ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.18, No.7, July 2007, pp.1563−1572
DOI: 10.1360/jos181563

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

支持术语公理约束的扩展模糊描述逻辑推理*

康达周[1], 徐宝文[1,2+], 陆建江[3], 李言辉[1]

[1](东南大学 计算机科学与工程系,江苏 南京 210096)
[2](江苏省软件质量研究所,江苏 南京 210096)
[3](解放军理工大学 指挥自动化学院,江苏 南京 210007)

Reasoning Within Extended Fuzzy Description Logic Supporting Terminological Axiom Restrictions

KANG Da-Zhou[1], XU Bao-Wen[1,2+], LU Jian-Jiang[3], LI Yan-Hui[1]

[1](Department of Computer Science and Engineering, Southeast University, Nanjing 210096, China)
[2](Jiangsu Institute of Software Quality, Nanjing 210096, China)
[3](Institute of Command Automation, PLA University of Science and Technology, Nanjing 210007, China)
+ Corresponding author: Phn: +86-25-52090882, Fax: +86-25-83794838, E-mail: bwxu@seu.edu.cn

Kang DZ, Xu BW, Lu JJ, Li YH. Reasoning within extended fuzzy description logic supporting terminological axiom restrictions. *Journal of Software*, 2007,18(7):1563–1572. http://www.jos.org.cn/1000-9825/18/1563.htm

Abstract: Extended fuzzy description logics are fuzzy extensions of description logics, which support representation and reasoning for expressive fuzzy knowledge. But they lack reasoning algorithms with the terminology axioms. This paper defines restricted TBoxes (terminological boxes) to describe terminology axioms in $EFALC_{R+}$(extended fuzzy attributive concept description language with complements and transitive roles), proposes and optimizes reasoning algorithms for $EFALC_{R+}$ with respect to restricted TBoxes. The optimized reasoning algorithm is proved sound, complete and with a worst complexity of exponential time. Its complexity has reached the lower bound, since the reasoning problem for $EFALC_{R+}$ with respect to restricted TBoxes is proved exponential time complete. So it is an efficient algorithm of reasoning for fuzzy knowledge bases with terminology axioms.
Key words: description logic; fuzzy; Tbox (terminological box); reasoning; semantic Web; knowledge representation

摘 要: 扩展模糊描述逻辑是对描述逻辑的一种模糊扩展,支持对复杂模糊知识的表示和推理,但该逻辑缺乏支持术语公理约束的推理算法.提出扩展模糊描述逻辑 $EFALC_{R+}$(extended fuzzy attributive concept description language with complements and transitive roles)的受限 TBox(terminological box)描述术语公理,给出受限 TBox 约束

* Supported by the National Natural Science Foundation of China under Grant Nos.60373066, 90412003 (国家自然科学基金); the National Natural Science Funds for Distinguished Young Scholar of China under Grant No.60425206 (国家杰出青年科学基金); the National Basic Research Program of China under Grant No.2002CB312000 (国家重点基础研究发展计划(973)); the High Technology Research Project of Jiangsu Province of China under Grant No.BG2005032 (江苏省高技术研究项目)

Received 2005-12-19; Accepted 2006-07-26

下的 $EFALC_{R+}$推理算法,并对该算法进行优化,证明优化后的算法是正确完备的,时间复杂性不超过指数,最后证明受限TBox约束下的$EFALC_{R+}$推理问题是指数时间完全问题.优化算法的最坏时间复杂性已达到该问题推理算法的复杂度下界,是实现术语公理约束下模糊知识库推理的有效算法.

关键词: 描述逻辑;模糊;TBox(terminological box);推理;语义 Web;知识表示

中图法分类号: TP301 文献标识码: A

描述逻辑是一种具有形式化语义、很强表达能力和可判定推理算法的知识表示语言,是语义 Web 表示知识和实现推理任务的逻辑基础[1].Web 应用中常常需要处理不确定的模糊知识.但描述逻辑仅支持精确知识的描述和推理,不适于处理模糊知识.Straccia 扩展描述逻辑 ALC(attributive concept description language with complements)提出模糊描述逻辑 FALC(fuzzy ALC),并给出推理算法[2].在此基础上,Hölldobler 等人引入模糊隶属度操纵算子[3],Sanchez 等人扩展模糊关系的数量约束[4],Straccia 引入模糊具体域[5].Stoilos 等人给出支持模糊关系的传递、蕴涵、逆和非限定数量约束的模糊描述逻辑的推理算法[6].国内学者近年来也开展了对描述逻辑扩展的研究.文献[7]提出一种新的动态描述逻辑,文献[8]在此基础上提出智能主体的心智状态模型.文献[9]提出基于区间模糊理论的描述逻辑系统.文献[10]扩展描述逻辑用于模糊 ER 模型.我们提出具有比模糊描述逻辑更强模糊知识表达能力的扩展模糊描述逻辑[11],给出支持数量约束的扩展模糊描述逻辑的推理算法[12].

描述逻辑知识库分为描述个体事实的 ABox(assertional box)和描述术语公理的 TBox(terminological box).个体事实是关于特定个体的具体知识,而术语公理则是不针对特定个体的一般性知识,它们结合起来才能完整地表示知识.但现有的模糊描述逻辑和扩展模糊描述逻辑推理算法都只能处理 ABox,而没有考虑 TBox 的约束.缺乏支持术语公理约束的推理算法制约了扩展模糊描述逻辑的应用.本文在扩展模糊描述逻辑中引入传递模糊关系得到 $EFALC_{R+}$(extended fuzzy attributive concept description language with complements and transitive roles),定义受限 TBox 描述术语公理,并给出受限 TBox 约束下的 $EFALC_{R+}$推理算法.

1 扩展模糊描述逻辑 $\mathbf{EFALC_{R+}}$

$EFALC_{R+}$引入模糊概念和模糊关系的截集作为原子概念和原子关系,语法上采用经典描述逻辑 ALC_{R+}的构造子构造概念和关系,并用模糊解释定义概念和关系的语义.

定义 1(**$\mathbf{EFALC_{R+}}$语法和语义**). 设个体名集 N_I、模糊概念名集 N_C和模糊关系名集 N_R为 3 个不相交的可数集,其中 N_R子集 N_{R+}为传递模糊关系名集.模糊解释是一个二元组 $I=(\Delta^I,\cdot^I)$,其中,论域Δ^I是一个非空集合,解释函数$\cdot^I$ 将每个 $a\in N_I$ 映射到Δ^I 中的元素 a^I,每个 $B\in N_C$ 映射到函数 $B^I:\Delta^I\to[0,1]$,每个 $R\in N_R$ 映射到函数 $R^I:\Delta^I\times\Delta^I\to[0,1]$.$EFALC_{R+}$的语法语义见表 1.

Table 1 Syntax and semantics of $EFALC_{R+}$

表 1 $EFALC_{R+}$的语法语义

Name	Syntax	Semantics
Atom concept	$B_{[n]}$	$\{x \mid x\in\Delta^I\wedge B^I(x)\geq n\}$
Atom role	$R_{[n]}$	$\{(x,y) \mid x,y\in\Delta^I\wedge R^I(x,y)\geq n\}$
Negation	$\neg C_{[n_1,\ldots,n_k]}$	$\Delta^I\setminus(C_{[n_1,\ldots,n_k]})^I$
Conjunction	$C_{[n_1,\ldots,n_h]}\sqcap D_{[n_{h+1},\ldots,n_k]}$	$(C_{[n_1,\ldots,n_h]})^I\cap(D_{[n_{h+1},\ldots,n_k]})^I$
Disjunction	$C_{[n_1,\ldots,n_h]}\sqcup D_{[n_{h+1},\ldots,n_k]}$	$(C_{[n_1,\ldots,n_h]})^I\cup(D_{[n_{h+1},\ldots,n_k]})^I$
Existential quantification	$\exists R_{[n_1]}.C_{[n_2,\ldots,n_k]}$	$\{x\in\Delta^I \mid \exists y\in\Delta^I, R^I(x,y)\geq n_1\wedge y\in(C_{[n_2,\ldots,n_k]})^I\}$
Value restriction	$\forall R_{[n_1]}.C_{[n_2,\ldots,n_k]}$	$\{x\in\Delta^I \mid \forall y\in\Delta^I, R^I(x,y)\geq n_1\to y\in(C_{[n_2,\ldots,n_k]})^I\}$

其中,$n_i\in(0,1)$,且对任何 $S\in N_{R+}$都有$(S_{[n]})^I=((S_{[n]})^I)^+$.$(\cdot)^+$是闭包运算,对任意 3 个元素 x,y,z 和二元关系 P 满足$(x,y)\in P\to(x,y)\in P^+$和$(x,y)\in P^+\wedge(y,z)\in P^+\to(x,z)\in P^+$.称 $C_{[n_1,\ldots,n_k]}$为概念,$R_{[n]}$为关系,n_i 为下标.概念常用下标向量缩写形式表示,如$\exists$friend$_{[0.7]}$.Strong$_{[0.9]}$可表示为$\exists$friend.Strong$_{[0.7,0.9]}$.

定义 2(**EFALC$_{R+}$**知识库). EFALC$_{R+}$知识库分为 TBox 和 ABox 两部分:TBox 是包含形如 $C_{[f_i(v),\ldots,f_j(v)]}\sqsubseteq D_{[f_l(v),\ldots,f_h(v)]}$,$v\in X$ 的术语公理的有限集合,其中,v 是变量,$X\subseteq(0,1)$是连通区间,$f_i(v)$等都是从 X 映射到(0,1)区间的线性函数.称 $C_{[f_i(v),\ldots,f_j(v)]}$ 这样含有变量的概念为可变概念.解释 I 满足 $C_{[f_i(v),\ldots,f_j(v)]}\sqsubseteq D_{[f_l(v),\ldots,f_h(v)]}$,$v\in X$ 当且仅当对任何常数 $n\in X$, $(C_{[f_i(n),\ldots,f_j(n)]})^I\subseteq(D_{[f_l(n),\ldots,f_h(n)]})^I$.解释 I 是 TBox T 的模型当且仅当 I 满足 T 中所有公理.ABox 是包含形如 a: $C_{[n_1,\ldots,n_k]}$ 的概念事实和形如(a,b):$R_{[n]}$的关系事实的有限集合,其中,$a,b\in N_I$.解释 I 满足 a: $C_{[n_1,\ldots,n_k]}$ 当且仅当 $a^I\in(C_{[n_1,\ldots,n_k]})^I$,满足$(a,b)$:$R_{[n]}$当且仅当$(a^I,b^I)\in(R_{[n]})^I$.解释 I 是 ABox A 的模型当且仅当 I 满足 A 中所有事实.

EFALC$_{R+}$知识库支持以下推理问题:(1) 概念可满足性:概念 $C_{[n_1,\ldots,n_k]}$ 关于 TBox T 可满足,当且仅当存在 T 的模型 I 满足$(C_{[n_1,\ldots,n_k]})^I\neq\varnothing$;(2) ABox 一致性:ABox A 关于 TBox T 是一致的,当且仅当存在 A 和 T 的模型 I.Abox 一致性可转化为概念可满足性问题.与文献[12]类似,EFALC$_{R+}$将概念可满足性扩展为可变概念可满足区间问题:在 TBox T 约束下,可变概念 $C_{[f_1(v),\ldots,f_k(v)]}$ 在 X 上的可满足区间为 $X_S\subseteq X$,当且仅当 $n\in X_S$ 时 $C_{[f_1(n),\ldots,f_k(n)]}$ 关于 T 可满足.一般的概念可视为下标都是常数函数的可变概念,概念可满足性是可满足区间问题的特例.

EFALC$_{R+}$ TBox 中公理可以包含变量和函数,有很强的表达能力,但难以进行推理.经典描述逻辑中通过将公理转化为概念实现 TBox 约束下的推理,如公理 $C\sqsubseteq D$ 可以等价表示为论域个体都是概念$\neg C\sqcup D$ 的实例.然而,这种方法无法用于包含变量和函数的公理.如公理 $VeryYoung_{[v]}\sqsubseteq Young_{[2v]}$,$v\in(0,0.5)$,对 v 的每个不同取值都可以转化为不同的概念,而 v 可能的取值数目是无限的.为此本文考虑受限 TBox,不允许在公理中出现变量:受限 TBox 定义为包含形如 $C_{[n_1,\ldots,n_i]}\sqsubseteq D_{[m_1,\ldots,m_j]}$ 的公理的有限集合.解释 I 满足 $C_{[n_1,\ldots,n_i]}\sqsubseteq D_{[m_1,\ldots,m_j]}$ 当且仅当 $(C_{[n_1,\ldots,n_i]})^I\subseteq(D_{[m_1,\ldots,m_j]})^I$.受限 TBox 足够描述常见的公理性知识,如"很喜欢运动的强壮的人会是不错的选手"可以表示为 $Strong_{[0.7]}\sqcap\exists likes_{[0.9]}.Sports_{[1]}\sqsubseteq GoodPlayer_{[0.8]}$.此外,还可以在一定程度上近似模拟非受限的 TBox.如含有变量的公理 $VeryYoung_{[v]}\sqsubseteq Young_{[2v]}$,$v\in(0,0.5)$ 可以转化为多个不含变量的公理:$VeryYoung_{[n_i]}\sqsubseteq Young_{[2n_i]}$,$i=\{1,2,\ldots,k\}$,其中,$n_i$ 为知识库中属于(0,0.5)的下标.下文中的 Tbox 均指受限 TBox.

2 可满足区间问题的推理算法

本文基于空 TBox 约束下 EFALCN 可满足区间推理算法[12]给出 TBox 约束下 EFALC$_{R+}$可满足区间问题的转化规则,新增了处理传递关系的$\forall^+$规则和处理 TBox 的 KB 规则,并引入阻塞条件保证算法的可终止性.对任意 ABox A 和区间 $X\subseteq[0,1]$,称形如(A,X)的二元对为可变 ABox.求 TBox 约束下 EFALC$_{R+}$概念 $C_{0[f_1(v),\ldots,f_k(v)]}$ 在 X_0 上可满足区间的算法从可变 Abox $(\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0)$出发,应用转化规则扩展可变 ABox,动态地将可变 ABox 的区间划分成不同的子区间,在每个子区间上分别检查可满足性.为此引入函数 comp(f,g,X)返回 X 的一个子区间 X'满足$\forall n\in X'$,$f(n)\le g(n)$;$\forall n\in X\backslash X'$,$f(n)>g(n)$.

为考虑 TBox 约束,引入概念 $C_T=\sqcap\{\neg C_{[n_1,\ldots,n_i]}\sqcup D_{[m_1,\ldots,m_j]}\mid C_{[n_1,\ldots,n_i]}\sqsubseteq D_{[m_1,\ldots,m_j]}\in T\}$.显然,一个解释 I 是 T 的模型当且仅当$\Delta^I=(C_T)^I$,即论域中所有元素都属于$(C_T)^I$.下面规定可变概念都是标准否定形式,即$\neg$符号仅出现在单个概念名之前.任何可变概念都可在多项式时间内改写为语义等价的标准否定形式,改写方法与经典情形相同[1].

定义 3(**EFALC$_{R+}$**转化规则). TBox T 约束下,对可变 Abox (A,X)中未被阻塞的个体 x 可应用以下转化规则:

1. $\neg$规则:如果$\{x:\neg B_{[f(v)]},x:B_{[g(v)]}\}\subseteq A$ 且 $X'\leftarrow$comp$(f,g,X)\neq\varnothing$,则令 $X\leftarrow X\backslash X'$.
2. $\sqcap$规则:如果 x: $C_{[f_i(v),\ldots,f_j(v)]}\sqcap D_{[f_l(v),\ldots,f_h(v)]}\in A$ 且 x: $C_{[f_i(v),\ldots,f_j(v)]}\notin A$ 或 x: $D_{[f_l(v),\ldots,f_h(v)]}\notin A$,则令
$$A\leftarrow A\cup\{x: C_{[f_i(v),\ldots,f_j(v)]}, x: D_{[f_l(v),\ldots,f_h(v)]}\}.$$
3. $\sqcup$规则:如果 x: $C_{[f_i(v),\ldots,f_j(v)]}\sqcup D_{[f_l(v),\ldots,f_h(v)]}\in A$ 且$\{x: C_{[f_i(v),\ldots,f_j(v)]}, x: D_{[f_l(v),\ldots,f_h(v)]}\}\cap A=\varnothing$,则令
$$A\leftarrow A\cup\{x: C_{[f_i(v),\ldots,f_j(v)]}\},\text{或者令 } A\leftarrow A\cup\{x: D_{[f_l(v),\ldots,f_h(v)]}\}.$$
4. $\exists$规则:如果 x: $\exists R_{[f(v)]}.C_{[f_i(v),\ldots,f_j(v)]}\in A$ 且 x 不存在 $R_{[f(v)]}$后继 z 满足 z: $C_{[f_i(v),\ldots,f_j(v)]}\in A$,则新建个体 y,令

$A \leftarrow A \cup \{(x,y): R_{[f(v)]}, y: C_{[f_i(v),\dots,f_j(v)]}\}$,称 y 是 x 的直接后继,且是 $R_{[f(v)]}$后继.

5. $\forall$规则:如果 $x: \forall R_{[f(v)]}.C_{[f_i(v),\dots,f_j(v)]} \in A$ 且存在 x 的 $R_{[g(v)]}$后继 y 满足 $y: C_{[f_i(v),\dots,f_j(v)]} \notin A$ 和 $X' \leftarrow \mathrm{comp}(f,g,X) \neq \varnothing$,则令 $A \leftarrow A \cup \{y: C_{[f_i(v),\dots,f_j(v)]}\}, X \leftarrow X'$;或者不改变 A 但令 $X \leftarrow X \backslash X'$.
6. $\forall^+$规则:如果 $x: \forall R_{[f(v)]}.C_{[f_i(v),\dots,f_j(v)]} \in A, R \in N_{R+}$且存在 x 的 $R_{[g(v)]}$后继 y 满足 $y: \forall R_{[f(v)]}.C_{[f_i(v),\dots,f_j(v)]} \notin A$ 和 $X' \leftarrow \mathrm{comp}(f,g,X) \neq \varnothing$,则令 $A \leftarrow A \cup \{y: \forall R_{[f(v)]}.C_{[f_i(v),\dots,f_j(v)]}\}, X \leftarrow X'$;或者不改变 A 但令 $X \leftarrow X \backslash X'$.
7. KB 规则:如果 $x: C_T \notin A$,则令 $A \leftarrow A \cup \{x: C_T\}$.

其中,$C_{[f_i(v),\dots,f_j(v)]}$ 和 $D_{[f_l(v),\dots,f_h(v)]}$ 是可变概念,$R_{[f(v)]}$和 $R_{[g(v)]}$称可变关系.应用规则时遵守如下条件:

- 在可能的执行动作不唯一时($\sqcup, \forall, \forall^+$规则),采用非确定性图灵机的执行方式,将自身复制为多个,每个执行其中一种动作.实际执行时可采用深度优先或广度优先遍历所有非确定分支.
- 对任意个体 $x, L(x)=\{C_{[f_i(v),\dots,f_j(v)]} \mid x: C_{[f_i(v),\dots,f_j(v)]} \in A\}$称为 x 在 A 中的标签.令个体间后继关系是直接后继关系的闭包.如果 x 的后继 y 满足 $L(y) \subseteq L(x)$,则称 y 被 x 阻塞.规定被阻塞个体的后继都是被阻塞的.不可对被阻塞的个体应用任何规则.
- 当 $X=\varnothing$时,不必再对(A,X)应用任何规则.

当任意规则都不能应用到可变 ABox (A,X)时,称(A,X)是完备的.由于存在非确定规则,可能得到多个完备的可变 ABox,将它们的区间合并即得到要求的可满足区间.

算法 **1**. TBox T 约束下 $\mathrm{EFALC_{R+}}$可变概念 $C_{0[f_1(v),\dots,f_k(v)]}$ 在 X_0 上的可满足区间问题算法.

输入:TBox T,可变概念 $C_{0[f_1(v),\dots,f_k(v)]}$,区间 X_0;

输出:可满足区间 X_S.

过程:初始化 $X_S \leftarrow \varnothing, A \leftarrow \{x_0: C_{0[f_1(v),\dots,f_k(v)]}\}, X \leftarrow X_0$;

在 T 约束下对(A,X)反复应用转化规则,直到(A,X)完备;

$X_S \leftarrow X_S \cup X$,停机.

当应用非确定规则时,算法 1 可分裂为多个过程,它们对同一个 X_S 和不同的(A,X)进行操作,只有在它们都停机时算法才结束.假设从$(\{x_0: C_{0[f_1(v),\dots,f_k(v)]}\}, X_0)$出发穷尽地应用规则,共可得到 m 个完备的$(A_1,X_1),\dots,(A_m,X_m)$,则最后返回的结果应为 $X_S=X_1 \cup \dots \cup X_m$.算法 1 的正确性和完备性证明类似于后文优化算法 2,但更简单.

定理 **1**. 算法 1 的时间复杂性为非确定二阶指数.

证明:本文复杂性描述(多项式、指数等)都是针对输入规模 $N=\|C_{0[f_1(v),\dots,f_k(v)]}\|+\|T\|$而言的.任意概念的规模(记作$\|\cdot\|$)为写下它们所需的符号串长度,其中写下任何一个个体名、概念名、关系名和构造子($\neg, \sqcap, \sqcup, \exists, \forall$)都占用长度为 1 的符号串.如$\exists$friend.Strong$_{[0.7,0.9]}$的规模为 3.TBox 的规模等于其中概念的规模之和.对任意概念 $C_{[f_i(v),\dots,f_j(v)]}$,定义其子概念集[12]为 $\mathrm{sub}(C_{[f_i(v),\dots,f_j(v)]})$.显然,对任意 $C_{[f_i(v),\dots,f_j(v)]}$,$|\mathrm{sub}(C_{[f_i(v),\dots,f_j(v)]})| \le \|C_{[f_i(v),\dots,f_j(v)]}\|$.对于任何个体的标签 L,由定义 3 可知,总有 $L \subseteq \mathrm{sub}(C_{0[f_1(v),\dots,f_k(v)]}) \cup \mathrm{sub}(C_T)$.由于$\|C_T\| \le 3\|T\|$,可知$|L|$的上界是 N 的多项式,记作 M.不同的标签最多不超过 2^M 个.由于对任意 $\mathrm{comp}(f,g,X)$,f 和 g 只能在最多 N 个线性函数中选取,因此,算法中的 comp 函数只能将 X_0 划分为不超过 $1+N^2$ 个不可再划分的子连通区间,记这些子区间的集合为 $\mathrm{sub}(X_0)$.记 $H=|\mathrm{sub}(X_0)| \le 1+N^2$.算法中出现的区间都是 $\mathrm{sub}(X_0)$中的元素或它们的合并,不同区间总数不超过 2^H.

对于算法中任意 ABox A,以 x_0 为根节点,其中,所有个体按直接后继关系可构成一棵树,称为个体树.由于不同的标签最多不超过 2^M,当个体树的任一分支长度超过 2^M 时,其上一定存在两个个体的标签相同;由阻塞条件可知,叶子节点无法应用任何规则,则该分支不可能新增任何个体.因此,个体树的深度上界是 $1+2^M$.由于个体的直接后继只能通过$\exists$规则产生,其数目不会超过 M,可知个体树的出度不超过 M.因此,A 中个体数不超过 $1+M^{2^M}$.因为每个个体最多应用 M 次规则,一个可变 ABox 应用规则不超过$(1+M^{2^M}) \times M$ 次,是二阶指数.算法 1 非确定地对可变 ABox (A,X)应用规则,时间复杂性是非确定二阶指数. □

经典 $\mathrm{ALC_{R+}}$在 TBox 约束下的概念可满足性问题是指数时间完全的.文献[13]提出 TBox 约束下 ALC 概念

可满足性问题的指数时间算法,其中的优化技术经过扩展后可用于优化算法 1.

3 可满足区间问题的优化算法和复杂性证明

3.1 可满足区间问题的优化算法

为优化算法的时间复杂性,对算法 1 进行 3 点主要改进得到算法 2:(1) 按深度优先顺序生成个体树及处理$\sqcup,\forall,\forall^+$规则中的非确定分支;(2) 裁减$\sqcup$分支:对 $C_{[f_i(v),\dots,f_j(v)]}\sqcup D_{[f_l(v),\dots,f_h(v)]}$,先计算 $C_{[f_i(v),\dots,f_j(v)]}$ 的可满足区间,如果它在 x_i 上可满足,则 $C_{[f_i(v),\dots,f_j(v)]}\sqcup D_{[f_l(v),\dots,f_h(v)]}$ 在 x_i 上可满足,不必再计算 $D_{[f_l(v),\dots,f_h(v)]}$ 在 x_i 上的可满足性;(3) 用两个集合 unsat 和 tested 记录中间结果,避免重复计算.为方便描述深度优先顺序和分析复杂性,算法 2 主体描述为递归过程 sat.每一次 sat 调用对可变 ABox 中的一个个体应用规则,直到区间或该个体的标签发生改变为止.

算法 **2**. TBox T 约束下 $EFALC_{R+}$可变概念 $C_{0[f_1(v),\dots,f_k(v)]}$ 在 X_0 上可满足区间问题的优化算法.

输入:TBox T,可变概念 $C_{0[f_1(v),\dots,f_k(v)]}$,区间 X_0;

输出:可满足区间 X_S.

过程:初始化两个集合:unsat←∅,tested←∅;返回 $X_S\leftarrow sat(x_0,\{x_0: C_{0[f_1(v),\dots,f_k(v)]}\},X_0)$.

过程 **sat**

输入:个体 x,ABox A,区间 X;

输出:区间 X.

过程:令 $X_f\leftarrow X, L\leftarrow L(x)$,其中,$L(x)$是 x 在 A 中的标签.

1. 如果 $X=\varnothing$或(A,X)中个体 x 被阻塞或 $X\subseteq tested(L)$,则称本次调用被阻止,返回 $X\backslash unsat(L)$
2. 如果 $X\cap unsat(L)\neq\varnothing$,则令 $X\leftarrow sat(x,A,X\backslash unsat(L))$,称调用被 $sat(x,A,X\backslash unsat(L))$替代,转到第 6 步
3. 如果在 T 约束下可对(A,X)中的个体 x 应用规则,则按以下优先级顺序任意应用一条规则:

   $\neg,\sqcap$,KB 规则:记(A_1,X_1)为应用规则后的可变 ABox,令 $X\leftarrow sat(x,A_1,X_1)$,转到第 5 步

   $\forall,\forall^+$规则:记规则产生的两个可变 ABox 为(A_1,X_1)和$(A,X\backslash X_1)$,

   如果 $X\backslash X_1=\varnothing$,则令 $A\leftarrow A_1$,重新执行第 3 步

   否则令 $X\leftarrow sat(x,A_1,X_1)\cup sat(x,A,X\backslash X_1)$,转到第 5 步

   $\sqcup$规则:记规则产生的两个 ABox 为 A_1 和 A_2,令 temp←tested,$X_1\leftarrow sat(x,A_1,X)$;对任何标签 G,

   如果$(G,X_i)\in tested\backslash temp, X_i\subseteq X_1$,则在 temp 中添加$(G,X_i)$;然后令 tested←temp,

   $X_2\leftarrow sat(x,A_2,X\backslash X_1)$;最后令 $X\leftarrow X_1\cup X_2$,转到第 5 步

   $\exists$规则:令(A,X)为应用规则后的可变 ABox,重新执行第 3 步
4. 如果不可对(A,X)中的个体 x 应用任何规则,且 x 没有被阻塞,记 x 在 A 中的直接后继为 $y_1,\dots,y_m$,则执行循环 for (i=1 to m) $X\leftarrow sat(y_i,A,X)$;每一次循环以上一次输出的 X 值作为输入的 X 值
5. 对任何 $X_i\in sub(X_0)$,如果 $X_i\subseteq X_f\backslash X$,则 $unsat\leftarrow unsat\cup\{(L,X_i)\}$
6. 对任何 $X_i\in sub(X_0)$,如果 $X_i\subseteq X_f$,则 $tested\leftarrow tested\cup\{(L,X_i)\}$
7. 返回 X

过程中,A 和 X 是 sat 过程的局部变量.集合 tested 和 unsat 是算法 2 的全局变量,包含形如(L,X_i)的元素.记 $unsat(L)=\cup\{X_i|(L,X_i)\in unsat\}$,$tested(L)=\cup\{X_i|(L,X_i)\in tested\}$.下面首先分析算法复杂性,然后证明完备性和正确性.

3.2 算法分析

3.2.1 可终止性与复杂性

引理 **1**(可终止性). 对任何有限规模的输入,算法 2 在有限时间内终止.

证明:称运行形如 $sat(x,A,X)$的语句为一次调用,x 为作用个体,X 为作用区间,L 为作用标签;调用中产生的新

调用称为直接子调用;子调用关系是直接子调用关系的闭包.以 $\mathrm{sat}(x_0,\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0)$为根节点,它及其子调用按直接子调用关系构成一棵树,记作调用树 PT.对于算法中任意 ABox A,以 x_0 为根节点的个体树的深度上界是 $1+2^M$,出度不超过 M.而且由第 4 步可知,一个个体最多只有一个直接后继在 A 中有后继.因此,A 中个体数不会超过 $1+2^M\times M$.对 PT 中任意节点,它的直接子调用数目不会超过 2(产生于第 3 步)或作用个体的直接后继数目(产生于第 4 步),即 PT 的出度不超过 $\max(2,M)$.对 PT 中任意调用 $\mathrm{sat}(x,A,X)$,它的直接子调用满足:如果作用个体不变,则作用标签内容增加或作用区间减小,否则作用个体变为 x 的直接后继.因此,PT 的深度上界为 $(M+H)\times(2^M+1)$.由于深度和出度都是有限的,PT 中的调用次数是有限的.每次调用至多仅能对一个个体 x 应用总共不超过 M^2 次规则;最多有 $\max(2,M)$个直接子调用.若不考虑子调用的内部耗费,每次调用的时间耗费不超过指数.因此,算法 2 执行有限次的、耗费时间有限的调用,是可终止的. □

引入 tested 和 unsat 可将调用次数控制在指数.由第 1 步,如果 tested 单调增加,则可防止重复执行具有同样作用标签和作用区间的调用,使得最多只有指数次调用.但是,tested 在第 3 步$\sqcup$规则可能被删除内容.unsat 用来记录不同标签的已知不可满足区间,在整个算法执行中只能增加.引理 2 将证明 unsat 约束了对 tested 的删除操作次数是指数,又因为两次删除操作之间的调用最多为指数次,因此算法 2 最多执行指数次调用.

引理 **2**. 算法 2 最多执行指数次调用.

证明:根据是否删除 tested 内容,PT 中的调用分为两类:一类称为分离调用,它从 tested 集合中删除内容;其他调用都称为连续调用.分离调用一定应用$\sqcup$规则.对 PT 中任意应用$\sqcup$规则的 $\mathrm{sat}(x,A,X)$,称 $\mathrm{sat}(x,A_1,X)$为左儿子,$\mathrm{sat}(x,A_2,X\backslash X_1)$为右儿子,作用标签分别是 L_1 和 L_2,$X_1=\mathrm{sat}(x,A_1,X)$.显然,$X\backslash X_1\subseteq\mathrm{unsat}(L_1(x))$.假设 $\mathrm{sat}(x,A_1,X)$没有在 unsat 中增加任何形如$(L_1(x),X_i)$的元素,根据第 2 步,$\mathrm{sat}(x,A_1,X)$或者被阻止或者满足 $X=X_1$ 或者被 $\mathrm{sat}(x,A_1,X_1)$替代.$\mathrm{sat}(x,A_1,X_1)$及其子调用只能在 tested 中增加满足 $X_i\subseteq X_1$ 的形如(G,X_i)的元素.由第 3 步$\sqcup$规则,这些元素都没有被删除,$\mathrm{sat}(x,A,X)$不是分离调用.因此对任意应用$\sqcup$规则的调用 $\mathrm{sat}(x,A,X)$,它为分离调用仅当左儿子在 unsat 中添加了元素.unsat 在整个算法执行中只能增加,且$|\mathrm{unsat}|$的上界为 $2^M\times H$. PT 中分离调用的数目不会超过 $2^M\times H$.

按以下方法将调用树 PT 划分为调用森林 PF:首先以 PT 为当前树,从根节点出发按实际调用执行的顺序遍历当前树,每次遇到一个分离调用,则将以左儿子为根的子树与当前树分离.然后对所有独立出来的子树都穷尽地做同样的操作.最后得到调用森林 PF.图 1 给出了一个划分例子,其中,实心节点表示分离调用,空心节点表示连续调用,短划线表示子树分离的位置.PT 中的每一次调用都在 PF 中,PF 中调用树的个数等于 PT 中分离节点个数加 1,至多为 $2^M\times H+1$.

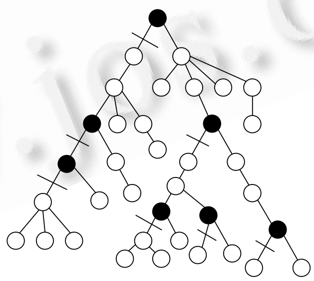

Fig.1 Partition of PT

图 1 调用树 PT 的划分

对于任意 $PT_i\in PF$,其中,原 PT 中分离调用删除的 tested 内容必是在以左儿子为根的子树中添加的,现在该子树已不在 PT_i 中,因此,PT_i 中的所有调用都可视作没有从 tested 集合中删除内容.对 PT_i 中任意两个调用 $\mathrm{sat}(x,A,X)$和 $\mathrm{sat}(y,A',X')$满足$(L,X)=(L',X')$:不妨设 $\mathrm{sat}(x,A,X)$先于 $\mathrm{sat}(y,A',X')$产生,

- 如果 $\mathrm{sat}(y,A',X')$是 $\mathrm{sat}(x,A,X)$的子调用.由于对任意调用 $\mathrm{sat}(x,A,X)$,分析 sat 过程可知,它的子调用 $\mathrm{sat}(y,A',X')$一定满足 y 为 x 在 A'中的后继,且 A'中的 x 标签 $L'(x)\supseteq L$;此时,y 在 A'中被 x 阻塞,不可能产生

子调用.

- 如果它们之间无子调用关系,由深度优先顺序,sat(y,A',X')一定在 sat(x,A,X)退出后才产生,根据第 6 步,sat(x,A,X)退出时一定已有 $X\subseteq$tested(L),则 sat(y,A',X')被阻止,不可能产生子调用.

由此可知,PT_i 中作用标签和作用区间相同的调用中仅有一个可能产生子调用.不同的作用标签数目不会超过 2^M,而作用区间的可能取值最多是 2^H.PT_i 中除了根节点外,任何调用都是其他调用的子调用.引理 1 中已知任意调用的直接子调用数目不超过 max(2,M).因此,对于任意 $PT_i\in PF$,PT_i 中最多只有 1+max(2,M)$\times 2^M\times 2^H$ 次调用.

由于 PF 中调用树数目和每棵调用树中的调用次数都是指数,因此 PT 中调用次数不超过指数. □

引理 1 中已证,任何一次调用都可在指数时间内完成,引理 2 保证算法 2 最多执行指数次调用.直接得到:

定理 **2**(复杂性). 算法 2 可在输入规模的指数时间内完成.

3.2.2 完备性和正确性

引理 **3**(完备性). 对任何 $X_i\in$sub(X_0),如果 $X_i\not\subseteq$ sat($x_0,\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0$),则 $C_{0[f_1(v),\ldots,f_k(v)]}$ 在 X_i 上不可满足.

证明:调用的返回区间 X 和作用区间 X_f 满足 $X_i\subseteq X_f\backslash X$,则称该调用排除 X_i.由第 5 步,(L,X_i)$\in$unsat 当且仅当存在作用标签为 L 的调用排除 X_i.对任何标签 L,如果对常数 $n\in X_i$ 有 T 的模型 I 使得

$$\cap\{(C_{[f_i(n),\ldots,f_j(n)]})^I \mid C_{[f_i(n),\ldots,f_j(n)]}\in L\}\neq\varnothing,$$

则称 I 是 L 的 n-模型.如果对任何 $n\in X_i$,都不存在 L 的 n-模型,则称 L 在区间 X_i 上不可满足.

下面用归纳法证明命题:对任何(L,X_i)$\in$unsat,L 在区间 X_i 上不可满足.

(i) 当 unsat=$\{(L,X_i)\}$时,一定存在$\{\neg B_{[f(v)]},B_{[g(v)]}\}\subseteq L$ 且 $X_i\subseteq$comp(f,g,X).对任何 $n\in X_i$,有 $f(n)\le g(n)$,不存在解释 I 使得$(\neg B_{[f(n)]})^I\cap(B_{[g(n)]})^I\neq\varnothing$.因此,$L$ 在区间 X_i 上不可满足.假设当|unsat|=r 时,命题成立.

(ii) 当|unsat|=r+1 时,在 unsat 中新添加的(L,X_i)都有 L 在区间 X_i 上不可满足.分情况讨论:

情形(a),应用$\neg$规则,发现$\{\neg B_{[f(v)]},B_{[g(v)]}\}\subseteq L$,且 $X_i\subseteq$comp(f,g,X).类似(i)可证 L 在区间 X_i 上不可满足.

情形(b),应用$\sqcap$,$\sqcup$和 KB 规则,产生子调用排除 X_i.设子调用的作用标签为 L',则 L'在区间 X_i上不可满足.易证,无论应用$\sqcap$,$\sqcup$和 KB 规则,都有 L 在区间 X_i 上也不可满足.以 KB 规则为例,假设存在 $n\in X_i$ 使得 L 具有 n-模型 I,则$\Delta^I=(C_T)^I$;KB 规则仅在 L 中增加 C_T 得到 L',显然,I 是 L'的 n-模型,和 L'在 X_i 上不可满足相矛盾.当应用$\neg$,$\forall$,$\forall^+$规则产生子调用时作用标签不变,因此,如果子调用排除 X_i,则已有(L,X_i)$\in$unsat,不可能在 unsat 中新添加(L,X_i).

情形(c),第 4 步中存在子调用排除 X_i.令子调用的作用标签为 L',L'是 L 通过$\exists$,$\forall$,$\forall^+$规则得到的,由 L'在区间 X_i 上不可满足,可证 L 在区间 X_i 上也不可满足.这里仅以证$\forall^+$规则为例:由规则定义,A 中必存在

$$x\colon \forall R_{[f(v)]}.C_{[f_i(v),\ldots,f_j(v)]},R\in N_{R+},(x,y)\colon R_{[g(v)]},\text{且 } X_i\subseteq\text{comp}(f,g,X).$$

假设存在 $n\in X_i$ 使得 L 具有 n-模型 I,则存在 $x\in(\forall R_{[f(n)]}.C_{[f_i(n),\ldots,f_j(n)]})^I$.对 I 中任意 y 满足$(x,y)\in(R_{[g(n)]})^I$.由 $f(n)\le g(n)$,$(x,y)\in(R_{[g(n)]})^I\subseteq(R_{[f(n)]})^I$;$R\in N_{R+}$,则$(R_{[f(n)]})^I=((R_{[f(n)]})^I)^+$;因为 x 满足$\forall z,(x,z)\in((R_{[f(n)]})^I)^+\to z\in(C_{[f_i(n),\ldots,f_j(n)]})^I$,即满足$\forall y\forall z,(x,y)\in(R_{[f(n)]})^I\wedge(y,z)\in((R_{[f(n)]})^I)^+\to z\in(C_{[f_i(n),\ldots,f_j(n)]})^I$,因此,$y$ 仍然满足$\forall z,(y,z)\in((R_{[f(n)]})^I)^+\to z\in(C_{[f_i(n),\ldots,f_j(n)]})^I$,由定义 1,$y\in(\forall R_{[f(n)]}.C_{[f_i(n),\ldots,f_j(n)]})^I$.$\forall^+$规则仅在 y 的标签 L'中添加 $\forall R_{[f(v)]}.C_{[f_i(v),\ldots,f_j(v)]}$,显然,$I$ 也是 L'的 n-模型,矛盾.因此,L 在区间 X_i 上不可满足.

由(i)、(ii)归纳可得,对任何(L,X_i)$\in$unsat,L 在区间 X_i 上不可满足.如果 $X_i\not\subseteq$ sat($x_0,\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0$),则有($\{C_{0[f_1(v),\ldots,f_k(v)]}\},X_i$)$\in$unsat, $C_{0[f_1(v),\ldots,f_k(v)]}$ 在区间 X_i 上不可满足. □

引理 **4**(正确性). 对任意 $X_i\in$sub(X_0),如果 $X_i\subseteq$sat($x_0,\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0$),则对任意 $n\in X_i$ 可以构造 T 的一个模型,使得 $C_{0[f_1(n),\ldots,f_k(n)]}$ 解释不为空,即 X_i 属于 $C_{0[f_1(v),\ldots,f_k(v)]}$ 在 X_0 上的可满足区间.

证明:若调用的返回区间包含 X_i,则称它支持 X_i.按如下方法在 PT 中选取支持 X_i 的调用:首先选取 sat($x_0,\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_0$)为支持调用;如果支持调用只有一个直接子调用或有多个由第 4 步生成的直接子调用,则令直接子调用都为支持调用;如果支持调用有两个由第 3 步生成的直接子调用,则仅可能有一个支持 X_i,令它为支持调用;直到新的支持调用都不存在子调用.最后将所有支持调用作用的 ABox 合并,记作 A.可知(A,X_i)是

从$(\{x_0: C_{0[f_1(v),\ldots,f_k(v)]}\},X_i)$出发,按照支持调用的执行顺序应用规则和选择非确定分支得到的可变 ABox.A 不一定是完备的,但所有还可以作用规则的个体都是被阻止的支持调用的作用个体.扩展定义 3 中的阻塞定义:如果作用个体为 x 的调用在 tested 中添加(L,X_i),而另一个作用标签为 L 的支持调用因为 tested 被阻止,则称它的作用个体 y 被 x 阻塞.将 A 中个体按直接后继关系构成个体树,可以发现,定义 3 中的阻塞发生在个体树的一条分支上的个体之间,而扩展后的阻塞还可以发生在不同分支上的个体之间.在构造模型时,所有被阻塞的个体都不包括在论域中,它们的作用被阻塞它们的个体所代替.图 2 给出了一个简单的处理阻塞个体的例子.

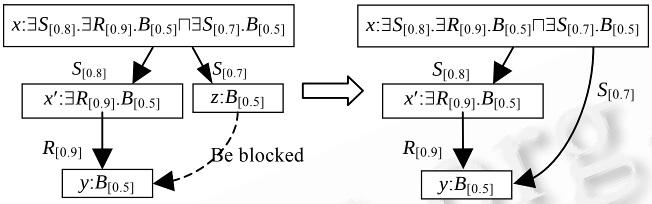

Fig.2 A simple example of handling blocked individuals

图 2 处理被阻塞个体的简单例子

但采用这种处理方法需要满足一个前提条件:A 中任何被阻塞的个体 y,都存在 A 中个体 x 阻塞 y,且 $L(y)\subseteq L(x)$.对于满足定义 3 中阻塞定义的被阻塞个体,显然满足前提.如果支持调用 sat(y,A',X')由于 tested(L)而被阻止,则一定存在一个调用将(L,X_i)添加到 tested.假设有一个非支持调用将(L,X_i)添加到 tested 且在 y 被阻塞前不被删除.分析选择过程可知,所有非支持调用都是由第 3 步生成的两个直接子调用之一为根的子树(称为非支持子树)中的调用,分 3 种情况:(i) 如果是应用∀和$\forall^+$规则,两个直接子调用作用区间不相交,则非支持子树中任何调用的作用区间不包括 X_i,不可能在 tested 中添加(L,X_i);(ii) 如果是应用⊔规则,且非支持子树的根是右儿子,则有 $X_i\subseteq X_1,X_i\cap X\backslash X_1=\varnothing$,非支持子树中的任何调用的作用区间同样不包括 X_i;(iii) 如果是应用⊔规则,且非支持子树的根是左儿子,则非支持子树中的调用可能在 tested 中添加(L,X_i);但因为 $X_i\subseteq X\backslash X_1,X_i\cap X_1=\varnothing$,$(L,X_i)$在非支持子树执行完毕后即被删除,不可能影响到 sat(y,A',X').综上所述,假设不成立.因此一定存在一个支持调用将(L,X_i)添加到 tested,令它的作用个体为 x,则 y 被 x 阻塞且满足 $L(y)\subseteq L(x)$.

对任意常数 $n\in X_i$,将 n 代入 A 中变量 v,将可变概念转化为一般概念,得到的 ABox 仍记作 A.因为 n 是常数,以下 $f_i(n)$等也都是常数.通过 A 可构造模型 $I=(\Delta^I,\cdot^I)$,Δ^I 为 A 中所有未被阻塞的个体构成的集合;对任意 $x\in\Delta^I$ 和模糊概念名 B,定义 $B^I(x)=\max(\{0\}\cup\{f_i(n)|\ B_{[f_i(n)]}\in L(x)\})$;对任意 $x,y\in\Delta^I$,模糊关系名 R,如果 $R\in N_{R+}$,则定义 $R^I(x,y)=\max(\{0\}\cup\{f_i(n)|(x,y)\in(\pi R_{[f_i(n)]})^+\})$,否则定义 $R^I(x,y)=\max(\{0\}\cup\{f_i(n)|(x,y)\in\pi R_{[f_i(n)]}\})$,其中,

$$\pi R_{[f(n)]}=\{(x,y)|(x,y):R_{[g(n)]}\in A\text{ 或}(x,z):R_{[g(n)]}\in A\text{ 满足 }z\text{ 被 }y\text{ 阻塞},f(n)\le g(n)\}.$$

对任意截关系 $R_{[f(n)]},x,y\in\Delta^I$,如果$(x,y):R_{[f(n)]}\in A$,则 $R^I(x,y)\ge f(n),(x,y)\in(R_{[f(n)]})^I$.如果 $R\notin N_{R+}$,则$(R_{[f(n)]})^I=\pi R_{[f(n)]}$,否则,$(R_{[f(n)]})^I=(\pi R_{[f(n)]})^+=((R_{[f(n)]})^I)^+$.因此,$I$ 满足 A 中所有关系事实和传递关系约束.

下面用归纳法证明对任意概念 $C_{[f_i(n),\ldots,f_j(n)]}$,$x\in\Delta^I$,如果 $C_{[f_i(n),\ldots,f_j(n)]}\in L(x)$,则 $x\in(C_{[f_i(n),\ldots,f_j(n)]})^I$:(i) 当$\|C_{[f_i(n),\ldots,f_j(n)]}\|=1$ 时,$C_{[f_i(n),\ldots,f_j(n)]}=B_{[f(n)]}$,如果 $B_{[f(n)]}\in L(x)$,则 $B^I(x)\ge f(n),x\in(B_{[f(n)]})^I$.当$\|C_{[f_i(n),\ldots,f_j(n)]}\|=2$ 时,$C_{[f_i(n),\ldots,f_j(n)]}=\neg B_{[f(n)]}$,如果$\neg B_{[f(n)]}\in L(x)$,则 $B^I(x)<f(n),x\in(\neg B_{[f(n)]})^I$.假设对于任何规模小于 r 的 $C_{[f_i(n),\ldots,f_j(n)]}$,$x\in\Delta^I$,都有 $C_{[f_i(n),\ldots,f_j(n)]}\in L(x)\rightarrow x\in(C_{[f_i(n),\ldots,f_j(n)]})^I$.(ii) 对任何 $C_{[f_i(n),\ldots,f_j(n)]}$ 满足$\|C_{[f_i(n),\ldots,f_j(n)]}\|=r$,仅有 $D_{[f_i(n),\ldots,f_l(n)]}\sqcap E_{[f_{l+1}(n),\ldots,f_j(n)]}$,$D_{[f_i(n),\ldots,f_l(n)]}\sqcup E_{[f_{l+1}(n),\ldots,f_j(n)]}$,$\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),\ldots,f_j(n)]}$ 和 $\exists R_{[f_i(n)]}.D_{[f_{i+1}(n),\ldots,f_j(n)]}$ 这 4 种情形.这里仅以 $\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),\ldots,f_j(n)]}$ 为例来作证明:因为 $\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),\ldots,f_j(n)]}\in L(x)$,如果 $R\notin N_{R+}$,则由∀规则对任何 y 满足$(x,y):R_{[g(n)]}\in A,g(n)\ge f_i(n)$都有 $y: D_{[f_{i+1}(n),\ldots,f_j(n)]}\in A$;对于任何 y 阻塞的个体 z 满足$(x,z):R_{[g(n)]}\in A,g(n)\ge f_i(n)$都有 $z: D_{[f_{i+1}(n),\ldots,f_j(n)]}$,因为 $L(z)\subseteq L(y)$,也有 $y: D_{[f_{i+1}(n),\ldots,f_j(n)]}$;由归纳假设,对任意$(x,y)\in\pi R_{[f_i(n)]}=(R_{[f_i(n)]})^I$ 都有 $y\in$

$(D_{[f_{i+1}(n),...,f_j(n)]})^I$,由定义 1,$x\in(\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),...,f_j(n)]})^I$.如果 $R\in N_{R+}$,对任意 y 满足$(x,y)\in(R_{[f_i(n)]})^I=(\pi R_{[f_i(n)]})^+$,一定存在序列 $x_0,x_1,x_2,...,x_h$ 满足 $0\le l<h,(x_l,x_{l+1})\in\pi R_{[f_i(n)]}$,$x_0=x,x_h=y$.由$\forall^+$规则和阻塞条件,对任何 $0\le l\le h$ 都有,$\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),...,f_j(n)]}\in L(x_l)$.又因为 $\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),...,f_j(n)]}\in L(x_{h-1})$和$(x_{h-1},y)\in\pi R_{[f_i(n)]}$,由$\forall$规则和阻塞条件,$D_{[f_{i+1}(n),...,f_j(n)]}\in L(y),y\in(D_{[f_{i+1}(n),...,f_j(n)]})^I$.因此,对任意 y 满足$(x,y)\in(R_{[f_i(n)]})^I$都有 $y\in(D_{[f_{i+1}(n),...,f_j(n)]})^I$,即 $x\in(\forall R_{[f_i(n)]}.D_{[f_{i+1}(n),...,f_j(n)]})^I$.

由于 x_0 不可能被阻塞,$x_0\in\Delta^I$ 且 $C_{0[f_1(n),...,f_k(n)]}\in L(x_0),x_0\in(C_{0[f_1(n),...,f_k(n)]})^I$.根据 KB 规则可知,对任意 $x\in\Delta^I,C_T\in L(x),x\in(C_T)^I$,即$\Delta^I=(C_T)^I$.因此,$I$ 是 T 的模型且使得$(C_{0[f_1(n),...,f_k(n)]})^I\neq\varnothing$. □

由引理 3 和引理 4 得到:对任何输入 T, $C_{0[f_1(v),...,f_k(v)]}$和 X_0,算法 2 返回 X_S 当且仅当 X_S 是 T 约束下 $C_{0[f_1(v),...,f_k(v)]}$在 X_0 上的可满足区间.再由定理 2 可知:

定理 3. 算法 2 是 EFALC$_{R+}$概念可满足区间问题的指数时间算法.

任何经典 ALC$_{R+}$知识库 K 可转化为等价的 EFALC$_{R+}$知识库 K':从经典到模糊的转化是平凡的,只需将 K 中概念名 B 和关系名 R 替换为概念 $B_{[1]}$和截关系 $R_{[1]}$,且转化得到的 TBox 不会含有变量,是受限的.已知 TBox 约束下 ALC$_{R+}$概念满足性是指数时间完全问题[1],且可多项式归约到受限 TBox 约束下 EFALC$_{R+}$概念可满足区间问题,因此,后者是指数时间难问题.由定理 3,后者存在指数时间算法(算法 2),从而得到:

定理 4. 受限 TBox 约束下的 EFALC$_{R+}$概念可满足区间问题是指数时间完全的.

4 结 论

本文考虑受限 TBox 约束下扩展模糊描述逻辑 EFALC$_{R+}$的推理,给出推理算法并进行优化,得到最坏时间复杂性为指数的优化推理算法(算法 2),并证明该算法的正确性和完备性.最后证明受限 TBox 约束下的 EFALC$_{R+}$概念可满足区间问题是指数时间完全的.算法 2 的最坏时间复杂性理论上已达到该问题推理算法的复杂性下界,是实现对术语公理约束下模糊知识库推理的有效算法.今后的研究工作是,一方面将在扩展模糊描述逻辑中引入更多算子增强表达能力,研究相应的推理算法;另一方面将考虑在含变量和函数的非受限 TBox 约束下的推理算法.

References:

[1] Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider PF. The description logic handbook: Theory, implementation and applications. Cambridge: Cambridge University Press, 2003.

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

[3] Hölldobler S, Störr HP, Khang TD. The fuzzy description logic ALCFH with hedge algebras as concept modifiers. Int’l Journal of Advanced Computational Intelligence and Intelligent Informatics, 2003,7(3):294−305.

[4] Sanchez D, Tettamanzi G. Generalizing quantification in fuzzy description logic. In: Proc. of the 8th Fuzzy Days. Dortmund, 2004. http://mago.crema.unimi.it/pub/SanchezTettamanzi2004.pdf

[5] Straccia U. Description logics with fuzzy concrete domains. In: Proc. of the 21st Conf. on Uncertainty in Artificial Intelligence. Edinburgh, 2005. http://faure.isti.cnr.it/~straccia/download/papers/UAI05/UAI05.pdf

[6] Stoilos G, Stamou G, Tzouvaras V, Pan J, Horrocks I. The fuzzy description logic f-SHIN. In: da Costa PCG, Laskey KB, Laskey KJ, Pool M, eds. Proc. of the Int’l Workshop on Uncertainty Reasoning for the Semantic Web. Galway, 2005. 67−76.

[7] Dong MK, Zhang HJ, Shi ZZ. An Agent model based on dynamic description logic. Journal of Computer Reaearch and Development, 2004,41(5):780−786 (in Chinese with English abstract).

[8] Shi ZZ, Jiang YC, Zhang HJ, Dong MK. Agent service matchmaking based on description logic. Chinese Journal of Computers, 2004,27(5):625−634 (in Chinese with English abstract).

[9] Hu H, Du XY. Description logics based on interval fuzzy theory. Journal of Huazhong University of Science and Technology (Nature Science), 2005,33(z1):275−277 (in Chinese with English abstract).

[10] Jiang YC, Tang Y, Wang J. Fuzzy ER modeling with description logics. Journal of Software, 2006,17(1):20−30 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/17/20.htm

[11] Li YH, Xu BW, Lu JJ, Kang DZ, Wang P. A family of extended fuzzy description logics. In: Proc. of the 29th Annual Int'l Computer Software and Applications Conf., Vol. 1. Edinburgh: IEEE Computer Society Press, 2005. 221−226. http://www.computer.org/portal/site/store/menuitem.41cf17dc879177c86ee948ce8bcd45f3/index.jsp?&pName=store_level1&path=store/p2005&file=p2413.xml&xsl=generic.xsl&

[12] Li YH, Xu BW, Lu JJ, Kang DZ. On the computational complexity of the extended fuzzy description logic with numerical constraints. Journal of Software, 2006,17(5):968−975 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/17/968.htm

[13] Donini FM, Massacci F. EXPTIME tableaux for ALC. Artificial Intelligence, 2000,124(1):87−138.

附中文参考文献:

[7] 董明楷,张海俊,史忠植.基于动态描述逻辑的主体模型.计算机研究与发展,2004,41(5):780−786.

[8] 史忠植,蒋运承,张海俊,董明楷.基于描述逻辑的主体服务匹配.计算机学报,2004,27(5):625−634.

[9] 胡鹤,杜小勇.一种基于区间模糊理论的描述逻辑系统.华中科技大学学报(自然科学版),2005,33(z1):275−277.

[10] 蒋运承,汤庸,王驹.基于描述逻辑的模糊 ER 模型.软件学报,2006,17(1):20−30. http://www.jos.org.cn/1000-9825/17/20.htm

[12] 李言辉,徐宝文,陆建江,康达周.支持数量约束的扩展模糊描述逻辑.软件学报,2006,17(5):968−975. http://www.jos.org.cn/1000-9825/17/968.htm

康达周(1980－),男,江苏南京人,博士生,主要研究领域为语义 Web.

陆建江(1968－),男,博士,副教授,主要研究领域为语义 Web,数据挖掘.

徐宝文(1961－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为软件工程,知识与信息获取技术.

李言辉(1981－),男,博士生,主要研究领域为语义 Web.