

面向对象数据库和关系数据库的比较

杨春才 徐铭欣 钟晓星

(中国人民解放军 63653 部队 新疆 841700)

[摘要] 对传统的关系数据库和面向对象数据库的数据模型、数据库系统进行对比分析,总结其各自优势及应用领域,指出在需求日益发展的今天,面向对象数据库相对于关系数据库的优势所在,并探讨其发展前景。

[关键词] 面向对象数据库 关系数据库 数据模型

[分类号] GB250.7

自数据库技术诞生以来,传统的数据模型共有三种类型:层次模型、网状模型和关系模型。以层次型和网状型为主的第一代数据库系统,使用了记录集合上的数据描述语言和数据操作语言,这些数据语言的过程化程度高,一般用户使用困难。以关系型为主的第二代数据库系统,使用了简单的二维表和非过程化的数据操作语言,具有高度的数据独立性,从而成为 20 世纪 70 年代到 80 年代中期数据库系统的主流。但是,在计算机技术日新月异的发展过程中,要求数据库系统处理的已不仅仅是一些简单的数据类型,还必须能够处理一些诸如图形、图像、声音等复杂的数据类型。在这种情况下,一种新型的数据模型——面向对象的数据模型诞生了。它的问世极大地推动了数据库技术的发展。

1 两种数据库的比较

1.1 数据模型

1.1.1 模型的基础 关系模型是建立在集合代数基础上的,具有严格的数学基础、较高的抽象级别,单一的数据结构便于用户的理解和使用,也有利于机器的实现与软件的管理。但是,过于简单的数据结构在描述复杂的现实世界方面又存在着明显的不足。

面向对象数据模型是建立在面向对象的程序设计方法基础之上的。面向对象方法是以对象为中心的方法,即在系统中用对象来表示现实世界中的任何实体。对象具有状态和行为;对象的状态由一组称为属性的值来描述,而对象的行为用一组在状态上进行搜索的运算来表示。对象将数据和有关操作封装在一起,更好地实现了数据的抽象。面向对象技术通过对现实世界中对象的抽象以及对各个对象之间的相互关系、相互作用的描述,来实现由现实世界到目标系统的映射。

这两种数据库最主要的区别在于,面向对象的数据模型使用了许多面向对象的概念,如对象、封

装、继承、多态等;而关系模型包含的概念是关系、属性、记录等。

1.1.2 实体的描述 关系模型用关系描述实体,它的基本特征是用二维数组表示所有实体,实体之间的联系也通过表示实体联系的关系或公共属性来体现。因此,对于最基本的实体间的联系描述不是显式的,而对公共属性的定义也不能从数据结构中反映出来。单从这种简单、单一的数据结构上不能区分出哪些是表示实体的关系,哪些是表示实体间的联系的关系。并且缺乏对实体间的联系的性质的描述,具有不同语义的实体间的联系被不加区别地用相同的数据结构来描述。

面向对象的数据模型将现实世界的任何实体都看作一个对象,它有唯一的标识符,叫对象标识。每个对象有一组属性表示它的状态,由若干方法描述它的行动,对象将属性和方法封装在一起,外界通过对象提供的方法对其进行操作,而对象与对象之间的互动是通过消息传递实现的,消息会调用适当的方法来执行或访问对象。概括起来,面向对象的数据模型不对每个实体一定义,它只定义对象的类,根据类定义对象的属性和方法。

1.2 数据库系统

1.2.1 基本结构 关系数据库的基本结构是记录,一个记录包含若干条属性;面向对象数据库的基本结构是对象。对象与记录的概念相似,但内容要比记录复杂得多,一个对象不仅包括描述属性的数据,而且还包括操作方法,每个对象由唯一的一个标识符把状态和行为封装起来。

1.2.2 支持的数据类型 关系数据库支持的基本数据类型包括整型、实型、字符串型、布尔型、日期型、货币型等等。目前关系模型还不能直接支持复杂的数据模型,只能表示结构化的数据,随着应用的日益发展,非结构化或半结构化的数

据也占有越来越重要的地位。

在面向对象的数据库中, 提供了可扩展性的方法, 允许用户使用程序设计语言定义新的数据类型, 因此面向对象的数据库不仅包含关系数据库中的基本数据类型, 还包括诸如图形、图像、声音、视频和动画等多媒体数据。

1.2.3 模式修改 在关系数据库中, 主要的模式修改有: ①创建或删除一个关系; ②在关系模式中增加或删除一个属性; ③在关系模式中, 修改完整性的约束条件。由于关系与关系之间是相互独立的, 对一个关系模式的修改仅限于一个关系, 并不牵扯其他的, 因此模式修改操作比较简单。

在面向对象的数据库系统中, 主要的模式修改有: ①类集的改变; ②已有类成分的改变; ③子类/超类成分的改变。面向对象模型包含有类的定义、类属性之间及类之间丰富的语义联系, 修改模式就有可能影响模式的一致性。例如: 删除一个类时, 不仅删除了这个类本身, 它所有的子类也都要被删除。因此, 对面向对象的数据库模式修改, 要比关系数据库的模式修改复杂得多。

1.2.4 查询操作 数据库查询是数据库的核心操作。关系数据库的查询操作有选择、投影、连接等, 查询结果是从一个或多个关系中得到一个新的关系。而且, 存取路径对用户而言是透明的, 查询是非导航式的。

在面向对象的数据库中, 查询操作是基于类进行的, 访问数据的操作是通过对象标识(指针)的移动进行的。由于封装性和对象标识, 在面向对象的数据库中查询是导航式的, 因此不易进行关系数据库那样的高层次描述性查询。按面向对象的观点, 任何对象的属性都要通过方法来访问, 如果仅为了提高查询的效率, 允许查询过程直接访问对象属性, 那么就有可能破坏了对对象的封装性。再者, 面向对象的数据库不像关系数据库那样, 它缺乏形式化的数学理论基础, 进一步的分析和演算都无法进行。

1.2.5 事务处理 在关系数据库中, 事务处理包括并发控制、从故障中恢复等。事务对数据的处理一般在几秒钟之内完成; 但对于面向对象的数据库来说, 它应用于诸如计算机辅助设计等新的领域, 事务处理可能需要很长时间, 关系数据库中的事务处理方法对它已不合适, 因此面向对象的数据库需要增强事务处理的功能, 如长事务处理等。

2 面向对象数据库相对于关系数据库的优势

根据以上的比较不难看出, 关系模型的数据结构简单, 容易理解和掌握, 对实体与实体之间的联

系描述简明、精确, 给用户使用数据库提供了很大的方便。其查询语言定义的集合操作, 语言精练, 功能强大, 而且存取路径对用户隐蔽, 这样就大大提高了数据的独立性。由于关系数据库表达的数据结构以及实体间的联系比较单一, 因此适用于人事档案管理、商业管理和仓库管理等简单的数据库管理系统。但是, 关系数据库不能表示复杂的对象, 也不具备定义复杂操作的能力。因此, 根据目前数据库技术发展的趋势, 关系模型还应扩充, 与面向对象模型相结合, 以提高其功能, 扩展应用领域。

面向对象数据库提供了比关系数据库更优越的数据模型, 能够支持关系模型无法支持或支持力度不够的复杂应用, 增强了程序的可设计性, 提高了导航访问能力, 简化了并发控制。面向对象数据库适用于复杂应用领域, 如需要表示抽象数据类型如地图、人口密度、植被、水源、建筑物等数据的领域。而且他对实体的描述比关系模型更加现实、自然, 也更加直观。面向对象数据库相比于关系数据库的优势主要体现在以下几点:

2.1 支持复杂应用

面向对象模型主要用于复杂应用, 应用中数据的复杂性越高, 数据间相互的关系越复杂, 性能的提高也越大。相关对象的寻找可通过类层次或其他相关关系来完成, 将特定对象放入高缓冲区或内存的技术, 可通过预测用户或应用程序可能存取的类及其实例而得到优化。当数据复杂性较高时, 聚集和缓冲技术对性能的提高更显著, 而这正是关系数据库无法达到的。

2.2 存储大型数据结构

面向对象数据库不仅能存储复杂的应用程序, 而且还能存储较大的数据结构。尽管关系数据库支持大量的元组, 但单个元组的大小受到限制, 虽然有些关系数据库在这方面有所放松, 但带来了数据库重组与管理低效等问题, 面向对象数据库不会因为大量对象的存在而降低了性能。因为不管对象的特性有多复杂, 应用程序都没有必要把对象分离或装配。

2.3 直接引用对象

面向对象的数据库支持对象的直接引用, 这样做不仅减少了系统的数据冗余, 提供了数据共享能力, 有利于数据完整性维护, 同时还大大提高了搜索和导航访问能力。在关系数据库中则相反, 复杂的数据集必须由应用程序组装, 如低效的连接运算。

2.4 长事务处理

在传统的意义下, 设计事务不是原子的。因为

事务具有较长的持续时间, 所以事务的部分结果应该是可见的。在事务进行阶段, 系统可以通过发布一个修改令牌来实现单一设计者对数据的更新, 也可以通过发布多个浏览令牌来实现多个设计者对数据的同时浏览。系统能够实现对系统崩溃点的保存, 进行系统恢复。除此之外, 面向对象的数据库将对不同版本的支持和事务管理集成于一体, 以此来简化当前的存取和恢复。

2.5 优良的应用开发环境

从应用程序的开发环境来看, 面向对象数据库又显示出其优势。关系数据库的应用开发离不开数据操作语言 DDL 和相应的宿主语言(如 C、Pascal 等), 程序员必须同时掌握这两种语言, 两者缺一不可, 否则不能建立完整的应用程序; 而面向对象数据库的应用开发并没有这些限制。

2.6 简化并发控制

在关系数据库中, 并发控制理论已经很成熟了, 但实现起来却比较复杂, 应用程序必须显式地对数据进行封锁, 封锁类型也要考虑。对某表或其中一部分元组的封锁往往带来对相关的其他表或元组的封锁; 而面向对象数据库的并发控制以对象为封锁单位, 相关数据由对象本身的结构决定, 一些控制可由对象的方法和触发器完成, 并发控制简单而有效。

2.7 完整性

面向对象数据库可以更好地支持完整性。在纯面向对象数据库中, 由于数据和过程被封装在一个对象中, 因而对一个对象的修改影响到数据库中其他对象完整性的可能性比较小, 应用程序可以在一个操作中锁住所有的相关数据, 保证引用的完整性。它通过指针对数据本身进行维护和修改, 以保证更好的支持完整性。而关系数据库却不能保证引用完整性。

2.8 直观性和人机交互性

在符合用户习惯方面, 面向对象数据库又显示出相当的优势。关系数据库中的二维表结构虽易于实现, 但在用户方却不是一种直观模型。而图像、图形、声音、文字是自然界与人类社会中常见的信息形式, 但这些常见的信息形式又不易用纯关系模型来描述与表达。面向对象数据库使用对象来描述现实中的实体, 对图像、图形、声音、视频等信息形式都能提供较自然和完整的模型。

3 面向对象数据库的发展现状

3.1 面向对象数据库的发展策略

近年来面向对象数据库的研究和开发取得了长足的发展, 在数据库系统中引入面向对象的概念、方

法和机制, 有多种策略和方法, 归纳起来为以下几个方法:

- 不借助于任何现有系统, 开发全新的面向对象的数据库程序设计语言 OOPBPL。
- 用数据库能力扩充 OOPL, 使其成为面向对象的数据库程序设计语言。
- 用面向对象能力扩充现有的数据库查询语言, 使之成为对象结构化查询语言。
- 提供可扩充的面向对象数据库管理系统的类库。
- 在 C++、SMALLTALK 等宿主语言中加入面向对象数据库语言结构。
- 将面向对象机制加入关系数据库, 使其扩充成面向对象数据库。

3.2 面向对象产品的研制和开发上存在着两大派别

一种是以现有的关系数据库为基础, 加入面向对象的特征, 即所谓的对象-关系数据库。对象-关系数据库派认为关系数据库具有成熟的理论基础, 主张对现有的关系数据库进行扩充和改进, 使之升级为对象-关系数据库。具有代表性的数据库产品有: DB2、SMALLTALK/SQL、DBKIT、COMMOMBASE、ORACLE8、SQLSERVER7.0 等。另一种是以面向对象技术为基础, 引入数据库技术, 即所谓的“纯”面向对象数据库。纯粹的面向对象数据库派则主张进行彻底的数据库革命, 即采用全新的数据模型和模式, 抛开现有的数据库系统, 从最底层做起, 使之成为真正的、纯粹的面向对象数据库。其代表性产品有 OBJECTSTORE、ONTOSDB、VERSANT、IRIS 及 ORION/ITAS-CA 等。无论是对象-关系数据库, 还是纯粹的面向对象数据库, 面向对象的概念和方法是他们不可缺少的组成部分。究竟哪一个更适合于存储和访问复杂的数据, 具有更优越性的性能, 理论界和工业界还在争论, 有待于在实际应用中加以比较和校验。

4 结语

从数据库技术的发展方向可以看出, 面向对象数据库具有更强、更自然的建模能力, 能对存储和管理复杂数据结构的信息提供有效的支持, 并能将数据和操作进行一致的管理。但是面向对象理论还不够成熟, 面向对象的数据模型目前还不象关系模型一样有完美的数学基础, 其数据库语言还缺乏形式化的基础, 如何将成熟的关系数据库技术, 如事务控制、完整性检查、数据复制等应用到面向对象数据库中, 这些问题均有待于进一步地研究和完

(下转第 134 页)

供了《华南理工大学各类型阅览室及分馆文献典藏暂行规定》、《华南理工大学中文图书分类标引细则》和《华南理工大学图书馆中文图书著录细则》等文档,让他们充分了解我馆的编目规则和工作方式,并由有经验的编目员对来我馆工作的外包公司人员进行岗前培训,让他们尽快熟悉 Horizon 系统编目模块的使用,使他们能对我们的具体要求有一个感性认识,这样我们的要求会给他们留下深刻的印象,对以后的顺利合作起到积极作用。合作中的沟通也非常重要,他们的疑问我们随时解答,我们的细节要求可以随时提出,协商解决,尽量满足,使整个工作达到和谐顺利。

3.3 进行严格的质量控制,确保编目外包达到预期目标并取得最大效益

在编目过程中进行质量控制是保证编目外包实现目标最重要的环节。没有严格管理的编目外包是具有很高风险的。可能会出现外包公司服务差,不能兑现所承诺的服务,如有的不会做医学类图书分类和医学主题标目、编目质量不能保证等问题。在编目外包中坚持监管,实行全程的质量控制,使外包的工作取得最大效益。

我馆派专门的编目员对外包公司的书目数据进行校对,这项工作的初期是非常困难的。外包公司初期有的不熟悉编目业务,有的不熟悉我馆规则,有的是由于初到一个新环境的紧张,造成书目数据质量非常差,校对人员的工作量非常大,不仅要检查质量,还要指明出错原因、改正理由,再让他们改正错误,然后再检查。有时同一种书更正了三、

四次还有错误出现。我馆编目员甚至发出了“还不如我们自己做数据效率高!”的感慨。虽然如此,但是我馆坚持高标准,高质量地把关,不放松监管,不降低要求,经过三、四个月的磨合,外包公司适应了我馆编目规则和工作流程,质量渐渐地提高了,效率渐渐地上来了,我馆编目人员才开始体会到业务外包的优势。我馆并没有因为外包公司数据质量提高了就放松了对质量的监控,全程严格地控制质量是我馆编目外包达到预期目标并取得最大效率的保证。在此基础上,我馆人员与外包人员同心协力进行业务钻研,共同将我馆的编目工作推向一个新的水平。

参考文献:

- 1 朱复成. 美国图书馆业务外包的得失及启示. 大学图书馆学报, 2000(1):22-23
- 2 孙丽芳. 信息技术外包策略在我国图书馆中的应用探讨. 情报杂志, 2002(8):15-16
- 3 万荣花. CALIS 联合目录数据库的质量. 图书馆论坛, 2004(4): 18-19
- 4 姚多岚. 高校扩招对图书馆的影响及其对策. 大学图书馆学报, 2002(9):23-25
- 5 招生信息. 中国教育招生考试网. [2006-01-17]. <http://www.gxzhaozheng.com>
- 6 Clare B. Dunkle. Outsourcing the Catalog Department: A Mediation Inspired by the Business and Library Literature. Journal of Academic Librarianship, 1996,22(1):33-43
- 7 印洪斌. 我馆图书业务外包情况介绍. 高校图书馆情报论坛, 2005(6):16-18

[作者简介] 万荣花,女,1967年生,馆员,发表论文数篇。

(上接第152页)

善。随着技术的不断发展和完善,面向对象数据库的研究会更加深入,它必将给数据库领域带来更加辉煌的发展前景。

参考文献:

- 1 袁俊红,罗梅等. 面向对象数据库的探讨. 信阳师范学院学报, 2002(7):360-362
- 2 应宏. 从 RDBMS 到 ORDBMS. 计算机应用研究, 1999(3): 36-37
- 3 李铮. 面向对象数据库的研究与发展. 计算机与现代化, 2000(4):40-43
- 4 星光. 任重道远——对象数据库. 计算机世界, 1998(3):40-43
- 5 张力,慕晓东. 面向对象的数据库和关系数据库的比较. 武警

工程学院学报, 2001(8):27-28

- 6 李振华. ORACLE 空间数据库的对象——关系模式初探. 地球科学-中国地质大学学报, 2002(5):333-337
- 7 陆晶. 面向对象数据库系统的特性及其发展现状. 管理信息系统, 1999(9):38-41
- 8 郑刚. 面向对象数据库中查询处理技术的研究. 微机发展, 2002(5):37-41
- 9 马春光,肖冬梅. 面向对象数据库通用性研究. 齐齐哈尔大学学报, 2000(9):66-69
- 10 刘义英,郝忠孝. 将关系数据模型转换为对象数据模型的研究. 哈尔滨理工大学学报, 2001(10):6-9
- 11 邵维忠,廖钢城等. 面向对象的设计. 北京:北京大学出版社, 1994

[作者简介] 杨春才,男,1966年生,副研究馆员,发表论文4篇。徐铭欣,女,1982年生,馆员,发表论文1篇。钟晓星,女,1979年生,馆员,发表论文2篇。