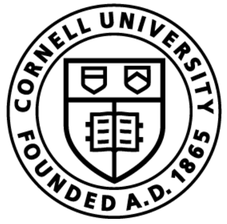
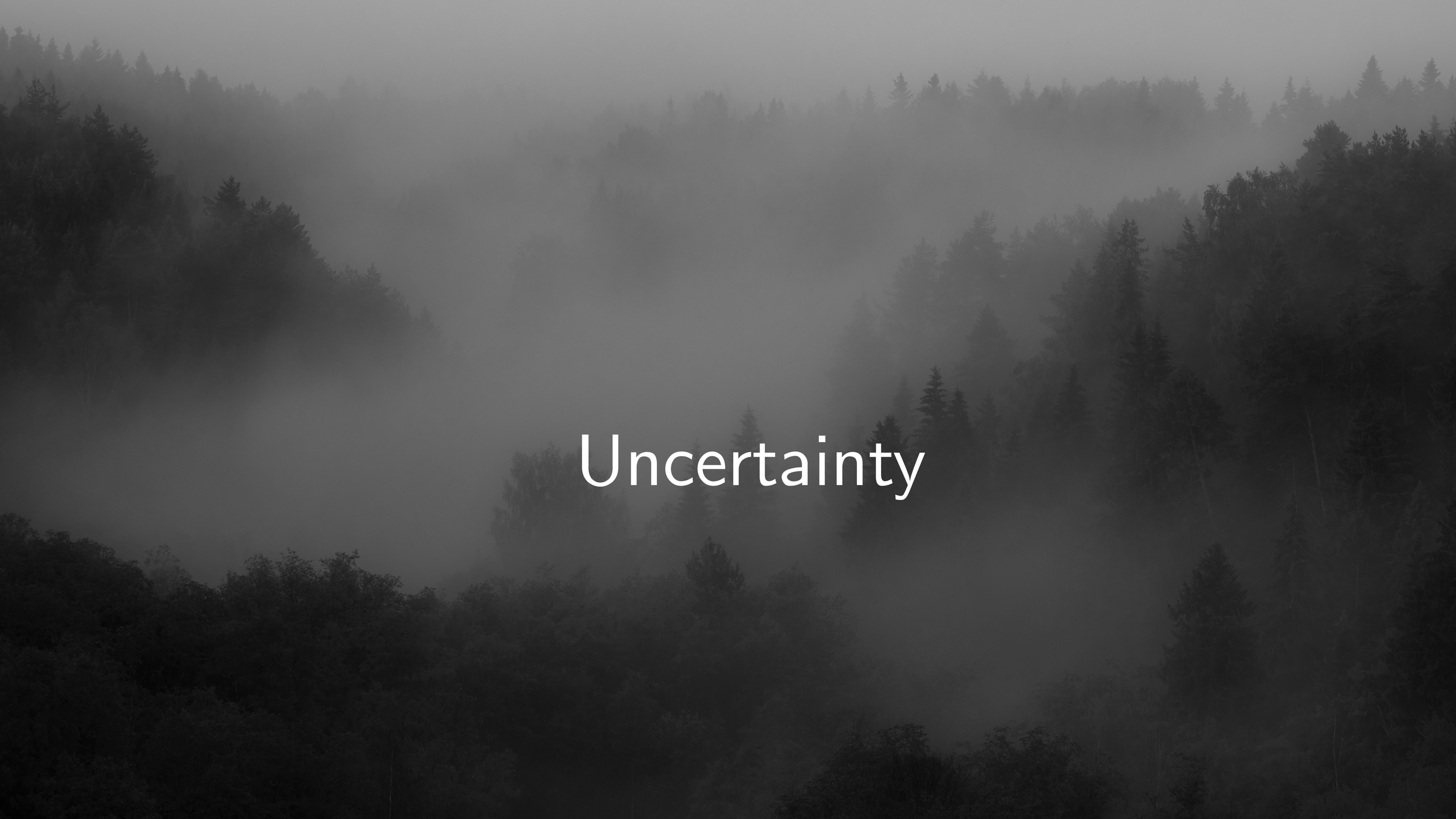
Dealing with Uncertainty: Part 2

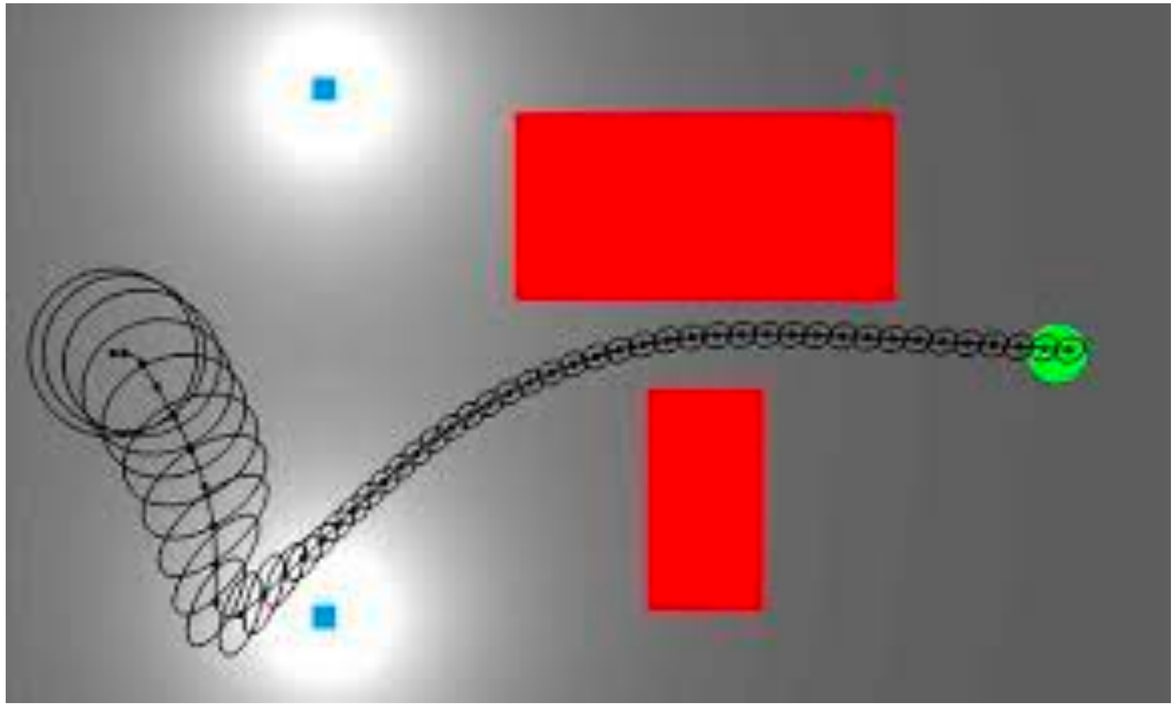
Sanjiban Choudhury



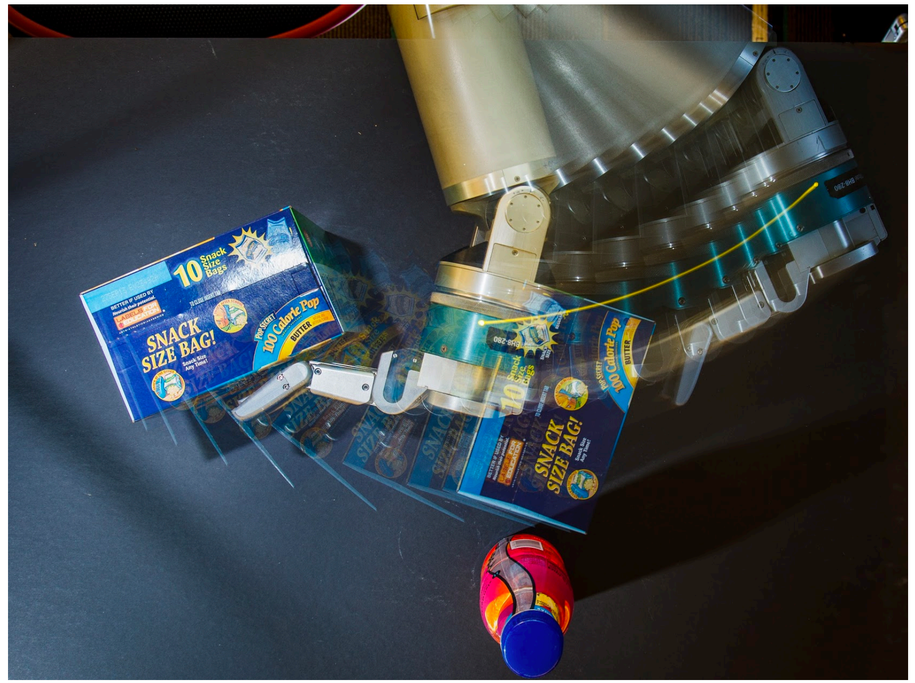
Cornell Bowers CIS
Computer Science


Uncertainty

Epistemic Uncertainty

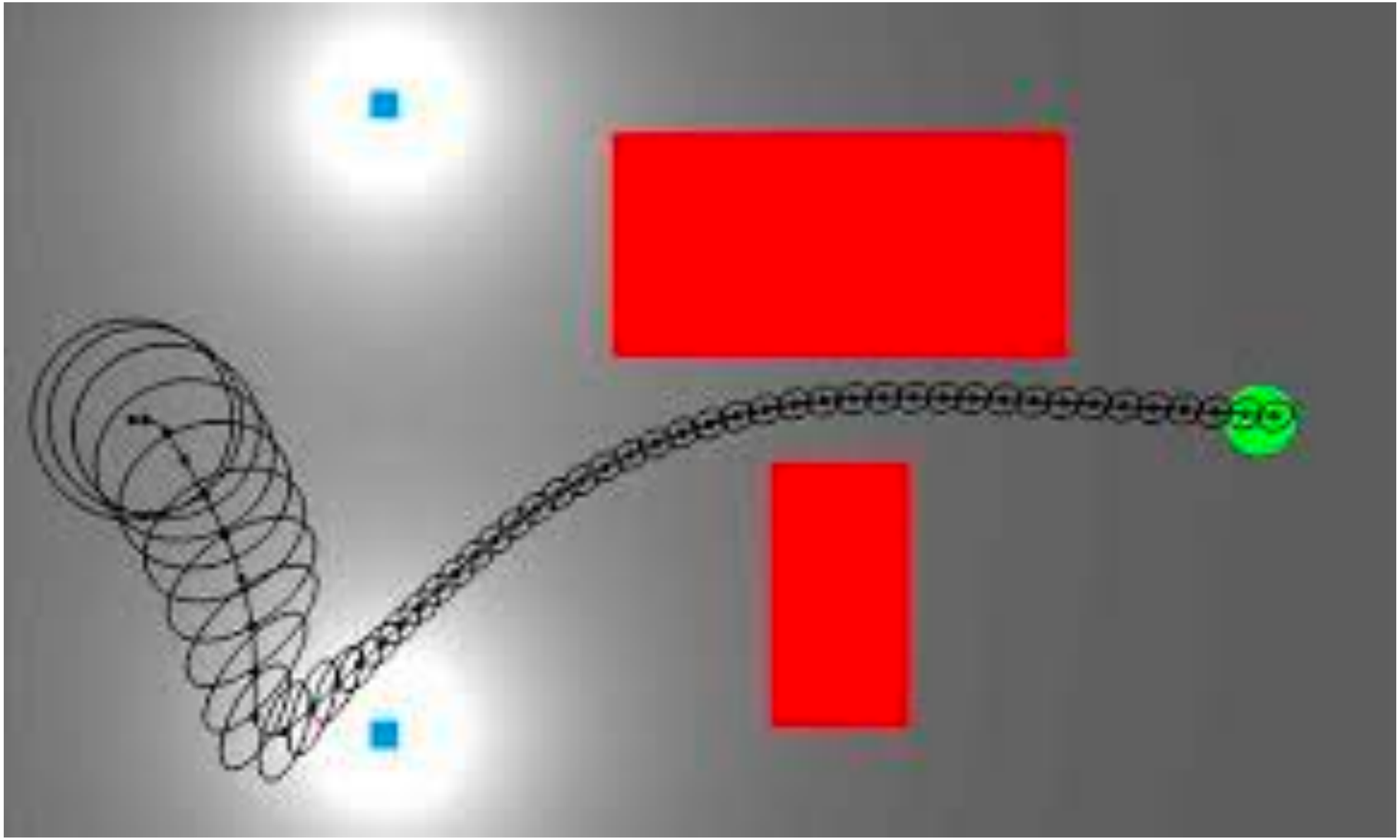


Uncertain about state

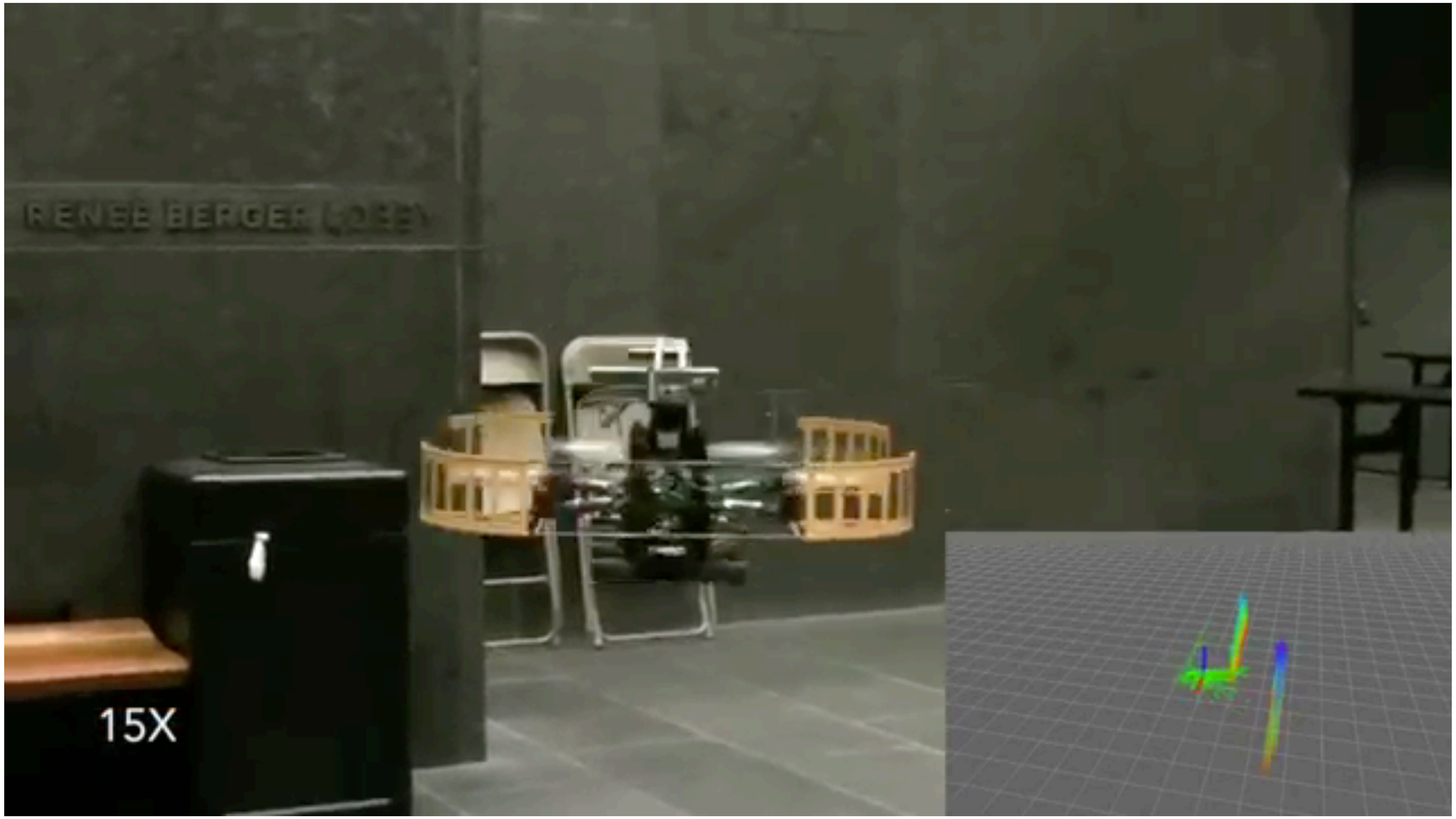


Uncertain about transitions

Uncertain about the robot pose

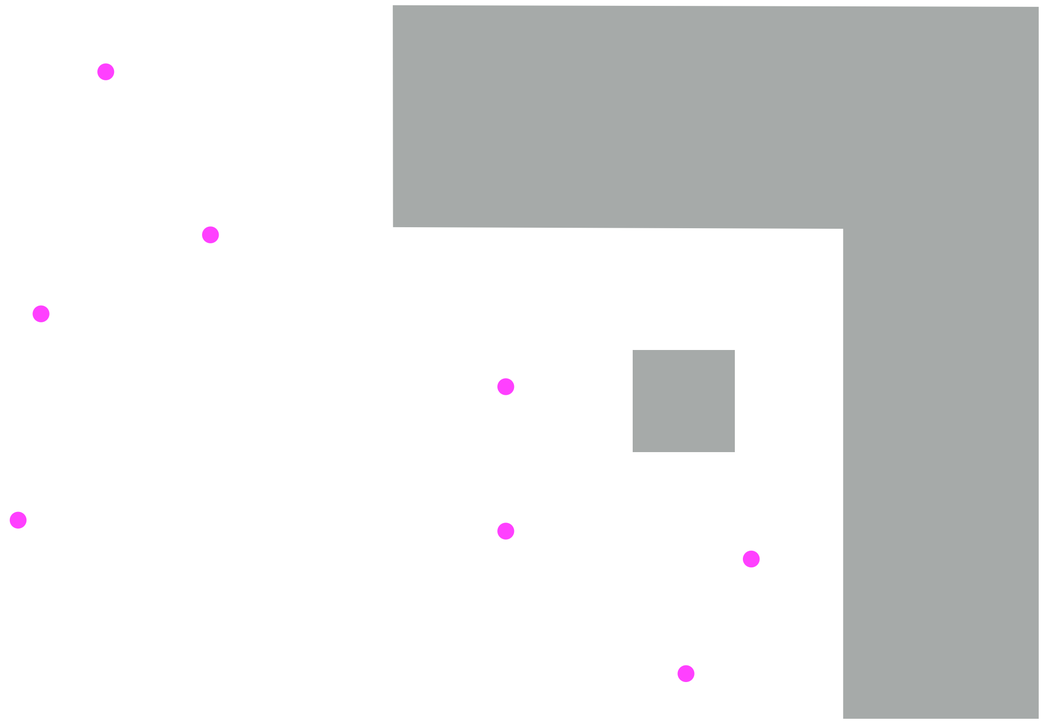


Uncertain about the world



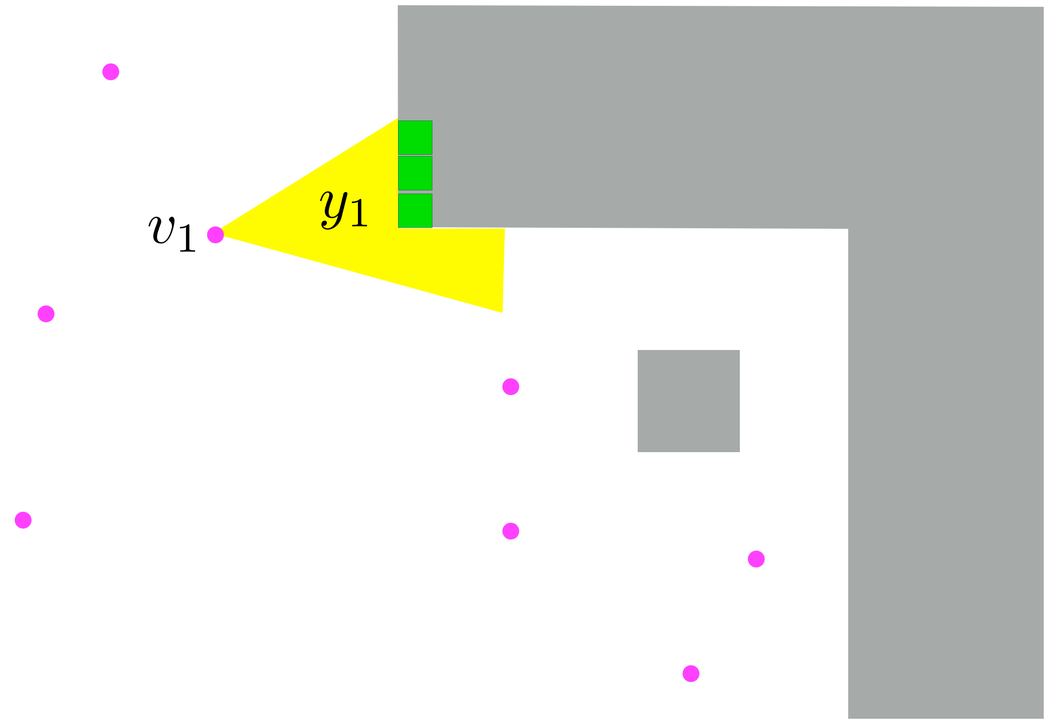
The coverage problem

$$\underset{\{v_1,\ldots,v_n\}}{\text{maximize}}\ F(v_1,\ldots,v_n)$$



The coverage problem

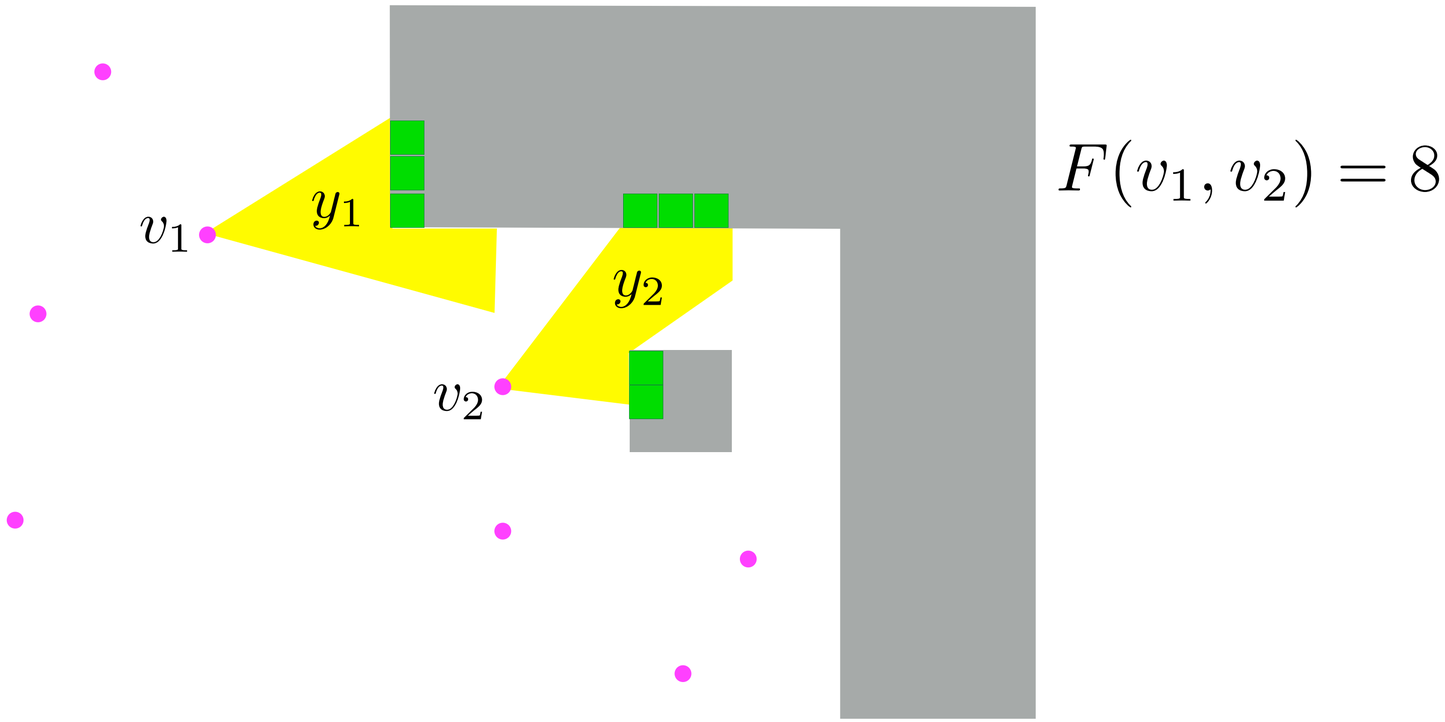
$$\underset{\{v_1,\ldots,v_n\}}{\text{maximize}}\ F(v_1,\ldots,v_n)$$



$F(v_1) = 3$

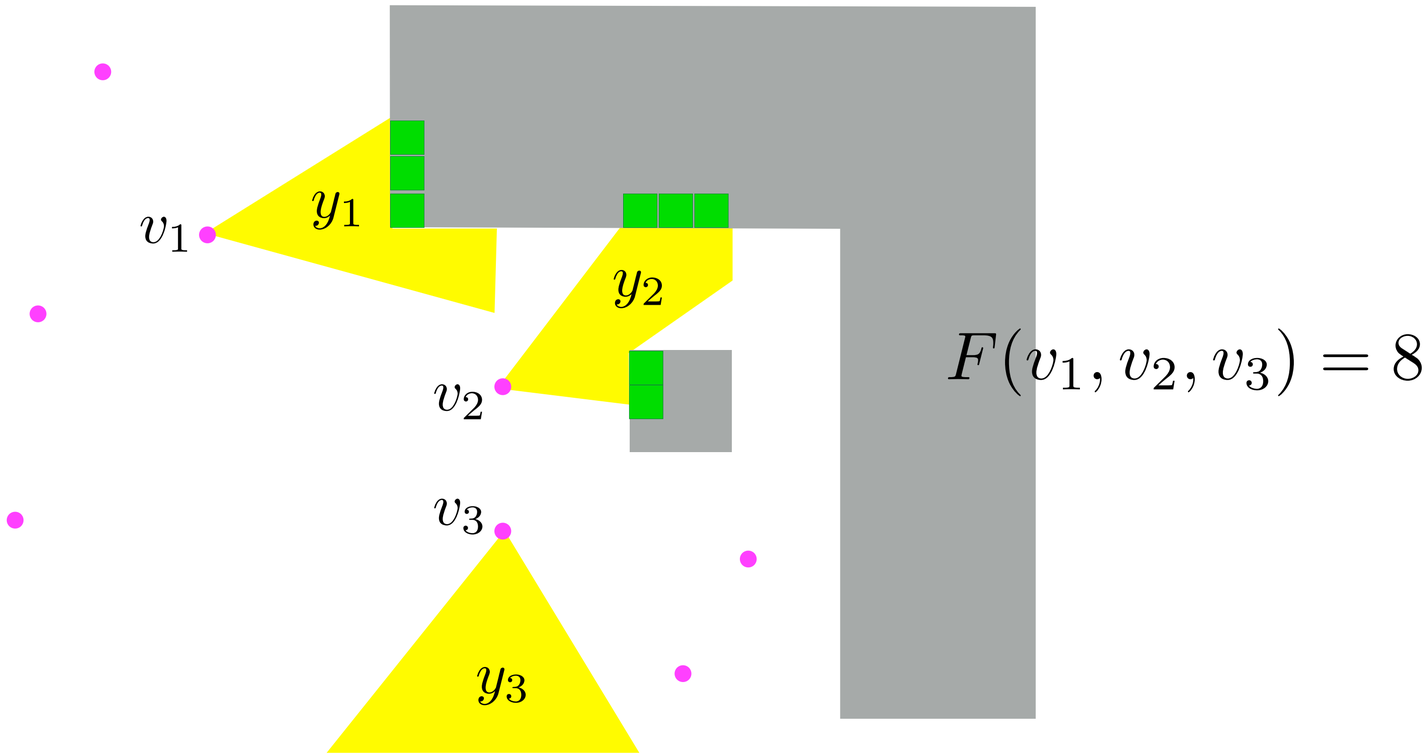
The coverage problem

$$\underset{\{v_1,\ldots,v_n\}}{\text{maximize}}\ F(v_1,\ldots,v_n)$$

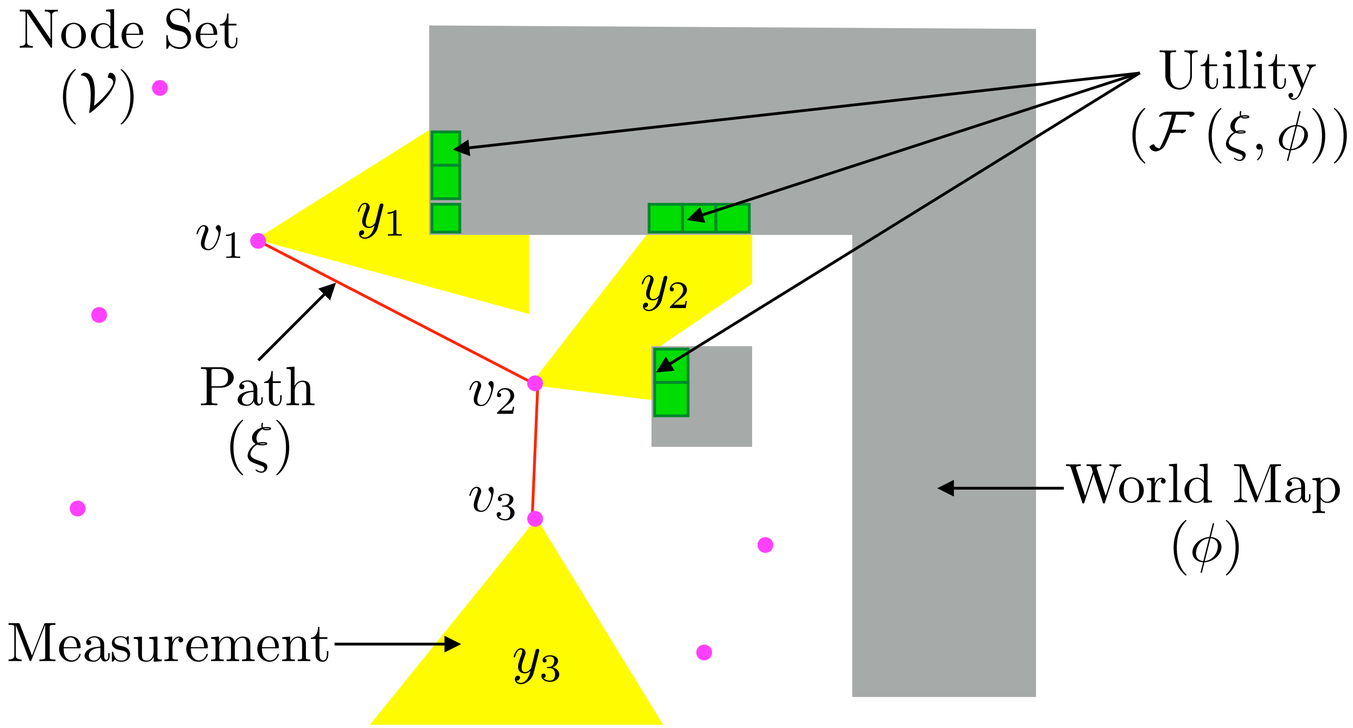


The coverage problem

$$\underset{\{v_1,\ldots,v_n\}}{\text{maximize}}\ F(v_1,\ldots,v_n)$$



The *budgeted* coverage problem



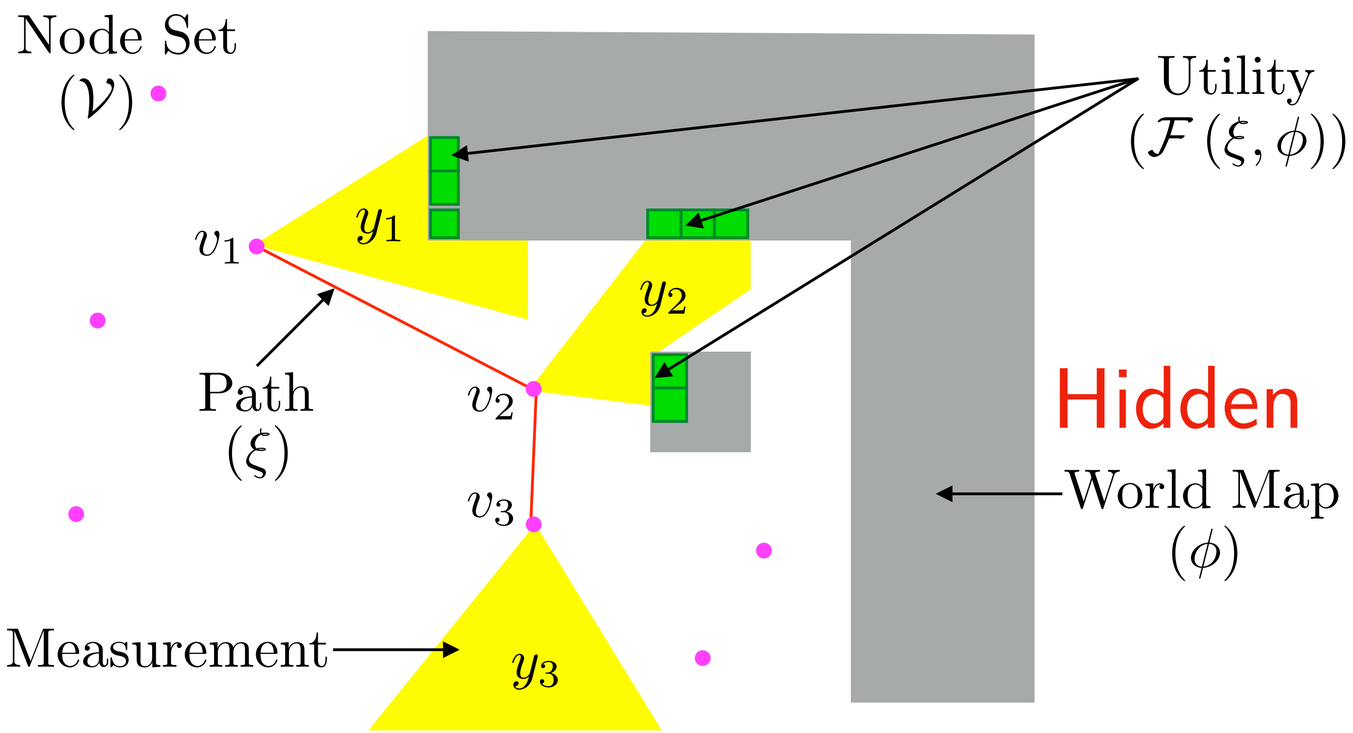
$$\underset{\xi \in \Xi}{\arg\max} \quad \mathcal{F}(\xi, \phi)$$

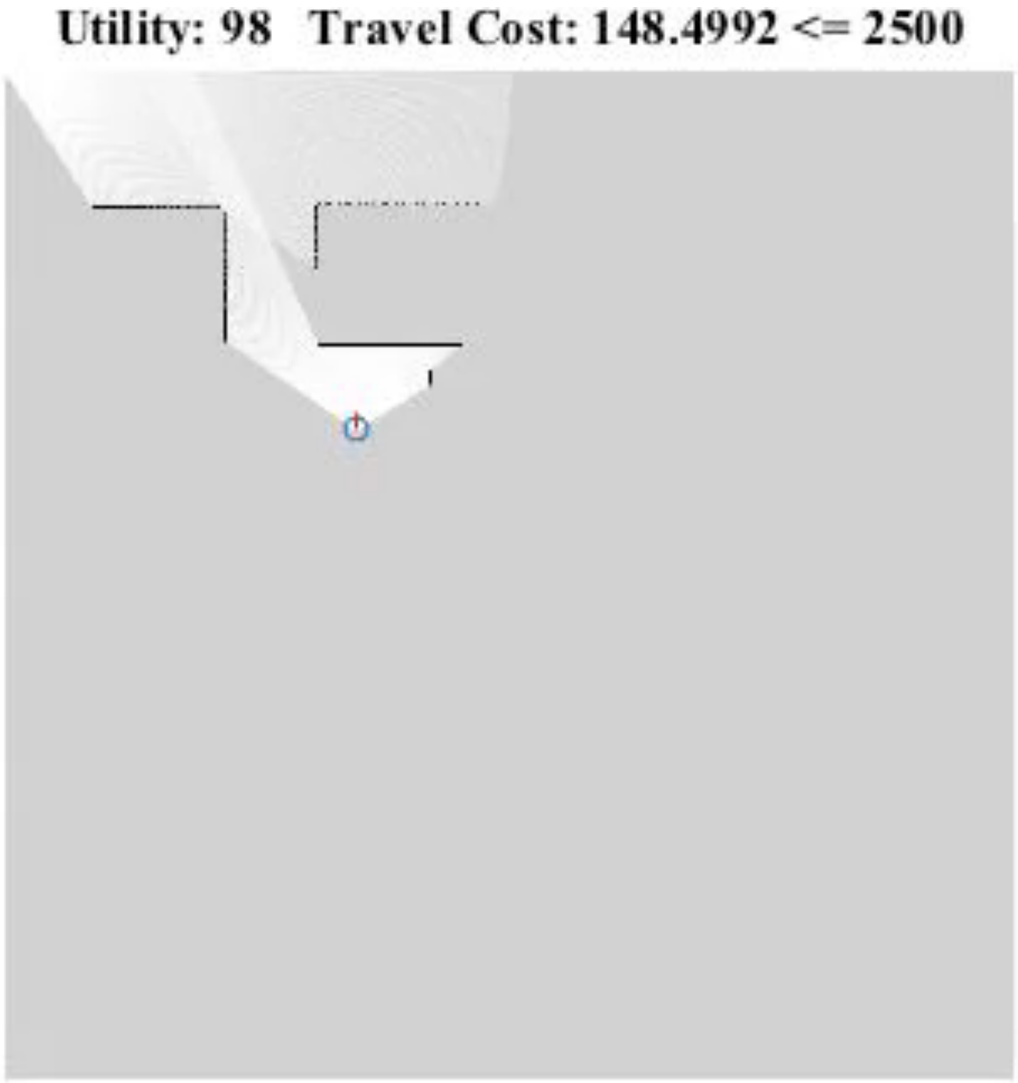
Cover as many cells

$$s.t. \quad \mathcal{T}(\xi, \phi) \leq B$$

Subject to travel cost!

The *budgeted* ~~coverage~~ info-gathering problem





$$\underset{\xi \in \Xi}{\arg\max} \quad \mathcal{F}(\xi, \phi)$$

$$s.t. \quad \mathcal{T}(\xi, \phi) \leq B$$

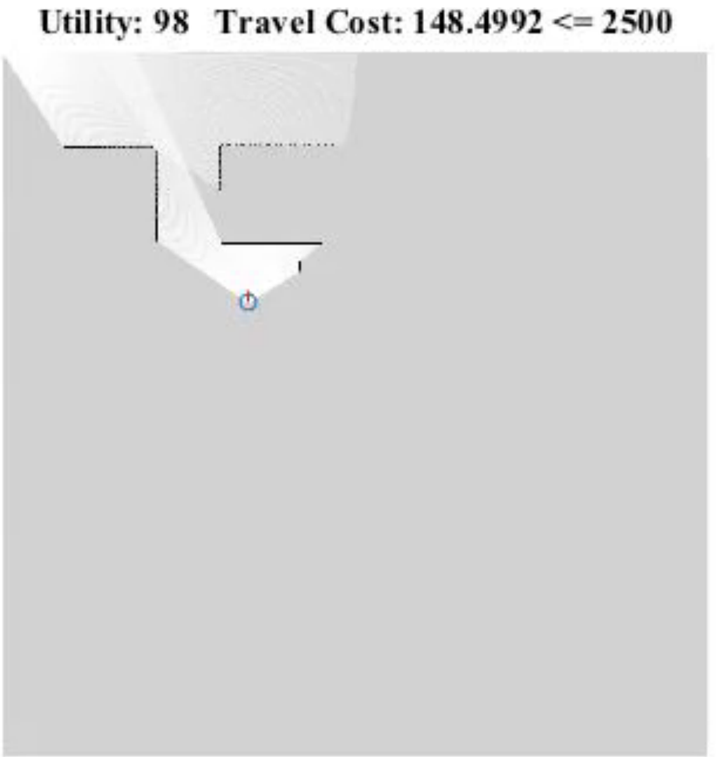
Activity!

Think-Pair-Share

Think (30 sec): Can you think of some heuristics for solving the budgeted information gathering problem?

Pair: Find a partner

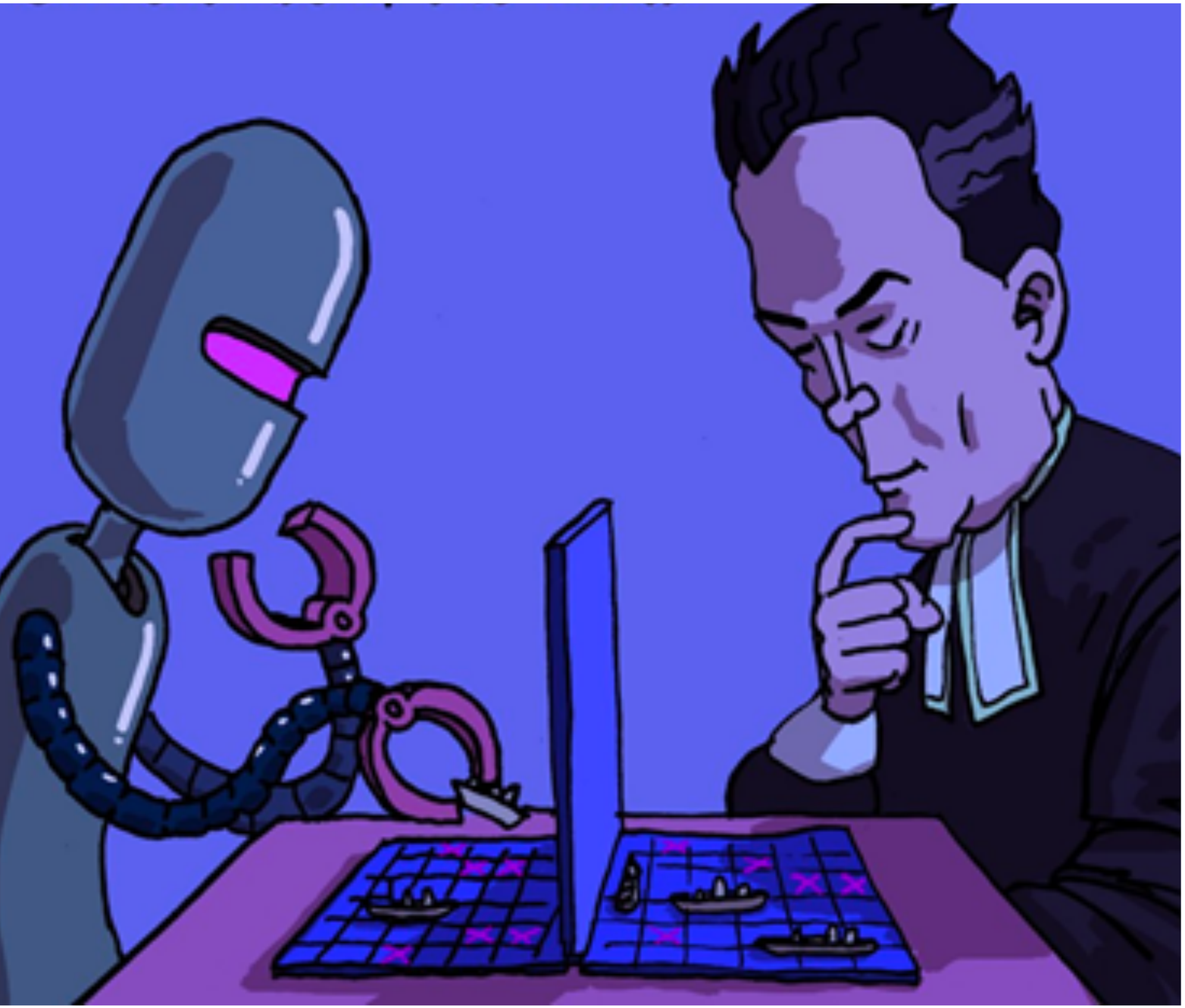
Share (45 sec): Partners exchange ideas



Belief Space Planning is NP-Hard
at best, undecidable at worst

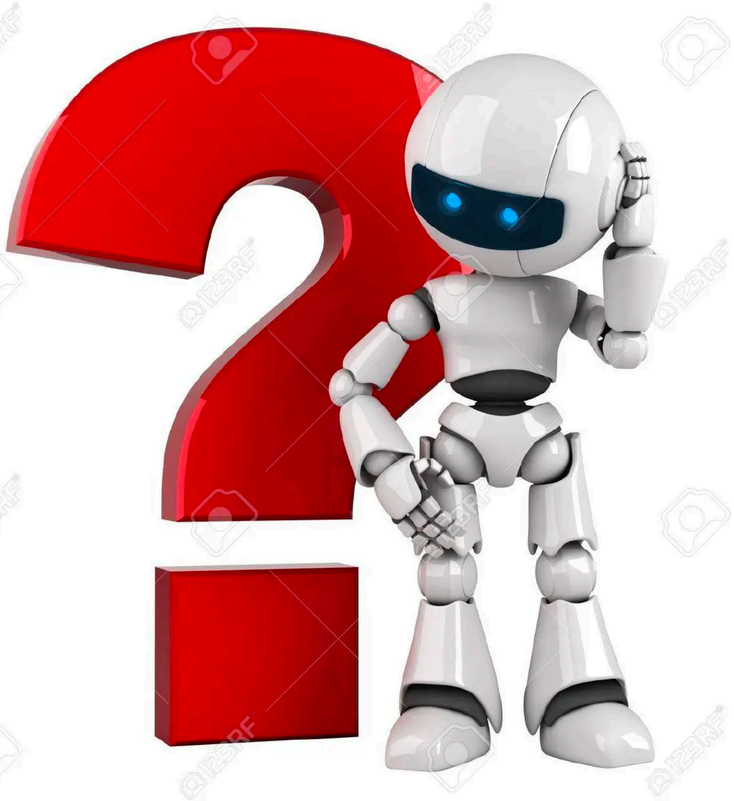
Need to relax our problem!

What if we wanted to explore as optimally as possible using prior information?



Information Gain

20 Questions



Let's say you have a set of hypotheses

$$\{\theta_1, \theta_2, \ldots, \theta_n\}$$

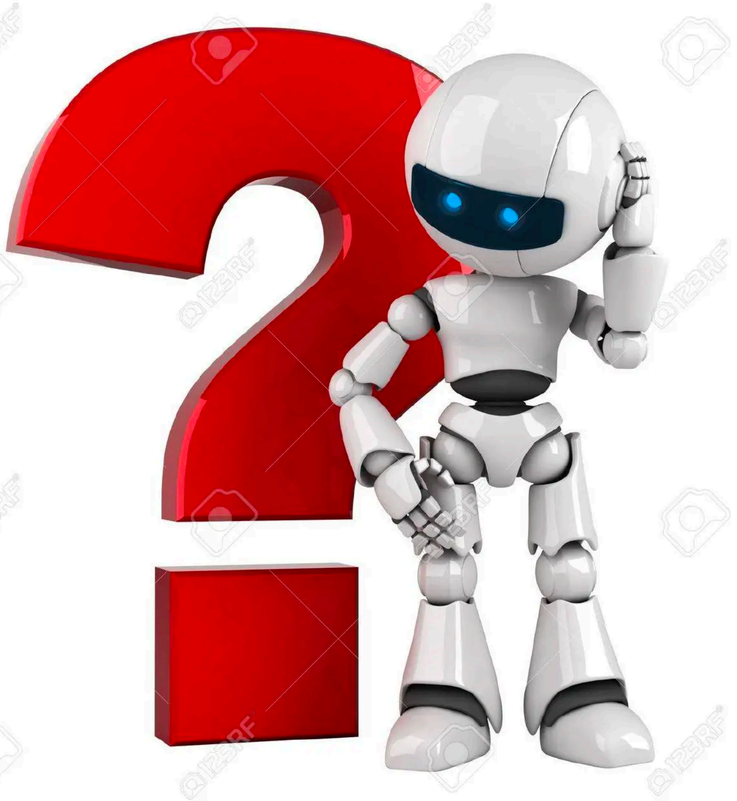
and a set of tests

$$\{t_1, t_2, \ldots, t_n\}$$

Given a prior over hypotheses $P(\theta)$

Find the minimal number of tests to identify hypothesis

20 Questions



Let's say you have a set of hypotheses

$$\{\theta_1, \theta_2, \ldots, \theta_n\}$$

and a set of tests

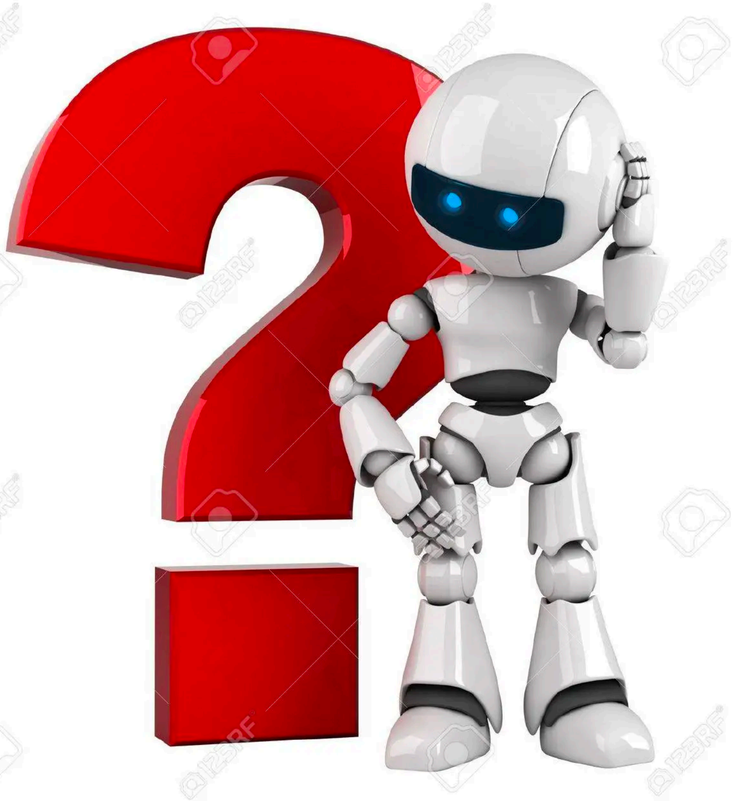
$$\{t_1, t_2, \ldots, t_n\}$$

NP-HARD

Given a prior over hypotheses $P(\theta)$

Find the minimal number of tests to identify hypothesis

A simple algorithm



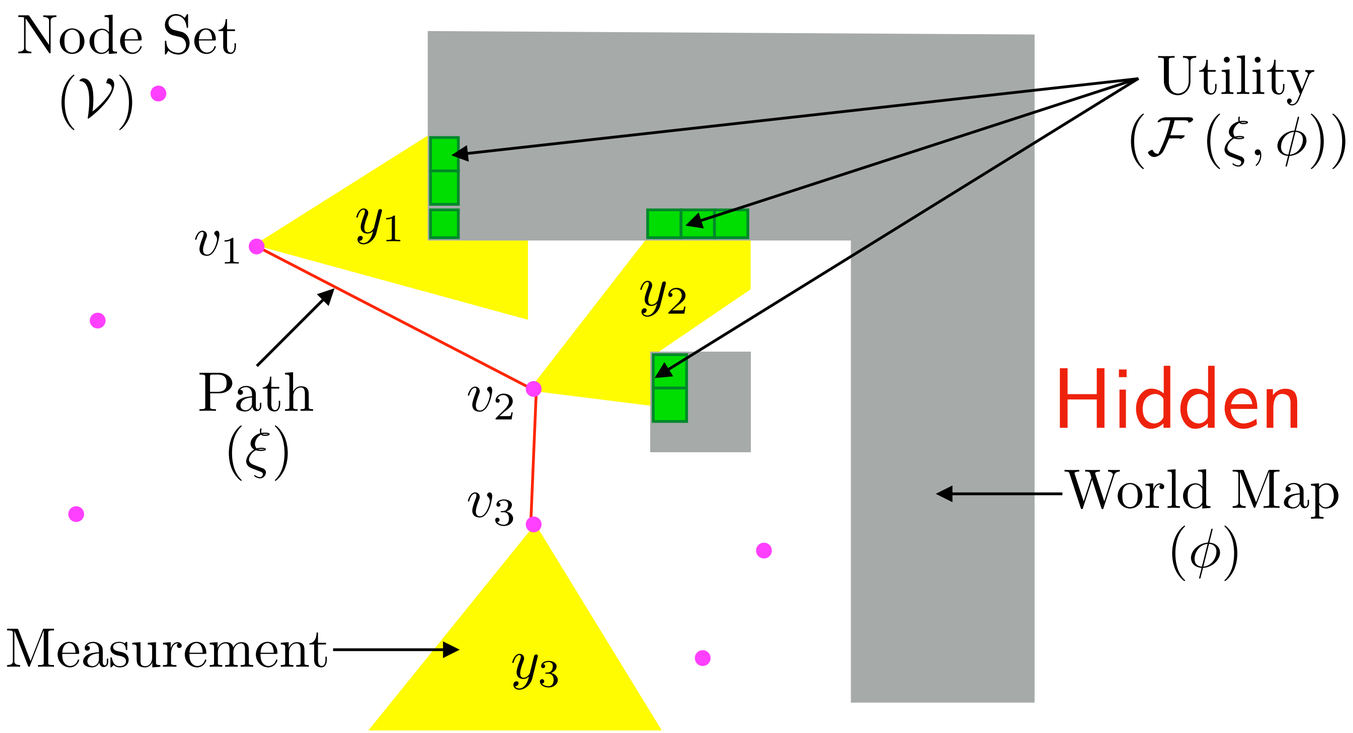
Greedily pick the test that maximizes information gain

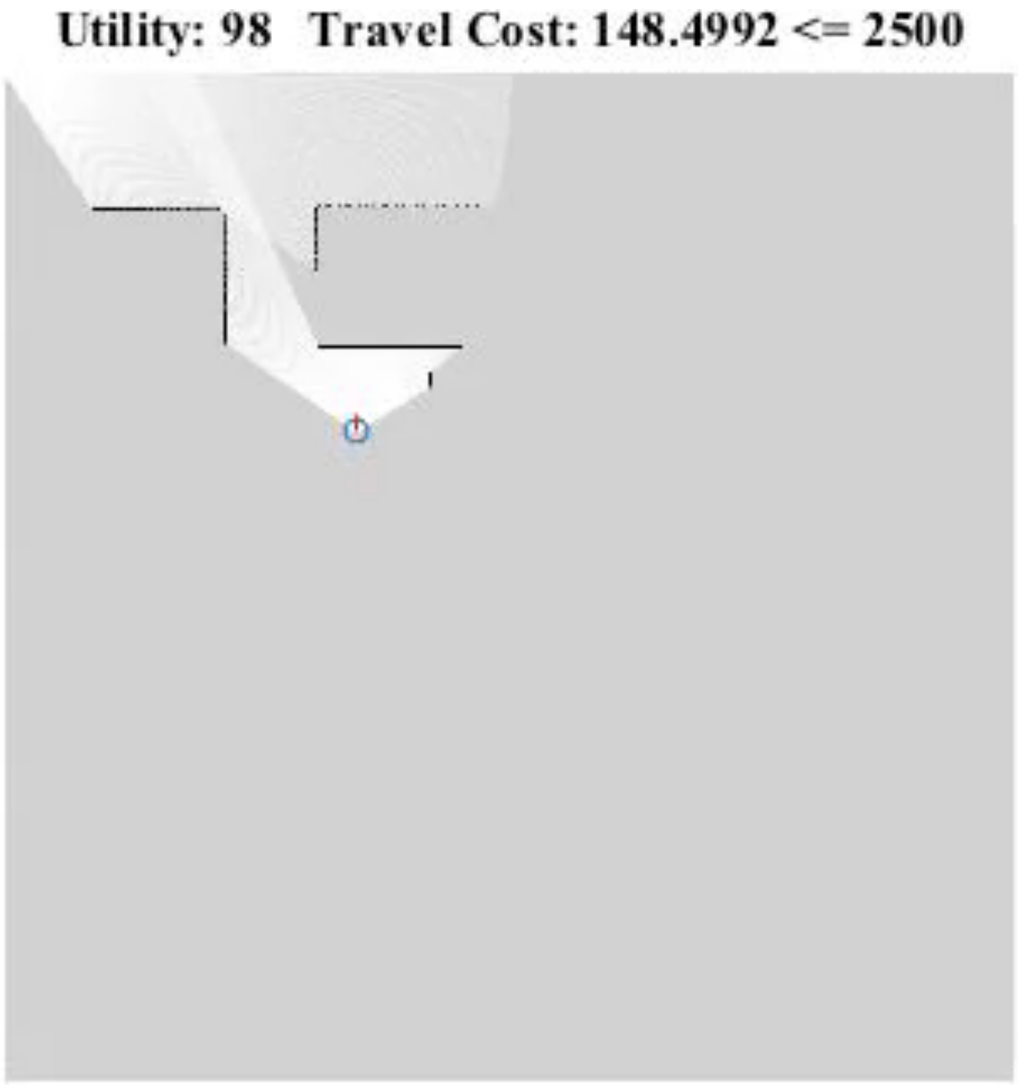
$$\max_t H(\theta) - \mathbb{E}_o H(\theta \mid t, o)$$

Entropy Posterior entropy

Entropy is adaptive sub modular => Greedy is near-optimal

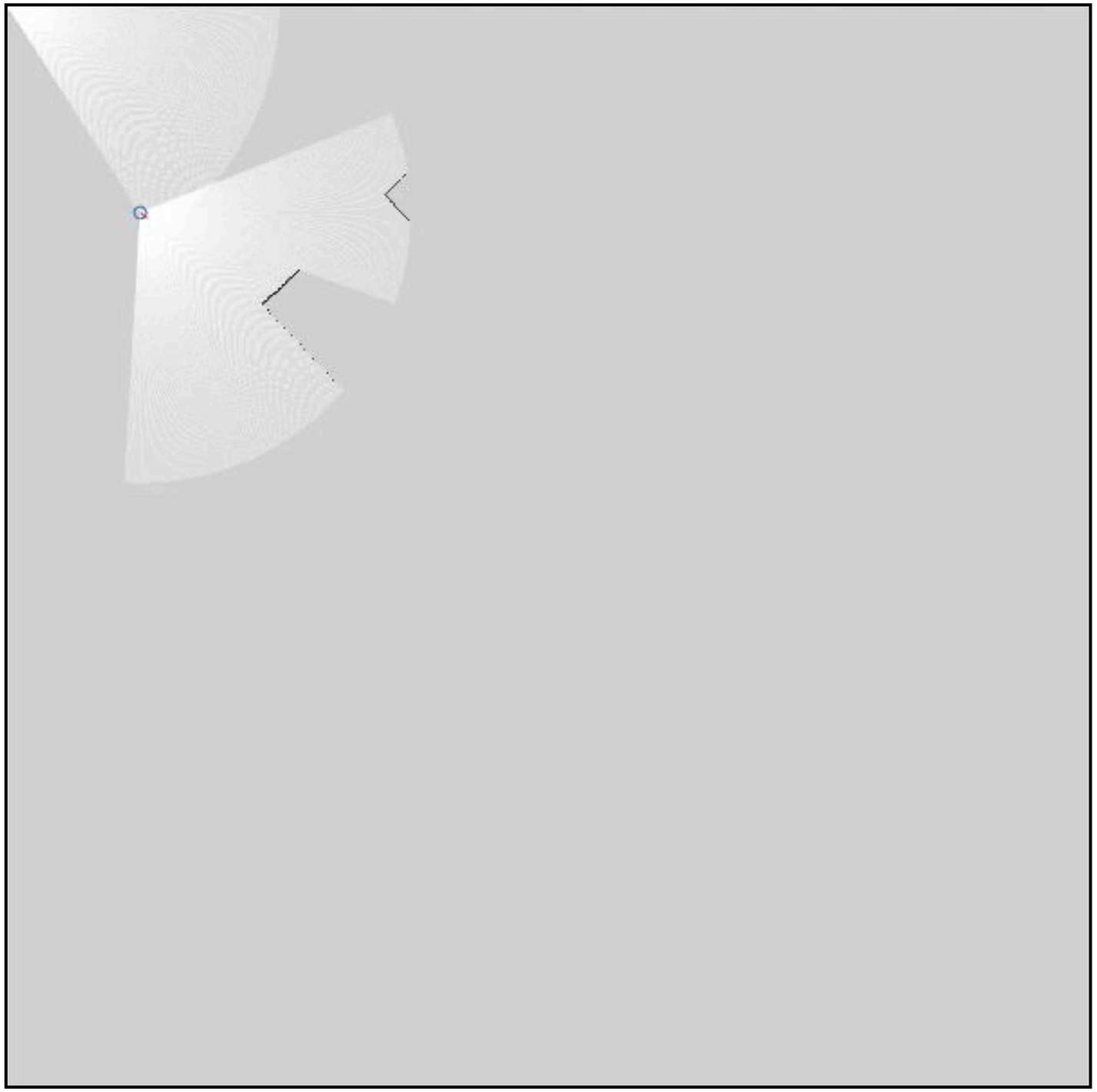
So does information gain work for this problem?





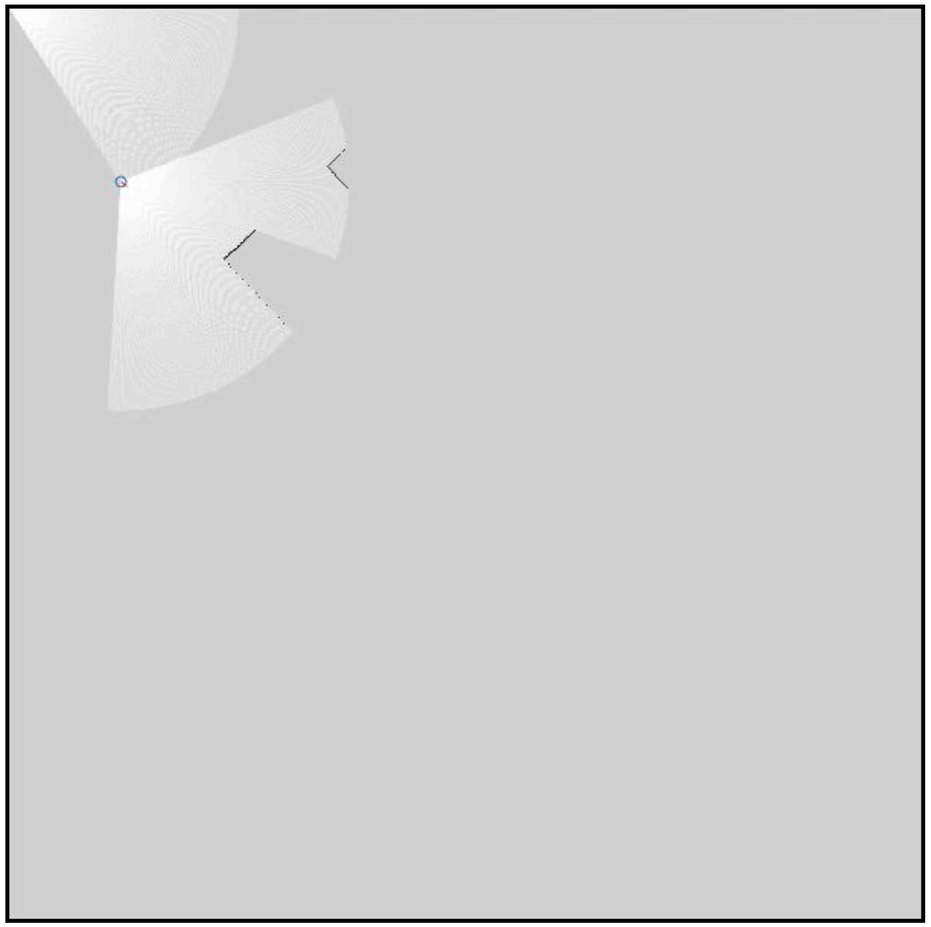
$$\underset{\xi \in \Xi}{\arg\max} \quad \mathcal{F}(\xi, \phi)$$

$$s.t. \quad \mathcal{T}(\xi, \phi) \leq B$$



Information Gain
overexplores!

Why does this happen?



1. Information Gain does not take into account travel cost!

2. Uniform Bernoulli prior may not be the best prior!

Can we find a better exploration / exploitation algorithm?

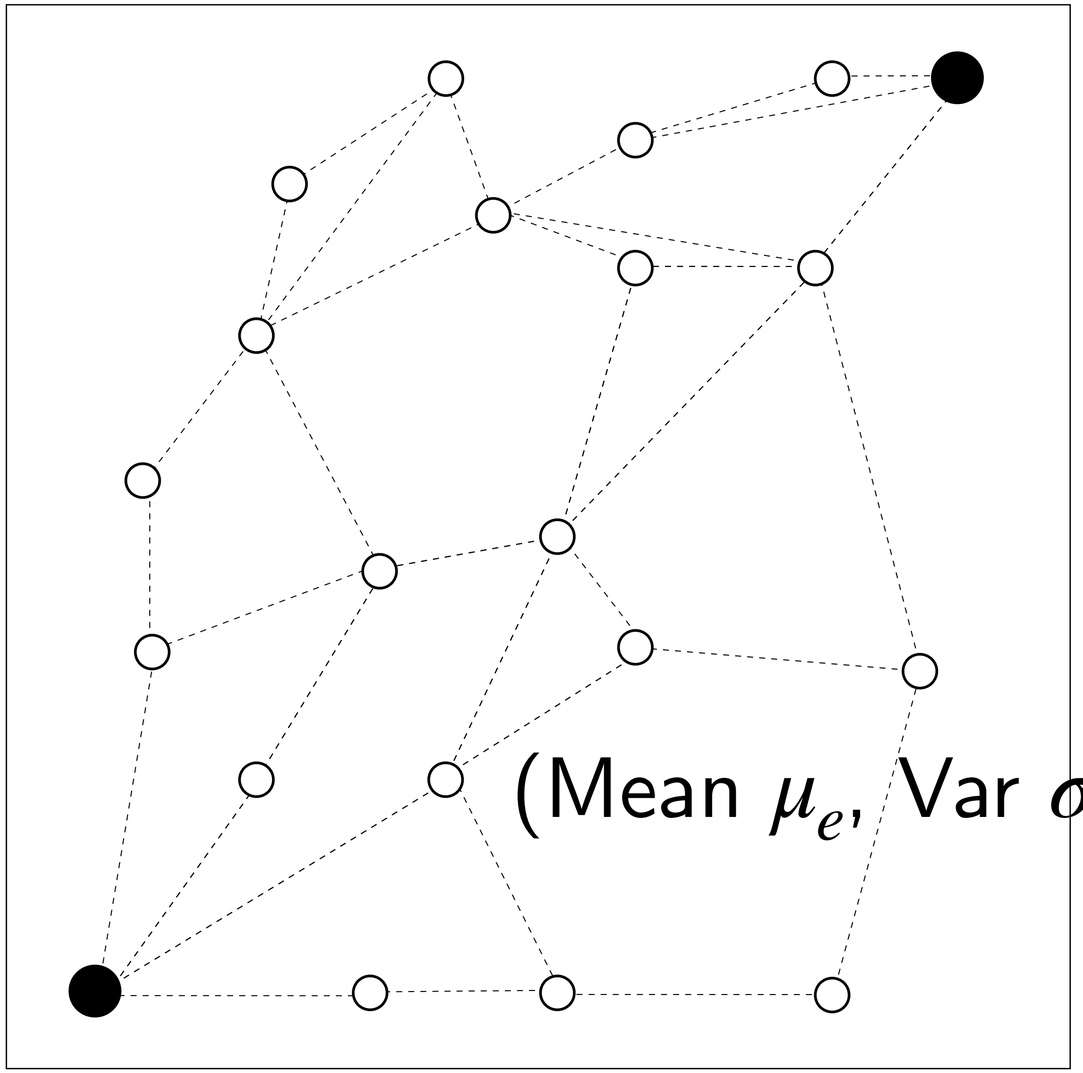
Posterior Sampling

The Online Shortest Path Problem

You just moved to Cornell and are traveling from office to home.

You would like to get home quickly but you are uncertain about travel times along each edge

Suppose we had a prior on travel time for each edge (Mean μ_e, Var σ_e)



What if ...

... we just sampled travel times from our prior and solved the shortest path?

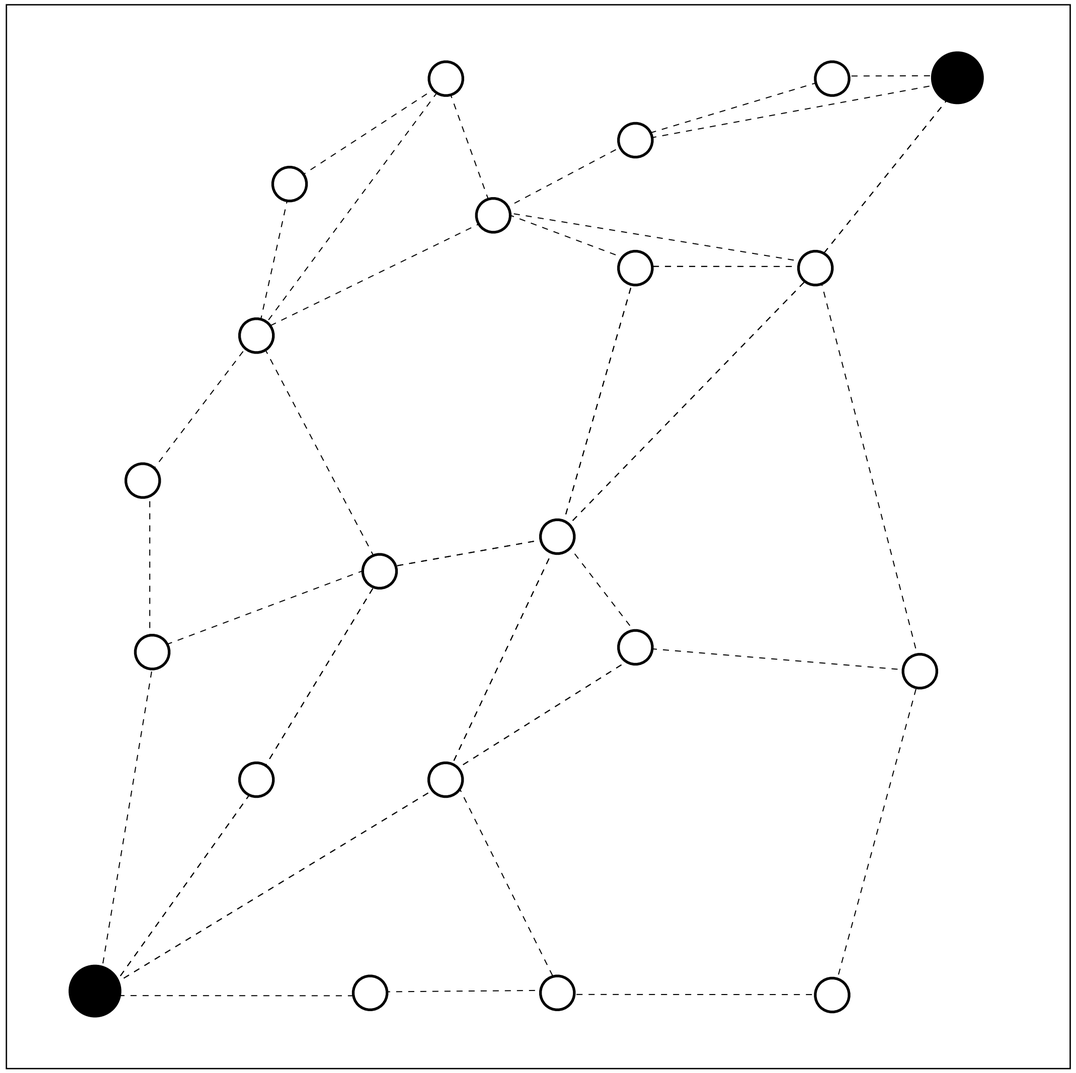
A suspiciously simple algorithm

Repeat forever:

Sample edge times from posterior

Compute shortest path

Travel along path, and update posterior



A suspiciously simple algorithm

Repeat forever:

Sample model from posterior

Compute optimal policy

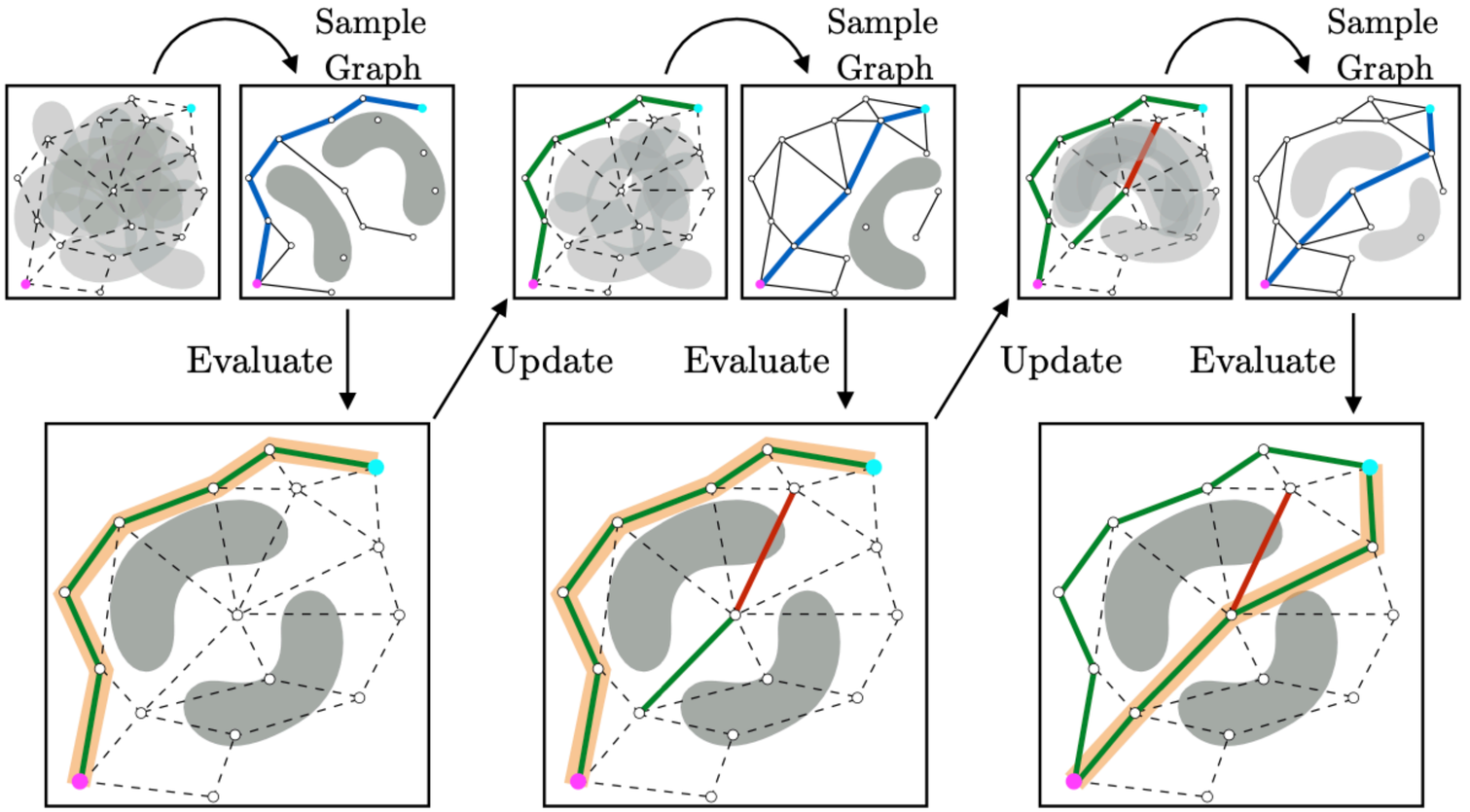
Execute policy, observe s,a,s',
Update model

A Tutorial on Thompson Sampling

Daniel J. Russo[1], Benjamin Van Roy[2], Abbas Kazerouni[2], Ian Osband[3] and Zheng Wen[4]

[1] *Columbia University*
[2] *Stanford University*
[3] *Google DeepMind*
[4] *Adobe Research*

Posterior Sampling for Motion Planning



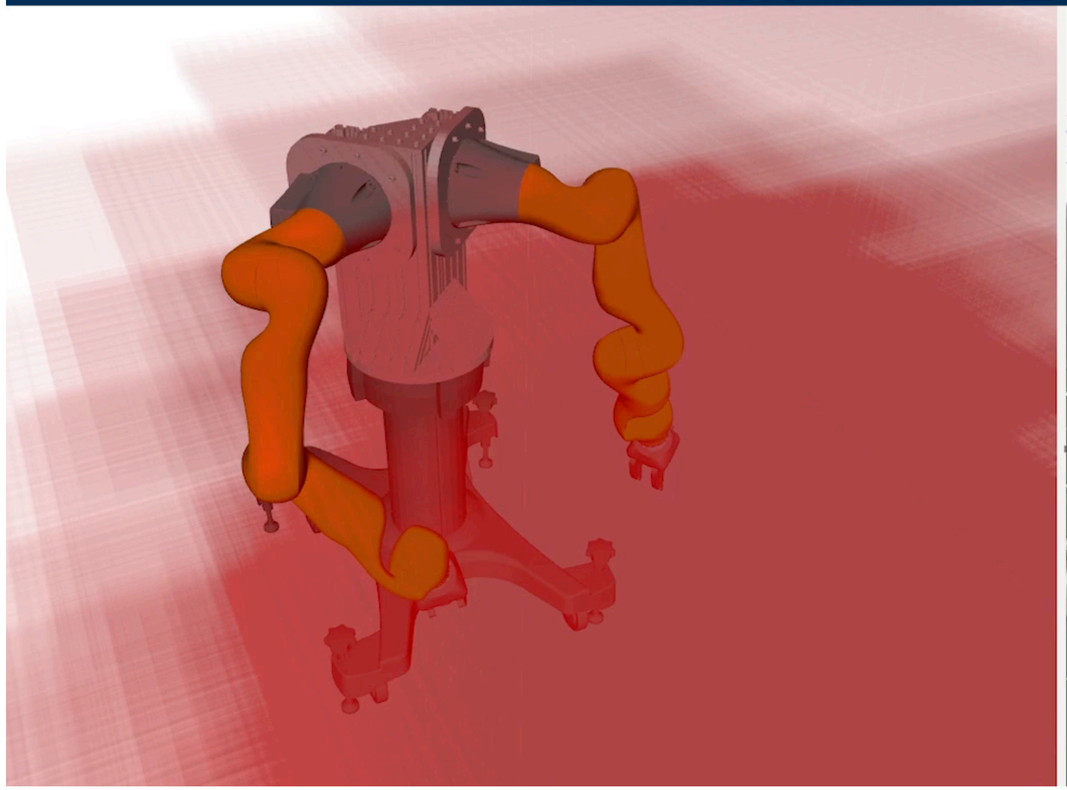
Posterior Sampling for Anytime Motion Planning on Graphs with Expensive-to-Evaluate Edges

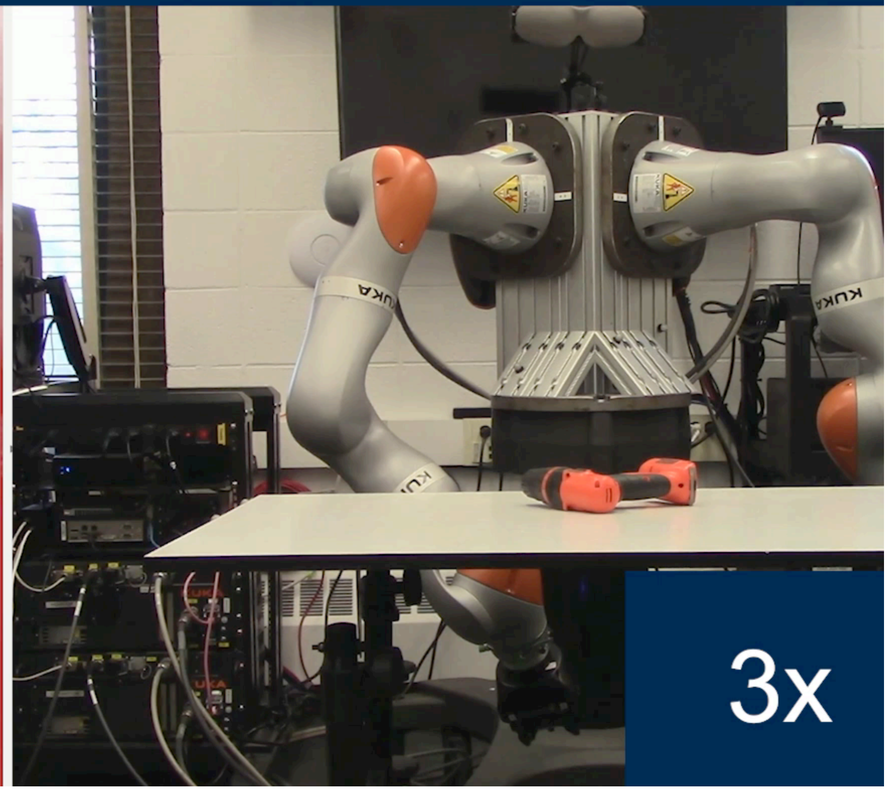
Brian Hou, Sanjiban Choudhury, Gilwoo Lee, Aditya Mandalika, and Siddhartha S. Srinivasa

Real Robot Problems!

ROBOTICS

The Blindfolded Robot:
Bayesian Planning with Contact Feedback
[ISRR'19]



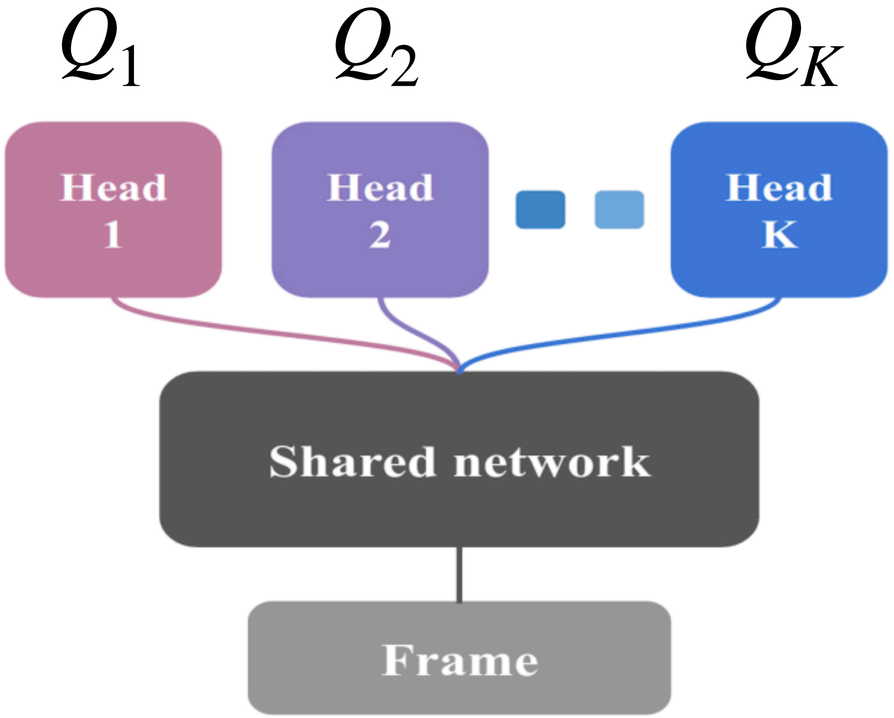


Posterior Sampling for Reinforcement Learning

1. sample Q-function Q from $p(Q)$
2. act according to Q for one episode
3. update $p(Q)$

Deep Exploration via Bootstrapped DQN

Ian Osband[1,2], Charles Blundell[2], Alexander Pritzel[2], Benjamin Van Roy[1]
[1]Stanford University, [2]Google DeepMind
{iosband, cblundell, apritzel}@google.com, bvr@stanford.edu



Bootstrapped Q Network

Posterior Sampling for Reinforcement Learning

1. sample Q-function Q from $p(Q)$
2. act according to Q for one episode
3. update $p(Q)$

Atari

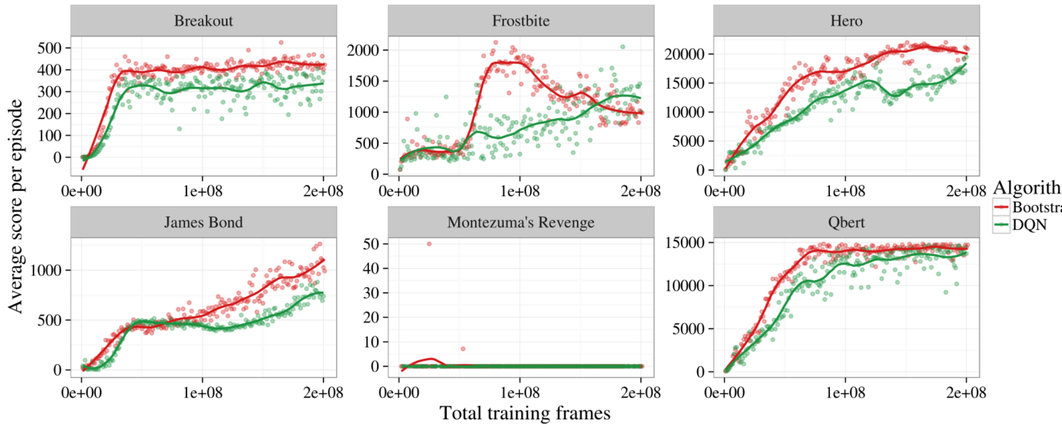


Figure 6: Bootstrapped DQN drives more efficient exploration.

Why does work better than taking random actions?

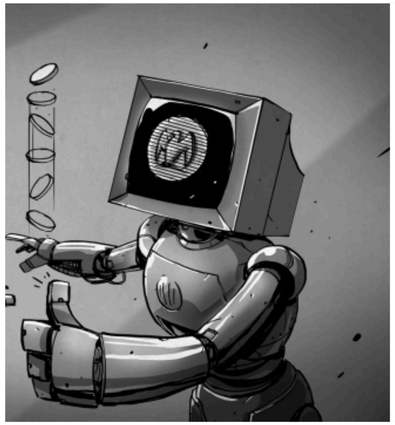
tl;dr

Belief Space Planning is NP-Hard at best, undecidable at worst

Need to relax our problem!

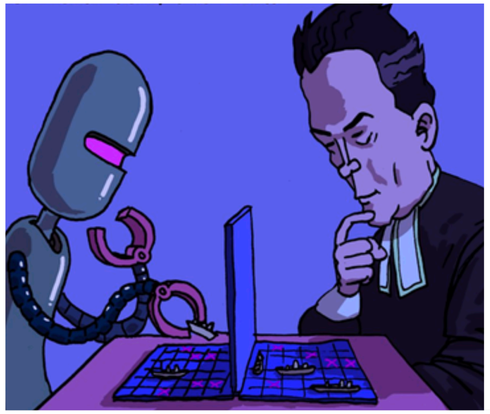
Optimism in the Face of Uncertainty (OFU)

17



Posterior Sampling

45



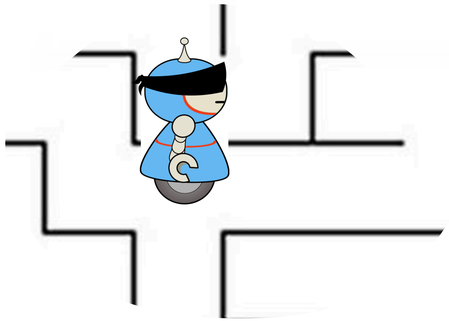
Information Gain

What is my prior is intractable to represent and sample from?

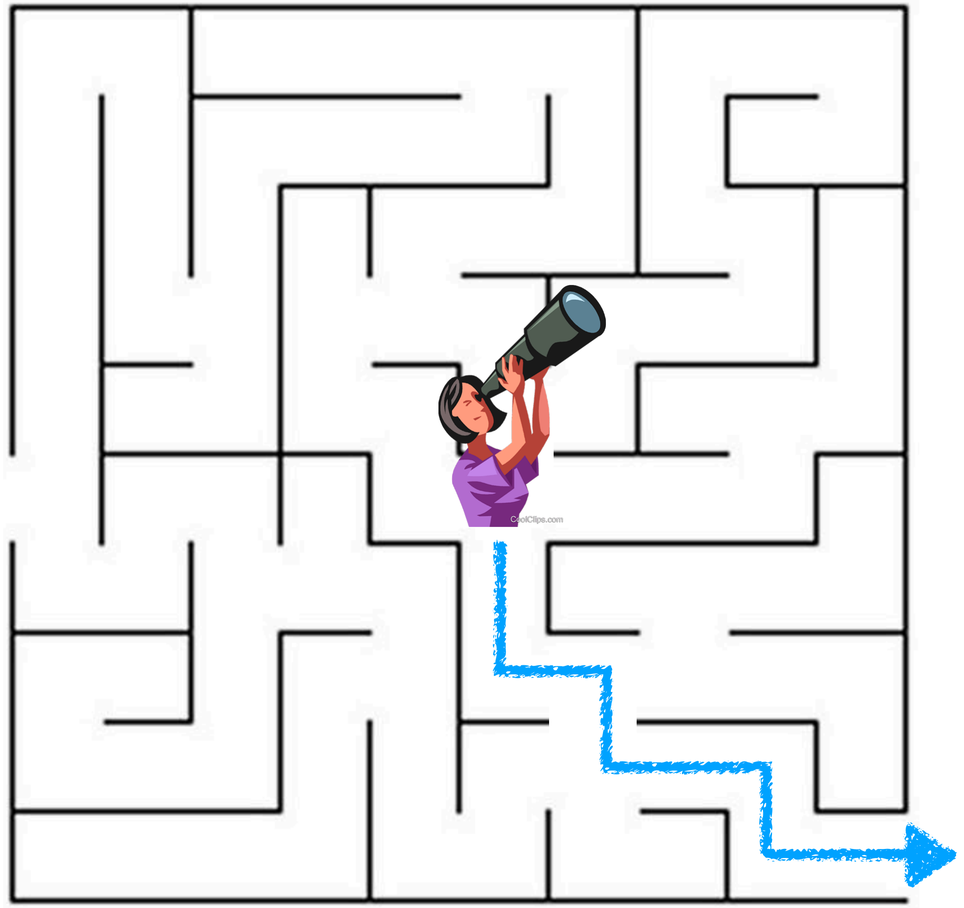
What if ...

... we trained a learner to imitate a clairvoyant oracle?

Imitating Experts with Privileged Information



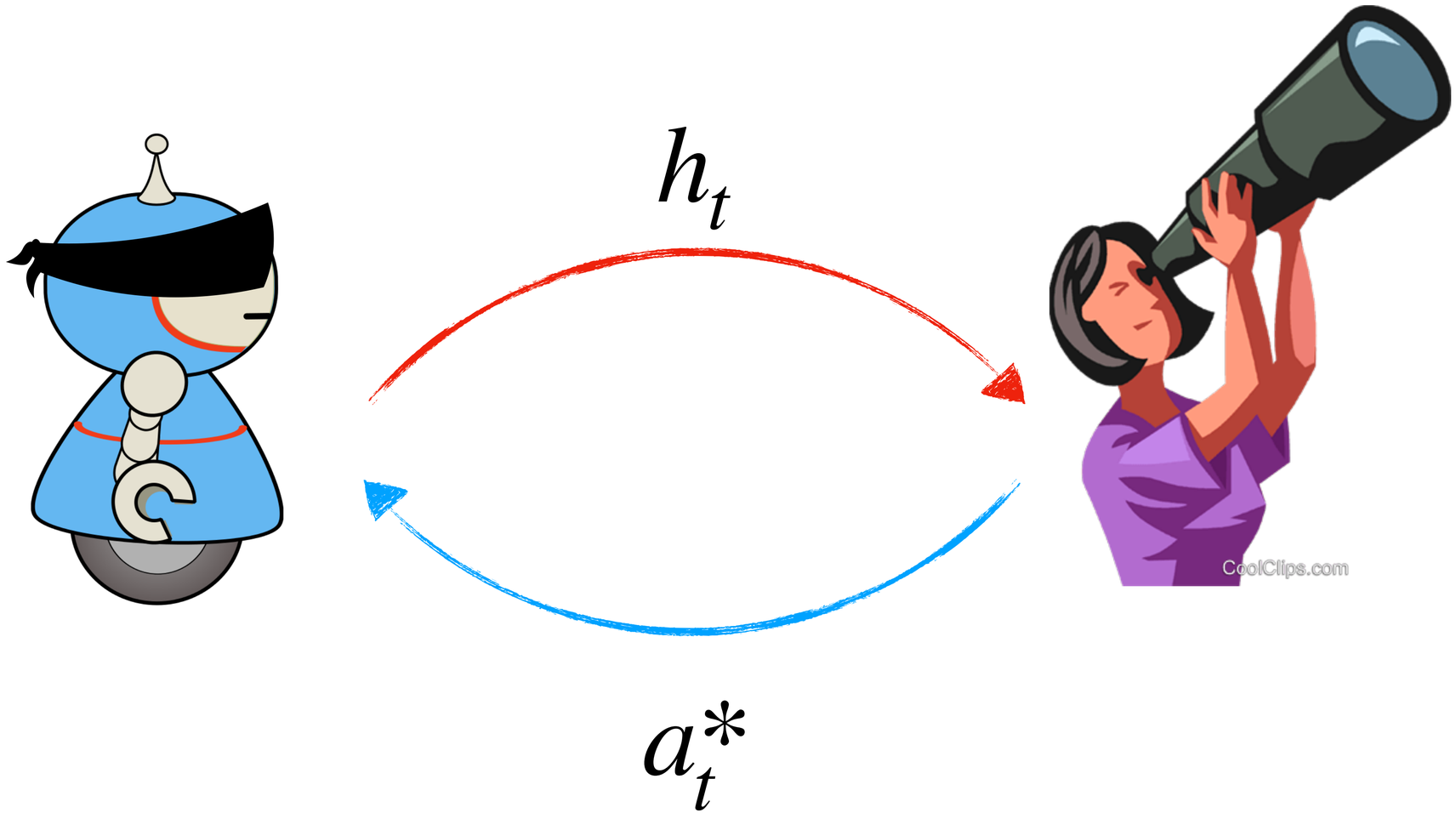
Imitate



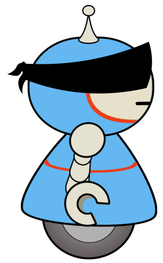
Learner
w/ limited sensing

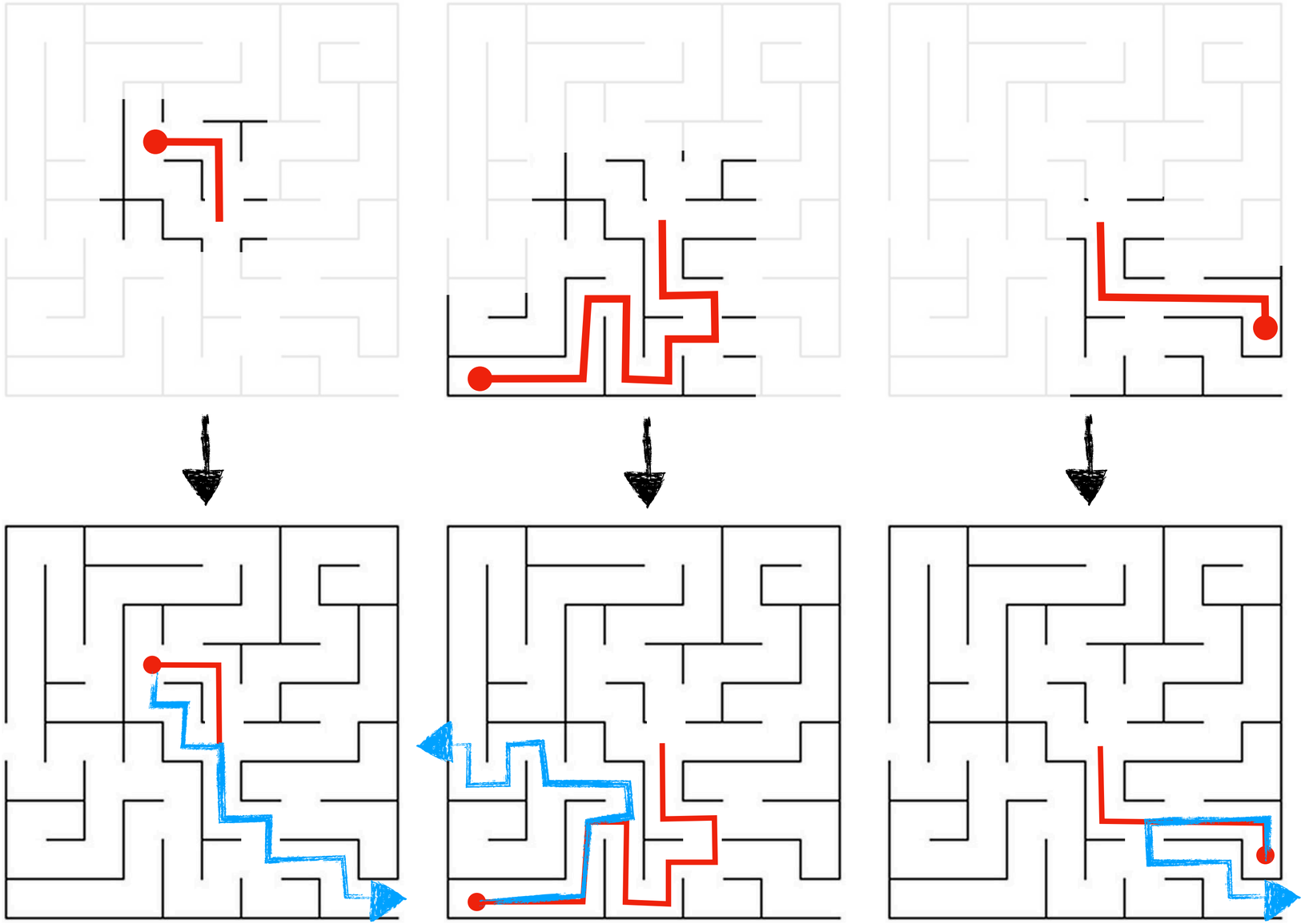
Expert
can see further

Solution: Interactively query expert



Solution: Interactively query expert



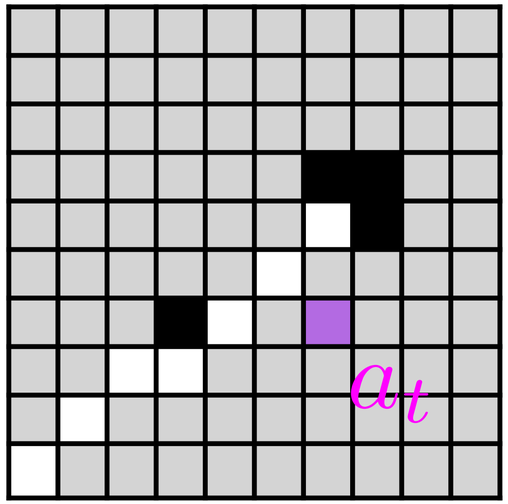


e.g DAGGER

1. Roll out learner
2. Query Expert
3. Aggregate Data

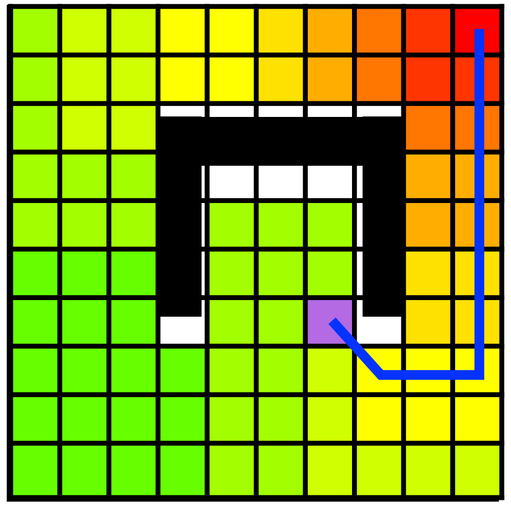
and repeat!

Privileged Information: Motion Planning

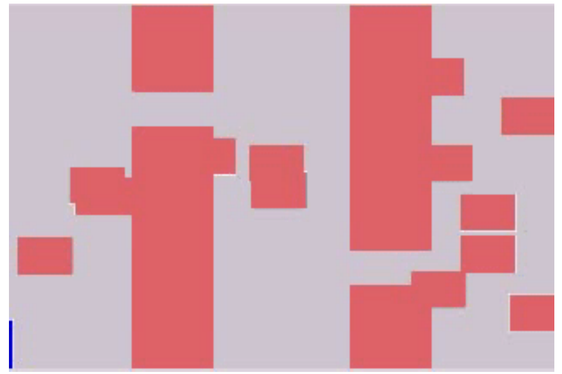


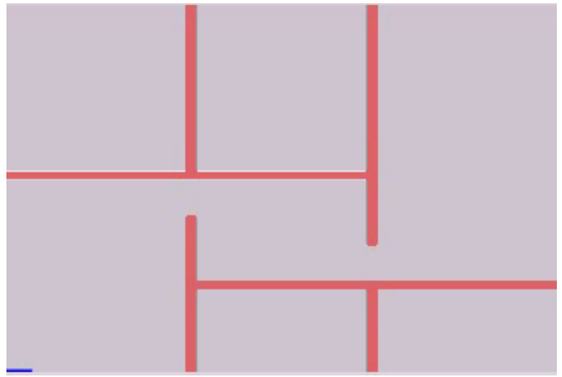
Learned
Search Heuristic

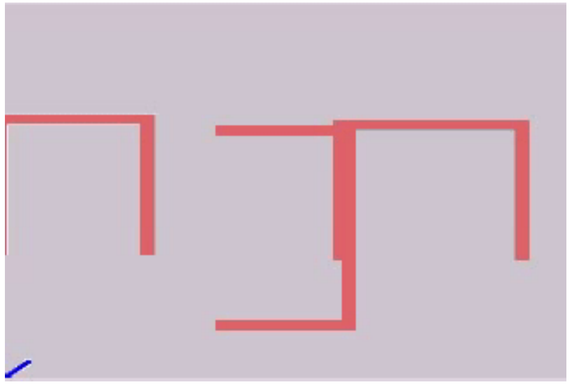
Imitate



Optimal
Value Function





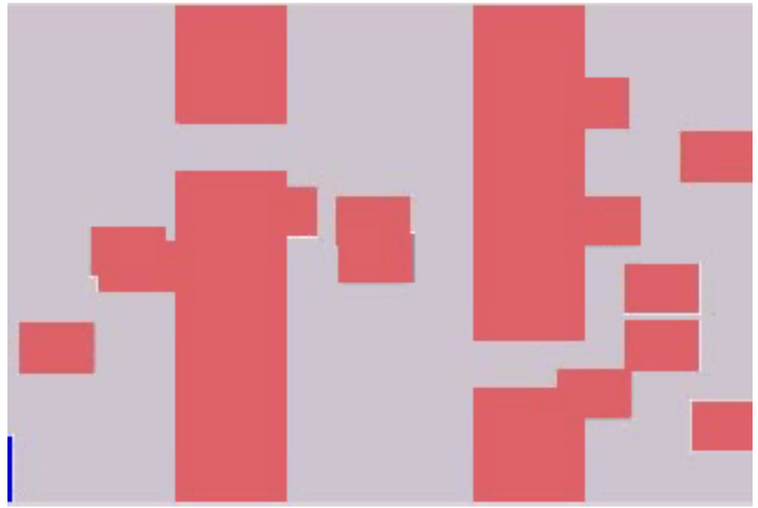


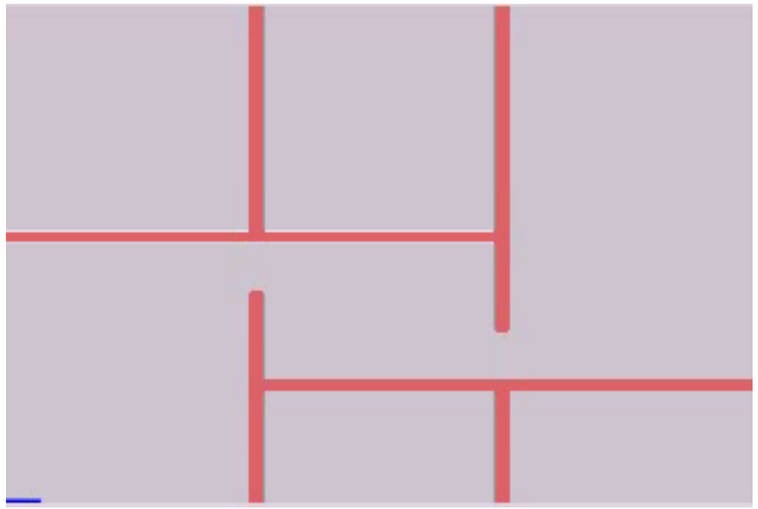
[Choudhury et al. ‘2018]

[Choudhury 2018]

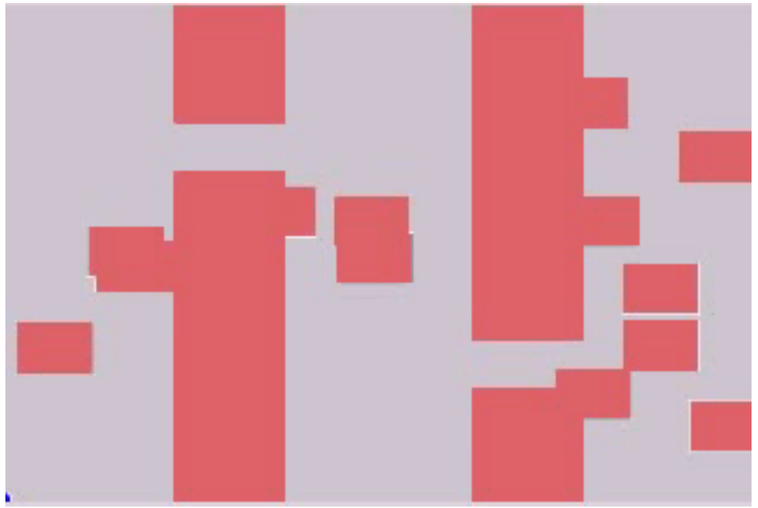
Example: Training search heuristics

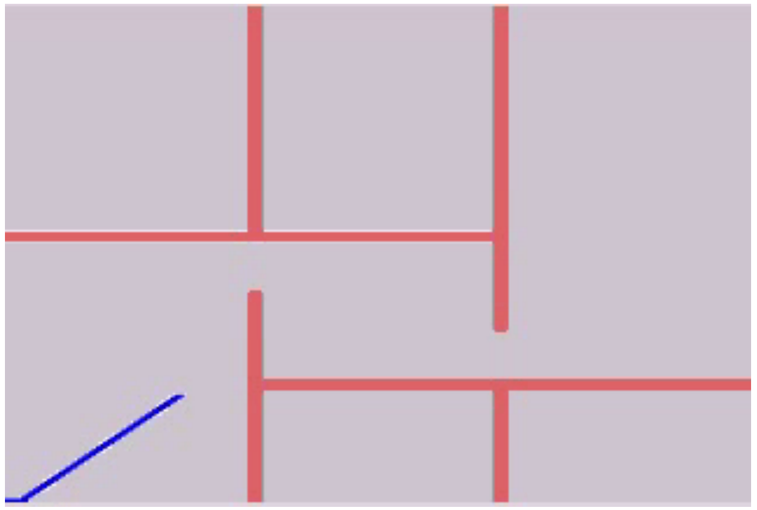
On-policy (Aggrevate)





Behavior Cloning





Why / When does this work?

Proved that this approximates Hindsight Optimization / QMDP

Fails when you need to explicitly explore (i.e. asymptotic realizability not hold)