

# 一种粒化理论的地理本体融合方法

李军利<sup>1,2</sup> 何宗宜<sup>1</sup> 朱乔利<sup>1</sup> 刘亚虹<sup>3</sup>

(1 武汉大学资源与环境科学学院,武汉市珞喻路 129 号,430079)

(2 滁州学院地理信息与旅游学院,滁州市丰乐大道 1528 号,239000)

(3 武汉市测绘研究院,武汉市万松园路 209 号,430022)

**摘要:**提出了一种粒化理论的地理本体融合方法。该方法首先运用形式概念分析地理本体,再引入粒计算理论,在不同粒度下约简概念格完成本体融合,认为地理融合过程是一个渐进过程。利用地理信息基础数据和水利空间数据进行融合实验,验证了此方法的有效性。

**关键词:**地理本体融合;形式概念分析;粒计算;空间数据;地理信息

**中图分类号:**P208

目前,多源空间数据融合中基于结构与语法的异构研究已有较多成果<sup>[1]</sup>,而基于语义的融合研究相对较少。地理本体被视为解决多源空间信息共享与互操作的有力工具,因分类标准和规范的不同,其复杂程度也不同,从不同视角构建的本体相互兼容性差,地理本体共享困难。

当前地理语义的数据融合研究已有一些阶段性成果<sup>[2-6]</sup>,但其主要侧重于顶层本体与同一领域内的异构数据集成,文献<sup>[7]</sup>虽为跨领域融合,但融合结果随着规模增大,概念空间异常复杂。本文针对地理本体融合的复杂性,应用形式概念分析与粒计算理论对地理本体进行融合实验,并对不同粒度下的融合结果进行了约简。

## 1 形式概念语义分析

事物的特征在哲学上描述就是概念,概念及其之间的关系形成形式概念(概念格)。形式概念依据对象与属性之间的二元关系建立所有概念之间的层次关系<sup>[8]</sup>。概念格的每一个节点被称为一个形式概念,包括内涵和外延两个部分。外延指属于该形式概念的所有对象(实体)集合。内涵指所有这些对象所共同具有的特征(属性)集合。

形式概念中外延与内涵同等重要,但地理本体的语义基础通常是指内涵<sup>[2]</sup>。地理概念的外延与

内涵一般要借助领域专业辞典与标准抽取。

地理概念语义可从持久特性、非持久特性、反持久特性、半持久特性,载体标识性与外部依赖性元概念特性方面抽取<sup>[9]</sup>。参照文献<sup>[5,9-10]</sup>,确定两个原则:自然观察实体、客观本体的地理要素时间、空间属性优于其他属性;社会本体、主观本体的要素人工属性优于其他属性。本文选取两个领域的部分要素,构造共同的形式概念表<sup>[8]</sup>(见表 1)。

## 2 地理本体粒化理论

粒计算是研究人类启示问题求解的理论及方法,主要侧重于不确定性处理、多粒度计算<sup>[11]</sup>。跨领域地理本体融合,必然产生一些内涵相似、功能相近、依存关系复杂的地理概念,特别是在形式概念分析中,随着概念内涵与外延的不断增大,构造概念格算法受时空复杂性影响,性能会逐渐降低。因此,将粒计算引入地理本体融合,通过选择合适的粒度,立足于某一个感兴趣的点、层或视角进行地理领域本体详尽分析,可寻找到一种较好的近似解决方案。在粗粒度下,忽略一些描述细节,从而降低问题求解的难度,还可以在不同粒度之间进行切换,以解决在原有层次上粒度难以表示与刻划的问题。参照文献<sup>[12-13]</sup>,给出地理本体粒化理论的几个定义。

收稿日期:2013-03-15。

项目来源:国家自然科学基金资助项目(41071290,41201463);安徽省高校省级科学研究资助项目(2010sk471)。

定义1 地理领域本体可以形式定义为序对  $K=(U,A)$ ,  $U$  是一非空有限集,称为地理领域本体的论域,  $A$  是一个二元等价关系。  $A \subseteq U \times U$ , 表示  $U$  的一个不可分辨关系。对于  $\forall X \subseteq A$ , 由  $X$  产生不可分辨关系为  $IND(X) = \bigcap_{A \in X} A$ , 表示领域本体的最高分辨与表示程度。

根据粗糙集理论,拥有领域知识的智能体不能分辨  $U$  上的每一个划分对象  $u$  与其所在的等价类  $[u]_A$ 。因此,下文引入一个对应于  $A$  的模糊等价矩阵,用于形式概念分析。

在形式背景  $K=(G,M,I)$  中,如果两个对象具有相同的属性,则两个对象等价。如果  $R$  是  $M$  上的一个模糊等价矩阵,那么  $\Sigma=(M,R)$  作为一个模糊粒度知识库。 $R$  可以通过以下方式获得。如果两个元素  $O_i, O_j \in M$ , 则其相似性为:

$$\text{sim}(i, j) = \left| \frac{O_i^{\triangleright} \cap O_j^{\triangleright}}{O_i^{\triangleright} \cup O_j^{\triangleright}} \right|, i, j \in \{i, j, \dots, n\} \quad (1)$$

其模糊关系矩阵为:令  $\tilde{R} = \text{sim}(i, j)_{|\Sigma| \times |\Sigma|}$  满足自反性与对称性,多数情况下不满足传递性  $\tilde{R}^{\circ} \tilde{R} \subseteq \tilde{R}$ , 仅为模糊相似矩阵。可通过传递闭包算法获得模糊等价关系  $R$ :

$$R = \tilde{R} \cup \tilde{R}^2 \cup \dots \cup \tilde{R}^n \quad (2)$$

表1 融合本体的部分形式概念表

Tab.1 Parts of Formal Contexts of Fused Ontologies

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
1 时令湖	*		*		*		*		*						*
2 地面河流	*			*	*		*		*						*
3 时令河	*		*		*		*			*					*
4 湖泊	*		*		*		*		*		*	*			
5 池塘	*		*		*		*			*	*	*			
6 水库	*			*	*		*							*	*
7 溢洪道	*			*	*				*						*
8 干堤		*		*			*	*							*

注:从  $a$  到  $o$  依次表达物质性:水、石;成因:天然、人工,空间形态:长条线、低洼地、建筑物;空间位置:地表、地下;时间性:常年、季节性,物质状态:流动性;功能:运输、防洪、蓄水。

改造模糊关系矩阵的过程中,设定限定值进行截取,对模糊等价关系进行粒化分类。

定义2 令  $\delta \in [0, 1]$ ,  $r_{ij} = \text{sim}(i, j)$ , 对应于限定值  $\delta$  截取的等价关系:

$$R_{\delta} = (r_{ij}^{\delta}) \text{ 当 } \begin{cases} 1, r_{ij} \geq \delta \\ 0, r_{ij} < \delta \end{cases} \quad (3)$$

再令  $\Sigma/R_{\delta} = \{X_1, X_2, \dots, X_m\}$ , 则粒度  $R_{\delta}$  定义如下:

$$\rho(R_{\delta}) = \frac{1}{|M|^2} \sum_{i=1}^m |X_i|^2, i \in \{i, j, \dots, m\} \quad (4)$$

式(4)用来划分融合后的概念格粒度。  $X_i \in \Sigma/R_{\delta}$  称之为一个粒度,  $\rho(R_{\delta})$  则表示对应粒度下的粒度概念集合。引入概念相似度度量两个或多个概念间的语义距离,引入  $\mu_{\delta}^{\theta}$  表示  $\theta$  粒度下的概念集合。

定义3 设  $K_{\theta}=(G_{\theta}, M_{\theta}, I_{\theta})$  为对应  $\theta$  粒度下的形式表,  $(A, B)$  是  $K_{\theta}$  的一个概念。如果  $\exists \forall X \in M_{\theta}$ , 都有  $\mu_{\delta}^{\theta} = \left| \frac{X \cap (A, B)_M}{X} \right|$ ,  $1 - \mu_{\delta}^{\theta} \leq \theta$  成立,且当  $X \cap (A, B)_{\Sigma} \neq \emptyset$  时,则  $(A, B)$  是  $\rho(R_{\delta})$  中的一个概念。令  $(A_1, B_1), (A_2, B_2)$  为  $\mu_{\delta}^{\theta}$  上的两个概念,两个概念相似度取值区间为  $[0, 1]$ , 表示为<sup>[12]</sup>:

$$\text{sim}_{(\delta, \theta)}((A_1, B_1), (A_2, B_2)) = \frac{\left| \bigcap_{i=1,2} (A_i, B_i)_{M(\delta, \theta)} \right|}{\left| \bigcup_{i=1,2} (A_i, B_i)_{M(\delta, \theta)} \right|} \quad (5)$$

### 3 实验与分析

各抽取基础地理信息与水利工程当中的几个要素,形成类似表1的形式背景,其中,湖泊、水库、溢洪道、干堤为跨领域共同对象。语义数据融合结果见图1。

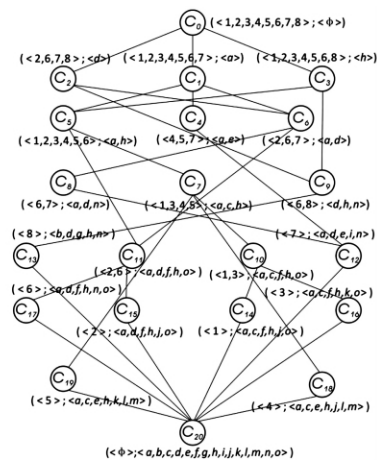


图1 融合概念格

Fig.1 Fused Concept Lattice

将表1形式背景转化为表2,根据定义1可得图2的模糊等价关系矩阵。根据定义2,获得不同融合粒度下的概念外延集合,表3中,  $\delta$  为切割模糊等价关系矩阵的阈值,  $\theta$  为粒度。当  $\delta$  增加时,  $\theta$  减小,当  $\delta$  减小时,  $\theta$  增大,可见不同粒度下,概念内涵不一样。根据定义3,分别计算  $(\delta, \theta) = (0.75, 0.16)$ ,  $(\delta, \theta) = (1.0, 0.13)$  粒度下的相似度,如下:

$$\text{sim}_{\rho(0.75, 0.16)}(C_1, C_7) = 0.33$$

$$\begin{aligned} \text{sim}_{c_{0.75,0.16}}(C_1, C_4) &= 0.67 \\ \text{sim}_{c_{1.0,0.13}}(C_1, C_7) &= 0.57 \\ \text{sim}_{c_{1.0,0.13}}(C_1, C_4) &= 0.86 \end{aligned}$$

在 $(\delta, \theta) = (0.75, 0.16)$ 粒度下,  $C_1$  是  $C_4$  的父结点,  $C_4$  是  $C_7$  的父结点。由于  $C_1$  与  $C_4$  的相似度大于  $C_1$  与  $C_7$  的相似度, 如果进行格树转化, 在同一个连通概念路径中, 父子结点相似度要高于子孙结点相似度,  $C_1$  与  $C_4$  语义距离小于  $C_1$  与  $C_7$ , 但概念格节点的父子关系和一般的树结构不同, 概念格的父节点是包含了所有具备此内涵的外延, 并且该所有外延的属性交集正好为该内涵。再比较 $(\delta, \theta) = (1.0, 0.13)$ 下的相似度, 结果表明, 地理本体概念相似度因不同粒度而有所变化, 细粒度下的概念相似度要高于粗粒度下的相似度, 这与粒计算理论在粗粒度下隐藏一些细节、简化问题求解的难度相一致。

表 2 融合领域本体库

Tab. 2 Fused Domain Ontology Base

对象	属性集( $O^D$ )	对象	属性集( $O^D$ )
$X_1$	$acfhIol$	$X_5$	$acehijklm$
$X_2$	$adfhIol$	$X_6$	$adfh nol$
$X_3$	$achjol$	$X_7$	$adehnl$
$X_4$	$acehIklm$	$X_8$	$bdghnl$

实际应用中, 随着融合本体概念数量的加大, 融合后的概念格也变得庞大复杂, 产生概念冗余。可以根据实际需求, 选择合适的粒度, 在不同粒度下约简概念格, 减少本体融合的复杂性。如表 3 中, 在 $(\delta, \theta) = (0.75, 0.16)$ 粒度下,  $X_4$  (地面河流) 与  $X_5$  (时令河) 可近似等同为一个粗概念, 而在 GB/T 13923-2006 分类中它们同属于陆地水系河流分支, 这表明  $X_4$  与  $X_5$  的内涵在此粒度下的一致性要高于其他概念, 可以将“ $X_4 X_5$ ”的联合体看作一个新的粗概念“河流”, 其内涵应该是地面河流与时令河的并集。同理, 在 $(\delta, \theta) = (0.71, 0.31)$ 下, 尽管  $X_1$  (湖泊)、 $X_2$  (池塘)、 $X_3$  (时令湖) 与  $X_6$  (水库) 占据两个领域概念, 但内涵上基本都具有在空间形态上是“低洼地”, 功能上具有“蓄水”, “ $X_1 X_2 X_3 X_6$ ”联合体也可以看作一个新的粗概念“广义湖泊”。图 3、4 分别为 $(\delta, \theta) = (0.75, 0.16)$ 、 $(\delta, \theta) = (0.71, 0.31)$ 粒度下的哈斯图, 显然在粗粒度下比细粒度下简化。因此, 在不同粒度下, 寻找属性与对象的等价特征组, 通过形式背景的全、交、并约简方法, 可去掉一些概念冗余。对融合需求比较高的情况, 可选择在细粒度下进行, 地理本体从粗粒度到细粒度的融合过程是一个逐步求精的渐进过程。

$R =$

1.00	0.71	0.71	0.44	0.44	0.71	0.57	0.50
0.71	1.00	0.71	0.44	0.44	0.71	0.57	0.50
0.71	0.71	1.00	0.44	0.44	0.71	0.57	0.50
0.44	0.44	0.44	1.00	0.75	0.44	0.44	0.44
0.44	0.44	0.44	0.75	1.00	0.44	0.44	0.44
0.71	0.71	0.71	0.44	0.44	1.00	0.44	0.50
0.57	0.57	0.57	0.57	0.44	0.44	1.00	0.50
0.50	0.50	0.50	0.44	0.44	0.50	0.50	1.00

图 2 模糊等价关系矩阵

Fig. 2 Fuzzy Equivalence Relation Matrix  $R$

表 3 不同粒度下的领域本体

Tab. 3 Domain Ontologies in Different Granulations

$\delta$	$\theta$	$M_\delta$
0.44	1	$X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8$
0.5	0.63	$X_1 X_2 X_3 X_6 X_7 X_8, X_4 X_5$
0.57	0.47	$X_1 X_2 X_3 X_6 X_7, X_4 X_5, X_8$
0.71	0.31	$X_1 X_2 X_3 X_6, X_4, X_5, X_7, X_8$
0.75	0.16	$X_4 X_5, X_1, X_2, X_3, X_6, X_7, X_8$
1	0.13	$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$

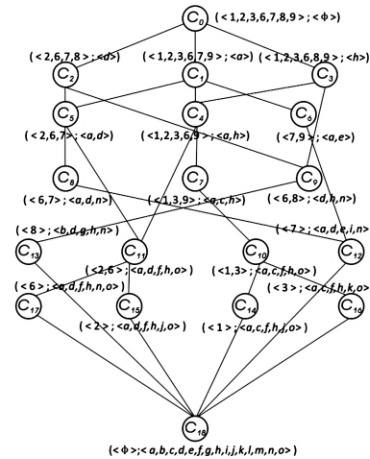


图 3  $(\delta, \theta)$ 为 $(0.75, 0.16)$ 粒度下的概念层次模型

Fig. 3 Concept Hierarchy Model Under Granularity  $(\delta, \theta)$  is  $(0.75, 0.16)$

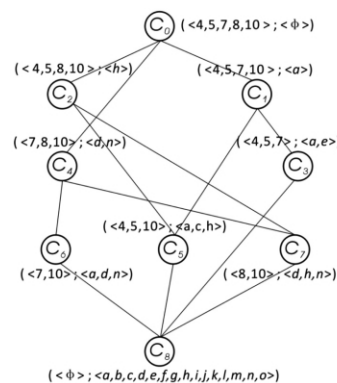


图 4  $(\delta, \theta)$ 为 $(0.71, 0.31)$ 粒度下的概念层次模型

Fig. 4 Concept Hierarchy Model Under Granularity  $(\delta, \theta)$  is  $(0.71, 0.31)$

## 4 结 语

本文针对跨领域地理本体共享问题,从语义角度给出了一种粒化理论的地理本体融合方法。该方法首先构建不同领域的形式概念表,运用概念格算法进行融合,借助粒计算理论划分粒度,再在不同粒度下约简融合后的概念格,完成不同粒度下的本体融合。此方法将形式概念分析与粒计算结合用于地理本体融合,打破了传统形式背景的二值局限性,增加了问题求解的灵活性。在粗粒度下,忽略一些描述细节,可削减融合问题的复杂程度。本文研究是约定地理本体内涵(属性)均匀且同等重要,领域本体构建部分依赖于专家。后续研究将考虑不同权值条件下的加权概念格应用,削减地理本体构建的主观性,结合其他本体融合方法进行比较研究。

### 参 考 文 献

- [1] Buccella A, Cechich A, Fillottrani P. Ontology-driven Geographic Information Integration: a Survey of Current Approaches[J]. Computers & Geosciences, 2009, 35(4): 710-723
- [2] Kokla M, Kavouras M. Fusion of Top-Level and Geographical Domain Ontologies Based on Context Formation and Complementarity[J]. International Journal of Geographical Information Science, 2001, 15(7): 679-687
- [3] Torres M, Quintero R, Moreno-Ibarra M, et al. GEONTO-MET: an Approach to Conceptualizing the Geographic Domain[J]. International Journal of Geographical Information Science, 2011, 25(10): 1 633-1 657
- [4] Fonseca F, Davis C, Camara G. Bridging Ontologies and Conceptual Schemas in Geographic Information Integration[J]. GeoInformatica, 2003, 7(4): 355-378
- [5] 李霖, 王红. 基于形式化本体的基础地理信息分类[J]. 武汉大学学报·信息科学版, 2006, 31(6): 523-526
- [6] 刘耀林, 李红梅, 杨淳惠. 基于本体的土地利用数据综合研究[J]. 武汉大学学报·信息科学版, 2010, 35(8): 883-886
- [7] 王晓璇, 陈鹏, 刘鹏, 等. 基于统计式机器学习的地理本体融合模型[J]. 同济大学学报(自然科学版) 2011, 39(5): 758-763
- [8] Ganter B, Wille R. Formal Concept Analysis, Mathematical Foundations[M]. New York: Springer, 1999
- [9] Guarino N, Welty C A. Formal Ontology of Properties[C]. The 12th International Conference on Knowledge Engineering and Knowledge Management, New York, 2000
- [10] Frank A U. Tiers of Ontology and Consistency Constraints in Geographical Information Systems[J]. International Journal of Geographical Information Science, 2001, 15(7): 667-678
- [11] Zadeh L A. Towards a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic[J]. Fuzzy Sets and Systems, 1997, 90: 111-121
- [12] Kang X P, Li D Y, Wang S G. Research on Domain Ontology in Different Granulations Based on Concept Lattice [J]. Knowledge-Based Systems, 2012, 27: 152-161
- [13] Liang J, Shi Z, Li D, et al. Information Entropy, Rough Entropy and Knowledge Granulation in Incomplete Information Systems [J]. International Journal of General Systems, 2006, 35(6): 641-654

第一作者简介: 李军利, 博士生, 讲师, 主要研究方向为地理信息本体论与地图制图。

E-mail: chzugis@gmail.com

## A Geographic Ontology Fusion Method Based on Granular Theory

LI Junli<sup>1,2</sup> HE Zongyi<sup>1</sup> ZHU Qiaoli<sup>1</sup> LIU Yahong<sup>3</sup>

(1 School of Resources and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 School of Geographic Information and Tourism, Chuzhou University, 1528 Fongle Road, Chuzhou 239000, China)

(3 Wuhan Geomatic Institute, 209 Wansongyuan Road, Wuhan 430022, China)

**Abstract:** A geographic ontology fusion method based on granular theory is presented. First,

(下转第 497 页)

## Construction of Business Ontology and Workflow Model for Land Consolidation Projects Management

LIU Yaolin<sup>1,2</sup> WANG Hua<sup>1,2</sup> TANG Xu<sup>1,2</sup> LIU Zhongqiu<sup>1,2</sup>

(1 School of Resources and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University,  
129 Luoyu Road, Wuhan 430079, China)

**Abstract:** An information system of land consolidation projects management (LCPM) based on the workflow model and business ontology is designed to solve the problem of that traditional information systems can't handle the complicated data and business of land consolidation projects simultaneously. The business ontology is established on the basis of related official documents of management measures and expert knowledge, and it is regarded as the metamodel of the workflow model of LCPM which can make the integration of workflow and dataflow, and satisfy the uncertainty of business of LCPM. Functional modules of the system with weak coupling and strong polymerization are designed in accordance with the workflow model of LCPM, and the software development pattern of plug-in framework is adopted to realize the system on computer. At last, taking the Guangxi Zhuang Autonomous Region as a case study, we find that the design of Information system of LCPM presented above is effective and feasible.

**Key words:** information system of land consolidation projects management; business ontology; workflow model; business node

---

**About the first author:** LIU Yaolin, professor, Ph.D supervisor, majors in GIS and LIS.

E-mail: yaolin610@163.com

.....  
(上接第 492 页)

Formal concept analysis (FCA) is used to formalize different geographic domain ontologies. Then granular computing is introduced to geo-ontologies fusion to implement concept lattice reduction in different granulations. Finally, experiments are conducted by combing fundamental geographic information data and spatial data in the hydraulic engineering domain. The results show that Geo-ontologies fusion is a progressive process, and the proposed method is valid in multi-source Geo-ontologies fusion. As well as merits and limitations are discussed.

**Key words:** geographic ontology fusion; formal concept analysis; granular computing; spatial data; geographical information

---

**About the first author:** LI Junli, Ph.D candidate, lecturer, majors in geo-ontology and cartography.

E-mail: chzugis@gmail.com