International Journal of Machine Intelligence, ISSN: 0975–2927, Volume 2, Issue 1, 2010, pp-93-100

Intelligent means of analyzing norm emergence in a homogeneous society of both biased and unbiased agents in the light of different bi-matrix games

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

Abstract- This paper deals with study of the evolution of social conventions or norms that selects one equilibrium over others based on repeated distributed interactions between agents in a society. 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 different learning algorithms.
Keywords- social conventions , agent , optimal payoff , learning algorithms

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. LITERATURE SURVEY

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.

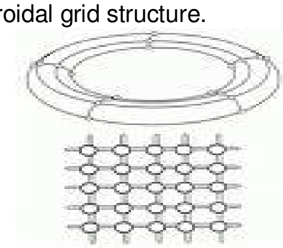

Fig. 1 Two forms of Toroidal Grid

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.

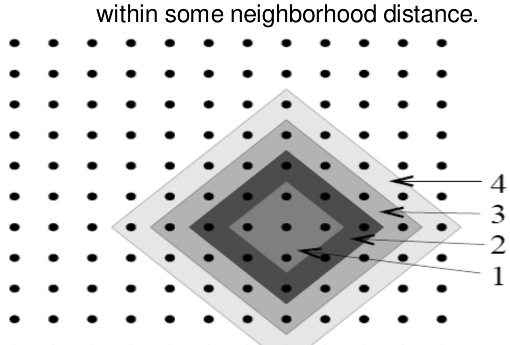

Fig. 2 : Agents located on a grid and allowed to interact only in a limited neighborhood

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. TYPES OF NORMS AND ITS IMMERGENCE

Due to multi-disciplinary interest in norms, several definitions for norms exist. Habermas , a renowned sociologist, identified norm regulated actions as one of the four action patterns in human behavior. A norm to him means fulfilling a generalized expectation of behavior, which is a widely accepted definition for social norms.
Researchers have divided norms into different categories. Tuomela has categorized norms into the following categories. r-norms (rule norms), s-norms (social norms), m-norms (moral norms), p-norms (prudential norms).
Rule norms are imposed by an authority based on an agreement between the members (e.g. one has to pay taxes). Social norms apply to large groups such as a whole society (e.g. one should not litter). Moral norms appeal to one's conscience (e.g. one should not steal or accept bribe). Prudential norms are based on rationality (e.g one ought to maximize one's expected utility). When members of a society violate the societal norms, they may be punished. Many social scientists have studied why norms are adhered to. Some of the reasons for norm adherence include: fear of authority or power, rational appeal of the norms, emotions such as shame, guilt and embarrassment that arise because of nonadherence, willingness to follow the crowd.
Elster categorizes norms into consumption norms (e.g. manners of dress), behavior norms (e.g. the norm against cannibalism), norms of reciprocity (e.g. gift-giving norms), norms of cooperation (e.g. voting and tax compliance) etc.
The role models are agents who the societal members may wish to follow. The inspiration is derived from human society where one might want to use successful people as a guide. Any agent in the society can become a role model agent if some other agent asks for its advice. The role model agent represents a role model or an advisor who provides normative advice to those who ask for help. In our mechanism, each agent will have at most one leader. An agent will

choose its role model depending upon the performance of its neighbors. We assume that agents that are connected know each other's performances. This is based on the assumption that people who are successful in the neighborhood are easily recognizable. In an artificial society, the agents which are involved in the norm emergence process interact with each other using a fixed learning algorithm [3]. The interaction is through playing some bi-matrix game. This game playing provides some rewards to the agents who ever achieve the goal. Every agent tries to achieve a higher reward or payoff value thus helping in the emergence of a norm.

V. AGENT INTERACTION: LEARNING ALGORITHM AND BI-MATRIX GAMES

This experiment is based on agent-agent interaction. The agents interact with each other using some bi-matrix games, learn the state of other agents using learning algorithms and decide the nest action they will take and next state they will go to.

By definition, a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience.

Applications of Machine Learning are like Learning to recognize spoken words SPHINX (Lee 1989), Learning to drive an autonomous vehicle ALVINN (Pomerleau 1989), Learning to classify celestial objects (Fayyad et al 1995), Learning to play world-class backgammon TD-GAMMON (Tesauro 1992), Designing the morphology and control structure of electro-mechanical artifacts GOLEM (Lipton, Pollock 2000).

There are several forms of learning algorithms. Among this reinforcement learning addresses the question of how an autonomous agent that senses and acts in its environment can learn to choose optimal actions to achieve its goals. This very generic problem covers tasks such as learning to control a mobile robot, learning to optimize operations in factories, and learning to play board games. Each time the agent performs an action in its environment, a trainer may provide a reward or penalty to indicate the desirability of the resulting state. For example, when training an agent to play a game the trainer might provide a positive reward when the game is won, negative reward when it is lost, and zero reward in all other states. The task of the agent is to learn from this indirect, delayed reward, to choose sequences of actions that produce the greatest cumulative reward. Reinforcement learning algorithms are related to dynamic programming algorithms frequently used to solve optimization problems. Some of the reinforcement learning algorithms are: Q learning, SARSA, SARSA-λ, etc.

VI. 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 homogeneous society of agents using same learning algorithm, which will interact with each other using different bi-matrix game. The learning algorithm used in our experiment is the SARSA reinforcement learning algorithm and bi-matrix games used are Prisoner's Dilemma, Chicken Game (Hawk- Dove Game), Assurance Game (Stag Hunt Game).

VII. RESULTS

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. 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. When the agents used Sarsa learning algorithm and interacted with different games the output graphs for the emergence of norms were as given below:

For the Prisoner's Dilemma game, different strategies such as alternate strategy, defect only strategy, etc were implemented. It was observed that

- For biased or defect only policy where the agents try to defect only, there was no emergence of norms. As the players didn't cooperate with each other.
- For alternate strategy or Tit-for-Tat policy the norm emerged quite smoothly.
- As the number of agents over the grid increased the social welfare value reached became lower.

For Assurance game the presence of two Nash equilibriums reduces the average payoff values. For that though the graph converges, it doesn't reach the social welfare value.

As chicken game contains two Nash equilibriums and it is a non- cooperative game the graph doesn't converge to any social welfare value. No emergence of any norm is seen.

VIII. CONCLUSION

The experiments carried out in this paper reveal that each agent learns from anonymous members of the society. This is in contrast to most results in multi-agent learning where two or more agents learn from repeatedly interacting with the same or different group. Norm emergences in real environments are likely to be influenced by both physical neighborhood effects imposed by mobility restrictions and biases as well as diverse learning and reasoning capabilities of members of the society. Our primary goal in this experiment is to evaluate the effect of dynamic homogeneous learning populations, which are sparsely located, on the rate and nature of norms that emerges through social learning. Due to the presence of more than one Nash equilibrium, in the bi-matrix games we find that a smooth emergence of norms doesn't occur in when the interaction is through Prisoner's Dilemma (Defect only policy) or Chicken Game or Assurance Game. This shows that for non-coordination games there is no emergence of norms. This proves the integrity of concept of norm emergence. This depicts a dynamic change of policies adapted by the electronic society. This change of policy is of function over time.

References

[1] Lewis D. (1969) *Convention, A Philosophical Study, Harvard University Press*.
[2] Axelrod R. (1997) *The complexity of cooperation, Agent-based models of conflict and cooperation, Princeton University Press, Princeton, NJ*.
[3] Sen S. and Airiau S. (2007) *Emergence of norms through social learning, In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI'07),* 1507–1512.
[4] Boella G. and van der Torre L. (2003) *Norm governed multiagent systems, The delegation of control to autonomous agents, In Proceedings of IEEE/WIC IAT Conference*, 329–335.
[5] Boella G. and Lesmo L. (2002) *A game theoretic approach to norms, Cognitive Science Quarterly*.
[6] Peyton H. Young (1996) *The Jounal of Economic Perspectives*, 10(2):105–122.
[7] Liviu Panait and Sean Luke (2005) *Autonomous, Agents and Multi-Agent Systems*, 11(3):387–434.
[8] Tuyls K. and Now´e A. (2006) *The Knowledge Engineering Review*, 20(1):63–906.
[9] Rob Powers and Yoav Shoham (2005) *New criteria and a new algorithm for learning in multiagent systems, In Proceedings of NIPS*.
[10] Kagan Tumer and David H Wolpert (2000) *Collective intelligence and Braess' paradox, In Proceedings of the Seventeenth National Conference on Artificial Intelligence*, 104–109, *Menlo Park, CA. AAAI Press.*
[11] Mukherjee P., Sen S. and Airiau S. (2008) *Norm emergence under constrained interaction in diverse societies. Proc of 7th Int. conf. on Autonomous Agents and Multiagent Systems(AAMAS 2008),* 779-786.

The graphs convergence pattern of different games are shown below

(i) For Prisoner's Dilemma - Alternate Strategy:

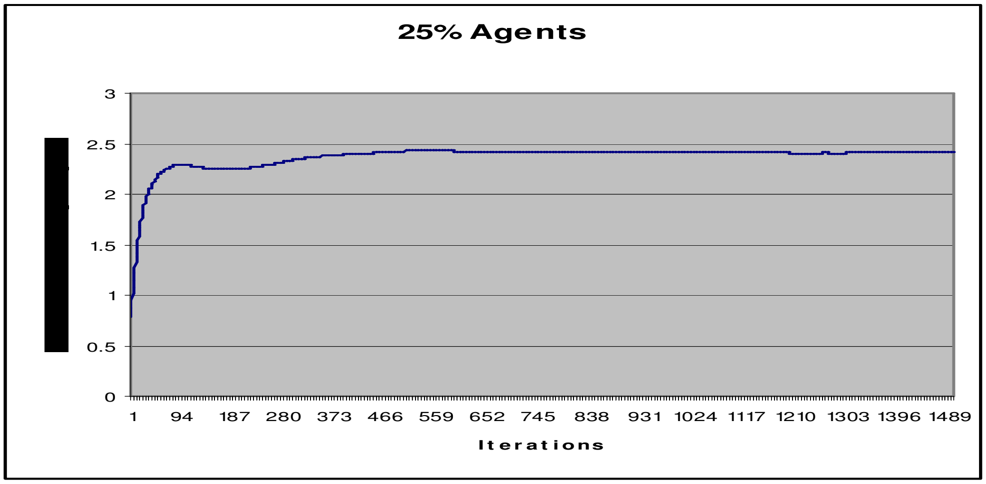

Fig. 3: Cumulative Average Payoff for 25% agents following alternate strategy

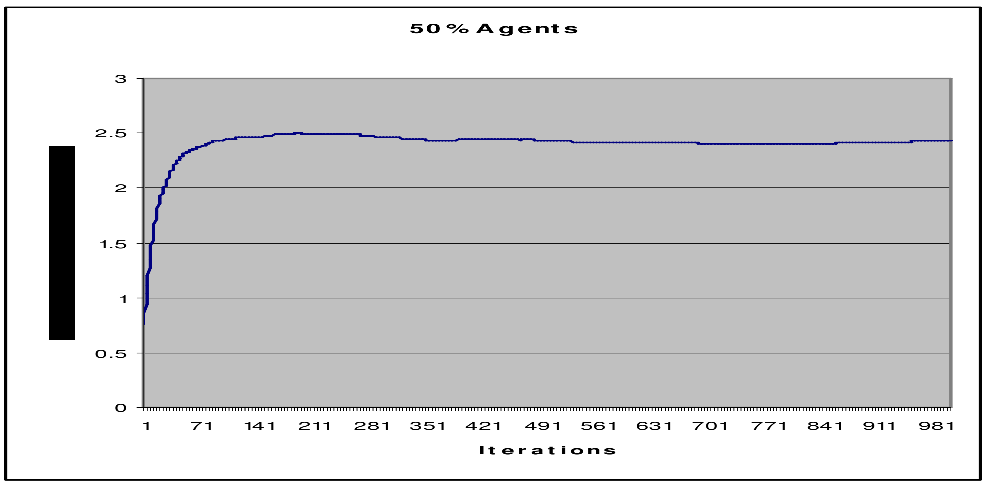

Fig. 4 :Cumulative Average Payoff for 50% agents following alternate strategy

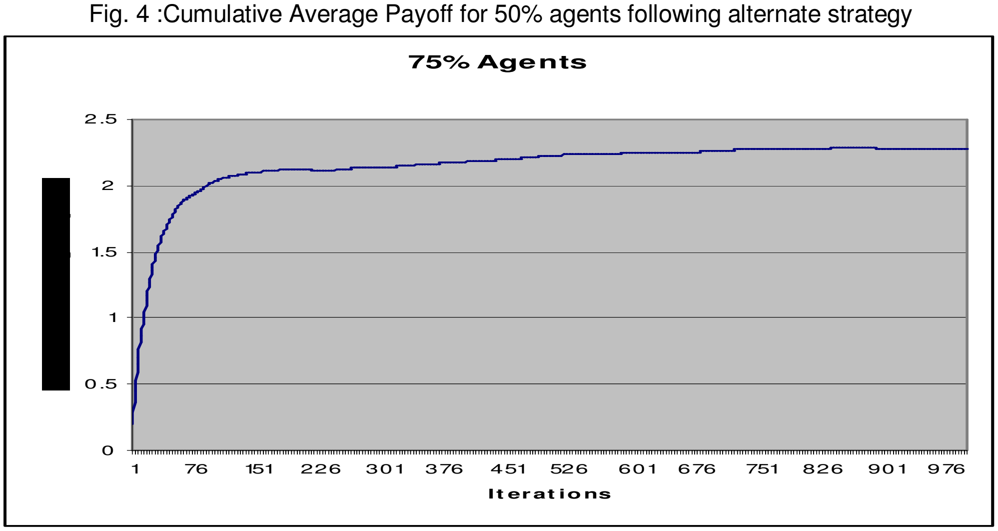

Fig. 5: Cumulative Average Payoff for 75% agents following alternate strategy

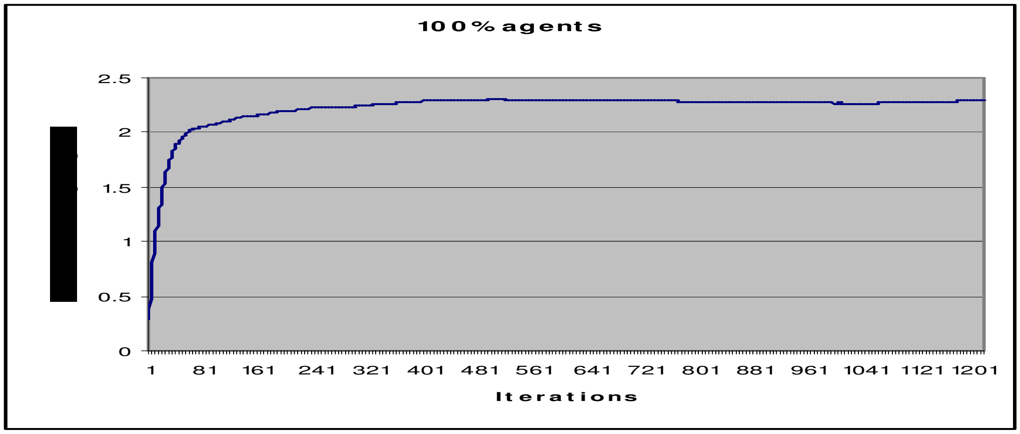

Fig. 6: Cumulative Average Payoff for 100% agents following alternate strategy
(ii) For Assurance Game

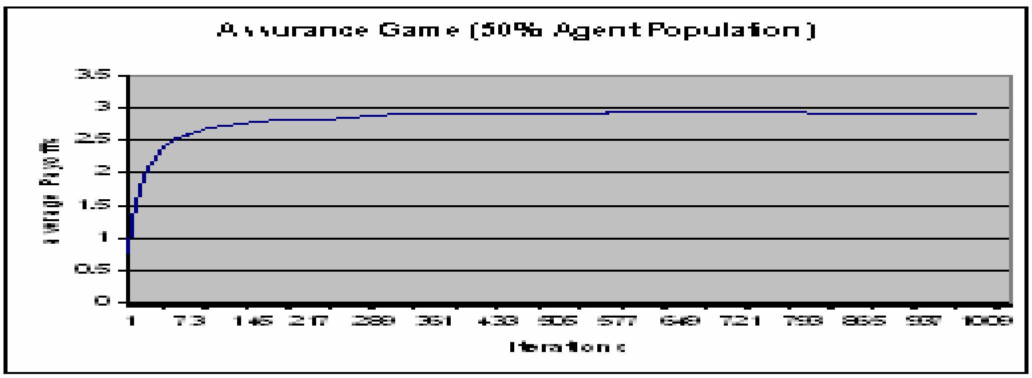

Fig. 7: Cumulative Average Payoff for 50% agents Assurance Game

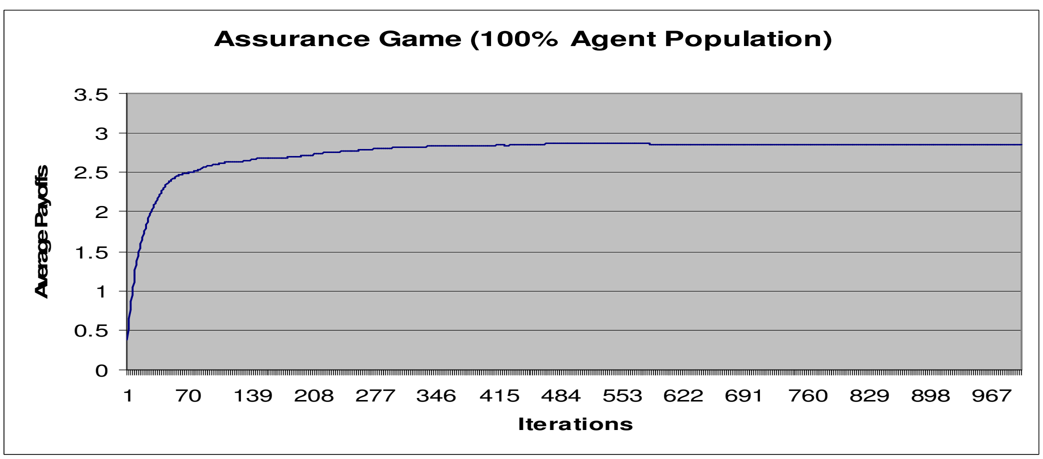

Fig. 8: Cumulative Average Payoff for 100% agents Assurance Game

(iii) For Chicken Game

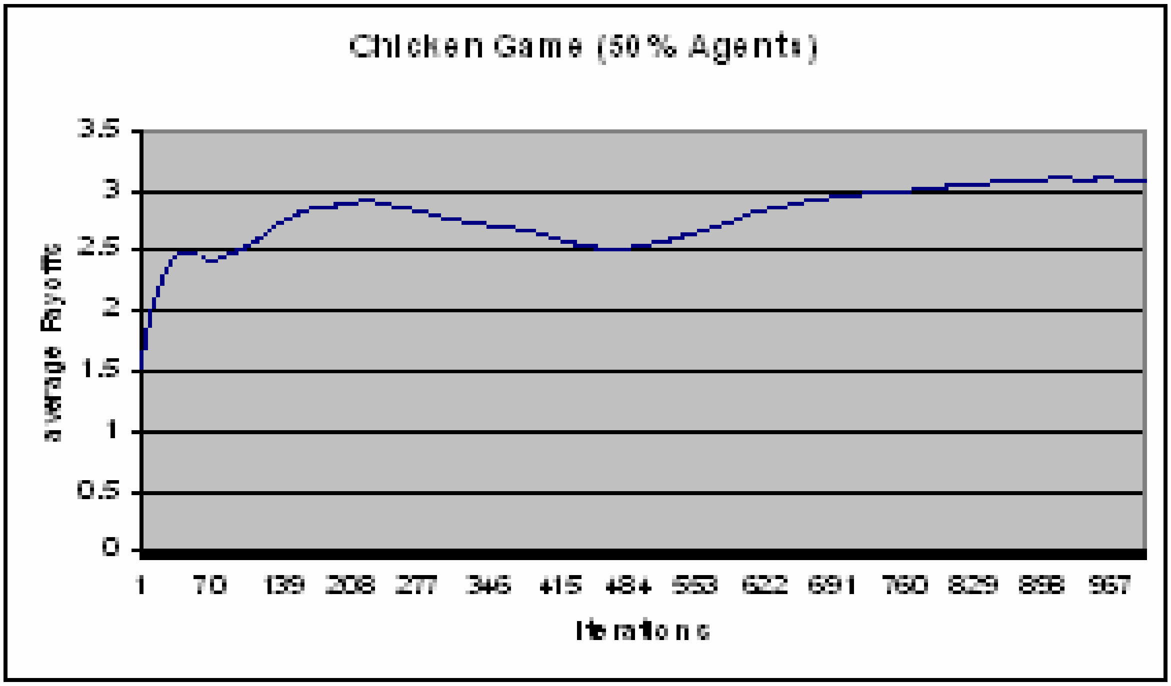

Fig. 9: Cumulative Average Payoff for 50% agents Chicken Game

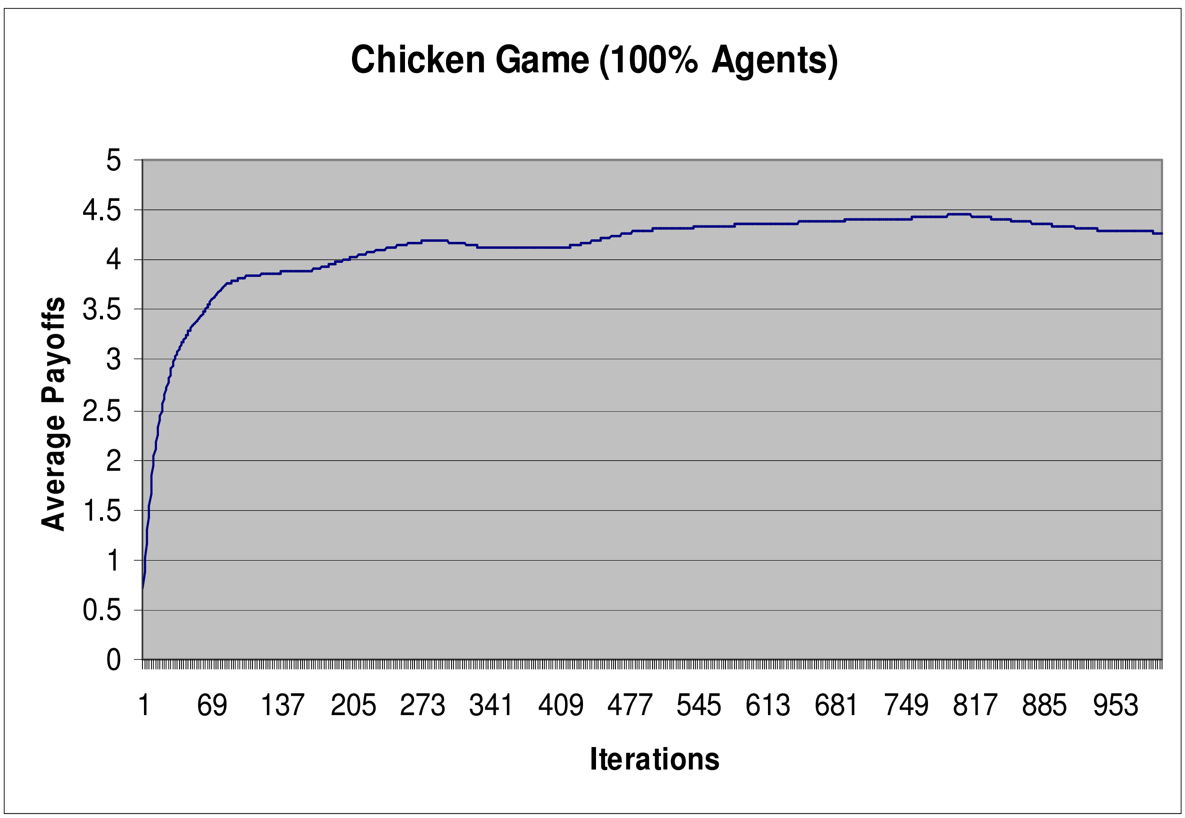

Fig. 10 Cumulative Average Payoff for 100% agents Chicken Game