浅入深,由具体到抽象,由实践到理论,应用驱动,理论与实践紧密相结合,使读者深刻了解什么是数据库及数据库要达到的目标,掌握数据库的设计方法并了解实际数据库应用系统的组成及实现的方法。

本书先对比了应用程序设计语言进行数据处理的方法、Excel 和数据库环境下数据存储与操作方法的不同,通过解剖一个数据库的数据文件的实验认识什么是数据库,及数据库理论中一些至关重要的概念,建立对顺序文件与链表文件、等长存放、逻辑结构、数据字段、数据类型等的感性认识,让读者更真实地体会要完成数据库系统的目标需要解决的矛盾。

本书以在实际应用系统中被广泛应用的 SQL SERVER 数据库管理系统为例组织教学,使关于数据库原理的学习与数据库应用实际紧密结合,使更容易地加深对数据库基本理论、基本概念的理解。

随本书发布应用“管理信息系统软部件库(Java 版)”实现的桌面系统和软件生产线程序。“管理信息系统软部件”不以具体的应用系统为目标,而是以抽象的“数据库”为目标,总结作者开发应用系统的经验,整合各类应用系统的模块设计特点,将界面设计与功能设计有机结合到一起。每个部件都集成许多功能,就如同电子产品中的智能插件、机械产品中的通用部件一样,只需设定初始参数,就能选择具体的功能,用于新的设计。以之作为工具用于实践教学,可以帮助读者更深入地理解数据库的基本理论与基本概念。也为建立软件生产线铺平道路。

我们基于软部件库设计了一条管理信息系统通用软件生产线,以参照 UML 语言、独立设计的建模工具驱动,只需全部或部分绘制用例图、数据结构类图、组件图、系统结构类图与时序图,就能自动生成菜单程序、调用部件库中部件程序,构建应用系统。该生产线目标是提高系统易维护性、易扩展性与开发效率,让非计算机人员和初学者也能很容易地学会开发应用系统,将理论与实践紧密相结合到一起,读者将不只是学习少数几个虚拟的应用系统的设计方法,而是要实地设计并建立一到多个真实的数据库应用系统,实战式的学习数据库系统的设计与实现方法,全面地学习数据库的理论技术。设计数据库应用系统的生产实践也将有利于提高读者实际动手能力。希望通过本书的学习,读者不只停留在数据库的概念、原理和理论上,而是要成为使用数据库的行家里手。

运行桌面系统,只要选择部件、选择数据表、设置参数,就能打开数据库,进行一项关于数据库应用的操作,界面丰富,较 T-SQL 语言及一般数据库内嵌语言简单且容易学习,可以更灵活、更自由地支配与使用数据库。

本书尽量采用通俗方式表述那些涉及专业太深的问题,对一些用得不多或较难懂的问题只做最简单的介绍。

本书由程学先教授主编,张慧萍、程传庆、顾梦霞、贺红艳、陈小娟、罗宏芳等参加编写并完成课件的制作。其中,张慧萍负责第 1 章编写;程传庆负责第 2 章编写;顾梦霞负责第 3、8 章编写;贺红艳负责第 4 章编写;陈小娟负责第 5 章编写;罗宏芳负责第 6 章编写;程学先负责其他章节编写。

管理信息系统软部件程序由程学先、龚晓明、王富强、程传庆、齐赛、王玉明、陈义、王嘉、程传慧、聂志恒、张卫东、杨帅等设计,参加其他电子文稿设计与制作的还有曾玲、杨晓艳、刘伟、陈永辉、程传慧等,在此一并表示感谢。

本书未详细介绍部件库与软件生产线功能与设计方法,有兴趣的读者可参阅清华大学出版社出版的本人拙著《数据库系统原理与应用》。软部件库还需要进一步发展、完善与规范化,欢迎有兴趣者公开署名发表新的或经修改更新的部件,以充分满足应用的需要。我们愿意积极宣传与推广这方面的进展,可以应用信件和我们联系或在我的博客中留言。通信地址:430068,湖北省湖北工业大学。博客地址:<http://blog.sciencenet.cn/?2551>。

作者
2015 年 5 月



基 础 篇

第 1 章	认识数据库	3
1.1	数据处理发展历程	3
1.1.1	人工管理阶段(20 世纪 50 年代中期以前)	4
1.1.2	文件系统阶段(20 世纪 50 年代后期至 60 年代)	4
1.1.3	Excel	6
1.1.4	数据库系统阶段(20 世纪 60 年代后期)	8
1.2	SQL Server 2008 可视化建库建表操作方法	9
1.2.1	建立数据库	9
1.2.2	建立数据表	10
1.2.3	修改表结构	12
1.3	数据维护的可视化操作方法	13
1.3.1	数据维护的概念	13
1.3.2	分析数据库数据文件以深入认识数据库	14
1.4	数据备份与恢复的可视化操作方法	17
1.4.1	数据转储的概念	17
1.4.2	创建备份设备	18
1.4.3	数据库备份与还原	19
1.4.4	数据库分离与附加	20
1.5	数据库管理系统	21
1.5.1	DBMS 的功能	21
1.5.2	DBMS 的组成	22
1.5.3	关系数据库初步	23
	小结	25
	习题	25
第 2 章	SQL 语言	27
2.1	SQL 语言概貌	28

2.2	SQL 数据定义功能	29
2.2.1	定义数据库	29
2.2.2	基本表的定义和修改	30
2.2.3	在 SQL Server 中执行 SQL 语句	32
2.3	SQL 数据维护语句	33
2.3.1	修改(UPDATE)语句	33
2.3.2	删除(DELETE)语句	33
2.3.3	插入(ININSERT)语句	34
2.4	SQL 数据查询语句	34
2.4.1	SQL 查询语句格式	34
2.4.2	对单一表查询语句	36
2.4.3	对两个以上表的连接查询	38
2.4.4	外部连接	39
2.4.5	嵌套查询	40
2.4.6	批量录入数据的语句	41
2.4.7	SELECT 语句的关联应用	41
2.5	视图	43
2.5.1	视图的概念	43
2.5.2	可视化方式建立视图	43
2.5.3	建立视图的语句	45
2.5.4	视图用于更新	46
2.6	SQL 数据控制功能	48
	小结	49
	习题	50

第 3 章 关系代数..... 52

本章学习目标	52
3.1 传统的集合运算	52
3.2 专门的关系运算	55
小结	59
习题	60

第 4 章 数据库基础理论..... 61

4.1 数据库系统基本概念	61
4.2 三层模式与二级映像	62
4.3 数据控制的概念	63
4.4 数据完整性保护的可视化实现	64
4.4.1 实体完整性保护的实现	64



4.4.2	参照完整性保护的实现	65
4.4.3	域完整性保护的实现	67
4.5	设置数据完整性保护的语句	68
4.5.1	建表语句中添加数据完整性约束	68
4.5.2	涉及数据完整性的数据表结构修改语句	71
4.6	索引文件组织	74
4.6.1	索引的概念	74
4.6.2	聚集索引与唯一索引	75
4.7	数据库安全性管理	76
4.7.1	主体与安全对象	76
4.7.2	身份验证模式	77
4.7.3	登录名的管理	78
4.7.4	创建架构	78
4.7.5	针对具体数据库创建用户名	79
4.7.6	服务器角色	80
4.7.7	数据库角色	81
4.7.8	权限管理	83
小结	83
习题	84
第5章	数据模型	86
本章学习目标	86
5.1	数据模型	86
5.2	E-R 数据模型	87
5.2.1	数据之间的联系	87
5.2.2	实体联系模型(E-R 图)	89
5.3	关系数据模型	91
5.3.1	关系数据模型基本概念	91
5.3.2	从 E-R 数据模型到关系数据模型	92
小结	93
习题	94
第6章	关系优化	96
本章学习目标	96
6.1	基本概念	96
6.2	候选关键字与主属性	97
6.2.1	候选关键字	97
6.2.2	主属性	98



6.3 关系规范化	98
6.3.1 问题的提出	98
6.3.2 范式	100
小结	104
习题	104

第7章 小型仓库管理系统设计 106

本章学习目标	106
7.1 管理信息系统设计概述	106
7.1.1 管理信息系统设计步骤	106
7.1.2 软件生产线	107
7.1.3 部件库	108
7.1.4 用例图	108
7.1.5 数据结构类图	110
7.1.6 组件图	112
7.1.7 系统结构类图	113
7.2 仓储管理系统需求分析	115
7.2.1 仓储管理系统功能要求	115
7.2.2 系统模块结构	117
7.2.3 实体分析	117
7.3 用例图设计	118
7.4 数据结构类图设计与建立数据表	120
7.5 应用系统设计与建立	122
7.5.1 系统结构类图设计	122
7.5.2 生成应用系统	128
小结	129
习题	130

第8章 数据库系统管理 133

本章学习目标	133
8.1 SQL Server 管理工具	133
8.2 SQL Server 中的存储过程	135
8.2.1 存储过程的概念	136
8.2.2 存储过程的优点	136
8.2.3 在 SQL Server 中,创建一个存储过程的两种方法	137
8.2.4 使用管理工作平台创建存储过程	137
8.2.5 重新命名存储过程	138
8.2.6 删除存储过程	138

8.2.7	执行存储过程	139
8.3	SQL Server 中的触发器	139
8.3.1	触发器的概念及作用	139
8.3.2	触发器的种类	140
8.3.3	创建触发器	140
8.3.4	触发器的原理	140
8.3.5	INSTEAD OF 触发器	142
8.3.6	触发器的应用	142
8.4	数据库接口技术	142
8.4.1	ODBC(Open Database Connectivity)	142
8.4.2	JDBC	145
8.5	事务处理	145
8.5.1	事务的基本概念	145
8.5.2	事务处理过程分析	146
8.5.3	SQL 的事务管理	147
8.6	并发控制	148
8.6.1	并发处理产生的三种不一致性	148
8.6.2	封锁	150
	习题	153

实 践 篇

第 9 章	基于 SQL Server 2008 的应用实践	157
9.1	SQL Server 2008 的可视化操作方法	157
9.1.1	可视化方式创建数据库	157
9.1.2	可视化方式创建数据表	159
9.1.3	可视化方式进行数据录入	161
9.1.4	分离和附加数据库	162
9.1.5	备份和恢复数据库	163
9.1.6	修改数据库数据文件	165
	课堂练习	166
9.2	SQL Server 文件组织分析	168
	课堂练习	170
9.3	SQL 数据定义功能	170
9.3.1	创建数据库的语句	170
9.3.2	修改数据库定义的语句	171
9.3.3	创建数据表的语句	174
	课堂练习	177
9.4	可视化定义数据完整性的方法	178
9.4.1	设置主键	178

9.4.2	设置“唯一”约束	180
9.4.3	设置参照关系	181
	课堂练习	182
9.5	应用“新建数据表实验程序.jar”程序建立数据表	183
	课堂练习	189
9.6	应用“结构维护实验.jar”程序辅助生成修改数据表结构语句	190
9.6.1	操作“数据表结构维护实验.jar”	190
9.6.2	生成数据结构维护语句	191
	课堂练习	193
9.7	认识 SQL 查询语句	193
9.7.1	单表查询	196
9.7.2	涉及聚集函数的查询	197
9.7.3	多表查询	198
	课堂练习	200
9.8	应用“查询实验程序.jar”辅助生成查询语句	200
9.8.1	查询实验程序.jar 功能与操作方法	200
9.8.2	应用实验程序生成 SELECT 语句	201
	课堂练习	203
9.9	应用关系运算.jar 程序学习关系代数理论	204
9.9.1	“关系运算.jar”程序操作方法	204
9.9.2	认识关系运算运行结果	205
	课堂练习	206
9.10	建立工资管理系统	206
9.10.1	工资管理系统的需求	206
9.10.2	系统模块结构	207
9.10.3	数据库设计	208
9.10.4	建立工资管理系统	208
9.10.5	运行系统	210
	课堂练习	211
9.11	视图、索引、存储过程、触发器设计方法	212
9.11.1	关于规则	213
9.11.2	索引	214
9.11.3	T-SQL 实验	214
9.11.4	存储过程	215
9.11.5	触发器	217
9.11.6	默认对象	219
9.11.7	安全	220
	参考文献	223



基 础 篇

本章学习目标

本章首先介绍“数据处理”发展历程,说明数据库面对的环境及所需要解决的问题,表述对数据库的基本要求。其中包括文件管理系统进行数据管理的缺点,数据库管理系统的优点,数据库管理系统的组成。继而介绍 SQL Server 2008 数据库的基本操作方法,包括建库、建表、数据维护、数据备份与恢复等可视化操作方法,建立对数据库的感性认识。描述了数据库管理系统 DBMS 的构成与功能。通过本章学习,读者应该掌握以下内容:

- 数据库管理系统与传统的文件管理系统的主要区别与各自特点
- 数据共享、数据安全、数据完整性、数据独立、数据冗余、数据不一致等基本概念
- SQL Server 2008 数据库的建库、建表、数据维护、数据备份与恢复等可视化操作方法
- 数据库管理系统(DBMS)的构成与功能

当前时代是信息技术飞跃发展的时代。

信息,是以数据为载体的客观世界实际存在的事物、事件或概念在人们头脑中的反映。

信息以数据为载体。

信息系统是以计算机为核心,以数据库为基础,对代表信息的数据进行收集、组织、储存、加工、传播、管理和使用的系统。

数据库能借助计算机保存和管理复杂的大量的数据,快速而有效地为多个不同的用户和各种应用程序提供需要的数据,以便人们能更方便、更快速、更充分地利用宝贵的信息资源。

1.1 数据处理发展历程

所有企业、所有单位都离不开管理。人工管理的主要工作是进行数据处理。

数据处理指数据的存储、管理、加工、通信与使用。其特点是数据量大、数据之间有着复杂的逻辑联系。数据处理矛盾的焦点是数据管理,数据管理包括对数据的收集、整理、