

Human and multi-agent collaboration in a human-MARL teaming framework

Neda Navidi, Francois Chabot, Sagar Kurandwad, Irv
Lustigman, Vincent Robert, Gregory Szriftgiser, Andrea
Schuch

AIR (AI Redefined)

Human-In-The-Loop ML

- imitation learning
- behavioral cloning
- generative adversarial imitation learning

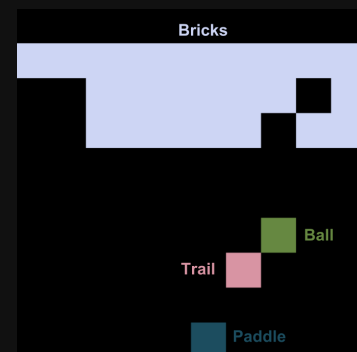
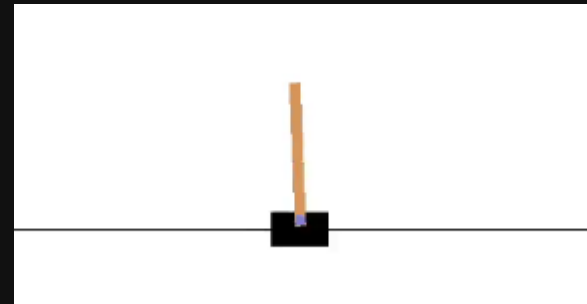
COGMENT: human-MARL teaming framework

environments

- OpenAI Gym
- PettingZoo
- MinAtar

frameworks

- TensorFlow
- PyTorch

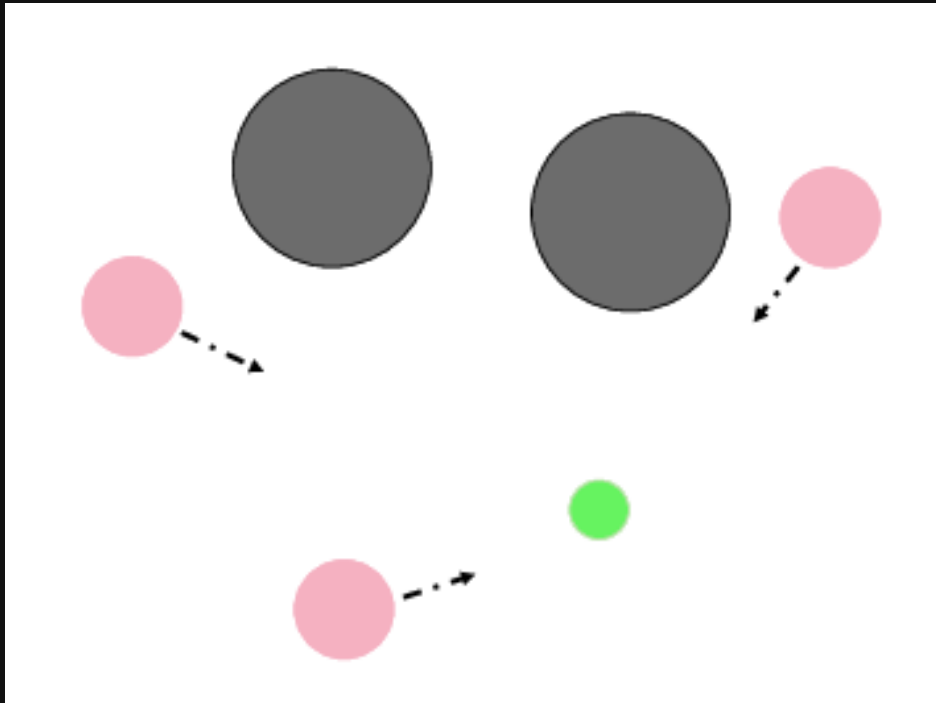


Components of COGMENT

- Actor(Agents or Clients)
- Orchestrator
 - manage all communications
 - handle rewards

Mixed up in a single use case

independent agent and human player



Monitoring agents

Experiments + < /benchmark_lander_hill Share

Search Experiments

- Default [↗](#) [🗑️](#)
- /connect_four_rainbow [↗](#) [🗑️](#)
- /cartpole_dqn [↗](#) [🗑️](#)
- /cartpole_rainbow [↗](#) [🗑️](#)
- /benchmark_lander_...** [↗](#) [🗑️](#)
- /play [↗](#) [🗑️](#)

Track machine learning training runs in experiments. [Learn more](#) ×

Experiment ID: 4

▶ Description [Edit](#)

[Refresh](#) [Compare](#) [Delete](#) [Download CSV](#) [Start Time](#) [All time](#)

[Columns](#) Only show differences [Search](#) [Filter](#) [Clear](#)

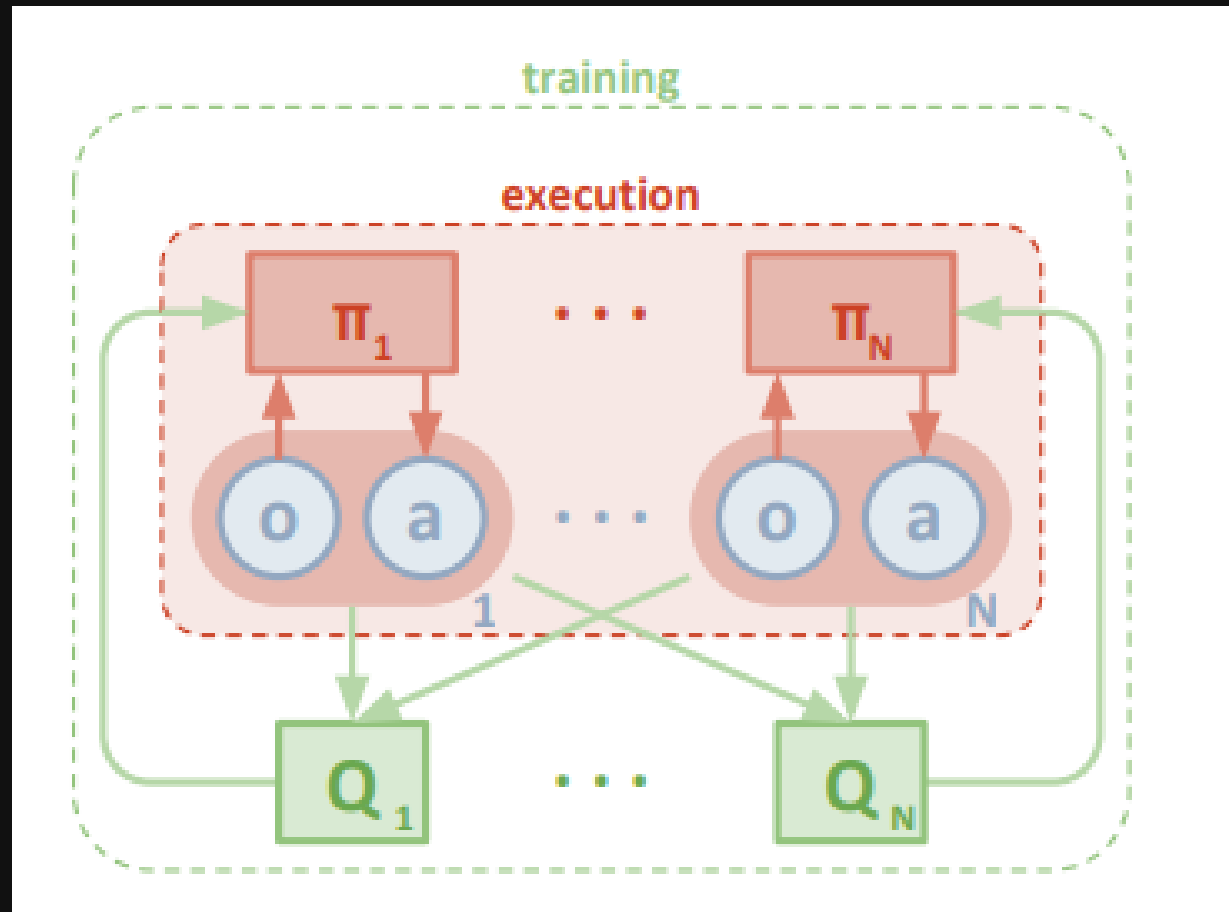
Showing 6 matching runs

								Metrics <						
<input type="checkbox"/>	↓ Start Time	Duration	Run Name	User	Source	Version	Models	batch_done	batch_reward	episodes_per_s	epsilon	loss	lr	mean_trial_rew
<input type="checkbox"/>	29 minutes ago		stupefied_...	unknown	-	-	-	0.016	-1.667	0.017	0.1	2.974	1.000e-4	-105.2
<input type="checkbox"/>	3 hours ago		funny_moore	unknown	-	-	-	0.012	-1.719	0.096	0.1	3.043	1.000e-4	-107.3
<input type="checkbox"/>	6 days ago		elastic_tharp	unknown	-	-	-	0	-0.031	0.003	0.1	2.759	1.000e-4	-167.8
<input type="checkbox"/>	7 days ago		peaceful_w...	unknown	-	-	-	0.004	-0.013	2.402e-4	0.1	2.575	1.000e-4	-180.3
<input type="checkbox"/>	11 days ago		clever_khor...	unknown	-	-	-	0.012	-1.035	0.141	0.1	2.4	1.000e-4	-178.3
<input type="checkbox"/>	11 days ago		objective_e...	unknown	-	-	-	0.012	-1.199	0.078	0.1	2.976	1.000e-4	-184.9

[Load more](#)

MADDPG

Centralized learning, Decentralized execution



D3-MADDPG

(Double Dueling Deep Q learning)

- Dueling: separate “state” value and “action” value
- Double: prevent overestimation



What I want to do

- Compare maddpg-based algorithm in several environments
- get many human play

Thank you!