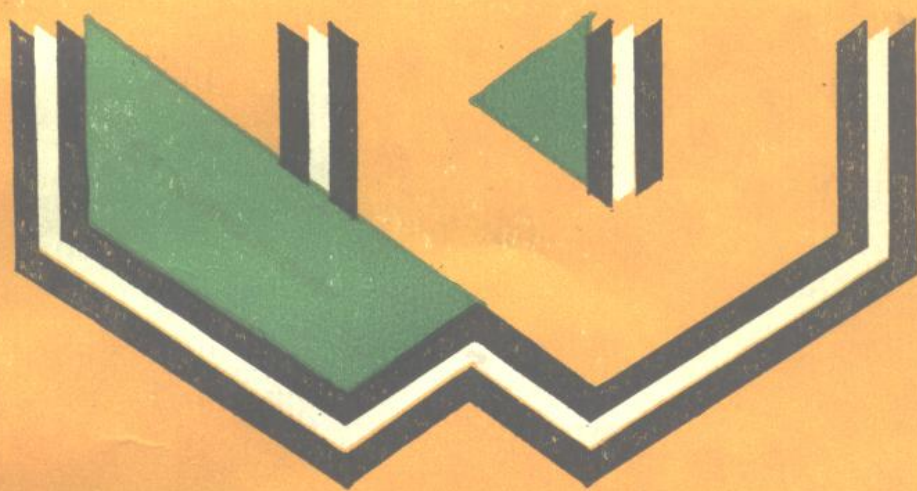


高等财经院校

试用教材



数据库方法

冯 关 源 主 编

中国财政经济出版社

311.132.3
GY/1

TP311.132.3

FGY/1

高等财经院校试用教材

数 据 库 方 法

冯关源 主编

中国财政经济出版社

35-10621

高等财经院校试用教材
数据库方法
冯关源 主编

中国财政经济出版社出版

(北京东城大佛寺东街8号)

新华书店北京发行所发行 各地新华书店经售

北京印刷二厂印刷

787×1092毫米 16开 17.25印张 414 000字

1988年6月第1版 1989年6月北京第1次印刷

印数：1—5 000 定价：3.15元

ISBN 7-5005-0563-9/F·0519(课)

编 审 说 明

本书是全国财经类通用教材。经审阅，我们同意作为高等财经院校试用教材版。出书中不足之处，请读者批评指正。

财政部教材编审委员会

1988年9月

前 言

数据库是数据处理的最新方法，也是信息管理中的一项非常重要的新技术。十多年来，这种新技术在我国发展很快，并在国防、国民经济、科学研究、文化教育和社会生活等各方面得到广泛应用。

数据库的发展至今仅有20多年，现在正处于青春时期，而且充满了活力。一方面，一些较成熟的技术，如各种数据库管理系统和传统的数据库设计方法已投入实用；另一方面，尚有许多理论及实际问题有待完善、开发和探索，如数据模型、关系数据库的新理论进展、分布式数据库、智能数据库、知识库等。所以说，数据库这门学科，目前处于既需普及，又有待提高的状态。

最近几年，财经院校的经济信息专业都把数据库列入必修课程。目的在于希望通过本课程的学习，使学生们能掌握这个方法，并很好地用于经济信息管理。为了提高这门课程的教学质量，1986年，财政部教育司委托上海财经大学信息系和中南财经大学信息系负责编写本教材。

本书由上海财经大学冯关源主编。其中，第一、二、三、四章及附录二，由上海财经大学华似燕编写；第五、六章及附录一，由中南财经大学吴盘珍编写；第七、八、九章，由冯关源编写。中南财经大学信息系胡久清、刘腾红两位同志详细地阅读了本书，提了不少宝贵意见，财政部教育司对本书的出版给以大力支持。在这里，我们对所有为本书的编写、审定、出版给予支持和帮助的同志表示衷心的感谢。

编 者

1987年8月

目 录

第一章 绪论	1
§1.1 什么是数据库	1
一、数据库的定义	1
二、数据库的特点	1
三、数据库与图书馆	2
§1.2 数据库系统的产生和发展	3
一、早期的人工数据管理	3
二、文件系统	3
三、早期的数据库系统	5
四、现在的数据库系统	6
§1.3 数据库系统与管理信息系统	6
一、信息系统	7
二、管理信息系统	7
三、数据库系统与管理信息系统	8
复习思考题一	8
第二章 数据模型	9
§2.1 实体、信息和数据	9
一、信息和数据	9
二、现实世界和观念世界	9
三、数据世界和数据模型	10
§2.2 实体之间的各种联系	10
一、一对一关系	10
二、一对多关系	11
三、多对多关系	11
§2.3 实体—联系(E—R)方法	11
一、E—R图	11
二、如何划分实体和实体属性	12
§2.4 三种基本数据模型	13
一、层次模型	13
二、网络模型	13
三、关系模型	14
§2.5 三种数据模型之间的转换和比较	16
一、模型的选用	16
二、模型的转换	16
三、三种数据模型的比较	17
复习思考题二	17

第三章 数据库系统简介	18
§3.1 数据库系统的组成	18
一、数据库系统的硬件组成	18
二、数据库系统的软件组成	18
§3.2 数据库的逻辑关系	19
§3.3 模式和子模式	20
一、模式	26
二、子模式	21
三、用户、模式、子模式之间的关系	21
§3.4 数据库管理系统 (DBMS)	21
一、DBMS 的功能	22
二、DBMS 的组成	22
§3.5 数据库管理员 (DBA)	24
§3.6 怎样建立数据库	24
§3.7 怎样使用数据库	25
复习思考题三	26
第四章 层次型数据库系统	27
§4.1 层次型数据模型	27
§4.2 IMS 系统的逻辑结构	28
*§4.3 IMS 的存贮结构及存取方法	29
一、IMS 存贮结构概述	29
二、HSAM (Hierarchical Sequential Access Method) 方法	30
三、HISAM (Hierarchical Indexed Sequential Access Method) 方法	30
四、HDAM (Hierarchical Direct Access Method) 方法	30
五、HIDAM (Hierarchical Indexed Direct Access Method) 方法	31
六、四种方法的比较	32
§4.4 IMS 的物理数据库 (PDB)	33
一、物理数据库	33
二、物理数据库描述	33
三、物理数据库描述举例	34
§4.5 IMS 的逻辑数据库 (LDB)	35
一、IMS 中 LDB 的第一种含义	35
二、IMS 中 LDB 的第二种含义	36
三、建立 LDB 的规则	33
四、逻辑数据库描述	39
§4.6 程序说明块 PSB	39
一、PSB 控制语句	39
二、PSB 举例	40
*§4.7 辅助索引	40
§4.8 IMS 的数据操纵语言	41
一、片段查找变量 SSA (Segment Search Argument)	41

二、DL/1操作命令	(42)
三、DL/1操作命令举例	(43)
§4.9 应用程序举例	(44)
一、应用程序与 MIS 的接口途径	(44)
二、应用程序举例	(44)
复习思考题四	(46)
第五章 网络型数据库	(48)
§5.1 网络模型的一般概念	(48)
一、网络数据结构的类型	(48)
二、网络数据结构的分解	(50)
§5.2 DBTG 系统的结构	(51)
§5.3 DBTG 系统的基本概念	(52)
一、DBTG 系 (Set)	(52)
二、系型和系值	(55)
三、奇异系	(57)
*§5.4 DBTG 的存取策略	(58)
一、域 (Area)	(58)
二、数据库码 (Data Base Key)	(58)
三、记录存放方式 (Location Mode)	(59)
四、系序 (Set Order)	(60)
五、系的实现方法	(61)
六、当前值 (Occurrence)	(62)
七、属籍类别 (Membership Class)	(63)
八、系值选择 (SET SELECTION)	(65)
§5.5 DBTG 模式描述语言	(67)
一、模式条目	(67)
二、域条目	(67)
三、记录条目	(67)
四、系条目	(68)
§5.6 DBTG 子模式描述语言	(70)
一、子模式与模式的区别	(70)
二、DBTG 子模式描述语言	(70)
三、一个 COBOL 子模式的实例	(72)
§5.7 DBTG 数据操纵语言	(73)
一、存取机制	(73)
二、DBTG 中程序的运行环境	(73)
三、DML 和主语言的接口	(75)
四、DML 语句	(75)
§5.8 应用程序例	(89)
一、标识部	(89)
二、数据部	(90)
三、过程部	(90)

复习思考题五	(93)
第六章 关系型数据库	(95)
§6.1 关系模型概述	(95)
§6.2 关系模型的基本概念	(95)
一、关系的数学定义	(95)
二、关系模型	(97)
§6.3 关系数据语言概述	(100)
§6.4 关系代数	(100)
一、通常的集合运算	(101)
二、特殊的关系运算	(103)
§6.5 关系演算	(107)
一、检索操作	(107)
二、存贮操作	(112)
§6.6 按例查询语言 (QBE)	(112)
一、QBE 概述	(113)
二、检索操作	(114)
三、存贮操作	(116)
*§6.7 结构查询语言 (SQL)	(121)
一、数据定义	(121)
二、数据操纵	(124)
三、数据控制	(125)
四、宿主型 SQL	(129)
*§6.8 关系数据库系统综述	(130)
一、关系数据库的特点	(130)
二、关系数据语言的比较	(131)
*§6.9 查询优化概述	(133)
一、优化概述	(133)
二、优化的一般策略	(134)
三、关系代数表达式的等价变换规则	(135)
复习思考题六	(136)
第七章 关系数据理论	(138)
§7.1 问题的提出	(138)
§7.2 函数依赖	(139)
一、属性间的联系	(139)
二、函数依赖	(139)
三、函数依赖与属性关系	(141)
§7.3 关系模式的规范化	(141)
一、基本概念	(141)
二、第一范式 (1NF)	(142)
三、第二范式 (2NF)	(143)
四、第三范式 (3NF)	(144)

五、规范化举例	(144)
六、BC 范式 (BCNF)	(147)
七、多值依赖与第四范式 (4NF)	(148)
八、规范化小结	(159)
*§7.4 函数依赖的公理系统	(150)
一、函数依赖的逻辑蕴涵	(150)
二、函数依赖的公理系统	(151)
三、闭包 F^+	(151)
四、属性集闭包 X_F^+	(152)
五、 F^+ 的计算	(153)
六、函数依赖集的等价和覆盖	(153)
七、函数依赖集的最小集 F_m	(154)
§7.5 模式的分解	(156)
一、模式分解的三种定义	(156)
二、无损联接的分解	(157)
三、保持函数依赖的分解	(159)
四、模式分解的算法	(160)
复习思考题七	(161)
第八章 数据库设计	(162)
§8.1 数据库设计综述	(162)
一、数据库设计过程	(162)
二、数据库设计的步骤	(162)
三、数据库设计方法的目标	(165)
四、数据库设计方法的内容	(165)
§8.2 概念设计	(168)
一、概念结构的特点	(168)
二、概念设计的步骤	(168)
三、结构化实体图 (Structured Entity Charts)	(174)
§8.3 逻辑设计	(176)
一、逻辑设计的步骤	(176)
二、E—R图向一般数据模型的转换	(177)
§8.4 物理设计	(178)
一、物理设计的内容	(178)
二、物理设计的步骤	(179)
三、性能评价	(180)
§8.5 数据库的实施和维护	(182)
一、模式和子模式的编译	(182)
二、载入数据	(182)
三、试运行	(183)
四、运行和维护	(183)
复习思考题八	(183)

第九章 数据库的保护.....	(184)
§9.1 数据库保护概述	(184)
§9.2 安全性	(185)
一、安全性保护策略	(185)
二、数据库安全检查纵览	(188)
三、现行系统中的安全措施	(189)
§9.3 完整性	(193)
一、完整性约束条件	(193)
二、现行系统的完整性约束	(194)
§9.4 并发控制	(197)
一、并发控制的基本问题	(197)
二、并发控制的实施	(198)
三、死锁的处理	(201)
§9.5 数据库的恢复	(202)
一、数据库恢复的基本技术	(202)
二、数据库恢复的原则	(203)
三、数据库的恢复方式	(203)
四、IMS 的系统恢复	(204)
复习思考题九	(204)
附录一 dBASE III 简介	(206)
§1 dBASE III 基本概念	(206)
一、dBASE III 中名字的约定	(206)
二、dBASE III 文件类型	(206)
三、数据库文件的结构	(207)
四、常数、变量、运算符、表达式及函数	(208)
五、命令结构	(210)
§2 数据库的基本操作	(211)
一、进入和退出 dBASE III	(211)
二、建立数据库文件的结构	(211)
三、打开数据库文件	(212)
四、数据库记录的输入	(213)
五、记录指针的移动	(214)
六、数据库内容的输出	(214)
七、数据库的分类、索引和查找	(215)
八、数据库的编辑	(217)
九、删除数据库记录	(219)
十、某些数值参数的处理	(219)
十一、多个数据库文件的操作	(220)
§3 数据库的辅助操作	(222)
一、列目录命令 DIR	(222)
二、清除文件命令 ERASE	(222)
三、文件复制命令 COPY	(222)

四、关闭文件命令 CLOSE	(222)
§4 dBASE III 程序设计	(223)
一、概述	(223)
二、命令文件的建立和执行	(223)
三、控制语句	(224)
四、过程调用	(225)
五、程序交互性命令	(226)
六、中止程序执行语句及注释语句	(226)
§5 输入输出格式设计	(227)
一、用于屏幕格式设计的命令	(227)
二、用于打印机输出的格式设计命令	(223)
§6 应用程序举例	(228)
§7 dBASE III 命令、函数及全屏幕编辑	(234)
一、命令一览表	(234)
二、函数一览表	(233)
三、全屏幕编辑控制键	(233)
附录二 System R 关系数据库系统	(211)
§1 System R 简介	(241)
一、System R 的背景	(241)
二、System R 的基本结构	(241)
§2 System R 的数据语言 SQL	(242)
一、检索操作	(242)
二、存贮操作	(245)
§3 System R 的外层	(246)
一、窗口	(246)
二、窗口上的数据操作	(247)
三、窗口和数据独立性	(250)
§4 System R 的里层结构	(251)
一、RDS	(251)
二、RSS	(253)
§5 小结	(256)
附录三 图表索引	(258)
附录四 参考资料	(264)

第一章 绪 论

§1.1 什么是数据库

数据库技术是计算机软件的一个重要的分支,它是计算机信息科学中迅速发展的一个领域。虽然不少计算机科学工作者对数据库的设计原则和实现方法作了多方面的研究,但它同计算机科学的其它分支比起来还是属于比较年轻的,它的概念、原理和方法仍然在不断地变化发展。直到60年代末,计算机厂商才开始为用户提供商品化的数据库管理系统。随着计算机应用的普及及信息科学的迅速发展,现在数据库技术在理论和实践上都已变得很重要,在一些计算机技术发达的国家,现在已有不少机关企业完全依赖于他们的数据库管理系统来进行他们的日常业务工作。

一、数据库的定义

近20年来,虽然数据库技术已有了相当的发展,但对数据库系统下一个简单而又确切的定义却仍然是困难的。这是因为数据库系统是一个很复杂、涉及面很广的系统。不少人曾对数据库做过各种定义,这些定义各有侧重,但比较可取的是如下这一定义:

数据库是存贮在一起的相关数据的集合,这些数据无有害或不必要的冗余,可为多种应用服务;这些数据的存贮独立于使用它的应用程序;对数据库插入新数据,修改和检索原有数据均能按一种公用的和可控制的方法进行;这些数据具有结构特性,因而为今后的应用研究提供了基础。当某个系统中存在结构上完全分开的若干数据库时,则说该系统包含一个数据库集合。

由于这个定义还是比较抽象的,所以我们还可以从数据库比较显著的特点来进一步认识数据库。数据库将具有一定关系的数据组成一个集合,它主要有结构化、独立性和共享性这三大特点。

二、数据库的特点

(一) 数据的结构化

数据库中的数据可以被用来进行一系列的操作,因此它就应具有比较复杂的数据模型。比如在一个大学中人事部门对教工需要一个人事基本记录;财会部门对教工需要一个工资历史记录;而技术部门对教工需要一个业务记录。因此我们需要设计一个教工记录,它不仅对某个部门具有应用灵活性,而且还要对包含多个部门的数个组织系统具有应用灵活性,这样我们就可以设计如图1-1所示的教工记录。这样的教工记录就把整个组织的数据结构化了。从这个教工记录中我们可以看到,这样的记录不仅描述了各项数据本身,而且还描述了数据之间的联系,这些联系是通过存取路径来实现的。通过存取路径来表示各类自然数据之间的联系,这是数据库与一般传统的数据文件的根本区别。

从图1-1的记录中,我们还可以看到,通过数据的结构化可以大大减少数据冗余度,既节约空间,又减少存取时间。另外,由于数据结构化要求从整体观点看数据,所以数据不是

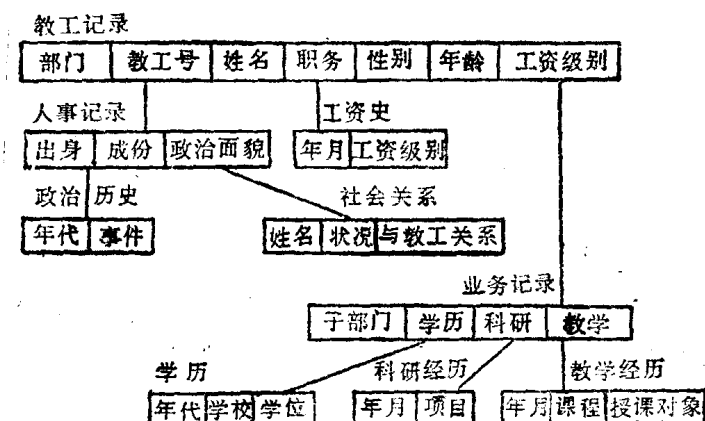


图1-1 教工记录

应用程序对数据的存贮结构和存取方法有较高的独立性。由于系统在数据的存贮结构与逻辑结构之间提供了映象，因此如果存贮结构或物理结构改变了，其逻辑结构可不改变，所以程序员可不必修改程序，这就是数据库的物理独立性。

由于系统在局部逻辑结构和整体逻辑结构之间提供了映象，所以总体逻辑结构变化了，其局部逻辑结构可以不变，而程序员是根据局部逻辑结构编写程序的，因此程序可以不改变，这就是数据的逻辑独立性。

(三) 数据的共享性

在早期的数据组织中，一组数据只是为一种应用而设计的。而数据库的目标是用同一组数据为尽可能多的应用服务。一个数据库的数据可以由几个部门，甚至由跨越行政边界的不同的应用系统共享。数据库的数据共享性是数据大量集成的直接结果。例如前面提到的教工记录就可为人力资源部门、工资部门、业务部门所共享，但各个部门对这个记录关心的是各个不同的部分。当然，不同的部门所关心的数据子集合也有重叠的可能，这就是说，不同的用户对数据库中的数据有不同的理解和解释方式。

共享性还包括了多个用户可同时使用数据库中数据的意义，几个用户可同时访问数据库，甚至同一个数据。能够提供这种共享性的数据库称为多用户数据库系统。

三、数据库与图书馆

为了能对数据库有更形象的理解，我们把它与大家都很熟悉的图书馆作个比较。图书馆是存贮和负责借阅图书的部门。而数据库则是存贮数据并负责用户访问数据的机构。正象图书馆不能简单地与书库等同起来一样，我们不能把数据库仅仅理解为大量数据的简单集合。一个图书馆要想很好的为读者服务，必须完成以下工作：

(1) 建立完善的书卡。书卡的格式内容可包括：书号、书名、作者名、出版社名、出版时间、内容摘要等。(2) 把图书有组织地存放在书库中。图书馆的书库中有很多房间和书架，分放图书需要按照一定的顺序和规则，并列出各类书籍的存放对应关系表，以便管理人员迅速查找。(3) 规定借阅权限。不同类型的图书其借阅的对象也不同，如机密图书只供有特权的人借阅；某些书只供读者在馆内翻阅。(4) 建立周密的借阅管理制度。读者借书要先出示借书证，图书管理员验明读者身份和借阅权限后，根据读者填写的借书单（访问

面向某种应用的，而是面向系统的，它弹性大，使用灵活，并且可扩充，即可增加一些数据来满足新的应用要求，因此，数据结构化是数据库的主要特征。

(二) 数据的独立性

在数据库中有两类独立性，即物理独立性和逻辑独立性。

数据库系统具有一种映象功能，因此在使用数据库时，

请求)，按书籍书架的对应关系表，到书库中去查找图书并交与读者（响应），并作某些登记（日志），还书时管理员要按对应关系表把交还的图书送回原来的书架。

对一个数据库来说，它所要完成的工作也类似于上述的图书馆工作。一个数据库系统和一个图书馆的相似性可以用表1-1来清楚地表示。

表1-1

数据库系统与图书馆

序 号	数 据 库	图 书 馆
1	数 据	图 书
2	外 存	书 库
3	用 户	读 者
4	用户标识	借 书 证
5	数据模型	书卡格式
6	数据库管理系统	图书馆管理员
7	数据的物理组织方法	图书的物理存放办法
8	用户对数据库的操作 （使用计算机语言 检索、插入、删除、修改）	读者对图书馆的访问 （用普通语言 借书、还书）
9	第8项独立于第7项	第8项独立于第7项

现在，什么是数据库就比较容易理解了。数据库系统是有组织地、动态地存贮大量关联数据，方便多用户访问的计算机软硬件组成的系统，它与文件系统的重要区别是数据的结构化、独立性及共享性。

§1.2 数据库系统的产生和发展

数据库系统是从早期的人工数据管理开始逐渐演化而成的。

一、早期的人工数据管理

在50年代中期以前的计算机系统中，没有专门的软件来管理数据。因为当时的计算机主要是用于科学计算，而不是数据处理。科学计算任务一般计算比较复杂，计算量大，而数据量相对就比较小。当时一般是一个程序对应一组数据。编写程序的程序员不仅要安排数据的逻辑结构，同时还必须在编程序时设计好数据的物理结构、存取方式和输出输入等方式。

所谓数据的逻辑结构是指数据的内容、功能及使用方式，而数据的物理结构则需考虑数据存放的具体位置，数据与数据之间的存放关系等等。在这个时期的人工数据管理中，程序员除了设计程序外，还必须花费大量的精力来具体安排数据的物理位置。并且由于数据和程序是连成一体的，因此只要数据在存贮上有所改变，则对应的程序也必须修改。

二、文件系统

60年代初期，在数据库这个术语还鲜为人知时，实际上已经出现了数据库管理方法的雏形，即文件系统。

（一）逻辑结构与物理结构一致的文件系统

早期的文件系统是按顺序方式组织的。它的外存物理结构与用户观点的逻辑结构是完全一致的，其主要功能是完成 I/O 设备的输入输出操作，如图1-2所示。这种文件只适用于批

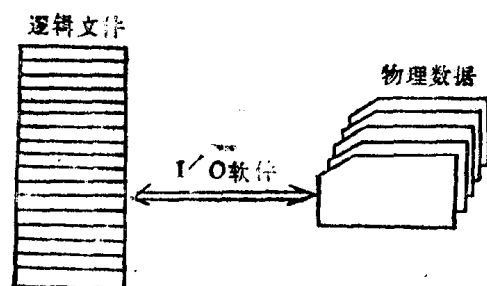


图1-2 早期的文件系统

处理的情况，而且为了更新一个文件就得重写一个新文件，同时原来的文件也被保留下来，称为父亲。因此同一个文件往往存有多份副本。那时候，应用程序员仍需要具体设计数据的物理结构，然后再把它们嵌入应用程序，因此数据与应用程序仍不具有独立性。如果数据组织或存贮装置稍有变动，应用程序员还是必须相应地修改程序，然后重新编译，重新调试。这时候的数据通常是作为一种应用设计的，如果其它应用也要使用相同的数据，但应用形式上可能有所不同，就需要创建几个不同的文件，从而产生了很高的冗余性。

在这个阶段也开始使用了一些随机存取文件，它允许用户随机地存取数据，而不必顺序地扫描整个文件。但是应用程序员在编写程序时必须规定具体的寻址方法，如果存贮设备发生改变，这部分包括寻址方式的程序也必须随之改变。此时的数据存贮方式仍属于初级数据文件方式。

(二) 逻辑结构与物理结构不同的文件系统

到60年代后期，磁盘和磁鼓等联机直接存取的外存设备性能有了较大的改变，从而使得这时的文件系统的物理结构和逻辑结构也开始有所区别，如图1-3所示。这种方法试图把应用程

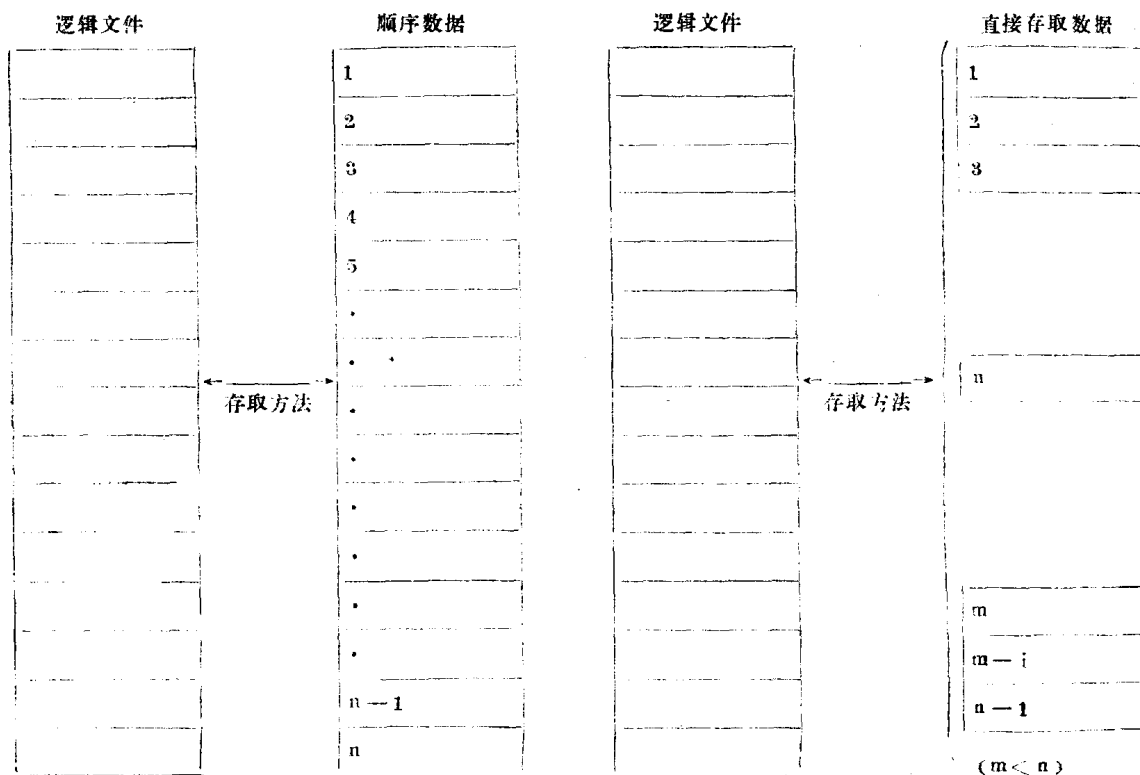


图1-3 按一定存取方法组织的文件系统

序与硬件变化的影响隔离开来，即数据的物理结构的改变不影响数据的逻辑结构。这时的文件存取方式提供了对记录顺序存取或随机存取的手段。显然，它不但适应批处理，也可以用于实时联机处理，并且系统更换外存也不影响应用程序，在某种程度上它也可以实现以文件为单位的数据共享，但这种文件系统仍然不够理想，因为同初级数据文件相类似，这种数据文件一般是为一种应用或紧密相关的应用设计的。

例如，一个公司的合同管理业务需要一组文件，程序员设计这些文件时，把合同管理中所需的数据集合包括在里面，文件的优化也尽可能为这项特殊的业务服务。如果这个公司中应付帐单业务需要许多与合同管理相同的信息，这时可以专门设计一个应付帐文件，也可以使用已做好的合同管理文件。但由于合同文件缺少一些应付帐单业务需要的数据，并且应付帐单业务使用合同文件中的数据总不如合同管理业务使用这些数据那么方便，因此一般总是设计一个专门的新文件。这样必然导致数据的大量冗余。并且，这时的文件系统也不便于增删新旧数据。这些亟待解决的问题，促使人们去研究一种新的更有效的数据管理技术。

三、早期的数据库系统

60年代末、70年代初，终于出现了初期的数据库系统，如图1-4所示。这时的数据系统具有如下特点：（1）对用户观点的数据进行严格细致的描述，使文件记录中数据项之间具有清晰的联系和简单的结构；（2）允许用户以记录或数据项为单位进行访问，允许文件之间的交叉访问；（3）数据的物理存储可以很复杂，同样的物理数据可以导致多个不同的逻辑文件，用户根据简单的逻辑结构来操作而不必考虑数据的存储情况，如果数据的物理位置和存储结构有所改动，用户的应用程序无需作相应的修改或重写，即用户逻辑数据与它们的物理存储之间的转换由数据管理软件来完成（见图1-4），这样就解决了数据的应用独立于数据的存储问题。

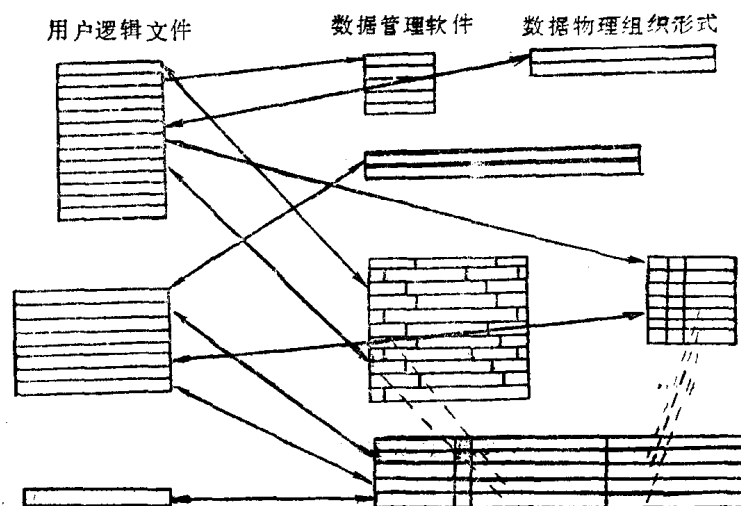


图1-4 早期的数据库系统

初期数据库系统的数据整体逻辑结构是用户逻辑文件的简单并集，随着用户的增多，数据库的整体逻辑结构也日趋杂乱。为了提高效率，常常需要对数据作一些增删，这样，数据的整体逻辑结构就会有所改变，从而导致用户逻辑文件的修改，进而导致用户应用文件的修