# CAPE: Corrective Actions from Precondition Errors using Large Language Models

## BACKGROUND

- **LLMs encode common-sense knowledge** useful for downstream tasks (Bosselut et al., 2019).
- We use **LLMs for decision-making with embodied agents performing long-horizon tasks** (e.g., cleaning a room).
- Previous works use zero-shot approaches, natural-language descriptions (Huang et al., 2022) and visual-language models (Ahn et al., 2022)
- Recent works use multi-modal approaches to detect precondition failures (Liu et al., 2023) (Zhang et al, 2023)
- We utilize **contextual information through precondition errors** and query LLMs to extract executable actions.

## METHODS



$$\mathcal{S}_g = \arg\max_{a_e} \left[ \max_{\hat{a}} \frac{C(f(\hat{a}), f(a_e)) + 1}{2} \cdot e^{P_\theta(\hat{a})} \right]$$
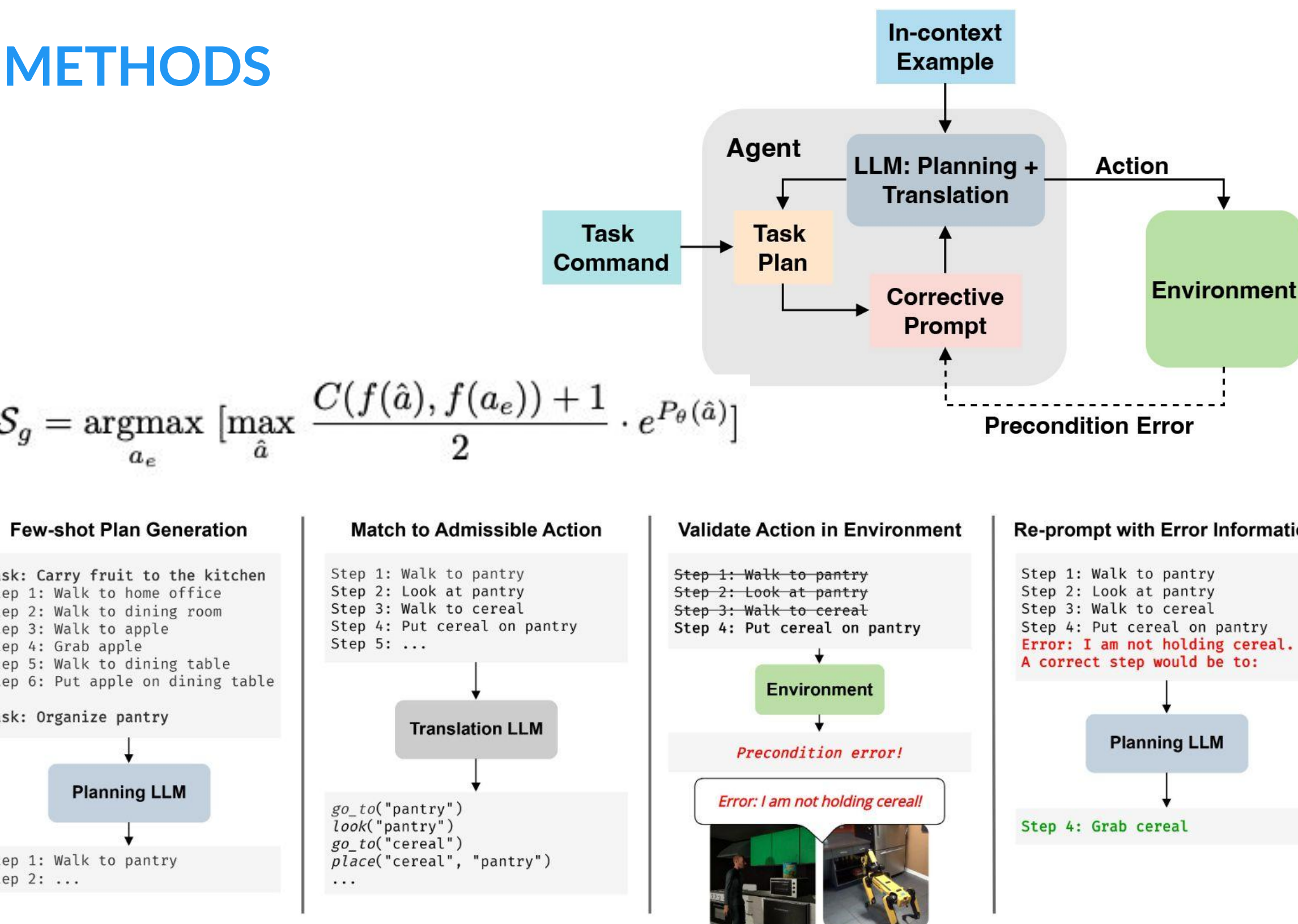


## RESULTS

| Method | % Exec | % Afford | Correct | % GS | Steps | Correction |
|---|---|---|---|---|---|---|
| Huang et al | 72.52 | 87.72 | 38.15 | 95.54 | 7.21 | - |
| Resampling | 76.47 | 75.24 | 38.89 | 95.65 | 6.87 | 7.67 |
| SayCan | 97.33 | 99.89 | 22.59 | 94.68 | 5.97 | - |
| Success Only | 97.57 | 90.46 | 41.11 | 95.49 | 7.68 | 1.08 |
| Implicit | 97.86 | 90.05 | 42.22 | 95.64 | 7.48 | 0.93 |
| Explicit | 98.57 | 91.69 | 42.59 | 96.23 | 8.81 | 0.72 |
| FS Explicit | 98.57 | 92.23 | 47.04 / 49.63 | 96.05 / 96.29 | 8.69 | 0.89 |

| Method | % Exec | % Afford | Correct | % GS | Steps | Correction |
|---|---|---|---|---|---|---|
| Huang et al | 87.72 | 56.46 | 16.67 | 66.03 | 2.40 | - |
| Resampling | 75.24 | 47.98 | 13.33 | 67.33 | 4.60 | 13.19 |
| SayCan | 99.89 | 79.13 | 16.67 | 67.54 | 6.80 | - |
| Success Only | 75.00 | 43.05 | 18.33 | 66.02 | 3.04 | 2.25 |
| Implicit | 75.00 | 52.37 | 20.00 | 66.25 | 3.14 | 1.83 |
| Explicit | 100 | 79.69 | 31.67 | 69.18 | 6.30 | 1.91 / 1.73 |
| FS Explicit | 100 | 81.36 | 45.00 / 50.00 | 77.91 | 11.7 | 2.91 |

**Explicit Reprompting** reduces executability error rate by **~50%** and improves semantic correctness by **~20%**


**View project website here!**

---

# Few-shot reasoning from action precondition error information helps LLMs generate substantially more semantically correct and executable plans.

### HUANG ET AL

**Task: Put away groceries**

Step 1: Walk to kitchen
Step 2: Walk to fridge
Step 3: Open fridge
Step 4: Put cereal in fridge

**Task: Eat snacks and drink tea**

Step 1: Walk to livingroom
Step 2: Sit on couch
Step 3: Drink tea
Step 4: Eat food snack

### SAYCAN

**Task: Put away groceries**

Step 1: Walk to kitchen
Step 2: Walk to fridge
Step 3: Find cereal
Step 4: Done

**Task: Eat snacks and drink tea**

Step 1: Walk to dining room
Step 2: Walk to cupboard
Step 3: Open cupboard
Step 4: Find tea
Step 5: Grab tea
Step 6: Drink tea
Step 7: Walk to freezer
Step 8: Open freezer
Step 9: Find juice
Step 10: Grab juice
Step 11: Drink juice
Step 12: Walk to cupboard

### OUR METHOD

**Task: Put away groceries**

Step 1: Walk to dining room
Step 2: Walk to kitchen
Step 3: Walk to fridge
Step 4: Open fridge
Step 5: Put cereal in fridge
Error: I am not holding the cereal. A correct step would be to
Step 5: Grab cereal
Step 6: Put cereal in fridge
Step 7: Put milk in fridge
Error: I am not holding the milk. A correct step would be to
Step 7: Grab milk
Step 8: Put milk in fridge
.......
Step 17: Grab pasta
Step 18: Put pasta in fridge

**Task: Eat snacks and drink tea**

Step 1: Walk to dining room
..........
Step 4: Find tea
Step 5: Grab tea
Step 6: Walk to table
Step 7: Pour tea into cup
Error: I am not near the cup. A correct step would be to
Step 7: Walk to cup
Step 8: Find cup
Step 9: Grab cup
Step 10: Pour tea into cup
Step 11: Drink tea
Step 12: Eat food snack
Error: I am not near cereal. A correct step would be to
Step 12: Walk to food cereal
Step 13: Open cereal
Error: I cannot open cereal. My hands are full. A ...would be to
Step 13: Put tea on table
Error: I am not near the table. A correct step would be to
Step 13: Walk to table

---

## ADDITIONAL RESULTS

### Re-prompting Strategies

- Re-prompting with Success-Only
  `Error: Task Failed`
- Re-prompting with Implicit Cause
  `Error: I cannot <action> the <object>`
- Re-prompting with Explicit Cause
  `Error: I cannot <action> the <object> because <templated reason>`

### Qualitative Example: Robot Execution
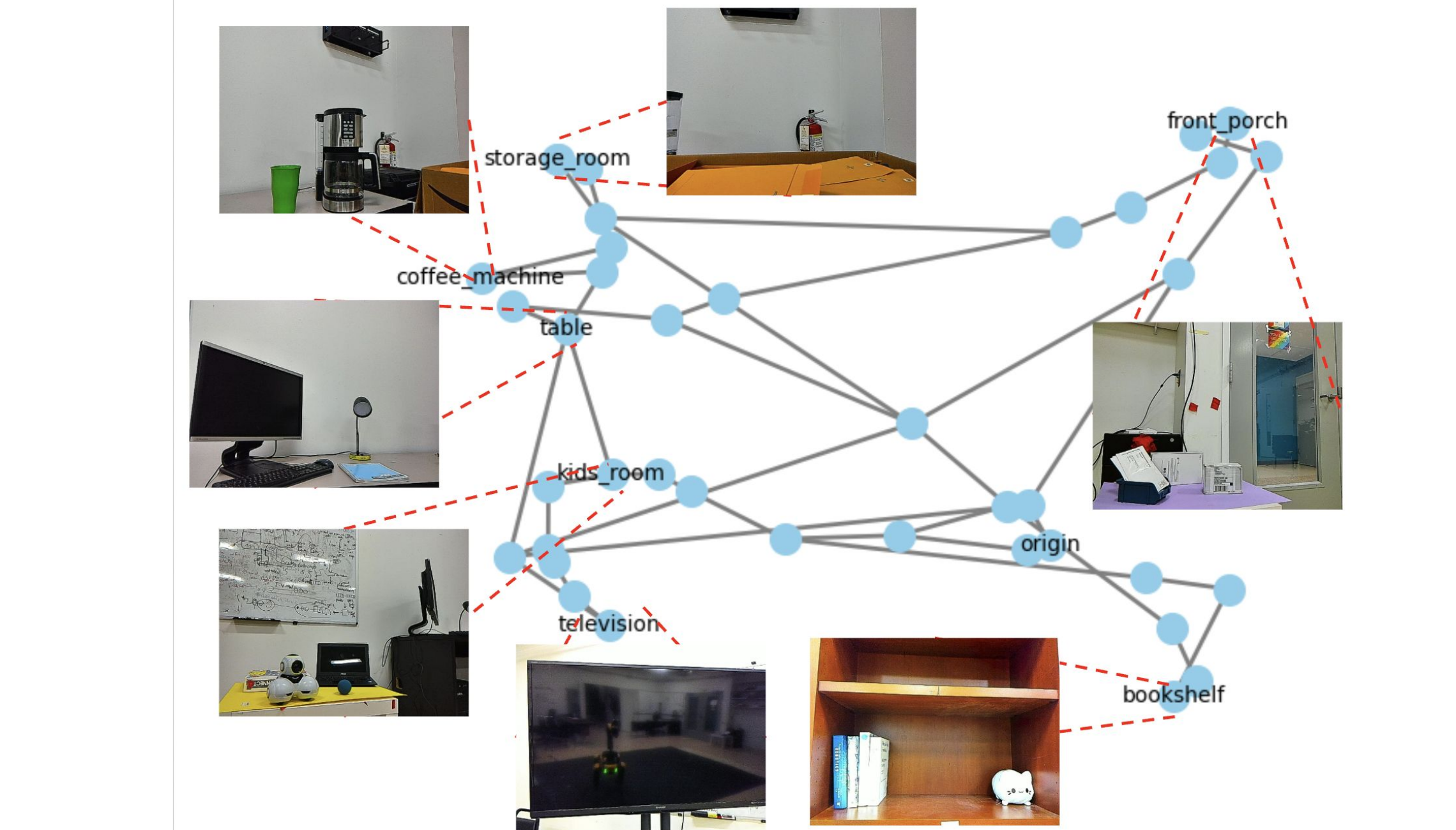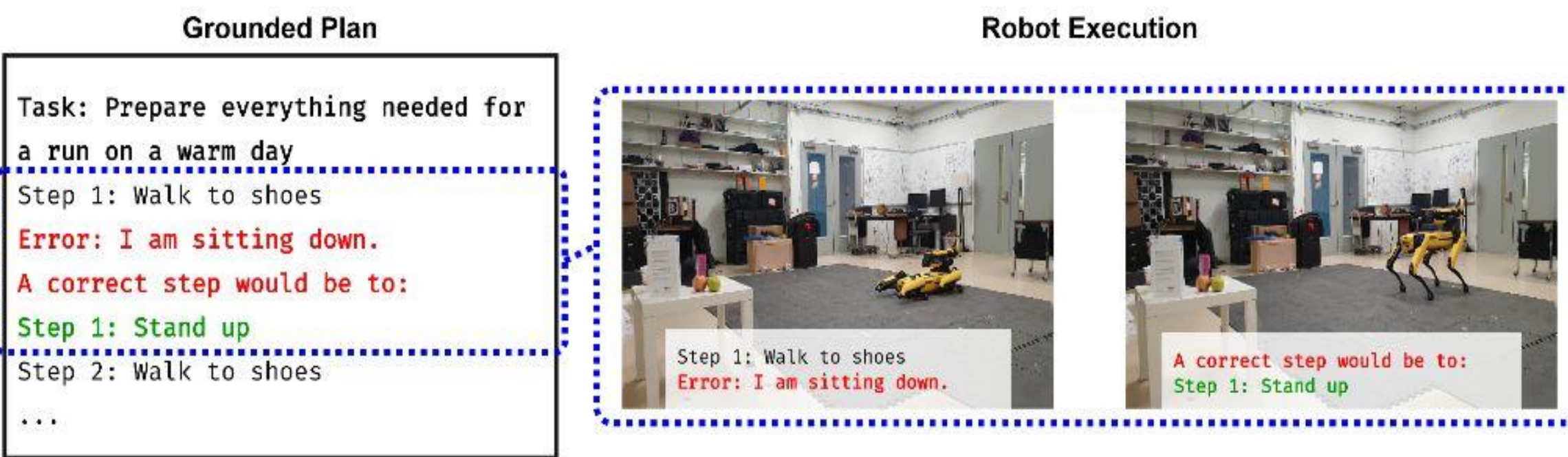


### Error Decomposition

| Missing Object (41.7%) | Unflipped Boolean (4.2%) |
|---|---|
| Empty Program (16.7%) | Absent from Room |
| Enclosed Object (12.5%) | Agent Over-occupied |
| Field of View (8.3%) | Enclosed Action |
| Agent Proximity (8.3%) | Other Pre-conditions |

Shreyas Raman, Vanya Cohen, David Paulius, Ifrah Idrees, Eric Rosen, Ray Mooney, Stefanie Tellex