

基于地理本体的地下管线空间知识表示方法研究

徐 寅¹, 吉根林², 张书亮¹

(1. 南京师范大学虚拟地理环境教育部重点实验室, 江苏 南京 210023)
(2. 南京师范大学计算机科学与技术学院, 江苏 南京 210023)

[摘要] 地下管线空间知识作为连接管线建设规范与管线空间关系的桥梁,在地下管线 GIS 系统的管线数据校验等方面的重要作用已被众多研究人员、开发厂商所认同. 而针对目前地下管线空间知识存在的语义缺失与可扩展性不强等问题,本文给出了一种基于地理本体的地下管线空间知识表示方法,通过本体描述语言 OWL 构建地下管线地理本体,在此基础上通过扩展的语义网规则语言 SWRL 实现地下管线空间知识的形式化表达. 与传统的产生式及三值权位表方法相比,本体的方法使空间知识包含丰富的语义信息,易于扩展且支持复杂计算,有利于地下管线空间知识在不同系统间的共享与重用.

[关键词] 地下管线地理本体,空间知识表示,OWL,SWRL

[中图分类号]TP391 [文献标志码]A [文章编号]1001-4616(2013)01-0127-06

Research on Spatial Knowledge Representation Method of
Underground Pipeline Based on Geo-ontology

Xu Yin¹, Ji Genlin², Zhang Shuliang¹

(1. Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing 210023, China)
(2. School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, China)

Abstract: As the bridge of connecting the construction standard and spatial relationship of the underground pipeline, the spatial knowledge of the underground pipeline has been well-accepted by researchers and developers in aspects of data verification of the Underground Pipeline Geographic Information System (GIS). However, methods for representation and reasoning of spatial knowledge of underground pipeline still have existed some problems like loss of semantics and poor extendibility. Therefore, in this paper we give a spatial knowledge representation method of underground pipeline based on geo-ontology. Besides, we compare this method with the production representation and three values weight table method. The results show that, the spatial knowledge of underground pipeline based on geo-ontology can include rich semantic information, is easily extendible and support complex computations, which will be conducive to the share and reuse of spatial knowledge of underground pipeline among different system.

Key words: geo-ontology of underground pipeline, representation of spatial knowledge, web ontology language (OWL), semantic web rule language (SWRL)

随着地理信息系统 (Geographic Information System, GIS) 技术的发展以及城市地下管线信息化管理需求的日益提升,地下管线 GIS 的建设与应用也愈加深入与普及. 作为连接管线建设规范与管线空间关系的桥梁,地下管线空间知识在地下管线 GIS 系统中的管线数据校验、管线空间分析、数据模型构建等方面的重要作用已被众多研究人员、开发厂商所认同.

地下管线空间知识主要是指管线与相关行业以及管线行业内各要素间的空间关系,在实际应用中一般选用国家或地方标准^[1-3]规定的地下管线规则型空间知识(简称规则). 国内外对于空间知识表示方法的研究由来已久,最初的空间知识往往表示为文字或地图,随着信息科学的引入,原有的手绘地图的空间

知识表示方法逐渐被 GIS 工具所代表的计算机制图替代,并且引入了形式化的表示方法,如 Kondaveeti 等^[4]采用产生式规则表示空间关联关系,用来从传感器获得的数据中检测美国与墨西哥边境的入侵行为. Yu 等^[5]基于 GeoAgent 提出了一种时空表达的框架,由领域、对象、时间、GeoAgent 以及关系等五元组构成. 韩春华等^[6]将铁路选线知识分为面向领域人员和面向计算机知识,用语义网络表示面向领域人员知识概念. 王生等^[7]基于经典描述逻辑 ALC,提出了空间语义知识的表示方法和可判定的推理算法,实现了空间关系和语义知识相合并且可判定的推理. Ballatore 等^[8]基于语义爬虫工具从 Wiki 页面中提取地理词汇并通过 RDF 保存提取的地理知识,并通过同被引(co-citation)算法进行语义相似度计算. 这些形式化的空间知识表示方法易于通过计算机实现自动识别与处理,在实际生产中得到了广泛应用. 而针对地下管线领域的空间知识表示,国内的一些学者研究了管线空间知识的获取和表达,并设计管线数据校验系统. 如许婷^[9]研究并构建了管线知识管理系统,采用“三值权位表”的形式存储并表达管线规则;王国荣等^[10]从管线数据空间关系的角度提取规则,规则的表示以数据形式存储于数据库中;陈继山等^[11]梳理了管线监理规则,并采用产生式规则表达管线规则. 可见,现有地下管线空间知识往往通过产生式^[11]或数据库二维表^[9]进行存储及表示,规则的表示形式不统一,并且在规则中含有缺乏语义支撑的自然语言描述,从而限制了地下管线空间知识在多个系统中使用的可适应性与可扩展性.

本体是概念的明确解释或表示,通过刻画领域概念及概念之间的关系,可以为构成知识内容的概念、实例及关系提供形式化的、唯一性的定义规范. 本体理论在知识的形式化表达与共享方面具有独特的优势. 用本体表达知识具有本体描述的准确性、本体语义的丰富性、以及本体规则公理的可扩充性等特点^[12-14]. 因此,本文给出了一种基于地理本体的地下管线空间知识的形式化表示方法,通过本体描述语言,构建了地下管线的地理本体,在此基础上通过扩展的本体规则描述语言实现了对地下管线空间知识的表示.

1 地下管线地理本体的构建

1.1 地下管线地理本体模型

在建立本体时,首先应明确本体所涉及的范围,以确定该领域所包含的概念、属性和关系. 本文参考地下管线相关行业标准^[1-3],结合与管线领域专家的交流,根据领域中的概念分类体系与数据组织方式,形成地下管线领域概念中的语义层次关系. 同时,地理本体作为一类特殊的领域本体,除了对概念之间的语义层次关系的定义,还应支持地理要素特有的几何形态、空间关系、坐标系等概念的描述. 本文在宋佳等^[15]提出的时空地理本体模型的基础上实现对几何特征与空间关系的描述. 形成的地下管线地理本体模型如图 1 所示.

(1) 要素模型 (Feature)

在地下管线的实际生产和管理过程中,管线数据往往是以管线和管点两类地理要素来组织. 因此,管线本体要素模型的概念体系也从管线和管点的角度进行层次划分,分类对象为跨越省市的长距离输送管线、城市地下管线以及企业的特种管线等. 其中,管线以级别、功能、功能主特征为分类依据,共分为电力管线、通信管线、给水管线、排水管线、燃气管线、热力管线、工业管线、综合管沟、不明管线共 9 类,每一类根据功能再细分为若干子类,如给水管线仍可继续分为原水管线、输水管线、中水管线、直饮水管线、消防水管线、绿化水管线等,共 53 个子类.

管点作为管线普查中的探测点,根据功能可分为设备、设施、配件等. 其中设备是指经过加工制造,由多种材料和部件组成的用于管线的建设的机械、装置,如阀门、水表、消火栓等. 设施是指为管线建设而建立的固定的和永久性的建筑、场地等,如水源井、泵站、调压房等. 配件是指为管线建设所需的原材料和经过工业加工的设备本体以外的零配件、附件、成品、半成品,如三通、四通等. 而一些特殊的管点,如管线探测中的隐蔽点,可归入其他类别.

(2) 几何模型 (Geometry)

在 GML 中的几何模型主要包括原生几何形状中的点图形 (Point)、直线图形 (LineString)、多边形图形 (Polygon),及它们的聚合几何形状 (MultiPoint, MultiLineString, MultiPolygon). 几何坐标模型以 GML 的坐标结构为基础. GML 提供了 DirectPosition、Coord、Coordinates 3 种坐标模式. 本体几何模型中的坐标本体采用了 GML 中的 Coordinates 坐标模式.

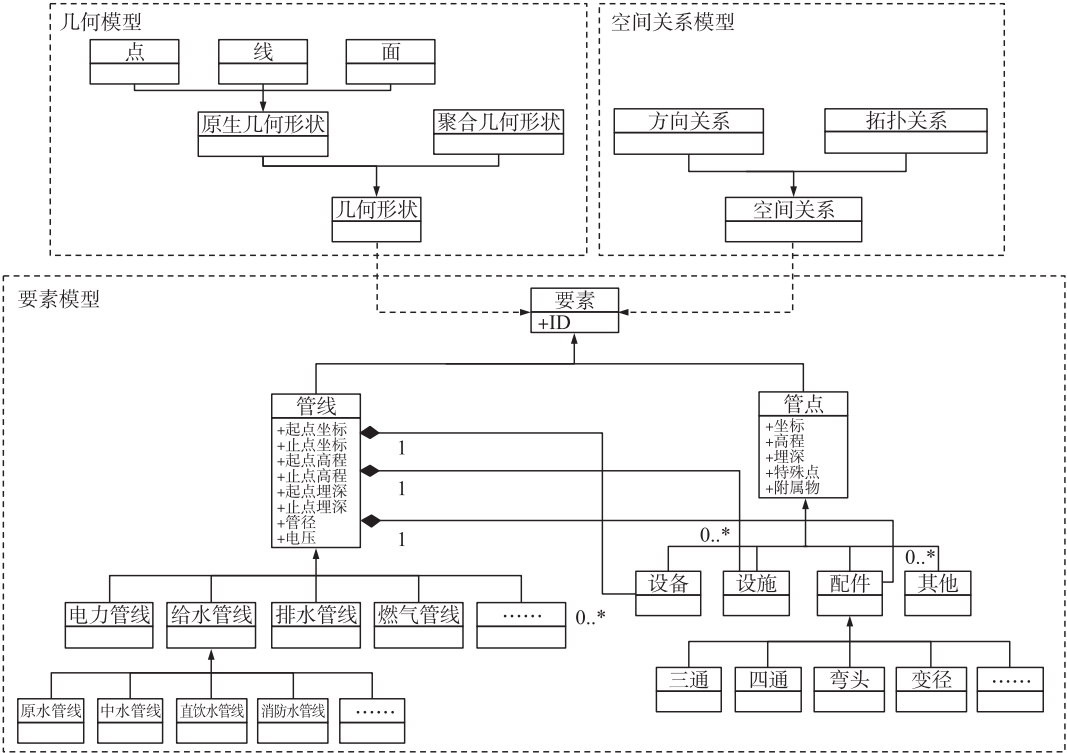


图1 地下管线地理本体模型图

Fig.1 Geo-ontology model of underground pipeline

(3)空间关系模型(SpatialRelationship)

空间关系模型主要表达拓扑关系和方向关系两种空间关系. 拓扑关系是空间关系中最重要的一类关系. 根据空间推理相关理论,主要有两类拓扑关系模型:求交模型和 RCC 逻辑模型,两者在描述两个简单区域间的面-面拓扑关系时,其结果是一致的. 因此拓扑关系本体构建的是两个简单区域间的 8 种面-面拓扑关系,即 { Disjoint, Touch, Overlap, Equal, CoveredBy, Cover, Within, Contain }.

方位关系是空间对象方位间的关系,它是一类重要的定性空间关系,描述的是空间对象间的一种顺序,如:前、后、左、右、上、下、东、南、西、北等. 空间方向有 5 个性质:传递性、反射性、完整性、平等性、对等性. 文中方向关系的构建以“基于投影”的划分方法为基础. 共有 N(北)、S(南)、E(东)、W(西)、NE(东北)、NW(西北)、SE(东南)、SW(西南)8 种关系.

1.2 地下管线地理本体的创建

(1)本体类的定义

类(Class)是领域内的概念,是有相似属性的元素集合,定义类的同时,还要考虑类之间的继承关系. 根据地下管线领域本体模型,继承自基类 owl:Thing 的分别有 Feature、Geometry 两类,并依次创建其子类,各子类通过 subClassOf 与父类关联. 对于类之间的关系,如各类地下管线是不相交的,因此可以用公理 disjointwith 来描述.

(2)属性的定义

属性(Property)代表一种二元关系,分为数据属性(Datatype Property)和对象属性(Object Property)两类,每种均有定义域(Domain)和值域(Range). 数据属性代表某类与数据类型之间的关系. 如管点拥有数据属性 hasCoordinate(坐标)、hasElevation(高程)、hasBurialDepth(埋深)等. 对象属性代表类与类之间的关系. 如一条管线上会附加各类设备设施与配件,因此管线与管点间存在 hasEquipment、hasFacility 和 hasAccessory 等关系.

本文基于网络本体语言(Web Ontology Language, OWL)^[16]来创建地下管线地理本体. 相对于其他本体描述语言,OWL 遵循面向对象的思想,按类和属性的形式描述领域概念所包含的结构,并具有逻辑描述和演算能力. OWL 分为 3 种子语言:OWL Full、OWL DL 和 OWL Lite. 考虑 3 种子语言的表达能力和推理能力,本文通过 Protégé 工具本体构建工具,采用 OWL DL^[17]对地下管线本体进行形式化描述. 部分类及属

性的 OWL 语言描述如下：

```
<owl:Class rdf:ID="Tee">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">三通</rdfs:comment>
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Accessory"/>
  </rdfs:subClassOf>
</owl:Class>
.....
<owl:DatatypeProperty rdf:ID="hasElevation">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">管点高程</rdfs:comment>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
  <rdfs:domain rdf:resource="#PipePoint"/>
</owl:DatatypeProperty>
```

2 基于地理本体的地下管线空间知识表示

2.1 基于本体规则描述语言的规则表示

本文中的地下管线空间知识主要来源于国家和地方的相关行业标准^[1-3]，涉及管线的取值范围、命名规范、最小覆土深度，地下管线间的最小水平净距、最小垂直净距等。以面向管线数据校验的地下管线空间知识为例，主要包括取值合理性规则、命名规范性规则、覆土深度规则、水平净距规则等。

OWL 本身可以作为一种知识的表示语言，其主要作用是将领域内的概念分类，通过约束、布尔组合来描述概念自身的语义含义，利用属性来表达概念与概念之间复杂的依赖，所以通过本体可以得到一个类的实例所具有的属性，以及类与类之间的关系。但 OWL 所拥有的描述能力主要来自于以类别为基础的关联性的推理，难以描述以 If...Then...表示的产生式规则表示。因此本文在已构建的地理本体的基础上，使用语义网规则语言(Semantic Web Rule Language, SWRL)^[18]对地下管线规则进行形式化表示，并通过 Protégé 本体工具实现对规则的可视化编辑及存储。

SWRL 主要由 4 部分组成，分别是 Imp、Atom、Variable 和 Built-In。Imp 中包含 antecedent (body) 推论前件和 consequent (head) 推论后件。head 和 body 中所有的例子都是由 Atom 的合取构成，不支持否定和析取操作。Atom 是由限制式构成，例如 Gas(?x)，hasLocation(?x, ?a) 等。可见限制式均由本体的类或属性构成，这就保证了规则描述中携带的语义。Variable 用来记录 Atom 中所使用的变量部分。Built-In 是 SWRL 中模块化的组件，赋予 SWRL 计算能力，SWRL 提供了基本的数学操作和字符串及日期操作，如表 1 所示，并允许用户自定义的扩展以实现复杂运算。

表 1 SWRL 的 Built-In 中的各种逻辑比较关系
Table 1 Logical comparison operation in Built-In of SWRL

类型	举例
数值比较	<swrlb:equal>,<swrlb:notEqual>,<swrl:lessThan>,<swrl:greaterThan>等
数学计算	<swrlb:add>,<swrlb:subtract>,<swrlb:mutiply>,<swrlb:deride>等
字符串处理	<swrlb:stringConcat>,<swrlb:stringLength>,<swrlb:normalizeSpace>等
日期处理	<swrlb:yearMonthDuration>,<swrlb:dayTimeDuration>,<swrlb:dateTime>,<swrlb:date>等
URI 处理	<swrlb:resolveURI>,<swrlb:anyURI>

例如，一个地下管线覆土深度规则“燃气管线的最小覆土深度为 0.6 m”，用 SWRL 语言表示为

```
PipeLine(?x) ^ has GXZL(?x, Gas) ^ hasQDMS(?x, ?qdms) ^ hasZDMS(?x, ?zdms) ^ swrlb:greaterThanOrEqual(?qdms, 0.6) ^ swrlb:greaterThanOrEqual(?zdms, 0.6) ^ PipeLine(?allX) ^ has GXZL(?allX, Gas) ^ sqwrl:makeSet(?s, ?x) ^ sqwrl:makeSet(?allS, ?allX) ^ sqwrl:difference(?leftS, ?allS, ?s) ^ sqwrl:element(?e, ?leftS) -> sqwrl:select(?e)
```

其含义为：找出在起点和止点埋深有值，并且皆大于或等于 0.6 m 的燃气管线 x，以及所有的燃气管线 allX，分别存储到集合 s 和 allS 中，利用集合 s 的补操作，则剩余集合 leftS 中的元素即为埋深小于 0.6 m

的管线数据.

2.2 扩展 Built-In 的地下管线复杂规则表示

由 SWRL 提供的 Built-In 可以实现基本的计算能力,但无法满足更复杂的规则表达. 如对于以下的水平净距校验规则:“给水管与排水管的最小水平净距为 1.0 m”,涉及空间中水平方向距离的计算,因此需要对 SWRL 的 Built-In 模块进行自定义扩展. Protégé 的 API 提供了 SWRLBuiltInBridge 机制,允许用户以 Java 语言实现自定义 Built-In 扩展^[19],包括以下两个步骤:

(1)在 OWL 本体描述中定义 Built-In 扩展. 由于在 SWRL 中定义每个 Built-In 均为类 `swrl:BuiltIn` 的实例,因此,首先创建一个描述 Built-In 扩展的 OWL 文件 `pipe.owl`,任意给定一个命名空间,如“`xmlns="http://mydomain.com/ontologies/pipe.owl#"`”,在其中声明计算线段距离的 Built-In“`<swrl:BuiltIn rdf:ID="distanceOfPipePoint"/>`”,最后通过建模原语 `owl:imports` 将其导入管线本体的定义中,使规则描述时能够加载自定义扩展.

(2)使用 Java 实现 Built-In 扩展. SWRLBuiltInBridge 要求所有 Built-In 的实现类限定为 `SWRLBuiltInLibraryImpl`,且 Java 包名的前缀必须为 `edu.stanford.smi.protege.owl.swrl.bridge.builtins`,具体的实现算法定义为类中相应的方法.

对于两条管段水平净距的计算,可转化为管段的起点或止点间水平距离的计算. 若起点和止点间距离均满足最小水平净距的要求,则管段之间也符合要求. 此外,由于水平净距指的是管段水平方向外边缘的距离,而管段起止点测量的是中心点的位置,因此在水平净距的计算中,端点的水平距离求出后还应减去各自管径的一半. 部分关键代码如下:

```
//方法定义,遵循固定格式
public boolean distanceOfPipePoint( List<BuiltInArgument>arguments) throws BuiltInException
{
    boolean result=true;
    //验证传入参数的数量
    checkNumberOfArgumentsEqualTo(7,arguments.size());
    //获取每个传入参数的值
    double distance=getArgumentAsADouble(0,arguments);
    double qdx1=getArgumentAsADouble(1,arguments);
    double qdy1=getArgumentAsADouble(2,arguments);
    .....
    //计算距离
    distance=Math.sqrt(Math.pow((qdx1-qdx2),2)+Math.pow((qdy1-qdy2),2))+(gj1+gj2)/2;
    return result;
}
```

由此,根据自定义的 Built-In 扩展 `pipe:distanceOfPipePoint`,上文的水平净距规则用 SWRL 语言表示为

```
PipeLine(?water) ^ hasGXZL(?water,WaterSupply) ^ hasGJ(?water,?gj1) ^ hasLocation(?water,?waterlocation) ^ hasLineCoor
(?waterlocation,?line) ^ X(?line,?qdx1) ^ Y(?line,?qdy1) ^ PipeLine(?dra) ^ hasGXZL(?dra,Drainage) ^ hasGJ(?dra,?gj2)
^ hasLocation(?dra,?dralocation) ^ hasLineCoor(?dralocation,?draline) ^ X(?draline,?qdx2) ^ Y(?draline,?qdy2) ^ pipe:
distanceOfPipePoint(?distance,?qdx1,?qdy1,?qdx2,?qdy2,?gj1,?gj2) ^ swrlb:lessThanOrEqual(?distance,1) ^ PipeLine
(?waterAll) ^ hasGXZL(?waterAll,WaterSupply) ^ sqwrl:makeSet(?s,?water) ^ sqwrl:makeSet(?allS,?waterAll) ^ sqwrl:
difference(?leftS,?allS,?s) ^ sqwrl:element(?e,?leftS) ^ sqwrl:select(?e)
```

本文与文献[3]给出的产生式方法(如表2所示)及文献[4]给出的三值权位表方法(如表3所示)进行了对比分析,首先由于本体对概念层次及概念间关系的明确描述,使得组成空间知识的概念也带有准确的语义信息,而非带有二义性的自然语言描述,并且可以通过本体的推理工具,如 Pellet 等检查空间知识定义的一致性和独立性;另外,通过自定义 Built-In 扩展,空间知识可以支持复杂的数值及逻辑运算,这也是传统的表示方法难以实现的;第三,通过标准的本体操作接口,可以供不同系统调用并实现基于规则的推理,方便规则的定制及重用.

表 2 基于产生式的地下管线空间知识表示方法

Table 2 Spatial knowledge representation method of underground pipeline based on production

规则	表述
Rule 1	IF 铺设方式=4 Or 铺设方式=9 THEN 地板厚>0
Rule 2	IF 铺设方式 In(2,3,7) THEN 已用孔数>0
Rule 3	IF 铺设方式=4 Or 铺设方式=9 THEN 起点覆土深度=起点内底埋深-0.001*(Last(断面尺寸)+顶板厚)

表 3 基于三值权位表的地下管线空间知识表示方法

Table 3 Spatial knowledge representation method of underground pipeline based on three values weight talbe

设备 1	设备 2	空间关系	最小值	最大值
给水管	雨水污水排水管	线线水平相离	1.5	∞
大灯	路灯线	点线相接	0	2
污水管	污水处理厂	点线相通	0	0

3 总结与展望

本文依据地理本体的构建原则,采用 OWL 本体描述语言构建并表达了地下管线地理本体.在此基础上,使用扩展的 SWRL 实现地下管线空间知识的形式化表达.通过与传统地下管线空间知识表示方法的对比分析,表明基于地理本体描述的地下管线空间知识包含丰富的语义信息,在实际应用中具有可扩展性与可适应性,有利于管线知识的共享与重用.不过,由于基于本体的空间知识描述的规范性,在表达复杂知识的同时也带来了表达式结构的冗长,不够直观与易理解,如何对基于本体的空间知识表达进行优化,也是下一步需要研究的问题.

[参考文献]

[1] 关增义,刘绍治,王健,等. GB50289 城市工程管线综合规划规范[S]. 北京:中国建筑工业出版社,1998.

[2] 戚盛豪,万玉成,于超英,等. GB50013 室外给水设计规范[S]. 北京:中国计划出版社,2006.

[3] 金石坚,李颜强,徐良,等. GB50028 城镇燃气设计规范[S]. 北京:中国建筑工业出版社,2006.

[4] Kondaveeti A,Runger G,Liu H,et al. Extracting geographic knowledge from sensor intervention data using spatial association rules[C]//Proceedings of 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services(ICSDM). Taiwan:IEEE Press,2011:127-130.

[5] Yu C,Peuquet D J. A GeoAgent-based framework for knowledge-oriented representation:Embracing social rules in GIS[J]. International Journal of Geographical Information Science,2009,23(7):923-960.

[6] 韩春华. 基于 GIS 的铁路选线系统智能环境建模方法研究[D]. 成都:西南交通大学土木工程学院,2008.

[7] 王生生,刘大有,谷方明,等. 基于 ALC 的空间语义知识库[J]. 吉林大学学报:工学版,2011,41(2):424-429.

[8] Ballatore A,Bertolotto M,Wilson D C. Geographic knowledge extraction and semantic similarity in OpenStreetMap[J]. Knowledge and Information Systems,2012:1-21.

[9] 许婷. 基于“知识与规则”的城市管线 GIS 数据校验研究[D]. 南京:南京师范大学地理科学学院,2005.

[10] 王国荣,孙毅中. 基于知识与规则的城市管线空间关系及数据校验[J]. 江苏城市规划,2006(1):40-43.

[11] 陈继山,戴悦. 基于知识与规则的城市地下管线数据监理[J]. 苏州科技学院学报:自然科学版,2009,26(2):76-80.

[12] Lutz M,Klien E. Ontology-based retrieval of geographic information[J]. International Journal of Geographical Information Science,2006,20(3):233-260.

[13] Janowicz K,Keßler C. The role of ontology in improving gazetteer interaction[J]. International Journal of Geographical Information Science,2008,22(10):1 129-1 157.

[14] 黄勇奇,牛振国,崔伟宏. 基于地理本体和 SWRL 的时空推理研究[J]. 武汉理工大学学报:交通科学与工程版,2009,33(6):1 175-1 178.

[15] 宋佳,诸云强,王卷乐,等. 基于 GML 的时空地理本体模型构建及应用研究[J]. 地球信息科学学报,2009,11(4):442-451.

[16] Antoniou G,Harmelen F. Web ontology language:OWL[J]. Handbook on Ontologies,2009:91-110.

[17] Baader F. The Description Logic Handbook: Theory, Implementation, and Applications [M]. London: Cambridge Univ Press,2003.

[18] Chi Y L,Chen I J. An ontological partner network model for composing suppliers[C]//Proceedings of 2011 International Symposium on Innovations in Intelligent Systems and Applications(INISTA). Fuzhou:IEEE Press,2011:475-479.

[19] Keßler C,Raubal M,Wosniok C. Semantic rules for context-aware geographical information retrieval[J]. Smart Sensing and Context,2009:77-92.

[责任编辑:丁 蓉]