SEED RL: Scalable and Efficient Deep-RL with Accelerated Central Inference

Ithan Velarde, Devarsi Rawal, Luis Batista,
Cedric Pradalier

Outline

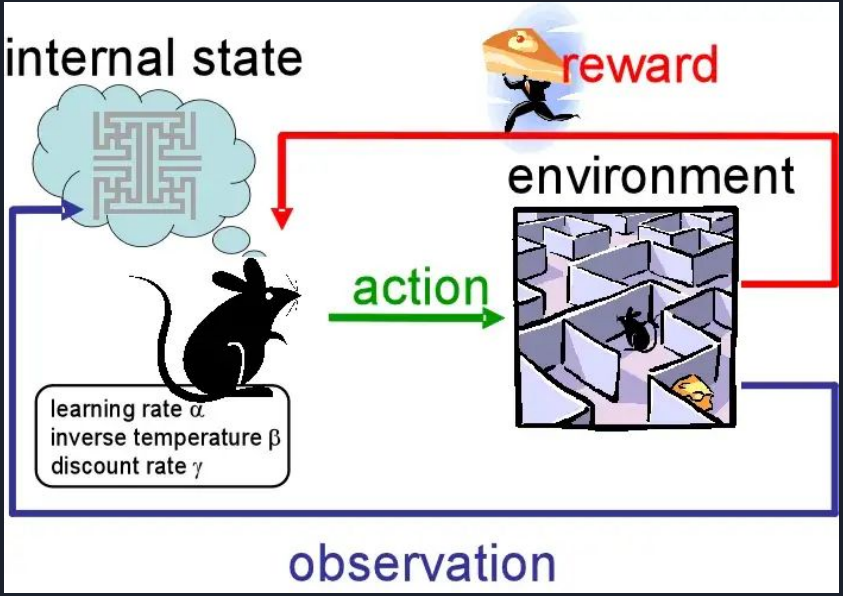
- Purpose
- What is Reinforcement Learning ?
- What is Deep Reinforcement Learning?
- What is this special project’s goal ?
- What is SEED-RL ?
- Experiments

Purpose

- Altiscan Crawler
  - Used for inspection and maintenance of marine vessels
  - Attached by an umbilical that bundles data, power, and water supply cables
- Irregular behavior as the crawler traverses larger structures
  - Currently controlled by inverse velocity kinematics
- Velocity control is necessary for localization, defect detection, and coordination

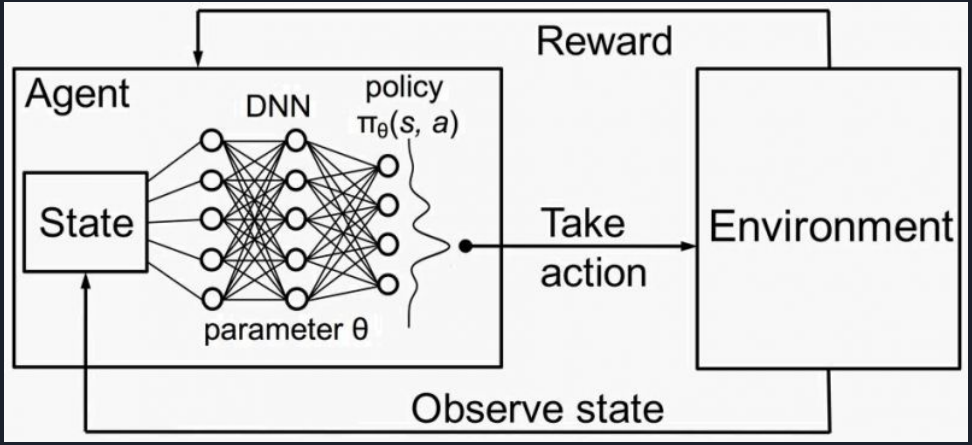
Source: https://www.directindustry.fr/prod/roboplanet/product-64387-1763534.html

What is Reinforcement Learning ?



What is Deep Reinforcement Learning (DRL)?

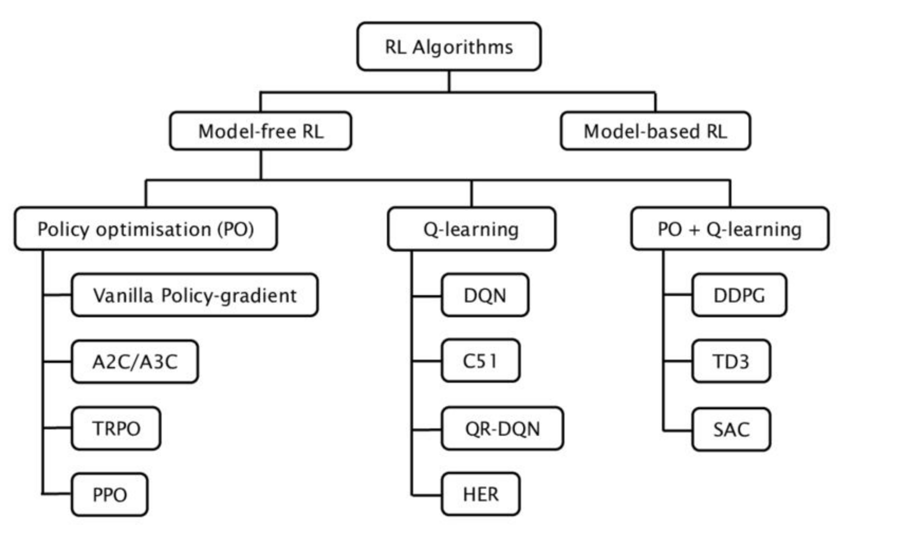
- Extends RL using deep neural networks (DNNs) to tasks with complex environments
  - Continuous state and action spaces



Source: https://medium.com/@vishnuvijayanpv/deep-reinforcement-learning-value-functions-dqn-actor-critic-method-backpropagation-through-83a277d8c38d

What is DRL?

- Policy Gradient
    - How do we update the DNN params to choose better actions?
- Actor-Critic
    - How do we improve learning stability?
- Trust-Region Policy Optimization (TRPO)
    - How do we account for large gradient steps?
- Proximal Policy Optimization (PPO)
    - How do we use less calculus?



Source: https://www.researchgate.net/figure/Class-of-model-free-RL-algorithms-adapted-from-21_fig2_351105046

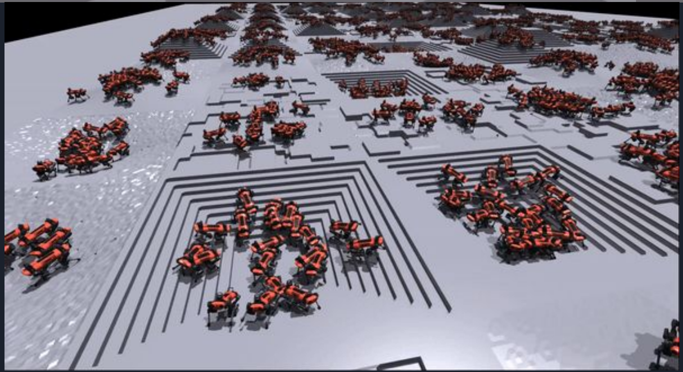
Learning Robotic Control with DRL

- DRL needs to take observations from the environment and produce actions to control a robot

- Training agents on physical robots can be costly and dangerous

- How do we efficiently train an agent for robotic control?
  - Use simulation!

Isaac Gym

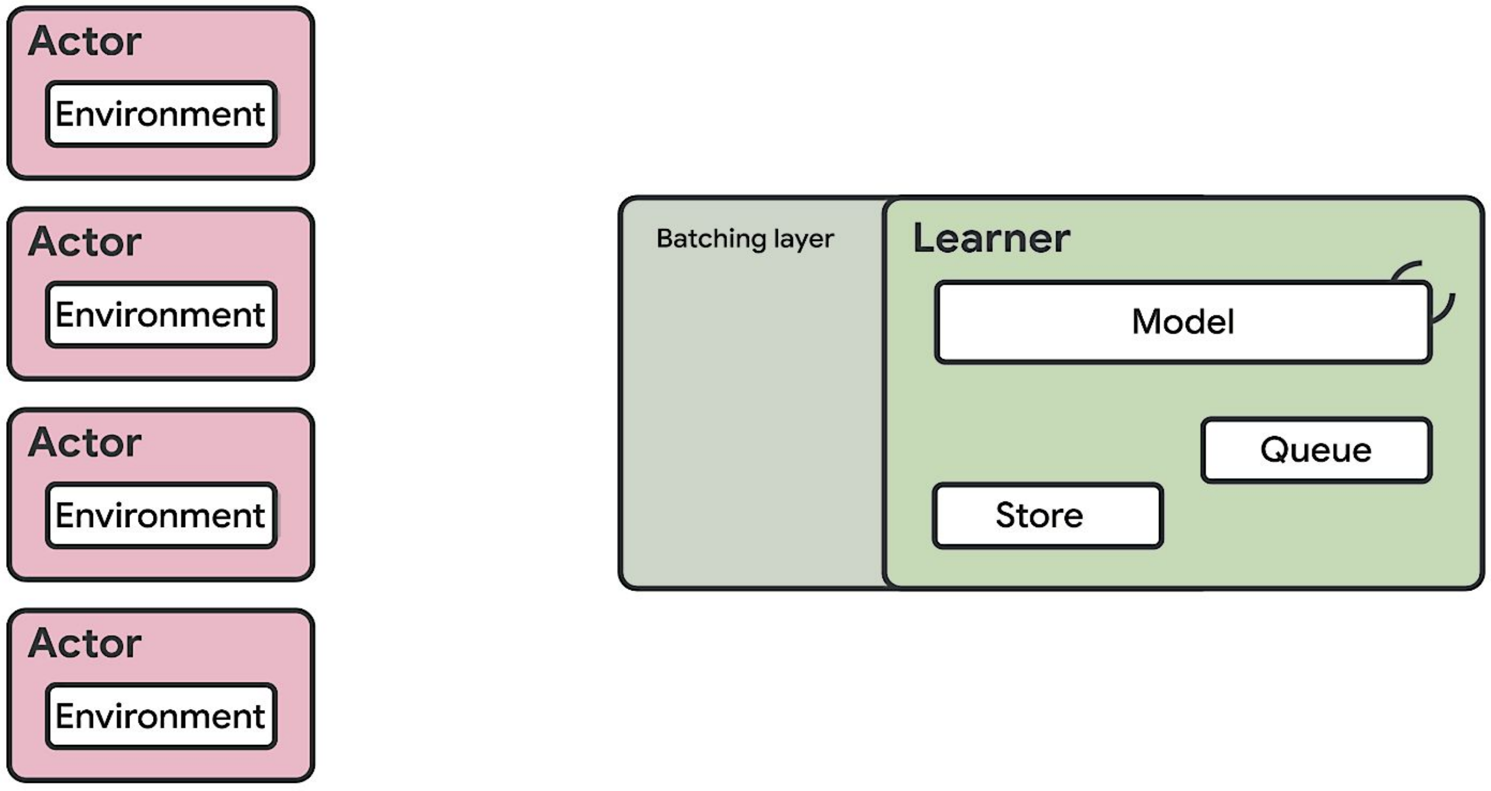


Source: https://developer.nvidia.com/isaac-gym

What is this special project's goal ?

- Reuse existing infrastructure
- Speed-up CPU-based environments
- Create a benchmark for GPU-based environments
- Make a reusable framework for future GTL projects

What is SEED-RL ?



SEED-RL vs Isaac Gym

- Isaac Gym
  - 1024 actors
  - cost < 0.296 $
  - 21 917 steps per second
  - 204 hours of simulation
- Small SEED-RL
  - 3.5 actors
  - cost = 0.807 $
  - 260 steps per second
  - 42 hours of simulation
- Big SEED-RL
  - Started with 26 actors, but we lost many
  - cost = 0.517 $
  - 2 988 steps per second
  - 44 hours of simulation

Each actor uses ~70% CPU

Resource	Cost per hour
CPU core	$0.0475
Nvidia Tesla P100	$1.46
TPU v3 core	$1.00

Table 3: Cost of cloud resources as of Sep. 2019.

Table 6.2: Runtime performance.

Simulation Software	DRL Library	Training Time
Isaac Gym	rl-games	12m10s
PyBullet	SEED RL	17h3m16s
PyBullet	**SEED-RLv2**	**1h39m12s**

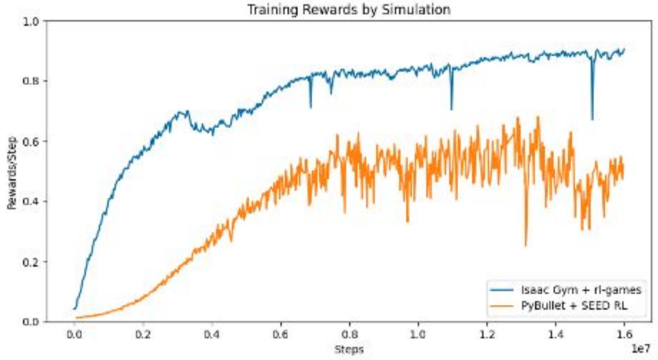


Figure 6.2: Training Reward for Isaac Gym and PyBullet

References

- SEED RL: Scalable and Efficient Deep-RL with Accelerated Central Inference. Lasse Espeholt and Raphael Marinier and Piotr Stanczyk and Ke Wang and Marcin Michalski. 2019
- The very basics of Reinforcement Learning, Aneek Das. https://becominghuman.ai/the-very-basics-of-reinforcement-learning-154f28a79071
- Deep Reinforcement learning for the velocity control of a magnetic, tethered differential-drive robot. Devarsi Rawal. 2022
- N. Rudin, D. Hoeller, P. Reist, and M. Hutter, *Learning to walk in minutes using massively parallel deep reinforcement learning*, 2021.
- V. Makoviychuk et al., *Isaac gym: High performance gpu-based physics simulation for robot learning*, 2021.
- J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, Proximal policy optimization algorithms, 2017.