

文章编号: 1005-0523(2007)02-0078-04

基于本体的语义推理研究

宋 岚, 黄兆华

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘要: 语义 Web 作为新一代 Web 正倍受关注, 而 Web 的资源描述语言 RDF、RDF 的词汇描述 RDF Schema, 以及最新的本体语言 OWL 等已成为现今语义 Web 上的规范语言. 文中概述了本体语言和描述逻辑及它们之间的关系, 并通过一个实例显示了一个本体文件中传递闭包的推理路径.

关键词: 语义 Web; RDF; OWL; 描述逻辑; 语义推理

中图分类号: TP291.72

文献标识码: A

0 本体语言概要

语义 Web (Semantic Web) 是当前万维网的扩展, 1999 年由万维网创始人 Tim Berners-Lee 等人提出, 其目的是通过结构化和形式化的方法来表示 Web

上的资源, 使得计算机程序能够对网络资源进行分析和推理. 这首先要求机器“理解”网络资源, 因此需要一系列的统一语言来表示 Web 上的资源. 于是, Tim Berners-Lee 在 2000 年提出了语义 Web 的体系结构^[1]. 如图 1 所示:

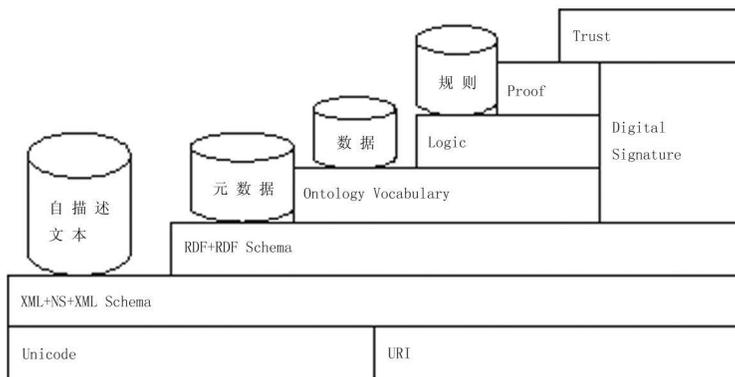


图 1 语义网框架图

构建语义 Web 体系结构的语义层首先由 RDF (Resource Description Framework, 资源描述框架) 开始^[2]. RDF 是用来表现万维网上各类资源的信息的一种语言. 它采用 Web 标识符 (统一资源标识符 URI) 来标识事物, 采用三元组 (资源, 属性, 值) 表示资源, 但这种说明性的语言没有提供机制来描述属

性或说明属性与其他资源之间的关系, 因此需要 RDF 的词汇描述语言 RDF Schema (RDFS). RDFS 是在 RDF 的基础上引进了类的概念、类之间的包含关系、属性之间的包含关系, 以及属性的定义域和值域^[3]. RDFS 中的 URI 都是以“rdfs:”为命名空间, 如: rdfs:resource, rdfs:domain, rdfs:range, rdfs:subProperty-

收稿日期: 2007-02-20

基金项目: 江西省自然科学基金项目 (0411035).

作者简介: 宋岚 (1978-), 女, 江西南昌人, 华东交通大学研究生, 研究方向: 人工智能.

Of, rdfs:subClassOf, rdfs:class, 等等。

但 RDFS 的表达能力并不足以表达 Web 上资源所属领域的概念模型,如 RDF(S)不能自定义新的概念,无法强调属性的传递性或对称性,这些局限性导致了 Web 本体语言的提出。

OWL(Web Ontology Language),是 W³C 推荐的语义 Web 中本体描述语言的标准^[2]。它在 RDFS 的基础上提供了更为丰富的类描述能力。

针对不同的需求 OWL 有三个子语言:OWL Lite, OWL DL 和 OWL Full^[4]。

OWL Lite 用于提供给那些只需要一个分类层次和简单约束的用户,比如辞典(thesauri)和分类系统(taxonomy)等。OWL Lite 对表达能力进行了约束,例如,虽然 OWL Lite 支持基数限制,但只允许基数为 0 或 1。相比 OWL DL, OWL Lite 具有更低的形式复杂度。

OWL DL 用于支持那些需要最强表达能力,并需要保持计算完备性(computational completeness,即所有的结论都能够确保被计算出来)和可判定性(decidability,即所有的计算都能在有限的时间内完成)。OWL DL 这么命名是因为它对应于描述逻辑。

OWL Full 用于支持那些需要尽管没有可计算性保证,但有最强的表达能力和完全自由的 RDF 语法的用户。例如,在 OWL Full 中,一个类可以被同时看成许多个体的一个集合以及本身作为一个个体。使用 OWL Full 相比于 OWL DL,对推理的支持是更难预测的,因为目前还没有完全实现在 OWL Full 上的推理。OWL Full 可以看成是对 RDF 的扩展,而 OWL Lite 和 OWL DL 可以看成是对一个受限的 RDF 版本的扩展。所有的 OWL 文档(Lite, DL, Full)都是一个 RDF 文档;所有的 RDF 文档都是一个 OWL Full 文档;但只有一些 RDF 文档是一个合法的 OWL Lite 和 OWL DL 文档。

OWL 中的 URI 都是以“owl:”为命名空间,例如,owl:SymmetricProperty, owl:FunctionalProperty, owl:inverseOf, owl:Restriction, owl:intersectionOf, owl:allValuesFrom, owl:hasValue, owl:equivalentClass, 等等。

1 逻辑描述语言

描述逻辑(Description Logic)是基于对象的知识表示的形式化,也称为概念表示语言或术语逻辑,是一阶谓词逻辑的一个可判定子集。

例如在 Mother 和 Parent 之间的关系说明“moth

ers are parents”,被看作是一种“IS-A”关系。描述逻辑的一个标志特征是它们能表示除 IS-A 关系外的更多的概念之间的关系^[5]。

描述逻辑是建立在概念(concept)和角色(role)之上,由构子(constructor)从简单概念和角色中构造出复杂概念和角色。概念对应于经典逻辑中的一元谓词,角色对应于二元谓词,构子决定着语言的表达能力。知识库主要由 TBox 和 ABox 组成,前者是描述概念关系的术语公理集,后者是描述个体实例的断言公理集,现已有 FaCT 和 RACER 等实用推理机^[6]。

描述逻辑在许多领域中被作为知识表示的工具,如信息系统、数据库、软件工程、网络智能访问、规划等。Horrocks 对表达能力较强的描述逻辑进行了研究,并建立了一些逻辑框架和系统,如 FaCT, SHIQ 等。他和 Dieter Fensel 等人将描述逻辑、语义网和 DAML 结合起来,提出了 DAML+OIL 语言,其以描述逻辑作为核心的表示和推理基础。并在 XML 及其 RDF 上面进行了扩展,用描述逻辑来研究语义网络和本体论。

OWL 的抽象句法与描述逻辑之对照见表 1 所示。OWL 的本体由公理和事实组成,这正对应着描述逻辑的知识库^[2],而后者是由 TBox 和 Abox 组成的。

表 1 OWL 与描述逻辑(DL)名称对照

OWL	类	属性	数据类型	公理	事实	本体
DL	概念	角色	有型域	Tbox	Abox	知识库

2 一个基于本体的语义推理实例

Jena 工具是一个 Java 开发工具包,由 HP 公司的 Brian Me Bride 开发,起源于早些时候 SiRPAC API 的工作。它允许应用程序解析、创建和查询 RDF 模型 model。Jena 定义了许多的接口来访问和处理 model 模型中的 RDF statement,包括 RDF Node 接口、资源接口、常量接口、属性接口、陈述接口等。由于 Jena 在数据库支持方面的可靠性和持续稳定性,它一直是各类应用系统在开发过程中的首选。在最近几年的 WWW 会议上,提出的许多开发成功的语义网工程也都是采用 Jena API 来处理 RDF。

以一个动物本体为例,我们定义了三个类 plant(植物)、Animal(动物)、Zoo(动物园)。它的类层次图如下所示:

上图中动物本体的中英文索引表如下所示:

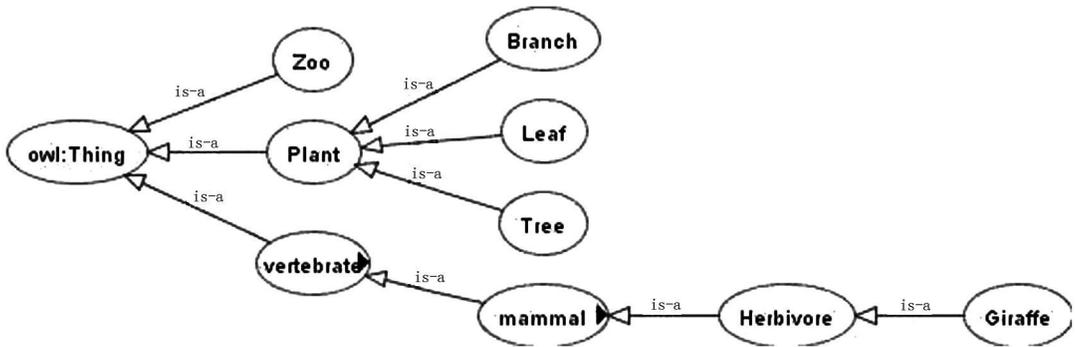


图2 一个动物本体

表2 动物本体中英文索引表

英文	中文
Vertebrate	脊椎动物
Mammal	哺乳动物
Herbivore	素食动物
Giraffe	长颈鹿

在下面程序中,我们用 a 代表脊椎动物,b 代表哺乳动物,c 代表素食动物,d 代表长颈鹿,p 代表包含关系.那么可以表示为如下几个三元组(其中 demo: 代表名称空间“http://jena.hpl.hp.com/demo#”):

```
demo:a demo:p demo:b
demo:c demo:p demo:a
demo:b demo:p demo:d
```

那么我们定义一个规则来支持关系 demo:p 的传递闭包属性:

```
rule1:(? a demo:p ? b)(? b demo:p ? c)
->(? a demo:p ? c)]
```

连接 Jena 自带的推理机(也可以连接 Racer, Fact 等推理机)后,我们可以查询出长颈鹿同时也是一种脊椎动物,哺乳动物.

```
for (StmtIterator i=inf.listStatements(a,p1,d);
i.hasNext();i.hasNext()) {
Statement s=i.nextStatement();
System.out.println(s);
for (Iterator id=inf.getDerivation(s);
id.hasNext();) {
Derivation deriv=(Derivation) id.next();
deriv.printTrace(out,true);
}
}
```

程序运行后输出结果如下所示:

```
[http://jena.hpl.hp.com/demo#a,
```

```
http://jena.hpl.hp.com/demo#p,
http://jena.hpl.hp.com/demo#d]
Rule rule1 concluded (demo:a demo:p demo:d)
< -
Fact (demo:a demo:p demo:b)
Rule rule1 concluded (demo:b demo:p demo:d)
< -
```

```
Fact (demo:b demo:p demo:c)
Fact (demo:c demo:p demo:d)
结果表明由事实集可以推出三元组 (demo:a
demo:p demo:d). 路径依次是:事实集 Fact (demo:b
demo:p demo:c)和 Fact (demo:c demo:p demo:d)推
出三元组(demo:b demo:p demo:d),再和 Fact (demo:
a demo:p demo:b)推出结论 (demo:a demo:p demo:
d).
```

本体中的推理包括:计算本体类层次关系;检查概念一致性;计算本体内部关系和隐含关系的合法性;检查实例个体是否是合法的实例等.

3 结束语

RDF 与 OWL(Lite 与 DL)语言可以认为是谓词逻辑的特例.通过逻辑公理的形式以公认的语义已经对他们进行了阐述.这些专用语言存在的理由就是它们提供了一个非常适合特定用途(在我们的环境中就是基于标签的网络语言)的语法.还有一个主要的理由就是它们定义了一个合理的逻辑子集.在特定的逻辑中,在表达能力与计算复杂性之间存在着一个交替情形:语言表达能力越强,相应的推理系统的效率就越低.OWL Lite 和 OWL DL 大致遵循描述逻辑^[2],这个描述逻辑是谓词逻辑的一个子集,利用它可构造一个有效的推理系统.基于 RDF 三元组句法的 OWL FULL 是与 RDF 图紧密相连的,而 RDF 图上允许出现匿名结点、环路结构等,这些都已超出

描述逻辑的常规语法.对此,OWL FULL并未做任何限制,同时也没约束属性的传递性和数量限定等条件,因此OWL FULL是不可判定的^[7].在实际运用中,现已有切实可行的描述逻辑系统FaCT和RACER等推理机,能为OWL Lite提供推理服务,而OWL DL需要相对较高的计算完备性和可判定性,因此OWL DL的推理实现还需进一步的研究.

参考文献:

- [1] T. Berners-Lee. The Semantic Web [J]. Scientific American, 2001, (6): 1-6.
 [2] Grigoris Antoniou. A Semantic Web Primer [M]. Cambridge,

- Massachusetts London, England: The MIT Press, 2004.
 [3] D. Brickley. RDF Vocabulary Description Language 1.0; RDF Schema [EB/OL]. <http://www.w3.org/TR/rdf-schema/>, 2004-02-10.
 [4] Michael K. Smith. OWL Web Ontology Language Guide [EB/OL]. <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>, 2004-02-10.
 [5] F. Baader. Handbook of Description Logics [M]. Cambridge University, 2003.
 [6] V. Haarslev. RACER system description [M]. Springer, 2001.
 [7] I. Horrocks. Reducing OWL entailment to description logic satisfiability [C], the 2003 International Semantic Web Conference (ISWC 2003), Springer, 2003; 17-29.

Study of Semantic Reasoning Based on Ontology

SONG Lan, HUANG Zhao-hua

(School of Information Engineer, East China Jiaotong Univ, Nanchang 330013, China)

Abstract: The next generation Web, Semantic Web, has recently been drawn considerable attention from both academia and industry. RDF, RDF Schema, OWL etc. are the recommendation languages in Semantic Web. In this paper, we show the relationships of ontology language and Description Logic, and the path on transitive closure of relation have been computed through an example.

Key words: semantic web; RDF; OWL; description Logic; semantic reasoning

补白:

科技论文中参考文献的类型

根据国标 3469 规定,科技论文最后的参考文献以单字母标识类型:

文献标识	M	J	C	N	D	R	S	P
文献类型	专著	期刊文章	论文集	报纸	学位论文	报告	标准	专利

对于专著和论文集中的析出文献,用单字母“A”,其他未说明的,用“Z”.例:

- [1] 刘国均. 图书馆目录[M]. 北京:高等教育出版社,1957.
 [2] 李西. 信息技术论文集[C]. 北京:中国社会科学出版社,1994.
 [3] 张平. 微分动力系统不变集[D]. 北京:北京大学数学研究所,1983.
 [4] 金一,王建中,等. 一种在线检测局部放电的数字技术[J]. 电子电路, 1993, 23(4): 28-32.
 [5] 谢希德. 创造学习的新思路[N]. 人民日报, 1998-12-23.
 [6] GB 50023/T-95. 建筑抗震鉴定标准[S].
 [7] 王信林. 建筑砌块联接件[P]. 中国专利:CN 1036800, 1997-09-29.
 [8] 李军. 循环图的自同构群[R]. 南昌:华东交通大学数学研究所,2003.