Emergence of Norms in a Society of Heterogeneous Agents Influenced by the Rules of Cellular Automata Techniques

P.Chakrabarti[1] , J.K. Basu[2]
[1]Sir Padampat Singhania University, Udaipur-313601,Rajasthan, India
[2]Heritage Institute of Technology, Kokata-700107 , West Bengal , India
Email_id: prasun9999@rediffmail.com , basu.jayanta@yahoo.co.in

Abstract- This paper deals with study Emergence of Norms in a Society of Heterogeneous Agents Influenced by the Rules of Cellular Automata Techniques.To study the phenomenon of emergence of social norms, we have assumed that the interactions between the agents are private, i.e.; not observable to the other agents not involved in the interactions. We consider a population of agents, where, in each interaction each agent is paired with another agent selected randomly from its neighborhood or from the population in a non-uniform manner. Each agent is learning concurrently over repeated interactions with selected opponents from the society. An agent learns a policy to play the game from repeated interactions with multiple agents. We are particularly interested in finding out if the entire population learns to converge to a consistent norm when multiple action combinations yield the same optimal payoff. In addition to this, we also would like to explore the effects of heterogeneous populations where different agents may be using cellular automata techniques.

I . INTRODUCTION

Our social learning framework considers a potentially large population of learning agents. At each time step, however, each agent interacts with a single opponent agent, chosen from the population and the opponent changes at each interaction. The payoff received by an agent for a time step depends only on this interaction as is the case when two agents are learning to play a game. The specific social learning situation that we consider is that of learning "rules of the road" so that the driver could decide which side he has to take to drive the car. Each interaction between two drivers can be modeled by 2-person 2-action stage game. When two cars arrive at an intersection, a driver will have another car sometimes on its left and sometimes on its right. These two situations can be mapped to two different roles an agent can perform: playing as a row and column player respectively. As a consequence each agent has two private matrices, one when it plays as row player and the other for its role as a column player. The agents have perfect but incomplete information: the identity and the payoff of their opponents are not known to them but they can observe the opponents' actions.

Norms or conventions are key influences on social behavior of humans. Conformity to norms reduces social frictions, relieves cognitive load on humans, and facilitates coordination. "Everyone conforms, everyone expects others to conform, and everyone has good reason to conform because conforming is in each person's best interest when everyone else plans to conform" [1]. Conventions in human societies range from fashions to tipping, driving etiquette to interaction protocols. Norms are ingrained in our social life and play a pivotal role in all kinds of business, political, social, and personal choices and interactions. They are self-enforcing: "A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way" [2]. Effective norms, emerging from sustained individual interactions over time, can complement societal rules and significantly enhance performance of individual agents and agent societies. We have used a model that supports the emergence of social norms via learning from interaction experiences [3]. Each interaction is framed as a stage game. Interactions between agents can be formulated as a stage game with simultaneous moves made by the players. Such stage games often have multiple equilibrium, which makes the coordination uncertain. An agent learns a policy to play the game from repeated interactions with multiple agents. We are particularly interested in finding out if the entire population learns to converge to a consistent norm when multiple action combinations yield the same optimal payoff. Here, we explore the effects of homogeneous populations where different agents will be using same learning algorithm in different bi-matrix games.

II. RELATED WORKS

The need for effective norms to control agent behaviors is well-recognized in multiagent societies

[4]. In particular, norms are key to the efficient functioning of electronic institutions. Most of the work in multiagent systems on norms, however, has centered on logic or rule-based specification and enforcement of norms. Similar to these research, the work on normative, game-theoretic approach to norm derivation and enforcement also assumes centralized authority and knowledge, as well as system level goals[5]. While norms can be established by centralized dictat, a number of real-life norms evolve in a bottom-up manner, via "the gradual accretion of precedent" [6]. We find very little work in multiagent systems on the distributed emergence of social norms. We believe that this is an important niche research area and that effective techniques for distributed norm emergence based on local interactions and utilities can bolster the performance of open multiagent systems. We focus on the importance for electronic agents solving a social dilemma efficiently by quickly adopting a norm. Centralized social laws and norms are not sufficient, in general, to resolve all agent conflicts and ensure smooth coordination. The gradual emergence of norms from individual learning can facilitate coordination in such situations and make individuals and societies more efficient. In one of the formulations, norms evolve as agents learn from their interactions with other agents in the society using multiagent reinforcement learning algorithms[7],[8]. Most multiagent reinforcement learning literature involves two agents iteratively playing a stage game and the goal is to learn policies to reach preferred equilibrium[9]. Another line of research considers a large population of agents learning to play a cooperative game where the reward of each individual agent depends on the joint action of all the agents in the population[10]. The goal of the learning agent is to maximize an objective function for the entire population, the world utility. The social learning framework we use to study norm emergence in a population is somewhat different from both of these lines of research. We are considering a potentially large population of learning agents. At each time step, however, each agent interacts with a single agent, chosen at random, from the population. The payoff received by an agent for a time step depends only on this interaction as is the case when two agents are learning to play a game. In the two-agent case, a learner can adapt and respond to the opponent's policy. In another work; each interaction is framed as a stage game. An agent learns a policy to play the game from repeated interactions with multiple agents. We are particularly interested in finding out if the entire population learns to converge to a consistent norm when multiple action combinations yield the same optimal payoff. In this extension, we explore the effects of heterogeneous populations where different agents may be using different learning algorithms. They investigate norm emergence when an agent is more likely to interact with other agents nearby it [11] In our framework, however, the opponent changes at each interaction. It is not clear a priori if the learners will converge to useful policies in this situation. A model was proposed that supports the emergence of social norms via learning from interaction experiences. In that model, individual agents repeatedly interact with other agents in the society over instances of a given scenario. Each interaction is framed as a stage game. An agent learns its policy to play the game over repeated interactions with multiple agents. The key research question was to find out if the entire population learns to converge to a consistent norm. In addition to studying such emergence of social norms among homogeneous learners via social learning, they studied the effects of heterogeneous learners, population size, multiple social groups, The goal of the learning agent is to maximize an objective function for the entire population, the world utility. This framework considers a potentially large population of learning agents. At each time step, however, each agent interacts with a single opponent agent chosen from the population, and the opponent changes at each interaction. The payoff received by an agent for a time step depends only on this interaction as is the case when two agents are learning to play a game. In the two-agent case, a learner can adapt and respond to the opponent's policy. In this framework, however, the opponent changes at each interaction. It is not clear a priori if the learners will converge to useful policies in this situation A model was proposed that supports the emergence of social norms via learning from interaction experiences. In that model, individual agents repeatedly interact with other agents in the society over instances of a given scenario. Each interaction is framed as a stage game. An agent learns its policy to play the game over repeated interactions with multiple agents. The key research question was to find out if the entire population learns to converge to a consistent norm. In addition to studying such emergence of social norms among homogeneous learners via social learning, they studied the effects of heterogeneous learners, population size, multiple social groups, etc.

III. SOCIAL LEARNING APPROACH

While implementing the situation of learning in society we have considered the learning rules of the road where the driver decides which side of the road to take to drive the vehicle. Each interaction between two drivers can be modeled by 2-person 2-action stage game. There are various topologies used for the depiction of a artificial agent society.

Here we have only used a single topology for our experiment. We have taken a toroidal grid structure.

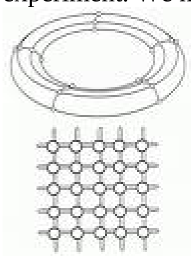

Fig. 1 Two forms of Toroidal Grid

There are two ways of looking at toroidal grids. In one way, the purpose is to ensure that each cell has the same number of neighbors. In the other, the grid is seen as repeating to infinity. The difference is that when computing neighborhood, the former way of looking at things might have the attitude that the set of neighbors of any point, whilst having the same number of elements, should never contain repeated cells.

For example, consider the 3 by 2 grid below:

```
+-----+-----+-----+
|     |     |     |
|  A  |  B  |  C  |
|     |     |     |
+-----+-----+-----+
|     |     |     |
|  D  |  E  |  F  |
|     |     |     |
+-----+-----+-----+
```

Using a distance d = 1, the neighbors of A are B, C and D, with D repeated since it is both above and below A if we are wrapping around. If we use a toroidal grid because we want each cell to have the same number of neighbors, then we could define things such that neighbors of a cell cannot be repeated so each cell would have 3 neighbors rather than 4 in the grid above. The agents are placed on the nodes of the grid. We have taken this structure as this structure enables each and every agent to interact with each other.

We consider the agents are distributed over space where each agent is located at a grid point. Each agent has a fixed location on the grid and hence a static set of neighbors. In our experiment we have considered two ways in which the agents are selected for interaction.

a) Uniform Selection: - Here any agent in any position of the grid can interact with any agent present on the grid irrespective of its neighborhood. The agents are randomly selected from the grid.
b) Non-uniform: - Here only those agents can interact with another only if it is within some neighborhood distance.

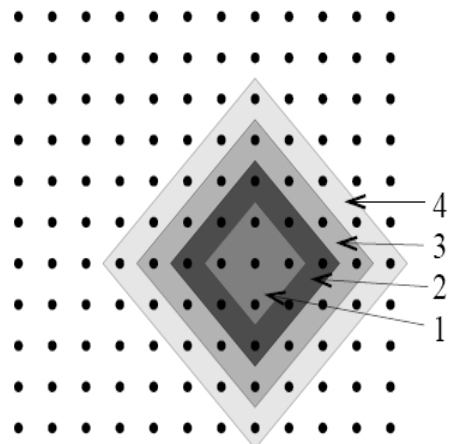

Fig. 2 : Agents located on a grid and allowed to interact only in a limited neighborhood.

The neighborhood of an agent is composed of all agents within a distance D of its grid location. We have used the Manhattan distance metric, i.e., $|x1 - x2| + |y1 - y2|$ is the distance between grid locations (x1, y1) and (x2, y2). Different D values are used to represent different neighborhood sizes. In each time period, each agent interacts with another agent in the society.

The non-uniform selection of opponents can be done in two modes:

a) Agents are chosen randomly from anywhere within the neighborhood distance D.
b) Agents are chosen from the neighborhood if they have the higher probability of being closer to each other within the neighborhood distance D.

IV. EXPERIMENTAL METHODOLOGY

In this experiment with a society of N agents have been placed in a sqrt (N) x sqrt(N) grid. Here we have used 225 agents placed on a 15 by 15 grid. Our main objective was to simulate an artificial heterogeneous society of agents using different learning algorithms, which will interact with each other using a bi-matrix game. The selection of the learning algorithms by the agents in the bi-matrix game will be done by the use of cellular automata technique. At the initialization phase the agents in the grid have selected the learning algorithms arbitrarily to interact each other. Then using cellular automata techniques the selections of the learning algorithms

have been optimized. The process of selection has been made in the following way.
Each time the agents look into their neighbours residing at left, right, top and bottom and determine their learning algorithm according the algorithm which is used by the most number of neighbouring agents. This way they determine their algorithm and updates their state-action table to get the optimized reward from the environment.
The generalized payoff scenarios for the Prisoner's Dilemma bi-matrix game are as given below.

		P1	
		confess	don't
P2	confess	3,3	0,5
	don't	5,0	1,1

P1,P2 – Player1,Player2

Table 1. Generalised Payoff in a Prisoner's Dilemma game

We have used Sarsa, Q-learning and WoLF PHC(Policy Hill Climbing) temporal difference learning algorithm for the agents in this experiment.
Cellular automata technique has been used considering neighbors are a selection of cells relative to the specified cell, and do not change (though the cell itself may be in its neighborhood, it is not usually considered a neighbor). Every cell has the same rule for updating, based on the values in this neighborhood. Each time the rules are applied to the whole grid a new generation is created. In our experiment we have used the Conway's game of life. Here the toroidal grid represents the universe, where each node may or may not contain agents. von Neumann's model takes into consideration the grid where any agent can have at most four neighbors. Conway's game of life model takes into account all 8 neighbors for an agent, considering the diagonal elements also.
In this experiment Conway's "game of Life" [12,13] has been implemented, with four neighbors (East, West, North and South) for simplicity.The rules for Conway's Game of Life are:

a) If any dead agent contains three living neighbors then it will become alive in the next generation.
b) If any dead or living agent contains two neighbors then it will remain in the same state in the next generation, irrespective of its current state.
c) Any agent having any other number of agents will become dead irrespective of its current state in the next generation.

V. RESULTS AND DISCUSSION

We implemented the learning algorithms over sparsely located agents over a toroidal grid. We implemented the program generating the average payoffs over iterations.
The agents were selected by uniform mode of selection without considering the neighborhood distance between them. It is found that the emergence of norm using cellular automata technique in heterogeneous society is more rapid than that of without using cellular automata .The experiment involves agents which are sparsely distributed over the grid. The number of agents were taken 25% at first and then increased by 25% until there were 100% agents located over the grid.

Implementation of Cellular Automata (Conway's Game of Life)

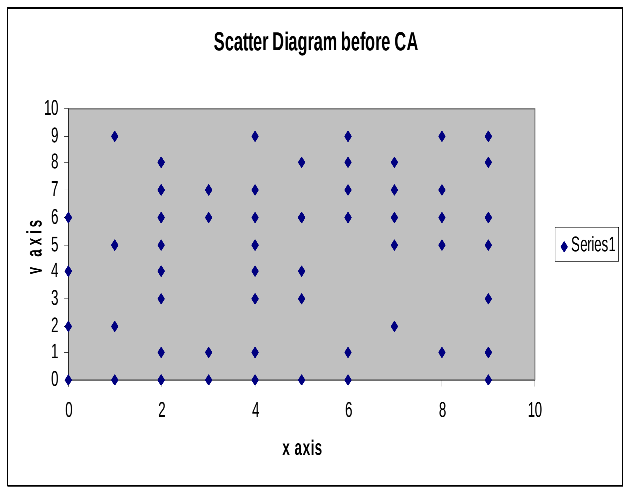

Fig. 3: Scatter diagram for 50% agents, before the implementation of Cellular Automata

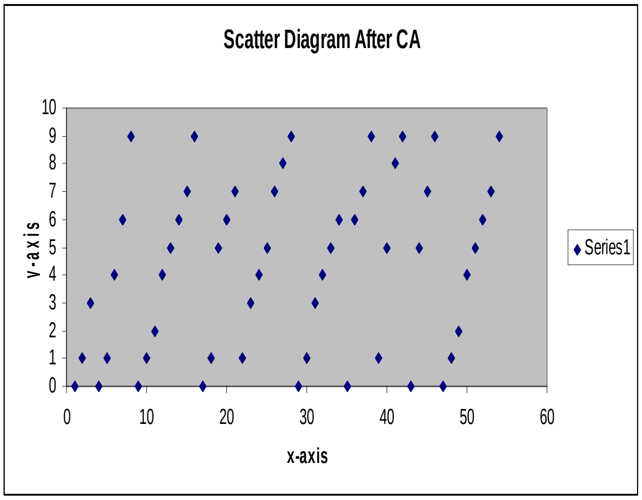

Fig. 4 : Scatter diagram for 50% agents, after the implementation of Cellular Automata

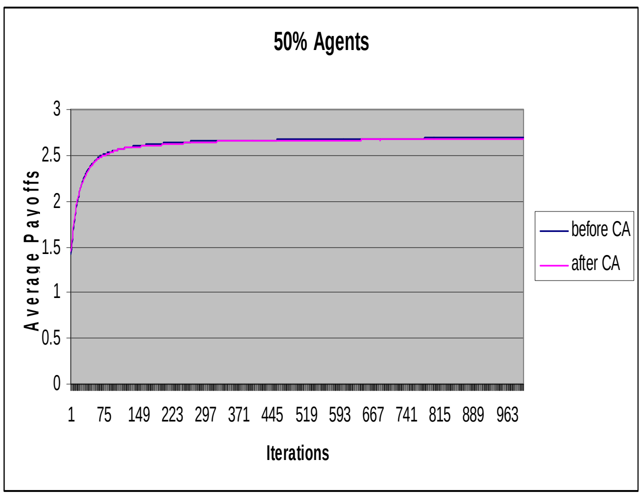

Fig. 5 : Rate of emergence of norm before implementation of CA and after implementation for 50% agent population.

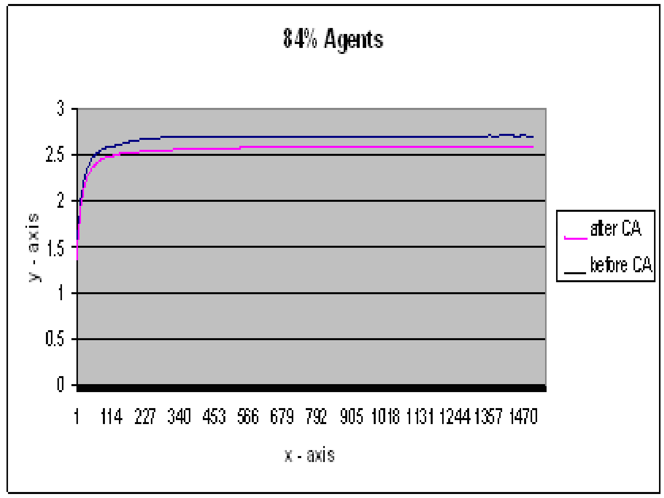

Fig. 6: Norm emergence for 84% agent population

VI. CONCLUSION AND FUTURE WORK

From the result this is evident that with the implementation of cellular automata we not only have achieved better social payoff but also the convergence to emerge a norm has been quicker and smoother. At last we have tested the learning algorithms using Coordination bimatrix game and noticed the significant difference in output. Though using Coordination game we have obtained better convergence but poor payoff. It is observed that for 50% of sparse agent population, a cellular automaton doesn't affect the density of population. This again doesn't put an effect over the rate of norm emergence. The rate of norm emergence, for the population before the implementation of Conway's game of life and after that, remains more or less the same. When the agent population is increased the number of agents after the implementation of the cellular automata decreases and becomes zero when the number of agents becomes 100%. The rate of emergence of norm also increases substantially increases with the implementation of Cellular Automata.

REFERENCES

[1] D. Lewis, "Convention", A Philosophical Study, Harvard University Press, 1969.
[2] R. Axelrod, "The complexity of cooperation", Agent-based models of conflict and cooperation, Princeton University Press, Princeton, NJ, 1997.
[3] S. Sen and S. Airiau, "Emergence of norms through social learning", In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI'07), pages 1507–1512, January 2007.
[4] G. Boella and L. van der Torre, "Norm governed multiagent systems", The delegation of control to autonomous agents, In Proceedings of IEEE/WIC IAT Conference, pages 329–335. IEEE Press, 2003.
[5] Boella, G. and Lesmo, L. "A game theoretic approach to norms", Cognitive Science Quarterly, 2002.
[6] Peyton H. Young, "The economics of convention", The Jounal of Economic Perspectives, 10(2):105–122, Spring 1996.
[7] Liviu Panait and Sean Luke, "Cooperative multi-agent learning", The state of the art. Autonomous, Agents and Multi-Agent Systems, 11(3):387–434, 2005.
[8] K. Tuyls and A. Now´e, " Evolutionary game theory and multi-agent reinforcement learning", The Knowledge Engineering Review, 20(1):63–90, March 2006
[9] Rob Powers and Yoav Shoham, "New criteria and a new algorithm for learning in multiagent systems", In Proceedings of NIPS, 2005.

[10] Kagan Tumer and David H Wolpert, “Collective intelligence and Braess’ paradox”, In Proceedings of the Seventeenth National Conference on Artificial Intelligence, pages 104–109, Menlo Park, CA, 2000. AAAI Press.
and Braess’ paradox”, In Proceedings of the Seventeenth National Conference on Artificial Intelligence, pages 104–109, Menlo Park, CA, 2000. AAAI Press.
[11] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained interaction in diverse societies. Proc of 7^{th} Int. conf. on Autonomous Agents and Multiagent Systems(AAMAS 2008) pp. 779-786.

[12] Niloy Ganguly, Biplab K Sikdar, Andreas Deutsch, Geo_rey Canright, P Pal Chaudhuri, “ A Survey on Cellular Automata”.

[13] Harald Niesche, “Introduction to Cellular Automata”, “Organic Computing”, Seminer, SS2006