Reinforced Grounded Action Transformation for Sim-to-Real Transfer

Siddharth Desai
UT Austin

Haresh Karnan
UT Austin

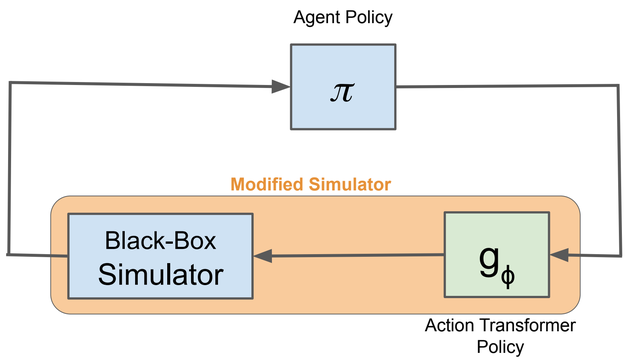
Garrett Warnell
Army Research Laboratory
UT Austin

Josiah P Hanna
School of Informatics
The University of Edinburgh

Peter Stone
UT Austin
Sony AI

Abstract

- Policies trained in simulation often do not perform as well in the real world.
- Prior grounded simulation learning algorithms like **GAT** use an action transformer to ground the simulator's transition dynamics to the real world.
- We show that learning the action transformer end-to-end is better than composing the action transformer with two independently learned dynamics models.
- We introduce **Reinforced Grounded Action Transformation (RGAT)**, which grounds the simulator using an action transformer, learned end-to-end using Reinforcement Learning.
- We perform experiments in several simulated domains with mismatched dynamics, and show that agents trained with ***RGAT*** achieve relatively higher returns in the target environment compared to other baseline algorithms.



Method

- During each **grounding step**, we collect a little real world experience and use it to ground the simulator.
- This modified simulator is constructed by prepending the original simulator with a learned **action transformer** function.
- In *RGAT*, the action transformer policy is learned end-to-end using Reinforcement Learning.
- A learned forward model is used to compute the reward function for the action transformer policy.
- A policy trained on the **grounded simulator** transfers better to the real world

Is your robot simulator's *dynamics* inaccurate ? 😐

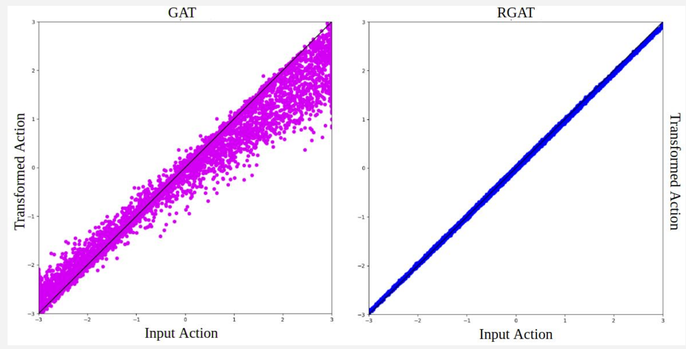
RGAT grounds simulators with data from the target domain, making it more accurate 🙂

https://arxiv.org/abs/2008.01279

Results

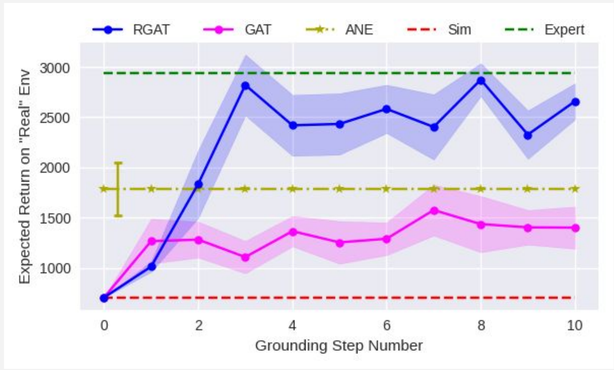
- We empirically validate our results on various Mujoco Gym tasks.
- In most domains, our method outperforms other similar algorithms and almost performs as well as training in the target environment directly.

End-to-End learning is better



In a Sim-to-Self experiment, we show that end-to-end learning in RGAT results in more precise action transformation

Experimental Results



Our method (*RGAT*) achieves higher returns than the baseline on several Mujoco Domains (full results in paper)