



Emergence of norms in interactions with complex rewards

Dhaminda B. Abeywickrama^{1,2} · Nathan Griffiths³ · Zhou Xu⁴ · Alex Mouzakitis⁴

Accepted: 27 September 2022
© The Author(s) 2022

Abstract

Autonomous agents are becoming increasingly ubiquitous and are playing an increasing role in wide range of safety-critical systems, such as driverless cars, exploration robots and unmanned aerial vehicles. These agents operate in highly dynamic and heterogeneous environments, resulting in complex behaviour and interactions. Therefore, the need arises to model and understand more complex and nuanced agent interactions than have previously been studied. In this paper, we propose a novel agent-based modelling approach to investigating norm emergence, in which such interactions can be investigated. To this end, while there may be an ideal set of optimally compatible actions there are also combinations that have positive rewards and are also compatible. Our approach provides a step towards identifying the conditions under which globally compatible norms are likely to emerge in the context of complex rewards. Our model is illustrated using the motivating example of self-driving cars, and we present the scenario of an autonomous vehicle performing a left-turn at a T-intersection.

Keywords Norm emergence · Agent interactions · Agent-based modelling · Reinforcement learning

✉ Dhaminda B. Abeywickrama
dhaminda.abeywickrama@bristol.ac.uk

Nathan Griffiths
nathan.griffiths@warwick.ac.uk

Zhou Xu
xzhou1@jaguarlandrover.com

Alex Mouzakitis
amouzak1@jaguarlandrover.com

¹ Trustworthy Systems Laboratory, Department of Computer Science, University of Bristol, Bristol, UK

² Bristol Robotics Laboratory, University of the West of England, Bristol, UK

³ Department of Computer Science, University of Warwick, Coventry, UK

⁴ Jaguar Land Rover, Coventry, UK

1 Introduction

Autonomous systems are becoming increasingly ubiquitous and are making significant impacts in various safety-critical systems, such as driverless cars [2, 9, 18], exploration robots [29], and unmanned aerial and ground vehicles [1, 43]. Such systems operate in highly dynamic and heterogeneous environments, resulting in complex agent behaviour and interactions that feature learning, adaptation and feedback loops. Norms, or conventions, are social rules or standards of behaviour that are agreed or expected by a set of individuals [11]. Norms may impose an obligation to act or to not act in a particular way, where violation incurs the risk of punishment. Alternatively, norms can be considered solely as expected forms of behaviour without punishment for violation [17]. In this paper, we consider the latter case where violation is not explicitly punished. The norms life cycle has three main stages: norm creation, propagation and emergence [34], where emergence implies that part of the population has adopted the norm. Kittock [24] examined the emergence of norms through co-learning, and considered a norm to have emerged if at least 90% of agents have converged to choose the same action. Such learning is typically in the form of reinforcement learning [26] or imitation learning [40].

There is significant interest within the AI community on norm emergence in multi-agent systems. For example, Sen and Airiau consider agents learning rules of the road [3, 35, 36, 48], using a coordination game (which side of the road to drive) and a social dilemma game (who yields if two drivers arrive at an intersection simultaneously from neighbouring streets). However, to our knowledge, existing research has focussed on agent interactions in which the combination of actions leads to a binary distinction between those that are either compatible or incompatible, and so the reward space is relatively small (e.g., the Coordination Game or Prisoner's Dilemma). To the best of our knowledge, little work exists on the emergence of norms where agent interactions are complex and nuanced, in the sense that there may be a range of compatible actions of varying desirability, which is the focus of this paper.

The contributions of this paper are as follows:

1. We propose a novel agent-based model of norm emergence in which agent interactions are more nuanced than in previous work. For this, while there is an ideal set of compatible actions which leads to an optimal norm there are also combinations that are compatible and have positive rewards (but lower than the optimal). Also, we introduce a complex reward function having clauses which correspond to the different cases the agents can encounter. Furthermore, our model differentiates between the notions of the *role* and *state* of an agent.
2. We identify conditions under which globally compatible conventions are likely to emerge in an environment containing nuanced interactions. Towards this end, we perform simulations to measure the structure of the topology that is required to provide equivalent convergence between completely non-mixed communities, or partitions, compared to mixed partitions in different proportions.
3. We consider a state (i.e., a configuration of the environment) as comprising both a role and features of the environment (e.g., distance between two vehicles). This allows us to consider how the appropriateness of an action depends on the role played by an agent in an encounter and the environmental circumstances of the agents.
4. We investigate the frequency of convergence to a norm, conformity and diversity [22] from two different viewpoints: network community (partitions) and agent type. We

measure the emergence of a norm by keeping track of the frequency of convergence in a given state, and base these measures on the states of the interacting agents.

The remainder of this paper is organized as follows. Section 2 discusses related work and we formally describe our model, using the language of reinforcement learning, in Sect. 3. In Sect. 4, we describe the case study scenario that we use to validate our approach. Section 5 describes the experimental setup used to validate our model, and provides discussion and analysis of the results. Finally, Sect. 6 concludes the paper.

2 Related work

Conventions are social rules or standards of behaviour which are expected by a set of individuals [11], and have no deontic aspect [17]. In contrast, norms impose an obligation to act in a particular way where violation incurs the risk of punishment. Conforming to norms and conventions has several benefits including encouraging cooperative behaviour and coordination, reducing social friction, and modelling the aggregate interactions of agents [11, 48]. In existing literature, the terms *norm* and *convention* have often been used interchangeably in the context of expected agent behaviour in the absence of obligations, punishments and actions [6, 11, 17, 19, 21, 34, 48] while the term *norm* is exclusively used where such mechanisms are present. In line with this literature, we also refer to such emergent behaviour as norms, even though there are no explicit obligations, punishments or sanctions in our context.

In multi-agent systems, there are two approaches to establishing norms: the prescriptive and emergent approaches [20]. In the prescriptive approach, a central authority specifies how agents should behave, while in the emergent approach conventions emerge as a result of interactions between agents. The norms life cycle has three main stages: norm creation, propagation and emergence [34]. The introduction of a norm into the system is norm creation. The distribution of norms to agents within the system is norm propagation, and norm emergence is where a part of the population is observed to be adopting the norms [34]. Norms have also been considered from an automated norm synthesis perspective which, while related to norm emergence, focuses on identifying and formalising norms within a system, in addition to proposing new norms to avoid conflicts [30–33]. IRON [30] and SIMON [31] identify undesirable states in a system at runtime (equivalent to incompatible actions in our context), and propose norms to avoid these in the future using mechanisms such as prohibition and obligation. LION [32] builds on this approach but also aims to maximise agents' freedom by minimising the number of constraints imposed. In this paper, we focus on emergent norms in the absence of notions of obligation, prohibition or sanctions etc., however, our work provides insights which could be applied at the norm emergence and detection stages of norm synthesis.

Shoham and Tennenholtz [49] discussed results and observations about the efficiency of norm evolution, and Kittock [24] examined the emergence of conventions through co-learning in a multi-agent system. According to Kittock, a norm is considered to have emerged in a system, if at least 90% of agents have converged to choose the same action. Later, in [50], Shoham and Tennenholtz defined the notion of social conventions in a standard game-theoretic framework.

Several existing approaches have explored how the topology of the connections between agents affects the emergence of norms (where such connections determine which agents

are able to interact). By using regular, small-world and scale-free networks, Kittock [24] and Delgado [8] found that agents in networks with larger diameters take longer to achieve global convention emergence. Pujol et al. [42] showed how a system converges to a pareto-efficient convention when communities are highly-clustered. This was demonstrated using different topologies including random, regular, small-world and scale-free graphs. Other research has also considered the role of topology in norm emergence [17, 19, 21, 22, 27, 46], but has not explicitly considered the role of community structure, with the notable exception of [21, 22] (discussed below). Moreover, such investigation of topology has focused on cases where there is a binary distinction between (optimal) compatible norms and incompatible norms, rather than our setting in which there are optimal, compatible and incompatible norms. Meanwhile, Epstein [10] and Villatoro et al. [54] identified the notion of sub-conventions in regular and scale-free networks where systems usually take a long time to change sub-conventions into a global convention [22]. Later, in order to dissolve the sub-conventions and facilitate the emergence of global conventions, Griffiths and Luck [13], and Villatoro et al. [53] proposed approaches that allow agents to change the network topology by rewiring their connections. Franks et al. [11] proposed a general methodology for learning the network value of an agent with respect to influence, which indicates its likely impact on convention emergence.

Existing literature has focused on two games in particular, namely the Coordination Game and the Cooperation Game (also known as Prisoner's Dilemma [50]). The Coordination Game describes a situation where the goal is to reach homogeneity in society [25]. Later, Sen and Airiau [3, 35, 36, 48] applied the Coordination Game in their social learning framework which uses norms to resolve social dilemmas. They proposed a social learning framework that supports the emergence of social norms via learning from interaction experiences. At each time step, each agent in a population interacts with another that is randomly selected, and the payoff received by the agents depends only on this interaction. The role of network topology in social learning has also been considered [47, 52–54]. For example, Sen and Sen [47] evaluated how different network topologies affect the emergence of norms through interaction-based social learning, and how the number of action choices impacts the rate of norm emergence. Villatoro et al. [52, 53] also investigated the impact of different topologies in reaching social conventions, defining *social instruments* to speed-up and support full convergence in scale-free networks.

In the majority of the literature, there is a binary distinction in terms of whether agents' actions are compatible or incompatible [3, 35, 36, 48]. For example, in the n -action Coordination Game, a positive reward is only achieved if both agents choose the same action, which is rather limited. Furthermore, in many investigations, such as that of Sen and Airiau [3, 48] the "states" and the "roles" are indistinguishable. Therefore, this does not allow one to consider how the appropriateness of an action depends on other factors like the role played by an agent or the environmental circumstances. In previous work, the possible sets of compatible actions or norms are typically associated with equal rewards, representing the optimal payoff that can be achieved [17, 34]. We introduce the notion of compatible, but non-optimal, actions which have a positive payoff less than that achieved in the optimal case. For clarity, in this paper we use the terms *optimal*, *compatible*, and *incompatible* to refer to the cases where the combination of agents' actions are compatible and give the highest (or optimal) reward, are compatible but with reward less than the optimal, or are incompatible giving zero or negative reward respectively. There has been limited consideration of such (non-optimal) compatible actions from a norms perspective, with a few notable exceptions. The case of multiple compatible actions has been considered in stochastic coordination games, having multiple and differently valued Nash equilibria [15, 16, 56].

Suboptimal norms have also been considered in the context of language learning in which the compatibility of words and concepts in a lexicon enables norms to be “ranked” in terms of effectiveness [45]. However, this previous work does not consider the different types and roles of agents, or the impact of the community structure of the underlying topology.

Meanwhile, Hu and Leung [5, 20–22] investigated local convention emergence, in contrast with most prior work which only considered a global norm [8, 24, 28, 36, 42, 46, 48, 51, 54, 57]. Hu and Leung investigate the effects of community structure on convention emergence where nodes are partitioned into communities. A local convention is defined as a restriction on behaviour that is imposed on a community of agents. The authors proposed two quantitative measures: local convention conformity and diversity in agents’ actions. The former measures the overall strength of local conventions in a system, while the latter measures the diversity of the agents’ actions throughout the system. However, their work differs to ours since, (i) their metrics do not consider interacting agents having different roles and potentially different states and, (ii) they assume a binary distinction between (optimally) compatible and incompatible actions. Furthermore, these metrics have been considered only from the whole network or underlying network community perspective. This is potentially limiting since while agents might converge to a high conformity level from a network community perspective, there may be a much lower level from other perspectives such as agents’ type.

Overall, while there is extensive literature on norm emergence in multi-agent systems, the focus has been on interactions with a binary distinction between (optimal) compatible and incompatible actions [3, 35, 36, 48]. To the best of our knowledge, little work has considered reinforcement learning for norm emergence in the context of more nuanced interactions, which include compatible but sub-optimal action combinations, and this provides motivation this paper.

3 The formal model

In this paper, we aim to investigate norm emergence in the context of agent interactions that are more nuanced than those which have typically been studied. Towards this goal, our model: (i) differentiates the notions of *role* and *state*; (ii) categorizes norms broadly into two types: *optimal norms* and *compatible norms*; and (iii) introduces a complex reward function with clauses that correspond to the different cases agents can encounter. Previous research has typically made a binary distinction into the cases of actions either being compatible (and optimal) or incompatible, the former potentially leading to a norm which, due to this distinction, will be an *optimal norm*. We introduce the notion of compatible, but non-optimal, actions which have a positive payoff less than that achieved in the optimal case. Such actions may also lead to the emergence of norms, and we refer to such norms as *compatible norms*.

We assume a population of *agents*, denoted by the set $\mathbf{A} = \{a_1, a_2, \dots, a_{|\mathbf{A}|}\}$, which, for example, in our case study represents the set of autonomous and human-driven vehicles. Agents can be of a particular *type* (or profile), denoted by the set $\mathbf{T} = \{t_1, t_2, \dots, t_{|\mathbf{T}|}\}$. For example, in our case study, agents can be of *assertive* or *non-assertive* type depending on whether they typically interact in a city or rural location. We employ this notion later in our modelling to show the measures of convergence, conformity and diversity from two different viewpoints.

In our model, we differentiate the notions of *role* and *state*. Agent interactions are often asymmetric, where each agent plays a different *role* in an interaction [11], for example, determining when a driver should yield at a junction [48]. Other real-world examples of asymmetric interactions can be found in auctions and task allocations [11]. Each agent can be associated with a set of roles. The set of roles is denoted by $\mathbf{R} = \{r_1, r_2, \dots, r_{|\mathbf{R}|}\}$, which in our case study corresponds to the roles of the vehicle turning left at the intersection or the vehicle travelling on the main road. An agent can be in a particular *state* which describes both the *role* played by the agent and any *environmental factors* (e.g., in our case study, these include the distance between the vehicles and their velocities). In our model, the set of states is denoted $\mathbf{S} = \{s_1, s_2, \dots, s_{|\mathbf{S}|}\}$, such that $s_i = \langle r_k, \hat{f}_l \rangle$ where r_k and \hat{f}_l correspond to the role and any environmental factors associated with the state s_i respectively.

At any time, an agent can choose an action from the set $\mathbf{P} = \{p_1, p_2, \dots, p_{|\mathbf{P}|}\}$ of *actions* or *strategies*. These combinations of these actions chosen by agents represent the possible norms that might emerge. For example, in our case study, the possible actions of a vehicle are *go* or *yield*.

In our model, we categorize norms broadly into two types. First, *optimal norms* correspond to those giving maximal possible rewards to the agents, and are the type of norm considered in previous work [17, 34]. Second, *compatible norms* return positive rewards, but not as high as those provided by optimal norms. This notion of compatible norms has received little attention in the literature, with some notable exceptions that do not consider types and roles of agents, or the impact of the community structure of the underlying topology [15, 16, 45, 56].

Each agent (a_i) in the population is paired with a random neighbour (a_j) for interaction. Both agents observe their state (s_i or s_j) and choose an action (p_i or p_j) for which a reward (g_i or g_j) is received. This reward reflects how good or bad the chosen action was for the corresponding agent. Rewards can be represented using H , M , and L to denote *high*, *medium* and *low* rewards respectively. In this context, an *optimal norm* is a combination of actions which yield the maximum positive reward denoted by H (i.e., $g(s, p) = H$, where $g(s, p)$ denotes the reward received from action p in state s). A *compatible norm* is an action which yields a positive reward but not as high as an optimal norm (i.e., $H > g(s, p) > 0$). Thus, we formalize the reward function for the two agents (a_i, a_j) using a function $g(s, p)$, which contains clauses that correspond to the different situations that might be encountered (refer to the following section for case study examples).

$$g(s, p) = \begin{cases} \text{case 1 : } g_i \text{ for } a_i, g_j \text{ for } a_j \\ \dots \\ \text{case } n : g_i \text{ for } a_i, g_j \text{ for } a_j \end{cases}$$

4 The case study

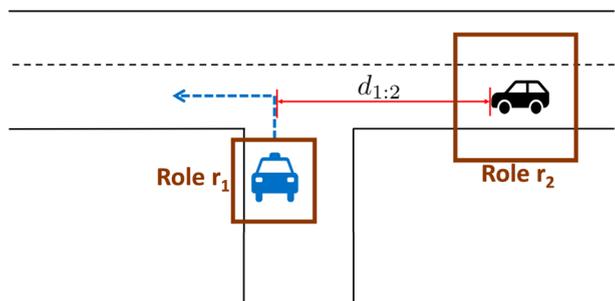
In this section, we describe the case study scenario that we use to illustrate our model. We begin by giving a general introduction to the scenario, and then discuss the possible cases that define the reward function.

4.1 Scenario description

Our case study scenario is that of an autonomous vehicle (AV) performing a left-turn (in the UK, where cars drive on the left) at a T-intersection, as illustrated in Fig. 1. Let us assume that the AV monitors the distance between itself and any vehicle approaching on the main road to be joined, and the speed of any vehicle on the main road. Using the model described in Sect. 3, we can describe this scenario as a 2-player social dilemma game by defining the sets of agents A , agent types T , roles R , a state space S , an action space P , and reward function $g()$.

- Agents: there are two agents in an interaction, a_1 representing the AV and a_2 representing the human-driven vehicle, so $A = \{a_1, a_2\}$.
- Types: agents can be of type t_1 or t_2 representing the assertive or non-assertive type respectively, $T = \{t_1, t_2\}$.
- Roles: The role of a vehicle turning left is represented by r_1 , while r_2 represents the vehicle travelling on the main road, so $R = \{r_1, r_2\}$. We assume that the agent playing role r_1 is the AV (a_1), and the agent playing role r_2 is a human-driven vehicle (a_2).
- States: The environmental factors of the state are represented by the *distances* and *velocities* of the two vehicles at the intersection and their roles, so $S = \{s_1, s_2\}$ where $s_1 = (r_1, d_{1:2}, v_1, v_2)$ and $s_2 = (r_2, d_{1:2}, v_2, v_1)$. Here, $d_{1:2}$ denotes the distance between the agent in role r_1 and the agent in role r_2 . The initial velocity (i.e., at the time of interaction or making decision) of the agent in role r_1 is denoted by v_1 , and v_2 is the initial velocity of the agent in role r_2 . The initial velocity (v_1) of the agent in role r_1 is zero. After actions are chosen by the agents, the post-action velocity of the agent in role r_1 (v'_1) is drawn from a distribution and not known at the time of action selection. Also, the agent in role r_2 may change its velocity (v'_2) after action selection.
- Actions: there are two actions, namely p_1 and p_2 representing go or yield respectively, and so $P = \{p_1, p_2\}$.
- The threshold for the desirable distance between vehicles for agent a_i is denoted λ_{Da_i} and γ_{Ca} denotes the critical distance threshold (which is shared by all agents). The desirable and critical thresholds are used to characterise the separation between the two vehicles where the former corresponds to no risk of collision, and the latter with a high risk of collision.
- The loss of velocity (or delay) of agent a_i (i.e., the agent in role r_2 in our scenario) due to the action of another agent (e.g., when the agent in role r_1 chooses to pull out) is denoted ΔV_i .

Fig. 1 Case study scenario as a 2-player social dilemma game



- We assume there is a small, positive constant, β , which can be used to reduce an agent's reward. For example, it can be used to provide a proportion (i.e., $-\beta \cdot \Delta V_i$) of the negative reward of the agent in role r_2 to the agent in role r_1 , as a result of any frustration caused due to the loss of velocity or delay.

4.2 Possible cases for determining rewards

We now describe the possible situations that can occur in our case study scenario, and how the rewards are calculated. Note that the situations are named according to the actions chosen by the agents (e.g., YY to indicate yield-yield, GG to denote go-go, etc.) and whether it describes an optimal or compatible norm (ON and CN respectively). The possible situations, or cases, that must be considered when defining rewards are as follows.

1. **Case YY:** If both the agents in role r_1 and role r_2 choose to *yield*, then both receive a *small negative* reward (-L). This case is specified in clause 1 of Equation 1, defined below.

If one agent chooses to *go* (i.e., the agent in role r_1 or r_2) and the other agent chooses to *yield*, then this can have several outcomes because the impact of yielding is different in difference situations. We characterise these situations below in cases YG, GY1, GY2 and GY3.

2. **Case YG:** If the agent in role r_1 yields (stops) and the agent in role r_2 goes, then the agent in role r_1 receives a *small* reward (+L) and the agent in role r_2 receives a *high* reward (+H). This case is defined in clause 2 of Equation 1 below.
3. If the agent in role r_1 goes and the agent in role r_2 yields, this can have several outcomes depending on the pull out distance and whether the agent in role r_2 *slows down* or *stops*. These are presented in cases GY1, GY2 and GY3 below. Note that if the agent in role r_1 goes or pulls out, we assume that it obtains an instantaneous steady velocity (v'_1).
 - (a) **Case GY1:** If the agent in role r_1 chooses to go and the agent in role r_2 chooses to yield, and if $d_{1:2}$ is maintained within the desirable threshold for a_1 ($d_{1:2} \geq \lambda_{Dd_1}$), then the agent in role r_1 receives a *high* reward (+H) and the reward for the agent in role r_2 will be a *medium* reward (+M). In case GY1, we do not consider whether *yielding* corresponds to slowing or to stopping, since $d_{1:2} \geq \lambda_{Dd_1}$. This case is represented in clause 3 of Equation 1 below.
 - (b) **Case GY2:** Let us consider another situation where the agent in role r_1 chooses to go and the agent in role r_2 chooses to yield. Here, the agent in role r_2 will *slow down*, if $d_{1:2}$ is between its desirable and critical thresholds ($\gamma_{Cd} < d_{1:2} < \lambda_{Dd_2}$). As a result, the agent in role r_1 receives a *high* reward (+H). On the other hand, the reward for the agent in role r_2 will be determined by the delay (ΔV_i) that the agent in role r_2 incurs. This is because it has to wait in its current position so that λ_{Dd_1} can be reached. Note that a level of frustration can be felt by the agent in role r_2 , if it has to slow down due to the loss of velocity or delay of the agent in role r_1 . In order to model the reward for the agent in role r_1 causing this situation, it is given a proportion of the negative reward of the agent in role r_2 (i.e., $-\beta \cdot \Delta V_i$). As a result, the agent in role r_1 's overall reward will be much lower compared to

a situation where there is no frustration caused to the agent in role r_2 . Case GY2 is represented in clause 4 of Equation 1 below.

Calculating delay (ΔV_i): The reward for the agent in role r_2 will be defined by the period of time (ΔV_i) before the desirable distance threshold can be achieved again. By using the pull out distance ($d_{1:2}$) and λ_{Dd_i} , we can calculate how much time (i.e., ΔV_i) it would take for the distance between the vehicles to regain λ_{Dd_i} . For example, let us consider that: $v'_1 = 50kph$, $v_2 = 20kph$, $d_{1:2} = 75m$, and $\lambda_{D-d_i} = 100m$. Thus, the required time is (in seconds): $((100-75/1000)/(50-20))*60$, which is the reward provided for the agent in role r_2 .

- (c) **Case GY3:** If the agent in role r_1 chooses to go and the agent in role r_2 chooses to yield, the agent in role r_2 will *stop* if $d_{1:2}$ is less than or equal to the critical threshold ($d_{1:2} \leq \gamma_{C-d}$). In this case, the agent in role r_1 receives a *high* reward (+H). Meanwhile, the agent in role r_2 will receive a reward that is determined by the delay (ΔV_i) that it has to incur to wait in its current position, so that λ_{D-d_i} can be reached.

As in the case GY2, frustration can be caused to the agent in role r_2 , if it has to slow down due to the delay of the agent in role r_1 . To model this situation, the agent in role r_1 can be awarded with a proportion of the negative reward of the agent in role r_2 (i.e., $-\beta \cdot \Delta V_i$). Thus, the agent in role r_1 's overall reward will be comparatively much lower than a situation where there is no frustration involved. Case GY3 is shown in clause 5 of Equation 1 below.

Calculating delay (ΔV_i): We calculate the reduction reward for the agent in role r_2 by calculating the delay (i.e., ΔV_i) that the agent in role r_2 has to incur, so that λ_{D-d_i} can be reached. That is, how much time the agent in role r_2 has to wait in its current position, which is the same as how much time the agent in role r_1 has taken to move. For example, let us consider that: $v'_1 = 30kph$, $v_2 = 20kph$, $d_{1:2} = 75m$, and $\lambda_{Dd_i} = 100m$. This results in a delay of $((100-75/1000)/30)*60$ seconds, which is the reward provided for the agent in role r_2 .

4. If both the agents in roles r_1 and r_2 choose to *go*, we consider that the agent in role r_1 obtains an instantaneous steady velocity (v'_1) and the agent in role r_2 's velocity is v_2 . The following two cases describe situations where optimal norms occur in the case study.
- (a) **Case GG-ON1:** If $v'_1 \geq v_2$ and if the agent in role r_1 pulls out maintaining the desirable threshold for $d_{1:2}$ ($d_{1:2} \geq \lambda_{Dd_i}$), then both agents receive a *high* reward (+H). This situation is specified in clause 6 of Equation 1 below.
- (b) **Case GG-ON2:** If $v'_1 < v_2$ and if the agent in role r_1 pulls out maintaining the desirable threshold for $d_{1:2}$ ($d_{1:2} \geq \lambda_{Dd_i}$) then the agent in role r_1 receives a *high* reward (+H) and the agent in role r_2 receives a *medium* reward (+M). This situation is specified in clause 7 of Equation 1 below.
5. If both the agents in roles r_1 and r_2 choose to *go*, and the agent in role r_1 pulls out between the desirable and critical thresholds for $d_{1:2}$ ($\gamma_{Cd} < d_{1:2} < \lambda_{Dd_i}$), then this can have three outcomes depending on the velocities of the agent in role r_1 (v'_1) and the agent in role r_2 (v_2), as outlined below (see Fig. 2). These three cases describe situations where compatible norms emerge in the case study.

We assume that if an agent chooses to *go*, it will *maintain* its velocity unless the critical threshold for distance is violated. We also describe how we calculate the nega-

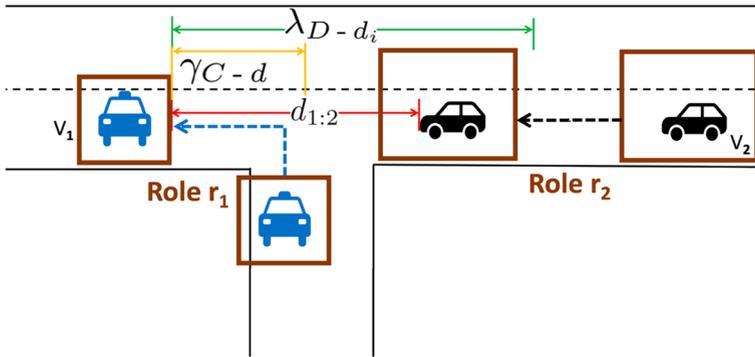


Fig. 2 The agent in role r_1 pulling out between the desirable and critical threshold values for distance.

tive rewards for the agent in role r_2 in cases GG-CN1, GG-CN2 and GG-CN3, as those rewards are determined by: (i) the time (i.e., ΔV_i) needed to reach the desirable threshold value for distance (λ_{Dd_i}) (case GG-CN1); or (ii) the delay (i.e., ΔV_i) which needs to be incurred to achieve λ_{Dd_i} (cases GG-CN2 and GG-CN3). In this regard, our intention is not to compute the actual trajectories of the two vehicles, such as their final velocities and acceleration. We assume that changes in agents' behaviour occur instantaneously. The time or delay to reach λ_{Dd_i} corresponds to the negative reward provided for the agent in role r_2 .

- (a) **Case GG-CN1:** If $v'_1 > v_2$ then it can be considered that at some point in time, the desirable threshold for $d_{1:2}$ (λ_{Dd_i}) can be achieved. In this situation, the reward for the agent in role r_2 will be defined by the period of time (ΔV_i) the desirable distance threshold can be achieved again, assuming agents maintain their current speeds. The reward given for the agent in role r_1 is *high* (+H). This situation is specified in clause 8 of Equation 1 below.

Calculating delay (ΔV_i): In this situation, the agent in role r_1 's pull out velocity (v'_1) is greater than the agent in role r_2 's velocity (v_2). If we assume that the vehicles maintain their current speeds, by using the pull out distance ($d_{1:2}$) and λ_{Dd_i} , we can then calculate how much time (i.e., ΔV_i) it would take for the distance between the vehicles to regain λ_{Dd_i} . For example, let us consider that: $v'_1 = 50kph$, $v_2 = 30kph$, $d_{1:2} = 75m$, and $\lambda_{Dd_i} = 100m$. Thus, the required time is (in seconds): $((100-75/1000)/(50-30))*60$, and this period of time defines the reward provided for the agent in role r_2 .

- (b) **Case GG-CN2:** Let us consider that the agent in role r_1 pulls out with a lower velocity than the agent in role r_2 's velocity (i.e., $v'_1 < v_2$). This means that regardless of the initial distance at some point the distance can be less than the desirable threshold of $d_{1:2}$, and later, it can be less than the critical threshold of $d_{1:2}$ which requires the agent in the role r_2 to stop. In this situation, the reward for the agent in role r_2 is calculated based on the delay (ΔV_i), it has to incur if it is to achieve the desirable threshold of $d_{1:2}$. More specifically, how much time the agent in role r_2 has to wait in its current position, which is essentially same as the time the agent in role r_1 has taken to move.

The reward given for the agent in role r_1 is *high* (+H). But, if the frustration level of the agent in role r_2 is also considered because it has to slow down due to the actions of the agent in role r_1 , then the agent in role r_1 will also be assigned with a proportion of the negative reward of the agent in role r_2 (i.e., $-\beta \cdot \Delta V_i$). This case is defined in clause 9 of Equation 1 below.

Calculating delay (ΔV_i): In this case, the agent in role r_1 pulls out with a lower velocity than the agent in role r_2 (i.e., $v'_1 < v_2$). The reward for the agent in role r_2 is calculated based on the delay (i.e., ΔV_i) the agent in role r_2 has to incur, if it is to achieve λ_{Dd_i} . That is, how much time the agent in role r_2 has to wait in its current position, which is the same as the time taken by the agent in role r_1 to travel. For example, let us consider that: $v'_1 = 20kph$, $v_2 = 30kph$, $d_{1:2} = 75m$, and $\lambda_{Dd_i} = 100m$. Then the required delay period is: $((100-75/1000)/20)*60$ seconds. This delay defines the reward provided for the agent in role r_2 .

- (c) **Case GG-CN3:** If $v'_1 = v_2$ then this means that the undesirable distance is maintained and the desirable threshold for $d_{1:2}$ cannot be achieved. In this case, the reward for the agent in role r_2 will be determined by the delay (ΔV_i) that the agent in role r_2 has to incur to wait, so λ_{Dd_i} can be reached. On the other hand, the reward provided for the agent in role r_1 is *high* (+H). This case is defined in clause 10 of Equation 1 below.

Calculating delay (ΔV_i): In this situation, the velocities of the two agents are same, which means that the undesirable distance is maintained. We calculate the negative reward for the agent in role r_2 by calculating the delay (i.e., ΔV_i) that the agent in role r_2 has to incur so λ_{Dd_i} can be reached.

That is, how much time the agent in role r_2 has to wait in its current position, and this time is same as how much time the agent in role r_1 has taken to move. For example, let us consider that: $v'_1 = 30kph$, $v_2 = 30kph$, $d_{1:2} = 75m$, and $\lambda_{Dd_i} = 100m$. This results with a delay of $((100-75/1000)/30)*60$ seconds, which defines the reward provided for the agent in role r_2 .

6. **Case GG-C:** If both agents, that is the agent in role r_1 and the agent in role r_2 , choose to *go*, and the agent in role r_1 pulls out less than or equals to the critical threshold for $d_{1:2}$ (i.e., $d_{1:2} \leq \gamma_{Cd}$), then both agents receive a *high negative* reward (-H). This case is defined in clause 11 of Eq. 1 below.

Finally, the reward function for the case study scenario is defined using Equation 1. The clauses of this equation correspond to the different situations the agents can encounter with different rewards. Clauses 6–7 correspond to situations where *optimal norms* can emerge, while clauses 8–10 correspond to cases in which *compatible norms* can emerge.

$$g(s, p) = \left\{ \begin{array}{ll} -L, -L & \text{if } (s_1 = r_1 \ \& \ A = \text{yield}) \ \& \ (s_2 = r_2 \ \& \ A = \text{yield}) \\ +L, +H & \text{if } (s_1 = r_1 \ \& \ A = \text{yield}) \ \& \ (s_2 = r_2 \ \& \ A = \text{go}) \\ +H, +M & \text{if } (s_1 = (r_1, d_{1:2} \geq \lambda_{D-d_i}) \ \& \ A = \text{go}) \ \& \\ & (s_2 = r_2 \ \& \ A = \text{yield}) \\ (+H - \beta \cdot \Delta V_i), \text{delay}(\Delta V_i) & \text{if } (s_1 = (r_1, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}) \ \& \ A = \text{go}) \ \& \\ & (s_2 = r_2 \ \& \ A = \text{yield}) \\ (+H - \beta \cdot \Delta V_i), \text{delay}(\Delta V_i) & \text{if } (s_1 = (r_1, d_{1:2} \leq \gamma_{C-d}) \ \& \ A = \text{go}) \ \& \\ & (s_2 = r_2 \ \& \ A = \text{yield}) \\ +H, +H & \text{if } (s_1 = (r_1, d_{1:2} \geq \lambda_{D-d_i}, v'_1 \geq v_2) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, d_{1:2} \geq \lambda_{D-d_i}, v'_1 \geq v_2) \ \& \ A = \text{go}) \\ +H, +M & \text{if } (s_1 = (r_1, d_{1:2} \geq \lambda_{D-d_i}, v'_1 < v_2) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, d_{1:2} \geq \lambda_{D-d_i}, v'_1 < v_2) \ \& \ A = \text{go}) \\ +H, \text{delay}(\Delta V_i) & \text{if } (s_1 = (r_1, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 > v_2) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 > v_2) \ \& \ A = \text{go}) \\ (+H - \beta \cdot \Delta V_i), \text{delay}(\Delta V_i) & \text{if } (s_1 = (r_1, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 < v_2) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 < v_2) \ \& \ A = \text{go}) \\ +H, \text{delay}(\Delta V_i) & \text{if } (s_1 = (r_1, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 = v_2) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, \gamma_{C-d} < d_{1:2} < \lambda_{D-d_i}, v'_1 = v_2) \ \& \ A = \text{go}) \\ -H, -H & \text{if } (s_1 = (r_1, d_{1:2} \leq \gamma_{C-d}) \ \& \ A = \text{go}) \ \& \\ & (s_2 = (r_2, d_{1:2} \leq \gamma_{C-d}) \ \& \ A = \text{go}) \end{array} \right. \quad (1)$$

5 Experimental results and discussion

In this section, we discuss the experimental setup and results of our study, showing the influence of different properties of the underlying network topology on the emergence of norms. First, we provide an overview of the agent-based modelling performed in the case study. Then, we discuss each of the four main experiments describing their hypothesis, experimental set up, and their individual results. Finally, we provide a discussion and analysis of the main results.

5.1 Overview of the agent-based model

In this subsection, we describe the modelling of the case study, with the main steps of the simulation being summarised in the algorithm shown in Fig. 3. We model the interaction between the agents as a single stage pairwise interaction. We model a population of m agents (i.e., 2000 or 5000) who are allocated into n partitions (i.e., 2 or 5) with 1000 agents in each partition. In all the experiments, we limited the execution of the simulations to 100 runs where each run contains 1000 time steps or iterations. In each interaction, pairing is performed between each agent and another randomly selected neighbour from the population. Then, learning is performed by both agents concurrently over repeated interactions. The agents learn following the Q-learning algorithm [55] described in [41] with ϵ -greedy exploration strategy (learning rate: 0.3 and exploration rate: 0.1). In Q-learning, the agent takes the maximum-valued action from the next state [14].

- 1: Assume a population of m agents. Graph G contains n partitions with 1000 agents each. In the non-mixed case, for each partition, an agent type is picked randomly and allocated. In the mixed agents case, all partitions are allocated using different proportions of assertive and non-assertive agents (i.e., 20-80%, 80-20%, 40-60%, 60-40%, and 50-50%).
- 2: **for** each iteration **do**
- 3: **for** each agent a_i in graph G **do**
- 4: Pair agent a_i with a random neighbour agent a_j
- 5: Assign agent a_i and a_j with role r_1 or r_2 randomly
- 6: Sample velocity (v_2) of the agent in role r_2 from distribution based on its agent type. The agent in role r_1 has zero initial velocity ($v_1 = 0$)
- 7: Draw distance $d_{i:j}$ from a distribution between two intervals. Each agent has its own λ_{D-d_i} based on its agent type, with γ_{C-d} being the same for all agents
- 8: Determine states of the agents using $d_{i:j}$, and λ_{D-d_i} and γ_{C-d}
- 9: Both agents a_i and a_j choose an action (go or yield)
- 10: If the agent in role r_1 chooses to go, draw its pull out velocity (v'_1) from a distribution based on the agent type
- 11: Both agents observe the rewards and learn
- 12: **end for**
- 13: **end for**

Fig. 3 Simulation main steps

In order to demonstrate the influence of different characteristics of the underlying network topology on the emergence of conventions, we conducted simulations to investigate four main questions:

1. What effect does the number of inter-partition edges have on the emergence of norms?
2. What effect does the number of intra-partition edges have on norms emergence?
3. What effect does the interoperable (overlapping) region of the distributions used to generate velocities (initial velocity of the agent in role r_2 and pull out velocity of the agent in role r_1) and distances have on norm emergence?
4. What effect does network community and agent type have on the overall strength of local conventions in a system (conformity) and diversity throughout the system?

In the experiments, we used different types of communities based on how mixed the agents are, as described below.

- Non-mixed case: we create a non-mixed community of agents where each partition has agents of one agent type only (i.e., assertive or non-assertive). Here, for each partition, an agent type is picked randomly and all agents are allocated to be that type.
- Mixed cases: we created a mixed community with different proportions of assertive and non-assertive agents (i.e., 20-80%, 80-20%, 40-60%, 60-40%, and 50-50%). In the mixed cases, both agents of assertive and non-assertive type are allocated in a random manner, so that agents are well mixed in the two partitions.

As mentioned in Sect. 3, the agents in the population are of a particular *type* (or profile) which is determined by the environment in which the agent typically interacts, namely *city* or *rural*. Because a city is associated with a more dynamic and crowded environment, we consider agents which originate from a city to be more *assertive* than those from a rural location which can be less assertive (or *non-assertive* for simplicity hereafter). Thus, we identify two types of agents that meet and interact: assertive agents and non-assertive agents.

In an interaction, the two paired agents are randomly assigned with the role r_1 (vehicle turning left) or r_2 (vehicle travelling on the main road). The agent in role r_1 has an initial velocity (v_1) of zero since it is stationary, and waiting to pull out to the main road from the side road. We draw the initial velocity of the agent in role r_2 (v_2) from a distribution with different lower and upper limits for assertive and non-assertive agent types, thus supporting agent heterogeneity. In the simulation, we differentiate the assertive agent type from the non-assertive type mainly in two ways. Refer to Sect. 5.4 in regards to the details of the interoperable regions defined in the case study.

1. First, since a city has crowded roads, the velocity intervals of the assertive agents are lower compared to the non-assertive agents. Therefore, the lower and upper limits used in the distribution to calculate the initial velocity of the agent in role r_2 (v_2) for an assertive agent are relatively low compared to a non-assertive agent (i.e., 10–60 kph compared to 30–80 kph). Also, if the agent in role r_1 chooses to pull out, the lower and upper velocity limits that are used to generate the initial pull out velocity (v'_1) for an assertive agent are lower compared to a non-assertive agent (i.e., 10–60 kph compared to 30–80 kph). In both above velocity cases, the interoperable region is 30–60 kph for the two agent types.
2. Second, each agent in the population has its own desirable threshold for distance (λ_{Dd_i}) which is sampled from a distribution based on the type of the agent. In the case study scenario, as we consider this threshold for the vehicle in front, it is the agent in role r_2 whose threshold is relevant. So, this threshold is applied to both agents involved in the interaction. Meanwhile, we assume that the critical threshold for distance (γ_{Cd}) is the same for all the agents (50m). In the case study, λ_{Dd_i} can be sampled between 60–90m for an assertive agent, and 70–100m for a non-assertive agent. Therefore, the interoperable region is 70–90m between the two agent types.

The distance between agents in roles r_1 and r_2 (i.e., $d_{i;j}$) is drawn from a distribution between two intervals (*lower*, *upper*) where *lower* is less than γ_{Cd} and *upper* is greater than λ_{Dd_i} .

We need to consider *state* because there can be a norm which can emerge very strongly in some states, but not as strong in other states. Therefore, similarly to Hu and Leung [22] who separate a population by the communities the individual agents are in, we additionally separate the action choices by their states.

5.1.1 States used in Q-learning

In general, Q-learning is applied to problems with discrete states and actions, since the state space can be too large when the variables are continuous. Therefore, to overcome this problem, the states used in Q-learning of our study are defined using a combination of *role* and whether the *distance* between the two vehicles is within one of the three constraints

related to the desirable and critical thresholds (see below). Instead of discretising distance directly, we discretise the state space by checking whether the distance is within one of the three constraints. Note that although our formal model (Sect. 3) for the state also considers velocity of the vehicles, we do not model velocity as part of the state in the Q-learning due to the large state space. So, conceptually, an agent (with respect to Q-learning) can be in six states: three states when the agent is in role r_1 (i.e., autonomous vehicle turning left), and three when it is in role r_2 (i.e., vehicle moving on main road). Thus, the agent in role r_1 can be in any of the following three states and can perform two actions (yield, go).

1. The agent in role r_1 's state is 0 when $d_{1:2} \geq \lambda_{Dd_i}$.
2. The agent in role r_1 's state is 1 when $\gamma_{Cd} < d_{1:2} < \lambda_{Dd_i}$.
3. The agent in role r_1 's state is 2 when $d_{1:2} \leq \gamma_{Cd}$.

The three possible states for the agent in role r_2 are defined analogously.

Actions are only chosen randomly when an agent is exploring. Otherwise, the actions are selected according to the learning method used by the agents, i.e., Q-learning. We use ϵ -greedy exploration strategy to trade off exploration and exploitation. In this strategy, an exploration rate is used to select the greedy action all but ϵ of the time and to select a random action ϵ of the time [41]. Each agent has its own Q-table which is indexed by states and actions, and it is initialized to zeros. A Q-table has three entries of Q-values for the two actions (i.e., yield, go) that can be performed in the three main states. In Q-learning, an agent begins in an initial state and performs an action. Then the agent sees what the highest possible reward is for taking any action from its new state, and updates its value for the state-action pair based on this new highest possible value [14]. As distance and role are part of the state, the agents will learn appropriate actions in the different states.

If the agent in role r_1 chooses to pull out, we assume that it will obtain a steady velocity instantaneously (v'_1), which is drawn randomly between a lower limit and an upper limit based on the agent type. In this manner, the drawn and calculated measures for $d_{1:2}$ and the two threshold values for $d_{1:2}$ will inform what state an agent is in, which along with their actions will determine the reward. Thus, it allows us to calculate the consequences of a particular action choice of the agent in role r_2 .

5.1.2 Optimal and compatible norms

As formally defined earlier in Sect. 3, we categorize *norms* broadly into two types:

- *Optimal norms* which correspond to those giving maximal possible rewards to the agents, similar to agents performing the same actions in a standard coordination game [48].
- *Compatible norms* which are norms that return positive rewards, but not as high as those provided by optimal norms.

The clauses of Equation 1 correspond to the eleven different situations in the case study which provide different rewards. The 6th and 7th clauses (case GG-ON1 and GG-ON2 of case study) correspond to the two situations where *optimal norms* emerge, while clauses 8–10 (case GG-CN1, GG-CN2 and GG-CN3) correspond to situations where *compatible norms* emerge.

As previously described in Sect. 4, the case GG-ON1 describes a situation where both agents in roles r_1 and r_2 go, and the agent in role r_1 pulls out maintaining the desirable threshold for distance. The relationship between pull out velocity of r_1 (v'_1) and initial velocity of r_2 (v_2) is: $v'_1 \geq v_2$. The situation in case GG-ON2 differs from GG-ON1, as $v'_1 < v_2$. In cases GG-CN1, GG-CN2, and GG-CN3 both agents in roles r_1 and r_2 go, and r_1 pulls out between the desirable and critical thresholds for distance. The three cases differ depending on the relationship between pull out velocity of r_1 (i.e., v'_1) and initial velocity of r_2 (i.e., v_2). In case GG-CN1, this is: $v'_1 > v_2$, whereas in case GG-CN2 it is: $v'_1 < v_2$. Meanwhile, in case GG-CN3, this relationship is $v'_1 = v_2$.

In the graphs of the experiment results, the average convergence time and average proportion of runs of the different *optimal* and *compatible norms* are shown in gold and silver respectively.

5.1.3 Metrics used

The results are analysed and discussed using the following three metrics (as used by [3, 22]):

- **Frequency of convergence (convergence rate)** of a norm is measured by dividing the number of times the norm emerges by the number of agents in the partition/system. We measure this for each 1000 iterations which is then averaged over the total number of runs (100).
- **Convergence time** is the number of iterations (time steps) required for the system to converge to a particular norm. We calculate convergence time using the following method. Note that, after analysing the results of all simulations, we have made the assumption that by iteration 800, all the norms in the system have essentially converged. When calculating convergence time, first, we find the *mean* and *standard deviation* of the frequency of convergence of a norm (averaged to 100 runs) after the point it can be certain that the system has converged (i.e., from 800–1000 iterations). Then, for each run, the frequency of convergence dataset for that particular norm is searched to find the time step where it has a frequency within one standard deviation for some window of consecutive iterations (i.e., 5 iterations). This is then averaged to 100 runs.
- **Proportion of runs** the system converges to different norms. This metric shows the proportion of runs each norm emerges after the system has converged. Before we calculate this measure, a test for convergence is performed to find at which iteration the system has converged to the norm which has the highest frequency of convergence. Then, we analyse the results of each run and find how many runs each optimal and compatible norm has emerged during that iteration.

5.1.4 Network topology used in the experiments

All the experiments were conducted using a topology generated using a *random_partition_graph* [37], which is a community-based generator for classes of graphs used in studying social networks. A partition graph is a graph of communities where nodes of same group and different groups are connected with probabilities p_{in} and p_{out} , respectively. In each of the experiments, we describe how we vary these parameters (p_{in} and p_{out}), and the

characteristics of the topology using several metrics (average degree of network, density, intra community and inter community edges).

5.2 Experiment 1: inter-partition edges on norms emergence

5.2.1 Experimental setup

This experiment was conducted, first to compare the effects of the probability of fewer inter-partition edges on convergence in different agent communities which can be non-mixed; or mixed using different proportions of assertive and non-assertive agents (i.e., 20-80%, 80-20%, 40-60%, 60-40% and 50-50%).

Parameters used in the topology: These were conducted using a topology generated using `random_partition_graph` [37] that has a relatively low probability of inter-partition edges. The probability of intra-partition–inter-partition edges ratio (p_{in} , p_{out}) which was used when creating the topology is 0.010, 0.005 respectively.

The **topological metrics** (averaged in 100 runs) are:

- Average degree of network: 15.0093; density: 0.0075; intra community edges: 10006.48; inter community edges: 5002.81.

Then, by taking the mixed agent type case that has an equal proportion of assertive and non-assertive agents, we show that the more the number of inter-partition edges are, the probability of both partitions converging to a global state is higher.

Parameters used: The relative number of inter-partition edges was gradually increased by using the following probability of intra-partition–inter-partition edges ratios (p_{in} , p_{out}) when defining the network topology: (i) 0.010, 0.005; (ii) 0.010, 0.010; (iii) 0.010, 0.015; (iv) 0.010, 0.020; and (v) 0.010, 0.025.

The **topological metrics** of the simulations conducted are provided below.

- Average degree of network: 14.9879; density: 0.0075; intra community edges: 9979.19; inter community edges: 5008.67.
- Average degree of network: 19.9957; density: 0.0100; intra community edges: 9998.88; inter community edges: 9996.78.
- Average degree of network: 24.9770; density: 0.0125; intra community edges: 9964.99; inter community edges: 15012.01.
- Average degree of network: 29.9833; density: 0.0150; intra community edges: 9988.82; inter community edges: 19994.47.
- Average degree of network: 34.9662; density: 0.0175; intra community edges: 9975.8; inter community edges: 24990.38.

Hypothesis

1. Non-mixed agent type case, inter-partition edges, and norms emergence:

In the non-mixed agent type case, in general, we expect a relatively quick emergence of norms within the partition and type, that is: time to convergence is low. However, the probability of the same norm emerging is low in both partitions. That is, the convergence may not be to a compatible norm.

If there are few inter-partition edges and if the partitions are not well mixed, we expect it will take a relatively long time to find the part of action space that is compatible for both agent types across partitions. We expect there will be local norm convergence in both partitions, but they will not be interoperable. Assume there are two completely non-mixed partitions (one partition with assertive agents and the other partition with non-assertive agents) with a sufficient number of intra edges within a partition, but with a very few inter-partition edges. In this situation, it is likely both partitions will achieve local convergence, but they will not be interoperable.

2. **Mixed agents, inter-partition edges, and norms emergence:**

When the agents are more mixed, in general, we expect the probability of the same norm emerging in both partitions to be higher, but learning and convergence can take longer. Here, the probability of getting a compatible norm that ends up in the interoperable region is expected to be higher.

3. **Relative number of inter-partition edges and norms emergence:**

We expect that there is a clear link between the probability of the same norm emerging and the relative number of inter-partition edges. A few inter-partition edges can result in a relatively small overlap area of the actions space. As a result, there will be a low probability of the same norm emerging. The more the number of inter-partition edges are, the higher the probability of both partitions converging to a global state.

We also expect when the number of inter-partition edges increases convergence can take a shorter time.

5.2.2 Results

Now we analyse the effects of low number of inter-partition edges on norms emergence when the communities are non-mixed and mixed. In the non-mixed case, for each partition, an agent type is randomly picked and all agents are allocated of that type only. In the mixed cases, agents are mixed using different proportions of assertive and non-assertive agents (i.e., 20–80%, 80–20%, 40–60%, 60–40%, and 50–50%). Figure 4 shows the proportion of runs partition 1 and 2 converge to different norms when the number of inter-partition edges is low and the partitions contain non-mixed and different proportions of mixed agents. Meanwhile, Fig. 5 shows the convergence times of the different norms in the two partitions.

As shown in Fig. 4, the gap between the compatible norms in cases GG-CN1 and GG-CN2 is narrowest in the non-mixed case, which means the probability of the same norm emerging is low in both partitions (refer to hypothesis 1 and 2). Also, it shows that the more the agents are mixed, the probability of the same norm emerging in both partitions is higher (note the gap widening).

As specified in hypothesis 1, we observed that when the number of inter-partition edges is low and the partitions contain non-mixed agents, both partitions achieved local convergence and converged to different norms. That is, partition 1 converged to the norm in case GG-CN1 with a frequency of convergence of 0.1892, and partition 2 to the norm in case GG-CN2 with 0.1913 (see Table 1). However, in our results, we did not observe a relatively low convergence time for the non-mixed case. The convergence time of the non-mixed case in partition 1 for the norm which has the highest frequency of convergence (GG-CN1) is 59, which is relatively high compared to the mixed cases (see Table 1 and Fig. 5). However, we noted that this variation is still within the relatively large standard deviations of convergence times in 100 runs we observed for this case in both partitions

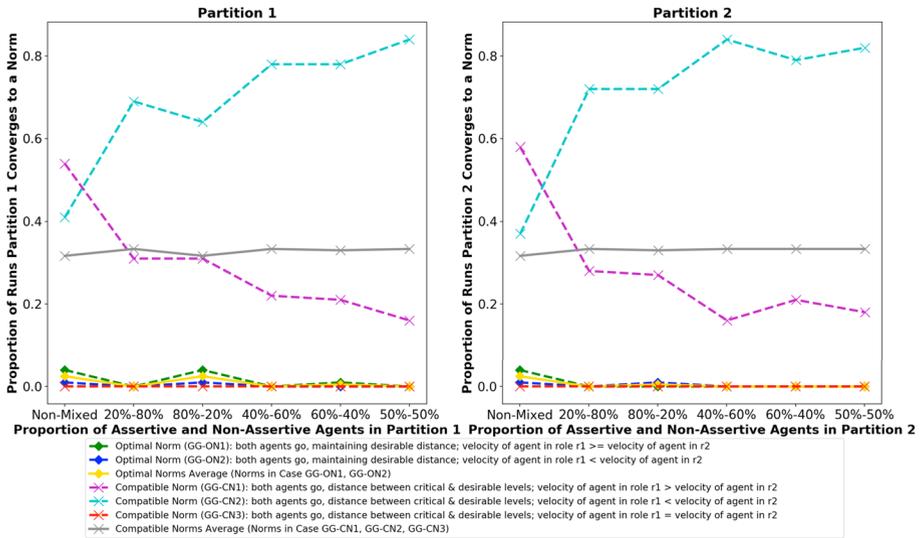


Fig. 4 Proportion of runs agents converged to different norms using different communities (non-mixed or mixed in different proportions) when the number of “inter-partition edges” is low

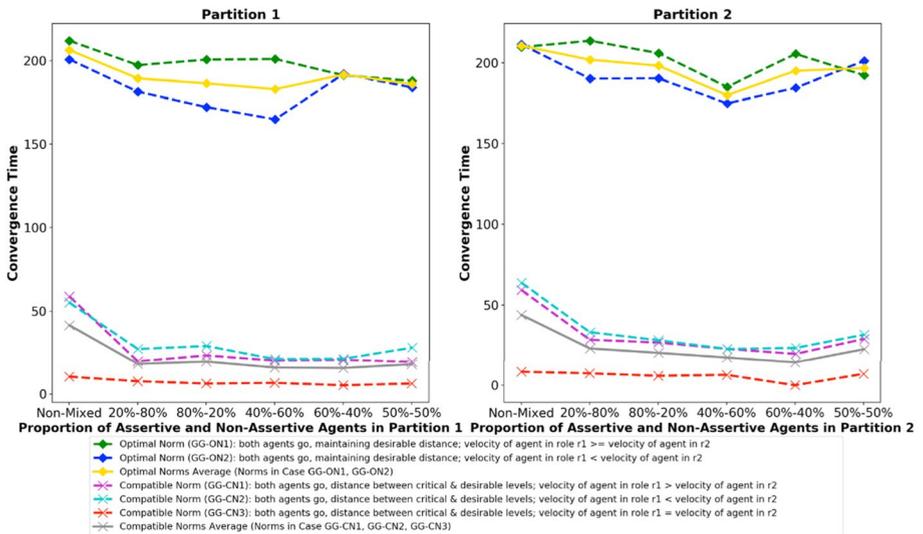


Fig. 5 Convergence times in different communities (non-mixed or mixed in different proportions) when the number of “inter-partition edges” is low

(50.336921 and 77.598847). In order to better understand these results, we performed the same experiment for the non-mixed case with both partitions containing only one agent type (assertive agents). This is to minimize any effect the agent type of the partition has on convergence time (note that assertive and non-assertive agents have different interoperable regions). In this variation, as expected in hypothesis 1, we noted that both partitions

Table 1 Frequency of convergence and convergence time of the norm with the highest frequency of convergence when the communities are non-mixed or mixed in different proportions

Assertive-non-assertive agents ratio	Partition 1		Partition 2	
	Frequency of convergence	Convergence time (standard deviation)	Frequency of convergence	Convergence time (standard deviation)
Non-mixed	0.1892 (highest frequency of convergence in compatible norm in case GG-CN1) Assertive agents only case: 0.1688 (highest frequency of convergence in optimal norm in case GG-ON1)	59 (50.336921)	0.1913 (highest frequency of convergence in compatible norm in case GG-CN2) Assertive agents only case: 0.1691 (highest frequency of convergence in optimal norm in case GG-ON1)	64 (77.598847)
20–80%	0.2111 (highest frequency of convergence in compatible norm in case GG-CN2)	27 (16.257488)	0.2150 (highest frequency of convergence in compatible norm in case GG-CN2)	33 (18.545231)
80–20%	0.1483 (optimal norm in case GG-ON1) 0.1762 (highest frequency of convergence in compatible norm in case GG-CN2)	197 (50.287474) 29 (19.901987)	0.1524 (optimal norm in case GG-ON1) 0.1724 (highest frequency of convergence in compatible norm in case GG-CN2)	214 (46.814767) 28 (17.664812)
40–60%	0.1718 (optimal norm in case GG-ON1) 0.2007 (highest frequency of convergence in compatible norm in case GG-CN2)	201 (43.007562) 21 (10.824509)	0.1715 (optimal norm in case GG-ON1) 0.1934 (highest frequency of convergence in compatible norm in case GG-CN2)	206 (47.715066) 23 (13.438914)
60–40%	0.1574 (optimal norm in case GG-ON1) 0.1854 (highest frequency of convergence in compatible norm in case GG-CN2)	201 (46.000977) 21 (11.077956)	0.1539 (optimal norm in case GG-ON1) 0.1878 (highest frequency of convergence in compatible norm in case GG-CN2)	185 (36.226708) 23 (12.409271)
50–50%	0.1617 (optimal norm in case GG-ON1) 0.1983 (highest frequency of convergence in compatible norm in case GG-CN2)	191 (39.098669) 28 (16.929749)	0.1648 (optimal norm in case GG-ON1) 0.1967 (highest frequency of convergence in compatible norm in case GG-CN2)	206 (41.573068) 32 (17.369041)
	0.1597 (optimal norm in case GG-ON1)	188 (40.744884)	0.1542 (optimal norm in case GG-ON1)	192 (39.701399)

Both partitions either converged to the norm in case GG-CN1 or GG-CN2. (Note: in GG-CN1 and GG-CN2, both agents in roles r_1 and r_2 go, and r_1 pulls out between the desirable and critical thresholds for distance. They differ in pull out velocity of r_1 (i.e., v_1') and initial velocity of r_2 (i.e., v_2) which are: $v_1' > v_2$ and $v_1' < v_2$, respectively. Meanwhile, in GG-ON1 (for variation), both agents in roles r_1 and r_2 go, and the agent in role r_1 pulls out maintaining the desirable threshold for distance. Also, $v_1' > v_2$)

converged more quickly to the norm in case GG-ON1 compared to the mixed cases (see Table 1). For example, in partition 1, the convergence time was 174, whereas in the mixed cases of 20–80%, 80–20%, 40–60%, 60–40% and 50–50%, it emerged at 197, 201, 201, 191 and 188. Table 1 reports results of both variations of this experiment.

As expected, when the number of inter-partition edges is low and the partitions are mixed (see hypothesis 2), we observed that the two partitions converged to the same compatible norm (norm in case GG-CN2) (see Table 1).

Now we analyse the effects of the relative number of inter-partition edges have on norms emergence (hypothesis 3). For this, we used an equally mixed population (50-50%) of assertive and non-assertive agents. Figure 6 shows the convergence time when the number of inter-partition edges in partitions is increased at individual partition level. Meanwhile, Fig. 7 shows the proportion of runs that the different norms emerge for the same case at the individual partition level.

We observed a clear pattern of both partitions converging to the same compatible norm, i.e., the compatible norm GG-CN2 of the case study in all cases. Also, we noted a gradual decrease of frequency of convergence with the increasing of the relative number of inter-partition edges (see Table 2).

The results also indicate of a clear trend of convergence time reducing when the relative number of inter-partition edges is increased (see hypothesis 3). For example, as seen in Fig. 6, convergence times for the optimal norm GG-ON1 in partition 1 reduces steadily from 206 to 77, and from 206 to 84 in partition 2. Note that the slight increases seen in convergence times at 0.010, 0.025 for the norm in case GG-CN2 in partitions 1 and 2 are still within the standard deviations of the convergence times, which are 20.071532 and 18.029695 respectively (see last row, Table 2).

Also, as expected, using the metric proportion of runs (see Fig. 7), we noted the probability of both partitions converging to a global state is relatively higher when the number of inter-partition edges is increased. When the number of inter-partition edges is increased,

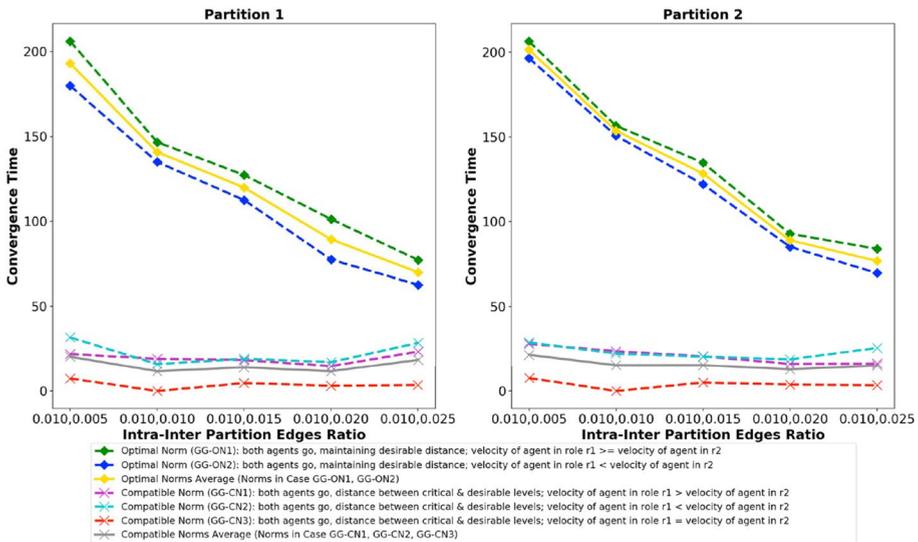


Fig. 6 Convergence times in different communities when the number of “inter-partition edges” in partitions is increased

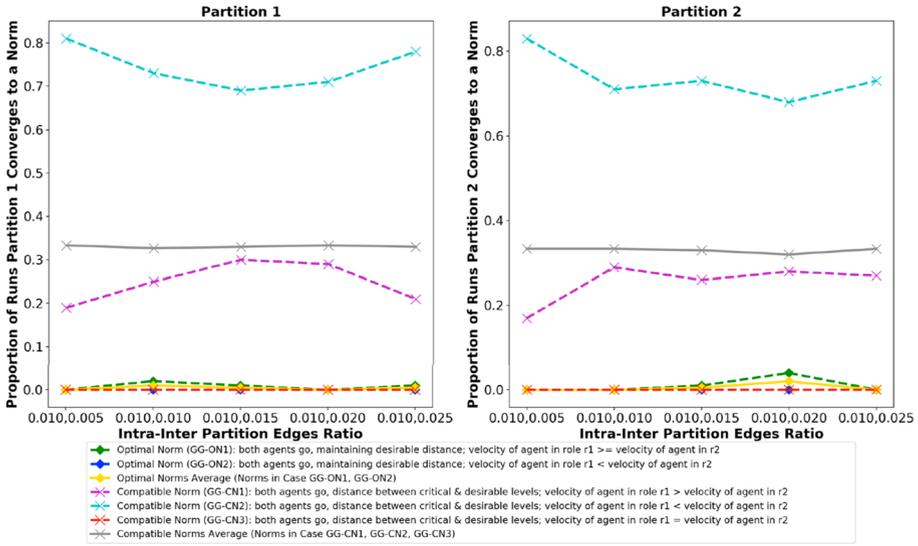


Fig. 7 Proportion of runs agents converged to different norms when the number of “inter-partition edges” in partitions is increased

Table 2 Frequency of convergence and convergence time of the norm with the highest frequency of convergence (i.e., compatible norm in the case GG-CN2) when the relative number of “inter-partition edges” is increased in a mixed community

Intra-inter partition edges ratio	Partition 1		Partition 2	
	Frequency of convergence	Convergence time (standard deviation)	Frequency of convergence	Convergence time (standard deviation)
0.010, 0.005	0.2015	32 (19.876559)	0.1972	29 (17.657859)
0.010, 0.010	0.1417	16 (9.756408)	0.1464	22 (12.976791)
0.010, 0.015	0.1163	19 (11.759949)	0.1137	20 (12.271834)
0.010, 0.020	0.0927	17 (11.227337)	0.0970	19 (11.668402)
0.010, 0.025	0.0868	28 (20.071532)	0.0818	26 (18.029695)

it shows the general trend that the proportion of runs the compatible norm in case GG-CN2 emerging is decreasing while the same measure for the norm in case GG-CN1 is increasing. That is, the gap between the two norms which have the highest relative frequency of convergence, is seen to be narrowing. The slight increase of gap seen at 0.010, 0.025 for partition 1 is likely because of the standard deviations of convergence times for the two norms (in 100 runs), which are used to calculate proportion of runs.

5.3 Experiment 2: intra-partition edges on norms emergence

5.3.1 Experimental setup

This experiment was conducted to show the influence of “intra-partition” edges on norms emergence.

Parameters used in the topology: For this, we used a non-mixed community of agents and gradually increased the relative number of intra-partition edges with following probability of intra-partition–inter-partition edges ratios (p_{in} , p_{out}): (i) 0.005, 0.010; (ii) 0.010, 0.010; (iii) 0.015, 0.010; (iv) 0.020, 0.010; and (v) 0.025, 0.010.

Meanwhile, the **topological metrics** of the simulations conducted are provided below, respectively.

- Average degree of network: 14.9998; density: 0.0075; intra community edges: 4993.7; inter community edges: 10006.14.
- Average degree of network: 19.9811; density: 0.0100; intra community edges: 9995.36; inter community edges: 9985.71.
- Average degree of network: 25.0019; density: 0.0125; intra community edges: 14991.41; inter community edges: 10010.53.
- Average degree of network: 29.9828; density: 0.0150; intra community edges: 19983.94; inter community edges: 9998.87.
- Average degree of network: 34.9381; density: 0.0175; intra community edges: 24957.23; inter community edges: 9980.88.

Hypothesis

The intra-partition edges increase the internal connectivity of agents in a partition. The more connected the agents are within a partition, we expect the two partitions to converge to different states more quickly. But, this may slow the overall system convergence, if there are sufficient number of inter-partition edges to achieve global convergence.

5.3.2 Results

Now we discuss the results to show the effects of the number of “intra-partition” edges on norms emergence. Figure 8 shows convergence times for different relative number of intra-partition edges.

Figure 9 presents the proportion of runs the system converged to the different norms when the number of intra-partition edges in partitions is gradually increased.

A clear pattern we observed is when the number of intra-partition edges is increased, the frequency of convergence to a norm in a partition increases as well. For example, in the case we maintained intra-inter probability ratio of 0.005, 0.010, we observed that partition 1 converged to the compatible norm described in case GG-CN2 of the case study with a frequency of convergence of 0.0893 (see Table 3). This gradually increased for the other four cases of probability ratios with a frequency of convergence of 0.1400, 0.1687, 0.1805 and 0.1938. The same pattern was observed in partition 2 as well. We further observed that in all the five different intra-partition edges cases, both partitions either converged to compatible norm in case GG-CN1 or case GG-CN2.

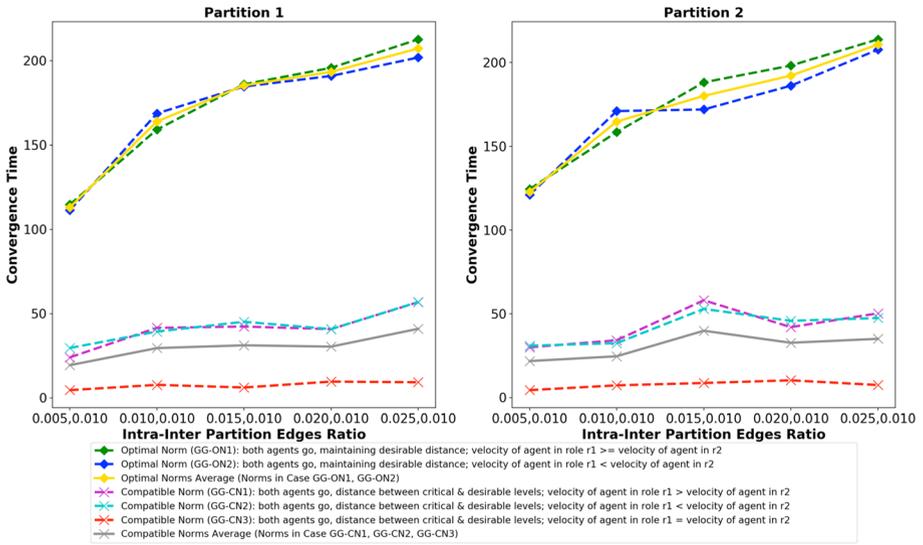


Fig. 8 Convergence times for different number of “intra-partition edges” in partitions when the number of intra-partition edges is increased

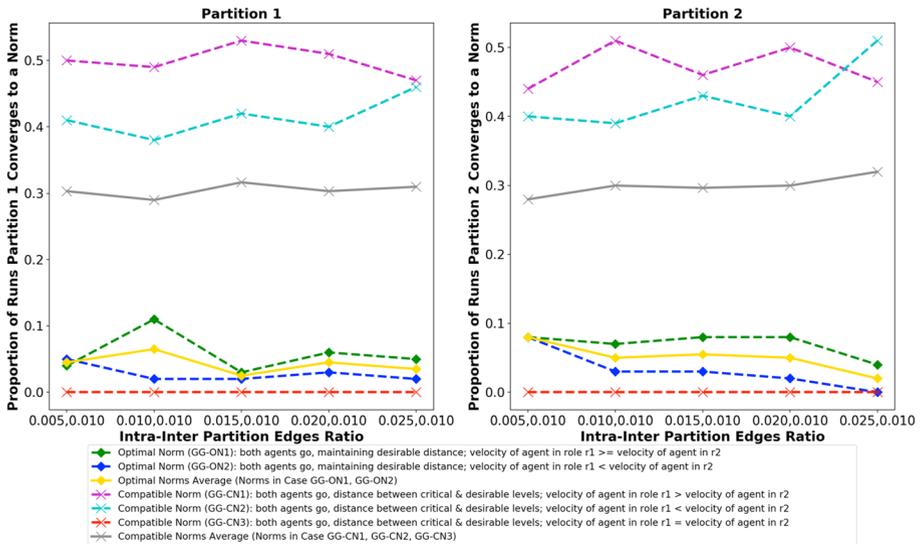


Fig. 9 Proportion of runs agents converged to different norms when the number of “intra-partition edges” in partitions is increased

The results also show a clear trend of convergence time increasing when the relative number of intra-partition edges is increased. This essentially is the opposite behaviour we observed from increasing the “inter-partition edges” in the previous experiment. As seen in Fig. 8, convergence times for the optimal norm GG-ON1 in partition 1 increases steadily

Table 3 Frequency of convergence and convergence time when the relative number of “intra-partition edges” is increased in a non-mixed community

Intra-inter partition edges ratio	Partition 1		Partition 2	
	Frequency of convergence	Convergence time (standard deviation)	Frequency of convergence	Convergence time (standard deviation)
0.005, 0.010	0.0893 (highest frequency of convergence in compatible case GG-CN2) 0.0873 (Compatible norm in case GG-CN1)	30 (25.845454) 24 (19.324728)	0.0858 (highest frequency of convergence in compatible case GG-CN1) 0.0855 (Compatible norm in case GG-CN2)	30 (27.267745) 31 (27.207166)
0.010, 0.010	0.1402 (highest frequency of convergence in compatible case GG-CN1)	42 (42.188387)	0.1326 (highest frequency of convergence in compatible case GG-CN2)	32 (28.661980)
0.015, 0.010	0.1400 (Compatible norm in case GG-CN2) 0.1687 (highest frequency of convergence in compatible case GG-CN2)	39 (37.862118) 45 (45.122565)	0.1301 (Compatible norm in case GG-CN1) 0.1683 (highest frequency of convergence in compatible case GG-CN2)	34 (31.060681) 53 (46.620987)
0.020, 0.010	0.1673 (Compatible norm in case GG-CN1) 0.1805 (highest frequency of convergence in compatible case GG-CN2)	42 (40.571300) 41 (37.059606)	0.1655 (Compatible norm in case GG-CN1) 0.1806 (highest frequency of convergence in compatible case GG-CN1)	58 (55.953642) 42 (36.726659)
0.025, 0.010	0.1775 (Compatible norm in case GG-CN1) 0.1966 (highest frequency of convergence in compatible case GG-CN1)	41 (36.859206) 57 (52.627631)	0.1785 (Compatible norm in case GG-CN2) 0.1970 (highest frequency of convergence in compatible case GG-CN1)	46 (42.847183) 50 (53.504486)
	0.1938 (Compatible norm in case GG-CN2)	57 (52.191555)	0.1965 (Compatible norm in case GG-CN2)	48 (45.067616)

Both partitions either converged to the compatible norm in the case GG-CN1 or GG-CN2

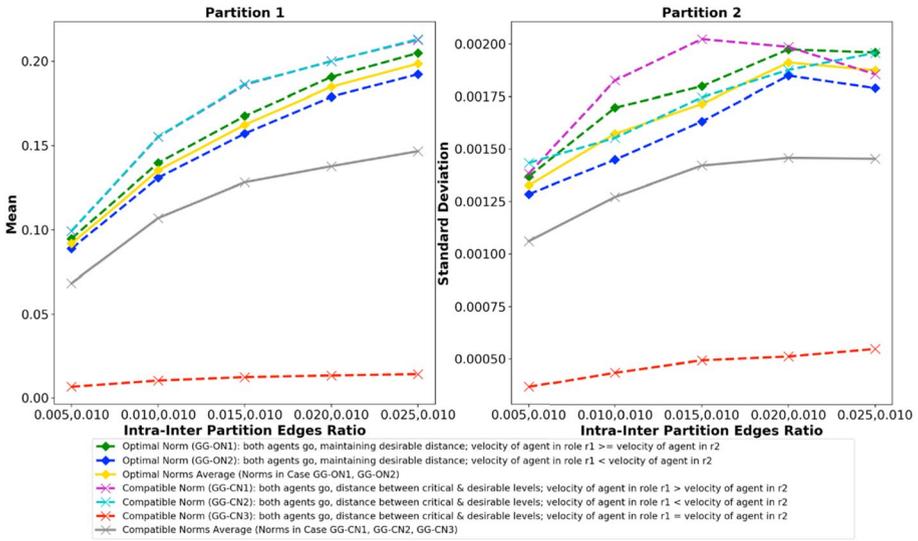


Fig. 10 Mean and standard deviation of different optimal and compatible norms in partition 1 when the number of “intra-partition edges” is increased. This is shown after all the norms in the partitions have converged

from 115 to 213, and from 125 to 214 in partition 2. Note that the decreasing seen in convergence times at 0.020, 0.010 for the norms in cases GG-CN1 and GG-CN2 in partition 2 are still within the standard deviation of the convergence time in 100 runs (36.726659 and 42.847183 respectively).

We note that we expected to see convergence time to decrease but our results do not show this pattern, and one reason to observe this can be explained as follows. As the number of intra-partition edges increases, we observe that the probability of different compatible norms emerging is higher (see Fig. 9 — the gap between the norms in cases GG-CN1 and GG-CN2 is seen to be reducing). So, even though the number of intra-partitions edges is increased, the effect of inter-partition edges can be making more compatible norms to emerge. Thus, this can be a reason for the increase in convergence time.

Figure 10 shows the mean and standard deviation of frequency of convergence of the different optimal and compatible norms in partition 1 when the number of intra-partition edges is increased. These are shown after all the norms in the partitions have converged (i.e., from 800–1000 iterations). As mentioned in Sect. 5.1.3, we consider that at iteration 800 all the norms in the system have essentially converged. From Fig. 10, it is clear that there is an increasing trend in the mean and standard deviation of frequency of convergence (after the system has converged) when the number of intra-partition edges is increased.

5.4 Experiment 3: interoperable region on norms emergence

5.4.1 Experimental setup

Using an equally mixed community of agents, this experiment was conducted to show how the relative size of the “interoperable region” (overlapping) can influence norms emergence. As mentioned in Sect. 5.1, we used separate distributions for assertive and non-assertive agents when generating velocities (i.e., initial velocity of the agent in role r_2 and the pull out velocity of the agent in role r_1), and the distance between the agent in roles r_1 and r_2 .

In this experiment, we gradually increased the interoperable region of the distributions used to generate velocities and distances, and compared the effects on norms emergence.

Parameters used in the topology: In all five cases of the experiment, we used the same network graph and maintained the probability of intra-partition–inter-partition edges ratio (p_{in}, p_{out}) of 0.005, 0.010 (see Table 4).

The **topological metrics** are:

- Average degree of network: 14.9904; density: 0.0075; intra community edges: 4996.03; inter community edges: 9994.34.

Hypothesis

We expect to see that if the interoperable region is relatively limited, then the communities to converge to the same norm. However, this is dependent on the relative high number of inter-partition edges. On the other hand, if the interoperable region is relatively wider, then the communities are expected to converge to different norms which are still interoperable.

5.4.2 Results

Now we analyse the effects of the “interoperable region” on the emergence of norms using the case study scenario. Figure 11 shows proportion of runs partition 1 and 2 converge to the different optimal and compatible norms, while Fig. 13 shows time to convergence in both partitions. Figure 12 shows the frequency of convergence of the norm in case GG-CN2 (the norm that emerged in most cases) when “interoperable region” is steadily increased.

As expected, the results indicate that when the interoperable region is “low” then both partitions converged to the same norm. In this case, both partitions converged to the norm that emerges in the compatible norm case GG-CN2 of the case study with a frequency of convergence of 0.1128 and 0.1131 (see Table 5). This is dependent on the relative high number of inter-partition edges (p_{in}, p_{out} is 0.005, 0.010). Also, we noted the probability of both partitions converging to a different norm is higher when the interoperable region is broader. We show this using proportion of runs metrics (see Fig. 11). It shows the general trend that when the interoperable region is increased, the proportion of runs the compatible norm in case GG-CN2 emerging is decreasing,

Table 4 “interoperable regions” created with different distributions of velocities and distances for assertive and non-assertive agents

Velocities	Interoperable region	Distances	Interoperable region
<i>Low interoperable region</i>			
Assertive: 10–50kph; Non-assertive: 40–80kph	40–50kph	Assertive: 60–85m; Non-assertive: 75–100m	75–85m
<i>Low-medium interoperable region</i>			
Assertive: 10–55kph; Non-assertive: 35–80kph	35–55kph	Assertive: 60–87.5m; Non-assertive: 72.5–100m	72.5–87.5m
<i>Medium interoperable region</i>			
Assertive: 10–60kph; Non-assertive: 30–80kph	30–60kph	Assertive: 60–90m; Non-assertive: 70–100m	70–90m
<i>Medium-high interoperable region</i>			
Assertive: 10–65kph; Non-assertive: 25–80kph	25–65kph	Assertive: 60–92.5m; Non-assertive: 67.5–100m	67.5–92.5m
<i>High interoperable region</i>			
Assertive: 10–70kph; Non-assertive: 20–80kph	20–70kph	Assertive: 60–95m; Non-assertive: 65–100m	65–95m

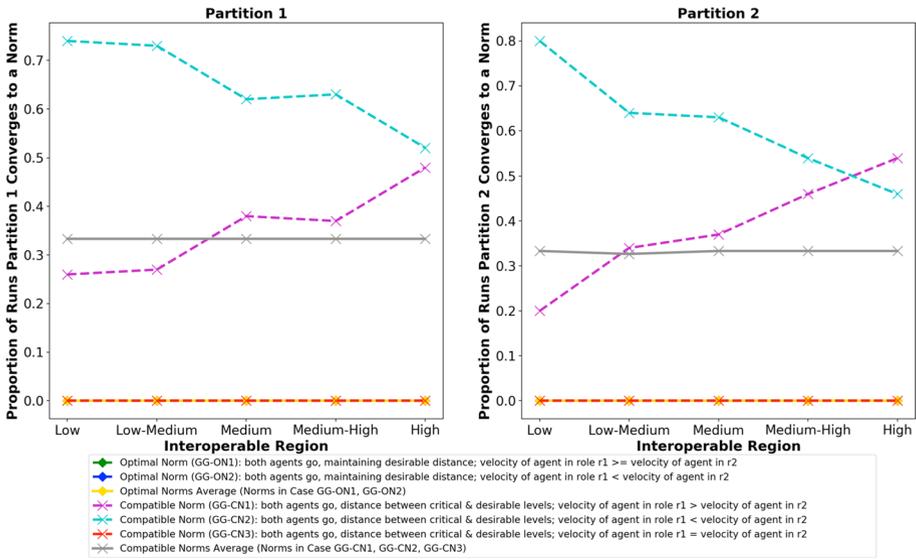


Fig. 11 Proportion of runs agents converged to different norms when the “interoperable region” in partitions is gradually increased

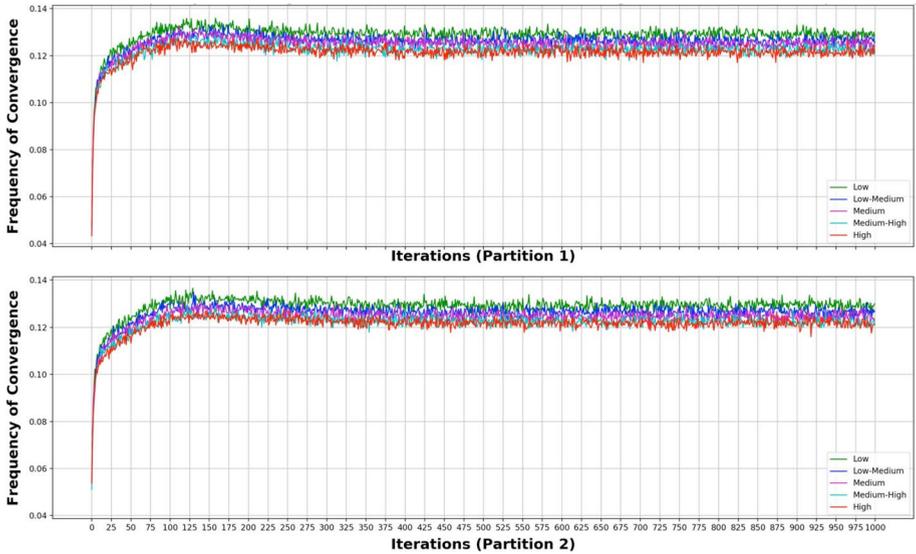


Fig. 12 Frequency of convergence of the norm in the case GG-CN2 when “interoperable region” is steadily increased in partition 1 and 2

while the same measure for the compatible norm in case GG-CN1 is increasing (gap between them reducing). Thus, it shows that if the interoperable region is relatively wider, then the communities are expected to converge to different norms which are still interoperable.

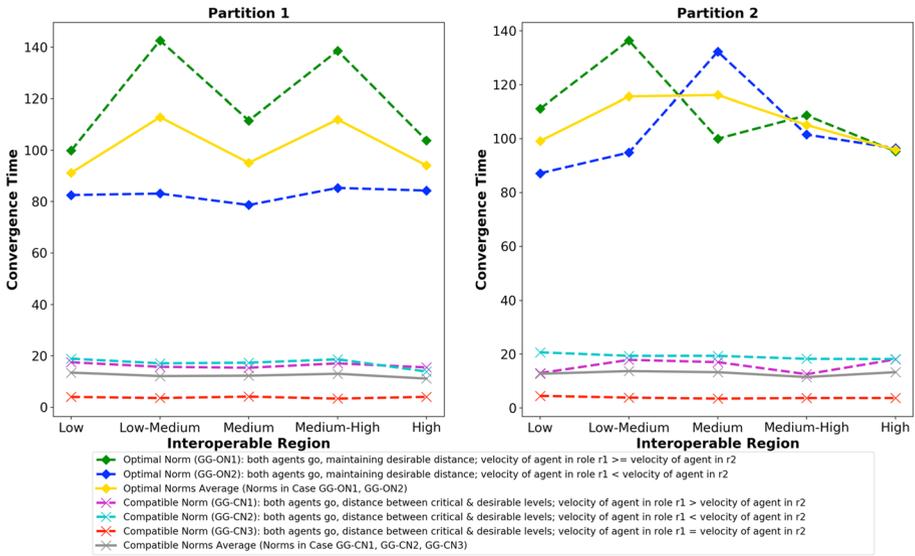


Fig. 13 Convergence time of partitions when the “interoperable region” is gradually increased

In this experiment, we observed that the both partitions converged to compatible norm in case GG-CN2 or GG-CN1. We also observed when the interoperable region is widened, the frequency of convergence of a norm decreased. Figure 12 shows the frequency of convergence of the norm in case GG-CN2 decreasing in both partitions when “interoperable region” is steadily increased. Also, as shown in Table 5, when the interoperable region is “low”, the frequency of convergence for the norm in case GG-CN2 in partition 2 is 0.1131, which gradually decreases to 0.1047 when the interoperable region is “high”. Furthermore, we observed that the convergence time is seen to decrease when the interoperable region is increased. As shown in Table 5, convergence time for the norm in case GG-CN2 is seen to be decreasing for both partitions. Note that the slight increase seen in the “medium-high” case in partition 1 is still within the standard deviation (11.451895) of convergence time.

5.5 Conformity and diversity based on network community and agent type

5.5.1 Experimental setup

The goal of this experiment is to measure the overall strength of local conventions in a system (“conformity”), and “diversity” of the agents’ actions throughout the system. The experiment was conducted based on two different perspectives: network community (partitions) and agent type. This is because sometimes there can be convergence to a higher conformity relatively more quickly in terms of the topological interpretation, but less quickly with respect to the agent type. Furthermore, compared to Hu and Leung’s work [22], we go a step further by basing our measures of conformity and diversity on the *state* of the two interacting agents. For example, see Figs. 14–15 where we measure conformity and diversity in a case where the partitions contain mixed agents with an intra-partition–inter-partition ratio of 0.010, 0.015.

Table 5 Frequency of convergence and convergence time of the norm with the highest frequency of convergence when “interoperable region” is steadily increased

Interoperable region	Partition 1		Partition 2	
	Frequency of convergence	Convergence time (standard deviation)	Frequency of convergence	Convergence time (standard deviation)
Low	0.1128 (compatible norm in case GG-CN2)	19 (12.706986)	0.1131 (compatible norm in case GG-CN2)	21 (13.883141)
Low-medium	0.1093 (compatible norm in case GG-CN2)	17 (10.885196)	0.1109 (compatible norm in case GG-CN2)	19 (14.036510)
Medium	0.1104 (compatible norm in case GG-CN2)	17 (12.626080)	0.1082 (compatible norm in case GG-CN2)	19 (9.850056)
Medium-high	0.1108 (compatible norm in case GG-CN2)	19 (11.451895)	0.1063 (compatible norm in case GG-CN2)	18 (11.703418)
High	0.1060 (compatible norm in case GG-CN2)	14 (9.239935)	0.1048 (Highest frequency-compatible norm in case GG-CN1)	18 (11.346449)
			0.1047 (compatible norm in case GG-CN2)	18 (11.110621)

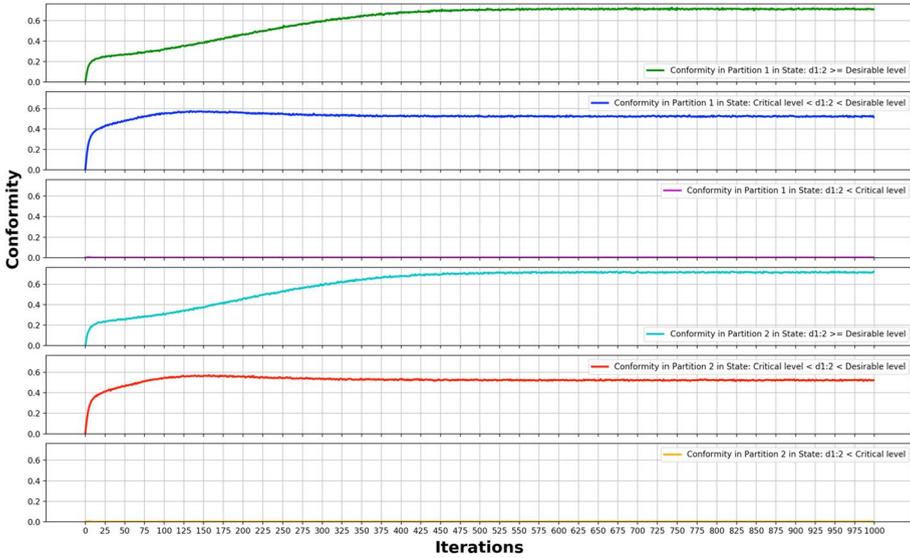


Fig. 14 Individual “conformity” of agents in both partitions for the three states

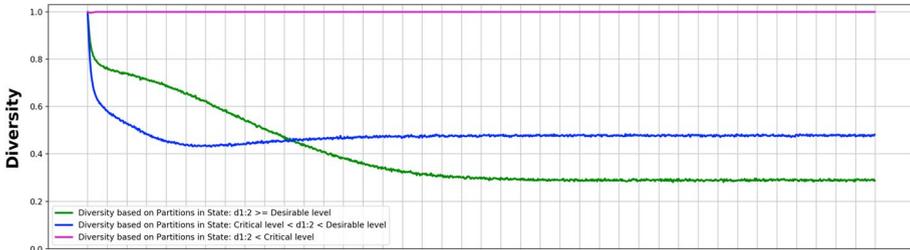


Fig. 15 “Diversity” of agents based on partitions in the global community in the three states

Hypothesis

At times we expect to see convergence to a higher conformity relatively more quickly in terms of the topological interpretation, but less quickly with respect to the agent type.

5.5.2 Results

Figure 14 shows that the individual conformity of agents in partition 1 and 2 in the three main states. It shows that the state $d_{1:2} \geq \lambda_{Dd_i}$ has the highest conformity followed by the states $\gamma_{Cd} < d_{1:2} < \lambda_{Dd_i}$ and $d_{1:2} \leq \gamma_{Cd}$. Figure 15 shows the effects of diversity of the global community in a given state based on partitions. It shows that the state $d_{1:2} \leq \gamma_{Cd}$ has the highest diversity followed by the states $\gamma_{Cd} < d_{1:2} < \lambda_{Dd_i}$ and $d_{1:2} \geq \lambda_{Dd_i}$. Similarly, we can show the measures of conformity and diversity based on the state of the interacting agents.

5.6 Experiment 5: five partitions case

We also conducted experiments using five partitions to check whether there is an effect on norms emergence when the number of partitions is different to the two partitions used in previous experiments. As an example, we conducted the experiment on intra-partition edges (Sect. 5.3) again, but with five partitions using a non-mixed community of agents. As in the experiment for two partitions, we observed that when the number of intra-partition edges is increased, the frequency of convergence to a norm and convergence time in a partition increase as well. Like in the two partitions case, here also we observed that all partitions converged either to the compatible norm in case GG-CN1 or GG-CN2 of the case study. The results essentially show that there is no or very little difference in the main patterns we observed when the number of partitions is different to two partitions.

5.7 Discussion

Based on the results we analysed from the experiments, we draw several conclusions, as presented below. We used several properties of the underlying network topology to show their influence on the emergence of conventions, such as inter-partition edges, intra-partition edges, and interoperable region. We also conducted an experiment on conformity and diversity from both viewpoints of the network community (partitions) and the agent type. Furthermore, by using a different number of partitions (five), it was investigated whether we obtained the same pattern of results, as when there are two partitions.

1. From the results, we conclude that when the communities are well mixed and if there are few *inter-partition edges*, the probability of the same norm emerging in the two partitions is greater compared to the case where the communities are not mixed.

A clear observable pattern we identified was when the relative number of inter-partition edges is increased, the *frequency of convergence* to a norm decreases.

Furthermore, we observed that when the number of inter-partition edges is increased, the system showed the general trend of converging more quickly (i.e., *convergence time* reduces).

From the results, we also conclude that increasing the inter-partition edges did have an effect on both partitions converging to a global state. This observed pattern of results can be explained as follows. A few inter-partition edges can result in a relatively small overlap of the action spaces. Thus, there will be a low probability of the same norm emerging. The more the number of inter-partition edges are, the higher the probability of both partitions converging to a global state.

2. From the results, we noted that when the number of intra-partition edges is increased, the *frequency of convergence* to a norm in a partition increases as well. This essentially was the opposite behaviour we observed from increasing the inter-partition edges. The results also showed convergence time increasing when the relative number of intra-partition edges is increased. Furthermore, we noted an increasing trend of the mean and standard deviation of frequency of convergence (after the system has converged).
3. From the results we also conclude that if the *interoperable region* is relatively limited, then the probability of communities converging to the same norm is higher. On the other hand, if the interoperable region is relatively wider, then the communities have a higher probability of converging to different norms which are still interoperable.

- We also observed that when the interoperable region is widened, the frequency of convergence of a particular norm and convergence time decrease.
4. Our results also indicate the importance of showing local conventions in a system (*conformity*), and *diversity* of the agents' actions based from different perspectives, such as network community (partitions) and agent type. This is because sometimes there can be convergence to a higher conformity relatively more quickly with respect to the topological interpretation, but less quickly in terms of the agent type. In addition, it is important to base our measures of conformity and diversity on the *state* of the two interacting agents.
 5. The results also indicate that essentially we found no or very little difference in the patterns we observed when the number of partitions is different to two partitions. This was evident by the results of the experiment on intra-partition edges with five partitions.

6 Conclusions and future work

In this paper, we have proposed a novel agent-based modelling approach of norm emergence based on reinforcement learning to influence decisions of users, such as drivers in the case study. The nature of this norm emergence is dependent on the context of the population of users and the system. We have evaluated our approach using a case study of self-driving cars focussing on the scenario of an autonomous vehicle performing a left-turn at a T-intersection. In our simulations, we have used several characteristics of the underlying network topology to show their influence on norms emergence, i.e., *inter-partition edges* (and how mixed the communities are), *intra-partition edges* and the *interoperable region*. We also investigated the significance of *conformity* and *diversity* from both the viewpoint of the network community and the agent type. Furthermore, we investigated whether the same pattern of results is seen when the number of partitions is varied.

We analysed the results of the experiments using three topological metrics to predict influence on driver behaviour, namely frequency of convergence, convergence time, and proportion of runs the system converges to a norm. We showed how the relative number of inter-partition edges, intra-partition edges and the size of the interoperable region affect the frequency of convergence, convergence time, and probability of the same norm (or global norms) emerging. From the simulations, we draw several conclusions: (i) if the communities are well mixed and if there are few inter-partition edges, then the probability of the same norm emerging in the two partitions is greater, compared to the case where the communities are not that mixed; (ii) frequency of convergence of a norm decreases when the relative number of inter-partition edges is increased or when the interoperable region is widened; in contrast, the convergence rate increases when the relative number of intra-partition edges is increased; (iii) the system is seen to converge more quickly when the relative number of inter-partition edges or the interoperable region is increased; (iv) the probability of both partitions converging to a global state increases with the relative number of inter-partition edges and the size of the interoperable region; (v) the significance of showing local conventions in a system (conformity) and diversity of the agents' actions for a particular state, based from both network community and agent type; and (vi) the number of partitions does not indicate to have an effect on the different patterns or trends we observed in the two partitions case.

These results have implications for understanding norm emergence in real-world systems, in particular where there is the potential for different norms to emerge in different

communities and where there is not a simple binary distinction between (optimal) compatible actions and incompatible actions. In applications such as driving, we see that different norms emerge in different geographical regions (e.g., assertiveness differences between rural and city drivers [38]) and that there is a range of compatible actions where individuals report different satisfaction levels for different driving behaviours (e.g., following distance in highway conditions [44]). Our results give some insight into how such factors might give rise to norm emergence in such settings, in particular (i) that where communities contain a mix of individuals (in terms of their preferences) there is an increased probability of the same norm emerging in different communities; (ii) if the interoperable region is relatively small then the probability of different communities arriving at the same norm is increased; and (iii) the number of communities has relatively little impact on the norms that emerge overall. In the context of driving, for example, this means that we might expect geographical areas that typically have a mix of rural-based and urban-based drivers to arrive at similar norms, compared to areas that are primarily rural or urban in terms of prevalent driver types. Furthermore, while we typically are not able to modify or directly influence the “rules of the environment” for real-world applications, in cases where this is possible then reducing the size of the interoperable region will increase the likelihood of a consistent global norm. Regulation and best practice for self-driving cars is an example of where this is possible, and where reducing the size of the interoperable region for actions such as pull-out distance and following distance, might increase the probability of single norm emerging. Where human controlled vehicles interact with autonomous vehicles, reducing the number of possible norms increases the likelihood of successful and satisfactory interactions.

As for further work, we identify several possible future directions. At present, the pay-offs or rewards for the all the situations in the case study are predefined. This implies that we have a set of prescribed agent interactions. In future, it can be useful to handle *dynamic* [23] and *off-diagonal reward matrices* [4], so the norms can be changed or destabilised at runtime. Future work can also include investigation of reinforcement learning of agents under *partial observability* [39], and the *dynamic population* of agents (i.e., the individual agents can leave and appear at runtime). The recent advancements in AI and machine learning in safety-critical settings and the use of increasingly complex and non-transparent algorithms have driven the need to make machine decisions transparent, understandable and *explainable* [12]. Agents need to be able to *explain* why one’s actions are morally right or wrong [7]. Doing so, generally increases acceptance and trust, which are essential in safety-critical systems. Therefore, in this context, it could be useful to investigate how two interacting agents can explain the reason for choosing a particular action to an individual or to policy makers.

Acknowledgements This work was supported by Jaguar Land Rover and the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/N012380/1 as part of the jointly funded Towards Autonomy: Smart and Connected Control (TASCC) Programme. D.A. is also supported by the UKRI Trustworthy Autonomous Systems Node in Functionality under EPSRC grant EP/V026518/1.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Abeywickrama, D. B., Cirstea, C., & Ramchurn, S. D. (2019). Model checking human-agent collectives for responsible AI. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, (pp. 1–8). doi: <https://doi.org/10.1109/RO-MAN46459.2019.8956429>.
2. Abeywickrama, D. B., Mamei, M., & Zambonelli, F. (2018). Engineering collectives of self-driving vehicles: The SOTA approach. In T. Margaria & B. Steffen (Eds.), *Leveraging applications of formal methods, verification and validation. distributed systems* (pp. 79–93). Cham: Springer International Publishing.
3. Airiau, S., Sen, S., & Villatoro, D. (2014). Emergence of conventions through social learning - heterogeneous learners in complex networks. *Autonomous Agents Multi Agent System*, 28(5), 779–804. <https://doi.org/10.1007/s10458-013-9237-x>.
4. Amaral, M. A., & Javarone, M. A. (2020). Strategy equilibrium in dilemma games with off-diagonal payoff perturbations. *Physical Review E*, 101, 6. <https://doi.org/10.1103/physreve.101.062309>.
5. Borglund, T., Hu, S., & Leung, H. F. (2018). The effects of fixed-strategy agents on local convention emergence in multi-agent systems. In Z. Shi, E. Mercier-Laurent, & J. Li (Eds.), *Intelligent information processing IX* (pp. 99–108). Cham: Springer International Publishing.
6. Chalub, F. A. C. C., Santos, F. C., & Pacheco, J. M. (2006). The evolution of norms. *Journal of Theoretical Biology*, 241, 233–240.
7. Conitzer, V., Sinnott-Armstrong, W., Borg, J. S., Deng, Y., & Kramer, M. (2017). Moral decision making frameworks for artificial intelligence. In *Proc. of the 31st AAAI Conference on Artificial Intelligence*, February 4–9, 2017, San Francisco, California, USA, (pp. 4831–4835).
8. Delgado, J. (2002). Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1), 171–185. [https://doi.org/10.1016/S0004-3702\(02\)00262-X](https://doi.org/10.1016/S0004-3702(02)00262-X)
9. Dogan, E., Chatila, R., Chauvier, S., Evans, K., Hadjixenophontos, P., & Perrin, J. (2016). Ethics in the design of automated vehicles: The AVEthics project. In *Proceeding of the 1st Workshop on Ethics in the Design of Intelligent Agents*, (pp. 10–13). The Hague, The Netherlands.
10. Epstein, J. M. (2001). Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18(1), 9–24. <https://doi.org/10.1023/A:1013810410243>
11. Franks, H., Griffiths, N., & Anand, S. S. (2013). Learning influence in complex social networks. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '13*, (pp. 447–454). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
12. Goebel, R., Chandler, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable AI: the new 42? In: Holzinger, A., Kieseberg, P., Tjoa, A. M., & Weippl, E. (Eds.), *2nd International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE), Machine Learning and Knowledge Extraction*, vol. LNCS-11015, (pp. 295–303). Springer, Hamburg, Germany. doi: https://doi.org/10.1007/978-3-319-99740-7_21.
13. Griffiths, N., & Luck, M. (2010). Norm emergence in tag-based cooperation. *Organization, Institutions and Norms in Multi-Agent Systems (COIN10)9th International Workshop on Coordination* (pp. 80–87). Toronto: Canada.
14. Habib, N. (2019). *Hands-on Q-learning with Python: Practical Q-learning with OpenAI Gym, Keras, and TensorFlow*. UK: Packt Publishing.
15. Hao, J., & Leung, H. F. (2013). The dynamics of reinforcement social learning in cooperative multiagent systems. In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, (pp. 184–190).
16. Hao, J., Sun, J., Chen, G., Wang, Z., Yu, C., & Ming, Z. (2017). Efficient and robust emergence of norms through heuristic collective learning. *ACM Transactions on Autonomous and Adaptive Systems*, 12, 4.
17. Haynes, C., Luck, M., McBurney, P., Mahmoud, S., Vitek, T., & Miles, S. (2017). Engineering the emergence of norms: A review. *Knowledge Engineering Review*. <https://doi.org/10.1017/S026988917000169>.
18. Hoch, N., Bensler, H. P., Abeywickrama, D., Bureš, T., & Montanari, U. (2015). *The E-mobility case study* (pp. 513–533). Cham: Springer International Publishing.
19. Hollander, C. D., & Wu, A. S. (2011). The current state of normative agent-based systems. *Journal of Artificial Societies and Social Simulation*, 14, 2.
20. Hu, S., & Leung, H. (2017). Local norm phenomena in multi-agent systems under community networks. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, São Paulo, Brazil, May 8–12*, (pp. 1570–1572).

21. Hu, S., & Leung, H. (2018). Do social norms emerge? the evolution of agents' decisions with the awareness of social values under iterated prisoner's dilemma. In: *2018 IEEE 12th International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, (pp. 11–19).
22. Hu, S., & Leung, H. F. (2017). Achieving coordination in multi-agent systems by stable local conventions under community networks. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI'17*, (pp. 4731–4737). AAAI Press.
23. Ji, Z., Kiani, A. K., Qin, Z., & Ahmad, R. (2020). Power optimization in device-to-device communications: A deep reinforcement learning approach with dynamic reward. *IEEE Wireless Communications Letters*. <https://doi.org/10.1109/LWC.2020.3035898>.
24. Kittock, J. (1995). Emergent conventions and the structure of multi-agent systems. In: *Lectures in Complex systems: Proceedings of the 1993 Complex systems summer school, Santa Fe Institute Studies in the Sciences of Complexity Lecture Volume VI, Santa Fe Institute*, (pp. 507–521). Addison-Wesley.
25. Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
26. Li, D., Zhao, D., Zhang, Q., & Chen, Y. (2019). Reinforcement learning and deep learning based lateral control for autonomous driving [application notes]. *IEEE Computational Intelligence Magazine*, 14(2), 83–98. <https://doi.org/10.1109/MCI.2019.2901089>
27. Mahmoud, S., Griffiths, N., Keppens, J., & Luck, M. (2017). Establishing norms with metanorms over interaction topologies. *Autonomous Agents and Multi-Agent Systems*, 31, 1344–1376.
28. Marchant, J., Griffiths, N., & Leeke, M. (2015). Convention emergence and influence in dynamic topologies. In: *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '15*, (pp. 1785–1786). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
29. Molnar, L., & Veres, S. M. (2009). System verification of autonomous underwater vehicles by model checking. In: *Proceeding of the Oceans 2009 - Europe Conference*, (pp. 1–10). IEEE. <https://doi.org/10.1109/OCEANSE.2009.5278284>.
30. Morales, J., Lopez-Sanchez, M., Rodriguez-Aguilar, J. A., Wooldridge, M., & Vasconcelos, W. (2013). Automated synthesis of normative systems. In: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, (pp. 483–490).
31. Morales, J., Lopez-Sanchez, M., Rodriguez-Aguilar, J. A., Wooldridge, M., & Vasconcelos, W. (2014). Minimality and simplicity in the on-line automated synthesis of normative systems. In: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, (pp. 109–116).
32. Morales, J., López-Sánchez, M., Rodríguez-Aguilar, J. A., Wooldridge, M., & Vasconcelos, W. (2015). Synthesising liberal normative systems. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, (pp. 433–441).
33. Morales, J., Wooldridge, M., Rodríguez-Aguilar, J. A., & Lopez-Sanchez, M. (2018). Off-line synthesis of autonomously stable normative systems. *Autonomous Agents and Multi-Agent Systems*, 32, 635–671.
34. Morris-Martin, A., De Vos, M., & Padget, J. (2019). Norm emergence in multiagent systems: A viewpoint paper. *Autonomous Agents and Multi-Agent Systems*, 33(6), 706–749. <https://doi.org/10.1007/s10458-019-09422-0>.
35. Mukherjee, P., & Sen, S. (2007). Norm emergence in spatially constrained interactions. In: *Working Notes of the Adaptive and Learning Agents Workshop at AAMAS*, (vol. 7).
36. Mukherjee, P., Sen, S., & Airiau, S. (2008). Norm emergence under constrained interactions in diverse societies. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2, AAMAS '08*, (pp. 779–786). International Foundation for Autonomous Agents and Multiagent Systems.
37. NetworkX: Graph Generators: random_partition_graph. https://networkx.org/documentation/stable/reference/generated/networkx.generators.community.random_partition_graph.html (2021). Access Date: 2022-02-04.
38. Nordfjærn, T., Jørgensen, H., & Rundmo, T. (2010). An investigation of driver attitudes and behaviour in rural and urban areas in Norway. *Safety Science*, 48(3), 348–356.
39. Omidshafiei, S., Pazis, J., Amato, C., How, J. P., & Vian, J. (2017). Deep decentralized multi-task multi-agent reinforcement learning under partial observability. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70, ICML'17*, (pp. 2681–2690). JMLR.org.
40. Pan, Y., Cheng, C. A., Saigol, K., Lee, K., Yan, X., Theodorou, E. A., & Boots, B. (2020). Imitation learning for agile autonomous driving. *The International Journal of Robotics Research*, 39(2–3), 286–302. <https://doi.org/10.1177/0278364919880273>
41. Poole, D. L., & Mackworth, A. K. (2017). *Artificial intelligence: Foundations of computational agents* (2nd ed.). New York: Cambridge University Press.

42. Pujol, J. M., Delgado, J., Sangüesa, R., & Flache, A. (2005). The role of clustering on the emergence of efficient social conventions. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05*, (pp. 965–970). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
43. Ramchurn, S. D., Huynh, T. D., Wu, F., Ikuno, Y., Flann, J., Moreau, L., et al. (2016). A disaster response system based on human-agent collectives. *Journal of Artificial Intelligence Research*, 57(1), 661–708. <https://doi.org/10.1613/jair.5098>.
44. Ritchie, O., Watson, D., Griffiths, N., Xu, Z., & Mouzakitis, A. (2021). Influence of traffic context and information presentation on evaluation of autonomous highway journeys. *Accident Analysis and Prevention*, 161, 106385.
45. Salazar, N., Rodriguez-Aguilar, J. A., & Arcos, J. L. (2010). Robust coordination in large convention spaces. *AI Communications*, 23, 357–371.
46. Savarimuthu, B. T. R., Cranefield, S., Purvis, M., & Purvis, M. (2007). Norm emergence in agent societies formed by dynamically changing networks. In *2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'07)*, (pp. 464–470). doi: <https://doi.org/10.1109/IAT.2007.57>.
47. Sen, O., & Sen, S. (2010). Effects of social network topology and options on norm emergence. In J. Padget, A. Artikis, W. Vasconcelos, K. Stathis, V. T. da Silva, E. Matson, & A. Polleres (Eds.), *Coordination, organizations, institutions and norms in agent systems V* (pp. 211–222). Berlin Heidelberg: Springer.
48. Sen, S., & Airiau, S. (2007). Emergence of norms through social learning. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence, IJCAI'07*, (pp. 1507–1512). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
49. Shoham, Y., & Tennenholtz, M. (1992). Emergent conventions in multi-agent systems: Initial experimental results and observations (preliminary report). In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning, KR'92*, (pp. 225–231). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
50. Shoham, Y., & Tennenholtz, M. (1997). On the emergence of social conventions: Modeling, analysis, and simulations. *Artificial Intelligence*, 94(1), 139–166. [https://doi.org/10.1016/S0004-3702\(97\)00028-3](https://doi.org/10.1016/S0004-3702(97)00028-3) (Economic Principles of Multi-Agent Systems).
51. Sugawara, T. (2011). Emergence and stability of social conventions in conflict situations. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One, IJCAI'11*, (pp. 371–378). AAAI Press.
52. Villatoro, D., Malone, N., & Sen, S. (2009). Effects of interaction history and network topology on rate of convention emergence. In *3rd International Workshop on Emergent Intelligence on Networked Agents (WEIN'09)*, (pp. 13–19).
53. Villatoro, D., Sabater-Mir, J., & Sen, S. (2011). Social instruments for robust convention emergence. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One, IJCAI'11*, (pp. 420–425). AAAI Press. doi: <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-078>.
54. Villatoro, D., Sen, S., & Sabater-Mir, J. (2009). Topology and memory effect on convention emergence. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2009)*. Milan, Italy.
55. Watkins, C. J. C. H. (1989). *Learning from delayed rewards*. Ph.D. thesis, King's College, University of Cambridge, England.
56. Yang, T., Meng, Z., Hao, J., Sen, S., & Yu, C. (2016). Accelerating norm emergence through hierarchical heuristic learning. In *In proceedings of the 22nd European Conference on Artificial Intelligence*, (pp. 1344–1352).
57. Yu, C., Zhang, M., Ren, F., & Luo, X. (2013). Emergence of social norms through collective learning in networked agent societies. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS '13*, (pp. 475–482). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.