

基于交互式动态影响图的对手建模

罗 键, 武 鹤

(厦门大学 信息科学与技术学院, 福建 厦门 361005)

摘 要: 在充满竞争的环境中, 资源有限导致智能体之间存在利益冲突, 有必要建立对手模型并对其行为进行准确预测, 从而制定对自身有利的策略. 利用交互式动态影响图对未知对手进行建模, 将对手的候选模型保存在模型节点并随时间更新其信度. 结合观测到的对手动作, 在模型空间中利用“观察-动作”序列逐步排除候选模型, 最终判定对手的真实模型. 实验结果表明, 所提出的算法取得了很好的效果, 验证了该算法的实用性.

关键词: 交互式动态影响图; 多智能体; 对手建模; 策略树

中图分类号: TP18

文献标志码: A

Opponent modeling based on interactive dynamic influence diagram

LUO Jian, WU He

(School of Information Science and Engineering, Xiamen University, Xiamen 361005, China. Correspondent: WU He, E-mail: wuhe@stu.xmu.edu.cn)

Abstract: In an environment that is full of competition, the limited resources lead to the conflict of interests which exists among the agents, therefore it is necessary to establish models of the opponent, and to accurately predict its behaviors in order to make strategies for our own advantage. The interactive dynamic influence diagrams are used to model the unknown opponent by keeping the candidate models of the opponent in the model node which can be updated over time. Then combining with the observed actions of the opponent, candidate models in the model space are pruned by using ‘observation-action’ sequences to finally identify the true model of the opponent. The experiment results show the effectiveness and feasibility of the proposed algorithm.

Keywords: interactive dynamic influence diagram; multi-agent; opponent modeling; policy tree

0 引 言

环境资源的日益消耗不可避免地带来了同一资源的竞争. 当环境中有多 Agent 时, 每个 Agent 都必须考虑到其他 Agent 可能采取的行动以及这个行动对自己决策的影响. 《孙子兵法》中记载“知己知彼, 方能百战不殆”, 那么如何在竞争环境中建立并更新对手的模型, 使其更好地了解对手, 并取得自身的最大利益是当前人工智能领域的研究热点之一. 近年来, 有不少国内外学者研究对手建模并取得了一定的成果. 国外方面: Sondberg 等^[1]利用概率图模型(动态影响图)构建了对手的模型; Carmel 等^[2]利用观察到的一系列对手过去的行为, 使用一种无监督的学习算法来推理对手将来的行为; Saha 等^[3]采用切比雪夫多项式来近似对手的效用函数, 进而推理竞争对手在协商过程中的行为; Southey 等^[4]利用贝叶斯概率模型, 根据先验分布推理对手动作的后验概率并制定相

应对策, 应用扑克游戏进行了实验; Sturtevant 等^[5]针对环境中有多对手的情况, 提出了一种“soft-maxn”算法以应对不同类型的对手. 而在国内, 对手建模领域的研究才刚刚起步, 常见于机器人足球比赛. 薛方正等^[6]解决了多机器人对抗系统的对手建模问题, 利用贝叶斯网络分析并判断对手的行为; 顿文力等^[7]在 Agent 的 BDI 模型基础上结合普遍的认识规律, 提出了一种对抗性多机器人系统中的对手建模问题, 并在机器人足球系统上得到了验证; 李毅等^[8]在信息不完全的动态环境下, 跟踪对手的思维状态进行意图与规划识别, 建立对手 Agent 模型并提出了相应的算法, 为决策提供了支持; 曾一峰等^[9]提出了基于朴素贝叶斯学习和互信息的方法来判断对手的真实模型, 并取得了不错的效果. 影响图是一种有效的决策工具^[10], 由 Doshi 等^[11]提出的交互式动态影响图(I-DID)是交互式部分可观测马尔科夫决策过程^[12](I-POMDP)的

收稿日期: 2015-03-14; 修回日期: 2015-07-21.

基金项目: 国家自然科学基金项目(61375070); 福建省重大科技项目(2011H6027).

作者简介: 罗键(1954—), 男, 教授, 博士生导师, 从事多智能体、系统控制与优化等研究; 武鹤(1986—), 男, 博士生, 从事多智能体决策建模的研究.

图形表达形式. I-DID 从个体 Agent 的视角对多 Agent 决策问题进行求解, 能在复杂的交互中对其他 Agent 的动作行为进行预测并动态地更新其他 Agent 的模型信度, 从而合理地制定自身策略, 有效解决多 Agent 决策问题^[13]. 田乐等^[14]针对 I-DID 面临的算法复杂度较高的问题提出了基于 KL 距离的近似解法, 大大提高了 I-DID 的求解速度, 而后提出了一种通信交互式动态影响图^[15-16], 将通信动作引入 I-DID, 使得传统的 I-DID 具有通信的能力, 让 Agent 可以分享自己的观察. 可以看出, 利用 I-DID 建立起对手模型是一种理想并且行之有效的手段.

首先, 本文采用交互式动态影响图对未知对手进行建模, 其特有的模型节点可以用来保存对手的所有候选模型, 同时随着时间的推移利用模型更新链对候选模型的信度进行更新; 然后, 提出一种用于识别对手的真实模型的算法, 在众多候选模型中, 根据策略树的差异, 利用“观察-动作”序列逐步排除大部分候选模型, 判断出对手使用的真实模型, 掌握对手的行动策略; 最后, 在两个著名问题——老虎问题和机器维修问题上对算法的有效性进行两方面的验证, 在找到对手真实模型的同时提高己方的收益值.

1 交互式动态影响图

1.1 交互式动态影响图概述

交互式影响图 (I-ID) 将一般意义上的影响图扩展到多智能体领域. 除了一般影响图所具有的机会节点、决策节点和效用节点以外, 交互式影响图还包含一种新型的节点称为模型节点. 除了模型节点外, I-ID 不同于一般影响图的地方还在于其拥有一个“策略链”, 图 1 中以虚线表示. 它表示 j 的模型给定时的 j 的动作分布情况. 对于两个以上智能体的情况, 只需要为每个智能体添加一个模型节点、一个机会节点和一个策略链即可. I-DID 再次扩展 I-ID 到 T 个时间片的动态环境上, 两个时间片的交互式动态影响如图 1 所示.

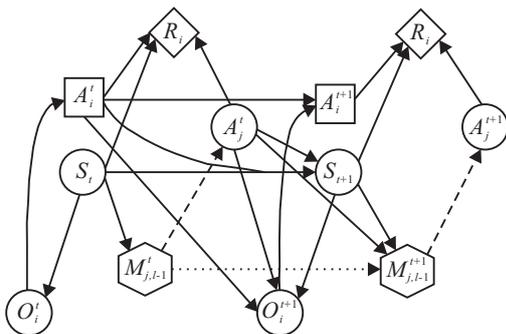


图 1 交互式动态影响 (两个时间片)

为了表述清晰, 图 1 中只包含两个智能体, 位于 l 层的主观智能体 i 和位于 $l-1$ 层的智能体 j . 模型节点所含有的正是由上层智能体 i 描述的下层智能体 j 的所有模型. 而从机会节点 S 指向模型节点 $M_{j,l-1}$ 的连

接表示的是 i 对 j 的模型的信度. 具体而言, 就是机会节点 $\text{Mod}[M_j]$ 的条件概率表 (CPT) 中的概率分布. 智能体 j 的模型为 $m_{j,l-1} = \langle b_{j,l-1}, \hat{\theta}_j \rangle$. 其中: $b_{j,l-1}$ 为该模型的信度; $\hat{\theta}_j$ 为该智能体的框架, 包含决策节点、观察节点和效用节点. 每一个模型 $m_{j,l-1}$ 可以是一个 $l-1$ 层的 I-DID 或者是一个 0 层的 DID. 模型的解是对 j 的行为的预测并通过一个虚线连接到机会节点 A_j , 这个连接就是上文提到的“策略链”. 在 I-DID 中, 机会节点 A_j 、模型节点 $M_{j,l-1}$ 和其他节点联合显示了智能体 j 的动作是如何影响智能体 i 的决策过程的. I-DID 的扩展包含了模型节点随着时间更新的过程. 图 1 中由 $M_{j,l-1}^t$ 到 $M_{j,l-1}^{t+1}$ 的点线称为模型更新链, 随着智能体 j 在每个时间步执行动作并获得观察, j 的模型也应该逐步更新. 在 t 时刻, 对于每个模型 $m_{j,l-1}^t$, 它的最优解也许包含所有可能的动作, 与此同时, 智能体 j 可能得到任意一种可能的观察. 这样, 在 $t+1$ 时刻, 更新后的模型集合可能包含最多 $|M_{j,l-1}^t| |A_j| |\Omega_j|$ 个模型. 其中: $|M_{j,l-1}^t|$ 为 j 在 t 时刻的模型数量, $|A_j|$ 为所有可能的动作个数, $|\Omega_j|$ 为所有可能获得的观察个数. 例如, 在 t 时刻, 模型节点包含两个模型, 每个模型有一个最优动作, 随后在 $t+1$ 时刻, 它可能获得两个观察. 于是可以通过计算得到, 此时模型的个数为 $|M_{j,l-1}^t| |A_j| |\Omega_j| = 2 \times 1 \times 2 = 4$ 个. 可以利用标准的依赖性连接和机会节点实现模型更新的整个过程, 与此同时将一个 I-DID 转换为一个一般意义上的动态影响图.

1.2 I-DID 的解

求解一个 l 层 T 个时间片的 I-DID 采用的是一种自底向上的方法. 为了解出智能体 i 的第 l 层的 I-DID, 必须首先解出所有低一层 ($l-1$ 层) 的智能体 j 的模型. 而 $l-1$ 层的模型 $m_{j,l-1}$ 的解就是智能体 j 的策略, 也就是给定 j 的信度 $b_{j,l-1}$ 后, 从 j 的观察 O_j 到最优动作 A_j 的映射. 将 j 的每一时间步的最优决策赋值给机会节点 A_j , 并根据 j 的每个“最优动作-观察”序列扩展 $\text{Mod}[M_{j,l-1}]$ 中 j 的模型. 将这个过程用于每个时间片, 每个 $l-1$ 层的 j 的模型, 从而获得完整的扩展后的 l 层模型. 求解 I-DID 的计算复杂度主要在于随时间呈指数增长的 $l-1$ 层 j 的模型. 虽然整个模型空间非常大, 但是不需要具体求解出每一个模型, 比如行为等价模型^[16], 下面给出行为等价的定义.

定义 1 行为等价 (BE). 对于模型空间中的两个模型 $m_{j,l-1}$ 和 $m'_{j,l-1}$, 如果满足条件

$$\text{OPT}^T(m_{j,l-1}) = \text{OPT}^T(m'_{j,l-1}), \quad (1)$$

则称这两个模型为行为等价模型, 其中 $\text{OPT}^T(\cdot)$ 为模型在 T 个时间片上的最优解.

将所有行为等价的模型归为一类, 称为行为等价类. 模型的解通常由策略树表示, 由于行为等价的模型具有完全相同的行为, 即其策略树是完全相同的,

而对于主观建模的智能体 i 而言, 更关心的是客观被建模的智能体 j 的动作, 以此来规划自己的动作, 因此没有必要区分行为等价模型. 要确定 j 的真实模型, 只需找到其真实模型所在的行为等价类即可. 而对于真实模型不在候选模型中的情况, 同样可以找到其在候选模型中的行为等价模型来代替真实模型.

2 对手真实模型识别

I-DID 中的底层模型可能有很多, 而找到真实的模型就需要进一步的判定. 本文采用的方法是利用通信过程中对方 (Agent j) 传递的观察以及上层智能体 i 观测到的对方的动作来判断其真实模型. 这里的智能体都是理性的, 其执行的最优动作均由最大效用理论得到. 因此, 底层不具有建模能力的智能体 j 的策略是可以由初始信度求解得到的. 由此可知, 若得到了对方传来的观察, 则可以推出其应该执行的最优动作, 与观测到的对方动作不一致的模型就可以删除.

2.1 实例

下面以老虎问题为例说明本文算法. 智能体 j 的 3 棵策略树 ($T = 3$) 如图 2 所示, 其初始信度为 TL = 0.2, 0.4, 0.85, 分别表示其认为老虎在左的概率.

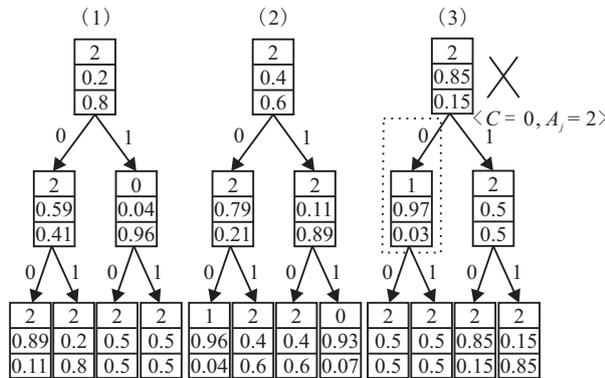


图 2 一个时间片后排除模型 (3)

在图 2 中: 每个节点最上方的数字表示其最优动作, 0 表示开左门, 1 表示开右门, 2 表示听; 中间和下方的数字分别表示其认为老虎在左和在右的概率; 连接节点之间的线段上的数字表示 j 的观察, 0 表示听到老虎在左, 1 在右. 假设 j 的真实模型存在于这 3 个模型之中, 且为模型 (2), 对应策略树 (2). 一个时间片之后, i 通过通信^[16] 获得了来自 j 的观察 0, 同时观测到了 j 的动作为 2. 由于 j 的观察是 0, 由 3 棵策略树可以看出, 前两棵树对应的节点的最优动作都是 2, 而第 3 棵树对应的节点的最优动作是 1, 如图 2 中虚线框中的部分所示. 图 2 中: C 为通信的内容, A_j 为观察到的 j 的动作. 由于 i 观测到的 j 的动作是 2, 第 3 棵树是真实模型的可能性被排除了. 如果第 3 棵树是真实模型的策略树, 则他会选择最优动作 1 而不是 2. 同样的, 在下一个时间片, i 通过通信获得了来自 j 的观察 0, 以及 i 自己观测到的 j 的动作 1, 利用上面相同的方法将策略树 (1) 排除, 见图 3 虚线框处, 最终得到

了 j 的真实模型为模型 (2).

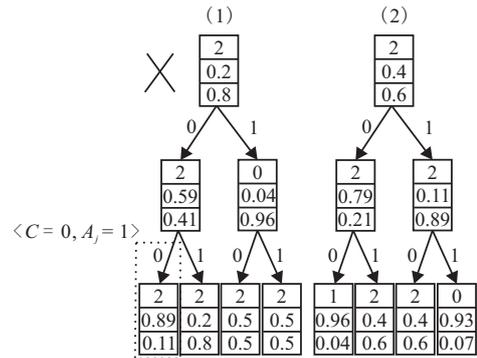


图 3 两个时间片后排除模型 (1)

2.2 算法描述与复杂度分析

下面给出判定模型真实模型的具体算法.

输入: I-DIDs ($l \geq 1$), 时间片数 T , 初始模型 $m_{j,0}$;

输出: T 个时间片上 i 的收益值.

1) 扩展阶段:

For t from 0 to $T-1$,

If $l \geq 1$,

Then, 首先求解其他 Agent 的底层模型; 对于节点 $M_{j,l-1}^t$ 中每个模型 m_j^t , 递归调用算法来处理 $l-1$ 层的 I-DID (或者 DID), 求解 t 时刻的最优决策 $a_j = \text{OPT}(m_j^t)$, 利用 a_j 和观察值 o_j 更新模型信度 $b_j^{t+1} \leftarrow \text{SE}(b_j^t, a_j, o_j)$, 得到以 b_j^{t+1} 作为初始信度的模型 m_j^{t+1} , 将得到的每个 m_j^{t+1} 赋值到模型节点 $M_{j,l-1}^{t+1}$; 添加相邻的两个模型节点 $M_{j,l-1}^t$ 与 $M_{j,l-1}^{t+1}$ 的模型更新链接; 添加 $t+1$ 时间片的机会、决策、效用节点, 以及相应的依赖链接; 为每个机会节点和效用节点建立 CPT.

2) 求解阶段:

If $l \geq 1$,

Then, 利用模型节点、策略链接和模型更新链接将 I-DID 扩展为 DID; 用向前搜索算法计算扩展后的 DID,

For t from 0 to $T-1$,

利用通信得到的对方观察删去候选模型空间中最优动作和观察到的对方动作不一致的模型; 逐步删去非真实模型并最终得到 i 的策略树和收益值.

应用本文算法将在有限步内找到对手的真实模型, 且有如下定理成立.

定理 1 假设对手的真实模型在模型空间中, 即 $m_j^* \in M_j$, 并且真实模型的初始信度不为 0, 则使用本文算法一定能找到该真实模型.

证明 在部分可观测的随机过程中, 使用“动作-观察”序列更新后的信度是一个对历史进行充分统计的结果. 根据初始信度就能预测智能体未来的行为, 所以候选模型的解即策略树是预先求解出来的且模型空间中隶属于不同行为等价类的模型的策略树也不相同. 因此, 利用本文算法逐步删去与得到的“动

作-观察”序列不一致的模型就能得到其真实模型。□

本文算法的复杂度取决于被比较的策略树的大小,且有如下命题成立。

命题 1 复杂度分析. 利用本算法进行真实模型的判定,在最坏的情况下复杂度为 $O(|\Omega|^{T-1} |m_{j,l-1}^0|)$. 其中: $|m_{j,l-1}^0|$ 为初始模型个数, T 为总时间片数, $|\Omega|$ 为智能体观察的个数。

证明 算法利用观察到的对手动作与候选模型 T 个时间片上的策略树进行比较,算法的复杂度与被比较的策略树的大小有关. 在 T 个时间片上,一棵策略树共有节点数 $\frac{|\Omega|^T - 1}{|\Omega| - 1}$, 初始模型空间中模型个数为 $|m_{j,l-1}^0|$. 因此,在最坏情况下,算法需要遍历所有策略树上的节点,复杂度为 $O(|\Omega|^{T-1} |m_{j,l-1}^0|)$. □

3 算法仿真与分析

在 Windows 下,使用 VC++ 编程,并调用 HUGIN Researcher 进行相关的推理. 首先,应用多 Agent 老虎问题验证算法的有效性. 对老虎问题进行一些有趣的改进. 这里采用的收益函数不同于文献 [11]. 具体而言,考虑两个智能体 i 和 j , 智能体 j 为合作性智能体,而智能体 i 为自私型智能体,力求自身的收益最大化,他认为 j 是自己的对手,会影响到自己的收益值. 就收益值而言,当两个智能体同时打开带有宝藏的门时,由于要分享宝藏, i 获得的收益值只有正常的一半,而当对手打开带有老虎的门时,对手的伤害值的一半将转为自己的收益,详细的收益函数见表 1.

表 1 老虎问题中智能体 i 和 j 的收益函数

| $\langle a_i^t, a_j^t \rangle$ | TL | TR |
|--|-------|-------|
| $\langle \text{OR}, \text{OR} \rangle$ | 5 | -50 |
| $\langle \text{OL}, \text{OL} \rangle$ | -50 | 5 |
| $\langle \text{OR}, \text{OL} \rangle$ | 60 | -105 |
| $\langle \text{OL}, \text{OR} \rangle$ | -105 | 60 |
| $\langle L, L \rangle$ | -0.5 | -0.5 |
| $\langle L, \text{OR} \rangle$ | -6 | 49 |
| $\langle \text{OR}, L \rangle$ | 10.5 | -99.5 |
| $\langle L, \text{OL} \rangle$ | 49 | -6 |
| $\langle \text{OL}, L \rangle$ | -99.5 | 10.5 |

假设初始时刻,智能体 j 有很多个模型,这些模型的差异在于其对环境的信度(老虎在左的概率)不同,而其真实的模型只有一个. 应用本文的算法判断这些模型中哪个才是 j 真正的模型,考虑 j 的初始模型个数分别为 50、30 和 10,随机分配 j 的初始模型的信度,在模型空间中任选一个模型作为 j 的真实模型. 经过 10 个时间片后,不是 j 真实的模型将被陆续删除,留下的则是与真实模型为一个行为等价类的模型,具体情况见图 4.

与此同时,在每个时间片上分别记录下此时智能体 i 的收益情况(每个点为执行算法 500 次获得的平均收益值),并将其与另外两种情况进行比较: 其一是利用贝叶斯方法进行真实模型的判断,其二是环境中

只存在真实的模型. 3 种情况收益值的比较见图 5.

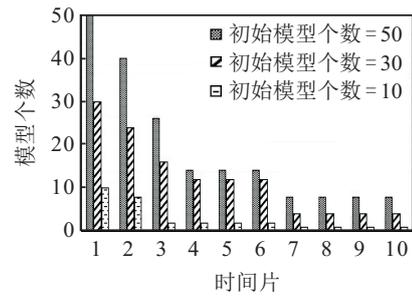


图 4 老虎问题模型个数随着时间片减少的情况

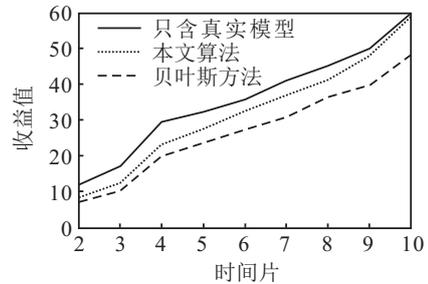


图 5 老虎问题收益值的比较

结合图 4 与图 5 可以看出,在前 3 个时间片上,当智能体 i 还区分不出哪个才是 j 的真实模型时,其取得的收益值并不高,而到了第 4 个时间片,由于 i 对 j 的慢慢了解,模型空间中那些不是真实的模型被慢慢删除,剩余的模型个数越来越少,其收益值慢慢提高,随着时间的推移,从第 7 个时间片开始,模型数量不再发生变化, i 知道了 j 的真实模型 (j 的行为等价类),其收益值逐渐接近于只含有真实模型的情况. 而使用贝叶斯方法进行模型判定的速度较慢,不能很快地找到真实模型,所以获得的收益值比较低. 可以看出,本文算法是非常有效的. 将算法应用于多智能体的机器维修问题 [17], 同样假设在初始时刻,智能体 j 有很多个初始模型 (50,30,10), 这些模型不同表现的初始信度不同,同样为随机生成,任取模型空间中的一个模型作为真实模型. 应用本文算法逐步剔除假模型,留下的则是真实模型,或者是真实模型的行为等价类. 如图 6 所示,算法在第 3 个时间片就取得了很好的效果,接下来的时间片模型数量没有进一步的减少,事实上剩下的模型为真实模型的行为等价类. 记录下在每个时间片智能体 i 获得的收益值,同样与另外两种情况进行比较,如图 7 所示.

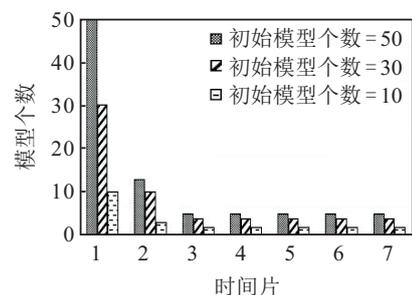


图 6 机器维修问题中模型个数随时间片减少的情况

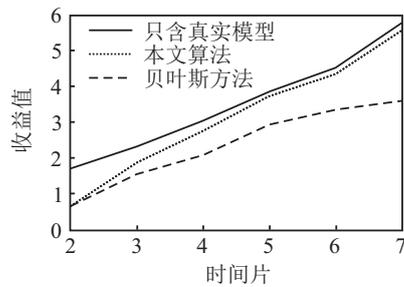


图7 机器维修问题收益值的比较

从图7中可以明显看出,在前两个时间片上,当智能体 i 对 j 的真实模型还知之甚少时,其取得的收益值并不高,几乎接近贝叶斯方法判定真实模型的情况,而在3个时间片之后,由于 i 知道了 j 的真实模型(j 的行为等价类),其收益值得到了显著的增长并接近于只含有真实模型的情况。同样,使用贝叶斯方法在较少的时间片内效果仍然较差。

4 结 论

本文以交互式动态影响图为工具,研究了竞争环境中的对手建模问题,提出了一种利用对手的观察和动作来逐步判定对手的真实模型的算法,并将此算法应用在多智能体老虎问题和机器维修问题上,取得了比较好的效果。在众多候选模型中确定了对手的真实模型或其所在的行为等价类。本文中使用的每个底层模型信度是不同的,而框架(framework)是一样的。今后的工作将继续研究,在对手的真实模型不一样时,算法该如何改进以获得对手的真实模型。

参考文献(References)

- [1] Sondberg-Jepesen N, Jensen F V, Zeng Y F. Opponent modeling in a PGM framework[C]. Proc of the 12th Int Conf on Autonomous Agents and Multiagent Systems. Saint Paul: IFAAMAS, 2013: 1149-1150.
- [2] Carmel D, Markovich S. Learning models of intelligent agents[C]. Proc of the 13th National Conf on Artificial Intelligence. Portland: AAAI, 1996: 62-67.
- [3] Saha S, Biswas A, Sen S. Modeling opponent decision in repeated one-shot negotiations[C]. Proc of the 4th Int Conf on Autonomous Agents and Multiagent Systems. Utrecht: ACM, 2005: 397-403.
- [4] Southey F, Bowling M, Larson B, et al. Bayes' bluff: Opponent modelling in poker[C]. Proc of the 21st Conf in Uncertainty in Artificial Intelligence. Edinburgh: AUAI Press, 2005: 550-558.
- [5] Sturtevant N, Bowling M. Robust game play against unknown opponents[C]. Proc of the 5th Int Conf on Autonomous Agents and Multiagent Systems. Hakodate: ACM, 2006: 713-719.
- [6] 薛方正, 方帅, 徐心和. 多机器人对抗系统仿真中的对手建模[J]. 系统仿真学报, 2005, 17(9): 2138-2141.
(Xue F Z, Fang S, Xu X H. Opponent modeling in

- adversarial multi-robot system simulation[J]. J of System Simulation, 2005, 17(9): 2138-2141.)
- [7] 顿文力, 孟庆春, 庄晓东. 对抗性多机器人系统对手建模的研究[J]. 计算机应用研究, 2004, 21(3): 53-55.
(Dun W L, Meng Q C, Zhuang X D. Study on opponent modeling in adversarial multi-robot system[J]. Application Research of Computers, 2004, 21(3): 53-55.)
- [8] 李毅, 石纯一. 基于BDI的对手Agent模型[J]. 软件学报, 2002, 13(4): 643-648.
(Li Y, Shi C Y. An opponent agent model based on BDI[J]. J of Software, 2002, 13(4): 643-648.)
- [9] Zeng Y F, Doshi P. Model identification in interactive influence diagrams using mutual information[J]. Web Intelligence and Agent Systems, 2010, 8(3): 313-327.
- [10] 吴江, 宋晗, 周锐, 等. 基于扩展影响图的超视距空战辅助决策方法[J]. 控制与决策, 2010, 25(11): 1669-1674.
(Wu J, Song H, Zhou R, et al. Extended influence diagram based decision aiding approach for beyond visual-range air combat[J]. Control and Decision, 2010, 25(11): 1669-1674.)
- [11] Doshi P, Zeng Y F, Chen Q Y. Graphical models for interactive pomdps: Representation and solutions[J]. J of Autonomous Agents and MultiAgent Systems, 2009, 18(3): 376-416.
- [12] Piotr J G, Doshi P. A framework for sequential planning in multi-agent settings[J]. J of Artificial Intelligence Research, 2005, 24(1): 49-79.
- [13] Zeng Y F, Pan Y H, Mao H, et al. Improved use of partial policies for identifying behavioral equivalence[C]. Proc of the 11th Int Conf on Autonomous Agents and Multiagent Systems. Valencia: IFAAMAS, 2012: 432-439.
- [14] 田乐, 罗键, 曹浪财, 等. 基于KL距离的交互式动态影响图近似算法[J]. 系统工程与电子技术, 2013, 35(1): 207-211.
(Tian L, Luo J, Cao L C, et al. Approximate algorithm of interactive dynamic influence diagrams based on KL distance[J]. Systems Engineering and Electronics, 2013, 35(1): 207-211.)
- [15] Tian L, Luo J, Huang Z L. Communication based on interactive dynamic influence diagram in cooperation multi-Agent systems[C]. The 8th Int Conf on Computer Science and Education. Sri Lanka: IEEE Press, 2013: 56-61.
- [16] Wu H, Luo J, Tian L. Exploring efficient communication in interactive dynamic influence diagrams[C]. Proc of 2013 Chinese Intelligent Automation Conference. Yangzhou: Springer Press, 2013: 317-324.
- [17] Zeng Y F, Doshi P. Exploiting model equivalences for solving interactive dynamic influence diagrams[J]. J of Artificial Intelligence Research, 2012, 43(1): 211-255.