Heuristic Collective Learning for Efficient and Robust Emergence of Social Norms

(Extended Abstract)

Jianye Hao and Jun Sun
Singapore University of Technology and Design
{jianye_hao,sunjun}@sutd.edu.sg

Dongping Huang and Yi Cai
South China University of Technology
{huang.dp,ycai}@scut.edu.cn

Chao Yu
Dalian University of Technology
cy496@uowmail.edu.au

ABSTRACT

In multiagent systems, social norms is a useful technique in regulating agents' behaviors to achieve coordination or cooperation among agents. One important research question is to investigate how a desirable social norm can be evolved in a bottom-up manner through local interactions. In this paper, we propose two novel learning strategies under the collective learning framework: *collective learning EV-l* and *collective learning EV-g*, to efficiently facilitate the emergence of social norms. Experimental results show that both learning strategies can support the emergence of desirable social norms more efficiently in a much broader range of multiagent interaction scenarios than previous work, and also are robust across different network topologies.

Categories and Subject Descriptors

I.2 [**ARTIFICIAL INTELLIGENCE**]: Distributed Artificial Intelligence—*Multiagent systems*

Keywords

Collective learning; Norm emergence

1. INTRODUCTION

In multiagent systems (MASs), social norms play an important role in regulating agents' behaviors to ensure effective coordination among agents, and have been applied in a wide variety of practical distributed systems such as electronic institutions.

In distributed multiagent environments, since there does not exist a centralized controller and the norm of the system may vary when the environment changes and is usually not available beforehand, it is not feasible to precompute any norm for agents to employ before their interaction starts. Thus investigating what mechanism can facilitate agents towards a consistent and desirable social norm through local interaction is important to ensure the effective coordination among agents and overall high performance in such kinds of distributed and dynamic systems. Until now much effort has been devoted by researchers from the normative MASs area to investigating the emergence of norms in agent societies through different manners of learning. Sen and Airiau [1] firstly proposed applying a number of existing multiagent learning algorithms to the *social learning framework* to investigate the emergence of social norms through learning in a population of agents. Later, a number of work [2, 3] extended this social learning framework by taking into consideration complex networks (e.g., small-world and scale free network) to model the underlying topology of the agent society, and a number of learning strategies and mechanisms have been proposed to facilitate the convergence to social norms through local interaction.

1's payoff, 2's payoff		Player 2's action: a	b	c
Player 1's action	a	10,10	0,0	-30,-30
	b	0,0	7,7	0,0
	c	-30,-30	0,0	10,10

(a)

1's payoff, 2's payoff		Player 2's action: a	b	c
Player 1's action	a	12/8, 12/8	5/-5,5/-5	-40/-20, -40/-20
	b	5/-5,5/-5	14/0, 14/0	5/-5,5/-5
	c	-40/-20, -40/-20	5/-5,5/-5	12/8, 12/8

(b)

Figure 1: Payoff matrices for (a) coordination game with high penalty (CGHP) (b) Fully stochastic coordination game with high penalty (FSCGHP)

One commonly adopted abstraction of a norm in previous work is that it corresponds to a Nash equilibrium where all agents choose identical actions, which usually can be modeled as a coordination game. However, the norms in practical scenarios may correspond to other forms of Nash equilibria requiring agents to choose different actions (a.k.a. anti-coordination game), which cannot be (efficiently) handled in previous work [3]. Moreover, more complicated scenarios involve multiple Nash equilibria while only some of them correspond to norms and may involve stochastic payoffs shown in Figure 1(a) and 1(b). In this kind of scenarios, it is very likely for agents to converge to the non-norm equilibrium to avoid the high mis-coordination cost, which has not been addressed by previous work [1, 3] yet.

To tackle the above challenges, we propose two novel learning strategies: *collective learning EV-l* and *collective-learning EV-g*, under the *networked collective learning framework* which is applicable to a wide variety of scenarios for norm emergence. We extensively evaluate the learning perfor-

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*
Copyright © 2015, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

mance of both *collective learning EV-l* and *collective learning EV-g* and show that both strategies enable agents to reach consistent norms more efficiently in a wider range of games than previous approaches. We also empirically find that the performance of *collective learning EV-l* and *EV-g* is robust across different topologies.

2. NETWORKED COLLECTIVE LEARNING FRAMEWORK

Under the networked collective learning framework, there are a population of n agents in which each agent's neighbors are determined by the underlying network topology. Three major topologies are considered: ring network, small-world network and scale-free network. Each agent i learns its policy (i.e., which norm to adopt) through repeated pairwise interactions with all its neighbors each round. The interaction between each pair of agents is modeled as a two-player strategic game. During each interaction, one agent is randomly assigned as the row player and the other agent as the column player. At the beginning of each round, each agent first determines its current best-response actions against each of its neighbors as either row and column player respectively. Following that, the sets of best response actions for each role are synthesized into a single best response action respectively, which will be used as the overall strategy to interact with all of its neighbors in the current round. This models people's collective decision-making process in which people make collective decisions based on multiple feedbacks. In each round, each agent i has the opportunity to interact with each of its neighbors once. At the end of each round, each agent updates its learning strategy towards each neighbor based on its previous experience.

2.1 Collective Learning Strategy

We propose that each agent i holds a Q-value $Q_{i,j}(s,a)$ for each action a under each state $s \in \{Row, Column\}$ against each of its neighbors j, which keeps a record of action a's past performance against neighbor j and serves as the basis for making decisions. At the end of each round t, each agent i updates its Q-values for each neighbor j as follows,

$$\begin{aligned} Q_{i,j}^{t+1}(s,a) &= Q_{i,j}^{t}(s,a) + \alpha_{i,j}^{t}(s)\times \\ &\quad [R_{i,j}^{t}(s,a) + \gamma max_{a\in A_i} Q_{i,j}^{t}(s',a) - Q_{i,j}^{t}(s,a)] \end{aligned} \tag{1}$$

We employ the Frequency Maximum Q-value heuristic (FMQ) as the updating heuristic to compute the estimated values of the actions as follows.

$$EV_{i,j}^{t+1}(s,a) = Q_{i,j}^{t+1}(s,a) + c \times f^{t+1}(s,a) \times R_{max}^{t}(s,a) \tag{2}$$

where $R_{max}^{t}(s,a)$ is the highest payoff in the interaction history with agent j, $f^{t+1}(s,a)$ is the frequency of receiving the reward of $R_{max}^{t}(s,a)$ by choosing action a, and c is the weighting factor.

Based on its corresponding set of EV-values, each agent chooses its best response action against each neighbor under each state using the ϵ-greedy mechanism denoted as $\mathcal{S}_r^t$ and $\mathcal{S}_c^t$. Finally, each agent synthesizes a single best-response action for both states based on $\mathcal{S}_r^t$ and $\mathcal{S}_c^t$ using majority voting. The action with the highest vote is selected as the final action against all neighbors. We distinguish two general ways of making explorations following the ϵ-greedy mechanism: local exploration and global exploration. Local exploration is made before synthesizing the overall best-response action, while the global exploration is made after the best-response action has been synthesized. We denote the collective learning strategy with local and global exploration as *collective learning EV-l* and *collective learning EV-g*.

3. EXPERIMENTAL SIMULATION

We compare the performance of *collective learning EV-l* and *collective learning EV-g* with previous work: collective learning-l/g [3] and pairwise learning [1] in a variety of games. Three representative network topologies are considered here: *ring network*, *small-world network* and *scale-free network*. We also consider the case of *random network* where there is no fixed network topology and each agent simply interacts with a fixed and same number of agents randomly selected from the system each round.

We briefly summarize the results across all network topologies as follows: (1) both *collective learning EV-l* and *collective learning EV-g* are always able to achieve coordination on norms for all four types of games; (2) both *collective learning-l* and *collective learning-g* can only succeed in coordination games, while fail in the rest of games; (3) *pairwise learning* can succeed in both coordination and AC games, while fail in the rest of games; (4) for all networks, it generally takes longer time to converge to norms (or fails to converge) when the game becomes more challenging; 5) agents converge to norms faster under collective learning framework than social learning framework; 6) agents converge to norms faster using local exploration than global exploration under the collective learning framework. Overall, we can observe that the performance of *collective learning EV-l* and *EV-g* is robust towards different network topologies. The norm emergence performance is independent of the network topology and the identities of agents that they interact with.

4. CONCLUSION

We proposed two novel learning strategies for norm emergence through local interactions under the collective learning framework. Extensive simulation shows that collective learning EV-l and EV-g can enable agents to reach consistent norms more efficiently and robustly across a wider variety of games and topologies compared with previous approaches.

5. ACKNOWLEDGMENTS

This work is supported by project "IDD11100102A/ IDG31100105A" from Singapore University of Technology and Design, National Natural Science Foundation of China (Grant NO. 61300137) and the Guangdong Natural Science Foundation, China (NO. S2013010013836)

REFERENCES

[1] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proceedings of IJCAI'07*, pages 1507–1512, 2007.

[2] D. Villatoro, J. Sabater-Mir, and S. Sen. Social instruments for robust convention emergence. In *Proceedings of IJCAI'11*, pages 420–425, 2011.

[3] C. Yu, M. J. Zhang, F. H. Ren, and X. D. Luo. Emergence of social norms through collective learning in networked agent societies. In *Proceedings of AAMAS'13*, pages 475–482, 2013.