

知识服务的 语义匹配机制研究

ZHISHI FUWU DE YUYI PIPEI JIZHI YANJIU

黄 涛 / 著

知识服务的语义匹配机制研究

黄 涛 著

华中师范大学出版社

新出图证(鄂)字 10 号

图书在版编目(CIP)数据

知识服务的语义匹配机制研究/黄涛 著. —武汉: 华中师范大学出版社, 2015. 3

ISBN 978-7-5622-6933-5

I. ①知… II. ①黄… III. ①检索语言—研究 IV. ①G254.0

中国版本图书馆 CIP 数据核字(2015)第 071861 号

知识服务的语义匹配机制研究

© 黄涛 著

责任编辑: 陈 艳 袁正科

责任校对: 刘 峥

封面设计: 胡 灿

编辑室: 第二编辑室

电 话: 027-67867362

出版发行: 华中师范大学出版社

社 址: 湖北省武汉市珞喻路 152 号

邮 编: 430079

销售电话: 027-67863426/67863280

邮购电话: 027-67861321

传 真: 027-67863291

网 址: <http://www.ccupress.com>

电子信箱: hscbs@public.wh.hb.cn

印 刷: 虎彩印艺股份有限公司

督 印: 章光琼

开 本: 710mm×1000mm 1/16

印 张: 9 字 数: 152 千字

版 次: 2015 年 4 月第 1 版

印 次: 2015 年 4 月第 1 次印刷

定 价: 20.00 元

敬告读者: 欢迎举报盗版, 请拨打举报电话 027-67861321

前 言

本体作为一种语义和知识层面上的概念共享模型，自提出以来就引起数据整合、P2P 系统、电子商务、语义 Web 服务、社会网络等应用领域科研人员的广泛关注，并得到了有效的研究与应用，一系列基于本体领域问题的解决方法也应运而生。

本体匹配是发现不同本体之间实体元素(包括本体的类、属性或者个体)映射关系的关键技术，已被国内外学者普遍认为是解决计算机系统语义异构问题的有效手段之一。

知识服务中的语义知识检索问题实际上是要解决用户检索条件和被检索的资源描述文本之间在语义级别上的匹配问题。最好的途径之一是将用户检索条件和资源描述文件都表示成可以被机器所理解的形式化知识表示下的语义信息，然后通过严格的逻辑推理和演算来评判两者之间的匹配程度。将查询条件和资源描述文本之间的相似度计算出来，按照相似度值将候选资源排序，最后将排序靠前的资源返回给用户。由于本体可以作为用户检索条件与资源描述文件的语义描述表现方式，而本体匹配可以解决本体的语义匹配程度问题，所以本书试图将本体匹配技术作为知识服务以及知识检索问题的解决方案之一，尝试通过对本体匹配技术的研究来解决知识服务中的语义匹配问题。

本书的研究主要包括以下几个方面：

第 3 章介绍教育领域学习资源本体建模与描述。针对目前学习资源元数据标准在表示学习资源时语义缺失的现状，对学习资源元数据以及学习资源元数据与本体结合的相关研究进行了调研，在 Dragan Gašević、Marek Hatala 等人提出的学习资源元数据与学习内容领域本体层次模型的基础上进行扩展。结合元数据本体和领域本体，形成了学习资源本体，并对元数据本体和领域本体进行了实例化。

第 4 章介绍基于上下文的元素层次本体匹配算法。在对分析独立元素和分

析元素本体结构这两种元素层次匹配方法进行归纳,对基于 WordNet 和文本集合概率的语义相似度算法和基于元素本体结构的语义相似度算法这两种语义相似度比较方法进行分析的基础上,对现有的语义相似度算法进行了拓展,提出改进 Hirst & St-Onge 语义相似度算法的元素独立语义相似性计算方法,并根据元素的上下文路径,进一步提出了基于上下文的元素层次本体匹配方法。该方法从元素以及元素上下文要素出发,将元素的前驱元素、属性和元素的后继元素看成元素的上下文,并组成元素的上下文路径,针对元素的上下文路径,在分别计算前驱元素、属性、元素本身以及后继元素独立语义相似性的基础上,进行组合加权,得到基于上下文的元素语义相似性。

第5章介绍基于加权本体的结构层次本体匹配算法。在对各种基于图形结构的匹配算法研究的基础上,分析了加权本体结构图形的形式化表示方法。同时,还在此基础上研究了基于加权本体的结构层次匹配算法。该算法采用自顶向下、分层加权的思想,根据学习资源本体的特点,给不同的层次实体分配不同的权重系数,将加权本体结构中所有实体元素结合在一起,计算实体绑定后的本体结构语义相似性,以此来完成加权本体的结构层次匹配。结构层次的匹配研究以基于上下文的元素层次的匹配算法为依据,通过二维矩阵进行存储和重用,在得到单个元素和属性的语义相似性的基础上,研究多个实体结合在一起的学习资源本体结构语义匹配。

第6章介绍本体匹配算法在教育知识语义检索系统中的应用与实现。将本体匹配技术集成到教育知识语义检索中,设计并开发教育知识语义检索原型系统来验证本体匹配算法在语义知识检索实践中的正确性与可行性。

第7章介绍基于知识地图的资源导航与检索系统。将“知识地图”应用在资源的组织与导航中,学习者除了能以知识地图为线索查找学习资源外,还能享有系统根据学习者学习情况智能推荐的相关学习资源。

本书的第1章到第6章是本人博士论文《知识服务语义匹配机制》的研究工作,第7章由杨华利老师编写,是杨华利老师论文《基于知识地图的资源检索的设计与实现》的研究成果,在此,对她表示感谢!

黄 涛

2015年1月于桂子山

目 录

第 1 章 知识服务语义匹配机制研究概述	1
1.1 研究背景与意义	1
1.2 研究内容与本书组织	4
1.3 本章小结	7
第 2 章 本体及本体语义匹配研究基础	8
2.1 语义 Web 与本体	8
2.2 本体匹配技术	15
2.3 本章小结	26
第 3 章 学习资源本体建模与语义表示	27
3.1 相关研究工作	28
3.2 学习资源本体	34
3.3 学习资源本体的语义表示	37
3.4 本章小结	38
第 4 章 基于上下文的元素层次本体匹配算法	39
4.1 相关研究工作	39
4.2 语义相似度	41
4.3 基于上下文的元素层次语义匹配算法	49
4.4 实验结果与分析	54
4.5 本章小结	59
第 5 章 基于加权本体的结构层次本体匹配算法	60
5.1 相关研究工作	60
5.2 加权本体(Weighted Ontology)的形式化表示	64
5.3 基于加权本体的结构层次匹配方法	66
5.4 实验结果与分析	76

5.5 本章小结	80
第6章 本体匹配算法在教育知识语义检索中的应用与实现	81
6.1 相关研究工作	81
6.2 教育知识语义检索系统	88
6.3 原型系统实现	93
6.4 实验结果与分析	94
6.5 本章小结	96
第7章 基于知识地图的资源检索与导航系统的设计与实现	98
7.1 相关研究工作	99
7.2 基于知识地图的资源导航系统的设计	105
7.3 基于知识地图的资源导航系统的实现与测试	112
7.4 本章小结	119
第8章 总结与展望	120
本章小结	122
参考文献	123

第1章 知识服务语义匹配机制研究概述

1.1 研究背景与意义

1.1.1 知识服务的提出

信息化是当今世界发展的大趋势，是我国经济加快发展和社会全面进步的重大战略机遇。“以信息化带动工业化，以工业化促进信息化”是全面建设小康社会和构建社会主义和谐社会的重要举措。教育信息化是国民经济和社会信息化的重要组成部分。以教育信息化为龙头，带动教育现代化，实现教育跨越式发展已成为我国教育事业发展的战略选择。

发达国家高度重视教育信息化发展，纷纷制订国家战略和行动计划，确保经费投入，试图在未来信息社会中让教育走在社会发展的前列。十六大明确提出新时期教育的任务是加强人才培养，造就数以亿计的高素质劳动者、数以千万计的专门人才和一大批拔尖创新人才；形成比较完善的现代国民教育体系，构建终身教育体系；形成全民学习、终身学习的学习型社会，促进人的全面发展，为全面建设小康社会提供高素质的人力资源。

在我国目前的教育信息化中还有许许多多亟待解决的关键问题，如：

- (1)资源整合不力，难以共享；
- (2)高等教育规模与质量矛盾突出；
- (3)农村教育基础薄弱，已成为制约全民族素质提高的关键；
- (4)在职教育、培训和终身学习需要重视和加强；
- (5)如何降低成本，保证学习的公平性。

此外还有开放的知识服务体系尚未形成等等问题。

上述问题解决的有效途径之一是：必须以大力推进数字资源整合、应用和服务，发展教育信息化为主线，集成和整合各种信息资源、各种网络设施，提

升资源到知识层面,建立基于内容的传输、存储和使用体系与结构,为用户提供开放、及时、准确、便捷、主动、智能、低成本、无处不在的知识服务体系。开放的、及时的、准确的、便捷的、主动的、智能的、低成本的、无处不在的知识服务对于建立现代教育服务体系,形成“全民学习、终身学习”的学习型社会具有重要的意义。

简单地讲,知识服务就是基于知识提供的服务。它关系到两个方面:知识相关的问题和服务相关的问题。知识相关的问题包括知识的提取(Extraction)、描述(Description)、组织(Organization)、集成(Integration)、融合(Fusion)、检索(Retrieve)等,处理知识相关问题的目标是为知识服务提供所需要的、适当和准确的知识。服务相关的问题包括基于知识的各种表现形式,以自动化的方式在网络上满足用户的个性化需求。

1.1.2 知识服务面临的问题

随着互联网(World Wide Web)在人们日常生活中的广泛应用,当前知识服务的运行机制主要是通过网络来获取用户的个性化需求,然后通过知识服务系统的运作,在互联网上以个性化资源推送或者资源检索的方式来满足用户的需求。众所周知,互联网的应用中,搜索引擎的使用占据了很大的份额,可以说没有搜索引擎(如 Google、百度、Yahoo 等)的成功应用,就没有今天互联网的巨大成功。因此,知识检索技术是基于互联网技术的知识服务的核心技术之一。

由于现有的网络资源大多是面向用户理解设计的,所以资源并不能够被计算机理解,知识服务系统主要是基于传统信息资源的 Web 服务,知识服务中的知识检索,无论是基于元数据,还是基于全文文本检索,归根到底都是基于关键字的检索,这种检索方式容易产生以下问题:

(1)高召回率(Recall),低准确率(Precision)。虽然返回大量含有输入关键字的检索结果,但这些结果中与用户输入请求语义相关的结果很少。

(2)低召回率。一旦关键字匹配失效,用户就检索不到任何结果,这样就会丢失与关键字语义相关的大量信息。

(3)检索结果基于用户输入关键字的词义而不是语义,现有的检索原理是基于关键字名称的词义匹配,导致与关键字具有相同意义或者相近意义的结果检索不到。

造成上述问题的主要原因是网络知识资源格式的异构性、知识资源语义的

多重性以及知识资源语义关系的缺乏性。要真正解决这些问题,就要求一个知识检索系统不仅仅依靠关键词义匹配技术,而是能在理解知识的语义内涵的基础上进行语义匹配,实现知识资源的语义检索和知识服务的智能化。

知识资源的语义检索强调的是“语义”。“语义”(Semantics)和“词义”(Syntax)是一个相对的概念,知识的“语义”表示的是知识所蕴藏的意义,即除了文本字面含义以外的内在含义,知识的“词义”表示的是文本字面上的含义。现在要获取知识的语义,往往需要人工的参与,通过人的学习和联想,才能够发掘知识在字面之外的其他意义(即语义)。但是这种人工获取知识语义的一个最大弊端是不可想象的“人力开销”。网络上知识资源的数量是巨大的,如果不能实现知识语义检索的自动化或者半自动化,而只依赖人的手动检索,这种“人力开销”将是不可想象的。因此,需要实现知识语义检索的自动化(半自动化)。目前存在的问题是现有的知识检索虽然可以实现自动化(关键字匹配),但并不智能,达不到实现知识语义检索的目的。

要实现计算机自动(半自动)的知识语义检索,首先要解决的一个问题是使计算机能够“理解”知识。解决这个问题最好的方法之一是对知识进行语义描述,使互联网变成语义互联网(语义 Web),这就是互联网创始人 Tim Berners Lee 提出语义 Web 概念的原因。本体(Ontology)作为语义 Web 的核心概念之一,自提出以来就引起了国外众多科研人员的关注。本体的主要作用是采用共享和抽象的方法,对各个领域普遍用到的知识进行建模,并且采用计算机可以“理解”其描述标签的知识描述形式,使本体成为知识描述的载体,这是解决目前知识语义检索问题的最好途径之一。

1.1.3 研究思路

本体作为一种语义和知识层面上的概念共享模型,自提出以来就引起数据整合、P2P 系统、电子商务、语义 Web 服务、社会网络(Social Networks)等应用领域科研人员的广泛关注,并得到了有效的研究与应用,一系列基于本体的领域问题解决方法也应运而生。

本体匹配(Ontology Matching)是发现不同本体之间实体元素(包括本体的类、属性或者个体)映射关系的关键技术,已被国内外学者普遍认为是解决计算机系统语义异构问题的有效手段之一。

知识服务中的语义知识检索问题实际上是要解决用户检索条件和被检索的

资源描述文本之间在语义级别上的匹配问题。最好的途径之一是将用户检索条件和资源描述文件都表示成可以被机器所理解的形式化知识表示下的语义信息,然后通过严格的逻辑推理和演算来评判两者之间的匹配程度。将查询条件和资源描述文本之间的相似度计算出来,然后按照相似度值将候选资源排序,最后将排序靠前的资源提交给用户。由于本体可以作为用户检索条件与资源描述文件的语义描述表现方式,而本体匹配可以解决本体的语义匹配程度问题,所以作者试图将本体匹配技术作为知识服务以及知识检索问题的解决方案之一,尝试通过对本体匹配技术的研究来解决知识服务中的语义匹配问题。

1.2 研究内容与本书组织

针对上述问题,研究知识服务中的语义匹配机制,在借鉴现有各种本体匹配算法与本体匹配系统优点的基础上,提出一种面向知识检索的多层次本体匹配算法。本书的研究围绕以下四个方面展开:

(1)领域本体建模与表示,即实现某一领域本体(以教育技术学为研究背景,本书研究教育领域学习资源本体),体现知识的语义表达,为后面的工作打下基础;

(2)基于上下文的元素层次的本体匹配算法,在传统语义相似度算法的基础上,提出了基于上下文的元素层次本体匹配算法;

(3)基于加权本体的结构层次的本体匹配算法,采用自顶向下、分层加权的思想,针对学习资源本体的特点,研究加权本体结构中所有实体元素绑定后的本体结构语义相似性计算方法;

(4)本体匹配算法在教育知识语义检索中的应用与实现,通过设计与开发基于本体匹配的教育知识语义检索系统 EKSR,并将其与关键字匹配检索系统进行实验比较,来验证和测试本书研究的本体匹配算法的有效性与可行性。

本书后续章节的组织如图 1-1 所示,首先对本体语义匹配技术的发展概况进行综述(第 2 章),然后从学习对象元数据标准出发,考虑学习资源本体的建模与描述问题(第 3 章)。接下来,从本体匹配的粒度出发,分别考虑基于上下文的元素层次本体匹配算法(第 4 章)和基于加权本体的结构层次本体匹配算法(第 5 章)。最后,针对知识服务中教育知识检索的需求,考虑本体匹配算法在知识检索中的应用与实践验证问题,并设计与实现教育知识的语义检索系统(第 6 章)。

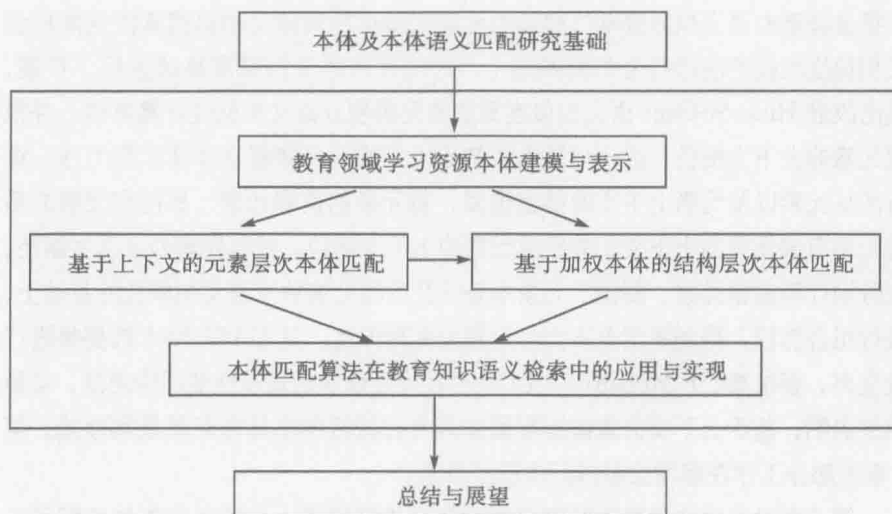


图 1-1 研究内容的组织结构

本书的主要工作体现在第3~6章，这4章的内容介绍如下：

第3章对教育领域学习资源本体建模进行分析和讨论。本章通过在建立传统本体的基础上进行扩展，针对目前学习资源元数据标准在表示学习资源时语义缺失的现状，对学习资源元数据以及学习资源元数据与本体结合的相关研究进行了调研，在 Dragan Gašević、Marek Hatala 等人提出的学习资源元数据与学习内容领域本体层次模型的基础上进行扩展。结合元数据本体和领域本体，形成了学习资源本体，并对元数据本体和领域本体进行了实例化，分别以 IEEE LOM 元数据本体和 ACM CSS 计算机学科领域本体作为具体的元数据本体和领域本体，结合本体建模原则，将二者有机地组合在一起形成面向 ACM CSS 计算机学科领域分类系统的学习资源本体。与现有的学习资源本体相比，本书提出的本体模型更加全面和具体，而且主要研究面向学习资源的语义检索。学习资源本体模型的提出将为本书后续章节本体匹配算法的研究以及本体匹配算法应用系统的实现提供理论基础。第3章的部分工作在参考文献[13]、[14]中已经发表。

第4章对本体元素层次的本体匹配算法进行研究。本章在对分析独立元素和分析元素本体结构这两种元素层次匹配方法进行归纳，对基于 WordNet 和文

本集合概率的语义相似度算法和基于元素本体结构的语义相似度算法这两种语义相似度比较方法进行分析的基础上,对现有的语义相似度算法进行了扩展,提出改进 Hirst-St-Onge 语义相似度算法的元素独立语义相似性计算方法,并根据元素的上下文路径,进一步提出了基于上下文的元素层次本体匹配方法。该方法从元素以及元素上下文的要素出发,将元素的前驱元素、属性和元素的后继元素看成元素的上下文,并组成元素的上下文路径,针对元素的上下文路径,在分别计算前驱元素、属性、元素本身以及后继元素独立语义相似性的基础上,进行组合加权,得到基于上下文的元素语义相似性。对 OAEI 2006 数据集进行查全率、查准率、F-Measure 以及 OverAll 等指标的质量和性能评价测试,实验结果表明,基于上下文的本体匹配算法具有比较好的本体匹配质量和性能。第 4 章的部分工作在参考文献[15]中已经发表。

第 5 章对本体结构层次的语义匹配算法进行研究。结构层次本体匹配可以看成对两个本体图形结构的匹配操作,产生两个图形元素之间、结构之间的语义映射结果。本章在对各种基于图形结构的匹配算法研究的基础上,分析了加权本体结构图形的形式化表示方法。同时,本章还在此基础上研究基于加权本体的结构层次匹配算法,该算法采用自顶向下、分层加权的思想,根据学习资源本体的特点,给不同的层次实体分配不同的权重系数。将加权本体结构中所有实体元素结合在一起,计算实体绑定后的本体结构语义相似性,以此来完成加权本体的结构层次匹配。结构层次的匹配研究以上一章基于上下文的元素层次的匹配算法为依据,通过二维矩阵进行存储和重用,在得到单个元素和属性的语义相似性的基础上,研究多个实体结合在一起的学习资源本体结构语义匹配。第 5 章的部分工作在参考文献[16]、[17]中已经发表。

第 6 章对本体匹配算法在教育知识语义检索中的应用与实现进行研究。本章将综合前面第 3、4、5 章的研究内容,将本体匹配技术集成到教育知识语义检索中,设计并实现教育知识语义检索原型系统来验证本体匹配算法在语义知识检索实践中的正确性与可行性。为了验证教育知识语义检索系统的效率,本章还将教育知识语义检索系统与我们开发的基于传统关键字检索的教育资源检索系统进行了检索实验结果的对比与分析,以此来验证该系统的有效性。第 6 章的部分工作在参考文献[18]中已经发表。

第 7 章将本体匹配算法运用在基于知识地图的数字资源检索与导航系统中,

研究知识地图的可视化技术、知识本体与资源实体的映射技术及个性化资源推荐技术,有效解决了当前在线资源服务系统中资源组织混乱的问题,提高了资源检索及用户学习的效率。

1.3 本章小结

本章首先介绍了本书的背景与意义,以及本书的研究内容和内容结构。本体作为一种能在语义和知识层次上描述信息系统的概念模型和建模工具,在匹配过程中导入概念的内容表述语义,使匹配过程由原来的关键词匹配进化为内容匹配,使知识服务向语义化、智能化方向发展。

第2章 本体及本体语义匹配研究基础

2.1 语义 Web 与本体

2.1.1 语义 Web

互联网(World Wide Web)的出现彻底改变了人们传统的生活方式,它将计算机从只能进行数学计算的计算器变成人们日常生活中必不可少的信息处理与交流的工具。现在互联网上的信息资源主要是通过结合自然语言和非结构化信息载体(如多媒体、文本、图像、声音等)进行表示,人们可以很好地理解并进行交流,但是存在的一个问题是计算机不能理解网络上的信息资源表达的“语义”。

为解决现有互联网的知识表示不能被机器理解的“语义”缺失问题, Tim Berners Lee 首先提出了语义 Web(Semantic Web)概念,并于 2000 年给出了如图 2-1 所示的语义 Web 层次“蛋糕”(Layer Cake),定义和描述了语义 Web 的主要层次。它的主要思想是基于 XML(XML Schema)和 RDF(RDF Schema),利用语义 Web 的核心——本体(Ontology)所定义的共享词汇(Ontology Vocabulary),来对互联网上的信息进行描述,并表示成机器可以理解和处理的形式。然后在机器可以理解的知识表示方式的基础上,进行语义 Web 层次结构中逻辑推理、推理证明、处理过程可信任等语义层面的智能操作。

在上述分类层次的最底部是互联网资源的 Unicode 编码和资源的全球统一标识符(Uniform Resource Identifiers, URI),这一层可以解决互联网上资源的定位和跨地区字符编码的标准格式问题。

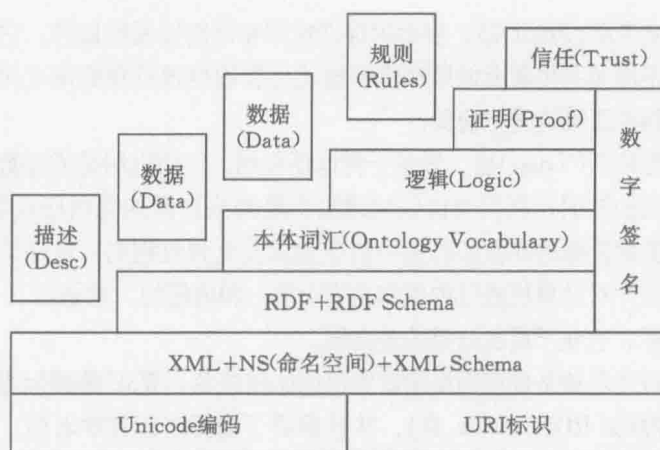


图 2-1 语义 Web 的层次分类方法(翻译自 Tim Berners Lee 层次模型)

第二层是 XML、NS(命名空间)与 XML Schema 层,这一层的目的是使用用户自己定义的词汇来描述非结构化的互联网资源,使资源结构化后,在分布式网络上自动传输。

第三层是 RDF 和 RDF Schema 层,这一层在第二层的基础上,加入了资源描述框架(Resource Description Framework)的语法规则。RDF 是一个类似于实体关系模型的基本数据模型,基于 XML 形式的词义,利用主语、谓语、宾语(三元组)的陈述语句方式来表示网络资源。RDF Schema 是抽象的 RDF 数据层次模式,它主要包含数据模型的类(Class)、属性(Property)、子类(Subclass)关系、子属性(SubProperty)关系以及领域(Domain)和值域(Range)限制。

第四层是本体词汇(Ontology Vocabulary)层,虽然第三层的 RDFS(RDF Schema)可以简单地表示本体(Ontology),但是从第三层的分析可以看出,RDFS 所表示的数据模式关系(如子类、子属性、领域、值域等)还不能包含领域的所有关系,因此需要一种包含更多语义关系的本体描述语言来对 RDFS 进行扩展。目前用得比较多的扩展语言是 W3C 提出的 OWL(Web Ontology Language),它在 RDFS 的基础上进行延伸,能够表示资源之间更多的语义关系。

第五层是逻辑(Logic)层,它是本体词汇层的延伸,针对具体的应用领域,利用本体描述语言来对领域知识进行表示。用户可以基于一定的规则(Rules),在本体描述的基础上进行逻辑推理,并验证推理的正确性和可信任性。

第六层是证明(Proof)层,表示对推理规则和评判过程的信赖,它包含推理过程的真实还原和利用前面低层次资源描述语言的推理过程的形式化表示,并且对推理过程的正确性进行检验。

第七层是信任(Trust)层,表示一种信任机制,信任机制是通过数字签名技术和其他相关的知识,利用可信任代理来实现网络资源的可信任获取与传输,它包括用户提供资源的安全和质量可信任性以及处理过程的可信任性。一种理想的状态是用户对计算机进行的操作是完全放心和信任的,这是语义 Web 的最高层次和目标,它处于层次分类的最高层。

目前的研究现状是前面四层的研究已经比较成熟,W3C 等国际组织提出了一系列的标准(如 RDF、OWL 等),并且得到了国际上研究学者的广泛研究和关注。第五层逻辑的研究是目前国内外研究的热点问题,其中一个引起广泛关注的议题是描述逻辑(Description Logic)。国内外对第六层、第七层的研究目前才处于起步阶段,相对而言还不太成熟,还有待国内外学者的进一步研究。

通过上面的分析可以看出,语义 Web 是当前互联网的延伸与扩展,能够促进人与计算机之间的语义互操作。语义 Web 研究的起点和核心问题就是本体(Ontology),它是语义 Web 研究的落脚点和研究基础,只有先弄清本体的概念及相关知识,才能在本体的基础上进行扩展和进一步探讨。

2.1.2 本体(Ontology)

本体(Ontology)这个概念来源于哲学界,它是一种存在论,是哲学家认识世界的一种思维方式。人工智能界引入这个概念,并赋予其新的含义。本体的定义比较多,被学术界广泛认可的定义是 Gruber 在 1993 年提出的“本体是概念模型的明确的规范说明”和 Studer 在 Gruber 本体定义的基础上,于 1998 年提出的“本体是共享概念模型的明确的形式化规范说明”。本体包含了四个方面的含义:共享(Share)、概念模型(Conceptualization)、明确(Explicit)、形式化(Formal)。

(1)共享:表示本体是某领域的公认的概念集,所反映的是大家共同认可的知识。

(2)概念模型:对客观世界物体的一种抽象模型和概念化表示。

(3)明确:抽象出来的概念及概念之间的关系都有显性和明确的定义。

(4)形式化:本体的表示是计算机可以理解的形式,是概念精确的数学表达。