ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.19, No.10, October 2008, pp.2483−2497
DOI: 10.3724/SP.J.1001.2008.02483

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

描述逻辑$\mathcal{EL}$混合循环术语集的 LCS 和 MSC 推理*

蒋运承[1,2], 王 驹[1], 周生明[1], 汤 庸[2+]

[1](广西师范大学 计算机科学与信息工程学院,广西 桂林 541004)
[2](中山大学 计算机科学系,广东 广州 510275)

LCS and MSC Reasoning of Hybrid Terminological Cycles in Description Logic $\mathcal{EL}$

JIANG Yun-Cheng[1,2], WANG Ju[1], ZHOU Sheng-Ming[1], TANG Yong[2+]

[1](College of Computer Science and Information Engineering, Guangxi Normal University, Guilin 541004, China)
[2](Department of Computer Science, Sun Yat-Sen University, Guangzhou 510275, China)
+ Corresponding author: E-mail: issty@mail.sysu.edu.cn

Jiang YC, Wang J, Zhou SM, Tang Y. LCS and MSC reasoning of hybrid terminological cycles in description logic $\mathcal{EL}$. *Journal of Software*, 2008,19(10):2483−2497. http://www.jos.org.cn/1000-9825/19/2483.htm

Abstract: Current research progresses and the existing problems of terminological cycles in description logics are analyzed in this paper. Based on the works of F. Baader, the LCS (least common subsumer) and MSC (most specific concept) reasoning of hybrid terminological cycles in description logic $\mathcal{EL}$ is further studied. The syntax and semantics of hybrid terminological cycles in description logic $\mathcal{EL}$ are given. Under the requirement of the LCS and MSC reasoning of hybrid terminological cycles in description logic $\mathcal{EL}$, TBox-completion is presented, and the description graph is redefined. The LCS and MSC reasoning algorithms of hybrid terminological cycles in description logic $\mathcal{EL}$ w.r.t. greatest fixpoint semantics are presented by using TBox-completion and description graph. The correctness of reasoning algorithms is proved, and it is also proved that the LCS and MSC reasoning w.r.t. greatest fixpoint semantics can be computed in the polynomial time. Theoretical foundation for the LCS and MSC reasoning for hybrid terminological cycles in description logic $\mathcal{EL}$ is provided through the reasoning algorithms.
Key words: description logic; hybrid terminological cycles; fixpoint semantics; descriptive semantics; LCS (least common subsumer); MSC (most specific concept)

摘 要: 分析了描述逻辑循环术语集的研究现状和存在的问题,在 F.Baader 工作的基础上进一步研究了描述逻辑$\mathcal{EL}$混合循环术语集的 LCS(least common subsumer)和 MSC(most specific concept)推理问题.给出了$\mathcal{EL}$混合循环术语集的语法和语义.针对$\mathcal{EL}$混合循环术语集 LCS 和 MSC 推理的需要,提出了 TBox-完全的概念,并重新定义了描述图.使用描述图和 TBox-完全给出了最大不动点语义下$\mathcal{EL}$混合循环术语集 LCS 和 MSC 的推理算法,证明了推理算法的

---

* Supported by the National Natural Science Foundation of China under Grant Nos.60663001, 60673135, 60373081, 60573010 (国家自然科学基金); the Postdoctoral Science Foundation of China under Grant No.20060400226 (中国博士后科学基金); the Program for New Century Excellent Talents University (新世纪优秀人才支持计划); the Natural Science Key Foundation of Guangdong Province of China under Grant No.04105503 (广东省自然科学重点基金); the Natural Science Foundation of Guangxi Province of China under Grant Nos.0640030, 0832103 (广西自然科学基金)

Received 2007-01-22; Accepted 2007-04-26

正确性,并证明了推理算法是多项式时间复杂的.该推理算法为εL混合循环术语集的LCS和MSC推理提供了理论基础.

关键词: 描述逻辑;混合循环术语集;不动点语义;描述语义;LCS(least common subsumer);MSC(most specific concept)

中图法分类号: TP301 **文献标识码**: A

循环术语集(也称循环定义)是描述逻辑长期以来的研究难点,它的最基本的问题即语义及推理问题没有得到合理的解决[1,2].但循环定义可以极大地扩充描述逻辑的表达能力,而且在许多实际应用中(如医学本体、语义数据模型),循环定义是不可避免的[3−5].另外,循环定义还能够方便用户建立描述逻辑知识库,并使所表示的知识或公理符合人们的直觉.如果没有循环定义,则只能用非循环定义来描述相应的循环定义,这样会使知识库变得非常复杂,用户也很难理解[5,6].因此,研究描述逻辑的循环定义相当有意义.

在循环定义的语义研究方面,Nebel最早深入研究了描述逻辑循环定义,提出了循环定义的3种语义:最大不动点语义、最小不动点语义和描述语义[5,6].大部分学者认为,这3种语义的选择取决于所定义的概念.Giacomo认为,描述语义不适合解释递归概念,最小不动点语义适合解释归纳定义的概念,最大不动点语义适合解释non-well-founded或co-inductive结构的概念[1].Buchheit将TBox分成两部分:模式和视图,对于模式中的循环定义用描述语义来解释,而对于视图中的循环定义用最小不动点语义或最大不动点语义来解释[2].

循环定义的推理包括标准推理和非标准推理两个方面[7],并且目前都针对一些很小的、不带否定构造算子的描述逻辑.在标准推理方面,Nebel[5]和Baader[3]利用自动机理论给出了描述逻辑TL和FL_0循环术语集的包含推理算法,其中,TL和FL_0只包含两个构造算子:交和全称量词.Baader[8]使用描述图及模拟关系给出了描述逻辑εL循环术语集的包含推理算法,其中εL只包含两个构造算子:交和存在量词.Nebel[5]和Baader[3,8]只允许TBox出现概念等价公理,不允许出现概念包含公理(GCI公理),而GCI公理是描述逻辑的主要功能之一[9].为了循环定义能够处理GCI公理,Brandt[9]提出了εL循环术语集混合TBox概念,将TBox分成两部分:基础TBox和术语TBox,其中,基础TBox允许出现GCI公理,用描述语义进行解释;术语TBox允许循环定义,用不动点语义进行解释,并给出了最大不动点语义下TBox约束的包含推理算法.在非标准推理方面,Brandt[7]给出了最大不动点语义下εL混合循环术语集匹配和近似推理的多项式时间算法.Baader[10]给出了最大不动点语义下εL循环术语集的LCS(least common subsumer)和MSC(most specific concept)推理的多项式时间算法,其中,LCS和MSC推理是描述逻辑重要的非标准推理[11],但Baader[10]只允许TBox出现等价公理,不允许TBox出现GCI公理,因此,Baader[10]的工作在实际应用中受到了很大的制约.

1 εL循环术语集

描述逻辑εL[8,10]的概念如下定义:

$$C,D\rightarrow\top|A|C\sqcap D|\exists R.C|,$$

其中,A表示原子概念,C,D表示概念,R表示关系名.

下面用N_C表示εL的所有概念名的集合,N_R表示εL的所有关系名的集合.

εL的语义将概念解释为解释论域的子集,关系是该论域上的二元关系.形式上,一个解释$I=(\Delta^I,\cdot^I)$由解释论域Δ^I和解释函数$\cdot^I$所构成,其中解释函数把每个原子概念$A\in N_C$映射到Δ^I的子集,而把每个关系$R\in N_R$映射到$\Delta^I\times\Delta^I$的子集,即εL的语义解释如下:

- $A^I\subseteq\Delta^I$;
- $R^I\subseteq\Delta^I\times\Delta^I$;
- $\top^I=\Delta^I$;
- $(C\sqcap D)^I=C^I\cap D^I$;
- $(\exists R.C)^I=\{x\in\Delta^I|\exists y\in\Delta^I,(x,y)\in R^I\wedge y\in C^I\}$.

$\mathcal{EL}$ 的 TBox TB 是有限个形如 $A\equiv D$ 的概念定义的集合,其中,A 是一个概念名,D 是一个概念描述.在 TBox 中不允许概念的重复定义.如果概念出现在概念定义的左边,则称为被定义概念,否则称为原始概念.

解释 $I=(\Delta^I,\cdot^I)$满足概念定义 $A\equiv D$,当且仅当 $A^I=D^I$.如果解释 I 满足 TBox TB 中的所有概念定义,即对任意概念定义 $A\equiv D\in TB,A^I=D^I$,则称 I 是 TB 的一个模型.如果 TBox TB 存在一个模型,则称 TB 是可满足的.

上述定义的语义就是 Nebel 提出的描述逻辑的描述语义[5],它接受所有的模型.在描述逻辑循环定义中,Nebel 已证明描述语义不能客观地刻画循环定义的语义,为此引入了循环定义的不动点语义,包括最大不动点语义和最小不动点语义[5].

给定关系 $R\in N_R,\cup_{n\geq 1}R^n$ 称为 R 的传递闭包,其中 $R^{n+1}=R\circ R^n$,$\circ$表示二元关系的合成运算.

如果 TBox 中存在一个被定义概念,该概念直接或间接地出现在它的概念定义中,则称 TBox 是循环的.形式上说,假设 A 是被定义概念,B 是概念(被定义概念或原子概念),A 的概念定义是 $A\equiv C$,其中,C 是一个概念描述,如果 B 在 C 中出现,则称 A 直接使用 B.将关系直接使用的传递闭包称为使用.如果 TBox 中存在一个被定义概念 A,A 使用 A,则称 TBox 是循环的.

例 1:TBox TB: $Human\equiv Mammal\sqcap\exists parents.Human$.该 TB 中,$Human$ 是被定义概念,并且 $Human$ 使用 $Human$,从而该 TB 是循环的.

给定描述逻辑$\mathcal{EL}$ 的 TBox TB,TB 中所有出现的关系名记为 N_{role},所有出现的原始概念记为 N_{prim},所有出现的被定义概念记为 N_{def},并且令 $N_{role}=\{R_1,\ldots,R_m\}$,$N_{prim}=\{P_1,\ldots,P_n\}$,$N_{def}=\{A_1,\ldots,A_k\}$,则 TB 的一个原始解释 $J=(\Delta^J,\cdot^J)$,其中,对任意 $R_i\in N_{role},P_i\in N_{prim},R_i^J\subseteq\Delta^J\times\Delta^J,P_i^J\subseteq\Delta^J$.也就是说,原始解释不对 N_{def} 中的被定义概念 A_i 进行语义解释.TBox TB 的一个解释 $I=(\Delta^I,\cdot^I)$是原始解释 $J=(\Delta^J,\cdot^J)$的扩充,当且仅当$\Delta^I=\Delta^J$, $P_1^I=P_1^J,\ldots,P_n^I=P_n^J$ 以及 $R_1^I=R_1^J,\ldots,R_m^I=R_m^J$.此时,也称 I 是基于 J 的解释.给定原始解释 J,基于 J 的解释 I 由元组($A_1^I,\ldots,A_k^I$)唯一确定,其中 $A_i\in N_{role}$.给定原始解释 J,定义集合 $Int(J)=\{I|I$ 是基于 J 的解释$\}$.

为 $Int(J)$上的解释定义偏序关系$\preceq_J$:如果 $I_1,I_2\in Int(J)$,则 $I_1\preceq_J I_2$,当且仅当 $A_i^{I_1}\subseteq A_i^{I_2}$,$1\leq i\leq k$.并且$\langle Int(J),\preceq_J\rangle$是一个完备格,即 $Int(J)$的任何子集都有一个最小上界和最大下界[8,10].由 Tarski 不动点定理[12]可知:如果存在单调函数 $O:Int(J)\rightarrow Int(J)$,使得当 $I_1\preceq_J I_2$ 时有 $O(I_1)\preceq_J O(I_2)$,则函数 O 存在不动点,即存在解释 $I\in Int(J)$,使得 $O(I)=I$,并且函数 O 存在最小不动点和最大不动点.

给定$\mathcal{EL}$ 的 TBox $TB=\{A_1\equiv D_1,\ldots,A_k\equiv D_k\}$、原始解释 J,可以为 $Int(J)$定义单调函数 $O_{T,J}:Int(J)\rightarrow Int(J),O_{T,J}(I_1)=I_2$,当且仅当 $A_i^{I_2}=D_i^{I_1}$,$1\leq i\leq k$.从而函数 $O_{T,J}$ 存在最小不动点和最大不动点.并且如果 I 是基于原始解释 J 的解释,则 I 是函数 $O_{T,J}$ 的不动点,当且仅当 I 是 TBox TB 的模型[13].因此,可以如下定义$\mathcal{EL}$ 的不动点语义:

定义 1. 给定$\mathcal{EL}$ 的 TBox TB,TB 的模型 I 称为最大不动点模型(最小不动点模型),当且仅当存在一个原始解释 $J,I\in Int(J)$,并且 I 是函数 $O_{T,J}$ 的最大不动点(最小不动点).由于 TBox TB 的模型可能有多个,最大不动点语义(最小不动点语义)仅接受最大不动点模型(最小不动点模型)作为 TB 的模型.下面将最大不动点模型(最小不动点模型)分别称为 gfp-模型(lfp-模型).

定义 2. 给定$\mathcal{EL}$ 的 TBox TB,A,B 是 TB 中的被定义概念,则:

- 在最大不动点语义下,B 包含 A(记为 $A\sqsubseteq_{gfp,TB}B$),当且仅当对 TB 的任意 gfp-模型 I,有 $A^I\subseteq B^I$;
- 在最小不动点语义下,B 包含 A(记为 $A\sqsubseteq_{lfp,TB}B$),当且仅当对 TB 的任意 lfp-模型 I,有 $A^I\subseteq B^I$;
- 在描述语义下,B 包含 A(记为 $A\sqsubseteq_{TB}B$),当且仅当对 TB 的任意模型 I,有 $A^I\subseteq B^I$.

由不动点理论[12]可知,如果函数不仅单调,而且向下连续或向上连续,则最大或最小不动点可以通过ω-迭代来获得;否则,仍然可以通过迭代来获得最大或最小不动点,只是迭代过程需要比ω更大的序数.

给定一个基于原始解释 J 的递增解释链 $chain$:$I_0\preceq_J I_1\preceq_J I_2\preceq_J\ldots$,由$\preceq_J$ 的定义可知,解释链 $chain$ 的上确界是一个基于 J 的解释 I,满足 $A_i^I=\cup_{j\geq 0}A_i^{I_j}$,其中,$1\leq i\leq k$.并且,函数 $O:Int(J)\rightarrow Int(J)$是向上连续的,当且仅当

$$O(lub(\{I_j|j\geq 0\}))=lub(\{O(I_j)|j\geq 0\}).$$

给定一个基于原始解释 J 的递减解释链 $chain$:$I_0\succeq_J I_1\succeq_J I_2\succeq_J\ldots$,同理可得解释链 $chain$ 的下确界是一个基

于 J 的解释 I,满足 $A_i^I=\cap_{j\geq 0}A_i^{I_j}$,其中,$1\leq i\leq k$.并且函数 $O:Int(J)\rightarrow Int(J)$是向下连续的,当且仅当

$$O(glb(\{I_j|j\geq 0\}))=glb(\{O(I_j)|j\geq 0\}).$$

定理 1[14]. 给定εL 的 TBox TB,J 是一个原始解释,则函数 $O_{T,J}$ 是向上连续的,但不一定是向下连续的.

下面给出函数 $O_{T,J}$ 的最大不动点(最小不动点)的构造方法.

定义 3. 给定εL 的 TBox TB,J 是一个原始解释,I_{top} 和 I_{bot} 分别是基于 J 的最大和最小解释,即 $A_i^{I_{top}}=\Delta^J$, $A_i^{I_{bot}}=\phi$,$1\leq i\leq k$,则对任意序数α:

- 如果$\alpha=0$,则 $I^{\uparrow\alpha}=I_{bot}$,$I^{\downarrow\alpha}=I_{top}$;
- $I^{\uparrow\alpha+1}=O_{T,J}(I^{\uparrow\alpha})$,$I^{\downarrow\alpha+1}=O_{T,J}(I^{\downarrow\alpha})$;
- 如果α是极限序数,则 $I^{\uparrow\alpha}=lub(\{I^{\uparrow\beta}|\beta<\alpha\})$,$I^{\downarrow\alpha}=glb(\{I^{\downarrow\beta}|\beta<\alpha\})$.

由定理 1 可知,函数 $O_{T,J}$ 是向上连续的,因而由不动点定理[12]可知:$I^{\uparrow\omega}$是函数 $O_{T,J}$ 的最小不动点;又因为函数 $O_{T,J}$ 不是向下连续的,因而 $I^{\downarrow\omega}$不是函数 $O_{T,J}$ 的最大不动点,但一定存在一个序数α,有 $I^{\downarrow\alpha}$是函数 $O_{T,J}$ 的最大不动点.

2 εL 混合循环知识库

2.1 语法和语义

εL 混合循环知识库包括 TBox TB 和 ABox AB 两部分,其中,TBox TB 又包括基础 TBox FT 和术语 TBox TT 两部分.与εL 的混合 TBox[9]相同,基础 TBox FT 是由定义在 N_{role} 和 N_{prim} 上的形如 $C\sqsubseteq D$ 的 GCI 公理所组成的有限集合,其中,C,D 是概念描述.术语 TBox TT 是由定义在 N_{role},N_{prim} 和 N_{def} 上εL 的循环术语集[8,10],它是有限个形如 $A\equiv D$ 的概念定义的集合,其中,A 是一个概念名,D 是一个概念描述,并且对 TT 中的任意原始概念 P(或存在约束概念$\exists R.P$),存在 GCI 公理 $C\sqsubseteq D\in FT$,使得 P(或$\exists R.P$)是 C 或 D 的子概念.ABox AB 是εL 的实例断言公理的有限集合,其中,实例断言公理有两种形式:$A(a)$和 $R(a,b)$,A 是概念,R 是关系,a,b 是个体.

定义 4. εL 的混合循环知识库 $HKB=(FT,TT,AB)$,其中,FT 是定义在 N_{role} 和 N_{prim} 上的基础 TBox,TT 是定义在 N_{role},N_{prim} 和 N_{def} 上的术语 TBox,AB 是 ABox.

例 2:图 1 是一个有关医疗术语的混合循环知识库(参考文献[9]):

FT: $Disease\sqcap\exists has\_loc.ConnTissue\sqsubseteq\exists acts\_on.ConnTissue$;
$Inflammation\sqsubseteq Disease$; $Pericardium\sqsubseteq ConnTissue$.
TT: $ConnTissDisease\equiv Disease\sqcap\exists acts\_on.ConnTissue$;
$BactInfection\equiv Infection\sqcap\exists causes.BactPericarditis$;
$BactPericarditis\equiv Inflammation\sqcap\exists has\_loc.Pericardium\sqcap\exists caused\_by.BactInfection$.
AB: $ConnTissDisease(a)$; $acts\_on(a,b)$; $ConnTissue(b)$; $BactInfection(c)$; $BactPericarditis(d)$.
Where *ConnTissDisease* means "connective tissue disease"; *ConnTissue* means "connective tissue";
BactInfection means "Bacterial Infection"; *BactPericarditis* means "Bacterial Pericarditis".

Fig.1 Hybrid cyclic knowledge base

图 1 混合循环知识库

由于基础 TBox FT 只包含原始概念和关系,没有被定义概念,因而 FT 利用描述语义来解释:

- 给定原始解释 $J=(\Delta^J,\cdot^J)$,J 是 FT 的模型(记为 $J\models FT$),当且仅当对任意 GCI 公理 $C\sqsubseteq D\in FT$,有 $C^J\subseteq D^J$.

由于术语 TBox TT 允许循环定义,因而需要利用不动点语义和描述语义来解释:

- 给定原始解释 $J=(\Delta^J,\cdot^J)$,I 是基于 J 的最大不动点解释,I 是 TT 的最大不动点模型(记为 $I\models_{gfp,TT}TT$),当且仅当对任意概念定义 $A\equiv D\in TT$,有 $A^I=D^I$;
- 给定原始解释 $J=(\Delta^J,\cdot^J)$,I 是基于 J 的最小不动点解释,I 是 TT 的最小不动点模型(记为 $I\models_{lfp,TT}TT$),当且仅当对任意概念定义 $A\equiv D\in TT$,有 $A^I=D^I$;
- 给定原始解释 $J=(\Delta^J,\cdot^J)$,I 是基于 J 的解释,I 是 TT 的模型(记为 $I\models_{TT}TT$),当且仅当对任意概念定义

$A\equiv D\in TT$,有$A^I=D^I$.

由于ABox AB是实例断言的集合,因而ABox AB的语义解释如下:

- 个体解释为论域的一个元素,即对任意个体a,解释$I=(\Delta^I,\cdot^I)$,则$a^I\in\Delta^I$.I满足$A(a)$,当且仅当$a^I\in A^I$.I满足$R(a,b)$,当且仅当$(a^I,b^I)\in R^I$;
- $I\models_{gfp,TT}TT$,I是AB的模型(记为$I\models_{gfp,TT}AB$),当且仅当I满足ABox AB所有实例断言公理,即对任意$A(a)\in AB$,有$a^I\in A^I$;对任意$R(a,b)\in AB$,有$(a^I,b^I)\in R^I$;
- $I\models_{lfp,TT}TT$,I是AB的模型(记为$I\models_{lfp,TT}AB$),当且仅当I满足ABox AB所有实例断言公理,即对任意$A(a)\in AB$,有$a^I\in A^I$;对任意$R(a,b)\in AB$,有$(a^I,b^I)\in R^I$;
- $I\models_{TT}TT$,I是AB的模型(记为$I\models_{TT}AB$),当且仅当I满足ABox AB所有实例断言公理,即对任意$A(a)\in AB$,有$a^I\in A^I$;对任意$R(a,b)\in AB$,有$(a^I,b^I)\in R^I$.

基于上面的分析,可以如下定义混合循环知识库的语义:

定义5. 给定εL的混合循环知识库$HKB=(FT,TT,AB)$、原始解释$J=(\Delta^I,\cdot^J)$,I是基于J的解释:

- 如果$J\models FT$,$I\models_{gfp,TT}TT$,$I\models_{gfp,TT}AB$,则I是HKB的最大不动点模型.如果HKB存在一个最大不动点模型,则称在最大不动点语义下HKB是可满足的;
- 如果$J\models FT$,$I\models_{lfp,TT}TT$,$I\models_{lfp,TT}AB$,则I是HKB的最小不动点模型.如果HKB存在一个最小不动点模型,则称在最小不动点语义下HKB是可满足的;
- 如果$J\models FT$,$I\models_{TT}TT$,$I\models_{TT}AB$,则I是HKB的模型.如果HKB存在一个模型,则称HKB是可满足的.

由于混合循环知识库HKB包含ABox,当ABox为空集时,混合循环知识库HKB就等价于文献[9]的混合TBox,这时,$HKB=(FT,TT)$.

定义6. 给定εL的混合循环知识库$HKB=(FT,TT)$、原始解释$J=(\Delta^I,\cdot^J)$,I是基于J的解释:

- 如果$J\models FT$,$I\models_{gfp,TT}TT$,则I是$HKB=(FT,TT)$的最大不动点模型;
- 如果$J\models FT$,$I\models_{lfp,TT}TT$,则I是$HKB=(FT,TT)$的最小不动点模型;
- 如果$J\models FT$,$I\models_{TT}TT$,则I是$HKB=(FT,TT)$的模型.

由于混合循环知识库有3种语义,因此对应有3种概念包含推理和实例检测推理,定义如下:

定义7. 给定εL的混合循环知识库$HKB=(FT,TT,AB)$、原始解释$J=(\Delta^I,\cdot^J)$、被定义概念$A\in N_{def}$以及实例个体a,则有:

- 在最大不动点语义下,B包含A(记为$A\sqsubseteq_{gfp,FT,TT}B$),当且仅当对$(FT,TT)$的任意最大不动点模型$I$有$A^I\subseteq B^I$;
- 在最小不动点语义下,B包含A(记为$A\sqsubseteq_{lfp,FT,TT}B$),当且仅当对$(FT,TT)$的任意最小不动点模型$I$有$A^I\subseteq B^I$;
- 在描述语义下,B包含A(记为$A\sqsubseteq_{FT,TT}B$),当且仅当对(FT,TT)的任意模型I有$A^I\subseteq B^I$;
- 在最大不动点语义下,a是A的实例个体(记为$AB\models_{gfp,FT,TT}A(a)$),当且仅当对$HKB$的任意最大不动点模型$I$有$a^I\in A^I$;
- 在最小不动点语义下,a是A的实例个体(记为$AB\models_{lfp,FT,TT}A(a)$),当且仅当对$HKB$的任意最小不动点模型$I$有$a^I\in A^I$;
- 在描述语义下,a是A的实例个体(记为$AB\models_{FT,TT}A(a)$),当且仅当对HKB的任意模型I有$a^I\in A^I$.

因为本文主要研究εL混合循环术语集的LCS和MSC推理,下面给出相关定义及性质.

定义8. 给定εL的混合循环知识库$HKB_1=(FT_1,TT_1)$,$HKB_2=(FT_2,TT_2)$,如果$HKB_1\subseteq HKB_2$,即$FT_1\subseteq FT_2$,$TT_1\subseteq TT_2$并且HKB_1和HKB_2有相同的原始概念和关系,则称HKB_2是HKB_1的保守扩充.

定理2. 给定εL的混合循环知识库$HKB_1=(FT_1,TT_1)$,$HKB_2=(FT_2,TT_2)$,HKB_2是HKB_1的保守扩充,A,B是TT_1中的被定义概念,即$A,B\in N_{def}^{TT_1}$,则$A\sqsubseteq_{gfp,FT_1,TT_1}B$,当且仅当$A\sqsubseteq_{gfp,FT_2,TT_2}B$.

证明:先证明$\Leftarrow$.因为HKB_2是HKB_1的保守扩充,从而$FT_1\subseteq FT_2$,$TT_1\subseteq TT_2$.令$FT=FT_2\backslash FT_1$,$TT=TT_2\backslash TT_1$,$HKB=(FT,TT)$,其中,TT的原始概念是TT_1的原始概念和被定义概念,TT的关系是TT_1(或TT_2)的关系.

用反证法,假设$A\not\sqsubseteq_{gfp,FT_1,TT_1}B$,从而存在$HKB_1=(FT_1,TT_1)$的最大不动点模型$I$,使得$A^I\not\subseteq B^I$,即存在原始解释$J$,$I$

是基于 J 的解释,并且 $J\models FT_1,I\models_{gfp,TT_1} TT_1$,使得 $A^I\not\subseteq B^I$.因为 TT 的原始概念是 TT_1 的原始概念和被定义概念,TT 的关系是 TT_1 的关系,从而 I 是 TT 的一个原始解释,因此,对 I 进行扩充可以得到 TT 的一个最大不动点模型 $\hat{I}$.又因为 I 和 $\hat{I}$ 对 TT_1 的原始概念和被定义概念的解释相同,I 是 TT_1 的模型,从而 $\hat{I}$ 是 TT_2 的模型.又因为 I 是 TT_1 的模型,$\hat{I}$ 是基于 I 的模型,HKB_1 和 HKB_2 有相同的原始概念和关系,从而 $\hat{I}\models FT_2$.因此,$\hat{I}$ 是 $HKB_2=(FT_2,TT_2)$的模型.下面证明 $\hat{I}$ 是 $HKB_2=(FT_2,TT_2)$的最大不动点模型.

假设 I' 也是基于 J 的 $HKB_2=(FT_2,TT_2)$的模型,并且 $\hat{I}\preceq_J I'$.因为 I 是 $HKB_1=(FT_1,TT_1)$的最大不动点模型,因此,I 和 I' 对 TT_1 的被定义概念解释相同.由 $\hat{I}$ 是基于 I 的模型可知,$\hat{I}$ 和 I' 对 TT_1 的被定义概念解释相同,从而 TT 中存在一个被定义概念 C 有 $C^{\hat{I}}\subseteq C^{I'}$,又因为 $\hat{I}$ 是 TT 的最大不动点模型,因此 $C^{\hat{I}}\supseteq C^{I'}$,从而 $C^{\hat{I}}=C^{I'}$,即 $\hat{I}$ 和 I' 对 TT 的被定义概念解释相同.由 $TT=TT_2\backslash TT_1$ 可知,$TT_2=TT+TT_1$,因此 $\hat{I}$ 和 I' 对 TT_2 的被定义概念解释相同,即 $\hat{I}=I'$,即 $\hat{I}$ 是 $HKB_2=(FT_2,TT_2)$的最大不动点模型.

又因为 $A,B\in N_{def}^{TT_1}$,$TT_1\subseteq TT_2$,从而 $A,B\in N_{def}^{TT_2}$.所以有 $A^{\hat{I}}=A^I$, $B^{\hat{I}}=B^I$.又因为 $A^I\not\subseteq B^I$,从而 $A^{\hat{I}}\not\subseteq B^{\hat{I}}$,即有 $A\not\sqsubseteq_{gfp,FT2,TT2}B$,与 $A\sqsubseteq_{gfp,FT2,TT2}B$ 矛盾.

再证明⇒.由于 $FT_1\subseteq FT_2,TT_1\subseteq TT_2$,因而只需证明对 $HKB_2=(FT_2,TT_2)$的任意最大不动点模型 I,I 在 $HKB_1=(FT_1,TT_1)$上的限制 I' 是 $HKB_1=(FT_1,TT_1)$的最大不动点模型即可.

首先证明 I' 是 $HKB_1=(FT_1,TT_1)$的模型.因为 I 是 $HKB_2=(FT_2,TT_2)$的模型,从而存在一个原始解释 J,I 是基于 J 的解释,使得 $J\models FT_2,I\models TT_2$.又因为 I' 是 I 在 $HKB_1=(FT_1,TT_1)$上的限制,从而存在 J 在 FT_1 上的限制 J'(由于 HKB_1 和 HKB_2 具有相同的原始概念和关系,因此可知 $J'=J$),I' 是基于 J' 的解释.由 $J\models FT_2,I\models TT_2,FT_1\subseteq FT_2,TT_1\subseteq TT_2$ 可知,$J'\models FT_1,I'\models TT_1$.因此,$I'$ 是 $HKB_1=(FT_1,TT_1)$的模型.下面再证明 I' 是 $HKB_1=(FT_1,TT_1)$的最大不动点模型.

假设 I' 不是 $HKB_1=(FT_1,TT_1)$的最大不动点模型,从而存在基于 J' 的 $HKB_1=(FT_1,TT_1)$的模型 I'',并且 $I'\preceq_{J'}I''$.由⇐的证明可知,对 I'' 进行扩充可以得到 $HKB_2=(FT_2,TT_2)$的一个最大不动点模型 I''',并且 I''' 和 I 具有相同的原始解释,因而 $I'''=I$.又因为 I' 是 I 在 $HKB_1=(FT_1,TT_1)$上的限制,I'' 是 I''' 在 $HKB_1=(FT_1,TT_1)$上的限制,由于 $I'''=I$,因此 $I'=I''$,所以 I' 是 $HKB_1=(FT_1,TT_1)$的最大不动点模型. □

定义 9. 给定 εL 的混合循环知识库 $HKB_1=(FT_1,TT_1),HKB_2=(FT_2,TT_2)$,$HKB_2$ 是 HKB_1 的保守扩充,A,B 是 TT_1 中的被定义概念,即 $A,B\in N_{def}^{TT_1}$,E 是 TT_2 中的被定义概念并且不是 TT_1 中的被定义概念,即 $E\in N_{def}^{TT_2}$,$E\notin N_{def}^{TT_1}$,则在最大不动点语义下,TT_2 中的 E 是 TT_1 中 A,B 的 LCS(记为 *gfp-LCS*),当且仅当满足下列两个条件:

(1) $A\sqsubseteq_{gfp,FT_2,TT_2}E,B\sqsubseteq_{gfp,FT_2,TT_2}E$;

(2) 如果 $HKB_3=(FT_3,TT_3)$是 $HKB_2=(FT_2,TT_2)$的保守扩充,F 是 TT_3 中的被定义概念,即 $F\in N_{def}^{TT_3}$,并且 $A\sqsubseteq_{gfp,FT_3,TT_3}F,B\sqsubseteq_{gfp,FT_3,TT_3}F$,则 $E\sqsubseteq_{gfp,FT_3,TT_3}F$.

对于 εL 混合循环术语集的 LCS,有下列性质:

定理 3. 给定 εL 的混合循环知识库 $HKB_1=(FT_1,TT_1),HKB_2=(FT_2,TT_2),HKB_2'=(FT_2',TT_2')$,$HKB_2$ 和 HKB_2' 是 HKB_1 的保守扩充,A,B 是 TT_1 中的被定义概念,即 $A,B\in N_{def}^{TT_1}$,并且:

- TT_2 中的 E 是 TT_1 中 A,B 的 *gfp-LCS*;
- TT_2' 中的 E' 是 TT_1 中 A,B 的 *gfp-LCS*;
- $NC=\{C|C\in N_{def}^{TT_2},C\notin N_{def}^{TT_1}\},NC'=\{C|C\in N_{def}^{TT_2'},C\notin N_{def}^{TT_1}\},NC\cap NC'=\varnothing$.

令 $FT_3=FT_2\cup FT_2',TT_3=TT_2\cup TT_2',HKB_3=(FT_3,TT_3)$,则 $E\equiv_{gfp,FT_3,TT_3}E'$,即 $E\sqsubseteq_{gfp,FT_3,TT_3}E'$ 和 $E'\sqsubseteq_{gfp,FT_3,TT_3}E$.

证明:因为 $NC\cap NC'=\varnothing,HKB_3,HKB_2$ 和 HKB_2' 的原始概念、关系相同,从而 $HKB_3=(FT_3,TT_3)$是 HKB_2 和 HKB_2' 的保守扩充.又因为 TT_2 中的 E 是 TT_1 中的 A,B 的 *gfp-LCS*,从而 $A\sqsubseteq_{gfp,FT_2,TT_2}E,B\sqsubseteq_{gfp,FT_2,TT_2}E$.由于 $HKB_3=(FT_3,TT_3)$是 HKB_2 的保守扩充,由定理 2 可知,$A\sqsubseteq_{gfp,FT_3,TT_3}E,B\sqsubseteq_{gfp,FT_3,TT_3}E$.又因为 TT_2' 中的 E' 是 TT_1 中 A,B 的 *gfp-LCS*,从而有 $A\sqsubseteq_{gfp,FT_2',TT_2'}E',B\sqsubseteq_{gfp,FT_2',TT_2'}E'$.由于 $HKB_3=(FT_3,TT_3)$是 HKB_2' 的保守扩充,由定理 2 可知,$A\sqsubseteq_{gfp,FT_3,TT_3}E'$, $B\sqsubseteq_{gfp,FT_3,TT_3}E'$.因为 TT_2 中的 E 是 A,B 的 *gfp-LCS*,所以有 $E\sqsubseteq_{gfp,FT_3,TT_3}E'$.又因为 TT_2' 中的 E' 是 A,B 的 *gfp-LCS*,所

以有 $E'\sqsubseteq_{gfp,FT_3,TT_3}E$.因此,$E\equiv_{gfp,FT_3,TT_3}E'$. □

由定理 3 可知,$\mathcal{EL}$混合循环术语集中概念的 *gfp-LCS* 是唯一确定的,即如果 E_1 和 E_2 是概念 C 和 D 的 *gfp-LCS*,则 $E_1\equiv E_2$.

定义 10. 给定$\mathcal{EL}$ 的混合循环知识库 $HKB_1=(FT_1,TT_1,AB_1)$,a 是 HKB_1 的实例个体,$HKB_2=(FT_2,TT_2)$是(FT_1,TT_1)的保守扩充,E 是 TT_2 中的被定义概念,即 $E\in N_{def}^{TT_2}$,则在最大不动点语义下,TT_2 中的 E 是 HKB_1 中 a 的 MSC(记为 *gfp-MSC*),当且仅当满足下列两个条件:

(1) $AB_1\models_{gfp,FT_2,TT_2}E(a)$;

(2) 如果 $HKB_3=(FT_3,TT_3)$是 $HKB_2=(FT_2,TT_2)$的保守扩充,F 是 TT_3 中的被定义概念,即

$$F\in N_{def}^{TT_3},AB_1\models_{gfp,FT_3,TT_3}F(a),$$

则 $E\sqsubseteq_{gfp,FT_3,TT_3}F$.

对于$\mathcal{EL}$ 混合循环术语集的 MSC,有下列性质:

定理 4. 给定$\mathcal{EL}$ 的混合循环知识库 $HKB_1=(FT_1,TT_1,AB_1)$,a 是 HKB_1 的实例个体,$HKB_2=(FT_2,TT_2)$和 $HKB_2'=(FT_2',TT_2')$是(FT_1,TT_1)的保守扩充,并且:

- TT_2 中的 E 是 HKB_1 中 a 的 *gfp-MSC*;
- TT_2' 中的 E'是 HKB_1 中 a 的 *gfp-MSC*;
- $NC=\{C|C\in N_{def}^{TT_2},C\notin N_{def}^{TT_1}\}$,$NC'=\{C|C\in N_{def}^{TT_2'},C\notin N_{def}^{TT_1}\}$,$NC\cap NC'=\varnothing$.

令 $FT_3=FT_2\cup FT_2'$,$TT_3=TT_2\cup TT_2'$,$HKB_3=(FT_3,TT_3)$,则 $E\equiv_{gfp,FT_3,TT_3}E'$,即 $E\sqsubseteq_{gfp,FT_3,TT_3}E'$和 $E'\sqsubseteq_{gfp,FT_3,TT_3}E$.

证明:因为 $NC\cap NC'=\varnothing$,HKB_3,HKB_2 和 HKB_2' 的原始概念、关系相同,从而 $HKB_3=(FT_3,TT_3)$是 HKB_2 和 HKB_2' 的保守扩充.又因为 TT_2 中的 E 是 HKB_1 中 a 的 *gfp-MSC*,从而 $AB_1\models_{gfp,FT_2,TT_2}E(a)$.由于 $HKB_3=(FT_3,TT_3)$是 HKB_2 的保守扩充,所以 $AB_1\models_{gfp,FT_3,TT_3}E(a)$.又因为 TT_2' 中的 E'是 HKB_1 中的 a 的 *gfp-MSC*,从而 $AB_1\models_{gfp,FT_2',TT_2'}E'(a)$.由于 $HKB_3=(FT_3,TT_3)$是 HKB_2' 的保守扩充,所以 $AB_1\models_{gfp,FT_3,TT_3}E'(a)$.因为 TT_2 中的 E 是 HKB_1 中的 a 的 *gfp-MSC*,所以有 $E\sqsubseteq_{gfp,FT_3,TT_3}E'$.又因为 TT_2' 中的 E'是 HKB_1 的 a 的 *gfp-MSC*,所以有 $E'\sqsubseteq_{gfp,FT_3,TT_3}E$.因此,$E\equiv_{gfp,FT_3,TT_3}E'$. □

由定理 4 可知,$\mathcal{EL}$ 混合循环术语集中个体的 *gfp-MSC* 是唯一确定的,即如果 E_1 和 E_2 是个体 a 的 *gfp-MSC*,则 $E_1\equiv E_2$.

2.2 知识库的正规化

$\mathcal{EL}$ 混合循环知识库包括 TBox TB 和 ABox AB 两部分.由于 ABox 是实例断言公理的有限集合,其中,实例断言公理有两种形式:$A(a)$和 $R(a,b)$,A 是概念,R 是关系,a,b 是个体.形如 $A(a)$和 $R(a,b)$的断言已是正规形式,不需要再进行正规化,因而下面仅给出 TBox(包括基础 TBox 和术语 TBox)的正规化方法.

定义 11. 给定$\mathcal{EL}$ 的 TBox $TB=(FT,TT)$,TB 是正规化的,当且仅当

(1) 对任意的概念定义 $A\equiv D\in TT,A\in N_{def},D$ 具有下列形式:

$$P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n,$$

其中,$m,n\geq 0,P_1,\ldots,P_m\in N_{prim},R_1,\ldots,R_n\in N_{role},B_1,\ldots,B_n\in N_{def}$.如果 $m=n=0$,则 $D\equiv\top$.

(2) 对任意的 GCI 公理 $C\sqsubseteq D\in FT,C\sqsubseteq D$ 是下列形式之一:

$$A\sqsubseteq B,A_1\sqcap A_2\sqsubseteq B,A\sqsubseteq\exists R.B,\exists R.A\sqsubseteq B,\text{其中},A,B,A_1,A_2\in N_{prim}\cup\{\top\},R\in N_{role}.$$

(3) 对 FT 中的任意原始概念 P 和存在约束概念$\exists R.P$,其中,$P\in N_{prim},R\in N_{role},TT$ 包含如下概念定义式:

$$A_P\equiv P,A_{\exists R.P}\equiv\exists R.A_P.$$

由于 ABox 是正规形式,不需要进行正规化,因而有下列定义:

定义 12. 给定$\mathcal{EL}$ 的混合循环知识库 $HKB=(FT,TT,AB)$,如果 FT 和 TT 是正规形式的,则 HKB 称为正规形式的(或 HKB 是正规化知识库).

例 3:图 2 是图 1 的混合循环知识库所对应的正规化混合循环知识库.

FT: $\exists has\_loc.ConnTissue \sqsubseteq HasLocConnTissue$;
$ActsOnConnTissue \sqsubseteq \exists acts\_on.ConnTissue$;
$Disease \sqcap HasLocConnTissue \sqsubseteq ActsOnConnTissue$;
$Inflammation \sqsubseteq Disease$; $Pericardium \sqsubseteq ConnTissue$.
TT: $ConnTissDisease \equiv Disease \sqcap \exists acts\_on.ConnectiveTissue$;
$ConnectiveTissue \equiv ConnTissue$;
$BactInfection \equiv Infection \sqcap \exists causes.BactPericarditis$;
$BactPericarditis \equiv Inflammation \sqcap \exists has\_loc.Pericardia \sqcap \exists caused\_by.BactInfection$;
$Pericardia \equiv Pericardium$;
$A_{ConnTissue} \equiv ConnTissue$; $A_{HasLocConnTissue} \equiv HasLocConnTissue$;
$A_{ActsOnConnTissue} \equiv ActsOnConnTissue$; $A_{Disease} \equiv Disease$;
$A_{Inflammation} \equiv Inflammation$; $A_{Pericardium} \equiv Pericardium$; $A_{ConnTissue} \equiv ConnTissue$;
$A_{\exists has\_loc.ConnTissue} \equiv \exists has\_loc.ConnTissue$; $A_{\exists acts\_on.ConnTissue} \equiv \exists acts\_on.ConnTissue$.
AB: $ConnTissDisease(a)$; $acts\_on(a,b)$; $ConnTissue(b)$; $BactInfection(c)$; $BactPericarditis(d)$

Fig.2 Normalized hybrid cyclic knowledge base
图 2 正规化混合循环知识库

由定义 11 可知,由于 TBox 包括基础 TBox 和术语 TBox,因而 TBox TB 的正规化对应两部分:基础 TBox 的正规化和术语 TBox 的正规化.由于 TBox TB 与文献[9]的混合 TBox 具有相同的形式,因而基础 TBox 的正规化方法与文献[9]的基础 TBox 的正规化方法相同,术语 TBox 的正规化方法与文献[8]的术语 TBox 的正规化方法相同.又因为 ABox 不需要正规化,从而由文献[8,9]的结论直接可得:

定理 5. 给定$\mathcal{EL}$ 的混合循环知识库 HKB,HKB 能在多项式时间内转化为一个等价的正规形式的混合循环知识库 HKB′.

由于$\mathcal{EL}$ 的混合循环知识库能够等价地转化为一个正规形式的混合循环知识库,下面不妨假设$\mathcal{EL}$ 的混合循环知识库都是正规形式的.

2.3 描述图

定义 13. 描述逻辑$\mathcal{EL}$ 的描述图 $G=(V,E,L)$,其中:

- V 是节点集合;
- $E \subseteq V \times N_{role} \times V$ 是有向边集合,每条有向边用$\mathcal{EL}$ 的关系标注;
- $L:V \rightarrow 2^{N_{prim}}$ 是将节点映射为原始概念集合的函数.

描述图分为语法描述图和语义描述图,其中,语法描述图不考虑实例个体,即只考虑 TBox;而语义描述图还要考虑 ABox 或解释论域.

给定描述逻辑$\mathcal{EL}$ 的术语 TBox TT,TT 能够转化为下列对应的一个语法描述图 G_{TT}.

定义 14. 给定描述逻辑$\mathcal{EL}$ 的术语 TBox TT,TT 对应的语法描述图 $G_{TT}=(N_{def}^{TT}, E_{TT}, L_{TT})$,其中:

- G_{TT}的节点集是 TT 的被定义概念集 N_{def}^{TT};
- 如果 A 是被定义概念,并且 A 在 TT 中的定义式为 $A \equiv P_1 \sqcap \ldots \sqcap P_m \sqcap \exists R_1.B_1 \sqcap \ldots \sqcap \exists R_n.B_n$,则 $L_{TT}(A)=\{P_1,\ldots,P_m\}$;$(A,R_1,B_1),\ldots,(A, R_n,B_n) \in E_{TT}$.

例 4:将图 2 中的 TT 如下部分可以转化为图 3 所示的描述图.

TT: $ConnTissDisease \equiv Disease \sqcap \exists acts\_on.ConnectiveTissue$;
$ConnectiveTissue \equiv ConnTissue$;
$BactInfection \equiv Infection \sqcap \exists causes.BactPericarditis$;
$BactPericarditis \equiv Inflammation \sqcap \exists has\_loc.Pericardia \sqcap \exists caused\_by.BactInfection$;
$Pericardia \equiv Pericardium$.

其中,图 3 中的术语见例 2.从图 3 可以看出,$\mathcal{EL}$ 的术语 TBox TT 与语法描述图 G_{TT}之间可以相互转化,即术语 TBox TT 可以等价地转化为一个语法描述图 G_{TT},语法描述图 G_{TT} 也可以等价地转化为一个术语 TBox TT.也就

是说,语法描述图 G_{TT} 是用图的形式来描述术语 TBox TT.

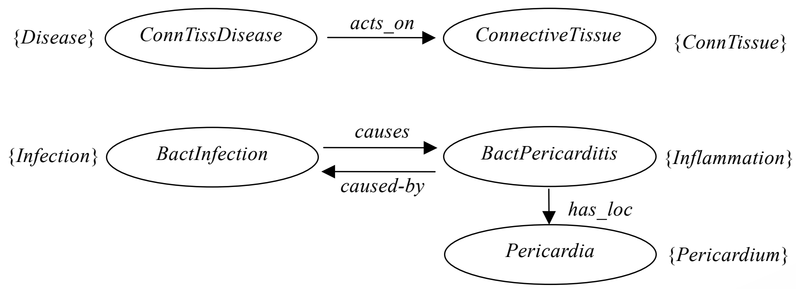

Fig.3 Description graph

图 3 描述图

定义 15. 给定描述逻辑$\mathcal{EL}$的正规化 TBox $TB=(FT,TT)$,对任意被定义概念 $A\in N_{def}^{TB}$,定义函数:

$$f(A)=(\sqcap_{P\in\{P'\in N_{prim}^{FT}|A\sqsubseteq_{FT\cup TT}P'\}}P)\sqcap(\sqcap_{R\in N_{role}^{FT}}\sqcap_{Q\in\{Q'\in N_{prim}^{FT}|A\sqsubseteq_{FT\cup TT}\exists R.Q'\}}\,_{P'}\exists R.A_Q),$$

则 TT 的 f-完全 $f(TT)=\{A\equiv C\sqcap f(A)|A\equiv C\in TT\}$.

例 5:对于例 3 给出的正规化 TBox $TB=(FT,TT)$,可以得到:$f(A_{Inflammation})=Disease$,$f(A_{Pericardium})=ConnTissue$, $f(A_{ActsOnConnTissue})=\exists acts\_on.ConnTissu$,$f(BactPericarditis)=Disease\sqcap HasLocConnTissue$,从而可以得到 TT 的 f-完全 $f(TT)$.以 $BactPericarditis\equiv Inflammation\sqcap\exists has\_loc.Pericardia\sqcap\exists caused\_by.BactInfection$ 为例,该公理修改为

$$BactPericarditis\equiv Inflammation\sqcap\exists has\_loc.Pericardia\sqcap\exists caused\_by.BactInfection\sqcap Disease\sqcap HasLocConnTissue.$$

引入 f-完全的作用是为了将混合循环术语集的推理转化为循环术语集的推理,即将基础 TBox 部分去除,在 TT 的 f-完全 $f(TT)$下研究混合循环术语集的 LCS 和 MSC 推理.由定义 15 可知,$f(TT)$仍然是一个正规化的术语 TBox,并且 $f(TT)$是根据 FT 对 TT 的扩充,从而可以如下给出 TBox $TB=(FT,TT)$对应的语法描述图 $G_{f(TT)}$.

定义 16. 给定描述逻辑$\mathcal{EL}$的 TBox $TB=(FT,TT)$,TB 对应的语法描述图 $G_{f(TT)}=(N_{def}^{f(TT)},E_{f(TT)},L_{f(TT)})$,其中:

- $G_{f(TT)}$的节点集是 $f(TT)$的被定义概念集 $N_{def}^{f(TT)}$;
- 如果 A 是被定义概念,并且 A 在 $f(TT)$中的定义式为 $A\equiv P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n$,则 $L_{f(TT)}(A)=\{P_1,\ldots,P_m\}$;$(A,R_1,B_1),\ldots,(A,R_n,B_n)\in E_{f(TT)}$.

由上可以看出,$f(TT)$与语法描述图 $G_{f(TT)}$之间也可以相互转化.

定义 17. 给定描述逻辑$\mathcal{EL}$的原始解释 $J=(\Delta^J,\cdot^J)$,J 对应的语义描述图 $G_J=(\Delta^J,E_J,L_J)$,其中:

- G_J 的节点集是论域Δ^J的元素;
- $E_J=\{(x,R,y)|(x,y)\in R^J\}$;
- $L_J(x)=\{P\in N_{prim}|x\in P^J\}$,对任意 $x\in\Delta^J$.

由上可以看出,$\mathcal{EL}$的原始解释 J 与语义描述图 G_J 之间也可以相互转化.

定义 18. 给定$\mathcal{EL}$的混合循环知识库 $HKB=(FT,TT,AB)$,TBox $TB=(FT,TT)$对应的语法描述图 $G_{f(TT)}=(N_{def}^{f(TT)},E_{f(TT)},L_{f(TT)})$,则 ABox AB 对应的语义描述图 $G_{AB}=(V_{AB},E_{AB},L_{AB})$,其中:

- G_{AB} 的节点集 V_{AB} 是 ABox AB 中的所有个体名和 TBox TB 中的所有被定义概念的集合,即

$$V_{AB}=N_{def}^{f(TT)}\cup\{a|a\text{ 是出现在 }AB\text{ 中的个体名}\};$$

- $E_{AB}=E_{f(TT)}\cup\{(a,R,b)|R(a,b)\in AB\}\cup\{(a,R,B)|A(a)\in AB,(A,R,B)\in E_{f(TT)}\}$;
- $L_{AB}(u)=\begin{cases}L_{f(TT)}(u), & \text{如果}u\in N_{def}^{f(TT)}\\ \{P\mid P(u)\in AB\}\cup\bigcup\limits_{A(u)\in AB}L_{f(TT)}(A), & \text{如果}u\in V_{AB}\setminus N_{def}^{f(TT)}\end{cases}$,其中,$P$ 是原始概念,A 是被定义概念.

下面给出如何利用描述图对$\mathcal{EL}$混合循环术语集的 LCS 和 MSC 进行推理.

3 εL 混合循环术语集的 LCS 和 MSC 推理

由于描述逻辑εL 的混合循环术语集有 3 种语义:最大不动点语义、最小不动点语义和描述语义,因而εL 混合循环术语集的 LCS 和 MSC 推理对应着也有 3 种:最大不动点语义下的 LCS 和 MSC 推理、最小不动点语义下的 LCS 和 MSC 推理以及描述语义下的 LCS 和 MSC 推理.因为最小不动点语义下εL 循环术语集解释为空集[8],因而研究最小不动点语义下εL 循环术语集的 LCS 和 MSC 推理没有实际意义.描述语义下εL 循环术语集的 LCS 和 MSC 不一定存在[13].因此,本文只给出最大不动点语义下的 LCS 和 MSC 推理.

3.1 最大不动点语义下的LCS推理

定理 6[9]. 给定εL 的混合循环知识库 $HKB=(FT,TT),A,B$ 是 TT 中的被定义概念,即 $A,B\in N_{def}^{TT}$,$f(TT)$是 TT 的 f-完全,则 $A\sqsubseteq_{gfp,FT,TT}B$,当且仅当 $A\sqsubseteq_{gfp,f(TT)}B$.

由定理 6 可知,最大不动点语义下εL 混合循环知识库 $HKB=(FT,TT)$的概念包含关系等价于循环知识库 $f(TT)$下的包含关系,即可以将含有基础 TBox 的混合循环知识库下的概念包含关系转化为不含基础 TBox 的循环知识库下的概念包含关系.

定理 7. 给定εL 的混合循环知识库 $HKB=(FT,TT),A,B$ 是 TT 中的被定义概念,即 $A,B\in N_{def}^{TT}$,$f(TT)$是 TT 的 f-完全,则 TT 中 A,B 的 gfp-LCS 等于 $f(TT)$中 A,B 的 gfp-LCS.

证明:不妨假设 $HKB'=(FT',TT')$是 $HKB=(FT,TT)$的保守扩充,TT'中的被定义概念 E 是 TT 中 A,B 的 gfp-LCS,则有:

(1) $A\sqsubseteq_{gfp,FT',TT'}E,B\sqsubseteq_{gfp,FT',TT'}E$;

(2) 如果 $HKB''=(FT'',TT'')$是 $HKB'=(FT',TT')$的保守扩充,F 是 TT''中的被定义概念,即 $F\in N_{def}^{TT''}$,并且 $A\sqsubseteq_{gfp,FT'',TT''}F,B\sqsubseteq_{gfp,FT'',TT''}F$,则 $E\sqsubseteq_{gfp,FT'',TT''}F$.

因为 $HKB'=(FT',TT')$是 $HKB=(FT,TT)$的保守扩充,从而 $f(TT')$是 $f(TT)$的保守扩充.因为 $A\sqsubseteq_{gfp,FT',TT'}E$, $B\sqsubseteq_{gfp,FT',TT'}E$,由定理 6 可知,$A\sqsubseteq_{gfp,f(TT')}E,B\sqsubseteq_{gfp,f(TT')}E$.

又因为 $HKB''=(FT'',TT'')$是 $HKB'=(FT',TT')$的保守扩充,从而 $f(TT'')$是 $f(TT')$的保守扩充.如果 $A\sqsubseteq_{gfp,f(TT'')}F$, $B\sqsubseteq_{gfp,f(TT'')}F$,由定理 6 可知,$A\sqsubseteq_{gfp,FT'',TT''}F,B\sqsubseteq_{gfp,FT'',TT''}F$,从而有 $E\sqsubseteq_{gfp,FT'',TT''}F$.又由定理 6 可知,$E\sqsubseteq_{gfp,f(TT'')}F$.因此,$E$ 是 $f(TT)$中 A,B 的 gfp-LCS.

反过来,不妨假设 $f(TT')$是 $f(TT)$的保守扩充,$f(TT')$中的被定义概念 E 是 $f(TT)$中 A,B 的 gfp-LCS,则:

(1) $A\sqsubseteq_{gfp,f(TT')}E,B\sqsubseteq_{gfp,f(TT')}E$;

(2) 如果$f(TT'')$是$f(TT')$的保守扩充,F 是$f(TT'')$中的被定义概念,即 $F\in N_{def}^{f(TT'')}$,并且 $A\sqsubseteq_{gfp,f(TT'')}F,B\sqsubseteq_{gfp,f(TT'')}F$,则 $E\sqsubseteq_{gfp,f(TT'')}F$.

因为$f(TT')$是$f(TT)$的保守扩充,从而 $HKB'=(FT',TT')$是 $HKB=(FT,TT)$的保守扩充,其中,FT'由定义 15 决定.因为 $A\sqsubseteq_{gfp,f(TT')}E,B\sqsubseteq_{gfp,f(TT')}E$,由定理 6 可知,$A\sqsubseteq_{gfp,FT',TT'}E,B\sqsubseteq_{gfp,FT',TT'}E$.

又因为$f(TT'')$是$f(TT')$的保守扩充,从而 $HKB''=(FT'',TT'')$是 $HKB'=(FT',TT')$的保守扩充,其中,FT''由定义 15 决定.如果 $A\sqsubseteq_{gfp,FT'',TT''}F,B\sqsubseteq_{gfp,FT'',TT''}F$,由定理 6 可知,$A\sqsubseteq_{gfp,f(TT'')}F,B\sqsubseteq_{gfp,f(TT'')}F$,从而 $E\sqsubseteq_{gfp,f(TT'')}F$.又由定理 6 可知, $E\sqsubseteq_{gfp,FT'',TT''}F$.因此,$E$ 是 TT 中 A,B 的 gfp-LCS. □

由定理 7 可以看出,εL 混合循环知识库 $HKB=(FT,TT)$的概念的 gfp-LCS 等于循环知识库 $f(TT)$的概念的 gfp-LCS,即可以将含有基础 TBox 的混合循环知识库下的概念的 gfp-LCS 推理转化为不含基础 TBox 的循环知识库下的概念的 gfp-LCS 推理.

定义 19. 给定εL 的描述图 $G_1=(V_1,E_1,L_1)$和 $G_2=(V_2,E_2,L_2)$,G_1 和 G_2 的乘积所得到的描述图 $G_1\times G_2=(V,E,L)$如下定义:

- $V=V_1\times V_2$;
- $E=\{((v_1,v_2),R,(v_1',v_2'))|(v_1,R,v_1')\in E_1\wedge(v_2,R,v_2')\in E_2\}$;

- $L(v_1,v_2)=L_1(v_1)\cap L_2(v_2)$.

例 6:假设图 3 表示的描述图为 G_{TT},则 G_{TT} 和 G_{TT} 的乘积所得到的描述图 $G_{TT}\times G_{TT}$ 如图 4 所示.

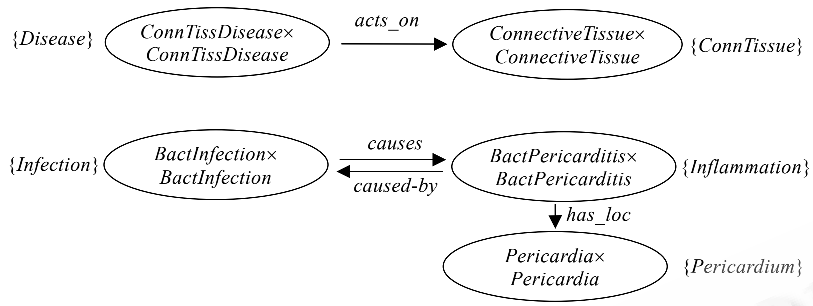

Fig.4 Product of description graphs

图 4 描述图的乘积

给定术语 TBox TT_1 及其对应的描述图 G_{TT_1},从而描述图 $G_{TT_1}\times G_{TT_1}$ 对应一个术语 TBox TT,并且 TT 所对应的描述图 $G_{TT}=G_{TT_1}\times G_{TT_1}$.令 $TT_2=TT_1\cup TT$,则 TT_2 是 TT_1 的保守扩充.因为 $G_{TT_1}\times G_{TT_1}$ 和 G_{TT_1} 基于相同的原始概念和关系,从而 TT_2 和 TT_1 具有相同的原始概念和关系.另外,TT 中的被定义概念集合是 $N_{def}^{TT_1}\times N_{def}^{TT_1}$,$TT_1$ 中的被定义概念集合是 $N_{def}^{TT_1}$,$N_{def}^{TT_1}\cap(N_{def}^{TT_1}\times N_{def}^{TT_1})=\varnothing$,$TT_1\subseteq TT_2$.

定理 8. 给定$\mathcal{EL}$的混合循环知识库 $HKB=(FT,TT)$,A,B 是 TT 中的被定义概念,即 $A,B\in N_{def}^{TT}$,$f(TT)$是 TT 的 f-完全,$f(TT)$所对应的描述图为 $G_{f(TT)}$,描述图 $G_{f(TT)}\times G_{f(TT)}$ 对应的术语 TBox 为 TT_1,令 $TT_2=f(TT)\cup TT_1$,则 TT 中 A,B 的 gfp-LCS 等于 TT_2 中的(A,B).

证明:由定理 7 可知,TT 中 A,B 的 gfp-LCS 等于 $f(TT)$中 A,B 的 gfp-LCS.由于 $f(TT)$是一个正规化的术语 TBox,从而由文献[10]的引理 13 可知,$f(TT)$中 A,B 的 gfp-LCS 等于 TT_2 中的(A,B). □

从定理 8 可以看出,$\mathcal{EL}$混合循环知识库 $HKB=(FT,TT)$的 gfp-LCS 推理能够转换为 TT 的 f-完全 $f(TT)$及其乘积所对应的术语 TBox 的 gfp-LCS 推理,而 TT 的 f-完全 $f(TT)$及其乘积所对应的术语 TBox 是一个不含基础 TBox 的循环知识库,因此,可以将含有基础 TBox 的混合循环知识库的 gfp-LCS 推理转化为不含基础 TBox 的循环知识库的 gfp-LCS 推理.

说明:为了求出混合循环知识库(FT,TT)中概念的 gfp-LCS,根据定理 8,首先需要求出 TT 的 f-完全 $f(TT)$所对应的描述图 $G_{f(TT)}$的乘积 $G_{f(TT)}\times G_{f(TT)}$,然后利用 $G_{f(TT)}\times G_{f(TT)}$所对应的 TBox 及 $f(TT)$即可求出 gfp-LCS.

文献[9]证明了将 $HKB=(FT,TT)$转化为 $f(TT)$可在多项式时间内完成.文献[10]证明了正规化术语 TBox 概念的 gfp-LCS 推理也能在多项式时间内完成.因为 $f(TT)$是一个正规化的术语 TBox,因而 $f(TT)$概念的 gfp-LCS 推理也能在多项式时间内完成.定理 5 给出了得到正规化的知识库是多项式时间的.因此有下面的定理:

定理 9. $\mathcal{EL}$混合循环知识库 $HKB=(FT,TT)$的 gfp-LCS 推理是多项式时间复杂的.

3.2 最大不动点语义下的MSC推理

定理 10. 给定$\mathcal{EL}$的混合循环知识库 $HKB=(FT,TT,AB)$,原始解释 $J=(\Delta^J,\cdot^J)$,被定义概念 $A\in N_{def}^{TT}$,实例个体 a,则 $AB\vdash_{gfp,FT,TT}A(a)$,当且仅当 $AB\vdash_{gfp,FT,f(TT)}A(a)$.

证明:令知识库 $HKB_f=(FT,f(TT),AB)$,首先证明基于 J 的 HKB 的最大不动点模型 I_{HKB} 等于基于 J 的 HKB_f 的最大不动点模型 I_{HKB_f},即证明 $I_{HKB}=I_{HKB_f}$,从而只需证明:(1) $I_{HKB}\preceq_J I_{HKB_f}$;(2) $I_{HKB_f}\preceq_J I_{HKB}$.

先证明(1).要证明 $I_{HKB}\preceq_J I_{HKB_f}$,只需证明 $I_{HKB}\vdash_{gfp,f(TT)}f(TT)$和 $I_{HKB}\vdash_{gfp,f(TT)}AB$ 即可.因为 I_{HKB} 是基于 J 的 HKB 的最大不动点模型,由定义 5 可知,$J\vdash FT$,$I_{HKB}\vdash_{gfp,TT}AB$,$I_{HKB}\vdash_{gfp,TT}TT$.不妨假设 A 在 TT 中的定义式是:$A\equiv P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n$.由 $I_{HKB}\vdash_{gfp,TT}TT$ 可知,$A^{I_{HKB}}=(P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n)^{I_{HKB}}$.又因为 $J\vdash$

FT,I_{HKB}是基于 J 的模型,由定义 15 可知,$A^{I_{HKB}} \subseteq f(A)^{I_{HKB}}$.从而有 $A^{I_{HKB}}=(P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n)^{I_{HKB}}\cap f(A)^{I_{HKB}}=(P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n\sqcap f(A))^{I_{HKB}}$.又因为 TT 的 f-完全 $f(TT)=\{A\equiv C\sqcap f(A)|A\equiv C\in TT\}$,从而有 $I_{HKB}\models_{gfp,f(TT)}f(TT)$.又因为 $I_{HKB}\models_{gfp,TT}AB$,所以对任意 $A(a)\in AB$,有 $a^{I_{HKB}}\in A^{I_{HKB}}$,从而有 $a^{I_{HKB}}\in(P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n)^{I_{HKB}}$.由 $A^{I_{HKB}}\subseteq f(A)^{I_{HKB}}$ 可知,$a^{I_{HKB}}\in f(A)^{I_{HKB}}$.因此,$a^{I_{HKB}}\in(P_1\sqcap\ldots\sqcap P_m\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_n.B_n\sqcap f(A))^{I_{HKB}}$.因为 $f(TT)=\{A\equiv C\sqcap f(A)|A\equiv C\in TT\}$,从而有 $I_{HKB}\models_{gfp,f(TT)}AB$.

再证明(2).要证明 $I_{HKB_f}\lesssim_J I_{HKB}$,只需证明对任意被定义概念 A,$A^{I_{HKB_f}}\subseteq A^{I_{HKB}}$ 和 $I_{HKB_f}\models_{gfp,TT}AB$ 即可.因为 I_{HKB_f} 是基于 J 的 HKB_f的最大不动点模型,由定义 5 可知,$J\models FT$, $I_{HKB_f}\models_{gfp,f(TT)}AB$, $I_{HKB_f}\models_{gfp,f(TT)}f(TT)$.并且由定理 1 和定义 3 可知,存在一个序数$\alpha$,使得 $I^{\downarrow\alpha}_{HKB_f}$ 是 HKB_f的最大不动点模型,即 $I_{HKB_f}=I^{\downarrow\alpha}_{HKB_f}$.同理可得 $I_{HKB}=I^{\downarrow\alpha}_{HKB_f}$ 是 HKB 的最大不动点模型.对序数α用超穷归纳法可以证明 $I_{HKB_f}\lesssim_J I_{HKB}$.

先证明⇒.因为 $AB\models_{gfp,FT,TT}A(a)$,所以对 HKB 的任意最大不动点模型 I,有 $a^I\in A^I$.因为 I 也是 HKB_f的最大不动点模型,所以 $AB\models_{gfp,FT,f(TT)}A(a)$.

再证明⇐.因为 $AB\models_{gfp,FT,f(TT)}A(a)$,所以对 HKB_f的任意最大不动点模型 I,有 $a^I\in A^I$.因为 I 也是 HKB 的最大不动点模型,所以,$AB\models_{gfp,FT,TT}A(a)$. □

由定理10可知,最大不动点语义下混合循环知识库(FT,TT,AB)的实例检测推理等价于混合循环知识库(FT, $f(TT)$,AB)的实例检测推理,即将混合循环知识库(FT,TT,AB)中的术语 TBox TT 替换成 TT 的 f-完全 $f(TT)$后得到的混合循环知识库(FT,$f(TT)$,AB)与混合循环知识库(FT,TT,AB)保持相同的实例隶属关系.

定理 11. 给定εL 的混合循环知识库 $HKB=(FT,TT,AB)$,原始解释 $J=(\Delta^J,\cdot^J)$,被定义概念 $A\in N_{def}$,实例个体 a,则 $AB\models_{gfp,FT,f(TT)}A(a)$,当且仅当 $AB\models_{gfp,f(TT)}A(a)$.

证明:先证明⇐.因为 $AB\models_{gfp,f(TT)}A(a)$,从而对 $f(TT)$的任意基于 J 的最大不动点模型 I,有 $a^I\in A^I$,并且 $J\models FT$,从而 $I\models FT$.因此,$AB\models_{gfp,FT,f(TT)}A(a)$.

再证明⇒.因为 $AB\models_{gfp,FT,f(TT)}A(a)$,从而对 $f(TT)$的任意基于 J 的最大不动点模型 I,有 $a^I\in A^I$,从而 $AB\models_{gfp,f(TT)}A(a)$. □

由定理 11 可以看出,最大不动点语义下混合循环知识库(FT,$f(TT)$,AB)的实例检测推理等价于循环知识库($f(TT)$,AB)的实例检测推理,即可以将含有基础 TBox 的混合循环知识库下的实例检测推理转化为不含基础 TBox 的循环知识库下的实例检测推理.由定理 10 和定理 11 可直接得到下列定理:

定理 12. 给定εL 的混合循环知识库 $HKB=(FT,TT,AB)$,原始解释 $J=(\Delta^J,\cdot^J)$,被定义概念 $A\in N_{def}$,实例个体 a,则 $AB\models_{gfp,FT,TT}A(a)$,当且仅当 $AB\models_{gfp,f(TT)}A(a)$.

由定理 12 可以看出,最大不动点语义下混合循环知识库(FT,TT,AB)的实例检测推理等价于循环知识库($f(TT)$,AB)的实例检测推理.

定理 13. 给定εL 的混合循环知识库 $HKB=(FT,TT,AB)$,a 是 HKB 的实例个体,$f(TT)$是 TT 的 f-完全,令 $HKB'=(f(TT),AB)$,则 HKB 中 a 的 gfp-MSC 等于 HKB'中 a 的 gfp-MSC.

证明:不妨假设 $HKB^*=(FT^*,TT^*)$是 $HKB=(FT,TT)$的保守扩充,TT^*中的被定义概念 E 是 HKB 中 a 的 gfp-MSC,则有:

(1) $AB\models_{gfp,FT^*,TT^*}E(a)$;

(2) 如果 $HKB^{**}=(FT^{**},TT^{**})$是 $HKB^*=(FT^*,TT^*)$的保守扩充,F 是 TT^{**}中的被定义概念,即 $F\in N_{def}^{TT^{**}}$,$AB\models_{gfp,FT^{**},TT^{**}}F(a)$,则 $E\sqsubseteq_{gfp,FT^{**},TT^{**}}F$.

因为 $HKB^*=(FT^*,TT^*)$是 $HKB=(FT,TT)$的保守扩充,从而 $f(TT^*)$是 $f(TT)$的保守扩充.因为 $AB\models_{gfp,FT^*,TT^*}E(a)$,由定理 12 可知,$AB\models_{gfp,f(TT^*)}E(a)$.

又因为 $HKB^{**}=(FT^{**},TT^{**})$是 $HKB^*=(FT^*,TT^*)$的保守扩充,从而 $f(TT^{**})$是 $f(TT^*)$的保守扩充.如果 $AB\models_{gfp,f(TT^{**})}F(a)$,由定理 12 可知,$AB\models_{gfp,FT^{**},TT^{**}}F(a)$,从而有 $E\sqsubseteq_{gfp,FT^{**},TT^{**}}F$.又由定理 12 可知,$E\sqsubseteq_{gfp,f(TT^{**})}F$.因此,$E$

是 HKB′中的 a 的 *gfp-MSC*.

反过来,不妨假设 $f(TT^*)$是 $f(TT)$的保守扩充,$f(TT^*)$中的被定义概念 E 是 $f(TT)$中 a 的 *gfp-MSC*,则:

(1) $AB \vDash_{gfp,f(TT^*)} E(a)$;

(2) 如果 $f(TT^{**})$是 $f(TT^*)$的保守扩充,F 是 $f(TT^{**})$中的被定义概念,即 $F\in N_{def}^{f(TT^{**})}$,并且 $AB \vDash_{gfp,f(TT^{**})} F(a)$,则 $E\sqsubseteq_{gfp,f(TT^{**})} F$.

因为 $f(TT^*)$是 $f(TT)$的保守扩充,从而 $HKB^*=(FT^*,TT^*)$是 $HKB=(FT,TT)$的保守扩充,其中,FT^*由定义 15 决定.因为 $AB \vDash_{gfp,f(TT^*)} E(a)$,由定理 12 可知,$AB \vDash_{gfp,FT^*,TT^*} E(a)$.

又因为 $f(TT^{**})$是 $f(TT^*)$的保守扩充,从而 $HKB^{**}=(FT^{**},TT^{**})$是 $HKB^*=(FT^*,TT^*)$的保守扩充,其中,FT^{**}由定义 15 决定.如果 $AB \vDash_{gfp,FT^{**},TT^{**}} F(a)$,由定理 12 可知,$AB \vDash_{gfp,f(TT^{**})} F(a)$,从而 $E\sqsubseteq_{gfp,f(TT^{**})} F$.又由定理 12 可知,$E\sqsubseteq_{gfp,FT^{**},TT^{**}} F$.因此,$E$ 是 HKB 中 a 的 *gfp-MSC*. □

由定理 13 可以看出,εL 混合循环知识库(FT,TT,AB)的个体的 *gfp-MSC* 等于循环知识库$(f(TT),AB)$的个体的 *gfp-MSC*,而$(f(TT),AB)$是一个不含基础 TBox 的循环知识库,因此,可以将含有基础 TBox 的混合循环知识库的 *gfp-MSC* 推理转化为不含基础 TBox 的循环知识库的 *gfp-MSC* 推理.

定理 14. 给定εL 的混合循环知识库 $HKB_1=(FT_1,TT_1,AB_1)$,a 是 HKB_1 的实例个体,$f(TT_1)$是 TT_1 的 f-完全,令 $HKB'=(f(TT_1),AB_1)$,AB_1 所对应的描述图为 G_{AB_1},并且将 G_{AB_1} 对应为术语 TBox TT_2 的描述图,即 TT_2 是一个术语 TBox,$G_{AB_1}=G_{TT_2}$,并且 AB_1 中的个体 b 对应于 TT_2 中的 C_b,则 HKB 中 a 的 *gfp-MSC* 等于 TT_2 中的 C_a.

证明:由定理 13 可知,HKB_1 中 a 的 *gfp-MSC* 等于 HKB′中 a 的 *gfp-MSC*.由于 $f(TT_1)$是一个正规化的术语 TBox,从而由文献[10]的引理 19 可知,HKB′中 a 的 *gfp-MSC* 等于 TT_2 中的 C_a. □

由定理 14 可知,εL 混合循环知识库(FT,TT,AB)的 *gfp-MSC* 推理能够转换为 TT 的 f-完全 $f(TT)$所对应的知识库$(f(TT),AB)$的 *gfp-LCS* 推理,而$(f(TT),AB)$是一个不含基础 TBox 的循环知识库,因此,可以通过不含基础 TBox 的循环知识库来计算混合循环知识库中个体的 *gfp-MSC*.

说明:为了求出混合循环知识库(FT_1,TT_1,AB_1)中个体的 *gfp-MSC*,根据定理 14,首先需要求出 TT_1 的 f-完全 $f(TT_1)$,再在混合知识库$(f(TT_1),AB_1)$下求出描述图 G_{AB_1},最后利用描述图 G_{AB_1} 所对应 TBox 即可求出 *gfp-MSC*.

文献[9]证明了将 $HKB=(FT,TT)$转化为 $f(TT)$能够在多项式时间内完成.文献[10]证明了正规化术语 TBox TT 和 ABox AB 所对应的知识库(TT,AB)的 *gfp-MSC* 推理也能在多项式时间内完成.因为 $f(TT)$是一个正规化的术语 TBox,因而 $f(TT)$和 ABox AB 所对应的知识库$(f(TT),AB)$的 *gfp-MSC* 推理也能在多项式时间内完成.定理 5 给出了得到正规化的知识库是多项式时间的.因此有下面的定理:

定理 15. εL 混合循环知识库 $HKB=(FT,TT,AB)$的 *gfp-MSC* 推理是多项式时间复杂的.

4 相关工作

目前,描述逻辑的研究集中在两方面:一是不带循环定义的描述逻辑;另一方面是描述逻辑循环定义[15].由于不带循环定义的描述逻辑的语义和推理相对比较简单,国内外许多学者对它都进行了深入研究[15−17].

描述逻辑循环定义是描述逻辑长期以来的研究难点[1,2],其中,描述逻辑循环定义最基本的问题(即语义及推理)都没有得到合理的解决,甚至在已给出的方法中存在错误.例如,不动点语义是循环术语集的理论基础,文献[15]给出了两个有关循环术语集不动点模型存在的命题(即命题 2.8 和命题 2.9),其中,命题 2.8 不含否定构造算子,命题 2.9 包含否定构造算子.这两个命题是循环术语集的非常基本的重要定理,已经证明命题 2.9 是错误的.在描述逻辑循环定义语义研究方面,目前采用 Nebel 提出的 3 种语义:最大不动点语义、最小不动点语义和描述语义[5,6].循环定义的推理包括标准推理和非标准推理[7].本文主要研究循环定义的非标准推理.

在标准推理方面,Nebel[5]和 Baader[3]利用自动机理论给出了描述逻辑 TL 和 FL_0 循环术语集的包含推理算法.Kusters 对 Baader[3]的工作做了进一步研究,使用自动机理论给出了描述逻辑 FLN 循环术语集的包含推理算

法[4].Baader[8]使用描述图及模拟关系给出了描述逻辑$\mathcal{EL}$循环术语集的包含推理算法.但 Nebel[5],Kusters[4]和Baader[3,8]只允许 TBox 出现概念等价公理,不允许出现 GCI 公理.为了循环定义能够处理 GCI 公理,Brandt[9]提出了描述逻辑$\mathcal{EL}$循环术语集混合 TBox 的概念,将 TBox 分成两部分:基础 TBox 和术语 TBox,其中,基础 TBox 允许出现 GCI 公理,用描述语义进行解释;术语 TBox 允许循环定义,用不动点语义进行解释,并给出了最大不动点语义下 TBox 约束的包含推理.

在非标准推理方面,Brandt[7]给出了最大不动点语义下描述逻辑$\mathcal{EL}$混合循环术语集匹配和近似推理的多项式时间算法.Baader[10]给出了最大不动点语义下描述逻辑$\mathcal{EL}$循环术语集的 LCS 和 MSC 推理问题的多项式时间算法,但 Baader[10]不允许 TBox 出现 GCI 公理,即 Baader F[10]没有给出$\mathcal{EL}$混合循环术语集的 LCS 和 MSC 推理算法.因此,Baader[10]的工作在实际应用中受到很大的制约,因为在实际应用中,TBox 完全由等价公理组成是不现实的,必须包含 GCI 公理[9].本文进一步研究了描述逻辑$\mathcal{EL}$混合循环术语集的 LCS 和 MSC 推理问题,给出了$\mathcal{EL}$混合循环术语集的语法和语义,针对混合循环术语集 LCS 和 MSC 推理的需求,提出了 TBox-完全和描述图的概念,利用 TBox-完全和描述图给出了最大不动点语义下$\mathcal{EL}$混合循环术语集 LCS 和 MSC 的推理算法,并给出了推理算法正确性证明和复杂性定理.可以看出,本文的工作是目前描述逻辑$\mathcal{EL}$循环术语集 LCS 和 MSC 推理最全面的结果,Baader[10]的结果是本文结果的特殊情况(不包含 GCI 公理).

5 结束语

本文分析了描述逻辑循环术语集的研究现状和存在的问题,在 Baader 的基础上,进一步研究了描述逻辑$\mathcal{EL}$混合循环术语集的 LCS 和 MSC 推理问题.给出了$\mathcal{EL}$混合循环术语集的语法和语义(不动点语义和描述语义).针对$\mathcal{EL}$混合循环术语集 LCS 和 MSC 推理的需要,提出了 TBox-完全的概念,并重新定义了描述图(语法描述图和语义描述图).使用 TBox-完全和描述图给出了最大不动点语义下$\mathcal{EL}$混合循环术语集 LCS 和 MSC 的推理算法,给出了推理算法正确性证明和复杂性性质.进一步的工作主要是扩充$\mathcal{EL}$循环术语集及混合循环术语集.

致谢 向对本文提出宝贵意见的评审专家表示衷心的感谢,对广西师范大学计算机科学与信息工程学院语义 Web 与描述逻辑讨论班上的老师和同学表示感谢.

References:

[1] Giacomo GD, Lenzerini M. A uniform framework for concept definitions in description logics. Journal of Artificial Intelligence Research, 1997,6(1):87−110.

[2] Buchheit M, Donini FM, Nutt W, Schaerf A. A refined architecture for terminological systems: Terminology=schema+views. Artificial Intelligence, 1998,99(2):209−260.

[3] Baader F. Using automata theory for characterizing the semantics of terminological cycles. Annals of Mathematics and Artificial Intelligence, 1996,18(2-4):175−219.

[4] Kusters R. Characterizing the semantics of terminological cycles in ALN using finite automata. In: Cohn AG, Schubert L, Shapiro SC, eds. Proc. of the 6th Int'l Conf. on Principles of Knowledge Representation and Reasoning. San Francisco: Morgan Kaufmann Publishers, 1998. 499−511.

[5] Nebel B. Terminological cycles: Semantics and computational properties. In: Sowa JF, ed. Principles of Semantic Networks. San Francisco: Morgan Kaufmann Publishers, 1991. 331−362.

[6] Nebel B. Reasoning and revision in hybrid representation systems. LNAI 422, Berlin: Springer-Verlag, 1990. 125−156.

[7] Brandt S. Standard and non-standard reasoning in description logics [Ph.D. Thesis]. Dresden: Dresden University of Technology, 2006.

[8] Baader F. Terminological cycles in a description logic with existential restrictions. In: Gottlob G, Walsh T, eds. Proc. of the 18th Int'l Joint Conf. on Artificial Intelligence (IJCAI 2003). San Francisco: Morgan Kaufmann Publishers, 2003. 325−330.

[9] Brandt S, Model J. Subsumption in $\mathcal{EL}$ w.r.t. hybrid TBoxes. In: Furbach U, ed. Proc. of the 28th Annual German Conf. on Artificial Intelligence. LNCS 3698, Berlin: Springer-Verlag, 2005. 34−48.

[10] Baader F. Least common subsumers and most specific concepts in a description logic with existential restrictions and terminological cycles. In: Gottlob G, Walsh T, eds. Proc. of the 18th Int'l Joint Conf. on Artificial Intelligence (IJCAI 2003). San Francisco: Morgan Kaufmann Publishers, 2003. 319–324.
[11] Baader F, Kusters R. Non-Standard inferences in description logics: The story so far. In: Gabbay DM, Goncharov SS, Zakharyaschev M, eds. Mathematical Problems from Applied Logic. Berlin: Springer-Verlag, 2004. 1–75.
[12] Tarski A. A lattice-theoretical fixpoint theorem and its applications. Pacific Journal of Mathematics, 1955,5(2):285–309.
[13] Baader F. Least common subsumers, most specific concepts, and role-value-maps in a description logic with existential restrictions and terminological cycles. LTCS-Report, 02-07, Dresden: Dresden University of Technology, 2002. 1–47.
[14] Baader F. Terminological cycles in a description logic with existential restrictions. LTCS-Report, 02-02, Dresden: Dresden University of Technology, 2002. 1–34.
[15] Baader F, Nutt W. Basic description logics. In: Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P, eds. The Description Logic Handbook: Theory, Implementation and Applications. Cambridge: Cambridge University Press, 2003. 47–100.
[16] 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
[17] Jiang YC, Tang Y, Wang J, Zhou SM. Semantic Web oriented description logic. Pattern Recognition and Artificial Intelligence, 2007,20(1):48–54 (in Chinese with English abstract).

附中文参考文献:

[16] 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
[17] 蒋运承,汤庸,王驹,周生明.面向语义 Web 的描述逻辑.模式识别与人工智能,2007,20(1):48–54.

蒋运承(1974－),男,广西桂林人,博士,教授,主要研究领域为描述逻辑,语义 Web,Web 智能.

王驹(1950－),男,博士,研究员,博士生导师,主要研究领域为描述逻辑,数理逻辑,人工智能.

周生明(1956－),男,副教授,主要研究领域为人工智能,描述逻辑,数理逻辑.

汤庸(1964－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数据库,知识工程,CSCW.