Using a hierarchy of coordinators to overcome the frontier effect in social learning*

(Extended Abstract)

Sherief Abdallah
British University in Dubai
University of Edinburgh
shario@ieee.org

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Performance

Keywords

Supervision, Hierarchy, Reinforcement Learning

1. INTRODUCTION

We propose in this paper the use of a hierarchy of coordinators to improve the convergence of a network of agents to a global norm. A norm or a convention is an unwritten law that a society of agents agree on. Social norms are used by humans all the time. Choosing on which side of the road to drive a car and the right-of-way at an intersection are well-known examples. In a multi-agent setting, a convention may refer to a dominant coordination strategy, a common communication language, or the right of way among a group of robots. Upon establishing a norm, the overhead of coordination drops and the reliability of the multi-agent system increases [2]. When studying the emergence of norms and conventions, we typically assume the interaction between agents is random: a pair of agents are selected randomly to interact with one another. The process repeats both concurrently (several pairs interact at the same time) and consecutively (each agent collects history of interactions). When agents are adaptive, the process is then referred to as social learning. The coordination game is perhaps the most widely used game for studying social learning as it presents an agent community with two equally plausible norms to choose from (i.e. two Nash equilibriums). It was shown that in the absence of any restriction on agent interactions, a norm is guaranteed to emerge in the simple social learning setting where agents play the coordination game [1].

*The work reported in this paper was supported by Emirates Foundation grant #2010/107 , Program: Science and Engineering Research Grants

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

Recently social leaning was studied in networks [3]. The main difference here is that an underlying network restricts the interactions between the agents. In such a setting, convergence to a global norm is no longer guaranteed as more than one (sub)convention might emerge concurrently and remain stable. A sub-convention is a convention that is not adopted by the vast majority of the agents. The reason for the emergence of multiple stable sub-conventions is the existence of a stable barrier that separates the sub-conventions from one another (or equivalently, prevents each convention from invading the other). Such a barrier creates a suboptimal equilibrium. The frontier effect was reported to either prevent or significantly slowdown the convergence to a global norm across variety of network types. One of the recent proposed solutions to overcome the frontier effect was the use of social instruments [3]. Although the social instruments were successful in overcoming the barrier in regular and random networks, the social instruments failed in the case of scale-free networks. Furthermore, these social instruments had several limitations.

2. PROPOSED SOLUTION: HIERARCHY OF COORDINATORS

When we looked at examples of the frontier effect in scale-free networks, it became apparent that the problem was in the strictly local view of agents on the frontier. If only agents had a more global view, they would have reached a global convention. The social instruments that were proposed before [3] effectively provided (implicitly) individual agents with slightly more global view. For example, the observation social instrument allowed an agent to observe another agent in the network without being restricted to the underlying network (i.e. an agent could observe what convention was adopted by a randomly chosen agent anywhere in the network). The re-wiring social instrument also allowed agents to extend their view beyond their immediate neighbors. Here we propose a more structured mechanism for agents to exchange information (with varying detail and range) about their current state and exchange advices about the best course of action to reach a convention. We propose the use of an organization (hierarchy) of coordinators, as follows. Agents are separated in-to clusters, where each cluster is assigned a coordinator from the agents in the cluster. The clusters are then grouped in to meta-clusters, again selecting one of the cluster members to be the coordinator. The process is repeated recursively until we end up with root coordinator (the hierarchy can stop at a lower level with a

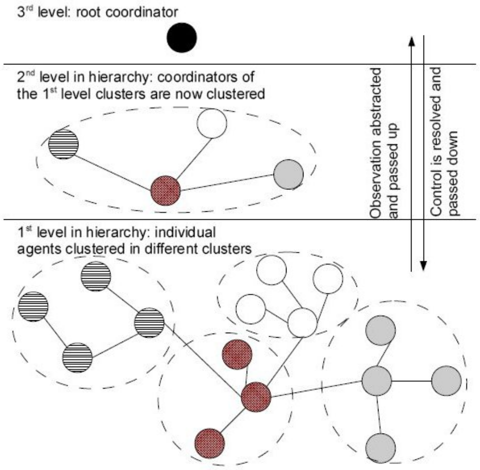

Figure 1: Illustration of the coordinators hierarchy. In the first level we have individual agents who are clustered into four clusters. In the second level we have the coordinators of the clusters in the first level (each coordinator has the same color as its cluster). In the final level we have the root coordinator (note that the hierarchy does not have to end with an individual root).

set of clusters at the top). Figure 1 illustrates the hierarchy of coordinators operating over a scale-free network. In our experiments we use a simple bottom-up hierarchical clustering to automatically build the coordinators' hierarchy, but more sophisticated clustering techniques can be used.

An important property of our solution is that it does not change the underlying network structure (no re-wiring) nor does it force individual agents to permanently adopt any convention. The coordinators only coordinate agent exploration of the state-action space so the agents can reach a global convention. Intuitively, our approach works as follows. Individual agents interact normally through the network that govern their interactions. Each group of agents is assigned a coordinator that observes their convergence. If the agents in a group does not converge after some period of time, the coordinator then asks its group to try a recommended convention for a short period of time. However, since each coordinator only observes its own group, it is still possible that different coordinators recommend different conventions. However, because there is a hierarchy of coordinators, inconsistencies are guaranteed to be discovered higher in the hierarchy. Such arrangement does appear in real-life. For example, when a new technology is discovered, private companies that produce the new technology are left unregulated. If after a while no industry-wide standard (convention) is adopted, individual states may start enforcing some standards, and if needed a federal law may be put in place to ensure the quick reach to a convention. Once a convention is established it becomes self-enforcing without an external enforcement.

3. EXPERIMENTS AND DISCUSSION

Our experiments focus on applying social learning in scale-free networks to reach one of the two possible norms of the coordination game. The state observed by coordinators in this case is the ratio of its subordinates that adopted the first norm. So for example if coordinator X controls a cluster of 5 individuals, and 4 which adopted the first norm, then X's state is 0.8. Each coordinator has two control actions: to ask its subordinates to try either norm 1 or norm 2 for short period of time T_{try} . To avoid conflicting control actions, a coordinator does not issue a control action if a control action from its superior is currently being executed. The strategy of the coordinator for choosing a control action is simple: if the ratio of subordinates that adopted the first norm is between δ and $1-\delta$ (we used $\delta = 0.25$ in our experiments) this means no norm is adopted yet within the cluster controlled by the coordinator. The coordinator then chooses a consistent control action (e.g. if the ratio is 0.3, the coordinator asks subordinates to try the second norm). The results we report here have been obtained by randomly generating 10 different scale-free networks. Every network consists of 225 agents and individual agents execute Q-learning with learning rate = 0.1 and exploration rate = 0.1. To avoid initial bias, the Q-learning action values are initialized to uniformly random values between 0 and 1. The coordinators hierarchy is generated randomly such that the hierarchy has 5 levels and each coordinator controls a cluster of at least 5 individuals. Figure 2 illustrates how using a hierarchy of coordinators ensured the convergence to a global norm.

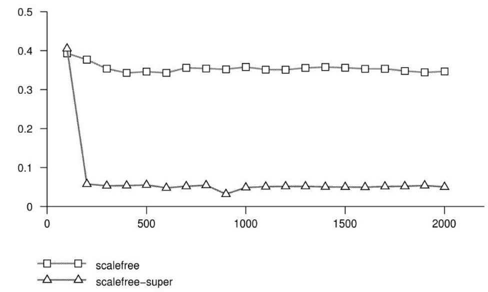

Figure 2: Comparing the percentage of the first norm adopted by agents for both social learning with and without the use of hierarchy coordinators in a sample run. Without coordinators no clear norm dominates the population, while with the hierarchy of coordinators a quick adoption of the second norm takes place.

4. REFERENCES

[1] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proceedings of the international joint conference on Artifical intelligence*, IJCAI'07, pages 1507–1512, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.

[2] T. Sugawara. Emergence and stability of social conventions in conflict situations. In *Proceedings of the international joint conference on Artifical intelligence*, pages 371–378, 2011.

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