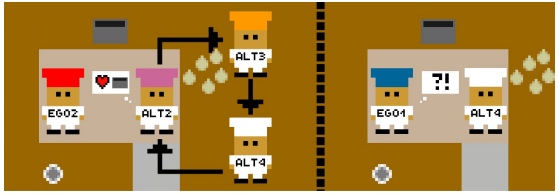
PantheonRL: A MARL Library for Dynamic Training Interactions

Bidipta Sarkar, Aditi Talati, Andy Shih, Dorsa Sadigh

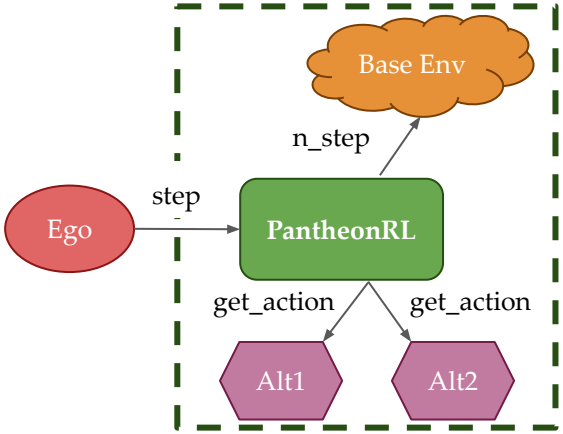
Overview

Supports Complex Training Interactions with existing single-agent RL algorithms



Round-Robin *Cross-Play*

PantheonRL Framework



Agent Design

Ego: Works with **any** gym-based RL algorithm

Partner: Any Stable-Baselines3 algorithm

- OnPolicyAgent: PPO, A2C
- OffPolicyAgent: DQN, SAC, …
- StaticPolicyAgent: ActorCriticPolicy

Can create custom partners with hard-coded get_action and update

Example Scripts

Round-Robin Training

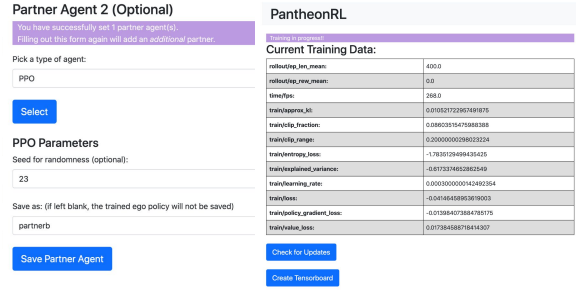
```
env = gym.make('OvercookedMultiEnv-v0', layout_name='simple')
partner_1 = PPO.load(partner_1_file)
env.add_partner_agent(StaticPolicyAgent(partner_1))
partner_2 = A2C('MlpPolicy', env)
env.add_partner_agent(OnPolicyAgent(partner_2))
ego = PPO('MlpPolicy', env, verbose=1)
ego.learn(total_timesteps=10000)
```

Ad-hoc Adaptation

```
env = gym.make('OvercookedMultiEnv-v0', layout_name='simple')
env.add_partner_agent(StaticPolicyAgent(partner_2.policy))
ego_b = PPO.load(ego_b_file, env=env)
ego_b.learn(total_timesteps=500000)
```

Web Interface

Allows for asynchronous training and monitoring

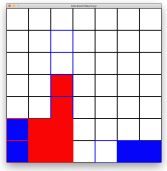


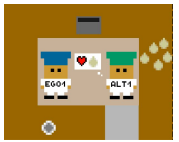
Use Cases

Decentralized MARL Training
Human-AI Interaction
Independent AI Collaboration

Supported Environments

Block World | Overcooked | PettingZoo





PettingZoo