Stochastic Grounded Action Transformation for Robot Learning in Simulation

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

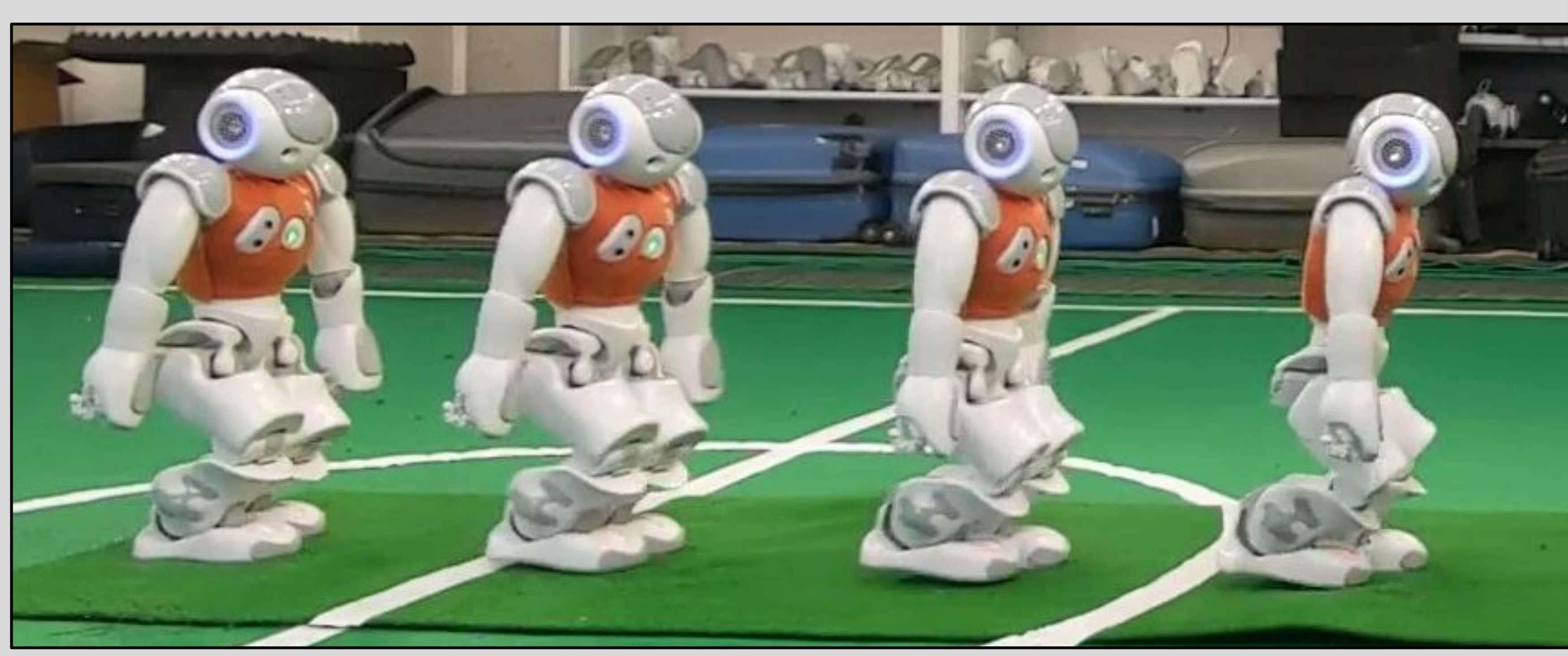
Peter Stone UT Austin, Sony AI

Abstract

- Policies trained in simulation often do not perform as well in the real world.
- Prior methods, such as Grounded Action Transformation (***GAT***), reduce the dynamics mismatch between sim and real, but do not address process noise.
- We show that methods like *GAT*, which assume a deterministic environment fail to close the reality gap in highly stochastic environments.
- We introduce **Stochastic Grounded Action Transformation (*SGAT*)**, which accounts for process noise as well as dynamics mismatch.
- We show several theoretical examples where *SGAT* succeeds, but *GAT* fails.
- We show that our method succeeds in learning a walk policy over uneven terrain on the humanoid NAO robot, even when the terrain in simulation is flat.

Method

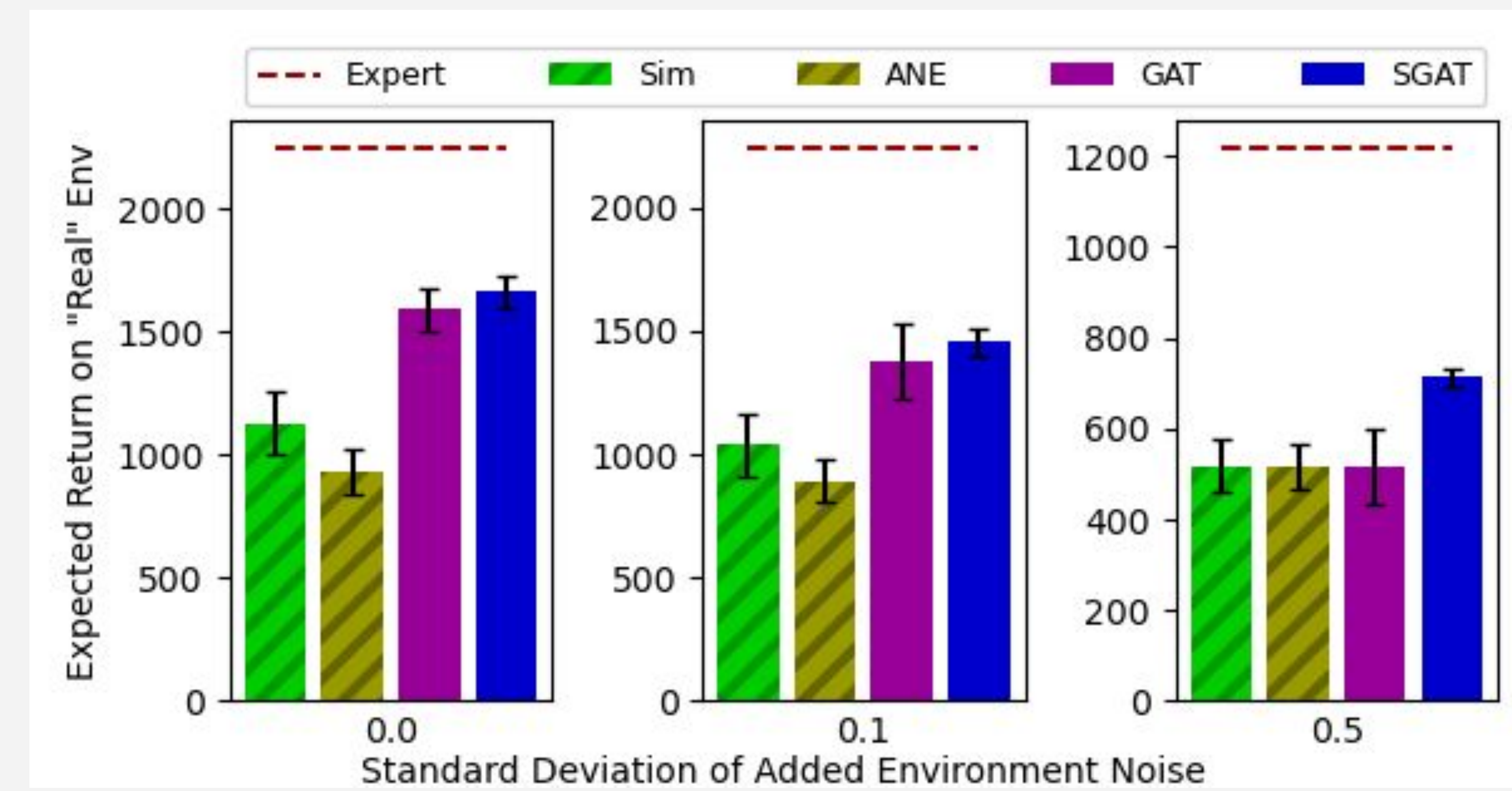
- During each **grounding step**, we collect a little real world experience and use it to ground the simulator.
- This modified simulator is constructed by prepending the original simulator with a learned **action transformer** function
- In *SGAT*, the action transformer is stochastic, capturing the variation observed in the real world data
- A policy trained on the **grounded simulator** transfers better to the real world



SGAT **makes robot simulators more realistic by correcting for dynamics mismatch and process noise.**

https://arxiv.org/abs/2008.01281

MuJoCo Results



SGAT outperforms *GAT,* a domain ranomization baseline, and a naive transfer approach (Half-Cheetah domain)

NAO Results

- We used *SGAT* to learn a walk policy in simulation for the NAO robots. The simulator has no model of the terrain, but using our method, the robot learns this variance from real-world data.

	Grounding Step 1		Grounding Step 2	
	Speed (cm/s)	Falls	Speed (cm/s)	Falls
GAT	15.7 ± 2.98	6/10	18.5 ± 3.63	10/10
SGAT	16.9 ± 0.678	**0/10**	18.0 ± 2.15	**1/10**

- Our method results in a walk that is much more stable without compromising its speed