Reinforcement Learning with OpenAI Gym

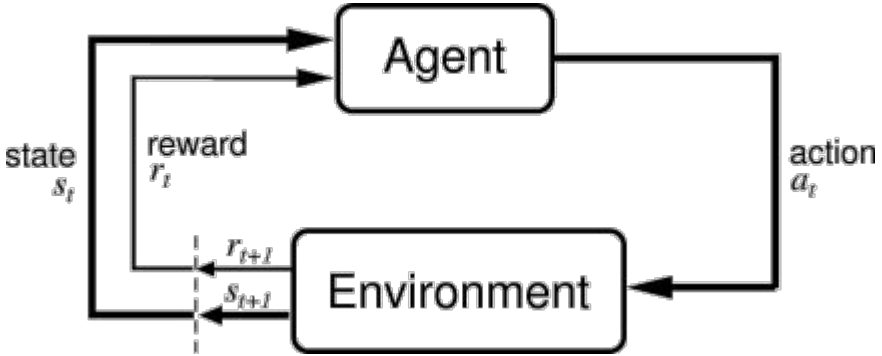
Dennis Tran
9/21/21

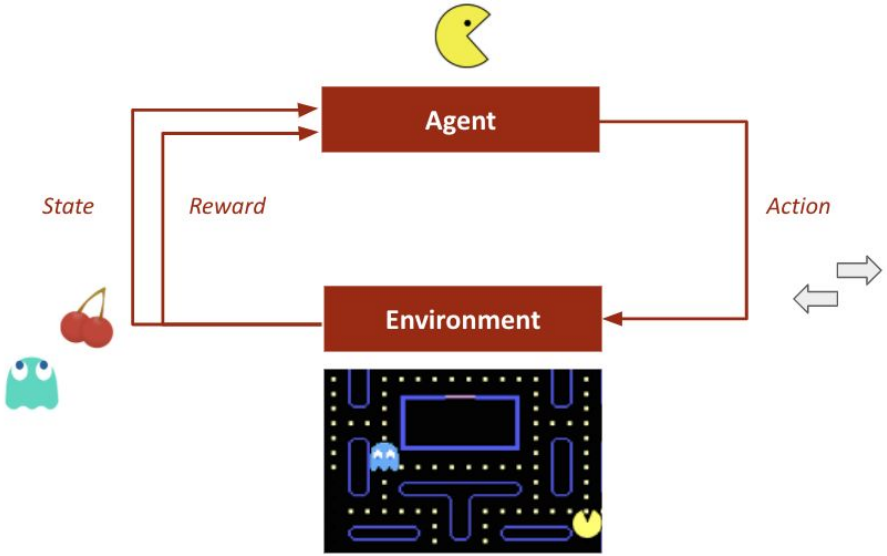
Problem Statement

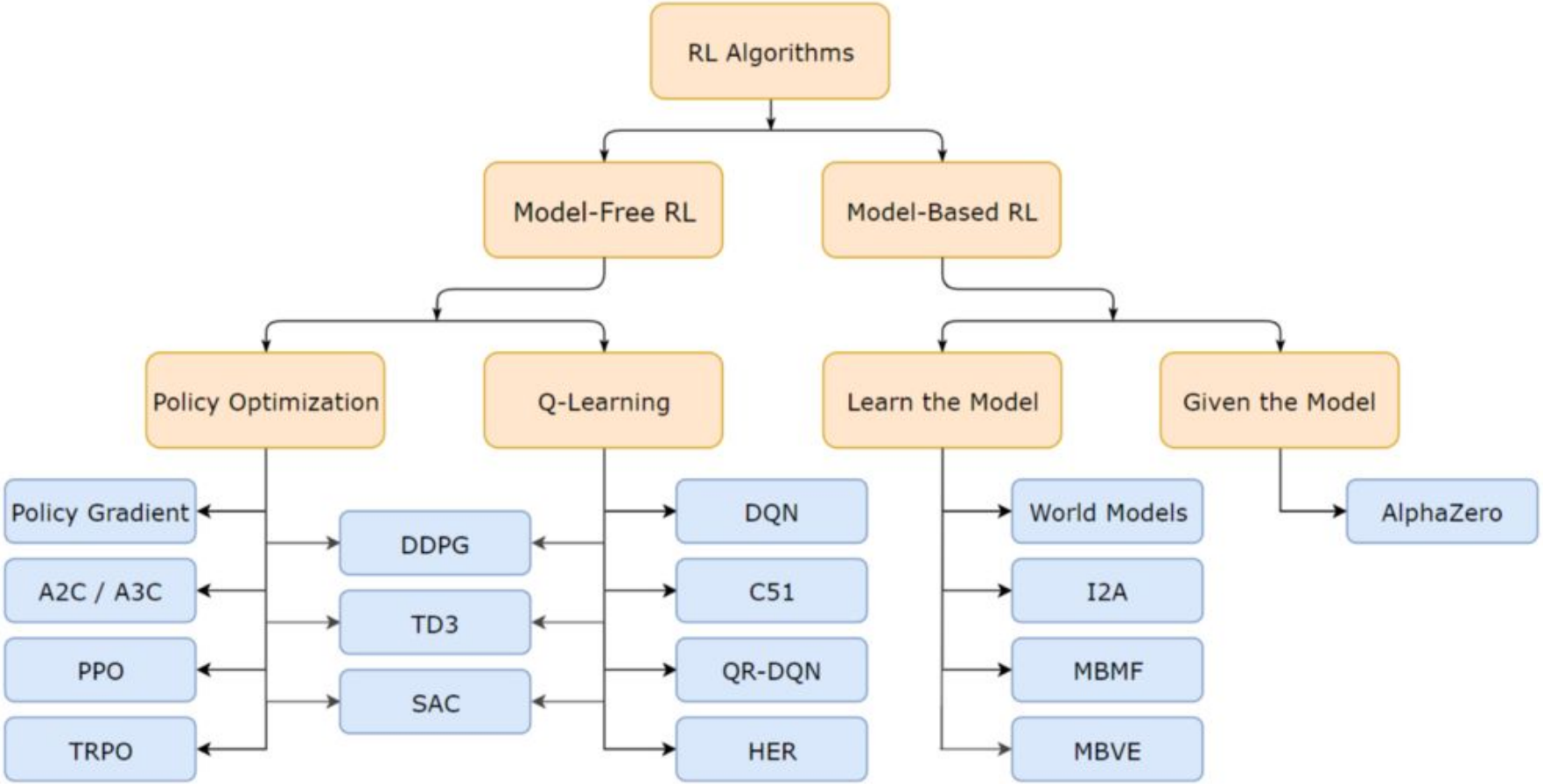
A Robotics company is looking for ways to reduce costs in training its robots. It is wondering whether or not Reinforcement Learning in simulated environments is a viable alternative for training.

What is Reinforcement Learning?

Reinforcement learning (**RL**) is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward.






RL Algorithms
Model-Free RL
Model-Based RL
Policy Optimization
Q-Learning
Learn the Model
Given the Model
Policy Gradient
A2C / A3C
PPO
TRPO
DDPG
TD3
SAC
DQN
C51
QR-DQN
HER
World Models
I2A
MBMF
MBVE
AlphaZero

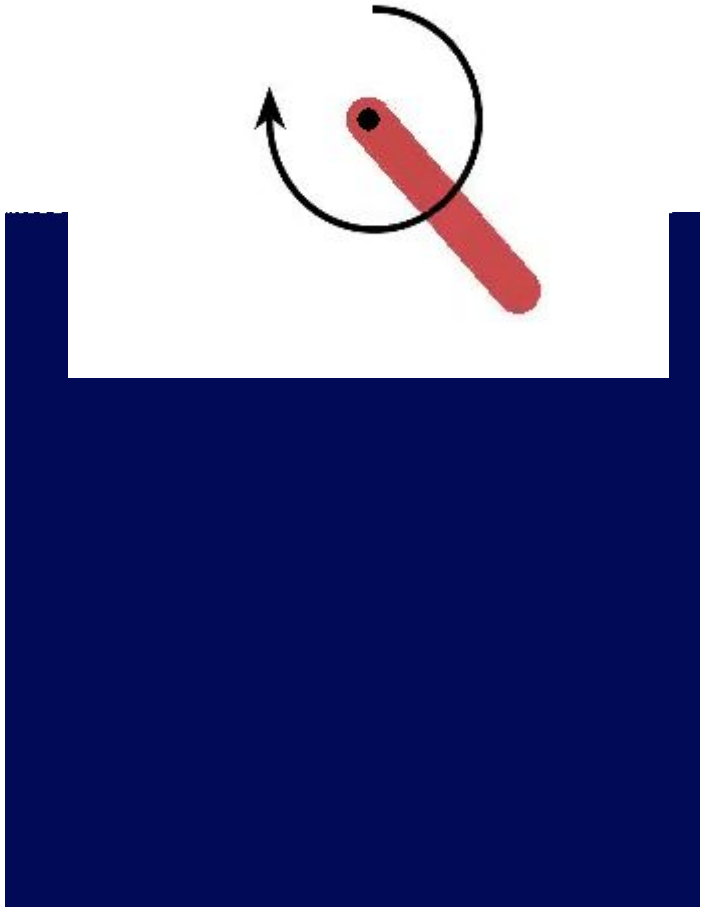
Classic Control

Cart Pole: balance pole

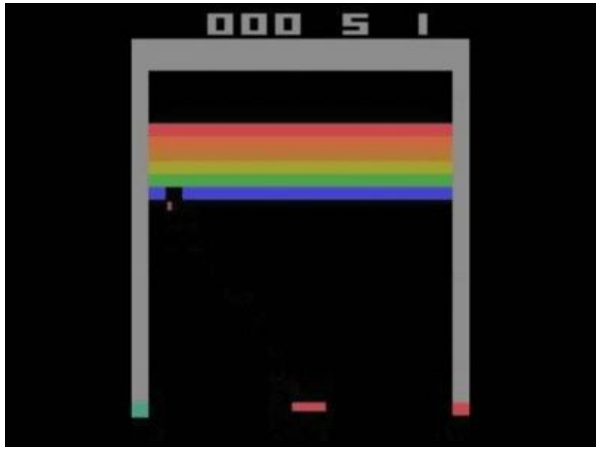
a. PPO (Proximal Policy Optimization)
b. Score: avg 200, std 0

Pendulum: swing pendulum so it stays upright

a. A2C (variant of Asynchronus Advantage Actor Critic)
b. Score: avg -175.399, std 56.19

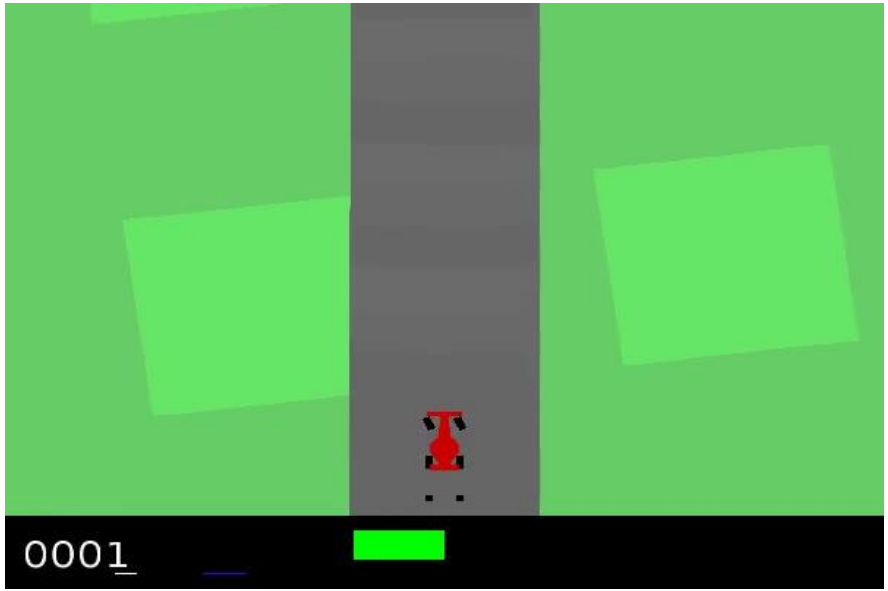


Atari Games



Breakout

a. DDPG (Deep Deterministic Policy Gradient)
b. Avg score 16.9, std 6.86



Car Racing

a. PPO
b. Avg score -76.54, std 2.75

Why Simulate?

1. Overcome hardware limitations
2. Generate more data/scenarios
    a. Tesla: https://www.youtube.com/watch?v=j0z4FweCy4M&t=5784s

Results and Recommendations

Simple environments are easily solvable, but mastery requires much more

More complex environments need more resources(time/computation)

a. Will also require custom implementations for greater functionality

b. Simulation may still be cheaper than real world training

Real world applications exist

a. Faster progress with human feedback?

b. https://venturebeat.com/2021/07/16/openai-disbands-its-robotics-research-team/