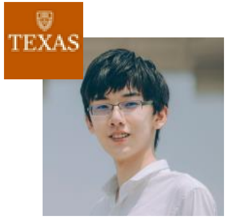
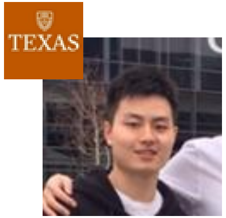
LARG
Learning Agents Research Group
The University of Texas at Austin

Adaptive Planner Parameter Learning from Interventions

Zizhao Wang[1], Xuesu Xiao[1], Bo Liu[1], Garrett Warnell[2], and Peter Stone[1, 3]





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

Motivation

Deploying a classical autonomous navigation system in new environments requires adaptivity to variety of environments by tuning the parameters (max speed, etc).

Otherwise, it may produce suboptimal behaviors or even fail.

Motivation

Manually re-tuning those parameters requires expert knowledge.

However, it’s easy for human to provide interventions via teleoperation.

Related Work

Learning for Navigation

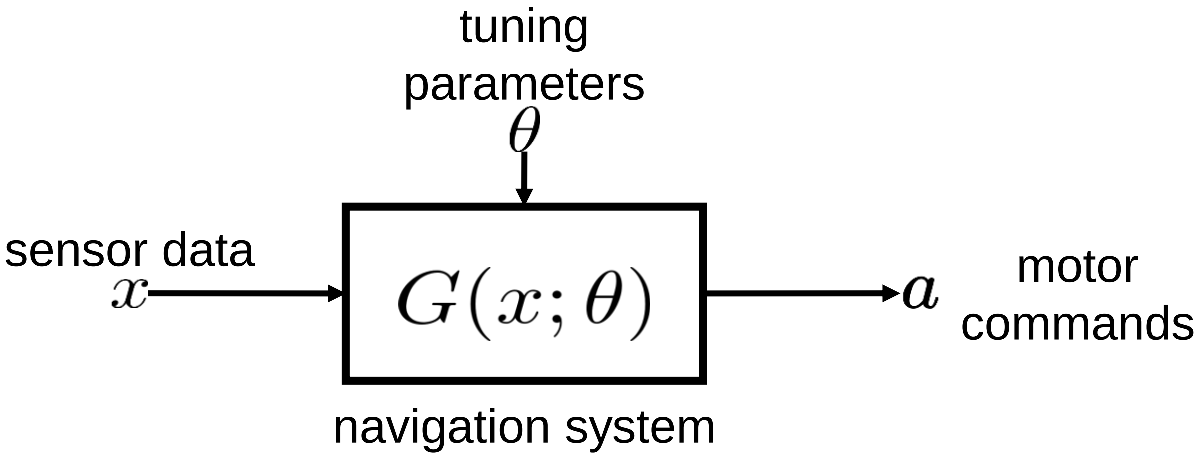
- End-to-end learning (Tai, et al., IROS'17; Zhang, et al., IROS'17)
- Learning for subsystems
  - Global planner (Yao, et al., IROS'19)
  - Local planner (Gao, et al., CoRL'17; Faust, et al., ICRA'18)
- Learning for components in subsystems
  - Cost function (Shiarlis, et al., ICRA'17), cost map (Luber, et al., IROS'12), ...
  - Planner parameters

Xiao, et al., *Motion Control for Mobile Robot Navigation Using Machine Learning: a Survey*, the International Journal of Robotics Research, 2020.

Background

APPL: Adaptive Planner Parameter Learning

- Use different human interactions to “tune” any navigation system.

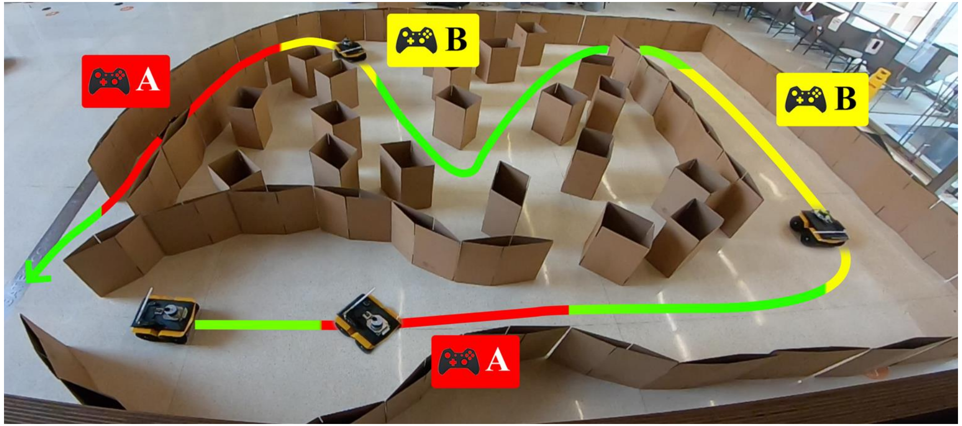


- APPLD: APPL from human demonstrations (Xiao, et al., IROS’20)
- APPLR: APPL from reinforcement (Xu, et al., ICRA’21)

APPLI: Adaptive Planner Parameter Learning From Interventions

Contributions

- Use a few **interventions** rather than full demonstrations to focus on challenging scenarios.
- Use a **confidence-based** context classifier to generalize to unseen environments.



A failures

B suboptimal behaviors

APPLI: Adaptive Planner Parameter Learning From Interventions

Procedures

1. Collect interventions.

$$\longrightarrow I_{1:N} \text{ where } I_i = \{(x_1, a_1), \ldots, (x_{T_i}, a_{T_i})\}$$

Here we also consider different human concentration/expertise levels:

Human supervisors that are not focused/experts may not recognize suboptimal behaviors and provide corresponding Type B interventions.

A failures

B suboptimal behaviors

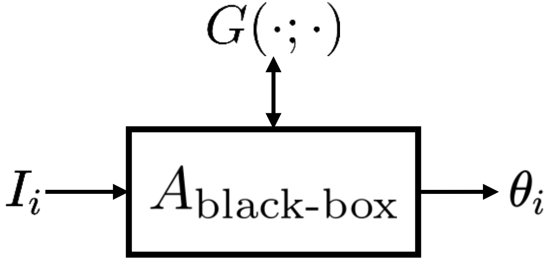
APPLI: Adaptive Planner Parameter Learning From Interventions

Procedures

1. Collect interventions.

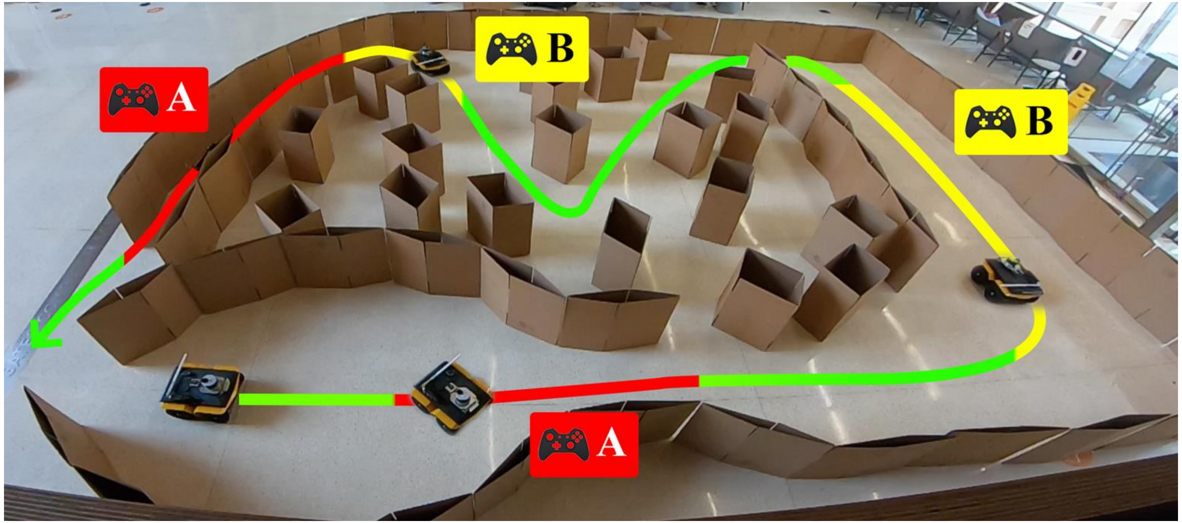
$\longrightarrow I_{1:N}$ where $I_i = \{(x_1, a_1), \ldots, (x_{T_i}, a_{T_i})\}$

2. Use black-box optimization to find navigation system parameters.



APPLI: Adaptive Planner Parameter Learning From Interventions

3. Train **Context Classifier with Confidence Measure** based on Evidential Deep Learning[1].



[1] M. Sensoy, L. Kaplan, and M. Kandemir, *Evidential deep learning to quantify classification uncertainty*, in Advances in Neural Information Processing Systems, 2018

APPLI: Adaptive Planner Parameter Learning From Interventions

3. Train **Context Classifier with Confidence Measure** based on Evidential Deep Learning[1].

Trust classifier's prediction when it is confident.

$$x \rightarrow \boxed{B_\phi} \rightarrow$$

Context	1	2	3	4
Prediction	**0.7**	0.1	0.1	0.1
Confidence	**0.8**			

$$\rightarrow \theta_1$$

Switch to the default parameter $\bar{\theta}$ when the classifier is inconfident.

$$x \rightarrow \boxed{B_\phi} \rightarrow$$

Context	1	2	3	4
Prediction	**0.85**	0.05	0.05	0.05
Confidence	**0.4**			

$$\rightarrow \bar{\theta}$$

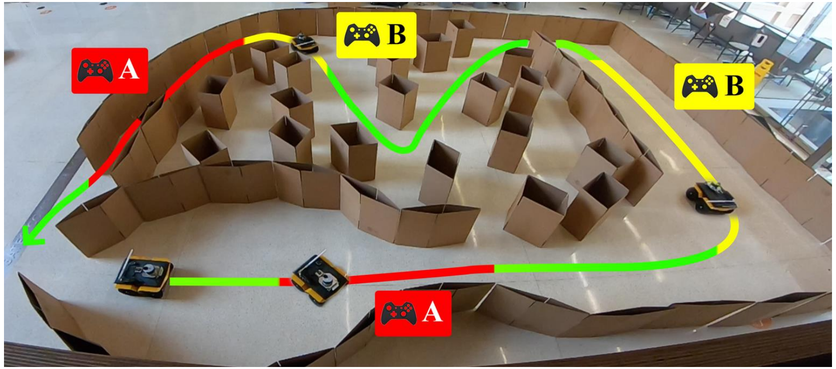
[1] M. Sensoy, L. Kaplan, and M. Kandemir, *Evidential deep learning to quantify classification uncertainty*, in Advances in Neural Information Processing Systems, 2018

APPLI: Adaptive Planner Parameter Learning From Interventions

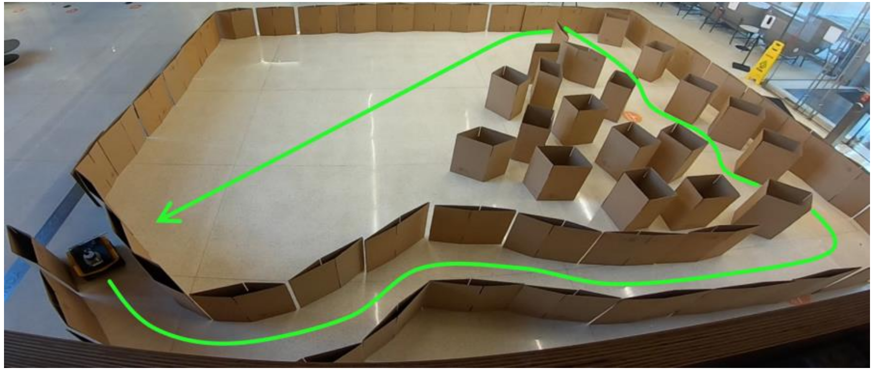
Real Experiment Setup

Robot: Clearpath Jackal (Velodyne Puck lidar)

Human Supervisor: an author (teleoperation via Xbox controller)



Training Environment



Unseen Environment

APPLI: Adaptive Planner Parameter Learning From Interventions

Real Experiment Results

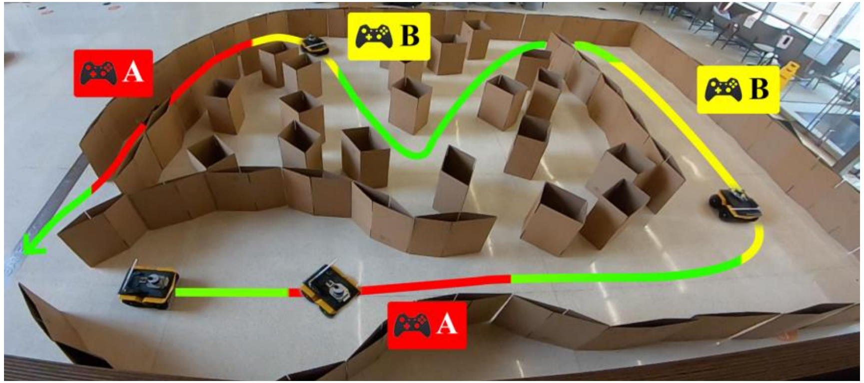
Traversal Time in Training Environment

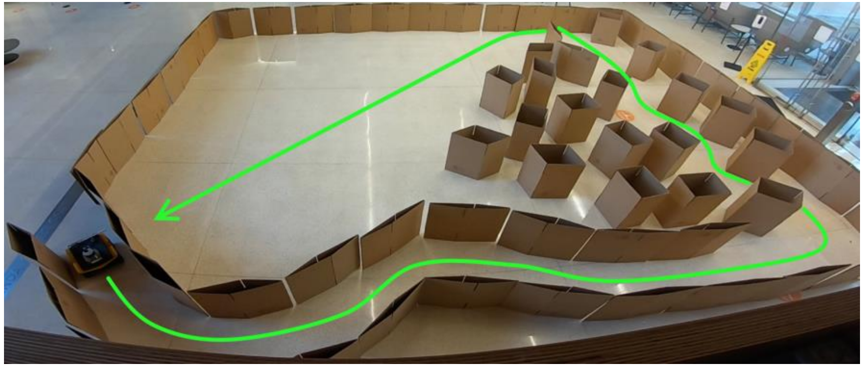
Default	Type A	Type A+B	Full Demo
134.0±60.6s	77.4±2.8s	70.6±3.2s	78.0±2.7s

Type A: APPLI learned from type A interventions (failures) only
Type A+B: APPLI learned from both type A and type B interventions (suboptimal behaviors)

Traversal Time in Unseen Environment

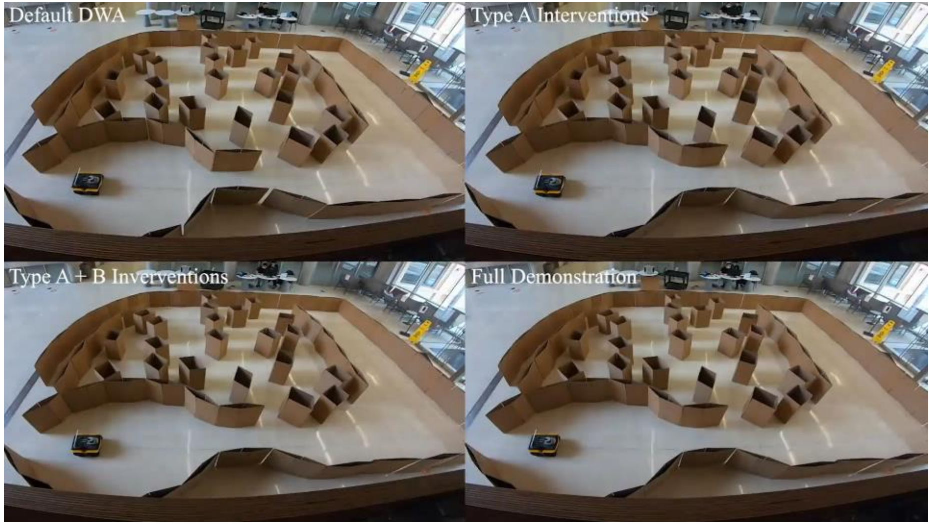
Default	Type A	Type A+B	Full Demo
109.2±50.8s	71±0.7s	59±0.7s	62.0±2.0s



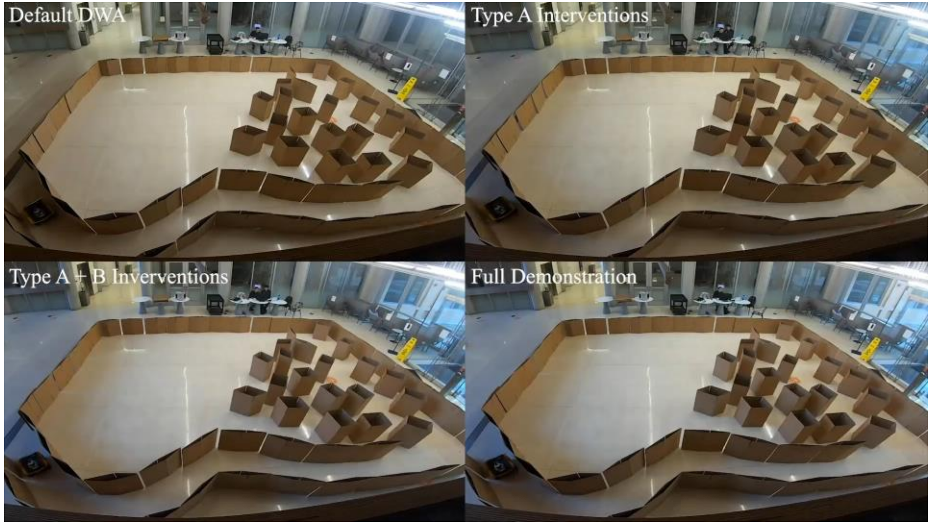


APPLI: Adaptive Planner Parameter Learning From Interventions

Real Experiment Results



Training Environment

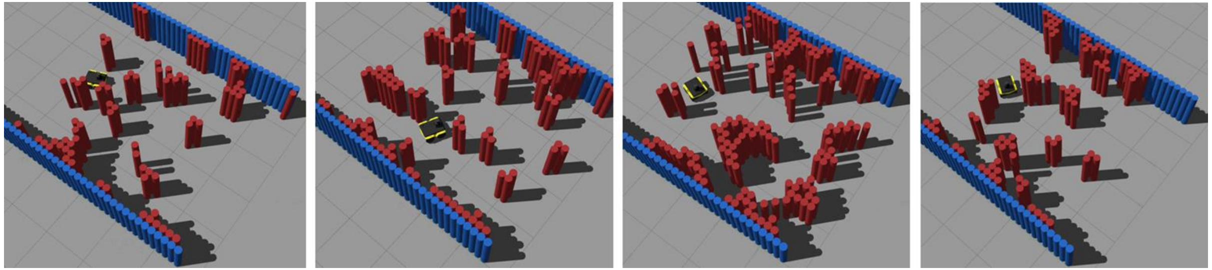


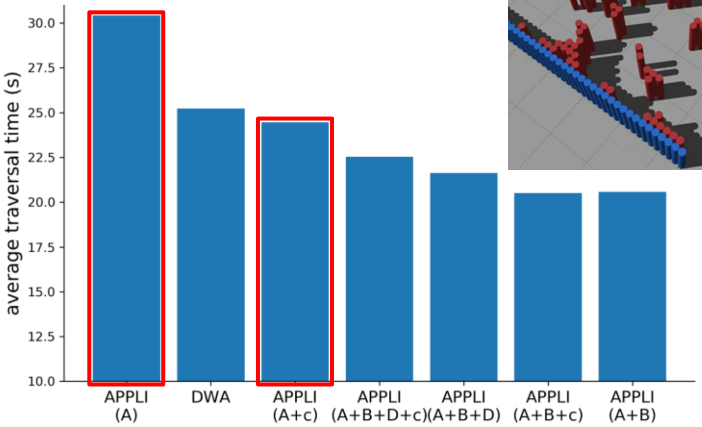
Unseen Environment

APPLI: Adaptive Planner Parameter Learning From Interventions

Simulated Experiment Results with Real-World Interventions

25K simulation trials in Benchmark Autonomous Robot Navigation (BARN) Dataset (Perille, et al., SSRR'20)





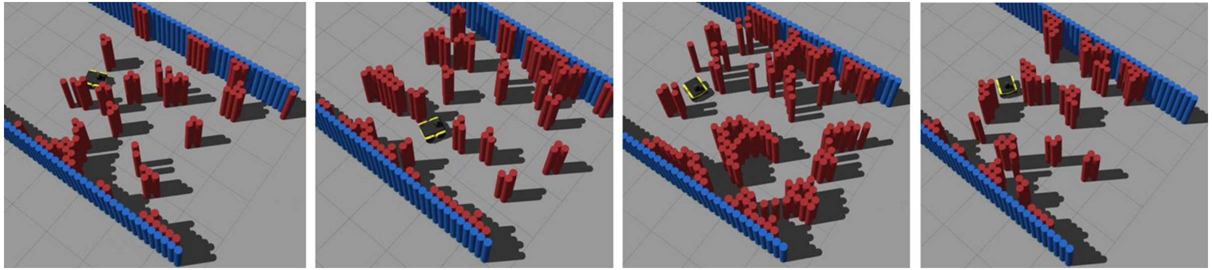
A: Type A interventions (failures)
B: Type B interventions (suboptimal behaviors)
D: Demonstrations in non-challenging scenarios
c: Use confidence measure in context predictor

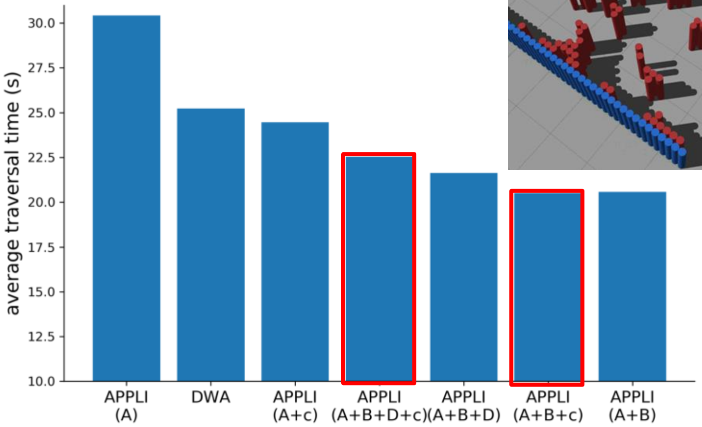
- Confidence is important when collected interventions are not sufficiently generalizable.

APPLI: Adaptive Planner Parameter Learning From Interventions

Simulated Experiment Results with Real-World Interventions

25K simulation trials in Benchmark Autonomous Robot Navigation (BARN) Dataset (Perille, et al., SSRR'20)



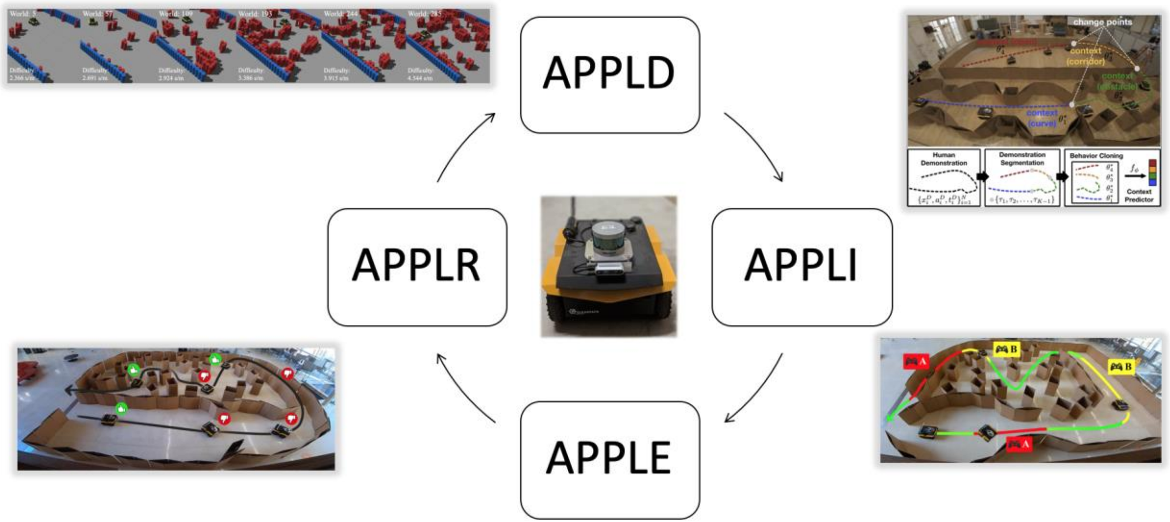


A: Type A interventions (failures)
B: Type B interventions (suboptimal behaviors)
D: Demonstrations in non-challenging scenarios
c: Use confidence measure in context predictor

- Confidence is important when collected interventions are not sufficiently generalizable.
- Unnecessary demonstrations may not help as they may be suboptimal (worse than default navigation system).

Future works

- Can we use other human interactions that are easier to collect, e.g., evaluative feedback?
- Can we integrate different approaches in APPL framework?

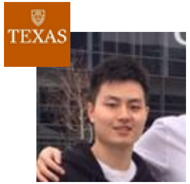


Cycle of Learning

Adaptive Planner Parameter Learning from Interventions

Zizhao Wang[1], Xuesu Xiao[1], Bo Liu[1], Garrett Warnell[2], and Peter Stone[1, 3]





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

Contact Information:
Zizhao Wang: zizhao.wang@utexas.edu
Xuesu Xiao: xiao@cs.utexas.edu

Link to the Paper: https://www.cs.utexas.edu/~xiao/papers/appli.pdf

Scan to access the paper