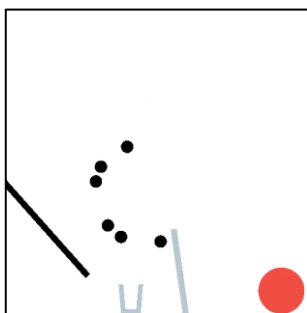
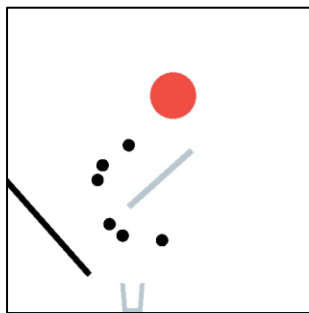
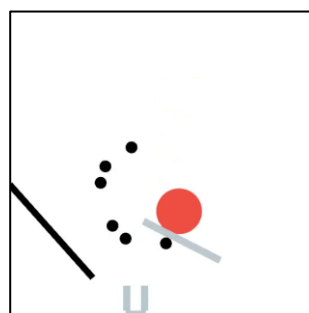
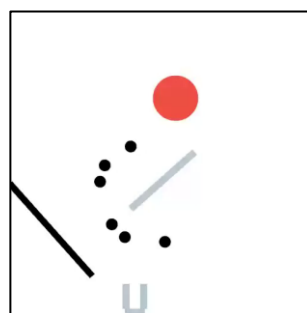


Case Study 6

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, 68], "position": { "center_x": 279, "center_y": 146 }, "bbox": { "x_min": 312, "y_min": 202, "x_max": ....  
{ "id": "line_segment", "category": "circle", "color_rgb": [192, 202, 208], "position": {  
"center_x": 257.14, "center_y": 285.48....
```

Model Analysis: The scene is a 2D physics puzzle contained within a 512x512 area. Static structures include a thick black diagonal barrier on the left acting as a ramp, a curved cluster of six small black circular pegs in the center-left, and a U-shaped container (cup) composed of three light-blue bars at the bottom center. A ground plane is implied at the bottom edge.....

GT Analysis: The scene appears to be a sparse environment with a white background. On the left side, there is a single thick black straight segment that is tilted upward from lower left toward upper right, acting as an inclined rigid bar or ramp. No clear ground line or enclosing boundaries are visible; the rest of the space is open. There are no obvious fixed supports other than this black bar. Dynamic objects include: (1) a large red circle located in the upper-right.....

Generated
Video

