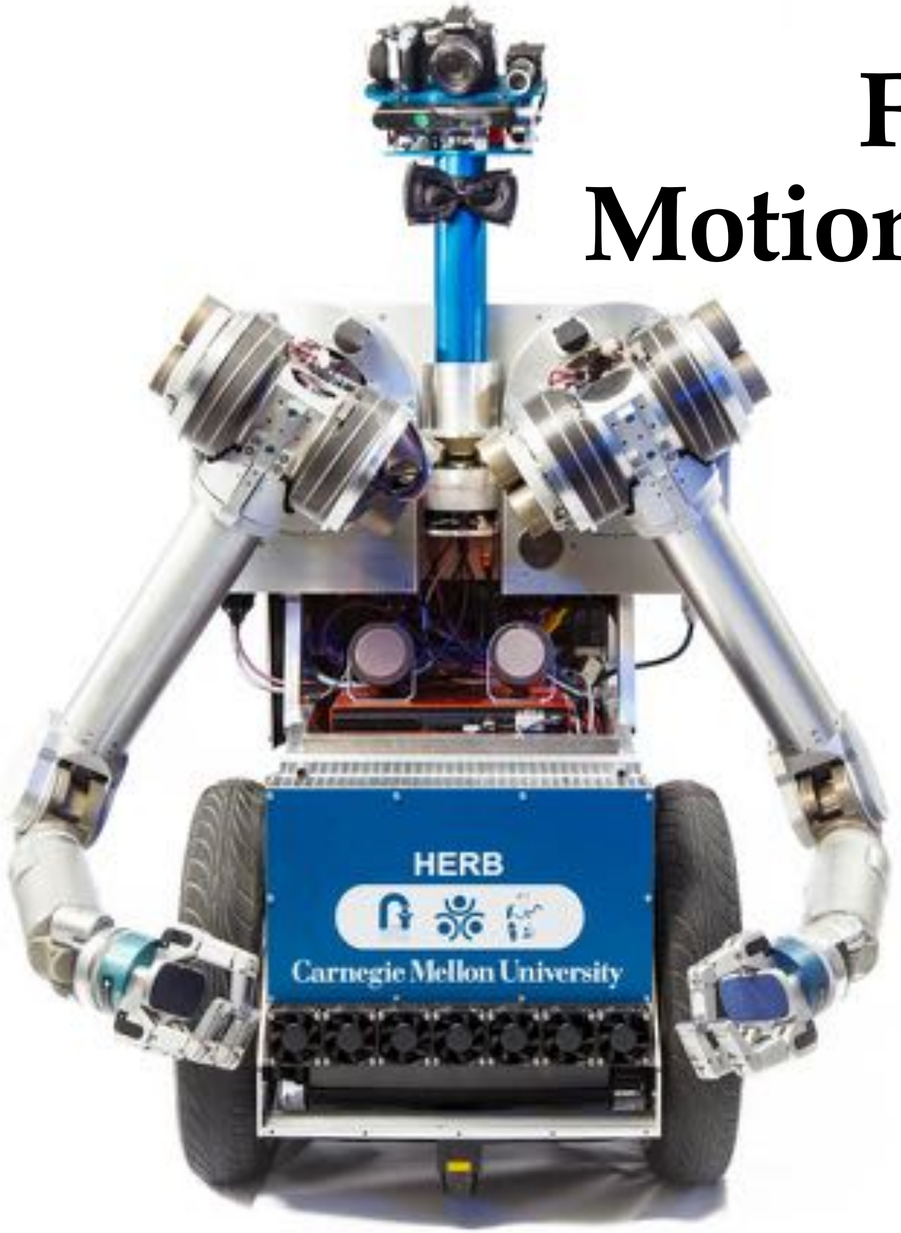


Formalizing Connections Between Motion Planning and Machine Learning

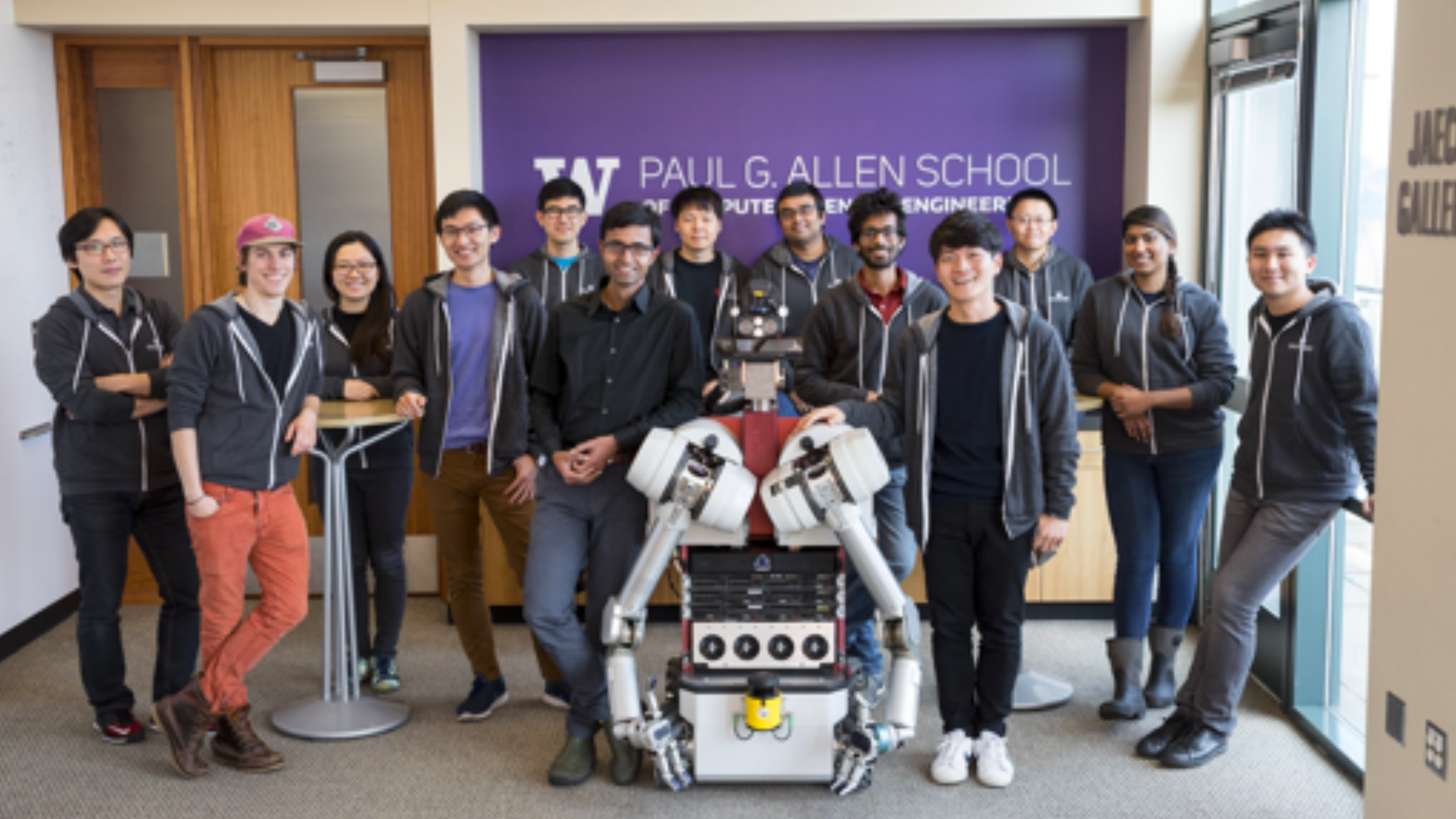
Siddhartha Srinivasa
Boeing Endowed Professor
University of Washington

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

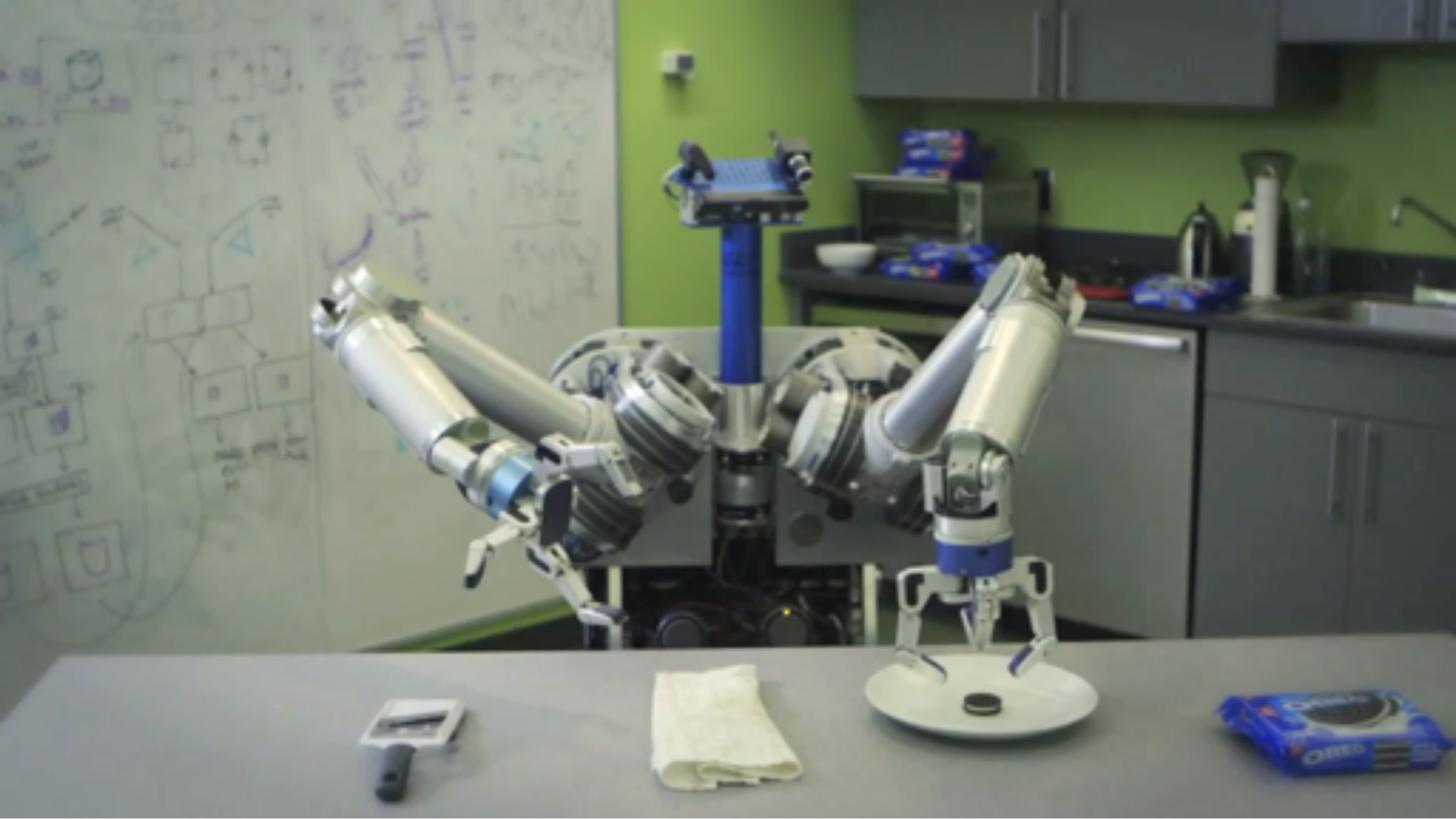
Problems I Want You to Solve So I can Retire

Siddhartha Srinivasa
Retired Boeing Endowed Professor
University of Washington

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING


PAUL G. ALLEN SCHOOL

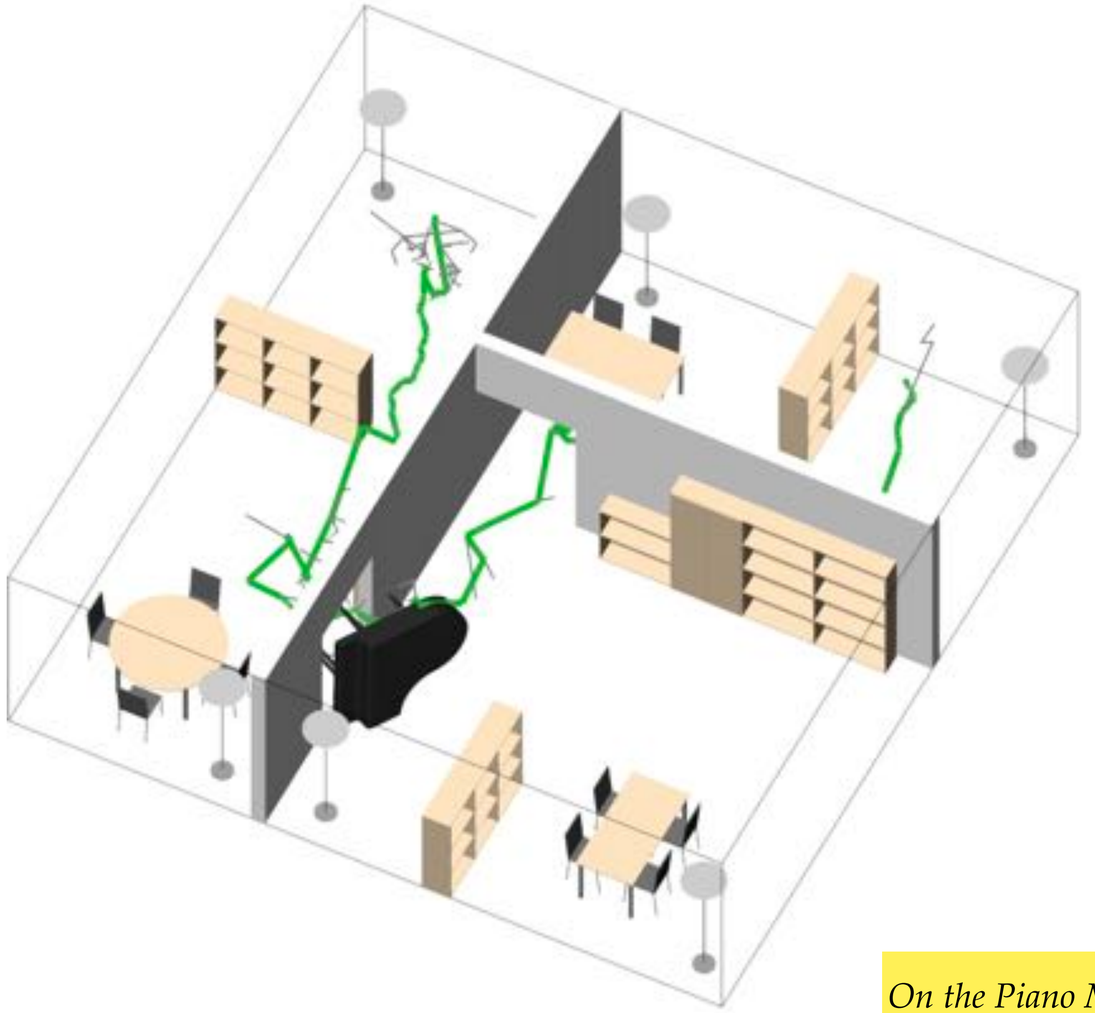
Motion Planning


OREO

Motion Planning is a technology

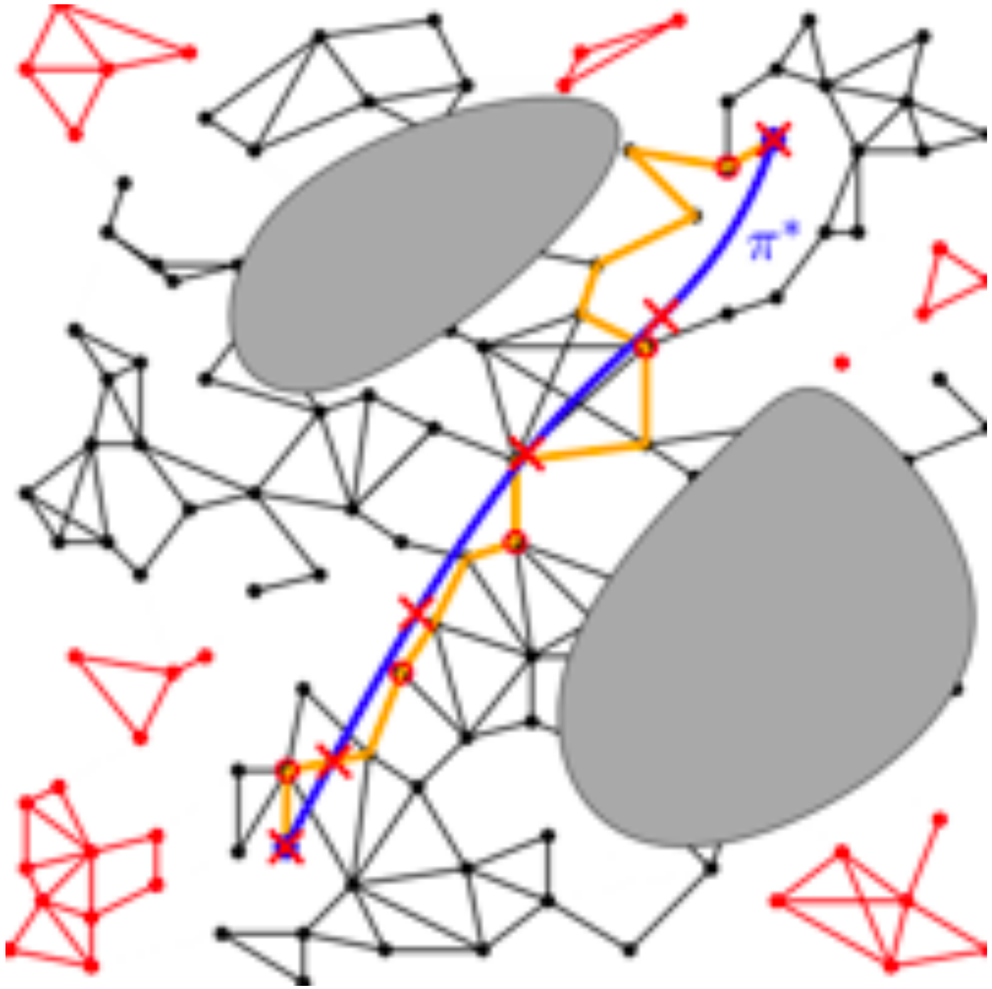
10-100X Improvement

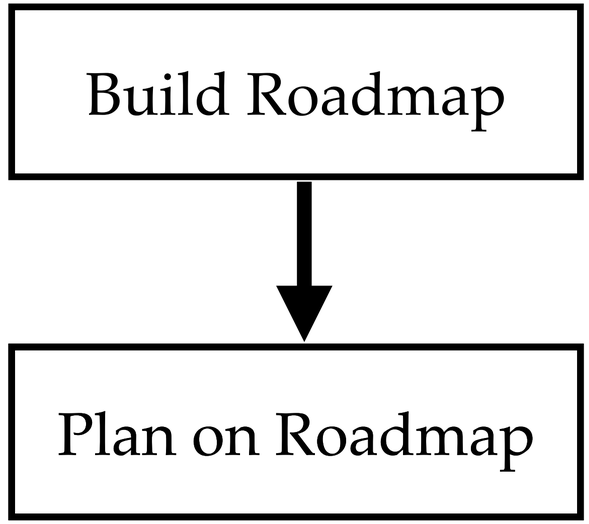
The Piano Movers' Problem



On the Piano Movers problem. I-III, Schwartz and Sharir, Comm. on Pure and Applied Math., 1983

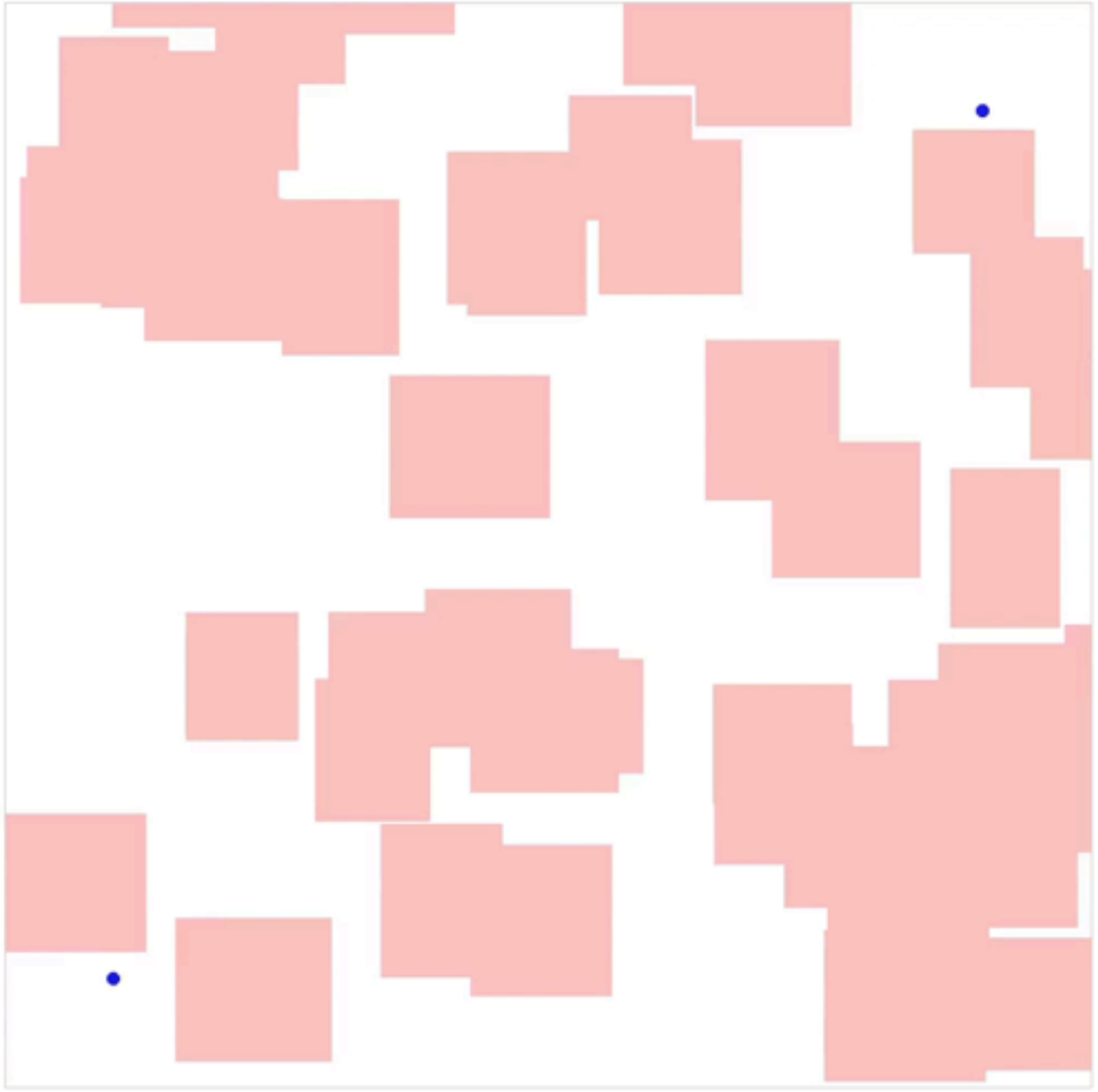
Roadmaps





Probabilistic roadmaps for path planning in high-dimensional configuration spaces, Kavraki et al., IEEE TRO, 1996.

A* Search



A* Search

OPTIMAL!!

Is it optimal over something we care about?

A* Search: A Personal Journey

Search for Optimal Solutions:
the Heart of Heuristic Search is Still Beating

Ariel Felner
ISE Department
Ben-Gurion University
ISRAEL
felner@bgu.ac.il

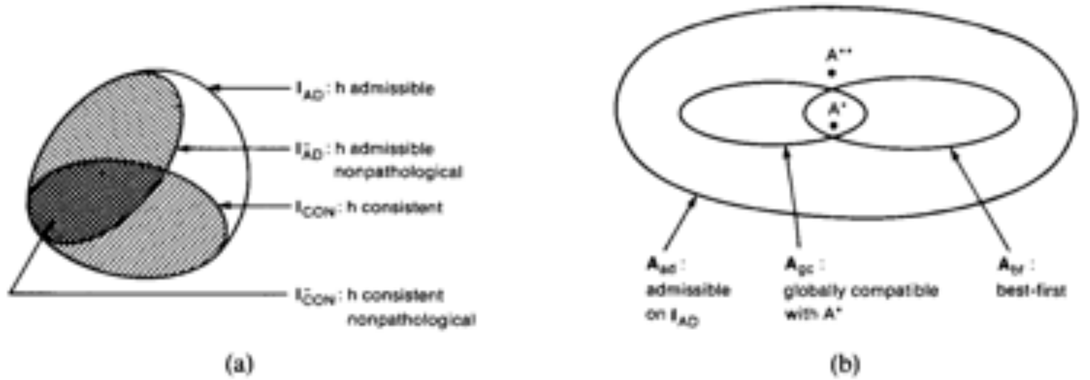
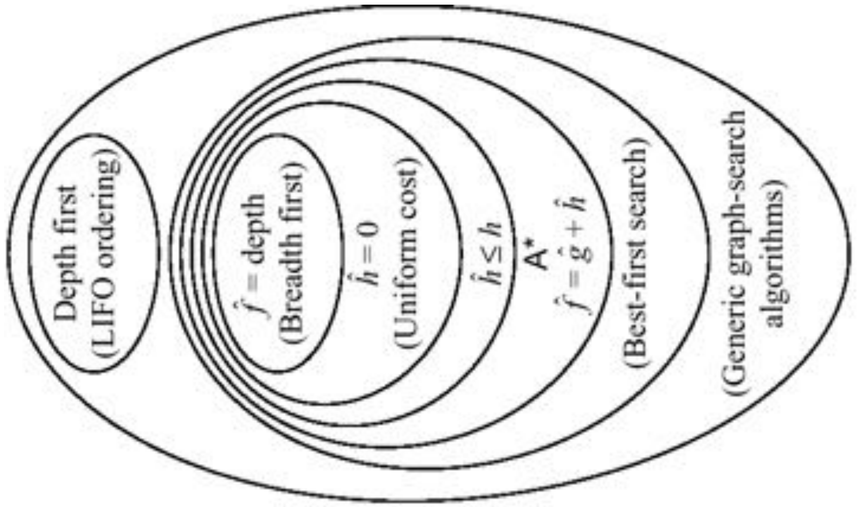
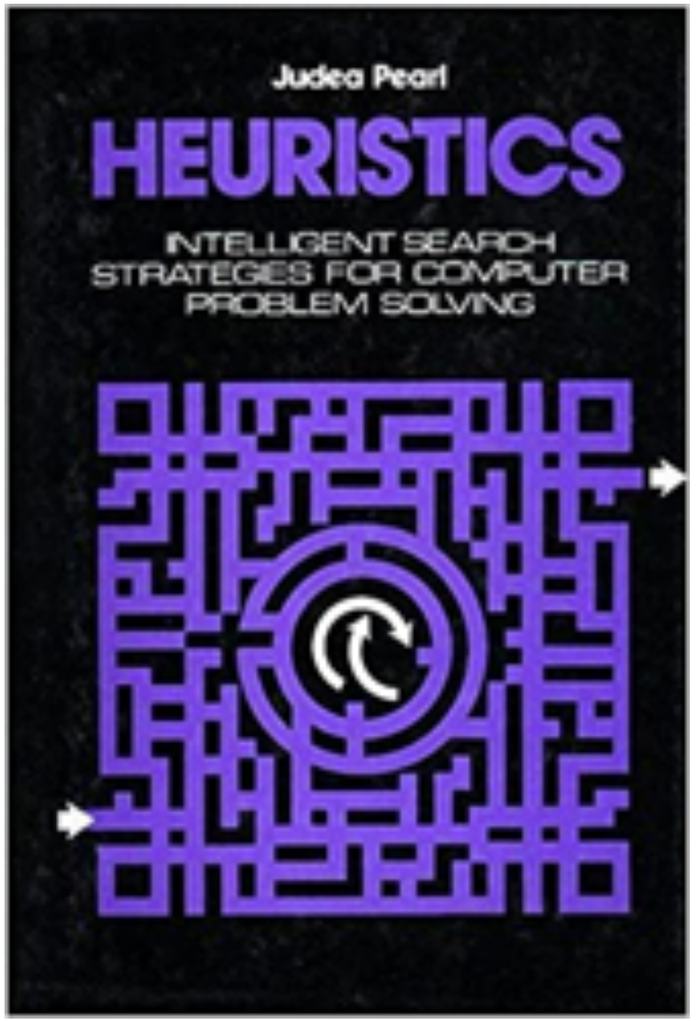
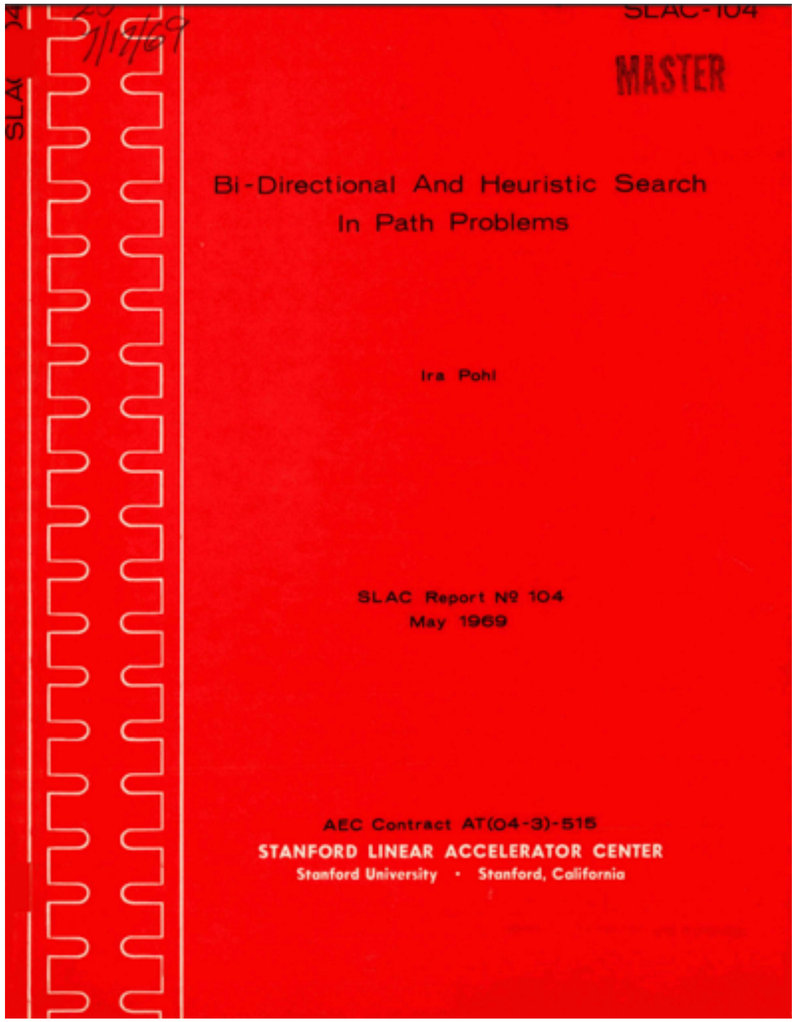


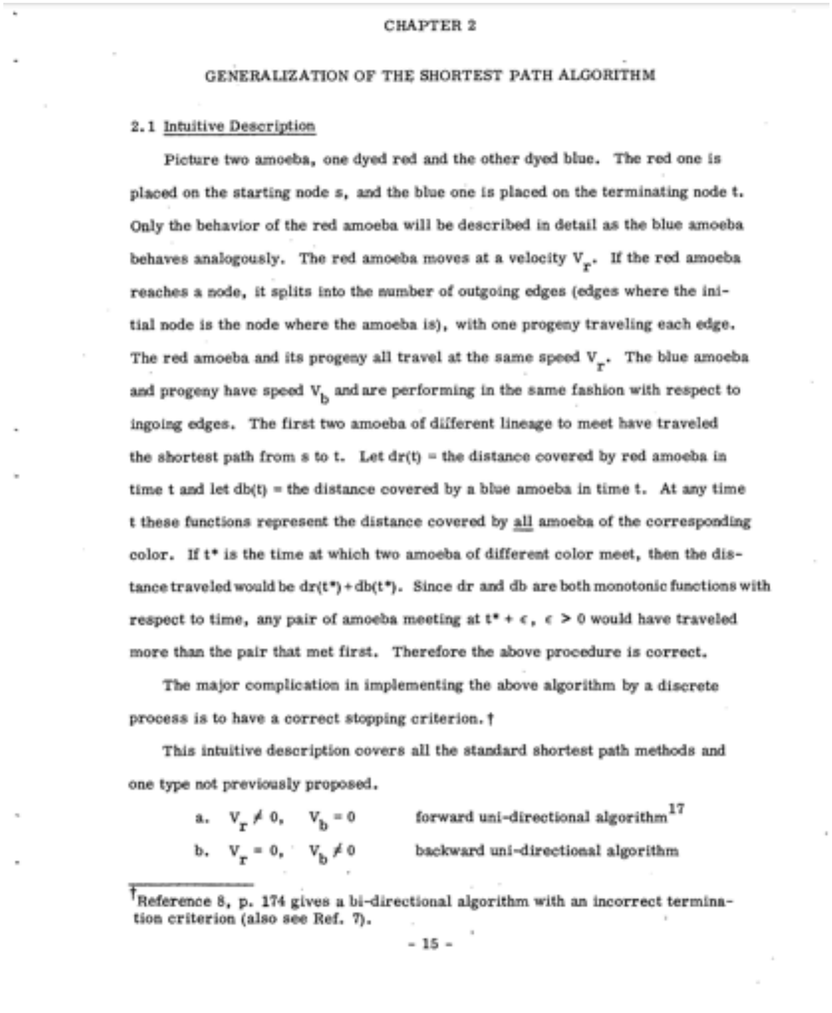
FIG. 4. Venn diagrams showing the classes of algorithms and the domains of problem instances relative to which the optimality of A* is examined. (a) Problem domains. (b) Classes of algorithms.



A* Search: A Personal Journey







CHAPTER 2

GENERALIZATION OF THE SHORTEST PATH ALGORITHM

2.1 Intuitive Description

Picture two amoeba, one dyed red and the other dyed blue. The red one is placed on the starting node s, and the blue one is placed on the terminating node t. Only the behavior of the red amoeba will be described in detail as the blue amoeba behaves analogously. The red amoeba moves at a velocity V_r. If the red amoeba reaches a node, it splits into the number of outgoing edges (edges where the initial node is the node where the amoeba is), with one progeny traveling each edge. The red amoeba and its progeny all travel at the same speed V_r. The blue amoeba and progeny have speed V_b and are performing in the same fashion with respect to ingoing edges. The first two amoeba of different lineage to meet have traveled the shortest path from s to t. Let dr(t) = the distance covered by red amoeba in time t and let db(t) = the distance covered by a blue amoeba in time t. At any time t these functions represent the distance covered by all amoeba of the corresponding color. If t* is the time at which two amoeba of different color meet, then the distance traveled would be dr(t*) + db(t*). Since dr and db are both monotonic functions with respect to time, any pair of amoeba meeting at $t^* + \epsilon$, $\epsilon > 0$ would have traveled more than the pair that met first. Therefore the above procedure is correct.

The major complication in implementing the above algorithm by a discrete process is to have a correct stopping criterion.†

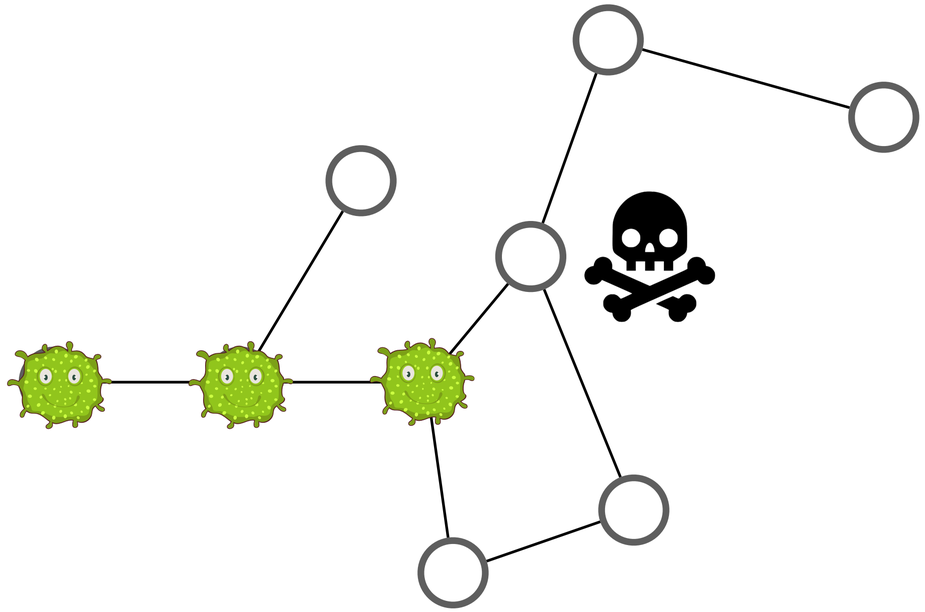
This intuitive description covers all the standard shortest path methods and one type not previously proposed.

a. $V_r \neq 0$, $V_b = 0$ forward uni-directional algorithm[17]

b. $V_r = 0$, $V_b \neq 0$ backward uni-directional algorithm

†Reference 8, p. 174 gives a bi-directional algorithm with an incorrect termination criterion (also see Ref. 7).

A* Search: Amoebas!



Optimal Substructure

$$f(a) < f(b) \implies f(a \circ x) < f(b \circ x) \forall x$$

You will never catch up.

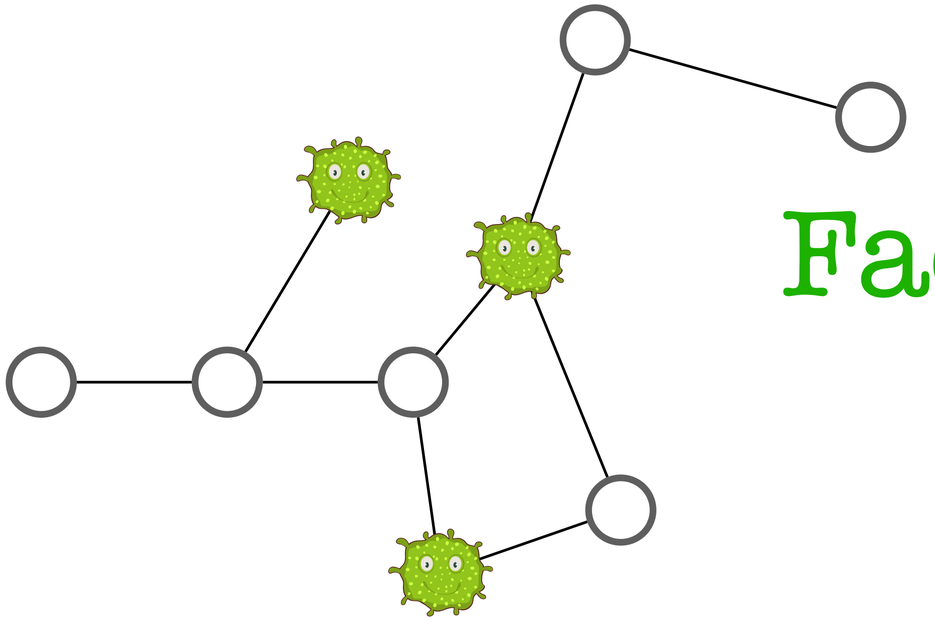
Bellman Condition

$$f^*(a) = \min_{x \in \text{succ}} \{c(a,x) + f^*(b)\}$$

Be best, locally.

Bacteria Vectors by Vecteezy

A* Search: Favoritism



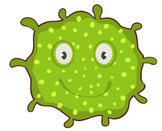
Optimism in the Face of Uncertainty (OFU)

$$\min_{x \in \text{open}} \; g(x) + h(x)$$

Always be optimistic under uncertainty.
You'll either be correct,
or learn something important if you're wrong.

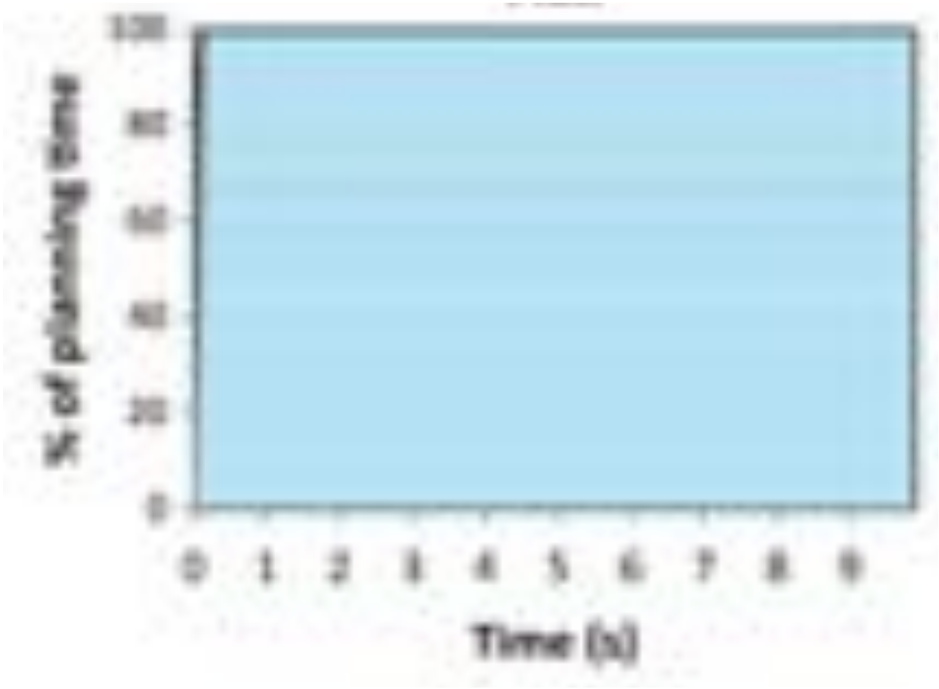
R-MAX: A general polynomial time algorithm for near-optimal reinforcement learning, Brafman and Tennenholtz, JMLR, 2002.

A* Search is Optimal …

Expands the Fewest Number of Vertices 

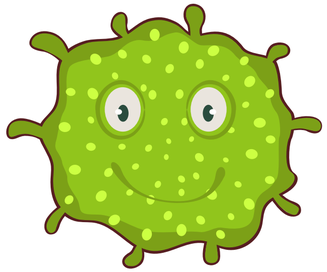
But is this what we
really want in Motion Planning?

Edge Evaluation Dominates Planning Time



Edge Evaluations

Other

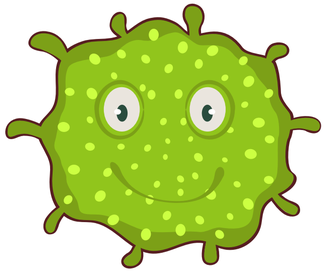


Amoebas are Cheap

Slime is Expensive

Lazy collision checking in asymptotically-optimal motion planning, Hauser, ICRA 2015.

Is there a Search Algorithm
that Minimizes
the Number of Edge Evaluations?



I don't care about amoebas.
What algorithm minimizes slime?

LazySP

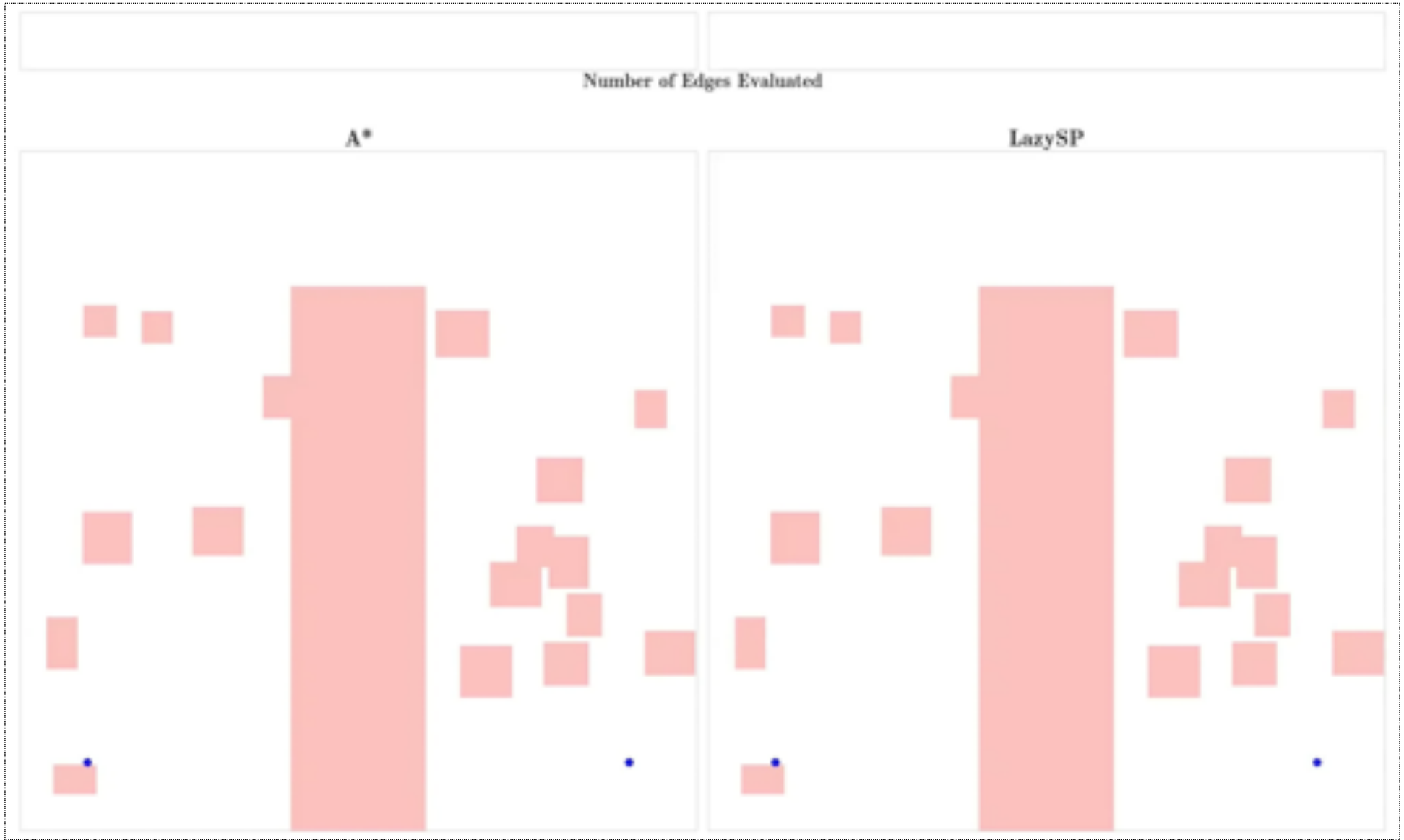
ICAPS 2018 [Best Conference Paper Award Winner]

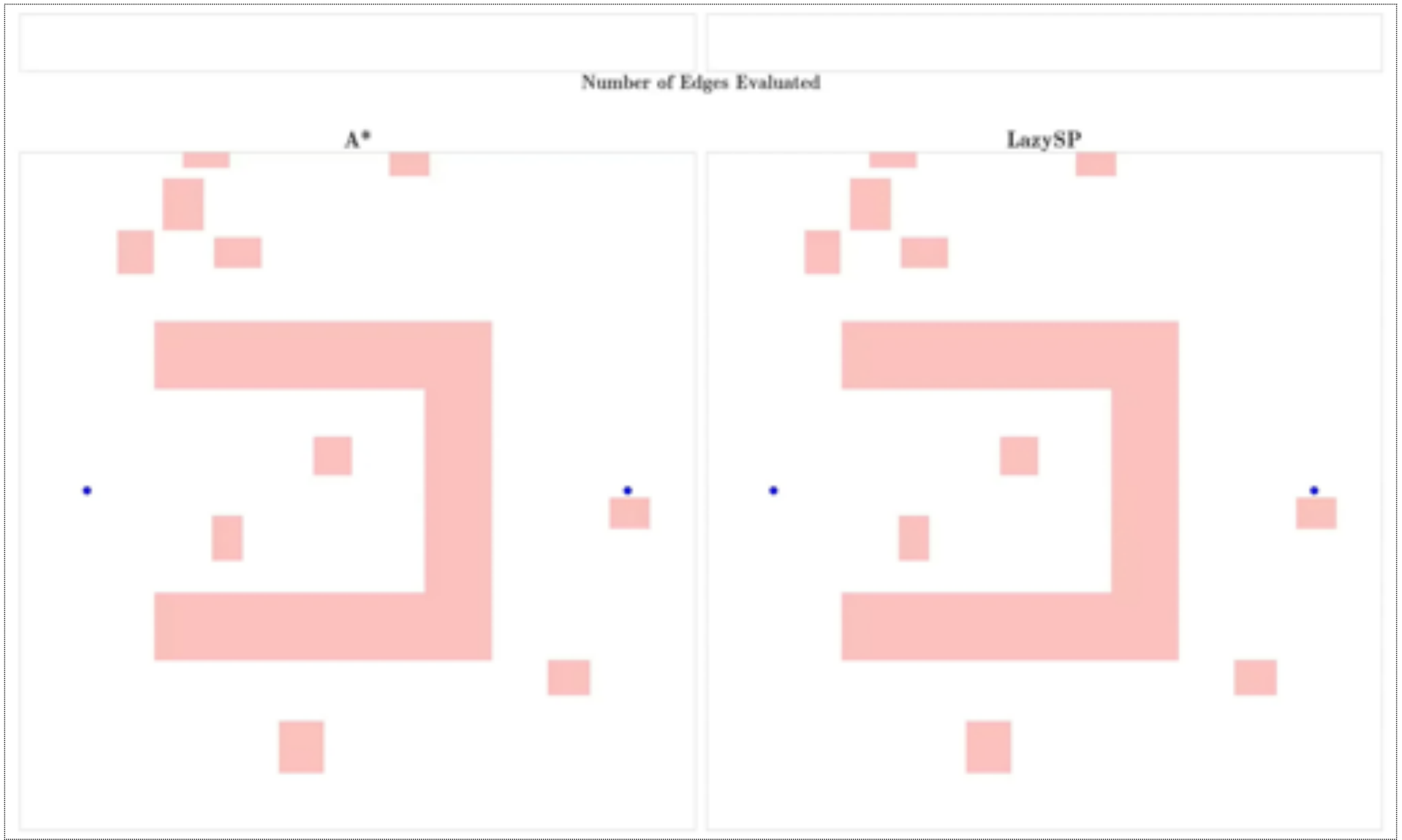
First Provably Edge-Optimal A*-like Search Algorithm

The Provable Virtue of Laziness in Motion Planning, Hagtalab et al., ICAPS 2018.


Number of Edges Evaluated
A*
LazySP


Number of Edges Evaluated
A*
LazySP


Number of Edges Evaluated
A*
LazySP


Number of Edges Evaluated
A*
LazySP

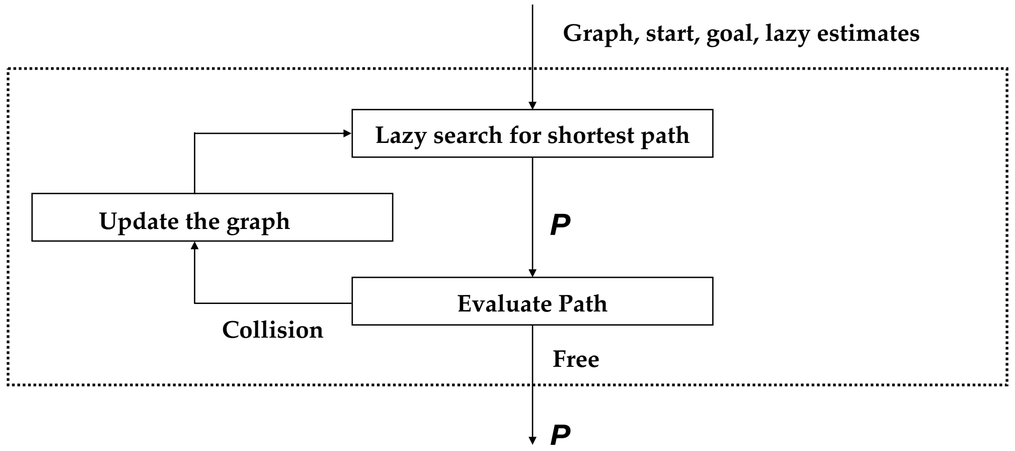
LazySP

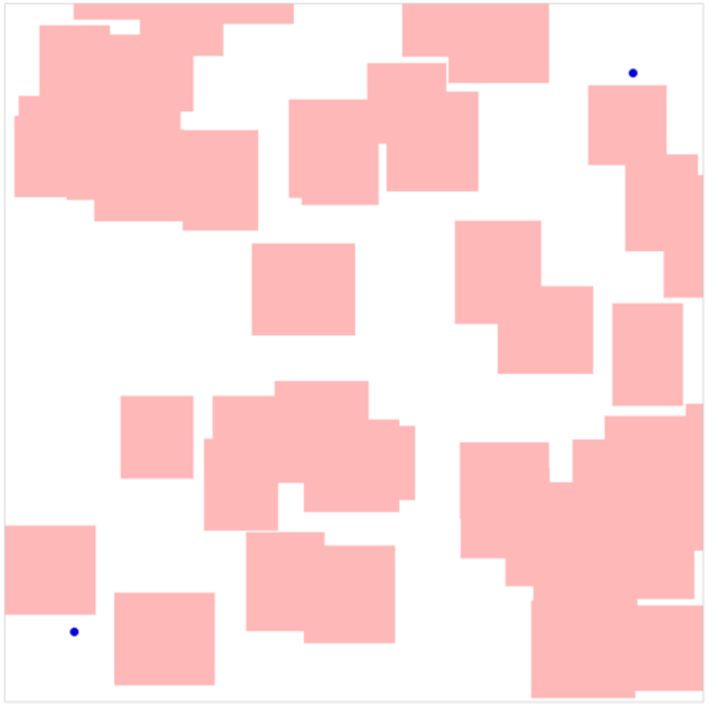
Greedy Best-first Search over Paths

To find the shortest path,
eliminate all shorter paths!

LazySP

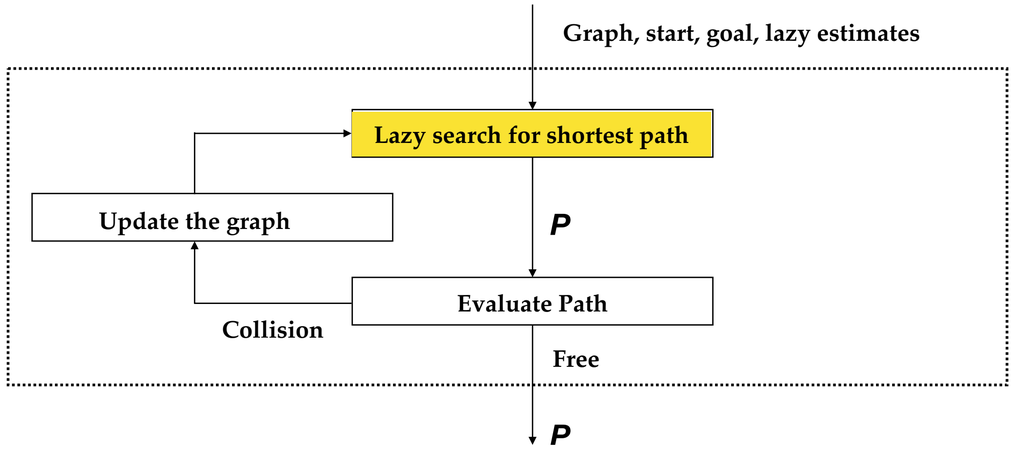
OFU on Steroids!

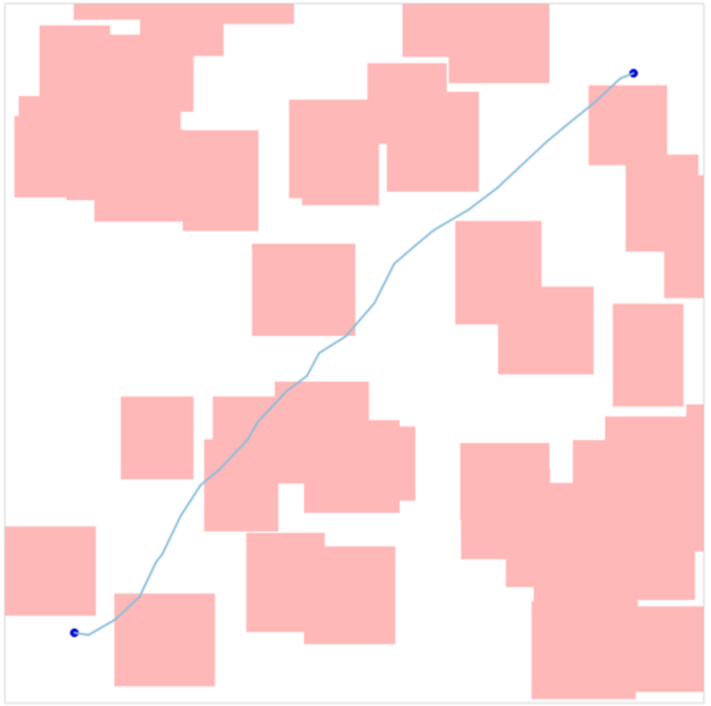


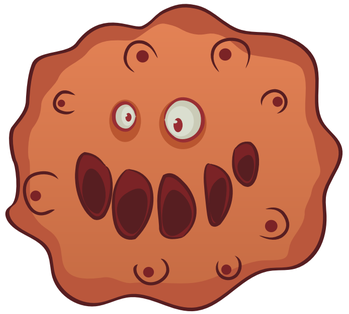


LazySP

OFU on Steroids!



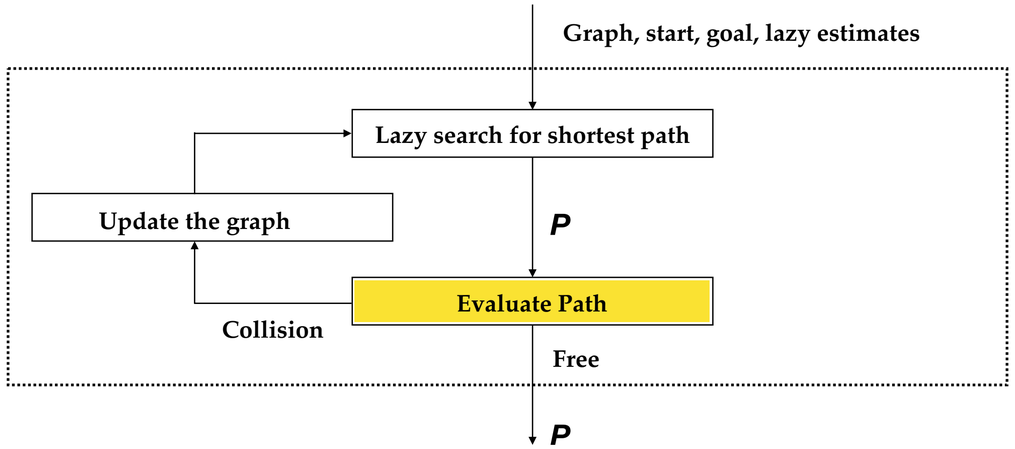


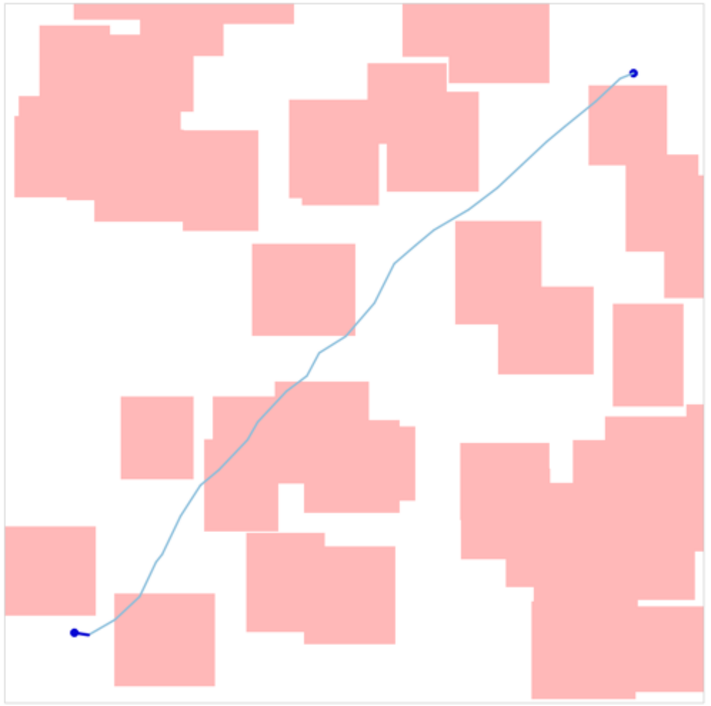


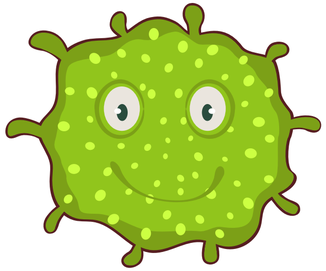
Send out the Ghost Amoebas

LazySP

OFU on Steroids!



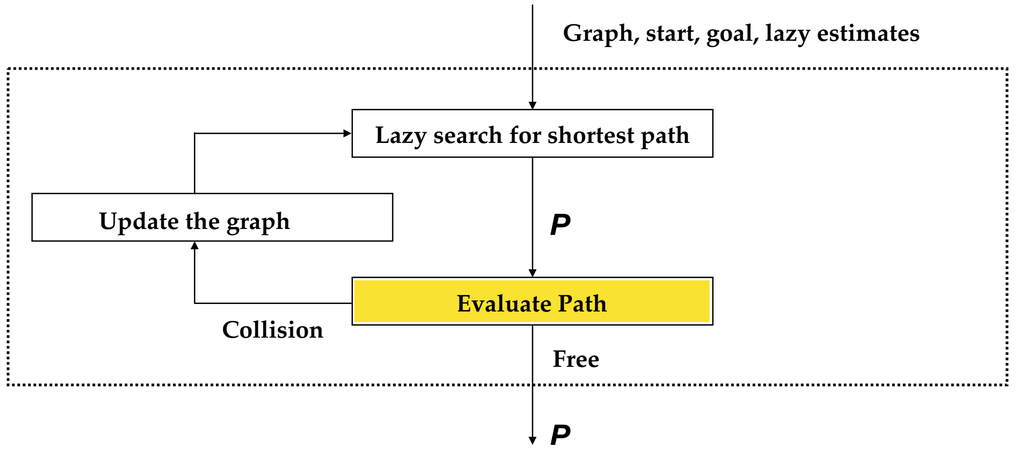


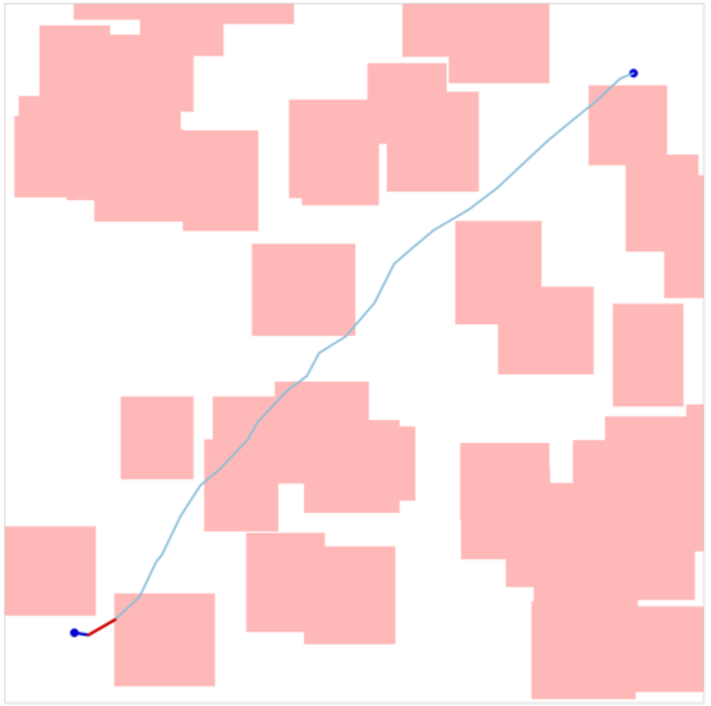


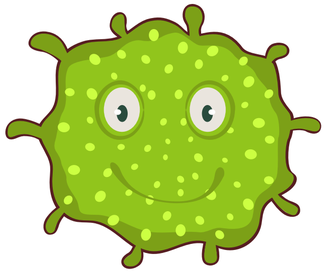
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



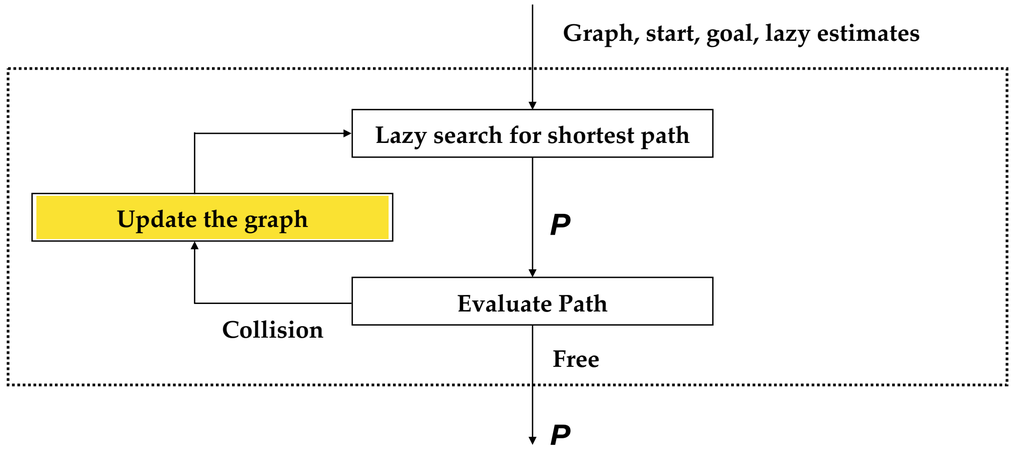


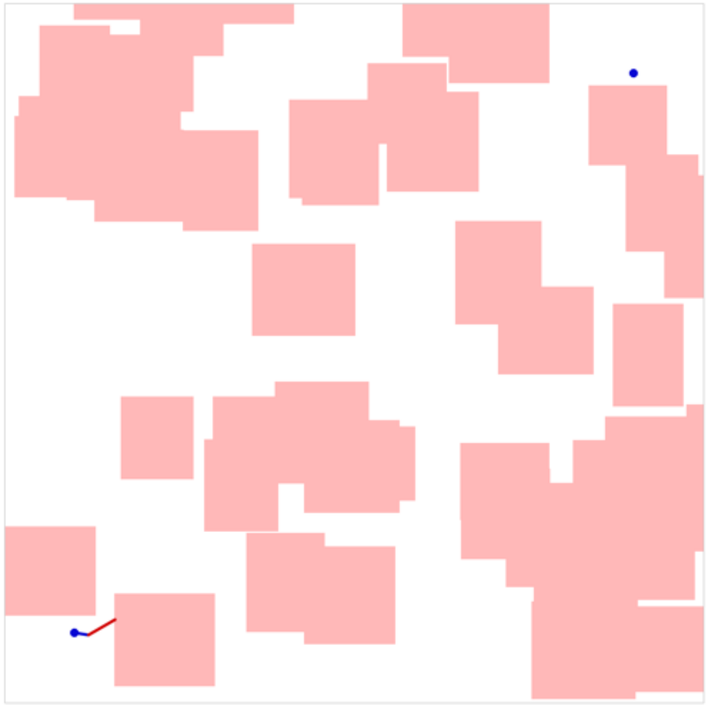


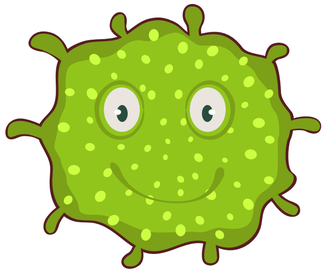
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



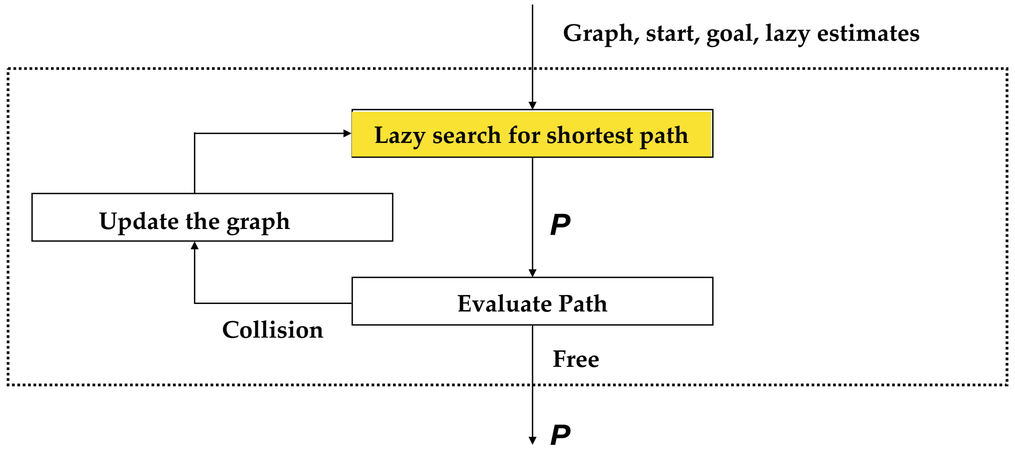


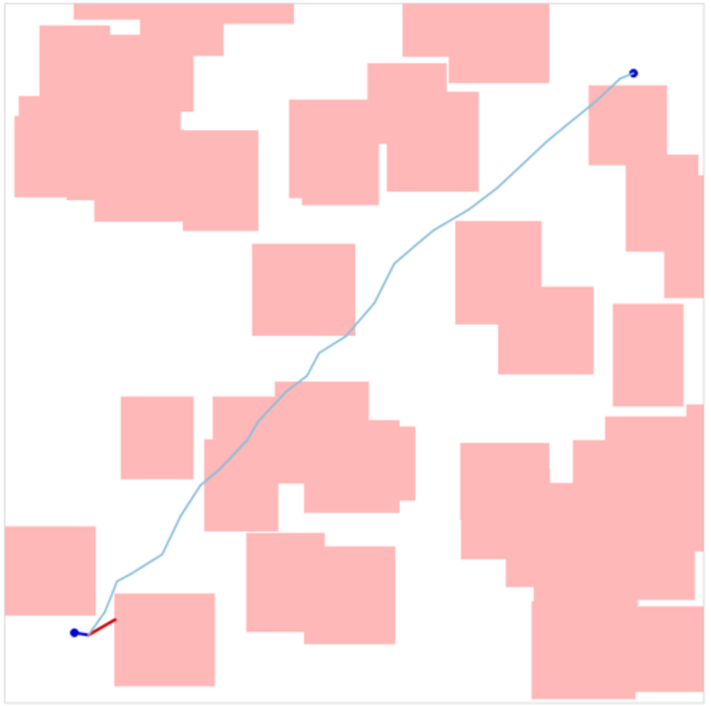


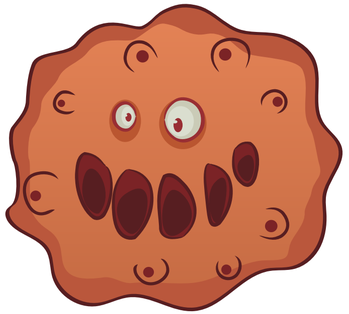
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



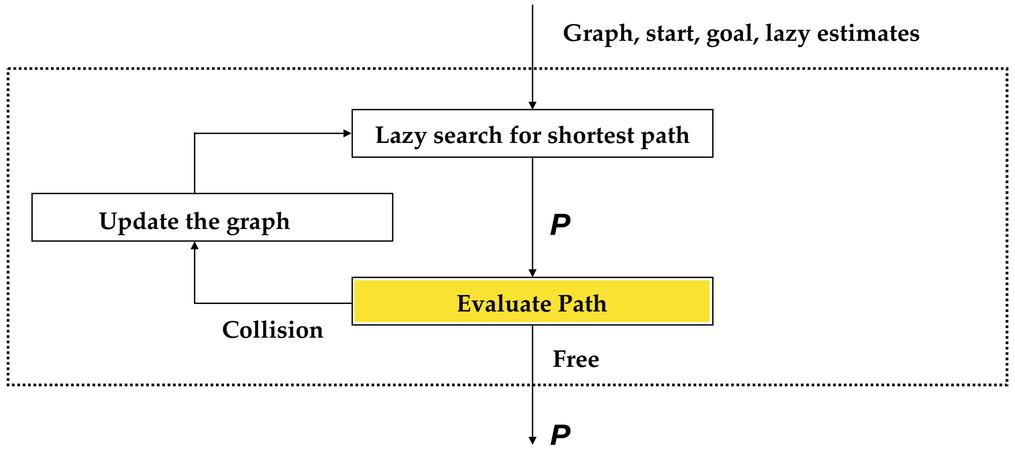


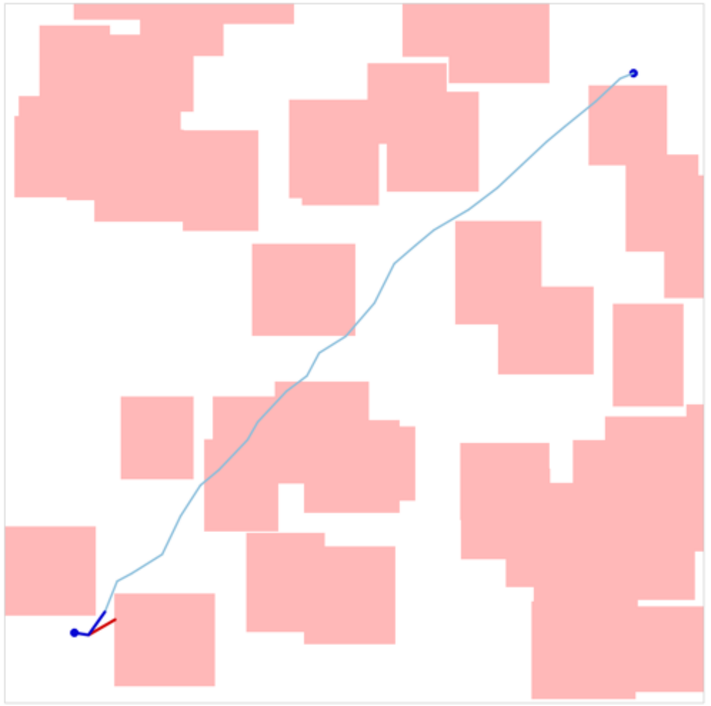


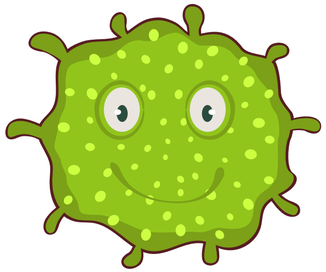
Send out the Ghost Amoebas

LazySP

OFU on Steroids!



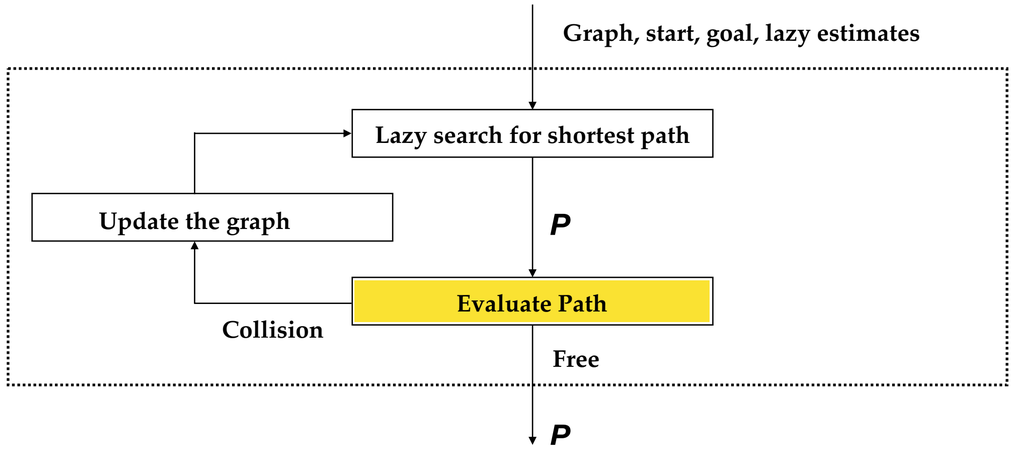


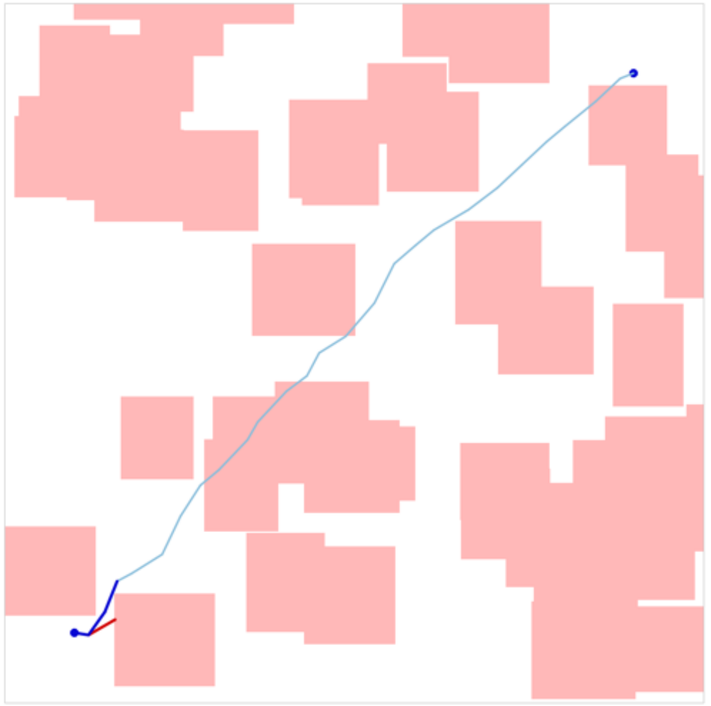


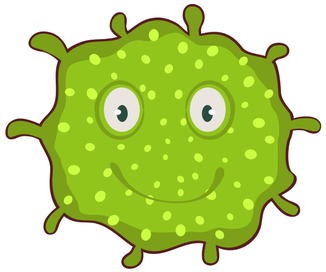
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



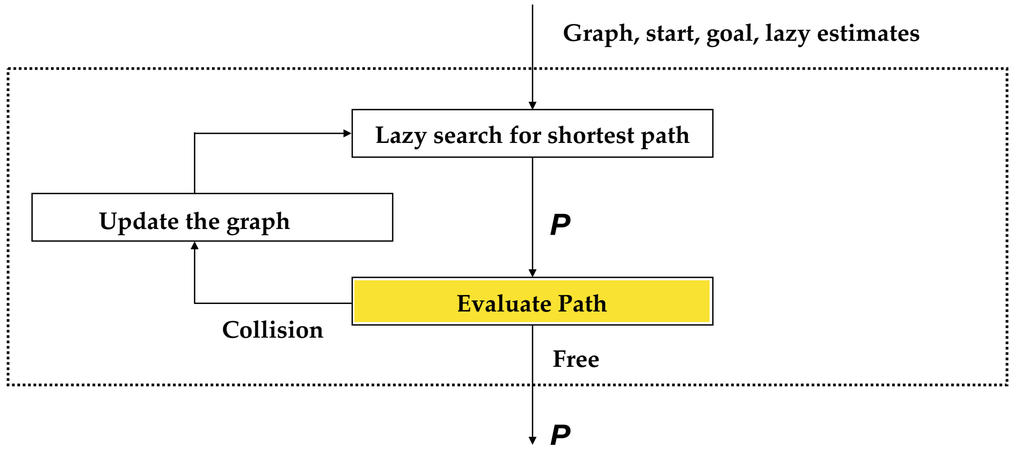


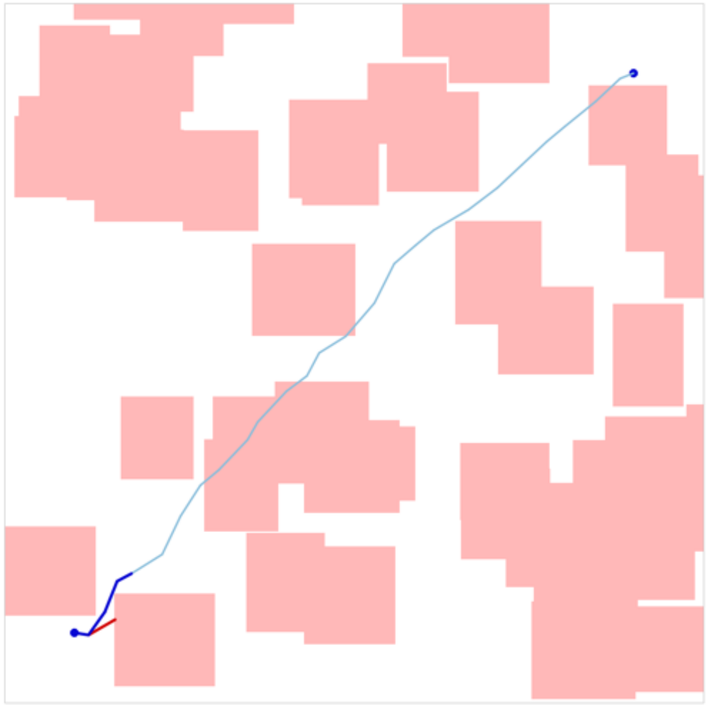


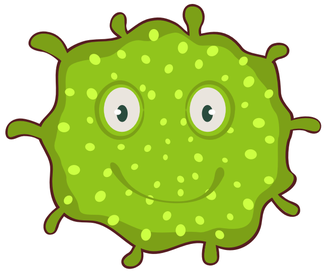
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



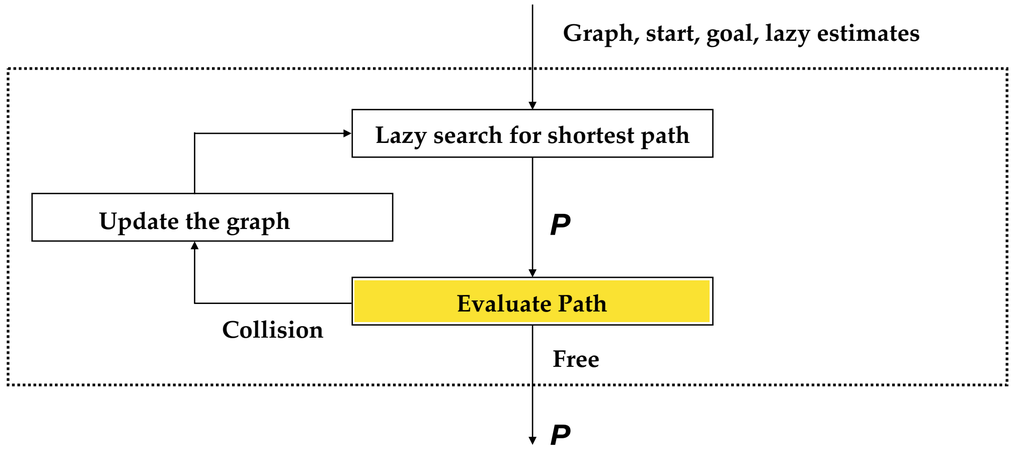


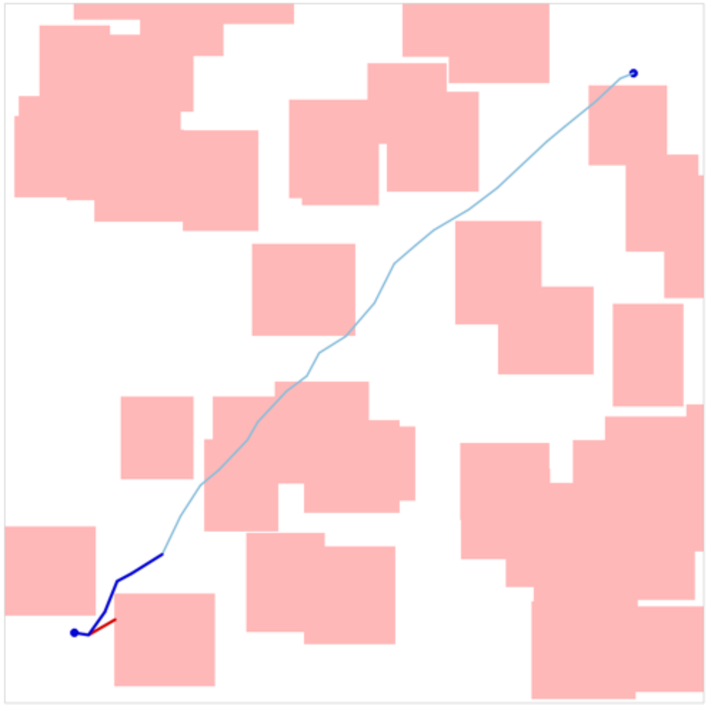


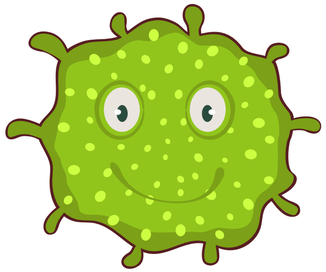
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



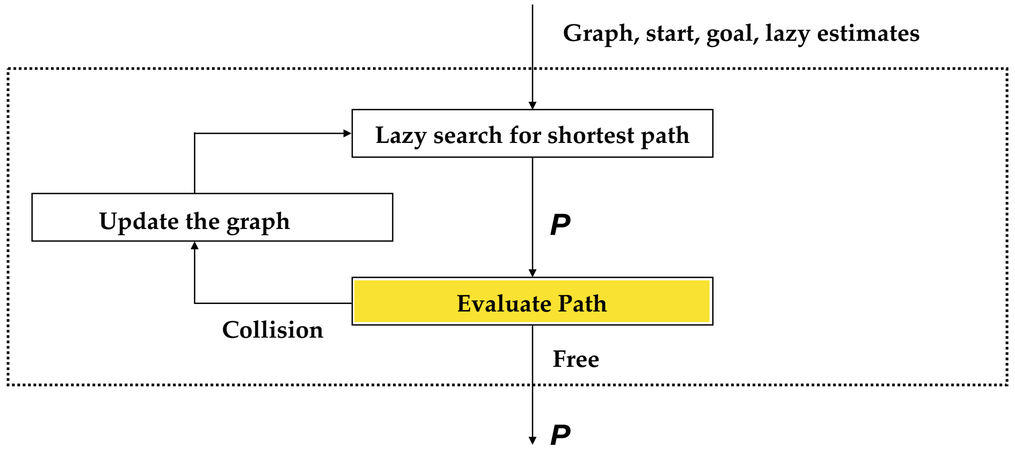


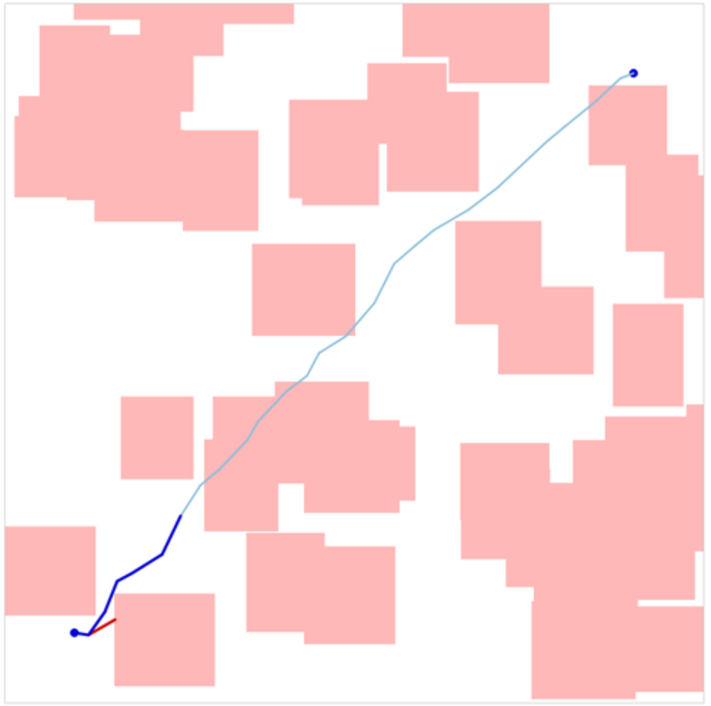


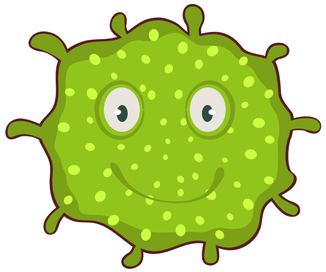
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



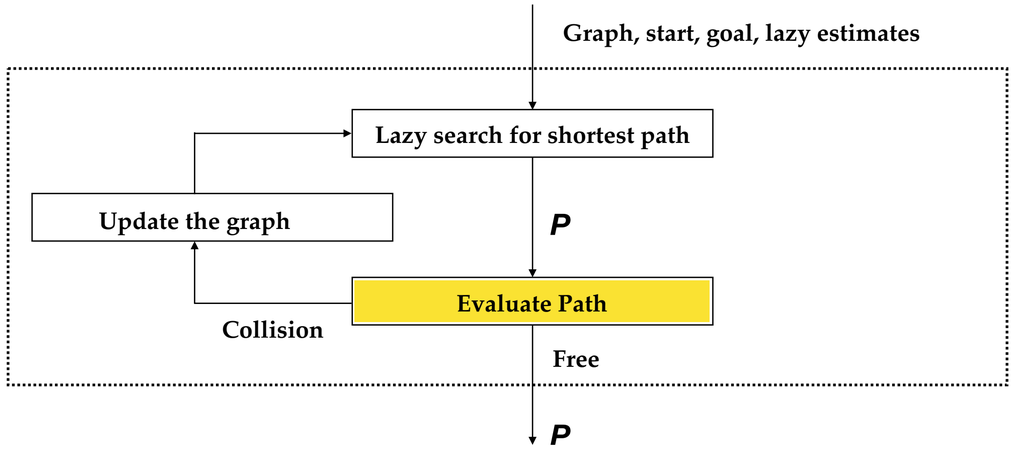


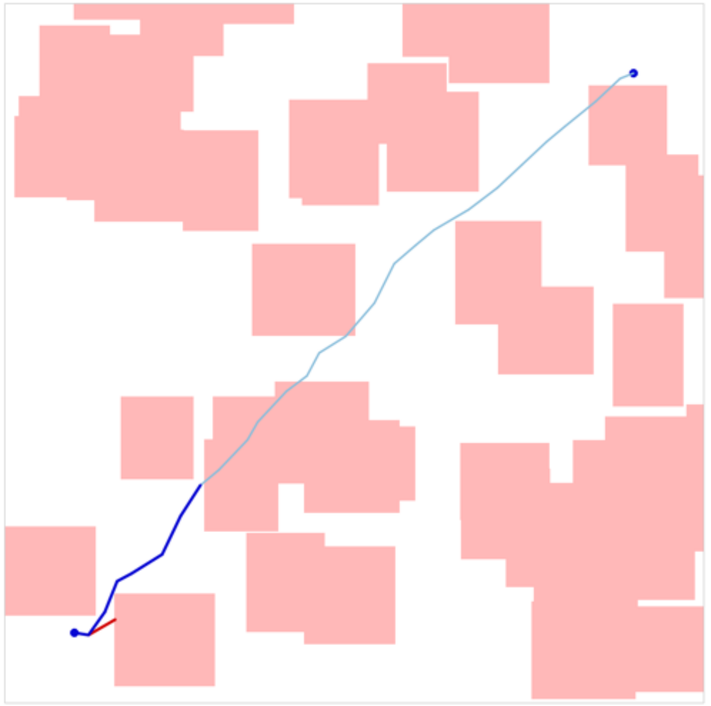


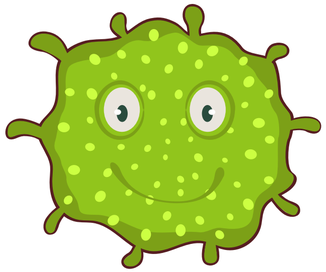
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



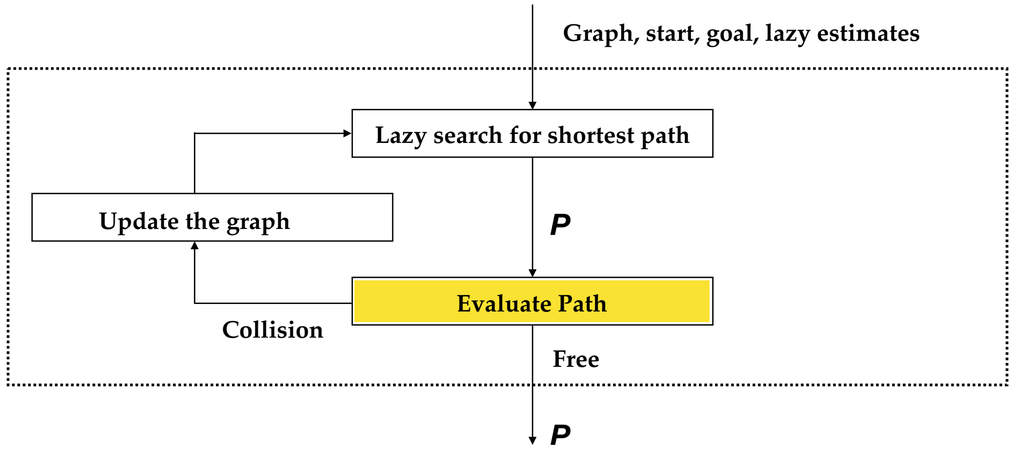


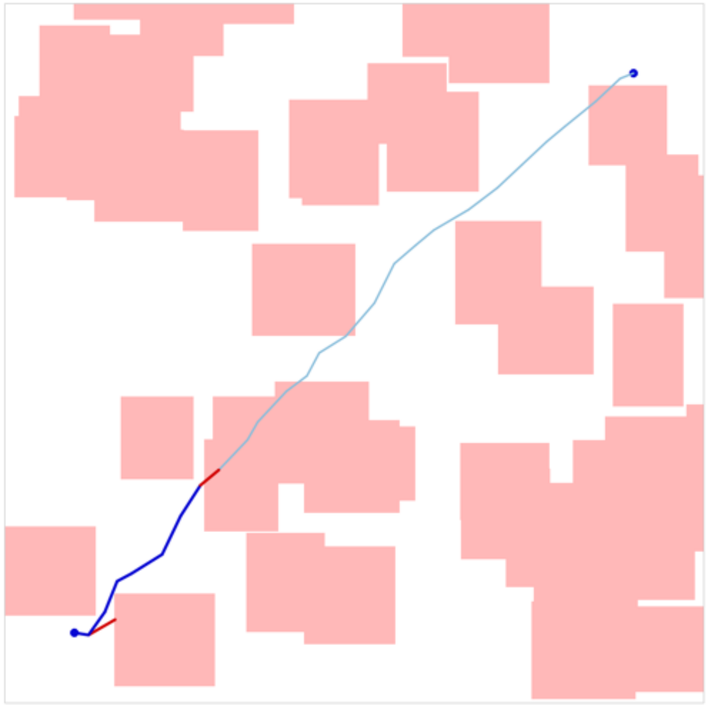


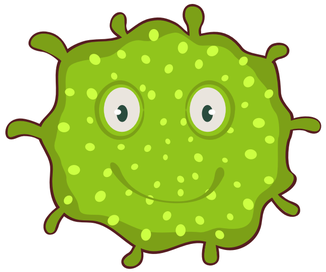
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



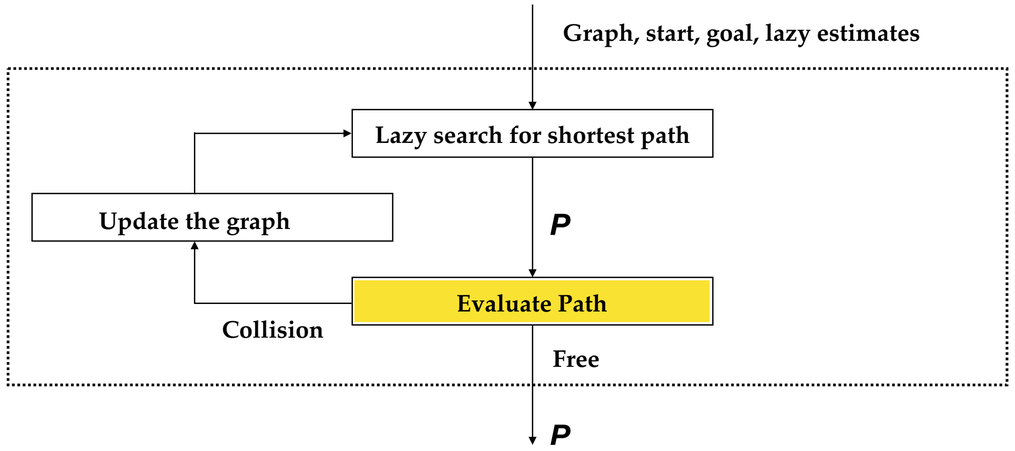


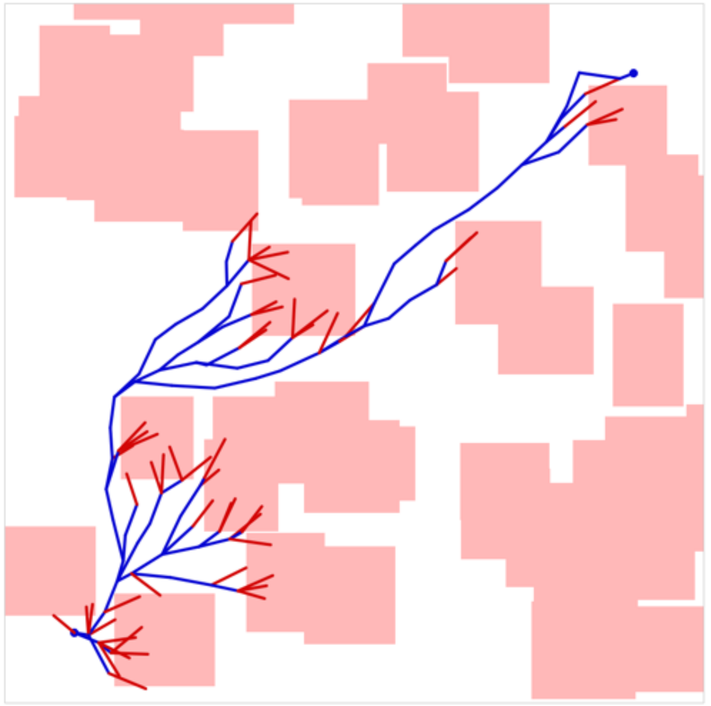


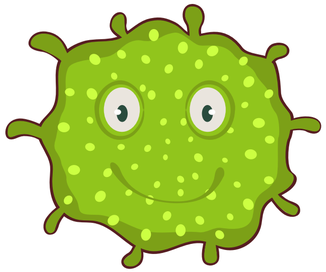
Only Slime Known Shortest Paths

LazySP

OFU on Steroids!



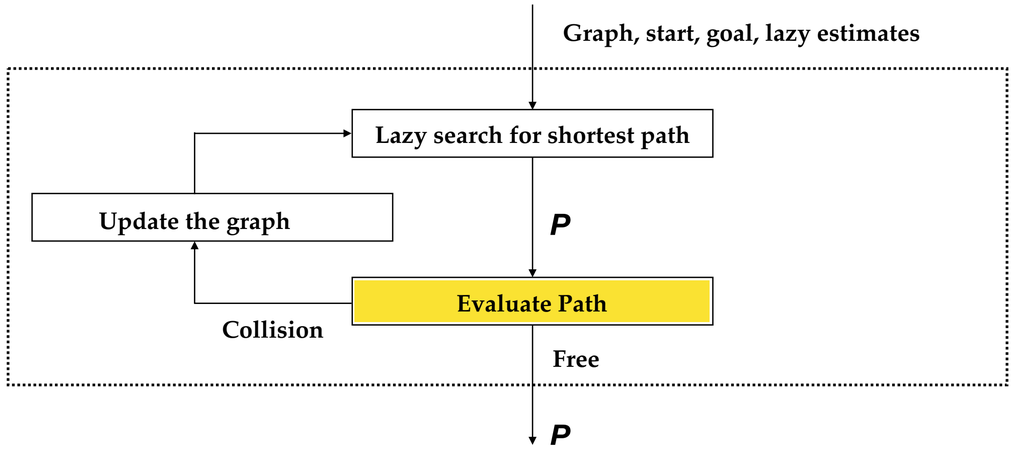


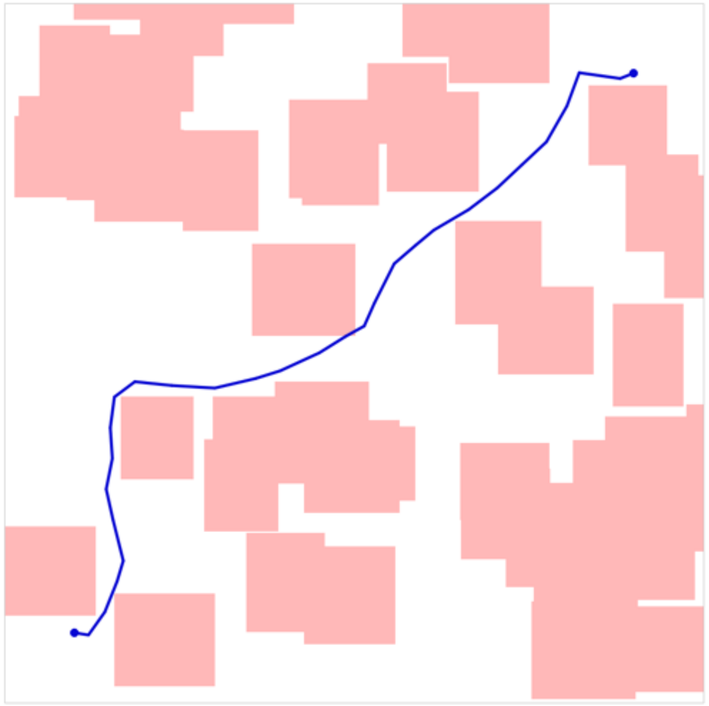


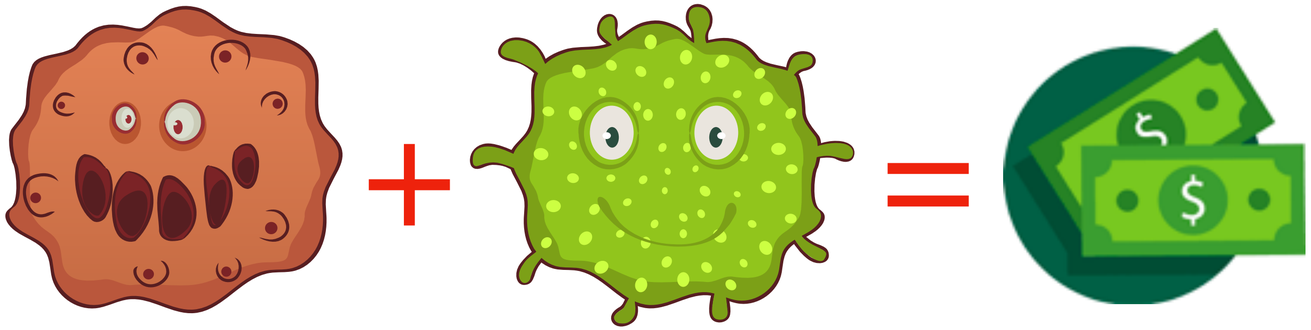
Optimal Slime!

LazySP

OFU on Steroids!

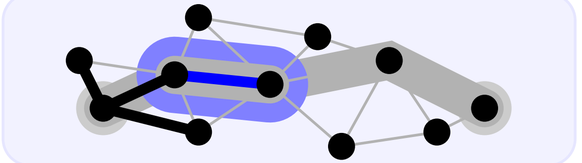




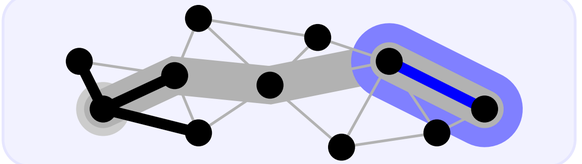


Edge Selectors

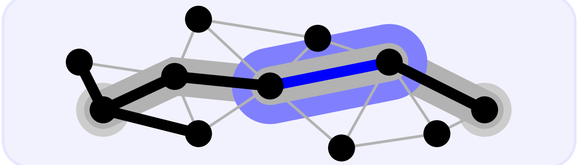
Forward
(first unevaluated edge)



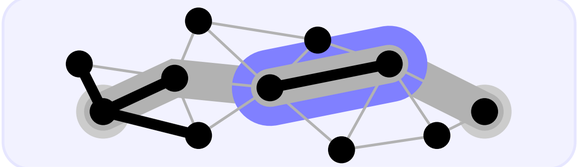
Reverse
(last unevaluated edge)



Alternate
(alternate Forward and Reverse)



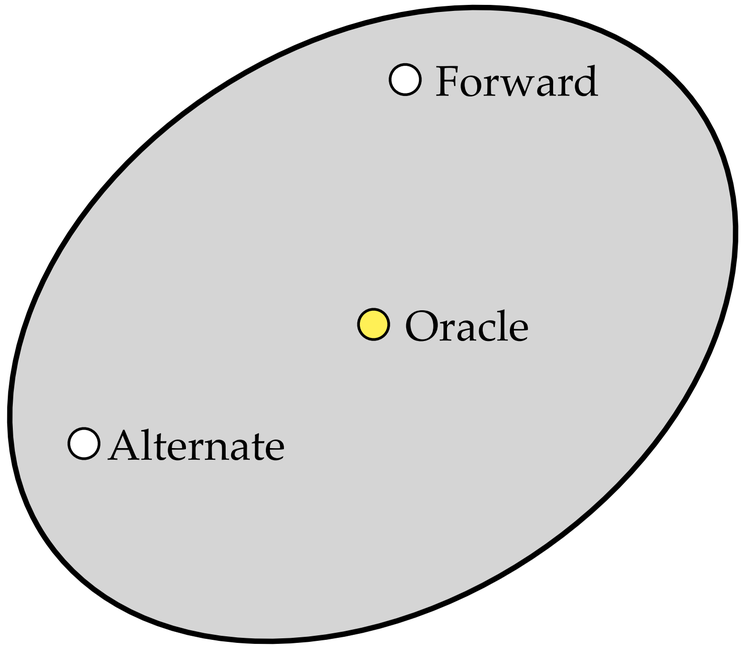
Bisect
(furthest from an unevaluated edge)



The Realizability Assumption

Can we Learn to Imitate the Oracle?

Leveraging experience in lazy search, Bhardwaj et al., RSS 2019.



Hypothesis Class
All LazySP Selectors

The Oracle is a LazySP Selector!

The Provable Virtue of Laziness in Motion Planning, Hagtalab et al., ICAPS 2018.

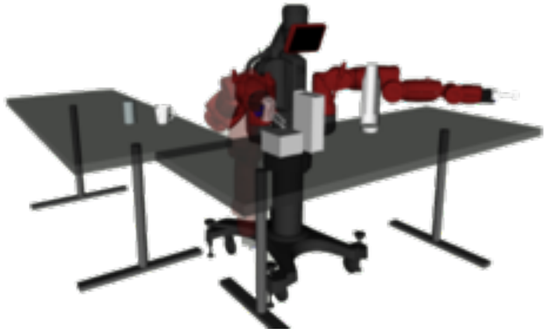
Is there a Search Algorithm
that Minimizes
the Number of Edge Evaluations?

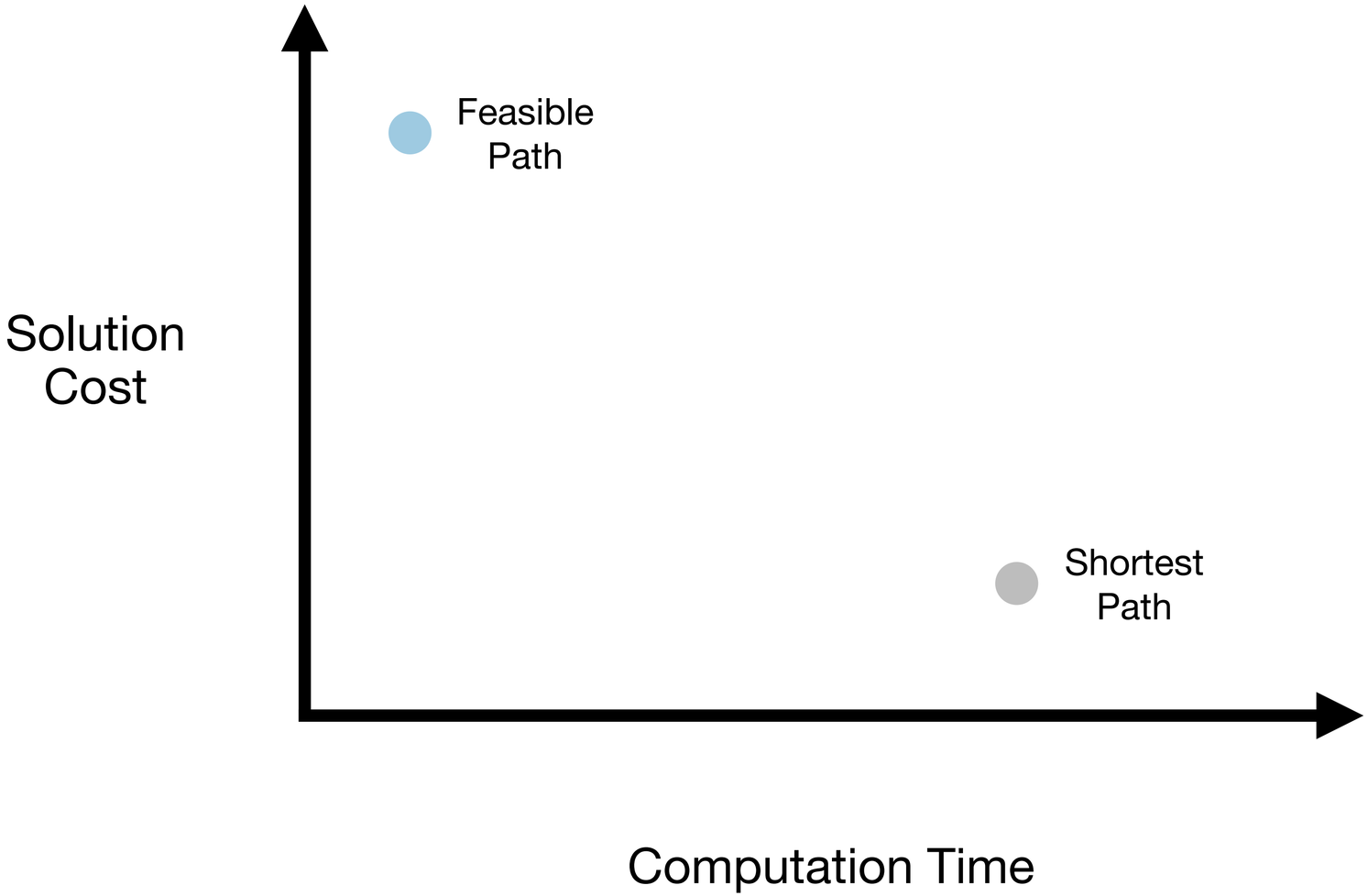
LazySP

ICAPS 2018 [Best Conference Paper Award Winner]

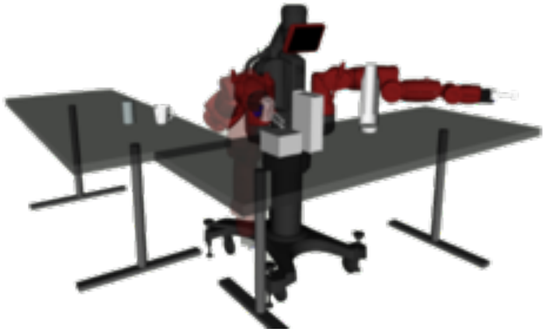
First Provably Edge-Optimal A*-like Search Algorithm

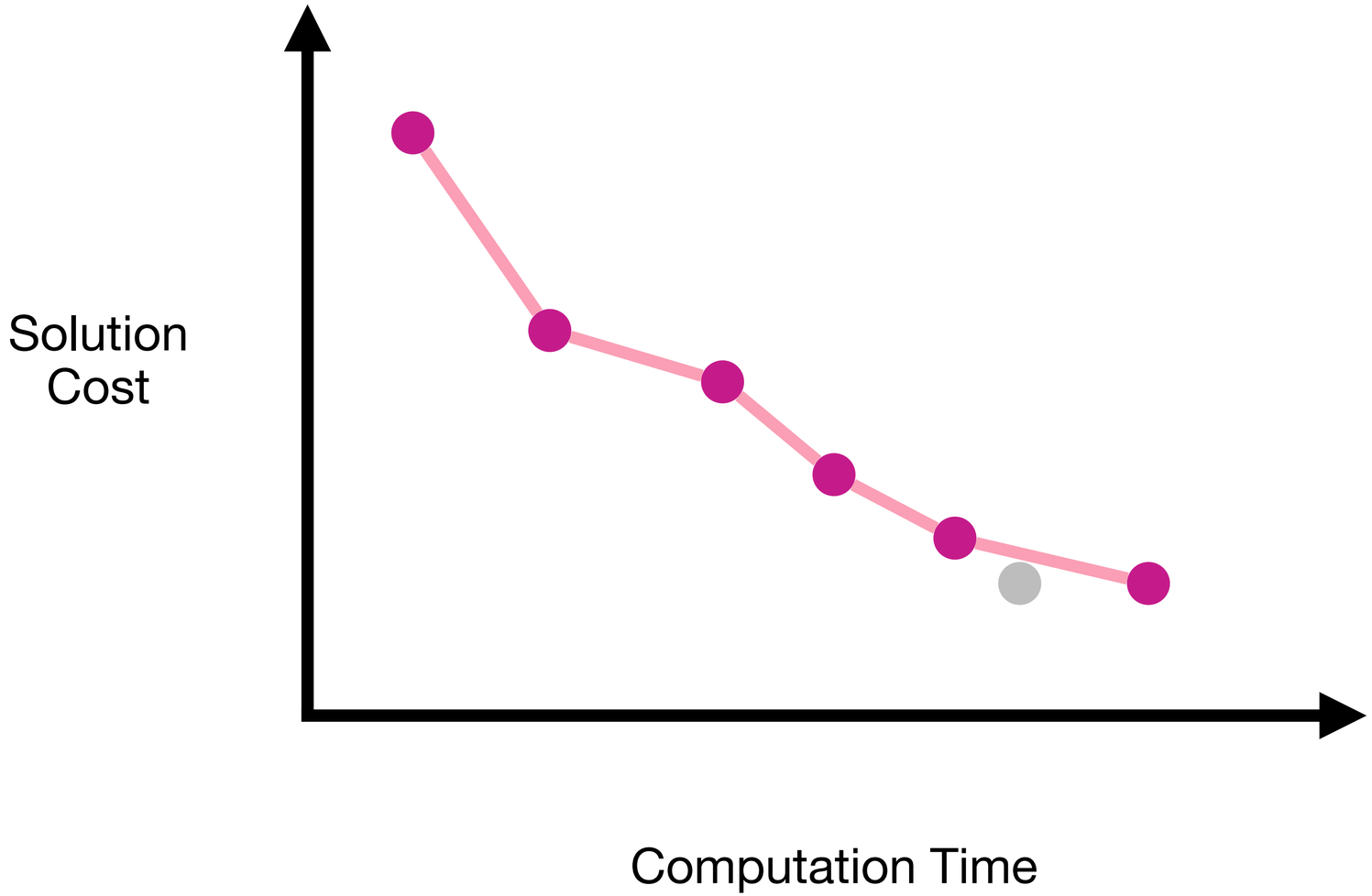
Anytime Motion Planning



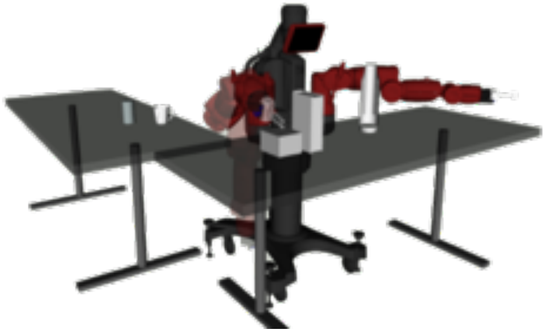


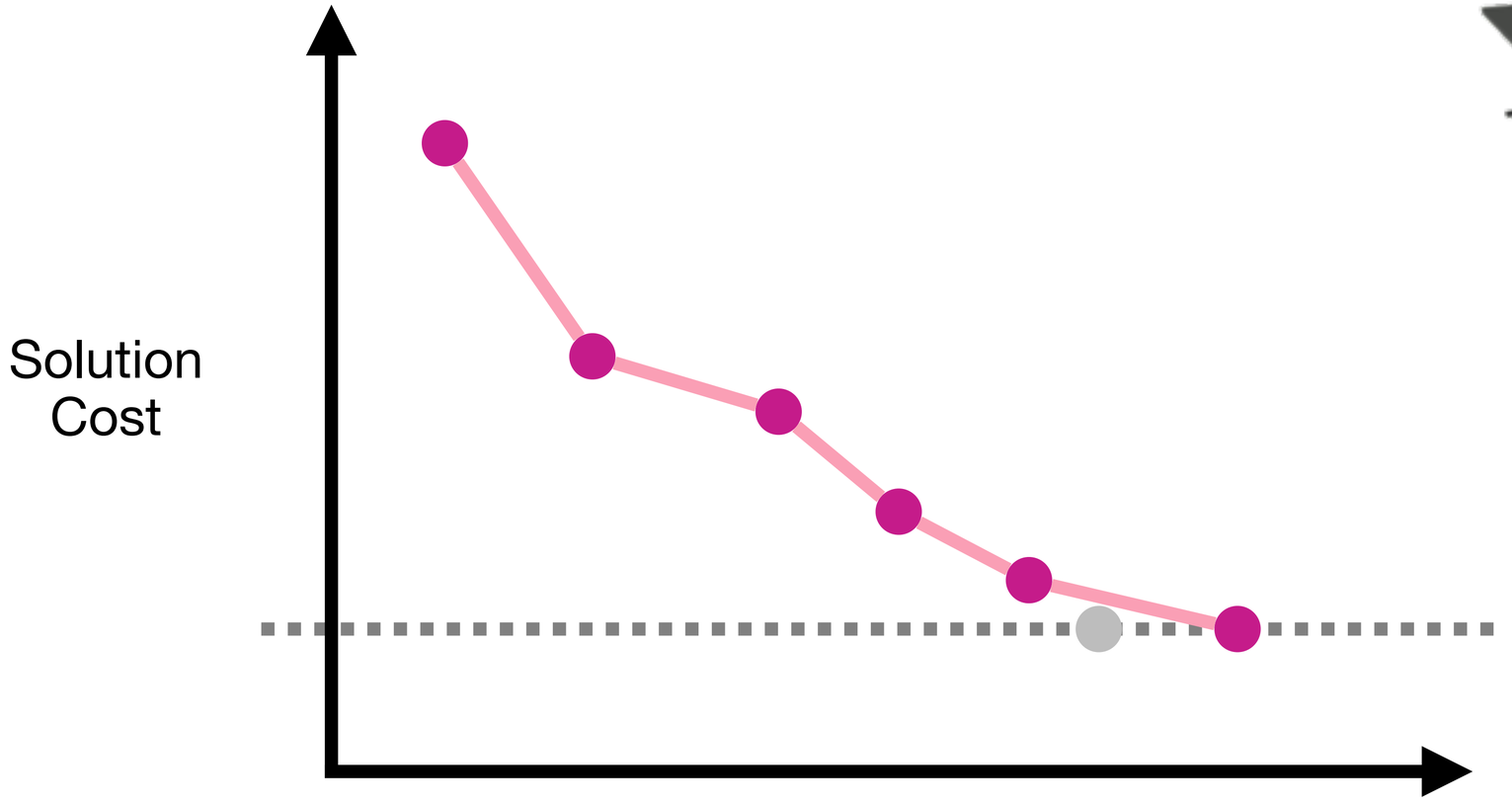
Anytime Motion Planning



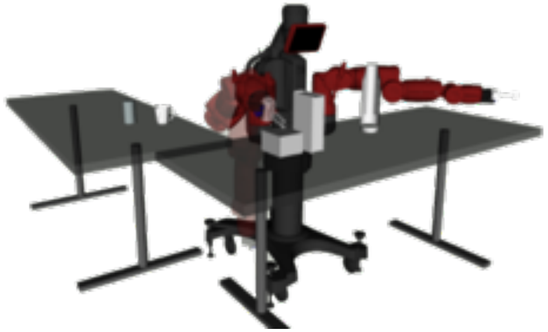


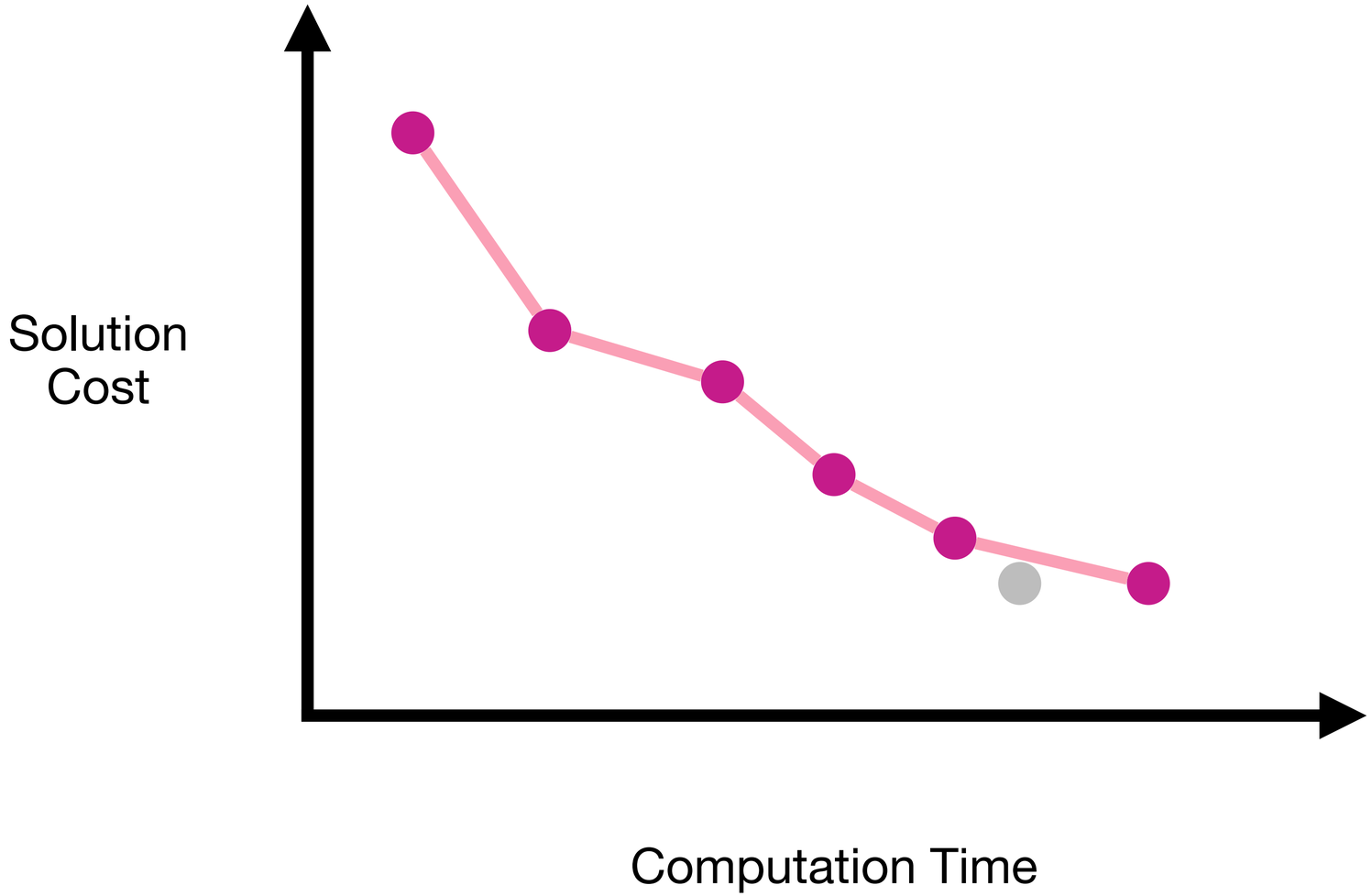
Will it converge to the shortest path?



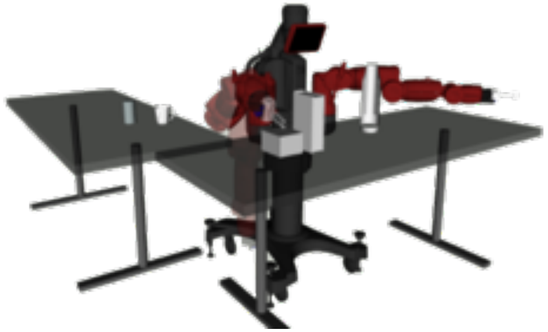


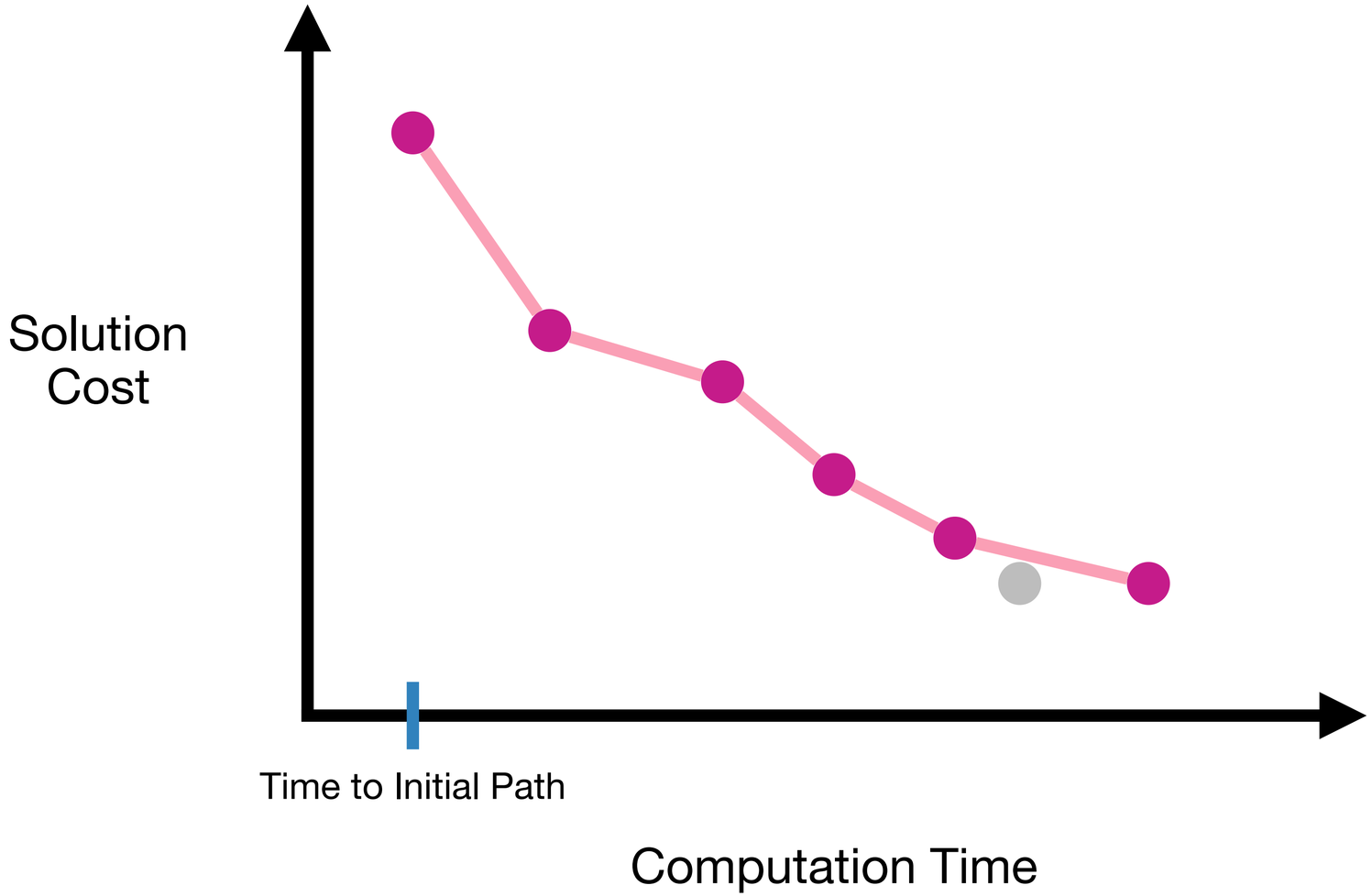
Beyond Asymptotic Optimality



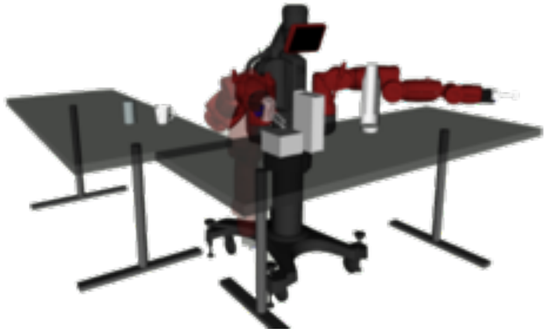


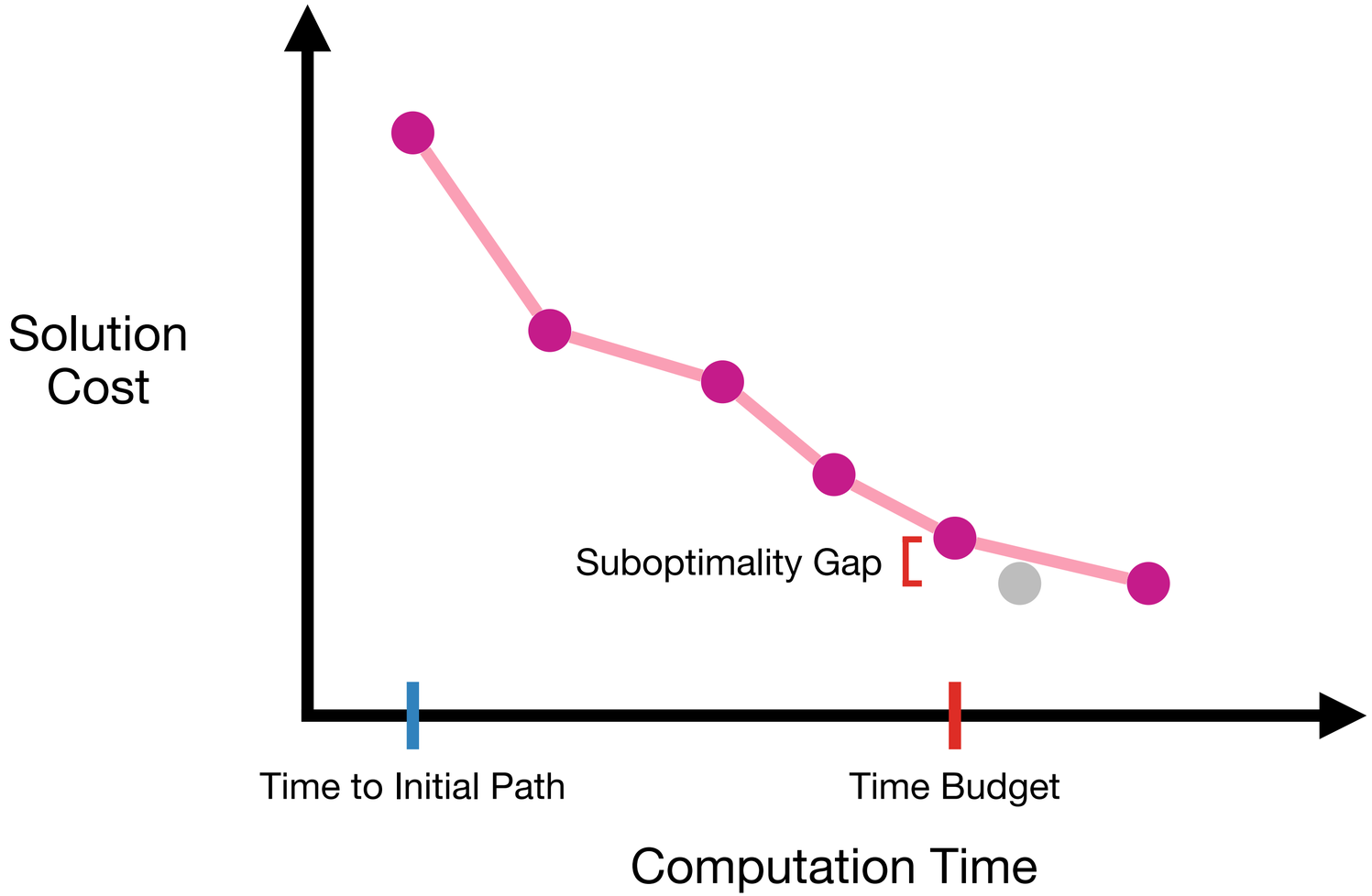
Beyond Asymptotic Optimality





Beyond Asymptotic Optimality



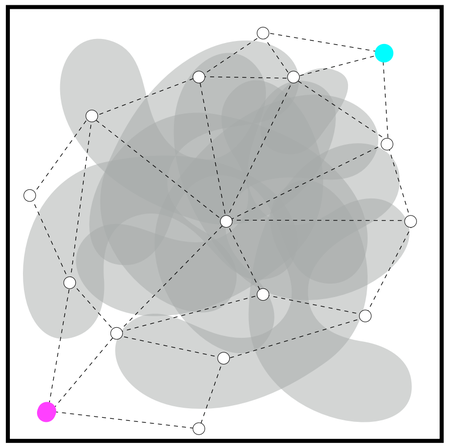


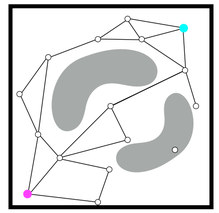
We **formalize anytime search** as Bayesian Reinforcement Learning

Posterior Sampling for Anytime Motion Planning on Graphs with Expensive-to-Evaluate Edges, Hou et al., ICRA 2020.

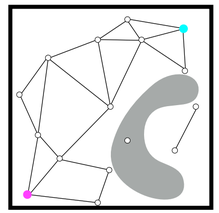
Bayesian Anytime Motion Planning

- Evaluating edges uncovers shorter paths ξ_t
  - **Anytime Objective**: cumulative path lengths $\sum_t w(\xi_t)$
- Given prior on collision statuses $P(\phi)$
  - **Bayesian Anytime Objective**: $\mathbb{E}_{P(\phi)} \sum_t w(\xi_t)$
- Bayesian planning algorithm uses edge evaluation history ψ_t to compute collision posterior $P(\phi|\psi_t)$

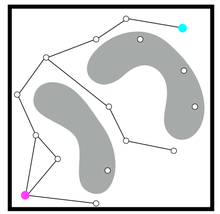




ϕ_1

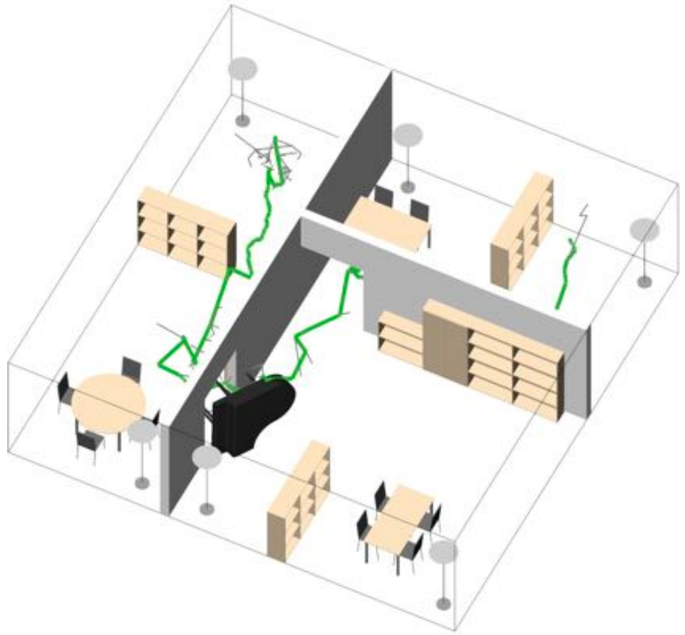


ϕ_2



ϕ_3

The Experienced Piano Movers' Problem



New Piano.
New House.
Same Mover.

Bayesian Anytime Motion Planning as Bayesian Reinforcement Learning

- Equivalence to episodic Bayesian RL [Osband et al, 2013]
- Infer unknown MDP $\mathcal{M}$ through m repeated episodes

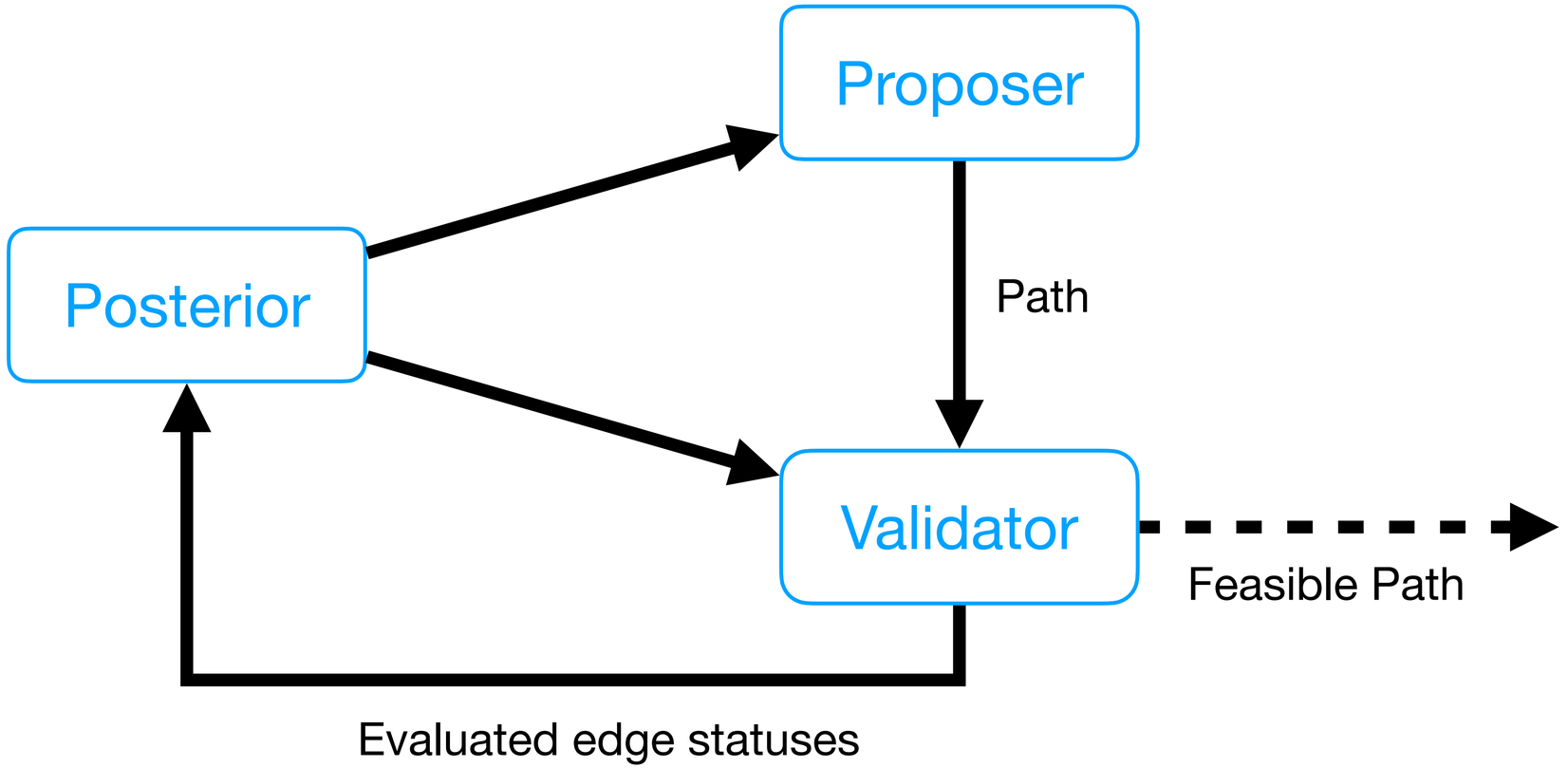
$$\mathbb{E}_{P(\mathcal{M})}\left[\textsc{Regret}(m)\right] = \mathbb{E}_{P(\mathcal{M})}\sum_{k=1}^{m}\left[V(\mu^*) - V(\mu_k)\right] = \mathbb{E}_{P(\phi)}\sum_{k=1}^{m}\left[w(\xi_k) - w(\xi^*)\right]$$

Minimizing Bayesian regret is **equivalent** to
minimizing the Bayesian anytime planning objective!

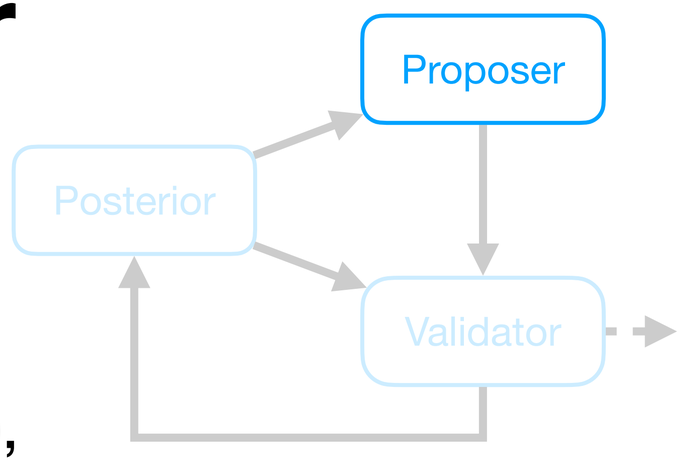
Extending rapidly-exploring random trees for asymptotically optimal anytime motion planning, Abbasi-Yadkori et al., IROS 2010.

"no regret" is equivalent to asymptotic optimality

Experienced Lazy Path Search



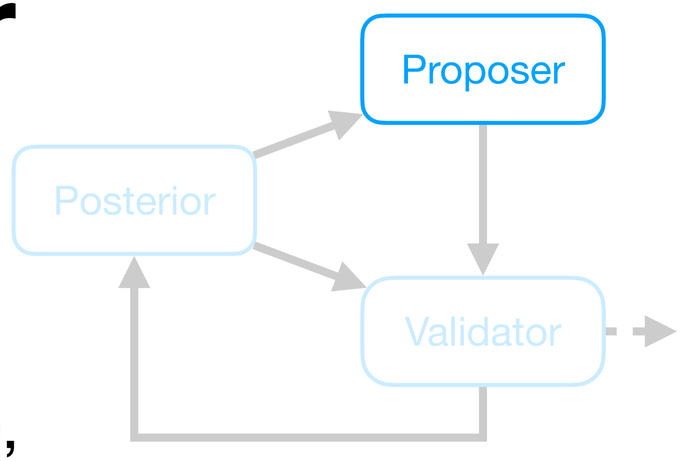
The Posterior Sampling Proposer



- **Posterior Sampling for Motion Planning** (PSMP): propose paths according to probability they are optimal
  - Idea from multi-armed bandits (as Thompson sampling), Posterior Sampling for RL

(More) efficient reinforcement learning via posterior sampling, Osband et al., N*IPS 2013.

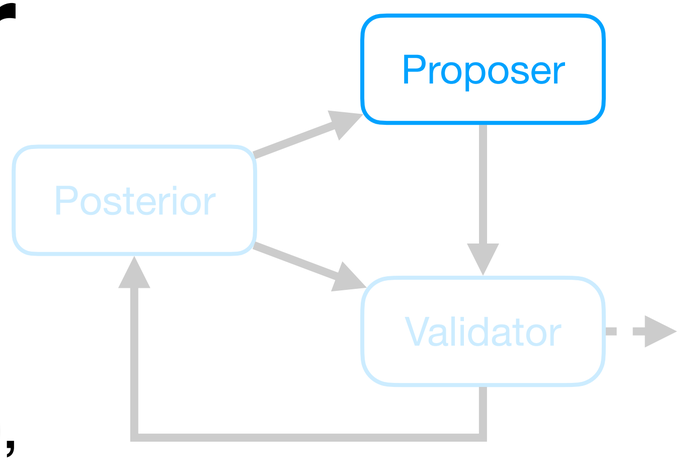
The Posterior Sampling Proposer



- **Posterior Sampling for Motion Planning** (PSMP): propose paths according to probability they are optimal
  - Idea from multi-armed bandits (as Thompson sampling), Posterior Sampling for RL [Osband et al, 2013]
  - First anytime motion planning algorithm with Bayesian regret bounds
    - Analysis adapts [Osband et al, 2013] for deterministic MDPs
    - Bound of $O(\tau\sqrt{|V||E|T\log(|V||E|T)})$ matches known lower bounds

(More) efficient reinforcement learning via posterior sampling, Osband et al., N*IPS 2013.

The Posterior Sampling Proposer

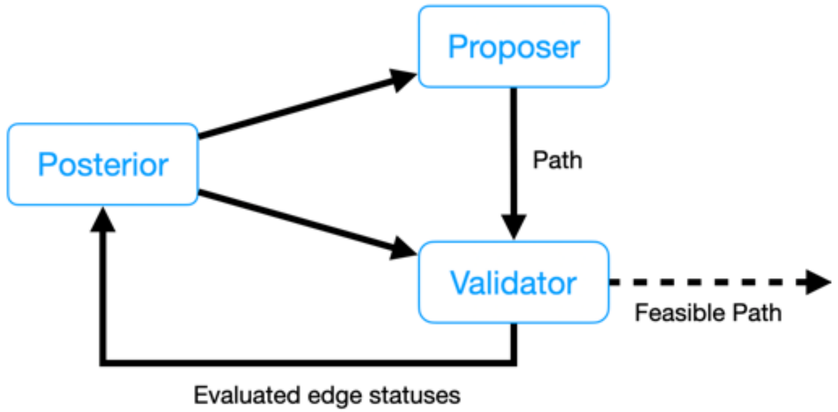


- **Posterior Sampling for Motion Planning** (PSMP): propose paths according to probability they are optimal
  - Idea from multi-armed bandits (as Thompson sampling), Posterior Sampling for RL [Osband et al, 2013]
  - First anytime motion planning algorithm with Bayesian regret bounds
    - Analysis adapts [Osband et al, 2013] for deterministic MDPs
    - Bound of $O(\tau\sqrt{|V||E|T\log(|V||E|T)})$ matches known lower bounds
  - Solves one shortest path problem per proposal

(More) efficient reinforcement learning via posterior sampling,
Osband et al., N*IPS 2013.

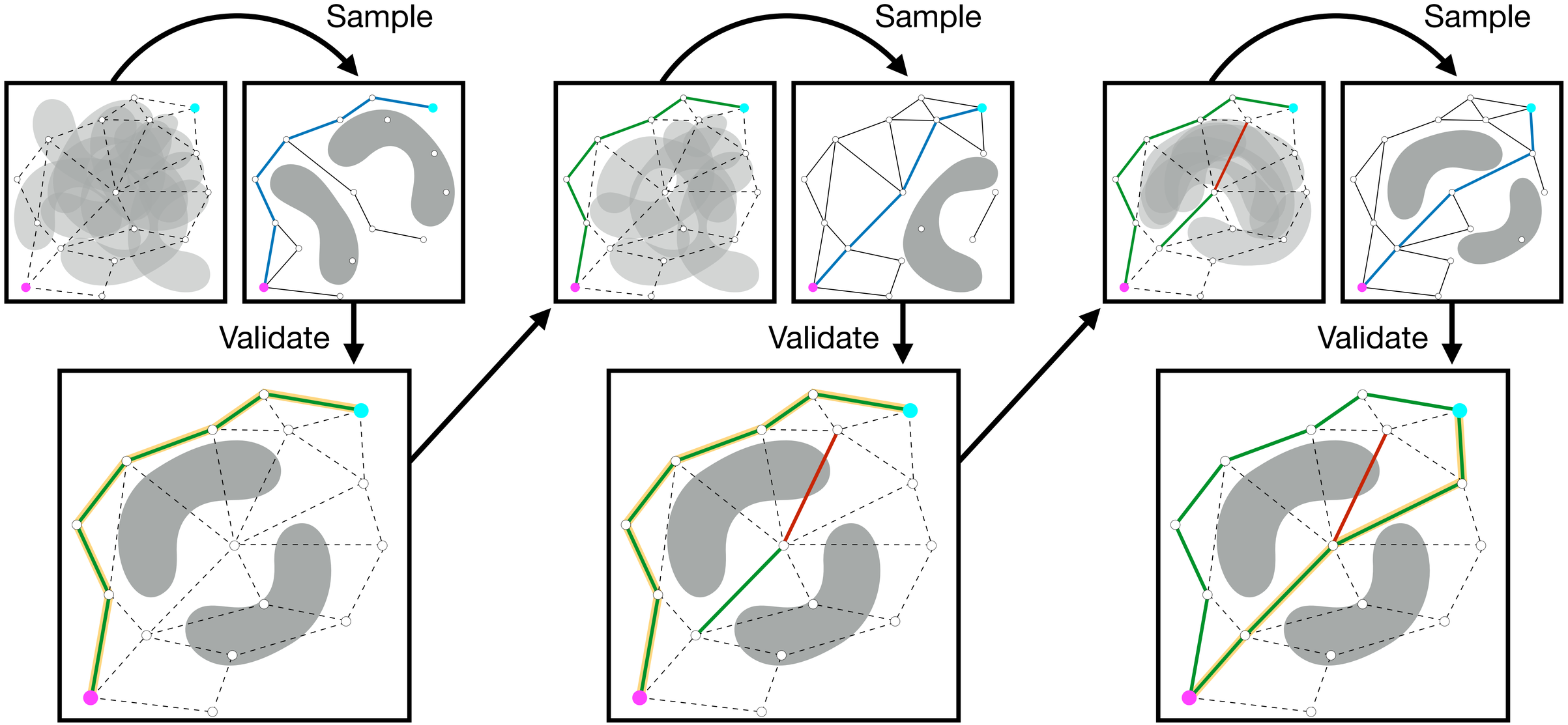
But Whatever Happened to Optimism?!

Optimism in the Face of Uncertainty (OFU) — Shortest Path



Bayesian Regret — Anytime Performance

Bayes Optimality — Feasible Path

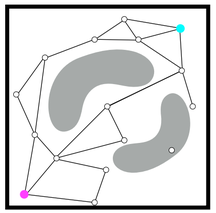
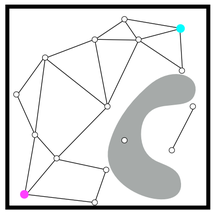
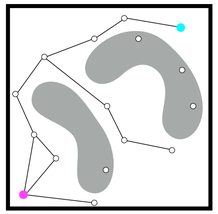

Sample
Sample
Sample
Validate
Validate
Validate

Bayesian Anytime Motion Planning via Posterior Sampling

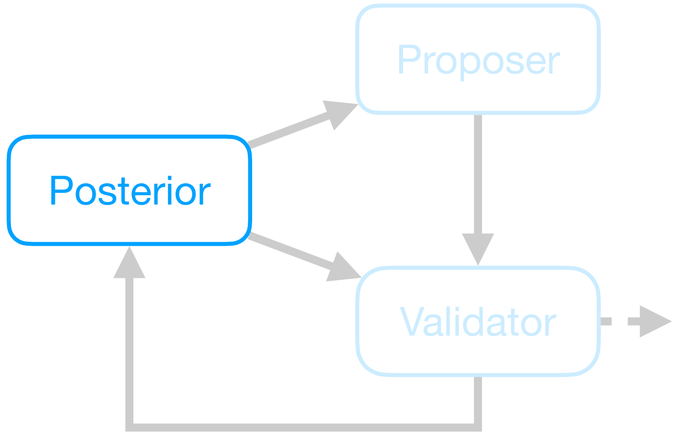
Posterior Distributions

$$\mathbb{E}_{P(\phi)} \sum_t w(\xi_t)$$

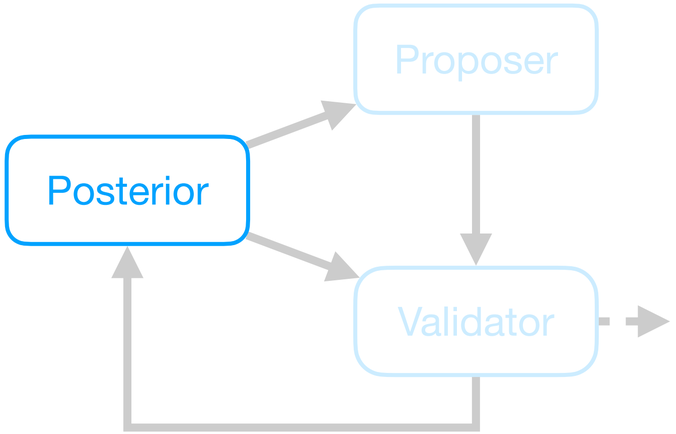
When you have eliminated the impossible, whatever remains, however improbable, must be the truth.

 ϕ_1  ϕ_2  ϕ_3

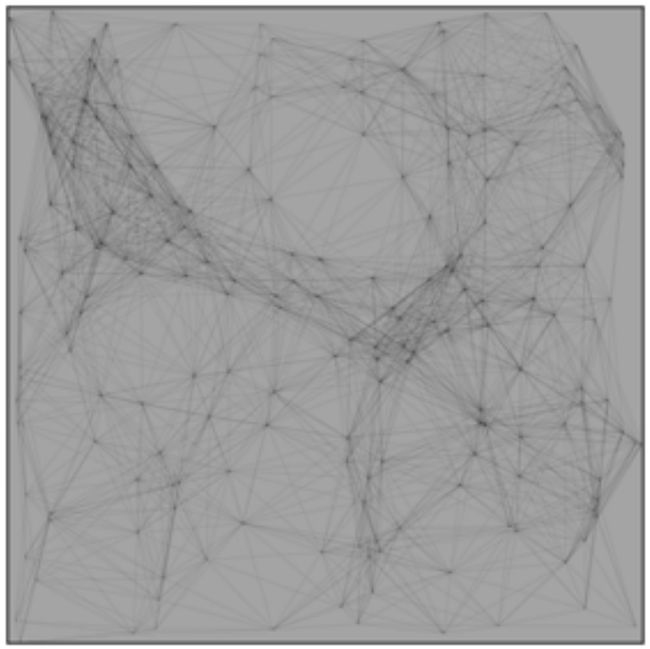
Learning Collision Posteriors



Learning Collision Posteriors

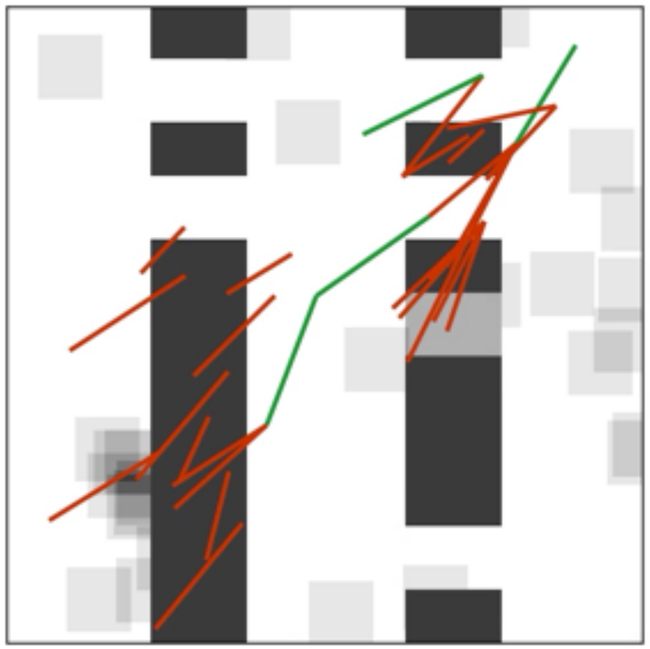


Nearest Neighbor (NN)



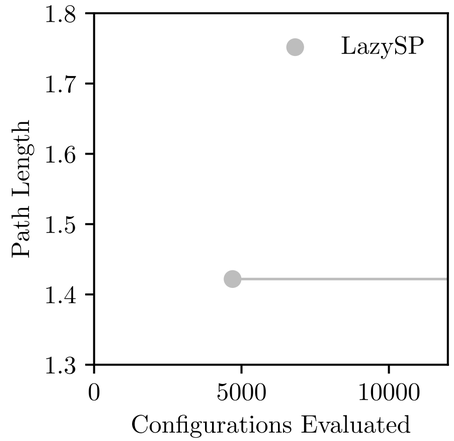
new environments with unknown structure

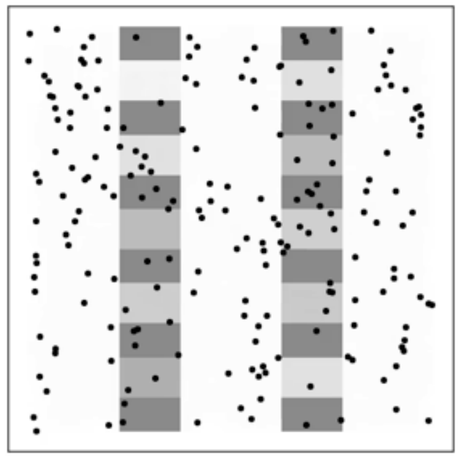
Finite Set (FS)

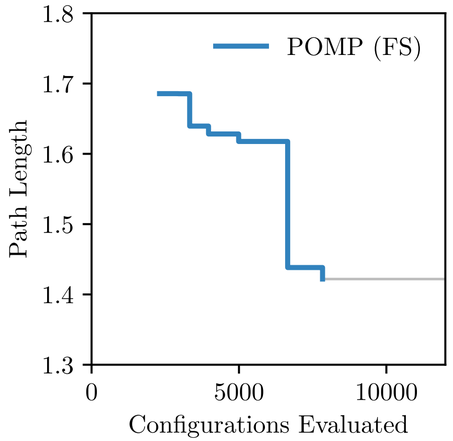


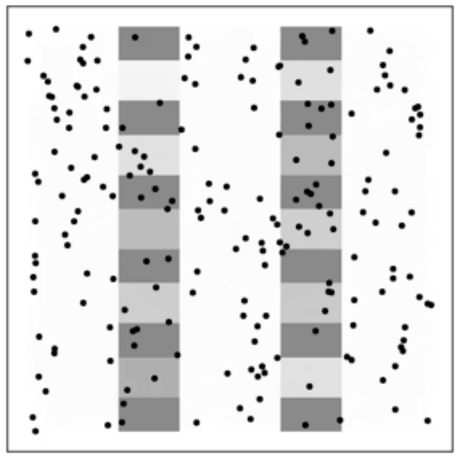
known structure from past experience

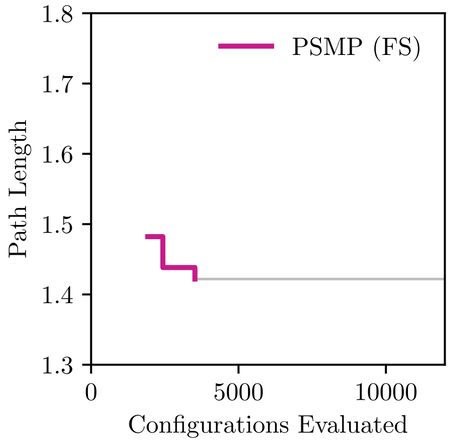
Shorter paths in fewer collision checks

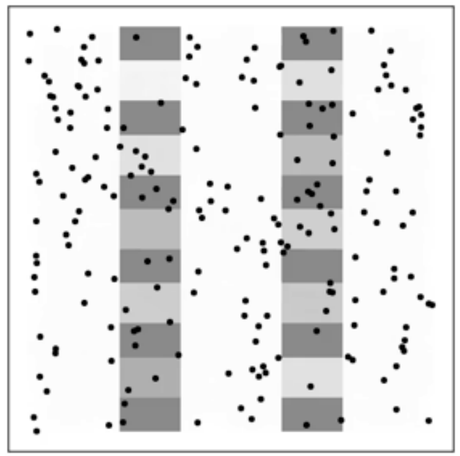






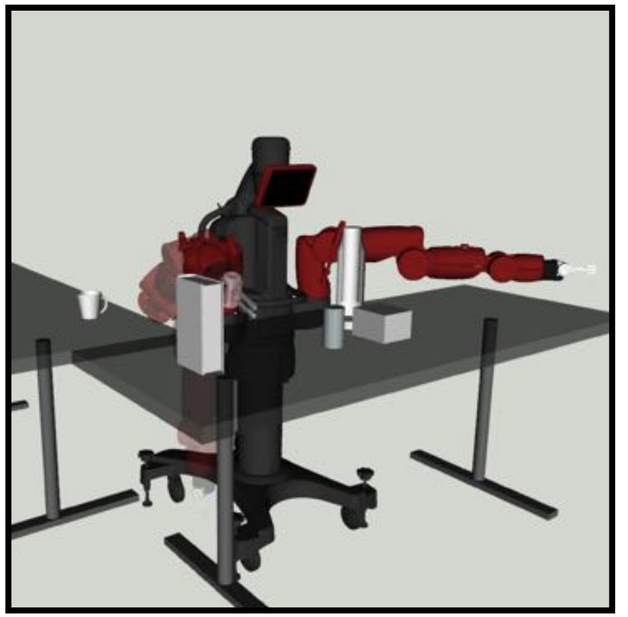
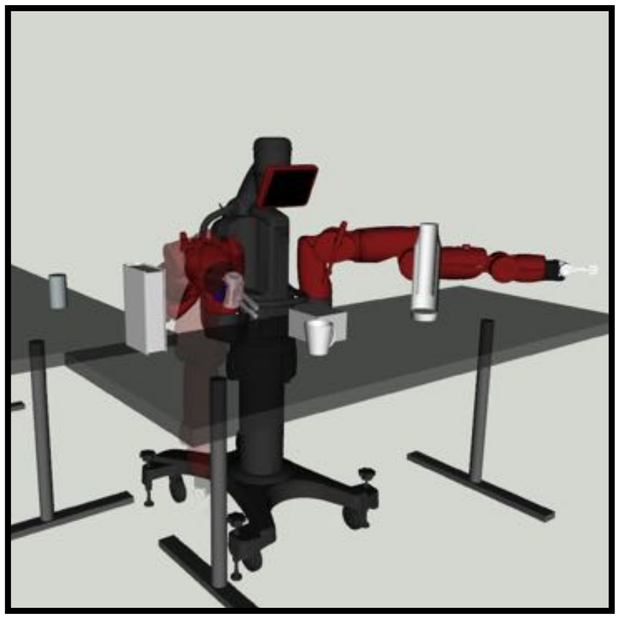






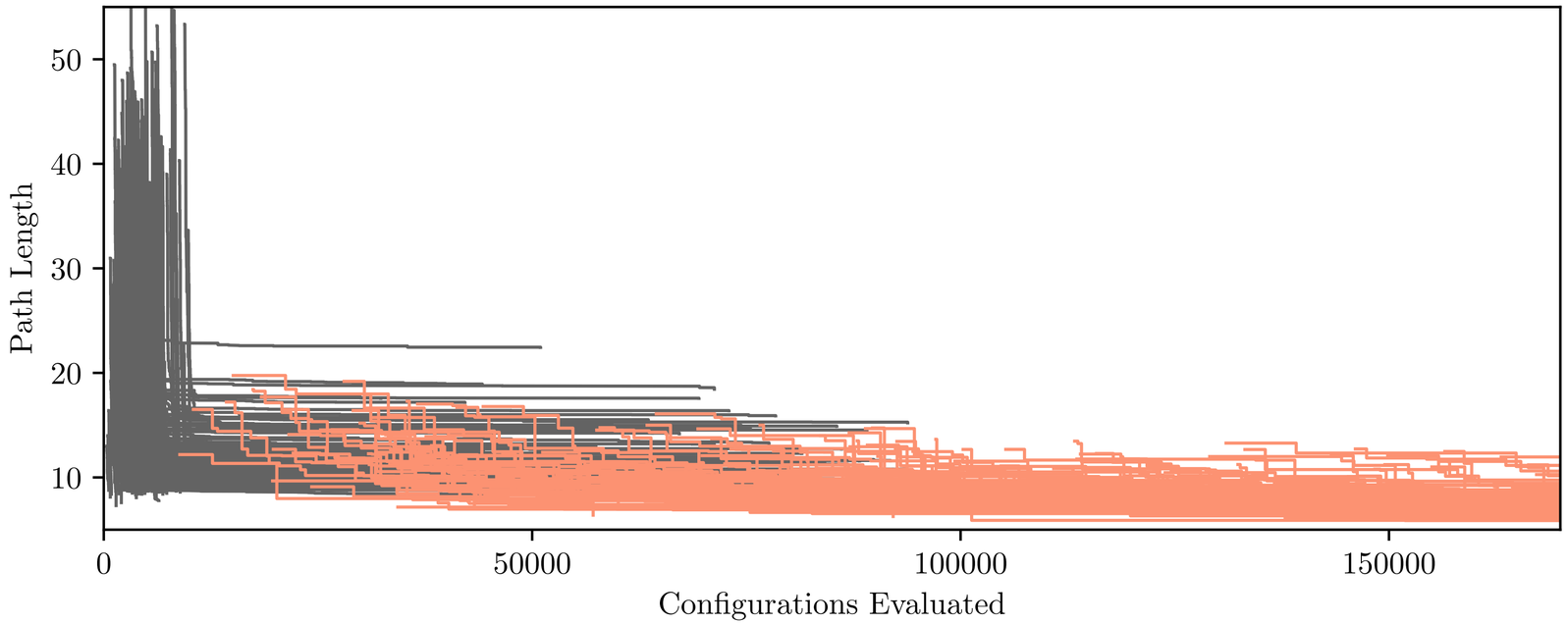
Pareto-Optimal Search over Configuration Space Beliefs for Anytime Motion Planning, Choudhury et al., IROS 2016.

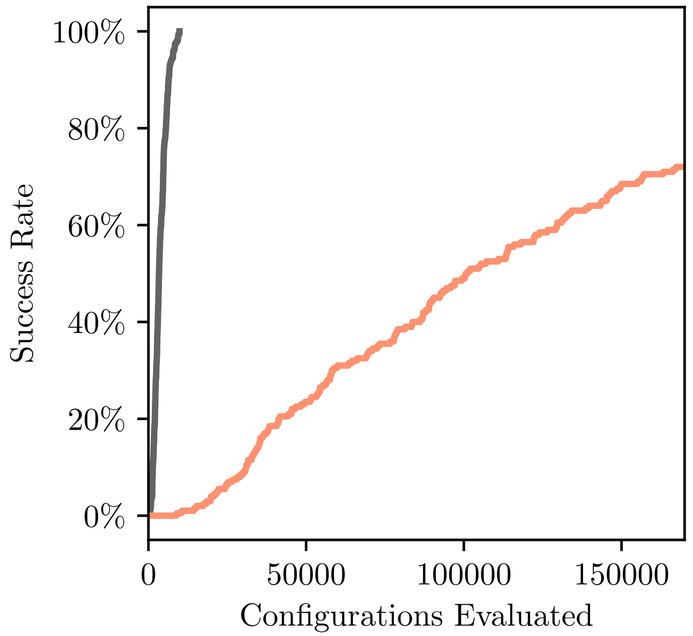
Posterior Sampling for Anytime Motion Planning on Graphs with Expensive-to-Evaluate Edges, Hou et al., ICRA 2020.

Motion Planning Networks, Qureshi et al., ICRA 2019.

RRT* requires many collision checks



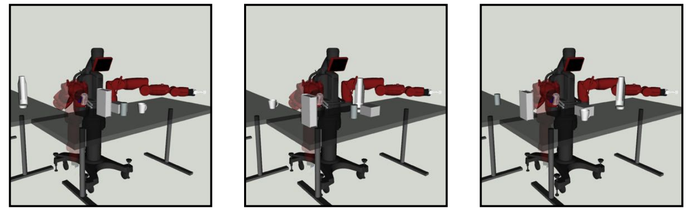


RRT-Connect: An efficient approach to single-query path planning, Kuffner and Lavalle, ICRA 2000.

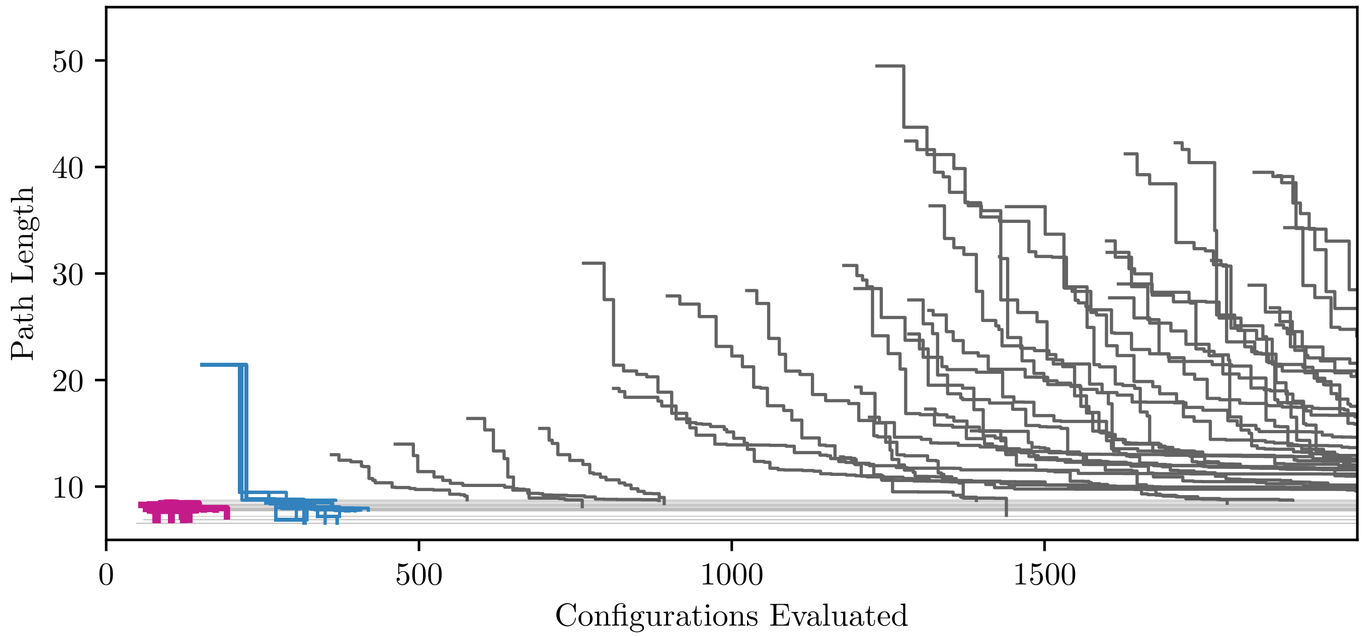
Sampling-based Algorithms for Optimal Motion Planning, Karaman and Frazzoli, IJRR 2011.

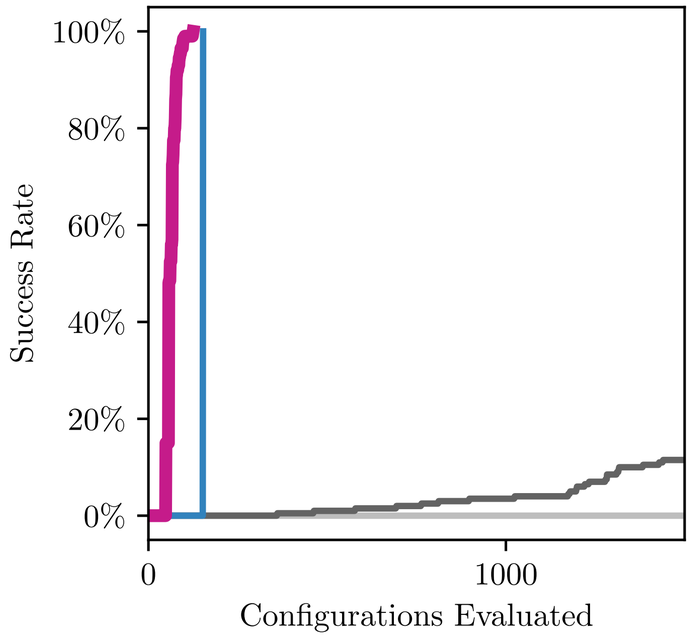
RRT*

RRT + PS



Outperforms common anytime heuristics



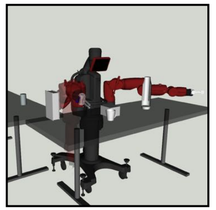


POMP (FS)

PSMP (FS)

LazySP

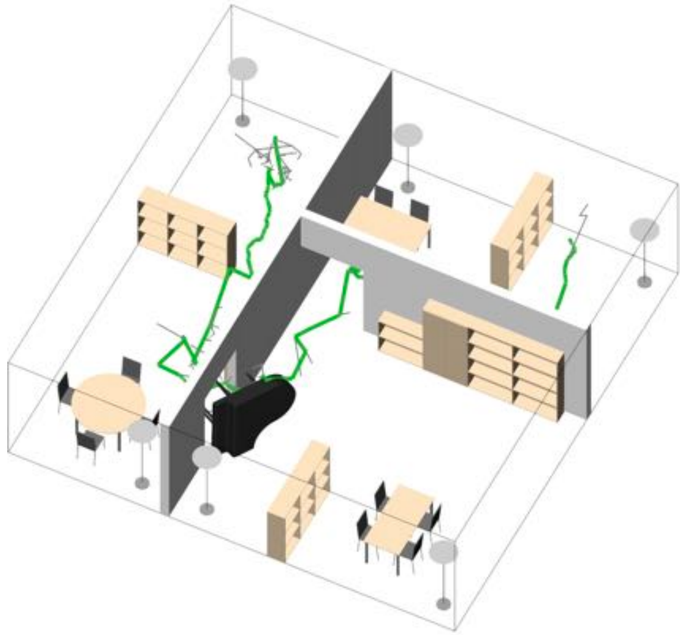
RRT + PS



We **formalize anytime search** as Bayesian Reinforcement Learning

Posterior Sampling for Anytime Motion Planning on Graphs with Expensive-to-Evaluate Edges, Hou et al., ICRA 2020.

The Experienced Piano Movers' Problem



New Piano.
New House.
Same Mover.

Search = Eliminating Paths

Optimal Substructure

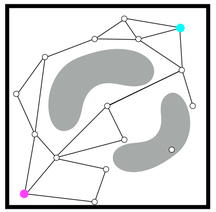
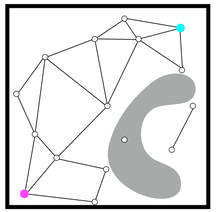
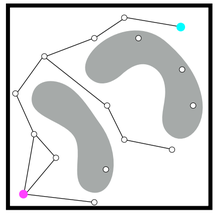
$$f(a) < f(b) \implies f(a \circ x) < f(b \circ x) \forall x$$

You will never catch up.

Posterior Distributions

$$\mathbb{E}_{P(\phi)} \sum_t w(\xi_t)$$

When you have eliminated the impossible, whatever remains, however improbable, must be the truth.

 ϕ_1  ϕ_2  ϕ_3

RL = Eliminating Policies

Optimal Substructure

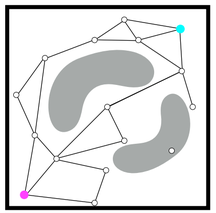
$$f(a) < f(b) \implies f(a \circ x) < f(b \circ x) \forall x$$

You will never catch up.

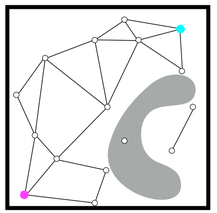
Posterior Distributions

$$\mathbb{E}_{P(\phi)} \sum_t w(\xi_t)$$

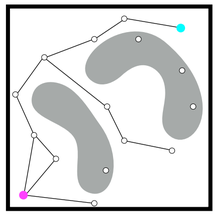
When you have eliminated the impossible, whatever remains, however improbable, must be the truth.



ϕ_1



ϕ_2



ϕ_3

Bayesian Residual Policy Optimization: Scalable Bayesian Reinforcement Learning with Clairvoyant Experts, Lee et al., arXiv:2002.03042

Search for Optimal Solutions: the Heart of Heuristic Search is Still Beating

Ariel Felner
ISE Department
Ben-Gurion University
ISRAEL
felner@bgu.ac.il

1

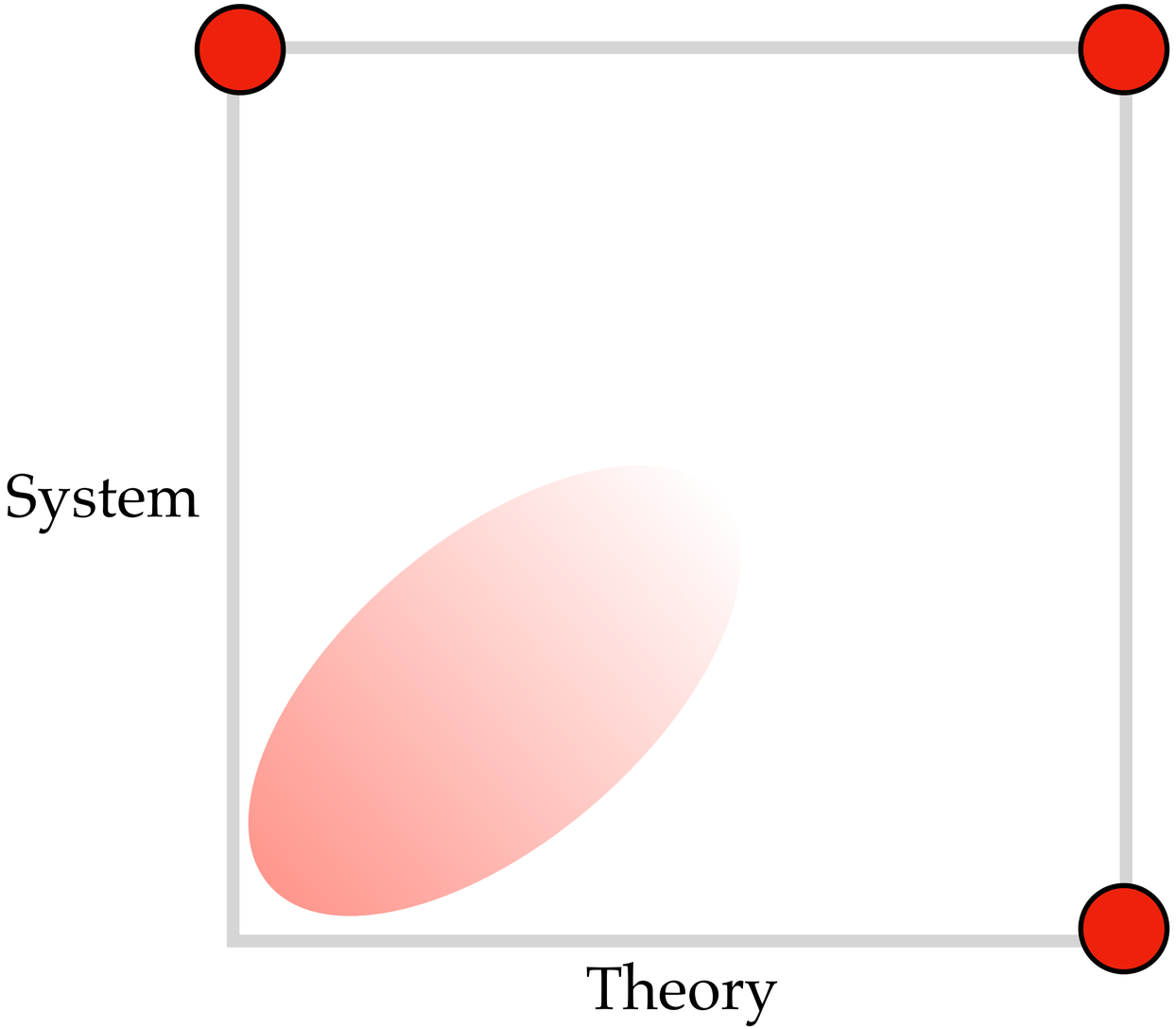
Data-driven Planning via Imitation Learning, Choudhury et al., IJRR 2018.

Exploit Structure

Embrace Laziness

Prove some Damn Theorems

Aim for the Corners!



Pareto-Optimal Search over Configuration Space Beliefs for Anytime Motion Planning, Choudhury et al., IROS 2016.

A Unifying Formalism for Shortest Path Problems with Expensive Edge Evaluations via Lazy Best-First Search over Paths with Edge Selectors, Dellin and Srinivasa, ICAPS 2016.

Near-Optimal Edge Evaluation in Explicit Generalized Binomial Graphs, Choudhury et al., N*IPS 2017.

Bayesian Active Edge Evaluation on Expensive Graphs, Choudhury et al., IJCAI 2018.

The Provable Virtue of Laziness in Motion Planning, Hagtalab et al., ICAPS 2018.

Lazy Receding Horizon A for Efficient Path Planning in Graphs with Expensive-to-Evaluate Edges,* Mandalika et al., ICAPS 2018.

Leveraging experience in lazy search, Bhardwaj et al., RSS 2019.

Generalized Lazy Search for Robot Motion Planning: Interleaving Search and Edge Evaluation via Event-based Toggles, Mandalika et al., ICAPS 2019.

Posterior Sampling for Anytime Motion Planning on Graphs with Expensive-to-Evaluate Edges, Hou et al., ICRA 2020.

Bayesian Residual Policy Optimization: Scalable Bayesian Reinforcement Learning with Clairvoyant Experts, Lee et al., arXiv: 2002.03042

Coauthors

Mohan Bhardwaj, Byron Boots, Sushman Choudhury, Sanjiban Choudhury, Chris Dellin, Nika Hagtalab, Brian Hou, Shervin Javdani, Gilwoo Lee, Simon Mackenzie, Aditya Mandalika, Ariel Procaccia, Oren Salzman, Sebastian Scherer.

Collaborators

Tim Barfoot, Dmitry Berenson, Jon Gammell, David Hsu, Brad Saund, Rahul Vernwal.

Smarty Pants

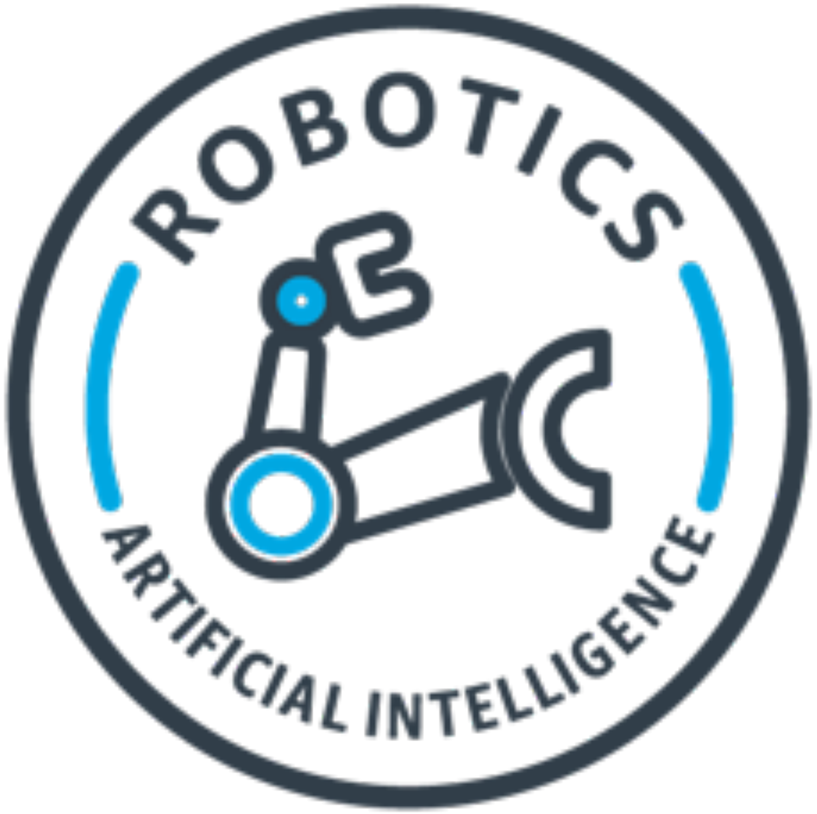
Drew Bagnell, Kostas Bekris, Dan Halperin, Kris Hauser, Sven Koenig, Max Likhachev.

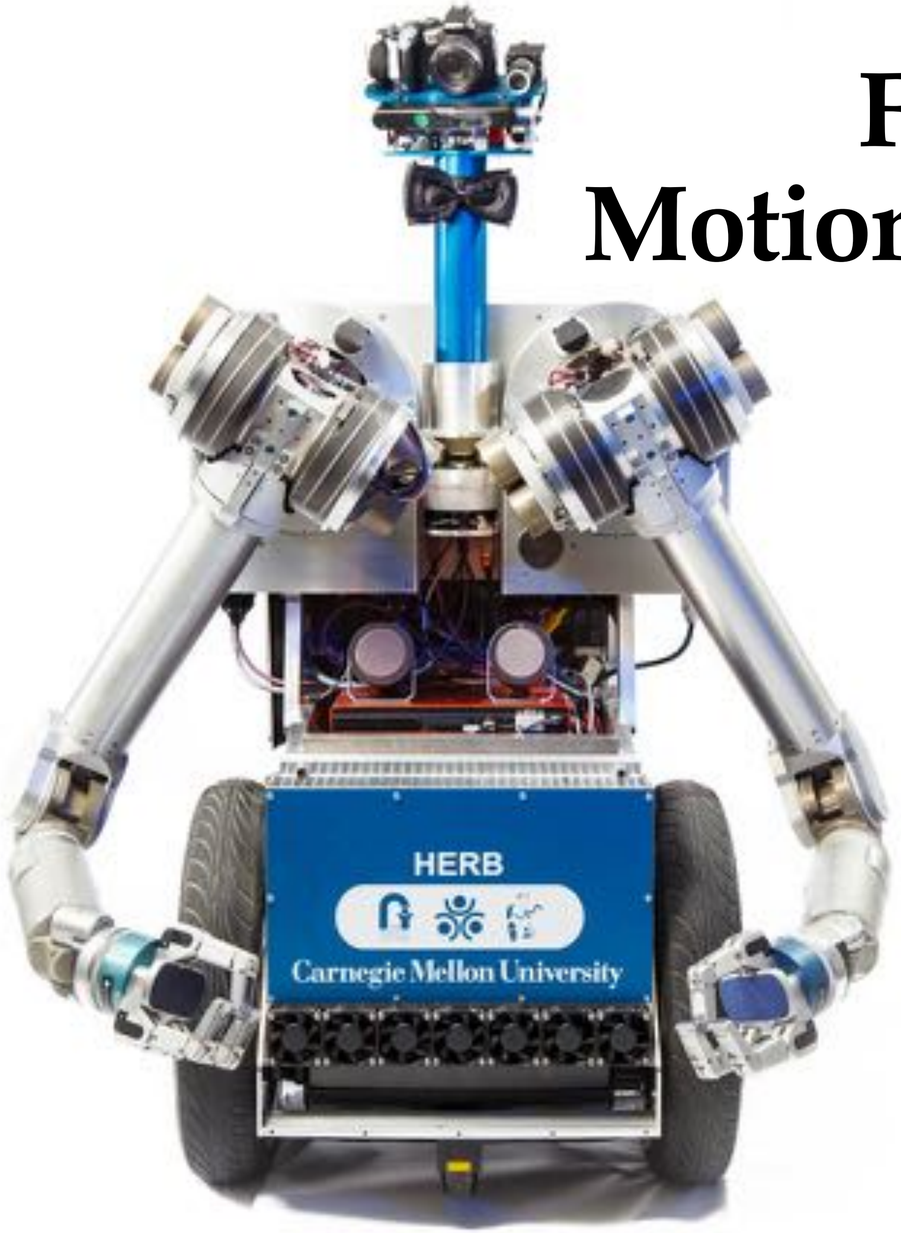
Funders

Army, DARPA, Honda, NIH, NSF, ONR.

https://personalrobotics.cs.washington.edu/publications/

https://www.amazon.jobs/en/teams/rai





Formalizing Connections Between Motion Planning and Machine Learning

Siddhartha Srinivasa
Boeing Endowed Professor
University of Washington

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING