Introspection for Long-Horizon Robot Planning under Uncertainty

Abhishek Paudel

***Abstract*— The next generation of household service robots must not only perform day-to-day tasks effectively, but also learn and improve over time by continuously evaluating their own performance. My research focuses on *introspection* as a technique that enables a robot to self-assess its behavior while performing long-horizon tasks in environments that may not be fully known. We show that such introspection enables improved performance in tasks such as navigation in unknown environments, deployment-time learning and adaptation, and large language model (LLM)-informed object search.**

***Keywords*—**planning under uncertainty, long-horizon planning, introspection, robot learning

I. Introduction

As service robots become part of our homes and offices assisting us in tasks like cleaning, making breakfast, moving things around or other routine chores, they must not only perform these tasks effectively but also learn and improve over time by continuously evaluating their own performance. Humans demonstrate remarkable abilities to reason about our effectiveness while performing tasks and identify ways to improve our skills. By imagining alternative actions and their outcomes, we can *introspect* without actually taking those actions, helping us avoid poor decisions. Such abilities are essential for household robots to continuously refine their behavior and avoid undesirable outcomes. Thus, *introspection* is a crucial capability for next generation of service robots.

While advances in machine learning have enabled robots to learn from prior experiences and get better over time, most are limited in their abilities to identify poor long-horizon planning performance. *Runtime monitoring* approaches aim to evaluate the deployment-time reliability of learning-guided robot behavior [1]–[4] to decide whether to continue using learning or switch to backup strategies when performance degrades. However, most such approaches focus on addressing issues related to local perception or raw model predictions rather than evaluating overall task performance [3], [4], limiting their effectiveness in assessing long-horizon behavior.

My research aims to improve deployment-time performance for long-horizon robot planning under uncertainty. To achieve this, my research focuses on developing techniques for *introspection* via which the robot can reason about good and poor behaviors without the robot having to go through them first-hand to realize their impacts. I aim to develop general-purpose introspection techniques that are not only applicable to navigation and task planning problems but are also agnostic learning paradigms (e.g. pre-trained models, LLMs, etc.) used to inform robot behavior. By leveraging these techniques, my work facilitates quick selection of best performing behaviors during deployment, enabling robots to achieve improved long-horizon performance despite uncertainty. I envision robots that seamlessly adapt to new environments and perform well during the course of deployment by leveraging the techniques for introspection to continuously assess their performance and learn from hindsight.

Abhishek Paudel is a PhD candidate advised by Dr. Gregory J. Stein in the Department of Computer Science at George Mason University, Fairfax, Virginia, USA. Email: apaudel4@gmu.edu

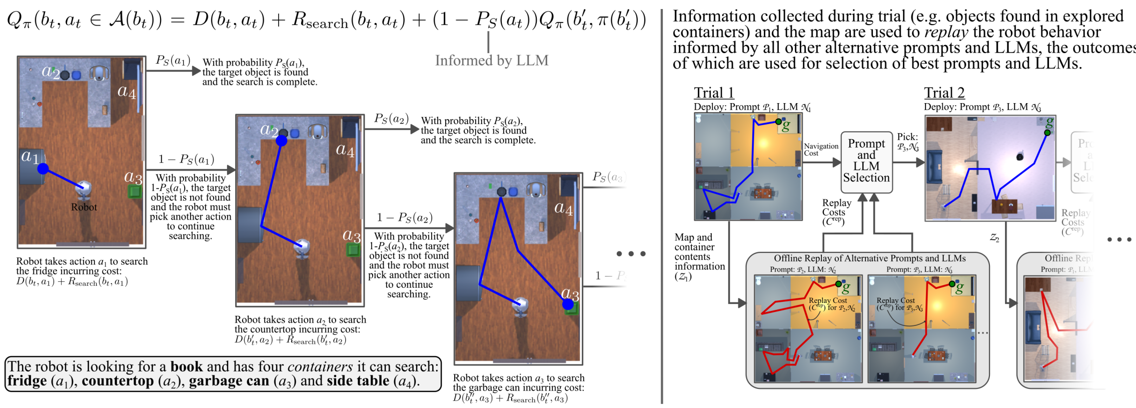
II. Research Contributions

During my PhD studies, I have developed algorithmic foundations for introspection in the context of long-horizon planning under uncertainty. Below are my key contributions.

Introspection via Offline Alt-Policy Replay Improving deployment-time performance for goal-directed navigation in unknown environments requires the robot to look back on its behavior and *introspect* whether a different strategy could have performed better in the same scenario. However, testing multiple strategies directly is costly and requires the robot to experience potentially poor behaviors before ruling them out. In our work [5], we introduce an introspection technique, called *offline alt-policy replay*, which uses data collected during deployment to perform counterfactual reasoning about alternative strategies the robot could have instead deployed. Such introspection computes lower bound cost of alternative strategies without deploying them, which is then used to rule out poor strategies and identify best ones, resulting in 5–11% reduction in average navigation costs during deployment.

Asymptotically Optimal Policy Selection In many scenarios, relying on a single strategy during deployment may not yield good performance, requiring the robot to select from multiple strategies to find the one that performs best in the deployed environment. To facilitate this selection, we introduce a policy selection formulation [5] based on the multi-armed bandit problem. However, bandit algorithms like upper confidence bound (UCB) [9] converge too slowly in this domain as they require many iterations of trial and error to identify the best performing policies. To accelerate convergence, we incorporate lower bound costs computed via our introspection technique [5] as an additional bound in UCB bandit-like selection strategy. This approach enables quick, data-efficient identification of best policies, achieving 67–96% improvements in cumulative regret while maintaining guarantees on asymptotic sub-linear regret.

LLM-informed Model-based Planning, and Introspection for Prompt Selection Effective object search in partially-known environments requires long-horizon reasoning and continuous self-evaluation. While LLMs offer valuable commonsense world knowledge for such tasks, directly



(a) LLM-informed Model-Based Planning for Object Search

(b) Deployment-time Selection of Prompts and LLMs

Fig. 1. Our work [6]–[8] enables improved object search in partially-known environments by (a) augmenting model-based planning with commonsense knowledge from LLMs, and (b) leveraging introspection via offline replay for deployment-time bandit-like selection of best prompts and LLMs.

prompting LLMs to decide where the robot should search often leads to suboptimal, myopic behavior [10]–[12] as their planning and reasoning abilities are often brittle [13]–[16]. Our work [7] introduces an LLM-informed model-based planning framework for object search (Fig. 1(a)), where high-level actions correspond to searching containers, and model-based planning is guided by LLM's predictions of object likelihood, leading up to 23% improvements in average navigation cost. Crucially, this high-level action abstraction enables fast deployment-time prompt selection [8] through introspection via offline replay (Fig. 1(b)), enabling the robot to quickly identify best choices during deployment and thus resulting in 40% improvements in cumulative regret.

Deployment-time Learning and Adaptation Robots deployed in unknown environments must learn and adapt their behavior during deployment to perform well in a wide variety of scenarios. While various deployment-time learning and adaptation strategies exist, there is no *one-size-fits-all* approach, and the robot must select best-performing strategies during deployment. However, selecting best strategies can be challenging due to the non-stationary nature of continually evolving policies. To avoid having to deploy such continually evolving strategies for assessing whether they should be deployed, we leverage introspection via *offline replay* to determine when the policies learned or adapted during deployment would yield better performance without actually deploying them. In this work [17], we demonstrate improvements up to 19% in average cost and 95% in cumulative regret with deployment-time selection among multiple policies: pre-trained, learned from scratch during deployment, and adapted during deployment via visual domain adaptation techniques like CycleGAN [18]. This work enables improved deployment-time performance for long-horizon goal-directed navigation in unknown environments by selecting the best strategies from an ensemble of continually evolving policies.

III. Ongoing and Future Directions

My prior works demonstrate the benefits of introspection in multiple scenarios including goal-directed navigation, object search, LLM-informed planning, and deployment-time learning and adaptation. However, such introspection techniques have the potential to solve more general and broader problems in robotics, which I am eager to explore.

Introspection for Effective Task Planning under Uncertainty Many real-world service robot tasks, such as assisting in homes and offices, often require executing complex, temporally-extended objectives. For example, a task like serving a coffee may involve sequential subtasks such as finding a cup, pouring the coffee, and placing it on the table, often under conditions of partial knowledge about the environment. Introspection of such plans executed by the robot, combined with the hindsight information about the environment, could enable robots to reason about and identify better planning strategies for effective performance of such long-lived household robots. My ongoing work focuses on developing introspection techniques for general task planning problems and how they can seamlessly be integrated into standard planning frameworks like planning domain definition language (PDDL) to enable reliable and effective planning performance for long-horizon tasks despite uncertainty—a capability that is crucial for service robots.

Reliability and Adaptability for LLM- and VLM-enabled Robotics As robots increasingly rely on LLMs or vision language models (VLMs) for decision-making, adaptation to new environments while demonstrating reliable task performance is critical. While my recent work on LLM-informed planning and prompt selection [6] makes progress towards reliable deployment of such LLM-based robotic systems, there remain significant challenges in making such systems adaptable to changes in environments and human preferences while still maintaining reliability. I aim to develop methods for LLM-based robotic systems to autonomously adapt how LLMs are used for informing robot behaviors, e.g. via automated improvements of prompting strategies to facilitate adaptation. This research direction addresses a key challenge in making LLM-guided robots reliable and practical for real-world applications.

Introspection for Multi-Robot Teams In multi-robot applications like warehouse automation and search-and-rescue, introspection not only about one's own behavior but also about the behaviors of other robots in the team has potential to improve collective strategies and enhance team performance. I am excited to explore these avenues in future.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 2232733.

References

[1] Q. M. Rahman, P. Corke, and F. Dayoub, "Run-time monitoring of machine learning for robotic perception: A survey of emerging trends," *IEEE Access*, 2021.

[2] P. Mallozzi, E. Castellano, P. Pelliccione, G. Schneider, and K. Tei, "A runtime monitoring framework to enforce invariants on reinforcement learning agents exploring complex environments," in *2019 IEEE/ACM 2nd International Workshop on Robotics Software Engineering (RoSE)*, 2019.

[3] W. Zhou, J. S. Berrio, S. Worrall, and E. Nebot, "Automated evaluation of semantic segmentation robustness for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[4] S. Daftry, S. Zeng, J. A. Bagnell, and M. Hebert, "Introspective perception: Learning to predict failures in vision systems," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016.

[5] A. Paudel and G. J. Stein, "Data-efficient policy selection for navigation in partial maps via subgoal-based abstraction," in *International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[6] A. Paudel, S. Hossain, and G. J. Stein, "LLM-informed object search in partially-known environments via model-based planning and prompt selection," *preprint*, 2025.

[7] S. Hossain, A. Paudel, and G. J. Stein, "Enhancing object search by augmenting planning with predictions from large language models," in *2nd CoRL Workshop on Learning Effective Abstractions for Planning*, 2024.

[8] A. Paudel and G. J. Stein, "Deployment-time selection of prompts for LLM-informed object search in partially-known environments," in *ICRA 2025 Workshop on Foundation Models and Neuro-Symbolic AI for Robotics*, 2025.

[9] T. L. Lai, H. Robbins, *et al.*, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, 1985.

[10] K. Zhou, K. Zheng, C. Pryor, Y. Shen, H. Jin, L. Getoor, and X. E. Wang, "ESC: Exploration with soft commonsense constraints for zero-shot object navigation," in *International Conference on Machine Learning*, 2023.

[11] V. S. Dorbala, J. F. Mullen Jr, and D. Manocha, "Can an embodied agent find your "cat-shaped mug"? LLM-guided exploration for zero-shot object navigation," *arXiv preprint arXiv:2303.03480*, 2023.

[12] B. Yu, H. Kasaei, and M. Cao, "L3MVN: Leveraging large language models for visual target navigation," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[13] K. Valmeekam, A. Olmo, S. Sreedharan, and S. Kambhampati, "Large language models still can't plan (a benchmark for LLMs on planning and reasoning about change)," in *NeurIPS 2022 Foundation Models for Decision Making Workshop*, 2022.

[14] K. Valmeekam, M. Marquez, S. Sreedharan, and S. Kambhampati, "On the planning abilities of large language models-a critical investigation," *Advances in Neural Information Processing Systems*, 2023.

[15] S. Kambhampati, "Can large language models reason and plan?" *Annals of the New York Academy of Sciences*, 2024.

[16] S. Kambhampati, K. Valmeekam, L. Guan, M. Verma, K. Stechly, S. Bhambri, L. P. Saldyt, and A. B. Murthy, "LLMs can't plan, but can help planning in LLM-modulo frameworks," in *International Conference on Machine Learning*, 2024.

[17] A. Paudel, X. Xiao, and G. J. Stein, "Multi-strategy deployment-time learning and adaptation for navigation under uncertainty," in *Conference on Robot Learning (CoRL)*, 2024.

[18] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017.