COLLECTIVE BEHAVIOR IN SPATIAL PRISONER'S DILEMMA GAME

[1]FRANCISZEK SEREDYNSKI, [2]JAKUB GĄSIOR

Cardinal Stefan Wyszynski University in Warsaw, Department of Mathematics and Natural Sciences, Woycickiego 1/3, 01-938 Warsaw, Poland
E-mail: [1]f.seredynski@uksw.edu.pl, [2]j.gasior@uksw.edu.pl

Abstract - In this paper we study conditions of emerging collective behavior of agents acting in the 2-dimensional (2D) cellular automata (CA) space, where each agent plays locally the Prisoner's Dilemma (PD) game. The system is modeled by 2D CA evolving in discrete moments of time, where each cell-agent changes its state according to a currently assigned to it rule. We assume existing of three rules: always cooperate (all-C), always defect (all-D), and k-D – defect only if more than k neighbors defect. Rules are initially assigned randomly to cells-agents, but during iterated game agents may replace their current rules by rules used by their neighbors. While each agent is oriented on maximization of its own profit in the game, we are interested in answering the question if and when a phenomenon of a global cooperation in a large set of agents is possible. We present results of experimental study showing conditions and degree of emerging such a cooperation.

Index Terms - Collective behavior, multi-agent systems, spatial prisoner's dilemma game, second order cellular automata.

I. INTRODUCTION

A phenomenon of emerging cooperation between members of human teams is a subject of study for a number of years. Understanding mechanisms and conditions of emerging cooperation is very important for different areas of science, such as political science, economy, sociology and today also for computer science, especially in parallel and distributed computing. Prisoner's Dilemma (PD) game [7] is one of the most accepted game-theoretical models where both cooperation (C) and defecting (D) of humans can be observed and studied. A. Tucker was a scientist who formalized the game as the 2-person game originally framed by M. Flood and M. Dresher around 1950. In the beginning of 1980's R. Axelrod organized the first tournament [1] to recognize competitive strategies in iterated PD (IPD), more exactly in the 2-person IPD game and the winner was a strategy Tit for Tat (TFT) submitted by A. Rapopport. TFT assumes cooperation of a player on the first move and subsequent repeating actions of the opponent player used in the previous move. Because genetic algorithms (GAs) became at that time a popular search technique in solving optimization problems, Axelrod proposed [2] to apply GAs to discover players strategies enabling cooperation in the 2-person PD game. GAs were able to discover TFT strategy and other specific for humans interesting strategies. While discovering by GAs cooperating strategies in the 2-person PD game was relatively easy task, discovering strategies enabling cooperation in N-person PD games (N>2) is not easy problem. The approach proposed by Axelrod to code an individual of GA was not effective in this case, therefore X. Yao and P. J. Darwen proposed in [9] another approach where GAs are still applied but the payoff function of the game was simplified, what enabled to reduce the size of an individual of GA. The main idea of the simplified payoff function was that a payoff of a given player in the game depends on a number of cooperating players among remaining N-1 participants of the game. Under these assumptions GA was able to find strategies of players enabling a global cooperation for a number of players equal to around 10. For more players such strategies were not discovered by GA. One of the main reasons of that is the form of the payoff function which in fact assumes participation of a player in a game with a "crowd" – a large number of anonymous persons without any structural relations between them. It is rather difficult to imagine a reasonable behavior in such a crowd. A concept of spatial games in 2D space with a neighbor relations between players is a solution of the crowd problem. Among the first concepts related to spatial IPD was a game on the ring considered by M. L. Tsetlin [8] in the context of games of learning automata, where a payoff of a given player depends on its action and actions of two immediate neighbors. A number of spatial IPD games on 2D grids has been studied recently. M. A. Nowak and R. M. May proposed [6] an original spatial IPD game on 2D grid where two types of players exist – players who always cooperate (all-C) and players who always defect (all-D). Players occupy cells of 2D space and each of them plays PD game with all immediate neighbors and depending on the total score he is replaced (or not) by the best performing player in the neighborhood. They showed that both types of players persist indefinitely in chaotically changing local structures. H. Ishibuschi and N. Namikawa studied in [4] a variant of spatial IPD game with two neighborhoods: one for playing locally defined PD game with randomly selected neighbors, and the second one for

mating strategies of the game by a locally defined GA. E. Howley and C. O'Riordan considered [3] N-person IPD game with all-to-all interaction between players and existence of a tagging mechanisms structure in subgroups of players. Y. Katsumata and Y. Ishida in [5] extended the model of spatial IPD games proposed by [6] by considering 2D space as the 2D cellular automata (CA) and introducing an additional strategy called k-D, which tolerates maximally D defects in neighborhood. They showed emergence of specific spatial structures called membranes created by players using k-D strategies, which separate cooperating and defeating players. In this paper we consider the model proposed in [5] but we focus on the study of conditions enabling emerging collective behavior in large teams of players. The structure of the paper is following. In the next Section some background concerning CA is given. Section 3 contains details of the model of studied spatial IPD game. Last section, Section 4 contains results of experimental study of the game.

II. CELLULAR AUTOMATA

A. One-dimensional CA

CA are spatially and temporally discrete computational systems originally proposed by S. Ulam and J. von Neumann in the late 1940s to study processes of self-reproduction of machines. Today CA are powerful tools used in computer science, mathematics and natural science to model different phenomena and develop parallel and distributed algorithms [10]. One-dimensional (1D) CA is in the simplest case a collection of two-state elementary cells arranged in a lattice of the length n, which locally interact and evolve in a discrete time t (see, Fig. 1).

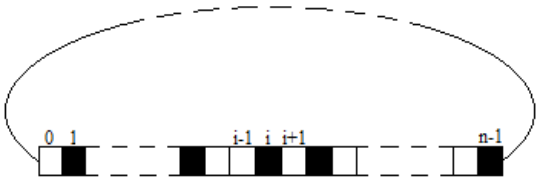


Fig. 1. Example of 1D CA: each cell is either in state 0 (white) or 1 (black).

For each cell i (0<= i <= n-1) called a central cell, a neighborhood of a radius r is defined. The neighborhood consists of $n_i = 2r+1$ cells (see, Fig. 2), including the cell i. (sometimes a central cell is not included). A cyclic boundary condition is applied to a finite size of CA, which results is in a circle grid.

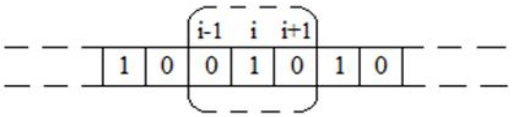


Fig. 2. Example of neighborhood of radius r=1 for 1D CA.

It is assumed that a state q_i^{t+1} of the i-th cell at the time t + 1 depends only on states of its neighborhood at the time t, i.e.:

$$q_i^{t+1} = f(q_{i-r}^t, \ldots, q_{i-1}^t, q_i^t, q_{i+1}^t, \ldots, q_{i+r}^t),$$

where f is a transition function called also a rule, which defines the way of updating the cell i. If CA update states of their cells according to a single rule assigned to all cells, such CA are called uniform CA. If two or more different rules are assigned to update cells, such CA are called non-uniform CA.

B. Two-dimensional CA

Two dimensional (2D) CA is 2D lattice consisting of n*n elementary cells. For each cell a local 2D neighborhood of a radius r is defined (see, Fig. 2). At a given discrete moment of time t each cell (i,j) is in some state $q_{i,j}^t$, in the simplest case it is a binary number 0 or 1.

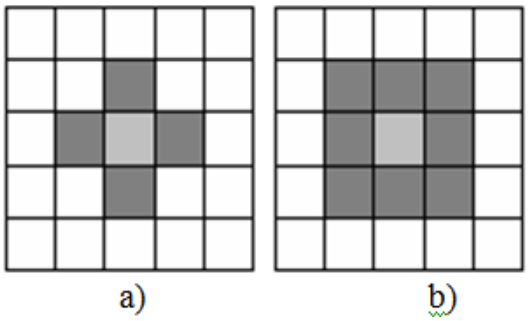


Fig. 2 Examples of a local neighborhood of the size r=1 of a central cell (in light grey) in 2D CA of the size 5x5: a) von Neumann neighborhood consisting of 4 neighbors, b) Moore neighborhood consisting of 8 neighbors .

At discrete moments of time all cells synchronously update their states according to assigned to them local rules (transition functions), which depend on states of their neighborhood.

A cyclic boundary condition is also applied to a finite size of 2D CA. Similarly like in 1D CA, 2D CA can be uniform or non-uniform. In classical CA rules assigned to cells are not changed during evolving CA cells.

III. SPATIAL PRISONER'S DILEMMA GAME

We consider 2D CA of the size n*n. Each cell of CA has a Moore neighborhood of radius r, and also has a rule assigned to it which depends on the state of its neighborhood.

Each cell of 2D CA will be considered as an agent – player participating in the iterated spatial Prisoner's Dilemma (ISPD) game [5]. Each player (a cell of CA) has two possible actions: C (cooperate) and D (defect). It means that at a given moment of time each cell is either in the state C or the state D. Payoff

function of the game is given in Tab. 1.

Tab. 1 Payoff function of a player participating in ISPD

Player	Opponent player		
		C	D
	C	R=1	S=0
	D	T=b	P=0

Each player playing a game with an opponent player in a single round (iteration) receives a payoff equal to R, T, S or P, where T > R > P > S. Values of these payoffs used in this study are specified in Tab. 1, where P=S=0, R=1, and T=b (1.1<= b <= 1.8). If a player selects action C and the opponent a player from the neighborhood also selects action C than the player receives payoff equal to 1. However, if a player has action D and the opponent player still has action C, the defecting player receives payoff equal to b. In two remaining cases a player receives a payoff equal to 0. In fact, each player associated with a given cell plays in a single round a game with each of his 8 neighbors and this way collects some total score. After q number of rounds (iterations of CA) each cell (agent) of CA can change its rule (strategy). We assume that considered 2D CA is non-uniform CA, and to each cell one of the following rules can be assigned: all-C (always cooperate), all-D (always defect), and k-D (cooperate until not more than k (0 <= k <= 7) neighbors defect). The strategy k-D is a generalized strategy TFT, and when k=0 it is exactly the strategy TFT. A player changes its current strategy into another one comparing collected during q rounds total score with scores collected by his neighbors. He selects as his new strategy the strategy of the best performing neighbor, i.e. the player whose the collected total score is the highest. This new strategy is used by a cell (player) to change its current state, and the value of the state is used in games during the next q rounds.

It is worth to notice that considered 2D CA differs from a classical CA, where rules assigned to cells do not change during evolving CA in time. The considered CA with a possibility of changing rules is called the second order CA. In opposite to classical CA the second order CA has potentials to solve optimization problems. The main research issue in [5] was to study conditions of appearing of specific structures called membranes created by cells using k-D rules which protected cooperated players from defected players. In this study we are interested in the collective behavior of large team of players, in particular in conditions of emerging a global cooperation in such teams.

IV. EXPERIMENTAL RESULTS

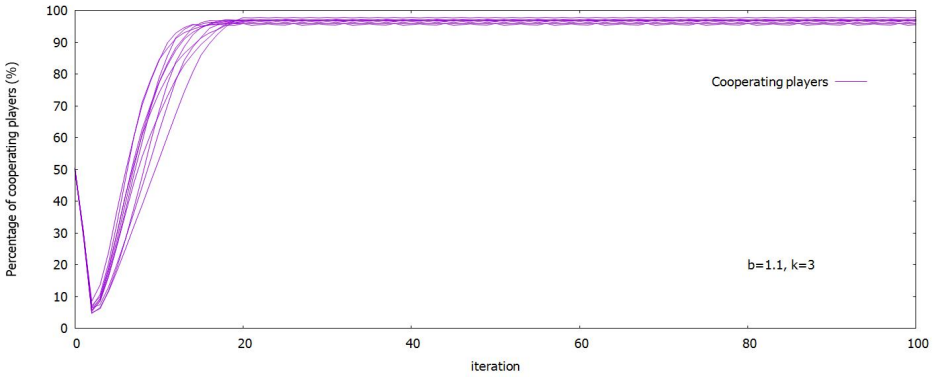
A. Experiment parameters setting

A number of simulation experiments have been conducted with the presented model of ISPD game to find out conditions of emerging a global cooperation in large teams of players. The following settings were used in the experiments:

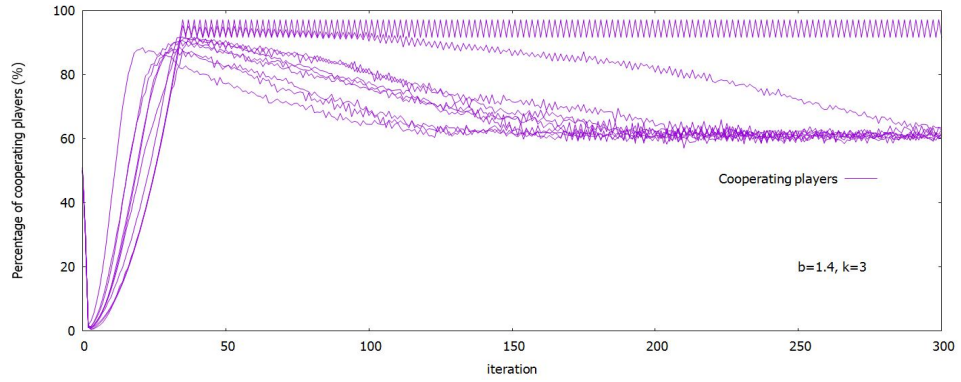
- 2D CA with a size of 100x100 was used, i.e. with a total number of players equal to 10000
- An initial state (C or D) of each cell was set with probability 0.5
- To each cell one of 3 rules (all-C, all-D, k-D) was assigned with probability 0.33; for a given run a constant value of k from the range [0..7] was set, but for a cell to which rule k-D was assigned, the value of k was an integer number randomly generated from the range [0..k]
- A neighborhood of each cell was set to r=1
- For a given run the value of b was from the range [1.1,…,1.8]
- Updating of rules assigned to cells was conducted after q=1 rounds (iterations)
- During a single run of an experiment, depending on values of b, a number of rounds ranged from 100 to 500
- If averaging of results was applied it was done on the base of 30 single runs.

B. Experiment #1: Sample runs of ISPD game under constant value of parameter k=3.

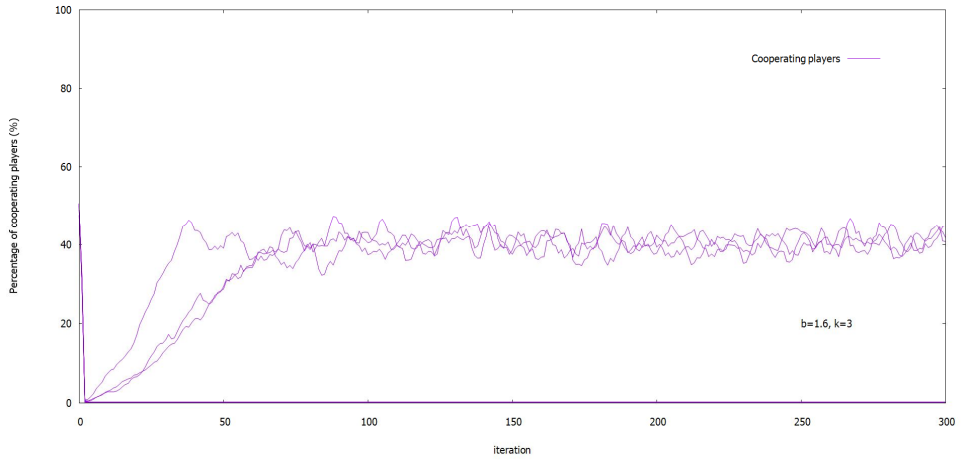
The purpose of the first set of experiments was to observe dynamics of the game for a constant value of k and different values of parameter b. For a given value of b and k 10 simple runs are shown in Fig. 3.



a)

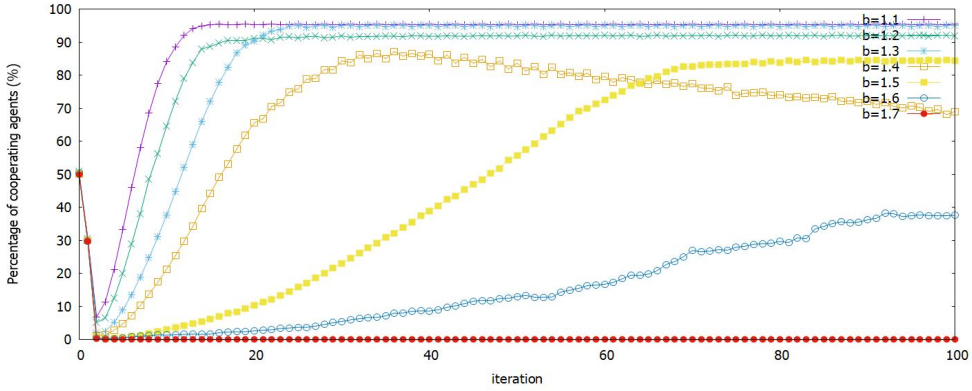


b)

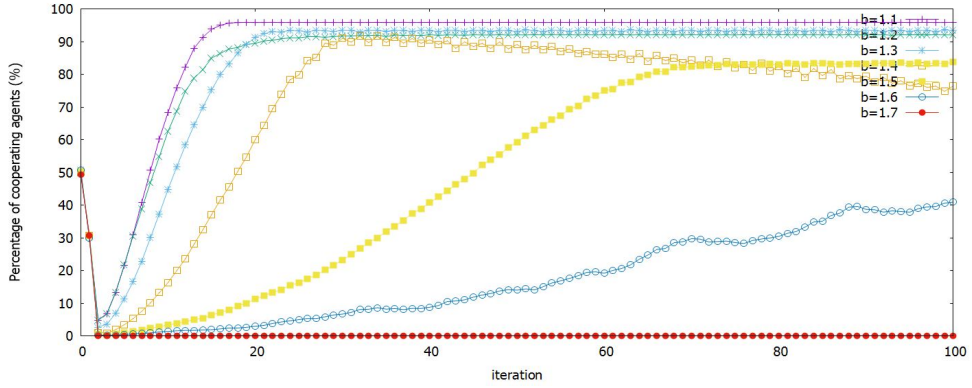


c)

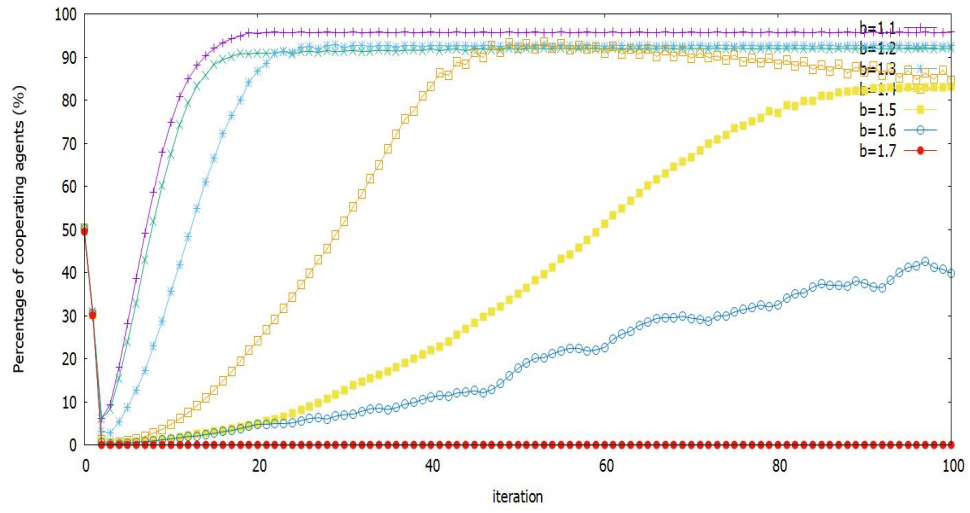
Fig. 3 Ten sample runs with k=3 and: a) b=1.1, b) b=1.4, and b=1.6.



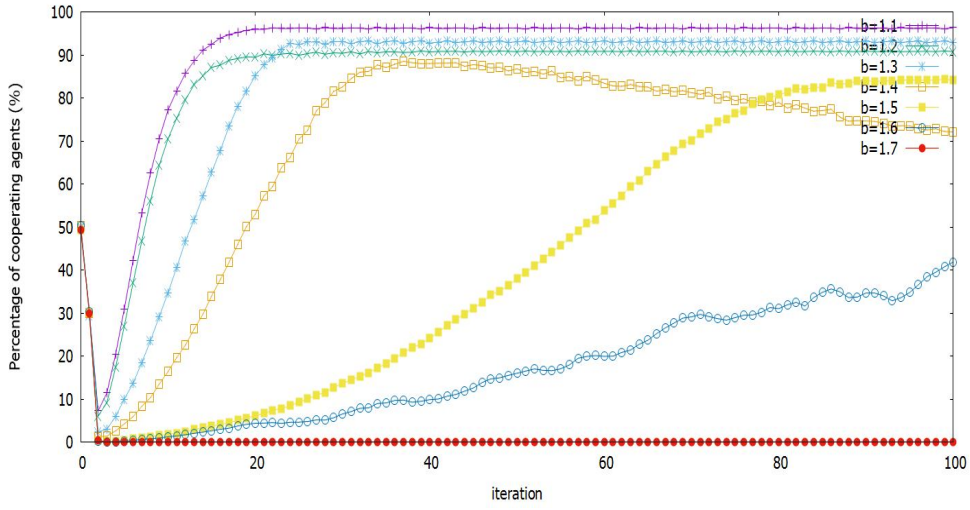
a)



b)



c)



d)

Fig. 4 Sample runs for a range of values b, and: a) TFT (k=0) strategy, b) k=3, c) k=5, and d) k=7.

One can see that for b=1.1 and k=3 (see, Fig. 3a) at each of 10 runs around 95% of players achieves a stable cooperation after about 20 iterations of the game. For the same value of k and b=1.4 (see, Fig. 3b) only in one run over 10 runs the team of players achieves a level of cooperation similar to the previous experiment.

A global cooperation on the level around of 60% is achieved after around 150 iterations in 8 out of 10 runs. When the value b=1.6 (see, Fig. 3c) only in 3 out of 10 runs the team of players achieves a global cooperation on the level of around 40%. In remaining runs no cooperation at all is observed.

C. Experiment #2: Sample runs of ISPD game under different values of parameters b and k.

The purpose of the second set of experiments was to observe a dynamics of the game for different values of parameter b and different values of k. Fig. 4 shows single runs of the game played during 100 iterations for a range [1.1, …,1.7] of the parameter b, and different values of k. One can see that behavior of the team of players does not depend too much on the value of k used in the strategy of k-D, but sufficiently depends on the parameter of b.

Indeed, players team achieves the level of near 95% cooperation in the range of values of b between 1.1 and 1.3 when strategy TFT is used (see, Fig. 4a), and also when strategies 3-D, (see, Fig. 4b), 5-D (see, Fig. 4c) and 7-D (see, Fig. 4c) are used. The players team is not able to cooperate when above strategies are used for values of b >= 1.7. When the value of b is in the range [1.4,..,1.5] a high level of cooperation is possible, but when b=1.6 cooperation grows very slowly.

D. Experiment #3: Cooperation as a function of parameters b and k

To find out relations between the phenomenon of a global cooperation and parameters b and k observed in the previous experiments a new set of experiments was conducted. For each pair of values b and k 30 runs

of the game was performed. Each game lasted 500 iterations, and

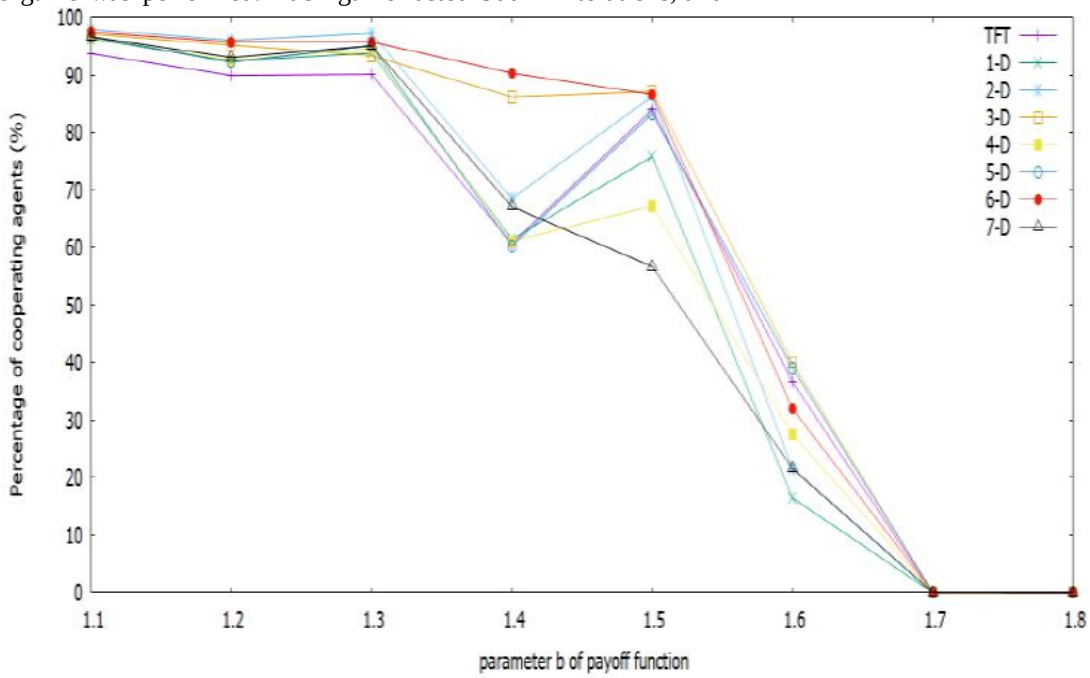


Fig. 5 Percentage of cooperating agents as a function of parameters b and k.

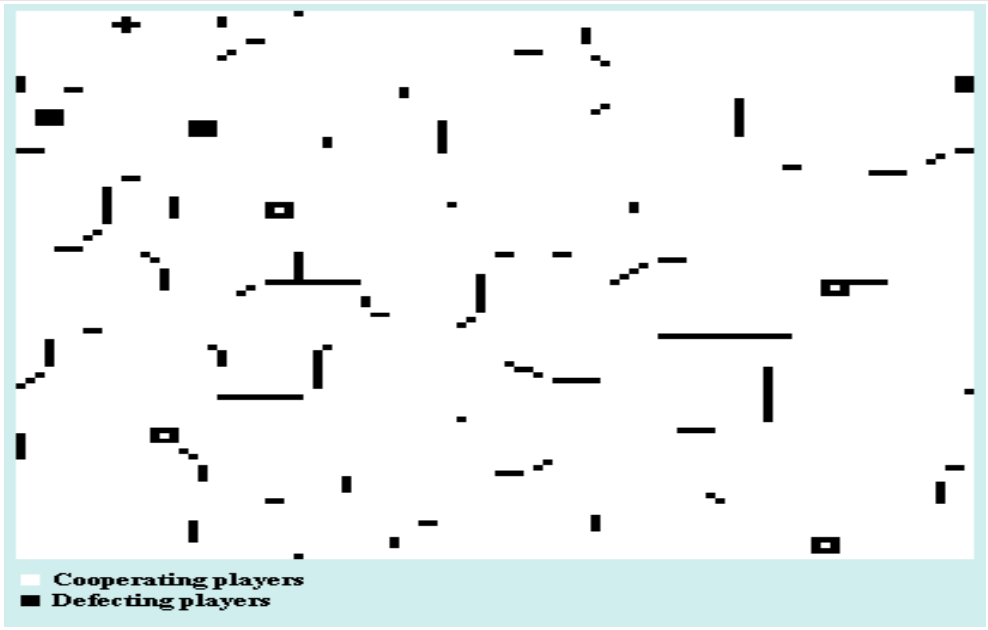
averaged results are presented in Fig. 5. These results generalize observations from previous experiments.
One can see that for b from the range [1.1,…,1.3] players team is able to achieve a cooperation of very high level (around 95% of cooperating players) for all strategies k-D.

Small differences are visible: TFT strategy is slightly less effective than other strategies, while the best performing strategies are 6-D and next 3-D. When b=1.4 the ability of emerging cooperation becomes much more dependent on the level k of tolerance of defecting. Strategies 6-D and 3-D ensure achieving by players team a level of cooperation around 90%, while remaining strategies provide much lower level of cooperation close to around 65%.
When b is decreased to the value 1.5 the ability of achieving a global cooperation is also decreased. Strategies 6-D and 3-D are still the most performing and provide a level of cooperation close to 90%. The remaining strategies provide a level of cooperation close to around 60%.
When b=1.6 possible cooperation level is in the range [15-40]%. When b reaches the value 1.7 or higher value, cooperation in the system will not emerge.

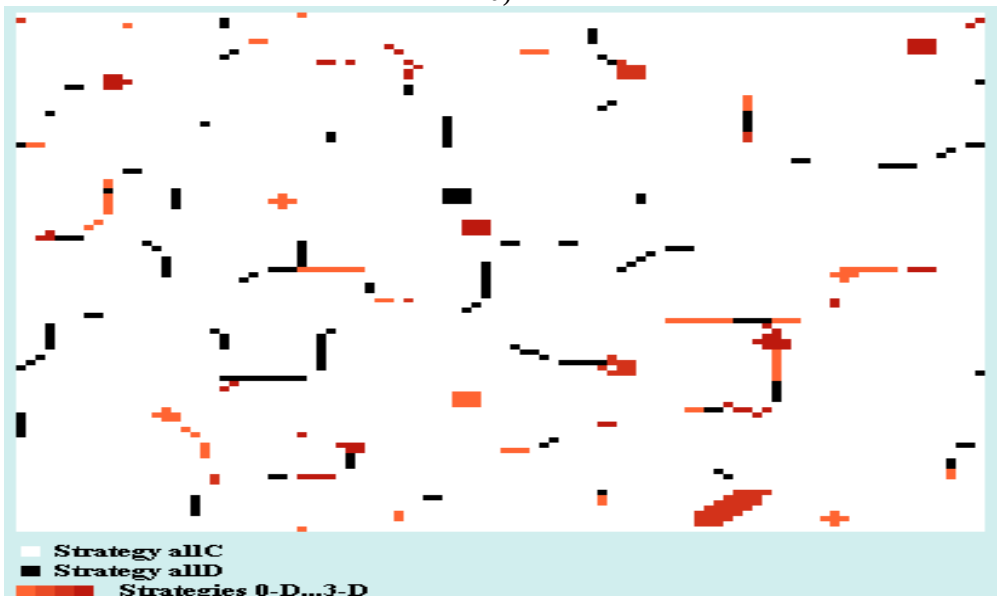
E. Experiment #5: Patterns of agents' actions and strategies.

For presented in Fig. 5 dependencies between the level of a global cooperation and parameters b and k accompanies spatial patterns of actions taken by players and patterns of strategies used by them. These patterns are shown in Figs. 6-12. Because the main factor influencing these dependencies is the parameter b, presented patterns are shown for a constant value of k=3.

Fig. 6 shows corresponding patterns of players' actions and strategies used by them for a game with parameter b=1.1. One can see that dominating action used by players (see, Fig. 6a) is action C (white color), and rarely action D (black color) is used. Actions of players which defect create short, isolated, sometimes vertical or horizontal regular structures. Dominating strategy used by players (see, Fig. 6b) is strategy all-C (white color), and some players use strategies all-D and k-D.
Spatial structure of strategies assignment is similar to actions structure. For games with values b=1.2 and b=1.3 (see, Figs. 7-8) spatial structures of actions and strategies are similar like for the game with b=1.1, but we can see an increasing length of regular vertical and horizontal structures, which are still isolated for b=1.2, but mostly connected for b=1.3. We can see also an increasing number of strategies k-D which replace strategy D. For b=1.4 (see, Fig. 9) spatial structures of both actions and strategies are created by connected rectangles of different sizes. For the actions pattern (see, Fig. 9a) rectangles are created by increasing number of action D, and for the strategies pattern (see, Fig. 9b) by similar number of strategies all-D and strategies k-D. When b=1.5 (see, Fig. 10) both actions and strategies spatial structures are created by irregular isolated chains, where in the strategies pattern more than half of strategies all-D was replaced by strategies k-D.

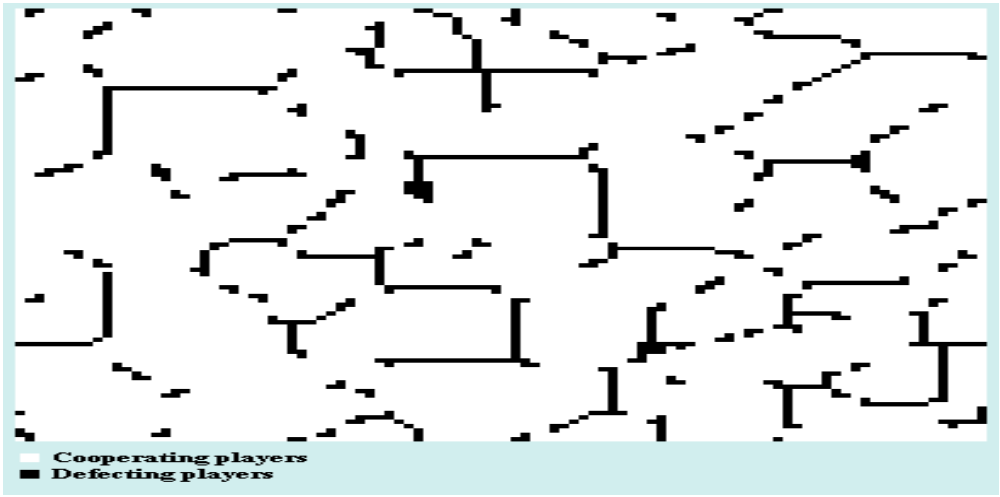


a)

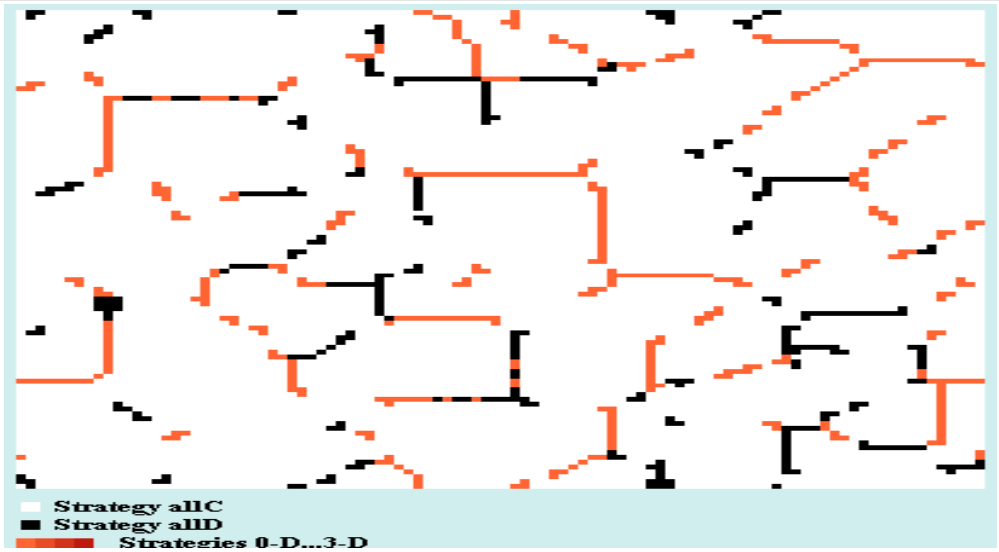


b)

Fig. 6 Agents' actions (a) and strategies (b), run with b=1.1, 3-D

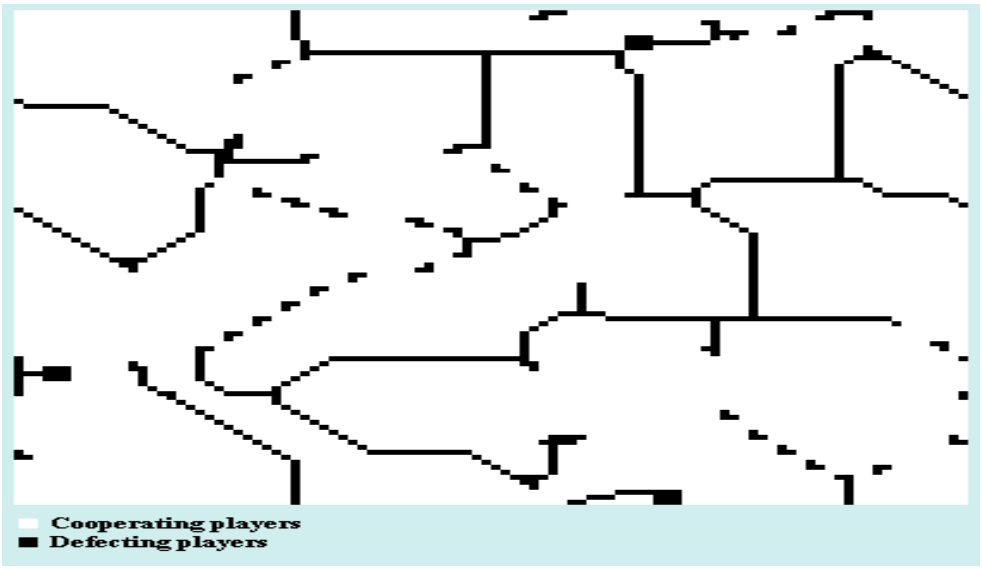


a)

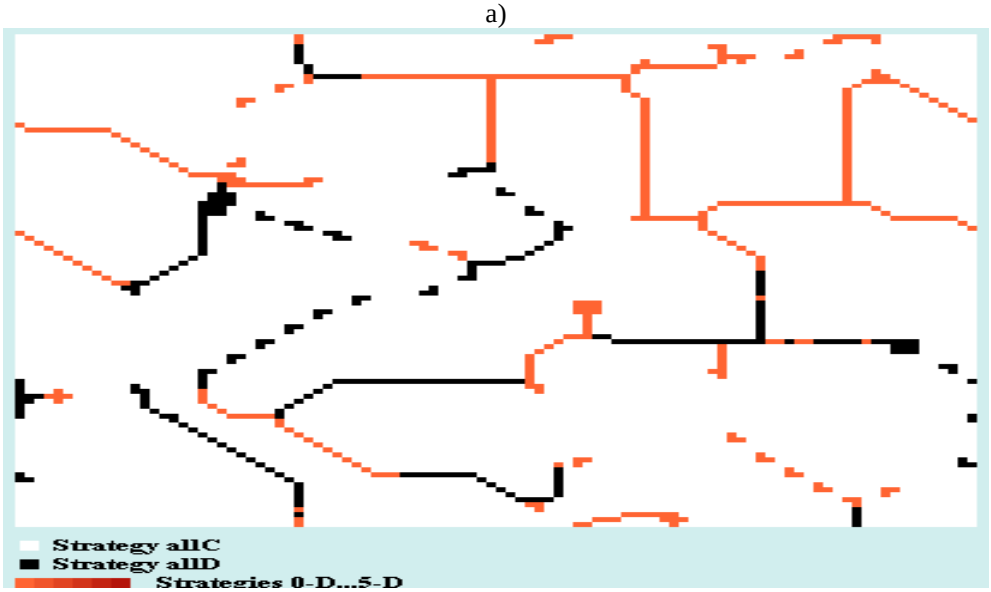


b)

Fig. 7 Agents' actions (a) and strategies (b), run with b=1.2, 3-D

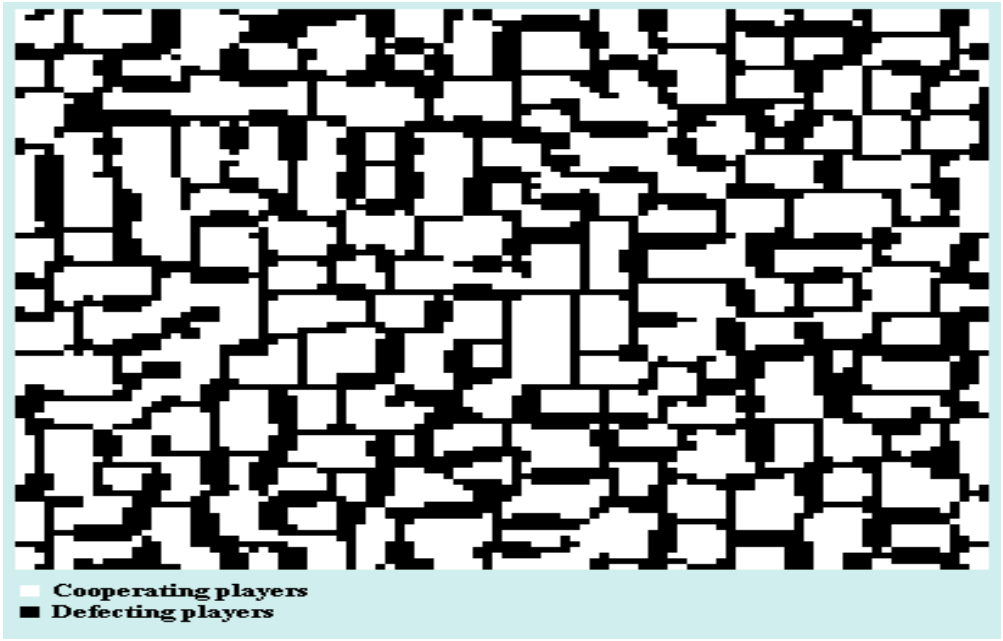


a)

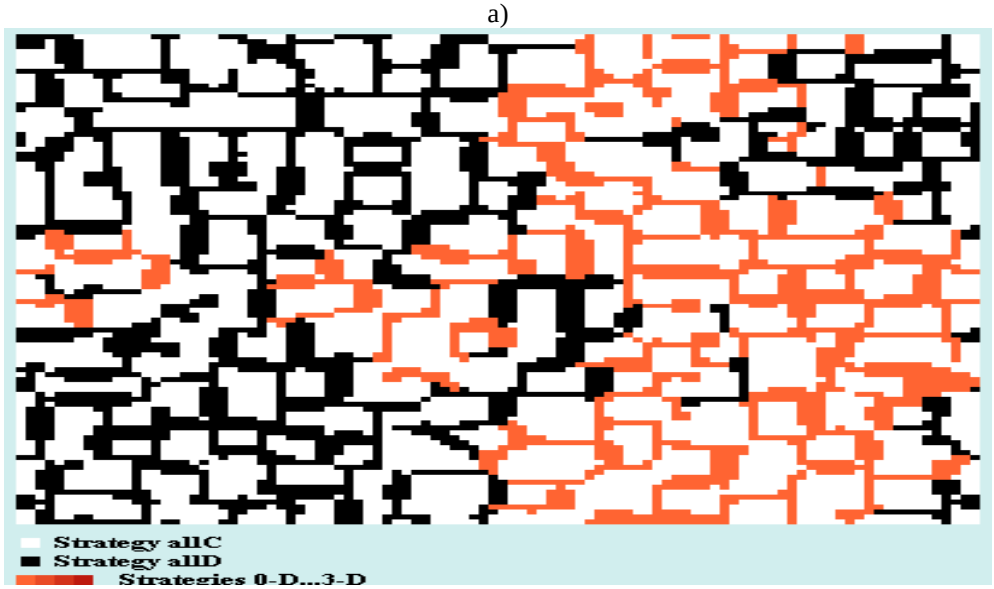


b)

Fig. 8 Agents' actions (a) and strategies (b), run with b=1.3, 3-D

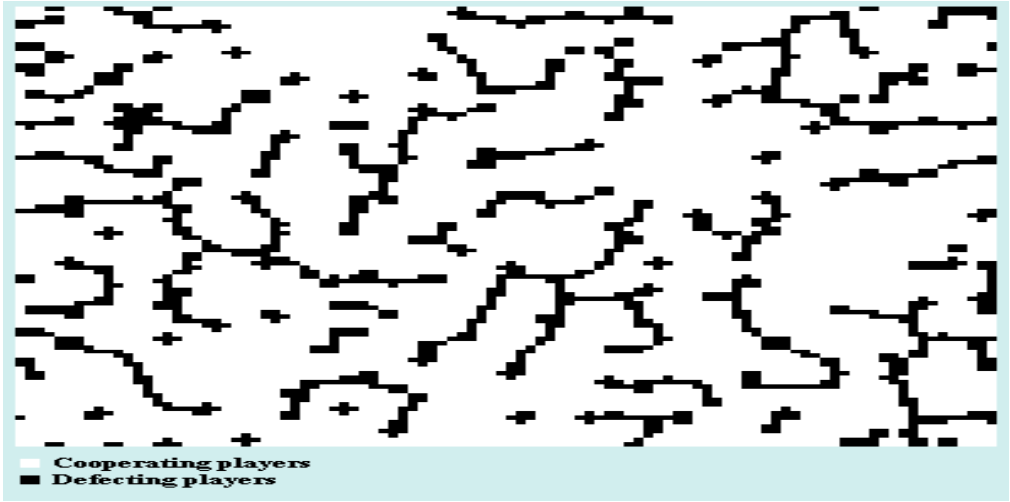


a)

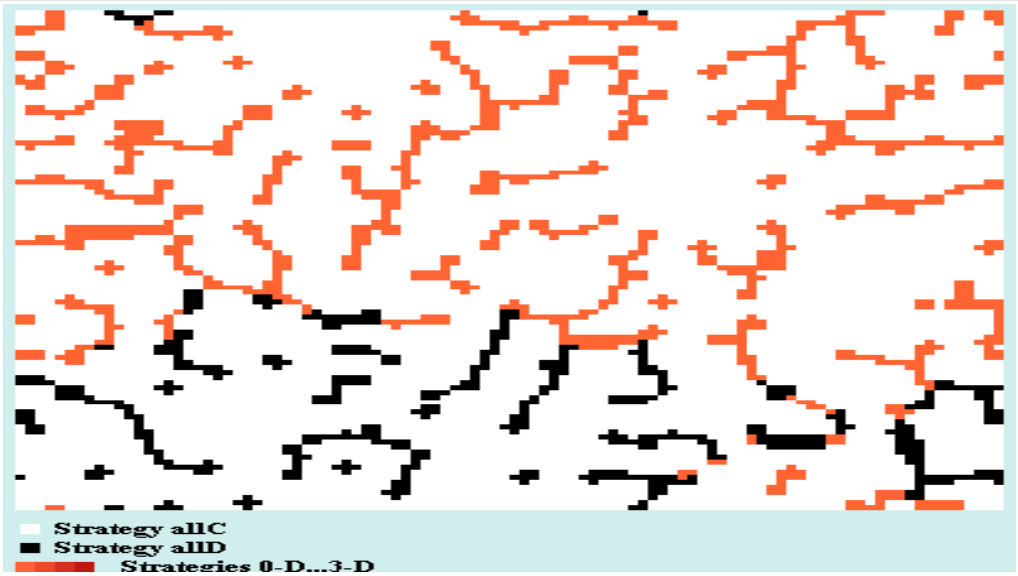


b)

Fig. 9 Agents' actions (a) and strategies (b), run with b=1.4, 3-D

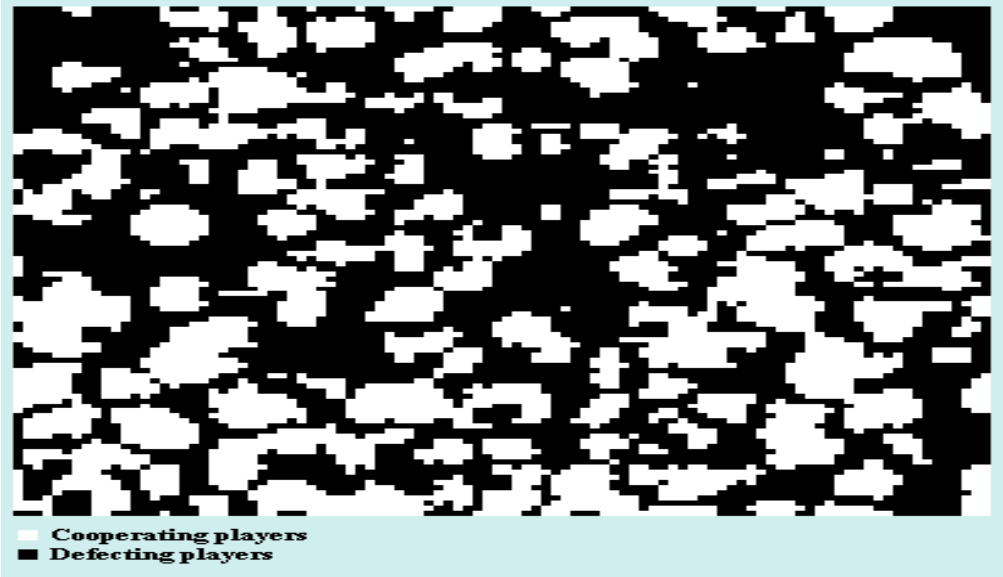


a)

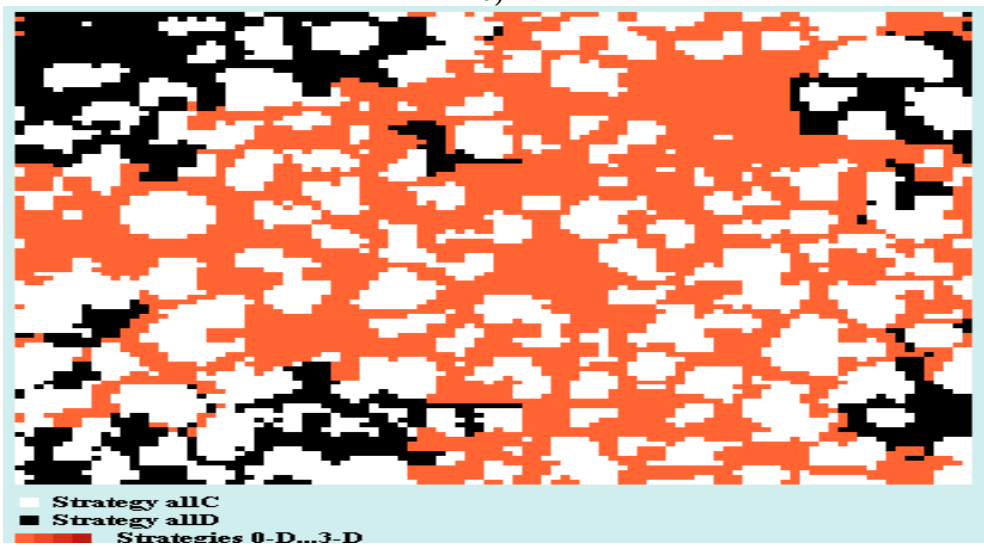


b)

Fig. 10 Agents' actions (a) and strategies (b), run with b=1.5, 3-D

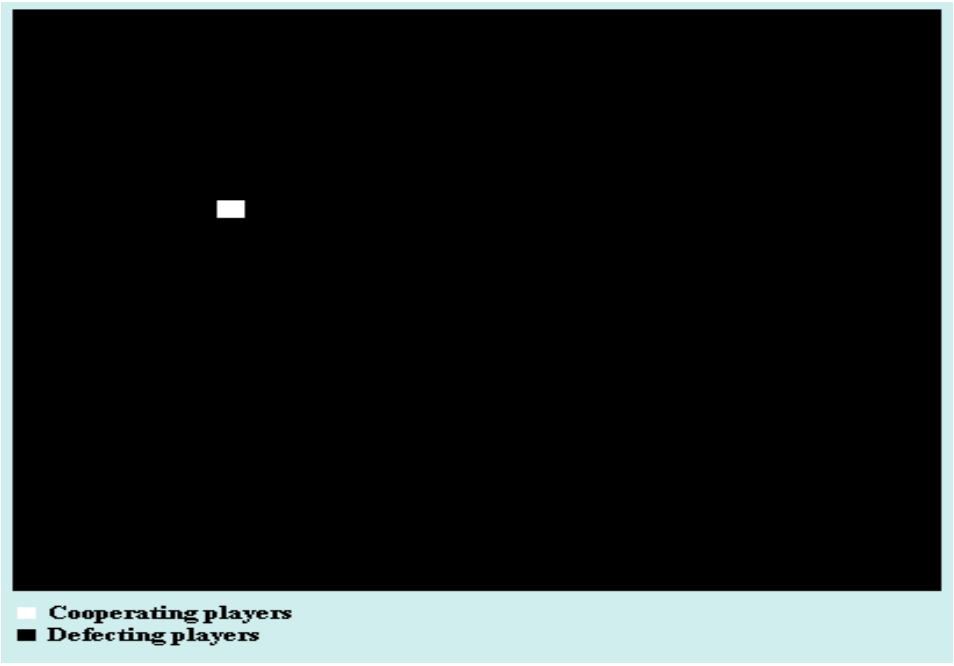


a)

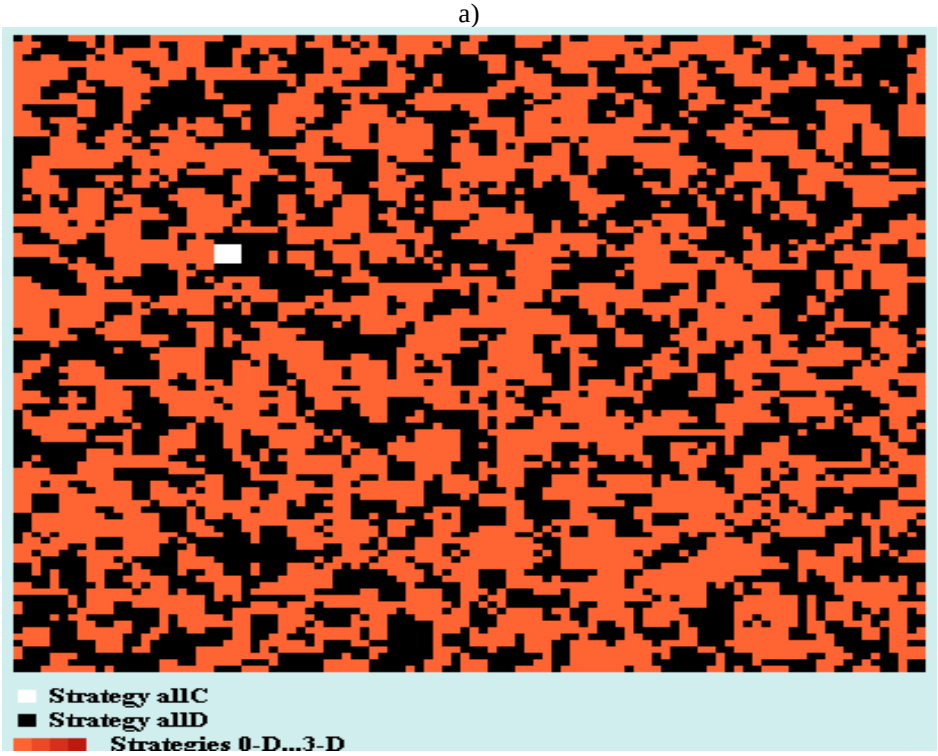


b)

Fig. 11 Agents' actions (a) and strategies (b), run with b=1.6, 3-D



a)



b)

Fig. 12 Agents' actions (a) and strategies (b), run with b=1.7, 3-D

When b=1.6 we can see the first time in the action pattern (see, Fig. 11a) that players playing action C create isolated structures surrounded by prevailing number of players playing action D. However, when we look at the strategies pattern (Fig. 11b) we notice that most strategies all-D was replaced by strategy k-D. Finally, when we look in Fig. 12 (b=1.7) we can see (Fig. 12a) that except a tiny island of players playing C, near 100% of players selects action D. However, dominating strategy performed by players is not strategy all-D but strategy k-D.

REFERENCES

[1] R. Axelrod, "The Evolution of Cooperation", Basic Books, 1984
[2] R. Axelrod, "The Evolution of Strategies in the Iterated Prisoner's Dilemma". In: L. Davis (Ed.) Genetic Algorithms and Simulated Annealing, Pitman, pp. 31-41, 1987.
[3] E. Howley, C. O'Riordan, "The Emergence of Cooperation among Agents using Simple Fixed Bias Tagging", IEEE Congress on Evolutionary Computation, vol. 2, pp. 1011-1016, 2005.
[4] H. Ishibushi, N. Namikawa, "Evolution of Iterated Prisoner's Dilemma Game Strategies in Structured Demes Under Random

Pairing in Game Playing", IEEE Transactions on Evolutionary Computation, vol. 9, No. 6, pp. 552-561, 2005.

[5] Y. Katsumata, Y. Ishida, "On a Membrane Formation in a Spatio-temporally Generalized Prisoner's Dilemma", In: H. Umeo, S. Morishita, K. Nishinari, T. Komatsuzaki, S. Bandini (Eds.), ACRI 2008, LNCS 5191, Springer, pp. 60-66, 2008.

[6] M. A. Nowak, R. M. May, "Evolutionary games and spatial chaos", Nature 359, pp. 826-829, 1992.

[7] M. Osborne, "An Introduction to Game Theory", Oxford University Press, 2009.

[8] M. L. Tsetlin, "Automaton Theory and Modelling of Biological Systems", Academic Press, 1973.

[9] X. Yao, P. J. Darwen, "An Experimental Study of N-Person Iterated Prisoner's Dilemma Games", In: X. Yao (Eds.) Progress in Evolutionary Computation, LNCS 956, Springer, pp. 90-108, 1995.

[10] S. Wolfram, "A New Kind of Science", Wolfram Media, 2002

★★★