Improving Human-Robot Interaction via a Population of Synthetic Human-like Teams

Siddharth Srikanth*, Varun Bhatt*, Charles Michael Lewis‡, Katia P. Sycara†, Aaquib Tabrez*, Stefanos Nikolaidis*
*University of Southern California †Carnegie Mellon University ‡University of Pittsburgh

***Abstract*—Understanding how humans collaborate and communicate is crucial for advancing human-robot teaming and decision support systems. However, collecting exhaustive human data and evaluating human-robot teaming algorithms through user studies is often impractical, underscoring the need for synthetic models of diverse human behavior. In our prior work, we introduced PLAN-QD, a novel framework that combines Quality Diversity (QD) optimization with LLM-powered agents to algorithmically generate diverse teaming and communication behaviors in collaborative tasks. We validated PLAN-QD through a user study ($n = 54$ participants) and showed that it effectively replicates trends from human teaming data while also capturing behaviors that are difficult to observe without large-scale data collection. In this work, we summarize the PLAN-QD framework and key findings, and propose its extensions to human-robot teaming: (1) generating synthetic human data for robot learning, (2) incorporating diverse personalities to support personalization, and (3) studying human-like communication and negotiation. Our source code is available at github.com/icaros-usc/plan-qd.**

I. Introduction and Motivation

Robots and autonomous systems deployed in the real world must collaborate with and adapt to humans who exhibit diverse behaviors, expectations, and communication styles. For example, consider a robot that assists chefs in a restaurant kitchen. Different chefs may want different robot behavior (e.g., only chop vegetables) or may require it to listen to different commands. For robots to adapt to such varied teams, they would need to *build an understanding of how different humans might operate when performing these tasks*.

One approach to generating such teaming behaviors is through learning models of large-scale human data [7, 33]. However, collecting a sufficiently large and diverse dataset from collaborative domains, especially ones involving multiple interacting humans, is expensive and challenging [36]. On the other hand, prior work has shown large language model (LLM)-powered agents to be a viable option for modeling human behavior [50, 21, 47, 48]. When prompted with personalities or strategies to bias their actions, LLMs are shown to exhibit human-like behavior in social domains [32].

In our prior work, we proposed **PLAN-QD** (**P**rompting **L**LM-powered **A**gents for **N**ovel Behavior via **Q**uality **D**iversity) [40], a framework that uses Quality Diversity (QD) optimization to algorithmically generate prompts that elicit diverse human-like teaming and communication behaviors in LLM-powered agents. Unlike manual prompt engineering, PLAN-QD iteratively discovers high-performing and behaviorally diverse agents by searching over user-defined diversity axes (e.g., workload distribution), referred to as measure functions [35]. Prompts found during the search act as *stepping stones* to find new prompts for agents with better performance or novel behavior.

In this paper, we summarize the PLAN-QD framework (Sec. II), its key empirical findings (Sec. III), and explore its potential applications in human-robot teaming (Sec. IV). First, we describe how the human-like agents generated by PLAN-QD can guide robot learning, improving its adaptability to different users when deployed. Second, we propose ways to include personality models into PLAN-QD's population generation, allowing robots to learn personalized strategies for different human personalities. Finally, we present some alternative communication methods, such as negotiation, and discuss their integration into PLAN-QD.

II. Summary of the PLAN-QD Framework

The key insight of PLAN-QD is that *QD optimization can be used to algorithmically generate prompts that elicit human-like teaming diversity in LLM-powered agents*, leveraging the idea that prompts discovered during the search serve as *stepping stones*. Prior works use this principle to diversify agent behaviors [34, 31, 45, 3] and text generation or red teaming LLMs [5, 39, 26]. For PLAN-QD, we combined these ideas to diversify LLM-powered collaborative agents.

The framework consists of the following components: (1) LLM-powered agents to interface between an LLM and the environment, along with a communication setup for agents to pass messages, and (2) QD optimization to find prompts that elicit diverse behavior in LLM-powered agents. Fig. 2 shows the overview of the complete framework.

A. LLM-powered agents

We formulate the environment as a decentralized Markov Decision Process (dec-MDP [4]) $\langle \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma \rangle$ with N agents, where $\mathcal{S}$ is the state space, $\mathcal{A} = \Pi_i^N \mathcal{A}_i$ is the joint action space of all agents, $\mathcal{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the common reward function that all agents receive, $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is the transition function, and γ is the discount factor. The agents' goal is to maximize the discounted sum of rewards, $J = \Sigma_t \gamma^t r_t$, where r_t is the reward obtained at timestep t. For PLAN-QD's LLM-powered agents, the state and the actions are provided and received via a text interface (**purple arrows** in Fig. 2), with the space of textual inputs/outputs to the LLM.

At the beginning of an episode, PLAN-QD queries the LLMs for all agents in a random sequence. The **input** to the LLM

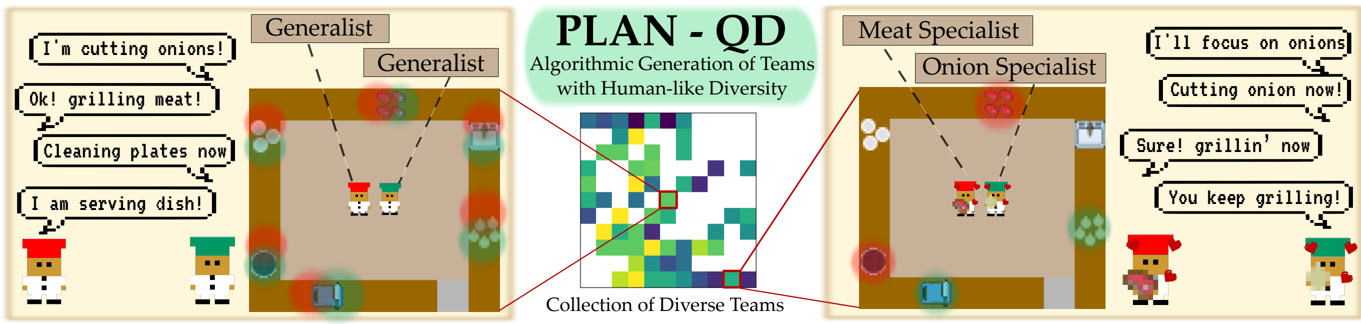


Fig. 1: PLAN-QD uses Quality Diversity (QD) optimization to generate a set of prompts to elicit human-like teaming diversity in LLM-powered agents. The resulting teams exhibit distinct collaboration strategies (e.g., meat specialist with onion specialist), enabling the systematic study of communication and coordination in complex environments.

provides textual context about the environment and its current state so that the LLM can make an informed choice about the next action to take. The LLM **outputs** a high-level action, which is converted into a sequence of low-level actions (e.g., "move left, up, and interact") by a motion planner, and an optional message. Once the corresponding agent completes a high-level action or fails after a timeout, PLAN-QD re-queries the agent's LLM for a new high-level action and a message.

B. QD optimization to generate prompts for diverse agents

To automate the process of finding personality prompts for diverse agents, PLAN-QD algorithmically searches for them using QD optimization (**green arrows** in Fig. 2). To begin, PLAN-QD maintains a prompt archive consisting of discretized cells, where each cell stores a list of high-quality personality prompts, one for each agent in the domain, found during the optimization. At the beginning of the optimization, an initial prompt is selected for all agents. The algorithm searches for new prompts that promote diversity along *measure axis* by querying a separate LLM (referred to as the *mutator LLM* to differentiate from LLM-powered agents) for new prompts.

The mutator LLM is provided with the prompt list and queried to mutate the prompt list in a random direction in the measure space (e.g., "more number of plates cleaned"). This process generates a new prompt list that is expected to induce behaviors aligned with that direction.

PLAN-QD evaluates this prompt list by simulating the corresponding LLM-powered agents in the environment, repeating each evaluation multiple times and taking the median of the resulting objective and measure values. The prompt list then replaces a cell in the archive if the cell is empty or if it achieves a higher objective value than the one currently stored.

Through this iterative process of ***prompt selection***, ***mutation***, ***evaluation***, and ***archive update***, PLAN-QD populates the archive with prompts that elicit high-quality and diverse behavior in LLM-powered agents.

III. Key Takeaways from PLAN-QD

We evaluated PLAN-QD on a collaborative domain, *Steakhouse* [16]. Inspired by the game Overcooked [1] and its simulation environment [7], Steakhouse introduces complex coordination challenges via larger and more varied layouts and multi-step recipes. Our experiments led to three key takeaways: (1) Humans exhibit diverse behaviors depending on communication and layout, (2) PLAN-QD's agents match communication trends from human data, and (3) PLAN-QD's agents are diverse.

A. Humans exhibit diverse behaviors depending on communication and layout

Experimental Setup. We conducted a 2×1 between-subjects user study with and without verbal communication, using four kitchen layouts to evaluate how spatial constraints and communication affect human coordination. We ran a study with 54 participants forming 27 teams of two, as part of IRB-approved study. We defined individual and team behavior via *subjective teaming measures* (e.g., trust, fluency, and workload [24, 14, 15, 38], from a 7-point Likert-scale questionnaire) and *teamwork measures* (fitness (objective), average action delay, percent contribution to the task, and specialization).

Results. We hypothesized that communication would affect subjective and teamwork measures in human teams. We discovered that while communication generally improved performance in asymmetric layouts, the impact was weaker in symmetric layouts where tasks could be executed more independently. We also observed that background knowledge and personality influenced teaming behavior with communication. Participants in the "without communication" condition validated these findings, reporting coordination difficulties and expressing a preference for some form of communication in their exit interviews.

B. PLAN-QD's agents match communication trends from human data

Experiment Setup. We generated a population of LLM agents with the PLAN-QD framework on each of the four Steakhouse layouts, and compared the aforementioned teamwork measures obtained by the LLM teams and the human teams from our data. The diversity axes in PLAN-QD were based on differences in sub-task completions (i.e., workload), a known indicator of team coordination behavior [13, 15, 11]

Results. We observed that behavioral trends exhibited by PLAN-QD's agents partially matched those observed in the user study between "with communication" and "without

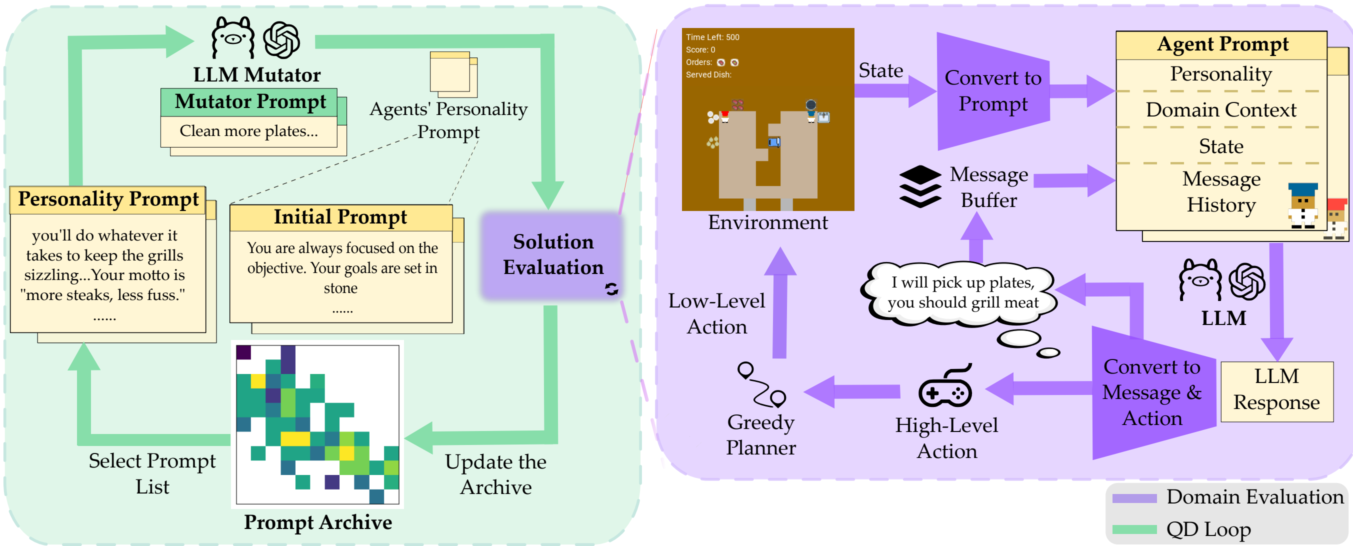


Fig. 2: Overview of the PLAN-QD framework, including the QD optimization (green arrows) and the LLM-powered agents. QD optimization repeatedly selects and mutates prompts to generate new prompts that are then evaluated in the environment (purple arrows). Only high-quality and diverse prompts are retained in the prompt archive.

communication" conditions. We observed, via a one-sample test of proportions, that 12 out of 16 layout–measure combinations matched human trends ($p = .038$).

C. *PLAN-QD's agents are diverse*

Experiment Setup. We tested PLAN-QD against a Random Mutation baseline, in which an LLM generated 100 prompts for diverse agents (the same number generated by PLAN-QD), but without an iterative process.

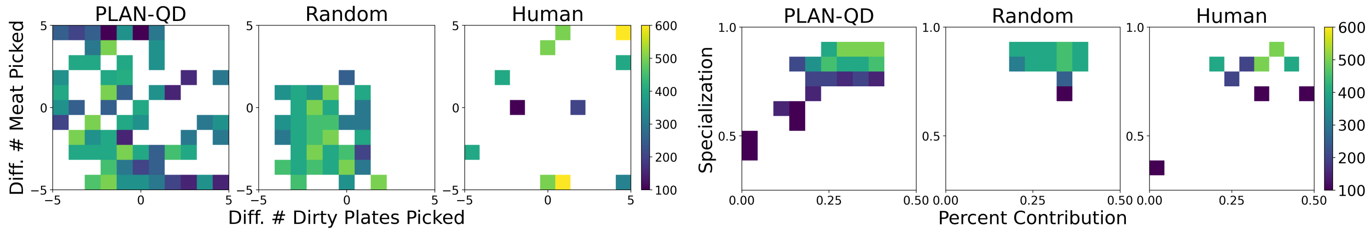
Results. We discovered that PLAN-QD achieves significantly higher coverage (number of cells covered in the prompt archive) than Random Mutation, based on a one-sample test of proportions ($p < .001$). Qualitatively, by plotting the heatmaps of the prompt archive (Fig. 3), we saw that PLAN-QD covers extreme behaviors (Fig. 3a), including those not found in our small-scale human data, highlighting its benefit in augmenting human datasets. Even when looking at measures that were not explicitly diversified by PLAN-QD (Fig. 3b), we still see a better coverage than Random Mutation, including certain rare behaviors such as low percent contribution and specialization, similar to those exhibited by human users who did not fully understand the rules of the game.

IV. Open Challenges and Future Approaches

Our findings demonstrate that humans exhibit diversity in teamwork strategies, influenced by spatial layout, communication, and individual differences such as personality and prior experience. PLAN-QD agents can replicate several of these behavioral trends and generate strategies that are difficult to observe without large data collection human user studies. In this section, we propose extensions to PLAN-QD toward more embodied, adaptive, and personalized applications in human-robot teaming: (1) *Robot Learning and Teaming*: enabling robots to train with a diverse set of synthetic human-like partners, (2) *Personality and Personalization*: incorporating psychologically grounded traits into agent behavior for simulating and adapting to varied teammate profiles, and (3) *Communication and Negotiation*: improving the communication framework to support richer, more human-like coordination and dialogue strategies. These directions aim to bridge the gap between synthetic agent simulation and training teaming capabilities in embodied autonomous systems.

A1: Robot Learning and Teaming. A major opportunity for future work is leveraging PLAN-QD to generate diverse, human-like populations in embodied robot simulators such as CALVIN [25] and RLBench [18], where a simulated human and robot (or human–human) share a physical workspace. Extending PLAN-QD to these domains would allow us to algorithmically explore variations in human behavior (e.g., role specialization) and communication style (e.g., ambiguity) within embodied shared tasks such as tabletop manipulation or tool-use scenarios. This offers a scalable alternative to collecting large-scale human–robot interaction data, which remains a crucial bottleneck for training and evaluating collaborative robotic agents [27].

PLAN-QD can generate a library of teaming trajectories that reflect a range of human strategies. The robot could then either learn to imitate these strategies [7] or learn the best collaborative strategies for each human strategy [41, 6, 29, 49]. This approach encourages the robot to generalize across a spectrum of teaming behaviors, rather than overfitting to a narrow set of collaborators. Additionally, the QD archive can support curriculum-based training, starting with cooperative or predictable teammates and progressively introducing more unpredictable or adversarial styles to build robustness. Such a progression could help the robot develop flexible collaboration strategies under controlled and scalable conditions. Finally, the PLAN-QD-based population could also serve as a test suite to thoroughly evaluate robotic systems, supporting safer and more reliable deployment in human-centric applications.



(a) Workload Measures on Open Layout **with communication**.

(b) Teamwork Measures on Forced Coordination Layout **without communication**.

Fig. 3: Example heatmaps of the archives resulting from human data and the agents generated by PLAN-QD and Random Mutation, colored by the corresponding fitness value. PLAN-QD generates agents covering a wider range of behavior compared to the Random Mutation baseline, including certain extremes observed in the human data.

A2: Personality and Personalization. Our user study showed that individual personality traits influenced team behavior and outcomes. For example, we observed that in the "without communication" condition, team performance was highly correlated with interpersonal trust, suggesting that some individuals were more effective at implicit coordination than others. These preliminary patterns suggest that individual personality traits can meaningfully shape collaborative behavior.

Recent work has shown that LLMs can be trained to reflect personality traits in human communication datasets [22]. PLAN-QD can extend this from text-based settings to embodied collaboration, generating synthetic agents with varied personality measures based on psychological models such as the Big Five [19]. For example, an extroverted agent might communicate more frequently or take initiative in ambiguous situations, while a neurotic agent may be more reactive or avoid uncertain actions. These traits could be incorporated as axes in the QD mutation space, guiding behavioral and communication variations during population generation. However, ensuring the actions of these embodied agents are aligned with their intended personalities remains an open challenge.

Another avenue for future work is exploring how personality-aware agents affect teaming outcomes when paired with real humans. As found in prior work [2, 28] and in our results, certain personality pairings may lead to better coordination, trust, or subjective benefits. For example, an extroverted robot could successfully pair users who prefer explicit communication instead of relying on implicit coordination. PLAN-QD could serve as a testbed for systematically studying these effects and training robots to adapt their behavior dynamically based on inferred human traits. Although personalization at this level remains an open area in HRI, PLAN-QD offers a scalable and controlled way to explore it.

A3: Communication and Negotiation. Prior work shows that communication timing, content, and modality significantly impact team performance, particularly in safety-critical or uncertain environments [20, 46, 23]. Yet, accounting for variation in communication strategies in human-robot teams remains challenging [44, 27]. While foundation models have enabled significant progress in language understanding and generation, deploying them on embodied agents is computationally expensive and not optimized for real-time, long-term human collaboration [17, 10].

PLAN-QD could enable the generation of human-like messages aligned with different coordination strategies or personality traits. One promising direction is to leverage this synthetic dialogue to train communication-aware robot policies that allow the robot to condition its policy on incoming teammate utterances [25]. Future work could also explore mechanisms that help the robot decide when to attend to language versus when to rely on other cues, particularly in scenarios involving unreliable or adversarial partners (e.g., agents that lie to or attempt to confuse their partners).

Beyond basic coordination, negotiation is a critical communication skill in collaborative interaction, where autonomous agents must often align on subgoals and resolve conflicts, particularly in dynamic, uncertain scenarios [43, 12, 37]. In human-robot teaming, enabling negotiation behaviors can facilitate shared mental models, flexible role assignment, and improved trust and robustness [8, 9]. The LLM-powered agents in PLAN-QD can be guided by frameworks from human factors (e.g., Joint Intention Theory) [42] and game-theoretic models of negotiation and cooperation [30], leading to diverse negotiation behaviors.

V. Conclusion

In our prior work, we introduced PLAN-QD, a framework that applies QD optimization to generate diverse, human-like teaming and communication behavior. Our results showed that PLAN-QD discovers diverse team behaviors, including certain extremes observed in human data, that replicate human-like effects of communication. As a step toward extending the framework for human-robot teaming, we outline three core directions for future work: (1) enabling robot learning from PLAN-QD's diverse population of synthetic partners, (2) incorporating personality models into PLAN-QD's team generation to support personalization, and (3) modeling communication timing and negotiation strategies that are more aligned with human teams. We are excited about the prospect of these extensions providing a foundation for training and evaluating adaptive, communication-aware, and trustworthy robot teammates.

Acknowledgments. This work was supported by NSF CAREER #2145077, NSF NRI #2024949, and DARPA EMHAT HR00112490409.

REFERENCES

[1] Overcooked 2, 2018. URL https://store.steampowered.com/app/728880/Overcooked_2/.

[2] Murray R Barrick, Greg L Stewart, Mitchell J Neubert, and Michael K Mount. Relating member ability and personality to work-team processes and team effectiveness. *Journal of applied psychology*, 83(3):377, 1998.

[3] Sumeet Batra, Bryon Tjanaka, Matthew Christopher Fontaine, Aleksei Petrenko, Stefanos Nikolaidis, and Gaurav S. Sukhatme. Proximal policy gradient arborescence for quality diversity reinforcement learning. In *Proceedings of the Twelfth International Conference on Learning Representations, ICLR*, 2024. URL https://openreview.net/forum?id=TFKIfhvdmZ.

[4] Daniel S. Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. The complexity of decentralized control of markov decision processes. *Math. Oper. Res.*, 27(4):819–840, 2002. doi: 10.1287/MOOR.27.4.819.297. URL https://doi.org/10.1287/moor.27.4.819.297.

[5] Herbie Bradley, Andrew Dai, Hannah Teufel, Jenny Zhang, Koen Oostermeijer, Marco Bellagente, Jeff Clune, Kenneth Stanley, Grégory Schott, and Joel Lehman. Quality-diversity through ai feedback. *arXiv preprint arXiv:2310.13032*, 2023.

[6] Rodrigo Canaan, Xianbo Gao, Julian Togelius, Andy Nealen, and Stefan Menzel. Generating and adapting to diverse ad hoc partners in hanabi. *IEEE Transactions on Games*, 15(2):228–241, 2023. doi: 10.1109/TG.2022.3169168.

[7] 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. *Advances in neural information processing systems*, 32, 2019.

[8] Christine T Chang, Mitchell Hebert, and Bradley Hayes. Collaborative planning and negotiation in human-robot teams. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 763–765, 2023.

[9] Christine T Chang, Maria P Stull, Breanne Crockett, Emily Jensen, Clare Lohrmann, Mitchell Hebert, and Bradley Hayes. Iteratively adding latent human knowledge within trajectory optimization specifications improves learning and task outcomes. *IEEE Robotics and Automation Letters*, 2024.

[10] Roya Firoozi, Johnathan Tucker, Stephen Tian, Anirudha Majumdar, Jiankai Sun, Weiyu Liu, Yuke Zhu, Shuran Song, Ashish Kapoor, Karol Hausman, et al. Foundation models in robotics: Applications, challenges, and the future. *The International Journal of Robotics Research*, pp. 02783649241281508, 2023.

[11] Matthew C Fontaine, Ya-Chuan Hsu, Yulun Zhang, Bryon Tjanaka, and Stefanos Nikolaidis. On the importance of environments in human-robot coordination. *arXiv preprint arXiv:2106.10853*, 2021.

[12] Matthew Gombolay, Ronald Wilcox, and Julie Shah. Fast scheduling of multi-robot teams with temporospatial constraints. 2013.

[13] Matthew C. Gombolay, Anna Bair, Cindy Huang, and Julie A. 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. doi: 10.1177/0278364916688255. URL https://doi.org/10.1177/0278364916688255.

[14] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pp. 139–183. Elsevier, 1988.

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

[16] Ya-Chuan Hsu, Michael Defranco, Rutvik Patel, and Stefanos Nikolaidis. Integrating field of view in human-aware collaborative planning. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*. IEEE, 2025.

[17] Yafei Hu, Quanting Xie, Vidhi Jain, Jonathan Francis, Jay Patrikar, Nikhil Keetha, Seungchan Kim, Yaqi Xie, Tianyi Zhang, Hao-Shu Fang, et al. Toward general-purpose robots via foundation models: A survey and meta-analysis. *arXiv preprint arXiv:2312.08782*, 2023.

[18] Stephen James, Zicong Ma, David Rovick Arrojo, and Andrew J Davison. Rlbench: The robot learning benchmark & learning environment. *IEEE Robotics and Automation Letters*, 5(2):3019–3026, 2020.

[19] Oliver P John, Eileen M Donahue, and Robert L Kentle. Big five inventory. *Journal of personality and social psychology*, 1991.

[20] Tobias Kaupp, Alexei Makarenko, and Hugh Durrant-Whyte. Human–robot communication for collaborative decision making—a probabilistic approach. *Robotics and Autonomous Systems*, 58(5):444–456, 2010.

[21] Huao Li, Yu Quan Chong, Simon Stepputtis, Joseph Campbell, Dana Hughes, Michael Lewis, and Katia Sycara. Theory of mind for multi-agent collaboration via large language models. *arXiv preprint arXiv:2310.10701*, 2023.

[22] Wenkai Li, Jiarui Liu, Andy Liu, Xuhui Zhou, Mona Diab, and Maarten Sap. Big5-chat: Shaping llm personalities through training on human-grounded data. *arXiv preprint arXiv:2410.16491*, 2024.

[23] Matthew B Luebbers, Aaquib Tabrez, Kyler Ruvane, and Bradley Hayes. Autonomous justification for enabling explainable decision support in human-robot teaming. In *Robotics: Science and Systems*, 2023.

[24] Daniel J McAllister. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal*, 38(1):24–59, 1995.

[25] Oier Mees, Lukas Hermann, Erick Rosete-Beas, and Wolfram Burgard. Calvin: A benchmark for language-

conditioned policy learning for long-horizon robot manipulation tasks. *IEEE Robotics and Automation Letters*, 7(3):7327–7334, 2022.

[26] Elliot Meyerson, Mark J. Nelson, Herbie Bradley, Adam Gaier, Arash Moradi, Amy K. Hoover, and Joel Lehman. Language model crossover: Variation through few-shot prompting, 2024. URL https://arxiv.org/abs/2302.12170.

[27] Manisha Natarajan, Esmaeil Seraj, Batuhan Altundas, Rohan Paleja, Sean Ye, Letian Chen, Reed Jensen, Kimberlee Chestnut Chang, and Matthew Gombolay. Human-robot teaming: grand challenges. *Current Robotics Reports*, 4(3):81–100, 2023.

[28] George A Neuman, Stephen H Wagner, and Neil D Christiansen. The relationship between work-team personality composition and the job performance of teams. *Group & organization management*, 24(1):28–45, 1999.

[29] Tianwei Ni, Huao Li, Siddharth Agrawal, Suhas Raja, Fan Jia, Yikang Gui, Dana Hughes, Michael Lewis, and Katia P. Sycara. Adaptive agent architecture for real-time human-agent teaming. *CoRR*, abs/2103.04439, 2021. URL https://arxiv.org/abs/2103.04439.

[30] Stefanos Nikolaidis, Swaprava Nath, Ariel D Procaccia, and Siddhartha Srinivasa. Game-theoretic modeling of human adaptation in human-robot collaboration. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*, pp. 323–331, 2017.

[31] Olle Nilsson and Antoine Cully. Policy gradient assisted map-elites. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2021. doi: 10.1145/3449639.3459304.

[32] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701320. doi: 10.1145/3586183.3606763. URL https://doi.org/10.1145/3586183.3606763.

[33] Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shan Zheng Tan, Ida Momennejad, Katja Hofmann, et al. Imitating human behaviour with diffusion models. *arXiv preprint arXiv:2301.10677*, 2023.

[34] Thomas Pierrot, Valentin Macé, Felix Chalumeau, Arthur Flajolet, Geoffrey Cideron, Karim Beguir, Antoine Cully, Olivier Sigaud, and Nicolas Perrin-Gilbert. Diversity policy gradient for sample efficient quality-diversity optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1075–1083, 2022.

[35] Justin K Pugh, Lisa B Soros, and Kenneth O Stanley. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40, 2016.

[36] Yvonne Rogers and Paul Marshall. *Research in the Wild*. Morgan & Claypool Publishers, 2017.

[37] Alessandro Roncone, Olivier Mangin, and Brian Scassellati. Transparent role assignment and task allocation in human robot collaboration. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1014–1021. IEEE, 2017.

[38] Richard M Ryan and Edward L Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55 (1):68, 2000.

[39] Mikayel Samvelyan, Sharath Chandra Raparthy, Andrei Lupu, Eric Hambro, Aram H. Markosyan, Manish Bhatt, Yuning Mao, Minqi Jiang, Jack Parker-Holder, Jakob Foerster, Tim Rocktäschel, and Roberta Raileanu. Rainbow teaming: Open-ended generation of diverse adversarial prompts, 2024. URL https://arxiv.org/abs/2402.16822.

[40] Siddharth Srikanth, Varun Bhatt, Boshen Zhang, Werner Hager, Charles Michael Lewis, Katia P Sycara, Aaquib Tabrez, and Stefanos Nikolaidis. Algorithmic prompt generation for diverse human-like teaming and communication with large language models, 2025. URL https://arxiv.org/abs/2504.03991.

[41] DJ Strouse, Kevin McKee, Matt Botvinick, Edward Hughes, and Richard Everett. Collaborating with humans without human data. *Advances in Neural Information Processing Systems*, 34:14502–14515, 2021.

[42] Rajah Annamalai Subramanian, Sanjeev Kumar, and Philip R Cohen. Integrating joint intention theory, belief reasoning, and communicative action for generating team-oriented dialogue. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, pp. 1501. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[43] Ilia Sucholutsky, Lukas Muttenthaler, Adrian Weller, Andi Peng, Andreea Bobu, Been Kim, Bradley C Love, Erin Grant, Iris Groen, Jascha Achterberg, et al. Getting aligned on representational alignment. *arXiv preprint arXiv:2310.13018*, 2023.

[44] Aaquib Tabrez, Matthew B Luebbers, and Bradley Hayes. A survey of mental modeling techniques in human–robot teaming. *Current Robotics Reports*, 1:259–267, 2020.

[45] Bryon Tjanaka, Matthew C. Fontaine, Julian Togelius, and Stefanos Nikolaidis. Approximating gradients for differentiable quality diversity in reinforcement learning. *CoRR*, abs/2202.03666, 2022. URL https://arxiv.org/abs/2202.03666.

[46] Floris van den Oever and Jan Maarten Schraagen. Team communication patterns in critical situations. *Journal of Cognitive Engineering and Decision Making*, 15(1):28–51, 2021.

[47] Chengxing Xie, Canyu Chen, Feiran Jia, Ziyu Ye, Shiyang Lai, Kai Shu, Jindong Gu, Adel Bibi, Ziniu Hu, David Jurgens, James Evans, Philip Torr, Bernard Ghanem, and Guohao Li. Can large language model agents simulate human trust behavior?, 2024. URL https://arxiv.org/abs/2402.04559.

[48] Ziyi Yang, Zaibin Zhang, Zirui Zheng, Yuxian Jiang, Ziyue Gan, Zhiyu Wang, Zijian Ling, Jinsong Chen, Martz Ma, Bowen Dong, et al. Oasis: Open agents social

interaction simulations on one million agents. *arXiv preprint arXiv:2411.11581*, 2024.

[49] Michelle Zhao, Reid Simmons, and Henny Admoni. Coordination with humans via strategy matching. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 9116–9123. IEEE, 2022.

[50] Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, and Maarten Sap. Sotopia: Interactive evaluation for social intelligence in language agents, 2024. URL https://arxiv.org/abs/2310.11667.