



高等学校应用型本科
“十三五”规划教材

GAODENG XUEXIAO YINGYONGXING BENKE
SHISANWU GUIHUA JIAOCAI

数据库原理 及应用

SHUJUKU YUANLI
JI YINGYONG

主 编 郭 鑫 陈秀玲
主 审 唐 友



重庆大学出版社

内容提要

本书以数据库技术在信息化社会及大数据金融背景下的应用为目标,深入浅出地设计了13个项目的内容,共79个任务,内容涵盖数据库基础知识、T-SQL语言基础、创建和管理数据库、创建和管理数据表、视图、触发器、高级数据操作、存储过程的创建与管理、游标、安全管理、函数、大数据金融及仓库管理系统案例实践等。

本书知识点全部通过实践任务引出,且任务设计强调相互衔接与梯度递进。从内容的组织和编写结构上看,它既可以作为高等院校数据库应用课程的教材,又可供社会各类计算机及金融领域应用人员阅读参考。

图书在版编目(CIP)数据

数据库原理及应用 / 郭鑫,陈秀玲主编. --重庆:
重庆大学出版社,2018.3

高等学校应用型本科“十三五”规划教材

ISBN 978-7-5689-1027-9

I. ①数… II. ①郭… ②陈… 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 034478 号

数据库原理及应用

主 编 郭 鑫 陈秀玲

副主编 白 玲 陈井霞 伞 颖

责任编辑:顾丽萍 版式设计:顾丽萍

责任校对:郭小梅 责任印制:张 策

*

重庆大学出版社出版发行

出版人:易树平

社址:重庆市沙坪坝区大学城西路21号

邮编:401331

电话:(023) 88617190 88617185(中小学)

传真:(023) 88617186 88617166

网址:<http://www.cqup.com.cn>

邮箱:fxk@cqup.com.cn(营销中心)

全国新华书店经销

重庆华林天美印务有限公司印刷

*

开本:787mm×1092mm 1/16 印张:22 字数:537千

2018年3月第1版 2018年3月第1次印刷

印数:1—3 000

ISBN 978-7-5689-1027-9 定价:53.00元

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书
制作各类出版物及配套用书,违者必究

前 言

人类已步入大数据、人工智能时代,数据被视为“新世纪的矿产与石油”,正成为巨大的经济资产,而数据应用及管理显得尤为重要。围绕数据进行的组织、存储、维护、统计和查询等工作意义深远,在“互联网+”背景下,减少数据冗余,提供更高的数据共享能力价值无限。

在计算机技术迅猛发展、社会信息化进程加快以及大数据金融风靡的背景下,广大企事业单位管理人员、工程技术人员以及各行各业的相关人员都迫切希望掌握数据管理技术,以提高工作效率和工作质量;而对于面向 21 世纪的高层次人才,广大高校学生都需要学习并掌握数据库的基本知识和数据管理的基本技能,并开发出实用的数据库应用系统。本书在传统经典的数据库应用基础上,以相关慕课课程内容为入口,介绍了大数据金融,提供了跨学科的应用介绍,丰富了学生的视野,开阔了学习思路。从本书的内容组织和编写结构上看,它既可以作为高等院校数据库应用课程的教材,又可供社会各类计算机及金融领域应用人员阅读参考。

全书采用任务驱动的方式组织内容,即采用情境设置、样例展示、任务实施、知识链接的编写模式,以项目“学生管理数据库”贯穿整本教材,将项目分解成若干个任务,通过解决具体任务学习对应的理论知识;使教、学、做紧密结合。全书内容上共设 13 个项目:数据库基础知识、Transact-SQL 语言基础、创建和管理数据库、创建和管理数据表、视图、触发器、高级数据操作、存储过程的创建与管理、游标、安全管理、函数、大数据金融及仓库管理系统案例实践。教材编写上贯彻理论够用、侧重实践的原则,力求做到结构合理、层次清晰、概念明确、突出应用。本书能够体现出课程的特点,注重应用技能的培养,每一部分理论知识均有与其相对应的实验内容,以典型事例为素材。

本书由郭鑫、陈秀玲担任主编,白玲、陈井霞、伞颖担任副主编。具体编写分工如下:项目 1、项目 2、项目 13 和附录 2 由郭鑫编写;项目 3、项目 4、附录 1 和附录 3 由陈秀玲编写;项目 5、项目 6 和项目 8 由伞颖编写;项目 7 和项目 9 由白玲编写;项目 10、项目 11 和项目 12 由陈井霞编写。全书由郭鑫最后定稿。本书在编写过程中还得到了各编者单位有关领导的大力支持,在此深表谢意。全书由唐友教授主审。

由于编者水平有限,书中不当之处在所难免,恳请广大读者提出宝贵意见和建议,以便修订时加以完善。

编 者
2018 年 1 月

目录

项目 1 数据库基础知识

任务 1 数据库的发展概况	1
任务 2 关系数据库	8
任务 3 数据库系统设计	12
任务 4 SQL Server 2012 数据库简介	13
任务 5 E-R 设计	21
项目小结	23
习题	23

项目 2 T-SQL 语言基础

任务 1 T-SQL 语法规则	26
任务 2 运算符及优先级	30
任务 3 T-SQL 中的常量	34
任务 4 T-SQL 中的变量	36
任务 5 流程控制语句	39
项目小结	48
习题	49

项目 3 创建和管理数据库

任务 1 创建“学生管理”数据库	52
任务 2 管理“学生管理”数据库	64
任务 3 分离和附加数据库	70
任务 4 备份和还原	73
项目小结	77
习题	78

项目 4 创建和管理数据表

任务 1 创建数据表	80
任务 2 管理数据表	89
任务 3 约束	95
任务 4 规则	102
任务 5 数据库关系图	105

项目小结	107
习题	107

项目 5 视图

任务 1 创建视图“stud_view”	109
任务 2 修改视图“stud_view”	114
任务 3 利用视图“stud_view”更新基本表中数据	118
任务 4 查看视图“学生_view”	122
任务 5 删除视图“学生_view”	124
项目小结	126
习题	127

项目 6 触发器

任务 1 创建 DML 触发器“trigger_student1”	128
任务 2 创建 DDL 触发器“safe”	139
任务 3 修改 DML 触发器“trigger_student1”	142
项目小结	149
习题	149

项目 7 高级数据操作

任务 1 创建查询	152
任务 2 条件子句	160
任务 3 排序子句	171
任务 4 连接查询	176
任务 5 聚合函数	183
任务 6 分组汇总	186
任务 7 子查询与嵌套查询	193
任务 8 UPDATE, INSERT 和 DELETE 语句中的子查询	201
任务 9 并运算	206
项目小结	208
习题	208

项目 8 存储过程的创建与管理

任务 1 创建不带参数存储过程“proc_student1”	210
任务 2 创建带参数存储过程“s_count”	215
任务 3 查看存储过程“s_count”的相关信息及其创建代码	220
任务 4 用 T-SQL 语句修改存储过程“proc_student1”	222

任务 5 使用 SMSS 删除存储过程	224
项目小结	226
习题	227
项目 9 游标	
任务 1 声明游标	228
任务 2 打开游标	233
任务 3 读取游标中的数据	234
任务 4 关闭游标	237
任务 5 释放游标	238
项目小结	239
习题	239
项目 10 安全管理	
任务 1 SSMS 创建 SQL 账户	241
任务 2 SSMS 修改、删除 SQL 账户	243
任务 3 T-SQL 创建 SQL 账户	245
任务 4 T-SQL 修改、删除 SQL 账户	246
任务 5 SSMS 创建数据库用户	248
任务 6 T-SQL 创建数据库用户	250
任务 7 SSMS 创建服务器角色	252
任务 8 T-SQL 创建服务器角色	255
任务 9 SSMS 管理数据库角色	256
任务 10 T-SQL 管理数据库角色	259
任务 11 SSMS 管理权限	260
任务 12 T-SQL 管理权限	262
项目小结	263
习题	263
项目 11 函数	
任务 1 字符串函数	264
任务 2 数学函数	269
任务 3 日期和时间函数	276
任务 4 文本和图像函数	279
任务 5 系统函数	280
任务 6 数据类型转换函数	284
任务 7 用户自定义函数	285

项目小结	288
习题	288

项目 12 大数据金融

任务 1 大数据金融概述	289
任务 2 大数据金融的优势	294
任务 3 大数据金融模式	295
任务 4 大数据时代金融行业受到的冲击与变革	296
任务 5 大数据金融前瞻	301
任务 6 大数据金融跨界应用典型案例	301
任务 7 大数据时代金融数据库管理系统	303
项目小结	306
习题	306

项目 13 仓库管理系统案例实践

任务 1 仓库管理系统概要设计 E-R 图	307
任务 2 创建“仓库管理系统”数据库	309
任务 3 创建数据表	310
任务 4 创建视图	311
任务 5 创建触发器	312
任务 6 创建存储过程	313
任务 7 创建游标	313
项目小结	314

附录

附录 1 简单存储过程命令	315
附录 2 复杂存储过程命令	327
附录 3 SQL Server Management Studio 键盘快捷键	340

参考文献

项目 1

数据库基础知识

【项目描述】

本项目主要阐述了数据库的发展概况,数据库的基本概念、技术以及关系数据库基础理论,为用户学习 SQL Server 2012 数据库打下坚实的基础。

本项目重点是掌握关系数据库的基本概念;难点是理解关系数据库的常用术语,并能领会和灵活应用。包含的任务如表 1.1 所示。

表 1.1 项目 1 包含的任务

名 称	任务名称
项目 1 数据库基础知识	任务 1 数据库的发展概况
	任务 2 关系数据库
	任务 3 数据库系统设计
	任务 4 SQL Server 2012 数据库简介
	任务 5 E-R 设计

任务 1 数据库的发展概况

1.1.1 情境设置

某学校为了规范管理在校学生信息,统一建立数据库,实现全方位、无纸化办公管理,领导要求小张为在校学生建立“学生管理”数据库,统一利用计算机实现实时管理,于是小张开始学习数据库知识并为规范数据库作准备。

1.1.2 知识链接

1) 数据库系统的概念

(1) 数据

数据是事实或观察的结果,是对客观事物的逻辑归纳,是用于表示客观事物的未经加工的原始素材,既可以用数字表示(如身高、体重、大小等数值型数据),也可以用非数字形式表示(如字符、文字、图表、图形、图像、声音等非数值型数据)。

(2) 信息

信息则是经过加工后的数据,也就是有用的数据,是客观事物的特征通过一定物质载体形式的反映。信息是经过整理并通过分析、比较得出的推断或结论,能够反映客观事物的状态,和形式无关。

提示:信息与数据紧密联系又有区别。

数据是具体的符号,信息是抽象概念。数据有如原始材料,比如用户买的一份报纸,报纸上所有的内容都是数据,可用户不会把报纸上所有的数据看完,用户所看的或者关心的内容就是信息。

(3) 数据库

数据库(Database, DB)是以一定的组织方式将相关数据组织在一起,存储在外部存储介质上所形成的、能为多个用户共享的、与应用程序相互独立的相关数据集合的文件。在信息系统中,数据库是数据和数据库对象(如表、视图、存储过程等)的集合。数据库中的大量数据必须按一定的逻辑结构加以存储,目的是提高数据库中的数据的共享性、独立性、安全性以及较低的数据冗余度,以便对数据进行各种处理,并保证数据的一致性和完整性。

(4) 数据库管理系统

数据库管理系统(Database Management System, DBMS)是管理数据库的软件工具,是帮助用户创建、维护和使用数据库的软件系统。它建立在操作系统的基础上,实现对数据库的统一管理和操作,满足用户对数据库进行访问的各种需要。目前广泛运用的大型数据库管理系统软件有 Oracle, Sybase, DB2 等,而在 PC 机上广泛应用的则有 SQL Server, Visual Foxpro, Access 等。

(5) 数据库管理员

数据库管理员(Database Administrator, DBA)负责全面管理和控制数据库系统。数据库管理员是支持数据库系统的专业技术人员。数据库管理员的任务主要是决定数据库的内容,对数据库中的数据进行修改、维护,对数据库的运行状况进行监督,并且管理账号,备份和还原数据,以及提高数据库的运行效率。

(6) 数据库系统

数据库系统(Database System)泛指引入数据库技术后的系统,指在计算机系统中引入数据库后构成的系统,一般由数据库、数据库管理系统(及其开发工具)、应用系统、数据库管

理员和用户构成。

数据库系统是一个由硬件、软件(操作系统、数据库管理系统和编译系统等)、数据库和用户构成的完整计算机应用系统。数据库是数据库系统的核心和管理对象。因此,数据库系统的含义已经不仅仅是一个对数据进行管理的软件,也不仅仅是一个数据库,数据库系统是一个实际运行的,按照数据库方式存储、维护和向应用系统提供数据支持的系统,整体之间的关系如图 1.1 所示。

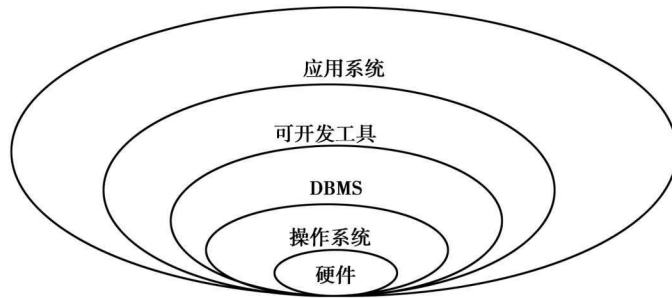


图 1.1 数据库在计算机中的地位

2) 数据管理技术的发展

数据管理技术是对数据进行分类、组织、编码、输入、存储、检索、维护和输出的技术。数据管理技术的发展大致经过了以下 3 个阶段:人工管理阶段、文件系统阶段、数据库管理系统阶段。

(1) 人工管理阶段

20 世纪 50 年代以前,计算机主要用于数值计算。从当时的硬件看,外存只有纸带、卡片、磁带,没有直接存取设备;从软件看(实际上,当时还未形成软件的整体概念),没有操作系统以及管理数据的软件;从数据看,数据量小,数据无结构,由用户直接管理,且数据间缺乏逻辑组织,数据依赖于特定的应用程序,缺乏独立性。特点:其一是数据不保存,只是在计算某一具体问题时将数据进行输入,运行后得到输出结果,输入、输出和中间结果均不保存;其二是数据不共享,冗余度大,一组数据只对应一个应用程序,即使多个应用程序使用相同的数据,也要各自定义,不能共享,导致冗余度大;其三是数据缺乏独立性,数据和程序是紧密结合在一起的,数据的逻辑结构、物理结构和存储方式都是由程序规定的,没有文件的概念,数据的组织形式完全是由程序员决定。

(2) 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,出现了磁鼓、磁盘等数据存储设备,出现了操作系统和专门的数据管理软件,称为文件系统。这种数据处理系统是把计算机中的数据组织成相互独立的数据文件,系统可以按照文件的名称对其进行访问,对文件中的记录进行存取,并可以实现对文件的修改、插入和删除。文件可以命名,应用程序可以“按文件访问、按记录进行读取”。文件系统实现了记录内的结构化,即给出了记录内各种数据间的关系,可以对文件进行修改、插入、删除操作。但是,文件从整体来看却是无结构的,其数据面向特定的应用程序,因此数据共享性、独立性差,且冗余度大,管理和维护的代价也很大。

(3) 数据库管理系统阶段

从 20 世纪 60 年代后期开始,计算机数据管理技术出现了数据库这样的数据管理技术阶段。硬件方面有了大容量的磁盘,软件方面出现了大量的系统软件;处理方式上,联机实时处理要求增多,并开始考虑和提出分布式处理。数据库的特点是数据不再只针对某一特定应用,而是面向全组织,具有整体的结构性,共享性高,冗余度小,并且实现了对数据进行统一的控制。

与文件系统不同的是,数据库系统是面向数据的而不是面向程序的,各个处理功能通过数据库管理软件从数据库中获取所需要的数据和存储处理结果。它克服了文件系统的缺点,为用户提供了一种更为方便、功能强大的数据库管理方法。

3) 数据库管理系统

数据库管理系统是以统一的方式管理、维护数据库中数据的一系列软件的集合,数据库管理系统在操作系统的支持与控制下运行。

用户一般不能直接加工和使用数据库中的数据,而必须通过数据库管理系统。数据库管理系统主要功能是维护数据库系统的正常活动,接受并响应用户对数据库的一切访问要求,包括建立及删除数据库文件,检索、统计、修改和组织数据库中的数据以及为用户提供对数据库的维护手段等。通过使用数据库管理系统,用户可以逻辑地、抽象地处理数据,而不必关心这些数据在计算机中的存放方式以及计算机处理数据的过程细节,把一切处理数据的具体而繁杂的工作交给数据库管理系统去完成。因此,在信息素养已经成为现代人的基本素质之一的信息社会里,学习并掌握一种数据库管理系统不但重要,而且必要。

数据库管理系统的功能归结起来主要有以下 4 点。

(1) 数据库定义(描述)功能

数据库管理系统提供数据描述语言(DDL)实现对数据库逻辑结构的定义以及数据之间联系描述。

(2) 数据库操纵功能

数据库管理系统提供数据操纵语言(DML)实现对数据库检索、插入、修改、删除等基本操作。DML 通常分为两类:一类是嵌入语言,如嵌入 C、VC++ 等高级语言中,这类 DML 一般不能独立使用,称为宿主型语言;另一类是交互命令语言,它语法简单,可独立使用,称为自含型语言。目前,数据库管理系统广泛采用的就是可独立使用的自含型语言,为用户和应用程序员提供操纵使用数据库的语言工具。本书介绍的 Visual FoxPro 6.0 提供的是自含型语言。

(3) 数据库管理功能

数据库管理系统提供了对数据库的建立、更新、结构维护以及恢复等管理功能。它是数据库管理系统运行的核心部分,所有数据库的操作都要在其统一管理下进行,以保证操作的正确执行,保证数据库的正确有效。

(4) 通信功能

数据库管理系统提供数据库与操作系统的联机处理接口以及用户与数据库的接口。作

为用户与数据库的接口,用户可以通过交互式 and 应用程序方式使用数据库。交互式直接明了,使用简单,通常借助 DML 对数据库中的数据进行操作;应用程序方式则是用户或应用程序员通过文本编辑器编写应用程序,实现对数据库中数据的各种操作。

4) 数据库系统

数据库系统是指在计算机系统中引入数据库后构成的系统。

数据库系统一般由 4 部分组成:数据库、数据库管理系统、计算机系统和人(数据库管理人员、用户)。

数据库系统的特点主要有以下 5 个方面:

(1) 数据共享

数据共享是数据库系统的目的,也是它的重要特点。数据共享是指多个用户可以同时存取数据而不相互影响,它包含 3 个方面的含义:所有用户可以同时存取数据;数据库不仅可以为当前用户服务,也可以为将来的新用户服务;可以使用多种语言完成与数据库的接口。

(2) 数据的独立性

数据独立是指数据与应用程序之间彼此独立,不存在相互依赖的关系。应用程序不必随数据存储结构的改变而改变,这是数据库的一个最基本的优点。

(3) 可控冗余度

数据冗余就是数据重复,数据冗余既浪费存储空间,又容易产生数据的不一致。在数据库系统中,由于数据集中使用,从理论上说可以消除冗余,但实际上出于提高检索速度等方面的考虑,常常允许部分冗余存在。这种冗余是可以由设计者控制的,故称为“可控冗余”。

(4) 数据的一致性

数据的一致性是指数据的不矛盾性。比如,在上述员工培训管理系统中,某员工的职称信息在员工基本信息中为“讲师”,而在员工培训需求信息中为“助讲”,这就称为数据不一致。如果数据有冗余,就容易引起数据的不一致。由于数据库能减少数据的冗余,同时提供对数据的各种检查和控制,保证在更新数据时能同时更新所有副本,维护了数据的一致性。

(5) 数据的安全性与完整性

数据库中加入了安全保密机制,可以防止对数据的非法存取。由于实行集中控制,有利于控制数据的完整性。数据库系统采取了并发访问控制,保证了数据的正确性。

5) 数据库系统的网络结构

(1) 大型数据库

大型数据库是由一台性能很强的计算机(称为主机或者数据库服务器)负责处理庞大的数据,用户通过终端机与大型主机相连,以存取数据。

(2) 本地小型数据库

在用户较少、数据量不大的情况下,可使用本地小型数据库。小型数据库一般是由个人建立的个人数据库。常用的个人数据库有 Access 和 FoxPro 等。

(3) 分布式数据库

分布式数据库是为了解决大型数据库反应缓慢的问题而提出的,它是由多台数据库服务器组成。数据可来自不同的服务器。

(4) 客户机/服务器数据库

在客户机/服务器数据库的网络结构中,数据库的处理可分为两个系统,即客户机(Client)和数据库服务器(Database Server),前者运行数据库应用程序,后者运行全部或部分数据库管理系统。在客户机上的数据库应用程序将请求通过网络发送给数据库服务器,数据库服务器对此请求进行搜索,并将用户所需的数据返回到客户机。

6) 数据模型

数据库中的数据是按照一定的逻辑结构存放的,这种结构是用数据模型来表示的。任何一种数据库管理系统都是基于某种数据模型的,目前比较普遍使用的数据模型有 3 种,即层次模型、网状模型和关系模型。

(1) 层次模型(Hierarchical Model)

层次模型犹如一棵倒置的大树,因此称为树形结构,用树形结构来表示数据以及数据之间的联系,数据对象之间是一种依次的一对一或一对多的联系,如图 1.2 所示。

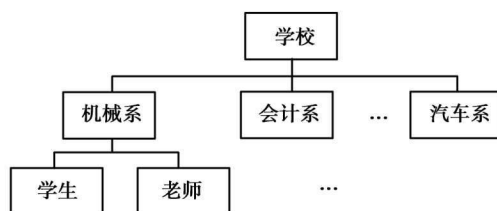


图 1.2 层次数据模型

层次结构的优点:结构简单,层次清晰,并且易于实现,适宜描述类似于目录结构、行政编制、家族关系等信息载体的数据结构。但层次模型不能直接表示多对多的联系,因而难以实现对复杂数据关系的描述。

(2) 网状模型(Network Model)

在网状模型中,各个实体之间建立的往往是一种层次不清的一对一、一对多或多对多的联系,用来表示数据之间复杂的逻辑关系。网状模型使用网状结构表示数据以及数据之间的联系。

网状模型的主要优点:表示数据之间的多对多联系时具有很大的灵活性,如图 1.3 所示。

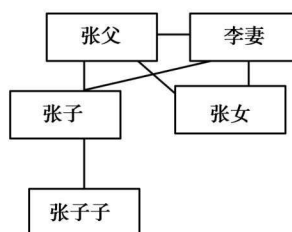


图 1.3 网状数据模型

(3) 关系模型 (Relational Model)

关系模型是一种理论最成熟、应用最广泛的数据模型。在关系模型中,数据存放在一种称为二维表格的逻辑单元中,整个数据库又是由若干个相互关联的二维表组成的。即用若干行和若干列构成的表格来描述数据集合以及它们之间的联系,如图 1.4 所示。

	学号	姓名	性别	出生日期	家庭住址	班级
1	20120102	李乐	0	1992-06-08 00:00:00	大庆市卧里屯	2012级计算机应用
2	20120203	赵小明	0	1905-06-13 00:00:00	绥化市	2012级计算机物联网
3	20120305	陈玲玲	1	1995-02-08 00:00:00	富裕县繁荣乡	2012级计算机网络
4	20120306	王晨瑞	0	1994-06-08 00:00:00	哈尔滨市南岗区	2012级计算机网络
5	20120308	王德凯	0	1994-08-09 00:00:00	伊春市	2012级计算机网络

图 1.4 关系数据模型

对于一个符合关系模型的二维表格,通常将表格中的每一行称为一条记录,而将每一列数据称为字段。一张二维表格若能构成一个关系模型的数据集合,通常具备以下条件:

- ①表中没有组合的列,也就是说每一列都是不可再分的。
- ②表中不允许有重复的字段。
- ③表中每一列数据的类型必须相同。
- ④在含有主关键字或唯一关键字字段的表中,不应该有内容完全相同的记录。
- ⑤在表中,行或列的顺序不影响表中各数据项之间的关系。

关系模型与层次模型、网状模型的主要区别在于:它描述数据的一致性,把每一数据子集都分别按照同一方法描述为一个关系,并让子集之间彼此独立,且不影响后续记录和字段的改变。在使用时,通过选择、筛选、投影等方法可以使数据之间或子集之间按照某种关系进行操作。因此关系数据库具有数据管理功能,数据表示能力较强,易于理解,使用更为方便。

7) 数据库管理系统功能

数据库管理系统的主要功能包括以下 4 个方面:

(1) 数据库定义功能

DBMS 提供了数据定义语言(Data Definition Language, DDL),用户通过它可以方便地对数据库中的数据对象进行定义。

(2) 数据操纵功能

DBMS 还提供了数据操纵语言(Data Manipulation Language, DML),用户可以使用 DML 数据操纵语言实现对数据库的基本操作,如查询、插入、删除和修改等。

(3) 数据库的控制和管理功能

数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制,以保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

(4) 数据字典

数据库管理系统通常提供数据字典功能,以便对数据库中数据的各种描述进行集中管理。数据字典中存放了系统中所有数据的定义和设置信息,如字段的属性、记录间的规则等。

任务 2 关系数据库

1.2.1 情境设置

小张了解了数据库的发展概况并掌握了数据库的基本功能,于是选择了关系数据库作为日后学习和使用的方向。

1.2.2 知识链接

1) 关系数据库

所谓关系数据库是指以关系数据模型为基础的数据库系统。

在关系理论中,有以下几个常见的关系术语:

①字段。它是关系数据库文件中最基本的、不可分割的数据单位。它用来描述某个对象的属性(在现实世界中,一个事物常常取若干特性来描述,这些特性成为属性),相当于二维表中的一列。

②记录。记录是描述某一个体的数据集合,它由若干个字段组成,相当于二维表中的一行。

③域。就是每个属性的取值范围,每个属性的取值范围对应一个值的集合,成为该属性的域(Domain)。

④关键字。在一个关系中有一个或多个字段的组合,其值能唯一辨别表格里的记录,便称为关键字。主关键字是用来唯一标识关系中记录的字段或字段组合;外部关键字是用于连接另一个关系,并且在另一个关系中为主关键字的字段。

⑤关系。一个关系就是一张二维表,每一列是一个相同属性的数据项,称为字段;每一行是一组属性的信息集合,称为记录。每个关系都有一个关系名,在 SQL Server 中关系称为表。

2) 关系的特点

在关系模型中,每一个关系都必须满足一定的条件,即关系必须规范化,一个规范化的关系必须具备以下几个特点:

①关系中的每个属性必须是不可分割的数据单元(即表中不能再包含表)。

②关系中的每一列元素必须是类型相同的数据。

③同一个关系中不能有相同的字段(属性),也不能有相同的记录。

④关系的行、列次序可以任意交换,不影响其信息内容。

3)关系的运算

把数据存入数据库是为了方便地使用这些数据。关系数据库管理系统为了便于用户使用,向用户提供了可以直接对数据库进行操作的查询语句。这种查询语句可以通过对关系(即二维表)的一系列运算来实现。

关系数据库系统至少应当支持3种关系运算,即选择、投影和连接。

(1)选择

选择是根据某些条件对关系作水平分割,即选择符合条件的记录,它是从行的角度对关系进行运算。

(2)投影

投影是从二维表中选出所需要的列,对关系进行垂直分割,消去某些列,并重新安排列的顺序,再删去重复元组。它是从列的角度对关系进行运算。

(3)连接

连接是同时涉及两个二维表的运算,它是将两个关系在给定的属性上满足给定条件的记录连接起来而得到的一个新的关系。

4)关系的完整性

数据完整性是指数据库中的数据在逻辑上的一致性和准确性。凡是数据都要遵守一定的约束,最简单的一个例子就是数据类型,如定义成整型的数据就不能是浮点数。由于数据库中的数据是持久和共享的,因此对于使用这些数据的单位来说,数据的正确与否显得非常重要。在关系数据库系统中,比较重要的完整性有实体完整性、域完整性、参照完整性和用户自定义的完整性等。

为了保证数据库的一致性和完整性,设计人员往往会设计过多的表间关联(Relation),尽可能地降低数据的冗余。表间关联是一种强制性措施,建立后,对父表(Parent Table)和子表(Child Table)的插入、更新、删除操作均要占用系统的开销。另外,最好不要用Identify属性字段作为主键与子表关联。如果数据冗余低,数据的完整性容易得到保证,但增加了表间连接查询的操作,为了提高系统的响应时间,合理的数据冗余也是必要的。使用规则(Rule)和约束(Check)来防止系统操作人员误输入造成数据的错误是设计人员的另一种常用手段。但是,不必要的规则和约束也会占用系统的非必要开销,需要注意的是,约束对数据的有效性验证要比规则快。所有这些,设计人员在设计阶段应根据系统操作的类型、频度加以均衡考虑。关系模型的完整性规则是对关系的某种约束条件。关系模型中可以有4类完整性约束,分别是实体完整性、域完整性、参照完整性和用户定义的完整性。

(1)实体完整性

一个基本关系通常对应现实世界的一个实体集。实体完整性又称行完整性,是指将行定义为特定表的唯一实体。要求表中有一个主键,并且其值不能为空且不允许有重复的值与之对应。实体完整性(通过索引、UNIQUE约束、PRIMARY KEY约束或IDENTITY属性来

实现)强制表的标识符列或主键的完整性。

(2) 域完整性

域完整性又称列完整性,是指给定列的输入有效性。强制域有效性的方法有:限制类型(通过数据类型)、格式(通过 CHECK 约束和规则)或可能值的范围(通过 FOREIGN KEY 约束、CHECK 约束、DEFAULT 定义、NOT NULL 定义和规则来实现)。

(3) 参照完整性

参照完整性又称引用完整性,是指主表中的数据与从表中的数据的一致性。在输入或删除其中一个表的记录时,另一个表对应的约束应满足,即参照完整性保持表之间已定义的关系。在 SQL Server 中,参照完整性基于外键与主键之间或外键与唯一键之间的关系(通过 FOREIGN KEY 和 CHECK 约束)。参照完整性确保键值在所有表中一致。这样的一致性要求不能引用不存在的值,如果键值更改了,那么在整个数据库中,对该键值的所有引用要进行一致的更改。

(4) 用户定义的完整性

实体完整性和参照性适用于任何关系数据库系统。除此之外,不同的关系数据库系统根据其应用环境的不同,往往还需要一些特殊的约束条件,用户定义的完整性就是针对某一具体关系数据库的约束条件,它反映某一具体应用所涉及的数据必须满足的语义要求。关系模型应提供定义和检验这类完整性的机制,以使用统一的系统方法处理它们,而不应由应用程序承担这一功能。

5) 关系模型的规范化

关系模型的范式有第一范式、第二范式、第三范式和 BCNF 范式等多种。

(1) 第一范式

第一范式(First Normal Form, 1NF)是其他范式的基础,是最基本的范式。它包括下列原则:

- ①数据组的每个属性只可以包含一个值。
- ②关系中的每个数组必须包含相同数量的值。
- ③关系中的每个数组一定不能相同。

如果关系模式 R 中的所有属性值都是不可再分解的原子值,那么就称此关系 R 是第一范式的关系模式。

(2) 第二范式

第二范式(Second Normal Form, 2NF)规定关系必须在第一范式中,并且关系中的所有属性依赖于整个候选键。候选键是一个或多个唯一标识每个数据组的属性集合。

(3) 第三范式

第三范式(Third Normal Form, 3NF)同第二范式一样依赖于关系的候选键。为了遵循第三范式的指导原则,关系必须在 2NF 中,非键属性相互之间必须无关,并且必须依赖于键。