ilpyt: Imitation Learning Research Code Base in PyTorch

Amanda Vu, Alex Tapley, Brett Bissey
The MITRE Corporation

Code base, model zoo, documentation, and datasets: https://github.com/mitre/ilpyt

Summary

ilpyt is an imitation learning (IL) research library which implements a variety of IL and reinforcement learning (RL) algorithm families in a shared infrastructure. The library contains benchmark implementations of common IL algorithms, written in a modular fashion for easy user customization, novel implementation, and benchmarking.

Imitation learning shows promise for teaching safe agent behaviors in increasingly dynamic environments by implicitly bounding behaviors to lay in the field of human demonstration and tackling the sample efficiency issues of RL methods.

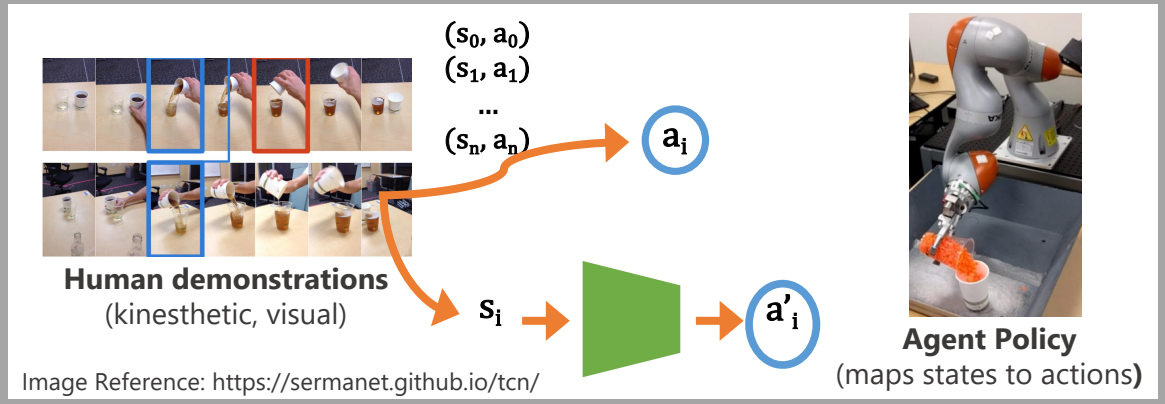
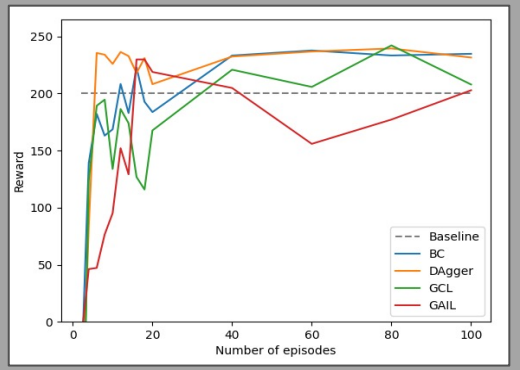
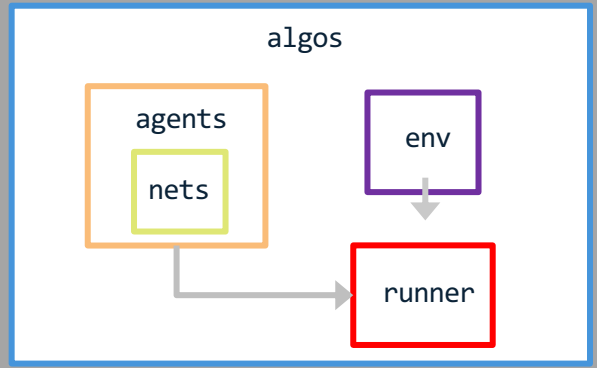


Image Reference: https://sermanet.github.io/tcn/

```python
import ilpyt
from ilpyt.agents.imitation_agent import ImitationAgent
from ilpyt.algos.bc import BC

env = ilpyt.envs.build_env(env_id='LunarLander-v2',  num_env=16)
net = ilpyt.nets.choose_net(env)
agent = ImitationAgent(net=net, lr=0.0001)
algo = BC(agent=agent , env=env)
algo.train(num_epochs=10000, expert_demos='demos/LunarLander-v2/demos.pkl')
algo.test(num_episodes=100)
```





Key Toolbox Features

- Unified framework for the training, testing, and eval of IL/RL algorithms
- Popular IL and RL baselines established, with model zoo
- Simple API exposing *train* and *test* methods for easy setup and usage
- Modular infrastructure for easy modification and reuse of existing components in novel algorithm development
- Parallel and serial environment modes for optimized training and eval
- Compatibility with the OpenAI Gym environment interface for access to most existing benchmark environments as well as custom environments

Benchmarks

	CartPole -v0	MountainCar -v0	MountainCar Continuous-v0	LunarLander -v2	LunarLander Continuous-v2
Threshold	200	-110	90	200	200
Expert (Mean/Std)	200.00 / 0.00	-98.71 / 7.83	93.36 / 0.05	268.09 / 21.18	283.83 / 17.70
BC (Mean/Std)	200.00 / 0.00	-100.800 / 13.797	93.353 / 0.113	244.295 / 97.765	285.895 / 14.584
DAgger (Mean/Std)	200.00 / 0.00	-102.36 / 15.38	93.20 / 0.17	230.15 / 122.604	285.85 / 14.61
GAIL (Mean/Std)	200.00 / 0.00	-104.31 / 17.21	79.78 / 6.23	201.88 / 93.82	282.00 / 31.73
GCL (Mean/Std) [2]	200.00 / 0.00	- [1]	- [1]	212.321 / 119.933	255.414 / 76.917
AppL(Mean/Std)	200.00 / 0.00	-108.60 / 22.843	- [3,5]	- [4]	- [3,4,5]
DQN (Mean/Std)	-	-	-	281.96 / 24.57	-
A2C (Mean/Std)	-		-	201.26 / 62.52	-
PPO (Mean/Std)	-	-	-	249.72 / 75.05	-