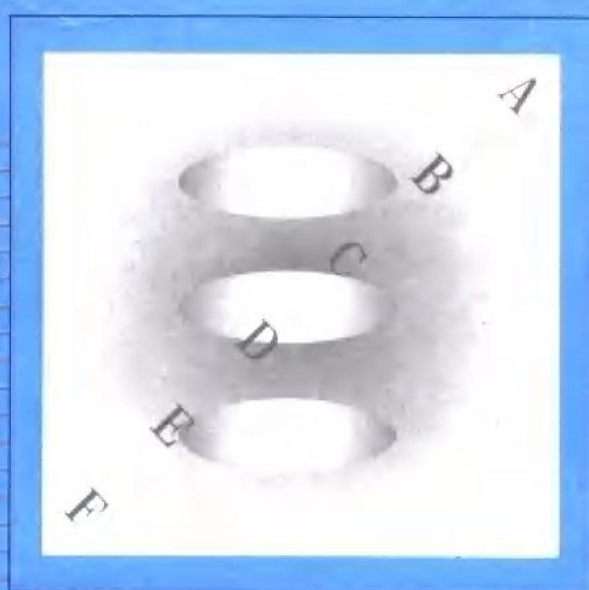


北京市高等教育学历文凭考试计算机专业教材

数据库原理与应用

李昭原 主编



2FC

科学出版社

内 容 简 介

本书是根据国务院发布的《高等教育自学考试暂行条例》，以及北京市高等教育自学考试委员会制定的《高等教育学历文凭考试课程大纲》编写的，其内容的深度和广度符合大纲要求。

本书以当前主流的关系数据库为主线，全面地介绍了数据库技术的基本内容。全书共九章，分别为：数据库系统概述；关系模型的基本概念；SQL语言；关系数据理论；数据库设计；数据库安全保护；FoxPro 数据库管理系统；FoxPro 应用程序设计；数据库技术的发展与展望。本书以 Oracle 为背景介绍了 SQL 语言，以 Visual FoxPro DBMS 为背景介绍了应用，理论联系实际，概念清晰，叙述简练易懂。各章后均备有习题。

本书由北京市高等教育自学考试委员会推荐使用，不仅可作为高等教育自学考试计算机专业文凭考试课程的理想教材，还可以作为各类高等专科学校、职工大学、职业大学、夜大学，以及函授大学等大专类“数据库原理与应用”课程的教材与教学参考书。

图书在版编目(CIP)数据

数据库原理与应用/李昭原主编. - 北京:科学出版社,1999.2

(北京市高等教育学历文凭考试计算机专业教材)

ISBN 7-03-006817-3

I. 数… II. 李… III. 数据库系统-理论-高等教育-教材

IV. TP311.13

中国版本图书馆 CIP 数据核字(98)第 16508 号

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

新华书店北京发行所发行 各地新华书店经售

*

1999年2月第一版 开本:787×1092 1/16

1999年2月第一次印刷 印张:14 1/2

印数:1—4 500 字数:323 000

定价:19.00元

(如有印装质量问题,我社负责调换〈环伟〉)

序

高等教育学历文凭考试，是我国高等教育事业发展过程中出现的一个新生事物。它是社会力量办学与国家考试相结合，以宽进严出、教考分离、全日制教学为特点的高等教育形式。这种形式从它产生之日起，就受到社会各界的重视和赞誉，认为它为我国社会主义经济建设对人才的大量需求又提供了一种培养手段，同时它也得到了国家有关领导同志的称赞，认为这是“穷国办大教育”的一条好的途径。北京在全国是最早进行这项试点工作的，目前有 24 所民办高校参加，开设的专业有 16 个，在校生 3 万余人，年均招生规模都在万人左右，其中计算机应用专业 1997 年成为招生规模最大的专业。经过五年的实践和总结，特别是结合国家为民办高等教育培养目标“应用性、职业性”定位的理解，我们感到，我市试点工作在发展过程中，基本建设做得还不够，这其中一个表现就是抓教材建设做得不够，特别是目前市场上还缺乏与高等职业技术教育相匹配的有关教材，导致目前参加文凭考试的民办学校在教学上基本上是借用普通高校的教材，因而教材与培养要求上的矛盾就尤显突出。

我们非常感谢的是，北京航空航天大学计算机系非常积极和认真地提出了要组织编写一套适合这种考试的教材的建议，并做了很具体的安排，北京航空航天大学的许多专家、教授克服了学校教学、科研任务繁忙的困难，按时、保质地完成了撰稿工作。同时此项工作也得到了科学出版社有关同志的大力配合，没有他们耐心、细致地做组稿和出版等工作，这套教材能这样快地面世是不可想象的。

经过各方的通力协作，今天这套教材终于可以奉献给大家了。我们觉得这套教材基本上体现了高等教育学历文凭考试计算机应用专业培养方向上的要求，在内容上也是科学、严谨的，我们同意把这套教材作为推荐使用的教材。同时，我们也希望社会各界对这套教材有什么意见和建议，能及时反馈给我们，以便使它不断完善。

谢谢大家。

北京市高等教育自学考试委员会办公室

1998 年 6 月 12 日

编委会名单

主任

怀进鹏 周 轩

委员

董存稳 沈旭昆 马殿富 李昭原

吴保国 刘 瑞 檀凤琴 何自强

唐发根 任爱华 熊桂喜 邵鸿余

郭俊美 于守谦

前 言

数据库技术是计算机领域发展最快的学科之一，也是应用很广、实用性很强的一门技术。它应用于各类管理信息系统（MIS）、决策支持系统（DDS）、办公自动化（OA）、地理信息系统（GIS）、计算机辅助设计/计算机辅助制造（CAD/CAM），计算机辅助软件工程（CASE）等领域。数据库技术与网络通信技术、面向对象技术、多媒体技术、人工智能技术等互相渗透、互相结合，成为当代数据库技术发展的主要特征。它使数据库领域中新的技术内容层出不穷，新的学科分支不断涌现，形成了新一代数据库系统的大家族。在 Internet 高速发展的信息化的今天，信息资源的经济价值和社会价值越来越明显，建设以数据库为核心的各类信息系统和应用系统，对提高企业的效益、改善部门的管理、改进人们的生活均具有实实在在的意义。因此，学习和掌握数据库技术的基本知识和基本技能已成为计算机专业、计算机应用专业、各类管理专业学生的必修课。

本书共分九章。第一章是数据库系统概述，从数据管理技术的发展阐述了数据库技术的优点，从总体上介绍了数据库技术的一般概念、原理和方法。第二章是关系模型的基本概念。由于关系数据库代替了网状、层次数据库而成为当代主流数据库，为此本书从第二章开始全部介绍关系数据库的基本原理和应用技术。在第二章中，介绍了关系的数据定义和性质、关系键、关系数据库模式和关系数据库等重要概念，同时用较大篇幅对关系代数进行了介绍。第三章是 SQL 语言。SQL 语言是应用广泛的关系数据库的国际标准语言，本章以 Oracle 数据库为背景，较详细地介绍了 SQL 语言的查询、定义、操作及控制等语句成分和使用举例。第四章是关系数据库理论，在介绍了函数依赖的定义和概念后，重点介绍了 2NF, 3NF, BCNF 规范化过程和方法。第五章是数据库设计，重点介绍了用工程化方法设计关系数据库的步骤、方法和关键技术。第六章是数据库安全保护，重点介绍了数据库的安全性、完整性、并发控制和数据库恢复的概念和方法。第七章是 FoxPro 数据库管理系统。FoxPro 数据库是应用最广泛的微机数据库系统，本章重点介绍了 FoxPro 数据库系统的组成、语言成分和运行环境等基本知识。第八章是 FoxPro 应用程序设计，主要介绍了 FoxPro 程序设计的主要语句和编程技术，最后给出了编写应用系统的实例。第九章是数据库技术的发展与展望，介绍了传统数据库的局限性，阐述了新一代数据库的发展。

本书以当前广泛应用和发展成熟的关系数据库技术作为讨论的主要内容。本书在叙述方法上是先从总体上介绍数据库的一般原理与方法，之后全面介绍关系数据库的基本概念、原理、技术、方法以及应用设计与实现等，特别是重点介绍了关系数据库国际标准化语言 SQL 和微机数据库 FoxPro DBMS 的实用技术。最后在分析关系数据库局限性的基础上，阐述了关系数据库的发展以及新一代数据库的发展方向。

本书各章后均附有适量的概念性和计算性的习题，可供学生书面作业和上机使用。

本书取材新颖，尽量反映国内外最近的研究成果，注意理论联系实际，叙述简练易

懂。

本书由北京航空航天大学李昭原教授任主编，负责全书的统编，并编写了第一、二、四、五、六章。吴保国副教授编写了第七、八章。刘瑞讲师编写了第三、九章。

本书是在北京航空航天大学计算机系领导的关怀和支持下编写的。在此谨致以衷心的感谢。

本书大部分文字录入和编辑工作由方武军同志完成，在此表示谢意。

由于编著者水平有限，书中难免有不妥之处，恳请读者批评指正。

编著者

1998年12月18日于北航

目 录

序

前 言

第一章 数据库系统概述	(1)
1.1 信息、数据与数据处理	(1)
1.1.1 数据与信息	(1)
1.1.2 数据处理	(2)
1.2 数据管理发展的三个阶段	(2)
1.2.1 人工管理阶段	(2)
1.2.2 文件系统阶段	(3)
1.2.3 数据库系统阶段	(3)
1.3 什么是数据库系统	(6)
1.3.1 数据(Data)	(6)
1.3.2 用户(Users)	(7)
1.3.3 软件(Software)	(8)
1.3.4 硬件(Hardware)	(8)
1.4 数据库管理系统(DBMS)	(9)
1.4.1 数据库管理系统的抽象层次	(9)
1.4.2 DBMS 的数据存取	(14)
1.4.3 DBMS 的主要功能	(14)
1.4.4 DBMS 的组成	(16)
1.5 数据模型	(17)
1.5.1 数据之间的联系	(17)
1.5.2 数据模型的分类	(19)
1.5.3 层次模型	(20)
1.5.4 网状模型	(20)
1.5.5 关系模型	(22)
习题	(23)
第二章 关系模型的基本概念	(25)
2.1 从格式化模型到关系模型	(25)
2.2 关系的数学定义	(25)
2.2.1 域	(26)
2.2.2 笛卡儿积	(26)
2.2.3 关系的数学定义	(27)
2.3 关系的性质	(29)
2.4 关系的键	(30)
2.4.1 关系键	(30)
2.4.2 候选键和主关系键	(30)

2.4.3	外部关系键	(31)
2.4.4	关系模型的完整性	(31)
2.5	关系数据库模式与关系数据库	(32)
2.5.1	关系模式和关系数据库模式	(32)
2.5.2	关系数据库	(33)
2.6	关系运算	(33)
2.7	关系代数	(34)
2.7.1	传统的集合运算	(34)
2.7.2	专门的关系运算	(35)
	习题	(40)
第三章	SQL 语言	(41)
3.1	SQL 语言概貌及特点	(41)
3.1.1	SQL 语言的发展及标准化	(41)
3.1.2	SQL 语言支持的关系数据库三级逻辑结构	(41)
3.1.3	SQL 语言的主要特点	(42)
3.2	SQL 数据查询功能	(43)
3.2.1	投影检索	(44)
3.2.2	选取检索	(45)
3.2.3	排序的检索	(46)
3.2.4	连表检索	(46)
3.2.5	子查询块嵌套检索	(47)
3.2.6	并、交和差检索	(51)
3.2.7	库函数检索	(52)
3.2.8	分组检索	(53)
3.2.9	算术表达式值的检索	(54)
3.2.10	部分匹配查询	(55)
3.2.11	空值检索	(55)
3.3	SQL 数据定义功能	(56)
3.3.1	定义,修改和撤消数据库用户	(56)
3.3.2	定义,修改和撤消基本表	(57)
3.3.3	定义和撤消视图	(63)
3.3.4	定义和撤消索引	(65)
3.3.5	聚集的建立和撤消	(67)
3.3.6	定义和撤消同义名	(69)
3.4	SQL 数据操纵功能	(70)
3.4.1	INSERT	(70)
3.4.2	DELETE	(71)
3.4.3	UPDATE	(71)
3.5	SQL 数据控制功能	(73)
3.5.1	特权和角色	(73)
3.5.2	系统特权与角色的授予与回收	(73)
3.5.3	对象特权的授予与收回	(74)
3.6	FoxPro SQL 语法规则	(75)

习题	(78)
第四章 关系数据库理论	(81)
4.1 什么是“不好”的数据库设计	(81)
4.2 函数依赖	(83)
4.2.1 函数依赖	(83)
4.2.2 完全函数依赖与部分函数依赖	(84)
4.2.3 传递函数依赖	(85)
4.3 关系键的形式定义	(85)
4.4 规范化	(86)
4.4.1 范式	(86)
4.4.2 第二范式(2NF)	(87)
4.4.3 2NF 规范化	(87)
4.4.4 第三范式(3NF)	(89)
4.4.5 3NF 规范化	(90)
4.4.6 Boyce-Codd 范式(BCNF)	(91)
4.4.7 规范化小结	(93)
习题	(94)
第五章 数据库设计	(96)
5.1 数据库设计概述	(96)
5.1.1 数据库设计方法简述	(96)
5.1.2 数据库设计步骤	(97)
5.2 系统需求分析	(98)
5.2.1 调查分析用户活动	(99)
5.2.2 确定系统边界	(99)
5.2.3 调查分析系统功能	(99)
5.2.4 调查分析系统数据	(99)
5.2.5 编写系统分析报告	(100)
5.3 概念设计	(102)
5.3.1 E-R 图方法	(103)
5.3.2 建立 E-R 模型	(105)
5.4 逻辑设计	(109)
5.4.1 E-R 模型向关系模型转换	(109)
5.4.2 关系规范化	(112)
5.4.3 模式优化	(112)
5.5 数据库物理设计	(113)
5.6 应用程序编码、调试、试运行	(113)
5.6.1 应用程序设计与编写	(113)
5.6.2 组织数据入库	(114)
5.6.3 应用程序的调试与试运行	(114)
5.6.4 整理文档	(114)
5.7 数据库的运行和维护	(114)
习题	(115)
第六章 数据库安全保护	(116)

6.1 数据库的安全性控制	(116)
6.1.1 视图定义和查询修改	(116)
6.1.2 访问控制	(117)
6.1.3 数据加密	(119)
6.1.4 跟踪审查	(119)
6.2 数据库完整性控制	(119)
6.2.1 完整性规则	(120)
6.2.2 完整性约束分类	(120)
6.3 数据库并发控制	(121)
6.3.1 封锁	(122)
6.3.2 死锁及消除的方法	(122)
6.4 数据库的恢复	(124)
6.4.1 预备性恢复	(124)
6.4.2 恢复过程	(125)
习题	(125)
第七章 FoxPro 数据库管理系统	(126)
7.1 Visual FoxPro DBMS 运行环境、组成和运行方式	(126)
7.1.1 Visual FoxPro DBMS 的运行环境	(126)
7.1.2 Visual FoxPro DBMS 的组成	(126)
7.1.3 Visual FoxPro DBMS 的运行方式	(128)
7.2 Visual FoxPro 语言	(129)
7.2.1 Visual FoxPro 命令语法规则	(129)
7.2.2 数据类型	(129)
7.2.3 函数与表达式	(132)
7.2.4 内存变量和数组的操作	(134)
7.2.5 数据定义	(136)
7.2.6 数据操作	(139)
7.2.7 数据库的统计汇总	(150)
7.2.8 多表连接操作	(152)
7.2.9 报表设计	(155)
7.2.10 环境参数设置与辅助操作	(157)
习题	(158)
第八章 FoxPro 应用程序设计	(161)
8.1 程序文件的建立、修改和执行	(161)
8.1.1 Visual FoxPro 的程序概念	(161)
8.1.2 程序文件建立和修改命令	(161)
8.1.3 程序文件的执行	(162)
8.2 输出和输入命令	(162)
8.2.1 输出命令	(162)
8.2.2 输入命令	(165)
8.2.3 程序调试命令和注释语句	(169)
8.3 内存变量和数组作用域的定义	(170)
8.3.1 定义全局内存变量或数组	(170)

8.3.2 定义私有内存变量或数组	(170)
8.3.3 定义局部内存变量或数组	(171)
8.3.4 数组定义和重新定义	(171)
8.4 程序的控制结构	(172)
8.4.1 顺序结构	(172)
8.4.2 选择分支结构	(173)
8.4.3 循环结构	(178)
8.5 主程序、子程序、过程和自定义函数	(183)
8.5.1 主程序与子程序	(183)
8.5.2 过程与过程文件	(185)
8.5.3 用户自定义函数	(188)
8.5.4 子程序、过程与函数的对比	(190)
8.5.5 捕捉按键信息	(190)
8.6 应用系统开发实例	(192)
8.6.1 应用系统的总体设计	(192)
8.6.2 编写应用程序	(192)
习题	(197)
第九章 数据库技术的发展与展望	(198)
9.1 传统数据库的局限性	(198)
9.2 新一代数据库技术	(199)
9.2.1 关系数据库系统的发展	(200)
9.2.2 数据库技术与其他相关技术相结合	(202)
9.2.3 面向应用领域的数据库新技术	(203)
9.2.4 面向对象数据库技术	(204)
习题	(207)
附录一 全屏幕光标控制键说明	(208)
附录二 北京市高等教育学历文凭考试“数据库原理与应用”课程考试大纲	(210)
参考文献	(218)

第一章 数据库系统概述

随着计算机技术的蓬勃发展,计算机应用从科学计算、过程控制进入数据处理,计算机已从少数科学家手中的珍品变成人们日常工作中处理数据的得力助手和有力工具。当今世界,在计算机的三大主要应用(科学计算、过程控制和数据处理)领域中,数据处理迅速上升为计算机应用的主要方面。数据处理的中心问题是数据管理。数据库系统技术是数据管理技术发展的最新研究成果。在本章中,主要介绍数据管理技术的发展、数据模型和数据库系统的基本概念等,为后面各章的学习打下基础。

1.1 信息、数据与数据处理

在信息社会中,信息是一种资源,其重要性可以与物质和能量相提并论,是一个企业生存和发展所必需的。按信息论的观点,任何社会实践活动都可以抽象为人流、物流、财流、能源流和信息流这五种流运动,其中起主导和支配作用的是信息流,它调节和控制着其他各流的数量、方向、速度和目标,从而使社会实践活动更具有目的性和规律性。信息是维持生产活动、经济活动和社会活动必不可少的资源,因此,信息是有价值的,信息成为构成客观世界的三大要素(信息、能源和材料)之一。人们为了获取有价值的信息用于决策,就需要对数据进行处理,进行管理。

人们把用计算机对数据进行处理的应用系统称为计算机信息系统。信息系统是“一个由人、计算机等组成的能进行信息的收集、传递、存储、加工、维护、分析、计划、控制、决策和使用的系统”。信息系统的核心是数据库。

1.1.1 数据与信息

信息的英文单词是“information”,数据的英文单词是“data”。信息与数据有着不同的含义。

1. 信息

“信息”是指现实世界事物存在方式或运动状态的反映。具体地说,信息是一种已经被加工为特定形式的数据,这种数据形式对接收者来说是有意义的,而且对当前和将来的决策具有明显的或实际的价值。

信息有如下一些重要特征:

- (1)信息传递需要物质载体,信息的获取和传递要消耗能量。
- (2)信息是可以感知的。不同的信息源有不同的感知方式(如感觉器官、仪器或传感器)。
- (3)信息是可以存储、压缩、加工、传递、共享、扩散、再生和增值的。

2. 数据

“数据”是将现实世界中的各种信息记录下来的、可以识别的符号,是信息的载体,是信息的具体表示形式。可用多种不同的数据形式来表示一种同样的信息,而信息不随它的数据形式不同而改变。数据的表现形式多种多样,不仅有我们熟知的数字和文字,还可以有图形、图像、声音等形式。

3. 数据与信息的联系

信息是各种数据所包括的意义,数据则是载荷信息的物理符号。例如,“今年大学招生录取线为 450 分,张红的高考成绩为 500 分”。这段文字(数据)提供了张红可能被某一个大学录取的信息。可见数据与信息是密切相关联的。因此,在许多场合下,对它们不做严格的区分,可互换使用。例如通常说的,“信息处理”与“数据处理”;“信息资源”与“数据资源”;“信息采集”与“数据采集”等就具有同义性。

1.1.2 数据处理

数据处理(也称为信息处理)实际上就是利用计算机对各种形式的数据进行处理。它包括:数据采集、整理、编码和输入,有效地把数据组织到计算机中,由计算机对数据进行一系列存储、加工/计算、分类、检索、传输、输出等操作过程。其目的是从大量的原始数据中抽取和推导出对人们有价值的信息,以作为行动和决策的依据。比如,一个企业,需要收集大量的有关市场产品销售的信息(数据),经存储、加工/计算,生成市场销售情况图表,从而获得哪种型号的产品最受欢迎的信息,由此指导生产计划,...

数据处理的一系列活动中,数据收集、存储、分类、传输等操作作为基本操作,这些基本操作环节称为数据管理,而加工/计算、输出等操作是千变万化的,不同业务有不同的处理。数据管理技术是解决上述基本环节的,而其他环节是由应用程序实现的。

随着计算机软硬件技术的发展,数据管理技术的发展大致经历了人工管理、文件系统和数据库系统三个阶段。

1.2 数据管理发展的三个阶段

1.2.1 人工管理阶段

50 年代中期以前,计算机主要用于科学计算。在这一阶段,计算机除硬件外,没有管理数据的软件。使用计算机对数据进行管理时,设计人员除考虑应用程序、数据的逻辑定义和组织外,还必须考虑数据在存储设备内的存储方式和地址。其特点是:

(1)数据不保存。因为计算机主要用于科学计算,不要求保存数据。每次计算机先将程序和数据输入主存,计算结束后,将结果输出,计算机不保存程序和数据。

(2)编写程序时要安排数据的物理存储。程序员编写应用程序时,还要安排数据的物理存储。程序和数据混为一体,一旦数据的物理存储改变,必须要重新编程,程序员的工作量大,繁琐,程序难以维护。

(3)数据面向程序。每个程序都有属于自己的一组数据,程序与数据相互结合成为一

体,互相依赖。各程序之间的数据不能共享,因此数据就会重复存储(冗余度大)。

数据与程序之间的关系如图 1.1 所示。

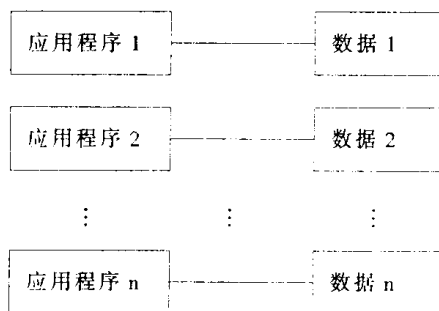


图 1.1 数据与程序之间的关系

1.2.2 文件系统阶段

在 50 年代后期至 60 年代中期,计算机外存已有了磁鼓、磁盘等存储设备,软件有了操作系统。

人们在操作系统的支持下,设计开发了一种专门管理数据的计算机软件,称之为文件系统。这时,计算机不仅用于科学计算,也已大量用于数据处理,其特点是:

(1)数据以文件的形式长期保存。由于计算机大量用于数据处理,数据需要长期保留在外存上反复处置,即经常对其进行查询、修改、插入和删除等操作。因此,在文件系统中,按一定的规则将数据组织为一个文件,存放在外存储器中长期保存。

(2)数据的物理结构与逻辑结构有了区别,但较简单。程序员只需用文件名与数据打交道,不必关心数据的物理位置,可由文件系统提供的读写方法去读/写数据。

(3)文件形式多样化。为了方便数据的存储和查找,人们研究了许多文件类型,如索引文件、链接文件、顺序文件和倒排文件等。数据的存取基本上是以记录为单位的。

(4)程序与数据之间有一定的独立性。应用程序通过文件系统对数据文件中的数据进行存取和加工,因此,处理数据时,程序不必过多地考虑数据的物理存储的细节,文件系统充当应用程序和数据之间的一种接口,这样可使应用程序和数据都具有一定的独立性。这样,程序员可以集中精力于算法,而不必过多地考虑物理细节。并且,数据在存储上的改变不一定反映在程序上,这可以大大节省维护程序的工作量。

尽管文件系统有上述优点,但是,这些数据在数据文件中只是简单地存放,文件中的数据没有结构,文件之间并没有有机的联系,仍不能表示复杂的数据结构;数据的存放仍依赖于应用程序的使用方法,基本上是一个数据文件对应于一个或几个应用程序;数据面向应用,独立性较差,仍然出现数据重复存储,冗余度大,一致性差(同一数据在不同文件中的值不一样)等问题。文件系统中程序与数据之间的关系如图 1.2 所示。

1.2.3 数据库系统阶段

从 60 年代末期开始,随着计算机技术的发展,数据管理的规模越来越大,数据量急剧增加,数据共享的要求越来越高。这时磁盘技术取得了重要进展,大容量和快速存取磁盘陆续进入市场,为数据库技术的发展提供了物质条件。人们研制出了一种新的、先进的数据管理方法,即数据库系统。数据库系统克服了以前所有数据管理方式的缺点,试图提供

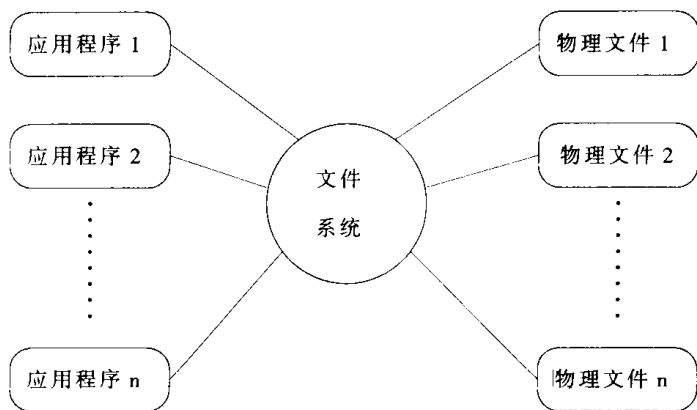


图 1.2 文件系统阶段

一种完美的、更高的数据管理方式。它的指导思想是对所有的数据实行统一的、集中的、

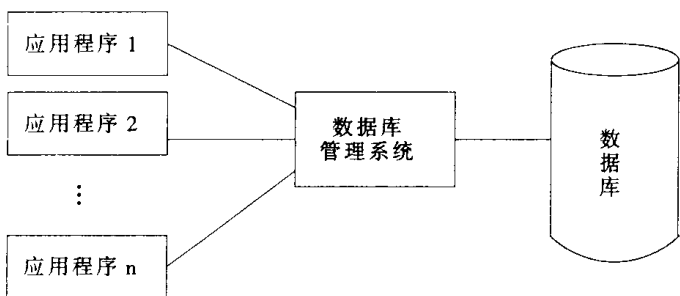


图 1.3 数据库系统阶段

独立的管理,使数据存储独立于使用数据的程序,实现数据共享。如图 1.3 所示,通过一个叫做数据库管理系统的软件集中管理独立存放的数据,并成为用户的应用程序和数据库打交道的接口。

数据库系统管理方式具有如下特点:

1. 数据共享

这是数据库系统区别于文件系统的最大特点之一,也是数据库系统技术先进性的重要体现。

共享是指多用户、多种应用、多种语言互相覆盖地共享数据集合,如图 1.4 所示。

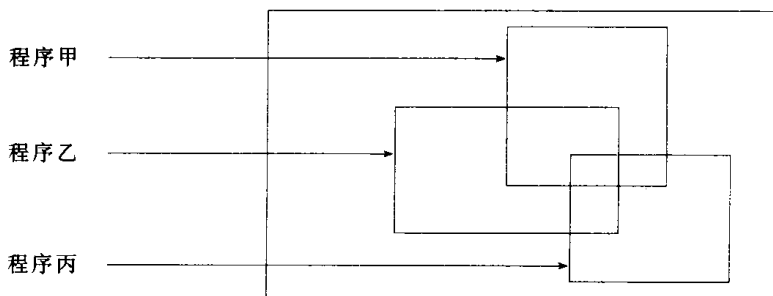


图 1.4 数据共享

当前的所有用户可同时存取数据库中的数据。可以是数据库中的一部分数据或相互交叉重叠的数据,可以实现文件级、记录级、数据项级的数据共享。

2. 面向全组织的数据结构化

数据库系统不再像文件系统那样从属于特定的应用,而是面向整个组织来组织数据,常常是按照某种数据模型,将整个组织的全部数据组织成为一个结构化的数据整体。它

不仅描述了数据本身的特性,而且也描述了数据与数据之间的种种联系,这使数据库能够描述复杂的数据结构。

全组织的数据结构化,有利于实现数据共享。比如,一个学校,可以把学校所有的各个应用(人事、学籍、科研、财务、后勤等)的数据组织到一个数据库中,并且结构化。实现集中、统一的存储与管理,那么各种应用可存取各自相关的数据子集,满足各种应用要求,实现数据共享。

全组织的数据结构化,使数据不是面向应用(程序),而是面向系统。这样系统弹性大,可扩充。在系统中加一部分数据,便可适应新的应用要求。

全组织的数据结构化,是数据库管理有别于文件管理的一个重要特征。文件系统是一个无结构的数据集合,文件之间是孤立的,不能反映现实事物之间的联系,数据面向专用,没有弹性,不宜扩充。

3. 数据独立性

文件系统管理中,应用程序严重依赖于数据文件,如果把应用程序使用的磁带顺序文件改成为磁盘索引文件,则必须对应用程序进行修改。而数据库技术的重要特征就是数据独立于应用程序而存在,数据与程序相互独立,互不依赖,不因一方的改变而改变另一方。这大大简化了应用程序的设计与维护的工作量。

4. 可控数据冗余度

数据共享、结构化和数据独立性的优点使数据存储不必重复,不仅可以节省存储空间,而且从根本上保证了数据的一致性,这又是有别于文件系统的重要特征。

从理论上讲,数据存储完全不必重复,即冗余度为零,但有时为了提高检索速度,常有意安排若干冗余,这种冗余由用户控制,称为可控冗余度。可控冗余要求任何一个冗余的改变都能自动地对其余冗余加以改变。这个过程叫做传播更新。

5. 统一数据控制功能

数据库是系统中各用户的共享资源,因而计算机的共享一般是并发的,即多个用户同时使用数据库。因此,系统必须提供以下四个方面的数据控制功能:

(1) 数据安全性控制。数据的安全性控制是指采取一定安全保密措施确保数据库中的数据不被非法用户存取,防止造成数据丢失、破坏或被盗。

(2) 数据完整性控制。数据的完整性指数据的正确性、有效性与相容性。所谓正确性是指数据的合法性;所谓有效性是指数据是否在定义的有效范围;所谓相容性是指表示同一个事实的两个数据应相同。系统要提供必要的功能,保证数据库中的数据在输入、修改过程中始终符合原来的定义和规定。

(3) 并发控制。当多个用户并发进程同时存取、修改数据库中的数据时,可能会发生互相干扰而得到错误结果,并使数据库完整性遭到破坏,因此必须对多用户的并发操作加以控制和协调。

(4) 数据恢复。当系统发生故障造成数据丢失或对数据库的数据操作发生错误时,系统能进行应急处理,把数据库恢复到正确状态。

从文件系统管理发展到数据库系统管理是信息处理领域的重大变化,人们由传统的关注系统功能设计(因为程序设计处于主导地位,数据服从于程序)转向关注数据的结构设计,数据的结构设计成为信息系统首要关心的中心问题。

1.3 什么是数据库系统

数据库技术已被认为是“计算机和信息科学增长最迅速的领域之一”。它很快发展成为实践上和理论上都相当重要且相当成熟的领域之一,它是计算机软件方面的一个独立分支,目前还在发展之中。自 60 年代末以来,有许多从事数据库研究的专家都给出了关于数据库系统的定义,从不同的角度来描述数据库系统。这里我们参照 C.J. Date1986 年出版的《数据库系统导论》(卷 1)第四版中关于什么是数据库系统的描述介绍如下:

数据库系统,从根本上说不过是计算机化的记录保持系统,也就是说,它的总目的是存储和产生所需要的有用信息。这些有用的信息可以是使用该系统的个人或组织的有意义的任何事情,换句话说,是对某个人或组织辅助决策过程中不可少的事情。图 1.5 是一幅最简单的数据库系统示意图。

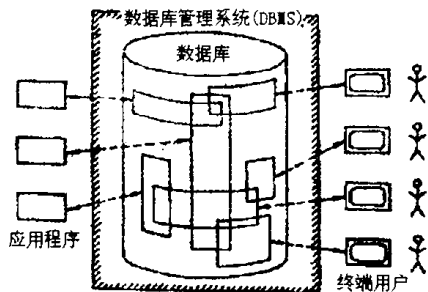


图 1.5 数据库系统简图

从图 1.5 中可以看出,一个数据库系统要包括四个主要部分:数据(库)、用户、软件、硬件。下面做简要介绍。

1.3.1 数据(Data)

数据,是数据库系统中集中存储的一批数据集合——即数据库。它是数据库系统的工作对象。

为了把输入、输出或中间数据加以区别,我们常把数据库数据称为“存储数据”或“工作数据”或“操作数据”。它们是某特定应用环境中进行管理和决策所必需的信息。

特定的应用环境,可以指一个制造公司,或一个银行,或一所医院,或一所学校,或一个商店,或一个政府部门,或一个城市,或某个团体或个人等等各种各样的应用环境。在这些各种各样的应用环境中,各种不同的应用可通过访问其数据库,获得必要的信息,以辅助进行决策,决策完成后,再将决策结果存储在数据库中。

特别需要指出,数据库中的存储数据是“集成的”和“共享的”。

所谓“集成”,是指把某特定应用环境中的各种应用相关的数据及其数据之间的联系(联系也是一种数据)全部地集中地并按照一定的结构形式进行存储,或者说,把数据库看成若干单个性质的数据文件的联合和统一的数据整体,并且在文件之间局部或全部消除了冗余。这使数据库系统具有整体数据结构化和数据冗余小的特点。

所谓“共享”,是指数据库中的一块块数据可为多个不同的用户所共享,即多个不同的用户,使用多种不同的语言,为了不同的应用目的,而同时存取数据库,甚至同时存取同一块数据。共享实际上是基于数据库是“集成”的这一事实的结果。

支持上述共享的数据库系统,有时称为多用户系统。大型机上的数据库系统是多用