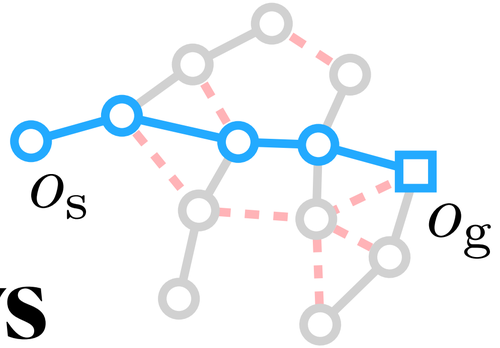


Learning from Graphical Replays

Ge Yang, Amy Zhang[1], Ari Morcos[1], Joelle Pineau[1], Pieter Abbeel[2], Roberto Calandra[1]

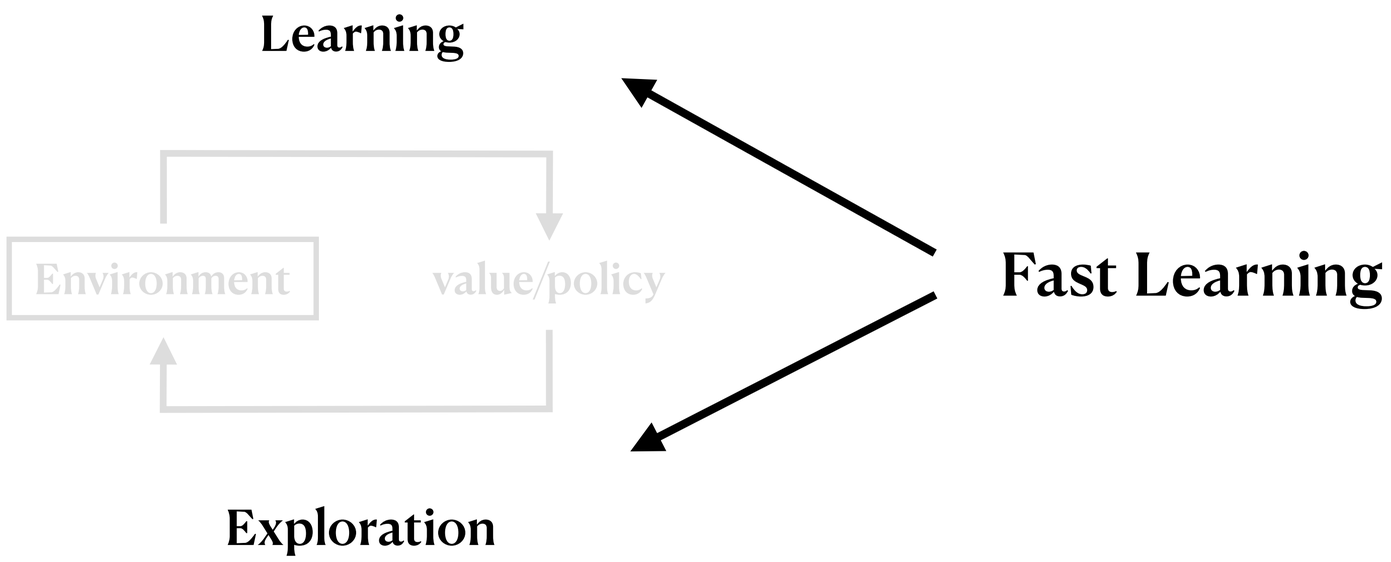
1. Facebook AI Research 2. UC Berkeley EECS

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

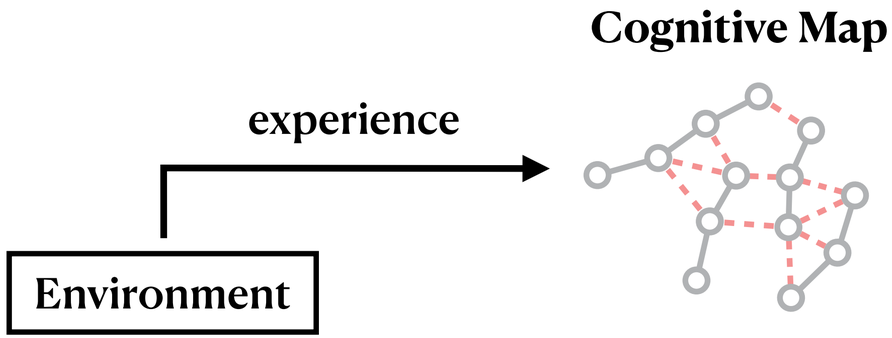
The ability to quickly adapt to a changing environment is a critical feature in general intelligence [1].

[1] Racanière et al, 2017

The ability to quickly adapt to a changing environment is a critical feature in general intelligence [1].



Topological Map as World Model



$\pi(a \mid o_s, o_g)$
value/policy

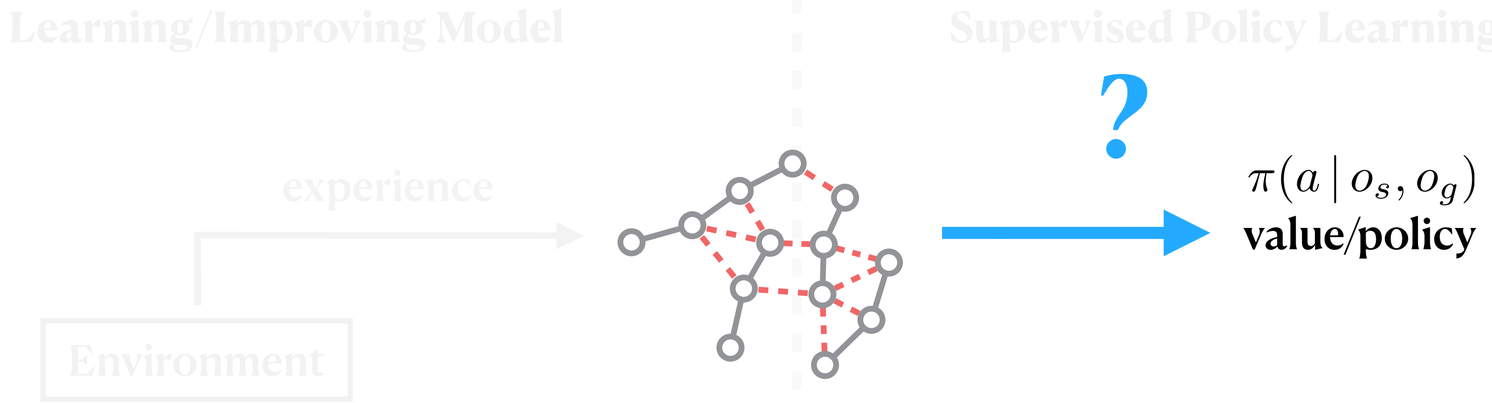
"**local**" policies

Episodic Control
+ Planning

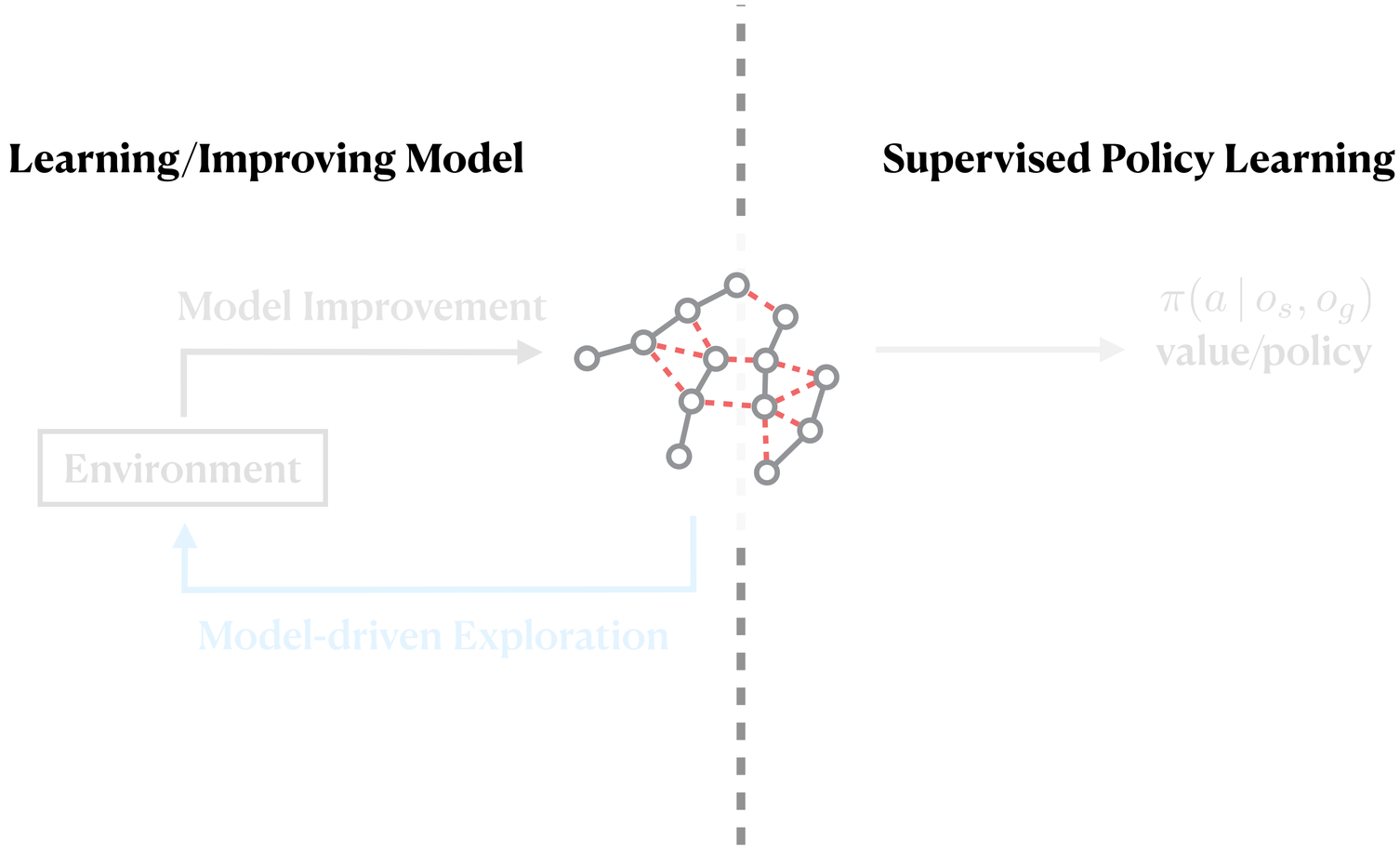
Great for adaptation

[2] Savinov et al, 2018

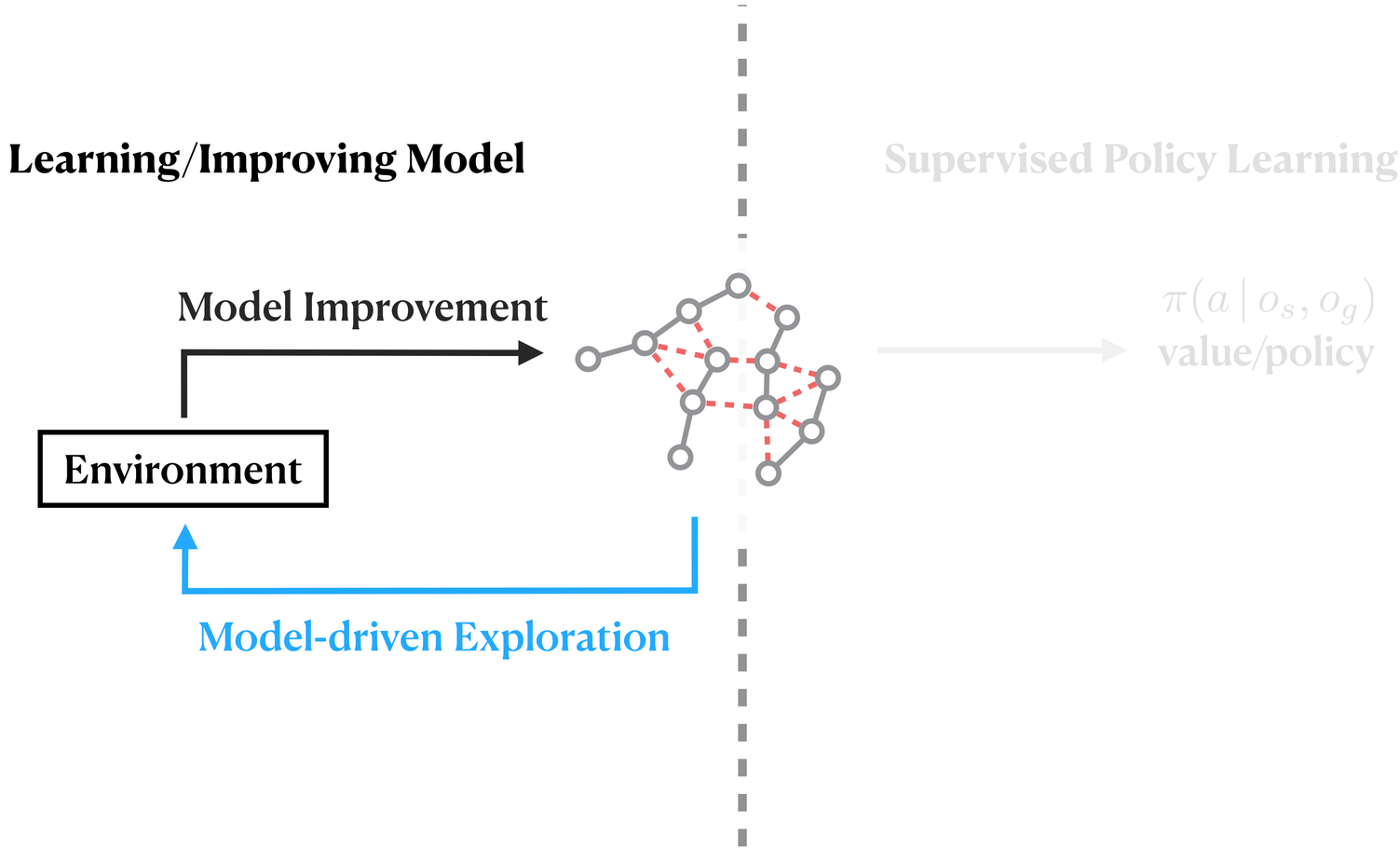
How can we learn long-horizon reactive policies in non-stationary environments?



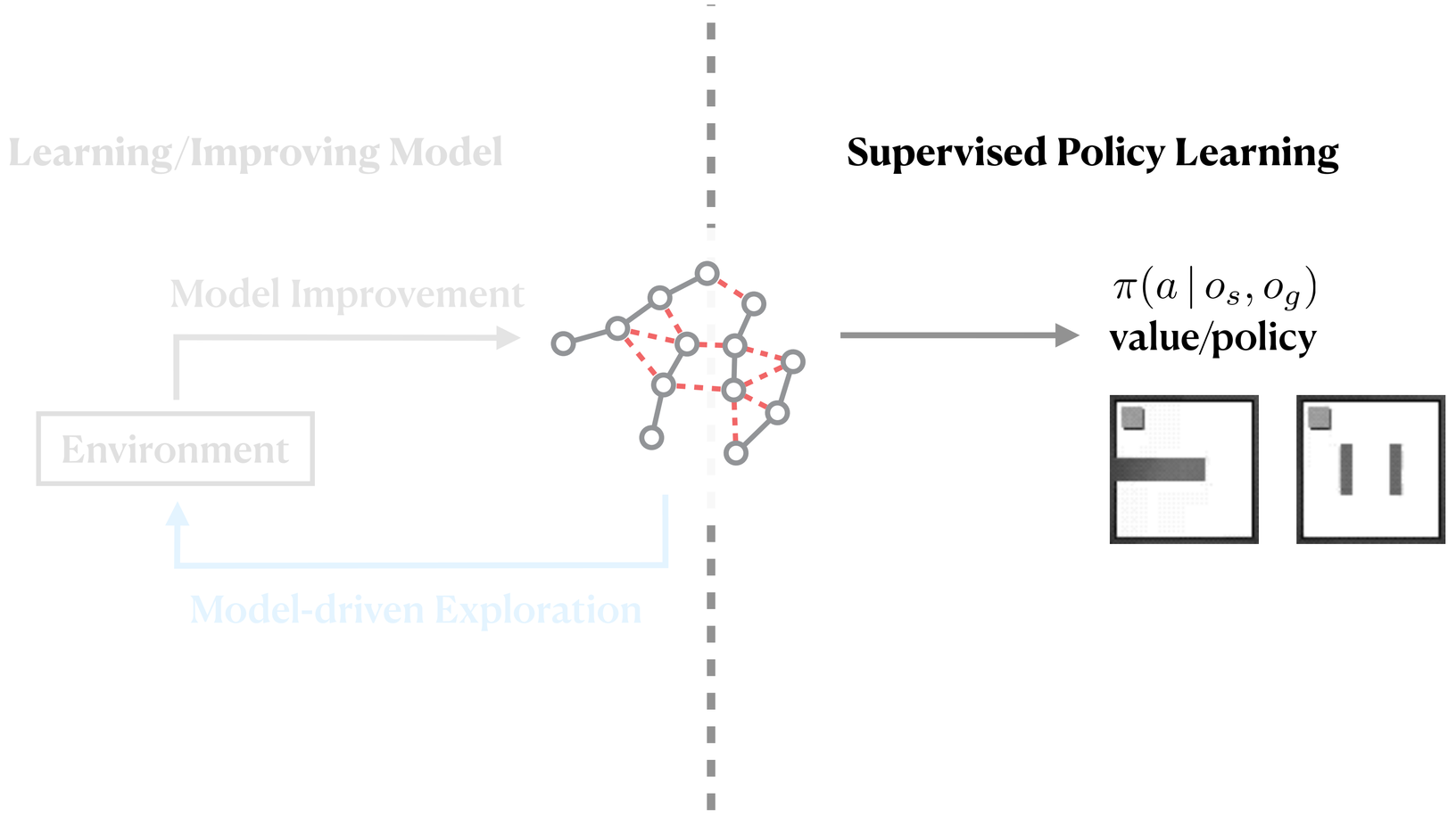
Learning from Graphical Replay (LfGR)



Learning from Graphical Replay (LfGR)

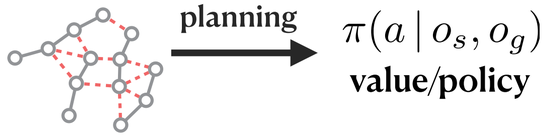


Learning from Graphical Replay (LfGR)

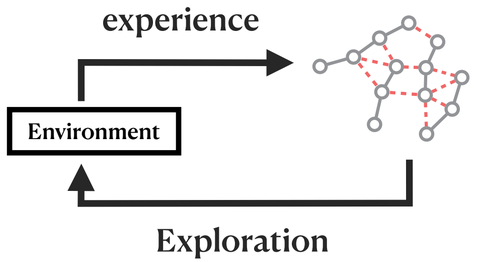


Learning from Graphical Replay (LfGR)

Policy Learning



Graph Improvement

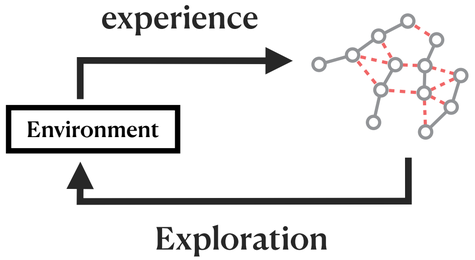


Learning from Graphical Replay (LfGR)

Policy Learning

1. Universal Value Prediction Network (**UVPN**)
2. Goal-Relabeled Expert Distillation (**GRED**)

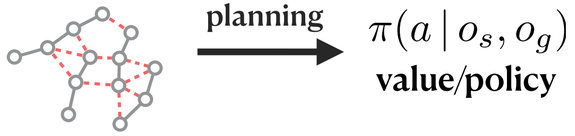
Graph Improvement



Learning from Graphical Replay (LfGR)

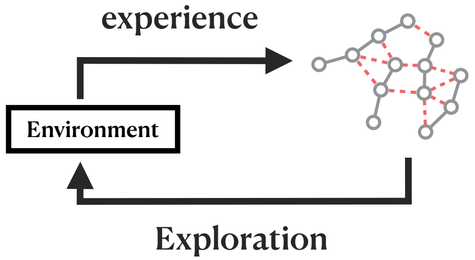
Policy Learning

1. Universal Value Prediction Network (**UVPN**)
2. Goal-Relabeled Expert Distillation (**GRED**)



Graph Improvement

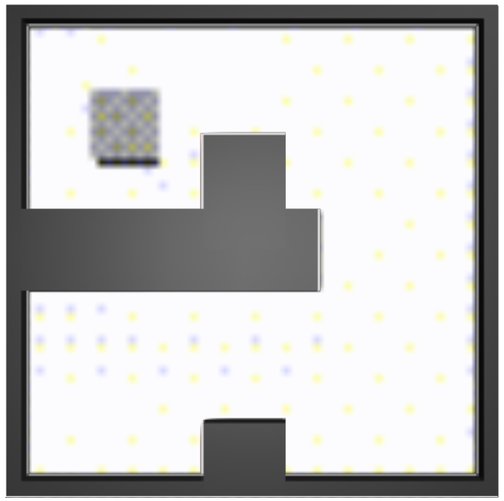
3. Causal Structure Learning on the Graph



Adapting to A Changing Environment

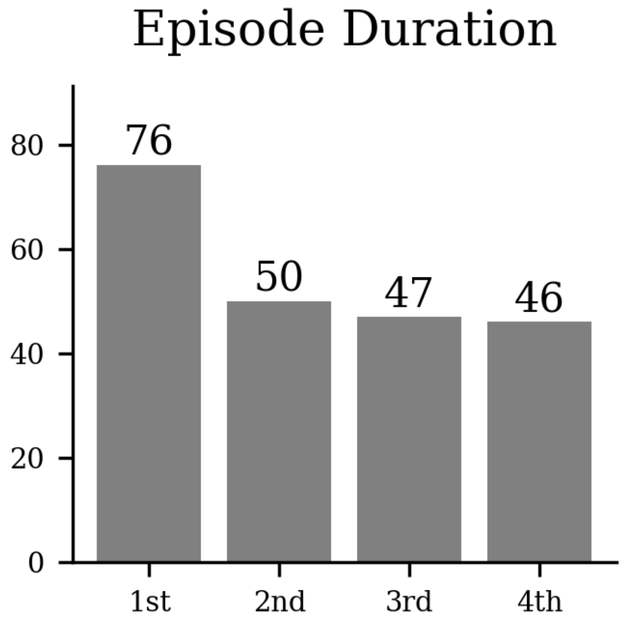
1. Graph enables robust episodic adaptation

Insert Walls



Model Update

Episode Duration

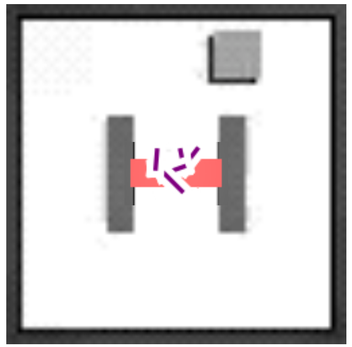


Sample Efficient Adaptation

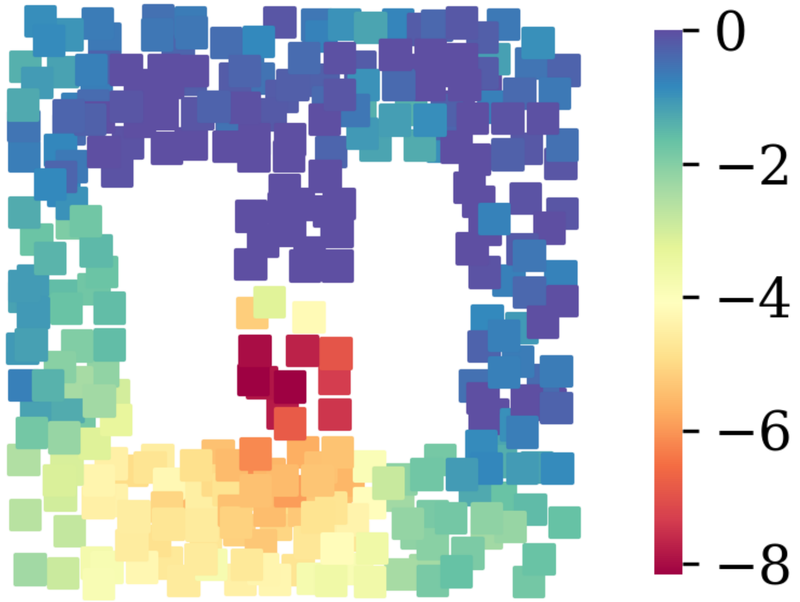
2. Learning from Graphical Model enables sample efficient adaptation

Before

Insert Wall



Δ max Q

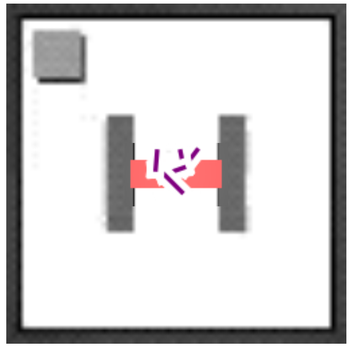


Sample Efficient Adaptation

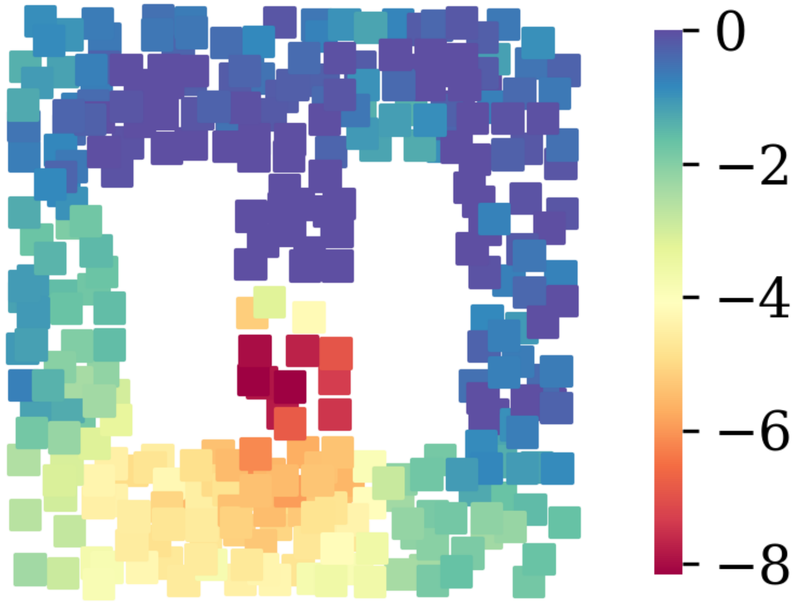
2. Learning from Graphical Model enables sample efficient adaptation

Before

Adapted

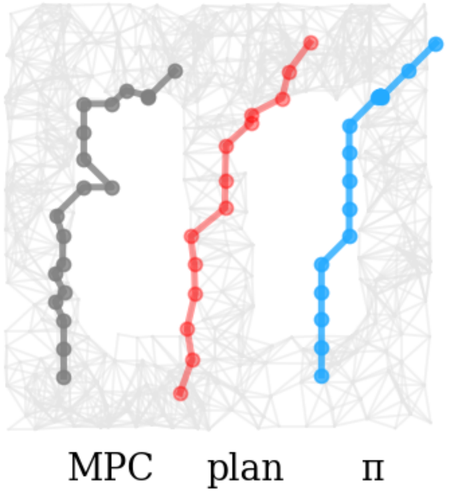


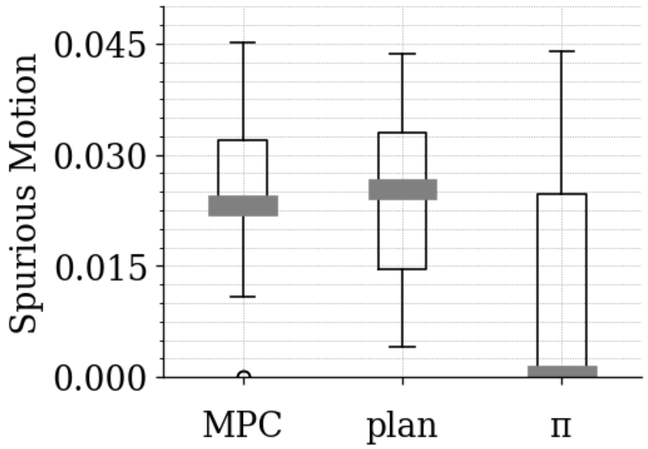
Δ max Q

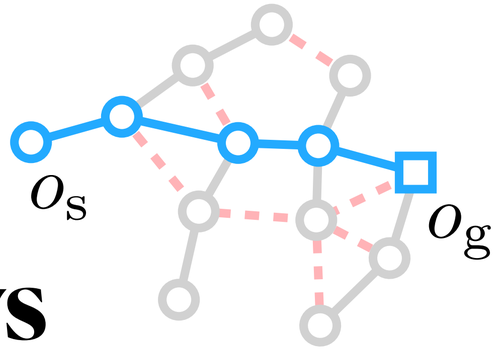


Generalization

3. Neural network policy enable generalization







Learning from Graphical Replays

Ge Yang, Amy Zhang[1], Ari Morcos[1], Joelle Pineau[1], Pieter Abbeel[2], Roberto Calandra[1]

1. Facebook AI Research 2. UC Berkeley EECS

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Key Takeaways

- We can learn long-horizon reactive policy from the graph
- Planning enable focused exploration
- Model Improvement is key