Explanations In, Explanations Out: Human-in-the-Loop Social Navigation Learning

Jakob Karalus, Amar Halilovic, Felix Lindner
Institute of Artificial Intelligence
Ulm University
Ulm, Germany
{jakob.karalus,amar.halilovic,felix.lindner}@uni-ulm.de

***Abstract*—Social navigation is a desirable capacity of every mobile robot that operates in a human-populated environment. A core challenge is the need to account for a vast number of possible configurations of social spaces, interaction contexts, and individual preferences. We assert that a mobile robot should be able to explain its navigational choices to humans in terms of social aspects of a situation. This way, humans can challenge the robot's decisions and give corrective feedback. In our approach, explanations are first-class citizens employed both as inputs and as outputs. As inputs, explanations enhance human feedback ("Robot, you should not go this way, because . . . "). Our preliminary results indicate that allowing humans to formulate explanations as feedback can speed up training. As outputs, explanations make navigational decisions transparent to humans. This way, humans can verify that the learnt model incorporates the intended social norm. We show how explanation-generation methods known from explainable AI (XAI) community can be adopted for this task. We sketch the project in its early stage and point out planned research directions.**

***Index Terms*—robot navigation, explainability, human-in-the-loop reinforcement learning**

I. Introduction

As robots are expected to become a greater part of humans' daily life in the future, there is a challenge to bridge the gap between a robot's actions and humans' understanding of what a robot is doing and how it makes its decisions. For mobile robots, a crucial part of decision-making and planning is concerned with navigational choices. Human-robot spatial co-presence imposes known challenges on robot navigation. Consequently, quite some work has already been done on robot social navigation aimed at equipping navigation algorithms with the capacity to account for social norms of spatial behavior [1]–[3]. Social norms of spatial behavior are known to be highly situation-dependent varying across a large number of possible interactions and situations, e.g., [4]. Hence, there is a chance that a robot is unaware of some particular social norm and thus its spatial behavior is perceived as inappropriate by humans. We want to address such situations by making robot spatial behavior explainable. To this end, we apply methods from explainable AI (e.g., [5]) to social navigation. This way, a robot can explicate the reasons why it exhibits some spatial behavior, e.g., it could explain its increasing the distance to humans in terms of respect for their personal space. By giving the explanation, the robot opens the opportunity for humans to challenge the robot's decision by giving corrective feedback, e.g., stating that it is permissible to approach humans closer for initiating a hand-over interaction. From this explanatory feedback, the robot can learn about social norms and update its policy. From the perspective of a robot engineer, explanations can help to verify that a trained policy actually incorporates the intended social norms.

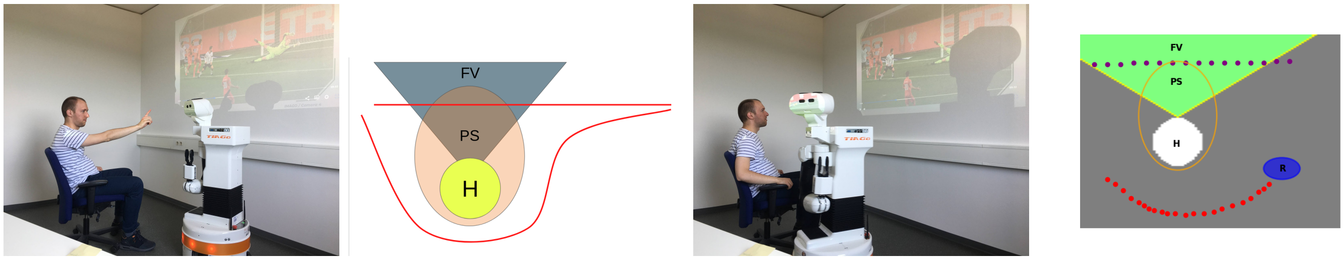
In the following, we briefly outline our research agenda which encompasses *explanation-informed human-in-the-loop reinforcement learning* and *explanation generation*. We comment on preliminary results and anticipated challenges.

II. Explanation-Informed Human-in-the-Loop Reinforcement Learning

The handling of social preferences is a highly challenging task due to its inherent complexity and vagueness. Therefore, learning these preferences with reinforcement learning is a natural approach. For example, Chen et al. [6] trained a social scenario (passing of humans) with reinforcement learning and a custom reward function. While it is feasible to provide a custom reward function for simple tasks, capturing the full complexity of social norms is infeasible. In fact, social robot navigation is an instance of the general value-alignment problem, and it inherits all of its challenges, cf., [7]. Human-in-the-loop reinforcement learning (HRL) is a possible solution. HRL enables humans to directly influence and shape the learning process of robots. Unlike writing reward functions for social preferences, giving feedback is relatively natural for humans.

Our system employs methods from HRL, like TAMER [8], to enable the human to interactively give feedback towards a robot. For example, in figure 1d the robot wants to move from the right to the left. Initially, its preferred route would go straight through the field of vision (FV) of the trainer, who is watching a game. The user then explains that this behavior is undesired and that it should not move through his field of vision. Internally, the robots learns with this explanatory feedback which of the possible features is most responsible for the feedback and updates its policy accordingly. In the example, the influential feature could be either the field of vision (FV) or the personal space (PS). The explanation clarifies towards the FV, and thus the updated policy makes a detour necessary (figure 1c).

Human feedback is costly. Therefore, it should be provided through a convenient interface and used as efficiently as



(a) Explanatory Feedback (b) Learning (c) Detour (d) Explanation

Fig. 1: Envisioned full lifecycle: 1a) A person is watching a soccer game, while a robot (R) is about to move through the person's cone of vision. The person gives the robot feedback that it should not move through the field of vision because he is watching the game. 1b) This explanatory feedback is then used by the HRL algorithm to learn a suitable policy respecting the person's needs. Here, the robot updates its policy to account for the human (H), its personal space (PS), and the field of view (FV). 1c) The robot then uses the updated policy to create a different path that respects the learnt social norm. 1d) Finally, the trainer can inspect the learnt policy of the robot with the help of XAI methods like LIME. Here, it can be verified that the robot's behaviour is based on the field of vision (and not on the person itself or his personal space).

possible. Including explanations in the learning process can mitigate both challenges. First, it has been shown that humans rarely just want to provide binary feedback (see [9]), so the inclusion of explanations should reduce frustrating experiences. One research challenge is how to design interfaces that allow humans to give explanatory feedback. Secondly, in preliminary work (see [10]), we have shown that the inclusion of counterfactual explanations speeds up the convergence of the learning process. This in turn results in a more efficient usage of human feedback. Further work will extend the system by other types of explanations and by more dynamic interfaces for their delivery.

III. Explainable Navigation

We explore the possibility to make robot navigation more understandable to humans using Explainable Artificial Intelligence (XAI) methods for explaining robot navigation algorithms. Particularly, black-box XAI methods can be used to explain decisions of both traditional and RL-based navigation algorithms. Figure 1d exemplifies such an explanation of an individual navigational decision of the robot. Using LIME [5], a region in the local cost map is highlighted as being most important for the navigational action (the local path) of the robot. As expected, in the depicted situation, the field of view by the person watching soccer explains the local path (shown as red dots): Without the field of view, the robot would choose a different local plan, viz., going straight (shown as purple dots). The role of the explanation module is two-fold: First, it allows humans to verify that the robot is making its navigational choices for the right reasons. This may increase understanding and trust. Such a scenario is outlined in Fig. 1. Second, the explanation module allows humans to detect situations in which corrective feedback is necessary, because some learnt social norm is not or incorrectly represented. Main challenges include automating the generation of explanations that are actually understood by humans, making explanation methods functioning in real-time, and designing multi-modal interfaces for communicating explanations to humans using natural communication modalities.

Conclusion and Future Work

We envision a system that learns social norms of spatial behavior naturally with the help of explanations. Explanations should be a crucial part of a system that is designed to work in human spaces. Our first results show that it is possible to include them as first-class citizens both as inputs for interactive learning, and as outputs for enabling humans to understand the robot behavior. Future work will expand on these results and include more modes of explanations and focus on full integration of all parts in a realistic benchmark.

References

[1] F. Lindner and C. Eschenbach, "Towards a formalization of social spaces for socially aware robots," in *Spatial Inform. Theory*, 2011, pp. 283–303.

[2] T. Kruse, A. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics Auton. Syst.*, vol. 61, pp. 1726–1743, 2013.

[3] F. Lindner, "A conceptual model of personal space for human-aware robot activity placement," *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5770–5775, 2015.

[4] E. Hall, *The Hidden Dimension, Man's Use of Space in Public and Private*. Bodley Head, London, England, 1966.

[5] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[6] Y. Chen, M. Everett, M. Liu, and J. How, "Socially aware motion planning with deep reinforcement learning," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1343–1350, 2017.

[7] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in AI safety," *ArXiv*, vol. abs/1606.06565, 2016.

[8] W. Knox and P. Stone, "TAMER: Training an agent manually via evaluative reinforcement," *2008 7th IEEE International Conference on Development and Learning*, pp. 292–297, 2008.

[9] S. Krening and K. M. Feigh, "Interaction algorithm effect on human experience with reinforcement learning," *J. Hum.-Robot Interact.*, vol. 7, no. 2, Oct. 2018.

[10] J. Karalus and F. Lindner, "Accelerating the convergence of human-in-the-loop reinforcement learningwith counterfactual explanations," 3^{rd} *ICML Human-in-the-loop learning workshop (ICML HILL)*, 2021.