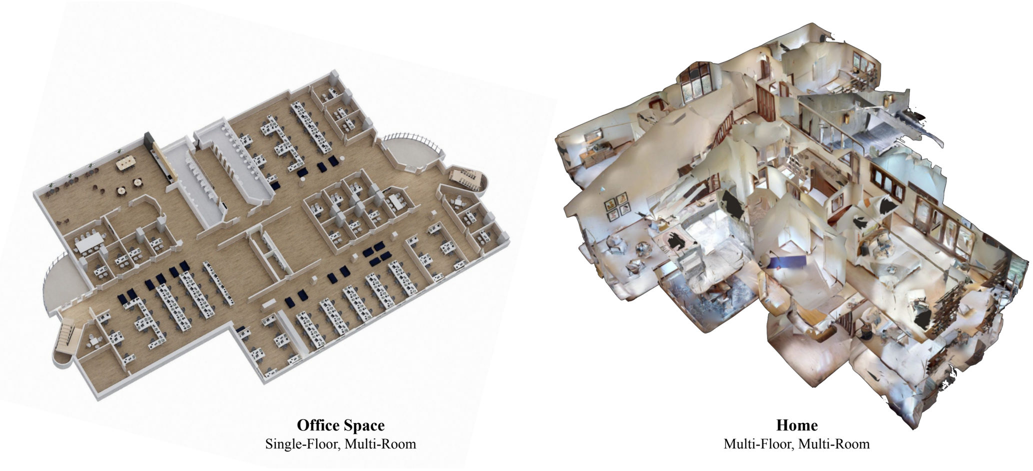
A Environments


Office Space
Single-Floor, Multi-Room

Home
Multi-Floor, Multi-Room

Figure 3: **Large-scale environments used to evaluate SayPlan.**

We evaluate SayPlan across a set of two large-scale environments spanning multiple rooms and floors as shown in Figure 3. We provide details of each of these environments below:

Office: A large-scale office floor, spanning 36 rooms and 150 assets and objects which the agent can interact with. This scene graph represents a real-world office floor within which a mobile manipulator robot is present. This allows us to embody the plans generated using SayPlan and evaluate their feasibility in the corresponding environment. A full visual 3D scene graph representation of the office floor is provided in Appendix C and D. We utilise generic names across our scene to maintain anonymity throughout the review process.

Home: An existing 3D scene graph from the Stanford 3D Scene Graph dataset [11] which consists of a family home environment (`Klickitat`) spanning 32 rooms across 3 floors and contains 121 assets and objects which the agent can interact with. A 3D visual of this environment can be viewed at the 3D Scene Graph project website.[2]

B Tasks

We evaluate SayPlan across 4 instruction sets which are classified to evaluate different aspects of its 3D scene graph reasoning and planning capabilities:

Simple Search: Focused on evaluating the semantic search capabilities of the LLM based on queries which directly reference information in the scene graph as well as the basic graph-based reasoning capabilities of the LMM.
Complex Search: Abstract semantic search queries which require complex reasoning. The information required to solve these search tasks is not readily available in the graph and has to be inferred by the underlying LLM.
Simple Planning: Task planning queries which require the agent to perform graph search, causal reasoning and environment interaction in order to solve the task. Typically requires shorter horizon plans over single rooms.
Long Horizon Planning: Long Horizon planning queries require multiple interactive steps. These queries evaluate SayPlan's ability to reason over temporally extended instructions to investigate how well it scales to such regimes. Typically requires long horizon plans spanning multiple rooms.

The full list of instructions used and the corresponding aspect the query evaluates are given in the following tables.

---

[2] 3dscenegraph.stanford.edu/Klickitat

B.1 Simple Search

B.1.1 Office Environment

Instruction	
Find me object K31X.	▷ unguided search with no semantic cue
Find me a carrot.	▷ semantic search based on node name
Find me anything purple in the postdoc bays.	▷ semantic search with termination conditioned on attribute
Find me a ripe banana.	▷ semantic search with termination conditioned on attribute
Find me something that has a screwdriver in it.	▷ unguided search with termination conditioned on children
One of the offices has a poster of the Terminator. Which one is it?	▷ semantic search with termination conditioned on children
I printed a document but I don't know which printer has it. Find the document.	▷ semantic search based on parent
I left my headphones in one of the meeting rooms. Locate them.	▷ semantic search based on parent
Find the PhD bay that has a drone in it.	▷ semantic search with termination conditioned on children
Find the kale that is not in the kitchen.	▷ semantic search with termination conditioned on a negation predicate on parent
Find me an office that does not have a cabinet.	▷ semantic search with termination conditioned on a negation predicate on children
Find me an office that contains a cabinet, a desk, and a chair.	▷ semantic search with termination conditioned on a conjunctive query on children
Find a book that was left next to a robotic gripper.	▷ semantic search with termination conditioned on a sibling
Luis gave one of his neighbours a stapler. Find the stapler.	▷ semantic search with termination conditioned on a sibling
There is a meeting room with a chair but no table. Locate it.	▷ semantic search with termination conditioned on a conjunctive query with negation

Table 5: **Simple Search Instructions.** Evaluated in Office Environment.

499 ### B.1.2 Home Environment

Instruction	
Find me a FooBar.	▷ unguided search with no semantic cue
Find me a bottle of wine.	▷ semantic search based on node name
Find me a plant with thorns.	▷ semantic search with termination conditioned on attribute
Find me a plant that needs watering.	▷ semantic search with termination conditioned on attribute
Find me a bathroom with no toilet.	▷ semantic search with termination conditioned on a negation predicate
The baby dropped their rattle in one of the rooms. Locate it.	▷ semantic search based on node name
I left my suitcase either in the bedroom or the living room. Which room is it in.	▷ semantic search based on node name
Find the room with a ball in it.	▷ semantic search based on node name
I forgot my book on a bed. Locate it.	▷ semantic search based on node name
Find an empty vase that was left next to sink.	▷ semantic search with termination conditioned on sibling
Locate the dining room which has a table, chair and a baby monitor.	▷ semantic search with termination conditioned on conjuctive query
Locate a chair that is not in any dining room.	▷ semantic search with termination conditioned on negation predicate
I need to shave. Which room has both a razor and shaving cream.	▷ semantic search with termination conditioned on children
Find me 2 bedrooms with pillows in them.	▷ semantic search with multiple returns
Find me 2 bedrooms without pillows in them.	▷ semantic search with multiple returns based on negation predicate

Table 6: **Simple Search Instructions.** Evaluated in Home Environment.

B.2 Complex Search

B.2.1 Office Environment

Instruction	
Find object J64M. J64M should be kept at below 0 degrees Celsius.	▷ semantic search guided by implicit world knowledge (knowledge not directly encoded in graph)
Find me something non vegetarian.	▷ semantic search with termination conditioned on implicit world knowledge
Locate something sharp.	▷ unguided search with termination conditioned on implicit world knowledge
Find the room where people are playing board games.	▷ semantic search with termination conditioned on ability to deduce context from node children using world knowledge ("board game" is not part of any node name or attribute in this graph)
Find an office of someone who is clearly a fan of Arnold Schwarzenegger.	▷ semantic search with termination conditioned on ability to deduce context from node children using world knowledge
There is a postdoc that has a pet Husky. Find the desk that's most likely theirs.	▷ semantic search with termination conditioned on ability to deduce context from node children using world knowledge
One of the PhD students was given more than one complimentary T-shirts. Find his desk.	▷ semantic search with termination conditioned on the number of children
Find me the office where a paper attachment device is inside an asset that is open.	▷ semantic search with termination conditioned on node descendants and their attributes
There is an office which has a cabinet containing exactly 3 items in it. Locate the office.	▷ semantic search with termination conditioned on the number of children
There is an office which has a cabinet containing a rotten apple. The cabinet name contains an even number. Locate the office.	▷ semantic search guided by numerical properties
Look for a carrot. The carrot is likely to be in a meeting room but I'm not sure.	▷ semantic search guided by used provided bias
Find me a meeting room with a RealSense camera.	▷ semantic search that has no result (no meeting room has a realsense camera in the graph)
Find the closest fire extinguisher to the manipulation lab.	▷ search guided by node distance
Find me the closest meeting room to the kitchen.	▷ search guided by node distance
Either Filipe or Tobi has my headphones. Locate it.	▷ evaluating constrained search, early termination once the two office are explored

Table 7: **Complex Search Instructions.** Evaluated in Office Environment.

B.2.2 Home Environment

Instruction	
I need something to access ChatGPT. Where should I go?	▷ semantic search guided by implicit world knowledge
Find the livingroom that contains the most electronic devices.	▷ semantic search with termination conditioned on children with indirect information
Find me something to eat with a lot of potassium.	▷ semantic search with termination conditioned on implicit world knowledge
I left a sock in a bedroom and one in the living room. Locate them. They should match.	▷ semantic search with multiple returns
Find me a potted plant that is most likely a cactus.	▷ semantic search with termination implicitly conditioned on attribute
Find the dining room with exactly 5 chairs.	▷ semantic search with termination implicitly conditioned on quantity of children
Find me the bedroom closest to the home office.	▷ semantic search with termination implicitly conditioned on node distance
Find me a bedroom with an unusual amount of bowls.	▷ semantic search with termination implicitly conditioned on quantity of children
Which bedroom is empty.	▷ semantic search with termination implicitly conditioned on quantity of children
Which bathroom has the most potted plants.	▷ semantic search with termination implicitly conditioned on quantity of children
The kitchen is flooded. Find somewhere I can heat up my food.	▷ semantic search guided by negation
Find me the room which most likely belongs to a child	▷ semantic search with termination conditioned on ability to deduce context from node children using world knowledge
15 guests are arriving. Locate enough chairs to seat them.	▷ semantic search with multiple returns
A vegetarian dinner was prepared in one of the dining rooms. Locate it.	▷ semantic search with selection criteria based on world knowledge
My tie is in one of the closets. Locate it.	▷ evaluating constrained search, termination after exploring closests

Table 8: **Complex Search Instructions.** Evaluated in Home Environment.

B.3 Simple Planning

Instruction
Close Jason's cabinet.
Refrigerate the orange left on the kitchen bench.
Take care of the dirty plate in the lunchroom.
Place the printed document on Will's desk.
Peter is working hard at his desk. Get him a healthy snack.
Hide one of Peter's valuable belongings.
Wipe the dusty admin shelf.
There is coffee dripping on the floor. Stop it.
Place Will's drone on his desk.
Move the monitor from Jason's office to Filipe's.
My parcel just got delivered! Locate it and place it in the appropriate lab.
Check if the coffee machine is working.
Heat up the chicken kebab.
Something is smelling in the kitchen. Dispose of it.
Throw what the agent is holding in the bin.

Table 9: **Simple Planning Instructions.** Evaluated in Office Environment.

B.4 Long Horizon Planning

Instruction
Heat up the noodles in the fridge, and place it somewhere where I can enjoy it.
Throw the rotting fruit in Dimity's office in the correct bin.
Wash all the dishes on the lunch table. Once finished, place all the clean cutlery in the drawer.
Safely file away the freshly printed document in Will's office then place the undergraduate thesis on his desk.
Make Niko a coffee and place the mug on his desk.
Someone has thrown items in the wrong bins. Correct this.
Tobi spilt soda on his desk. Throw away the can and take him something to clean with.
I want to make a sandwich. Place all the ingredients on the lunch table.
A delegation of project partners is arriving soon. We want to serve them snacks and non-alcoholic drinks. Prepare everything in the largest meeting room. Use items found in the supplies room only.
Serve bottled water to the attendees who are seated in meeting room 1. Each attendee can only receive a single bottle of water.
Empty the dishwasher. Place all items in their correct locations
Locate all 6 complimentary t-shirts given to the PhD students and place them on the shelf in admin.
I'm hungry. Bring me an apple from Peter and a pepsi from Tobi. I'm at the lunch table.
Let's play a prank on Niko. Dimity might have something.
There is an office which has a cabinet containing a rotten apple. The cabinet name contains an even number. Locate the office, throw away the fruit and get them a fresh apple.

Table 10: **Long-Horizon Planning Instructions.** Evaluated in Office Environment.

505 ## C Full 3D Scene Graph: Office Environment

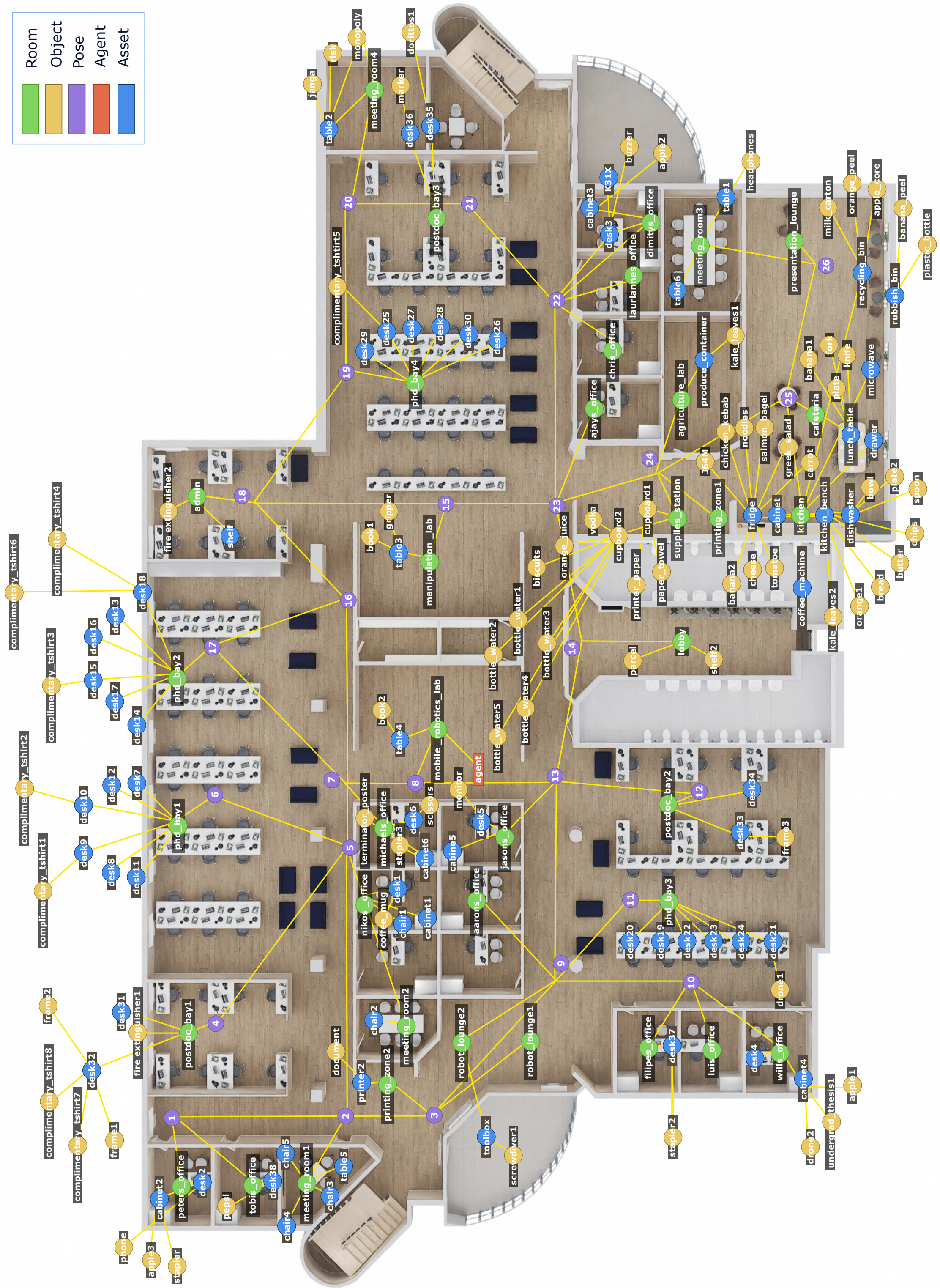


Figure 4: **3D Scene Graph - Fully Expanded Office Environment.** Full 3D scene graph exposing all the rooms, assets and objects available in the scene. Note that the LLM agent never sees all this information unless it chooses to expand every possible node without contraction.

D Contracted 3D Scene Graph: Office Environment

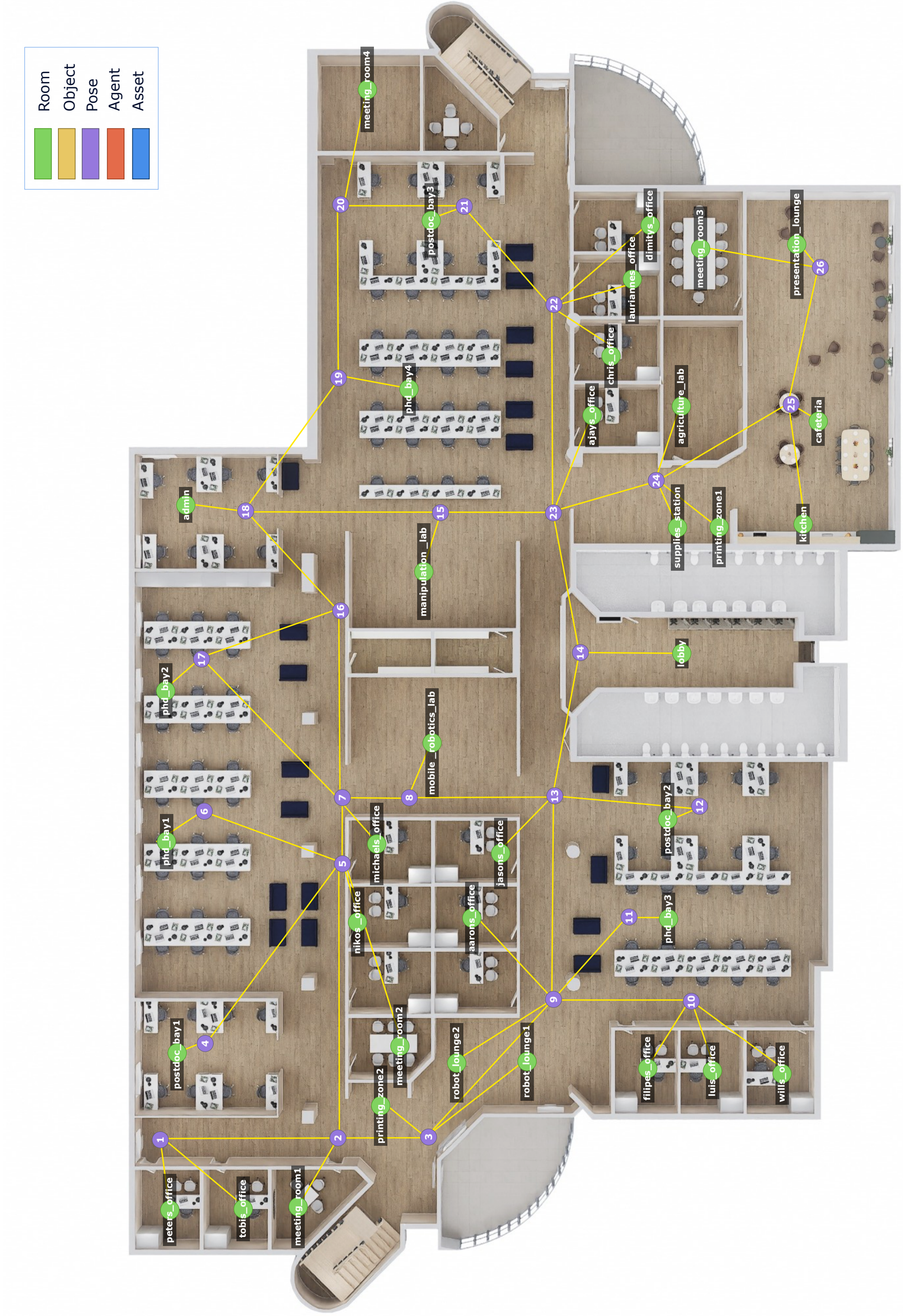
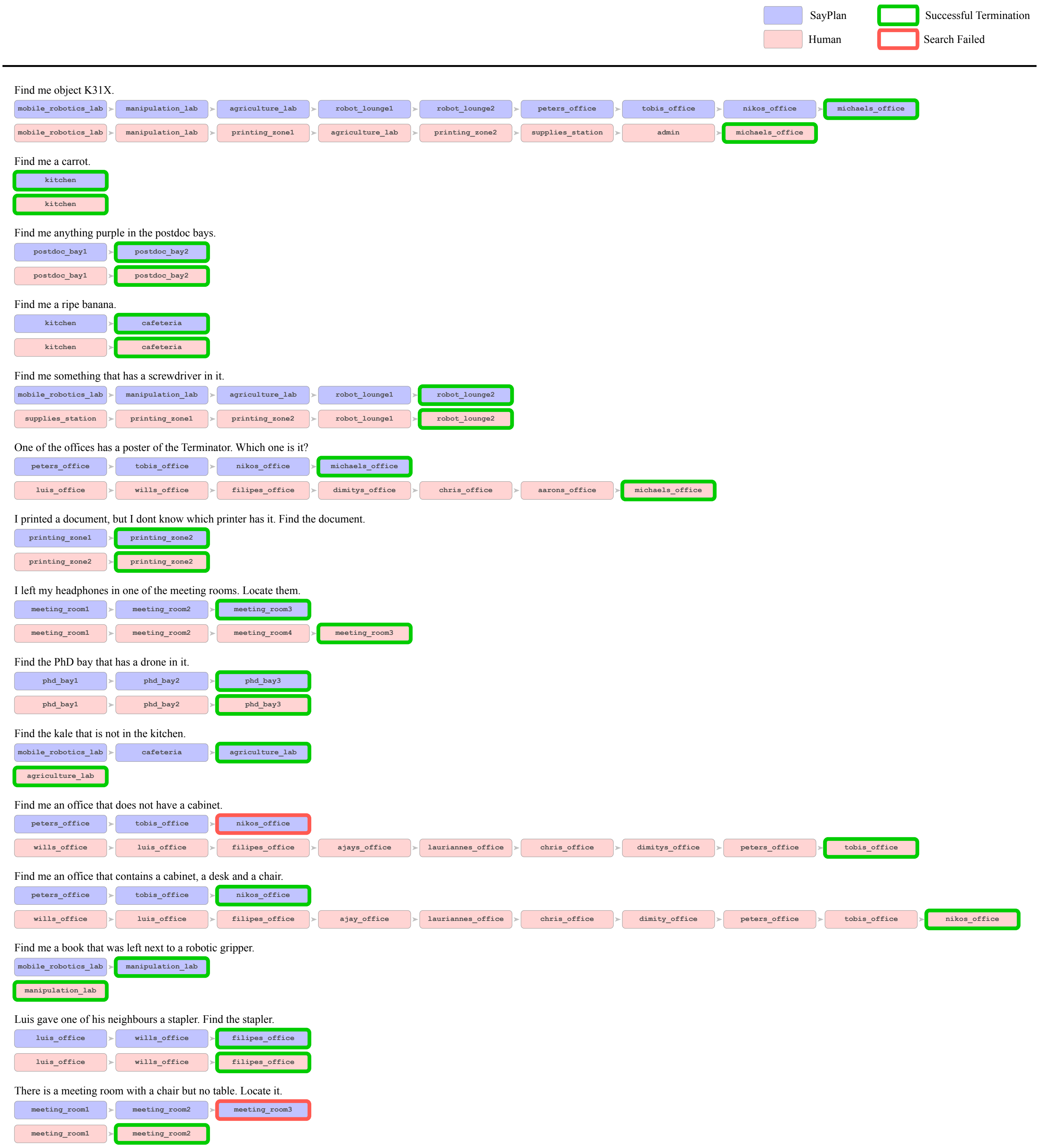


Figure 5: **3D Scene Graph - Contracted Office Environment.** Contracted 3D scene graph exposing only the highest level within the hierarchy - room nodes. This results in an 82.1% reduction in the number of tokens required to represent the scene before the exploration phase.

E Semantic Search Evaluation Results

- Full listings of the generated semantic search sequences for the evaluation instruction sets are provided on the following page -



Find me object K31X.
mobile_robotics_lab ▸ manipulation_lab ▸ agriculture_lab ▸ robot_lounge1 ▸ robot_lounge2 ▸ peters_office ▸ tobis_office ▸ nikos_office ▸ michaels_office
mobile_robotics_lab ▸ manipulation_lab ▸ printing_zone1 ▸ agriculture_lab ▸ printing_zone2 ▸ supplies_station ▸ admin ▸ michaels_office

Find me a carrot.
kitchen
kitchen

Find me anything purple in the postdoc bays.
postdoc_bay1 ▸ postdoc_bay2
postdoc_bay1 ▸ postdoc_bay2

Find me a ripe banana.
kitchen ▸ cafeteria
kitchen ▸ cafeteria

Find me something that has a screwdriver in it.
mobile_robotics_lab ▸ manipulation_lab ▸ agriculture_lab ▸ robot_lounge1 ▸ robot_lounge2
supplies_station ▸ printing_zone1 ▸ printing_zone2 ▸ robot_lounge1 ▸ robot_lounge2

One of the offices has a poster of the Terminator. Which one is it?
peters_office ▸ tobis_office ▸ nikos_office ▸ michaels_office
luis_office ▸ wills_office ▸ filipes_office ▸ dimitys_office ▸ chris_office ▸ aarons_office ▸ michaels_office

I printed a document, but I dont know which printer has it. Find the document.
printing_zone1 ▸ printing_zone2
printing_zone2 ▸ printing_zone2

I left my headphones in one of the meeting rooms. Locate them.
meeting_room1 ▸ meeting_room2 ▸ meeting_room3
meeting_room1 ▸ meeting_room2 ▸ meeting_room4 ▸ meeting_room3

Find the PhD bay that has a drone in it.
phd_bay1 ▸ phd_bay2 ▸ phd_bay3
phd_bay1 ▸ phd_bay2 ▸ phd_bay3

Find the kale that is not in the kitchen.
mobile_robotics_lab ▸ cafeteria ▸ agriculture_lab
agriculture_lab

Find me an office that does not have a cabinet.
peters_office ▸ tobis_office ▸ nikos_office
wills_office ▸ luis_office ▸ filipes_office ▸ ajays_office ▸ lauriannes_office ▸ chris_office ▸ dimitys_office ▸ peters_office ▸ tobis_office

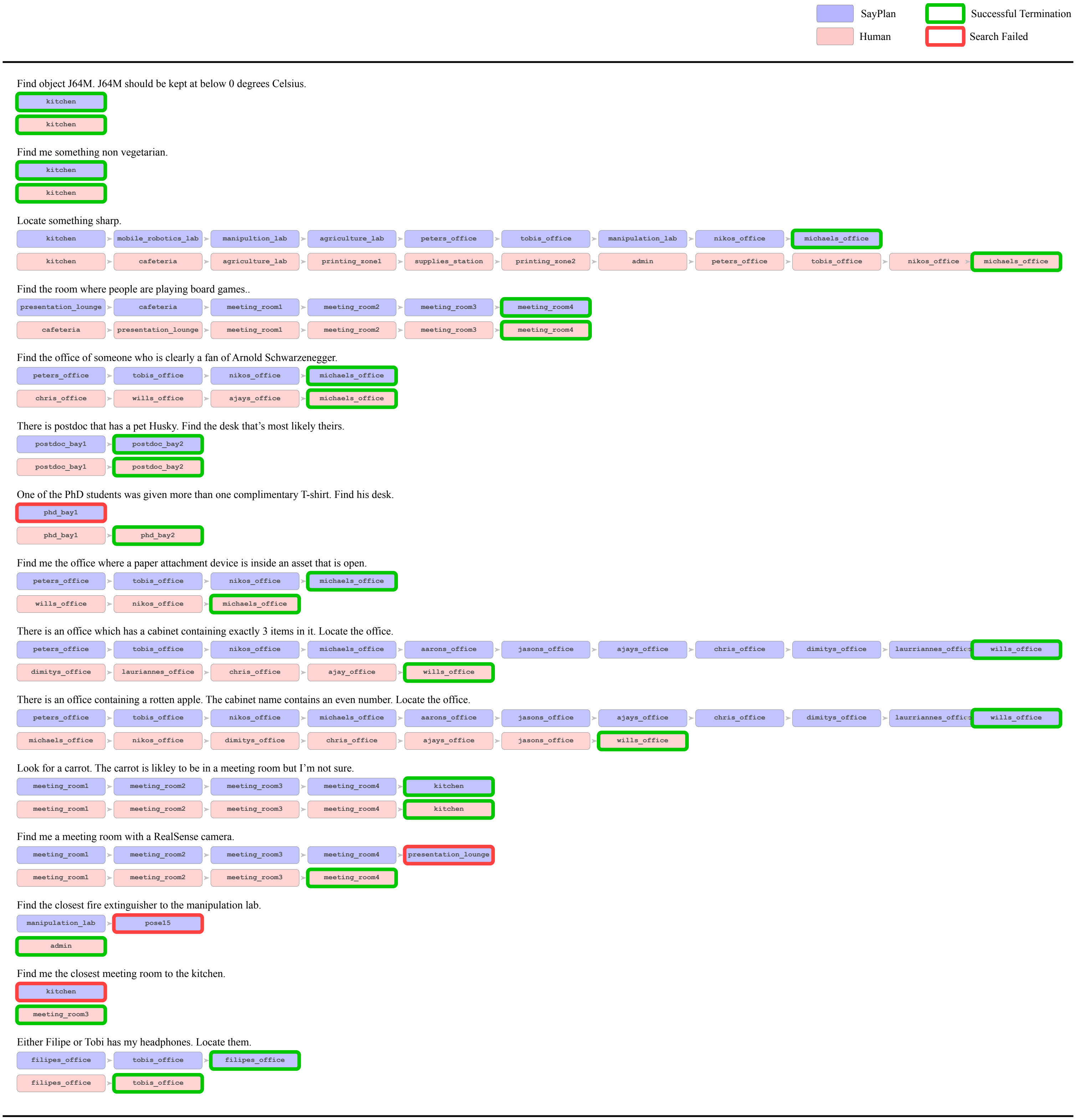
Find me an office that contains a cabinet, a desk and a chair.
peters_office ▸ tobis_office ▸ nikos_office
wills_office ▸ luis_office ▸ filipes_office ▸ ajay_office ▸ lauriannes_office ▸ chris_office ▸ dimity_office ▸ peters_office ▸ tobis_office ▸ nikos_office

Find me a book that was left next to a robotic gripper.
mobile_robotics_lab ▸ manipulation_lab
manipulation_lab

Luis gave one of his neighbours a stapler. Find the stapler.
luis_office ▸ wills_office ▸ filipes_office
luis_office ▸ wills_office ▸ filipes_office

There is a meeting room with a chair but no table. Locate it.
meeting_room1 ▸ meeting_room2 ▸ meeting_room3
meeting_room1 ▸ meeting_room2

Table 11: **Simple Search Office Environment Evaluation.** Sequence of Explored Nodes for Simple Search Office Environment Instructions.



🟦 SayPlan
🟧 Human
🟩 Successful Termination
🟥 Search Failed

Find object J64M. J64M should be kept at below 0 degrees Celsius.
🟦 🟩 kitchen
🟧 🟩 kitchen

Find me something non vegetarian.
🟦 🟩 kitchen
🟧 🟩 kitchen

Locate something sharp.
🟦 kitchen > mobile_robotics_lab > manipultion_lab > agriculture_lab > peters_office > tobis_office > manipulation_lab > nikos_office > 🟩 michaels_office
🟧 kitchen > cafeteria > agriculture_lab > printing_zone1 > supplies_station > printing_zone2 > admin > peters_office > tobis_office > nikos_office > 🟩 michaels_office

Find the room where people are playing board games..
🟦 presentation_lounge > cafeteria > meeting_room1 > meeting_room2 > meeting_room3 > 🟩 meeting_room4
🟧 cafeteria > presentation_lounge > meeting_room1 > meeting_room2 > meeting_room3 > 🟩 meeting_room4

Find the office of someone who is clearly a fan of Arnold Schwarzenegger.
🟦 peters_office > tobis_office > nikos_office > 🟩 michaels_office
🟧 chris_office > wills_office > ajays_office > 🟩 michaels_office

There is postdoc that has a pet Husky. Find the desk that's most likely theirs.
🟦 postdoc_bay1 > 🟩 postdoc_bay2
🟧 postdoc_bay1 > 🟩 postdoc_bay2

One of the PhD students was given more than one complimentary T-shirt. Find his desk.
🟦 🟥 phd_bay1
🟧 phd_bay1 > 🟩 phd_bay2

Find me the office where a paper attachment device is inside an asset that is open.
🟦 peters_office > tobis_office > nikos_office > 🟩 michaels_office
🟧 wills_office > nikos_office > 🟩 michaels_office

There is an office which has a cabinet containing exactly 3 items in it. Locate the office.
🟦 peters_office > tobis_office > nikos_office > michaels_office > aarons_office > jasons_office > ajays_office > chris_office > dimitys_office > laurriannes_offic > 🟩 wills_office
🟧 dimitys_office > lauriannes_office > chris_office > ajay_office > 🟩 wills_office

There is an office containing a rotten apple. The cabinet name contains an even number. Locate the office.
🟦 peters_office > tobis_office > nikos_office > michaels_office > aarons_office > jasons_office > ajays_office > chris_office > dimitys_office > laurriannes_offic > 🟩 wills_office
🟧 michaels_office > nikos_office > dimitys_office > chris_office > ajays_office > jasons_office > 🟩 wills_office

Look for a carrot. The carrot is likley to be in a meeting room but I'm not sure.
🟦 meeting_room1 > meeting_room2 > meeting_room3 > meeting_room4 > 🟩 kitchen
🟧 meeting_room1 > meeting_room2 > meeting_room3 > meeting_room4 > 🟩 kitchen

Find me a meeting room with a RealSense camera.
🟦 meeting_room1 > meeting_room2 > meeting_room3 > meeting_room4 > 🟥 presentation_lounge
🟧 meeting_room1 > meeting_room2 > meeting_room3 > 🟩 meeting_room4

Find the closest fire extinguisher to the manipulation lab.
🟦 manipulation_lab > 🟥 pose15
🟧 🟩 admin

Find me the closest meeting room to the kitchen.
🟦 🟥 kitchen
🟧 🟩 meeting_room3

Either Filipe or Tobi has my headphones. Locate them.
🟦 filipes_office > tobis_office > 🟩 filipes_office
🟧 filipes_office > 🟩 tobis_office

Table 12: **Complex Search Office Environment Evaluation.** Sequence of Explored Nodes for Complex Search Office Environment Instructions.

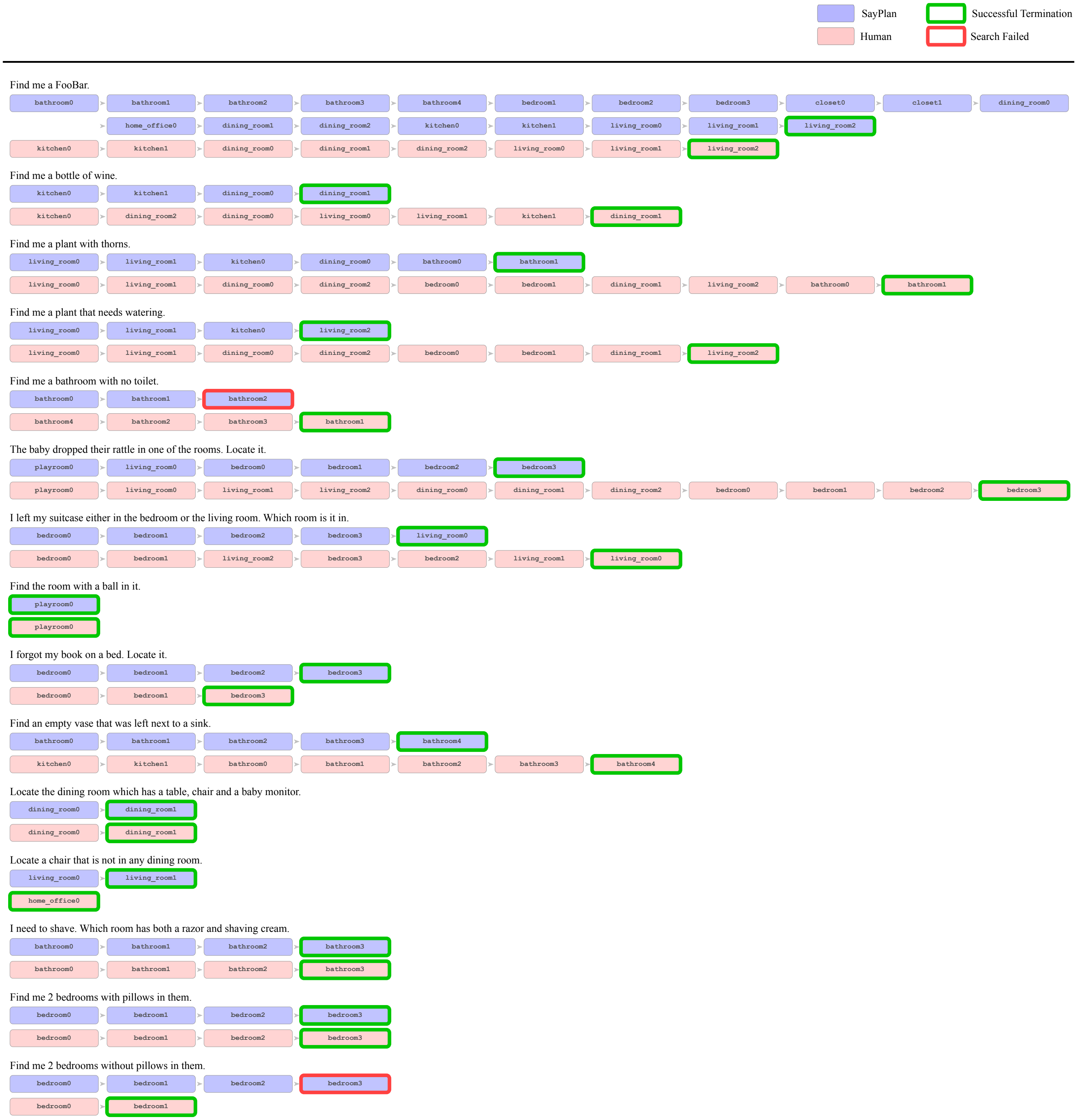
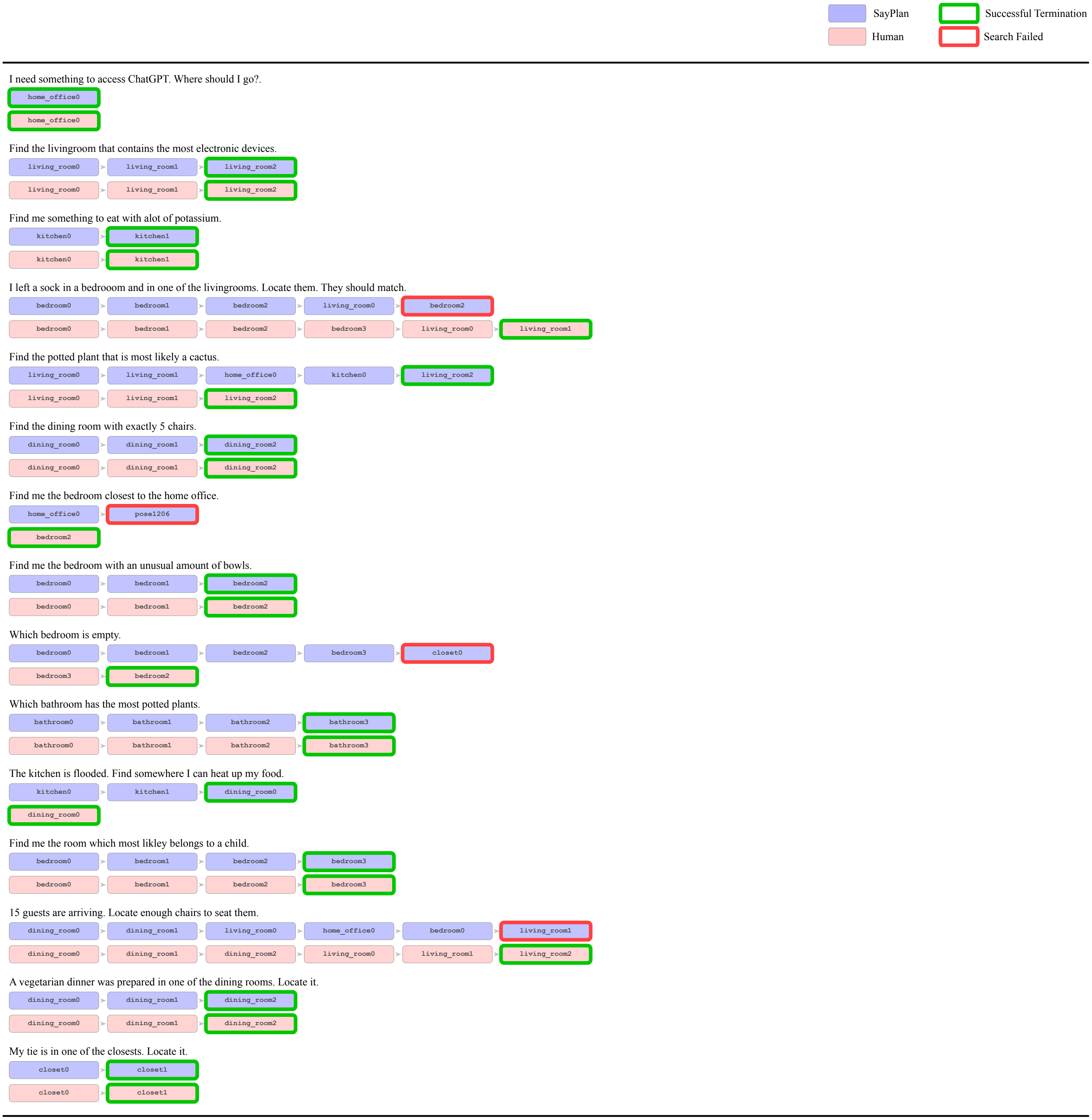

SayPlan
Human
Successful Termination
Search Failed
Find me a FooBar.
bathroom0 > bathroom1 > bathroom2 > bathroom3 > bathroom4 > bedroom1 > bedroom2 > bedroom3 > closet0 > closet1 > dining_room0
> home_office0 > dining_room1 > dining_room2 > kitchen0 > kitchen1 > living_room0 > living_room1 > living_room2
kitchen0 > kitchen1 > dining_room0 > dining_room1 > dining_room2 > living_room0 > living_room1 > living_room2
Find me a bottle of wine.
kitchen0 > kitchen1 > dining_room0 > dining_room1
kitchen0 > dining_room2 > dining_room0 > living_room0 > living_room1 > kitchen1 > dining_room1
Find me a plant with thorns.
living_room0 > living_room1 > kitchen0 > dining_room0 > bathroom0 > bathroom1
living_room0 > living_room1 > dining_room0 > dining_room2 > bedroom0 > bedroom1 > dining_room1 > living_room2 > bathroom0 > bathroom1
Find me a plant that needs watering.
living_room0 > living_room1 > kitchen0 > living_room2
living_room0 > living_room1 > dining_room0 > dining_room2 > bedroom0 > bedroom1 > dining_room1 > living_room2
Find me a bathroom with no toilet.
bathroom0 > bathroom1 > bathroom2
bathroom4 > bathroom2 > bathroom3 > bathroom1
The baby dropped their rattle in one of the rooms. Locate it.
playroom0 > living_room0 > bedroom0 > bedroom1 > bedroom2 > bedroom3
playroom0 > living_room0 > living_room1 > living_room2 > dining_room0 > dining_room1 > dining_room2 > bedroom0 > bedroom1 > bedroom2 > bedroom3
I left my suitcase either in the bedroom or the living room. Which room is it in.
bedroom0 > bedroom1 > bedroom2 > bedroom3 > living_room0
bedroom0 > bedroom1 > living_room2 > bedroom3 > bedroom2 > living_room1 > living_room0
Find the room with a ball in it.
playroom0
playroom0
I forgot my book on a bed. Locate it.
bedroom0 > bedroom1 > bedroom2 > bedroom3
bedroom0 > bedroom1 > bedroom3
Find an empty vase that was left next to a sink.
bathroom0 > bathroom1 > bathroom2 > bathroom3 > bathroom4
kitchen0 > kitchen1 > bathroom0 > bathroom1 > bathroom2 > bathroom3 > bathroom4
Locate the dining room which has a table, chair and a baby monitor.
dining_room0 > dining_room1
dining_room0 > dining_room1
Locate a chair that is not in any dining room.
living_room0 > living_room1
home_office0
I need to shave. Which room has both a razor and shaving cream.
bathroom0 > bathroom1 > bathroom2 > bathroom3
bathroom0 > bathroom1 > bathroom2 > bathroom3
Find me 2 bedrooms with pillows in them.
bedroom0 > bedroom1 > bedroom2 > bedroom3
bedroom0 > bedroom1 > bedroom2 > bedroom3
Find me 2 bedrooms without pillows in them.
bedroom0 > bedroom1 > bedroom2 > bedroom3
bedroom0 > bedroom1

Table 13: **Simple Search Home Environment Evaluation.** Sequence of Explored Nodes for Simple Search Home Environment Instructions.



🟦 SayPlan
▨ Human
🟩 Successful Termination
🟥 Search Failed

I need something to access ChatGPT. Where should I go?.

🟦 home_office0 🟩

▨ home_office0 🟩

Find the livingroom that contains the most electronic devices.

🟦 living_room0 → living_room1 → living_room2 🟩

▨ living_room0 → living_room1 → living_room2 🟩

Find me something to eat with alot of potassium.

🟦 kitchen0 → kitchen1 🟩

▨ kitchen0 → kitchen1 🟩

I left a sock in a bedrooom and in one of the livingrooms. Locate them. They should match.

🟦 bedroom0 → bedroom1 → bedroom2 → living_room0 → bedroom2 🟥

▨ bedroom0 → bedroom1 → bedroom2 → bedroom3 → living_room0 → living_room1 🟩

Find the potted plant that is most likely a cactus.

🟦 living_room0 → living_room1 → home_office0 → kitchen0 → living_room2 🟩

▨ living_room0 → living_room1 → living_room2 🟩

Find the dining room with exactly 5 chairs.

🟦 dining_room0 → dining_room1 → dining_room2 🟩

▨ dining_room0 → dining_room1 → dining_room2 🟩

Find me the bedroom closest to the home office.

🟦 home_office0 → pose1206 🟥

▨ bedroom2 🟩

Find me the bedroom with an unusual amount of bowls.

🟦 bedroom0 → bedroom1 → bedroom2 🟩

▨ bedroom0 → bedroom1 → bedroom2 🟩

Which bedroom is empty.

🟦 bedroom0 → bedroom1 → bedroom2 → bedroom3 → closet0 🟥

▨ bedroom3 → bedroom2 🟩

Which bathroom has the most potted plants.

🟦 bathroom0 → bathroom1 → bathroom2 → bathroom3 🟩

▨ bathroom0 → bathroom1 → bathroom2 → bathroom3 🟩

The kitchen is flooded. Find somewhere I can heat up my food.

🟦 kitchen0 → kitchen1 → dining_room0 🟩

▨ dining_room0 🟩

Find me the room which most likley belongs to a child.

🟦 bedroom0 → bedroom1 → bedroom2 → bedroom3 🟩

▨ bedroom0 → bedroom1 → bedroom2 → bedroom3 🟩

15 guests are arriving. Locate enough chairs to seat them.

🟦 dining_room0 → dining_room1 → living_room0 → home_office0 → bedroom0 → living_room1 🟥

▨ dining_room0 → dining_room1 → dining_room2 → living_room0 → living_room1 → living_room2 🟩

A vegetarian dinner was prepared in one of the dining rooms. Locate it.

🟦 dining_room0 → dining_room1 → dining_room2 🟩

▨ dining_room0 → dining_room1 → dining_room2 🟩

My tie is in one of the closests. Locate it.

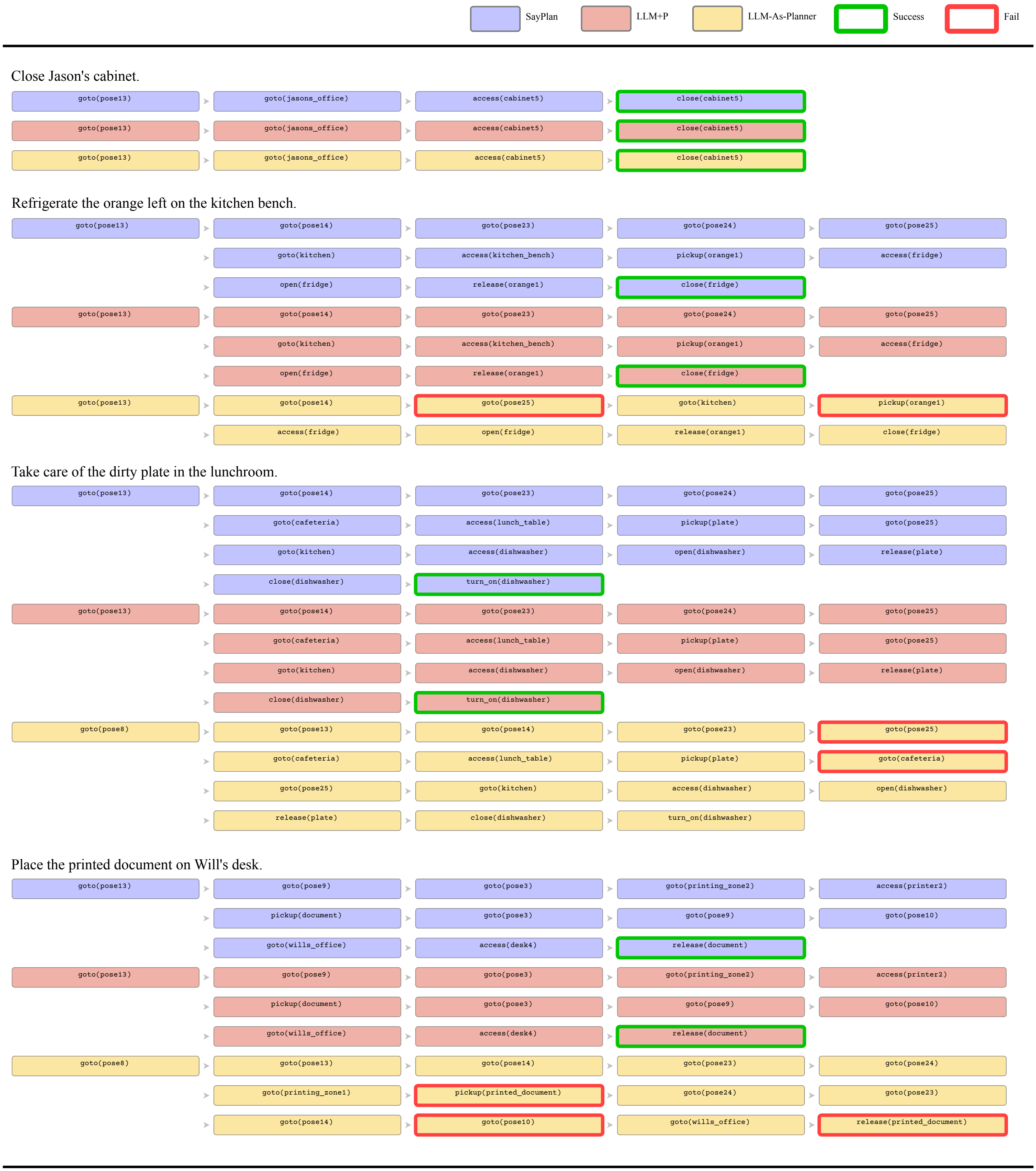
🟦 closet0 → closet1 🟩

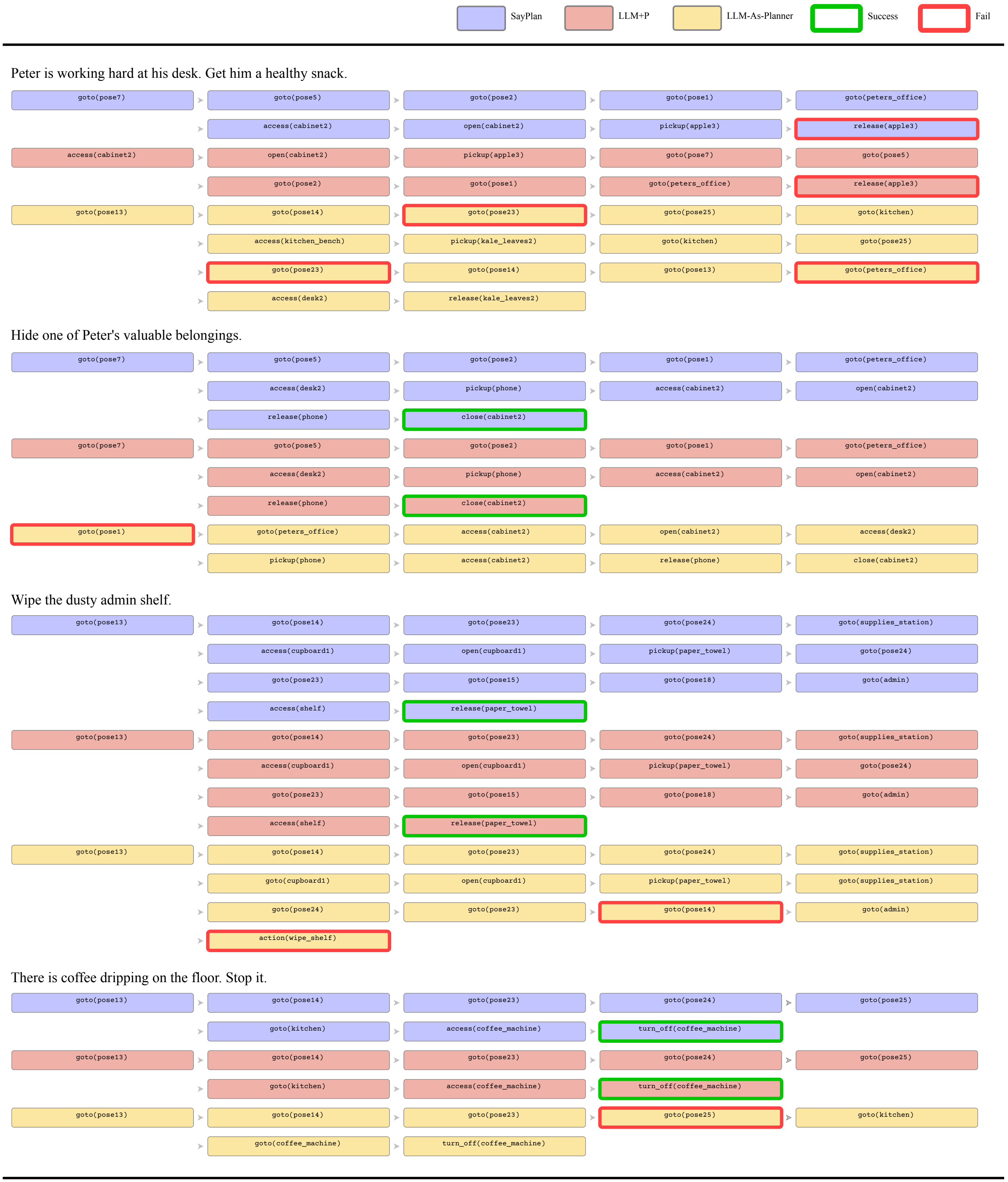
▨ closet0 → closet1 🟩

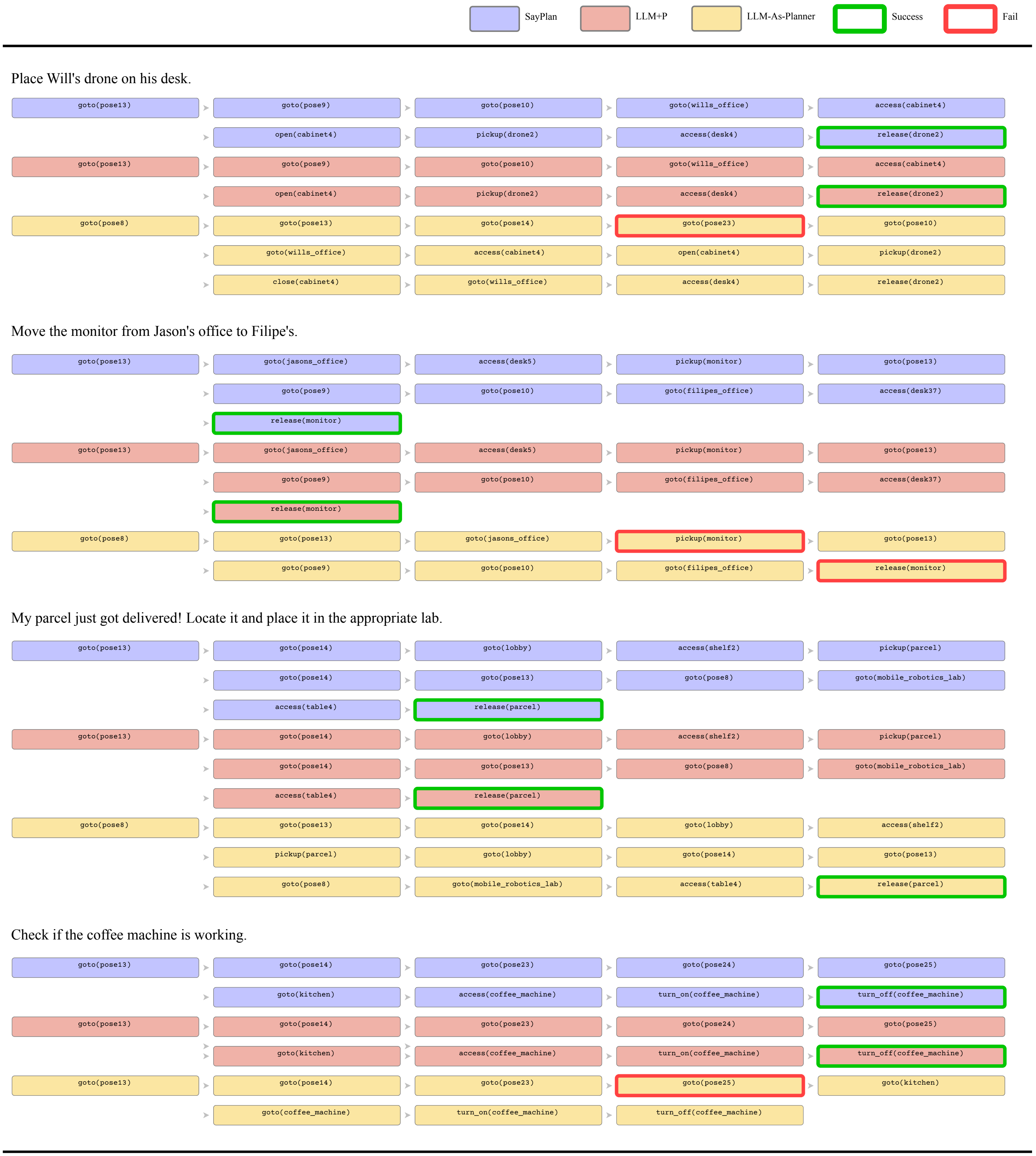
Table 14: **Complex Search Home Environment Evaluation.** Sequence of Explored Nodes for Complex Search Home Environment Instructions.

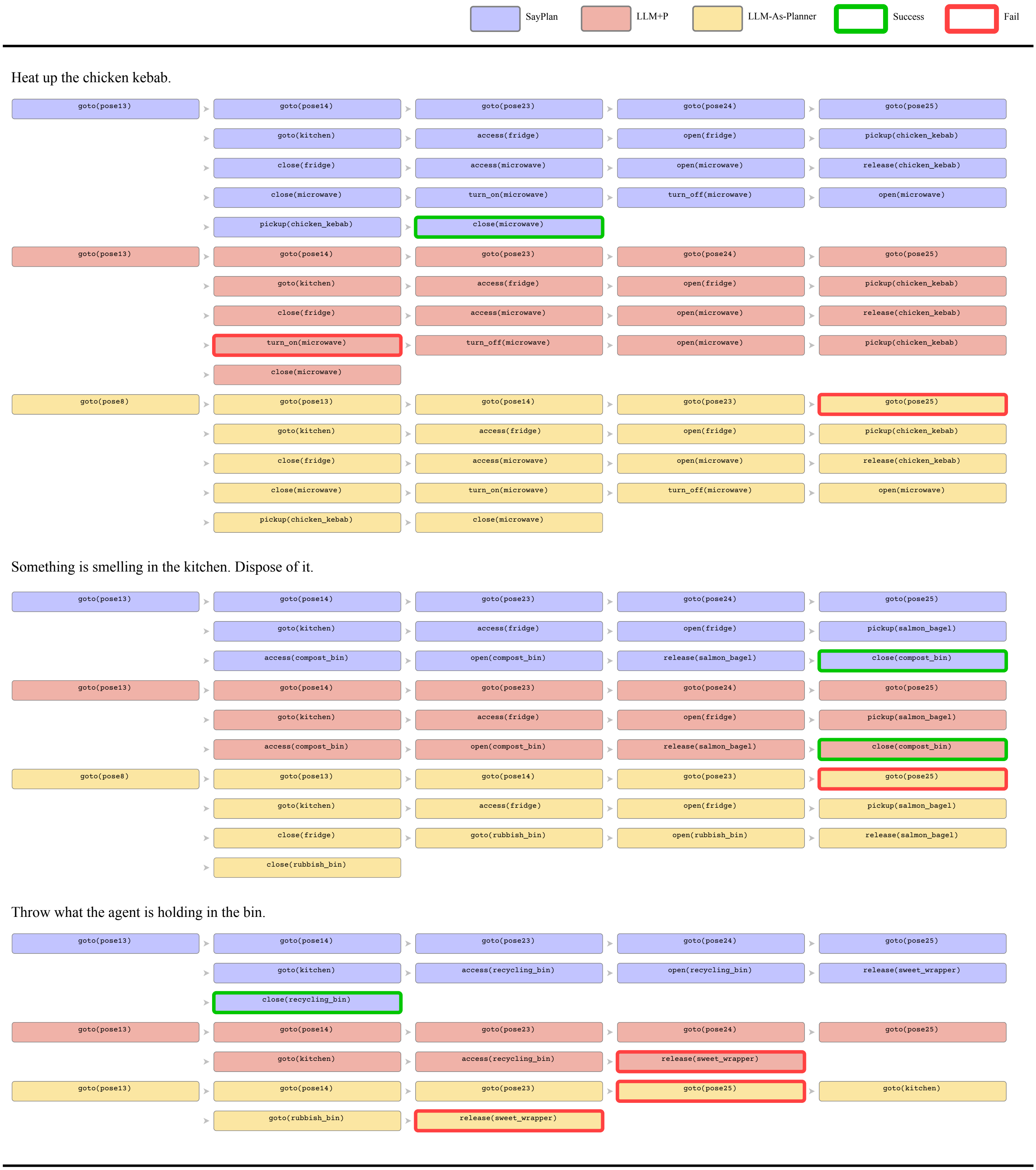
F Causal Planning Evaluation Results

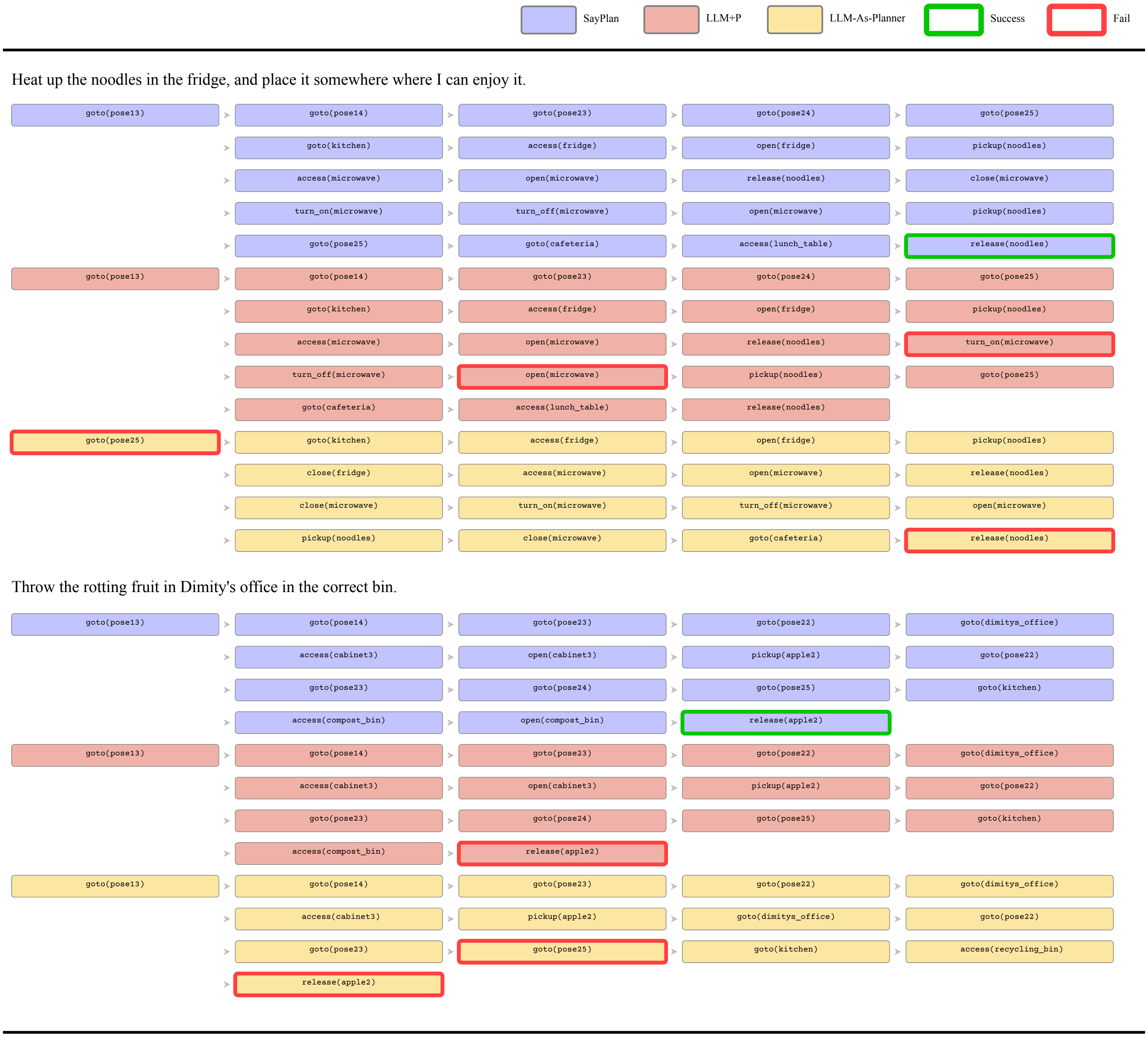
- Full listings of the generated planning sequences for the evaluation instruction sets are provided on the following page -

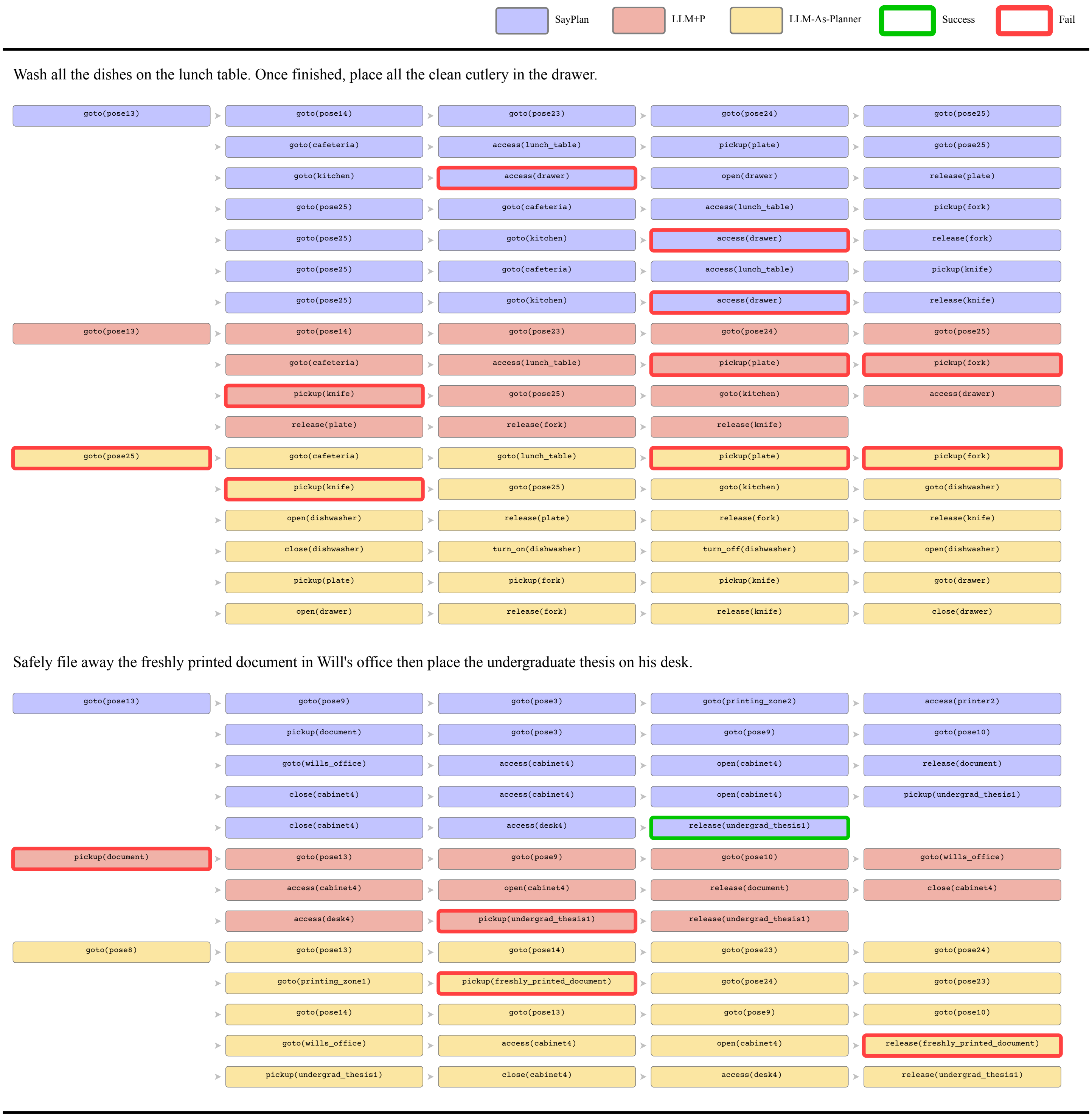

SayPlan
LLM+P
LLM-As-Planner
Success
Fail
Close Jason's cabinet.
goto(pose13)
goto(jasons_office)
access(cabinet5)
close(cabinet5)
goto(pose13)
goto(jasons_office)
access(cabinet5)
close(cabinet5)
goto(pose13)
goto(jasons_office)
access(cabinet5)
close(cabinet5)
Refrigerate the orange left on the kitchen bench.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(orange1)
access(fridge)
open(fridge)
release(orange1)
close(fridge)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(orange1)
access(fridge)
open(fridge)
release(orange1)
close(fridge)
goto(pose13)
goto(pose14)
goto(pose25)
goto(kitchen)
pickup(orange1)
access(fridge)
open(fridge)
release(orange1)
close(fridge)
Take care of the dirty plate in the lunchroom.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(cafeteria)
access(lunch_table)
pickup(plate)
goto(pose25)
goto(kitchen)
access(dishwasher)
open(dishwasher)
release(plate)
close(dishwasher)
turn_on(dishwasher)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(cafeteria)
access(lunch_table)
pickup(plate)
goto(pose25)
goto(kitchen)
access(dishwasher)
open(dishwasher)
release(plate)
close(dishwasher)
turn_on(dishwasher)
goto(pose8)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(cafeteria)
access(lunch_table)
pickup(plate)
goto(cafeteria)
goto(pose25)
goto(kitchen)
access(dishwasher)
open(dishwasher)
release(plate)
close(dishwasher)
turn_on(dishwasher)
Place the printed document on Will's desk.
goto(pose13)
goto(pose9)
goto(pose3)
goto(printing_zone2)
access(printer2)
pickup(document)
goto(pose3)
goto(pose9)
goto(pose10)
goto(wills_office)
access(desk4)
release(document)
goto(pose13)
goto(pose9)
goto(pose3)
goto(printing_zone2)
access(printer2)
pickup(document)
goto(pose3)
goto(pose9)
goto(pose10)
goto(wills_office)
access(desk4)
release(document)
goto(pose8)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(printing_zone1)
pickup(printed_document)
goto(pose24)
goto(pose23)
goto(pose14)
goto(pose10)
goto(wills_office)
release(printed_document)


SayPlan
LLM+P
LLM-As-Planner
Success
Fail
Peter is working hard at his desk. Get him a healthy snack.
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
access(cabinet2)
open(cabinet2)
pickup(apple3)
release(apple3)
access(cabinet2)
open(cabinet2)
pickup(apple3)
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
release(apple3)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(kale_leaves2)
goto(kitchen)
goto(pose25)
goto(pose23)
goto(pose14)
goto(pose13)
goto(peters_office)
access(desk2)
release(kale_leaves2)
Hide one of Peter's valuable belongings.
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
access(desk2)
pickup(phone)
access(cabinet2)
open(cabinet2)
release(phone)
close(cabinet2)
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
access(desk2)
pickup(phone)
access(cabinet2)
open(cabinet2)
release(phone)
close(cabinet2)
goto(pose1)
goto(peters_office)
access(cabinet2)
open(cabinet2)
access(desk2)
pickup(phone)
access(cabinet2)
release(phone)
close(cabinet2)
Wipe the dusty admin shelf.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(supplies_station)
access(cupboard1)
open(cupboard1)
pickup(paper_towel)
goto(pose24)
goto(pose23)
goto(pose15)
goto(pose18)
goto(admin)
access(shelf)
release(paper_towel)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(supplies_station)
access(cupboard1)
open(cupboard1)
pickup(paper_towel)
goto(pose24)
goto(pose23)
goto(pose15)
goto(pose18)
goto(admin)
access(shelf)
release(paper_towel)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(supplies_station)
goto(cupboard1)
open(cupboard1)
pickup(paper_towel)
goto(supplies_station)
goto(pose24)
goto(pose23)
goto(pose14)
goto(admin)
action(wipe_shelf)
There is coffee dripping on the floor. Stop it.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(coffee_machine)
turn_off(coffee_machine)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(coffee_machine)
turn_off(coffee_machine)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
goto(coffee_machine)
turn_off(coffee_machine)


SayPlan
LLM+P
LLM-As-Planner
Success
Fail
Place Will's drone on his desk.
goto(pose13) goto(pose9) goto(pose10) goto(wills_office) access(cabinet4)
open(cabinet4) pickup(drone2) access(desk4) release(drone2)
goto(pose13) goto(pose9) goto(pose10) goto(wills_office) access(cabinet4)
open(cabinet4) pickup(drone2) access(desk4) release(drone2)
goto(pose8) goto(pose13) goto(pose14) goto(pose23) goto(pose10)
goto(wills_office) access(cabinet4) open(cabinet4) pickup(drone2)
close(cabinet4) goto(wills_office) access(desk4) release(drone2)
Move the monitor from Jason's office to Filipe's.
goto(pose13) goto(jasons_office) access(desk5) pickup(monitor) goto(pose13)
goto(pose9) goto(pose10) goto(filipes_office) access(desk37)
release(monitor)
goto(pose13) goto(jasons_office) access(desk5) pickup(monitor) goto(pose13)
goto(pose9) goto(pose10) goto(filipes_office) access(desk37)
release(monitor)
goto(pose8) goto(pose13) goto(jasons_office) pickup(monitor) goto(pose13)
goto(pose9) goto(pose10) goto(filipes_office) release(monitor)
My parcel just got delivered! Locate it and place it in the appropriate lab.
goto(pose13) goto(pose14) goto(lobby) access(shelf2) pickup(parcel)
goto(pose14) goto(pose13) goto(pose8) goto(mobile_robotics_lab)
access(table4) release(parcel)
goto(pose13) goto(pose14) goto(lobby) access(shelf2) pickup(parcel)
goto(pose14) goto(pose13) goto(pose8) goto(mobile_robotics_lab)
access(table4) release(parcel)
goto(pose8) goto(pose13) goto(pose14) goto(lobby) access(shelf2)
pickup(parcel) goto(lobby) goto(pose14) goto(pose13)
goto(pose8) goto(mobile_robotics_lab) access(table4) release(parcel)
Check if the coffee machine is working.
goto(pose13) goto(pose14) goto(pose23) goto(pose24) goto(pose25)
goto(kitchen) access(coffee_machine) turn_on(coffee_machine) turn_off(coffee_machine)
goto(pose13) goto(pose14) goto(pose23) goto(pose24) goto(pose25)
goto(kitchen) access(coffee_machine) turn_on(coffee_machine) turn_off(coffee_machine)
goto(pose13) goto(pose14) goto(pose23) goto(pose25) goto(kitchen)
goto(coffee_machine) turn_on(coffee_machine) turn_off(coffee_machine)

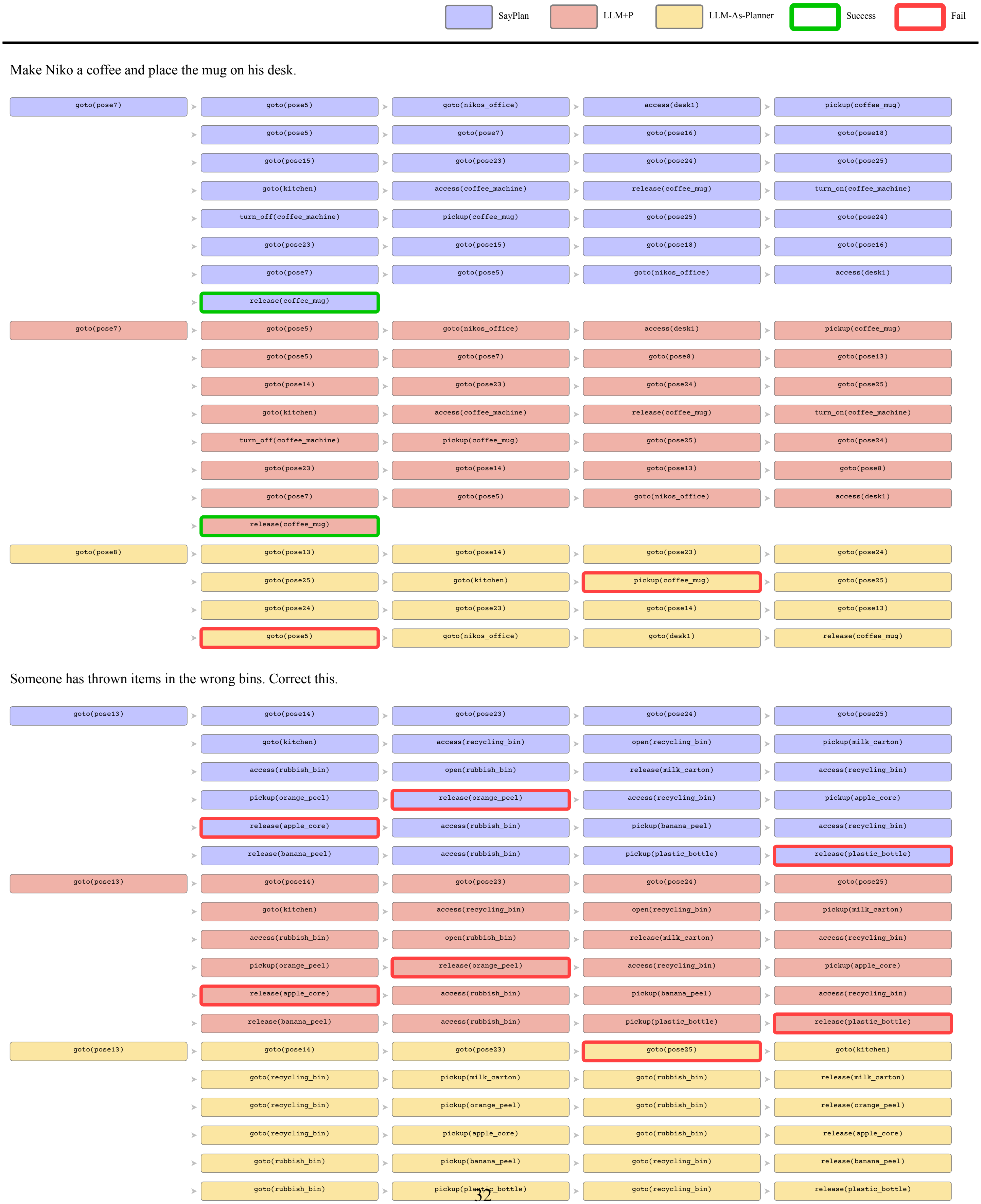

SayPlan
LLM+P
LLM-As-Planner
Success
Fail
Heat up the chicken kebab.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(chicken_kebab)
close(fridge)
access(microwave)
open(microwave)
release(chicken_kebab)
close(microwave)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(chicken_kebab)
close(microwave)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(chicken_kebab)
close(fridge)
access(microwave)
open(microwave)
release(chicken_kebab)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(chicken_kebab)
close(microwave)
goto(pose8)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(chicken_kebab)
close(fridge)
access(microwave)
open(microwave)
release(chicken_kebab)
close(microwave)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(chicken_kebab)
close(microwave)
Something is smelling in the kitchen. Dispose of it.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(salmon_bagel)
access(compost_bin)
open(compost_bin)
release(salmon_bagel)
close(compost_bin)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(salmon_bagel)
access(compost_bin)
open(compost_bin)
release(salmon_bagel)
close(compost_bin)
goto(pose8)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(salmon_bagel)
close(fridge)
goto(rubbish_bin)
open(rubbish_bin)
release(salmon_bagel)
close(rubbish_bin)
Throw what the agent is holding in the bin.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(recycling_bin)
open(recycling_bin)
release(sweet_wrapper)
close(recycling_bin)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(recycling_bin)
release(sweet_wrapper)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
goto(rubbish_bin)
release(sweet_wrapper)


SayPlan
LLM+P
LLM-As-Planner
Success
Fail
Heat up the noodles in the fridge, and place it somewhere where I can enjoy it.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(noodles)
access(microwave)
open(microwave)
release(noodles)
close(microwave)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(noodles)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(noodles)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(noodles)
access(microwave)
open(microwave)
release(noodles)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(noodles)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(noodles)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(noodles)
close(fridge)
access(microwave)
open(microwave)
release(noodles)
close(microwave)
turn_on(microwave)
turn_off(microwave)
open(microwave)
pickup(noodles)
close(microwave)
goto(cafeteria)
release(noodles)
Throw the rotting fruit in Dimity's office in the correct bin.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(cabinet3)
open(cabinet3)
pickup(apple2)
goto(pose22)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(compost_bin)
open(compost_bin)
release(apple2)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(cabinet3)
open(cabinet3)
pickup(apple2)
goto(pose22)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(compost_bin)
release(apple2)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(cabinet3)
pickup(apple2)
goto(dimitys_office)
goto(pose22)
goto(pose23)
goto(pose25)
goto(kitchen)
access(recycling_bin)
release(apple2)



Wash all the dishes on the lunch table. Once finished, place all the clean cutlery in the drawer.

Safely file away the freshly printed document in Will's office then place the undergraduate thesis on his desk.

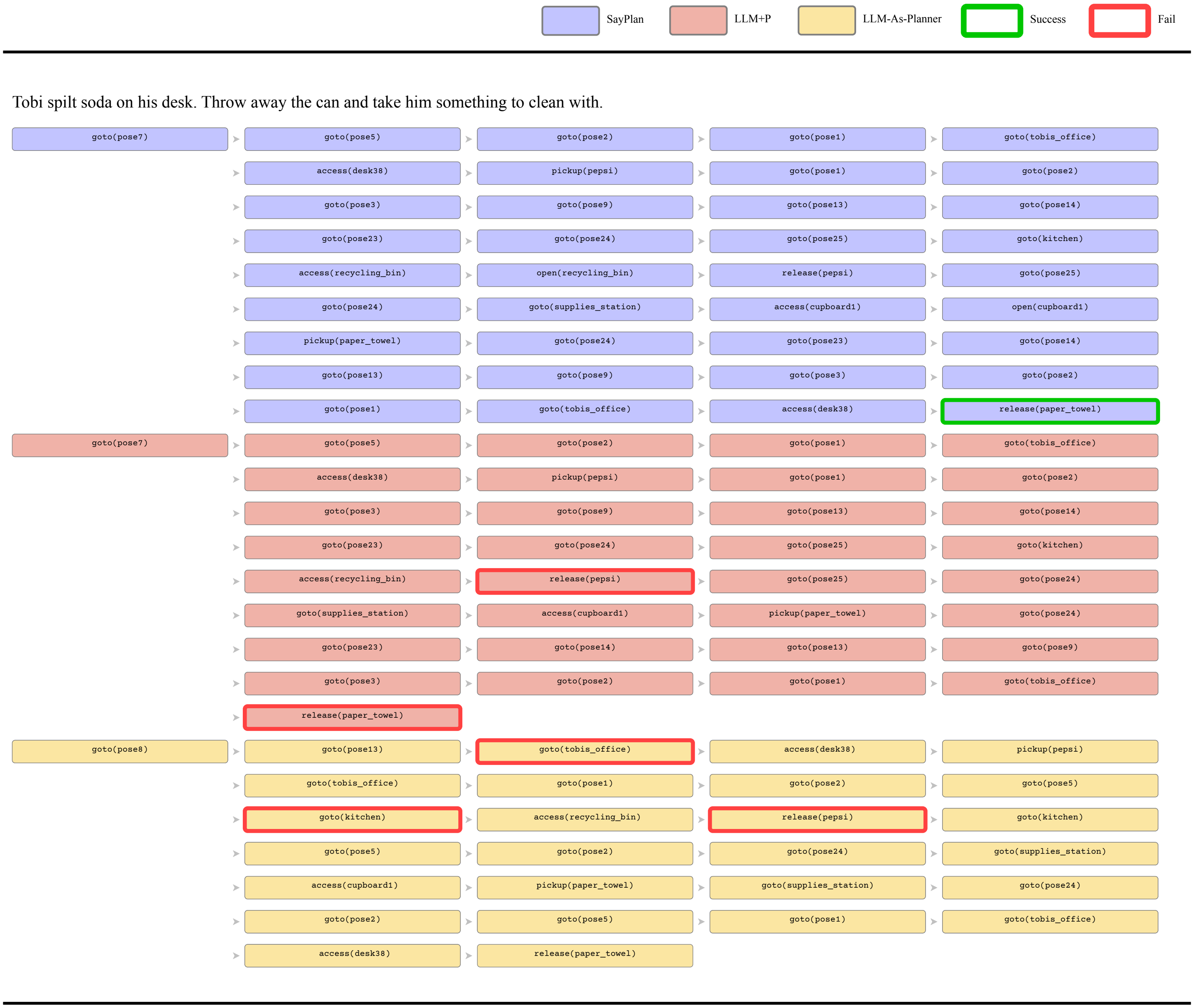


Make Niko a coffee and place the mug on his desk.

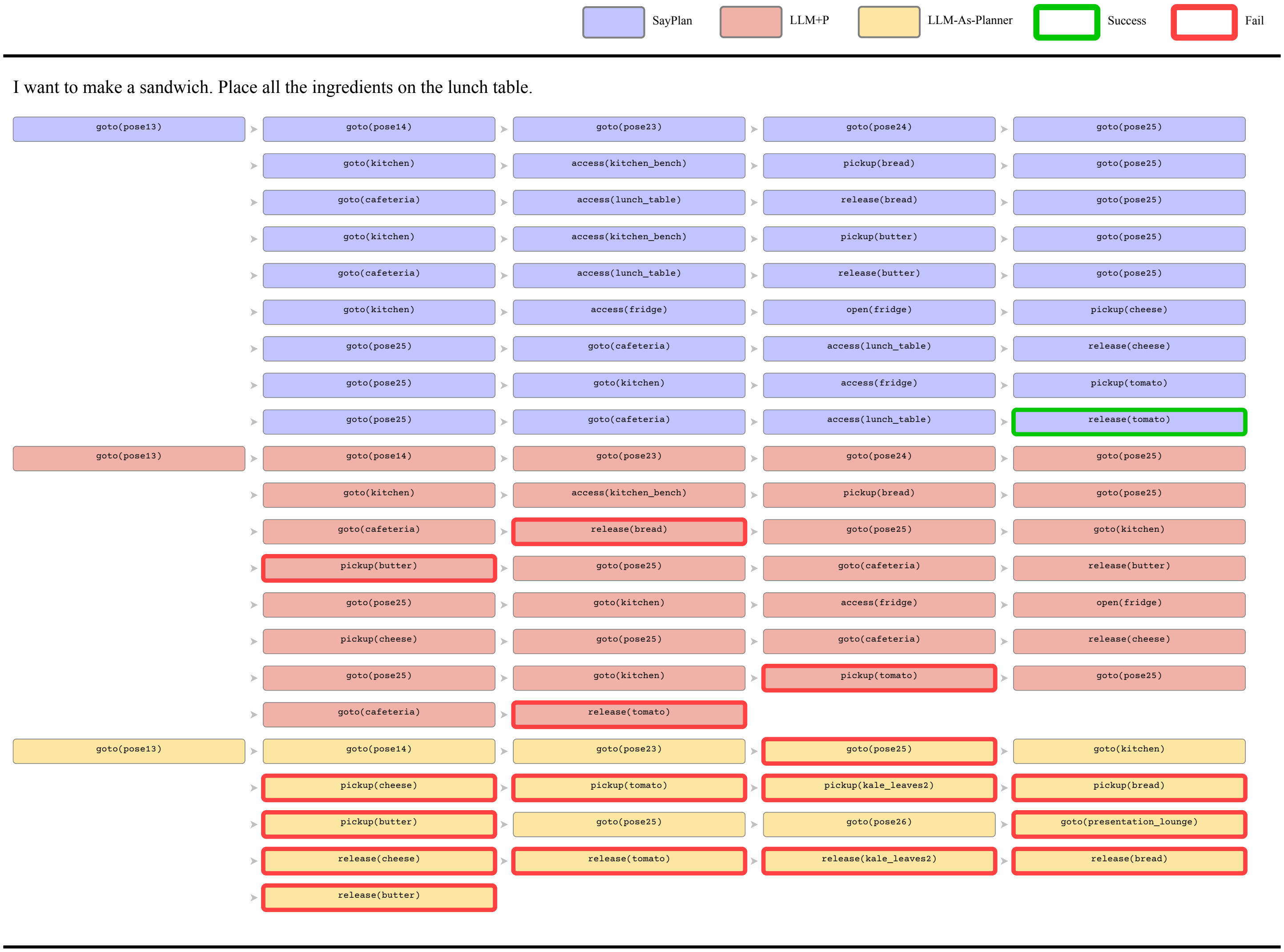
```
goto(pose7) > goto(pose5) > goto(nikos_office) > access(desk1) > pickup(coffee_mug)
> goto(pose5) > goto(pose7) > goto(pose16) > goto(pose18)
> goto(pose15) > goto(pose23) > goto(pose24) > goto(pose25)
> goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine)
> turn_off(coffee_machine) > pickup(coffee_mug) > goto(pose25) > goto(pose24)
> goto(pose23) > goto(pose15) > goto(pose18) > goto(pose16)
> goto(pose7) > goto(pose5) > goto(nikos_office) > access(desk1)
> release(coffee_mug)
goto(pose7) > goto(pose5) > goto(nikos_office) > access(desk1) > pickup(coffee_mug)
> goto(pose5) > goto(pose7) > goto(pose8) > goto(pose13)
> goto(pose14) > goto(pose23) > goto(pose24) > goto(pose25)
> goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine)
> turn_off(coffee_machine) > pickup(coffee_mug) > goto(pose25) > goto(pose24)
> goto(pose23) > goto(pose14) > goto(pose13) > goto(pose8)
> goto(pose7) > goto(pose5) > goto(nikos_office) > access(desk1)
> release(coffee_mug)
goto(pose8) > goto(pose13) > goto(pose14) > goto(pose23) > goto(pose24)
> goto(pose25) > goto(kitchen) > pickup(coffee_mug) > goto(pose25)
> goto(pose24) > goto(pose23) > goto(pose14) > goto(pose13)
> goto(pose5) > goto(nikos_office) > goto(desk1) > release(coffee_mug)
```

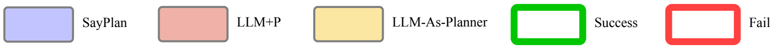
Someone has thrown items in the wrong bins. Correct this.

```
goto(pose13) > goto(pose14) > goto(pose23) > goto(pose24) > goto(pose25)
> goto(kitchen) > access(recycling_bin) > open(recycling_bin) > pickup(milk_carton)
> access(rubbish_bin) > open(rubbish_bin) > release(milk_carton) > access(recycling_bin)
> pickup(orange_peel) > release(orange_peel) > access(recycling_bin) > pickup(apple_core)
> release(apple_core) > access(rubbish_bin) > pickup(banana_peel) > access(recycling_bin)
> release(banana_peel) > access(rubbish_bin) > pickup(plastic_bottle) > release(plastic_bottle)
goto(pose13) > goto(pose14) > goto(pose23) > goto(pose24) > goto(pose25)
> goto(kitchen) > access(recycling_bin) > open(recycling_bin) > pickup(milk_carton)
> access(rubbish_bin) > open(rubbish_bin) > release(milk_carton) > access(recycling_bin)
> pickup(orange_peel) > release(orange_peel) > access(recycling_bin) > pickup(apple_core)
> release(apple_core) > access(rubbish_bin) > pickup(banana_peel) > access(recycling_bin)
> release(banana_peel) > access(rubbish_bin) > pickup(plastic_bottle) > release(plastic_bottle)
goto(pose13) > goto(pose14) > goto(pose23) > goto(pose25) > goto(kitchen)
> goto(recycling_bin) > pickup(milk_carton) > goto(rubbish_bin) > release(milk_carton)
> goto(recycling_bin) > pickup(orange_peel) > goto(rubbish_bin) > release(orange_peel)
> goto(recycling_bin) > pickup(apple_core) > goto(rubbish_bin) > release(apple_core)
> goto(rubbish_bin) > pickup(banana_peel) > goto(recycling_bin) > release(banana_peel)
> goto(rubbish_bin) > pickup(plastic_bottle) > goto(recycling_bin) > release(plastic_bottle)
```

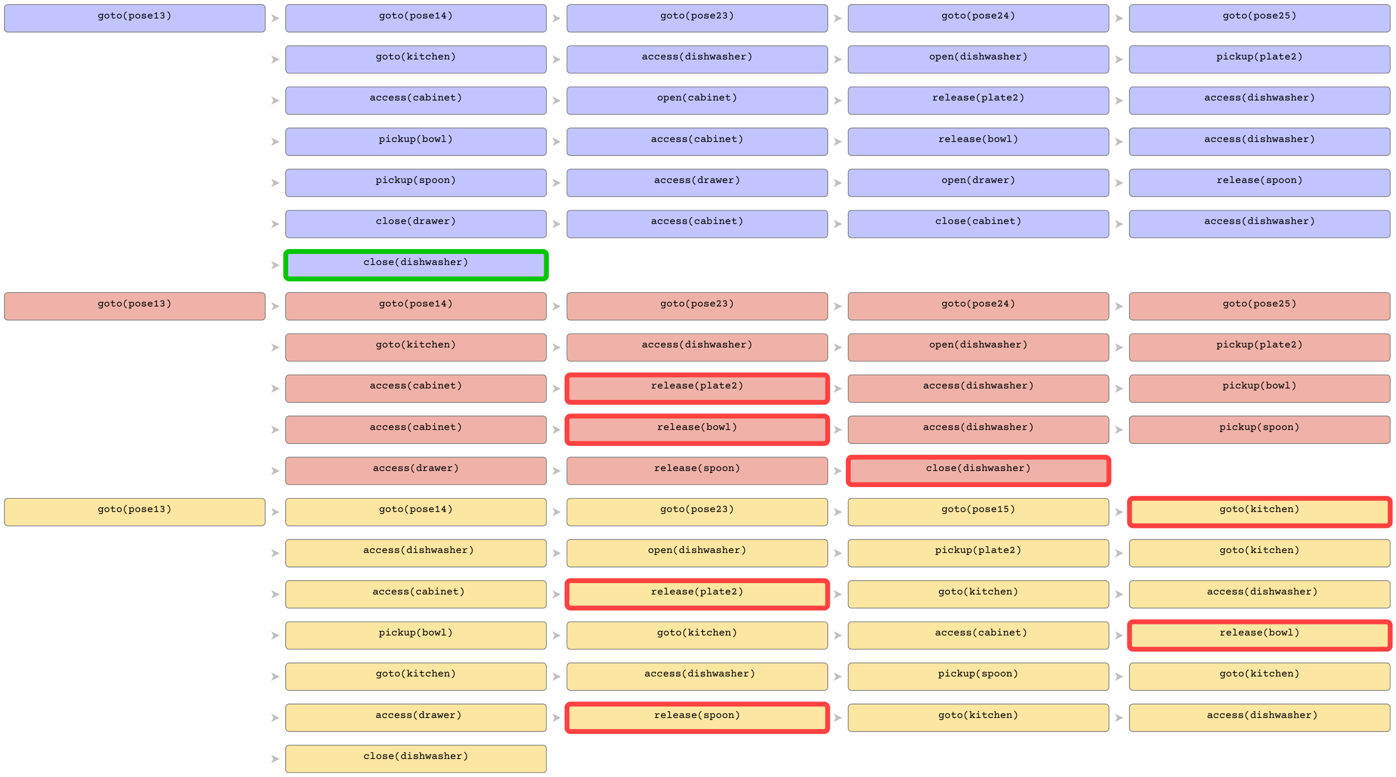


Tobi spilt soda on his desk. Throw away the can and take him something to clean with.


SayPlan
LLM+P
LLM-As-Planner
Success
Fail
I want to make a sandwich. Place all the ingredients on the lunch table.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(bread)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(bread)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(butter)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(butter)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(cheese)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(cheese)
goto(pose25)
goto(kitchen)
access(fridge)
pickup(tomato)
goto(pose25)
goto(cafeteria)
access(lunch_table)
release(tomato)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(kitchen)
access(kitchen_bench)
pickup(bread)
goto(pose25)
goto(cafeteria)
release(bread)
goto(pose25)
goto(kitchen)
pickup(butter)
goto(pose25)
goto(cafeteria)
release(butter)
goto(pose25)
goto(kitchen)
access(fridge)
open(fridge)
pickup(cheese)
goto(pose25)
goto(cafeteria)
release(cheese)
goto(pose25)
goto(kitchen)
pickup(tomato)
goto(pose25)
goto(cafeteria)
release(tomato)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose25)
goto(kitchen)
pickup(cheese)
pickup(tomato)
pickup(kale_leaves2)
pickup(bread)
pickup(butter)
goto(pose25)
goto(pose26)
goto(presentation_lounge)
release(cheese)
release(tomato)
release(kale_leaves2)
release(bread)
release(butter)



Empty the dishwasher. Place all items in their correct locations

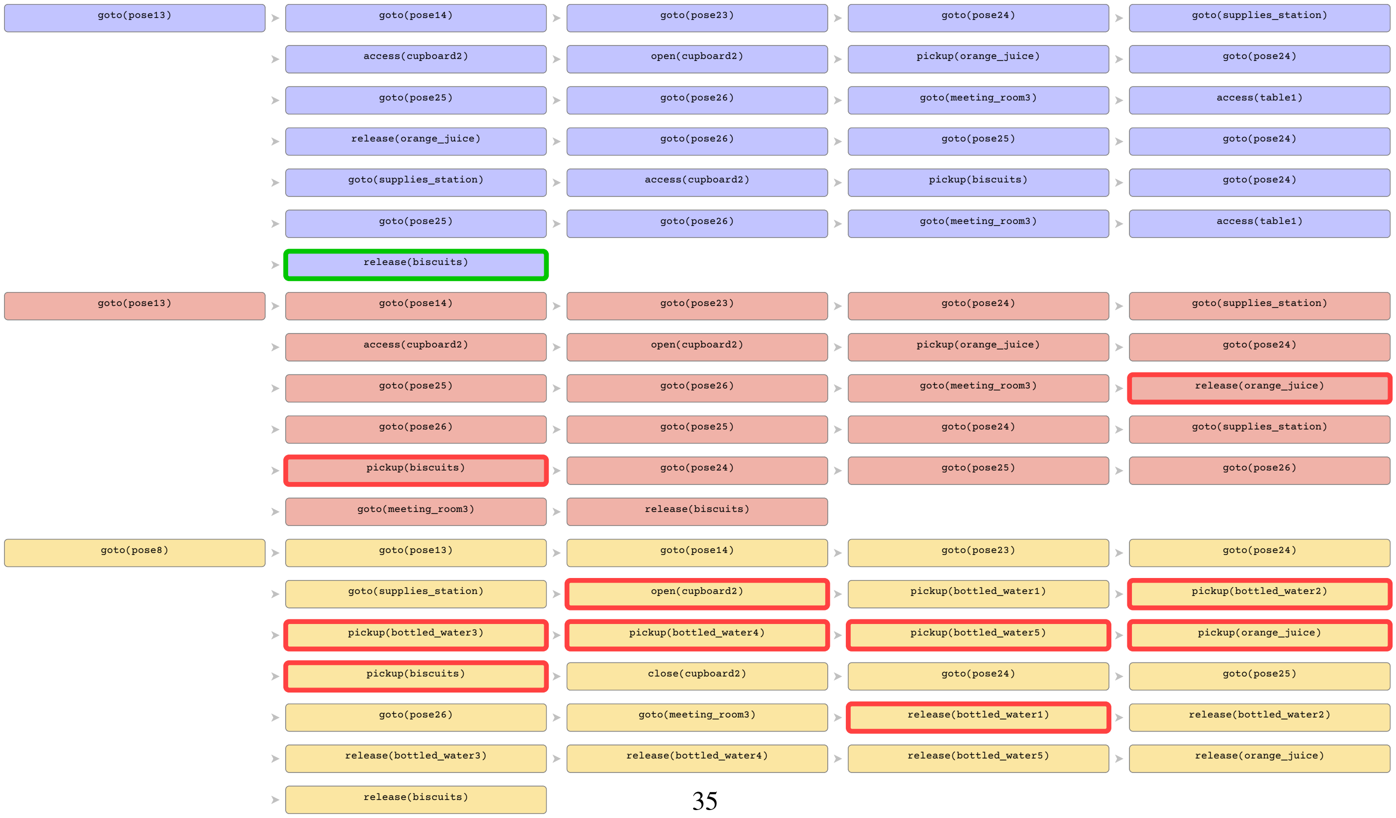


goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose24) ➤ goto(pose25)
➤ goto(kitchen) ➤ access(dishwasher) ➤ open(dishwasher) ➤ pickup(plate2)
➤ access(cabinet) ➤ open(cabinet) ➤ release(plate2) ➤ access(dishwasher)
➤ pickup(bowl) ➤ access(cabinet) ➤ release(bowl) ➤ access(dishwasher)
➤ pickup(spoon) ➤ access(drawer) ➤ open(drawer) ➤ release(spoon)
➤ close(drawer) ➤ access(cabinet) ➤ close(cabinet) ➤ access(dishwasher)
➤ close(dishwasher)

goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose24) ➤ goto(pose25)
➤ goto(kitchen) ➤ access(dishwasher) ➤ open(dishwasher) ➤ pickup(plate2)
➤ access(cabinet) ➤ release(plate2) ➤ access(dishwasher) ➤ pickup(bowl)
➤ access(cabinet) ➤ release(bowl) ➤ access(dishwasher) ➤ pickup(spoon)
➤ access(drawer) ➤ release(spoon) ➤ close(dishwasher)

goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose15) ➤ goto(kitchen)
➤ access(dishwasher) ➤ open(dishwasher) ➤ pickup(plate2) ➤ goto(kitchen)
➤ access(cabinet) ➤ release(plate2) ➤ goto(kitchen) ➤ access(dishwasher)
➤ pickup(bowl) ➤ goto(kitchen) ➤ access(cabinet) ➤ release(bowl)
➤ goto(kitchen) ➤ access(dishwasher) ➤ pickup(spoon) ➤ goto(kitchen)
➤ access(drawer) ➤ release(spoon) ➤ goto(kitchen) ➤ access(dishwasher)
➤ close(dishwasher)

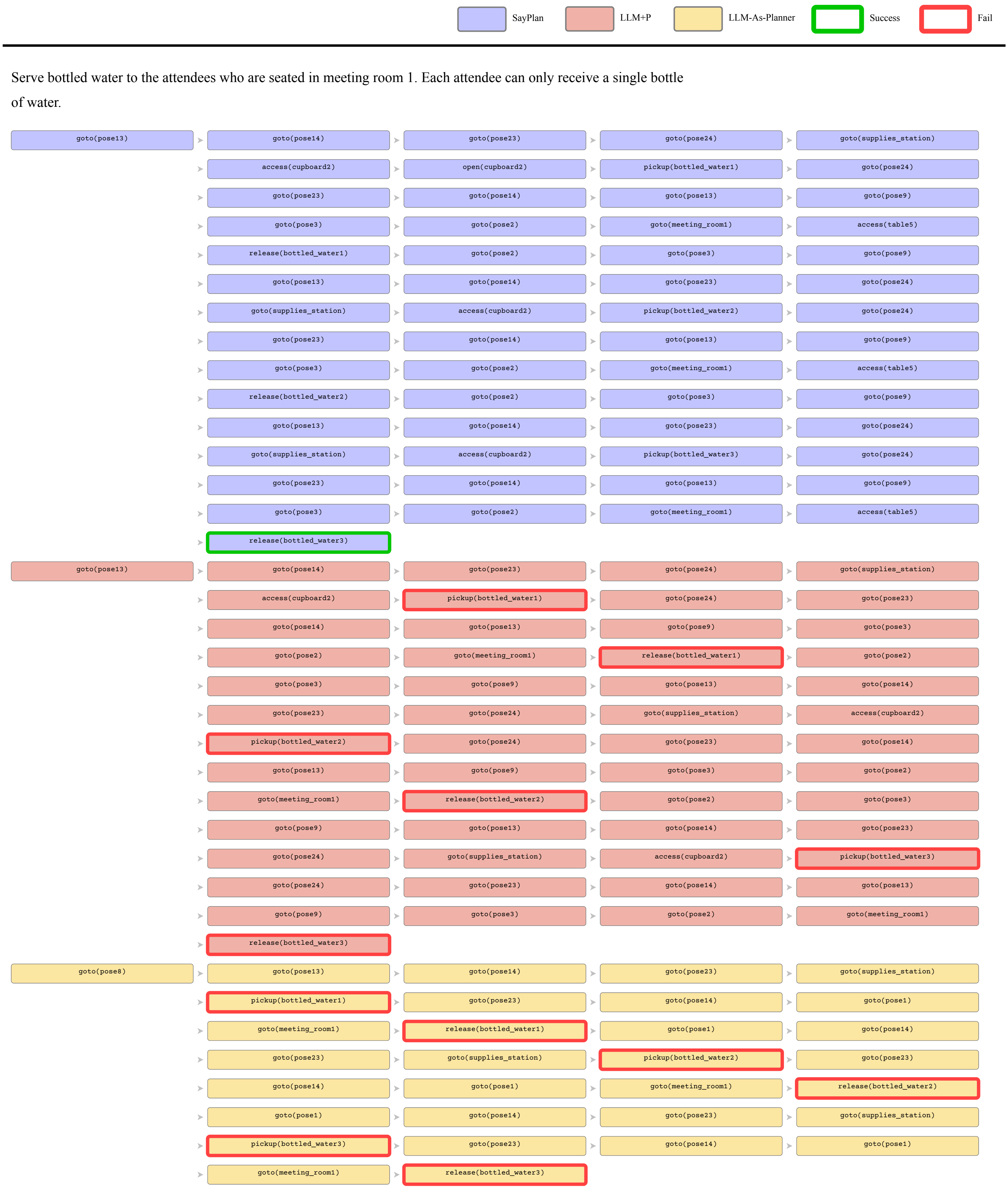
A delegation of project partners is arriving soon. We want to serve them snacks and non-alcoholic drinks. Prepare everything in the largest meeting room. Use items found in the supplies room only.



goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose24) ➤ goto(supplies_station)
➤ access(cupboard2) ➤ open(cupboard2) ➤ pickup(orange_juice) ➤ goto(pose24)
➤ goto(pose25) ➤ goto(pose26) ➤ goto(meeting_room3) ➤ access(table1)
➤ release(orange_juice) ➤ goto(pose26) ➤ goto(pose25) ➤ goto(pose24)
➤ goto(supplies_station) ➤ access(cupboard2) ➤ pickup(biscuits) ➤ goto(pose24)
➤ goto(pose25) ➤ goto(pose26) ➤ goto(meeting_room3) ➤ access(table1)
➤ release(biscuits)

goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose24) ➤ goto(supplies_station)
➤ access(cupboard2) ➤ open(cupboard2) ➤ pickup(orange_juice) ➤ goto(pose24)
➤ goto(pose25) ➤ goto(pose26) ➤ goto(meeting_room3) ➤ release(orange_juice)
➤ goto(pose26) ➤ goto(pose25) ➤ goto(pose24) ➤ goto(supplies_station)
➤ pickup(biscuits) ➤ goto(pose24) ➤ goto(pose25) ➤ goto(pose26)
➤ goto(meeting_room3) ➤ release(biscuits)

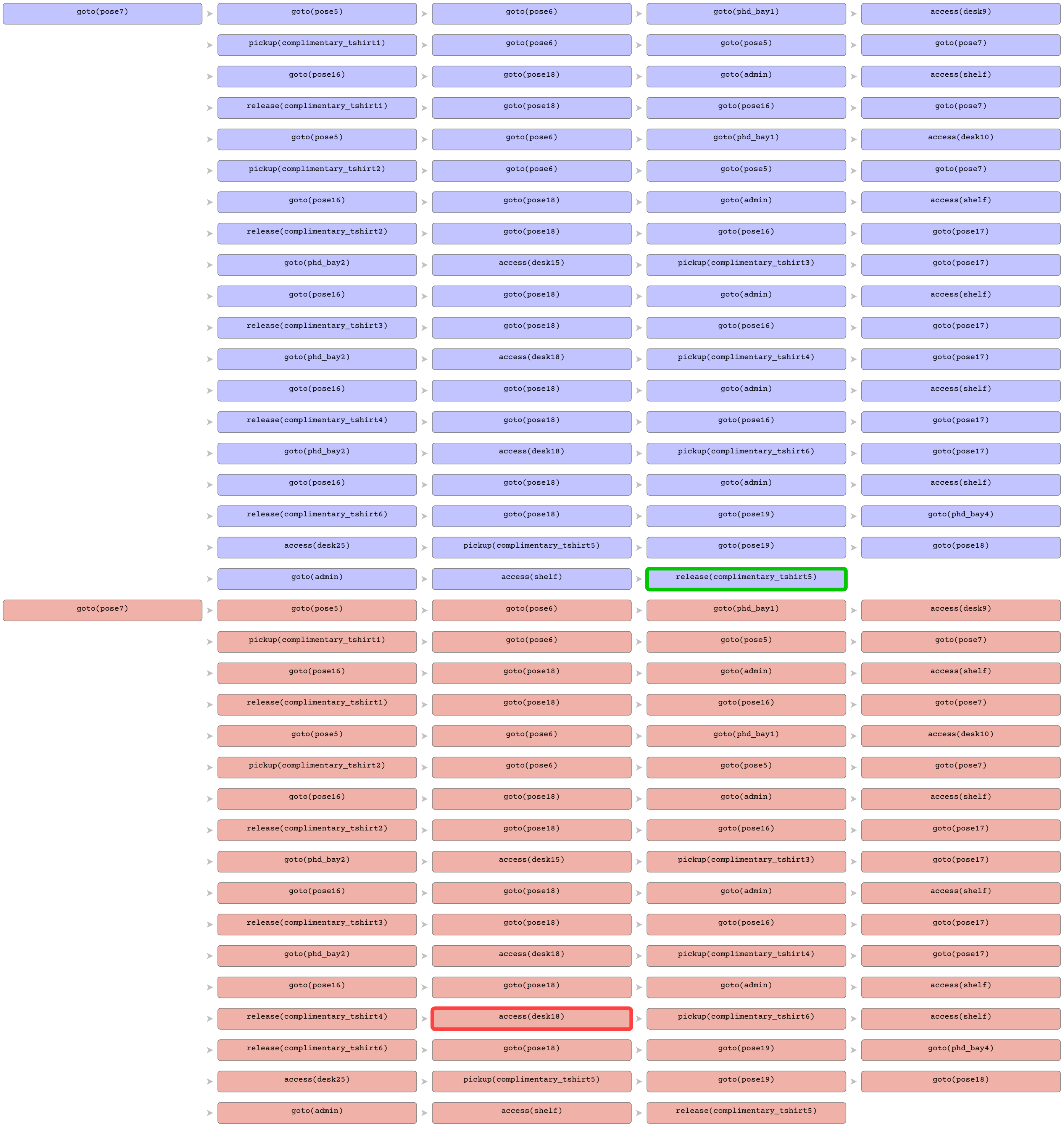
goto(pose8) ➤ goto(pose13) ➤ goto(pose14) ➤ goto(pose23) ➤ goto(pose24)
➤ goto(supplies_station) ➤ open(cupboard2) ➤ pickup(bottled_water1) ➤ pickup(bottled_water2)
➤ pickup(bottled_water3) ➤ pickup(bottled_water4) ➤ pickup(bottled_water5) ➤ pickup(orange_juice)
➤ pickup(biscuits) ➤ close(cupboard2) ➤ goto(pose24) ➤ goto(pose25)
➤ goto(pose26) ➤ goto(meeting_room3) ➤ release(bottled_water1) ➤ release(bottled_water2)
➤ release(bottled_water3) ➤ release(bottled_water4) ➤ release(bottled_water5) ➤ release(orange_juice)
➤ release(biscuits)

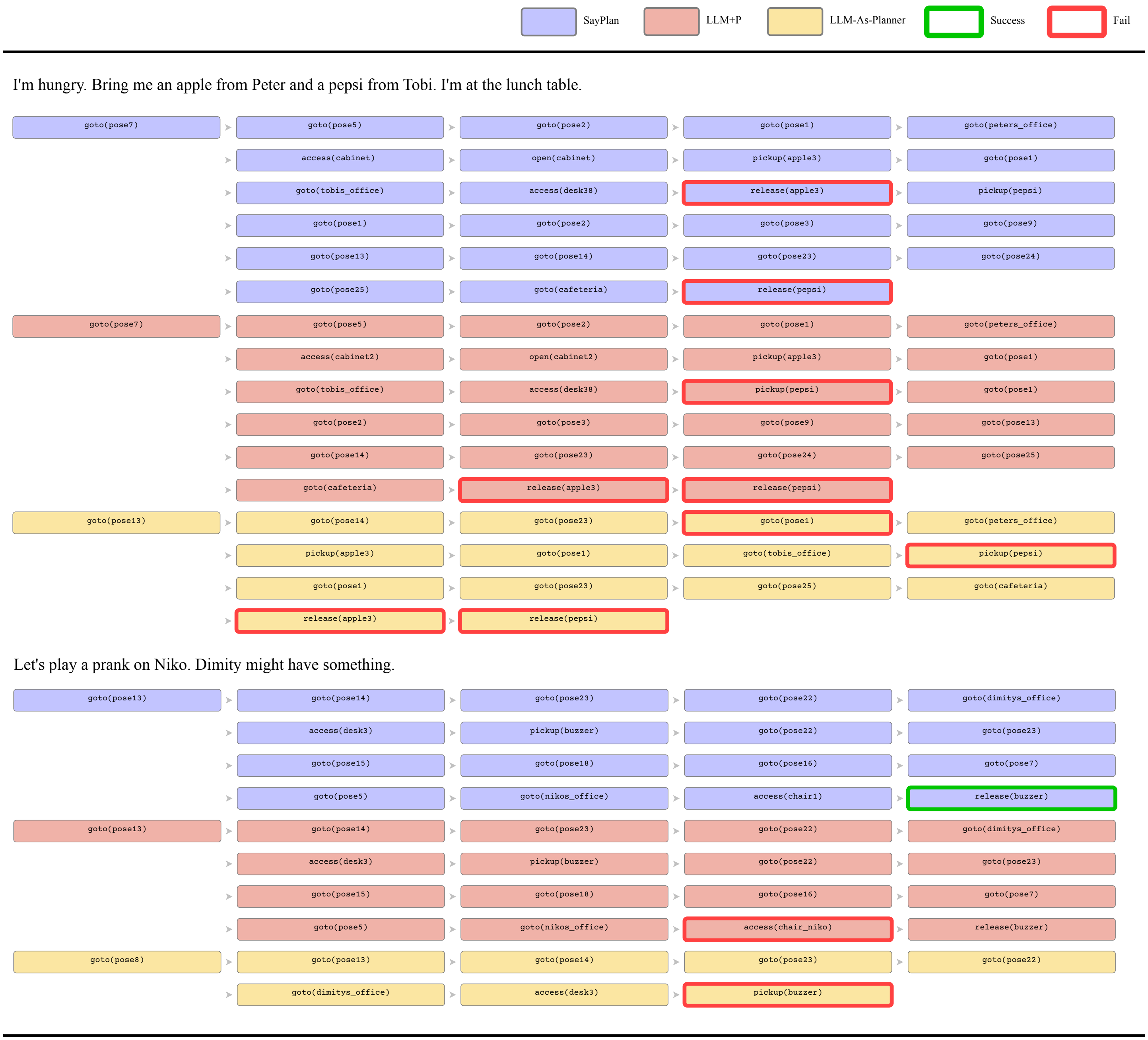


Serve bottled water to the attendees who are seated in meeting room 1. Each attendee can only receive a single bottle of water.

🟦 SayPlan 🟥 LLM+P 🟨 LLM-As-Planner 🟩 Success 🟥 Fail

Locate all 6 complimentary t-shirts given to the PhD students and place them on the shelf in admin.




SayPlan
LLM+P
LLM-As-Planner
Success
Fail
I'm hungry. Bring me an apple from Peter and a pepsi from Tobi. I'm at the lunch table.
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
access(cabinet)
open(cabinet)
pickup(apple3)
goto(pose1)
goto(tobis_office)
access(desk38)
release(apple3)
pickup(pepsi)
goto(pose1)
goto(pose2)
goto(pose3)
goto(pose9)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(cafeteria)
release(pepsi)
goto(pose7)
goto(pose5)
goto(pose2)
goto(pose1)
goto(peters_office)
access(cabinet2)
open(cabinet2)
pickup(apple3)
goto(pose1)
goto(tobis_office)
access(desk38)
pickup(pepsi)
goto(pose1)
goto(pose2)
goto(pose3)
goto(pose9)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose24)
goto(pose25)
goto(cafeteria)
release(apple3)
release(pepsi)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose1)
goto(peters_office)
pickup(apple3)
goto(pose1)
goto(tobis_office)
pickup(pepsi)
goto(pose1)
goto(pose23)
goto(pose25)
goto(cafeteria)
release(apple3)
release(pepsi)
Let's play a prank on Niko. Dimity might have something.
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(desk3)
pickup(buzzer)
goto(pose22)
goto(pose23)
goto(pose15)
goto(pose18)
goto(pose16)
goto(pose7)
goto(pose5)
goto(nikos_office)
access(chair1)
release(buzzer)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(desk3)
pickup(buzzer)
goto(pose22)
goto(pose23)
goto(pose15)
goto(pose18)
goto(pose16)
goto(pose7)
goto(pose5)
goto(nikos_office)
access(chair_niko)
release(buzzer)
goto(pose8)
goto(pose13)
goto(pose14)
goto(pose23)
goto(pose22)
goto(dimitys_office)
access(desk3)
pickup(buzzer)

G Real World Execution of a Generated Long Horizon Plan.

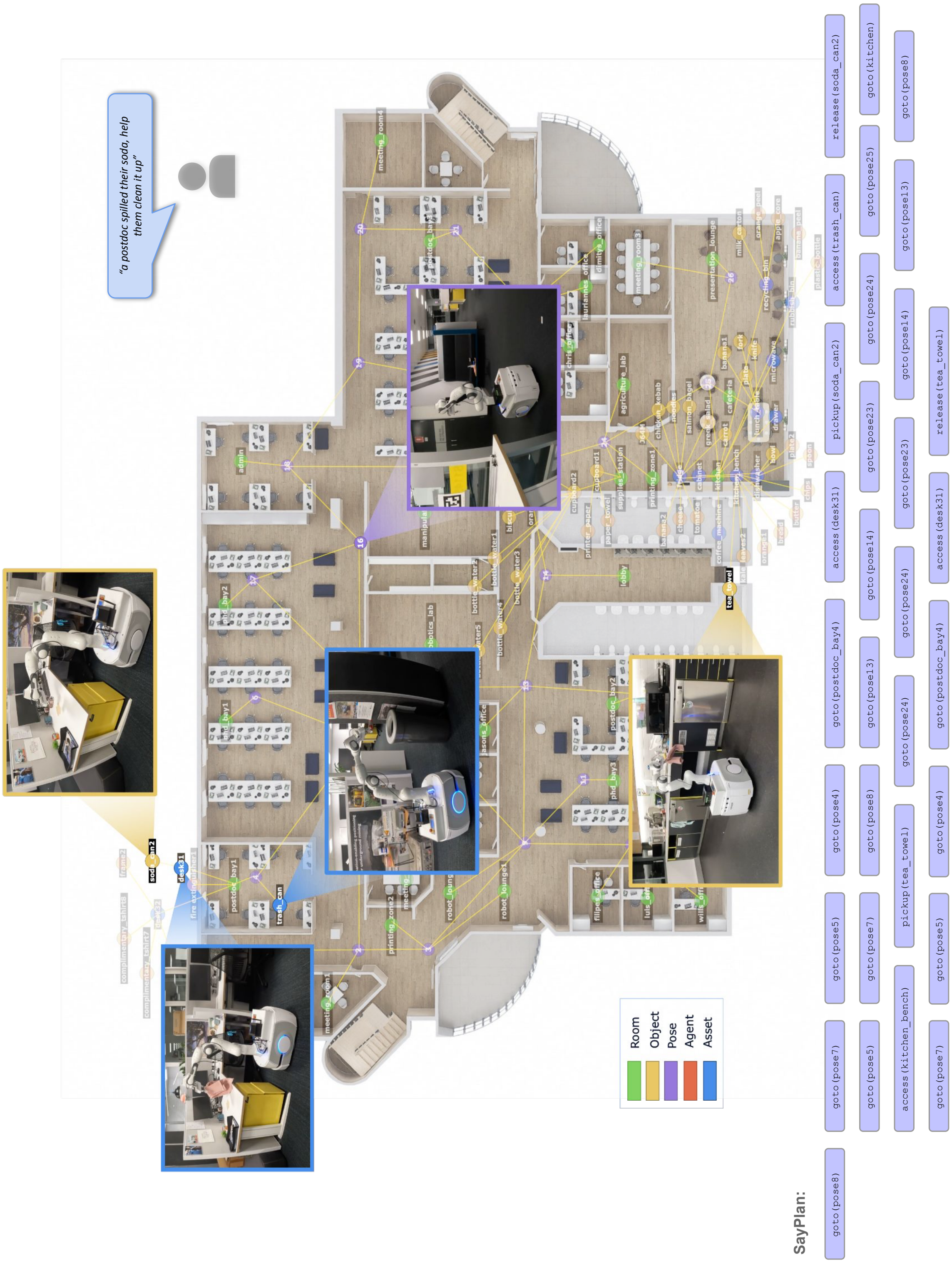


Figure 6: **Real World Execution of a Generated Long Horizon Plan.** Execution of a generated plan on a real world mobile manipulator robot.

H Input Prompt Structure

Input prompt passed to the LLM for SayPlan. Note that the components highlighted in violet represent static components of the prompt that remain fixed throughout both the semantic search and iterative re-planning phases of SayPlan.

Agent Role: You are an excellent graph planning agent. Given a graph representation of an environment, you can explore the graph by expanding nodes to find the items of interest. You can then use this graph to generate a step-by-step task plan that the agent can follow to solve a given instruction.

Environment Functions:

goto(<pose>): Move the agent to any room node or pose node.

access(<asset>): Provide access to the set of affordances associated with an asset node and its connected objects.

pickup(<object>): Pick up an accessible object from the accessed node.

release(<object>): Release grasped object at an asset node.

turn_on/off(<object>): Toggle object at agent's node, if accessible and has affordance.

open/close(<asset>): Open/close asset at agent's node, affecting object accessibility.

done(): Call when the task is completed.

Environment State:

ontop_of(<asset>): Object is located on <asset>

inside_of(<asset>): Object is located inside <asset>

inside_hand: Object is currently being grasped by the robot/agent

closed: Asset can be opened

open: Asset can be closed or kept open

on: Asset is currently on

off: Asset is currently off

accessible: The object is not accessible if it is inside an asset and the asset state is "closed".

Environment API:

expand_node(<node>): Reveal assets/objects connected to a room/floor node.

contract_node(<node>): Hide assets/objects, reducing graph size for memory constraints.

verify_plan(): Verify generated plan in the scene graph environment.

Output Response Format:

{*chain_of_thought:* break your problem down into a series of intermediate reasoning steps to help you determine your next command,

reasoning: justify why the next action is important

mode: "exploring" OR "planning"

command: {"command_name": Environment API call
          "node_id": node to perform an operation on
          "plan": task plan if in planning mode}}

Example: <see Appendix I and J>

Instruction: Natural language description of the task

3D Scene Graph: Text-serialised JSON description of a 3D scene graph

Memory: History of previously expanded nodes

Feedback: External textual feedback from scene graph simulator

Ensure the response can be parsed by Python json.loads.

I Example Input-Output Interaction Between LLM and Scene Graph Simulator During Semantic Search

Instruction: make a coffee for Tom and place it in his room

3D Scene Graph: {nodes: {room: [{id: bobs_room}, {id: toms_room}, {id: jacks_room}, {id: kitchen}, {id: livingroom}], pose: [{id: pose1}, {id: pose2}, {id: pose3}, {id: pose4}, {id: pose5}], agent: [{location: bobs_room, id: agent}]}, links: [bobs_room↔pose1, bobs_room↔agent, toms_room↔pose1, toms_room↔pose2, toms_room↔pose5, jacks_room↔pose2, jacks_room↔pose3, kitchen↔pose3, kitchen↔pose4, kitchen↔pose5, livingroom↔pose4]}

Memory: []

SayPlan [Mode->Search]: chain-of-thought: explore graph for coffee mug, coffee machine and somewhere to place the coffee -> perform common sense exploration first -> explore toms room -> explore kitchen -> if items not found then explore other nodes
reasoning: i will expand toms room
command: expand(toms_room)

Scene Graph Simulator: <Updated 3D scene graph>

Memory: [toms_room]

SayPlan [Mode->Search]: chain-of-thought: i have found a wardrobe in tom's room -> leave this node expanded -> the coffee mug is not in his room -> still have not found the coffee machine -> kitchen might have coffee machine and coffee mug -> explore this node next
reasoning: i will expand the kitchen next
command: expand(kitchen)

Scene Graph Simulator: <Updated 3D scene graph>

Memory: [toms_room, kitchen]

SayPlan [Mode->Search]: chain-of-thought: i've found the coffee machine in the kitchen -> the coffee mug however is not in the kitchen -> i will explore other unexpanded nodes now for the coffee mug -> jacks room, bobs room and the living room are unexplored yet -> i will try jack"s room next
reasoning: i will expand jacks room next
command: expand(jacks_room)

Scene Graph Simulator: <Updated 3D scene graph>

Memory: [toms_room, kitchen, jacks_room]

SayPlan [Mode->Search]: chain-of-thought: jacks room does not have anything relevant towards solving this task -> I should contract this node to save token memory -> i will explore other unexplored nodes after doing this
reasoning: contract jack's room to save memory
command: contract(jacks_room)

Scene Graph Simulator: <Updated 3D scene graph>

Memory: [toms_room, kitchen, jacks_room]

SayPlan [Mode->Search]: chain-of-thought: i have found a wardrobe in tom's room and coffee machine in the kitchen -> the coffee mug is still required for the task -> explore a new room next
reasoning: i will expand bob's room as it is unexplored
command: expand(bobs_room)

Scene Graph Simulator: <Updated 3D scene graph>

Memory: [toms_room, kitchen, jacks_room, bobs_room]

SayPlan [Mode->Search]: chain-of-thought: i have found the coffee mug, the coffee machine and tom's wardrobe on the graph -> search complete -> I will switch to planning mode-> collect coffee mug -> generate plan for making coffee -> place coffee mug on Tom's wardrobe
reasoning: I will generate a task plan using the identified subgraph

J Example Input-Output Interaction Between LLM and Scene Graph Environment During Iterative Re-planning

Instruction: make a coffee for Tom and place it in his room

3D Scene Graph: {nodes: {room: [{id: bobs_room}, {id: toms_room}, {id: jacks_room}, {id: kitchen}, {id: livingroom}], pose: [{id: pose1}, {id: pose2}, {id: pose3}, {id: pose4}, {id: pose5}], agent: [{location: bobs_room, id: agent}], asset: [{room: toms_room, state: free, affordances: [release], id: bed2}, {room: toms_room, state: closed, affordances: [open, close, release], id: wardrobe2}, {room: kitchen, state: closed, affordances: [open, close, release], id: fridge}, {affordances: [turn_on, turn_off], state: off, id: coffee_machine}, {room: bobs_room, state: free, affordances: [release], id: bed1}, {room: bobs_room, state: closed, affordances: [open, close, release], id: wardrobe1}], object: [{affordances: [pickup], state: inside_of(wardrobe1), attributes: "blue", id: coffee_mug}]}, links: [bobs_room↔pose1, bobs_room↔agent, bobs_room↔bed1, bobs_room↔wardrobe1, toms_room↔pose1, toms_room↔pose2, toms_room↔pose5, toms_room↔bed2, toms_room↔wardrobe2, jacks_room↔pose2, jacks_room↔pose3, kitchen↔pose3, kitchen↔pose4, kitchen↔pose5, kitchen↔fridge, kitchen↔coffee_machine, livingroom↔pose4, wardrobe1↔coffee_mug]}

Memory: [toms_room, kitchen, jacks_room, bobs_room]

SayPlan [Mode->Planning]: chain-of-thought: "i have found the coffee mug, the coffee machine and tom's wardrobe on the graph -> search complete -> I will switch to planning mode-> collect coffee mug -> generate plan for making coffee -> place coffee mug on Tom's wardrobe"
reasoning: I will generate a task plan using the identified subgraph
plan: [goto(bobs_room) > access(wardrobe1) > pickup(coffee_mug) > goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine) > turn_off(coffee_machine) > pickup(coffee_mug) > goto(toms_room) > access(wardrobe2) > release(coffee_mug) > done]

Path Planner:plan: [goto(pose3) > goto(pose4) > goto(pose1) > goto(bobs_room) > access(wardrobe1) > pickup(coffee_mug) > goto(pose2) > goto(pose5) > goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine) > turn_off(coffee_machine) > pickup(coffee_mug) > goto(pose5) > goto(pose2) > goto(toms_room) > access(wardrobe2) > release(coffee_mug) > done]

Scene Graph Simulator: "*coffee mug is not accessible*"

Memory: [toms_room, kitchen, jacks_room, bobs_room]

SayPlan [Mode->Planning]: chain-of-thought: "the coffee mug was not accessible -> the wardrobe it is in is currently closed -> I update the plan to open the wardrobe first -> generate plan for making coffee -> place coffee mug on Tom's wardrobe"
reasoning: I will generate a task plan using the identified subgraph and the feedback from the simulator
plan: [goto(bobs_room) > access(wardrobe1) > open(wardrobe1) > pickup(coffee_mug) > goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine) > turn_off(coffee_machine) > pickup(coffee_mug) > goto(toms_room) > access(wardrobe2) > release(coffee_mug) > done]

Path Planner:plan: [goto(pose3) > goto(pose4) > goto(pose1) > goto(bobs_room) > access(wardrobe1) > open(wardrobe1) > pickup(coffee_mug) > goto(pose2) > goto(pose5) > goto(kitchen) > access(coffee_machine) > release(coffee_mug) > turn_on(coffee_machine) > turn_off(coffee_machine) > pickup(coffee_mug) > goto(pose5) > goto(pose2) > goto(toms_room) > access(wardrobe2) > release(coffee_mug) > done]

Scene Graph Simulator: Plan Verified