Effects of interaction history and network topology on rate of convention emergence

Daniel Villatoro
Artificial Intelligence Research Institute (IIIA)
Spanish Scientific Research Council (CSIC)
Bellatera, Barcelona, Spain
dvillatoro@iiia.csic.es

Nick Malone
Department of Mathematical and Computer Science
University of Tulsa
Tulsa, Oklahoma, USA
nick-malone@utulsa.edu

Sandip Sen
Department of Mathematical and Computer Science
University of Tulsa
Tulsa, Oklahoma, USA
sandip-sen@utulsa.edu

ABSTRACT

Social conventions are useful self-sustaining protocols for groups to coordinate behavior without a centralized entity enforcing coordination. The emergence of such conventions in different multi agent network topologies has been investigated by several researches. Although we will perform an exhaustive study of different network structures, we are concerned that different topologies will affect the emergence in different ways. Therefore, the main research question in this work is comparing and studing effects of different topologies on the emergence of social conventions. While others have investigated memory for learning algorithms, the effects of memory on the reward have not been investigated thoroughly. We propose a reward metric that is derived directly from the history of the interacting agents. The reward metric is the majority rule, thus the emerging convention becomes self propagating in the society. Agents are proportionally rewarded based upon their conformity to the majority action when interacting with another agent. Another research question to be answered is what effect does the history based reward function have on convergence time in different topologies. We also investigate the effects of history size, agent population size and neighborhood size proving their effects by agent-based experimentation.

1. INTRODUCTION

Social norms such as driving on the left side of the road or not stepping in front of other people in line are prevalent in human groups and societies. Such norms are conflict resolution strategies that develop from the population interactions instead of a centralized entity dictating agent protocol. History of interaction is then an instrumental in norm evolution. Learning algorithms incorporate history of interaction into their calculations, but reward metrics are typically static and independent of the agent histories. Norm evolution is dependent upon the exertion of social pressure by the group on aberrant individuals. It is through learning via repeated interactions that social pressure is applied to individuals in the group. However, a static reward metric does not necessarily model the nature of social pressure in human societies. We propose a reward structure based upon the agent's interaction history as a more appropriate alternative to the static reward metric normally used. In our model agents are rewarded based upon the conformity of action between two agents, such that the agent who has the most of the majority interaction receives higher reward. In our system both agents' history is used to calculate the payoff. In addition we investigate how memory about this history and memory size affects different types of society structure. Societal connections are represented by different network types in which the network links constrain interaction between agents. Specifically we investigate a one-dimensional lattice with variable neighborhood size, and scale free networks. Agents are then randomly selected via their constraints to play a two-player game with the reward for interaction based on their history of interaction. We believe that the underlying topology of the society is a key factor in determining the process of convention emergence. In this work we will **experiment on different types of topologies in order to observe, compare and analyze their effects and dynamics of reaching social conventions**. Mainly we are interested in two types of topologies: a one-dimensional lattice, and, a scale free network.

Cite as: Effects of interaction history and network topology on rate of convention emergence, D. Villatoro, N. Malone and S. Sen, *Proc. of 3rd Int. Workshop on Emergent Intelligence on Networked Agents (WEIN'09)*, May, 12, 2009, Budapest, Hungary, pp. XXX-XXX.

The structure of this article is as follows: we start reviewing the previous and related work in the area of multi-agent emergence in social conventions in Section 2; after that, in Section 3 we present the model that is going to be used in this work and in the experimentation; the experimental results are presented in section 4; after that some conclusions are presented in Section 5, and finally we present the future work we plan to perform to continue this research in section 6.

2. PREVIOUS WORK

In the work of Sen and Airiau [5, 3], the authors explored norm emergence where interaction rewards were not dependent upon previous interactions. That work is focused on the problem of coordination of two cars arriving at an intersection. Each agent can choose to "go" or "yield" to the other agent. The reward metric is designed so that if each agent chooses the same action, they receive small payoff but if agents choose opposite actions, they receive a large payoff. So if the row and the column agents both "go" they both receive a poor payoff. Each agent in an interaction was randomly chosen from the population. Agents learned the social norm from repeated interactions with other agents in the population. The history of interaction does not directly affect the reward agents receive. Reward is only affected by the agents' action choice in the current interaction. However, learning takes place via social pressure from repeated interaction, thus the history of interaction influences agent's action choice.

Delgado et al. [1] investigate a similar norm emergence scenario with several key differences. The agents in their research are restricted in their interactions to their neighbors in a scale free graph.

Furthermore, their agents are playing a coordination game in which payoff is high if both agents chose the same action and low if both agents chose different actions. The authors formulate their action choice in terms of history. Each agent keeps a history of interactions and the corresponding reward. The agents then utilize the history to select the best payoff action. However, the history does not determine the reward they receive.

Kittock's research [2] is very similar to the research done by Delgado et al. Kittock also utilizes the same style of payoff metric used in Delgado's work as well as using a graph to restrict to interactions of agents. His agents also utilize memory of interaction payoffs to select their actions in future interactions. His work is different in that he investigates several graph topologies and payoff matrices.

3. MODEL

The social learning situation for norm emergence that we are interested in is that of learning to reach a social convention. We borrow the definition of a social convention from [6]: A social law is a restriction on the set of actions available to agents. A social law that restricts agents' behavior to one particular action is called a social convention. In our case, as in the case in [1], a social convention will be reached if all the N agents are either in state A or in state B. From now on, it is equivalent for us that an agent chooses an action to perform or an state to be.

We represent the interaction between two agents as an n-person m-action game. At each time step, each agent is paired with another agent and decides in which state it wants to be (A or B).

In this article we restrict agents to be located in two different types of environments: a *one-dimensional lattice* with connections between all neighboring vertex pairs (one example can be seen in Fig. 1(a)) ; and, a *scale-free network*, whose node degree distribution asymptotically follows a power law (one example can be seen in 1(b)) [1]. In both environments, agents are represented by the nodes in the network and the links represent the possibility of interaction between nodes (or agents). The *one-dimensional lattice* provides a structure in which agents are connected with their *n* closest neighbors. Different values of the neighborhood size (*n*) produces different network structures; for example, when $n = 2$ the network will have a ring structure and agents will only be connected with their direct neighbors (those at left and right if we imagine a ring topology). On the other hand, when $n = PopulationSize$, the network will be a fully connected network where all the agents are connected with all the others. On the other hand, in the *scale-free network* there are many vertices with small degrees and only few of vertices with large degrees. This makes the network diameter (average minimum distance between pairs of nodes) significantly small with respect to the *one-dimensional lattice*.

As in [2] agents have a memory of size M (same size for all the agents). For agent k, the memory M_k will record some information on the history of its decisions: The value of the position *i* of the memory M_k will be a tuple $\langle a_k^i, t^i \rangle$ where t^i is the time the *i*-th took place, and a_k^i is the decision taken by agent *k* at time t^i ($1 \leq i \leq M$). Thus, the memory of each agent will work as a history record for the last *memory size* actions taken.

Agents do not know before interaction the payoff that they will get for any of the states they choose. Agents cannot observe the other agent's memory, current decision or immediate reward. The payoff given by the system depends on both agent's decision and their memories. When two agents y and z interact, the instantaneous reward that, for example, agent *y* gets is calculated based on the action it selected and the previous history of both agents as follows: where A_x and B_x are the number of *A* and *B* actions in

$TotalAActions = A_y + A_z$;
$TotalBActions = B_y + B_z$;
if $TotalAActions \leq TotalBActions$ **then**
  $reward_y = \frac{B_y}{TotalBActions}$
else
  $reward_y = \frac{A_y}{TotalAActions}$
end

Algorithm 1: Reward Calculation based on action history.

memory that agent *x* has taken. Agents use a learning algorithm to estimate the worth of each action. Agents will choose their state in each interaction in a semi-deterministic fashion. A certain percentage of the decisions will be chosen randomly, representing the exploration of the agent, and the rest of te decisions will be chosen deterministically corresponding to the action estimated to be of higher utility. In this article, the exploration rate has been fixed at 25 %. Therefore, one out of four decisions will be taken randomly.

The learning algorithm used here is a simplified version of the Q-Learning algorithm [7]: The Q-Update function is as follows:

$$Q^t(a) \leftarrow (1 - \alpha) \times Q^{t-1}(a) + \alpha \times reward$$

Algorithm 2: Q-update function

Therefore, when agents decide not to explore, they will choose the action with higher Q value.

The protocol of interaction is presented in Algorithm 3

for *timesteps* **do**
  forall *agents* **do**
    Select another agent from population;
    Ask agents to select an action;
    Send the joint action for policy update;
    Update;
  end
end

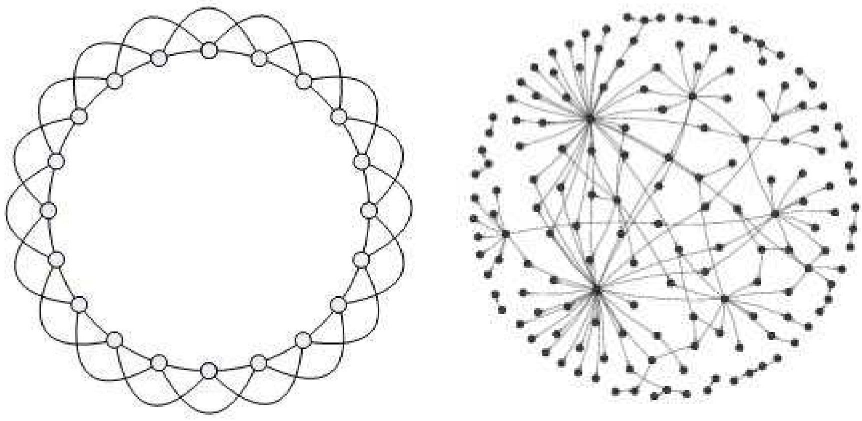
Algorithm 3: Interaction Protocol

4. EXPERIMENTS

We present experiments that have been run on societies with different configurations using the following parameters:

- Memory Size: We vary the number of past interaction stored, so we can contrast the effects of memory sizes.
- Population Size: We want to extract conclusions that are unrelated with the population size, therefore, we need to experiment on different population sizes so we can extrapolate the results.
- Neighborhood Size: We are interested in observing how different neighborhood sizes in a one dimensional lattice affect the process of emergence of conventions.

[1]In [4] we can find several examples of usage of scale-free networks (like citations between scientific articles, a 300 million vertex subset of the World Wide Web, the power grid of the western United States, or, the interaction network of proteins in the metabolism of the yeast S. cerevisiae).



(a) One-dimensional Lattice (b) Scale-Free Network

Figure 1: Underlying Topologies

- Underlying Topology: We observe the dynamics of the process of emergence of conventions depending on the underlying topology, scale free or one dimensional lattice.

Some details about how the experiments have been run are the following: All the experiments performed in this section have been run 50 times and the results are averaged. All agents are initialized with uniformly distributed values in their memories, and initially do not have any preference for any action. We define that a social convention has been reached when 90% of the population are in the same state [2].

The hypotheses that we want to test are the following:

1. Different neighborhood sizes affect the convergence time in one dimensional lattices.

2. Convergence time is proportional to memory size: We have the intuition that when agents have a bigger memory size, it will take longer for them to reach a convention.

3. Convergence time is proportional to population size: We believe that the bigger the population is, the longer it takes the system to converge.

Though some aspects of results from our simulated agent society can be transferred to human situations (with additional mechanisms), our results are targeted towards a better understanding of how to develop self-adaptive agent societies.

4.1 Effect of Neighborhood Size

For the first experiment, we observe the effects of neighborhood size on the convergence time. The memory size has been fixed to 25, and the underlying topology used is just the one-dimensional lattice, as scale-free networks predetermine the neighborhood for each node. The results in Figure 2. show a comparison between 4 different population sizes and the corresponding convergence times for changing neighborhood sizes.

We can notice from the experimental results that when increasing the neighborhood size, the convergence time is reduced, until a certain point where it stabilizes. This effect is produced due to the dynamics of the system. When agents have a small neighborhood size, on average, they need a higher number of interactions to transmit their decisions from any two randomly chosen agents.

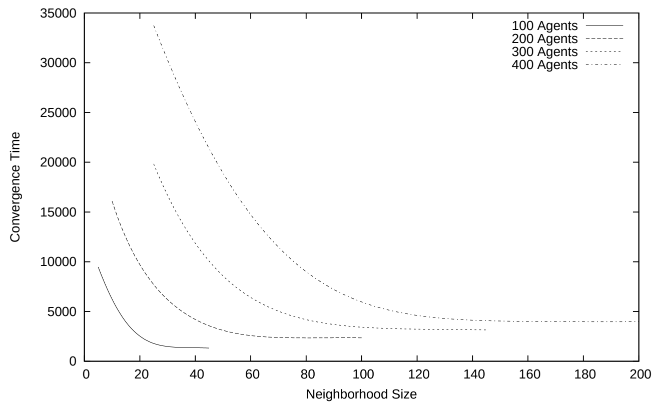
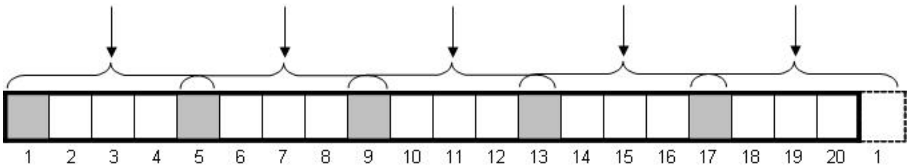


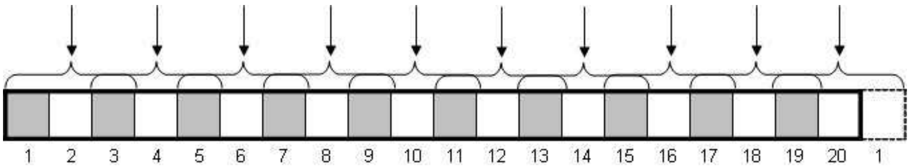
Figure 2: Convergence rates with different Neighborhood Sizes (Memory Size = 25)

Consequently, the *"interaction paths"* between any two agents are larger. For example, for the knowledge of agent in cell 2 in Figure 3(a) to reach the agent in cell 12, it would be necessary, on average, to have 5 interactions between intermediate agents. In a similar way, when increasing the neighborhood size, we reduce the average number of interactions for knowledge to be transmitted between any random pair of agents. However, for the knowledge of agent in cell 2 in Figure 3(b) to reach the agent in cell 12, it would be necessary, on average, to have 2 interactions between interactions between intermediate agents.

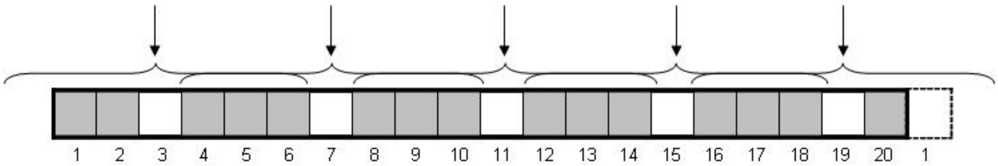
Moreover, we can observe in Figure 2 that in all the scenarios tested, the pattern of the system convergence is replicated(apparently following a geometric distribution). The neighborhood size beyond which the convergence time does not significantly decrease anymore is when the neighborhood size is half of the population size. The knowledge of any agent can reach another agent in just one interaction, facilitating knowledge dissemination in the system and reducing the number of conflicting conventions among interacting agents. For higher values of the neighborhood sizes, the results will be the same as the same set of agents can be reached with any value for the neighborhood size above $\frac{Population}{2}$.



(a) 20 Agents with Neighborhood Size = 2



(b) 20 Agents with Neighborhood Size = 4



(c) 20 Agents with Neighborhood Size = 6

Figure 3: Different neighborhood sizes in one dimensional lattice produce different average interactions to reach conventions. (shaded cells indicate agents that overlap two neighborhoods)

4.2 Effect of memory size

In this experiment, we want to observe the effect of different memory sizes on convention emergence in both underlying topologies. Therefore, we keep the parameters constant. The population size will be fixed at 100 agents, and in the case of the one-dimensional lattice, the neighborhood size is fixed at 8. We present the results in Figure 4. The information plotted represent the average convergence time[2] for a society with the previously specified parameters. The memory size is given on the x-axis and the convergence time can be observed on the y-axis. The experimental results confirm that the memory size matters, and with longer memory convergence time increases. The reason for this phenomenon is due to the configuration of the reward function and the learning algorithm. When having a small memory size, the proportional reward that an agent gets is higher due to the design of the reward function defined in Algorithm 1, therefore, the learning algorithm receives larger reinforcements for the actions performed. Consequently, if the action is reinforced to a higher degree, convergence will be reached faster. Convergence is accelerated in this situation because higher rewards have a larger impact on the Q value updated by the learning algorithm 2. On the other hand, when dealing with higher history windows, the proportional reward is much smaller, and therefore, the reinforcement will be smaller. Due to this smaller reinforcement, a higher number of interactions, which translates into higher number of timesteps, will be needed to reinforce that action to same degree, and hence, convergence times will increase.

In Figure 5 we can observe two independent executions per scenario (Figure 5(a) shows the results for the one dimensional lattice and in Figure 5(b) the results for the Scale-Free Network). One of the runs shows the evolution of action selection with a memory size limited to 25, and the other one to 150. These results confirms the statement that smaller memory sizes lead to faster convergence (though these results are just from a single execution). We can observe in Figure 5(a) that the system, when fixed to a higher memory size (150) initially converges towards action (or state) B, but then oscilates, finally converging to action A. So initial biases for a convention can be reversed in the long-run.

Furthermore, in Figure 6 we can observe (also for the same two independent executions for which results are shown in Figure 5) the

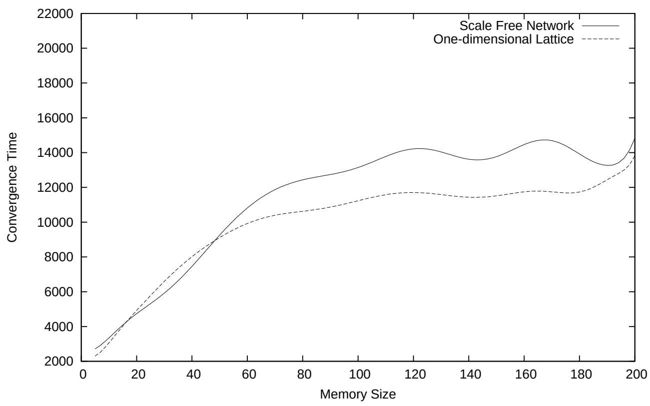
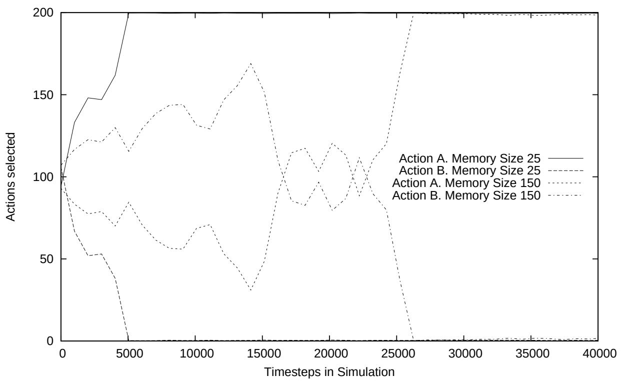
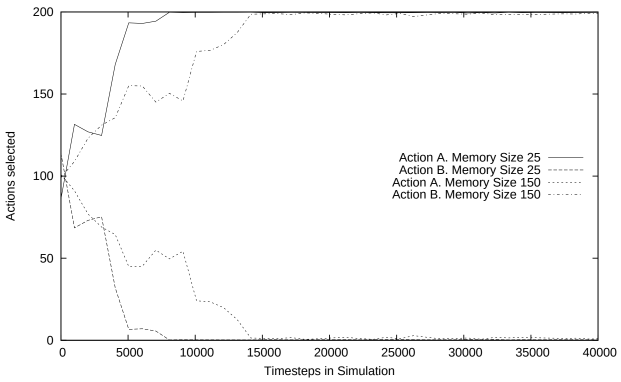


Figure 4: Effect of Memory Size of Convergence Time. (100 Agents)

[2]We remind the reader than the convergence time is the number of timesteps needed so 90% of the population are in the same state.It does not matter which state the agents have converge to as long as 90 % of the agents are in that same state.

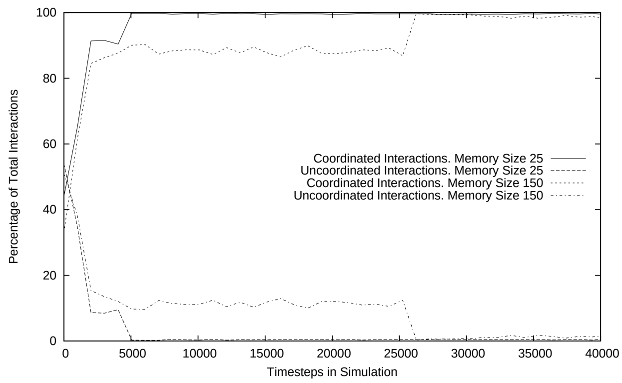


(a) One dimensional Lattice

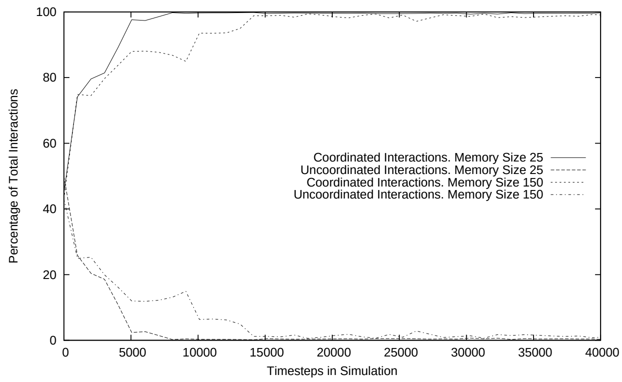


(b) Scale Free Network

Figure 5: Frequency of different action selection by agents in different environments and variable memory size.



(a) One dimensional Lattice



(b) Scale Free Network

Figure 6: Evolution of Coordinated and Uncoordinated Actions by agents in different environments and variable memory size.

evolution of coordinated interactions. A coordinated interaction is one interaction where both agents have chosen the same action and this action is also the majority action for both agents' memory. In case agents do not chose the same action, it will count as an uncoordinated interaction. The figures suggest that history size affects both network topologies equally.

4.3 Effect of Population Size

In this set of experiments we want to observe how the convergence time is affected by the population size. It is intuitive that the larger the societies are, the harder it should be for the agents to reach a convention. However, we also want to observe the effects of different connection topologies, specifically scale free networks and one dimensional lattices. In these experiments we use a memory size of 25, and for the one-dimensional lattice we use a neighborhood size of 8. In Figure 7 we can observe how the convergence time is directly proportional to population size.

However, different effects can be observed depending on the environment. When the population increases in the one-dimensional lattice, the convergence time is much higher than in the scale-free network. This phenomena is due to the structure of the society and the characteristics of both networks. In the one-dimensional lattice, when society size increases, the convergence takes longer to reach because agents will create islands with local conventions (one example can be seen in Figure 8), and as it was explained before in Section 4.1, a higher number of interactions will be needed for knowledge to be transmitted. When these local conventions are

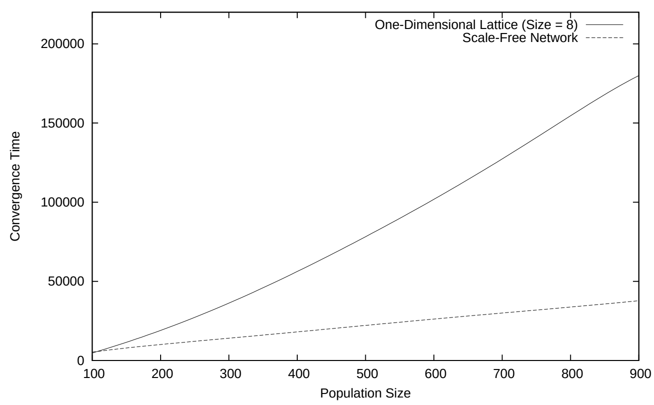


Figure 7: Population Size Comparison in different environments

created it takes time to break out of them and reach a global convention. The amount of time depends on the number of agents that have reached the same local convention. This is the reason why when increasing population sizes (without increasing the neighborhood size) in one-dimensional lattices environments, convergence time also increases.

On the other hand, we can observe how the Scale-Free Network

Figure 8: Example of one-dimensional lattice with local conventions.

reduces the factor of growth convergence time with respect to the one-dimensional lattice. The explanation for this is due to the inherent structure and characteristic of the scale-free network: the average distance between two vertices in the network is very small relative to a highly ordered network such as a lattice. In a scale free network knowledge dissemination is still fast and effective even with larger populations. This structure helps in the process of reaching a convention. Although we can observe than when the population size increases, it also does the convergence time, the growth factor is much smaller than that observed in the lattice.

5. CONCLUSIONS

We have presented a set of experiments to study the emergence of social conventions based not only on direct interactions but also on the memory (and previous history) of each of the agents. This social learning framework requires each agent to learn from repeated interaction with anonymous members of the society. Norm emergence in real environments are likely to be influenced by both physical neighborhood effects imposed by mobility restrictions and biases as well as diverse learning, memory and reasoning capabilities of members of the society. Our main goal in this paper was to study the effects of these features on the rate of norm emergence. Our initial hypotheses were that different characteristics of the topology in which agents are located would produce different results in terms of the convergence time of reaching a social convention. Experimental results have confirmed our hypotheses. We have shown how in a one-dimensional lattice (although results can be extrapolated to any spatial environment), when agents are allowed to interact with other agents located farther away from them (and therefore not in their immediate neighborhoods) conventions are reached in less time. The reason for this acceleration is that agents interact with a larger percentage of the population, which prevents small local conventions from forming. Moreover, and affecting both types of constraint interaction topologies, we have observed that the memory size does affect the emergence of conventions. Systems where agents have larger memory sizes take longer to reach conventions. The reason for this hindrance to convergence is due to the design of the reward function: the reward that a certain action receives is inversely proportional to the size of the memory, and therefore, with higher memory sizes, the reward actions received will be smaller. Due to this, we will need a higher number of interactions for the convention to be reached. Finally, we have experimented with different population sizes. The main conclusion that we can extract is that for both types of environments, the larger the population, the longer it takes to reach a convention. However, we could observe that scale-free network structures are not affected as adversely by the increasing the population as one-dimensional lattices. The resistance of the scale free network is due to the inherent structure of the scale-free network, where the hubs facilitate knowledge dissemination and the network diameter is relatively small even with large populations.

6. FUTURE WORK

This experimental framework has served as a proof-of-concept testbed for our initial hypotheses in emergence of social conventions. In the process, however, issues have arisen. One of the most immediate experiments that we plan to perform concerns the reward function. Up to now, this reward function assigns a reward proportional to the number of majority actions that a certain agent has performed with respect to the sum of the majority action in history taken by both agents. In the next set of experiments we want to vary this reward function so it will only assign a reward to those agents that have performed the majority action, and not to all agents as it is done now. We surmise that such a change will produce a speed-up in the convergence times.

Another question that we plan to answer in future versions of this work is under what circumstances and configuration of parameters the one-dimensional lattice behaves similarly to the scale-free network for large population sizes. We could see in Section 4.3, when the population size increases, the convergence times in the one-dimensional lattice increases at a much faster rate compared to scale-free networks. We believe that a dynamic adjustment of the neighborhood size on a one-dimensional lattice will produce similar dynamics to those obtained in scale-free networks. We also want to experiment with heterogeneous populations, as it is done in the work of Mukherjee [3]. So far all the agents are initialized with the same parameters and with the same distribution of initial memory. We want to observe the resulting dynamics of different types of populations, for example: in a scale-free network, initialize the hubs with a specific bias towards a certain state, and observe the speed of convergence of the rest of the population. Another interesting experiment to be carried out is when agents in the same population are initialized with different memory sizes.

Furthermore, as a generalization of this work, we are interested in comparing the results we have presented here with different connection topologies, for example, other type of two dimensional graphs, random networks and small world networks.

Finally, to make the model more general, we want to extend the game from the actual 2-action game to a n-action game. This extension will give us a more generalized game, closer to real life situations.

Acknowledgments.

This work was supported by the European Community under the FP6 programme [eRep project CIT5-028575]; the Spanish Education and Science Ministry [AEI project TIN2006-15662-C02-01, AT project CONSOLIDER CSD2007-0022, INGENIO 2010]; Proyecto Intramural de Frontera MacNorms [PIFCOO-08-00017] and the Generalitat de Catalunya [2005-SGR-00093]. Daniel Villatoro is supported by a CSIC predoctoral fellowship under JAE program. Nick Malone is supported by A Graduate Assistanship from the University of Tulsa. Sandip Sen is partially supported in part by a DOD-ARO Grant #W911NF-05-1-0285.

7. REFERENCES

[1] J. Delgado, J. M. Pujol, and R. Sangüesa. Emergence of coordination in scale-free networks. *Web Intelli. and Agent Sys.*, 1(2):131–138, 2003.

[2] J. E. Kittock. Emergent conventions and the structure of multi-agent systems. In *Lectures in Complex systems: the*

proceedings of the 1993 Complex systems summer school, Santa Fe Institute Studies in the Sciences of Complexity Lecture Volume VI, Santa Fe Institute, pages 507–521. Addison-Wesley, 1993.

[3] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence with biased agents. *International Journal of Agent Technologies and Systems (IJATS)*, 1(2):71–84, January 2009.

[4] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[5] S. Sen and S. Airiau. Emergence of norms through social learning. *Proceedings of IJCAI-07*, pages 1507–1512, 2007.

[6] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94:139–166, 1997.

[7] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.