An Adaptive Learning Framework for Efficient Emergence of Social Norms

(Extended Abstract)

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

Hongtao Lv
School of Computer Science and Technology
Dalian University of Technology
Dalian, 116024, China
lvhongtao@mail.dlut.edu.cn

Sandip Sen
Department of Mathematical & Computer Sciences
The University of Tulsa
Tulsa, Oklahoma 74104, United States
sandip-sen@utulsa.edu

Jianye Hao
School of software
Tianjin University
Tianjin, 300072, China
jianye.hao@tju.edu.cn

Fenghui Ren
School of Computer Science and Software Engineering
University of Wollongong
Wollongong, 2500, Australia
fren@uow.edu.au

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

ABSTRACT

This paper investigates how norm emergence can be facilitated by agents' adaptive learning behaviors. A general learning framework is proposed, in which agents can dynamically adapt their learning behaviors through social learning of their individual learning experience. Experimental results indicate that the proposed framework outperforms the static learning framework in various comparison criteria.

Keywords

Norm Emergence; Reinforcement Learning; Social Networks

1. INTRODUCTION

Learning from individual experience has been shown to be a robust mechanism to facilitate norm emergence in multiagent systems (MASs). A great deal of work has studied norm emergence achieved through agent learning behaviors [1, 2, 3]. The focus of these existing studies is to examine general mechanisms behind efficient emergence of norms while agents interact with each other using basic learning (particularly reinforcement learning) methods. These mechanisms include the social learning strategy [1], the collective interaction protocol [3, 4], and the utilization of topological knowledge [2], etc. Learning parameters in these studies, however, are often fine-tuned by hand and thus cannot be adapted according to the varying norm emerging situations. A key question then arises that how agents can adapt their learning behaviors dynamically during the process of norm emergence, and how this kind of adaptive learning behaviors can influence the final emerging outcomes?

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

This paper provides another perspective in the research of norm emergence by simply focusing on the role of learning itself in affecting the process of norm emergence. A double-layered adaptive learning framework is proposed, in which agents interact with each other using basic Reinforcement Learning (RL) methods in the local layer learning, and generate guiding policies by exploiting historical learning experience in the upper layer learning. To generate guiding policies, the historical learning experience of each agent is synthesized into a strategy that competes with other strategies in the population based on the principle of Evolutionary Game Theory (EGT). The generated guiding policies are then passed down to the local layer learning in order to adapt agents' learning behaviors based on the consistency between agents' behaviors and the guiding policies. Experiments show that the proposed framework enables norms to emerge more efficiently and with higher convergence levels than the static learning framework, and some critical parameters such as norm spaces and network topologies can have significant influences on norm emergence.

2. THE PROPOSED FRAMEWORK

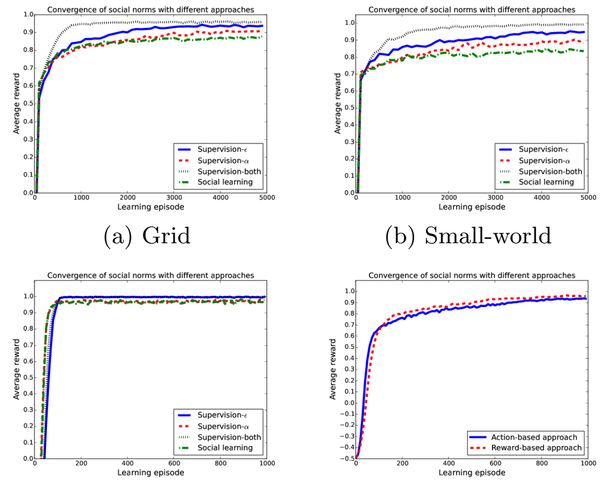
Each agent is equipped with a capability to memorize a certain period of learning experience in terms of the chosen action and the corresponding reward. Agent i then synthesizes its past learning experience into two tables $TA_i^t(a)$ and $TR_i^t(a)$. $TA_i^t(a)$ denotes the frequency of choosing action a in the last M steps and $TR_i^t(a)$ denotes the overall reward of choosing action a. The past learning experience in terms of table $TA_i^t(a)$ and $TR_i^t(a)$ indicates how successful the strategy of choosing action a is in the past. The upper layer learning makes use of this information in order to generate a supervision policy for local layer learning. To realize the supervision policy generation, each agent learns from other agents by comparing their learning experience. The motivation of this comparison comes from the EGT, which provides a powerful methodology to model how strategies evolve overtime based on their performance. Two different approaches are proposed in this paper to realize the EGT concept in the upper layer learning process, depending on how to define the competing strategy and the corresponding performance evaluation criteria (i.e., fitness) in EGT. In the *reward-based approach*, the strategy in EGT is represented by the most profitable action, and the fitness is represented by the corresponding reward of that action; On the contrary, the *action-based approach* considers the action which has been most adopted in the past to be the strategy in EGT, and the corresponding reward of that action to be the fitness in EGT. After synthesizing the historical learning experience, agent i then gets an action strategy of a'_i and its corresponding fitness of $TR(a'_i)$. It then interacts with other agents through social learning (i.e., imitation rule) in EGT.

The new strategy a'_i indicates the most successful strategy in the neighborhood and therefore should be integrated into the local layer learning in order to entrench its influence. By comparing its action at time step t, a^t_i, with the supervision strategy a'_i, agent i can evaluate whether it is performing well or not so that its learning behavior can be dynamically adapted to fit the supervision strategy. The concept of "winning" or "losing" in the well-known Multi-Agent Learning (MAL) algorithm WoLF (Win-or-Learn-Fast) [5] is elegantly borrowed to indicate whether an agent's behavior is consistent with the guiding policy. According to the "winning" or "losing" situation, agents then can dynamically adapt their learning behaviors (in terms of learning and/or exploration rate) in the local layer learning.

3. EXPERIMENTS AND RESULTS

We compare the proposed learning framework with the static social learning framework [1] in order to demonstrate the merits of adaptive learning behavior of agents. The three adaptive learning approaches under the proposed framework are: (1) *Supervision-α* (adapting learning rate), *Supervision-ϵ* (adapting exploration rate), and *Supervision-both* (adapting learning rate and the exploration rate at the same time). The performance of the four different approaches in the corresponding learning frameworks is plotted in Figure 1(a)-1(c). The results show that the three adaptive learning approaches outperform the static social learning approach in all three networks in terms of a higher level of convergence and a faster convergence speed. Through dynamically adapting their learning behaviors, agents are able to reach an agreement more easily, and thus norms can emerge more quickly to a higher level of convergence in the proposed framework. Result in Figure 1(d) shows that norms emerge faster with the action-based approach in the beginning. As the process moves on, the reward-based approach catches up with the action-based approach, and then brings about a higher level of norm emergence afterwards.

We also investigate the influence of several key factors on norm emergence under the proposed framework: (1) In terms of norm space size, a larger number of available actions results in a delayed convergence of norms; (2) The norm emergence process is hindered as the population is growing larger; (3) Network topologies such as network diameter and cluster coefficient can have significant influences on norm emergence. It is more efficient for a norm to emerge in a network with smaller network diameter (i.e., higher network randomness), and norm emergence is steadily promoted when cluster coefficient of network is increased (i.e., larger average number of neighbors). In all cases, the adaptive learning approaches can bring about more robust and efficient norm emergence than the social learning approach.



(a) Grid (b) Small-world (c) Scale-free (d) Ensemble method

Figure 1: Performance of norm emergence.

4. CONCLUSION AND FUTURE WORK

In this paper, a novel learning framework was proposed to investigate how agents' adaptive learning behaviors can facilitate norm emergence. The highlight of the framework is the integration of social learning into the local individual learning in order to dynamically adapt agents' learning behaviors for a better performance of norm emergence. Experimental results illustrated that this kind of interplay between individual learning and social learning is indeed helpful in facilitating the emergence of social norms among agents.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61502072), Post-Doctoral Science Foundation of China (No. 2014M561229 and 2015T80251), and National Undergraduate Training Programs for Innovation and Entrepreneurship (No. 2015101410154).

REFERENCES

[1] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proc. of 20th IJCAI*, pages 1507–1512, 2007.

[2] D. Villatoro, J. Sabater-Mir, and S. Sen. Social instruments for robust convention emergence. In *Proc. of 22nd IJCAI*, pages 420–425, 2011.

[3] C. Yu, M. Zhang, F. Ren, and X. Luo. Emergence of social norms through collective learning in networked agent societies. In *Proc. of AAMAS2013*, pages 475–482, 2013.

[4] C. Yu, M. Zhang, and F. Ren. Collective learning for the emergence of social norms in networked multiagent systems. *IEEE Transactions on Cybernetics*, 44(12):2342–2355, 2014.

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