

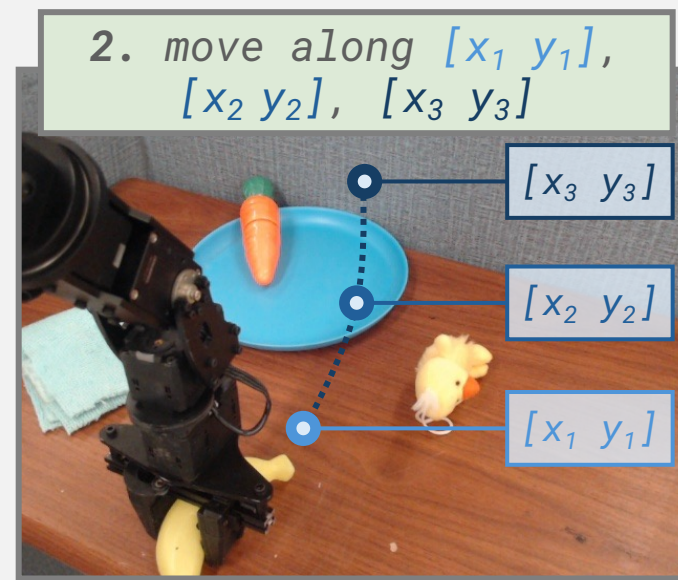
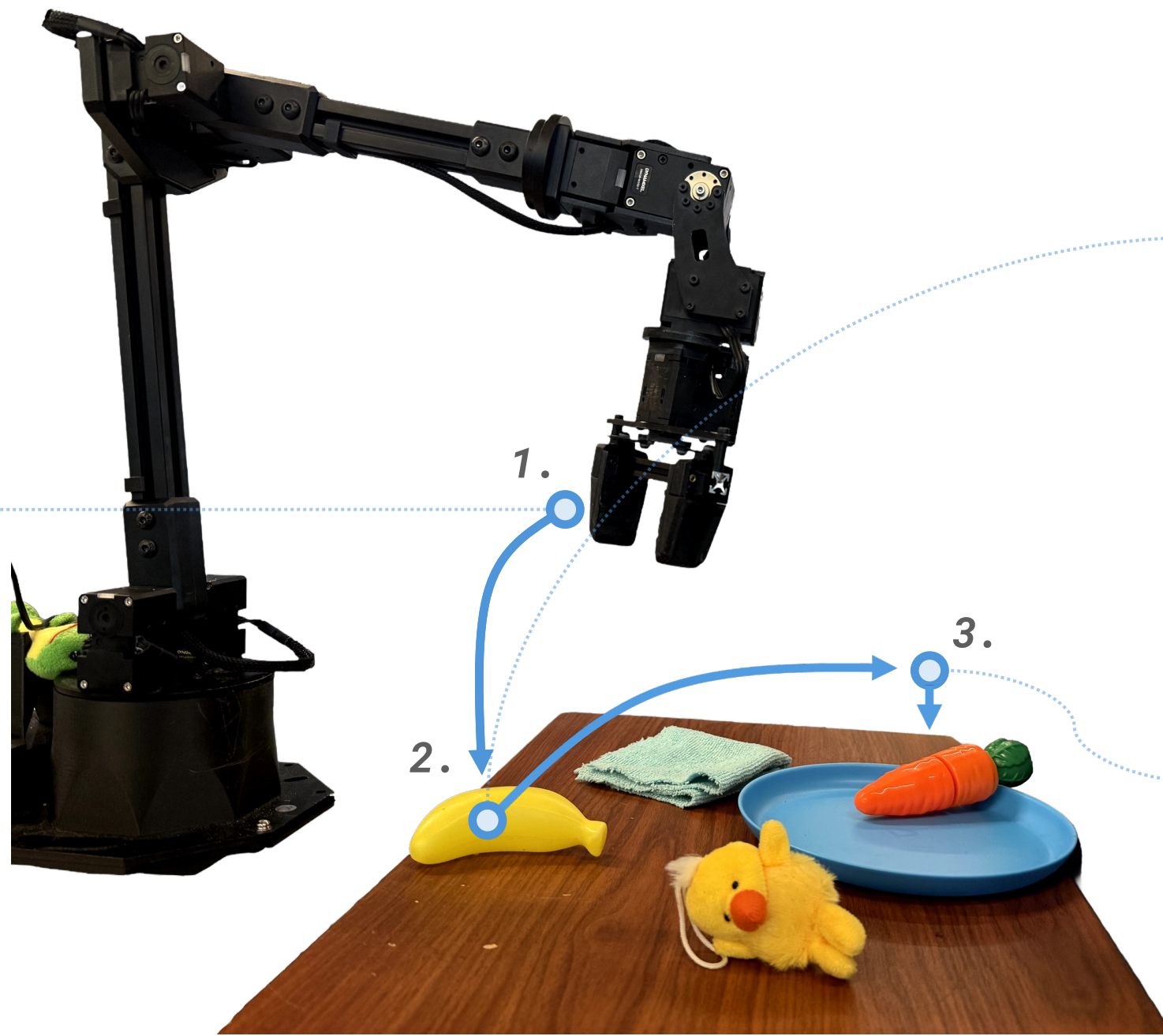
Put all the food on the plate

High Level VLM:

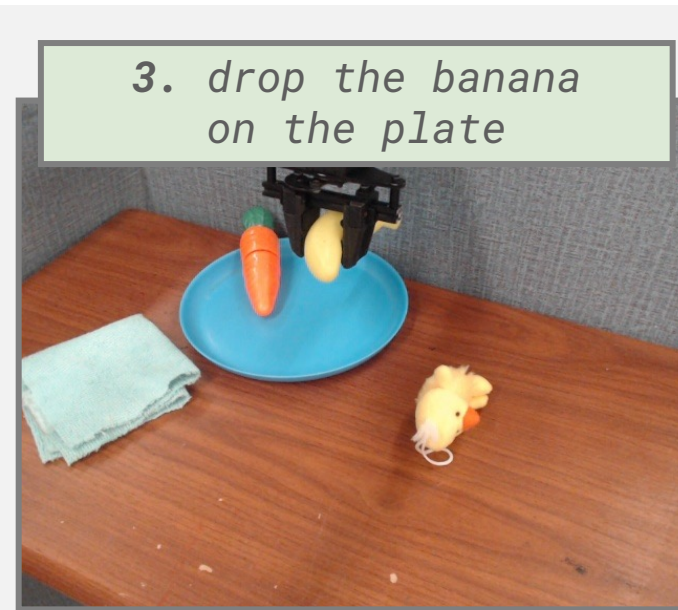
The carrot is already on the plate. The robot must now go to the banana.

To avoid confusing the banana with the yellow duck toy, a **pointing** and **motion** command is best.

1. move back to the banana at $[x\ y]$



2. move along $[x_1\ y_1]$, $[x_2\ y_2]$, $[x_3\ y_3]$



3. drop the banana on the plate

Old "Standard" Language

Task-level Commands

- Put the carrot on the cloth
- Wipe the table with the towel
- Put the banana next to the duck
- Stack the plate on the towel

New Detailed Steering Command Styles

Subtasks

- Wipe with cloth
- Lift the duck
- Grasp the carrot on the plate

Atomic Motions

- Move right
- Rotate clockwise and grasp
- Open gripper

Points

- Grasp at <carrot pixel position>
- Reach for <banana pixel position>

Traces

- Move from $[x_1\ y_1]$ to $[x_2\ y_2]$
- Go along $[x_1\ y_1]$, $[x_2\ y_2]$, $[x_3\ y_3]$