

Category 1: Numerical

Question 1 (1 point): How many red pillows are on the sofa?



Response: Print 2 in terminal.

Category 2: Object Reference

Question 2 (1 point): Find the lamp on the nightstand that has the photo on it.



Response: Bounding box of the object

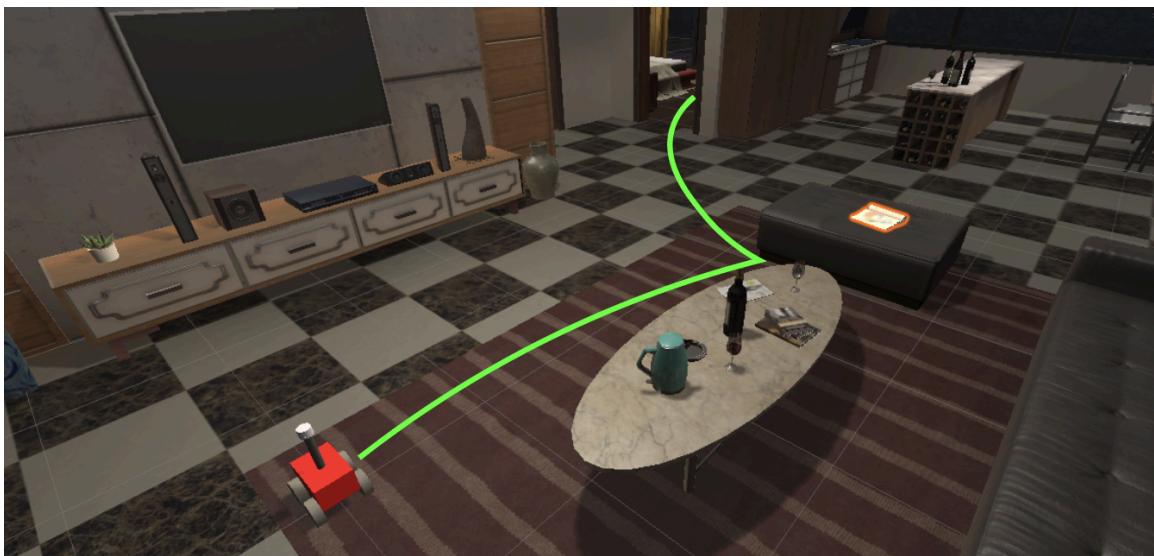
Question 3 (1 point): Find the speaker on the TV cabinet closest to the potted plant on the TV cabinet.



Response: Bounding box of the object

Category 3: Instruction Following

Question 4 (2 points): Go near the magazine on the ottoman, then go to the potted plant on the dressing table.

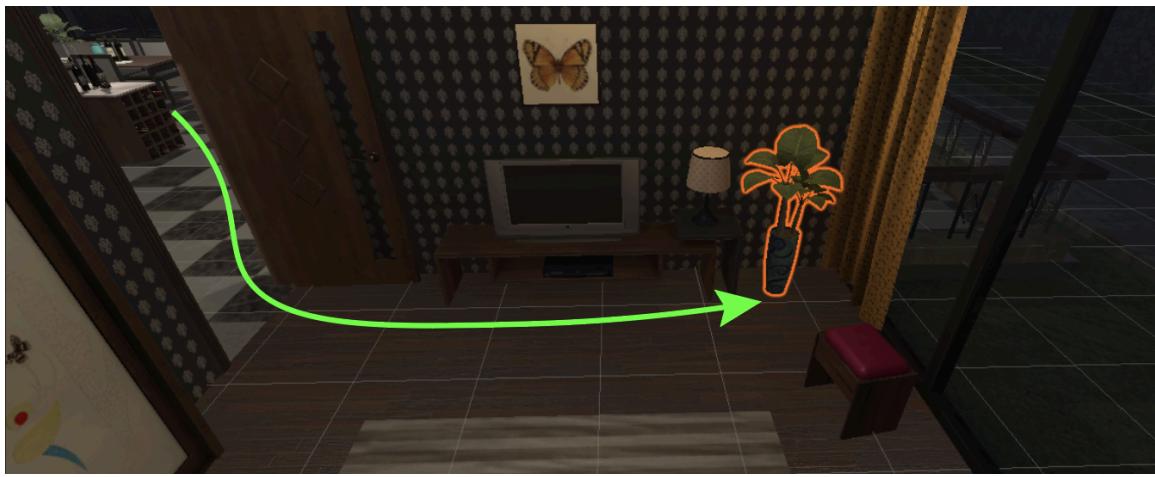




Response: trajectory (series of waypoints)

Question 5 (3 points): Take the path between the sofa and the coffee table and go to the kettle on the dining table, then go to the potted plant between the curtain and the TV.





Response: trajectory (series of waypoints)

Note: The number of points of a question is based on the number of steps in the response. For example, Question 5 expects 3 steps in the response and is worth 3 points. The score is calculated as the number of correctly responded steps.