

基于面向对象技术和标准关系数据库的 Geodatabase 的研究

刘艳丰, 姜耀鹏

(中南大学 信息物理工程学院, 湖南 长沙 410083)

摘 要:面向对象技术为空间数据模型的发展提出了新的要求。本论文在此背景下,回顾了空间数据模型的发展历程,从整体上介绍了第三代基于面向对象技术和关系数据库的 Geodatabase 数据模型,并详细介绍了其体系结构及在关系数据库中的存储。最后本文通过总结 Geodatabase 模型的优缺点,提出了未来空间数据模型的发展趋势。

关键词:面向对象;空间数据模型; Geodatabase; 关系数据库

中图分类号: TP311.132 P208 **文献标识码:** B **文章编号:** 1672-5867(2009)04-0139-04

The study of Geodatabase Based on Object-Oriented Technology and Standard Relational Database

LIU Yan-feng JIANG Yao-peng

(School of Info-Physics and Geomatics Engineering, Central South University, Changsha 410083, China)

Abstract: The object-oriented technology proposes new requirements for the development of spatial data model. Under this background, this paper reviews the development history of spatial data model, introduces the third generation spatial data model—Geodatabase based on Object-Oriented technology and standard relational database as a whole, and explicates the system structure and storage method in relational database of Geodatabase. After discussing the advantages and disadvantages of Geodatabase, this paper comes to a conclusion that the development trend of the future space data model is to realize the object-oriented space data physical model.

Key words: Object-Oriented; Spatial data model; Geodatabase; relational database

0 引 言

地理信息系统 (GIS) 是一种采集、存储、管理、分析、显示与应用地理信息的计算机系统, 既是表示、模拟现实空间世界和进行空间数据处理分析的工具, 又是一门关于空间信息处理分析的科学技术^[1]。GIS 将地理数据划分为空间数据和属性数据两种类型, 目前, 大多数商业关系数据库不支持空间数据, 因为该数据库模式是以 2 维表格为基础, 不支持变长数据类型, 也不支持用户自定义数据类型, 无法表示像空间数据这样的复杂对象。空间数据只能存储在二进制索引文件中, 属性数据存储 in 关系数据库中, 它们之间通过用户码 (ID) 关联起来。这种分开存储的方式不利于空间数据库的整体管理, 数据的一致性也容易遭到破坏, 查询速度慢, GIS 的开放性、可扩张性与可维护性都受到了限制。因此, 迫切需要开发一种能集空间数据和属性数据与一体的空间数据库管理系

统。面向对象技术的出现, 为解决这个问题提供了契机^[5]。

1 空间数据模型及其发展历程

空间数据模型是关于 GIS 中空间数据组织的概念, 反映现实世界中空间实体及其相互之间的联系, 为空间数据组织和空间数据库模式设计提供了基本的概念和方法, 是空间数据库的基础和核心, 也是地理信息系统研究的核心。因此, 对空间数据模型的深入研究直接影响着新一代 GIS 系统的发展。随着地理信息系统软件技术的不断发展和完善, 空间数据模型的发展经历了从 CAD 数据模型、Coverage 数据模型到 Geodatabase 空间数据模型^[2-6,9] 的转换。

1.1 基于文件系统的 CAD 数据模型

在 20 世纪 60 年代和 70 年代, 地图主要由计算机辅助设计软件 (CAD) 绘制。CAD 数据模型以二进制文件存

储以点、线、面表达的地理数据。这些文件缺乏对属性信息的描述,地图图层和注记符号是主要的属性表达方式。空间数据之间也不能建立拓扑关系或实现空间分析。这种模型采用文件系统来存储和管理空间数据,便于操作。然而,因其不能完全表达空间要素的属性信息,不能建立拓扑关系或实现空间分析,便在地理信息系统的发展中被逐步淘汰了。

1.2 基于文件与数据库混合方式的 Coverage数据模型

1981年 ESR公司推出其第一个商业化软件—— ArcInfo,它采用了 Coverage数据模型,该模型是基于地理关联的矢量数据模型,因此也称地理相关数据模型。它空间数据存储为索引化的二进制文件,并为显示和存取做了优化。属性数据存储在表里,表的行数等于二进制文件中的要素数量,属性数据和要素通过共同的标识符 (ID)相连接,实现了空间数据和属性数据的结合;另外,它存储了矢量数据的拓扑关系,这就意味着一条线段的空間数据记录包含了以下信息:哪些节点界定了这条线,通过推断可知该线与哪些线相连,以及哪些多边形在其左边或右边。所以说 Coverage数据模型提高了对地理空间信息的表达能力和数据的分析能力。

Coverage数据模型强调了几何特征和空间要素的拓

扑关系,并兼顾了空间数据和非空间数据的特点,使得高性能的 GIS成为可能。但该模型的不足之处表现在属性数据存储在关系数据库中,而空间数据单独存在二进制文件中,在操作空间数据时不得不与一个个二进制文件对话,这样不仅增加了数据转换成本,也限制了海量数据的存储与管理。另一个缺陷是要素由类似的点、线、多边形集合聚合而成,不具有真实世界对象的特殊行为,在描述丰富复杂的地理对象时很受限制。

1.3 基于面向对象技术和关系数据库的 Geodatabase数据模型

Geodatabase是 ArcInfo8的创新,是采用面向对象的思想而提出的一种适用于关系型数据库管理系统的、统一的、智能化的空间数据模型,它将现实世界抽象为若干对象类 (Objectclass),对象类本身并没有空间特征,但每个对象类都有其属性、行为和规则;相同属性集、行为和规则的空间对象集合体现为要素类 (Featureclass);要素数据集 (Feature Dataset)是具有相同空间参考特征的要素类集合;关联类 (Relationshipclass)是 Geodatabase定义的用来描述要素类或对象类之间的关联关系;同时,这种数据模型还定义了拓扑一致性、属性有效范围的域 (Domains)以及要素类行为取值的约束规则 (Rules)等。Geodatabase内部结构见表 1。

表 1 Geodatabase内部结构
Tab 1 The interior structure of Geodatabase

要素数据集	要素数据集中的所有要素必须具有相同的坐标系统。因为在要素数据集中存储了 Geodatabase的拓扑关系,而
空间参考	空间参考又是维护拓扑关系的关键。要素数据集中可以存储对象类、要素类和关联类。
对象类,子类	对象类中存储的是一组类型相同的对象;要素类存储的是同种类型要素的集合。对象类不包含空间实体,而
要素类,子类	要素类却相反。
关联类	关联类存储了对象类和要素类之间的关联信息。关联可以是对象类之间、要素类之间或者是要素类与对象之间的。关联类可以将要素和对象关联到一起。对象类、要素和关联类可以直接存储在 Geodatabase中,而不是必须存放在要素集中。
几何网络	几何网络是用来模拟线性系统,如道路交通网络等,以支持丰富的网络跟踪和分析功能。
平面拓扑	拓扑表达地理对象之间的相邻、包含、关联等空间关系。一个拓扑关系存储了三个参数规则 (rules),等级 (ranks)和拓扑容限 (cluster tolerance)。
域	域是对象属性的有效值集合。其类型可以是文本型,也可以是数值型。
验证规则	验证规则是通过关系和连接性规则确保数据完整性的规则。

Geodatabase数据模型中融入了面向对象核心技术,因而具有封装性、继承性和多态性等特征。其将所有空间地物以对象的形式封装,并将对象外部行为语义和内部执行之间显著分离,根据操作行为来定义封装;继承性是指对象类的定义能够包含其他对象类的行为以及拥有新增的行为,通过它们很容易得到空间目标间的关系;面向对象的可扩充性,使得 Geodatabase数据模型的预定义类型无需明显改变即可成为用户类型。

2 Geodatabase数据模型在关系数据库中的存储

Geodatabase是一种采用标准关系数据库技术来表现

地理信息的数据模型,支持多种 DBMS结构和多用户访问,且大小可伸缩。目前,有两种 Geodatabase结构:个人 Geodatabase和多用户 Geodatabase。

1个人 Geodatabase对于 ArcGIS用户是免费的,它使用 Microsoft Jet Engine数据文件结构,将 GIS数据存储在小数据库 Access中,数据库存储量最大为 2 GB只允许单用户编辑,不支持版本管理。

2多用户 Geodatabase通过 ArcSDE支持多种数据库平台,包括 IBM DB2、Informix、Oracle(有或没有 Oracle Spatial)都可以和 SQL Server。多用户 Geodatabase支持海量的、连续的 GIS数据库;支持多用户并发访问;支持长事务和版本管理的工作流。

Geodatabase 是一个多层的体系结构, 由 GIS 应用软件和关系数据库共同完成地理数据的管理。实际上, DBMS 是专门用来存放地理数据的, 而不是用来定义地理数据的行为。数据的存取是通过存储层 (DBMS), 由简单表来实现, 而应用层 (GIS 软件) 则通过定义 RDBMS 表完成高级的数据完整性维护和信息处理的功能。也就是说, Geodatabase 对象作为具有唯一标识的表中的记录进行存储, 其行为通过 Geodatabase 应用逻辑实现。

Geodatabase 相对于其所选用的 DBMS 而言, 扮演了“中间件”的角色, 通过中间件的作用, 将不同的操作系统平台和数据库平台的差异屏蔽在中间件之中, 将面向空间数据管理和应用所需的技术高度专业地呈现出来, 供不同的客户端高效地共享和互操作。

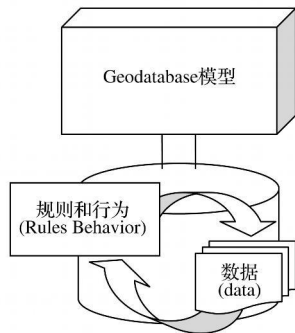


图 1 Geodatabase 模型的体系结构

Fig 1 The system structure of Geodatabase model

空间数据一般存储为矢量要素、栅格数据和属性表。一个 DBMS 表可以用来存放一个要素的集合, 表中的每行可以用来存储一个要素。每行中的 shape 字段存储要素的空间几何或形状信息, shape 字段的类型一般有两种: 一种是普通 DBMS 的 BLOB 类型; 另一种是某些 DBMS 所支持的空间扩展类型。栅格和图像数据也存放在关系表中, 栅格数据通常很大, 需要副表用于存储, 并且通常需要将其切成小片, 称为块 (block), 存放在单独的块表记录中。在关系数据库中存储的不只是空间数据和属性数据, 还包括地理数据的模式和规则, 如要素类的属性、拓扑、网络、关系、域等, 该部分内容通过 Geodatabase 元数据表来维护^[3]。

3 Geodatabase 数据模型的优点

Geodatabase 数据模型在实现上融合了面向对象技术和关系数据库技术, 大大扩展了带有行为关系和属性的表达能力, 允许在数据模型中使用规则定义更多的 GIS 应用功能。它也提供了大型数据库系统在数据库管理方面的优势, 这就意味着地理数据可以利用信息技术的发展集中管理。与以往的数据模型相比具有无可比拟的优势^[2, 4, 6]:

1 实现了地理数据的统一存储从而形成了严格意义上的地理空间数据库

Geodatabase 在逻辑上统一了以往的空间数据模型

(包括 CAD 影像、矢量数据、栅格数据、TN 地址数据等) 实现了在同一数据库系统中无须进行格式转换而统一管理存储和处理各种模型的空间数据, 为上层应用提供了统一的数据接口; 它不仅实现了将不同格式的数据集成在一个数据库系统中, 也实现了空间数据和属性数据的集成存储, 改变了传统模型中通过 1 联系的状态。

2 允许多用户通过使用版本管理和长事务处理访问数据库

多个用户可以分别在本地编辑要素 workflow, 对出现的冲突进行处理, 使之达成一致。

3 实现了要素的连续无缝存储

Geodatabase 能容纳非常巨大的要素集而不需要进行数据分片或分块。

4 要素的几何特征能够更精确地被描述出来

Geodatabase 数据模型用直线、圆弧、椭圆弧和贝塞尔曲线等定义特征的外形。

5 用户可以更为直观地处理数据模型

Geodatabase 包含的数据对象可与用户数据的模型保持一致。用户操作的不再是通常的点、线和面, 而是具有实际意义的变压器、湖泊和道路等。

6) Geodatabase 中的要素有更丰富的相邻关系

通过拓扑关系、空间关系和一般关系, 用户不仅可以定义要素的性质, 而且可以定义它与其他要素的相邻关系。这种相邻关系能够让用户详细说明一个要素在相关要素被移动、改变或删除时所发生的行为, 而且能让用户定位或检查有关要素。

7 数据录入与编辑更加精确

通过智能化的校验行为, 用户可以避免许多数据录入和编辑的错误。

8 对象的属性和行为智能化结合

GIS 中的属性可以赋予自然行为, 属性间任何类型的关系都可以在 Geodatabase 中定义。由于对象的可继承性和可扩充性, 用户完全可以构建或拓展更符合需求的对象。

4 Geodatabase 数据模型存在的问题

Geodatabase 最基本的一个优点是打破了关系模型范式的限制, 直接支持对象的嵌套和变长记录。由于面向对象数据库技术尚不成熟, 而空间数据本身又相当复杂, 故 Geodatabase 仍有其局限性, 有待进一步完善^[4, 7]。

首先, Geodatabase 数据模型仅是一种逻辑模型, 它仅在代码级实现了面向对象。由于目前面向对象数据库技术尚不成熟, 只能将面向对象的空间实体存储于对象—关系型数据库中, 因此, 空间实体的存储需中间件 (空间数据引擎) 将其属性和规则分解后才能存储, 而空间实体的组合也是通过中间件来处理。而这种复杂繁琐的分解和组合操作将降低系统的工作效率, 同时, 阻碍了空间实体规则的继承和派生。

其次, Geodatabase 数据模型中不存在时间维, 不能表达时空数据, 与传统的数据模型一样, 描述的仅仅是静态

的空间数据。而描述地理现象和过程相关的时空变化信息在空间决策分析系统中起着不可忽视的作用,因此,支持时空数据的表达与处理是空间数据模型发展的必然趋势。

再次, Geodatabase数据模型不能充分描述空间实体复杂的组合情况,因而对其约束规则不能详尽表达。 Geodatabase数据模型可以定义两个要素类之间的组合关系,但对于组合与被组合对象间空间位置和属性信息的约束规则没有给予描述。如:某个省由若干地区组成,当省界发生变化时,各相应的地区也要发生相应的变化,且省人口总数是各地区人口总数的和。像这样的组合规则在现实世界里极为常见,而 Geodatabase模型却不能提供详尽的约束规则,给空间分析带来了一定的困难。

5 结束语

通过以上的总结可以看出, Geodatabase数据模型存在的问题归根结底是因为没有实现真正意义上的面向对象的数据模型。

所谓面向对象的数据模型是指人们可以按照对现实空间世界的认识对地理空间进行任意层次的分解,可以根据所描述的不同空间实体选择最基本的操作单元,从而最大限度地接近人们的思维习惯,对现实世界进行结构和行为的模拟。它应该完全摒弃原有的关系理论,具有完全面向对象的特征。在计算机存储的不应该是表中的某一条记录,而应是对象的整个属性和行为(时间状态

也作为对象的一个属性而存在)以及与其相关的约束规则,而这些信息应该是被封装起来使设计者和使用者分开。 GIS应用软件可以直接与数据库中的对象对话,而无需任何中间件。因此,实现面向对象的空间数据物理模型是未来发展的趋势。

参考文献:

[1] 陈述彭,鲁学军,周成虎.地理信息系统导论[M].北京:科学出版社,2000

[2] Michael Zeiler. Modeling Our World[M]. ESRI PRESS,1998

[3] Robert West. Understanding ArcSDE[M]. ESRI PRESS,1999

[4] 程昌秀,周成虎,陆锋. ArcInfo中面向对象空间数据模型的应用[J]. 地球信息科学,2002,4(1):86—90

[5] 何新东,陈曦.面向对象技术在GIS数据库中的应用研究[J]. 干旱区地理,2001,24(4):371—375

[6] 陈静,张树文.面向对象空间数据模型——Geodatabase及其实现[J]. 国土与自然资源研究,2003(2):20—23

[7] 宋杨,万幼川.一种新型空间数据模型 Geodatabase[J]. 测绘通报,2004(11):31—33

[8] 余应刚,包世泰.新型的GIS数据库——Geodatabase[J]. 广西师院学报(自然科学版),2000,17(4):65—69

[9] 李国标,庄雅平,王钰华.面向对象的GIS数据模型——地理数据库[J]. 测绘通报,2001(6):37—40

[责任编辑:王丽欣]

(上接第135页)

区域基础平面控制,还是方格网放样与测量,最后艾丁湖洼地最低点确定及联测,都体现了全球定位系统(GPS)测量技术的先进性和优越性。利用GPS RTK作业方式测量放样点,采用逐级布设方格网,逐渐缩小范围来趋近艾丁湖洼地最低点位置,既保证了艾丁湖洼地最低点位置的准确性和可靠性,又提高了工作效率,取代了传统的边角测量仪器和布网方式给测量成果带来的复杂过程 and 不便,是测量技术的一大进步,为浅水地形测量积累了经验。

(上接第138页)

4 结束语

由于GIS强大的数据管理和空间分析功能,因此,在城市环境可持续发展研究中应用GIS方便了环境信息的采集、管理、共享、分析利用等,为城市环境管理、监测、规划、决策和评价等提供了强有力的工具,建立基于GIS的环境信息系统集成管理多源环境目标是必要手段,在实际应用中将有助于城市环境的可持续发展。

参考文献:

[1] 杜培军,唐宏. GIS的发展及其在城市环境管理中的应用[J]. 江苏环境科技,1999,12(3):30—31

参考文献:

[1] 李玉冰. 多边形重心的计算方法[M]. 北京:科学出版社,2005

[2] 肖学年,岳建利,张鹏. 全球定位系统(GPS)测量规范[S]. 北京:中国标准出版社,2009

[3] 中国有色金属工业协会. 工程测量规范(GB50026—2007)[S]. 北京:中国计划出版社,2008

[责任编辑:李颖]

[2] 汤国安,赵牡丹. 地理信息系统[M]. 北京:科学出版社,2001

[3] 余洁,边馥苓,胡炳清. 基于GIS和3D方法的社会经济发展与生态环境响应动态模拟预测研究[J]. 武汉大学学报:信息科学版,2003,28(1):19—24

[4] 崔宏伟. 区域可持续发展决策支持系统研究[M]. 北京:宇航出版社,1995

[5] 杜培军,高并祥. “3S”技术的城市环境监测与管理系统研究[J]. 环境监测管理与技术,2000,12(2):20—21

[6] 宋振宇,刘永清. 基于GIS的城镇可持续发展决策支持系统[J]. 系统工程理论与实践,1997(11):36—40

[7] 龚强. 城市综合功能GIS框架结构研究[J]. 测绘软科学研究,2002,8(1):18—22

[责任编辑:李颖]