

数据库系统引论

徐洁磐 王银根 编著



南京大学出版社

网络操作系统引论

网络操作系统引论



网络操作系统引论

441963

TP311.13
X76-2

数据库系统引论

徐洁磐 编著
王银根

南京大学出版社
1996·南京

ISBN 7-305-03020-1/TP·157

数据库系统引论

徐洁磐 王银根 编著

*

南京大学出版社出版

(南京大学校内 邮政编码:210093)

江苏省新华书店发行 江苏省地质测绘院印刷厂印刷

*

开本 787×1092 1/16 印张 16 字数 396 千

1997 年 2 月第 1 版 1997 年 2 月第 1 次印刷

印数 1—4000

ISBN 7-305-03020-1/TP·157

定价: 21.00 元

(南大版图书若有印、装错误可向承印厂退换)

前 言

数据库技术在我国日渐普及,数据库应用也日渐拓宽,据不完全统计,在计算机应用中涉及到数据库应用的百分比已占 60%以上,数据库不仅在传统的商业领域中、管理领域中发挥着主要作用而且在非传统应用中也起到越来越大的作用,如在工程领域中,在数据集成领域中,在科学统计领域以及在图形、图像、CAD、多媒体等领域中起着关键作用。

随着技术的逐渐成熟,数据库自身也起着显著变化,关系数据库技术已成为当前数据库发展的主流,层次与网络数据库已退出舞台,另外,分布式数据库系统及面向对象数据库系统正在兴起。

鉴于以上情况,有必要编写一本与数据库新技术有关材料以适应当前数据库应用的发展及数据库自身的发展,本书正是为此目的编写的,它包括数据库技术的最新内容,其主要有以下几点:

1. 突出以关系数据库技术为主要内容。
2. 增加分布式数据库及面向对象数据模型的内容。
3. 增加数据库开发工具的内容。
4. 增加数据库设计的内容。
5. 减少及压缩层次、网络模型的内容。

以上述五点为基础,再经过结构上的调整构成了一本具有能适应当前数据库技术发展的具有最新内容的数据库书稿。

本书可适用于作为高等学校数据库课程教材,也可适用于数据库应用开发人员参考之用。本书内容深入浅出,文字浅显易懂,因此也适用于各类自学人员使用。

本书以 DB₂ 为背景数据库并同时兼顾 ORACLE、SYBASE 等流行数据库的背景要求。

本书由徐洁磐编写一、二、三、四、五、七、八、十等共八章,由王银根编写六、九两章,柏文阳编写十一章。

在本书编写过程中得到众多同仁的支持,如王静英老师,汪辉、朱树春老师等,在此特表示感谢,由于作者水平有限,书中难免有不少缺点和错误,殷切希望读者批评指正。

作 者

1995. 11. 26

目 录

第一章 概论	1
1.1 数据、数据库、数据库管理系统与数据库系统	1
1.2 数据库系统的发展	2
1.3 数据库系统的基本特点	3
1.4 数据库系统的构成	5
1.5 数据库系统的工作过程	6
习题	7
第二章 数据模型	8
2.1 数据模型基本概念	8
2.2 数据模式	9
2.3 数据模型的四个世界.....	10
2.4 概念世界与概念模型.....	11
2.5 信息世界与逻辑模型.....	22
2.6 计算机世界与物理模型.....	32
习题.....	34
第三章 关系数据库系统	35
3.1 关系数据库系统概述.....	35
3.2 关系数据库系统的衡量准则.....	36
3.3 关系代数——关系模型数学理论之一.....	38
3.4 关系演算——关系模型数学理论之二.....	46
习题.....	49
第四章 关系数据库系统标准语言 SQL	50
4.1 SQL 概貌	50
4.2 SQL 数据定义功能	51
4.3 SQL 的数据操纵功能	53
4.4 视图.....	61
4.5 嵌入式 SQL	63
4.6 SQL 的特点	64
习题.....	65
第五章 数据库开发工具	66

5.1	概论	66
5.2	数据库应用开发工具内容	66
5.3	数据库应用开发工具发展的三个阶段	67
5.4	Power Builder 介绍	68
	习题	72
第六章	分布式数据库管理系统	73
6.1	引言	73
6.2	分布式数据库系统的分类和结构	76
6.3	数据分布策略	83
6.4	查询分解和优化	86
6.5	分布式数据库系统中的并发控制与恢复技术	95
6.6	ORACLE 数据库管理系统的分布式功能	102
6.7	SYBASE 数据库管理系统的分布式功能	106
	习题	110
第七章	数据库保护	112
7.1	数据库的安全性保护	112
7.2	数据库的完整性保护	117
7.3	数据库的并发控制技术	118
7.4	数据库恢复技术	121
	习题	122
第八章	关系数据库规范化理论	124
8.1	引言	124
8.2	规范化理论	126
8.3	规范化所引起的一些问题	136
	习题	137
第九章	数据库设计	138
9.1	数据库设计概述	138
9.2	数据库设计的需求分析	145
9.3	概念模型设计	163
9.4	逻辑设计	191
9.5	物理设计	200
	习题	207
第十章	关系型数据库系统——DB₂	212
10.1	RDBMS 简介	212
10.2	DB ₂ 的基本结构	212
10.3	DB ₂ 的数据定义	216
10.4	DB ₂ 的数据操纵	226
10.5	DB ₂ 系统的数据存储	233
10.6	DB ₂ 系统的内部特性	236

10.7	DB ₂ 小结	240
第十一章	数据库技术的新发展	241
11.1	面向对象数据库系统	241
11.2	工程数据库与多媒体数据库	242
11.3	知识库系统	243

第一章 概 论

1.1 数据、数据库、数据库管理系统与数据库系统

计算技术的发展,计算机应用的进一步深入与拓广,使得数据在计算机应用中的地位与作用日益重要,它在商业中、事务处理中占有主导地位,近年来在统计领域、工程领域,在图形、图像、声音等多媒体领域以及在智能领域等的地位与作用也变得十分重要,因此,数据已成为构成一个计算机系统的重要支柱。数据同时也成为一种财富与信息资源,它具有独立地向社会提供信息的能力。

由于数据作用的日渐重要,它在计算机软件中的地位也发生变化,过去,系统是以程序为主体,数据是程序附属品,这样使得数据在系统内是分散、凌乱的。这也造成了管理上的混乱,使数据大量冗余,产生了数据的不一致性以及数据安全性、完整性方面的问题等。

但是,随着数据在软件中的地位的改变,系统是以数据结构及数据为主体而程序已退居于附属地位,在此种情况下,需要对数据作集中统一的管理并以数据共享为目标,这就出现了数据库系统(database system)。

数据库系统由下面几个部分组成:

1. 数据库(database,简称 DB)

数据库是数据的集合,它具有一定的组织形式并存放于统一的存贮介质上,它是多种应用的数据集成,并可被应用所共享。

2. 数据库管理系统(database management system,简称 DBMS):

数据库管理系统是管理数据库的机构,它是一种系统软件,它负责:

▲数据库中数据的组织;

▲数据库中数据的维护以及保护数据不受破坏;

▲数据库中数据的操作。

近年来,有些数据库管理系统还提供构造多种用户界面的工具软件或第四代语言以及构造应用的开发系统。

3. 数据库应用(database application)

数据库提供了数据共享的功能,因此多个应用程序可以使用数据库,应用程序使用数据库是通过 DBMS 实现的。

4. 数据库管理员(database administrator,简称 DBA)

由于数据库具有共享性,因此对数据库设计、规划、协调需专职人员负责进行此项工作,这些人员或集体统称为数据库管理员。

5. 系统硬件(system hardware)

数据库、数据库管理系统及数据库应用都是建立在计算机上的,这种计算机要求有较大容

量的外存储器(如磁盘存储器)以及较强 I/O 通道能力,此外,还需要有一个较好的操作系统以及建立在该操作系统上的文件系统。

系统硬件提供了数据库系统的基本物理支撑,在目前大多数数据库系统中,可以适应多种不同硬件环境(包括多种计算机与多种操作系统)使系统更为灵活与方便。

上述五个部分构成了一个完整的数据库系统,有时为方便起见也可简称为数据库。图 1-1 给出了数据库系统五个部分的关系

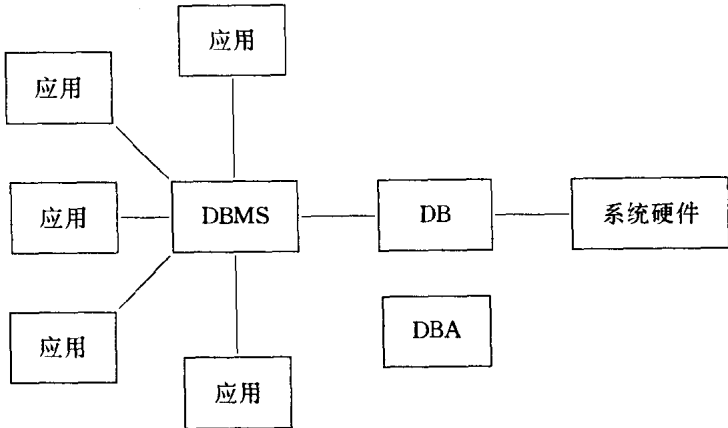


图 1-1 数据库系统示意图

1.2 数据库系统的发展

数据库系统的发展分为若干个阶段：

1. 文件系统阶段

文件系统是数据库系统的初级阶段,它出现于 50 年代至 60 年代中,它提供了简单的数据共享与数据管理功能,但它无法提供完整统一管理与数据共享能力,由于它的功能较为简单,一般它均附属于操作系统而并不成为独立的软件。目前,一般已将它看成为仅是数据库系统雏形而并不是真正的数据库系统。

2. 初级数据库系统阶段

自 60 年代起各种数据库系统相继问世,它们是层次模型与网络模型的数据库系统,它们为统一管理与共享数据提供了有力支撑。这个时期由于数据库系统的蓬勃发展,而形成了著名的“数据库时代”。当然,这两种类型的数据库系统也有其一定的不足,最主要是它们均脱胎于文件系统,因此,受文件中的物理结构影响较大,用户在使用数据库时需要对数据的物理结构有详细的了解,这对使用数据库带来了诸多的麻烦,同时,数据库中表示数据模式的结构方式过于烦琐,这也影响了数据库应用中越来越多的对数据结构的复杂要求的实现。

3. 关系数据库系统阶段

关系型数据库系统的理论出现于 70 年代初,其系统形成于 70 年代中并在 80 年代得到了充分的发展,它具有简单的结构方式与较少的物理表示,使用与操作又极为方便,因此,在 80 年代它逐步取代层次与网络型数据库成为占主导地位的数据库,到目前为止,关系型数据库系统仍占领数据库应用的主要地位。

4. 高级数据库系统阶段

由于计算机应用的发展,计算机已从传统的科学计算、事务处理等领域逐步扩展到工程设计统计、人工智能、多媒体、分布式等领域,这些新的领域需要有新的数据库支撑,而传统关系数据库系统是以商业应用、事务处理为背景而发展起来的,它并不完全适用于新领域的应用,因此,需要有新的数据库系统以满足不同领域的要求,自 80 年代中开始,各种适应不同领域的新型数据库系统不断涌现,如工程数据库、多媒体数据库、CAD 数据库、图形数据库、图像数据库、智能数据库、分布式数据库以及面向对象数据库等,其中特别是面向对象数据库系统由于其通用性强,适应面广而受到青睐,预计 90 年代后期将会形成多种数据库系统共同支撑应用的局面。当然在那时,关系数据库系统仍将是其中一支强有力的主力军,而面向对象数据库将会起主导作用。

在我国,数据库的发展也经历了上述的类似的四种阶段,但是略有不同。由于我国计算机应用发展较技术先进国家略晚一步,因此,文件系统阶段较为短暂,而初级数据库系统阶段基本没有经历,自 80 年代中起即进入关系数据库系统阶段。特别是以 dBASE-Ⅲ 与 Foxbase 为代表的初级关系型数据库系统阶段和近年来以 ORACLE、Sybase 为代表的成熟的关系型数据库系统阶段。目前,我国也已逐渐向高级数据库系统方向靠拢。

1.3 数据库系统的基本特点

下面给出数据库系统的几个基本特点,这些特点也给出了数据库系统与文件系统以及其他的本质区别。

1. 数据的集成化

在数据库系统中数据按一定模式组织与存放,这就称为数据的集成化或称为数据结构化,集成化的数据能较为自然的反映数据间的内在的联系。它也是对数据的集中控制和减少数据冗余的前提和保证。

由于数据库是从一个组织的全部应用的全局考虑并集成其数据结构,因此,数据库中的数据已不再是面向个别应用,而是面向整个系统,各不同应用所需数据只是整个全局数据的局部子集。

2. 数据独立性

数据独立性即是数据库中数据独立于应用程序而不依赖于应用程序,也就是说数据的逻辑结构、存储结构与存取方式等不因应用程序的变化、修改而有所改变。

数据独立性一般分为物理独立性与逻辑独立性两级。

1) 物理独立性

物理独立性即是数据的物理结构(包括存储结构、存取方式等)的改变,如存储设备的更换、物理存储的更换、存取方式改变等都不影响数据库的逻辑结构,从而不致引起应用程序的变化。

2) 逻辑独立性

数据库总体逻辑结构的改变,如修改数据模式、增加新的数据类型、改变数据间联系等,不需要相应修改应用程序,这就是数据的逻辑独立性。遗憾的是到目前为止数据的逻辑独立性还无法做到完全的实现。

总之,数据独立性即是数据与程序间的互不依赖性。一个具有数据独立性的系统可称为以数据为中心的系统或称为面向数据的系统。显然,数据库系统是以数据为中心的系统或是面向数据的系统。

3. 数据共享

数据资源是一种十分重要的信息资源,充分利用这种资源的重要手段即是数据共享,数据共享是数据库技术先进性的重要体现。数据库中的数据共享主要体现在以下几方面:

- ▲数据库中的数据可供多个应用使用;
- ▲可以开发新的应用而不增加新的数据;
- ▲数据库中数据可直接对外开改,向社会提供服务。

4. 减少数据冗余

在非数据库系统中,每个应用均有其相应的数据文件,因此数据反复重复,冗余量大,而在数据库系统对数据进行集中、统一管理,因此从本质上解决了数据的冗余问题。数据的冗余不仅大量浪费空间,增加系统维护副本的开销,而且容易造成数据的不一致,因此数据冗余的减少可为数据库减少空间、提高效率、保证安全提供支撑。

当然,在数据库系统中绝对消除冗余是不可能的,也是不必要的,适当保留一定冗余有时对提高系统效率会起一定作用。

5. 集中管理

数据库系统负责对其数据作全面、集中的管理,这些管理包括提供统一数据模式,统一数据操纵方式,此外,还提供集中的数据安全性、完整性、一致性以及并发处理、故障恢复等管理功能。

6. 保证数据的一致性

数据的一致性即是在数据库中同一数据的不同出现应保持其值的一致。数据的不一致即是指同一数据在数据库内重复出现且具有不同的值。如职工工资在工资单上和人事档案中有不同的值,这就是数据的不一致性。在数据库内由于数据冗余度的减少,避免了大量的不一致性的产生,同时,数据库系统提供各种控制、检查,保证在更新数据时同时更新所有副本,从而保证了数据的一致性。

7. 数据的完整性、安全性保护

数据库系统提供各种手段以防止数据的非法使用及保证数据的正确性。

8. 数据的并发控制与故障恢复

数据库系统提供多个用户共享数据的能力,它通过并发控制避免多个用户“同时”对同一数据操作时所产生的错误,这种错误的最著名的例子是民航订票问题。有两个终端同时操作购买一张去北京的机票,而此时在数据库内仅剩1张北京机票,甲终端首先操作并预订了该票,但是当它准备修改而尚未修改数据库时(即将北京机票由1变为0),乙终端抢入并发现仍有1张北京机票,因此,也预订了该票,这样就造成了一张票两人订的错误结果,这种错误的出现并非由机器故障造成,而是因对并发操作未施行控制所造成。并发控制可以避免这种错误的发生,保证数据的正确性。

数据库中的数据在系统运行过程中可能会受到破坏,这包括硬件、软件故障、周围环境刺激以及用户操作失误等多种原因所造成,其中有的是不可避免的。而数据库系统有一套完整的措施能及时发现故障并迅速将其恢复至正确状态,这就是数据库系统的故障恢复能力。

9. 系统标准化

由于数据库中数据的集中管理与集中控制,因此能够保证数据的表示与操纵的标准化,尤其是数据格式的标准化以利于数据交换与系统间数据沟通。

上述九点构成了数据库系统有别于文件系统及其他数据处理系统的主要标志。特别要注意的是数据库系统与文件系统间的差别,这两者既有紧密联系又有显著差别,其主要方面如下:

(1)数据库系统与文件系统有一定的数据管理功能,这是其主要相同点。

(2)数据库中数据是由数据库系统统一规划并按照一定数据模式组织和建立的,由系统统一管理和集中控制,而文件系统则是由各应用分别建立,数据分散,无法集中、统一管理与控制。

(3)文件系统中数据结构简单,不能反映现实世界中数据间的内在联系,而数据库系统所提供的数据结构(叫数据模式)具有灵活、方便的反映客观世界中数据间的多种关系的能力。

(4)文件系统中数据文件是面向应用的,它依赖于应用程序的存在而存在,应用程序的改变必须重新组建文件,而数据库系统则不然,它是面向数据的,它不依赖于应用程序,具有数据的物理独立性与逻辑独立性。

(5)数据库系统的数据冗余小,具有保证数据安全、完整的能力,还提供并发控制,故障恢复的功能,而文件系统的冗余大,数据的安全性、完整性能差,并发控制与故障恢复功能也几乎没有。

(6)文件系统是数据管理的低级工具而数据库系统则是高级工具,在物理实现上大多数数据库系统还是以文件系统为其实现基础的,数据库中数据最终是以文件形式存放在物理设备上的。因此,文件系统是数据库系统的基础,数据库系统是文件系统的提高与发展。

1.4 数据库系统的构成

数据库系统是一个复杂的系统,它由五个部分组成,它们是:

- ▲集成化、结构化的数据集合——数据库 DB;
- ▲管理数据的软件——数据库管理系统 DBMS;
- ▲系统应用;
- ▲硬件;
- ▲数据库管理员 DBA。

下面对这几个部分中的数据库、数据库管理系统及数据库管理员作介绍。

1. 数据库

数据库存放数据。数据按数据库所提供的数据模式存放,数据模式能构造复杂的数据结构以建立数据间的内在联系与复杂的关系,从而构成数据的全局结构模式以适应数据共享要求。而数据库的每个应用则仅是取全局模式中的一个局部子模式。

数据库中的数据具有“集成”、“共享”之特点,亦即是数据库集中了各种应用的数据,进行统一构造与存储,同时,这些数据可为不同应用使用。

2. 数据库管理系统

数据库管理系统是一种管理数据库的系统软件,它是数据库系统的核心,数据库管理系统

主要有以下几方面功能：

1)数据的定义

数据库管理系统一般均提供数据定义语句(data definition language,简称 DDL)用以定义数据模式,它包括全局模式,局部模式以及部分物理存贮结构。

2)数据的操纵

数据库管理系统一般均提供数据操纵语言(data manipulation language,简称 DML)用以对数据库中数据作检索、插入、修改和删除等操作。

DML 一般分两种：

(1)交互式命令语言:它语法简单,可在终端上即席操作,它又称为自含型或自主型语言,这种语言有时还包含部分控制语句并能独立编程。

(2)宿主型语言:它一般可嵌入于某些主语言中,如可嵌入于 FORTRAN、C、PASCAL 等高级语言中。这种语言本身不能独立使用,因此,称为宿主型语言。

3)数据的控制

数据的控制主要完成数据安全性、完整性及一致性的定义与检查,它们由相应的 DDL 或 DML 中的有关语句实现。此外,数据库的控制还完成多用户访问过程中的并发控制。

4)数据的服务

它包括数据库中初始数据的录入,数据库的转储、重组、性能监测、分析以及系统恢复等功能,它们大都由 DBMS 中的服务性程序完成。

3. 数据库管理员

数据库系统是一个复杂的系统,它可供多个用户使用,它的数据可作为信息资源向社会开放,因此,对它的管理极为重要。需要由一组专门人员进行管理,这就是数据库管理员 DBA, DBA 的主要工作有下面几点:

1)数据库设计(database design)

数据库设计即是数据模式设计。由于数据库中的集成与共享性,因此,需要有 DBA 对数据库作全局与局部的模式设计。

2)数据库的维护

DBA 必须对数据库中数据的安全、完整、并发控制及系统恢复进行实施与维护。

3)改善系统性能提高系统效率

DBA 必须随时监视数据库运行状况,不断调整物理结构(如重建索引与集簇、系统重组与转储等)使系统保持最佳状态与最高效率。

1.5 数据库系统的工作过程

为了解数据库系统的工作原理,我们以查询为例,对数据库的工作过程作了一个了解,其工作步骤如下:

(1)某应用程序用相应 DML 语句向 DBMS 发出查询请求并递交必要的参数,此后控制即转至 DBMS。

(2)DBMS 分析应用程序提交的命令及参数,并进行合法性检查,如不通过则拒绝执行并向应用程序返回出错信息。

- (3)合法性检查通过后,取出相应模式,执行相应程序,最后确定应访问的物理记录区域。
- (4)DBMS 向 OS 发出“读”物理记录的命令
- (5)OS 执行物理读/写操作,将数据从外存读至 I/O 缓冲区。
- (6)DBMS 将 I/O 缓冲区数据传送至其工作区作适当的组织加工后,再传送至用户工作区。
- (7)DBMS 将命令执行结果状态信息传送给应用程序。
- (8)查询执行完毕控制转向应用程序。

习 题

1. 试解释下列术语:
 - 数据库.
 - 数据库管理系统.
 - 数据库系统.
2. 试述数据库系统的组成内容。
3. 试比较文件系统与数据库系统的异同。
4. 什么叫数据的物理独立性与逻辑独立性? 并说明它的重要性。
5. 试述数据库系统的优点。
6. 什么叫数据的冗余与数据的不一致性?
7. 什么叫数据库管理员? 它的主要工作是什么?
8. 试述数据库系统发展的几个阶段。

第二章 数据模型

2.1 数据模型的基本概念

数据是现实世界“量”的抽象,而数据模型(data model)则是数据特征的抽象。在数据库系统中,数据模型是它的核心与基础。数据模型描述数据的结构,定义在其上的操作以及约束条件。它从概念层次上描述了系统的静态特征、动态特征和约束条件,为数据库系统的信息表示与操作提供一个抽象框架。

数据模型按不同的应用层次分成三种类型,它们是概念数据模型(conceptual data model)、逻辑数据模型(logic data model)、物理数据模型(physical data model)。

概念数据模型又称概念模型,它是一种面向客观世界,面向用户的模型,它与具体的数据库管理系统无关,与具体的计算机无关。概念模型着重于对客观世界复杂事物的结构描述及它们间的内在联系的刻画,而将与 DBMS、计算机有关的物理的、细节的描述留给其他类型的模型。因此,概念模型是整个数据模型的基础。目前,较为有名的概念模型有 E-R 模型、扩充的 E-R 模型、面向对象模型等。

逻辑数据模型又称数据模型,它是一种面向数据库系统的模型,该模型着重于在数据库系统一级的实现。它是客观世界到计算机间的中介模型,具有承上启下的功能。概念模型只有在转换成数据模型后才能在数据库中表示。目前,数据模型很多,较为成熟并被人们所大量使用的有:层次模型、网状模型及关系模型。目前占主导地位的是关系模型。面向对象模型既是概念模型又可作为数据模型,由于其表示上的优越性,目前越来越被人们所注意。

物理数据模型又称物理模型,它是一种面向计算机物理表示的模型,此模型给出了数据模型在计算机上真正的物理结构的表示。

数据模型所描述的内容有三个部分,它们是数据结构、数据操作与数据约束。

1. 数据结构

数据模型中的数据结构主要描述数据类型、内容、性质的有关情况以及描述数据间的联系的有关情况。数据结构是数据模型的基础,数据操作与约束均建立在数据结构上。不同数据结构有不同的操作与约束。因此,一般数据模型的分类均以数据结构的不同而分。

2. 数据操作

数据模型中的数据操作主要描述在相应数据结构上的操作类型与操作方式。

3. 数据约束

数据模型中的数据约束主要描述数据结构内数据间的语法,语义联系,它们间的制约与依存关系,以及数据动态变化的规则以保证数据的正确、有效与相容。

2.2 数据模式

在数据模型中有关数据结构及相互间关系的描述这部分有一个专用名词称为数据模式(data schema)。数据模式是数据模型的一部分,但是由于其重要性,因此作为一个专门术语并作专门的讨论。

数据库的数据模式可分成为三级,它们分别是:概念模式(conceptual schema)、外模式(external schema)和内模式(internal schema)。

1. 概念模式

概念模式给出了整个数据库的全局逻辑结构,它是全体用户的公共数据视图。概念模式所描述的结构是面向数据库系统的,它一般以某种数据模型为基础,描述数据的类型、长度、特征、数据间的联系以及有关安全性、完整性要求。数据库系统一般提供模式描述语言用以描述概念模式。概念模式对应于数据模型中的逻辑数据模型。

2. 外模式

外模式又称子模式或用户模式,它是用户的数据视图,亦即是用户所见到的模式的一个部分或由概念模式推导而出。概念模式给出了全局的数据描述而外模式则给出局部的描述。一个概念模式可以有若干个外模式,每个用户只关心与它有关的外模式,这样可屏蔽大量无关信息且有利于数据保护,因此对用户极为有益。数据库系统一般提供外模式描述语言用以描述外模式。

3. 内模式

内模式又称物理模式,它表示数据库的内部结构即其内部存储方式和物理结构,如数据存储的文件、块结构、索引、簇集、hash 等存取方式及存取路径等,数据库系统一般提供内模式描述语言用以描述内模式。

数据模式给出了数据库的结构框架,而数据库中的数据才是真正的实体,但是这些数据必须按框架所描述的结构组织,以模式为框架的数据库叫概念数据库(conceptual database),以子模式为框架的数据库叫用户数据库(user's database),以内模式为框架的数据库叫物理数据库(physical database)。这三种数据库中只有物理数据库是真实存在于计算机的外存中,其他两种并不真正存在于外存而是通过物理数据库由 DBMS 构造而成。

美国国家标准化协会(ANSI)在 1978 年的 ANSI /X3 / SPARC 最终报告中将数据模式分成为三级,它们是概念模式、外模式与内模式。其中外模式与内模式的含义和内容与上面讲的基本一致,而概念模式则与上面讲的概念模式则有一定区别,概念模式是一种抽象的模式,它不涉及具体的 DBMS 与具体计算机,它一般用概念模型中的 E-R 方法,扩充 E-R 方法或面向对象方法描述,而无法用具体数据库系统的语言描述,这是 ANSI 的报告与目前一般所公认的概念模式的区别。

数据模式的三个级别层次反映了模式的不同环境的不同要求,其中内模式处于最低层,它反映了数据的计算机外存储器的实际存储的形式,概念模式处于中层,它反映了设计者的数据全局逻辑要求,而外模式则处于最外层,它反映了用户的局部数据逻辑要求。

这三个模式之间可以有二个映射将概念模式映射至内模式以及将外模式映射至概念模式。