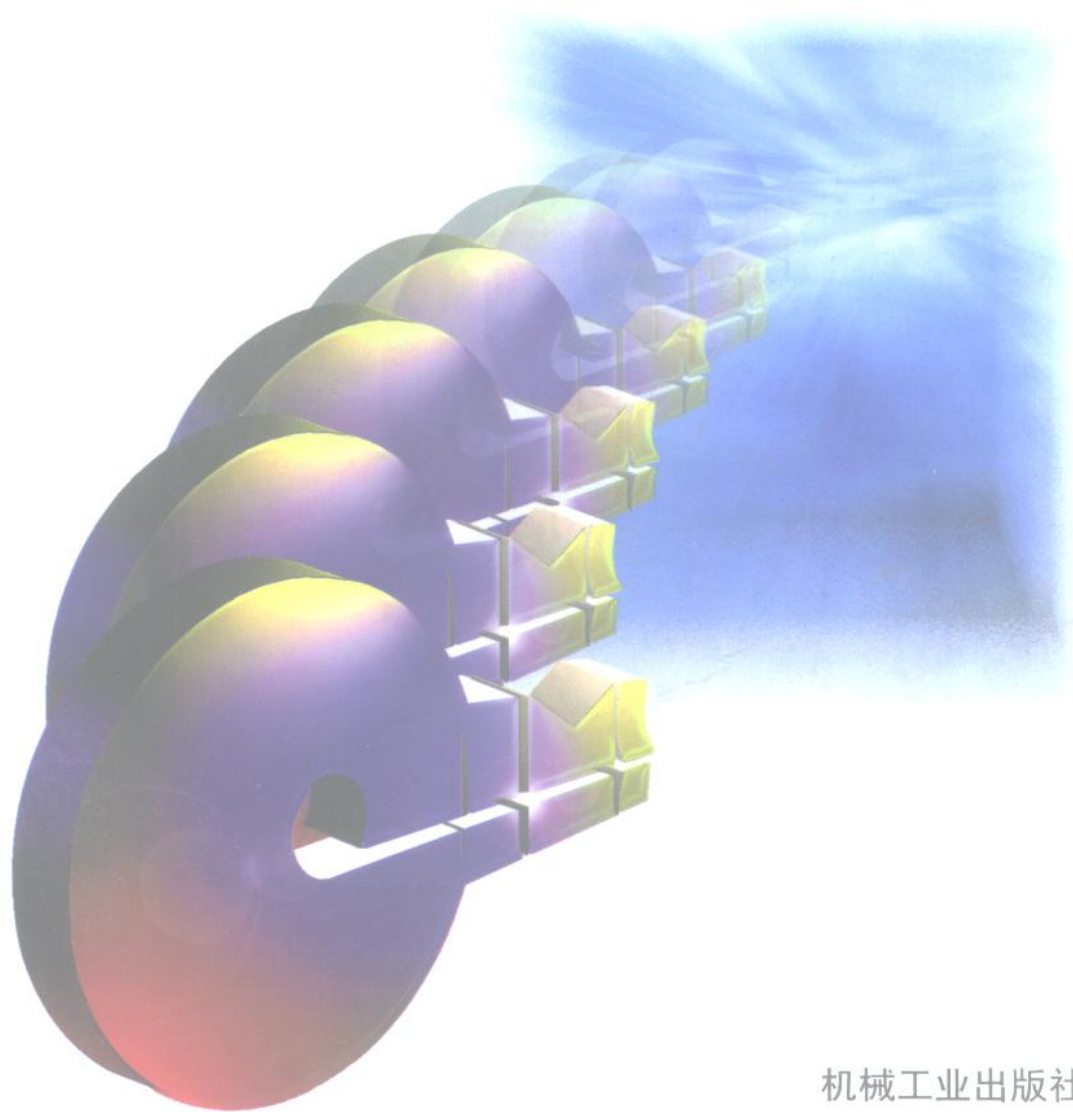




CIMS系列培训教材

CIMS 中的数据库技术

董逸生 等 编著



机械工业出版社

163
78

CIMS 系列培训教材

CIMS 中的数据库技术

董逸生 王茜 姜浩 编著



机械工业出版社

(京)新登字 054 号

内 容 简 介

CIMS—计算机集成制造系统是加强制造技术、改造传统制造业、支持新兴制造业、提高企业竞争能力的一种高技术。CIMS 包括了经营决策、生产计划、产品设计、制造与控制等功能子系统,数据库技术是 CIMS 中实现集成的主要手段。本书介绍了 CIMS 对数据库技术的要求、有关数据库的基本概念、数据库设计的方法和技术、数据库应用系统的体系结构及数据库的集成技术。通过本书可使读者对 CIMS 中的数据库技术有一个全面的了解。

图书在版编目 (CIP) 数据

CIMS 中的数据库技术/董逸生主编. —北京:机械工业出版社, 1996. 12

CIMS 系列培训教材

ISBN 7-111-05497-0

I. C… II. 董… III. 计算机集成制造-数据库-技术培训-教材 N. TP392

中国版本图书馆 CIP 数据核字 (97) 第 02013 号

出 版 人: 马九荣 (北京市百万庄南街 1 号 邮政编码 100037)

责任编辑: 刘 岗 版式设计: 李松山 责任校对: 卢锦宝

封面设计: 姚 毅 责任印制: 赵永洪

林业大学出版社印刷厂印制·新华书店北京发行所发行

1997 年 4 月第 1 版·1997 年 4 月第 1 次印刷

787mm×1092mm 1/16·12.125 印张·301 千字

0 001—5000 册

定价: 21.00 元

编辑委员会

主任委员：吴 澄

副主任委员：周 济

委 员：田连会 刘 飞 祁国宁 孙家广
张申生 李芳芸 李伯虎 李美莺
娄勤俭 柴天佑 顾冠群 徐晓飞
崔德刚 曾庆宏 穆世诚 薛劲松
戴国忠

序

这一套《CIMS 系列培训教材》是参加国家高技术研究和发展计划（863 计划）的科技人员经过 10 年研究和应用实践之后奉献给广大读者的。

二十世纪世界的一个重大变革是形成了一个统一的全球市场，每一个国家都不可能离开这个全球市场求得自身的发展。统一的全球市场形成的直接后果是市场竞争更加激烈，表现为产品更新换代加快、质量更好、价格更便宜、产品的技术含量更高，并且围绕产品的服务也越做越好。这就要求企业（制造商）能尽快响应市场的变化，制造出性能价格比优良的、满足用户各种要求的产品，并提供良好的服务。对于技术含量高的产品，竞争更为激烈，因为技术含量高就有可能取得产品的市场独占性，从而获取高利润。

我们把市场对产品的压力归结为：时间 T （即开发新产品的时间或成熟产品的上市时间），质量 Q ，成本 C 和服务 S 。 T 、 Q 、 C 、 S 是制造业的一个永恒主题，企业不断完善 T 、 Q 、 C 、 S 是一个永无止境的过程，对各国企业都一样，但我国企业的压力大得多。

国家 863 计划的 CIMS 主题，以促进我国企业的信息化、现代化为宗旨；近 10 年来，以提高企业市场竞争力、提高企业效益为目标，以信息技术、现代管理技术改造国有企业、支持新兴企业为切入点，用高技术解决我国企业的难点和热点问题。

CIMS，即计算机集成制造系统（Computer Integrated Manufacturing Systems）是用信息技术（包括计算机技术、自动化技术、通信技术等）和现代管理技术，加强制造技术，改造传统制造业，支持新兴制造业，提高企业市场竞争能力的一种高技术。具体地说，以企业的产品为龙头，在产品的设计过程、管理决策过程、加工制造过程、产品的质量管理与控制过程等方方面面，采用各种计算机辅助技术和先进的科学管理方法，使企业优化运行，达到产品上市快、质量好、成本低、服务好，从而提高企业的效益和市场竞争能力。

立足国情，按照企业的需求开展对 CIMS 的研究和关键技术攻关，是应用技术研究和一个重要指导思想。它也使我国的 CIMS 得到了长足的进展和国际同行的关注。

1994年,清华大学国家CIMS工程研究中心获得了美国SME(制造工程师学会)的CIMS“大学领先奖”,这标志着我国CIMS的研究水平进入国际先进行列。1995年,北京第一机床厂荣获SME的CIMS“工业领先奖”,这标志着我国一些试点企业的CIMS应用达到国际先进水平。“大学领先奖”和“工业领先奖”一般每年在世界范围内各评选一名。中国已经成为除美国外唯一获得此两项大奖的国家。

更加重要的是,CIMS的应用在我国取得了显著的经济效益和社会效益。成都飞机公司、沈阳鼓风机厂、北京第一机床厂、东风汽车集团模具厂、山西经纬纺机厂和杭州三联电子有限公司等不同的大、中、小企业是其中的代表。当前,CIMS的进一步试点推广应用已扩展到机械、电子、航空、航天、轻工、纺织、石油、化工、冶金、通信、煤炭等行业的60多家企业。CIMS在我国正面临着一个发展的好势头。

进一步的推广应用,并且能取得预期的效益,关键之一在于人才,在于培训。这一套教材正是为这目的而写的。我们希望它们能为我国制造业的发展,“圆我工业强国之梦”作出贡献。

由于参加编写的作者都是第一线的科技工作者,任务繁忙,时间不足,加上编写的经验不够以及学术上的不足,使这套CIMS教材难免有错误和不足。我们愿意和广大读者一起,使之精益求精。

国家863计划CIMS主题专家组组长

吴 澄

一九九六年十月

前言

数据库技术是 CIMS 中实现集成的主要手段。CIMS 是一个十分复杂的集成系统,它包括了经营决策、生产计划、产品设计、制造与控制等功能子系统,所以 CIMS 的数据库技术,要能支持大量事务数据的组织与管理、支持产品设计与生产控制数据的组织与管理,也要能支持各应用领域的数据的集成,通过数据的集成实现各个功能子系统的集成。另外, CIMS 是一个不断发展的系统,其功能及所采用的技术将随着应用的扩大以及技术本身的进步会不断地扩充更新,因而也要求数据库系统能适应这种发展的要求,具有良好的开放性及灵活的可扩充性。作为一门知识, CIMS 中的数据库技术涉及各类数据库系统的基本概念、功能原理及组成结构,涉及数据库的设计、数据库系统的体系结构以及集成技术等。基于以上的理解,本书是这样组织的:

第 1 章为数据库系统概论,概略地介绍了数据库技术的发展史,阐述了数据库系统的特征、结构以及数据模型和数据模式等概念,较为详细地介绍了 SQL 数据库语言,简要地叙述了 DBMS 的功能和结构以及分布式数据库系统的概念。第 2 章是数据库设计,主要从实用的角度介绍了数据库设计的需求分析、概念设计、逻辑设计以及物理设计的有关概念和方法步骤,也阐述了新老数据库系统的集成问题以及分布式数据库的设计方法。第 3 章是 CIMS 中的数据集成技术,阐述了 CIMS 中数据集成的特点、实现集成的目标,重点叙述了数据集成的各种机制与技术,也介绍了工程数据管理的一般概念以及典型的工程数据管理系统产品的功能特点。第 4 章为数据库系统的客户/服务器(C/S)结构,阐述了 C/S 结构的基本概念、组成原理以及主要特点,较为详细地叙述了 C/S 结构各个组成成分的功能、作用机制以及某些典型产品的技术特点,最后介绍了 C/S 结构的设计与应用开发。

实事求是地说,到底怎样组织这本教材,我们的体会还很肤浅,想法也不太成熟,再加上时间的紧迫,所以编集在本书中的材料,无论在深度、广度还是在合理性、系统性等方面,一定存在不少问题,恳请广大读者批评指正。

本书的第 1 与第 3 章由王茜和吕云松编写,第 2 章由董逸生编写,第 4 章由姜浩编写,夏松涛参与了第 2 章的部分编写工作。在本书的编写过程中引用了书后有关文献中的材料或思想,谨向这些文献的作者表示谢忱。

全书由南京航空航天大学计算机系林钧海教授主审,对他提出的宝贵意见表示衷心感谢。

编 者
一九九六年九月

目 录

序

前言

第 1 章 数据库系统概论	1
1.1 什么是数据库系统	1
1.1.1 计算机数据管理技术的发展	1
1.1.2 数据库系统	3
1.1.3 数据库系统的主要特征	4
1.1.4 数据库系统的体系结构	5
1.2 数据模型与数据模式	6
1.2.1 数据模型	6
1.2.2 数据模式	6
1.2.3 常见的数据模型	6
1.3 数据库语言	19
1.3.1 SQL 语言概述	19
1.3.2 SQL 数据定义语言 (DDL)	20
1.3.3 SQL 数据查询语言	23
1.3.4 SQL 数据操纵语言 (DML)	31
1.3.5 视图 (VIEW)	33
1.3.6 SQL 数据控制语言 (DCL)	35
1.3.7 数据库开发工具	37
1.3.8 嵌入式 SQL	37
1.4 数据库管理系统	42
1.4.1 DBMS 的基本功能	42
1.4.2 DBMS 的层次结构	43
1.4.3 DBMS 的进程结构	44
1.4.4 DBMS 的查询处理及优化	45
1.4.5 DBMS 的分类	46
1.5 分布式数据库系统	46
1.5.1 什么是分布式数据库系统	46
1.5.2 发展分布式数据库的动机	48
1.5.3 分布式数据库系统的系统结构	49
1.5.4 分布式数据库系统的模式结构	49
1.5.5 分布事务管理	50
1.5.6 分布式查询优化	53
1.6 数据库应用	53
第 2 章 数据库的设计	55
2.1 数据库设计概述	55
2.2 数据库设计的需求分析	57
2.2.1 确认设计范围	58
2.2.2 面向数据的方法	58
2.2.3 面向过程的方法	64
2.3 数据库的概念设计	67
2.3.1 引言	67
2.3.2 局部视图设计	69
2.3.3 视图集成概述	78

2.3.4	预集成	81
2.3.5	实体类的集成	82
2.3.6	联系类的集成	85
2.3.7	新老数据模式的集成	85
2.4	数据库的逻辑设计	86
2.4.1	引言	86
2.4.2	E-R 数据模型到关系数据模型的转换	87
2.4.3	关系模式的规范化	91
2.4.4	关系模式的分解及其问题	94
2.4.5	改善性能的调整	95
2.4.6	节省存储空间的调整	95
2.4.7	外模式的设计	95
2.5	数据库的物理设计	97
2.6	分布式数据库的设计	105
第 3 章	CIMS 中的数据集成技术	110
3.1	CIMS 中的信息集成的需求和特点	110
3.1.1	CIMS 的主要功能子系统信息集成的需求	110
3.1.2	各功能子系统的信息特点	111
3.2	数据集成的目标	112
3.2.1	数据集成的理想目标	112
3.2.2	数据集成的近期目标	116
3.3	现有共享数据的存储机制	116
3.3.1	计算机系统体系结构的变化	116
3.3.2	支持数据共享的存储机制	118
3.3.3	支持数据共享的集成环境建立	119
3.4	分布式环境下的数据分布技术	119
3.4.1	数据分布的六种基本形式	120
3.4.2	数据分布的技术	123
3.5	分布环境下的多库集成	124
3.5.1	开放体系结构下的数据集成技术	125
3.5.2	多数据库的集成技术	133
3.6	工程数据管理	135
3.6.1	CAD/CAPP/CAM 系统集成的工程数据管理	136
3.6.2	产品数据管理系统	137
第 4 章	Client/Server 体系结构	145
4.1	发展背景	145
4.2	基本定义	147
4.3	客户/服务器模式的结构和原理	150
4.4	数据库服务器	157
4.4.1	硬件平台	157
4.4.2	网络操作系统	158
4.4.3	基于网络运行环境的数据库管理系统	161
4.5	典型的客户机平台	164
4.6	客户/服务器结构的设计与应用开发	167
4.7	可用的开发环境和工具	176
4.8	建立客户/服务器模式的注意点	182
参考文献		185

第 1 章 数据库系统概论

1.1 什么是数据库系统

随着现代科学技术的迅猛发展,人类社会已经进入到一个充分利用信息资源的信息社会,信息工程已成为涉及到各行各业的一门崭新的学科。数据库系统则是其中最重要的成果与工具之一。本节将从数据管理发展过程简要地阐明这个问题,并概括地介绍数据库系统以及所涉及的基本概念。

1.1.1 计算机数据管理技术的发展

早期的计算机技术,主要用于数值计算。计算机应用从科学计算进入数据处理是一个划时代的转折。数据处理是指对大量的需要持久保存的数据进行收集、储存、加工和传播的一系列活动的总和。利用计算机进行数据处理就是把原始数据送入计算机系统进行统一管理并按照给定的算法对原始数据进行加工,最后产生结果数据作为输出,其过程如图 1-1 所示。在这个过程中,数据管理技术对数据处理方法和处理效率有着直接的影响。

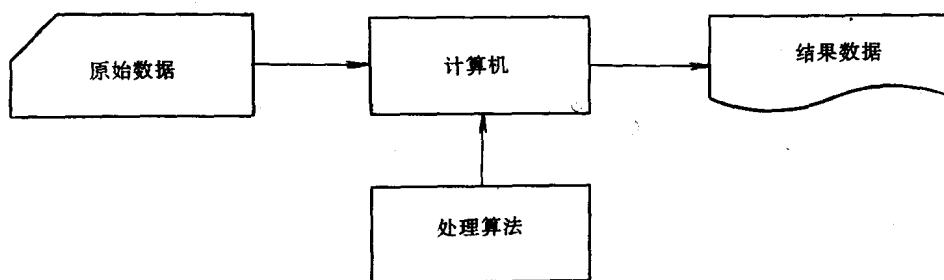


图 1-1

早期的数据管理都采用文件系统,数据是以文件的形式组织和保存的。文件是一组具有相同结构的记录的集合,记录由某些相关数据项组成。数据被组织成文件之后,就可离开处理它的程序而单独存在,并将由一个称做文件管理系统的专用程序对其进行管理和维护。文件系统按其对数据管理的方式,又可分为:独立文件方法和共享文件方法。图 1-2 是独立文件方法的文件系统工作示意图。其中,文件管理系统是应用程序和文件之间的一个接口,应用程序通过文件管理系统建立和访问文件。如图所示,应用程序 A,通过文件管理系统可以使用数据文件 A1;应用程序 B,通过文件管理系统可以使用数据文件 B1、B2;同样,应用程序 C,通过文件管理系统可以使用数据文件 C1、C2、C3。采用这种方法的优点在于:用户在设计应用程序时,只需考虑数据的逻辑结构和特征,并按规定的组织方式和存储方法,建立并使用相应的数据文件,用户可以较少地考虑数据物理存储方面的各种实现的细节问题,由此可以简化应用程序设计的复杂性。但是这种管理方法存在着以下弊病:

①文件的专用性 图 1-2 方式中的文件管理方法中,文件是用户专有的,它可能仅属于

个别的用户，甚至只属于个别用户的个别应用程序，这样会导致数据文件分散的状态，造成大量的数据冗余及不一致性。

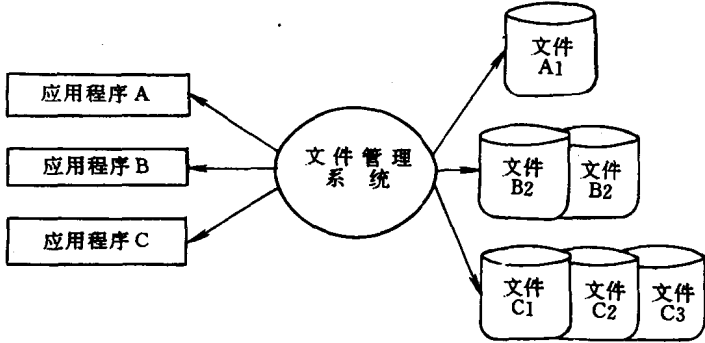


图 1-2

②数据和应用密切相关 文件中数据的逻辑定义、物理存储设备的指定、组织方式和存取方法都须由程序员在应用程序中设定，因此文件一旦离开它所依附的程序便失去存在的意义。而且文件和程序的变化必须是相应的，否则，相互之间因不能呼应而导致程序运行失败。

③数据的无结构性 独立文件中的数据，往往只表示了客观世界中单一事物的信息，而不反映各种事物之间的联系。

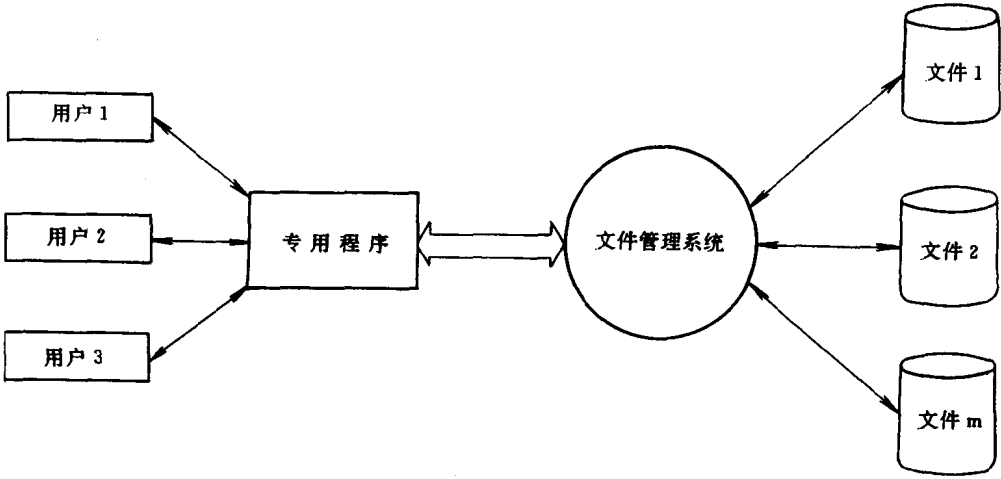


图 1-3

图 1-3 是共享文件方法的工作示意图。这方式中所建立的数据文件包括了多个用户可能用到的数据，这使得文件中的数据可为多个用户的应用程序所共享。但这种方法同样属于文件管理方法的范畴，存在着数据冗余度大，程序和数据独立性差等缺点。针对文件管理系统的一些缺点，人们逐步发展了以统一管理和共享数据为主要特征的数据库系统。

1.1.2 数据库系统

在数据库的发展过程中，人们对“数据库”的定义和解释是众说纷纭的，从不同的角度对数据库有不同的理解。在本书中，我们认为数据库是为满足某一组织中多个用户的多种应用需要，在计算机系统中按照一定的数据模型组织、存储和使用的相互联系的数据集合。所谓组织，是指一个独立存在的单位，如公司、工厂、学校、银行、商场等。所谓数据集合，是指关于企事业运行的业务数据的集成，如公司和工厂的订单数据、库存数据、各种计划数据、生产数据及成本核算数据等。这些数据，可以通过各种原始单据、测试或统计计算得到。

数据库系统则是由与运行数据库有关的计算机硬件、软件、数据集合和应用人员等所组成，它是一个为用户提供信息服务的系统，图 1-4 是这个系统的示意图。

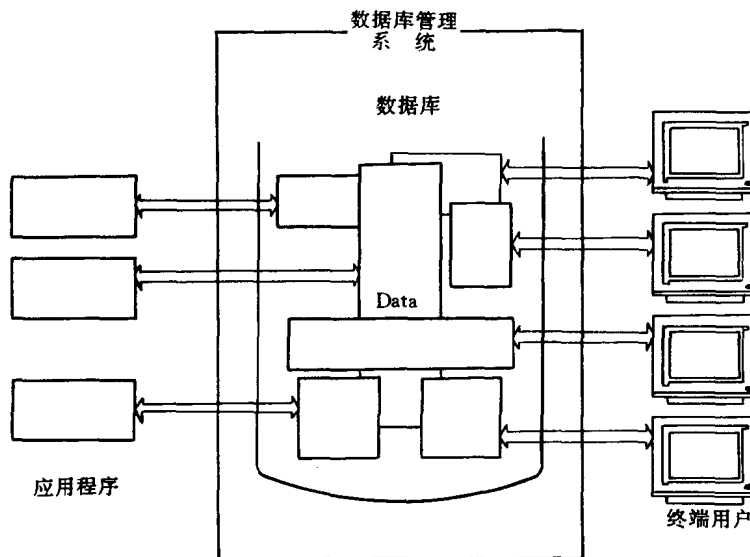


图 1-4

实际的数据库系统由下面几部分构成：

- ①数据库。
- ②计算机硬件。
- ③计算机软件 包括操作系统、数据库管理系统（DBMS）、应用软件及应用开发工具软件等。
- ④人员 包括数据库管理员（Data Base Administrator 简记为 DBA）、系统分析员、应用程序员和用户。

下面我们分别介绍这几部分的内容。

1. 数据库 数据库是某一组织中各种应用所需的业务数据集合，具有一定的结构和组织方式，并为所有的应用所共享。在数据库系统中，数据的定义与应用程序分开，对数据库的描述是独立的，因此数据库可以为多种任务所使用。在构造数据库时，我们可以完全地或部分地消除有关文件中的数据重复，减少数据的冗余存储。

2. 计算机硬件 用于数据库管理的计算机硬件资源包括中央处理器 CPU，用以存放系统

软件和数据文件的大容量存储器（目前主要指硬盘），以及其它一些相关的硬件设备。

3. 计算机软件 数据库系统的软件主要包括支持 DBMS 运行的操作系统以及 DBMS 本身，此外，为了开发应用系统，还要有各种高级语言及其编译系统。这些高级语言应该具有和数据库的接口。如 Microsoft 开放数据库联通（Microsoft Open Database Connectivity——ODBC）软件标准，使基于 Windows 的应用程序可方便地访问多种数据库系统中的数据。Microsoft 开放数据库连通标准不但定义了 SQL 语法规则，而且还定义了 C 语言同 SQL 数据库之间的程序设计接口。经过编译的单个 C 或 C++ 程序就有可能对任何带有 ODBC 驱动程序的 DBMS 进行访问。

应用开发工具软件是系统为应用开发人员和最终用户提供的高效率、多功能的应用生成器、第四代语言等各种软件系统。例如，数据字典，报表书写系统，表格软件，图形系统等等。它们为数据库系统的开发和应用提供了良好的环境。这些软件系统都是基于数据库的。

数据库管理系统（DBMS - Database Management System）是物理数据库本身（即实际存储的数据）和系统用户之间的界面。本书将在以后部分加以详细介绍。

4. 人员 管理、开发和使用数据库系统的人员主要是数据库管理员、系统分析员、应用程序员和用户。

1.1.3 数据库系统的主要特征

目前，数据库系统已成为数据处理业务部门用作管理其内部数据的主要工具，其主要原因是因为它具有独特的优点。数据库方法不仅克服了文件管理方法中存在的主要问题，而且还提供了更强有力的数据管理能力和维护能力，显示了很强的生命力。主要体现在以下几个方面。

1. 具有较高的数据独立性 数据库系统提供了两方面的映象功能。一个是数据的存储结构与逻辑结构之间的映象，一个是数据总体逻辑结构（概念模式）与某类应用所涉及的局部逻辑结构之间的映象。因此数据独立性通常分为两级：第一级为物理独立性。数据库的存储结构发生改变时，不影响数据的逻辑结构，也不致引起应用程序的修改；第二级是逻辑独立性，这是指数据库局部逻辑结构独立于总体逻辑结构概念模式。因此，当总体逻辑结构发生改变，局部逻辑结构可以不变，而基于这些逻辑结构的程序也可不变或改变很小。同样，当用户需要修改应用程序时，也不强求数据结构作相应的改变。数据库系统是“以数据为中心的”或“面向数据”的系统，它的数据和应用程序之间依存小，具有一定的数据独立性。

2. 数据冗余度小 数据冗余度指数据库中数据的重复程度。由于数据库从整体观点看待和描述数据，数据是面向整个系统的，而不是面向某一应用，因此可以大大降低数据冗余度，一方面节约了存储空间，减少存取时间，另一方面又可避免数据之间的不相容性和不一致性。必须指出，在一个数据库系统中，数据的冗余并不是都须消除的。在实际应用中，有时为了某种原因，在一定控制条件下的多次存储还是必要的。

3. 统一的数据管理和控制 数据库系统提供统一的数据定义、删除、检索及更新手段，并统一控制数据的安全性、完整性、保密性和并发性，使对数据的应用更加有效和可靠。

• 安全性保护 数据安全性是指保护数据以防止不合法的使用所造成的数据破坏和泄密。这需要采取一定的安全保密措施。例如，对一个应用用户，数据库系统将对他的权限进行安全性检查，只有合法的用户才能以规定的方式对允许他使用的那一部分数据进行相应的操作。

• 并发控制 并发操作是指多个用户进程在同一时刻期望存取同一数据时发生的事件。为了避免并发进程之间相互干扰而导致错误的结果或使数据完整性遭到破坏，必须对多用户的并发操作加以控制、协调。

• 数据完整性控制 数据完整性指数据的正确性、有效性和相容性。数据库系统提供了必要的功能，保证数据在输入、修改过程中始终符合原来的定义和规定，如工人每天的工时不能超过“8”小时或工人的年龄必须在18到65之间。数据完整性检查对用户系统尤为重要，数据库系统必须对完整性进行必要的维护。

• 故障的发现和恢复控制 数据库系统运行过程中，由于用户操作失误及硬件/软件的故障，可能使数据库遭到局部性或全局性破坏，系统能进行应急处理，把数据库恢复到正确状态。

1.1.4 数据库系统的体系结构

1975年，美国ANSI/SPARC数据库管理系统研究组的研究报告提出了数据库三级结构的建议，即一个数据库的数据模式结构从逻辑上可分为外部级、概念级和内部级三级。该结构已成为目前世界上各种类型和各种应用数据库系统的体系结构的依据。

数据库系统的三级模式结构是概念模式、外模式和内模式。如图1-5所示。

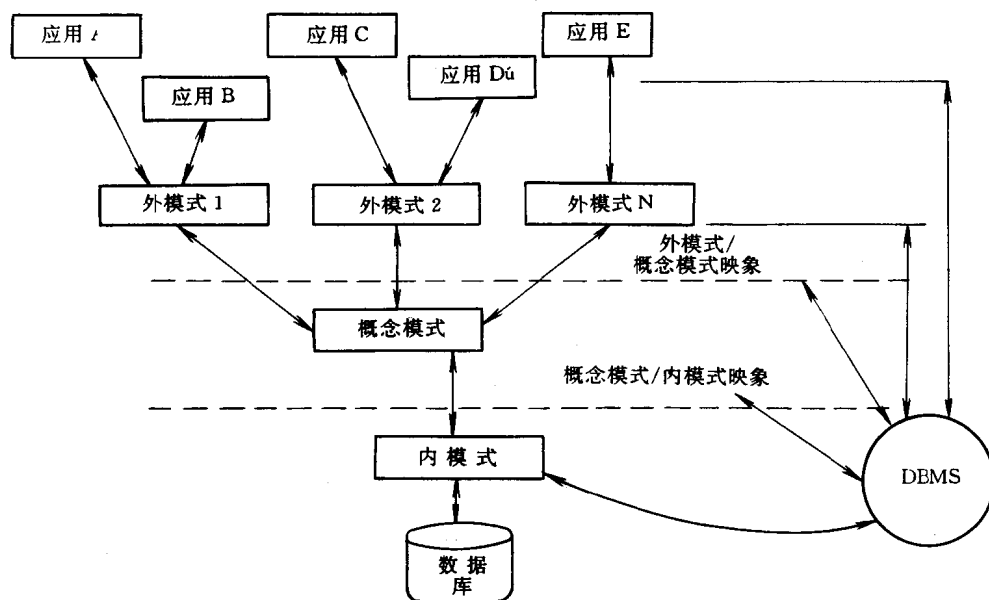


图 1-5

外模式又称子模式或用户模式，是数据库用户看到的数据视图，它是三级结构中最接近于用户的级。每个用户必须以自己的观点使用一个外部模式来对数据库的部分数据结构进行逻辑描述。即描述用户视图中记录的组成、相互联系、数据项的特征、数据的安全性和完整性约束条件等。

概念模式又称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述。一个数据库允许有许多不同的外部模式，但只有一个概念模式，它描述整个数据库的逻辑结构，包括现实世界中实体的性质与联系，定义记录、数据项、数据的完整性约束及记录之间的联系。概念

模式通常简称模式。子模式是模式的一个子集，从一个模式可以导出不同的子模式。

内模式又称存储模式，是数据在数据库系统内部的表示，即为数据的物理结构和存储方式的描述，包括各种存储记录的类型、索引及索引字段、记录的物理存储顺序等。

数据库系统的三级模式是数据的三个抽象级别，使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的表示和存储。为了实现三个抽象层次的联系和转换，数据库系统在三级模式中提供两个映象：外模式/概念模式映象和概念模式/内模式映象。

1.2 数据模型与数据模式

1.2.1 数据模型

数据模型 (Data Model) 是描述现实世界的工具，也是实现 DBMS 的基础。一般地讲，数据模型是严格定义的一些概念的集合。这些概念精确地描述了数据、数据间的联系、对数据的操作以及有关的语义约束规则。著名的 DB 专家 E. F. Codd 提出，一个数据模型实质上是一组向用户提供的规则，这组规则规定数据结构如何组织以及相应地允许进行何种操作。一般来说，一个数据模型由三个部分组成：数据结构、操作集合和完整性约束规则。

数据结构是所研究的对象类型的集合，它是数据模型中最基本的部分。对象类型有两类：一类是指数据类型、内容、性质，另一类是数据之间的联系。通过对这些对象的描述，来确定符合所选模型的数据库的逻辑结构，即数据在逻辑上是如何组织。

操作集合是指对数据库中各种对象的实例允许执行的操作的集合，包括操作及有关的操作规则。数据库的操作主要包括增、删、改、查四种。数据模型要定义这些操作的确切含义、操作符号、操作规则以及实现操作的语言。

完整性约束规则是定义数据的约束条件，即给定数据模型中的数据及其联系所具有的制约和依存规则，用于限定数据库的状态及变化，以保证数据的正确、有效和相容。

1.2.2 数据模式

数据模式和数据模型是两个不同的概念，不应混淆。数据模式是以某一种数据模型为工具，对一个具体单位的数据结构的描述。如前面所述，在 DBMS 中有三级数据模式的表示机制，数据库系统中不同的人员涉及到不同的数据模式的表示形式，如图 1-6 所示。

1.2.3 常见的数据模型

现有的数据模型按其目的，可分为两大类。一类是信息模型（也称概念模型），这类模型用于将现实世界抽象为信息世界，这种信息结构不依赖于具体的计算机系统，不是由某一个 DBMS 支持的数据模型所表示的，而是用概念模型来描述。此类模型主要包括实体—联系模型 (ER 模型)、扩展的实体联系模型 (EER 模型)、IDEF1X 模型、面向对象模型 (OO 模型)、语义网络模型等。信息模型用于信息世界的建模，它强调其语义表达能力，能较方便、直接地表达应用中各种语义知识。这类模型应当概念简单、清晰，易于为用户理解。另一类模型则是直接面向数据库中数据的逻辑结构的，为了和概念模型区别，称之为基本数据模型或结构数据模型。任何一个数据库管理系统都有它自身支持的基本数据模型。基本数据模型通常需要严格的形式化定义，以便于机器上实现，它是用于机器世界的模型。

1. 实体—联系模型 (ER 模型) 实体—联系模型是最常用的概念模型。该模型是 1976 年由 Peter pin-shan chen 提出的一种对现实世界进行抽象的方法。ER 模型描述整个组织的概念模式，该模式不考虑效率和物理数据库的设计。它能充分地反映现实世界，易于理解，将

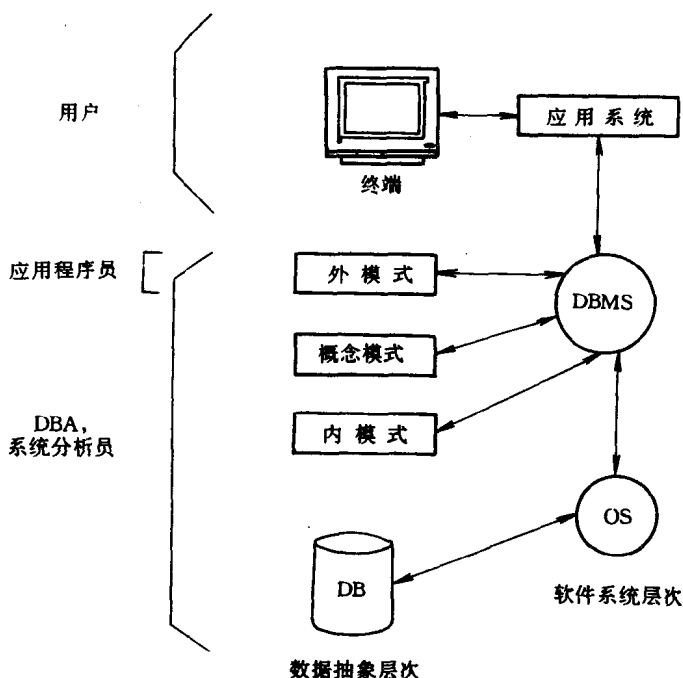


图 1-6

现实世界的状态以信息结构的形式很方便地表示出来。ER 模型中涉及的主要概念有：

(1) 实体 实体是客观存在并可相互区分的事物，它可以是人、物等实际的对象，也可以指某些概念；可以是事物本身，也可以指事物与事物之间的联系。例如一个学生、一本书、一个部门、学生的一次选课、部门的一次定货等都是实体。

(2) 属性和码 每个实体具有的特性，称为属性。一个实体可以由若干个属性来刻画。例如，学生实体可以由学号、姓名、年龄、性别、系、年级等属性组成。

唯一地标识实体的属性称为实体的码或关键字。

例 (2520994, 夏青, 23, 男, 计算机系, 硕士研究生)，这些属性组合起来表征了一个学生，其中学号是学生实体的码。

(3) 实体集 (Entity Set) 实体集是具有相同属性的实体的集合。例如在图书馆管理系统中全体书实体组成书的实体集，全体学生组成学生实体集等等。

(4) 联系 现实世界中事物之间是有联系的，这种联系必然要在信息世界中给以反映。实体间的联系可分为三类：

一对一联系 (1:1)，若对于实体集 A 中的每一个实体，实体集 B 中至多有一个实体与之联系，反之亦然，则称实体集 A 与实体集 B 具有一对一联系，记为 1:1。

一对多联系 (1:n)，若对于实体集 A 中的每一个实体，实体集 B 中有 n 个实体 ($n > 0$) 与之联系，反之，对于实体集 B 中的每一个实体，实体集 A 中至多只有一个实体与之联系，则称实体集 A 与实体集 B 有一对多的联系。记为 1:n。

多对多联系 (m:n)，若对于实体集 A 中的每一个实体，实体集 B 中有 n 个实体 ($n > 1$) 与之联系。反之，对于实体集 B 中的每一个实体，实体集 A 中也有 m 个实体 ($m > 1$) 与之联系，则称实体集 A 与实体集 B 具有多对多联系，记为 m:n。

ER 模型是用 ER 图来表示的。ER 图中的一些约定如下：

①用长方形表示实体型，在框内写上实体名。

②用椭圆形表示实体的属性，并用无向边把实体与其属性连接起来。双椭圆表示该属性为实体的码。

③用菱形表示实体间的联系，菱形内写上联系名。用无向边把菱形分别与有关实体相连接，在无向边旁标上联系的类型。若实体之间联系也具有属性，则把属性和菱形也用无向边连接上。

下面我们用 ER 图来表示某个工厂的库存管理。如图 1-7 所示。该库存管理涉及的实体集有：

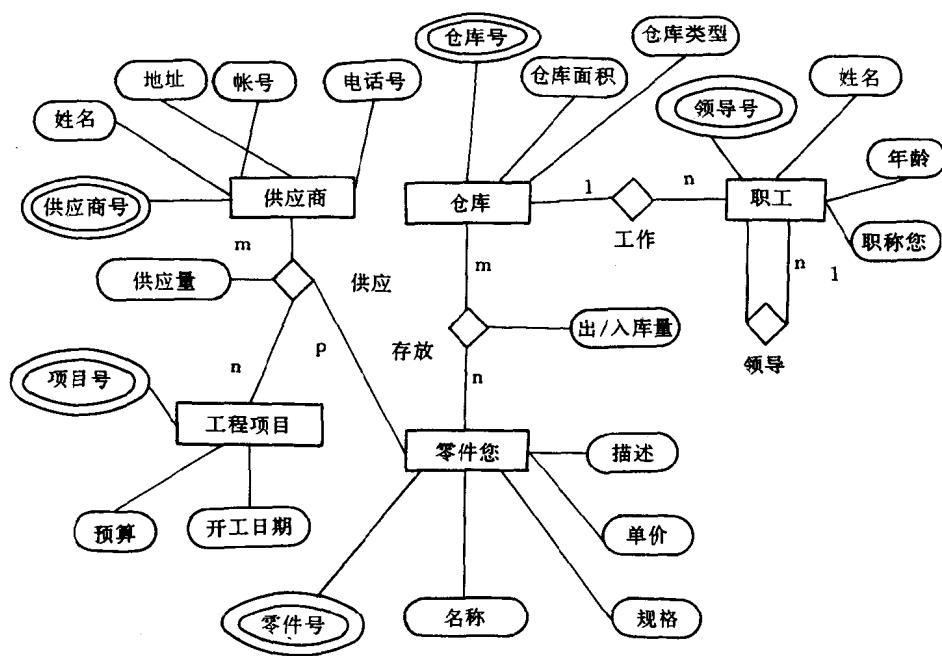


图 1-7

仓库实体集、零件实体集、供应商实体集、工程项目实体集和职工实体集。这些实体集属性如下：

仓库的属性包括仓库号、仓库面积、仓库类型，其中仓库号为主码；

零件的属性包括零件号、名称、规格、单价、描述，其中零件号为主码；

供应商的属性包括供应商号、姓名、地址、电话号、帐号，其中供应商号为主码；

工程项目的属性包括项目号、预算、开工日期，其中项目号为主码；

职工的属性包括职工号、姓名、年龄、职称，其中职工号为主码。

这些实体之间的联系如下：

①仓库和零件为多对多关系 ($m:n$)，一个仓库可以存放多种零件，一种零件可以存放在多个仓库中。

②仓库和职工之间为一对多的联系，一个仓库有多个职工当仓库保管员，一个职工只能在一个仓库工作。