Manipulating Conventions in a Particle-based Topology

James Marchant and Nathan Griffiths

Department of Computer Science, University of Warwick,
Coventry, CV4 7AL, United Kingdom
{james, nathan}@dcs.warwick.ac.uk

Abstract. Coordination is essential to the effective operation of multi-agent systems. Convention emergence offers a low-cost and decentralised method of ensuring compatible actions and behaviour, without requiring the imposition of global rules. This is of particular importance in environments with no centralised control or where agents belong to different, possibly conflicting, parties. The timely emergence of robust conventions can be facilitated and manipulated via the use of fixed strategy agents, who attempt to influence others into adopting a particular strategy. Although fixed strategy agents have previously been investigated, they have not been considered in dynamic networks. In this paper, we explore the emergence of conventions within a dynamic network, and examine the effectiveness of fixed strategy agents in this context. Using established placement heuristics we show how such agents can encourage convention emergence, and we examine the impact of the dynamic nature of the network. We introduce a new heuristic, Life-Degree, to enable this investigation. Finally, we consider the ability of fixed strategy agents to manipulate already established conventions, and investigate the effectiveness of placement heuristics in this domain.

Keywords: Dynamic networks · Conventions · Social Norms · Influence

1 Introduction

Within multi-agent systems (MAS) cooperation and coordination of individuals' actions and goals are required for efficient interaction. Incompatible actions result in clashes that often incur a resource cost, such as time, to the participating agents. The predetermination of which actions clash is not always possible, particularly for large action spaces and dynamic populations.

The emergence of *conventions* is often used to solve these problems. Conventions represent socially-adopted expected behaviour amongst agents and thus facilitate coordinated action choice without the dictation of rules. Convention emergence has been shown to be possible in static networks with minimal requirements, namely agent rationality and the ability to learn from previous interactions [5, 25]. This adds little design overhead, and is of particular importance in open MAS where agent modification is likely to be impractical or impossible.

Fixed strategy (FS) agents continue to choose the same action regardless of its efficacy or the choices of others in the system. Their presence has been shown to affect the direction and speed of convention emergence in static networks. Small numbers of these agents are able to influence much larger populations [21], especially when placed using appropriate heuristics [7, 10]. Fixed strategy agents can also be used to cause a system to abandon an already established convention in favour of an alternative [13, 15].

In many domains, the nature of the relationships between agents is not static. Agents may leave the system, new agents can enter, and the links between agents may change over time. These dynamic interaction topologies induce different system characteristics than those found in static networks. Relatively little work has studied the nature of convention emergence in these types of network.

This paper considers the emergence and manipulation of conventions within dynamic topologies. We introduce a new heuristic, LIFE-DEGREE, to support this investigation, which considers aspects of the dynamic nature of the system when placing fixed strategy agents. We examine the importance of dynamic topology characteristics by comparing the performance of LIFE-DEGREE against previously used heuristics based on network metrics. We then consider the efficacy of the various heuristics when fixed strategy agents are used to destabilise or remove an established convention.

The remainder of this paper is organised as follows: Section 2 discusses the related work on convention emergence, fixed strategy agents and dynamic topologies. Section 3 describes the model of convention emergence being used, as well as the simulation model used to generate the topologies. Additionally this section introduces the heuristics used to place fixed strategy agents. Our results are shown in Section 4 and, finally, we present our conclusions in Section 5.

2 Related Work

A *convention* is a form of socially-accepted rule regarding agent behaviour and choices. Conventions can be viewed as "an equilibrium everyone expects in interactions that have more than one equilibrium" [26]. No explicit punishment exists for going against a convention, nor is there any implicit benefit in the action represented by the convention over other possible actions. Members of a convention expect others to behave a certain way, and acting against the convention increases the likelihood of incompatible action choices and the costs associated with these. Conventions have been shown to emerge naturally from local agent interactions [5, 12, 23, 25] and enhance agent coordination by placing *social constraints* on agents' action choices [22].

Although the terms are often used interchangeably in the literature [17, 21], in this paper we differentiate between conventions and *norms*. Norms typically imply an obligation or prohibition on agents with regards to a specific action. Failure to adhere to norms and exhibit the expected behaviour is often associated with punishments or sanctions [1, 3, 11, 19]. Alternatively, agents may be explicitly rewarded for adherence to norms. Thus, norms generally require additional system or agent capabilities as well as incurring a system-level overhead for punishment/reward. In this paper, we assume that agents do not have the capability to punish one another, nor can they observe defection in others. Instead, we use conventions as a lightweight method of increasing coordination.

We make only minimal assumptions about agent architecture and behaviour; we assume that agents are rational and that they have access to a (limited) memory of previous interactions. Numerous studies have focused on convention emergence with these assumptions [5, 10, 21, 25] and have shown that they allow rapid and robust convention emergence. Walker and Wooldridge [25] investigated convention emergence whilst making few assumptions about agent capabilities. In their model, agents select actions based on the observed choices of others, and global convention emergence is shown to be possible.

Expanding on this, Sen and Airiau [21] investigated social learning for convention emergence, where agents receive a payoff from their interactions which informs their learning (via Q-Learning). They showed convention emergence can occur when agents have no memory of interactions and only observe their own rewards. However, their model is limited in that agents are able to interact with any other member of the population rather than being situated in a network topology. Additionally, the convention space considered is restricted to only two possible actions. In more realistic settings larger convention spaces and more restrictive connecting network topologies are likely. The network topology agents are situated in has been shown to have a significant effect on convention emergence [4,5,12,24], affecting the speed with which emergence occurs. Recent work has shown that a larger number of actions typically slows convergence [7, 10, 18].

The use of fixed strategy agents, who always choose the same action regardless of others' choices, to influence convention emergence has also been explored. Sen and Airiau [21] show that a small number of such agents can cause a population to adopt the fixed strategy as a convention over other equally valid choices. This indicates that small numbers of agents can affect much larger populations.

In Sen and Airiau's model, due to the lack of connecting topology, all agents are identical in terms of their ability to interact with others. However, in many domains, agent interactions may be limited to neighbours in the network. As such, some agents will have larger sets of potential interactions than others. In the context of static topologies, Griffiths and Anand [10] establish that which agents are selected and *where* they are in the topology is a key factor in their effectiveness as fixed strategy agents. They show that placement by simple metrics such as degree offers better performance than random placement.

Franks *et al.* [6, 7] investigated fixed strategy agents where interactions are constrained by a static network topology and agents are exposed to a large convention space. They found that topology affects the number of fixed strategy agents required to increase convergence speed. This also expanded on the work of Griffiths and Anand [10] by investigating the effectiveness of placing by more advanced metrics such as eigenvector centrality.

Few studies have focused on convention emergence in dynamic topologies, with most work focusing on static networks. Savarimuthu et al. [20] consider the related phenomenon of norm emergence in a dynamic topology. They show that norms are able to emerge under a number of conditions, but their work differs from ours due to the requirements placed on agents. The interaction model used requires agents to maintain an internal norm as well as being able to query other agents. We make minimal assumptions about agent internals or the information available. Additionally, our work investigates the manipulation of convention emergence, something not considered by Savarimuthu et al.

Mihaylov et al. [16] briefly consider convention emergence in dynamic topologies using the coordination game. However, their work focuses on a new proposed method of learning, rather than on the emergence itself. In particular, they do not consider fixed strategy agents, or the action that emerges as a convention. In this paper, we consider both convention emergence in dynamic topologies and the use of fixed strategy agents to understand the impact of network dynamics.

Relatively little work has considered destabilising established conventions, with previous investigations of fixed strategy agents typically inserting them at the beginning of interactions. We have previously [13,15] investigated using fixed strategy agents in static topologies to cause members of the dominant convention to change their adopted convention and hence *destabilise* it. We found that this required substantially more fixed strategy agents than is needed to influence conventions before emergence. This paper expands on this work to examine aspects of dynamic networks when selecting fixed strategy agents for destabilisation. We also expand on [14] and consider the general nature of convention emergence in dynamic topologies, particularly without the use of fixed strategy agents, and the effect of topology features on convention emergence time. Finally, we explore the relationship between placement heuristics, number of fixed strategy agents and the speed of convention emergence.

3 Convention Emergence Model

Our experimental setup consists of three main components, introduced below: the network topology, the interaction regime used by agents and the heuristics used for placing fixed strategy agents.

3.1 Dynamic Topology Generator

Similar to Savarimuthu et al. [20] we utilise a particle-based simulation, developed by González et al. [8, 9], to model dynamic network topologies with characteristics comparable to those observed in real-world networks. Agents are represented as colliding particles and the topology is modified by collisions creating links between the agents. A population of N agents, represented as a set of particles with radius r, is placed within a 2D box with sides of length L. Initially, all agents are distributed uniformly at random within the space and are assigned a velocity of constant magnitude v_0 and random direction.

Each timestep, agents move according to their velocity and detect collisions with other agents. When two agents collide, an edge is added between them in the network topology if one does not already exist. Both agents then move away in a random direction with a speed proportional to their degree. Thus, higher degree nodes have an increased probability of further collisions, which in turn further increases their degree. In this way, the model exhibits preferential attachment, a characteristic found in static scale-free networks [2]. Such networks are often studied in the field of convention emergence [5, 7, 10, 18] due to characteristics that are representative of real-world networks.

Additionally, all agents are assigned a Time-To-Live (TTL) when created. This is drawn uniformly at random between zero and the maximum TTL, T_l. After each timestep agents' TTLs are decremented by one. When an agent's TTL $= 0$ the agent and all its edges are removed. A new agent is placed at the same location within the simulation with the randomised initial properties discussed above. In this manner, the topology is constantly changing.

Different topologies can be characterised by the value of T_l/T_0 where T_0 is the characteristic time between collisions. This can be expressed as:

$$\frac{T_l}{T_0} = \frac{2\sqrt{2\pi} r N v_0 T_l}{L^2} \tag{1}$$

González et al. show that this value dictates key characteristics of the generated topology, primarily the average degree and degree distribution.

The concept of a quasi-stationary state (QSS) is discussed by González et al., such that a QSS emerges after a number of timesteps and is characterised by macro-scale stability of network characteristics. Micro-scale characteristics, for individual agents, remain in flux. In [8] it is shown that the QSS can be described as any timestep, t, where $t \gtrsim 2T_l$. Our approach here differs from Savarimuthu et al. [20] as we consider agent interactions starting from $t = 0$ rather than waiting for the QSS. This allows us to mimic scenarios where agents have been placed in a new environment rather than only considering already established networks.

3.2 Interaction Regime

Agents within the system interact with one another and, learning from these interactions, converge to a shared behaviour in the form of a convention. Agent interactions occur during each timestep of the regime. In each timestep, every agent chooses one of its neighbours in the network at random. These agents play a round of the n-action pure coordination game. In this game, both agents are given a choice from a set of n-actions, A. Agents do not know what their opponent has chosen. The payoff that each agent receives depends on the combination of the chosen actions: if both chose the same action they receive a positive payoff, otherwise a negative payoff. Alternative payoff matrices and their effect on the effectiveness of the intervention strategies are discussed in Section 4.4.

Each agent monitors their expected payoff for each action, based on the previous payoffs they have received when choosing that action. We adopt the approach of Villatoro et al. [24] in this regard by using a simplified form of Q-Learning. For each action, $a \in A$, the agent maintains a Q-Value which is updated by $Q^i(a) = (1 - \alpha) \times Q^{i-1}(a) + \alpha \times \mathit{payoff}$ where α is a parameter known as the learning rate and i represents the number of times a has been chosen. All agents start with $Q^0(a) = 0, \forall a \in A$. To combat the issue of local optima, we allow each agent, with probability $p_{explore}$ to randomly select an action. Otherwise, as each agent is rational, they will always select the action with the highest Q-Value, selecting randomly between ties.

In the formulation proposed by Kittock [12], a convention is considered to have emerged when a high proportion (90%) of non-fixed strategy agents, when not exploring, would choose the same action. We adopt this definition of a convention but modify it to better fit the dynamic nature of the network topology. Instead of considering the entire population, we monitor adoption within the largest connected component. This follows from the findings of Gonzalez et al. [8] that in most simulations a giant cluster consisting of nearly all agents will emerge. Agents not within this cluster are likely to be recently created agents and, as such, should not be included in the adoption rate calculation as they have not interacted. This is reinforced by our simulations which showed that most agents not within the largest connected component had degree zero. Similarly, 100% adoption is unlikely due to new agents joining.

The Kittock criteria sets a high threshold and measures nearly pervasive conventions. If a convention does not emerge at this threshold there is often still a highly dominant strategy in the system. By considering a different threshold and defining these as conventions, we can examine the effectiveness of the heuristics in situations not normally considered. This approach can be seen in Section 4.4.

Fixed strategy agents will be placed within the network to study the effect on convention emergence. These agents will replace selected agents upon insertion, keeping all of that agent's edges. This can be justified in real-world scenarios as persuading the agents to act in a desired manner via some reward mechanism. Such agents will be assigned the same fixed strategy and their placement will be determined heuristically as discussed below. If a fixed strategy agent's TTL should reach zero, a new agent will be selected using the same heuristic.

We consider two different scenarios: placing fixed strategy agents at the beginning of a system's life, to encourage and direct initial convention emergence, and inserting fixed strategy agents once a convention has emerged to attempt to change it. In the former case, the fixed strategy will be randomly chosen from the available actions. In the latter, it will be randomly chosen from the available actions excluding the already established convention. Initial insertion will occur once a connected component of size greater than $N/2$ has emerged. This prevents convention emergence being declared prematurely for a non-giant cluster. Additionally, placement heuristics which rely on network metrics (such as degree) may select sub-optimal agents if used before a main cluster has emerged.

3.3 Placement Heuristics

Previous work has utilised placement heuristics to enhance the effect of fixed strategy agents. Metrics such as degree, eigenvector centrality and betweenness centrality have been used with greater efficacy than random placement [6,10]. In this paper, we focus on degree-based placement. However, the dynamic nature of the topology introduces a number of ways to apply it. All heuristics are calculated with respect to the largest connected component.

Our initial heuristic, Static Degree, corresponds to the equivalent heuristic for static networks. At the time of insertion, agents are chosen to be fixed strategy agents in descending order of degree. This selection is static once chosen, only being modified upon agent expiration as detailed above. This simplistic approach is computationally cheap, a factor of importance in settings where gathering or computing this information is expensive. However, this risks selected agents potentially becoming sub-optimal choices as the simulation progresses. The static nature of this heuristic means that if another agent acquires a larger degree it will not be selected until one of the current agents expires. Depending on the TTL of the current fixed strategy agents, this could be a substantial period.

To address this issue we propose another degree-based heuristic: Updating Degree. This approach is sensitive to the dynamic nature of the topology and reselects the fixed strategy agents each timestep, based on highest current degree. Whilst this offers a solution to the potential sub-optimality of Static Degree it suffers from two problems. Firstly, the ability to acquire this information each timestep in a timely manner may be infeasible in many domains. Secondly, there is the potential that the fixed strategy agents will not remain in a given location long enough to influence the local area before being replaced.

The Static and Updating Degree heuristics do not fully consider the dynamic network context. Whilst high degree agents are likely to be influential due to their ability to interact with many others, additional dimensions may affect their applicability. Agents close to expiring may be less desirable than younger agents as their expected number of interactions before replacement is lower. However, the youngest agents, those newly created, cannot be guaranteed to become influential later on. Hence, the age of an agent adds an additional consideration. We propose a new heuristic, LIFE-DEGREE, that allows exploration of the effect of age in addition to degree on a fixed strategy agent's efficacy.

In many settings it may be impossible to *know* an agent's TTL. However, we can estimate an agent's remaining life. Given the upper bound, T_l, and the uniformly distributed nature of TTL, the normalised expected remaining TTL, E_{rTTL}, for an agent $n \in N$ is:

$$E_{rTTL}(n) = 1 - \frac{age(n) \times 2}{T_l} \tag{2}$$

We can also calculate the normalised degree of a node within the largest connected component as:

$$deg_{norm}(n) = \frac{deg(n)}{\max_{n' \in LCC} deg(n')} \tag{3}$$

The LIFE-DEGREE heuristic is then defined as:

$$\text{LIFE-DEGREE}(n) = \omega \times deg_{norm}(n) + (1 - \omega) \times E_{rTTL}(n) \tag{4}$$

In this, $0 \leq \omega \leq 1$ is a weight, determining the relative contributions of degree and expected TTL.

Life-Degree allows combination of the relevant information, normalised against theoretical maximums, in a manner that allows exploration of the importance of both. Two variations of Life-Degree will be used, Static and Updating, to compare against the heuristics discussed above.

4 Results and Discussion

In this section we present our findings on convention emergence in dynamic topologies and consider the effect of agent age via our proposed heuristic, Life-Degree. Unless otherwise mentioned, all experiments used 1000 agents, the 10-action coordination game and an exploration and Q-Learning rate of 0.25. Results were averaged over 100 runs. A payoff of +4 for coordinated actions and −1 for conflicting actions was used. This was found to rapidly emerge thorough and robust conventions. Additional payoff schemes are considered in Section 4.4.

4.1 Characterising Topology

We initially consider convention emergence without external manipulation in dynamic topologies. This gives insight into the impact of network dynamics on convention emergence and provides a baseline. Additionally, it allows us to quantify the point at which a stable convention will have emerged for later experiments that focus on destabilisation.

The features of the dynamic topology can be manipulated by varying the parameters of the network model, and are encapsulated in different values of T_l/T_0. González et al. [9] show that the features of the topology thus only depend on the ratio T_l/T_0 and the density, $\rho \equiv N/L^2$. Additionally, they show that the average degree is a non-linear function of T_l/T_0 that depends on the chosen ρ. As such, for all experiments we use a constant $\rho = 0.625$ (i.e. $N = 1000$, $L = 40$) to allow meaningful comparisons of the T_l/T_0 values.

Parameter settings were chosen that generated values of T_l/T_0 between 0 and 20. These were rounded to the nearest integer to combine similar T_l/T_0 values, with each bucket containing 10 values. The average time taken, over 30 rounds, for convention emergence to occur was measured on the generated topologies and the average time over the bucketed values was then calculated. Values which did not result in convention emergence after 20,000 timesteps were discounted from the second average as they were unlikely to result in conventions emerging. Only runs with $T_l/T_0 \lesssim 4$ are affected by this. Simulations with a higher T_l/T_0 exhibited convention emergence for all runs. With $T_l/T_0 \lesssim 4$ as much as 80% of the runs for a given simulation did not result in convergence. The transition is notable and is discussed below.

It is clear that convention emergence is successful in the dynamic topology, and for most values of T_l/T_0 there is little variation in the average time for convention emergence as shown in Figure 1. Values of $T_l/T_0 \gtrsim 5$ all have a

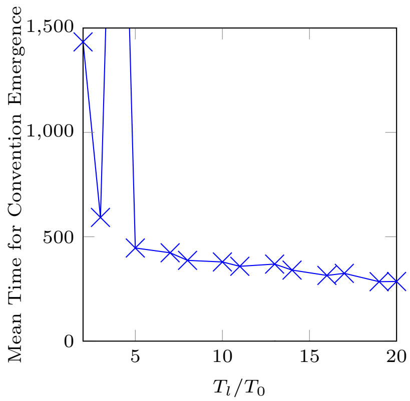


Fig. 1: Average convention emergence time for different values of T_l/T_0 with no fixed strategy agents.

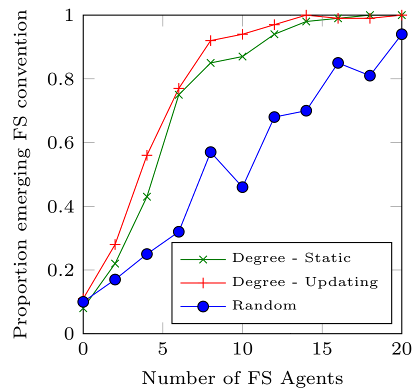


Fig. 2: Proportion of runs in which the fixed strategy emerged as a convention after initial intervention using standard heuristics

convention emergence time of around $t = 500$ with little variation between runs. However, values of $T_l/T_0 \lesssim 4$ displayed significant variation and, in general, much more time was required for convention emergence to occur if it occurred at all. Higher values of T_l/T_0 did not exhibit this.

At low T_l/T_0 values the topology either did not generate a giant cluster or agents were found to expire before meaningful convention emergence could occur. This follows from the parameter settings required to give a small T_l/T_0 and means that there is a lower threshold for the topology to experience convention emergence. In particular, there is a minimum level of connectedness and lifespan that must be present. Below this threshold the network will be partially disconnected and not representative of real-world topologies. However, once this is achieved the time required for convention emergence is mostly independent of T_l/T_0. As such, we select parameter settings that are used for all following simulations that give $T_l/T_0 = 4.7$ which was found to provide stable convention emergence times. For completeness, additional T_l/T_0 values in the range 20 to 200 were also examined. There was a slight decrease in the average time at higher values, although the low variation remained. As the real-world networks examined by González et al. had T_l/T_0 values around 5-6 these results were purely to determine the impact of high T_l/T_0 values, and have not been included.

4.2 Initial Intervention

Having established that convention emergence occurs in dynamic topologies, we now examine the effect of fixed strategy agents. We start by considering the scenario where fixed strategy agents are introduced early in a system's lifespan to manipulate convention emergence. As discussed in Section 3, this initial insertion is delayed until a cluster of size greater than $N/2$ has emerged. This was found empirically to always have occurred by $t = 200$. Fixed strategy agents are inserted after this "burn-in" period has elapsed.

We begin by considering the initial heuristics discussed in Section 3: Static Degree and Updating Degree. We also consider random placement of the fixed strategy agents as a baseline. The fixed strategy agents were inserted into the system at $t = 200$ and the simulation allowed to run for 5000 timesteps. Prior simulations showed that conventions always emerged well before this time even without the presence of fixed strategy agents. The number of fixed strategy agents inserted into the system was varied from zero to twenty and the proportion of simulations in which the fixed strategy emerged as the convention was monitored. The results of this setting are shown in Figure 2.

As expected, given the size of the action space (10), when no fixed strategy agents were inserted, the proportion of times the fixed strategy emerged as the convention is approximately 0.1. With the introduction of only a few fixed strategy agents placed at targeted locations we are able to readily manipulate the emerged convention more than 50% of the time. The results also show that even randomly placed fixed strategy agents are able make a large difference in convention emergence. This corroborates the findings in previous work on static networks [10, 21], although larger numbers of fixed strategy agents are needed comparatively. As the number of inserted agents increases, the difference between the targeted heuristics and random placement becomes more pronounced. The targeted heuristics are able to cause convention emergence in nearly 100% of cases with only 12 agents whilst random placement requires 20.

Importantly, there is little difference between the two targeted heuristics. Updating Degree slightly outperforms Static Degree although in most cases this is not statistically significant (only 4 and 10 FS agents exhibited differences at a 10% significance level with most showing a Z-Score less than 1.0). Given this, and the additional complexity and resource requirements for calculating the Updating Degree heuristic, Static Degree is likely sufficient in most cases.

Having established the efficacy of the traditional heuristics, we now examine the effect of considering agent age using our new heuristic, Life-Degree. We begin by examining Static Life-Degree, contrasting this to Static Degree. Various weightings of Life-Degree were considered and the results are presented in Figure 3. The results of Static Degree have also been included for comparison.

When given equal weighting between expected life and degree ($\omega = 0.5$), Life-Degree performs markedly worse than Static Degree for nearly all numbers of fixed strategy agents. This is due to the fact that such a weighting is heavily biased to much younger agents. The range of possible ages is larger than that of degree and as such, even when normalised, age was found to be the primary selector. A weighting of 0.7 in favour of degree exhibits the same performance as Static Degree (within a 90% confidence interval). Further increasing the weighting offers no further improvement in performance, with $\omega = 0.9$ also performing the same as Static Degree. Additional weightings of 0.95 and 0.99 were also considered and similarly offered no improvements.

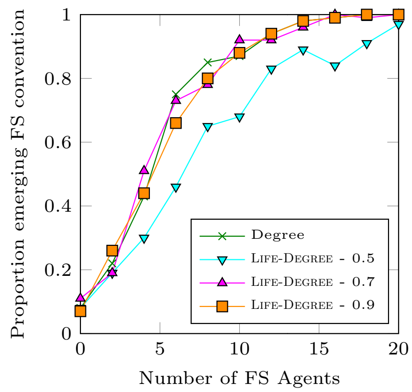


Fig. 3: Proportion of runs in which the fixed strategy emerged as a convention after initial intervention using Static Degree and Life-Degree

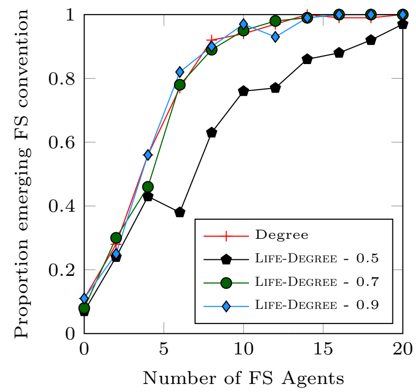


Fig. 4: Proportion of runs in which the fixed strategy emerged as a convention after initial intervention using Updating Degree and Life-Degree

These results show that an agent's connectivity, indicated by its degree, is a much larger contributor to its ability to influence others than how long that agent will remain in the system. The fact that considering age can only decrease the effectiveness of the chosen agents indicates that agents' short-term influence is a larger factor in convention emergence than choosing long-term targets.

Life-Degree was also used in an updating manner, such that the set of fixed strategy agents was recalculated each iteration. The results from this and, for comparison, Updating Degree are shown in Figure 4. Similar to the Static Life-Degree experiments, the performance of Updating Life-Degree depends heavily on the value of ω being used. As before, giving equal weighting to each factor results in poor performance, far below that of pure degree. Increasing the weighting again enhances performance but only to that of Updating Degree. This mirrors the results of Static Life-Degree and shows that, regardless of the ability to continuously assess an agent's remaining lifespan, choosing agents with numerous connections is the most important factor. This indicates that, even in the extreme case where an agent is expected to expire in a few timesteps, on average equal performance can be achieved when selecting them compared to selecting an agent who remains in the system much longer.

Static Life-Degree and Updating Life-Degree, like their pure degree counterparts, have only slight differences in performance, with Updating Life-Degree performing slightly better. At each weight, Updating Life-Degree outperforms Static Life-Degree at a 10% significance level for several numbers of FS agents. This is most pronounced when $\omega = 0.9$ where Updating Life-Degree performs significantly better between 4-8 FS agents. However, the constant information updates may make Updating Life-Degree untenable in many domains. In domains where this information is readily available, we have

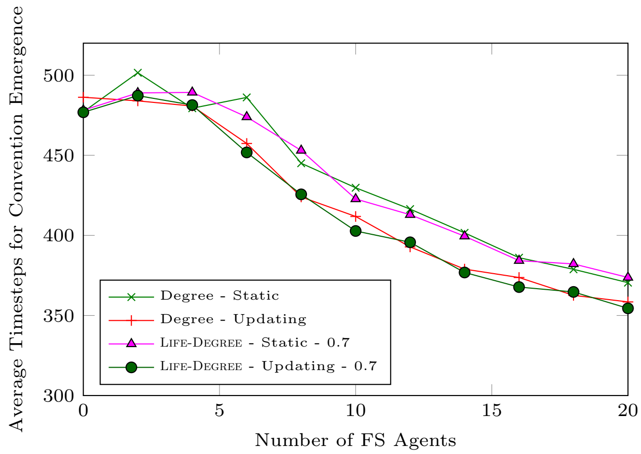


Fig. 5: Effect of fixed strategy agents on convention emergence speed

shown that using up-to-date estimates of degree is sufficient to offer improved outcomes from fixed strategy agent selection.

The results presented above show that it is possible to influence the direction of convention emergence in dynamic topologies. Another commonly used metric of the efficiency of fixed strategy agents is the effect they have on the *speed* of convention emergence [7, 10]. Figure 5 shows how time for convention emergence varies for different numbers of fixed strategy agents using the heuristics. As is to be expected, given the asymptotic behaviour exhibited above, consideration of age, depending on weighting, causes either an increase in the average time required or results in similar times to the equivalent pure degree heuristics. Omitted from the graph for clarity, a value of $\omega = 0.5$ requires more time for convention emergence to occur for any number of fixed strategy agents. Values higher than 0.7 perform similarly to 0.7 and hence have also been omitted.

The standard deviation of the convention emergence time also decreases rapidly as the number of fixed strategy agents rises, from up to 100 with zero agents to around 20 with 20 agents. The standard deviation of the results from the Life-Degree simulations are equivalent to those of the pure degree heuristics except for $\omega = 0.5$ which exhibits much larger variance. Thus, consideration of age has a negative effect both in establishing conventions as well as the time it takes to do this. This indicates that, in all aspects, degree is the factor that contributes most to how influential a given agent will be.

4.3 Late Intervention

We now look to the related use of fixed strategy agents in *destabilising* and replacing an already established convention [13, 15]. This requires a convention to already have emerged within the system. So that the results are representative of the general case, we allow a convention to naturally emerge without the use

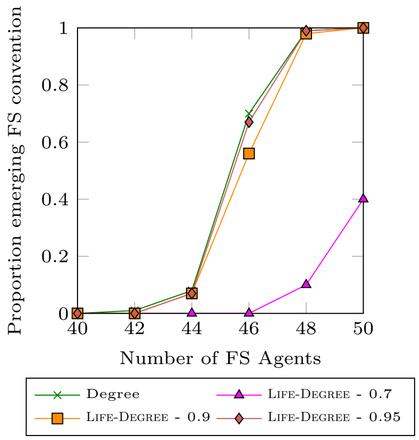


Fig. 6: Proportion of runs in which the fixed strategy emerged as a convention after late intervention using Static Degree and LIFE-DEGREE

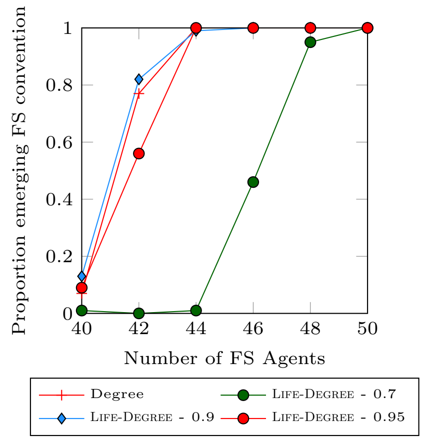


Fig. 7: Proportion of runs in which the fixed strategy emerged as a convention after late intervention using Updating Degree and LIFE-DEGREE

of fixed strategy agents to encourage it. It was found that conventions always emerged before timestep $t = 1500$ and, as such, insertion of fixed strategy agents occurs at this time. This also means that the system will have entered the QSS. The action of the fixed strategy is chosen uniformly at random from the actions that exclude the established convention.

In common with the findings of Marchant et al. [13,15] for static networks, our initial experiments showed a much larger number of fixed strategy agents was required to affect the established convention compared to the number needed when inserted into a system earlier. However, a relatively small set of fixed strategy agents are still able to effect a change. In contrast to static networks, the transition between no effect and guaranteed change occurs over a much smaller range of fixed strategy agents. For nearly all heuristics (excluding random) there is little or no effect at 40 fixed strategy agents (4% of the population), whilst 50 fixed strategy agents (5% of the population) results in the targeted convention supplanting the established convention in almost 100% of cases. This narrow window indicates that there is a critical number of fixed strategy agents that is required to guarantee replacement of a convention in dynamic topologies.

Figure 6 shows the proportion of runs in which the convention represented by the fixed strategy became established when using the static heuristics: Static LIFE-DEGREE and Static Degree. In common with initial intervention, consideration of age induces poorer performance here. With $\omega = 0.7$, LIFE-DEGREE is substantially outperformed by Static Degree for any non-trivial proportion, in contrast to the case in initial intervention when such a weighting produced similar performances. Even when increasing the weighting to 0.9, previously equivalent to the performance of pure degree, Static LIFE-DEGREE is still slightly outperformed by Static Degree though this is within the margin of error (only 46 FS agents produce significant differences at a 10% level). The performance of higher weights asymptotically approached that of Static Degree.

Similar results are presented in Figure 7 for updating heuristics. The difference between Updating LIFE-DEGREE and Updating Degree in this scenario is even more pronounced. A weighting of 0.7 is again substantially worse than the pure degree heuristic with the higher weightings, 0.9 and 0.95, being of similar quality to Updating Degree.

Of note, the difference in performance between static heuristics and updating heuristics is more pronounced here than in initial interventions; the updating heuristics consistently require significantly fewer fixed strategy agents to effect a change. This indicates that inclusion of up-to-date information of agent state is more important when combating an existing convention and makes a larger contribution compared to establishing a convention from a state of neutral agents.

These findings indicate that destabilisation of an existing convention is even more sensitive to the consideration of agent longevity than initial convention emergence. Indeed, the age or expected lifespan of an agent can be safely ignored with no detrimental effects to the performance of the fixed strategy agents. This strongly implies that the major factor in destabilising conventions is instead choosing agents with high degree, regardless of how long that agent will last. High degree is more effective at spreading influence than choosing a lower degree agent with longer life. The difference between Static and Updating Degree, not present in initial intervention, also supports this view; the importance of choosing the current highest degree agents is far more pronounced.

4.4 Alternative Payoffs

We now turn our attention to the effect the payoff matrix has on intervention effectiveness. In particular, we examine whether the positive and negative rewards the agents receive (and the symmetry or asymmetry of these) changes the relationship or relative performance of the various placement heuristics.

This exploration uses 3 different payoff matrices: *4v-1* (positive reinforcement), *1v-1* (neutral reinforcement) and *1v-4* (negative reinforcement) where the first number represents the payoff for coordinated strategy choice, the second the payoff for conflicting strategy choice. *4v-1* is the payoff structure that has been used in all previous experiments and represents situations where coordination is more beneficial than conflict is harmful, or where coordination is more encouraged. For example, attempting to find a mutual radio channel over which to communicate; whilst there is an expenditure of time for each failure, it is not necessarily very harmful whilst correctly communicating is very beneficial. This structure has been used in previous work [21] and has been shown to allow rapid and thorough convention emergence. *1v-1* can instead represent situations where there is symmetry between the benefit and harm, such as choosing which side of a corridor to walk on; there are both minor inconveniences and minor benefits but neither of a larger scale than the other. Finally, *1v-4* represents situations where conflicting action choices could be very detrimental and should

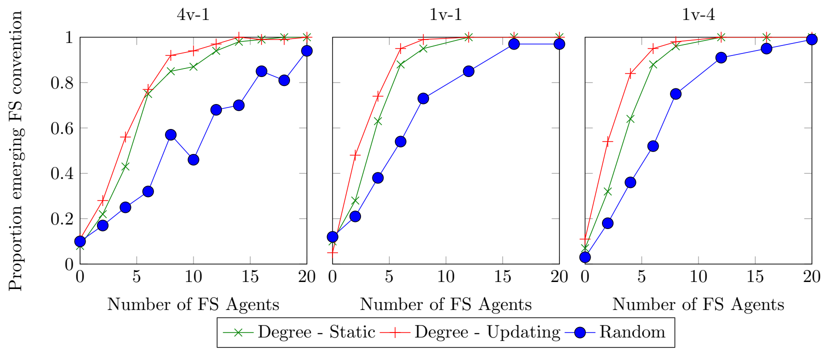


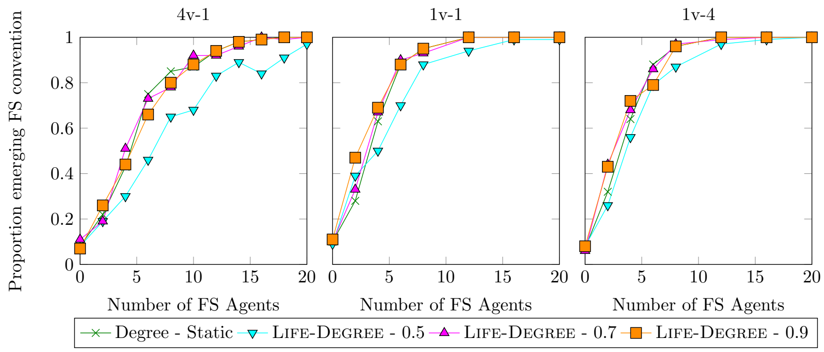
Fig. 8: A comparison of the effect of game payoff on the effectiveness of the standard heuristics in encouraging initial convention emergence.

be discouraged rapidly. An example of this is which side of the road to drive on (although this is often described using a symmetric payoff); the negative effects of a crash are substantial.

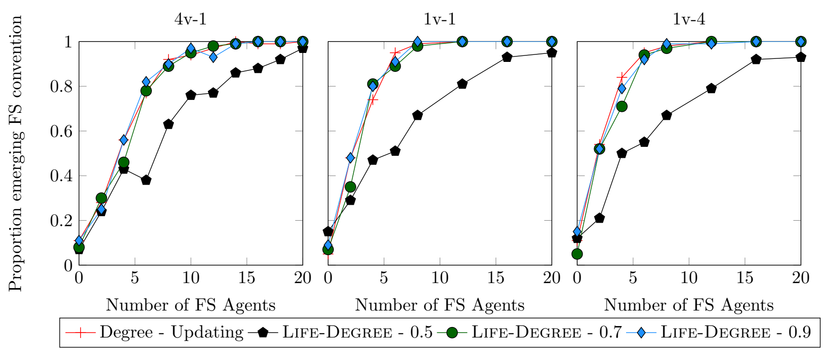
As this paper is concerned with the effectiveness of strategies to either direct or replace convention emergence, we primarily concern ourselves with the payoff matrix that best enables conventions to emerge rapidly and thoroughly, so the strategies can be studied. The *4v-1* payoff matrix performs best in this regard. Indeed, the other payoff matrices nearly always fail to reach the Kittock criteria of 90% for convention emergence, even when the simulation is run for 50000 timesteps. This is related to both the number of strategies available and the payoff matrix. As the number of strategies increases, the average percentage of agents adhering to the primary convention decreases and, with the alternative payoff structures, falls below the 90% Kittock threshold for our strategy space of 10. Whilst the positive reinforcement system teaches agents which choice is best, the other payoff structures instead teach agents which choices are worst. Due to the asymmetry of this, and the fact that coordination is not as heavily rewarded, the level of coordination is lower.

However, although the 90% threshold of the Kittock criteria is not met, there is in general still a singular strategy that dominates agent choice and if we reduce the threshold to 80% we can view this as convention emergence. Lowering the threshold of convention emergence enables us to compare the effectiveness of the strategies under different payoff matrices whilst still considering situations where the system is heavily dominated by a single strategy. All the results in this subsection use the 80% threshold, with all other parameters kept as defined at the beginning of this section.

Initial Intervention We begin by considering the payoff matrices as applied to initial intervention. Using the same heuristics and weightings as before the simulations were run with the three different payoff matrices and, using a threshold of 80%, the proportion of runs in which the fixed strategy emerged above



(a) Static Heuristics Comparison

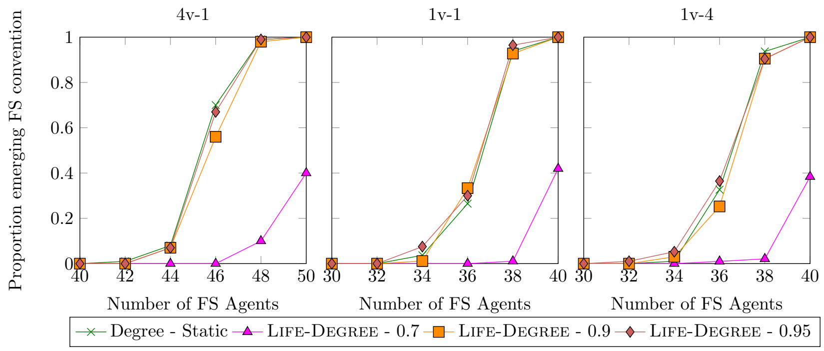


(b) Updating Heuristics Comparison

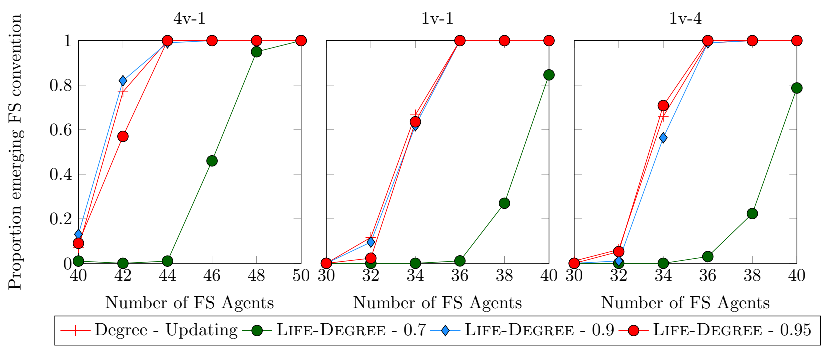
Fig. 9: A comparison of the effect of game payoff on the effectiveness of the advanced heuristics in encouraging initial convention emergence.

this threshold was compared. Figure 8 shows the comparison when using the standard heuristics. As this figure shows, the relative performance between the random, static degree and updating degree placement strategies never changes. However, the *absolute* performance of each of the heuristics increases in both *1v-1* and *1v-4* with fewer FS agents needed to enact the same change. This is likely due to it being easier for the system to overcome any partial convention that has started to emerge by the time the FS agents are inserted ($t = 200$) as the reward for perpetuating this emerging convention is lower and, in the case of *1v-4*, the negative payoff for conflicting with the FS agents is higher. Additionally, whilst the other payoff structures may provide easier to manipulate systems, *4v-1* is the only one of those examined that reached the Kittock criteria.

Figure 9 shows the comparison for both static and updating placement heuristics. The findings in Figure 8 are also present here: the relative performance between the heuristics does not change as the payoff is altered but performance for all heuristics increases. Additionally, the poor performance of the 0.5 weighting of LIFE-DEGREE has a reduced penalty compared to the other weightings.



(a) Static Heuristics Comparison



(b) Updating Heuristics Comparison

Fig. 10: A comparison of the effect of game payoff on the effectiveness of the advanced heuristics in encouraging replacement of an already established convention via late intervention. Note the different x-axes between the original payoff structure and the others.

Late Intervention Having examined the effect of the payoff matrix on initial interventions we now investigate the effect on late interventions and destabilisation. For these experiments the threshold for both destabilisation and considering a new convention to have replaced the old are both 80%. Figure 10 shows the results for both static and updating heuristics. Of particular note is the difference in x-axis range between *4v-1* and the other payoffs: the former ranges from 40 to 50, the latter from 30 to 40.

Similar to the findings for initial intervention, the relative performances amongst the heuristics are the same across the different payoff matrices. However the absolute performance, in the number of agents needed, is substantially smaller for the latter payoff matrices. This provides additional evidence for the hypothesis discussed above, that it is easier to get agents to switch away from the established convention as the reward for continuing to use it is less compared to switching to the introduced strategy. As with other aspects discussed in this paper the effect is amplified in late interventions compared to similar effects present in initial intervention.

Overall, changing the payoff matrix, either from positive asymmetry to neutral symmetry or negative asymmetry caused no change on the relative effectiveness of the various interventions. Degree placement still performs best in both initial and late interventions and, depending on weighting, can offer large improvements over the consideration of agent age. However, the absolute performance change is interesting and future work will further explore this difference.

5 Discussion and Conclusions

Convention emergence is often used in multi-agent systems to encourage efficient and coordinated action choice. It provides a mechanism through which such behaviour can naturally occur without requiring changes to, or assumptions about, underlying agent capabilities. How best to facilitate robust convention emergence in a timely manner is an area of ongoing research. Fixed strategy agents can be used to speed up and direct emergence. In particular, placing small numbers of fixed strategy agents at targeted locations within the network topology connecting agents has been shown to better facilitate convention emergence than untargeted placement. The heuristics used to choose these locations often make use of metrics derived from an agent's location within the topology.

In this paper, we initially considered uninfluenced convention emergence in a dynamic network, using the topology model proposed by González et al. [8,9]. We showed that conventions emerge in a dynamic environment and that the average time taken for this is largely independent of the parameter settings used in the network model provided the value of T_l/T_0 is above a threshold of approximately 4. Below this, the topology or agent lifespans are not conducive to any convention emergence occurring at all. This indicates that there is a minimum connectedness required in dynamic topologies for conventions to emerge.

We proposed a new placement heuristic, Life-Degree, that utilises information unique to dynamic topologies in its decision making process, allowing us to test the importance of that information. We contrasted this to the performance of the traditionally used placement heuristics. We examined the scenario where fixed strategy agents are introduced early in the life of the system to direct and encourage faster convention emergence. We showed that, as in static networks, targeted placement offers better performance than untargeted. A small number of agents are able to influence a population much larger than themselves. We established that, in domains where it is possible to change the fixed strategy agents after selection, doing so offers small improvements in performance. In both settings, the most important aspect of selected agents was found to be their degree, ignoring their longevity. This both increased the probability of a specific convention emerging as well as increasing the speed of that emergence.

Additionally, we considered the destabilisation of already established conventions in dynamic networks. We found that destabilisation is more sensitive to the inclusion of agent lifespan than when using fixed strategy agents to establish a convention at the beginning of simulation. Choosing locations that will maximise an agent's influence, regardless of how long they will remain, is the most important aspect to consider when destabilising conventions in dynamic networks. Future work will investigate this further and examine if other features of dynamic networks offer beneficial information when selecting fixed strategy agents. We showed that the updating heuristics cause more destabilisation than the static heuristics and that this effect was much larger than the equivalent difference when encouraging initial convention emergence.

Finally we explored the effect that different payoff schemes had on the effectiveness of the heuristics. We showed that the ordering of performance was not affected by the payoff scheme but that the overall effectiveness of all heuristics is sensitive to the rewards the agents receive.

Overall, we have shown that convention emergence is possible in dynamic topologies and that many characteristics have direct parallels in static networks. We have shown that the degree of an agent is a major factor when choosing them and can be used to cause rapid convention emergence and destabilisation.

References

1. Axelrod, R.: An evolutionary approach to norms. American Political Science Review 80, 1095–1111 (1986)
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science 286(5439), 509–512 (1999)
3. Bicchieri, C., Jeffrey, R.C., Skyrms, B.: The Dynamics of Norms. Cambridge Studies in American Literature and Culture, Cambridge University Press, Cambridge (1997)
4. Delgado, J.: Emergence of social conventions in complex networks. Artificial Intelligence 141(1-2), 171–185 (2002)
5. Delgado, J., Pujol, J.M., Sangüesa, R.: Emergence of coordination in scale-free networks. Web Intelligence and Agent Systems 1(2), 131–138 (2003)
6. Franks, H., Griffiths, N., Anand, S.: Learning agent influence in MAS with complex social networks. Autonomous Agents and Multi-Agent Systems 28(5), 836–866 (2014)
7. Franks, H., Griffiths, N., Jhumka, A.: Manipulating convention emergence using influencer agents. Autonomous Agents and Multi-Agent Systems 26(3), 315–353 (2013)
8. González, M.C., Lind, P.G., Herrmann, H.J.: Networks based on collisions among mobile agents. Physica D: Nonlinear Phenomena 224(1-2), 137–148 (2006)
9. González, M.C., Lind, P.G., Herrmann, H.J.: System of mobile agents to model social networks. Physical Review Letters 96, 088702 (2006)
10. Griffiths, N., Anand, S.S.: The impact of social placement of non-learning agents on convention emergence. In: 11th International Conference on Autonomous Agents and Multiagent Systems. vol. 3, pp. 1367–1368. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2012)
11. Kandori, M.: Social norms and community enforcement. The Review of Economic Studies 59(1), 63–80 (1992)
12. Kittock, J.: Emergent conventions and the structure of multi–agent systems. In: Lectures in Complex Systems: the Proceedings of the 1993 Complex Systems Summer School. pp. 507–521. Addison-Wesley, Reading, MA (1995)

13. Marchant, J., Griffiths, N., Leeke, M.: Destabilising conventions: Characterising the cost. In: 8th IEEE International Conference on Self-Adaptive and Self-Organising Systems (2014)
14. Marchant, J., Griffiths, N., Leeke, M.: Convention emergence and influence in dynamic topologies. In: 14th International Conference on Autonomous Agents and Multi-Agent Systems (2015)
15. Marchant, J., Griffiths, N., Leeke, M., Franks, H.: Destabilising conventions using temporary interventions. In: 17th International Workshop on Coordination, Organisations, Institutions and Norms (2014)
16. Mihaylov, M., Tuyls, K., Nowé, A.: A decentralized approach for convention emergence in multi-agent systems. Autonomous Agents and Multi-Agent Systems 28(5), 749–778 (2014)
17. Mukherjee, P., Sen, S., Airiau, S.: Norm emergence with biased agents. International Journal of Agent Technologies and Systems 1(2), 71–84 (2009)
18. Salazar, N., Rodriguez-Aguilar, J.A., Arcos, J.L.: Robust coordination in large convention spaces. AI Communications 23(4), 357–372 (2010)
19. Savarimuthu, B.T.R., Arulanandam, R., Purvis, M.: Aspects of active norm learning and the effect of lying on norm emergence in agent societies. In: Kinny, D., Hsu, J.Y.J., Governatori, G., Ghose, A.K. (eds.) PRIMA 2011, vol. 7047, pp. 36–50. Springer, Heidelberg (2011)
20. Savarimuthu, B.T.R., Cranefield, S., Purvis, M., Purvis, M.: Norm emergence in agent societies formed by dynamically changing networks. In: 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology. pp. 464–470 (2007)
21. Sen, S., Airiau, S.: Emergence of norms through social learning. In: 20th International Joint Conference on Artifical Intelligence. pp. 1507–1512. Morgan Kaufmann, San Francisco (2007)
22. Shoham, Y., Tennenholtz, M.: On the emergence of social conventions: modeling, analysis, and simulations. Artificial Intelligence 94(1-2), 139–166 (1997)
23. Villatoro, D., Sabater-Mir, J., Sen, S.: Social instruments for robust convention emergence. In: 22nd International Joint Conference on Artificial Intelligence. pp. 420–425. AAAI Press (2011)
24. Villatoro, D., Sen, S., Sabater-Mir, J.: Topology and memory effect on convention emergence. In: 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. pp. 233–240. IEEE Computer Society, Washington, DC (2009)
25. Walker, A., Wooldridge, M.: Understanding the emergence of conventions in multi-agent systems. In: International Conference on Multi-Agent Systems. pp. 384–389 (1995)
26. Young, H.P.: The economics of convention. The Journal of Economic Perspectives 10(2), 105–122 (1996)