Simulation of the Enhanced Version of Prisoner's Dilemma Game

Tatyana Vlasova, Maryna Vladymyrova and Dmitry Shabanov

V.N. Karazin Kharkiv National University, 4, Svobody Sqr., Kharkiv, 61077, Ukraine
vlasova.tatyana@bk.ru, vladymyrova@gmail.com, d.a.shabanov@gmail.com

Abstract. This paper presents the model and software to explore pair interactions of objects with different behavior and their patterns. The research is based on the enhanced version of a classic prisoner's dilemma game. The non-cooperative finite and infinite pair games with non-zero sums are investigated. Pure and mixed strategies with finite and infinite memory developed by Biology School of V. N. Karazin Kharkiv National University are used to analyze the results.

1 Introduction

This paper presents the model and software to explore pair interactions of biological objects with specified behavior. The research was carried out on request of the herpetology department of V. N. Karazin Kharkiv National University and its results are used in educational process to illustrate some topics of the ecology discipline as well as in biology students' research for studying patterns of pair behavior of some biological species. Our model is based on the classic prisoner's dilemma game [1,2,3] as it is widely accepted as the model to study the pair interactions and behavior of different agents from interactions of animals in nature up to economical transactions in human world [4, 5, 6, 7, 8]. The main goal of the game in its classic version is to get maximal score doing preset number of steps with given values of the fine, the cooperation bonus and the cooperation award. Each participant can remember not more than two its own and the opponent's previous actions.

In distinction to classic case we enhance the rules of the game by allowing more complex behavior of agents depending on their ability to remember their own or opponent's previous steps and on values of fines, bonuses and awards.

The important factor which influences the evolution of living organisms including humans is their communications with environment. It is only typical for such communications to have conflicts of interest between opponents, absence of information about future actions of an opponent and need to foresee its future actions only on the base of the prehistory of similar interactions. The paradox of the game clearly shows the contradiction between group interests and individual ones: what is optimal for the group of two is not good for each member of the group. In fact the same is true for multilevel systems with optimization functions on different levels: their behavior is intuitively unpredictable and even paradoxical.

It is absolutely obvious that the outcome of the game fully depends on the participants' strategies. Here we define strategy in a slightly different way than it is done in the game theory.

Since the objective of this research is to explore the pair interactions in real biological environment it is natural to define strategy as a set of rules used by a participant to make its next step. Because of the infinite variety of strategies for different biological objects we use simulation with strategies constructed by the experimenter. The developed software allows experimenter to set different strategies as input information for simulation. The main goal of experimenter is to find optimal strategy and the value of the game in their classical sense.

2 Model and simulation description

We consider the pair non-cooperative finite and infinite games with a non-zero sum [11; 12; 14] so the general result can be both positive (in the case when participants cooperate during the whole game) and negative. We take the game in its normal form $\Gamma_N = < N; X_1,...,X_i,...,X_n; K_1,...,K_i,...,K_n >$, where Γ_N is the notion of the game, $N = \{1,...,i,...,n\}$ is a set of participants, $X_i = \{x_i\}$ is a set of strategies for the *i*-th participant and $K_i(x_1,...,x_i,...,x_n)$ is the gain function for the *i*-th participant, the value of which is the gain obtained by the *i*-th participant if participants use strategies $(x_1,...,x_i,...,x_n)$. In our model we allow participants to use pure or mixed strategies. Pure strategy is just definite sequence of steps i. e. it can be represented by any element $x_i \in X_i$. Mixed strategy can be a simple set of pure strategies or a set of pure strategies with given probabilities distribution.

Any strategy as a set of rules depends on the participant's memory depth or its type of memory. Here we suppose that each participant remembers its previous actions or the previous actions of the opponent. If the participant has zero memory depth it can make predefined actions during the whole game or make spontaneous actions. If the participant has finite memory depth it uses pure strategy depending on the actions of its opponent. If the participant has infinite memory depth only the absolute value of its current gain influences its actions.

The pure strategy is a response of the participant to the actions of its opponent. We consider the following cases:

No response (leads to spontaneous actions)

The response to the definite set of actions

The response to the small value of the own gain or to the big value of the opponent's gain

The response to the big value of the own losses or to the small value of the opponent's losses

In this case we achieve the goal of maximization of the participant's own gain.

Similarly when the participant remembers only its opponent's history of actions and responds to them we achieve the goal of minimization of the opponent's gain.

We use finite automation with two states which are "accept" and "reject" to model the game [8, 12]. The input information for the automation is as follows: participants' strategies, values of the gain function, transition rules for a participant's behavior, and initial conditions. Transition rules depend on the type of a participant's strategy which can be pure or mixed. Simulation repeats the preset number of times or till the winning of one of its participants.

There exist four scenarios of simulation each corresponding to four different goals.

First scenario is the experiment between the two chosen strategies ("pure-pure", "mixed-pure", "mixed-mixed"), when the values of gain/bonus/fine (M, L, and K, respectively) are fixed. The simulation results are presented by the graph, where x-axis denotes the sequence of steps and y-axis denotes the quantitative characteristics of the gain/loss (Fig.1).

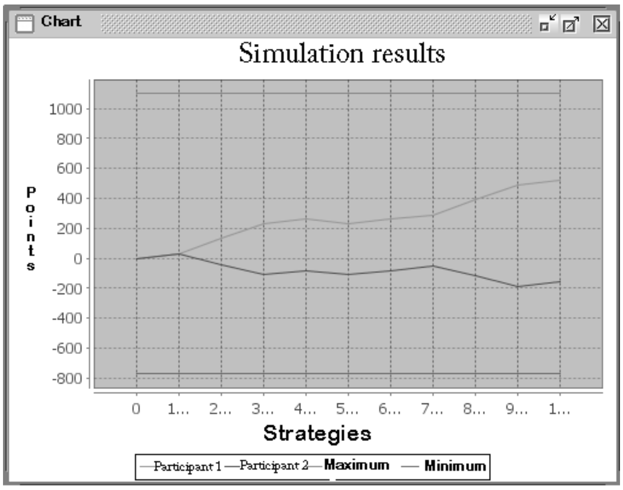

Fig. 1. Simulation results with two strategies (one is random).

The result data of this scenario show that among the strategies with zero and finite memory, the strategy of random steps and zero memory wins in most cases. Among the strategies with finite and infinite memory the finite memory strategy wins in most cases. Analogous conclusions were made previously by R. Dockins [6] when he analyzed the classical version of the game.

Analyzing the obtained results of pair interaction between the pure and mixed strategies we may conclude that the mixed strategy is more advantageous than any of pure strategies. This can be explained by the flexible behavior of the participant with the mixed strategy.

The second scenario is the evaluation of competition between the one fixed strategy and the variety of others. The results of the experiment are presented by the bar chart displaying the number of win points over each of the chosen strategies.

Besides the graphical visualization one can look through the steps history (absolute values of gains or losses) of each of the simulation participants.

The third scenario allows experimenter to approve or disprove the hypothesis that the strategy of a participant depends on the gain function. In order to do it we implemented the feature which allows experimenter to conduct the simulation with fluctuating parameters M, L, K. In this case the results are presented by the graph and the table showing the data of each participant's results for chosen strategies with respect to varying values of K, L, and M (Fig.2).

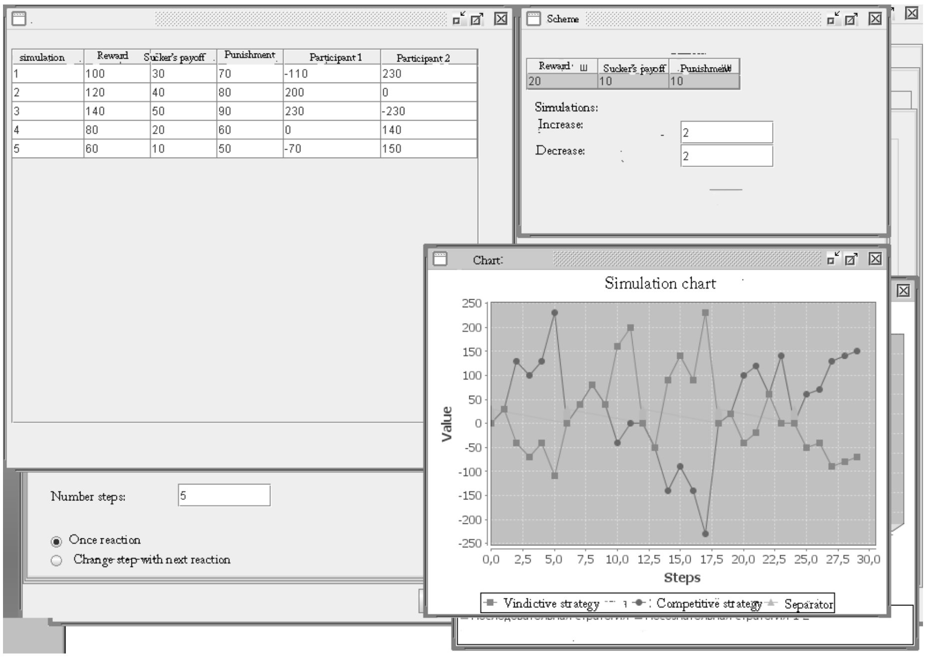

Fig. 2. Simulation results with fluctuating values of K, L, and M.

The results of the experiments show that the outcome of the game depends on the values of the gain function. So one can find experimentally the values of the gain function for any given strategy to be optimal.

Fourth scenario deals with a strategy as an element of the given set. We do not use the formal methods of working with such sets but accepted the way used by experimenters (biologists) to form them. In this case our goal was to find the probability for a strategy from one set to win a strategy from another one. For example our experimenters divided strategies into three sets namely provocative, forgiving and neither provocative nor forgiving exactly in the same way as it is done in R. Axelrod's experiment [4]. In our case a provocative strategy means immediate change of behavior in condition of the participant's own loss (or the opponent's gain) and keeping it till the next loss or till the end of the game. A forgiving strategy means that the participant changes its behavior under the same conditions but keeps it only some limited time (definite number of steps). In some way one can see it as follows: in forgiving strategy the participant only fights back as a response to the smack while in provocative strategy the participant not only fights back as a response to the smack but retaliates. Such division is just conditional as it reflects the view of experimenters.

The results of experiments show that the provocative strategies win in more cases than forgiving ones.

3 Conclusion

The software for simulation the enhanced version of the prisoner's dilemma game to set up experiments and explore pair interactions of objects with different behavior has been developed. The software allows experimenter to set pure and mixed strategies with finite and infinite memory. The simulation depends on its goal: whether it is maximization of the participant's gain or minimization of the opponent's loss. The participant's gain depends on the values of the gain function. If the participants' strategies are fixed then one can find experimentally such values of the gain / bonus / fine that the strategy of one of the participants is optimal. The results of the simulation comply with those given in literature in the case of classic game which can be accepted as an adequacy of the model.

The software was tested and operates at Biology School of V. N. Karazin Kharkiv National University but it can be successfully used at other schools and fields such as economy, sociology, etc.

References

1. Merrill, M. Flood: Some Experimental Games. Research Memorandum RM 789
2. Tucker A. W. On Jargon: The Prisoner's Dilemma. UMAP Journal 1 (1950): 101
3. Rheingold, H.: The Virtual Community – Perennial (2006)
4. Axelrod, R.: The evolution of cooperation. New York: Basic Books (1984)
5. Axelrod, Robert and Hamilton, William D.: The Evolution of Cooperation. Science, 211 (1981) 1390-1396
6. Dawkins, R.: The selfish gene (1993)
7. The Iterated Prisoner's Dilemma Competition [Electronic resource]. – Mode of access: http://www.prisoners-dilemma.com/
8. Adaptive modeler [Electronic resource]. – Mode of access: http://www.softsoft.ru/business/investment-tools/3210.htm
9. Vorobjev, N.N.: Game theory. (1976)
10. Grabovski, B.I.: Cellular automata, as simple models of complex systems. (1995) 412–419.
11. Karlin, S.: Mathematical methods in the game theory, programming and economy (1964).
12. Tseitlin, M.: Research on the theory of state machines and modeling of biological systems. (1969)