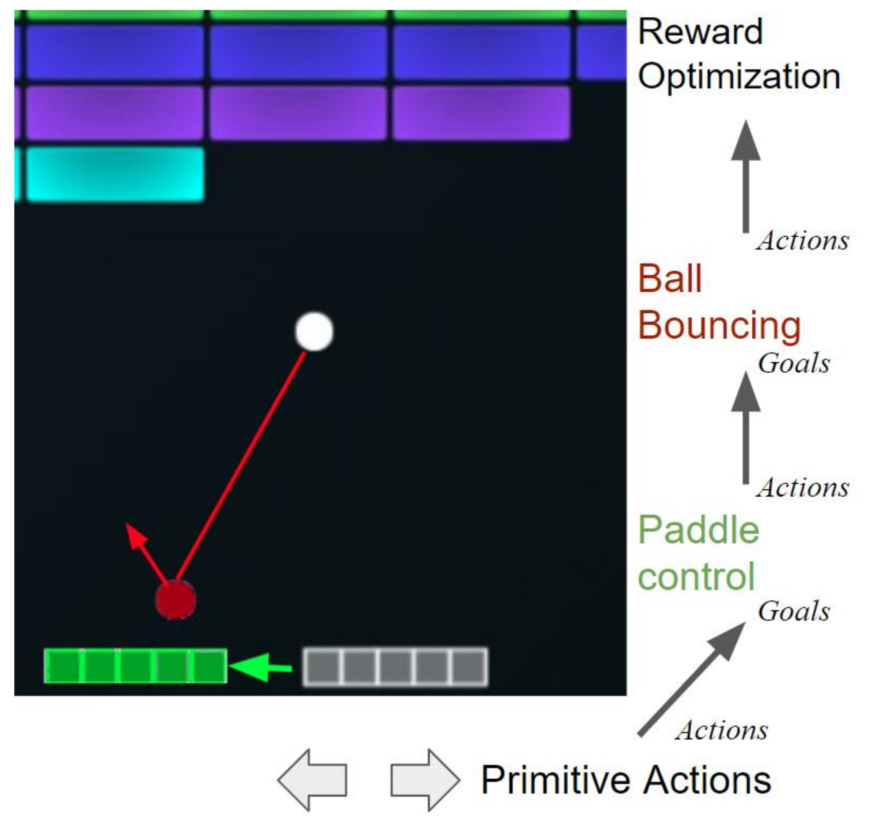
Granger Causal Interaction Skill Chains

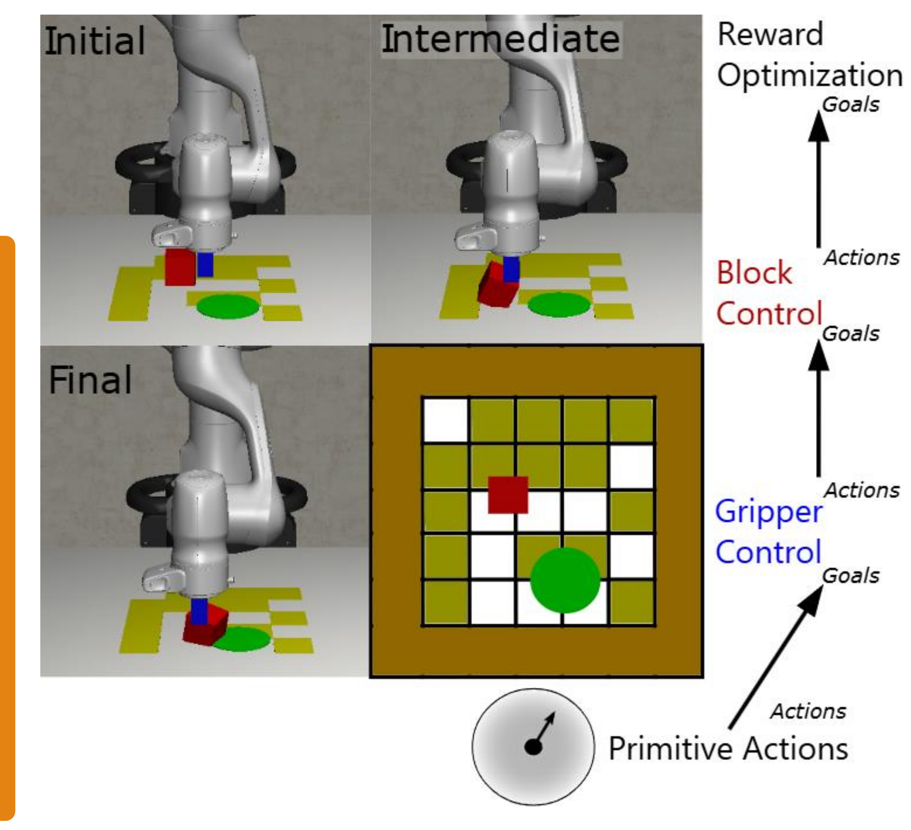
Caleb Chuck*, Kevin Black, Aditya Arjun, Yuke Zhu*, Scott Niekum**

Code: https://github.com/CalCharles/object-options Contact: calebc@cs.utexas.edu Paper:

Overview

Core Takeaway: Goal-conditioned skills that induce **interactions** offer improved **sample efficiency, overall performance** and **transferability** in long-horizon factored environments.



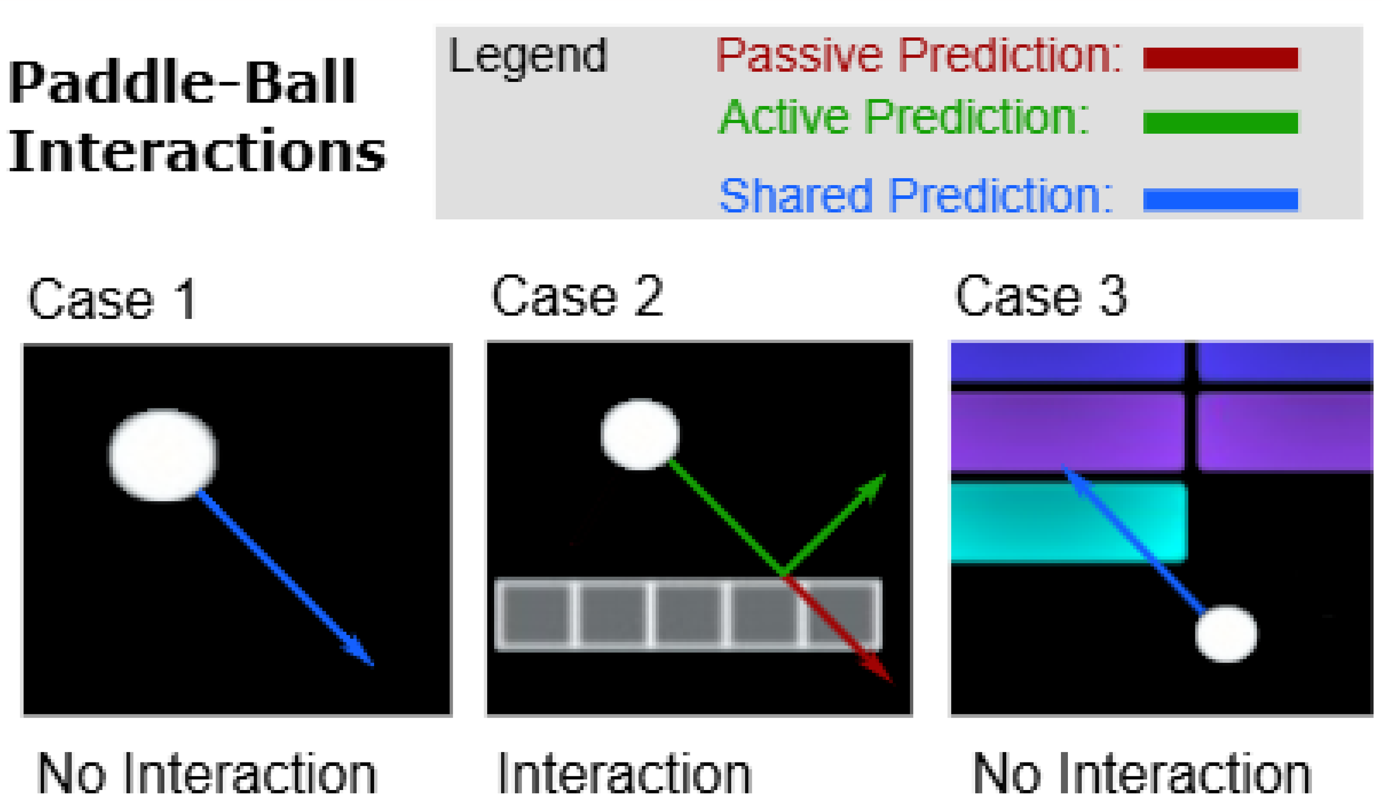


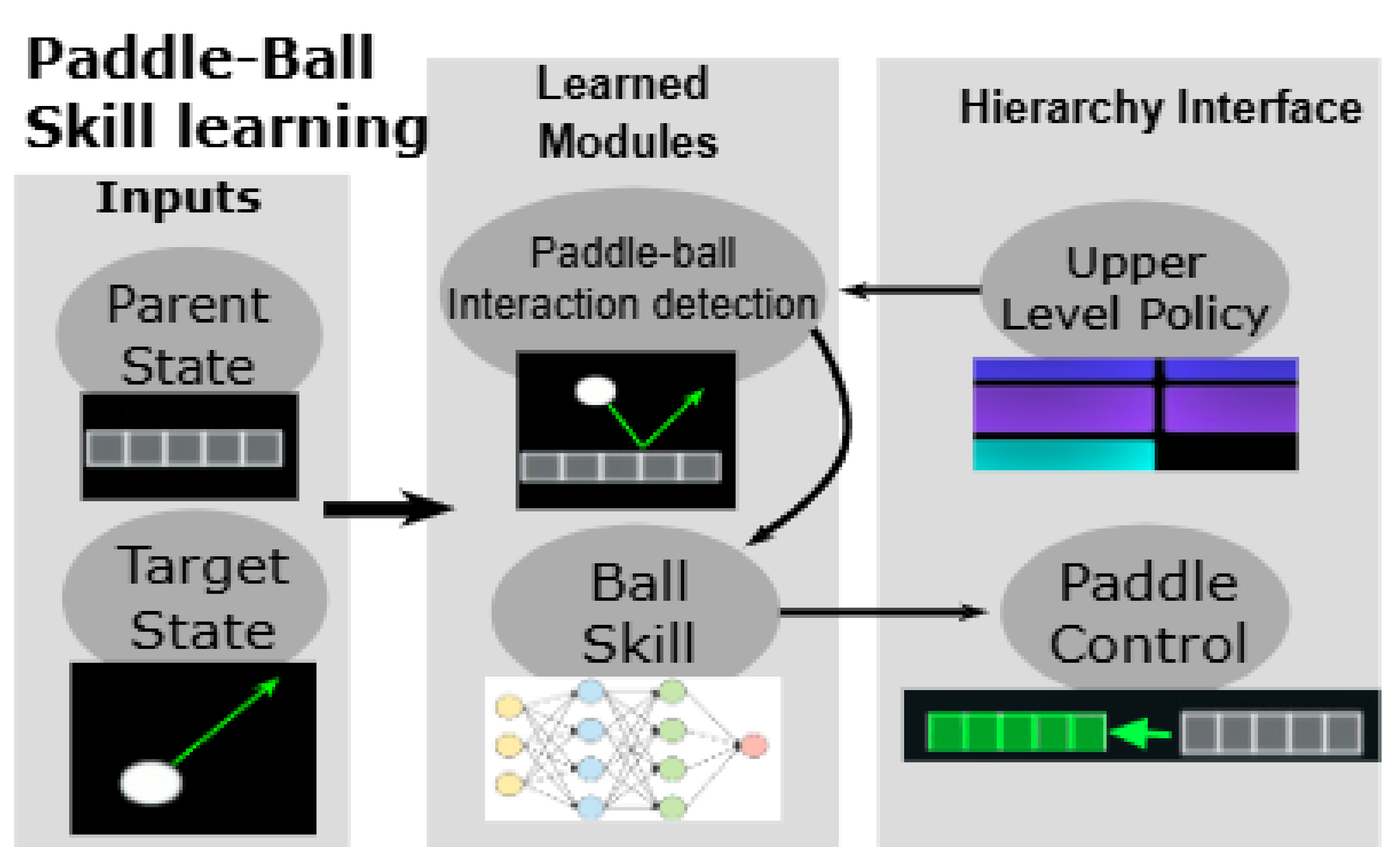
Step 1
Learn Granger-Causal factored dynamics models: "passive" autoregressive model and "active" pairwise model.

Step 2
Identify desired interactions for skill goals: states with disagreement between the passive and active models.

Step 3
Learn Goal-conditioned policy with hindsight: reward desired interactions using previously learned policies as actions

Granger Causal Interactions





Interaction Definition

$$m^{\mathrm{pas},G}(\mathbf{s}_b^{(t-w)},\ldots,\mathbf{s}_b^{(t-1)};\theta) = \theta^0 + \left[\sum^{w} \theta^i \mathbf{s}_b^{t-i}\right] + \epsilon_b$$

$$m^{\mathrm{act},G}(\mathbf{s}_a^{(t-w)},\ldots,\mathbf{s}_a^{(t-1)},\mathbf{s}_b^{(t-w)},\ldots,\mathbf{s}_b^{(t-1)};\psi_b,\psi_a) = \psi^0 + \left[\sum_{i=1}^{w} \psi_b^i \mathbf{s}_b^{t-i} + \psi_a^i \mathbf{s}_a^{t-i}\right] + \epsilon_a$$

$$\ell_{\mathrm{pas}}(\mathbf{s}_a,\mathbf{s}_b';\theta) := \log m^{\mathrm{pas}}(\mathbf{s}_b;\theta)[\mathbf{s}_b']$$

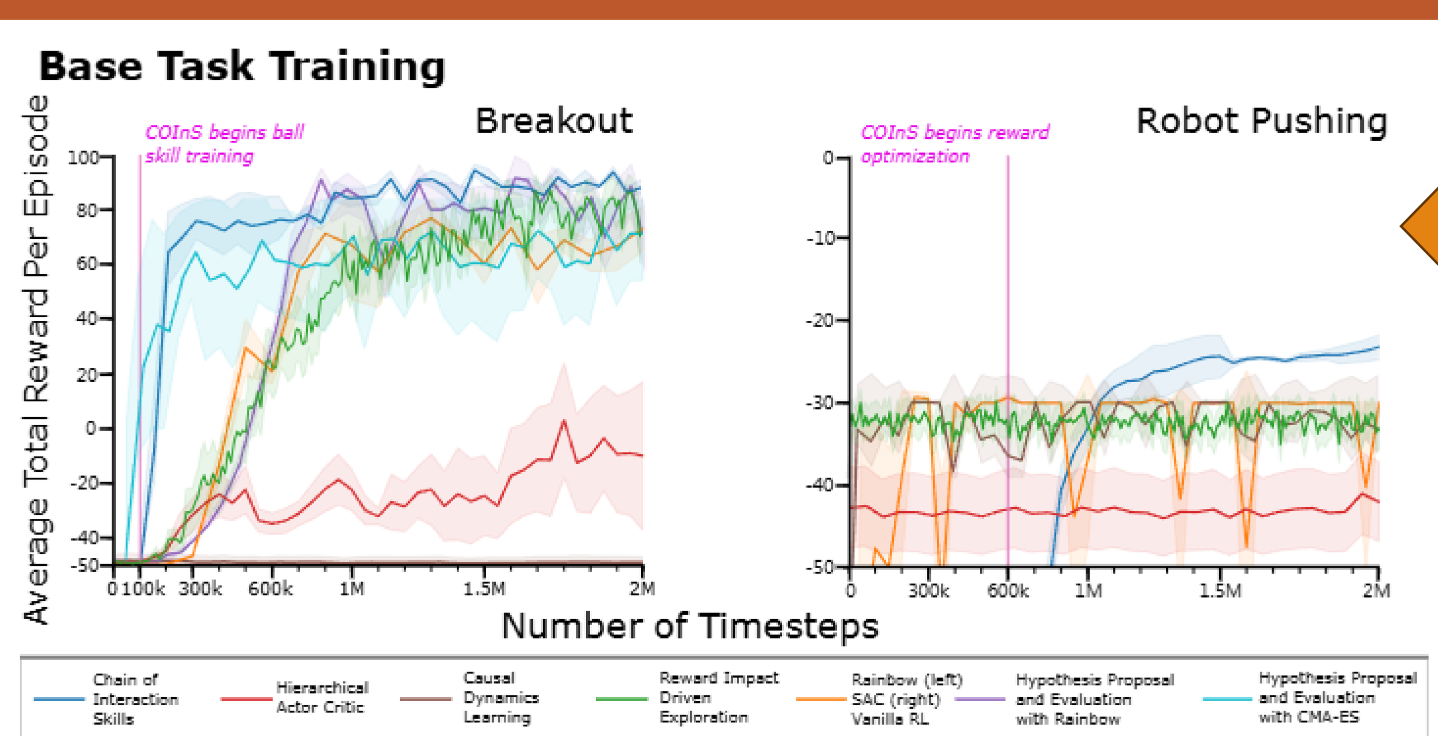
$$\ell_{\mathrm{act}}(\mathbf{s}_a,\mathbf{s}_b,\mathbf{s}_b';\psi) := \log m^{\mathrm{act}}(\mathbf{s}_a,\mathbf{s}_b;\psi)[\mathbf{s}_b']$$

$$h_{a,b}(\mathbf{s}_a,\mathbf{s}_b,\mathbf{s}_b';\psi,\theta) := (\ell_{\mathrm{act}}(\mathbf{s}_b,\mathbf{s}_a,\mathbf{s}_b';\psi) > \epsilon_{\mathrm{act}}) \wedge (\ell_{\mathrm{pas}}(\mathbf{s}_b,\mathbf{s}_b';\theta) < \epsilon_{\mathrm{pas}})$$

Skill Termination Condition

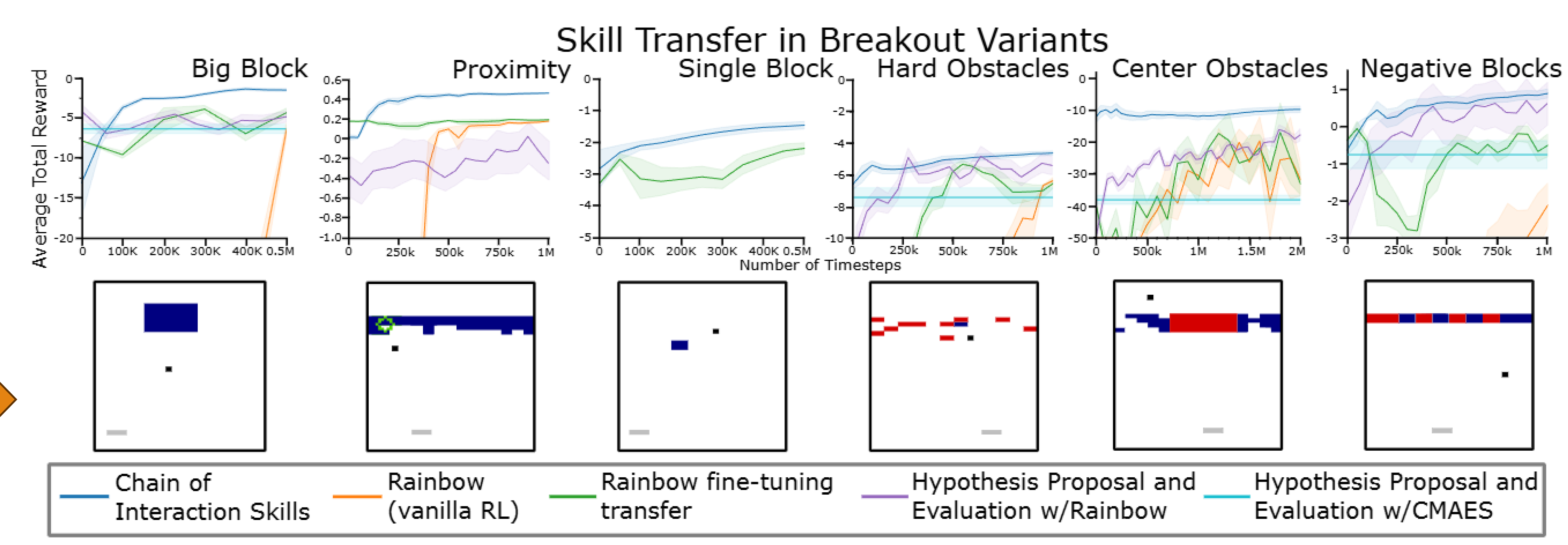
$$\phi_b(\mathbf{s}_a,\mathbf{s}_b,\mathbf{s}_b',\mathbf{c}_b) := \begin{cases} 1 & h_{a,b}(\mathbf{s}_a,\mathbf{s}_b,\mathbf{s}_b';\psi,\theta]) \wedge \|\eta_b \mathbf{s}_b' - \mathbf{c}_b\|_1 < \epsilon_c \\ 0 & \text{otherwise.} \end{cases}$$

Evaluation



Sample Efficiency and Overall Performance on long-horizon tasks

Skill transfer to challenging in-domain tasks



SCALAR LAB
MIDI lab
RPL Robot Perception and Learning Lab
TEXAS The University of Texas at Austin
TEXAS Robotics
University of Massachusetts Amherst

Affiliations:

- *UT Austin, ** Also UMass Amherst