Extended Abstract:
Adaptive Planner Parameter Learning for Mobile Robot Navigation in the Wild

Xuesu Xiao[1], Bo Liu[1], Garrett Warnell[2], Zizhao Wang[1], Zifan Xu[1], Gauraang Dhamankar[1], Anirudh Nair[1], and Peter Stone[1, 3]

[1]University of Texas at Austin, [2]Army Research Laboratory, [3]Sony AI

xiao@cs.utexas.edu, bliu@cs.utexas.edu, garrett.a.warnell.civ@mail.mil, zizhao.wang@utexas.edu, zfxu@utexas.edu, gauraang@cs.utexas.edu, ani.nair@utexas.edu, pstone@cs.utexas.edu

Introduction

Mobile robot navigation has been studied by the robotics community for decades (Fox, Burgard, and Thrun 1997). In relatively controlled and uniform environments, e.g., indoor laboratories, existing navigation systems can produce robust navigation behaviors with verifiable guarantees that the robot will not collide with obstacles while moving.

However, when facing the variety of navigation environments encountered during deployment in the wild, existing navigation systems require robotics experts to make in-situ adjustments to adapt systems through sensor calibration (Xiao et al. 2017) or parameter tuning (e.g. maximum speed, sampling rate, inflation radius)) (Zheng 2017). This dependency on expert roboticists onsite during deployment makes it difficult for non-expert users to successfully deploy mobile robots in the wild.

On the other hand, these non-expert users often are able to have valuable interactions with the platform that could be used to teach robots to adapt to a variety of deployment scenarios. For example, non-expert users can provide a teleoperated demonstration to show the robot the desirable navigation behaviors, intervene when the robot performs suboptimally, or simply give evaluative feedback, from which the robot can learn.

To this end, we have proposed *Adaptive Planner Parameter Learning* (APPL), which utilizes different interaction modalities with non-expert users to *learn* adaptive planner parameters and enable desirable navigation behaviors in the wild. We have introduced a suite of APPL methods that utilize human demonstration (APPLD), corrective intervention (APPLI), evaluative feedback (APPLE), and unsupervised reinforcement learning (APPLR). In contrast to end-to-end learning for navigation, the APPL paradigm inherits the benefits from classical navigation approaches, e.g. safety and explainability, and simultaneously enjoys the flexibility and adaptivity of learning methods (Xiao et al. 2020b; Liu, Xiao, and Stone 2021). All these methods are tested on the Robot Operating System `move_base` navigation stack with a physical mobile robot in the real world.

Adaptive Planner Parameter Learning (APPL)

We present four APPL methods to enable desirable navigation behaviors in the wild, which are based on a *parameter library* or a *parameter policy* learned from different non-expert human interaction modelities.

APPL from Demonstration (APPLD)[1]

While they may not be able to perform expert-level system tuning, non-expert users can typically provide a teleoperated demonstration to show the robot the desirable navigation behavior during deployment in the wild (Xiao et al. 2020a). The overview of APPLD is shown in Fig. 1.

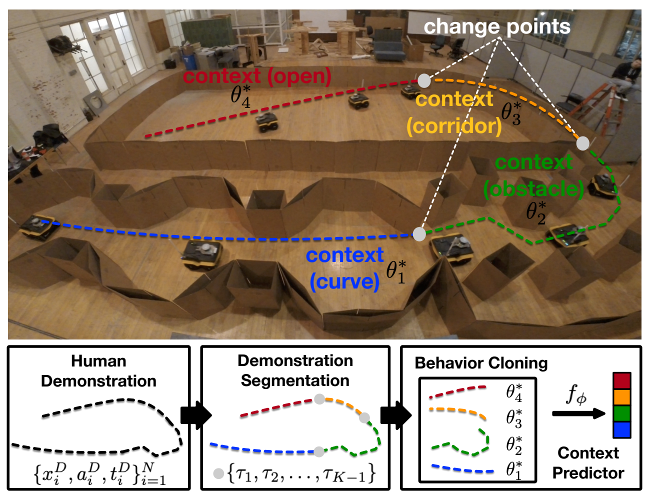

Figure 1: APPLD. Human demonstrations are segmented into different contexts, for each of which a set of parameters θ_k^* is learned via behavior cloning. During deployment, the system parameters are dynamically selected from the learned parameter library using an online context predictor.

APPL from Interventions (APPLI)[2]

In many places, the default navigation system performs well (shown in green in Fig. 2), but it may fail (red) or suffer from suboptimal behavior (yellow) in others. The robot learns

[1]https://www.youtube.com/watch?v=u2xxPTZA0DY

[2]https://www.youtube.com/watch?v=aRAJ1Dl69gI&t=10s

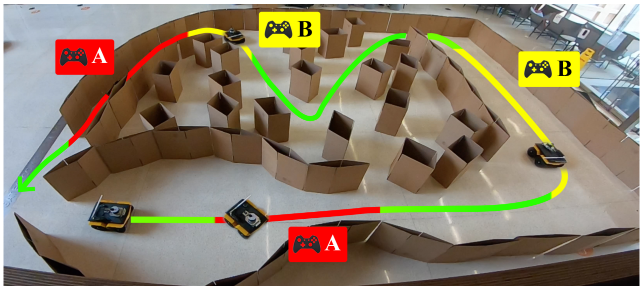

Figure 2: APPLI. Human interventions in places where failure (Type A) and suboptimal navigation (type B) occur are used to learn adaptive planner parameters and, based on a confidence measure, select them during deployment.

new parameters from the interventions provided by the non-expert users at these challenging places and applies them in similar scenarios based on a confidence measure during deployment (Wang et al. 2021).

APPL from Evaluative Feedback (APPLE)

In cases for which non-expert users are unable to provide teleoperated demonstrations or corrective interventions, robot performance can still be improved using APPL with interactions that are easier to provide, e.g., evaluative feedback. APPLE uses either discrete feedback (the robot's action is good or bad, i.e. $\{-1, 1\}$) to select a parameter set from a pre-built parameter library, or learns a parameter policy (explained in detail in the following paragraph) from continuous feedback (e.g. a score within the range $[-1, 1]$).

APPL from Reinforcement (APPLR)[3]

In contrast to all aforementioned APPL variants, APPLR (Xu et al. 2021) does not require any human interactions and is pre-trained before any deployment. APPLR learns a *parameter policy* (Fig. 3) that is trained using reinforcement learning in the simulated Benchmark for Autonomous Navigation (BARN) dataset (Fig. 4) (Perille et al. 2020) to make planner parameter decisions in such a way that allows the system to take suboptimal actions at one state in order to perhaps perform even better in the future. For example, while it may be suboptimal in the moment to slow down or alter the platform's trajectory before a turn, doing so may allow the system to carefully position itself so that it can go much faster in the future than if it had not.

Conclusions

In this abstract, we present an overview of APPL, which provides a suite of methods to enable non-expert users to deploy mobile robot with desirable navigation behaviors in the wild. APPL inherits benefits from classical planning approaches and also enjoys adaptivity from machine learning. Video links of the physical experiments of the APPL methods are provided, along with references to the detailed papers.

[3]https://www.youtube.com/watch?v=JKHTAowdGUk&t=134s

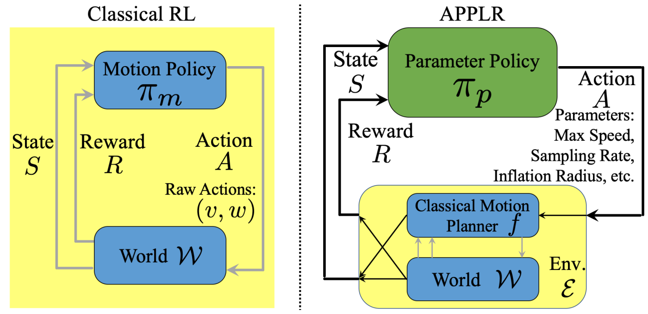

Figure 3: APPLR's Parameter Policy vs. Classical RL's Motion Policy

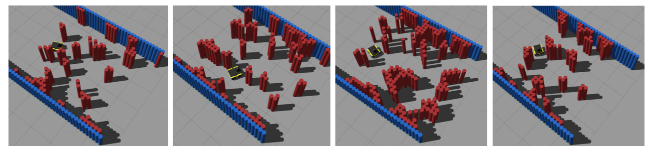

Figure 4: APPLR Training in BARN Dataset

References

Fox, D.; Burgard, W.; and Thrun, S. 1997. The dynamic window approach to collision avoidance. *IEEE Robotics & Automation Magazine* 4(1): 23–33.

Liu, B.; Xiao, X.; and Stone, P. 2021. A Lifelong Learning Approach to Mobile Robot Navigation. *IEEE Robotics and Automation Letters* 6(2): 1090–1096.

Perille, D.; Truong, A.; Xiao, X.; and Stone, P. 2020. Benchmarking Metric Ground Navigation. In *2020 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*. IEEE.

Wang, Z.; Xiao, X.; Liu, B.; Warnell, G.; and Stone, P. 2021. Appli: Adaptive planner parameter learning from interventions. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.

Xiao, X.; Dufek, J.; Woodbury, T.; and Murphy, R. 2017. UAV assisted USV visual navigation for marine mass casualty incident response. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6105–6110. IEEE.

Xiao, X.; Liu, B.; Warnell, G.; Fink, J.; and Stone, P. 2020a. APPLD: Adaptive Planner Parameter Learning From Demonstration. *IEEE Robotics and Automation Letters* 5(3): 4541–4547.

Xiao, X.; Liu, B.; Warnell, G.; and Stone, P. 2020b. Motion Control for Mobile Robot Navigation Using Machine Learning: a Survey. *arXiv preprint arXiv:2011.13112* .

Xu, Z.; Dhamankar, G.; Nair, A.; Xiao, X.; Warnell, G.; Liu, B.; Wang, Z.; and Stone, P. 2021. Applr: Adaptive planner parameter learning from reinforcement. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.

Zheng, K. 2017. Ros navigation tuning guide. *arXiv preprint arXiv:1706.09068* .