Social Learning in Networked Agent Societies

Chao Yu
School of Computer Science and Technology
Dalian University of Technology
Dalian, China, 116024
Email:cy496@dlut.edu.cn

Rui Liu
School of Computer Science and Technology
Dalian University of Technology
Dalian, China, 116024
Email: 913917030@qq.com

Yuliang Wang
School of Computer Science and Technology
Dalian University of Technology
Dalian, China, 116024
Email: 429482349@qq.com

***Abstract*—In networked multiagent societies, how global social order (i.e., social norm) can be achieved through agents' local interactions is a critical research problem in multiagent systems. It has been shown that learning from individual local interactions is an effective mechanism to facilitate norm emergence. Most of the existing work, however, mainly focuses on studying norm emergence via agent learning from its individual experience. The role of social learning, i.e., learning directly from others, has been comparatively less investigated. This paper steps forward the state-of-the-art by investigating how social learning can impact the emergence of social norms in networked agent societies. Experiments are carried out to show the impact of different strategies of choosing between individual and social learning on norm emergence. Experimental results reveal some significant insights into the manipulation and control of norm emergence in networked agent societies achieved through agent local behaviors.**

I. Introduction

How global social order (social norm) can be established automatically as an emerging process is a key problem in the research of Multiagent Systems (MASs). It has been well recognized that learning from individual experience is a robust mechanism to facilitate emergence of social norms [1], [2]. For this reason, a number of researchers have focused on equipping agents with a learning capability to establish a norm for an agent society [3], [4], [5], [6], [7], [8]. All these studies, however, focus on studying norm emergence based on agent learning from its own experience. This kind of individual learning indicates that each agent must interact with another agent (or other agents), randomly or preferentially selected in the population, so that the agent can directly learn from this interaction (or these interactions).

However, agents not only can learn from their individual trial-and-error experiences, but also can learn from the information provided by other agents [9], [10]. Generally, learning directly from others (as opposed to one's own experiences) is referred to as social learning [11]. By exploiting the information provided by other experienced individuals, social learning is potentially an efficient way of acquiring valuable information to boost the consensus among agents and therefore to facilitate norm emergence in a society. However, it is still not clear how social learning can be valuable for an efficient emergence of social norms, and how the different relationships between individual and social learning can influence the norm emergence.

Against this background, this paper studies the emergence of social norms in networked agent societies, where agents learn from both their individual experience and social information. Two sets of experiments are carried out to show the impact of different strategies to choose between individual and social learning on norm emergence. The first set of experiments show that individual and social learning have different roles in facilitating norm emergence, that is, social learning is less valuable than individual learning in dynamic environments. The second set of experiments show that agents can dynamically choose between social and individual learning using an adaptive strategy in order to facilitate norm emergence among agents, and different network types and topologies have significant impacts on the emergence of social norms.

The remainder of the paper is organized as follows. Section II describes the definitions of social norms and networked agent societies. Section III presents the individual and social learning strategies. Section IV gives the experimental studies. Section V discusses related work. Finally, Section VI concludes the paper with some directions for future research.

II. Networked Agent Societies and Social Norms

This section gives descriptions of networked agent societies and social norms.

A. Networked agent societies

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 (N is the number of nodes).

(2) **Small-world networks**. This kind of network is to represent the small-world phenomenon in many natural, 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. We use $SW_N^{k,\rho}$ to denote a small-world network, where k is the average size of the neighborhood of a node, ρ is the re-wiring

probability to indicate the different orders of randomness of the network, and N is the number of nodes.

(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 $k^{-\gamma}$. We use $SF_N^{k,\gamma}$ to denote a scale-free network (N is the number of nodes).

B. Social norms

A social norm is said to have been established when all (or at least the majority of) agents in the society have complied with the same action. This research uses learning "rules of the road" [4], [12] as a metaphor to study the emergence of norms. In this scenario, agents strive to establish a convention of driving either on the left (L) or on the right (R) of the road. This interaction can be viewed as a 2-person 2-choice symmetric coordination game [3], with the payoff matrix displayed in Table I.

TABLE I
PAYOFF MATRIX OF THE SYMMETRIC COORDINATION GAME ($x > v$ AND $y > u$)

	Left (L)	Right (R)
Left (L)	x, x	u, u
Right (R)	v, v	y, y

This paper uses the pure coordination game to formulate the interaction between two neighboring agents, with $x = y = +1$ and $u = v = -1$ in Table I. Although its payoff matrix appears simple, the coordination game poses a very challenging puzzle for human beings to solve efficiently. 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 efficiently because they can rely on some contextual cues to agree on a particular equilibrium [12]. One such contextual cue is social norms (i.e., conventions and laws) that can be used to guide human behaviors when moral or rational reasoning does not provide a clear guidance because of the myopic behavior and the limited processing ability of individuals.

III. INDIVIDUAL AND SOCIAL LEARNING STRATEGIES

This section introduces the individual and social learning strategies for agent interaction in networked agent societies.

A. The individual learning strategy

Many previous studies have focused on equipping agents with an individual learning capability to establish a norm for an unstructured agent population [4] or a networked agent society [5], [6]. All these works, however, are based on a simple pairwise interaction protocol that, at each time step, each agent is randomly paired with one of its neighbors, randomly or preferentially, for interaction so that the agent can directly learn from this interaction either through a best response rule [6] or a memory-based rule [5]. This interaction protocol simplifies real-life situations when individuals can collectively make a decision from multiple alternatives before them. To reach a group consensus, people often interact with others at the same time and learn collectively from all these interactions.

Algorithm 1: The individual learning strategy in view of agent i

```
for each neighbor j ∈ N(i) of agent i do
    Agent i chooses best-response action a_{i→j} regarding
    neighbor j using a learning strategy;
end
Agent i aggregates all the actions a_{i→j} into a final action
a_i using ensemble learning methods;
for each neighbor j ∈ N(i) of agent i do
    Agent i plays action a_i with neighbor j and receives
    reward r_{i→j};
    Agent i updates learning information regarding
    neighbor j using ⟨a_i, r_{i→j}⟩;
end
```

To model the opinion aggregation process in human decision making, an individual learning strategy [8] has been proposed to study the impact of agent learning from local collective interactions on norm emergence in networked agent societies. Algorithm 1 gives the interaction protocol of the individual learning strategy, in which norms evolve as agents learn over repeated interactions with their neighbors using multiagent reinforcement learning algorithms [13]. 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 different 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. Experimental results showed that this kind of learning strategy was more efficient and robust for an emergence of social norms [8].

Several ensemble learning methods were proposed in [8]. The majority voting method simply counts the number of each action as the preference for corresponding action so that the most preferred action is simply the one that is suggested by most of the neighbors. The weighted voting method considers the "social ranks" of each neighbor (structural position or the neighbor's performance in past interactions) in the calculation of the preference for each action.

B. The social learning strategy

Agents not only learn from their individual trial-and-error experiences, but also learn from the information provided by other agents. Social learning is potentially a cheap way of acquiring valuable information. By exploiting the information provided by other experienced individuals, agents might shortcut the time required to develop the skills that are hard to be developed through individual learning. It has been shown that social learning plays an important role in the evolution of complex social behaviors in humans and animals [9].

Copying, either through observation or communication, is a major form of social learning [9]. Most work on social learning

Algorithm 2: The social learning strategy in view of agent i

1 **for** *each neighbor* $j \in N(i)$ *of agent* i **do**
2 | Agent i observes the action that neighbor j has played;
3 **end**
4 Agent i counts the number of actions from all the neighbors;
5 Agent i determines a_i using the majority voting strategy;
6 Agent i plays action a_i with all neighbors;

focuses on proposing strategies to dictate the circumstances under which agents should copy others and whom these agents should copy from [11]. Algorithm 2 shows a simplified social learning strategy based on observation, in which agent i simply observes the action taken by each of its neighbors (Line 2) and copies the action which is played by most of its neighbors for next round play (Line 4-6).

IV. Individual and Social Learning for Norm Emergence

This paper uses the widely used reinforcement learning algorithm Q-learning [13] for agent interaction. Learning rate α is set to 0.1, and ε is set to 0.1 in the ε-exploration strategy. The majority voting method is adopted as the basic ensemble learning method in the individual learning strategy. The Barabasi-Albert model [14] is used to generate a scale-free network by starting with $m_0 = 5$ agents and adding a new agent with $m = 1$ edge to the network at every time step. This network evolves into a scale-free network $SF_N^{k,3}$ following a power law with an exponent $\gamma = 3$. The Watts-Strogatz model [15] is used to generate a small-world network. All results are averaged in 100 independent runs.

Two sets of experiments are carried out to show the impacts of different strategies to choose individual and social learning on the emergence of social norms. In the first set of experiment (Subsection IV-A), each agent has a fixed strategy, namely, each agent either adopts one pure strategy (individual or social learning), or has a fixed mixed strategy to choose between these two learning strategies. In the second set of experiment (Subsection IV-B), the strategy to choose between individual and social learning is adaptive, which means that each agent adapts its strategy to choose between individual and social learning based on the expected future reward received using the corresponding learning strategy. Again, Q-learning is used as the reinforcement learning algorithm to adapt this kind of meta-level strategy.

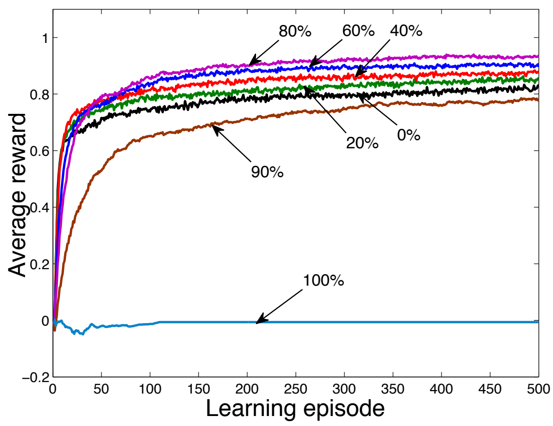
A. Fixed strategy

When agents use a fixed strategy, the agent population can be heterogeneous or homogeneous. The population is heterogeneous when each agent has a different strategy for interaction (i.e., either social learning or individual learning), and homogeneous when each agent has the same mixed strategy to choose between social and individual learning.

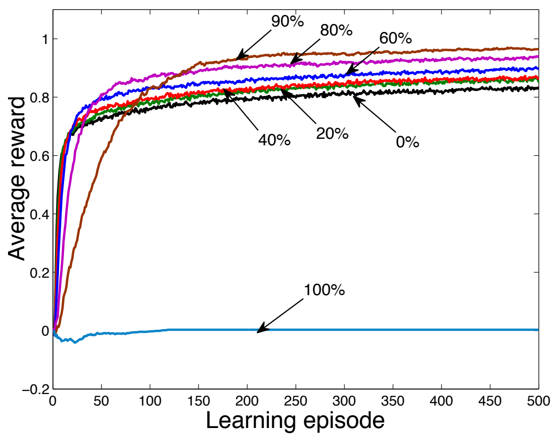
Fig. 1 shows the norm emergence when agents use a fixed strategy in closed environments. Fig. 1(a) shows the impact of different proportions of agents using a pure social learning strategy on norm emergence in network GR_{100}. The results show that the norm emerges more quickly when the proportion of social learning agents gets larger at the beginning (i.e., from 0% to 80%). However, increasing the proportion further from 90% to 100% significantly hinders the convergence process. When the population consists of 100% social learning agents, the system is fully in chaos because all the agents simply observe their neighbors in order to copy the most chosen action. Since no agents use individual learning, the agents cannot remember the past learning experience to make a reasonable decision. That is why the learning curve for the system of 100% social learning agents fluctuates at the beginning and then stabilizes at a reward of 0 afterwards. However, incorporating a small number of individual learning agents, e.g., 10% (i.e., curve 90% in Fig. 1(a)), can drastically boost the norm convergence. This is because the individual learning agents can take advantage of the interaction experience to exploit other agents for a better outcome. The same patter of results can be observed in a homogeneous population in Fig. 1(b), where agents adopt different levels of social learning. As can be seen, a low level of social learning can significantly boost the convergence of social norms.

In real applications, equipping agents with an individual learning capability often means a cost to either the agents themselves or the whole system [9]. For example, the agents might need physical space to store the learning information, thus imposing a managerial cost on the agents, or for some safety-critical environments, a fatal decision caused by the trial-and-error process during individual learning could bring about disastrous consequences to the whole system. It is therefore more efficient to deploy as few individual learning agents as possible in the whole system in order to decrease the side-effects caused by learning. The results in Fig. 1 indicate that it is possible to achieve a maximal performance by either incorporating only a small proportion of individual learning agents into a large group of social learning agents, or letting agents use a mixed strategy with a low probability of individual learning and high probability of social learning. These two principles can be helpful for the efficient mechanism design of large-scale norm-governed systems, in which the coexistence of millions of agents makes it inefficient or even unfeasible to achieve a global optimal performance through each agent's individual learning.

Fig. 2 shows the norm emergence in open environments, in which existing agents can leave the society and new agents can enter into the society at running time. Notation D denotes that, at each learning episode, D existing agents in the society are replaced by new agents. Therefore, D indicates different levels of population dynamics in an open environment. Fig. 2(a) shows the impact of different population dynamics on norm emergence when agents use a pure strategy of individual learning. As can be seen, a higher population dynamic causes a slower convergence speed and a lower level

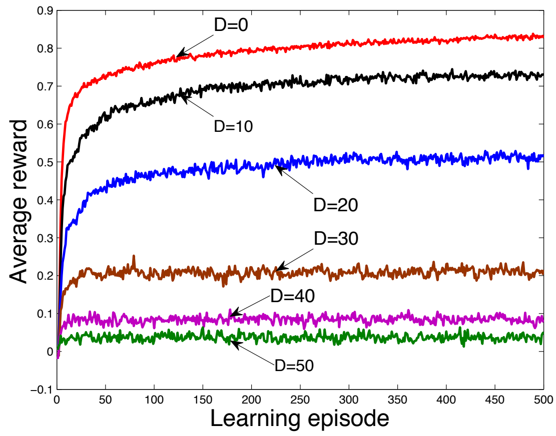


(a) Heterogeneous population

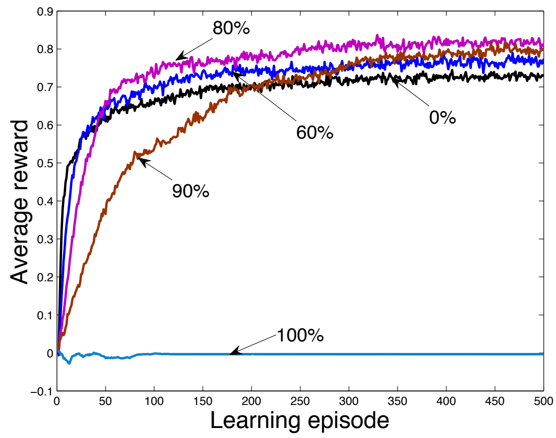


(b) Homogeneous population

Fig. 1. Norm emergence in closed environments.



(a) Impact of population dynamics



(b) Learning dynamics when $D = 10$

Fig. 2. Norm emergence in open environments.

of convergence ratio of social norms. This is because it is more difficult for the agents in societies with higher population dynamics to distinguish the effects of their actions on the environment, and these uncertainties can hinder the agents from reaching consensuses among them. Fig. 2(b) shows the learning dynamics when agents adopt different probabilities of social learning in an open environment with population dynamic of 10 (i.e., $D = 10$). As in the closed population in Fig. 1(b), a small probability of social learning can significantly boost the convergence of social norms. However, in the open population, using around 80% social learning can bring about the most efficient emergence of social norms, which is against in the closed population,where around 90% social learning is needed to bring about the most efficient emergence of social norms. This result reveals an important relationship between individual learning and social learning in norm emergence, i.e., individual learning is more valuable than social learning for an efficient emergence of social norms in dynamic environments. This is because the information received through social learning might be outdated, misleading or inappropriate in dynamic environments. As a result, when the dynamic of the environment increases, the value of social learning (e.g., copying) declines [9]. In other words, social learning, although it is a cheaper way of acquiring knowledge about the environment, is more error-prone than individual learning. Therefore, a higher level of individual learning is required to sample the environment in order to provide more reliable information for social learning in a more dynamic environment, so that an efficient emergence of social norms can be guaranteed.

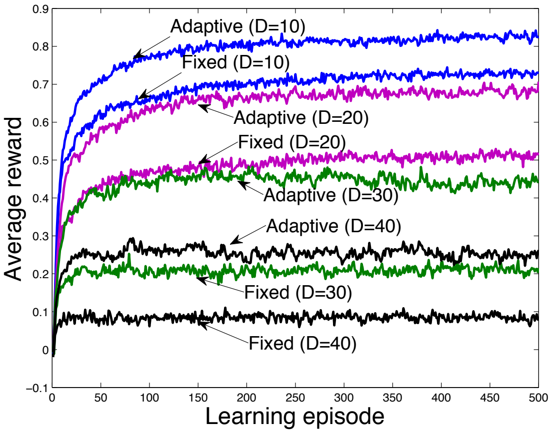
B. Adaptive strategy

A focus of research on social learning is how to take advantage of social learning, while managing the risks and uncertainties associated with the social information. In this subsection, the strategy to choose between social and individual learning is not fixed, but is adapted during running-time. General reinforcement learning algorithms can be used for this adaptation. Algorithm 3 gives a sketch of the adaptive strategy based on Q-learning, in which $Q_t(s)$ and $Q_t(i)$ are the Q values for choosing social learning and individual learning, respectively.

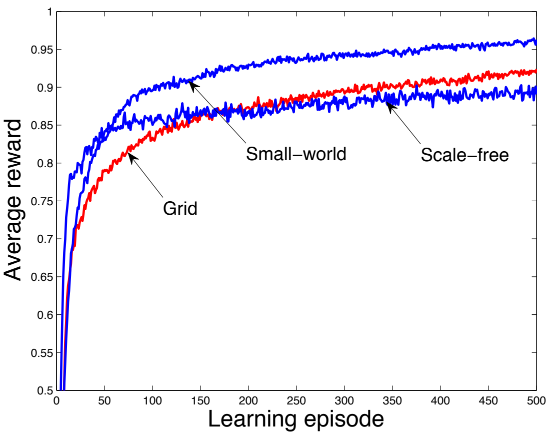
Fig. 3(a) shows the impact of population dynamics on norm emergence when agents use the adaptive strategy in network GR_{100}. As can be seen, in populations with different levels of dynamics, the adaptive strategy enables an more efficient emergence of social norms than the fixed strategy (using a pure strategy of individual learning). This result verifies that

Algorithm 3: The adaptive strategy

$Q_0(s) \leftarrow 0$, $Q_0(i) \leftarrow 0$;
for *each step t (t=1,...,T)* **do**
  if $Q_t(s) \leq Q_t(i)$ **then**
    Chooses individual learning (Algorithm 1) with ε-exploration strategy;
  else Chooses social learning (Algorithm 2) with ε-exploration strategy Interacts with neighbors and receives average reward r;
  if *Chooses individual learning* **then**
    Updates Q value:
    $Q_{t+1}(i) \leftarrow Q_t(i) + \alpha(r - Q_t(i))$;
  else Updates Q value:
  $Q_{t+1}(s) \leftarrow Q_t(s) + \alpha(r - Q_t(s))$
end



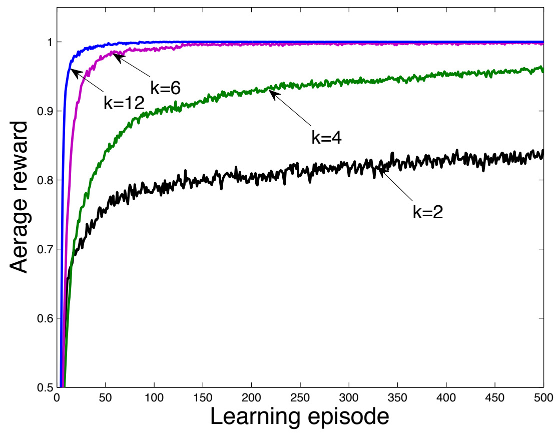
(a) Benefits of using adaptive strategy



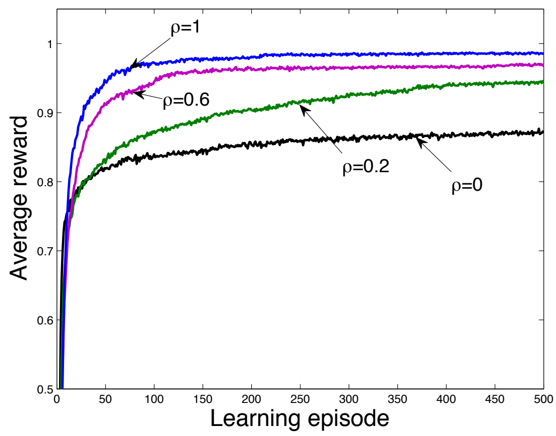
(b) Different kinds of social networks

Fig. 3. Learning dynamics using the adaptive strategy.

the adaptive strategy enables agents to dynamically choose between social and individual learning in order to facilitate norm emergence among agents. Fig. 3(b) shows norm emergence using the adaptive strategy in the three deferent kinds of networks (GR_{100}, $SW^{4,0.8}_{100}$ and $SF^{k,3}_{100}$). As can be seen, norms emerge fast at the beginning and then gradually when the learning moves on in the scale-free network. Norms emerge slower in the small-world and grid network than that in the scale-free network at first and then faster later on. Small-world network enables a more efficient emergence of social norms than the grid network throughout the whole learning process.



(a) Impact of neighborhood size k



(b) Impact of network randomness ρ

Fig. 4. Impacts of network topologies on norm emergence.

Fig. 4 shows the impacts of network topologies (i.e., neighborhood size and network randomness) on norm emergence. As can be seen from Fig. 4(a), 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. Clustering coefficient is a measure of degree to which nodes in a graph tend to cluster together. When the average number of neighbors increases, the clustering coefficient also increases, and therefore agents located in different parts of the network only need a smaller number of interactions to reach a consensus. On the other hand, when agents have a small neighborhood size, they only interact with their neighbors, which account for a small proportion of the whole population. This results in diverse sub-norms formed at different regions of the network. Such sub-norms conflict with each other in the network, and thus more interactions are needed to solve these conflicts and achieve a uniform norm for the whole society. Fig. 4(b) shows the influence of network randomness on norm emergence. When $\rho = 0$, network $SW^{k,\rho}_N$ is reduced to a regular ring lattice. Increasing rewiring probability ρ produces a small network with increasing randomness. When $\rho = 1$, the network becomes a fully random network. The results indicate that it is more efficient for a norm to emerge in a network with higher randomness. This is because the increase in randomness can reduce the network diameter (i.e., the largest number of hops in order to traverse from one vertex to another [5]), and the smaller a network diameter is, the more efficient for the network to evolve a social norm [8].

V. Related Work

A number of researchers have studied learning for norm emergence in networked MASs. Airiau *et al.* [7] and Sen *et al.* [6] evaluated how varying topologies of social networks affected the emergence of norms through agent individual learning. Three different kinds of network topologies 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.* [5] 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 network topology could produce different convergence rates for a social norm. All the above studies focused on norm emergence based on agent individual learning from its trial-and-error experience. This is in contrast to our work, in which agents learn from both their individual experiences and social information from others.

Several studies also investigated the impact of social learning on norm emergence. For example, Savarimuthu *et al.* [1] demonstrated the usefulness of combining both individual learning (i.e., experiential) and two other forms of social learning (i.e., observational and communication-based learning) to boost the convergence of social norms; Villatoro *et al.* [16] used social learning (i.e., observation) as an efficient social instrument to effectively address the *frontier effect* problem so as to facilitate norm emergence in networked agent societies; and Verhagen [17] proposed a simple simulation model to test the impact of different proportions of individual learning (represented by updating an agent's self-model through its own experience) and social learning (represented by updating an agent's group model through communicating with other group members) on the processes of norm spreading and internalization in the whole group. Different from all these studies, our work focuses on investigating the relationship between social and individual learning in order to understand the different roles of these two learning strategies in an efficient emergence of social norms.

VI. Conclusion and Future Work

This paper studies individual and social learning for norm emergence in networked agent societies. Two sets of experiments were carried out to show the impact of different strategies to choose between individual and social learning on the emergence of social norms.The first set of experiments showed that individual and social learning had different roles in the facilitation of norm emergence, that is, social learning was less valuable than individual learning in dynamic environments. The second set of experiments showed that the adaptive strategy enabled agents to dynamically choose between social and individual learning in order to facilitate norm emergence among agents, and different network types and topologies had significant impacts on the emergence of social norms.

For future work, it is necessary to associate a cost with individual learning in the adaptive strategy so that agents can learn to use individual learning only when the associated cost can potentially bring about higher future rewards. When the norm is entrenched, agents can reduce the use of individual learning to save cost. This can model a phenomenon in real-life situations that individual learning is often inversely related to the strength of a social norm, that is, once a norm is entrenched, we conform thoughtlessly.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant 61502072, 61572104 and 61403059, Fundamental Research Funds for the Central Universities of China under Grant DUT16RC(4)17, and Post-Doctoral Science Foundation of China under Grants 2014M561229 and 2015T80251.

References

[1] Savarimuthu, B., Arulanandam, R. and Purvis, M.: Aspects of Active Norm Learning and the Effect of Lying on Norm Emergence in Agent Societies. LNAI 7047. 36–50 (2011)
[2] Yu, C., Zhang, M. and Ren, F.: Collective learning for the emergence of social norms in networked multiagent systems. IEEE T. Cybernetics. 44(12): 2342–2355 (2014)
[3] Shoham, Y. and Tennenholtz, M.: On the emergence of social conventions: modeling, analysis, and simulations. Artif. Intel. 94: 139–166 (1997)
[4] Sen, S. and Airiau, S.: Emergence of norms through social learning. Proc. of 20th IJCAI, pp. 1507–1512 (2007)
[5] Villatoro, D., Sen, S. and Sabater-Mir, J.: Topology and memory effect on convention emergence. Proc. of WI-IAT'09, pp. 233–240 (2009)
[6] Sen, O. and Sen, S.: Effects of social network topology and options on norm emergence. LNAI 6069, 211–222 (2010)
[7] Airiau, S., Sen, S. and Villatoro, D.: Emergence of conventions through social learning. Auton. Ag. Multi-Agent Syst. 1–26 (2013)
[8] Yu, C., Zhang, M., Ren, F. and Luo, X.: Emergence of social norms through collective learning in networked agent societies. In: AAMAS 2013, pp. 475–482 (2013)
[9] Rendell, L., et al.: Cognitive culture: theoretical and empirical insights into social learning strategies. Trends in cognitive sciences. 15(2): 68–76 (2011)
[10] Tamura, K., Kobayashi, Y. and Ihara, Y.: Evolution of individual versus social learning on social networks. Journal of the Royal Society Interface. 12: 20141285 (2015)
[11] Laland, K.N.: Social learning strategies. Animal Learn. & Behavior. 32(1): 4–14 (2004)
[12] Young, H.P.: The economics of convention. J. Econ. Pers.. 10(2): 105–122 (1996)
[13] Busoniu, L. and Babuska, R. and De Schutter, B.: A comprehensive survey of multiagent reinforcement learning. IEEE T. Syst. Man Cy. C. 38(2): 156–172 (2008)
[14] Barabási, A.L. and Albert, R.: Statistical mechanics of complex networks. Rev. Modern Phys. 74: 47–97 (2002)
[15] Watts, D.J. and Strogatz, S.H.: Collective dynamics of small-world networks. Nature. 393: 440–442 (1998)
[16] Villatoro, D., Sabater-Mir, J. and Sen, S.: Social instruments for robust convention emergence. Proc. of IJCAI, pp. 420–425 (2011)
[17] Verhagen, H.: Simulation of the Learning of Norms. Social Science Comput. Rev. 19(3): 296–306 (2001)