基于动态时序描述逻辑的动作理论

孙永新[1,2] 赵希顺[1]

(中山大学逻辑与认知研究所 广州 510275)[1] (仲恺农业工程学院信息科学与技术学院 广州 510225)[2]

摘 要 动态时序描述逻辑($DLTL_{DL}$)是一类描述逻辑的动态时序扩展。提出一种基于 $DLTL_{ALCIO}$ 的动态域建模方法,利用该方法可构造出刻画动态域知识的 $DLTL_{ALCIO}$ 理论,并解决动作推理中的框架问题和分支问题。动作推理问题,如动作可执行性和投影问题等,可归结为关于 $DLTL_{ALCIO}$ 理论的推理问题,并最终归结为 $DLTL_{ALCIO}$ 的公式可满足性问题。$DLTL_{ALCIO}$ 公式可表达动作和时间约束,相对于其他基于描述逻辑的动作形式,基于 $DLTL_{ALCIO}$ 的动作形式在需要执行复杂查询,尤其是含时间或动作的查询的应用场合具有更好的适用性。

关键词 动态时序描述逻辑,动作推理,动态域,动作理论

中图法分类号 TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.09.039

Action Theory Based on Dynamic Linear Temporal Description Logic

SUN Yong-xin[1,2] ZHAO Xi-shun[1]

(Institute of Logic and Cognition, Sun Yat-sen University, Guangzhou 510275, China)[1]

(College of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)[2]

Abstract Dynamic linear temporal description logics ($DLTL_{DL}$ s) are a family of dynamic and temporal extensions of description logics. Based on $DLTL_{ALCIO}$, a method for modeling dynamic domains was presented. By application of the modeling method, a $DLTL_{ALCIO}$ theory for describing a giving domain can be constructed, and frame problem and ramification problem can be solved. Action reasoning problems, such as executability problem and projection problem, can be reduced to reasoning problems in the $DLTL_{ALCIO}$ theory, and can eventually be reduced to the satisfiablity problem of $DLTL_{ALCIO}$ formulas. Since constraints with action and time can be expressed with $DLTL_{ALCIO}$ formulas, so compared with the other action formalisms based on description logics, the action formalism based on $DLTL_{ALCIO}$ is more applicable in the case of complex queries, especially queries containing time or action, to be executed.

Keywords Dynamic linear temporal description logic, Reasoning about action, Dynamic domain, Action theory

1 引言

语义 Web 服务等新型动态应用的出现,对动作知识表示和推理形式的可判定性和表达能力提出新的要求,传统动作形式面临两难困境。例如:情景演算[1]和流演算[2]包含了完全一阶逻辑,有足够强的表达能力描述语义 Web 服务,然而其推理一般不可判定,其他基于一阶动态逻辑和一阶时态逻辑的动作形式也存在类似问题;而基于命题动态逻辑或命题时态逻辑的动作形式虽然可判定,但它们的表达能力太弱,不足以描述语义 Web 服务。

面对以上问题,近年来研究人员将描述逻辑引入动作推理研究领域。描述逻辑是一族基于逻辑的知识表示形式,一般用于表达静态域知识,例如语义 Web 的标准语言 OWL 就以描述逻辑作为其形式基础。语义 Web 服务是一类语义 Web 上的动态应用,其标准描述语言 OWL-S 建立在 OWL 基础上。描述逻辑是一阶逻辑子集,具有较强的表达能力,又具有可判定性,通过扩展描述逻辑构建新的动作形式,是解决以上两难问题的一种自然途径。

基于描述逻辑,近年研究人员已提出一些可判定的动作形式。例如,Baader 等[3]提出用特定形式的描述逻辑公式表示动作的前提和效果,采用可能模型方式定义动作语义,动作可执行性和投影推理可采用类似情景演算的回归运算的方式归结到描述逻辑公式可满足性推理。常亮等在文献[4]中提出基于动态描述逻辑 DDL 的动作理论,除了用描述逻辑公式表示动作的前提和效果外,DDL 还支持了原子动作上的正规表达式,可以用于表达和推理具有选择、迭代等结构的组合动作。笔者曾提出一类描述逻辑的动态时序扩展 $DLTL_{DL}$ (Dynamic Linear Temporal Description Logic)[5],$DLTL_{DL}$ 可看作增加了动作表示的时序描述逻辑,也以正规表达式表示动作组合,利用 $DLTL_{DL}$ 公式还可以很方便地表达动态系统的时态特性,然而文献[5]中并没有提出如何将 $DLTL_{DL}$ 应用于动作表示和推理。

本文研究以 $DLTL_{ALCIO}$ 为形式基础的动作表示和推理。第 2 节提出 $DLTL_{ALCIO}$ 的语法和语义;第 3 节提出构造动态域描述和基于 $DLTL_{ALCIO}$ 的域理论的方法;第 4 节给出一个第 3 节方法应用的例子。

到稿日期:2013-11-05 返修日期:2014-03-08 本文受国家自然科学基金项目(61272059),教育部基地重大项目(11JJD7200020)资助。
孙永新(1972—),男,博士生,工程师,CCF 会员,主要研究领域为描述逻辑、语义 Web、软件工程、人工智能,E-mail:syx2000@tom.com;赵希顺(1964—),男,教授,博士生导师,主要研究领域为可计算性与计算复杂性理论、数理逻辑、人工智能逻辑。

2 $DLTL_{ALCIO}$

令 $\Sigma=\{a_1,a_2,\cdots,a_n\}$ 为原子动作集合。Σ 上正规程序的生成产生式为：

$\pi,\pi'::=a|\pi;\pi'|\pi+\pi'|\pi^*$

其中，$a\in\Sigma$ 称为原子动作，;，+，* 分别表示动作的顺序、选择和迭代操作。正规程序 π 表示的正规集合记作 $RS(\pi)$（$RS(\pi)$ 的定义可参见文献[5]）。Σ 上所有正规程序的集合记作 $prg(\Sigma)$，Σ 上无穷长的动作序列集合记作 Σ^ω。文中用 $\tau_0\leqslant\tau_1$、$\tau_0<\tau_1$ 分别表示符号串 τ_0 是 τ_1 的前缀和严格前缀，符号串 σ 的前缀的集合记作 $prf(\sigma)$。

定义 1（$DLTL_{ALCIO}(\Sigma)$ 语法） 给定原子动作集合 Σ，概念名集合 $N_C=\{A_1,A_2,\cdots,A_m\}$，角色名集合 $N_R=\{R_1,R_2,\cdots,R_n\}$ 和个体名集合 $N_I=\{a_1,a_2,\cdots,a_k\}$，$DLTL_{ALCIO}(\Sigma)$ 角色由产生式 $R::=R_i|R_i^-$ 生成，其中 R_i^- 为 R_i 的逆角色，$R_i\in N_R$，文中用 $R^{\ominus}$ 表示 R_i^-（R 为 R_i）或 R_i（R 为 R_i^-）。$DLTL_{ALCIO}(\Sigma)$ 概念由产生式 $C,D::=\{a_i\}|A_i|\top|\neg C|C\sqcap D|\exists R.C|CU^{\pi}D$ 生成，其中 $a_i\in N_I$，形如 $\{a_i\}$ 的概念称为枚举概念，$A_i\in N_C$，R 为 $DLTL_{ALCIO}(\Sigma)$ 角色，$\pi\in prg(\Sigma)$。$DLTL_{ALCIO}(\Sigma)$ 公式由产生式 $\alpha,\beta::=\mathrm{true}|C=D|C(a)|R(a,b)|\neg\alpha|\alpha\wedge\beta|\alpha U^{\pi}\beta$ 生成，其中 $a,b\in N_I$，C,D 为 $DLTL_{ALCIO}(\Sigma)$ 概念，R 为 $DLTL_{ALCIO}(\Sigma)$ 角色。

定义 2（$DLTL_{ALCIO}(\Sigma)$ 语义） $DLTL_{ALCIO}(\Sigma)$ 模型是一个三元组 $M=(\sigma,<,I)$，其中 $(\sigma,<)$ 是关于 $\sigma\in\Sigma^\omega$ 的线性序列，$<$ 是 $prf(\sigma)$ 上的严格前缀关系，I 是一个解释函数，将每个点 $\tau\in prf(\sigma)$ 映射为一个 $ALCIO$ 模型（$\Delta^{I(\tau)},a_1^{I(\tau)},\cdots,R_1^{I(\tau)},\cdots,A_1^{I(\tau)},\cdots$）。$\Delta^{I(\tau)}$ 是 M 在 τ 点的非空域，对任意 $\tau_1<\tau_2$ 有 $\Delta^{I(\tau_1)}\subseteq\Delta^{I(\tau_2)}$；在 τ 点，每个 $R_i\in N_R$、$A_i\in N_C$ 和 $a_i\in N_I$ 分别映射为 $\Delta^{I(\tau)}$ 上的二元关系 $R_i^{I(\tau)}$，$\Delta^{I(\tau)}$ 的子集 $A_i^{I(\tau)}$ 和 $\Delta^{I(\tau)}$ 中元素 $a_i^{I(\tau)}$，对任意 $i\neq j$ 有 $a_i^{I(\tau)}\neq a_j^{I(\tau)}$，对 $\tau_1,\tau_2\in prf(\sigma)$ 有 $a_i^{I(\tau_1)}=a_i^{I(\tau_2)}$，$a_i^{I(\tau)}$ 称为有名对象，$\Delta^{I(\tau)}$ 中的其他元素称为匿名对象。按以下方式扩展上述解释，逆角色 $R_i^-(R_i\in N_R)$ 和任意概念在 τ 点将分别被解释为 $\Delta^{I(\tau)}$ 上的二元关系和 $\Delta^{I(\tau)}$ 上的子集：

$(R_i^-)^{I(\tau)}=\{(d,d_1)|(d_1,d)\in R_i^{I(\tau)}\}$；$\top^{I(\tau)}=\Delta^{I(\tau)}$；

$\{a\}^{I(\tau)}=\{a^{I(\tau)}\}$；$(\neg C)^{I(\tau)}=\Delta^{I(\tau)}\backslash C^{I(\tau)}$；

$(C\sqcap D)^{I(\tau)}=C^{I(\tau)}\cap D^{I(\tau)}$；

$(\exists R.C)^{I(\tau)}=\{d\in\Delta^{I(\tau)}\mid\exists d_1\in C^{I(\tau)}\wedge(d,d_1)\in R^{I(\tau)}\}$；

$(CU^{\pi}D)^{I(\tau)}=\{d\in\Delta^{I(\tau)}\mid\exists\tau_1\in RS(\pi)(\tau\tau_1\in prf(\sigma),d\in D^{I(\tau\tau_1)},\forall\tau_2(\varepsilon\leqslant\tau_2<\tau_1\Rightarrow d\in C^{I(\tau\tau_2)}))\}$。

$DLTL_{ALCIO}(\Sigma)$ 真值关系 $M,\tau\models\phi$ 定义如下：

$M,\tau\models\mathrm{true}$；$M,\tau\models C=D$ 当且仅当 $C^{I(\tau)}=D^{I(\tau)}$；

$M,\tau\models C(a)$ 当且仅当 $a^{I(\tau)}\in C^{I(\tau)}$；

$M,\tau\models R(a,b)$ 当且仅当 $(a^{I(\tau)},b^{I(\tau)})\in R^{I(\tau)}$；

$M,\tau\models\neg\alpha$ 当且仅当非 $M,\tau\models\alpha$；

$M,\tau\models\alpha\wedge\beta$ 当且仅当 $M,\tau\models\alpha$，$M,\tau\models\beta$；

$M,\tau\models\alpha U^{\pi}\beta$ 当且仅当 $\exists\tau_1\in RS(\pi)(\tau\tau_1\in prf(\sigma),M,\tau\tau_1\models\beta,\forall\tau_2(\varepsilon\leqslant\tau_2<\tau_1\Rightarrow M,\tau\tau_2\models\alpha))$。

文中，在不影响阅读的前提下我们一般将 $DLTL_{ALCIO}(\Sigma)$ 写作 $DLTL_{ALCIO}$。给定 $DLTL_{ALCIO}$ 公式 α，判断是否存在模型 M，$M,\varepsilon\models\alpha$ 的问题称为公式可满足问题，如模型存在，则称 α 可满足。

由于篇幅限制，文中没有给出 $DLTL_{ALCIO}$ 公式可满足算法，不过扩展文献[5]中的 tableau 算法不难设计出 $DLTL_{ALCIO}$ 的公式可满足算法（增加部分关于 I，O 算子的 tableau 规则，并将局部阻塞方式改为动态等集阻塞（参见文献[6]的 6.2.2）），采用与文献[5]中定理 1 证明类似的方式也不难证明扩展后的算法可终止。$DLTL_{ALCIO}(\Sigma)$ 公式可满足问题是可判定的。

对一个 $DLTL_{ALCIO}$ 公式集合 T，如果存在 $DLTL_{ALCIO}$ 模型 M 使得 $M,\varepsilon\models\bigwedge_{\alpha\in T}\alpha$，则称 M 是 T 的模型，T 可满足或 T 有模型。

定义 3（$DLTL_{ALCIO}$ 后承关系） 给定一个 $DLTL_{ALCIO}$ 公式集合 T 和一个 $DLTL_{ALCIO}$ 公式 ϕ，如对任意使得 $M,\varepsilon\models\bigwedge_{\alpha\in T}\alpha$ 的 $DLTL_{ALCIO}$ 模型 M 都有 $M,\varepsilon\models\phi$，则称 ϕ 为 T 的后承，记作 $T\models\phi$。

显然，$T\models\phi$ 当且仅当 $\neg\phi\wedge(\bigwedge_{\alpha\in T}\alpha)$ 不可满足。文中用 $\perp$，$C\sqcup D$，$\forall R.C$，$\alpha\vee\beta$，$\alpha\rightarrow\beta$，$C\sqsubseteq D$ 分别作为形如 $\neg\top$，$\neg(\neg C\sqcap\neg D)$，$\neg\exists R.C$，$\neg(\neg\alpha\wedge\neg\beta)$，$\neg\alpha\vee\beta$ 和 $(\neg C\sqcup D)=\top$ 的概念或公式的缩写，不含 U^{π} 算子的概念、公式分别称为 $ALCIO$ 概念和 $ALCIO$ 公式，其中形如 $C(a)$，$\neg C(a)$，$R(a,b)$ 或 $\neg R(a,b)$ 的 $ALCIO$ 公式称为 $ALCIO$ 断言（简称断言），如一个断言中可能出现的概念必形如 $A_i\in N_C$，则该断言又称为一个 $ALCIO$ 文字（简称文字），不含 $\neg$ 的文字为正文字（又称原子），反之则为负文字。

利用 U^{π} 算子，传统动态算子 $\langle\pi\rangle$、$[\pi]$ 和时态算子 U(until)、$\bigcirc$(next)、$\Diamond$(sometimes)、$\Box$(always) 可以分别按以下方式定义（在未特别说明情况下 θ 可为公式或概念）：$\langle\pi\rangle\theta\equiv\mathrm{true}U^{\pi}\theta$（$\theta$ 为公式），$\langle\pi\rangle\theta\equiv\top U^{\pi}\theta$（$\theta$ 为概念）；$[\pi]\theta\equiv\neg\langle\pi\rangle\neg\theta$；$\theta U\gamma\equiv\theta U^{\Sigma^*}\gamma$（$\Sigma^*$ 表示 $(\underset{a\in\Sigma}{+}a)^*$）；$\bigcirc\theta\equiv\bigvee_{a\in\Sigma}\mathrm{true}U^a\theta$（$\theta$ 为公式），$\bigcirc\theta\equiv\bigsqcap_{a\in\Sigma}\top U^a\theta$（$\theta$ 为概念）；$\Diamond\theta\equiv\top U\theta$；$\Box\theta\equiv\neg\Diamond\neg\theta$。

3 基于 $DLTL_{ALCIO}$ 的域描述和动作理论

本节提出开放世界假设下，基于 $DLTL_{ALCIO}$ 公式形式的动态系统建模方法。我们假设域中对象包含有名和匿名对象，任意有名对象具有唯一名字，域在任意时刻的“状态”为一个 $ALCIO$ 解释。与其他形式的动作推理类似，文中还假设域的变化遵循极小法则，动态域模型是一个符合以下定义的极小模型：

定义 4（极小模型） 称 $DLTL_{ALCIO}$ 公式集合 Ω 的模型 $M=(\sigma,<,I)$ 为极小模型当且仅当不存在满足以下条件的 $\tau\in prf(\sigma)$ 和 Ω 的其他模型 M'：

1）M 与 M' 仅解释函数不同，即 $M'=(\sigma,<,I')$ 且对任意 $\tau'\in prf(\sigma)$ 有 $\Delta^{I(\tau')}=\Delta^{I'(\tau')}$。

2）对在 Ω 中出现的任意个体名 a 有：$a^{I(\tau)}=a^{I'(\tau)}$，对在 Ω 中出现的任意概念名 A、角色名 R 和 $d_1,d_2\in\Delta^{I(\tau)}$ 有：$d_1\in A^{I(\tau)}$ 当且仅当 $d_1\in A^{I'(\tau)}$，$(d_1,d_2)\in R^{I(\tau)}$ 当且仅当 $(d_1,d_2)\in R^{I'(\tau)}$。

3）a 为原子动作，$\tau a\in prf(\sigma)$，对在 Ω 中出现的任意概念名 A、角色名 R 和 $d_1,d_2\in\Delta^{I(\tau)}$ 有：如果 $d_1\in A^{I(\tau)}$，$d_1\in A^{I(\tau a)}$，则 $d_1\in A^{I'(\tau)}$，$d_1\in A^{I'(\tau a)}$；如果 $(d_1,d_2)\in R^{I(\tau)}$，$(d_1,d_2)\in R^{I(\tau a)}$，则 $(d_1,d_2)\in R^{I'(\tau)}$，$(d_1,d_2)\in R^{I'(\tau a)}$。

4）a 为原子动作，$\tau a\in prf(\sigma)$，存在 $d_1,d_2\in\Delta^{I(\tau)}$ 和 Ω 中

出现的某概念名 A 或角色名 R 满足：$d_1\in A^{I(\tau)}$，$d_1\notin A^{I(\tau a)}$ 但 $d_1\in A^{I'(\tau)}$，$d_1\in A^{I'(\tau a)}$；或，$(d_1,d_2)\in R^{I(\tau)}$，$(d_1,d_2)\notin R^{I(\tau a)}$，但 $(d_1,d_2)\in R^{I'(\tau)}$，$(d_1,d_2)\in R^{I'(\tau a)}$。

直观上，以上定义中的极小模型 M 刻画了一种在动作执行的后续状态，域中对象的属性和联系与动作前相比，除应 Ω 中约束不得不发生改变的部分外，其他的保持不变的模型（满足 Ω 中约束、"变化"更小的模型 M' 不存在），因而是一种"极小变化"模型。文中关于域描述的动作推理是基于极小模型的推理。

给定应用域，我们按照以下步骤 1－5 构造域描述 Ω，关于 Ω 的极小模型的推理是一种非单调的推理，在步骤 6 我们由 Ω 构造基于 $DLTL_{ALCIO}$ 的动作理论 T_Ω，将关于 Ω 的极小模型的推理问题转化为关于 T_Ω 的 $DLTL_{ALCIO}$ 模型的推理问题。

与文献[3,4]等类似，文中假设域的动态变化仅发生在有名对象上。导致变化的动作效果，包括动作的直接效果和间接效果，分别由以下步骤 1、2 中的效果公理和因果规则给出。

步骤 1　效果公理。对每一个动作 a，用一组以下形式的公理刻画 a 的直接效果：

$\Box(\langle a\rangle\text{true}\wedge\varphi_{a,l}\rightarrow\bigcirc l)$

其中，$\varphi_{a,l}$ 为 $ALCIO$ 公式，l 为 $ALCIO$ 文字。

$\varphi_{a,l}$ 称为动作 a 效果 l 的后置条件，l 为正文字的效果公理又称为正效果公理，否则称为负效果公理。例如，在硬币域中，假设 $flip$ 表示翻转硬币动作，$Up(c)$ 表示硬币正面朝上，则可用 $\Box(\langle flip\rangle\text{true}\wedge\neg Up(c)\rightarrow\bigcirc Up(c))$，$\Box(\langle flip\rangle\text{true}\wedge Up(c)\rightarrow\bigcirc\neg Up(c))$ 分别表示翻转反面朝上、正面朝上的硬币的直接效果。

动作 a 的效果公理集合记作 $D_{a,ef}$，所有效果公理的集合记作 D_{ef}。为了保证动态域有无限长的线性模型，文中假设 D_{ef} 中包含有空动作 nop 的效果公理：对任意文字 l，均有 $\{\Box(\langle nop\rangle\text{true}\wedge l\rightarrow\langle nop\rangle l)\}\subseteq D_{ef}$。

步骤 2　因果规则。可导出动作间接效果的状态约束常称为分支约束[7]，最常见的分支约束是因果关系，文中仅考虑可用以下形式描述的因果关系：

$\Box(\alpha_{c,p}\wedge\bigcirc\phi_{c,l}\rightarrow\bigcirc l)$

其中，$\alpha_{c,p}$ 为 $ALCIO$ 公式，$\phi_{c,l}$ 为 $ALCIO$ 文字的合取，l 为 $ALCIO$ 文字。

例如，在一个关于"动物走动"的域中用 $\Box(\bigcirc Dead(x)\rightarrow\bigcirc\neg Walking(x))$ 表示因果关系"x 死去导致 x 不再走动"。更多关于因果关系的描述及其例子可参见文献[8,9]。

因果规则集合记作 D_{ca}。

步骤 3　前提公理。文中不考虑条件约束（可间接导出动作前提的状态约束，参见文献[7]），假设动作的前提完全由前提公理显式给出。对每一个动作 a，其前提公理形式如下：

$\Box(\langle a\rangle\text{true}\rightarrow\varphi_a)$

其中，φ_a 为 $ALCIO$ 公式，表示动作 a 执行的前提条件。

例如，在硬币域中，如 $Stuck(c)$ 表示硬币被黏住，则可用 $\Box(\langle flip\rangle\text{true}\rightarrow\neg Stuck(c))$ 表示硬币没有被黏住是翻转硬币的前提条件。

前提公理集合记作 D_{pr}。

步骤 4　域约束公理。文中假设除 D_{ca} 中公式所描述的状态约束以外，其他状态约束都是对应用域的背景描述，称为域约束，可以用以下形式的公式描述它们：$\Box\varphi$，其中 φ 为 $ALCIO$ 公式。

域约束公理集合记作 D_{do}。为了构造下文的后续状态公理，我们规定 D_{do} 中包含一个公理 $\Box(Nom=\{a_1\}\sqcup\{a_2\}\sqcup\cdots\sqcup\{a_n\})$，其中 $a_1,a_2,\cdots,a_n$ 是域中所有的有名对象名，Nom 表示有名对象概念。

步骤 5　初始状态描述公式。域初始状态描述可以由一个 $ALCIO$ 公式 φ_i 表示，根据开放世界假设，φ_i 描述的是关于域初始状态的不完全知识。

令域描述 $\Omega=D_{ef}\cup D_{ca}\cup D_{pr}\cup D_{do}\cup\{\varphi_i\}$。在 Ω 任意元素中出现过的原子的集合记作 P_Ω，出现过的非负 $ALCIO$ 概念的集合记作 C_Ω，令 $CL_\Omega=C_\Omega\cup\{\exists R^{\ominus}.\{a\}\mid\exists R.C(a)$ 在 Ω 任意元素中出现过$\}$。

定义 5(分层)　令 p,p' 为正文字，如果 Ω 中存在因果规则 $\Box(\alpha_{c,p}\wedge\bigcirc\phi_{c,l}\rightarrow\bigcirc l)$ 满足 p 出现 $\Box$ 限定的公式的 $\rightarrow$ 的右端，而 p' 出现在 $\rightarrow$ 的左端，则称 p 直接依赖于 p'（记作 $p\mapsto p'$）。如 $p_0\mapsto p_1\mapsto\cdots\mapsto p_n$，则称 p_0 传递依赖于 p_n，如果 Ω 中不存在 p 满足 p 传递依赖 p，则称 Ω 是分层的，否则称 Ω 中存在循环或递归依赖。

文中假定 Ω 分层，关于非分层的讨论可参见文献[8]等。

步骤 6　后续状态公理。在我们的假设中，动态域是一个满足 Ω 的极小模型，域变化完全由 D_{ef} 和 D_{ca} 刻画的动作效果决定。从 D_{ef} 和 D_{ca} 中公式形式看，$\Box$ 限定的公式的后件均为 $ALCIO$ 文字，动作的效果只影响有名对象的属性和联系，匿名对象的属性和联系则在动作前后保持不变。我们可以按以下方式 a)、b) 分别构造一组关于有名对象和匿名对象的后续状态公理。

a) 有名对象的后续状态公理。对每个 $p\in P_\Omega$，假设 D_{ef} 中分别有 m,n 个直接效果为 $p,\neg p$ 的效果公理，其形式分别为：

$\Box(\langle a_i\rangle\text{true}\wedge\varphi_{a_i,p}\rightarrow\bigcirc p)$，$1\leqslant i\leqslant m$

$\Box(\langle a_i'\rangle\text{true}\wedge\varphi_{a_i',\neg p}\rightarrow\bigcirc\neg p)$，$1\leqslant i\leqslant n$

D_{ca} 中分别有 m',n' 个间接效果为 $p,\neg p$ 的因果规则，其形式分别为：

$\Box(\alpha_{c_i,p}\wedge\bigcirc\varphi_{c_i,p}\rightarrow\bigcirc p)$，$1\leqslant i\leqslant m'$

$\Box(\alpha_{c_i',\neg p}\wedge\bigcirc\varphi_{c_i',\neg p}\rightarrow\bigcirc\neg p)$，$1\leqslant i\leqslant n'$

对每个 $\bigcirc p$，应用 CLARK 完备计算出以下形式的后续状态公理：

$\Box(\bigcirc p\leftrightarrow(\varphi_{ep}\vee\varphi_{cp})\vee(p\wedge\neg(\varphi_{e\neg p}\vee\varphi_{c\neg p})))$

其中，φ_{ep}，$\varphi_{e\neg p}$ 分别表示 $\bigvee_{i\in[1,m]}(\langle a_i\rangle\text{true}\wedge\varphi_{a_i,p})$，$\bigvee_{i\in[1,n]}(\langle a_i'\rangle\text{true}\wedge\varphi_{a_i',\neg p})$，$\varphi_{cp}$，$\varphi_{c\neg p}$ 分别表示 $\bigvee_{i\in[1,m']}(\alpha_{c_i,p}\wedge\bigcirc\varphi_{c_i,p})$，$\bigvee_{i\in[1,n']}(\alpha_{c_i',\neg p}\wedge\bigcirc\varphi_{c_i',\neg p})$。

所有有名对象的后续状态公理的集合记作 D_{sn}。

b) 匿名对象的后续状态公理。对任意非枚举概念 $C\in CL_\Omega$，其匿名对象的后续状态公理形式为：

$\Box(C\sqcap\neg Nom=\bigcirc(C\sqcap\neg Nom))$

所有匿名对象的后续状态公理的集合记作 D_{sa}，后续状态公理的集合 $D_s=D_{sn}\cup D_{sa}$。

经过步骤 1－6，最后可得到与域描述等价的（单调）$DLTL_{ALCIO}$ 理论 $T_\Omega=D_s\cup D_{ca}\cup D_{pr}\cup D_{do}\cup\{\varphi_i\}$。

定理 1　如果 Ω 是分层的，则 Ω 有极小模型当且仅当 T_Ω 有模型。

证明：由 D_s 的构造可看出，Ω 分层和 Ω 有极小模型显然可以推出 T_Ω 有模型。下面要证明 Ω 分层和 T_Ω 有模型可推得 Ω 有极小模型。

由假设 T_Ω 有模型 $M=(\sigma,<,I)$，则由 $D_{sa}\models\phi$（ϕ 为 D_{ef} 中的任意公式）可知 M 是 Ω 的模型，以下证明分有名对象和匿名对象的变化两种情形来讨论：

a)有名对象的变化符合极小模型中对象变化的定义。假设不然，则存在 Ω 的模型 $M'=(\sigma,<,I')$ 和 $\tau,\tau a\in prf(\sigma)$ 满足极小模型定义中条件 1)－4)，从而有

· 对任意 $ALCIO$ 公式 α，均有 $M,\tau\models\alpha$ 当且仅当 $M',\tau\models\alpha$；

· 存在 $p_1\in P_\Omega$，满足：(1)$M,\tau\models p_1,M,\tau a\models\neg p_1,M',\tau a\models p_1$ 或 $M,\tau\models\neg p_1,M,\tau a\models p_1,M',\tau a\models\neg p_1$；(2)不存在 $p_0\in P_\Omega$，使得 p_1 传递依赖于 p_0 且 p_0 满足(1)。

下面只讨论 $M,\tau\models p_1,M,\tau a\models\neg p_1,M',\tau a\models p_1$ 的情形，p_1 的另一种情形的讨论与此类似。T_Ω 中关于 p_1 的后续状态公理为 $\square(\bigcirc p_1\leftrightarrow(\varphi_{ep_1}\vee\varphi_{cp_1})\vee(p\wedge\neg(\varphi_{e\neg p_1}\vee\varphi_{c\neg p_1})))$，由 $M,\tau\models p_1,M,\tau a\models\neg p_1$ 可推得：或者析取式 $\varphi_{e\neg p_1}$ 中包含子公式 $\langle a\rangle\text{true}\wedge\varphi_{a,\neg p_1}$ 且 $M,\tau\models\varphi_{a,\neg p_1}$，或者 $\varphi_{c\neg p_1}$ 包含子公式 $\alpha_{c_i,\neg p_1}\wedge\bigcirc\varphi_{c_i,\neg p_1}$ 且 $M,\tau\models\alpha_{c_i,\neg p_1},M,\tau a\models\varphi_{c_i,\neg p_1}$，前者说明 D_{ef} 中包含公理 $\square(\langle a\rangle\text{true}\wedge\varphi_{a,\neg p_1}\rightarrow\bigcirc\neg p_1)$ 且 $M',\tau\models\varphi_{a,\neg p_1}$，后者说明 D_{ca} 中包含 $\square(\alpha_{c_i,\neg p_1}\wedge\bigcirc\varphi_{c_i,\neg p_1}\rightarrow\bigcirc\neg p_1)$ 且 $M',\tau\models\alpha,M',\tau a\models\varphi_{c_i,\neg p_1}$（由 p_1 满足条件(2)可推知），以上不管哪一种情形都可推得 $M',\tau a\models\neg p_1$，与假设矛盾。

b)考虑 M 中匿名对象的变化情况。由 D_{sa} 中公理可推出，对任意 $\tau,\tau a\in prf(\sigma)$，匿名对象 $x\in\Delta^{I(\tau)}$ 和有名对象 a，概念 $C\in CL_\Omega$ 和角色 R，有 $x\in C^{I(\tau a)}$ 当且仅当 $x\in C^{I(\tau)}$，$(x,a)\in R^{I(\tau a)}$ 当且仅当 $(x,a)\in R^{I(\tau)}$，$(a,x)\in R^{I(\tau a)}$ 当且仅当 $(a,x)\in R^{I(\tau)}$，可能使得 M 不符合 Ω 的极小模型定义的情形是：存在 $\tau,\tau a\in prf(\sigma)$，角色 R 和匿名对象 $x,y\in\Delta^{I(\tau)}$，满足(1)$(x,y)\in R^{I(\tau)},(x,y)\notin R^{I(\tau a)}$ 或(2)$(x,y)\notin R^{I(\tau)},(x,y)\in R^{I(\tau a)}$。

我们可以按以下方式修改 M，构造符合极小模型定义的 Ω 的模型 $M'=(\sigma,<,I')$：I' 和 I 基本相同，只是当情形(1)出现时，我们修改 R 在 τa 点的解释，使得 $(x,y)\in R^{I'(\tau a)}$，在情形(2)出现时，我们修改 R 在 τ 点的解释，使得 $(x,y)\in R^{I'(a)}$，其他保持不变。修改后的模型 M' 仍可使得 T_Ω 可满足，不然，以对情形(1)的修改为例，假设 $(x,y)\in R^{I'(\tau a)}$ 在 τa 点违背 T_Ω 中某约束 ϕ，则 ϕ 中必然出现概念 $\neg\exists R.C$ 且 $x\in(\neg\exists R.C)^{I(\tau a)}$，$y\in C^{I(\tau a)}$。由此可知，$\exists R.C\in CL_\Omega$，再由 $y\in C^{I(\tau a)}$，$(x,y)\in R^{I(\tau)}$ 可先后推得 $y\in C^{I(\tau)}$，$x\in(\exists R.C)^{I(\tau)}$，最后得到 $x\in(\exists R.C)^{I(\tau a)}$，与 $x\in(\neg\exists R.C)^{I(\tau a)}$ 矛盾，假设不成立。

综合 a)、b)可知，Ω 有极小模型。

在对动态域建模之后，通常需要检查域描述 Ω 中定义的动作是否一致，正确的建模应保证一旦动作可执行（前提条件可满足），则执行动作后的后续状态应是良定义的。因此，如果根据域描述 Ω 进行动作推理，出现了执行某个原子动作之后，动作效果和域约束相矛盾的情况，则说明域建模出了问题，Ω 中存在动作不一致（类似讨论还可参见文献[9]）。

定义 6（动作一致）　给定域描述 Ω，a 为 Ω 中出现的动作，其效果公理集合为 D_{ef} 的子集 $D_{a,ef}$，前提公理为 $\square(\langle a\rangle\text{true}\rightarrow\varphi_a)$，称 a 关于 D_{ca}，D_{do} 动作一致当且仅当对任意满足 $D_{a,ef}\cup D_{ca}$ 的极小模型 $M=(\sigma,<,I)$ 和 $\tau,\tau a\in prf(\sigma)$ 均有：如 $M,\tau\models\bigwedge_{\square\varphi\in D_{do}}\varphi\wedge\varphi_a$，则 $M,\tau a\models\bigwedge_{\square\varphi\in D_{do}}\varphi$。

定理 2　给定 Ω，令 $\Omega'=D_{a,ef}\cup D_{ca}\cup\{\varphi\mid\square\varphi\in D_{do}\}\cup\{\varphi_a\}$，$\Omega$ 中动作 a 关于 D_{ca}，D_{do} 动作一致当且仅当 $T_{\Omega'}\models\langle a\rangle\text{true}\rightarrow\langle a\rangle(\bigwedge_{\square\varphi\in D_{do}}\varphi)$。

动作可执行性和投影是动作推理的两个基本问题，给定域描述 Ω，动作可执行性和投影定义如下：

定义 7（动作可执行性和投影）　可执行性：假设动作 a_i $(1\leqslant i\leqslant n)$ 的前提公理形式为 $\square(\langle a_i\rangle\text{true}\rightarrow\varphi_{a_i})$，称动作序列 $a_1;a_2;\cdots;a_n$ 可执行当且仅当对任意 $\tau\in prf(a_1a_2\cdots a_{n-1})$ 和满足 $D_{ef}\cup D_{ca}\cup D_{do}\cup\{\varphi_i\}$ 且 $\tau\in prf(\sigma)$ 的极小模型 $M=(\sigma,<,I)$，均有 $M,\tau\models\varphi_{a_{|\tau|+1}}$。

投影：称 $ALCIO$ 公式 φ_g 为动作序列 $a_1a_2\cdots;a$ 的一个逻辑后承当且仅当对任意满足 $a_1a_2\cdots a_n\in prf(\sigma)$ 的 Ω 的极小模型 $M=(\sigma,<,I)$ 均有 $M,a_1a_2\cdots a_n\models\varphi_g$。

对域描述 Ω，可以按以上步骤 1－6 计算出理论 T_Ω，进而将域的动作可执行问题和投影问题归结为 $DLTL_{ALCIO}$ 的公式可满足问题。

定理 3　给定域描述 Ω，动作序列 $a_1a_2\cdots a_n$ 可执行当且仅当 $T_\Omega\models\varphi_{a_1}$，$T_\Omega\models\langle a_1\rangle\text{true}\rightarrow\langle a_1\rangle\varphi_{a_2}$，…，$T_\Omega\models\langle a_1;\cdots;a_{n-1}\rangle\text{true}\rightarrow\langle a_1;\cdots;a_{n-1}\rangle\varphi_{a_n}$。

文中假定 φ_g 中不会出现在任意 Ω 元素中没有出现过的非负概念和原子，则关于动作投影，有以下定理：

定理 4　给定域描述 Ω，$ALCIO$ 公式 φ_g 为动作序列 $a_1;a_2;\cdots;a_n$ 执行后的一个投影当且仅当 $T_\Omega\models\langle a_1;\cdots;a_n\rangle\text{true}\rightarrow\langle a_1;\cdots;a_n\rangle\varphi_g$。

4　一个例子

本节将介绍一个关于某公司员工医疗保险的例子。假设智能体能执行的原子动作只有 $hireJOHN$ 和 $fireJOHN$（为简化我们省略了 nop）。我们按上节 6 个步骤建立该域的 $DLTL_{ALCIO}$ 理论，其中 PAICS 和 CPICS 分别表示中国平安保险公司和中国太平洋保险公司。

步骤 1　效果公理

$\square(\langle hireJOHN\rangle\text{true}\rightarrow\bigcirc Employee(JOHN))$

$\square(\langle fireJOHN\rangle\text{true}\rightarrow\bigcirc\neg Employee(JOHN))$

步骤 2　因果规则

$\square(\neg\exists InsuredBy.MedInsuranceCompany(JOHN)\wedge\bigcirc Employee(JOHN)\rightarrow\bigcirc InsuredBy(JOHN,PAIC))$

该因果规则描述的是：如果 JOHN 以前没有办过医疗保险，那么一旦 JOHN 成为公司员工，即为其在 PAICS 办理医疗保险。

步骤 3　前提公理

$\square(\langle hireJOHN\rangle\text{true}\rightarrow\neg Employee(JOHN))$

$\square(\langle fireJOHN\rangle\text{true}\rightarrow Employee(JOHN))$

步骤 4　域约束公理

$\square(Employee\sqsubseteq\exists InsuredBy.MedInsuranceCompany)$

$\square(\{PAIC\}\sqcup\{CPIC\}\sqsubseteq MedInsuranceCompany)$

$\square(Nom=\{JOHN\}\sqcup\{PAIC\}\sqcup\{CPIC\})$

步骤 5　初始状态描述公式

$\neg Employee(JOHN)$

我们假设在初态时，智能体只有关于 JOHN 的不完全知识，对 JOHN 是否办过医疗保险一无所知。

步骤 6　由域描述可计算出以下后续状态公理

$\Box(\bigcirc Employee(JOHN) \leftrightarrow \langle hireJOHN \rangle \text{true} \vee (Employee(JOHN) \wedge \neg \langle fireJOHN \rangle \text{true}))$

$\Box(\bigcirc InsuredBy(JOHN, PAIC) \leftrightarrow (\neg \exists InsuredBy.MedInsuranceCompany(JOHN) \wedge \bigcirc Employee(JOHN)) \vee InsuredBy(JOHN, PAIC))$

$\Box(Employee \sqcap \neg Nom = \bigcirc(Employee \sqcap \neg Nom))$

$\Box(MedInsuranceCompany \sqcap \neg Nom = \bigcirc(MedInsuranceCompany \sqcap \neg Nom))$

$\Box(\exists InsuredBy.MedInsuranceCompany \sqcap \neg Nom = \bigcirc(\exists InsuredBy.MedInsuranceCompany \sqcap \neg Nom))$

$\Box(\exists InsuredBy^{-}.\{JOHN\} \sqcap \neg Nom = \bigcirc(\exists InsuredBy^{-}.\{JOHN\} \sqcap \neg Nom))$

经以上步骤 1－5 建立的域描述记作 Ω，经步骤 6 后建立的域理论记作 T_Ω，显然 Ω 是分层的，因此根据定理 1，T_Ω 是 Ω 的等价 $DLTL_{ALCIO}$ 理论。根据定理 2 判定 Ω 中没有不一致的动作。基于 T_Ω 可以执行一些动作可执行性和动作投影的查询，如(1)“动作序列 *hireJOHNfireJOHN* 是否可执行”和(2)“执行 *hireJOHNfireJOHN* 后 JOHN 是否在 PAIC 办有医疗保险”，利用 $DLTL_{ALCIO}$ 公式可满足算法，根据定理 3，可以给出查询(1)的结果为 *hireJOHN*;*fireJOHN* 可执行，根据定理 4，查询(2)的结果为不能肯定 JOHN 在 PAIC 办有医疗保险（*InsuredBy*(*JOHN*, *PAIC*) 不是动作序列 *hireJOHN*;*fireJOHN* 的逻辑后承）。

5　相关研究

在动作推理研究领域中，有一类研究涉及动态域的时间特性的表示和推理，本文属该类研究，并受以下介绍的 *DLTL* 的影响。

DLTL[10] 是 Henriksen 等人提出的一种命题时序逻辑扩展，Henriksen 等人在 *DLTL* 中使用正规表达式标记的 until 算子 U^π 将命题动态逻辑的动态表示成分与命题时序逻辑的时态表示成分结合在一起，从而使得 *DLTL* 兼具动态和时态特性表达能力。Giordano 等人提出基于 *DLTL* 的动作理论[11]并将其应用于智能体通信验证[12]。基于 *DLTL* 的动作理论支持时间特性的表示和推理。

如引言所述，出于表达能力和计算特性综合考虑，近年来人们将描述逻辑引入动作推理领域，提出了一些可判定的、基于描述逻辑的动作形式，其中有些就明确考虑到时间特性的表示和推理。

在文献[13]中，Baader 等扩展了文献[3]中的动作形式，提出了一种可用于无限长动作序列的时间特性查询的动作形式，其中动作序列由 Büchi 自动机产生（一个 Büchi 自动机即一个计算程序），查询目标用时序描述逻辑 *ALCO-LTL* 公式表示。Baader 等证明对初态用 *ALCO-ABox* 表示，域约束用无环 *ALCO-TBox* 表示，运行在条件动作上的 Büchi 自动机来说，判断其运行时间特性的可满足性或有效性的难度在 ExpSpace 内。

在文献[14]中，常亮等引入上文中的 U^π 算子扩展动态描述逻辑 *DDL*，在其结果逻辑 TD_{ALCQIO} 中，U^π 算子用于限制 TD_{ALCQIO} 公式。TD_{ALCQIO} 具有与动态逻辑类似的语义结构，文献[12]在对其时态成分的语义处理方式上又具有分支时态逻辑的特征。常亮等在文献[14]中设计了 TD_{ALCQIO} 的公式可满足性算法并证明 TD_{ALCQIO} 具有可判定性，利用 TD_{ALCQIO} 可刻画和推理动态域的动作和时间特性。在文献[15]中，Wang 等提出 LTD_{ALCO}，LTD_{ALCO} 与 TD_{ALCQIO} 类似，同样是引入 U^π 算子扩展 *DDL*，不同之处在于 LTD_{ALCO} 中的 U^π 算子仅用于限制不含等号的简单 *ALCO* 公式，因此 LTD_{ALCO} 的表达能力相对弱一些，LTD_{ALCO} 同样有可判定性和与 TD_{ALCQIO} 类似的表达特性。Wang 等将 LTD_{ALCO} 应用于视频事件的建模和检测。

$DLTL_{ALCIO}$ 是 *DLTL* 与描述逻辑 *ALCIO* 结合的产物，$DLTL_{ALCIO}$ 保持有与 *DLTL* 类似的线性语义结构，U^π 算子可以限制 $DLTL_{ALCIO}$ 概念和公式。文中提出的 $DLTL_{ALCIO}$ 动作理论可用于动态域的时间特性推理，不过，与文献[13]不同，本文动作推理主要是面向可由正规程序生成的动作序列的推理，而且，与文献[13-15]不同，文中动态域知识完全用逻辑公理（公式）描述，采用与情景演算[1,7,8]类似的方式解决框架和条件问题。

结束语　本文提出了基于 $DLTL_{ALCIO}$ 公式形式建立动态域描述以及构造与之等价的 $DLTL_{ALCIO}$ 理论的方法，利用该方法可将非单调、关于域描述的推理问题转化为单调的关于 $DLTL_{ALCIO}$ 理论的推理问题，从而使得动作推理问题最终可以用 $DLTL_{ALCIO}$ 可满足算法求解。

文中的动作理论构造方法还存在一些不足之处。对包含 *Q* 算子的 $DLTL_{DL}$，由于暂时找不到合适方法用后续状态公理表达动作（极小变化）语义，本文的动作理论构造方法还无法进一步推广，也许对基于 $DLTL_{ALCQIO}$ 等的动作形式，还需要设计专门的推理算法。

此外，对于动作推理中的 3 大问题：框架、分支和条件问题，文中方法主要考虑了其中前两个问题，对条件问题暂时还没有好的解决思路，这还有待以后进一步研究。

参 考 文 献

[1] Reiter R. Knowledge in action: logical foundations for specifying and implementing dynamical systems[M]. Cambridge University Press, 2001

[2] Thielscher M. FLUX: a logic programming method for reasoning agents[J]. Theory and Practice of Logic Programming, 2005, 5(4/5): 533-565

[3] Baader F, Lutz C, Milicic M, et al. A description logic based approach to reasoning about web services[C]//Proceedings of the WWW 2005 Workshop on Web Service Semantics (WSS2005). Chiba City, Japan, 2005: 10-14

[4] 常亮，陈立民. 基于动态描述逻辑 DDL 的动作理论[J]. 计算机科学，2011，38(7)：203-208

[5] 孙永新，赵希顺，符志强. 描述逻辑的动态时序扩展[J]. 计算机应用研究，2012，29(2)：536-541

[6] Baader F, Sattler U. An overview of tableau algorithms for description logics[J]. Studia Logica, 2001, 69(1): 5-40

（下转第 238 页）

结束语 传统的在线最小二乘策略迭代方法对在线生成的样本仅利用一次，用完后就丢弃，其经验利用率低、收敛速度慢。本文将批量更新方法与在线最小二乘策略迭代方法相结合，提出了一种批量最小二乘策略迭代算法。该算法以四元组(x,u,r,x')的形式保存在线生成的样本数据，并多次重复利用这批样本。重复利用的过程中，尽管用的都是同一批四元组形式的样本(x,u,r,x')，但更新策略参数用的是五元组$(x,u,r,x',h(x'))$，而每次更新用的策略h并不相同，因此同一个样本可以用来多次更新策略参数，提高了经验利用率。使用LSTD-Q评估策略的BLSPI算法在处理完一批样本后才更新策略参数，性能较稳定；而采用LSPE-Q评估策略的BLSPI算法每处理一个样本都更新策略参数，并且依赖于上一轮的评估结果，是一个增量式算法，通过给定较好的初始值并逐步减小步长参数也可以取得较好的性能。由于求解策略参数需要进行矩阵求逆运算，其计算复杂度较高，计算量较大，在算法中应用了Sherman-Morrison公式，消去了矩阵求逆运算，降低了计算复杂度，减少了计算量。倒立摆实验表明，重复利用以前的经验，可以提高经验利用率，加快收敛速度，并且重复次数越多，效果越明显，但计算量也越大，适量增大批量更新间隔可以在保证算法一定性能的基础上降低计算量，因此应根据具体情况选择合适的算法参数。

当状态划分比较细致、径向基函数维度比较大时，最小二乘方法将会带来巨大的计算量。如何减少计算量，提高计算效率及如何更加高效地筛选、利用经验知识是下一步研究的方向。另外，本文考虑的是离散动作空间的情况，对连续动作空间的研究也是下一步研究的课题。

参 考 文 献

[1] Sutton R S, Barto A G. Reinforcement learning: An introduction [M]. Cambridge: MIT Press, 1998

[2] 刘全，闫其粹，伏玉琛，等. 一种基于启发式奖赏函数的分层强化学习方法 [J]. 计算机研究与发展，2011，48(12)：2352-2358

[3] Kaelbing L P, Littman M L, Moore A W. Reinforcement learning: A survey [J]. Journal of Artificial Intelligence Research, 1996, 4(2): 237-285

[4] 刘全，傅启明，龚声蓉，等. 最小状态变元平均奖赏的强化学习方法 [J]. 通信学报，2011，32(1)：66-71

[5] Gao Yang, Chen Shi-fu, Lu Xin. Research on reinforcement learning technology: A review [J]. Journal of Acta Automatica Sinica, 2004, 30(1): 86-100

[6] Geist M, Pietquin O. Parametric value function approximation: A unified view [C] // Proc of the IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning. NJ: IEEE, 2011: 9-16

[7] Bradtke S J, Barto A G. Linear least-squares algorithms for temporal difference learning[J]. Journal of Machine Learning, 1996, 22: 33-57

[8] Boyan J. Technical update Least-squares temporal difference learning [J]. Journal of Machine Learning, 2002, 49: 233-246

[9] Maei H R, Szepesvari C, Bhatnagar S, et al. Toward off-policy learning control with function approximation [C] // Proc of the 27th International Conference on Machine Learning. Haifa: Omnipress, 2010: 719-726

[10] Sutton R S. Learning to predict by the method of temporal differences [J]. Journal of Machine Learning, 1988, 22: 33-57

[11] Sutton R S, Szepesvari Cs, Maei H R. A convergent O(n) algorithm for off-policy temporal-difference learning with Linear function approximation[C] // Proc of the 25th Annual Conference on Neural Information Processing Systems. Granada, 2008: 1609-1616

[12] Lagoudakis M, Parr R, Littman M. Least-squares methods in reinforcement learning for control[J]. Methods and Applications of Artificial Intelligence, 2002, 2308: 249-260

[13] Lagoudakis M, Parr R. Least squares policy iteration [J]. Journal of Machine Learning Research, 2003(4): 1107-1149

[14] Busoniu L, Babuska R, Schutter B D, et al. Reinforcement Learning and Dynamic Programming using Function Approximators [M]. New York: CRC Press, 2010

[15] Kalyanakrishnan S, Stone P. Batch reinforcement learning in a complex domain[C] // Proc of the 6th International Conference on Autonomous Agents and Multiagent Systems. New York, 2007: 650-657

[16] Jung T, Polani D. Kernelizing LSPE (λ) [C] // Proc of the IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning. NJ: IEEE, 2007

[17] Jung T, Polani D. Least squares SVM for least squares TD learning[C] // Proc of the 17th European Conference on Artificial Intelligence. Riva del Garda, 2006: 499-503

(上接第 214 页)

[7] Lin F, Reiter R. State constraints revisited[J]. Journal of Logic and Computation, 1994, 4(5): 655-677

[8] Lin F. Embracing causality in specifying the indirect effects of actions[C] // Proceedings of the Thirteenth National Conference on Artificial Intelligence. Oregon, Portland: AAAI Press, 1996 (1): 670-676

[9] Baader F, Lippmann M, Liu H. Using causal relationships to deal with the ramification problem in action formalisms based on description logics [C] // Proceedings of the 17th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning. Berlin, Heidelberg: Springer-Verlag, 2010: 82-96

[10] Henriksen J G, Thiagarajan P S. Dynamic linear time temporal logic[J]. Annals of Pure and Applied Logic, 1999, 96(1-3): 187- 207

[11] Giordano L, Martelli A, Schwind C. Reasoning about actions in dynamic linear time temporal logic[J]. Logic Journal of the IGPL, 2001, 9(2): 273-288

[12] Giordano L, Martelli A, Schwind C. Specifying and verifying interaction protocols in a temporal action logic[J]. Journal of Applied Logic, 2007, 5(2): 214-234

[13] Baader F, Liu H, Mehdi A. Verifying properties of infinite sequences of description logic actions [C] // ECAI 2010: 19th European Conference on Artificial Intelligence Proccedings. Lisbon, Portugal: IOS Press, 215, 2010: 53-58

[14] 常亮，史忠植，古天龙，等. 可判定的时序动态描述逻辑[J]. 软件学报，2011，22(7)：1524-1537

[15] Wang Xiao-feng, Chang Lang, Li Zhi-xin, el al. A dynamic description logic based system for video event detection[J]. Frontiers of Electrical and Electronic Engineering in China, 2010, 5 (2): 137-142