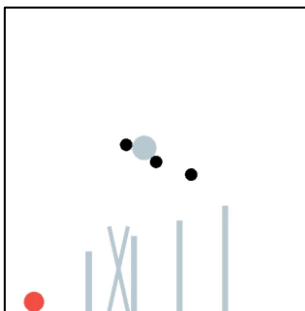
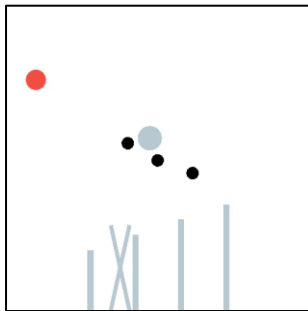
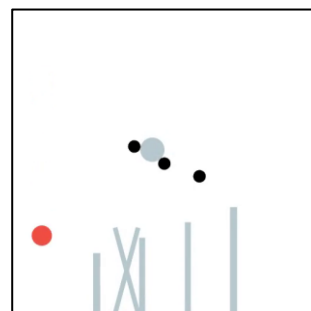
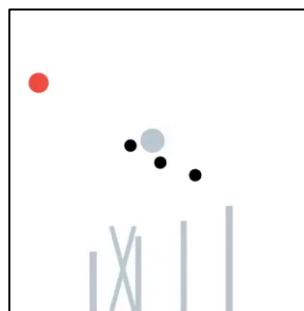


Case Study 5

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": 48.52, "center_y": 123.48}, "bbox": { "x_min": 262, "y_min": 222, "x_max": ....  
{ "id": "gray_ball", "category": "circle", "color_rgb": [0, 0, 0], "position": { "center_x": 241.7, ....
```

Model Analysis: The scene presents a 2D physics setup with a mix of static structures and dynamic spheres. The static environment at the bottom consists of four gray vertical pillars of increasing height from left to right, and a tall, narrow X-shaped structure to their left. Suspended in the center are three small, fixed black circles arranged in a descending diagonal, acting as rigid pivots or obstacles

GT Analysis: The scene is set in a mostly empty vertical space with a white background and no visible ground plane. At the lower right region there are several tall, thin, vertical light-gray bars acting as fixed obstacles or posts. One bar is on the far right, another slightly left of it, and two more on the left side that cross each other in an "X" pattern, forming a loose barrier or gate. There are no ramps or horizontal platforms; these vertical elements are the primary

Generated
Video

