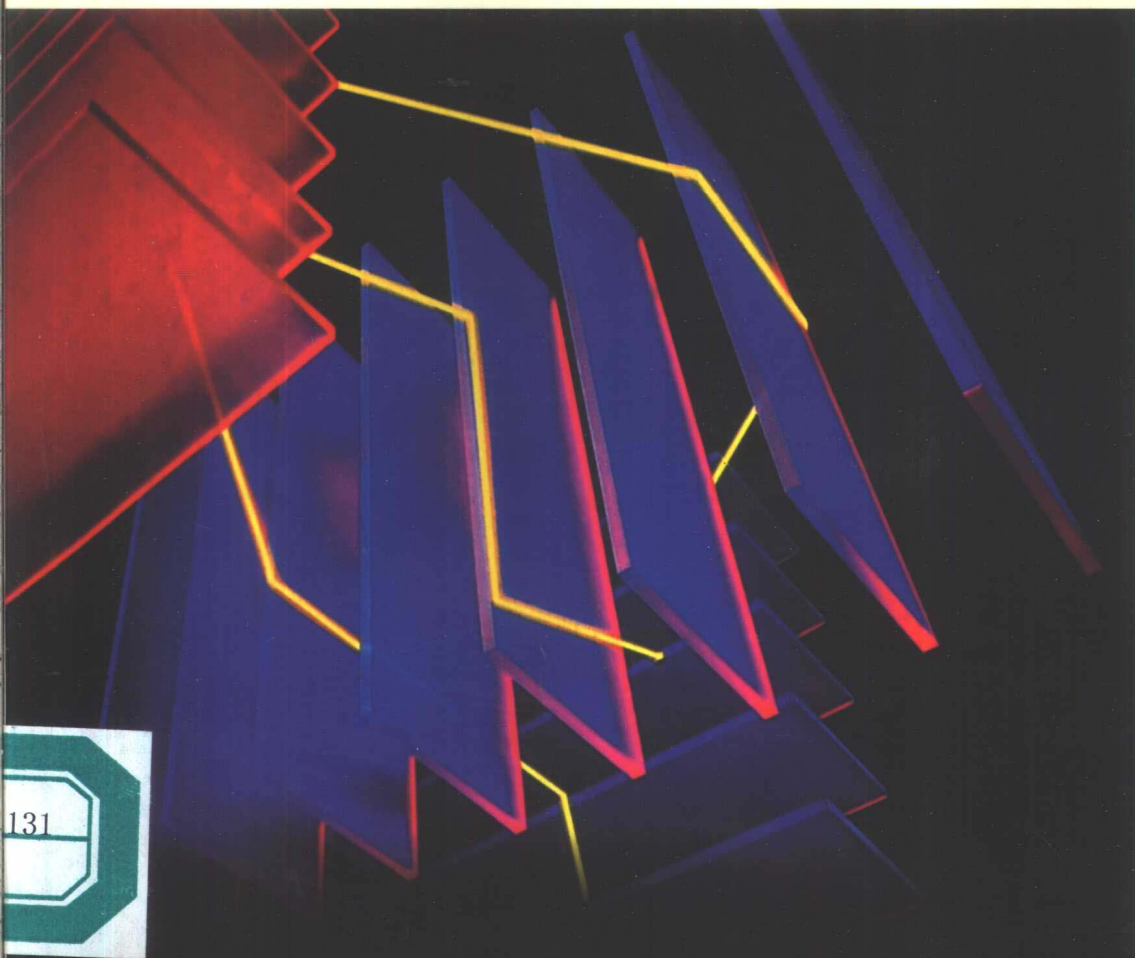


信息管理与信息系统专业系列教材

数据库原理与设计

金银秋 主编

金银秋 肖慎勇 黄杰 编著



科学出版社

信息管理与信息系统专业系列教材

数据库原理与设计

金银秋 主编

金银秋 肖慎勇 黄杰 编著

科学出版社

2000

内 容 简 介

本书汇集了数据库学科近年来一些最新的研究成果,如面向对象模型、SQL 递归查询和数据库逻辑描述语言。同时全面地、系统地介绍了数据库的基础理论和基本原理。本书配备了大量习题和例题,以便读者能更好地理解基本理论和系统地进行自学。

本书可作为信息管理与信息系统专业和计算机专业本科学生的教材或教学参考书。

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

北京双青印刷厂印刷

科学出版社发行 各地新华书店经销

*

2000年10月第一版 开本:787×1092 1/16

2000年10月第一次印刷 印张:14 3/4

印数:1—4 000

字数:326 000

ISBN 7-03-006943-9/TP·999

定价:25.00元

(如有印装质量问题,我社负责调换<环伟>)

《信息管理与信息系统专业系列教材》

编 委 会 名 单

主 任

邱家武

副 主 任

刘康泽 胡乾顺

委 员

(按姓氏笔划排序)

冯发石 刘康泽 刘腾红 杨开汉

杨怡光 邱家武 余尚智 周 岩

金银秋 胡乾顺 贾启禹 贾希辉

钱 渝 彭勇行 童涌泉

总 序

中南财经大学是财政部直属的一所以经济学科、管理学科为主，兼有法学、文学、哲学、理学等6个一级学科的，具有50年历史的高等学校。中南财经大学经济信息管理系始建于1978年，1980年开始招收本科生，是继中国人民大学之后在全国高校第二个建立信息管理专业的系，并于1990年，经国务院学位委员会批准建立信息经济硕士点，是全国首批设立的该专业4个硕士点之一。

改革开放20年，正是信息管理与信息系统专业不断建设成长的20年。中南财经大学信息系经过不断的探索和建设，在教学研究、师资队伍建设、教材建设、实验室建设及教学管理等方面均打下了良好的基础。

在专业发展和教材建设中，我们遵循教育必须为社会主义建设服务和必须面向现代化、面向世界、面向未来的要求，20年来，无论是专业目录调整前的管理信息系统专业，还是专业目录调整后的信息管理与信息系统专业，我们都努力在专业建设的深度以及市场经济建设的应用力度上下功夫，力求学生所学的专业知识在实际工作中能派上用场，在教学体系建设及教材建设中力求体现本专业的特色。经过20年艰苦奋斗与教学科研实践，中南财经大学信息管理与信息系统专业已经建立起规模适当，多层次多形式的办学体系；初步形成多学科有机结合，互相渗透的专业特色；建立了结构合理的教师队伍；具备了比较完善的办学条件；取得了一批先进水平的科研成果，为国家培养了大批受社会欢迎的信息管理专门人才。

为了建设一套有信息管理与信息系统专业特色的教材，我们长期以来在加强基础、拓宽知识面、增强适应性、建立主动适应社会主义建设需要和适应现代科学技术、文化发展趋势的教学内容以及课程结构等方面搜集了大量的素材和案例，特别是在理论联系实际，面向经济建设主战场，强化学生的动手能力，结合最新的科技发展以及在教材中融进各位教师的研究成果上花了不少的精力。1998年我们按照教育部公布调整后的新专业目录，组织了两个小组到兄弟学校调查研究，进行了多次座谈和研讨，进一步明确了信息管理与信息系统专业的性质是以系统的方法、现代信息处理技术来研究人类管理活动规律及其应用的学科。它融合了管理学、经济学、计算机科学与技术等学科的知识，以系统观点为指导，运用定性与定量结合的方法及相关学科的研究手段，深入研究并有效地解决社会中各类信息管理问题。本专业的目标是：培养具备现代管理学理论基础、计算机科学技术知识及应用能力，掌握系统思想和信息系统分析与设计方法以及信息管理等方面的知识与能力，能在国家各级管理部门、工商企业、金融机构、科研单位等部门从事信息管理以及信息系统分析、设计、实施管理和评价等方面的高级专门人才。本专业的培养要求是：学生主要学习经济、管理、数量分析方法、信息资源管理、计算机及信息系统方面的基本理论和基本知识，接受系统和设计方法以及信息管理方法的基本训练，具备综合运用所学知识去分析和解决问题的能力。本专业的毕业生应具备以下的知识和能力：（1）掌握信息管理和信息系统的基本理论、基本知识；（2）掌

握管理信息系统的分析方法、设计方法和实现技术；(3) 具有信息组织、分析研究、传播与开发利用的基本能力；(4) 具有综合运用所学知识分析和解决问题的能力；(5) 了解本专业相关领域的发展动态；(6) 掌握文献检索、资料查询、收集的基本方法，具有一定的科研和实际工作能力。

基于上述思想，我们修订了信息管理与信息系统专业教学计划，相应地修订了相关课程的教学大纲，组织人员编写出有信息管理与信息系统专业特色的教材，供教学之需。经反复讨论，确定出版 18 种图书作为信息管理与信息系统专业系列教材，即：该套丛书包括以下 18 种教材：

《计算机实用技术基础》

《离散数学》

《数据结构》

《数据库原理与设计》

《计算机网络》

《计算机操作系统》

《管理信息系统分析与设计》

《计算机组成原理》

《多媒体与信息管理》

《管理决策分析》

《信息管理学》

《应用数理统计》

《运筹学（一）》

《运筹学（二）》

《经济预测方法》

《高等数学》

《线性代数》

《语言程序设计》

本套教材得以顺利出版，得到了科学出版社的大力支持，我代表本套教材的各位编写人员向科学出版社表示由衷的感谢！

由于水平所限，在陆续出版的系列教材中错误难免。望读者不吝赐教，以资改进，在此一并致谢！

邱家武

1999 年元旦于中南财经大学

前 言

数据库理论和技术在近 10 多年得到了广泛的应用。利用成熟的技术建造大型数据库已成为趋势，数据库技术已成为管理信息系统中最成熟的技术，也是管理信息系统能否成功的关键。因此，数据库理论和技术已成为计算机专业和信息管理与信息系统专业的必修课。

本书是在多年数据库课程讲授经验基础上写成的。书中收集了数据库基础理论方面的许多最新成就，如面向对象数据模型、数据逻辑描述语言 Datalog 和 SQL 递归查询语言。本书配有大量例题和练习题，以加深读者对基本原理、基本技巧的理解，并可以熟练地运用。

编著本书的作者中，黄杰编写了第 2 章的内容，肖慎勇编写了第三章和第九章的内容，金银秋编写了第一、二、四、五、六、七章的内容。中南财经大学信息系主任邱家武教授负责全书的主审工作，金银秋副教授担任本书主编。

在编著本书的过程中得到校、系领导和同事的大力支持、帮助，在此深表谢意。

由于时间仓促，加上水平有限，难免存在许多错误和不足，恳请读者批评指正。

最后希望广大读者、计算机界同仁和我们共同努力，为把我国的数据库研究、教学和产品开发提高到新的水平做出应有的贡献。

编著者
2000 年元月

目 录

总序

前言

第一章 绪论	(1)
第一节 数据库发展历程	(1)
一、手工数据处理	(1)
二、文件方式的数据处理	(1)
三、数据库方式的数据处理	(1)
第二节 数据库系统的构成	(2)
一、数据库	(2)
二、数据库管理系统	(2)
三、数据库应用系统	(4)
四、数据库管理员	(4)
第三节 数据库系统中的数据存取过程	(5)
一、数据库的三层模式	(5)
二、数据独立性	(6)
三、查询处理的一般流程	(6)
第四节 数据库分类	(7)
一、传统数据库	(7)
二、多媒体数据库	(7)
三、集中式数据库	(8)
四、分布式数据库	(8)
习题一	(9)
第二章 数据模型	(10)
第一节 E-R 模型	(10)
一、E-R 模型的基本概念	(10)
二、E-R 模型的构成	(10)
三、E-R 模型图	(12)
四、E-R 模型举例	(13)
第二节 面向对象模型及 ODL 语言	(13)
一、面向对象模型的基本思想	(13)
二、面向对象模型概念及结构	(14)
三、联系及逆联系	(15)
四、子类和继承性	(17)
五、ODL 模型举例	(17)

第三节 关系数据模型	(18)
一、关系数据模型的概念	(18)
二、关系数据库	(19)
三、关系模式	(20)
四、关系数据模型的操作	(20)
第四节 E-R 模型向关系数据模型和面向对象模型转化	(21)
一、实体、属性向关系模式、属性转化	(21)
二、实体集转化为相应的关系	(21)
三、联系图向关系模式的转化	(22)
四、联系值向关系的转化	(22)
五、E-R 模型向面向对象模型转化	(22)
第五节 面向对象模型向关系模型的转化	(24)
一、类属性的转化	(24)
二、类联系的转化	(24)
三、子类的转化	(25)
第六节 早期的数据模型	(26)
一、层次数据模型	(26)
二、物理数据库	(27)
三、层次序列	(27)
四、层次数据模型的存储结构	(28)
五、树型 E-R 模型象层次数据模型的转化	(29)
六、网状数据模型	(29)
七、系	(30)
八、系的存储结构	(32)
习题二	(33)
第三章 关系运算	(35)
第一节 关系代数	(35)
一、关系	(35)
二、传统集合运算	(38)
三、投影	(39)
四、选择	(40)
五、笛卡尔积	(40)
六、 θ 连接	(41)
七、自然连接	(42)
八、除法	(42)
九、复合运算与运算树	(43)
十、改名	(44)
十一、应用举例	(44)
第二节 元组谓词演算	(44)

一、谓词演算的基本概念	(44)
二、元组变量和元组分量	(45)
三、合式公式	(45)
四、关系代数到谓词演算的转换	(46)
五、安全表达式与安全限制	(47)
六、ALPHA 的查询功能	(47)
第三节 域关系谓词演算	(49)
第四节 数据库逻辑查询语言 Datalog	(50)
一、谓词和原子	(50)
二、算术谓词	(51)
三、Datalog 查询规则	(51)
四、关系代数向 Datalog 的转换	(52)
五、Datalog 的递归查询	(56)
习题三	(59)
第四章 SQL 语言	(62)
第一节 SQL 简介	(62)
一、SQL 语言产生的历史	(62)
二、SQL 功能	(62)
三、SQL 使用方式	(63)
第二节 关系模式和索引的建立	(64)
一、属性的基本类型	(64)
二、建立模式	(65)
三、删除表	(66)
四、修改表结构	(66)
五、建立与删除索引	(67)
第三节 SQL 简单查询	(68)
一、简单查询的基本格式	(68)
二、简单查询举例	(68)
第四节 多关系连接查询	(70)
一、多关系连接查询的一般格式	(71)
二、连接查询举例	(71)
第五节 嵌套查询	(73)
一、子查询的基本概念	(73)
二、带 IN, NOT IN 的子查询	(73)
三、带 ALL, ANY 的子查询	(74)
四、带 EXISTS, NOT EXISTS 的子查询	(75)
五、全称量词向 NOT EXISTS 的转化	(76)
第六节 分组统计查询	(77)
一、SQL 统计函数	(77)

二、分组统计的基本概念	(79)
三、分组统计语句	(80)
四、分组统计过程	(81)
五、输出控制条件	(82)
第七节 数据更新	(83)
一、更新的定义	(83)
二、修改	(83)
三、插入	(83)
四、删除	(84)
第八节 SQL 中的集合运算	(85)
一、查询的差集	(85)
二、查询的交集	(86)
三、查询的并集	(87)
第九节 视图	(88)
一、视图的基本概念	(88)
二、单关系生成的视图	(89)
三、由多关系生成的视图	(89)
四、视图查询	(89)
五、视图的更新	(90)
六、视图清除	(91)
第十节 递归查询	(91)
一、递归查询	(91)
二、递归查询定义语句	(93)
三、递归查询的执行过程	(94)
四、间接递归查询	(94)
习题四	(95)
第五章 关系数据库理论	(98)
第一节 关系模式	(98)
一、关系	(98)
二、关系中允许的基本类型	(99)
三、关系模式	(99)
四、元组及分量的表示方法	(101)
第二节 函数依赖	(101)
一、函数依赖的定义	(101)
二、关于函数依赖的说明	(102)
三、完全函数依赖	(102)
四、部分函数依赖	(102)
五、传递函数依赖	(103)
六、平凡函数依赖	(103)

七、码、超码、主码	(103)
八、高级关系模式	(104)
第三节 函数依赖的公理系统	(105)
一、Armstrong 公理	(105)
二、公理的有效性	(107)
三、由 Armstrong 公理导出的三条推理规则	(107)
第四节 多值依赖	(108)
一、问题的提出	(108)
二、多值依赖的定义	(109)
三、多值依赖的基本性质	(110)
四、多值依赖的公理系统	(110)
习题五	(112)
第六章 关系模式分解理论	(113)
第一节 关系范式	(113)
一、关系规范化	(113)
二、第一范式	(114)
三、第二范式	(115)
四、第三范式	(115)
五、BCNF 范式	(117)
六、第四范式	(118)
七、范式之间的关系	(119)
第二节 传递闭包	(119)
一、闭包的基本概念	(119)
二、求 X_F^+ 的算法	(120)
三、最小函数依赖集	(121)
四、Armstrong 公理是完备的	(124)
第三节 模式分解的基本概念	(125)
一、函数依赖集的投影	(125)
二、模式分解	(125)
三、保持函数依赖的分解	(126)
四、无损连接分解	(127)
五、关于分解的一些性质	(128)
六、无损连接分解的判别算法	(129)
第四节 第三范式分解算法及其证明	(134)
一、第三范式保持函数依赖的分解算法	(134)
二、分解算法正确性证明	(135)
三、保持函数依赖和无损连接的 3NF 分解算法	(136)
四、保持函数依赖且无损连接第三范式分解算法的正确性证明	(137)
第五节 BCNF 范式分解算法及其证明	(138)

第六节 第四范式分解算法	(140)
一、多值依赖的二项分解定理	(141)
二、第四范式分解算法	(142)
习题六	(143)
第七章 SQL 程序设计	(146)
第一节 嵌入式 SQL 概述	(146)
一、SQL 语言的独立执行方式	(146)
二、SQL 语言的宿主执行方式	(147)
三、宿主 SQL 语句的嵌入方式	(148)
第二节 用 SQL * C 程序建立表	(149)
一、共享变量及其描述	(149)
二、利用 SQL * C 建立股票交易表	(150)
第三节 用 SQL * C 程序进行关系查询处理	(151)
一、游标	(151)
二、利用游标进行程序设计的基本步骤	(153)
第四节 用 SQL * C 进行关系更新操作	(155)
一、简单的数据更新	(156)
二、利用游标对查询结果逐行进行修改	(157)
习题七	(158)
第八章 数据保护	(159)
第一节 数据安全	(159)
一、安全性控制的一般方法	(159)
二、SQL 授权语句	(162)
三、ORACLE 数据库的安全性措施	(164)
第二节 数据完整性	(167)
一、数据完整性定义	(167)
二、实体完整性	(169)
三、参照完整性	(169)
四、域完整性	(172)
第三节 SQL 实现数据完整性的方法	(172)
一、实体完整性定义	(172)
二、参照完整性定义	(173)
三、域完整性定义	(174)
四、其他约束	(175)
第四节 触发器	(176)
一、触发器基本概念	(177)
二、触发器的定义方法	(178)
三、触发器在数据保护中的作用	(179)
第五节 并行事务的一致性	(181)

一、并行事务的基本特征·····	(181)
二、并行事务的控制·····	(182)
三、死锁的发生、预防和解除·····	(186)
四、可串行调度及二段锁协议·····	(188)
习题八·····	(191)
第九章 数据库设计基础 ·····	(192)
第一节 数据库设计概述·····	(192)
一、数据库设计概念·····	(192)
二、数据库设计步骤·····	(194)
三、需求分析·····	(195)
第二节 概念数据模型设计·····	(200)
一、概述·····	(200)
二、局部 E-R 图设计·····	(201)
三、局部 E-R 模型的集成·····	(205)
第三节 数据库逻辑设计及物理设计和实现·····	(208)
一、概念模型向关系模型的转换·····	(208)
二、关系规范化设计与优化·····	(209)
三、数据库实现与维护·····	(210)
第四节 数据安全设计·····	(212)
一、用户权限·····	(213)
二、口令·····	(214)
三、数据加密设计·····	(215)
第五节 数据恢复设计·····	(215)
一、数据恢复的概念·····	(216)
二、定期转储法及设计·····	(217)
三、日志法·····	(217)
习题九·····	(219)
参考文献 ·····	(220)

第一章 绪 论

信息技术在国民经济中的作用和地位已经越来越受到人们的重视。信息是以数字方式存储的，因此，合理地组织数据是进行信息处理的前提和关键步骤。沿着历史的足迹，我们可以看到：数据的管理方式从简单的数据管理方式发展到复杂的数据库方式，由单一载体的存储方式发展到多种载体的组织方式。这可以为今后的数据管理方式提供各种有益的启示。

第一节 数据库发展历程

一、手工数据处理

在计算机未发明之前，人们对数据的处理依赖于笔、纸和算盘。这种处理方式无法做到对大量数据进行及时处理。原因包括两个方面：一方面是指数据量大使算盘变得无用；另一方面是指用算盘进行处理，时间很长，无法做到对数据进行及时地处理。例如，在手工处理阶段，银行工作人员通常要花费大量时间来处理一笔取款业务。首先，银行人员要上千份帐单中进行查找，然后进行核对，在经过很长的等待时间之后，取款人才能拿到款项。等待时间直接与帐单数量的多少成正比。

二、文件方式的数据处理

计算机早期主要用于科学计算，需要保存的数据存放在文件中。由于文件之间是独立的，早期的计算不提供文件之间的运算，当大量相关的数据要分散存放于多个文件中时，人们不得不编写大量程序去处理多个文件的连接。当人们为了减少文件的数量时，又必须面对大量的数据冗余（这是因为要对文件进行合并），过多的数据冗余对信息系统的长期运行带来灾难性的后果。当我们用文件方式进行数据处理时，由于早期的计算机语言主要是用于文件处理，是与文件的组织方式相关的。这就是说当我们改变文件的组织方式的时候，比如，将非排序文件变成排序文件时，相应的程序就要做出相应的改变。

综上所述：文件方式处理数据有如下几点缺陷：

- ① 造成了程序的复杂性，使得程序设计必须过多地关心文件的细节；
- ② 出现大量数据冗余，使得修改和并发控制变得困难；
- ③ 使得程序和文件过于相关，即缺少程序和数据的独立性。

三、数据库方式的数据处理

用文件和高级语言处理大量数据的查询、修改、插入和删除等操作，显然很不协调。人们通过研究和探索，认识到应该用专门适用于数据处理的软件进行数据处理，因此，一门专门用于研究数据处理的学科——数据库，便应运而生了。数据库系统的目的

在于：为人们科学地组织数据提供方法、原理，并为人们提供对数据进行定义、操作、控制的工具。在用数据库方法对数据进行处理时，人们首先详细分析对象之间的关系，找出哪些数据是描述对象本身的，哪些数据是描述对象之间联系的。从而确定哪些应归为一类，并且指出这种归类的原理是什么，好处是什么，力图达到数据组织的最优化。当今社会生活越来越依赖于计算机，因此，对数据进行保护显得尤为重要。数据库系统在这方面比早期的文件处理方式显得更为优越。它提供了各种层次的数据安全保护，既能有效防止数据破坏，也能有效防止犯罪分子对数据库的恶意攻击。

假设，我们要建立学生成绩管理信息系统。若采用文件处理方式，可以只建一个文件，它包含了学生、课程、成绩等数据。这样做，学生选了几门课，该学生的数据就要重复几次，课程的数据也要重复几次，这就造成了大量的数据冗余。由于我们是用高级语言进行处理，因而在程序中必须指明文件的存储方式，因为不同的文件处理方式是不一样的。如果文件中数据的存储方式发生改变，数据的次序发生改变，用高级语言所编写的程序就必须作出相应的修改，这是因为缺少数据的独立性。若采用数据库方式进行处理，可以把该数据库设计成三个文件：学生文件、选课文件、课程文件。这样做，减少了不必要的数据冗余。在数据库系统管理下，数据操作与文件存储方式无关，这样就实现了程序(由数据操作语句所构成的程序)与数据存储方式之间的独立，由于数据库系统提供了数据演算语言，所以应用程序很容易实现对数据库文件的各种操作。

第二节 数据库系统的构成

数据库系统通常是指数据库和相应的软件系统。任何信息系统都至少包含这两个部分，只有这两部分完整，才能保证信息系统是完整的。

一、数据库

从数学的角度看，数据库可以定义为：数据库是所有数据文件的集合。但这一定义只强调了数据库中的内容。例如人们说的证券交易数据库是指人们在进行股票交易时的有关股票价格、数量、买入、卖出、股票代码、帐号、资金量等数据的集合。这个定义是一个大众化的定义。从学术的角度看，数据库可以定义为：数据库是关于数据组织的科学，通过这种组织，使得数据存储最优，数据最容易操作，最能有效地避免数据失真和实现数据安全保护。从这个定义上看，我们可以看到如下特点：数据库的设计是综合了各种应用领域的知识(如证券交易、企业财务处理、学生学籍管理)和计算机领域的知识(如数据库理论、网络技术、程序设计)的一门学科。同时，我们看到数据库是计算机人员提供给用户的最终产品。因而，产品应在各种环境下都能正常运行，并且具有完善的自我保护能力和数据恢复能力。数据库是一组文件的集合，每一个文件又是记录的集合，而记录是进行删除、插入操作的最小单位。

二、数据库管理系统

数据库管理系统(DBMS)是指：对外，它是向数据库的使用者提供数据服务的软件系统，使用户能够很方便地远离数据的具体细节去使用数据库；对内，它实现对数据

的存储管理,解决多用户同时使用数据库时可能产生的冲突,保证数据是正确的、一致的、完整的。数据库管理系统应具备如下功能:

1. 数据定义功能

该功能以语句形式提供,也可以通过填表方式提供。数据放入计算机中要占用一定的存储空间,因此,DBMS 应提供关于描述数据长度的语句或表项。除此以外,由于数据是按名存取的,因此,DBMS 还应提供对数据命名的语句。由于数据的运算特性不一样,数据被归为不同的运算类。而这是以运算特性来划分的,例如:能够进行算术运算的数据集合称为数值类型,能够进行字符连接运算的数据集合称为字符类型。定义了数据的类型就决定了数据的运算特性。

2. 数据操作功能

数据操作由以下四大功能构成:

(1) 数据查询

查询语句能够在数据库中找到满足一定条件的数据,这些数据构成查询结果集。查询是数据操作中使用最多,变化最多的一类操作。数据查询是从文件中读取数据,不改变数据以及数据的存储位置,因而不会破坏数据库。数据查询操作是通过 SQL 查询语句实现的,对 SQL 查询语句不能实现的查询,只有通过高级语言编程来实现。

(2) 数据更新

数据更新是指只改变当前数据库中已有的数据,不增加也不减少数据库中的数据。例如:将学生张三的学号由 970608 改为 970708。数据更新是通过更新语句实现的。它通常由两部分构成:查询条件和修改值。查询条件指明被修改的数据对象应满足的条件。

(3) 数据插入

数据插入是指不改变当前数据库中的任何数据,只增加一项或多项数据(或者一条或多条记录)。数据的插入可能引起数据有效位置的改变。新的数据只有通过插入操作才能进入数据库中。例如:在证券交易数据库中,若新增一名客户,则相应的客户文件中应增加有关该客户的数据。数据插入通过数据插入语句实现。

(4) 数据删除

数据删除是指从数据库中删掉现有的若干项数据(或者若干条记录)。使数据库中的数据减少。删除可分为逻辑删除和物理删除:逻辑删除是可恢复删除;物理删除是永久性删除。数据库操作人员进行数据删除时必须十分小心,以免删除不应删掉的数据。

3. 数据控制功能

数据控制是指事务的并发控制和数据安全控制。数据库中的数据可以同时为多个用户服务。例如:证券交易数据库需要同时为若干客户服务,客户对数据库的操作是多种多样的,这就不可避免地出现数据冲突。数据冲突是指数据操作出现异常。例如:修改不正确,读取数据不正确等。因此,DBMS 必须对数据共享操作进行控制,避免出现上