



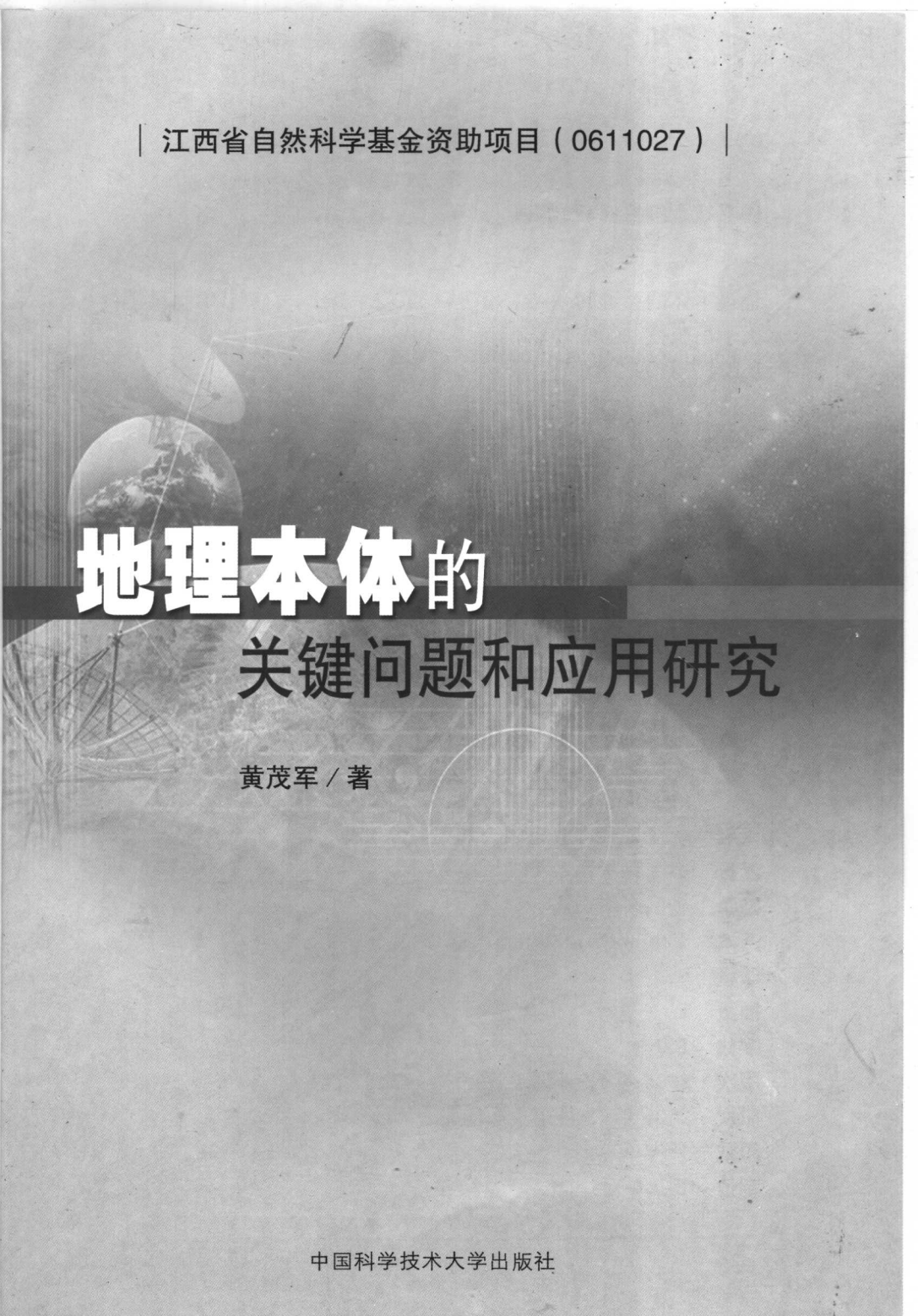
# 地理本体的 关键问题和应用研究

Study on Key Issues and Application of Geographic Ontology

黄茂军 著

中国科学技术大学出版社

| 江西省自然科学基金资助项目 ( 0611027 ) |



# 地理本体的 关键问题和应用研究

黄茂军 / 著

中国科学技术大学出版社

## 图书在版编目(CIP)数据

地理本体的关键问题和应用研究/黄茂军著. —合肥: 中国科学技术大学出版社, 2006. 12

ISBN 7-312-02026-7

I. 地… II. 黄… III. 地理学—研究 IV. K90

中国版本图书馆 CIP 数据核字(2006)第 152332 号

出版 中国科学技术大学出版社  
安徽省合肥市金寨路 96 号, 邮编: 230026  
网址: <http://press.ustc.edu.cn>  
印刷 合肥现代印务有限公司  
发行 中国科学技术大学出版社  
经销 全国新华书店  
开本 880 mm×1230 mm 1/32  
印张 5  
插页 1  
字数 150 千  
版次 2006 年 12 月第 1 版  
印次 2006 年 12 月第 1 次印刷  
印数 1~3 000 册  
定价 15.00 元

## 前 言

本体论原本是一个哲学上的概念,然而近十多年来,本体论研究在计算机科学中日益受到重视,其重要性在知识工程、数据库设计、信息建模、信息检索与提取、知识管理与组织等不同领域得到体现。

作为计算系统的地理信息系统也认识到了本体论的重要性,地理本体的研究在 20 世纪 90 年代中后期逐渐得到重视。目前,国外已经有不少研究机构和学者对地理本体极为关注,做了一些探索性的工作,而国内的地理信息学界,对这一问题还鲜有涉及。本体论观点在空间信息的概念模型、空间数据的共享和互操作、地理类别的研究等方面具有十分重要的意义,其最大的意义是对空间信息语义理论的丰富,潜在的优点在知识工程领域已初见端倪,本体论研究会成为我们加深对地理信息和地理信息系统认识的重要方式(杜清运, 2001)。

地理本体是一个非常复杂而又难以把握的概念,作为一个从哲学引入信息科学进而引入地理信息科学的概念,本体在其哲学源头上就是歧义最多、意义最为模糊的范畴之一。从现有研究来看,大多数 GIS 领域的学者在研究地理本体时,仅仅是把本体在信息科学中的含义直接移植到地理信息科学中去,没有考虑地理信息自身的特殊性。

本书对于地理本体的探讨就是从其概念展开的,尤其关注地理本体与一般信息本体的不同之处。本书把地理本体纳入到哲学、信息科学以及地理信息科学这些大的学科背景中去考察,重点研究了地理本体的形式化表达机制,它包括属性特征和空间特征两种不同的形式化表达机制。

全书共分 7 章,第 1 章为绪论,介绍了国内外地理本体研究的现

状,重点介绍了典型的研究情况.第2章为哲学本体、信息本体与空间本体,提出了地理本体的深刻内涵,指出地理本体涵盖了哲学本体、信息本体以及空间本体3个层面的含义,其中空间本体主要表现为与地理信息空间特征相关的本体,空间本体是地理本体所独有的,也是地理本体区别于其他一般信息本体的本质所在.第3章为地理本体的逻辑结构与建立方法,总结分析了现有的本体以及地理本体的逻辑结构,结合构建地理本体的实践,提出了地理本体宏观和微观两个不同层次的逻辑结构;在全面总结现有的本体构造方法的基础上归纳出建立地理本体的一般方法.第4章为OWL表达地理本体语义能力的剖析,总结了OWL的构成要素在语义表达上的优势,通过构造地理本体实例验证了这些特点;阐明了OWL具有表达地理本体属性特征的语义能力,但缺乏对地理本体空间特征的支持,指出OWL是表达地理本体的理想语言,但必须对其加以扩展.第5章为地理本体空间特征的形式化表达机制,指出部分一整体学、拓扑学和位置论这3个理论工具为地理本体空间特征和空间关系的表达提供了强大的理论基础,提出借助这3个理论工具,构造出形式化的空间特征以及空间关系公理,加入到OWL的建模原语当中,从而能够在OWL构建的地理本体之中表达其空间特征.第6章为基于本体的地图服务及其改进,分析了地理信息服务中的语义异质问题,提出了基于地理本体的解决思路,并以实例验证了这种方法的独到之处;指出网络地图服务面临不能提供有效查询服务的突出难题,提出综合利用本体和Web挖掘技术来改善网络地图的查询服务质量.第7章是总结和展望.

本书由黄茂军独立撰写,书中的实验部分得到了吴运超博士的大力相助.

本书是江西省自然科学基金项目“基于本体的地理信息服务若干关键问题研究”的最终研究成果之一.在写作和出版过程中,得到了江西省自然科学基金委员会和江西财经大学科研处的大力支持,

也得到杜清运教授、吴运超博士、Frederico Fonseca 博士、杜晓初博士、侯涛博士和解吉波博士等老师和朋友的帮助,在此表示衷心感谢.还要感谢江西财经大学用友软件学院的各位领导和老师,尤其是夏家莉院长、钟元生教授、尹爱华副院长、李华旻副院长、林勇教授和刘细发主任等老师,感谢他们的关心和帮助.最后,衷心感谢妻子李凤丹女士和亲人们多年来的鼓励和支持.

由于作者学术水平有限,疏漏与错误在所难免,诚恳希望读者批评指正.

黄茂军

2006 年 9 月

# 目 录

|                            |        |
|----------------------------|--------|
| 前言 .....                   | ( i )  |
| 第 1 章 绪论 .....             | ( 1 )  |
| 1.1 地理本体研究概况 .....         | ( 1 )  |
| 1.1.1 国内外研究概述 .....        | ( 2 )  |
| 1.1.2 典型研究介绍 .....         | ( 4 )  |
| 1.1.3 现有研究的分析与思考 .....     | ( 6 )  |
| 1.2 本书研究内容及意义 .....        | ( 9 )  |
| 第 2 章 哲学本体、信息本体与空间本体 ..... | ( 12 ) |
| 2.1 哲学本体 .....             | ( 13 ) |
| 2.1.1 本质论 .....            | ( 13 ) |
| 2.1.2 是论 .....             | ( 14 ) |
| 2.1.3 本体承诺论 .....          | ( 15 ) |
| 2.1.4 哲学本体演变的思考 .....      | ( 16 ) |
| 2.1.5 地理本体的哲学体现 .....      | ( 17 ) |
| 2.2 信息本体 .....             | ( 19 ) |
| 2.2.1 信息本体的一般概念 .....      | ( 20 ) |
| 2.2.2 信息本体的经典解释 .....      | ( 21 ) |
| 2.2.3 哲学本体与信息本体的比较 .....   | ( 23 ) |
| 2.2.4 信息本体的分类 .....        | ( 28 ) |
| 2.2.5 信息本体的应用 .....        | ( 33 ) |
| 2.2.6 信息本体和语义网 .....       | ( 35 ) |
| 2.2.7 地理本体的信息体现 .....      | ( 37 ) |
| 2.3 空间本体 .....             | ( 38 ) |
| 2.4 地理本体的内涵 .....          | ( 40 ) |
| 第 3 章 地理本体的逻辑结构与建立方法 ..... | ( 44 ) |
| 3.1 地理本体的逻辑结构 .....        | ( 44 ) |

|              |                                 |                |
|--------------|---------------------------------|----------------|
| 3.1.1        | 本体及地理本体逻辑结构分析 .....             | ( 44 )         |
| 3.1.2        | 地理本体的宏观和微观逻辑结构 .....            | ( 47 )         |
| 3.2          | 地理本体的构造方法 .....                 | ( 50 )         |
| 3.2.1        | 本体的构造准则 .....                   | ( 50 )         |
| 3.2.2        | 本体的开发方法 .....                   | ( 51 )         |
| 3.2.3        | 地理本体的建立方法 .....                 | ( 55 )         |
| <b>第 4 章</b> | <b>OWL 表达地理本体语义能力的剖析 .....</b>  | <b>( 58 )</b>  |
| 4.1          | OWL 概述 .....                    | ( 59 )         |
| 4.1.1        | 几种本体表达语言的比较 .....               | ( 59 )         |
| 4.1.2        | OWL 与描述逻辑 .....                 | ( 61 )         |
| 4.1.3        | OWL 的 3 个子语言 .....              | ( 62 )         |
| 4.2          | OWL 构成要素及其语义表达能力分析 .....        | ( 64 )         |
| 4.2.1        | OWL Lite 构成要素分析 .....           | ( 64 )         |
| 4.2.2        | OWL DL 和 OWL Full 增加的语言要素 ..... | ( 74 )         |
| 4.2.3        | OWL 表达地理本体能力总结 .....            | ( 77 )         |
| 4.3          | OWL 与 GML 比较 .....              | ( 79 )         |
| <b>第 5 章</b> | <b>地理本体空间特征的形式化表达机制 .....</b>   | <b>( 83 )</b>  |
| 5.1          | 问题的提出 .....                     | ( 84 )         |
| 5.2          | 空间关系的描述模型 .....                 | ( 86 )         |
| 5.2.1        | 空间关系及其分类 .....                  | ( 87 )         |
| 5.2.2        | 拓扑关系的描述及推理模型 .....              | ( 88 )         |
| 5.3          | 基于 3 个理论工具的地理本体形式化表达机制 .....    | ( 100 )        |
| 5.3.1        | 3 个理论工具 .....                   | ( 100 )        |
| 5.3.2        | 基于 3 个理论工具的定义、公理表达空间特征<br>..... | ( 102 )        |
| 5.3.3        | 应用实例 .....                      | ( 104 )        |
| <b>第 6 章</b> | <b>基于本体的地图服务及其改善 .....</b>      | <b>( 110 )</b> |
| 6.1          | 从地理信息系统到地理信息服务 .....            | ( 110 )        |
| 6.1.1        | Web 服务 .....                    | ( 110 )        |
| 6.1.2        | OGC Web 服务框架及规范 .....           | ( 112 )        |

---

|                                    |              |
|------------------------------------|--------------|
| 6.2 基于本体的地理信息服务 .....              | (114)        |
| 6.2.1 地理信息服务中的语义异质 .....           | (114)        |
| 6.2.2 基于本体的解决思路 .....              | (115)        |
| 6.2.3 实例应用——以地图服务为例 .....          | (116)        |
| 6.3 利用本体和 Web 挖掘技术改善网络地图查询服务 ..... | (120)        |
| 6.3.1 Web 挖掘 .....                 | (121)        |
| 6.3.2 改善查询服务的原理 .....              | (122)        |
| 6.3.3 实验分析 .....                   | (124)        |
| <b>第 7 章 总结与展望 .....</b>           | <b>(129)</b> |
| 7.1 总结 .....                       | (129)        |
| 7.2 主要创新 .....                     | (130)        |
| 7.3 研究展望 .....                     | (132)        |
| <b>参考文献 .....</b>                  | <b>(134)</b> |

# 第 1 章 绪 论

在地理信息科学领域中,把本体论作为一个突出的主题来研究还是近几年的事情.本体论在西方哲学史上占据统治地位长达两千多年,直到今天仍然是哲学家热烈争论的对象.时至今日,本体论不再只是哲学家关心的问题,它已经广泛应用到其他学科之中,像自然语言处理、生物科学、医学,尤其是信息科学领域,诸如人工智能、知识工程、地理信息科学等.将本体论引入地理信息科学领域主要反映了一种中心的转移,即由过去片面强调计算模型的形式化,到目前对空间目标域本身的关注(杜清运,2001).地理信息科学中的本体论研究是高度跨学科的交叉研究,与地理信息的认知、表达、互操作、尺度和不确定性密切相关,其最重要的一点是研究空间信息的语义理论,或者更一般地说,就是研究人类思维、信息系统与地理现实世界之间的关系(Mark 等,2000).地理信息科学中的本体论研究主要包含 3 个层次上的内容:① 研究建立完善详尽的地理本体所涉及的概念问题;② 研究地理本体形式化的方法和工具,使所建立的本体能够方便地被共享和重用,研究不同本体之间的转换方法和不同本体的集成方法,提供不同本体之间互操作的手段;③ 以所建立的地理本体为基础,开发出本体驱动的地理信息系统或进行基于本体的地理信息集成和互操作等应用研究.

## 1.1 地理本体研究概况

虽然学者们关注地理本体的时间还不算长,但有关地理本体的研究却已从多个方面展开.目前,国外已有不少研究机构和学者关注地理本体的研究,而国内在这方面的研究才刚刚起步.

### 1.1.1 国内外研究概述

国外比较有代表性的工作有:Smith 和 Mark(1998,2001,2003)对地理信息的认知类型和地理目标的本体特征的研究;Bittner(2001-2004)对地理尺度、不确定性以及部分一整体理论的本体关注;Frank(2001,2003)对时空数据库的基础本体的研究;Fonseca(1999-2002)则采用了 Guarino(1998a)的“本体论驱动的信息系统”的理念,重点探讨了本体论在互操作环境下信息系统的设计和实现阶段的应用;Kavouras 和 Kokla(2001,2002)对不同的地理本体进行了融合,在基于本体的地理信息集成方面做了一些工作。

本体论被引入地理信息科学始于 Egenhofer 和 Mark(1995)对于常识地理学(Naive Geography)的研究,常识地理学的研究能够为建立智能化和大众化的地理信息系统提供基础.在常识地理学研究当中,地理本体被认为是一个极其重要的研究领域,被列为常识地理学的 47 个研究问题之首.在 1996 年召开的美国国家地理信息与分析中心(NCGIA)专家会议上,首次提出了一个基于常识的地理本体框架。

美国大学地理信息系统联合会(UCGIS)显然也认识到了地理本体对于本学科的意义.在其 2000 年提出的 4 个新的研究领域中,有一个就是地理信息科学的本体论基础.在其 2002 年给出的研究日程表中,地理本体排在十大长期研究挑战之首。

受美国国家自然科学基金委员会(NSF)和其他一些组织机构的资助,美国有不少学者对地理本体进行了深入研究(Mark 等,2000;孙敏等,2004).1999 年 Egenhofer 主持的“地理空间信息的集成研究”项目研究重心是地理空间数据库的语义互操作,以此项目为依托,其学生 Rodriguez 于 2000 年完成了一篇题为“空间实体类的语义相似性评价”的博士论文;Fonseca 于 2001 年完成了一篇题为“本体论驱动的地理信息系统”的博士论文;1999 年 Findler 和 Malyankar 主持的题为“空间知识的表达与分发”的项目,研究建立沿海实体(如海岸线、潮汐面等)的本体;1999 年开始的 Mark 和 Smith 主持的项目“地理类别:本体论调查”,试图确立与认知关联的地理实体和地理类别

的实体,该项目研究的主题是一般性常识地理学,重点测试不同语言条件下人们对一些地理概念的认知,采用了主体实验的方法对地理认知种类的本体进行了系统的研究,通过设计一些典型的实验,得到了一些调查结果。

在欧洲,尤其是德国 Munster 大学的 Kuhn 教授(2001,2002,2003)、雅典工业大学的 Kavouras 教授(2000-2005)、维也纳工业大学的 Frank 教授(1997,2001,2003)、英国 Cardiff 大学的 Jones 教授(2002,2003,2004)都组织了专门的研究团队对地理本体进行了深入的研究。

从最近几年地理信息科学领域重要国际会议的主题来看,地理本体也受到了前所未有的关注。此外,还召开了几次专门的地理本体讨论会。2000 年,由 Schuurman 和 Mark 组织,在 Pittsburgh 召开了关于地理本体的美国地理学家协会(AAG)年会分会。2000 年 9 月,由 Winter 组织,在法国召开了名为“空间数据标准的本体论与认识论”的欧洲会议,会议就空间本体论、空间认知、空间不确定性、空间层次(Hierarchies)、空间一般化(Generalization)、标准与标准化等主题进行了研讨,该会议的优秀论文 2001 年在地理信息科学领域的顶尖刊物——国际地理信息科学杂志出版了本体特刊。2000 年、2002 年以及 2004 年,连续三次地理信息科学会议(GIScience)均对地理信息本体论进行了深入探讨。2001 年以及 2003 年的空间信息理论会议(COSIT)也有地理本体的专场讨论。此外,在近几年的欧洲地理信息实验室联合会(AGILE)年会上也有地理本体的主题讨论。从 1998 年开始的两年一届的信息系统中的形式化本体国际会议(FOIS)每次都有地理本体的相关主题。2002 年 9 月,在英国召开了地理本体(Geo-ontology)的专题讨论会,包括 Frank、Mark 和 Kuhn 等知名教授在内的学者对地理本体的关键问题以及研究的重点范围进行了深入讨论,对地理本体的未来研究方向提出了一些有益的建议。上述研究与学术研讨活动表明,地理本体的研究近年来在国外受到了广泛的关注,已有众多学者和学术团体投身于地理本体的研究之中。

国内,武汉大学的崔巍博士 2004 年完成了地理本体方面的第一篇博士论文,其研究主题是基于本体的地理信息系统语义集成和互

操作. 此外, 中科院地理所何建邦教授(2003)在进行资源环境信息分类编码研究时, 考虑到了与地理本体进行关联. 北京大学遥感所的孙敏博士等(2004)对地理本体的研究进行了综述.

### 1.1.2 典型研究介绍

下面重点介绍 Fonseca, Kavouras 领导的地理本体研究小组 OntoGeo 以及国内崔巍博士的研究工作.

Fonseca 在基于本体的地理信息集成方面做了大量的研究工作 (Fonseca, 1999-2005). 他认为集成地理信息首先需要形式化定义人们对于现实世界的概念模型, 而且为了能够在不同的学术共同体内部获得基本的一致, 这些概念模型需要依据学术共同体的不同而分类. 对每个学术共同体的概念模型的形式化定义就是本体.

他提出了一个本体驱动的地理信息系统 (Ontology-Driven Geographic Information Systems, ODGIS) 框架 (如图 1-1 所示). 在 GIS 中使用本体, 并把本体转换为信息系统中能起作用的一个组成部分就产生了 ODGIS. ODGIS 是利用衍生自各种本体的软件组件建立起来的, 而这些组件可以用来开发新的应用的各种类. 因为本体驱动

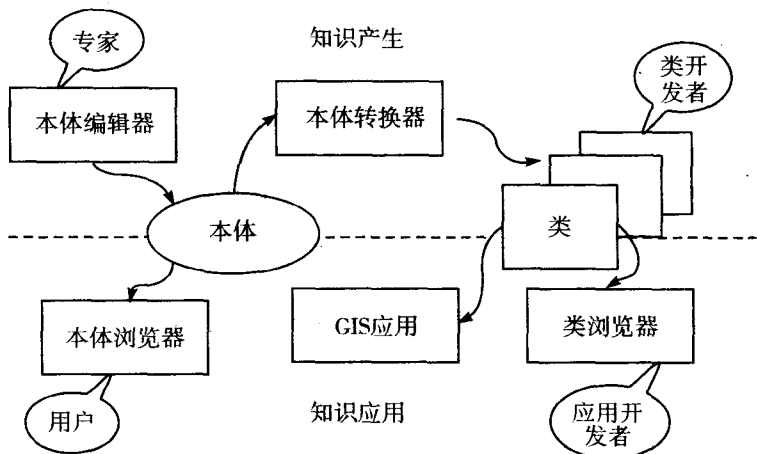


图 1-1 本体驱动的地理信息系统的框架 (Fonseca, 2001, 2002a)

的,这些类嵌入了从本体中提取的知识。

ODGIS 结构包含两个主要部分:知识产生和知识使用。知识产生阶段包括用本体编辑器定义本体,利用已有本体定义新的本体以及把本体转化为软件组件。知识使用阶段依赖于知识产生阶段的产品,也就是用形式化语言定义的一组本体和一组类。本体能够被终端用户浏览,它们提供了可获得数据的元数据信息。类包含了数据和操作,构成了系统的主要功能,并且这些类通过中间件和信息源联系。通过集成与信息源紧密联系的本体,使得地理信息的集成是基于其语义信息的。同时,因为信息的集成可能跨越不同的尺度,他还提出了不同尺度下的地理信息集成的机制。

雅典工业大学的 Kavouras 教授领导的地理本体研究小组 OntoGeo制订了一个宏伟的研究计划(OntoGeo 网站),他们认为本体问题对地理信息科学领域中的很多研究方向具有根本性的影响。目前的研究涵盖了地理本体的理论方面和实践方面。理论方面的研究主要包括地理学中的概念理论、地理类别、概念映射、本体的模糊性、时空建模、相似性和粒度;实践方面的研究主要包括形式化本体、本体工程、本体集成、语义互操作、形式化概念分析和概念格(Concept Lattices)。

该研究分为 4 个大的方向,涉及地理本体的理论、方法与应用的各个层面:① 本体的建立(Building Ontologies);② 本体的可视化(Visualizing Ontologies);③ 本体的实现和应用(Implementing and Using Ontologies);④ 本体的集成(Integrating Different Ontologies)。

由于其研究计划非常庞大,必须分步实施,大部分研究尚在进行之中,正在进行的研究主要依托博士生和硕士生完成,包括:① 集成时空框架的开发(Development of an Integrated Spatio-temporal Framework, Panopoulos);② 本体集成(Ontology Integration, Kokla);③ 本体模糊性(Ontological Vagueness, Tomai);④ 语义转换(Semantic Translators, Karalopoulos);⑤ 社会经济单元的本体论(the Ontology of Socioeconomic Units, Darra);⑥ 概念影射和本体空间化(Concept Mapping and Ontology Spatialization, Kontaxaki);⑦ 本体编辑器的开发(Development of an Ontology Editor, Per-

dikis); ⑧ 时空归纳 (Spatio-temporal Generalization, Stamatopoulos).

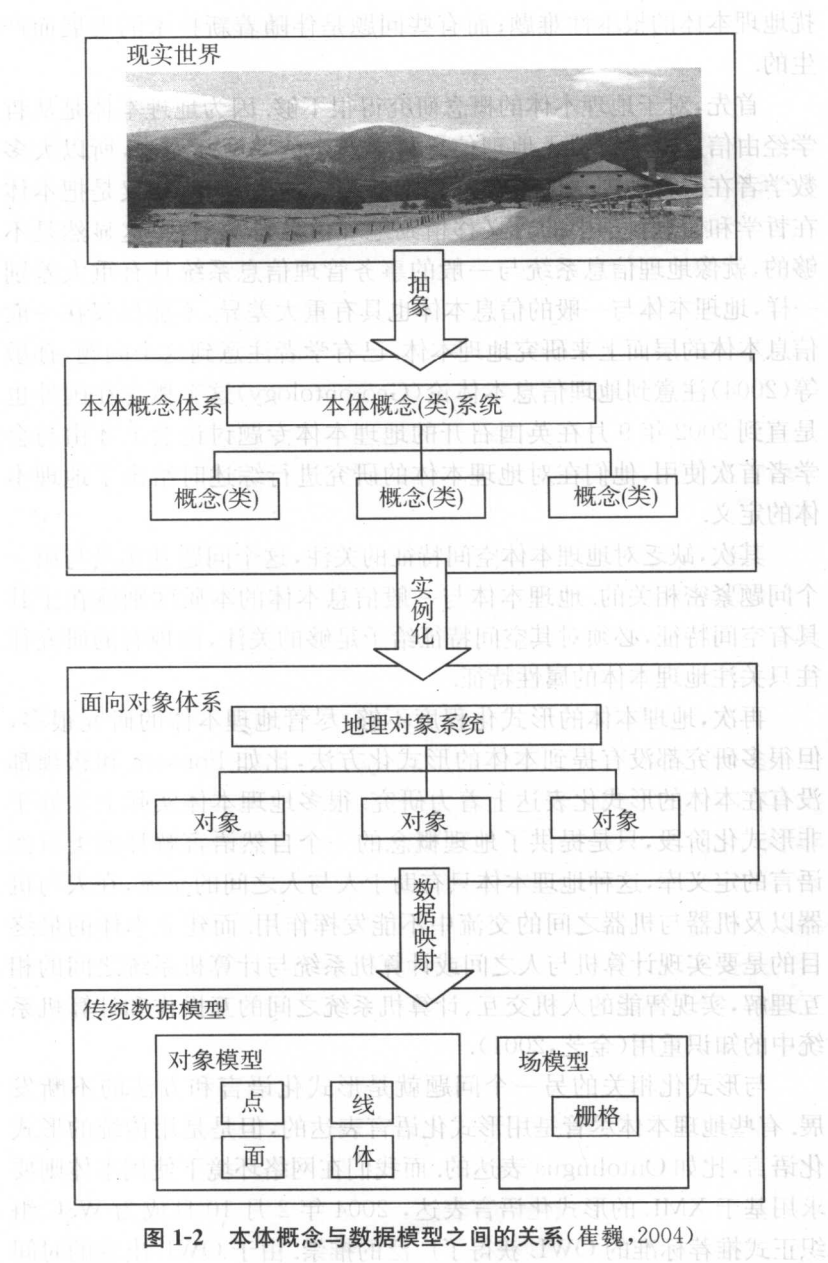
从已发表的文章来看,主要在基于本体的地理信息集成方面做了一些工作,用概念格的方法对不同的地理本体进行了融合研究(Kavouras 和 Kokla, 2002, 2003). 在地理本体的建立方面提出了一个产生健壮(Robust)地理本体的框架,能够保持具体领域概念的语义信息(Tomai 和 Kavouras, 2004).

国内,武汉大学的崔巍博士对基于本体的地理信息系统语义集成和互操作进行了深入研究,提出并建立了基于本体的空间信息语义网格的较为完整的理论体系框架和实现技术(崔巍, 2004b). 他把本体、网络技术和轻量目录访问协议有机地融合在一起,构造了一个基于本体的空间信息语义网格系统结构,为地理信息社会化以及其他领域信息系统的融合做出了有益的探索.

在其论文中,研究了本体概念与类、对象以及传统地理信息系统数据模型之间的关系(如图 1-2 所示),说明了基于本体的地理信息系统与传统的地理信息系统之间的联系与区别. 他认为本体概念和面向对象的类分别是现实地理世界在人类思维和计算机软件系统中的表达,是同一客观世界在人类思维和计算机世界中的映射. 面向对象技术的类和对象可以用于显式地和形式化地描述本体概念的语义和概念之间的层次关系. 在基于本体的空间信息语义网格中,本体系统为面向对象技术的类、对象、关系、事件和过程等概念术语建立了标准的语义解释. 本体一旦建立,很容易转换为面向对象的类库,本体也可以借助类的继承和实例化机制延续生命周期,使本体从过去作为系统的静态分析工具扩展为动态运行系统的一个主要组成部分.

### 1.1.3 现有研究的分析与思考

上述研究表明,地理本体的研究近年来受到了广泛的关注,但在这种现象的背后还是能够发现一些问题. 地理本体出现的时间还不长,有些问题现在被忽略了,还没有引起足够的重视;有些问题是因



扰地理本体的根本性难题;而有些问题是伴随着新技术的发展而产生的。

首先,对于地理本体的概念研究得很不够。因为地理本体是从哲学经由信息科学被引入地理信息科学产生的一个新概念,所以大多数学者在研究地理本体或与地理本体相关的内容时,仅仅是把本体在哲学和信息科学中的含义移植到地理信息科学中去。这显然是不够的,就像地理信息系统与一般的事务管理信息系统具有重大差别一样,地理本体与一般的信息本体也具有重大差异,不能仅仅在一般信息本体的层面上来研究地理本体。已有学者注意到这个问题,孙敏等(2004)注意到地理信息本体论(Geo-ontology)这个概念在国外也是直到2002年9月在英国召开的地理本体专题讨论会上才由与会学者首次使用,他们在对地理本体的研究进行综述时给出了地理本体的定义。

其次,缺乏对地理本体空间特征的关注,这个问题其实是与第一个问题紧密相关的。地理本体与一般信息本体的本质区别就在于其具有空间特征,必须对其空间特征给予足够的关注,而现有的研究往往只关注地理本体的属性特征。

再次,地理本体的形式化程度不够。尽管地理本体的研究很多,但很多研究都没有提到本体的形式化方法,比如 Fonseca 和崔巍都没有在本体的形式化表达上着力研究。很多地理本体实际上是处于非形式化阶段,只是提供了地理概念的一个自然语言解释或类自然语言的定义库,这种地理本体只有助于人与人之间的交流,在人与机器以及机器与机器之间的交流中不能发挥作用。而建立本体的最终目的是要实现计算机与人之间或计算机系统与计算机系统之间的相互理解,实现智能的人机交互、计算机系统之间的互操作和计算机系统中的知识重用(金芝,2001)。

与形式化相关的另一个问题就是形式化语言和方法的不断发展。有些地理本体尽管是用形式化语言表达的,但是是用传统的形式化语言,比如 Ontolingua 表达的。而我们在网络环境下使用本体则要求用基于 XML 的形式化语言表达,2004年2月10日成为 W3C 组织正式推荐标准的 OWL 获得了广泛的推崇。由于 OWL 出现的时间