Title
A Memory-Based Account of the Spatial Prisoner's Dilemma

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Author
Cotla, Chenna Reddy

Publication Date
2011

Peer reviewed

A Memory-Based Account of the Spatial Prisoner's Dilemma

Chenna Reddy Cotla (CCOTLA@Gmu.Edu)
Department of Computational Social Science, 4400 University Drive,
Fairfax, VA 22030 USA

Abstract

After the seminal work of Nowak and May (1992), the Spatial Prisoner's Dilemma has become a common metaphor for studying the dynamics of cooperation in a spatially structured population. In contrast to the widely employed evolutionary model, which studies the dynamics of cooperation in a population of primitive players that lack memory, this paper examines the problem of cooperation in a population of memory-based players. Using computational simulations, it is shown that partial cooperation is maintained in a spatially structured population of players whose decision-making is effectuated by the adaptive nature of memory embodied in the ACT-R cognitive architecture (Anderson & Lebiere, 1998).

Key Words: Cooperation; Spatial Prisoner's Dilemma; ACT-R; Cognitive Modeling; Memory.

Introduction

Prisoner's Dilemma (PD) has long been used as a paradigm to study the problem of cooperation faced by unrelated individuals in the absence of central authority (Axelrod, 1984, Nowak and May, 1992). In its classical form, it characterizes a strategic situation between two players, in which both the players can be better off by mutual cooperation but their pursuits of rational self-interest leads to mutual defection and mutually inefficient outcome. Classical Game theory (Fudenberg & Tirole, 1991) provides a rigorous mathematical framework to represent and analyze the strategic interactions like PD and others among rational players. Game theory formalizes PD as a mixed-motive two-person game with two moves: Cooperate (C) and Defect (D) (Macy & Flache, 2002). The normal form representation of PD game given in Table 1 conveniently summarizes the moves available to the players and corresponding payoff each player would get for each of the four possible outcomes.

TABLE 1: Payoff Matrix of PD

		Player 2	
		C	D
Player 1	C	*R, R*	*S, T*
	D	*T, S*	*P, P*

In the above formalization, the conflict between social and selfish interests is captured with the payoff ordering: $T > R > P > S$. The payoffs: *R, P, T,* and *S*, characterize Reward for mutual cooperation, Punishment of mutual defection, Temptation to cheat, and Sucker payoff.

The main solution concept of game theory, the Nash equilibrium (Nash, 1950), predicts the socially inefficient mutual defection as the rational outcome for PD when it is played only once. However, many real world phenomena involve repeated PD type interactions. In the repeated version of the game, additional constraint of $2R > T + S$ is enforced on the payoff structure to prohibit the players from periodic alternations of cooperation and defection. In the finitely repeated version of the game, game theoretic solution is again mutual defection in each round. However, results from various behavioral experiments with repeated PD shows that partial cooperation is often sustained in the finitely repeated PD (Coleman, 1995). According to game theory, the socially efficient mutual cooperation is possible only in the infinitely repeated PD; one of the Folk theorems (Fudenberg & Tirole, 1991) of game theory predicts that in the infinitely repeated case, sustenance of mutual cooperation is possible that is Nash equilibrium. However, there are so many such possibilities and game theory cannot predict how players choose among them (Macy & Flache, 2002). In addition to the imprecise prediction problem, these forward-looking predictions make implausible assumptions about computational abilities of the players. Such complications led game theorists to explore cognitive and evolutionary alternatives to traditional game-theoretic solution concepts (Macy & Flache, 2002).

Since decision-making in a strategic situation is an important instance of cognitive behavior, more recently there have been several attempts to use general theories of cognition to study game-theoretic situations. Importantly, researchers from cognitive modeling had considerable success in matching human data obtained from various game contexts like Rock-Paper-Scissors (West & Lebiere, 2001) and repeated PD (Lebiere, Wallach, & West, 2000). By using the memory model offered by ACT-R cognitive architecture (Anderson & Lebiere, 1998), Lebiere et al. (2000) could reproduce important results like bimodal character of outcomes and the phenomenon of strategy shift observed in the experimental studies with human subjects by Rapaport, Guyer, and Gordon (1976). Another important achievement of this model is its success in explaining the rationale behind cooperative move by a player in the model without relying on any assumptions about altruism or fairness on the player's part.

Cognitive modelers argue that humans lack the required capabilities to be optimal players from the game theory perspective (West et al. 2006). According to them, human behavior in a strategic situation can be characterized as maximal rather than optimal. Instead of solving a game-theoretic problem on the grounds of rationality assumptions and then making the optimal move, maximal players use their cognitive mechanisms to process the past interactions, and respond with a move that exploits the perceived pattern of opponent's play. Such a maximal play, in which both the players are learning about weakness of opponent and are responding dynamically to exploit it in the best possible way, leads to the formation of a dynamically coupled system that has various interesting emergent regularities at the system level (West et al., 2006).

This paper extends the memory based model of repeated PD proposed in Lebiere et al. (2000) to the context of large population of interacting players. Consideration of large population of interacting players leads to further assumptions concerning the structure of the interaction. Mean-field approximation and rigid spatial structures are often considered as the two limiting cases of interaction topologies (Hauert, 2002). In the simplest mean-field approximation, each player interacts with everybody else with equal probability. The other extreme, the rigid spatial structure, represents the case where players are situated on a spatial structure like regular lattice and interact only with their local neighbors. The seminal work by Nowak and May (1992) has established that this crude approximation of interaction topologies observed in real world could be used to explain the emergence and maintenance of cooperation in a population of extremely simple players with no memory or any other cognitive capabilities. In this paper, the problem of cooperation is analyzed in a finite (more than two) population of so-called maximal players that are situated on a regular lattice.

Spatial Prisoner's Dilemma

Nowak and May (1992) used a spatial version of PD, which is commonly referred to as "Spatial Prisoner's Dilemma" (SPD), to analyze the emergence and sustenance of cooperation in a spatially structured population. Players in this model are extremely simple entities having no memory or any other cognitive capabilities. Only two basic strategies are available to players: always cooperate or always defect. Players are placed at each site of a square lattice and each of them plays PD with itself and its neighbors. In each time step or generation, every player obtains a score that is the sum of payoffs received from these interactions and each player imitates the strategy of highest scoring player in the neighborhood (including itself). In the case of a tie, a random highest scoring individual is imitated. Substantial results about emergence and maintenance of cooperation have been obtained from such a simple evolutionary spatial game. Importantly, for certain parameter values of the model, the fraction of cooperators always reaches the same proportion almost independently of the initial configuration like size of the lattice and the initial fraction of cooperators in the population. Figure 1 depicts an instance of the behavior of such a model over *2000* generations. Players are bound to the lattice sites of a 100×100 square lattice, and interact with their eight adjacent neighbors and also with themselves. It can be easily seen that the fraction of cooperators, f_C, fluctuates slightly around *0.318* and agrees

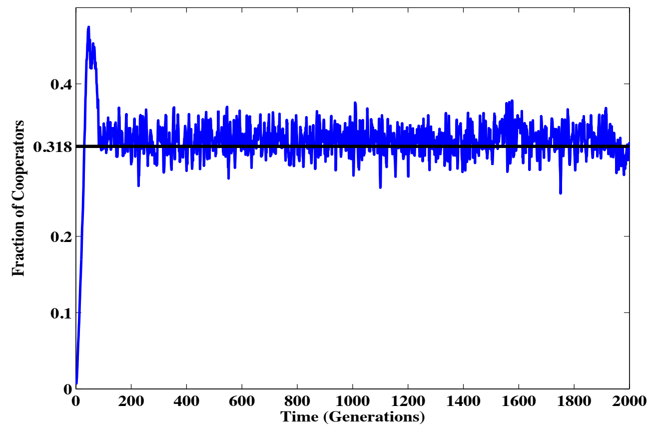

Figure 1: Fraction of cooperators in a simulation of evolutionary SPD on a square lattice with fixed boundary conditions. Synchronous updating and interactions with eight nearest neighbors are considered.

with the observation made in Nowak and May (1992). Note that the payoff matrix considered here has the structure given by $R = 1$, $T = b$, $S = 0$, and $P = 0$. The value of b is set at *1.9* for this particular instance, but the asymptotically stable value of proportion of cooperators will be the same for any initial conditions of the model with $1.8 < b < 2$. Due to their deterministic nature, the asymptotic cooperation levels achieved in these models can often be determined analytically using difference equations (Huberman & Glance, 1993). This most basic version of the model is augmented along many dimensions by considering the different definition of neighborhoods, probabilistic winning, spatial irregularity, and asynchronous updating (Nowak, Bonhoeffer, & May, 1994). Computational experiments with these variations confirmed the robustness of the claim that interaction with local neighbors in an evolutionary SPD can promote coexistence of cooperators and defectors in a population of memory-less players (Nowak et al., 1994).

Evolutionary spatial frameworks like the one considered in Nowak and May (1992) provides an apt template for modeling strategic interactions and explaining the maintenance of cooperation in a spatially structured population of simple biological or physical entities that lack memory or any other cognitive capabilities. However, straightforward adaptation of these results to explain the cooperative dynamics in a social system may not be appropriate. If the rationality assumption of classical game theory may be interpreted as one extreme, the overly simplistic characterization of players' behavior in the evolutionary spatial models may be interpreted as the opposite extreme. Human decision-making in repeated strategic interaction appears to be more sophisticated than the pure imitation in the evolutionary models. This paper attempts to strike a balance between the evolutionary and rational paradigms by explicitly taking into account the character of human memory in the analysis of the problem of cooperation in a spatially structured population. It is worth mentioning that

Qin, Chen, Zhao, & Shi (2008) provides an account of the role of memory in evolutionary SPD. However, the memory model adopted is quite primitive in its character and lacks many of the important characteristics of human memory that are observed in experimental studies. In contrast, the adaptive memory model considered in this paper is an integral part of the ACT-R cognitive architecture (Anderson & Lebiere, 1998) that is developed with the aim of replicating the behavior of actual human memory observed in extensive experimental studies.

Model

Similar to the evolutionary spatial framework considered in Nowak and May (1992), players in the current model are located on a regular lattice and play PD game with other players in their neighborhoods to receive the corresponding payoff. The principal divergence from the evolutionary SPD model is that the players in the current model use decision-making mechanism offered by ACT-R memory model to choose a strategy rather than imitating the strategy of the best scoring neighbor. The representations of declarative and procedural components of ACT-R memory model in the present model are largely derived from the memory-based account of two-person Prisoner's Dilemma game proposed in Lebiere et al. (2000). In this transition from the two-person game playing model to the spatial model, the game playing logic is kept intact; a player looks at its two possible moves, determine the most likely outcome given each move, and make the move associated with the best likely outcome. The following production rule captures the decision-making logic of a player in the model:

Spatial Prisoner's Dilemma

IF the goal is to play Spatial Prisoner's Dilemma
  and the most likely outcome of making move C is outcomeC
  and the most likely outcome of making move D is outcomeD
THEN make the move associated with the largest of outcomeC
    and outcomeD
    Note the actual outcome, and push new goal to make the
    next play

The number of possible outcomes for each player in a spatial game depends upon the notion of the neighborhood under consideration. If each player has *n* neighbors then there are in total 2^{n+1} possible outcomes for each player. For representational simplicity, a totalistic representation of the outcomes is adopted that is inspired by the totalistic approach adopted in Ishida and Mori (2005) to represent spatial strategies. The symbol *kC* is used to represent a scenario where *k* neighbors of a given player have chosen to cooperate and *n–k* neighbors have chosen to defect, where *n* is the size of the neighborhood and $0 \le k \le n$. With this notation of specifying the neighbors' moves, the outcome where the player under consideration has cooperated, and the configuration of neighbors' moves is *kC,* is denoted with *C-kC,* and the outcome where the player has chosen to defect for the same configuration of neighbors' moves with *D-kC*. In addition to simplifying the representational matters, such a totalistic representation of outcomes explicitly takes into account the spatial phenomenon that is an important characteristic of spatial games. In the remainder of this section representation of the outcomes as declarative chunks and the decision making process are illustrated when four orthogonal neighbors are considered for each player.

In a SPD on a square lattice with periodic boundary conditions, there are five possible outcomes for each possible move when four orthogonal neighbors are considered for each player. All these possible outcomes are represented as declarative chunks of type *outcome* with three slots: *p-move* that encodes player's action; *N-config* that encodes the choice of moves by the neighbors; and, *payoff* that encodes payoff received by the player for that particular outcome calculated from payoff matrix in Table 1. The ten chunks necessary to encode the possible outcomes for a given player in the model are given below:

(C-4C isa outcome p-move C N-config 4C payoff 4R)
(C-3C isa outcome p-move C N-config 3C payoff 3R+S)
(C-2C isa outcome p-move C N-config 2C payoff 2R+2S)
(C-1C isa outcome p-move C N-config 1C payoff R+3S)
(C-0C isa outcome p-move C N-config 0C payoff 4S)
(D-4C isa outcome p-move D N-config 4C payoff 4T)
(D-3C isa outcome p-move D N-config 3C payoff 3T+P)
(D-2C isa outcome p-move D N-config 2C payoff 2T+2P)
(D-1C isa outcome p-move D N-config 1C payoff T+3P)
(D-0C isa outcome p-move D N-config 0C payoff 4P)

For a given player, the first clause of the production will retrieve one of the five chunks associated with the player making move C, i.e. one of the five chunks: *C-4C, C-3C, C-2C, C-1C, C-0C,* and the retrieved chunk is denoted as *OutcomeC*, and the second clause will retrieve one of the five chunks associated with the player choosing to defect, i.e. one of the five chunks: *D-4C, D-3C, D-2C, D-1C, D-0C*, and the retrieved chunk is denoted as *OutcomeD*. The payoffs associated with these two outcomes, *OutcomeC* and *OutcomeD,* are compared and the *p-move* associated with the chunk with the highest payoff is taken. Similar to the model in Lebiere et al. (2000), by systematically selecting the move associated with the expected outcome that has the largest payoff, a player in this model attempts to maximize its own payoff. In this way, no assumptions about altruism or fairness are needed to explain cooperative move of a player.

The production rule retrieves the most likely outcome for each move by retrieving chunk with the highest activation for each move. The activation of declarative chunks is calculated using the following equation:

$$B = \ln\left[\sum_{i=1}^{k} t_i^{-d} + \frac{(n-k)\left(t_n^{1-d} - t_k^{1-d}\right)}{(1-d)(t_n - t_k)}\right] + N\left(0, \frac{\pi \,.\, s}{\sqrt{3}}\right)$$

The first part of the sum, known as base level activation, accounts for the adaptive nature of the human memory observed in various psychological experiments reported in

Anderson and Schooler (1991). t_i in the sum refers to the time since *jth* reference, n is the total number of references, and d is the forgetting rate. This computationally efficient approximation of the original formula proposed in Anderson and Schooler (1991) is due to Petrov (2006). Petrov has shown that by keeping the most recent k references, the base level activation can be approximated with great accuracy. In the actual implementation we used $k = 1$ for computational efficiency. The second part of the equation accounts for the stochasticity and is calculated as noise that is normally distributed with the mean of zero and the standard deviation determined by the activation noise parameter s (Lebiere et al., 2000). As in Lebiere et al. (2000), the same default values are considered for the forgetting rate, $d = 0.5$, and the activation noise parameter $s = 0.25$. The initial references of declarative chunks are uniformly distributed such that on average each chunk would get *100* references. It has been observed from simulations that results are qualitatively unchanged when we varied the number of initial references from *10* to *100*.

Simulation Results

The first simulation is carried out on a square lattice of the size 50×50 with periodic boundary conditions. Memory based players with the procedural and declarative memory components described in the previous section are placed at each lattice site and each of them interacts with its four orthogonal neighbors (self interaction is not considered). Standard PD payoff matrix (Axelrod, 1984) with $R = 3, S = 0, T = 5,$ and $P = 1$ is considered. In each generation (time step), all the players simultaneously make their choice of moves using the production rule, receive payoffs determined by the corresponding outcomes, and update their declarative memories. To characterize the macroscopic dynamics of the model, the fraction of cooperators (f_C) in the population at each generation is considered. Since the model involves stochastic elements, simulation output from a single realization may be misleading and some statistical treatment would be more appropriate. The simulation is carried out *30* times with a different random seed each time to ensure statistical independence across the runs. The model is considered to be asymptotically stable in a given run when the difference in the mean values of f_C over two consecutive windows of 10^4 generations is less than 10^{-3} in absolute value. After the model is considered as asymptotically stable, the mean value of f_C over next 10^4 generations is taken as the asymptotic f_C of the run. Figure 2 depicts the behavior of the mean value of f_C over the consecutive time windows of 10^4 generations in a sample run of the model that is run for a total of *400,000* generations. It can be easily seen that the magnitude of the slope of the curve is rapidly approaching zero with time, indicating that f_C is approaching an asymptotic value. In this run, after *120,000* generations, the model meets the asymptotic stability criteria and the asymptotic f_C is the mean fraction of cooperators over the next 10^4 generations, which is 0.3226 for this case. The *insert* in Figure 2 depicts frequencies of different values of f_C in the window of 10^4 generations after the model is considered stable. Almost normally distributed frequencies imply that f_C is fluctuating around the mean. The *95%* confidence interval obtained for asymptotic f_C from *30* different runs is [*0.3200*,*0.3228*]. However, the important insight here is that cooperators and defectors coexist in the model in such a way that the fraction of cooperators in the population is asymptotically stabilized. To better understand the microscopic dynamics of the model leading to the constant asymptotic cooperation levels, densities of the four following groups in the population are analyzed: individuals who cooperated in the previous and current generations (f_{CC}); individuals who cooperated in the previous generation but chose to defect in the current one (f_{CD}); individuals who defected in the previous generation but chose to cooperate in the current generation (f_{DC}); and, individuals who defected in both the previous and current generations (f_{DD}). Figure 3 depicts dynamics of these four densities for the same simulation run considered in Figure 2, and the emergent regularity is evident. All of these four densities stabilize after the system reaches asymptotic stability and more interestingly, f_{CD} and f_{DC} are varying in almost identical manner — the graphs of f_{CD} and f_{DC} are overlapping. This implies, an almost equal number of cooperators and defectors are changing their choice in such a way that f_C is asymptotically stable. In other words, the population is in a kind of dynamic equilibrium with cooperators and defectors coexisting in a chaotically shifting balance to keep f_C asymptotically stable.

In every generation, each of the maximal players considered in the model makes use of its memories of outcomes in the past generations to construct an expectation about the neighbors' moves conditional upon its own choice of move. Such an expectation is entirely experience based and is facilitated in a

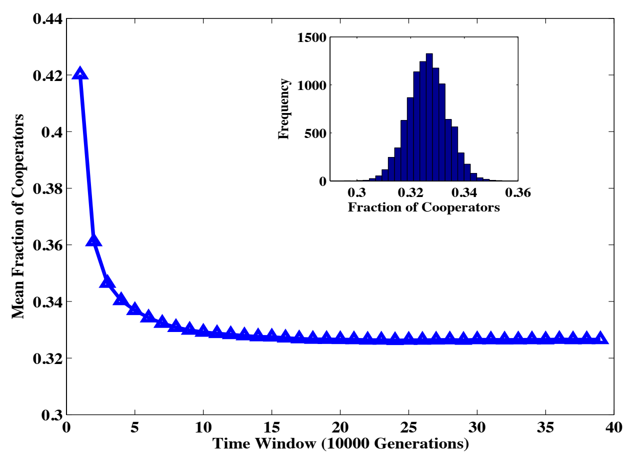

Figure 2: Mean f_C over time windows of 10^4 generations in a simulation run of *400,000* generations. The *insert* depicts the histogram of frequencies of different f_C values over the 10^4 generations after the model is considered as asymptotically stable.

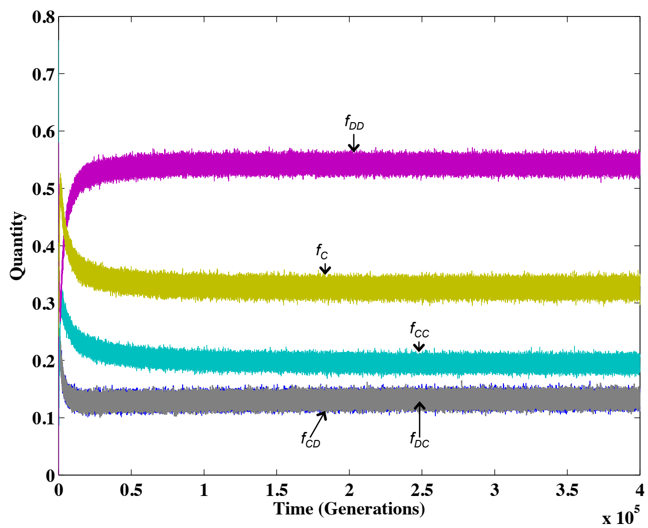

Figure 3: Time series dynamics of the densities f_{CC} , f_{CD} , f_{DC} , f_{DD} and f_C over *400,000* generations.

unique manner by the adaptive character of the memory of a player. The adaptive nature of the player's memory captures the observed pattern of neighbors' play using the past occurrences of the outcomes. The maximizing move taken by a player in a given generation is the best response to such an observed pattern. A given player's choice of a move in turn affects the adaptive memories of player's neighbors and their future moves. Such an interaction leads to the formation of so-called reciprocal causation (Clark, 1997) between a player and its neighborhood. Reciprocal causation often produces unexpected, macro-level regular patterns of behavior in a complex system of many interacting entities (West et al., 2000). In the present context, the coexistence of cooperators and defectors in a balance leading to the asymptotically stable fraction of cooperators can be understood as such a system level regularity in a spatially situated population of maximal players. It is often hard to express the behavior of a reciprocal causation system in terms of mathematical equations (Clark, 1997). Agent based models like the one discussed in this paper prove useful when the equation-based analytical treatment of the system characteristics is impractical due to the underlying complexity (Axtell, 1999).

Discussion

Agent based computational models facilitate explicit representation of the individual members of a system and the direct interactions among them (Axtell, 1999). Before deriving conclusions from these models, it is often important to be cautious about certain representational matters. Even though the system level regularities emerging out of complex microscopic interactions are interesting in the first simulation, further investigation is needed to make an overall claim about such an emergent regularity. One implicit but very crucial assumption in the previous simulation model is that activation scheme is synchronous. In other words, players are updated in unison at each time step as if a global clock governs them. In a social system, such an assumption about the existence of a global clock that synchronizes the players is often inappropriate (Huberman & Glance, 1993; Axtell, 2001). An asynchronous activation scheme is generally used to realistically model these systems where the members act at different and uncorrelated times (Huberman & Glance, 1993). Consideration of asynchronous updating resulted in strikingly different results for some multi agent models (Huberman & Glance, 1993). A generation in an asynchronous updating scheme consists of N micro time steps, where N is the number of players, and a single player is active during each micro time step. By active it is meant that the corresponding player chooses its move and receives a payoff based on the outcome. One important feature of this scheme is that often some players in the neighborhood of an active player have not made their decision in the current generation. In these cases, it is generally assumed that these players are still playing their choice of moves that are made during the most recent generation. Two variations of asynchronous updating are commonly used in agent-based simulations: Uniform activation (UA) and Random Activation (RA) (Axtell, 2001). In UA, each player is active exactly once every generation. To eliminate artifacts due to the spurious agent-agent correlations, the players are activated in a randomized order in each generation (Axtell, 2001). In RA, each player is active once on average in a generation; some players may be active more than once while the others may not be active at all in a given generation. For each of these two schemes, thirty statistically independent runs of the simulation described earlier were run and it has been observed in all these runs that model has reached asymptotic stability criterion discussed earlier with cooperators and defectors coexisting in a shifting balance. The asymptotic f_C values were also very close to the synchronous updating case. For UA, the resulting *95%* confidence interval for f_C is [*0.3211, 0.3265*], and for RA, it is [*0.3205, 0.3235*]. However, as mentioned before, the main emphasis of the current discussion is not about the quantity of asymptotic f_C but the emergent regularity developing out of complex microscopic dynamics.

Further investigations using computational simulations confirmed that results obtained in this framework are almost independent of the size of the lattice. Lattice size is varied for the first simulation from *20 × 20* to *400 × 400,* and it is observed that asymptotic value of f_C is almost independent of these parameters for a given neighborhood definition and updating scheme. Such independence may have important implications such as the existence of a universal constant governing the PD interactions among memory-based players on a square lattice (Huberman & Glance, 1993). The conclusion about maintenance of cooperation, that model reaches an asymptotically stable state where cooperators and defectors can coexist, remains valid when eight nearest neighbors are considered. These experiments confirm, on a more general level, the claim that cooperation can be maintained in a SPD of maximal players.

Recent behavioral experiments involving SPD have shown that humans do not unconditionally imitate the best scoring neighbor as assumed in many evolutionary game theoretic models (Traulsen et al., 2010). We believe that cognitive framework proposed in this paper to analyze spatial games may be an important candidate among the possible alternatives that can complement the evolutionary framework in understanding the dynamics of strategic puzzles in social sciences. In our future studies, we intend to validate the model output with the data from behavioral experiments on SPD. We also intend to study the effects of different payoff structures, and network topologies on the cooperative dynamics in the memory-based framework considered here.

Agent based models have gained significant appreciation in explaining the relation between macro level outcomes and micro level dynamics in complex social systems. However, often these models employ *ad hoc* behavioral specifications at the agent level that lack empirical underpinnings (Axtell, 2007). In contrast, the current model uses a behavioral specification derived from the memory model of a successful cognitive architecture that is developed from thorough experimental studies. Using empirically justified behavioral specifications instead of *ad hoc* formulations of behavior may render agent based models to account for social phenomena more convincingly.

Conclusion

This paper has investigated the role of the adaptive nature of human memory in the sustenance of cooperation in the context of Spatial Prisoner's Dilemma. Computational experiments showed that individually maximizing behavior facilitated by the memory mechanism implemented in the ACT-R cognitive architecture promotes the coexistence of cooperators and defectors in such a way that fraction of cooperators in the populations is asymptotically stable. Further investigations confirmed that such an emergent system level regularity remains effective for various definitions of neighborhoods and updating schemes. This work may be relevant in understanding the dynamics of cooperation in a social system where the memory processes that facilitate and constrain decision-making of individuals may not be ignored.

Acknowledgement

The author would like to thank Omar Guerrero, Timothy Gulden, William G. Kennedy, Sarah Wise and anonymous reviewers for their comments.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of the thought.* Mahwah, NJ: Erlbaum.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 49,* 7-15

Axelrod, R. (1984). *Evolution of cooperation*. New York: Basic Books.

Axtell, R. (1999). Why agents? On the varied motivations for agent computing in the social sciences. *Proceedings of the Workshop on Agent Simulation: Applications, Models, and Tools* (pp. 3-24).

Axtell, R. (2001). Effects of interaction topology and activation regime in several multi-agent systems. *Multi-Agent-Based Simulation*, 33–48.

Axtell, R. (2007). What economic agents do: How cognition and interaction lead to emergence and complexity. *Review of Austrian Economics, 20(2)*, 105-122.

Coleman, A. (1995). *Game theory and its applications.* Oxford: Butterworth-Heinemann.

Clark, A. (1997). *Being there*: *putting brain*, *body and world together again*. Cambridge, MA: The MIT Press

Fudenberg, D., & Tirole, J. (1991). *Game theory*. Cambridge, MA: The MIT Press.

Hauert, C. (2002). Effects of space in 2 x 2 games. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering, 12,* 1531-1548.

Huberman, B.A., & Glance, N. S. (1993). Evolutionary games and computer simulations. *PNAS, 90*, 7716.

Ishida, Y., & Mori, T. (2005). Spatial strategies in a generalized spatial prisoner's dilemma. *Artificial Life and Robotics, 9(3),* 139-143.

Lebiere, C., Wallach, D., & West, R. L. (2000). A memory based account of the prisoner's dilemma and other 2x2 games. *Proceedings of International Conference on Cognitive modeling* (pp. 185-193).

Macy, M. W., & Flache, A. (2002). Learning dynamics in social dilemmas. *PNAS, 99*, 7229-7236.

Nash, J. F. (1950). Equilibrium points in N-person games. *PNAS, 36*, 48-49.

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos, *Nature, 359*, 826–829.

Nowak, M. A., Bonhoeffer, S., & May, R. M. (1994). Spatial games and the maintenance of cooperation. *PNAS, 91*, 4877.

Petrov, A. A. (2006). Computationally efficient approximation of the base-level learning equation in ACT-R. *Proceedings of the seventh international conference on cognitive modeling* (pp. 391–392).

Qin, S. M., Chen, Y., Zhao, X., & Shi, J. (2008). Effect of memory on the prisoner's dilemma game in a square lattice. *Physical Review E*, *78(4),* 1-5.

Rapaport, A., Guyer, M., & Gordon, D. G. (1976). *The 2×2 game*. Ann Arbor, MI: University of Michigan Press.

Traulsen, A., Semmann, D., Sommerfeld, R., Krambeck, H., & Milinski, M. (2010). Human strategy updating in evolutionary games. *PNAS, 107(7),* 2962.

West, R. L. & Lebiere, C. (2001). Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Cognitive Systems Research, 1(4),* 221-239.

West, R. L., Lebiere, C., & Bothell, D. J. (2006). Cognitive architectures, game playing, and human evolution. In R. Sun (ed.), *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, 103-123. New York: Cambridge University Press.