

Category 1: Numerical

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



Response: Print 2 in terminal

Category 2: Object Reference

Question 2 (1 point): The blue chair that is closest to the cup of coffee



Response: Bounding box of the object

Question 3 (1 point): Find the potted plant between a vase and the cabinet with a TV on it.



Response: Bounding box of the object

Category 3: Instruction Following

Question 4 (2 points): Go to the cup near the TV remote and avoid the path near the cabinet



Response: trajectory (series of waypoints)

Question 5 (3 points): Go near the fireplace, pass by the stairs, then stop at the sphere decoration on the cabinet.



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.