Zero-Shot Adaptation of Behavioral Foundational Models to Unseen Dynamics

M.Bobrin, I.Zisman, A.Nikulin, V.Kurenkov and D.Dylov

Intro & Motivation

➔ Behavioral Foundational Models (BFMs) are generalist RL agents that extract *optimal* policies to unseen tasks in zero-shot manner at test time

⛔ **However**, existing approaches **can not handle dynamics shifts** in data or work in POMDP setting

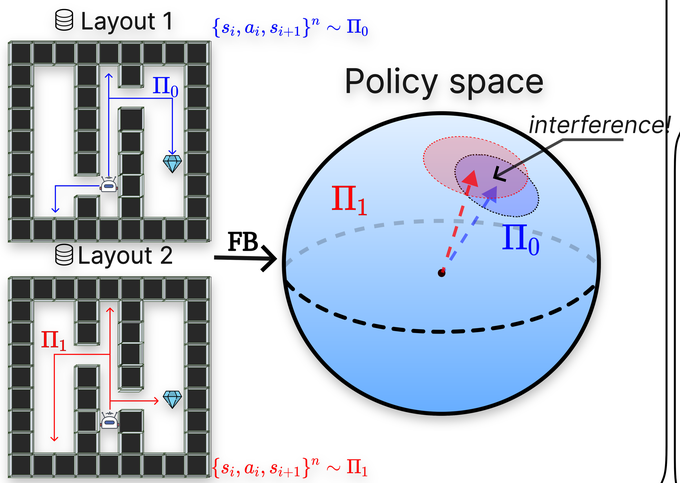
➔ This limits development of generalists agents, capable of adapting to both dynamics mismatches and tasks

➔ In our work, we focus on improving methods, based on **Forward-Backward** (FB) representation

Problem

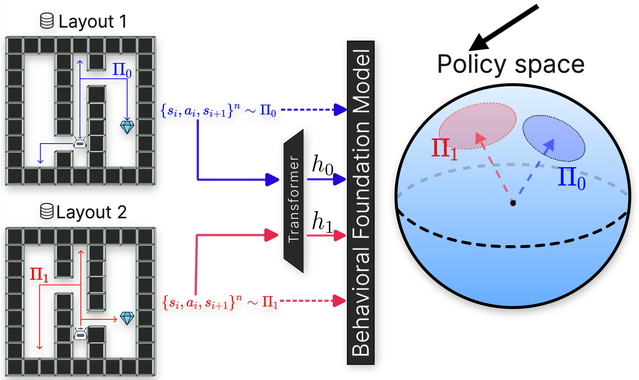
➔ When trained on offline data from different Contextual MDPs, possible futures average

➔ This leads to **interference** between policy representations π_z



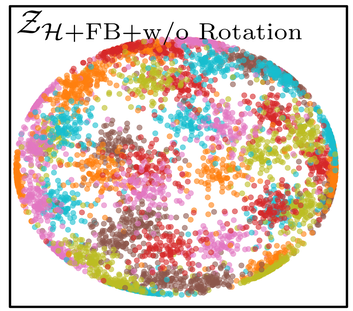
Proposed Solution

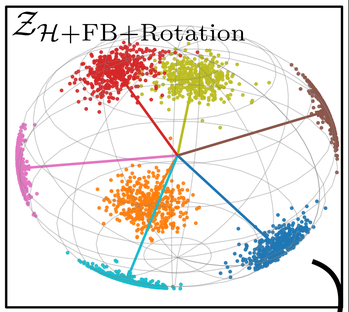
Idea: Infer context latent representations h from data and condition FB: $F(s, a, z^{\pi}_{FB}, h)^T B(s')$



(1) Belief-FB (BFB)

Conditioning on $h = f_{dyn}(\{s_i, a_i, s_{i+1}\})$ **disentangles** and **clusters** space π_z





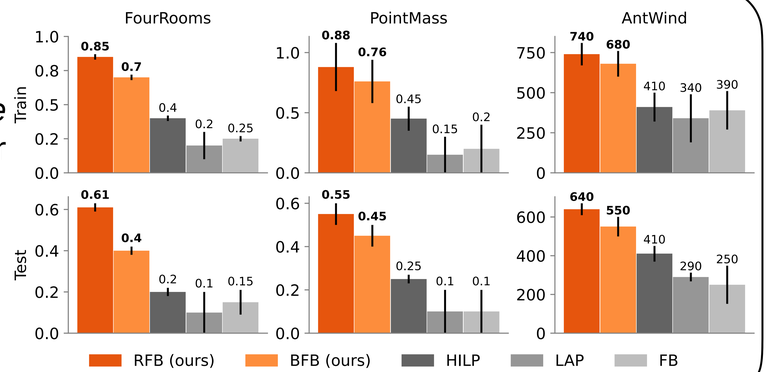
(2) Rotation-FB (RFB)

Even more structure with **dynamics specific policy-regions**!
(Same color = same MDP, points = policies)

❗ Policies diversity can be controlled!

Experiments

➔ **RFB** and **BFB** *enable generalization* under both seen and unseen dynamics mismatches, maintaining zero-shot abilities!



ICLR

Project Website