

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

Xuesu Xiao<sup>1</sup>, Bo Liu<sup>1</sup>, Garrett Warnell<sup>2</sup>, Zizhao Wang<sup>1</sup>, Zifan Xu<sup>1</sup>,  
Gauraang Dhamankar<sup>1</sup>, Anirudh Nair<sup>1</sup>, and Peter Stone<sup>1,3</sup>

<sup>1</sup>University of Texas at Austin, <sup>2</sup>Army Research Laboratory, <sup>3</sup>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 modalities.

### APPL from Demonstration (APPLD)<sup>1</sup>

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.

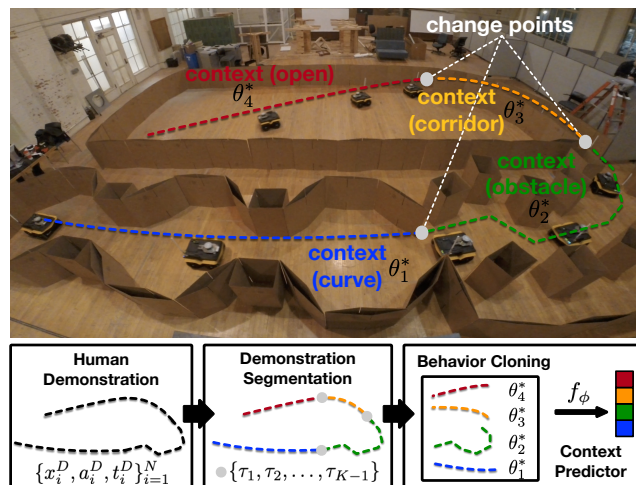


Figure 1: APPLD. Human demonstrations are segmented into different contexts, for each of which a set of parameters  $\theta_k^*$  is learned via behavior cloning. During deployment, the system parameters are dynamically selected from the learned parameter library using an online context predictor.

### APPL from Interventions (APPLI)<sup>2</sup>

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

<sup>1</sup><https://www.youtube.com/watch?v=u2xxPTZA0DY>

<sup>2</sup><https://www.youtube.com/watch?v=aRAJ1D169gI&t=10s>

