

数据库管理系统剖析

[美] James A. Larson 著

马克平 徐秋元 译

DBMS

西北工业大学出版社

2
16

数据库管理系统剖析

〔美〕James A. Larson 著

马克平 徐秋元 译

郭宜斌 校

西北工业大学出版社

数据库管理系统剖析

[美] James A. Larson 著

马克平 徐秋元 译

责任编辑 朱丹墀

*

西北工业大学出版社出版

陕西省新华书店发行

陕西省富平县印刷厂印刷

*

开本787×1092毫米 1/32 印张 6.875 字数 140 千字

1986年4月第一版

1986年4月第一次印刷

印数: 0001—4000册

统一书号: 15439·020 定价: 1.45元

译者的话

近年来,随着计算机在各个行业中的普及应用,从事系统分析、数据库管理、数据库设计及维护的广大科技工作者,都迫切需要更深入地了解数据库管理系统的内部结构。计算机专业的大学生和研究生也需要在掌握了数据库基本原理的基础上学习有关数据库管理系统构造的知识,以便更好地应用、设计和维护自己的数据库管理系统。译者认为本书的内容正好满足这些需要。

本书系作者在为大学高年级和研究生编写的讲义之基础上整理而成。作者 James A. Larson 是美国国家标准协会 ANSI/X3/SPARC/数据库关系任务组的成员及 CODASYL 数据描述语言委员会的成员,具有从事大型数据库管理系统研制工作的经验。作者以一个跨跃关系和网状模型的小型数据库管理系统为实例,对数据库管理系统的内部结构和实现方法进行了详细剖析。本书的特点是概念清楚,深入浅出。

参加本书翻译工作的有西北工业大学计算机系教师马克平同志(第一、二、三、四、五、七、八章)和徐秋元副教授(第六、九、十章)。全书由电子工业部第十五研究所数据库研究室主任郭宜斌同志进行了总校。西北工业大学研究生部朱丹墀同志对本书的译稿提出了许多中肯的修改意见。在此表示感谢。

由于译者水平有限，译文中错误和不当之处在所难免，
恳望读者批评指出。

译 者

1985年8月于西安

前 言

这是一本有关设计和建立数据库管理系统的书。它汇集了数据结构、编译原理、操作系统和软件工程诸领域的内容。作者尽力避免用数学方法论述数据库理论，而是以通俗方式描述数据库领域的最新研究成果及这些成果的实际应用。本书不要求读者具有数据库管理系统的预备知识。

本书通过设计和建立一个称作 BARE 的简单的数据库管理系统来使学生了解如何着手设计和建立数据库管理系统。本书前几章概述了 BARE 系统，这样可以使学生在习本课程不久就开始设计和实现自己的系统；后面几章详细介绍了扩充特性和供选用的特性。如果时间允许的话，学生可以自己实践这些内容。

第一章介绍了有关数据库管理的术语和概念的实用知识。第二章描述了 BARE 数据库管理系统，它支持网络和关系数据库管理系统的特性。第三章概述了 BARE 的实现过程，但保留了某些要求学生去完成细节。BARE 可以扩充，扩充之后所能提供的更优越的特性将在本书的其余章节中介绍。第四章涉及到数据库设计方法和自动帮助数据库管理员定义数据库的内容。第五章推荐了几种不同的描述数据库的方法。第六章描述了可用于访问数据库的不同的工具。第七章提出了一个数据库管理系统，这个系统可在单一的数据库中以各种方式描述和操纵数据。第八章讨论了数据库管理系统的发展和如何实现它们。第九章回顾了由螺帽螺栓问

题引起的数据的保密性、完整性、可用性和一致性问题。第十章使用一种抽象化结构总结了数据库和数据库管理系统。

有很多人对本书提出过好的意见和见解。明尼苏达大学1980年春季学期选修计算机科学5199课程的同学曾经修订过这本书的原始材料。值得特别提一下的是T. B. 威尔逊、H. R. 约翰逊和J. A. 劳伦斯，本书后几章中有很多新概念和新见解是他们发展出来的。

James A. Larson

目 录

前 言

第一章 数据库管理引论	1
1.1 企业	1
1.2 数据库	2
1.3 数据字典	4
1.4 数据模型	6
1.5 数据结构	10
1.6 抽象框架	12
1.7 数据库管理系统的任务描述	14
1.8 系统特征	19
1.9 小结	21
1.10 习题	22
第二章 BARE 数据库管理系统	23
2.1 数据库管理系统结构	23
2.2 BARE数据模型	23
2.3 BARE数据描述语言	27
2.4 BARE数据操纵语言	31
2.5 把数据操纵命令嵌入宿主程序设计语言	37
2.6 对数据库的交互式访问	40
2.7 小结	41
2.8 习题	41

第三章 BARE 数据库管理系统的实现44

3.1 元语言.....	44
3.2 BARE 数据描述处理程序	47
3.3 信息的表示	57
3.4 BARE数据字典描述.....	73
3.5 小结	78
3.6 文献书目注记	79
3.7 习题	79

第四章 数据库设计.....82

4.1 需求说明	82
4.2 概念化和调整	83
4.3 协调	94
4.4 自动设计工具	95
4.5 小结	97
4.6 文献书目注记	97
4.7 习题	97

第五章 描述数据 101

5.1 文件数据模型	102
5.2 关系数据模型	105
5.3 CODASYL网络数据模型	114
5.4 小结	129
5.5 文献书目注记	130
5.6 习题.....	130

第六章 操纵数据 131

6.1 面向线性关键字的语言	131
----------------------	-----

6. 2	非线性二维语言	142
6. 3	小结	147
6. 4	文献书目注记	147
6. 5	习题	147
第七章	多模型数据管理	150
7. 1	关系与网络数据库的比较	150
7. 2	多模型数据库管理系统	151
7. 3	网络数据的关系视图	152
7. 4	实现任选	155
7. 5	小结	158
7. 6	文献书目注记	158
7. 7	习题	158
第八章	数据库系统的体系结构	161
8. 1	三级数据体系结构	162
8. 2	多模型数据库管理	164
8. 3	其它模式类型	166
8. 4	数据体系结构的演变与发展	171
8. 5	一种典型的数据体系结构	173
8. 6	命令的联编	176
8. 7	小结	179
8. 8	文献书目注记	180
8. 9	习题	180
第九章	数据的完整性与保密性	182
9. 1	数据的可用性	182
9. 2	数据的合法性	188
9. 3	数据的一致性	189

9.4 数据的保密性	194
9.5 小结	197
9.6 习题	197
第十章 数据库管理系统的结构框架	199
10.1 数据库管理设计决策	199
10.2 逐步求精方法	201
10.3 结构框架的约束	203
10.4 灵活的数据库系统	204
10.5 小结	205
10.6 习题	205
参考文献	206

第一章 数据库管理引论

数据库管理系统涉及人、数据库以及一个允许人们对数据库进行存取的系统。数据库中含有描述人们感兴趣的信息的数据。借助数据库管理系统，人们可以对数据库中的数据插入、删除、修改和检索。

这个能处理与希望存取数据的人们的利益有关的数据的神秘系统究竟是怎样构成的呢？数据库管理系统有哪些功能？它是如何工作的？

1.1 企业

为了完成靠个人单独做所不能完成的事情，人们组成了企业。一个企业是使用共同的资源和工具为一个目的而工作的一群人。一个师范学院、一支保安部队、一所医院、一个民航公司、一个玩具制造厂商、一个食品分配组织和一个犯罪集团都是包括使用共同的资源为一个目的而工作的人们的企业的例子。这些企业的资源可以是教室、小汽车、医疗设施、飞机、未浇铸的石膏、装小麦的箱子及偷来的电视机。企业与企业的目标是不同的，大部分私营企业的目标是通过生产商品和提供服务来积累资金，而大部分政府的企业的目标是提供服务和把从纳税者那里收集来的钱用于公共事业。

企业又可以是一更大企业的一个组成部分。例如，大商店中的玩具部和鞋部可以被看作独立的企业，各自都在努力完成自己的销售额。玩具部和鞋部，可能还有其它一些部可

以合起来组成一个叫作公司的较大企业。作为整体的公司也有它自己要达到的目标，例如在财政年结束时的总利润。几个公司可以组成一个大公司，大公司照例又有它自己的目标；这些目标是组成大公司的各个小公司的目标的结合。每个分部、每个小公司或大公司都可视为一个努力实现它自己的目标的企业。所以，当我们谈论企业时，精确地给“企业”下定义是重要的，这使人们不致于对企业是什么和它如何与其下级组织和上级组织发生联系等概念产生混淆。

随时了解一个企业的所有资源情况可能会是个难题，特别对于大企业更是如此。资源和人力是否得以充分利用这一点通常可以区分达到目标获得成功的组织和因没有达到目标而濒临破产的组织。大部分企业有一个负责资源和人力的利用的管理组织。管理者的任务包括弄清资源的位置和总量，并分析可供选择的资源的利用等。已研制出了用于帮助管理者完成这些任务的多种工具，其中就包括电子计算机。

电子计算机可用于存贮与企业利益有关的信息。这些信息可由计算机组织、概括、提炼和插入，并提供给企业经理。用这些结果，经理可以对企业的经营作出决策。以计算机内部数据所表示的信息可认为是如同设备、资金和雇员等企业资源同等重要的资源。在企业的经营中对表示为计算机内部数据的信息的存取非常频繁。这些信息对要作出有关企业前途的决策是十分重要的。为了改善经营，许多企业需要这些逐日统计的信息。必要时这些信息的利用可给企业带来许多好处，并可很好地确定成功与失败之间的差异。

1.2 数据库

描述与企业利益有关的信息的数据的集合叫做企业的数

数据库。数据库并不总是驻留在可由计算机快速存取的电子或磁性存贮装置内的。一张处方卡片文件、一个书库中的书、史前期山洞里墙壁上的图案以及某个成年人的孩提时代的回忆和故事都可认为是数据库。然而，在本书中我们仅涉及那些可由电子计算机快速存取的数据库。

一个数据库是一个以上的存储于计算机中的数据集合。数据必须以某些易于以任何所希望的数据段存贮的结构来组织。正如排列书库中书架上的图书那样，计算机中的数据也必须被有条不紊地安排与组织。负责数据库中数据组织的人叫数据库管理员，他不仅要决定如何存贮数据库中的数据，而且也要决定存贮哪些数据以及谁可以存取这些数据。

有必要使用一些工具来帮助数据库管理员管理数据库中的数据。数据库管理系统就是用来描述、插入、删除和检索数据库中的数据的通用软件。数据库管理员用这个系统描述数据库中的数据并确定存取数据的规则。数据库用户按数据库管理员所指定的存取规则使用数据库管理系统来插入、删除、修改和检索数据库中的数据。数据库管理系统是管理企业的主要资源之一——由存贮在计算机中的数据所描述的与企业利益密切相关的信息——的一种重要工具。

数据库由描述客体、概念和与企业利益相关的事件的各种数据所组成。每个数据都可说是它所描述的企业客体、概念或它所表示的事件的一个外延。因而数据库又是由客体、概念或事件的外延所组成的，这些外延可以由数据库管理系统的用户检索、修改、插入或删除。本书中的术语企业数据和企业的企业外延是等价的。

数据库中的企业外延是随时间而变化的；典型的是数据库都随数据的增加而变化。在数据库的设计阶段，在数据库

管理员手中通常不掌握数据库生存期中所有外延的集合。所以，最初放进数据库中的企业外延标识是数据分析的唯一依据。这样的数据分析不能得到改变后值的结果。因此，通常的作法是用一种比较抽象的方式去考察一个企业，将企业的内容归为实体、概念和事件。

1.3 数据字典

正如一个成长中的孩子要将世间的事物分类一样，数据库管理员也必须将企业外延进行分类和组织，以使得更容易地管理和综合数据库。企业外延可按实体、概念或与企业利益相关事件为类别分组。这样的类别叫意图（intention），例如，企业外延 TONSILLITIS（扁桃体炎）、HERNIA（疝）和 INFLUENZA（流感）可以抽象为企业内涵 ILLNESS（疾病）。

识别企业外延中的相似性乃是组织数据时所使用的 important 方法。一个设计得好的企业内涵一般不应在数据库引进新的外延时受到影响。

企业的外延就存贮于该企业的数据库中。企业的各种内涵则存于与数据库分离的贮藏室中。如果这个贮藏室仅可由数据库管理系统访问，它通常就称作数据目录。若它也能由系统的其它部分访问，则通常称它为数据字典。数据字典这个术语隐含它不但存贮关于数据库的信息，也可能含有其它系统所使用的字典信息。本书以后将经常用到数据字典这个术语。

数据字典对于数据库管理系统来说是很重要的。它含有数据库中所有数据的描述和定位信息。数据库管理系统用数据字典使对数据库中的数据提出的要求生效。数据字典含有

能用于确定是否要求的数据存在以及数据存贮在数据库中的那个位置上的信息。数据字典也用于确定要求是否能够实现。

数据字典也可以包含一些附加的信息。这些信息可供很多其它的系统部分使用。例如，数据字典可以含有：

▲用户轮廓。计算机的各级用户及这些用户所拥有的特权的描述。例如，仅允许初学的用户检索数据，而允许高级用户修改数据。每个用户级别中可有一个或多个用户。

▲程序名。每个存取数据库中数据的程序、模块、子程序或软件部件的名称以及它们的状态（例如，可以是开发的、测试的、生产的或是已废弃不用的）。

▲计算机各级用户间的关系。例如，允许访问所有工资单数据的用户仅是允许访问雇员姓名和被雇用日期的用户的子类。

▲程序名间的关系。例如，工资单计算程序可以调用分类和检查打印子程序。

▲数据类型间的关系。例如，雇员记录包含雇员名和被雇用日期数据字段。

▲程序和数据间的关系。例如，工资调整程序要检索雇员记录并仅修改那些记录的工资数据字段。

▲各级用户和程序间的关系。例如，某级用户可以调用程序X和Y，但它仅能修改程序X。

▲各级用户和数据间的关系。例如，A级用户负责定义雇员记录类型，而B级用户负责确定雇员记录中的数据字段的值。

数据字典中不仅含有管理数据库中的数据的信息，而且也含有管理用户和程序的各种信息。程序管理系统用于控制

软件的开发和维护。它起着编目和存贮每个软件模块的所有版本的图书馆的作用。安全系统用于控制计算机用户对于程序管理系统中的程序模块和数据库中的数据可以做什么和不可以做什么。程序管理系统和安全系统可以不包括在数据管理系统中,然而,这三个系统可以共用一个共同的数据字典。

1.4 数据模型

模型这个术语常用于说明一个现实世界系统的抽象。例如,经济学家研究描述经济领域的理论模型。这个模型把它所模型化的系统的某些方面孤立起来,并试图解释这些方面,甚至对未来事件进行预测。这些理论模型可能包括一些对现实世界系统模拟或模型化的复杂公式。虽然计算机经常被用来对这些理论模型进行处理,但本书要讨论的却是另一种模型,即数据模型。这是对数据库的数据进行描述和操纵的考察的方式。数据模型的许多形式化定义都是基于诸如关系、集合和图论等数学结构的。某些形式定义包括按这些数学结构可实现的运算,而其它形式定义不描述数据操纵命令。因为我们对数据模型的兴趣包括修改由这种数据模型所描述的数据,所以我们把数据操纵看作数据模型的一个重要方面。

不同的数据模型在描述数据实体和数据实体间的关系方面的格式是有差别的。它们在对所表示的数据实体的限定和对所表示的数据库进行的操作方面的格式有所不同。

已提出了许多种数据模型,每一种都有其自己的概念和术语。每种数据模型都是由表示企业的各种语义和结构的特定的构型能力的要求所推动而产生的。每种模型对应用范围都有所限制,看来还没有一种数据模型能满足所有应用要求。

数据模型的使用有几个重要原因。数据模型有助于交流