

基于 Multi-agent 的对策论协商模型

A Game Theory Negotiation Model Based on Multi-Agent System

86-88

张虹 李晓^{*} 邱玉辉

0225

(西南师范大学计算机科学与信息管理学院 重庆 400715)

Abstract In traditional Game Theory, there do not exist or exist many Nash balance points. So it is very difficult to find a solution, which could make every agent to be satisfied. At a view of human's daily life, this article tries to draw into a trusty third party to deal with some comparatively difficult game and condition, and draw into two communication behaviors: Asking Guarantee and Offering Compensation. That will make multi-agent find a compromise, stable, comparatively good solution in negotiation, which would satisfy every agent.

Keywords Multi-agent, Negotiation, Game theory, Nash equilibrium, Trusted Third Party

1 引言

对策论的研究是以每个 Agent 的效用和赢利为基础的, 每个 Agent 都能够在协商过程中求得 Nash 平衡的策略组合, 使得那些基于理性假设的 Agent 均可获得一个稳定并且满意的解。当任何理性 Agent 达到 Nash 平衡点后, 它们都不会试图离开, 这是因为只有 Nash 平衡点, 才能获得较大的赢利。可惜的是, Nash 平衡不一定总是存在, 或者可能存在多个, 在这些情况下, 如何寻找满意解, 是人们一直试图解决的问题。在人们的早期研究中, 往往不需要通信, 就可以利用对策论的决策机制进行 Agent 的协调和协作^[2,3]。但这样至少存在三个缺点: 1) Agent 在错误的理论假设下, 有可能陷入一个错误的结果中, 这就需要改变理性假设或者重新进行协商。2) 有一类对策它可能没有 Nash 平衡点, 那么它可能陷入无止境的循环推理中。3) 在那些有多个 Nash 平衡点的策略中, 又很难选出最好的一个并且保持稳定的解。

在此, 我们在研究过程中, 将两种通信行为引入协商过程, 这两种通信行为分别是: 请求保证 (Asking Guarantee) 和提供补偿 (Offering Compensation), 这两种通信行为都需要引入一个可信的第三方参与两个 Agent 之间的协调, 在这里可信的第三方扮演着一个可存储的银行或者说是一个仲裁者的角色。在这两种通信行为下, 一个 Agent 可以要求其它的 Agent 给它提供保证, 以确保某些可能导致较差结果的策略不被采用, 如果这个 Agent 没有采用此策略而是其他已经获得承诺的策略, 则保证将会在可信的第三方监督下实现; 同样, 一个 Agent 也可以提供一些补偿给其它的 Agent, 试图使其接受自己提出的那些可能带来较好结果的策略, 其他 Agent 若同意, 补偿也将会在可信的第三方监督下实现。

但是由于经验和时间的限制, 目前我们只针对那

些一次有效问题 (one-shot problem) 进行了研究, 我们假设 Agent 只进行一次对策协商并且协商过程必定会产生一个使 Agent 满意的解。接着我们将说明基本符号及某些概念的定义; 给出详细协商过程和一个具体的协商实例。然后将给出相关的定理及其证明, 最后是结论和展望。

2 协商模型

2.1 基本符号和概念定义

定义 1 一次有效对策论定义为四元式 $\langle A, S, P, N \rangle$, 其中:

A: 表示 Agent 集: $A = \{Agent_1, Agent_2, \dots, Agent_i, \dots, Agent_n\}$, 其中 $i = 1, 2, \dots, n$ 。

S: 每个 Agent 有一个纯策略的有限集:

$S = \{S_1(i), \dots, S_l(i), \dots, S_m(i)\}$, 其中 $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ 。

P: 每个 Agent_i 有一个赢利函数 $P_i, i = 1, 2, \dots, n$, 其中定义 $P_i(S(i), S(j))$ 表示当 Agent_i 采用一种策略, Agent_j 采用一种策略时, Agent_i 所获得的赢利值, 那么每个 Agent 在每个对策下的赢利值, 形成了一个对策矩阵。

N: 协商机制 (在后面会详细给出过程)。

定义 2 (Nash 平衡) 如果 Agent_i 和 Agent_j 采用一个策略组合 $(S^*(i), S^*(j))$, 且这个策略组合满足下面这个公式, 则我们就称它是 Nash 平衡:

$$\forall i P_i(S^*(i), S^*(j)) > P_i(S'(i), S^*(j)), \forall S'(i)$$

定义 3 (Pareto 有效) 如果存在一个策略组合 $(S^*(i), S^*(j))$, 使得其他任意一个策略都不可能在不减少 $P_i(S^*(i), S^*(j))$ 的前提下, 提高 $P_j(S^*(i), S^*(j))$, 那么我们就称这个策略组为 Pareto 有效。

定义 4 (Pareto 占优) 假设我们有策略组合分别定义为 X 和 Y, 如果所有 Agent 都在策略 X 下获得的赢利值比在策略 Y 下获得的赢利值高, 我们就称策

略 X 强 Pareto 占优于策略 Y; 如果有一个 Agent 在策略 X 下获得较高的赢利值, 并且没有 Agent 在策略 X 下获得较低的赢利值, 我们就称策略 X 弱 Pareto 占优于策略 Y。

定义 5 (请求保证机制) 将可信的第三方加入对策中, Agent 可先将一些保证放于第三方处, 并要求对方 Agent 承诺一定的策略下实现, 如果承诺兑现, 那么保证也将会实现。在对策矩阵中, 它的效果在于除承诺的策略外, 其它的策略平均地减少赢利值。这样, 这个策略就会成为 Pareto 占优策略。

定义 6 (提供补偿机制) 同样, 将可信的第三方加入对策中, Agent 会为了实现某一个策略, 将一些补偿放于第三方处, 并期望对方 Agent 承诺此策略实现, 如果承诺兑现, 那么补偿也将会实现。在对策矩阵中, 它的效果在于补偿的策略平均地增加赢利值。这样, 这个策略就会成为 Pareto 占优策略。

2.2 协商过程

我们将协商过程分为一个总的过程及其四个子过程, 其中我们用 N 表示 Nash 平衡集, PE 表示 Pareto 有效集, pe_i 和 pe_j 分别表示 PE 中的一个策略组合。

过程 P1 (总的协商过程)

1. 根据所有 Agent 的赢利函数, 求得相应的赢利值, 并且根据赢利值构建策略矩阵。
2. 在策略矩阵中按照前面所给的定义进行相应的推理, 找到以下的集合。
 - 2.1 Nash 平衡集 N。
 - 2.2 Pareto 有效集 PE。
3. 循环进行下面的过程: 请求、相反请求、协商。
 - 3.1 Agent_i 提出一个策略组合的请求 (过程 P2)。
 - 3.2 另一个 Agent_j 接受请求或者提出另一个不同的请求 (过程 P3)。
 - 3.3 得到一个折中或满意的结果或不再需要请求、协商结束。

过程 P2 (得到请求的子过程)

1. 在 PE 集中选择一个策略组合 pe_i , 它能使得 Agent_i 得到最高的赢利值。
2. 如果 pe_i 不是一个 Nash 平衡, 就计算能够 pe_i 通过过程 P4 达到 Nash 平衡的数量。通过这个计算过程可使 Agent_i 争取运用策略 pe_i , 然后将保证传给其它的 Agent。
3. 如果 pe_i 是一个 Nash 平衡, 则通过过程 P5 计算给 pe_i 的额外补偿的最小数量, 以便劝说另外的 Agent 接受 pe_i 。如果 pe_i 仍然是 Pareto 有效, 那么将请求把补偿发给另一个 Agent; 否则将 pe_i 从 PE 中删除, 并且重新选择其它的策略组合。

过程 P3 (接受请求或做出相反请求的子过程)

在 Pareto 有效集 PE 中选择策略组合 pe_i, pe_j , 能够使 Agent_i 获得最高的赢利值, 如果 pe_i 的赢利值比 pe_j 的赢利值少, 放弃 pe_i , 转过程 P2, 提出请求; 否则, 接受 pe_i 。

过程 P4 (计算保证数量的子过程)

1. 考虑一个特别的值, 在策略矩阵中比较 Agent_i 在其它策略下的赢利值, 找到最大的距离 $d_i = \max P_i(S^*(i), S^*(j)) - P_i(S^*(i), S^*(j)) \forall S^*(i) \neq S^*(j)$, 如果 d_i 大于 0, 则 Agent_i 所能提供的保证数量为 $d_i + \delta$, 否则没有必要给另一个 Agent 提供任何保证, 而这里的 δ 则定义为边缘保证, 它是使 Agent 改变占优的赢利最小值。
2. 同样的过程还可以找到另一个 Agent_j 的保证数量。只需要满足公式

$$d_j = \max P_j(S^*(i), S^*(j)) - P_j(S^*(i), S^*(j)) \forall S^*(i) \neq S^*(j)$$

(j), 即可。

过程 P5 (计算补偿数量的子过程)

在这个过程中, 我们一般采取选用赢利值的基本定量的最小值 0, 如果赢利值没有基本定量, 那么就特别选用两个最近的赢利值之间最小差值的一半作为赢利值的基本定量 0。由于在一般情况下, 赢利值的基本定量可能都不存在, 而补偿不能小于其它 Agent 提出的赢利值。因此, 这个定量必须满足这样一个原则, 适当定量原则: 避免一个 Agent 提出的补偿太小, 以至于协商需要很长的过程。

2.3 协商实例

在此, 我们根据上面提出的定义和协商模型, 给出具体的实例。在此我们运用一个有关协作 (cooperate) 和告密 (fink) 的例子, P 和 Q 分别表示两个 Agent, 如图 1。

Q \ P	Cooperate	Fink
Cooperate	-1	-10
Fink	-10	-8

(a) 协商前

Q \ P	Cooperate	Fink
Cooperate	-1.1	-10
Fink	-10	-9.1

(b) 协商后

图 1 策略矩阵

现在, 我们一起来看一下它是如何求得解的。在此我们先规定边缘保证 δ 为 0.1, 并且赢利值的基本定量 0 为 1, 我们先看到图 1 中的 (a), 这里三个结果是 Pareto 有效, 分别是策略组合 (Cooperate, Cooperate), (Cooperate, Fink), 以及 (Fink, Cooperate), 但 Nash 平衡的策略组合 (Fink, Fink) 不在其中, 下面就看它的协商过程。

1. Agent_Q 提出策略 (Cooperate, Fink)。由于, 策略组合 (Cooperate, Fink) 能给 Agent_Q 带来最高的盈利值, 但由于此策略组合不是一个 Nash 平衡, 计算过程 P4 计算出 d_i 为 2, 则 $d_i + \delta$ 为 2.1, 将保证值 2.1 放于可信的第三方处

2. 由于策略 (Cooperate, Fink) 不能给自己带来最

大的盈利值,所以 Agent₂ 拒绝 Agent₁ 提议。

3. Agent₂ 提出策略组合 (Fink, Cooperate), 并同样提供保证 2.1, 放于可信的第三方处。

4. Agent₁ 由于策略组合 (Fink, Cooperate) 不能使自己获得盈利值最大, 也拒绝。

5. 这是 Agent₁ 考虑策略组合 (Cooperate, Fink), 发现将其盈利值和补偿值 1 做组合变为 (-1, -9), 仍然被策略组合 (Cooperate, Cooperate) Pareto 占优, 因此, 将 (Cooperate, Fink) 从策略集中删除, 并选择另外的策略组合。

6. Agent₁ 提出策略 (Cooperate, Cooperate), 并通过过程 P4 算出保证为 1.2, 放于可信的第三方处。

7. Agent₂ 依然拒绝这个提议, 由于这时策略组合 (Fink, Cooperate) 在它的策略集中能使他带来更高的盈利值。

8. 这时 Agent₂ 考虑策略组合 (Fink, Cooperate), 发现将其盈利值和补偿值 1 做组合变为 (-9, -1), 仍然被策略组合 (Cooperate, Cooperate) Pareto 占优, 因此, 将 (Fink, Cooperate) 从策略集中删除, 并选择另外的策略组合。

9. Agent₂ 提出策略 (Cooperate, Cooperate), 并通过过程 P4 算出保证为 1.1 (如图 1 中的 (b))。

10. Agent₁ 接受提出, 协商完成。

在这个例子中, 我们可以看出在协商过程中, 请求保证和提供补偿这两种通信行为帮助 Agent 实现了一个稳定的较好解。在某些例子中, 可能对通信行为有不同的要求, 当然要具体情况, 具体分析, 在此我们不一一详述。

3 有关定理及其证明

定理 1 (Nash 平衡存在定理) 在一次有效的协商对策中, 在请求保证和提供补偿的机制下, 必定存在并且至少存在一个 Nash 平衡的策略组合, 它 Pareto 弱占优于其它的策略组合。

证明: 将每一个 Agent 的盈利值排序, 然后选出一个最大的盈利值策略组合 $(S^*(i), S^*(j))$, 这个策略组合肯定是 Pareto 有效的, 如果这里仍然有其他的策略组合 Pareto 占优于这个策略组合, 则它的盈利值比大于 $(S^*(i), S^*(j))$, 这与前面矛盾, 并且, 我们也可以通过请求保证来使得这个策略称为 Nash 平衡。

定理 2 (顺序无关或者顺序独立) 在一次有效的协商对策中, 在请求保证和提供补偿的机制下, 最终的协商结果和最初的协商顺序无关, 所有的 Nash 平衡最终都会得到同样的结果。

证明: 在定理 1 中我们可以看出, 协商过程的最后结果肯定属于那些能给每个 Agent 带来较大的盈利值协商组合的排序之列。在协商刚开始的时候, 每个 Agent 都会根据自己的情况和协商目的, 选择不同的结果进行协商。假设 Agent₁ 选择的是策略组合 $(S^*$

$(i), S^*(j))$, Agent₂ 选择的是策略组合 $(S^{**}(i), S^{**}(j))$, 如果它们中有策略组合不能带来较大盈利值, 则在协商中提供补偿以后, 这个策略组合就会被其他的策略组合 Pareto 占优, 否则, 这两个策略组合在多次协商并提供补偿后, 协商将会使每个 Agent 都在这两个策略组合下取得相同的盈利值 $P_i(S^*(i), S^*(j)) = P_i(S^{**}(i), S^{**}(j))$, $P_j(S^*(i), S^*(j)) = P_j(S^{**}(i), S^{**}(j))$ 而结束。如不其然, 协商将会下终止, 因此, 协商过程的最后选择策略有可能不同, 但它们所得的盈利值是相同的, 并且与协商顺序无关。

定理 3 (协商过程的收敛性) 在一次有效的协商对策中, 在请求保证和提供补偿的机制下, 协商过程必定能在有限步内停止。

证明: 由于在协商过程中, 每一次提供补偿的数量必定是单调递增的, 因此较高的补偿值必定会使策略组合被其他的 Pareto 占优, 所以, 根据前面的定义 Pareto 有效集中策略组合的数量肯定是逐步减少的。协商过程会在获得满意解或者 Pareto 有效集中元素为 0 时结束, 即可知, 无论怎样, 协商过程必定会在完成有限步后停止, 协商过程收敛。

结论和展望 在对策论中, 当所有参与者都不遵守承诺时, 这时对策就称为无协作的对策。Nash 证明了在无协作对策中, 固定策略平衡的存在性^[6]。但这种策略不适用于一次有效的问题。在本文中, 我们提出的一个通过改变盈利值而寻找策略平衡的协商模型。在现实世界的运用中, 人们都会试图寻找更可靠, 更理性的功能, 其中可行的第三方在人类社会中起着举足轻重的作用, 能够帮助人们处理许多复杂的协商问题。在此, 我们将两者相结合, 并定义了两种通信行为: 请求保证和提供补偿, 使其在协商过程有助于 Agent 在没有 Nash 平衡或有多个 Nash 平衡时, 寻找到最优解。这种结合不但使 Agent 的协商简化, 而且提高了协作的机制。我们相信这种协商模型必定是可行的和有效的。由于时间和经验的不足, 目前我们只在一次有效问题的基础上进行了一些相关的讨论, 在未来的时间, 我们必会逐步扩大问题的讨论范围, 并完善通信语言和协商机制。

参考文献

- 1 Brans S J. Negotiation, Games, Applying Game Theory to Bargaining and Arbitration. Roulledge, New York, 1990
- 2 Axelrod R. The Evolution of Cooperation. Basic Books Inc., New York, 1984
- 3 Genesereth M R, et al. Cooperation without commutation. In: Proc. of the National Conf. on Artificial Intelligence (AAAI-86). 1986. 51~57
- 4 Wu Shih-huang, Soo You-Wun. Game Theoretic Reasoning in Multi-agent Coordination by Negotiation with a Trusted Third Party, ACM, 1999
- 5 Rasmusen E. Games and Information: An Introduction to Game Theory. Basil Blackwell, and Oxford, 1989
- 6 Nash J F. Non-cooperative games, Ann. of Math. 1951. 286~295