Red Team

Kunal Lad

Goal

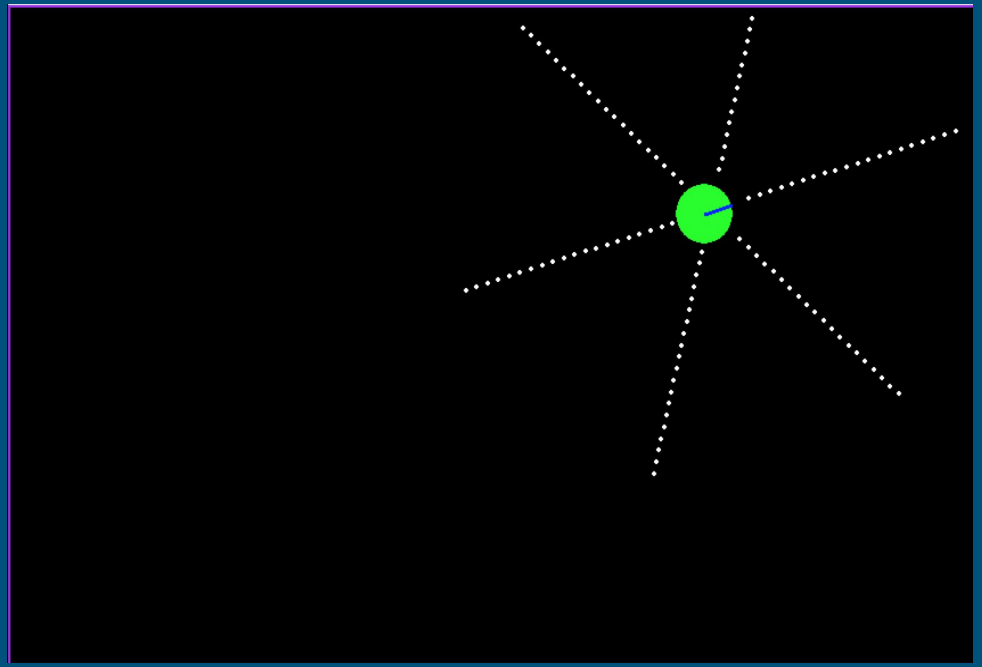
- Adversarial agent which learns to exploit loopholes in agent's policy during training.
- Demonstrate effectiveness of Red Teams in Reinforcement Learning approaches.

Motivation

- Many RL environments have adversarial events which occur rarely
- Epsilon greedy agents might not encounter them enough
- Red Team can train to generate adversarial rare inputs

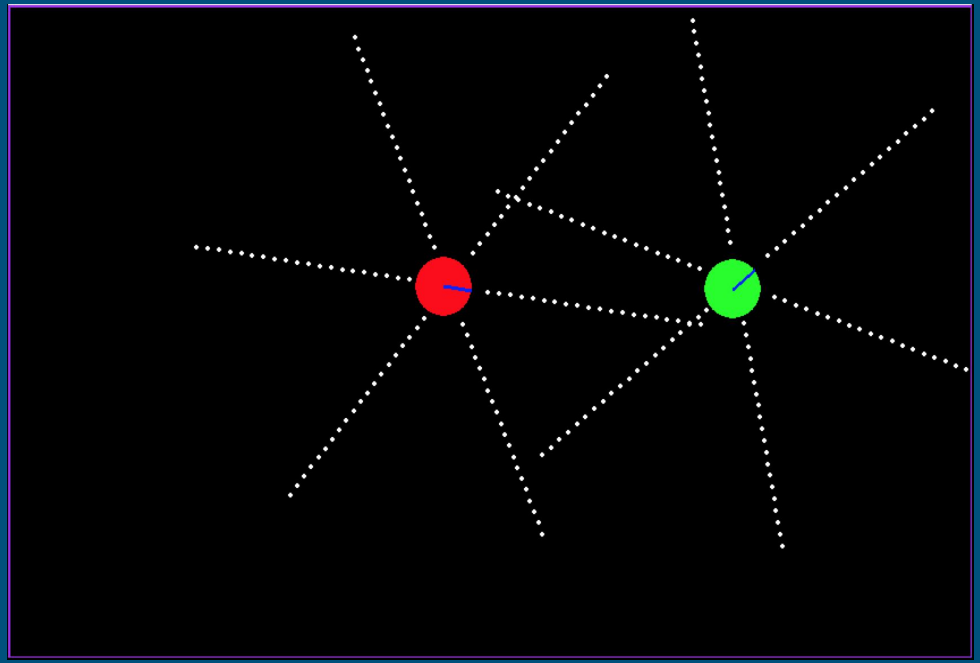
Training Environment

- Partially Observable Environment
  - State: Distance measured along 6 sensory arms
- Discrete Actions
  - Turn Left (+0.2 radians)
  - Turn Right (-0.2 radians)
  - Straight (0 radians)
- Transitions: Physics based simulator
- Reward (Discounted Episodic, $\gamma = 0.9$)
  - -50000 on crash
  - Min Distance on all sensor arms OR Sum of distances on all sensor arms



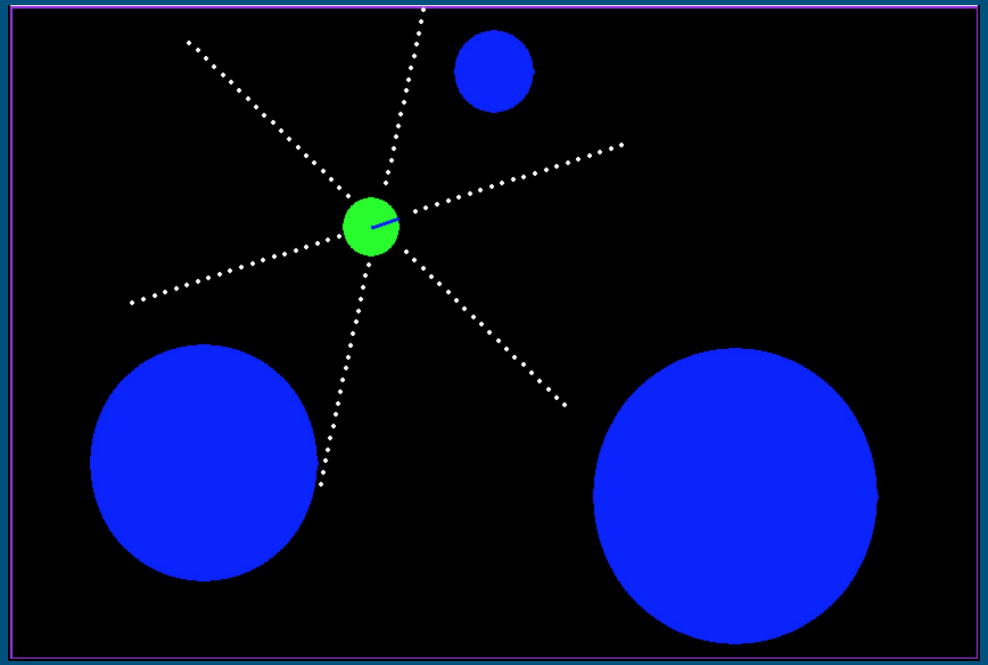
Training Environment (Red Team)

- Partially Observable Environment
  - State: Distance measured along 6 sensory arms
- Discrete Actions
  - Turn Left (+0.2 radians)
  - Turn Right (-0.2 radians)
  - Straight (0 radians)
- Transitions: Physics based simulator
- Reward (Discounted Episodic, $\gamma = 0.9$)
  - -50000 on crash



Test Environment

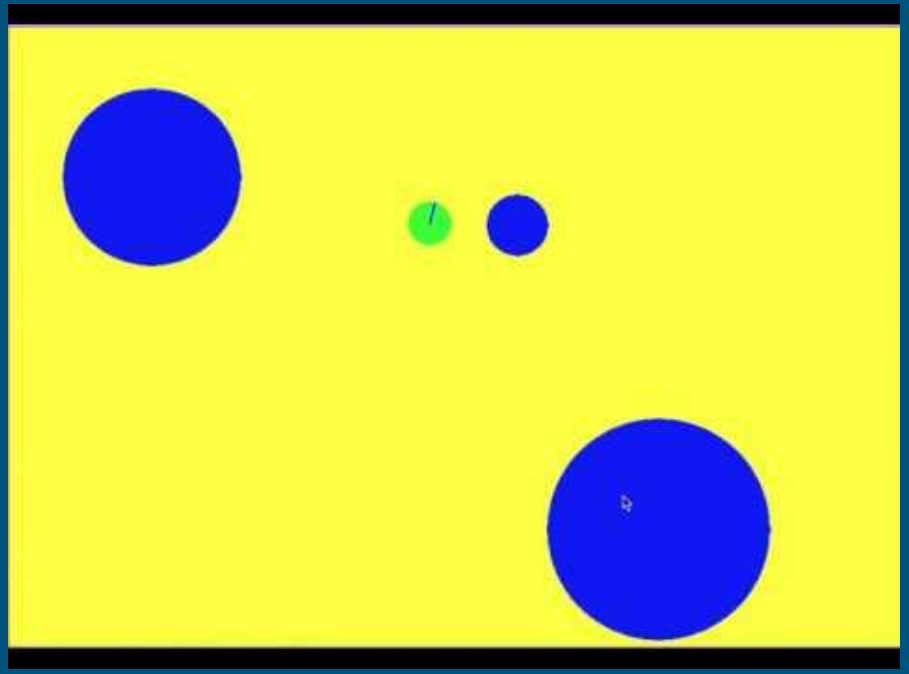
- Partially Observable Environment
  - State: Distance measured along 6 sensory arms
- Discrete Actions
  - Turn Left (+0.2 radians)
  - Turn Right (-0.2 radians)
  - Straight (0 radians)
- Transitions: Physics based simulator
- Reward (Discounted Episodic, $\gamma = 0.9$)
  - +1000 on collision with agent
  - - (Max Distance on all arms) OR - (Sum of distances on all arms)

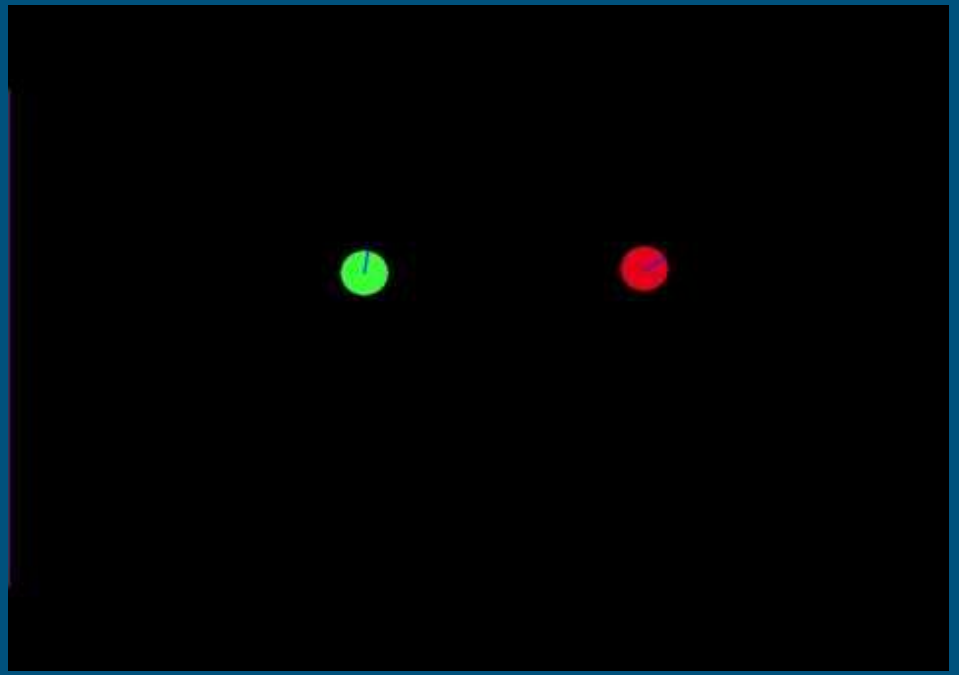


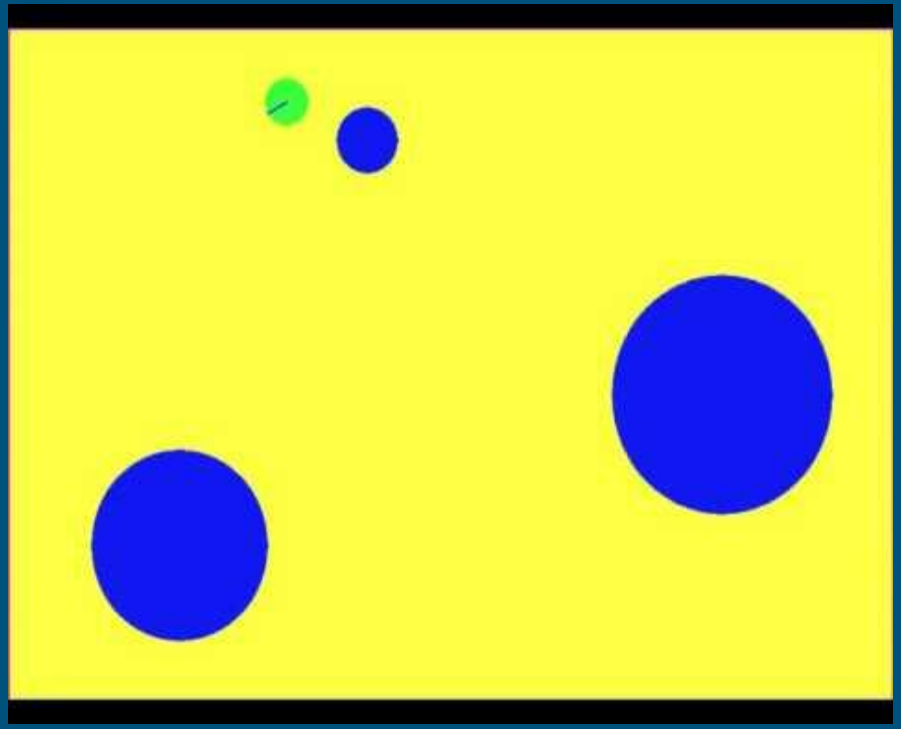
Methodology

- Q Learning with function approximation
- Train separate models for agent and red team (trained jointly)
  - 2 Hidden Layers
  - Relu Activations
  - Adam Optimizer
  - Replay (Buffer size 50000)

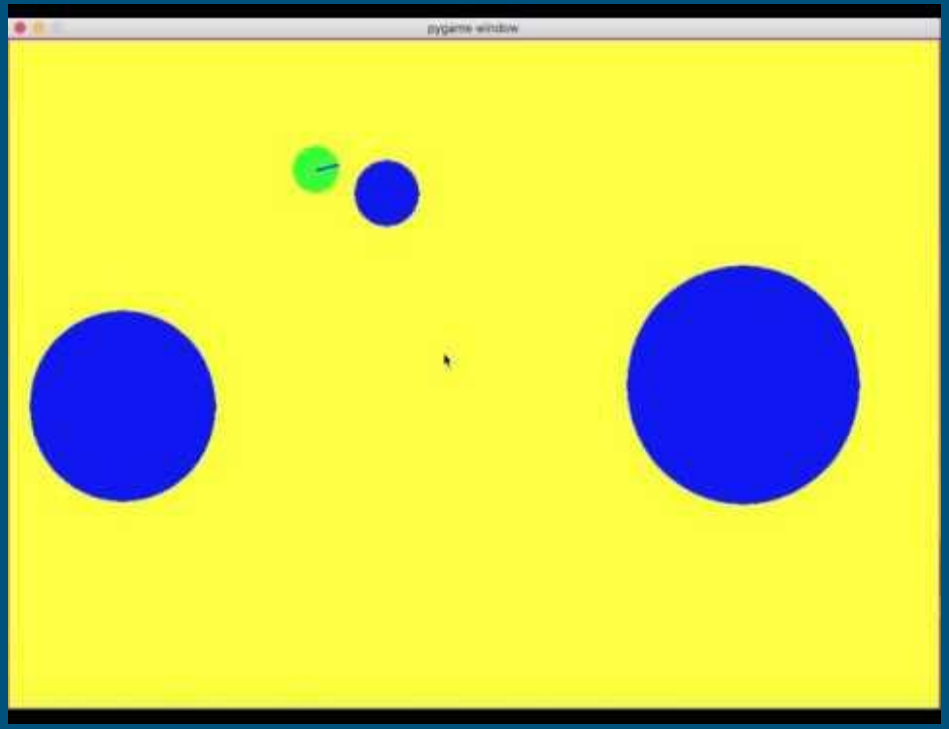
Results







Normal Agent



Red Team

Conclusion

“If you know the enemy and yourself, you need not fear the result of hundred battles”

- Sun Zhu, The Art of War