

●徐国虎 许芳董慧

基于描述逻辑的本体检错推理研究*

——以国共合作领域本体为例

摘 要 本体检错推理在本体形式化汉语言的描述逻辑基础上,运用 Tableau 算法对本体的概念层次,声明的实例以及实例间复杂的语义关系进行检测,以保证本体库结构的逻辑一致性和知识描述的正确性。以国共合作领域本体库为例,探讨了本体检错推理的逻辑基础及 Tableau 算法的原理,分析本体检错推理的内容,并采用描述逻辑推理接口协议 DIG 集成描述逻辑推理引擎 Racer 实现了国共合作领域本体检错推理。图 1。参考文献 8。

关键词 描述逻辑 本体检错 领域本体 本体推理 本体库结构 知识描述

分类号 G254

ABSTRACT In this paper, the authors introduce the definition of ontology error checking reasoning. Taking the domain ontology database of the Cooperation between Kuomintang and the Communist Party of China as an example, the authors analyze the logical basis of ontology error checking reasoning and the principles of Tableau algorithm, analyze the contents of ontology error checking reasoning, and realize the error checking reasoning for the domain ontology of the Cooperation between Kuomintang and the Communist Party of China by using the description logic reasoning interface protocol DIG and integrated description logic reasoning engine Racer. 1 fig. 8 refs.

KEY WORDS Description logic. Ontology error checking. Domain ontology. Ontology reasoning. Ontology database structure. Knowledge description.

CLASS NUMBER G254

应用本体来组织和描述知识存在着本体构建工作复杂繁重、依赖于领域专家、知识组织易出错而且错误难以发现等问题。这就促使我们思考,能否从本体形式化的逻辑基础入手,借助于逻辑检错推理引擎对本体知识库进行推理检测,从而解决或缓解这些问题。本文以国共合作领域本体库为例,对本体检错推理的逻辑基础、推理算法原理、检错推理内容及其推理实现等问题进行探讨。

1 本体检错推理的逻辑基础:描述逻辑

本体知识推理是以(多)领域本体知识库中已有的形式化知识为根据,按照一定的逻辑公理和规则来检验本体形式化知识的逻辑严密性和发现蕴涵的、未显式描述的知识。本体知识推理包括基于逻辑的知识检错推理和基于关系的蕴涵知识发现推理。其中前者在本体形式化语言的逻辑基础上,运用 Tableau 算法对(领域)本体的概念层次、声明的实例以及实例间复杂的语义关系进行检测,以保证本体库结构的逻辑一致性和知识描述的正确性,为后续的蕴涵知识推理、本体库的拓展和与其他领域本体库的集成提供

逻辑保证。

本体作为一种知识表示方法,其逻辑基础、形式化语言以及模型验证问题都与描述逻辑(description logic, DL)紧密相关。目前主流的本体描述语言 DAML、OIL 和 OWL 都以描述逻辑作为形式化的基础,主流的本体推理引擎也主要是基于描述逻辑实现的,如 FaCT、RACE 和 Racer 就分别是基于描述逻辑 SH、描述逻辑 SHN 与描述逻辑 SHIQ^[1]。从这个角度看,描述逻辑可以被认为是本体检错推理的逻辑基础。

1.1 描述逻辑的语义元素

描述逻辑是人工智能领域的一个研究分支,是一种用来描述概念和概念层次关系的知识表示方法,可以看成是谓词逻辑的子语言,具有严格的以逻辑为基础的语义和良好的推理机制。描述逻辑通过定义应用领域的概念及其结构关系,刻画领域内的个体信息。描述逻辑建立在概念和关系之上,由构造因子从简单概念和关系构造出复杂概念和关系。描述逻辑包含 5 种基本的语义元素:原子概念、原子关系、构造因子集、TBox 概念定义集和 ABox 实例断言集。一个

* 本文系国家自然科学基金资助项目(批准号 70373047)和教育部基地重大课题(批准号 05JJD870004)研究成果之一,并得到中南财经政法大学博士引进科研启动基金资助。

描述逻辑系统的表示能力和推理能力取决于对这5种语义要素的选择组合以及不同的假设^[2]。

概念具有内涵和外延,可分为原子概念(atomic concepts)和复杂概念(complex concepts)。原子概念是一种一元谓词(unitary predicates),表示最基本的概念,复杂概念则由原子概念通过构造因子以及关系组合而成。

关系表示概念的属性或不同概念之间存在的联系,原子关系是一种二元谓词(binary predicates),与复杂概念一样,复杂关系由原子关系通过构造因子以及概念组合而成。

构造因子可以理解为逻辑运算符,用于构造复杂概念和复杂关系,从而增强描述逻辑的表达力。不同的描述逻辑系统,其表达能力的强弱和推理的复杂性在很大程度上取决于它所拥有的构造因子的不同。描述逻辑常见的构造因子包括:合取因子、析取因子、否定因子、存在量词、全称量词、数量限量因子、枚举因子等等。

TBox 是一个描述领域结构的公理集,包含内涵性知识,用概念来表示,通过一组描述概念的一般性属性的声明来实现。内涵性公理知识被认为是不变的,它一般包括概念定义公理和概念包含公理。ABox 是一个描述关于具体个体事实的公理集,包含的是外延性知识。外延性知识通常被认为是偶然的,依赖于某一特定的环境,因此它是经常变动的。ABox 中的断言包括实例断言和关系断言。

1.2 基本描述逻辑系统 ALC

描述逻辑的发展起源于1977年的KL-ONE系统,至今已发展为一个庞大的逻辑家族,其中ALC(attributive concept description language with complements)是最基本的一种描述逻辑,它包含原子概念、原子关系以及合取、析取、否定、存在量词和全称量词等构造因子^[3]。在ALC的基础上再添加不同的构造因子,则构成不同表达能力的描述逻辑系统。由原子概念、原子关系与构造因子进行复杂运算形成的ALC语法与语义如下所述。

定义1 设 N_c 是一个概念集, N_r 是一个关系集,则 ALC—概念集为满足下列条件的最小集合:

- (1) 若概念 $C \in N_c$, 则 C 是 ALC—概念;
- (2) 若 C, D 是 ALC—概念, $R \in N_r$ 是关系名, 则 $C \cap D, C \cup D, \neg C, \neg D, \forall R. C, \exists R. C$ 都是 ALC—概念。

定义2 设 ALC—解释 $I = (\Delta^I, \bullet^I)$ 由非空集合 Δ^I 和函数 $\bullet^I$ 组成, Δ^I 称为 I 的论域。函数 $\bullet^I$ 将每个概念映射为 Δ^I 的一个子集, 将每个关系映射为 $\Delta^I \times \Delta^I$ 的一个子集。同时满足下列语义: $(C \cap D)^I = C^I \cap D^I, (C \cup D)^I = C^I \cup D^I, (\neg C)^I = \Delta^I \setminus C^I, (\forall R. C)^I$

$= \{x \in \Delta^I \mid \text{对任意 } y \in \Delta^I, \text{若 } (x, y) \in R^I, \text{则 } y \in C^I\},$
 $(\exists R. C)^I = \{x \in \Delta^I \mid \text{存在 } y \in \Delta^I, \text{满足 } (x, y) \in R^I, \text{而且 } y \in C^I\}。$

定义3 概念 C 是可满足的, 当仅存在解释 I 使得 $C^I \neq \emptyset$, I 称为 C 的模型。

定义4 概念 D 包含概念 C , 当仅对所有解释 I 使 $C^I \subseteq D^I$ 成立, 记作 $C \subseteq D$ 。

定义5 概念 C 与 D 是等价的, 当且仅当 C 与 D 互包含, 即 $C \subseteq D$, 而且 $D \subseteq C$, 记作 $C = D$ 。

定义6 对于一个解释 I , 称个体 $x (x \in \Delta^I)$ 是概念 C 的实例, 当且仅当 $x \in C^I$ 。

由于存在否定因子, 我们可以将包含关系的推理转化为可满足性的证明。这是因为: $C \subseteq D$ 成立当且仅当 $C \cap \neg D$ 是不可满足的。反之, 可满足性推理也可演算为包含关系的证明, 这是因为: C 是可满足的当且仅当 $C \subseteq (E \cap \neg E)$ 不成立, 其中 E 是任一概念名^[4]。

由于 ALC 仅仅包括合取、析取、否定、存在量词和全称量词等构造因子, 其描述能力也非常有限, 还不足以刻画纷繁复杂的大千世界。在 ALC 的基础上对描述逻辑进行扩展的方式主要有经典扩展和非经典扩展两类。描述逻辑经典扩展主要是通过添加概念构造因子与关系构造因子来实现; 而描述逻辑的非经典扩展主要包括具体领域限制, 添加模态、时态算子, 概率和模糊逻辑扩展以及缺省逻辑扩展。

2 本体检错推理的 Tableau 算法

可满足性问题是描述逻辑推理的核心问题, 因为其他许多问题(如包含性检测、一致性问题等)都可化为可满足性问题。应用于描述逻辑推理的 Tableau 算法最早由 Schmidt-Schaub 和 Smolka 为检验描述逻辑 ALC 概念的满足性而提出, 该算法能在多项式时间内判断描述逻辑 ALC 概念的可满足性问题^[5]。

2.1 Tableau 算法的基本原理

Tableau 算法主要用于逻辑系统中对概念间的定义、关系进行满足性测试。任何 Tableau 算法的基本思想都试图通过构建一个概念 C 的模型来证明其满足性。它试图建立一个概念 C 的类似于树的模型 I , 通常是在否定范式的概念上工作, 它按照语法分解概念 C , 推导关于模型 I 上元素的约束。算法在一个树(或树集合)上运行, 各节点被标注为 C 的子概念集, 边被标注为 C 上发生的关系集。通常一个单独的树从被标注为 $\{C\}$ 的根节点初始化。

该算法竭尽全力应用各种 Tableau 算法规则来对描述逻辑概念或关系定义中的构造因子进行分解, 从而分解节点上定制概念的语法结构, 或者扩展节点标志, 向树中添加新的边和节点, 或者合并边和节点。

一个规则的应用有效地说明了从概念到应用的规则暗示了推断的约束。例如,如果 $A \cap B$ 是一个节点 x 的标注,那么 $\cap$ 规则被添加到 A 和 B 的标注中,这说明:如果 $x \in (A \cap B)^I$,那么 $x \in A^I$ 且 $x \in B^I$ 。或者说,如果 $\exists B.A$ 是一个节点 x 的标注,那么 $\exists$ 规则添加一个标注为 $\{A\}$ 的新节点 y ,它在 x 和 y 之间带有一个标注为 $\{R\}$ 的边。这也说明:如果 $x \in (\exists R.A)^I$,那么就必须在存在一个节点 y ,使得 $(x, y) \in R^I, y \in A^I$ 。

如果存在明显的矛盾,建立模型的尝试就失败了,这通常被称为产生了冲突。在 Tableau 算法中,如果没有新的规则可被应用,而且一直没有冲突产生,就成功建立了模型。当且仅当 Tableau 规则可以按照模型被成功构建的方式应用,证明一个概念是可满足的就相对简单。但部分 Tableau 规则,比如规则与规则,在算法运行应用中是不确定的,这就意味着搜索,而各种 Tableau 算法的计算复杂度则由于这些规则不确定的事实而增长。事实上当冲突产生时,处理方式是回溯试图找到另一个不确定的规则进行应用。当冲突发生时或没有规则被使用时,Tableau 算法停止运行操作,而循环检验用于保证 Tableau 算法的可终止性。

2.2 Tableau 算法在本体工程中的应用

Tableau 算法对于采用基于描述逻辑的本体描述语言来形式化领域知识的本体系统有着多方面的应用。Tableau 算法及其应用(如 FaCT、Racer、Pellet 等)在本体工程中的应用,除了可优化本体描述,集成和融合相关(相似)本体,最主要的应用还是辅助本体构建者检测本体冲突从而验证本体知识的正确性。

本体库中除了实例声明体系存在冲突,概念定义体系中也可能存在冲突,而且相比之下更加致命。因为如果 TBox 存在内在冲突,则在 TBox 上构建的 ABox 就存在隐含的逻辑错误。冲突检测的基本步骤是按照知识层次的顺序(可以是类和属性体系以及实例体系)运用相关的 Tableau 算法规则检测出可能产生矛盾冲突的信息。但因为本体体系通常十分庞大,关系复杂,而且在本体的构建与使用过程中会不停地变化,本体冲突检测如果不借助基于 Tableau 算法的推理引擎,那么人工检测本体冲突的工作量将是巨大的,而且非常低效。在我们的项目实施中,就是借助 Racer 推理引擎再辅以人工来对国共合作本体库进行检错。

3 基于描述逻辑的本体检错推理的内容

在描述逻辑里,本体知识库被分成了两个部分,即 $KB = \langle TBox, ABox \rangle$,因此基于描述逻辑的本体

检错推理主要包括 TBox 概念检验推理和 ABox 实例检验推理,检验推理的内容主要包括包含性检验、可满足性检验、一致性检验、等价性检验与非相交性检验等^[6]。

3.1 TBox 概念检验推理

本体知识库 TBox 部分引进术语(Terminology),即一个应用领域的主要概念,它是描述应用领域结构的公理集合。TBox 里的命题的基本形式是概念定义和概念包含,前者是指新的概念根据其他先前定义的概念来定义,而后者主要定义两个概念之间的语义包含关系,比如 $A = C, A \subseteq C; \text{Mother} = \text{Woman} \cap \exists \text{has-child. Human}, \text{Human} \subseteq \text{Animal} \cap \text{Biued}$ 。

TBox 概念检验推理主要对本体中的概念定义进行可满足性检验、包含性检验、等价性检验与非相交性检验。给定本体的形式化定义 $0 = \langle T, X, TD, XD \rangle$,其中 T 是概念集, D 与 E 为概念公式, T 中的概念又被称为原子概念; TD 是概念定义集 TBox,用来定义 T 中的概念; X 为实例集; XD 为实例声明集 ABox,用来声明概念类的实例。

概念的可满足性检验:一个解释 I 是 TD 的一个模型,当且仅当对每个 $C_1 \subseteq C_2 \in T, C_1^I \subseteq C_2^I$,成立并且对每个 $C_1 = C_2 \in T, C_1^I = C_2^I$ 成立(C_1, C_2 为 C 的子概念),此时记为 $I \models TD$ 。一个术语 $C \in T$ 是关于 TD 可满足的当且仅当有一个具有 $I \models TD$ 和 $C^I \neq \emptyset$ 的 I ,否则称 C 关于 TD 是不可满足的。概念的可满足性用于评价概念公式对于概念集是否有意义。

概念的包含性检验:如果对任意一个本体解释 I ,有 $D^I \subseteq E^I$,则称 E 包含 D ,记为 $I \models DMR$ 。如果对任意一个 TD 的模型 I ,有 $D^I \subseteq E^I$,则称 TD 涵蕴着 E 包含 D ,记为 $TD \models DME$ 。包含性检验用于评价本体中概念之间的包含关系,可以检验本体概念类 $Is-a$ 层次结构的合理性。

概念的非相交性检验:如果对任意一个本体解释 I ,有 $D^I \cap E^I = \emptyset$,则称 D 与 E 非交,记为 $I \models D - \cap E$ 。如果对任意一个 TD 的模型 I ,有 $D^I \cap E^I = \emptyset$,则称 TD 涵蕴着 D 与 E 非交,记为 $TD \models D - \cap E$ 。非相交性检验用于判断两个概念类是否具有相同的实例。

概念的等价性检验:如果对任意一个本体解释 I ,有 $D^I = E^I$,则称 D 与 E 等价,记为 $I \models D = E$ 。如果对任意一个 TD 的模型 I ,都有 $D^I = E^I$,则称 TD 涵蕴着 D 与 E 等价,记为 $TD \models D = E$ 。等价性检验用于检验两个概念类是否具有相同的实例集。

实际上由下面命题可知,上述 4 类概念检验问题是可以归结为可满足性检验的:给定本体 $0 = \langle T, X, TD, XD \rangle$, D 与 E 为概念公式,包含性、等价性、非相交性检验均可转换为可满足性检验:

$TD \models DME$,当且仅当 $D \cap \neg E$ 关于 TD 是不可

满足的;

$TD \models D \sqcap E$, 当且仅当 $D \cap \neg E$ 与 $\neg D \cap E$ 关于 TD 是不可满足的;

$TD \models D - \cap E$, 当且仅当 $D \cap E$ 关于 TD 是不可满足的。

3.2 ABox 实例检验推理

本体知识库 ABox 部分是描述具体个体事实断言状态的集合, 包括概念实例声明断言与实例关系断言。概念实例声明断言一般形如: $a:C, John:Man \cap \exists has-child. Female$ 。实例关系断言一般形如: $\langle a, b \rangle : R, \langle John, Rose \rangle : has-child$ 。实例检验主要包括实例声明一致性检验和实例声明集一致性检验, 而实例声明一致性检验又可转换为对实例集的一致性检验。

实例声明的一致性检验: 给定本体 $O = \langle T, X, TD, XD \rangle$, 若存在本体解释 I 是实例声明 A 的一个模型, 则称 A 是一致的, 否则称 A 是不一致的。若 I 是 A 的一个模型, 又是 TD 的一个模型, 则称 A 关于 TD 是一致的。若 I 是 XD 的一个模型, 则称 XD 是一致的。若 I 是 XD 的一个模型, 又是 TD 的一个共有模型, 则称 XD 关于 TD 是一致的。

实例声明的一致性检验推理问题追根究底就是

判定 ABox XD, 个体 a 和概念 C 之间是否 $XD \models C(a)$ 。它主要用来检验个体实例的声明是否矛盾。假设本体中存在这样的实例声明集 $XD = \{ \text{抗日英雄}(\text{汪精卫}), \text{卖国汉奸}(\text{汪精卫}) \}$, 由于事先已定义“抗日英雄”与“卖国汉奸”这两个角色类是非交的, 即对任何一个本体解释, “抗日英雄”与“卖国汉奸”都被解释为是两个不相交的集合, 它们不存在共同的模型, 因此个体“汪精卫”不可能被同时声明为“抗日英雄”与“卖国汉奸”的实例, 所以 XD 关于 TD 是不一致的。这就是本体实例声明的一致性错误。

实例声明集一致性检验: 给定本体 $O = \langle T, X, TD, XD \rangle$, 如果 XD 的每一个模型 I 都满足实例声明 A, 则称 XD 蕴涵着 A, 记为 $XD \models A$ 。

实际上, 由下面两个命题可知, 对某个实例声明一致性的检验是可以转换为对实例集的一致性检验。命题 1: 给定本体 $O = \langle T, X, TD, XD \rangle$, $XD \models C(a)$, 当且仅当 $XD \cup \{ \neg C(a) \}$ 是不一致的。命题 2: 给定本体 $O = \langle T, X, TD, XD \rangle$, 有 (i) C 关于 TD 是可满足的, 当且仅当存在个体 a, 使得实例声明 $C(a)$ 关于 TD 是一致的; (ii) C 关于 TD 是可满足的, 当且仅当存在个体 a, 使得实例声明集 $\{ C(a) \}$ 关于 TD 是一致的。

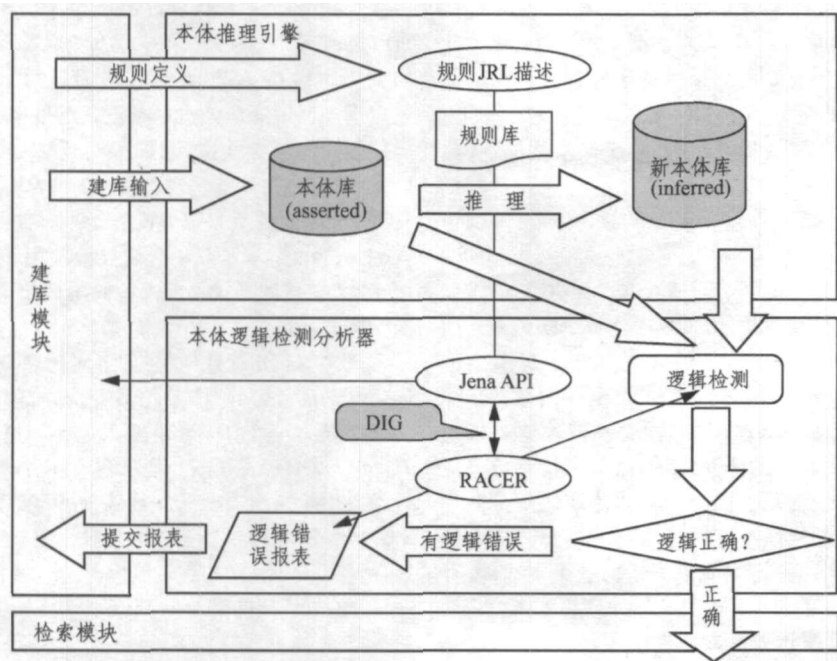


图1 国共合作领域本体推理模型结构

3.3 TBox 推理与 ABox 推理的关系

对整个本体知识库推理而言, 知识库的一致性是其基本的推理问题: 对 ABox A 和 TBox T, 本体知识

库 $\langle T, A \rangle$ 是否一致, 是否有一个 $\langle T, A \rangle$ 的模型 I。根据模型理论语义, 本体知识库的一致性的形式化定义为: 如果存在一个 A 并且是 T 的模型的解释, 那

么一个 ABox A 是关于一个 TBox T 一致的。实际上由定理:概念 C 是可满足的 $\Leftrightarrow \{C(a)\}$ 是一致的,可知对一个知识库的一致性推理问题能被简化成关于概念的可满足性问题^[7],这也就是说 TBox 推理与 ABox 推理是紧密相关的,两者一起构成了完整的本体知识库检错推理,但检错推理的重心在于 TBox 推理,因为对实例的检错推理是可以转化为对概念的可满足性检测。

4 基于 Racer 的本体检错推理实现

4.1 系统构架

国共合作领域本体检错推理只是整个国共合作领域本体推理模型的一部分,其主要结构如图 1 所示。在参考国外研究成果并与领域专家商量的基础上,决定基于描述逻辑的领域知识检错推理直接在 Jena 中通过 DIG 协议 (Description Logic Inference Gateway) 来集成外部的 Racer 推理引擎插件在原始的本体库与蕴涵推理后的本体库中进行逻辑检错,并依据其检错报表进行手工修改。

Racer (Renamed ABox and Concept Expression Reasoner) 是加拿大 Concordia 大学的 Haarslev 等人开发的一个语义网推理引擎,它可以对 RDFs、OIL + DAML 和 OWL 的本体知识库进行推理^[8]。Racer 推理引擎实现了基于描述逻辑系统 SHIQ 的 TBox 和 ABox 推理,可以提供概念一致性推理、包含性推理、父子概念判定推理、实例检测推理等推理服务。Racer 与它的前身 RACE 一样,采用的都是优化 Tableau 算法。它支持标准的描述逻辑推理接口协议 DIG,可以与其他推理引擎或其他应用集成使用。

4.2 具体实现

语义工具 Jena 在本体模型和实现了 DIG 接口的外部推理引擎之间提供一个透明网关,通过这样一种机制,就可以在 Jena 中集成外置逻辑推理引擎来对本体知识库进行检错推理^[9]。DIG 接口是一个逐渐发展形成的标准,它通过基于 HTTP 协议的接口来提供对描述逻辑推理引擎的访问,从而完成一个分布式的推理过程。我们使用 DIG 来调用外部的 Racer 进行逻辑检错的主要代码如下:

```
Model cModel = ModelFactory.createDefaultModel();
Resource conf = cModel.createResource();
conf.addProperty(ReasonerVocabulary.EXT_REASONER_URL,
    cModel.createResource("http://localhost:2004"));
DIGReasonerFactory drf =
    (DIGReasonerFactory) ReasonerRegistry.
```

```
theRegistry()
    .getFactory(DIGReasonerFactory.
        URI);
DIGReasoner r = (DIGReasoner) drf.create(conf);
OntModelSpec spec = new OntModelSpec(OntModelSpec.OWL_DL_MEM);
spec.setReasoner(r);
OntModel m = ModelFactory.createOntologyModel
    (spec, null);
```

使用 ReasonerRegistry.theRegistry().getFactory(DIGReasonerFactory.URI) 方法创建 DIGReasonerFactory 对象,再用 DIGReasonerFactory 的 create(conf) 方法创建 DIGReasoner 对象。这里的 conf 对象是一个预定义的 Resource 对象,用来设置 DIGReasonerFactory 对象的参数,比如代码中的 ReasonerVocabulary.EXT_REASONER_URL 参数,用来设置外置推理引擎的访问地址。得到 DIGReasoner 对象后,就可以进一步创建 OntModel 对象了,创建完后就可以进行下面的检测步骤了。

```
ValidityReport validity = infmodel.validate();
if (validity.isValid()) {
    System.out.println("\nOK");
} else {
    System.out.println("\nConflicts");
    for (Iterator i = validity.getReports(); i.hasNext(); ) {
        ValidityReport.Report report = (ValidityReport.Report) i.next();
        System.out.println(" - " + report);
    }
}
```

调用 ValidityReport 对象的 isValid() 方法就可以得到检测结果。如果有冲突,再调用 ValidityReport 对象的 getReports() 方法,这里将返回一个迭代集合,迭代遍历这个集合后,就可得到完整的检测结果信息。由于在前期的本体库构建工作中,我们参考了相关领域叙词表,反复听取了领域专家的意见,所以在运用 Racer 对国共合作领域本体库检测时,除了发现 3 处概念定义错误和 21 处实例声明错误,没有发现大的知识描述错误。这些检错推理工作既为后续的蕴涵知识推理开发提供了必需的逻辑保障和原始知识准备,也增强了开发人员发现本体库中逻辑错误的能力,减小了本体构建工作中由于本体构建工作人员认识局限性或疏忽大意而导致本体描述的领域知识发生错误的可能性,减轻了本体库管理和维护的工作量。

4.3 本体检错推理实现的局限性

目前的本体检错推理完全是通过 DIG 接口由 Racer 推理引擎完成的,Racer 性能的好坏对本体库

检错效果影响很大。虽然 Racer 是推理能力较为强大的商业化推理工具,但是其局限性也显而易见。由于 Racer 是基于描述逻辑 Tableau 算法实现,所以对 TBox 概念层次的推理有较好支持,而对 ABox 实例推理支持不太好。而且对于 OWL 的一些属性,如 sameAs、基数约束集等也支持不太好。这就意味着使用 Racer 进行本体库检错推理时并不能完全检测出本体库中的知识描述错误;而且 Racer 是个商业化推理工具,不提供开源服务,我们无法进行二次开发和算法优化,不能为系统提供更好的个性化支持。

参考文献

- 1 R. Moller, V. Haarslev, The Description Logic Systems. Cambridge University Press, 2003: 289 ~ 310
- 2 Franz Baader. The Description Logic Handbook: Theory, Implementation and Applications, Cambridge University Press, 2003. 47 ~ 56
- 3 Schmidt-Schaup M, Smolka G, Attributive concept descriptions with complements. Artificial Intelligence, 1991. 48(1)
- 4 文斌. 基于描述逻辑的语义网知识推理研究. [学位论文]

文], 云南师范大学, 2005: 123 ~ 133

- 5 Baader F, Sattler U, An Overview of Tableau Algorithms for Description Logics. Studia Logica, 2001(69)
- 6 徐国虎. 基于本体的国共合作领域知识推理研究. [博士学位论文]. 武汉大学, 2006
- 7 I Horrocks, S Tobies. Reasoning with Axioms: Theory and Practice. Proc. of the 7th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR' 2000), 2000: 286 ~ 294
- 8 Volker Haarslev, Ralf Möller. Description of the Racer System and its Applications, Workshop on Description Logics (DL - 2001), Stanford, USA, August 2001

徐国虎 中南财经政法大学工商管理学院教师。
通讯地址: 武汉中南财经政法大学(南湖校区)工商
管理学院物流教研室。邮编 430064。

许芳 中南民族大学管理学院教师。通讯地
址: 武汉中南民族大学管理学院。邮编 430074。

董 慧 武汉大学信息管理学院教授、博士生导师。
通讯地址: 武汉大学。邮编 430072。

(来稿时间: 2007-03-06)

(上接第78页)

2.2 SKOS 的局限性

SKOS 采用 RDF 图形式描述, 与本体语言相比有一定的局限, 它并不规定语义, 只为每一个资源描述体系提供一个能够描述其特定需求的语义结构的能力, 缺乏对属性值加以限制, 不能像本体那样采用描述逻辑进行推理, 无法进行一致性检验或者对新类进行归类。因此采用 SKOS 描述的叙词表只是叙词表形式化表示的过渡阶段, 揭示的语义关系有待进一步明确, 离真正转换为本体还需一定的路程, 需要进一步的人工参与, 尤其在语义关系上作进一步调整, 需建立与其他叙词表、本体的映射, 尽可能利用已构建成功的本体的语义关系, 利用各词表及本体的优点, 通过相互补充相互作用发挥本体在语义网中的作用。

3 结束语

W3C 组织明确提出传统叙词表的存在为发展语义网中的本体提供了起点。如果不借助当前叙词表的成果, 语义网的发展将会变得缓慢^[4]。语义网的关键元素是使用本体来定义概念及关系, 通过本体提供数据的上下文环境, 信息检索及搜索引擎能够利用这些上下文信息来执行基于概念的语义搜索, 同时本体所获得的丰富语义可将简单事实合并, 推理出新事实, 并推导出本体所隐含的知识。如何在语义层面共享或重用叙词表, 并使之与本体融合适应当前语义检索的要求, 成为目前叙词表在语义网环境下发展所面临的主要问题。W3C 组织适时发布并维护的简单知

识组织系统为叙词表向形式化表述迈进了一大步, 为叙词表的形式化表达提供了概念模型, 使各叙词表之间、叙词表与本体之间的映射成为可能^[5]。但这并不是叙词表发展的最终目标, 如何利用 SKOS, 采用 Web 上表达机器可理解内容更强的语言描述叙词表, 叙词表如何在语义检索中得到更好应用, 从语义层面组织信息、检索信息, 实现用户检索请求的语义本体化, 将是叙词表研究者面临的更大挑战。

参考文献

- 1 贾君枝. FrameNet 叙词表与传统叙词表语义关系比较研究. 情报理论与实践, 2006(5)
- 2 W3C Working Draft 2. SKOS Core Guide. [2007-01-12] <http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/> SKOS.
- 3 中国科学技术情报研究所, 北京图书馆. 汉语主题词表. 北京: 科学技术文献出版社, 1980
- 4 Salvador Sanchez-alonso, Elena Garcia-barriocanal. Making Use of Upper Ontologies to Foster Interoperability Between Skos Concept Schemes. Online Information Review, 2006, 30(3)
- 5 B. J. Wielinga A. Th. Schreiber J. Wielemaier J. A. C. Sandberg From Thesaurus to Ontology. [2006-07-14]. <http://www.cs.vu.nl/guus/papers/Wielinga01a.pdf>.

贾君枝 山西大学管理学院副教授, 博士。通讯地址:
太原山西大学管理学院。邮编 030006。

(来稿时间: 2007-03-26)