Project2 Part1

State Space Definition and Justification

The **state space** in this project consists of the positions of all the objects in the container relative to the container's position. Each object's position is represented by its x and y coordinates in a two-dimensional space (since the movement within the container is primarily horizontal). The state space can be described as a vector that includes the relative positions of all objects.

The **justification** for this state space representation is that the arrangement and distribution of objects are critical for determining how well-mixed the objects are. Since the goal is to achieve a uniform mix of two types of objects, tracking the exact positions of all objects at each step provides essential information for the reinforcement learning agent to make effective decisions about how to move the container.

Action Space Definition and Justification

The **action space** consists of four possible actions: moving the container **Up**, **Down**, **Left**, and **Right**. These actions correspond to discrete movements of the container in the horizontal plane, which affect the positions of the objects within the container.

The **justification** for this action space is that these four basic movements are sufficient to induce a mixing effect within the container. By repeatedly moving the container in different directions, the objects inside will collide and spread out, promoting a more even distribution. Using a discrete set of actions simplifies the learning process for the reinforcement learning agent, as it does not have to choose from an infinite range of possible movements.

Reward Function and Justification

The **reward function** is designed to measure how well-mixed the objects are within the container. One possible approach is to define a reward based on the difference in the density of each type of object across different regions of the container. A more uniform distribution yields a higher reward, while a less uniform distribution yields a lower reward.

The **justification** for this reward function is that the primary objective of the reinforcement learning agent is to achieve a uniform distribution of objects. By assigning higher rewards to states where the objects are more evenly mixed, the agent is incentivized to take actions that promote mixing. This approach directly aligns the reward function with the ultimate goal of the project, ensuring that the agent learns behavior that optimizes the mixing process.