

Effects of Social Network Topology and Options on Norm Emergence

Onkur Sen¹ and Sandip Sen²

¹ Oklahoma School of Science and Mathematics
onkursen@gmail.com

² Department of Computer Science
University of Tulsa
sandip@utulsa.edu

Abstract. A social norm is a behavior that emerges as a convention within society without any direction from a central authority. Social norms emerge as repeated interactions between individuals give rise to biases toward actions or behaviors which spread through the society until one behavior is adapted as the default behavior, even when multiple acceptable behaviors exist. Of particular interest to us is how and when norms emerge in social networks, which provide a framework for individuals to interact routinely. We study how quickly norms converge in social networks depending on parameters such as the topology of the network, population size, neighborhood size, and number of behavior alternatives. Our research can be used to model and analyze popular social networks on the Internet such as Facebook, Flickr, and Digg. In addition, it can be used to predict how norms emerge and spread in human societies, ranging from routine decisions like which side of the road to drive on to social trends such as the *green* phenomenon.

1 Introduction

Recent literature in multiagent systems show a significant increase in interest and research on normative systems which are defined as [6]:

A normative multiagent system is a multiagent system organized by means of mechanisms to represent, communicate, distribute, detect, create, modify, and enforce norms, and mechanisms to deliberate about norms and detect norm violation and fulfillment.

Norms or conventions routinely guide the choice of behaviors in human societies and plays a pivotal role in determining social order [19]. 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” [21]. Norms or conventions can therefore be substituted as external correlating signals to promote coordination (all coordination is choosing a solution from a space of possible solutions).

While these aspects of norms or conventions have merited in-depth study of the evolution and economics of norms in social situations [14, 28, 38, 39], 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." [39]. This observation has particular significance for the study of norms³ in the context of computational agents.

As computational agents in a multiagent system often have to coordinate their actions, adoption and adherence to norms can improve the efficiency of agent societies. A large class of interactions between self-interested agents (players) can be formulated as stage games with simultaneous moves made by the players [18]. Such stage games often have multiple equilibria [23], which makes coordination difficult. While *focal points* [31] 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 [39] that reduces disagreement and promote coherent behavior in societies with minimal oversight or centralized control [9]. Norms can therefore have economic value to agents and help improve their efficiency.

To study the important phenomenon of emergence of social norms via private interactions, we use 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* [22, 33] to distinguish from learning in iterated games against the same opponent [15].

Our experiments involve symmetric games with multiple pure-strategy equilibria with the same payoff. While some research on norm emergence assumed uniform interaction probabilities between agents [33, 34] others have studied the effect of topologies of agent relationships [11, 20, 22]. This body of research can be further divided into *interaction-based learning* approaches to norm emergence [22] and *observation-based adoption* approaches [11]. In the interaction-based learning mode, agents learn utility estimates of their behavior choices and over time converge on a particular behavior that becomes the norm in the society. In this mode, agents actually do not even need to observe others' behaviors. In the observation-based learning mode, agents' behaviors must be fully observable and typically there is no direct consideration of utilities.

We are interested in studying how different network topologies affect the rate of interaction based norm emergence. In particular, we experiment with (a) scale-free networks, (b) fully-connected networks, and (c) ring networks. Within ring networks we further consider the effect of neighborhood sizes, where a neighborhood size of δ implies that any node in the ring can interact with the δ nearest nodes on the ring. Such ring networks represent realistic situations where the agents are physically situated in space and are more likely to interact with other

³ Henceforth we use the term norm to refer to both social norms and conventions.

agents in their physical proximity⁴. Scale-free networks, on the other hand, represent logical connectivity in social networks [2], and represent situations where the node degrees of the network follow a power law distribution [24]. With the explosion in interest in social networking sites like Facebook, MySpace, Flickr, Digg, etc., understanding and exploiting how information is disseminated and how choices are adopted by individuals in social networks have assumed critical significance both to social scientists who want to study this fascinating social phenomena and to companies who want to benefit from mining these interaction data to produce better marketing and advertising tools.

While our previous work [22, 33] has studied the effect of learning algorithms, population biases, physical proximity based interaction likelihood, etc. we have not addressed the issue of scale-up. Whereas studying scaling-up properties of rate of norm convergence for larger agent societies is of obvious interest given rapid growth of user base of social networking sites, we believe that studying norm emergence scale-up properties of different network topologies in face of increasing number of alternate choices or behaviors can also offer key insights to the working of these systems. For example, users face an increasingly diverse set of choices in their everyday interactions, ranging from usage of software and web-based products (email clients, chat facilities, browsers, social networking/blogging sites), availability of TV shows, interaction styles (emoticons, acronyms, etc.). It would be of significant interest and value to better understand how and when the entire society, or any sub-group thereof, adopt a particular choice given a large set of initial choices. Who are the drivers of this adoption, e.g., are the hubs in a social network the drivers or facilitators of the rise of common choices or conventions? While this paper does not purport to answer all of these complex, interrelated issues, we will evaluate how increasing the number of alternative available behaviors affect the rate of norm emergence in the different network topologies described above.

The rest of the paper is organized as follows: in Section 2 we briefly overview relevant literature; in Section 3 we present the network topologies we use in our study; in Section 4 we describe how individuals behave in the network; in Section 5 we analyze our experimental results and their implications; in Section 6, we summarize our results and propose future work.

2 Related work

Norms may be adhered to in human societies because they facilitate the functioning of individuals, or because of the threat of social disapproval [29] or acceptance by individuals of desired conduct [13]. They are self-enforcing: “A norm exists in a given social setting to the extent that individuals usually act in a certain way

⁴ 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 well-recognized in social sciences [26] and, more recently, in multiagent systems literature [32].

and are often punished when seen not to be acting in this way” [3]. 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.

Hence, the systematic study and development of robust mechanisms that facilitate emergence of stable, efficient norms via learning in agent societies promises to be a productive research area that can improve coordination in and thereby functioning of agent societies. Establishment of social norms may come about by top-down influences like official edicts and role models, bottom-up processes driven by local customs, and lateral diffusion of established norms between related interaction types [37]. Most research on norms in multiagent systems focus on the *legalistic view* where norms are used to shape the behavior of open systems without using sanctions to enforce desirable behavior. In this approach norms are typically logically specified using a normative language [16] from which rules of behavior can be automatically derived [10]. Our approach to norm emergence from personal interactions is based on the *interactionist view*, which adopts a bottom-up view of individual adoption of norms because of alignment of goals and utilities between agents in a population [7, 8].

While researchers have studied the emergence of norms in agent populations, they typically assume access to significant amount of global knowledge [14, 28, 38, 39]. For example, all of these models assume that individual agents can observe 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.

Amaral *et al.* study the topology of various “small-world” networks, which encompass scale-free networks [2]. Noble *et al.* study how the topology of a network affects the rate of spreading of information [25]. The need for effective norms to control agent behaviors is well-recognized in multiagent societies [5, 11]. In particular, norms are key to the efficient functioning of electronic institutions [17]. Most of the work in multiagent systems on norms, however, has centered on logic or rule-based specification and enforcement of norms [12]. 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 [4, 5]. While norms can be established by centralized diktat, norms in real-life often evolve in a bottom-up manner, via “the gradual accretion of precedent” [39]. In our formulation, norms evolve as agents learn from their interactions with other agents in the society using multiagent reinforcement learning algorithms [27]. Most multiagent reinforcement learning literature involve two agents iteratively playing a stage game and the goal is to learn policies to reach preferred equilibrium [30]. 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 [35]. The goal of the learning agent is to maximize an objec-

tive function for the entire population, the world utility. While these learning approaches consider the same set of individuals repeatedly interacting and learning, in our framework an agent learns by interacting with different individuals at each time step.

3 Network Topologies

An important property of the topology of a social network is its *diameter*. A social network's diameter is defined as the largest distance between any two nodes in a network. The diameter represents the largest path within the network and characterizes the compactness and connectivity of the network. A network with a small diameter is very well-connected, and thus the average path length of the network will be small. On the other hand, a network with a large diameter will be very sparsely-connected, and the average path length can be large. In addition, a network with a small diameter is more likely to have many different paths between nodes, but a network with a large diameter will have many longer paths between nodes.

Scale-free networks have the structural property that the connectivity of the network follows a power law distribution. This means that the network has a small number of nodes, designated as *hubs*, which have a very high connectivity. However, the most of the nodes in the network are sparsely-connected. A familiar example is the current airport system: small cities do not have very busy airports, but cities like Atlanta, Chicago, and Los Angeles are analogous to the hubs in the network. The diameter of a scale-free network can be approximated as the largest distance among hubs plus 2 since this is the distance between a neighbor of one hub of the longest path and a neighbor of the other hub of the longest path. We use the algorithm presented by Albert and Barabási [1] to generate scale-free networks (the parameters used were as follows: number of initial nodes = 10, number of links to be added or rewired at each step = 3, probability of adding links = 0.4, probability of rewiring links = 0.4).

We also examine *ring networks*, where nodes are connected in a ring. We actually consider generalized rings where each node is linked to all other nodes within a certain distance δ , the *neighborhood size*. Unlike scale-free networks, in which certain nodes dominate the network, in a ring network, all nodes have the same connectivity and thus are equally important in the network.

A special case of a ring network is a *fully-connected network* (also known in graph theory as a *clique*), in which every node is connected to every other node. Therefore, a fully-connected network is a ring network with diameter 1 or neighborhood size equal to the size of the network.

4 Individual Behavior in Networks

In our framework, an agent interacts with a random neighbor at each time interval. An interaction consists of both agents selecting an action (behavior). The first agent receives a payoff based on the action chosen by both agents: if the

actions are identical, the payoff is +4; else, the payoff is -1. 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). Note that only one player gains experience from each interaction. This is to ensure that all agents learn at the same rate, as opposed to agents that are randomly chosen more often learning quicker than those who are chosen less often. We present the protocol of interaction for the entire agent society in Algorithm 1.

```

for a fixed number of time intervals do
  repeat
    remove randomly agents  $p_a$  from the population
    randomly choose  $p_b$ , one of the neighbors of  $p_a$ ,
     $p_a$  and  $p_b$  choose their respective actions;
     $p_a$  updates its utility estimate for the chosen action based on the
    reward received from the joint action
  until all agents have been selected during this time interval ;
end

```

Algorithm 1: Interaction protocol.

Upon receiving the payoff from an interaction, an agent adjusts its estimate of the utility of the action chosen using *Q-Learning* [36]:

$$Q_t(a) = \alpha R + (1 - \alpha)Q_{t-1}(a),$$

where $Q_t(a)$ represents the agent’s Q-valuation of the action a at time t , R is the reward received, and α is a learning rate that weights the current reward with the previous valuation of the action to produce a more accurate approximation to the true valuation. Next, the agents semi-deterministically choose the action estimated to be the most profitable, i.e., the agents will choose the action with the highest valuation most of the time, but with a small probability ϵ (we use $\epsilon = 0.2$) they will instead choose a random action. Note that the choice of action does not change their opinion of which action is the most profitable. Algorithm 2 details the behavior followed by an agent at each time interval:

```

Data: Q-Table  $Q[ ]$ , action  $a$ , Reward  $R$ , Learning Rate  $\alpha$ 
 $R = \text{playRandomNeighbor}()$ ;
 $Q[a] \leftarrow \alpha R + (1 - \alpha)Q[a]$ ;
 $\text{probabilityOfRandomAction} = \text{generateRandomDouble}()$ ;
if  $\text{probabilityOfRandomAction} \leq \epsilon$  then
  |  $\text{action} = Q[\text{generateRandomInteger}(\text{numberOfActions})]$ ;
else
  |  $\text{action} = Q.\text{indexOf}(\max(Q))$ ;
end

```

Algorithm 2: Action selection and Learning algorithm.

5 Results

The performance metric for our experiments was the number of time intervals necessary for convergence, i.e., how quickly a norm emerges. As an example, for

a population $N = 100$ and the number of actions $A = 5$, we examined how many agents were following each action over time until 100% of the population view the same action to be the most profitable, i.e. that action emerged as the norm in the society. We performed experiments for scale-free and ring networks (in the following, unless otherwise specified, we include the results from the completely connected network as a special case of ring network with diameter 1). The set of experiments that we ran are as follows:

Scale-free networks: We studied how varying the number of actions as well as the population size would affect the rate of norm emergence. We varied A over the set $\{2, 5, 10, 20\}$, and N over the set $\{250, 500, 750, 1000\}$.

Ring networks: Within ring networks, we also studied the variation of the number of actions A over the set $\{2, 5, 20\}$, but instead of varying the population size (fixed at 500), we varied the diameter D of the network over the set $\{1, 2, 3, 4\}$.

Comparing Topologies: We compared the convergence speeds of scale-free and ring networks for $A = 2$ and $A = 20$.

5.1 The Norm Emergence Process

Figure 1 shows graphs for how norms emerge in a society of 100 social learners with 5 action choices. Initially, approximately the same number of agents play each action. Over time, however, through agent-agent interactions, a bias toward one action spreads through the entire network until 100% of the population believes that this action is the most profitable, i.e., its Q-value is the highest. At this point (approximately time 70 for the scale-free network and time 100 for the fully-connected network), we say that a norm has emerged. However, this process may sometime demonstrate surprising complexity. For example, in the graph for fully-connected networks, the blue and orange actions rise at approximately the same rate until about time 35 when the blue action dominates the orange action and later emerges as the norm.

5.2 Scale-Free Networks

Figure 2 show the iterations required to reach different levels of consensus in the network, measured as the largest percentage of agents to prefer the same action, for our experiments on scale-free networks. All scale-free network systems converged to a norm; however, the speed of norm emergence varied. A larger number of available actions resulted in a delayed convergence of the system. This is to be expected as a larger number of actions may produce local norms dispersed throughout the network which will lead to many clashes across localities. These conflicting norms must then be resolved and consensus slowly spreads throughout the network until one action emerges as the norm. In addition, a larger population size also delays the emergence of a norm in the network. This is because a larger population size entails that more nodes have to converge to that action based on interaction with other agents.

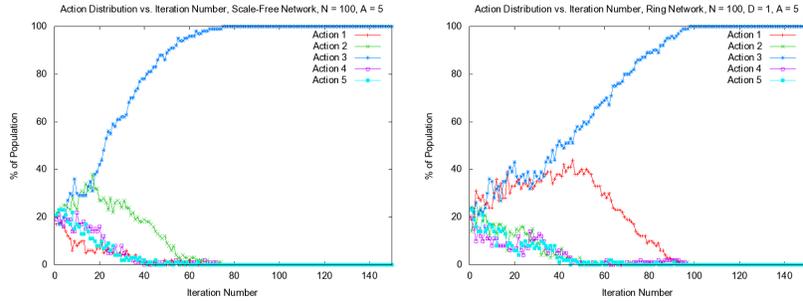


Fig. 1. The process of norm emergence in scale-free (left) and fully-connected (right) networks.

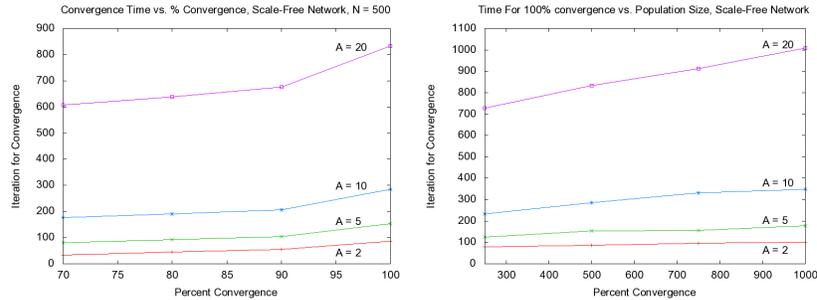


Fig. 2. Iterations required to reach different levels of consensus in scale-free networks with varying number of action options (left) and network size (right).

5.3 Ring Networks

Similar results from ring networks (see Figure 3) shows similar results as scale-free networks in that a larger number of actions resulted in a slower convergence time. In addition, an increase in diameter of the graph causes a near-exponential growth in convergence time. This is because as the diameter of the ring network grows, the spread of bias takes more intermediaries. Hence, biases tend to be confined within a locality until the localities converge on an action, at which point the bias toward an action can spread to other areas of the network. However, a notable exception to this is a comparison between $D = 1$, i.e., a fully connected network, and $D = 2$, where the convergence rates are approximately the same. This is because interacting with everybody may, in some cases, lead to more time taken to form a consensus.

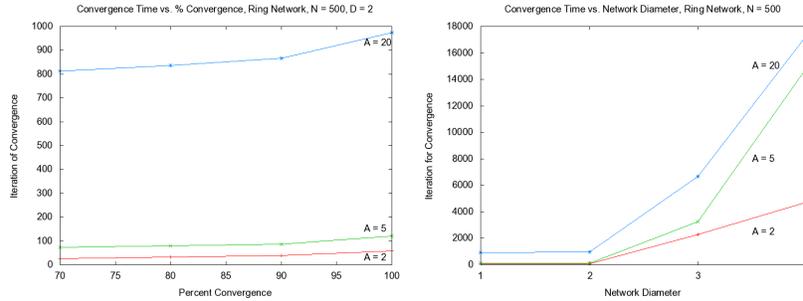


Fig. 3. Iterations required to reach different levels of consensus in ring networks with varying number of action options (left) and network diameter (right).

5.4 Comparing Scale-Free and Ring Networks

Figure 4 shows the comparison of scale-free, fully-connected, and normal ring networks for both a small and large number of actions. For a small number of actions, the ring network converges the fastest, followed by the fully-connected network, and the scale-free network converges last. However, for a large number of actions, the scale-free network converges much faster than the ring and fully-connected networks. The slight difference between the ring and fully-connected networks can be explained as mentioned above in that interacting with *almost* everybody can strike the right balance between developing and propagating biases in some cases when compared to interacting with everybody. The fact that

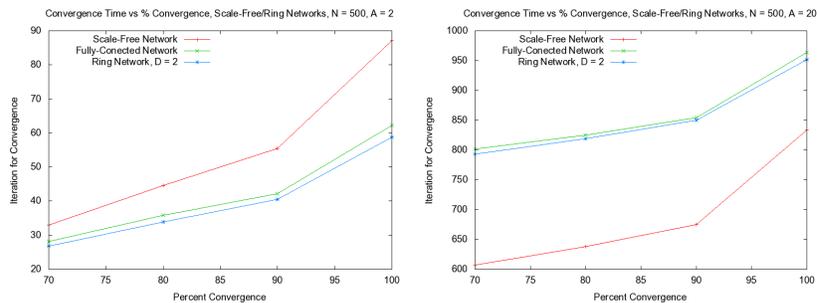


Fig. 4. Comparing convergence speeds of scale-free, normal ring, and fully-connected networks for $A = 2$ (left) and $A = 20$ (right).

scale-free is not that effective for small number of actions but is clearly more efficient for a large number of action choices is a very surprising and intriguing

result. We conjecture a hypothesis that needs to be further refined and tested to explain this phenomenon. Note that the structure of scale-free networks is based around a small number of hubs which dominate the network. However, in a ring network, every node is equally important since each node has the same degree of connectivity. For a small number of actions, the ring networks converge in a relatively small amount of time as clashes have a smaller probability of occurring. However, as the number of actions available increases, the convergence time for the ring networks will increase rather quickly since clashes are more likely to occur within localities. This delays convergence of a node's locality, and the spread of bias throughout the network is slowed, resulting in a longer time for the norm to emergence. However, for scale-free networks, after a certain threshold, the number of actions becomes insignificant since the actions the hubs choose are the ones that will drive the rest of the network. The sparsely-connected nodes will follow the hubs on what action to choose. Therefore, the convergence for a scale-free network is determined by which actions the hubs choose and how quickly the hubs converge and that rate is less affected by an increase in the number of actions available.

6 Conclusions and Future Work

Our research goal was to evaluate how varying topologies of social networks would affect the emergence of norms through interaction-based social learning in these networks. We chose to study scale-free, fully-connected, and ring networks as the topologies. We were particularly interested in understanding the influence of the number of action choices on the rate of norm emergence. An important, counter-intuitive result from our experiments is that although ring networks converge faster for a fewer number of actions, scale-free networks are able to converge faster for a larger number of actions. In addition, we saw the general trend that for both topologies, a larger population size and more actions to choose from delays the emergence of a norm. For future work we first plan to evaluate the hypothesis about the observed phenomena of relative performance of ring and scale-free networks when number of actions is increased. Another interesting set of experiments would be to weigh the experience with each node based on its "status" in the network, e.g., its connectivity. This would mean, e.g., that hub nodes in scale-free networks have more influence than tertiary nodes. We also plan to study the emergence of norms in more topologies, e.g., small-world networks.

References

1. R. Albert and A.-L. Barabási. Topology of evolving networks: Local events and universality. *Phys. Rev. Lett.*, 85(24):5234–5237, Dec 2000.
2. L. Amaral, A. Scala, M. Barthélémy, and H. Stanley. Classes of small-world networks. In *Proceedings of the National Academy of Sciences of the United States of America*, pages 11149–11152, 2000.

3. R. Axelrod. *The complexity of cooperation: Agent-based models of conflict and cooperation*. Princeton University Press, Princeton, NJ, 1997.
4. G. Boella and L. Lesmo. A game theoretic approach to norms. *Cognitive Science Quarterly*, 2(3-4):492-512, 2002.
5. 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.
6. G. Boella, L. van der Torre, and H. Verhagen. Introduction to the special issue on normative multiagent systems. *Autonomous Agents and Multiagent Systems*, 17(1):1-10, 2008.
7. C. Castelfranchi. Modeling social action for AI agents. *Artificial Intelligence*, 103(1-2):157-182, 1998.
8. C. Castelfranchi. Formalising the informal? Dynamic social order, bottom-up social control, and spontaneous normative relations. *Journal of Applied Logic*, 1(1-2):47-92, 2003.
9. J. S. Coleman. Norms as social capital. In G. Radnitzky and P. Bernholz, editors, *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*. Paragon House, New York, NY, 1987.
10. V. T. da Silva. From the specification to the implementation of norms: an automatic approach to generate rules from norms to govern the behavior of agents. *Autonomous Agents and Multiagent Systems*, 17(1):113-155, 2008.
11. J. Delgado, J. M. Pujol, and R. Sanguesa. Emergence of coordination in scale-free networks. *Web Intelligence and Agent Systems*, 1:131-138, 2003.
12. F. Dignum, D. Kinny, and L. Sonenberg. From desires, obligations and norms to goals. *Cognitive Science Quarterly*, 2(3-4):407-430, 2002.
13. J. Elster. Social norms and economic theory. *Journal of Economic Perspectives*, 3(4):99-117, 1989.
14. J. M. Epstein. Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18:9-24, 2001.
15. D. Fudenberg and K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.
16. A. Garcia-Camino, J. Rodriguez-Aguilar, and C. Sierra. Implementing norms in electronic institutions. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 667-673, New York, NY, 2005. ACM Press.
17. A. Garcia-Camino, J. Rodriguez-Aguilar, C. Sierra, and W. Vasconcelos. A rule-based approach to norm-oriented programming of electronic institutions. *ACM SIGecom Exchanges*, 5(5):33-41, 2006.
18. 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.
19. D. Hume. *A Treatise of Human Nature*. Oxford University Press, 1978.
20. J. E. Kittock. The impact of locality and authority on emergent conventions: Initial observations. In *AAAI*, pages 420-425, 1994.
21. D. Lewis. *Convention: A Philosophical Study*. Harvard University Press, Cambridge, MA, 1969.
22. P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained interactions in diverse societies. In *Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 779-786, 2008.
23. R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

24. M. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.
25. J. Noble, S. Davy, and D. W. Franks. Effects of the topology of social networks on information transmission. In *From Animals to Animats 8: Proceedings of the Eighth International Conference on Simulation of Adaptive Behavior*, pages 395–404, 2004.
26. M. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature(London)*, 359:826–829, 1992.
27. L. Panait and S. Luke. Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11(3):387–434, 2005.
28. M. Posch. Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory*, 65:383–415, 1995.
29. E. Posner. *Law and Social Norms*. Harvard University Press, Cambridge, MA, 2000.
30. R. Powers and Y. Shoham. New criteria and a new algorithm for learning in multi-agent systems. In *Proceedings of NIPS*, 2005.
31. T. C. Schelling. *The Strategy of Conflict*. Harvard University Press, 1960.
32. F. Schweitzer, J. Zimmermann, and H. Muhlenbein. Coordination of decisions in a spatial agent model. *Physica A*, 303(1–2):189–216, 2001.
33. S. Sen and S. Airiau. Emergence of norms through social learning. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence(IJCAI'07)*, pages 1507–1512, January 2007.
34. Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94:139–166, 1997.
35. K. Tumer and D. H. Wolpert. Collective intelligence and Braess’ paradox. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, pages 104–109, Menlo Park, CA, 2000. AAAI Press.
36. C. J. C. H. Watkins and P. D. Dayan. Q-learning. *Machine Learning*, 3:279 – 292, 1992.
37. H. P. Young. Social norms. In S. N. Durlauf and L. E. Blume, editors, *The New Palgrave Dictionary of Economics, Second Edition*. Macmillan, London, 2008.
38. P. H. Young. The evolution of conventions. *Econometrica*, 61:57–84, January 1993.
39. P. H. Young. The economics of convention. *The Journal of Economic Perspectives*, 10(2):105–122, Spring 1996.