The Impact of Social Placement of Non-Learning Agents on Convention Emergence

(Extended Abstract)

Nathan Griffiths
Department of Computer Science
University of Warwick
Coventry, CV4 7AL, UK
nathan.griffiths@warwick.ac.uk

Sarabjot Singh Anand
Department of Computer Science
University of Warwick
Coventry, CV4 7AL, UK
s.s.anand@dcs.warwick.ac.uk

ABSTRACT

Social conventions are important for establishing and maintaining coordination in groups of agents, especially where there is no centralised control. As individuals interact, learn, and update their strategies, effective coordination can be achieved through the emergence of suitable conventions. In this paper we (i) show how the structure of a population affects convention emergence, (ii) demonstrate how fixed strategy agents can manipulate emergence, and (iii) evaluate strategies for inserting fixed strategy agents.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence — *Multiagent systems*

General Terms

Experimentation

Keywords

Conventions, Norms, Emergence, Social Influence

1. INTRODUCTION

Social conventions are behaviours or strategies that are generally accepted in a society as describing how to act in a particular situation, and effective conventions can facilitate effective coordinated action. Where centralised control is lacking, conventions can emerge from the local interactions and observations of self-interested individuals [1]. This is a form of *social learning* in which individuals learn from repeated interactions with multiple agents in the population. Many previous investigations assume that agents can perceive the actions, strategy and payoffs of those with whom they interact. Although sometimes possible, in general we cannot make such an assumption, and so we *limit an agent's perception to knowledge of its own payoff.* There has been little exploration of settings in which observations are restricted in this way, with some notable exceptions such as the work of Sen *et al.* [2, 3] and Villatoro *et al.* [4]. Previous work on convention emergence has also typically given little consideration to the importance of the network topology that constrains interactions, the size of the action space (i.e. the number of possible actions, or candidate conventions), the effect of previous interactions on the rewards received, and the effect of fixed strategy non-learning agents.

Where agents in a population learn and adapt based on interactions with others, inserting a small number of non-learning individuals can influence the direction in which the population evolves [2]. In this paper we investigate the effect of fixed strategy (FS) agents on convention emergence, while addressing some of the limitations of previous work.

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.

2. THE SOCIAL LEARNING MODEL

We consider agents that are situated in a network topology, with agents' interactions being restricted to their neighbours. Many previous investigations have considered completely connected or regular networks. However, in most social networks the degree distribution of nodes is typically highly skewed, with a few nodes having an unusually high degree. In this paper we explore topologies that represent properties observed in real-world environments, namely *scale-free* and *small-world* networks, along with *random* networks as a base case for comparison[1].

We modify the interaction game defined by Villatoro *et al.* [4] to support m actions ($m > 2$). The reward for an interaction depends on the current and previous choices, modelling the social pressure that arises from the history of interactions. Each agent x has a fixed length FIFO memory M_x recording the most recent l actions that it has selected. Each time step each agent randomly selects one of its neighbours, and both agents choose which of the m actions they will take. If an agent selects the majority action, as represented in the combination of the two memories, then its payoff is equal to the proportion of the majority actions that it was responsible for, otherwise it receives nothing. Specifically, when an agent x interacts with another y, the reward r_x it receives for action a_x is given by:

$$r_x = \begin{cases} \frac{M_x^{a'}}{M_x^{a'} + M_y^{a'}}, & \text{if } a_x = a' \\ 0, & \text{otherwise} \end{cases}$$

where $M_x^{a'}$ is the number of times action a' appears in agent x's memory and a' is the majority action. An agent's per-

[1] We use the generator implementations provided by JUNG (v2.0.1): http://jung.sourceforge.net/

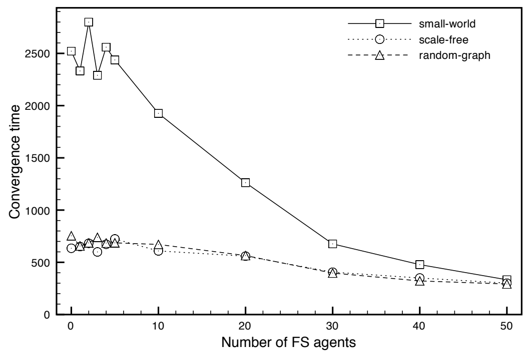

Figure 1: Time for convergence with random, scale-free and small-world topologies using an action space of size $m = 2$ and random placement of FS agents.

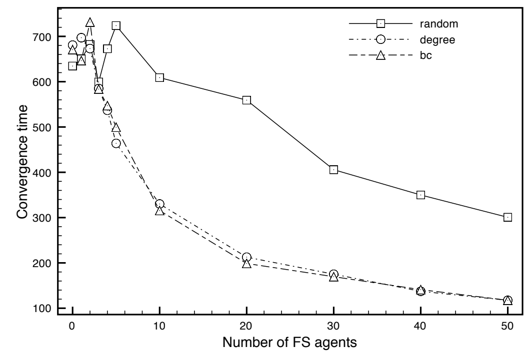

Figure 2: Time for convergence using FS agents placed according to random, degree, and bc with $m = 2$ in a scale-free topology.

ception is restricted to the payoff received in an interaction; agents cannot observe others' actions, memories, or payoffs.

In order to select an action agents use a learning algorithm to estimate the desirability of each possible action. We adopt the approach taken by Villatoro *et al.* [4] of using a simplified Q-Learning algorithm. For each action $a \in A$ each agent maintains an estimate of the utility of choosing that action (a Q-value), which is updated according to: $Q^t(a) = (1 - \alpha) \times Q^{t-1}(a) + \alpha \times reward$ where $Q^t(a)$ is the estimated utility of action a after selecting it t times, α is the learning rate, and $reward$ is the payoff received from the current interaction. With some probability $p_{explore}$ an agent will explore by selecting an action at random, otherwise it selects the action that has the highest Q-value.

In this setting we consider the effect of non-learning *fixed strategy* FS agents, which are each given one of the possible actions as a fixed strategy. We explore two alternatives: (i) all FS agents have the same strategy, with the motivation of reducing the convergence time and (ii) each of the m-actions are uniformly distributed among the FS agents, with the motivation of slowing convergence and maintaining diversity.

3. EXPERIMENTAL RESULTS

In the simulations described below we use a learning rate of $\alpha = 0.5$ and an exploration probability of $p_{explore} = 0.25$. The Q-values for each action are initialised to zero, and each agent's memory is of length $l = 5$ and is initially empty. We use a population of $N = 500$ agents (we see similar trends for $N = \{100, 1000\}$). Each topology was generated to have approximately the same number of edges (1500), using the following parameters: (i) random-graph: $p = 0.012$, (ii) scale-free: $v = 25$ and $e = 3$, and (iii) small-world: $c = 1$ and $\alpha = 2.0$. We adopt Kittock's convergence criteria [1], considering the population to have converged when 90% of the regular agents (non-FS agents), when not exploring, select the same action. Simulations are run for 10000 learning steps, and results are averaged over 100 simulation runs.

Figure 1 shows the effect of the network topology on convergence time as the number of fixed strategy agents increases. In all cases, convergence time reduces as the number of fixed strategy agents increases. Interestingly agents in a small-world network converge to a single convention at a much slower rate than those in scale-free or random graphs. While in the absence of fixed strategy agents the difference in convergence time between small-world and other networks is largest, the convergence times tend to become more similar as the number of fixed strategy agents increases. The difference is insignificant once the number of fixed strategy agents reaches 50, while in the absence of FS agents small-world networks take approximately four times as long to converge when compared with random and scale-free topologies.

Figure 2 shows the results, for a scale-free topology, of placing FS agents by degree and betweenness centrality (bc), along with the baseline random placement strategy as used in Figure 1. As with random placement, increasing the number of fixed strategy agents decreases the convergence time when using degree and bc for placement. Once the number of FS agents is greater than 5 or 6, the degree and bc strategies outperform random placement. The difference in performance between degree and bc is insignificant for random and scale-free topologies, while for small-world degree outperforms bc once the number of FS agents is greater than 5 (graphs for random and small-world topologies are omitted due to space). This is explained by the values for Pearson's correlation between degree and bc for agents within random, scale-free and small-world networks of 0.95, 0.95 and 0.79 respectively, meaning that the same agents are typically selected by degree and bc in random and scale-free networks.

We have performed further experiments that show increasing the size of the action space m increases the time for convergence, and that this increase is fairly consistent across topologies. We have also explored the impact of FS agents having different fixed strategies, and our results show that giving FS agents different strategies can be effective in delaying convergence, with scale-free and random topologies being more manipulable than small-world.

4. REFERENCES

[1] J. E. Kittock. Emergent conventions and the structure of multi-agent systems. In *Lectures in Complex systems*, pages 507–521. Addison-Wesley, 1993.

[2] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained interactions in diverse societies. In *AAMAS-08*, pages 575–582, 2008.

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

[4] D. Villatoro, S. Sen, and J. Sabater-Mir. Topology and memory effect on convention emergence. In *WI-IAT-09*, pages 233–240, 2009.