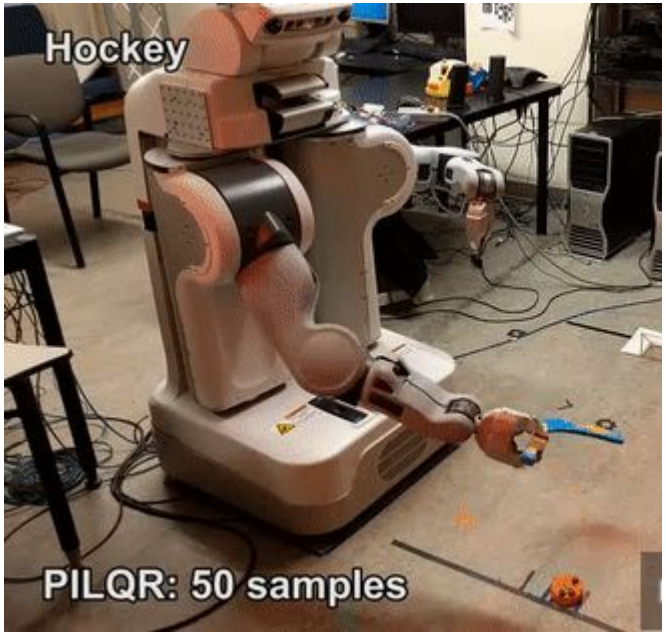
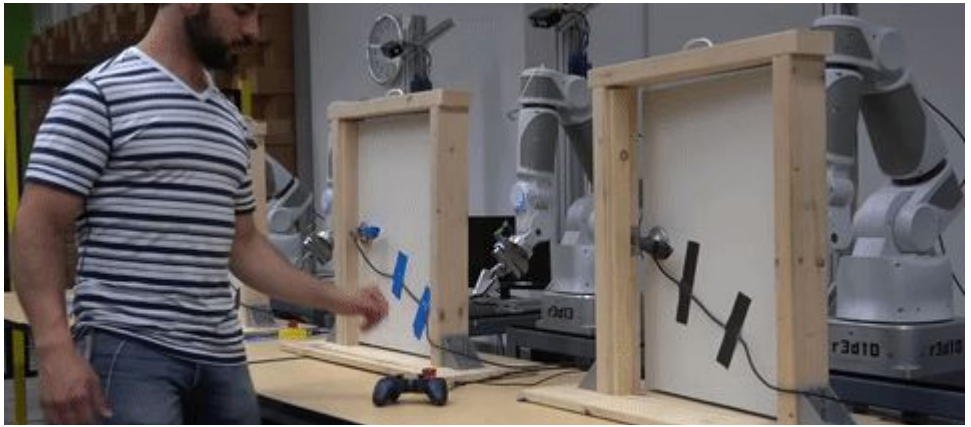
Autonomous Reinforcement Learning

Archit Sharma

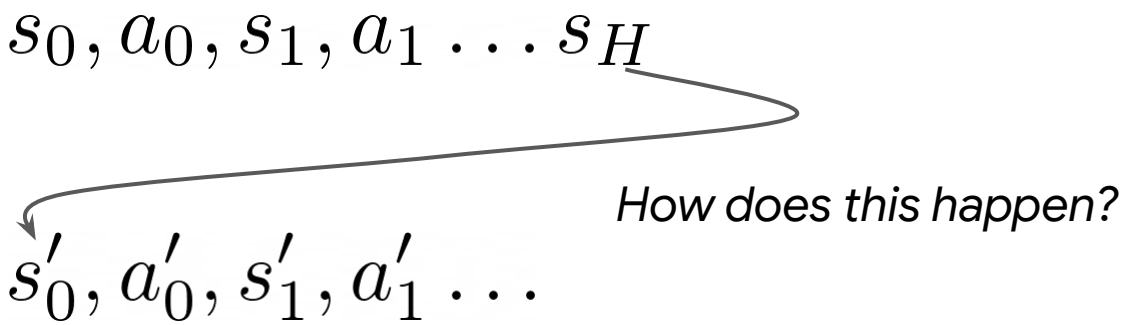
h/t: Rolf Hegrand (https://www.youtube.com/watch?v=JxH9ogP5STQ&t=185s&ab_channel=RolfHegrand)



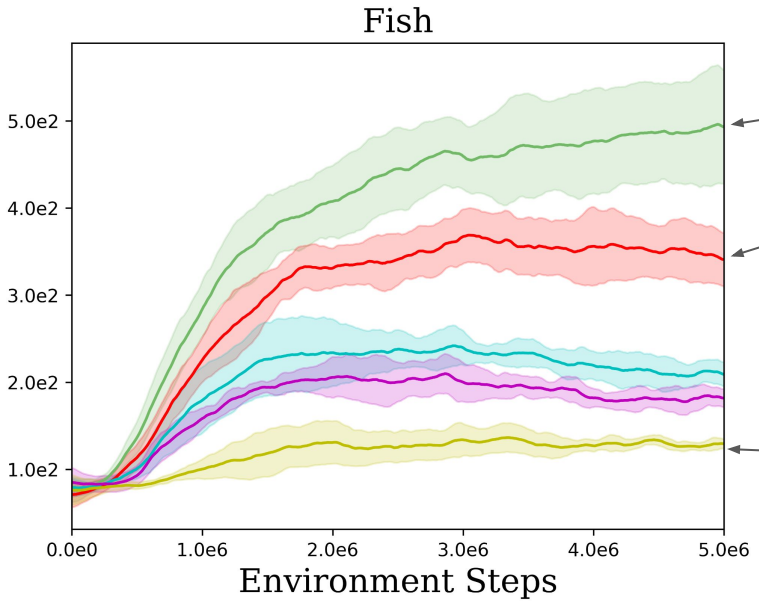
[Combining model-based and model-free updates for trajectory-centric reinforcement learning, Chebotar et al. 2017]



[Collective Robot Reinforcement Learning with Distributed Asynchronous Guided Policy Search, Yahya et al. 2016]


$s_0, a_0, s_1, a_1 \ldots s_H$
How does this happen?
$s'_0, a'_0, s'_1, a'_1 \ldots$

From Standard RL to Autonomous RL [1]



resets given every 1000 steps

resets given every 2000 steps

.
.
.

resets given every 200,000 steps

[1] Sharma*, Xu*, Sardana, Gupta, Hausman, Levine, Finn. *Autonomous Reinforcement Learning: Formalism and Benchmarking*.

Autonomous Reinforcement Learning [1]

$$\mathbb{A} : \{s_i, a_i, s_{i+1}\}_{i=0}^{t-1} \mapsto (a_t, \pi_t)$$

$$s_0, a_0, s_1, a_1 \ldots$$

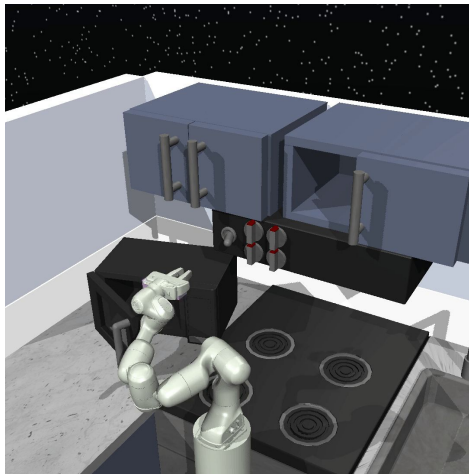
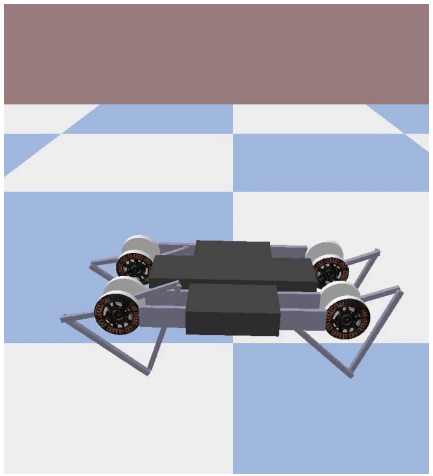
Initialize ***once*** at the beginning

Deployed Policy Evaluation

$$J(\pi) = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \right]$$

[1] Sharma*, Xu*, Sardana, Gupta, Hausman, Levine, Finn. *Autonomous Reinforcement Learning: Formalism and Benchmarking.*

EARL Benchmark


Reach out to me!

Some references:
Autonomous Reinforcement Learning via Subgoal Curricula, NeurIPS 2021
Autonomous Reinforcement Learning: Formalism and Benchmarking (releasing soon)

Autonomous Learning => Robust Learning?