Video Prediction	**Real-world Execution**

Grasp the Red Apple: There are multiple red objects on the table. The robot is instructed to grasp the red apple using its left arm. The robot needs to: **(0) identify the red apple among all red objects**, (1) secure its hold on the apple, (2) lift the apple as required.

Grasp the Blue Cube: There are multiple cubes of different colors on the table. The robot is instructed to grasp the blue cube using its right arm. The robot needs to: **(0) identify the blue cube among all cubes**, (1) secure its hold on the cube, (2) lift the cube as required.

Lift the dish: There is **a green dish** on the **unseen table** in front of the **unseen cupboard**. The robot is instructed to lift the dish using both arms. The robot needs to: (0) align the grippers with the edge of the dish, (1) secure its hold on the dish, **(2) lift upwards simultaneously**.

Clean the bowl: There is **a pink bowl and a sponge** on the **unseen table** in front of the **unseen cupboard**. The robot is instructed to clean the bowl using both arms. The robot needs to: (0) align the grippers with two objects, (1) secure its hold, **(2) lift upwards simultaneously**.