Attraction and Cooperation in Space

Elpida S. Tzafestas, *Member, IEEE*

Abstract—**This paper presents an attraction mechanism and a study of its behavior in noisy spatial IPD games. The attraction mechanism is coupled with a regular IPD strategy to produce the final agent's behavior. We study the mechanism in 2D space to understand how it influences the emerging spatial structures. We find that any agent strategy, even a plainly irrational one, may become stable in a spatial game given appropriate attraction relations with the neighbors. Also, often the removal of attraction after stabilization causes no qualitative effect. Then we study the effect of introduction of "extreme" agents, such as a "Don Juan" agent that is attracted by all others. Such agents can change dramatically the structure of emerging spatial strategy blocks. Especially the addition or removal of agents after stabilization almost always leads to re-organization and re-stabilization to a new configuration. Various parameters of this model are studied and finally the notion of mobility in relation to evolution is rediscussed. All our results suggest that psychological mechanisms external to the actual "problem" (here IPD) can interfere with it and can actually lead to enhanced cooperation and social stability despite environmental noise and agent irrationality. Moreover, all social/spatial variants that induce more interactions between agents with non-reciprocal attraction relations (such as the extreme agents), are bound to lead to a better social average and to more complex structures.**

I. Introduction

A major research theme in both evolutionary computation and theoretical biology is the emergence and prevalence of cooperative behavior between selfish agents. The cooperation problem states that each agent has a strong personal incentive to defect, while the joint best behavior would be to cooperate. This problem is traditionally modeled as a special two-party game, the Iterated Prisoner's Dilemma (IPD).

At each cycle of a long interaction process, the agents play the Prisoner's Dilemma. Each of the two may either cooperate (C) or defect (D) and is assigned a payoff defined by table I.

TABLE I : IPD

AGENT	OPPONENT	PAYOFF
C	C	3 (= Reward)
C	D	0 (= Sucker)
D	C	5 (= Temptation)
D	D	1 (= Punishment)

Manuscript received April 10, 2007.

E. S. Tzafestas is with the Institute of Communication and Computer Systems, National Technical University of Athens, Zographou Campus, 15773 GREECE (phone: 30-210-772-2489; fax: 30-210-772-2490; e-mail: brensham@softlab.ece.ntua.edu). Web page: http://www.softlab.ece.ntua.gr /~brensham (includes a full color version of this paper)

The first notable behavior for the IPD designed and studied by Axelrod [1][2] is the Tit For Tat behavior (TFT, in short):

Start by cooperating,
From there on return the opponent's previous move

This behavior has achieved the highest scores in early tournaments and has been found to be fairly stable in ecological settings. TFT demonstrates three important properties, shared by most high scoring behaviors in IPD experiments.

- *It is good (it starts by cooperating)*
- *It is retaliating (it returns the opponent's defection)*
- *It is generous (it forgets the past if the defecting opponent cooperates again).*

Further strategies include stochastic ones ([9]), the Pavlov strategy ([11]) that cooperates when it has played the same move as its opponent etc. In the literature we may also find studies in a purely evolutionary perspective ([5]), theoretical or applied biological studies ([3][4][7]) and studies of modified IPD versions ([12]).

We adopt the noisy version of IPD in which there is a nonzero probability that an agent's action will be switched to the opposite, i.e. from *COOPERATE* to *DEFECT* or vice versa. It has been shown that retaliating strategies such as TFT can score quite badly in the presence of noise, despite their superiority in the non-noisy domain [6][8]. This happens because even accidental defections may lead to a persistent series of mutual defections by both players, thus breaking cooperation. The usual approach is to introduce some degree of explicit generosity to account for opponent's misbehaviors or to attempt opponent modeling.

Our approach purports to show that an independent psychological or social factor can allow agents in a society to cooperate fairly well despite noise and without explicit opponent modeling or other intricate reasoning behavior. The organization of the paper is as follows: in section 2 we introduce the attraction mechanism and we discuss the motivation behind it and its conceptual implications. In sections 3 to 5 we describe experiments done with various configurations, perturbation tests and personality extremes. Based on our findings, we revisit in section 6 the notion of mobility and evolution in the context of the spatial IPD. We finally sum up and give potential directions of future research in the last section.

II. Rationality and Cooperation

Our motivation behind the introduction of an attraction mechanism is the general observation that in human societies, and especially in economic contexts, the agents' behavior can be heavily influenced by external psychological and social factors and also many times it can be driven to behaviors outside their normal scope. By "external" we mean a factor or process that is not influenced itself by the primary agent task and does not normally participate in it. We are using the benchmark iterated prisoner's dilemma (IPD) in its noisy version as a study vehicle with a stronger bias toward defection, where we feel it could make sense to introduce such an external attraction factor. More specifically, we believe that biological evolution or, equivalently, social experience would spontaneously exploit any external factor that would induce better agent scores. This is particularly true for noisy environments where agent scores may degrade abruptly, and especially when interactions are lengthier.

The attraction mechanism relies on our everyday experience that people tend to be good and cooperative with other people that attract them and tend to be "regular" with the rest. This translates in our model as:

If (attracted by the opponent) then play ALLC
(always cooperate),
Else play as usually (for example, TFT)

We should note that noise is applied to the outcome of this behavior as well. We performed tournament experiments with populations of agents playing a noisy IPD. The agents are interconnected via a "web of attraction" where each agent is connected to (attracted by) a number of others. The normal behavior of an agent is usually one of ALLC, ALLD (always defect), TFT and Adaptive TFT [13], but we have also experimented occasionally with STFT (Suspicious TFT) or other strategies. We experimented with both uniform or mixed populations, whose agents have the same or diverse normal strategies. The reason we use mostly ALLC, ALLD, TFT and Adaptive TFT is that we want to make sure we explore the limits of our attraction mechanism by studying its effect on the extreme behaviors (ALLC and ALLD that act without feedback) as well as on the most intelligent ones (TFT that retaliates immediately and Adaptive TFT that tries to make sense of a situation).

In previous work [13], we have classified usual IPD strategies in two categories: "retaliating" (or "rational") and "irrational". Retaliating strategies are those mostly TFT-derived strategies that basically seek cooperation in the long run, but may start by exploring the opponent's reaction to a few initial D moves and will certainly retaliate the opponent's defections in some possibly intricate way (apparently, this general behavioral organization is the best choice so as to achieve maximum scores in the long run). For example, the suspicious tit-for-tat (or STFT) strategy starts by defecting, and then plays usual tit-for-tat. On the contrary, irrational strategies are those that do not employ any feedback from the game and play blindly using some innate law (although in some cases this can work, in the general case such strategies lose in the long run). For example, periodic strategies repeat patterns of C's and D's, such as CDD, CCD, CDCD etc. In this sense, ALLC and ALLD are irrational strategies whereas TFT and Adaptive TFT are retaliating ones. We have shown that the Adaptive TFT strategy manages to differentiate between retaliating and irrational strategies and especially between a retaliating strategy and an irrational one that has initially the same behavior. For example it manages to converge to total defection against CDCD, that resembles STFT in the beginning. This feature allows it to achieve much higher scores than TFT in usual diverse social environments. As in real life, we would expect irrational strategies to profit more from our attraction mechanism or other similar mechanisms and rational ones to be less dependent on such add-ons. In sum, we expect cooperation to be able to emerge in social interactions even in the absence of rationality and good reason.

Before proceeding to describe the experimental setup and the results obtained, we should stress the fact that the attraction mechanism described is in our own terms irrational in that it does not depend on any real feedback of the agent. Our results then suggest that the coupling of reasoning mechanisms with reactive ones (such as attraction, be it physical, emotional, social or other) may be advantageous to social behavior and this is in line with current trends in cognitive and social science.

III. Spatial Games I: Basic Model

In accompanying work of ours [15] we have experimented with populations of uniform or diverse agents by running tournaments for varying factors of attraction (number of agents that an agent is attracted to) and pairing types. Our results showed that the bigger the frequency and diversity of interactions within non-reciprocal attraction relations, the better the average score of the overall society. We retain the same experimental setup of noisy IPD games between pairs of agents with the length of interactions set to 100 and the degree of noise set to 10%. Each agent is attracted by N of the surrounding agents (N ranges from 0 to 9, including oneself). We call N the attraction factor of the system and in the experiments that follow we have fixed it to 2.

As a verification of the power of attraction to achieve high-scoring agents, we have run spatial games in two-dimensional grids consisting of 100 uniform agents, i.e., agents having the same IPD strategy, organized in a 10x10 grid. In these games, all agents run a 100-cycle noisy IPD game (with 10% noise) against their neighbors at each cycle and no evolution is allowed. The average scores are collected and compared for the cases of attraction factors 0 (without attraction) and 2 (with attraction). A few indicative results are given in Table II, that confirm the superiority of attraction by default. For reference, we are giving also the scores for the no noise case as well as a set of scores for three random configurations (the first one is a mix of TFT and Adaptive TFT, while the other two are random mixes ike the ones given in figures 2, 4 and 5). Notice that ALLD profits from the introduction of attraction because it is given the chance to exploit others. On average an ALLD strategy exploits and is being exploited in a web of attraction so that its average score rises beyond its theoretical equilibrium. This is why the scores with noise improve for the ALLD unlike what happens with retaliating agents (TFT and Adaptive TFT). Both TFT and ADAPTIVE degrade with noise but remedy this a little with the help of attraction – the improvement in the case of TFT is much more pronounced than in the case of Adaptive TFT which is more robust to noise. In random populations we also have significant score gains; scores are comparable to the theoretical score without noise (random 1 and 2) or even surpass it (random 3 that contains a high proportion of ALLD). As a result, and because point-to-point scores vary greatly depending on agent strategies and presence of attraction links, we expect the local uneven relations between agents to influence the evolution of spatial structures in unforeseen ways.

TABLE II
SCORES IN UNIFORM AND RANDOM CONFIGURATIONS

STRATEGIES	ATTRACTION YES	ATTRACTION NO	NO NOISE
100 ALLD	175.51	114.77	100
100 TFT	258.86	232.33	300
100 ADAPTIVE	272.43	255.39	300
RANDOM 1	255.70	228.44	262.35
RANDOM 2	245.75	215.51	244.57
RANDOM 3	236.92	204.66	213.49

We have then run spatial games ([10]) in two-dimensional grids consisting of diverse agents, i.e., agents having different IPD strategies. As before, all agents run a 100-cycle noisy IPD game (with 10% noise) against their neighbors at each cycle and with an attraction factor set to 2, unless otherwise stated. After all games are finished, the agent at each place assumes the winner strategy among itself and its neighbors, i.e., the strategy whose total score is greatest. The spatial game continues in the next cycle with the new generation of strategies and so on. Such experiments show the evolution of different types of strategies in spatial grids.

Experiments starting with regular or random configurations of agents of the four strategies presented in the previous section (ALLC, ALLD, TFT and Adaptive TFT), as well as irrational ones, have been found to quickly settle on a stable configuration where any agent strategy or any combination of strategies can survive and be present in high proportions, even plainly irrational ones, such the ALLD or the periodic ones. In the correspoding no noise, no attraction case [14], the final stable configurations consist of high percentages of Aadaptive TFT and very low percentages of ALLC with occasional small stable territories occupied by ALLD. Figure 1 shows such an initial configuration and two final stable ones, for the no noise, no attraction case and for the noisy case with attraction. We start without noise and attraction and run the system for 10 cycles. Then we reinitialize the system with noise and attraction and rerun the system for 25 cycles. In the first case, we observe a clear domination of the Adaptive TFT strategy that drives ALLD to extinction, whereas in the second case the Adaptive TFT is being invaded by ALLC that comes thus in direct contact with the ALLD strategy and is finally taken over by ALLD.

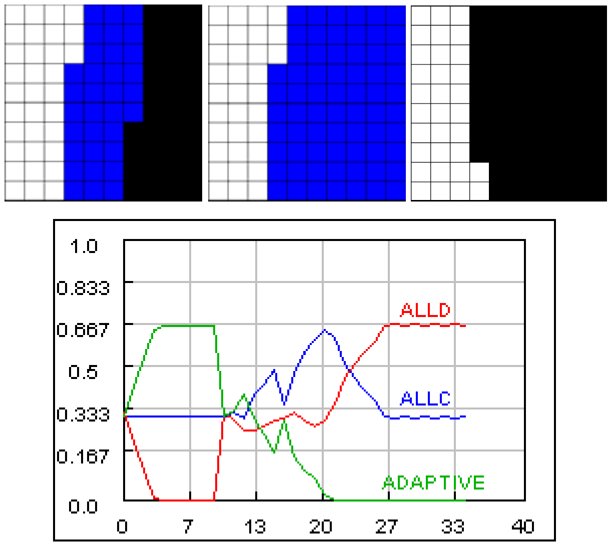
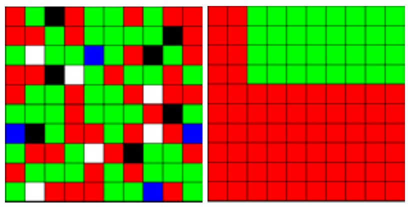


Fig. 1. (a, top row left) Initial configuration of Adaptive TFT (blue), ALLC (white) and ALLD (black) in a spatial IPD game (10x10 grid). (b, top row middle) Final configuration after 10 cycles without noise and without attraction. (c, top row right) Final configuration after reinitialization and another 25 cycles with noise and attraction. (d, bottom row) Evolution of strategies in the population.

Other unusual results are taken in the case of totally random configurations that often enough manage to stabilize to configurations comprising high proportions of irrational agents. For example, figure 2 presents an experiment where an initial random configuration leads to a stable final configuration of only TFT and irrational agents.



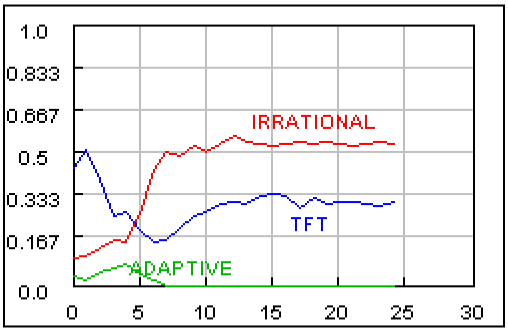


Fig. 2. (a, top row left) Initial random configuration of TFT (green), Adaptive TFT (blue), ALLC (white), ALLD (black) and CDD (red) in a spatial IPD game (10x10 grid). (b, top row right) Final configuration after 25 cycles with noise and attraction. (c, bottom row) Evolution of strategies in the population.

IV. Spatial Games II: Perturbations

In order to examine the potential of the systems with attraction for self-organization and resistance to perturbations, we have performed three kinds of perturbation studies. First, we have performed the "shuffle test", where, after the population has stabilized, we shuffle the agents, i.e. we move each one of them in a random place on the grid. In this case, we have generally found that the system restabilizes to a new configuration not too different qualitatively, unless Adaptive TFT is present in the original configuration in high proportions, in which case it is favored by shuffling. Figure 3 shows the system of figure 2 after shuffling and re-stabilization. Contrary to what one might expect and to results in the non-noisy case [14], the CDD agents don't die out; on the contrary, they proliferate during re-organization. The conclusion from these results is that the attraction web has the power to maintain stability once reached, without radical deviations in population composition (such as extinction of a strategy), because it defines a "spreading" mechanism that counter-balances shuffling.

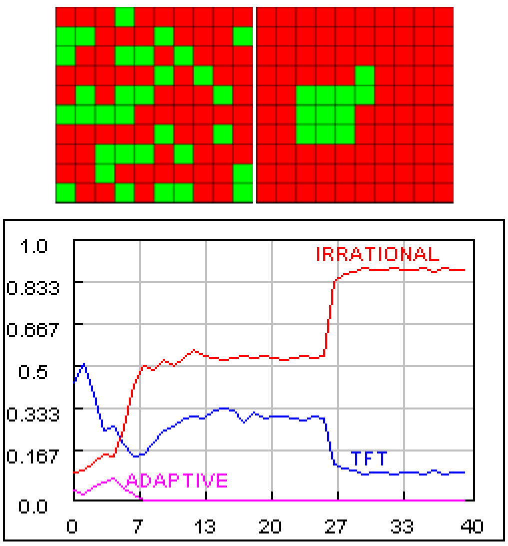


Fig. 3. (a, top row left) Initial configuration of TFT (green) and CDD (red) in a spatial IPD game (10x10 grid), shuffled from figure 2(b). (b, top row middle) Final configuration after 15 cycles with noise and attraction. (c, bottom row) Evolution of strategies in the population.

The second kind of test we have performed is the "injection test", where, after the population has stabilized, we inject a uniform block of agents (i.e., all bearing the same strategy) somewhere in the grid. In this case, we have found that the system restabilizes to a new configuration that can be radically different. Figure 4 presents an experiment where an initial random configuration leads to a stable final configuration of AdaptiveTFT, ALLD and irrational agents. There we inject a block of ALLC agents that invade the Adaptive TFT population and thus allow the irrational and ALLD agents to quickly take over. In general, we have found much easier to distort a stable configuration by injection than by shuffling, where distortion means to produce a new system qualitatively different as far as population composition is concerned.

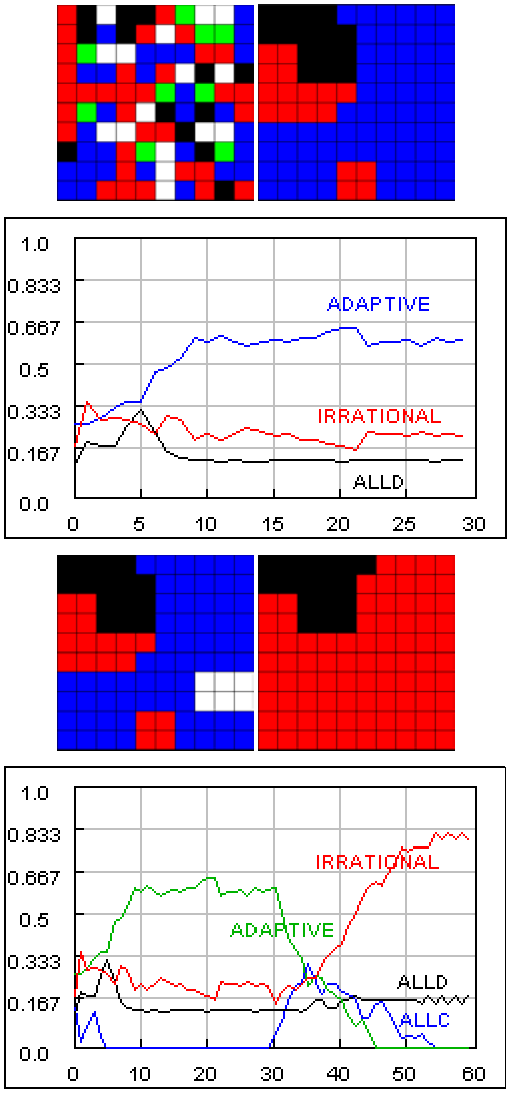


Fig. 4. (a, top row left) Initial configuration of TFT (green), Adaptive TFT (blue), ALLC (white), ALLD (black) and CDD (red) in a spatial IPD game (10x10 grid). (b, top row right) Final configuration after 30 cycles with noise and attraction. (c, second row) Evolution of strategies in the population. (d, third row left) The system of (b) with a block of ALLC

agents injected on the right side. (e, third row right) Final configuration after another 30 cycles with noise and attraction. (f, bottom row) Evolution of strategies in the population.

The final kind of test we have performed is the "attraction in-out test" where we have run a system with attraction and we have removed it after stabilization or vice versa. In both cases, the results do not change significantly qualitatively after stabilization. In the case of attraction removal, this is a little surprising because we would expect a system to degrade without attraction. However, this does not happen, because apparently attraction has led the system to a stable configuration and is not useful anymore. Attraction is useful when there is high diversity in the population as far as spatial relations are concerned (mixed populations instead of adjacent, for example). On the other hand, a system without attraction can settle, although sometimes after quite a long time, in very "poor" configurations as far as population diversity is concerned (most often in "flat", single-strategy populations). In this case, the insertion of attraction has generally no effect. Again, attraction makes sense when a system appears to be in a "chaotic" or random state, and it manages to have it settle fairly quickly to a stable configuration. In cases where the stabilized system contains more than one type of agents, the insertion of attraction can indeed make a difference. For example, figure 5 presents an experiment where an initial random configuration leads to a stable final configuration of ALLC and irrational agents (system with noise but without attraction). At this point, we enable attraction thus allowing CDD to invade ALLC and quickly take over. The issue of attraction in-out deserves its own focused study, especially because the noisy spatial games are very rich in emergent phenomena, most notable being the presence of ecological-type although instable cycles between strategy-species for long periods.

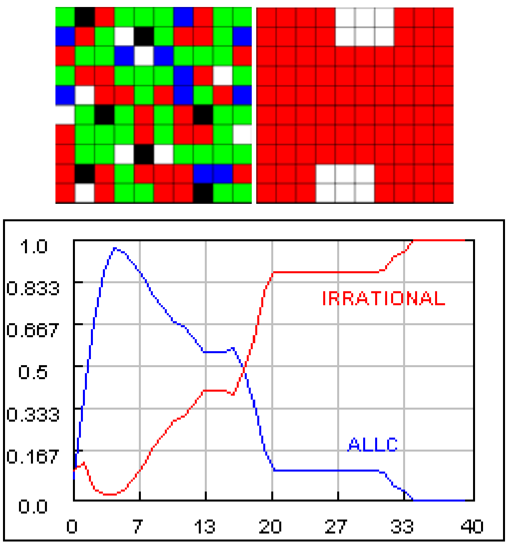


Fig. 5. (a, top row left) Initial configuration of TFT (green), Adaptive TFT (blue), ALLC (white), ALLD (black) and CDD (red) in a spatial IPD game (10x10 grid). (b, top row right) Stable configuration after 30 cycles with noise but without attraction. At this point, we enable attraction and ALLC dies out in a few cycles. (c, bottom row) Evolution of strategies in the population.

V. Spatial Games III: Personality Experiments

Finally, we have experimented with "extreme" agent personalities as far as attraction is concerned. More specifically we defined two new types of agents, the *"Don Juan"* and the *"Sex Symbol"* agents. The first one is attracted by every other agent in the system just like a real womanizer, while the second one is an object of attraction for all other agents, hence it behaves as a sex symbol. Again, these apply to the game setting of previous settings, thus a Don Juan agent is attracted by all the agents around him while the Sex Symbol attracts all agents around him. The results of the introduction of those extreme agents in tournaments are detailed in [15]. In summary, the Don Juan agent has consistently very low scores, while the Sex Symbol agent has consistently very high ones. This is because the Don Juan is easily exploited while the Sex Symbol is a dedicated exploiter. The Don Juan profile can lead an agent to become extremely exploited by fellow agents in particular social settings, and thus obtain very low scores to the benefit of other agents in the system. Such an agent may easily serve as a scapegoat.

In the spatial games of the present study, it is interesting to investigate whether the introduction of extreme agents influences the types of spatial structures that emerge. We distinguish two cases: either the "extreme personality gene" (the property of Don Juan or Sex Symbol) will be copied together with strategy during evolution from generation to generation of agents, or it will not.

In the first case, we observe that the Sex Symbol gene spreads very rapidly in the population and finally takes over the whole grid. If both Don Juan and Sex Symbol genes are present in the population, then the Sex Symbol always prevails. The spreading of the Sex Symbol gene should be expected because this agent achieves the highest scores, thus its neighbors copy its strategy and gene. The Don Juan gene spreads often but very slowly and generally stabilizes to a subpopulation. This is a side-effect of the fact that the Don Juan gene induces high scores to its neighbors and not to itself, thus it is not immediately copied. On the contrary the Don Juan agent will copy the highest-scoring neighbor strategy thus coming to balance with the neighbors; this can indirectly allow him to obtain high scores later and have its gene spread in the population. The Don Juan gene seldom takes over the whole space in random populations and very often dies out. Figures 6 and 7 present experiments where an initial random configuration, with a few Don Juans or a few Sex Symbols and Don Juans respectively, leads to a stable final configuration with only Don Juans or with all Sex Symbols, respectively. In both cases, the final configurations are composed of block of uniform locks both in strategy and in extreme personality gene.

Let us now turn to the case when the extreme personality gene is not transmitted. This is more reasonable, because at first we expect agents to be able to copy "playing" strategies of neighbors, but to be practically unable to copy reactive, uncontrollable (or even unconscious) attraction behavior that can be thought of as purely emotional or psychological in general. In this case, we observe small clusters of uniform strategy to emerge around the extreme agents and to rapidly stabilize. Figure 8 (a,b,d) presents such an experiment where an initial random configuration with a few Don Juans and a few Sex Symbols leads to a final stable configuration with visible clustering around the extreme agents. These results indicate that the introduction of extreme agents to the spatial system leads to finer grain emergent spatial structures, i.e. to smaller structures of simpler shape than in the absence of extreme behaviors.

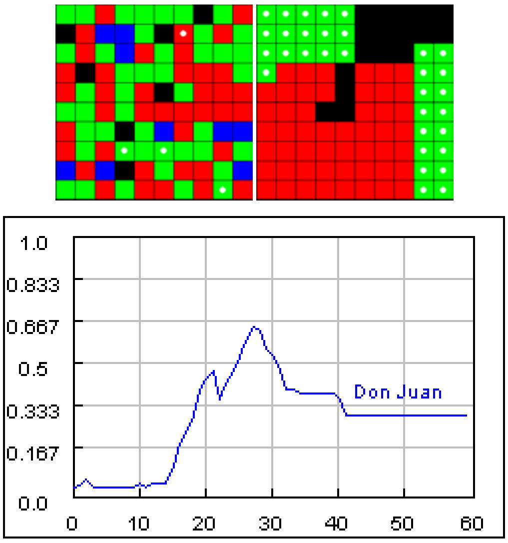


Fig. 6. (a, top row left) Initial configuration of TFT, Adaptive TFT, ALLC, ALLD and CDD in a spatial IPD game (10x10 grid) with 4 Don Juan agents. Color coding: Blue (Adaptive TFT), White (ALLC), Black (ALLD), Green (TFT), Red (Irrational, here CDD), White bullet (Don Juan). (b, top row right) Stable configuration after 60 cycles. (c, bottom row) Evolution of Don Juan agents in the population. The strategies relative populations have also stabilized.

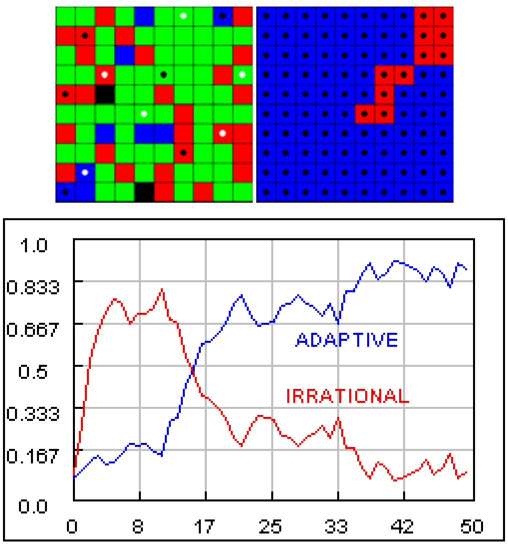


Fig. 7. (a, top row left) Initial configuration of TFT, Adaptive TFT, ALLC, ALLD and CDD in a spatial IPD game (10x10 grid) with 6 Don Juan and 6 Sex Symbol agents. Color coding: Blue (Adaptive TFT), White (ALLC), Black (ALLD), Green (TFT), Red (Irrational, here CDD), White bullet (Don Juan), Black bullet (Sex Symbol). (b, top row right) Stable configuration after 50 cycles. (c, bottom row) Evolution of strategies in the population.

To finish the study of extreme agents, we try a few perturbations. First, we remove the extreme agents after stabilization, as in figure 8(c,d), where after the system has stabilized at t=20 we remove all extreme agents. We observe that the fine structure of figure 8b disappears and the system settles down to the more uniform and "flat" structure of figure 8c, with two big areas one for ALLC and one for ALLD.

Second, we introduce extreme agents in a stabilized system without extremes, as in figure 9, where after the system has stabilized at t=20 we introduce a few extreme agents. We observe that the relative proportions of the three populations change drastically and the irrational agents multiply quickly.

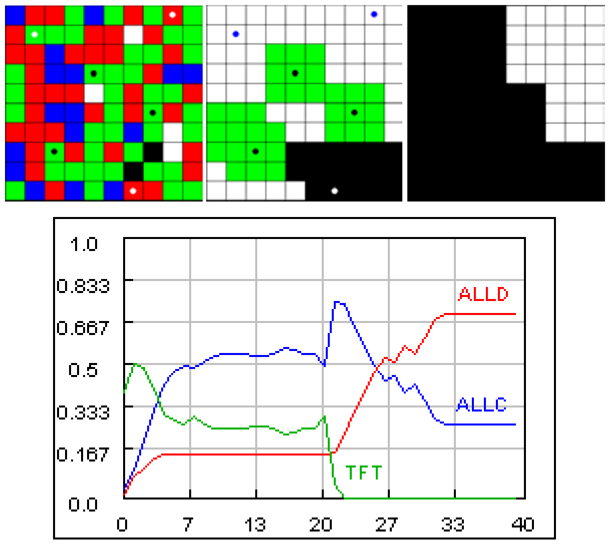
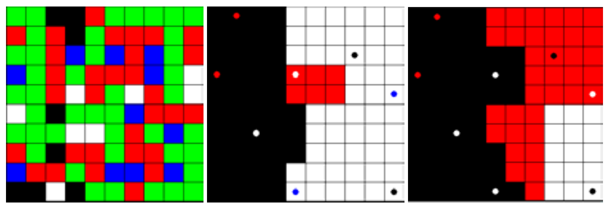


Fig. 8. (a, top row left) Initial configuration of TFT, Adaptive TFT, ALLC, ALLD and CDD in a spatial IPD game (10x10 grid) with 6 Don Juan and 6 Sex Symbol agents. Color coding: Blue (Adaptive TFT), White (ALLC), Black (ALLD), Green (TFT), Red (Irrational, here CDD), White bullet (Don Juan, blue in white background), Black bullet (Sex Symbol, white in black background). (b, top row middle) Stable configuration after 20 cycles. At this point, we remove all extreme agents. (c, top row right) New stable configuration after another 20 cycles. (d, bottom row) Evolution of strategies in the population.



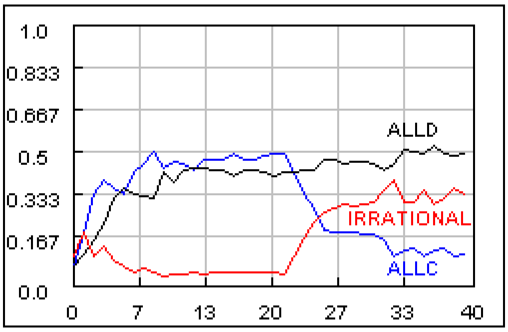


Fig. 9. (a, top row left) Initial configuration of TFT, Adaptive TFT, ALLC, ALLD and CDD in a spatial IPD game (10x10 grid) without extreme agents. Color coding: Blue (Adaptive TFT), White (ALLC), Black (ALLD), Green (TFT), Red (Irrational, here CDD), White bullet (Don Juan, blue in white background), Black bullet (Sex Symbol, white in black background). (b, top row middle) Stable configuration after 20 cycles. At this point, we introduce the shown 4 Don Juan and 4 Sex Symbol agents. (c, top row right) New stable configuration after another 20 cycles. (d, bottom row) Evolution of strategies in the population.

VI. Discussion: Evolution and "Real" Mobility

The potential of attraction to achieve high scores and to remedy for noise, as well as to stabilize spatial structures, acts as evidence supporting the hypothesis that such a mechanism would be evolutionarily advantageous. However, the spatial model used in this paper is the abstract one introduced by Nowak and May [10]. While this captures the main evolutionary feature of transmission from generation to generation, it fails to even crudely model spatial or "social" mobility, despite its name. In a more realistic model, we would expect agents to move around in the grid and cluster where they achieve highest scores. This way, in a system with attraction we would see agents clustering in groups interconnected with dense relations of attraction. Under evolutionary pressure, attraction as well as mechanisms to *develop* it would evolve and spread in the population.

To this end, we performed an initial set of experiments with a moving population of 225 agents in a 7x7 grid with unlimited carrying capacity per position. The pairwise game setup is as before, but the attraction pairing differs because an agent can be attracted by any other agent on the grid (we set the attaction factor to 20, i.e. to approximately 1/10 of the population). This way, we expect the agents to actively seek partners by moving on the grid. There are at least two parameters to consider: whether agents change strategy (update as before, after each step) and how they move around (toward the position of the highest scoring neighbor, with or without noise).

We tested both the move-without-update and move-with-update case, with and without motion noise. In the first case, with absence of noise the system settles very quickly to a configuration of isolated blocks of agents with slightly improved average actual attraction factor (i.e. the fraction of attractive neighbors), as figure 10 shows. With added noise, the system does not manage to improve the average actual attraction factor and does not settle either or does so very slowly. The same dynamical process when coupled with the feature of update-on-move (i.e. agents assuming the highest scoring neighbor's strategy as well as moving to its position) leads to mono-block configurations (uniform strategy blocks), but the results do not change qualitatively. None of the cases models exactly what we expected, i.e. an ever-moving system with agents self-organizing and clustering dynamically. Because the spatial structures are indifferent to the update-on-move feature, the mobility model remains as the one to blame for the premature and unnatural stabilization of the system. The lesson drawn is that, in order to model "true" mobility in social-spatial IPD games, we should consider introducing some other, possibly personal, agent behavior interacting with the social behavior (IPD) in some intricate way, not pure approach of agents. More specifically, we feel it is necessary to endow agents with a mechanism to assess continuously the expected utility of approaching or moving away from each one of the neighbor agents. Somehow then, the introduction of attraction would help agents to move closer to attractive neighbors, but indirectly, in order to maximize overall personal score and not because of attraction *per se*.

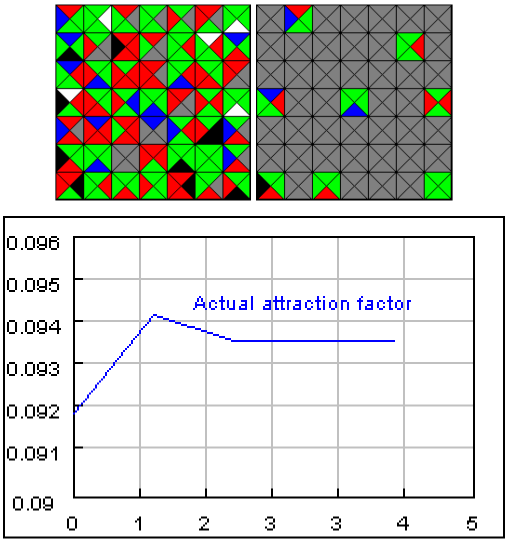


Fig. 10. (a, top row left) Initial configuration of TFT (green), Adaptive TFT (blue), ALLC (white), ALLD (black) and CDD (red) in a spatial IPD game with mobility (225 agents in a 7x7 grid, with gray background and agents occupying one quarter of the grid cell for visualization purposes). (b, top row right) Stable configuration after 5 cycles. (c, bottom row) Evolution of actual attraction factor in the system (average attraction factor with neighbors).

VII. Conclusion

In this paper we presented an attraction mechanism and a study of its behavior in noisy spatial IPD games. We implement attraction as a mechanism that changes the usual agent's behavior (strategy). More specifically, an agent follows its regular strategy unless it faces an attractive agent. In the latter case, she becomes unconditionally cooperative. We study the mechanism in 2D space to understand how it influences the emerging spatial structures. We find that any agent strategy may become stable in a spatial game given appropriate attraction relations with the neighbors. This is especially interesting for irrational behaviors, that would die out quickly in a no noise, no attraction environment. We have performed a number of perturbation tests (shuffling, injection, attraction in-out) and we have confirmed that any kind of strategy, and especially irrational ones, may benefit from any perturbation scheme. We have repeated the experiments for the case of agents with an extreme personality, namely, a "Don Juan" agent that is attracted by all others, and a "Sex Symbol" to whom all others are attracted. Again we have found similar results. In sum, the attraction mechanism has been found to promote cooperation in "difficult" environments, for example when there is noise or when irrational agents are present. The benefits are more pronounced when the attraction relations are widely spread in the population, non-reciprocal and uneven (as in the case of extreme agents). We have also briefly discussed why the spatial IPD game model is not a good vehicle for social-spatial study, for example it does not allow attraction to produce ever-moving, self-organizing structures with dynamically created clusters as expected.

Future directions of study include the investigation of the emergence of groups in a system with attraction, especially with a modified spatial mobility model. Groups can also preexist in the system, in which case it would be interesting to explore how they influence the structure of the interactions. Another result that deserves further study is the intricacy of the attraction in-out experiment that prompts for a study of how attraction could actually *develop* within a spatial IPD context. Therefore, the results of these and other assorted studies seem to have a huge potential for social and cognitive modeling, for example in issues such as ethnic mobility, cultural epidemiology and social microeconomics.

REFERENCES

[1] R. Axelrod and W. D. Hamilton, "The evolution of cooperation", *Science*, Vol. 211, pp. 1390-96, 1981.

[2] R. Axelrod, *The evolution of cooperation*. Basic Books, 1984.

[3] R. Axelrod and D. Dion, "The further evolution of cooperation", *Science*, Vol. 242, pp. 1385-90, 1988.

[4] M. W. Feldman and E. A. C. Thomas, "Behavior-dependent contexts for repeated plays of the prisoner's dilemma II: Dynamical aspects of the evolution of cooperation", *Journal of Theoretical Biology*, Vol 128, pp. 297-315, 1987.

[5] D. Fogel, "Evolving behaviors in the iterated prisoner's dilemma", *Evolutionary Computation*, Vol. 1, pp. 77-97, 1987.

[6] D. Kraines and V. Kraines, "Evolution of learning among Pavlov strategies in a competitive environment with noise", *Journal of Conflict Resolution*, Vol. 39, Issue 3, pp. 439-466, 1995.

[7] M. Milinski, "Tit for tat in sticklebacks and the evolution of cooperation", *Nature*, Vol. 325, pp. 433-435, 1987.

[8] P. Molander, "The optimal level of generosity in a selfish, uncertain environment", *Journal of Conflict Resolution*, Vol. 31, Issue 4, pp. 692-724, 1987.

[9] M. A. Nowak and K. Sigmund, "Tit for tat in heterogeneous populations", *Nature*, Vol. 355, pp. 250-53, 1992.

[10] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos", *Nature*, Vol. 359, pp. 826-29, 1992.

[11] M. A. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game", *Nature*, Vol. 364, pp. 56-58, 1993.

[12] E. A. Stanley, D. Ashlock and L. Tesfatsion, "Iterated prisoner's dilemma with choice and refusal of partners", in *Artificial Life III*, Addison-Wesley, 1994.

[13] E. Tzafestas, "Toward adaptive cooperative behavior", Proceedings of the Simulation of Adaptive Behavior Conference, Paris, September 2000.

[14] E. Tzafestas, "Spatial games with adaptive tit-for-tats", Proceedings of the Simulation on Parallel Problem Solving from Nature (PPSN), Paris, September 2000.

[15] E. Tzafestas, "Attraction and cooperation in noisy environments: Individuals and groups", Proceedings ESHIA 2007 conference.