

# 多 agent 系统中 agent 的行为决策<sup>\*</sup>)

毛新军<sup>1</sup> 王怀民<sup>1</sup> 楚蓓蓓<sup>2</sup> 王逸欣<sup>3</sup>

(国防科学技术大学计算机系 长沙410073)<sup>1</sup>

(信息工程学院计算机系 郑州)<sup>2</sup> (第一军医大学网络中心 广州)<sup>3</sup>

## The Behaviors Decision of Agent in Multi-agent System

MAO Xin-Jun<sup>1</sup> WANG Huai-Min<sup>1</sup> CHU Bei-Bei<sup>2</sup> WANG Yi-Xin<sup>3</sup>

(Department of Computer, National University of Defence Technology, Changsha 410073)<sup>1</sup>

(Department of Computer, College of Information Engineering, Zhengzhou)<sup>2</sup>

(Campus Network Center, First Military Medical University, Gangzhou)<sup>3</sup>

E-mail: xjmiao21@21cn.com

**Abstract** The behavior decision of agent in multi-agent system is dependent on the task of individual agent, the cooperation with other agents, and the multi-agent system's constraints. The paper presents three abstract concept models of achievement intention, joint intention and maintenance intention representing the factors that affect agent's behaviors respectively, to investigate autonomous agent's behaviors decision in support of agent-oriented software development. The paper discusses and analyzes how they will affect agent's behavior decision and the relationship among them, defines their formal and rigorous semantics, finally specifies and proves a number of important properties.

**Keywords** Agent, Achievement intention, Joint intention, Maintenance intention

## 1. 介绍

agent 是指在某一环境下能够持续自主运行, 具有社会性、反应性、自发性等特征的计算实体<sup>[5,11]</sup>。多 agent 系统由一组具有一定资源和能力、相对独立且交互合作的 agent 组成。由于多 agent 系统提供了更高层次的抽象模型, 能够自然、贴切、直观地表示现实世界中的计算实体及其问题求解方式, 提供了有效的工具和手段来控制和管理软件的复杂度, 因而目前有关 agent 理论和技术的研究引起了学术界和产业界的高度关注和重视。

agent 自主行为决策是利用 agent 技术分析和设计多 agent 系统的关键。在多 agent 系统中, 由于 agent 驻留于多 agent 系统环境之中, agent 间动作的相关性、单个 agent 能力的有限性、资源的分布性以及为了满足全局性约束, agent 的行为决策受多方面因素的影响。为了支持多 agent 系统的开发, 我们必须系统、全面地研究影响 agent 行为决策的因素以及这些因素如何影响 agent 的行为, 以提供有效的工具支持多 agent 系统的描述和设计。

计算机科学所研究的通常作法是寻求抽象的概念和工具来促进问题的解决, 如过程抽象、进程、对象等等。在人工智能领域, 人们通常基于意向观点来研究 agent, 即将 agent 视为由一组认知部件构成的意向系统, 代表性工作是 agent 的 BDI 体系结构。意向观点为我们研究 agent 提供了一组高层的抽象认知概念如信念、意愿等等。基于这些概念, 我们可以独立于 agent 的内部结构和具体实现细节来构造 agent 体系结构, 定义 agent 状态, 分析 agent 行为的规律性特征。

在多 agent 系统中, 个体 agent 的行为决策受其所具有的

任务、与其它 agent 间的合作以及多 agent 系统的约束三个不同抽象层次的影响和限制。基于意向观点的思想, 本文提出了三个不同的抽象概念模型: 实现型意图、联合意图和维护型意图, 分别表示影响 agent 行为的三个不同抽象层次, 来研究影响自主 agent 行为决策的因素。论文讨论和分析了实现型意图、联合意图和维护型意图对 agent 行为决策的影响以及它们之间的关系, 基于多 agent 系统的逻辑框架, 给出了其严格、形式化的语义定义, 描述和验证了一些重要的属性, 最后对本文的工作做了总结。

## 2. agent 行为决策

在多 agent 系统中, 根据影响范围、作用域以及作用性质上的差异, 个体 agent 的行为决策受其所具有的任务、与其它 agent 间的合作以及多 agent 系统的社会约束三个不同抽象层次的影响和限制。

agent 是一个目标指导(goal-directed)的计算实体, agent 总是根据其所拥有的任务和选择动作, 执行规划。在多 agent 系统中, agent 基于其任务和目标的行为决策是理性的, 一方面, agent 总是根据其任务和选择动作, 因而 agent 所拥有的任务应该是一致、非冲突和可实现的, 另一方面, agent 的动作选择并不是盲目的, 一旦发现任务已经实现或者已经不可能实现, agent 将终止其任务。agent 任务对 agent 行为的影响将作用于拥有该任务的 agent, 并贯穿 agent 的整个生命周期。为了研究 agent 的任务对 agent 行为决策的影响, 我们引入了实现型意图抽象概念模型。agent 具有实现型意图  $\varphi$  是指 agent 意图使  $\varphi$  成立, 它是对 agent 的任务和目标的抽象表示。

<sup>\*</sup>) 本文得到国家自然科学基金 60003002、973 项目 G1999032700 的支持。毛新军 博士, 副教授, 研究方向: agent 理论和技术的新型软件开发方法学等。

在多 agent 系统中,由于 agent 间动作的相关性、单个 agent 能力的有限性、资源的分布性以及为了满足全局性约束而实现问题求解的自然性,agent 间合作是必须的,它揭示了多 agent 系统的社会性行为<sup>[10]</sup>。agent 间的合作必将对参与合作的个体 agent 的行为产生影响,这种影响将贯穿 agent 的整个合作过程,并随 agent 间合作的形成而产生,随合作的结束而终止。从本质上看,agent 间合作将影响参与合作 agent 的内部状态,从而影响相关 agent 的行为。为了研究 agent 间合作对 agent 行为决策的影响,我们引入了联合意图抽象概念模型。agent 间的联合意图体现了多个 agent 的联合行为选择以共同实现任务和目标,它对 agent 间合作社会性行为的抽象表示。

在多 agent 系统中,每个 agent 的资源和能力都是有限的而且并非是完全自主和独立的,它们构成了一个社会。为了使多 agent 系统是有序的,能够有效地促进 agent 间的协同与合作、减少冲突,多 agent 系统往往具有相应的社会规则。agent 处于这样的社会环境中,因而必定受社会规则的约束和限制。例如在某一环境中存在一组机器人 agent。为了减少冲突,这些机器人的设计人员在开发它们的过程中提出了一组它们必须遵循的社会规则:每个机器人在移动过程中必须靠左行走。又如自动售货机的主人希望在任意时刻自动售货机所接收到的硬币价值大于或等于它所出售的巧克力的价值。这些社会规则显式或隐式地存在于多 agent 系统之中,它们影响和约束系统中各个 agent 的行为。多 agent 系统中的社会规则将影响处于该系统中的所有 agent,并贯穿系统的整个生命周期。这些社会行为约束对 agent 而言是外在、具体的,为了指导 agent 计算以及规范和描述 agent,必须对它们进行抽象和内部化。从本质上看,多 agent 系统的社会性规则将影响系统中各个 agent 的内部状态,从而影响它们的行为决策。为此,我们引入了维护型意图的概念以抽象地表示多 agent 系统社会性规则。

表1 影响 agent 行为决策的三个不同抽象层次

	作用域	作用周期	抽象表示
agent 的任务和目标	个体 agent	agent 生命周期	实现型意图
agent 间的合作	参与合作的 agent	合作过程	联合意图
多 agent 系统的社会性约束	系统中的所有 agent	多 agent 系统的生命周期	维护型意图

### 3. 逻辑模型

我们将基于多 agent 系统的逻辑框架来定义和分析实现型意图、联合意图和维护型意图概念,研究、描述和验证它们之间的关系。逻辑框架包括语法、模型和语义解释三个部分。

形式化语言 L 是对命题分枝时序逻辑 CTL<sup>\*</sup><sup>[11]</sup>的扩充。语言 L 由状态公式集 L<sub>s</sub> 和路径公式集 L<sub>p</sub> 二部分组成。设 Φ 是原子命题符号集合,Const<sub>ag</sub> 是 agent 符号集合。为简化说明,文中有下列符号约定:p, q, ... 为原子命题;φ, ψ, ... 为公式;x, y, ... 为 agent。

定义 3.1(语言 L 的语法) 形式化语言 L 是由下列规则定义的最小封闭集合:

- (1) 如果 p ∈ Φ, 则 p ∈ L,
- (2) 如果 φ, ψ ∈ L<sub>s</sub> 且 x, y ∈ Const<sub>ag</sub>, 则 ¬φ, φ ∧ ψ, Bel(x, φ), MB(x, y, φ), AI(x, y, φ), MI(x, y, φ), MAI(x, y, φ),

MAB(x, y, φ), WAC(x, y, φ), MAC(x, y, φ), JI(x, y, φ) ∈ L<sub>p</sub>

(3) L<sub>s</sub> ⊆ L<sub>p</sub>,

(4) 如果 φ, ψ ∈ L<sub>s</sub>, x ∈ Const<sub>ag</sub>, 则 ¬φ, φ ∧ ψ, φ Until ψ, ψ Until φ ∈ L<sub>s</sub>,

(5) 如果 φ ∈ L<sub>s</sub>, 则 Aφ ∈ L<sub>p</sub>

语言 L 的一个模型是 M = (T, <, U<sub>ag</sub>, π, [], B, C)。T 是时刻集, T 中的每一时刻对应于世界的一个状态。< 是 T 上的偏序关系, 它描述了时刻间的先后次序。任一时刻的过去是确定和线性的, 它的将来可能是分枝的, 形式化模型呈图 1 所示的树形结构。U<sub>ag</sub> 是 agent 集合, π: Φ → ℘(T), ℘ 是幂集符号, π(p) 定义了使原子公式 p 成立的时刻集。[] 是对 agent 符号的赋值。B: U<sub>ag</sub> → ℘(T × T), (t, t') ∈ B(x) 是指 agent<sub>x</sub> 在 t 时刻认为 t' 时刻是可能的, B 用于定义 agent 的信念。

时刻 t 的一条路径是指始于该时刻, 由 t 的将来时刻构成的一条线性分枝, 它刻画了世界的某种发展轨迹。

定义 3.2(路径) 时刻 t 的一条路径是指集合 S ⊆ T 且满足: (1) t ∈ S; (2) ∀ t<sub>1</sub>, t<sub>2</sub> ∈ S: (t<sub>1</sub> < t<sub>2</sub>) ∨ (t<sub>2</sub> < t<sub>1</sub>) ∨ (t<sub>1</sub> = t<sub>2</sub>); (3) ∀ t<sub>1</sub>, t<sub>2</sub> ∈ S; t<sub>3</sub> ∈ T: (t<sub>1</sub> < t<sub>3</sub> < t<sub>2</sub>) ⇒ (t<sub>3</sub> ∈ S); (4) ∀ t<sub>1</sub> ∈ S; t<sub>2</sub> ∈ T: (t<sub>1</sub> < t<sub>2</sub>) ⇒ (∃ t<sub>3</sub> ∈ S: (t<sub>1</sub> < t<sub>3</sub>) ∧ ¬(t<sub>3</sub> < t<sub>2</sub>)); (5) ∀ t<sub>1</sub> ∈ S: (t = t<sub>1</sub>) ∨ (t < t<sub>1</sub>)。

其中, (1) 表示时刻 t 的路径包含该时刻; (2) 刻画了路径的线性特征; (3) 指出了路径的稠密性; (4) 描述了路径的相对最大性; (5) 刻画了路径的初始性。设 S<sub>t</sub> 表示时刻 t 的所有路径的集合, S<sub>s</sub> 是所有路径的集合。

在多 agent 系统中, 各个 agent 的动作并发、异步地发生。在任一时刻 agent 可能执行各种动作并通过动作的执行来影响和控制世界的发展, 然而这种影响和控制是有限的, 世界发展的轨迹还受其它 agent 动作执行事件的影响, 所有 agent 动作执行事件和环境事件共同确定世界的发展。考虑图 1 所示的由二个 agent 构成的多 agent 系统的形式化模型。图中的结点表示时刻, 边表示多个 agent 的动作并发地发生。假设“||”左侧符号表示 agent<sub>1</sub> 的动作, 右侧符号表示 agent<sub>2</sub> 的动作。在 t<sub>0</sub> 时刻 agent<sub>1</sub> 通过执行动作 a 使得世界沿 t<sub>1</sub> 或 t<sub>2</sub> 方向发展, 但世界发展的将来时刻是 t<sub>1</sub> 还是 t<sub>2</sub> 还取决于 agent<sub>2</sub> 的动作。当 agent<sub>2</sub> 执行动作 c 时则世界沿 t<sub>1</sub> 方向发展, 当 agent<sub>2</sub> 执行动作 d 时则世界沿 t<sub>2</sub> 方向发展。

C: U<sub>ag</sub> × T → ℘(S<sub>s</sub>), S ∈ C(x, t), 是指在 t 时刻 agent<sub>x</sub> 选择的路径, 因而有 C(x, t) ⊆ S<sub>s</sub>, C 用于定义联合意图概念。为简化研究, 在模型中限定: ∀ t ∈ T, x ∈ U<sub>ag</sub>: C(x, t) ≠ ∅。

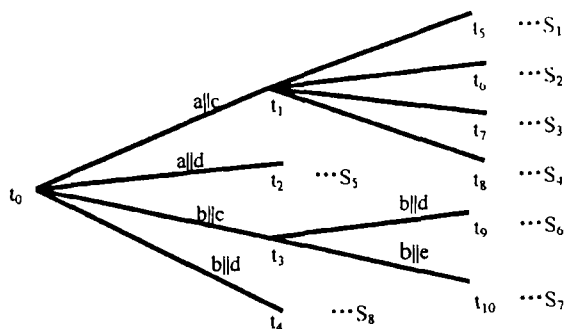


图1 多 agent 系统的形式化模型

L<sub>s</sub> 中公式的语义定义由模型 M 和时刻 t 给出。M ⊨<sub>t</sub> φ 表示模型 M 在时刻 t 满足公式 φ。L<sub>p</sub> 中公式的语义由模型 M, 路径 S 和时刻 t 加以定义。M ⊨<sub>s,t</sub> ψ 表示模型 M 在路径 S 的

时刻  $t$  满足公式  $\psi$ 。

定义3.3 (语言  $L$  的语义)

- (1)  $M \models p$  iff  $t \in \pi(p)$
- (2)  $M \models \psi \wedge \varphi$  iff  $M \models \psi$  且  $M \models \varphi$
- (3)  $M \models \neg \varphi$  iff  $M \not\models \varphi$
- (4)  $M \models A \varphi$  iff  $\forall S: S \in S_i \Rightarrow M \models_{s,t} \varphi$
- (5)  $M \models Bel(x, \varphi)$  iff  $\forall t': (t, t') \in B(x) \Rightarrow M \models_{t'} \varphi$
- (6)  $M \models_{s,t} \psi \wedge \varphi$  iff  $M \models_{s,t} \psi$  且  $M \models_{s,t} \varphi$
- (7)  $M \vdash_{s,t} \neg \varphi$  iff  $M \not\models_{s,t} \varphi$
- (8)  $M \models_{s,t} \psi \text{Until} \varphi$  iff  $\exists t' \in S: (t \leq t') \text{ 且 } (M \models_{s,t'} \varphi) \text{ 且 } (\forall t'': t \leq t'' < t' \Rightarrow M \models_{s,t''} \psi)$
- (9)  $M \models_{s,t} \psi \text{Until}_V \varphi$  iff  $\forall t' \in S: (\forall t'': t \leq t'' \leq t \Rightarrow (M \models_{s,t''} \neg \varphi) \Rightarrow M \models_{s,t''} \psi)$
- (10)  $M \models_{s,t} \varphi$  iff  $M \models \varphi$ , 其中  $\varphi \in L$ 。
- (11)  $M \models MB(x, y, \varphi)$  iff  $M \models Bel(x, \varphi)$  且  $M \models Bel(y, \varphi)$

根据上述语义定义,我们可以派生出其它命题连接词和算子。 $Until$  是“until”算子, $Until_V$  是弱“until”算子。 $F\varphi = true \text{Until} \varphi$ ,  $F$  是存在时序量词。 $G$  是  $F$  的对偶算子即  $G\varphi = \neg F(\neg\varphi)$ , 因而  $G$  是全称时序量词。 $A$  是全称路径算子,  $A\varphi$  在某一时刻  $t$  成立当且仅当  $\varphi$  在该时刻的所有路径上均成立。 $E$  是  $A$  的对偶算子即  $E\varphi = \neg A(\neg\varphi)$ , 因而  $E$  是存在路径算子。 $Bel(x, \varphi)$  表示 agent 具有信念  $\varphi$ , 我们假定  $B(x)$  满足自反性和传递性, 因而  $Bel$  对应于 S4 系统中的正规模态算子。

定理3.1 算子  $Bel$  有以下性质 (1)  $\models Bel(x, \varphi) \rightarrow \varphi$ , (2)  $\models Bel(x, \varphi) \rightarrow Bel(x, Bel(x, \varphi))$ , (3)  $\models Bel(x, \varphi) \wedge Bel(x, \varphi \rightarrow \psi) \rightarrow Bel(x, \psi)$ , (4) 如果  $\models \varphi$ , 则  $\models Bel(x, \varphi)$

#### 4. 实现型意图、联合意图和维护型意图

下面我们将给出实现型意图、联合意图和维护型意图的语义定义, 分析这三个抽象概念模型在 agent 行为决策过程中它们之间的相互关系。

agent 的实现型意图是指 agent 意图实现某个命题, 它对应于 agent 的任务和目标。多 agent 系统的形式化模型是一个树形的分枝结构, 模型中任一时刻的分枝表示 agent 在该时刻所具有的各种可能选择。agent 具有实现型意图  $\varphi$  是指 agent 对世界发展轨迹(即路径)的选择, 在这些被选择的世界发展轨迹中, agent 知道  $\varphi$  将最终成立。

定义4.1(实现型意图的语义)  $M \models AI(x, \varphi)$  iff  $M \models Bel(x, \neg\varphi)$  且  $(\forall S: S \in C(x, t) \Rightarrow M \models_{s,t} FBel(x, \varphi))$

上述语义定义揭示了实现型意图的最本质特性即选择性。不同于已有方法, 我们没有基于可能世界间的可达关系来定义实现型意图概念, 而是将 agent 意图视为 agent 对世界发展轨迹的选择。在形式化模型中, 世界发展轨迹与 agent 的动作是密切相关的, 任一世界发展轨迹对应于多 agent 系统中各个 agent 动作执行序列的组合。我们将 agent 的意图解释为 agent 对世界发展轨迹的选择不仅清晰地刻画了意图概念的选择特征, 揭示了在多 agent 系统中 agent 意图与 agent 的行为之间的关系, 反映了 agent 意图对 agent 未来动作的影响和限制, 而且有助于我们进一步定义维护型意图概念的语义。

agent 的维护型意图是指 agent 意图维持某个命题使之恒成立。不同于实现型意图, 我们将 agent 的维护型意图解释为 agent 在实施社会性行为过程中所具有的系统约束。这些约束显式或隐式地存在于 agent 之中, 它们将影响和约束 agent

的行为。

定义4.2(维护型意图的语义)  $M \models MI(x, \varphi)$  iff  $(\forall S: S \in C(x, t) \Rightarrow M \models_{s,t} GBel(x, \varphi))$

联合意图体现多个 agent 的联合行为选择, 选择性和联合性是联合意图概念的本质属性。下面我们将基于这两个本质属性给出联合实现型意图的语义定义。

agent<sub>x</sub> 和 agent<sub>y</sub> 具有联合意图  $\varphi$  的前提条件之一是它们具有共同的选择。

定义4.3(共同选择)  $MAI(x, y, \varphi) = AI(x, \varphi) \wedge AI(y, \varphi)$

二个 agent 具有共同的选择来实现  $\varphi$  并不意味着它们就具有联合意图  $\varphi$ 。在具备共同选择的基础上, agent<sub>x</sub> 和 agent<sub>y</sub> 要形成联合意图, 还必须满足联合特征。agent 间的联合特征包括二方面的内容: 意图互知和合作。

定义4.4(互知意图)  $MAB(x, y, \varphi) = MB(x, y, AI(x, \varphi) \wedge AI(y, \varphi))$

上述概念定义表明, agent<sub>x</sub> 和 agent<sub>y</sub> 互知意图是指它们都知道双方都具有实现型意图  $\varphi$ 。

为了定义合作概念的语义, 我们首先引入“弱合作”概念。

定义4.5(弱合作)  $WAC(x, y, \varphi) = (Bel(x, \varphi) \wedge \neg Bel(x, Bel(y, \varphi)) \rightarrow AI(x, MB(x, y, \varphi))) \wedge (Bel(x, AG \rightarrow \varphi) \wedge \neg Bel(x, Bel(y, AG \rightarrow \varphi)) \rightarrow AI(x, MB(x, y, AG \rightarrow \varphi)))$

$AG \rightarrow \varphi$  表示对于所有的路径,  $\varphi$  在这些路径上的任何时刻均不成立。agent<sub>x</sub> 关于  $\varphi$  与 agent<sub>y</sub> 具有弱实现型合作是指如果 agent<sub>x</sub> 知道  $\varphi$  已经成立, 并且不知道  $y$  知道  $\varphi$  已成立, 则 agent<sub>x</sub> 意图让双方都互知  $\varphi$  已成立; 并且如果 agent<sub>x</sub> 知道  $\varphi$  已经不可能成立, 并且不知道  $y$  知道  $\varphi$  已不可能成立, 则 agent<sub>x</sub> 意图让双方都互知  $\varphi$  已不可能成立。基于弱合作概念, 下面我们给出合作概念的语义定义。

定义4.6(合作)  $M \models MAC(x, y, \varphi)$  iff  $(\forall S: S \in C(x, t) \Rightarrow M \models_{s,t} (MB(x, y, WAC(x, y, \varphi) \wedge WAC(y, x, \varphi))) \text{Until} \rightarrow AI(x, \varphi))$  且  $(\forall S: S \in C(y, t) \Rightarrow M \models_{s,t} (MB(x, y, WAC(x, y, \varphi) \wedge WAC(y, x, \varphi))) \text{Until} \rightarrow AI(y, \varphi))$

agent<sub>x</sub> 和 agent<sub>y</sub> 具有合作  $\varphi$  是指在联合意图的实施过程中, 它们都知道自己会和对方合作, 对方也会和自己合作, 及至它们放弃意图  $\varphi$ 。

定义4.7(联合实现型意图)  $JAI(x, y, \varphi) = MAI(x, y, \varphi) \wedge MAB(x, y, \varphi) \wedge MAC(x, y, \varphi)$

上述概念定义表明, agent<sub>x</sub> 和 agent<sub>y</sub> 具有联合意图  $\varphi$  当且仅当它们具有共同的选择实现  $\varphi$ , 它们相互知道双方都有这样的选择, 并且在实施联合实现型意图的过程中, 双方都互信会有必要的合作。联合实现型意图的上述语义定义清晰、简洁、准确地描述了联合实现型意图的本质特征: 选择性和联合特征。

基于实现型意图、维护型意图和联合意图的上述语义定义, 我们可以进一步分析它们在 agent 行为决策过程中的相互关系。

定理4.1(实现型意图与维护型意图的一致性)  $\models \neg(AI(x, \varphi) \wedge MI(x, \varphi)) \wedge \neg(AI(x, \varphi) \wedge MI(x, \neg\varphi))$

上述定理表明 agent 的实现型意图与维护型意图间是一致的。agent 不可能既有维护型意图同时又有实现型意图  $\varphi$ 。而且 agent 也不可能既有维护型意图  $\neg\varphi$  同时又有实现型意图  $\varphi$ 。可根据实现型意图和维护型意图的语义定义来证明该定理。  
(下转第74页)

理内置知识库中的调度或计划模块中的规则实现信息代理的智能行为。

3. 应用集成代理主要实现遗留应用的物理封装,基于 KQML 和 KIF 的应用代理之间的互操作,并以公共服务代理实现软件应用的注册与管理。

**总结** 多代理系统研究相互独立而又有联系的个体(智能代理)在分布式环境下的协调合作,其研究及成果为支持企业集成的协同信息系统的描述和实现提供了有利的支持。本文首先从人工智能领域对 Agent 理论研究的内容和进展进行了讨论。在此基础上,分别从计划与调度等企业应用集成、敏捷供应链管理及基于 Web 的企业信息综合管理等不同的角度对多代理技术应用于企业集成以支持企业实现智能业务运作的相关项目及应用进行介绍,并对其中重要问题做相关讨论,给出了一种应用多代理技术实现协同信息系统的方案框架。

## 参考文献

- Jennings, Sycara, Wooldridge. *Autonomous Agents and Multi-Agent Systems*. Kluwer Academic Publishers, Boston, 1998. 275~

(上接第70页)

**定理4.2(联合意图与实现型意图的对应关系)**  $\models JI(x, y, \varphi) \rightarrow AI(x, \varphi) \wedge AI(y, \varphi)$

上述定理表明,如果 agent<sub>x</sub> 和 agent<sub>y</sub> 具有联合意图  $\varphi$ , 那么 agent 间的联合意图将反应到参与其中的所有 agent 的内部状态,从而影响和引导 agent 的行为决策,即参与合作的所有 agent 将具有实现型意图  $\varphi$ 。

**定理4.3(联合意图与维护型意图的一致性)**  $\models \neg(JI(x, y, \varphi) \wedge MI(x, \varphi)) \wedge \neg(JI(x, y, \varphi) \wedge MI(x, \neg\varphi))$

上述定理表明 agent 间的联合意图与维护型意图间是一致的,agent 不可能既有联合意图  $\varphi$  同时又有维护型意图  $\varphi$ 。而且 agent 也不可能既有联合意图  $\varphi$  同时又有维护型意图  $\neg\varphi$ 。可根据定理4.1和4.2来证明该定理。

**定理4.4(实现型意图与联合意图的非冲突性)**  $\models JI(x, y, \varphi) \wedge AI(x, \psi) \rightarrow E(F Bel(x, \varphi) \wedge F Bel(x, \psi))$

上述定理表明 agent 间的联合意图与实现型意图是非冲突的,如果 agent 既有联合意图  $\varphi$  同时又有实现型意图  $\psi$ , 那么 agent 至少存在一条路径,在该路径上 agent 知道  $\varphi$  和  $\psi$  将以某种时序关系得以实现。可根据联合意图和实现型意图的语义定义以及模型限定:  $\forall t \in T, x \in U_x: C(x, t) \neq \emptyset$  来证明该定理。

**定理4.5(实现型意图与维护型意图的非冲突性)**  $\models AI(x, \varphi) \wedge MI(x, \psi) \rightarrow E(Bel(x, \psi) U Bel(x, \varphi))$

上述定理指出如果 agent 具有实现型意图  $\varphi$  同时具有维护型意图  $\psi$ , 则存在一条路径使得 agent 知道  $\varphi$  将最终成立并且在此过程中 agent 知道  $\psi$  始终成立。可根据实现型意图和维护型意图的语义定义证明该定理。

**定理4.6(联合意图与维护型意图的非冲突性)**  $\models JI(x, y, \varphi) \wedge MI(x, \psi) \rightarrow E(Bel(x, \psi) U Bel(x, \varphi))$

上述定理指出如果 agent 具有联合意图  $\varphi$  同时具有维护型意图  $\psi$ , 则存在一条路径使得 agent 知道  $\varphi$  将最终成立并且在此过程中 agent 知道  $\psi$  始终成立。可根据联合意图和维护型意图的语义定义证明该定理。

**定理4.7(实现型意图、联合意图与维护型意图的可满足性)**  $\models JI(x, y, \varphi) \wedge MI(x, \psi) \wedge AI(x, \phi) \rightarrow E(G Bel(x, \psi) \wedge F Bel(x, \varphi) \wedge F Bel(x, \phi))$

306

- 徐晋晖. Agent 模型与联盟机制研究:[博士学位论文]. 北京:清华大学计算机系, 2000
- Rao A S, Georgeff M P. BDI agents: from theory to practice. In: Georgeff M P, ed. Proc. of the 1st Intl. Conf. on Multi-Agent Systems(ICMAS-95). San Francisco. ACM Press, 1995. 312~319
- Smith R, Davis R. Frameworks for Cooperation in Distributed Problem Solving. IEEE Trans. on S. M. C. 1981. 61~80
- Cheyner A, Martin D. The Open Agent Architecture. Journal of Autonomous Agents and Multi-Agent Systems, March 2001. 143~148
- Chu WJ and Tolone, et al. Integrating Manufacturing Softwares for Intelligent Planning Execution: a CIIMPLEX Perspective. http://www.cs.umbc.edu/~finin/papers/papers/spie96.pdf
- Shen W, Norrie D H. A Hybrid Agent-Oriented Infrastructure for Modeling Manufacturing Enterprises. In: Proc. of KAW'98, Banff, Canada, 1998(Agent-5: 1~19)
- Sycara K. In-Context Information Management through Adaptive Collaboration of Intelligent Agents. Intelligent Information Agents Matthias Klusch(Ed.) Springer, 1999

上述定理指出如果 agent 具有联合意图  $\varphi$  同时具有维护型意图  $\psi$  以及实现型意图  $\phi$ , 则存在一条路径使得 agent 知道  $\varphi$  以及 agent 知道  $\phi$  将最终成立, 并且在该路径上 agent 知道  $\psi$  始终成立。可根据实现型意图、联合意图和维护型意图的语义定义以及模型限定:  $\forall t \in T, x \in U_x: C(x, t) \neq \emptyset$  来证明该定理。

**结论** agent 的行为决策是一个非常复杂的过程,它需要考虑诸多的因素。本文分析了个体 agent 的任务、agent 间的合作以及多 agent 系统的社会约束对 agent 行为决策的影响,提出了实现型意图、联合意图、维护型意图抽象概念来分别表示影响 agent 行为决策的三个因素,给出了它们严格的语义定义,分析了它们之间的关系,描述和验证一些重要属性。我们将根据该成果进一步系统、全面地研究 agent 在多 agent 系统中的运行机制和执行模型。

## 参考文献

- Chen P R, Levesque H J. Intention is choice with commitment. Artificial Intelligence, 1990, 42(2-3): 213~261
- Chen P R, Levesque H J. Teamwork. Nous, 21, 1991
- Tuomela R. Collective and joint intention. In: Proc. of cognitive theory of social action, 1998
- Tuomela R. Philosophy and distributed artificial intelligence: The case of joint intention. Foundation of distributed artificial intelligence, John Wiley&Sons, 1996: 487~504
- 毛新军,王怀民. Agent 在多 Agent 系统中计算的意愿理论. 软件学报, 1999, 10(1): 43~48
- Singh M P. Multiagent System: A Theoretical framework for Intentions, Know-how, and Communications. Berlin, Heidelberg: Springer-Verlag, 1994
- Chellas B. Modal logic: an introduction. Cambridge University Press, Cambridge, MA, 1980
- Dunin-Keplicz B, Verbrugge R. Collective Commitment. In: Proc. of the second intl. conf. on multi-agent system, Menlo Park, CA, 1996
- Panzarasa P, Jennings N R. Social mental shaping: Modelling the impact of sociality on the mental state of autonomous agents. Computational Intelligence, 2001
- Jennings N R. Specification and implementation of belief-desire-joint intention architecture for collaborative problem solving. Journal of Intelligent and Cooperative Information Systems, 1993, 2(3): 289~318
- Jennings N R. Building complex software system: the case for an agent-based approach. Communication of ACM, 2001