

一种支持自治计算的基于可废止逻辑的柔性 Agent *

廖备水¹, 黄华新¹, 高济²

¹浙江大学 语言与认知研究中心, 浙江 杭州 310028)

²(浙江大学 计算机学院, 浙江 杭州 310027)

A Defeasible Logic-Based Flexible Agent for Autonomic Computing

LIAO Bei-Shui¹⁺, HUANG Hua-Xin¹, GAO Ji²

¹(Center for the Study of Language and Cognition, Zhejiang University, Hangzhou 310028, China)

²(College of Computer Science, Zhejiang University, Hangzhou 310027, China)

+ Corresponding author: Phn: +86-571-88273089, E-mail: baisesliao@zju.edu.cn

Liao BS, Huang HX, Gao J. A defeasible logic-based flexible agent for autonomic computing. *Journal of Software*, 2008,19(3):605-620. <http://www.jos.org.cn/1000-9825/19/605.htm>

Abstract: In the context of autonomic computing, by taking advantage of the non-monotonic knowledge representation and reasoning mechanisms of defeasible logic, a flexible Agent model is proposed, which is capable of accepting the real-time rule modifications, flexibly handling the run-time rule conflicts, and providing efficient non-monotonic reasoning functions. The flexible agent is both autonomous and controllable, and is able to cooperate with other Agents via contracts in an open and dynamic environment.

Key words: Agent; autonomic computing; defeasible logic; non-monotonic reasoning; distributed system management

摘要: 以自治计算的研究为背景,利用可废止逻辑理论的非单调知识表征和推理机制,提出一种能够动态接受规则变更、灵活处理实时发生的规则冲突,并进行高效的非单调推理的柔性 Agent 模型.这种 Agent 既是自主的,又是可控的,而且可以在开放、动态的环境中通过合同与其他 Agent 进行协同工作.

关键词: Agent;自治计算;可废止逻辑;非单调推理;分布式系统管理

中图法分类号: TP18 文献标识码: A

Agent 又称为智能主体或智能代理,具有反应性、自治性和社会性等特点,已被普遍认为是支持大规模、开放和分布的信息系统实现动态服务集成和协同工作的关键技术^[1-3].典型的 Agent 系统基于 BDI(信念-愿望-意图)逻辑^[4].该逻辑采用信念、愿望和意图这三类意识态度来刻画 Agent 的结构,并在此基础上描述其抽象性质和推理过程.目前,基于 BDI 逻辑的 Agent 已经得到了广泛而深入的研究,在个体 Agent 模型、多 Agent 协商、协同与协作等领域都有丰富的研究成果.但是,传统的基于 BDI 逻辑的 Agent 模型在体系结构和推理决策机制

* Supported by the National Natural Science Foundation of China under Grant No.60773177 (国家自然科学基金); the China Postdoctoral Science Foundation under Grant No.20060400316 (中国博士后科学基金); the National Social Science Foundation of China under Grant No.04BZX045 (国家社会科学基金)

Received 2006-11-26; Accepted 2007-01-24

上仍然存在一些重大的缺陷:

(1) 传统 Agent 的信念修正(包括信念的扩充、修正和收缩等)是 Agent 基于对环境和自身的认识而对当前知识进行修改,并保持信念的一致性,面向的是 Agent 的信息部分.而对于动机部分,即 Agent 对动作选择的知识(规则)是静态确定的,不能动态变更.因此,Agent 不能接受外部动机(如来自 IT 管理者的动态管理策略和目标),并依此调整其对动作的选择.

(2) 尽管包括基于合同网在内的协作机制为 Agent 之间的协作奠定了基础,但是对个体 Agent 而言,现有的 BDI 逻辑不支持 Agent 之间动态协作关系的显式表示和推理,即尚未考虑在引入新的规则(反应协作需求)后,如何对其行为的选择进行调整.

(3) 传统的推理过程本质上是单调的,不具备非单调的推理能力.

(4) 传统 Agent 没有讨论多种动机来源,因此也缺乏各种动机之间的冲突处理机制.

这些缺陷从根本上限制和阻碍了它在新领域的应用.本文以自治计算的研究为背景,针对现有 Agent 模型的不足,以 BDI Agent 理论为基础,采用具有非单调推理能力的可废止逻辑理论作为知识表征和推理的工具,提出一种经过革新的基于规则的柔性 Agent 模型.

1 应用背景和相关工作

与本文相关的应用背景是“自治计算(autonomic computing)”^[5].它作为一个新的热点研究领域,于 2001 年 10 月由 IBM 公司正式发起,旨在克服阻碍未来 IT 工业进步的主要障碍,即目前已隐约可见的软件复杂性危机.随着计算机技术和网络技术的发展,大型分布式信息系统的管理复杂性问题已经成为其发展过程中的瓶颈.为了使系统正常运行,IT 企业通常要花费高出设备成本 4~20 倍的管理费用,而且单靠 IT 专家和技术人员的努力,越来越难以驾驭开放、动态和异构的信息系统.在这种背景下,以降低分布式系统管理复杂性问题为目标的自治计算正在成为当前的研究热点.自治计算的思想受启发于人体自主神经系统的管理机制:人的心律、血压和体温等由自主神经系统自行控制,不需要人脑的直接参与,使得人脑可以用于解决宏观的问题.与之相仿,我们希望具备自治计算特性的信息系统及其各个部分拥有自我配置、自我修复、自我优化和自我保护等自我管理功能,尽量减少人工的参与,以大幅度降低系统的管理费用.

自治计算系统由自治元素构成,每个自治元素作为自主工作的实体,在 IT 管理者的动态管理策略和目标的指导下,自主地管理内部资源,并依据一定的方式参与协作,实现整个系统的自我管理.由于 Agent 具有反应性、自治性和社会行为能力等特点,刚好与自治元素的主要特征相吻合,它被认为是担任自治元素、支持信息系统实现自治计算的理想候选^[6-8].然而,现有的 BDI Agent 的上述固有缺陷,限制和阻碍了它在自治计算系统中的应用:

缺陷(1)使得 Agent 只能根据系统的设计目标和当前的环境状态进行推理决策,而不能动态接受来自 IT 管理者的动态管理策略和目标.这对于自治计算元素来说是重大的缺陷之一,因为 IT 管理者一方面希望系统能够根据环境的状态自主决策和运作,另一方面又需要根据当前的商业目标和系统状态制定适当的管理策略,以调整 Agent 的行为从而满足当前的目标.也就是说,希望 Agent 既是自主工作的实体,又是柔性的和可控的.

缺陷(2)使得 Agent 在开放、动态的 Internet 环境中通过合同进行协作时,难以合适地处理内部工作动机与外部协作动机之间的关系(包括动机冲突).

Agent 推理过程的单调性以及缺乏动机的冲突处理机制,使其在根本上不支持行为选择知识(即目标规则)的变更,这是因为这种变更可能造成规则的冲突,使得动机之间相互抵触.

针对自治计算系统的特性需求以及现有 BDI Agent 的上述缺陷,廖备水等人^[9-12]提出了一种初步满足自治元素基本特性的扩展 BDI Agent 模型.与传统 BDI Agent 不同的是,它把基于政策的管理方法和基于合同的协作机制有机地与个体 Agent 的知识表征和推理决策机制结合起来,以便能够接受和处理来自 IT 管理者的管理目标和策略以及与其他 Agent 因协作需要而签订的合同,并在此基础上依据环境信息,进行自主决策和运作.由政策(policy)和合同(contract)引起的行为动机用道义逻辑来表示,分别称为政策型义务和合同型义务;同时,由环境

信息引起的行为动机对应于传统 BDI Agent 的愿望(desire).这样,Agent 的目标就不仅由内部愿望所决定,还应考虑政策和合同所产生的义务,因此,把这种 Agent 简称为 PDC-Agent^[9,12].三类行为动机的冲突处理基于优先级的方法:当同类动机之间或不同类动机之间产生冲突时,选择优先级较高者.关于 PDC-Agent 的进一步阐述,详见相关文献[9-12].

以 PDC-Agent 为自治元素,廖备水等人^[12]提出了一个面向服务的自治计算系统模型.该模型由资源层、服务层和自我管理层组成.资源层包含各种资源,诸如计算、存储、应用软件和数据库等;服务层通过 Web 服务(或网格服务)基础设施提供的标准接口和协议,实现异构资源的共享和互操作.自我管理层(如图 1 所示)是系统的核心,实现在 IT 管理者指导下的系统自我管理功能.它由政策管理框架和可以在政策指导下进行资源自主管理和动态协同工作的 PDC-Agent 组成.其中,政策用于表示来自 IT 管理者的管理策略和目标.政策分为抽象政策(对应于 IT 管理者的管理策略和目标)和可实行政策,它们都由通用的形式语言加以表征,并通过细化、映射和部署机制,把抽象政策转化为具体的可实行政策后,指派给特定的 Agent.政策可以在系统的运行过程中动态变更,并以此来配置 Agent 的行为,从而在不改变底层软件编码的条件下,动态调整 Agent 的行为.关于政策管理框架以及面向服务的自治计算系统的详细内容请见文献[12,13].

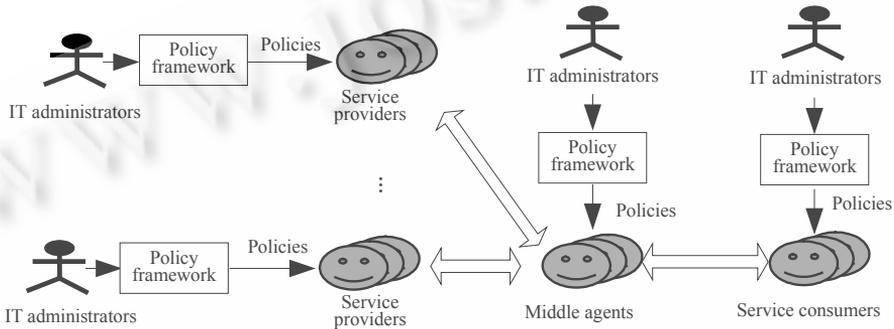


Fig.1 A multi-agent based self-management layer of autonomic computing system

图 1 基于多 Agent 的自治计算系统自我管理层

上述的 PDC-Agent 模型阐明了用政策和合同来扩展传统 BDI Agent 的基本原理,但尚未深入研究并阐明在引入动态规则后,如何有效解决各种规则之间的冲突问题,以及 Agent 推理过程的非单调性问题.本文将从该问题入手,利用 Governatori 和 Dastani 等人提出的运用可废止逻辑进行 Agent 建模的思想^[14],阐明一种能够动态接受规则变更、灵活处理实时发生的规则冲突,并进行高效的非单调推理的柔性 Agent 模型.我们把这种 Agent 称为“柔性 Agent”,意指其用于行为选择的知识(即目标规则)可以动态变更,并具备灵活地适应这种变更的有效机制.

2 柔性 Agent 体系结构

柔性 Agent 的体系结构如图 2 所示,它包含 3 个主要功能模块:信念修正、目标产生和意图产生.(1) 信念修正模块依据环境信息对信念进行扩充、修正或收缩,在修改信念的同时保持了信念的一致性.因此,该模块在功能上与传统的 BDI Agent 的类似;(2) 目标产生模块是柔性 Agent 区别于其他 Agent 的核心部分.在这里,目标的产生来源于三方面的动机:愿望、由政策产生的义务和由合同产生的义务,三者组成了 3 类目标;(3) 意图产生模块依据当前信念和目标,产生意图以及与该意图对应的规划.上述 3 个模块的推理逻辑均表示为规则,分别来自于下列规则集:信念规则集(R^B)、合同规则集(R^C)、愿望规则集(R^D)、政策规则集(R^P)和意图规则集(R^I).其中,合同规则集、愿望规则集和政策规则集统称为目标规则集,记作 R^G ;另外, R^C 和 R^P 可以动态变更,图中的合同维护和政策维护模块分别用于合同规则和政策规则的添加与删除.

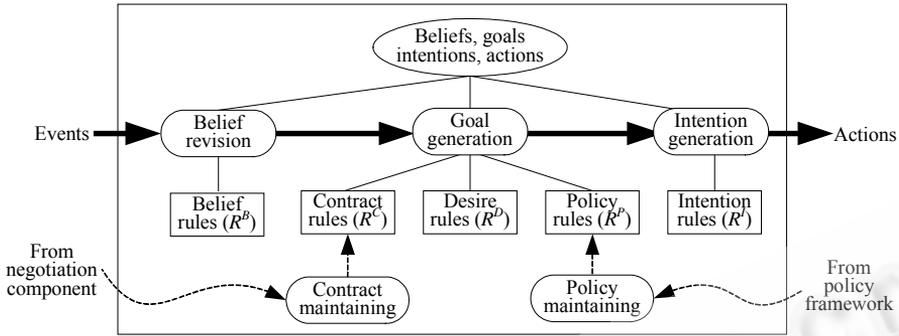


Fig.2 The architecture of defeasible logic based flexible agent

图2 基于可废止逻辑的柔性 Agent 体系结构

合同规则和政策规则可以动态变更,这不仅有可能引起规则的冲突,而且使得 Agent 的推理过程呈现出非单调的特征.为了满足该要求,需要引入非单调逻辑来替换传统 BDI 中的经典谓词逻辑.在非单调推理方面,现有的理论包括限定推理、缺省逻辑、自认识逻辑^[15]和可废止逻辑^[16]等.相对而言,可废止逻辑的突出优点是它具有线性的计算复杂性,而其他非单调推理系统都具有很高的计算复杂性^[17].因此,可废止逻辑理论虽然表达能力有限,但具备简单、高效和易于实现等突出优点,近年来得到了快速的发展.另外,在可废止理论中,基于优先级的规则冲突处理机制恰好为柔性 Agent 的规则变更提供了天然的支持.鉴于此,本文将可废止逻辑作为柔性 Agent 表征和推理的基本工具.

3 基于可废止逻辑的柔性 Agent 模型

3.1 预备知识:可废止逻辑

可废止逻辑于 1987 年由 Nute 首先提出^[16],它不像缺省逻辑或限定推理^[15]那么有名,不过近年来却很受重视,这主要得益于它所具备的两个突出优点^[18]:1) 可计算性:在多项式的时间里得出结论;2) 有效的规则冲突处理机制:有一套内嵌的优先级处理手段.可废止推理与缺省推理很相似,不同点在于对规则的使用.对于缺省推理,如果某条规则的所有前提均满足,结论就成立;而对于可废止推理,这个结论则可以被其他规则推翻.

一个可废止理论^[19]由 5 种不同的知识组成:事实(fact)、硬性规则(strict rule)、可废止规则(defeasible rule)、废止者(defeater)和规则上的优越关系(superiority relation).事实是不可辩驳的陈述,例如,“计算是资源”可表示为:资源(计算).规则由前提(身体)、箭头符号(“ \rightarrow ”,“ \Rightarrow ”和“ \sim ”分别表示硬性规则、可废止规则和废止者)以及结论(头)组成,其中,前提是一组文字的集合,头是一个文字.硬性规则与传统的规则意义相同,即只要前提是不可辩驳的,那么结论也是.例如,“资源具有属性”可表示为:资源(X) \rightarrow 有属性(X);可废止规则是指那些可以被反面证据废止的规则,例如“典型地,资源可以被使用”,资源(X) \Rightarrow 可使用(X),意指一般情况下资源可以使用,除非有证据表明其不可使用;废止者是指那些不能用于得出任何结论,只用于防止某些结论发生的规则,例如“未通过认证的用户,可能不可使用资源”,可表示为:未认证(X) \sim 可使用(X),意指“未通过认证的用户”这个证据不足以得出其“不能使用资源”的结论,而只是其“可能不能使用资源”的证据;规则之间的“优越关系”用于定义规则之间的优先级,即某条规则可以推翻另一条规则的结论.例如,给定两条规则:

$$r_1: \text{资源}(X) \Rightarrow \text{可使用}(X); r_2: \text{异常}(X) \Rightarrow \neg \text{可使用}(X).$$

这两条规则相互抵触,不能得出任何结论.但如果引入优越关系“ $>$ ”,且 $r_2 > r_1$,那么,我们可以得出该资源不可使用的结论.优越关系必须是非循环的.

这样,可废止理论 T 表示为三元组 $(F, R, >)$,其中, F 是事实, R 是规则的有限集合, $>$ 是 R 上的优越关系.作为一种非单调的方法,可废止逻辑采用怀疑论的观点来构造结论,即结论不是定义为外延或问答集的交集. T 的结论是一个带标签的文字,并有如下 4 种形式:

- + Δq :指 q 在 \mathcal{T} 中是明确可证的(即只使用事实和硬性规则);
- Δq :指我们已经证明 q 在 \mathcal{T} 中是明确不可证的;
- + ∂q :指 q 在 \mathcal{T} 中是以可废止的方式可证的;
- ∂q :指我们已经证明 q 在 \mathcal{T} 中是以可废止的方式不可证的。

一个证明是 \mathcal{T} 中的一个推导,它是由推理规则构造的带标签文字的有限序列 $P=(P(1),\dots,P(n))$,对应于上述 4 种结论,有 4 种推理规则分别说明推导是如何展开的(如图 3 所示)。图中, R_s, R_{sd}, R_d 和 R_{df} 分别表示给定规则 R 中的所有硬性规则集合、所有硬性和可废止规则的集合、所有可废止规则以及所有废止者集合; $R[q]$ 表示 R 中结论为 q 的规则集合; $P(1\dots i)$ 表示序列 P 的长度为 i 的初始部分。

- + Δ : We may append $P(i+1)=+\Delta q$ if either
 - $q \in F$ or
 - $\exists r \in R_s[q] \forall a \in A(r): +\Delta a \in P(1\dots i)$
- Δ : We may append $P(i+1)=-\Delta q$ if
 - $q \notin F$ and
 - $\forall r \in R_s[q] \exists a \in A(r): -\Delta a \in P(1\dots i)$
- + ∂ : We may append $P(i+1)=+\partial q$ if either
 - (1) $+\Delta q \in P(1\dots i)$ or
 - (2) (2.1) $\exists r \in R_{sd}[q] \forall a \in A(r): +\partial a \in P(1\dots i)$ and
 - (2.2) $-\Delta \sim q \in P(1\dots i)$ and
 - (2.3) $\forall s \in R[\sim q]$ either
 - (2.3.1) $\exists a \in A(s): -\partial a \in P(1\dots i)$ or
 - (2.3.2) $\exists t \in R_{sd}[q]$ such that
 - $\forall a \in A(t): +\partial a \in P(1\dots i)$ and $t > s$
- ∂ : We may append $P(i+1)=-\partial q$ if
 - (1) $-\Delta q \in P(1\dots i)$ and
 - (2) (2.1) $\forall r \in R_{sd}[q] \exists a \in A(r): -\partial a \in P(1\dots i)$ or
 - (2.2) $+\Delta \sim q \in P(1\dots i)$ or
 - (2.3) $\exists s \in R[\sim q]$ such that
 - (2.3.1) $\forall a \in A(s): +\partial a \in P(1\dots i)$ and
 - (2.3.2) $\forall t \in R_{sd}[q]$ either
 - $\exists a \in A(t): -\partial a \in P(1\dots i)$ or $t \succ s$

Fig.3 Inference rules for conclusions of a defeasible theory

图 3 可废止理论的结论推理规则

第一,为了证明+ Δq ,我们必须仅使用事实和硬性规则来建立一个证明,这是传统意义上的一个演绎。第二,为了证明- Δq ,必须说明所有证明+ Δq 的努力都是失败的,因此,用于- Δ 的推理规则是用于+ Δ 的推理规则的结构上的补。第三,由于可废止逻辑的可废止特性,可废止结论的推理规则要更加复杂,即为了说明 q 是以可废止的方式可证明的,我们两种选择:(1) 说明 q 已经是明确可证的;或者(2) 使用 \mathcal{T} 的可废止部分来证明。对于后者,我们不仅需要有一个可以应用的头为 q 的硬性规则或可废止规则(2.1),而且还要考虑可能的“攻击”,即支持 $\sim q$ 的推理链。详细地说,为了可废止地证明 q ,我们必须说明 $\sim q$ 是明确不可证的(2.2),同时,对于那些头为 $\sim q$ 的所有规则,它们或者是不可证的,或者被优先级更高的规则所反驳(2.3)。第四,用于- ∂ 的推理规则是用于+ ∂ 的推理规则的结构上的补,具体解释略。

3.2 柔性Agent理论

柔性 Agent 是一种基于可废止逻辑的 Agent,其知识分为事实和规则两部分。与传统 Agent 类似,它的意图也通过规划来实现。下面,首先给出用于表征柔性 Agent 知识的文字、规则以及用于实现意图的规划;然后给出柔性 Agent 理论的定义。

假设 $Prop=\{p,q,\dots\}$ 是原子命题集合, $\{r_1,r_2,\dots\}$ 是规则标签集合, $\mathcal{F}=\{t_1,t_2,\dots\}$ 是离散时间点的全序集合,用 $t_2 > t_1$ 表示 t_2 在 t_1 之后, $M=\{BEL,P-OBL,DES,C-OBL,INT\}$ 为模态算子集合(分别表示信念算子、由政策引起的义务算子、愿望算子、由合同引起的义务算子和意图算子), $\{\rightarrow,\Rightarrow,\rightsquigarrow\}$ 为规则箭头符号集合, $Act=\{\alpha,\beta,\gamma,\dots\}$ 为柔性

Agent 动作符号集合, $\Pi = \{\pi_1, \pi_2, \dots\}$ 为规划集合, “ \neg ” 为符号非.

定义 1 (文字).

- (1) 如果 $p \in Prop$, 那么 p 和 $\neg p$ 是平凡文字, 记作 l ;
- (2) 如果 l 是平凡文字, $t \in \mathcal{T}$ 是时间点, 则 $l:t$ 是时间平凡文字;
- (3) 如果 l 是时间平凡文字, $X \in M$ 是模态算子, 那么 Xl 和 $\neg Xl$ 是模态文字;
- (4) 文字包括平凡文字、时间平凡文字和模态文字.

文字的补用文字之前加符号“ \sim ”表示, 即如果 l 是文字, 则它的补表示为 $\sim l$, 因此, 若 $l=p$, 则 $\sim l = \neg p$; 若 $l = \neg p$, 则 $\sim l = p$. 文字也称为公式.

定义 2 (规则).

在可废止理论中, 规则形如 $r:A(r) \rightarrow C(r)$, 其中, r 为规则标签; $A(r)$ 为规则的前提, 也称为规则的身体, 它是一组文字的有限集合; 箭头符号 $\rightarrow \in \{\rightarrow, \Rightarrow, \rightsquigarrow\}$, 分别表示硬性规则、可废止规则和废止者; $C(r)$ 为规则的结论, 也称为规则的头. 本文区分下列不同类型的规则:

- (1) 设 A 是文字的有限集, C 是文字, 则 $A \rightarrow_{BEL} C$ 是信念规则, 式中, $\rightarrow_{BEL} \in \{\rightarrow_{BEL}, \Rightarrow_{BEL}, \rightsquigarrow_{BEL}\}$;
- (2) 设 A 是文字的有限集, C 是时间平凡文字, 则 $A \rightarrow_{GOAL} C$ 是目标规则, 其中, $GOAL \in \{P-OBL, DES, C-OBL\}$, $\rightarrow_{GOAL} \in \{\rightarrow_{GOAL}, \rightsquigarrow_{GOAL}\}$, 即目标规则不含硬性规则;
- (3) 设 ϕ_1, \dots, ϕ_n 是平凡文字或时间平凡文字, $GOAL \psi:t$ 是目标, 其中, $GOAL \in \{P-OBL, DES, C-OBL\}$, 则 $\phi_1, \dots, \phi_n: GOAL \psi:t \Rightarrow_{INT} \psi:t$; π 是意图规则, 其中 π 是规划 (见定义 3). 意图规则只有可废止的形式.

如果 R 是规则集, 则 R_s, R_{sd}, R_d, R_{df} 和 $R[q]$ 的含义与第 3.1 节介绍的相同; R^B, R^P, R^D, R^C 和 R^I 与第 2 节规定的相同. 对于意图规则 r , 用 $A^B(r)$ 表示集合 r 的信念条件, 用 $A^G(r)$ 表示 r 的目标条件 (意图规则的定义参考了文献 [14]), 用 $C^I(r)$ 表示 r 的意图结论, 用 $C^P(r)$ 表示 r 的规划结论. 于是, 我们用 $R^I[C^I(r) = \psi:t]$ 表示 R^I 中意图结论为 $\psi:t$ 的规则集合. 规则中可以出现自由变量, 带自由变量的规则解释为规则模式, 它是所有规则实例的集合.

定义 3 (规划).

- (1) 如果 $\alpha \in Act$ 是动作, 那么 α 是基本动作规划;
- (2) 如果 l 是平凡文字或时间平凡文字, 那么 $l?$ 是测试动作规划;
- (3) 如果 π 和 π' 是规划, 那么 π, π' (顺序), $\pi \pi'$ (选择), $\pi | \pi'$ (并列), π^* (重复) 也是规划, 称为合成规划;
- (4) 规划包括基本动作规划、测试动作规划、合成规划和空规划 (ϵ).

定义 4. 柔性 Agent 理论是一个十元组: $M = (\mathcal{T}, F, \Theta, R^B, R^P, R^D, R^C, R^I, Act, >)$, 其中:

- $\mathcal{T} = \{t_1, t_2, \dots\}$ 是离散时间点的全序集合;
- F 为事实集, 是指 Agent 对环境的观察信息集, 表示为平凡文字或时间平凡文字;
- $\Theta = (Bel, Obl_p, Des, Obl_C, Int)$ 是指 Agent 的心智状态, 包括当前的信念集 (Bel)、由政策规则产生的义务集 (Obl_p)、愿望集 (Des)、由合同规则产生的义务集 (Obl_C) 和意图集 (Int), 它们表示为模态文字;
- R^B 是信念规则集, 即 Agent 关于世界的理论;
- R^P, R^D 和 R^C 属于目标规则集, 分别表示政策规则、愿望规则和合同规则;
- R^I 是意图规则, 用于产生意图和相应的规划;
- $>$ 是规则上的优越关系, $> \subseteq (R^B \cup R^P \cup R^D \cup R^C \cup R^I)^2$. 优越关系满足传递性;
- Act 为原子动作集, 表示为平凡文字.

3.3 柔性 Agent 基本模块工作原理

柔性 Agent 模型基于可废止理论, 其主要模块的工作机制均在该理论之下进行, 包括信念修正、目标产生和意图产生. 其中, 柔性 Agent 的信念修正可以基于传统的信念修正理论 (AGM 理论), 包括 3 种信念变化形式: 扩充、修正和收缩. 但是, 与 AGM 理论不同, 柔性 Agent 的信念修正是建立在可废止逻辑理论的基础之上的, 因此, 必须对 AGM 假设和信念修正算子进行相应的革新. 与经典命题的逻辑推论相对应, 在可废止理论之下, 信念可以作为结论被推导出来, 称为信念产生. 由于信念修正是一个相对独立的部分, 且比较复杂, 本文只介绍信念的

产生,而略去信念修正这一部分,相似的工作可见相关文献[20].

本文的重点在以下 3 个方面:目标规则的动态变更、规则的冲突处理以及基于可废止理论的信念产生、目标产生和意图产生.下面分别加以叙述.

3.3.1 规则的动态变更

柔性 Agent 的目标规则包括政策规则、愿望规则和合同规则.其中,愿望规则是静态设计的,而政策规则和合同规则可以动态变更,具体定义如下:

定义 5. 目标规则的动态变更是指柔性 Agent 在运行的过程中,政策规则和合同规则及其优越关系的改变,分别建模为两个函数(δ_p 和 δ_c),对应于政策维护模块和合同维护模块(如图 2 所示)的功能:

$$\delta_p:(R_a^p, \succ_a^p):t_i \mapsto (R_b^p, \succ_b^p):t_j \tag{1}$$

$$\delta_c:(R_a^c, \succ_a^c):t_i \mapsto (R_b^c, \succ_b^c):t_j \tag{2}$$

式中, a 和 b 为自然数且 $b > a$; $t_i, t_j \in \mathcal{T}$ 且 $t_j \geq t_i$; R_a^p 和 \succ_a^p 是变更之前的政策规则及其上的优越关系, R_a^c 和 \succ_a^c 是变更之前的合同规则及其上的优越关系, R_b^p 和 \succ_b^p 是变更之后的政策规则及其上的优越关系, R_b^c 和 \succ_b^c 是变更之后的合同规则及其上的优越关系.由于二者类似,下面讨论时略去上标.函数 δ_p (或 δ_c)由两个操作来实现:添加和删除规则.当进行规则变更时,可能存在规则冲突,因此,必须依据特定的要求对规则的优越关系进行适当的更改.下面首先定义冲突规则集的概念.

定义 6. 设在时刻 $t \in \mathcal{T}$ 有某规则集 R 和待添加或删除的规则 r , $C(r)$ 是 r 的结论,那么, R 与 r 在时刻 t 的冲突规则集为 R 中结论为 $\sim C(r)$ 的规则集合,记作 $C(R, r, t) = R[\sim C(r)]$.

用符号 $+_t$ 和 $-_t$ 分别表示规则添加算子和删除算子,其中, $t \in \mathcal{T}$; 用符号 $>^A$ 表示需要添加或删除的规则优越关系集,用 $>^T$ 表示优越关系 $>$ 上的传递闭包.下面给出规则添加和删除操作的定义.

定义 7. 设 $(R, >):t$ 为时刻 t 某规则集及其上的优越关系,在时刻 $t'(t' \geq t)$ 将一条规则 r 加入 $(R, >):t$ 中,称为规则的添加操作,形式化表示为

$$(R, >):t +_{t'} r = \begin{cases} (R \cup \{r\}, >):t', & C(R, r, t) = \emptyset \\ (R \cup \{r\}, > \cup >^A):t', & C(R, r, t) \neq \emptyset \end{cases} \tag{3}$$

式中, $>^A$ 必须满足下列条件:

(1) $>^A \subseteq C(R, r, t) \times \{r\}$.

(2) $\forall \text{sup}_1, \text{sup}_2 \in (> \cup >^A)^T$, 如果 $\text{Left}(\text{sup}_1) = \text{Right}(\text{sup}_2)$, 则 $\text{Left}(\text{sup}_2) \neq \text{Right}(\text{sup}_1)$, 其中, $\text{sup}_i (i=1 \text{ 或 } i=2)$ 表示优越关系, $\text{Left}(\text{sup}_i)$ 和 $\text{Right}(\text{sup}_i)$ 分别表示优越关系 sup_i 的左边和右边项.例如,对于优越关系 $\text{sup} = r_1 > r_2$, 有 $\text{Left}(\text{sup}) = r_1, \text{Right}(\text{sup}) = r_2$.

定义 8. 设 $(R, >):t$ 为某规则集及其上的优越关系,在时刻 $t'(t' \geq t)$ 将一条规则 r 从 $(R, >):t$ 中删除,称为规则的删除操作,形式化表示如下:

$$(R, >):t -_{t'} r = \begin{cases} (R - \{r\}, >):t', & C(R, r, t) = \emptyset \\ (R - \{r\}, >^T - >^A):t', & C(R, r, t) \neq \emptyset \end{cases} \tag{4}$$

式中, $>^A$ 必须满足条件: $>^A = \{\text{sup} | \text{sup} \in >^T, \text{且 } \text{Left}(\text{sup}) = r, \text{或 } \text{Right}(\text{sup}) = r\}$.

为了保证可废止理论的良好性,要求在执行上述的规则变更操作后得到的优越关系是非循环的.下面首先定义可废止理论的良好性,再给出相关的定理.

定义 9. 一个可废止理论 $T = (F, R, >)$ 是良构的,当且仅当 R 上的优越关系 $>$ 是非循环的,即:假设 $>^T$ 是 $>$ 的传递闭包,那么对于任意的 r_i 和 r_j , 如果 $r_i > r_j \in >^T$, 则 $r_j > r_i \notin >^T$.

定理 1. 一个良构的可废止理论,按照上述方法进行规则的变更后,得到的可废止理论仍然是良构的.

证明:设 $T = (F, R, >)$ 是一个良构的可废止理论,所以, $>$ 是非循环的,依据定义 9, $>^T$ 也是非循环的.首先,在执行规则的添加操作时,如果 $C(R, r, t) = \emptyset$, 则 $>$ 不变;如果 $C(R, r, t) \neq \emptyset$, 则 $>$ 变为 $> \cup >^A$, 依据式(3)的第 2 个条件可知,修改后的优越关系是非循环的.其次,在执行规则的删除操作时,被删除的优越关系集 $>^A$ 是 $>^T$ 的子集,因此,

$\succ^T \rightarrow^A \subseteq \succ^T$. 由于 \succ 和 \succ^T 是非循环的, 于是可知 $\succ^T \rightarrow^A$ 也是非循环的. 否则, 如果 $\succ^T \rightarrow^A$ 是循环的, 则 $\exists r, r' \in \succ^T \rightarrow^A \subseteq \succ^T$, 使得 $r \succ r' \in \succ^T$, 且 $r' \succ r \in \succ^T$, 即 \succ^T 是循环的, 矛盾. 综上所述, 规则变更后得到的优越关系是非循环的, 因此与之对应的可废止理论是良构的. \square

3.3.2 规则之间的冲突处理机制

柔性 Agent 的规则冲突分为两类: 同类规则冲突和异类规则冲突. 在对带标签文字(结论)的证明过程中, 规则的冲突处理基于规则之间的优越关系. 对于同类规则, 优越关系是明确指定的; 对于异类规则, 优越关系则是基于 Agent 类型, 由元规则确定, 即在信念规则、政策规则、愿望规则、合同规则和意图规则等 5 种异类规则之间发生冲突时, 应该如何指派优先级. 目前, 解决这个问题的基本方法是依据 Agent 的类型来确定规则之间的优先次序^[21]. 例如, 对于现实型 Agent, 信念规则优越于所有其他类型规则; 对于社会型 Agent, 义务规则强于除了信念规则之外的所有目标规则. 考虑到本文的柔性 Agent 担任自治元素, 它应该是综合类型的, 这些类型包括: 遵从型、社会型和现实型. 其中, 遵从型 Agent 定义如下:

定义 10. 遵从型 Agent 特指能够遵从 IT 管理者的政策指导, 并在此基础上自主工作的 Agent.

依据定义 10 和自治计算系统中自治元素的要求, 规定对于遵从型 Agent, 政策规则优越于除了信念规则之外的其他所有目标规则. 于是, 对于柔性 Agent, 首先有以下 3 条元规则来限定信念规则、政策规则、愿望规则和合同规则等 4 种异类规则之间的优越关系:

$$\begin{aligned} (MR_1) r \in R^B[\varphi], r' \in (R^P \cup R^D \cup R^C)[\sim\varphi] &\rightarrow r \succ r'; \\ (MR_2) r \in R^P[\varphi], r' \in (R^D \cup R^C)[\sim\varphi] &\rightarrow r \succ r'; \\ (MR_3) r \in R^D[\varphi], r' \in R^C[\sim\varphi] &\rightarrow r' \succ r. \end{aligned}$$

MR_1, MR_2 和 MR_3 分别对应于现实型、遵从型和社会型 Agent 的特点.

其次, 意图规则与其他各类规则之间的优越关系由 Agent 的承诺策略来确定. 对于担任自治元素的柔性 Agent, 拟采用开放型承诺策略, 旨在保证 Agent 具备较好的敏捷性. 不过, 也因此会存在 Agent 工作的稳定性问题. 对此, 本文暂不作深入探讨. 所谓开放型承诺策略是指只要意图仍然是目标, 它就维护这些意图. 柔性 Agent 通过意图规则来实现这种承诺策略. 依据定义 2 的意图规则: $\phi_1, \dots, \phi_n: GOAL \psi.t \Rightarrow_{INT} \psi.t; \pi$, 其中, $GOAL \in \{P-OBL, DES, C-OBL\}$, 目标作为与该目标对应的意图的前提条件之一. 这样, 意图规则与其他类型规则之间没有直接的冲突, 它与其他类型规则之间的间接冲突通过目标这一层面来处理.

依据上述的元规则, 可以得到柔性 Agent 的规则冲突处理策略, 见表 1. 表中, 从第 2 列开始, 每一列表示一条属于某类的支持文字 φ 成立的规则在受到可能的攻击后所得的结论. 例如, 当一条支持 φ 成立的信念规则遇到一条支持 $\sim\varphi$ 成立的政策规则的攻击时, 将战胜它, 得到结论 $+\partial_{BEL}\varphi$, 这是因为信念规则比较优越; 而当两条信念规则互相攻击时, 则依赖于同类规则的内部优越关系.

Table 1 Strategies for rule conflict treatment in flexible Agent

表 1 柔性 Agent 规则冲突处理策略

	$\Rightarrow_{BEL}\varphi$	$\Rightarrow_{P-OBL}\varphi$	$\Rightarrow_{DES}\varphi$	$\Rightarrow_{C-OBL}\varphi$
$\Rightarrow_{BEL}\sim\varphi$	Depends on internal superiority relations	$-\partial_{P-OBL}\varphi$ $+\partial_{BEL}\sim\varphi$	$-\partial_{DES}\varphi$ $+\partial_{BEL}\sim\varphi$	$-\partial_{C-OBL}\varphi$ $+\partial_{BEL}\sim\varphi$
$\Rightarrow_{P-OBL}\sim\varphi$	$+\partial_{BEL}\varphi$ $-\partial_{P-OBL}\sim\varphi$	Depends on internal superiority relations	$-\partial_{DES}\varphi$ $+\partial_{P-OBL}\sim\varphi$	$-\partial_{C-OBL}\varphi$ $+\partial_{P-OBL}\sim\varphi$
$\Rightarrow_{DES}\sim\varphi$	$+\partial_{BEL}\varphi$ $-\partial_{DES}\sim\varphi$	$+\partial_{P-OBL}\varphi$ $-\partial_{DES}\sim\varphi$	Depends on internal superiority relations	$+\partial_{C-OBL}\varphi$ $-\partial_{DES}\sim\varphi$
$\Rightarrow_{C-OBL}\sim\varphi$	$+\partial_{BEL}\varphi$ $-\partial_{C-OBL}\sim\varphi$	$+\partial_{P-OBL}\varphi$ $-\partial_{C-OBL}\sim\varphi$	$-\partial_{DES}\varphi$ $+\partial_{C-OBL}\sim\varphi$	Depends on internal superiority relations

3.3.3 信念的产生

定义 11. 对于柔性 Agent 理论 $M=(\mathcal{J}, F, \mathcal{O}, R^B, R^P, R^D, R^C, R^I, Act, \succ)$, 公式 $\varphi.t_b (t_b \in \mathcal{J})$ 是信念, 当且仅当 $M \vdash +\Delta_{BEL}\varphi.t_b$ 或 $M \vdash +\partial_{BEL}\varphi.t_b$ 成立. 这时, 我们称信念 $\varphi.t_b$ 是由 M 产生的.

依据可废止逻辑理论, 带标签的文字 $+\Delta_{BEL}\varphi.t_b, +\partial_{BEL}\varphi.t_b$ 以及 $-\Delta_{BEL}\varphi.t_b$ 和 $-\partial_{BEL}\varphi.t_b$ 的证明基于推导规则, 下

面给出相关定义.

定义 12. 带标签文字 $+\Delta_{BEL}\varphi:t_b, -\Delta_{BEL}\varphi:t_b, +\partial_{BEL}\varphi:t_b$ 和 $-\partial_{BEL}\varphi:t_b$ 的推导规则定义如下:

$+\Delta_{BEL}$:我们可以添加 $P(i+1)=+\Delta_{BEL}\varphi:t_b$, 如果

$\varphi:t_b \in F$, 或

$\exists r \in R_s^B [\varphi:t_b]$, 使得

$\forall a:t_a \in A(r): \Delta_{BEL}a:t_a \in P(1\dots i)$.

$+\partial_{BEL}$:我们可以添加 $P(i+1)=+\partial_{BEL}\varphi:t_b$, 如果

(1) $+\Delta_{BEL}\varphi:t_b \in P(1\dots i)$, 或

(2) (2.1) $\exists r \in R_{sd}^B [\varphi:t_b]$, 使得

$\forall a:t_a \in A(r): \partial_{BEL}a:t_a \in P(1\dots i)$, 且

(2.2) $-\Delta_{BEL}\sim\varphi:t_b \in P(1\dots i)$, 且

(2.3) $\forall s \in R^B[\sim\varphi:t_b]$:

(2.3.1) $\exists a:t_a \in A(s)$, 使得

$-\partial_{BEL}a:t_a \in P(1\dots i)$, 或

(2.3.2) $\exists u \in R_{sd}^B [\varphi:t_b]$, 使得

$\forall a:t_a \in A(u)$:

$+\partial_{BEL}a:t_a \in P(1\dots i)$ 且 $u \succ s$.

$-\Delta_{BEL}$:我们可以添加 $P(i+1)=-\Delta_{BEL}\varphi:t_b$, 如果

$\varphi:t_b \notin F$, 且

$\forall r \in R_s^B [\varphi:t_b]$:

$\exists \varphi:t_a \in A(r)$, 使得 $-\Delta_{BEL}\varphi:t_a \in P(1\dots i)$.

$-\partial_{BEL}$:我们可以添加 $P(i+1)=-\partial_{BEL}\varphi:t_b$, 如果

(1) $-\Delta_{BEL}\varphi:t_b \in P(1\dots i)$, 且

(2) (2.1) $\forall r \in R_{sd}^B [\varphi:t_b]$:

$\exists a:t_a \in A(r)$, 使得 $-\partial_{BEL}a:t_a \in P(1\dots i)$, 或

(2.2) $+\Delta_{BEL}\sim\varphi:t_b \in P(1\dots i)$, 或

(2.3) $\exists s \in R^B[\sim\varphi:t_b]$, 使得

(2.3.1) $\forall a:t_a \in A(s)$:

$+\partial_{BEL}a:t_a \in P(1\dots i)$, 且

(2.3.2) $\forall u \in R_{sd}^B [\varphi:t_b]$:

$\exists a:t_a \in A(u)$, 使得

$-\partial_{BEL}a:t_a \in P(1\dots i)$, 或 $u \succ s$.

在定义 12 中,第一,要使得 $M \vdash +\Delta_{BEL}\varphi:t_b$ 成立(即 $\varphi:t_b$ 是信念),必须具备下列两个条件之一:公式 $\varphi:t_b$ 属于 Agent 的观察信息集;或存在一条结论为 $\varphi:t_b$ 的硬性信念规则,其所有前提均明确可证.第二,如果 $\varphi:t_b$ 不属于 Agent 的观察信息集,且对于每一条结论为 $\varphi:t_b$ 的硬性规则,至少有 1 个前提不可证,那么, $M \vdash -\Delta_{BEL}\varphi:t_b$ 成立.第三,为了表明 $\varphi:t_b$ 是以可废止的方式可证的,即 $M \vdash +\partial_{BEL}\varphi:t_b$, 我们两种选择:其一, $\varphi:t_b$ 已经是明确可证的;其二,使用 M 的可废止部分来证明,这时需要考虑 3 个子条件:存在一条可应用的硬性的或可废止的信念规则(2.1);互补公式 $\sim\varphi:t_b$ 是明确不可证的(2.2);对于所有结论为 $\sim\varphi:t_b$ 的信念公式,要么前提不可证,要么被优越等级较高的规则所反驳(2.3).另外,上一节已说明,信念规则优越于所有其他类型的规则,所以,在考虑反面规则(结论为 $\sim\varphi:t_b$ 的规则)时略去了其他类型规则,包括 3 种目标公式和意图规则.第四,用于 $-\partial_{BEL}\varphi:t_b$ 的推理规则是用于 $+\partial_{BEL}\varphi:t_b$ 的推理规则的结构上的补,具体解释略.

3.3.4 目标的产生

柔性 Agent 目标的产生来源于 3 个方面的动机:愿望、由政策产生的义务和由合同产生的义务.其中,产生义务的合同规则和政策规则可以动态变更(第 3.3.1 节),同时,3 类动机之间以及动机与信念之间可能存在冲突,对此,我们使用第 3.3.2 节介绍的冲突处理机制.对于目标的产生,我们有以下基本假定:

(1) 目标只能由目标规则产生;

(2) 目标的持久性:目标产生之后,在没有出现中断事件(包括规则变更和信念变化)使之不成立之前,该目标一直成立.

另外,我们假定目标规则都是可废止的.因此,对于每个目标公式 $\varphi:t_g$,只有两个带标记的公式 $+\partial_{GOAL}\varphi:t_g$ 和 $-\partial_{GOAL}\varphi:t_g$,其中, $GOAL \in \{P-OBL, DES, C-OBL\}$.于是,我们有如下定义:

定义 13. 对于柔性 Agent 理论 $M=(\mathcal{I}, F, \Theta, R^B, R^P, R^D, R^C, R^I, Act, \succ)$,公式 $\varphi:t_g (t_g \in \mathcal{T})$ 是目标,当且仅当 $M \vdash +\partial_{GOAL}\varphi:t_g$.这时,我们称 $\varphi:t_g$ 是由 M 产生的.下面给出 $+\partial_{GOAL}\varphi:t_g$ 和 $-\partial_{GOAL}\varphi:t_g$ 的推导规则定义(用 $GOAL$ 表示 $P-OBL, DES$ 或 $C-OBL$;用 R^G 表示 R^P, R^D 或 R^C).

定义 14. 带标签文字 $+\partial_{GOAL}\varphi:t_g$ 和 $-\partial_{GOAL}\varphi:t_g$ 的推导规则如下:

$+\partial_{GOAL}$:我们可以添加 $P(i+1)=+\partial_{GOAL}\varphi:t_g$,如果

(1) $\exists r \in R^G[\varphi:t_g]$,使得

(1.1) $\forall a:t_a \in A(r):+\partial_{BEL}a:t_a \in P(1\dots i)$,且

(1.2) $\forall s \in R[\sim\varphi:t_g]$,其中, $R=R^B \cup R^P \cup R^D \cup R^C$:

(1.2.1) $\exists a:t_a \in A(s):-\partial_{BEL}a:t_a \in P(1\dots i)$,或

(1.2.2) $\exists u \in R[\varphi:t_g]$,使得

$\forall a:t_a \in A(u):+\partial_{BEL}a:t_a \in P(1\dots i)$,且 $u \succ s$;或

(2) (2.1) $\exists t_h < t_g, +\partial_{GOAL}\varphi:t_h \in P(1\dots i)$,且

(2.2) $\forall s \in R[\sim\varphi:t_k], t_h < t_k < t_g$:

(2.2.1) $\exists a:t_a \in A(s)$,使得 $-\partial_{BEL}a:t_a \in P(1\dots i)$,或

$-\partial_{GOAL}$:我们可以添加 $P(i+1)=-\partial_{GOAL}\varphi:t_g$,如果

(1) $\forall r \in R^G[\varphi:t_g]$:

(1.1) $\exists a:t_a \in A(r)$,使得 $-\partial_{BEL}a:t_a \in P(1\dots i)$,或

(1.2) $\exists s \in R[\sim\varphi:t_g]$,其中, $R=R^B \cup R^P \cup R^D \cup R^C$,使得

(1.2.1) $\forall a:t_a \in A(s):+\partial_{BEL}a:t_a \in P(1\dots i)$,且

(1.2.2) $\forall u \in R[\varphi:t_g]$:

$\exists a:t_a \in A(u)$,使得 $-\partial_{BEL}a:t_a \in P(1\dots i)$,或 $u \succ s$;且

(2) (2.1) $\forall t_h < t_g: -\partial_{GOAL}\varphi:t_h \in P(1\dots i)$,或

(2.2) $\exists s \in R[\sim\varphi:t_k], t_h < t_k < t_g$,使得

(2.2.1) $\forall a:t_a \in A(s): \partial_{BEL}a:t_a \in P(1\dots i)$,且

(2.2.2) $\forall u \in R[\varphi:t_k]$:

$\exists a:t_a \in A(u): -\partial_{BEL}a:t_a \in P(1\dots i)$,或 $u \succ s$.

定义 14 表明,可废止地证明一个公式是目标,包括两条途径:(1) 存在一条前提均成立的支持该公式成立的目标规则;同时,对于每个结论为 $\sim\varphi:t_g$ 的规则,要么它的前提不成立,要么存在另一条可实施的支持该公式成立的规则有着更高的优越等级;(2) 在时间 t_g 之前的某个时刻 t_h ,该目标公式已经成立;同时,未出现使该目标公式不成立的规则,即对于任何具有互补结论的规则,要么它的前提不成立,要么存在另一条可实施的支持该公式成立的规则有着更高的优越等级.用于 $-\partial_{GOAL}\varphi:t_g$ 的推理规则是用于 $+\partial_{GOAL}\varphi:t_g$ 的推理规则的结构上的补,具体解释略.

3.3.5 意图的产生

柔性 Agent 依据意图规则以及当前的信念和目标,产生意图和规划.对于意图和规划的产生,我们有以下基本假定:

- (1) 意图的产生伴随着规划的产生;
- (2) 意图只能由意图规则产生;
- (3) 意图的持久性:意图产生之后,在支持该意图成立的目标和信念仍然成立之前,该意图一直成立.

另外,意图规则之间的冲突包括意图冲突和动作冲突,因此,先给出动作冲突和意图规则冲突的概念,再定义意图产生和相关的推导规则.

定义 15. 动作冲突:两个动作 α 和 α' ($\alpha, \alpha' \in Act$) 冲突,当 α 与 α' 之间存在资源冲突或逻辑冲突时,记为 $c(\alpha, \alpha')$. Act 中所有动作冲突的集合记作 C ,它满足关系 $C \subseteq Act \times Act$.

所谓资源冲突是指在执行动作 α 和 α' 时,由于时间或其他资源的限制,使得它们不能同时执行.例如,写入动作和删除动作不能同时执行;若某服务 s 是唯一的,则不能同时将其提供给两个不同的服务使用者等.另外,我们规定在动作之上没有非操作.

定义 16. 两条意图规则 $r: \phi_1, \dots, \phi_m: GOAL \psi: t \Rightarrow_{INT} \psi: t; \pi_r$ 和 $s: \phi_1, \dots, \phi_n: GOAL \xi: t \Rightarrow_{INT} \xi: t; \pi_s$ 冲突,当且仅当两个规则的意图互补或两个规划中存在动作冲突.

为了区分两种不同的冲突来源,我们规定 R^I 中与某规则 $r: \phi_1, \dots, \phi_m: GOAL \psi: t \Rightarrow_{INT} \psi: t; \pi_r$ 的意图互补的冲突规则集记作 $R^I[C^I(r) = \sim \psi: t]$,而与 r 中规划 π 存在动作冲突的规则集记作 $R^I[C_\pi]$,其中 $C_\pi \subseteq C$,是指引起冲突的动作对的集合.

定义 17. 对于柔性 Agent 理论 $M = (\mathcal{J}, F, \Theta, R^B, R^P, R^D, R^C, R^I, Act, >)$,公式 $\varphi: t; \pi$ ($t \in \mathcal{T}$) 是意图,当且仅当 $M \vdash \partial_{INT} \varphi: t; \pi$.这时,我们称意图 $\varphi: t; \pi$ 是由 M 产生的.下面给出 $\partial_{INT} \varphi: t; \pi$ 和 $-\partial_{INT} \varphi: t; \pi$ 的推导规则定义.

定义 18. 带标签文字 $\partial_{INT} \varphi: t; \pi$ 和 $-\partial_{INT} \varphi: t; \pi$ 的推导规则如下:

∂_{INT} : 我们可以添加 $P(i+1) = \partial_{INT} \psi: t; \pi$, 如果

- (1) $\exists r \in R^I[C^I(r) = \psi: t]$, 使得
 - (1.1) $\forall \phi: t \in A^B(r): \partial_{BEL} \phi: t \in P(1 \dots i)$, 且
 $\forall GOAL \psi: t \in A^G(r): \partial_{GOAL} \psi: t \in P(1 \dots i)$, 且
 - (1.2) $C^P(r) = \pi$, 且
 - (1.3) $\forall s \in R^I[C^I(r) = \sim \psi: t] \cup R^I[C_\pi]$:
 - (1.3.1) $\exists \phi: t \in A^B(s)$, 使得 $-\partial_{BEL} \phi: t \in P(1 \dots i)$, 或
 $\exists GOAL \psi: t \in A^G(s)$, 使得 $-\partial_{GOAL} \psi: t \in P(1 \dots i)$, 或
 - (1.3.2) $\exists u \in R^I[C^I(r) = \psi: t]$, 使得 $u > s$, 且
 $\forall \phi: t \in A^B(u): \partial_{BEL} \phi: t \in P(1 \dots i)$, 且
 $\forall GOAL \psi: t \in A^G(u):$
 $\partial_{GOAL} \psi: t \in P(1 \dots i)$; 或
- (2) $\exists v \in R^I[C^I(r) = \psi: t']$, 其中 $t' < t$, 使得
 - (2.1) $\partial_{INT} \psi: t'; \pi' \in P(1 \dots i)$, 且
 - (2.2) $\pi' = \pi$, 且
 - (2.3) $\forall \phi: t' \in A^B(v): \partial_{BEL} \phi: t' \in P(1 \dots i)$, 且
 $\forall GOAL \psi: t' \in A^G(v): \partial_{GOAL} \psi: t' \in P(1 \dots i)$.

$-\partial_{INT}$: 我们可以添加 $P(i+1) = -\partial_{INT} \psi: t; \pi$, 如果

- (1) $\forall r \in R^I[C^I(r) = \psi: t]$:
 - (1.1) $\exists \phi: t \in A^B(r)$, 使得 $-\partial_{BEL} \phi: t \in P(1 \dots i)$, 或
 $\exists GOAL \psi: t \in A^G(r)$, 使得 $-\partial_{GOAL} \psi: t \in P(1 \dots i)$, 或

(1.2) $C^P(r) \neq \pi$, 或

(1.3) $\exists s \in R^I[C^I(r) = \sim \psi : t] \cup R^I[C_\pi]$, 使得

(1.3.1) $\forall \phi : t \in A^B(s) : +\partial_{BEL}\phi : t \in P(1 \dots i)$, 且

$\forall GOAL \psi : t \in A^G(s) : +\partial_{GOAL}\psi : t \in P(1 \dots i)$, 且

(1.3.2) $\forall u \in R^I[C^I(r) = \psi : t] : u \not\prec s$, 或

$\exists \phi : t \in A^B(u)$, 使得 $-\partial_{BEL}\phi : t \in P(1 \dots i)$, 或

$\exists GOAL \psi : t \in A^G(u)$, 使得

$-\partial_{GOAL}\psi : t \in P(1 \dots i)$; 且

(2) $\forall v \in R^I[C^I(r) = \psi : t']$, 其中 $t' < t$:

(2.1) $-\partial_{INT}\psi : t'$; $\pi' \in P(1 \dots i)$, 或

(2.2) $\pi' \neq \pi$, 或

(2.3) $\exists \phi : t' \in A^B(v)$, 使得 $-\partial_{BEL}\phi : t' \in P(1 \dots i)$, 或

$\exists GOAL \psi : t' \in A^G(v)$, 使得 $-\partial_{GOAL}\psi : t' \in P(1 \dots i)$.

定义 18 表明,可废止地证明一个公式 $\phi : t$; π 是意图,可以有两条途径:(1) 存在一条前提均成立的意图结论为 $\phi : t$ 的意图规则 $r(1.1)$;同时,该规则的规划结论为 $\pi(1.2)$,且对于每个意图结论为 $\sim \psi : t$ 的或与 r 存在动作冲突的规则,要么它的前提不成立,要么存在另一条可实施的支持该公式成立的规则有着更高的优越等级(1.3);(2) 在时刻 t 之前的某时刻 t' 存在某条规则 v ,使得 $+\partial_{INT}\psi : t'$ 成立(2.1);同时, v 的规划与 r 的规划相同(2.2),且在时刻 t ,规则 v 的前提仍然成立,即表示一个意图产生后,一直有效直到产生该意图的目标不成立,或者 Agent 的信念不支持它成立为止.用于 $-\partial_{INT}\phi : t$; π 的推理规则是用于 $+\partial_{INT}\phi : t$; π 的推理规则的结构上的补,具体解释略.

前面详细定义了信念、目标和意图的产生,我们要求依据上述定义产生的信念、目标和意图必须是一致的,即对于公式 $\phi : t$,不允许 $M \vdash +\partial_X \phi : t$ 和 $M \vdash -\partial_X \phi : t$ 同时成立,式中 $X \in \{BEL, P-OBL, DES, C-OBL, INT\}$.对此,我们有以下定理:

定理 2. 对于一个柔性 Agent 理论,其信念、目标和意图的产生具有一致性.

证明:以意图产生为例,用反证法:假设存在公式 $\phi : t$; π ,使得 $M \vdash +\partial_{INT}\phi : t$; π 和 $M \vdash -\partial_{INT}\phi : t$; π 同时成立.分为两种情况:

第一,假设 $M \vdash +\partial_{INT}\phi : t$; π 成立是因为定义 18 的第 1 个条件满足,即存在一条意图结论为 $\phi : t$ 的意图规则 r ,使得以下 3 个条件同时成立: r 的所有前提均成立; r 的规划部分为 π , r 的结论不会被反面规则所推翻.而依据 $M \vdash -\partial_{INT}\phi : t$; π ,对于所有意图结论为 $\phi : t$ 的意图规则 r ,上述 3 个条件至少有 1 个不成立,产生矛盾.

第二,假设 $M \vdash +\partial_{INT}\phi : t$; π 成立是因为定义 18 的第 2 个条件满足,同理可以得出矛盾.因此,意图的产生满足一致性.

同理可证信念和目标的产生具有一致性(略). □

另外,在可计算性方面, Maher 等人提出的算法证明了可废止理论的结论可以在 $O(N)$ 的时间内得出,其中, N 为可废止理论中的符号数^[18].对于本文提出的柔性 Agent,其具体实现和算法将在进一步的研究中给出.类似的实现和算法请参见文献[18].

4 实 例

用一个提供计算服务的 Agent ag_0 作为自治元素,可以在 IT 管理者的政策指导下管理内部服务,进行服务合成,并通过合同对外开展协作(即向外提供计算服务 s_c). ag_0 是一个柔性 Agent,对应的模型是 $M = (\mathcal{I}, F, \Theta, R^B, R^P, R^D, R^C, R^I, Act, >)$, 在时刻 $t_0 \sim t_2$, M 中的各部分知识如下:

• $\mathcal{I} = \{t_0, t_1, t_2, \dots\}$;

• $F = \{golden(g_1):t_0, \neg golden(g_2):t_0, pay(g_1):t_0, profit(g_2):t_1, request(g_1):t_1, request(g_2):t_1, reward(g_2):t_1, punish(g_1):t_1,$

$acquaintance(g_1):t_1,request(g_1):t_2,request(g_2):t_2\}$, 式中, 文字 $golden(g_i):t_k$ 表示在时刻 t_k , Agent g_i 是金卡用户; $pay(g_i):t_k$ 表示在时刻 t_k , Agent g_i 向 ag_0 支付了服务费用; $profit(g_i):t_k$ 表示在时刻 t_k , ag_0 为 Agent g_i 提供服务获得了较高的利润; $request(g_i):t_k$ 表示在时刻 t_k , ag_0 接收到 Agent g_i 的服务请求; $reward(g_i):t_k$ 表示在时刻 t_k , Agent g_i 获得奖励; $punish(g_i):t_k$ 表示在时刻 t_k , Agent g_i 受到惩罚; $acquaintance(g_i):t_k$ 表示在时刻 t_k , Agent g_i 是熟人;

- Θ 是 Agent 的心智状态, 见表 2;
- $R^B = \{b_1: golden(g_i):t_k \Rightarrow_{BEL} golden(g_i):t_{k+1},$
 $b_2: \neg golden(g_i):t_k \Rightarrow_{BEL} \neg golden(g_i):t_{k+1},$
 $b_3: golden(g_i):t_k, punish(g_i):t_k \Rightarrow_{BEL} \neg golden(g_i):t_{k+1},$
 $b_4: \neg golden(g_i):t_k, reward(g_i):t_k \Rightarrow_{BEL} golden(g_i):t_{k+1}\};$
- $R^P = \{p_1: \neg golden(g_i):t_k, request(g_i):t_k \rightsquigarrow_{P-OBL} \neg provide(s_c, g_i):t_k$
 $p_2: acquaintance(g_i):t_k, request(g_i):t_k \Rightarrow_{P-OBL} provide(s_c, g_i):t_k\};$
- $R^D = \{d_1: profit(g_i):t_k, request(g_i):t_{k+1} \Rightarrow_{DES} provide(s_c, g_i):t_{k+1}\};$
- $R^C = \{c_1: pay(g_2):t_0, t_k > t_0, \neg INT provide(s_c, g_2):t_{k-1} \Rightarrow_{C-OBL} provide(s_c, g_2):t_k\};$
- $R^I = \{i_1: golden(g_i):t_k, GOAL provide(s_c, g_i):t_k \Rightarrow_{INT} provide(s_c, g_i):t_k, do(s_1) || do(s_2); do(s_3)^*$
 $i_2: \neg golden(g_i):t_k, GOAL provide(s_c, g_i):t_k \Rightarrow_{INT} provide(s_c, g_i):t_k, do(s_1) || do(s_2)\};$
- $Act = \{do(s_1), do(s_2), do(s_3)\}$, $do(s_i)$ 是个原子动作: 执行服务 s_i ;
- $> = \{b_3 > b_1, b_4 > b_2, p_1 > p_2\}$.

上式中, $GOAL provide(s_c, g_i)$ ($GOAL \in \{P-OBL, DES, C-OBL\}$) 表示一个目标: 向 Agent g_i 提供服务 s_c ; g_i ($i=1, 2, \dots$) 为外部服务使用者. 另外, 假设在时刻 t_2 , 发生了目标规则变更: 删除政策 p_1 , 添加政策 p_3 ; $punish(g_i):t_k, request(g_i, s_c):t_{k+1} \Rightarrow_{P-OBL} \neg provide(s_c, g_i):t_{k+1}$. 另外, 上述的信念规则、政策规则、愿望规则和意图规则都是规则模式, 它们在运行中将被实例化(具体表示略).

Table 2 The working status of the basic modules of Agent ag_0 from t_0 to t_2

表 2 自 t_0 至 t_2 , Agent ag_0 基本模块工作状况

Time	Rule modification	Belief generation	Goal generation	Intention generation
t_0		$M^+ + \Delta_{BEL} golden(g_1):t_0$		
		$M^+ + \Delta_{BEL} \neg golden(g_2):t_0$		
		$M^+ + \Delta_{BEL} pay(g_2):t_0$		
t_1		$M^+ + \Delta_{BEL} profit(g_2):t_1$		
		$M^+ + \Delta_{BEL} request(g_1):t_1$		
		$M^+ + \Delta_{BEL} request(g_2):t_1$		
		$M^+ + \Delta_{BEL} reward(g_2):t_1$		
		$M^+ + \Delta_{BEL} punish(g_1):t_1$	$M^+ + \hat{c}_{P-OBL} provide(s_c, g_1):t_1$	$M^+ + \hat{c}_{INT} provide(s_c, g_1):t_1;$ $do(s_1) do(s_2); do(s_3)^*$
		$M^+ + \Delta_{BEL} acquaintance(g_1):t_1$		
		$M^+ + \hat{c}_{BEL} golden(g_1):t_1$ $M^+ + \hat{c}_{BEL} \neg golden(g_2):t_1$		
t_2		$M^+ + \Delta_{BEL} request(g_1):t_2$		
	$-t_2 p_1$	$M^+ + \Delta_{BEL} request(g_2):t_2$	$M^+ + \hat{c}_{DES} provide(s_c, g_2):t_2$	$M^+ + \hat{c}_{INT} provide(s_c, g_2):t_2;$ $do(s_1) do(s_2); do(s_3)^*$
	$+t_1 p_3$	$M^+ + \hat{c}_{BEL} \neg golden(g_1):t_2$	$M^+ + \hat{c}_{C-OBL} provide(s_c, g_2):t_2$	
		$M^+ + \hat{c}_{BEL} golden(g_2):t_2$		

从时刻 t_0 至 t_2 , Agent ag_0 基本模块工作状况见表 2(表中只列出了可证的文字), 包括规则的动态变更以及信念产生、目标产生和意图产生.

(一) 规则变更: 规则变更发生在时刻 t_2 , 在变更发生之前(时刻 t_1), 政策规则及其上的优越关系为 $(R^P, >^P):t_1$,

其中, $R_p = \{p_1, p_2\}, >^P = \{p_1 > p_2\}$. 首先, 执行规则删除操作 $-_{t_2} p_1$, 这时, $C(R^P, p_1, t_2) = R^P[\sim C(p_1)] = R^P[provide(s_C, g_i):t_k] = \{p_2\} \neq \emptyset, >^A = \{p_1 > p_2\}$, 于是, 执行删除操作 $-_{t_2} p_1$ 后的结果为 $(R_1^P, >_1^P):t_2 = (R^P, >^P):t_1 -_{t_2} p_1 = (\{p_2\}, \{\}) : t_2$. 然后执行规则添加操作 $+_{t_2} p_3$, 这时, $C(R_1^P, p_3, t_2) = R_1^P[\sim C(p_3)] = R_1^P[provide(s_C, g_i):t_k] = \{p_2\} \neq \emptyset$, 取 $>^A = \{p_3 > p_2\}, >$ 满足定义 7 的两个条件, 于是执行添加操作 $+_{t_2} p_3$ 后的结果为 $(R_2^P, >_2^P):t_2 = (R_1^P, >_1^P):t_2 +_{t_2} p_3 = (\{p_2, p_3\}, \{p_3 > p_2\}):t_2$.

(二) 信念产生: 在 t_0, t_1 和 t_2 , 可证的信念见表 2 第 3 列. 明确可证的信念来源于 ag_0 的观察信息集合 F , 依据定义 12, 显然成立. 以可废止方式可证的信念, 我们举例证明: $M\vdash +_{\partial BEL} golden(g_1):t_1$.

证明: (1) M 中有一条信念规则 $b_1 \in R_{sd}^B[golden(g_1):t_1]$ (该规则由信念规则实例化而来, 下列其他规则类似), 且 b_1 的前提 $golden(g_1):t_0$ 可证, 即 $M\vdash +_{\Delta BEL} golden(g_1):t_0$ 成立; (2) 互补的文字 $\neg golden(g_1):t_1$ 明确不可证, 即 $M\vdash -_{\Delta BEL} \neg golden(g_1):t_1$ 成立; (3) $R^B[\neg golden(g_1):t_1] = \{b_2, b_3\}$, 且在 b_2 和 b_3 的前提中, 分别存在一个条件不成立, 即可证 $M\vdash -_{\Delta BEL} \neg golden(g_1):t_0$ 和 $M\vdash -_{\partial BEL} punish(g_1):t_0$; 依据定义 12, 可得 $M\vdash +_{\partial BEL} golden(g_1):t_1$. \square

(三) 目标产生: 在时刻 t_0 , 未产生目标; 在时刻 t_1 , 有 $M\vdash +_{\partial P-OBL} provide(s_C, g_1):t_1$, 即产生了目标 $provide(s_C, g_1):t_1$ (在时刻 t_1, ag_0 有目标为 Agent g_1 提供服务 s_C); 在时刻 t_2 , 发生了规则变更, p_1 被删除, 且信念也发生了变化, 这时有 $M\vdash +_{\partial DES} provide(s_C, g_2):t_2$ 和 $M\vdash +_{\partial C-OBL} provide(s_C, g_2):t_2$, 即产生了目标 $provide(s_C, g_2):t_2$ (在时刻 t_2, ag_0 有目标为 Agent g_2 提供服务 s_C). 另外可以证明, $M\vdash -_{\partial C-OBL} provide(s_C, g_2):t_1$ 和 $M\vdash -_{\partial P-OBL} provide(s_C, g_1):t_2$ 等.

下面, 我们以 $M\vdash +_{\partial P-OBL} provide(s_C, g_1):t_1$ 为例, 证明如下.

证明: (1) $R^{P-OBL}[provide(s_C, g_1):t_1] = \{p_2\}, p_2$ 的所有前提均成立, 即可证 $M\vdash +_{\Delta BEL} acquaintance(g_1):t_1$ 和 $M\vdash +_{\Delta BEL} request(g_1):t_1$; (2) $R[\neg provide(s_C, g_1):t_1] = \{p_1\}$, 且在 p_1 的前提中, 存在一个条件不成立, 即可证 $M\vdash -_{\partial BEL} \neg golden(g_1):t_1$; 依据定义 14, 可得 $M\vdash +_{\partial P-OBL} provide(s_C, g_1):t_1$. \square

(四) 意图产生: 在时刻 t_0 , 未产生意图; 在时刻 t_1 , 有 $M\vdash +_{\partial INT} provide(s_C, g_1):t_1; do(s_1)||do(s_2); do(s_3)^*$, 即产生了意图 $provide(s_C, g_1):t_1$ 和相应的规划 $do(s_1)||do(s_2); do(s_3)^*$; 在时刻 t_2 , 有 $M\vdash +_{\partial INT} provide(s_C, g_2):t_2; do(s_1)||do(s_2); do(s_3)^*$, 即产生了意图 $provide(s_C, g_2):t_2$ 和相应的规划 $do(s_1)||do(s_2); do(s_3)^*$. 另外我们发现, 对于 t_1 时刻产生的意图 $INT provide(s_C, g_1):t_1; do(s_1)||do(s_2); do(s_3)^*$, 在 t_2 时刻被放弃, 这是由于支持该意图的目标在 t_2 时刻不成立 ($M\vdash -_{\partial P-OBL} provide(s_C, g_1):t_2$), 因此有 $M\vdash -_{\partial INT} provide(s_C, g_1):t_2; do(s_1)||do(s_2); do(s_3)^*$.

下面我们以 $M\vdash +_{\partial INT} provide(s_C, g_1):t_1; do(s_1)||do(s_2); do(s_3)^*$ 为例, 证明如下.

证明: (1) $R^I[provide(s_C, g_1):t_1] = \{i_1, i_2\}$, 其中, i_1 可应用, 即对于 $A^B(i_1) = \{golden(g_1):t_1\}$, 可证 $M\vdash +_{\partial BEL} golden(g_1):t_1$, 对于 $A^G(i_1) = P-OBL provide(s_C, g_1):t_1$, 可证 $M\vdash +_{\partial P-OBL} provide(s_C, g_1):t_1$; (2) $C^P(i_1) = do(s_1)||do(s_2); do(s_3)^*$; (3) $R^I[C^I(r) = \sim \psi : t] \cup R^I[C_{\pi}] = \{\}$; 依据定义 18, 可得 $M\vdash +_{\partial INT} provide(s_C, g_1):t_1; do(s_1)||do(s_2); do(s_3)^*$. \square

5 结论和进一步的工作

在传统的 BDI(信念-愿望-意图)模型中加入社会思维属性(规范、义务等)以反映实际需求, 提高 Agent 的协作效率, 改善 Agent 之间的协作可靠性和安全性, 已经成为新的研究趋势. 例如, 马光伟等人的把信念、愿望和义务作为基本思维属性的 Agent 模型^[22]、Broersen 的 BOID(信念-义务-意图-愿望)模型^[21]、Dignum 等人的用规范和义务扩展的 BDI Agent 模型^[23]和 Boella 的基于合同的 Agent 模型^[24]等. 另一方面, 运用可废止逻辑的非单调推理机制来建立 Agent 模型目前已经有了初步进展, 例如, 在 Governatori 和 Dastani 等人的工作^[14]中提出了一种可废止 Agent 理论, 阐明了如何运用偏好(preference)来表达和处理规则之间的冲突. 这些研究成果表明, 把规范和义务引入 Agent 模型, 并让 Agent 具有非单调的推理能力, 是使得 Agent 能够适应开放、分布和异构的信息系统特点的必然选择. 不过, 上述工作虽然在 Agent 模型中加入了社会思维属性, 引入了基于可废止逻辑的非单调推理机制, 但尚未考虑如何使得 Agent 能够动态变更规则(反映 IT 管理策略和协作关系), 并基于动态规则进行非单调推理.

与上述工作相比, 本文的创新点主要体现在以下 3 个方面:

(1) 提出了一种柔性 Agent 体系结构和模型,这种 Agent 能够接受动态变更的 IT 管理策略(表示为政策规则)的指导,并结合协作合同进行非单调推理,其动机包括 3 个方面:由政策规则产生的义务(反映来自 IT 管理者的动态管理目标和策略)、内部愿望以及由合同规则产生的义务(反映协作关系),其突出的特点是政策规则可以动态变更,使得柔性 Agent 既是自主工作的,又是可控的。

(2) 提出了一种规则动态变更机制和一种基于元规则的冲突处理机制,使得规则能够恰当地加入现有的规则库,或者从规则库中删除,并使得 Agent 可以利用可能冲突的规则进行有效的推理。

(3) 对信念公式、目标公式和意图公式的推导规则进行了新的定义,提出了目标和意图是可废止的且是持久的这一思想(即当目标或意图产生之后,在没有遇到使得目标或意图不成立的事件之前,它们一直有效),并运用该思想给出了目标公式和意图公式推导规则的定义,证明了在该定义下产生的信念、目标与意图是一致的。

本文提出的基于可废止逻辑的柔性 Agent 在体系结构、知识表征和推理决策机制上所具备的上述特点为 Agent 在新领域的应用奠定了基础。不过,作为一个新的理论体系,基于可废止逻辑的柔性 Agent 模型仍然不够成熟。进一步的研究工作包括建立对应的语义模型,给出系统的具体实现和算法,进行系统的复杂性分析等。

References:

- [1] Jennings NR. On agent-based software engineering. *Artificial Intelligence*, 2000,117(2):277–296.
- [2] Lü J, Tao XP, Ma XX, Hu H, Xu F, Cao C. A study of agent-based internetware model. *Science in China (Series E)*, 2005,35(12): 1233–1253 (in Chinese with English abstract).
- [3] Gao J, Yuan CX, Wang J. SASA5: A method system for supporting agent social activities. *Chinese Journal of Computers*, 2005, 28(5):838–848 (in Chinese with English abstract).
- [4] Rao AS, Georgeff M. BDI agents: From theory to practice. In: Georgeff MP, ed. *Proc. of the 1st Int'l Conf. on Multi-Agent Systems (ICMAS'95)*. San Francisco: ACM Press, 1995. 312–319.
- [5] Kephart JO, Chess DM. The vision of autonomic computing. *IEEE Computer*, 2003,36(1):41–50.
- [6] Bonino D, Bosca A, Corno F. An agent based autonomic semantic platform. In: Kephart J, Parashar M, eds. *Proc. of the Int'l Conf. on Autonomic Computing*. New York: IEEE Computer Society, 2004. 189–196.
- [7] Tesauro G, Chess DM, Walsh WE, Das R, Segal A, Whalley I, Kephart JO, White SR. A multi-agent systems approach to autonomic computing. In: Sierra C, Sonenberg L, eds. *Proc. of the 3th Int'l Conf. on Autonomous Agents and Multi-Agent Systems*. New York: IEEE Computer Society, 2004. 464–471.
- [8] Zhang HJ, Shi ZZ. Software engineering for autonomic computing. *Mini-Micro Systems*, 2006,27(6):1077–1082 (in Chinese with English abstract).
- [9] Liao BS, Gao J. Dynamic self-organizing system supported by PDC-agent. *Journal of Computer-Aided Design & Computer Graphics*, 2006,18(2):217–224 (in Chinese with English abstract).
- [10] Liao BS, Huang HX, Gao J. An extended BDI agent with policies and contracts. In: Shi ZZ, Sadananda R, eds. *Proc. of the Agent Computing and Multi-Agent Systems*. LNAI 4088, Berlin, Herdelberg: Springer-Verlag, 2006. 94–104.
- [11] Liao BS, Gao J. A model of multi-agent system based on policies and contracts. In: Pěchouček M, Petta P, Varga LZ, eds. *Proc. of the Multi-Agent Systems and Applications IV*. LNAI 3690, Berlin, Herdelberg: Springer-Verlag, 2005. 62–71.
- [12] Liao BS. Service-Oriented autonomic computing based on PDC-agent [Ph.D. Thesis]. Hangzhou: Zhejiang University, 2006 (in Chinese with English abstract).
- [13] Liao BS, Gao J, Hu J, Chen JJ. Ontology-Based conceptual modeling of policy-driven control framework:oriented to multi-agent system for Web services management. In: Chi CH, Lam KY, eds. *Proc. of the Content Computing*. LNCS 3309, Berlin, Herdelberg: Springer-Verlag, 2004. 346–356.
- [14] Dastani M, Governatori G, Rotolo A, van der Torre L. Programming cognitive agents in defeasible logic. In: Sutcliffe G, Voronkov A, eds. *Proc. of the LPAR 2005*. LNAI 3835, Berlin, Herdelberg: Springer-Verlag, 2005. 621–636.
- [15] Lu RQ. *Artificial Intelligence*. Beijing: Science Press, 2000. 617–659 (in Chinese).
- [16] Nute D. Defeasible logic. In: Bartenstein O, *et al.*, eds. *Proc. of the INAP 2001*. LNAI 2543, Berlin, Herdelberg: Springer-Verlag, 2003. 151–169.

- [17] Cadoli M, Schaerf M. A survey on complexity results for non-monotonic logics. *Journal of Logic Programming*, 1993,17:127-160.
- [18] Maher MJ. Efficient defeasible reasoning systems. *Int'l Journal on Artificial Intelligence Tools*, 2001,10(4):384-392.
- [19] Antoniou G, Billington D. Representation results for defeasible logic. *ACM Trans. on Computational Logic*, 2001,2(2):255-287.
- [20] Billington D, Antoniou G, Governatori G, Maher M. Revising nonmonotonic theories: The case of defeasible logic. In: Burgard W, Christaller T, Cremers AB, eds. *Advances in Artificial Intelligence. LNCS 1701*, Berlin, Herdelberg: Springer-Verlag, 1999. 101-112.
- [21] Broersen J, Dastani M, Hulstijn J, van der Torre LWN. Goal generation in the BOID architecture. *Cognitive Science Quarterly*, 2002,2(3-4):428-447.
- [22] Ma GW, Xu JH, Shi CY. Modeling social agents in BDO logic. *Chinese Journal of Computers*, 2001,24(5):521-528 (in Chinese with English abstract).
- [23] Dignum F, Morley D, Sonenberg E, Cavendon L. Towards socially sophisticated BDI agents. In: Durfee EH, ed. *Proc. of the 4th Int'l Conf. on Multi-Agent Systems*. Washington: IEEE Computer Society, 2000. 111-118.
- [24] Boella G, van der Torre L. Regulative and constitutive norms in normative multiagent systems. In: Dubois D, Welty CA, eds. *Proc. of the 9th Int'l Conf. on the Principles of Knowledge Representation and Reasoning (KR 2004)*. Whistler: AAAI Press, 2004. 255-265.

附中文参考文献:

- [2] 吕建,陶先平,马晓星,胡昊,徐锋,曹春.基于 Agent 的网构软件模型研究. *中国科学(E)*,2005,35(12):1233-1253.
- [3] 高济,袁成祥,王进.支持 Agent 社交活动的方法体系 SASA5. *计算机学报*,2005,28(5):838-848.
- [8] 张海俊,史忠植.自主计算软件工程方法. *小型微型计算机系统*,2006,32(7):1-3.
- [9] 廖备水,高济.PDC-Agent 支持的动态自组织系统. *计算机辅助设计与图形学学报*,2006,18(2):217-224.
- [12] 廖备水.基于 PDC-Agent 的面向服务的自治计算研究[博士学位论文].杭州:浙江大学,2006.
- [15] 陆汝钤. *人工智能*.北京:科学出版社,2000.617-659.
- [22] 马光伟,徐晋晖,石纯一.社会 Agent 的 BDO 模型. *计算机学报*,2001,24(5):521-528.



廖备水(1971-),男,福建古田人,博士,主要研究领域为人工智能,Agent 与多 Agent 系统,自治计算,逻辑学.



高济(1946-),男,教授,博士生导师,主要研究领域为人工智能,网络计算,软件工程,信息智能.



黄华新(1959-),男,教授,博士生导师,主要研究领域为语言逻辑,科学逻辑,“语言与认知”交叉领域.