

基于历史行为的 agent 联盟策略

蒋建国, 尹 翔, 夏 娜, 苏兆品
(合肥工业大学计算机与信息学院, 安徽合肥 230009)

摘 要: 协商是多 agent 系统中联盟形成的重要手段, 很少有对协商初始提议的讨论. 提出了一种基于个体历史行为确定协商起点的方法, 根据 agent 在历史任务中的各维能力贡献率, 计算其在当前任务中的期望获利因子, 并据此给出协商时的初始提议, 在此基础上设计了一种联盟形成策略, 试验表明了该方法的有效性.

关键词: 初始提议; 能力贡献率; 期望获利因子; 历史行为

中图分类号: TP301 文献标识码: A 文章编号: 0372 2112 (2007) 08 1485 05

Strategy of Agent Coalition Based on Historic Behavior

JIANG Jian guo, YIN Xiang, XIA Na, SU Zhao pin

(Department of Computer and Information Science, Hefei University of Technology, Hefei, Anhui 230009)

Abstract: Negotiation is an important method to coalition formation in MAS. Most of existing work seldom studies how to generate initial proposals. A strategy based on historic behavior to decide initial proposals is proposed. By evaluating each agent's contribution ratio in historic tasks, the expected profit factor to the current job can be attained and based on it the initial proposal is obtained. Moreover, a coalition formation strategy is presented. Experimental results demonstrate that this method is efficient.

Key words: initial proposals; contribution ratio in capabilities; expected profit factor; historic behavior

1 引言

由于单个 agent 无法独立完成某个任务或者通过多个 agent 协作能提高求解效率, 获得额外效用, 因此 agent 为增大自身效用, 会选择与其他 agent 结成联盟. 作为自利的个体, agent 加入联盟的目的是获得更多的收益, 联盟形成主要考虑如何在联盟内 agent 间实现效用的合理划分, 而协商则是效用分配中的核心问题.

经典的协商策略是基于博弈论的^[3-6], 主要考虑联盟收益划分的公平性和稳定性, 其最明显的不足就在于假设 agent 有无限的计算资源并且所有可能的结果都是已知的, 这就导致了算法的复杂度是指数级的, 在实际系统中难以实现.

文献[7]给出了一种在超加环境下通过协商来形成联盟的算法, 当有形成联盟的可能时, 双方的初始策略是将额外效用均分, 再分别同第三方进行协商, 并将此结果作为下一步协商时要价的“筹码”, 为自己争取更多的收益. 此方法更加符合现实情况, 但由于在协商中缺乏有效的指导, 尤其是初始时对额外效用的均分没有考虑个体和任务的实际情况, 需要很长的时间才能得到一

个双方都满意的结果.

文献[8]中为了提高协商的效率, 建立了以 Bayes 学习为基础的学习系统(LS), agent 以自己的信息状态和接收到的提议, 利用 LS 产生下一时刻的反提议, 通过不断地学习使自身的行为更加理性, 从而提高协商效率. 但此方法需要一系列的前提假设和先验概率分布, 这在实际系统中是难以满足的.

迄今为止, 大部分针对协商的讨论主要集中在如何改进协商中双方的策略以提高效率, 减少交互次数^[9-11], 虽然文献[12, 13]已经指出, 协商的初始提议对协商过程和结果有着非常重要的影响, 但相关研究还较少. 基于此, 本文给出一种新的方法, 考虑个体 agent 在历史任务中的行为来预测其在当前任务中的表现, 得到期望获利因子, 并以此为依据产生模糊数作为协商的起点, 为协商提供有效的指导, 从而大大缩短协商时间, 促使联盟快速有效地形成.

2 协商初始提议的确定

定义 1 $\Theta_i = \langle \pi_i^1(1), \pi_i^1(2), \dots, \pi_i^1(r) \rangle$ 是 agent A_i 在完成任务 t_i 时的能力发挥率. 其中 $\pi_i^1(k) = b_i^1(k) / b_i^0(k)$,

($k = 1, 2, \dots, r$), $b_i^l(k)$ 是 A_i 完成任务 t_l 时在第 k 维能力上实际发挥的能力值. 特别的, Θ_i 表示 A_i 在当前任务中能力发挥率的预测值.

定义 2 期望获利因子 α_i 表示 agent A_i 在联盟形成时期望自己在任务完成之后所获收益与联盟整体效用的比值, 此因子正比于 Θ_i .

2.1 Agent 各维能力贡献率的预测

一个 agent 在联盟中所获得的收益主要应该取决于它对完成任务所做贡献的大小, 也即在任务完成过程中, 此 agent 所贡献的能力和任务需求能力的比值, 比值越大代表此 agent 的相对能力越强, 在协商中也处于有利地位, 其所获得的收益应该越多. 而在联盟形成阶段, 各 agent 在执行此次任务时所实际贡献的能力无法具体确定, 只能综合历史信息来进行预测.

为了对每个 agent 在本次任务中的表现进行预测, 需要记录整个系统在一段历史任务中的行为, 我们对其建立一个数据库, 定义为历史任务集.

定义 3 历史任务集是指 MAS 在一段时间内所完成任务的集合, $T_h = \{t_h^1, t_h^2, \dots, t_h^L\}$, L 是系统所记录历史的长度. 每个 $t_h^l (l = 1, 2, \dots, L)$ 对应一个联盟 C_h^l , 称为历史联盟, 对于 C_h^l 的每个成员 A_i 记录其在本次任务中所发挥的能力值 B_i^l .

历史任务集以任务作为索引, 并由系统中一组不会加入任何联盟的个体, 称之为专家组所保存. 专家组成员根据自己所保存的数据库对当前联盟中的每个 agent 在本次承接任务时的能力贡献率进行预测, 具体步骤如下:

Step 1 遍历数据库得出当前联盟中的成员曾在历史任务集中参与了哪些任务, 假设为 p 个, 注意: 当前联盟的成员不一定参与所记录的全部历史联盟, 因此 p 可能不等于 L .

Step 2 由历史记录得到 agent A_i 在所参与的每个历史任务中实际发挥的能力向量 $B_i^l = \langle b_i^l(1), b_i^l(2), \dots, b_i^l(r) \rangle$, ($l = 1, 2, \dots, p$), 并计算出其在相应任务中每维能力发挥率 $\Theta_i^l(k) = \langle \pi_i^l(k), \pi_i^l(k), \dots, \pi_i^l(k) \rangle$, ($l = 1, 2, \dots, p$).

Step 3 对每个 Θ_i^l 赋予相应的权值 ω_l , $\sum_{l=1}^p \omega_l = 1$, 可以根据 t_h^l 在 T_h 中的位置来确定 ω_l , 越靠前的任务其对应权值越小, 即 $\omega_1 < \omega_2 < \dots < \omega_p$. 得出个体 A_i 在当前联盟中每维能力发挥率的预测 $\Theta_i(k) = \langle \pi_i(k), \pi_i(k), \dots, \pi_i(k) \rangle$, 其中 $\pi_i(k) = \sum_{l=1}^p [\omega_l \times \pi_i^l(k)]$ ($k = 1, 2, \dots, r$).

2.2 Agent 的整体贡献因子

由上面的各维能力贡献率得到当前任务中 agent A_i

能力发挥的预测值: $B_i^c = \langle b_i^c(1), b_i^c(2), \dots, b_i^c(r) \rangle$, $b_i^c(k) = b_i^l(k) \times \pi_i^l(k)$, $k = 1, 2, \dots, r$, 其中 $b_i^l(k)$ 为 A_i 的第 k 维能力值.

设 $W = \langle w_1, w_2, \dots, w_r \rangle$ 是需求向量权重矩阵, 由任务本身的性质决定, 表示任务需求向量每一维的相对重要性, 那么 A_i 对任务求解的整体贡献因子为 $c_i = B_i \times W^T$.

2.3 期望获利因子

A_i 的期望获利因子为: $\alpha_i = \frac{c_i}{\sum_{A_j \in C} c_j}$, 则 A_i 的预期收益

$O_i = \alpha_i \times V(C)$. 在接下来的协商中, A_i 将以 O_i 为基础产生初始提议.

2.4 初始提议

上一步得到的预期收益 O_i 是一个确定的值, 系统中每个 agent 都知道此值, 若协商双方都以此值作为初始提议, 则不需要协商过程. 另一方面, agent 作为自利的个体, 总是追求自身利益的最大化, 这使得 agent 都会给出对自己有利的提议, 具体表现为一方会尽量压低报价而另一方则会努力提高报价, 但这是在一定范围内的选择, 若提议过于“离谱”, 产生了“非严肃报价”^[14], 有可能不会形成联盟, 谁也得不到收益, 这是双方不愿看到的.

因此, 在这里初始报价的基准是 O_i 但并不一定就等于 O_i , 即具有一定的不确定性, 我们可以用模糊集理论中模糊数^[15]的概念来描述.

定义 4 如果模糊集 \tilde{M} 是定义在实数集 R 上的正则凸模糊集, 且满足下列条件: (1) 存在点 $x_0 \in R$, 使得 $\mu_{\tilde{M}}(x_0) = 1$. 此时 x_0 被称为 \tilde{M} 的平均值. (2) $\mu_{\tilde{M}}(x)$ 是左、右连续的, 则称 \tilde{M} 为模糊数. 模糊数 \tilde{M} 的含义是“近似于 x_0 的实数”.

在协商开始时, agent A_i 和 A_j 将分别产生模糊数 \tilde{L} 和 \tilde{R} , 满足 $\mu_{\tilde{L}}(O_j) = \mu_{\tilde{R}}(O_i) = 1$, 表示 \tilde{L} 以 O_j 为平均值在一定范围内变化, 而 \tilde{R} 以 O_i 为平均值在一定范围内变化. 此时虽然 O_i 和 O_j 是公共知识, 但 \tilde{L} 和 \tilde{R} 却是各自的私有信息, 对方并不知道. 接下来, A_i 和 A_j 在各自的模糊数范围内选取一定值作为初始提议.

3 基于历史行为的联盟形成策略

Step 1 个体从外界感知到任务, 我们称此 agent 为盟主 agent, 并向系统中的其余 agent 广播有关消息, 有合作意愿的 agent 进行回复.

Step 2 作为联盟的发起者, 盟主 agent 建立多个协商线程 $TH = \{thread_1, thread_2, \dots\}$, 每个线程 $thread_j$ 对

应于一个有意愿加入联盟的 agent A_j , 且 TH 并发执行。

Step 3 $thread_j$ 向 A_j 发出形成联盟的邀请, 按上节的方法给出初始提议 $prop_j$, 主要包括 A_j 所能获得收益 u_j , 并等待答复。

Step 4 基于自己的期望收益, A_j 评价所收到的 $prop_j$, 若满意, 则接收邀请加入联盟, $thread_j$ 结束, 转 step 6; 否则给出初始反提议 $anti-prop_j$ 。

Step 5 双方进入协商阶段, 即不断产生提议与反提议, 直到形成一个双方都能接受的提议, 协商成功; 或者产生了双方都不能接受的提议, 退出协商。协商线程 $thread_j$ 结束。

Step 6 TH 中的所有协商线程执行完毕, 协商成功的 agent 构成潜在联盟。盟主 agent 评价所得到的潜在联盟, 若能完成任务, 则形成该联盟; 若不能完成任务, 则宣布此次联盟形成失败。

4 分析与讨论

4.1 历史任务中能力贡献率的模糊评判

由上面的过程可以看出, 期望获利因子的获得主要取决于 agent 在所参与的历史任务中实际发挥的能力向量 B_i^t 。由于任务执行是由多种因素共同起作用的结果, 例如: 自身的能力、工作的态度、任务的性质等。同时, 联盟中各 agent 的行为和决策是自治的, 其根本目的是实现自己利益的最大化。这些都导致了 agent 实际发挥的能力向量可能是一个模糊的范围, 只能用模糊语言来评价, 如“高、较高、较低、低”。因此, B_i^t 可以在任务完成之后, 由专家组成员根据每个 agent 在任务执行过程中的实际表现, 通过模糊评判的方法来得到。作为中立的可信第三方, 专家组成员将不会对任何个体 agent 存有偏见, 这就保证了评判的公平公正性, 得到的结果也会接近于真实情况。评判时, 应主要从以下几个方面考虑个体 agent 的表现:

(1) A_i 本身能力的大小: 个体 agent 能力越强, 对其评价越优。

(2) A_i 在任务中的具体表现: 表现越优秀, 对其评价越高。这里的优秀有两个方面的含义:

(a) 当有形成联盟的可能时, 个体越早加入联盟, 其表现越优秀。这保证了联盟形成的时效性。

(b) 在执行任务时, 个体工作越努力越积极, 其表现越优秀。这保证了系统的效率。

(3) A_i 的经验: A_i 在历史中参加的任务越多, 其经验越丰富, 对其评价越高。

可以看出, 各 agent 为了获得更多的利益, 会积极地加入任务, 在工作中也会竭尽全力, 以求得专家组的认可, 这样在进行模糊评判时自己就可得到较好的评价, 在下次协商中占据有利地位, 进而获得多的收益。这也

体现了本文策略对个体行为具有激励作用。

4.2 初始提议对协商步数的影响

假设作为盟主的 agent A_i 和有合作意愿的 agent A_j 就 A_j 应得的收益进行协商, 双方在不完全信息下进行协商, 即互不知道对方将采取何种策略。协商过程可以描述为: $NEG_i = \{O_i^1, O_i^2, \dots, O_i^{2m-1}\}$ 表示在 t_e 时刻 A_i 向 A_j 发送的协商提议, $NEG_j = \{O_j^1, O_j^2, \dots, O_j^{2m}\}$ 表示在 t_{e+1} 时刻 A_j 向 A_i 发出的反提议, 则 $NEG = NEG_i \cup NEG_j$ 构成了整个协商过程。我们对这一过程建立一个数学模型, 设每一时刻只有一方发出一个协商提议, 经过 $2m$ 步之后, 达到一个双方均能接受的提议, 则:

$$\begin{cases} O_i^{t+} - m \times \delta_1 < O_j^{t-} - m \times \delta_2 \\ O_i^{t+} + (m+1) \times \delta_1 > O_j^{t-} - (m+1) \times \delta_2 \\ O_i^{t+} < O_j^{t-}, \delta_1 > 0, \delta_2 > 0 \end{cases}$$

其中 δ_1 和 δ_2 分别表示 A_i 和 A_j 在提议不能被对方接受的情况下对原提议进行调整的步长。

由上面的不等式组可以确定 m 的范围: $\frac{O_i^{t-} - O_i^t}{\delta_1 + \delta_2} -$

$1 < m < \frac{O_i^{t-} - O_i^t}{\delta_1 + \delta_2}$, 可以清楚地看到, 协商所需要的步数直接取决于双方初始提议与反提议的差值。若没有有效的指导, A_i 和 A_j 为了获得更多的利益, 都会产生有利于自己的提议, 加上信息的私有性, 这些都将增加交互的轮次, 甚至使协商失败而无法形成联盟, 造成系统资源的巨大浪费。采用本文所提出的方法, 由于在产生初始提议前, A_i 和 A_j 对 A_j 应得的收益均有一个合理的估计, 即 O_i^t 和 O_j^t 趋向一致, 可以大大减少协商步数, 缩短联盟形成时间, 提高系统效率。

5 仿真试验

为了验证本文所提策略的有效性, 我们进行了仿真试验。设待求解任务集共包括 20 个任务 $T = \{t_1, t_2, \dots, t_{20}\}$, 给系统设定一个初始的历史任务集 T_h , 共包含 8 个任务, 之后每完成一个任务 t_k , $T_h = T_h \cup \{t_k\}$ 。在联盟 C_k 能完成任务 t_k 的前提下, 联盟成员 A_i 在任务 t_k 中第 k 维需求上所贡献的能力由一随机数 $b_i^k(k) \in [0, b_i^k(k)]$ 来确定。根据历史任务距当前任务的远近来确定其权值, 满足归一化条件。在每次形成联盟时, 我们随机指定一个 agent 为盟主 agent, 其余的个体以一定的概率 α 期望加入该任务, 根据第 4 节中的方法产生个体 A_i 的预期收益 O_i 。盟主 agent 和成员 agent 分别产生各自的模糊数, 在这里选用常见的三角模糊数。我们对两种不同的方法分别进行试验, 并将结果做比较。第一种方法是本文所提出的基于历史行为的联盟策略 (CF-H), 第二种方法是随机的联盟策略 (CF-S), 即协商时初

始提议随机产生.

第一组试验主要验证本文策略对联盟形成时效性的影响,具体表现为在联盟形成过程中协商次数的多少,试验中 $\alpha = 0.35$. 对每个任务共进行 10 次试验,结果取其平均值. 试验结果如图 1 所示. 可以看出,本文的策略由于考虑了 agent 在历史任务中的行为,对其在当前任务中的表现进行了预测,从而得到较为合理的初始提议,使协商在一个较高的起点开始,有效减少了协商所需步数,这同我们的分析是一致的.

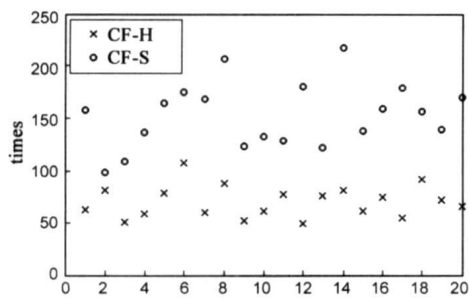


图 1 不同策略下的协商次数

其次考察在两种策略下参与协商的个体数目对协商次数的影响,通过取不同的 α 值来改变参与协商的 agent 数目,取 20 个任务的平均值. 试验结果如图 2 所示. 从图 2 可知,随着 α 的变大,参与协商的个体数目逐渐增大,两种策略所需协商次数都有所增加. 但相比本文策略,CF-S 由于初始提议产生的随机性,每对 agent 间需要较多的次数才能达到一致,因此随着参与协商个体数的增加,其协商次数的增加也更为明显. 另一方面,本试验中的协商模型是一对多的,若是多对多的协商,减少两两之间的交互将会显著降低系统的协商总次数,此时本文策略将更显其优越性.

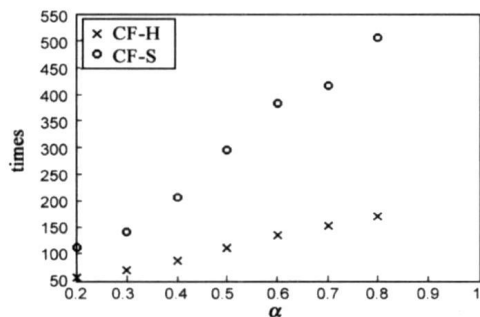


图 2 α 对协商次数的影响

6 结论

本文从一个全新的角度研究了协商的基本问题—如何产生初始提议,给出了能力贡献率和期望获利因子的概念,然后根据 agent 的历史行为对其进行预测,以此作为先验知识,确定协商初始时的提议与反提议,并给出了详细的分析. 通过仿真试验,可以看出这种方法

能够使协商在一个比较高的起点上开始,从而有效减少联盟形成过程中 agent 间的通信开销,缩短协商时间,并能提高协商的成功率.

另一方面,针对联盟问题中存在的 uncertainty,我们采用了专家模糊综合评判,从模糊数中得到初始提议等方法,使结果更加符合现实情况.

在初始提议之后的协商中,agent 的行为可能会受多个因素的影响. 如任务本身的性质,完成任务的环境、agent 自身的状态等等. 如何利用 agent 的智能,通过对对手状态和周边环境的学习以得到一个更优的协商策略是值得进一步研究的问题.

参考文献:

- [1] Rosenschein J S. Consenting agents: negotiation mechanisms for multi agent systems [A]. Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence [C]. Chambéry, France, 1993. 792– 799.
- [2] Zoltnik G, Rosenschein J S. Coalition, cryptography and stability: Mechanisms for coalition formation in task oriented domains [A]. Proceedings of the National Conference on Artificial Intelligence [C]. Seattle, Washington, 1994. 432– 437.
- [3] Hu W B, Wang S M. Research on the negotiation mechanism of multi agent system based on game theory [A]. Proceedings of the Ninth International Conference on Computer Supported Cooperative Work in Design [C]. Coventry, United Kingdom, 2005. 396– 400.
- [4] Bonnevey S, Kabachi N, Lamure M. Agent based simulation of coalition formation in cooperative games [A]. 2005 IEEE/WIC/ ACM International Conference on Intelligent Agent Technology [C]. Compiegne, France, 2005. 136– 139.
- [5] 张虹, 邱玉辉. 一个基于对策论的协商模型 [J]. 南京大学学报, 2001, 37(2): 159– 164.
Zhang Hong, Qiu Yuhui. A negotiation model that based game theory [J]. Journal of Nanjing University (Natural Sciences), 2001, 37(2): 159– 164. (in Chinese)
- [6] Hsiang C J, Ming C K, Nick G, Colin R, Peter S. An automated negotiation mechanism based on co evolution and game theory [A]. Applied Computing 2002: Proceedings of the 2002 ACM Symposium on Applied Computing [C]. Madrid, Spain, 2002. 63– 67.
- [7] Shehory O, Kraus S. Coalition formation among autonomous agents: Strategies and complexity [A]. Proceedings of MAAW-93 [C]. Neuchatel, 1993. 57– 72.
- [8] 王黎明, 黄厚宽. 一个基于多阶段的多 Agent 多问题协商框架 [J]. 计算机研究与发展, 2005, 42(11): 1849– 1855.
Wang Liming, Huang Houkuan. A multi stage based framework for multi agent multi issue negotiation [J]. Journal of Computer Research and Development, 2005, 42(11): 1849–

1855. (in Chinese)
- [9] John R, Raith M. G. Optimizing multi stage negotiation [J]. Journal of Economic Behavior and Organization, 2001, 45(2): 155- 173.
- [10] Fatima S S, Wooldridge M, Jennings N. R. An agenda based framework for multi issue negotiation [J]. Artificial Intelligence, 2004, 152(1): 1- 45.
- [11] Faratin P, Sierra C, Jennings N R. Negotiation decision functions for autonomous agents [J]. Robotics and autonomous systems, 1998, 24(3): 159- 182.
- [12] Ketchpel S. Forming coalitions in the face of uncertain rewards [A]. Proceedings of the National Conference on Artificial Intelligence [C]. Seattle, Washington, 1994. 414- 419.
- [13] 王立春, 陈世福. 多 agent 多问题协商模型 [J]. 软件学报, 2002, 13(8): 1637- 1643.
Wang Li chun, Chen Shi fu. A multi agent multi issue negotiation model [J]. Journal of Software, 2002, 13(8): 414- 419. (in Chinese)
- [14] 詹文杰, 汪寿阳. 评“ Smith 奥秘”与双向拍卖的研究进展 [J]. 管理科学学报, 2003, 6(1): 1- 12.
Zhan Wen jie, Wang Shou yang. Review on the Smith' s

mystery and development of double auctions [J]. Journal of Management Sciences, 2003, 6(1): 1- 12. (in Chinese)

- [15] 李贵安, 张志宏, 孟艳, 顾春. 模糊数学及其应用 [M]. 北京: 冶金出版社, 2006.

作者简介:



蒋建国 男, 1955 年生于安徽黄山, 合肥工业大学计算机与信息学院院长, 教授, 博士生导师, 研究方向为信号与信息处理、传感与智能控制等.



尹 翔 男, 1980 年生于江苏扬州, 博士研究生, 研究方向为分布式人工智能、多媒体信息处理等.