ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, No.3, March 2009, pp.477−490
doi: 10.3724/SP.J.1001.2009.03215

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

描述逻辑εLN 循环术语集的不动点语义及推理*

蒋运承[1,2+], 王 驹[1], 史忠植[3], 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)
[3](中国科学院 计算技术研究所,北京 100190)

Fixpoint Semantics and Reasoning of Terminological Cycles in Description Logic εLN

JIANG Yun-Cheng[1,2+], WANG Ju[1], SHI Zhong-Zhi[3], 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)
[3](Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)
+ Corresponding author: E-mail: ycjiang@mailbox.gxnu.edu.cn

Jiang YC, Wang J, Shi ZZ, Tang Y. Fixpoint semantics and reasoning of terminological cycles in description logic εLN. *Journal of Software*, 2009,20(3):477−490. http://www.jos.org.cn/1000-9825/3215.htm

Abstract: Terminological cycles have been a very hard problem in description logics for a long time, and their essential problems, i.e. semantics and reasoning problems, have not been solved reasonably. Current research progress and the existing problems of terminological cycles in description logics are analyzed in this paper. Based on the work of Baader, a new direction in terminological cycles is put forward. Aiming at more expressive description logic, the semantics and reasoning mechanism of terminological cycles are studied. The number restrictions are added to description logic εL, and the descripion logic εLN is presented. The semantics (fixpoint semantics and descriptive semantics) of εLN are given. To meet the requirement of εLN, the description graphs (syntax description graph and semantics description graph) are redefined. The satisfiability and subsumption reasoning algorithms of terminological cycles in description logic εLN w.r.t. fixpoint semantics are presented with simulation between description graphs. It is proved that the satisfiability and subsumption reasoning algorithms of terminological cycles in εLN are polynomial.
Key words: description logics; εLN; terminological cycle; description graph; simulation; fixpoint semantics

摘 要: 循环术语集是描述逻辑长期以来的研究难点,其最基本的问题即语义及推理问题没有得到合理的解决.

∗ 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 National Basic Research Program of China under Grant No.2003CB317004 (国家重点基础研究发展计划(973)); the Postdoctoral Science Foundation of China under Grant No.20060400226 (中国博士后科学基金); 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 Program for New Century Excellent Talentsin University of China (新世纪优秀人才支持计划); the Natural Science Foundation of Guangxi Province of China under Grant Nos.0640030, 0832103 (广西自然科学基金)

Received 2006-11-10; Accepted 2007-10-25

分析了描述逻辑循环术语集的研究现状和存在的问题,将 Baader 的工作扩展到新的方向.针对更大的描述逻辑系统研究了循环术语集的语义及推理机制,即在描述逻辑εL 的基础上添加数量约束构造算子,提出了描述逻辑εLN,给出了εLN 的语义(包括不动点语义和描述语义).针对εLN 的需要,重新定义了描述图(包括语法描述图和语义描述图).使用描述图之间的模拟关系给出了不动点语义下εLN 循环术语集的可满足性和包含关系推理算法,并证明了推理算法是多项式时间复杂的.

关键词: 描述逻辑;εLN;循环术语集;描述图;模拟关系;不动点语义
中图法分类号: TP301 **文献标识码**: A

循环术语集(也称为循环定义)是描述逻辑长期以来的研究难点,其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].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,经典的描述逻辑论著《The Description Logic Handbook: Theory, Implementation and Applications》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)[3],其中命题 2.8 不包含否定构造算子,命题 2.9 包含否定构造算子,这两个命题是循环术语集非常基本的重要定理,我们已经证明命题 2.9 是错误的.

由于描述逻辑循环术语集的语义及推理问题没有得到合理的解决,因此在目前已实现的描述逻辑推理系统中都给出了强制规定:描述逻辑知识库的 TBox 中不允许出现循环定义.但循环定义可以大大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[4−6].例如,二叉树必须如下定义:$Binary\text{-}tree \equiv Tree \sqcap (\geq 2\ branch) \sqcap (\leq 2\ branch) \sqcap \forall branch.Binary\text{-}tree$.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[6,7].因此,无论从理论上还是从实际应用上讲,研究描述逻辑循环定义都相当有意义.

给定一个带循环定义的 TBox,主要有 3 个问题需要解决:(1) 它是否有模型;(2) 如何构造一个(或一类)模型;(3) 如何给出判断两个被定义概念 A,B 之间包含关系的推理算法.目前,关于(1)的最好结果是:如果 TBox 的定义式中不包含否定构造算子,则 TBox 有不动点模型.如何将此结果推广到包含否定构造算子的情形,仍然是一个未解决的问题.对于(2)和(3),目前只针对一些很小的不带否定构造算子的描述逻辑系统.Nebel 利用自动机给出了 TL 循环术语集的可满足性和包含推理算法,其中 TL 只包含两个构造算子:交和全称量词,但在描述语义下,Nebel 的工作是针对有限解释的[6].针对无限解释,Baader 也利用自动机给出了 FL_0 循环术语集的可满足性和包含推理算法,其中 FL_0 也只包含两个构造算子:交和全称量词[4].Kusters 使用自动机给出了 FLN 循环术语集的可满足性和包含推理算法,其中,FLN 只包含 4 个构造算子:交、全称量词、最大数量约束和最小数量约束[5].Baader 使用描述图及描述图之间的模拟关系给出了εL 循环术语集的可满足性和包含推理算法,其中,εL 只包含两个构造算子:交和存在量词[8,9].可以看出,描述逻辑 TL,FL_0 和εL 都是很小的系统,并且都不带否定构造算子,这说明描述逻辑循环定义理论框架研究的困难.

本文将对描述逻辑循环定义作进一步研究,在更大的描述逻辑系统中研究循环术语集的语义及推理机制,即在描述逻辑εL[8]的基础上添加数量约束构造算子[10,11],得到的描述逻辑称为εLN,给出εLN 的语法和语义.针对εLN 的需要,重新定义描述图.使用描述图之间的模拟关系给出不动点语义下εLN 循环术语集的可满足性和包含推理算法,并证明推理算法是多项式时间复杂的.显然,本文的工作是εL[8]不动点语义下的扩充.

1 εLN 循环术语集

描述逻辑εLN 是εL[8]的扩充,增加了最大和最小数量约束构造算子,即εLN 的概念可以如下定义:

$C,D \rightarrow \top \mid A \mid C \sqcap D \mid \exists R.C \mid \geq k\ R \mid \leq k\ R$,其中 A 表示原子概念,C,D 表示概念,R 表示关系名,$k \in N$.

下面用 N_C 表示εLN 的所有概念名的集合,N_R 表示εLN 的所有关系名的集合.

εLN 的语义将概念解释为一定论域的子集,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εLN 的语义解释如下:

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

εLN 的 TBox 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,即不允许出现下列情形:对于概念名 A,存在两个不相同的概念描述 D_1 和 D_2,使得 $A\equiv D_1$ 和 $A\equiv D_2$.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.并且约定对 TBox 中任意出现的数量约束中的关系集合与存在约束中的关系集合的交集为空集.

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 T 中的所有概念定义,即对任意的概念定义 $A\equiv D\in T,A^I=D^I$,则称 I 是 TBox T 的一个模型.如果 TBox T 存在一个模型,则称 TBox T 是可满足的.

上述定义的语义就是 Nebel 提出的描述语义[6].Nebel 已证明:在没有循环定义的描述逻辑中,给定所有原始概念的解释后能够得到唯一一个模型,但在循环定义中可能存在多个模型[6].

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,Nebel 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[6].

给定关系 $R\in N_R$,$\bigcup_{n\geq 1}R^n$ 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,或 TBox 存在循环定义(循环术语集).形式上说,假设 A 是被定义概念,B 是概念,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,或 TBox 存在循环定义(循环术语集).

例 1:TBox T: $Human\equiv Mammal\sqcap(\geq 2\ parents)\sqcap(\leq 2\ parents)\sqcap\exists parents.Human$.该 TBox T 中,$Human$ 是被定义概念,并且 $Human$ 使用 $Human$(因为 $Human$ 在定义式的右边也出现),从而该 TBox T 是循环的.

给定描述逻辑εLN 的 TBox T,T 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 T 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 T 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 $A_i\in N_{def}$.给定原始解释 J,定义集合 $Int(J)=\{I|\ I$ 是基于 J 的解释$\}$.

与描述逻辑 ALCN 循环定义[3]类似,可以如下为 $Int(J)$上的解释定义偏序关系$\preceq_J$:若 $I_1,I_2\in Int(J)$,则 $I_1\preceq_J I_2$,当且仅当 $A_i^{I_1}\subseteq A_i^{I_2}$,$1\leq i\leq k$.

由于εLN 是 ALCN 的子系统,所以由 Baader 的结果[3]可知:$\langle Int(J),\preceq_J\rangle$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.由 Tarski 不动点定理[12]可知:若存在单调函数 $O:Int(J)\rightarrow Int(J)$,使得当 $I_1\preceq_J I_2$ 时有 $O(I_1)\preceq_J O(I_2)$,则 O 存在不动点,即存在解释 $I\in Int(J)$,使得 $O(I)=I$,并且 O 存在最小不动点和最大不动点.

与描述逻辑 ALCN 循环定义[3]类似,给定εLN 的 TBox $T=\{A_1\equiv D_1,\ldots,A_k\equiv D_k\}$,以及原始解释 J,可以如下为 $Int(J)$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.又由 Baader 的结果[3]可知,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 T 的模型.因此,可如下定义εLN 的不动点语义:

定义 1. 给定εLN 的 TBox T,T 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 T 的模型可能有多个,最大不动点语义(最

小不动点语义)仅仅接受最大不动点模型(最小不动点模型)作为 T 的模型.下面将最大不动点模型(最小不动点模型)称为 *gfp*-模型(*lfp*-模型).

说明:I 是函数 $O_{T,J}$ 的最大不动点,当且仅当 $O_{T,J}(I)=I$,并且对任意 I',如果 $O_{T,J}(I')=I'$,则有 $I'\lesssim_J I$. I 是函数 $O_{T,J}$ 的最小不动点,当且仅当 $O_{T,J}(I)=I$,并且对任意 I',如果 $O_{T,J}(I')=I'$,则有 $I\lesssim_J I'$.

定义 2. 给定εLN 的 TBox T,A,B 是 T 中的被定义概念,则:

- 在最大不动点语义下,B 包含 A(记为 $A\sqsubseteq_{gfp,T}B$),当且仅当对 T 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 $A\sqsubseteq_{lfp,T}B$),当且仅当对 T 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 $A\sqsubseteq_{T}B$),当且仅当对 T 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(downward continuous)或向上连续(upward continuous),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,其中ω表示第 1 个无限序数,即非负整数的序类型[13].

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

(2) 向下连续:给定一个基于原始解释 J 的递减解释链 *chain*:$I_0\gtrsim_J I_1\gtrsim_J I_2\gtrsim_J\ldots$,同理可得解释链 *chain* 的下确界(*glb*)是一个基于 J 的解释 I,满足 $A_i^I=\bigcap_{j\ge 0}A_i^{I_j}$,其中 $1\le i\le k$,则函数 $O:Int(J)\rightarrow Int(J)$是向下连续的,当且仅当 $O(glb(\{I_j|\ j\ge 0\}))=glb(\{O(I_j)|j\ge 0\})$.

在给出函数 $O_{T,J}$ 最大不动点(最小不动点)的构造方法之前,先给出一个预备定理,该定理是 Baader 的结果的扩充[9],即将 Baader 的结果推广到描述逻辑εLN 中.

定理 1. 给定εLN 的 TBox T,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.

证明:有关函数 $O_{T,J}$ 的向下连续性,Baader[9]已证明在描述逻辑εL 中函数 $O_{T,J}$ 不是向下连续的.由于εL 是εLN 的子系统,从而在描述逻辑εLN 中函数 $O_{T,J}$ 也不是向下连续的.

下面证明函数 $O_{T,J}$ 是向上连续的.

假设 $I_0\lesssim_J I_1\lesssim_J I_2\lesssim_J\ldots$基于 J 的递增解释链,$I_0,I_1,I_2,\ldots\in Int(J)$,$I$ 是它的上确界.T 包含原始概念$\{P_1,\ldots,P_n\}$、关系$\{R_1,\ldots,R_m\}$和被定义概念$\{A_1,\ldots,A_k\}$,并且 $I_j=(P_1^{I_j},\ldots,P_n^{I_j},R_1^{I_j},\ldots,R_m^{I_j},A_1^{I_j},\ldots,A_k^{I_j})$.由于$\{P_1,\ldots,P_n\}$或$\{R_1,\ldots,R_m\}$是原始概念或关系,因而 $P_i^{I_0}=P_i^{I_1}=\ldots$, $R_i^{I_0}=R_i^{I_1}=\ldots$ 又因为 $I_0\lesssim_J I_1\lesssim_J I_2\lesssim_J\ldots$,所以由$\lesssim_J$的定义可知 $A_i^{I_0}\subseteq A_i^{I_1}\subseteq\ldots$ 要证明 $O_{T,J}$ 是向上连续的,只需证明 $O_{T,J}(lub(\{I_j|\ j\ge 0\}))=lub(\{O_{T,J}(I_j)|\ j\ge 0\})$即可.

由于 $O_{T,J}(I_j)=I_{j+1}$,所有 $lub(\{O_{T,J}(I_j)|\ j\ge 0\})=lub(\{I_j|\ j\ge 1\})=lub(\{I_j|\ j\ge 0\})=\left(\bigcup_{j\ge 0}P_1^{I_j},\ldots,\bigcup_{j\ge 0}P_n^{I_j},\bigcup_{j\ge 0}R_1^{I_j},\ldots,\bigcup_{j\ge 0}R_m^{I_j},\bigcup_{j\ge 0}A_1^{I_j},\ldots,\bigcup_{j\ge 0}A_k^{I_j}\right)$.又因为 I 是 $I_0,I_1,I_2,\ldots$的上确界,因而 $O_{T,J}(lub(\{I_j|j\ge 0\}))=O_{T,J}(I)=(P_1^I,\ldots,P_n^I,R_1^I,\ldots,R_m^I,A_1^I,\ldots,A_k^I)$.由于 $P_i^{I_0}=P_i^{I_1}=\ldots$ 和 $R_i^{I_0}=R_i^{I_1}=\ldots$,所以有 $\bigcup_{j\ge 0}P_i^{I_j}=P_i^I$, $\bigcup_{j\ge 0}R_i^{I_j}=R_i^I$.

因而下面只需对任意被定义概念 D,证明 $D^I=\bigcup_{j\ge 0}D^{I_j}$ 即可.下面对概念 D 进行归纳证明:

如果 $D\equiv P$,其中 P 是原始概念,则 $D^I=P^J=D^{I_j}$,$j\ge 0$,从而有$\bigcup_{j\ge 0}D^{I_j}=P^J=D^I$.

如果 $D\equiv C_i$,其中 C_i 是被定义概念,则 $D^I=C_i^I$,并且对任意的 $j\ge 0$, $D^{I_j}=C_i^{I_j}$.又因为 I 是基于 J 的解释,从而由归纳假设有 $C_i^I=\bigcup_{j\ge 0}C^{I_j}$,所以 $D^I=\bigcup_{j\ge 0}C^{I_j}=\bigcup_{j\ge 0}D^{I_j}$.

如果 $D\equiv E\sqcap F$,其中 E 和 F 是概念描述,则 $D^I=(E\sqcap F)^I=E^I\cap F^I$.由归纳假设可知,$E^I=\bigcup_{j\ge 0}E^{I_j}$ 和 $F^I=\bigcup_{j\ge 0}F^{I_j}$.因此 $D^I=(\bigcup_{j\ge 0}E^{I_j})\cap(\bigcup_{j\ge 0}F^{I_j})=\bigcup_{j\ge 0}(E^{I_j}\cap F^{I_j})=\bigcup_{j\ge 0}D^{I_j}$.

如果 $D\equiv\exists R.C$,其中 R 是关系,C 是概念描述,则$\exists R.C$ 的语义解释可知,$D^I=(\exists R.C)^I=\{x\in\Delta^I|\ \exists y\in\Delta^I,(x,y)\in R^I\wedge y\in C^I\}$.由归纳假设可知,$D^I=(\exists R.C)^I=\{x\in\Delta^J|\ \exists y\in\Delta^J,(x,y)\in R^J\wedge y\in\bigcup_{j\ge 0}C^{I_j}\}$.也就是说,$x\in D^I$,当且仅当

$\exists y((x,y)\in R^J\wedge\forall j(y\in\bigcup_{j\ge0}C^{I_j}))$.由于$j$没有在$R$中出现,因此有$x\in D^I$,当且仅当$\forall j\exists y((x,y)\in R^J\wedge y\in\bigcup_{j\ge0}C^{I_j})$.又因为$\{x\in\Delta^J|\ \exists y\in\Delta^J,(x,y)\in R^J\wedge y\in C^{I_j}\}=D^{I_j}$,所以$D^I=\{x\in\Delta^J|\ \exists y\in\Delta^J,(x,y)\in R^J\wedge y\in\bigcup_{j\ge0}C^{I_j}\}=\bigcup_{j\ge0}(\{x\in\Delta^J|\ \exists y\in\Delta^J,(x,y)\in R^J\wedge y\in C^{I_j}\})=\bigcup_{j\ge0}D^{I_j}$.

如果$D\equiv\ge k\ R$,其中R是关系,则$\ge k\ R$的语义解释可知,$D^I=(\ge k\ R)^I=\{x\in\Delta^I|\ |\{y|(x,y)\in R^I\}|\ge k\}$.又因为$I$是基于$J$的解释,所以$D^I=\{x\in\Delta^J|\ |\{y|(x,y)\in\bigcup_{j\ge0}R^{I_j}\}|\ge k\}=\bigcup_{j\ge0}(\{x\in\Delta^J|\ |\{y|(x,y)\in R^{I_j}\}|\ge k\})=\bigcup_{j\ge0}D^{I_j}$.

如果$D\equiv\le k\ R$,其中R是关系,则$\le k\ R$的语义解释可知,$D^I=(\le k\ R)^I=\{x\in\Delta^I|\ |\{y|(x,y)\in R^I\}|\le k\}$.又因为$I$是基于$J$的解释,所以$D^I=\{x\in\Delta^J|\ |\{y|(x,y)\in\bigcup_{j\ge0}R^{I_j}\}|\le k\}=\bigcup_{j\ge0}(\{x\in\Delta^J|\ |\{y|(x,y)\in R^{I_j}\}|\le k\})=\bigcup_{j\ge0}D^{I_j}$.

所以对任意被定义概念D,证明$D^I=\bigcup_{j\ge0}D^{I_j}$成立,从而有:$\left(\bigcup_{j\ge0}P_1^{I_j},\dots,\bigcup_{j\ge0}P_n^{I_j},\ \bigcup_{j\ge0}R_1^{I_j},\dots,\bigcup_{j\ge0}R_m^{I_j},\ \bigcup_{j\ge0}A_1^{I_j},\dots,\bigcup_{j\ge0}A_k^{I_j}\right)=(P_1^I,\dots,P_n^I,\ R_1^I,\dots,R_m^I,\ A_1^I,\dots,A_k^I)$.即$O_{T,J}(lub(\{I_j|\ j\ge0\}))=lub(\{O_{T,J}(I_j)|\ j\ge0\})$成立. □

定义 3. 给定εLN的TBox T,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$A_i^{I_{top}}=\Delta^J$, $A_i^{I_{bot}}=\phi,1\le i\le k$,则对任意序数$\alpha$:

- 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$I^{\uparrow\alpha}=lub(\{I^{\uparrow\beta}|\ \beta<\alpha\}),I^{\downarrow\alpha}=glb(\{I^{\downarrow\beta}|\ \beta<\alpha\})$.

因为函数$O_{T,J}$是向上连续的,由Tarski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,$I^{\downarrow\alpha}$是函数$O_{T,J}$的最大不动点.

2 描述图

2.1 TBox的正规化

给定εLN的TBox T,N_{def},N_{prim}和N_{role}分别表示T中被定义概念、原始概念和关系的集合,其中$N_{prim}=N_{nprim}\cup N_{gprim}\cup N_{lprim}$,$N_{nprim}\cap N_{gprim}\cap N_{lprim}=\varnothing$,$N_{nprim}$表示原始概念名的集合$\{P_1,\dots,P_w\}$,$N_{gprim}$表示最大数量约束原始概念的集合$\{\ge k_1\ R_1,\dots,\ge k_u\ R_u\}$,$N_{lprim}$表示最小数量约束原始概念的集合$\{\le k_1\ R_1,\dots,\le k_v\ R_v\}$.$N_{gprim}\cup N_{lprim}$称为数量约束原始概念,记为$N_{rprim}$,即$N_{rprim}=N_{gprim}\cup N_{lprim}$.

定义 4. 给定εLN的TBox T,T是正规化的,当且仅当对任意的概念定义$A\equiv D\in T,A\in N_{def},D$具有下列形式:

$$P_1\sqcap\dots\sqcap P_w\sqcap GP_1\sqcap\dots\sqcap GP_u\sqcap LP_1\sqcap\dots\sqcap LP_v\sqcap\exists R_1.B_1\sqcap\dots\sqcap\exists R_z.B_z,$$

其中$w,u,v,z\ge0,P_1,\dots,P_w\in N_{nprim},GP_1,\dots,GP_u\in N_{gprim},LP_1,\dots,LP_v\in N_{lprim},R_1,\dots,R_z\in N_{role},B_1,\dots,B_z\in N_{def}$.

如果$w=u=v=z=0$,则$D\equiv\top$.

例 2:给定εLN的TBox T如下(参考了文献[9]):

$A_1\equiv P_1\sqcap(\ge k_1\ R_1)\sqcap A_2\sqcap\exists R_2.\exists R_3.A_3$;

$A_2\equiv P_2\sqcap(\le k_2\ R_4)\sqcap A_3\sqcap\exists R_3.\exists R_2.A_1$;

$A_3\equiv P_3\sqcap(\ge k_3\ R_5)\sqcap A_2\sqcap\exists R_2.(P_1\sqcap P_2\sqcap(\le k_4\ R_6))$.

显然,上述TBox T不是正规化的.通过引入新的被定义概念,T可以转化为如下的T_1:

$A_1\equiv P_1\sqcap(\ge k_1\ R_1)\sqcap A_2\sqcap\exists R_2.B_1$;

$B_1\equiv\exists R_3.A_3$;

$A_2\equiv P_2\sqcap(\le k_2\ R_4)\sqcap A_3\sqcap\exists R_3.B_2$;

$B_2\equiv\exists R_2.A_1$;

$A_3\equiv P_3\sqcap(\ge k_3\ R_5)\sqcap A_2\sqcap\exists R_2.B_3$;

$B_3\equiv P_1\sqcap P_2\sqcap(\le k_4\ R_6)$.

由于在A_1的定义式右边出现了A_2,而A_2是被定义概念.A_2,A_3的定义式的右边也出现了被定义概念A_3,A_2.

因而 T_1 也不是正规化的.

因为 $A_2\equiv P_2\sqcap(\le k_2\ R_4)\sqcap A_3\sqcap\exists R_3.B_2$,所以 $A_2\sqsubseteq A_3$.又因为 $A_3\equiv P_3\sqcap(\ge k_3\ R_5)\sqcap A_2\sqcap\exists R_2.B_3$,所以 $A_3\sqsubseteq A_2$.因此 A_2 与 A_3 等价,即 $A_2\equiv A_3$.又因为 $A_2\sqsubseteq P_2\sqcap(\le k_2\ R_4)\sqcap P_3\sqcap(\ge k_3\ R_5)\sqcap\exists R_3.B_2\sqcap\exists R_2.B_3$,因此,在最大不动点语义下,$T_1$ 可以转化为如下正规化的 T_2:

$A_1\equiv P_1\sqcap P_2\sqcap P_3\sqcap(\ge k_1\ R_1)\sqcap(\le k_2\ R_4)\sqcap(\ge k_3\ R_5)\sqcap\exists R_2.B_1\sqcap\exists R_3.B_2\sqcap\exists R_2.B_3$;

$B_1\equiv\exists R_3.A_2$;

$A_2\equiv P_2\sqcap P_3\sqcap(\le k_2\ R_4)\sqcap(\ge k_3\ R_5)\sqcap\exists R_3.B_2\sqcap\exists R_2.B_3$;

$B_2\equiv\exists R_2.A_1$;

$A_3\equiv P_2\sqcap P_3\sqcap(\le k_2\ R_4)\sqcap(\ge k_3\ R_5)\sqcap\exists R_3.B_2\sqcap\exists R_2.B_3$;

$B_3\equiv P_1\sqcap P_2\sqcap(\le k_4\ R_6)$.

因为 $A_2\sqsubseteq P_2\sqcap(\le k_2\ R_4)\sqcap P_3\sqcap(\ge k_3\ R_5)\sqcap\exists R_3.B_2\sqcap\exists R_2.B_3$,所以在最小不动点语义下,$A_2$ 是不可满足的,即 A_2 的解释是空集$\varnothing$.从而由 TBox T_1 中的概念定义式可知:A_1,A_3,B_1,B_2 的解释都为空集$\varnothing$.因此,在最小不动点语义下,T_1 可以转化为如下正规化的 T_3:

$B_3\equiv P_1\sqcap P_2\sqcap(\le k_4\ R_6)$.

当然,T_3 也可以具有 T_2 一样的形式.

将例 2 的正规化过程进行一般化推广,可以得到如下 TBox T 的正规化方法:

(1) 引入新的被定义概念,对任意的概念定义 $A\equiv D\in T,A\in N_{def},D$ 具有下列形式:

$$P_1\sqcap\ldots\sqcap P_w\sqcap GP_1\sqcap\ldots\sqcap GP_u\sqcap LP_1\sqcap\ldots\sqcap LP_v\sqcap C_1\sqcap\ldots\sqcap C_t\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_z.B_z,$$

其中 $w,u,v,z\ge 0,P_1,\ldots,P_w\in N_{nprim},GP_1,\ldots,GP_u\in N_{gprim},LP_1,\ldots,LP_v\in N_{lprim},R_1,\ldots,R_z\in N_{role},C_1,\ldots,C_t,B_1,\ldots,B_z\in N_{def}$.

把得到的 TBox 记为 T_1.

(2) 将 TBox T_1 转化为如下的图 $G=(V,E)$:

- 图 G 的节点集 V 是 T_1 中所有的被定义概念;
- 图 G 的有向边如下定义:从被定义概念 A 到被定义概念 B 存在一条边,当且仅当 B 出现在 A 的定义式的右边.所有有向边的集合构成图 G 的边集 E.

利用图 G,引入下列定义:

- $B\precsim A$,当且仅当在图 G 中存在一条从 A 到 B 的路径;
- $A\cong B$,当且仅当 $A\precsim B$ 和 $B\precsim A$;
- $B\prec A$,当且仅当 $B\precsim A$,并且 $A\cong B$ 不成立.

定理 2. ① 关系$\precsim$是拟序关系;② 关系$\cong$是等价关系;③ 假设 $G/\cong$ 是由等价关系$\cong$所确定的等价类的集合(即商集),则$(G/\cong,\precsim)$是偏序集,定义$[A]\precsim[B]$,当且仅当 $A\precsim B$,其中$[C]=\{C'|C\cong C'\}$,即$[C]$表示 C 的等价类.

证明:① 要证明关系$\precsim$是拟序关系,只需证明关系$\precsim$满足自反性和传递性即可.因为对任意被定义概念 A,有 $A\equiv A$.因而在图 G 中对任意节点 A,存在一条从 A 到 A 的路径,所以有 $A\precsim A$.假设 $B\precsim A,A\precsim C$,即存在路径 $A\to\ldots\to B$ 和 $C\to\ldots\to A$.因而 $C\to\ldots\to A\to\ldots\to B$ 也是图 G 的一条路径,从而有 $B\precsim C$.

② 要证明关系$\cong$是等价关系,只需证明关系$\cong$满足自反性、对称性和传递性即可.由①可知 $A\precsim A$,从而有 $A\cong A$.假设 $A\cong B$,从而有 $A\precsim B$ 和 $B\precsim A$,即 $B\precsim A$ 和 $A\precsim B$ 成立,所以有 $B\cong A$.假设 $A\cong B,B\cong C$,则有 $A\precsim B$ 和 $B\precsim A$,以及 $B\precsim C$ 和 $C\precsim B$.由①可知关系$\precsim$满足传递性,从而由 $A\precsim B$ 和 $B\precsim C$ 可得 $A\precsim C$;由 $C\precsim B$ 和 $B\precsim A$ 可得 $C\precsim A$.即有 $A\precsim C$ 和 $C\precsim A$ 成立,从而有 $A\cong C$.

③ 要证明$(G/\cong,\precsim)$是偏序集,只需证明关系$\precsim$满足自反性、反对称性和传递性即可.因为 $A\precsim A$,所以有$[A]\precsim[A]$.假设$[A]\precsim[B]$和$[B]\precsim[A]$,所以对任意 $C\in[A]$、$D\in[B]$有 $C\precsim D$ 和 $D\precsim C$,即 $C\equiv D$,从而有$[A]=[B]$.假设$[A]\precsim[B]$和$[B]\precsim[C]$,从而对任意 $E\in[A],F\in[B],G\in[C]$,有 $E\precsim F$ 和 $F\precsim G$ 成立.由①可知 $E\precsim G$,从而有$[A]\precsim[C]$. □

(3) 在图 G 中判断是否存在等价类.如果没有,则此时的 TBox T_1 就是正规化的;否则,即 G 中存在等价类,则对任意一个等价类,进行下列转化.不妨假设该等价类包含被定义概念 $A_1,\ldots,A_h$,由等价类的定义可知:

$A_1 \equiv P_{11} \sqcap \ldots \sqcap P_{w1} \sqcap GP_{11} \sqcap \ldots \sqcap GP_{u1} \sqcap LP_{11} \sqcap \ldots \sqcap LP_{v1} \sqcap A_2 \sqcap \ldots \sqcap A_h \sqcap \exists R_{11}.B_{11} \sqcap \ldots \sqcap \exists R_{z1}.B_{z1}$;

…

$A_i \equiv P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap A_1 \sqcap \ldots \sqcap A_{i-1} \sqcap A_{i+1} \sqcap \ldots \sqcap A_h \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi}$;

…

$A_h \equiv P_{1h} \sqcap \ldots \sqcap P_{wh} \sqcap GP_{1h} \sqcap \ldots \sqcap GP_{uh} \sqcap LP_{1h} \sqcap \ldots \sqcap LP_{vh} \sqcap A_1 \sqcap \ldots \sqcap A_{h-1} \sqcap \exists R_{1h}.B_{1h} \sqcap \ldots \sqcap \exists R_{zh}.B_{zh}$,

则在最大不动点语义下被定义 $A_1,\ldots,A_h$ 都可以转化为下列统一的概念定义式:

$A_i \equiv P_{11} \sqcap \ldots \sqcap P_{w1} \sqcap \ldots \sqcap P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap \ldots \sqcap P_{1h} \sqcap \ldots \sqcap P_{wh} \sqcap GP_{11} \sqcap \ldots \sqcap GP_{u1} \sqcap \ldots \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap \ldots \sqcap GP_{1h} \sqcap \ldots \sqcap GP_{uh} \sqcap LP_{11} \sqcap \ldots \sqcap LP_{v1} \sqcap \ldots \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap \ldots \sqcap LP_{1h} \sqcap \ldots \sqcap LP_{vh} \sqcap \exists R_{11}.B_{11} \sqcap \ldots \sqcap \exists R_{z1}.B_{z1} \sqcap \ldots \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi} \sqcap \ldots \sqcap \exists R_{1h}.B_{1h} \sqcap \ldots \sqcap \exists R_{zh}.B_{zh}$.

在最小不动点语义下被定义 $A_1,\ldots,A_h$ 都不需要概念定义式.

例 3:上述例 2 的 T_1 中,$A_2 \equiv P_2 \sqcap (\leq k_2\ R_4) \sqcap A_3 \sqcap \exists R_3.B_2$,$A_3 \equiv P_3 \sqcap (\geq k_3\ R_5) \sqcap A_2 \sqcap \exists R_2.B_3$,从而 A_2 和 A_3 等价,因此在最大不动点语义下可以将 A_2 和 A_3 转化为下列正规形式:

$A_2 \equiv P_2 \sqcap P_3 \sqcap (\leq k_2\ R_4) \sqcap (\geq k_3\ R_5) \sqcap \exists R_3.B_2 \sqcap \exists R_2.B_3$;

$A_3 \equiv P_2 \sqcap P_3 \sqcap (\leq k_2\ R_4) \sqcap (\geq k_3\ R_5) \sqcap \exists R_3.B_2 \sqcap \exists R_2.B_3$.

下面证明经过上述 3 步,能够将描述逻辑εLN 的 TBox T 转化为一个等价的正规化的 TBox T'.

定理 3. 给定描述逻辑εLN 的 TBox T,T 能转化为一个等价的正规化的 TBox T'.

证明:由于第(1)步是通过引入新的被定义概念对 TBox T 中的概念定义式进行等价替换,显然得到的 TBox T_1 与 TBox T 等价.第(2)步得到图 G,定理 2 已经证明关系≅ 是等价关系,从而可以得到等价类.下面证明第(3)步能将 TBox T_1 转化为一个等价的正规化的 TBox T'.

如果 G 中存在等价类$\{A_1,\ldots,A_h\}$,其中 $A_1,\ldots,A_h$ 是被定义概念.由等价类的定义可知,对任意 $A_i,A_j \in \{A_1,\ldots,A_h\}$,$A_i \neq A_j$,有 $A_i \cong A_j$,即 $A_i \lesssim A_j$ 和 $A_j \lesssim A_i$,也就是说,在图 G 中存在一条从 A_i 到 A_j 的路径和一条从 A_j 到 A_i 的路径.因此,$A_1,\ldots,A_h$ 的定义式有如下形式:

$A_1 \equiv P_{11} \sqcap \ldots \sqcap P_{w1} \sqcap GP_{11} \sqcap \ldots \sqcap GP_{u1} \sqcap LP_{11} \sqcap \ldots \sqcap LP_{v1} \sqcap A_2 \sqcap \ldots \sqcap A_h \sqcap \exists R_{11}.B_{11} \sqcap \ldots \sqcap \exists R_{z1}.B_{z1}$;

…

$A_i \equiv P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap A_1 \sqcap \ldots \sqcap A_{i-1} \sqcap A_{i+1} \sqcap \ldots \sqcap A_h \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi}$;

…

$A_h \equiv P_{1h} \sqcap \ldots \sqcap P_{wh} \sqcap GP_{1h} \sqcap \ldots \sqcap GP_{uh} \sqcap LP_{1h} \sqcap \ldots \sqcap LP_{vh} \sqcap A_1 \sqcap \ldots \sqcap A_{h-1} \sqcap \exists R_{1h}.B_{1h} \sqcap \ldots \sqcap \exists R_{zh}.B_{zh}$.

即对被定义概念 A_1,被定义概念 $A_2,\ldots,A_h$ 出现在 A_1 的右边,从而有 $A_1 \sqsubseteq A_2,\ldots,A_1 \sqsubseteq A_h$.对被定义概念 $A_2,\ldots,A_h$ 也有类似的结果.因此可得:$A_1 \equiv \ldots \equiv A_h$,从而 $A_1,\ldots,A_h$ 具有相同的概念定义式.

因为 A_i 的定义式是:$A_i \equiv P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap A_1 \sqcap \ldots \sqcap A_{i-1} \sqcap A_{i+1} \sqcap \ldots \sqcap A_h \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi}$,所以有 $A_i \sqsubseteq P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi}$.从而对被定义概念 A_i,可以等价地转化为

$A_i \sqsubseteq P_{11} \sqcap \ldots \sqcap P_{w1} \sqcap \ldots \sqcap P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap \ldots \sqcap P_{1h} \sqcap \ldots \sqcap P_{wh} \sqcap GP_{11} \sqcap \ldots \sqcap GP_{u1} \sqcap \ldots \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap \ldots \sqcap GP_{1h} \sqcap \ldots \sqcap GP_{uh} \sqcap LP_{11} \sqcap \ldots \sqcap LP_{v1} \sqcap \ldots \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap \ldots \sqcap LP_{1h} \sqcap \ldots \sqcap LP_{vh} \sqcap \exists R_{11}.B_{11} \sqcap \ldots \sqcap \exists R_{z1}.B_{z1} \sqcap \ldots \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi} \sqcap \ldots \sqcap \exists R_{1h}.B_{1h} \sqcap \ldots \sqcap \exists R_{zh}.B_{zh}$.

由于最大不动点语义将被定义概念 A_i 进行最大解释,在包含约束下将包含关系改为等价关系就是对 A_i 进行最大解释,因而对被定义概念 A_i,有:

$A_i \equiv P_{11} \sqcap \ldots \sqcap P_{w1} \sqcap \ldots \sqcap P_{1i} \sqcap \ldots \sqcap P_{wi} \sqcap \ldots \sqcap P_{1h} \sqcap \ldots \sqcap P_{wh} \sqcap GP_{11} \sqcap \ldots \sqcap GP_{u1} \sqcap \ldots \sqcap GP_{1i} \sqcap \ldots \sqcap GP_{ui} \sqcap \ldots \sqcap GP_{1h} \sqcap \ldots \sqcap GP_{uh} \sqcap LP_{11} \sqcap \ldots \sqcap LP_{v1} \sqcap \ldots \sqcap LP_{1i} \sqcap \ldots \sqcap LP_{vi} \sqcap \ldots \sqcap LP_{1h} \sqcap \ldots \sqcap LP_{vh} \sqcap \exists R_{11}.B_{11} \sqcap \ldots \sqcap \exists R_{z1}.B_{z1} \sqcap \ldots \sqcap \exists R_{1i}.B_{1i} \sqcap \ldots \sqcap \exists R_{zi}.B_{zi} \sqcap \ldots \sqcap \exists R_{1h}.B_{1h} \sqcap \ldots \sqcap \exists R_{zh}.B_{zh}$.

由于最小不动点语义将被定义概念 A_i 进行最小解释,在包含约束下空集就是对 A_i 进行最小解释,因而对被定义概念 A_i 可以不需要定义式. □

定理 4. 将描述逻辑εLN 的 TBox T 转化为一个在最大不动点语义(最小不动点语义)下等价的正规化的 TBox $T'(T'')$是多项式时间复杂的.

证明:由定理 3 可知,经过下列 3 步可以将 TBox T 转化为最大不动点语义(最小不动点语义)下等价的正规化的 TBox $T'(T'')$:(1) 对 TBox T 中的概念定义式进行等价替换;(2) 将 TBox T 转化为图 G;(3) 在图 G 中判断是否存在等价类,并合并等价类中的概念.下面证明这 3 步都能在多项式时间内完成.

第(1)步采用替换的方法可以完成(对被定义概念进行替换),并且概念定义式中仅含有交构造算子(⊓),不会产生分支的情形,即不会导致指数级的扩充,因而第(1)步能在多项式时间内完成.

经过第(1)步,对任意的概念定义 $A\equiv D\in T, A\in N_{def}, D$ 具有下列形式:$P_1\sqcap\ldots\sqcap P_w\sqcap GP_1\sqcap\ldots\sqcap GP_u\sqcap LP_1\sqcap\ldots\sqcap LP_v\sqcap C_1\sqcap\ldots\sqcap C_t\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_z.B_z$,因而对每个定义式可以直接转化为图 G 的节点和边.不妨假设 T 有 q 个定义式,从而用 q 次循环可以将 TBox T 转化为图 G,即第(2)步能在多项式时间内完成.

在图 G 中判断等价类的方法与文献[9]相同,能在多项式时间内完成.至于等价类中概念的合并,不妨假设等价类中有 p 个定义式,最大不动点语义下的概念合并是 p 次字符串合并操作,而最小不动点语义下的概念合并为空,因而第(3)步能在多项式时间内完成. □

2.2 描述图

由于εLN 的 TBox T 能够等价地转化为一个正规化的 TBox,下面不妨假设εLN 的 TBox T 都是正规化的,并且 N_{def}, N_{prim} 和 N_{role} 分别表示 T 中被定义概念、原始概念和关系的集合,其中 $N_{prim}=N_{nprim}\cup N_{gprim}\cup N_{lprim}, N_{nprim}\cap N_{gprim}\cap N_{lprim}=\varnothing$, N_{nprim} 表示原始概念名的集合$\{P_1,\ldots,P_w\}$, N_{gprim} 表示最大数量约束原始概念的集合$\{\geq k_1\,R_1,\ldots,\geq k_u\,R_u\}$, N_{lprim} 表示最小数量约束原始概念的集合$\{\leq k_1\,R_1,\ldots,\leq k_v\,R_v\}$. $N_{gprim}\cup N_{lprim}$ 称为数量约束原始概念,记为 N_{rprim}.

定义 5. 描述逻辑εLN 的语法描述图 $G=(V,E,L)$,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εLN 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.

给定描述逻辑εLN 的 TBox T,T 能够转化为下列对应的一个语法描述图 G_T.

定义 6. 给定描述逻辑εLN 的 TBox T,T 对应的语法描述图 $G_T=(N_{def},E_T,L_T)$,其中:

- G_T 的节点集是 T 的被定义概念集 N_{def};
- 如果 A 是被定义概念,并且 A 的定义式为

$$A\equiv P_1\sqcap\ldots\sqcap P_w\sqcap GP_1\sqcap\ldots\sqcap GP_u\sqcap LP_1\sqcap\ldots\sqcap LP_v\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_z.B_z,$$

则 $L_T(A)=\{P_1,\ldots,P_w,GP_1,\ldots,GP_u,LP_1,\ldots,LP_v\}$;$(A,R_1,B_1),\ldots,(A,R_z,B_z)\in E_T$.

例 4:上述例 2 中的正规化 TBox T_2 对应的语法描述图如图 1 所示.

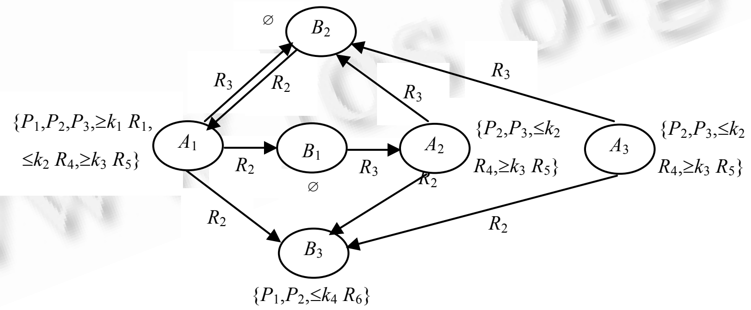

Fig.1 Syntax description graph G_T corresponding to TBox T_2 in Example 2

图 1 例 2 的 TBox T_2 对应的语法描述图 G_T

εLN 的 TBox T 与语法描述图 G_T 之间可以相互转化.

定义 7. 给定描述逻辑εLN 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 $x\in\Delta^J$.

εLN 的原始解释 J 与语义描述图 G_J 之间也可以相互转化.

例 5:如果将图 1 看成是一个语义描述图 G_J,则对应的原始解释 J 如下:

- $\Delta^J=\{A_1,A_2,A_3,B_1,B_2,B_3\}$;
- $P_1^J=\{A_1,B_3\}$,$P_2^J=\{A_1,A_2,A_3,B_3\}$,$P_3^J=\{A_1,A_2,A_3\}$,$(\geq k_1\ R_1)^J=\{A_1\}$,$(\leq k_2\ R_4)^J=\{A_1,A_2,A_3\}$,$(\geq k_3\ R_5)^J=\{A_1,A_2,A_3\}$,$(\leq k_4\ R_6)^J=\{B_3\}$;
- $R_2^J=\{(A_1,B_1),(A_1,B_3),(A_2,B_3),(A_3,B_3),(B_2,A_1)\}$;$R_3^J=\{(A_1,B_2),(A_2,B_2),(A_3,B_2),(B_1,A_2)\}$.

2.3 模拟关系

先给出原始概念集包含关系的定义.

定义 8. 给定描述逻辑εLN 的原始概念集 $AS=\{P_{11},\dots,P_{1r},\geq k_{11}\ R_{11},\dots,\geq k_{1s}\ R_{1s},\leq k_{21}\ R_{21},\dots,\leq k_{2t}\ R_{2t}\}$,$BS=\{P_{21},\dots,P_{2w},\geq k_{31}\ R_{31},\dots,\geq k_{3u}\ R_{3u},\leq k_{41}\ R_{41},\dots,\leq k_{4v}\ R_{4v}\}$,如果 $BS\subseteq AS$,当且仅当:

- $\{P_{21},\dots,P_{2w}\}\subseteq\{P_{11},\dots,P_{1r}\}$;
- $3u\leq 1s$,并且对任意$(\geq k_{3i}\ R_{3i})\in BS$,$1\leq i\leq u$,存在$(\geq k_{1j}\ R_{1j})\in AS$,$1\leq j\leq s$,使得 $R_{3i}=R_{1j}$,$k_{1j}\leq k_{3i}$;记为$\{\geq k_{31}\ R_{31},\dots,\geq k_{3u}\ R_{3u}\}\subseteq\{\geq k_{11}\ R_{11},\dots,\geq k_{1s}\ R_{1s}\}$;
- $4v\leq 2t$,并且对任意$(\leq k_{4i}\ R_{4i})\in BS$,$1\leq i\leq v$,存在$(\leq k_{2j}\ R_{2j})\in AS$,$1\leq j\leq t$,使得 $R_{4i}=R_{2j}$,$k_{2j}\leq k_{4i}$;记为$\{\leq k_{41}\ R_{41},\dots,\leq k_{4v}\ R_{4v}\}\subseteq\{\leq k_{21}\ R_{21},\dots,\leq k_{2t}\ R_{2t}\}$.

例 6:假设εLN 的原始概念集 $AS=\{P_1,P_2,P_3,\geq 2\ R_1,\geq 3\ R_2,\leq 1\ R_3,\leq 2\ R_4\}$,$BS=\{P_1,P_2,\geq 3\ R_1,\geq 5\ R_2,\leq 1\ R_3,\leq 3\ R_4\}$,因为$\{P_1,P_2\}\subseteq\{P_1,P_2,P_3\}$,$\{\geq 3\ R_1,\geq 5\ R_2\}\subseteq\{\geq 2\ R_1,\geq 3\ R_2\}$,$\{\leq 1\ R_3,\leq 3\ R_4\}\subseteq\{\leq 1\ R_3,\leq 2\ R_4\}$,所以有 $BS\subseteq AS$.

利用原始概念集的包含关系可以给出描述图之间的模拟关系.

定义 9. 给定描述逻辑εLN 的描述图(语法描述图 G_T 或语义描述图 G_J)$G_i=(V_i,E_i,L_i)(i=1,2)$,二元关系 $Z\subseteq V_1\times V_2$ 是 G_1 到 G_2 的模拟关系,当且仅当

- 如果$(v_1,v_2)\in Z$,则 $L_1(v_1)\subseteq L_2(v_2)$;
- 如果$(v_1,v_2)\in Z$,$(v_1,R,v_1')\in E_1$,则存在一个节点 $v_2'\in V_2$,使得$(v_1',v_2')\in Z$,并且$(v_2,R,v_2')\in E_2$.

Z 是 G_1 到 G_2 的模拟关系,记为 $Z:G_1\Rightarrow G_2$.

例 7:给定描述逻辑εLN 的 TBox T 如下(参考了文献[9]):

$B\equiv P_1\sqcap(\geq k_1\ R_1)\sqcap\exists R.C$,$C\equiv P_1\sqcap(\geq k_1\ R_1)\sqcap\exists R.D$,$D\equiv P_1\sqcap(\geq k_1\ R_1)\sqcap\exists R.C$,$A\equiv P_1\sqcap(\geq k_1\ R_1)\sqcap\exists R.E$,$E\equiv P_1\sqcap(\geq k_1\ R_1)\sqcap\exists R.D$.该 TBox T 对应的语法描述图 $G_T=(N_{def},E_T,L_T)$如图 2 所示.

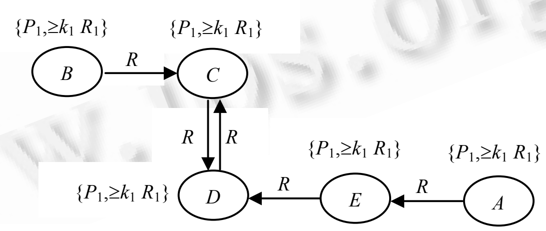

Fig.2 Syntax description graph G_T corresponding to TBox T in Example 7

图 2 例 7 的 TBox T 对应的语法描述图 G_T

其中,$N_{def}=\{A,E,B,C,D\}$,$E_T=\{(B,R,C),(C,R,D),(D,R,C),(E,R,D),(A,R,E)\}$,$L_T(B)=L_T(C)=L_T(D)=L_T(E)=L_T(A)=\{P_1,\geq k_1\ R_1\}$.

给定二元关系 $Z=\{(B,B),(C,C),(D,D),(E,E),(A,A)\}\subseteq N_{def}\times N_{def}$,则 Z 是 G_T 到 G_T 的模拟关系.

下面给出有关计算描述图之间模拟关系的时间复杂性性质.

定理 5. 给定描述逻辑εLN 的描述图 $G_1=(V_1,E_1,L_1)$和 $G_2=(V_2,E_2,L_2)$,Z_1 和 Z_2 是 G_1 到 G_2 的模拟关系,则 $Z_1\cup Z_2$ 也是 G_1 到 G_2 的模拟关系,即模拟关系在并操作($\cup$)下是封闭的.

证明:对任意的$(v_1,v_2)\in Z_1\cup Z_2$,一定有$(v_1,v_2)\in Z_1$ 或$(v_1,v_2)\in Z_2$.如果$(v_1,v_2)\in Z_1$,由 Z_1 是 G_1 到 G_2 的模拟关系可知,如果$(v_1,v_2)\in Z_1$,则 $L_1(v_1)\subseteq L_2(v_2)$.并且如果存在$(v_1,R,v_1')\in E_1$,则存在一个节点 $v_2'\in V_2$,使得$(v_1',v_2')\in Z_1$,并且$(v_2,R,v_2')\in E_2$.所以满足模拟关系的两个条件.同理可证$(v_1,v_2)\in Z_2$的情形.所以 $Z_1\cup Z_2$ 也是 G_1 到 G_2 的模拟关系. □

由定理 5 可知,如果描述图 G_1 和 G_2 是有限的,则一定存在一个从 G_1 到 G_2 的最大模拟关系,并且最大模拟关系可按下列方法获得:

令 $Z_0=\{(v_1,v_2)\in V_1\times V_2|\ L_1(v_1)\subseteq L_2(v_2)\}$,即 Z_0 是满足模拟关系定义第 1 个条件的集合.然后从 Z_0 中删除不满足模拟关系定义第 2 个条件的所有元素.

上述两步都可以在多项式时间内完成.事实上,Henzinger 已给出求两个有限描述图 G_1 和 G_2 的最大模拟关系的方法,并证明了从 G_1 到 G_2 的最大模拟是多项式时间的[14].

定理 6. 给定描述逻辑εLN 的描述图 $G_1=(V_1,E_1,L_1)$和 $G_2=(V_2,E_2,L_2)$,$v_1\in V_1$,$v_2\in V_2$,则可以在多项式时间内判断是否存在模拟关系 $Z:G_1\Rightarrow G_2$,$(v_1,v_2)\in Z$.

证明:由定理 5 可知,从 G_1 到 G_2 存在一个最大模拟关系 Z_g.要判断是否存在模拟关系 Z,使得$(v_1,v_2)\in Z$,只需判断是否有$(v_1,v_2)\in Z_g$ 即可.而构建最大模拟关系 Z_g 可以在多项式时间内完成,所以定理得证. □

与εL 的描述图[8,9]相同,εLN 的描述图之间的模拟关系还满足下列性质:

定理 7. 给定描述逻辑εLN 的描述图 $G_1=(V_1,E_1,L_1)$,$G_2=(V_2,E_2,L_2)$和 $G_3=(V_3,E_3,L_3)$,$Z_1:G_1\Rightarrow G_2$ 和 $Z_2:G_2\Rightarrow G_3$ 是模拟关系,则 $Z_1\circ Z_2=\{(v,v'')|$存在 v'使得$(v,v')\in Z_1,(v',v'')\in Z_2\}$也是模拟关系,即 $Z_1\circ Z_2:G_1\Rightarrow G_3$.

证明:如果$(v,v'')\in Z_1\circ Z_2$,则存在 v'使得$(v,v')\in Z_1$,$(v',v'')\in Z_2$.由于 Z_1 和 Z_2 是模拟关系,所以有 $L_1(v)\subseteq L_2(v')$和 $L_2(v')\subseteq L_3(v'')$,从而有 $L_1(v)\subseteq L_3(v'')$.即 $Z_1\circ Z_2$ 满足模拟关系定义的第 1 个条件.

如果$(v,v'')\in Z_1\circ Z_2$,$(v,R,v_1)\in E_1$,由$(v,v'')\in Z_1\circ Z_2$ 可知存在 v'使得$(v,v')\in Z_1$,$(v',v'')\in Z_2$.因为 Z_1 是模拟关系,所以由$(v,v')\in Z_1$ 和$(v,R,v_1)\in E_1$ 可知,存在 $v_2\in V_2$,使得$(v_1,v_2)\in Z_1$,并且$(v',R,v_2)\in E_2$.又因为 Z_2 是模拟关系,所以由$(v',v'')\in Z_2$ 和$(v',R,v_2)\in E_2$ 可知,存在 $v_3\in V_3$,使得$(v_2,v_3)\in Z_2$,并且$(v'',R,v_3)\in E_3$.所以由$(v_1,v_2)\in Z_1$ 和$(v_2,v_3)\in Z_2$ 可知,$(v_1,v_3)\in Z_1\circ Z_2$.从而证明,如果$(v,v'')\in Z_1\circ Z_2$,$(v,R,v_1)\in E_1$,则存在一个节点 $v_3\in V_3$,使得$(v_1,v_3)\in Z_1\circ Z_2$,并且$(v'',R,v_3)\in E_3$.即 $Z_1\circ Z_2$ 满足模拟关系定义的第 2 个条件.所以 $Z_1\circ Z_2:G_1\Rightarrow G_3$. □

3 εLN 循环术语集的推理

3.1 最大不动点语义下的推理

首先给出最大不动点语义下可满足性推理算法.

定理 8. 给定描述逻辑εLN 的 TBox T,原始解释 $J=(\Delta^J,\cdot^J)$,I 是 T 的基于 J 的最大不动点模型,语法描述图 $G_T=(N_{def},E_T,L_T)$,语义描述图 $G_J=(\Delta^J,E_J,L_J)$,对任意 $A\in N_{def}$,$x\in\Delta^J$,则 $x\in A^I$,当且仅当存在一个从 G_T 到 G_J 的模拟关系 $Z:G_T\Rightarrow G_J$,使得$(A,x)\in Z$.

证明:先证明⇒,假设 $x\in A^I$,二元关系 $Z\subseteq N_{def}\times\Delta^J$ 如下定义:

$$Z=\{(B,y)\in N_{def}\times\Delta^J|y\in B^I\}.$$

下面证明 Z 是从 G_T 到 G_J 的模拟关系,并且满足$(A,x)\in Z$.

因为 $x\in A^I$,所以有$(A,x)\in Z$.因而下面只需证明 Z 是从 G_T 到 G_J 的模拟关系,即证明 Z 满足定义 9 的两个条件.

给定$(B,y)\in Z$,并且不妨假设 B 具有如下正规形式:

$$P_1\sqcap\ldots\sqcap P_w\sqcap GP_1\sqcap\ldots\sqcap GP_u\sqcap LP_1\sqcap\ldots\sqcap LP_v\sqcap\exists R_1.B_1\sqcap\ldots\sqcap\exists R_z.B_z,$$

其中 $w,u,v,z\ge 0$,$P_1,\ldots,P_w\in N_{nprim}$,$GP_1,\ldots,GP_u\in N_{gprim}$,$LP_1,\ldots,LP_v\in N_{lprim}$,$R_1,\ldots,R_z\in N_{role}$,$B_1,\ldots,B_z\in N_{def}$.

由语法描述图的定义可知 $L_T(B)=\{P_1,\ldots,P_w,GP_1,\ldots,GP_u,LP_1,\ldots,LP_v\}$.

由语义描述图的定义可知 $L_J(y)=\{P\in N_{prim}|\ y\in P^J\}$.又因为$(B,y)\in Z$,所以 $y\in B^I$,从而有 $y\in P_i^I=P_i^J$,$y\in GP_j^I=GP_j^J$,

$y\in LP_k^I=LP_k^J$,其中 $1\le i\le w,1\le j\le u,1\le k\le v$.因此有:$\{P_1,\dots,P_w,GP_1,\dots,GP_u,LP_1,\dots,LP_v\}\subseteq L_J(y)$,从而有 $L_T(B)\subseteq L_J(y)$,即 Z 满足模拟关系定义的第 1 个条件.

假设$(B,R_i,B_i)\in E_T$,因为 $B\equiv P_1\sqcap\dots\sqcap P_w\sqcap GP_1\sqcap\dots\sqcap GP_u\sqcap LP_1\sqcap\dots\sqcap LP_v\sqcap\exists R_1.B_1\sqcap\dots\sqcap\exists R_z.B_z$,所以 $B^I\subseteq(\exists R_i.B_i)^I$,其中 $1\le i\le z$.由于 $y\in B^I$,所以 $y\in(\exists R_i.B_i)^I$.由$\exists R_i.B_i$的语义解释可知:存在 $y_i\in\Delta^J$,使得$(y,y_i)\in R_i^J$和 $y_i\in B_i^I$成立.从而有$(y,R_i,y_i)\in E_J$和$(B_i,y_i)\in Z$,即 Z 满足模拟关系定义的第 2 个条件.

再证明$\Leftarrow$,假设 $Z:G_T\Rightarrow G_J$ 是从 G_T 到 G_J 的模拟关系,并且$(A,x)\in Z$.因为 I 是 T 的基于 J 的最大不动点模型,所以由定理 1 和定义 3 可知,存在一个序数α,使得 $I=I^{\downarrow\alpha}$.下面用超穷归纳法[13]证明:给定元组(B,y,β),如果$(B,y)\in Z$,则 $y\in B^{I^{\downarrow\beta}}$,其中 $B\in N_{def},y\in\Delta^J$,$\beta$是一个序数.

用反证法,假设$(B,y)\in Z$,并且 $y\notin B^{I^{\downarrow\beta}}$.因为$\beta$是一个序数,所以有两种情况:$\beta$是一个极限序数,或$\beta$是一个后继序数.

如果β是一个极限序数,则由定义 3 可知,$B^{I^{\downarrow\beta}}=B^{\mathrm{glb}(\{I^{\downarrow\gamma}|\gamma<\beta\})}=\bigcap_{\gamma<\beta}B^{I^{\downarrow\gamma}}$.因为$(B,y)\in Z,y\notin B^{I^{\downarrow\beta}}$,所以存在一个序数$\gamma<\beta$,使得$(B,y)\in Z,y\notin B^{I^{\downarrow\gamma}}$.根据超穷归纳法假设,对任意序数$\gamma<\beta$,有 $y\in B^{I^{\downarrow\gamma}}$,与 $y\notin B^{I^{\downarrow\gamma}}$ 矛盾.

如果β是一个后继序数,即$\beta=\gamma+1$.不妨假设 B 具有如下正规形式:

$$P_1\sqcap\dots\sqcap P_w\sqcap GP_1\sqcap\dots\sqcap GP_u\sqcap LP_1\sqcap\dots\sqcap LP_v\sqcap\exists R_1.B_1\sqcap\dots\sqcap\exists R_z.B_z,$$

其中 $w,u,v,z\ge 0,P_1,\dots,P_w\in N_{nprim},GP_1,\dots,GP_u\in N_{gprim},LP_1,\dots,LP_v\in N_{lprim},R_1,\dots,R_z\in N_{role},B_1,\dots,B_z\in N_{def}$.

由定义 3 可知:$B^{I^{\downarrow\beta}}=O_{T,J}(B^{I^{\downarrow\gamma}})=(P_1\sqcap\dots\sqcap P_w)^{I^{\downarrow\gamma}}\cap(GP_1\sqcap\dots\sqcap GP_u)^{I^{\downarrow\gamma}}\cap(LP_1\sqcap\dots\sqcap LP_v)^{I^{\downarrow\gamma}}\cap(\exists R_1.B_1\sqcap\dots\sqcap\exists R_z.B_z)^{I^{\downarrow\gamma}}$.

由语法描述图的定义可知,$L_T(B)=\{P_1,\dots,P_w,GP_1,\dots,GP_u,LP_1,\dots,LP_v\}$.

由语义描述图的定义可知,$L_J(y)=\{P\in N_{prim}|\ y\in P^J\}$.又因为$(B,y)\in Z,Z$ 是从 G_T 到 G_J 的模拟关系,从而有 $L_T(B)\subseteq L_J(y)$,即$\{P_1,\dots,P_w,GP_1,\dots,GP_u,LP_1,\dots,LP_v\}\subseteq\{P\in N_{prim}|\ y\in P^J\}$.从而有 $y\in P_i^J,y\in GP_j^J,y\in LP_k^J$,其中 $1\le i\le w,1\le j\le u,1\le k\le v$.

又因为 $y\notin B^{I^{\downarrow\beta}}$,由 $y\in P_i^J,y\in GP_j^J,y\in LP_k^J$,其中 $1\le i\le w,1\le j\le u,1\le k\le v$,可知:存在 $j,1\le j\le z$,使得 $y\notin(\exists R_j.B_j)^{I^{\downarrow\gamma}}$.

又因为$(B,y)\in Z$,并且由语法描述图的定义可知$(B,R_j,B_j)\in E_T$,由于 $Z:G_T\Rightarrow G_J$ 是从 G_T 到 G_J 的模拟关系,所以存在 $y_j\in\Delta^J$,使得$(y,R_j,y_j)\in E_J,(B_j,y_j)\in Z$.从而有$(y,y_j)\in R_j^J,y_j\in B_j^{I^{\downarrow\gamma}}$.所以 $y\in(\exists R_j.B_j)^{I^{\downarrow\gamma}}$,与 $y\notin(\exists R_j.B_j)^{I^{\downarrow\gamma}}$ 矛盾.

因此,证明了命题:给定元组(B,y,β),如果$(B,y)\in Z$,则 $y\in B^{I^{\downarrow\beta}}$,其中 $B\in N_{def},y\in\Delta^J$,$\beta$是一个序数.

所以对于元组(A,x,α),因为$(A,x)\in Z$,从而有 $x\in A^{I^{\downarrow\alpha}}$.又因为 $I=I^{\downarrow\alpha}$,所以 $x\in A^I$.所以定理得证. □

定理 9. 给定描述逻辑εLN 的 TBox T,原始解释 $J=(\Delta^J,\cdot^J)$,I 是 T 的基于 J 的最大不动点模型,语法描述图 $G_T=(N_{def},E_T,L_T)$,语义描述图 $G_J=(\Delta^J,E_J,L_J)$,对任意 $A\in N_{def}$,则 A 在最大不动点语义下是可满足的,当且仅当存在 $x\in\Delta^J$,以及存在从 G_T 到 G_J 的模拟关系 $Z:G_T\Rightarrow G_J$,使得$(A,x)\in Z$.

证明:因为 A 在最大不动点语义下是可满足的,当且仅当存在 $x\in\Delta^J$,使得 $x\in A^I$.由定理 8 可知,$x\in A^I$,当且仅当存在从 G_T 到 G_J 的模拟关系 $Z:G_T\Rightarrow G_J$,使得$(A,x)\in Z$. □

下面给出最大不动点语义下包含关系推理算法.

定理 10. 给定描述逻辑εLN 的 TBox T,A,B 是 T 中的被定义概念,即 $A,B\in N_{def}$,则 $A\sqsubseteq_{gfp,T}B$,当且仅当存在一个从 G_T 到 G_T 的模拟关系 $Z:G_T\Rightarrow G_T$,使得$(B,A)\in Z$.

证明:先证明$\Leftarrow$,给定原始解释 J,I 是基于 J 的最大不动点模型,要证明 $A\sqsubseteq_{gfp,T}B$,只需要证明对任意 $x\in\Delta^J$,如果 $x\in A^I$,则 $x\in B^I$即可.

如果 $x\in A^I$,由定理 8 可知,存在一个从 G_T 到 G_J 的模拟关系 $Y:G_T\Rightarrow G_J$,使得$(A,x)\in Y$.又因为 $Z:G_T\Rightarrow G_T$ 是从 G_T 到 G_T 的模拟关系,并且$(B,A)\in Z$,所以由定理 7 可知,$X=Z\circ Y$ 是从 G_T 到 G_J 的模拟关系,并且$(B,x)\in X$.又由定理 8 可知,$x\in B^I$.

再证明$\Rightarrow$,假设 $A\sqsubseteq_{gfp,T}B$,不妨将语法描述图 G_T 看成是语义描述图,从而由 G_T 可以构造一个原始解释 J,并且 $G_T=G_J$.假设 I 是基于 J 的最大不动点模型.

下面证明恒等关系 $Id\subseteq N_{def}\times N_{def}$,即 $Id(A)=A$,是从 G_T 到 G_T 的模拟关系 $Id:G_T\Rightarrow G_T$.

如果$(A,A)\in Id$,因为有 $L_T(A)=L_T(A)$,所以 $L_T(A)\subseteq L_T(A)$,即 Id 满足模拟关系定义的第 1 个条件.

对任意节点 $B\in N_{def}$,有$(B,B)\in Id$.所以如果$(A,A)\in Id,(A,R,B)\in E_T$,则有$(B,B)\in Id$,并且$(A,R,B)\in E_T$.即 Id 满足模拟关系定义的第 2 个条件.所以恒等关系 $Id\subseteq N_{def}\times N_{def}$ 是从 G_T 到 G_T 的模拟关系.

由于 $Id:G_T\Rightarrow G_T=G_J$ 是模拟关系,所以$(A,A)\in Id$,由定理 8 可知 $A\in A^I$.又因为 $A\sqsubseteq_{gfp,T}B$,所以 $A\in B^I$.又由定理 8 可知,存在一个从 G_T 到 G_T 的模拟关系 $Z:G_T\Rightarrow G_J=G_T$,使得$(B,A)\in Z$. □

下面给出最大不动点语义下可满足性推理和包含关系推理算法的时间复杂性性质.

定理 11. 最大不动点语义下描述逻辑εLN 的可满足性推理和包含关系推理都能在多项式时间内完成.

证明:由定理 9 和定理 10 可知,最大不动点语义下εLN 的可满足性推理和包含关系推理都等价于描述图之间模拟关系的构造.从而由定理 6 可知,εLN 的可满足性推理和包含关系推理都能在多项式时间内完成. □

3.2 最小不动点语义下的推理

给定描述逻辑εLN 的 TBox T,语法描述图 $G_T=(N_{def},E_T,L_T)$,$A,B\in N_{def}$,并引入下列标记:

- $A\xrightarrow{*}_T B$ 表示在 G_T 中存在一条从 A 到 B 的路径;
- $A\xrightarrow{+}_T B$ 表示在 G_T 中存在一条从 A 到 B 的非空路径;
- $Cyc_T=\{A|$ 在 G_T 中存在节点 B,使得 $A\xrightarrow{*}_T B\xrightarrow{+}_T B\}$.

首先给出最小不动点语义下可满足性推理算法.

定理 12. 给定描述逻辑εLN 的 TBox T,语法描述图 $G_T=(N_{def},E_T,L_T)$,如果 $A\in Cyc_T$,则存在一个被定义概念 $A'\in Cyc_T$ 和关系 R,使得$(A,R,A')\in E_T$.

证明:因为 $A\in Cyc_T$,所以在 G_T 中存在节点 B,使得 $A\xrightarrow{*}_T B\xrightarrow{+}_T B$,从而在 TBox T 中一定存在下列形式的概念定义式:

$A\equiv\ldots\sqcap\exists R_1.B_1\sqcap\ldots;$

$B_1\equiv\ldots\sqcap\exists R_2.B_2\sqcap\ldots;$

…

$B_i\equiv\ldots\sqcap\exists R_{i+1}.B\sqcap\ldots;$

$B\equiv\ldots\sqcap\exists R_{i+2}.B_{11}\sqcap\ldots;$

$B_{11}\equiv\ldots\sqcap\exists R_{i+3}.B_{12}\sqcap\ldots;$

…

$B_{1j}\equiv\ldots\sqcap\exists R_{i+j+2}.B\sqcap\ldots$.其中 $i>0,j>0$,

则令 $A'=B_1,R=R_1$.从而有 $A'\xrightarrow{*}_T B\xrightarrow{+}_T B$,即 $A'\in Cyc_T$,并且$(A,R,A')\in E_T$. □

定理 13. 给定描述逻辑εLN 的 TBox T,原始解释 $J=(\Delta^J,\cdot^J)$,I 是 T 的基于 J 的最小不动点模型,$A\in N_{def}$,如果 $A\in Cyc_T$,则在最小不动点语义下 A 是不可满足的,即 $A^I=\varnothing$.

证明:因为 I 是 T 的基于 J 的最小不动点模型,所以由定理 1 和定义 3 可知,$I=I^{\uparrow\omega}$,并且 $A^I=\bigcup_{n\ge 0}A^{I^{\uparrow n}}$.下面用归纳法证明对任意 $n\ge 0$, $A^{I^{\uparrow n}}=\varnothing$.

当 $k=0$ 时,由定义 3 直接可得 $A^{I^{\uparrow 0}}=\varnothing$.

假设 $k=n$ 时有 $A^{I^{\uparrow n}}=\phi$.下面证明 $k=n+1$ 时有 $A^{I^{\uparrow n+1}}=\varnothing$.

因为 $A\in Cyc_T$,由定理 12 可知,存在一个被定义概念 $B\in Cyc_T$ 和关系 R,使得$(A,R,B)\in E_T$.不妨假设 A 的定义式是 $A\equiv D$,由$(A,R,B)\in E_T$ 可知,D 一定具有下列形式:

$$D\equiv\ldots\sqcap\exists R.B\sqcap\ldots$$

由归纳假设可知 $B^{I^{\uparrow n}}=\varnothing$,从而 $D^{I^{\uparrow n}}=\varnothing$,因此 $A^{I^{\uparrow n+1}}=O_{T,J}(A^{I^{\uparrow n}})=D^{I^{\uparrow n}}=\varnothing$.

所以对任意 $n\ge 0$,有 $A^{I^{\uparrow n}}=\varnothing$.由于 $A^I=\bigcup_{n\ge 0}A^{I^{\uparrow n}}$,所以 $A^I=\varnothing$. □

定理 14. 给定描述逻辑εLN 的 TBox T,原始解释 $J=(\Delta^J,\cdot^J)$,I 是 T 的基于 J 的最小不动点模型,$A\in N_{def}$,如

果在最小不动点语义下 A 是可满足的,即 $A^I \neq \varnothing$,则 $A \notin Cyc_T$.

证明:反证法.假设 $A \in Cyc_T$,由定理 13 可知,在最小不动点语义下 A 是不可满足的,即 $A^I = \varnothing$,与 $A^I \neq \varnothing$ 矛盾. □

由定理 13 可以看出,最小不动点语义下 Cyc_T 中所有被定义概念的定义式都可以从 TBox T 中去除.

下面给出最小不动点语义下包含关系推理算法.

定理 15. 给定描述逻辑εLN 的 TBox T,Cyc_T 中的所有被定义概念的定义式从 T 中去除后得到 TBox T',A,B 是 T 中的被定义概念,则 $A \sqsubseteq_{lfp,T} B$,当且仅当

- $A,B \notin Cyc_T$,$A \sqsubseteq_{gfp,T'} B$;
- $A \in Cyc_T$.

证明:对 T 中任意被定义概念 A,B,一共有 3 种情况:(1) $A \notin Cyc_T, B \notin Cyc_T$;(2) $A \in Cyc_T$;(3) $A \notin Cyc_T, B \in Cyc_T$.

(1) 如果 $A \notin Cyc_T, B \notin Cyc_T$.因为 T' 是从 T 中去除 Cyc_T 中的所有被定义概念的定义式后得到的,而由 Cyc_T 的定义可知,Cyc_T 包含 T 中的所有循环定义,因此,T' 不包含循环定义,即 T' 是一个无环 TBox.又由定理 13 可知,如果 $A \in Cyc_T$,则在最小不动点语义下 A 是不可满足的,即对任意最小不动点模型 I,有 $A^I = \varnothing$.因此 $A \sqsubseteq_{lfp,T} B$,当且仅当 $A \sqsubseteq_{lfp,T'} B$.Nebel 已证明:在没有循环定义的条件下,最大不动点模型和最小不动点模型是相等的[6].因此,$A \sqsubseteq_{lfp,T'} B$,当且仅当 $A \sqsubseteq_{gfp,T'} B$.所以有:$A \sqsubseteq_{lfp,T} B$,当且仅当 $A \sqsubseteq_{gfp,T'} B$.

(2) 如果 $A \in Cyc_T$.由定理 13 可知,在最小不动点语义下 A 是不可满足的,即 $A^I = \varnothing$.因此,对任意 B,$B \in Cyc_T$ 或 $B \notin Cyc_T$,都有 $A \sqsubseteq_{lfp,T} B$.

(3) 如果 $A \notin Cyc_T, B \in Cyc_T$.下面证明 $A \sqsubseteq_{lfp,T} B$ 不成立,即 $A \not\sqsubseteq_{lfp,T} B$.因为 $B \in Cyc_T$,由定理 13 可知,在最小不动点语义下 B 是不可满足的,即对任意最小不动点模型 I,有 $B^I = \varnothing$.因而要证明 $A \not\sqsubseteq_{lfp,T} B$,只需证明存在一个最小不动点模型 I',使得 $A^{I'} \neq \varnothing$ 即可.

由定理 13 可知,如果 $A \in Cyc_T$,则在最小不动点语义下 A 是不可满足的,即对任意最小不动点模型 I,有 $A^I = \varnothing$.因此,如果将 Cyc_T 中的被定义概念解释为空集$\varnothing$,则 T' 的最小不动点模型可以扩充为 T 的最小不动点模型.又因为 T' 的最大不动点模型和最小不动点模型相同,因此,由 G_T 可以构造一个原始解释 J.假设 I' 是基于 J 的最大不动点模型,因为 T' 不包含循环定义,所以 I' 也是基于 J 的最小不动点模型.由于恒等关系 Id 是模拟关系,即有 $Id: G_T \Rightarrow G_T$.因为 $(A,A) \in Id$,由定理 8 可知 $A \in A^{I'}$,即 $A^{I'} \neq \varnothing$. □

可以看出,最小不动点语义下εLN 循环术语集的可满足性和包含关系推理机制与εL[8,9]相同.

定理 16. 最小不动点语义下描述逻辑εLN 的可满足性推理和包含关系推理都能在多项式时间内完成.

证明:由定理 13 和定理 14 可知,最小不动点语义下εLN 的可满足性推理等价于 Cyc_T 的构造,而 Cyc_T 是语法描述图 G_T 中的循环路径的集合,在图论中已经有构造有向图中循环路径集合的多项式时间复杂的算法[15],所以最小不动点语义下εLN 的可满足性推理能在多项式时间内完成.

由定理 15 可知,最小不动点语义下εLN 的包含关系推理等价于 Cyc_T 的构造和最大不动点语义下的包含关系推理.由定理 11 可知,最大不动点语义下的包含关系推理可以在多项式时间内完成,又因为 Cyc_T 的构造也可以在多项式时间内完成,所以最小不动点语义下εLN 的包含关系推理能在多项式时间内完成. □

4 结束语

在描述逻辑εL 的基础上添加数量约束构造算子,提出了描述逻辑εLN,给出了εLN 的语法和语义.针对εLN 的需要,重新定义了描述图.使用描述图之间的模拟关系给出了不动点语义下εLN 循环术语集的可满足性和包含关系推理算法,并证明了推理算法是多项式时间复杂的.进一步的工作主要有:研究描述语义下描述逻辑εLN 循环术语集的可满足性和包含关系推理算法;以及扩充εLN 循环术语集(可以添加并构造算子).这应该是一个很有难度、但相当有意义的课题.

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, 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.

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

[5] 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.

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

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

[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] Baader F. Terminological cycles in a description logic with existential restrictions. LTCS-Report 02-02, Dresden, Dresden University of Technology, 2002. 1−34.

[10] Jiang YC, Shi ZZ, Tang Y, Wang J. Fuzzy description logic for semantics representation of the semantic Web. Journal of Software, 2007,18(6):1257−1269 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/18/1257.htm

[11] 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

[12] Tarski A. A lattice-theoretical fixpoint theorem and its applications. Pacific Journal of Mathematics, 1955,5(2):285−309.

[13] Zhang JW. Intoduction to Axiomatic Set Theory. Beijing: Science Press, 1999. 28−61 (in Chinese).

[14] Henzinger MR, Henzinger TA, Kopke PW. Computing simulations on finite and infinite graphs. In: Prabhakar R, ed. Proc. of the 36th Annual Symposium on Foundations of Computer Science. New York: IEEE Computer Society Press, 1995. 453−462.

[15] Yang H. Selections from Graph Theory Common Algorithms. Beijing: China Railway Publishing House, 1988. 15−21(in Chinese).

附中文参考文献:

[10] 蒋运承,史忠植,汤庸,王驹.面向语义 Web 语义表示的模糊描述逻辑.软件学报,2007,18(6):1257−1269. http://www.jos.org.cn/1000-9825/18/1257.htm

[11] 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

[13] 张锦文.公理集合论导引.北京:科学出版社,1999.28−61.

[15] 杨洪.图论常用算法选编.北京:中国铁道出版社,1988.15−21.

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.

史忠植(1941－),男,研究员,博士生导师,CCF 高级会员,主要研究领域为人工智能,机器学习,分布式人工智能.

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.

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.