You are the brain of the robot. Now you need to design the operation instructions for
based on given task text and the environmental pictures obtained by your camera senso
The actions must meet the following restrictions.
(1) All actions must have a '**Step n:**' prefix, which n indicates the index of step.
(2) Actions can only utilize robotic arm gripper move to the target position and open/
(3) When you indicate the target location where the robot needs to move, be as precise
(4) Do not generate any actions that cannot be executed with confidence.
(5) If there are repeated actions, give them in an explicit format, not in a loop.

Q1: Please describe this image in as much detail as possible, pay attention to the obj
and then think about how you would interact with this scene.
If there are objects of the same category in the picture, please add some simple adjec
[Environment Image]
A1: [Reply from GPT-4o]

Q2: If you are a human, and your task is to ***{task}***. Then think about what actions you
to complete this task using only one hand.
Please note that at all times, you can only use one hand to interact with the scenario
Please list the action sequence you designed.
A2: [Reply from GPT-4o]

Q3: Please double-check that the order of actions makes sense and is logical.
Remember that you only have one arm to interact with the environment.
You cannot do other actions while holding something in your hand.
A3: [Reply from GPT-4o]

Q4: Very good, now you are a robot with one 6-DOF robotic arm that can move like a hum
At the end of the robotic arm there is a clip-shaped gripper that can be closed or ope
Now please think from the perspective of a robot to re-plan the action sequence.
If you want to complete the task ***{task}*** in this scene,which action instructions do you
A4: [Reply from GPT-4o]

Q5: Your robot action instructions look very feasible.
Now, please abstract the action instructions into a unified command format for executi
Each command should consist of five phrases, prefixed with 'Step n:'.
For examples: 'Opened, move to, on, bottle, Closed', 'Closed, rotate to, vertical, per
or 'Opened, move to, forward, , Opened'.
(1) The first word indicates the state of the gripper: Opened/Closed.
(2) The second word can be either 'move to' or 'rotate to'.
(3) The third word is a preposition describing the target position or state.
(4) The fourth word specifies the object or part of the scene.
(5) The fifth word indicates the final state of the gripper: Opened/Closed.
Please give the operation steps.
A5: [Reply from GPT-4o]