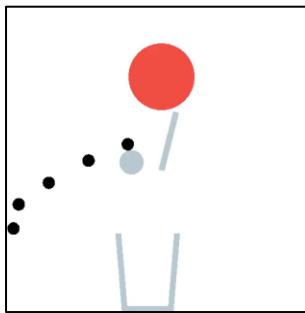
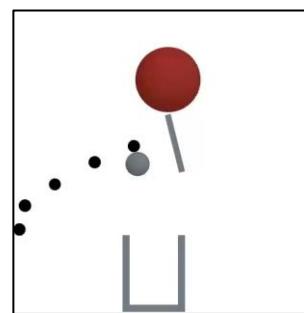
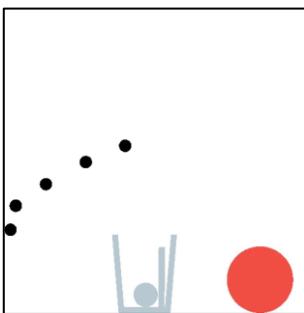


Case Study 4

Task: You are a coding-focused assistant whose primary job is to **OUTPUT RUNNABLE CODE**. Your response **MUST ALWAYS** include executable code. Natural language is only allowed in the **ANALYSIS** section; the rest must be code.....



GT Input frame 1 &10



Generated Output frame 1 &10

First-frame JSON:

```
{ "id": "red_ball", "category": "circle", "color_rgb": [240, 78, 67], "position": { "center_x": 261.02, "center_y": 116.06 }, "bbox": { "x_min": 288, "y_min": 256, "x_max": .... }, "id": "line_segment", "category": "circle", "color_rgb": [184, 200, 209], "position": { "center_x": .... }
```

Model Analysis: The scene is set up as a 2D physics puzzle on a vertical plane. Static structures include a U-shaped gray container resting at the bottom center and a decorative arc of five small black circles on the left side, which act as fixed obstacles or markers. The ground plane is effectively the bottom edge of the image.

GT Analysis: The scene is set against a plain white background with no visible floor or ground platform. Near the bottom center there is a U-shaped static gray structure, resembling a container or goal, formed by two vertical gray bars with an open top. Near the upper central region is a large red circular object connected to a short vertical gray bar or rod that appears fixed or hinged; this rod suggests a possible obstacle or paddle-like structure attached to

