

基于 SWRL 元模型的本体公理建模方法

王小龙, 李 明

(兰州理工大学计算机与通信学院, 兰州 730050)

摘 要: 针对传统 SWRL 元模型难以表达规则前件中 OWL 类实例的问题, 提出改进的元模型, 采用逻辑运算符连接并区分规则的 UML 模型中不同的 OWL 属性和 OWL 类的实例。为规则设计了基于 UML profile 的建模机制。通过引入 horn 短句集、描述逻辑等知识表示方式, 论证了用描述逻辑对部分基于 UML 建模的规则形式化的可行性, 并给出基于描述逻辑对其形式化的方法。

关键词: 本体; 统一建模语言; 元模型; 规则

Modeling Method for Ontology Axioms Based on SWRL Metamodel

WANG Xiao-long, LI Ming

(College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050)

【Abstract】 The traditional UML metamodel of SWRL is difficult to express individuals of OWL class in antecedents of rules. To solve this problem, this paper represents an improved metamodel, using boolean operator to connect and to distinguish OWL properties and individuals of different OWL classes in a UML model of rules. A method to model rules with UML profile is represented. For a part of rules modeled by UML, the feasibility of formalizing them into description logics is proved, and a technique to realize that is represented.

【Key words】 ontology; Unified Modeling Language(UML); metamodel; rules

1 概述

本体已在知识建模、语义 Web、基于 Agent 的系统设计等领域发挥了重要作用。为了将本体用于工业化的软件设计开发当中, 统一建模语言(Unified Modeling Language, UML)被尝试用于本体建模, 从而扩展本体的应用领域, 提高本体的建模效率^[1]。

公理是本体的基本组成部分, 本体的推理机制基于该本体的公理体系。SWRL(Semantic Web Rule Language)是通过结合 OWL 与 RuleML 形成的语义 Web 的规则描述语言^[2]。文献[3]为 SWRL 提出了基于 UML 的元模型, 用以对本体公理集中的规则进行图形化建模, 但基于它的建模方法难以描述规则前件中 OWL 类的实例。针对文献[3]中建模方法的缺点, 本文提出了改进的元模型, 增加逻辑运算符的建模元素来连接不同的原子, 进而设计了规则的描述方法。

2 语义网规则语言(SWRL)

SWRL 是 W3C 于 2004 年提出的一种用于描述规则的语言。SWRL 是基于 OWL 的 2 种子语言(OWL Lite, OWL DL)和 RULE ML 的结合体。SWRL 扩展了 OWL 的公理集, 加入了类似 horn 短句的规则, 增强了对本体公理集中一般的规则的描述能力。SWRL 描述规则的语法为: antecedent consequent, antecedent 和 consequent 写成一组原子的合取。变量用标准前缀问号来标示, 例如 ?x。原子写成 $C(x)$ 或 $P(y,z)$ 的形式。 C 可以是 OWL 类(class), P 是 OWL 属性(property)。另一类原子是具有函数作用的关系(Built-In)。根据语法规范, 表达双亲关系合取兄弟关系蕴含叔侄关系的规则表达为

$$parent(?x, ?y) \wedge brother(?y, ?z) \Rightarrow uncle(?x, ?z)$$

$I = \langle R, EC, ER, L, S, LV \rangle$ 是 SWRL 的解释规则; R 为资源集合; $LV \subseteq R$, 是文字的集合; EC 是从 OWL 的类或数据类

型到 R 的子集或 LV 的映射; ER 是从 OWL 的属性到 R 上二元关系集合的映射; L 是给定文字到 LV 上元素的映射; S 是从 OWL 的个体名到 $EC(owl:Thing)$ 的元素的映射。

3 基于 UML 的规则建模

3.1 SWRL 元模型的构建

文献[3]提出了 SWRL 基于 UML 的元模型; 采用 UML profile 对 SWRL 的规则进行图形化建模。对于规则的前件, 文献[3]中的方案强调 OWL 属性的描述, 出现在规则中的原子以 OWL 属性为主。如果某一 OWL 类作为原子独立出现在规则中, 用该方法难以描述。这限制了基于 UML 建模规则的表达能力。本文在元模型中引入布尔运算符, 泛化出与运算。这使得用 UML 建模规则的前件中不仅包含 OWL 属性, 也可以出现 OWL 类。两者可以被清楚地区分开来。这种描述机制更符合 SWRL 的语法定义。

OWL 的基本建模元素包括 class, property, individual。SWRL 在 OWL 的基础上增加了对规则的描述。所以, 基于 UML 的 SWRL 的元模型包括 4 种基本概念: 分类器(classifier), 属性(property), 实例(instance), 规则(rule)。

分类符(classifier)源于 UML 中的同名概念。由它泛化出类(class)的概念。这个概念表示的是 OWL 的类, 描述了一些个体的共有特征。分类符的细化如图 1 所示。UML 中类的概念实质上是一些具有共有特征的个体的集合。两者的含义是

基金项目: 甘肃省自然科学基金资助项目(0803RJZA025); 甘肃省教育厅科研基金资助项目(0803-07)

作者简介: 王小龙(1981 -), 男, 硕士研究生, 主研方向: 智能信息处理; 李 明, 教授

收稿日期: 2009-08-15 **E-mail:** home@lut.cn

有区别的。经过元模型的定义，在语法上用 UML 的类图表示 OWL 的类。

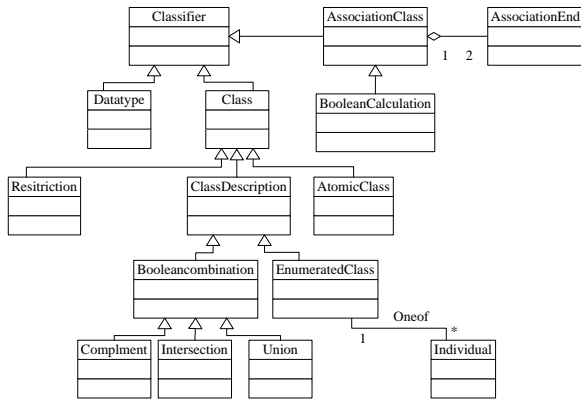


图 1 分类符的细化

为了用 UML 明确描述 SWRL 规则中的各类原子，引入二元逻辑运算符。在分类符中泛化出 BooleanCalculation，它的父类是关联类(AssociationClass)。根据 UML 对关联类的定义，它可以连接 2 个分类符，在语法上满足表达二元逻辑运算符的要求。

如图 2 所示，由 BooleanCalculation 泛化出与(and)运算符。按照 SWRL 的语法规则，与运算有 2 个原子作为操作数。这些原子可以是属性、概念、数据类型或 SWRL 提供的关系 (Built-In)。

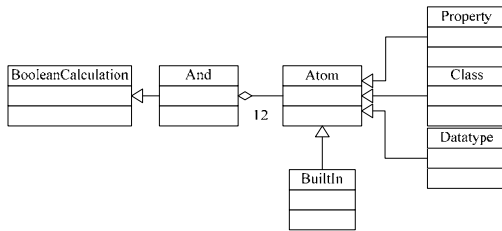


图 2 二元逻辑运算符的细化

图 3 所示是改进后的 SWRL 公理集中规则的元模型。为了形成基于一阶逻辑的规则，由逻辑操作符蕴含连接规则的前件与后件。SWRL 的规则由一阶逻辑的与运算连接前件和后件中的原子。与运算具有 2 个操作数。这 2 个操作数可以是基本的数据类型、概念、属性(关系)。

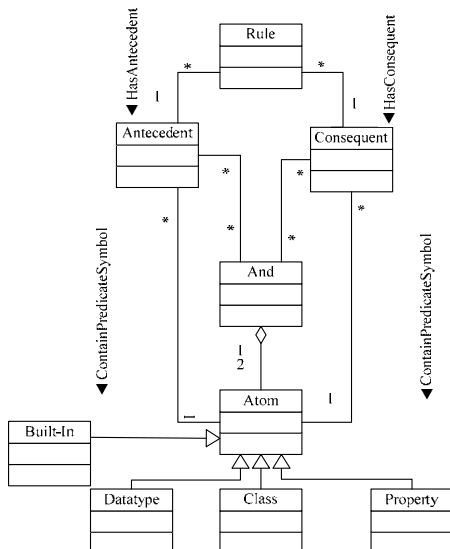


图 3 规则的细化

3.2 基于 UML 的规则描述方法

依据元模型对建模元素的定义，利用 UML profile 对规则进行建模。建立对象造型 variable 定义规则中的变量。对于原子中的 OWL 类，在对象图中标明类名。如果某个变量只出现在 OWL 属性中，则在对象图中不指明该变量所属的类。否则，在相应的对象图中给出完整的描述。对于明确定义的 OWL 个体或数据值，用一般的对象图就可以描述。作为原子出现在规则当中的 OWL 属性主要是对象属性或数据类型属性，由 UML 的关联关系连接表示定义域的和值域的类型从而描述一个属性。关联关系连接 2 个 variable 表达原子中的属性或 Built-In 关系。按照元模型的定义，不同原子之间用以 and 命名的关联类连接。关联类两侧可以任意出现类的命名变量、属性等原子。

4 规则的形式化语义

通过对基于 UML 建模的规则进行形式化的讨论，可以说明改进后的建模方法有更强的表达能力。同时，为利用本体检测实现 UML 检验提供了形式化的逻辑基础。

用 UML 建模的 SWRL 规则，可以直接用 SWRL 的语法规则形式化地表达。对于各种原子，通过解释规则 I 映射为一阶逻辑的一元或二元谓词。原子之间的与关系映射为一阶逻辑中的与运算。对于变量或 OWL 类的命名实例等，用一阶逻辑中的常量或变量来表达。设 s' 是 s 的基于一阶逻辑的形式化表达。在语法上 s' 与 horn 短句集是相似的。设 $Q = \{a_i; 1 \leq i \leq L\}$ 表示 s' 中的常量， $C = \{c(x_j); 0 \leq j \leq m\}$ 表示 s' 中的一元谓词， $P = \{p(x,y)_k; 0 \leq k \leq n\}$ 表示 s' 中的二元谓词。由于 s' 中不出现函数， s' 的 Herbrand 通域为 $H_0 = H_1 = H_2 = \dots = \{a_0\} \cup Q$ 。 s' 的 Herbrand 基为 $A = \{c(a), p(b,c) | a,b,c \in \{a_0\} \cup Q; c(a) \in C; p(b,c) \in P\}$ 。在此基础上，可获得 s' 的 Herbrand 解释。这样，可以用类似 horn 短句集的语法来形式化基于 UML 描述的规则。

本文的元建模方案扩展了基于 UML 建模的规则的表达能力。一元谓词可以明确出现在规则前件中。在此基础上，寻求使用不同于 horn 短句集但同样适合知识描述的逻辑语言对其进行形式化。

设 A 为描述逻辑中的原子概念(atomic concept)， R 为描述逻辑中的原子关系(atomic role)， I 为解释函数， Δ^I 为论域。 A^I 是 Δ^I 中的一元关系， R^I 是 Δ^I 中的二元关系。对任意 A ，有一元谓词 $A(x)$ ，使得 $A(x) \equiv A^I$ 。同理， R 对应于二元谓词 $R(x,y)$ ， $R(x,y) \equiv R^I$ 。设 $\phi \equiv \forall R.C(y)$ 是 $\forall R.C(y)$ 的基于一阶谓词逻辑的描述， C 为原子概念。 $(\forall R.C)^I = \{x \in \Delta^I | \forall y. \langle x,y \rangle \in R^I \rightarrow C^I\}$ ，所以 $\phi \equiv \forall R.C(y) \equiv \forall x. R(y,x) \rightarrow C(x)$ 。 $(\exists R.C)^I = \{x \in \Delta^I | \exists y. \langle x,y \rangle \in R^I \wedge C^I\}$ ，所以 $\phi \equiv \exists R.C(y) \equiv \exists x. R(y,x) \wedge C(x)$ 。设 C, D 是原子概念， $(C \cap D)^I = C^I \cap D^I$ ，所以 $\phi \equiv C \cap D \equiv C(x) \wedge D(x)$ ；同理， $\phi \equiv C \cup D \equiv C(x) \vee D(x)$ 。 $(C \subseteq D)^I = C^I \subseteq D^I$ ，所以 $\phi \equiv C \subseteq D \equiv D(x) \leftarrow C(x)$ 。

基于以上的翻译，设 $B(x), C(x), D(x)$ 为原子谓词公式，则有下式：

$$D(x) \leftarrow B(x) \wedge C(x) \equiv B \cap C \subseteq D \quad (1)$$

$$\begin{cases} B(x) \leftarrow D(x) \\ C(x) \leftarrow D(x) \end{cases} \Rightarrow B(x) \wedge C(x) \leftarrow D(x)$$

对于以上 horn 短句集，有下式：

$$B(x) \wedge C(x) \leftarrow D(x) \equiv D \subseteq B \cap C \quad (2)$$

$$\begin{cases} D(x) \leftarrow B(x) \\ D(x) \leftarrow C(x) \end{cases} \Rightarrow D(x) \leftarrow C(x) \vee B(x)$$

对于以上 horn 短句集，有下式：

$$D(x) \leftarrow C(x) \vee B(x) \equiv B(x) \cup C(x) \subseteq D(x) \quad (3)$$

设 $D(x), C(y), R(x,y)$ 为原子谓词公式，有：

$$C(y) \leftarrow R(x,y) \wedge D(x) \equiv (C(y) \leftarrow R(x,y)) \leftarrow D(x)$$

借助上面 $\phi \forall R.C$ 的结论得到：

$$C(y) \leftarrow R(x,y) \wedge D(x) \equiv D \subseteq \forall R.C \quad (4)$$

借助上面 $\phi \exists R.C$ 的结论得到：

$$D(x) \leftarrow R(x,y) \wedge C(y) \equiv \exists R.C \subseteq D \quad (5)$$

基于以上的映射规则，对于用 UML 建模的规则采用 AL-languages 中的 ALUE 作为形式化的描述手段。

(1)属于某个类构造型 variable 的变量名泛指任意满足该规则的对象，用原子概念形式化描述。

(2)关联表达的属性，用原子关系形式化描述。

(3)如果出现在关联当中的变量没有定义其所在类，在基于 DL 的形式化描述中，论域 Δ^I 要对这个变量所属的概念进行说明。

(4)某些规则会用到具体的对象或常量，在 DL 公理的基础上引入 ABox 对其形式化。ABox 的解释 I 的论域 Δ^I 必须包含 UML 图中出现的所有对象，解释函数将 DL 的对象映射到相应的 UML 对象。

例 1

$$\begin{cases} men(y) \leftarrow haschild(x,y) \wedge motherwithson(x) \\ mother(x) \leftarrow motherwithson(x) \end{cases} \Rightarrow$$

$$mother(x) \wedge (men(y) \leftarrow haschild(x,y)) \leftarrow motherwithson(x)$$

根据式(2)、式(4)及形式化描述规则，以上规则集可转化为：

$$motherwithson \subseteq mother \wedge \forall has Child.men$$

$$men^I = \{y | y \text{ 是 UML 定义的类 } men \text{ 的对象}\}$$

$$mother^I = \{x | x \text{ 是 UML 定义的类 } mother \text{ 的对象}\}$$

$$\Delta^I = men^I \cup mother^I$$

对例 1 的形式化描述如图 4 所示。

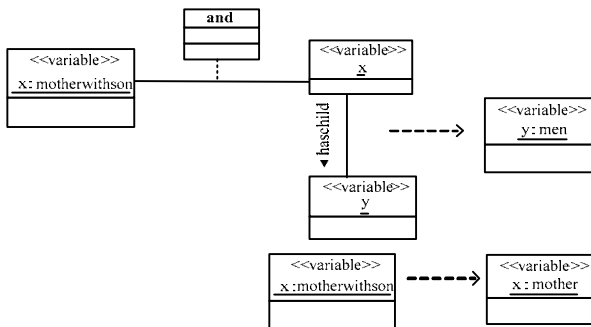


图 4 例 1 形式化描述

例 2

$$heart(z) \wedge haslocation(x,z) \leftarrow aorticvalve(y) \wedge haslocation(x,y)$$

以上规则可用描述逻辑形式化地表达为

$$\exists haslocation.aorticvalve \subseteq \exists haslocation.heart$$

$$aorticvalve^I = \{y | y \text{ 是类 } aorticvalve \text{ 的对象}\}$$

$$heart^I = \{z | z \text{ 是类 } heart \text{ 的对象}\}$$

$$\Delta^I = aorticvalve^I \cup heart^I \cup \{x | x,y \text{ 满足 } haslocation \text{ 关联 } \wedge y \in aorticvalve^I\}$$

对例 2 的形式化描述如图 5 所示。

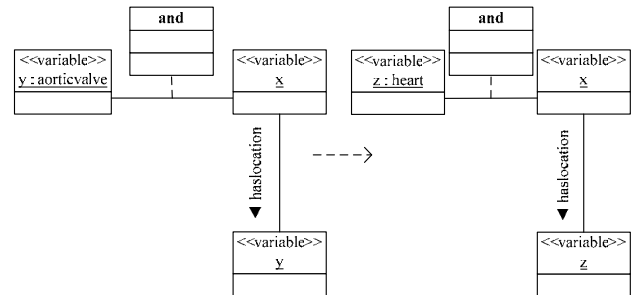


图 5 例 2 形式化描述

5 结束语

将 UML 用于本体建模的研究可以围绕特定的本体描述语言展开。公理集是基于本体的知识库的重要组成部分。实现本体公理集的图形化建模有助于扩展本体的应用范围，提高本体的建模效率。本文设计了改进的 SWRL 元模型，为规则设计了基于 UML 的建模方法，解决了文献[4]中建模方法难以描述规则前件中 OWL 类实例的问题。对于部分图形化建模的规则，论证了用描述逻辑对其形式化的可行性。通过将这部分基于 UML profile 建模的规则翻译为描述逻辑，说明本文的建模方法有更强的表达能力。下一步的工作包括如何针对图形化建模的规则进行推理，以便利用本体检测实现 UML 检验，给出了基于描述逻辑对其形式化的方法，为这项研究工作提供了形式化的逻辑基础。

参考文献

- [1] 钟凌燕, 陈 岗. OWL DL 的 UML 建模方法[J]. 计算机工程, 2006, 32(21): 58-60.
- [2] Horrocks I, Boley H. SWRL: A Semantic Web Rule Language Combining OWL and RuleML[EB/OL]. (2004-05-20). <http://www.w3c.org/Submission/2004/SUBM-SWRL-20040521/>.
- [3] Brockmans S, Haase P. A Metamodel and UML Profile for Rule-extended OWL DL Ontologies[C]//Proc. of the 3rd Annual European Semantic Web Conference. [S. l.]: Springer-Verlag, 2006: 303-316.
- [4] 苏依拉, 郑 磊, 刘椿年. 基于 OWL 和 Horn 子句的 β -PSML 语言框架[J]. 计算机工程, 2005, 31(23): 65-66, 78.

编辑 顾逸斐

(上接第 52 页)

参考文献

- [1] 常 伟. SQL 中查询语句的优化方法[J]. 重庆工学院学报, 2006, 20(5): 85-88.
- [2] 王振辉, 吴广茂. 查询语句优化研究[J]. 计算机应用, 2005, 25(12): 32-34.
- [3] 金洪伟, 苏厚勤. 查询表达式优化的研究与实现[J]. 计算机应用

与软件, 2008, 25(6): 123-125.

- [4] Pentaris F, Ioannidis P. Query Optimization in Distribute Networks of Autonomous Database Systems[J]. ACM Transactions on Database Systems, 2006, 31(2): 538-538.
- [5] Subramanian D K. Query Optimization in Multi_database Systems[J]. Distribute and Parallel Databases, 1998, 6(2): 183-210.

编辑 顾逸斐