Solving Continuous MDPs: The Linear Quadratic Regulator (LQR)

Sanjiban Choudhury

Cornell Bowers CIS
Computer Science

The Big Picture

Case 1: Known MDP

Planning!

Case 2: Unknown MDP, requires feedback from environment

Reinforcement Learning!

Case 3: Unknown MDP, requires feedback from expert

Imitation Learning!

The Big Picture

Case 1: Known MDP

Planning!

Case 2: Unknown MDP, requires feedback from environment

Reinforcement Learning!

We explored this a bit ...

Case 3: Unknown MDP, requires feedback from expert

Imitation Learning!

The Big Picture

Case 1: Known MDP

Now let's go deeper here!

Planning!

Case 2: Unknown MDP, requires feedback from environment

Reinforcement Learning!

Case 3: Unknown MDP, requires feedback from expert

Imitation Learning!

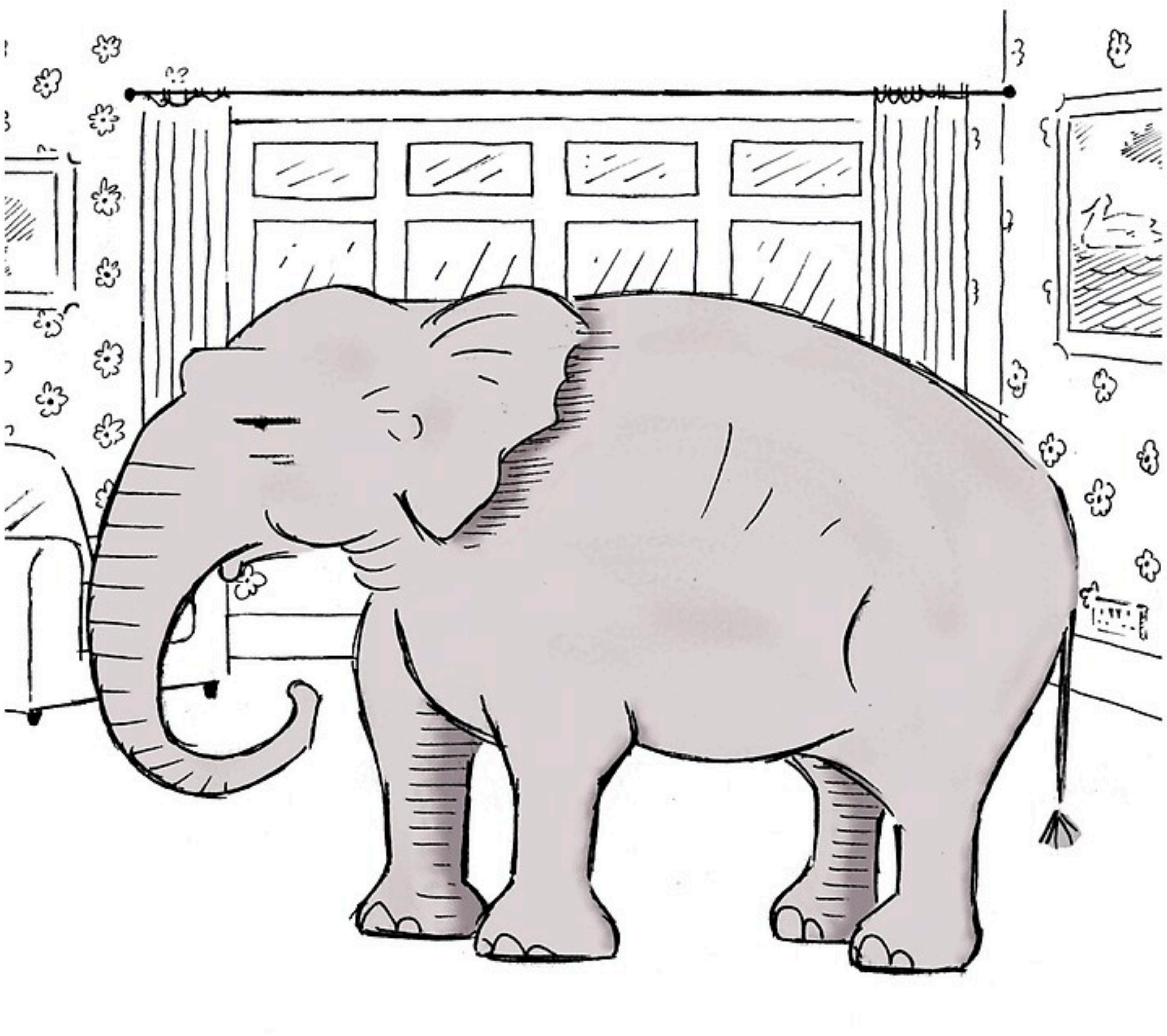
RL

=

Learn model

+

Plan with model



"Just pretend I'm not here..."

Model-based Planning

Step 0: Build a robot

Step 1: Collect data of your robot doing stuff in the world

Step 2: Use data to learn a dynamics model for your robot

Step 3: Plan with the model to compute an optimal policy

Today's Challenge!

Step 0: Build a robot

Step 1: Collect data of your robot doing stuff in the world

Step 2: Use data to learn a dynamics model for your robot

Step 3: Plan with the model to compute an optimal policy

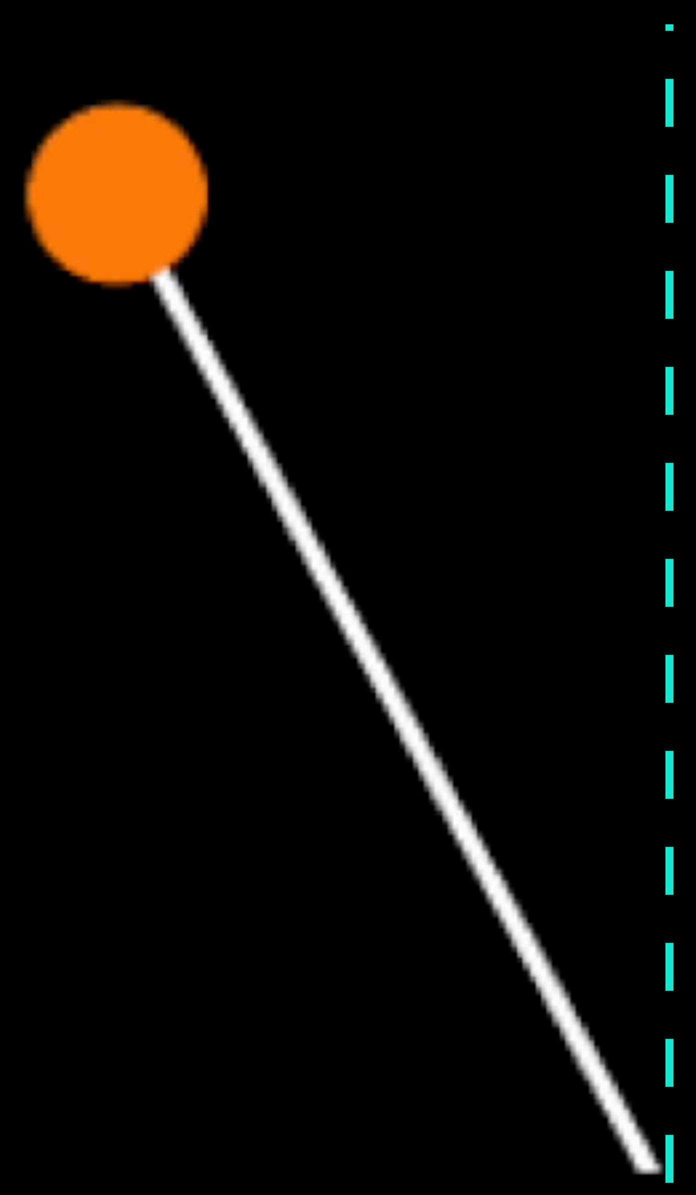
How do we do this for robots with continuous state-actions?


BostonDynamics

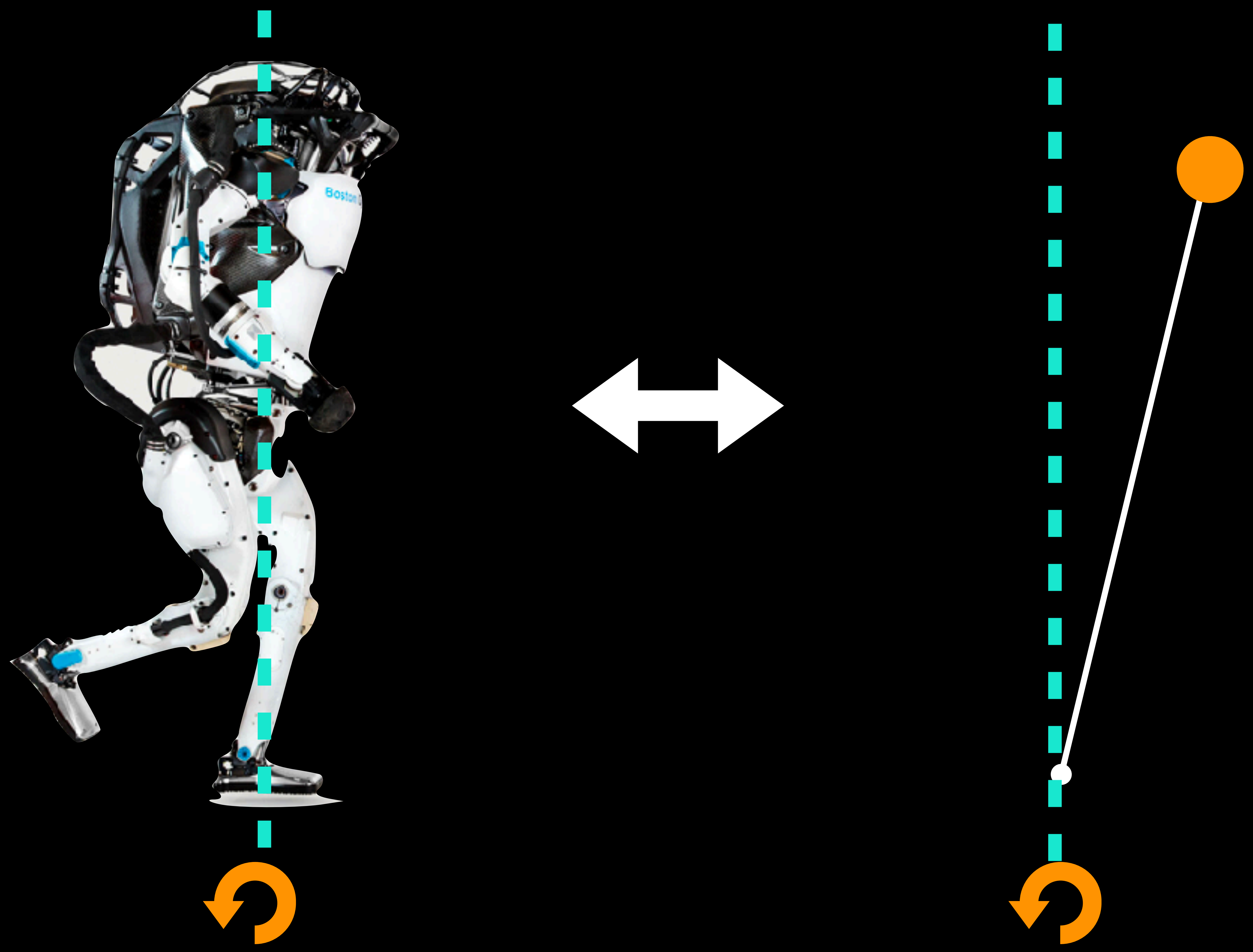
Brainstorm

How do we model the Atlas backflip as a Markov Decision Problem <S, A, C, T>?

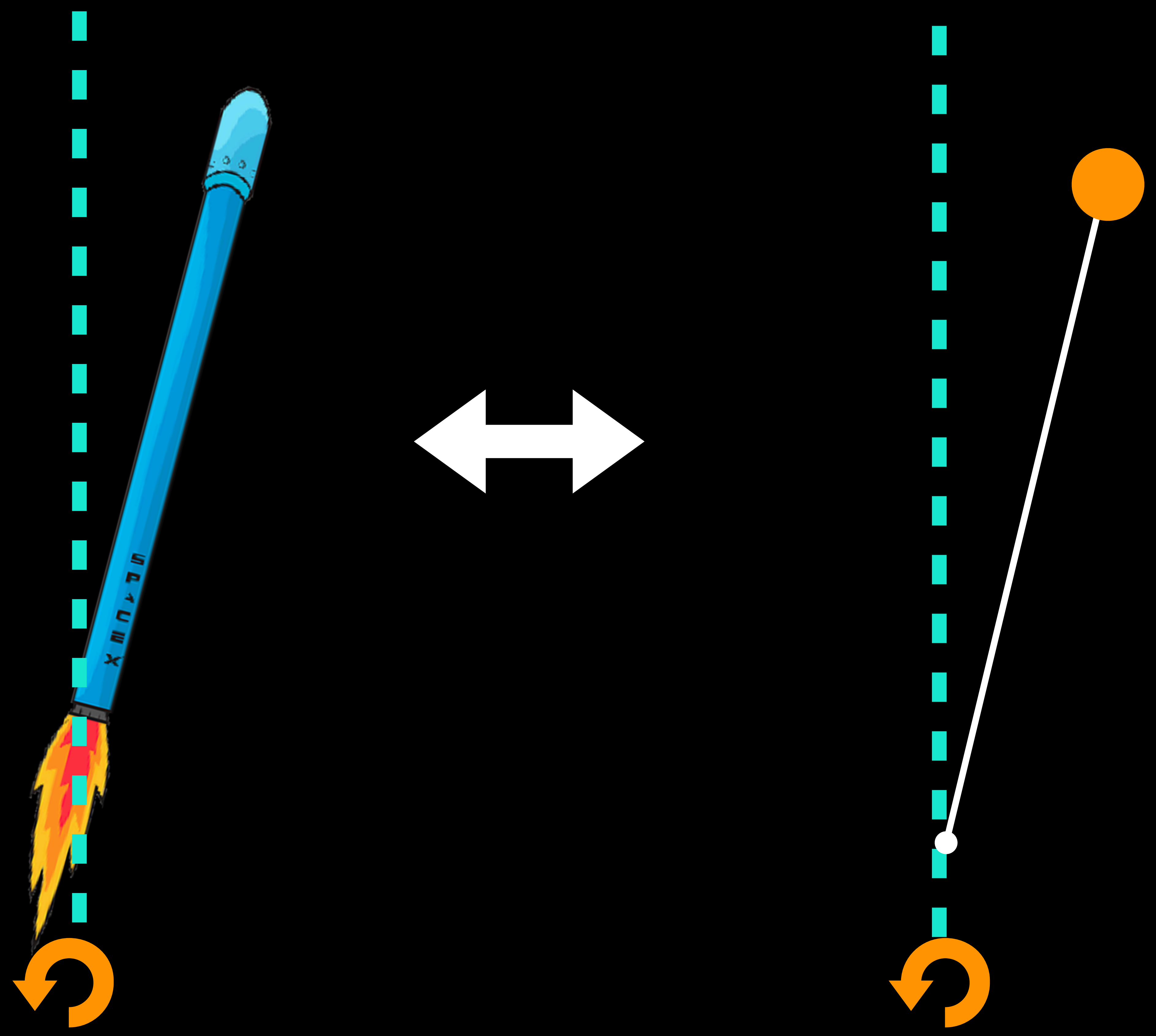
The Inverted Pendulum: A fundamental dynamics model



Humanoid balancing



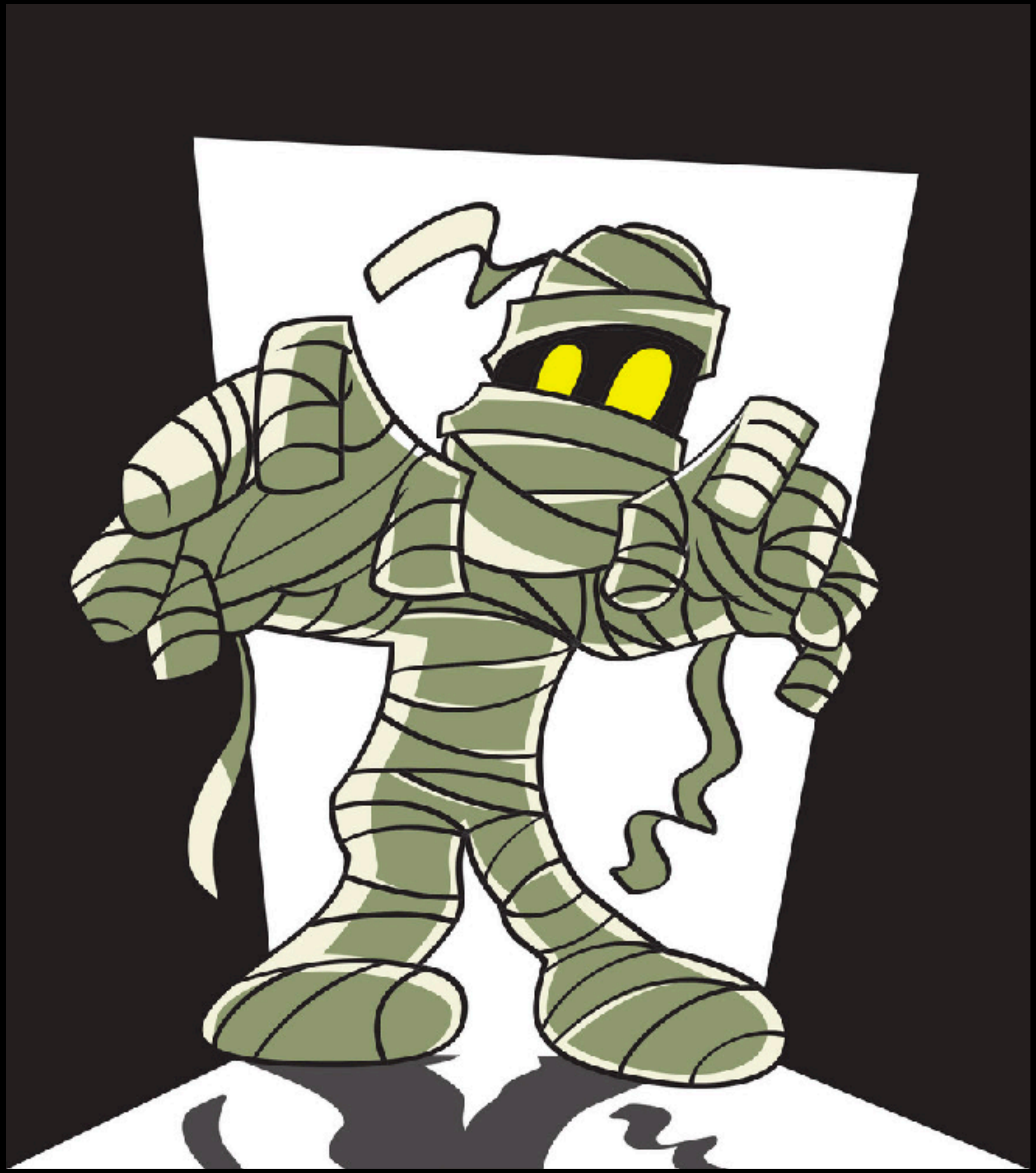
Rocket landing

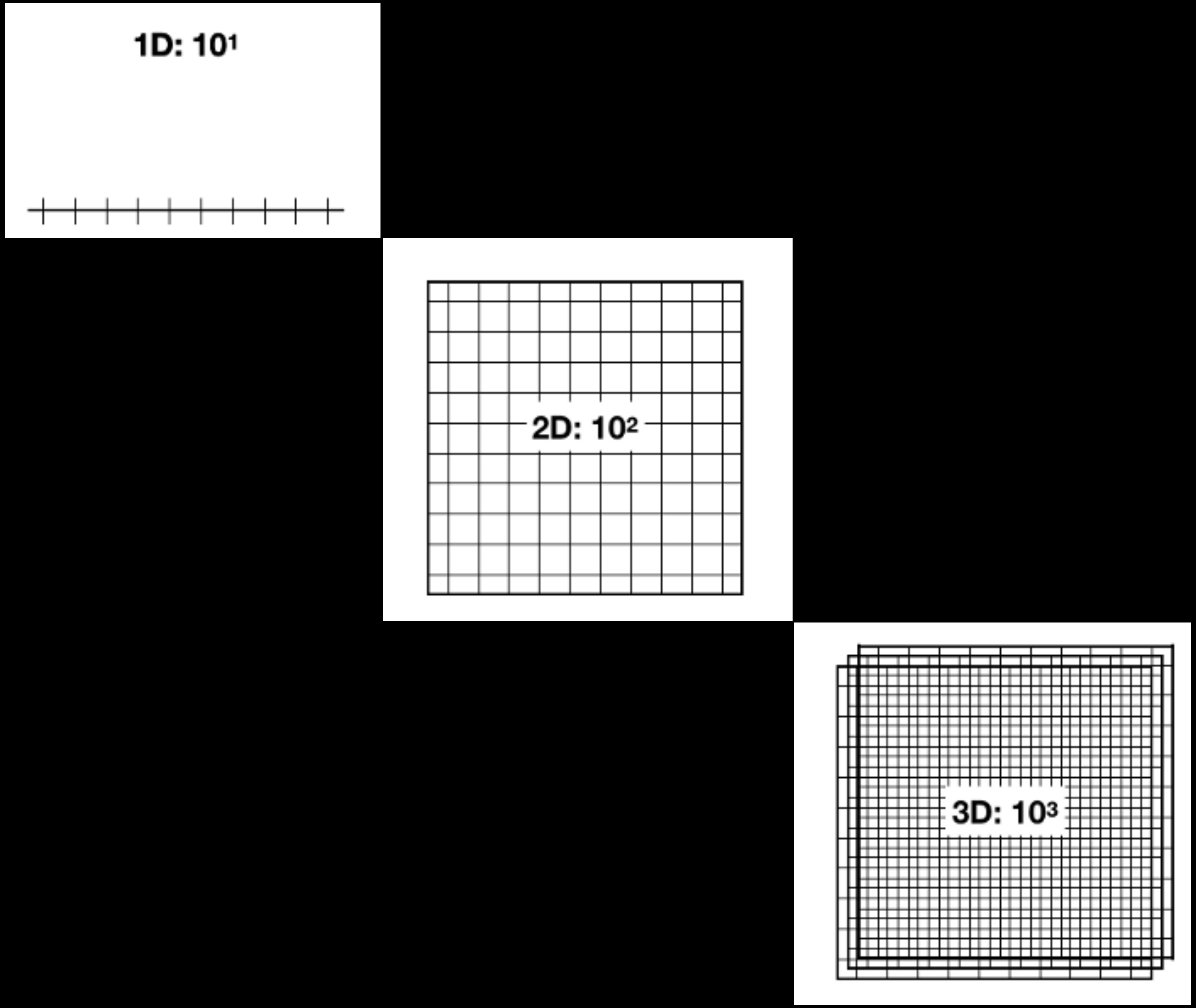


Can we apply value iteration to solve this MDP?

$$V^*(s) = \min_a \left[c(s,a) + \gamma \mathbb{E}_{s' \sim \mathcal{T}(s,a)} V^*(s')\right]$$

The Curse of Dimensionality





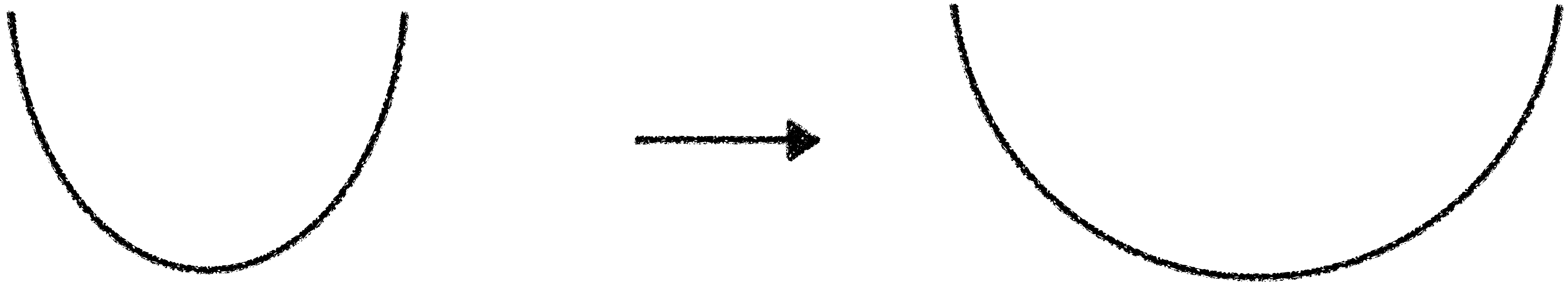
No Discretization!

Can we analytically *represent* and *update* the value function?

$$V^*(s) = \min_a \left[c(s,a) + \gamma \mathbb{E}_{s' \sim \mathcal{T}(s,a)} V^*(s') \right]$$

Can we analytically represent and *update* the value function?

Yes*



*linear dynamics, quadratic costs

Check out notebook

cs4756_robot_learning / notebooks / inverted_pendulum_lqr.ipynb

jren44 Initial commit a6c9feb · on Jan 18 History

Preview Code Blame Raw

Illustrated Linear Quadratic Regulator

Companion to courses lectures from CS6756: Learning for Robot Decision Making and Chapter 2 of Modern Adaptive Control and Reinforcement Learning.

In [3]:

```
import numpy as np
import autograd.numpy as np
from autograd import grad, jacobian
import matplotlib.pyplot as plt
from matplotlib.animation import FuncAnimation
from matplotlib import rc
from IPython.display import HTML, Image
from matplotlib.patches import Circle
rc('animation', html='jshtml')
```

Dynamics of an Inverted Pendulum

Let’s formalize!