

# Emergence of Social Norms through Collective Learning in Networked Agent Societies

Chao Yu, Minjie Zhang and Fenghui Ren  
School of Computer Science and Software  
Engineering  
University of Wollongong  
Wollongong, NSW, 2522, Australia  
cy496@uowmail.edu.au,  
{minjie,fren}@uow.edu.au

Xudong Luo  
Institute of Logic and Cognition  
Sun Yat-sen University  
Guangzhou, Guangdong, 510275, China  
luoxd3@mail.sysu.edu.cn

## ABSTRACT

Social norms play a pivotal role in sustaining social order by regulating individual behaviors in a society. In normative multiagent systems, social norms have been used as an efficient mechanism to govern virtual agent societies towards cooperation and coordination. In this paper, we study the emergence of social norms via learning from repeated local interactions in networked agent societies. We propose a collective learning framework, which imitates the opinion aggregation process in human decision making, to study the impact of agent local collective behaviors on norm emergence in different situations. In the framework, each agent interacts repeatedly with all of its neighbors. At each step, an agent first takes a best-response action towards each of its neighbors and then combines all of these actions into a final action using ensemble learning methods. We conduct extensive experiments to evaluate the framework with respect to different network topologies, learning strategies, numbers of actions, and so on. Experimental results reveal some significant insights into norm emergence in networked agent societies achieved through local collective behaviors.

## Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—*Multiagent systems*; I.2.6 [Artificial Intelligence]: Learning; J.4 [Computer Applications]: Social and Behavioral Sciences—*Sociology*

## General Terms

Experimentation, Design

## Keywords

Norm Emergence, Multiagent Learning, Collective Decision Making, Social Networks

## 1. INTRODUCTION

Social norms, such as driving on a particular side of the road, tipping in restaurants and not littering in parks, are

**Appears in:** *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May 6–10, 2013, Saint Paul, Minnesota, USA.

Copyright © 2013, International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). All rights reserved.

ubiquitous in our daily life, and play a pivotal role in regulating and guiding individual behaviors in human societies. Conformity to norms can reduce social conflicts, mitigate cognitive load on humans and thus sustain social order in the whole society [21]. In recent years, researchers from normative Multiagent Systems (MASs) have used social norms as an efficient mechanism to regulate agent behaviors in virtual societies such as electronic institutions [6], agent-supported virtual enterprises [7] and norm-governed Ad-hoc networks [1]. Due to the expense and inefficiency of having a centralized policing enforcer to formulate and specify social norms in a prescriptive manner, it is more desirable to enable social norms to evolve and emerge on their own without relying on any centralized authority. Understanding what mechanisms can contribute to the emergence of social norms is of great interest in the research of normative MASs [9, 21, 25].

Learning is a robust mechanism to facilitate the emergence of social norms in a bottom-up manner for distributed agents [19]. Some researchers have thus focused on equipping agents with a learning capability to establish a norm for an agent society. For example, Sen *et al.* [21, 14] proposed a framework for the emergence of norms through social learning based on private local interactions; and Savarimuthu *et al.* [18] discussed three aspects of active learning of norm emergence in agent societies. All these investigations handled the issue of norms in the context of an unstructured agent society and showed that a random learning by agents can be efficient for the emergence of robust social norms without relying on a centralized authority.

In the real world, however, people often interact with each other under some physical constraints. In other words, who meets whom is not random, but is determined by some spatial relationship or social network [15]. Actually, social network provides the topology and the infrastructure through which social norms can be exchanged and influenced, and thus are fundamental in determining the process of norm emergence. For this reason, a number of researchers have studied norm emergence by considering the underlying network topology of the agents. For example, Sen *et al.* [20] studied how varying topologies of social networks would affect norm emergence in scale-free, fully-connected and ring networks; Villatoro *et al.* [25, 26] investigated the effects of the memory of past activities during learning on norm emergence in different network structures, and used two social instruments (i.e., rewiring and observation) to facilitate norm emergence in networked agent societies.

All these studies, however, are based on a simple interaction protocol: each agent must be paired for interaction with one of its neighbors, randomly or preferentially, so that this agent can directly learn from the interaction. This interaction protocol simplifies real-life situations when individuals can collectively make a decision from multiple alternatives before them. This collective decision making is inherent in human nature because people often seek several opinions before making a final decision [16]. To reach a group consensus, people often interact with others at the same time and learn simultaneously from all these interactions. It is not clear, however, if in this collective decision making context, a norm will still emerge successfully, and what impacts this collective decision making will have on norm emergence under different conditions (e.g., different network topologies or heterogeneity of agents).

To answer these questions, this paper proposes a collective learning framework to study the emergence of social norms in networked agent societies. In this framework, norms evolve as agents learn over repeated interactions with their neighbors using multiagent reinforcement learning algorithms [23]. Each interaction is framed as a stage game, which has multiple equilibria. These multiple equilibria make coordination between agents uncertain. At each time step, an agent chooses a best-response action for each of its neighbors and aggregates all of these actions into an overall action using a number of ensemble techniques. The agent then plays the aggregated action with all its neighbors and receives a corresponding reward towards each neighbor. Finally, the learning information regarding each neighbor will be updated using the reward. The agent cannot observe its neighbors' payoffs but it can observe the neighbors' current actions (perfect but incomplete information). This makes our work differ from most existing work on norm emergence in the framework of evolutionary game theory [12, 24], in which each agent can observe its neighbors' payoffs so that the population can evolve according to specific strategies (e.g., Imitate-Best-Neighbor). We investigate a number of key issues such as neighborhood and population size, agent cognitive deficiency and learning strategy, and their influences on norm emergence under the collective learning framework.

The remainder of the paper is organized as follows. Section 2 describes the collective learning framework in networked agent societies. Section 3 presents the detailed ensemble learning methods. Section 4 shows the experimental studies. Section 5 discusses related work. Finally, Section 6 concludes the paper with some directions for future research.

## 2. COLLECTIVE LEARNING FRAMEWORK

Before introducing the proposed learning framework, we first give formal descriptions of a networked agent society and an agent's interacting neighbors.

**Definition 1.** A **networked agent society** can be represented as an undirected graph  $G = (V, E)$ , where  $V = \{v_1, \dots, v_n\}$  is a set of vertices (agents), and  $E \subseteq V \times V$  represents a set of edges, each of which connects two interacting vertices (agents).

**Definition 2.** Given a networked agent society  $(V, E)$ , the **neighbors** of agent  $i$ , which are denoted as  $N(i)$ , are a set of agents so that  $N(i) = \{v_j \mid (v_i, v_j) \in E\}$ , and  $N(i) \subset V$ .

This paper focuses primarily on the following three types of topologies to represent a networked agent society.

(1) **Grid networks.** A grid network is a two-dimensional lattice with four neighbors for each inner node, three neighbors for each boundary node and two neighbors for each corner node. In reality, parallel computing clusters and multi-core processors are usually organized as a grid network. We use  $GR_N$  to denote a grid network, where  $N$  stands for the number of nodes.

(2) **Small-world networks.** This kind of network is proposed to represent the small-world phenomenon in many social and computer networks, where each node has only a small number of neighbors and yet can reach any other node in a small number of hops. Small-world networks feature a high clustering coefficient and a short average path length. This kind of networks appears in many real-world social networks such as the collaboration networks of film actors and the friendship networks of students [3]. We use  $SW_N^{k,\rho}$  to denote a small-world network, where  $k$  is the average neighborhood size of a node and  $\rho$  is the re-wiring probability to indicate the different orders of network randomness.

(3) **Scale-free networks.** This kind of network is characterized by the power law of degree distribution of nodes, which means that a few "rich" nodes have high connectivity degrees, while the remaining nodes have low connectivity degrees. The probability that a node has  $k$  neighbors is roughly proportional to  $k^{-\gamma}$ . Examples of scale-free networks in the real-world include the network of citations of scientific papers [17] and links between web pages on the World Wide Web [3]. These real networks exhibit the feature of "preferential attachment", which means that the likelihood of connecting to a node depends on the connectivity degree of this node. We use  $SF_N^{k,\gamma}$  to denote a scale-free network.

---

### Algorithm 1: The collective learning framework

---

```

1 Initialize network and learning parameters;
2 for each step  $t$  ( $t=1, \dots, T$ ) do
3   for each agent  $i$  ( $i=1, \dots, n$ ) do
4     for each neighbor  $j \in N(i)$  of agent  $i$  do
5       Agent  $i$  chooses a best-response action  $a_{i \rightarrow j}$ 
       regarding agent  $j$  using a learning policy
       with exploration;
6       Agent  $i$  combines all the actions  $a_{i \rightarrow j}$  into
       action  $a_i$  using ensemble methods;
7       Agent  $i$  plays action  $a_i$  with all neighbors and
       receives reward  $r_i^j$  for each interaction;
8       Agent  $i$  updates learning information towards
       each neighbor using action-reward pair  $(a_i, r_i^j)$ ;

```

---

The sketch of our collective learning framework is given by Algorithm 1. All agents in the society interact repeatedly and simultaneously with all their neighbors. An agent is not assumed to know its neighbors' payoff. The only available information is the immediate reward from the environment after each interaction and each neighbor's played action. At each time step, an agent chooses a best-response action for each neighbor. The actions for all the neighbors are then aggregated to an overall action using a number of ensemble methods (which will be described in detail in Section 3). The agent then plays the aggregated action with all of its neighbors and receives a corresponding reward. Finally, the

learning information for each neighbor will be updated by using the aggregated action and the corresponding reward.

Our collective learning framework is significantly different from the pairwise learning framework that has been adopted in most previous studies [20, 25, 26]. In the pairwise learning framework, at each time step, each agent is randomly paired with one of its neighbors for interaction and the agent directly learns from this interaction either through a best response rule [20] or a memory-based rule [26]. Our collective learning framework, however, imitates the opinion aggregation process in human decision making because people usually seek several opinions before making a final decision [16]. As the final decision of an agent is affected by all its neighbors, this collective learning can have a significant influence on the emergence of social norms in the whole society in different conditions (e.g., different network topologies, ensemble methods or heterogeneities of agents).

To ensure an optimal policy, exploration is required in the learning process. In Algorithm 1, the exploration process is conducted during an agent's local interactions with each of its neighbors. This means that the agent conducts exploration for each neighbor before combining all the actions together into a final deterministic action. We call this kind of exploration *local exploration*. The exploration process can also be conducted globally, which means that the agent determines the greedy action according to the learning information for each of its neighbors and then aggregates all these actions into an overall action. Exploration is then conducted when the agent chooses a final action based on the aggregated overall action. We call this kind of exploration *global exploration*. In this paper, we adopt the  $\varepsilon$ -greedy policy as the basic exploration policy.

### 3. ENSEMBLE METHODS IN AGENT LEARNING

The basic idea behind ensemble methods is to weigh several individual classifiers first and then combine them in order to make a final decision that will be better than the one made by each of them separately [16]. Although ensemble methods are used with the same aim of increasing learning speed and improving final performance, they have been employed in different forms in a reinforcement learning setting. The different ways in which ensemble learning is used lies in the different ways of defining the action choices. For example, the action choices can be defined as different learning algorithms [29], diversified function approximations in terms of neural network topologies and weights [8], or state-value functions [11]. In our framework, however, the actions that need to be aggregated are the focal agent's best-response actions for each of its neighbors in every interaction. The agent needs to combine all these actions to make a final decision by considering each neighbor's position (e.g., degree of connectivity) as well as the neighbor's performance in past interactions. The ensemble learning imitates the human collective decision making process in that people usually consult with many others before they make a final decision. They consider not only these people's individual characteristics, such as intelligence and knowledge, but also their reputation, their position and power in the society.

Formally, let  $a_{t,j}^*$  be the best-response action regarding neighbor  $j$  at time  $t$  and let  $a_t$  be the aggregated final action. We enumerate the set of possible actions for each agent as

$A = \{a[1], \dots, a[m]\}$ , where  $m$  is the number of actions available. The selection of this final action is then denoted as  $\pi_t(a[i])$ ,  $a[i] \in A$ . The value of  $p_t(a[i])$  represents the focal agent's preference for the action  $a[i]$ . The final action  $a_t$  can then be determined by:

$$\pi_t(a[i]) = \begin{cases} 1 & \text{if } a[i] = \arg \max_a p_t(a), \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The calculation of  $p_t(a[i])$  uses the following methods:

(1) **Majority voting.** The preference values are calculated by the majority voting ensemble method as follows:

$$p_t(a[i]) = \sum_{j=1}^n I(a[i], a_{t,j}^*) \quad (2)$$

where  $n$  is the number of neighbors and  $I(a[i], a_{t,j}^*)$  is an indicator function defined by:

$$I(a[i], a_{t,j}^*) = \begin{cases} 1 & \text{if } a[i] = a_{t,j}^*, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

The most preferred action is simply the one that is suggested by most of the neighbors. The principle of this method reflects the fact that people are social beings and can be influenced by each other so that people are more prone to accept the opinion or strategy that is adopted by the most/majority of their neighbors.

(2) **Weighted voting.** The majority voting method simply counts the number of each action as the preference for corresponding action. However, each agent in the network can occupy different positions and thus can play a different role in shaping the norms of the whole society. For example, in scale-free networks, the power law distribution of the degree of connectivity of the nodes means that a few "rich" nodes can have high degrees of connectivity. The decisions of these powerful nodes are pivotal in the society. Thus, it is necessary to consider the "social ranks" of different agents in the calculation of the preference for each action. Assume that the decision from learner (neighbor)  $j$  is weighed by weight  $w_{t,j}$ . The weighted voting method can be given by:

$$p_t(a[i]) = \sum_{j=1}^n w_{t,j} I(a[i], a_{t,j}^*) \quad (4)$$

Several different ways can be used to determine weights  $w_{t,j}$ . Here, we propose two approaches as follows:

- **Structure-based approach.** This approach considers the different structural position of each agent in the network. A straightforward way of defining the structure-based weight of each agent is to use the agent's degree of connectivity. Therefore, the weight  $w_{t,j}$  of neighbor  $j$  can be calculated as follows:

$$w_{t,j} = \frac{N_{nei,j}}{\sum_i N_{nei,i}} \quad (5)$$

where  $N_{nei,j}$  is the connectivity degree of agent  $j$  and  $n$  is the number of neighbors of the focal agent.

- **Performance-based approach.** This approach determines each neighbor's weight according to past interaction experience between this neighbor and the focal agent. If a neighbor's action is always consistent

with the agent's own action, the agent will then consider the neighbor to be more trustworthy and accordingly assign a higher weight to this neighbor. This is driven by the fact that in the real world a person who has a higher reputation will have a greater influence on biasing the opinions in a society. So, we have:

$$w_{t,j} = w_{t-1,j} + \beta(s - w_{t-1,j}) \quad (6)$$

where  $w_{0,j} = \frac{1}{N_{nei,j}}$  ( $N_{nei,j}$  is the number of neighbors of agent  $j$ );  $\beta$  is a learning rate to adjust the weight; and  $s = 1$  if interaction at time  $t - 1$  is successful, otherwise  $s = 0$ .

## 4. EXPERIMENTAL STUDIES

This section presents the experimental studies. First, we give the basic settings of the experiments. We then present the results and analysis by evaluating the proposed framework in a number of different settings.

### 4.1 Experimental Setting

The learning framework for norm emergence is proposed to study how agents can learn to establish a social convention/law in a networked agent society via local collective decision making. A social convention/law (i.e., a restriction on the set of actions available to agents) is said to have been established when all (or at least the majority of) agents in the society have complied with the same action [22]. In this study, we use learning "rules of the road" [21, 30] as a metaphor to study the emergence of norms. In this scenario, agents strive to establish a convention/law of driving either on the left (L) side or on the right (R) side of the road. This interaction can be viewed as a pure coordination game [30] with the payoff matrix as displayed in Table 1.

**Table 1: Payoff matrix of the Coordination Game**

	Left (L)	Right (R)
Left (L)	1,1	-1,-1
Right (R)	-1,-1	1,1

Although its payoff matrix appears simple, the coordination game poses a very challenging puzzle for human beings to solve efficiently. The game has two pure Nash-equilibria: both agents drive left or both agents drive right. Classical game theory, however, does not give a coherent account of how people would play a game like this. The problem is that there is nothing in the structure of the game itself that allows the players (even purely rational players) to infer what they ought to do. In reality, people can play such games because they can rely on some contextual cues to agree on a particular equilibrium [30]. One such contextual cue is social norms (i.e., conventions and laws) that can be used to guide agent behavior when moral or rational reasoning does not provide clear guidance because of the myopic behavior and the limited processing ability of individuals [5].

The purpose of this experiment is to study the emergence of social norms in different agent societies (e.g., with different topologies, agent learning strategies or population sizes) by using the proposed collective learning framework. The performance metrics are the convergence ratio of the social norms (i.e., how many agents in the society can reach a final consensus after a period of interaction) and the time needed to reach such a consensus (i.e., how quickly a social norm emerges). A social norm is said to be established when at

least 90% of the agents have adopted the same action. We let  $T_{90}$  denote the convergence time when such a norm emerges. We use the Watts-Strogatz model [28] to generate a small-world network, and use the Barabasi-Albert model [3] to generate a scale-free network. We start with 5 agents and add a new agent with 1 edge to the network at every time step. This network evolves into a scale-free network following a power law with an exponent  $\gamma = 3$ . In this study, unless stated otherwise, we use the small-world network as the default network topology due to the variety of this kind of network, Q-learning as the learning strategy with learning rate of 0.1, local exploration as the exploration mode and majority voting as the ensemble method.

## 4.2 Results and Analysis

### 4.2.1 Convergence of social norms

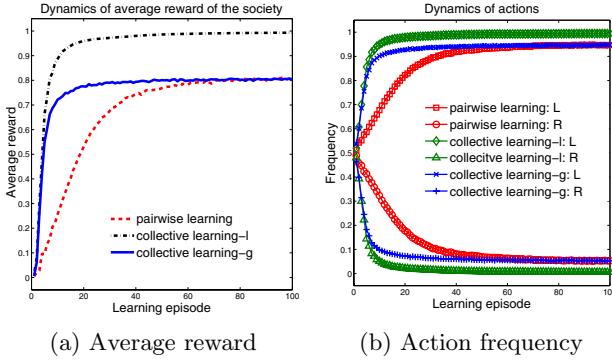
Some previous studies (e.g., [14, 21]) have shown that a social norm can always emerge when each agent learns randomly from another agent in the population. Other studies (e.g., [15, 20, 25, 26]), which consider a networked interaction topology of the agents, have demonstrated that a robust social norm can also evolve successfully when each agent plays in pairs with a random neighbor. It is not clear, however, whether a social norm can successfully emerge in the whole society under the collective learning framework. If a social norm does emerge, what is the ratio of such an emergence? Can the social norm emerge more quickly when using collective learning than when using pairwise learning? To answer these questions, we first test the proposed framework in small-world network  $SW_{100}^{12,0.8}$  and compare it with the pairwise learning framework to demonstrate the merits of our collective learning framework. We ran 1000 independent runs and the overall results are shown in Table 2.

**Table 2: The norm emergence during 1000 runs in a small-world network with 100 agents**

	Norm(L)	Norm(R)	Success(%)	Reward	Speed( $T_{90}$ )
Pairwise learning	506	494	100	0.80	33
Collective learning-l	493	507	100	0.99	7
Collective learning-g	522	478	100	0.80	10

Table 2 shows the frequency and success ratio of converging to a social norm, the final average agent reward for the society and the time steps needed to evolve a norm. From Table 2, we can see that a social norm can emerge under the three different frameworks (with 100% probability) and that agents converge to norm (Left) and norm (Right) with an almost equal probability. This makes sense because the game structure itself does not give any preference for any particular action to be the norm. The norm to which the whole population of agents converges depends on which norm is detected by the agents at first glance during the dynamics of agent interaction. The average reward in the society using collective learning with local exploration (which is denoted as collective learning-l) is almost 1, which means that nearly all the agents have reached a consensus on which action should be the norm. The rewards using collective learning with global exploration (which is denoted as collective learning-g) and pairwise learning are much lower than that using collective learning with local exploration because agents are exploring the environment with a probability of  $\varepsilon = 0.1$ . However, as agents using collective learning-l explore the environment locally and make a final decision collectively, the uncertainties caused by the exploration de-

crease. The time steps needed for 90% of the agents to choose the same action as a social norm under three learning frameworks differ dramatically. The collective learning framework is able to evolve a norm much faster than the pairwise learning. This faster emergence of norms indicates the merits of collective learning against pairwise learning.

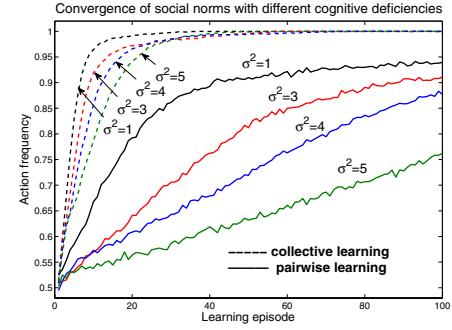


**Figure 1: Emergence of social norms in  $SW_{100}^{12,0.8}$  (Results are averaged over 1000 independent runs).**

Figure 1 (a) shows the dynamics of the average reward of the whole population. Figure 1 (b) shows the frequency of each action adopted by the agents when norm (L,L) emerges in the population. Initially, each agent randomly chooses an action, so there are about 50% of the agents choosing action L and the other 50% choosing action R. This results in the average payoff of 0 in the whole population. As the learning process moves on, however, the number of agents who choose action L as the norm increases. This means that more and more agents have reached a consensus on which action should be the norm, and this consensus correspondingly increases the average payoff dramatically. From Figure 1, we can also see that the norm emerges faster under the collective learning framework than under the pairwise learning framework. This is because agents using collective learning can learn directly from all of their neighbors at the same time to decrease the diversity among the agents. These results confirm that our collective learning framework is an efficient mechanism for the emergence of social norms.

#### 4.2.2 Influence of agents' cognitive deficiencies

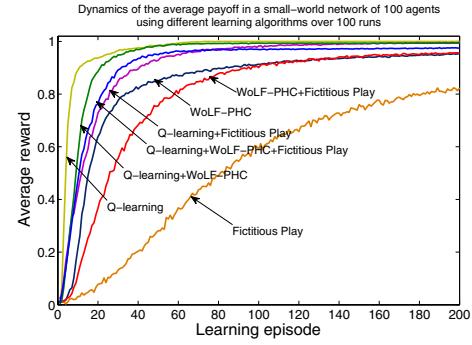
To better understand the advantages of the collective learning framework over the pairwise learning one, we test both frameworks in heterogeneous societies where the agents have varying cognitive capabilities by receiving feedbacks from the environment with different levels of uncertainty. Each agent has a probability of  $p_c$  to receive a transformed payoff, which is  $r \pm \sigma^2$  ( $r = \pm 1$  is the original payoff). The probability of  $p_c$  and the variant of payoff value  $\sigma^2$  indicate the different cognitive capabilities of the agents. We set  $p_c$  to 0.2 and choose  $\sigma^2$  from the set of {1, 3, 4, 5} in this study. Figure 2 shows the dynamics of the action frequency in  $SW_{100}^{12,0.8}$  with different cognitive deficiency  $\sigma^2$ . From the results, we can see that a society in the collective learning framework is able to maintain a high level of convergence ratio and quick convergence speed, whereas a society in the pairwise learning framework is only able to achieve a low level of convergence ratio and slow convergence speed. In both cases, as the cognitive deficiency  $\sigma^2$  increases, the convergence speed slows down. This is because it is more difficult for the agents in



**Figure 2: Dynamics of the action frequency with different cognitive deficiencies.**

societies with larger values of  $\sigma^2$  to distinguish the effects of their actions on the environment, and these uncertainties can hinder the agents from reaching consensuses. However, the difference in the convergence speed using collective learning is not as significant as that using pairwise learning. This indicates that the collective learning framework can mitigate the uncertainties caused by the agents' cognitive deficiency, and is more efficient and robust for norm emergence compared with the pairwise learning framework.

#### 4.2.3 Influence of agent learning strategies



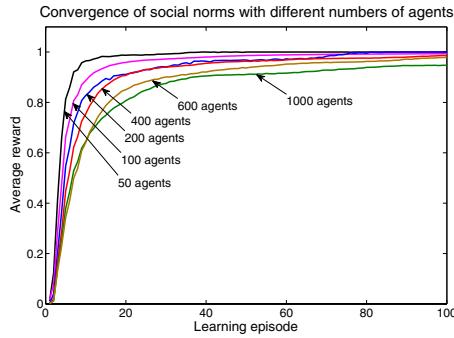
**Figure 3: Dynamics of the average reward when using different learning strategies.**

We are also interested in whether the proposed framework is robust enough for different learning strategies adopted by the agents. Thus, we use three basic learning strategies for agent interaction: Q-Learning [27] with  $\epsilon$ -greedy exploration, WoLF-PHC (Win or Learn Fast-policy hill climbing) [4] and Fictitious Play (FP) [10]. Q-learning has been widely used in MAs, but converges only to pure strategies. The learning rate  $\alpha$  is set to 0.1 in this study. WoLF-PHC can learn mixed strategies and is guaranteed to converge to a Nash equilibrium in a 2-person, 2-action game against a given opponent. However, it is not clear whether it is guaranteed to converge in the collective learning framework. The learning rate  $\alpha_w$  is set to be 0.04 when the agent is winning and the learning rate  $\alpha_l$  is set to be 0.01 when the agent is losing. Finally, an FP player uses the historical frequency count of its opponent's past actions and tries to maximize the expected payoff by playing a best response to that mixed strategy, represented by this frequency distribution. The learning rate is set to be 0.1 for an FP player. We test the three different learning strategies in a homogeneous and a

heterogeneous society of 100 agents. In the homogeneous society, all the agents use the same learning strategy when interacting with their neighbors, while in the heterogeneous society, agents are equally divided according to the strategies they adopt. This heterogeneity of society models the real-life situations, in which people have different learning capabilities in the same circumstance.

Figure 3 shows the dynamics of the average reward in the society using the three learning strategies. As we can see, societies in the collective learning framework can evolve a social norm using all three learning strategies. The quickest one is using Q-learning, followed by WoLF-PHC and Fictitious play. Norms evolve very slowly using Fictitious play because agents need a great deal of time to estimate the frequency distribution of neighbors' past actions. The time in the heterogeneous society to evolve a norm falls between the time taken by the corresponding homogeneous societies. These results are consistent with the previous study [21], in which agents learn randomly in an unstructured population.

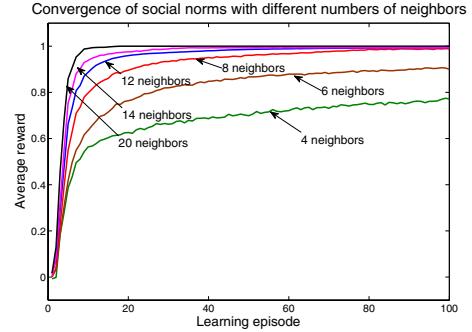
#### 4.2.4 Influence of population size, number of neighbors and actions



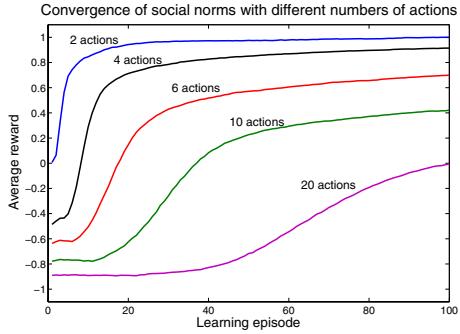
**Figure 4: Dynamics of the average payoff with different numbers of agents.**

The population size of the society and the number of neighbors are important factors that can influence the emergence of social norms [21, 26]. To show the effects of these two factors on norm emergence under the collective learning framework, we vary the agent number  $N$  in  $SW_N^{12,0.8}$  in  $[50, 1000]$ , and the number of neighbors  $k$  in the set of  $\{4, 6, 8, 12, 20\}$  in  $SW_{100}^{k,0.8}$ . The dynamics of average agent reward with different agent populations is shown in Figure 4, from which we can see that the more agents in the society, the longer it takes for the entire society to converge to a social norm. This result occurs because the larger the society, the more difficult to diffuse the effect of local learning to the whole society. This phenomenon can be seen in human societies where small groups and clans can more easily establish social norms than those larger societies, as argued in [21].

Figure 5 shows the dynamics of average agent reward with different neighborhood size  $k$ , from which we can see that when the average number of neighbors is increased, the convergence time is steadily reduced. This effect is due to the clustering coefficient of the network. When the average number of neighbors increases, the number of links between agents also increases, and therefore agents located in different parts of the network only need a smaller number of interactions to reach a consensus.



**Figure 5: Dynamics of the average reward with different numbers of neighbors.**



**Figure 6: Dynamics of the average reward with different numbers of actions.**

It has been shown that the number of actions available to the agents is an important factor in norm emergence [21, 20]. Thus, we vary the number of actions in the set  $N_a = \{2, 4, 6, 10, 20\}$  to investigate its impact on norm emergence. Only when two agents choose the same action will they receive a payoff of 1. Otherwise, they receive a payoff of  $-1$ . Figure 6 shows the effect of different numbers of actions on norm emergence. As can be seen, a larger number of available actions results in a delayed convergence of norms. This is because a larger number of actions may produce more varied local sub-norms, leading to diversity across the society. It thus takes a longer time for the agents to eliminate this diversity to achieve a final consensus, and thus norm emergence is prolonged throughout the network.

#### 4.2.5 Influence of ensemble methods

Figure 7 shows norm emergence using different ensemble methods as well as pairwise learning method in three different kinds of networks. In the grid and small-world network, the majority voting method and structure-based method outperform the pairwise learning method throughout the whole learning period. Norms using the performance-based method converge very slowly at the beginning and then quickly outperform those using the pairwise learning method afterwards. In the scale-free network, however, the three methods under the collective learning framework have almost the same performance, and all outperform the pairwise learning method throughout the whole learning period. These results show that the proposed ensemble methods can bring about different patterns of norm emergence in the three different kinds of networks, and further confirm that our collective learning framework is more efficient for norm emergence than the pairwise learning framework.

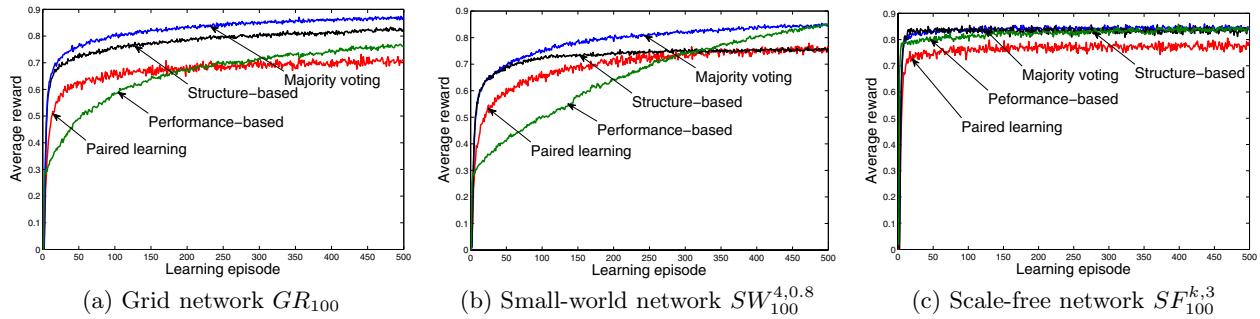


Figure 7: Norm emergence in three networks using different ensemble methods (averaged over 100 runs).

## 5. RELATED WORK

The emergence of social norms has gained increasing attention in the area of MAs. Much work on norms has been done using simulation models, analysis methods and theoretical concepts from MAs [19]. Most of the work, however, focuses on the logic, rule-based specification and declaration of norms by assuming a centralized authority and complete knowledge [21]. Comparatively, not much work has been done on the decentralized emergence of social norms in a bottom-up manner via learning from local interactions. Shoham and Tennenholtz [22] pioneered this research area by proposing a reinforcement learning approach based on the Highest Cumulative Reward (HCR) rule to study the emergence of social norms. According to this rule, an agent chooses the strategy that has yielded the highest reward in the past  $m$  iterations. The history of the strategies chosen and the rewards for each strategy are stored in a memory of a certain size. Their experiments showed that the rate of updating strategy and interval between memory flushes had a significant impact on the efficiency of norm emergence. Sen *et al.* [21, 14] proposed a mechanism for the emergence of norms through social learning in which agents learn norms based on private interactions. They experimented with three reinforcement learning algorithms and studied the influence of the population size, the set of possible actions and the heterogeneity of the population on norm emergence. More recently, Savarimuthu *et al.* [18] discussed three aspects of active learning (i.e., experiential, observational and communication-based learning) of norm emergence in an agent society, and demonstrated the usefulness of combining these three aspects of norm learning to boost the convergence of social norms. All these studies handled the issue of norm emergence via learning in the context of an agent population, in which each agent can interact randomly with other agents. Our work differs from all these studies because we focus on the emergence of norms under a networked agent society in which the interactions of agents are physically constrained. In addition, agents in our model learn simultaneously with all their neighbors to achieve a final consensus by using ensemble techniques. This is in contrast to the sequential learning process in all the previous studies, in which each agent is selected sequentially to interact with another agent in the population.

A number of researchers have studied norm emergence by examining the underlying network topology of agents. Sen *et al.* [20] evaluated how varying topologies of social networks affected the emergence of norms through social learning in these networks. Three different kinds of network topolo-

gies (i.e., scale-free, fully-connected and ring networks) were studied to show how quickly norms converged in social networks depending on parameters such as the topology of the network, the population size and the number of actions available. Villatoro *et al.* [26] investigated the effects of memory and the history of past activities during learning on the success and rate of emergence of social norms in different network structures. The authors confirmed that different characteristics of the topology in which agents are located could produce different convergence rates for reaching a social norm. Later, Villatoro *et al.* [25] used two social instruments (i.e., rewiring and observation) to effectively address the frontier effect problem caused by the sub-conventions so as to facilitate norm emergence in the whole network. Recently, Mahmoud *et al.* [13] further extended Axelrod’s seminal model [2] by considering the topological structures, in particular, scale-free networks. All these studies, however, were based on a simple interaction protocol: each agent must be paired with one of its neighbors for interaction so that this agent can learn directly from this interaction. This interaction protocol simplifies real-life situations when individuals can collectively decide among multiple alternatives. In our study, an agent interacts with all of its neighbors simultaneously and learns from these interactions collectively. The focus is to study the impact of local collective behaviors on the overall emergence of norms in the whole society in a number of different conditions. This focus differentiates our work from all these previous studies.

In the area of reinforcement learning, ensemble techniques have been widely used to boost learning efficiency and to improve learning performance. Wiering and Hasselt [29] used different ensemble methods to combine multiple independently reinforcement learning algorithms to choose the best action. Hans and Udluft [11] used ensemble techniques to make reinforcement learning more robust and less dependent on the various parameters. Various ways of aggregating single learners are proposed to learn a combined parameterized state-value function of multiple agents. In a more recent work, Fauer and Schwenker [8] proposed several ensemble methods to learn a combined parameterized state-value function of multiple agents. All these studies employed ensemble learning in reinforcement learning with the aim of enhancing learning speed and improve final performance. In our work, however, the ensemble methods are used to combine the focal agent’s best-response actions for each of its neighbors to make a final decision. Our focus is not on the learning efficiency, but on the different patterns of norm emergence achieved through local collective behaviors of agents in networked agent societies.

## 6. CONCLUSION AND FUTURE WORK

In this paper, we studied the emergence of social norms through collective learning from local interactions in networked agent societies. The goal of this work is to investigate whether such collective learning can successfully establish a social norm as a bottom-up process so that coordination can be achieved across the whole society. We showed that the collective learning framework is more efficient and robust than the pairwise learning framework which has been adopted in most previous studies. We investigated the influence of various essential issues on the ratio and speed of norm emergence. The experimental studies confirmed that the collective learning framework is robust and efficient for evolving stable norms in networked agent societies.

The long-term goal of this research is to design robust mechanisms capable of predicting social behavior and facilitating social order in agent societies through analyzing local collective behaviors. Such robust mechanisms can not only provide us with a better understanding of the formation and evolution process of opinions and conventions in human society, but also enable us to build and design robust MASs such as electronic institutions. This paper is a step towards this goal, however, much more work still remains to be done. For example, to better imitate real-life societies, relationships between agents need to be added to the network structure. In addition, more varied local learning behaviors can be defined and further investigated, for example, by considering the multiple transitive states of agents.

## 7. REFERENCES

- [1] A. Artikis, L. Kamara, J. Pitt, and M. Sergot. A protocol for resource sharing in norm-governed ad hoc networks. *Declarative agent languages and technologies II, LNAI 3476*, pages 221–238, 2005.
- [2] R. Axelrod. An evolutionary approach to norms. *Am. Polit. Sci. Rev.*, 80(4):1095–1111, 1986.
- [3] A. Barabási and R. Albert. Statistical mechanics of complex networks. *Rev. Modern Phys.*, 74:47–97, 2002.
- [4] M. Bowling and M. Veloso. Multiagent learning using a variable learning rate. *Artif. Intel.*, 136:215–250, 2002.
- [5] L. Brooks, W. Iba, and S. Sen. Modeling the emergence and convergence of norms. In *Proc. of 22nd IJCAI*, pages 97–102, 2011.
- [6] N. Criado, E. Argente, A. Garrido, J. Gimeno, F. Igual, V. Botti, P. Noriega, and A. Giret. Norm enforceability in electronic institutions? *Coordination, Organizations, Institutions, and Norms in Agent Systems VI, LNAI 6541*, pages 250–267, 2011.
- [7] P. Davidsson and A. Jacobsson. Aligning models of normative systems and artificial societies: Towards norm-governed behavior in virtual enterprises. In *Proc. of NormAS07*, 2007.
- [8] S. Faußer and F. Schwenker. Ensemble methods for reinforcement learning with function approximation. *Multiple Classifier Systems, LNCS 6713*, pages 56–65, 2011.
- [9] H. Franks, N. Griffiths, and A. Jhumka. Manipulating convention emergence using influencer agents. *Auton. Agent Multi-Agent Syst.*, DOI:10.1007/s10458-012-9193-x, 2012.
- [10] D. Fudenberg and D. Levine. *The theory of learning in games*. The MIT press, 1998.
- [11] A. Hans and S. Udluft. Ensembles of neural networks for robust reinforcement learning. In *Proc. of 9th ICMLA*, pages 401–406, 2010.
- [12] L. Hofmann, N. Chakraborty, and K. Sycara. The evolution of cooperation in self-interested agent societies: a critical study. In *Proc. of 10th AAMAS*, pages 685–692, 2011.
- [13] S. Mahmoud, N. Griffiths, and J. Keppens. Norm emergence: Overcoming hub effects in scale free networks. In *Proc. of COIN*, pages 136–150, 2012.
- [14] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained interactions in diverse societies. In *Proc. of 7th AAMAS*, pages 779–786, 2008.
- [15] H. Ohtsuki, C. Hauert, E. Lieberman, and M. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505, 2006.
- [16] R. Polikar. Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.*, 6(3):21–45, 2006.
- [17] S. Redner. How popular is your paper? an empirical study of the citation distribution. *Eur. Phys. J. B.*, 4(2):131–134, 1998.
- [18] B. Savarimuthu, R. Arulanandam, and M. Purvis. Aspects of active norm learning and the effect of lying on norm emergence in agent societies. *Agents in Principle, Agents in Practice, LNAI 7047*, pages 36–50, 2011.
- [19] B. Savarimuthu and S. Cranefield. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multi-Agent Grid Syst.*, 7(1):21–54, 2011.
- [20] O. Sen and S. Sen. Effects of social network topology and options on norm emergence. *Coordination, Organizations, Institutions and Norms in Agent Systems V, LNAI 6069*, pages 211–222, 2010.
- [21] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proc. of 20th IJCAI*, pages 1507–1512, 2007.
- [22] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artif. Intel.*, 94(1-2):139–166, 1997.
- [23] R. Sutton and A. Barto. *Reinforcement learning: An introduction*. The MIT press, 1998.
- [24] G. Szabo and G. Fáth. Evolutionary games on graphs. *Phys. Rep.*, 446(4-6):97–216, 2007.
- [25] D. Villatoro, J. Sabater-Mir, and S. Sen. Social instruments for robust convention emergence. In *Proc. of 22nd IJCAI*, pages 420–425, 2011.
- [26] D. Villatoro, S. Sen, and J. Sabater-Mir. Topology and memory effect on convention emergence. In *Proc. of WI-IAT'09*, pages 233–240, 2009.
- [27] C. Watkins and P. Dayan. Q-learning. *Mach. Learn.*, 8(3):279–292, 1992.
- [28] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.
- [29] M. Wiering and H. van Hasselt. Ensemble algorithms in reinforcement learning. *IEEE T. Syst. Man Cy. B.*, 38(4):930–936, 2008.
- [30] H. Young. The economics of convention. *J. Econ. Pers.*, 10(2):105–122, 1996.