

总主编 王会斌

成人高等教育学习指导用书（专科）

计算机 应用技术（下册）

主编 王会斌
贾元江

山东大学出版社

成人高等教育学习指导用书(专科)

计算机应用技术

(下册)

总主编 王会斌
主 编 王会斌 贾元江

山东大学出版社

图书在版编目(CIP)数据

计算机应用技术. 下册/王会斌, 贾元江主编.

—济南: 山东大学出版社, 2009. 6

成人高等教育学习指导用书. 专科

ISBN 978-7-5607-3862-8

I. 计…

II. ①王…②贾…

III. 电子计算机—成人教育: 高等教育—教学参考资料

IV. TP3

中国版本图书馆 CIP 数据核字(2009)第 085746 号

山东大学出版社出版发行

(山东省济南市山大南路 27 号 邮政编码: 250100)

山东省新华书店经销

日照日报印务中心印刷

787×1092 毫米 1/16 51.5 印张 1184 千字

2009 年 6 月第 1 版 2009 年 6 月第 1 次印刷

定价: 106.00 元(上、下册)

版权所有, 盗印必究

凡购本书, 如有缺页、倒页、脱页, 由本社营销部负责调换

《计算机应用技术》编委会

主 任 王志民

副主任 王会斌 刘 弘

委 员 郭玉锋 刘 浩 孙伯恒

贾元江 苑兆忠 宋忠义

主 编 王会斌 贾元江

副主编 苑兆忠 孙伯恒 张永胜

目 录

第七编 数据库技术.....	(417)
第1章 数据库系统概述.....	(417)
第2章 关系数据库系统模型.....	(421)
第3章 关系数据库的设计理论.....	(422)
第4章 数据库设计方法.....	(423)
第5章 结构化查询语言 SQL	(426)
第6章 SQL Server 关系数据库系统	(428)
第7章 SQL Server 2000 数据库应用与管理	(430)
第8章 Transact-SQL 语言结构	(430)
第9章 Transact-SQL 程序设计	(432)
第10章 数据库安全与维护	(434)
第11章 现代数据库技术	(436)
实验1 SQL Server 2000 的安装	(439)
实验2 企业管理器的使用	(440)
实验3 查询分析器的使用	(441)
实验4 创建数据库和表	(442)
实验5 表数据插入、修改和删除	(445)
实验6 数据库查询	(448)
实验7 自定义数据类型的使用	(452)
实验8 自定义函数的使用	(455)
实验9 索引的建立	(457)
实验10 默认值约束及默认值对象的应用	(457)
实验11 利用 CHECK 约束及规则对象实现数据的域完整性	(458)
实验12 利用 PRIMARY 约束实现数据的实体完整性	(460)
实验13 利用主键与外键关联实现数据的参照完整性	(460)
实验14 触发器的创建	(461)
实验15 存储过程的使用	(464)
实验16 数据库用户权限的设置	(466)
实验17 服务器角色的应用	(468)
实验18 数据库角色的应用	(469)

实验 19	数据库的备份	(469)
实验 20	数据库的恢复	(472)
实验 21	数据库的导入导出	(473)
第八编	面向对象技术	(478)
第 1 章	概 述	(478)
第 2 章	类与对象	(484)
第 3 章	继承和派生	(492)
第 4 章	多态性和虚函数	(497)
第 5 章	字符串和文件输入输出	(504)
第 6 章	模板与异常处理	(509)
实验 1	简单的 C++ 程序	(514)
实验 2	类与对象	(520)
实验 3	继承和派生	(523)
实验 4	多态性和虚函数	(526)
实验 5	字符串和文件输入输出	(530)
实验 6	模板与异常处理	(533)
第九编	计算机网络	(543)
第 1 章	概 述	(543)
第 2 章	物理层	(547)
第 3 章	数据链路层	(554)
第 4 章	网络层	(561)
第 5 章	运输层	(572)
第 6 章	应用层	(575)
第十编	多媒体技术	(583)
第 1 章	多媒体技术概论	(584)
第 2 章	多媒体音频技术	(589)
第 3 章	图形与图像处理	(596)
第 4 章	计算机动画技术	(607)
第 5 章	多媒体视频技术	(609)
实验 1	使用 Cool Edit Pro 2.1 实现卡拉 OK 歌曲录制	(614)
实验 2	图像处理和图像文件格式转换	(619)
实验 3	Flash 动画制作	(626)
实验 4	使用 Premiere Pro 1.5 实现电子相册	(634)

第十一编 软件工程	(645)
第1章 概 述.....	(645)
第2章 可行性研究.....	(651)
第3章 需求工程.....	(652)
第4章 设计工程.....	(655)
第5章 结构化分析.....	(657)
第6章 结构化设计.....	(668)
第7章 详细设计.....	(677)
第8章 面向对象的分析与设计.....	(685)
第9章 人机界面设计.....	(693)
第10章 程序设计语言与编码	(694)
第11章 软件测试	(695)
第12章 软件维护与再工程	(703)
第十二编 Java 程序设计	(712)
第1章 概 述.....	(712)
第2章 Java 语言语法基础	(714)
第3章 对象和类.....	(718)
第4章 继承和多态.....	(723)
第5章 抽象类、接口与内部类	(727)
第6章 数组、字符串及容器类	(731)
第7章 异常处理.....	(734)
第8章 Java I/O 系统	(737)
第9章 多线程.....	(741)
第10章 图形用户界面(GUI)	(749)
实验1 Java 运行环境的安装、配置与运行	(754)
实验2 Java 基本语法练习	(755)
实验3 面向对象编程练习	(758)
实验4 包、接口与异常处理(exception)	(759)
第十三编 网络管理与维护	(771)
第1章 计算机网络管理的基本概念.....	(772)
第2章 计算机网络管理的基本技术.....	(773)
第3章 网络管理协议.....	(774)
第4章 Windows 2000 网络操作系统	(777)
第5章 Windows Server 2003 网络操作系统简介	(780)
第6章 集线器与交换机的管理.....	(782)

第 7 章 路由器的管理.....	(784)
第 8 章 网络故障的诊断与网络维护.....	(786)
实验 1 双绞线的制作	(788)
实验 2 Windows 操作系统常用网络指令的使用	(790)
实验 3 交换机的基本配置方法	(795)
实验 4 交换机的常用配置命令	(797)
实验 5 交换机的端口配置	(798)
实验 6 无线局域网实验	(800)
后 记.....	(807)

第七编 数据库技术

【课程简介】

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。目前可见的绝大多数计算机应用系统都离不开数据库的支撑。

由于数据库技术在计算机技术中的重要作用,决定了“数据库技术”课程在计算机应用人才培养中的重要意义。尤其在当今知识经济的时代,信息化、数字化建设的决定因素是这方面的人才培养,信息处理为整个计算机应用的主要领域,其基础和核心属于数据库范畴。因此,“数据库技术”课程是计算机专业学生的一门专业性、技术性、实用性都非常重要的课程,是高职/大专计算机专业学生毕业后专业能力的重要体现。

本课程的重点要落实在数据库基础理论和应用两个方面,通过本课程的学习,要使学生理解数据库的基本概念、基本理论,掌握 SQL Server 数据库管理系统的基本操作技能,为日后的工作打下必要的专业能力基础。

本课程以 SQL Server 数据库管理系统为主要内容,在介绍关系数据库基础理论与基本知识的同时,主要讲解 SQL Server 2000 的安装与配置、用企业管理器及查询分析器对数据库进行管理与操作、Transact-SQL 编程、安全性管理、视图的使用、存储过程及触发器的管理等。

第 1 章 数据库系统概述

1.1 本章要点

本章阐述了数据库系统的基本概念,介绍了数据库管理技术的进展情况,数据库技术产生和发展的背景,数据库系统的组成以及数据库技术的应用等。本章还介绍了信息描述及三要素,三种常见的数据模型,数据库系统的三级模式结构。

本章的重要概念包括:①DB、DBMS、DBS 的定义;②数据库管理技术的发展阶段,包括人工管理阶段、文件系统阶段、数据库系统阶段和高级数据库技术阶段等各阶段的特点;③在数据描述方面,主要有概念设计、逻辑设计和物理设计等各阶段中数据描述的术语,概念设计中实体间二元联系的描述(1:1,1:N M:N);④在数据模型方面,主要有数据模型的定义,两类数据模型,逻辑模型的形式定义,ER 模型,层次模型,网状模型,关系模型和面向对象模型的数据结构以及联系的实现方式;⑤有关 DB 的体系结构,重要概念有

三级结构,两级映象,两级数据独立性;⑥DBMS的工作模式、主要功能和模块组成;⑦DBS的组成,DBA的含义,DBS的全局结构与DBS结构的分类。

1.2 典型习题与答案

1.2.1 试述数据管理技术发展的几个阶段及其特征。

【答】数据管理技术的发展分为:人工管理阶段、文件系统阶段、数据库系统阶段和高级数据库技术阶段。

(1)人工管理阶段

20世纪50年代中期以前,计算机主要用于科学计算。可使用的外部存储设备只有磁带、卡片、纸带,没有磁盘等直接存取设备。而且此时的计算机没有操作系统,没有管理数据的软件,数据处理方式是批处理,所有的数据完全由人工进行管理,因此这个阶段被称为人工管理阶段。

人工管理阶段主要有4个特点:数据不保存在计算机内;没有专用的软件对数据进行管理;只有程序的概念,没有文件的概念;数据面向程序。

(2)文件系统阶段

20世纪50年代后期到60年代中期,计算机不仅用于科学计算,还大量用于管理。磁盘、磁鼓等直接存取设备也相继使用,软件中有了高级语言和操作系统,在操作系统中已经有了专门的管理数据软件,称为文件系统,数据处理方式不再是单一的批处理,产生了联机实时处理的方式。

文件系统阶段主要有5个特点:数据以“文件”形式长期保存;数据的逻辑结构与物理结构有了区别;文件组织已多样化;数据面向应用;对数据的操作以记录为单位。

文件系统阶段主要有4个缺陷:数据冗余度大;数据不一致性;数据联系弱;数据和程序缺乏独立性。

(3)数据库系统阶段

20世纪60年代后期,计算机用于管理的规模越来越大,存储技术也取得了很大发展,有了大容量的磁盘存储数据;对联机实时处理的要求越来越多,并开始提出和考虑分步处理。在这样背景下,为了解决多用户、多应用共享数据的需求,人们着手开发和研制更加完美的数据管理模式,出现了数据库这样的数据管理技术,数据库的概念由此产生。进入20世纪70年代后,数据库技术得到了迅猛发展,成为计算机科学的一个重要分支。

进入数据库阶段的标志是20世纪60年代末发生的3件事情:1968年IBM公司研制的IMS系统是一个典型的层次DBS;1969年美国CODASYL组织DBTG报告,提出网状DBS的概念;1970年美国IBM公司的E. F. Codd发表论文,提出关系模型的思想。

数据库阶段的数据管理主要有5个特点:采用数据模型可以表示复杂的数据结构;有较高的数据独立性;为用户提供了方便的用户接口;提供了4个方面的数据控制功能:数据库的并发控制,数据库的恢复,数据的完整性,数据安全性;对数据的操作以数据项为单位,增加了系统的灵活性。

与文件结构相比,数据库结构主要有3点不同:①数据的结构化。文件系统中,文件由记录组成,但各文件之间缺乏联系。数据库中的数据在磁盘中仍以文件形式组织,但这

些文件之间有着广泛的联系。数据库的逻辑结构用数据模型来描述,整体结构化。数据模型不仅描述数据本身的特点,还要描述数据之间的联系。②数据独立性。文件只有设备独立性,而数据库还具有逻辑独立性和物理独立性。③访问数据的单位。访问文件中的数据,以记录为单位。访问数据库中的数据,以数据项(字段)为单位,增加了系统的灵活性。

(4)高级数据库技术阶段

20 世纪 80 年代以来,数据库技术在不同需求的驱动下得到了很大的发展,出现了分布式数据库、面向对象数据库和智能型知识数据库等等,通常被称为高级数据库技术。

分布式数据库主要有 3 个特点:数据物理上分布在各地,但逻辑上是一个整体;每个场地既可以执行局部应用,也可以执行全局应用;各地的计算机由数据通信网络相连接。

面向对象数据库主要有两个特点:面向对象数据模型能完整地描述现实世界的数据结构,能表达数据间嵌套、递归的联系;具有面向对象技术的封装性和继承性的特点,提高了软件的可重用性。

1.2.2 简述 ER 模型、层次模型、网状模型、关系模型和面向对象模型的主要特点。

【答】ER 模型直接表示实体类型及实体间联系,与计算机系统无关,充分反映用户的需求,用户容易理解。

层次模型的数据结构为树结构,记录之间联系通过指针实现,查询较快,但 DML 属于过程化的语言,操作复杂。

网状模型的数据结构为有向图,记录之间联系通过指针实现,查询较快,并且容易实现 M:N 联系,但 DML 属于过程化的语言,编程较复杂。

关系模型的数据结构均为二维表格,容易为初学者理解。记录之间联系通过关键码实现。DML 属于非过程化语言,编程较简单。

面向对象模型能完整描述现实世界的数据结构,具有丰富的表达能力,能表达嵌套、递归的数据结构。但涉及的知识面较广,用户较难理解,这种模型尚未普及。

1.2.3 数据之间的联系在各种结构数据模型中是怎样实现的?

【答】在层次、网状模型中,数据之间的联系通过指针实现的;在关系模型中,数据之间的联系通过外键和主键之间的联系实现的;在面向对象模型中,数据之间嵌套、递归联系通过对象标识符(OID)实现。

1.2.4 试述层次、网状、关系数据库的优缺点。

【答】(1)层次模型的优点主要有:①模型简单,对具有一对多层次关系的部门描述非常自然、直观,容易理解,这是层次数据库的突出优点;②层次模型的应用系统性能好,特别是对于那些实体间联系是固定的且预先定义好的应用,采用层次模型来实现,其性能优于关系模型;③层次数据模型提供了良好的完整性支持。

(2)层次模型的缺点主要有:①现实世界中很多联系是非层次性的,如多对多联系、一个结点具有多个双亲等,层次模型不能自然地表示这类联系,只能通过引入冗余数据或引入虚拟结点来解决;②对插入和删除操作的限制比较多;③查询子女结点必须通过双亲结点。

(3)网状数据模型的优点主要有:①能够更为直接地描述现实世界,如一个结点可以

有多个双亲;②具有良好的性能,存取效率较高。

(4)网状数据模型的缺点主要有:①结构比较复杂,而且随着应用环境的扩大,数据库的结构就变得越来越复杂,不利于最终用户掌握;②其 DDL、DML 语言复杂,用户不容易使用。由于记录之间联系是通过存取路径实现的,应用程序在访问数据库时必须选择适当的存取路径。因此,用户必须了解系统结构的细节,加重了编写应用程序的负担。

(5)关系数据库具有下列优点:①关系模型与非关系模型不同,它是建立在严格的数据概念的基础上的。②关系模型的概念单一,无论实体还是实体之间的联系都用关系表示,操作的对象和操作的结果都是关系,所以其数据结构简单、清晰,用户易懂易用。③关系模型的存取路径对用户透明,从而具有更高的数据独立性、更好的安全保密性,也简化了程序员的工作和数据库开发建立的工作。

(6)关系数据模型的主要缺点是:由于存取路径对用户透明,查询效率往往不如非关系模型。因此为了提高性能,必须对用户的查询请求进行优化,因此,对于数据量较大的应用状况,增加了开发数据库管理系统的难度。

1.2.5 数据库三级模式结构是什么? 三级模式结构描述了什么问题? 三级模式与用户、磁盘上的物理文件之间有什么联系和不同?

【答】(1)数据库的三级模式结构是:外模式、模式和内模式。

(2)数据库的三级模式结构是对数据的一个抽象级别,分别从外部(用户)级、概念级和内部级去观察数据库。

外部级是用户使用的局部数据库的逻辑结构,其描述称为外模式;概念级是数据库的整体逻辑结构,其描述称为概念模式;内部级是数据库的物理结构,其描述称为内模式。

(3)用户、外模式、概念模式、内模式和物理文件中的记录分别称为用户记录、外部记录、概念记录、内部记录和物理记录。

用户记录与外部记录的结构是一致的,它们之间只是数据传输问题;外部记录、概念记录和内部记录之间的结构可能是不一致的,除了数据传输问题,还有格式转换问题;内部记录和物理记录的结构是一致的,它们之间只是数据传输问题。

1.2.6 什么是 DBA? DBA 应具备什么素质? DBA 的职责是什么?

【答】DBA 是控制数据整体结构的一组人员,负责 DBS 的正常运行,承担创建、监控和维护 DB 结构的责任。

DBA 必须具备下列 4 方面素质:熟悉企业全部数据的性质和用途;对所有用户的需求有充分的了解;对系统的性能非常熟悉;兼有系统分析员和运筹学专家的品质和知识。

DBA 的主要职责有 8 点:定义模式,即定义数据库的信息内容和结构;定义内模式,即决定数据库的存储结构和存取策略;与用户的联络;定义安全性规则;定义完整性规则;DB 的转储与恢复;监督和控制数据库的使用和运行;改进和重组数据库系统。

1.2.7 系统分析员、数据库设计人员、应用程序员的职责是什么?

【答】系统分析员负责应用系统的需求分析和规范说明,系统分析员要和用户及 DBA 相结合,确定系统的硬件、软件配置,并参与数据库系统的概要设计。

数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。数据库设计人员必须参加用户需求调查和系统分析,然后进行数据库设计。在很多情况下,数据库设计

人员就由数据库管理员担任。

应用程序员负责设计和编写应用系统的程序模块,并进行调试和安装。

第2章 关系数据库系统模型

2.1 本章要点

本章系统讲解关系数据库的重要概念,并对关系模型进行阐述。关系模型包括关系数据结构、关系操作集合以及关系完整性约束三个组成部分。通过学习,要让学生掌握关系运算和关系演算的基本概念与基本方法,包括传统的关系运算、专门的关系运算、元组关系演算语言和域关系演算语言 QBE。其中域关系演算语言是选讲内容。

本章基本概念包括:关系模型,关键码(主键与外键),关系的定义和性质,3类完整性规则,ER模型到关系模型的转换规则,过程性语言和非过程性语言;本章介绍的关系代数包括5个基本操作,4个组合操作,7个扩充操作;本章的关系演算包括元组关系演算和域关系演算的原子公式、公式的定义。

本章的难点在于关系代数。关系代数较为抽象,学习过程中一定要结合具体的实例进行学习。同时要注意把握由具体语言到抽象语言的原则,即通过对具体语言如 ALPHA 和 QBE 的学习过渡到对抽象的关系演算的把握。

2.2 典型习题与答案

2.2.1 试述关系数据语言的特点和分类。

【答】关系数据语言可以分为3类,即:①关系代数语言,如 ISBL;②关系演算语言,包括元组关系演算语言(如 ALPHA, QUEL 等)和域关系演算语言(如 QBE);③具有关系代数和关系演算双重特点的语言,如 SQL。

这些关系数据语言的共同特点是:具有完备的表达能力;是非过程化的集合操作语言;功能强;能够嵌入高级语言中使用。

2.2.2 试述关系模型的完整性规则。

【答】关系模型的完整性规则是对关系的某种约束条件。关系模型中可以有三类完整性约束:实体完整性、参照完整性和用户完整性。

其中实体完整性和参照完整性是关系模型必须满足的完整性约束条件,被称做是关系的两个不变性,应该由关系系统自动支持。

(1)实体完整性规则:若属性 A 是基本关系 R 的主属性,则属性 A 不能取空值。

(2)参照完整性规则:若属性(或属性组) F 是基本关系 R 的外码,它与基本关系 S 的主码 K_s 相对应(基本关系 R 和 S 不一定是不同的关系),则对于 R 中每个元组在 F 上的取值必须为:①或者取空值(F 的每个属性值均为空值);②或者等于 S 中某个元组的主码值。

(3)用户定义的完整性是针对某一具体的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。

2.2.3 在参照完整性中,为什么外部码属性的值也可以为空?什么情况下才可以为空?

【答】在参照完整性中,外部码属性的值可以为空,它表示该属性的值尚未确定,但前提条件是该外部属性不是其所在关系的主属性。

例如,在下面“学生”表中,“专业号”是一个外部码,不是学生表的主属性,可以为空,其语义是,该学生的专业尚未确定:

学生(学号,姓名,性别,专业号,年龄)

专业(专业号,专业名)

而在下面的“选修”表中的“课程号”虽然也是一个外部码属性,但它同时也属于选修表的主属性,所以不能为空:

课程(课程号,课程名,学分)

选修(学号,课程号,成绩)

2.2.4 笛卡儿积、连接、等值连接、自然连接三者之间有什么区别?

【答】笛卡儿积是一个基本操作,而连接、等值连接和自然连接是组合操作。

设关系 R 的属性个数(元数)为 r ,元组个数(基数)为 m ;关系 S 的属性个数为 s ,元组个数为 n 。那么 $R \times S$ (笛卡儿积)的属性个数为 $r+s$,元组个数为 $m \times n$; $R \bowtie_{\theta} S$ (连接)和 $R \bowtie S$ (等值连接)的属性个数都为 $r+s$,元组个数都是小于或等于 $m \times n$; $R \bowtie S$ (自然连接)的属性个数(元数)小于或等于 $r+s$,元组个数小于或等于 $m \times n$ 。

第3章 关系数据库的设计理论

3.1 本章要点

通过本章学习需要掌握实体类型的属性关系,即一对一关系、一对多关系和多对多关系;要掌握实体属性值之间相互联系和相互的关系,数据的函数依赖与多值依赖的含义,及其有关的概念;对于关系数据库模式的规范化理论,需要掌握第一范式到第三范式的定义及其表现形式,对更高级的范式只需知道既可;在关系模式的分解算法方面,只需要掌握针对具体实例的分解方法即可,至于相关的理论及一般性算法,属于选讲内容。

3.2 典型习题与答案

3.2.1 对函数依赖 $X \rightarrow Y$ 的定义加以扩充, X 和 Y 可以为空属性集,用 ϕ 表示。那么 $X \rightarrow \phi, \phi \rightarrow Y, \phi \rightarrow \phi$ 的含义是什么?

【答】据推理规则的自反律可知: $X \rightarrow \phi, \phi \rightarrow \phi$ 是平凡的 FD,总是成立。

$\phi \rightarrow Y$ 表示在当前关系中,任意两个元组的 Y 值相等,也就是当前关系的 Y 值都是相等的。

3.2.2 设关系模式 R 有 n 个属性,在模式 R 上可能成立的函数依赖有多少个?其中平凡的 FD 有多少个?非平凡的 FD 有多少个?

【答】FD形为 $X \rightarrow Y$, 从 n 个属性值中选择属性组成 X , 共有 $C_n^0 + C_n^1 + \cdots + C_n^n = 2^n$ 种方法; 同理, 组成 Y 也有 2^n 种方法。因此组成 $X \rightarrow Y$ 形式应该有 $2^n \times 2^n = 4^n$ 种方法, 即可能成立的 FD 有 4^n 个。

平凡的 FD 要求 $Y \subseteq X$ 组合 $X \rightarrow Y$ 形式的选择有:

$$\begin{aligned} & C_n^0 C_n^0 + C_n^1 (C_n^0 + C_n^1) + \cdots + C_n^n (C_n^0 + C_n^1 + \cdots + C_n^n) \\ &= C_n^0 2^0 + C_n^1 2^1 + \cdots + C_n^n 2^n = (1+2)^n = 3^n \text{ 个} \end{aligned}$$

3.2.3 为什么要进行关系模式的分解? 分解的依据是什么?

【答】由于数据之间存在着联系和约束, 在关系模式的关系中可能会存在数据冗余和操作异常现象, 因此需把关系模式进行分解, 以消除冗余和异常现象。

分解的依据是数据依赖和关系模式的范式。

3.2.4 在数据库的设计实践中, 是不是关系设计的规范化标准越高越好? 在实践中数据库设计有没有什么原则?

【答】不是。在某些情况下, 关系分解前的非规范化的表可能更适用一些, 因为当所需数据必须从多个表组合而来时, DBMS 需要做一些额外的连接工作, 延迟了系统响应的时间。所以, 在数据库的设计实践中, 关系有时故意保留成非规范化的模式, 甚至在规范化后又进行反规范化处理, 这样做通常是为了改善数据库的性能。因此将关系分解到什么程度, 要根据实际情况来决定。

对大多数商业系统来说, 一般分解到 3NF 就够了, 所以常把 3NF 称为“工程范式”。

从数据库设计实践的角度, 有下面几条经验原则: ①部分函数依赖和传递函数依赖的存在是产生数据冗余、更新异常的重要原因。因此, 在关系规范化中, 应尽可能消除属性间的这些依赖关系。②1NF、2NF 以至非规范化的模式, 由于它们性能上的弱点, 一般不宜作为数据库模式。③由于 3NF 中不存在非主属性对关键字的部分依赖和传递依赖关系, 因而消除了很大一部分冗余和更新异常, 具有较好的性能。所以一般要求数据库设计达到 3NF。

第 4 章 数据库设计方法

4.1 本章要点

本章讲解数据库系统的设计过程、设计方法和技术, 内容的实践性较强。数据库系统的设计实质上是软件工程中有关数据库部分的实施过程, 所以它的设计过程与设计方法要遵循软件工程的一般性准则, 即包括需求分析、设计、编码, 以及数据库的实施、运行与维护等全过程。具体“设计”过程, 数据库设计可细分为概念设计、逻辑设计与物理设计。

读者需要熟练掌握的是 ER 图的设计, 包括 ER 图向关系模型的转换。

本章技术上的难点即是 ER 图的设计, 数据模型的优化。真正的难点是理论与实际的结合。读者一般缺乏实际经验, 缺乏解决实际问题的能力, 特别是缺乏应用领域的知识。而实际软件开发中的数据库设计, 需要设计人员对应用环境、专业业务有具体深入的了解, 才能设计出符合具体领域要求的数据库及其应用系统。

4.2 典型习题与答案

4.2.1 简述数据库设计各个阶段的设计内容。

【答】规划:主要进行问题定义,确定问题性质;同时进行可行性研究,包括技术可行性、经济可行性、政策可行性等;还要确定待建立的数据库系统在组织和信息系统中的地位,以及各个 DBS 之间的联系。

需求分析:准确了解与分析用户需求,包括数据需求和处理需求。需求分析阶段的工作主要包括:分析用户活动,产生用户活动图;确定系统范围,产生系统流程图;分析用户活动所涉及的数据,产生数据流图;分析系统数据,产生数据字典。需求分析的目标是对系统的整个应用情况作全面地、详细地调查,确定企业组织的目标,收集支持系统总的设计目标的基础数据和对这些数据的要求,确定用户的需求;并把这些需求写成用户和数据库设计者都能接受的文档。

概念设计的主要步骤分为:进行数据抽象,设计局部概念模式;将局部概念模式综合成全局概念模式;对结果进行评审,即对用户需求进行综合、归纳和抽象,形成一个独立于具体 DBMS 的概念模型。

逻辑设计的目的是把概念设计阶段设计好的基本 ER 图转换成与选用的具体机器上 DBMS 所支持的数据模型相符合的逻辑结构。这个结构在功能、性能、完整性和一致性约束、数据库的可扩充性等方面均应满足用户要求。主要步骤是:形成初始模式、设计子模式、设计应用程序梗概、评价模式和修改模式。

物理设计:为逻辑数据模型选取一个最适合应用环境的物理结构,包括存储结构和存取方法。主要分为五步:确定 DB 的存储记录结构;确定数据存储安排;存取方法的设计;完整性和安全性的设计;应用程序设计。

实现阶段也就是数据库实施阶段,这个阶段中数据库设计人员运用 DBMS 提供的数据库语言、工具及宿主语言,根据逻辑设计和物理设计的结果建立数据库,编制与调试应用程序,组织数据入库,并进行试运行。

数据库系统投入运行后就进入运行维护阶段,即在计算机系统运行过程中对其进行评价、调整与修改。主要工作有四项:维护 DB 的安全性、完整性及系统的转储和恢复;DB 性能的监督、分析和改进;增加 DB 新功能;改正运行中发现的系统错误。

4.2.2 试述数据库设计过程中结构设计部分形成的数据库模式。

【答】数据库结构设计的不同阶段形成数据库的各级模式,即:

(1)在概念设计阶段形成独立于机器特点,独立于各个 DBMS 产品的概念模式,即 ER 图。

(2)在逻辑设计阶段将 ER 图转换成具体的数据库产品支持的数据模型,如关系模型,形成数据库逻辑模式,然后在基本表的基础上再建立必要的视图,形成数据的外模式。

(3)在物理设计阶段,根据 DBMS 特点和处理的需要,进行物理存储安排,建立索引,形成数据库内模式。

其中,概念模式是面向用户和设计人员的,属于概念模型的层次;逻辑模式、外模式、内模式是 DBMS 支持的模式,属于数据模型的层次,可以在 DBMS 中加以描述和存储。

4.2.3 为什么说 ER 模型具有客观性和主观性两重含义?

【答】ER 模型是人们认识客观世界的一种方法和工具,它具有客观性和主观性两重含义。

ER 模型是在客观事物或系统的基础上形成的,在某种程度上反映了客观现实,反映了用户的需求,因此 ER 模型具有客观性。但 ER 模型又不等同于客观事物本身,它往往反映事物的某一方面,至于选取哪个方面或哪些属性,如何表达则决定于观察者本身的目的与状态,从这个意义上说,ER 模型又具有主观性。

4.2.4 概念模型的意义是什么? 概念模型有什么特点? 概念模型一般用什么工具描述?

【答】概念设计的任务是在需求分析阶段产生的需求说明书的基础上,按照特定的方法把它们抽象为一个不依赖于任何具体机器的数据模型,即概念模型。概念模型使设计者的注意力能够从复杂的实现细节中解脱出来,而只集中在最重要的信息的组织结构和处理模式上。

概念模型具有以下特点:①概念模型是对现实世界的抽象和概括,它真实、充分地反映了现实世界中事物和事物之间的联系,能满足用户对数据的处理要求。②由于概念模型简洁、明晰、独立于机器、很容易理解,因此可以用概念模型和不熟悉计算机的用户交换意见,使用户能积极参与数据库的设计工作,保证设计工作进行顺利。③概念模型易于更新,当应用环境和应用要求改变时,容易对概念模型修改和扩充。④概念模型很容易向关系、网状、层次等各种数据模型转换。

描述概念模型的有力工具是 ER 图。

4.2.5 概念模型向关系模型的转换规则是什么?

【答】关系模型是由一组关系(二维表)的结合,而 ER 模型则是由实体、实体的属性、实体间的联系三个要素组成。所以要将 ER 模型转换为关系模型,就是将实体、属性和联系都要转换为相应的关系模式。具体的转换规则是:

(1)一个实体类型转换为一个关系模式:将每种实体类转换为一个关系,实体的属性就是关系的属性,实体的关键字就是关系的关键字。

(2)一个一对一(1:1)联系有以下两种转换方式:①转换为一个独立的关系模式。联系名为关系模式名,与该联系相连的两个实体的关键字及联系本身的属性为关系模式的属性,其中每个实体的关键字均是关系模式的候选键。②与任意一端的关系模式合并。可将相关的两个实体分别转换为两个关系,并在任意一个关系的属性中加入另一个关系的主关键字。

(3)一个一对多(1:n)联系也有两种转换方式:①转换为一个独立的关系模式。联系名为关系模式名,与该联系相连的各实体的关键字及联系本身的属性为关系模式的属性,关系模式的关键字为 n 端实体的关键字。②将 1:n 联系与 n 端关系合并。1 端的关键字及联系的属性并入 n 端的关系模式即可。

(4)一个多对多(m:n)联系转换为一个关系模式。关系模式名为联系名,与该联系相连的各实体的关键字及联系本身的属性为关系模式的属性,关系模式的关键字为联系中各实体关键字的并集。