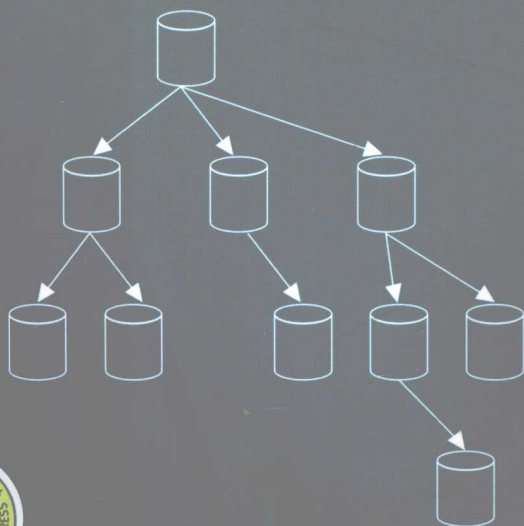


数据库系统原理 习题与解析

葛洪伟 罗海驰 王华君 姜代红 编著

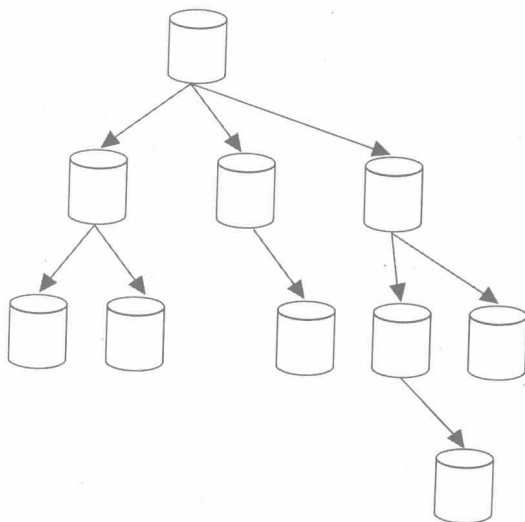


中国电力出版社

www.cepp.com.cn

数据库系统原理 习题与解析

葛洪伟 罗海驰 王华君 姜代红 编著



中国电力出版社

www.cepp.com.cn

内 容 提 要

本书为《数据库系统原理与应用》一书的教学参考书籍。全书共分六章,每章包括复习纲要、例题分析和习题三部分。书中所有习题均给出了答案,便于学生练习和自测。本书内容丰富,习题覆盖面广,难易程度层次分明;在形式和内容上既突出了基础知识、基本概念的复习巩固,也注重学生数据库应用能力的培养。

本习题集可作为计算机、信息管理、电子商务及相关专业的本、专科学生学习数据库系统原理课程的参考书,也适用于参加计算机等级考试者研习。

图书在版编目(CIP)数据

数据库系统原理习题与解析 / 葛洪伟等编著. —北京: 中国电力出版社, 2008
21 世纪高等学校规划教材
ISBN 978-7-5083-7237-2

I. 数… II. 葛… III. 数据库系统 - 高等学校 - 习题 IV. TP311.13-44

中国版本图书馆 CIP 数据核字 (2008) 第 100214 号

责任编辑: 夏华香

责任校对: 崔燕菊

责任印制: 郭华清

书 名: 数据库系统原理习题与解析

编 著: 葛洪伟 罗海驰 王华君 姜代红

出版发行: 中国电力出版社

地址: 北京市三里河路 6 号 邮政编码: 100044

电话: (010) 68362602 传真: (010) 68316497

印 刷: 汇鑫印务有限公司

开本尺寸: 185mm × 260mm 印 张: 12.25 字 数: 293 千字

书 号: ISBN 978-7-5083-7237-2

版 次: 2008 年 8 月北京第 1 版

印 次: 2008 年 8 月第 1 次印刷

印 数: 0001—4000 册

定 价: 19.00 元

敬 告 读 者

本书封面贴有防伪标签,加热后中心图案消失

本书如有印装质量问题,我社发行部负责退换

版 权 专 有 翻 印 必 究

前 言

数据库技术是计算机领域发展最快的学科之一，也是实用性很强且应用最为广泛的技术之一，它已成为计算机信息系统与应用系统的核心技术和重要基础。而作为数据库技术教育的“数据库系统原理”课程也历来是我国高等院校计算机科学与技术、信息管理等相关专业必修的一门主干课程，也是电子商务及其他相关专业的重要课程。

数据库系统原理课程的学习有一定的难度。为了给学生一些启发和帮助，作者根据长年讲授该课程的经验并充分考虑目前普通高校学生的知识水平、能力、素质特点和实际教学情况，编写了本书。该习题集是中国电力出版社出版的《数据库系统原理与应用》一书的配套书籍。希望大家在注重课堂学习的同时，通过对习题的解答，能够深化对基本概念的理解，掌握该课程求解问题的思路与方法，提高分析问题和解决问题的能力。

本书遵循数据库系统原理课程教学大纲的要求，将内容分为6章，每章包括复习纲要、例题分析和习题三部分。复习纲要主要是对各章中知识点进行简明扼要的讲解与综述；例题分析部分主要是针对各章中的重点、难点举例进行详细解析，加深学生对相关章节主要内容的理解和消化；而习题部分包括选择、填充等基本题和简答、设计、综合等解析题，所有题目均给出答案，以便学生对照检查自身的不足。

本书在编写过程中，得到了江南大学信息学院领导的大力支持和许多同事的鼓励与帮助，李婷和黄向前老师也参与了部分编写和校对工作，在此向他们表示衷心的感谢。另外，书末所列的参考文献及有关资料给了我们极大的帮助和启发，在此向与这些文献及资料的著作权人一并表示由衷的谢意。

限于习题较多、时间仓促以及知识水平，解答中可能存在不准确和不妥之处，敬请广大读者批评指正。

作 者

2008年7月于江南大学

目 录

前 言

第 1 章 数据库系统概述	1
第一部分 复习纲要	1
第二部分 例题分析	9
第三部分 习题与解答	11
▶ 基本题	11
▶ 解析题	27
第 2 章 关系模型和关系运算理论	33
第一部分 复习纲要	33
第二部分 例题分析	35
第三部分 习题与解答	39
▶ 基本题	39
▶ 解析题	48
第 3 章 关系数据库语言 SQL	58
第一部分 复习纲要	58
第二部分 例题分析	62
第三部分 习题与解答	66
▶ 基本题	66
▶ 解析题	75
第 4 章 关系的规范化理论	99
第一部分 复习纲要	99
第二部分 例题分析	107
第三部分 习题与解答	113
▶ 基本题	113
▶ 解析题	119
第 5 章 关系数据库设计	132
第一部分 复习纲要	132
第二部分 例题分析	140

第三部分 习题与解答	145
▶ 基本题	145
▶ 解析题	151
第6章 数据库安全保护	160
第一部分 复习纲要	160
第二部分 例题分析	166
第三部分 习题与解答	170
▶ 基本题	170
▶ 解析题	179
参考文献	187

第 1 章 数据库系统概述

第一部分 复习纲要

1. 数据管理技术的演变

数据处理是对数据进行加工,使其成为有用信息的过程。随着应用领域的不断扩大,计算机所处理的数据的类型和数量都大幅度增加,所用数据处理技术也不断发展进步。

数据管理是数据处理工作中的一项核心任务,主要包括数据的编码、存储、查询、维护等操作。从诞生迄今,计算机上的数据管理技术已经历了人工管理、文件系统、数据库系统三个阶段。

20 世纪 50 年代中期以前为人工管理阶段,该阶段的主要特点:各应用程序独立使用并管理对应的数据集;数据不在计算机系统内长期保存,不在应用程序间共享;若要对数据的逻辑结构或物理结构进行修改,则应用程序也要做出相应的修改。

20 世纪 50 年代后期到 60 年代中期为文件系统阶段,该阶段的主要特点:数据以数据文件的形式长期保存在计算机系统内;数据文件由操作系统进行统一管理,应用程序在操作系统的支持下按文件名对数据文件进行访问;数据物理结构的一些变动,可由操作系统进行处理调整,不影响应用程序的调用,但数据逻辑结构的变化仍将导致应用程序的修改需求;程序间的部分数据共享仍比较困难,数据冗余的状况依旧存在。

20 世纪 60 年代后期以来为数据库系统阶段,该阶段的主要特点:数据以数据库文件的形式长期保存在计算机系统内;数据库文件由专门的数据库管理系统进行管理,应用程序在数据库管理系统的支持下可对库中海量数据进行高效访问;库中数据具有物理独立性和逻辑独立性,数据物理结构和逻辑结构的变化不会影响建于其上的应用程序,从而大幅度减少了应用程序的维护和修改的工作量。

2. 数据库文件与传统数据文件间的主要区别

(1) 数据文件中的数据虽然已经以记录的形式进行组织,但记录之间的联系还并没有能够体现出来,而数据库文件则不仅以记录的形式来组织数据,还能表达记录之间各种复杂的联系。

(2) 数据文件仅面向单个应用组织数据,并可针对该应用的实际情况、具体要求进行优化,但如果想在该数据文件的基础上建立其他应用则会觉得不方便,数据库文件是面向数据本身及数据间的内在联系来组织数据的,并不特意面向某一应用,也因而便于在其上建立各种应用,数据在多用户多应用间共享,极大地降低了数据的冗余度,并提高了数据的一致性和可信度。

(3) 数据文件中,最小的存取单位为记录,而在数据库文件中,最小的存取单位可以是数据项。

3. 数据库系统的组成

数据库系统是引入了数据库技术的计算机系统,除了一般计算机系统的组成部分外,还包

含了数据库、数据库管理系统、数据库应用系统以及与系统开发、使用相关的各类人员。

数据库是按一定结构组织起来的数据的集合。数据库以数据库文件的形式保存在计算机系统的存储设备中，可为多个应用系统或同一应用系统的多个用户所共享。

数据库管理系统是运行于操作系统之上的，专用于对数据库进行管理的系统软件，主要负责在创建、使用和维护数据库时对库进行统一管理、控制和保护，通常还会提供一套支持用户进行应用开发的工具。

数据库应用系统是基于数据库创建的，能实现用户相关实际需求的应用系统。数据库应用系统一般采用高级程序设计语言进行开发，并通过符合某种规范的接口与数据库管理系统进行通信，借助数据库管理系统的功能，实现对库中数据的访问。

数据库系统相关人员指的是那些与数据库的设计、实现、管理、使用相关的人员，包括系统分析员、数据库设计人员、应用程序员、数据库管理员和最终用户。

4. 数据库管理员的主要任务

数据库管理员是用户单位中专门负责数据库的建立、使用和维护的人员，其主要任务包括如下几项。

(1) 参与数据库设计各个阶段的工作，协调各最终用户对数据、数据处理、数据的安全控制和完整性约束的需求。

(2) 与数据库设计人员一起完成数据库管理系统的安装、设置，配合数据库设计人员按既定设计，借助数据库管理系统提供的开发工具，建立数据库及库中各对象，加载初始数据，配合应用程序员安装调试所开发的数据库应用系统。

(3) 使用各类监测工具，对数据库应用系统的日常运行情况及数据库的访问情况进行监督，通过对监测数据的分析，寻找改进系统性能、提高数据库访问效率、提高数据库安全性的方法。

(4) 制定合理、适用的数据库备份策略，并按既定策略及时做好数据库的备份工作。

(5) 及时处理数据库运行过程中发生的各类突发事件，一旦系统发生故障，在排除故障后应利用备份数据尽快将数据库恢复到故障发生前的某个一致状态，尽量降低因故障而造成的损失。

(6) 及时了解用户对数据库及数据库应用系统的意见、建议及新的需求，帮助他们解决相关的技术问题，指导他们正确、充分地利用数据库。

(7) 确定库中各数据对象的保密级别，定义各用户对数据库的访问权限，制定必要的规章制度并负责实施。

5. 数据库系统的三级模式结构

数据库系统的三级模式结构指数据库系统是由外模式、逻辑模式和内模式三级构成的。

逻辑模式也称模式，是对数据库中全体数据的逻辑结构、完整性约束条件、安全性要求等内容的描述。一个数据库只有一个逻辑模式。

外模式又称用户模式或子模式，是对数据库中面向某数据库应用系统（或者说是面向某些数据库用户）的部分数据的逻辑结构、完整性约束条件、安全性要求等内容的描述。一个数据库可以有多个外模式。

内模式也称存储模式，是对数据库中全体数据的物理结构、存储方式、存储策略等的描述。一个数据库只有一个内模式。

6. 三级模式间的两级映像

数据库管理系统支持外模式与逻辑模式间以及逻辑模式与内模式间的两层映像，以便在这

三个抽象层次间进行数据的映射和转换。

一个数据库可以有多个外模式，但有且仅有一个逻辑模式。对于每一个外模式，数据库系统都会建立相应的外模式—逻辑模式映像，这种映像定义了数据的局部逻辑结构与全局逻辑结构之间的对应关系。

一个数据库有且仅有一个逻辑模式，也有且仅有一个内模式，所以逻辑模式—内模式映像是最唯一的，它定义了数据的全局逻辑结构与存储结构之间的对应关系。

7. 数据的逻辑独立性和物理独立性

数据库系统的三级模式两级映像实现了数据的逻辑独立性和物理独立性。

当逻辑模式发生变化时，可通过修改外模式—逻辑模式映像使外模式（即数据的局部逻辑结构）保持不变，这样基于外模式设计的应用程序就不必修改，从而保证了数据的逻辑独立性。

当内模式发生变化时，可通过修改逻辑模式—内模式映像使逻辑模式（即数据的全局逻辑结构）保持不变，那么外模式以及基于外模式设计的应用程序也就可以保持不变，从而保证了数据的物理独立性。

8. 信息的三个世界

信息的三个世界指现实世界、信息世界和数据世界。

现实世界指的是人类生存于其中的那个物质的、可以感知的客观世界。信息世界反映了人对现实世界的理解和认识。数据世界即计算机世界，以 01 数字串映像了客观世界中的万事万物。三个世界间的联系如图 1-1 所示。

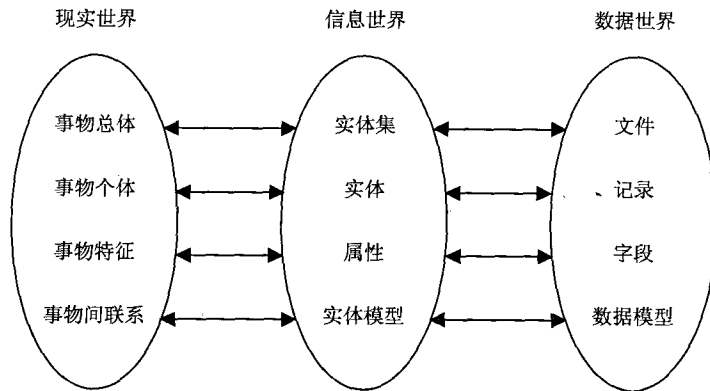


图 1-1 三个世界间的联系

9. 实体、属性、域、实体型、实体集、关键字、实体集间的联系

在信息世界中，客观存在并可相互区别的事物称为实体。实体的特征用属性来描述。属性的取值范围称为属性的域。实体型是对同类实体的抽象，记为：实体类别名（属性 1，属性 2，…，属性 n ）。同型实体的集合称为实体集。可用以标识同一实体集中不同实体的属性集称为关键字。

实体集之间的联系有一对一、一对多和多对多三种。设有实体集 X 和实体集 Y ，若对于实体集 X 中的每个实体，实体集 Y 中有 0 或 1 个实体与之关联，而对于实体集 Y 中的每个实体，实体集 X 中也有 0 或 1 个实体与之关联，则称实体集 X 和实体集 Y 间具有一对一联系；若对于实体集 X 中的每个实体，实体集 Y 中有 0 或 1 或多个实体与之关联，而对于实体集 Y 中的每个实体，实体集 X 中只有 0 或 1 个实体与之关联，则称实体集 X 和实体集 Y 间具有一对多

联系；若对于实体集 X 中的每个实体，实体集 Y 中有 0 或 1 或多个实体与之关联，而对于实体集 Y 中的每个实体，实体集 X 中也有 0 或 1 或多个实体与之关联，则称实体集 X 和实体集 Y 之间具有多对多联系。

10. 概念模型

概念模型用于表达人对现实世界的认识，其主要特点如下。

(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的事物及事物间的联系，应有较强的语义表达能力。

(2) 在数据库设计过程中，数据库设计人员用概念模型表达自己对系统所处理的数据及数据间联系的理解，并用这一模型与用户进行交流沟通，使自己的理解与实际情况相符，所以概念模型应该直观、简洁、清晰、易懂，可以用来与不具备计算机专业知识的用户交流意见，便于用户参与数据库的设计工作，从而保证设计工作的效率和质量。

(3) 概念模型便于维护，当现实世界或用户需求发生变化时，可以方便地对概念模型进行修改和扩充。

(4) 概念模型不依赖于开发系统时打算采用的数据库管理系统，所以概念模型还应当能够方便地转换为各种数据库管理系统所支持的逻辑模型，如关系模型、网状模型、层次模型、关系—对象模型等。

11. 逻辑模型三要素

逻辑模型从数据库管理系统实现的角度出发，对所有可能保存在数据库中的数据及数据间的联系进行分析、建模。逻辑模型的三要素是数据结构、数据操作和完整性约束条件。

数据结构是库中数据对象类型的集合，描述了数据的静态特征。这些数据对象可分为两大类，一类是对数据的内容、类型等的描述，另一类是对数据间联系的描述。逻辑模型三要素中，数据结构是最重要也是最根本的一个，因此常以之为逻辑模型命名。

数据操作是可以对库中数据对象执行操作的集合，描述了数据的动态特征。这些操作主要包括两大类，一类是查询，另一类是修改。逻辑模型应对操作的符号、含义、规则以及语言进行定义。

完整性约束条件是库中数据对象必须满足的完整性规则的集合。完整性规则是库中数据及数据间联系应满足的制约和依存规则，用以限定数据库的状态以及状态的变化，以保证库中数据正确、有效和相容。逻辑模型应提供定义完整性约束条件的手段。

12. 层次模型

层次模型以树型结构来表示各类实体以及实体间的联系。模型中的每个记录由若干字段组成，每个字段描述实体的一个属性，在这些字段中可以选择一个作为排序字段。这些记录只有按路径查看时，才能显出其全部意义。

层次模型可以自然地表示一对多联系，一对一联系作为一对多联系的特例，也可以方便地表示，而多对多联系则需分解为若干一对多联系才能表示，常用的分解方法有冗余节点法和虚拟节点法两种。

层次模型支持的数据操作主要有查询、插入、删除和更新，其中执行插入、删除、更新操作时要满足层次模型的完整性约束条件，包括：①进行插入操作时，如果没有对应的双亲节点值就不能插入孩子节点值；②进行删除操作时，如果删除了双亲节点值，则相应的孩子节点值也被同时删除；③进行更新操作时，应更新所有相应记录，以保证数据的一致性。

13. 网状模型

网状模型去除了层次模型中的两个限制，一是允许模型中存在多个没有双亲的节点；二是允许模型中存在有多个双亲的节点。

网状模型的典型代表 DBTG 模型中，实体以记录型来描述，记录型由数据项、向量、组项、重复组等数据单位组成。实体间的一对多联系以联系型描述，其中的双亲节点称为系主，孩子节点称为成员。多对多联系不能直接描述，而是需要引入连接记录型，把一个多对多联系转换为两个一对多联系。

网状模型的数据操作主要包括查询、插入、删除和更新。其中，插入操作允许插入尚未确定双亲节点值的孩子节点值；删除操作允许只删除双亲节点值；更新操作可对指定记录进行；查询操作有多种方式，可根据实际需要进行选择。

网状模型完整性约束条件主要包括系型约束和系值约束。系型约束指一个记录型不能同时作为同一个联系型的系主和成员，确实需要时可引入连接类型。系值约束指一个记录值不能出现在同一联系型的不同系值中。

14. 关系模型

关系模型以关系来表示实体及实体间的联系。一个关系相当于一张二维表，表中的一行称为一个元组，表中的一列称为一个属性。关系模型要求关系必须是规范化的，最基本的条件是关系的每一个属性都不能再进行有意义的分割，即不允许表中还有表。

关系模型所支持的数据操作主要包括选择、投影、连接、除、并、交、差等查询操作和对数据的插入、删除、更新等操作，这些操作的操作对象和操作结果都是关系。

关系模型所支持的完整性约束条件包括实体完整性、参照完整性和用户定义的完整性三类。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，由关系数据库管理系统自动支持；用户定义的完整性则由数据库设计人员根据实际应用的需要而拟定，数据库管理系统提供手段支持这类完整性的定义，并在运行中对数据库的状态进行监督，对不符合完整性约束要求的操作予以拒绝。

15. 数据库管理系统的主要功能

数据库管理系统运行于操作系统之上，负责对数据库进行统一的管理和控制，以保证数据库的安全性和完整性，用户对数据库的任何操作都要通过数据库管理系统来完成。数据库管理系统的主要功能如下。

(1) 数据库的定义和建立。数据库管理系统提供数据定义语言以支持数据库管理员对数据库系统的外模式、概念模式和内模式以及这三级模式间的两级映像进行定义，把各源模式翻译成目标模式后存入数据库字典以供系统查阅。数据库管理系统还支持数据库安全性要求、完整性规则的定义，并能根据已定义的概念模式和内模式，把初始数据装入数据库中，完成数据库的建立工作。

(2) 数据库的操作。数据库管理系统提供数据操作语言以支持数据库管理员、用户、应用程序等表达需要对数据库执行的操作，包括查询、插入、删除和修改等。

(3) 数据库的控制。这是数据库管理系统的核心功能，包括控制整个数据库的运行；控制用户对数据库的并发操作，以保证数据库的一致性；执行对数据库的安全性检查，以防止库中数据被非法访问、修改，甚至恶意破坏；执行对数据的完整性控制，以保证库中数据是正确的、相容的、符合完整性约束条件的。

(4) 数据库的维护与故障恢复。主要包括自动维护数据库的索引、数据字典等；在数据库运行时记录运行日志，监视和分析数据库的性能；当数据库系统发生故障时，提供有效的措施和工具，尽量减少故障造成的损失。

(5) 数据通信。数据库管理系统通常还会提供与其他数据库管理系统或文件系统的接口，从而能将数据库里的数据转换为另一种数据库管理系统或某应用程序能够处理的格式，也能由其他数据库管理系统中的数据库或某种格式的数据文件生成本系统的数据库。许多数据库管理系统还提供了许多基于图形界面（GUI）的用户接口软件，如查询管理器、报表生成器、统计图形生成器等。

16. 数据库管理系统的组成

数据库管理系统是由许多部件构成的大型系统软件，一个数据库管理系统通常由以下几个部分组成。

(1) 语言及其编译处理程序，包括数据定义语言及其编译处理程序、数据操作语言及其编译处理程序。

(2) 系统运行控制程序，负责在系统运行过程中实现对数据库的控制与管理，主要包括系统总控程序、安全性控制程序、完整性检查程序、并发控制程序、数据存取和更新程序，以及通信控制程序。

(3) 系统建立、维护程序，主要包括数据初始装入程序、数据库重组程序，以及数据库恢复程序。

17. 应用程序访问数据库的过程

应用程序访问数据库的一般过程为：①应用程序向数据库管理系统发出从数据库中读数据记录的命令；②数据库管理系统对该命令进行语法检查、语义检查，并调用该应用程序对应的外模式，检查其存取权限，决定是否执行该命令，如果拒绝执行，则向用户返回出错信息；③在决定执行该命令后，数据库管理系统依据外模式—逻辑模式映像的定义，确定应读入逻辑模式中的哪些记录；④数据库管理系统依据逻辑模式—内模式映像的定义，决定应从哪个文件、用什么存取方式、读入哪个或哪些物理记录；⑤数据库管理系统向操作系统发出执行读取所需物理记录的命令；⑥操作系统执行读数据的有关操作；⑦操作系统将数据从数据库的存储区送至系统缓冲区；⑧数据库管理系统依据逻辑模式—内模式映像以及外模式—逻辑模式映像的定义，由物理记录导出应用程序所要读取的记录格式；⑨数据库管理系统将数据记录从系统缓冲区传送到应用程序的用户工作区；⑩数据库管理系统向应用程序返回命令执行情况的状态信息。

18. 分布式数据库系统的特点和主要优缺点

分布式数据库系统是数据库技术与计算机网络技术相结合的产物。

分布式数据库系统的三大基本特点是物理分布性、逻辑整体性和场地自治性。物理分布性指分布式数据库中的数据分散存放在以网络相连的多个节点上，每个节点中所存储数据的集合即为该节点上的局部数据库。逻辑整体性指系统中分散存储的数据在逻辑上是一个整体，各节点上的局部数据库组成一个统一的全局数据库，能支持全局应用。场地自治性指系统中的各个节点上都有自己的数据库管理系统，能对局部数据库进行管理，响应用户对局部数据库的访问请求。

分布式数据库系统的主要优点：体系结构灵活，适应分布式的管理和控制机构；可扩展性

好,容易实现对现有系统的集成,经济性能优越;既支持全局应用,也支持局部应用,局部应用的响应速度快;系统可靠性高,可用性好。

分布式数据库系统的主要缺点是存取结构复杂,通信开销较大,数据安全性较差。

19. 同构分布式数据库系统和异构分布式数据库系统

根据分布式数据库系统建立的原则,可以将分布式数据库系统分为同构分布式数据库系统和异构分布式数据库系统。

同构分布式数据库类似于一个集中式数据库,只不过同构分布式数据库将数据存放在分布于网络中的多个不同节点内,而集中式数据库则将数据存放在一个节点上,该节点不必在网络中。同构分布式数据库系统还可按节点是否自治分为自治型同构分布式数据库系统和非自治型同构分布式数据库系统。

异构分布式数据库系统的特点是在网络中的各个不同节点上运行着不同的数据库管理系统。异构系统又可以分为两个子类:一个是完全在本系统中进行集成,另一个是还要通过网关与其他系统实现连接。

20. 多媒体数据库系统的基本功能

多媒体数据库系统是数据库技术与多媒体技术相结合的产物,其基本功能应包括如下几项。

(1) 对各种媒体数据进行表示和处理。

(2) 对各种媒体数据间的时空关联进行表示和管理。

(3) 在保障数据的物理独立性和逻辑独立性外,还应支持数据的媒体独立性,即用户或应用程序的操作应不受数据所依赖的媒体的影响和制约,即使所使用的媒体改变了,应用程序也可以不变。

(4) 支持对多媒体数据的特殊操作,包括:对非格式化数据的整体、部分查询;对非格式化数据的模糊查询;通过分析建立非格式化数据的索引;支持举例查询;支持主题描述查询;支持对数据库信息的目录结构的浏览;能根据给定的应用约束和触发条件解决大容量数据访问问题和数据库一致性问题,提供演绎和推理功能,提供过程或函数;支持各类媒体所要求的专门操作,如图形图像视频数据的覆盖、邻接、镶嵌、交接、比例、剪裁、颜色转换、定位等,音频数据的合成、声音信号调度、声调调节、音强调节等。

(5) 支持对多媒体数据的事务管理和版本管理。

21. 多媒体数据库管理系统的组织结构

多媒体数据库管理系统的组织结构有集中型、主从型和协作型三种。集中型由一个多媒体数据库管理系统来负责各类媒体数据库的创建和管理。主从型由一个主多媒体数据库管理系统管理若干从多媒体数据库管理系统,而每个从多媒体数据库管理系统管理一类媒体数据库。协作型由多个成员多媒体数据库管理系统对各类媒体数据库进行管理,这些成员多媒体数据库管理系统间没有主次之分,它们借助外部处理软件模块实现协作。

22. 主动数据库系统

主动数据库系统是数据库技术与人工智能技术相结合的产物,其特点是主动性、快速性和智能性。

传统数据库系统只能被动地响应用户的操作请求,而实际应用中可能希望数据库系统在特定条件下能根据数据库的当前状态,主动地做出一些反应,如执行某些操作,或显示相关信息等。为达到这一目的,通常采用的方法是在传统数据库系统的基础上添加事件检测器以及一个

事件驱动的规则库，该规则库中包含一组“事件—条件—动作”规则，当事件检测器监测到事件发生时，就把事件信息传递给规则管理器，规则管理器根据规则库中预定义的“事件—条件—动作”规则，确定由该事件触发的规则，再由条件评价器对这些规则中指定的条件进行评估，所有条件评估成功的规则形成规则冲突集，交由执行模块执行其中的动作，在动作执行过程中也可能导致新的触发事件形成。

23. 面向对象数据库系统

面向对象数据库系统是数据库技术与面向对象技术相结合的产物，兼顾了对象表达方面的长处与关系数据库系统事务处理方面的优越性，其核心是面向对象的数据模型。

在面向对象数据模型中，现实世界里客观存在且相互区别的事物被抽象为对象，一个对象由三个部分构成，即变量集、消息集和方法集。变量集中的变量是对事物特性的数据抽象，消息集中的消息是对象所能接收并响应的操作请求，方法集中的方法是操作请求的实现方法，每个方法就是一个程序段。

面向对象数据模型的主要优点包括：①消息集是对象与外界的唯一接口，方法和变量的改变不会影响对象与外界的交互，从而使应用系统的开发和维护变得容易；②相似对象的集合成类，而类具有继承性，从而使程序复用成为可能；③支持复合对象，即允许在一个对象中包含另一个对象，从而使数据间诸如嵌套、层次等复杂关系的描述变得更容易。

面向对象数据库系统一般有两种实现途径，一是在面向对象程序设计语言的基础上增加数据库功能；二是对传统的关系数据库系统进行改造，在其中加入面向对象特性。

24. 数据仓库

数据仓库是面向主题的、集成的、随时间变化的、非易失的数据的集合，用于支持管理层的决策过程。

“面向主题”是组织数据仓库中数据的基本原则。一个主题就是与决策者的某一决策分析对象相关的数据的集合，利用这些数据，决策者可对该决策分析对象进行正确地分析，从而做出科学合理的决策。这些数据中既包括用以描述决策分析对象的分析数据，也包括用以描述分析数据间关系的数据。

“集成”是指来自不同数据源的数据在进入数据仓库之前，需按决策分析的要求，依据统一的标准进行抽取、选择、整理、综合等预处理。

“随时间变化”是指数据仓库中的数据应当适时进行刷新。数据的刷新包括跟踪数据源中数据的变化情况，每隔一定时间，将相关信息追加到数据仓库中；搜索对决策分析有用的新数据源，并将其中相关数据添加到数据仓库中；将使用率极低的数据从数据仓库中移出；将超过保存期的数据从数据仓库中删除；与时间有关的概括数据需定时重新概括。

“非易失”是指对数据仓库中的数据一般只做加载和查询操作，而不进行插入、删除、修改等操作。

“数据集合”是指数据仓库必然以某种数据集合的形式存储，现在一般采用多维数据库或关系数据库或两者结合的方式存储。

“支持决策”是指数据仓库的使用目标是为决策者对数据的分析提供支持。

25. 数据挖掘

数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的实际数据中，提取隐含在其中的、人们不知道的、但又是潜在有用的信息和知识的过程，是基于数据库的知识发现过程中

最关键的步骤。

常用数据挖掘技术包括线性分析、非线性分析、回归分析、逻辑回归分析、单变量分析、多变量分析、时间序列分析、最近邻算法、聚类分析等统计分析类技术,还有人工神经网络、决策树、遗传算法、粗糙集、关联规则等知识发现类技术,以及文本数据挖掘、Web 挖掘、分类系统、可视化系统、空间数据挖掘和分布式数据挖掘等其他类技术。

第二部分 例题分析

例 1-1 文件系统中的文件与数据库系统中的文件有何本质上的不同?

解析:

文件系统中的文件与数据库系统中的文件是有着本质上的不同的,主要表现在以下几个方面。

(1) 文件系统中的文件和数据库系统中的文件虽然都是以记录的形式对数据进行组织,但数据库系统中的文件已能描述数据集间的一对一、一对多、多对多联系,而文件系统中的文件仅仅是数据的集合,还不能体现数据间的复杂联系。

(2) 文件系统中的文件通常是面向单个应用进行数据组织的,也因而可以有针对性地向该应用进行优化,但如果要在别的应用中直接使用该数据文件,则会带来不便,而如果为另一应用再建一个数据文件,却又可能带来数据的冗余、不一致等问题。数据库文件不是面向特定应用建立的,而是面向数据本身及数据间的内在联系来进行数据组织的,有利于多用户、多应用共享数据,降低了数据的冗余度,也提高了数据的一致性和可信度。

(3) 对文件系统中的文件进行访问时,最小的存取单位为记录,而在数据库文件中,则可对组成记录的数据项进行访问,数据的存取粒度更细,有利于对并发访问的支持。

例 1-2 什么是数据独立性?数据库系统是如何实现数据独立性的?

解析:

数据独立性包括数据的逻辑独立性和数据的物理独立性。数据的逻辑独立性是指在数据库系统中,数据库的逻辑模式与基于该数据库创建的应用程序相互独立。数据的物理独立性是指在数据库系统中,数据库的内模式与基于该数据库创建的应用程序相互独立。

数据库系统的外模式、逻辑模式、内模式这三级模式结构以及介于其间的外模式—逻辑模式、逻辑模式—内模式这两级映像实现了数据的逻辑独立性和物理独立性。

逻辑模式是对数据库中全体数据的逻辑结构、完整性约束条件、安全性要求等内容的描述,一个数据库只有一个逻辑模式;外模式是对数据库中面向具体应用的部分数据的逻辑结构、完整性约束条件、安全性要求等内容的描述,一个数据库可以有多个外模式;内模式是对数据库中全体数据的物理结构、存储方式、存储策略等的描述,一个数据库只有一个内模式。当逻辑模式发生变化时,通过修改外模式—逻辑模式映像可使外模式保持不变,因而基于外模式设计的应用程序也不必修改,从而保证了数据的逻辑独立性。当内模式发生变化时,通过修改逻辑模式—内模式映像可使逻辑模式保持不变,因而外模式以及基于外模式设计的应用程序也就可以保持不变,从而保证了数据的物理独立性。

例 1-3 层次模型、网状模型和关系模型等三种基本数据模型是根据什么来划分的?

解析:

层次模型、网状模型和关系模型是三种常见的逻辑模型。逻辑模型的三要素是数据结构、

数据操作和完整性约束条件,其中数据结构是最重要也是最根本的一个,常以之为逻辑模型命名。层次模型、网状模型和关系模型就是根据它们所采用的不同的数据结构来划分的。

例 1-4 数据库系统由哪几部分组成?

解析:

数据库系统是引入了数据库技术的计算机系统,除了一般计算机系统的组成部分外,还包含了数据库、数据库管理系统、数据库应用系统以及与系统开发、使用相关的各类人员。

例 1-5 什么是数据库的子模式?给用户使用子模式有何好处?

解析:

数据库的子模式是对数据库中面向某数据库应用系统的部分数据的逻辑结构、完整性约束条件、安全性要求等内容的描述。一个数据库可以针对不同的用户设置多个子模式。

给用户使用子模式的好处主要包括①子模式通常是逻辑模式的子集,逻辑模式中的同一个数据对象,在不同的子模式中所映射成的数据对象的类型、长度、保密级别等都可以不同,从而可满足多个应用的不同需求;②数据库管理系统保证了不同的应用系统只能操作它所对应的子模式中的数据,从而对用户的数据库访问权限进行了控制,是保障数据库安全性的一个有力措施;③子模式作为数据库的三级模式、两级映射结构的组成部分,支持了数据的逻辑独立性和物理独立性。

例 1-6 简述关系模型的主要特点。

解析:

关系模型以关系来表示实体及实体间的联系。一个关系相当于一张二维表,表中的一行称为一个元组,表中的一列称为一个属性。关系模型要求关系必须是规范化的,最基本的条件是关系的每一个属性都不能再进行有意义的分割,即不允许表中还有表。

关系模型所支持的数据操作主要包括选择、投影、连接、除、并、交、差等查询操作和对数据的插入、删除、更新等操作,这些操作的操作对象和操作结果都是关系。

关系模型所支持的完整性约束条件包括实体完整性、参照完整性和用户定义的完整性三类,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,由关系数据库管理系统自动支持,用户定义的完整性则由数据库设计人员根据实际应用中的需要而拟定,数据库管理系统提供手段支持这类完整性的定义,并在运行中对数据库的状态进行监督,对不符合完整性约束要求的操作予以拒绝。

例 1-7 用传统数据库系统存储和管理多媒体数据存在什么主要问题?

解析:

多媒体数据不仅包含数字、字符等格式化数据,还包括文本、图形、图像、声音、视频等非格式化数据,其中非格式化数据的数据量一般都比较大小,结构也比较复杂,有些数据还带有时间顺序、空间位置等属性,这就给数据的存储和管理带来了较大的困难。

对多媒体数据的查询要求往往也各不相同,系统不仅应当能支持一般的精确查询,还应当能支持模糊查询、相似查询、部分查询等非精确查询。

各种不同媒体的数据结构、存取方法、操作要求、基本功能、实现方法等一般也各不相同,系统应能对各种媒体数据进行协调,正确识别各种媒体数据之间在时间、空间上的关联,同时还应提供特种事务处理和版本管理能力。

例 1-8 什么是主动数据库?它与传统数据库有何主要不同?

解析:

主动数据库系统是数据库技术与人工智能技术相结合的产物,其特点是主动性、快速性和智能性。

传统数据库系统只能被动地响应用户的操作请求,而实际应用中可能希望数据库系统在特定条件下能根据数据库的当前状态,主动地做出一些反应,如执行某些操作,或显示相关信息等。

例 1-9 试举例说明实体集间的一对多联系。

解析:

以图书馆信息系统中所涉及的实体集“读者”和实体集“图书”间的“借阅”联系为例。

对于实体集“图书”中的每个实体,即该图书馆中的每本馆藏图书,实体集“读者”中只有 0 或 1 个实体与之关联,因为一本馆藏图书当前或正在馆,即当前未被读者借阅,或当前正被某一位读者借阅,但不可能同时被多位读者借阅。

对于实体集“读者”中的每位读者,实体集“图书”中可以有 0 或 1 或多个实体与之关联,因为一位读者当前可能并没有借书,也可能只借了一本书,也可能借了多于一本书。

由上述分析可见,实体集“读者”和实体集“图书”间的“借阅”联系为一对多联系。

第三部分 习 题

◎ 基本题

一、选择题

1. 描述事物的符号记录称为_____。

- A. 信息 B. 数据 C. 记录 D. 记录集合

答: B

2. 数据是信息的载体,信息是数据的_____。

- A. 符号化表示 B. 载体 C. 内涵 D. 抽象

答: C

3. _____是长期存储在计算机内的有组织、可共享的数据集合。

- A. 数据库管理系统 B. 数据库系统 C. 数据库 D. 文件组织

答: C

4. 数据的收集、整理、组织、存储、查询、维护和传送等操作,统称为_____。

- A. 数据操作 B. 数据处理 C. 数据维护 D. 数据管理

答: B

5. 在人工管理阶段,数据是_____。

- A. 有结构的 B. 无结构的
C. 整体无结构,记录有结构 D. 整体结构化的

答: B

6. 在数据库系统阶段,数据是_____。