

Category 1: Numerical

Question 1 (1 point): How many sofas are below a window?



Response: Print 2 in terminal.

Category 2: Object Reference

Question 2 (1 point): Find the pillow closest to the book on the stool.



Response: Bounding box of the object

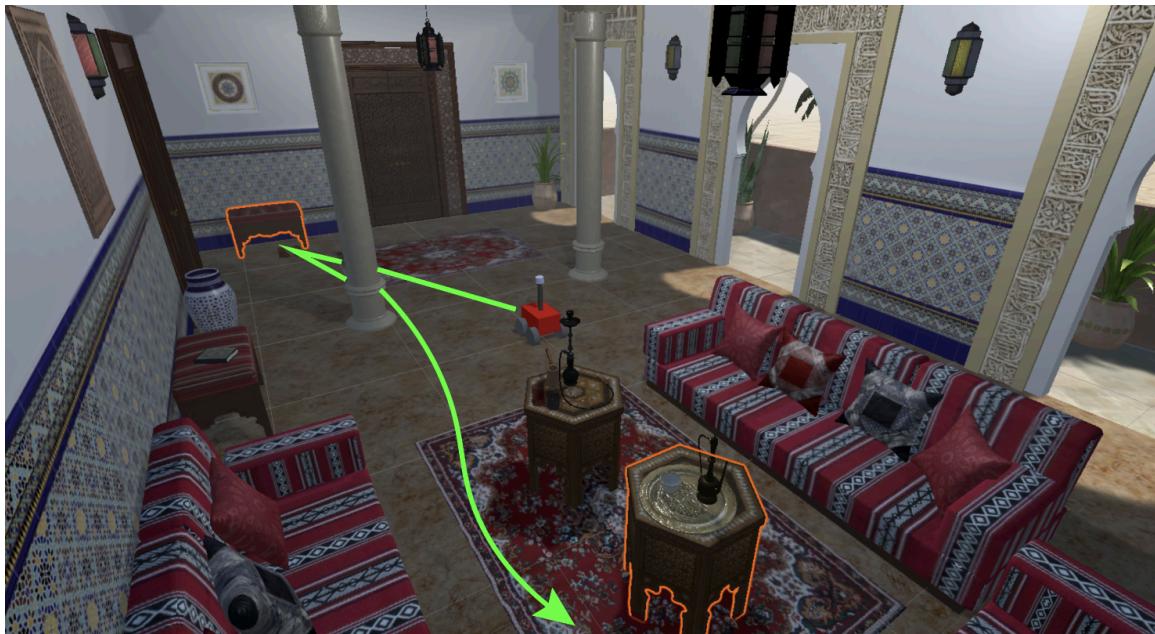
Question 3 (1 point): Find the wall lamp that is between a door frame and a window.



Response: Bounding box of the object

Category 3: Instruction Following

Question 4 (2 points): Go near the stool under the picture and stop at the small table farthest from the columns.



Response: trajectory (series of waypoints)

Question 5 (3 points): First, go to the potted plant furthest from the hookah, then take the path between the two columns, and stop at the tray on the table.



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.