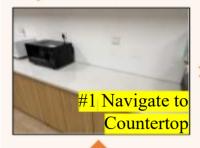
Task Instruction: Can you help me find the coffee and heat it up?















Embodied-Reasoner Success Rate: **56.** 7%

#1 Thought: Let me think..., my first step is to find the coffee and then heat it up.Looking around the room, I notice a **countertop**, a dining table, and a small coffee table. Hmm... I suppose the **countertop is** the most likely place to start. I should check there first.

#2 Thought: I noticed **#9 Thought:** that there is **no coffee** on the countertop. It seems that I should look for it elsewhere. Wait, it seems that I have seen a microwave oven. This before... can be used to heat up.

After I find coffee.

Next, I will heat up coffee. I have already seen the microwave on the countertop



OpenAI o3-mini Success Rate: 43.4%



Lacking planning:

o3-mini first came to the microwave instead of looking for coffee.



Forgot Task:

o3-mini forgot to open the cabinet door and check.



Repetitive Search:

o3-mini repeatedly comes to the same area for inspection.

