**ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY**
Volume **16**, Numbers 2–3, 2013, 192–202

# A Game Theoretical Perspective on Cooperation in Low Confidence Environments

Marcel CREMENE[1], Ligia CREMENE[1], Dumitru DUMITRESCU[2]

[1]Technical University of Cluj-Napoca
E-mail: cremene@com.utcluj.ro
[2] Babes-Bolyai University, Cluj-Napoca
E-mail: ddumitr@cs.ubbcluj.ro

**Abstract.** Cooperation is a key behaviour for common welfare. Spontaneous order occurs only under special conditions and often remains immaterialized or leads to situations that are not favorable to the society as a whole. The information society revolution, overlapping the particular political and economic context creates a highly dynamic, unpredictable social context. Emergence of cooperative behaviour in such a context is analyzed within the framework of Computational Game Theory. The Prisoner's Dilemma game yields a non-cooperative behaviour if played only ones. When the game is repeated players learn that they can gain more if they cooperate. But repetition does not seem to be favored by a rapidly changing environment. Our main hypothesis is that the lack of a reasonable expectation concerning the continuation of the game induces selfish, short-time thought actions, or actions based on altered perception, which do not favor social cooperation. Normal game and spatial scenarios are analyzed. Results indicate that low confidence conditions increase the incentive to defect.

## 1. Introduction

The information society revolution, overlapping the particular political and economic context, creates a highly dynamic, unpredictable social context. This is especially the case of developing countries like Romania for instance.

In order to ensure the persistence of a somehow natural social order, institutions and cultural norms have a crucial role in facilitating a trustworthy and predictable environment. But weak institutions means social insecurity and this seems to be the case of developing countries [8]. In such countries corruption is favored by the insecure environment [8].

Trust is a necessary base for cooperation. Yet, trust is more likely to emerge through a bottom-up institutional scheme, rather than a top-down set of actions, directed from a higher authority. This is mainly because an authority is often suspected of corrupt, hidden interests.

Cooperation is a key behaviour for common welfare. Spontaneous order occurs only under special conditions. Order often remains immaterialized or leads to situations that are not favorable to the society as a whole [14]. These considerations lead to the idea of studying the conditions for emergence of cooperation through self-organization. Computational Game Theory provides a fertile framework and the tools to analyze the emergence of the social cooperative behaviour. The role of the authority may be to finely tune some rule parameters that influence the players' payoffs in various situations. The Tragedy of the Commons [7], Prisoner's Dilemma [4], and Collective Action Logic [13] are the main considered models when approaching the governing of the commons [14].

The paper proposes two game models. The first model, based on a two player iterated prisoner's dilemma game (IPD), studies the effect of the lack of confidence on the strategies preferences. The second model, based on a spatial PD game, studies another form of a lack of confidence: a player is more or less confident that a strategy will be successful in the future if it was successful before. An imitation probability is used to describe this situation. In this case, 2D (latticeal) IPD game scenarios are studied.

**Outline.** The remainder of this article is organized as follows. Section 2 briefly discusses the Prisoner's Dilemma (PD) game with its various forms and its relevance to cooperation. Section 3 introduces the proposed game model based on the Continuation Confidence Degree (CCD) concept and presents three implementation scenarios for the proposed game model. Section 4 is dedicated to the spatial PD models. Final remarks conclude our paper.

## 2. Cooperation and the Prisoner's Dilemma game

Prisoner's Dilemma (PD) game has been extensively used for analyzing social interactions [4, 5]. This game describes, in a simple way, a situation when cooperation is the best strategy - if both players are using it. A selfish, non-cooperative action, as the Nash equilibrium indicates, leads to a situation far from optimal, for both players. It is well known that Prisoner's Dilemma leads to a non-cooperative behaviour if played only once [15], [3], [1]. Despite the mutual benefits of cooperation, the fear of treason blocks cooperation [6]. The game situation where both players defect is known as the Nash equilibrium of the PD game [15].

However, when the PD game is repeated, players learn that they can gain more if they cooperate. The iterative solution to Prisoner's Dilemma was presented by Axelrod [3]. According to this, applying a simple 'Tit-for-Tat' (TFT) strategy, where the player answers to cooperation by cooperation and to treason by treason, results in a learning process at the end of which each player becomes aware of the fact that, on the long run, the cooperation strategy ensures a higher payoff than the defect strategy, and is thus rationally optimal. The simplest explanation for this fact may be that a defecting player could be punished in the next round of the game.

Novak and May present a series of experiments based on a spatial version (lattice) of the PD game [11, 12]. Their game may be described as follows: each player is represented as an element of a matrix. The player may be a defector or a cooperator. Players do not have memory. Each player plays PD with its adjacent eight neighbours including with himself. The game is synchronous. The player payoff is equal to the sum of the payoffs obtained in each of these games. In the next round, each player will use the strategy of the neighbour having the highest payoff (imitation). Some fractal patterns, sometimes chaotic ones, are obtained function of the game parameters. The most important parameter is the defector advantage when playing against a cooperator: when this advantage is high enough, the defectors are invading the whole space.

Novak and Sigmund [9] investigate how the TFT strategy evolves in a population. It is shown that, in a population dominated by defectors, a small fraction of TFT players can change the situation and create a new robust population that will prove difficult to invade by another strategy. The player considers the last round and takes a decision with a certain probability (short memory, uncertain decision). A TFT weakness is that mistakes give rise to sequences of alternative defections. A solution is to introduce forgiveness with a specific probability. The GTFT (generous TFT) strategy seems to be the optimal solution after the TFT has changed the defectors domination. However, it is TFT that starts the evolution towards cooperation.

In [10] comparative experiments with different spatial PD games are conducted. Several directions are investigated: deterministic versus stochastic rules, discrete versus continuous time, different geometries of interaction in regular or random spatial arrays (2D and 3D lattice) and different levels of self-interaction. Interaction with local neighbours can promote co-existence of different strategies (defectors and cooperators) whereas, in the case of homogeneous and random interactions, a single strategy wins. A correlation between the ratio of changing over non-changing players and the cooperators rate was observed. Another observation is that, in certain conditions, a strategy (i.e. cooperation) may extinct even if the players using that strategy have a higher average payoff. The main conclusions were the same for both continuous (simultaneous interaction and reproduction) and discrete (sequential interaction and reproduction) instances of the game. Symmetry is lost when probabilistic strategy update is used.

Our main hypothesis is that the lack of a reasonable expectation concerning the continuation of the game induces selfish, short-time thought actions, or actions based on altered perception, which do not favour social cooperation. Such a scenario may be analyzed in the framework of the evolutionary game theory [2].

## 3. Repeated Prisoner's Dilemma with Continuation Confidence Degree – Proposed Model

The proposed model can be described as a repeated PD game with a Continuation Confidence Degree (CCD). A continuation confidence degree – confidence that a next round is possible – is associated to each player. This CCD reflects the subjective probability that the game will have a next round. The confidence degree may be considered as constant or variable over time. If the confidence degree of a player is lower than a certain threshold, the player tends to consider that the game ends at the present stage and therefore the Nash solution is the reasonable choice. Nash equilibrium is the non-cooperative (defect, defect) solution of the PD. If the continuation confidence degree of a player is higher than the threshold, the player takes into account future actions (including the possibility of being punished for a 'defect' strategy) and will thus have an incentive to cooperate.

The proposed game theoretical model may be implemented as a repeated Prisoner's Dilemma game, where the CCD factor is represented by a probability of choosing to cooperate, for each player. A high CCD indicates the choice to cooperate of, knowing that this is the best individual strategy when the game is repeated (indefinitely). In our experiments the game is repeated 1,000 times and the payoffs are added during these repeated games. Both players start by cooperating. The PD normal game is described in Table 1.

**Table 1.** Prisoner's Dilemma normal game

| | **Cooperate** | **Defect** |
|---|---|---|
| **Cooperate** | *R* = *3*<br>*R* = *3* | *S* = *0*<br>*T* = *5* |
| **Defect** | *T* = *5*<br>*S* = *0* | *P* = *1*<br>*P* = *1* |

When both players cooperate, each player's payoff is 3. If one player defects and the other one cooperates the defector's payoff is higher, 5. If both players defect their payoffs are minimum, 1.

**Scenario no. 1.** A first scenario considers that each player may choose to cooperate or to defect. Let p1 be the probability for Player 1 to defect, and p2 the probability for Player 2 to defect. The total payoff of Player 1, obtained in 1000 rounds, is depicted in Fig. 1.

As it can be observed in Fig. 1, a player may be tempted to increase her payoff by increasing the probability to defect, regardless of the probability to defect of the other player. When player 2 is doing the same thing, both payoffs are dramatically reduced.

A low continuation confidence degree means a high probability to defect and leads to a lower global payoff. If a player insists to cooperate, the other player will just take advantage of this situation and will increase her payoff by defecting. Thus, in a low confidence context, the “sucker” strategy (‘always cooperate’) is the worst strategy. The ‘always defect’ strategy makes things worse also but is more successful than the ‘always cooperate’ strategy.

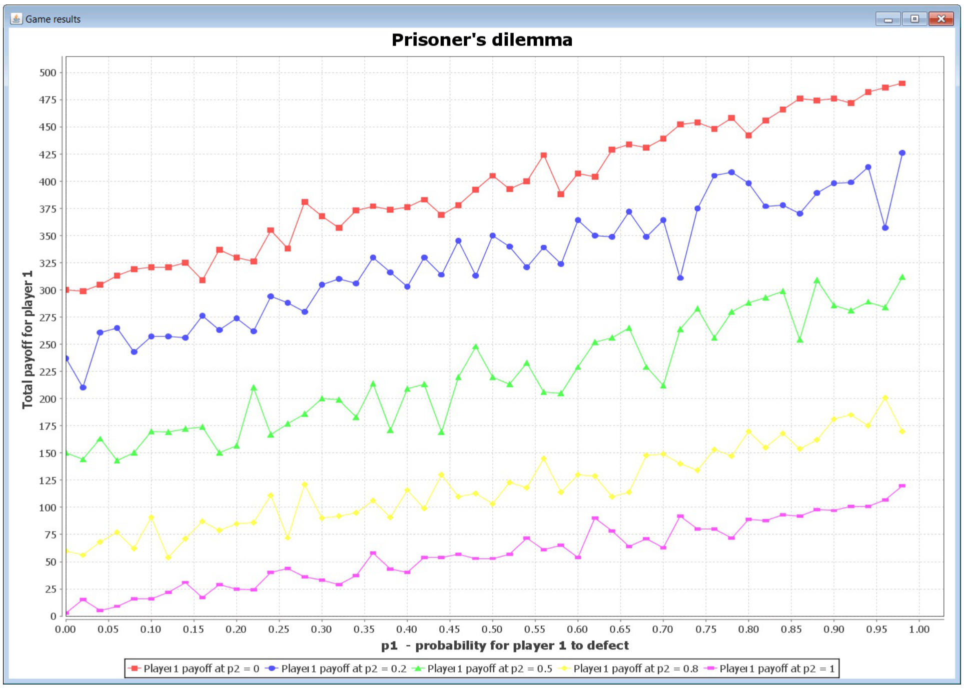


**Fig. 1.** Total payoff of Player 1. Player 1 defects with probability p1 or cooperates. Player 2 defects with probability p2 or cooperates. The value of p1 varies from 0 to 1 with a step of 0.01; p2 is set to: 0, 0.2, 0.5, 0.8, and 1, respectively.

**Scenario no. 2.** In a second scenario, the same game is played but with a small change: the first player decides to defect with a probability p1, and to play the ‘tit-for-tat’ strategy with a probability 1-p1.

‘Tit-for-tat’ was proved to be one of the best strategies in repeated PD game (Axelrod, 1984). The results are depicted in Fig. 2. It may be noticed that, if Player 2 is rather cooperative, a more probably defecting attitude will increase Player 1’s payoff. If Player 2 becomes unconfident and tends to defect more, Player 1 cannot increase her payoff by defecting more than by playing ’tit-for-tat’. When comparing the payoffs in Fig. 1 and Fig. 2, a conclusion is that it is definitely preferable to be a TFT player than a “sucker” (‘always cooperate’), especially in a low to medium confidence context.

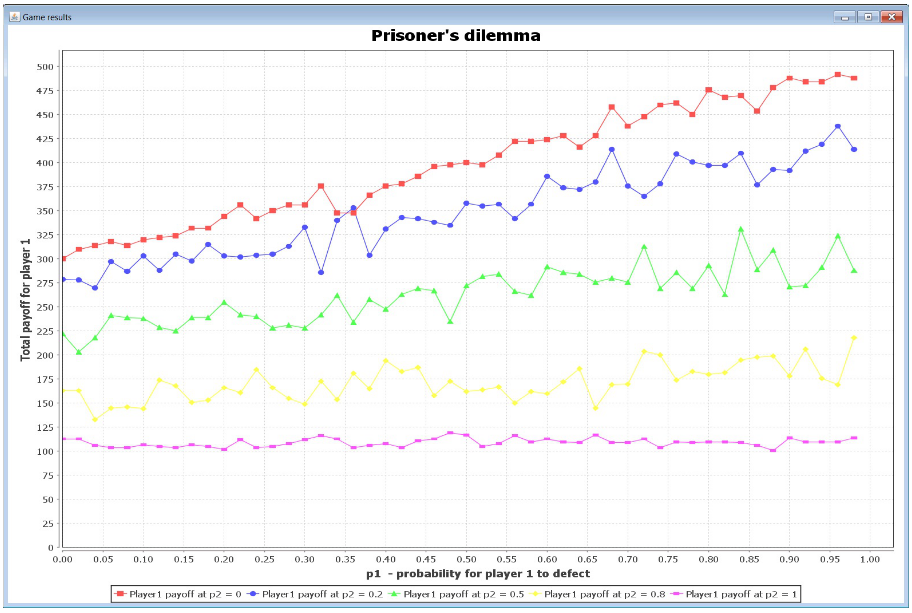


**Fig. 2.** Total payoff of Player 1. Player 1 defects with probability p1 or plays TFT. Player 2 defects with probability p2 or cooperates. The value of p1 varies from 0 to 1 with a step of 0.01; p2 is set to: 0, 0.2, 0.5, 0.8, and 1, respectively.

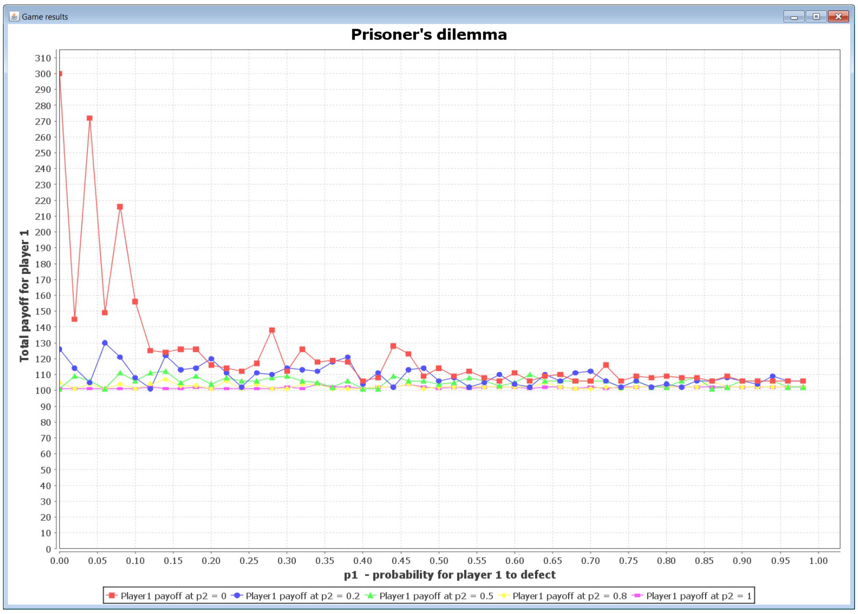


**Fig. 3.** Total payoff of Player 1. Player 1defects with probability p1 or plays TFT. Player 2 defects with probability p2 or plays TFT. the value of p1 varies from 0 to 1 with a step of 0.01; p2 is set to: 0, 0.2, 0.5, 0.8, and 1, respectively.

**Scenario no 3.** In a third scenario, both players decide to defect or to play TFT, according to a probability p1 or p2, respectively. As the payoffs captured in Fig. 3 reveal, in a low confidence context (high p1 and/or p2 values meaning a preference for defecting instead of playing TFT), Player 1 gets a low payoff but not as low as the "sucker". Player 1 has no reason to defect, rather than play TFT, in a situation where the other player also plays TFT. A small deviation from TFT, toward a defecting behaviour, has a high risk of payoffs fading. It is expected from both players to keep a pure TFT strategy. This may naturally lead to fully cooperative behaviour.

## 4. Spatial Prisoner's Dilemma

The spatial version of the PD game [11] is considered for the second model. The game space is described by a 200×200 grid. A player is placed in each cell of the grid and plays PD with her neighbours. A player accumulates the sum of payoffs from each individual game played with all neighbours.

### 4.1. Imitating the most successful neighbour

The default update mechanism is to imitate the best neighbour in the next round. In order to have a reference, we analyze first the results obtained in [11]. The player is considered memoryless and plays pure 'defect' or pure 'cooperate' strategies. After each round, a player adopts the strategy of the most successful neighbour (the one who gets the highest payoff). In other words, each player imitates the strategy of the most successful neighbour. The payoffs for a single-round game are captured in Table 2. The parameter b (b > 1) represents the advantage of defectors when playing against cooperators. The dynamical behaviour of the system depends very much on the value of b.

**Table 2.** Spatial Prisoner's Dilemma normal game proposed in [11]

| | **Cooperate** | **Defect** |
|---|---|---|
| **Cooperate** | $R = 1$<br>$R = 1$ | $S = 0$<br>$T = b$ |
| **Defect** | $T = b$<br>$S = 0$ | $P = 0$<br>$P = 0$ |

### 4.2. Proposed variation – imitating any better neighbour

In a low confidence context, the players may be less confident in imitating the others. This is due to the fact that the game conditions may change. Imitating any better neighbour mirrors the short term action behaviour, without a long term analysis. Thus, we propose a different imitation mechanism. A player may imitate not only the most successful neighbour, but also any better performing neighbour, with a probability p.

The strategy update mechanism is implemented according to the following algorithm: a player imitates a neighbour's strategy with a probability p if the neighbour's payoff is higher than her own payoff, and will memorize the new payoff. When another neighbour, with a higher payoff than the memorized value is found the player will imitate this neighbour's strategy with the same probability p. The process is iterated for each neighbour.

### 4.3. Spatial PD with probabilistic imitation mechanisms – numerical experiments comparison

We compare the results of the proposed imitation mechanism with the results reported in [11]. Figure 4 describes the situation after 100 rounds. The initial population contains 90% cooperators and 10% defectors.

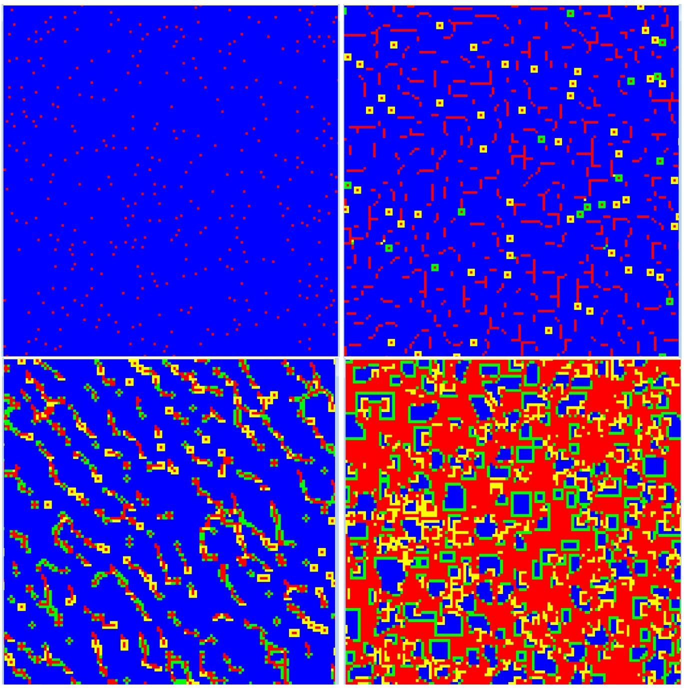

**Fig. 4.** Results reported in [11]: initial population with 90% cooperators and 10% defectors, after 100 rounds. b = 1.1, upper left – stable dot patterns, b = 1.25, upper right – stable line patterns, b = 1.5, lower left – stable line pattern with some oscillations, and b = 1.85, lower right – quasi-chaotic patterns. Colour code is the following: blue – is cooperating, did cooperate, red – is defecting, did defect, green – is cooperating, did defect, yellow – is defecting, did cooperate.

The value of b is a critical one as illustrated in Fig. 4. For b = 1.1 there is a cooperators' majority with some stable isolated defectors which do not disappear over time. Other values of b (1.25, 1.5) lead to stable patterns with some small random fluctuations. For b = 1.85 quasi-chaotic patterns exhibiting lots of fluctuations appear. Higher values, like b = 2, lead to a majority of defectors and some small islands of cooperators, while for b = 3 all population will defect after a number of rounds.

The results reported in [11] indicate that, if the defector's advantage when playing with cooperators, b, is high enough, then all players will defect. Some particular values of b may lead to a static or dynamic equilibrium between defectors and cooperators. Even for small values of b (1.1), defectors do not disappear completely. We should keep in mind that this model was based on the idea that every player tends to imitate the best neighbour strategy at each moment, without memory.

Figure 5 depicts the results obtained using the proposed strategy update algorithm, which is based on a stochastic preference for imitating better neighbours. The imitation probability is p = 0.1. The initial population contains 90% cooperators and 10% defectors. Figure 5 depicts the results after 100 rounds (the upper snapshots) and 20 runs (the bottom snapshots). Constant b takes different values: 1.1 – converge to cooperators, 1.25 – some static patterns with small oscillations appear, 1.5 (bottom left) – converging to defectors, and 1.85 (bottom right) – also converging to defectors in less than 100 rounds.

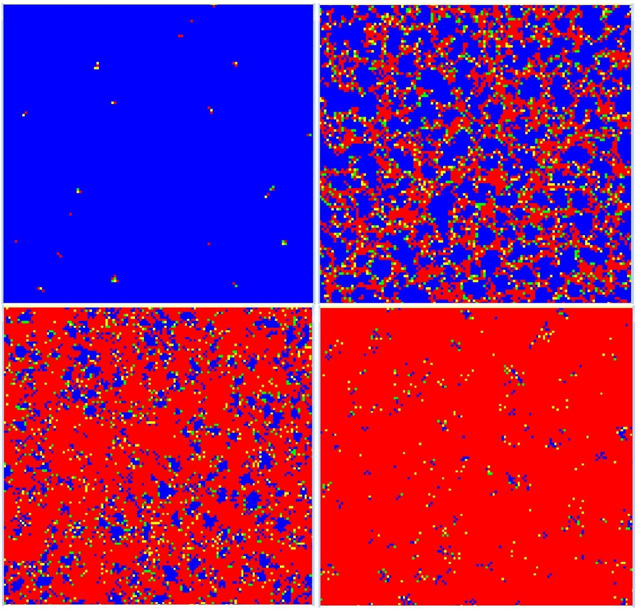

**Fig. 5.** Results obtained with the proposed algorithm: initial population with 90% cooperators and 10% defectors, after 100 rounds (upper row), and 20 runs (bottom row). b – advantage of defectors when playing against cooperators; b = 1.1, upper left – converge to cooperators, b = 1.25, upper right – some static patterns with small oscillations, b = 1.5, lower left – converging to defectors and b = 1.85, lower right – converging to defectors.

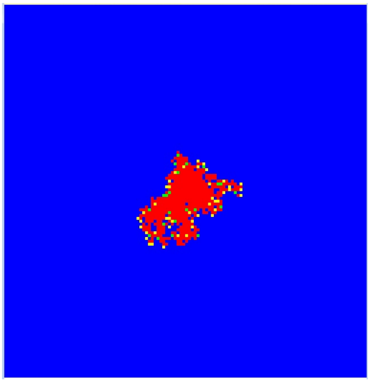

**Fig. 6.** Results obtained with the proposed algorithm: initial population with one defector in the middle, after 60 rounds. b – advantage of defectors when playing against cooperators; b = 1.5, in time the whole system converges to defectors.

The results are compared to the spatial PD proposed in [11]. It may be observed that even a small probability (p = 0.1) to imitate a neighbour with a higher payoff leads to faster convergence to a defectors' domination, even for small values of b (advantage of defectors when playing against cooperators), b = 1.5. For even smaller values of b, b = 1.1, the proposed mechanism (imitating any better neighbour) convergences towards a cooperators universality, whereas in the standard case some isolated defectors survive.

The quasi-chaotic dynamic behaviour appears in both cases, but, for different values of b and with different dynamics. Both experiments show that defector domination is easily obtained when the value of the advantage of defectors when playing against cooperators increases. A probabilistic imitation of the better neighbours is even more effective, in terms of strategy spreading, than the pure imitation of the best neighbour.

Other experiments were conducted with the proposed strategy update algorithm, starting with a single defector in the middle of a population of cooperators. Figure 6 depicts an invasion started by only one defector in the middle of a sea of cooperators, and b = 1.5. It was observed that a value of b = 1.5 (the advantage of defectors when playing against cooperators) is high enough to start an epidemic of defective behaviour in the entire population.

## 5. Conclusions

Experiments indicate that defective behaviour is very easily spread in a population with a majority of cooperative players when the defectors advantage is high enough. In a low confidence environment the TFT strategy proves to be better than 'always defect', which is also better than 'always cooperate'. According to our results, an insecure and unstable environment favours defective behaviour and discourages cooperation. This is consistent with the theory exposed in [8], which says that, in less developed countries, where institutions are weak, in order to fight against corruption the first priority should be to create a more stable, safe and predictable social context. An authority may be also preoccupied with finding situations that restrict/discourage the defective behaviour (i.e. punishment against defectors) by lowering the defectors advantage.

**Acknowledgments.** This research was supported by the Romanian national program PN II TE, code TE 252, financed by CNCS-UEFISCDI. The authors' contribution is equal.

## References

[1] Ahn T.K., Ostrom E., Schmidt D., Shupp R., Walker J., *Cooperation in PD Games: Fear, Greed, and History of Play*, Public Choice, **106**(1):137–155, Jan. 2001.

[2] Ainslie G., Gintis H., *The bounds of reason: Game theory and the unification of the behavioral sciences*, Princeton University Press, Princeton, NJ (2009). xviii + 286 pp., ISBN: 978-0-691-14052-0 (hc). Journal of Economic Psychology, **32**(1):201–204, February 2011.

[3] Axelrod R.M., *The evolution of cooperation*, Basic Books, New York, 1984.

[4] Dawes R., Eugene O.R.I., *Formal Models of Dilemmas in Social Decision-Making*, Defense Technical Information Center, 1974.

[5] Dawes R.M., McTavish J., Shaklee H., *Behavior, communication, and assumptions about other peoples behavior in a commons dilemma situation*, Journal of Personality and Social Psychology, **35**(1):1–11, 1977.

[6] Fukuyama F., *The Great Disruption: Human Nature and the Reconstitution of Social Order*, The Free Press, New York, NY, USA, 1999.

[7] Hardin G., *The tragedy of the commons*, Science, **XX**:1243-47, 1968.

[8] Johnston M., *Syndromes of Corruption: Wealth, Power, and Democracy*, Cambridge Univ. Press, 4 edition, 2010.

[9] Nowak M., Sigmund K., *Tit for tat in heterogeneous populations*, Nature, **355**:250–253, 1992.

[10] Nowak M. A., Bonhoeffer S., May R. M., *More spatial games*, Inter- national Journal Of Bifurcation And Chaos, **4**(1):33–56, 1994.

[11] Nowak M. A., May R. M., *Evolutionary games and spatial chaos*, Nature, **359**:826, 1992.

[12] Nowak M. A., May R. M., *The spatial dilemmas of evolution*, Interna- tional Journal of Bifurcation and Chaos (IJBC), **3**(1):35–78, Feb. 1993.

[13] Olson M., *The logic of collective action; public goods and the theory of groups*, Harvard University Press, 1971.

[14] Ostrom E., *Governing the Commons: The Evolution of Institutions for Collective Action (Political Economy of Institutions and Decisions)*, Cambridge University Press, Nov. 1990.

[15] Roth A. E., Murnighan J. K., *Equilibrium behavior and repeated play of the prisoners dilemma*, Journal of Mathematical Psychology, 1978.