

面向语义 Web 的逻辑描述原语扩展

姚绍文^{1,2}, 余 江¹, 周明天²

(1. 云南大学信息学院, 云南昆明 650091; 2. 电子科技大学计算机学院, 四川成都 610054)

摘 要: 语义 Web 旨在利用知识工程中成熟的技术为数据和知识交换提供语义上的互操作性, 为了针对 Web 这种标记语言环境描述一阶命题逻辑公式、Horn 子句和推理规则, 本文对 OIL/RDFS 进行了扩展, 定义了相关的逻辑描述原语. 通过典型的命题规则表示示例, 说明了本文的原语扩展为命题公式、推理规则表示提供了可行的方法.

关键词: 语义 Web; 知识工程; 本体推理层; 资源描述框架 (模式); Horn 子句

中图分类号: TN304 **文献标识码:** A **文章编号:** 0372-2112 (2002) 12A-2115-04

Semantic Web-Oriented Specification of Logic Descriptive Primitives

YAO Shao-wen^{1,2}, YU Jiang¹, ZHOU Ming-tian¹

(1. School of Information and Electronic Science, Yunnan University, Kunming, Yunnan 650091, China;

2. School of Computer Sci. and Eng., University of Elec. Sci. and Tech., Chengdu, Sichuan 610054, China)

Abstract: By employing the flow-blown techniques of KE, semantic Web is targeted to provide the semantic interoperability for data and knowledge exchange. As Web is characterized by markup language, this paper extends OIL/RDFS descriptive primitives to broadly represent first-order propositional formulae, Horn Clauses and inference rules. The representation example of typical propositional rule shows the extended specification is feasible.

Key words: semantic Web; knowledge engineering (KE); OIL; RDF(S); horn clauses

1 引言

类似于 DBMS 和 KBS, Web 为了支持描述、搜索、管理、交换和更新等准确的处理, 就需要以某种方式进行元信息结构化甚至语义化处理. 传统的网络、计算机技术和 KE 定义了多种描述元数据的标准, 而 Web 计算的特征就是对资源或对象进行标记的过程, 因此传统的标准难以直接应用于 Web 计算模式中. 但 Web 的元信息处理和语义化发展中, DBMS 和 AI 成熟的研究成果, 特别是 KE 成熟的理论、方法和技术, 将起到关键的作用.

元语言 XML 提供了定义和表示标记语言的能力, 促进了基于 Web 的各种新型的软件体系结构、数据交换、资源管理等应用开发. 采用 XML 作为表示语法, W3C 开发了在 Web 上进行元数据处理和交换的标准 RDF(S)^[2]. RDF(S) 的出现使得 KE 成熟的技术可以应用在 Web 计算中, 从而使得 Web 转变为机器可理解的知识库成为可能. 基于 XML 和 RDF 的下一代 Web 被 Tim Berners Lee 称之为语义 Web^[5].

为了在 Web 中进行本体建模, 欧洲 OntoKnowledge 联盟基于 RDFS 定义了本体建模语言 OIL^[1]. OIL 提供了 NOT、AND、OR 三种基本的逻辑描述原语和本体建模方法, 可以建立应用领域的本体模型、逻辑公式和规则. 在以标记语言为中心的

Web 计算环境中, 为了更广泛地描述逻辑公式、规则甚至公理, 需要扩展定义 OIL/RDFS 的描述原语. 本文面向命题逻辑公式和推理规则的本体模型, 在介绍相关技术标准基础上, 形式化地扩展了 OIL/RDFS 原语的本体包容关系, 定义了描述 Horn 子句的逻辑原语. 本文的扩展定义在描述逻辑公式、规则方面具有广泛的适应性.

2 本体建模技术与语义 Web 技术基础

2.1 KE 与本体建模

虽然 AI 在形式化方法、推理机制以及相关的工具开发上取得了一些具有相当前景的成果, 但 AI 技术应用在构建商业 KBS 并不成功^[4]. 这促使了 KE 学科的产生, KE 主要研究知识提取和建模技术、表示知识的形式化理论、以及实现自动推理机制的工具箱. 类似于软件工程的目标, KE 旨在提供一些方法和工具, 以一种系统的、可控制的方式构建各种 KBS 系统. 近十五年来, KE 的研究开发了一些关于建模方法和建模框架的基本概念, 取得了一些关于知识模型的构造、知识模型体系的构建和重用方面的研究成果. 这些研究成果的最新发展主要集中在本体建模和问题求解方法 (PSM) 两个主要方面.

本体论是哲学中形而上学中研究事物存在性质的一个分支, 近年来被 AI 采用, 系指一系列的概念和术语抽象^[3]. 由于

跨越不同的系统进行知识重用或共享越来越重要,本体论越来越受到关注.因为不同的系统使用不同的概念和术语来描述不同的领域,而这些不同使得从一个体系获得的知识很难应用到另一个体系,所以这种不同就成为知识共享的主要障碍.如果能够开发出本体模型,使它成为一种多种体系的基础,通过本体来共享共同的术语,从而方便地实现共享和重用,这样将大大提高知识重用的潜力.

PSM 目的是通过定制规则和规则的前提条件,对推理过程控制知识进行编码,使控制知识显式化;PSM 对通用推理引擎进行细化,把 KBS 的推理任务分解为许多子任务和推理行为,通过知识任务联系在一起,实现对推理过程更直接的控制;PSM 描述了独立于应用领域的控制知识,使得不同的领域和应用可以重用这种策略知识;最后,PSM 抽象表示形式与依赖于特定知识表示的通用推理引擎是分开的.这样,不仅对推理过程进行更为直接的控制,而且在构建新应用时,可以重用 PSM 方法库.因此,PSM 被视为 KBS 的重要知识部分,是用于描述 KBS 推理部分的特殊软件体系类型.

2.2 RDF(S)和语义 Web 的结构

RDF 定义了资源、属性、事实三种基本实体.任何可以通过 URI 表示的对象都是 RDF 资源.资源之间可相互联系或通过属性-值对联系,这种三元组模型表示了一个事实.此外事实本身可被看作一个资源,即可以建立具体化的 RDF 数据模型.

RDFS 定义了一种模式定义语言,提供了一个定义在 RDF 之上抽象的词汇集,如: `rdfs:Resource`, `rdfs:Class` 和 `rdfs:Property` 类, `rdfs:subClassof` 和 `rdfs:subPropertyof` 属性关系,以及 `rdfs:domain` 和 `rdfs:range` 的属性约束关系.该词汇集构成了一种基本类型系统及该系统的基础类,这些基础类包括定义新类必需的元对象类型,因此 RDFS 具有扩展特性,适合于显式地定义对象语义进行 Web 上资源描述的本体建模和对对象模型建模. RDFS 为这种建模做出了两个重要贡献:编写 ontology 的标准化的语法、一组标准的类似于“instance-of”和“subclass-of”关系的建模原语.

语义互操作性不仅需要针对文档语法形式的标准,而且还需要针对语义内容的标准.作为交换数据和知识的平台,语义 Web 是一种体系结构,如图 1 所示.其中,Unicode 和 URI 是整个语义 Web 的语法表示基础,

Unicode 提供了资源编码,而 URI 用于标识资源.第二层 XML + NS + XMLS 主要提供了表示数据内容和结构的语法.第三到第五层是语义 Web 提供语义交互支持的核心, RDF/RDFS 主要针对语义上的互操作性提供了描述 Web 资源的数据模型及类型原语, Ontology 层是在 RDFS 基础上定义的概念和关系抽象描述,用于描述应用领域的知识,而第五层主要定义逻辑描述原语,为更高层的逻辑推理提供语义上的支持.第六和第七层主要是在下面各层提供

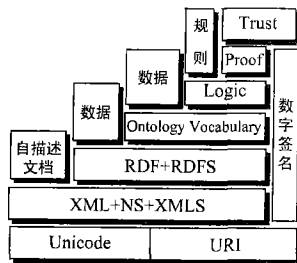


图 1 语义 Web 的体系结构

的功能基础上进行逻辑推理、证明等操作.

2.3 面向 Web 的本体描述语言 OIL

OIL 主要综合了三个方面的研究工作:描述逻辑(DL)、基于框架的系统(FL)、基于 XML 与 RDF 语法的 Web 语言标准.这样,与 RDF 兼容的 OIL 可以结合 FL 的模型原语和 DL 提供的推理服务,并精确地描述术语的语义.运用 OIL 定义面向应用的本体时,定义的模型可以用 XML 的 DTD、XML 模式或 RDF 数据模型进行描述,在 Web 环境中实现透明的交换与共享.

OIL 对 RDFS 建模原语的扩展主要集中在类扩展定义、槽定义两个方面,类之间的约束关系通过槽约束表示,槽的定义通过 RDFS 的属性扩展定义实现. OIL 通过 RDF 数据模型扩展定义类,类定义把类同名字、文档、类型、类的超类、为该类定义的属性联系在一起.图 2 示出了 OIL 对 RDFS 扩展的包容结构关系.在 OIL 中,类可以说成是某个类表达式的子类,类表达式是类的一个布尔表达式. OIL 三种布尔运算符的类: AND、OR、NOT 均被定义为类 `oil:ClassExpression` 的子类.运算符 AND、OR 和 NOT 通过利用 `oil:hasOperand` 属性同运算符的操作数连接起来. `oil:hasOperand` 属性只用于具体类表达式的事实陈述,事实用 `rdfs:domain` 构造机制进行表达,这种类型的建模直接源自于 RDF 的以属性为中心的方法.

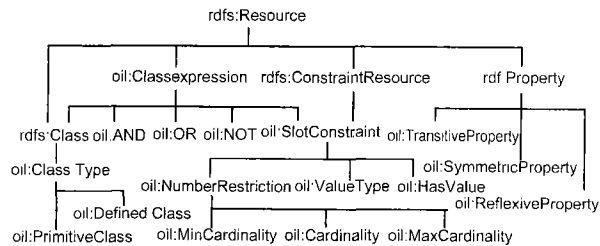


图 2 包容结构关系中 OIL 对 RDFS 的扩展

为了表示槽约束关系, OIL 引入了 `oil:SlotConstraint` 属性作为保留类,面向槽约束的某些特定类,如 `has-value`, `value-type`, `cardinality` 等等,都是 `oil:hasSlotConstraint` 子类. OIL 和 RDFS 都允许槽作为本体的首类成员,故 OIL 定义的槽可以很好地映射到 RDFS 的属性定义.并且, OIL 的槽定义允许定义槽之间的关系描述等多个方面.

3 逻辑描述原语本体扩展定义

作为 Web 的本体描述语言,在描述逻辑公式方面, OIL 提供了一定的描述能力.为了更加通用地描述逻辑公式,需要在 OIL 基础上扩展定义逻辑描述原语.

3.1 对 OIL 进行扩展的结构包容关系

图 3 是对 OIL/RDFS 进行扩展后的包容结构关系.为了描述命题逻辑中的等价和蕴涵运算,本文引入保留类 `oil:Formula`,在此基础上引入类 `oil:EQU` 和 `oil:IMP`,对 OIL 扩展后的五种运算均为 `oil:Formula` 子类,这里“oil”是对 OIL 扩展定义的名字空间.针对 Horn 公式的描述,本文引入保留类 `oil:HornClause` 作为 `oil:Formula` 的子类,同时为了表示无条件和无结论的 Horn 子句,引入特殊的类 `oil:EMPTY`.因为规则库是规则的集合,规则是具有前提和结论的蕴涵重言式.所以本

文引入保留类 `oil:RuleBase`、`oil:INF`、`oil:Premise` 和 `oil:Conclusion`。类 `oil:Premise` 和 `oil:Conclusion` 都是由公式组成,它们通过多重继承关系作为 `oil:INF` 和 `oil:IMP` 的子类。作为 `oil:IMP` 的子类,`oil:Premise` 和 `oil:Conclusion` 分别表示蕴涵的头部和尾部公式。

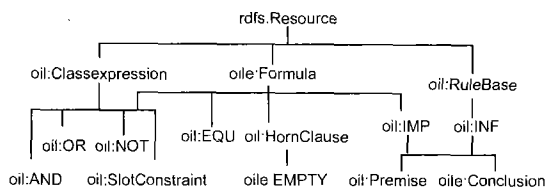


图3 对 OIL/RDF 扩展后的包结构关系

3.2 对 OIL 扩展的形式化定义

图3中对 OIL/RDFS 的逻辑描述原语扩展的形式化定义如下:

定义1 定义一种命题逻辑运算为 `oil:IMP`,表示逻辑“蕴含(\rightarrow)”运算。

定义2 定义一种命题逻辑运算为 `oil:EQU`,表示逻辑“等价(\equiv)”运算。

定义3 定义基于蕴含重言式的推理规则为 `oil:INF`,表示推理“ $P \vdash Q$ ”。

定义4 定义规则库为 `oil:RuleBase`,是推理规则的集合。

定义5 Horn子句(Horn Clause):Horn子句是具有如下形式的子句:

$$A_n \leftarrow (A_1, A_2, \dots, A_k \text{ 或 } A_n: \neg A_1, A_2, \dots, A_k)$$

定义6 定义 Horn子句为 `oil:HornClause`,是 `oil:Formula` 的子类。

定义7 定义一个类型为“EMPTY”,表示内容为空的逻辑公式。

根据以上的定义,扩展的命题逻辑和 Horn子句描述原语用 XML 语法表示如下:

命题公式 ::= 原子命题公式 | 复合命题公式;

原子命题公式 ::= 任何符合 XML 语法的文字描述的最小命题;

复合命题公式 ::= NOT 运算公式 | AND 运算公式 | OR 运算公式 | IMP 运算公式 | EQU 运算公式;

NOT 运算公式 ::= ‘<oil:NOT>’命题公式‘</oil:NOT>’

AND 运算公式 ::= ‘<oil:AND>’命题公式 命题公式 + ‘</oil:AND>’

OR 运算公式 ::= ‘<oil:OR>’命题公式 命题公式 + ‘</oil:OR>’

IMP 运算公式 ::= ‘<oil:IMP>’前提 结论‘</oil:IMP>’

EQU 运算公式 ::= ‘<oil:EQU>’命题公式 命题公式‘</oil:EQU>’

Horn子句 ::= ‘<oil:HornClause>’前提 结论‘</oil:HornClause>’;

前提 ::= ‘<oil:Premise>’命题公式‘</oil:Premise>’

结论 ::= ‘<oil:Conclusion>’命题公式‘</oil:Conclusion>’

规则库 ::= ‘<oil:RuleBase>’推理规则‘</oil:RuleBase>’

推理规则 ::= ‘<oil:INF>’前提 结论‘</oil:INF>’

上面定义的逻辑描述原语可以转化为用 RDFS 定义的 RDF 模式描述。

3.3 逻辑公式的描述实例

基于上面的扩展定义,我们可以表示所有的一阶命题逻辑公式、Horn子句和规则。图4是基于 OIL 及其扩展定义的描述原语对典型的蕴涵重言式公式“ $P_1, P_2, \dots, P_m \vdash Q_1, Q_2, \dots, Q_n$ ”的描述表示。

```
< oil: InferenceRule >
  < oil: Premise >
    < oil: AND >
      < oil: hasOperand rdf: resource = " # P1" />
      < oil: hasOperand rdf: resource = " # P2" />
      .....
      < oil: hasOperand rdf: resource = " # Pm" />
    < oil: AND >
  < oil: Premise >
  < oil: Conclusion >
    < oil: OR >
      < oil: hasOperand rdf: resource = " # Q1" />
      < oil: hasOperand rdf: resource = " # Q2" />
      .....
      < oil: hasOperand rdf: resource = " # Qn" />
    < oil: OR >
  < oil: Conclusion >
< oil: InferenceRule >
```

图4 命题逻辑公式描述示例

4 结束语

随着 XML 和 RDF 相继成为 W3C 标准,语义 Web 才刚刚引起众多研究者的注意,有待研究的相关问题还很多。传统的本体描述语言在应用领域的概念抽象、以及逻辑公式、规则和公理描述方面具有完备的理论基础,但不是面向以标记为特征的 Web 计算环境,因此语义 Web 中描述知识、规则、公理以及查询推理公式等需要定义相关的原语,定义的原语语义需要本体模型进行形式化。

虽然 OIL 定义的逻辑运算可以描述逻辑公式、规则和公理,但是只能表示一些代数属性而不支持广泛地公理描述。即使 OIL 借用 FL 的方法,在概念和槽的定义方面具有比较强大的功能,可以弥补一些 OIL 在公理化描述的不足,但对于复杂的公理、公式和规则的处理,由于 OIL 在公理和逻辑代数本身对象化方面存在不足,加之要尽量考虑 FL 和 DL 功能的继承,因此,在公理的对象化处理上显得薄弱而导致其难以描述类似于蕴涵、等价、Horn 逻辑的支持等。本文基于 OIL/RDFS 扩展了描述一阶命题逻辑公式、规则以及 Horn 子句的描述原

语,并进行了形式化的定义,这样可以在基于 RDF/XML 的语义 Web 中广泛地描述知识、规则和公理,并且定义的原语在语义上同语义 Web 的现有标准达到最大程度地兼容。

在本文研究的基础上,紧接着的研究就是定义相应的一阶谓词、逻辑公理以及查询推理描述原语,这样就可以广泛地描述一阶逻辑公式、规则、公理和查询推理公式,甚至把 PSM 的研究成果应用于 Web 推理过程的控制,从而使 Web 成为一个可控的分布式知识库系统。同时智能信息集成、机器学习、知识获取等都是语义 Web 走向成熟需要研究的各种应用和问题。

参考文献:

- [1] I Horrocks, D Fensel, J Broekstra, et al. The Ontology Inference Layer OIL[EB/OL]. <http://www.ontoknowledge.org/oil>.
- [2] W3C. Resource Description Framework(RDF) Model and Syntax[S]; Resource Description Framework(RDF) Schema Specification[S]. W3C Recommendation, <http://www.w3c.org/TR/>.
- [3] B Chandrasekaran, et al. What are ontologies, and why do we need them[J]. IEEE Intelligent System, 1999, 1:20-26.
- [4] Rudi Studer, et al. Knowledge engineering: Survey and future directions[A]. In F. Puppe, et al. (eds.), Lecture Notes in Artificial Intelligence (LNAI)[C]. Wurzburg, German: Springer-Verlag, 1999, LAI 1570.
- [5] Tim Berners-Lee. Semantic Web Architecture[DB/OL]. <http://www.w3.org/2000/talks/1206-xml2k-tbl/slide1-0.html>.

作者简介:



姚绍文 男,博士,1966年8月生,湖南省永顺县人,云南大学信息学院教授、网络智能计算研究室主任、中国计算机学会西南网络与 MIS 专委会副主任委员,1998年至1999年为澳大利亚南澳大利亚大学电信研究所访问学者,主要研究方向为语义 Web 技术、网络协议工程、网络分布式计算、着色 Petri 网(CPN)建模、知识工程技术等,近年来,先后在国内外发表学术论文 30 余篇。



余江 男,1961年3月出生,电子科技大学通信工程专业毕业,云南大学信息学院教授,1998年至1999年赴澳大利亚留学,长期从事电子信息领域的教学、科研及产品开发工作,主要研究方向为计算机网络、移动通信、图像处理、光电子技术,主持或参加国家自然科学基金项目、云南省自然科学基金项目等科研项目 8 项,在学术期刊及学术会议上发表论文 20 多篇,曾获得省部级科技成果一等奖 1 项。

周明天 男,1939年3月出生于广西容县,电子科技大学计算机学院教授、博士生导师,主要研究方向为网络信息系统、分布式处理系统、对象中间件技术等,在国内外学术期刊及会议上发表学术论文 100 余篇。