Learning Structured Interaction Frameworks for Robot Navigation in Human Environments

Shuijing Liu
Postdoc at The University of Texas at Austin
Ph.D. from University of Illinois Urbana-Champaign
https://shuijing725.github.io

I. Introduction

Robots are becoming increasingly prevalent in our daily lives. However, these autonomous agents are usually designed to navigate in isolation such as restricted areas in warehouses, or in a non-interactive manner, such as stopping and waiting when humans are around. A main barrier that hinders seamless human-robot interaction is the highly dynamic and complex nature of humans: As shown in Fig. 1 (a) and (b), in real world navigation scenarios, human behaviors are driven by subtle intentions that are difficult to predict. In addition, humans interact with each other, robots, and static obstacles frequently, which causes abrupt behavior changes.

To address the above challanges, my research aims to enable mobile robots to interact and cooperate with people in dynamic environments. To accomplish this goal, the robot must predict intricate human behavior, reason about the interactions, and adjust its plans accordingly. Fortunately, interactive human behaviors are not completely random; instead, they follow inductive biases that originate from common sense knowledge and social norms. My key insight is that *we can leverage the underlying structures behind human behaviors to design robot learning paradigms for effective human-robot interaction.* Following this principle, I propose **structured machine learning frameworks for safe, efficient, and socially aware robot navigation alongside humans**. My main contributions include the following two aspects:

(1) A structured system that incorporates predictions of human behaviors in planning, which leads to longsighted robot plans that align with human intentions (Fig. 1 (c)).
(2) A principled approach to build structured network architectures with spatial and temporal reasoning capabilities, which improves the robustness of robot policies with respect to changes in human environments (Fig. 1 (d)).

My methods are motivated by and successfully deployed in **various application domains**, including navigation of mobile robots in human crowds, autonomous driving in dense traffic, and a navigation guide for blind people (Fig. 1 (a)).

II. Past and Ongoing Research

A. *Structured prediction and planning system [18, 19, 3, 4, 20]*

Humans express their intentions through explicit language instructions and subtle non-verbal motions. To cooperate with humans in a proactive way, robots must infer these intentions and use the predictions to make longsighted plans. A line of previous work decouples prediction and planning, both of which are trained separately and combined together in test time [11, 24, 25]. Although interpretable and easy to trouble shoot, this decoupled framework leads to compounding errors and information loss between the predictor and the planner [15, 23]. To deal with this issue, end-to-end (e2e) systems combines the two modules and train them with multi-task learning [5, 10, 22]. However, multi-task learning easily results in negative transfer, where the gradients of each objective conflicts with each other [15, 19]. As a result, the robot learns suboptimal behaviors that misalign with human intentions, such as intruding into pedestrians' personal spaces or failing to follow instructions [3, 19].

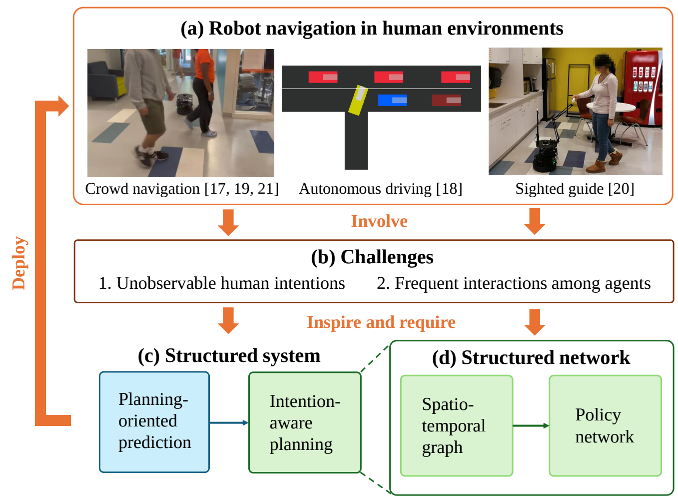

Fig. 1. **Overview of my research.** To address the challenges of robot navigation in human spaces, I propose (1) a structured prediction and planning system and (2) structured networks to learn robot policies. My method leads to successful robot deployment in various human-centered navigation tasks.

To combine the best of both worlds, we propose a structured robot stack to avoid collisions with humans and fulfill human intents, where the predictor and planner are modular but highly correlated (Fig. 1(c)). More specifically, our intention aware reinforcement learning (RL) framework incorporates predictions of human intentions into the robot policy learning process. To develop the prediction module, we start what needs to be predicted for downsteam planning tasks. For example, future trajectories and traits of pedestrians and drivers are important information for crowd navigation, whereas desired destinations is necessary for command following robots. After the predictors are trained, we use its output to modify the Markov Decision Process (MDP) formulation for policy learning. The predictor intentions are used as part of observation space of the MDP. Meanwhile, they are also used to calculate reward: robots are penalized for intruding into humans' future trajectories, and awarded for reaching humans' desired goals. By doing this, the robot gets immediate reward feedback for its actions that respond to human intentions during training. Thus, the robot is able to take proactive and context-aware actions during deployment. The predictor weights are frozen during RL training to minimize negative transfer and improve data efficiency.

Our system is decoupled but effectively uses predicted information to influence RL policy learning, maximizing the synergy between the prediction and planning modules. In all three application domains in Fig. 1(a), simulation and/or real-world experiments demonstrate that intention-aware RL leads to longsighted and more intention-aligned robot behaviors, such as avoiding the personal spaces of pedestrians, passing through an intersection when other drivers slow down and yield, and guiding a blind person to their desired detinations. Our work suggests the importance of maximizing synergies among different modules for system design. We also illustrate the necessity of modifying the MDP formulation for RL to succeed in real-world robotic problems. Our work has inspired follow-up research that co-designs prediction and planning modules in human-centered robotics [16, 12].

B. Structured policy networks [17, 18, 19, 21]

In dynamic environments, humans, robots, and static obstacles interact with each other in various ways. Robots must reason about these interactions that occur in both space and time to avoid collisions and unnecessary delays. To model interactions, previous works use mathematical models [14, 26], which heavily rely on assumptions, or black-box neural networks [7, 10], which leads to overfitting in low data regime. Thus, both lines of works result in undesirable robot behaviors such as freezing in front of a large crowd.

To learn robust policies with the limited data, for the first time, we propose a structured robot policy network with both principled interaction models and trainable hyperparameters (Fig. 1(d)). To this end, we first convert a dynamic scene with a heterogeneous spatio-temporal graph (st-graph, where agents and entities are nodes and their relations are edges. Based on the st-graph, we derive a novel neural network to parameterize navigation policies, which consists of two components as follows. First, we use attention networks to represent spatial interactions among agents at the same timestep. The attention networks enable the robot to pay more attention to important agents, which ensures good performance when the number of humans increases and the graph becomes complex. Second, we use recurrent neural networks to represent the temporal interactions, which model the rapid evolution of dynamic scenarios.

In crowd navigation and intersection driving tasks, our spatio-temporal network learns more robust robot policies when the density of humans and obstacles changes, compared with model-based methods and black-box methods. We also successfully deploy a mobile robot to challenging indoor navigation environments with continuous human flows, various furniture, and narrow corridors (Fig. 1(a)). This result demonstrates the power of injecting graph structures into neural networks. By doing so, complex multi-agent problems are decomposed into smaller components and become easier to solve. Our work has spurred a new line of research efforts that expand the usage of spatio-temporal graph attention networks in interactive navigation tasks [13, 28, 1, 6].

III. Future Work

Moving forward, to align robots with human values in our daily lives, I hope to enable robots to continually learn and adapt after deployment. To conquer performance degradation, I aim to develop interfaces and algorithms to enable end-users rather than engineers to finetune robots, which is more scalable and allows for customization of robots to individual preferences. Meanwhile, robots will collect a large amount of unstructured data from deployment. To turn these data into generalizable knowledge, I aim to contribute to large-scale HRI benchmarks and foundation HRI models.

Lifelong learning from non-expert feedback: When a robot is deployed in everyday environments, its performance drops inevitably due to domain shifts. Relying solely on engineers for finetuning after large-scale deployment would be both costly and impractical. Thus, we need data-efficient and intuitive fine-tuning algorithms that allow non-experts with little domain expertise to customize and improve the robot. Building on my previous work on a real-world finetunable RL pipeline [3] and recent advances in RL from human feedback [8], I aim to propose (1) intuitive user interfaces for non-experts to provide finetuning data from their own devices such as cellphones; (2) data-efficient algorithms for robot to self-improve with minimal data [4, 27].

Datasets and benchmarks: In many HRI fields, datasets and benchmarks are scattered among individual works, preventing a systematic comparison of new methodologies and training more generalizable foundation models. Inspired by the standardized benchmarks of autonomous driving and robot manipulation [2, 9], I aim to contribute to the benchmarking efforts of HRI fields with an emphasis on human-centered tasks such as spatio-temporal reasoning and intention prediction. By categorizing the difficulty levels of tasks, identifying the source of task difficulties, and collecting comprehensive scenarios and associated labels, we will be able to understand the merits and limitations of HRI models in a more objective way, which provides lessons and guidance for future research.

Foundation models in HRI: My previous work trains predictors and planners using a separate dataset or simulator for each task, which results in limited generalization. To improve the training data efficiency and generalization of my method, I plan to adopt large language/vision language models (LLMs/VLMs) to reason about dynamic scenarios and make plans for robots to execute. In this way, robots will be able to perform common sense reasoning and chain-of-thought reasoning [29]. These abilities are important to infer diverse human intents and assist humans with long-horizon mobile manipulation tasks in unstructured environments, such as helping a human chef prepare meals.

References

[1] Junaid Ahmed Ansari, Satyajit Tourani, Gourav Kumar, and Brojeshwar Bhowmick. Exploring social motion latent space and human awareness for effective robot navigation in crowded environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–8, 2023.

[2] Holger Caesar, Theo de Boer, Ronald Sim, Jonas Lange, Yuchen Yang, Kashyap Chitta, Simon Rosa, Steve Luu, Sammy Omari, Aseem Behl, and et al. nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles. *arXiv preprint arXiv:2106.11810*, 2021.

[3] Peixin Chang, Shuijing Liu, and Katherine Driggs-Campbell. Learning visual-audio representations for voice-controlled robots. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[4] Peixin Chang, Shuijing Liu, Tianchen Ji, Neeloy Chakraborty, Kaiwen Hong, and Katherine Rose Driggs-Campbell. A data-efficient visual-audio representation with intuitive fine-tuning for voice-controlled robots. In *Conference on Robot Learning (CoRL)*, 2023.

[5] Devendra Singh Chaplot, Kanthashree Mysore Sathyendra, Rama Kumar Pasumarthi, Dheeraj Rajagopal, and Ruslan Salakhutdinov. Gated-attention architectures for task-oriented language grounding. In *Conference on Artificial Intelligence (AAAI)*, pages 2819–2826, 2018.

[6] Bolei Chen, Haina Zhu, Shengkang Yao, Siyi Lu, Ping Zhong, Yu Sheng, and Jianxin Wang. Socially aware object goal navigation with heterogeneous scene representation learning. *IEEE Robotics and Automation Letters*, 9(8):6792–6799, 2024.

[7] Yu Fan Chen, Michael Everett, Miao Liu, and Jonathan P How. Socially aware motion planning with deep reinforcement learning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[8] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30, 2017.

[9] Open X-Embodiment Collaboration et al. Open x-embodiment: Robotic learning datasets and rt-x models. *arXiv preprint arXiv:2310.08864*, 2023. URL https://arxiv.org/abs/2310.08864.

[10] Daniel Dugas, Olov Andersson, Roland Siegwart, and Jen Jen Chung. Navdreams: Towards camera-only rl navigation among humans. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2504–2511, 2022.

[11] Stuart Eiffert, He Kong, Navid Pirmarzdashti, and Salah Sukkarieh. Path planning in dynamic environments using generative rnns and monte carlo tree search. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2020.

[12] Aamir Hasan, Neeloy Chakraborty, Haonan Chen, Jung-Hoon Cho, Cathy Wu, and Katherine Driggs-Campbell. PeRP: Personalized residual policies for congestion mitigation through co-operative advisory systems. In *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2023.

[13] Haodong He, Hao Fu, Qiang Wang, Shuai Zhou, Wei Liu, and Yang Chen. Spatio-temporal transformer-based reinforcement learning for robot crowd navigation. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2023.

[14] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 1995.

[15] Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, and Hongyang Li. Planning-oriented autonomous driving. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[16] Kanghoon Lee, Jiachen Li, David Isele, Jinkyoo Park, Kikuo Fujimura, and Mykel J. Kochendorfer. Robust driving policy learning with guided meta reinforcement learning. In *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pages 4114–4120, 2023.

[17] Shuijing Liu, Peixin Chang, Weihang Liang, Neeloy Chakraborty, and Katherine Driggs-Campbell. Decentralized structural-rnn for robot crowd navigation with deep reinforcement learning. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2021.

[18] Shuijing Liu, Peixin Chang, Haonan Chen, Neeloy Chakraborty, and Katherine Driggs-Campbell. Learning to navigate intersections with unsupervised driver trait inference. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2022.

[19] Shuijing Liu, Peixin Chang, Zhe Huang, Neeloy Chakraborty, Kaiwen Hong, Weihang Liang, D. Livingston McPherson, Junyi Geng, and Katherine Driggs-Campbell. Intention aware robot crowd navigation with attention-based interaction graph. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[20] Shuijing Liu, Aamir Hasan, Kaiwen Hong, Runxuan Wang, Peixin Chang, Zachary Mizrachi, Justin Lin, D Livingston McPherson, Wendy A Rogers, and Katherine Driggs-Campbell. Dragon: A dialogue-based robot for assistive navigation with visual language grounding. *IEEE Robotics and Automation Letters*, 2024.

[21] Shuijing Liu, Haochen Xia, Fatemeh Cheraghi Pouria, Kaiwen Hong, Neeloy Chakraborty, and Katherine Driggs-Campbell. Height: Heterogeneous interaction graph transformer for robot navigation in

crowded and constrained environments. *arXiv preprint arXiv:2411.12150*, 2024.

[22] Zhijian Liu, Haotian Tang, Alexander Amini, Xinyu Yang, Huizi Mao, Daniela L. Rus, and Song Han. Bevfusion: Multi-task multi-sensor fusion with unified bird's-eye view representation. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2774–2781, 2023.

[23] Christoforos Mavrogiannis, Francesca Baldini, Allan Wang, Dapeng Zhao, Pete Trautman, Aaron Steinfeld, and Jean Oh. Core challenges of social robot navigation: A survey. *ACM Transactions on Human-Robot Interaction*, 12(3), 2023.

[24] Rohan Paul, Jacob Arkin, Derya Aksaray, Nicholas Roy, and Thomas M Howard. Efficient grounding of abstract spatial concepts for natural language interaction with robot platforms. *The International Journal of Robotics Research*, 2018.

[25] Mohit Shridhar, Dixant Mittal, and David Hsu. Ingress: Interactive visual grounding of referring expressions. *The International Journal of Robotics Research*, 2020.

[26] Jur Van Den Berg, Stephen J Guy, Ming Lin, and Dinesh Manocha. Reciprocal n-body collision avoidance. In *Robotics research*. 2011.

[27] Chen Wang, Linxi Fan, Jiankai Sun, Ruohan Zhang, Li Fei-Fei, Danfei Xu, Yuke Zhu, and Anima Anandkumar. Mimicplay: Long-horizon imitation learning by watching human play. In *Conference on Robot Learning (CoRL)*, 2022.

[28] Weizheng Wang, Ruiqi Wang, Le Mao, and Byung-Cheol Min. Navistar: Socially aware robot navigation with hybrid spatio-temporal graph transformer and preference learning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11348–11355, 2023.

[29] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:24824–24837, 2022.