Simple Decision Heuristics in Perfect Information Games

Naoki Konno and Kyoichi Kijima
Graduate School of Decision Science and Technology, Tokyo Institute of Technology
2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan
nkonno@valdes.titech.ac.jp
kijima@valdes.titech.ac.jp

ABSTRACT

In the real world, players generally recognize the possibility of misperception in their own internal decision model. Therefore, decision makers sometimes adopt some kinds of simple decision heuristics even when they play a game of perfect information. To describe such situations, we first propose a rough reasoning model and show the difference by applying it to the centipede games. Then we propose 4 kinds of simple decision heuristics in order to describe more irrational decision making situations, where the players adopt rational / optimistic/ pessimistic / risk neutral decision heuristics. We, then, investigate that effectiveness of each decision heuristics by applying them to the constant-sum games and centipede game. We first show that rational decision heuristics is powerful in the constant-sum situations. However, optimistic societies that consist of optimistic player achieve more desirable social welfare than rational societies in the centipede game situations.

Keywords: bounded rationality, simple decision heuristics, game theory, perfect information game, expected utility.

1. INTRODUCTION

The purpose of this paper is two-folds: The first is to propose 4 kinds of simple decision heuristics based on backward induction. The second is to examine the effectiveness of the heuristics by applying them to the constant-sum games and centipede game. In the real world, it is well known that decision makers not always compute equilibrium strategies. They sometimes simplify the decision model and consider only a little parts of the original game even if they play a game of perfect information. For example though backward induction provides reasonable equilibrium, expert chess players seldom consider all possible actions. It is because a sequence of seeking for a checkmate may lead to a fatal result if she makes a mistake. In this paper, we propose models of simple decision heuristics where the players only consider some parts of the tree instead of bakward induction.

In the traditional game theory, it is usually assumed that players completely recognize the game situation so as to compare all the results without error and to choose rational strategies. However, even in the extensive form games with perfect information, as Selten [2] pointed out by using chain store paradox, not only Nash equilibrium but also subgame perfect Nash equilibrium may lead to strange outcomes. Indeed, in perfect information games, though it is theoretically able to calculate reasonable equilibria by backward induction, it is practically difficult to realize them due to various complexity and the limitation of abilities.

In order to describe such kinds of limitation of abilities, Myerson [4] proposed a concept of trembling-hand equilibrium in normal form games. He refined equilibrium from a viewpoint that players make infinitesimal errors and small errors are more likely to occur than big errors. Along with the spirit of trembling-hand equilibrium, many refined equilibrium are proposed. [8] [9] On the other hand, noninfinitesimal errors has been studied by Mckelvey and Palfrey [5]. They examined developed the quantal response equilibria by substituting quantal response for best response in the sense that the players are more likely to choose better strategies than worse strategies but do not play a best response with probability 1. He also examined the further property by using quantal logit functions. They [6] transformed it in extensive form games by using agent normal form, and interpret quantal response equilibria as steady states by bounded rational players.

We should notice that, in these previous papers They describe action errors. A represented kind of action error is typing errors. For example, though she tries to type a error, displayed word is eror. On the other hand, we also make mistake in our reasoning. It is represented by miscalculation in chess. Though it may be difficult to formulate reasoning structure in normal form games, Heifetz and Pauzner [7] proposed a model with possibility of wrong reasoning in binary action games. Their model is also based on agent normal forms and error rate is given by exogenous manner. It is also important to emphasize that players do not always try to implement equilibrium strategies therefore they replace the equilibrium approach with heuristics approach.

In our model, it is assumed that players try to evaluate each action by rough reasoning, then choose the best action in their internal model, while they make wrong evaluation with certain probabilities in it. We look objectively at these game situations, then describe the result of rough reasoning decision making from the viewpoints of outside observers.

We characterize player's rough reasoning by following two sides. One is the payoff, while the other is the depth of the tree.

.First, as Mckelvey and Palfrey [6] argued, we assume reasoning accuracy depends on the difference of the payoffs in such a way that error rate is a decreasing function of the difference of payoffs. In addition, we also suppose that as the depth of decision tree increases, reasoning accuracy tends to decrease. This property describes why it is difficult to compare actions in the far future. We call the model, a general rough reasoning model. In addition we adopt two more assumptions about reasoning error rate. The first is that as payoff difference decreases, error rate increases exponentially. The second is as depth of tree increases, error rate increases exponentially.

This paper claims that the backward induction solution concept hinges on the assumption that players are absolutely certain of their reasoning. We consider a slight deviation from this assumption and show that, in some cases, the prediction of the model change considerably.

We first show the result of the rough reasoning model in the centipede games in order to show the difference between rational player paradigm and rough reasoning paradigm.

The centipede game is investigated by Rosenthal. [7] This is known for the discrepancy between equilibrium obtained by backward induction and actual experimental results. According to Mckelvey and Palfrey [4], the unique subgame perfect Nash equilibrium outcome is not so often observed. They tried to rationalize the results by mathematical model in which some of the players have altruistic preferences. Aumann [10] [11] insisted that incompleteness of common knowledge assumption causes cooperative behaviors. Although these factors may work in the centipede game, we claim rough reasoning is also an essential factor leading to cooperative behaviors.

If players understand their imperfectness of reasoning ability, they may not choose the rational action in their internal model. We propose 4 kinds of simple decision heuristics where partially rational, maxmax, maxmin, and average evaluation heuristics.

In partially rational decision heuristics, players characterized as a rational agent. In maxmax heuristics, players characterized as a optimistic agent. In maxmin heuristics, players characterized as a pessimistic agent. In average evaluation heuristics, players characterized as a risk neutral agent.

To examine effectiveness of the heuristics, we apply them to the constant-sum game and centipede game. Constant-sum game is one of the most basic kind of games. In constant sum games, the sum of all players' payoffs is the same for any outcome. Hence, a gain for one participant is always at the expense of another, such as in most sporting events. Since payoffs can always be normalized, constant sum games may be represented as (and are equivalent to) zero sum game in which the sum of all players' payoffs is always zero. If players' reasoning is complete, obviously backward induction provides equilibrium strategies.

This paper is organized as follows. In Section 2, we provides rough reasoning model which describe the limitation of our reasonig ability. In Section 4, we provide results of the rough reasoning model in centipede game and show that the difference from the previous works. Section 4 presents 4 types of simple reasoning heuristics based on rough reasoning model. We then apply them to constant sum games and centipede games and show that the effectiveness of each type of simple decision heuristics in Section 5. Finally some conclusions and remarks are given in Section 6.

2. GENERAL ROUGH REASONING MODEL

In the traditional game theory, it is usually assumed that all players perceive situation precisely, and essentially compare all the strategies without error. However, such kind of rationality is quite unrealistic in most actual decision situations due to the players' abilities. Especially, reasoning process is quite complication in usual, we deal here with only perfect information games. It is because that reasoning process in perfect information games is easy to study. We first define the true objective game.

Definition2.1. True objective game is a finite perfect information extensive form game given by

$$G = (I, N, A, D, P, (r_i))$$

where I is the set of players, while N is the set of nodes. N_T and N_D are partitions of N, where N_T is the set of terminal nodes and N_D is the set of decision nodes. A is the set of actions D is the function from nodes except initial node n_1 to the prior nodes. P is the player function that determines the player who chooses an action at the node. (r_i) is the payoff function that determines the payoffs of each player.

Since G is a perfect information game, subgame perfect equilibria are obtained by backward induction. However, since the players can not compare all the result without error in the actual situations, we assume that players choose actions by the following heuristics.

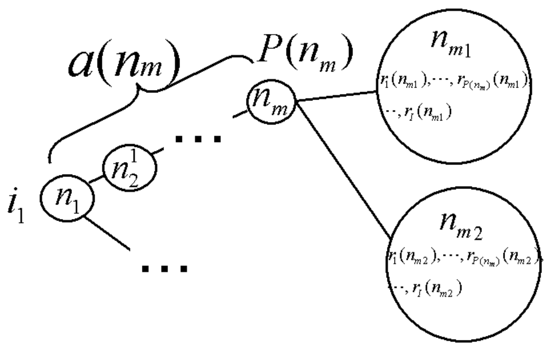

Fig.1 Notations of rough reasoning model

To implement it, we need some notations:

N_2: The set of attainable nodes from n_1

N_m^1: The set of the last decision nodes of G.

a(n) : The depth from reasoned node to the initial node.

$r(n_d) = (r_1(n_d),...,r_j(n_d),...,r_I(n_d))$: a payoff vector that the player reasons to achieve the optimal choices are taken at every stage

Then the heuristics are described as follows

(1) Let i_1 be the player that chooses an action at the initial node n_1. i_1 tries to reason the estimated payoff vector at $n_2 \in N_2$ by backward induction
(2) Indeed, i_1 tries to reason estimated payoff vector at node $n_m \in N_m^1$. Let *a* be the depth form the initial node to n_m. Let *b* be the payoff difference between two nodes. We assume that the error probability is an increasing function of *a* and a decreasing function of *b*. If there are some best responses, each best action is taken with same probability.
(3) When the above operations have been finished for every, $n_m \in N_m^1$. i_1 identifies every $n_m \in N_m^1$ with terminal nodes. Then i_1 generates a set of last decision nodes of a new truncated game. Start to reason next reasoning process. This process is iterated until n_2. By this process, $i_1$generates a payoff vector at n_2.
(4) Finally, i_1 compares the payoff vector of $n_2 \in N_m^2$ and chooses a best action. (This heuristics is a kind of backward induction with errors.)
(5) The players implement these processes until they reach a terminal node.

This process can be considered as a situation that all players are try to implement rational choices as possible as their reasoning abilities. The result is non-deterministic, we only take probability distribution over N_T. It is note that even if players' reasoning ability is not equal, they tries to reason by their reasoning abilities. Furthermore, if one player chooses actions more than once in the true game, reasoning at the subsequent nodes may contradict to that at the prior node. Our model can also describe such situations.

3. RESULTS OF ROUGH REASONING MODEL IN THE CENTIPEDE GAMES

3.1 Introduction of the centipede game

To examine difference between rational paradigm and rough reasoning paradigm, we focus on the Rosenthal's [7] centipede game by using more specific models. Centipede game is well known as an example illustrating differences between results by backward induction and those by actual experiments.

The centipede game is two-person finite perfect information game. We call player 1 is "she", and player 2 is "he". Each player alternately chooses Pass (P) or Take (T) in each decision node. If she chooses action P, her payoff decreases while his payoff increases by more than his decrease. If she chooses action T, the game is over and they receive payoffs at that node. Symmetrically if he chooses action P, his payoff decreases while her payoff increases by more than his decreases. If the game has n decision nodes, we call it the n-move centipede game. (Show Fig.2)

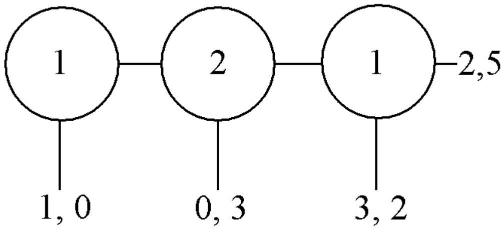

Fig.2 3-move centipede game

In the rational player paradigm, pair of strategies that both the players choose T at every decision node is unique subgame perfect equilibrium. It is because the centipede game has the last period, the player choose T is the rational action on the last decision nodes. Hence opponent considers that rational action at the (n-1)-th decision node is T because the rational action at the last node is T. This equilibrium leads to the result that the game is over at the first period.

The centipede game has many variants about payoff structures. However we adopt the original Rosenthal's structure, where if she chooses P, her payoff is decreased by 1 and his payoff is increased by 3.

3.2. Rough reasoning model based on logit function

In order to examine implications of rough reasoning, we propose a specific models, rough reasoning model based on logit function.

Suppose that player i at node n_k reasons about the decision node n_l. Then, we need the following notations.

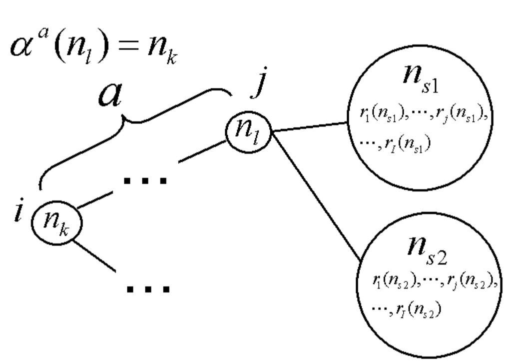

Fig. 3: Notations concerning rough reasoning with logit function.

j: The decision player at node n_l
N_s: A set of attainable nodes from n_l. i.e.

$$N_s = \{n \mid n \in N, \alpha(n) = n_l\}$$

σ : reasoning ability parameter with payoff scale.
We should notice that σ works as a fitting parameter with respect to the unit. For example, if description about payoffs change from dollar to cent, σ will be 1/100. Furthermore, if payoff scale is fixed equivalent, as the rationality of agent is increased, σ will be increased.

Then, the rough reasoning model based on logit function with parameter σ is defined, as follows:

Rough reasoning model based on logit function is a reasoning model that assigns n_{s1} to n_l with probability

$$\frac{e^{\frac{r_{P(n_m)}(n_{m1})}{a(n_m)}\sigma}}{\sum e^{\frac{r_{p(n_m)}(n_{mj})}{a(n_{mj})}\sigma}}$$

The probability essentially depends on the ratio of payoff against *a* in such a way that if *a* is sufficiently large, then the choice can be identical with random choice. If *a* is sufficiently small and *b* is sufficiently large, the choice can be seem as by the best response.

3.3 Rationality promotes rational action and social welfare?

In order to examine frequency of cooperative behavior P with relation with FCPF, we calculated several simulations, where FCPF denotes frequencies of choice P at first period. We focus on the choice at the first period, because if P is chosen at the first period, the remained subgame can be considered as the n-1move centipede game. Figure 4 shows the simulation results of FCPF.

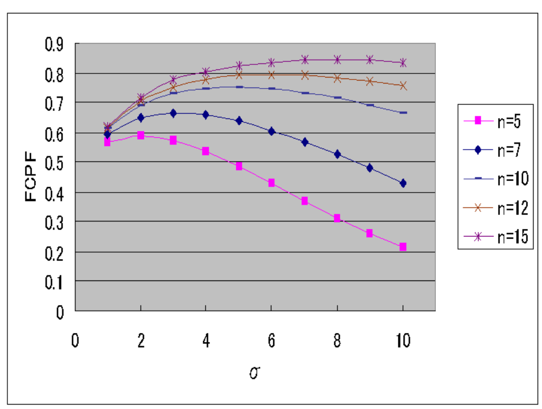

Fig.4: FCPF of logit function model

Note that in the Figure 4, larger σ implies the player has well reasoning ability. Figure 4 shows that for every n, it is observed that there is a turning point. Until the turning point, as the rationality is increased, cooperative behavior P tends to increase. However if reasoning ability exceeds the turning point, as the rationality is increased, cooperative behavior P tends to decrease.

Figures 4 gives following implications about the relation between FCPF and the reasoning ability: Moderate rationality may maximize the frequency of irrational actions. This property can be showed in the both case, so that, we think that this property holds that if human reasoning abilities proportional to the exponential rate of the depth of the tree.

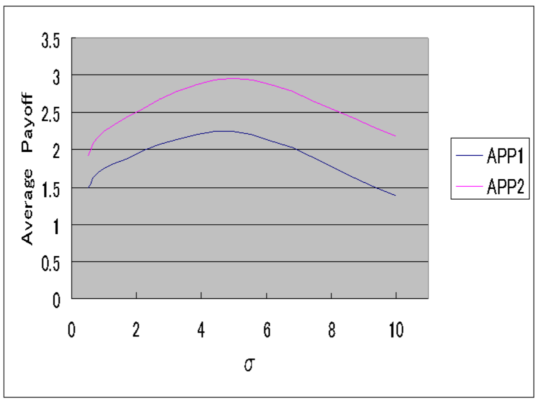

Fig.5: APP in the rough reasoning model

However other implication is obtained concerning relation between social welfare and the reasoning ability. Figure 5 shows that the relation among the APP1 and the APP2 and reasoning ability, where APP1 is the average payoff of player 1 and APP2 is that of player 2. Figure 5 is the result of 10-move centipede game with logit function model. The case around σ =5 in the logit function model, in other words, moderate rational society may maximize the social welfare.

This result implies that the centipede game can be considered as a kind of situation that cooperation is desired. Since cooperative behavior is not always increase their payoffs, Pareto efficiency is not guaranteed. To implement Pareto optimal results with certainly, we need to introduce a certain penalty system.

However, since introduction of such a penalty system inevitably requires social cost, it does not always increase social welfare in the real world. These arguments indicate severe penalty system may not required to implement cooperative state in the real situations. In addition, repetition of cooperative actions may generate a kind of moral so that the players may perceives the centipede game as if it were a game which the cooperative actions are equilibrium strategies.

4. SIMPLE DECISION HEURISTICS

4.1 Introduction of simple decision heuristics

Our main motivation is modelling of imperfect reasoning and player's internal model itself. Real players seldom consider all posiible result and sometimes miscalculate and they know it. It is also important to emphasize that players do not always try to implement equilibrium strategies. In the previous section we examined the case that players try to calculate as much as possible with there reasoning abilities. In this section, we leave from the assumption that players basically choose the best action in their internal models. We examine the properties four types of simple reasoning.

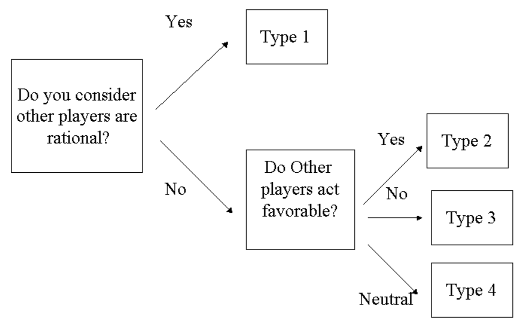

Fig.6: Simple Decision Heuristics

(1) The first type describes the players that they want to behave rational. i.e. player considers that other players also choose rational actions and she maximize her payoff under the assumption in every steps in their internal model. We call it type 1. This heuristics can be considered as simple version of backward induction and same type with the type that examine in the previous section.
(2) The second type describes the players that they want to behave optimistic. i.e. player considers that other players choose best actions for her and she maximize her payoff under the assumption move in every steps in their internal model. We call it type 2. This heuristics can be considered as a kind of maxmax criterion.
(3) The third type describes the players that they want to behave pesimstic. i.e. player considers

that other players choose worst actions for her and she maximize her payoff under the assumption in every steps in their internal model. We call it type 3. This heuristics can be considered as a kind of maxmin criterion.

(4) The forth type describes the players that they want to behave risk neutral. i.e. player considers that other players choose random actions for her and she maximize her payoff under the assumption in every steps in their internal model. We call it type 4. This heuristics can be considered that she plays with not rational player but with nature.

5. EFFECTIVENESS OF EACH HEURISTICS

5.1 Simple decision heuristics in zero-sum game

We then aplly them to some typical games and examine the effectiveness of each heurisics. The first example is zero-sum games since general analysis is quite difficult.

In 1944 von Neumann and Morgenstern proved that any zero-sum game involving *n* players is in fact a generalised form of a zero-sum game for two persons; and that any non-zero-sum game for *n* players can be reduced to a zero-sum game for *n* + 1 players, the (*n* + 1) th player representing the global profit or loss. This means that the zero-sum game for two players forms the essential core of mathematical game theory.

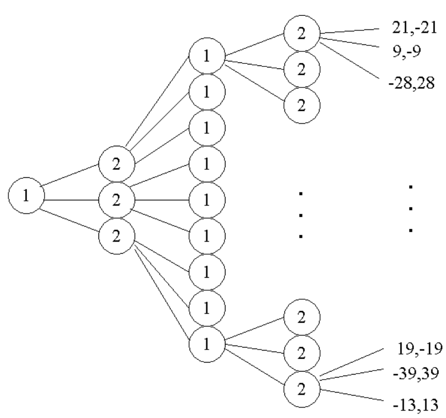

Fig.7: A example of Zero-sum Game

In this paper we examine the case that there are two players and each player has two decision nodes and three alternatives are given in each decision nodes. We should notice that there are 81 outcomes in this game. We assumed that payoffs are assigned from -40 to +40 at 1 unit intervals and uniformaly distributed.

If players' reasoning abilities are perfect, this kind of games, such as chess, reversi, go, etc. , would be quite boring. However these games are still played in today. Then, rational strategies are also effectve against other heuristics?

We made round robin competition and take the average by numerical simulations.

Follower / Leader	Type1	Type2	Type3	Type4
Type1	-4.6, 4.6	2.3, -2.3	-1.0,1.0	2.2, -2.2
Type2	-13.1, 13.1	-12.7, 12.7	-6.4, 6.4	-5.9, 5.9
Type3	-7.4, 7.4	-6.6, 6.6	-3.8, 3.8	-3.1, 3.1
Type4	-8.6,8.6	-4.5, 4.5	-4.6, 4.6	-5.2, 5.2

Fig.8: Results of round robin competition

Let us assume that we call player 1 in the Fig.7 is leader and player 2 in the Fig.7 is follower.

In this game setting, we should notice that follower have advantage over leader. Furthermore type 1's reasoning and type 3's reasoning are equivalent concerning oppnent moves, however they are opposite concerning future moves of theirselves.

We have 2 properties about the competition table. The first is overall properties. Rational heuristics takes the best score and maxmax heuristics takes the worst score, maxmin and average heuristics take moderate scores.

The second is indeed maxmax heuristics is the worst strategy in 4 types, in the case of type 2 follower take better score against type 2 leader than type 3 and type 4 follower.

Easygoing attitude is inferior to the rational attitude if opponent is rational. Thus maxmax heuristics can be considered as quite irrational. However, if opponennt is not rational, this result implies that less rational attitude may superior to more rational attitude. Therefore if we leave from rational paradigm, we need other axis in order to evaluate heuristics.

5.2 Simple decision heuristics in centipede game

The second example is centipede game. Since centipede game has only two alternatives in every decision nodes, we can examine by analytical way in the result of 4 kinds of heuristics. We now assume that the game has enough length. In this game, action of type 1 and type 3 is identical where they always choose T, and that of type 2 and type 4 is also identical where they always choose P. If both players are in former types, the game is over at the first period. If one player is former type and the other is latter type, then game is over at the first decision node of former type player. If both players are latter type, game is over at the last period by all P. As a result they receive large profit. Therefore latter type has larger payoff averagely though former type wins to the latter type in one specific game. This result shows that if there is a discrepancy with individual rationality and social rationality, optimistic or risk adverse decision heuristics may be superior to backward induction.

6. CONCLUSIONS

The main contributions of this paper are as follows: First, we proposed a dynamic mathematical models expressing rough reasoning. Reasoning ability is defined as dependent not only on the payoff but also on the depth of decision tree. Second, new interpretation of our intuition in centipede game was proposed. Third, we pointed out some implications of rough reasoning from two sides, frequency of rational action and social welfare. Finally, we formulate 4 kinds of simple decision heuristics and then apply it to constant-sum game and centipede game. In constant-sum decision situations, we have the result that backward induction is the best and risk averse is second is selfish is the worst. This result can be seen correspondent to the accepted theory of real games. On the other hand, in the centipede game, optimistic and risk averse heuristics get higher average score even if they lose in some specific games. This result would mean that we have risk averse and optimistic preference in some extent if preferences are inheritable.

In this paper, we only discussed cases where each of players is equally rational. It was shown that the increase of agent's rationality is not necessarily connected with the rise of social welfare. It is future task to analyze what strategy is stabilized from an evolutional viewpoint by assuming a social situation is repeated

REFERENCES

[1]. Axel rod, R. Effective Choices in the Prisoner's Dilemma. J. Conflict Resolution, 24, 3-25, 1980.

[2]. Selten, R. The Chain-Store Paradox. Theory and Decisio*n*, 9, 127-159, 1978.

[3]. Rosenthal, R. Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox, Journal of Economic Theory, 25, 92-100, 1981.

[4]. Myerson, R. Refinements of the Nash Equilibrium Concept. Int. Journal of Game Theory, 7, 73-80, 1978.

[5]. Mckelvey, R. Palfrey, T. Quantal Response Equilibria for Normal Form Games. Games and Economic Behavior, 10, 6-38, 1995.

[6]. Mckelvey, R., Palfrey, T. Quantal Response Equilibria in Extensive Form Games. Experimental Economics, 1, 9-41,1998.

[7]. Heifetz, A. Pauzner, A. Backward Induction with Players Who Doubt Others' Faultlessness. Mathematical Social Science, 50, 252-267, 2005.

[8]. Kreps, D., Wilson, R. Sequential Equilibria. Econometrica 50, 863-894, 1982.

[9]. Kreps, D., Reputation and Imperfect Imformation, Journal of economic Theory, 27, 253-79 1982

[10]. Aumann, R. Correlated Equilbrium as an Expression of Beysian Rationality. Econometrica 55 1-18, 1992

[11]. Aumann, R. On the Centipede game: Note. Games and Economic Behavior 23, 97-105, 1998.