

一种带物元的动态描述逻辑

王 静 刘 群 石 磊

(哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

摘 要 针对动态描述逻辑框架中只有概念和关系,在表述由于动作作用而引起的概念或个体的属性及值的变化和变化后的影响方面能力不强的问题,本文引入物元的概念及其发散规则扩充动态描述逻辑,给出了一种新的带物元的动态描述逻辑(MDDL)。文中按照传统描述逻辑的语义解释方法给出了物元的语义解释,然后引入物元及物元“一物多征”的发散推理规则,扩充动态描述逻辑的 Tableau 算法,生成一种新的 Tableau-M 算法。最后根据该算法深入研究了 MDDL 的基本推理问题,即实例断言集的一致性检测问题和概念与物元的可满足性检测问题。

关键词 动态描述逻辑,物元,发散规则,可扩展性

A Dynamic Description Logic with Matter Element

WANG Jing LIU Qun SHI Lei

(School of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

Abstract Dynamic description logic framework has weakness in describing the meanings and changes of concepts and its effects after changing because there are only concepts and relations in dynamic description logic framework. To solve this problem, the matter element and its divergence rules were injected to dynamic description logic, a new dynamic description logic with matter element (MDDL) was proposed. The paper gives the matter element's semantic explanation according to traditional description logic's. Then matter element and its divergence rules with a matter with multi-feature were injected to extend the Tableau algorithm in dynamic description logic, and a new Tableau-M algorithm was generated. At last, according to this algorithm, the paper deeply researches the basic reasoning problem of MDDL, which is the problem checks consistency of instance assertion set and the satisfiabilities of concept and matter element.

Keywords Dynamic description logic, Matter element, Divergence rules, Extensibility

描述逻辑(description logic, DL)^[1]是基于实体(对象)的知识表示的形式化,它吸收了 KL-ONE 的主要思想,是一阶谓词逻辑的一个可判定子集。描述逻辑的重要特征是具有很强的表达能力和可判定性,它保证推理过程总能停止,并返回确定的结果。除了知识表示以外,描述逻辑还用在其它许多领域,它被认为是以往表示工具的逻辑重构和统一形式化,所以在众多知识表示的形式化方法中,它越来越受到人们的特别关注^[2]。

但是,传统描述逻辑也有不足的地方,即传统描述逻辑框架中只有概念和关系,当需要描述某个概念或者个体具有的属性时就不很方便,只能将它转换为概念的包含断言或关系断言的形式来描述。因为事物是由于属性的不同而形成了各种不同的类,从而产生了不同的概念。判断一个个体是否属于某个概念,就要看它是否满足了概念的本质属性,而个体与个体之间存在什么样的关系也是通过属性来决定的。当人们要了解某个事物,就表现在对该事物各个属性的了解,所以知识的表示方法还应该具有描述概念属性的能力与手段。

另外,传统的描述逻辑只能描述静态知识,不能有效地对动态知识进行表示和推理。针对这个问题,中科院史忠植老师对传统描述逻辑进行了拓展,提出了动态描述逻辑(DDL)^[3]。它将静态和动态的知识表示与推理有机地整合在一起,形成了一种统一的形式化框架,从而弥补了传统描述逻辑在这方面的不足。但是,该描述逻辑在表示概念的属性与属性值时仍然与传统描述逻辑相同,没有明确、清晰的表示方法。而且,动作对某个个体作用时,可能仅仅是对该个体的属性或属性值发生了作用,而不是个体的所有属性。那么在考虑由此动作引起的领域状态的变化时,只需考虑与发生变化

的属性和属性值相关联的个体或者个体之间的关系所发生的变化,无需考虑与未发生变化的属性相关联的其它个体以及关系。所以,在动态描述逻辑中,属性的描述也是很重要的。

基于以上原因,本文对动态描述逻辑进行了扩充,引入了“物元”^[4-6]的概念,来统一描述概念及概念具有的各个属性,给出了一种新的描述逻辑,即一种带物元的动态描述逻辑(MDDL)。

“物元”是一种形式化表示事物的方法。为了使人们处理问题的时候既要考虑量,又要考虑质,物元将事物、特征和量值集于一个统一体中,并根据“一物多征”的特点,实现了知识拓展,称之为物元的发散性,这一点适合于人们解决问题时的联想思维方式。由此可知,物元也是描述事物可变性的基本工具。根据以上物元的特点,在动态描述逻辑中引入物元,不仅可以实现概念及其属性的统一,还可以利用物元的发散性弥补传统描述逻辑在处理知识不完全或存在隐性知识时的不足。

1 描述逻辑基础

描述逻辑^[1]是建立在概念(concept)和关系(role)之上,其中概念解释为对象的集合,关系解释为对象之间的二元关系。一般地,描述逻辑依据提供的构造算子,在简单的概念和关系之上构造复杂的概念和关系。通常描述逻辑至少包含以下构造算子:交(\sqcap),并(\sqcup),非(\neg),存在量词(\exists),全称量词(\forall)。例如,假设 Person 和 Female 是原子概念,hasChild 是一个原子关系,那么利用构造算子构造 Mother 和 MotherWithoutDaughter 概念为: $Mother \equiv (Person \sqcap Female) \sqcap \exists hasChild$, $Person$; $MotherWithoutDaughter \equiv Mother \sqcap \forall hasChild$, $(\neg(Person \sqcap Female))$ 。

一个描述逻辑系统包含四个组成部分:1)表示概念和关系的构造集;2)Tbox 包含断言;3)Abox 实例断言;4)Tbox 和 Abox 上的推理机制。一个描述逻辑系统的表示能力和推理能力取决于对以上几个要素的选择以及不同的假设。

一个描述逻辑知识库 $K = \langle T, A \rangle$ 由两个部分组成: Tbox T 和 Abox A。其中 Tbox 是一个关于包含断言的有限集合,具有如下形式: $C \sqsubseteq D$, 这里 C 和 D 为概念。通常用 $C \equiv D$ 作为 $C \sqsubseteq D$ 和 $D \sqsubseteq C$ 的缩写。Abox 是实例断言的有限集合,形为 $C(a)$, 其中 C 是一个概念, a 是一个个体(或者变元)的标识;或者形为 $P(a, b)$, 其中 P 为一个原子关系, a, b 为两个个体(或者变元)的标识。

动态描述逻辑(DDL)^[3]是对传统描述逻辑 ALC 的扩充,它增加了对动作的表示和推理。下面给出动态描述逻辑的几个重要定义。

定义 1 在 DDL 中,“概念”定义如下:

- (1)原子概念 P、全概念 \top 和空概念 \perp 都是概念;
- (2)如果 C 和 D 都是概念,则 $\neg C, C \sqcap D, C \sqcup D$ 都是概念;
- (3)如果 R 为关系, C 为概念,则 $\exists R.C, \forall R.C$ 都是概念;
- (4)如果 C 是概念, α 是动作,则 $[\alpha]C$ 也是概念。

定义 2 DDL 的公式定义如下,其中 C 为任意概念, R 为关系, a, b 为个体常元, x, y 为个体变元, α 是动作:

- (1)形如 $C(a), R(a, b)$ 和 $[\alpha]C(a)$ 的表达式称为断言公式,它们是不带变元的;
- (2)形如 $C(x), R(x, y)$ 和 $[\alpha]C(x)$ 的表达式称为一般公式,它们是带变元的;
- (3)断言公式和一般公式都是公式;
- (4)如果 φ 和 ψ 是公式,则 $\neg\varphi, \varphi \wedge \psi, \varphi \rightarrow \psi, \forall x\varphi$ 都是公式;
- (5)如果 φ 是公式,则 $[\alpha]\varphi$ 也是公式。

定义 3 DDL 的动作定义如下:

- (1)原子动作 $\dot{A}(a_1, \dots, a_n)$ 是动作;
- (2)如果 α 和 β 为动作,则 $\alpha; \beta, \alpha \cup \beta, \alpha^*$ 都是动作,其中“;”表示“合成”,“ \cup ”表示“交替”,“ $*$ ”表示“反复”;
- (3)如果 φ 为断言公式,则 $\varphi?$ 也是动作,其中“?”表示测试。

定义 4 一个动作描述^[7]是形如 $\dot{A}(x_1, \dots, x_n) \equiv (P_A, E_A)$ 的表达形式,其中,

- (1) \dot{A} 为动作名;指示动作表示符;
- (2) x_1, \dots, x_n 为个体变元,指定动作的操作对象,因此也称之为操作变元;
- (3) P_A 为前提公式集,指定动作执行前必须满足的前提条件;
- (4) E_A 为结果公式集,指定动作执行后得到的结果集。

关于动态描述逻辑中的基本推理问题,主要包括概念的可满足性问题、实例断言集的一致性检测、动作描述的一致性以及动作可实现性和动作之间包含关系的判定等。

2 动态描述逻辑的物元扩展

动态描述逻辑作为知识表示的形式化工具,不仅可以描述静态知识,而且可以对动作进行表示和推理,所以具有很强的知识表示和推理能力。但是,由于动态描述逻辑框架中只有概念和关系,当需要描述某个概念或者个体具有的属性时,只能将它转换为概念的包含断言或关系断言的形式来描述。例如,当描述“雪是白色的”,即雪具有“白色”这个属性时,就

要转化为“雪 \sqsubseteq 白色的事物”这样一个概念断言的形式。或者,当描述物体 a 的长度为 10m 时,表示为关系断言 $\text{hasLength}(a, 10\text{m})$ 的形式。如果将属性表示为概念的包含断言形式,不仅复杂而且容易使概念和属性混为一谈,无法分清楚哪些是概念,哪些是概念所具有的本质属性,这样也就很难通过个体具有的属性判断该个体是否属于某个概念。如果将属性描述为关系的形式,那么当需要通过对象(个体)的属性来定义概念时就变得很不方便。而且对象的各个属性之间往往是相互依赖、相互制约的,上述两种形式描述属性时均不能体现这种关系。不仅如此,当一个动作对某个个体作用时,首先发生变化的是个体的某个或某些属性值。当属性值变化到一定程度时发生质变,才导致个体的变化。而上面两种方式根本无法清晰地表现个体的哪个属性及值发生了变化。所以针对这个问题,本文引入了物元来扩充动态描述逻辑。下面首先对物元的基本概念^[4]进行简要介绍。

定义 5 物元 $W = (N, A, V)$, 其中 N 既可以表示概念也可以表示个体, A 表示 N 具有的属性, V 表示 N 关于属性 A 的值,这里称它们为物元的三要素。

例如“雪的颜色是白色的”表示为 $W = (\text{雪}, \text{颜色}, \text{白色})$ 。

定义 6 概念或个体都具有多个属性,所以定义多维物元:

$$W = \begin{bmatrix} N, & A_1, & V_1 \\ & A_2, & V_2 \\ & & \dots \\ & A_n & V_n \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_n \end{bmatrix}$$

由以上定义可知,物元将概念或个体、属性和属性值集于一个统一体中。当判断某个个体是否属于某概念时,只要个体具有的属性及属性值满足概念定义的范围即可。概念与概念、个体与个体之间的区别与联系一目了然。所以用物元表示概念及概念的属性可以解决动态描述逻辑在描述属性时的不足。

定义 7 物元的“一物多征”发散推理规则,符号“ \dashv ”表示发散推理。

- 1) $N \dashv (N, A, V) \dashv \{(N, A_1, V_1), (N, A_2, V_2) \dots (N, A_n, V_n)\}$, 表示由一个事物可以发散出该事物具有的多个属性。

下面根据物元表示知识的特点对动态描述逻辑 DDL 进行扩充,首先给出 MDDL 的语法和语义。

定义 8 在 MDDL 的语言中包括以下基本符号:

- 概念名: $C_1, C_2 \dots$;
- 物元名: $W_1, W_2 \dots$;
- 个体常元: $a_1, a_2 \dots$;
- 个体变元: $x_1, x_2 \dots$;
- 关系名: $R_1, R_2 \dots$;
- 概念运算: \sqcap, \sqcup, \neg 及量词: \exists, \forall ;
- 物元运算: \sqcap, \sqcup, \neg 及量词: \exists, \forall ;
- 公式运算: $\neg, \wedge, \rightarrow$, 以及量词 \forall ;
- 动作名: $\dot{A}_1, \dot{A}_2 \dots$;
- 动作构造: 如;(合成), \cup (交替), $*$ (反复), $?$ (测试);
- 动作变元: $\alpha, \beta \dots$;
- 公式变元: $\varphi, \psi \dots$;
- 状态变元: $u, v, w \dots$ 。

其中,个体常元“ $a_1, a_2 \dots$ ”的集合用 N_C 表示,个体变元“ $x_1, x_2 \dots$ ”的集合用 N_X 表示。 N_C 与 N_X 集合中的元素既可以是概念,也可以是物元。

物元的“ \sqcap ”,“ \sqcup ”运算分别表示两个物元同时成立和两个物元至少有一个成立。

物元的“ \rightarrow ”运算是指对概念的属性值取非。假设给定物元 $W=(C,A,V)$, 则 $\rightarrow W=(C,A,\rightarrow V)$ 。

物元的“ \exists ”, “ \forall ”运算分别写作 $\exists R.W, \forall R.W$ 。例如: $R=\text{haschild}, W=(\text{people, age, } 5)$, 则 $\exists R.W, \forall R.W$ 分别表示有一个 5 岁的孩子的人和所有有孩子并且孩子都是 5 岁的人。

此时, Tbox 中的包含断言包括以下几种形式: 1) $C \sqsubseteq D$; 2) $C \sqsubseteq W$; 3) $W \sqsubseteq C$; 4) $W_1 \sqsubseteq W_2$; 5) $A_1 \sqsubseteq A_2$ 。

Abox 中的实例断言则包括概念实例断言 $C(a), W(a)$, $[\alpha]C(a), [\alpha]W(a)$, 以及关系实例断言 $R(a,b)$ 。这些统称为 MDDL 中的断言公式, 它们是不带变元的。而 $C(x), W(x)$, $[\alpha]C(x), [\alpha]W(x), R(x,y)$ 这些表达式统称为 MDDL 中的一般公式, 它们是带变元的。MDDL 的公式定义与 DDL 的公式定义类似, 只是增加了物元公式, 这里不再介绍。

带物元的动态描述逻辑 MDDL 的有关动作的定义可以参考文献[3]中的动态描述逻辑 DDL 的定义。

定义 9 MDDL 的语义

下面按照传统描述逻辑的语义解释方法解释 MDDL 的语义。首先, 对于 MDDL 的状态集合 ω 中的两个状态 u 和 v , 有在状态 u 下应用动作 $\alpha=(P_\alpha, E_\alpha)$ 能产生状态 v (记为 $u \rightarrow_\alpha v$) 的关系, 状态 u 和 v 下的两个解释形式分别为 $I(u)=(\Delta, \cdot^{I(u)})$ 和 $I(v)=(\Delta, \cdot^{I(v)})$ 。其中解释函数把每个原子概念映射到 Δ 的子集, 把每个原子物元映射到既满足物元表示的概念 C 又满足物元中的属性及属性值的 Δ 的子集, 而把每个原子关系映射到 $\Delta \times \Delta$ 的子集。概念、物元、关系、动作的语义表示如下:

- 全概念 \top 的语义为论域 Δ , 即 $\top^{I(u)} = \Delta$;
- 空概念 \perp 的语义为空集, 即 $\perp^{I(u)} = \emptyset$;
- 若 W 为物元, C 为 W 表示的概念, 则 $W^{I(u)} \sqsubseteq C^{I(u)} \sqsubseteq \Delta$;
- 若 R 为关系, 则 $R^{I(u)} \sqsubseteq \Delta \times \Delta$;
- $(\neg C)^{I(u)} = \Delta - C^{I(u)}$;
- $(\neg W)^{I(u)} = (C, A, \rightarrow V)^{I(u)} = C^{I(u)} - W^{I(u)}$;
- $(C \sqcap D)^{I(u)} = C^{I(u)} \cap D^{I(u)}$;
- $(W_1 \sqcap W_2)^{I(u)} = W_1^{I(u)} \cap W_2^{I(u)}$;
- $(C \sqcup D)^{I(u)} = C^{I(u)} \cup D^{I(u)}$;
- $(W_1 \sqcup W_2)^{I(u)} = W_1^{I(u)} \cup W_2^{I(u)}$;
- $(\exists R.C)^{I(u)} = \{x \mid \exists y, (x,y) \in R^{I(u)} \wedge y \in C^{I(u)}\}$;
- $(\exists R.W)^{I(u)} = \{x \mid \exists y, (x,y) \in R^{I(u)} \wedge y \in W^{I(u)}\}$;
- $(\forall R.C)^{I(u)} = \{x \mid \forall y, (x,y) \in R^{I(u)} \Rightarrow y \in C^{I(u)}\}$;
- $(\forall R.W)^{I(u)} = \{x \mid \forall y, (x,y) \in R^{I(u)} \Rightarrow y \in W^{I(u)}\}$;
- $([\alpha]C)^{I(u)} = \{x \mid u T_\alpha v, x \in C^{I(v)}\}$;
- $\alpha = \{\langle u, v \rangle \mid u, v \in \omega, u \rightarrow_\alpha v\}$, 其中 γ 是对动作 α 中变量的赋值: $N_X \rightarrow \Delta$;
- $\alpha; \beta = \{\langle u, v \rangle \mid u, v, w \in \omega, u \rightarrow_\alpha w \wedge w \rightarrow_\beta v\}$;
- $\alpha \cup \beta = \{\langle u, v \rangle \mid u, v \in \omega, u \rightarrow_\alpha v \vee u \rightarrow_\beta v\}$;
- $\alpha * = \{\langle u, v \rangle \mid u, v \in \omega, u \rightarrow_\alpha v \vee u \rightarrow_{\alpha; \alpha} v \vee u \rightarrow_{\alpha; \alpha; \alpha} v \vee \dots\}$;
- $\varphi? = \{\langle u, u \rangle \mid u \in \omega, u \vdash \varphi\}$ 。

上述对动作的语义解释参考了文献[3]。

3 MDDL 的基本推理问题

关于带物元的动态描述逻辑 MDDL 中的基本推理问题, 主要包括概念和物元的可满足性问题、实例断言集的一致性、动作描述的一致性以及对动作可实现性和动作之间包含关系的判定等。关于动作的推理问题参考文献[3]。

3.1 实例断言集 Abox 的一致性问题

实例断言集的一致性问题^[8]是描述逻辑推理中的基本问

题, 其它推理问题都可以转换为这种推理或者利用这种推理结果。下面首先给出一些相关定义。

定义 10 一个冲突是具有如下情形之一的集合, 其中, C 为任意概念, W 为某概念具有的属性值 (即物元), R 为任意关系, a, b 为任意个体常元:

- $\{\perp(a)\}$;
- $\{C(a), \neg C(a)\}$;
- $\{C(a), \neg W(a)\}$, 其中 $W=(C,A,V)$;
- $\{W(a), \neg C(a)\}$, 其中 $W=(C,A,V)$;
- $\{R(a,b), \neg R(a,b)\}$ 。

定义 11 一个实例断言集 Abox 是一致的, 当且仅当 Abox 中不含冲突, 否则称 Abox 是不一致的。

下面给出实例断言集 Abox 的一致性检测算法, 该算法是动态描述逻辑的一致性检测算法^[3,9]的扩充, 我们称为 Tableau-M 算法。它主要利用动态描述逻辑的推理规则及物元的“一物多征”发散推理规则对断言公式集进行扩充, 然后检测是否有冲突出现。Tableau-M 算法的具体步骤如下:

1) 按顺序使用下述规则对断言公式集 A' 进行扩充, 直到没有规则可用。这里 G 既可以表示概念 C , 又可以表示物元 W ; x, y 表示概念个体或物元个体:

- (1) \sqcap 规则: 若 $G_1 \sqcap G_2(x) \in A'$, 且 $G_1(x) \notin A', G_2(x) \notin A'$, 则将 $\{G_1(x), G_2(x)\}$ 并入 A' 中;
- (2) \sqcup 规则: 若 $G_1 \sqcup G_2(x) \in A'$, 且 $G_1(x) \notin A', G_2(x) \notin A'$, 则将 $\{G_1(x)\}$ 或 $\{G_2(x)\}$ 并入 A' 中;
- (3) \exists 规则: 若 $\exists R.G(x) \in A'$, 且没有 y 使得 $R(x,y) \in A'$ 且 $G(y) \in A'$, 则将 $\{G(y), R(x,y)\}$ 并入 A' 中;
- (4) \forall 规则: 若 $\forall R.G(x) \in A', R(x,y) \in A'$ 且 $G(y) \notin A'$, 则将 $G(y)$ 并入 A' 中;
- (5) 动作 α 规则: 若 $[\alpha]G \in A'$, 其中 $\alpha=(P_\alpha, E_\alpha)$, 则将 G 并入 A' 中, 将 P_α 所有的条件从 A' 中删除, 将执行 E_α 后得到的结果添加到 A' 中;
- (6) $N \vdash (N, A, V) \vdash \{(N, A_1, V_1), (N, A_2, V_2) \dots (N, A_n, V_n)\}$, 将发散后的结果都并入 A' 中。

规则(1)-(5)是将动态描述逻辑的算法规则进行物元扩展后而形成的。

规则(6)是利用“一物多征”的发散规则对断言公式集 A' 进行扩充。例如, 断言公式集 A' 中有 $C(a)$ 或 $W(a)$, 则个体 a 一定满足概念 C 的各个属性, 所以要将表示个体 a 的各个属性的物元加入到断言公式集 A' 中。

2) 检测 A' 中是否包含有冲突, 若没有冲突, 则 A' 是一致的, 否则 A' 是不一致的。算法终止。

在上述几条规则中, “ \sqcup 规则”是不确定的, 即可能得到两个分支 A_1' 和 A_2' 。若多次使用该规则, 则将得到一个有穷集合 $\{A_1', A_2', \dots, A_n'\}$ 。此时只要 A_1', \dots, A_n' 中有一个公式集是一致的, 即不包含冲突, 则 A' 就是一致的; 反之, 若 A_1', \dots, A_n' 中都包含有冲突, 则 A' 就是不一致的。在 Tableau-M 算法运行的过程中, 只要有一个分支出现了冲突, 就可以将此分支去掉。

上述算法是动态描述逻辑算法的推广, 发散推理规则(6)可以将实例断言集中的个体具有的各个属性显示化地表示出来, 它适合处理知识库不完全或存在隐性知识的问题。

当公式通过上述步骤进行扩充后, 由于规则(6)是在给定的知识库中进行发散推理, 因此它是确定可以终止的, 而在得到的断言公式集中查找冲突的过程也是确定的。要么有冲突, 是不一致的; 要么没有冲突, 是一致的。因此实例断言集的一致性判定过程是确定可停止的, 即是可判定的。

(下转第 170 页)

符串中 e 为切分字符, 去掉 e 后的切分子串集合为 {"bcd", "bcd", "bcd", "bcd", "bcd", "bcd"}, 这些切分子串的长度(等于 3)均小于模式字符串的长度(等于 4), 所以不用再进行字符串的比较。而如果按照 BM 算法进行匹配需要比较字符 $4 \times 7 = 28$ 次, 移动比较位置 7 次。采用本文的新算法则只比较 7 次左右判断切分字符, 没有字符移动操作。在 BM 算法最优情况执行时间近似相等, 如 $T = \text{"slkjdfsjajdlshhkhjsaklhashedsekshfkskajdaj"} , n = \text{len}(T) = 60, p = \text{"mmmmnnnn"} , m = \text{len}(p) = 8, \text{BM 算法作字符比较 } n/m = 7 \text{ 次, 移动比较位置 } n/m = 7 \text{ 次, 采用新算法作切分 } 1 + 2 + 4 = 7 \text{ 次, 没有字符比较操作。}$

② 新算法便于并行编程, 提高算法具体实现程序的执行速度。BM 算法中字符比较和移位操作只能串行操作, 总的时间花销在最佳情况下为 $O([n/m])$, 而新算法可以将切分操作并行进行, 这样, 实现总的时间开销最佳情况下为 $O(\log_2 [n/m])$, 在 $[n/m]$ 较大情况下, 时间差别将很大, 比如 $[n/m] = 256$, 则 $\log_2 [n/m] = 8$, BM 算法的执行时间是改进算法的 32 倍。

5 算法测试及结果

为了评测该算法的性能, 随机地抽取一段文本和模式串, 并在同一台计算机上用不同的算法进行匹配。测试文本 $T = \text{"From automated teller machines and atomic clocks to mam-mograms and semiconductors, innumerable products and services rely in some way on technology, measurement and standards provided by the National Institute of Standards and Technology"} , \text{模式串 } P = \text{"products and services"} . \text{ 分别用}$

BM 算法和 split_BM_in_frequence_match 算法在同一台计算机上进行匹配计算, 并统计每种算法匹配时总的字符匹配次数。测试结果如表 4。

表 2 匹配算法与 BM 算法比较实验结果

	BM	split_BM_in_frequence_match
一次匹配的总的字符匹配次	129	109

实验结果表明了 split_BM_in_frequence_match 算法能取得比 BM 算法更快的速度。

结束语 本文研究了当前使用非常广泛且性能较好的 BM 算法, 提出了基于字符使用频率和分治思想的字符串模式匹配 split_BM_in_frequence_match 算法, 算法结合 BM 算法并利用字符统计使用频率和分治思想来提高坏字符的失配率, 减少不必要的比较次数。同时, split_BM_in_frequence_match 算法更容易并行编程实现, 可以进一步提高匹配的速度。比较试验进一步证实了 split_BM_in_frequence_match 算法的有效性。

参考文献

- [1] Knuth D, Morris J, Pratt P. Fast pattern matching in strings. SIAM Journal on Computing, 1977, 6(2):323
- [2] Boyer R S, Moore J S. A fast string searching algorithm. Communications of the ACM, 1977, 20(10):762
- [3] Nigel H R. Practical fast searching in strings. Software Practice and Experience, 1980, 10(6):501-506
- [4] Aho A, Corasick M. Efficient string matching: An aid to bibliographic search. Communications of the ACM, 1975, 18(6):333-343
- [5] <http://crypt.math.fju.edu.tw/key/09/Freq.html>

(上接第 157 页)

3.2 概念和物元的可满足性问题

定义 12 某个概念 C 或物元 W 在 MDDL 结构中是可满足的, 当且仅当 C 或 W 不包含于 \perp 。

上述定义的语义为, 若在 MDDL 结构中不存在一个解释 I 使得 $C^I \neq \emptyset$ 或 $W^I \neq \emptyset$, 则称概念 C 或物元 W 是可满足的。

对于概念或物元的可满足性问题, 关键就是要构造出关于概念或物元的一个解释, 使得该解释是非空的, 即不包含冲突。

以物元 W 为例, 令 W 为一个要检测的物元, 并假定某个个体 a 为物元 W 的实例, 于是得到断言公式 $W(a)$ 。现在令断言公式集 $A' = \{W(a)\}$, 利用实例断言集的一致性检测算法来判断 A' 是否一致。如果 A' 中不包含冲突, 则 A' 是一致的。此时, 实际上就已经找到了物元 W 的一个解释 W^I , 使得至少有 $a \in W^I$, 即物元 W 不包含于 \perp , 于是物元 W 是可满足的。否则如果 A' 是不一致的, 则说明 W 包含于 \perp , 此时 W 是不可满足的。

由上可知, 概念或物元的可满足性问题可以直接转化为实例断言集的一致性检测问题, 所以概念和物元的可满足性问题是可判定的。

结束语 本文针对动态描述逻辑在描述概念的属性及其变化时的不足, 提出了一种带物元的动态描述逻辑 MDDL。该描述逻辑采用物元来统一描述概念和个体、属性、属性值, 并按照传统描述逻辑的语义解释方法给出了物元的语义解释。这种描述事物的方法不仅充分表达了事物与属性之间的密切关系, 而且可以体现执行动作时个体、属性和属性值的动态变化, 使人们在处理问题的时候既考虑量的变化, 又考虑质

的变化。另外, 描述逻辑 MDDL 中还利用物元“一物多征”的发散推理规则, 扩充动态描述逻辑 DDL 的语法规则, 提出一种新的 Tableau-M 算法, 并根据该算法深入研究了 MDDL 的基本推理问题, 即实例断言集的一致性检测和概念与物元的可满足性检测。由于物元的发散规则是为了模仿人类的发散性思维而建立的, 因此利用它可以对现有知识库进行扩展, 从而得到更完备的知识。

参考文献

- [1] Baader F, Nutt W. Basic description logic[G]//Baader F, et al. eds. The Description Logic Handbook: Theory, Implementation and Applications. Cambridge University Press, 2003:47-100
- [2] Baader F, Horrocks I, Sattler U. Description logics as ontology languages for the semantic Web//Dieter Hutter, Werner Stephan, eds. Festschrift in honor of Jörg Siekmann, Berlin: Springer, 2003
- [3] 史忠植, 董明楷, 蒋运承, 等. 语义 Web 的逻辑基础[J]. 中国科学, E 辑, 2004, 34(10):1123-1138
- [4] 蔡文, 杨春燕, 林伟初. 可拓工程方法[M]. 北京: 科学出版社, 2000
- [5] 刘巍, 张秀芳. 基于可拓信息的知识表示[J]. 系统工程理论与实践, 1998, 18(1):104-107
- [6] 蔡文, 杨春燕, 何斌. 可拓逻辑初步[M]. 北京: 科学出版社, 2003
- [7] Lutz C, Sattler U. A Proposal for Describing Services with DLs//Horrocks T S, ed. Proceedings of the 2002 International Workshop on Description Logics, Aachen: CEUR-WS, 2002:129-140
- [8] Baader F, Sattler U. An Overview of Tableau Algorithms for Description Logics. Studia Logica, 2001, 69:5-40
- [9] Schmidt-Schauß M, Smolka G. Attributive concept description with complements. Artificial Intelligence, 1991, 48(1):1-26