

本体在概念建模中的应用研究

杨 斌¹, 齐玉东¹, 孟凡磊², 王亚宁³

- (1. 海军航空工程学院 兵器科学与技术系, 山东 烟台 264001;
2. 海军航空工程学院 飞行器工程系, 山东 烟台 264001;
3. 海军航空工程学院 研究生管理大队, 山东 烟台 264001)

摘 要:概念建模是指创建概念模型的行为, 这些模型用来描述问题, 与具体用于解决问题的技术和策略独立无关。在过去数十年中, 大量概念建模方法和工具纷纷涌现, 许多理论如本体论、语言学和认知学的引入, 增强了概念建模的理论基础, 基于本体的概念建模研究得到了充分重视。探讨了基于本体的概念建模理论、概念建模语言中本体的作用、上层本体与领域本体的关系; 对基于本体的概念模型以及本体语言及其逻辑基础在概念模型的一致性检测的应用进行了研究。

关键词:概念建模; 概念建模语言; 本体; 描述逻辑; 一致性检测

中图分类号: N945. 12

文献标识码: A

文章编号: 1673-629X(2011)05-0246-04

Application Research on Ontology in Conceptual Modeling

YANG Bin¹, QI Yu-dong¹, MENG Fan-lei², WANG Ya-ning³

- (1. Department of Ordnance Science and Technology, Naval Aeronautical and Astronautical University, Yantai 264001, China;
2. Department of Architectural Engineering, Naval Aeronautical and Astronautical University, Yantai 264001, China;
3. Graduate Students' Brigade, Naval Aeronautical and Astronautical University, Yantai 264001, China)

Abstract: Conceptual modeling is the activity of creating models. Models that describe problems are independently from the technology and strategy used to solve the problem. In the past few decades, a large number of conceptual modeling methods and tools have emerged. Many theories such as ontology linguistics and cognitive science enhance the concept modeling theoretical basis. Ontology based conceptual modeling got full attention. In the article, ontology based theory of conceptual modeling, the role of ontology for conceptual model language and relationship between upper and domain ontology have been discussed. Ontology based conceptual modeling and consistency checking of the conceptual model have been also studied.

Key words: conceptual modeling; CML; ontology; description logics; consistency checking

0 引 言

概念建模 (Conceptual Modeling, CM) 最早如 Mylopoulos^[1]所言: “是一项为理解和沟通的目的而形式化地描述我们周围的客观世界和社会的某些方面的活动。” Kilov 与 Ross^[2]将概念建模定义为 “一种创建可理解的、雅致的关于企业业务规则的说明的过程。” 其在信息系统设计, 人工智能中的知识表达、组织环境建模、业务处理、软件开发过程和软件需求等计算机科学的不同领域中扮演着重要的角色。本体 (Ontology) 是一个来源于哲学的概念, 它是研究存在的本质的哲学

问题。Gruber 给出的定义得到了许多同行的认可, 即本体论是对概念化的精确描述, 本体论用于描述事物的本质, 首次在计算机相关学科中被提及后, 被广泛用于信息系统研究、领域工程、人工智能和知识表达等诸多领域^[3,4]。

将本体引入概念建模, 对于概念模型的语义及模型的形式化具有非常好的效果。本体作为共享概念模型的明确的形式化规范, 其本质是概念模型, 表达的是概念及概念之间的关系, 本体理论可以有效地解决传统建模理论所存在的低复用、低共享、语义表达能力差的缺点。基于本体的概念建模可以实现领域知识的重用、共享和使用。因此在进行概念建模前, 需要进行本体建模, 且从领域工程和本体工程的角度来看, 本体建模其实是一种特殊类型的概念建模过程, 因此可以将本体建模与概念建模过程统一起来。基于本体的概念建模是指用本体理论或模型支持概念建模的方法, 最

收稿日期: 2010-09-14; 修回日期: 2010-12-17

基金项目: 海装军械部科研项目 (415147W8)

作者简介: 杨 斌 (1980-), 男, 山东烟台人, 硕士研究生, 讲师, 研究方向为知识表示与推理、概念建模、描述逻辑; 齐玉东, 博士研究生, 副教授, 研究方向为语义网、知识管理、软件工程、描述逻辑。

早由 Wand 和 Weber 两位学者提出。他们基于 Bunge 的本体开发出了适用于信息系统的本体,称之为 BWV 本体。目前,BWV 本体已成为最突出和最先进的建模理论,此外,Chisholm 本体,Guizzardi 本体的应用也十分广泛。

1 本体与概念建模

概念建模是在系统实施前,能够对现实世界进行抽象,帮助领域人员、系统分析者、设计者及用户清楚地表达抽象的概念,进行相互交流。由于概念建模的对象、目标和阶段有所差异,结构化建模理论、原型法建模理论、面向对象建模理论、UML 建模理论、IBM 可视化建模理论、基于体系结构的建模理论、基于本体的建模理论等是概念建模理论发展至今,出现的许多具有代表性的一些理论。Ontology 这个哲学范畴,引入信息科学中被人工智能界赋予了新的定义。然而信息科学界对 Ontology 的理解也是逐步发展才走向成熟的。1991 年 Neches 等人最早给出 Ontology 在信息科学中的定义:“给出构成相关领域词汇的基本术语和关系,以及利用这些术语和关系构成的规定这些词汇外延规则的定义。”后来在信息系统、知识系统等领域,随着越来越多的人研究 Ontology,产生了不同的定义。1993 年 Gruber 定义 Ontology 为“概念模型的明确的规范说明”。1997 年 Borst 进一步完善为“共享概念模型的形式化规范说明”。Studer 等人对上述两个定义进行了深入研究,认为 Ontology 是共享概念模型的明确的形式化规范说明,这也是目前对 Ontology 概念的统一看法。尽管定义有很多不同的方式,但是从内涵上来看,不同研究者对于 Ontology 的认识是统一的,都把它当作是领域(领域的范围可以是特定应用中,也可以是更广的范围)内部不同主体(人、机器、软件系统等)之间进行交流(对话、互操作、共享等)的一种语义基础,即由 Ontology 提供一种共识。

本体对于一个领域内的概念模型来说十分重要。一方面,保证了模型元素之间的概念结构一致性和概念层次的清晰性,使得概念模型可以正确地映射为设计模型,再到最后的具体实现模型;另一方面,保证了整个领域内的概念模型能够正确地反映现实世界中的事物及其联系,使得领域专家、建模人员和各种模型使用人员能够就一些领域知识达成一致理解,具有清晰而完整的语义。

领域概念种类繁多且其间关系也极为复杂,基于本体的概念建模可以实现领域知识的重用、共享和使用。因此在进行概念建模前,需要进行本体建模,且从领域工程和本体工程的角度来看,本体建模其实是一种特殊类型的概念建模过程,因此可以将本体建模与

概念建模过程统一起来。领域概念种类繁多且其间关系也极为复杂,基于本体建模的本质就是建立“元”模型,对于一个领域内的概念模型来说十分重要。一方面,保证了模型元素之间的概念结构一致性和概念层次的清晰性,使得概念模型可以正确地映射为设计模型,再到最后的具体实现模型;另一方面,保证了整个领域内的概念模型能够正确地反映现实世界中的事物及其联系,使得领域专家、建模人员和各种模型使用人员能够就一些领域知识达成一致理解,并具有清晰而完整的语义。

针对不同的应用目的,本体的定义也不尽相同,所使用的元本体也来自于不同的哲学定义。在概念建模领域,使用较为广泛的本体有 BWV 本体模型^[5-8], Chisholm 本体^[9], Guizzardi 本体^[4]。参照文献[10]对本体的分类,可以将本体模型与领域概念模型的关系表示为如图 1 所示的结构。

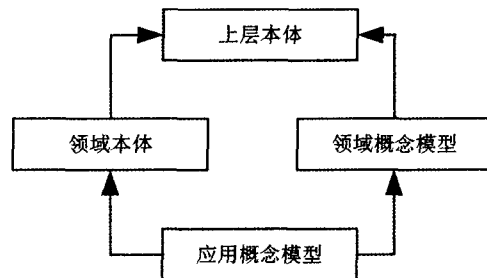


图1 本体模型与领域概念模型的关系

其中:上层本体定义一般化的概念及分类体系,对领域本体提供本体原语,领域本体定义现实世界的特定部分,对领域概念模型和应用概念模型提供明确的概念支持。领域概念模型表达领域内的任务、事件、条件、动作和交互等要素;应用概念模型是针对某一具体领域应用过程进行建模的成果。

目前作为上层本体的成果有 BWV^[5-8]、SUMO^[11]、DOLCE^[12]等可供选择来为领域本体建模和概念建模提供本体原语支持。

2 本体论与概念建模语言

概念建模语言(conceptual modeling language, CML)是关于概念模型的语言表示,是对概念模型的描述,也是概念建模的核心。其方法是把与概念模型相关的数据、信息抽取出来之后,经过整理,最后形成知识。CML 的组成包括:基础建模结构体和一些规则,将这些结构体有目的地组织在一起形成对被建模领域的陈述。CML 结构体的语义确定了现象的本质,并可以描述它。

对于概念建模来说,一个至关重要的问题是如何定义一种合适的概念建模语言。一种建模语言的能力在于它的结构体的语义。这种语义定义了单个的结构

体和它们的组合对被建模领域的支持。因此建模语言结构体的语义相当重要^[13]。

文献[5]对概念建模的基础与人类知识模型相关性进行了研究,提出了三种模型:本体论,概念理论和语言学。并揭示了这三个模型如何被一起使用来构成概念建模的基础。本体论是哲学的分支,用来解决存在的问题,因此本体能够被用来定义概念,且用一种建模语言表达,即这种语言的语义。概念理论处理人类以分类或概念化的形式组织知识,用来指引如何在概念模型中挑选富有意味的结构体。语言学处理人类如何为了沟通的目的而表达知识和如何进行沟通。特别关注语言学中的一个特定领域—言论动作理论。言论动作理论根据说话者的意图和对接受者可能产生的影响,研究和分析说话的方式。言论动作理论在概念建模中,可被用于捕获通信代理(包括人和机器)之间交互的动态细节。这三个模型一起使用来构成概念建模的基础,每个模型覆盖了其它模型没有涉及的方面。

文献[14]针对概念模型的结构形式介绍了一种面向对象的本体语言。其方法就是用本体对横向知识表示进行加强,并用面向对象的方法描述本体结构。把面向对象的本体语言分为两大部分:本体部分与对象部分。其中,前者作为整个语言的核心部分,采用的是框架结构进行描述,产生式规则隶属框架,作为框架行为处理的描述;后者采用面向对象的方法将框架用对象来处理,最后成为整个概念模型的模型表现结构。这种方法克服了不便于表示有横向联系的知识这一局限性,并充分利用了面向对象方法的适合于表示模块化和有纵向继承关系的知识的优点,最终达到了知识的重用、共享与操作。

3 本体在概念模型正确性检测的应用研究

概念模型的校核中,形式化方法是基于形式数学证明数学上的正确性,是最精确的模型 V&V 方法,但使用形式化技术具有很高的要求。用数学证明正确性尽管在原理上可能,但在实践中往往不都是可行的。另外,形式化方法本身能够描述的内容也是有限的。采用任何一种形式化技术进行 V&V 都要求建模过程拥有良好的结构和定义,使用形式化的语法和语义,完全应用形式化技术来完成 V&V 工作基本上是很难的^[14]。因而当概念模型用语义、语法都很清晰的表示法来描述时,一般只对其内部一致性进行逻辑检测。

将概念模型转换为描述逻辑^[15](Description Logics, DLs)或本体表示语言^[16]进行推理是目前概念模型一致性检测研究的热点。描述逻辑是基于对象的形式化知识表示工具,是一阶谓词逻辑的一个可判定子集,具有很强的表达能力和可判定性,具有清晰的模型

-理论语义,对概念性知识的处理,特别是对概念分层的处理非常有效,非常适合知识表示和自动推理。而描述逻辑则是本体表示语言的逻辑基础,其基本的推理问题包括以下几类:包含检测、一致性检测、等价性检测、实例检测。一般而言,所有的推理问题都可以归结为一致性检测(或称为可满足性检测)^[17]。

文献[18]为解决 UML 类图一致性检测问题,给出了 UML 类图转化为 OWL DL 本体语言的方法,并利用描述逻辑的推理功能实现 UML 类图的一致性检测。文献[19]提出了一种基于描述逻辑的具有自动推理功能的扩展概念图,针对概念图的特点和需求,给出了将概念图的一个子集转化为描述逻辑知识库的方法,并证明了该方法的正确性。同时给出了知识库的一致性、包含关系的检测方法也证明了检测方法的正确性。文献[20]提出了一种从 ER 模式到 OWL DL 本体的语义保持翻译方法,该方法在形式化表示 ER 模式的基础上,建立 ER 模式和 OWL DL 本体之间精确的概念对应,通过一个翻译算法按照一组预定义的映射规则实现模式翻译。基于此基础,可以实现对 ER 模型的正确性检测研究。文献[21]阐述了如何将 ORM(Object Role Modeling)模型用描述逻辑 DLR 形式化表示,并实现 ORM 模型的推理检测。

上述文献中采用的各种研究方法主要针对如何将模型映射为描述逻辑或本体语言,然后使用 RACER、FACT 或 LOOM 等作为推理引擎实现模型的正确性检测。目前研究来看,采用任何一种形式化技术进行 V&V(Verification&Validation)都要求建模过程具有良好的定义和结构,使用形式化的语法和语义^[14]。而完全应用形式化技术来完成模型的校验是很难达到的。当用清晰的语义、语法表示方法来描述概念模型时,主要对其内部的一致性进行逻辑分析和检测。

4 结束语

基于本体的概念建模可以明确概念的定义和概念之间的关系。引入上层本体对概念模型中的概念及其关系进行分类和约束,保证模型在逻辑和本体上的正确性,更加准确地反映领域的事实,提高模型的质量。在概念模型正确性检测中应用 OWL DL 本体语言进行推理实现模型一致性检测。概念建模的实践中充满了本体论的问题,概念建模如果要成为一个成熟的学科,就必须发展它的哲学基础。

参考文献:

- [1] Mylopoulos J. Conceptual modeling and telos[M] // Conceptual modeling, databases, and case: an integrated view of information systems development. New York: Wiley, 1992:

- 49-68.
- [2] Kilov H, Ross J. *Information modeling: an object - oriented approach*[M]. Englewood Cliffs, N. J.: Prentice-Hall, 1994.
- [3] Mealy G H. Another Look at Data[C]//Proceedings of the Fall Joint Computer Conference. Washington, DC: Thompson Books; London: Academic Press, 1967:525-534.
- [4] Guizzardi G. *Ontological Foundations for Structural Conceptual Models*[M]. The Netherlands: Universal Press, 2005.
- [5] Wand Y, Weber R. On the Ontological Expressiveness of Information Systems Analysis and Design Grammars[J]. *Journal of Information Systems*, 1993, 3:217-237.
- [6] Wand Y, Weber R. Mario Bunge's ontology as a formal foundation for information systems concepts[C]//Studies on Mario Bunge's Treatise. Amsterdam: Rodopi, 1990: 123 - 149.
- [7] Wand Y, Weber R. On the deep structure of information systems[J]. *Information Systems Journal*, 1995, 5(3): 203 - 223.
- [8] Weber R. *Ontological Foundations of Information Systems* [M]. Melbourne: Coopers & Lybrand, 1997.
- [9] Milton S K, Kazmierczak E, Keen C. Data modeling languages: An ontological study[C]//Proceedings of the 9th European Conference on Information Systems. [s. l.]: [s. n.], 2001:304-315.
- [10] Guarino N. *Formal Ontology and Information Systems*[C]//Proceedings of the International Conference on Formal Ontology and Information Systems (FOIS). Trento, Italy: IOS Press, 1998:3-15.
- [11] Niles I, Pease A. Towards a Standard Upper Ontology[C]//Proceedings of FOIS. Ogunquit, Maine, USA: [s. n.], 2001:17-19.
- [12] Colomb R M. *Ontology and the Semantic Web*[M]. Netherlands: 150 Press, 2007:87-90.
- [13] Carasik R P, Johnson S M, Patterson D A, et al. Towards a Domain Description Grammar: An Application of Linguistic Semantics [C]//Proceedings of the Fourth International Workshop on Computer-Aided Software Engineering. Irvine, CA: [s. n.], 1990:410-419.
- [14] 王杏林,曹晓东. 概念建模[M]. 北京:国防工业出版社, 2007.
- [15] 石 莲,孙吉贵. 描述逻辑综述[J]. *计算机科学*, 2006, 33(1):194-225.
- [16] 邓志鸿,唐世渭,张 铭,等. Ontology 研究综述[J]. *北京大学学报*, 2002, 38(5):730-738.
- [17] Doninifm, Lenzerinim, Nardid, et al. Reasoning in description logics[C]//Studies in Logic, Language and Information. [s. l.]: CLSI Publications, 1996:193 - 238.
- [18] Cali A, Calvanese D, De Giacomo G, et al. Reasoning on UML class diagrams in description logics[C]//Proc. of IJ-CAR Workshop on Precise Modeling and Deduction for Object-oriented Software Development (PMD 2001). [s. l.]: [s. n.], 2001.
- [19] 王 军,王继军,甘 丹,等. 基于描述逻辑的概念图推理[J]. *计算机科学*, 2008, 35(8):176-179.
- [20] 许卓明,董逸生,陆 阳. 从 ER 模式到 OWL DL 本体的语义保持的翻译[J]. *计算机学报*, 2006, 29(10):1786-1796.
- [21] Jarrar M. Towards Automated Reasoning on ORM Schemes-Mapping ORM into the DLRidf Description Logic[C]//Proc. of the 26th Int. Conf. on Conceptual Modeling (ER2007). [s. l.]: Springer, 2007:181-197.

(上接第 123 页)

- [2] 李晓明,刘建国. 搜索引擎技术及趋势[EB/OL]. 2003-04. <http://www.Se-express.com/se/se07.html>.
- [3] 蔡建超,郭一平,王 亮. 基于 LUCENE. Net 校园网搜索引擎的设计与实现[J]. *计算机技术与发展*, 2006, 16(11): 74-80.
- [4] 李 尉,霍 涛,惠勇侠. WWW 的信息检索技术研究[J]. *济南大学学报(自然科学版)*, 2001, 15(3):284-286.
- [5] 赵 汀,孟祥武. 基于 LUCENE API 的中文全文数据库的设计与实现[J]. *计算机工程与应用*, 2003(20):179-183.
- [6] 张校乾,金玉玲,侯玉波. 一种基于 lucene 检索系统的一种全文数据库的设计与实现[J]. *现代图书情报技术*, 2005(2):40-44.
- [7] 彭洪汇,林作铨. Internet 上的搜索引擎和元搜索引擎[J]. *计算机科学*, 2002, 29(9):1-12.
- [8] Cospodnetic O, Hatcher E. *Lucene in Action*[M]. [s. l.]: Manning Publications Co, 2005:15-18.
- [9] Owens S J. *Lucene Tutorial* [EB/OL]. 2009-11. <http://dark.Sleep/m/LUCENE/>.
- [10] 曹元大,贺海军,涂哲明. 中文 web 文档全文检索系统的设计与实现[J]. *北京理工大学学报*, 2002, 22(1):68-71.
- [11] 颜维龙,盖 杰,武港山,等. 面向网络的全文检索中索引文件的组织[J]. *计算机应用研究*, 2002, 11(2):124-126.
- [12] 李 刚,宋 伟,邱 哲. 征服 Ajax 和 Lucene 构建搜索引擎[M]. 北京:人民邮电出版社, 2006.
- [13] Du Lin, Zhang Yibo, Sun Yufang. The design and implementation of Web-based Chinese text retrieval system SEARCH 2000 [J]. *Journal of Chinese Information Processing*, 2000, 14(6):14-20.
- [14] Rogers W, Candela G, Harman D. Space and time improvements for indexing in information retrieval[C]//The Fourth Annual Symposium on Document Analysis and Information Retrieval (SDAIR'93). Las Vegas: [s. n.], 1995.
- [15] Sahon G, Wong A. On the specification of term value in automatic indexing [J]. *Journal of Documentation*, 1973, 29(4):351-372.