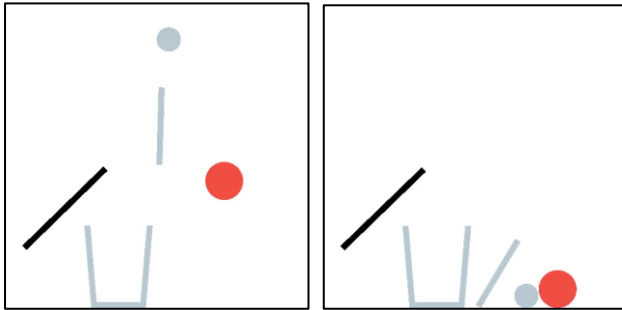
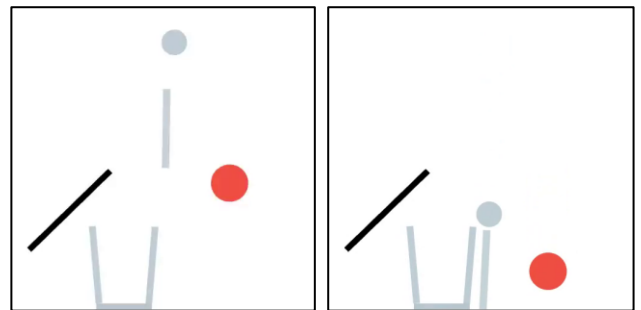


Case Study 1

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": "lightgray_ball", "category": "circle", "color_rgb": [192, 206, 214], "position": { "center_x": 275.16, "center_y": 58.34 }, "bbox": { "x_min": 254, "y_min": 36, "x_max": ....  
  { "id": "red_ball", "category": "circle", "color_rgb": [240, 78, 67].....
```

Model Analysis: The scene has a white background with a static slanted black bar on the left, a tall slim gray vertical bar near the middle, and a gray cup at the bottom made of two slightly tilted side walls and a flat base..... The black bar is....

GT Analysis: The scene is set in a mostly empty vertical space with no visible ground plane, but there is a U-shaped container near the bottom made of two light-gray..... there is a single black diagonal bar, slanted upward from left to right.....

Generated
Video

