On the Importance of Environments in Human-Robot Coordination

Matthew C. Fontaine*, Ya-Chuan Hsu*, Yulun Zhang*, Bryon Tjanaka, Stefanos Nikolaidis,
Department of Computer Science, University of Southern California, Los Angeles, CA, USA
{mfontain, yachuanh, yulunzha, tjanaka, nikolaid}@usc.edu

***Abstract*—When studying human-robot collaboration, people focus on improving robot policies to create fluent coordination with human teammates. However, the effect the environment has on human-robot interaction is often overlooked, further limiting enhancement in robot policies to accommodate environments. To thoroughly explore environments that result in diverse behaviors, we propose a framework for procedural generation of environments that are (1) stylistically similar to human-authored environments, (2) guaranteed to be solvable by the human-robot team, and (3) diverse with respect to coordination measures. We analyze the procedurally generated environments in the Overcooked benchmark domain via simulation and an online user study. Results show that the environments result in qualitatively different emerging behaviors and statistically significant differences in collaborative fluency metrics, even when the robot runs the same planning algorithm.**

I. Introduction

In this work, we focus on human-robot coordination in a kitchen environment. A robot and a human team up to cook and serve dishes. An important aspect of collaboration is the workload distribution between teammates. Human factors research has shown that too light or too heavy workload can affect human performance and situational awareness [6]. The perceived robot's contribution to the team is a crucial metric of fluency [4], and human-robot teaming experiments found that the degree to which participants were occupied affected their subjective assessment of the robot as a teammate [3].

Our key insight is that *changing the environment can result in significantly different coordination actions* despite running the same coordination algorithm. We propose a framework for generating environments that induces diverse human-robot interaction. The domain we apply the framework on is the increasingly popular Overcooked domain [1] for researching in the coordination of agent behaviors.

II. Approach

The proposed framework consists of three main components: 1) A generative adversarial network (GAN) that generates human-design alike environments. 2) A mixed-integer linear programming (MIP) repair procedure to ensure generated environments are playable. 3) A quality-diversity algorithm that searches the latent space of GAN to explore diverse environments that induce various human-robot collaboration behaviors. We evaluate our approach through an extensive experiment where we simulate and examine human-robot interaction and task completeness in different environments.

*Equal contribution

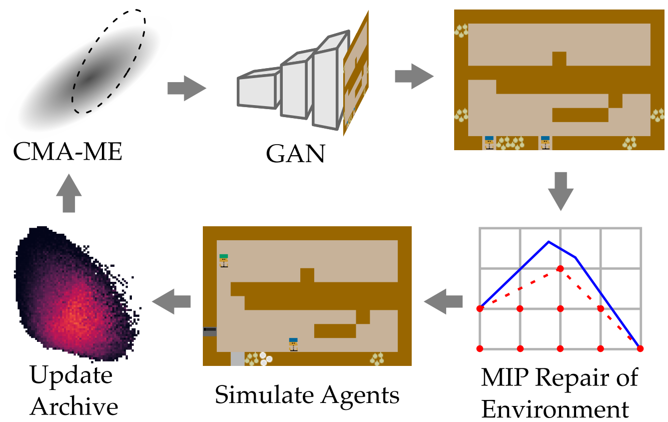

Fig. 1: An overview of the framework for procedurally generating environments that are stylistically similar to human-authored environments. Our environment generation pipeline enables the efficient exploration of the space of possible environments to procedurally discover environments that differ based on provided metric functions.

III. Experiments

We performed four experiments to demonstrate that our proposed framework generates a variety of environments that result in a diverse set of coordination behaviors. In the following paragraph, we focus on presenting an experiment that explores environments that captures failure cases caused by designed robot policies. The remaining experiments are documented in the full paper.

In this experiment, the robot executes a QMDP policy, which chooses actions that maximizes performance while also considering the actions taken by the human. The human executes a myopic policy, which myopically selects the highest priority task based on the current dish prepping progress. We generate environments that *minimize* the performance metric, which is useful for searching failure cases of developed algorithms [2]. We are specifically interested in drops in performance that arise from the assumptions of the QMDP formulation, rather than, for example, poor performance because objects are too far from each other. Therefore, we use a robot executing an MDP policy that fully observes the human subtask as a baseline performance of human-robot interaction. We *maximize the difference* in performance between simulations with the MDP policy and the QMDP policy.

Fig. 2 shows the generated archive: we illustrate the 3D behavior space as a series of five 2D spaces, one for each value of the difference in orders. Each colored cell represents an environment with workload distribution computed by simulating the two agents in that environment. Lighter colors indicate *lower* performance of the team of the QMDP robot and the myopic human compared to an MDP robot and a myopic human. We are particularly interested in the environments where the team fails to complete the task.

In environment (1) of Fig. 2, the simulated human picks up an onion at the same time step the robot delivers the third onion to the pot. There is now no empty pot to deliver the onion, so the human defaults to going to the pot and waiting there, blocking the path of the robot. The environment leads to an edge case that was not accounted for in the hand-designed human model but revealed by attempting to minimize the performance of the agents.

In environment (2) of Fig. 2, the two agents get stuck in the narrow corridor in front of the rightmost onion dispenser. Due to the "auto-unstuck" mechanism, the simulated human randomly picks an action and goes backward towards the onion dispenser. The QMDP planner, which uses the change of distance to the subtask goal location as observation (see appendix in full paper), erroneously infers the human subtask is to reach the onion dispenser, and does not move backwards to allow the human to go to the dish dispenser. This environment highlights a limitation of the distance-based observation function in the robot policy design since it is not robust to random motions that occur when the two agents get stuck.

Overall, *we observe that when minimizing performance, the generated environments reveal edge cases that can help a designer better understand, debug, and improve the agent models.*

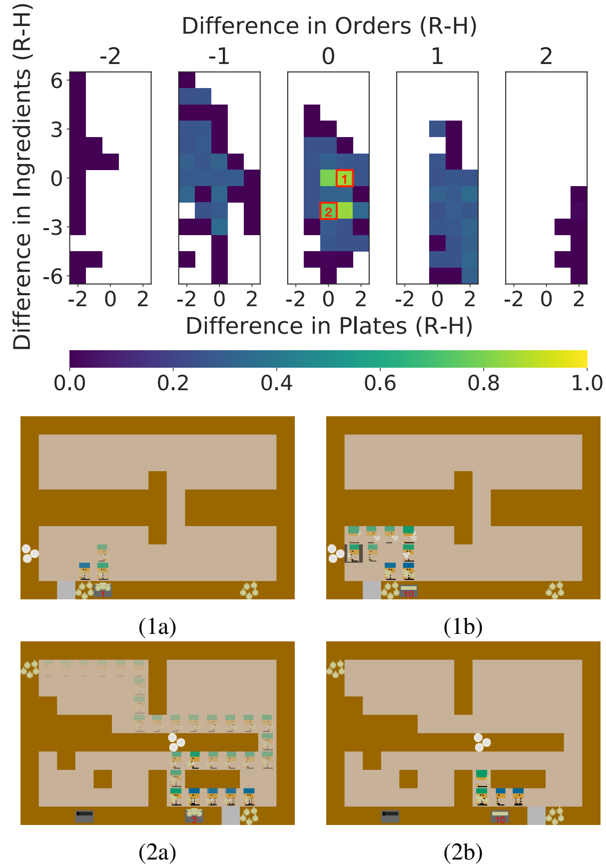

Fig. 2: Archive of environments aiming to *minimize* performance of a QMDP robot (green agent) and a simulated myopic human (blue agent). Lighter color indicates lower performance. (1a) and (1b) show successive frame sequences for environment (1), and so as (2a), (2b) for environment (2).

IV. Implications

We envision our framework as a method to assist human-robot interaction (HRI) planning in the future. The framework is capable to facilitate the understanding of complex human-aware algorithms or other adaptive agents [5] in complex task settings [7]. We are excited about future work that highlights diverse behaviors in different settings where coordination is essential, such as manufacturing and assistive care. Finally, we hope that our work will guide future HRI planning research to consider the environment as a significant factor in coordination problems.

References

[1] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. On the utility of learning about humans for human-ai coordination. In *Advances in Neural Information Processing Systems*, pages 5174–5185, 2019.

[2] Matthew Fontaine and Stefanos Nikolaidis. A quality diversity approach to automatically generating human-robot interaction scenarios in shared autonomy. *arXiv preprint arXiv:2012.04283*, 2020.

[3] Matthew Gombolay, Anna Bair, Cindy Huang, and Julie Shah. Computational design of mixed-initiative human–robot teaming that considers human factors: situational awareness, workload, and workflow preferences. *The International journal of robotics research*, 36(5-7):597–617, 2017.

[4] Guy Hoffman. Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems*, 49(3):209–218, 2019.

[5] Hengyuan Hu, Adam Lerer, Alex Peysakhovich, and Jakob Foerster. "other-play" for zero-shot coordination. In *International Conference on Machine Learning*, pages 4399–4410. PMLR, 2020.

[6] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of cognitive engineering and decision making*, 2(2):140–160, 2008.

[7] Tonghan Wang, Tarun Gupta, Anuj Mahajan, Bei Peng, Shimon Whiteson, and Chongjie Zhang. Rode: Learning roles to decompose multi-agent tasks. *arXiv preprint arXiv:2010.01523*, 2020.