Risk Control 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.
shwu@cs.nthu.edu.tw

Von-Wun Soo
Department of Computer Science
National Tsing Hua University
Hsin-Chu City, 30043 Taiwan, R.O.C.
soo@cs.nthu.edu.tw

Abstract

In multi-agent coordination, the uncertainty may come from two major sources: the moves of the nature agent and the unpredictable behavior of other autonomous agents. The uncertainty may affect the expected payoff and the risk of an agent. A rational agent would not always play the strategy that gives the highest expected payoff if the risk is too high. To tackle the uncertainty in multi-agent coordination, a risk control mechanism is necessary in multi-agent decision making. We assume agents may have different risk preferences, e.g. risk-averse, risk-neutral, and risk-seeking, and separate the risk preference from the utility function of a given strategy. Taking agent's risk preference into account extends the notions of the dominant strategy, the Nash equilibrium, and the Pareto-efficiency in traditional game theory. We show how the risk control can be carried out by a negotiation protocol using communication actions of asking guarantee and offering compensation via a trusted third party.

1 Introduction

Multi-agent coordination is necessary for a multi-agent community to prevent anarchy or chaos. Negotiation is a way to coordinate rational agents under a game theoretical deal-making mechanism [Rosenschein and Genesereth, 1985; Rosenschein, 1994, Haynes and Sen, 1996, Wu and Soo, 1998a, 1999]. The underlying assumption is that agents model each other as rational agents [Axelrod, 1984; Brafman and Tennenholtz, 1997; Koller and Pfeffer, 1997; Shehory and Kraus, 1998; Tennholtz 1998]. If there are different types of agents, a recursive modeling method may help in decision making [Gmytrasiewicz et al., 1991; Durfee et al., 1993;Vidal and Durfee, 1995]. Fuzzy theory is a way to deal with uncertainty in a fuzzy game [Wu and Soo, 1998b]. However, the issues on the risk in an uncertain game are rarely addressed.

In multi-agent coordination, the uncertainty may come from two major sources: the nature agent moves and the unpredictable behavior of other autonomous agents. The uncertainty may affect the expected payoff and the risk that an agent can seek. According to the rationality assumption of a traditional game theory, it is assumed that rational agents will try to play a strategy that gives the highest expected payoff. But in an uncertain game, human agents may not always play the strategy that gives the highest expected payoff if the risk is too high to be accepted. This kind of behavior seems to violate the rationality assumption of agents and deserves further investigation. One of possible reasons is that people may have different risk preferences. Conservative people tend to accept no risk, no matter how high the expected payoff is.

Although the risk preferences in principle can be included in the utility function of an agent, it is difficult to be encoded. We treat risk taking as a separated dimension of decision making from that of seeking the highest expected payoff. There are psychological justifications of how agents might want to take risk in terms of danger, slack, the aspiration level, assimilation and self-confidence [March and Shapira, 1992]. Risk measurement is quite a difficult task. The risk that a person feels in a game situation depends not only on the expected payoff in the game but also on the total amount of ones possession [Bernoulli 1954]. However, to model the detailed and the precise risk function of psychological parameters is not the purpose of this paper. We simplify the risk model into a measurable and comparable quantity in terms of payoff loss, so the agents can reason it.

In game theory, rational agents will try to reach Nash equilibrium, which is a stable outcome for each agent. All rational agents will not leave the Nash equilibrium they have reached, because whoever leaves alone will get a less payoff. However, the unique Nash equilibrium point maybe not always exists, even it does, it may not be the optimal solution. If the risk is taken into account other than the expected payoff, the stable outcome will be different, therefore, be more complicated to find the Nash equilibrium. In

order to coordinate both agents in an uncertain game, the risk needs to be controlled in addition to the expected payoff in a game.

Wu and Soo proposed two communication actions of asking guarantee and offering compensation that can help to coordinate rational agents [Wu and Soo, 1998a, 1999]. The mechanism is involves a trusted third party into the conventional game. The guarantee communication action can prevent an agent from playing strategies that will lead to a worse result for other agents. If P doesn't want Q to play strategies that will lead to a less payoff result for P. P may ask Q to deposit some guarantee at a trusted third party to ensure the strategies that lead to undesirable states will not be played. If Q keeps his commitment, the guarantee will be returned. The compensation communication action, on the other hand, is used to persuade an agent to play certain strategy that can lead to a desirable state. P may offer some compensation to Q in order to persuade him play certain strategy toward a better state for P. The compensation is deposited at a trusted third party and will be sent to Q if Q does play the strategy asked. With this negotiation protocol, Wu and Soo showed that rational agents can reach a Pareto-efficient and Nash equilibrium in all possible games.

In this paper, we extend the two communication actions to the new notion of the risk control under uncertain games. These communication actions also require involving a trusted third party in the game to coordinate the two agents.

In section 2, we introduce the new notion of the risk model in the game theory and then use it to redefine the strategy dominance, the Nash equilibrium and the Pareto efficiency concepts in traditional game theory. In section 3, we develop negotiation procedures with a trusted third party that allows agents to control the risk and expected payoff at the same time. We make discussion in section 4 and conclusion in the final section.

2 Risk Models in the Game Theory

Supposed that in a game, there are two strategies for an agent to play that give the same expected payoff. If the agent plays the first strategy, he will get a payoff of 5 with probability 0.5 and a payoff of 15 with probability 0.5, then the expected payoff is 10. If the agent plays the second strategy, he may get a payoff of 1000 with probability 0.01 and a payoff of 0 with probability 0.99, then the expected payoff is also 10. The two strategies are indifferent to a rational agent who makes decision based merely on the expected payoff. However, the risk-averse agent will play the first strategy even if the second strategy can give the highest payoff with probability 0.01. Suppose a risk-averse agent will always play the conservative strategy even if other strategies may give higher expected payoff. On the other hand, a risk-seeking agent may choose to play the second strategy for the highest payoff 1000. The risk-seeking agent will always play a strategy that might give the highest payoff, no matter how high the risk is. The rational agent assumption that maximizes the expected payoff seems not to be enough to explain the behavior. It is necessary to establish some risk model in order to clarify and explain the decision-making behavior of rational agents under risk.

2.1 Assumption on Risk Preference

We assume an agent may have different risk preferences. Here we discuss three typical types of risk preference for agents: risk-averse, risk-neutral, and risk-seeking. The risk-averse agent prefers to play the strategy that takes the least risk. A risk-seeking agent prefers to play the strategy that gives the highest payoff without taking the risk into account. A risk-neutral agent only tries to get the highest expected utility. In real situation, the risk preference of human agents may lay somewhat among the three extreme types of agents. But, for simplicity, we only analyze the three typical types of agents and leave the interpolation between the types into discussion and future works.

2.2 The Definitions and Notations

The expected payoff for each agent in a combination of strategies is denoted as $EP_i(s_i, s_{-i})$, the least payoff of a strategy is $LP_i(s_i, s_{-i})$, the highest payoff of a strategy is $HP_i(s_i, s_{-i})$, where i is the index of agents, and $-i$ denotes the other agents; agent i plays strategy s_i while other agent $-i$ plays strategy s_{-i}. The risk for each agent in a combination of strategies is denoted as $R_i(s_i, s_{-i})$. Without loss of generality, we define the risk of taking certain strategy combination as the expected payoff minus the least payoff, denoted as: $R_i(s_i, s_{-i}) = EP_i(s_i, s_{-i}) - LP_i(s_i, s_{-i})$. A further discussion about the risk is in section 4.

The dominance of a strategy: In traditional game theory, a dominant strategy for a rational agent is defined as a strategy that will always give a higher payoff no matter what strategies the other agent plays [Rasmusen 1989]. However, the notion of a dominant strategy can be extended to take the risk preference into consideration. For a risk-averse agent, the strategy that takes the least risk has the highest dominance. For a risk-seeking agent, the strategy that may give the highest payoff has the highest dominance regardless of the risk and the expected payoff. For a risk-neutral agent, the dominant strategy is the one that gives the highest expected payoff.

The Nash equilibrium under the risk model: In traditional game theory, a strategy combination (s_i^*, s_{-i}^*) is a Nash equilibrium if any agent gets less *payoff* when he deviates from the strategy combination alone [Nash, 1951; Rasmusen; 1989]. Under the risk model, a strategy combination (s_i^*, s_{-i}^*) is Nash equilibrium if any agent gets less *dominance* when he deviates from the strategy combination alone. That is, for a risk-neutral agent, deviates from the

Q \ P	C	D
A	15/0.5,25/0.5 5/0.5,15/0.5	0/0.9,1000/0.01 5/0.5,15/0.5
B	5/0.5,15/0.5 0/0.9,1000/0.01	0/0.9,2000/0.01 0/0.9,1000/0.01

(a)

Q \ P	C	D
A	20 10	10 10
B	10 10	20 10

(b)

Q \ P	C	D
A	5 5	10 5
B	5 10	20 10

(c)

Q \ P	C	D
A	25 15	1000 15
B	15 1000	2000 1000

(d)

Figure 1. An uncertain game represented in (a) payoff/probability (b) expected payoff (c) risk, and (d) highest payoff matrix forms respectively.

Nash equilibrium will decrease the expected payoff. For a risk-averse agent, deviates from the Nash equilibrium will increase the risk. For a risk-seeking agent, deviates from the Nash equilibrium will decrease the highest possible payoff.

The Pareto-efficiency under the risk model: In traditional game theory, 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 [Rasmusen, 1989]. Under the risk model, a strategy combination is Pareto-efficient if no other strategy combination increases the dominance of one agent without decreasing the dominance of the other agent.

Pareto-dominant under the risk model: The similar to the Nash equilibrium and the Pareto-efficiency concept, under the risk model, we may define the Pareto-dominant in terms of dominance instead of payoff. A strategy combination X strongly Pareto-dominates another strategy combination Y, then all agents have higher dominance at X.

2.3 Agent Decision-Making under of the New Dominance Concept

Take the uncertain game matrix in Figure 1(a) as an example. The probability associated with each payoff reflects the uncertainty of a nature move. Each entry X/Y in the game matrix denotes payoff X with probability Y. We can simplify the game matrix into an expected payoff form as shown in Figure 1 (b), which is a game that gives two Nash equilibrium points. Since the expected payoffs are the same in all the outcomes in Figure 1(b), Q makes no difference on each strategy. If Q has no risk preference, the uncertainty in Q's decision is an unavoidable risk for P. However, if Q is risk-averse, it will play strategy A to reduce the risk by reasoning on the game matrix in the risk form in Figure 1(c). Each entry is calculated in terms of the risk definition in section 2.2. If P knows that Q is risk-averse, P should play strategy C in order to get the maximum expected payoff. By doing so, both agents could get a strategy combination that gives the highest expected utility while reducing risk at the same time. However, if Q is risk-seeking, it will play strategy B for the highest possible payoff as in the game matrix in Figure 1(d). If P knows that Q is risk-seeking, P should play strategy D in order to get the maximum expected payoff. By doing so, both agents could get a strategy combination that gives the highest expected utility while seeking for the possible highest payoff at the same time. Therefore by knowing other agent's risk preference, each agent would know which

Q \ P	C	D
A	0/0.5, 40/0.5 5/0.8, 40/0.2	10/0.5, 20/0.5 0/0.33, 30/0.67
B	15/0.4, 30/0.6 10/0.1, 20/0.9	15/0.4, 40/0.6 0/0.75, 40/0.25

(a)

Q \ P	C	D
A	20 12	15 20
B	24 19	30 10

(b)

Q \ P	C	D
A	40 40	20 30
B	30 20	40 40

(c)

Q \ P	C	D
A	20 7	5 20
B	9 9	15 10

(d)

Figure 2. An uncertain game matrix (a) and its (b) expected payoff (c) highest payoff (d) risk matrix forms respectively

strategy has the highest dominance of other agents. However, there is still a problem, how agents reach some particular outcome that is the Pareto-efficient and Nash equilibrium? It needs a negotiation protocol to coordinate agents to reach a stable outcome.

3 Risk Control in a Multi-agent Negotiation Protocol

By taking the risk preference into consideration, the negotiation decision making will be based on the new notion of dominance instead of payoff as in the previous work. Here, we describe how the two communication actions can be used in a negotiation protocol to achieve the Pareto-efficient Nash equilibrium in the sense of dominance.

3.1 An Example of Negotiation Protocol

For simplicity, in the following communication protocol descriptions, the trusted third party is literally ignored.

Consider an uncertain game matrix in Figure 2(a). Figure 2(b)(c)(d) are the corresponding expected payoff, the highest payoff and the risk matrix forms respectively. The circles in Figure 2(b)(c)(d) indicate the most desirable states in its corresponding game matrix. There is no Nash equilibrium in Figure 2(b) and Figure 2(d) while two in Figure 2(c). Therefore, it is necessary for rational agents to negotiate a stable result. If both P and Q are risk-neutral, then this game can be solved according to the guarantee/compensation negotiation protocol based on the expected payoff matrix in Figure 2(b). Q will pay guarantee 8 not to play strategy A and will also pay a compensation 7 to ask P to play strategy C. This resulted in a new expected payoff game matrix in Figure 3(a). It can be shown that the strategy combination (B, C) is Pareto-efficient and the Nash equilibrium.

If Q is risk-averse and P is risk-seeking. Q may pay a guarantee 8 not to play strategy B and this results in a new game matrix in Figure 3(b). In Figure 3(b), the strategy combination (A, C) is a Pareto-efficient and Nash equilibrium subjected to the risk preference conditions that Q takes the minimum risk in Figure 2(d) and P gets the highest possible payoff in Figure 2(c). If Q is risk-natural and P is risk-averse. Q may pay a guarantee 8 not to play strategy B and P may pay a guarantee 6 not to play D and it forms the strategy combination (A, D) a Pareto-efficient and Nash equilibrium. If Q is risk-seeking and P is risk-neutral, Q may pay a guarantee 11 not to play strategy A and P may pay a guarantee 6 not to play D and it forms the strategy combination (B, D) also Pareto-efficient and the Nash equilibrium.

It turns out that in the above example, with the negotiation protocol, all strategy combinations (A, C), (A, D), (B, C), (B, D) can be made as a unique Pareto-

efficient and Nash equilibrium point, depending on different combination of the risk preference of both agents. Therefore, the rational agent can control the final negotiated state they will reach. The final state may satisfy the different risk preference of each agent. But, of course, the expected payoff may be affected in each encounter of agents who have different risk preference from risk-neutral one.

3.2 General Negotiation Protocol for Risk Control

The general protocol of the negotiation is:

Procedure: Negotiation Protocol for Risk Control

Input: the game matrix with payoff/probability entries

Output: a game matrix with a unique Pareto-efficient and Nash equilibrium

1. Construct the three game matrices (Expected Payoff, Risk, Highest Payoff)
2. Make a proposal that using guarantee or compensation communication actions based on the preference and rational assumption of other agents.
3. If the game still doesn't have a unique Pareto-efficient and Nash equilibrium, then go to step 2.
4. Play the game.

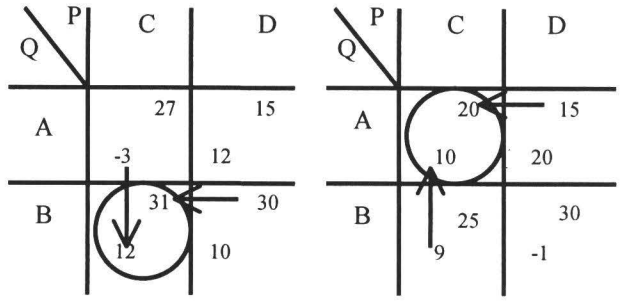
End of the Protocol

Note that **Step 2** is different from the previous work. In the previous work, agents just need to analyze one matrix, but in this paper, an agent must analyze on an appropriate game matrix under the risk preference and find out the highest dominance that is consistent with risk preference. If no dominant strategy can be found under the new dominance concept, then the agent needs to adopt asking guarantee or offering compensation negotiation actions to create one.

4. Discussion

In this paper, we define the risk in terms of a quantity that is calculated as the difference between the expected payoff and the least possible payoff. The definition of the risk can be a problem since the risk of losing or gaining certain amount of payoff may mean differently and non-linear to people. In fact, the risk measurement of certain amount of payoff was pointed out be a logarithm function [Bernoulli 1954]. The model of the risk preference discussed only three extreme types of the risk preference. It is too simple to reflect the true human risk preference in reality. Different functions of mapping payoff to utility will lead to different risk preference agents. A risk-averse agent maps less utility for a marginal payoff while a risk-seeking agent maps more utility for a marginal payoff [Hertz, 1983]. However, it is hard to obtain such an individual non-linear mapping function.

A risk preference can be a mixed weighting combination of three criteria: 1. Maximizing the expected value, 2. Minimizing the expectation variance, and 3 Minimize the probability of an unacceptable low value [Fabrycky 1980]. A moderate risk-averse agent, for example, may also have a tolerance bound on risk. If the risk is under the tolerance bound then he will take the risk to seek for a higher expected payoff. On the other hand, if the risk is over her tolerance bound, he will not play the strategy. Including such a tolerance bound into agent's decision making will indeed cause a more complex reasoning for the negotiation protocol to come up with a stable outcome. We leave it as future work.



(a) Q pays guarantee 8 not to play A and pays compensation 7 for asking P to play C

(b) Q pays guarantee 8 not to play B

Figure 3. The resulting game matrix after negotiation to get a Pareto-efficient and Nash equilibrium based on the expected payoff, if (a) both P and Q are risk-neutral. (b) Q is risk-averse and P is risk-seeking.

The assumption of rationality of agents is redefined in the new notion of dominance instead of the expected payoff. Therefore an agent is still rational if he always maximizes the expected payoff that is subjected to his risk preference. But he sometimes would be treated as irrational in the old sense of rationality that is trying to maximize the expected payoff only.

We can prove that the proposed negotiation procedure with risk control will end within finite time. A similar convergence theorem as well as the one of the feasibility of creating a unique Pareto-efficient and Nash equilibrium using the guarantee/compensation negotiation protocol via trusted third party can be proved [Wu and Soo, 1999].

5 Conclusions

We separated the risk preference from utility function as another dimension other than the expected payoff. To model the preference of taking risk, we defined three types of agents with different risk preferences. In this way, in game decision making, we redefined the concept of dominance so

that the optimal strategy that agents select can take risk into account. As illustrated in the above examples, we showed that different risk preference agents in the same uncertain game could lead to different outcomes that satisfy the agent's risk preference. Therefore the agent can in some sense control the risk of the outcomes of his decision making.

The purpose of agent negotiation is for the agents to reach a stable and efficient result. Under the risk preference model, agents negotiate to control not only the expected payoff but also the risk. The mechanisms of communication actions of guarantee and compensation help agents to reduce the risk due to unpredictable behaviors of other agents in games, especially in situations of there is no unique Nash equilibrium or there is a unique Nash equilibrium but not Pareto-efficient, e.g. prisoner's dilemma. The separation of the risk from the payoff and the redefinition of dominance in terms of different risk preferences allow agents to take risk control into decision making can reduce the risk due to the uncertainty of a nature move.

Acknowledgements

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

Reference

[Axelrod, 1984] R. Axelrod, *The Evolution of Cooperation*, Basic Books Inc., New York, 1984.

[Brafman and Tennenholtz, 1997] R.I. Brafman and M. Tennenholtz, Modeling Agents as Qualitative Decision Makers, *Artificial Intelligence*, Vol.94, pp217-268, 1994

[Bernoulli, 1954] D. Bernoulli, Exposition of a New Theory of the Measurement of Risk, *Econometrica*, Vol. 22, pp. 23-36.

[Durfee et al., 1993] 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.

[Fabrycky, 1980] W.J. Fabrycky and G.J. Thuesen, *Economic Decision Analysis* 2nd ed. 1980.

[Gmytrasiewicz et al., 1991] P.J. Gmytrasiewicz, E.H. Durfee and D.K. Wehe, The Utility of Communication in Coordinating Intelligent Agents, in *Proceedings of the Ninth National Conference on Artificial Intelligence* (AAAI-91), pp.166-172, 1991.

[Haynes and Sen, 1996] T. Haynes and S. Sen, Satisfying User Preferences while Negotiating Meetings, *in Proceeding of the Second International Conference on Multi-Agent Systems (ICMAS-96)*, 1996.

[Hertz, 1983] D. B. Hertz, *Risk Analysis and its Applications, Chichester*, John Wiley & Sons, 1983.

[Koller and Pfeffer, 1997] D. Koller and A. Pfeffer, Representations and Solutions for Game-theoretic Problems, *Artificial Intelligence*, Vol.94, pp 167-215, 1997.

[March and Shapira, 1992] James G. March and Zur Shapira, Variable Risk Preferences and the Focus of Attention, *Psychological Review*, 99(1), pp. 172-183, January, 1992.

[Mor and Rosenschein, 1995] 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.

[Nash 1951] J. F. Nash, Non-cooperative Games, *Ann. of Math.* 54, pp.286-295, 1951.

[Rasmusen 1989] E. Rasmusen, *Games and Information: An Introduction to Game Theory*, Basil Blackwell, Oxford, 1989.

[Rosenschein and Genesereth, 1985] J.S. Rosenschein, M.R. Genesereth, Deals among Rational Agents, *in Proceedings of the ninth international conference on Artificial Intel*ligence(*IJCAI-85)*, pp.91-99. 1985.

[Rosenschein and Zlotkin, 1994] J.S. Rosenschein and G. Zlotkin, *Rules of Encounter*, MIT Press, Cambridge, 1994.

[Sandholm and Lesser, 1995] T.W. Sandholm and V.R. Lesser, Issues in Automated Negotiation and Electronic Commerce: Extending the Contract Net Framework, in *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pp.328-335, MIT Press, 1995.

[Shehory and Kraus, 1998] O. Shehory and S. Kraus, Methods for Task Allocation via Agent Coalition Formation, *Artificial Intelligence*, Vol.101, pp.165-200, 1998.

[Tennenholtz 1998] M. Tennenholtz, On Stable Social Laws and Qualitative Equilibria, *Artificial Intelligence*, Vol.102, pp.1-20, 1998.

[Vidal and Durfee, 1995] J.M. Vidal and E.H. Durfee, Recursive Agent Modeling using Limited Rationality, in *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pp. 376-383, MIT Press, 1995.

[Wu and Soo, 1998a] S. Wu, V. Soo, Escape from a Prisoners' Dilemma by Communication with a Trusted Third Party, in *Proceeding of the Tenth International Conference on Tools with Artificial intelligent* (ICTAI'98), pp.58-65, 1998.

[Wu and Soo, 1998b] S. Wu, V. Soo, A Fuzzy Game Theoretic Approach to Multi-Agent Coordination, *Pacific Rim international workshop of on Multi-Agent* (PRIMA'98), 1998.

[Wu and Soo, 1999] S. Wu, V. Soo, Game Theoretic Reasoning in Multi-agent Coordination by Negotiation with a Trusted Third Party, in *Proceeding of the Third International Conference on Autonomous Agents* (Agents'99), Seattle, Washington, 1999.