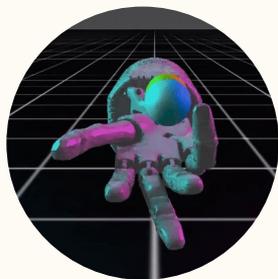


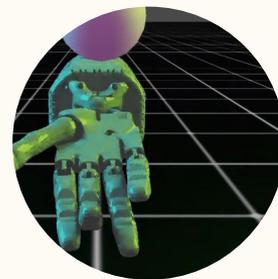
Is RL enough?

Achieving blind superhuman dexterity
with the help of World Models



BeNeRL Seminar

March 12 2026



Elle Miller



THE UNIVERSITY
of EDINBURGH

talk overview

1. introductions
2. "Is RL for robots still just a lab thing or is it real?"
3. **our work:** Enhancing Tactile-based RL for Robotic Control (*via world models*)
4. **practical RL advice**

part 1: introductions



THE UNIVERSITY OF
SYDNEY

Australia



NASA JPL, Los Angeles



DLR, Munich



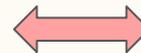
Scotland



THE UNIVERSITY
of EDINBURGH



Waseda University, Tokyo



reinforcement learning



i don't know much about it but am learning



i love it



i like it



its complicated



i dislike it

RL for robotics



This is my research!



I am interested/have used robotic envs for my research



Not interested



Too complex/difficult, imitation learning is easier



I have been hurt

RL can make a robot do ANYTHING *



Technically possible



Unsure



No

* In the realm of continuous control e.g. run, swim, fly, handstand, backflip, paint, knit, bake cookies

talk overview

1. ~~introductions~~

2. "Is RL for robots still just a lab thing or is it real?"

3. **our work:** Enhancing Tactile-based RL for Robotic Control (*via world models*)

4. **practical RL advice**

**part 2: "Is RL for robots still just a lab
thing or is it real?"**



locomotion

manipulation

“World’s first fully autonomous humanoid robot cluster Kung Fu performance” - **Unitree 2026**



Lin et al.
Sim-to-Real Reinforcement Learning for Vision-Based Dexterous Manipulation on Humanoids, **CoRL 2025**



Miki et al.
Learning to walk in confined spaces using 3D representation, **ICRA 2024**

Robot Synesthesia In-Hand Manipulation with Visuotactile Sensing



Ying Yuan*
Haichuan Che*
Yueha Chen*
Binghao Huang
Zhao-Hang Yin
Kang-Won Lee
Yi Wu
Soo-Chul Lim
Xiaolong Wang

UC San Diego
Tsinghua University
University of Mass. Lowell
Chungnam University
UC Berkeley
Dongguk University

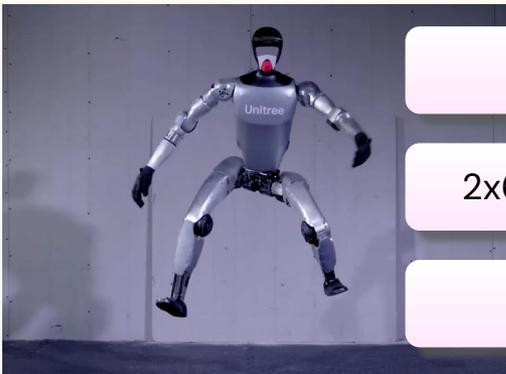
* Equal contributions

Yuan & Che & Qin et al. Robot Synesthesia: In-Hand Manipulation with Visuotactile Sensing, **ICRA 2024**

so... what's the hold up with
manipulation????

Definition 5. Manipulation refers to an agent's control of its environment through selective contact.

why RL for manipulation is hard



1 task/reward fn → **infinite** tasks/reward fns

2x6 or 4x3 ~ **12 actions** → 2x(arm + hand) ~ **50 actions**

short-horizon → **long-horizon**

"free space" sim easy → contact sim **not accurate** (yet)

doesn't need scene info → relies on **scene info**

can recover from error → minimal room for **error** 🐱





OpenAI, 2018

RL agents typically rely on
ground-truth information

(privileged/scene/state)

...but humans can do manipulation without
this information?

so why can't an RL agent?

talk overview

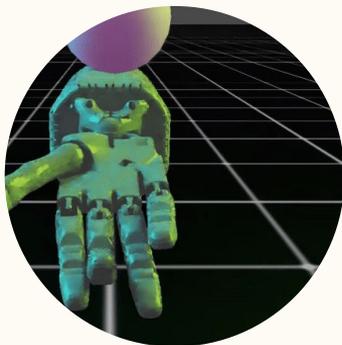
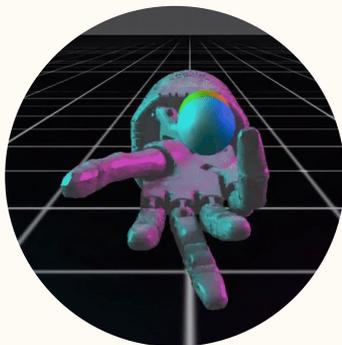
1. ~~introductions~~

2. ~~"Is RL for robots still just a lab thing or is it real?"~~

3. **our work:** Enhancing Tactile-based RL for Robotic Control (*via world models*)

4. **practical RL advice**

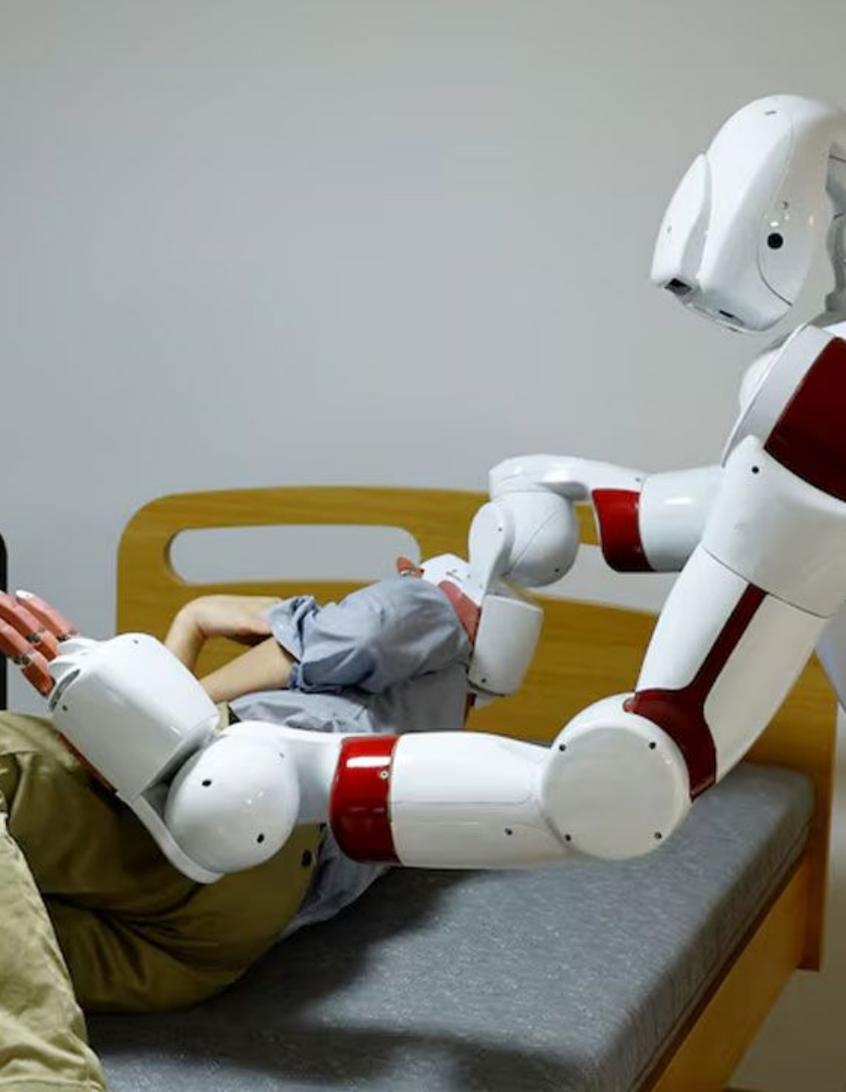
part 3: our contribution



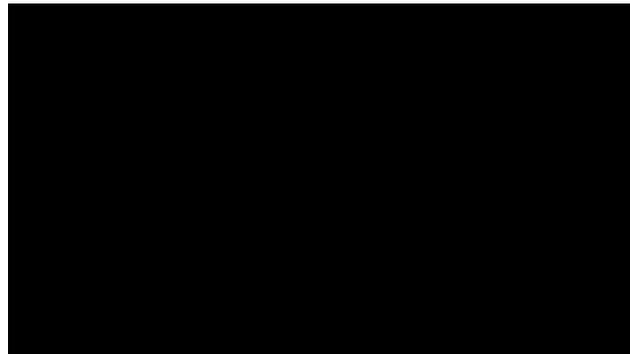
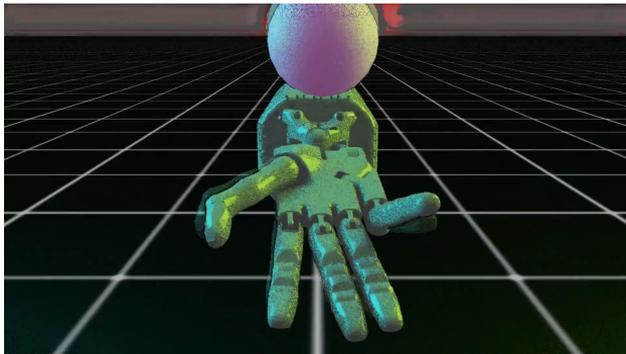
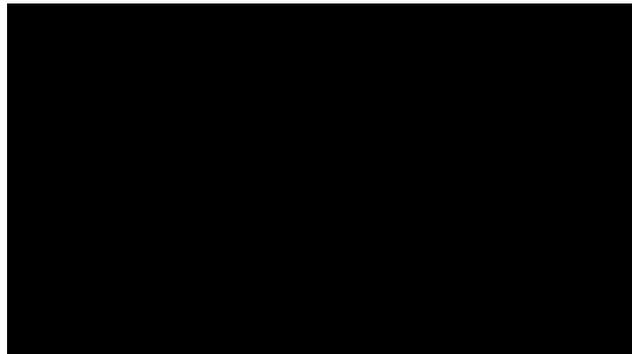
Super-human blind dexterity in simulation:

- No ground truth information
- Just proprioception + binary tactile sensing (17)

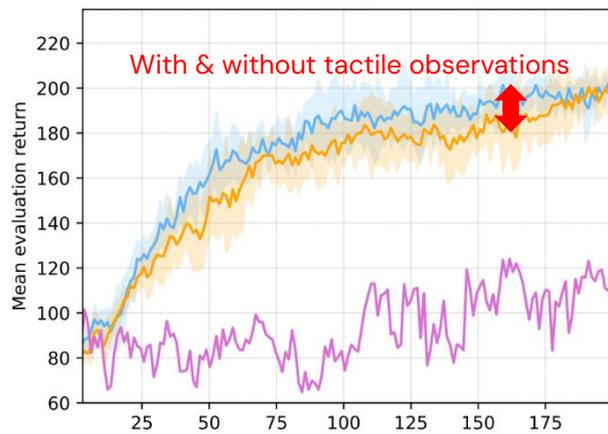
Recipe: **RL + self-supervised world model**



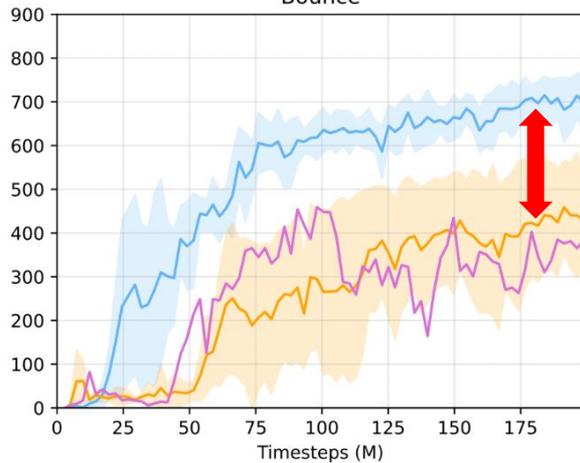
- Tactile-based RL holds significant promise for enabling real-world manipulation robots
- Progress has been hindered
 - Lack of groundbreaking demos
 - Many unanswered questions
 - Tactile data is not great for ML
 - Community hasn't converged on a sensor type



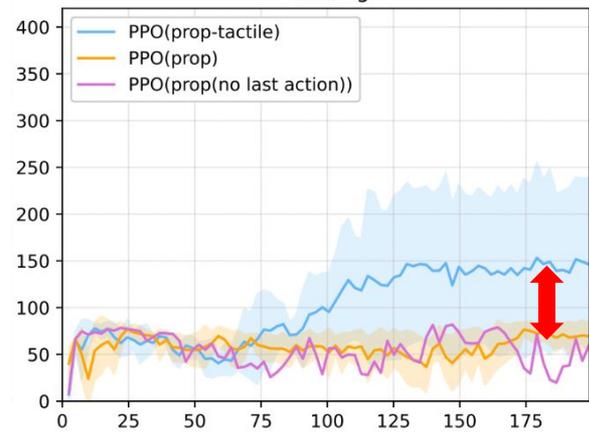
Find

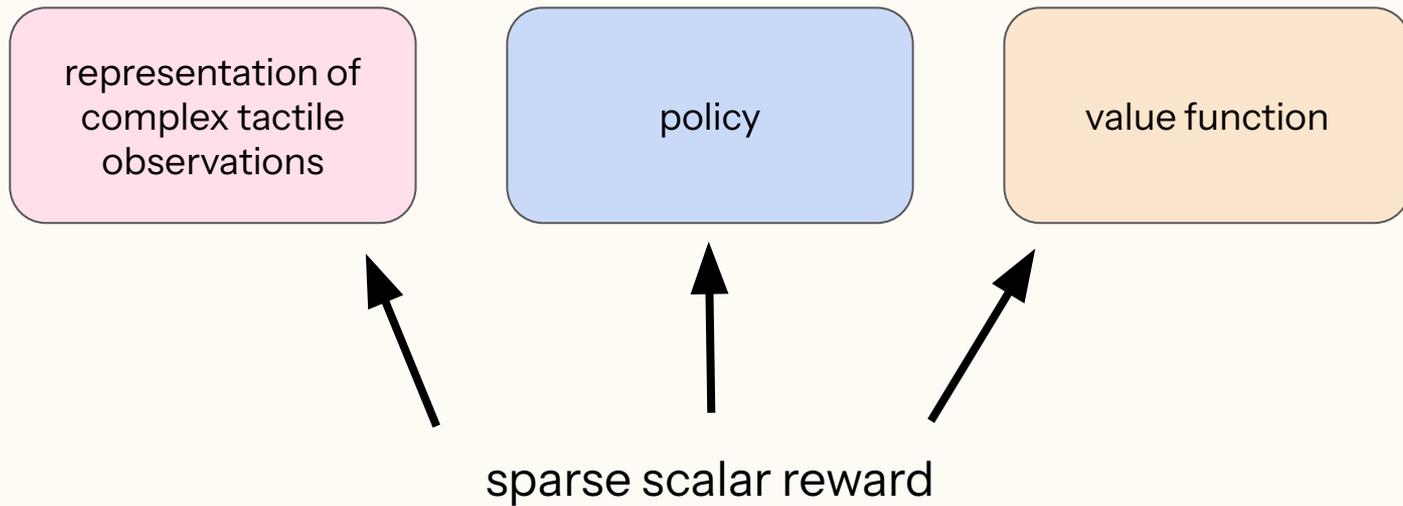


Bounce



Baoding





Is RL enough?



Good policy

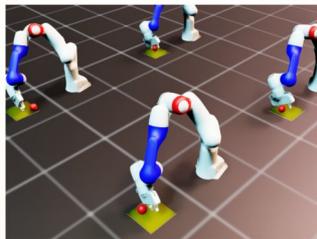
Needed to visit useful states to gather
data for representation learning



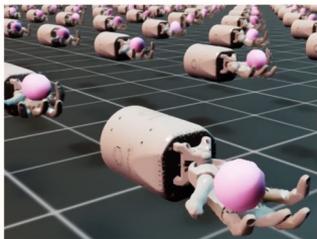
Good state representation

Needed to extract useful features for a
sophisticated policy

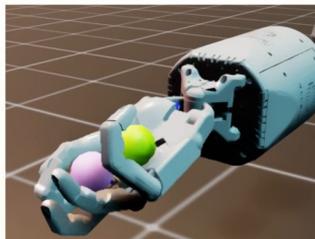
The Robot Tactile Olympiad (RoTO)



Find



Bounce



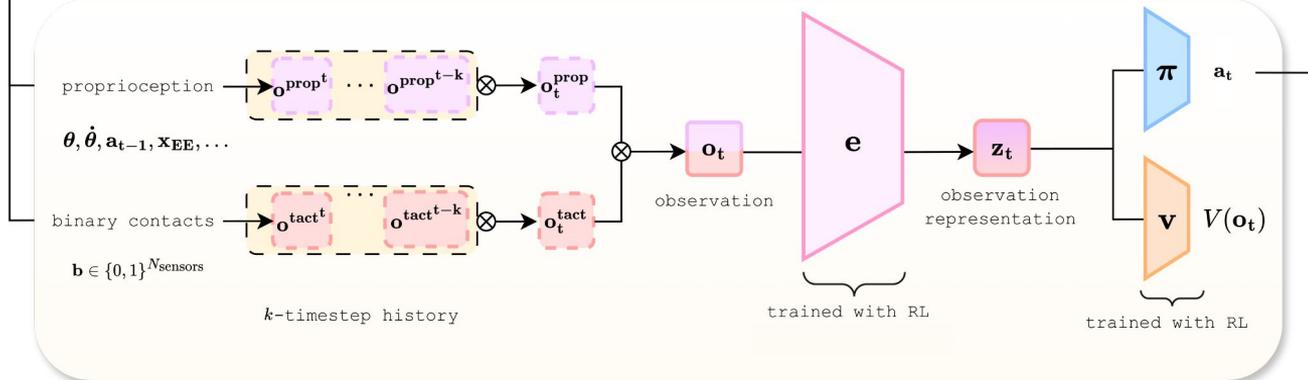
Baoding

shadow hand
tactile sensing

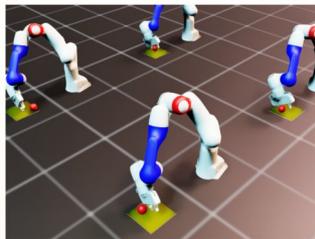


$b = 1$

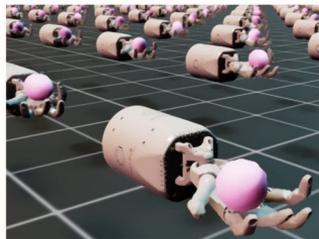
$b = 0$



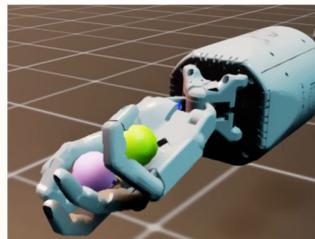
The Robot Tactile Olympiad (RoTO)



Find



Bounce

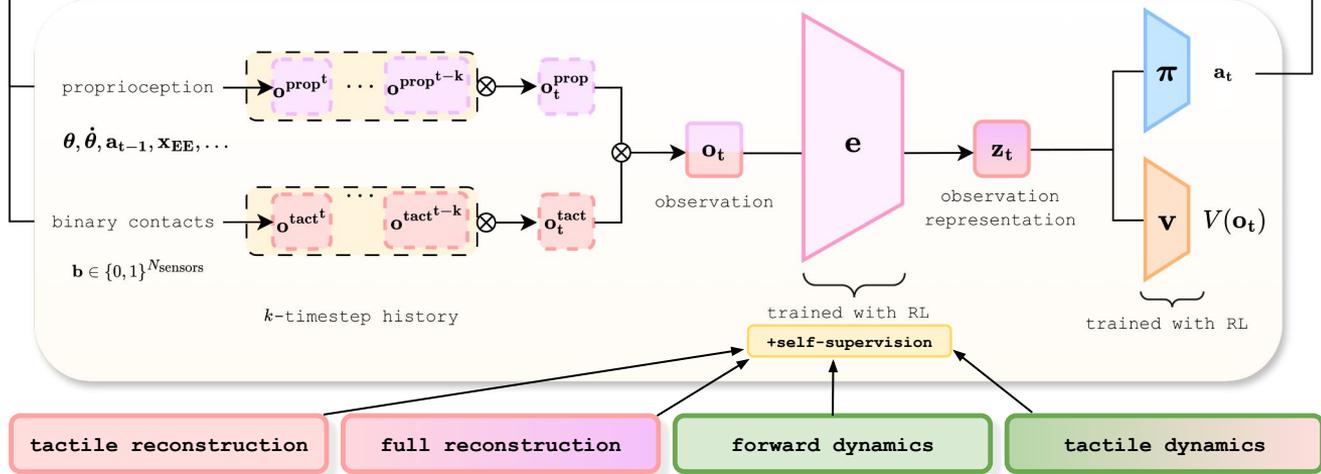


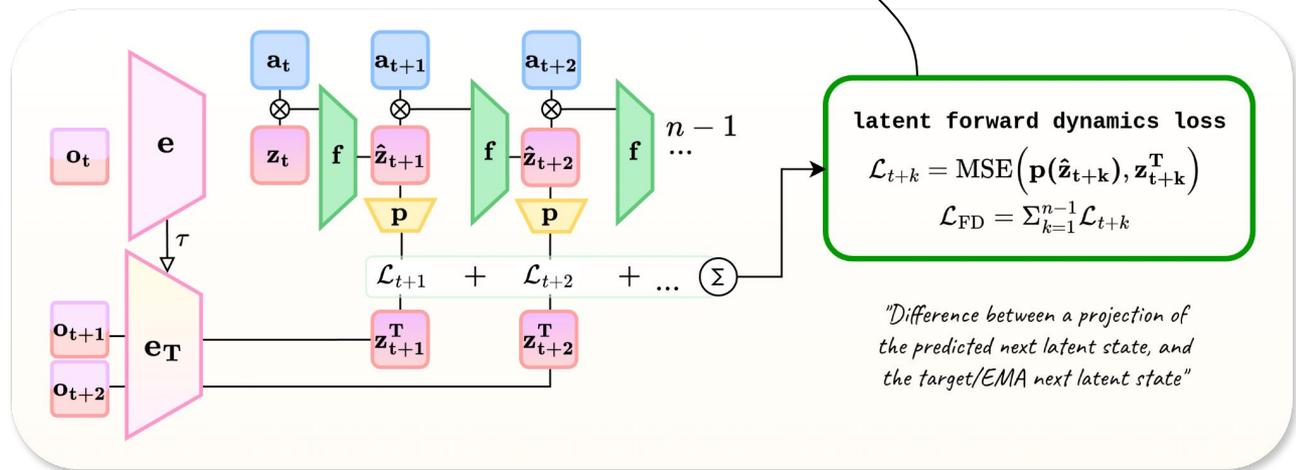
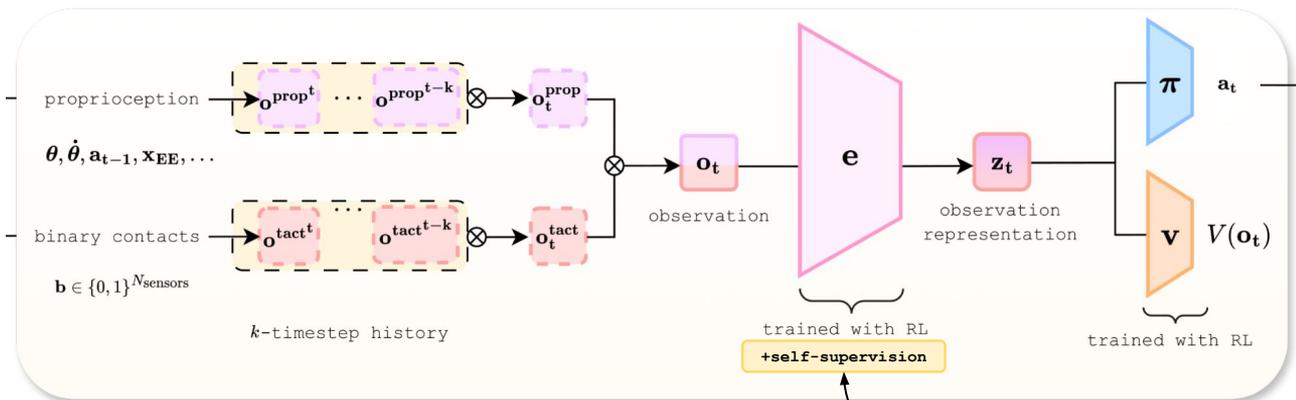
Baoding

shadow hand
tactile sensing

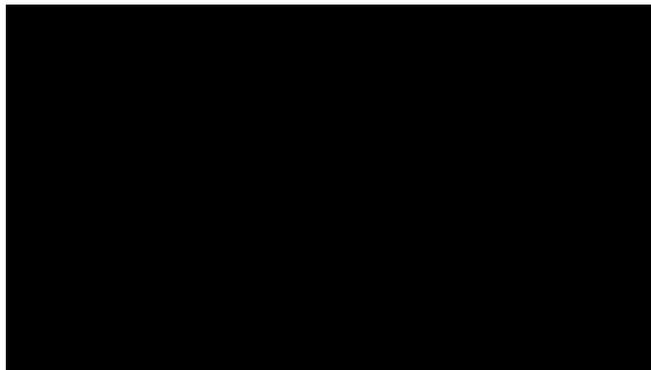


■ $b = 1$
■ $b = 0$

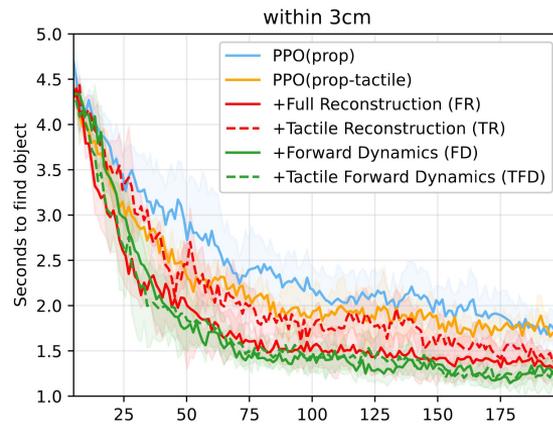
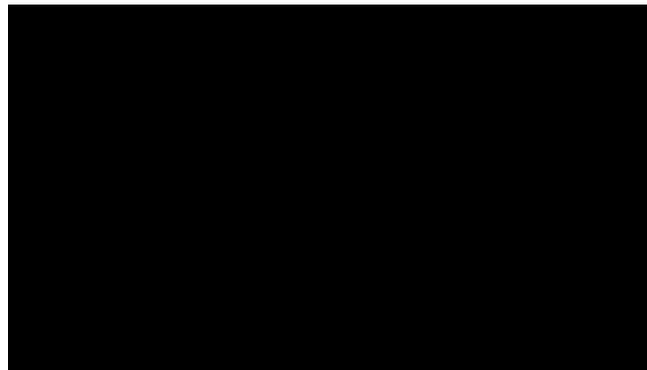




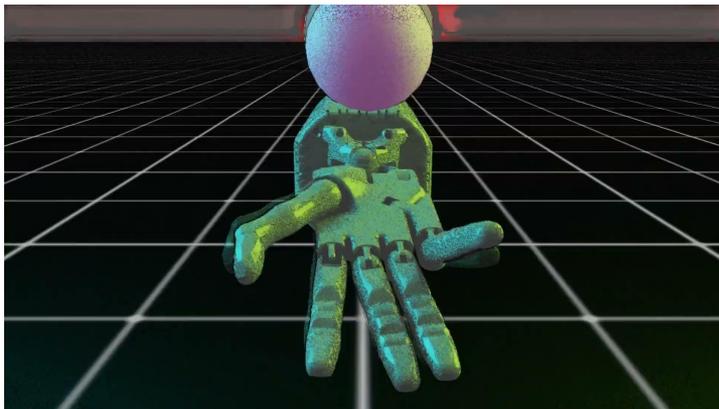
RL only



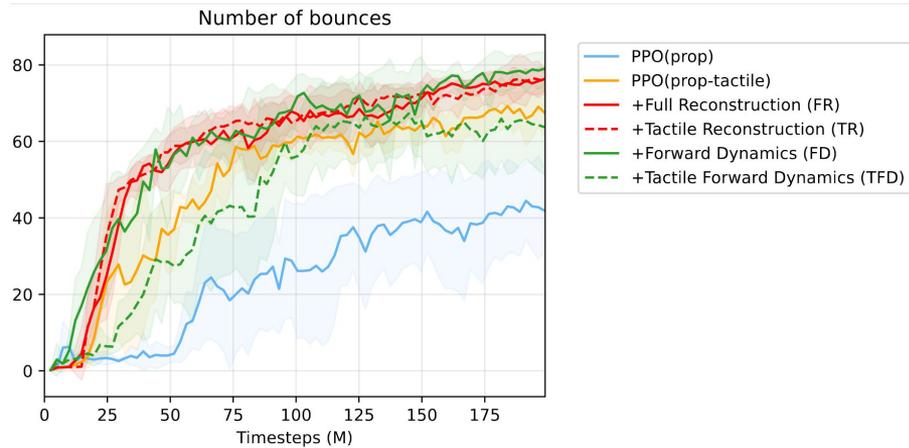
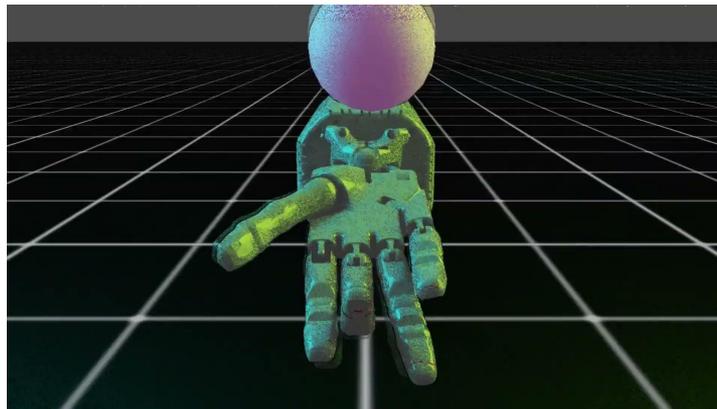
RL+Forward Dynamics



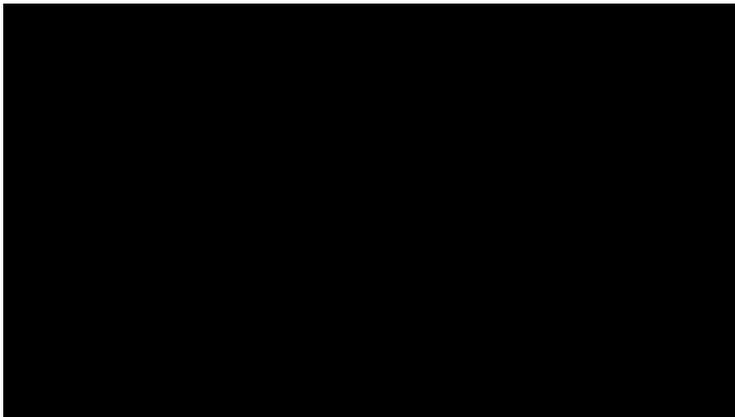
RL only



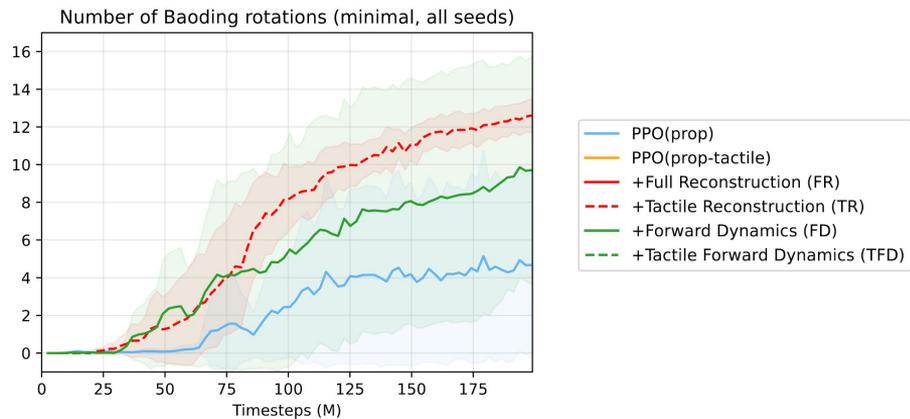
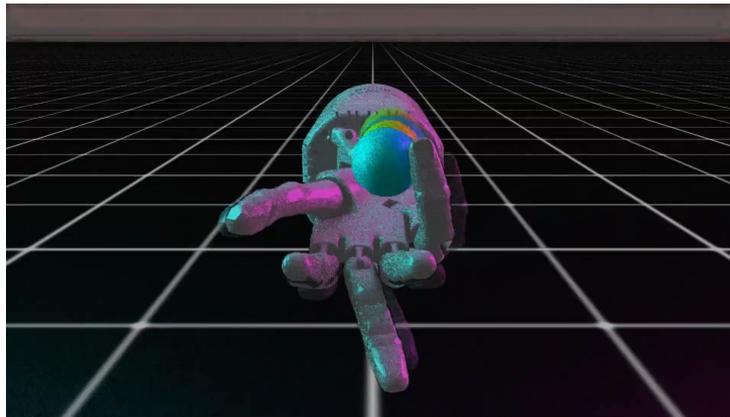
RL+Forward Dynamics

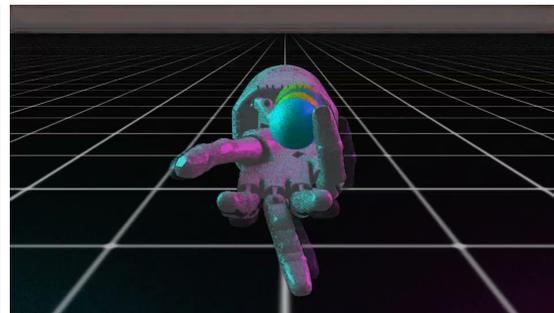
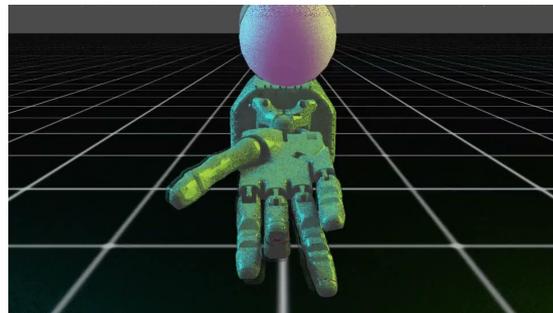
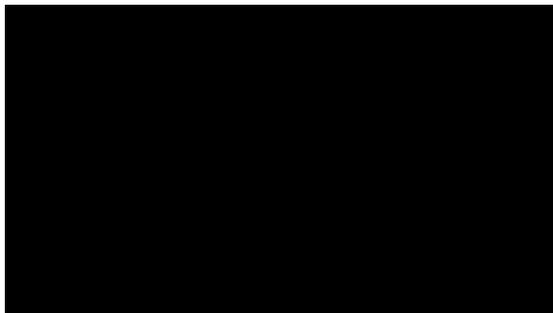
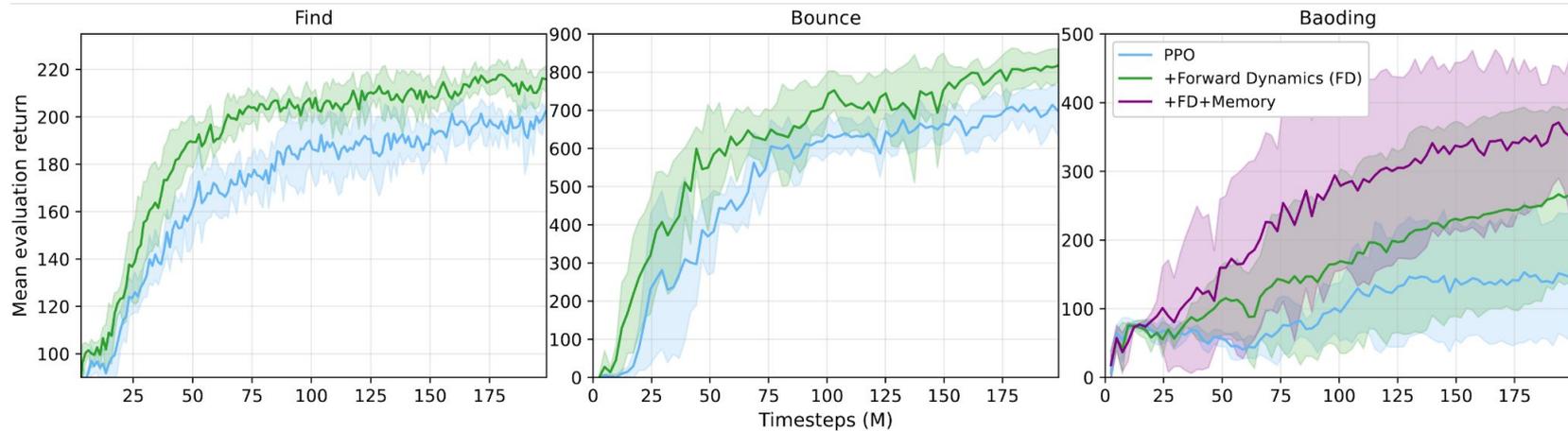


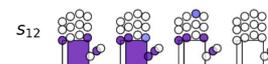
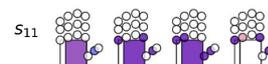
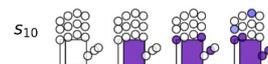
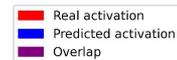
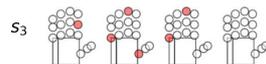
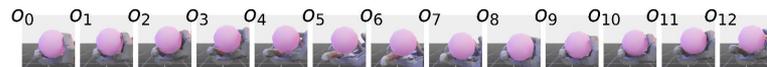
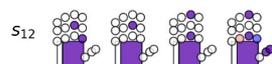
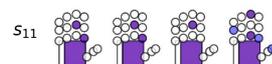
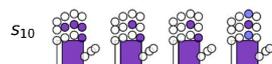
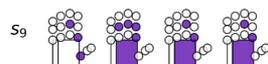
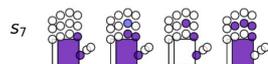
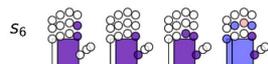
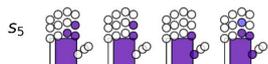
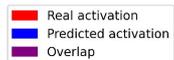
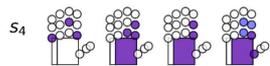
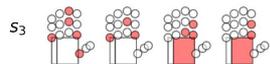
RL only



RL+Forward Dynamics+Memory







why does it work?

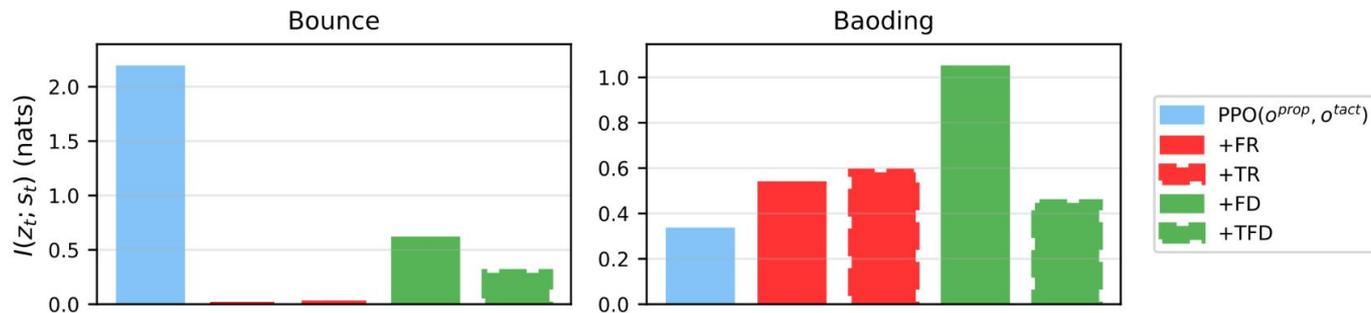
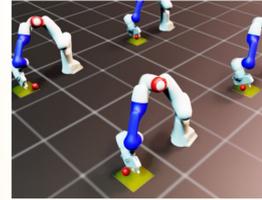
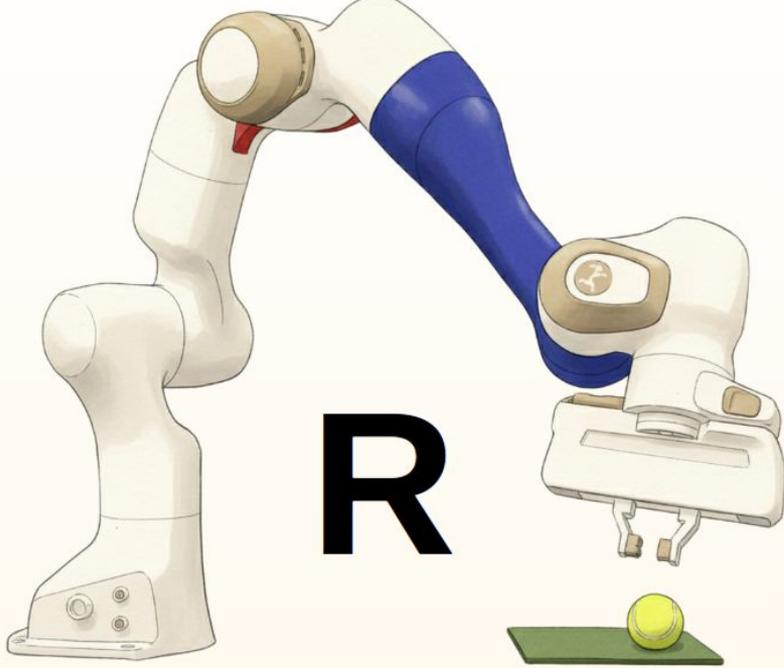
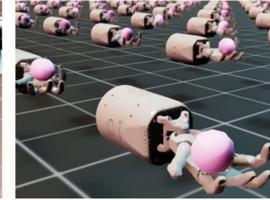


Figure 6: **Mutual information estimation.** We estimate $I(z_t; s_t)$ between the reduced observation representation z_t and the ground-truth state s_t to evaluate the representation quality.

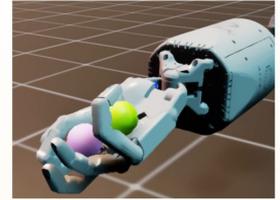
RoTO: Robot Tactile Olympiad



Find

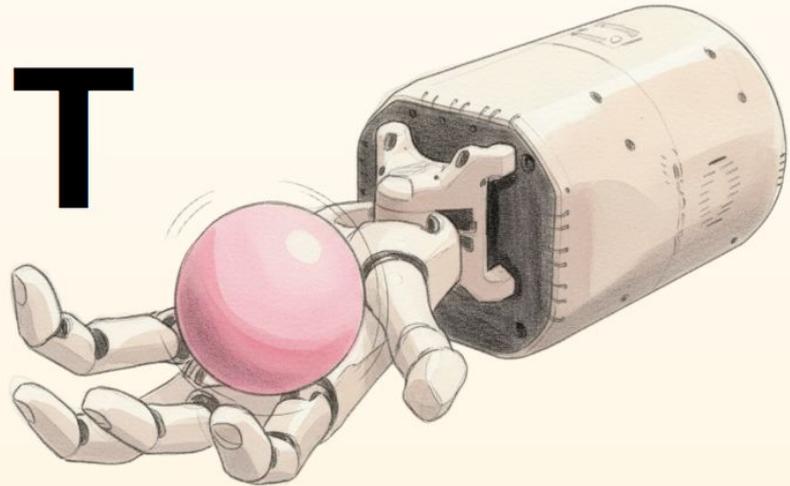


Bounce



Baoding

T



[github.com/
elle-miller/
roto](https://github.com/elle-miller/roto)

Enhancing Tactile-based Reinforcement Learning for Robot Control



Elle Miller, Trevor McInroe, David Abel, Oisin Mac Aodha, Sethu Vijayakumar



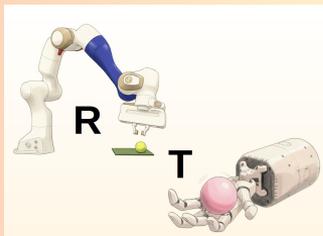
Contribution 1: Propose 4x self-supervised learning objectives for tactile RL, finding dynamics to be best

Contribution 2: Demonstrate blind superhuman dexterity with only proprioception + binary contacts

Contribution 3: New environments in RoTO for tactile RL research

RoTO: Robot
Tactile Olympiad

[github.com/
elle-miller/
roto](https://github.com/elle-miller/roto)



Stay up to date:

 [ellemiller101](https://www.linkedin.com/in/ellemiller101)



elle-miller.github.io/tactile_rl

talk overview

~~1. introductions~~

~~2. "Is RL for robots still just a lab thing or is it real?"~~

~~3. our work: Enhancing Tactile based RL for Robotic Control (*via world models*)~~

4. practical RL advice

part 4: practical RL advice

1. Your agent does not live on Earth
2. Don't fly by thermometer: Mastering the MDP
3. RL for robotics
4. Hyperparameters: Friend not foe
5. The long game



1. Your agent does not live on Earth

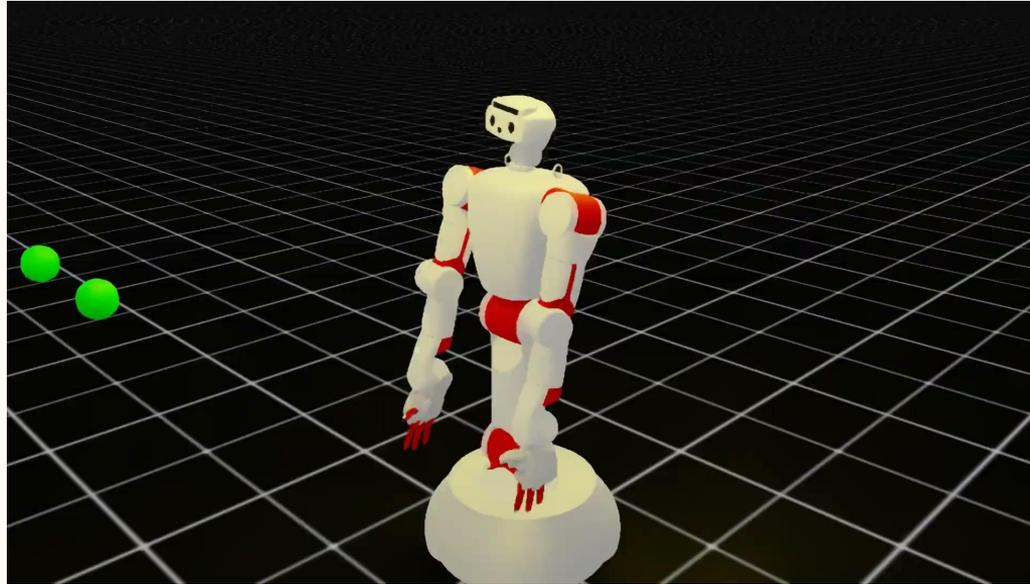
- Your simulation is a mathematical approximation of reality
- You cannot help an agent navigate a world whose fundamental laws are a mystery to you
- **Tip:** spend 1 day reading the documentation for your physics backend (like PhysX or MuJoCo) to save months debugging an agent that "should" work but doesn't



2. Don't fly by thermometer: master the MDP

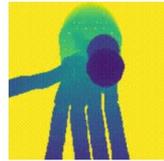
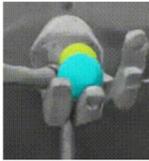
- A poorly designed MDP is like asking a pilot to fly a plane using only a thermometer and a playstation controller
- **Observations:** history, normalisation, translations/rotations relative, standard scalar
- **Actions:** only actuate what you need, check physics vs RL frequency
- **Rewards:** keep it simple as possible, build skills iteratively, avoid penalties
- **Resets:** Use lots of terminations, reset with small randomisations

Start by stretching

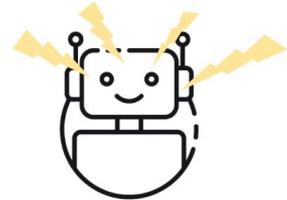
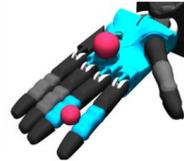


3. RL / robotics

- Don't code from scratch, but re-implement as much as possible
- Single file RL implementation
- Start with ground-truth observations, then move to “real-world” conditions
- BUT know that RL agents will struggle with multimodal observations (try self-supervision)

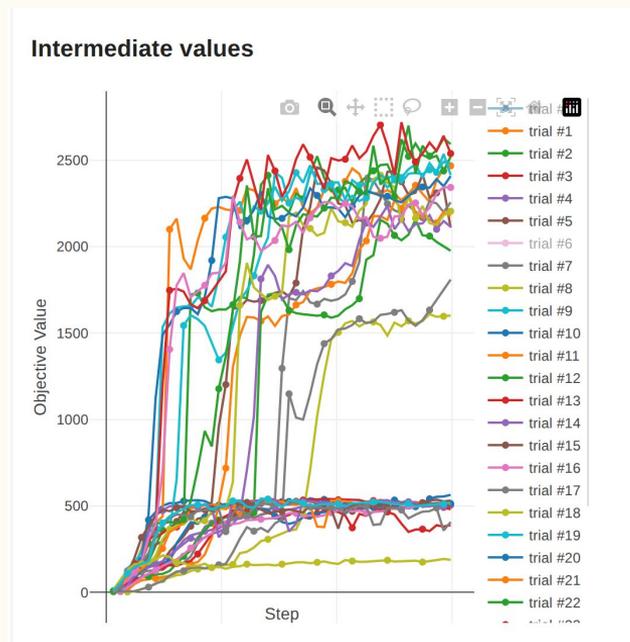


$\theta, \dot{\theta}, \mathbf{a}_{t-1},$
 \mathbf{x}_{EE}, \dots



4. Hyperparameters: Friend not Foe

- “Maybe RL doesn’t work for this task” – optimise hyperparameters to be sure
- Start early!!
- Eimer et al. “Hyperparameters in Reinforcement Learning and How To Tune Them” (ICML 2023)



5. The Long Game

isaac_lab

2025-10-08

👥 Team

6588

0

