



# GOAT: GO to Any Thing

Matthew Chang\*, Theophile Gervet\*, Mukul Khanna\*, Sriram Yenamandra\*, Dhruv Shah, Tiffany Min, Kavith Shah, Chris Paxton, Saurabh Gupta, Dhruv Batra, Roozbeh Mottaghi, Jitendra Malik\*, Devendra Singh Chaplot\*



- 1. GOAT Problem**
- 2. GOAT System Architecture**
- 3. Results**
- 4. Applications**
  - a. Pick & Place**
  - b. Social Navigation**
  - c. Platform Agnostic**

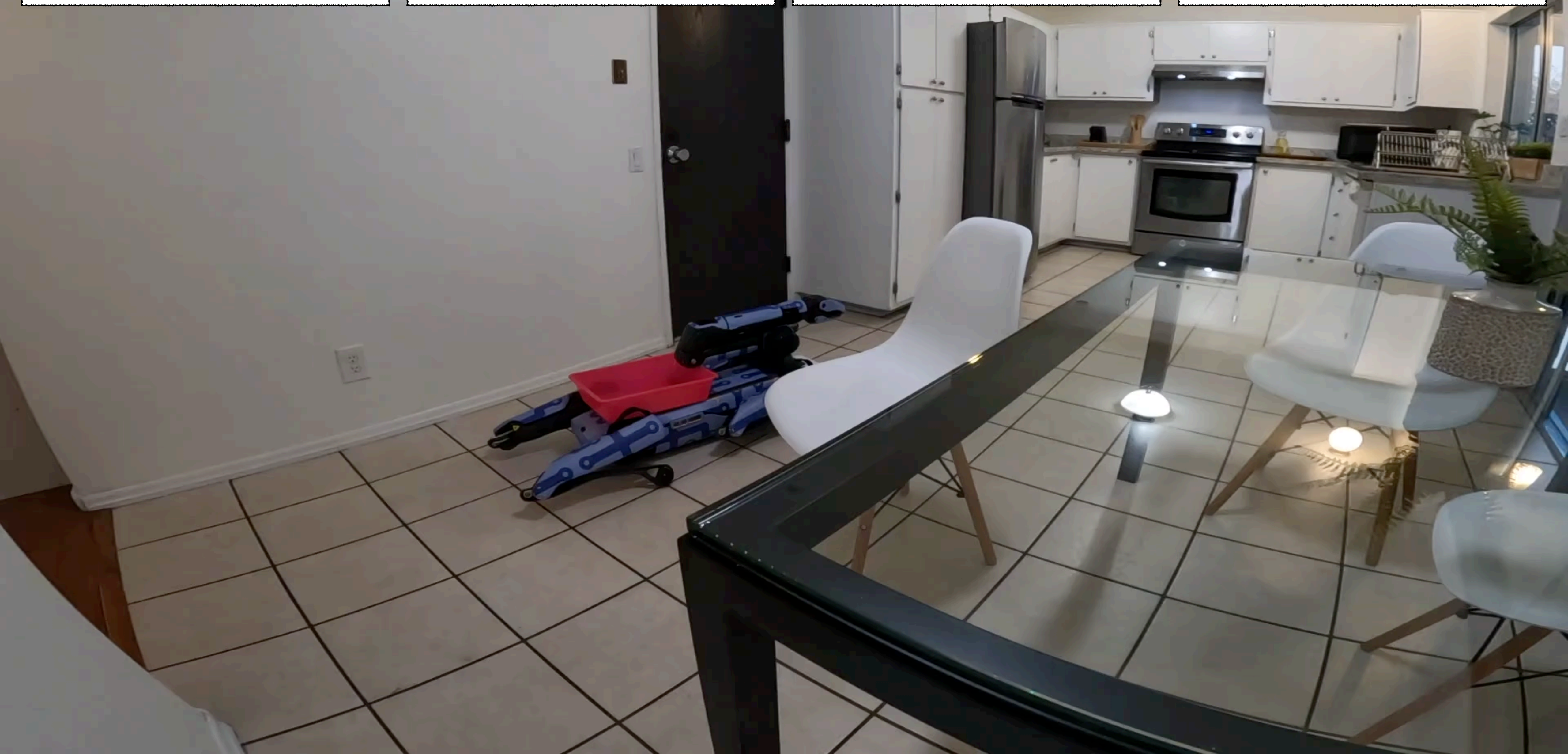


**Unknown Environment**  
*Explore*

**Perception**  
*Detect and Localize  
Objects*

**Lifelong Memory**  
*Remember Object  
Locations*

**Control**  
*Navigate to / Pick & Place  
Objects*





# Multimodal:

*Reach Any Object Specified in Any Way*

Image



Language

Find **the fruit basket on the kitchen counter**

Category

Bring me **a CUP**

# Lifelong:

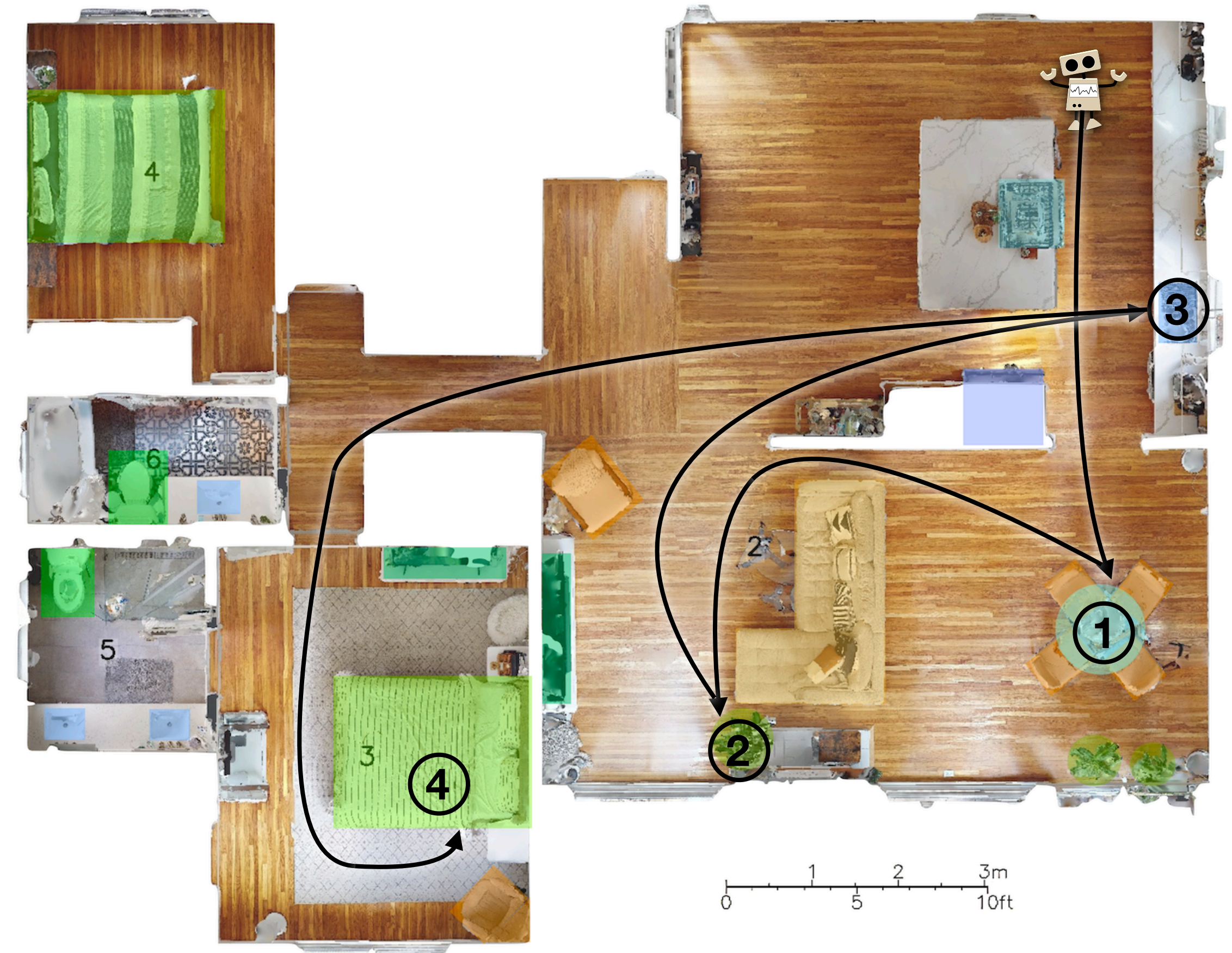
*Remember Object Locations*



② Go to the potted plant next to the couch

③ Go to a SINK

④ Go to the black and white striped bed





1. GOAT Problem
- 2. GOAT System Architecture**
3. Results
4. Applications
  - a. Pick & Place
  - b. Social Navigation
  - c. Platform Agnostic



### Language Goal

*Brown office chair*  
*Bed with white sheets*  
*Cup on the kitchen counter*

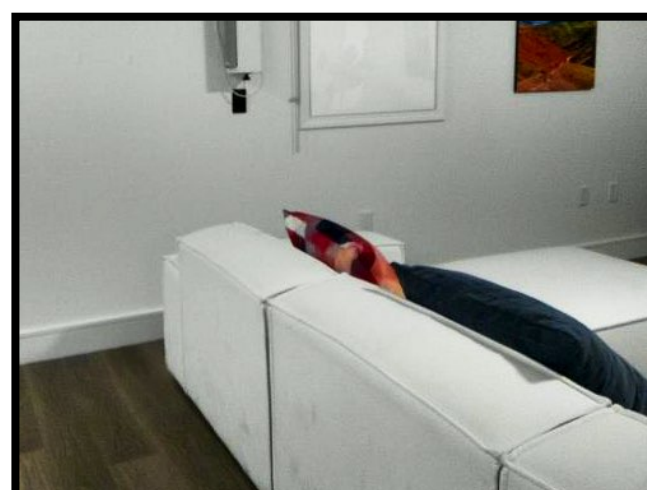
### Image Goal



### Category Goal

*CHAIR*  
*BOOK*  
*CUP*

Sensor Pose Reading ( $x_t$ )



Observation ( $s_t$ )  
(RGB-D)

#### Perception

- Instance Segmentation
- Depth Estimation
- Geometric Projection
- Dynamic Instance Mapping

#### $\pi_G$ Global Policy

#### $\pi_L$ Local Policy

Long-term goal ( $g_t$ )

#### Semantic Map ( $m_t$ )

#### Object Instance Memory ( $o_t$ )

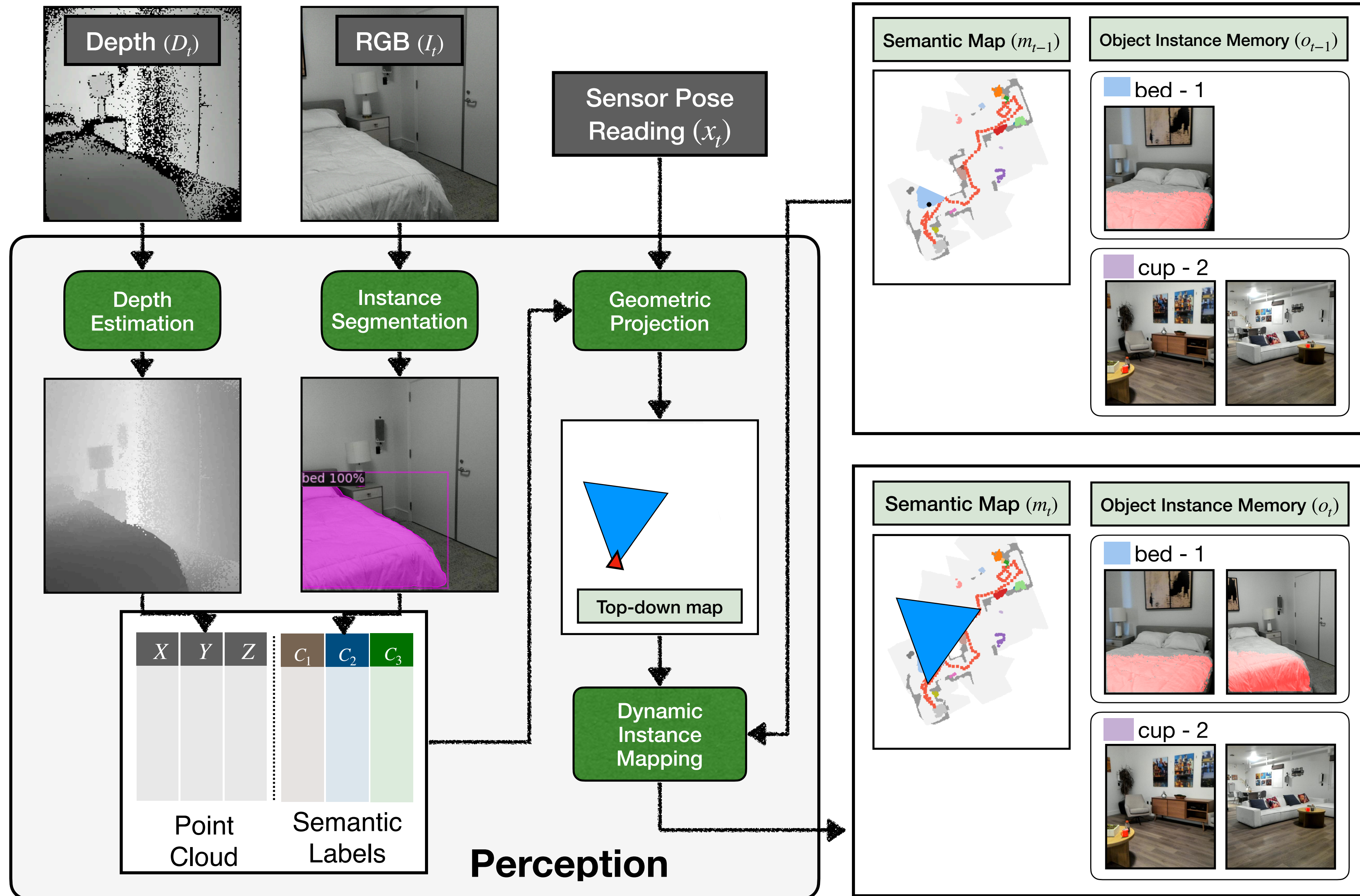
- chair - 3
- couch - 1
- ...

Action ( $a_t$ )

# GOAT: GO to Any Thing



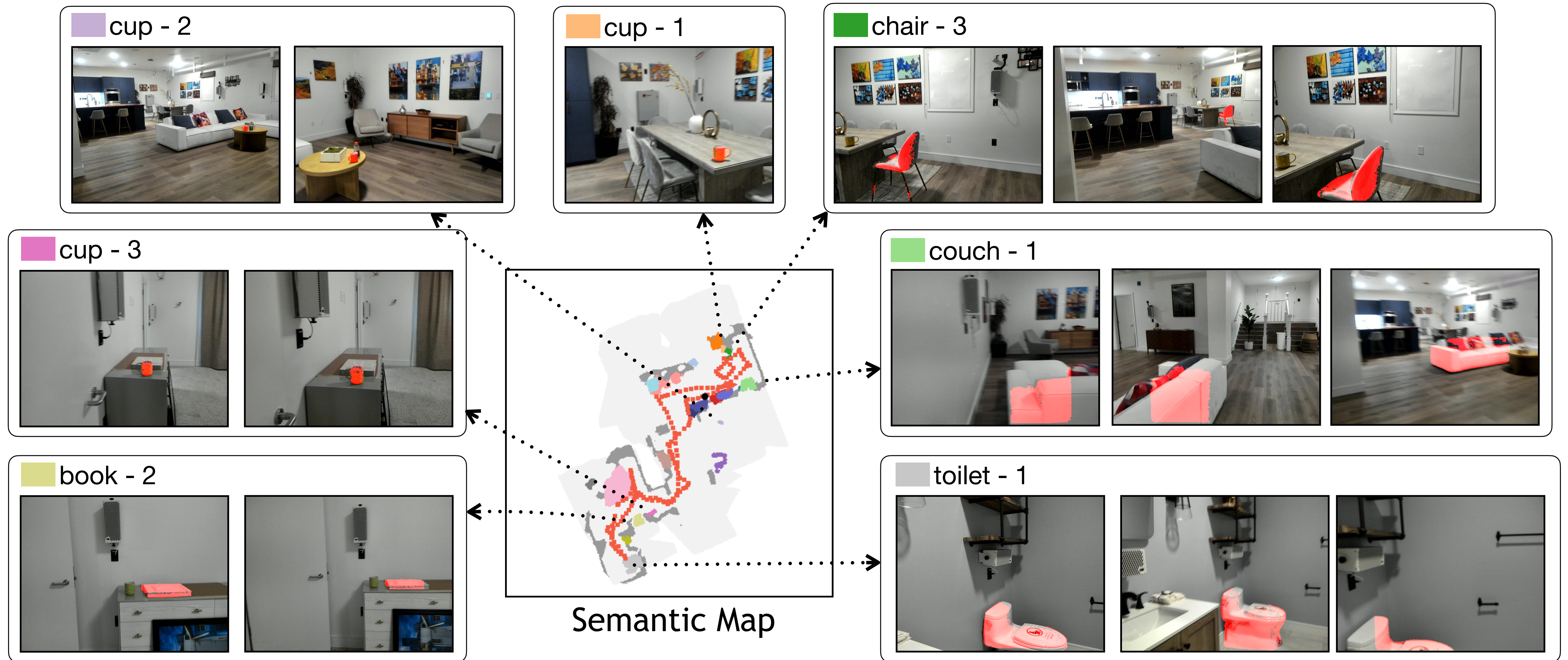
# Perception System





# GOAT Memory Representation

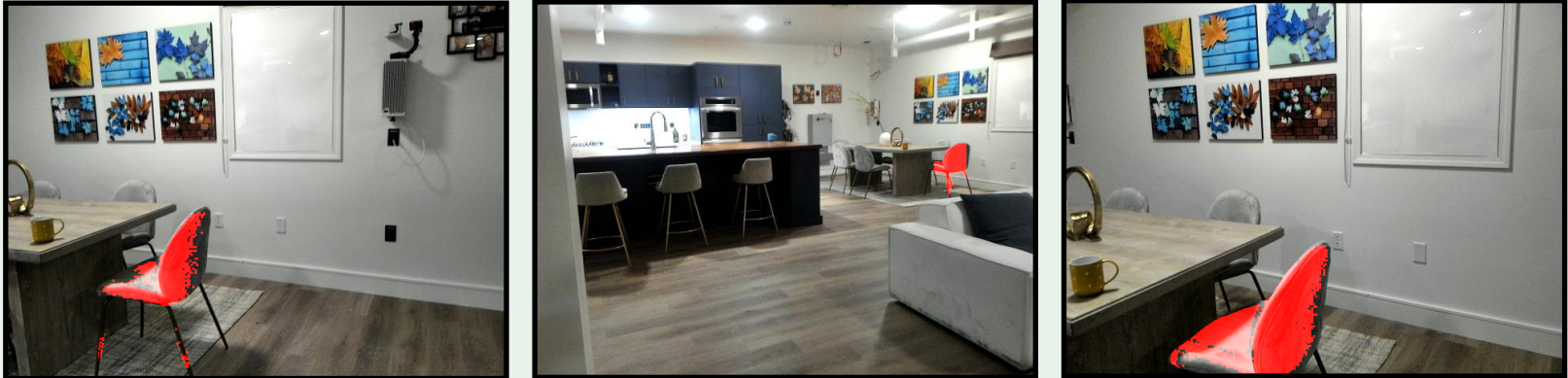
*Semantic map with associated Object Instance memory*






# Object Instance Memory


**chair - 3**




**couch - 1**



**toilet - 1**

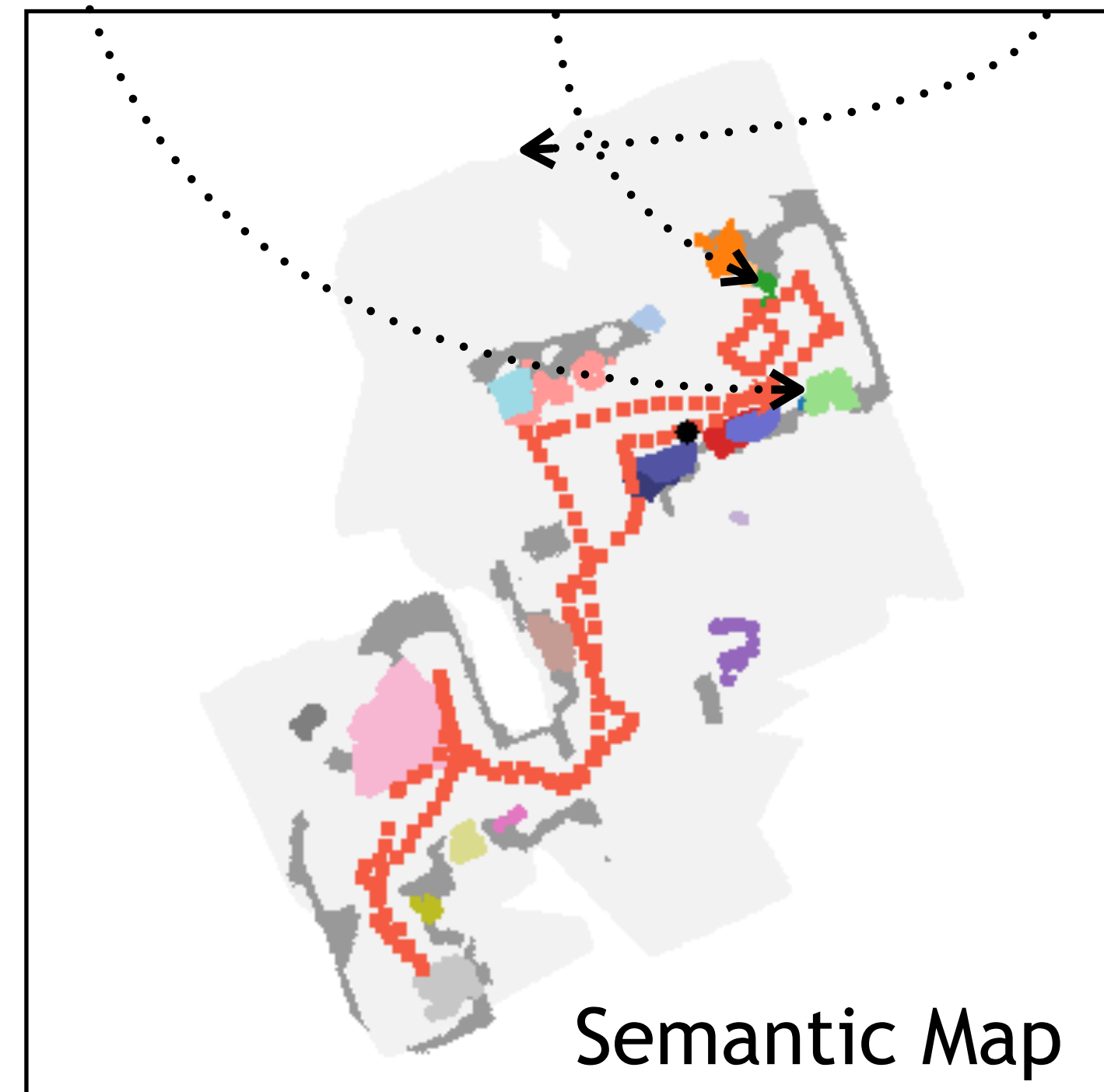
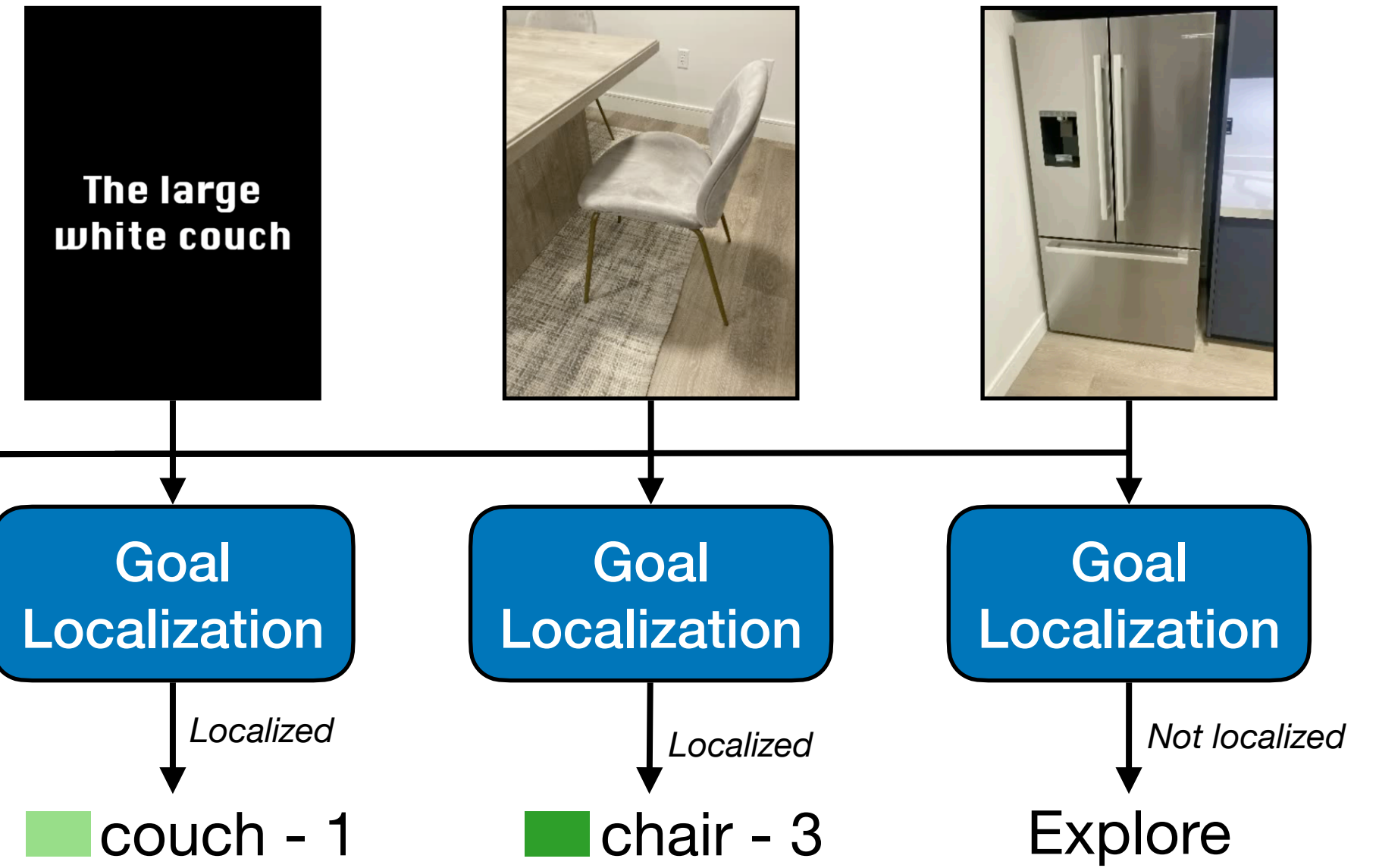


**book - 2**



•  
•  
•

# Goals





1. GOAT Problem
2. GOAT System Architecture
- 3. Results**
4. Applications
  - a. Pick & Place
  - b. Social Navigation
  - c. Platform Agnostic





**“In the Wild”  
Empirical Evaluation**  
*9 Unseen Homes*  
*4 Methods*  
*10 Trajectories per Home*  
*5-10 Goals per Trajectory*  
*~90h of Experiments*

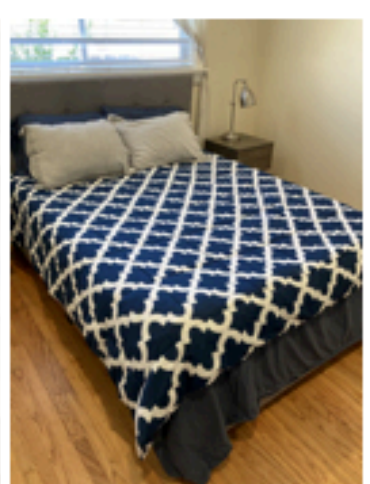




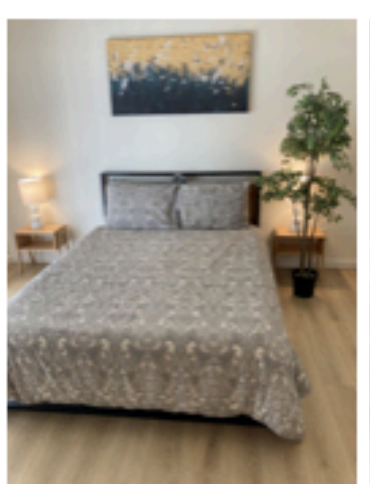
The bed with the white blanket pulled back halfway and grey sheets.



The bunk bed with stars on the blanket.



The bed with blue and white sheets.



The bed with the blue and yellow painting above it and plant on the right.



The bed with the blue blanket and blue pillows.



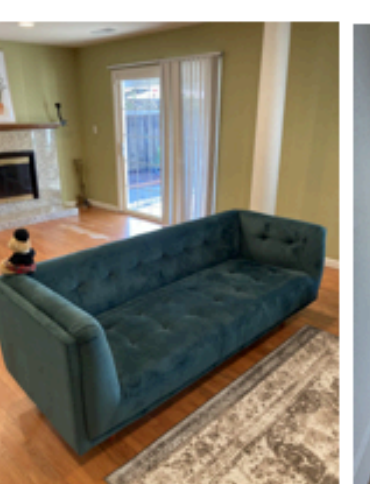
The bed with grey sheets and light green trim.



The bed with the white blanket and red pillows.



The large grey living room couch with many pillows.



The green couch.



The white rectangular couch with no pillows.



The large grey living room couch with many pillows.



The light brown couch with blue pillows.



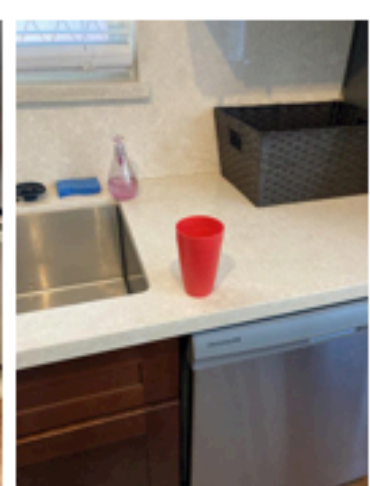
The large grey couch in front of the yellow wall.



The couch covered in a white blanket.



The green cup on the kitchen counter.



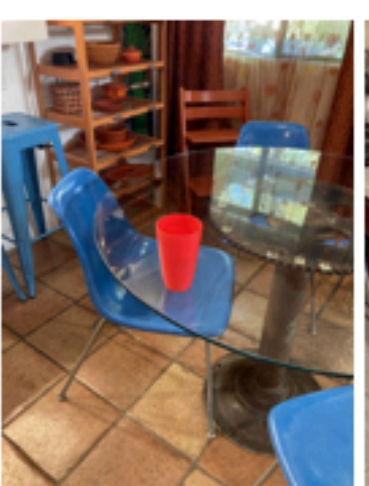
The red cup on the kitchen counter.



The green mug on the plastic chair.



The grey cup on the nightstand.



The red cup on top of the glass table.



The red cup on the kitchen counter.



The light blue cup.



The plant in front of the window.



The potted plant next to the stairs.



The large potted plant in the bedroom.



The group of plants in front of the curtain.



The large potted plant next to the foosball table.



The large potted plant in front of the mirror.



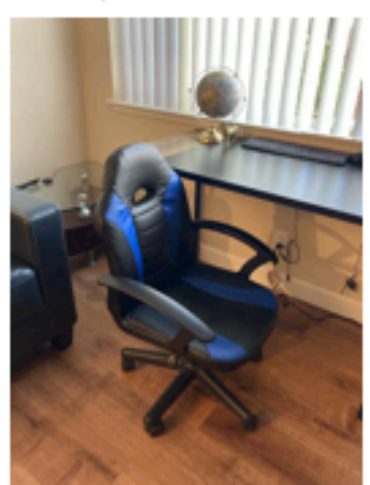
The small potted plant on the hallway table.



The grey dining table chair.



The green lawn chair.



The black office chair.

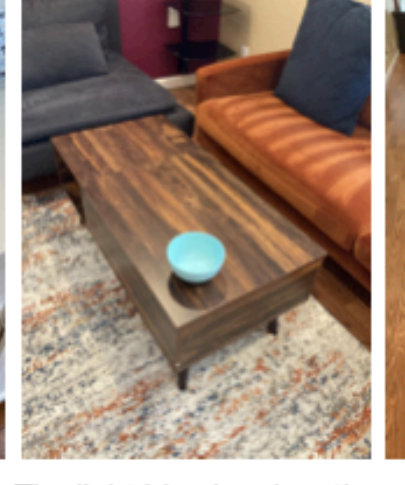


The small wooden chair at the desk.



The black leather chairs in the kitchen.

# "In the Wild" Empirical Evaluation 200+ Object Instances



The light blue bowl on the living room coffee table.



The lavender bowl on the kitchen counter.



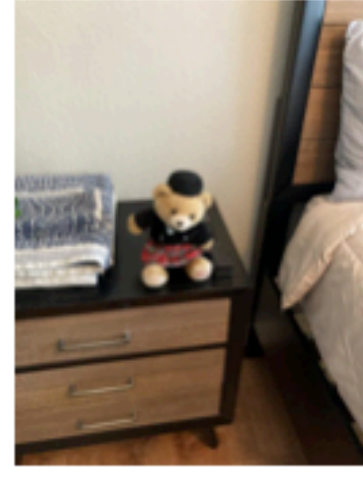
The light blue bowl on the couch.



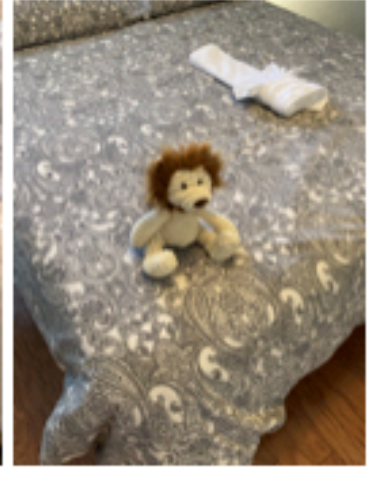
The green cover book on the coffee table.



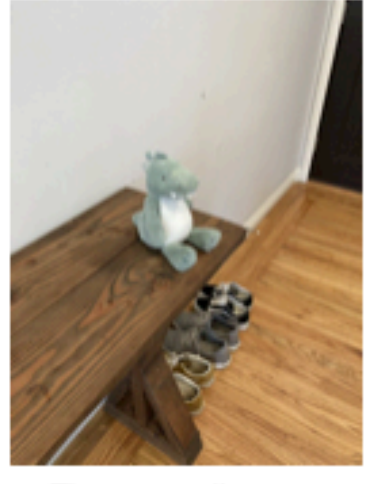
The book on the desk. It has a car on the cover.



The beige teddy bear.



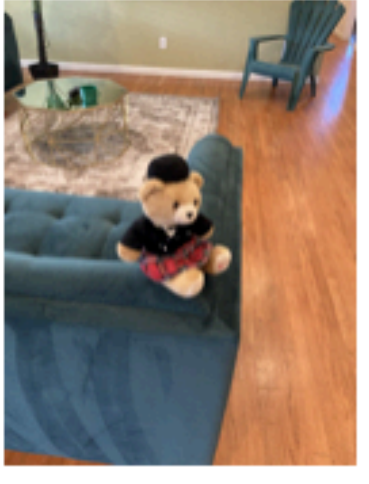
The stuffed lion toy.



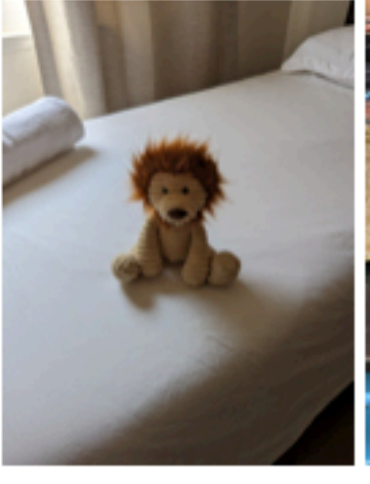
The green dinosaur stuffed toy.



The stuffed lion.



The beige teddy bear.



The stuffed lion toy.



The green dinosaur stuffed toy.



The refrigerator.



The refrigerator.



The refrigerator.



The refrigerator.



The toilet next to the shower curtain with blue fish.



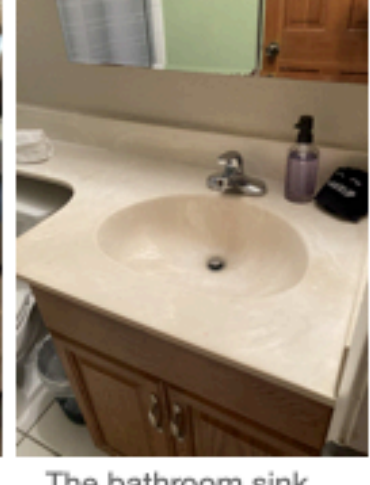
The toilet.



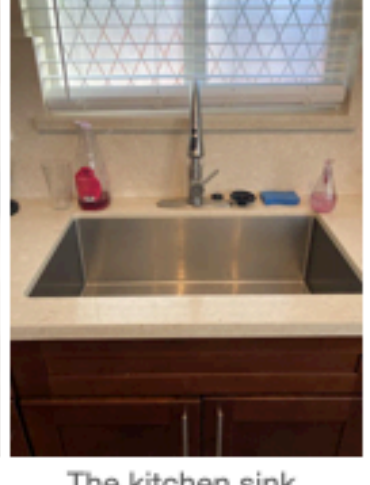
The toilet next to the glass shower door.



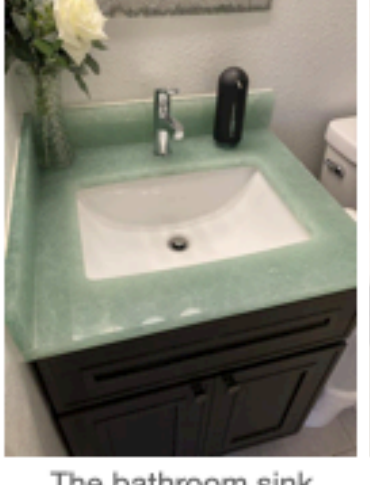
The kitchen sink.



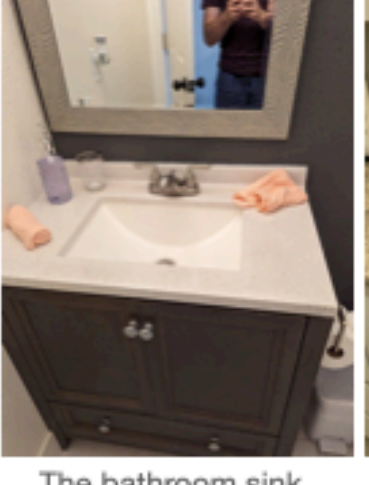
The bathroom sink with marble top.



The kitchen sink.



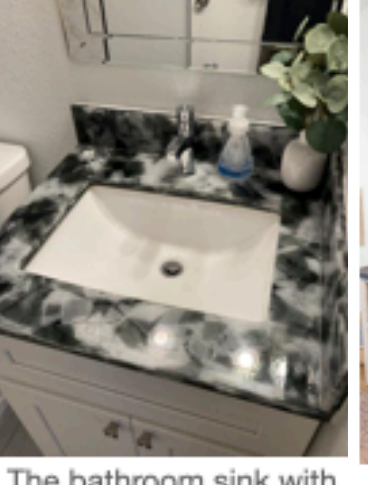
The bathroom sink with green counter.



The bathroom sink.



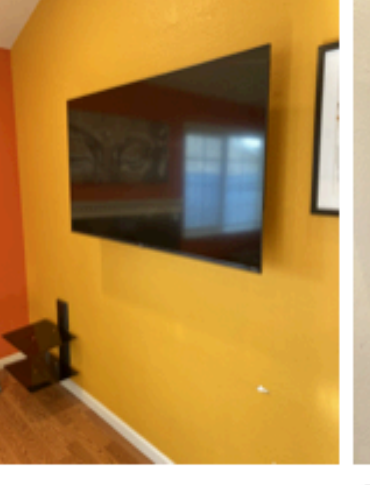
The kitchen sink.



The bathroom sink with black and white counter.



The television.



The television mounted on a yellow wall.



The television mounted on a white wall.



The oven.



The oven.



The oven.



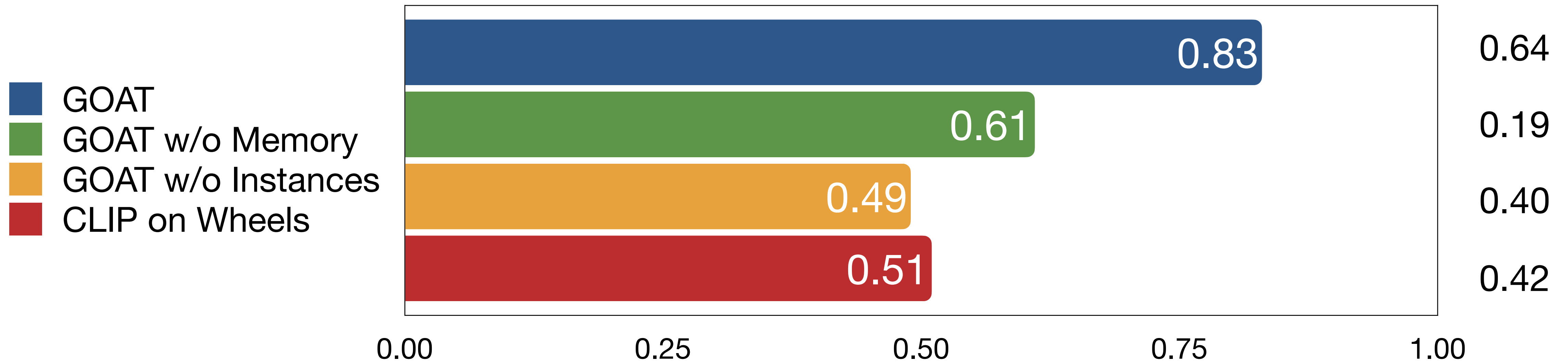
The oven.



# Results

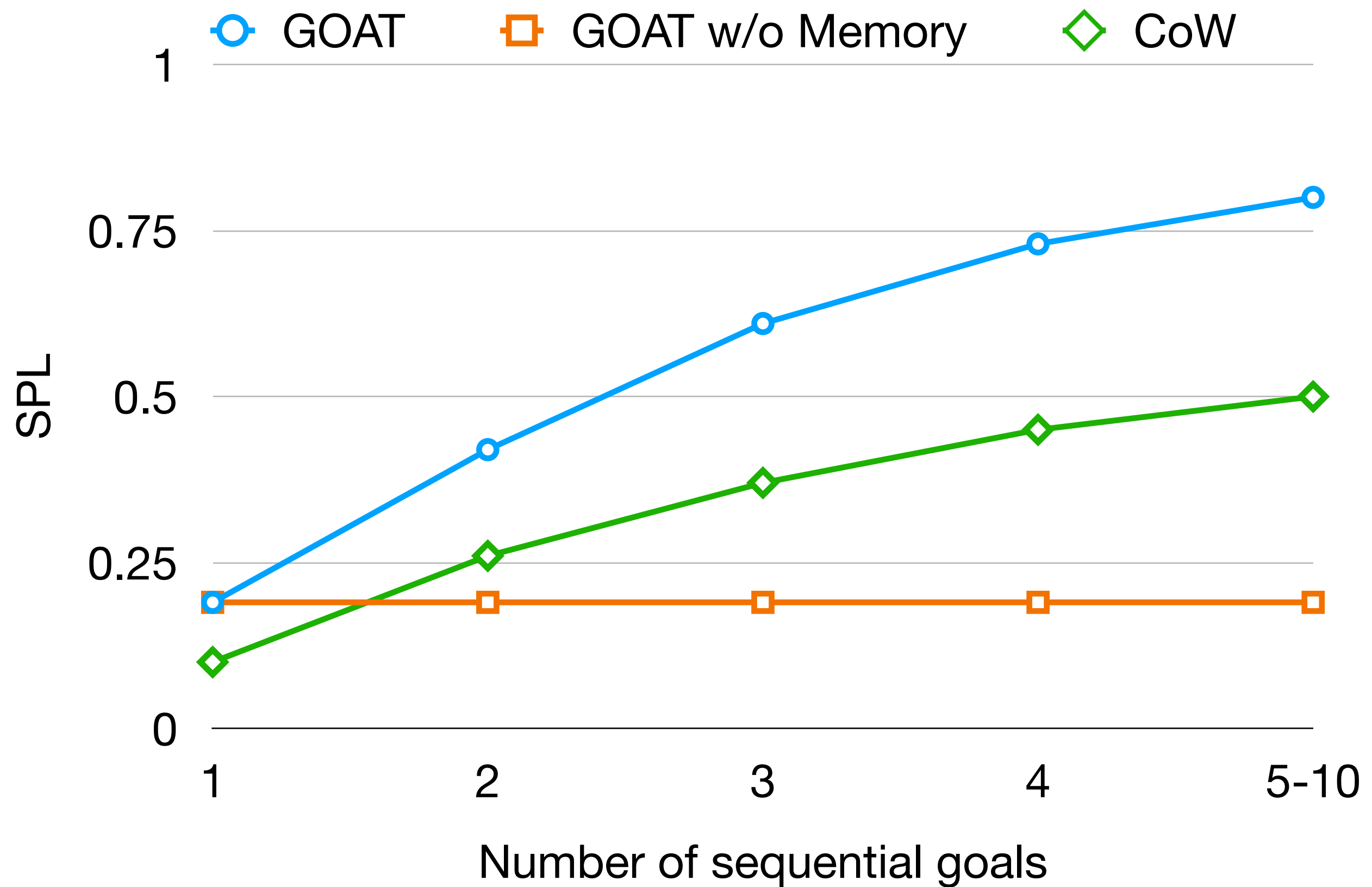
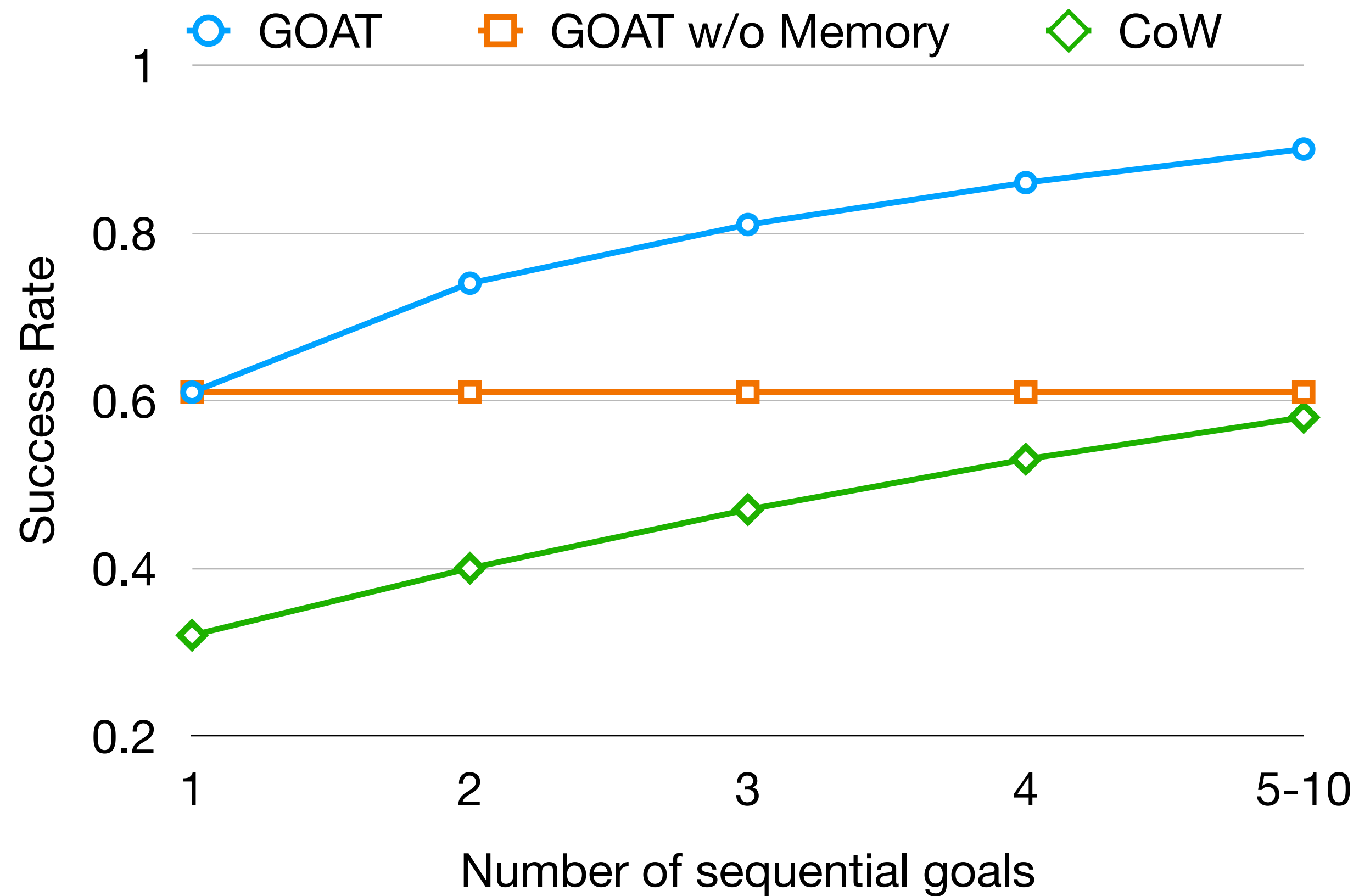
**Success Rate**

**SPL**





# Performance Across Episode



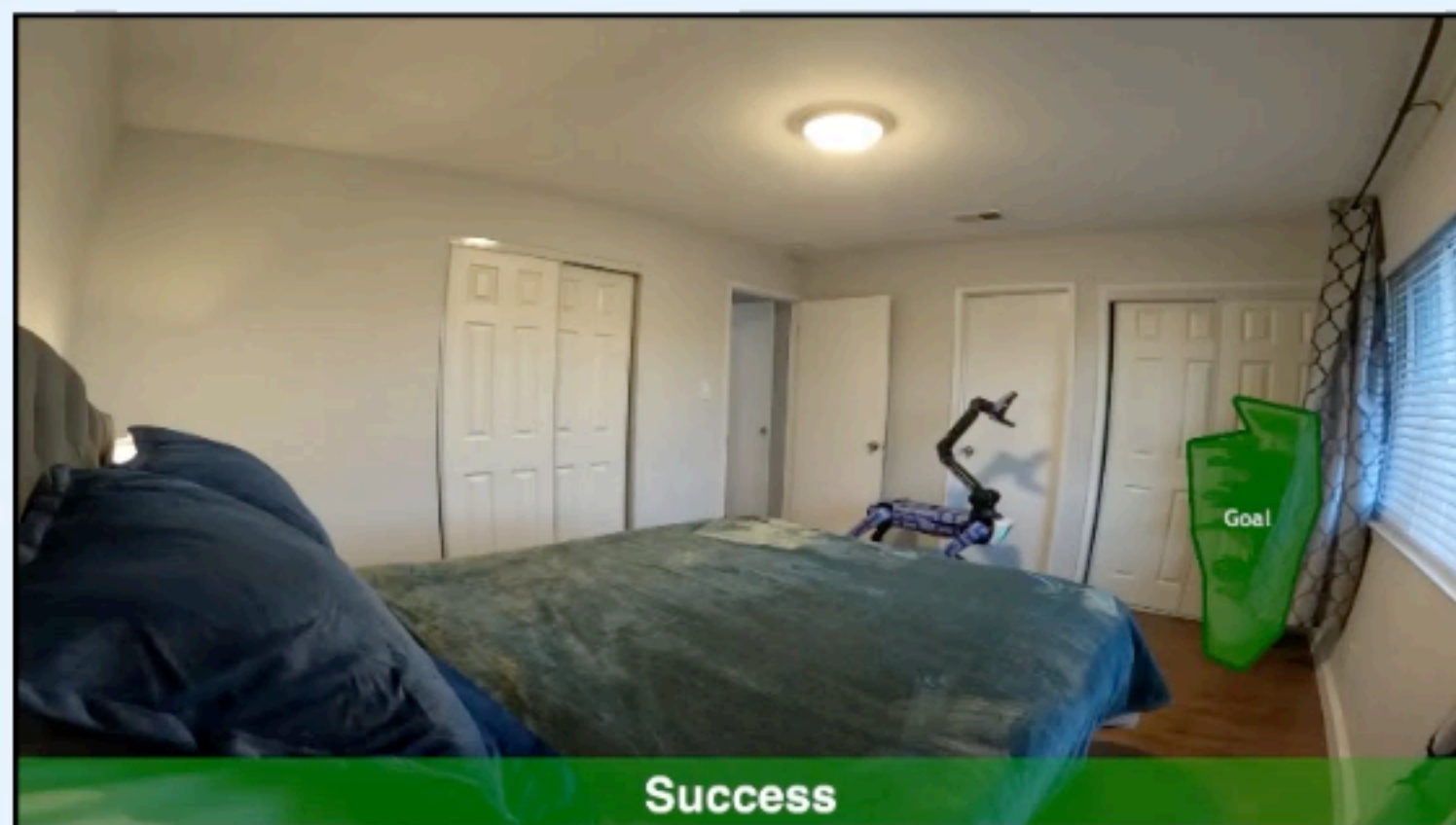


Goal:

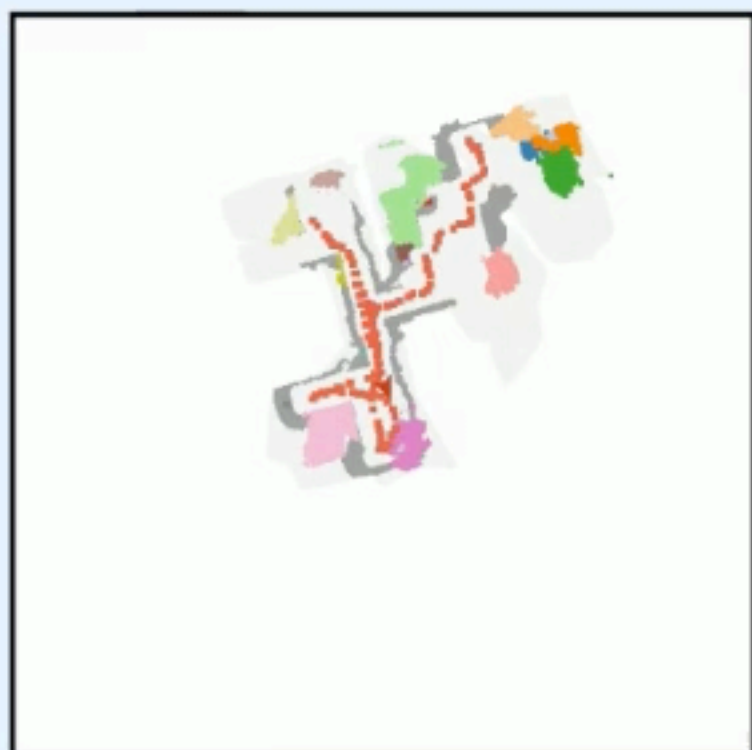


# Baselines

Ours



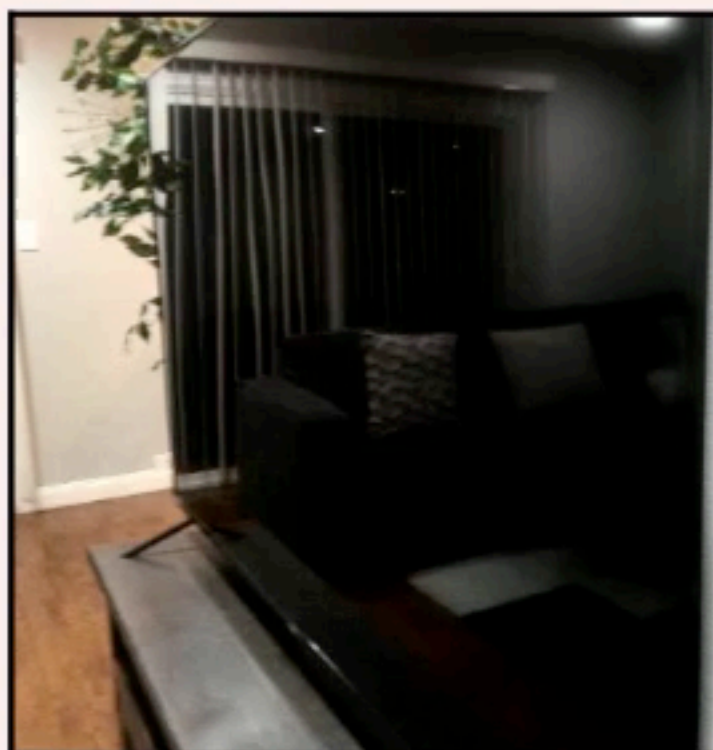
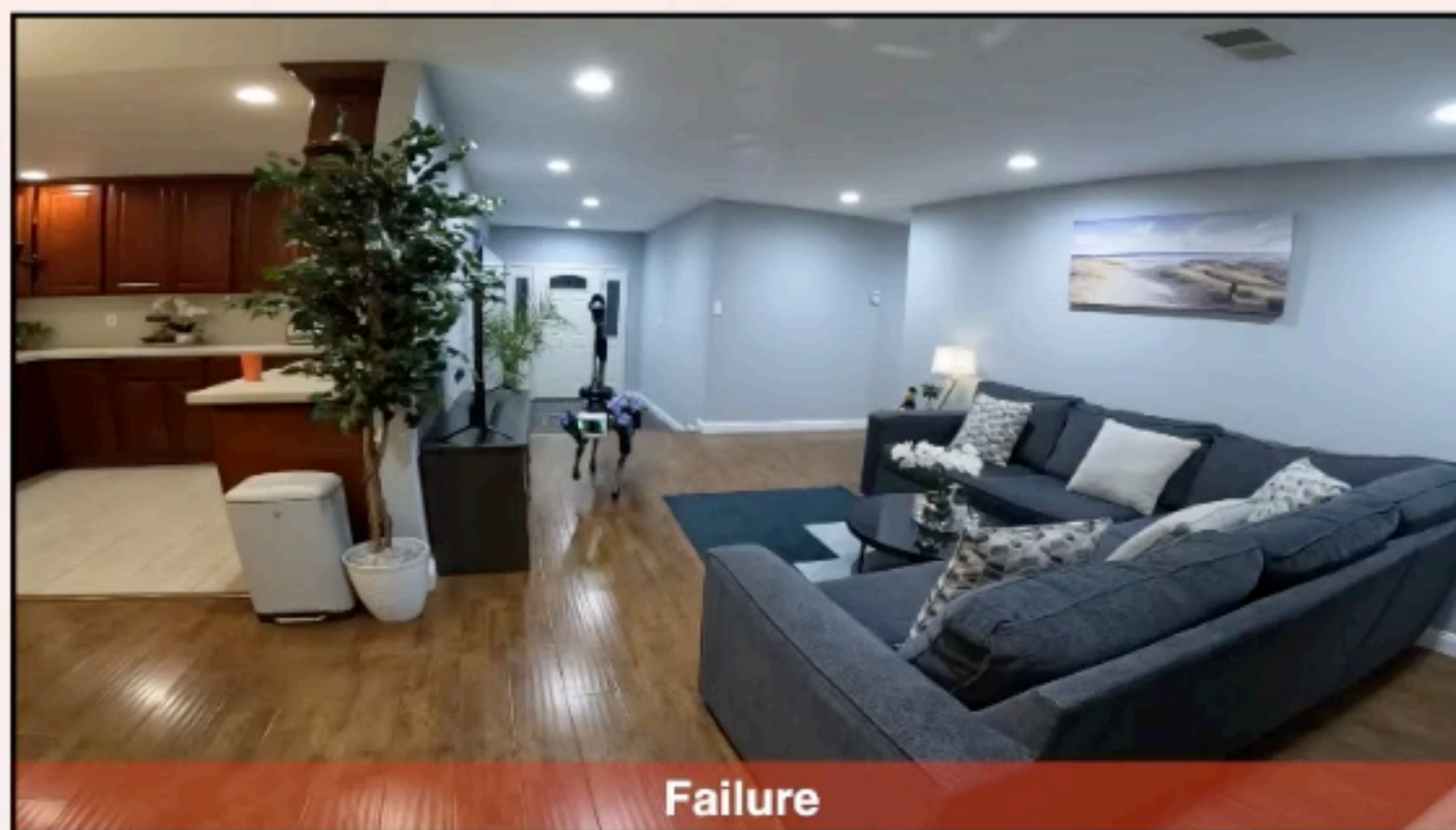
Observation



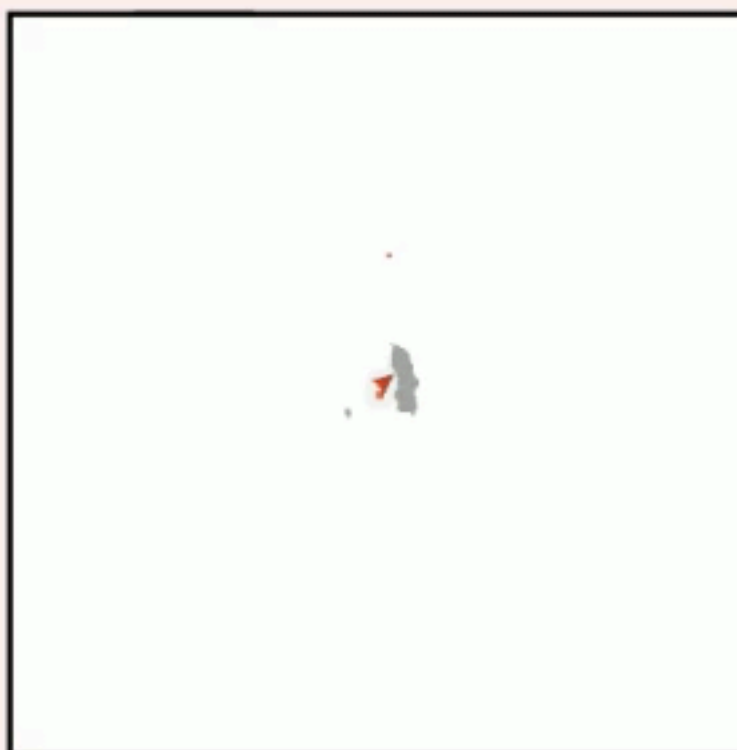
Instance Map

Success: 6/6 SPL: 0.78

No Memory



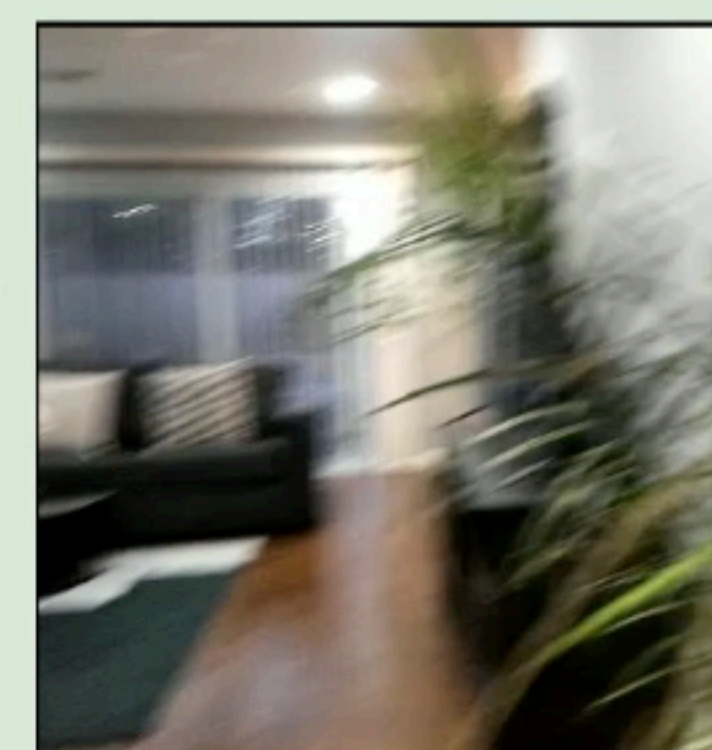
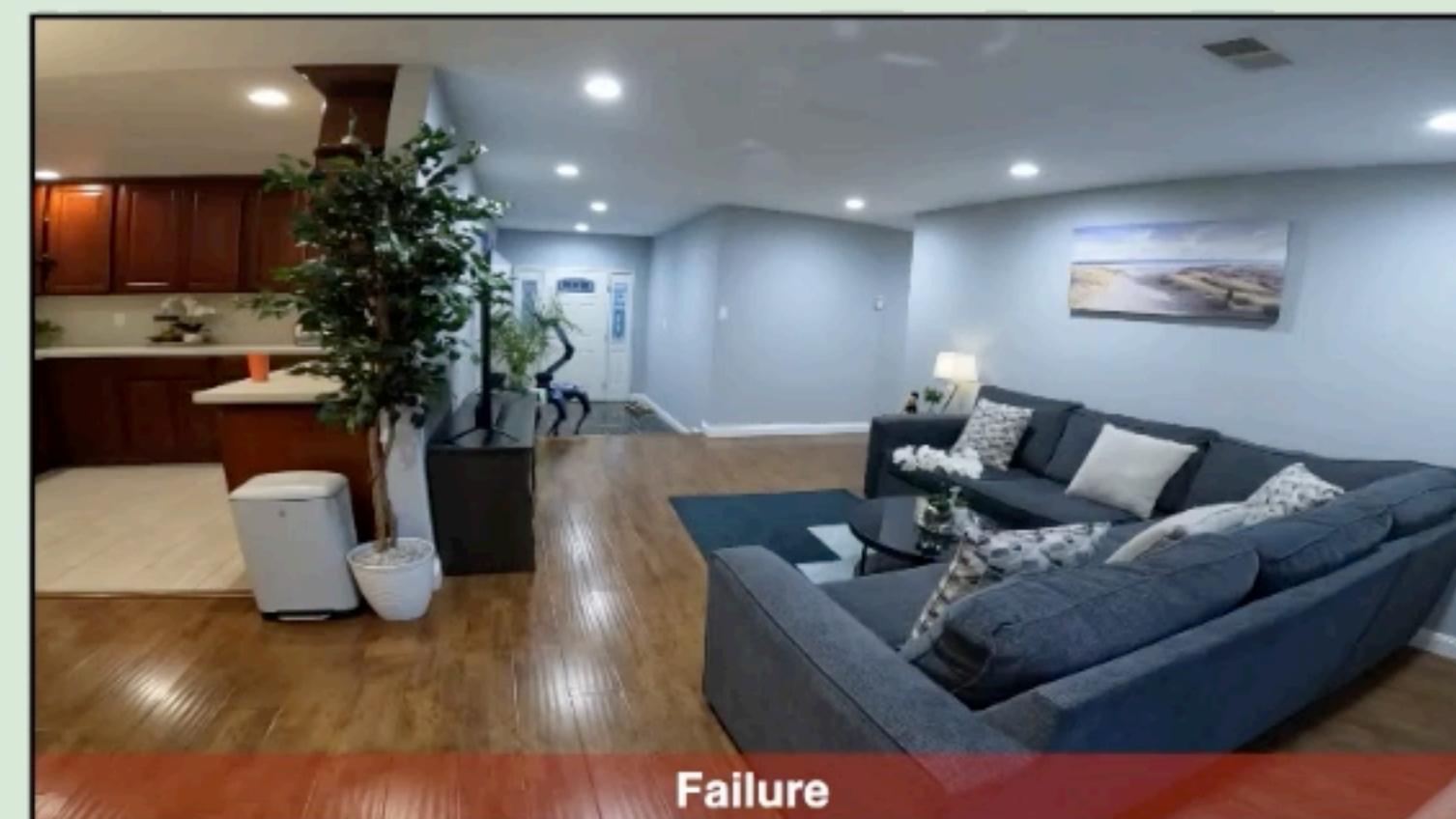
Observation



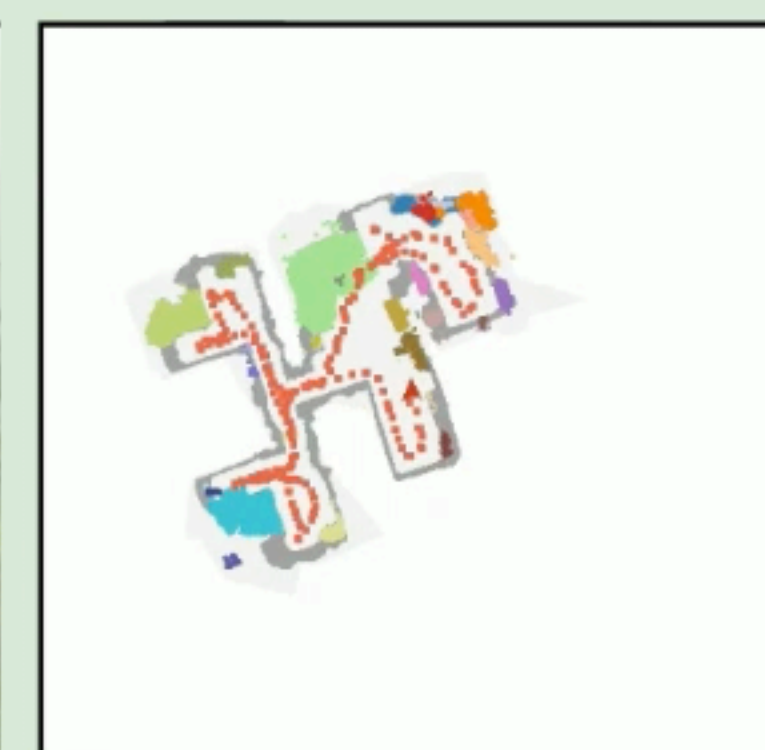
Instance Map

Success: 4/6 SPL: 0.40

CoW



Observation



Instance Map

Success: 1/6 SPL: 0.16

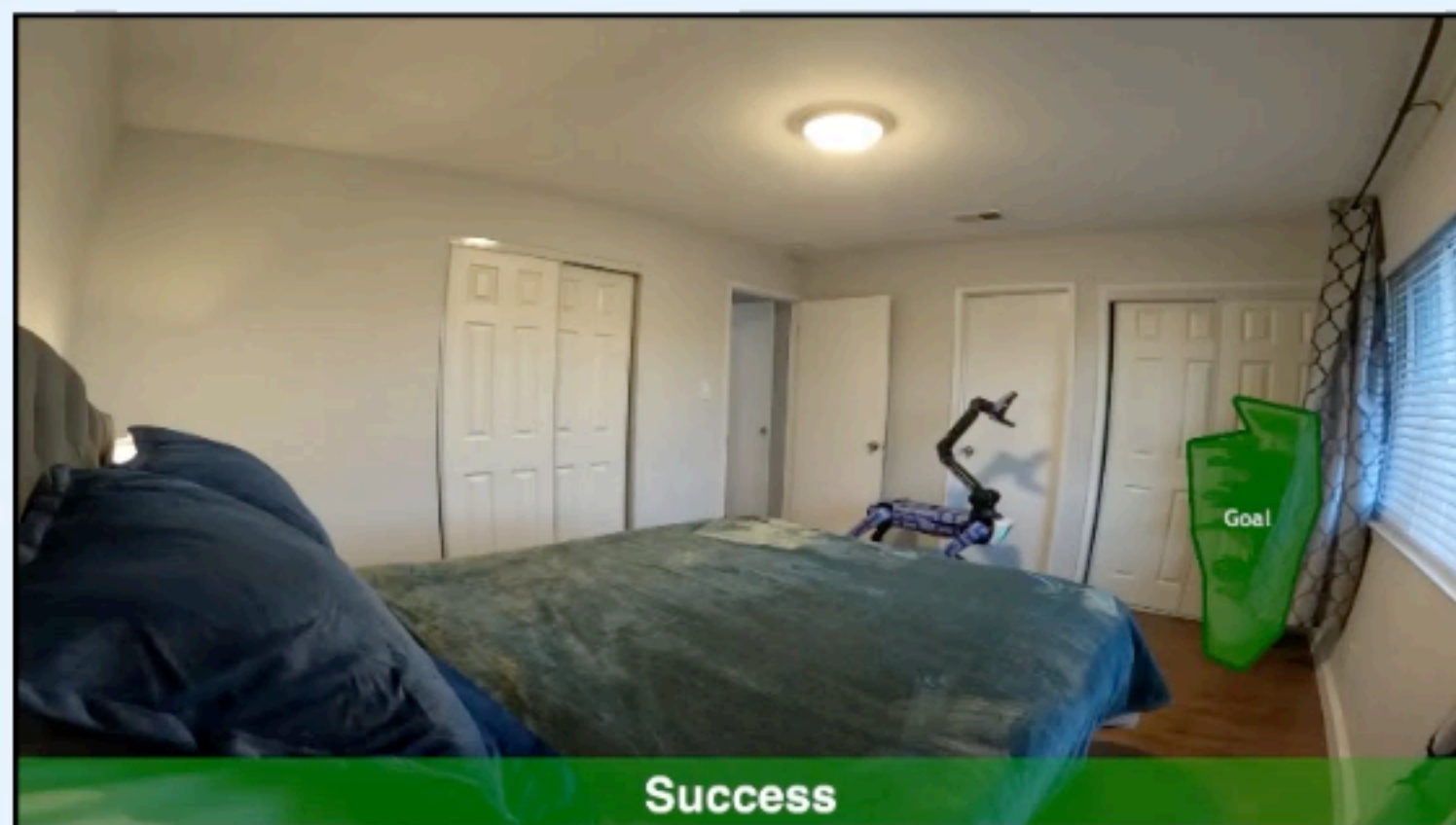


Goal:

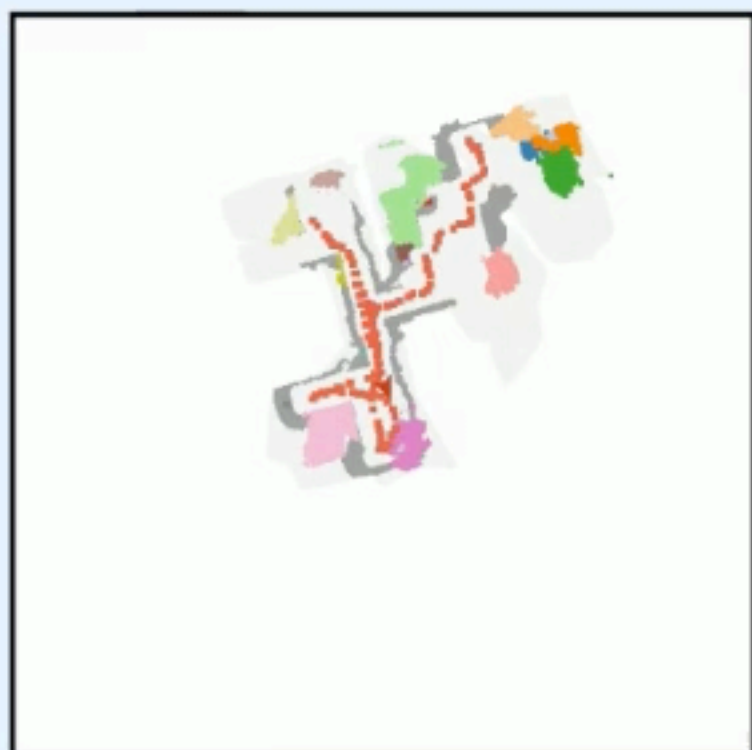


# Baselines

Ours



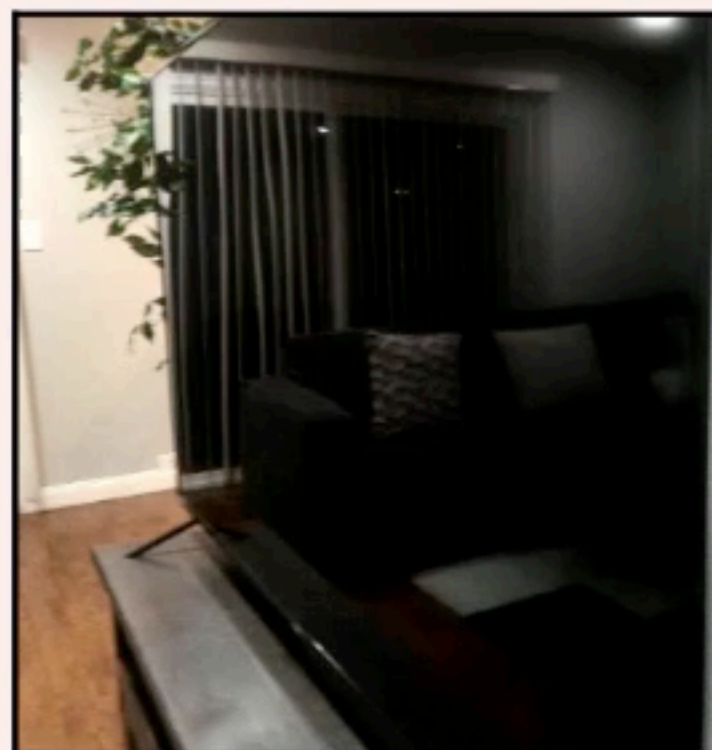
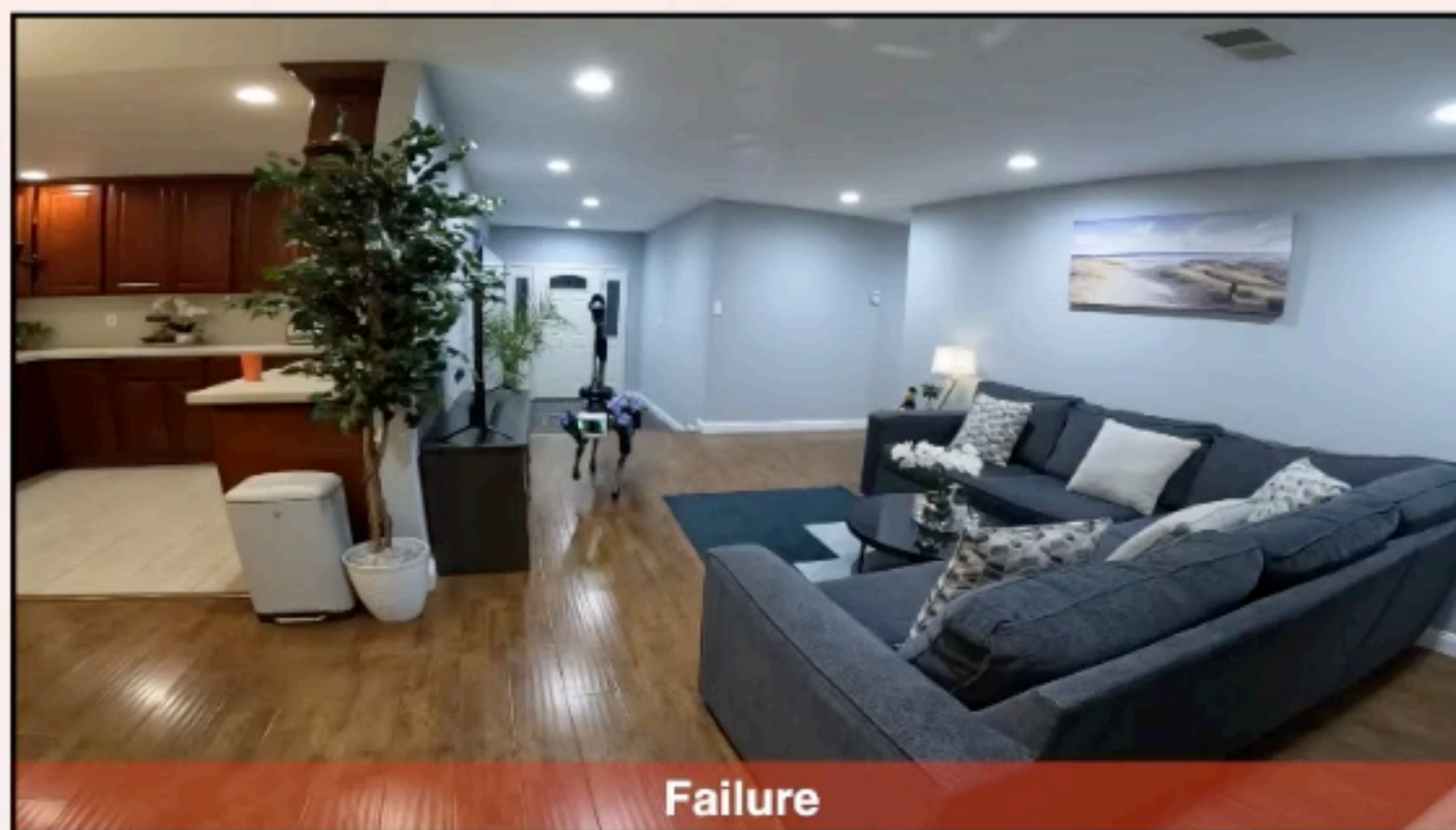
Observation



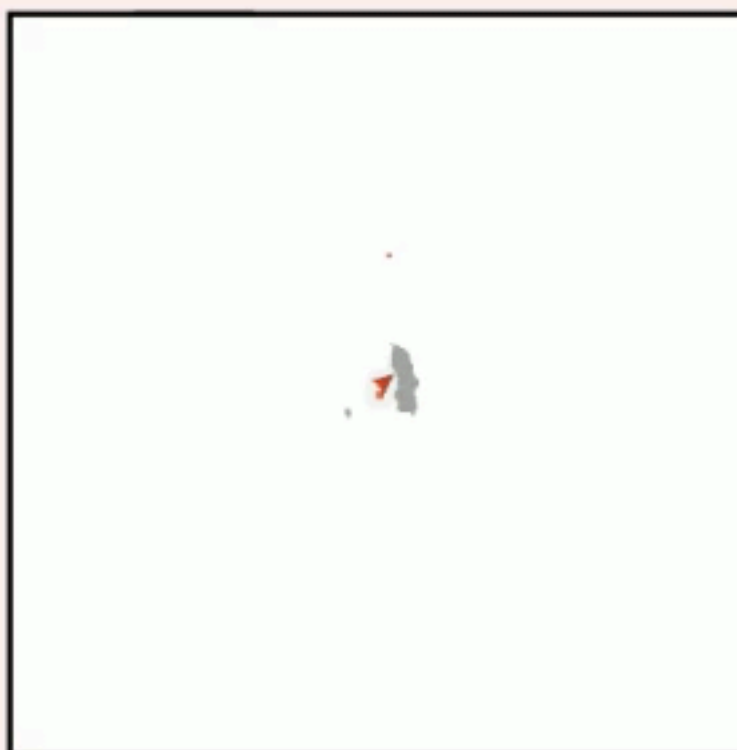
Instance Map

Success: 6/6 SPL: 0.78

No Memory



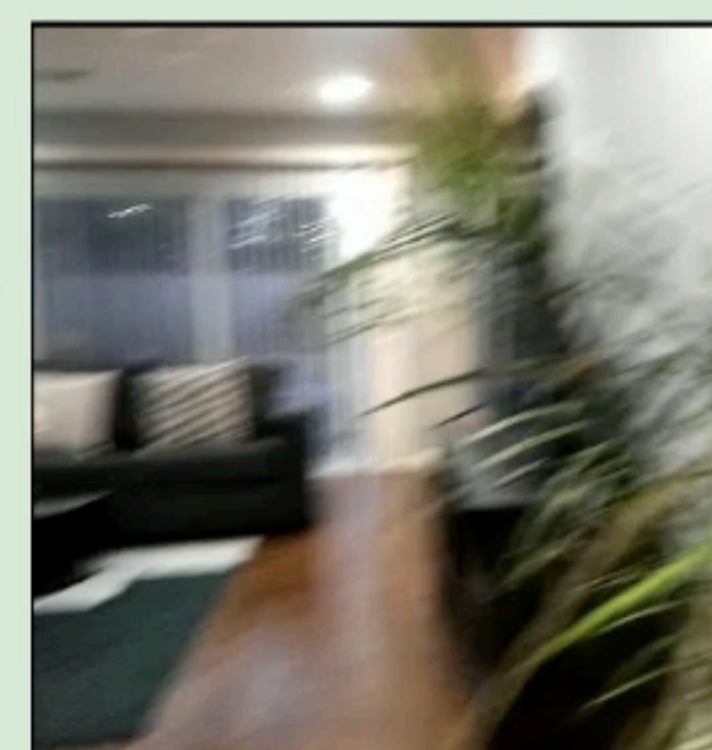
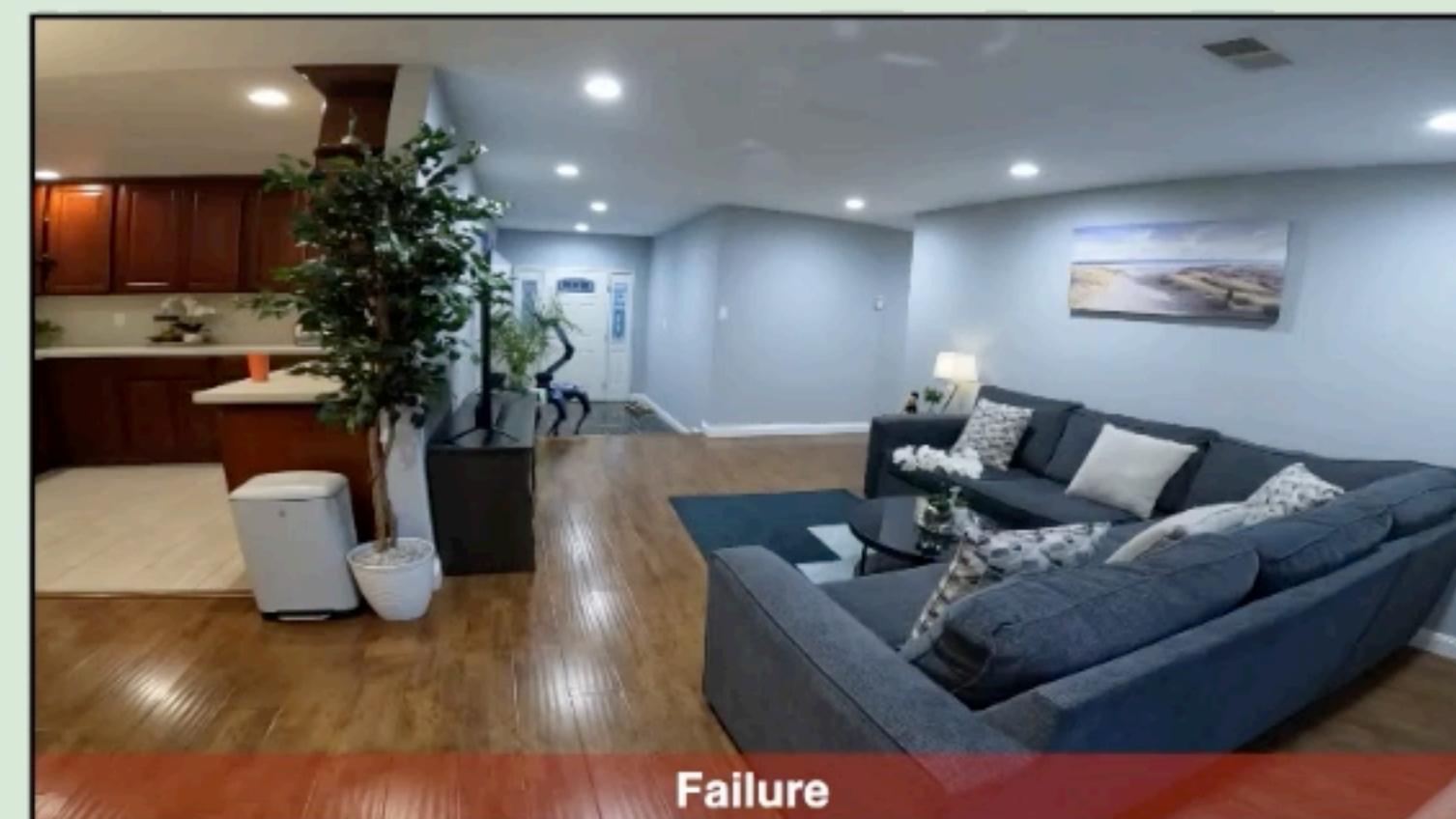
Observation



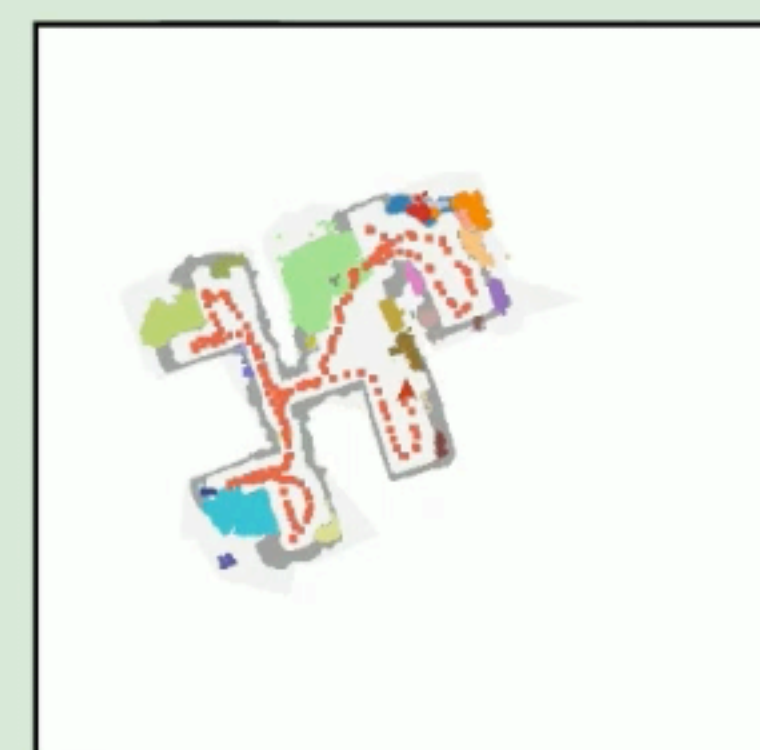
Instance Map

Success: 4/6 SPL: 0.40

CoW



Observation



Instance Map

Success: 1/6 SPL: 0.16

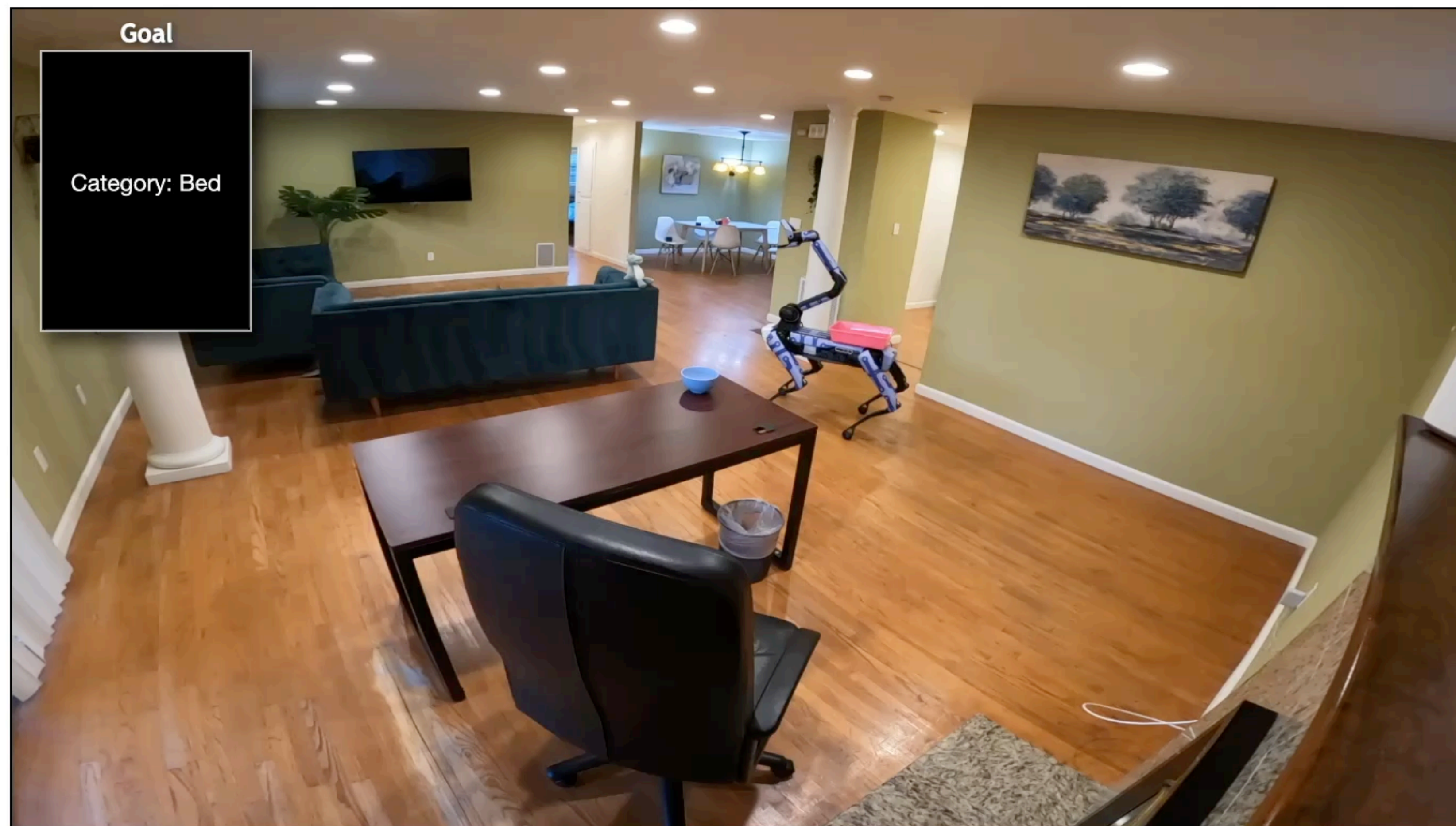


1. GOAT Problem
2. GOAT System Architecture
3. Results
4. Applications
  - a. Pick & Place
  - b. Social Navigation
  - c. Platform Agnostic



# Pick & Place

Third-person view





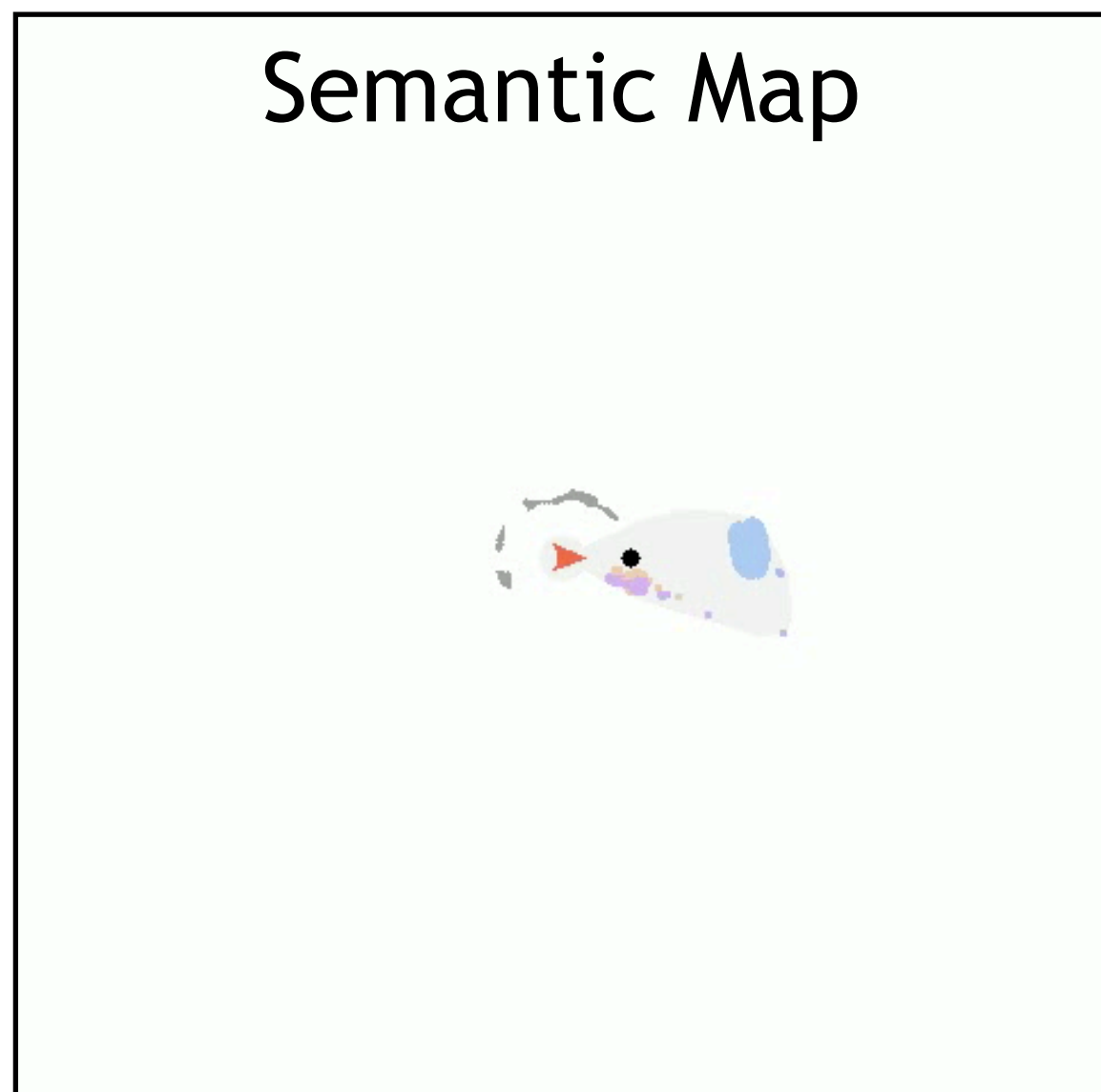
Observation



Third-person view



Semantic Map





# Social Navigation

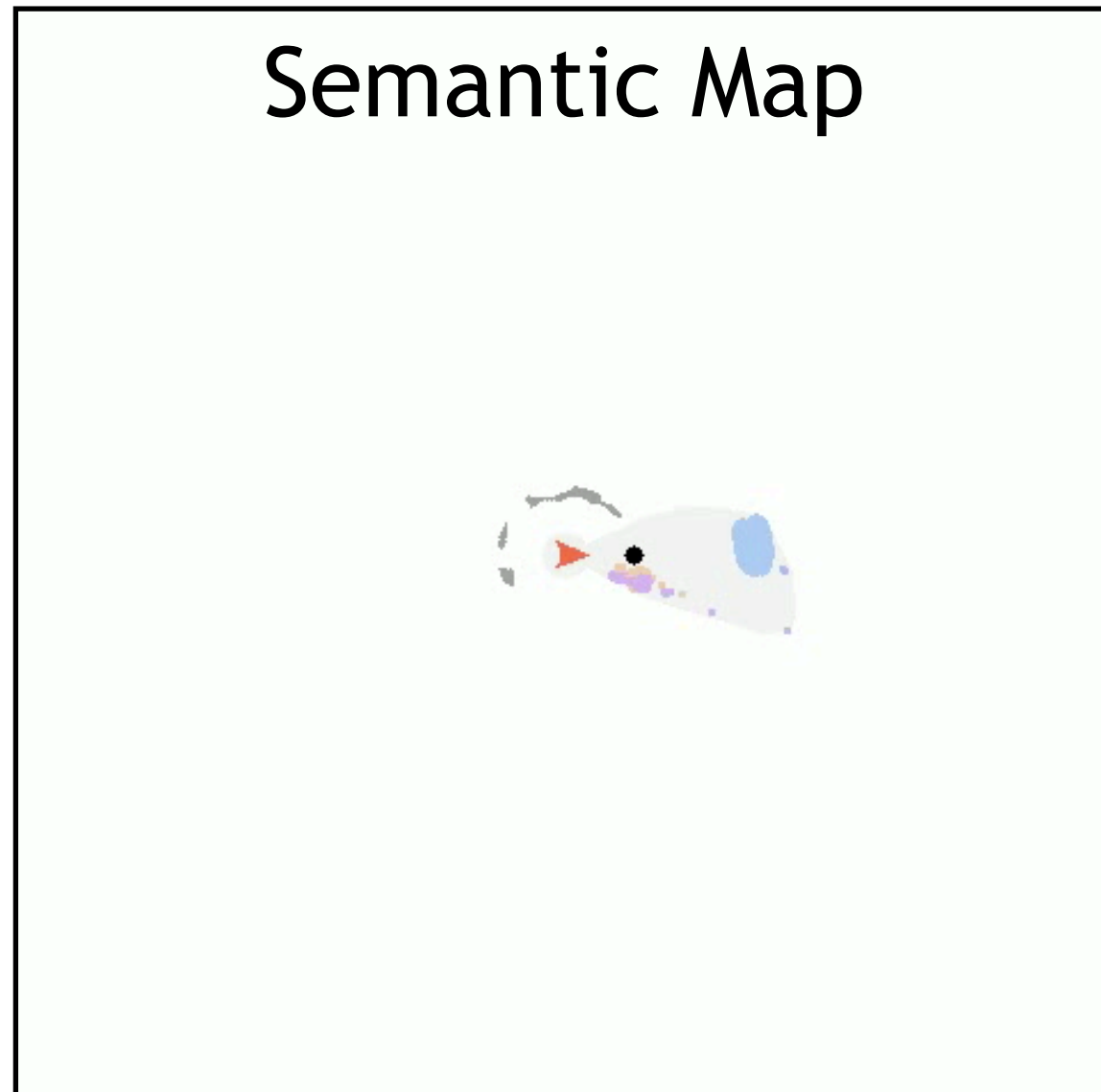
Observation



Third-person view



Semantic Map

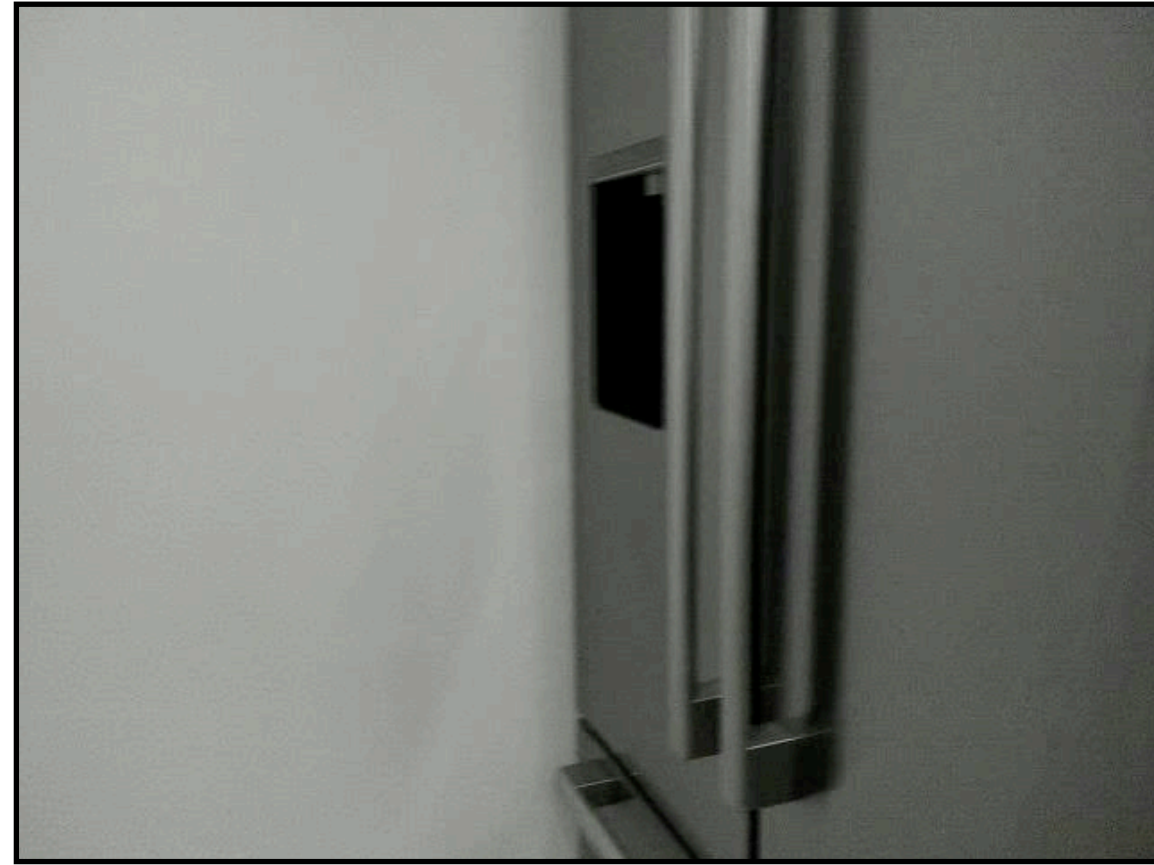


Robot plans around the dynamic obstacle (person) to go to the refrigerator



# Social Navigation

Observation



Third-person view



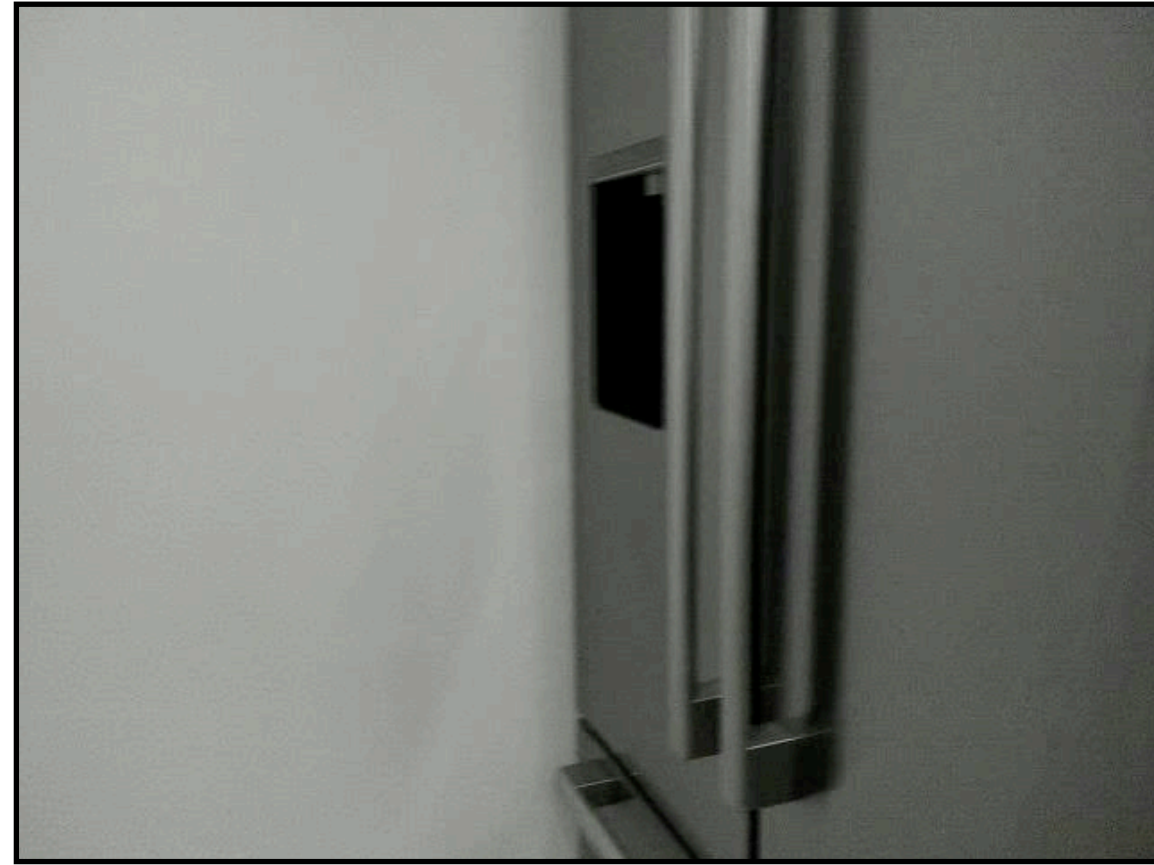
Semantic Map





# Social Navigation

Observation



Third-person view



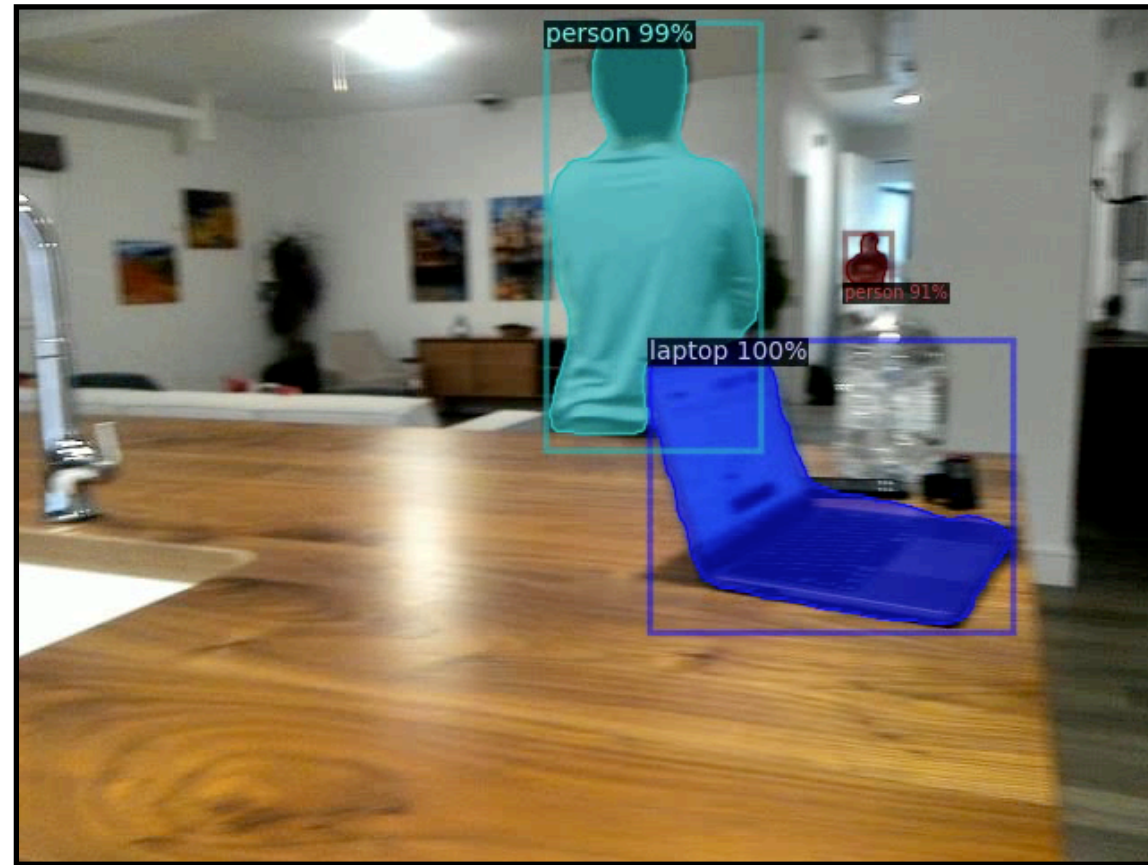
Semantic Map





# Social Navigation

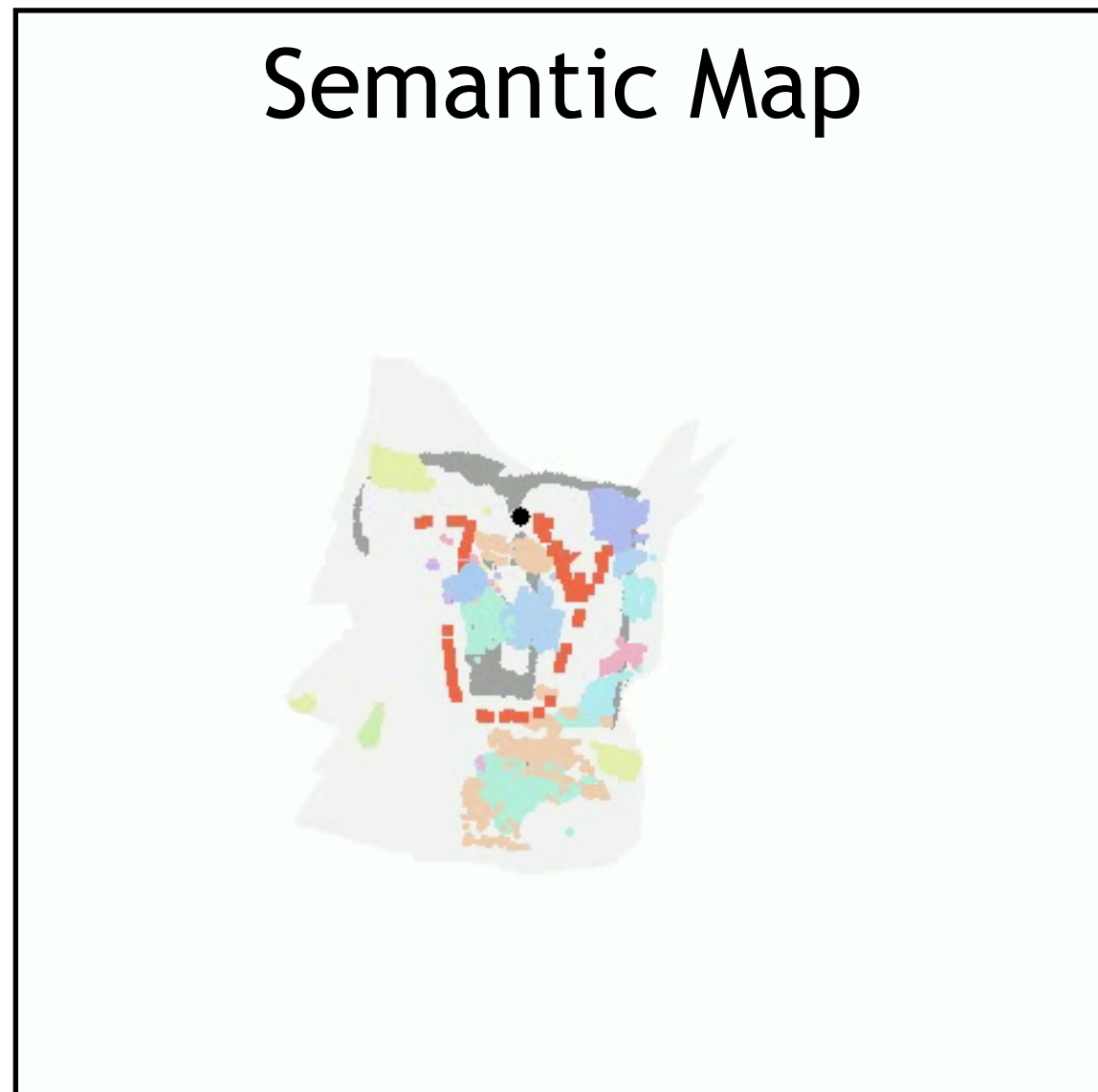
Observation



Third-person view



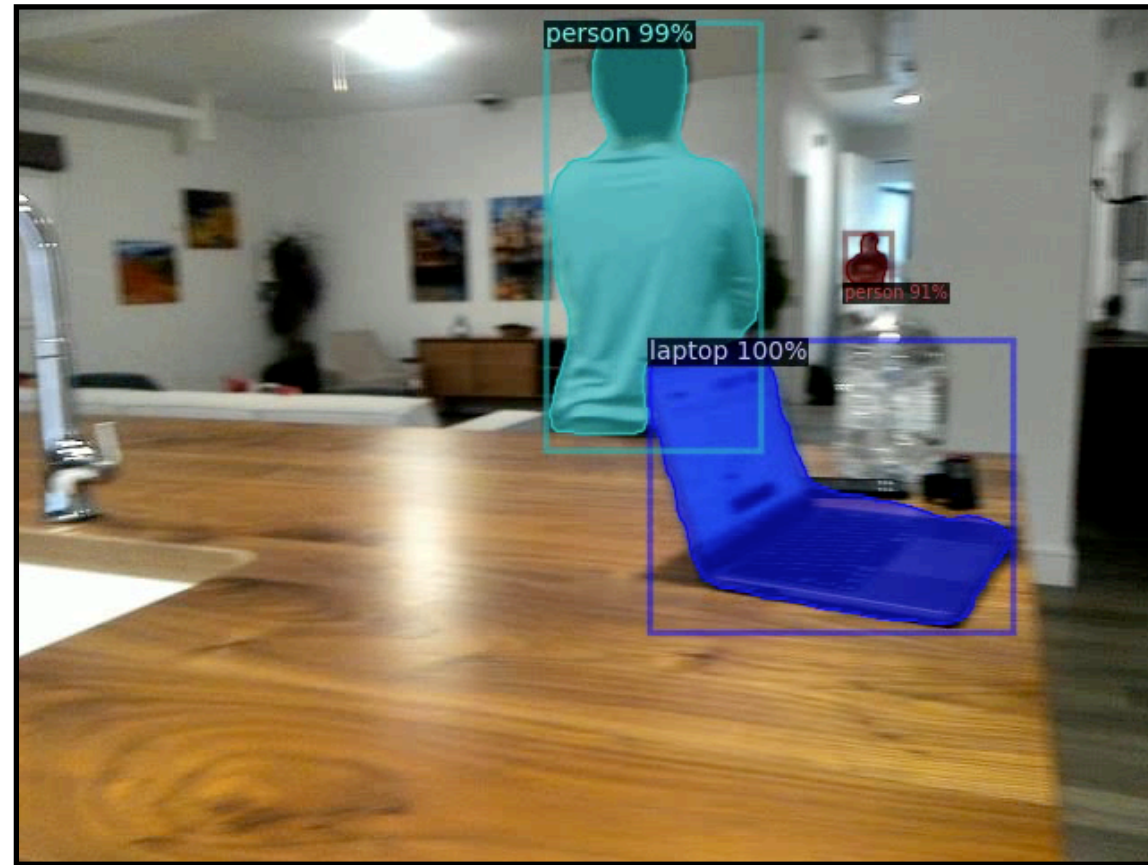
Semantic Map





# Social Navigation

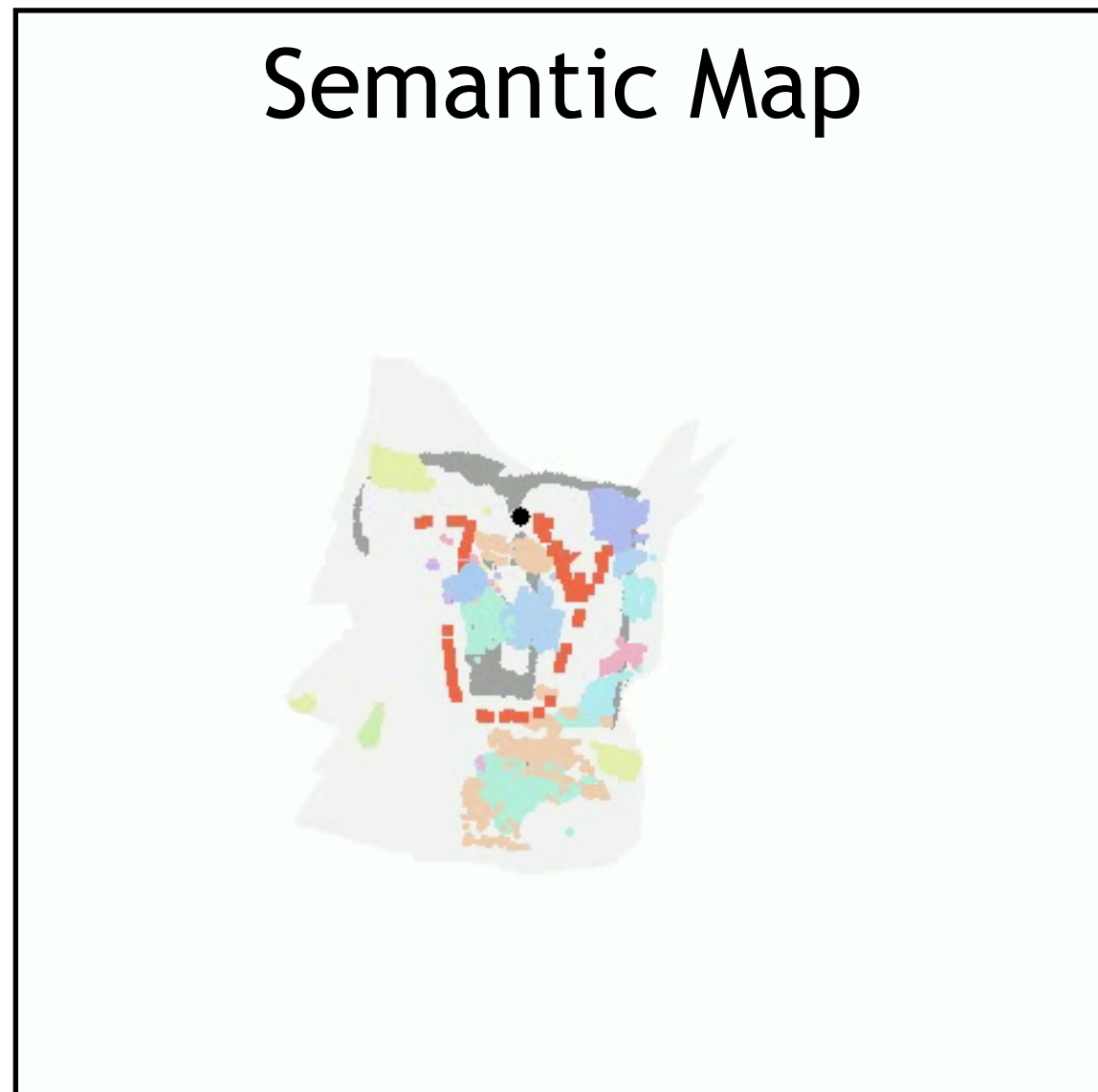
Observation



Third-person view



Semantic Map



Robot follows the person while updating their location



# Platform Agnostic

Third-person view



Goal

the large  
potted plant  
next to  
the couch



# Summary

## Universal navigation

- Multimodal

Image



Language

*Find **the fruit basket on the kitchen counter***

Category

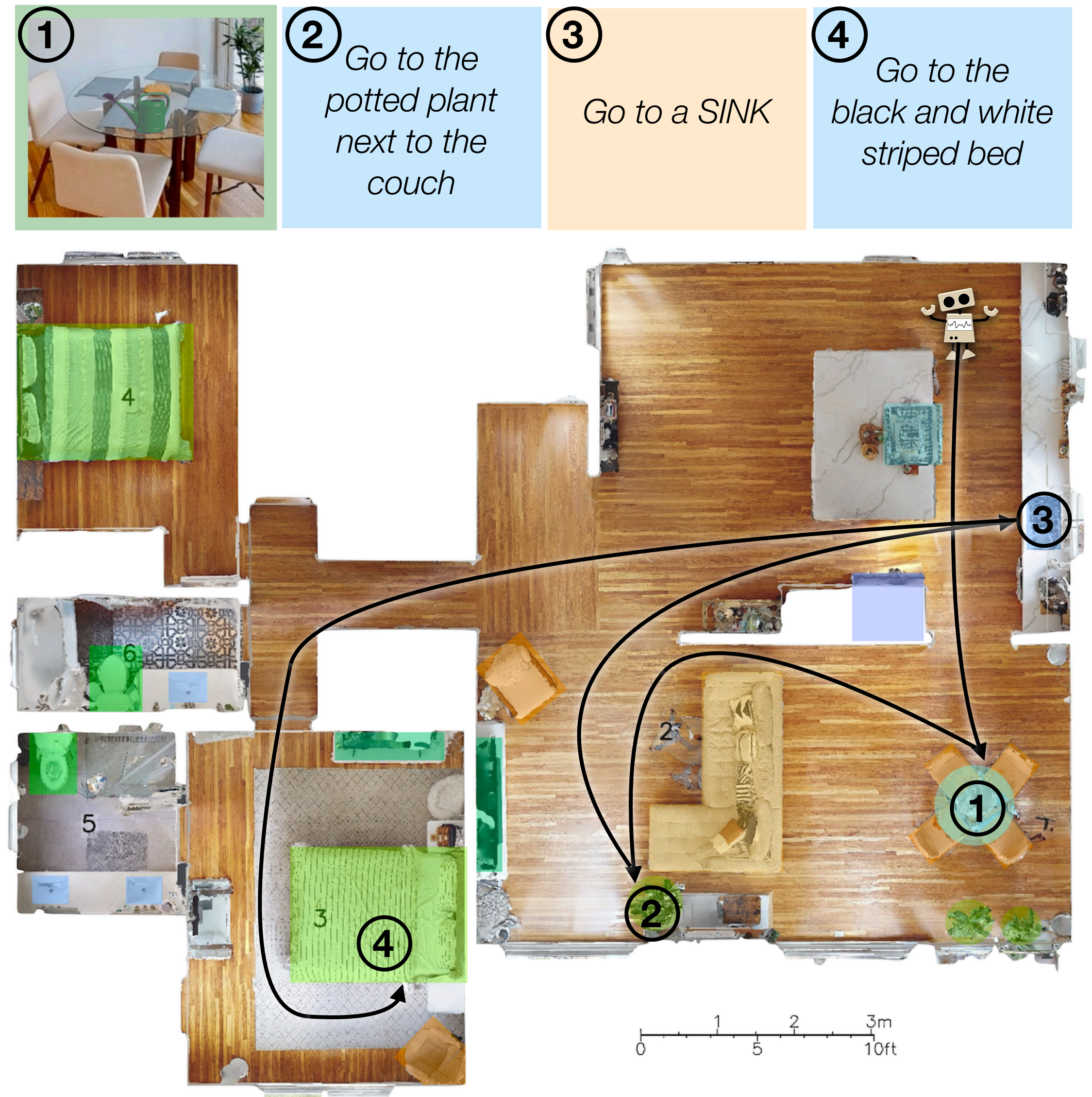
*Bring me a **CUP***



# Summary

## Universal navigation

- Multimodal
- Lifelong

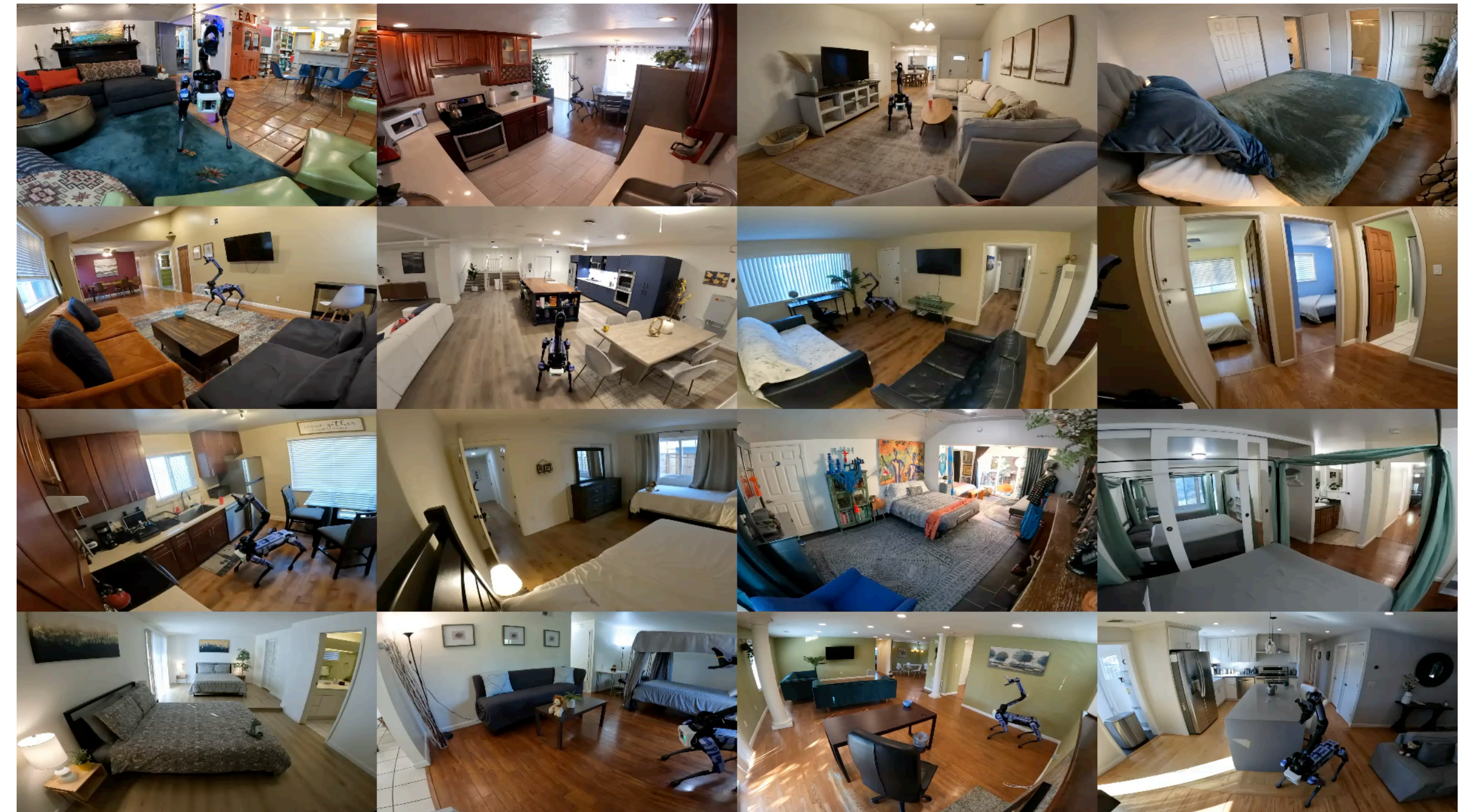




# Summary

## Universal navigation

- Multimodal
- Lifelong
- Unseen environments





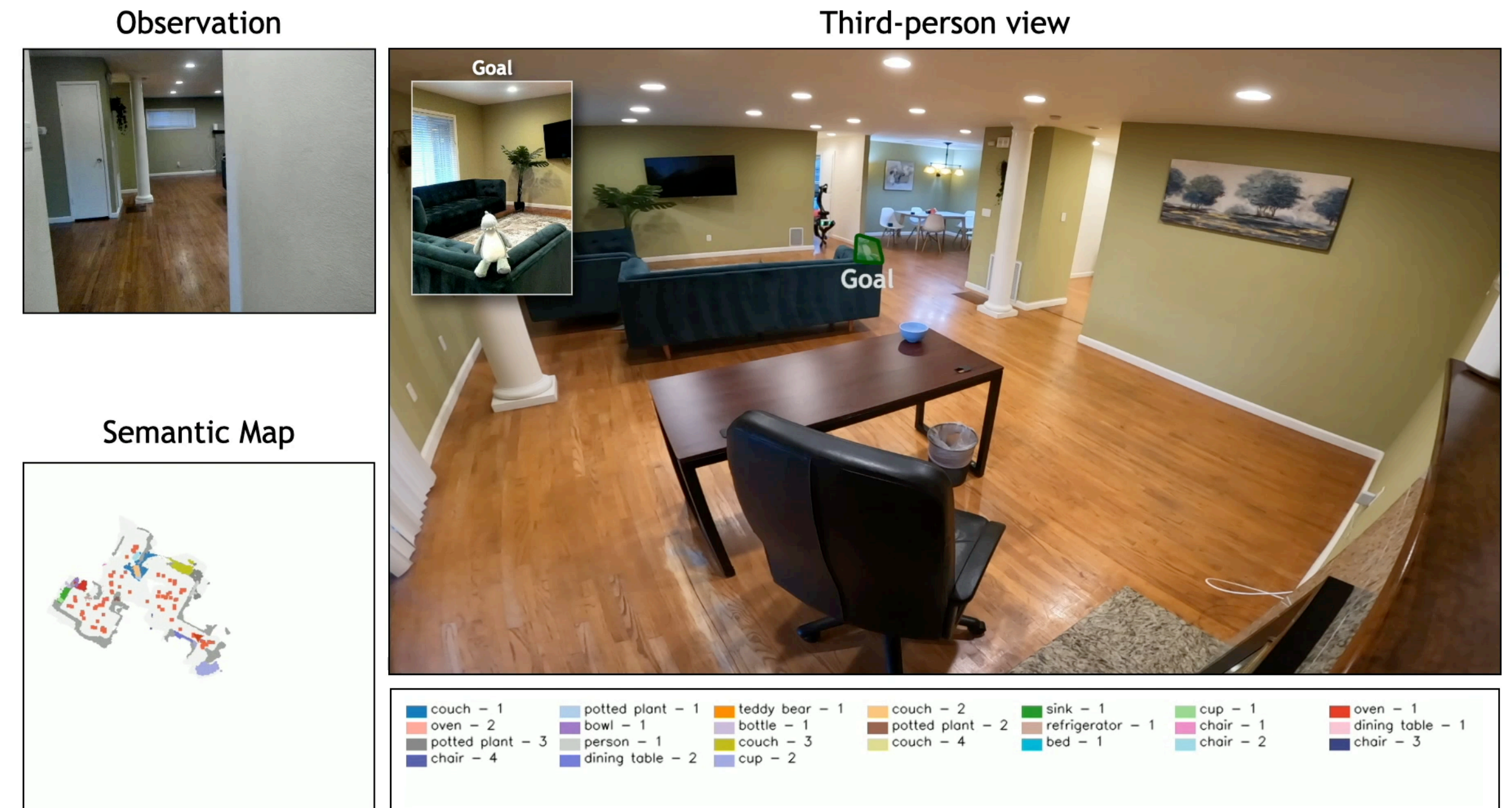
# Summary

## Universal navigation

- Multimodal
- Lifelong
- Unseen environments

## Applications

- Pick & Place





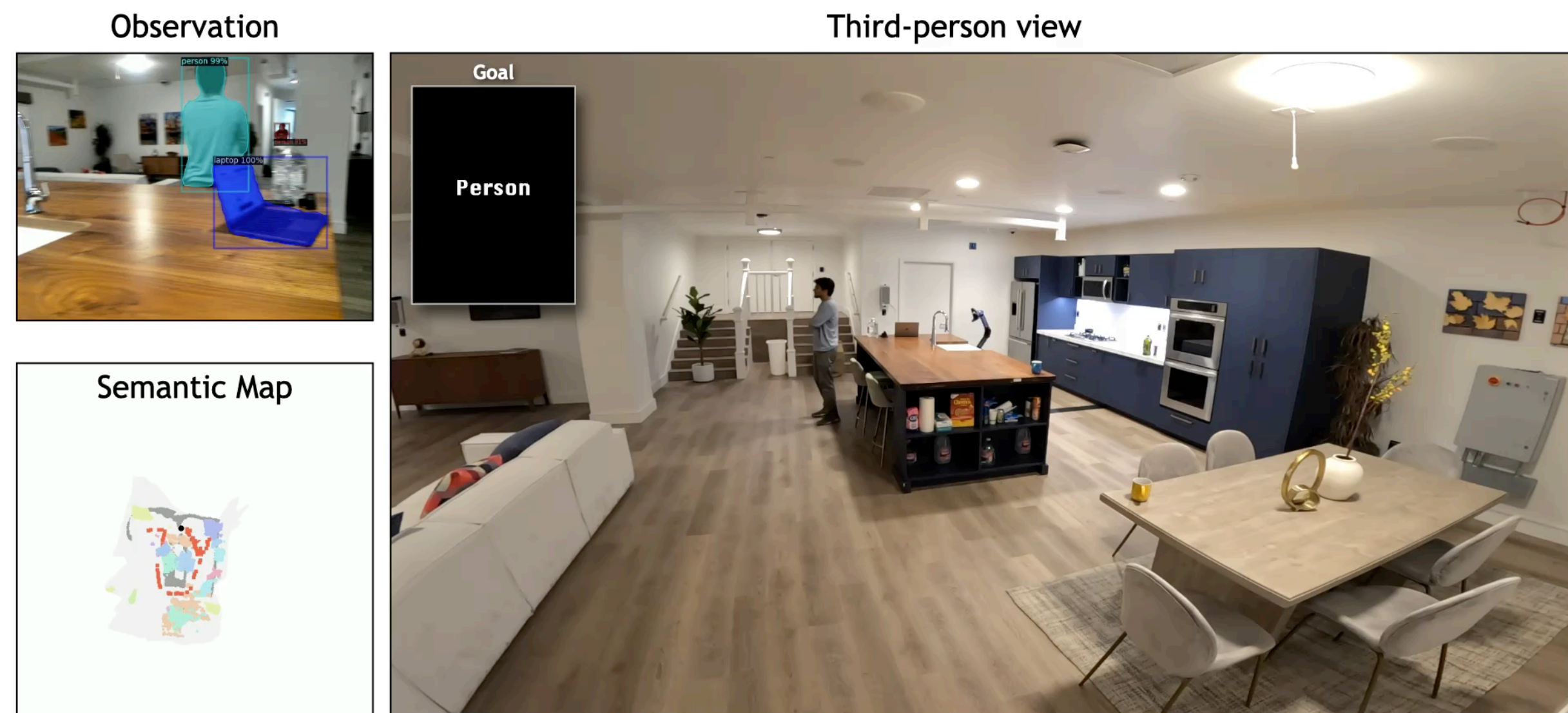
# Summary

## Universal navigation

- Multimodal
- Lifelong
- Unseen environments

## Applications

- Pick & Place
- Social Navigation





# Summary

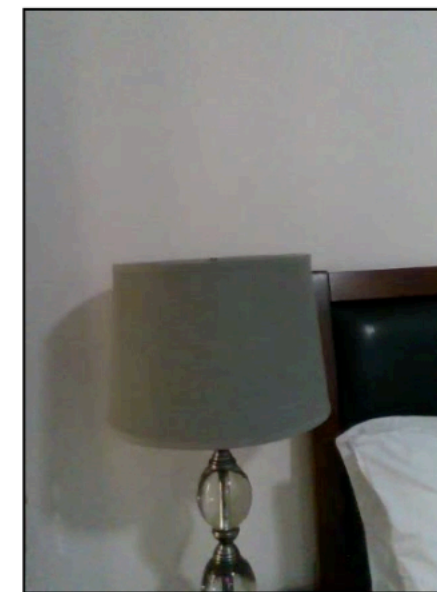
## Universal navigation

- Multimodal
- Lifelong
- Unseen environments

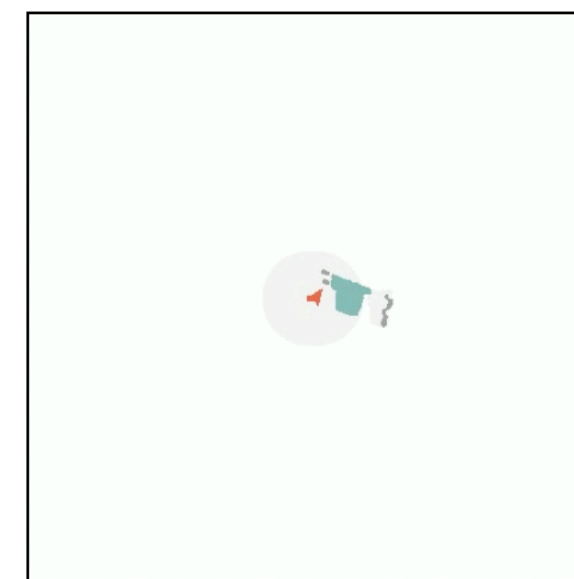
## Applications

- Pick & Place
- Social Navigation
- Platform Agnostic

Observation



Semantic Map



Third-person view



bed - 1





**Thank you!**

Webpage: <https://theophilegervet.github.io/projects/goat>