

# 基于 BDI 的对手 Agent 模型\*

李毅, 石纯一

(清华大学 计算机科学与技术系, 北京 100084)

E-mail: tiger\_li@sjdd.com.cn

http://www.tsinghua.edu.cn

**摘要:** MAS 中建立对手模型, 并进行意图和规划识别是有效交互的必要条件. BDI 是描述思维的有效手段, 可用来建立通用的对手思维状态模型. 当前对手模型的研究偏重于对单一行为结果的推断, 不能充分表示并发多意图及其关系, 也不利于实时地为决策提供支持. 提出的基于 BDI 建立通用对手 Agent 模型的算法改进了 Anand S.Rao 和 Milind Tambe 等人有代表性的研究成果.

**关键词:** Agent; 对手模型; 规划识别; BDI; MAS

**中图分类号:** TP18 **文献标识码:** A

MAS 社会性体现于 Agent 交互, 建立对手模型并进行意图和规划识别是有效交互的必要条件. “对手”是一个泛化概念, 协商伙伴也被称为“对手”. 在实际应用中, 由于存在开放、动态、对抗的环境以及 Agent 资源有限等条件, 对建立对手模型的算法提出了更高的要求.

在 MAS 研究中, BDI 理论描述了 Agent 思维状态的基本特征和相互关系, 利用 BDI 可建立通用的对手的思维状态模型; 在对抗等信息不完全的情况下, 按 BDI 间的约束关系推理判断, 消除矛盾, 对于提高模型的准确性以及建立的速度是有效的<sup>[1]</sup>.

对手模型的核心在于意图的识别, 因为意图对应于实际的行为规划. 当前, 在已有算法中, Anand S.Rao<sup>[2]</sup>和 Milind Tambe<sup>[3,4]</sup>等人的工作具有代表性. Milind Tambe 在 RESC(real-time situated commitments)模型中提出了跟踪 Agent 行为, 判断其目标的算法, 利用行为树来建立对手模型, 并采用单状态回溯机制来修正错误. 但其忽略了对手的思维状态, 将行为与思维分离; 某一时刻, 只存在一个对手操作树, 不能同时跟踪多个意图. Anand S.Rao 实现了一种基于 BDI 的对手思维状态识别, 采用事件触发机制, 可保存多个可能意图树, 并利用观察到的对方行动和 BDI 约束将其逐步减少为单一意图. 但 Rao 的算法须建立对方完整的规划模型, 没有考虑多个可能意图树相互之间的联系和比较, 也不利于为决策提供实时的支持.

本文建立了基于 BDI 的对手 Agent 模型和相应算法, 可以达到在动态开放环境下, 特别是信息不完全的情况下, 实时跟踪对手的思维状态, 进行意图与规划识别, 并为决策提供支持. 改进了 Anand S.Rao 和 Milind Tambe 等人的结果, 并以足球 Agent 为例进行了分析.

## 1 对手模型描述

将 Agent 自身视为 agent1, 被跟踪的 Agent 视为 agent2, 则 agent1 所要建立的对对手 agent2 模型:

$$O = \langle B, D, I_S, I_P, I_W, H \rangle,$$

$O$  在 agent1 的 BDI 模型中应是一组 Belief, 即  $Bel(\text{agent1}, O)$ . 其中

\* 收稿日期: 2000-03-08; 修改日期: 2001-06-20

基金项目: 国家自然科学基金资助项目(69973023; 69733020)

作者简介: 李毅(1974 - ), 男, 河北高阳人, 博士, 主要研究领域为分布式人工智能; 石纯一(1935 - ), 男, 河北秦皇岛人, 教授, 博士生导师, 主要研究领域为人工智能应用基础.

$B: \forall \psi \in B \Rightarrow \text{Bel}(\text{agent2}, \psi)$ , 集合  $B$  包含了 agent1 知道的 agent2 所具有的 Belief.

$D: \forall \psi, \text{Des}(\text{agent2}, \psi) \Rightarrow \psi \in D$ , agent2 的 Desire 集合是  $D$  的一个子集, 即  $D$  包含了 agent2 所有可能的 Desire, 也可将  $D$  表示为  $D_{\text{all}} - D_{\text{ban}}$ ,  $D_{\text{all}}$  表示 Desire 的全集,  $D_{\text{ban}}$  表示 agent2 所不具有的 Desire.

$I_S: \forall \psi \in I_S \Rightarrow \text{Intention}(\text{agent2}, \psi)$ ,  $I_S$  表示可以确定的 agent2 的 Intention, 包括先验的 Intention 和通过观察、通信等手段获得的 Intention.

$I_P$ : 对手的可能 Intention, 是通过已有的 BDI 信息推断出的当前 agent2 可能具有但无法确定的 Intention.

$I_W$ : 表示在 agent2 所有可能的 Intention 中 agent1 所关心的部分, 可称为 Intention 的观察窗口. 不考虑  $I_W$  之外的 Intention (在交互过程中, 对手对于自身的作用往往只在某一特定领域, 因此, 忽略对手在其他领域的 Intention, 并不影响系统运行).

$H: H = \{a_i\}$ ,  $a_i$  表示 agent2 在特定时间所执行的特定动作, 记录一定时间内的对手事件, 如历史信息.

约定  $O$  中的  $B, D, I_S$  满足典型的 BDI 约束关系, 对于  $\forall \psi \in I_P, B, D, I_S \cup \{\psi\}$  也满足典型的 BDI 约束关系. 但模型中,  $B, D, I_P$  无须满足 BDI 约束, 即  $I_P$  集中的 Intention 可以相互矛盾.

## 2 算法描述

模型的运行算法可划分为 5 步骤:

Step1:

初始化操作, 将模型各元素置初始状态:  $B, D, I_S, H$  的值取决于 agent1 对 agent2 的初始了解,  $I_P$  置为  $\emptyset$ ,  $I_W$  取决于 agent1 的自身需要. 初始化后, 系统按 Step2~Step5 的顺序循环执行.

Step2:

获得外界信息, 修正对手模型中的  $B, D, I_S$ . Agent1 获得对手信息可依靠通信和视觉、触觉的感知. 同时信息获取也可分为主动和被动两种类型. 在主动类型中, Agent 利用自身手段获得对手信息, 如向对方询问, 作试探性动作以观察对方反应等; 在被动类型中, Agent 接受对手所表达、传递的信息. 获得信息后, Agent 将其分解为对手模型中的 BDI 信息, 并更新  $B, D, I_S$  元素的取值.

Step3:

根据  $B, D, I$  的约束关系, 修正  $I_P$ , 消除思维状态的内部矛盾.

对  $\forall \psi \in I_P$ ,

if  $\neg \psi \in B$  then  $I_P = I_P - \{\psi\}$  (2.1)

if  $\psi \in B$  then  $I_P = I_P - \{\psi\}$  (2.2)

if  $\psi \notin D$  then  $I_P = I_P - \{\psi\}$  (2.3)

if  $\exists \psi' \in I_S \neg \text{Int}(\text{agent2}, \psi'; \psi)$  then  $I_P = I_P - \{\psi\}$  (2.4)

式(2.1)表示已确认为假的, 即无法实现的状态不能作为 Intention; 式(2.2)表示为真的, 即已实现的状态不可作为 Intention; 式(2.3)表示 Intention 必须有对应的 Desire; 式(2.4)表示 Intention 之间不能相互矛盾.

当  $D$  表示为  $D_{\text{all}} - D_{\text{ban}}$  时, 式(2.3)应为

if  $\psi \in D_{\text{ban}}$  then  $I_P = I_P - \{\psi\}$

Step4:

将可能的 Intention 加入  $I_P$

对  $\forall \psi \in I_W$ , if  $\psi \in I_S$  and  $\psi \notin I_P$  and  $\neg \psi \notin B$  and  $\psi \notin B$  and  $\psi \in D$

and  $\forall \psi' \in I_S, \text{Int}(\text{agent2}, \psi'; \psi)$

then  $I_P = I_P \cup \{\psi\}$

(2.5)

表示将与  $B, D, I_S$  不矛盾的属于  $I_W$  的 Intention 加入到  $I_P$  中.

当  $D$  表示为  $D_{\text{all}} - D_{\text{ban}}$  时, 式(2.5)应为

对  $\forall \psi \in I_W$ , if  $\psi \in I_S$  and  $\psi \notin I_P$  and  $\neg \psi \notin B$  and  $\psi \notin B$  and  $\psi \notin D_{\text{ban}}$

sports fields split figure and  $\forall \psi' \in I_S, \text{Int}(\text{agent2}, \psi'; \psi)$

then  $I_p = I_p \cup \{\psi\}$

Step5:

在  $I_p$  中存在多个对手 Intention,均是可能但不确定的,为了有效跟踪对手,应对其进行排序,以便为自身决策提供依据.对可能意图排序的方法应依不同具体情况而定,通常,我们需要考虑意图实现的代价和对手的历史信息等因素.

对  $\forall \psi \in I_p$ ,可分别计算

$C_\psi$ : 对手实现  $\psi$ 所需要的最小代价,可利用自身的推理机制计算,该方法运用于对手与自身相似的场合;也可利用对手模型中的 belief 等知识计算.

$H_\psi$ : 历史评估,若对手的历史动作序列为  $\{a_1, a_2, \dots, a_n\}$ ,当前的对手动作为  $a'$ ,

设存在序列  $\{k_1, k_2, \dots, k_m\}$ ,其中对  $\forall i, 1 \leq i \leq m, a_{k_i} = a'$ ,则当  $m > 0$  时

$$H_\psi = \left( \sum_{i=1}^m C_{\psi k_i} - C_\psi \right) / m,$$

而当  $m=0$  时,  $H_\psi=0$ .

其中  $C_{\psi k_i}$  表示对手在当前状态下,执行运作序列  $a_{k_{i+1}}, a_{k_{i+2}}, \dots, a_{k_{i+d}}$  后实现  $\psi$ 所需要的最小代价,  $H_\psi$ 表示在历史情况下,对手 Agent 执行动作  $a'$ 后的若干连续动作是否具有实现  $\psi$ 的作用,  $H_\psi$ 越大表示依据历史信息,执行  $a'$ 后具有意图  $\psi$ 的可能性越大.上式中  $d$  为常量,其取值应视具体情况而定.  $d$  值偏小会导致系统易被欺骗行为所迷惑;而  $d$  值偏大会导致系统负担加重,也会使  $H_\psi$ 值与  $a'$ 相关性降低,失去意义.

依据  $C_\psi$ 和  $H_\psi$ 可判定  $\psi$ 实现的可能性的,并将  $I_p$ 中的元素排序.

### 3 算法分析

算法的有效性表现为  $I_p$ 中的可能意图是正确的,且  $I_p$ 中元素的排序也是合理的,本质上则应表现为对 Agent 决策支持的效果.有效性取决于初始值、自身能力和交互方式等内、外因素.

对模型而言,初始  $B, D, I$  值的准确性是保证  $I_p$  集合的有效性的关键因素,而初始值的充分性,则决定了  $I_p$  集合与对方真实意图的差距以及求解速度.当  $B, D, I_s$  的初始值充分且准确时,  $I_p$  集合能准确表示对方的可能意图;否则将无法对  $I_p$  进行正确裁剪(Step3),也无法准确地将  $I_w$  中的意图加入到  $I_p$  集合(Step4).模型中,如何判定  $I_w$  范围,并对  $I_p$  集合中的意图排序,属于 Agent 自身推理能力,小而充分的  $I_w$  以及准确的排序结果,也对模型工作有重要作用.

在运行的 Step2 阶段, Agent 将利用观察和通信等方式获取信息,以修正对手的  $B, D, I_s$ ,该过程的准确,除与自身能力有关外,还取决于 Agent 与对手间的交互方式是协商还是对抗,诚实还是欺骗.在对抗、欺骗的情况下,将导致信息的不完全和不准确,并进而导致  $I_p$  的不准确和排序的不准确,但利用模型提供的 BDI 约束和历史信息等手段可减少这种情况.

Tambe 的 RESC 算法可实时提供对手的当前 Intention,但无法保证其正确性;而 Rao 的算法同时存在多个 Intention,并逐步消简,但在计算过程中难于为决策提供支持.本文提出的算法将对手所有可能意图保存于  $I_p$  中,并将其实时排序,使实时性与准确性得以统一.

以足球 agent 为例加以说明.在足球比赛中,如 agent 自身是守门员,对方一名球员带球运动,此时其意图可为直传、横传、射门、前、后、左、右带球等等.利用 RESC 算法时,只能假设对方具有某一特定意图,并根据假设采取对策,发现错误后再回溯处理.利用 Rao 的算法,初始时会建立多个不同的可能意图树,并根据不断的观察,逐步推断出真实意图,但在判断过程中, Agent 无法得到可能意图,并采取对策.本算法中,  $D = \{\text{直传, 横传, 射门, 前带球, 后带球, 左带球, 右带球}\}$ , 因为 Agent 是守门员,则不会对将球回带等动作感兴趣,故  $I_w = \{\text{直传, 横传, 射门, 前带球, 左带球, 右带球}\}$ .在 Step4 中,分别将  $I_w$  中元素代入式(2.5),其中因对手球员无队友,可知直传、横传等动作不合理,故  $I_p = \{\text{射门, 前带球, 左带球, 右带球}\}$ .进而根据意图实现代价和历史信息对意图排序,如:带球者左侧

空档较大,且此人以往多次向左侧过人,则向左带球的意图排序较高,在决定对策时应重点考虑,但由于也存在传给其他球员等可能性,故 Agent 可作出略微向球门左侧移动的意图.可见本模型可适应复杂的动态对抗下的对手思维状态实时识别,并为 Agent 决策提供有效信息.

此外,区别于其他算法侧重于观察对手行为<sup>[5]</sup>,本算法充分利用 agent 的感知能力,广泛获取外界信息,如对手行动、环境变化,并将其转化为对手模型中的 BDI 信息.在实际环境下,观察行为是感知对手的有效手段,但不是充分的.如在足球比赛中,足球向处于有利于射门位置的无球队员传来,此时该队员尚未做动作,但可能已产生射门意图,而该意图只靠观察对手行为是不会及时获得的,利用本算法则可快速产生对手的射门意图,并作出相应反应.

#### 4 实验分析

为了验证算法的有效性,我们以足球为背景进行了实验分析<sup>[6]</sup>.实验中,AgentK 是足球守门员,AgentF 是足球前锋,AgentF 将从本方后场接到一个长传球后可自由带球运动,目的是在特定时间内将足球射入由 AgentK 把守的球门.对于足球场只考虑半场区域,并划分为 18 个区域,其中,AgentF 所接到的传球的落点随机地分布在 1~6 区域.如图 1 所示.

9	10	11	12	13
8	4	5	6	14
7	1	2	3	15

Goal keeper  
球门.

Fig.1 Sports fields split figure

图 1 球场区域划分图

AgentK 的基本动作集合为{移动,扑球,抢球},AgentF 的基本动作集合为{左脚停球,右脚停球,左脚带球,右脚带球,移动,左脚传球,右脚传球,左脚射门,右脚射门},AgentF 的所有的可能 Intention 为{左脚射门,右脚射门,带球向某区域,向某区域传球}.实验中的 AgentK 和各项指标是不变的.

为了提高 AgentK 守门的成功率,必须准确判断 AgentF 的 Intention,为此引入算法,建立 AgentK 对于 AgentF 的对手模型,并按获得外部信息的方式、 $I_w$  和历史记录等因素分情形进行实验讨论.

实验 1. 研究按获得外部信息的方式对于算法有效性的影响.

A: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 } 1\sim 6\},$$

考虑历史记录;

B: 足球的运动不导致对手模型中 B 的变化,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 } 1\sim 6\},$$

考虑历史记录.

实验分类: A B (各进行 1 000 次)

防守成功次数: 364 297

在 B 情形中,由于未考虑足球的运动,使得大量直接传到区域 2 的球被前锋直接射门,而守门员却无法提前判断以截获足球或作出及时的扑救.

实验 2. 研究  $I_w$  集合对于算法有效性的影响.

C: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 } 2\},$$

考虑历史;

D: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 } 1\sim 6\},$$

考虑历史;

E: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 1~18, 向区域 1~8 传球}\},$$

考虑历史.

实验分类: C D E (各进行 1 000 次)

防守成功次数: 259 370 366

C 中  $I_w$  偏小, 导致防守成功率下降, E 中  $I_w$  偏大, 导致算法计算量增加, 但由于 AgentK 决策未受无关的 Intention 影响, 且机器性能良好, 故对成功防守次数基本没有影响.

实验 3. 研究了历史记录对于算法有效性的影响.

F: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 1~6}\},$$

考虑历史记录;

G: 按算法接收信息,

$$I_w = \{\text{左脚射门, 右脚射门, 带球向区域 1~6}\},$$

不考虑历史记录.

实验分类: F G (各进行 1 000 次)

防守成功次数: 380 355

实验环境为 WindowsNT4.0, 开发环境为 Microsoft VC++5.0.

## 5 结 语

本文提出的对手 Agent 模型建立和对手意图与规划识别算法不仅将 BDI 用于对手模型表示, 而且在推理过程中, 充分使用了 BDI 间的约束关系, 引入了 Intention 窗口的概念, 提高了算法效率和准确性, 并将对手历史信息 and 当前状态相结合, 可表达并发存在的多意图及其关系. 改进了 Anand S.Rao 和 Milind Tambe 等人的结果, 使得在对抗等信息不完全的条件下建立对手模型, 并为 Agent 决策提供支持成为可能.

在未来工作中, 应将该算法引入具有实际应有背景的 MAS 领域, 使其进一步完善, 在此基础上, 考虑对手群体模型的建立算法, 以适用于更广泛的 MAS 环境<sup>[7]</sup>.

## References:

- [1] Haddadi, A., Sundermeyer, K. Belief-Desire-Intention agent architectures. In: O'Hare, G.M.P., Jennings, N.R., eds. Foundations of Distributed Artificial Intelligence. New York: John Wiley&Sons Inc., 1996. 169~185.
- [2] Anand, S., Rao. Multi-Agent mental-state recognition and its application to air-combat modelling. In: Proceedings of the Workshop on Distributed Artificial Intelligence. 1994. 283~304.
- [3] Milind, Tambe. RESC: an approach for real-time, dynamic agent tracking. In: Proceedings of the Joint Conference on Artificial Intelligence. 1995.
- [4] Milind, Tambe. Building agent teams using an explicit teamwork model and learning. Artificial Intelligence, 1999, (110):215~239.
- [5] Hill, R., Johnson, W.L. Situated plan attribution for intelligence tutoring. In: Proceedings of the National Conference on Artificial Intelligence. Menlo Park, CA: AAAI Press, 1994.
- [6] Peter, Stone, Manuela, Veloso. Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork. Artificial Intelligence, 1999, (110):241~273.
- [7] Kamnla, G.A., Milind, Tambe. What is wrong with us? improving robustness through social diagnosis. In: American Association for Artificial Intelligence. 1998. 97~104.

## An Opponent Agent Model Based on BDI\*

LI Yi, SHI Chun-yi

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

E-mail: tiger\_li@sjdd.com.cn

<http://www.tsinghua.edu.cn>

**Abstract:** Developing on opponent model, recognizing intentions and plans are necessary for effective communication in MAS. BDI is an effective way to describe mental state, which can be used to develop a generic opponent mental state model. As current research on the opponent mode focus on implications from unilateral action, they neither express simultaneous intentions and other relations, nor provide support to making decisions making at real time. A generic opponent agent model based on BDI developed in this paper improves the research done by Anand S. Rao and Milind Tambe.

**Key words:** Agent; opponent model; plan recognizing; BDI; MAS

\* Received March 8, 2000; accepted June 20, 2001

Supported by the National Natural Science Foundation of China under Grant Nos.69973023, 69733020

---

### Pacific Graphics 2002

#### 征文通知

Pacific Graphics 2002 (PG2002) 国际会议将于 2002 年 10 月 9~11 日在北京清华大学召开. Pacific Graphics 是国际上有重要影响的图形学学术会议, 会议论文集由 IEEE Computer Society 出版社出版, 会议的优秀论文在 Graphical Models 和 Computer Graphics Forum 上以专辑形式发表. 会议的语言为英语. 有关论文投稿信息, 请查阅会议网页. 欢迎投稿和参加会议.

会议网页: <http://ncc.cs.tsinghua.edu.cn/pg2002>

征文截止时间: 2002 年 4 月 30 日

联系方式: 清华大学计算机科学与技术系 胡事民 邮编: 100084

电话: 010-62782052

传真: 010-62782052

E-mail: [pg2002@tsinghua.edu.cn](mailto:pg2002@tsinghua.edu.cn)