Game Theoretic Reasoning in Multi-agent Coordination by Negotiation with a Trusted Third Party

Shih-Hung Wu

Department of Computer Science National Tsing Hua University Hsin-Chu City, 30043 Taiwan, R.O.C. 886-3-5715131 ext. 3532

shwu@cs.nthu.edu.tw

Von-Wun Soo

Department of Computer Science National Tsing Hua University Hsin-Chu City, 30043 Taiwan, R.O.C. 886-3-5715131 ext. 1068

soo@cs.nthu.edu.tw

ABSTRACT

In multi-agent coordination, one would normally like to find a satisfactory solution that is stable, fair and optimal to all agents. According to traditional game theory, Prisoner's dilemma, no or more than one Nash equilibrium games are situations that are difficult to find such a satisfactory solution. In human society, it often involves a trusted third party in the negotiation process among agents to ensure the cooperation and commitment of agents. In this paper, we describe how the trusted third party can be involved in the negotiation of multi-agent coordination to deal with many difficult game situations. We introduce two communication actions into the traditional game-theoretical reasoning: asking guarantee and offering compensation for agents to use in negotiation. Asking guarantee from agents and depositing it at the trusted third party can ensure the agents to keep their commitments, while offering compensation can allow finding a fair and compromised solution for all agents. We show how the negotiation communication protocols can be proceeded using these two communication actions to reach a compromised and stable agreement in all different game situations.

Keywords

Multi-agent coordination, Game theory, Negotiation protocol, A Trusted third party, Nash equilibrium, Prisoner's dilemma.

1. INTRODUCTION

According to game theory, agents will try to reach Nash equilibrium, a stable outcome for each agent based on the assumption of rationality. No rational agents will leave the Nash equilibrium they have reached, because any agent who leaves alone will get fewer payoffs. However, the Nash equilibrium may not always exist, even it does, it may not be the optimal solution. In order to coordinate both agents, we need to change the payoff matrix, which can not only escape from a Prisoners' dilemma but also can reach a stable solution in the games that do not have a Nash equilibrium.

Previous work showed that rational agents are able to coordinate and cooperate with a game theoretical deal-making mechanism, even without communication [1][4]. However, there are three different faults in game theoretical decision making without communication. First is the problem of prisoner's dilemma, agents will run into a worse result under the assumption of rationality. It needs to change the assumption of rationality [4][5] or to allow negotiation [8][9]. Second, a class of games that give no equilibrium point will cause an endless cyclic reasoning [3]. Third, when there are more than one equilibrium, it is hard to select the best one and remain stable.

In our previous work, we have introduced the ability of asking guarantee as the first kind of communication action into the conventional game theory to escape the from prisoner's dilemma situations [9], but the games with no equilibrium remain unsolved. In this paper, we introduce a second kind of communication action, offering compensation, into the game theory. These communication actions need a trusted third party in the game to coordinate with the two agents involved. The trusted third party plays as the role of a bank or a mediator. One agent may ask another agent to pay some guarantee to ensure the strategies that lead to an undesirable state will not be picked up. The guarantee will be returned if the agent does not play any strategy other than the committed one. One agent may also pay compensation to another agent, in order to persuade the agent to play certain strategy that can lead to an acceptable state. The compensation will be sent to the agent who does play the strategy as was asked.

Take the game matrix in Figure 1(a) as an example, where the arrows represent the driving force to move form one outcome to another. If both agents deposit at the trusted third party the defection-free guarantee, say 1.1, then the game matrix can be changed into a dilemma-free one as shown in Figure 1 (b). In this game, both agents play "Cooperate" and thus get a payoff -1 and take back the guarantee from the trusted third party. By doing so, both agents could escape from the prisoners' dilemma. Although asking guarantee can help agents to make a particular outcome a Nash equilibrium, there is still a problem, why agents choose the particular outcome. We propose that there is a need of negotiation involving the action of offering compensation. In the subsequent sections of the paper, we will describe in detail how agents perform the reasoning that involves communication with a trusted third party.

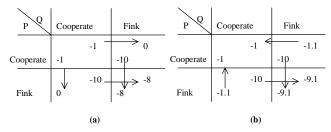


Figure 1: Prisoner's Dilemma games matrix (a) A special case of a PD game matrix. (b) A dilemma-free game matrix.

In section 2, we present the definitions in conventional game theory, and describe how to recognize an undesirable state and find a desirable one in all possible 2 by 2 games. In section 3, we develop procedures for agents to perform negotiation. In section 4, we make some discussion and the conclusion.

2. REASONING ON THE EQUILIBRIUM STATES IN GAMES

The game theory can serve as a tool of reasoning about the moves for rational agents. We first introduce some basic definitions and notations used in the game theory. The payoff value for each player in a combination of strategies is denoted as $\pi_i(s_i, s_{-i})$, where i is the index of player, and -i denotes the other players; player i plays strategy s_i while player -i plays strategy s_{-i} . A dominant strategy for a player is a strategy that will always give a higher payoff no matter what strategies the other player chooses.

The Nash equilibrium A strategy combination (s_i^*, s_{-i}^*) is a Nash equilibrium if any agent will get less its payoff when it deviates from this strategy combination alone. Mathematically: $\forall i, \pi_i(s_i^*, s_{-i}^*) > \pi_i(s_i^*, s_{-i}^*), \forall s_i^*$.

The Pareto-efficient: A strategy combination is Pareto efficient if there is no other strategy combination increases the payoff of one agent without decreasing the payoff of another agent.

Pareto-dominant: A strategy combination X is strongly Pareto-dominates another strategy combination Y, then all agents have higher payoff under X. If strategy combination X weakly Pareto-dominates strategy combination Y, then one agent has higher payoff under X, and no agent has lower payoff.

We can use the concept of Pareto-efficient to form the acceptable solution set. Because taking any strategy combination that is Pareto-dominated by other strategy combination is not smart. However, in the prisoner's dilemma, it does happen, because the (Fink, Fink) strategy combination is a Nash equilibrium. The strategy combination (Cooperate, Cooperate) Pareto-dominates (Fink, Fink) but it is not a Nash equilibrium; any agent that commits (Cooperate, Cooperate) will take a risk of trapping into the (Cooperate, Fink) or (Cooperate, Fink) situation due to the rationality of its opponent. The other problem is that there could be many strategy combinations, which are not Pareto-dominated by any other strategy combination. How agents would negotiate

in a way that a "better" strategy combination can be picked and ensured to be carried out by all agents is an important issue.

Asking guarantee is a method we proposed to make a particular outcome a Nash equilibrium. Previous work showed that it can help to escape the prisoner's dilemma. In section 2.1, we show that it can also help agents to find a stable solution when there is no Nash equilibrium exists in a game. However, asking guarantee cannot help agents to make a decision when there are multiple Nash equilibrium in a game. Offering compensation is a possible technique to deal with the case. In section 2.2, we show that it is helpful to have the ability of offering compensation in selecting among multiple Nash equilibria.

2.1 Finding a Desirable State in a Game Where no Nash Equilibrium Exists

The welfare game is a well-known problem for the government to choose aid or no aid for the pauper. Take the game matrix in Figure 2(a) as an example. There is no dominant strategy for both agents and no Nash equilibrium exists, if the agents try to simulate the opponent's behavior, it will run into an endless cycle. In addition, we can find that the two strategy combinations that involve "no aid" are Pareto-dominated by the strategy combination (Aid, Work). Nevertheless (Aid, Idle) is not Paretodominated by any strategy combination. Therefore, according to the traditional game theory the agents could try to find a mixed strategy Nash equilibrium by associating the probability 0.2 with the "work", 0.8 with "idle", 0.5 for "aid" and 0.5 with "no aid". The mixed strategy results in the expected payoff for agent G being -0.2 and that for agent P being 1.5. However, if agent G asks agent P to deposit guarantee 2 not to play "idle" at the trusted third party, then the game matrix will become the one in Figure 2 (b). The strategy combination (Aid, Work) becomes a Nash equilibrium, so rational agents will accept this strategy combination and will not leave alone. Since if agent P plays "Idle", the guarantee 2 will be forfeited by the trusted third party, then the payoff agent P can get will decrease from 3 to 1. Due to rationality, of course, agent P will cooperate to play "Work" in order to get the payoff 2. In addition, the guarantee will return to the agent P from the trusted third party. In this game, both agents get a higher payoff than the expected payoff of mixed strategy.

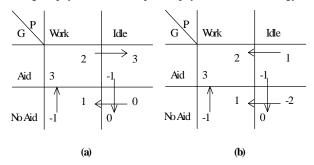


Figure 2: Welfare Game (a) The original game matrix. (b)

The game matrix after negotiation.

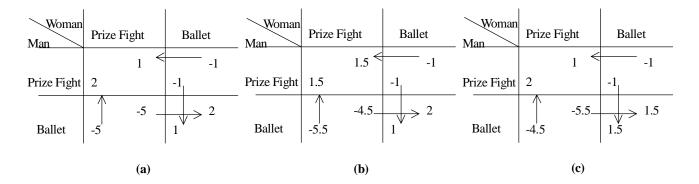


Figure 3: The Battle of the Sexes (a) The original game matrix. (b)The game matrix after Man offering compensation 0.5 to Woman for her to play "Prize Fight". (c) The game matrix after Woman offering 0.5 compensation to Man for him to play "Ballet".

2.2 Find a Desirable State in a Game Where Multiple Equilibria Exist

Here is another well known problem in game theory is the battle of the sexes. Take the game matrix in Figure 3(a) as an example. There is no dominant strategy for both agents but the game has two Nash equilibria. Moreover, we can find that the states that they play different strategies are Pareto-dominated by the states that they play the same strategy. Traditional game theory suggests a first-mover advantage concept. Namely if an agent moves first, then the other agent must follow toward a Nash equilibrium that has been chosen by the first agent. This is unfair for the agent being the second player. This will make both agents to rush into selecting the first move. Moreover, the result of the game will not be predictable unless we know who gets to move first. In order to keep the result of the game predicable, we suggest that the first mover should offer compensation to the second mover. For example the Man may suggest the Nash equilibrium (Prize Fight, Prize Fight) and offer a compensation 0.5 to woman, see Figure 3(b). How the amount of compensation is calculated is to be discussed in section 3. The Woman may accept the offer or propose another suggestion on (Ballet, Ballet) with a counter compensation 0.5 to Man, see Figure 3(c). The Man may accept the counter offer or propose another offer with a higher compensation. As the amount of compensation is getting higher, one of the Nash equilibria will become Paretodominating the other ones. Under the assumption of utility maximization, both Man and Woman will accept the new negotiated state. In this example, after the Woman proposed a counter offer, the payoff of each agent is 1.5 no matter they choose (Prize Fight, Prize Fight) or (Ballet, Ballet).

3. THE NEGOTIATION PROCEDURE TO CREATE A UNIQUE NASH EQUILIBRIUM

We present the procedure by two stages, the first one is a general procedure of negotiation, and the second one is the details of each sub-procedure. Note that both agents use the same procedures and the guarantee and compensation must be deposited to a trusted third party other than the two agents. The result is independent of which agent initiating the negotiation.

Stage 1:

Procedure P1 (The general negotiation procedure)

- 1. Getting all payoff entries and constructing the game matrix.
- 2. Reasoning on the game matrix, finding the sets
 - 2.1. The Nash equilibrium set N.
 - 2.2. The Pareto-efficient set P.
- 3. Loop of offer, counter offer, negotiation.
 - 3.1. agent *i* offers a particular strategy combination.
 - 3.2. agent -i accepts the offer or proposes a counter offer.
 - 3.3 Negotiation ends on a compromised outcome or there exists no more offers.

Stage 2:

Procedure P2 (The sub-procedure of making an offer)

- 1. Select the strategy combination p that gives the highest payoff for agent *i* from the Pareto-efficient set P.
- 2. If p is not a Nash equilibrium then calculate the amount of guarantee that needs to make p a Nash equilibrium by Procedure P2. The procedure finds the amount for asking agent -i to play p and for agent i itself not deviating form p. Then send it to agent -i.
- 3. If p is a Nash equilibrium, then offer a minimal amount of additional compensation to p by procedure P5, in order to persuade agent -i to accept p. If p is still Pareto-efficient, then send the offer to agent -i, otherwise delete p from P and got to 1 reselect another strategy combination.

Procedure P3 (The sub-procedure of accepting an offer or making a counter offer)

1. Select the strategy combination q that gives the highest payoff for agent i from the Pareto-efficient set P. If the payoff of p is less than the payoff of q, reject p, making another offer by **procedure P2**.

Procedure P4 (The sub-procedure of calculating the amount of guarantee)

1. For a particular outcome $(s_i^*, s_{\cdot i}^*)$, compare the payoffs of agent i in all its other strategy $s_i^{'}$ in the game matrix, find the maximum distance dl =

- $\max \pi_i(s_i^*, s_{-i}^*) \pi_i(s_i^*, s_{-i}^*), \forall s_i^* \neq s_i^*$. If d1 greater than 0, the guarantee that agent i should deposit is $d1 + \delta$ otherwise there is no need for agent -i to deposit any guarantee. Where the marginal guarantee δ is small amount of payoff that makes agent change the order of dominant.
- 2. The same procedure can find the amount of guarantee for agent -i. For a particular outcome (s_i^*, s_{-i}^*) compare the payoffs in all agent -i other strategy s_{-i} in the game matrix, find the maximum distance $d2 = \max \pi_{-i}(s_i^*, s_{-i}^*) \pi_{-i}(s_i^*, s_{-i}^*), \forall s_{-i}^* \neq s_{-i}^*$. If d2 greater than 0, the guarantee that agent i should deposit is $d2 + \delta$ otherwise there is no need for agent -i to deposit any guarantee.

Procedure P5 (The sub-procedure of calculating the amount of compensation)

1. Add the minimal amount of payoff quantum θ . If the payoff has no basic quantum, half of the minimal difference between the two closest distinct payoffs can be treated as the payoff quantum.

Proper quantum principle: Since the basic quantum of payoff may not exist in general cases. The compensation should not be less than the amount that another agent offered. This principle is necessary to prevent that one agent may offer so small compensation that causes a lengthy negotiation process.

In section 1 we showed that (Cooperate, Cooperate) could be a stable solution in the prisoner's dilemma game, but we did not give the reason of why agents choose Cooperate. In the following example, we give the reason by tracing the negotiation process of asking guarantee and offering compensation.

Example (The negotiation process for a Prisoner's Dilemma game)

Now we trace the negotiation process of finding (Cooperate, Cooperate) a Nash equilibrium in a prisoner's dilemma game. Assume the amount of marginal guarantee δ is 0.1 and payoff quantum θ is 1. In the game matrix, see Figure 1(a), there are three outcomes which are Pareto-efficient (Cooperate, Cooperate), (Cooperate, Fink), and (Fink, Cooperate), but the Nash equilibrium (Fink, Fink) is not one of them.

- 1. Agent Q suggests (Cooperate, Fink) because (Cooperate, Fink) gives the highest payoff for agent Q and asking guarantee 2.1 from agent P to play Cooperate because it is not a Nash equilibrium.
- Agent P rejects the suggestion, because it does not give the highest payoff for itself.
- 3. Agent P suggests (Fink, Cooperate) instead and asks guarantee 2.1 for agent Q to play Cooperate.
- 4. Agent Q rejects the suggestion, because it does not give the highest payoff for itself.
- 5. However, the outcome (Cooperate, Fink) that gives the highest payoff for agent Q has been proposed, need to offer some compensation. The payoff of (Cooperate, Fink) with compensation 1 become (-1, -9) which is Pareto-dominated by the payoff of (Cooperate, Cooperate), i.e., (-1, -1). So agent Q delete the (Cooperate, Fink) from its Pareto-efficient set and reselect another offer.

- Agent Q suggests (Cooperate, Cooperate) while offering a guarantee 1.1 that for playing Cooperate itself, and asking a guarantee 1.1 for agent P to play Cooperate.
- 7. Agent P rejects the suggestion. Because it still has an outcome (Fink, Cooperate) in its Pareto-efficient set that can get a better payoff.
- 8. Agent P finds that the payoff of (Fink, Cooperate) with compensation 1 becomes (-9, -1) which is Pareto-dominated by the payoff of (Cooperate, Cooperate), i.e., (-1, -1). It eliminates the state from its Pareto-efficient set.
- 9. Agent P suggests (Cooperate, Cooperate) while offering a guarantee 1.1 that for playing Cooperate itself, and asking a guarantee 1.1 for agent P to play Cooperate.
- 10. Agent Q accepts the suggestion. Because it gives no lower payoff among all the candidates.

As we can see, in this example, asking guarantee and offering compensation are used as communication actions that help agents to reach a stable state. In this particular case, no compensation is needed. However, there are cases that compensation must be deposited, such as the game of the Battle of the Sexes.

4. DISCUSSIONS AND PROOFS

The scope of this paper is limited in the one-shot games. We assume that agents play a game only once and a negotiation mechanism may help agents to achieve a better outcome. A negotiation game is extended from the traditional game by incorporating a negotiation mechanism. In this game, agents not only reason on the game matrix formed by the payoff functions but also attempt to find an equilibrium using given negotiation mechanisms. Negotiation games under different notions and representation can be found in [2].

Definition 1 (The One-Shot Negotiation Game):

A one-shot negotiation game can be defined as a tuple <A, S, P, N>. An agent set A, a strategy set S, payoff functions P and a negotiation mechanism N. The payoff functions of each agent on every combinations of strategies form an game matrix. A negotiation mechanism is a procedure of communication actions.

Definition 2 (The Asking Guarantee Mechanism):

A trusted third party agent is added into the game. An agent may commit to a certain strategy by depositing a guarantee to the trusted third party. The guarantee will be returned if the agent keeps the commitment. The effect on the game matrix is equivalent to decreasing all payoffs of the strategies other than the committed one. Therefore, the committed strategy becomes a dominant strategy. The amount of guarantee is determined by procedure **P4**.

Definition 3 (Prisoner's dilemma): A general definition of the prisoner's dilemma can be a game that has a unique Nash equilibrium but the Nash equilibrium is Pareto-dominated by some other strategy combination.

Lemma 1(Escape from the Prisoner's Dilemma)

In one-shot negotiation games, with the asking guarantee mechanism, we can escape from the prisoner's dilemma.

Proof: If the game is in the prisoner's dilemma, the guarantee mechanism can change the dominant strategy to be another

strategy for each agent. Since the combination of each agent's dominant strategy must be a unique Nash equilibrium. This mechanism changes the game matrix, so the position of Nash equilibrium may move to any chosen outcome.

Definition 3 (The Offering Compensation Mechanism):

A trusted third party agent is added into the game. This mechanism changes the game matrix by the procedure **P5**, so that at least one of the multiple Nash equilibria may Pareto weakly dominate others. In order to avoid a lengthy negotiation process, there should be a principle of the proper amount of compensation.

Theorem 1(Existence of Nash and Pareto Equilibrium):

In one-shot negotiation game, with the asking guarantee and offering compensation mechanisms, there exists at least one Nash Equilibrium which Pareto weakly dominates all other outcomes.

Proof: Sort all outcomes by the total payoff of every agent. Pick up the largest outcome (S_i^*, S_j^*) among the largest ones. This outcome must be Pareto-efficient, if there is an outcome which Pareto-dominates (S_i^*, S_j^*) , then the total payoff must be larger than the total payoff of (S_i^*, S_j^*) , this contradicts the pick up rule. Moreover, by the asking guarantee mechanism, we can make this strategy combination a unique Nash equilibrium.

Theorem 2(Order independent):

In one-shot negotiation game, with the asking guarantee and offering compensation mechanisms, the final negotiated utility is independent of the order of the initiating the negotiation. All the resulting Nash Equilibrium outcomes which Pareto weakly dominate all other outcomes will end up with the same utility values.

Proof: As shown in the previous theorem, the final outcome of the negotiation must belong to the set of outcome that gives the largest total payoff. At the initial state of the negotiation, each agent may pick up different outcome as the destination of the negotiation. Suppose that agent i picks up (S_i^*, S_j^*) and agent j picks up (S_i^*, S_j^*) , if one is not in the set of outcomes which gives the largest total payoff, then it will be Pareto-dominated by the other after the compensation is offered. If both outcomes belong to the set of outcomes that gives the largest total payoff, then after the iteration of compensation offering, the payoff of each agent will end up with the same value, i.e., $P_i(S_i^*, S_j^*) = P_i(S_i^{*'}, S_j^{*'})$. Otherwise, the negotiation will not be ended. There the negotiation is order-independent in the sense the final payoffs are the same although the final strategy combination can be different.

An example of theorem 2 can be found in Fig. 3 where no matter who initiating the compensation, the outcome will be (1.5, 1.5) regardless the final outcome is (Prize Fight, Prize Fight) or (Ballet, Ballet).

Theorem 3(Convergence of the negotiation procedure):

In one-shot negotiation game, with the offering compensation mechanisms, the negotiation procedure will end in finite time.

Proof: Since the amount of compensation for each outcome is a monotonic increasing number, higher compensation will cause the outcome Pareto-dominated by other outcomes. Therefore, the number of elements in the Pareto-efficient set is always

decreasing. If the negotiation reaches an acceptance state or the size of the Pareto-efficient set becomes zero, then the procedure ends.

5. CONCLUSION

In the game theory, when both players can't make a binding commitment, i.e., no players would bind their commitment to the decision they make; the game is called a non-cooperative game. Nash proved the existence of the mixed strategy equilibrium in all non-cooperative games [6]. However, mixed strategy is hard to use in one-shot game, i.e. the game played only once or without history of previous games. We present a constructive procedure to find pure strategy equilibrium by changing the payoff. Instead of using mixed-strategy, our negotiation protocol has an advantage of no need to take chances. In real-world application, people tend to prefer performing transactions in more reliable and rational fashion. The trusted third party plays important roles in human society to deal with various complicated negotiations.

We define two types of communication actions of agents to ensure the commitment. The fact that agents can deposit a guarantee to a trusted third party is very important in our approach. This method is generally adopted in human society. In agent society, there can be many different ways to implement the concept. Since agents are assumed to be rational, the forfeit of guarantee can in fact never happen. Therefore, we can replace the guarantee to some kind of punishment that will have the same utility value. The compensation is much harder mechanism to implement than the guarantee. If the payoff cannot be transferred partially, the compensation may be impossible. However, in human society, we have money as a tool for the guarantee and the compensation.

We have proposed two communication actions: asking guarantee and offering compensation into multi-agent negotiation process, which can help the agents not only escape from the prisoner's dilemma but also make a decision in no or multiple Nash equilibrium cases that cannot be solved in traditional game theory without a trusted third party. We find that involving a trusted third party in the game will make it easier to make rational deals. In this framework, agents can make rational decisions by depositing guarantee or compensation to the trusted third party. The involvement of a trusted third party not only can facilitate the coordination of multi-agent but also can enhance multi-agent cooperation.

6. ACKNOWLEDGMENTS

This work was financially supported by the National Science Council of Taiwan, Republic of China, under grant No. NSC88-2213-E-007-057.

7. REFERENCES

- R. Axelrod, The Evolution of Cooperation, Basic Books Inc., New York, 1984.
- [2] S.J. Brams, Negotiation Games: Applying Game Theory to Bargaining and Arbitration, Routledge, New York, 1990.
- [3] E.H. Durfee, J. Lee, and P.J. Gmytrasiewicz, Overeager Reciprocal Rationality and Mixed Strategy Equilibria, *In*

- Proceedings of the Eleventh National Conference on Artificial Intelligence(AAAI-93), pp.225-230, 1993.
- [4] M.R. Genesereth, M.L. Ginsberg, and J.S. Rosenschein, Cooperation without communication, *In Proceedings of the National Conference on Artificial Intelligence(AAAI-86)*, pp.51-57, Philadelphia, Pennsylvania, 1986.
- [5] Y. Mor and J. S. Rosenschein, Time and the Prisoner's Dilemma, In Proceeding of the First International Conference of Multi-Agent System (ICMAS-95), pp.276-282, 1995
- [6] J. F. Nash, Non-cooperative games, Ann. Of Math. 54, pp.286-295, 1951.

- [7] E. Rasmusen, *Games and Information: An Introduction to Game Theory*, Basil Blackwell, Oxford, 1989.
- [8] J.S. Rosenschein, M.R. Genesereth, Deals among Rational Agents, *IJCAI-85*, pp.91-99. 1985.
- [9] S.H. Wu, V.W. Soo, Escape from a Prisoners' Dilemma by Communication with a Trusted Third Party, ICTAI-98, 1998.Schwartz, M., and Task Force on Bias-Free Language. Guidelines for Bias-Free Writing. Indiana University Press, Bloomington IN, 1995.