

统一对象模型——对 GIS 融入 IT 的认识

梁 军 何建邦

(中国科学院地理科学与资源研究所 北京100101)

Unified Object Model——Understanding of GIS Fusion into IT

LIANG Jun HE Jian-Bang

(Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101)

Abstract The process of manufacturing tools by human being has experienced three stages: material tools, energy tools and information tools. Information tools compose information system in the form of systems. Information systems are mapping of entities and relationships between them in real world. It composes of objects and relationships, objects can be classified into Facility Objects and Domain Objects, Domain Objects are inheritance and extension of Software Facility Objects, Software Facility Objects are generalization and abstraction of Domain Objects. Information systems are digital model of the real world; the core problems of IT development are modeling the real world. The development of GIS extends the current object model, to construct a Unified Object Model with spatial property. It makes IT modeling the real world more effectually, information is organized and analyzed in the spatial framework, it will improve the unification of spatial and temporal recognition in IT and promote the development of knowledge tools.

Current GIS technology is suitable for resolving the macro spatial problems, Micro GIS must be developed to solve the micro spatial problem.

Keywords Unified object model, Information tools, GIS, Micro GIS

现实世界是由物质、能量和信息构成的。作为能够制造和使用工具的动物,人类制造工具的过程经历了物质工具、能量工具到信息工具三个阶段,信息工具(Information Tools)的广泛使用、发展与完善,是 IT 技术发展的结果,IT 技术的发展使人类社会进入了信息时代。

信息社会主要的特征是软件或信息系统(即信息工具)的广泛使用,信息工具在信息社会中无所不在。信息工具处理的对象是数据,数据是现实世界或系统运动状态的反映,经信息工具处理后,便成为对人类有用的信息。人类制造和使用信息工具的目的,在于认识与了解现实世界(包括人类自身),利用信息通过物质和能量的工具对现实世界或系统进行调控,以达到人类社会和经济可持续发展的目标。信息工具实质是现实世界和系统的模型,对现实世界的建模^[1]是 IT 发展的核心。

信息工具作为现实世界和系统的模型,不可能不反映现实世界和系统的空间现象与过程,“现代物理学的发展,特别是相对论的提出和得到的证实,更证明空间和时间同运动着的物质有着不可分割的联系,既没有脱离物质运动的空间和时间,也没有不在空间和时间中运动的物质。……各种各样的空间都是不同物质间的运动关系、结构形式,不存在独立于物质之外的空间体系。”^[2]GIS 作为一种空间信息处理技术,其发展和融入 IT,使 IT 能够更好地对现实世界建模,从而促进 IT 的发展。

1 信息工具的构成

信息工具主要以软件或信息系统的形式存在。按照面向对象的观点,软件或信息系统由相互联系和相互作用的对象构成^[3]。

$$S = \{O, R\}$$

其中, S 为软件或信息系统,由对象的集合与关系的集合构成。O 为对象的集合, $O = \{o_1, o_2, \dots, o_n\}$, o 表示对象。R 为对象之间的相互关系的集合, $R = \{r_1, r_2, \dots, r_n\}$, r 表示对象之间的关系。

信息工具以系统的形式存在,构成了信息系统,其组成如图 1 所示。构成信息系统的对象可以分为设施对象(Facility Object, FO)和领域对象(Domain Object, DO)两大类。

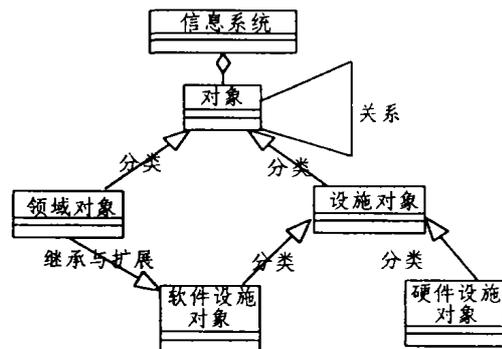


图1 信息系统的组成

1) 设施对象(FO)

设施对象是信息系统的基础设施,可以分为硬件设施对象(Hardware Facility Object, HFO)和软件设施对象(Software Facility Object, SFO)。硬件设施对象(HFO)包括计算机、网络和相关设备及其它硬件设备;软件设施对象(SFO)包括操作系统、数据库管理系统、应用开发工具及软件库等。它们为信息系统的构建提供了基础的软、硬件平台。IT 发展是设施对象不断发展与完善的过程,它体现为设施对象模型的

梁 军 博士,主要从事城市 GIS 研究。何建邦 教授,博士生导师。

完善、功能的增强、性能和可靠性的提高。

2) 领域对象

领域对象(DO)是现实世界特定领域(Domain)的实体(或对象)在信息空间的影射,是对这些实体及其相互关系进行建模和数字编码的结果。领域对象是对软件设施对象(SFO)的继承与扩展,是信息系统处理的核心对象。由领域对象及其相互关系构成的领域对象模型(Domain Object Model)能否有效地反映现实世界特定领域实体及其关系,是信息系统开发能否成功的关键,它一方面决定于作为基础的软件设施对象模型是否完善;另一方面决定于系统分析与设计人员对现实世界的特定领域组成对象及其相互关系的分析和利用所提供的软件设施对象进行建模和数字编码的能力。

领域对象是对软件设施对象的继承与扩展,而软件设施对象则是领域对象的抽象和概括,对领域对象研究的深入,将促进软件设施对象的模型的发展与完善,从而促进 IT 的发展。

模型工程(Model Engineering)^[4]和模型驱动的体系结构(Model Driven Architecture)^[5]的出现与发展反映了人们对建模技术在信息工具构建中的重要性的认识。人们开始从本体论(ontology)的角度认识信息工具^[6]。

2 统一对象模型

从本体论的角度,现实世界是由实体与关系构成,而信息世界是由对象与关系构成。信息世界中的对象与关系是现实世界中实体与关系的影射,信息工具是现实世界或系统的数字模型。面向对象技术(Object Oriented Technology, OOT)的发展反映了 IT 在认识论和方法论上的突破,同时为模型工程与 MDA 的发展奠定了基础,对象可以表示为图的顶点,关系表示为边,可以用图对现实世界与信息世界进行可视化建模。

传统的实体关系模型(ER 模型)就是以图的形式反映数据实体及其相互关系,建立信息系统的数据模型。

ER 模型可以表示为:

$$dm = \{E, R\}$$

其中, dm 为数据模型,由实体集合和关系集合构成。E 为实体的集合, $E = \{e_1, e_2, \dots, e_n\}$, e 表示实体。R 为实体关系的集合, $e = \{a_1, a_2, \dots, a_n\}$, a 表示属性。

$a = \{name, value, Domain\}$, 属性由属性名、属性值和属性值的集合,属性域构成, $value(a) \in Domain$ 。

ER 模型反映了实体、实体的属性及实体之间的相互关系,没有反映实体(或对象)的行为,主要用于表示数据对象(Data Object)及其相互关系。

OOT 则将属性与行为封装为一体,构成了一个完整的对象。

对象模型(Object Model)是对实体-关系模型的扩展,除了反映实体的属性外,还反映了实体(或对象)的行为。对象模型可以用 UML 的类图(Class Diagram)表示^[7,8]。

对象表示为: $o = \{A, B\}$ 。其中 A 为对象(或实体)属性的集合, $A = \{a_1, a_2, \dots, a_n\}$, a 表示属性。 $B = \{b_1, b_2, \dots, b_n\}$, b 表示行为,对应于方法(method)或操作(operation)。

GIS 技术的发展进一步完善与扩展了对象模型,从而建立了一个能真实反映现实世界的统一对象模型(Unified Object Model)。统一对象模型可以表示如下:

$$o = \{\{A, G, O\}, B\}$$

其中, A 为对象的非空间属性(一般属性)的集合, G 为对象的空间属性(Geometry)的集合:

$$G = \{g_1, g_2, \dots, g_n\}, g \text{ 为对象的空间属性。}$$

为了简化,通常设 $G = \{g\}$, 即对象只有一个空间属性,对象的其它空间属性可用于对象的空间属性表示。

空间属性可以分为点(Point)、线(Line)、面(Polygon)、体(Body)等基本类型和基本类型组合构成的复合类型,可以是二维或三维形式。

空间属性基于一定坐标系,由离散的点(x,y)或(x,y,z)及其序列表示。坐标系可以是地理坐标或其它相对的坐标系,反映对象在地理空间或其它相对空间的属性。

O 为构成对象的子对象的集合。

$$O = \{o_1, o_2, \dots, o_n\}, o \text{ 为构成对象的子对象。}$$

对象具有递归的特点,即对象包含子对象(sub-objects),子对象也可以包含其子对象。

B 为对象的集合,可以分为空间行为和非空间行为,对象的属性是通过其行为(b)来改变的, B 反映了对象的运动与变化特征。

统一对象模型可以用于描述现实世界或信息世界中的所有对象,包括空间对象和非空间对象、软件对象与数据对象和系统与组件(Component)。

基于统一对象模型可以对对象进行分类,见表1。

表1 统一对象模型对象分类

属性			行为	对象分类
A	G	O	B	
1	1	1	1	复合空间对象
1	0	0	1	空间对象(SO)
1	0	1	1	复合对象(CO)
1	0	0	1	一般对象(GO)
1	1	1	0	复合空间数据对象(CSDO)
1	1	0	0	空间数据对象(SDO)
1	0	1	0	复合数据对象(CDO)
1	0	0	0	一般数据对象(GDO)

对象间的关系可以表示为 $r = \{Ra, Rg, Ro, Rb\}$ 。其中, Ra 为对象与对象之间的属性关系集合; Rg 为对象与对象之间的空间属性关系的集合,即空间关系集合; Ro 为对象及其组成对象的关系集合; Rb 为对象与对象之间的行为关系集合。

对象间的关系如表2所示。

表2 统一模型中的对象关系

属性			行为	对象关系
A	G	O	B	
1	1	1	1	继承与泛化关系
1	0	0	0	关联关系
0	1	0	0	空间关系
0	0	1	0	聚集关系
0	0	0	1	行为关系

现实世界中实体的空间属性和空间关系是普遍存在的, 80%的信息与空间相关。GIS 主要处理和反映信息世界对象的空间属性、空间关系和空间行为。它扩展了对象模型,从而使 IT 能更好地对现实世界建模。

3 GIS 融入 IT 的意义

GIS 发展依赖于 IT 的发展,如组件式 GIS (Mapx、

MapObjects、ArcObjects 和 SuperMap)、Web GIS (Map Extreme, Arc IMS, SuperMap IS)、采用关系数据库和对象关系数据库管理空间数据 (Oracle Spatial 和 ESRI 的 Geo-Database^[1]) 等,它是组件技术、Internet 技术和数据库技术发展的结果。而 GIS 技术的发展又促进了 IT 技术的发展, GIS 对 IT 技术最主要的贡献在于解决了 IT 对现实世界建模过程中涉及的空间问题,促进了统一对象模型的发展,从而促进了信息系统的空间化 (Spatially Enable),空间化只是在信息系统原有对象或实体中增加空间属性,并通过 GIS 软件加以处理的过程,“空间对象 (Feature) 在关系表中是一个具有几何属性 (Geometry) 的行 (Row),空间对象类 (Feature Class) 是具有几何属性列 (Geometry Column) 的表”^[1]。GIS 使 IT 所处理信息能够纳入到一个统一空间框架体系中,发现其存在的空间关系和模式,这种空间框架不局限于现实地理空间,也可以是一种虚拟的空间,从而促进了基于空间关系的知识挖掘技术的发展。

现有采用关系数据库或对象关系数据库管理空间数据的 GIS 技术适合于解决宏观的空间问题,难于解决微观的空间问题,即反映微观空间对象 (Micro Spatial Objects)。在现实世界中,微观空间对象往往是一个复杂空间对象,如一个房屋的内部结构,并且往往基于独立、局部的坐标系。采用现有的方法,一个微观空间对象用一个表来表示,将大大增加数据库中表的数量,在实现上是不可行的;用一个表来存贮多个微观空间对象,由于一个表对应一个统一的坐标系,难于解决微观对象独立的、局部坐标系问题,而且实现效率较低。因此,需要发展一种微观 GIS 技术,采用表中的一个列 (Column) 来存贮微观空间对象,并记录微观空间对象的独立坐标系及其与统

一坐标系之间的关系,可以微观 GIS 视图 (Micro GIS View) 的方式对一个微观 GIS 对象进行操作,也可以在宏观 GIS 视图中通过坐标转换进行操作,微观 GIS 将成为现有 GIS 的一个扩展。

随着时态数据库 (TDB)^[9] 和时态 GIS (Temporal GIS) 技术的发展与成熟,基于 IT 技术,人们将进一步从时空的角度认识现实世界,最终实现人类对时空认知的统一。GIS 与 IT 技术的发展将使人类从制造和使用信息工具阶段,进入到制造和使用知识工具的时代。

参考文献

- 1 Zeler M. Modeling Our World-The ESRI Guide to Geodatabase Design. Environment Systems Research Institute, Inc. 1999
- 2 白光润. 地理学的哲学贫困. 地理学报, 1995, 50(3)
- 3 Amler S W. The Object Prime. Second Edition. Cambridge University Press, 2001
- 4 Bézin J. Who's Afraid of Ontologies?. OOPSLA'98 Workshop #25, Model Engineering, Methods and Tools Integration with CDIF
- 5 Architecture Board MDA Drafting Team. Model Driven Architecture--A Technical Perspective. Review Draft 14 February 2001, Document Number ab/2001-02-01
- 6 Nicola Guarino Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press
- 7 Object Management Group (OMG). OMG Unified Modeling Language Specification Version 1.3, June 1999
- 8 Rumbaugh J, Jacobson I, Booch G 著. UML 参考手册. 姚淑珍, 唐发银等译. 机械工业出版社, 2001. 1
- 9 何新贵, 唐常杰, 李霖, 刘云生著. 特种数据库技术. 科学出版社, 2000

(上接第39页)

SRB V1. 2^[8], 并在全国14个站点上安装了 SRB 软件。

4. 应用

涉及核模拟、气象、地球科学、高能物理学、医学、生物学等领域的科学计算是网格信息计算技术的典型应用。其中,生物学的发展关系到一个国家国民经济的发展,甚至整个人类社会的进步。

生物学的发展与精确的数据表示和高效的数据处理息息相关。目前,生物数据的表示和查询技术的发展已经能够适应大规模的软件分布。为此,美国的 NPACI 计划试图利用信息计算技术为生物学的研究提供支持,推动生物学研究工作的广泛开展。SDSC、斯坦福大学、德克萨斯大学联合推出了“生物数据表示与查询” (Biological Data Representation and Query, 简称 BDRQ) 计划。BDRQ 主要研究分子生物序列结构数据库的高效查询方法。为了支持对序列结构数据库的分析,需要从核酸数据库和蛋白质数据库中提取三维结构信息。因此,需要联合数据库进行统一查询,并进一步研究新的查询算法,以允许生物学家通过类比已知的分子来从序列和结构信息中推断出功能。

斯坦福大学已经开发了一种自动注释蛋白质三维结构中功能部位和结构部位的方法,并将通过 NPACI 向外提供 WEB 服务。德克萨斯大学正在研究解决物种分析中的复杂问题的工具,编程实现时将充分利用 NPACI 的并行计算资源。目前,已经利用 SDSC 的 SRB 对 SDSC 和德克萨斯大学等站点的数据库进行管理,并向有关生物学的研究机构提供信息服务。除了 BDRQ 计划外, NPACI 还推出了一系列相关的研

究计划,如,“增强型生物成像” (Enhanced Biological Imaging) 计划等。

结论 从社会发展到国防建设,从地球科学、高能物理学、分子科学、神经系统科学、医学、生物学等学科到作战指挥、战场仿真,各领域明确提出了分布式异构环境下的海量信息高效复杂处理的应用需求,迫切需要加强网格的信息计算能力,因此,针对应用需求开展信息计算关键技术的研究迫在眉睫,这对各领域的发展将起到极大的推动作用。

参考文献

- 1 Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International J. Supercomputer Applications, 2001, 15(3)
- 2 Foster I, Kesselman C. The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann Publishers, 1999
- 3 Stockinger H, Rana O, Moore R, Merzky A. Data Management for Grid Environments. HPCN Europe 2001. 151~160
- 4 Moore R. Massive Data Analysis Systems: [Technical Report]. San Diego Supercomputer Center, 1997
- 5 Hoschek W, Jaen-Martinez J. Data Management in an International Data Grid Project, ACM International Workshop on Grid Computing (Grid'2000), 17-20, Dec. 2000, Bangalore, India
- 6 Chervenak A, Foster I. The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets, Journal of network and computer applications, 2000
- 7 Foster I, Kesselman C. The Globus Project: A Status Report. Proc. IPPS/SPDP '98 Heterogeneous Computing Workshop, 1998. 4~18
- 8 Rajasekar A K, Wan M. SRB & SRBRack - Components of a Virtual Data Grid Architecture, Advanced Simulation Technologies Conference (ASTC02) San Diego, 2002. 15~17