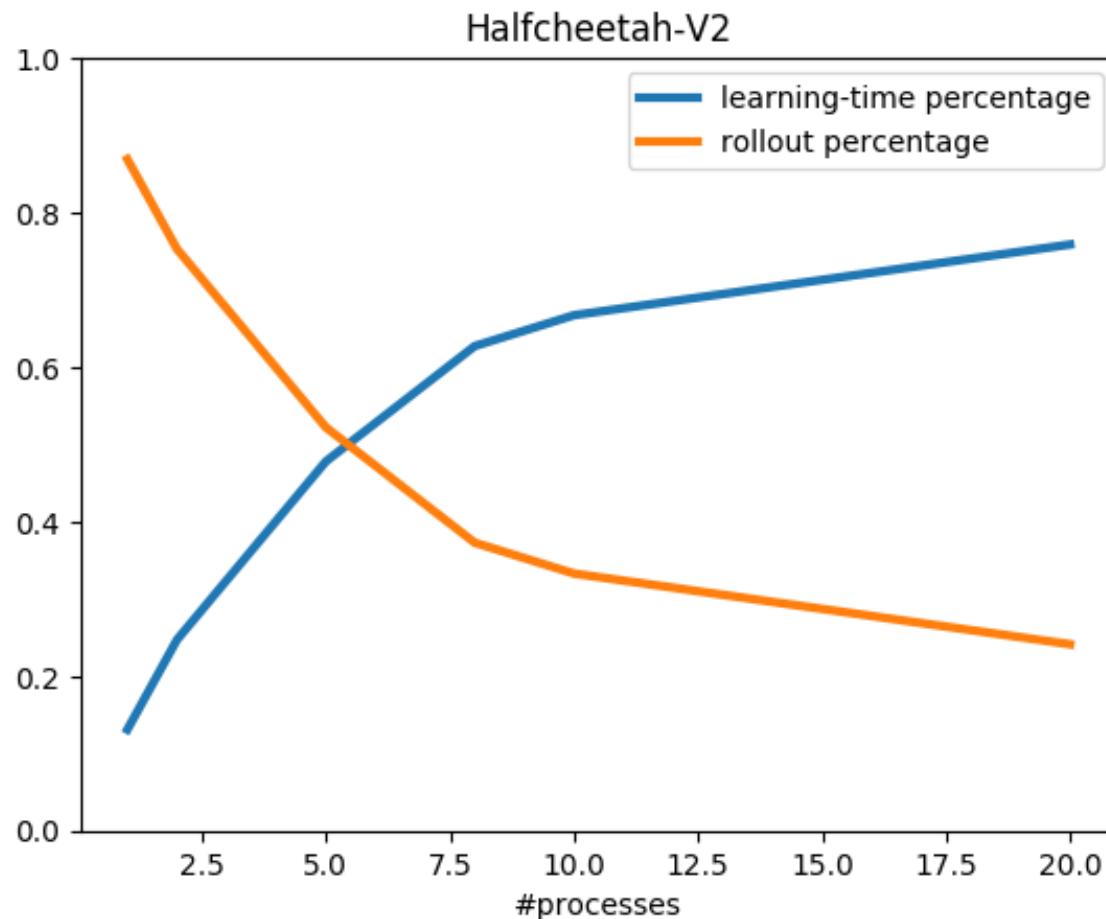


Speedup Reinforcement Learning Framework

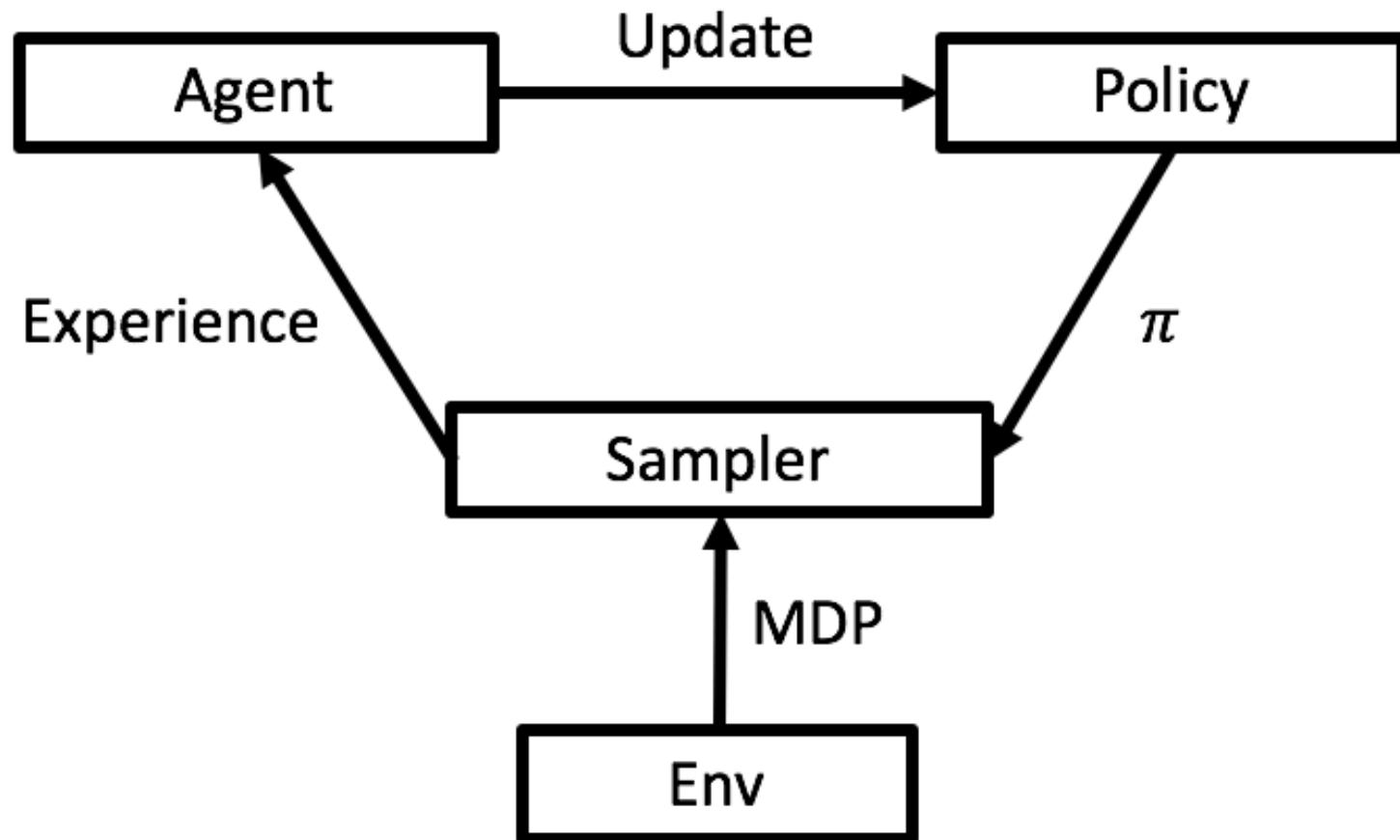
Achieve *near linear* speedup
for Experience Collection Time with multi-process support

Codebase: <https://github.com/tianbingsz/WALL-E>

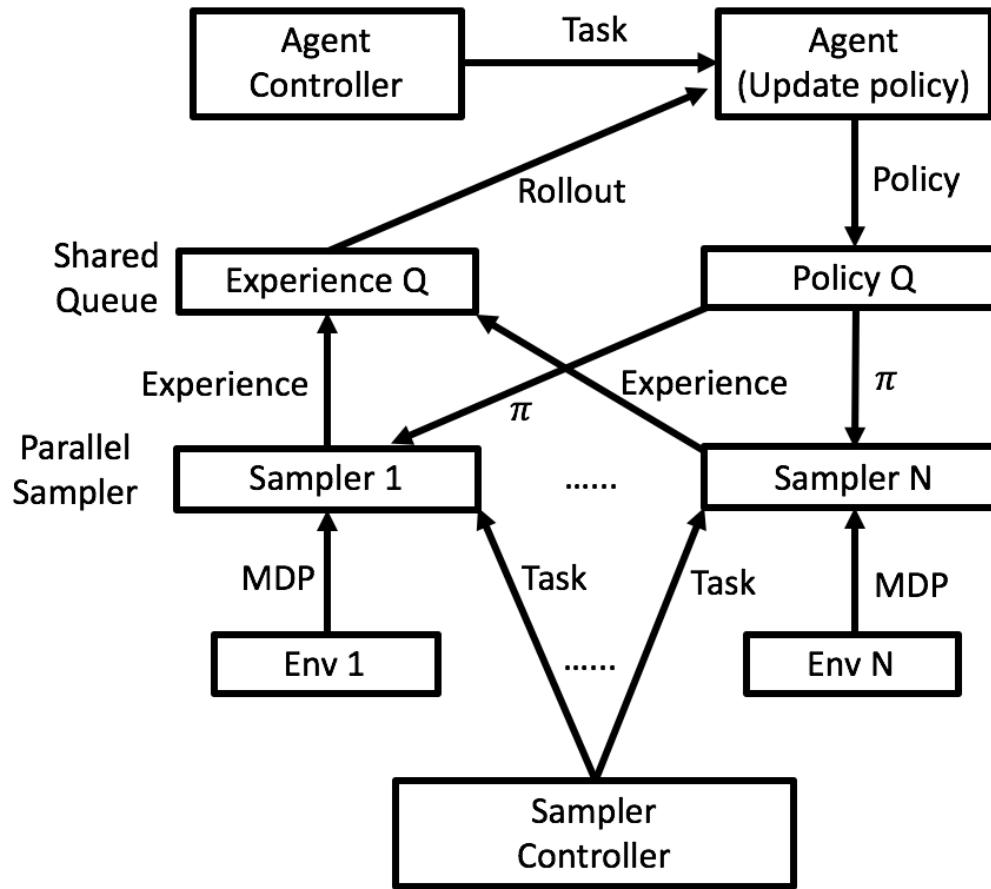
Running Time = Experience Collection+ Policy Learning Time



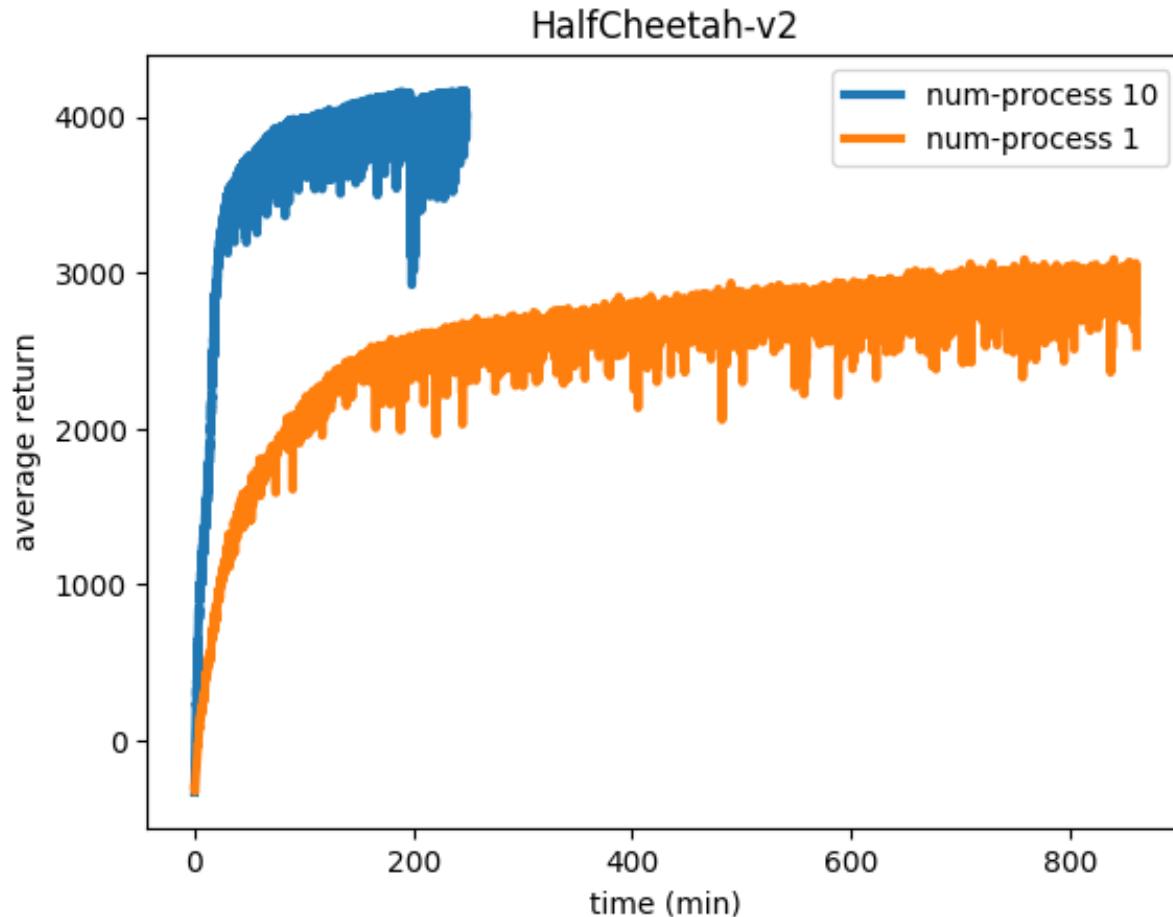
General Reinforcement Learning Framework



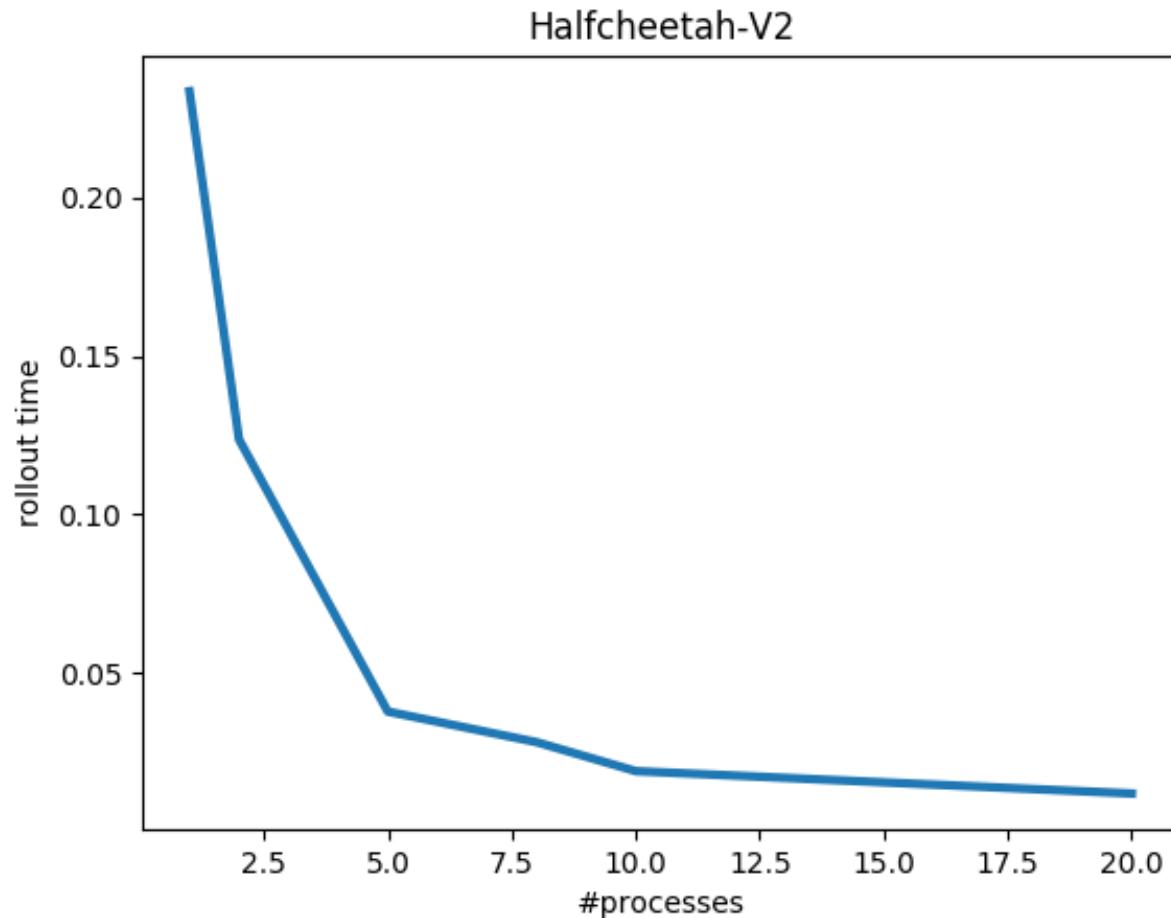
Multi-Process Experience Collection



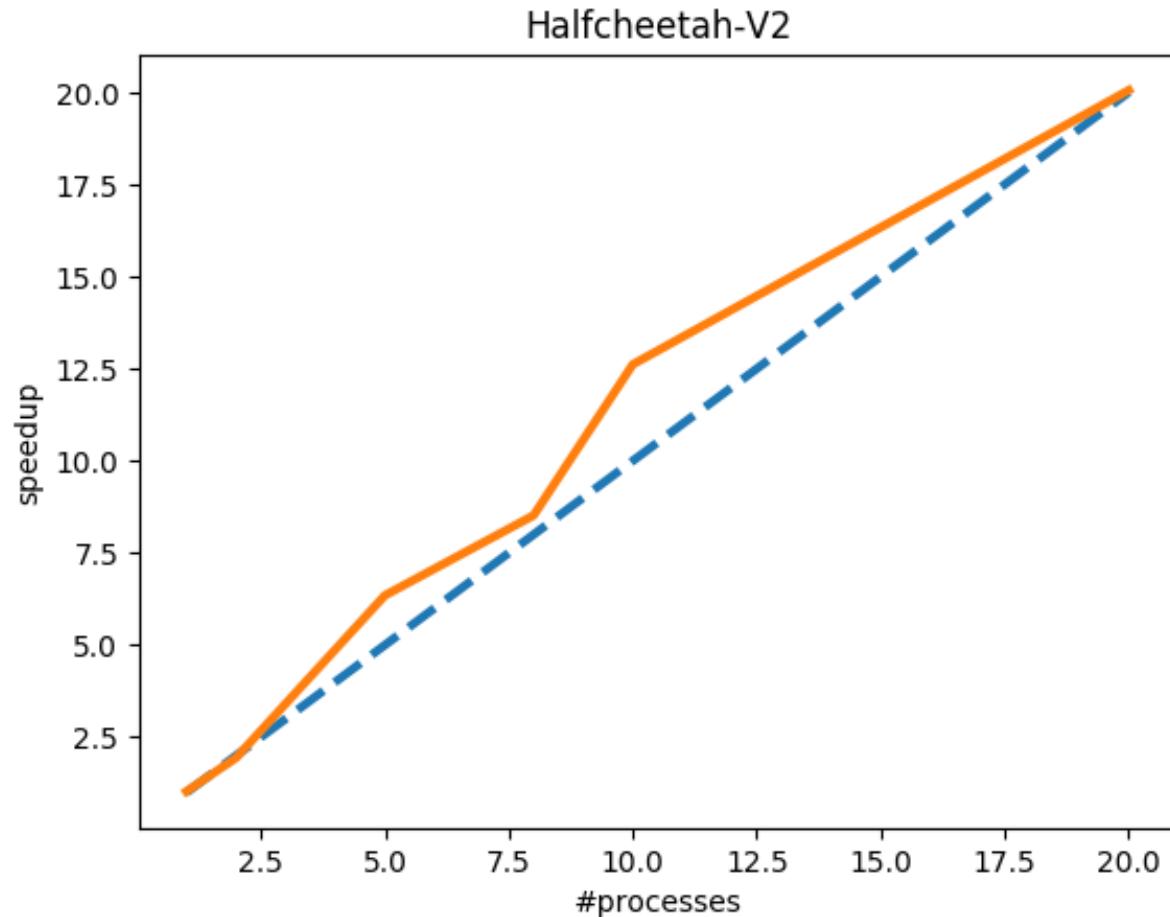
Performance on Mujoco Control task



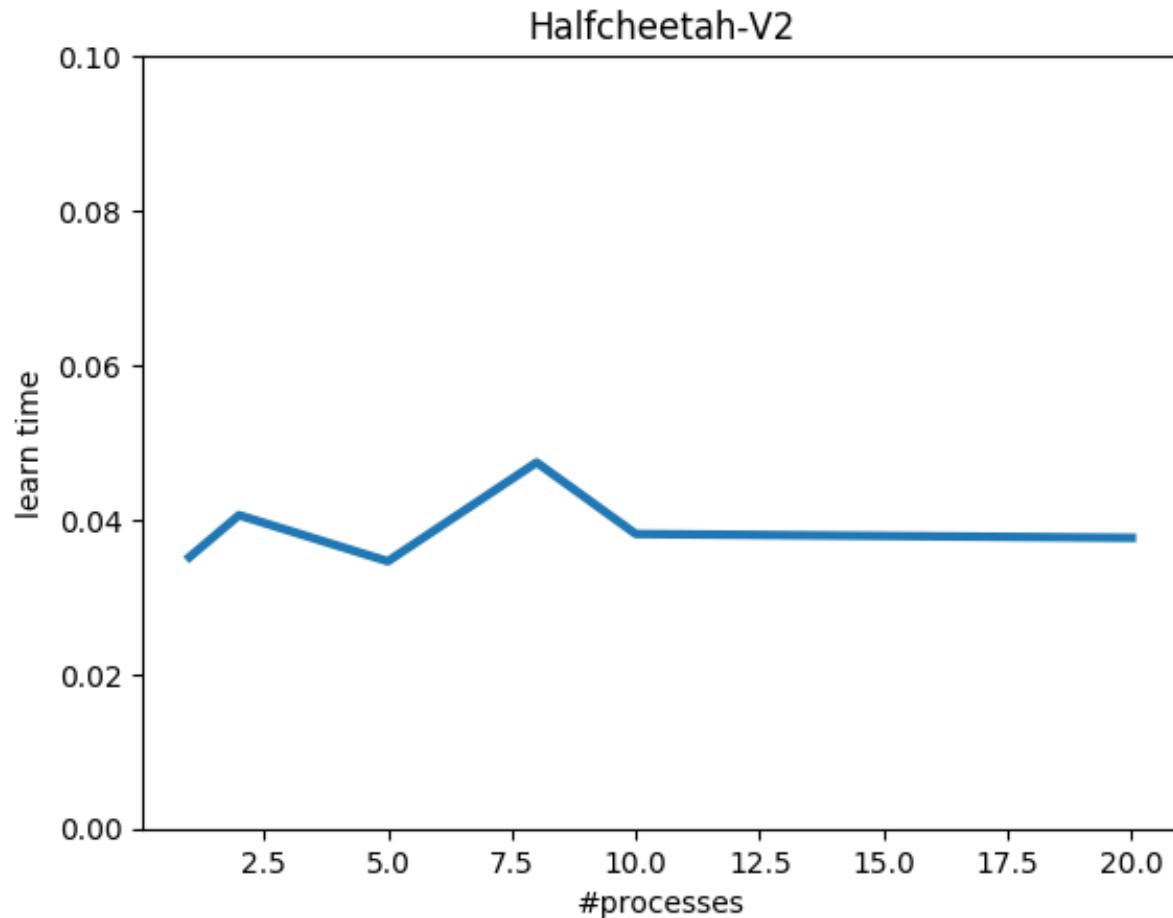
Experience Collection Time Reduction



Experience Collection Time Speedup



Policy Learning Time w.r.t. Processes

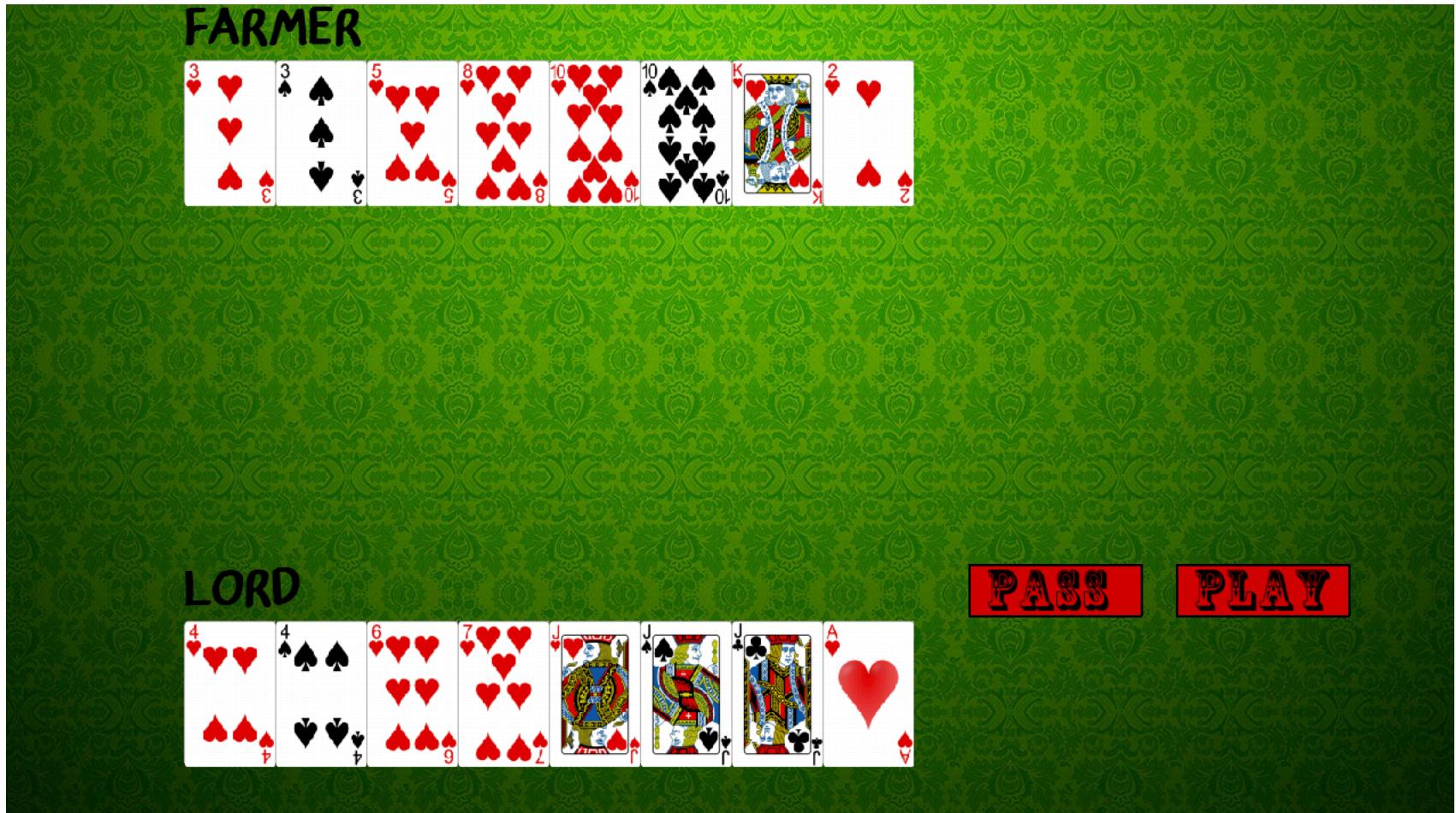


Future work

- Speedup Policy Learning (Perception to Control Policy Network) with Multi-GPU support
- Generalize the codebase to support off-policy learning
- Model-based RL, Planning, Exploration and more for Robotics

Game AI (Lord vs Farmer)

Codebase (C++/Python) <https://github.com/tianbingsz/LordGame>



Agents Playing Poker with Human

- min max tree search
- Speedup computation 100X with Memorization

$$\min_{\pi_F} \max_{\pi_L} R(\text{Lord}, \text{Farmer})$$

Auto Cards Generation

- Generate cards by interactive adversarial self-play
 - Pre-train Lord/Farmer with imitation learning
 - Generate cards via agents playing with each other
 - Improve the agents performance iteratively with generated cards and real human data