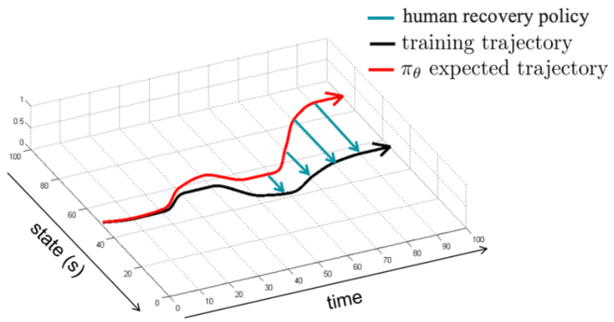
Robust Imitation Learning for Robotics Tasks through Planning Mode Switching

Hristo Todorov, Joey Hejna, Dorsa Sadigh

Stanford University

Background

Imitation learning has been a popular method for training machine learning algorithms for robotics, as, unlike reinforcement learning, a reward function does not have to be defined, and a given policy is trained to simply mimic expert behavior. However, one of the most serious problems it suffers from is the problem of distribution shifts, where small errors during evaluation time compound over time, leading to usually previously unobserved states, that could often lead to sub-optimal actions for achieving the goal.

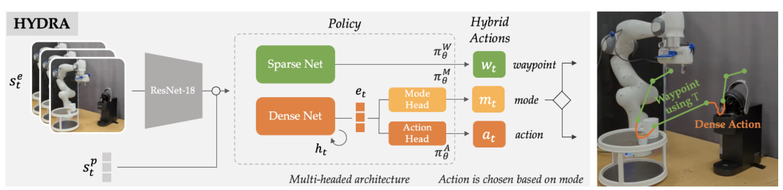


Most of the efforts to prevent distribution shifts could be separated into two categories:

- **Smaller action abstractions**: by reducing the dimensionality of the action space, we reduce the possible error size;
- **Reduce task horizon**: by executing less total number of actions, the errors are less likely to compound;

Related Work: Hydra

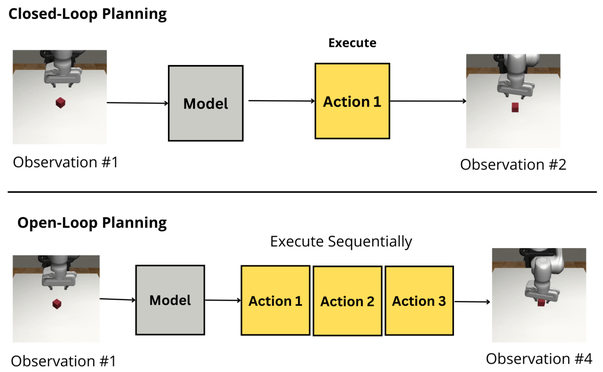
One approach that tackles both solutions to distribution shifts is Hydra. It switches between two modes of operation: sparse (not precise, e.g. get to object) and dense (precise, grab the object carefully). It uses smaller action abstractions for the sparse mode (waypoints) and reduces the horizon by combining multiple actions in a single waypoint.



Research Goal

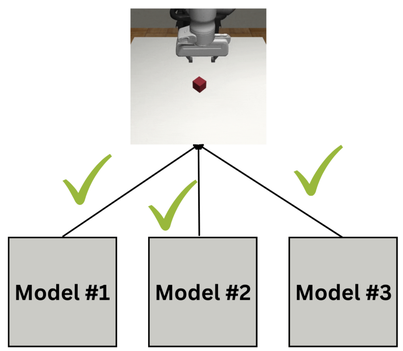
Can we reduce the horizon of robotics tasks by interpolating between open-loop and closed-loop planning based on current observation "difficulty"?

Open-Loop vs Closed-Loop



Observation Difficulty

As a metric of current observation difficulty, we use an ensemble of models. We compute the predictions of each model for the current observation and then compute the standard deviation of those predictions. The lower the standard deviation is, the easier we rank the observation.



Approach

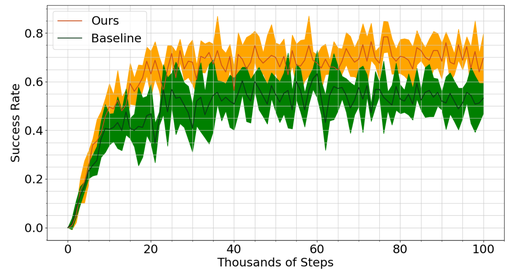
Use open-loop planning for "easy" observations and closed-loop planning for "hard" observations

Experiments & Results

Setup

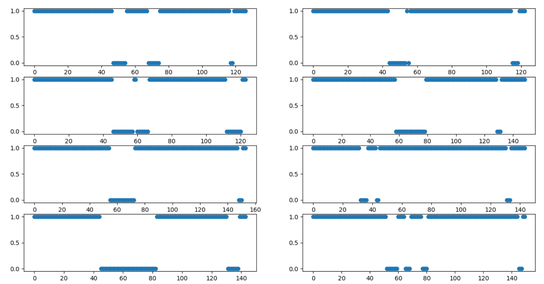
We evaluate our tasks on the Robomimic Square task. We use simple MLPs for the models (reproducing Robomimic paper) and use Behavioural Cloning (BC) to train both of our models.

Evaluation on Square



Dataset Difficulty Visualization

We show which parts of the datasets belong to which category: "easy" (0) or "hard" (1).



Continuous regions form within the dataset, which implies that certain sub-tasks are considered easier/harder than others.

Future Work

- Plot what the "easy"/"hard" actions look like in a simulation;
- Evaluate our approach in different environments;
- Implement dynamic choice of a number of actions to predict during open-loop policy based on task difficulty;