Norm emergence in spatially constrained interactions

Partha Mukherjee
University of Tulsa
partha-mukherjee@utulsa.edu

Sandip Sen
University of Tulsa
sandip@utulsa.edu

Stéphane Airiau
University of Tulsa
stephane@utulsa.edu

ABSTRACT

Behavioral norms are key ingredients that allow agent coordination where societal laws do not sufficiently constrain agent behaviors. Whereas social laws need to be enforced in a top-down manner, norms evolve in a bottom-up manner and are typically more self-enforcing. While effective norms can significantly enhance performance of individual agents and agent societies, there has been little work in multiagent systems on the formation of social norms. We have recently used a model that supports the emergence of social norms via learning from interaction experiences. In our 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. We term this mode of learning *social learning*, which is distinct from an agent learning from repeated interactions against the same player. We are particularly interested in situations where multiple action combinations yield the same optimal payoff. The key research question is to find out if the entire population learns to converge to a consistent norm. In this extension to our prior work we study the emergence of norms via social learning when agents are physically distributed in an environment and are more likely to interact with agents in their neighborhood than those that are further away. The key new results include the surprising acceleration in learning with limited interaction ranges. We also study the effects of pure-strategy players, i.e., non-learners in the environment.

1. INTRODUCTION

Norms or conventions routinely guide the choice of behaviors in human societies. 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" [10][1]. Conventions in human societies range from fashions to tipping, driving etiquette to interaction protocols. Norms are ingrained in our social milieu 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" [1].

While these aspects of norms or conventions have merited in-depth study of the evolution and economics of norms in social situations [6, 13, 19, 20], we are particularly interested in the following characterization: "... we may define a convention as an equilibrium that everyone expects in interactions that have more than one equilibrium." [20]. This observation has particular significance for the study of norms[2] in the context of computational agents. Computational agents often have to coordinate their actions and such interactions can be formulated as stage games with simultaneous moves made by the players [9]. Such stage games often have multiple equilibria [12], which makes coordination uncertain. While *focal points* [14] can be used to disambiguate such choices, they may not be available in all situations. Norms can also be thought of as focal points evolved through learning [20]. Hence, the emergence of norms via learning in agent societies promises to be a productive research area that can improve coordination in and hence functioning of agent societies.

While researchers have studied the emergence of norms in agent populations, they typically assume access to significant amount of global knowledge [6, 13, 19, 20]. For example, all of these models assume that individual agents can observe sizable fraction of interactions between other agents in the environment. While these results do provide key insights into the emergence of norms in societies where the assumption of observability holds, it is unclear if and how norms will emerge if all interactions were private, i.e., not observable to any other agent not involved in the interaction.

To study the important phenomenon of emergence of social norms via private interactions, we have recently used the following interaction framework. We consider a population of agents, where, in each interaction, each agent is paired with another agent randomly selected from the population. Each agent then is learning concurrently over repeated interactions with randomly selected members from the population. We refer to this kind of learning *social learning* to distinguish from learning in iterated games against the same opponent [7]. Most of our experiments involve symmetrical games with multiple pure-strategy equilibria with the same payoff. In previous work on learning in games, the opponent is fixed but in our work, the opponent is different at each iteration. In addition, the opponent may not use the same learning algorithm. It is unclear, *a priori*, if and how a social norm will emerge from such a social learning framework. In a recent paper, we have investigated the ef-

[1]Conventions can therefore be substituted as external correlating signals to promote coordination.

[2]Henceforth we use the term norm to refer to social norms and conventions.

ALAg-07 May 14–18 2007, Honolulu, Hawai'i, USA.

fect of population size, number of choices available, heterogeneous population with multiple learning algorithms, effect of non-learners in shaping norm adoption, etc. when each agent randomly interacts with any other agent in the population. Our experimental results and concomitant analysis throws light on the dynamics of the emergence of norm via social learning with private interactions [16].

The current paper builds on this work and studies norm emergence in more realistic situations where agents are physically distributed in space. In physical environments, e.g., real-life physical interactions between humans in the society, agents are much more likely to interact with those in close physical proximity compared to others located further away. Such physical or spatial interaction constraints or biases have been long well-recognized in social sciences [11] and, more recently, in the multiagent systems literature [15]. In this paper, then, we focus on agents located in a grid world where they interact predominantly with agents in their physical neighborhood. The goal is to evaluate the effects of neighborhood sizes on the rate and pattern of norm emergence. We believe that these results, influenced by spatial interaction constraints, are more representative of real-life phenomena of evolution of norms.

2. RELATED WORK

The need for effective norms to control agent behaviors is well-recognized in multiagent societies [3, 17]. In particular, norms are key to the efficient functioning of electronic institutions [8]. Most of the work in multiagent systems on norms, however, has centered on logic or rule-based specification and enforcement of norms [5, 17]. 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 [2, 3]. 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" [20]. 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.

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 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.

3. SOCIAL LEARNING FRAMEWORK

The specific social learning situation for norm evolution that we consider is that of learning "rules of the road". In particular, we will consider the problem of which side of the road to drive in [3]. We will represent each interaction between two drivers as a 2-person, m-action stage game. These stage games typically have multiple pure strategy equilibria. In each time period each agent is paired for interaction with a randomly selected agent from a subset of the population. An agent is randomly assigned to be the row or column player in any interaction. We assume that the stage game payoff matrix is known to both players, but agents cannot distinguish between other players in the population. Hence, each agent can only develop a single pair of policies, one as a row player and the other as a column player, to play against any other player from the agent population. The learning algorithm used by an agent is fixed, i.e. an intrinsic property of an agent.

When two cars arrive at an intersection, a driver will sometimes have another car on its left and sometimes on its right. These two experiences can be mapped to two different roles an agent can assume in this social dilemma scenario and corresponds to an agent playing as the row and column player respectively. Consequently, an agent has a private bimatrix: a matrix when it is the row player, one matrix when it is the column player. Each agent has a learning algorithm to play as a row player and as a column player and learns independently to play as a row and a column player. An agent does not know the identity of its opponent, nor its opponent's payoff, but it can observe the action taken by the opponent (perfect but incomplete information).

For an iteration of the simulation, each agent plays with one of its neighbors: for each agent i in the population, an opponent is randomly chosen in its neighborhood, the role of i is selected at random (play as a row or a column agents); then agents play the game once, and agent i observe the action of its opponent and updates its learning algorithms. At each iteration, each player plays at least once, and update its learning mechanism exactly once.

4. RESULTS

In this paper we run experiments using the coordination game, where agents receive high payoff for using the same action and otherwise receive a low-payoff (see Table 1). Note that either action combinations (0,0) or (1,1) would work equally well. This matrix can model the problem of which side of the road to drive in. When both agents decide to drive on the same side, there is no collision, which is modeled by a high reward. Otherwise, a collision occur, yielding a low payoff. The goal is then for all agent to develop a norm of choosing the same action consistently.

	0	1
0	4, 4	-1, -1
1	-1, -1	4, 4

Table 1: Payoff in a coordination game.

The agents are distributed over space where each agent is located at a grid point (see Figure 1). An agent is allowed to interact only with agents located within its neighborhood. We consider that the world is a toroid, i.e. agents on one edge are adjacent to the agents located on the opposite edge. The neighbor 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., $|x_1 - x_2| + |y_1 - y_2|$ is the distance between grid locations (x_1, y_1) and (x_2, y_2)). We vary the value of D to allow for different neighborhood sizes.

In this paper we have experimented with a society of N agents placed in a $\sqrt{N} \times \sqrt{N}$ grid. For the figures presented in this pa-

[3]It might seem to the modern reader that "rules of the road" are always fixed by authority, but historical records show that "Society often converges on a convention first by an informal process of accretion; later it is codified into law." [20].

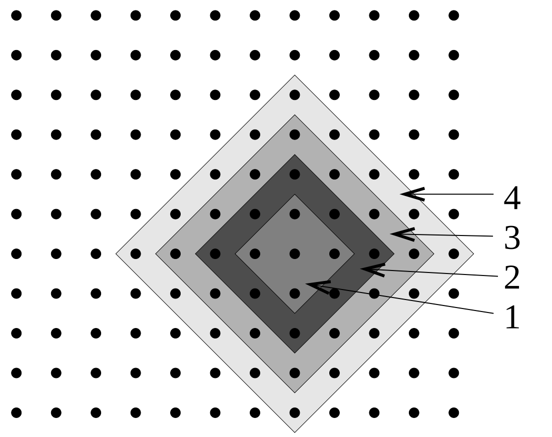

Figure 1: Agents located on a grid and allowed to interact only in a limited neighborhood.

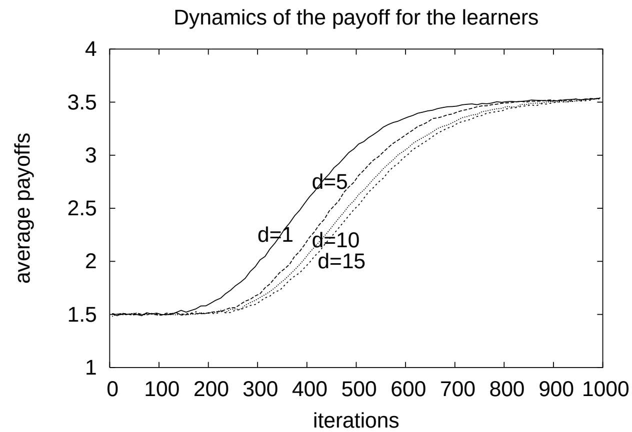

Figure 2: Influence of neighborhood size on learning rate. All agents are learning.

per, we use 225 agents placed on a 15 by 15 grid. We use WoLF-PHC (Win or Learn Fast - policy hill climbing [4]), which can learn mixed strategies. Though WoLF is guaranteed to converge to a Nash equilibrium of the repeated game in a 2-person, 2-actions game against a given opponent, it is not clear whether it is guaranteed to converge in social learning. We have empirically shown convergence when all agents are neighbor of each other in [16].

The results presented are averaged over 50 runs of the experiment with different random seeds.

4.1 Effect of neighborhood size

In this section, we have experimented by varying the neighborhood distance of the agents and the effect of the neighborhood size on learning of agents is observed. We have tested with four neighborhood distances, D, for each agent (the distances are 1, 5, 10, and 15 respectively). When $D = 4$, only an adjacent agent is a neighbor (there are 4 neighbors in that case). For an arbitrary D value, an agent has $(D + 1)^2 - 1$ neighbors. When the distance is 15, every agent is a neighbor of every agent.

We present in Figure 2 the dynamics of the average payoff of the population over a run when all agents are learning concurrently. A payoff of 1.5 is achieved when the agents use a uniform distribution when playing the game. The maximum payoff achievable is 4, and is obtained when the agents play the joint action (0,0) or (1,1). However, as our agents use the ϵ-greedy exploration scheme, they cannot reach 4. We conclude that a norm has emerged in the population when the average payoff of the population reaches 3.5. From Figure 2 we observe that the smaller the neighborhood distance, the faster the emergence of a norm.

When an agent has four neighbors ($D = 1$), the agents learn to coordinate faster by driving on the same side of the roads than when it has 35 or 99 neighbors ($D = 5$ and 10 respectively). For a given number of iterations, the agents interact more often with a particular neighbors for smaller neighborhoods. This means that the impact an agent has on another agent is larger when the neighborhood size is small. In addition, an agent with few neighbors will encounter few different behaviors from its neighbors, and it is *a priori* easier to coordinate with a small set of agents rather than a larger one. As the neighborhood distance increases, an agent has to coordinate with many other agents, and in addition, interactions between two particular neighbors in the network become less frequent. This decreasing interaction frequency between pairs of learners increases the time for exploration of the behavior space and thereby influences the learning patterns of the agents in the network. This problem is exaggerated when every agent is everyone's neighbor ($D = 15$) which further reduces the rate of learning.

Figure 3 represents, for largest ($D = 15$) and smallest ($D = 1$) neighborhoods, the policy of each agent in the population at different iterations in a single run. Each cell represent the policy of an agent: the darker it is, the higher the probability of driving on the left, whereas lighter colors denote higher probability of driving on the right. When a cell is completely dark, or white, it means that the learning algorithm of the agent has converged. In the particular run we present, the norm of "driving on the right" emerges (over different runs "driving on the left" and "driving on the right" norms were evolved in roughly the same number of runs). At iteration 145, the agents are exploring and are receiving low payoff (see corresponding payoff dynamics in Figure 2). At iteration 355, for $D = 1$, we are close to the inflection point for the curve of the payoff dynamics: the agents start to favor one norm over the other. For $D = 15$, however, there is a lesser bias favoring one action. We can see that, on the average, the snapshot for $D = 1$ is lighter than that with $D = 15$. At iteration 480, we can see that many more agents have converged for the smallest compared to the largest neighborhood. So smaller neighborhoods induce faster learning among agents on a grid.

The above effect of agent neighborhood size on learning rate was somewhat surprising. A priori, it was unclear whether smaller neighborhoods will engender divergent norms to initially form over the agent space, which would subsequently delay the convergence of the population to a consistent norm. Such effects, however, were overshadowed by the effects of increased interaction frequencies between neighbors in our framework.

4.2 Influence of non-learning agents

So far, we have observed that all norms with equal payoffs were evolved roughly with the same frequency over multiple runs. This is expected because the payoff matrix for the coordination game (Table 1) has no preference for one norm over the other. Extraneous effects, however, can bias a society of learners towards a particular norm. For example, some agents may not have learning capabilities and always choose a pre-determined action. We now study the influence of agents playing a fixed pure strategy (FPS agent) on the emergence of a norm. We are interested in the effect of multiple pure strategy players with the same or different fixed strategies.

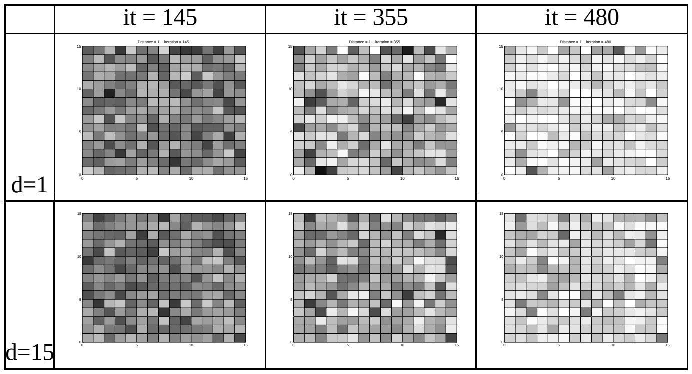

Figure 3: The probabilities of agents driving on the left (Whiter cells represent probabilities close to 1). All agents are learning.

4.2.1 Non-learners use same strategy

In the first experiment, we replace some learning agents by FPS agents and we study the effect of the speed of emergence of a norm. When there are no FPS agents, as the learners explore early in the run, they should encounter each joint action in the same proportion on average. When FPS agents are present, however, learners that have an FPS agent in their neighborhood should observe a bias towards one strategy which the FPS agent always chooses. As agents start to exploit, a learners i that has an FPS agent f in its neighborhood should exploit this bias and consequently, it is more likely to play the action played by f. This bias should also be boosted by i's neighbors which are also in the neighborhood of f. Our hypothesis is that with more FPS agents that play the same action, e.g., all FPS agent wants to drive on the right, the corresponding norm would emerge faster in the population. In Figure 4, we compare the results when there are no FPS agents and either 1, 2, 3, or 4 FPS agents in the population[4]. For these experiments, we used $D = 5$. Note that all the FPS agents play the same action (driving on the right).

The first observation from Figure 4 is that norms do not emerge any faster with only one FPS: the local effect of a single FPS agent is insufficient to expedite convergence to a norm. When there are two or more FPS agents, however, we observed the expected faster norm emergence. With our choice of locations for the two FPS agents, no learner has both FPS agents as neighbors. However, the speed of emergence is faster than with one FPS agent in the population. When there are three FPS agents, some agents have two FPS agents in their neighborhood, which could help them to converge faster. However, this is not the case as we observe a minor effect on the speed of emergence. When there are four FPS agents, more learners have two FPS agents in their neighborhood, and we do observe a positive impact on the speed of emergence. As we had expected, the speed of emergence increases with the number of FPS agents. However, we cannot yet accurately predict the variation of the speed of emergence with number of FPS agents, and we plan to further investigate this issue.

4.2.2 Non-learners use different strategies

In the previous experiment, all FPS agents were playing the same fixed strategy (driving on the right), and they are able to speed up

[4]When there are multiple FPS agents, we located them as far as possible from each other. When there are two FPS agents, they are located at (4,4) and (11,11). When there are three FPS, they are located at (4,4), (7,8) and (11,11). When there are four, they are located at (4,4), (11,11), (4, 11) and (11,4)

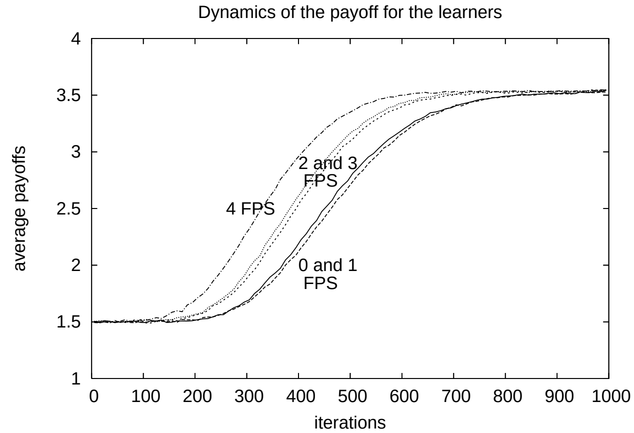

Figure 4: Influence of of non-learners, using identical strategy, on learning rate ($D = 5$).

the emergence of a norm. But FPS agents in practice may be unrelated and adopt conflicting behavior, e.g., some agents alway “drive on the right” and some others always “drive on the left”. In this case, they are likely to decrease the speed of emergence, or even prevent the convergence of a norm in the entire population. In [16], we have observed that two populations that interact infrequently can develop different norms. Hence, it may be possible that FPS agents influence other agents in their neighborhood, hence, different norms emerge in different neighborhoods. In the next set of experiments, we used two FPS agents playing different strategy R (for driving on right) and L (driving on left).

In Figure 5 we present snapshots representing the state of the policy of the agents in the population at different stages of the simulation. The two FPS agents are located at locations (4,11) for R and (11,4) for L. In the two runs, for $D = 1$ and $D = 5$, presented in Figure 5, “driving on the right” is the norm that emerges. We notice that the emergence is faster when the size of the neighborhood is smaller. When the simulation is at iteration 45, the agents are exploring, and the policies of the agents are close to $< 0.5, 0.5 >$. When the simulation is at 535, the population starts to learn and a norm starts to be preferred by a majority of agents. We were expecting that neighbors of the FPS agents will converge to the policy of the near-by FPS agent. But we do not observe this phenomenon, even when the size of the neighborhood is equal to one (for example the agent that is just below the agent choosing L has converged to the norm of R). This may be due to the fact that even with $D = 1$, three of the neighbors are learners, who might ultimately lead the neighbor of L to choose R. We plan to run further experiments to explain this phenomenon. When we ran multiple runs, we observe that each time, the entire population of learning agents converges to a norm: the norms driving on the right and driving on the left emerges with equal frequency. Hence, we did not observe the establishment of multiple norms in these population. This is particularly significant since, with the payoffs we chose (see Table 1), using a single norm in the population maximizes social welfare[5]. Hence, social learning is able to produce social welfare maximizing outcomes even in the presence of non-learners.

[5]If two regions of the population were to adopt distinct norms, the agents at the border and their neighbors would suffer a loss of payoff. When a single norm emerges, only the neighbor of the FPS agents suffer a loss of payoff.

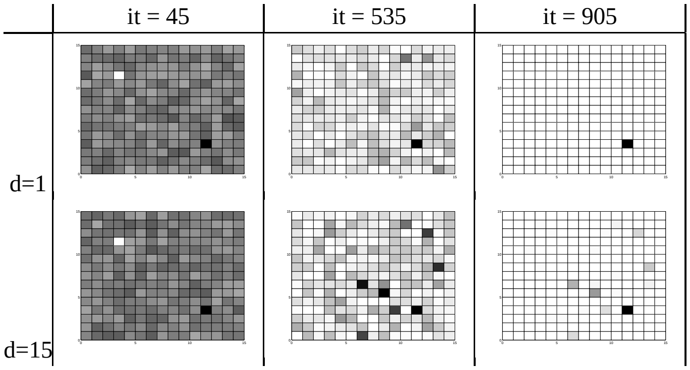

Figure 5: Probabilities of agents driving on the left . Two FPS players play different fixed strategis.

5. CONCLUSIONS

We investigated a bottom-up process for the evolution of social norm that depends exclusively on individual experiences rather than observations or hearsay. Our proposed social learning framework requires each agent to learn from repeated interaction with anonymous members of the society. This is in contrast to most results in multiagent learning where two or more agents learn from repeatedly interacting with the same group. These results confirm that only private experience is sufficient for the emergence of a norm in a society of learning agents. This is in contrast with prior work on norm evolution which requires agents to have knowledge about non-local interactions between other agents and their strategies [6, 13, 19]. Our primary goal in this paper was to evaluate the effect of spatial interaction restrictions on the speed and nature of norms that emerges through social learning. We realized that limiting interactions may isolate sub-populations, thus allowing for different norms to evolve in different parts of the space. Resolving of such emerging conflicts that may reduce social welfare and producing a consensus norm could have been time-consuming. Experimental results, however, clearly demonstrate that agent populations with more restrictions, i.e., those with smaller agent neighborhoods actually produce faster convergence to social norms! This is very likely due to the increased number of interactions between neighbors which allow them to quickly identify mutually-agreed behavior. This neighbor interaction frequency is found to overshadow the effect of time taken to resolve divergent norms. We plan to study this tradeoff more closely to better understand the observed phenomena. We also observed that the social learning framework is able to produce social welfare maximizing policies even in the presence of divergent non-learners.

In this paper, interaction restrictions were binary. Agents were equally likely to interact with any agent in its neighborhood and never interacted with anyone outside. Actually, the neighborhood topology has an interesting characteristic. Within its neighborhood, the agents were more likely to interact with agents at a larger distance than agents situated closer. This is because there are more agents at a larger distance than a smaller distance within a neighborhood, e.g. for $D = 3$, for any agent there are 4 agents at distance 1, 8 agents at distance 2, and 12 agents at distance 3). This is contrary to normal intuition which suggests that the likelihood of interaction with another agent should decrease with the distance to that agent. We plan to run experiments with corresponding interaction preferences and compare results with those presented in this paper. Other interesting experiments include study of different network topologies and the influence of the topology on the speed of emergence of norms. In particular, we would like to evaluate the emergence of norms in social networks.

Acknowledgment: US National Science Foundation award IIS-0209208 partially supported this work.

6. REFERENCES

[1] R. Axelrod. *The complexity of cooperation: Agent-based models of conflict and cooperation*. Princeton University Press, Princeton, NJ, 1997.

[2] G. Boella and L. Lesmo. A game theoretic approach to norms. *Cognitive Science Quarterly*, 2(3–4):492–512, 2002.

[3] 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.

[4] M. Bowling and M. Veloso. Multiagent learning using a variable learning rate. *Artificial Intelligence*, 136:215–250, 2002.

[5] F. Dignum, D. Kinny, and L. Sonenberg. From desires, obligations and norms to goals. *Cognitive Science Quarterly*, 2(3–4):407–430, 2002.

[6] J. M. Epstein. Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18:9–24, 2001.

[7] D. Fudenberg and K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.

[8] A. Garcia-Camino, J. Rodriguez-Aguilar, and C. Sierra. A rule-based approach to norm-oriented programming of electronic institutions. *ACM SIGecom Exchanges*, 5(5):33–41, 2006.

[9] M. Genesereth, M. Ginsberg, and J. Rosenschein. Cooperation without communications. In *Proceedings of the National Conference on Artificial Intelligence*, pages 51–57, Philadelphia, Pennsylvania, 1986.

[10] D. Lewis. *Convention: A Philosophical Study*. Harvard University Press, 1969.

[11] N. M. and M. R. M. Evolutionary games and spatial chaos. *Nature(London)*, 359:826–829, 1992.

[12] R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

[13] M. Posch. Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory*, 65:383–415, 1995.

[14] T. C. Schelling. *The Strategy of Conflict*. Havard University Press, 1960.

[15] F. Schweitzer, J. Zimmermann, and H. Muhlenbein. Coordination of decisions in a spatial agent modela. *Physica A*, 303(1–2):189–216, 2001.

[16] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*, p. 1507–1512, 2007.

[17] J. Vázquez-Salceda, H. Aldewereld, and F. Dignum. Norms in multiagent systems: From theory to practice. *International Journal of Computer Systems & Engineering*, 20(4):225–236, 2005.

[18] C. J. C. H. Watkins and P. D. Dayan. Q-learning. *Machine Learning*, 3:279 – 292, 1992.

[19] P. H. Young. The evolution of conventions. *Econometrica*, 61:57–84, January 1993.

[20] P. H. Young. The economics of convention. *The Jounal of Economic Perspectives*, 10(2):105–122, Spring 1996.