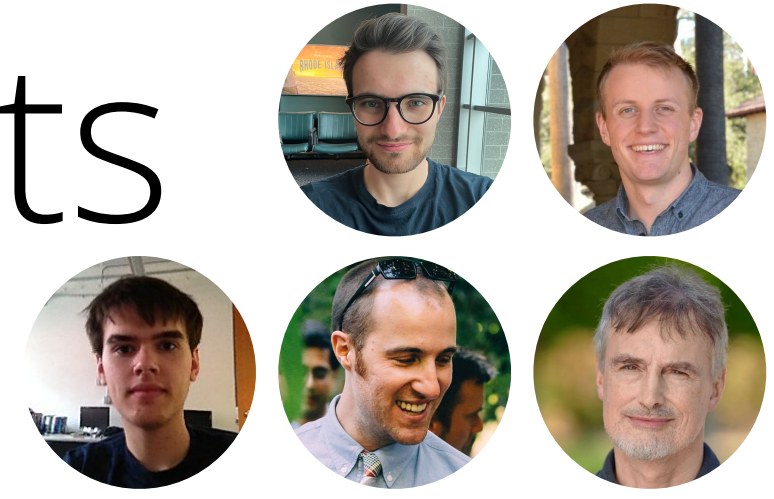


Towards General-Purpose In-Context Learning Agents

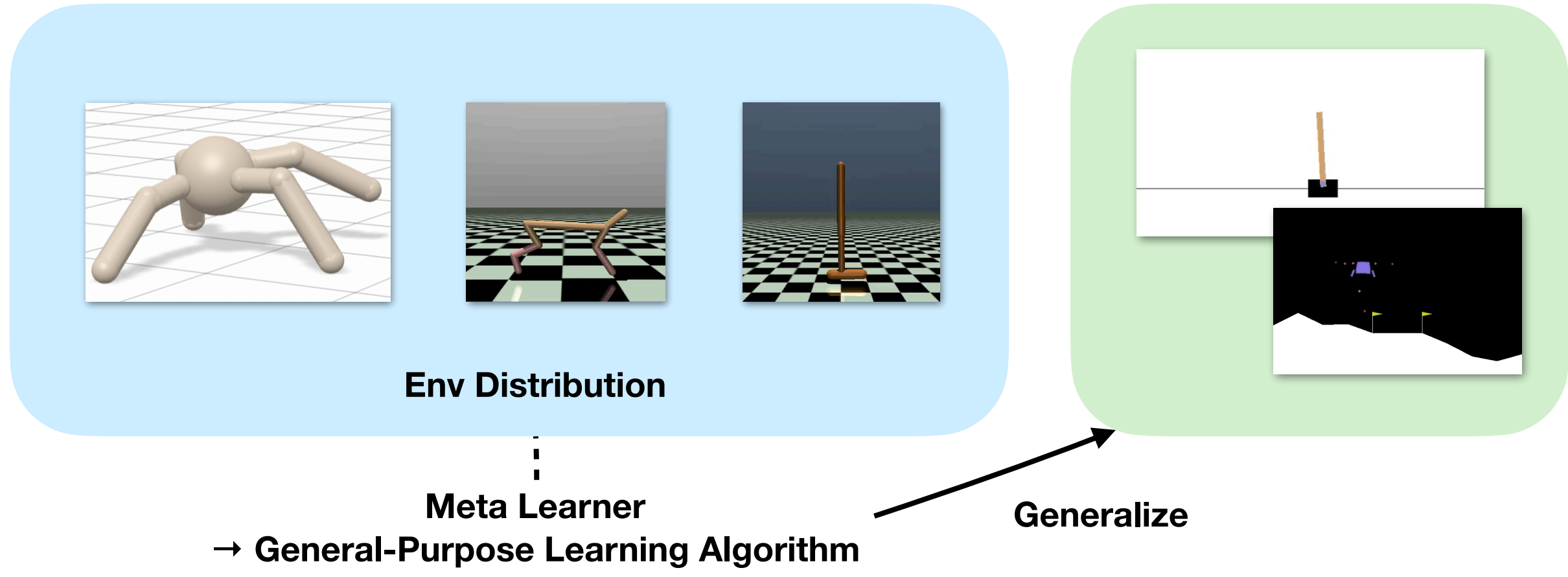
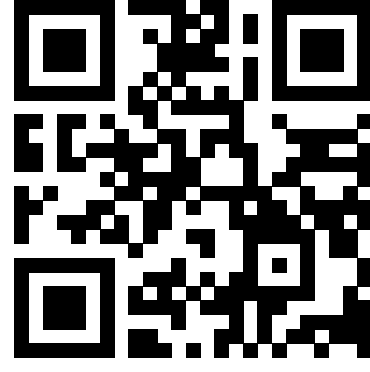


@LouisKirschAI

Louis Kirsch¹, James Harrison², C. Daniel Freeman²,
Jascha Sohl-Dickstein², Jürgen Schmidhuber^{1,3}

Motivation

General-Purpose Meta Learning
Can we train an agent that can efficiently in-context **learn and act in any environment?**



In-Context Learning

MetaGenRL, LPG, etc

In-context meta-RL

$$\{(s_i, a_i, r_i)\}_{i=1}^{N_D}$$

$$\{s_i, a_i, r_i\}_{i=1}^{N_D}$$

Gray-box Learning
 $\phi \leftarrow \phi - \nabla_{\phi} L_{\alpha}$

Black-box function approximator
e.g. LSTM, Transformer, FWP
 $\pi(a | s, \{(s_i, a_i, r_i)\}_{i=1}^{N_D})$
= in-context learning

Better π_{ϕ}

Better π

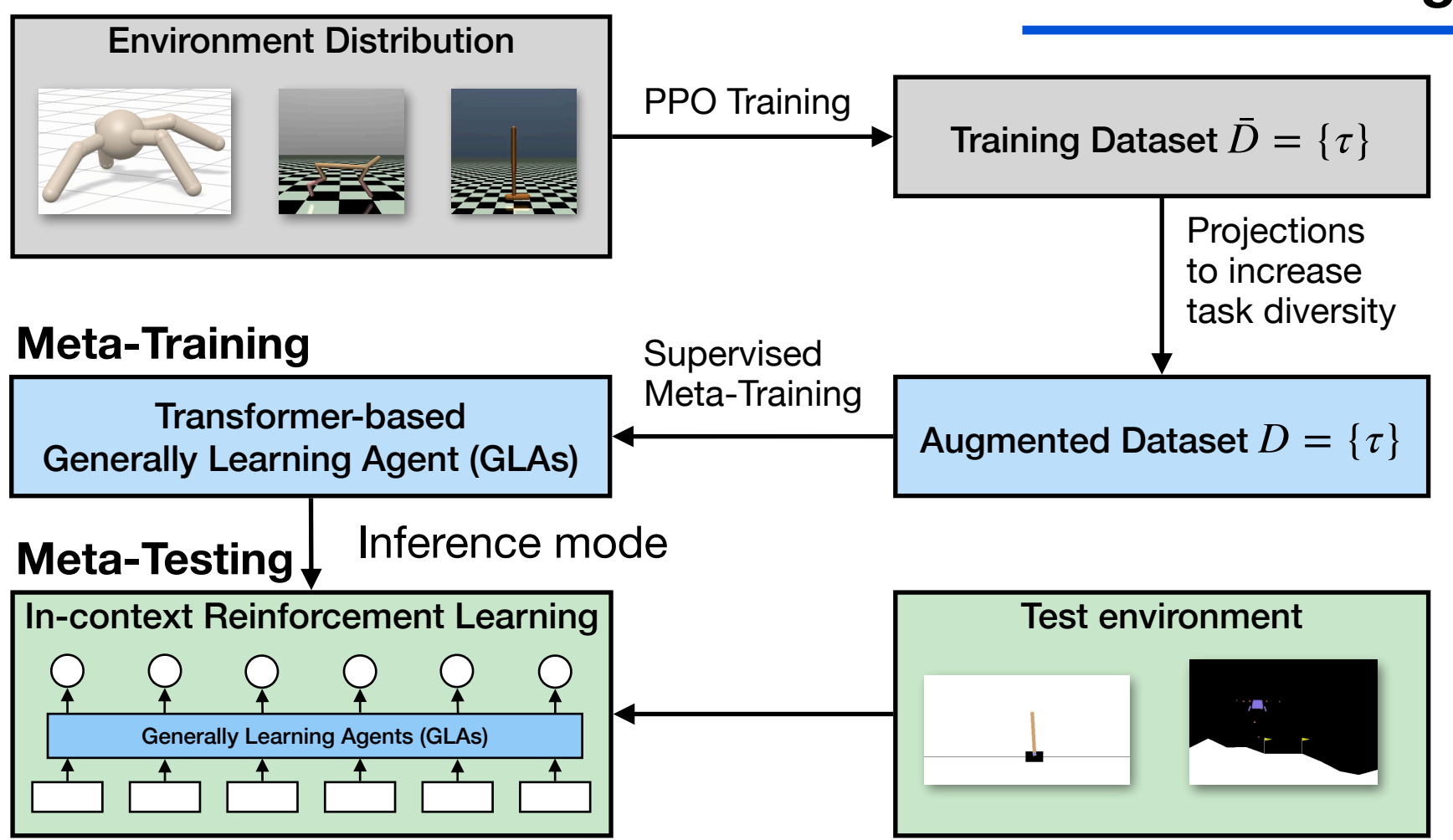
Good Generalization

Difficult Generalization
This paper: How to fix this?

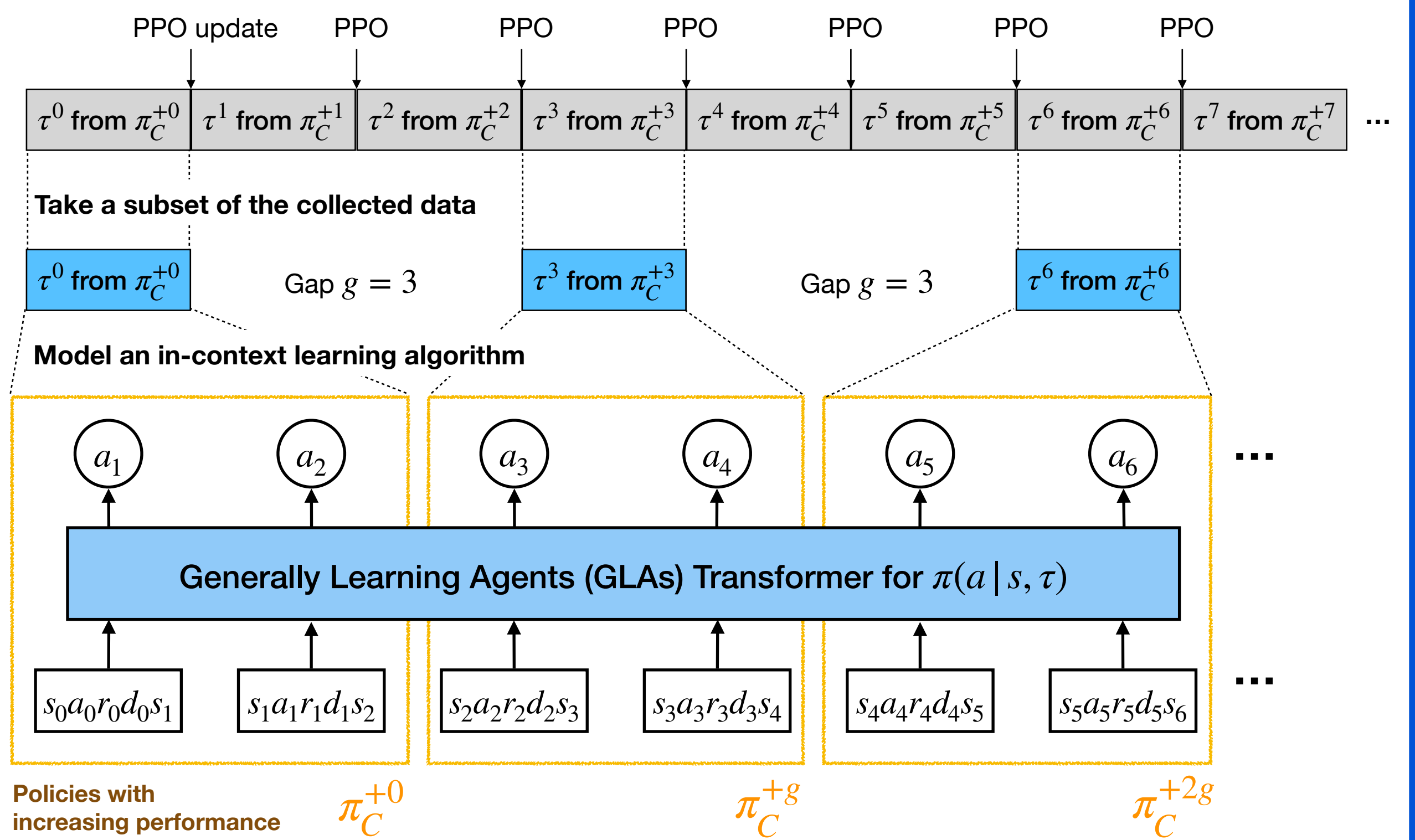
Generally Learning Agents (GLAs)

Data Collection

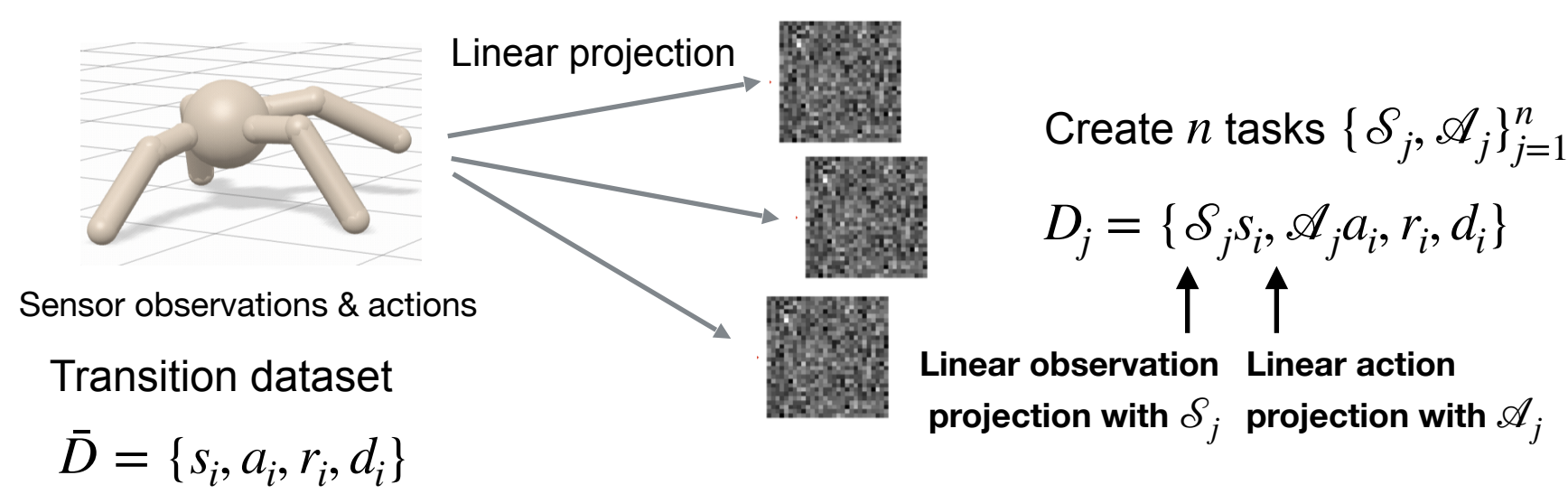
Offline training + augmented RL data → Generalization



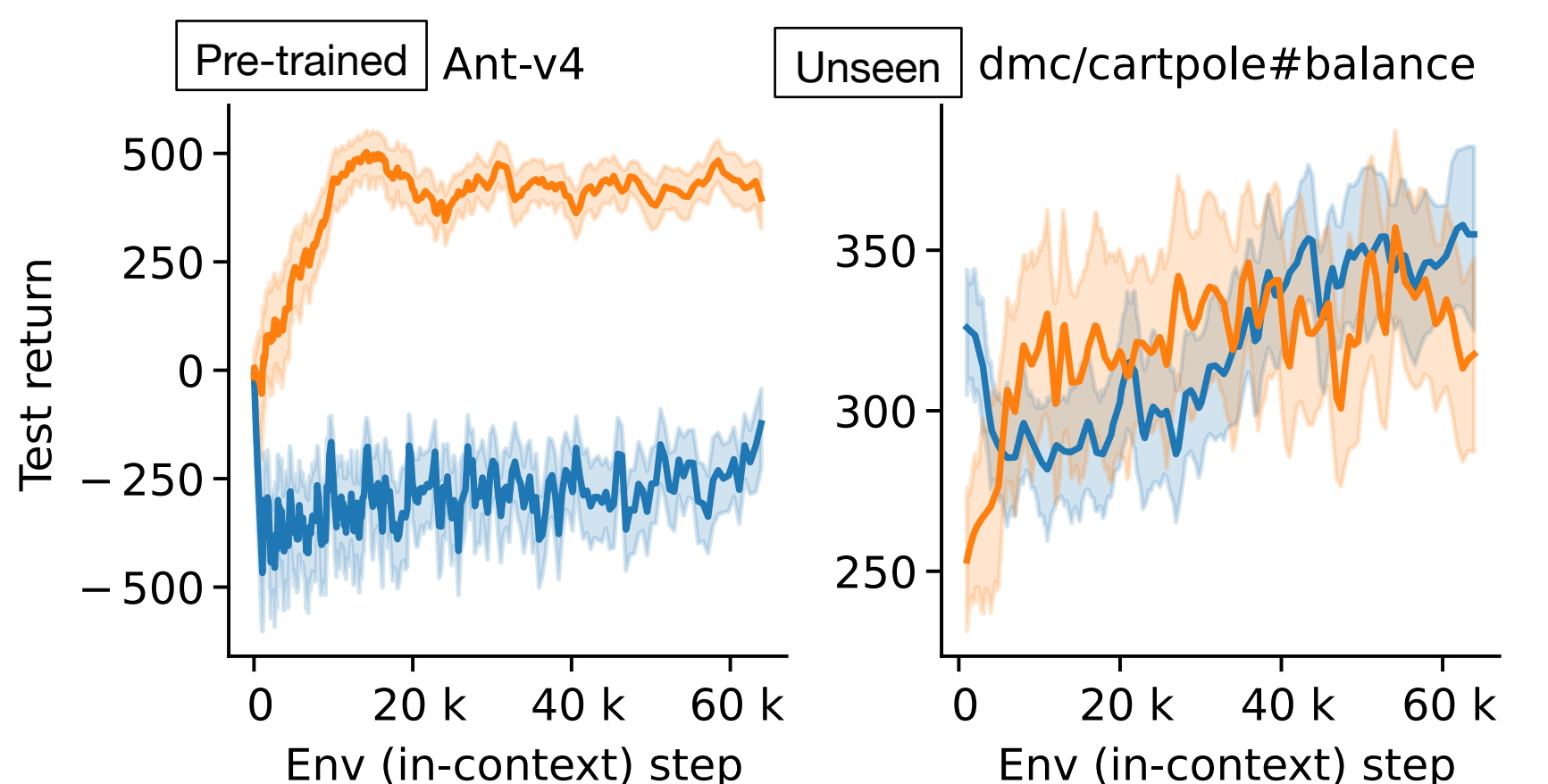
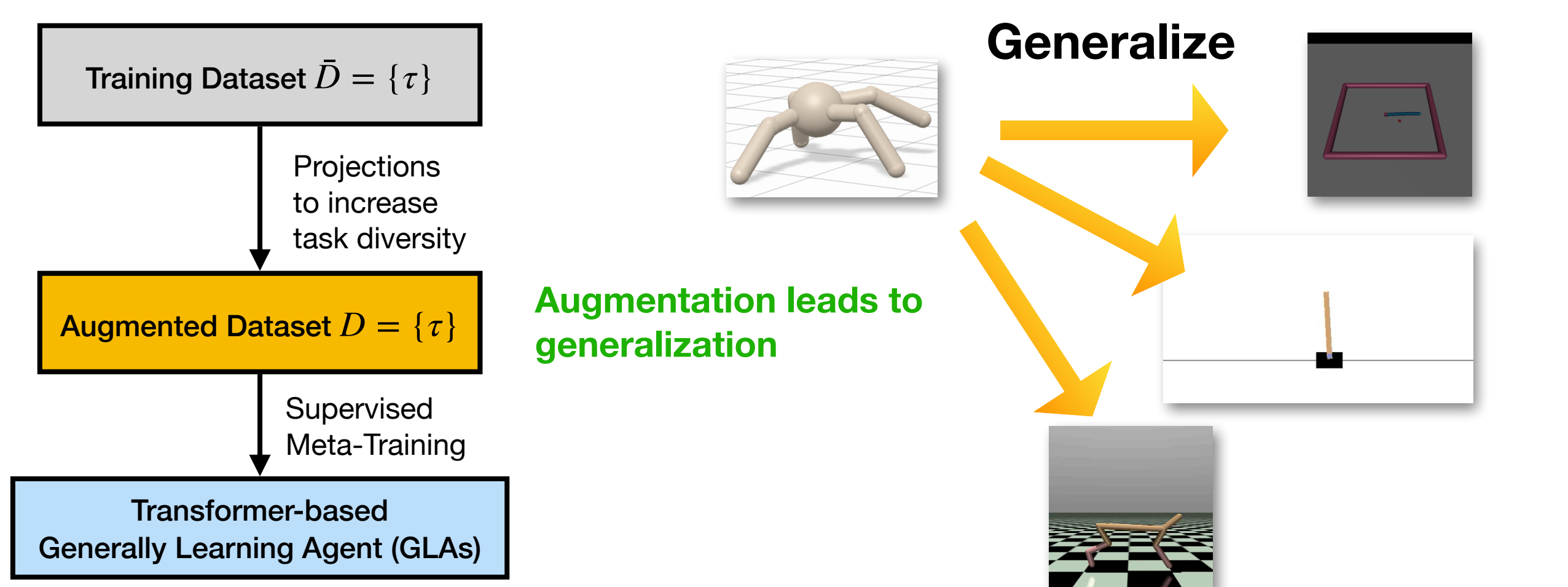
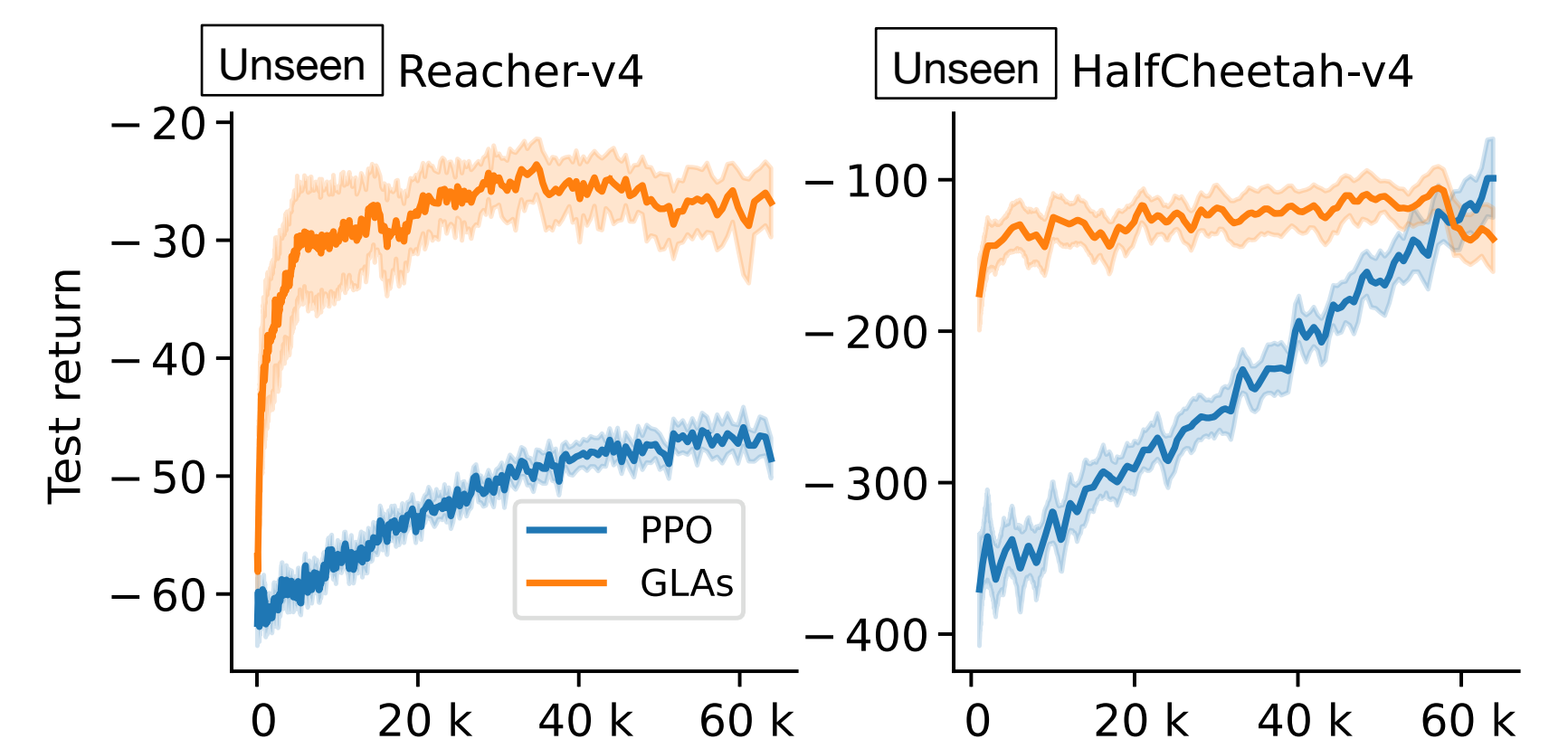
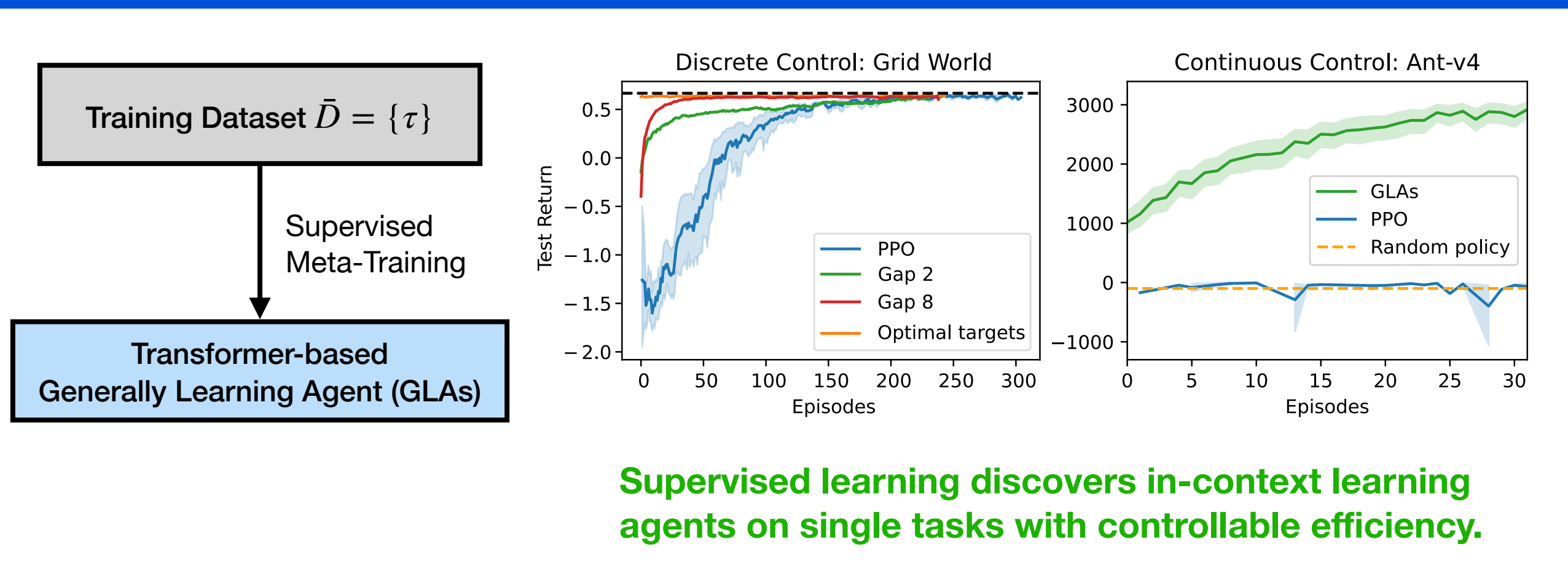
Data Collection (like Laskin et al 22)



Use data augmentation for generalization



Results



GLAs generalize to novel domains via in-context RL.