

Finding Risk Dominance Strategy in Imperfect Game, A Theory Perspective

Seyyed Mohammad Reza Farshchi¹, Irene Kehvazadeh²
1,2- Department of Data Mining Advanced Research Center, Mashhad Branch,
Iran. Email: Shiveex@Gmail.Com

ABSTRACT:

Game theoretic reasoning pervades economic theory and is used widely in other social and behavioral sciences. It is widely used in decision of economic behavior and other game conditions with risk characters. A model is an abstraction we use to understand our observations and experiences. This paper proposes an imperfect information multi-player system using risk dominance strategy for its decision module. In this article, the Influence of the strategy in different conditions is discussed. A model is unlikely to help us understand a phenomenon if its assumptions are wildly at odds with our observations. The cases that risk dominance worked and the process of the module to exclude risk is illustrated. In the end, basing on the results of experiments, the performance of the strategy is discussed.

KEYWORDS: Game theory, Imperfect game, Decision tree.

1. INTRODUCTION

Risk dominance and payoff dominance are two related refinements of the Nash Equilibrium (NE) solution concept in game theory, defined by John Harsanyi and Reinhard Selten[2]. Nash equilibrium is considered payoff dominant if it is Pareto superior to all other Nash equilibrium in the game. When faced with a choice among equilibrium, all players would agree on the payoff dominant equilibrium since it offers at least as much payoff as each player's best alternative. Conversely, Nash equilibrium is considered risk dominant if it has the largest basin of attraction [3], meaning the more uncertainty players have about the actions of the other player(s), the more likely they will choose the risk dominant strategy [4].

A game with imperfect information means the players face risks not only come from the uncertain strategies of other players but also from the uncertain game conditions [5]. Thus, risk dominance strategy is imported as the decision module of an imperfect information game called Military Chess.

The main contribution of the system that is presented is the realization of a decision module basing on risk dominance strategy which is used for the selection of optimal step in the Military game. The performance of the strategy is also discussed in the following experiments.

This paper is organized as follows. Section 2 briefly introduces the related theories and works about the risk dominance strategy in game theory. Section 3 gives a primary introduction of Military Chess game. And then, in Section 4, a specific description is presented about the risk dominance strategy and the realization details of the decision module in our game system.

Section 5 analyzes the performance of the new system that is tested locally and on the Internet. And finally, section 6 contains our conclusion.

2. ITERATED ELIMINATION OF WEAKLY DOMINATED ACTIONS

A strictly dominated action is clearly unattractive to a rational player. Now consider an action a_i is weakly dominated in the sense that there is another action that yields at least as high a payoff as does a_i whatever the other players choose and yields a higher payoff than does a_i for some choice of the other players. In the game in Figure 1, for example, the action T of player 1 weakly (though not strictly) dominates B.

A weakly dominated action that is not strictly dominated, unlike a strictly dominated one, is not an unambiguously poor choice: by Lemma 356.1 such an action is a best response to some belief. For example, in the game in Figure 1, if player 1 is sure that player 2 will choose R then B is an optimal choice for her. However, the rationale for choosing a weakly dominated action is very weak: there is no advantage to a player's choosing a weakly dominated action, whatever her belief. For example, if player 1 in the game in Figure 1 has the slightest suspicion that player 2 might choose L then T is better than B, and even if player 2 chooses R, T is no worse than B.

If we argue that it is unreasonable for a player to choose a weakly dominated action then we can argue also that each player should work under the assumption that her opponents will not choose weakly dominated actions, and they will assume that she does not do so, and so on. Thus, as in the case of strictly dominated

actions, we can argue that weakly dominated actions should be removed iteratively from the game. That is, first we should mark actions of player 1 that are weakly dominated; then, without removing these actions of player 1, mark actions of player 2 that are weakly dominated, and proceed similarly with the other players. Then we should remove all the marked actions, and again mark weakly dominated actions for every player. Once again, having marked weakly dominated actions for every player, we should remove all the actions and go through the process again. We should repeat the process until no more actions can be eliminated for any player. This procedure, however, is less compelling than the iterative removal of strictly dominated actions since the set of actions that survive may depend on whether we remove all the weakly dominated actions at each round, or only some of them, as the two-player game in Figure 1 shows. The sequence in which we first eliminate L (weakly dominated by C) and then T (weakly dominated by B) leads to an outcome in which player 1 chooses B and the payoff profile is (1, 2). On the other hand, the sequence in which we first eliminate R (weakly dominated by C) and then B (weakly dominated by T) leads to an outcome in which player 1 chooses T and the payoff profile is (1, 1).

	L	C	R
T	1, 1	1, 1	0, 0
B	0, 0	1, 2	1, 2

Figure 1. A two-player game in which the set of actions that survive iterated elimination of weakly dominated actions depends on the order in which actions are eliminated.

In the games studied in this chapter, the players are representatives from an evolving population of organisms (humans, animals, plants, bacteria, ...). Each player's payoffs measure the increments in the player's biological fitness, or reproductive success (e.g. expected number of healthy offspring), associated with the possible outcomes, rather than indicating the player's subjective feelings about the outcomes.

Each player's actions are modes of behavior that the player is programmed to follow.

The players do not make conscious choices. Rather, each player's mode of behavior comes from one of two sources: with high probability it is inherited from the player's parent (or parents), and with low (but positive) probability it is assigned to the player as the result of a mutation. For most of the models in this article, inheritance is conceived very simply: each player has a single parent, and, unless it is a mutant, simply takes the same action as does its parent. This model of inheritance captures the essential features of both

genetic inheritance and social inheritance: players either follow the programs encoded in their genes, which come from their parents, or learn how to behave by imitating their parents. The distinction between genetic and social evolution may be significant if we wish to change society, but is insignificant for most of the models considered in this article.

Members of a single large population of organisms are repeatedly randomly matched in pairs. The set of possible modes of behavior of each member of any pair is the same, and the consequence of an interaction for an organism depends only on the actions of the organism and its opponent, not on its name. As an example, think of a population of identical animals, pairs of which periodically are engaged in conflicts (over prey, for example). The actions available to each animal may correspond to various degrees of aggression, and the outcome for each animal depends only on its degree of aggression and that of its opponent. Each organism produces offspring (reproduction is asexual), to each of whom, with high probability, it passes on its mode of behavior; with low probability, each offspring is a mutant that adopts some other mode of behavior.

Harsanyi and Selten's risk dominance is based on what they refer to as a tracing procedure [2], the details of which are beyond the scope of this paper. As an alternative, Selten proposes a measure of risk dominance that is easy to calculate for our games [6], and is indicative of the outcome of the tracing procedure. Let $u_1(X, X)$ be the payoff of player 1 with strategy pair (X, X) , Selten's [6] weighted average log measure of risk dominance of the equilibrium (A, A) over (B, B) is given by:

$$R = \log \left[\frac{u_1(A, A) - u_1(B, A)}{u_1(B, B) - u_1(A, B)} \right] \quad (1)$$

If R is positive, Harsanyi and Selten's tracing procedure selects (A, A) as risk dominance. If R is zero, the mixed strategy Nash equilibrium is risk dominant. If R is negative, (B, B) is risk dominant. Notice that any affine transformation of the payoffs in the game would not change either the sign or the magnitude of R .

This criterion can explain the experimental result of Paul Straub. In Game 1, R of player 1 can be calculated by (1) that equal to $\log(2.5)$, so (A, A) is both payoff and risk dominance. In Game 2, R of player 1 changes to $\log(0.67)$ which is a negative value, so the strategy B has risk dominance. That is why the players picked strategy B in the final three rounds of the session utilizing Game 2.

There are two main factors that affect the risk characteristics of the games [7]. One is the trust level between the players as showing in table 2. Generally speaking, the coordinate game will keep a comparative higher trust level which weakens the influence of risk dominance. However, in uncoordinated games and zero

sum games, there is almost no trust between players. In another side, the imperfect information game situation will greatly enhance the uncertainty of the game, thus the effect of risk dominance strategy should be carefully considered in decision process.

The notion of Nash equilibrium requires only that each player's strategy be optimal in the whole game, given the other players' strategies; after histories that do not occur if the players follow their strategies, the actions specified by a player's Nash equilibrium strategy may not be optimal. In some cases we can think of the actions prescribed by a strategy for histories that will not occur if the players follow their strategies as "threats"; the notion of Nash equilibrium does not require that it be optimal for a player to carry out these threats if called upon to do so. In the previous chapter we studied the notion of subgame perfect equilibrium, which does impose such a requirement: a strategy profile is a subgame perfect equilibrium if every player's strategy is optimal not only in the whole game, but after every history (including histories that do not occur if the players adhere to their strategies).

Table 1. Payoff and risk dominance in two games condition

Game1	Player2			
Player 1	A	A	8	3
		B	4	6
	B	A	4	6
		B	3	6
Game2	Player 2			
Player 1	A	A	8	4
		B	0	6
	B	A	0	6
		B	4	6

Table 2. Games with low trust level, risk dominance will be very important in decision process

Game	Player2			
Player1	A	A(20%)	8	4
		B(80%)	0	6
	B	A(20%)	0	6
		B(80%)	4	6

Table 3. The four strategies of player 2 in the game in Figure 1.

	Action assigned to history C	Action assigned to history D
Strategy 1	E	G
Strategy 2	E	H
Strategy 3	F	G
Strategy 4	F	H

3. MILITARY CHESS GAME SYSTEM WITH MONTE-CARLO SAMPLING

In this section, the imperfect information game system basing on a game called Military Chess is presented.

3.1. Introduction for Proposed Game

The Military Chess, which is also called Kriegspiel, has been a very popular game for many years. The rules of the Military Chess are quite complex. Simply speaking, the basic process of the game is to move pieces, attack opponent's pieces and finally occupy the position of enemy's flag. Maybe it sounds quite like Chess but there are at least three main differences between them as following.

Opposite and two opponents sits aside, like Bridge. Second, the pieces which have a certain military rank can only destroy those with lower ranks. Last but most important, player cannot see the ranks of other player's pieces, even though it belongs to his associate. As the result, players have to guess the ranks of other pieces. This is the most difficult and complex point of the development of Military Chess game system. A Military Chess game system is developed which is called mygames [11].

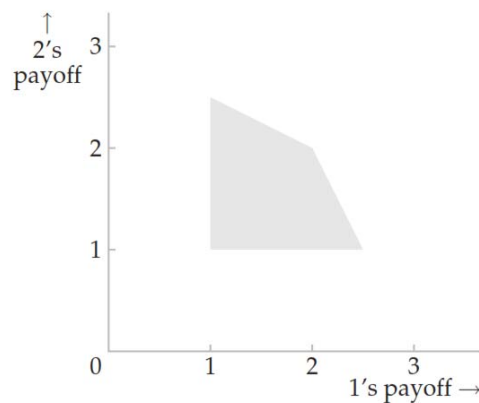


Figure 2. The approximate set of Nash equilibrium discounted average payoffs for the infinitely repeated Prisoner's Dilemma with one-shot payoffs to 1.

As in the case of a finite horizon game, if a strategy profile is a subgame perfect equilibrium then certainly it satisfies the one-deviation property, since no player

must be able to increase her payoff by any change in her strategy. What we need to show is the converse: if a strategy profile is not a subgame perfect equilibrium then there is some subgame in which the first-mover can increase her payoff by changing only her initial action.

3.2. Monte-Carlo Sampling

In imperfect information game, none of the players in the game have the whole knowledge about the current game state. In order to solve such problem in practice, one has to either resort to special cases, or fall back heuristics which is created by Monte-Carlo sampling [5].

Monte-Carlo sampling is an algorithm that deals with imperfect information games. When evaluating the choices of one player, the Monte-Carlo method randomly guesses the unseen pieces of the other players. The game is then searched using min-max game tree as if these pieces are known to each player. As demonstrated in figure 2, the square represents the max node and the circle represents the min node. Only the root node is the random (the diamond) and all the other nodes are board states with perfect information.

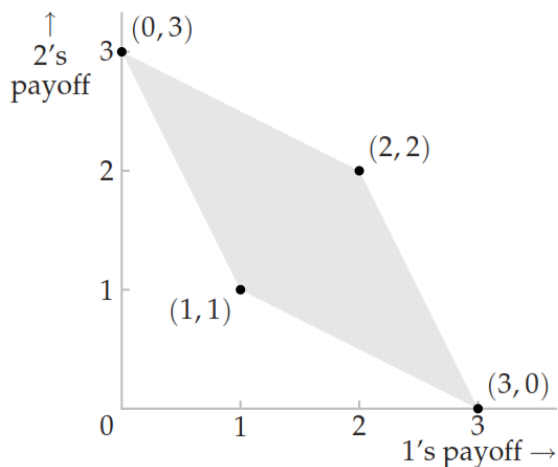


Figure 3. Game tree of imperfect information using Monte-Carlo sampling

3.3. Realization with Monte-Carlo Sampling

The set of Nash equilibria consists of the action profile (Hare,...,Hare) in which all hunters catch hares, and any action profile in which exactly k hunters pursue the stag and the remaining hunters catch hares.

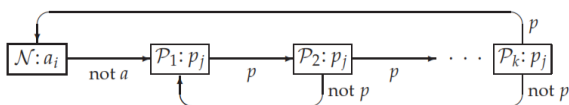


Figure 4. A subgame perfect equilibrium strategy for player i in a two-player infinitely repeated game.

The outcome p is that in which each player's action is one that holds the other player down to her minmax payoff.

In the process of game, when it is the agent's turn to move, the current state that agent faces is an imperfect information state. Monte-Carlo sampling process transforms the imperfect information state to perfect information state. Next, basing on the sampled perfect information, move generator module generates moves list and build the game tree. And then, tree search algorithm is used searching the optimal move from the game tree. When a move has been chosen, the system check the end condition, which could be set for a limit of running time or a limit times for system's iteration. This process will iterate many times which is set by users. Each iteration generates a candidate move. When the end condition is fulfilled, system will collect all the candidate moves in before iterations and decision module will take an analysis about them and finally choose the optimal move.

4. DECISION MODULE IN THE SYSTEM WITH RISK DOMINANCE STRATEGY

The task of decision module is to select one optimal move from tens of candidates. The criterion is the expect payoff of the moves. However, in the case of imperfect information, the uncertainty of the game condition leads the payoff for one move is uncertain. Decision module builds a suppositional game condition to deal with the problem. Just like the table 5 illustrates, player 1 denotes the player for whom the system make decision, and player 2 is set to denotes different Monte-Carlo sample results from the uncertain game circumstance which greatly influences the payoff of the player1's strategies.

For concisely express, set the numbers of candidate move and different conditions are both three. $E(X)$ means the expect payoff of move X , P_a means the appear probability of condition a and V_{Ab} means the payoff when condition b appears and player1 adopts move A . Let M means the cluster of all conditions and the system calculates the $E(X)$ as following:

$$E(X) = \sum_{i \in M} V_{Xi} P(i) \quad (2)$$

And the final optimal result will be decided by $\max\{E(A), E(B), E(C)\}$.

The system calculates the E value with equation (2) of each move, and then chooses the largest to be the optimal strategy. The process of decision can deal with most of the conditions well except some special cases which is also the reason the risk dominance strategy be adopted.

There are some cases that terrible departure lost appears just like table 6 shows. When condition b happens, candidate move B will leads a terrible lost.

The low appearance probability of condition b will smooth the loss of B when calculating $E(B)$. Just like no one wants to take the risk of low probability to eat candy in a bowl which is full of hundreds of candies and only one of them is poison, the task of risk dominance strategy in this case is to exclude B -like candidate moves from the final result.

The decision module creates the game condition like table 7 shows. Value (A, a) means the average payoff of other conditions and (B, b) means the payoff in the dangerous condition. Using equation (1) to calculate the value R is Log (5) which is a positive value and it can be concluded that the strategy A is risk dominance. That is the process that risk strategy B is excluded.

REFERENCES

- [1] Straub, P.G., "Risk dominance and coordination failures in static games", *Quart. Rev. Econ. Finance* 35, 1995, pp.339–363.
- [2] John C. Harsanyi and Reinhard Selten, "A General Theory of Equilibrium Selection in Games", MIT Press, 1988, ISBN 0262081733.
- [3] David Schmidt, Robert Shupp, James M. Walker, and Elinor Ostrom, "Playing safe in coordination games: the roles of risk dominance, payoff dominance, and history of play", *Games and Economic Behavior*, 2003, vol. 42, issue 2, pp.281-299.
- [4] Andreoni, J., Croson, R., "Partners versus strangers: random rematching in public goods experiments". Mimeo. Forthcoming in: Plott, C.R., Smith, V.L. (Eds.), *Handbook of Results in Experimental Economics*, 1998. Forthcoming in: Plott, C.R., Smith, V.L. (Eds.), *Handbook of Results in Experimental Economics*, 1998.
- [5] Howard James Bampton, "Solving imperfect information games using the Monte Carlo heuristic [Master thesis]", University of Tennessee, Knoxville, 1994.
- [6] Selten, R., "An axiomatic theory of a risk dominance measure for bipolar games with linear incentives", *Games Econ. Behav.*, 8, 1995, pp.213–263.
- [7] LIU Zongqian and WANG Feng, "Risk Dominance in Multi-equilibrium Selection and Mixed Strategy", Department of Mathematics, Capital Normal University, Beijing 100037, China, 2005.
- [8] Xuan Wang, Zhaoyang Xu, and Xiao Ma, "TD(A) optimization of imperfect information games evaluation function", Intelligence Computing Research Center Harbin Institute of Technology Shenzhen Graduate School, WCCGC, Japan, 2007.