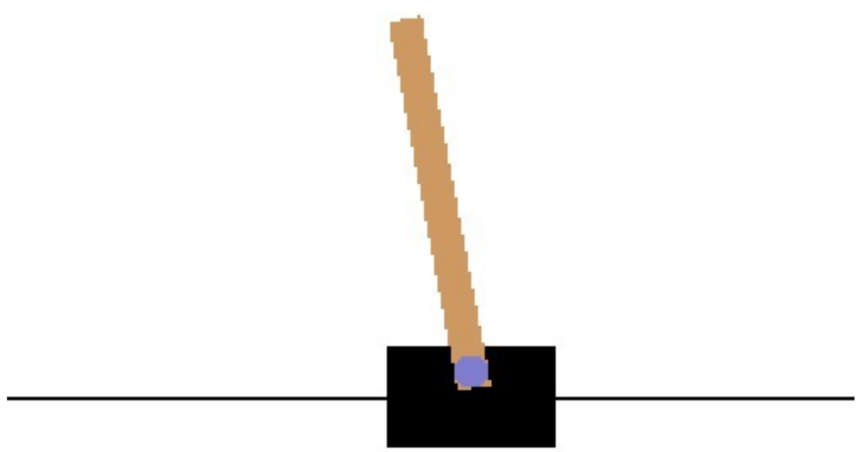
Learning Kendama

Problem Statement:

We used a series of algorithms to learn to swing-up and balance of an inverted pendulum on a cart using the OpenAI Gym, and then extended those algorithms to learn on the more complex game of Kendama (a cup-in-ball game).

Motivation:

Traditional controllers require models of system dynamics and expertise in control tuning. RL can be used to learn a model-free controller of a simple system and then easily expand it to a complex system.



Modified OpenAI simulator used to test swing-up and balance of inverted pendulum.

Kendama, a cup-in-ball game

Explored Algorithms:

Random Search: Choose random parameters that give the highest reward.

Policy Gradient: Given the value of a state, pick the best policy. Update policy by computing its gradient.

Actor-Critic: First, learn a value function given a series of states. Use that to learn an optimal policy.

Value Iteration (Q-learning): Given a policy, update the value of each state to maximize the reward over time.

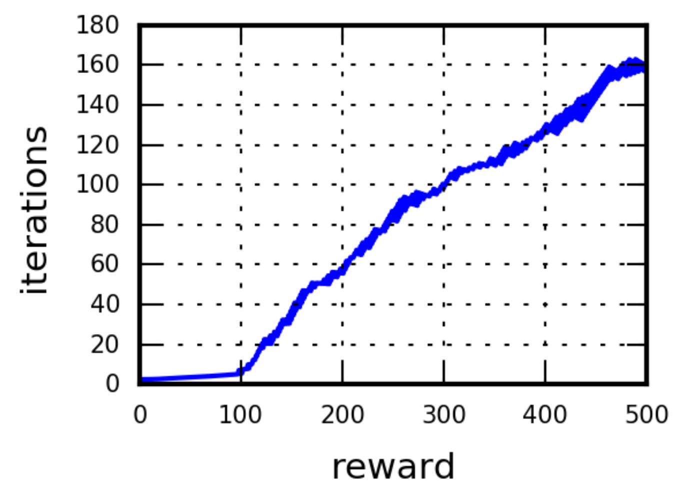
Results:

We found our **policy gradient** algorithm yielded both the fastest learning and the most robust policy.

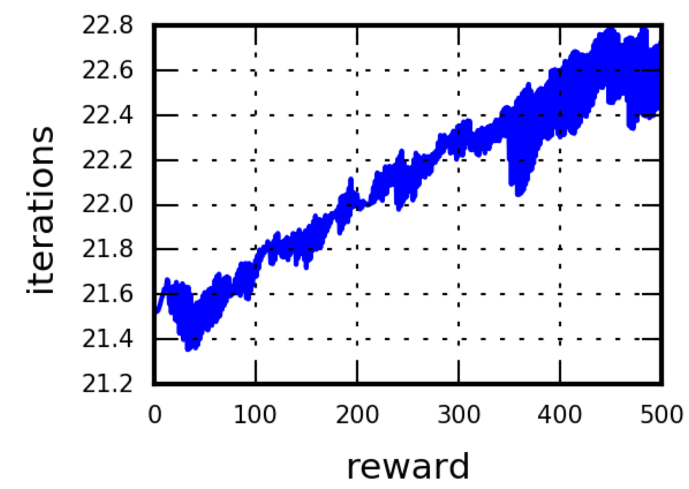
Success Rates:

86/100 trials cartpole

89/100 trials Kendama

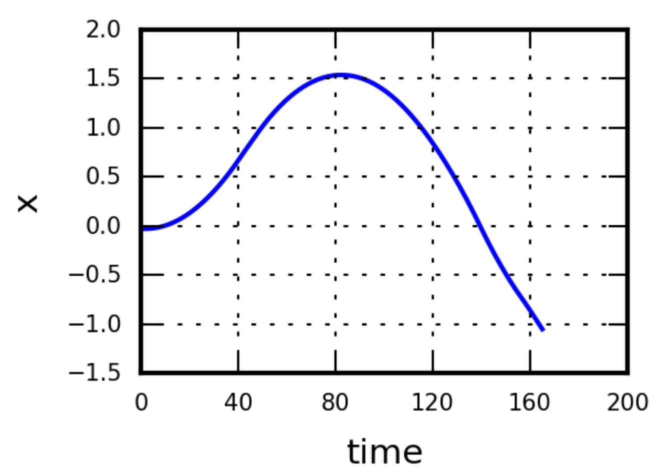


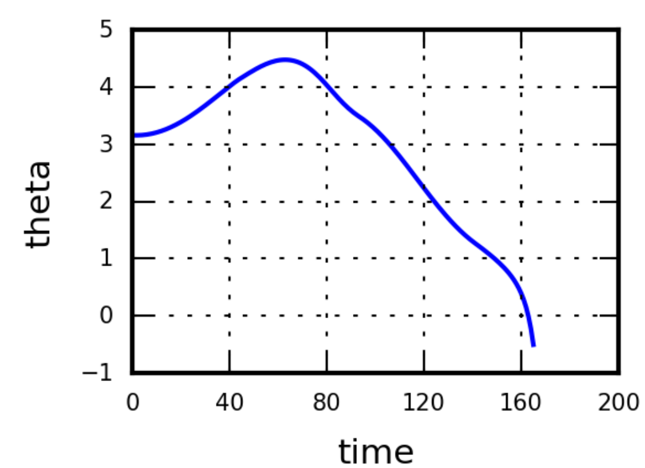
Policy gradient learning curve for Kendama swing-up.

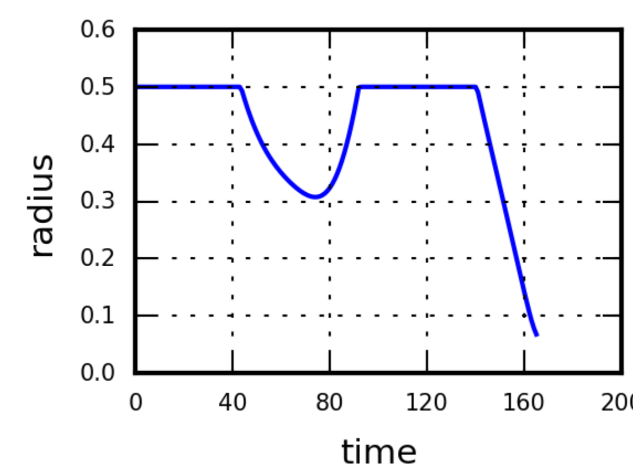


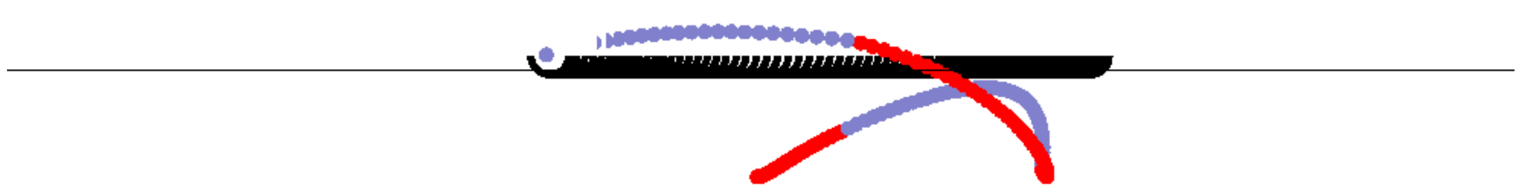
Policy gradient learning curve for Kendama catch.

Simulated Policy:









Plots of *x*, *r*, and θ over time for a successful Kendama catch (top) and a blurred image of that catch (bottom).