# Abstraction-Guided Policy Recovery from Expert Demonstrations

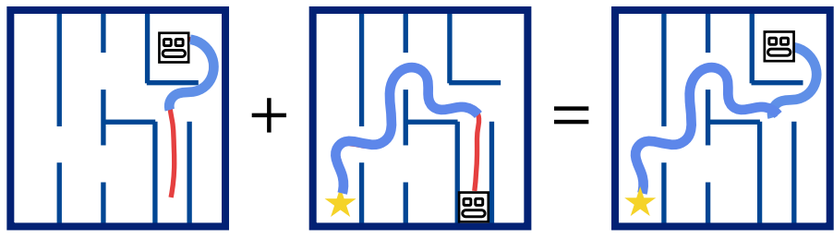
Canmanie T. Ponnambalam, Frans A. Oliehoek, Matthijs T. J. Spaan

Delft University of Technology, The Netherlands

*Learning from expert data:* Given expert demonstrations, copy the behavior of the expert online without any additional learning or exploration.


Which action should the agent take from a state that is not in the expert demonstrations?

ABSTRACTION-GUIDED POLICY **RECO**VERY

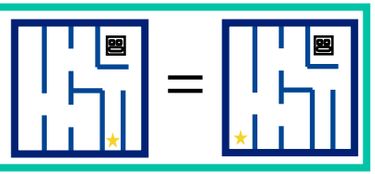


**RECO** uses abstraction to recover a lost agent back to a known state

**Problem setup:** We only have expert demonstrations, and do not know the task to be optimized. What if we end up in a state outside our expert demonstrations?

1. Define the recovery task: Reward is given when states in our expert trajectories are reached, the transition function is defined by projecting expert transitions in abstract space back onto the state space.
2. Solve the recovery task: This produces a policy that takes us from unknown states to states in the expert trajectory.

**ABSTRACTION** By ignoring the location of the goal, we map states that only differ by the goal location to the same abstract state.

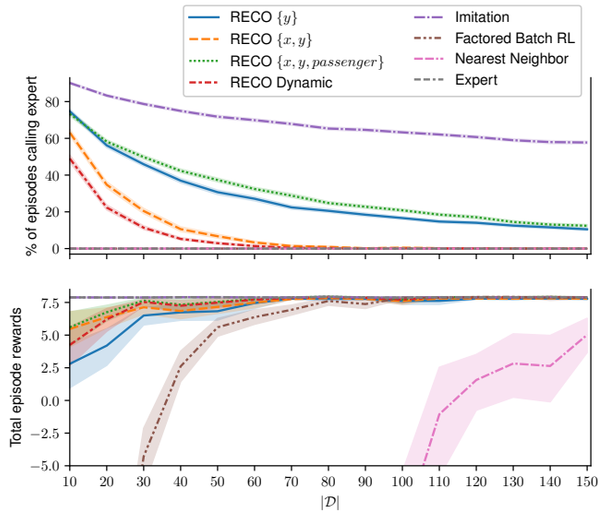


**Choosing abstractions:** In the example on the left, navigation through the maze is the same regardless of the goal location; i.e. we can ignore the goal location in the recovery task.

**Verifying the recovery policy:** The recovery task provides reward only if states in the expert trajectories are reached. Thus, our recovery policy will recover the agent from state *s* only if the value of the policy is greater than 0 at state *s*.

**Dynamic RECO:** Successful recovery depends on both the abstraction and the data given. When several abstractions are possible, we can define and solve multiple recovery tasks, increasing our chances of finding one that can recover the agent from a current state!

**RECO** results in a behavior cloning agent that **requires fewer expert trajectories** while **maintaining good performance**




Check out the paper for:

- Formalization of the recovery task
- Assumptions for correct abstractions
- Stochastic & continuous extensions

TU Delft