**Input:** "A chat between a curious human and an artificial intelligence assistant. The assistant gives helpful, detailed, and polite answers to the human's questions. USER: <im\_start><image><im\_end>\nCan you control the robot to take the blue block and rotate it to the right? ASSISTANT: <ACT>"



