维普资讯 http://www.cqvip.com

基于本体语言 OWL 的模糊扩展*)

赵德新[1,2]　冯志勇[1]

(天津大学计算机科学与技术学院　天津 300072)[1]　(天津理工大学计算机科学与技术学院　天津 300191)[2]

摘　要　本体能够对特定领域的概念、术语以及关系提供一种形式化的描述方法。尽管本体在知识表示上有很强的能力,但是有一个缺陷,即不能表达不确定和不精确的信息。而这些信息在语义网和多媒体应用中,又是至关重要的。针对模糊信息的本体表示问题,本文对本体语言 OWL DL 进行了基于模糊逻辑的扩展,给出了形式化的语法和语义,并通过一个实例说明了该方法在表达能力上的灵活性。

关键词　本体,OWL,模糊集合,描述逻辑

Fuzzy Extension of OWL Ontology Language

ZHAO De-xin[1,2]　FENG Zhi-yong[1]

(School of Computer Science and Technology, Tianjin University, Tianjin 300072, China)[1]

(School of Computer Science and Technology, Tianjin University of Technology, Tianjin 300191, China)[2]

Abstract　Ontology is intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain. Although ontology has a powerful expressive ability on knowledge, it has no capability to represent the uncertainty and imprecise information. Vague or imprecise information is most important in the context of semantic Web and multimedia applications. This paper proposes an extension of OWL DL with fuzzy logic to deal with the problem. The formal syntax and semantics are presented and an instance shows that fuzzy information can be expressed well in the proposed ontology language.

Keywords　Ontology, OWL, Fuzzy set, Description logics

语义网[1]的目标是数据、信息和知识交换的通用媒介,将现在的适合人类阅读(human readable)的网络内容,转变为机器可理解(machine understandable)的内容。为了实现机器可理解,就必然需要网络内容能够形式化地通过元数据描述自身。本体将在元数据的精确定义和共享资源上起到至关重要的作用。在开发本体语言上,人们已做了很多工作。目前的本体语言由于是建立在二值逻辑的基础上,因此限制了其对于模糊信息的表达。然而,模糊信息在客观世界中是普遍存在的,它在人类的认知过程中起到了重要的作用。模糊信息实际是对数据的一种压缩,例如:一个"漂亮的"女孩,一次"舒适的"旅行,这些日常生活中用到的词语包含着大量的信息。本文通过增强本体语言的逻辑基础,将模糊逻辑引入其中,从而为本体语言提供了能够表达和处理模糊信息的能力。

1　OWL 简介

OWL[2]是由 W3C 提出的最有发展前景的本体语言。OWL 的逻辑基础是描述逻辑,由于 OWL 在保证更加强大的语义表达能力的同时,还要保证描述逻辑的可判定推理,因此 W3C 的设计人员针对各类特征的需求制定了三种相应的 OWL 的子语言,即 OWL Lite, OWL DL 和 OWL Full,这三种子语言的表达能力递增。相比较而言,OWL DL 兼顾了语言的表达能力和计算的完备性和可判定性,因此本文针对 OWL DL 进行基于模糊逻辑的扩展,将这种新的扩展的本体语言称为 FOWL(Fuzzy OWL)。

2　FOWL 语法

1965 年美国控制理论家 L. A. Zadeh 发表了《模糊集合论》,从而创立了模糊集理论[3],模糊逻辑随后也诞生。它是研究模糊命题演算和模糊推理的一种非布尔逻辑,其理论和方法非常适合对不完备信息的处理。形式上,模糊集合 A 是论域 X 上的一个模糊子集,由其隶属函数来表征,记为 μ_A,其定义 $\mu_A(x): x\in X\rightarrow[0,1]$,表示元素 x 属于 A 的程度。模糊集合 A 可以表示成 $A=\{(x,\mu_A(x)), x\in X\}$。当隶属函数的值域由区间[0,1]蜕化为两个数{0,1}时,相应的模糊子集就成为通常的集合。在 FOWL 中,隶属度的概念可视为对公理描述客观事务的真实程度。如可以指派给 $\mu_A(x)$ 在[0,1]之间的一个值,以表达断言"x *is* A"的程度。由于 OWL 的逻辑基础是描述逻辑,为了保持叙述的简洁,我们使用描述逻辑的形式来表示 FOWL 的语法和语义。

概念(Concept)、角色(Role)和个体常量(Individual Constant)是构建描述逻辑的基础。概念是用来描述对象的集合,例如"教授"。角色给定了二元性质,如"朋友"。个体常量是对个体的简单命名,如"赵四"。从 OWL 数据建模的观点看,概念(Concept)对应 OWL 的类(Class),角色(Role)对应 OWL 的属性(Properties)。

在 FOWL 中,复杂的类、属性、个体可以由基本的类、属性、个体通过构造子得到,构造子与 OWL 的相同,这些构造子被用于创建类表达式。如支持基本的类操作,即并、交和补

*)国家"863"计划基金资助项目"面向解决方案的服务架构和支撑环境"(2007AA01Z130)。**赵德新**　博士生,主研方向为语义网技术、服务发现;**冯志勇**　教授,博导。

维普资讯 http://www.cqvip.com

运算。它们分别被命名为 fowl:unionOf,fowl:intersectionOf 和 fowl:complementOf。类可以是枚举的、用 fowl:oneOf 构造子来显示的声明。此外,还有类的存在性限定和值限定等构造子。通过这些构造子来表现类与类之间的关系,这些关系限制式可以单独使用,也可以混合使用,这样能灵活表达出某种特定的概念。下面给出 FOWL 相应公理的定义。

定义 1 设 a,b 是个体,C,C_1,C_2 是 FOWL 的类,P,P_1,P_2 是属性,M 代表修饰符,$m\in M$ 是修饰算子,$n\in[0,1]$,FOWL 有以下公理:

类公理 包含模糊概念、包含公理的有限集,有以下公理:

$\langle C_1\hat{o}C_2\geqslant n\rangle$(moreOrEquivalent),$\langle C_1\hat{o}C_2\leqslant n\rangle$(lessOrEquivalent),$\langle C_1\hat{o}C_2>n\rangle$(moreThan),$\langle C_1\hat{o}C_2<n\rangle$(lessThan),$m(C)$(modifyClass);

属性公理 包含一个模糊属性公理的有限集,有以下公理:$\langle P_1\hat{o}P_2\geqslant n\rangle$(moreOrEquivalent),$\langle P_1\hat{o}P_2\leqslant n\rangle$(lessOrEquivalent),$\langle P_1\hat{o}P_2>n\rangle$(moreThan),$\langle P_1\hat{o}P_2<n\rangle$(lessThan),$P_1\equiv P_2^-$(inverseOf),$P^+\hat{o}P$(Transitive),$P_1\equiv P_1^-$(Symmetric),$\cdot\hat{o}\leqslant 1P$(Functional),$\cdot\hat{o}\leqslant 1P^-$(InverseFunctional);

个体公理 包含一个模糊概念和模糊角色断言的有限集,有以下公理$\langle a:C\geqslant n\rangle$,$\langle a:C\leqslant n\rangle$,$\langle a:C>n\rangle$,$\langle a:C<n\rangle$,$\langle(a,b):P\geqslant n\rangle$,$\langle(a,b):P\leqslant n\rangle$,$\langle(a,b):P>n\rangle$,$\langle(a,b):P<n\rangle$.$\{a\}\equiv\{b\}$(sameIndividualAs),and $\{a\}\hat{o}\neg\{b\}$(differentFrom)。

按照以上的定义可知,FOWL 将 OWL DL 进行了基于类、属性、个体的模糊扩展。

3 FOWL 语义

语义是推理支持的先决条件,形式化的语义推理支持常常是通过将本体语言映射到某个形式化的逻辑系统上。OWL 是映射到谓词逻辑的子集——描述逻辑[4]上。描述逻辑是基于框架的知识表示语言的逻辑基础,也是语义网本体语言的逻辑基础。虽然 FOWL 的许多公理与 OWL 有同样的语法,但是在 FOWL 中,概念 C 不再是一个精确的定义,因为它被解释成模糊子集,而不是通常 OWL 的经典集合。对于"a is C"的断言,将有一个在[0,1]上的真值(truth-value)来表示个体 a 属于概念 C 的隶属度。

定义 2 FOWL 的语义解释是一个二元组 $I=(\Delta^I,\cdot^I)$,由称为论域的非空集合 Δ^I 和模糊解释函数 $\cdot^I$ 组成,其中模糊解释函数定义如下。

(1)对个体常量的解释:个体常量 a 映射为域 Δ^I 中的元素,这与在 OWL 中的解释相同;

(2)对概念的解释:概念 C 被映射到 Δ^I 上的一个隶属度函数,即 $C^I:\Delta^I\to[0,1]$;

(3)对属性的解释:属性 P 被映射到 $\Delta^I\times\Delta^I$ 上的一个隶属度函数,即 $P^I:\Delta^I\times\Delta^I\to[0,1]$;

(4)修饰算子的解释:修饰算子是将隶属度在[0,1]的概念,映射为另一个隶属度在[0,1]的概念,即 $m:[0,1]\to[0,1]$。

从上面的定义可以看出,每一个解释函数 I 均是一个具有在[0,1]上的隶属度的模糊解释,而不是经典的 0 或 1。具体的语义由以下公式给出($x,y\in\Delta^I$):

$\cdot^I(x)=1$

$\perp^I(x)=0$

$(C_1\hat{o}C_2)^I(x)=C_1^I(x)\wedge C_2^I(x)$

$(C_1\grave{o}C_2)^I(x)=C_1^I(x)\vee C_2^I(x)$

$\{a_1,\cdots,a_n\}^I(x)=\vee_{i=1}^n a_i^I=x$

$(\neg C)^I(x)=\neg(C^I(x))$

$(m(C))^I(x)=m(C^I(x))$

$(\forall P.C)^I(x)=\inf_{y\in\Delta^I}P^I(x,y)\to C^I(y)$

$(\exists P.C)^I(x)=\sup_{y\in\Delta^I}P^I(x,y)\wedge C^I(y)$

$(\geqslant nP)^I(x)=\sup_{y_1,\cdots,y_n\in\Delta^I}\wedge_{i=1}^n P^I(x,y_i)$

$(\leqslant nP)^I(x)=\neg(\geqslant n+1P)^I(x)$

$(C_1\hat{o}C_2)^I=\inf_{x,y\in\Delta^I}C_1^I(x,y)\to C_2^I(x,y)$

$(a:C)^I=C^I(a^I)$

$((a,b):P)^I=P^I(a^I,b^I)$

4 修饰算子

修饰算子在 FOWL 中有着重要的作用,通过修饰算子作用于模糊集合来改变其隶属函数。例如,在自然语言中,有些词如"很"、"稍许"、"非常"、"有点"等,起到程度修饰的作用,我们可以用程度修饰算子来定义这类修饰语。

定义 3 设 X 是域,A 是模糊集合,$\mathscr{F}(X)$ 表示 X 上的全集,$\lambda\in R^+$,有如下映射:

$D^{(\lambda)}:\mathscr{F}(X)\to\mathscr{F}(X)$

$D^{(\lambda)}(A)(x)\triangleq(A(x))^\lambda$

其中 $D^{(\lambda)}$ 称为程度算子。根据 λ 所起的作用不同,当 $\lambda>1$ 时,$D^{(\lambda)}$ 称为收敛算子;当 $\lambda<1$ 时,$D^{(\lambda)}$ 称为弱化算子。

在 FOWL 中,我们可以根据上下文具体情况定义该算子的形式。一般情况下,可以参考如下设定:[极度(extremely)]$=D^{(4)}$,[非常(very)]$=D^{(2)}$,[稍微(little)]$=D^{(0.5)}$。

5 一个例子

FOWL 与 OWL 一样是建立在 RDF 和 RDF Schema 基础上的,使用基于 XML 的语法来表达。考虑到表示的简洁性,我们用抽象语法的形式来给出一个简单的实例。该例子是关于"Tom"的 FOWL 本体,是对 Tom 的简单介绍。内容是:"Tom 是一个很年轻的男子,他在美国上大学,他喜欢学中文"。在上面的介绍中,有一些模糊的信息,如"很"、"年轻",这些在 OWL 中是无法表达的。而在 FOWL 中这些词通过修饰算子和隶属度函数就能恰当地表达出来。

其中 $youngperson=person\hat{o}\exists age.Young$,Young 可以通过隶属度函数来表征。在模糊集合的理论和实践中,有许多种隶属度函数。我们利用 L 函数来界定模糊谓词"年轻"。L 函数定义如下:

$$L(x;a,b)=\begin{cases}1 & \text{if}\quad x\leqslant a\\ (b-x)/(b-a) & \text{if}\quad x\in[a,b]\\ 0 & \text{if}\quad x\geqslant b\end{cases}$$

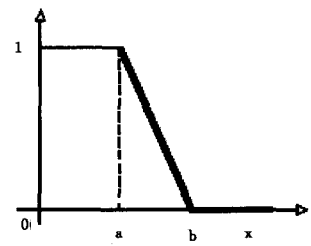

图 1 L 函数

(下转第 175 页)

维普资讯 http://www.cqvip.com

(1)信息来源层:包括非结构化或半结构化的文本和其他信息,它是系统信息的提供者,如数据库、企业人员信息、电子邮件和文档信息等。

(2)组织记忆层:该层表示信息仓库,存储结构化的信息。借助本体论,将信息体加上元信息(来自于本体论),经标注器处理后,变成为结构化的信息项,最终表示成为 XML 文档,存储于 PMIIS 信息库中。

(3)信息服务层:该层提供了信息的服务引擎,包括信息推送引擎、信息检索引擎和信息过滤引擎。其中信息推送指依据被服务者的个人爱好和信息订制,系统及时向其主动提供信息。这些服务引擎依据特定的协议和协作模型,负责不同应用程序数据请求格式(不一定是 XML 形式)的信息发布,同时既要通过必要的策略保证数据的一致性,又要实现必要的访问权限控制。

(4)项目应用层:即项目实施应用层面。根据对信息系统的具体应用需求和用户计算环境,利用系统提供的服务引擎,开发合适的项目应用,如项目计划、项目实施、项目验收等等。

结束语 虽然国内外对于本体论的研究一直呈增长趋势,但是本体的研究和应用很多都还处于不成熟的探索阶段。在探索的过程中,越来越多的新的问题呈现出来,本文在本体论的项目管理信息集成的基础上给出了系统的理论分析和策略方法,目前正处在系统实施和项目整体集成的试行阶段。通过系统实施验证了文章中的理论和构想,但很多其它问题如本体的形式化程度不够、集成工具的缺乏、本体获取的方法的不成熟等等还有待进一步研究,这种种问题都说明本体论的研究和应用还需要该领域的研究者们共同努力,相互交流,以取得更大的进展。

参 考 文 献

[1] Bandini S, Colombo G, Sartori F. Towards the Integration of Ontologies and SA-Nets to Manage Design and Engineering Core Knowledge // First Workshop on Ontology, Conceptualizations and Epistemology for Software and Systems Engineering 2005. University of Alcalá, Alcalá de Henares, Madrid, Spain, 2005

[2] 余明晖,费奇. 基于本体论的工程信息管理分解与建模. 华中科技大学学报(城市科学版),2004,21(1):15-19

[3] 韩宗海,刘振元,包晓春. 项目管理信息系统集成及其发展趋势. 计算机科学,2006(11)

[4] Bianchini D, Antonellis V D, Melchiori M. Ontology-based Interoperability for Interorganizational Applications. in Enterprise Modelling and Ontologies for Interoperability 2004. Riga, Latvia, 2004

[5] Koschmider A, Oberweis A. Ontology Based Business Process Description. in Enterprise Modelling and Ontologies for Interoperability 2005. Porto, Portugal, 2005

[6] Zhu Y, Augenbroe G. A Conceptual Model for Supporting the Integration of Inter-Organizational Information Processes of AEC Projects. Automation in Construction, 2006, 15(2): 200-211

[7] Han Z. An Ontology-based Transaction Model for the Exchange of Infrastructure Information. Technical Report. Construction and Project Management Group, The University of British Columbia: Vancouver, BC, Canada, 2005: 1-26

[8] McNally K R, et al. An Ontology-Driven Approach to Representing and Visualizing Critical Infrastructure Interdependencies. 2005. Available at: http://www.acsm.net/cagis/mcnally.pdf

(上接第 171 页)

使用 L 函数,修饰词"年轻"可以被定义为 $Young(x)=L(x;20,35)$,所以可以用该函数来描述 Tom 的年龄。在我们的例子中,假定 $Young(x)=L(x;20,35)=0.8$。对于另一个模糊词语"非常"的形式化描述就用修饰算子来表达。按照定义 3,$[非常(very)]=D^{(2)}$,则有

$[veryYoung]:\rightarrow[Young]^2$,即 $veryYoung:\rightarrow 0.8^2$

下面就是该例子基于 FOWL 的抽象语法(假设该本体的命名空间是"S"):

```
ObjectProperty (S:learns)
ObjectProperty (S:studiedIn)
ObjectProperty (S:livesIn)
ObjectProperty (S:likes)
ObjectProperty(S:age
range (S:Natural))
Class (owl:Thing partial)
Class (S:person partial)
Class (S:college partial)
Class (S:undergraduate complete
annotation (rdfs:comment "Someone who studies in a college.")
intersectionOf ( restriction (S: studiedIn someValuesFrom (S:
college))
S:person))
Class (S:male partial
annotation (rdfs:comment "The class of all male things."))
Class (S:young partial
annotation (rdfs:comment "Person whose age is less than 35")
domain (S:person)
range (S:(Young,0.8)))
Class (S:Chinese partial
S:languages)
Class (S:Chinese_liker complete
intersectionOf ( restriction (S: likes someValuesFrom (S: Chi-
nese)) S:person))
Class (S:person partial)
Individual (S:Tom
annotation (rdfs:comment "Tom is a male undergraduate in A-
merica. He is very young, and he likes learning Chinese.")
type (S:undergraduate)
type (S:Chinese_liker)
type (S:male)
value (S:age S:(veryYoung,0.8²)
value (S:livesIn S:America))
```

结束语 本文关注于对不精确的或模糊的信息的本体表示问题,这些模糊信息在我们现实生活中广泛存在,对这类信息进行形式化的表示和推理,是值得深入研究的问题。对于描述逻辑的针对处理不精确信息的扩展也有很多研究成果。Straccia 不但研究了如何将模糊知识引入到描述逻辑中,而且对于逻辑推理问题进行了深入的研究,提出了一种模糊描述逻辑 FALC[5]。但是 FALC 的表达能力还比较有限,也没有对大量存在的模糊性的修饰词语给出定义。在本文中,我们通过模糊逻辑扩展 OWL DL,将这种新的扩展的本体语言称为 FOWL(Fuzzy OWL)。在 FOWL 中,模糊信息可以通过抽象语法表达,其含义可以由语义来给出解释。另外,我们定义了修饰算子来增强语言的表达能力。实际上,模糊逻辑和相关的概念及技术能在语义网的开发和应用中起到积极作用,它有望在计算机和人类之间架起一道桥梁。

参 考 文 献

[1] Berners-Lee T, Hendler J, Lassila O. The semantic web. Scientific American, 2001

[2] McGuinness D L, van Harmelen F. OWL Web Ontology Language Overview. World Wide Web Consortium (W3C) Candidate Recommendation. http://www.w3.org/TR/owl-features/ August 2003

[3] Zadeh L A. Fuzzy Sets. Information and Control, 1965, 8: 338-353

[4] Baader F. The Description Logic Handbook: Theory, Implementation and Applications. New York: Cambridge University Press, 2003

[5] Straccia U. Reasoning within fuzzy description logics. Journal of Artificial Intelligence Research, 2001, 14(1): 137-166