Subgoal Generator

You are a helper to define subgoals for an Agent in order to complete long horizon tasks. Based on the given task description and environmental information, you are required to output a sequence of goals that the Agent needs to achieve at the moment.
Remember:
1. The subgoals should depict the states of objects in the environment instead of actions to do.
2. Output the subgoals in a list, e.g. [subgoal1, subgoal2, ...].
3. Try to generate minimum but key subgoals and you do not need to make them detailed steps.
4. You should output the subgoal as simple as possible without any explanation.

Food Preparation

In this setting, there is an Agent in a virtual home environment. The task is to prepare food, specifically, the Agent needs to get the pancake and put it in the microwave which are both located in the kitchen.
Initially, Agent is in the kitchen and not grabbing anything in hands, Agent notices the pancake but it is not nearby, the microwave is not opened.

Entertainment

In this setting, there is an Agent in a virtual home environment. The task is to relax in the living room. Specifically, the Agent needs to fetch the chips and the milk in the kitchen, then turn on the TV and sit on the sofa in the living room.
Initially, Agent is in the kitchen and not grabbing anything in hands, Agent notices the milk and chips but they are not nearby.

ROMAN Robot Manipulation

In this setting, there is a robot arm in a laboratory, in the scene there is a cabinet, a drawer, a box, a cover, a rack, a vial, a table, a button, and a conveyor belt. The aim of this task to use the conveyor belt to transport the rack with the vial to another place for production.
The initial state of the environment is as follows:
The vial is stored inside the box with a cover on top, which is located in the closed drawer. The rack is inside the closed cabinet under the table, where the rack can be stably placed. The conveyor is next to the table with a nearby button to start running.
Remember: During the process the vial can't be placed anywhere other than inside the rack, it should be taken out from the box only after the rack is stably placed.

Policy Generator

You are helper to generate policies for an agent to reach subgoals in completing long horizon tasks. Based on the environment information, given current state of the environment and the subgoal for the moment, you are required to output a sequence of actions to reach this goal.
Remember: You should output the action(s) in a list.
Answer template:
Analysis: <brief the reason to do so>
Action sequence: [act1, act2, ...]

Food Preparation

In this setting, there is an Agent in a virtual home environment. The task is to prepare food, specifically, the Agent needs to get the pancake and put it in the microwave which are both located in the kitchen.
The available actions are as follows:
["walk to the living room", "walk to the kitchen", "walk to the bathroom", "walk to the bedroom", "walk to the pancake", "walk to the microwave", "grab the pancake", "put the pancake in the microwave", 'open the microwave', 'close the microwave']
The information about the environment:
1. There are four rooms in the home, living room, kitchen, bathroom and bedroom.
2. The Agent initially holds nothing in hands.
3. The pancake and the microwave are located in the kitchen.

Entertainment

In this setting, there is an Agent in a virtual home environment. The task is to relax in the living room. Specifically, the Agent needs to get the chips and the milk located, then turn on the TV and sit on the sofa.
The available actions in current state are as follows:
["walk to the living room", "walk to the kitchen", "walk to the bathroom", "walk to the bedroom", "walk to the chips", "walk to the milk", 'walk to the coffee table', 'walk to the TV', 'walk to the sofa', "grab the chips", "grab the milk", 'put the chips on the coffee table', 'put the milk on the coffee table', "turn on the TV", "turn off the TV", "sit on the sofa", "stand up from the sofa"]
The information about the environment:
1. There are four rooms in the home, the living room, kitchen, bathroom and bedroom.
2. The agent initially holds nothing in hands.
3. The chips and the milk are located in the kitchen, the sofa and TV are located in the living room.
4. Agent can hold one object in each hand, consider if the Agent needs to free one before interaction.
5. Agent needs to get near the object before using it.

ROMAN Robot Manipulation

In this setting, there is a robot arm in a laboratory, in the scene there is a cabinet, a drawer, a box, a cover, a rack, a vial, a table, a button, and a conveyor belt.
The information about the environment is as follows: The vial is stored inside the box with a cover on top, which is located in the closed drawer. The rack is inside the closed cabinet under the table, where the rack can be stably placed. The conveyor is next to the table with a nearby button to start running. During the process the vial can't be placed anywhere other than inside the rack, it should be taken out from the box only after the rack is stably placed.
The available actions are as follows:
["Pull the drawer to open", "Take off the box cover", "Rotate the cabinet door to open", "Place the rack on the table", "Insert the vial into the rack", "Push the rack onto the conveyor belt", "Press the button"]

The current state is described as: {caption of current state},
In order to achieve: {subgoal for the current state},
what should you do next?

Tab 1. PPO hyperparameter settings for VirtualHome Tasks

Hyperparameter	Value
Max training step	5e4
Max episode steps	32
Update timestep	128
Number of mini-batches	4
Discount factor	0.99
Actor learning rate	3e-4
Critic learning rate	1e-3
Surrogate clipping coefficient	0.2
Entropy coefficient	0.01
Value loss coefficient	0.5

Tab 2. PPO hyperparameter settings for ROMAN Tasks

Hyperparameter	Value
Max training step	5e4 / 1e7
Max episode steps	32 / 2400
Number of steps for update	64 / 9600
Number of mini-batches	2 / 4
Discount factor	0.99
Actor learning rate	3e-4
Critic learning rate	1e-3
Surrogate clipping coefficient	0.2
Entropy coefficient	0.01
Value loss coefficient	0.5