

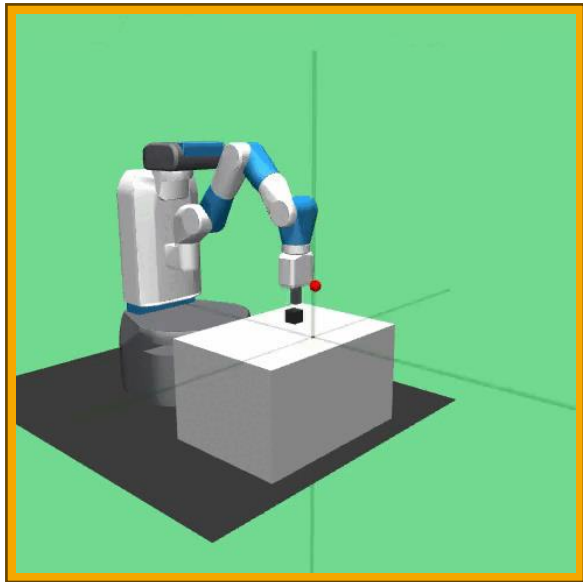
Reinforcement Learning For Grasping Top Open Liquid Containers



Adam McCutcheon

Peter Driscoll

Coming up!



Robotics.farama.org

RL (Reinforcement Learning)

DDPG (Deep Deterministic Policy Gradient)

HER (Hindsight Experience Replay)

Reinforcement Learning

General Idea

- To learn a policy that can complete the task (Picking up an object and moving it to a goal without spilling its contents).
- Policies are usually a neural network.
- Rewards can be sparse (goal state or not) or dense (every time step, complicated)

Problem!

Can you get to the goal state efficiently focusing on just a short sighted next best step?

Probably not!

Helpful Terms

- State (s) - { Joint positions, Goal position, object positions }
- Action (a) - Action to take change joint positions/velocities.
- Policy (π) - Given a state output an action $\pi(s) \rightarrow a$
- Reward (r) - Given a state and action outputs a reward $r(s,a) \rightarrow val$

Reward Solution (Bellman's)

The expected reward of a state \mathbf{s} and action \mathbf{a} is the current reward plus the expected future reward

$$Q(s, a) = \mathbb{E} \left[r + \gamma \max_{a'} Q(s', a') \mid s, a \right]$$

Current Action
Reward

Discount Factor

Expected Future
Reward

DDPG (Deep Deterministic Policy Gradient) 2 Neural Networks

Actor

- A deterministic policy that outputs an action given a state.
- Maximized Reward aka Minimize Critics loss function

Critic

- Trained to approximate Bellman's **Q(s,a)**. Minimize error to Bellman's

$$L(\theta^Q) = \frac{1}{N} \sum_{i=1}^N (Q(s_i, a_i | \theta^Q) - y_i)^2$$

Only N
Timesteps
Forward

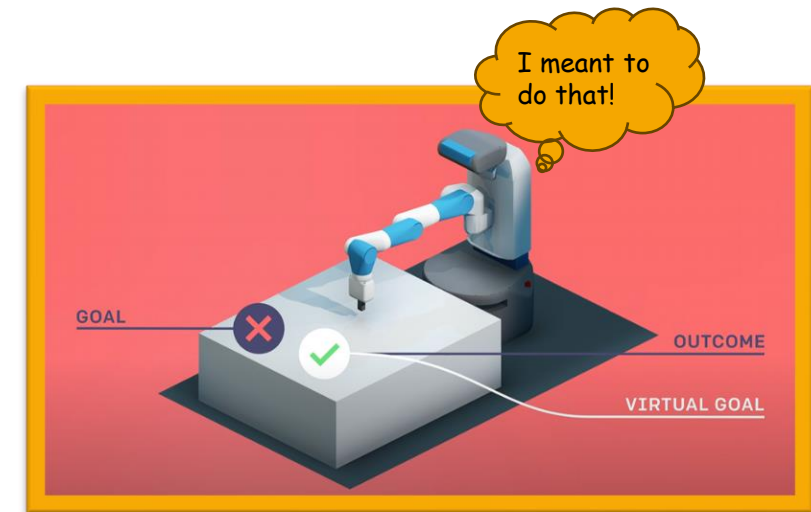
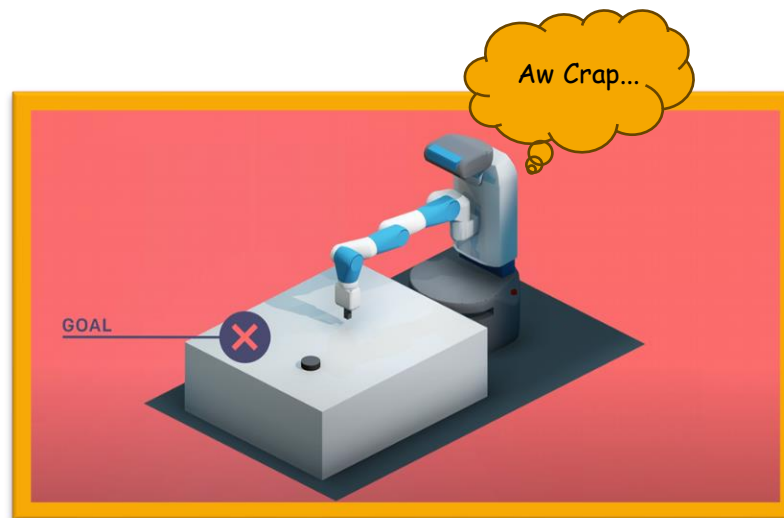
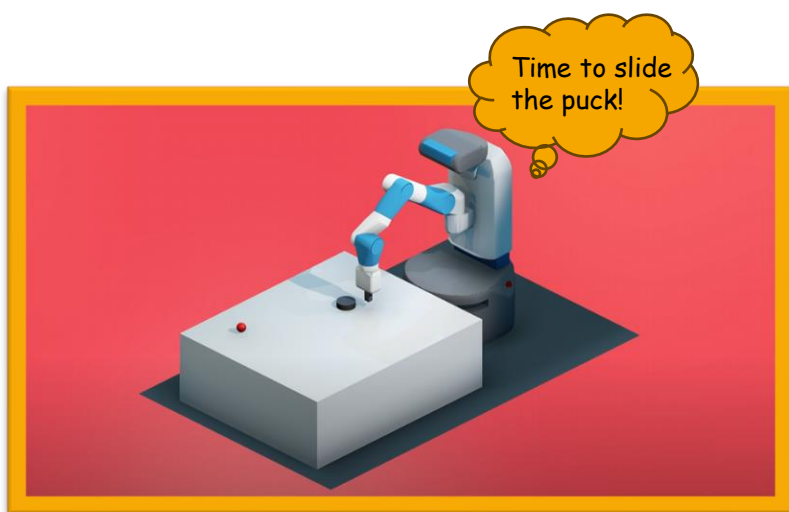
Predicted
Reward

Bellman Actual
Reward

HER (Hindsight Experience Replay)

Switches failed attempts into successes then learns from them.

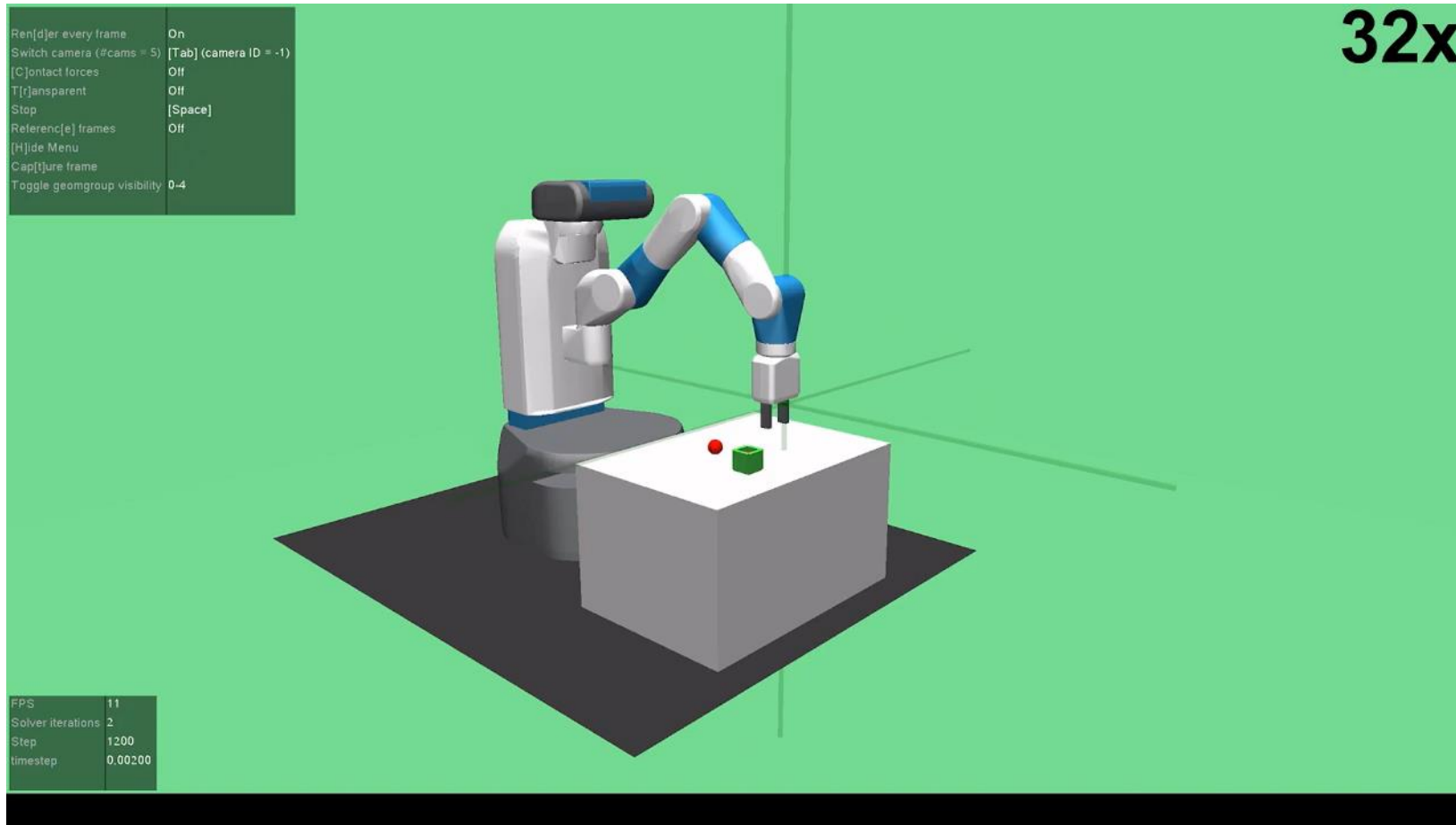
Helps the model quickly learn general ideas of how to reach different goal states by not throwing away "failed" past attempts.



Short Demo

MuJoCo
Physics
Simulation

Gymnasium
Environment



Stable
Baselines 3
RL Algorithms

Fetch Mobile
Manipulator



Questions?