RGB image from Simulator

Semantic Map

Semantic mapping → *Extract objects information* →

Perform actions to the environment

Action policy

Skill **Explore** —>
(Produced by action planner) LookDown_15, MoveAhead, ..., RotateRight

Skill **NavigateToObject**—>
(Produced by action planner) MoveAhead, ..., RotateLeft, MoveAhead, MoveAhead

...

Feed output to action policy to obtain executable actions

Executor output
— — — — — — —

Thought: I will explore the room to have a general idea of the environment.
Action: Play[Explore,None]

Executor input
— — — — — — —

Observation: You are in a kitchen. There are no objects currently in your view. You are not holding any objects.
Found objects: None.
Objects seeing in current observation: None.
Previous steps: steps=[]
Current objective: Explore the room to have a general idea of the environment.

Skill library
— — — — — — —

Explore, LookAround, NavigateTo, RequireReplan,
...

The OPEx framework
Planner: decompose a task into sub-tasks.
Observer: collect information from environment and prior experience.
Executor: produce actions based on Planner's plan and Observer's observation guidance.

Observer input
— — — — — — —

Setup: You are a helpful assistant helping .... Please output the description of the environment with the header 'Observation:'.

Room type: Kitchen.
Task description: Explore the room to have a general idea of the environment.
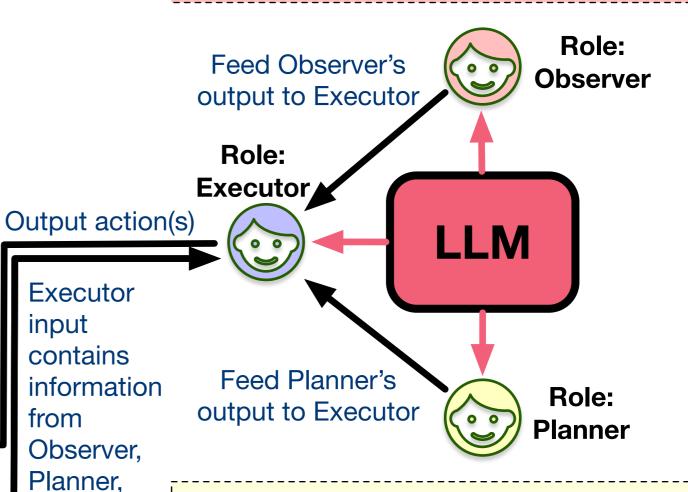**Previously found objects:** None. (From semantic map)
Objects seen in current observation: None. (From RGB image)
Holding object: None.
Error message: None.

Observer output
— — — — — — —

Observation: You are in a kitchen. There are no objects currently in your view. You are not holding any objects.

Role: Observer

Feed Observer's output to Executor

Role: Executor

LLM

Output action(s)

Executor input contains information from Observer, Planner, and skill library.

Feed Planner's output to Executor

Role: Planner

Planner input
— — — — — — —

Setup: You are an intelligent planner for a household robot agent .... Please output the task type starting with the header 'Task type:'. .... Please output the plan starting with the header 'Plan' and .... At the end of your plan, say '<END_OF_PLAN>'.

Task: put a washed bowl away in a kitchen cabinet.

Planner output
— — — — — — —

Task type:
PICK_CLEAN_THEN_PLACE_IN_RECEP
Thought: To complete the task, the robot needs to....
Plan:
0. Explore the room to have a general idea of the environment.
1. Find a bowl in the room.
...
6. Put the bowl in the cabinet.
<END_OF_PLAN>