

Category 1: Numerical

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



Response: Print 6 in terminal.

Category 2: Object Reference

Question 2 (1 point): Find the picture closest to a window.



Response: Bounding box of the object

Question 3 (1 point): Find the fossil decoration closest to the phone.



Response: Bounding box of the object

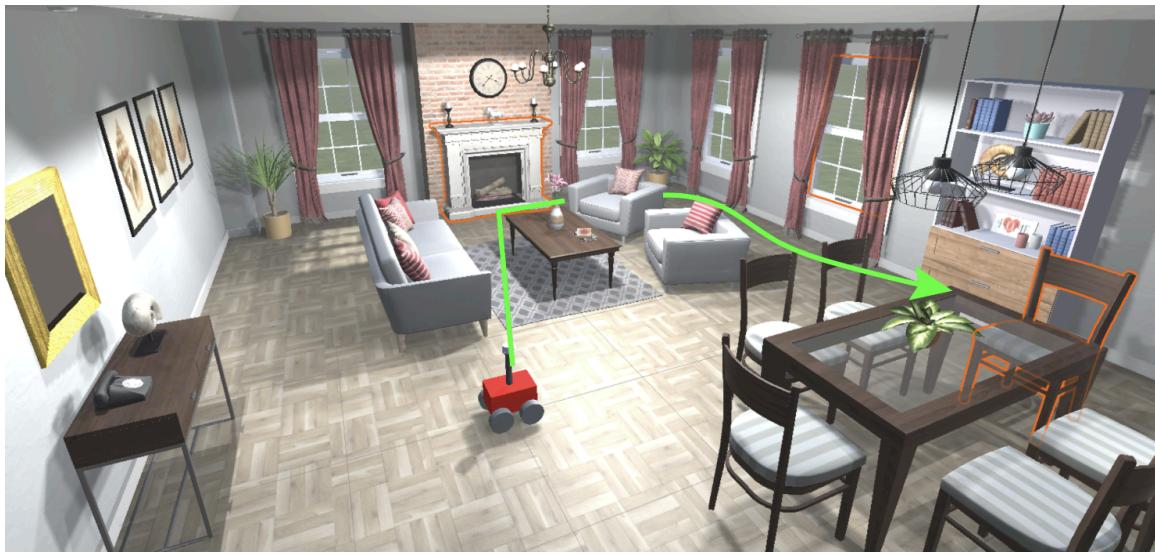
Category 3: Instruction Following

Question 4 (2 points): Go near the chair closest to the bookcase and stop at the table with the flowers on it.



Response: trajectory (series of waypoints)

Question 5 (3 points): First, go near the fireplace, then go to the window closest to the bookcase, and stop at the chair farthest from the mirror.



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.