计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200300188

循环描述逻辑系统 $\mathcal{FL}_0$ 最大不动点模型的有穷基

郑添健[1] 侯金宏[2] 张 维[3] 王 驹[4]
1 黔南民族师范学院计算机与信息学院 贵州 都匀 558000
2 南宁上林县第一中学 南宁 530500
3 黔南民族师范学院数学与统计学院 贵州 都匀 558000
4 广西师范大学计算机科学与信息工程学院 广西 桂林 541004

摘 要 描述逻辑和形式概念格都是以概念为基础的形式系统，它们都有各自的优缺点，近年来人们开始将两者结合起来，取长补短。文中将形式概念分析的方法引入描述逻辑，研究了循环描述逻辑系统 $\mathcal{FL}_0$ 在最大不动点语义下有限模型的有穷基。在形式概念分析中，属性有限的形式背景都存在 Duguenne-Guigues 基，文章将循环 $\mathcal{FL}_0$ 的最大不动点模型作为描述背景，将 $\mathcal{FL}_0$ 概念作为属性，定义了描述背景下的属性蕴含，证明了循环描述逻辑系统 $\mathcal{FL}_0$ 的最大不动点模型也存在有穷基底。

关键词：描述逻辑；形式概念分析；属性探测；有穷基
中图法分类号 TP181

Finite Basis of Implicational System Associated with Finite Models of Description Logic $\mathcal{FL}_0$ Under the Greatest Fixed Point Semantics

ZHENG Tian-jian[1], HOU Jin-hong[2], ZHANG Wei[3] and WANG Ju[4]
1 School of Computer and Information, Qiannan Normal University for Nationalities, Duyun, Guizhou 558000, China
2 The NO. 1 High Middle School of Shanglin Coutry, Nanning 530500, China
3 School of Mathematics and Statistics, Qiannan Normal University for Nationalities, Duyun, Guizhou 558000, China
4 School of Computer Science and Information Engineering, Guangxi Normal University, Guilin, Guangxi 541004, China

Abstract Description logic and formal concept analysis are two different formalisms based on concept, each has its own advantages and disadvantages. Researchers begin to combine them together recently. In this paper, methods of formal concept analysis are introduced into research in description logic to analyze the finite basis of finite models of $\mathcal{FL}_0$ under greatest fixed point semantics. In formal concept analysis, there always exists the Duguenne-Guigues basis as long as the attribute set is finite. The finite model of cyclic $\mathcal{FL}_0$ terminology under greatest fixed point semantics is taken as description context, $\mathcal{FL}_0$ concept as attribute and implications are defined in this context. It is proved that there exists a finite basis of the finite model, which is also sound and completed.

Keywords Description logic, Formal concept analysis, Attribute exploration algorithm, Finite basis

1 引言

描述逻辑是一种面向对象的形式化工具，是一阶逻辑的可判定性子集[1]，与其他形式化工具相比，其具有推理机制，能实现知识之间的自动推理。因此，它成为近年来人工智能研究的热点。在过去的几十年里，研究者们研究了描述逻辑的推理机制问题，得到了一系列推理算法并在实际应用中运行良好，如描述逻辑系统 KRIS（1991）、FaCT（1998）等。随着研究的深入，其应用也越来越广泛，在语义 Web[2-3]、本体（ontology）[4]、数据库[5-6]和软件工程[7]等领域都有很好的应用。当然描述逻辑最杰出的应用还是在语义 Web 方面，由于其良好的形式化语义和可靠、完备的推理算法，使它成为了Web 本体描述语言的逻辑基础，目前通用的本体语言 OWL 都是以描述逻辑为基础的。

形式概念分析是 20 世纪 80 年代初由德国 Wille 教授提出的，作为一种形式化方法，它为数据分析提供了有效的工具。形式概念分析是一种成熟的数学模型，其特点是小而全，能非常清晰地处理信息。给定一个特定的形式背景，它能自动地搜寻出所有的形式概念构成一个形式概念格，层次非常清晰。因此，形式概念分析被应用到了许多领域，如 Neuss，Kent 等[8]将形式概念分析应用到数字图书馆和文献检索方面，Deogun，Raghavan，Sever 等将其用于数据挖掘及知识发现方面，得到了一些新的挖掘算法。

描述逻辑和形式概念分析都是以概念为基础，但在概念

基金项目：贵州省教育厅自然科学基金项目（黔教科 20090047）；贵州省科学技术基金资助项目（黔科合 J 字[2012]2310）；黔南民族师范学院一般项目（qnsy2018023）
This work was supported by the Natural Science Foundation of Guizhou Education Department(20090047), Science and Technology Foundation of Guizhou Province([2012]2310) and Natural Science Foundation of Qiannan Normal University for Nationalities(qnsy2018023).
通信作者：郑添健（ztjpsc@163.com）

的获取上却有着显著的不同。在描述逻辑中，概念是由原子概念和原子角色符号在构造子的帮助之下递归形成，概念都解释成域的子集，但是概念的形成与解释域没有关系，描述语言含有角色符号，能描述域中个体之间的关系。在形式概念分析中，概念都与一个特定的形式背景相关，不同的形式背景形成不同的概念。所谓的形式背景，其实就是一个三元组 $K=\langle G,M,I\rangle$，其中 G 是个体（对象）的集合，M 是属性的集合，I 是 $G\times M$ 上的关系，即 $I\subseteq G\times M$。最简单的形式背景就是一个二维表，它的行代表属性，列代表个体也称为对象。形式概念分析中的概念就是一个二元组 (A,B)，A 是 G 中具有属性集 B 中所有属性的个体的集合，称作概念的外延；B 是 A 中所有个体的所有共同属性构成的集合，称作概念的内涵。正是由于这些差异，使得它们各有优缺点。描述逻辑具有强的表达能力和推理机制，能描述域中个体之间的关系，但是却很难全部找出一个给定域上的所有概念，形式概念分析能给出一个域上的所有概念，但却受限于表达力，而且概念依赖于具体的背景，不具有通用性。近年来研究者们开始尝试在它们之间取长补短，发挥其各自的优势。具体如下：用描述语言来定义形式背景中的属性，利用描述语言丰富的构造子来弥补形式背景中属性表达力弱的缺点，如 Prediger 和 Stumme 在 1999 年将描述逻辑引进概念信息系统（Conceptual Information System），其数据分析工具为形式概念分析（FCA），利用描述语言来弥补属性表达力弱的不足[8]；2007 年 Stumme 等将描述逻辑与形式概念分析结合形成一种称作关系概念分析的形式系统[9]。另一方面，将形式概念分析的方法引入描述逻辑的研究中，来处理一些构建描述逻辑知识库的问题，如 Baader 在 1995 年利用形式概念分析中的属性探测算法给出了描述逻辑系统中一个概念组关于交运算的层次计算[10]；1996 年 Stumme 扩充了 Baader 的结果，并给出了描述逻辑系统中一组概念关于交和并运算的层次计算[11]；2000 年 Baader 等使用属性探测算法给出了描述逻辑系统中一个概念组的最小公共包含概念的层次计算[12]；2006 年 Rudolph 将描述逻辑结合形式概念分析提出了一种新的方法——关系探索（Relation Exploration），并将此方法应用到描述逻辑系统 $\mathcal{FLE}$ 中，得到了 $\mathcal{FLE}$ 的一个完备知识模型[13]；2008 年 Baader 利用形式概念分析中的属性探测算法给出了描述逻辑系统 εL 有限模型的有穷基[14]；2010 年 Tang 等利用形式概念分析中的属性探测算法给出了循环描述逻辑系统 $\mathcal{FLE}$ 有限模型的有穷基，为构建一个完备非冗余的描述逻辑知识库提供了理论基础[15]。本文在前人的基础上研究了循环描述逻辑系统 $\mathcal{FL}_0$ 最大不动点语义下有限模型的有穷基问题，解决了 Baader 在文献[14]中提出的 $\mathcal{FL}_0$ 有限模型的有穷基存在性问题，具其有理论和现实意义。

2 预备知识

定义 1[1] 描述语言 $\mathcal{FL}_0$ 含构造子交 $\sqcap$，全称量词 $\forall$，$\mathcal{FL}_0$ 中概念递归地定义如下：$C,D\rightarrow\bot|\top|P|A\sqcap B|\forall r.A$，其中 P 是原子概念，r 是角色。

定义 2[1] 二元组 $I=(\Delta,\cdot^I)$ 是一个模型，其中 Δ 为解释域，是一个非空集合，$\cdot^I$ 是解释函数，它将概念解释成 Δ 的子集，将角色解释成 Δ 上的关系，具体如下：

(1) $(\bot)^I=\emptyset$；

(2) $(\top)^I=\Delta$；

(3) $(P)^I=P^I\subseteq\Delta$；

(4) $(r)^I=r^I\subseteq\Delta\times\Delta$；

(5) $(C\sqcap D)^I=C^I\cap D^I$；

(6) $(\forall r.C)^I=\{d\in\Delta_i\mid\forall e(d,e)\in r^I\rightarrow e\in C^I\}$。

定理 1[1] $C\sqsubseteq D$ 当且仅当对于所有的解释 I，有 $C^I\subseteq D^I$，$C\equiv D$ 当且仅当 $C\sqsubseteq D$ 且 $D\sqsubseteq D$。

$\mathcal{FL}_0$ 术语集公理（$Tbox$）就是有限条形如 $C\equiv D$ 或 $C\sqsubseteq D$ 的式子，前者称为包含公理，后者称为等价公理。若 $Tbox$ 中任意概念都不直接或间接地引用自己，则称它是非循环的，否则称之为循环的 $Tbox$。例如：$Momo=Man\sqcap\forall haschild.Momo$，$Momo$ 同时出现在等式两端，就是一个循环定义。

Nebel 在文献[16]中指出，不含循环定义的描述逻辑系统的解释（模型）是唯一确定的，而循环的情况不唯一。目前公认的是最大不动点语义、最小不动点语义和描述语义 3 种。本文只考虑最大不动点语义。

定义 3[1] 模型 I 被称为一个循环 $Tbox$ T 的最大不动点模型当且仅当存在一个原始解释 J，$I\in\mathrm{int}(J)$ 并且 I 是 $O_{T,J}$ 的最大不动点。

不是每一个循环的描述逻辑 $Tbox$ 都有最大不动点模型，$Tarski$ 的不动点定理告诉我们：完备格上的单调映射一定存在不动点。文献[1]指出，若描述语言中不含构造子非（$\neg$），$O_{T,J}$ 在 $Int(J)$ 上就是单调的。由定义 1 可知，描述语言 $\mathcal{FL}_0$ 不含非，因此循环的 $\mathcal{FL}_0$-$Tbox$ 一定具有不动点模型。Baader 在文献[17]中还指出循环的 $\mathcal{FL}_0$-$Tbox$ 同时具有最大不动点模型和最小不动点模型。

定义 4[17] A,B 是循环 $\mathcal{FL}_0$-$Tbox$ 中的两个概念，$A\sqsubseteq_{gfp,T}B$ 当且仅当对于任意的最大不动点模型 I 都有 $A^I\subseteq B^I$ 成立。

Baader[17]还指出一个循环的 $\mathcal{FL}_0$-$Tbox$ 可以表示成一部自动机，其中概念表示成自动机的状态，初始概念是可接受状态，角色符号作为自动机的接受语言的字母。在最大不动点语义解释下概念 A,B 之间的包含关系可以用自动机接受语言间的包含关系来刻画。

定理 2[17] T 是一个循环的 $\mathcal{FL}_0$-$Tbox$，A 是 T 的概念，A_T 是 T 对应的自动机表示，I 是 T 的最大不动点模型，$\forall d\in I$，$d\in A^I$ 当且仅当 $\forall e\in dom(I)$，若 $(d,e)\in W^I$，则有 $e\in P^I$，对任意的 P 和任意的 $W\in L(A,P)$ 成立。

定理 3[17] T 是一个循环 $\mathcal{FL}_0$-$Tbox$，A_T 是 T 对应的自动机表示，A,B 是 T 中的两个概念，$A\sqsubseteq_{gfp,T}B$ 当且仅当 $L(B,P)\subseteq L(A,P)$ 对任意的 P 成立。

定义 5[18] 形式背景 (G,M,I) 是由对象集 G、属性集 M 以及关系 I 组成。$(g,m)\in I$ 或 $g\mathrm{I}m$ 表示对象 g 具有属性 m。在 $X\subseteq G$ 和 $Y\subseteq M$ 上定义一对对偶算子：

$$X'=\{m\in M\mid\forall g\in X,g\mathrm{I}m\}$$

$$Y'=\{g\in G\mid\forall m\in Y,g\mathrm{I}m\}$$

X' 表示 X 中所有对象共同属性的集合，Y' 表示具有 Y 中所有属性的对象构成的集合。

定义 6[18] (G,M,I) 为一个形式背景，如果二元组 (X,Y) 满足 $X'=Y$，且 $Y'=X$，则称 (X,Y) 是 (G,M,I) 的一个形式概念，简称概念，其中 X 称为概念的外延，Y 称为概念的内涵。

定理 4[18] (G,M,I)是一个形式背景，$X,X_1,X_2\subseteq G,Y,Y_1,Y_2\subseteq M$，则有：

(1)$X_1\subseteq X_2\Rightarrow X_2{}'\subseteq X_1{}'$；

(1′)$Y_1\subseteq Y_2\Rightarrow Y_2{}'\subseteq Y_1{}'$；

(2)$X\subseteq X''$；(2′)$Y\subseteq Y''$；

(3)$X'=X'''$；(3′)$Y'=Y'''$；

(4)$X\subseteq Y'\Leftrightarrow Y\subseteq X'\Leftrightarrow X\times Y\subseteq I$。

定义 7[18] 蕴含式 $Y_1\rightarrow Y_2$ 在(G,M,I)中成立，当且仅当$\forall o\in G,o\in Y_1'$，则有 $o\in Y_2'$，即 $Y_1{}'\subseteq Y_2{}'$。$Y_1\rightarrow Y_2$ 在(G,M,I)中不成立，当且仅当$\exists o\in G,o\in Y_1{}'\backslash Y_2{}'$。

定义 8[18] 设 $T\subseteq M$，若 $Y_1\nsubseteq T$ 或 $Y_2\subseteq T(Y_1\subseteq T\rightarrow Y_2\subseteq T)$，则 T 与 $Y_1\rightarrow Y_2$ 相关。$\mathscr{L}$是一个蕴含式的集合，若 T 与 $\mathscr{L}$ 中每一个属性蕴含都相关，则称 T 与 $\mathscr{L}$相关。

由定义 7 可知，$Y_1\rightarrow Y_2$ 在(G,M,I)中成立当且仅当 $Y_1{}'\subseteq Y_2{}'$，即 $Y_2\subseteq Y_1{}''$成立。蕴含式集合 $\mathscr{L}$中的每一个蕴含都成立，称 $\mathscr{L}$在(G,M,I)中是可靠的。对任意属性集 Y，由定理 4 可知 $Y\subseteq Y''$，则蕴含 $Y\rightarrow Y''$在(G,M,I)中恒成立，因此 $\mathscr{L}:=\{Y\rightarrow Y''|Y$ 是任意属性集$\}$是可靠的。记 $\mathrm{Im}\ p(K)$为(G,M,I)中成立的蕴含式构成的集合，显然有 $\mathscr{L}\subseteq \mathrm{Im}\ p(K)$，$\mathrm{Im}\ p(K)$在很多时候是一个非常大的集合，而真正让人们感兴趣的是它的生成集。

定义 9[18] $\mathscr{L}$是一个蕴含式集合，若 M 的每一个与 $\mathscr{L}$相关的子集都与 $Y_1\rightarrow Y_2$ 相关，则称 $Y_1\rightarrow Y_2$ 可由 $\mathscr{L}$生成。若蕴含式集合 $\mathscr{L}_1$ 中的每一个蕴含都可以由 $\mathscr{L}$生成，则称 $\mathscr{L}$是 $\mathscr{L}_1$ 的生成集。若 $\mathrm{Im}\ p(K)$可以由集合 $\mathscr{L}$生成，则称 $\mathscr{L}$是完备的。

定义 10[18] 若(G,M,I)的蕴含式集合 $\mathscr{L}$ 中的每一个蕴含都不能由其真子集生成，则称 $\mathscr{L}$是非冗余的。

定义 11[18] 属性集 Y 被称为形式背景(G,M,I)的伪内涵，当且仅当 $Y\neq Y''$且其任意的真子集 Q 都有 Q''真包含于 Y''。

定理 5[18] 形式背景(G,M,I)中的蕴含式集合 $\mathscr{L}:=\{Y\rightarrow Y''|Y$ 是伪内涵$\}$是可靠的、完备的和非冗余的。

定义 12[18] 蕴含式集合 $\mathscr{L}$被称为形式背景(G,M,I)的 Duguenne-Guigues 基当且仅当它是可靠的、完备的和非冗余的。

文献[18]指出，任给一个属性集有限的形式背景，都可以通过属性探测算法找出它的 Duguenne-Guigues 基。

3 描述逻辑系统 $\mathscr{FL}_0$ 的有穷基

定义 13[14] 描述背景是一个二元组$\langle L,\Delta^i|i\in I\rangle$，$L$ 是 $\mathscr{FL}_0$ 所有概念的集合，$i=(\Delta^i,\cdot^i)$是 L 的一个解释(模型)，其中 Δ^i 是解释域，它是一个非空的有限集。L 相当于形式背景中的属性集，Δ^i 相当于对象集。

在给定的描述背景$\langle L,\Delta^i\rangle$下，可以类似地定义两个对偶算子，对任意的 $f\subseteq L,X\subseteq\Delta^i$：

(1)$f^i=\{x\in\Delta^i|\forall f_k\in f,x\in f^i\}$；

(2)$X^i=\{f\in L|\forall x\in X,x\in f^i\}$。

任意的概念集 $f_k\in f\subseteq L$，在解释 $i=(\Delta^i,\cdot^i)$下都有 $f_k^i\subseteq\Delta^i$，因此 $f^i\subseteq\Delta^i$。反过来，对于给定的 $X\subseteq\Delta^i$，却不一定存在概念集 f，使得 $X=f^i$。由于描述语言对概念交封闭，在描述背景中所有的概念集 f 都可以用一个概念$\sqcap f=\sqcap f_i$，$f_i\in f$ 表示。为了陈述的方便，给出了几个要用到的定义和定理：

定义 14[19] E 是描述逻辑系统 L 中的概念，I 是它的一个模型，$x\in dom(I)$，E 被称为是 x 基于模型 I 的最具体概念当且仅当满足条件：(1)$x\in E^I$；(2)$\forall F\in L$，若 $x\in F^I$，则 $E\sqsubseteq F$。

E 是 x 基于模型 I 的最具体概念就是满足条件(1)最小的概念。Distel 在文献[19]中指出，非循环的 $\mathscr{FL}_0$，$\varepsilon\mathscr{L}$ 系统不一定都存在基于模型的最具体概念，而在有循环定义并且基于最大不动点语义的情况下是存在的，并且有如下定理。

定理 6[19] I 是循环 $\mathscr{FL}_0$-$Tbox$ T 的最大不动点模型，A_I 是模型 I 的自动机表示，A_T 是 T 的自动机表示，对任意的 $x\in dom(I)$都存在基于模型 I 的最具体概念 E，并且有 $L(E,P)=L(x,P)$，对任意的初始概念 P 成立。

定义 15[1] L 是一个描述逻辑系统，A,B 是 L 中的概念，E 被称为 A,B 的最小公共包含当且仅当满足条件：1)$A\sqsubseteq E,B\sqsubseteq E$；2)$\forall F\in L$，若 $A\sqsubseteq F,B\sqsubseteq$;，则 $E\sqsubseteq F$。此时记 $E=lcs(A,B)$。

Baader 在文献[20]中指出，循环的 $\mathscr{FLN}$ 系统存在最小公共包含概念，是 $\mathscr{FLN}$ 的子系统，因此在循环的 $\mathscr{FL}_0$ 中概念的最小公共包含也是存在的。

在基于模型最具体概念和最小公共包含概念都存在的情况下，可以定义：$X^i=lcs\{C_x\in L|\forall x\in X,C_x$ 是 x 基于模型 i 的最具体概念$\}$，这样就完成了形式背景到描述背景的转换。

定理 7[14] 描述背景$\langle L,\Delta^i\rangle$，$X,Y\subseteq\Delta^i$，概念 $C,D\in L$，则有：

(1)$X\subseteq Y\Rightarrow X^i\sqsubseteq Y^i$；

(2)$C\Rightarrow C^i\subseteq D^i$；

(3)$X\subseteq X^{ii}$；

(4)$C^{ii}\sqsubseteq C$；

(5)$X\equiv X^{iii}$；

(6)$C\equiv C^{iii}$；

(7)$X\subseteq C^i\Rightarrow X^i\sqsubseteq C$。

定义 16[14] 在描述背景$\langle L,\Delta^i\rangle$中，$f_1,f_2\in L$，$f_1\rightarrow f_2$ 成立当且仅当 $f_1^i\subseteq f_2^i$ 成立，此时称 $f_1\rightarrow f_2$ 在$\langle L,\Delta^i\rangle$中有效，蕴含式集合 A 是有效的当且仅当 A 中的每一个蕴含式是有效的。

由定理 1 可知，对任意的模型 I，若 $A^I\subseteq B^I$，则 $A\sqsubseteq B$，因此 $A\rightarrow B$ 是有效的。反之则不然，由定义 7 可知，$A\rightarrow B$ 在$\langle L,\Delta^i\rangle$中有效，只需 $A^i\subseteq B^i$ 即可，并不代表在任意模型中都成立。

定义 17[14] 蕴含式 $f_1\rightarrow f_2$ 是可以由蕴含式集合 A 导出当且仅当对任意的背景$\langle L,\Delta^i\rangle$，若 A 中的每一个蕴含式在$\langle L,\Delta^i\rangle$中有效，则 $f_1\rightarrow f_2$ 也在$\langle L,\Delta^i\rangle$中有效。若蕴含集 B 的每一个蕴含式都可以由 A 导出，则称 A 导出 B。

定义 18[14] 蕴含式集合 A 被称为背景$\langle L,\Delta^i\rangle$的一个有穷基，当且仅当 A 是有穷的并且满足以下性质：

(1)可靠性：A 在$\langle L,\Delta^i\rangle$中是有效的；

(2)完备性：$\langle L,\Delta^i\rangle$中所有有效的蕴含式都可以由 A 导出；

(3)非冗余性：A 的任意真子集都不具备完备性。

定理 8[14] 若 $f_1\rightarrow f_2$ 在$\langle L,\Delta^i\rangle$中有效，则它可以由集合$\{f_1\rightarrow f_1^{ii}\}$导出，由形式概念分析的知识可知，$f_1\rightarrow f_1^{ii}$ 在任意背景中都是有效的，这样只要掌握了集合$\{f\rightarrow f^{ii}|f\in L\}$，也就掌握了$\langle L,\Delta^i\rangle$中所有有效的蕴含式。但是 L 中的概念有

无穷多个,人们希望能找到它的一个有限子集,使得它能被这个子集导出。

定义 19[14] $D\subseteq L$ 被称为背景 $\langle L,\Delta^i\rangle$ 的一个支配集当且仅当对于任意的 $f\in L$ 存在 $\bar{f}\in D$,使得 $f\sqsubseteq\bar{f}$,$f^i=\bar{f}^i$ 成立。

定理 9[14] $D\subseteq L$ 是背景 $\langle L,\Delta^i\rangle$ 的支配集,则蕴含式集合 $A=\{f\rightarrow f^{ii}\mid f\in D\}$ 是 $\langle L,\Delta^i\rangle$ 的可靠完备集。

由定理 9 可知,只要找出循环 $\mathcal{FL}_0$ 描述背景的一个有限支配集,就找到了一个兼具有限性、可靠性和完备性的蕴含式集合,然后在此基础上去找它的一个非冗余子集,也就找到了该背景下的一个有穷基。

对于任意的循环 $\mathcal{FL}_0-TboxT$,可以构造一个非循环的 T' 来逼近它。如:

T:

$B=P\sqcap\forall r.B$,

T':

$B_0=P\sqcap\forall r.\top\equiv P$;

$B_1=P\sqcap\forall r.B_0\equiv P\sqcap\forall r.P$;

$B_2=P\sqcap\forall r.B_1\equiv P\sqcap\forall r.P\sqcap\forall r.(\forall r.P)$;

$\vdots$

$B_{k+1}=P\sqcap\forall r.B_k$。

T' 中所有的被定义概念形成一条"$\sqsubseteq$-降链",即:$B_0\sqsupseteq B_1\sqsupseteq B_2\sqsupseteq\cdots\sqsupseteq B_k\sqsupseteq B_{k+1}$ 成立,并且 T' 中任意的 B_k 都有 $B\sqsubseteq B_k$。

定理 10 对于任意的 $\mathcal{FL}_0$ 循环定义 B,都可以在非循环的 $\mathcal{FL}_0$ 中找到一族概念 $\{B_k\mid k\in N\}$,并且使得 $B\sqsubseteq B_k$。

证明:对于循环 $\mathcal{FL}_0-TboxT$ 中的概念 B,可以通过上述示例的过程找到一组概念 $B_0,B_1,\cdots,B_k$ 构成 T',将 T 表示成一部自动机 A_T,记它的接受语言集合为 $L(B,P)$。由定理 2 可知,$\forall d\in I,\forall P$ 若 $d\in B^I$ 则有对每一个的 $W_k\in L(B,P)$,$\forall e\in I:(d,e)\in(W_k)^I$ 都有 $e\in P^I$ 成立,由定义 1 可知 $d\in(\forall W_k.P)^I$,当 $W_k=\varepsilon$ 时,$d\in P^I$。由 T' 的构造过程可以知道每一个 B_k 都可以表示成某些 $\forall W_k.P$ 的交,因此,$\forall d\in B^I$,则有 $d\in(B_k)^I$,从而有 $B\sqsupseteq B_k$。证毕。

定理 11 对于任意的 $\mathcal{FL}_0$ 循环定义 B,都可以在非循环的 $\mathcal{FL}_0$ 中找到概念 U,使得 $B^i=U^i$,i 是最大不动点模型。

证明:由定理 10 可知,对任意的循环定义 B 都存在一族非循环的概念 $\{B_k\mid k\in N\}$ 使得 $B\sqsubseteq B_k$,并且 $\{B_k\mid k\in N\}$ 构成一族"$\sqsubseteq$-降链"。对于给定的模型 i,由于 Δ^i 有限,则 $\wp(\Delta^i)$ 也有限,记 $\wp(\Delta^i)=\{D_1,D_2,\cdots,D_m\}$,其中 $D_i\subseteq\Delta^i$。由于 T' 是非循环的,因此,每一个 B_k 在模型 i 下的解释 B_k^i 是 Δ^i 的一个确定的子集,即 $B_k^i\in\wp(\Delta^i)=\{D_1,D_2,\cdots,D_m\}$,因此,可以将 T' 中所有的概念分成 m 类,记作 $[f_1],\cdots,[f_m]$,$\forall B_k\in[f_h]$ 当且仅当 $B_k^i=D_h\subseteq\Delta^i$,$1\leqslant h\leqslant m$。很容易验证这是一个等价分类,那么任意的 $B_k\in T'$ 一定唯一属于 $[f_1],\cdots,[f_m]$ 中的一类。由于循环 $\mathcal{FL}_0$-$Tbox$ T 一定具有最大不动点模型 i,则有 $B^i=D_l\in\wp(\Delta^i)$。因此,在 T' 中一定存在 U,使得 $U^i=B^i=D_l$。证毕。

蕴含式集合 $\{B\rightarrow B^{ii}\}$ 在任何背景下都是可靠的和完备的,但却不一定有限。记 $\mathcal{O}=\{B^i\mid B$ 是 $\mathcal{FL}_0$ 的循环定义,i 是 gfp 模型$\}$,$\mathcal{K}=\{U\mid U\in\mathcal{C},U^i=X,\forall X\in\mathcal{O}\}$,由于 Δ^i 有限,从而集合 $\mathcal{O},\mathcal{K}$ 都是有限的。定义蕴含式集合 $\Omega=\{P\rightarrow P^{ii}\mid P\in N_{Prim}\cup\{\top\}\}\cup\{(\forall r.U)\rightarrow(\forall r.U)^{ii}\mid r\in N_r,U\in\mathcal{K}\}\cup\{(U_1\sqcap U_2\rightarrow(U_1\sqcap U_2)^{ii}\mid U_1,U_2\in\mathcal{K}\}$,很显然 Ω 是一个有限集,并且 $\{B\rightarrow B^{ii}\}$ 可以由 Ω 导出。

定理 12 蕴含式集合 $\{B\rightarrow B^{ii}\}$ 可以由 Ω 导出。

证明:每一个概念 B 都是 $\mathcal{FL}_0$ 中的概念,施归纳于公式的形成过程:

(1)当 $B\in N_{Prim}\cup\{\top\}$ 时,$P\rightarrow P^{ii}\in\Omega$,显然 $P\rightarrow P^{ii}$ 可以由 Ω 导出。

(2)当 $B=\forall r.V$ 时,由归纳可知 $V\rightarrow V^{ii}$ 是可以由 Ω 导出的。令 j 是使得 Ω 成立的模型,则有 $V^j\subseteq(V^{ii})^j$。由定义 2 可知,$B^j=(\forall r.V)^j=\{x\in\Delta^j\mid\forall y(x,y)\in r^j\rightarrow y\in V^j\}$,则 $B^j\subseteq\{x\in\Delta^j\mid\forall y(x,y)\in r^j\rightarrow y\in(V^{ii})^j\}$。取 $U\in K$,$U^i=V^i$,由定理 7 可知 $V^{ii}\sqsubseteq U$,则 $B^j\subseteq\{x\in\Delta^j\mid\forall y(x,y)\in r^j\rightarrow y\in U^j\}=(\forall r.U)^j$。由 Ω 的构成可知 $(\forall r.U)\rightarrow(\forall r.U)^{ii}\in\Omega$,由假设可知 $\forall r.U\rightarrow(\forall r.U)^{ii}$ 在 j 中成立,则 $(\forall r.U)^j\subseteq((\forall r.U)^{ii})^j$,则有 $B^j\subseteq(\forall r.U)^j\subseteq((\forall r.U)^{ii})^j$,而 $((\forall r.U)^{ii})^j=(\{x\mid\forall y(x,y)\in r^i\rightarrow y\in U^i\}^i)^j=(\{x\mid\forall y(x,y)\in r^i\rightarrow y\in V^i\}^i)^j=((\forall r.V)^{ii})^j=(B^{ii})^j$,则有 $B^j\subseteq(B^{ii})^j$,由定义 16 和定义 17 可知 $B\rightarrow B^{ii}$ 可以由 Ω 导出。

(3)当 $B=V_1\sqcap V_2$ 时,由归纳可知 $V_1\rightarrow V_1^{ii}$,$V_2\rightarrow V_2^{ii}$ 可由 Ω 导出,对任意使得 Ω 有效的模型 j,有 $(V_1)^j\subseteq(V\alpha\ ii_1)^j$,$(V_2)^j\subseteq(V_2^{ii})^j$。$B^j=V_1^j\cap V_2^j\subseteq(V_1^{ii})^j\cap(V_2^{ii})^j$,令 $U_1^i=V_1^i$,$U_2^i=V_2^i$,则 $V_1^{ii}\sqsupseteq U_1$,$V\ ii_2\sqsupseteq U_2$,则 $(V_1^{ii})^j\subseteq U_1^j$,$(V_2^{ii})^j\subseteq U_2^j$,则 $B^j\subseteq U_1^j\cap U_2^j$。又由于 $(U_1\sqcap U_2)\rightarrow(U_1\sqcap U_2)ii\in\Omega$,则 $B^j\subseteq C_1^j\cap C_2^j\subseteq((C_1\sqcap C_2)^{ii})^j=((C_1^i\cap C_2^i)^i)^j=((V_1^i\cap V_2^i)^i)^j=((V_1\sqcap V_2)^{ii})^j=(B^{ii})^j$。由定义 16 和定义 17 可知 $B\rightarrow B^{ii}$ 可以由 Ω 导出。

综合(1)-(3),$\{B\rightarrow B^{ii}\}$ 可以由 Ω 导出。证毕。

由定理 12 可知,Ω 是一个有限的、可靠的和完备的蕴含公式集,则可以由 Ω 得到循环 $\mathcal{FL}_0$ 描述背景的一个有穷基。

定理 13 任意的循环 $\mathcal{FL}_0$ 描述背景 i 都存在一个有穷的、可靠的和完备的蕴含式集合作为其有穷基。

证明:由定理 12 可知,Ω 的一个极小真子集就是背景 i 的有穷基 Ω^*,具体步骤如下:1)令 $\Omega^*=\Omega$;2)对任意的蕴含式 $U\rightarrow V\in\Omega^*$,若 $U\rightarrow V$ 能被 $\Omega^*/\{U\rightarrow V\}$ 导出,则令 $\Omega^*=\Omega^*/\{U\rightarrow V\}$;否则,$\Omega^*=\Omega^*$。由于 Ω 有限,在有限步内一定可以得到一个 Ω 的一个极小真子集 Ω^*。证毕。

结束语 本文将形式概念分析的方法引入描述,逻辑研究了循环描述逻辑系统 $\mathcal{FL}_0$ 在基于最大不动点语义下有限模型的有穷基问题,证明有限的 $\mathcal{FL}_0$ 的 gfp 模型存在有穷基,这为构建描述逻辑知识库提供了强有力的理论依据和技术支持。在实际应用当中,人们更关注如何设计有效的算法来计算一个有限的 $\mathcal{FL}_0$ 系统的 gfp 模型的有穷基,可以从有穷基出发,在一个具体域上构建一个本体。本文的后续工作就是设计计算有穷基的算法。

参 考 文 献

[1] BAADER F,NUTT W. Basic Description Logics[M]//The Description Logic Handbook:Theory,Implementation and Applications. Cambridge:Cambridge University Press,2003:47-100.

[2] SHI Z Z,DONG M K,JIANG Y C,et al. Logical basis of Semantic Web [J]. Chinese Science Series E: Information Science,

2004,34(10):1123-1138.

[3] JIANG Y C, TANG Y, WANG J, et al. Description logic for Semantic Web [J]. Pattern recognition and artificial intelligence, 2007, 20(1):48-54.

[4] FBAADER, HORROCKS I, SATTLER U. Description Logics [M]// Handbook on Ontologies: International Handbooks on Information Systems. Springer, 2004:3-28.

[5] ABORGIDA, LENZERINI M, RROSATI M. Description Logics for Data Bases[M]// The Description Logic Handbook: Theory, Implementation and Applications. Cambridge: Cambridge University Press, 2003:472-494.

[6] JIANG Y C, TANG Y, WANG J, et al. ER model with attribute dependent timing based on description logic [J]. Computer Research and Development, 2007, 44(10):1765-1773.

[7] DBERARDI, DCALVANESE, GIACOMO D. Reasoning on UML Class Diagrams[J]. Artificial Intelligence, 2005, 168(1/2): 70-118.

[8] PREDIGER S, STUMME G. Theory-driven logical scaling: Conceptual information systems meet description logics[C]// Proceeding of the 6th International Workshop on Knowledge Representationmeets Databases (KRDB'99). 1999:14-21.

[9] ROUANE M H, HUCHARD, NAPOLI A, et al. A Proposal for Combining Formal Concept Analysis and DescriptionLogics for Mining Relational Data[C]// Proceedings of the 5th International Conference on Formal Concept Analysis (ICFCA 2007). Springer-Verlag, 2007:14-22.

[10] BAADER F. Computing a Minimal Representation of the Subsumption Lattice of all Conjunctions of Concepts Defined in a Terminology[C]// Knowledge Retrieval, Use and Storage for Efficiency: Proceedings of the 1st International KRUSE Symposium, 1955:168-178.

[11] STUMME G. The Concept Classification of a Terminology Extended by Conjunction and Disjunction[C]// Proceedings of the 4th Pacific Rim International Conference on Artificial Intelligence (PRICAI'96). Springer-Verlag, 1996:121-131.

[12] BAADER F, MOLITOR R. Building and Structuring Description Logic Knowledge Bases Using Least Common Subsumers and Concept Analysis [C] // Proceedings of the 8th International Conference on Conceptual Structures (ICCS 2000). Springer-Verlag, 2000:292-305.

[13] RUDOLPH S. Relational exploration: Combining Description Logics and Formal Concept Analysis for knowledge specification [D]. dissertation, Fakultät Mathematik und Naturwissenschaften, TU Dresden, Germany, 2006.

[14] BAADER F. A finite basis for the set of EL-implications holding in a finite model[C]// Proceedings of the 6th International Conference on Formal Concept Analysis (ICFCA 2008). Berlin: Springer, 2008:46-61.

[15] TANG S Q, CAI Z X, WANG J, et al. The finite basis of the description logic system $\mathscr{FL}\varepsilon$ based on the maximum fixed point model [J]. Computer Research and Development, 2010 (9): 1514-1521.

[16] NEBEL B. Terminological cycles: Semantics and computational properties [M]. SanFrancisco: Morgan Kaufmann Publishers, 1991:331362.

[17] BAADER F. Using automata theory for characterizing the semantics of terminological cycles [C] // Annals of Mathematics and ArtificialIntelligence. 1996:175-219.

[18] GANTER B, WILLE R, FRANZKE C. Formal concept analysis: mathematical foundations [M]. Springer-Verlag New York, Inc. , 1997.

[19] FELIX D. Model-based Most Specific Concepts in Description Logics with Value Restrictions [R]. Institute for Theoretical Computer Science, Dresden, Germany, 2008.

[20] KÜSTERS R. Characterizing the semantics of terminological cycles in ALN using finite automata[C]// Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning(KR'98). 1998.

ZHENG Tian-jian, born in 1976, associate professor, is a member of China Computer Federation. His main research interests include computer network, database theory and mathematical logic.