Interaction, observance or both? Study of the effects on convention emergence

Daniel VILLATORO [a,1], Jordi SABATER-MIR [a] and Sandip SEN [b]

[a] *IIIA - CSIC. Spain*

[b] *University of Tulsa. USA.*

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 researchers. In the literature we can observe two different approaches taken by agents to reach conventions: direct interactions with other agents that affects its payoff, or, an observance approach based on a majority rule, where agents change their state depending on their neighbors states. However, we proposed a mixed strategy, where agents are able to do both: interact and observe. The research question to be answered in this work is the percentage of each approach an agent should take in order to reach conventions faster.

Keywords. Conventional Norms, Emergence, Social Networks, Reinforcement Learning

1. Introduction

Conventions 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. Conventions are a special type of social norm. Coleman [1] claims that conventions are a type of norms that are directed at certain *focal actions*. The term *focal action* is directly borrowed from game theory where it exists the concept of *focal point*. A focal point is defined as a solution that players will tend to use in the absence of communication, because it seems natural, special or relevant to them. [2] All the focal actions are equally good, as long as it is the same one chosen by all the players.

Such conventions are conflict resolution strategies that develop from the population interactions instead of a centralized entity dictating agent protocol. However interactions

[1]Corresponding Author: Daniel Villatoro, Artificial Intelligence Research Institute (IIIA), Spanish National Research Council (CSIC), Campus UAB, Bellatera, Barcelona, Spain ; E-mail: dvillatoro@iiia.csic.es.

[2]For example, imagine that you and your partner are visiting Paris. It is the first time for both of you in that city. Unfortunately, you are not in the same hotel and you have no means to communicate with each other, although you know that you have to meet each other at a certain time in a public place. You can choose between all the public places in Paris. Using a common sense reasoning, you might choose to go to the Eiffel Tower, the Pyramid of the Louvre Museum, or the Arc de Triomphe. Those would be focal points in the decision game. Consequently the focal action will be the action taken by the agents in the absence of communication.

are not the only mechanism available to agents in order to reach conventions. In everyday life there are a number of scenarios where agents, before proceeding and interact with other agents, stop and observe how the others are behaving. This observation provides agents with extra information of the conventions followed in the society. However, this observance in a multiagent society might imply a cost that is not reflexed in human societies. In several works of the literature, like [2] or [3], a *Majority Rule* is used. This majority rule makes agents change to the state of the majority of their neighbors. This observance provides to the agent a more general view of the actual state of its neighborhood than the direct interactions do.

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 reward metric based on the current interaction does not necessarily model the full context or capture the persistent nature of social pressure in human societies. In particular, society often uses past history to judge individuals and hence actions have future consequences in addition to immediate effects. Accordingly, we propose a reward structure based upon the agent's interaction history as a more appropriate alternative to the single interaction 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. Hence, both interaction agents' history of actions are used to calculate each individuals' payoff from an interaction. We investigate in this work how a certain percentage of observance in the norm emergence process (on top of the direct interactions) affects in the convergence time of this process.

We are also keenly interested in understanding how agent relationships and social connections affect the success and rate of adoption of social norms. We represent different societal connection topologies by different network types in which the network links represent interactions between agents.

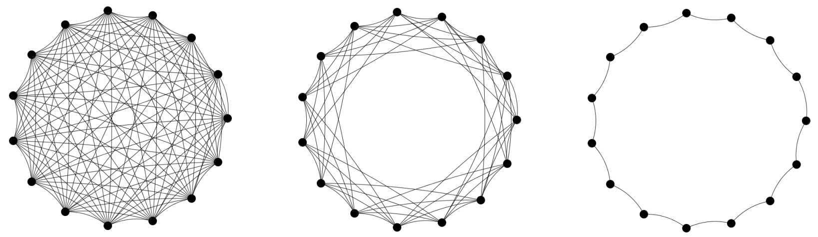
Given a connection topology, agents repeatedly play a two-player game with the reward for interaction based on their respective action histories. We believe that the underlying topology of the society is a key factor in determining the convention emergence process. In this work we will experiment on different types of topologies in order to observe, compare and analyze their effects and dynamics of reaching social conventions.

The structure of this article is as follows: in Section 2 we present the agent interaction and reward model that we have used; experimental results are presented in Section 3; conclusions from the analysis of the results are presented in Section 4, and finally we present the future work in Section 5.

2. 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 [4]: *A social law is a restriction on the set of actions available to agents*. A social law that restricts agents' behavior to one particular focal action is called a social convention.

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



(a) Fully Connected Network (b) One-dimensional Lattice with Neighborhood Size 8 (c) Ring Network or One-dimensional Lattice with Neighborhood Size 2

Figure 1. Underlying Topologies

to be. In our case, as in the case in [5], a social convention will be reached if all the n agents are in the same state, i.e., the actual state chosen is immaterial. For our purpose, an agent choosing a particular action is equivalent to it being in a corresponding state. In this paper, we consider only binary interactions, i.e., n=2, although we will study different values of the number of focal actions available to the agents. We would like to remember to the reader that in this work it will not matter the focal action where agents reach a convention at; all the focal actions are equally good, as long as they all converge to the same.

As we specified, agents will be located in a network topology that will represent the social network. In this work we will focus on the effects of a *one-dimensional lattice* with connections between all neighboring vertex pairs (examples can be seen in Figures 1(a), 1(b) and 1(c))

Each agent is represented by a node 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* nearest neighbors. Different values of the neighborhood size (*n*) produces different network structures; for example, when $n = 2$ the network will have a ring structure (as in Figure 1(c)) 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 is a fully connected network (as in Figure 1(a)) where each agent is connected with all other agents.

As in [6], we use agents with a memory M_k 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 memory event 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 record of the history for the last *memory size* actions taken by the agent.

Agents cannot observe the other agent's current decision, or immediate reward, and hence cannot calculate the payoff for any action before actually interacting with the opponent.

When an agent is chosen to execute an action, two actions can be chosen: *Observance* or *Interaction*. The action to be executed is randomly chosen by the agent following a standard distribution and the probability of observance assigned to each simulation.

In the experiments presented in this work, this probability will vary in order to study the effects of different *observance probabilities*.

When an agent have to *observe*, it *observes* the preferences of its neighboring agents. The observing agent will reinforce with 1 the action preferred by the majority. In this work we will assume that the observance action is cost-free, although it will be interesting studying situations where this observance implies a cost.

On the other hand, if an agent has to *interact*, it will randomly choose one of its neighboring agents to interact with. When two agents *interact*, the instantaneous reward that an agent receives is calculated based on the action it selected and the action history of both agents as shown in Algorithm 1, where $Action_x$ is the last action taken by agent x, and for which it is rewarded, $MajorityAction$ is selected to be whichever action is played most by the two players combined, $MajorityActions_x$ is the number of actions equal to the majority action that agent x has previously taken, and $TotalMajorityActions$ is the number of times the majority action was chosen by both players in their finite histories.

Algorithm 1 Memory Based Reward Function.

(*First, we select the majority action*)
$FindMajorityAction$
(*Then, we calculate the reward depending on the agents action selection and on the majority action*)
if $Action_1 == MajorityAction$ **then**
  $reward_1 = \frac{MajorityActions_1}{TotalMajorityActions}$
else
  $reward_1 = 0$
end if

This reward function has been designed like this in order to provide a reinforce to the more frequent actions, taking into account also agents history.

Agents use a learning algorithm to estimate the worth of each action. Agents will choose their action 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 for the rest of the decisions, the agents deterministically choose the action estimated to be of higher utility. In all the experiments presented in this article, the exploration rate has been fixed at 25%, i.e., one-fourth of the actions are chosen randomly.

The learning algorithm used here is a simplified version of the Q-Learning algorithm [7]. The Q-Update function for estimating the utility on an action is:

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

where $reward$ is the payoff received from the current interaction and $Q^t(a)$ is the utility estimate of action a after selecting it t times. When agents decide not to explore, they will choose the action with higher Q value.

The simulation process for repeated interactions in the agent society is presented in Algorithm 2.

Algorithm 2 Simulation Process.

```
for timesteps do
    for all agents do
        if OBSERVATION then
            Observe neighbors
            Majority Action amongst neighbors determines payoff
        else if INTERACTION then
            Select another partner agent from population
            Selected agents choose an action
            The joint action from the selected agents and their history determines payoffs
            Selected agent(s) use payoff received to update action estimates
        end if
    end for
end for
```

3. Experiments

To evaluate the rate and success of norm emergence we ran experiments with different societal configurations by varying the following system and agent properties:

- **Neighborhood Size:** We study how different neighborhood sizes in a one dimensional lattice affect the process of emergence of conventions.
- **Number of Focal Actions:** We study how the number of focal actions available to the agents affect to the process of emergence of conventions.

The key element to be observed is how by varying the probability of observance of the agents, the convergence time is also affected.

Results reported here have been averaged over 25 runs in simulation with 100 agents. Agents are initialized with uniformly random memories, and initially are unbiased in their action choice. We conclude that a social convention has been reached when 100% of the population choose the same action. Other authors in the literature such [6] or [5] fixed the convergence rate at 90%. However we have observed that with certain reward functions on certain topologies, even after 90% of the society has converged to a convention, it can still switch back to the other convention.

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.

3.1. Focal Actions Effect

3.1.1. 2 Focal Actions

In this first experiment we have set a population located in a one-dimensional lattice with neighborhood size 10 and two focal actions to reach a convention. Experimental results are shown in Fig. 2. On the horizontal axis we can observe different measures of *observance probability*, and the vertical axis represent the convergence time for convention emergence.

These results show that neither extreme option is optimal. We can observe that in situations of full observance or zero observance, the system takes longer to converge. The phenomenon is produced due to the design of the reward function in both situations. In both extreme scenarios, subconventions are created and hard to break. Only in the scenarios where a combined approach is taken, conventions are reached in a faster way. The subconventions that are created when agents interact one on one are easily broken when agents observe the others around and reinforce the majority action.

3.1.2. 10 Focal Actions

However, we can see in Fig 2 that when increasing the number of focal actions, the most efficient strategy is providing agents with the maximum observance capability possible. In general a higher number of focal actions makes the system converge in a slower fashion. Nonetheless, in the situation of full observance the effectiveness of the system is improved with respect to that of 2 focal actions. The explanation for this phenomenon is that when having more than two focal actions, it is harder to obtain a draw in the decision of the majority action of the agents.

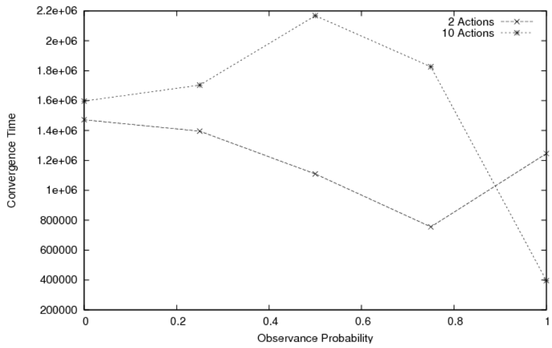


Figure 2. Convergence Time with different Observance probabilities in a One Dimensional Lattice with N=10. Axis y in Decimal Scale.

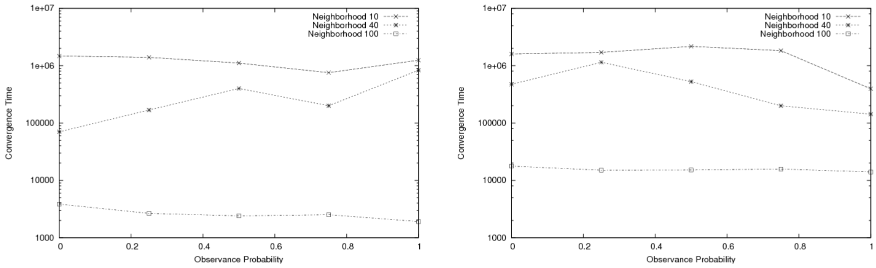
3.2. Neighborhood Size Effect

In this experiment we want to observe the effects of the network topology. We can observe in the experimental results shown in Fig. 3 that the network topology affects directly to the emergence of conventions.

In a previous work [8], we showed that the emergence of conventions in one-dimensional lattices was strongly affected by the neighborhood size: the convergence times are drastically reduced when the diameter [3] of the network (inversely proportional to the neighborhood size) is reduced. Therefore, in the following experiments we have selected three different neighborhood sizes that we believe represent the search space of the underlying topologies: 10, 40 and 100 (representing a fully connected network).

We can observe that the fully connected network is not importantly affected by the observance probability. The main reason is because of the design of the reward function of the observance action: the focal action reinforced will be that shared by the majority action. This reward function in a fully connected network will improve monotonically

[3]The diameter of a graph is the largest number of vertices which must be traversed in order to travel from one vertex to another



(a) 2 Focal Actions

(b) 10 Focal Actions

Figure 3. Convergence Time with different Observance probabilities in a One Dimensional Lattice. Axis y in Logarithmic Scale.

when increasing the observance probability, as agents will take into consideration all agents' preferences.

On the other hand, with less connected topologies ($N = 10$) we can observe how a high observance probability speeds up the convergence with respect to a lower observance probability.

However, the most interesting case is that of $N = 40$: in the 2-focal actions scenario, a higher observance probability increases the convergence time, whereas, in the 10-focal actions scenario, a higher observance probability reduces the convergence time. The main reason why the 10-focal action scenario is reduced when increasing the observance is because there are more focal actions where agents can reach subconventions. Therefore, we can imagine that the same amount of subconventions will be created in the 2-focal actions scenarios, however, in the case of the 10-focal actions these subconventions might be in different focal actions. Having different subconventions on more than 2 focal actions will produce that the alternatives to the majority action (in the 10-focal action scenario) to be in minority, therefore, discriminating rapidly most of the options (as they will be reinforced very rarely).

4. 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 observance of neighboring agents. This social learning framework requires that each agent learns from repeated interaction with anonymous members of the society and from the observances. 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 different observance probabilities on the rate of norm emergence.

Our initial hypothesis was that different observance probabilities would produce different effect on the process of norm emergence. Experimental results confirmed this hypothesis. We have shown that conventions emerge in less time when agents are allowed both to interact and to observe in the right proportion. This proportion will be determined by the diameter of the network and the number of focal actions available.

5. Future Work

The most inminent future work that we plan to perform is to explore more exhaustively the search space of variables. A deeper study will allow us to fully understand under which topological conditions agents need a different value of the observance probability.

One question that we plan to answer in future versions of this work is how different observance ratios will affect in the emergence of conventions. In the experiments presented in this work, the ratio of observance is limited to all the neighbors of each agent; however, we would like to observe how the emergence of conventions is affected when agents can observe further away than its neighbors.

Finally, we plan to extend this research to different topologies. Up to now we have only analyzed different versions of one-dimensional lattices, although we want to observe how topologies like scale free or small world networks respond to this combined approach of interaction and observance.

Acknowledgments

This work was supported by 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. Sandip Sen is partially supported in part by a DOD-ARO Grant #W911NF-05-1-0285.

References

[1] J. Coleman, "Foundations of social theory," August 1998. [Online]. Available: http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0674312260

[2] A. Walker and M. Wooldridge, "Understanding the emergence of conventions in multi-agent systems," in *Proceedings of the First International Conference on Multi–Agent Systems*, V. Lesser, Ed. San Francisco, CA: MIT Press, 1995, pp. 384–389. [Online]. Available: citeseer.ist.psu.edu/walker95understanding.html

[3] J. Delgado, "Emergence of social conventions in complex networks," *Artificial Intelligence*, vol. 141, no. 1-2, pp. 171 – 185, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/B6TYF-46RDC1K-5/2/6a3018fe523d5093b5b4ea60f6053ee9

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

[5] J. Delgado, J. M. Pujol, and R. Sangüesa, "Emergence of coordination in scale-free networks," *Web Intelli. and Agent Sys.*, vol. 1, no. 2, pp. 131–138, 2003.

[6] 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*. Addison-Wesley, 1993, pp. 507–521.

[7] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992. [Online]. Available: http://jmvidal.cse.sc.edu/library/watkins92a.pdf

[8] D. Villatoro, N. Malone, and S. Sen, "Effects of interaction history and network topology on rate of convention emergence," in *Proceedings of Third International Workshop on Emergent Intelligence on Networked Agents (WEIN09) at AAMAS'09*, 2009.