



万水最新数据库应用技术精通指南丛书

# 数据库原理及应用系统开发

张健沛 主编

张健沛 编著

中国水利水电出版社

## 内 容 提 要

本书系统地阐述了数据库系统的基本概念、基本原理和基本理论，并较为详细地介绍了ORACLE关系数据库系统和数据库应用系统的设计方法。第一章~第四章主要以关系数据库为主线介绍数据库原理。包括：数据库系统导论、关系数据库、关系数据库设计理论、查询优化；第五章为数据库管理系统及其实现技术；第六章介绍数据库保护；第七章介绍ORACLE关系数据库系统；第八章、第九章介绍数据库设计和数据库应用系统的设计开发方法；第十章介绍数据库系统的研究与发展。

本书理论与应用并重，概念清楚，系统性强，内容丰富，力图反映数据库领域的最新研究成果。

本书可作为高等院校计算机各专业数据库课程的教材，也可供从事计算机软件工作的科技人员及其他有关人员参考。

### 图书在版编目(CIP)数据

数据库原理及应用系统开发/张健沛编著. -北京:中国水利水电出版社, 1999.4

(万水最新数据库应用技术精通指南丛书)

ISBN 7-80124-764-7

I. 数… II. 张… III. 数据库系统-理论 IV. TP311.13

中国版本图书馆CIP数据核字(1999)第09178号

书 名	数据库原理及应用系统开发
作 者	张健沛 编著
出版、发行	中国水利水电出版社(北京市三里河路6号 100044) 网址: www.waterpub.com.cn E-mail: sale@waterpub.com.cn 电话: (010)63202266(总机)、68331835(发行部)
经 售	全国各地新华书店
排 版	北京万水电子信息有限公司
印 刷	北京牛山世兴印刷厂印刷
规 格	787×1092毫米 16开本 23.5印张 542千字
版 次	1999年4月北京第一版 1999年4月北京第一次印刷
印 数	0001—5000册
定 价	30.00元

凡购买我社图书，如有缺页、倒页、脱页者，本社发行部负责调换

版权所有·翻版必究

## 《万水最新数据库应用技术精通指南丛书》编委会

主 编 张健沛

委 员 刘大昕 于之硕 杨 静 王凤儒 宋克欧

马光胜 刘 群 刘远生 刘 江 马宗舜

## 《万水最新数据库应用技术精通指南丛书》总序

数据库技术是 60 年代末作为数据管理的最新技术登上数据处理舞台的。近三十年来，数据库技术得到迅速的发展，现在已经成为计算机软件的一个重要分支。

数据库系统的推广使用，使得计算机应用深入到工农业生产、商业、行政管理、科学研究、国防建设、工程技术的各个领域。当今的管理信息系统、办公自动化技术、计算机集成制造系统（CIMS）和知识库系统等，大都离不开数据库技术。近年来，我国在邮电、金融、电力能源、铁路交通、气象、财税、情报、军事、公安等行业装备的大型计算机系统都是以数据库为基础的。因此，对数据库系统原理、数据库应用开发方法、目前流行的典型数据库管理系统进行比较全面的介绍是国内各个行业的共同需要。

本丛书是为了适应国内广大计算机用户学习、使用数据库技术，为推动数据库技术在国内更加深入、更加广泛地应用，加快我国各个行业信息化的进程而编写的。本丛书编写的指导思想是突出实用技术、突出流行系统、突出理论与实践相结合。在本丛书中既有数据库系统原理、数据库应用系统开发方法这样理论性较强的图书，力图提高数据库应用人员的理论水平，使数据库应用系统的开发能够在理论的指导下进行，从而提高驾驭数据库技术的能力和数据库技术的应用水平。又包括了介绍当今世界上作为商品出售、技术成熟、性能稳定、又广泛应用的各种数据库管理系统的图书。这里既有技术先进、性能完善且在国内已经得以应用的大型关系数据库产品如 Oracle、Sybase、Informix 等大型数据库管理系统，又有已经在国内广泛应用的微型机数据库管理系统如 FoxPro 的各种版本，还有随着 Windows 平台应用的日益广泛而产生的 Visual Foxpro、Access、SQL/Server 以及近年出现的新一代数据库应用快速开发工具 PowerBuilder、Delphi 等。

本丛书以国内各行业计算机应用人员为主要对象，强调以应用为出发点，以实际应用为目的，在写法和内容组织上尽量做到突出实用、通俗易懂，力求做到科学性、先进性和通俗性的完美统一。本丛书的作者都是高等学校中多年从事数据库系统教学和科研的专业人员，他们又都从事过数据库应用系统的开发。在本书的编写过程中，我们本着突出实用、力图反映最新技术、注重风格统一、每本书都自成体系的原则，力求编写出适合国内特点、满足国内计算机用户需要的数据库技术丛书，为我国数据库技术的应用和普及作出贡献。

丛书主编张健沛

# 前 言

数据库技术自六十年代后期出现以后，一直活跃在数据处理领域。二十多年来，数据库技术得到了迅速的发展，已经成为计算机科学的一个重要分支。由于数据库技术已经成为开发各种信息系统的重要工具，因它的推广使用，使得计算机应用深入到工农业生产、商业管理、办公自动化、科学研究和工程技术的各个领域。随着社会各个领域的迫切需要和广泛应用，人们对数据库技术重要性的认识越来越明确，对它的应用和学习要求也越来越迫切。

本书是作者多年从事数据库系统教学和科研的基础上编写出来的。在编写过程中，本着由易到难，循序渐进，理论与实践并重，突出重点的原则，力求做到概念清楚，深入浅出，理论联系实际，面向应用。全书结构合理，选材精炼，重点突出应用最广泛的关系数据库系统，对它的概念、原理作了较为深入的讨论。为了突出实际应用，本书增加了数据库设计和数据库应用系统设计与开发的内容，并且还较为详细地介绍了目前国内外较为流行的ORACLE关系数据库，以使学生在学习过程中能更好地把理论与实际结合起来。从而达到更好的学习效果。另外，在内容选材上，我们还注意尽量反映这一领域的新方法，新技术，以使学生能对数据库领域的前沿动态能有初步的了解。

全书共分十章，第一章介绍数据库系统的基本概念，基本原理。第二章至第四章介绍当前最流行、应用最广泛、理论研究最多的关系数据库系统，对其基本概念、关系代数理论、查询语言、规范化理论、查询优化等都作了充分的介绍。第五章介绍数据库管理系统及其实现技术。第六章介绍数据库保护技术。第七章介绍ORACLE关系数据库系统。第八、九章介绍数据库设计和数据库应用系统的开发与设计，以期指导应用。第十章介绍数据库系统的研究与发展，力图反映数据库领域的最新研究成果。

本书的撰写工作得到了黑龙江省自然科学基金的资助。在本书的编写过程中还得到了哈尔滨建筑大学杨静副教授的帮助，在此表示感谢。由于水平有限，加之时间匆忙，书中难免有错误和不妥之处，敬请读者批评指正。

张健沛

1999年元月于哈尔滨

# 目 录

总序

前言

<b>第一章 数据库系统导论</b> .....	1
1.1 引言 .....	1
1.1.1 数据与数据处理 .....	1
1.1.2 数据管理技术的发展过程 .....	3
1.1.3 什么是数据库 .....	5
1.1.4 数据库系统的主要特征 .....	8
1.2 现实世界的描述 .....	12
1.2.1 现实世界 .....	12
1.2.2 信息世界 .....	12
1.2.3 计算机世界 .....	12
1.2.4 常用术语的说明 .....	13
1.3 数据模型 .....	14
1.3.1 层次模型 .....	14
1.3.2 网状模型 .....	15
1.3.3 关系模型 .....	16
1.3.4 对传统数据模型的评价 .....	18
1.3.5 E-R 数据模型 .....	20
1.3.6 面向对象数据模型 .....	25
1.4 数据库系统结构 .....	30
1.4.1 数据库系统结构的一些概念 .....	30
1.4.2 数据库系统的组成 .....	32
1.4.3 数据库管理系统 .....	34
1.4.4 数据库语言 .....	36
1.4.5 数据库管理员 .....	39
1.4.6 数据字典 .....	40
1.4.7 数据库系统的工作流程 .....	42
<b>第二章 关系数据库</b> .....	45
2.1 关系数据库的基本概念 .....	45
2.1.1 笛卡儿乘积 .....	45
2.1.2 关系 .....	46

2.1.3	关键字 .....	46
2.1.4	关系模式 .....	47
2.1.5	关系模型 .....	47
2.1.6	关系数据库 .....	47
2.2	关系模型及其描述 .....	48
2.2.1	关系模型 .....	48
2.2.2	关系模式 .....	49
2.2.3	关系子模式 .....	49
2.2.4	关系模型的完整性规则 .....	50
2.3	关系数据库语言 .....	51
2.3.1	关系代数 .....	52
2.3.2	关系演算语言 .....	61
2.3.3	QBE 语言 .....	66
2.3.4	关系数据库语言的评价 .....	68
<b>第三章</b>	<b>关系数据库设计理论 .....</b>	<b>70</b>
3.1	引言 .....	70
3.2	关系模式的规范化 .....	72
3.2.1	函数依赖 .....	72
3.2.2	第一范式 .....	75
3.2.3	第二范式 .....	75
3.2.4	第三范式 .....	77
3.2.5	BCNF .....	78
3.3	函数依赖的公理系统 .....	79
3.3.1	引言 .....	79
3.3.2	Armstrong 公理 .....	80
3.3.3	函数依赖集的等价与覆盖 .....	86
3.4	关系模式的分解 .....	89
3.4.1	模式分解的等价问题 .....	89
3.4.2	无损联接分解 .....	90
3.4.3	分解的函数依赖保持性 .....	96
3.4.4	模式分解算法 .....	98
3.5	多值依赖和关系的 4NF .....	103
3.5.1	多值依赖定义 .....	103
3.5.2	多值依赖公理 .....	105
3.5.3	第四范式 (4NF) .....	106
3.6	联接依赖和关系的 5NF .....	108
3.6.1	联接依赖 .....	108



3.6.2	第五范式 .....	110
<b>第四章</b>	<b>优化查询 .....</b>	<b>113</b>
4.1	查询优化概述 .....	113
4.2	查询优化的一般策略.....	114
4.3	基于关系代数表达式的优化算法.....	115
4.3.1	关系代数表达式的等价变换规则.....	115
4.3.2	关系代数表达式的优化算法.....	116
4.4	分解查询的优化方法.....	119
4.4.1	分解处理 .....	119
4.4.2	结局处理 .....	124
4.5	连接运算的优化 .....	124
4.5.1	实现二元连接运算的几种常用方法.....	125
4.5.2	n 元连接的优化.....	127
4.6	选择和投影操作的实现和优化.....	134
4.6.1	选择操作的实现和优化.....	134
4.6.2	投影操作的实现.....	136
4.6.3	集合操作的实现.....	137
4.6.4	组合操作 .....	138
<b>第五章</b>	<b>数据库管理系统及其实现技术.....</b>	<b>140</b>
5.1	数据库管理系统概述.....	140
5.1.1	数据库管理系统结构简介.....	140
5.1.2	数据库管理系统的目标.....	141
5.1.3	数据库管理系统的基本功能.....	142
5.1.4	数据库管理系统与操作系统的关系.....	143
5.2	数据库管理系统系统结构.....	144
5.2.1	事务 .....	144
5.2.2	DBMS 程序模块的组成 .....	145
5.2.3	DBMS 的进程结构 .....	147
5.2.4	DBMS 的系统结构 .....	149
5.3	数据库语言的处理 .....	152
5.3.1	语言翻译处理层的任务和工作步骤.....	152
5.3.2	解释方法 .....	155
5.3.3	预编译方法.....	155
5.4	数据存取层 .....	157
5.4.1	存取层的系统结构.....	157
5.4.2	存取层的功能子系统.....	158
5.6	缓冲区管理 .....	160

5.7	数据库物理组织 .....	161
5.7.1	数据库物理组织概述.....	161
5.7.2	数据库实现联系的方法.....	163
<b>第六章</b>	<b>数据库保护 .....</b>	<b>168</b>
6.1	引言 .....	168
6.2	数据库的安全性 .....	168
6.2.1	用户标识和鉴定.....	169
6.2.2	授权规则 .....	169
6.2.3	数据分级法.....	170
6.2.4	数据加密 .....	171
6.3	数据库的完整性 .....	172
6.3.1	完整性的破坏原因.....	172
6.3.2	完整性规则.....	172
6.4	数据库的并发控制 .....	173
6.4.1	排它型封锁和 P X 协议.....	175
6.4.2	可串行化、死锁和活锁.....	176
6.4.3	时间标志法.....	177
6.5	数据库的恢复 .....	179
6.5.1	数据库故障和数据库恢复.....	180
6.5.2	恢复方法 .....	180
6.5.3	如果出现了故障，有下面两种可能性： .....	181
6.6	数据库的再组织 .....	181
6.6.1	再组织的必要性.....	181
6.6.2	影响再组织点的因素.....	182
6.6.3	再组织点的选取.....	182
<b>第七章</b>	<b>ORACLE 关系数据库系统.....</b>	<b>183</b>
7.1	ORACLE 关系数据库系统概述.....	183
7.1.1	ORACLE 系统软件结构.....	183
7.1.2	ORACLE 产品结构.....	186
7.2	SQL*PLUS 语言 .....	188
7.2.1	概述 .....	188
7.2.2	数据定义 .....	189
7.2.3	数据操纵 .....	193
7.2.4	数据查询 .....	195
7.2.5	视图 .....	214
7.3	ORACLE 应用程序设计与动态 SQL.....	216
7.3.1	概述 .....	216

7.3.2	预编译接口.....	216
7.3.3	PRO*C.....	218
7.4	ORACLE 数据库管理.....	234
7.4.1	ORACLE 数据库存储组织.....	234
7.4.2	ORACLE 安全管理.....	240
7.4.3	ORACLE 数据库并发控制.....	248
7.4.4	ORACLE 数据库恢复.....	250
<b>第八章</b>	<b>数据库设计</b> .....	<b>255</b>
8.1	数据库设计概述.....	255
8.1.1	数据库设计的内容和特点.....	255
8.1.2	数据库设计方法简介.....	257
8.1.3	数据库设计步骤.....	259
8.2	需求分析.....	260
8.3	概念模型设计.....	262
8.3.1	设计局部 E-R 模型.....	264
8.3.2	将各局部 E-R 模型综合成总体 E-R 模型.....	264
8.4	逻辑结构设计.....	268
8.4.1	概念模型向网状模型转换.....	268
8.4.2	从 E-R 图向关系模型转换.....	270
8.4.3	规范化理论的应用.....	273
8.4.4	模型的评价与改进.....	274
8.5	数据库物理设计.....	275
8.5.1	数据簇集设计.....	276
8.5.2	索引的选择.....	279
8.6	数据库的实施与维护.....	282
8.6.1	数据库数据的装入.....	282
8.6.2	数据库的试运行.....	283
8.6.3	数据库的运行和维护.....	283
<b>第九章</b>	<b>数据库应用系统开发与设计</b> .....	<b>286</b>
9.1	数据库应用系统开发概述.....	286
9.1.1	数据库应用系统的特点.....	286
9.1.2	数据库应用系统的开发步骤.....	287
9.2	应用程序设计.....	289
9.2.1	应用程序设计步骤.....	289
9.2.2	数据流图.....	292
9.3	数据库应用系统设计实例.....	296
9.3.1	系统分析.....	296

9.3.2	系统设计 .....	298
<b>第十章</b>	<b>数据库系统的研究与发展 .....</b>	<b>329</b>
10.1	分布式数据库系统 .....	329
10.1.1	分布式数据库系统概述 .....	329
10.1.2	分布式数据库系统的组成 .....	334
10.2	客户/服务器结构的数据库系统 .....	339
10.2.1	概述 .....	339
10.2.2	客户/服务器结构数据库系统 .....	341
10.2.3	客户/服务器结构的数据库系统实现技术 .....	346
10.2.4	客户/服务器结构数据库系统的优点 .....	351
10.3	面向对象数据库系统 .....	352
10.3.1	传统的关系数据库系统的不足之外 .....	352
10.3.2	“对象”概念的实质 .....	353
10.3.3	面向对象数据库技术的产生 .....	354
10.3.4	面向对象模型 (object - oriented model) .....	354
10.3.5	面向对象数据库管理系统 (OODBMS) .....	355
10.4	一些新型的数据库系统 .....	355
10.4.1	演绎数据库 .....	355
10.4.2	模糊数据库 (fuzzy database) .....	357
10.4.3	模糊演绎数据库 .....	359
10.4.4	巨型数据库系统 .....	359
10.4.5	主动数据库 (active database) .....	359
10.4.6	多媒体数据库 (multimedia databases) .....	360
10.4.7	集成的工程数据库 (integrated engineering DB) .....	361
10.4.8	联邦数据库 (federal databases) .....	362
10.4.9	时态数据库 (temporal databases) .....	362
<b>参考文献</b>	.....	<b>364</b>

# 第一章 数据库系统导论

## 1.1 引言

### 1.1.1 数据与数据处理

我们在日常生活中接触到大量数据，如银行帐目数据、学校的教学管理数据、企业的生产管理和产品销售数据等等。这些数据都是人们生活中所离不开的，它们可以被视为人类社会中一种极其重要的资源。

然而这些数据之所以有价值，就是因为这些数据是表现信息，是载荷信息的物理符号。简单地说，信息是向人们提供关于现实世界新的事实的知识，它反映了客观事物的物理状态。例如，某一国家的汽车产量是 30 万辆，这个 30 万辆是一个数据，而它又表示了这个国家生产汽车的能力这一信息。由此可见，信息是用数据来表示的，但二者是不可分离而又有一定区别的概念。一方面并非任何数据都能表示信息，信息只是消化了的数据，信息是依赖于数据而存在的；另一方面，信息是更基本的、直接反映现实的概念，数据是信息的具体表现。

对于数据进行综合推导，得出新的数据，这些新的数据则表示了新的信息。例如，在商品生产中，生产管理者根据某种商品历年销售数量及最近的市场需求调查，获得了许多数据。再对这些数据进行加工，就得出了有关商品的市场预测信息。生产管理者就可以根据这些信息进行分析和评价，作出对该产品是增产、减产还是停产的决策，并指挥生产部门的生产活动。这种从收集数据、到加工成信息、进行评价和决策、再指挥实践活动，从而产生新的数据的循环过程称为信息循环。

诸如此类围绕信息所做的一系列工作，称为信息处理。一般地说，信息处理就是指信息的收集整理、加工、存储和传播等一系列活动的总和。因为信息是用数据表示的，所以对信息的处理又具体地体现在对数据的处理上。通过对原始数据（即尚未作评价的各种信息）的处理，产生新的数据（新的信息），作为结果。这一处理，包括对数据的收集、记录、分类、排序、存储、计算 / 加工、传输、制表和递交等处理，这就是数据处理的概念。经过处理的数据是精炼的数据，是能够反映事物或现象的本质和特征及其内在联系的数据。

数据处理的历史可以追溯到远古时代。原始人类的结绳记事，累石记数便是数据处理的雏形。随着社会生产和文明的日益发展，信息的概念就趋于越来越复杂和深化，信息已经支配着人类的整个社会活动，所以有人把现代社会称为信息社会。研究信息的形态、传输、处理和存储理论的信息科学也就应运而生。随着信息概念的深化和发展，数据处理也就得到了相应的发展，就整个数据处理的发展过程来看，大致可以分成三个阶段。

第一阶段是数据的手工处理阶段。这是自远古时代到 19 世纪末一直使用的数据处理方法。由于当时的社会生产力低下，科学技术不发达，因此，数据处理的方法也就保持着简单、低级的手工操作。人们只能借助于如算盘、手摇计算机、计算尺等低级的计算工具处理数据。这种方法的效率低，能处理的数量少，而且可靠性差。随着生产力和科学技术的发展，依靠手工操作来处理数据的方法，已经不能适应社会的需要，因而人们就着手计算工具的改革和发明创造。

第二阶段是数据的机械处理阶段。这一阶段的主要特征是使用了比第一阶段先进得多且比较有效的工具。例如 H·霍勒内斯在 19 世纪 80 年代发明了卡片制表机。这种机器能以自动方式进行卡片的穿孔、校验、分类、整理和制表等操作。这是跨入机械数据处理阶段的一个重要标志。机械设备的使用，使数据处理的能力有了很大的改进和提高。但是这种设备的性能和使用，仍然受着很大的限制。这一阶段仍然有着较多的手工操作。直到 20 世纪 40 年代中期，第一台电子计算机 ENIAC 问世以后，数据处理才进入了自动化的电子数据处理时代。

第三阶段是电子数据处理阶段。电子计算机的出现，为数据处理展现了广阔的前景。它不仅处理速度快、存储容量大、输入输出灵活，而且把人的手工操作降低到最小程度。电子数据处理方法不仅适应了不断提高的社会生产力的需要，而且给社会生产力的发展以推动和促进。因此，电子计算机的应用是数据处理领域中的一场革命。

尽管这三种数据处理的方法一个比一个先进，但是目前世界上仍然存在着手工和机械处理方法，特别是手工处理方法还是屡见不鲜的。而机械方法相对来说处于停顿了。这一方面是因为某些领域的某些部门还不适宜使用电子计算机，另一方面是因为世界上各个国家和地区的发展还不平衡。这种局面将随着超小型和微型计算机的发展而得到改善。

人类活动的整个历史，离不开对信息和数据的收集、保存和处理。像银行、统计、仓库等部门的工作人员，无时无刻不在和大量的数据打交道。然而，随着社会的不断进步，人类社会积累的信息，正以“几何级数”的速度增长。因此人们过去传统的、落后的数据处理方法，已经远远适应不了形势发展的需要了。所以长期以来，如何妥善地保存和管理数据，一直是人们十分关注的课题。人们对数据处理现代化的要求日益迫切。

电子计算机的出现，为人类处理数据自动化提供了可靠的物质基础。开始时，计算机只是单纯为科学计算服务，所以人们又把它称为数字计算机。随着计算机硬件和软件的发展，计算机硬件成本的不断降低和各种新型外部设备的问世，特别是大容量磁盘存储器的生产和应用，人们逐渐地认识到，除了科学计算之外，计算机还具有强大的处理数据的能力，而且计算机的外存储器还是保存数据资料的理想场所。于是，计算机就不单纯地应用于科学计算，而且还在数据处理这个广阔的领域大显身手。到了 60 年代后期，数据库技术出现了。数据库技术研究对象，就是高效地、巧妙地进行数据处理，而又开销最小。

目前，在整个计算机应用领域中，数据处理还是最庞大、最活跃的应用领域。在发达国家中，数据处理约占 80%~90%。计算机工业历史说明，电子计算机是在科学计算领域中诞生，却在数据处理领域中站稳了脚跟，找到了真正的立足点。而数据库系统是数据处理的核心机构，数据库管理系统的最新成果，反映了数据处理技术的最新水平。C·J·Date

下面这段话，说明了数据库系统的地位。“数据库管理是现代计算机系统提供的最重要的功能。事实上，其重要性已经达到了这样的程度，即它已普遍成为购买计算机的主要出发点”。

一个计算机化的电子数据处理系统（Electronic Date Processing System，简称 EDP 系统），包括了人和电子计算机系统。它执行输入、处理、存储、输出和控制功能，用于完成对数据的一系列操作过程。一个计算机化的电子数据处理系统，一般由三个基本系统组成，它们是电传处理系统、数据管理系统、科学计算系统。如图 1-1 所示。

电传处理系统的基本功能是，在适当的地点和时间提供必要的的数据，即由它完成数据提供、通信（传输）和传递。

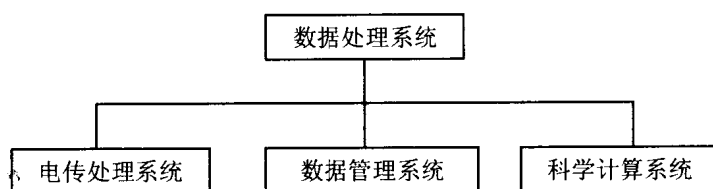


图 1-1 数据处理系统

数据管理系统完成对数据的组织、编目、定位、存储、检索和维护等功能。

科学计算系统作为数据处理系统的一部分，会大大地加强系统的处理能力。它强调使用推理过程产生信息，好像纸和笔一样，成为人们在分析问题时的有力工具。但是有些复杂的推理问题，电传处理系统 数据管理系统科学计算系统单靠纸和笔是不能解决的，而要靠科学计算系统来解决，因此它有助于提高人们的逻辑推理和通信的功能。例如在一个企业中，它能够用来解决某些管理决策、计划编制和资源最优分配等问题。

由于数据处理系统的核心是数据，因此数据管理系统是计算机数据处理系统中最重要的组成部分，而数据库管理系统则是一种新的数据管理软件，它提供了一种新的数据管理方法。

### 1.1.2 数据管理技术的发展过程

数据以及数据之间的关系可以从逻辑的和物理的两方面进行描述和组织。数据的逻辑描述是指按用户要求的数据间逻辑关系来组织和表达数据的，而物理描述则将涉及数据在计算机内的存储方式，是以存储数据的时空概念来描述数据间的关系的，前者可称为数据的逻辑结构，后者则称为数据的物理结构。

如何用计算机对数据进行组织、编目、定位、存储、检索和维护是数据管理的任务。数据管理是数据处理的中心问题。数据管理技术与数据处理方式有密切的关系，并且直接影响着数据处理的效率。在数据处理的手工处理阶段和机械处理阶段，数据管理技术也是手工的。但是，在计算机进入数据处理领域之后，原来的那套手工管理方式就不能适应计算机自动处理数据的需要了。而且，在跨入信息社会的今天，需要管理的数据越来越多，

如果仍然沿用原来那套方式，不仅需要耗费大量的人力，而且很难使这些数据发挥应有的作用。为此，许多计算机专家，特别是软件工作者，对数据管理技术做了大量的研究工作，使数据管理技术经历了一个由简单到复杂的不断完善的发展过程。

#### 1.1.2.1 人工管理阶段（50年代中期以前）

在这一阶段中，计算机除了硬设备之外没有任何软件可供使用，放在用户面前的只有硬件的裸机。用户在这种裸机上进行数据处理，除了编制自己的课题程序外，还必须考虑数据的逻辑定义和组织，以及数据在计算机存储设备内的物理存储方式和地址。数据的引用是按物理地址进行的。这时，表示处理流程的程序和作为处理对象的数据，相互结合成一个整体。数据的管理、基本上是手工的、分散的，计算机还没有在数据管理中发挥应有的作用。因此，严重地影响了计算机的使用效率。这一时期数据与程序之间的关系可以用图 1-2 表示。

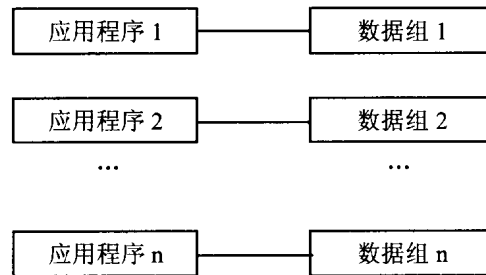


图 1-2 人工管理阶段数据和程序之间的关系

#### 1.1.2.2 文件系统阶段（50年代后期至60年代中期）

把数据组织成文件的形式，是计算机数据管理的重大进展，也是计算机在数据管理中直接发挥作用的开始。众所周知，无组织的一大堆数据是毫无意义的，也是没有实用价值的。该软件将数据按照一定的规则组织起来，成为一个有效的数据组合体，赋予它一个名字，称为文件名或文件标识，用户通过文件名来访问该文件。这时用户就不必过多地考虑物理细节了，而是由文件系统充当应用程序和数据文件之间的接口，这一时期应用程序和数据的关系如图 1-3 所示。

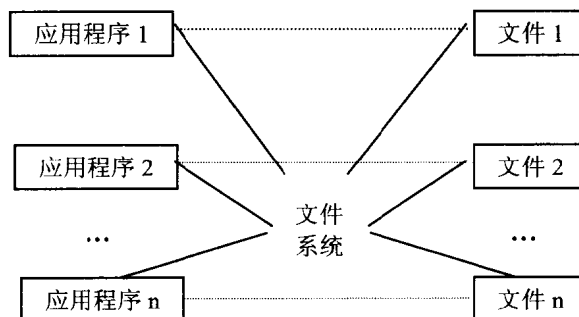


图 1-3 文件系统阶段数据和程序之间的关系



虽然这一阶段较人工管理阶段有了很大的改进，但仍存在很大的弱点，即文件本身还是基本上对应于一个或几个应用程序，或者说数据还是面向应用的。从逻辑结构上讲，还与前一阶段没有什么变化，尽管程序不必要直打接与文件交道而有软件做为接口，但它仍然是一个不具有弹性的无结构的信息集合，存在着冗余度大，空间浪费、文件不易扩充等缺点，从而反映了不了现实世界事物之间的广泛的内在联系。

### 1.1.2.3 数据库系统阶段（60年代后期开始）

在这一时期，磁盘技术已经成熟，并作为主要外存而广泛使用。计算机硬件的价格大幅度下降，可靠性增强，为数据管理技术的发展奠定了物质基础。另外，计算机用于管理，其规模更加庞大，从而使数据量急剧增加。对数据进行集中控制，充分提供数据共享的要求日益迫切。

在这样的背景下产生了一种新的数据管理技术即数据库技术。数据库技术克服了以前所有管理方式的缺点，试图提供一种完善的、更高级的数据管理方式。它的基本思想是解决多用户数据共享的问题，实现对数据的集中统一管理，具有较高的数据独立性，并为数据提供各种保护措施。这一时期，数据库管理软件作为用户与数据的接口，程序和数据的的关系如图 1-4 所示。

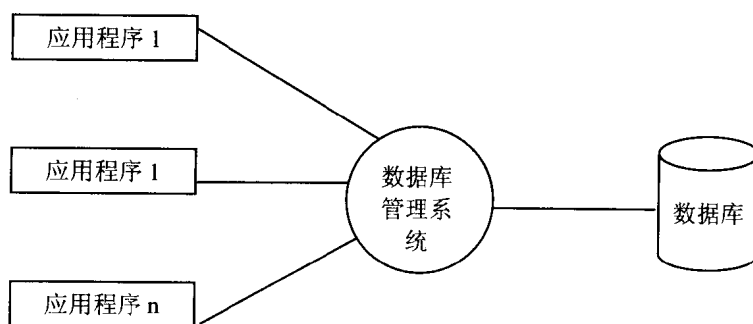


图 1-4 数据库阶段数据与程序之间的关系

### 1.1.3 什么是数据库

数据库系统是计算机软件的一个重要分支，是近 20 年来发展起来的一门新兴学科。它和计算机网络、人工智能被称为当今计算机技术界的三大热门技术。目前，虽有人对数据库的设计原则和方法进行总结和探讨，使之通用化、标准化和理论化。但总的说来，它的概念、原理和方法还处于从工程实践向理论过渡的阶段，它的概念，原理和方法还在继续变化和发展。另外，数据库是一个很复杂的系统，涉及面很广，难以用简练的语言准确地概括其全部特征。因此，给它下一个确切的定义是困难的，所以我们先从简单分析入手，逐步认识什么是数据库。

库，这个名词在日常生活中经常接触到。我们知道各种各样的库。例如：仓库、书库、金库、血库等。稍微留心一下，我们就会发现，这些库都具有这样几个特点：