Dynamics of Cooperation in Spatial Prisoner's Dilemma of Memory-Based Players

Chenna Reddy Cotla

Abstract In a population of extremely primitive players with no memory, interaction with local neighbors in a spatial array can promote the coexistence of cooperators and defectors, which is not possible in the well mixed case (Nowak, Bonhoeffer, and May, 1994). However, the applicability of this insight is unclear in the context of a social system where memory plays a significant role in the conscious decision-making of the members. In this paper, the problem of cooperation is analyzed in a population of players with the memory model embodied in the ACT-R cognitive architecture (Anderson and Lebiere, 1998). Using agent-based simulations, it is shown that in a population of memory-based agents, spatial structure supports higher levels of cooperation in comparison to the well mixed paradigm.

1 Introduction

The 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 [4, 16]. Even though, the PD offers an invaluable framework to study the problem of cooperation in the context of two self-regarding players, realistic investigation of cooperation problems at societal level involves a population of more than two players [18]. Consideration of a population of more than two interacting players leads to the 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 [9]. In the simplest mean-field approximation, each player interacts with every other player with equal probability. This represents the well-mixed scenario and here the interaction has no structure at all. The other extreme, the rigid spatial structure, represents the case where players are situated on a regular lattice and interact only with their local neighbors. The seminal work by

Department of Computational Social Science, George Mason University.
e-mail: ccotla@gmu.edu

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_16,

Nowak and May [16] has established that this crude approximation of interaction topologies observed in the real world could be used to explain the emergence and maintenance of cooperation in a population of extremely simple players with no memory. Nowak and May [16] have employed a spatial version of evolutionary PD, which is commonly referred to as the "Spatial Prisoner's Dilemma" (SPD), to show that cooperators and defectors can coexist in a chaotically shifting balance when the interactions are restricted to local neighbors on a square lattice. Computational experiments with numerous variations of such an evolutionary game confirmed the robustness of the claim that interaction with local neighbors in a spatial array can promote coexistence of cooperators and defectors in a population of memory-less players which is not possible in the well-mixed case [15] .

The evolutionary spatial framework considered in Nowak and May [16], and in Nowak et al. [15] 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. However, straightforward adaptation of these results to reason about the dynamics of cooperation in a social system may not be appropriate. Human decision-making in repeated strategic interaction appears to be more sophisticated than the pure imitation in the evolutionary models. Moreover, some researchers argue that the evolutionary template may not be an appropriate framework to study learning and adaptation processes at the cognitive level [3]. This paper studies the problem of cooperation in a spatial setting by explicitly taking into account the adaptive character of the human memory. By using the memory model embodied in the ACT-R cognitive architecture [1], Lebiere et al. [13] were able to reproduce important experimental observations in the context of the iterated PD. The current paper extends this successful two-person memory-based game playing model to the context of SPD and studies dynamics of cooperation in such a framework.

The paper is organized as follows: Section 2 presents the representational details of the decision-making model of players in the model. Section 3 presents details about computational simulations with SPD involving memory-based players described in Section 2 and corresponding results about the dynamics of cooperation in such a framework. A discussion about the decision-making process of players in the model and some perspectives for future research is given Section 4. Finally, concluding remarks are pointed out in Section 5.

2 Model

Similar to the evolutionary spatial framework considered in Nowak and May [16], 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. In ACT-R, higher level decision-making is embodied in the declarative memory of facts and the procedural memory of production rules []Taatgen2006. The production rules in the procedural memory are condition-action rules that encode the potential actions to be taken when certain conditions are met. The facts are encoded in declarative memory using items called *chunks*. Chunks encode knowledge as structured, schema-like configurations of labeled slots. Each chunk has a level of activation that depends on its previous usage, its relevance to the current context, and a noise component. ACT-R memory model retrieves the chunk with highest activation and applies the relevant production rule to achieve a goal [19].

The representations of a player's declarative and procedural memory components in the present model are largely derived from the memory based account of two-person Prisoner's Dilemma game proposed in Lebiere et al. [13] . In this transition from two-person game playing context to the spatial game, the procedural component of the player is kept intact: a player looks at its two possible moves, determines the most likely outcome given each move, and makes the move associated with the best likely outcome. This logic is captured in the following production rule:

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 larger of outcomeC and outcomeD
	Note the actual outcome and push a new goal to make the next play

The number of possible outcomes for each player in a spatial game depends upon the definition of the neighborhood under consideration. If each player has n neighbors then there are 2^{n+1} possible outcomes per player. For simplicity, a totalistic representation of the outcomes is adopted that is inspired by the totalistic approach used in Ishida and Mori [12] to represent spatial strategies. The symbol kC is used to represent a configuration 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 $0 \leq k \leq n$. Using this notation, 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 is denoted with D-kC. In addition to simplifying the representational matters, such a totalistic representation of outcomes explicitly takes into account the spatial phenomena that is an important characteristic of spatial games, in contrast to the well known SPD frameworks that use either very simple strategies (e.g. Nowak and May [16]) or use conventional two person strategies in repeated PD games which depend upon the past actions of a single opponent (e.g. Axelrod [4]). With this totalistic representation scheme, there are $(n+1)$ possible outcomes for each possible move when a neighborhood of size n is considered. All these possible outcomes are represented in a player's declarative memory as chunks of type *outcome* with three slots: *p-move* that encodes the 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 Reward (R), Temptation (T), Sucker (S), and Punishment (P) payoffs of the PD game. The $2(n+1)$ chunks necessary to encode all possible outcomes for a given player in the model are given below:

(C-nC isa outcome p-move C N-config nC payoff nR)
(C-$(n-1)$C isa outcome p-move C N-config $(n-1)$C payoff $(n-1)R+S$)
.
.
(C-kC isa outcome p-move C N-config kC payoff $kR+(n-k)S$)
.
.
(C-1C isa outcome p-move C N-config 1C payoff $R+(n-1)S$)
(C-0C isa outcome p-move C N-config 0C payoff nS)
(D-nC isa outcome p-move D N-config nC payoff nT)
(D-$(n-1)$C isa outcome p-move D N-config $(n-1)$C payoff $(n-1)T+P$)
.
.
(D-kC isa outcome p-move D N-config kC payoff $kT+(n-k)P$)
.
.
(D-1C isa outcome p-move D N-config 1C payoff $T+(n-1)P$)
(D-0C isa outcome p-move D N-config 0C payoff nP)

For a given player, the first clause of the production rule will retrieve one of the $(n+1)$ chunks associated with the player making the move C, and the retrieved chunk is denoted as *OutcomeC*, and the second clause will retrieve one of the $(n+1)$ chunks associated with the player choosing to defect, 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.

The production rule retrieves the most likely outcome for each move by retrieving the outcome chunk with the highest activation. For simplicity, relevance factor is not considered in calculating the activation of a chunk in the current model. The activations of a declarative chunks are calculated using the following equation:

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

The first part of the sum accounts for the adaptive nature of the human memory observed in various psychological experiments reported in [2]. This quantity, known as "Base level activation", increases with each reference and decreases with time justifying the power of learning and forgetting [2]. t_j 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 [2] is due to Petrov [17]. 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 [13]. The same default values are considered as in Lebiere et al. [13] for the forgetting rate of $d = 0.5$, and the activation noise parameter of $s = 0.25$. The initial references of declarative chunks are uniformly distributed such that on average each chunk would get 100 references. It has been observed from computational experiments that results are qualitatively unchanged when we varied the number of initial references from 10 to 100.

3 Simulation Results

The first computational experiment 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 von Neumann Neighbors (self interaction is not considered here). The standard PD payoff matrix [4] 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. As such the underlying dynamics of the model are synchronous. 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. In each run, it was observed that f_C in the model asymptotically stabilizes and fluctuates around a constant after some number of generations. 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 magnitude []gomez2007. 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 1 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 slope of the curve is continuously diminishing in magnitude indicating that mean f_C is approaching an asymptotic value. In this run, after 120,000 generations, the model

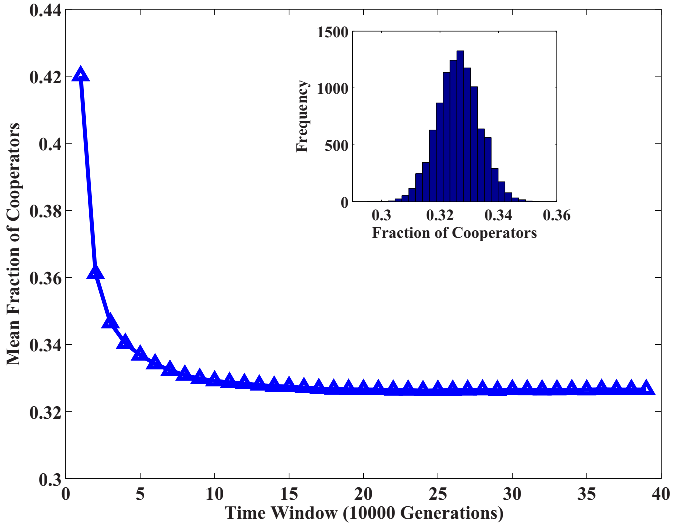

Fig. 1 Mean f_C over time windows of 10^4 generations in a simulation run of $400,000$ generations. The *insert* depicts the histogram of frequencies of different f_C values over the 10^4 generations after the model is considered as asymptotically stable.

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 1 depicts frequencies of different values of f_C in the window of 10^4 generations after the model is considered as asymptotically stable. Almost normally distributed frequencies suggest that f_C is fluctuating around the mean. The 95% confidence interval obtained for asymptotic f_C from 30 different runs is obtained as $[0.3220, 0.3228]$. Such a narrow confidence interval indicates that f_C is fluctuating around a constant. Also the experiments were repeated for lattice sizes from 20×20 to 400×400 and it was observed that asymptotic cooperation levels are almost independent of lattice size. These emergent stable cooperation levels independent of the initial configuration of the model are very interesting.

The second experiment considered the mean-field interaction scenario for a population size of 2500 players so that the effect of spatial structure on cooperation levels can be evaluated. Since the interactions in well-mixed case are bilateral, game playing model proposed in lebiere et al. [13] is directly used. A generation in this case consists of $N/2$ micro time steps, where N is the size of the population. In each micro time step, two randomly selected distinct players play a bilateral PD game. Thirty statistically independent runs of this well mixed scenario were carried out with stopping times determined by the asymptotic stability criteria discussed earlier. The 95% confidence interval of the asymptotic f_C was calculated as [0.2446 0.2448]. Here too the asymptotic f_C values were almost independent of size of the population. Also, this is in contrast with evolutionary PD where cooperation is not possible in mean-field interaction case. Comparison between cooperation lev-

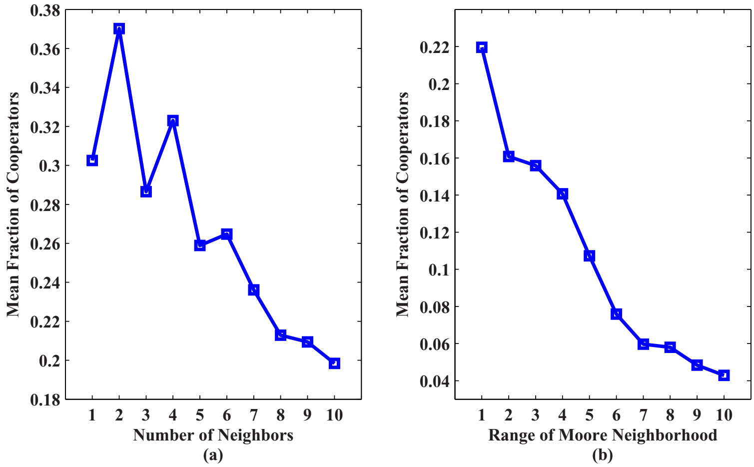

Fig. 2 (a). Mean asymptotic f_C of 30 independent runs of the model for different sizes of neighborhood. Players are located on a 50×50 square lattice. (b) Mean asymptotic f_C of 30 independent runs of the model for different ranges of Moore neighborhood. Players are located on a 50×50 square lattice.

els in the spatial context involving von Neumann neighborhoods and well-mixed case suggests that spatial structure can support higher levels of cooperation than that of well-mixed case in a population of memory-based players. However, further experiments with different notions and sizes of neighborhoods suggested that spatial structure does not always lead to the higher asymptotic cooperation levels than that of the well-mixed scenario. For example, the mean value of asymptotic f_C obtained from 30 independent runs of the spatial model considering Moore neighborhood was 0.2196, which is less than the mean asymptotic f_C observed in well-mixed scenario. To further characterize the relation between the neighborhood size and asymptotic cooperation levels, two more computational experiments were carried out. The first of these explored the effect of number of neighbors on the asymptotic cooperation levels. In this experiment, the spatial game is run with neighborhood sizes from 1 to 10. For a given neighborhood size, the corresponding number of distinct neighbors are randomly initialized for each player. Figure 2(a) plots the mean of the asymptotic f_C of the 30 independent runs against the neighborhood size from 1 to 10 for SPD with memory-based players located on a 50×50 lattice. Even though, neighborhood sizes between 1-6 support higher asymptotic cooperation levels than well-mixed case, neighborhood sizes bigger than 6 support asymptotic cooperation levels smaller than that of the well-mixed case. The other experiment explored the effect of the range of Moore neighborhood on the asymptotic cooperation level in the model. Figure 2(b) plots the mean asymptotic f_C observed in 30 independent runs against the range of Moore neighborhood for SPD with memory-based players located on a 50×50 lattice. Again, it can be observed that asymptotic cooperation levels decrease with larger neighborhood sizes and they are all smaller than the asymptotic cooperation levels observed in the well-mixed case. These simulation results suggest that in a SPD with memory-based players as described in this paper, neighborhoods larger than 6 in size lead to lower asymptotic cooperation levels compared to the well-mixed case. Thus in a population of memory-based players spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes. Furthermore, it is interesting to note that for all these experiments, the asymptotic cooperation levels were observed to be almost independent of the size of the population.

4 Discussion

In each generation, players in the current model make a decision to cooperate or defect through the use of accumulated experience. The particular decision-making model used in this paper is referred to as Instance based decision-making model in the cognitive science literature [8]. This model of decision making is based on two principles: a. storage of instances of experiences and b. decision making logic that involves selecting the most promising action by reviewing past experiences [8]. In the current model instances are encoded as the outcome chunks in the declarative memory and decision-making logic is encoded as the production rule in the procedural memory of a player. Since these two principles are implemented using memory model of a validated cognitive architecture, in this case ACT-R, this model is also cognitively plausible. Coming back to the decision making process, a player considered in the model makes use of its memories of outcome instances 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 unique manner by the adaptive memory of a player. In this way, the adaptive nature of the memory of a player 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 a best response to such an observed pattern of the neighbors' play. A given player's choice of a move in turn affects the adaptive memories of the player's neighbors and their future moves. Due to the presence of such causal loops that couple players and their neighborhoods analysis by reduction may not be practical in this model. Agent based models are very useful to model and analyze the emergent characteristics arising out of interactions among the elements of a complex adaptive system like the one discussed here [14]. Agent based models facilitate explicit representation of individual members of a complex adaptive system and the direct interactions among them to systematically analyze various emergent aspects [5]. In the current model, attainment of constant asymptotic cooperation levels for a given neighborhood independent of initial configuration of the model can be considered as such a system level emergent property. Before deriving conclusions from agent models, it is often important to be cautious about certain representational matters. Importantly, the consideration of different updating schemes proved to elicit significantly different emergent behaviors from certain agent based models [6, 11].

To examine the effect of different updating schemes, multiple runs of the previous experiments were carried out for each of the two asynchronous updating schemes of Random Activation(RA) and Uniform Activation (UA) [6]. It was observed that the previous conclusion that spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes remains valid for these asynchronous updating schemes too.

Behavioral experiments on spatial games have been very recent. Results from these recent experiments call for alternative efforts are needed to explore the mechanisms underlying the dynamics of decision-making in these games [10]. Especially, experiments on spatial Prisoner's dilemma with human subjects have established that strategy change of a player in these experiments is different from unconditional imitation of the best scoring neighbor as assumed in many evolutionary game theoretical models [20]. 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. Furthermore, it is quite straightforward to extend this model to study other social dilemmas by changing the relevant payoff quantities in the model. In our future studies we intend to extend the model considering various other social dilemmas, interaction topologies, and validate the model output with the experimental data on spatial games.

5 Conclusion

This paper has investigated the effect of the adaptive nature of human memory on the dynamics of cooperation in Spatial Prisoner's Dilemma. Computational experiments showed that fraction of cooperators in such a framework stabilizes around a constant almost independently of initial configuration of the model for a given definition of neighborhood and updating scheme. Furthermore, it is shown from the simulation results that spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes in a population of memory-based players. 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.

References

1. Anderson, J. and Lebiere, C. (1998). *The atomic components of thought*. Lawrence Erlbaum Associates.
2. Anderson, J. and Schooler, L. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6):396.
3. Aunger, R. (2001). *Darwinizing culture: The status of memetics as a science*. Oxford University Press, USA.

4. Axelrod, R. (1984). *The evolution of cooperation*. Basic Brook, New York.
5. Axtell, R. (2000). Why agents? on the varied motivation for agent computing in the social sciences. *Brookings Institution CSED Technical Report*.
6. Axtell, R. (2001). Effects of interaction topology and activation regime in several multi-agent systems. *Multi-Agent-Based Simulation*, pages 33–48.
7. Gómez-Gardenes, J., Campillo, M., Floría, L., and Moreno, Y. (2007). Dynamical organization of cooperation in complex topologies. *Physical Review Letters*, 98(10):1–4.
8. Gonzalez, C. and Lebiere, C. (2005). Instance-based cognitive models of decision making. In D., Z. and Courakis, A., editors, *Transfer of Knowledge in Economic Decision Making*. New York: Palgrave McMillan.
9. Hauert, C. (2002). Effects of space in 2×2 games. *Int. J. Bifurcat. Chaos*, 12:1531–1548.
10. Helbing, D. and Yu, W. (2010). The future of social experimenting. *Proceedings of the National Academy of Sciences*, 107(12):5265–5266.
11. Huberman, B. and Glance, N. (1993). Evolutionary games and computer simulations. *Proceedings of the national academy of sciences of the United States of America*, 90(16):7716.
12. Ishida, Y. and Mori, T. (2005). Spatial strategies in a generalized spatial prisoner?s dilemma. *Artificial Life and Robotics*, 9(3):139–143.
13. Lebiere, C., Wallach, D., and West, R. L. (2000). A memory-based account of the prisoner's dilemma and other 2×2 games. In *Proceedings of the 3rd International Conference on Cognitive Modeling, Groningen, Netherlands*, pages 185–193.
14. Miller, J. and Page, S. (2007). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press.
15. Nowak, M., Bonhoeffer, S., and May, R. (1994). Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11):4877.
16. Nowak, M. and May, R. (1992). Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829.
17. Petrov, A. (2006). Computationally efficient approximation of the base-level learning equation in act-r. In *Proceedings of the Seventh International Conference on Cognitive Modeling*, pages 391–392.
18. Szilagyi, M. (2003). An investigation of n-person prisoners' dilemmas. *Complex Systems*, 14(2):155–174.
19. Taatgen, N., Lebiere, C., and Anderson, J. (2006). Modeling paradigms in act-r. In Sun, R., editor, *Cognition and multi-agent interaction: From cognitive modeling to social simulation*, pages 29–52. Cambridge University Press.
20. Traulsen, A., Semmann, D., Sommerfeld, R., Krambeck, H., and Milinski, M. (2010). Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences*, 107(7):2962.