

语义 Web 与推理系统的应用研究*

朱创录

(渭南师范学院 计算机科学系, 陕西 渭南 714000)

摘要: 分析了语义 Web 的体系结构及其相关知识, 对本体论的逻辑基础进行了研究。在此基础上提出了一种有效的推理系统。

关键词: 语义 Web; 描述逻辑; 推理引擎

中图分类号: TP391

文献标识码: A

The research of semantic Web and reasoning system

ZHU Chuang Lu

(Department of Computer Science, Weinan Teachers University, Weinan 714000, China)

Abstract: This paper analyzed the framework of the semantic Web and interrelated knowledge, and then did some researches about the basic logic of ontology. At last, an effective reasoning system based on that research is proposed.

Key words: semantic Web; description logic; inference engine

语义 Web 发展很重要的目的就是要建立人与机器都可以理解的知识系统, 将信息用语义 Web 本体论的思想来表达可以部分解决机器无法理解人类所描述知识的问题。在语义 Web 本体论的基础上建立机器进行自动推理的方法, 将所描述的知识进行推理并重新组织加工。目前, 具有通用性的推理系统主要是专家系统, 在专家系统中所使用的知识呈现多元化, 其中包含语义网络、框架、法则、逻辑等。虽然这些知识的表示方法各有其优点, 但在建立推理系统时都存在各种各样的缺点和不足。传统的框架表示知识无法表现的交集、合集等概念; 语义网络表示的知识由于表现方式不是结构化, 所以带来了运算不便的问题; 法则表示知识通常很难表现类的层次关系, 存在知识表现不紧凑的问题; 逻辑的方式表示知识存在的问题是无法将类别与实例分开, 使得在推理时无法单独从类别中推导公理性的知识, 使推理的效率难以提高^[1]。另外, 通过传统的知识表现形式所表达的知识只能应用于特定的应用环境中, 很难做到被其他系统重复利用, 而语义 Web 却完全可以做到这一点, 因为语义 Web 本体论的出发点就是要建立人和机器都可以理解的知识体系。

语义 Web 本体层能够解决机器推理存在的知识表

示问题, 并且基于本体描述的知识可以多次重复利用于相关领域的推理, 充分利用所建立的知识库。本文主要就是依据这两点来探究关于语义 Web 本体论应用于传统的专家系统进行推理时需要解决的问题。

1 语义 Web 的技术

为了实现语义 Web 信息服务智能化与自动化的目的, 语义 Web 研究者们开发了许多新技术并提出了一系列的技术标准。从某种程度上讲, 语义 Web 的发展就是建立在一定逻辑基础上的信息表示语言和标准的发展。由于语义 Web 的知识表示具有创建上的分散性; 同时又具有应用上的通用性, 所以需要有一个统一的框架, 这个框架应该能够满足这种分散性, 以及由这种分散性所带来的安全性, 以满足知识跨应用、跨领域的可互操作性。Berners Lee 在 XML2000 大会上描述了这个体系结构^[2], 如图 1 所示。

在这个层式结构中, 上层将下层的语言机制作为本层的支撑语言, 如表 1 所示。通过分析器, 从合法有效的下层描述中抽取本层所能理解的模型, 实现更多的语义处理功能。因此, 语义 Web 是向下兼容的, 比如, 当用 RDF/XML 表达 OWL 时, 如果只有 RDF 分析器, 虽然仅仅可以部分解释该本体内的类结构, 但能完全解释该本体内的所有实例。因为这与 RDF(S) 语法相关。由于

* 基金项目: 渭南师范学院专项科研基金(07YKZ035)

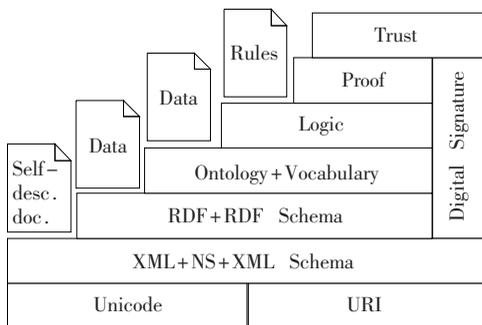


图1 语义Web的体系结构

表1 语义Web的层次依赖关系

层次	名称	描述
第1层	Unicode 和 URI	整个语义Web的基础,Unicode 处理资源的编码,URI 负责标识资源。
第2层*	XML+NS+XML Schema	用于表示数据的内容和结构。
第3层*	RDF+RDF Schema	用于描述Web的资源及其类型。
第4层*	Ontology Vocabulary	用于描述各种资源之间的联系。
第5层	Logic	在下面4层的基础上进行的逻辑推理操作。
第6层	Proof	
第7层	Trust	

注：* 核心层,用于表示Web信息的语义。

Web的海量信息不可能在短时间内全部建立语义标记,而且语义的标记方法也将不断发展,将会提供更广泛的语义支持能力。这种上层兼容下层的方式,为语义Web的发展提供了良好的基础。

2 基于语义Web的专家推理系统设计

传统的语义Web推理形式本身虽然具有很多优点(尤其是在描述逻辑推理方面较强),但其自身也存在相应不足,主要是由描述逻辑本身的不足造成的,因此采用描述逻辑结合SWRL规则来表示知识和规则。本设计中采用的推理方式包括2种:(1)基于描述逻辑的本体推理,应用描述逻辑产生推理的目的是为了使建立的本体有更好的层次结构,同时不产生冲突定义^[1];(2)基于专家系统的推理,主要目的是实现本体与基于本体的规则相结合产生推理。

在语义Web中,目前对于本体的表示多采用OWL DL进行描述,采用OWL DL建立起来的本体完全具有描述逻辑的推理能力^[4],因此可以采用Tableaux算法推理机对所建立的本体进行推理,发现其隐藏知识,并且发现知识中的矛盾之处,方便对本体进行修改和完善;另外从语义Web引入本体论的目的来看,采用本体论的观点,可以使人与机器都完全理解人类所建立起来的本体知识,便于知识的共享和再次利用。以OWL DL建立起来的本体具有丰富的表达能力,所以在此基础上寻找相应的规则表示方法来进行基于专家系统的推理,本文采用的规则表示方法是语义Web规则描述语言SWRL,其优点是SWRL能够很好地结合OWL DL本体

知识。因为SWRL自身的建立就是基于OWL本体之上,在定义规则时,可以充分结合OWL本体中丰富的表达能力达到推理目的。在用OWL DL建立本体基础上,采用Tableaux算法推理机对所建立的本体进行推理。通常使用RACER等推理机作为本体的推理工具,生成一个新的没有语义冲突的层次化本体。在此基础上定义SWRL规则,完成了知识库和规则的设计。由于现在的专家系统均不能直接处理OWL表示的知识和SWRL定义的规则,所以需要进行相应的格式转换。本文通过JessTab和XSLT分别对本体知识和SWRL规则进行格式转换,最后,将转换后的Jess格式的知识和规则导入到Jess推理引擎进行基于专家系统的推理。

根据以上分析,建立一种基于语义Web的推理系统模型,如图2所示。其中,“推理1”使用RACER,表示具有描述逻辑推理功能的推理机,因为基于Tableaux算法的RACER不仅可以与Protégé很好地进行描述逻辑的推理,而且可以使用专家系统Jess通常的推理引擎,很方便地进行以OWL DL描述的本体推理。推理系统中的知识库使用语义Web的本体描述语言OWL DL进行表示;使用语义Web规则语言SWRL进行定义。图2中,“本体1”表示根据领域知识建立起来的本体;“本体2”经过“推理1”进行基于描述逻辑的推理后建立的层次化更加良好,并且满足一致性要求,没有冲突发生的本体。其中,SWRL规则的定义使用“本体2”中的TBox和ABox,也就是“本体2”的概念、属性或者实例,在这些OWL DL基本元素的基础上,根据相应的推理

要求结合SWRL语法格式可以产生与本体具有良好结合的SWRL规则。

将该推理系统分为3个步骤进行设计:

(1)采用OWL进行领域本体的设计以及在领域本体基础上的SWRL规则设计。

本体的建立可以结合应用,根据具体的本体论工程模型进行本体的设计,通常使用的工具是Protégé。Protégé是一个开放的开发环境,可提供多种插件进行功能的扩充。建立SWRL规则时要充分考虑并且结合已经

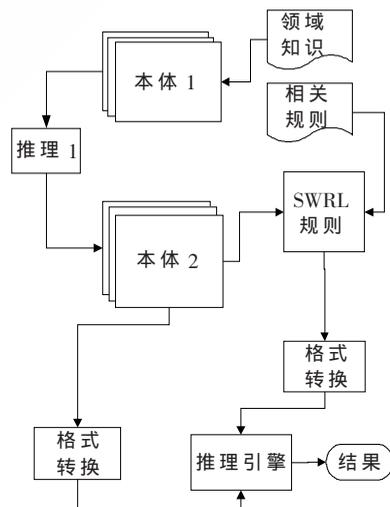


图2 总体设计

要求结合SWRL语法格式可以产生与本体具有良好结合的SWRL规则。

将该推理系统分为3个步骤进行设计:

(1)采用OWL进行领域本体的设计以及在领域本体基础上的SWRL规则设计。

本体的建立可以结合应用,根据具体的本体论工程模型进行本体的设计,通常使用的工具是Protégé。Protégé是一个开放的开发环境,可提供多种插件进行功能的扩充。建立SWRL规则时要充分考虑并且结合已经

技术与方法 Technique and Method

设计好的本体进行相关规则的设计,避免出现规则冲突。建立 SWRL 规则仍然可以使用 Protégé,只需在 Protégé 下增加 SWRL Tab 的插件,OWL 本体建立与 SWRL 规则建立在同一开发环境下的好处是在制定规则时可以自动调用本体中的知识,避免人工写入时可能发生的错误。领域本体建立后,通过将本体知识导入到描述逻辑推理机上进行推理,可以对本体进行分类整理,发现其中的隐含知识,建立一套基于语义 Web 本体论的知识和规则。OWL DL 知识库具有描述逻辑的特性,可以更好地表现知识、进行推理。SWRL 规则可以充分地结合本体进行规则设计,使得在设计规则的时候更方便。

(2)将 OWL 知识和 SWRL 规则转化为推理引擎可以接受的格式。

现在的专家系统通常不接受 OWL 和 SWRL 表示的规则,因此需要进行格式的转化。本体知识库格式的转换,首先将本体导入到描述逻辑的推理机,通过描述逻辑的推理机将本体进行推理。因为 ObjectProperty 属性实例是对象关系的实例,通常需要推理机首先将其客体连接到具体的 DataProperty,这样就可以在推理机中形成具体的三元组,将三元组导入格式转换程序进行格式转化,转换后的格式推理引擎可以接受。因为 SWRL 文件本身就是 XML 文件,并且具有固定的语法格式规范(通过 Schema 定义)。所以可以通过 XSLT 的方式进行格式的转化,从而使推理引擎可以接受规则。

(3)专家系统的推理。

将知识和规则导入到专家系统的推理引擎,进行模式匹配运算。最终将运算的结果通过专家系统的 API 和用户应用程序进行交互。推理引擎的选择种类很多,可以根据具体情况进行选择。

语义 Web 的知识表达形式性,是在传统的描述逻辑基础上建立起来的,因此继承了描述逻辑的一些优点。但由于受描述逻辑的限制,也存在一些描述逻辑的自身逻辑缺陷,本文使用 OWL DL 结合 SWRL 可以对相关推理进行必要的扩充,从而形成一个较完整的推理系统。语义 Web 作为一项新的信息表达形式,在目前的研究中出现了许多分支,表达方式出发点是更好地进行信息描述,使人与机器都可以理解信息。而针对目前 Web 的信息量十分庞大的情况,语义 Web 的重点是要解决机

器对语义理解的问题。目前的语义 Web 在描述现实世界的静态事物上已经表现得很好了,这主要取决于基于描述逻辑本体理论的应用,正是这种应用推动了机器对所描述事物的理解。

该推理系统主要基于 OWL 本体和 SWRL 规则。本体的建立只是在描述逻辑的基础上建立了一个知识库,虽然这种知识库改进了传统知识表示方法的一些缺点和不足,但这种知识库并不能充分表达推理层面上的复杂规则,这主要是由于描述逻辑的缺陷所造成的(例如对于通常的条件判断,描述逻辑是无法表达的),这同样体现在本体中。所以,在定义一些限制规则的时候,往往无法完全地应用本体表示。为了能够充分地表达这些规则,采用 SWRL 作为规则的表现形式,因为 SWRL 规则的建立是在所建立的本体基础之上的,所以 SWRL 作为规则的表现形式可以充分与本体知识库结合。当需要对本体所建立的知识进行某种推理的时候,可以根据本体建立 SWRL 法则,实现本体的描述和规则的分离。因为本体是作为一般意义上的定义,采用本体建立的知识库在应用时并不需要进行大的修改,只要增加所需要部分的知识信息,并且在此基础上设计相应的 SWRL 推理规则就可以完成推理服务,提高本体知识的利用率。另外,基于 OWL 的知识本身采用 XML 的形式,从而为知识库的共享和重复使用带来了极大的便利。

参考文献

- [1] VAN K R W, VISSER P R S, BENCH C T J M, et al. A principled approach to developing legal knowledge systems[J]. International Journal of Human-Computer Studies, 1999, 51(6):1127-1154.
- [2] 朱创录.基于语义 Web 的自动推理研究[D].西安:西北大学,2006.
- [3] HORROCKS I, PETER F, SCHNEIDER P, et al. From SHIQ and RDF to OWL: The making of a Web ontology language[J]. Web Semantics, 2003, 1(1):7-26.
- [4] 邓志鸿,唐世渭.Ontology 研究综述[J].北京大学学报(自然科学版),2002, 38(5):124-127.

(收稿日期:2009-09-10)

作者简介:

朱创录,男,1977年生,硕士,讲师,主要研究方向:网络应用与网络安全。