Emergence of Norms in Interactions with Complex Rewards

JAAMAS Track

Dhaminda B. Abeywickrama
Department of Computer Science, University of Bristol, Bristol, United Kingdom
dhaminda.abeywickrama@bristol.ac.uk

Nathan Griffiths
Department of Computer Science, University of Warwick, Coventry, United Kingdom
nathan.griffiths@warwick.ac.uk

Zhou Xu
Jaguar Land Rover, Coventry, United Kingdom
xzhou1@jaguarlandrover.com

Alex Mouzakitis
Jaguar Land Rover, Coventry, United Kingdom
amouzak1@jaguarlandrover.com

ABSTRACT

Autonomous systems are becoming pervasive, and as they become applied to highly dynamic and heterogeneous environments there is a need to model and understand more complex and nuanced agent interactions than have been previously studied. This paper proposes an agent-based modelling approach, based on norm emergence, to investigate such interactions. While there is typically an ideal set of compatible actions which lead to an optimal norm, in complex environments there may also be combinations that are *compatible* and yield positive (but not optimal) rewards. We illustrate our model of such scenarios using the case of an autonomous vehicle performing a manoeuvre at a T-intersection.

KEYWORDS

Agent Interactions; Norm Emergence; Reinforcement Learning

ACM Reference Format:
Dhaminda B. Abeywickrama, Nathan Griffiths, Zhou Xu, and Alex Mouzakitis. 2023. Emergence of Norms in Interactions with Complex Rewards: JAAMAS Track. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Autonomous systems are becoming increasingly prevalent, and are making an impact in safety-critical systems such as autonomous cars [1] and unmanned aerial vehicles [9]. Such highly dynamic and heterogeneous environments can result in complex agent behaviour, and interactions that contain learning, adaptation, and feedback loops. Norms are standards of behaviour that are agreed or expected by a set of individuals [4], and can help manage such complexity by imposing obligations to act or to not act in a particular way, where violation incurs the risk of punishment. Alternatively, norms (or *conventions*) can be considered simply as expected forms of behaviour without punishment for violation [6], as is the interpretation in this paper. There are three stages in the norms life cycle: creation, propagation, and emergence [8], where emergence implies that part of the population has adopted the norm.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Norm emergence in multi-agent systems is widely studied and has been used to model, for example, agents learning the rules of the road via coordination games and social dilemma games [3, 10]. However, existing studies have focused on agent interactions in which the combination of actions leads to a binary distinction between those that are compatible or incompatible. Therefore, the reward space is relatively small (e.g., in coordination games and the Prisoner's Dilemma). There is little consideration of the emergence of norms in which the agent interactions are complex and nuanced, such that there may be a range of compatible actions of varying desirability, which is the focus of this paper.

The main contribution of this work, presented in detail in the full paper [2], is a novel agent-based model of norm emergence in which agent interactions are more nuanced than those studied previously. To this end, while there is an ideal set of compatible actions which lead to an *optimal* norm, there are also combinations that are *compatible* and have positive rewards which are lower than the optimal. A complex reward function is introduced with clauses corresponding to the different cases the agents can encounter. Our model also differentiates between the notions of the *role* and *state* of an agent. We identify the conditions under which globally compatible norms are likely to emerge in an environment containing nuanced interactions. For this, simulations are performed to characterise 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.

2 THE MODEL

We assume a set of agents A (human-driven and autonomous vehicles), agent types T, and roles R, along with a state space S, an action space P, and reward matrices G. Each agent is associated with a role, which in our case study corresponds to the roles of a vehicle travelling on a main road or turning left at an T-intersection. In addition, we characterise the *type* of an agent (e.g., *assertive* or *non-assertive* depending on whether the agent typically interacts in a city or rural location). The *state* describes both the *role* played by the agent and any *environmental factors*. Agents choose actions (e.g., *go* or *yield*) in pairwise interactions, the combination of which represents the possible *norms* that can emerge. In our model, we categorise norms into two types, namely *optimal norms* which give maximal possible rewards to the agents and *compatible norms* which return positive rewards, but not as high as those provided by optimal

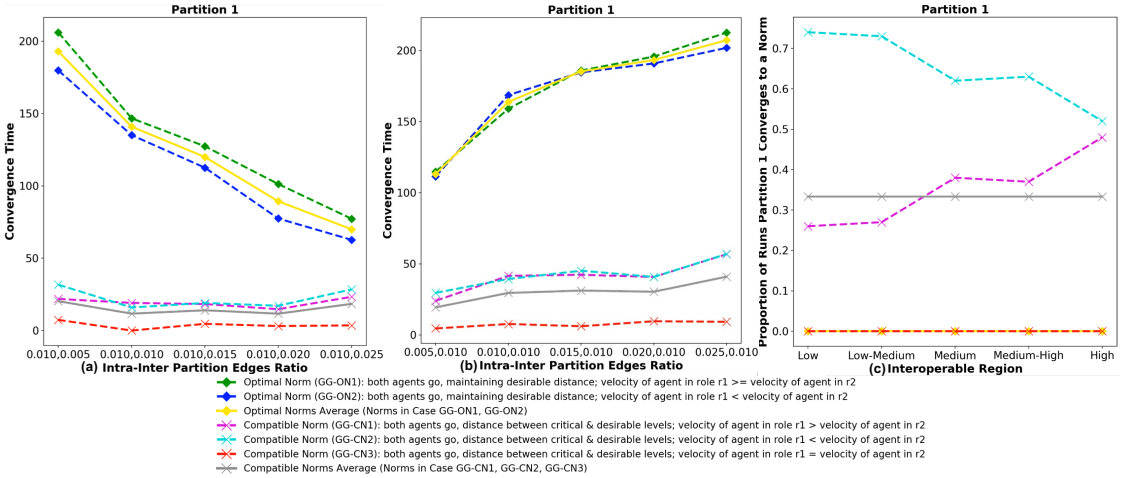

Figure 2: Effect of the number of (a) inter-partition and (b) intra-partition edges on convergence time, and (c) the size of the interoperable region on the probability of a partition converging to a global state.

norms. Finally, a reward function for the two agents participating in an interaction is defined as a function of the agents' states and actions, with clauses corresponding to the different situations that might occur.

As a case study, consider an autonomous vehicle performing a left-turn (role r_1) at a T-intersection in front of another vehicle (role r_2), as shown in Figure 1. In this scenario, the distance between the vehicles, their speeds, and their desirable and critical distance thresholds (to the car in front) determine the reward each vehicle receives. Details of the different possible outcomes and resulting rewards, given the different action choices and states, can be found in the full paper [2].

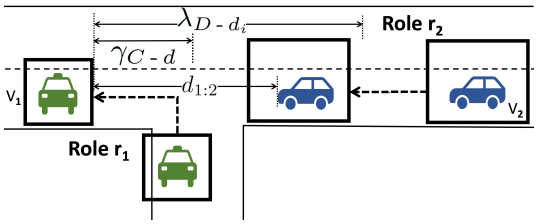

Figure 1: The agent in role r_1 pulling out between the desirable and critical threshold values for distance.

3 EXPERIMENT RESULTS AND CONCLUSION

Our experiments model populations of $m \in \{2000, 5000\}$ agents who are allocated to $n \in \{2, 5\}$ partitions, with 1000 agents in each partition [2]. Agents use Q-learning [11] and are situated in a random partition graph topology [5]. We conducted four main experiments to: (i) compare the effects of the probability of fewer inter-partition edges on convergence in different agent communities which can be non-mixed or mixed types, using different proportions of assertive and non-assertive agents; (ii) show the influence of intra-partition edges on norm emergence; (iii) show how the relative size of the interoperable region can influence norm emergence; and (iv) measure the overall strength of local conventions in a system in terms of the conformity and diversity of norms [7]. We analysed the experimental results using three metrics, namely frequency of convergence, convergence time, and the proportion of runs that converge to a norm. For example, Figure 2a shows that the system converges more quickly when the inter-partition edges are increased (the convergence time of both optimal norms, in green and blue, reduces). In contrast, Figure 2b shows that the system takes longer to converge when the intra-partition edges are increased. Meanwhile, Figure 2c shows that the probability of a partition converging to a global norm increases with the relative size of the interoperable region (i.e., the gap between the two compatible norms narrows).

This work proposed a novel agent-based modelling approach for norm emergence that considers compatible as well as optimal norms. Several conclusions can be drawn from our experimental results: (i) if there are few inter-partition edges and communities are well mixed, then the probability of the same norm emerging in the two partitions is greater, compared to the case where the communities are not well mixed; (ii) the frequency of convergence of a norm decreases when the relative number of inter-partition edges is increased or when the interoperable region is widened, while 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; and (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.

ACKNOWLEDGMENTS

This work was supported by Jaguar Land Rover and the UK-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 Grant EP/V026518/1.

REFERENCES

[1] Dhaminda B. Abeywickrama, Nicola Bicocchi, Marco Mamei, and Franco Zambonelli. 2020. The SOTA Approach to Engineering Collective Adaptive Systems. *International Journal on Software Tools for Technology Transfer* 22, 4 (aug 2020), 399–415. https://doi.org/10.1007/s10009-020-00554-3

[2] Dhaminda B. Abeywickrama, Nathan Griffiths, Zhou Xu, and Alex Mouzakitis. 2023. Emergence of Norms in Interactions with Complex Rewards. *Autonomous Agents and Multi-Agent Systems* 37, 2 (2023). https://doi.org/10.1007/s10458-022-09585-3

[3] Stéphane Airiau, Sandip Sen, and Daniel Villatoro. 2014. Emergence of Conventions through Social Learning. *Autonomous Agents and Multi-Agent Systems* 28, 5 (2014), 779–804. https://doi.org/10.1007/s10458-013-9237-x

[4] Henry Franks, Nathan Griffiths, and Sarabjot Singh Anand. 2013. Learning Influence in Complex Social Networks. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems* (St. Paul, MN, USA) *(AAMAS '13)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 447–454.

[5] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. 2008. Exploring Network Structure, Dynamics, and Function using NetworkX. In *Proceedings of the 7th Python in Science Conference*. 11–15.

[6] Chris Haynes, Michael Luck, Peter McBurney, Samhar Mahmoud, Tomas Vitek, and Simon Miles. 2017. Engineering the Emergence of Norms: A Review. *Knowledge Engineering Review* 32 (22 8 2017). https://doi.org/10.1017/S0269888917000169

[7] Shuyue Hu and Ho-Fung Leung. 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* (Melbourne, Australia) *(IJCAI'17)*. AAAI Press, 4731–4737.

[8] Andreasa Morris-Martin, Marina De Vos, and Julian Padget. 2019. Norm Emergence in Multiagent Systems: A Viewpoint Paper. *Autonomous Agents and Multi-Agent Systems* 33, 6 (01 Nov 2019), 706–749. https://doi.org/10.1007/s10458-019-09422-0

[9] Sarvapali D. Ramchurn, Trung Dong Huynh, Feng Wu, Yuki Ikuno, Jack Flann, Luc Moreau, Joel E. Fischer, Wenchao Jiang, Tom Rodden, Edwin Simpson, Steven Reece, Stephen Roberts, and Nicholas R. Jennings. 2016. A Disaster Response System Based on Human-agent Collectives. *Journal of Artificial Intelligence Research* 57, 1 (Sept. 2016), 661–708. https://doi.org/10.1613/jair.5098

[10] Sandip Sen and Stéphane Airiau. 2007. Emergence of Norms Through Social Learning. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (Hyderabad, India) *(IJCAI'07)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1507–1512.

[11] Christopher J. C. H. Watkins. 1989. *Learning from Delayed Rewards*. Ph.D. Dissertation. King's College, University of Cambridge, England.