

# 一种领域本体构建方法及其在相片管理中的应用<sup>①</sup>

梁婷婷, 李春青

(广西民族师范学院, 崇左 532200)

**摘要:** 自 20 世纪 90 年代以来, 本体逐渐被应用到很多领域, 本体的构建方法论也成为国内外机构的热门研究课题。在分析和总结现有的本体构建经验的基础上, 提出了一种领域本体的构建方法, 它符合人们的思维认知规律, 可操作性和扩展性强。并使用该方法为相片管理领域进行本体建模, 实验表明了该方法的科学性和有效性。

**关键词:** 本体; 领域本体; 构建方法; 相片管理

## Construction Method of Domain Ontology and Its Practice in Photo Management

LIANG Ting-Ting, LI Chun-Qing

(Guangxi Normal University for Nationalities, Chongzuo 532200, China)

**Abstract:** Since 1990s, ontology has been applied to many domains gradually. Construction method of domain ontology has become a very hot research topic. On the basis of analyzes and summarizes the existing experience of ontology construction, this paper proposes a domain ontology construction method. The methodology is in accordance with human's thoughts, and it owns strong operability and expansibility. Moreover, it uses the method and constructs photo domain ontology. The experiments show that the proposed approach is very scientific, effective and promising.

**Key words:** ontology; domain ontology; construction method; photo management

自 20 世纪 90 年代以来, 本体 (Ontology) 在通信、互操作、系统工程领域和信息检索方面有很大应用<sup>[1]</sup>。在通信 (Communication) 方面, 本体提供的统一术语使得人们和组织的相互交流准确无歧义; 要实现不同系统之间的互操作 (Inter-operability), 不同的软件、语言或模型方法可用本体进行转换映射; 在系统工程 (Systems Engineering) 方面, 由于本体对重要的类、属性、过程及其关系等进行了形式化表示, 使得软件系统的可重用性、共享性、可靠性、规范性得到了保证和提高; 在信息检索 (Information Retrieve) 方面, 本体的使用使得计算机检索的层次从传统的关键字检索提升到了语义检索, 而 (Onto)2Agent, Ontobroker 就是在信息检索中应用本体的著名项目。总的来说, 作为“共享概念模型的明确的形式化规范说明”<sup>[2]</sup>的本体是通过统一的术语和概念实现某程度上的知识共享和重用, 甚至使得计算机从语义层次上理解信息或语言。

## 1 知名的本体构建方法

随着本体的逐渐应用, 本体的构建方法论成为国内外机构的热门研究课题。大多数比较成功的本体构建方法是通过逆工程总结出来的, 并借鉴了软件工程方法, 知名的本体构建方法有: 七步法、生命周期法 (Methontology)、骨架法 (Skeletal Methodology)、企业建模法 (Gruninger&Fox 的本体建立模式 TOVE)、Meth 方法、IDEF-5 方法等等。以下详细介绍相对完整且成熟度较高的七步法、生命周期法, 以及本体构建指导方针的骨架法。

### 1.1 七步法

斯坦福大学开发的七步法<sup>[3]</sup>包括七个步骤:

- (1) 确定本体的专业领域、范围和场景;
- (2) 考察现有的本体并考虑复用;
- (3) 列出本体中的重要概念;
- (4) 定义本体中的类和类的等级体系;

<sup>①</sup> 收稿时间:2011-08-09;收到修改稿时间:2011-09-19

- (5) 定义类的相关属性;
- (6) 定义属性的分面 (Facets);
- (7) 创建实例。

七步法对构建本体的相关技术和方法细节做出了详细描述,且考虑到了软件复用但是没有提到本体的评价和优化步骤。

## 1.2 骨架法(Skeletal Methodology)

Mike Uschold & King 的骨架法(Skeletal Methodology)<sup>[4]</sup>是爱丁堡大学从开发企业本体(Enterprise Ontology)的经验中产生,此法只能作为开发本体的指导方针。过程如下:

- (1) 根据任务情况确定本体的应用目的和范围。

(2) 构造本体阶段:首先对本体进行分析,捕获相关领域中的概念和关系,产生无歧义的精准定义,然后利用某种形式化语言对这些概念和关系进行编码表示,最后集成该阶段的概念和关系以及来自其他领域的概念。

(3) 本体的评价阶段:根据前面确定的本体需求和能力问题对本体以及软件环境和相关文档进行评价,若符合要求,则形成文档;否则,重新回到构造阶段。

骨架法对构建框架和各阶段的指导方针有参考价值,且要求文档化,有本体评估的步骤,但每个阶段都没有具体的方法(包括本体评估方法)和技术。

## 1.3 生命周期法 (Methontology)

生命周期法 (Methontology)<sup>[5]</sup>是西班牙马德里理工大学人工智能实验室在开发人工智能图书馆时使用的,生命周期的概念贯穿整个本体的开发过程,此思想更接近软件开发方法。过程如下:

(1) 管理阶段:系统规划任务的进展情况、资源情况以及质量保证问题等。

(2) 开发阶段:这一阶段包括:规范说明本体的目的、形式化程度、范围等,并产生本体规格说明书;通过专家、图书等各种渠道获取本体的相关知识;识别任务的领域术语,将知识概念化;从已建好的其他本体中选择可用的术语定义,并集成到本本体中;用合适的语言形式化编码本体。

(3) 评价和维护阶段:在本体生命周期的各个阶段,评价本体和文档的正确性和有效性;维护范围包括知识获取、系统集成、评价、文档说明、配置管理等等。

生命周期法首次提出撰写本体的规格说明书,详

细描述了相关技术工具和知识获取方法,适合大型本体的开发,但是没有明确提出文档化过程,也没有考虑软件的复用。

尽管本体已经逐渐得到应用,但是在不同的学科领域和具体工程,本体的构建方法不同,也没有权威标准化机构认证的方法。由于不同的应用需求,不存在某种全部适用的途径或模式,也正是如此,各种方法体系都有它各自的特点及与其适用的领域。

## 2 构建领域本体方法的创新

领域本体(Domain Ontologies)是用于描述特定领域(如教育、医药等)中的概念及概念之间的关系,及其原理和基本法则有关的词汇,是一种专门本体。大家普遍认为构造特定领域的本体需要领域专家积极参与指导,然而,本体的相关知识形式化编码及推理又离不开 IT 人员的工作,通常,这两者的合作是不容易的。现实中,一般是由对两方面知识技术都有所了解又非领域专家的知识人员来做。为此,在领域专家无法大力支持及参与的情况下,可借助原型法的思想,注意考虑本体建设的可扩展性,先让知识人员建立领域本体框架,再在进化阶段扩展和完善本体。在选择本体构建方法时,本研究没有简单地选用某一种方法,而是在现有的本体构建经验的基础上,以七步法的思路为基础,综合生命周期法 (Methontology) 和软件工程的原型法,并在整个生命周期中以骨架法为本体构建的指导方针。可称之为原型七步法,具体的开发流程见图 1。该法构建本体的过程是螺旋式上升的,且符合人们的思维认知规律,可操作性及扩展性强。

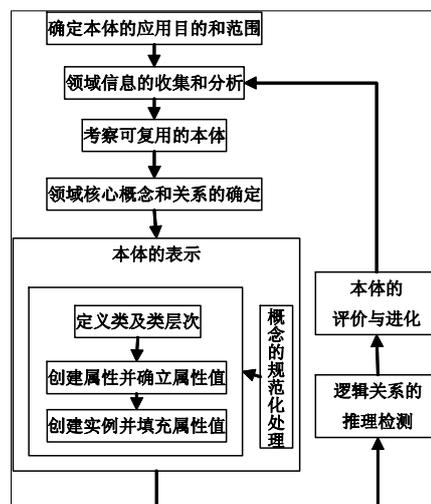


图 1 领域本体构建流程图





义和公理作为一种标准。在家庭数码照片领域,此过程只适用于构建本体的主要框架和基本的类层次时,却不太适合创建实例的过程中,因为用户更有可能需要个性化。因此,把创建实例的管理留给用户操作。

本体的表示即是以上定义类及类层次和创建属性这一系列过程,通过本体开发工具和本体表示语言来建立本体,使得计算机系统可理解其领域知识。本文则选取开发者首选的本体开发工具 Protégé,并采用 W3C 推荐的 OWL 语言描述本体概念体系。

### 3.6 逻辑关系的推理检测

本体的概念及概念间的关系通过本体表示语言描述和揭示出来,这可以说是一种显式的说明。而含在这其中的隐性知识(隐含的关系或新的分类),则靠一种处理机制——本体的推理将其提取出来。本文在分析清楚本体推理的应用需求基础上,采用 OWL DL(Description Logic)语言编写,基于 SWRL 推理规则,有效地组织推理机制。比如,已知有小孩,并且是女性,则可推理出是的女儿。那么推理规则可这么写:

由于本体中的各个概念及概念之间的关系是复杂的,很多时候有多人协作操作,所以可能会出现逻辑上的错误、不一致或冗余过多的问题,因此,有必要对构建出的本体进行逻辑上的检验和推理。在家庭数码照片领域本体中,利用 Racer 推理机与 Protégé 的结合,对本体模型进行一致性、包含性、实例和冗余检测,以确保本体的逻辑正确性。

### 3.7 本体的评价与进化

对本体的评价涉及本体是否满足需求所设定的目标,类及类层次的定义是否清晰、客观、完全、一致和可扩展,虽然目前没有统一的评价标准,毋庸置疑的是本体在实际应用中效果好才是真正有价值的。本文采用 T.R.Gruber 的 5 条规则<sup>[7]</sup>对领域本体进行评价,还将本体应用于数码照片标注与检索系统,以验证本体的有效性和科学性。

领域的知识无穷尽,领域的边界更是模糊的,领域之间也总是存在交叉,对于一般的研究机构和应用单位来说,只有通过原型法,先建立出核心的可用本体,然后在使用过程中不断扩展,比如通过集成新的本体或定义新的概念和关系等方式进行完善本体。本文就是在语义标注或语义检索的过程中发现了新的原本体没有的知识,从而实现对本体的进化的。

以上是本研究所采用的领域本体构建的七个主要

步骤,本体的开发应该是一个反复迭代、螺旋式上升的过程,初步构建好本体后,还需要相关人员进行长期的持续的本体完善工作,直至满意。

## 4 本体的应用

本文通过采用 PhotoStuff 工具进行基于本体的照片标注和检索操作,以验证所构建的本体的有效性。PhotoStuff 是一个图片标注工具,它允许用户使用本体中的概念来标记图片的某个区域。它的主要功能是导入图像、本体、实例、标注图片和导出已标注的结果到本体磁盘或网络上。而使用 PhotoStuff 基本功能的前提是本体和图像。图 6(左)显示了其主要工作界面。用户可通过“File”菜单载入本体和预处理的图片,然后使用本体中的概念为该图片进行标注,如拖拽右下角的 Class Tree 标签中所列的“父亲”概念到图片中,并创建实例,最后可将所标注的信息导出并存储到本体磁盘或网络上。当用户需要查询某些照片时,可在界面的下方输入要检索的信息,可选择限制的概念组,如用户输入了“qiang”并进行检索,PhotoStuff 返回了所有包含标注为“qiang”的相片的文件列表,见图 6(右),点击相应的文件名,即可查看相应照片。家庭数码照片领域本体在 PhotoStuff 中的成功应用,充分展示了该本体的有效性。

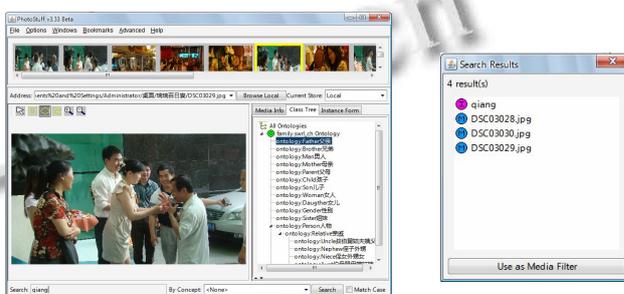


图 6 基于本体的照片标注和检索工具 PhotoShuff 截图

## 参考文献

- 1 何海芸,袁春风.基于 Ontology 的领域知识构建技术综述. 计算机应用研究,2005,3:14-18.
- 2 Studer R, Benjamins VR, Fensel D. Knowledge Engineering, Principles and Methods. Data and Knowledge Engineering, 1998,25(122):161-197.
- 3 Noy NF, Grun IM. Ontology development 101: a guide to

(下转第 104 页)

总而言之,采用 FME 进行数据转换,转换结果数据较以前采用直接导入源目标文件或使用各软件台自带的转换接口进行转换,然后进行大量的人工处理制作,其效率有了很大提高,能够进行大批量的自动处理,并且没有数据丢失的情况。基于 FME 平台的空间数据语义转换方法通过在转换过程中重构数据,改变数据表示方法,可以进行空间几何数据和属性数据的同时转换,基本上实现 GIS 应用程序间数据的快速、无损地移植,解决多年来制约 GIS 空间数据共享和标准化的瓶颈问题。但因为各 GIS 件平台自身对不同数据元素的定义格式、表达方式以及功能的差异使得转换不能做到完全一致。

### 参考文献

- 1 陈述彭.地理信息系统导论.北京:科学出版社,2000.
- 2 承继成,李琦.数字城市理论与应用.北京:科学出版社,2001.
- 3 于洪伟,林富明,李梅.哈尔滨市基础地理信息数据库建设.测绘与空间地理信息,2004,12.
- 4 奚长元.城市空间地理数据基础设施的建设与应用.行业论坛,2004.
- 5 蒋捷,陈军.基础地理信息数据库更新的若干思考.测绘通报,2000(05).
- 6 陈建光,吴洪举.城市规划建设与管理信息系统信息规范化探讨.北京测绘,2000,4.
- 7 杨正华.城市 GIS 空间数据共享初探.期刊论文-测绘通报,2003(5).
- 8 黄林进.城市地理信息系统空间基础数据建设探讨.期刊论文-城市勘测,2003(4).
- 9 黄坚,王丹,丁军城.市公共基础空间数据库建设方法初探.期刊论文-测绘通报,2003(2).
- 10 苏峰,黄正军.GIS 空间数据管理模式探讨.期刊论文-计算机仿真,2003(8).
- 11 戴玫.FME 在 DLG 数据入库中的应用.期刊论文-甘肃农业,2006(6).
- 12 徐昌荣,沈晶,施魁元.基于 FME 的 GIS 互操作讨论.江西理工大学学报,2006(6).
- 13 李刚,朱庆杰,张秀彦,王志涛.基于 FME 的城市 GIS 基础空间数据格式转换,2006(4).
- 14 熊登亮,贵仁义.基于 FME 的空间数据处理实现.四川测绘,2007(4).
- 15 徐景中.基于 FME 的空间数据语义转换技术的应用.昆明理工大学学报,2005(4).
- 16 钱业宏.基于 FME 实现 AutoCAD 数据格式相互转换-城市勘测,2007(4).
- 17 WEN Xue-dong.Based on the CAD to Geodatabase data storage of ARCGIS-SCIENCE OF SURVEYING AND MAPPING,2006,31(6).
- 18 CHEN Ying.Using FME for GIS data transformation- SCIENCE OF SURVEYING AND MAPPING,2007,32(2).
- 19 Zeng Qiaoling.Using FME to achieve the semantics of GIS and CAD conversion-COMPUTER ENGINEERING AND APPLICATIONS,2005,41(13).
- 20 FME Readers and Writers(FME 2007) ,Safe Software, Inc.May. 2007.

(上接第 144 页)

- creating your first ontology. Technical Report KSL201205 [R]. 2001.
- 4 Usehold M. Ontologies principles, methods and applications. Knowledge Engineering Review, 1996,6(11):2-3.
- 5 Femandez M, Gomez-Perez A. METHONTOLOGY: From Ontological Art Towards Ontological Engineering, AAAI-97 Spring Symposium on Ontological Engineering, Stanford University, 1997.3.
- 6 袁小娟.动画素材的领域本体模型与语义推理研究[硕士学位论文].长沙:湖南师范大学,2009.
- 7 Gruber TR. Towards principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies, 1995,43:907-92.