

# 本体推理研究

王常宇

作业	分数[70]
得分	

2020 年 12 月 18 日

# 本体推理研究

王常宇

(大连海事大学 信息科学技术学院 大连 116026)

**摘 要** 本体作为概念模型的明确的规范说明越来越受到重视, 本文在阐述本体的基础上, 从语义知识和推理以及逻辑推理的角度上, 研究本体推理。并重点探究了描述逻辑、描述逻辑与 OWL 的关系、描述逻辑到一阶谓词逻辑的映射、OWL 本体的建立以及本体和规则的一阶逻辑映射。

**关键词** 本体; 本体推理; 描述逻辑; 一阶谓词逻辑

## The research of Ontology Reasoning

Wang Chang Yu

(School of Information Science and Technology, Dalian Maritime University, Dalian 116026)

**Abstract** Ontology is an explicit specification of a conceptualization receiving increasing attention, the paper elaborated on the basis of the Ontology, from the perspective of semantic knowledge and reasoning and logical reasoning, studies Ontology reasoning. And focus on exploring the Description Logic, the relationship between Description Logic and OWL, the Description Logic to First-Order Logic, the building of Ontology and how to establishing rules and OWL Ontology mappings from first-order logic.

**Key words** Ontology, Ontology reasoning, Description Logic, First-Order Logic

## 1 引言

自从 2000 年 Tim Berners-Lee 提出了语义 Web 的理念和框架结构后, 这个被称之为下一代网络的技术就受到了普遍的关注, 在过去的近十年时间内, 人们对相关层面的知识架构、表示、推理等都进行了深入的研究, 并不断的引入相关领域的技术思想应用于语义 Web 的开发和实现<sup>[1]</sup>。Time Berners-Lee 指出语义网是当前万维网的扩展和延伸。语义网的信息具有充分的、完备的语义定义, 能够在人与计算机之间建立语义上的理解与合作。研究语义网就要从当前浩如烟海的网络数据中发掘与产生具有语

义本质的结构和模式; 并能建立能体现自然语义的、并与网络与一结构像协调的基础语义构架。本文重点介绍利用逻辑推理将自然语言通过本体作为概念的方法建立机器能够理解的具有语义本质的结构和模式。

## 2 本体

在人工智能界, 最早给出本体定义的是 Neches 等人, 他们将本体定义为“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的规则的定义”。Neches 认为: “本体定义了组成主题领域的词汇表的基本术语及其关系, 以及结合这些术语和关系来定义词汇表

外延的规则”。这个定义认为本体的概念包括四个主要方面：(1)概念化：客观世界的抽象模型；(2)明确：概念及它们之间联系都被精确定义；(3)形式化：精确的数学描述；(4)共享：本体中反映的知识是使用者共同认可的。

### 3 本体推理

#### 3.1 语义知识和推理

虽然计算机视觉、听觉研究，计算机语言学等都涉及机器理解的领域，但是计算机对于理解信息资源的语义仍然欠缺。[1]而语义网和本体的方式则能很好的解决这个问题。人们在交流时需要使用词语或者符号，这些词语或者符号都是有所指的（指向事物），但“词语”到事物的映射不是直接的，两者之间以“概念”作为桥梁。当人们看到特定上下文、语境中的一个“词语”或者一种“符号”时，之所以能够准确、迅速地对应到相应的“事物”，是因为经过了“概念”这一中间层。对于机器而言，如果要让它们能够准确、迅速的理解“词语”、“语义三角”和“事物”之间的映射关系，也要借助“概念”。事实上，如果仅仅给机器提供“词语”、“符号”，它们确实理解的很不好，甚至理解不了。为了让机器理解“词语”、“符号”中蕴含的语义知识，让机器学会推理，给出如图1所示的语义连续体，来说明从人们的尝试理解和表达到达机器的理解和表示的过程。[1]



图1 语义连续体

知识是语言和推理机制的结合，推理则建立在逻辑的基础之上，逻辑可分为命题逻辑、一阶逻辑、描述逻辑和框架逻辑。

#### 3.2 一阶谓词逻辑

一阶谓词逻辑（FOL, First-Order Logic）在计算机科学中有着广泛的应用,它不仅是

程序设计理论、程序逻辑研究的重要基础，而且还是程序正确性证明、定理机器证明和知识表示技术的有力工具。

一阶谓词逻辑建立在命题的基础上。其取值为真或假（表示是否成立）的句子称作命题。一阶逻辑中，命题被分解为个体和谓词两部分。带有一个变量(个体)的谓词叫一目谓词，带有两个变量(个体)的谓词叫二目谓词。

##### (1) 一阶谓词逻辑知识表示体系

在一阶谓词逻辑中使用的符号一般包括：常量、变量符号、函数符号、谓词符号、函数和谓词符号。以上符号中常量用来表示个体；变量表示非特定的个体；谓词用来表达  $n$  个实体之间的关系或属性；其取之为真或假；函数仅实现个体域中  $n$  个个体到某一个体的映射，没有真假取值。还有连接词和量词利用这两个符号，可把单个谓词组合成复杂的谓词公式，表达复杂的领域知识。一阶谓词逻辑语言的直观意义容易理解为：

“符号表”相当于自然语言中的字母表，“公式”代表完整的句子。而其中的函数符号构造复杂的个体，谓词符号则构造最原子的公式。

##### (2) 一阶谓词逻辑的推理

称蕴涵式  $(A_1 \wedge A_2 \wedge \dots \wedge A_k) \rightarrow B$  为推理的形式结构， $A_1, A_2, \dots, A_k$  为推理的前提， $B$  为推理的结论。若  $(A_1 \wedge A_2 \wedge \dots \wedge A_k) \rightarrow B$  为永真式，则称从前提  $A_1, A_2, \dots, A_k$  推出结论  $B$  的推理正确（或说有效）， $B$  是  $A_1, A_2, \dots, A_k$

的逻辑结论或称有效结论，否则称推理不正确。若从前提  $A_1, A_2, \dots, A_k$  推出结论  $B$  的推理正确，则记为  $(A_1 \wedge A_2 \wedge \dots \wedge A_k) \Rightarrow B$ 。[] 在推理过程中主要通过以下四种推理规则和其它

一些等价定律完成推理：前提引入规则、结论引入规则、置换规则、假言推理规则。推理过程就是反复使用谓词演算的推理规则和基本等价定律，对已知谓词公式进行变换，以得到所需逻辑结论的过程。[3]

##### (3) Horn 子句

定义不含任何连接词和量词的谓词公式是原子公式，如  $Computer(x)$ 、 $hasMom(x,y)$  都是原子公式。一些原子公式的析取（或）式称为子句，如  $\neg P(x) \vee Q(x)$  就是子句。这

样我们称最多含一个非否定原子公式的谓词子句为Horn 子句, 其一般形式为:

$$\neg A_1 \vee \neg A_2 \vee \dots \vee \neg A_n \vee B$$

此式等价于

$$A_1 \wedge A_2 \wedge \dots \wedge A_n \rightarrow B$$

式中,  $B$  均为谓词子句,  $i \in 1, 2, \dots, n$

可见, Horn 子句是包含若干个条件(前提)子句, 仅含一个结论子句的谓词公式。由于运用 Horn 子句表达知识较方便, 推力比较简单, 已成为人工智能问题求解的重要工具。

SWRL 语言描述的规则就采用类似Horn 子句的形式来描述。

### 3.3 描述逻辑

#### 3.3.1 描述逻辑概述

描述逻辑(DL, Description Logics)是一种基于对象的知识表示的形式化工具, 是一阶谓词逻辑的一个可判定的子集, 能够提供可判定的推理服务, 并且具有语义特征。从某种意义上说, 描述逻辑是语义 Web 的逻辑基础。描述逻辑建立在概念(concept)和角色(role 即二元关系)之上, 由构造算子(constructor)从简单概念和角色中构造出复杂概念和角色。

一个描述逻辑表示知识库主要由 TBox 和 ABox 两部分组成。TBox 定义了特定知识领域的结构并包含一系列公理, 可以通过已有概念构成新的概念。ABox 包含了 TBox 中概念的实例。TBox 具有分类(Taxonomy)的能力, 分类是系统化建立知识的第一步, 通过分类的动作可以将事物的本体集成成共通的概念, 这些概念又可组合成更广泛的概念。ABox 是实例断言的集合, 用于指明个体的属性或个体之间的关系。由概念断言和关系断言组成。

描述逻辑作为一种知识表示方法, 它不但可以表示领域中的知识, 也具有推理机制, 能够推导出隐含的知识。在描述逻辑中的推理分成概念推理和实例检测两种。概念推理包括判定概念的一致性(可满足性)、包含关系、等价关系和不相交性等; 实例检测指判定一个给定的个体是否是一个给定概念的实例。

在描述逻辑推理应用方面, 现已有 Racer、FaCT++等实用的 DL 推理机, DL 推

理机可以推导出两个概念之间的包含关系, 并完成一致性检验。

#### 3.3.2 OWL 与描述逻辑的关系

采用描述逻辑作为知识的表示方式是 Tbox 与 Abox 的架构, 将概念与实例分开表示, 描述逻辑语法若采用基于 XML 的表示方式便是 OWL 的语法, 通过 OWL 语法可以将描述逻辑的关系体现出来。而本体表示的主要目的就是希望机器能够理解其所描述的资源含义, XML 具有人机可读的特点, 语义 Web 是希望进一步实现人机可理解, 达到在语义层次上的理解。因为具有人机可理解的关系和层次架构, OWL 便可实现以类为基础的推理, 从已分类的知识推导出未分类的隐含知识。

OWL 是以描述逻辑为逻辑基础的本体语言, 也就意味着 OWL 中的类构造算子(类约束)及公理具有与描述逻辑相应地表示。

#### 3.3.3 逻辑到一阶谓词逻辑的映射

采用描述逻辑表述的知识完全可以映射到一阶谓词逻辑上, 而 OWL 描述的本体又与描述逻辑知识库等价, 因此就在理论上证明了将 OWL 本体映射到一阶谓词逻辑上的可行性。

因此可以将采用 OWL 表示的本体和采用 SWRL/SWRL FOL 描述的规则无损的映射到一阶谓词逻辑上, 通过一阶谓词逻辑推理系统进行推理。这就为本研究的基于 OWL 本体和 SWRL/SWRL FOL 规则的一阶谓词逻辑推理平台的可行性提供了理论支持。

### 3.4 OWL 本体和 SWRL/SWRL FOL 规则的推理平台

以上从理论上分析了将 OWL 本体和 SWRL/SWRL FOL 规则结合, 通过映射为一阶谓词逻辑公式进行推理的可行性。在此基础上提出基于 OWL 本体和 SWRL/SWRL FOL 规则的推理平台。

#### 3.4.1 平台框架

如图 2 所示给出了基于 OWL 本体和 SWRL/SWRL FOL 规则的一阶谓词逻辑推理平台框架。可见该平台主要由四部分组成: Ontology、SWRL/SWRL FOL、一阶谓词逻辑映射和 FOL 推理引擎。

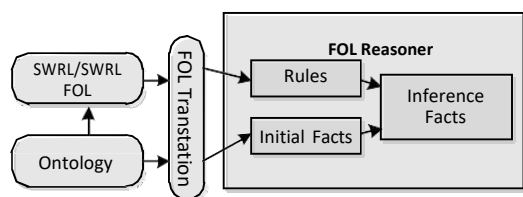


图2 基于本体和规则的 FOL 推理平台

根据对本体论的分析知道本体在知识呈现方面具有非常大的优势,已成为语义 Web 的重要组成部分,因此采用 OWL 描述的本体对推理平台中涉及到的领域知识进行建模,定义领域中的知识概念,并将概念与概念之间的关系清晰的表示出来,使知识推理更加灵活。当用户查询知识本体时,能够根据本体中定义的概念与概念之间的关系返回与用户需求真正语义相关的结果。

### 3.4.2 OWL 本体的建立

本体作为一种有效表现概念层次结构和语义的模型,目标是捕获相关领域的知识,提供对该领域知识的共同理解,确定该领域内共同认可的词汇,并从不同层次的形式化模式上给出这些词汇和词汇间相互关系的明确定义。从而为本领域中的应用服务和共享。

在建立本体过程中,主要考虑的元素有类(概念)、属性、关系及实例。

### 3.4.3 SWRL/SWRL FOL 规则的建立

OWL 具有强大的知识描述能力,同过 OWL 可以提供定义类、属性限制和描述逻辑的特点,但 OWL 所拥有的描述能力主要来自于以类为基础加上关系的推理,如果某项知识难以用类的方式表达,则 OWL 就难以描述。例如 If .... Then.... 关系是很普通的逻辑判定,但是 OWL 以规则为基础的描述和推理能力并不足,因此必须添加用户自定义的规则才能使推理能力较为完整。而传统的规则语法无法和知识呈现方式结合,在知识呈现方面所制订的词汇和类均无法在规则中再利用;虽然在规则中必须使用相同的词汇才能运作,但目前缺乏将知识库中的内容直接对应到规则中,进行规则的运算。为了解决上述问题,为 OWL 本体扩充逻辑规则,采用 SWRL(Semantic Web Rule Language)来描述用户自定义规则。SWRL

是建构在 OWL 之上,可以和 OWL 定义的和属性互相整合,在描述逻辑的基础上进行规则运算。另外 SWRL 以 XML 为实体语法,具有更高的跨平台、分享和再利用等特性。

SWRL 描述的规则形如(采用 human readable 语法表示):

前提(Body)→结论(Head)

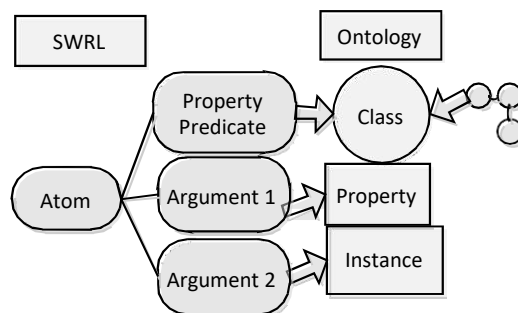
具体体现为:

$$B_1, B_2, \dots, B_m \rightarrow H_1, H_2, \dots, H_n$$

其中  $B_i$  和  $H_j$  都是规则中的 Atom,可见 SWRL 规则是由规则原子组成的。规则原子主要有如下几种类型,如表 1 所示。WRL 与 SOWL 本体的关系如图 3 所示。

表1 Atom 的基本类型

Atom	Type
C(i)	Class Atom
R(i,j)	Property Atom
$i=j$	Equality Atom
$i \neq j$	Inequality Atom

图3 SWRL 与 OWL 本体的关系<sup>[2]</sup>

SWRL FOL 是在 SWRL 的基础上,向一阶谓词逻辑公式扩展。因为 SWRL 描述的规则是类似于一阶谓词逻辑中的 Horn 子句,每一个 Atom 都是原子公式,因此在 SWRL 中无法显式的反映存在量词( $\exists$ ),非( $\neg$ )以及规则嵌套等可以用 FOL 表示的更加复杂的规则关系。这就需要 SWRL FOL 来补充。

### 3.4.4 本体和规则的 FOL 映射

OWL 描述的本体和 SWRL/SWRL FOL 表示的规则建立好之后,现有的本体推理引

擎并不能实现结合二者的推理,因此根据第三章的理论分析,将它们都转换为一阶谓词逻辑公式,利用现存的一些功能强大的 FOL 理论证明机来实现推理。在这个过程中主要涉及两个方面,即 OWL 本体的 FOL 映射和 SWRL 规则的 FOL 映射。

(1) OWL 本体向 FOL 映射,需要将描述逻辑映射为一阶谓词逻辑,转化步骤分为:  
C, 即 OWL 本体中的类映射为 FOL 中的一目谓词:  $C(x)$ ,  $x$  为实例变量;

P, 即 OWL 本体中的属性映射为 FOL 中的二目谓词:  $P(x,y)$ ,  $x, y$  为实例变量;

I, 即 OWL 中的实例映射为 FOL 中的实例常量  $i$ ;

$a:C$ , 即 OWL 本体中  $a$  是类  $C$  的一个实例映射为:  $C(a)$ ,  $a$  为实例常量;

$a:P$ , 即 OWL 本体中  $a$  的  $P$  属性的值为  $b$  映射为:  $P(a,b)$ ,  $a, b$  为实例常量;

(2) SWRL 规则向 FOL 映射,只需把 SWRL 规则中类似 Horn 子句转换为 FOL 谓词公式,即将类似 Horn 子句表示的规则中所有的自由变量转换为全局变量,将全称量词(显式的表现出来。形式上 SWRL 规则:

$$B_1, B_2, \dots, B_n \rightarrow H_1, H_2, \dots, H_n$$

被转化为 FOL 谓词公式:

$$\forall x_1, x_2, \dots, x_i T(B_1) \wedge T(B_2) \wedge \dots \wedge T(B_m) \rightarrow T(H_1) \wedge T(H_2) \wedge \dots \wedge T(H_n)$$

其中  $x_1, x_2, \dots, x_k$  是所有出现在规则 Atom  $B_i$  和  $H_j$  中的变量,  $T$  表示将每一个 Atom 转换为相应地 FOL 原子公式。具体地,将 SWRL 规则:

$$visited(u,r) \Rightarrow Acquired(r,u)$$

映射到 FOL 后得到谓词公式:

$$\forall u, r visited(u,r) \Rightarrow Acquired(r,u)$$

SWRL FOL 由于本身就以 FOL 谓词公式的形式来描述规则,因此在逻辑上无需转化。经过上述的映射原则,我们便可以通过一些语法解析器(Parser)将 OWL 本体和 SWRL/SWRL FOL 描述的规则的 XML 实体语法转换成 FOL 推理机可以识别的方式,进行推理。

OWL 本体和 SWRL/SWRL FOL 规则都

建立好之后,通过映射为一阶谓词逻辑公式,就可以利用 FOL 推理引擎进行推理了。

## 4 总结与展望

文章重点介绍了语义知识和推理以及本体推理的方法和原则。并介绍了一阶逻辑和描述逻辑,主要介绍了描述逻辑在本体推理上的应用,并将描述逻辑和 OWL 的关系阐述的较为清楚。由于在研究生阶段,开设有数理逻辑课程,因此一阶逻辑和描述逻辑的具体知识没有详细介绍。本体作为语义网的一个重要部分,非常值得我们去探究。目前研究本体的方面有很多,语义 Web 构建的核心技术之一是创建具有统一规范的领域本体,如何维护本体的正确性及有效的组织推理规则获取隐性信息是其中的主要工作,基于此点,本体推理研究在语义 Web 技术中具有重要意义<sup>[2]</sup>。作为一种富含逻辑信息的知识表示语言,描述逻辑<sup>[4]</sup>主要用于表征事物及它们之间的关联属性,对特定应用领域,它首先定义概念,然后用概念去表征关系和属性,最终实现对个体和对象进行表达的目的。它不同于其它的知识表示方式

(例如面向对象的方法),有着形式化的、基于逻辑的语义规范,并且概念语言推理可以从显式表达推出隐含的内容。因此描述逻辑在语义网本体推理上的应用非常值得研究。

由于时间和精力有限,本文只是接触与前任研究的成果加上个人的思考简单的阐述了本体推理的理论和基本的框架架构,并没有具体的去实现应用。由于互联网的发展已经不能满足人们对于知识检索和分类的要求,语义网的研究会越来越受到重视。本体推理作为语义网研究的一个重要方面在未来相信会取得很大的成果。

## 参考文献

- [1] 宋炜, 张铭. 语义网简明教程[M]. 高等教育出版社, 2004.
- [2] 潘超, 谷辉. 本体推理机及其应用[J]. 计算机系统应用, 2010, 19(2): 3—5.
- [3] 董美霞. 基于语义网模糊本体的知识推理研究[D]. 大连: 大连海事大学信息科学技术学院, 2011

董美霞。基于语义网模糊本体的知识推理研究

- [4] 孔为民, 涂中群. 语义网的技术及其应用[J]农业图书情报学刊, 2010, 01(3):1
- [5] 刘静, 胡越黎. 基于语义网络表示的皮肤症状模糊推理系统的研究[J]. 应用科学学报, 2006, 9(5):2.