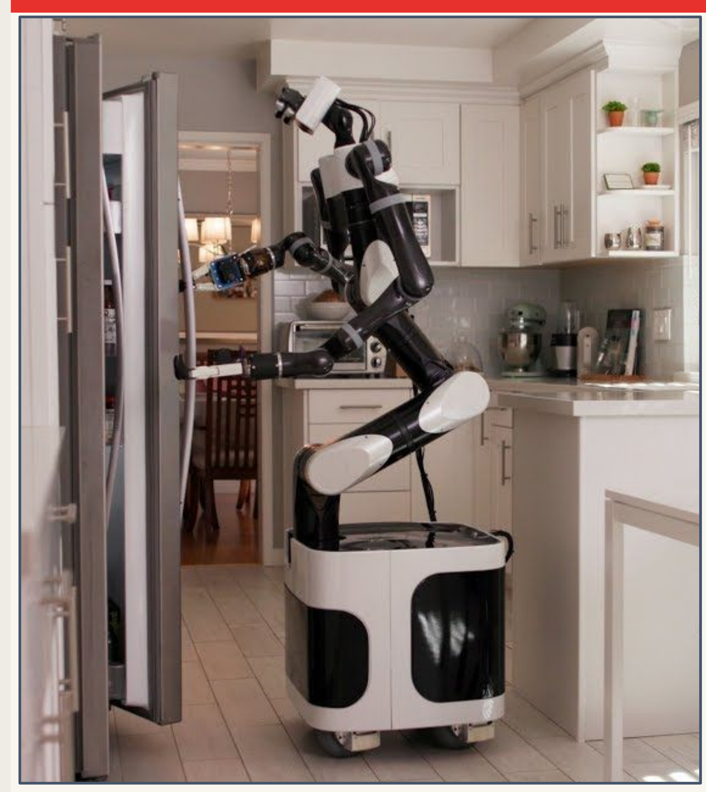
Information Seeking Macro-Actions for POMDPs

BROWN

Max Merlin | Thomas Ottaway | Vadim Kudlay | Calvin Luo

Motivation



Q: How can we design robots that can more autonomously understand and interact with the world around them?

A: By learning to handle two critical sources of uncertainty:

- **High-level state uncertainty** from partially observable environments.
- **Low-level observation uncertainty** from noisy or faulty sensor data.

We propose a framework to resolve both levels of uncertainty, enabling planning and interaction in the world.

Problem Definition

High-Level State Uncertainty

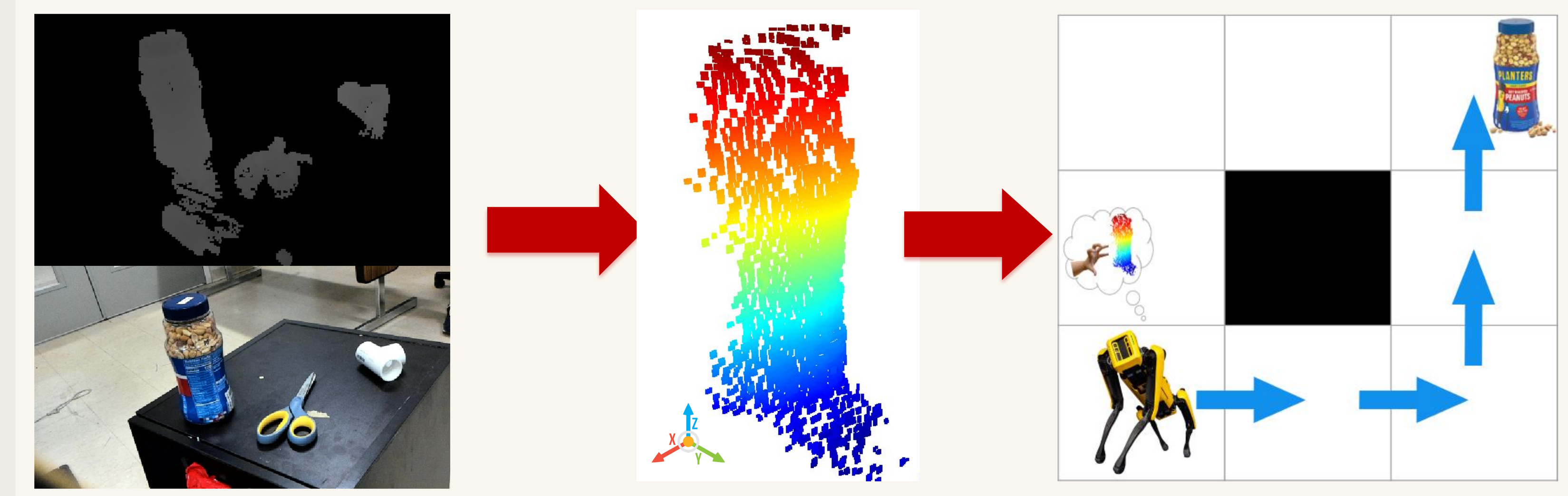
- POMDPs can model all levels of uncertainty together.
- This combined uncertainty increases state complexity and the total number of possible states necessary to simulate.
- Classical POMDPs are therefore intractable in all but the smallest domains.

Low-Level Sensor Uncertainty

- Raw sensor data is often noisy.
- Using such noisy observations propagates uncertainty into the high-level state.
- We would instead like to perform low-level observation confidence estimation and refinement outside of the POMDP model.

Goal

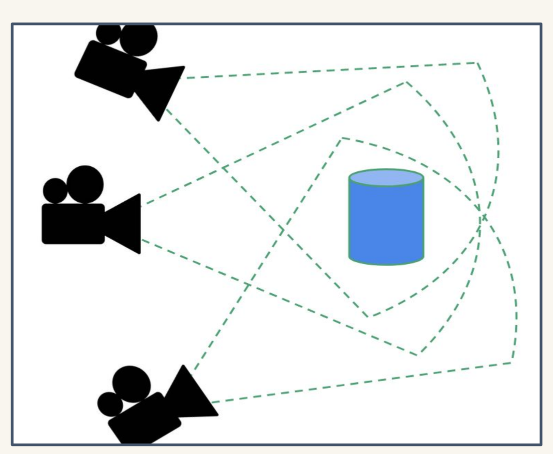
1. Create a skill that identifies the label and pose of a target object.
2. Plan using that skill to achieve a goal in a partially observable domain.



Our Approach

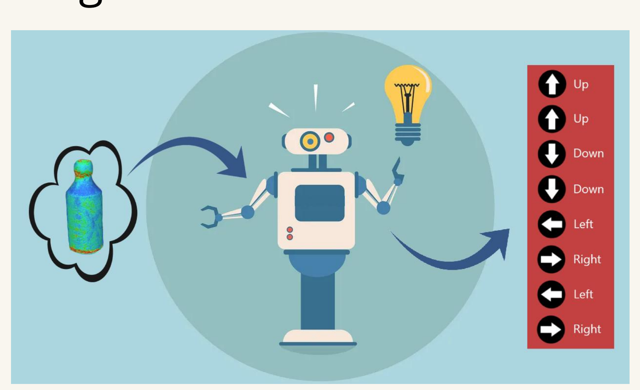
Skill Creation

We create a skill that gets multiple images of a scene local to the robot and calculates the certainty that a desired object is present.



POMDP Planning

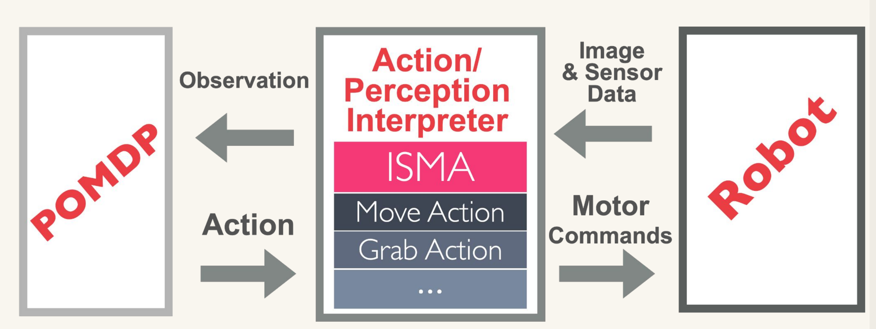
The POMDP receives an **abstract observation** from the skill. This can be treated as a proxy for observing the state directly and accurately for use in planning using POMCP.



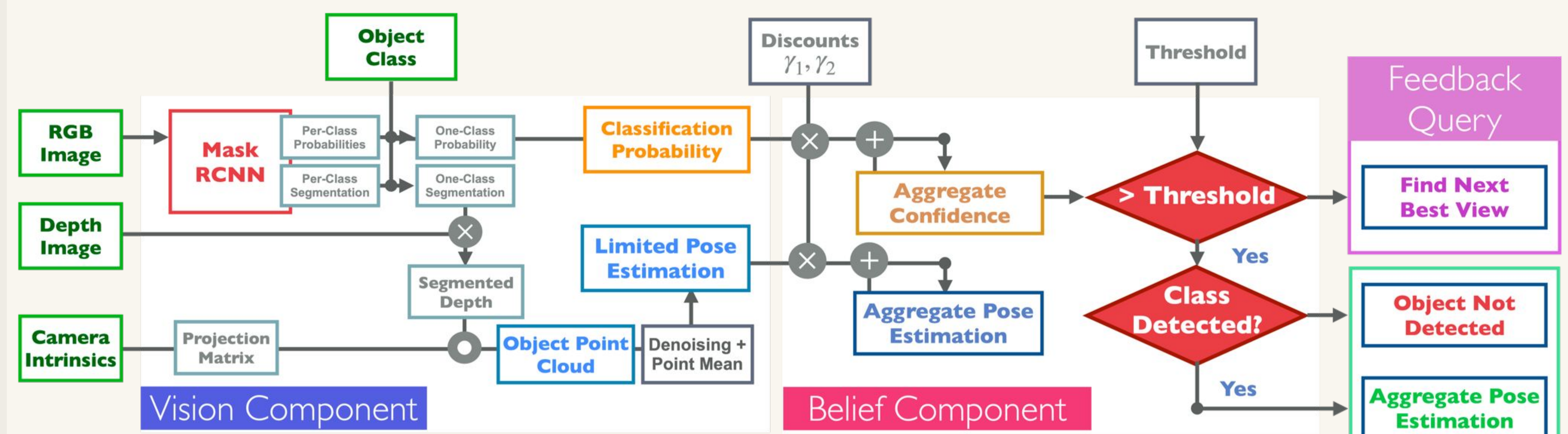
POMDP to Real World

We utilize both state and action abstraction to model our real world domain as a POMDP.

An action perception interpreter handles the low level action execution & sensor data parsing



ISMA Architecture



Information-Seeking Macro-Action **(ISMA)**

Results

Benefits

Providing the POMDP with the ISMA reduced the complexity inducing branching factors along **planning horizon** and **belief update.**

These are two of the vectors which traditionally make POMDPs intractable, so trying to alleviate them is of interest to roboticists.

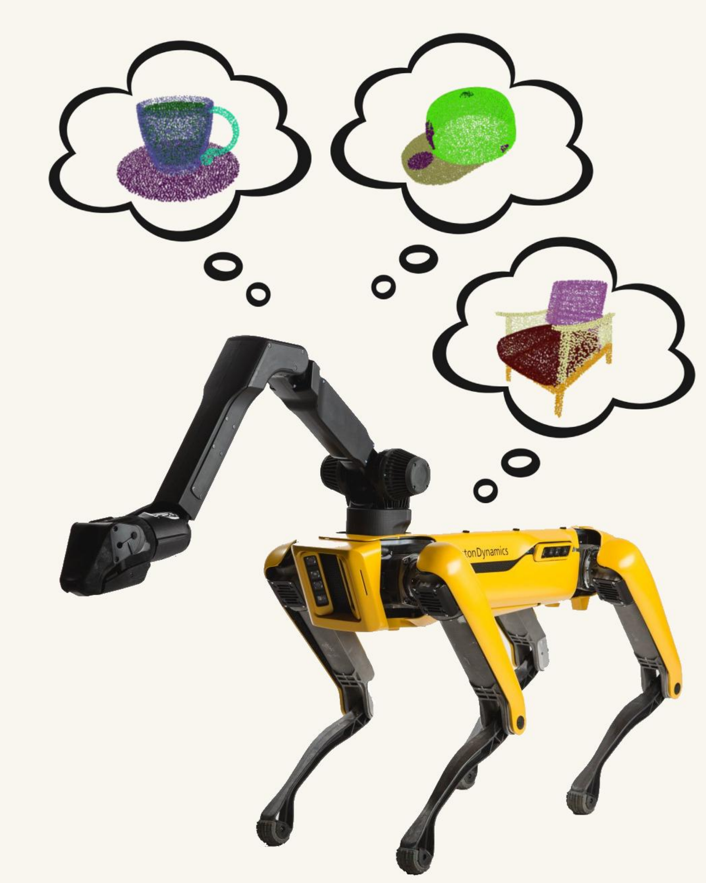
Taking an observation within the Information Seeking Macro-Action

A happy spot shows off its fetched toy

Future Research Directions

Future Work

- Using a more informative Next-Best-View algorithm
- Exhaustive identification loop for all objects in the local scene
- Exploring different base CV algorithms for object detection and pose estimation
- Exploring different methods for aggregating distinct observations
- Integration and deployment into other domains and problem classes



References

[1] Kaelbling, Leslie Pack, Michael L. Littman, and Anthony R. Cassandra. "Planning and acting in partially observable stochastic domains." Artificial intelligence 101.1-2 (1998): 99-134.
[2] Sutton, Richard S., Doina Precup, and Satinder Singh. "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning." Artificial intelligence 112.1-2 (1999): 181-211.

Acknowledgements

We would like to thank Eric Rosen, Kaiyu Zheng, Nicholas DeMarinis, George Konidaris and Boston Dynamics