

SAPIEN

A SimulAted Part-based Interactive ENvironment

Fanbo Xiang¹ Yuzhe Qin¹ Kaichun Mo² Yikuan Xia¹ Hao Zhu¹
Fangchen Liu¹ Minghua Liu¹ Hanxiao Jiang³ Yifu Yuan⁵ He Wang² Li Yi⁴
Angel X. Chang³ Leonidas J. Guibas² Hao Su¹

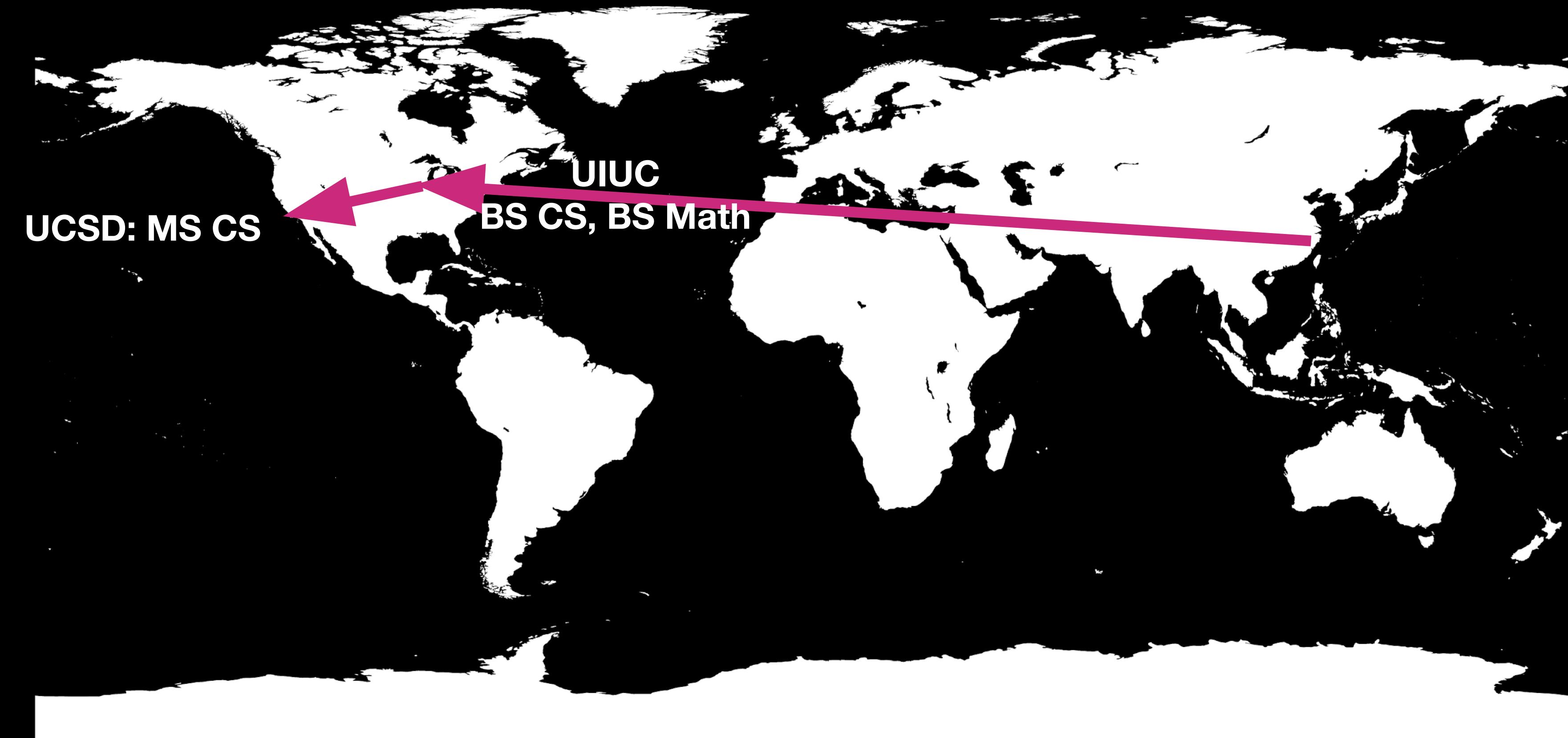
¹UC San Diego ²Stanford University ³Simon Fraser University ⁴Google Research ⁵UC Los Angeles

About Me

Fanbo Xiang, UCSD

Advisor: Hao Su, UCSD

Research: Graphics,
Vision, Robotics

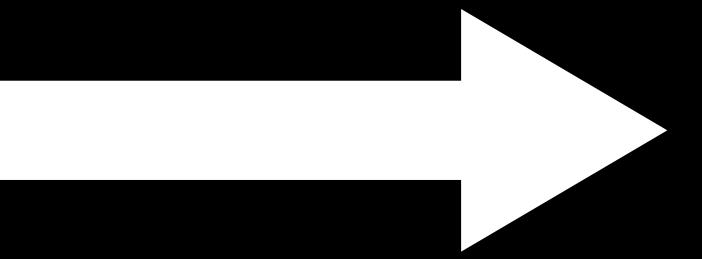


Outline

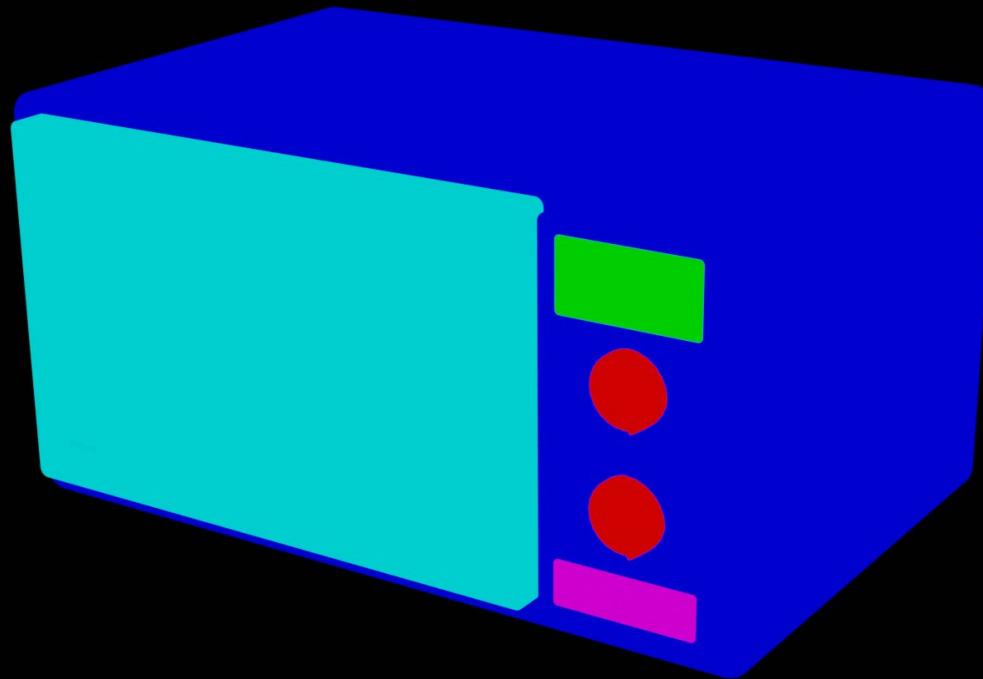
- Background
 - Intelligence, vision and robotics
 - Simulated environment
- SAPIEN architecture
 - Physics and robotics
 - Renderer
 - Assets
- Future research problems



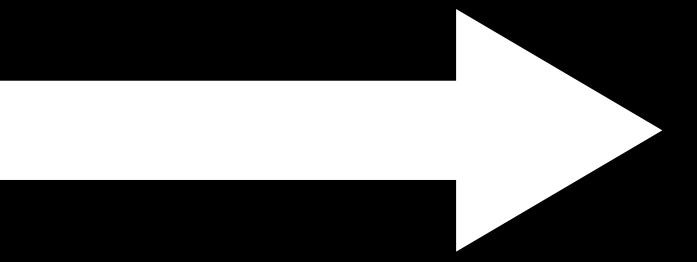
Intelligent Agent



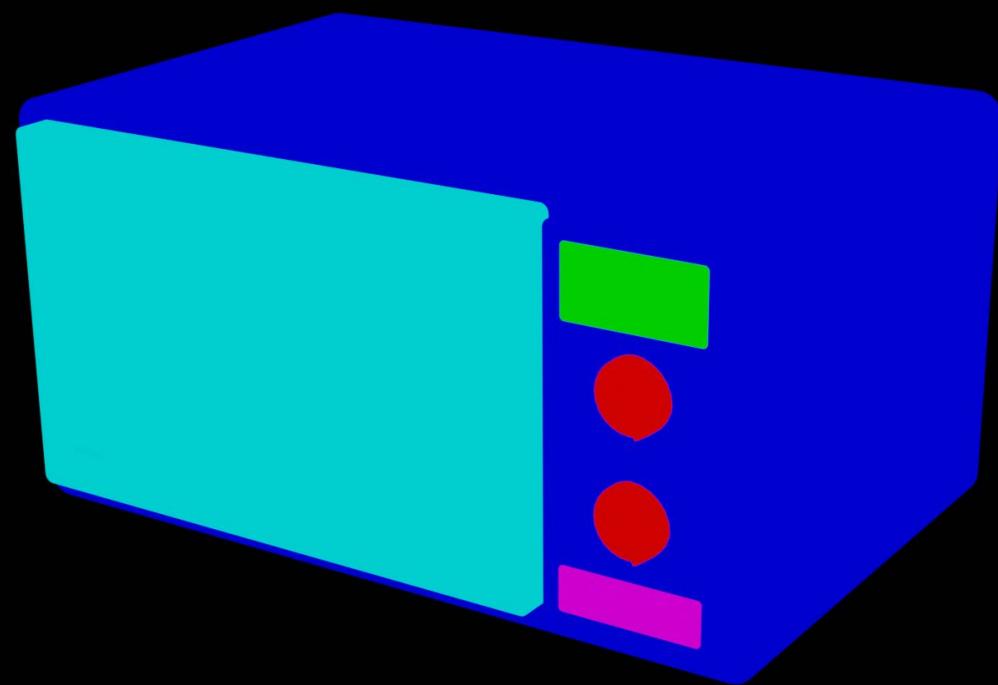
Perceive information



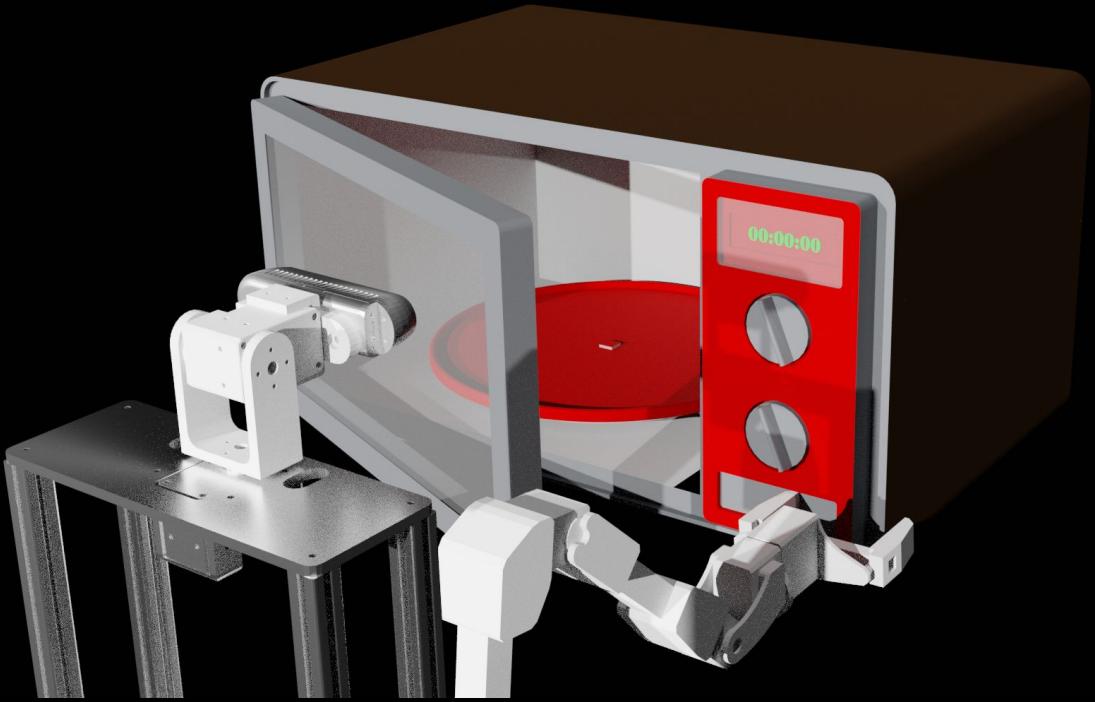
Intelligent Agent



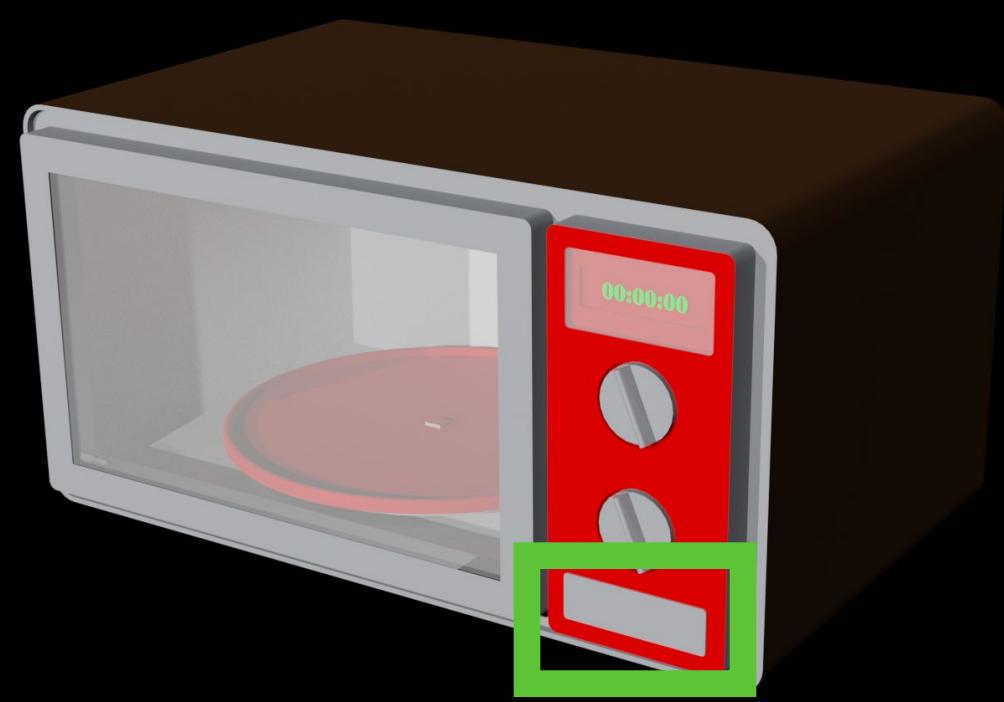
Perceive information

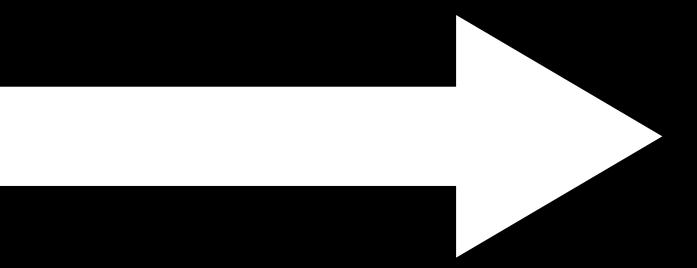


Intelligent Agent

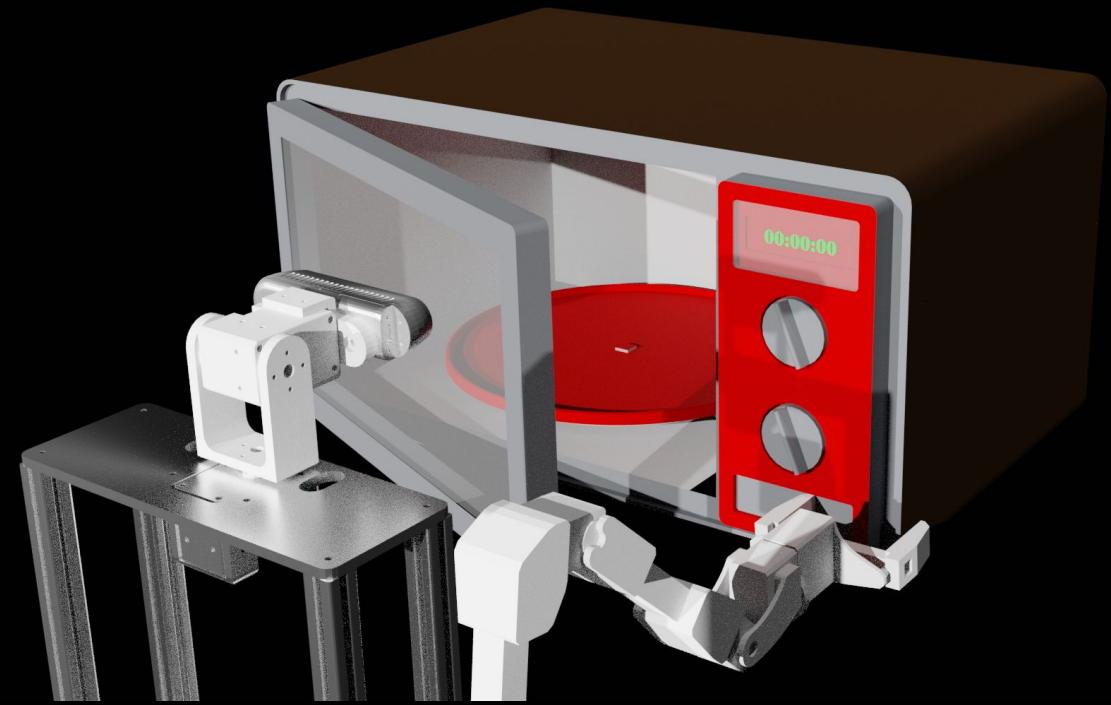
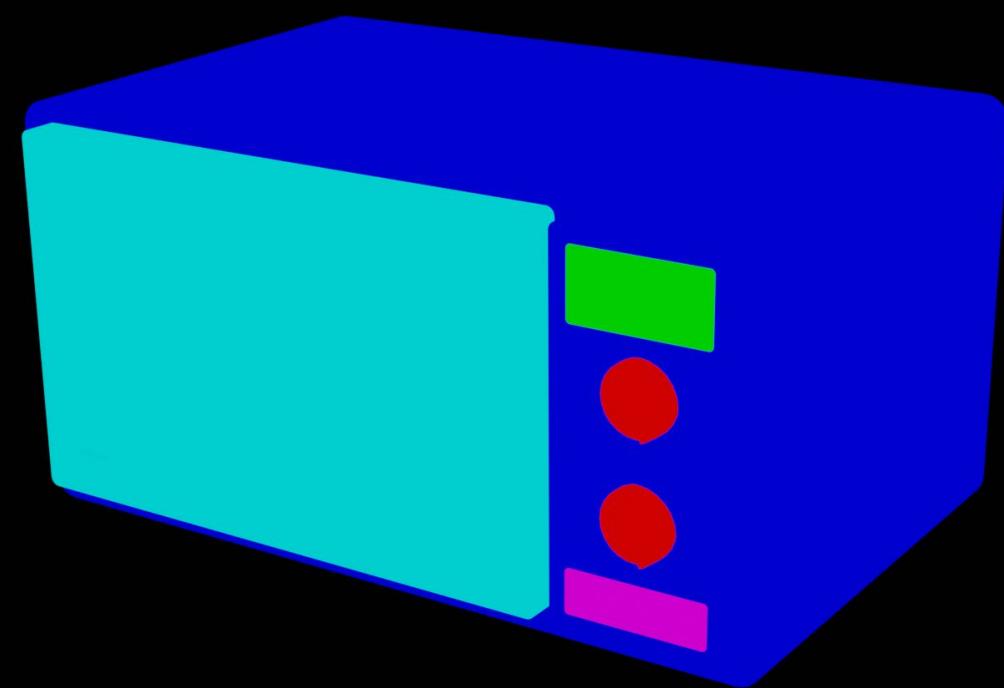


Adapt behavior

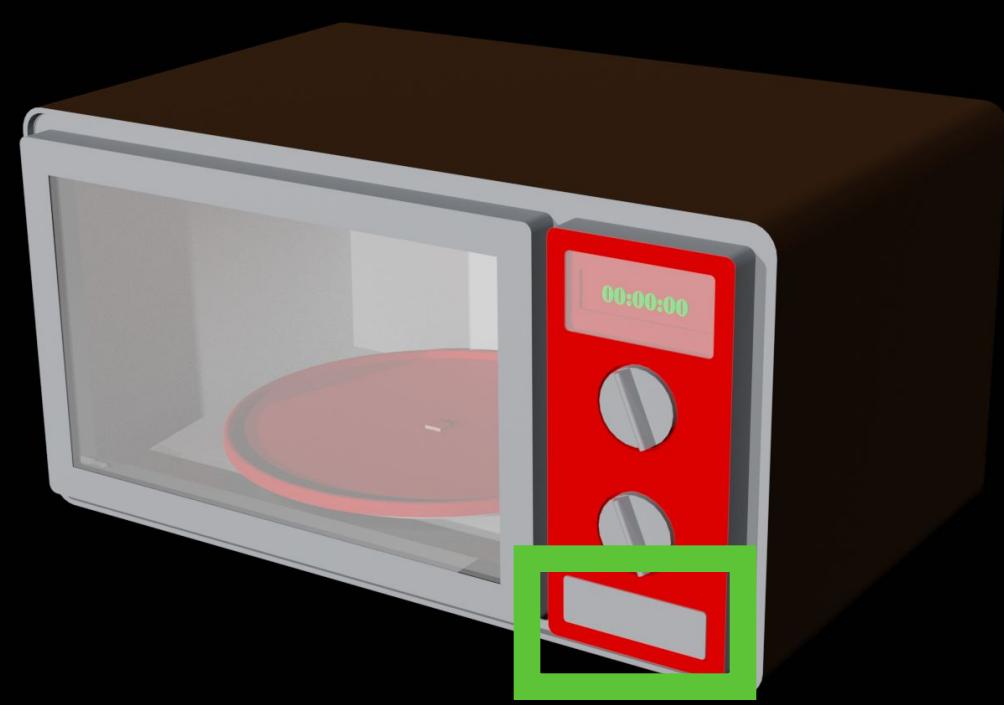


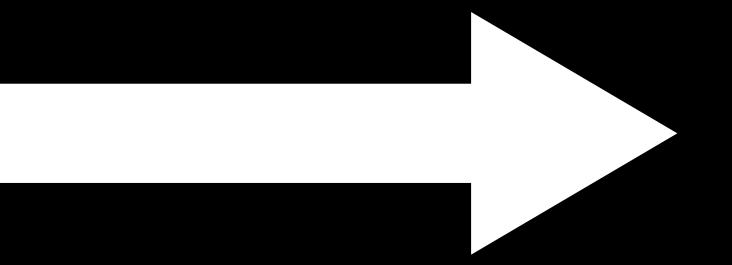
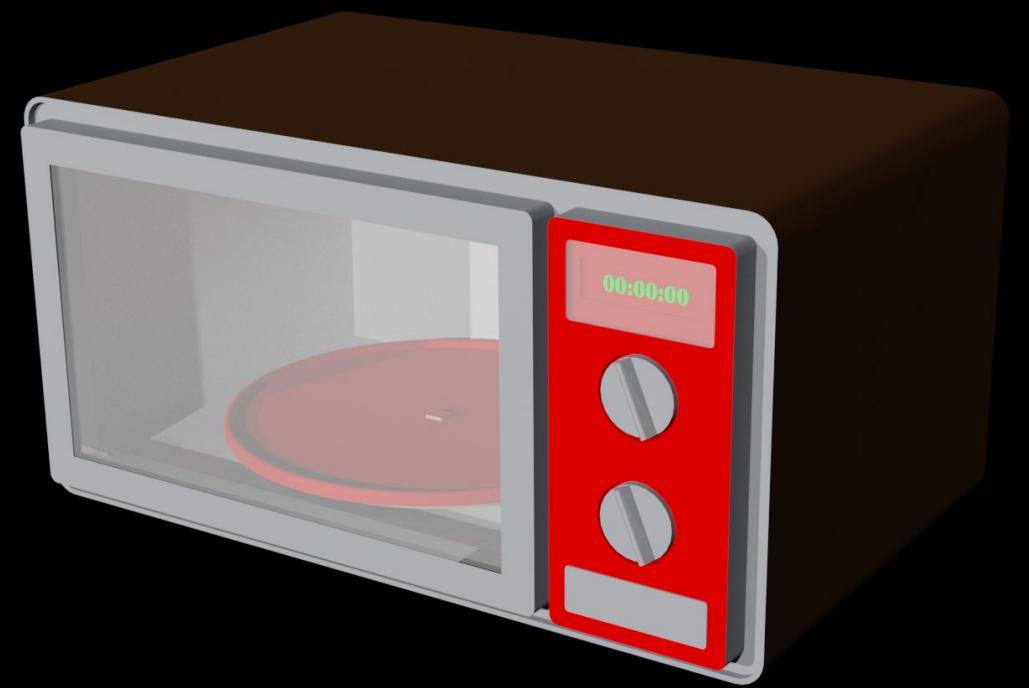


Perceive information

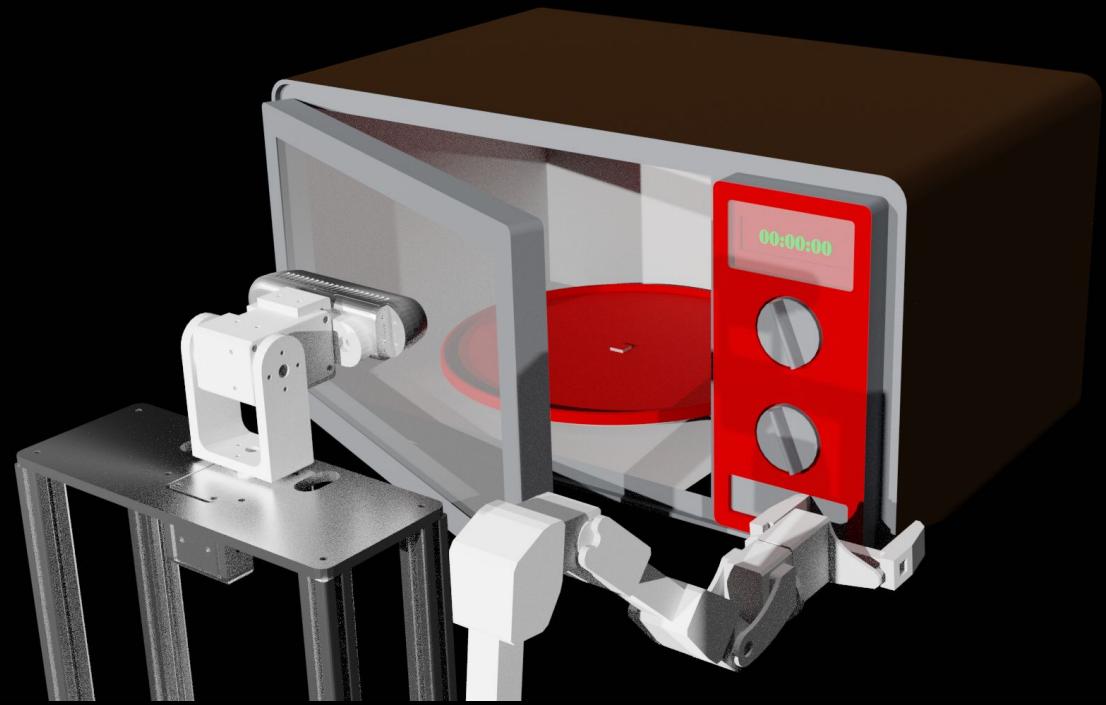
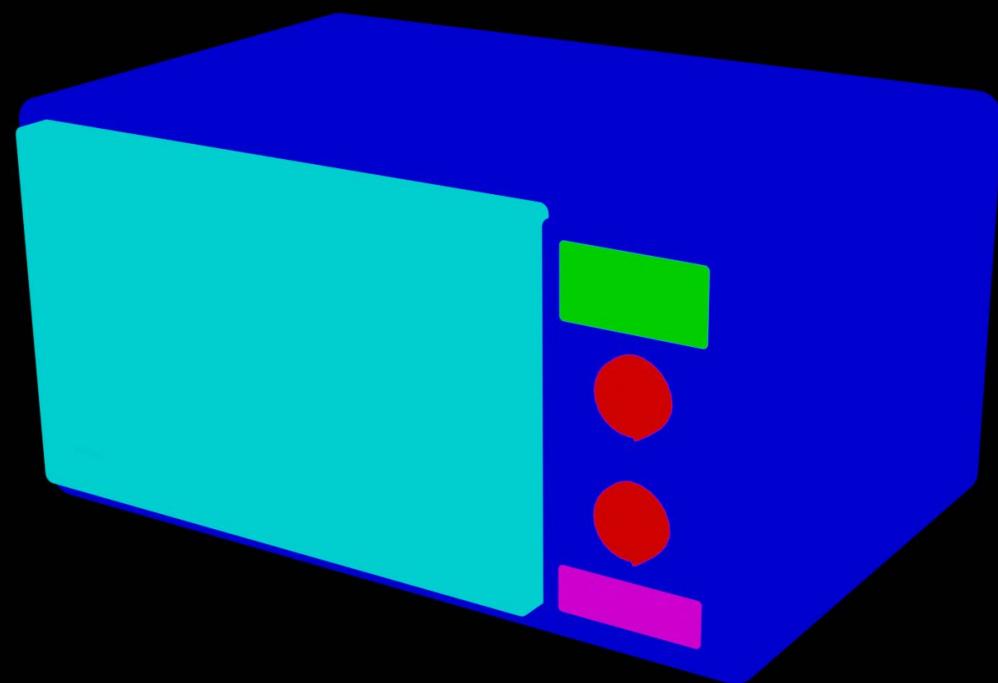


Adapt behavior

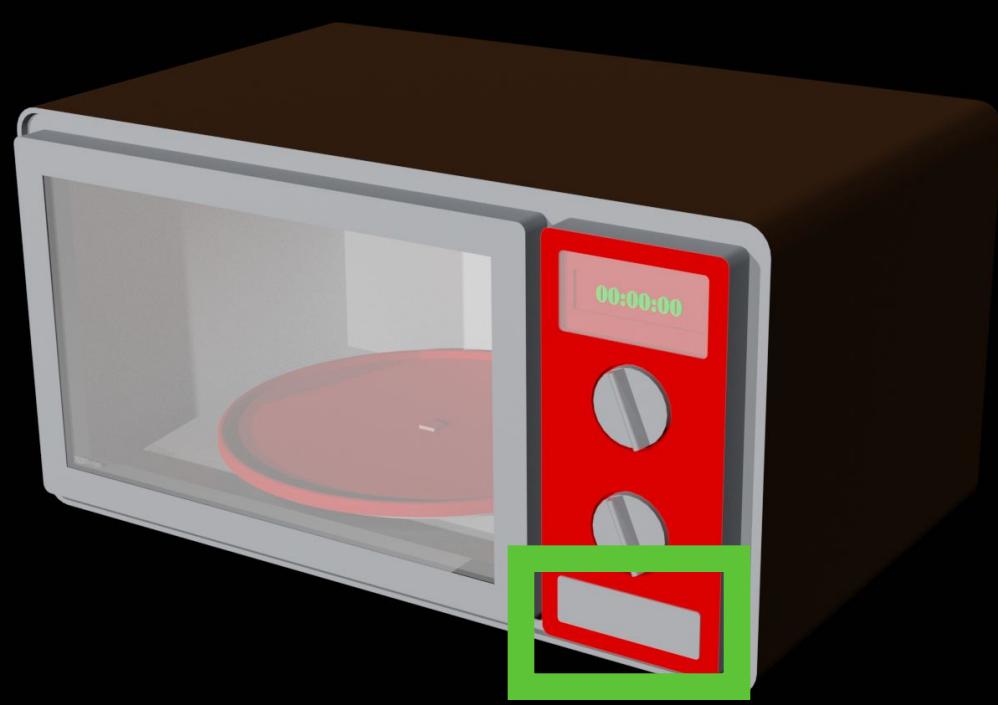


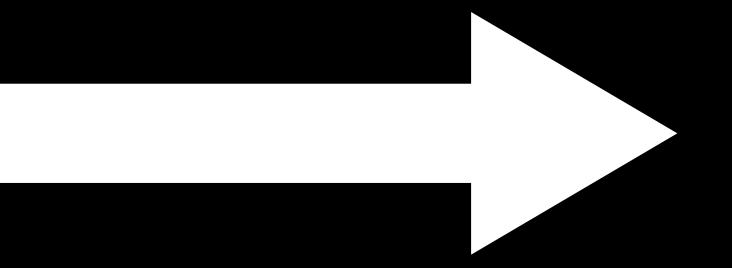
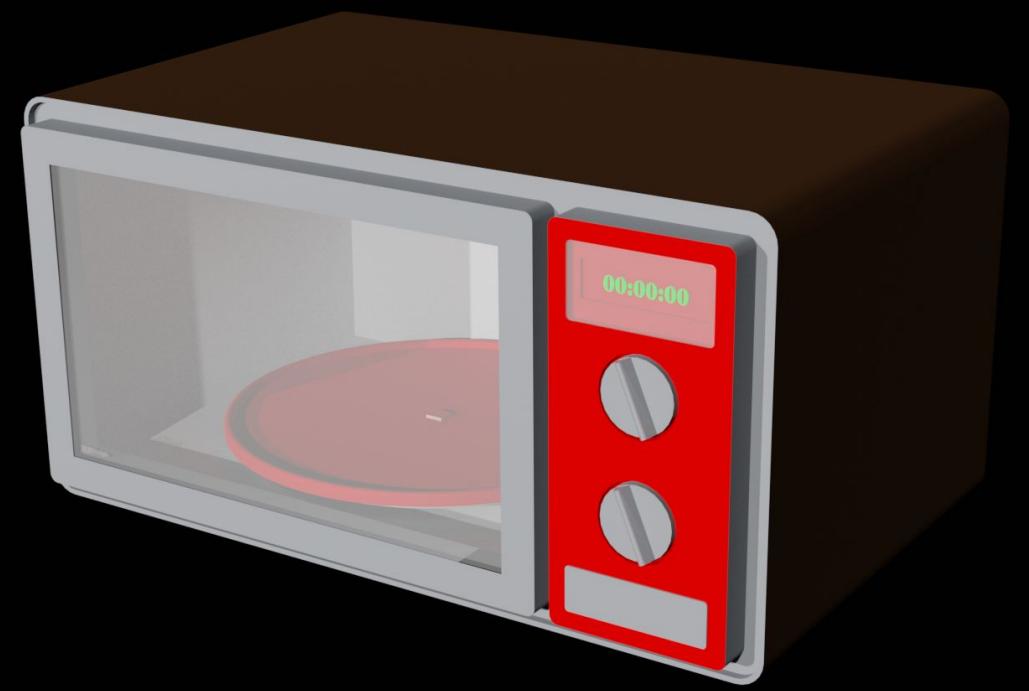


Perceive information
Computer vision

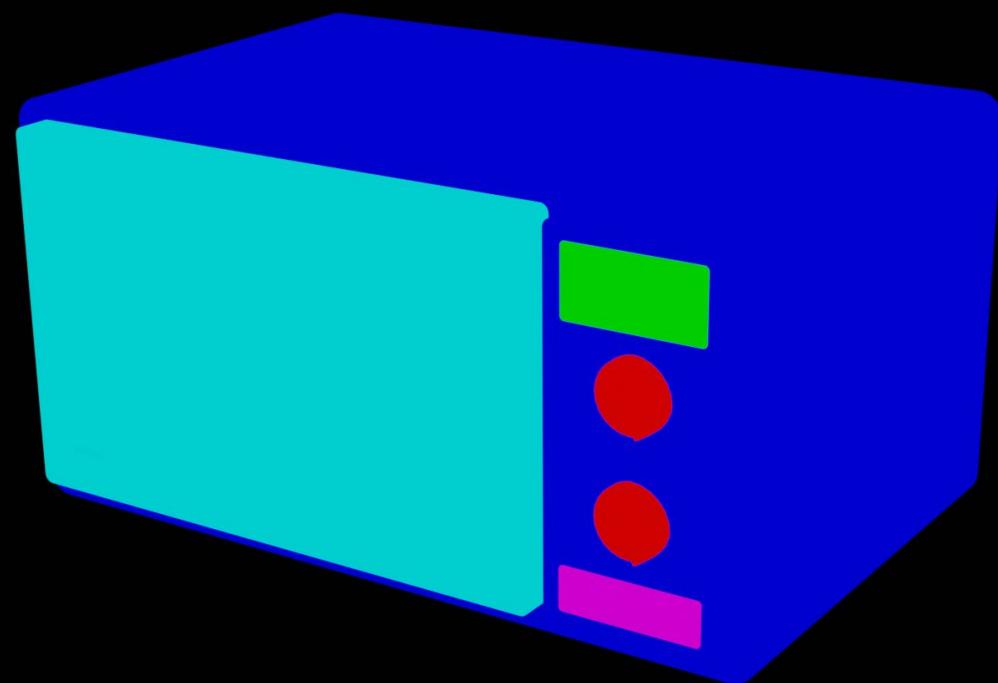


Adapt behavior

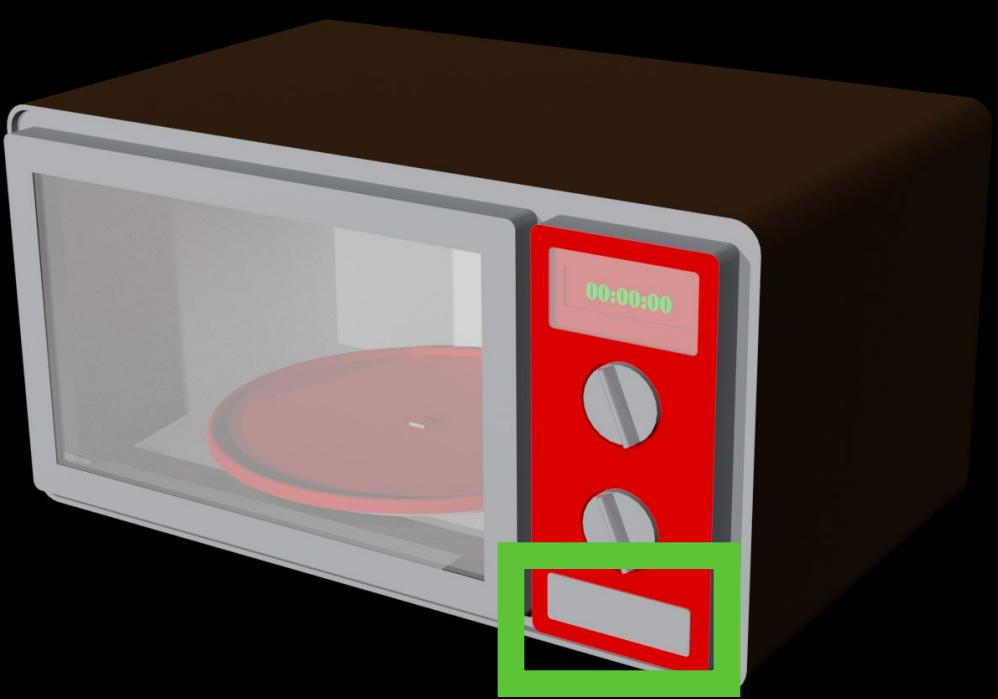
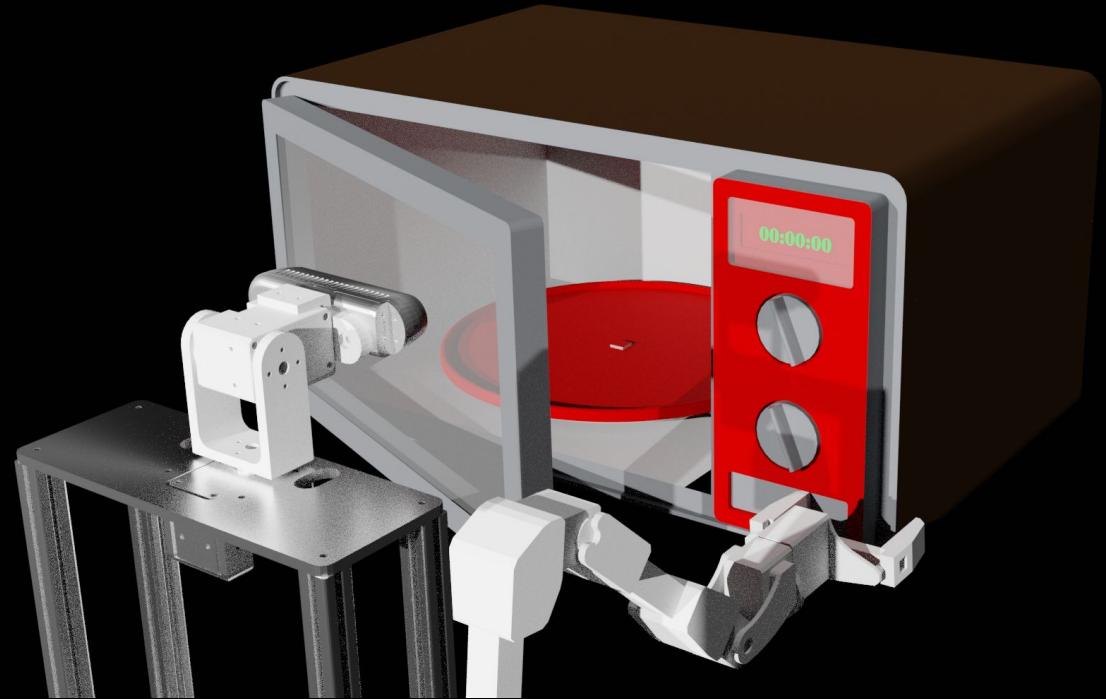


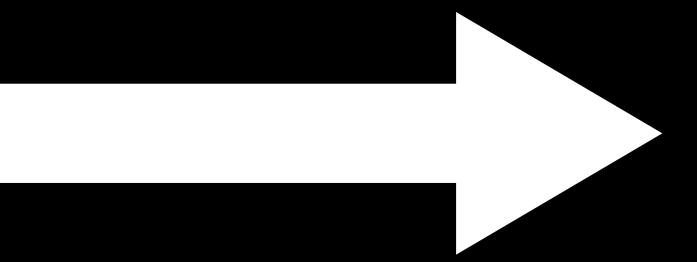


Perceive information
Computer vision

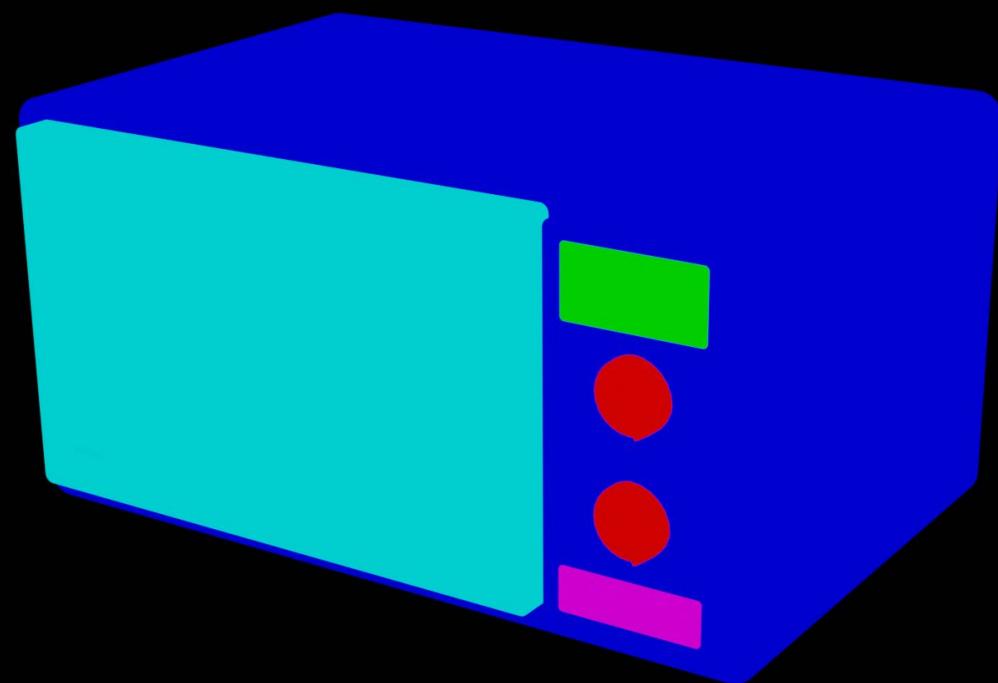


Robotics control
Adapt behavior

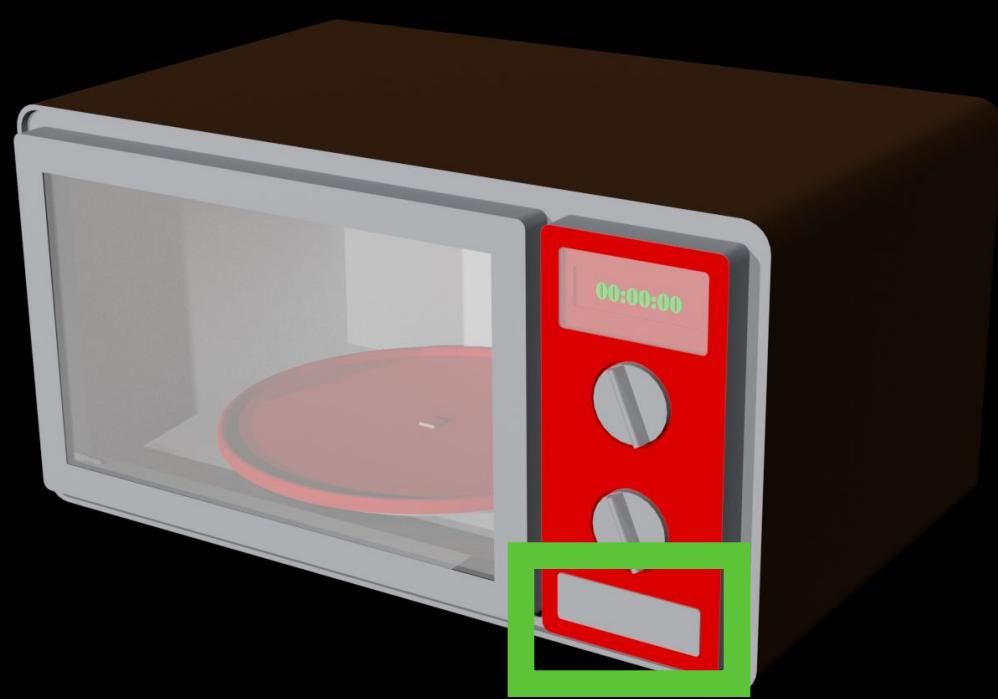
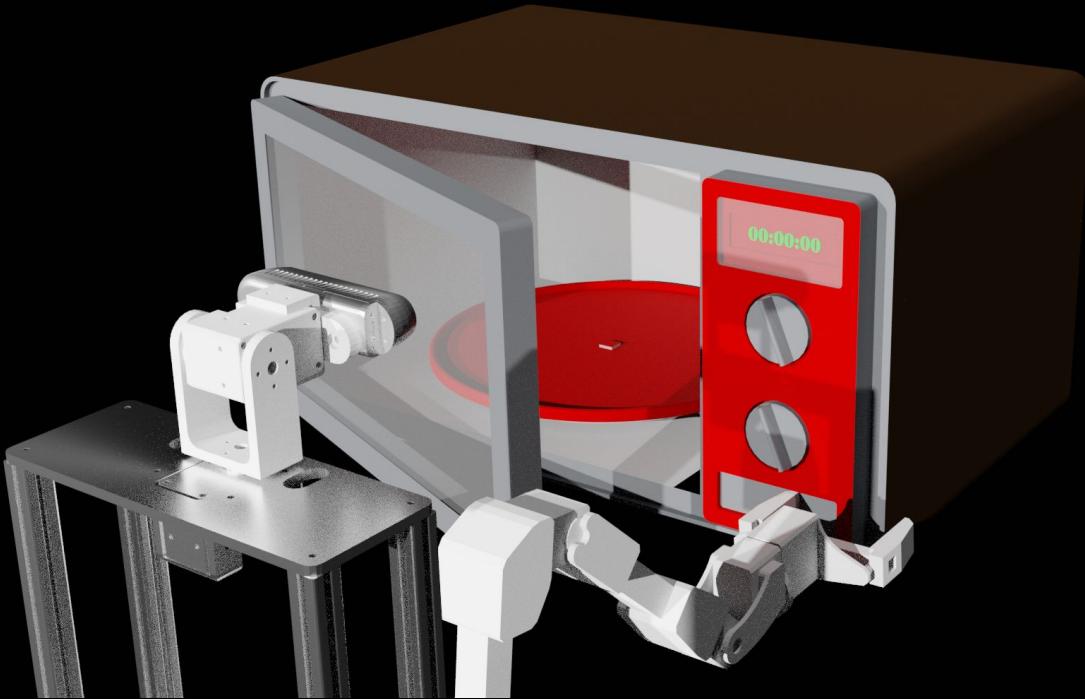


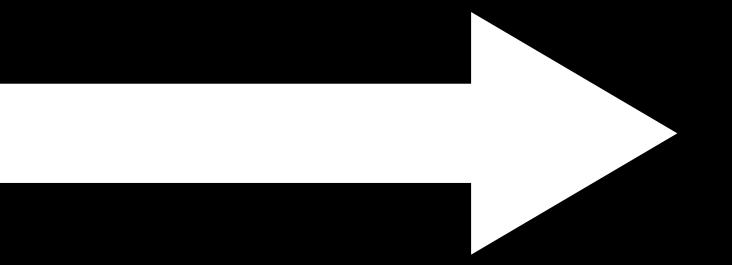
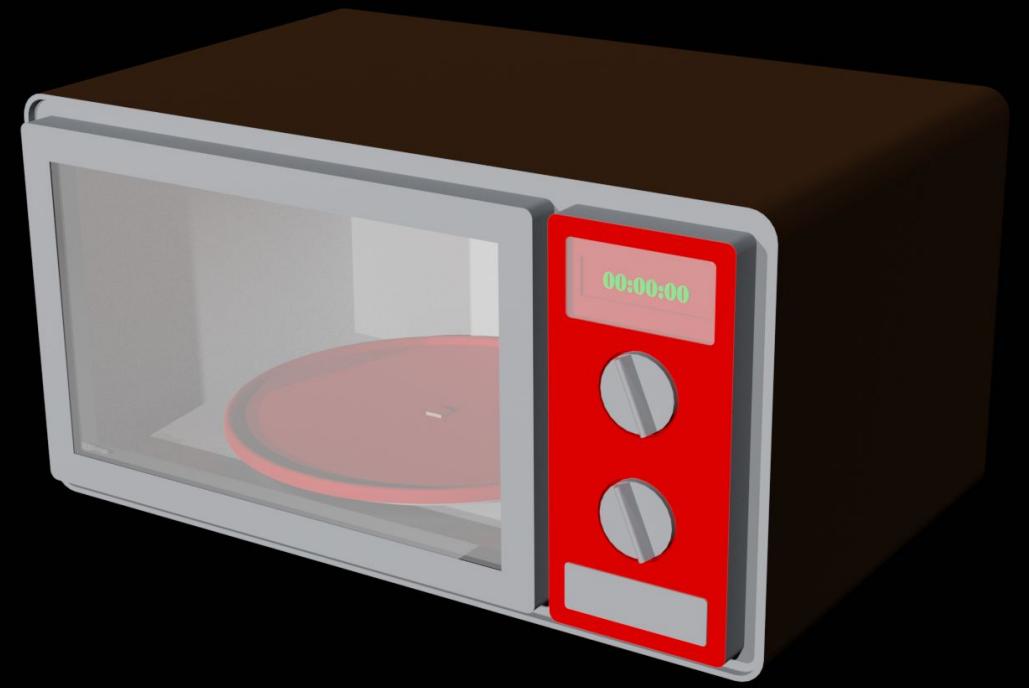


Perceive information
Computer vision

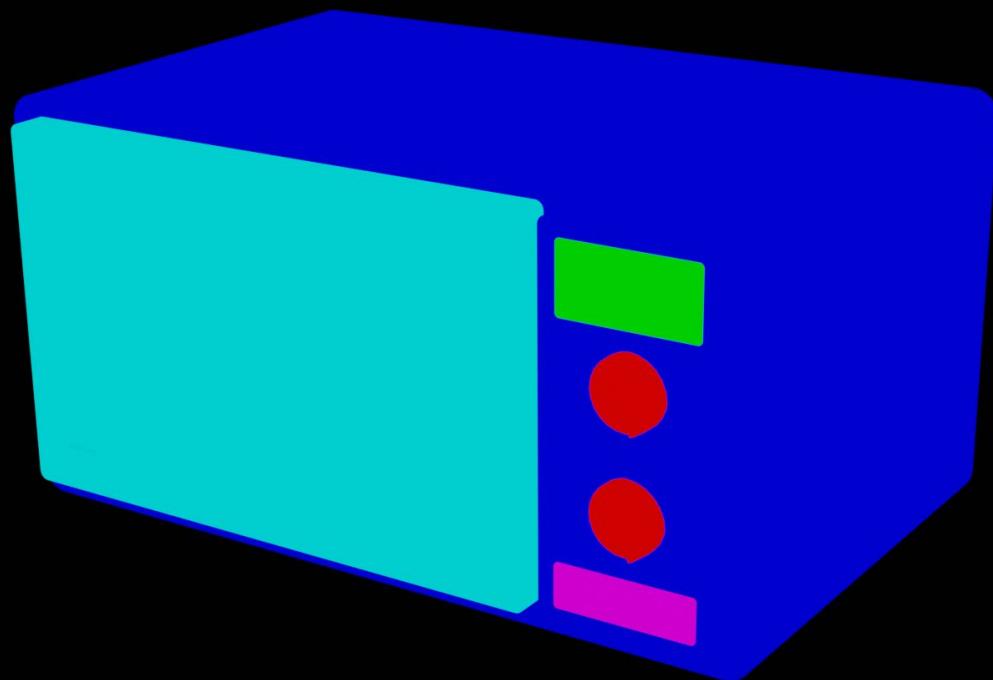


Robotics control
Adapt behavior

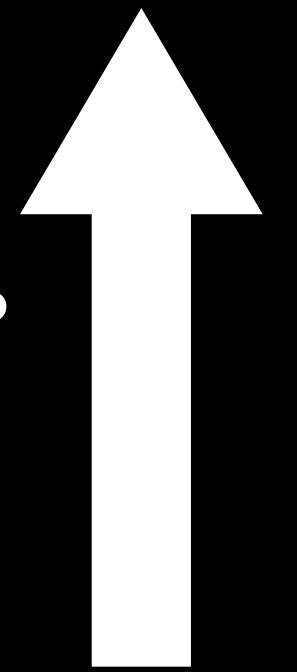




Perceive information
Computer vision



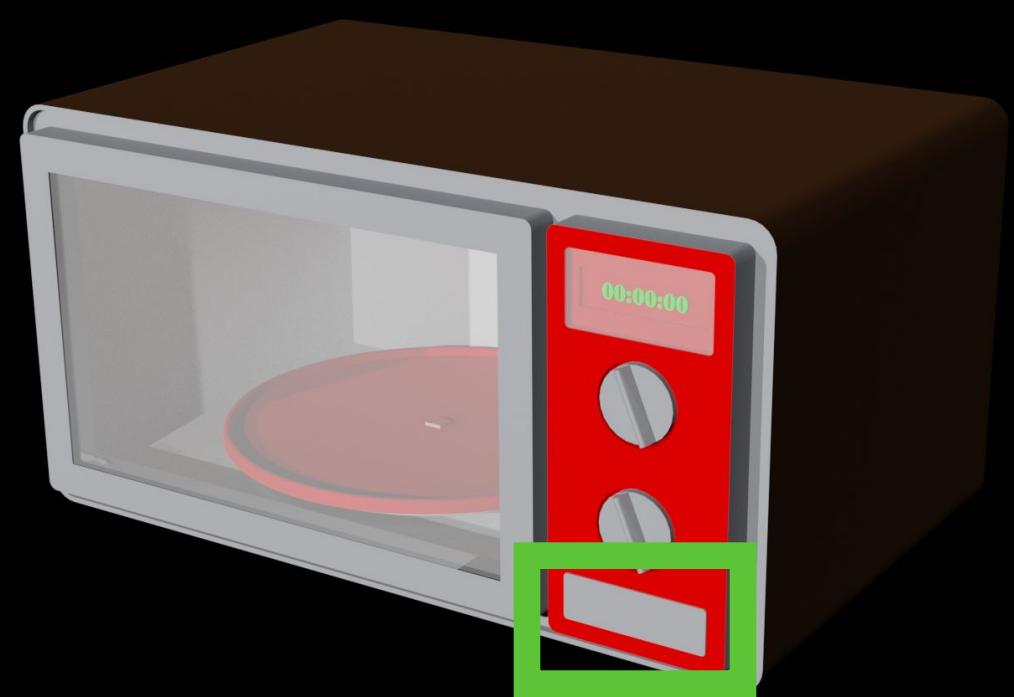
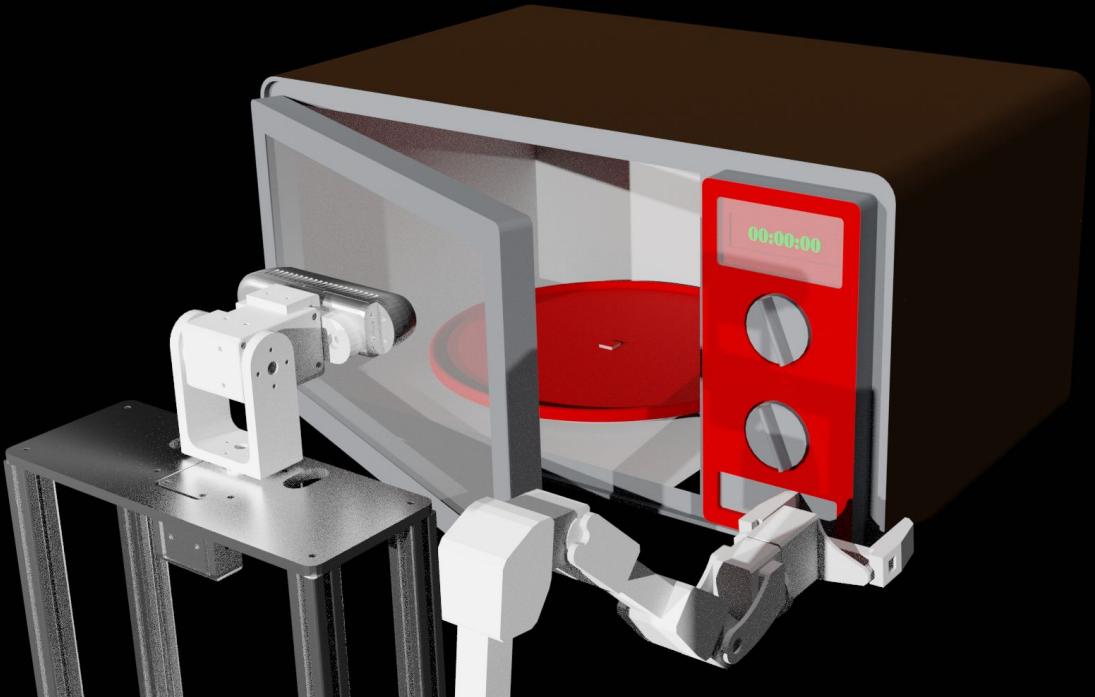
Active perception?

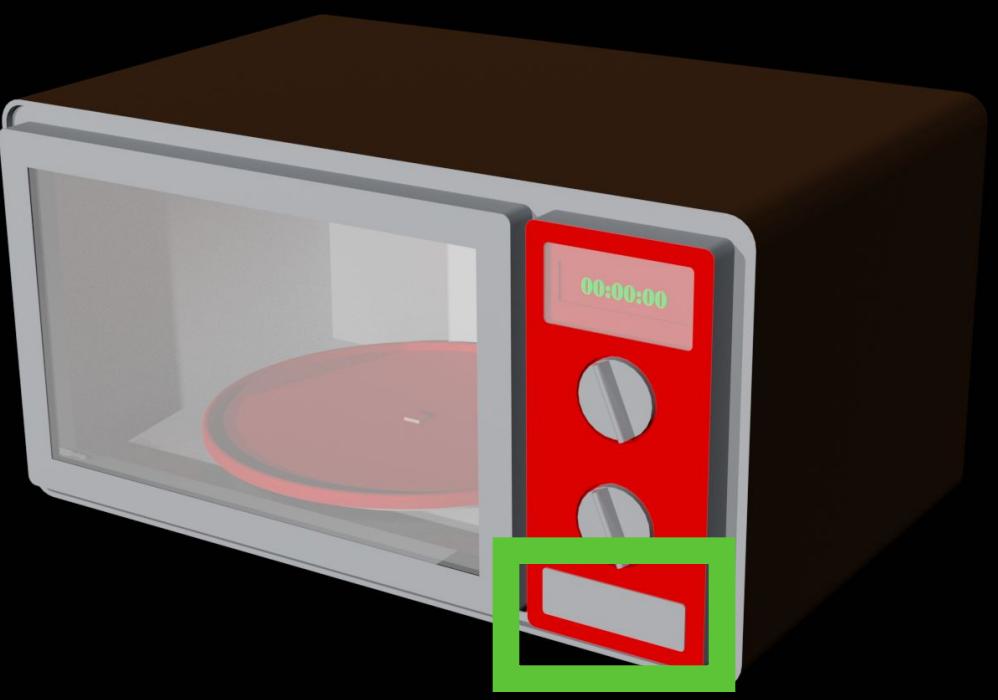
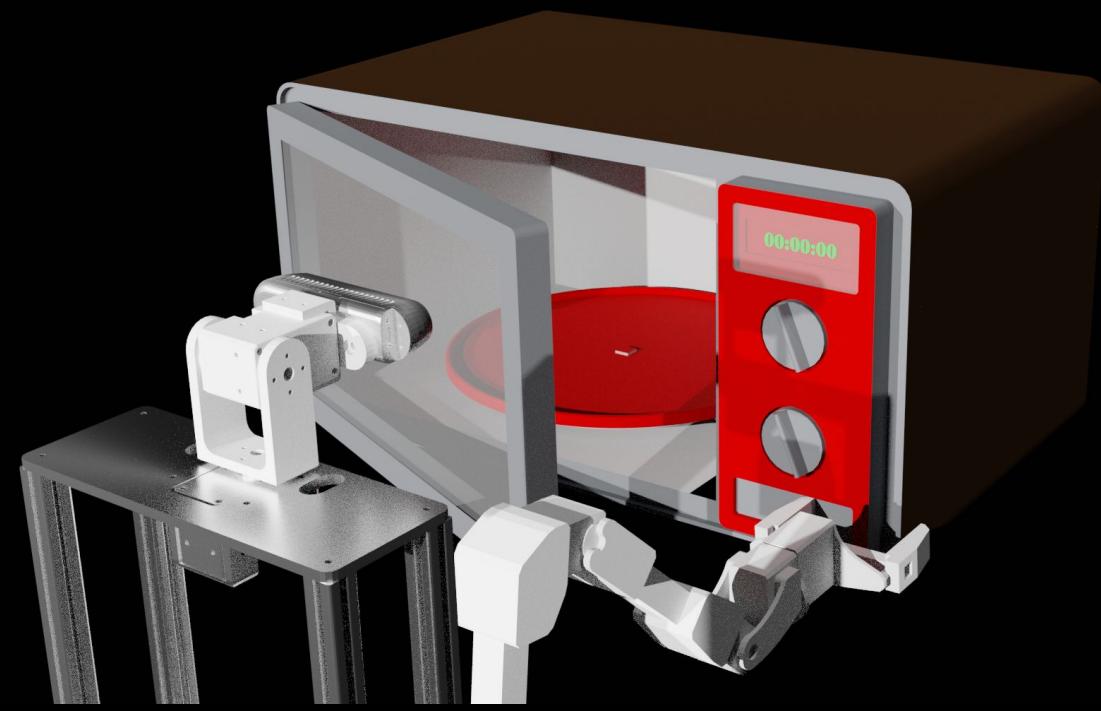
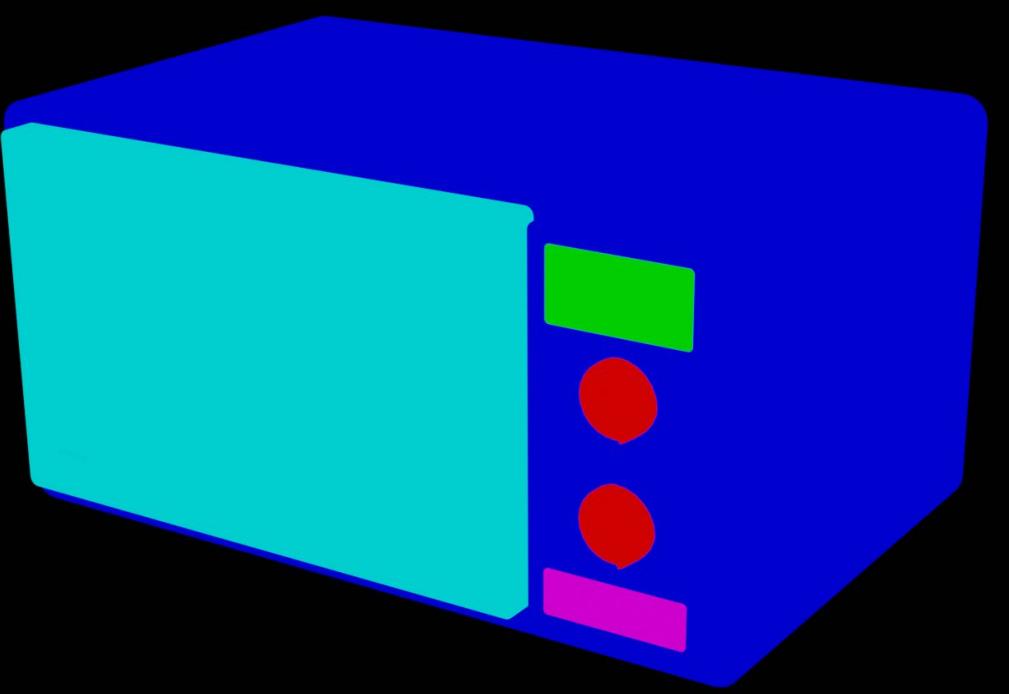
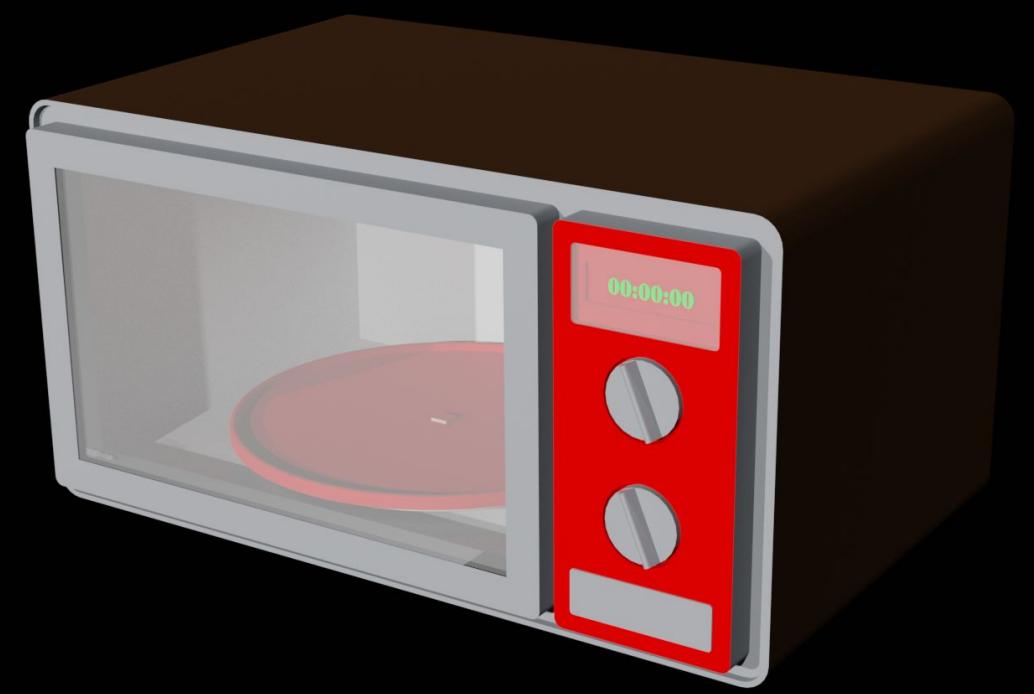


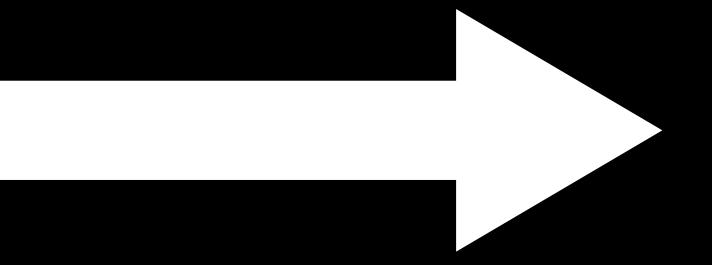
Planning?

Robotics control

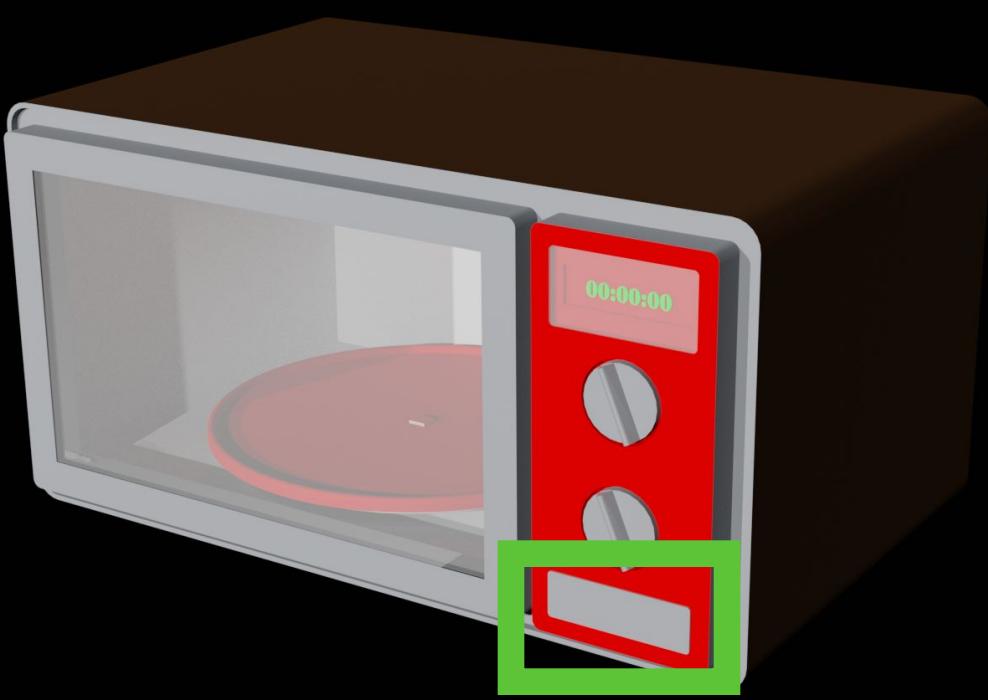
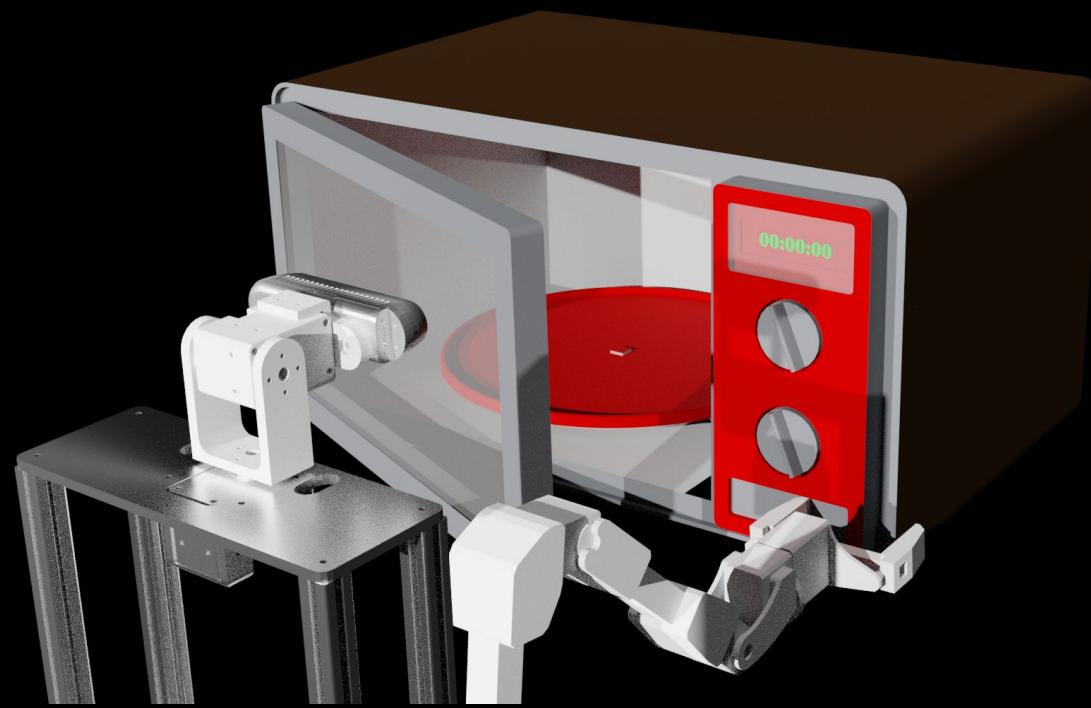
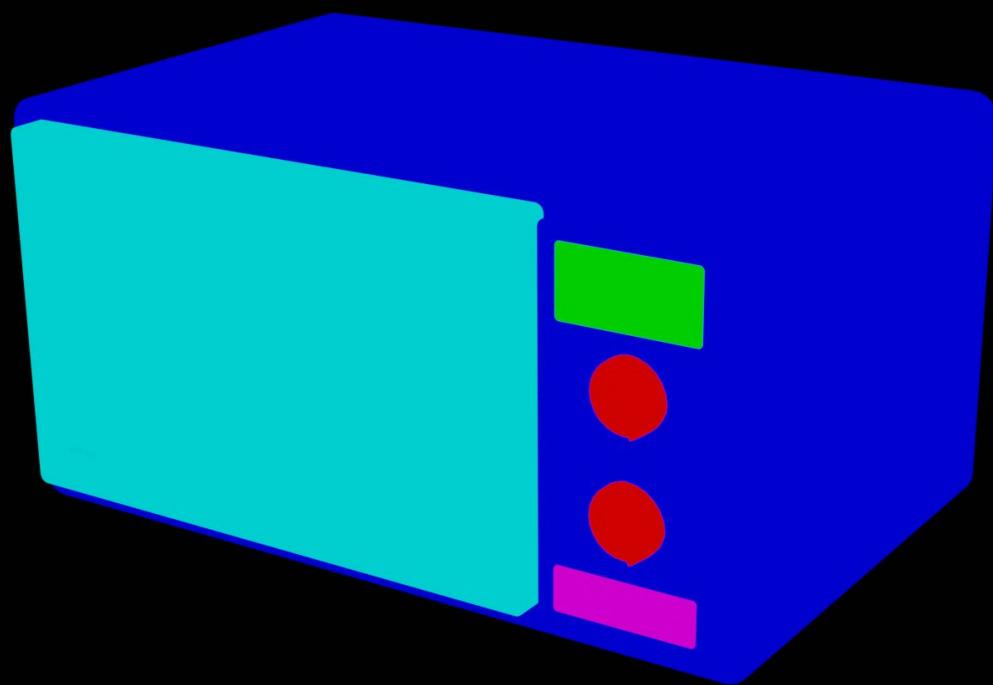
Adapt behavior

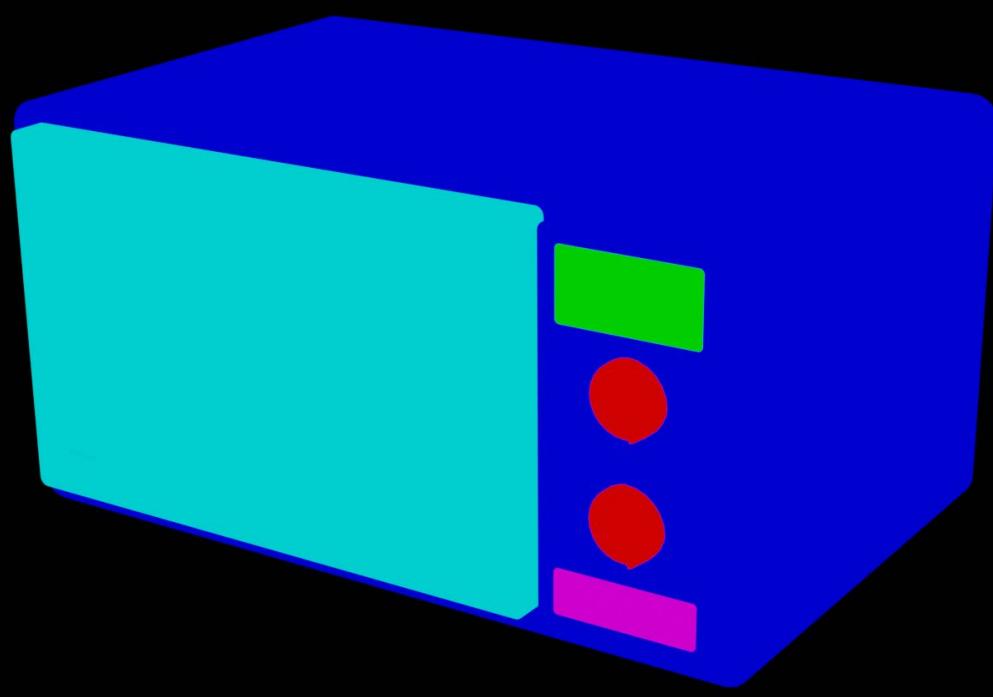
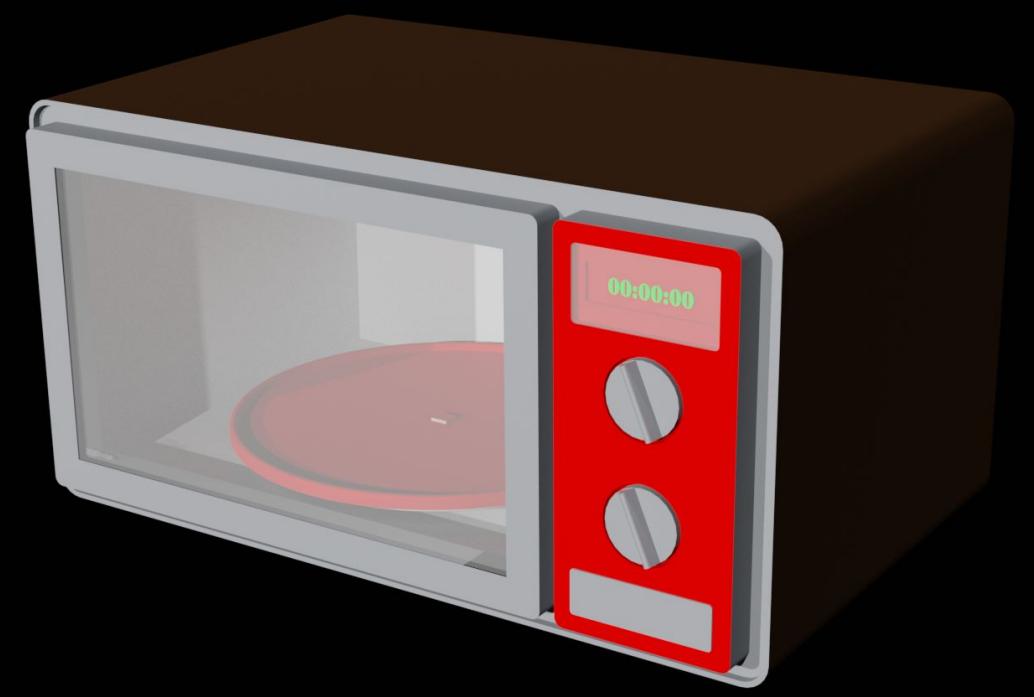
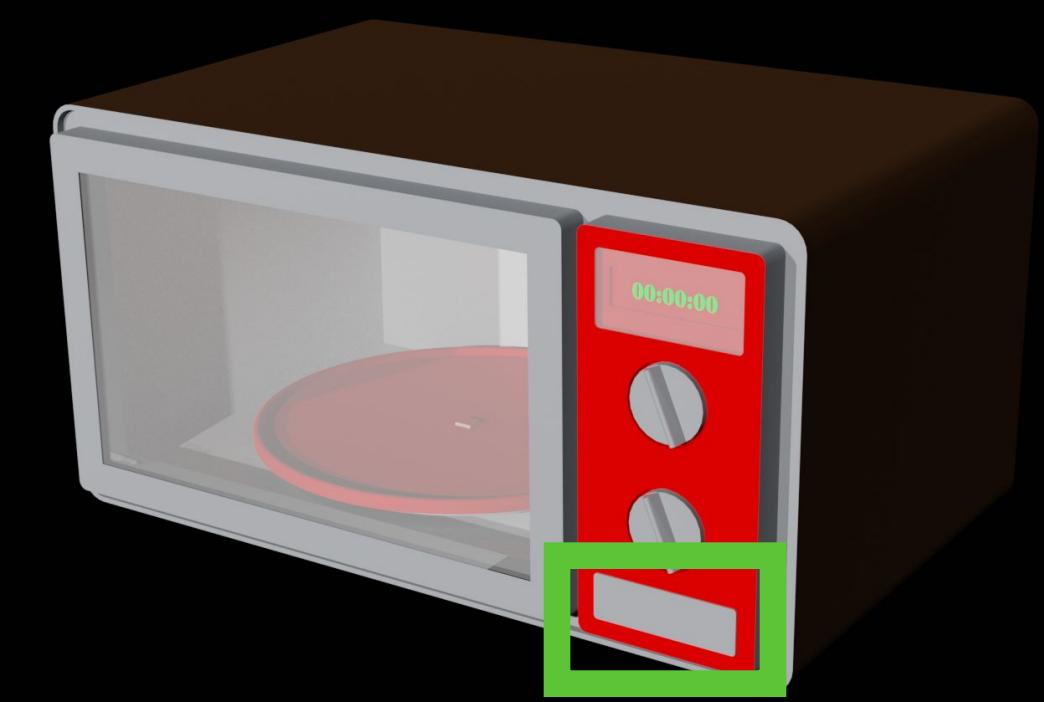
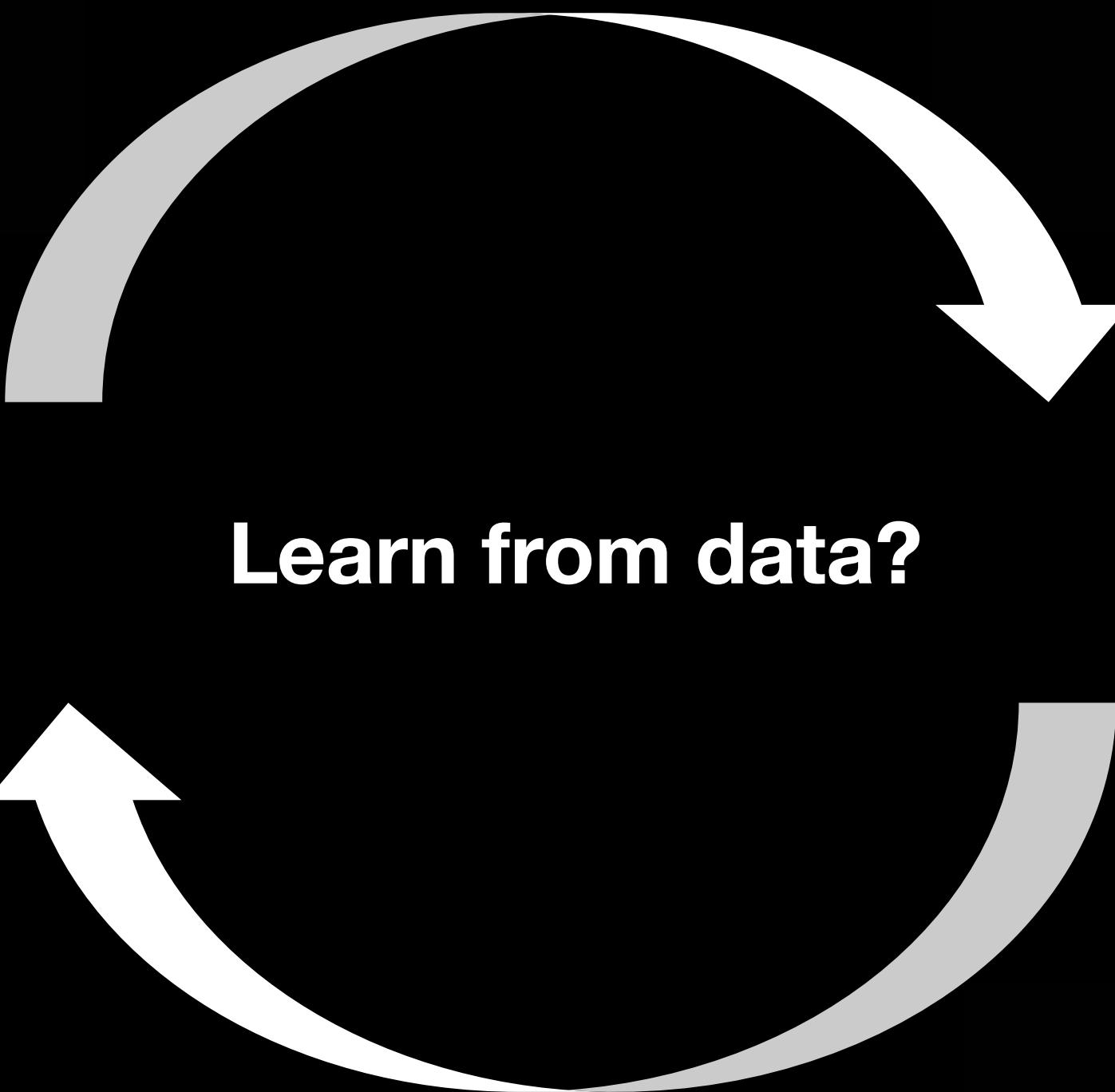
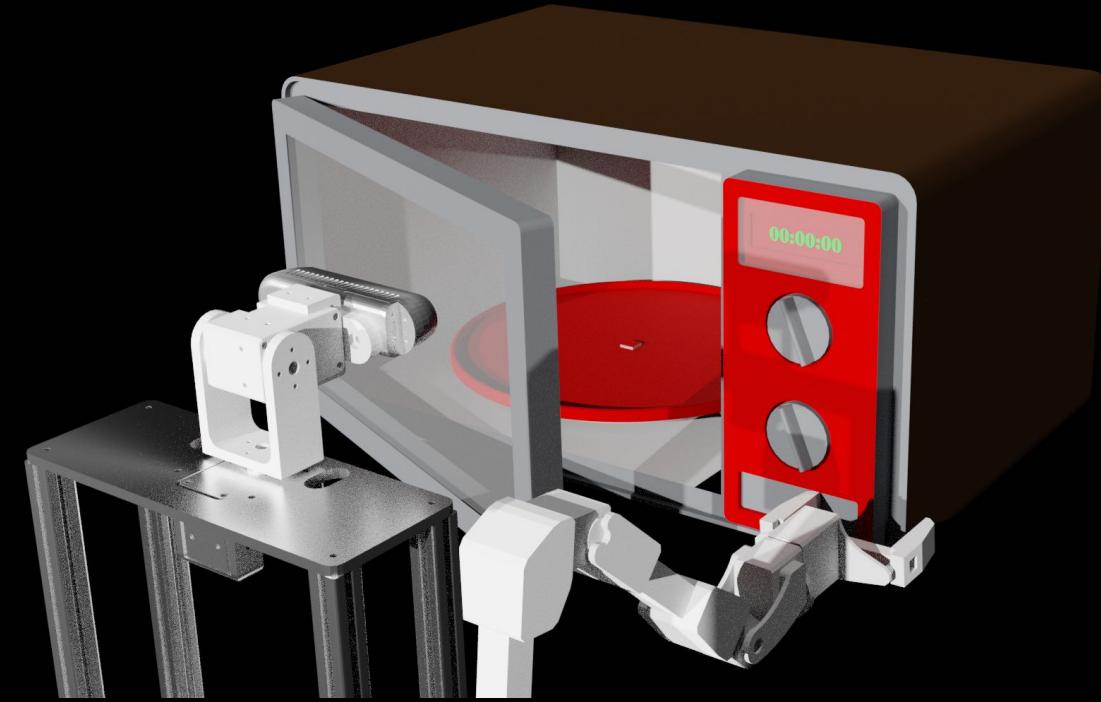


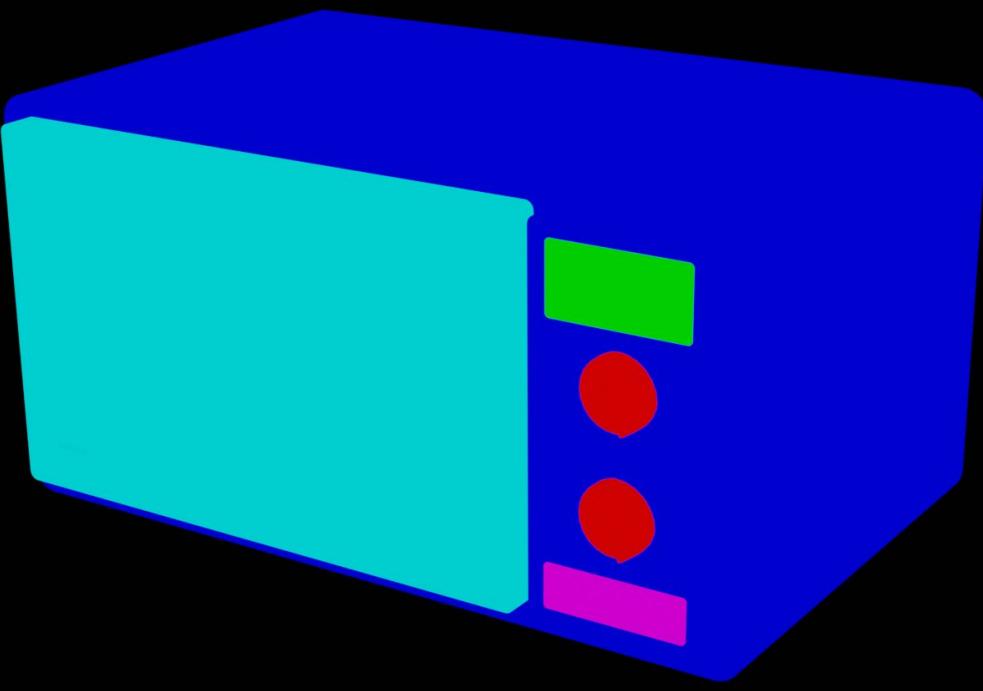
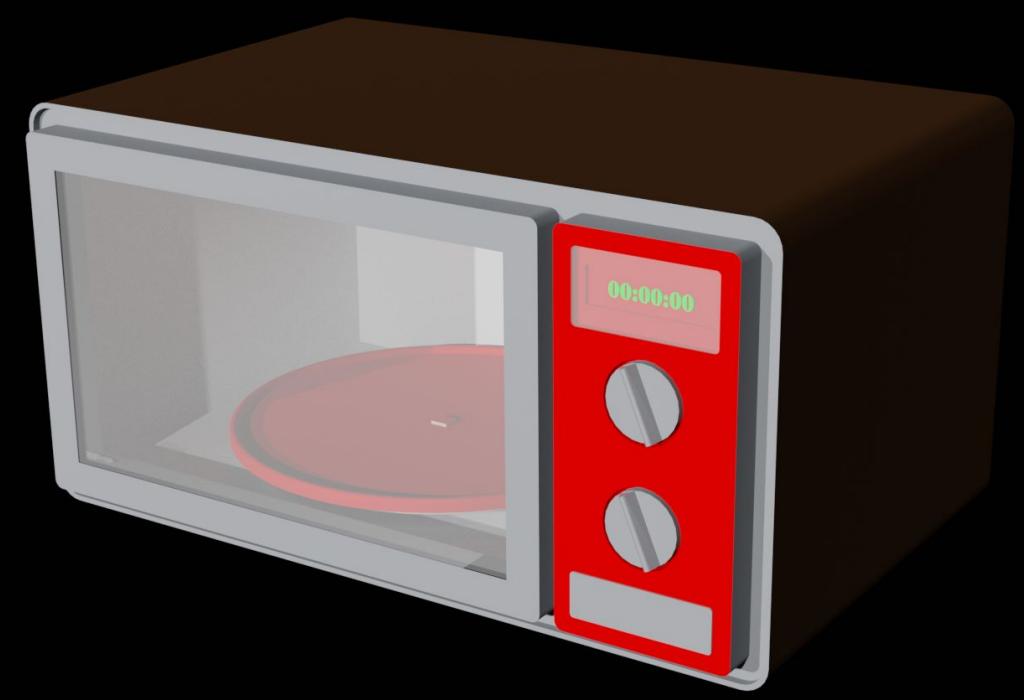
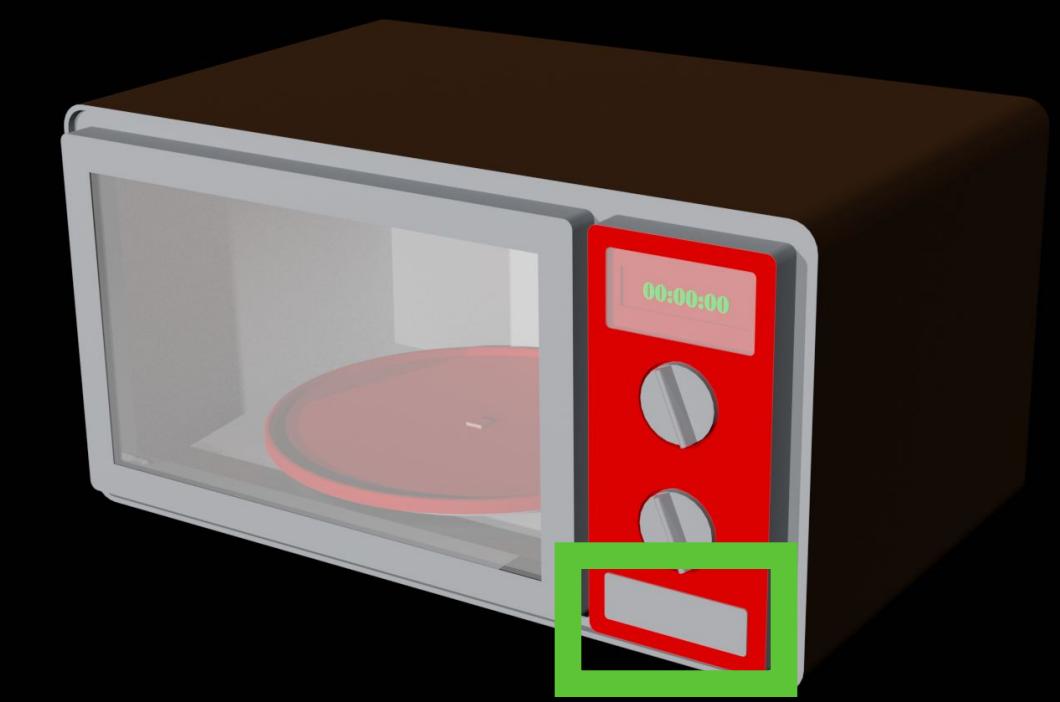
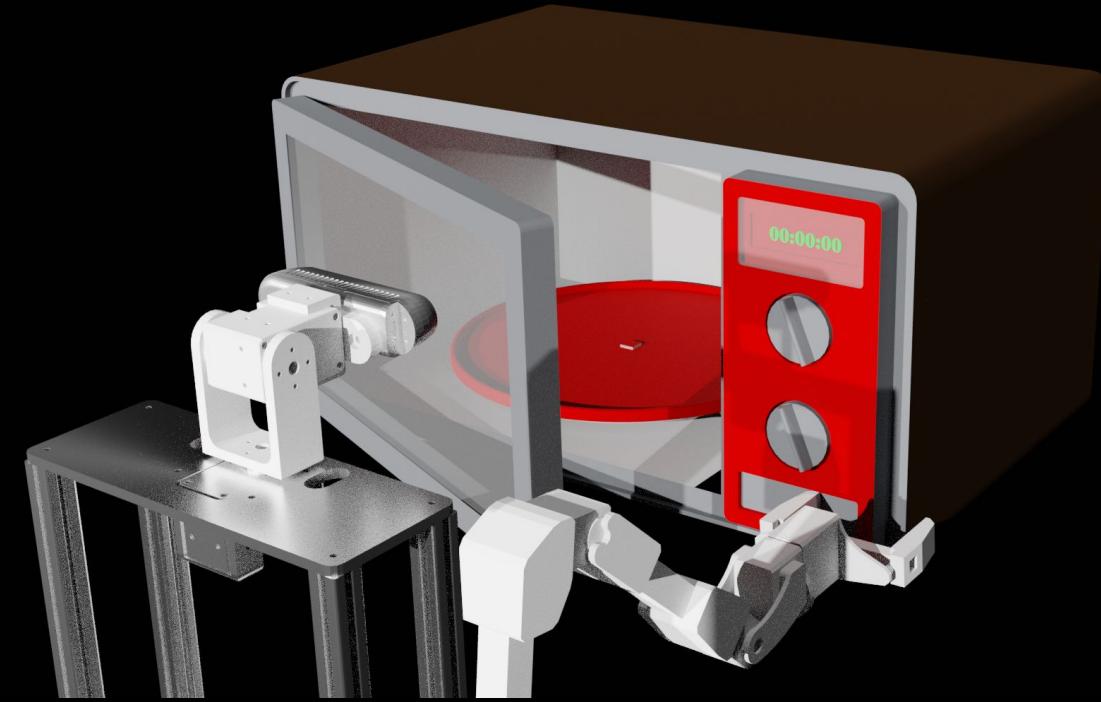


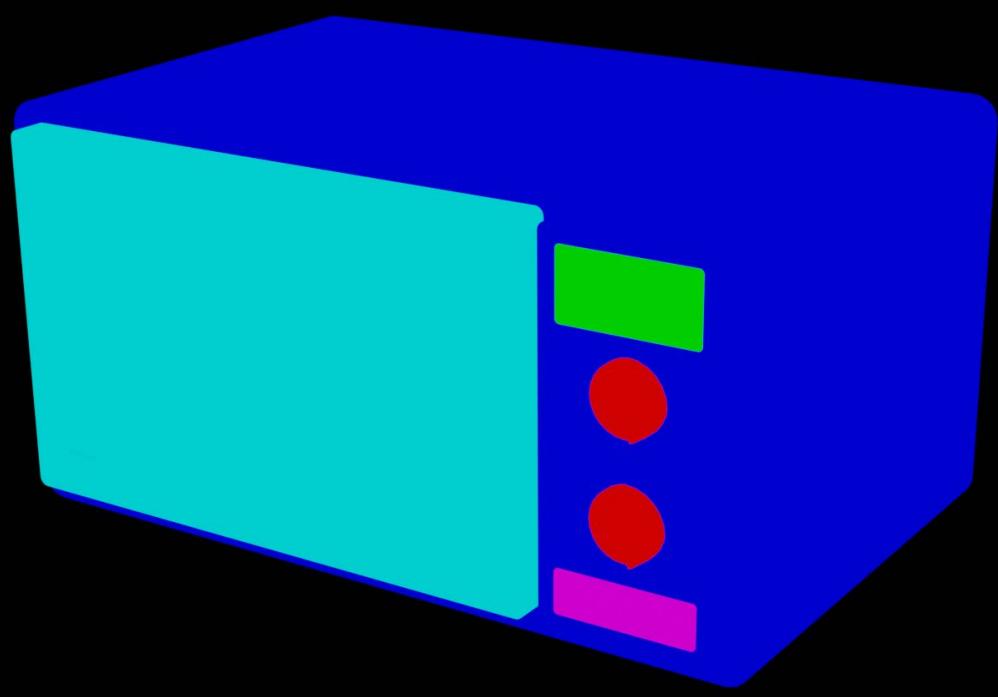
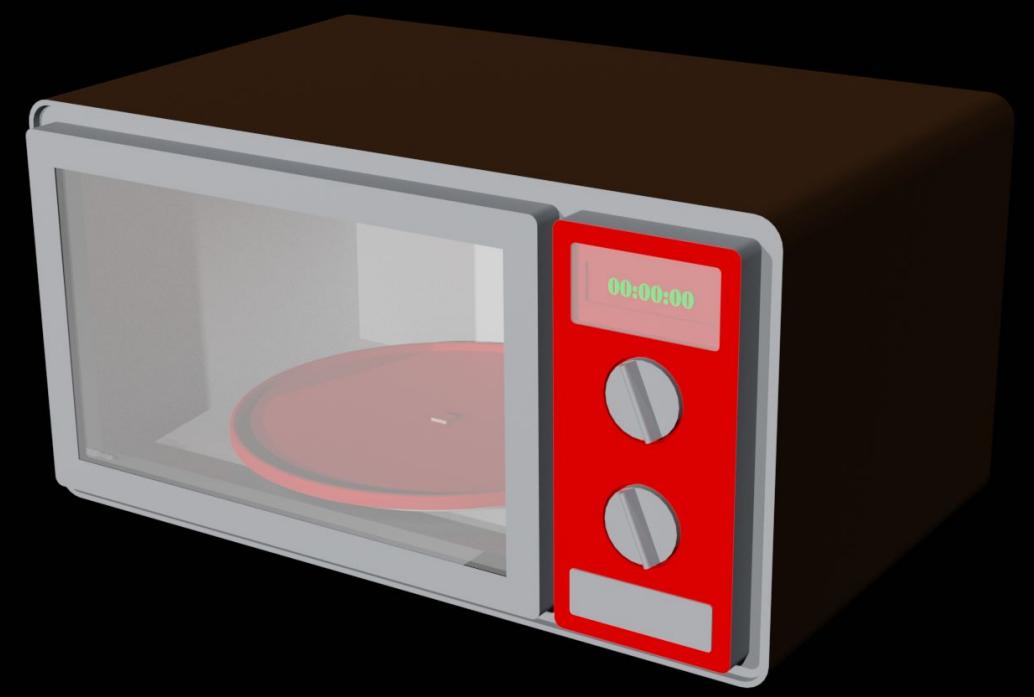
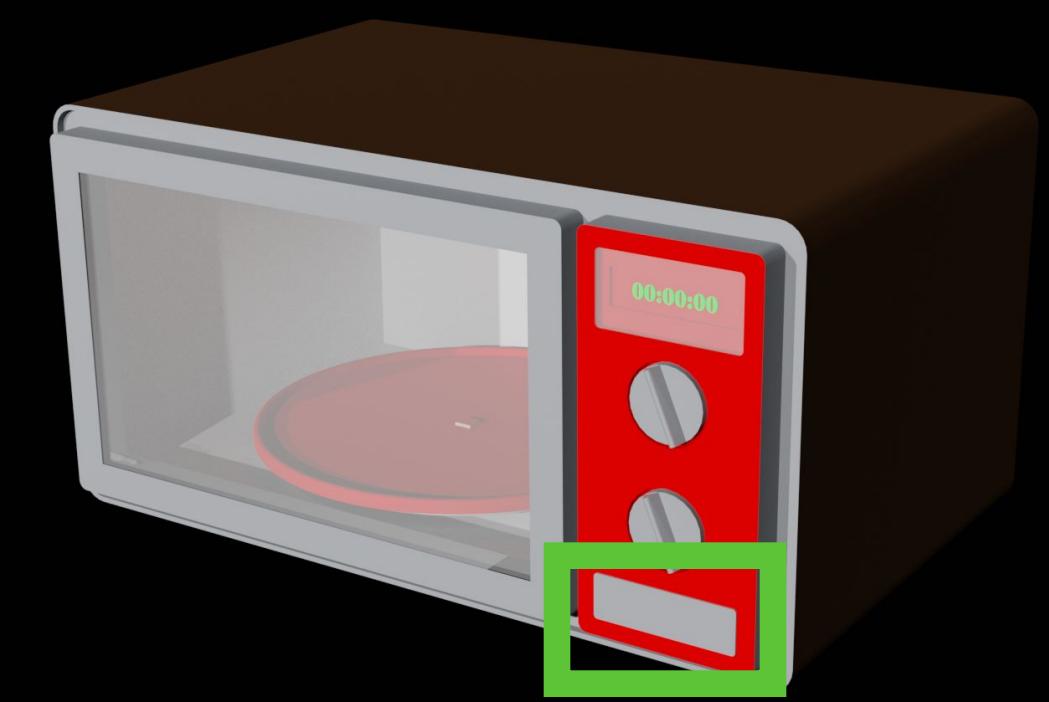
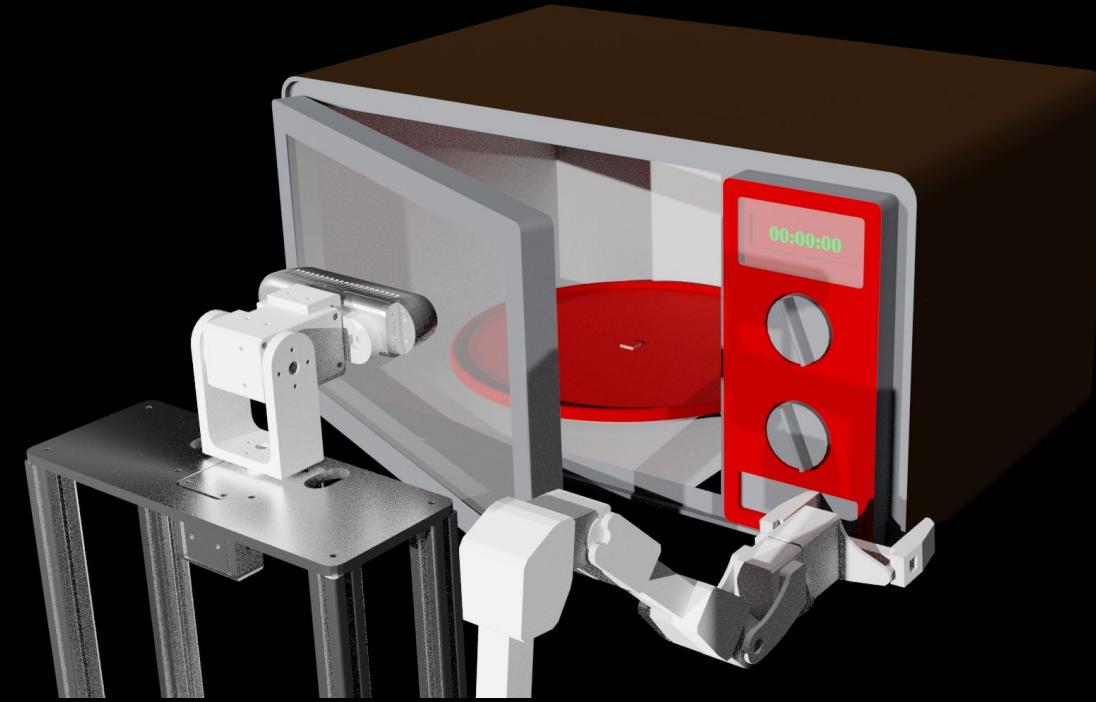


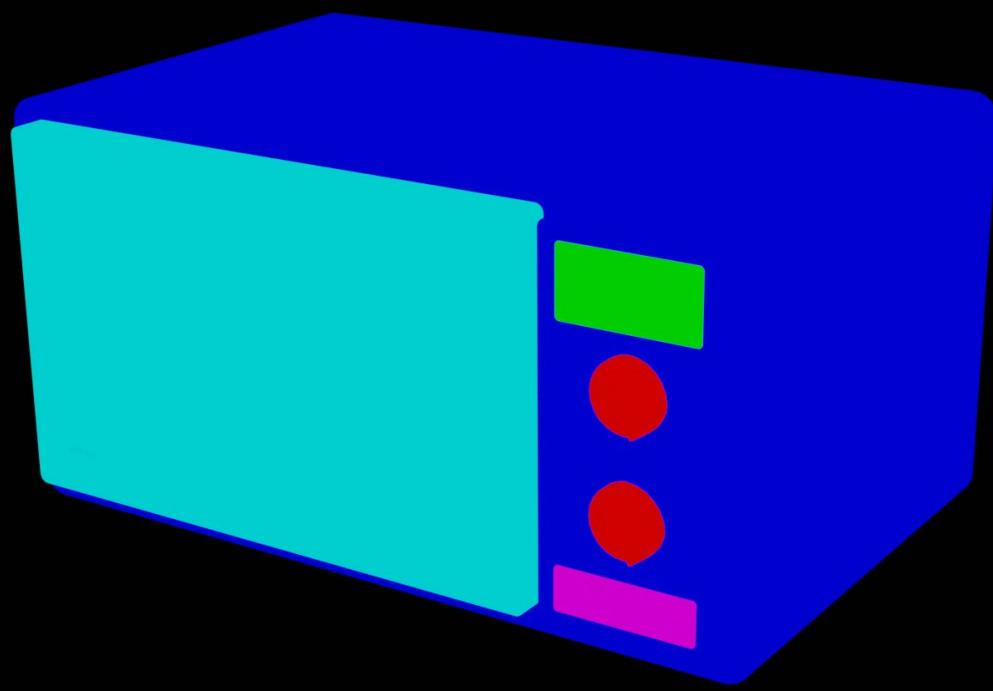
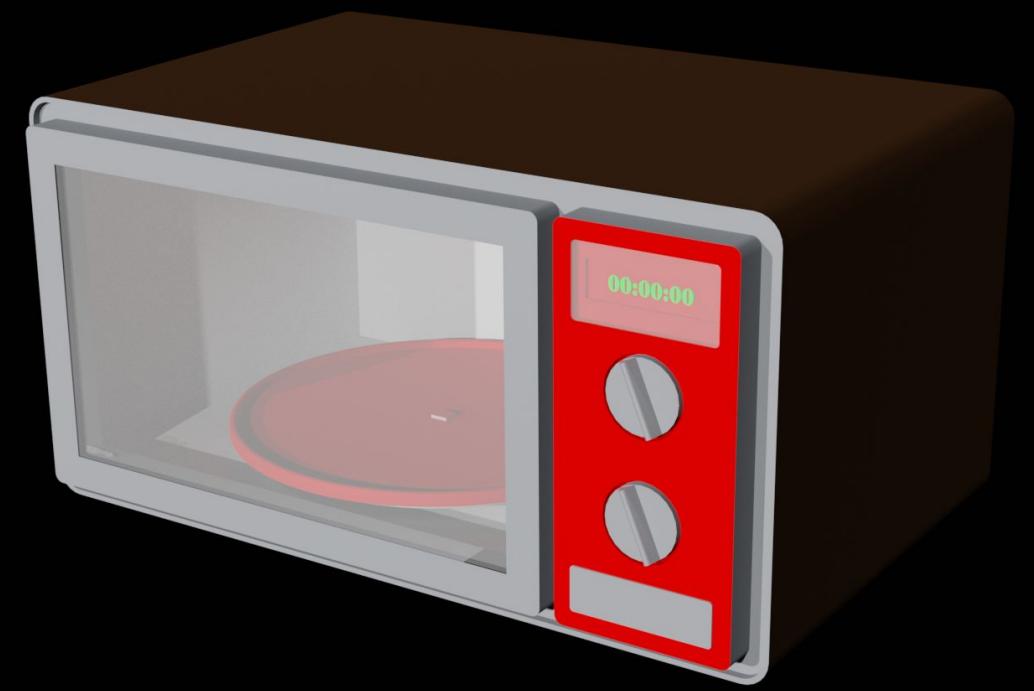
Data driven approaches



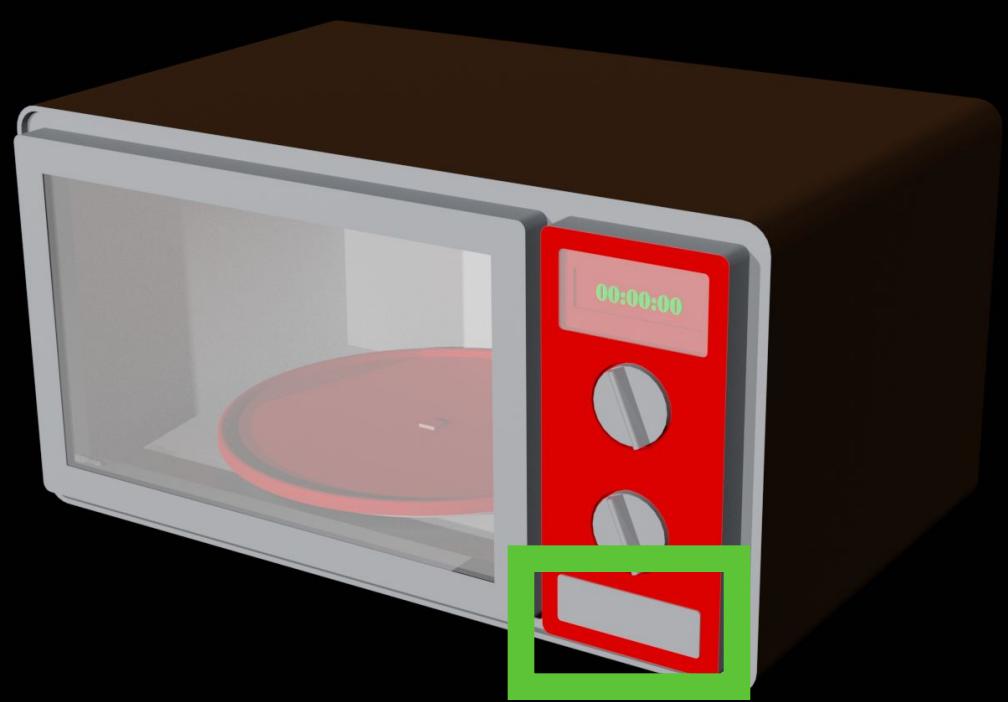
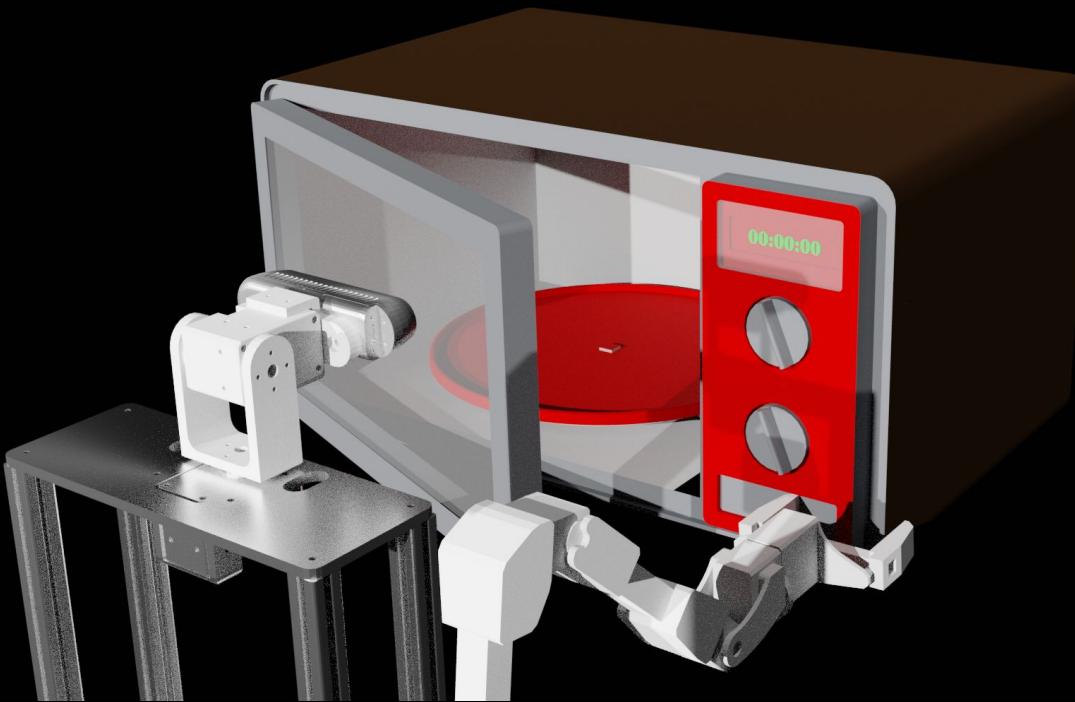


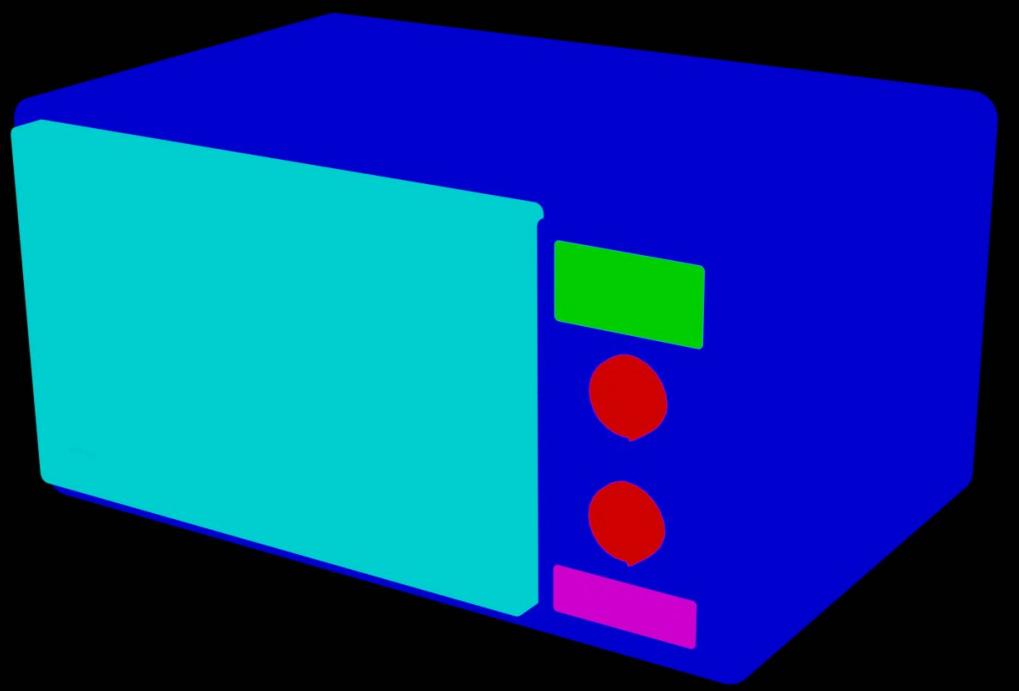
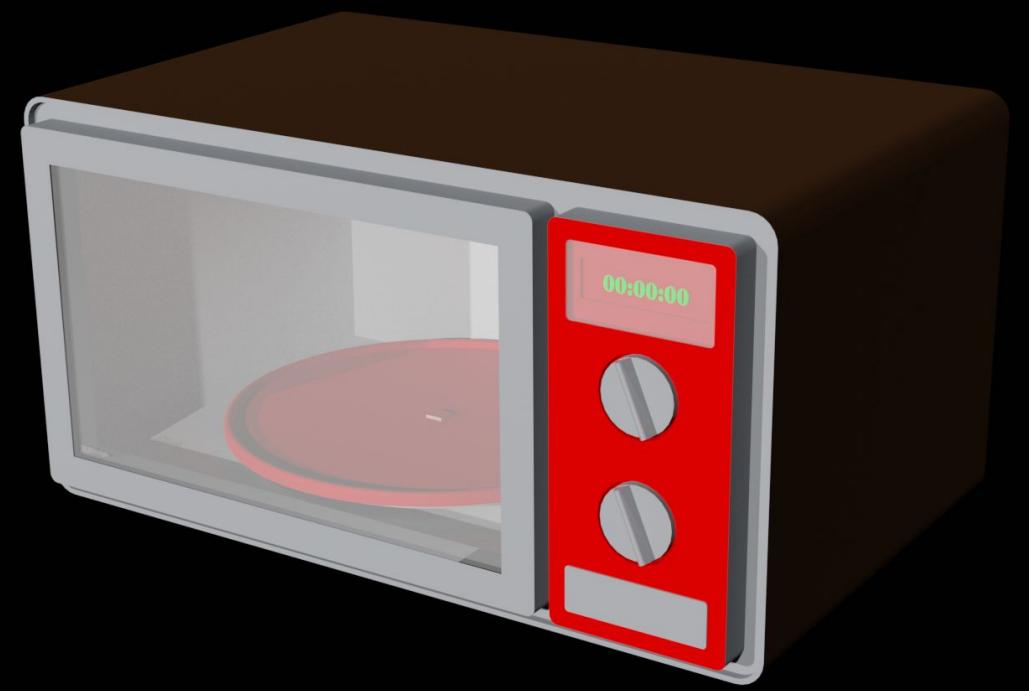




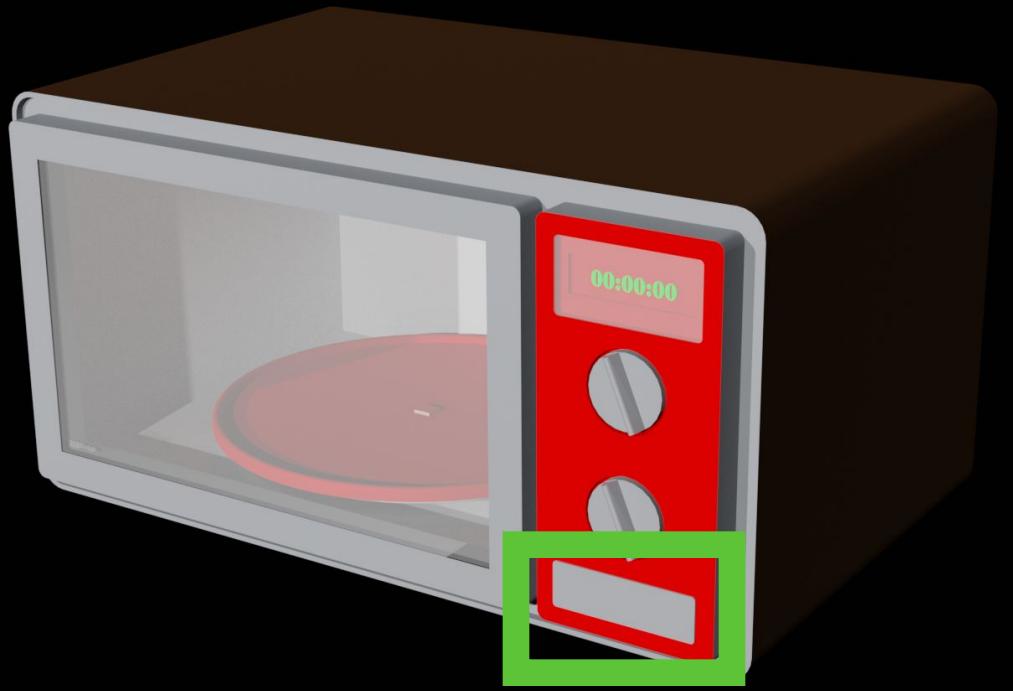
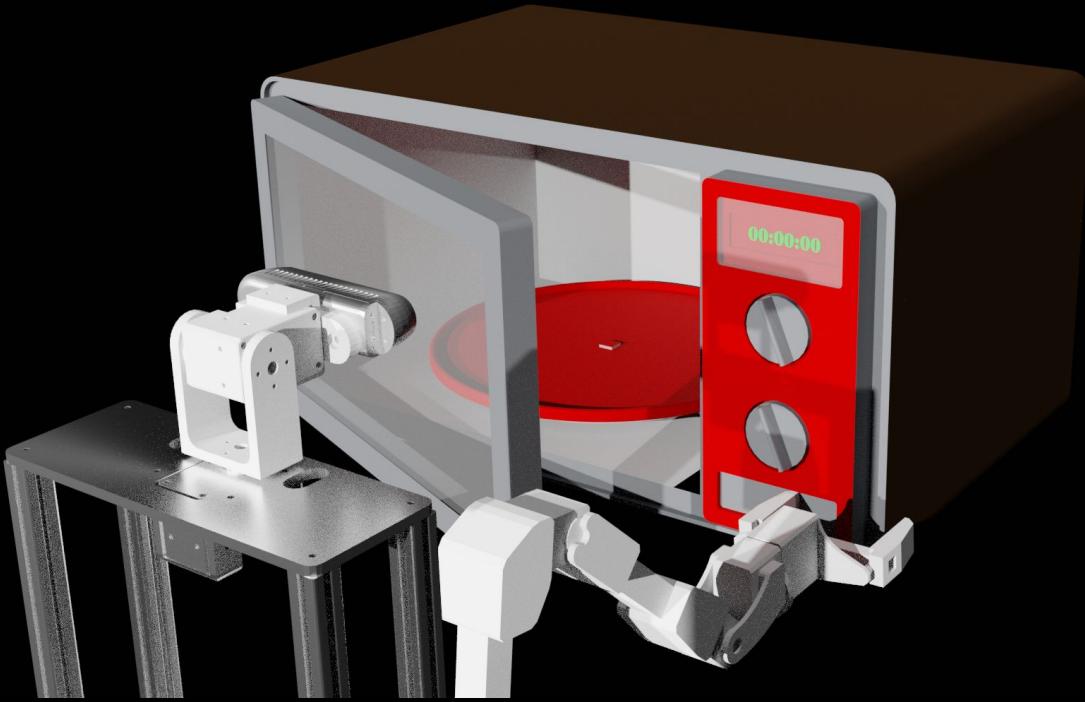


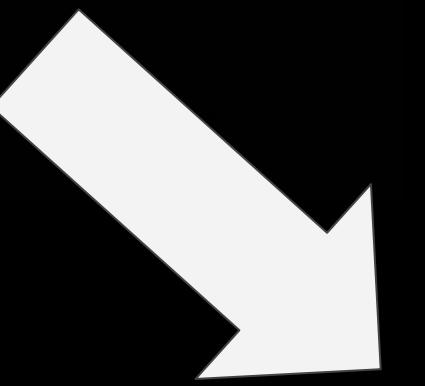
Reinforcement Learning





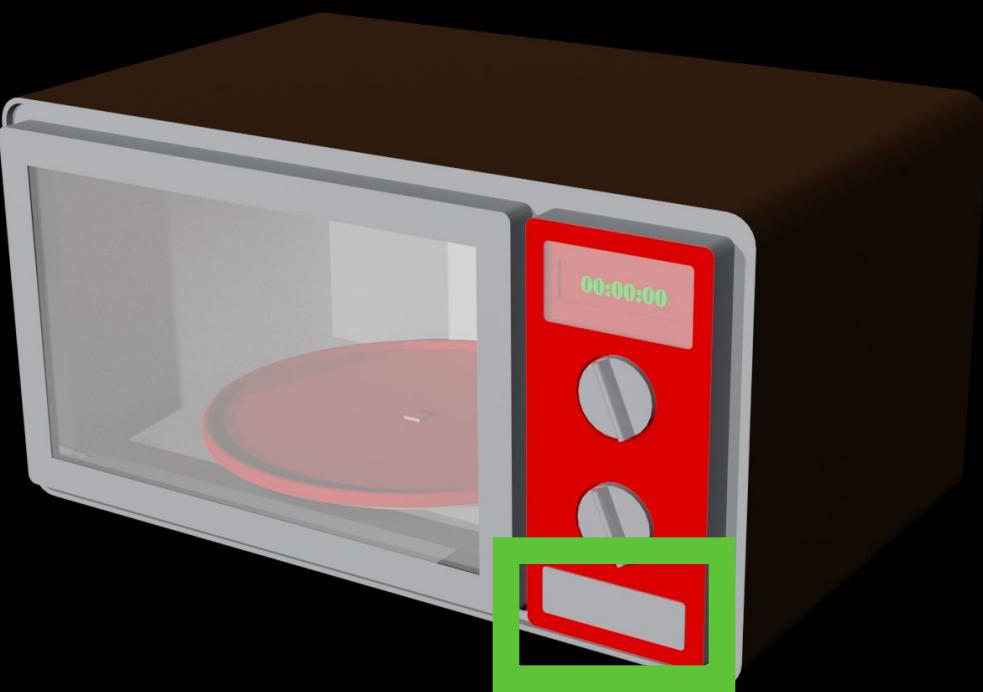
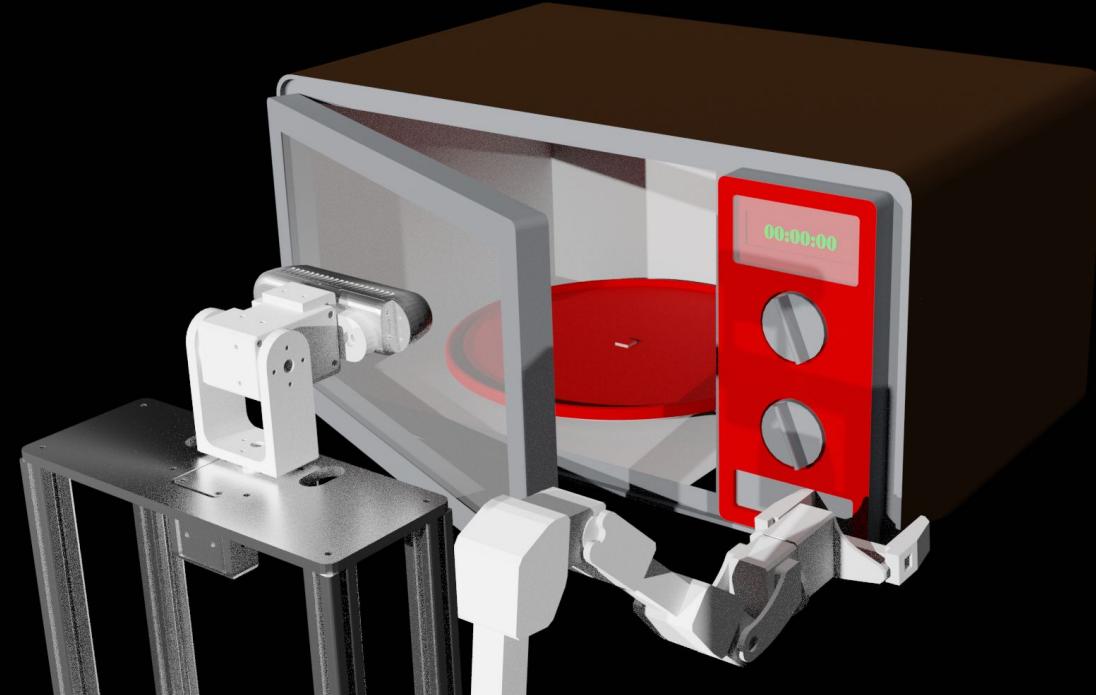
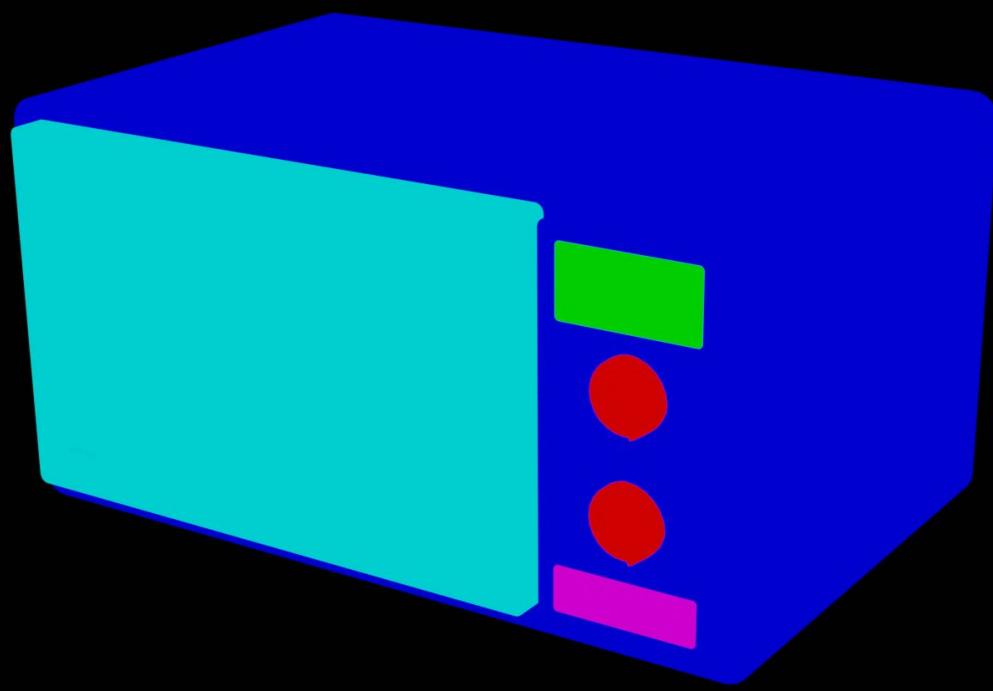
RL Agent

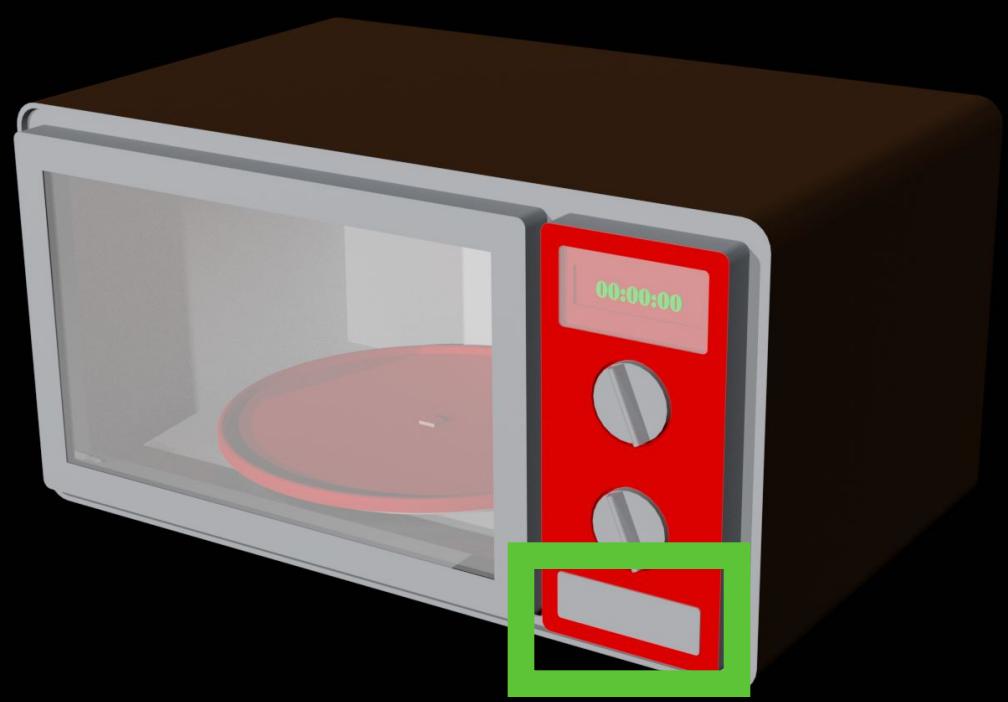
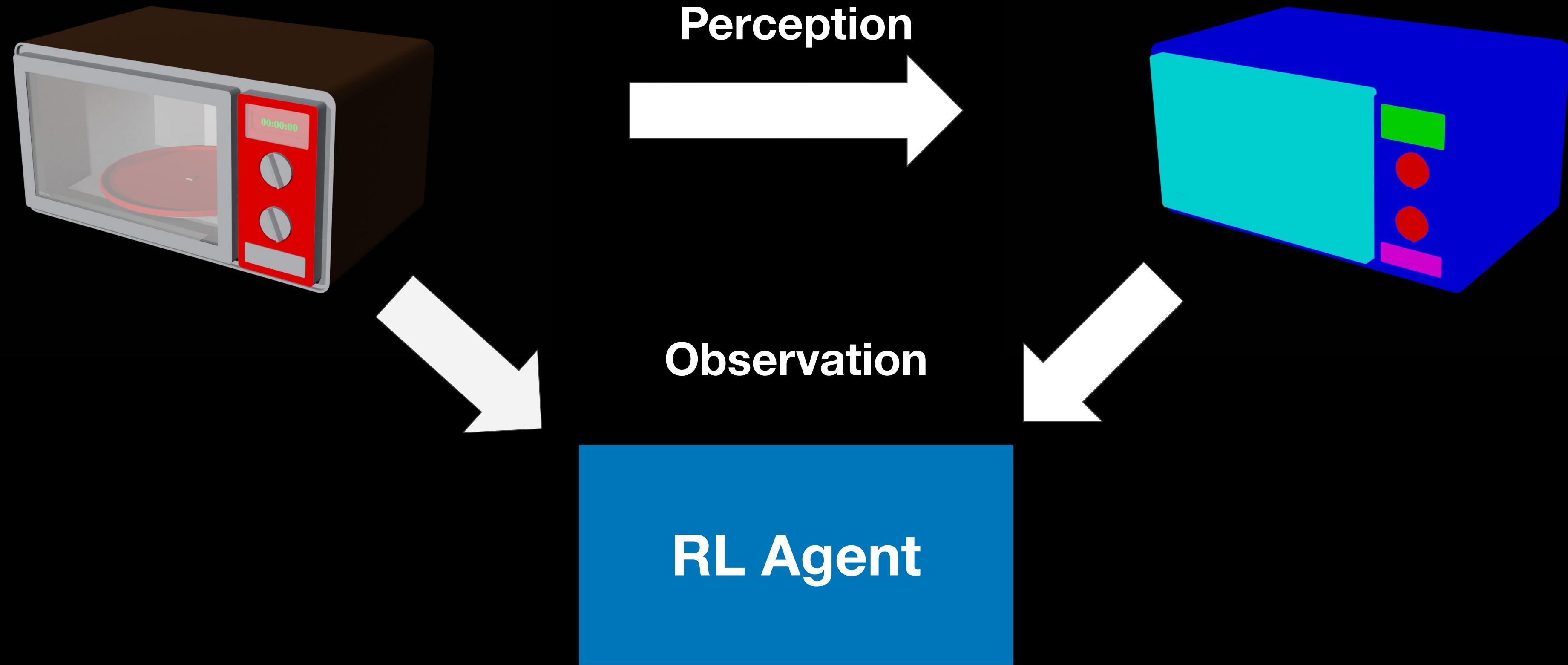
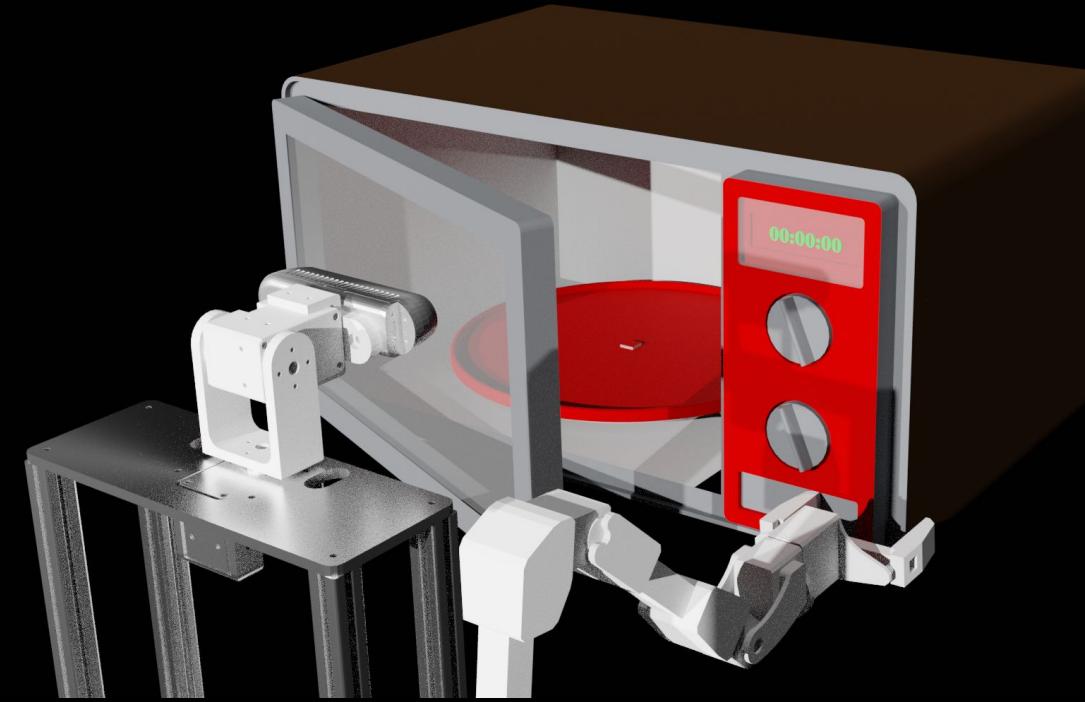


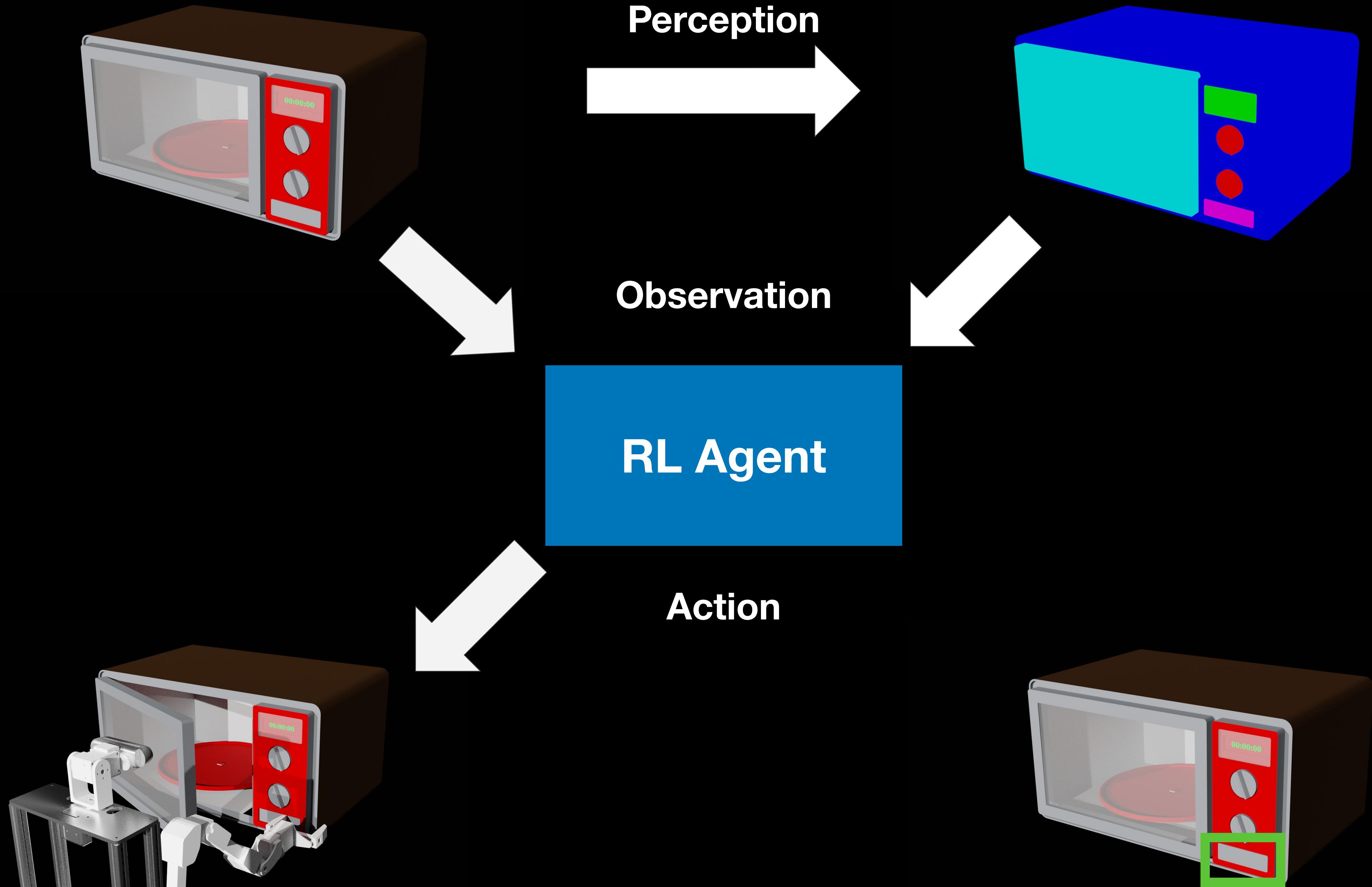


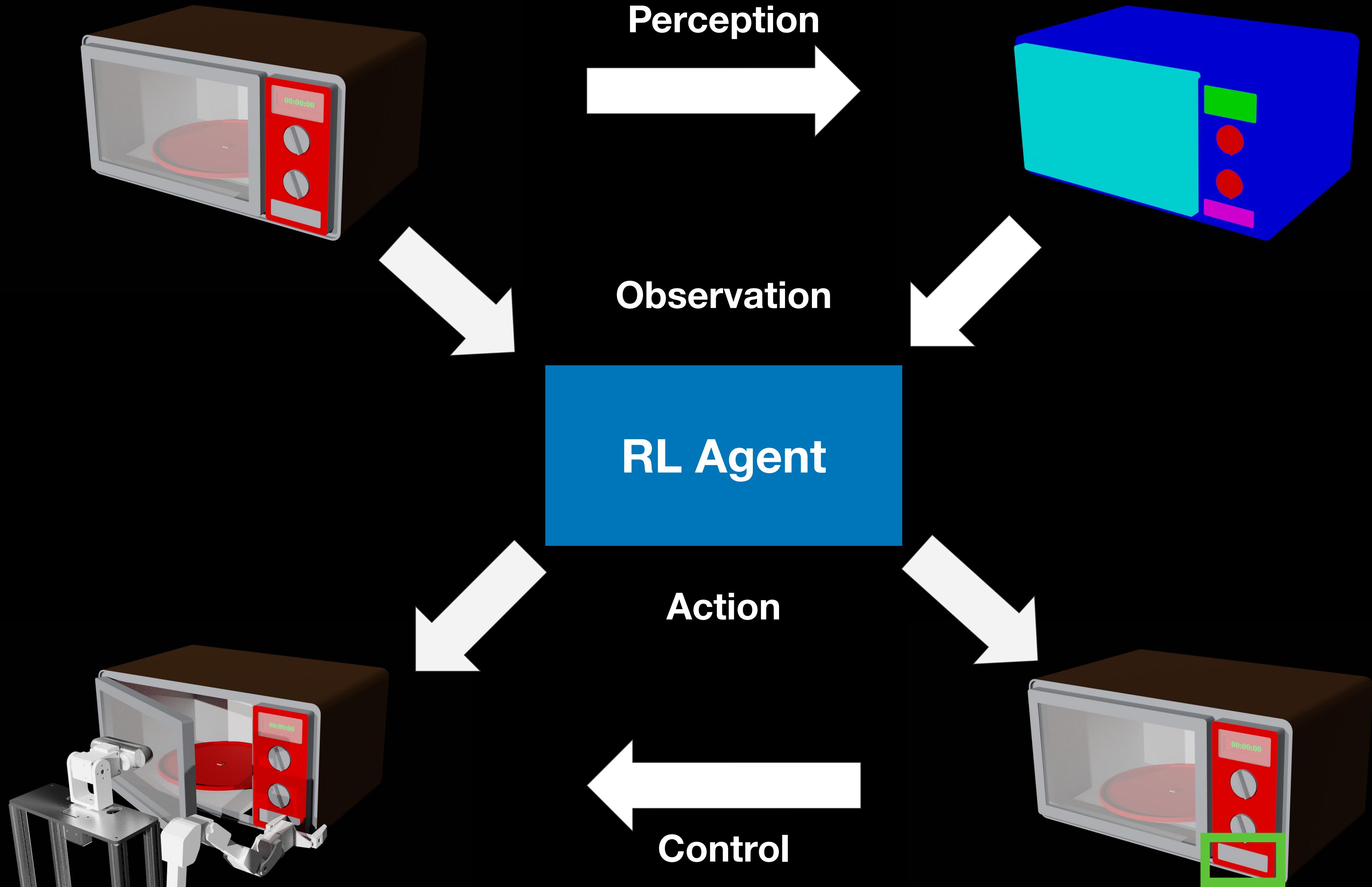
Observation

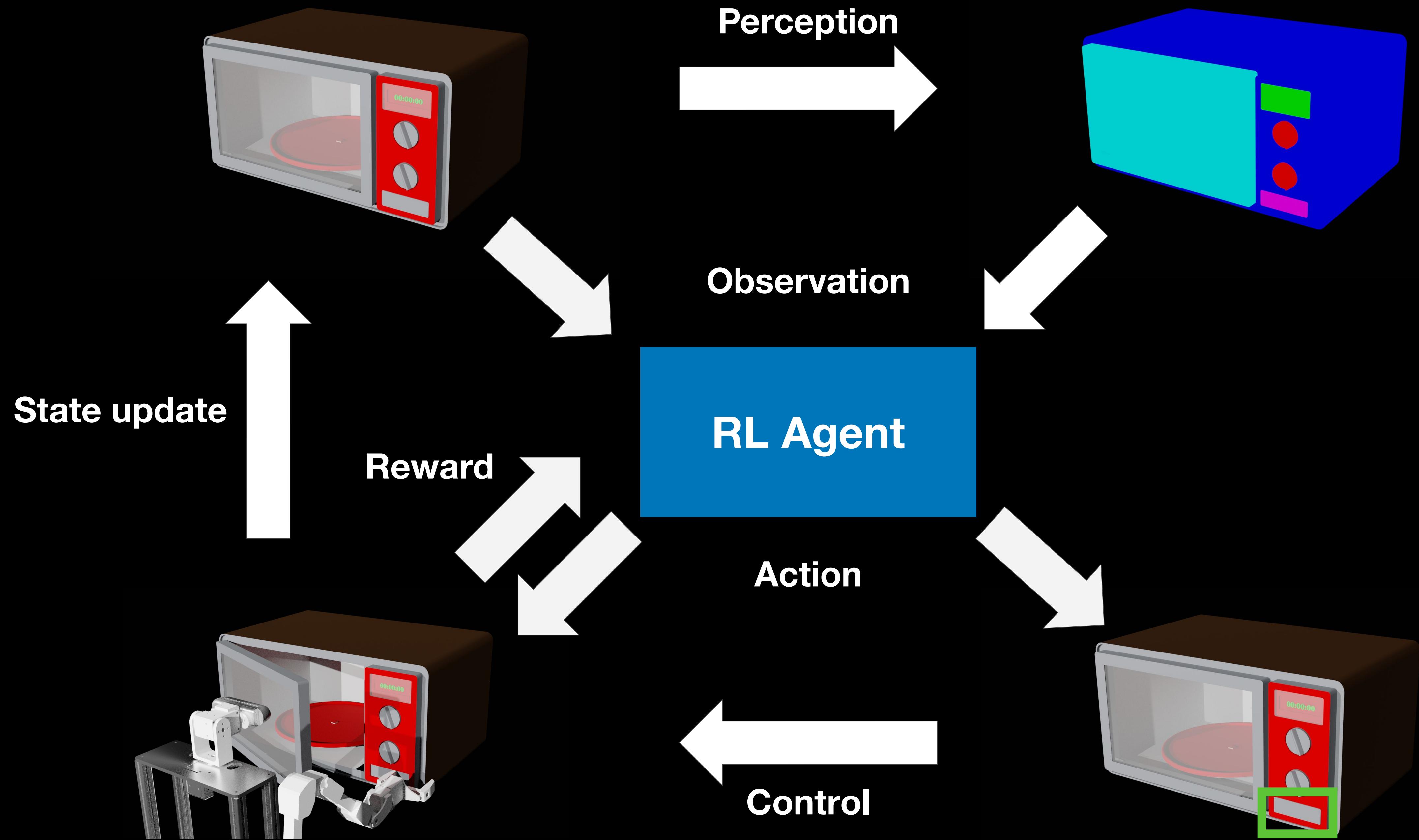
RL Agent

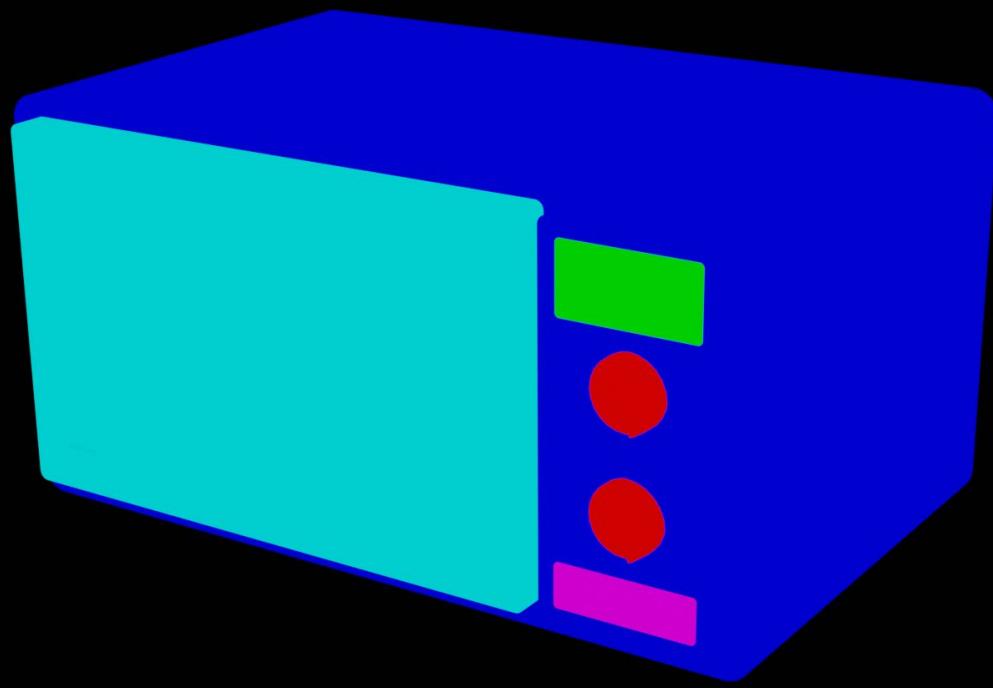
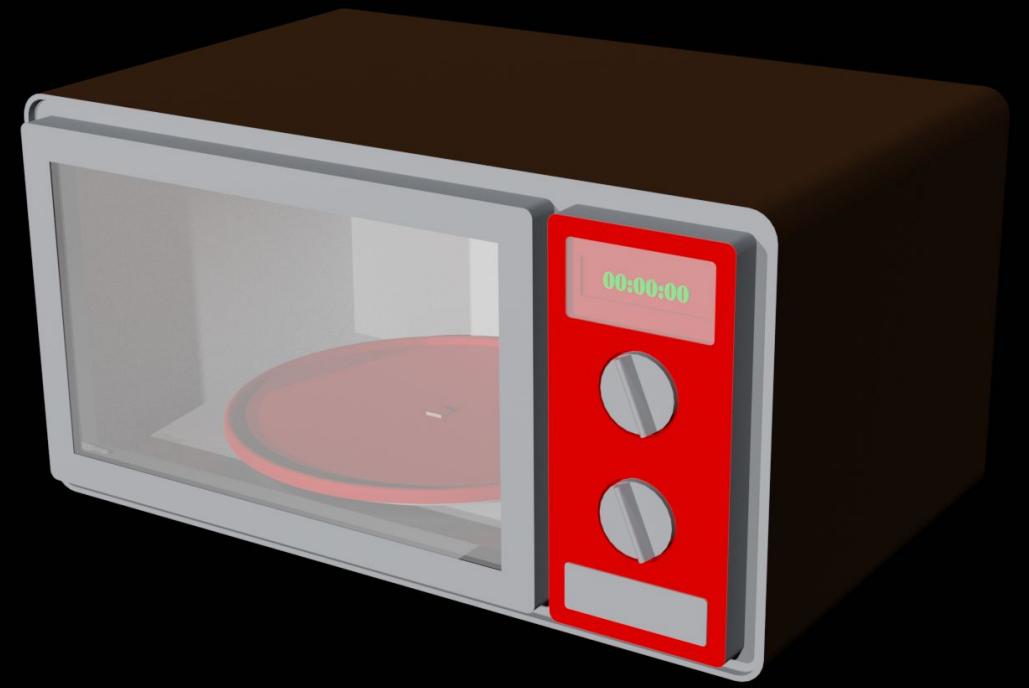






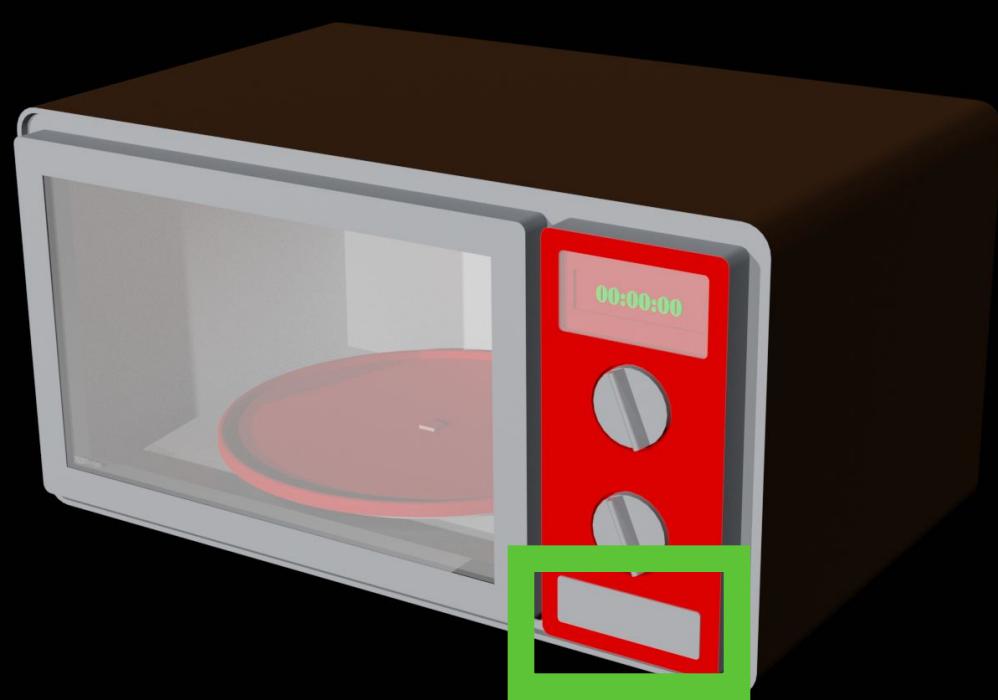
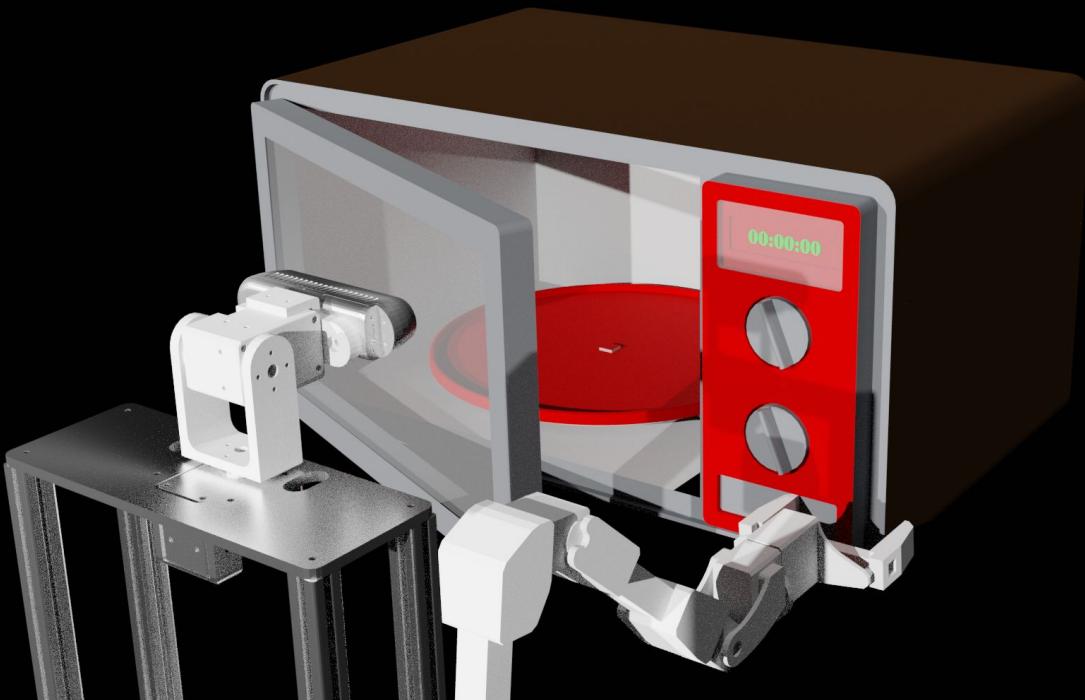


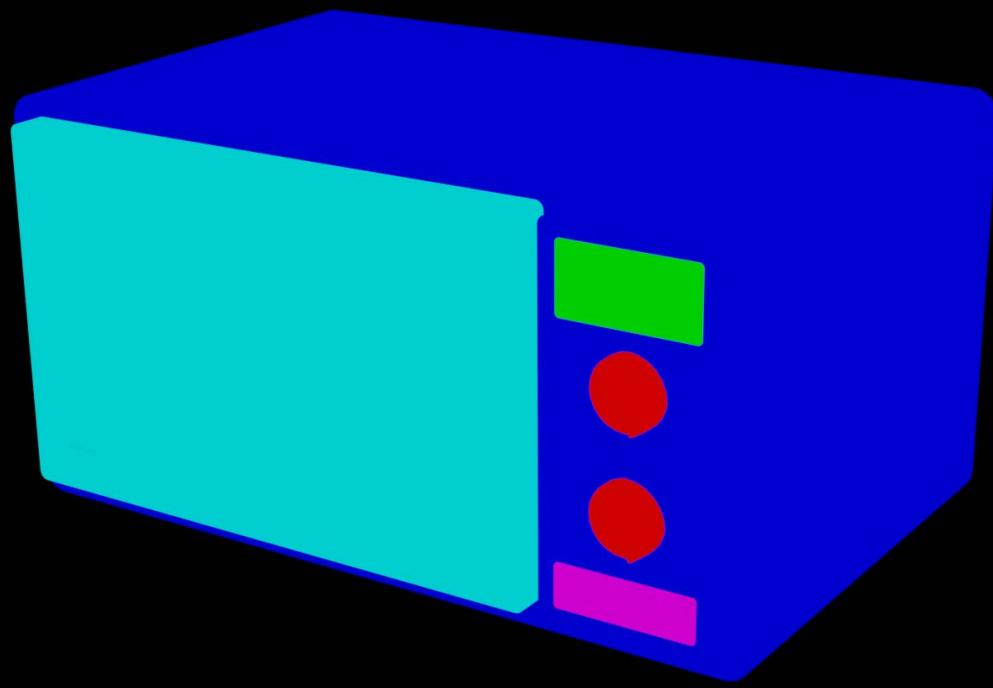




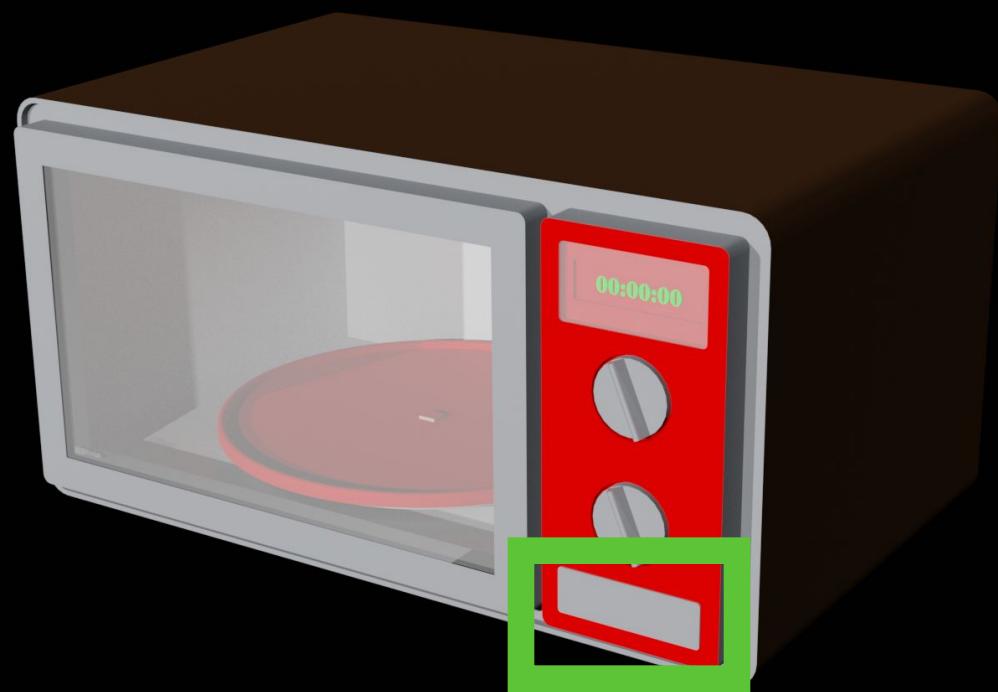
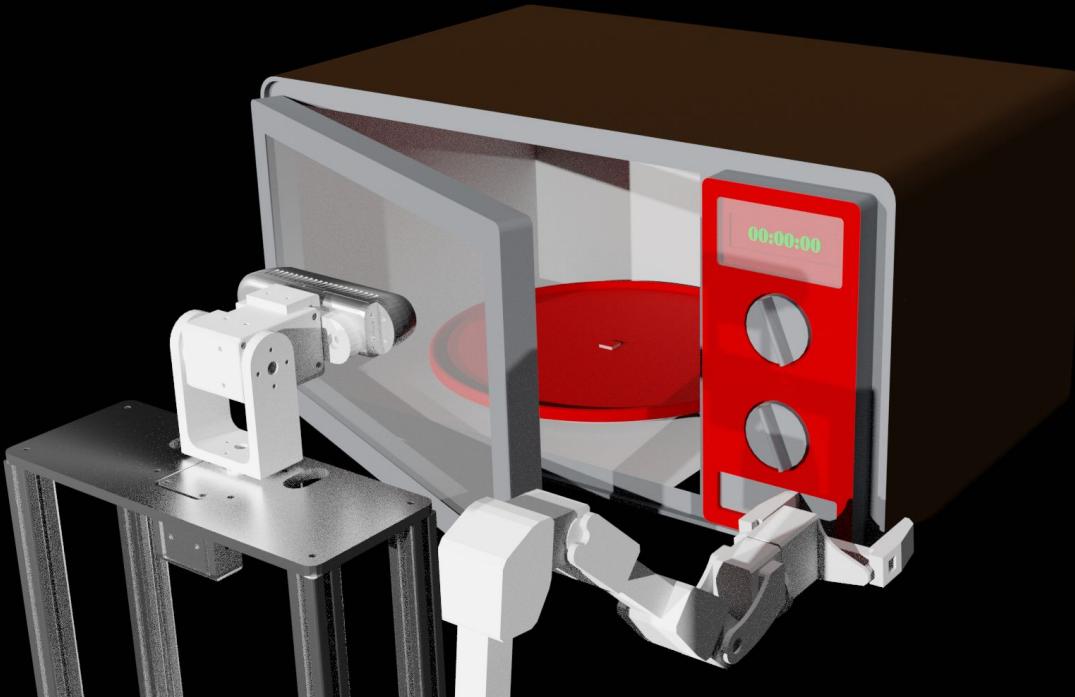
RL Problems

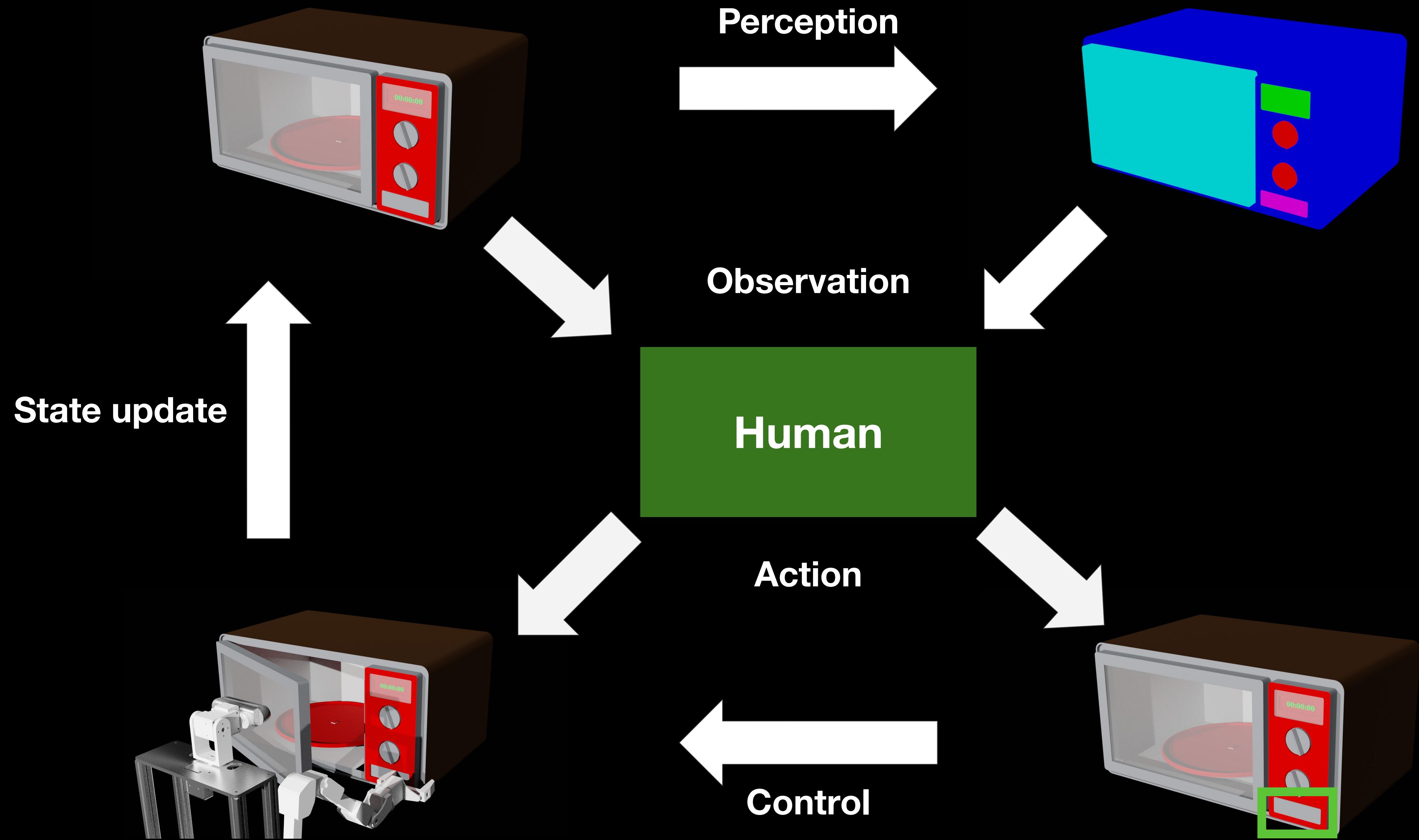
- **Expensive hardware**
- **Slow data collection**
- **Low sample efficiency**
- **Overfit to specific agent**

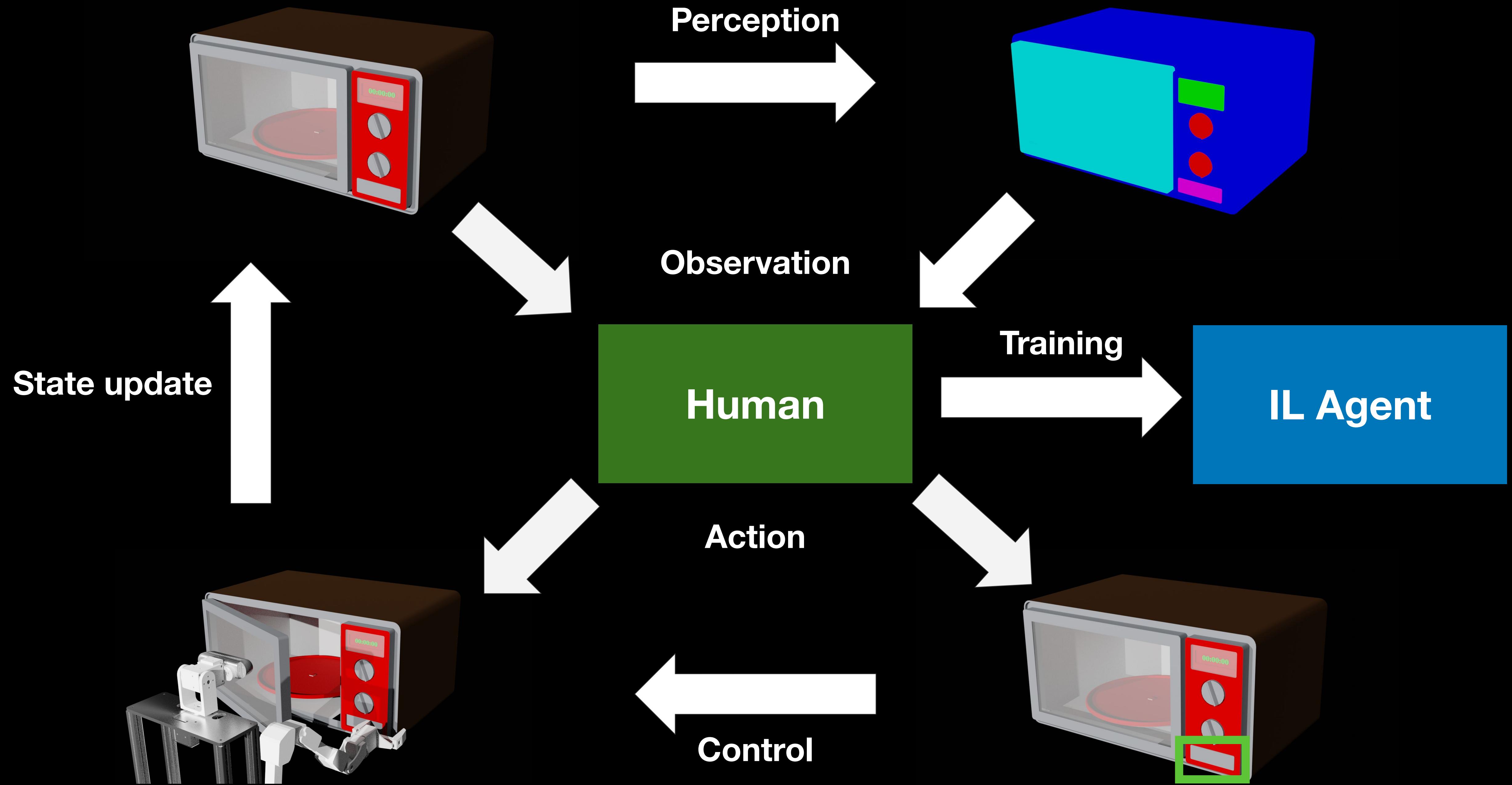


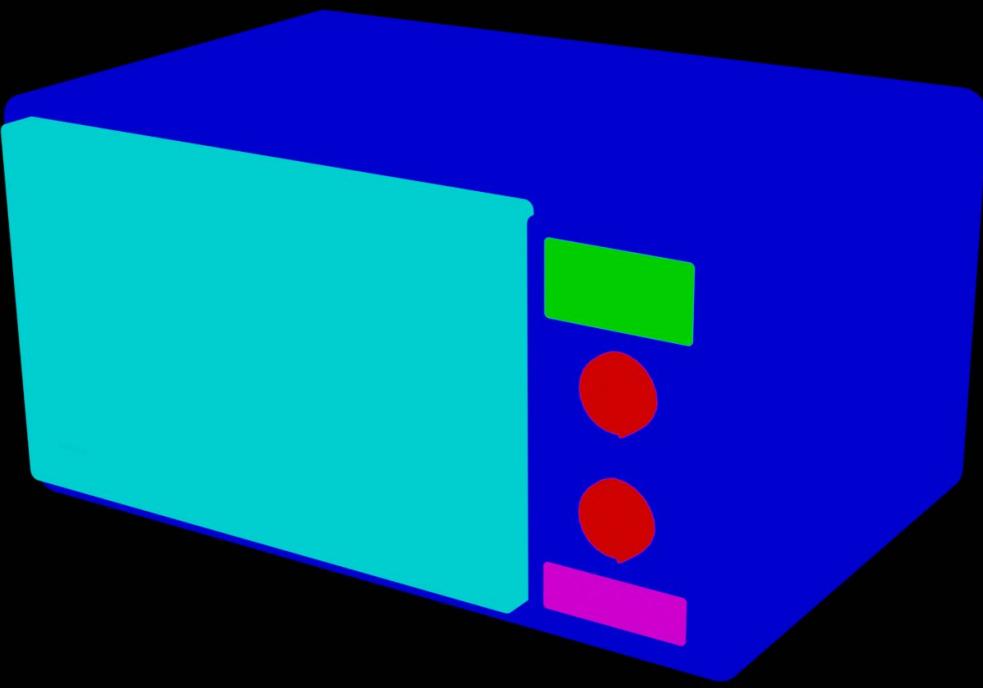
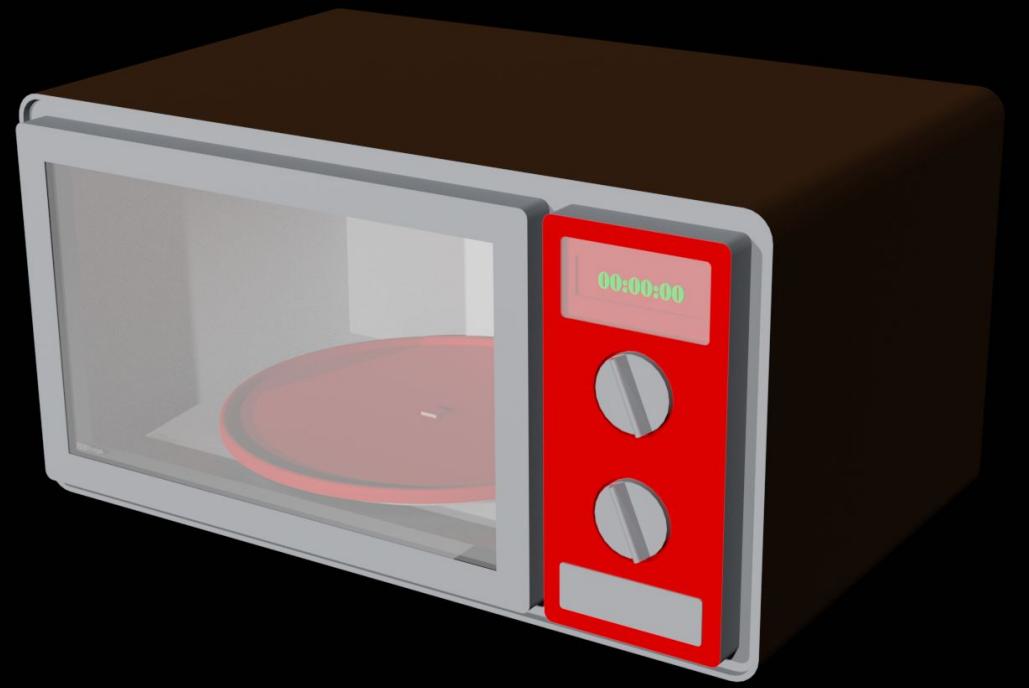


Imitation Learning (Manual Data Collection)



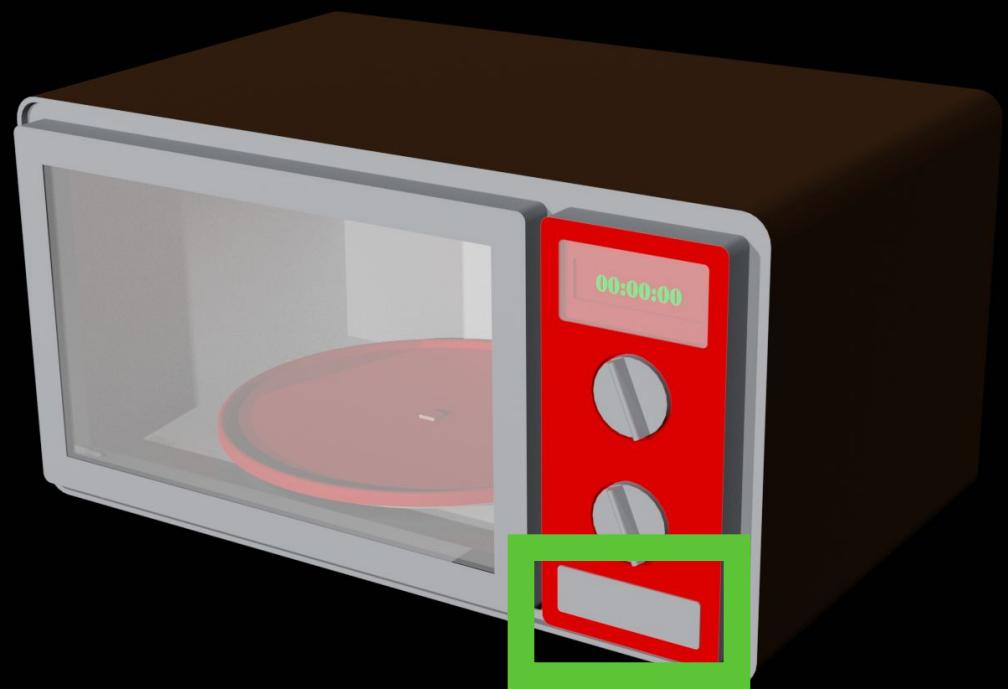
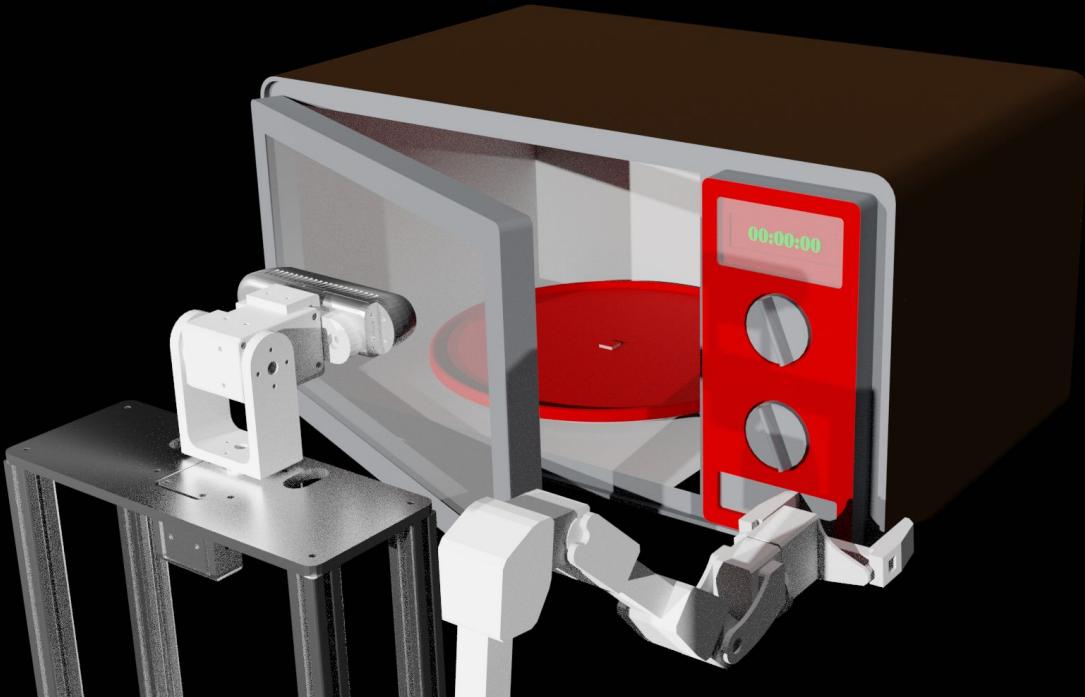


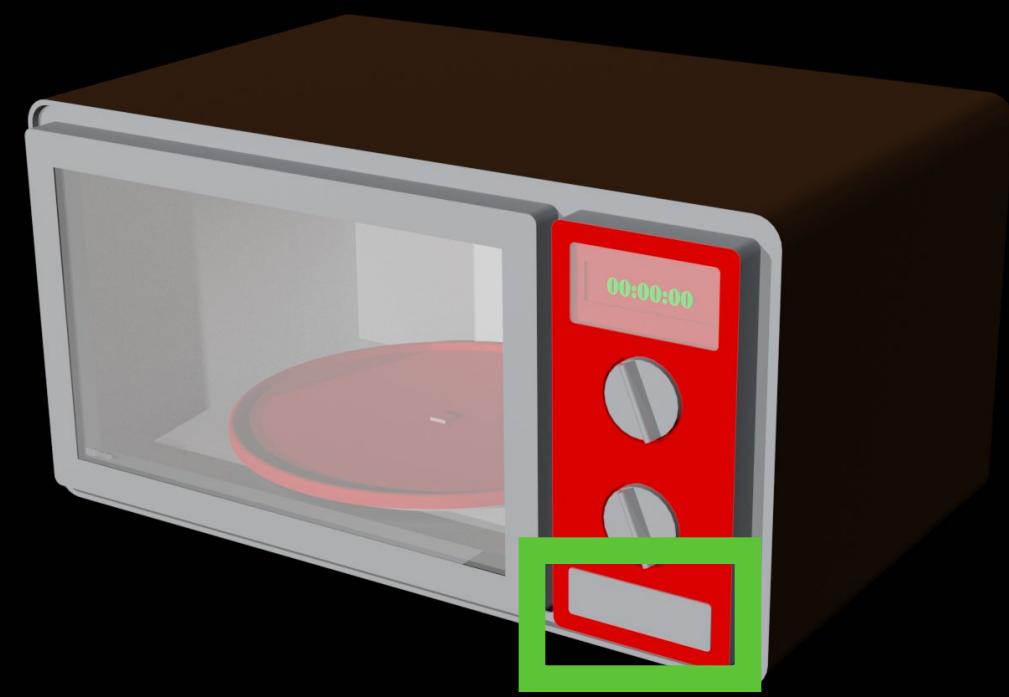
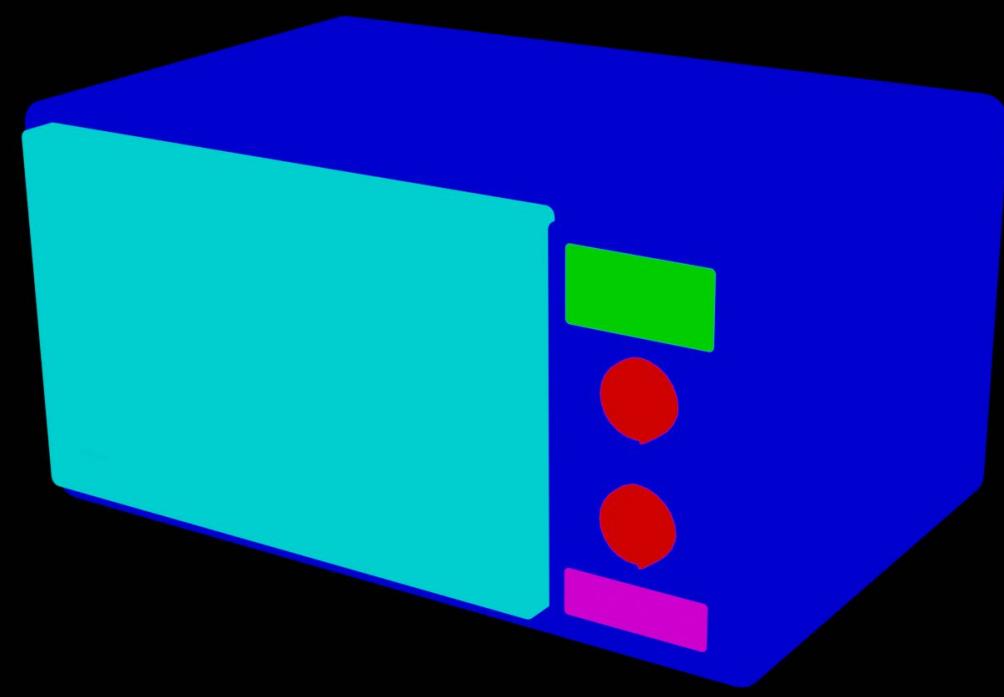
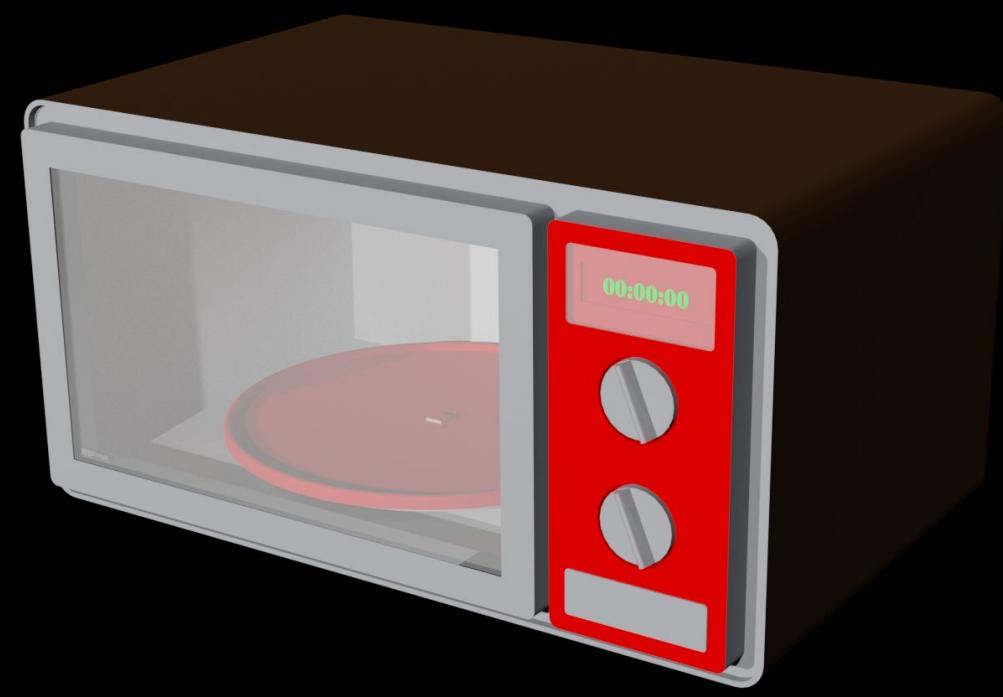
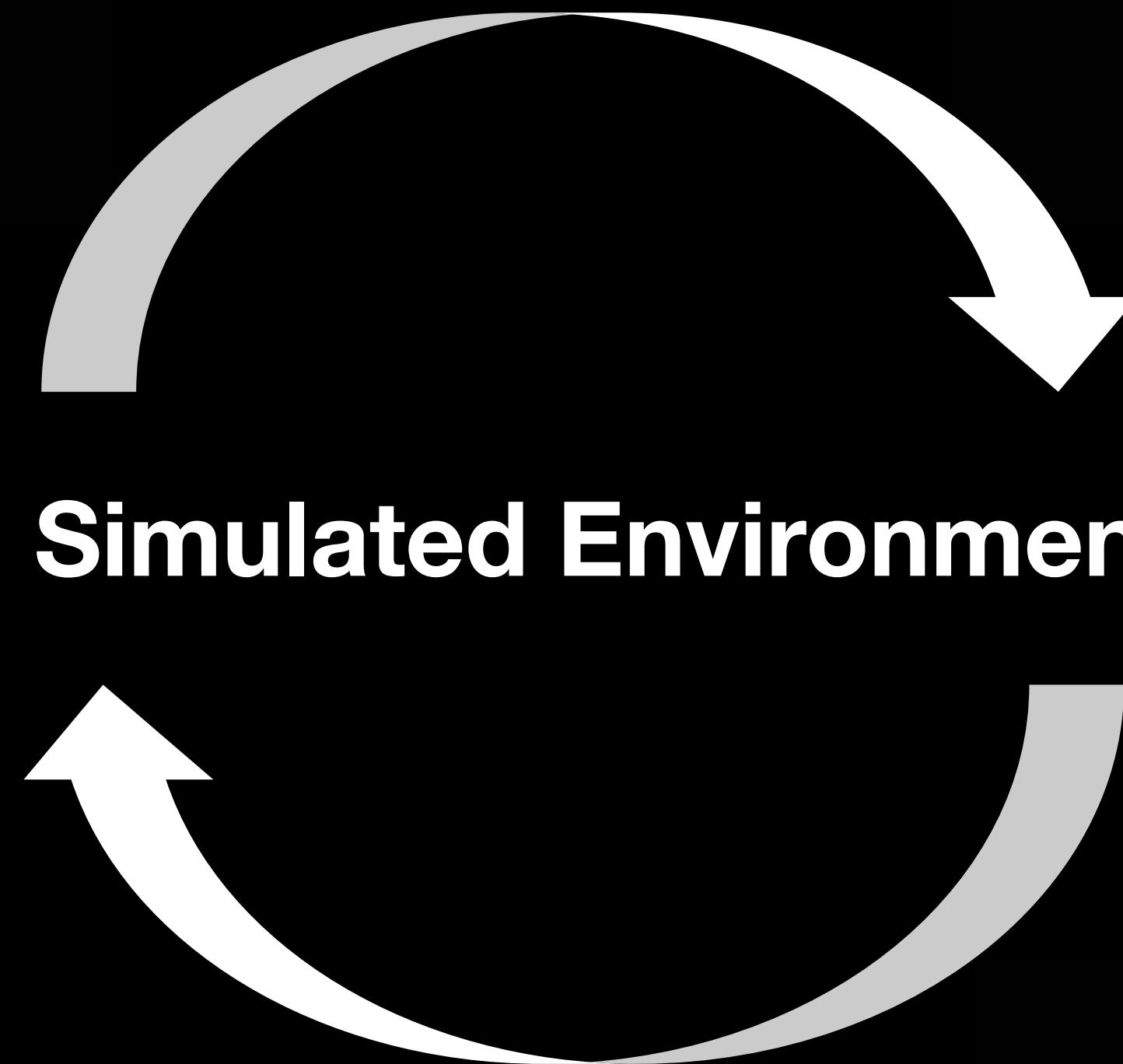
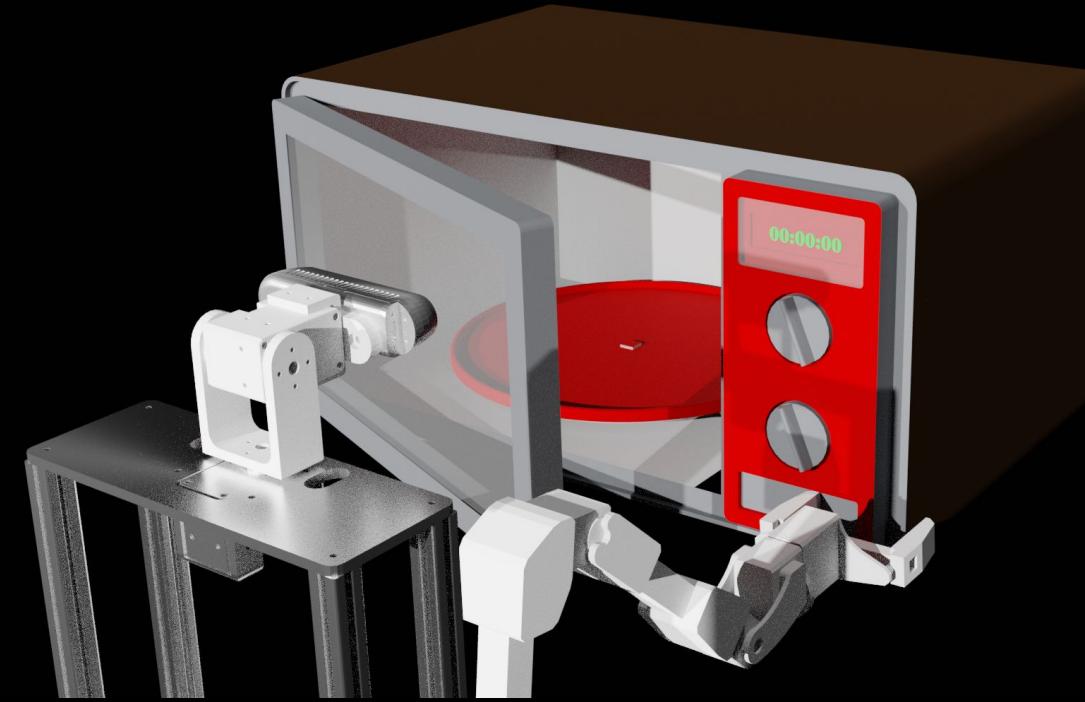




IL Problems

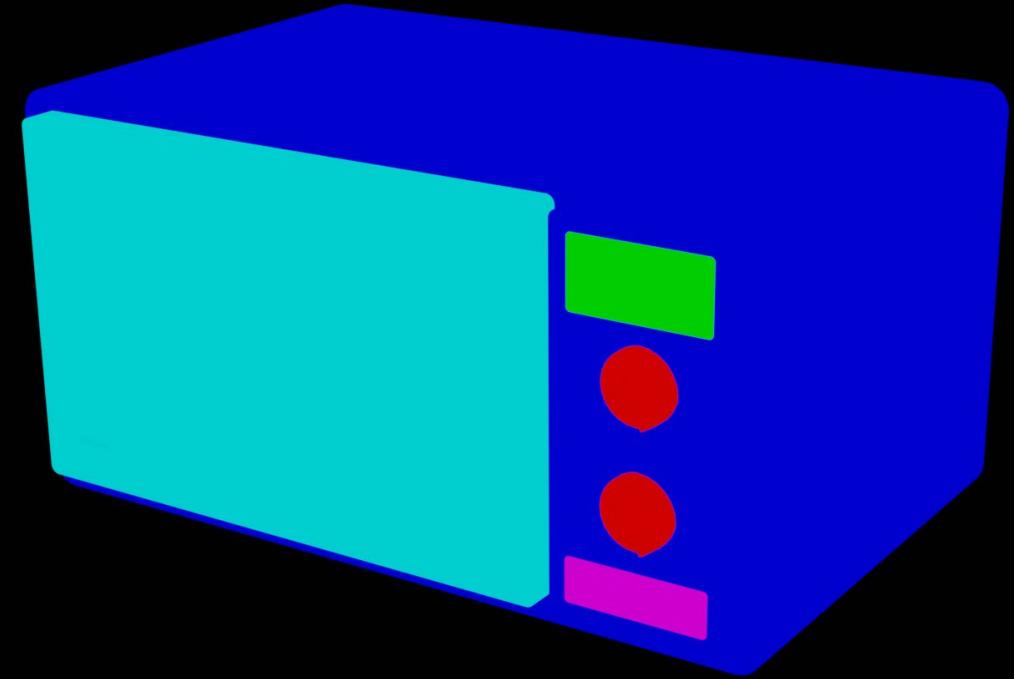
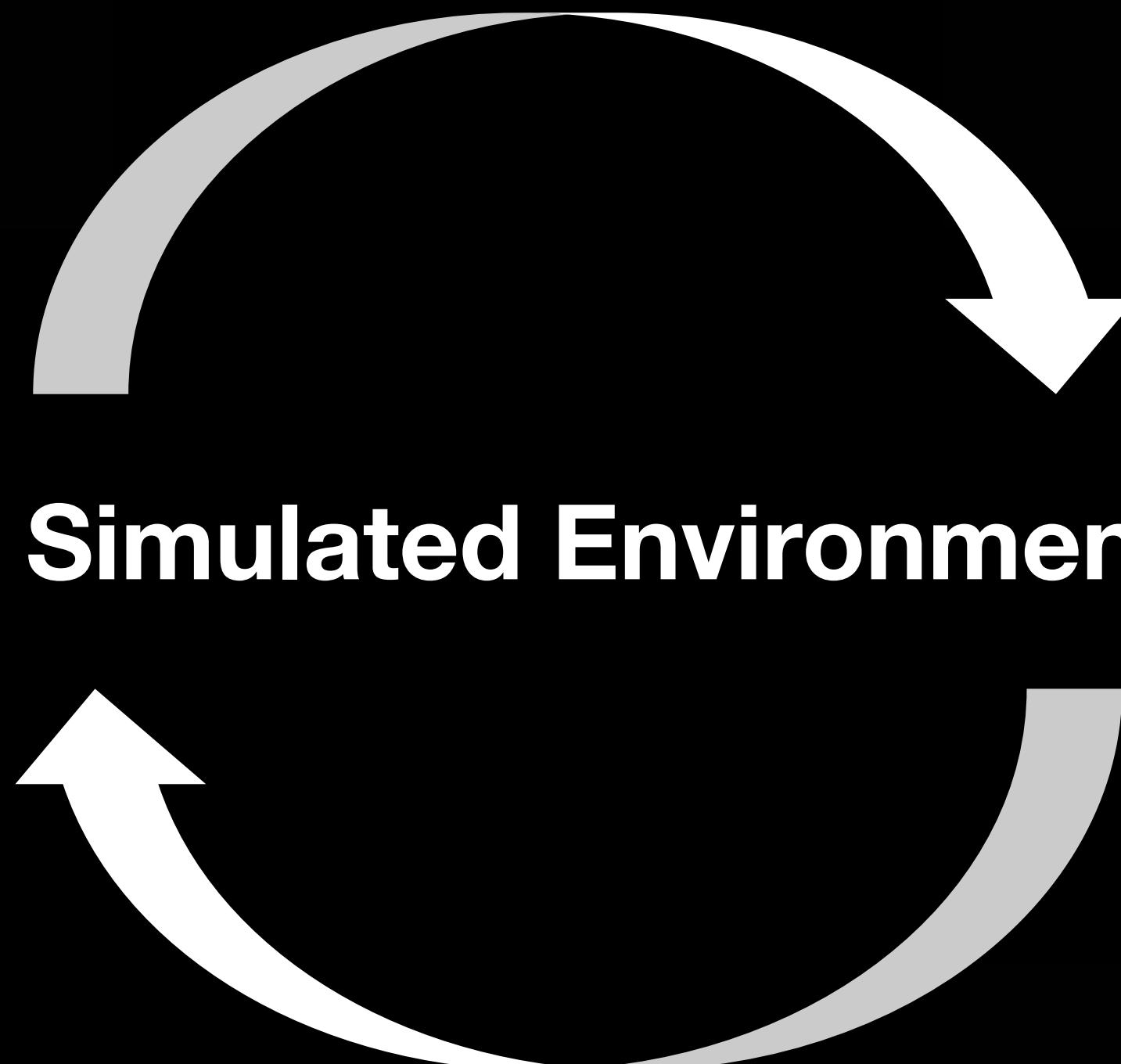
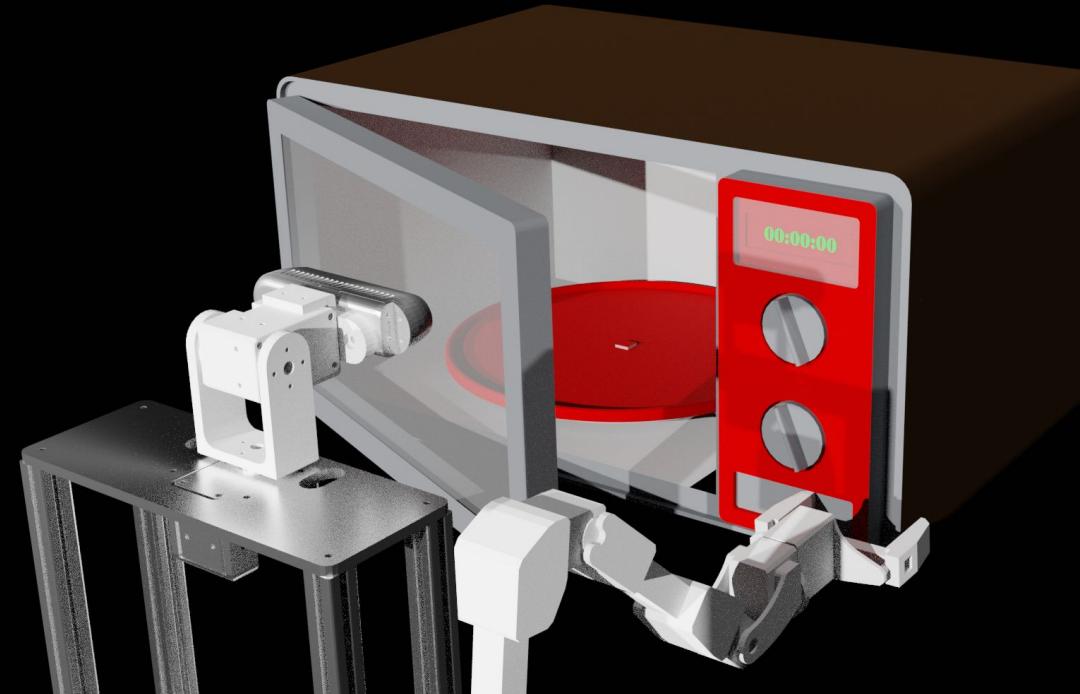
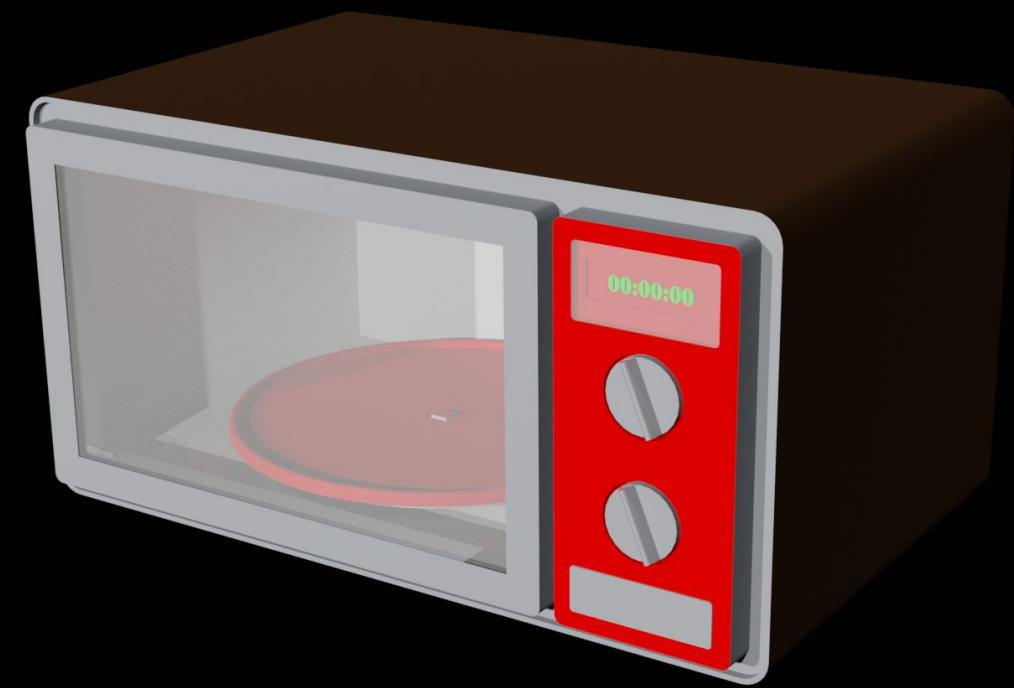
- **Unintuitive control**
- **May not be optimal**
- **Human expertise**





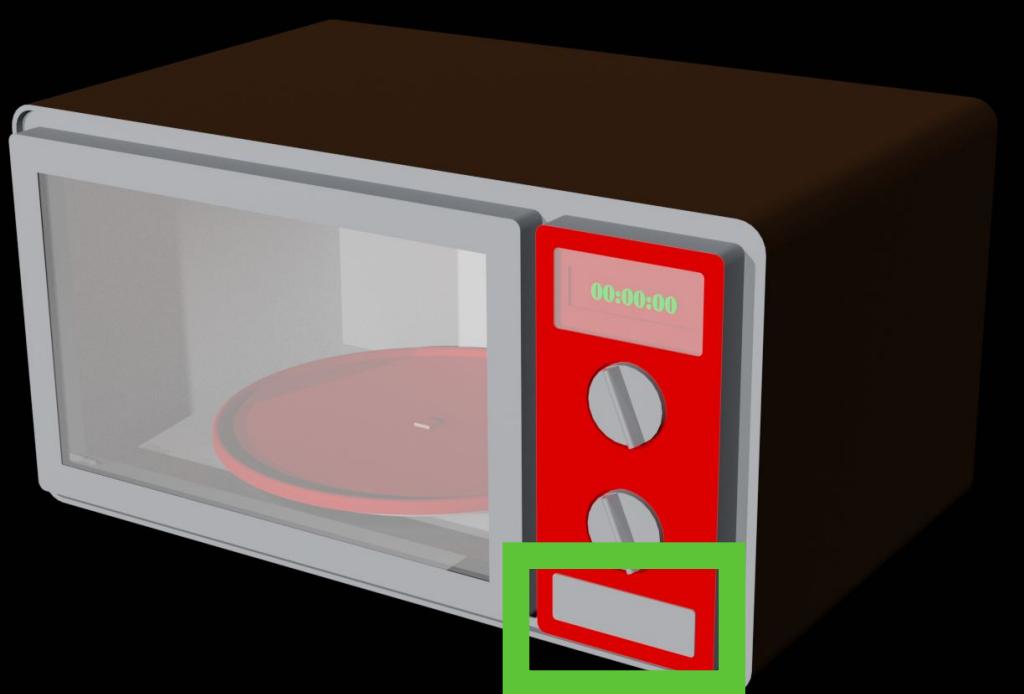
RL

- Will not break hardware
- Easy to scale



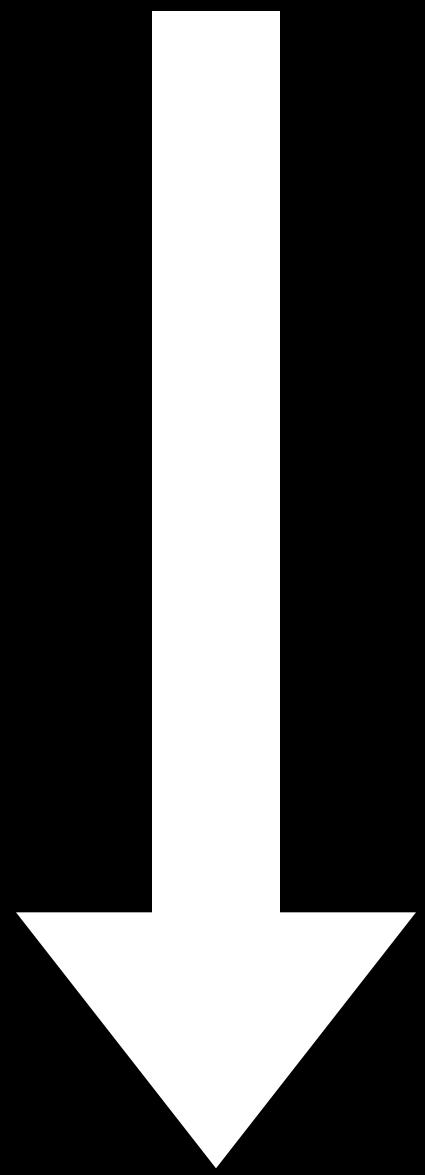
Manual data collection

- More freedom in controller design
- Less expertise required



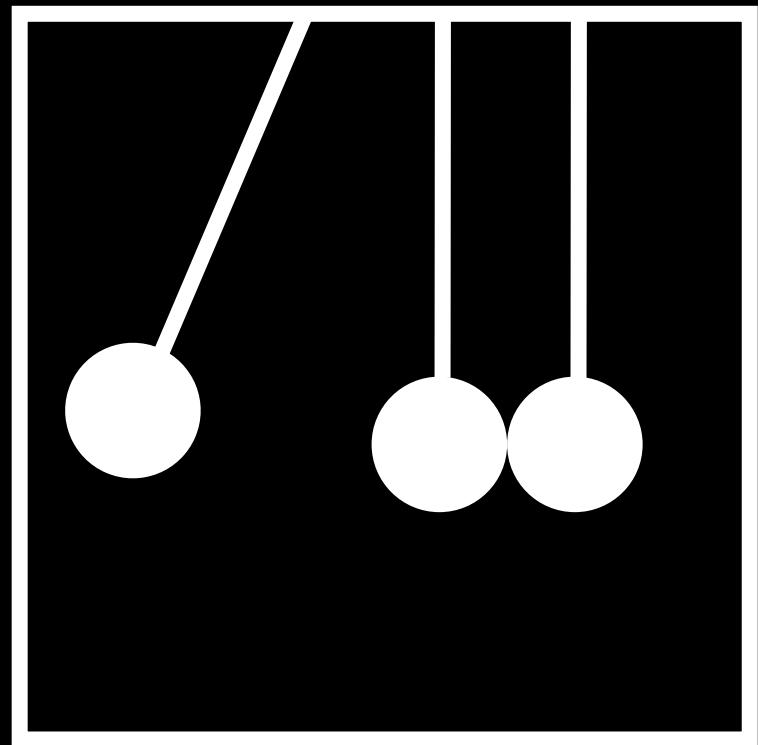
Simulated Environment

Simulated Environment



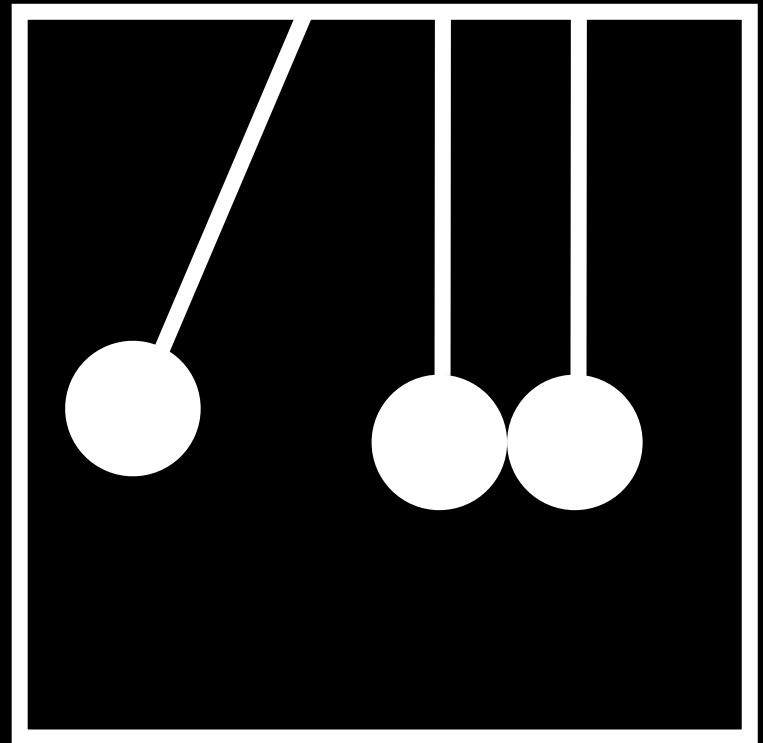
Real World

Simulated Environment

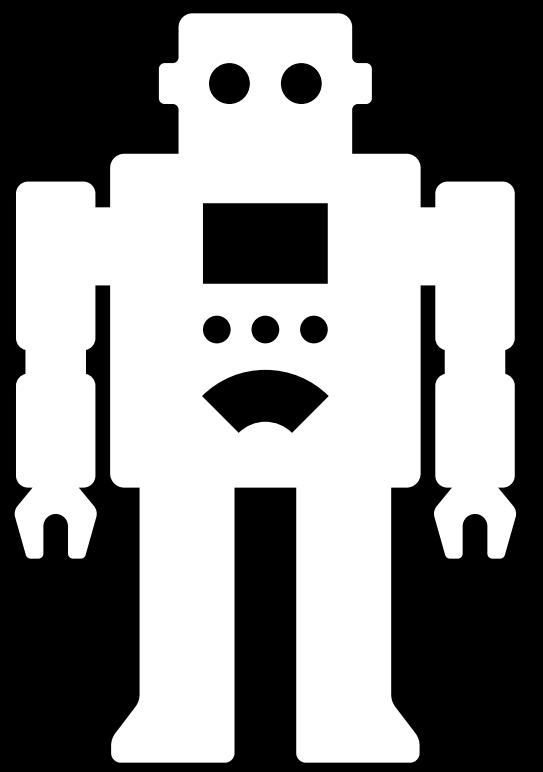


Physics

Simulated Environment

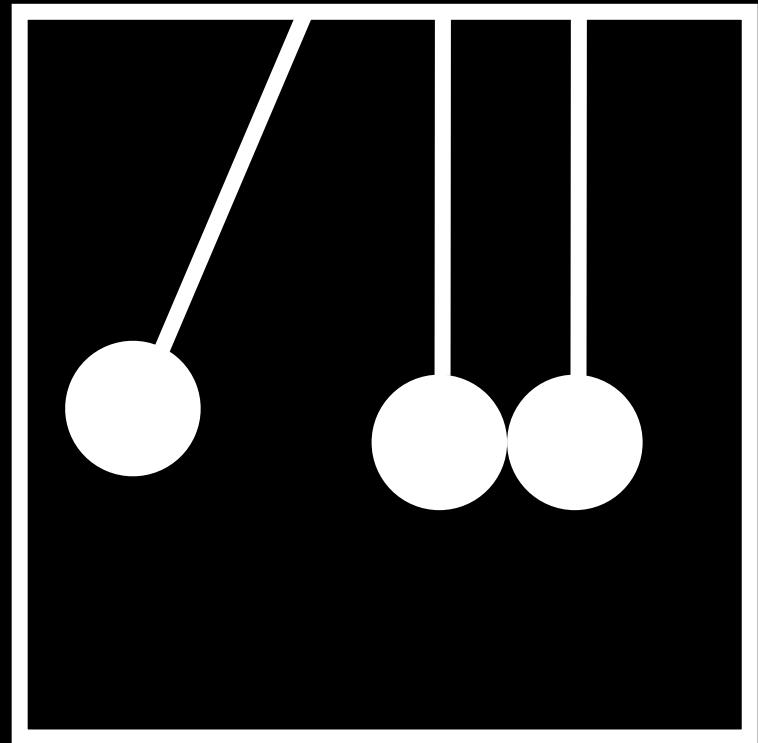


Physics

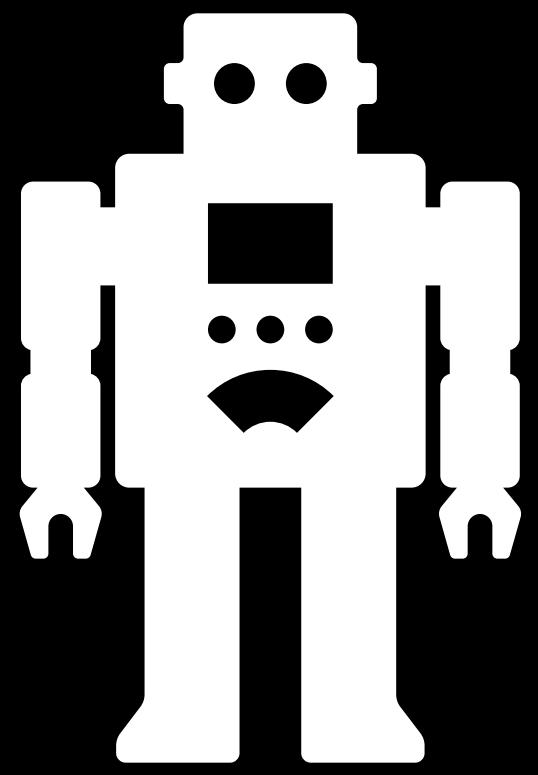


Robotics

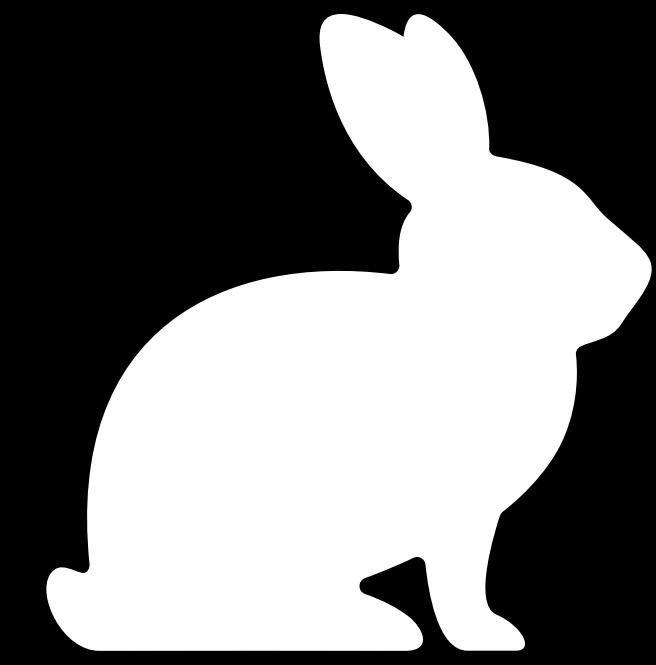
Simulated Environment



Physics

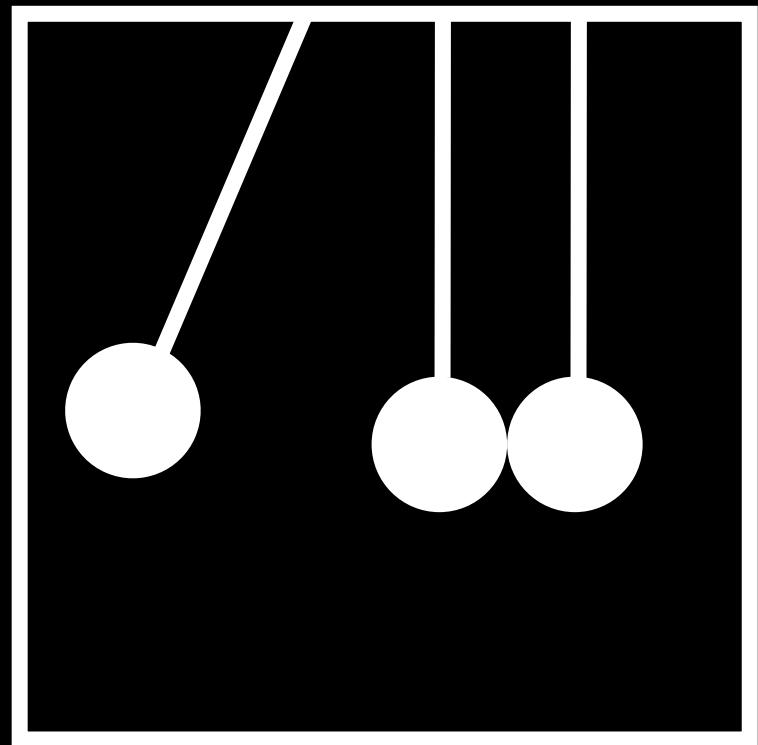


Robotics

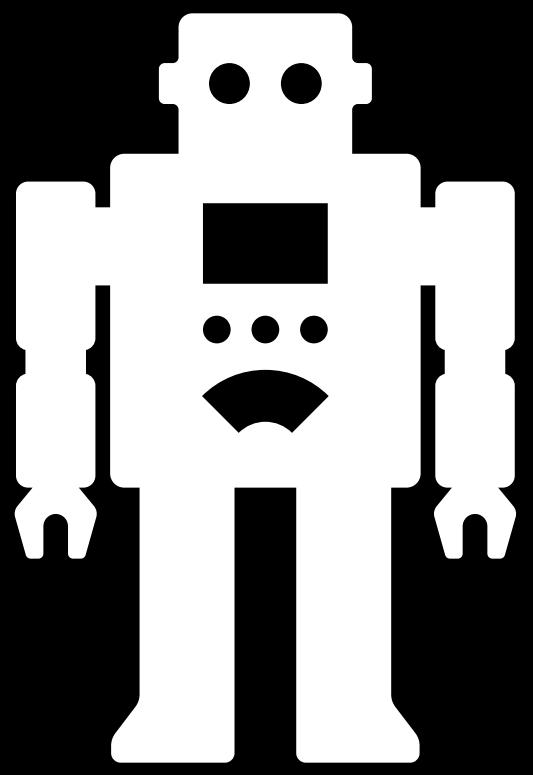


Rendering

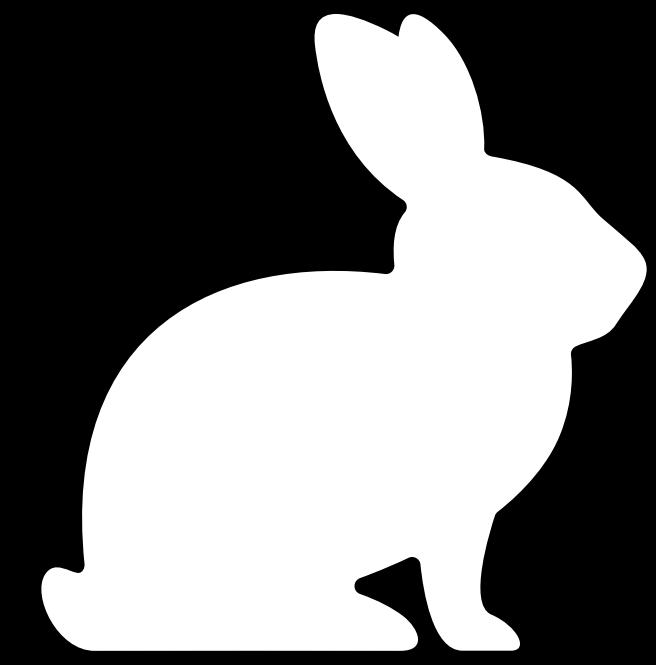
Simulated Environment



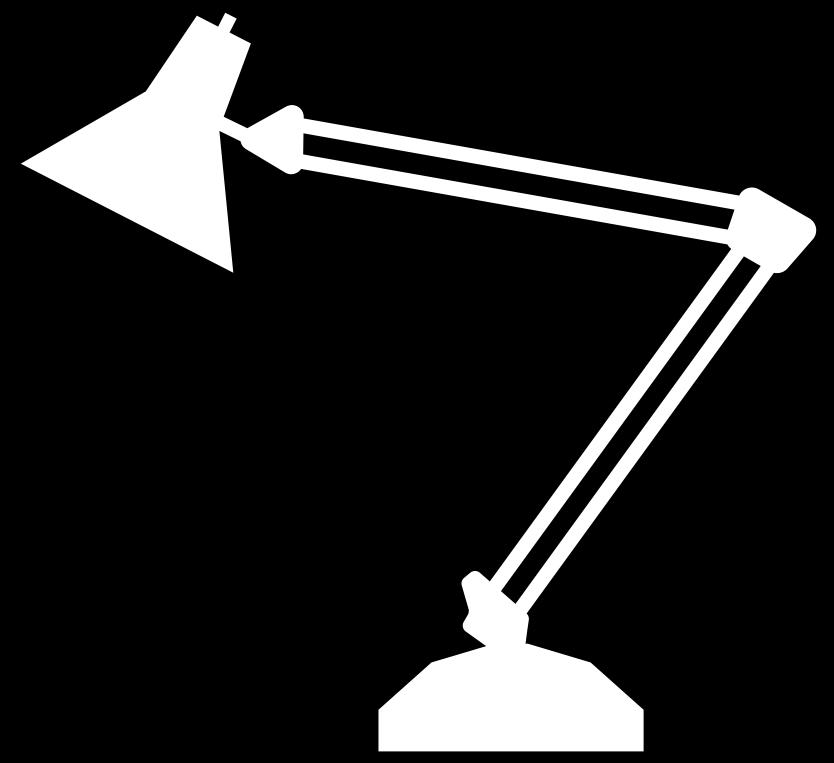
Physics



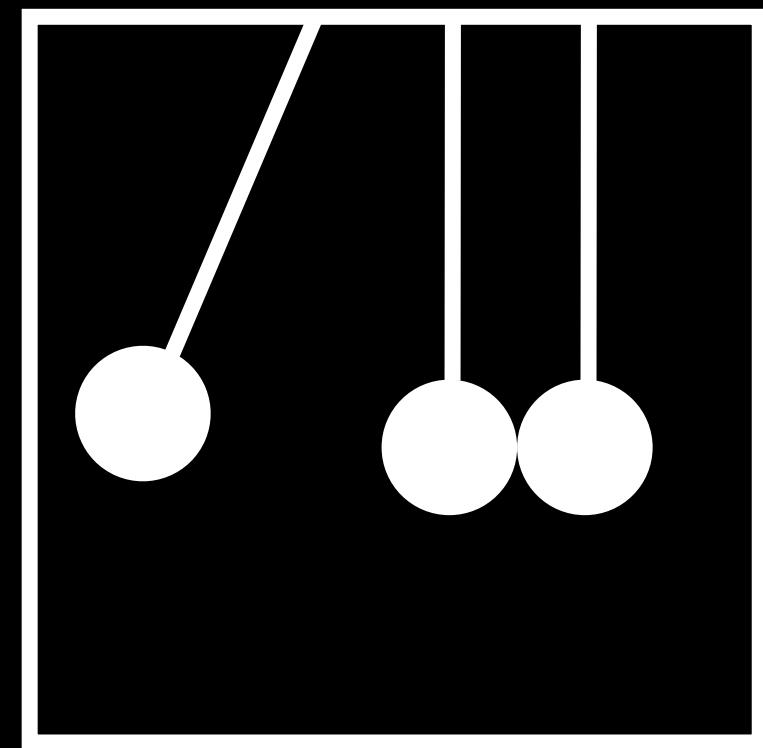
Robotics



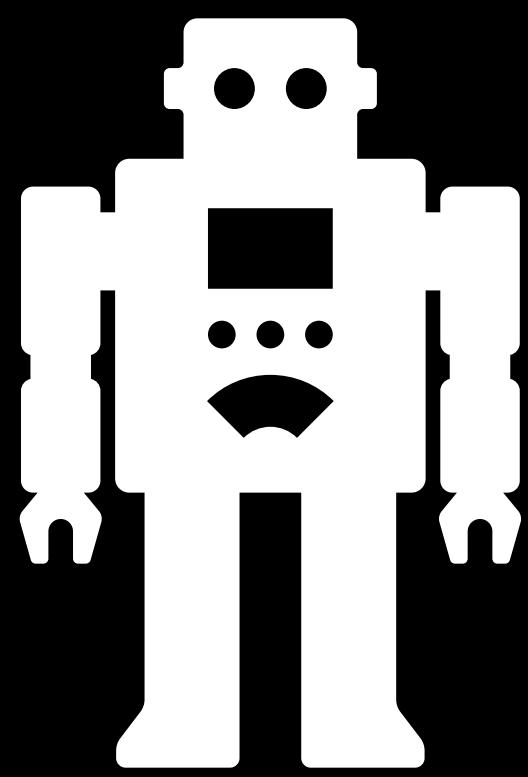
Rendering



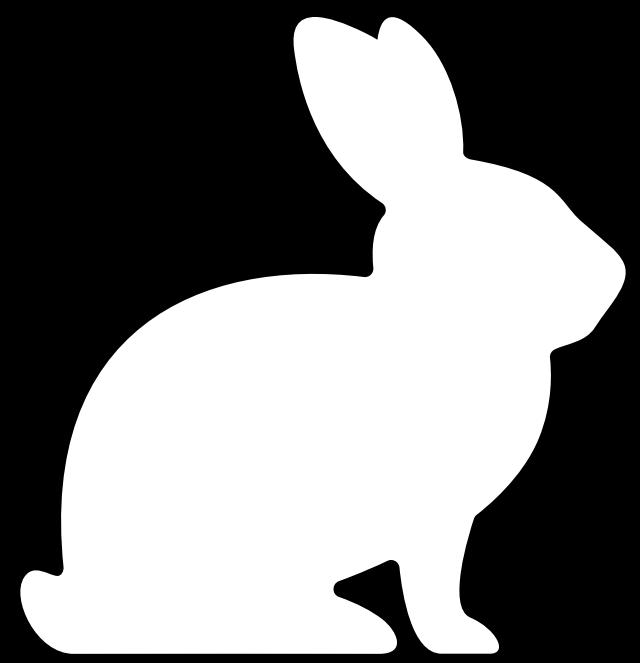
Simulation Content



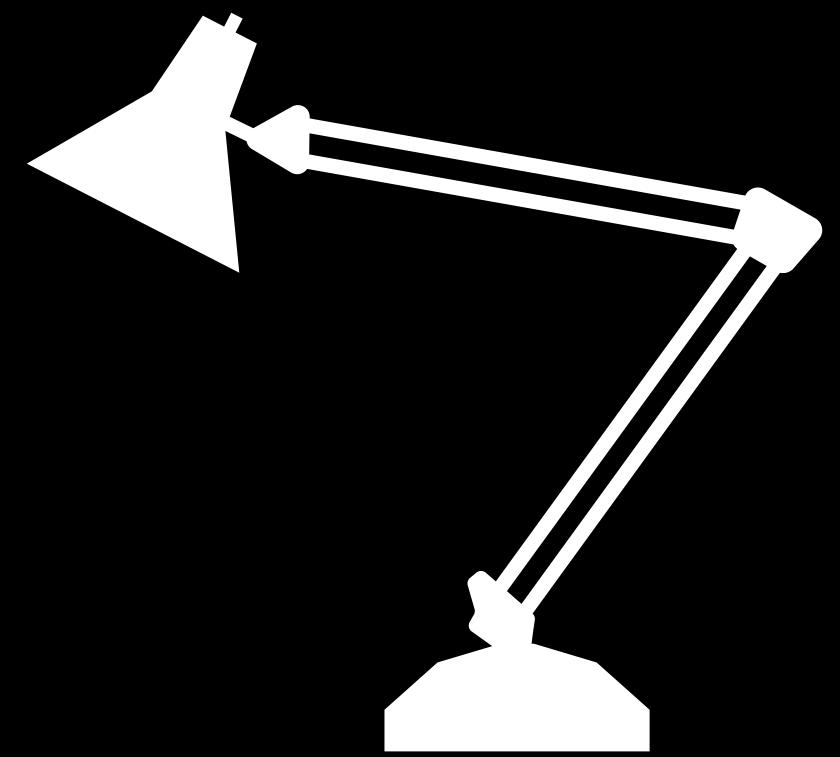
Physics



Robotics



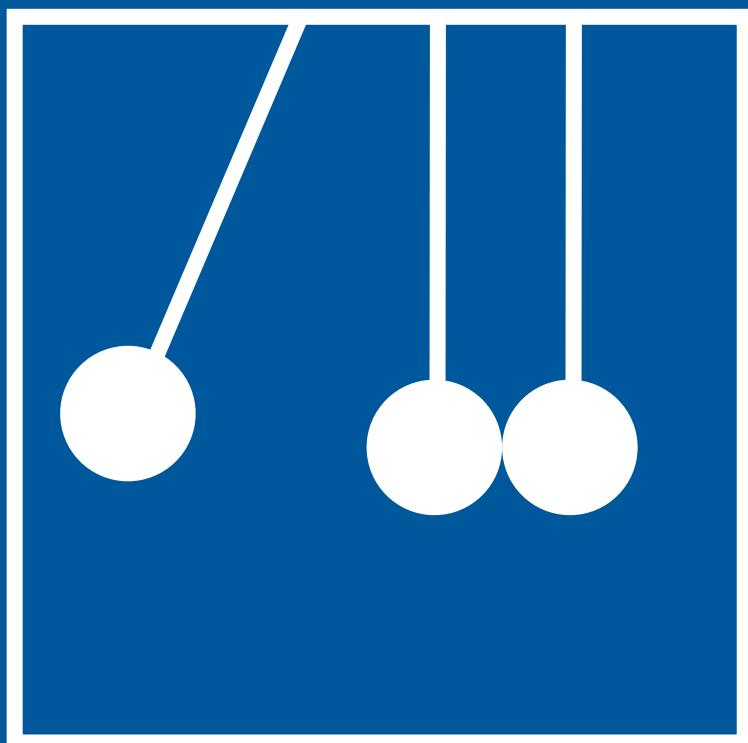
Rendering



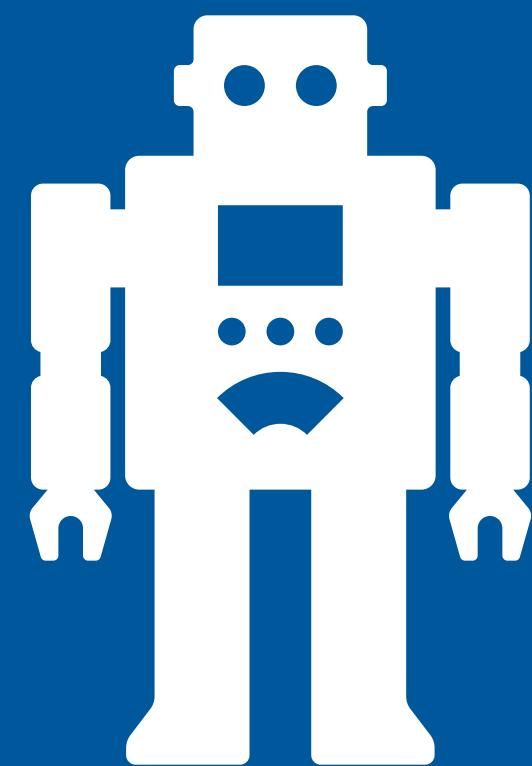
Simulation Content

SAPIEN

SAPIEN Engine

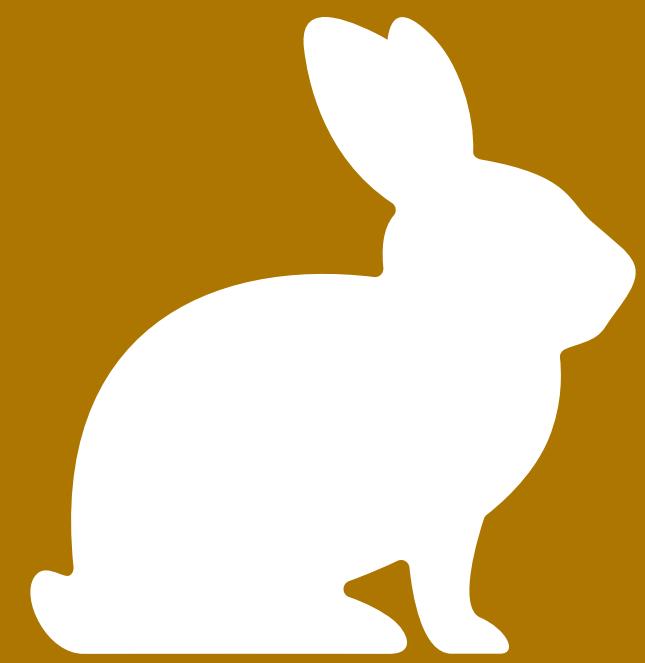


Physics



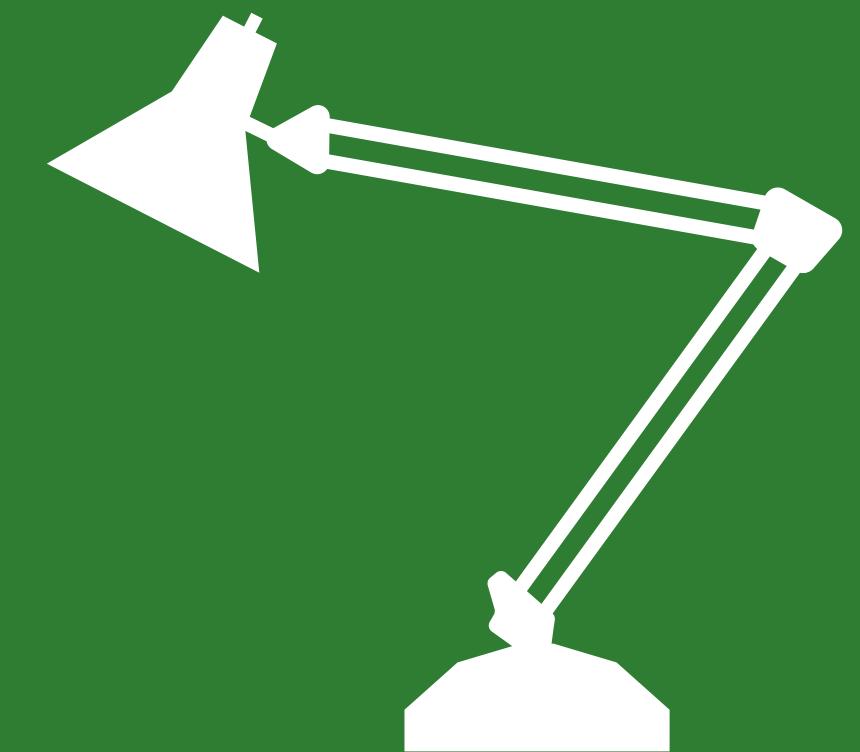
Robotics

SAPIEN Renderer



Rendering

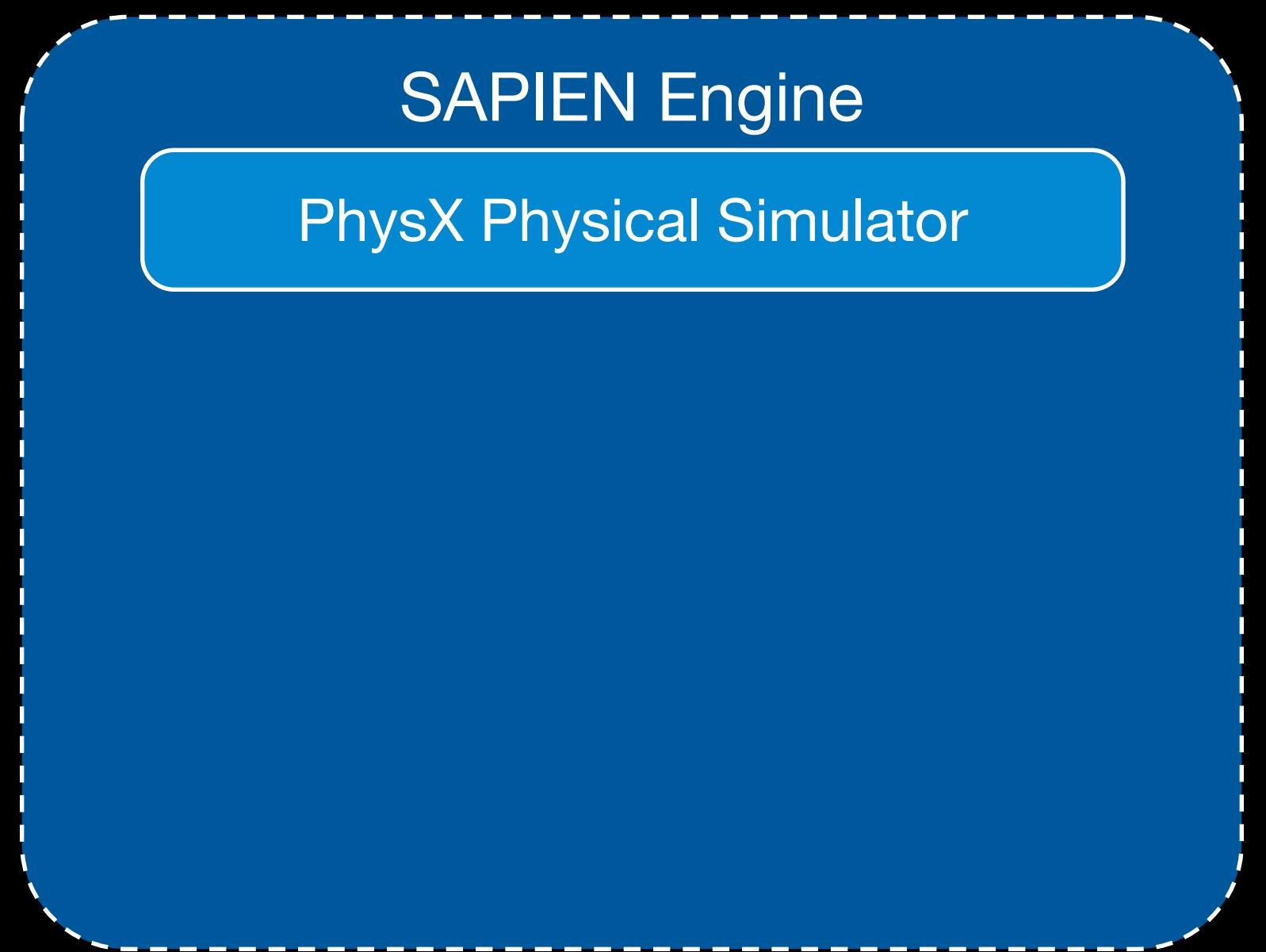
SAPIEN Asset



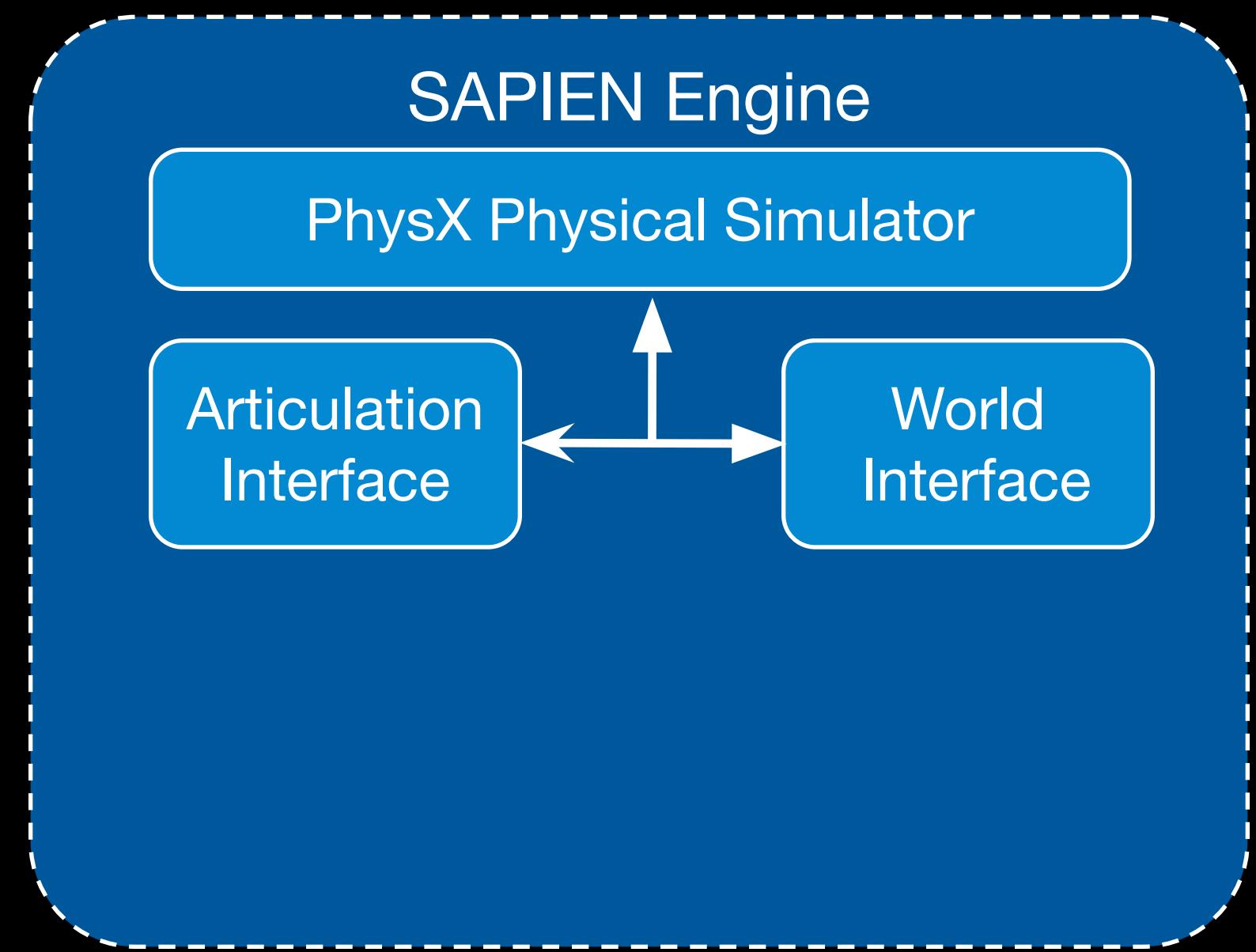
Simulation Content

SAPIEN Engine

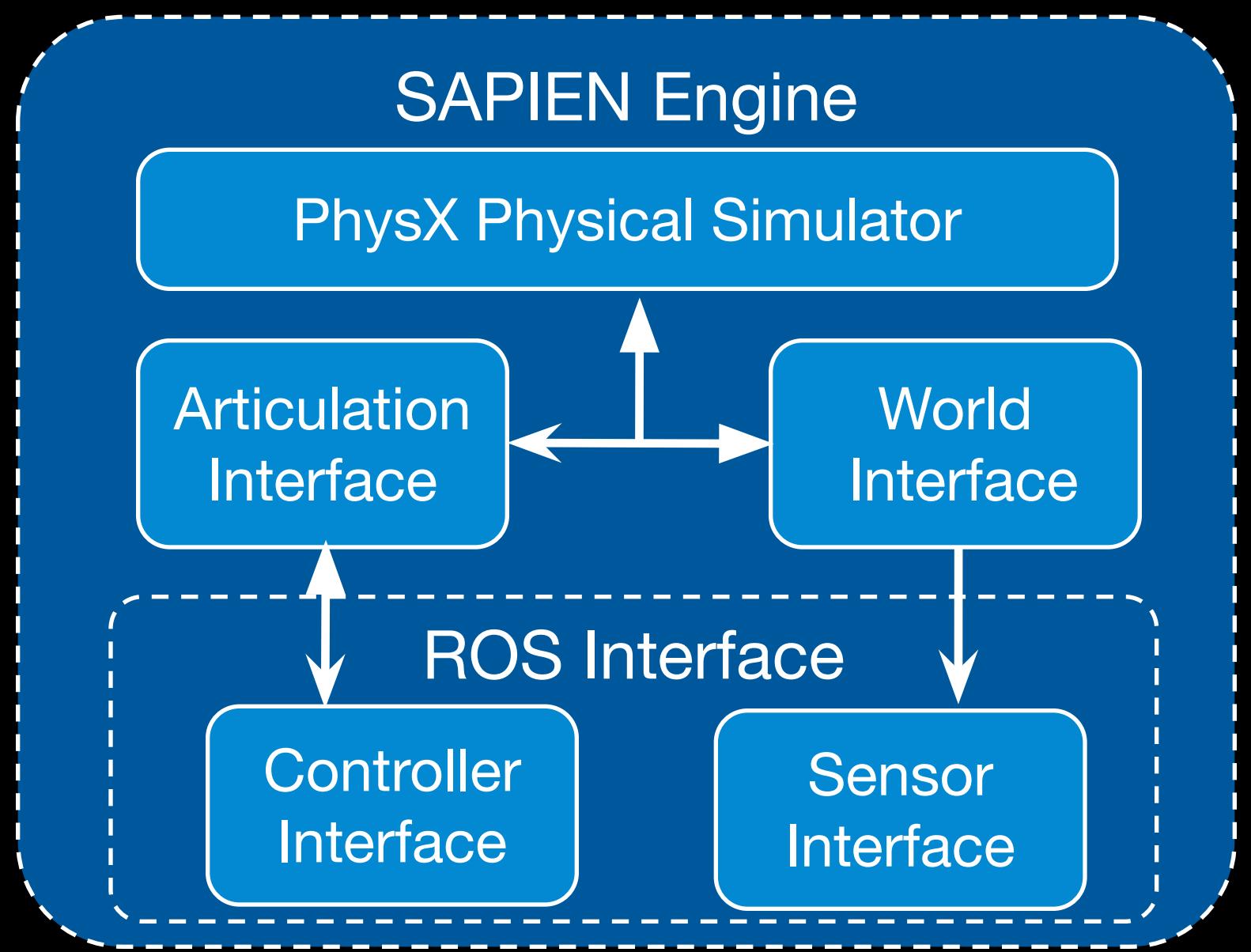
SAPIEN Engine



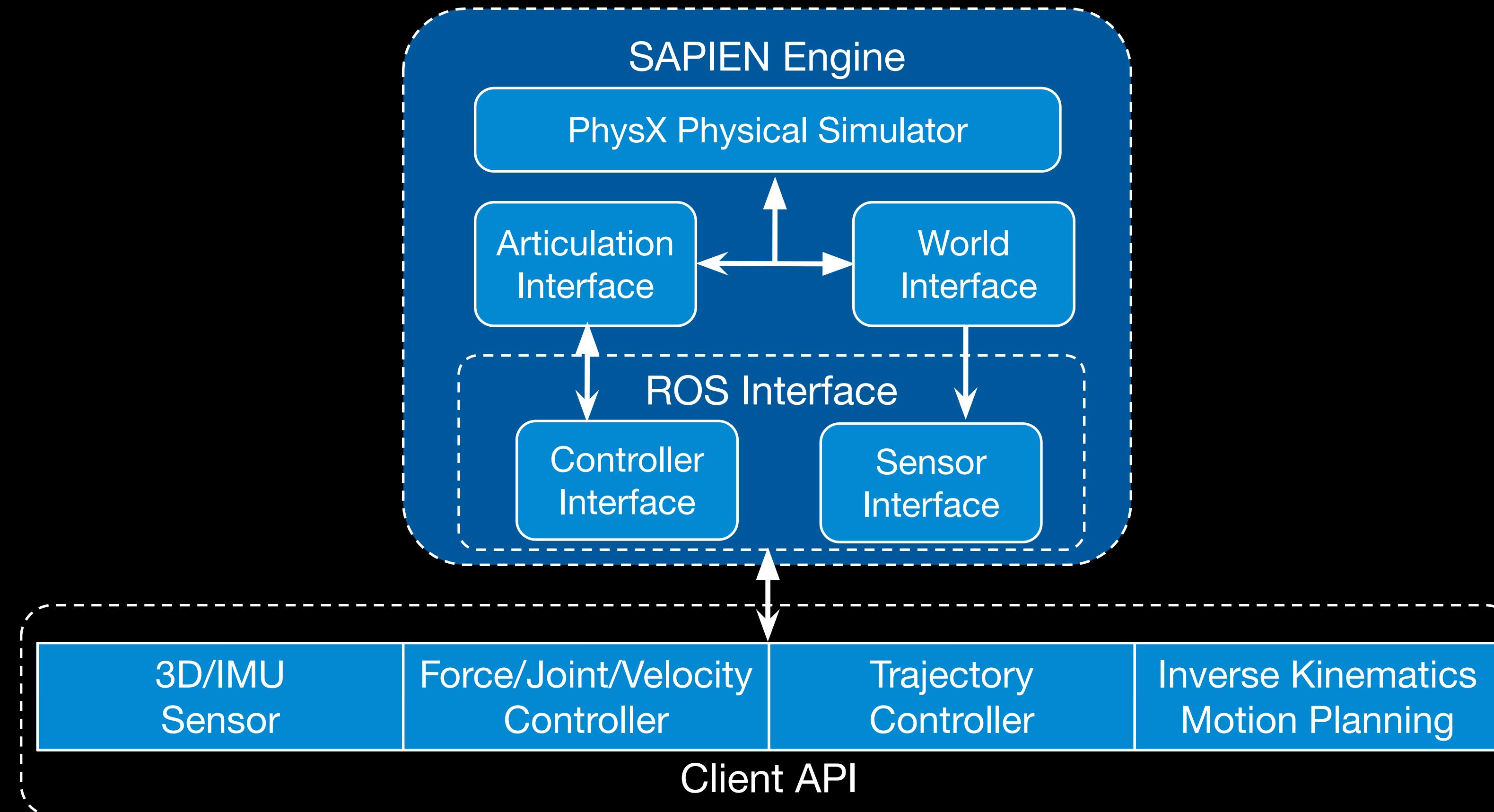
SAPIEN Engine



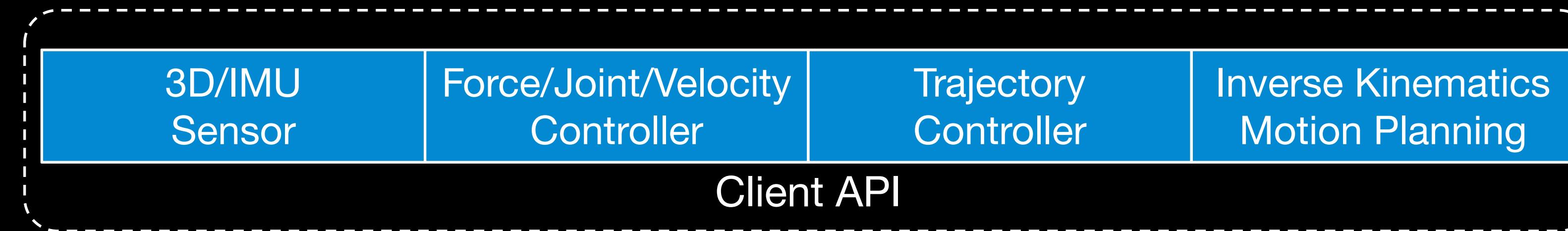
SAPIEN Engine



SAPIEN Engine



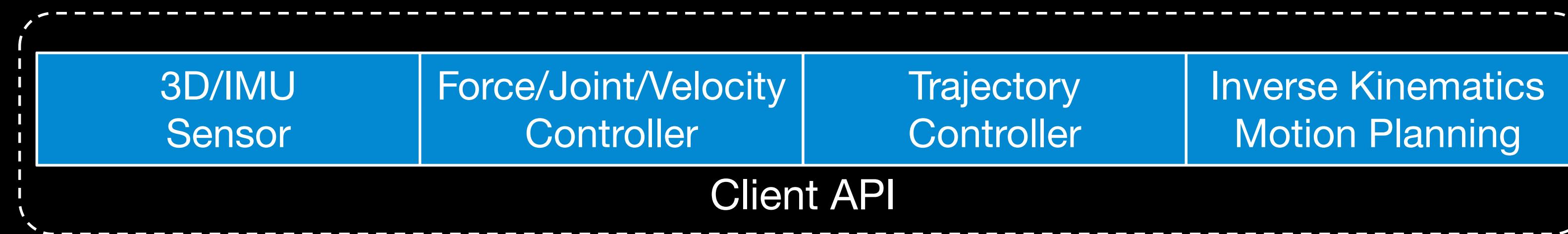
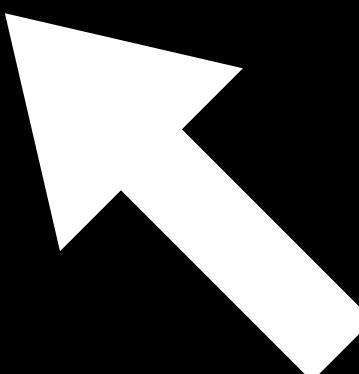
SAPIEN Engine



SAPIEN Engine



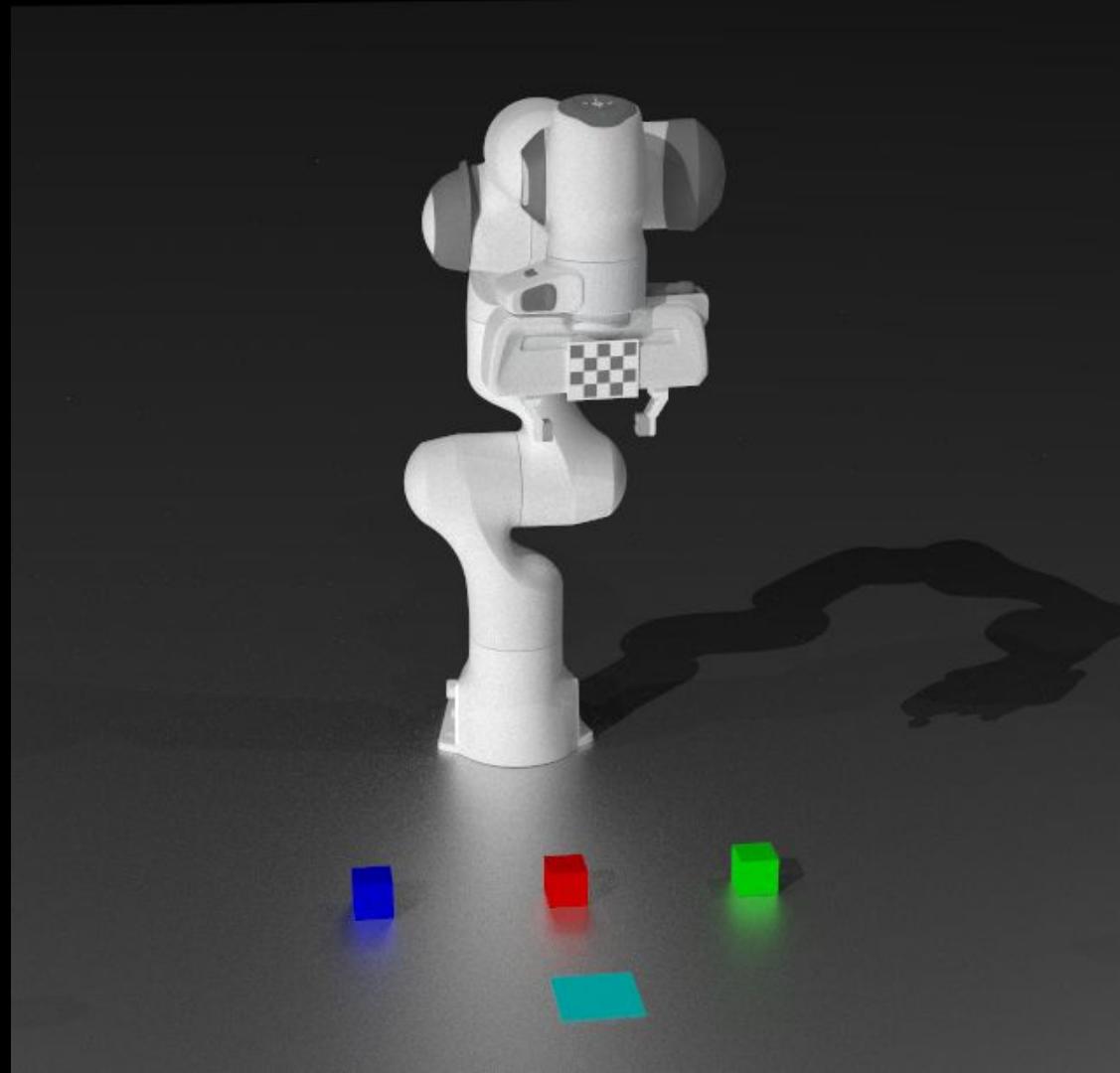
Reinforcement Learning



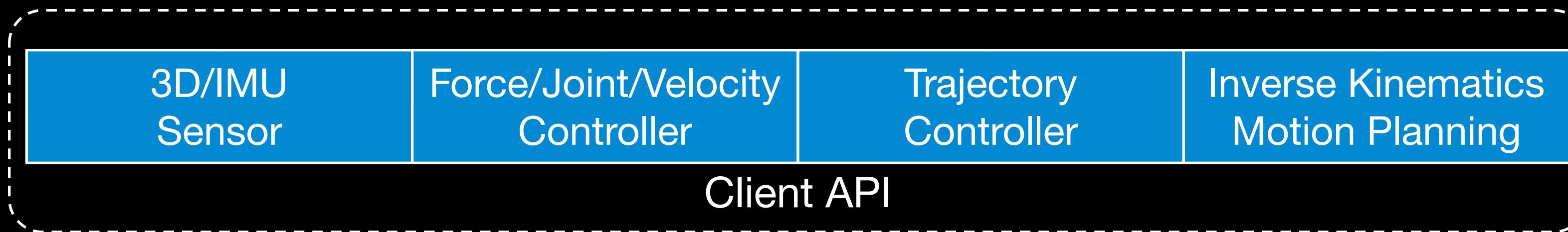
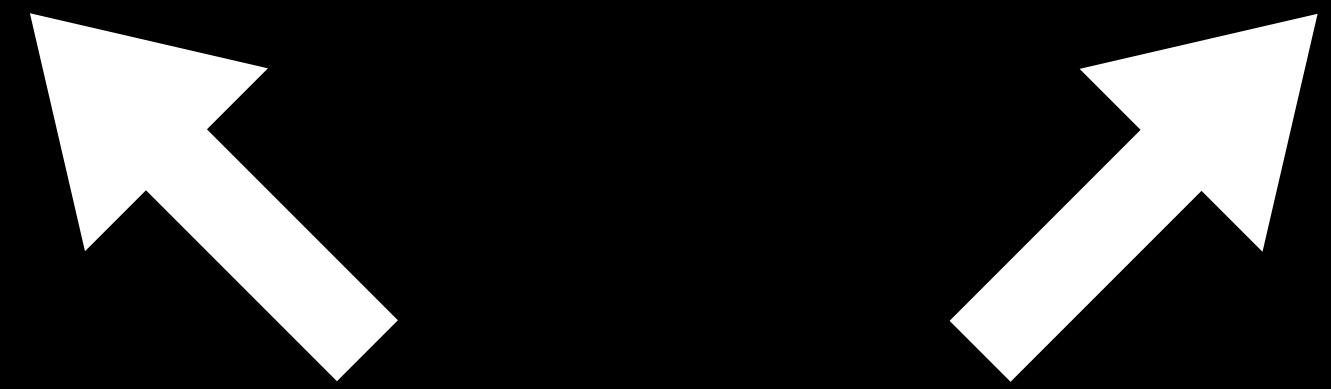
SAPIEN Engine



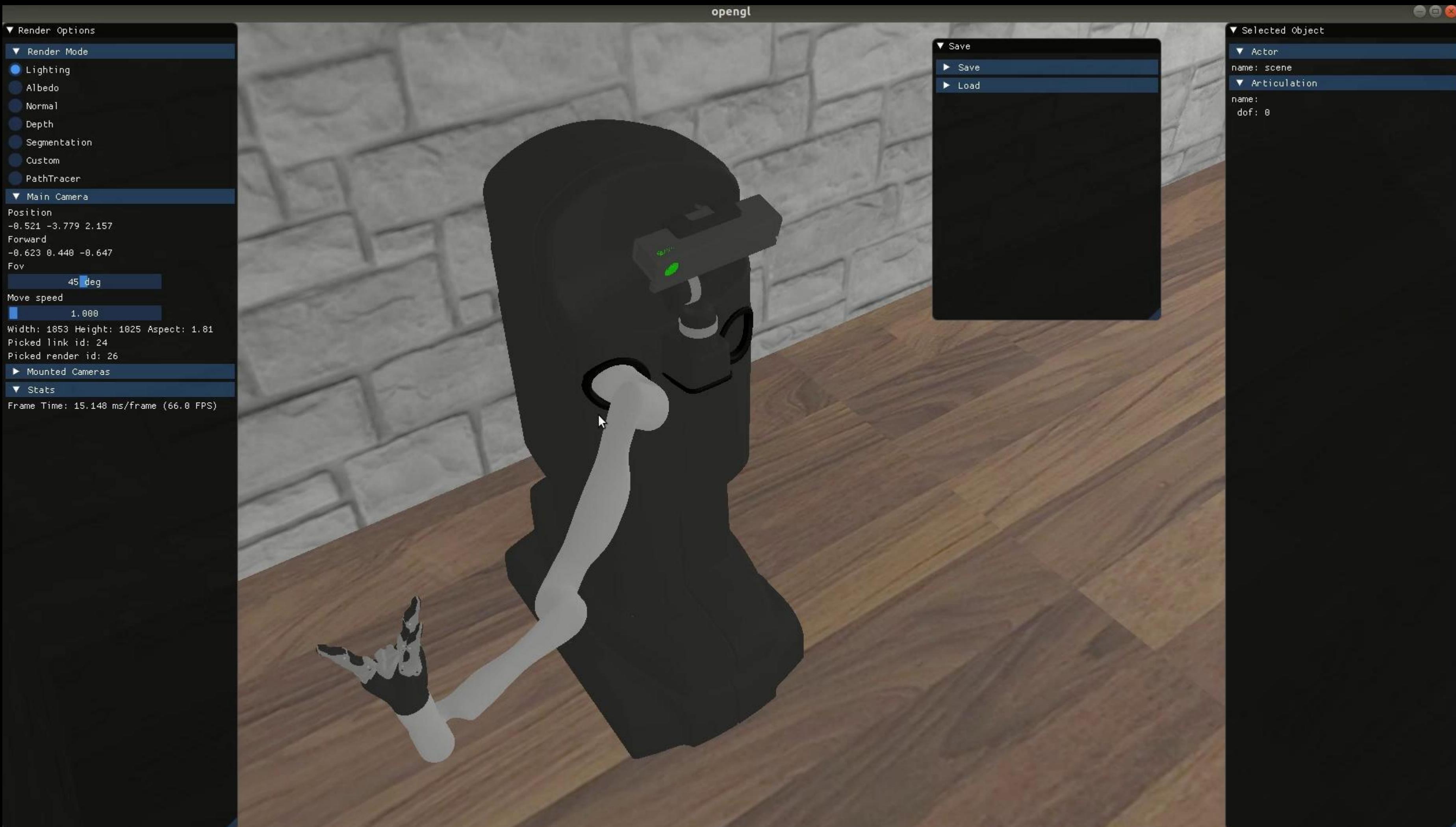
Reinforcement Learning



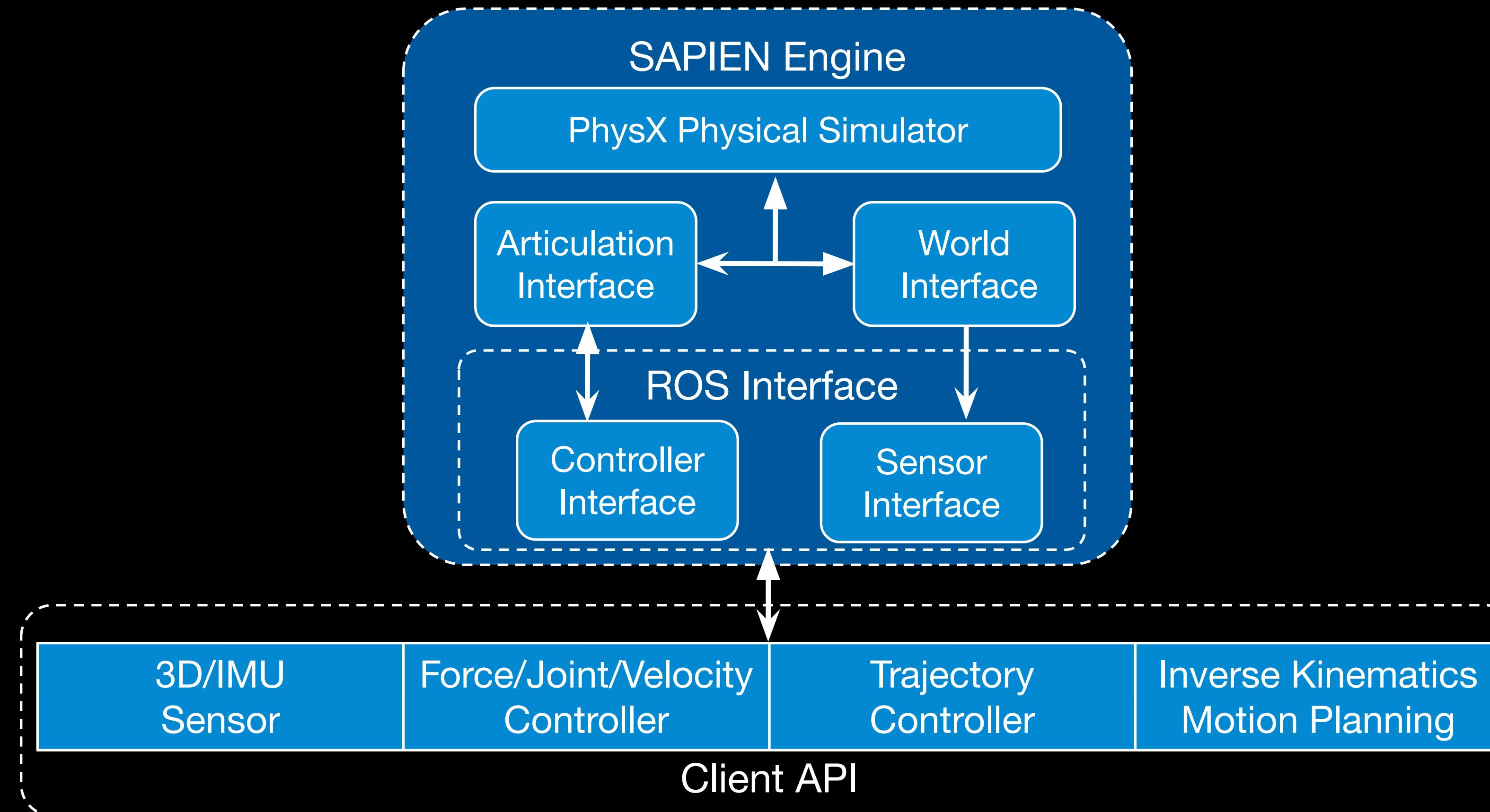
Robot Tasks



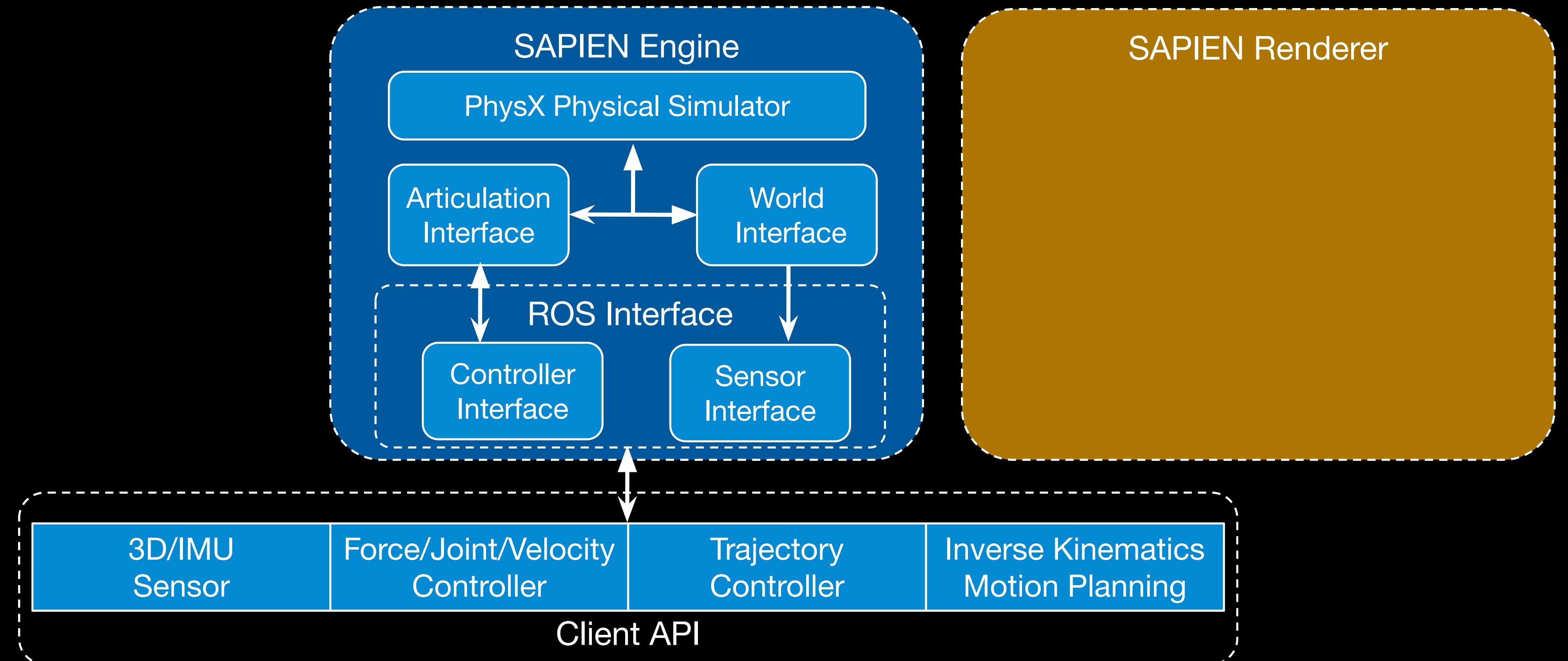
SAPIEN Engine



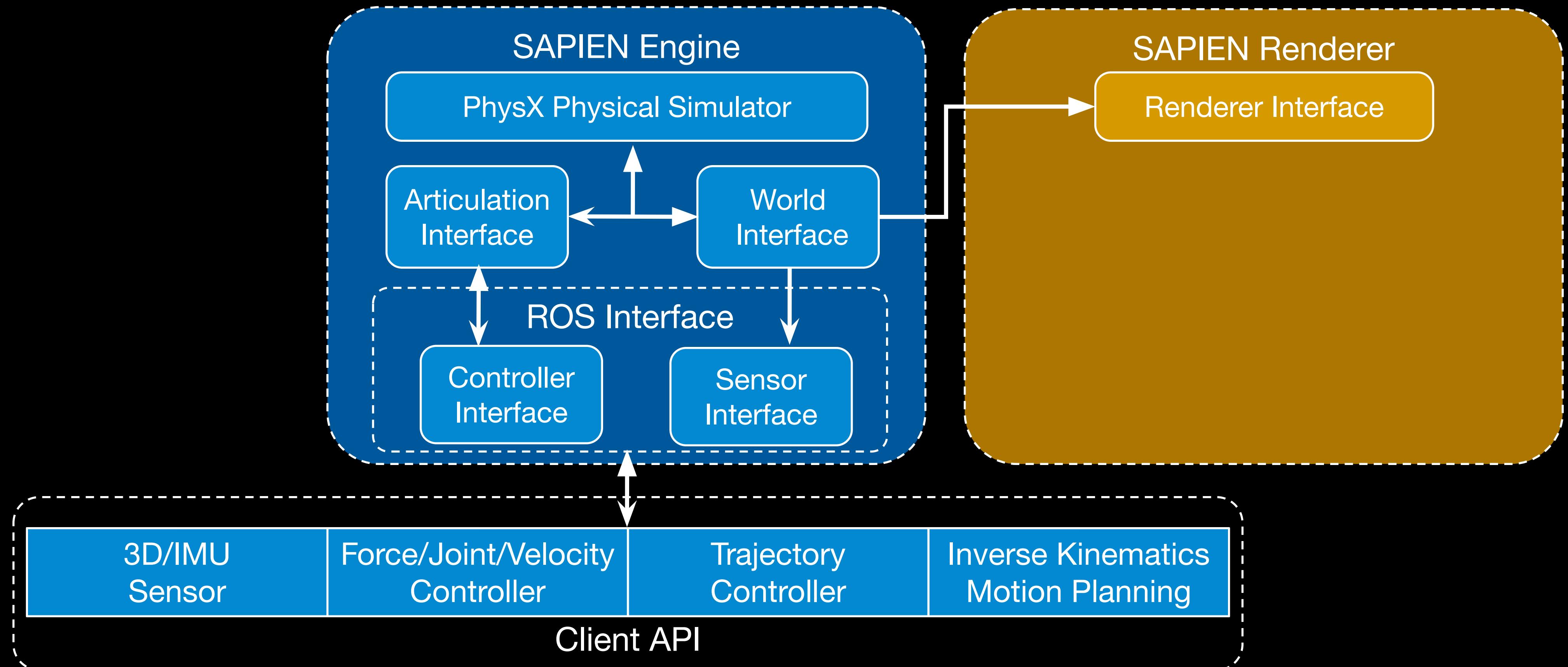
SAPIEN Renderer



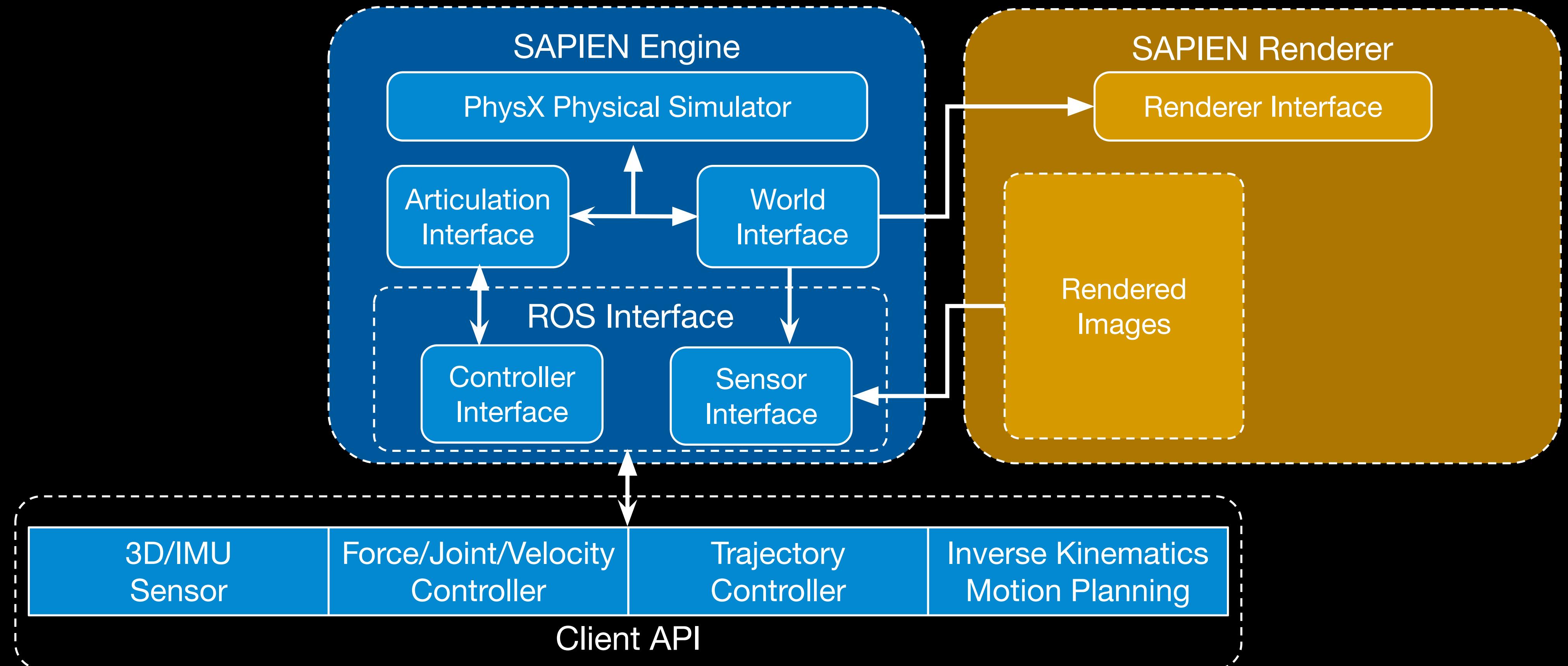
SAPIEN Renderer



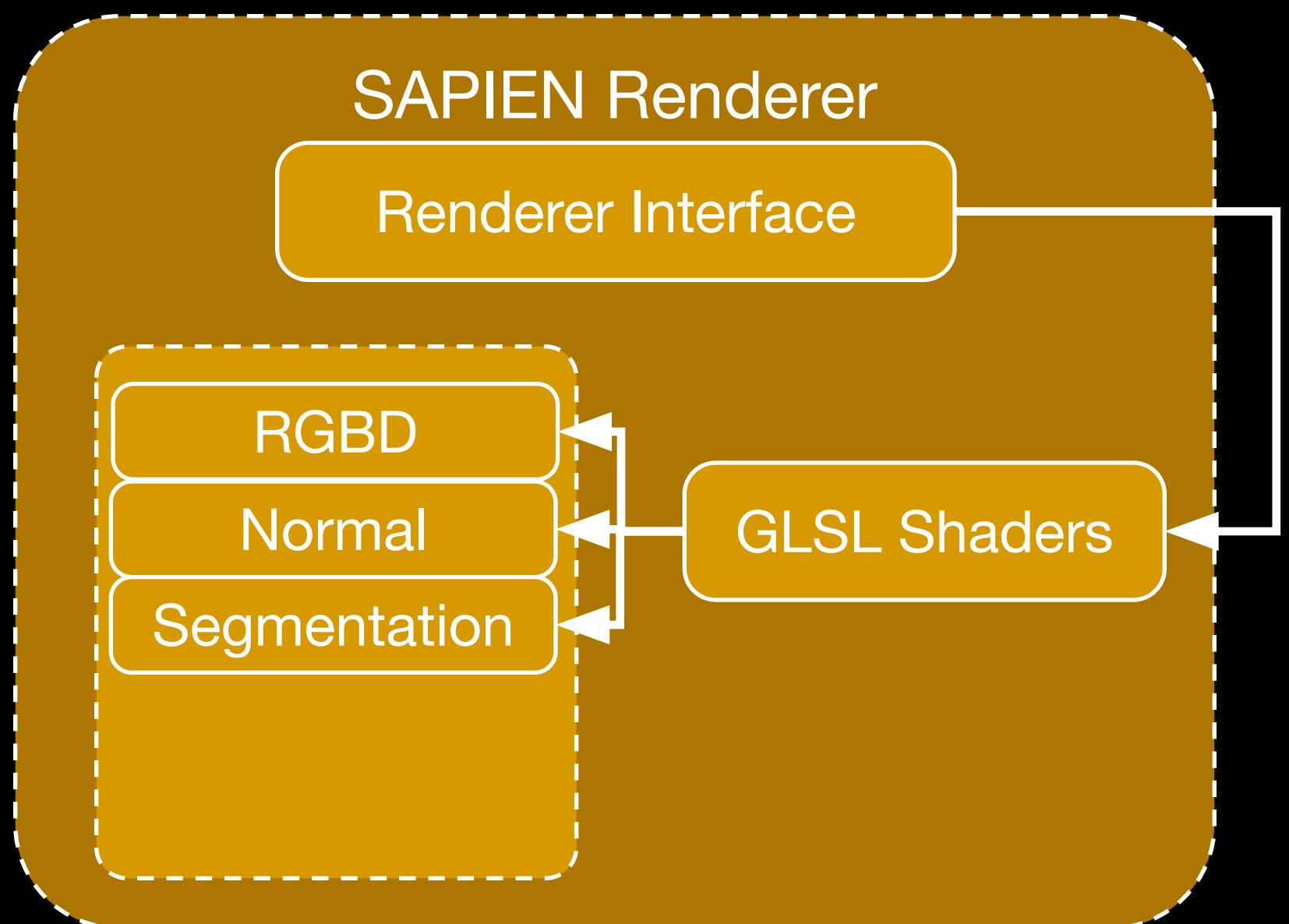
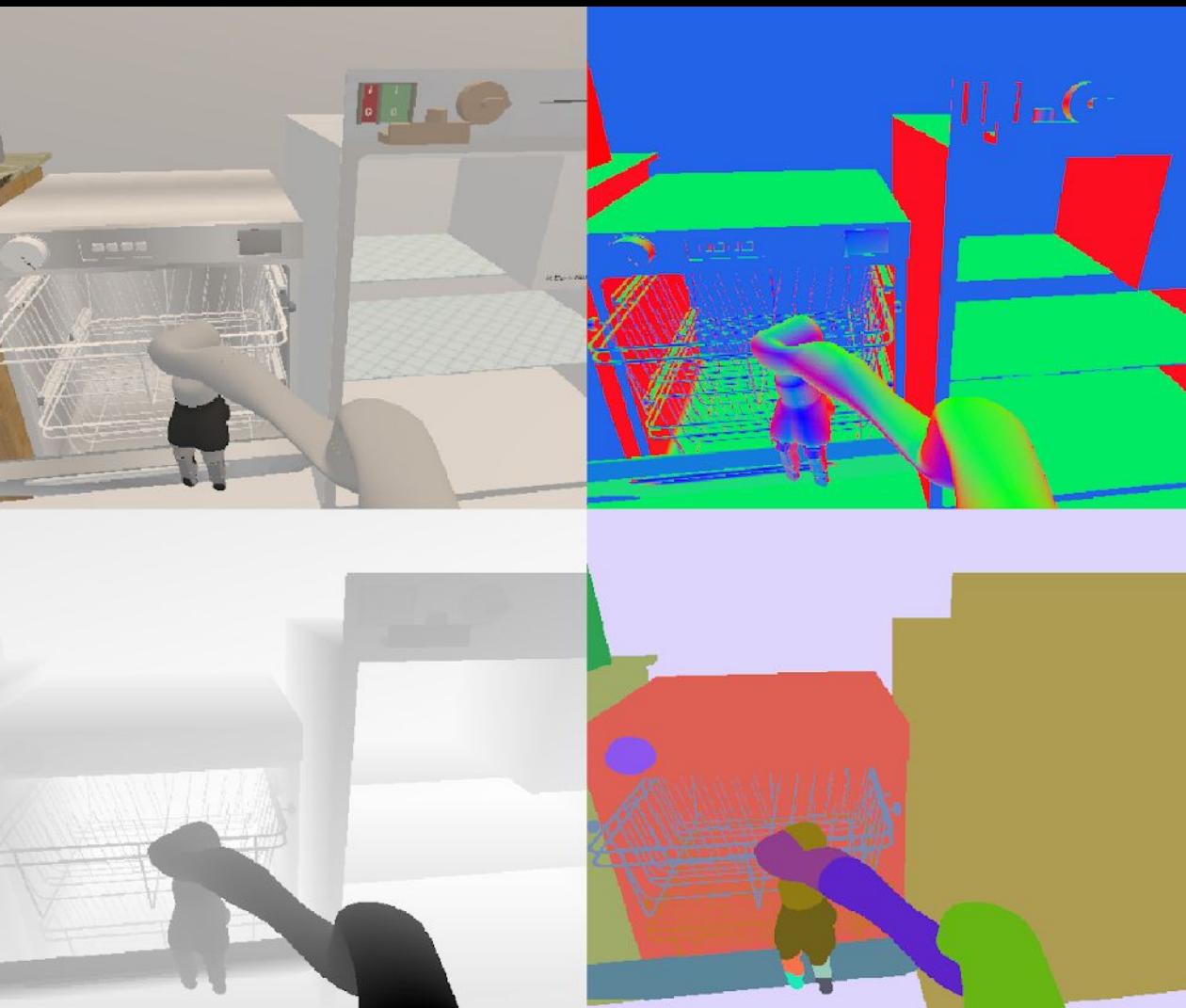
SAPIEN Renderer



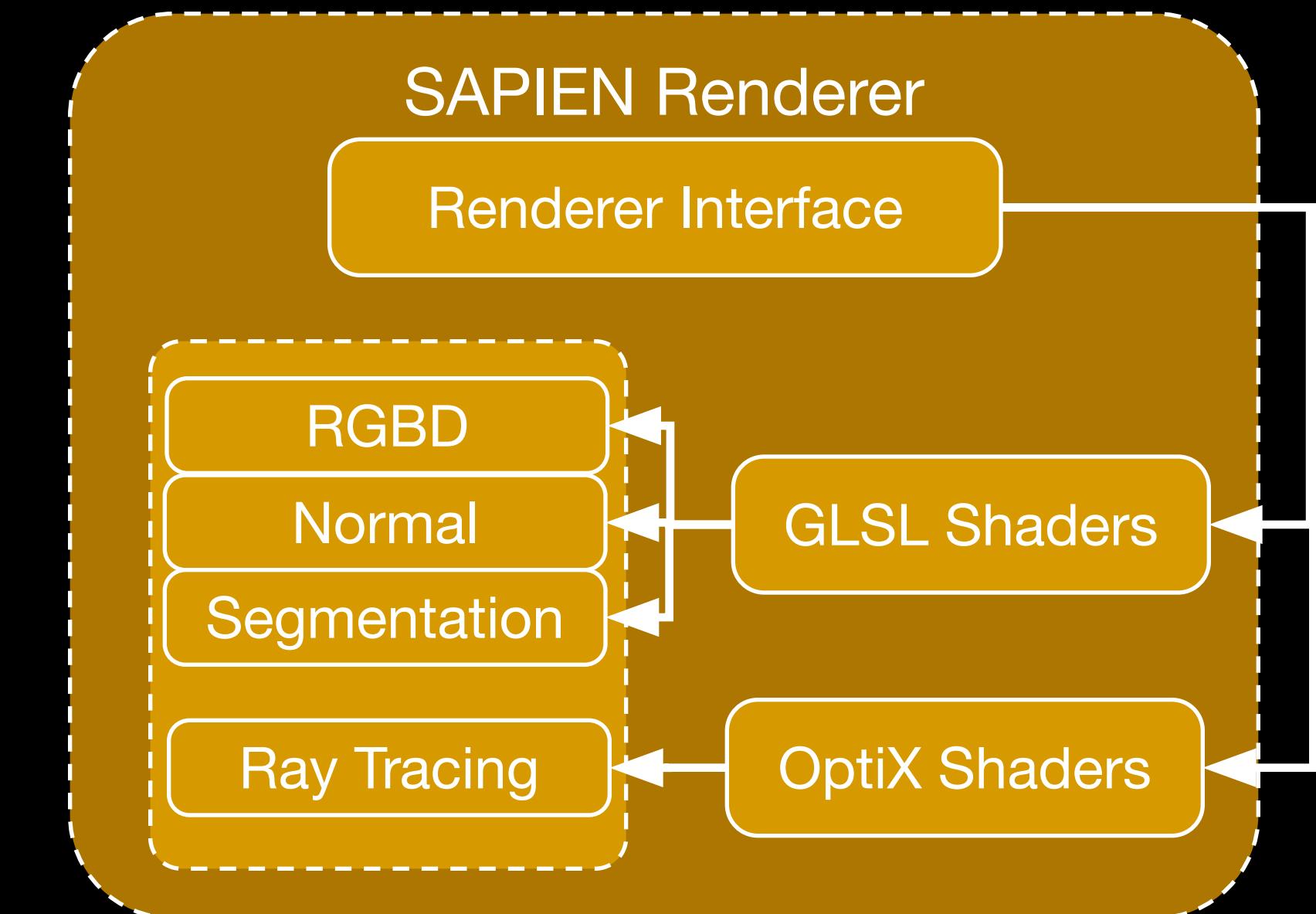
SAPIEN Renderer



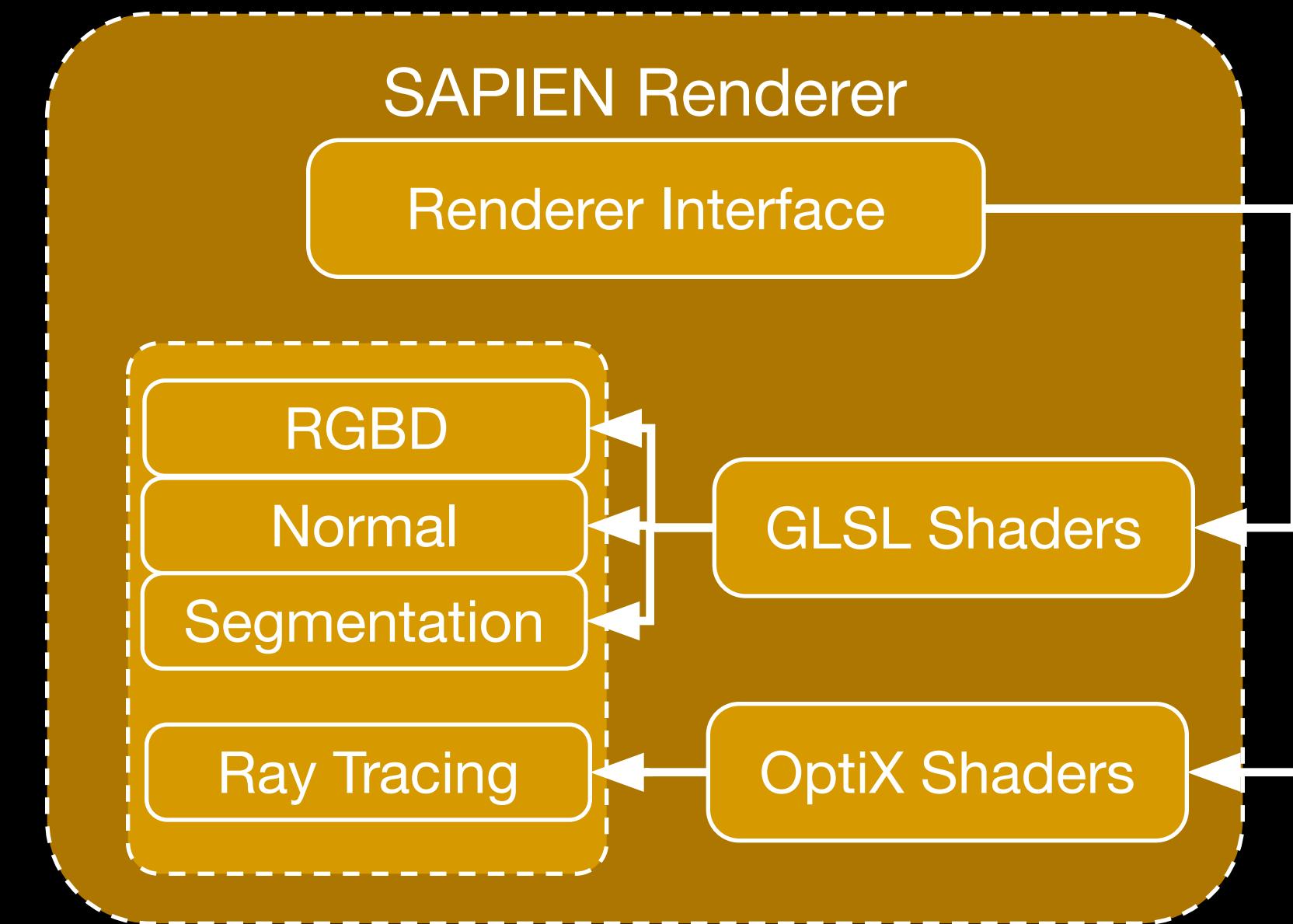
SAPIEN Renderer



SAPIEN Renderer

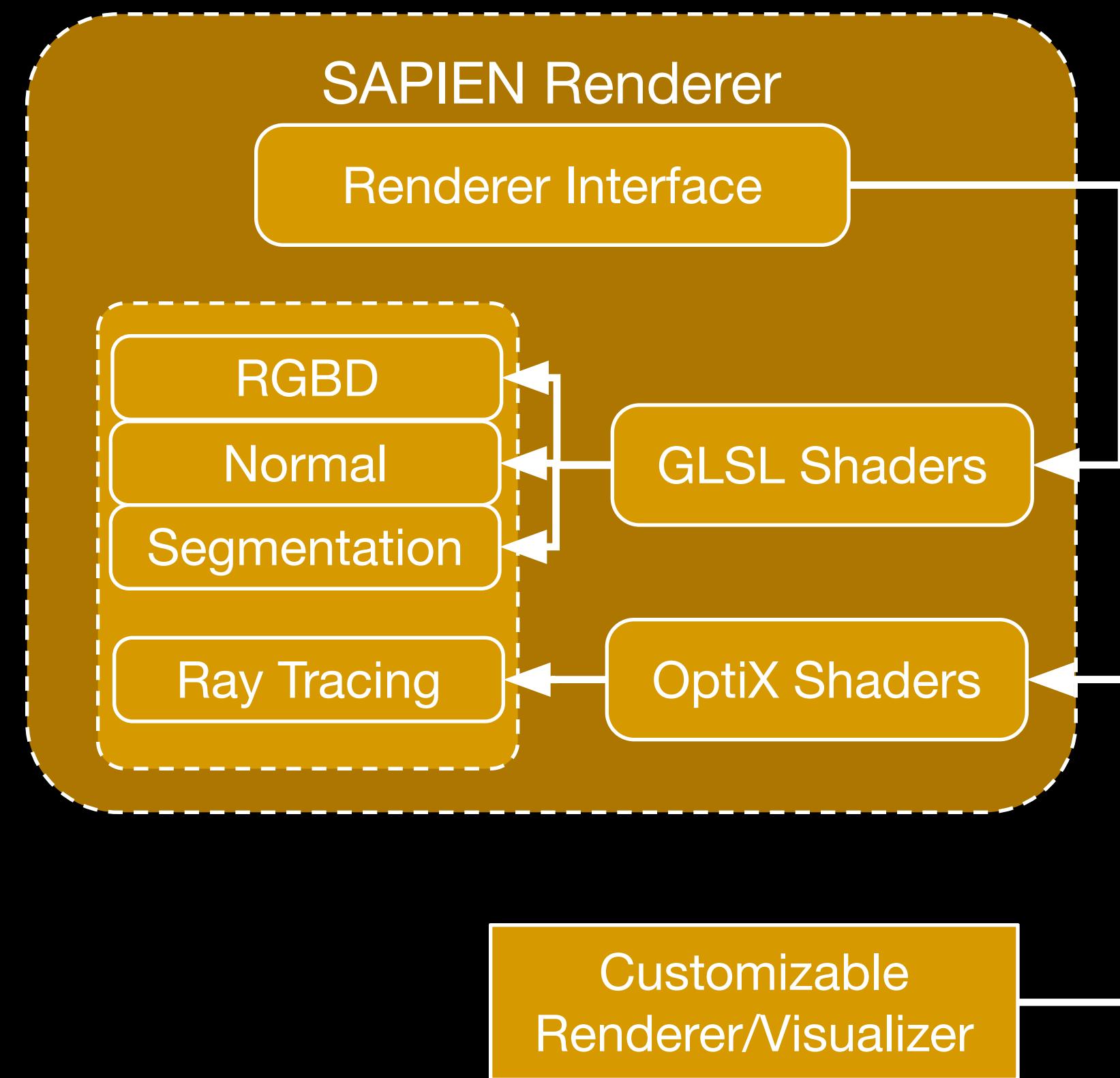


SAPIEN Renderer



**4 spp, 3 bounce, 512x512
OptiX denoise, < 20 FPS**

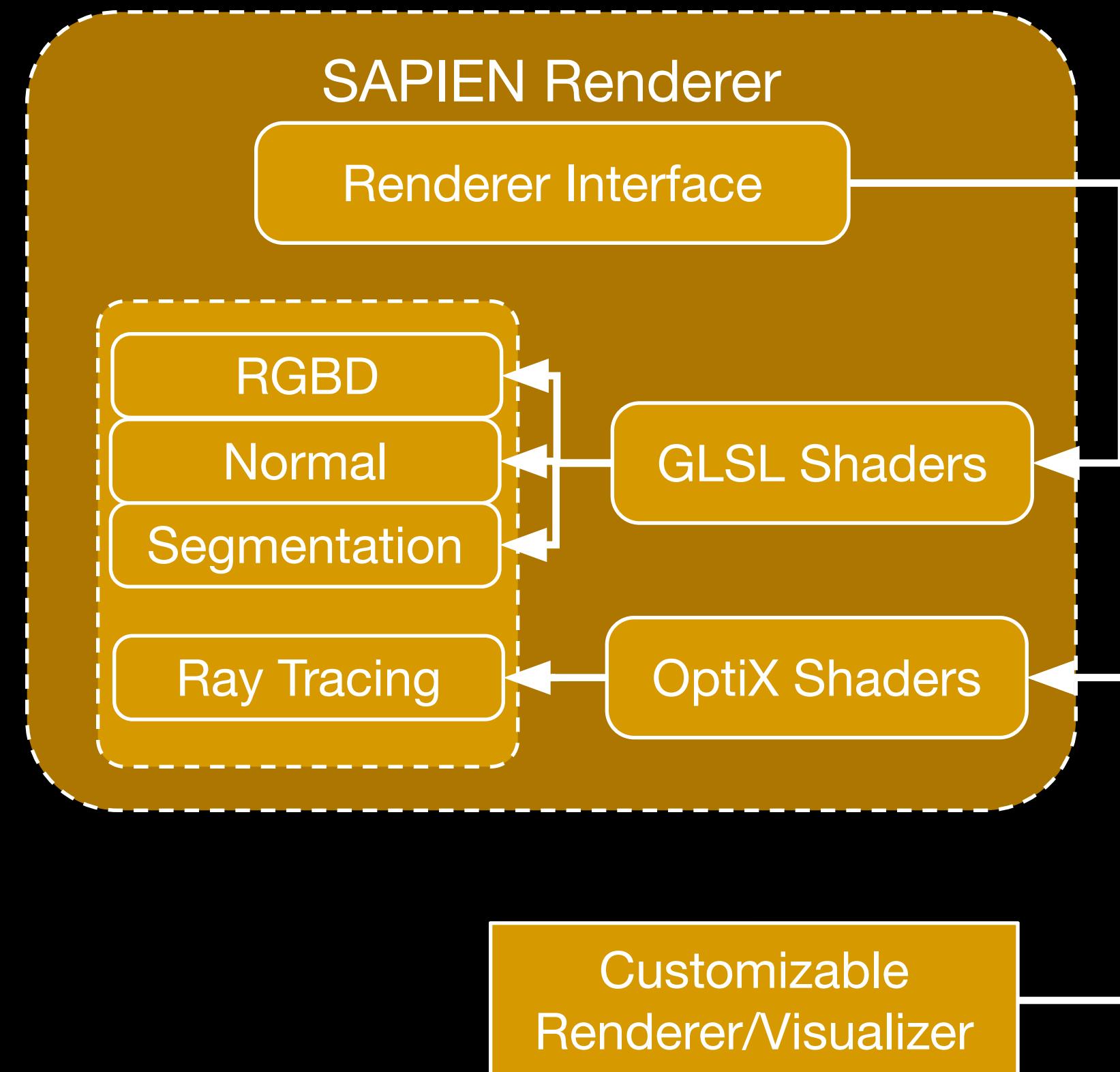
SAPIEN Renderer



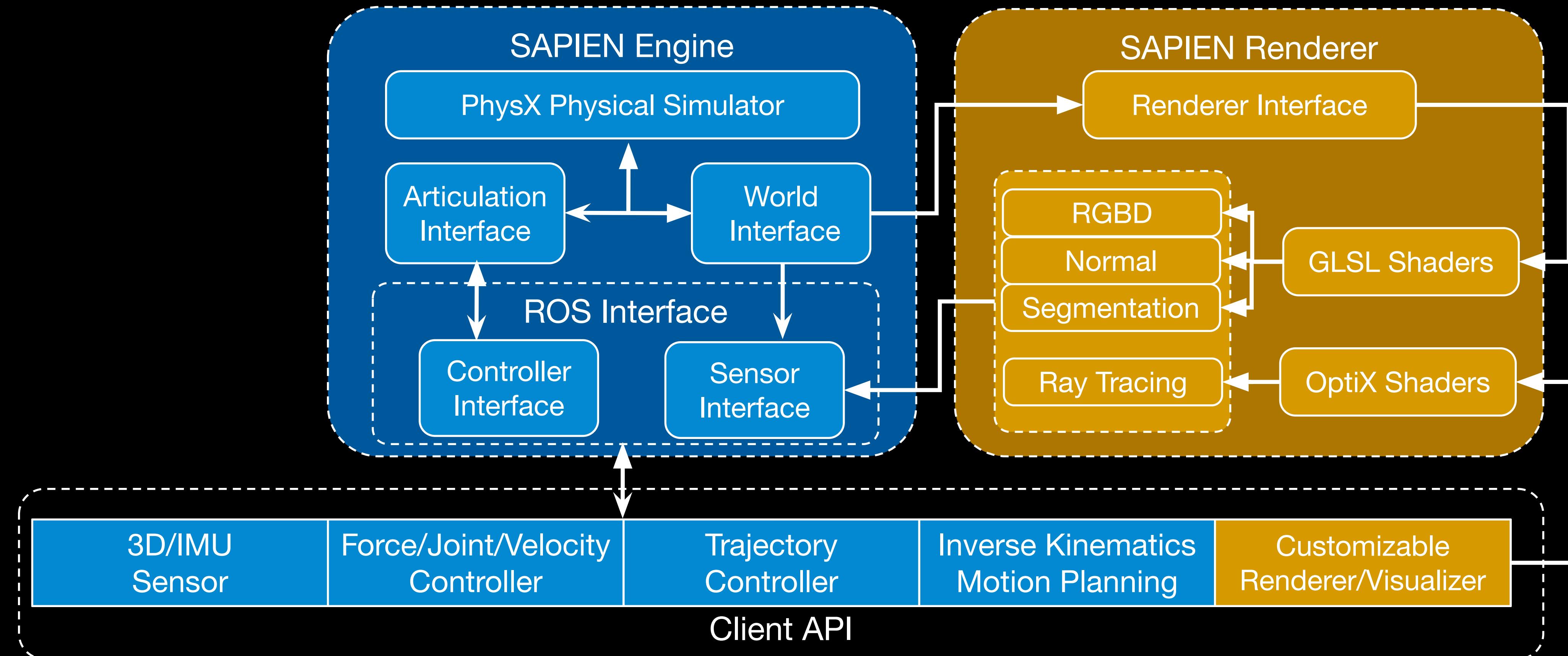
SAPIEN Renderer

Considerations

- **Needs to run at real time (~100 FPS)**
 - **Rasterizer**
 - **Ray tracer denoise? (1spp max)**
 - **GPU-CPU transfer? (>10 ms)**
 - **Prebaking? (interaction)**

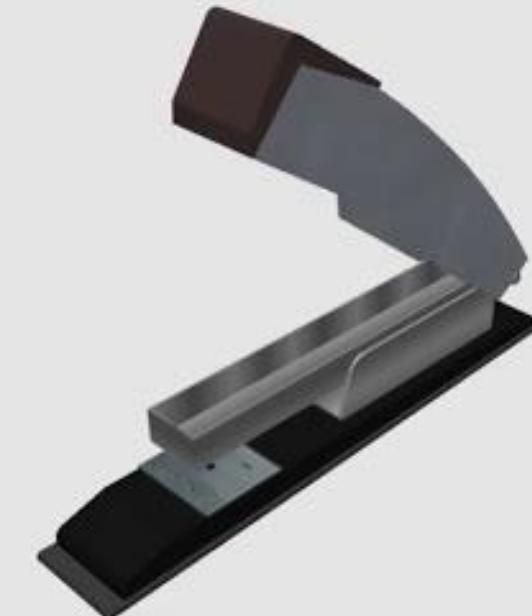
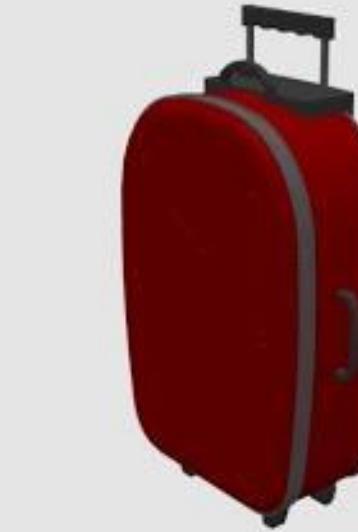


SAPIEN Renderer



SAPIEN Asset





PartNet-Mobility Annotation



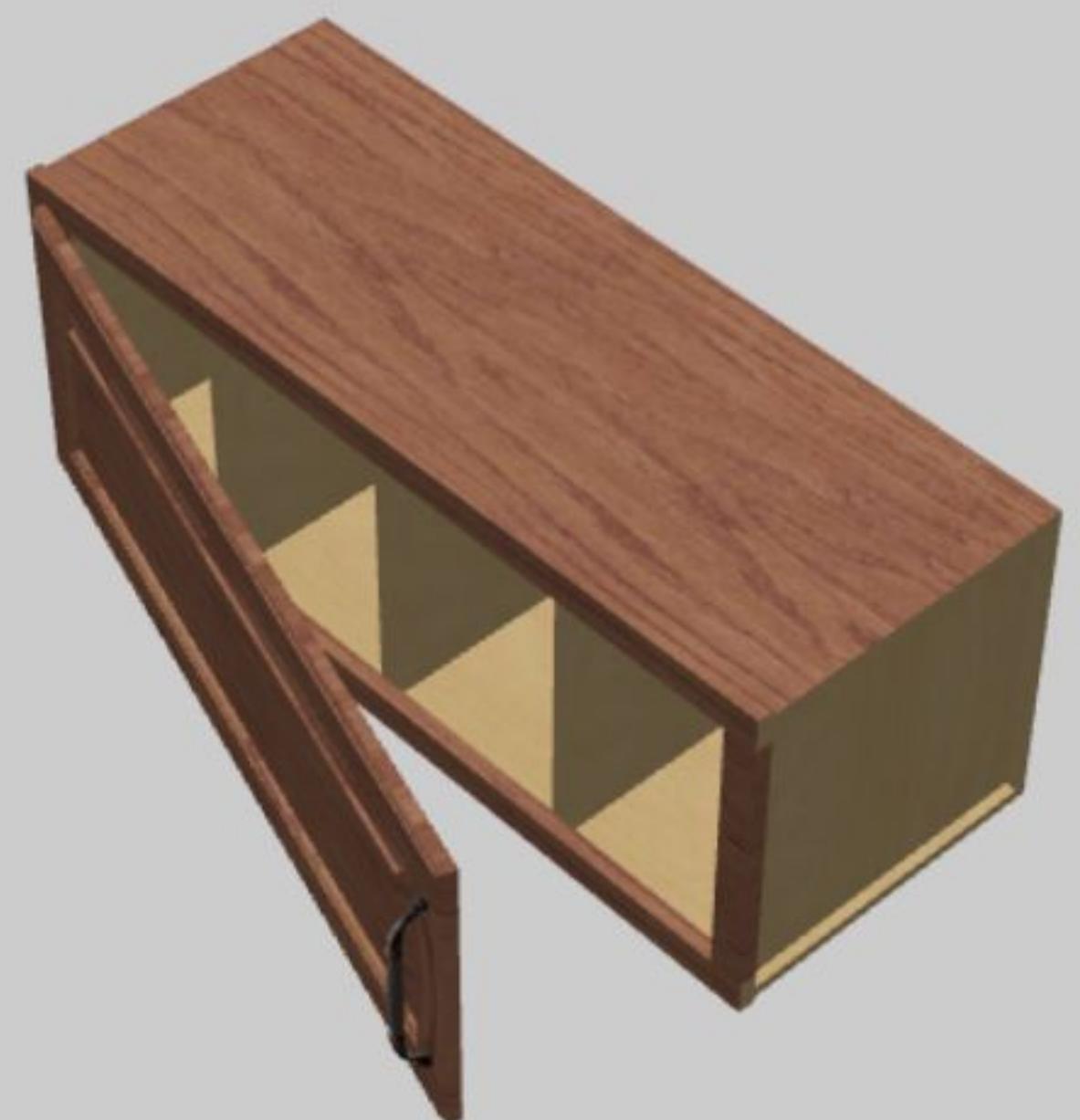
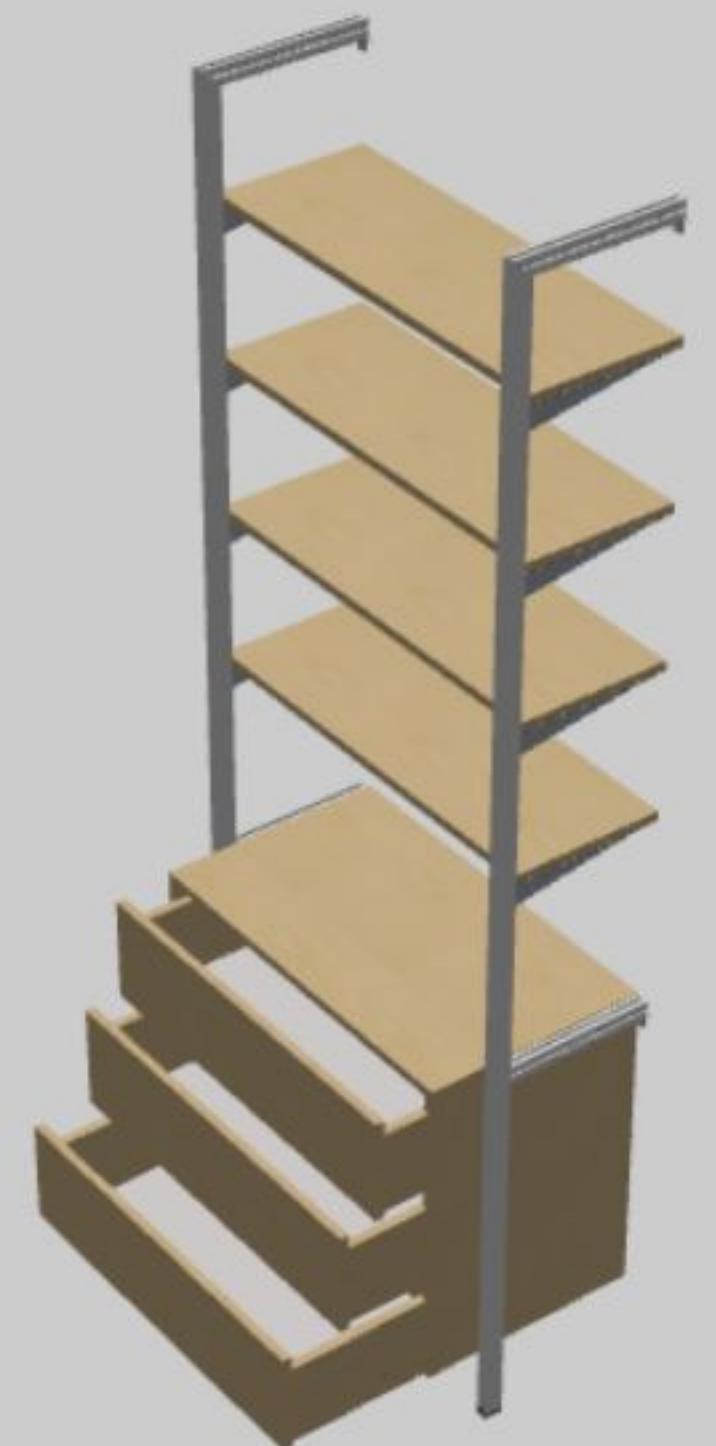
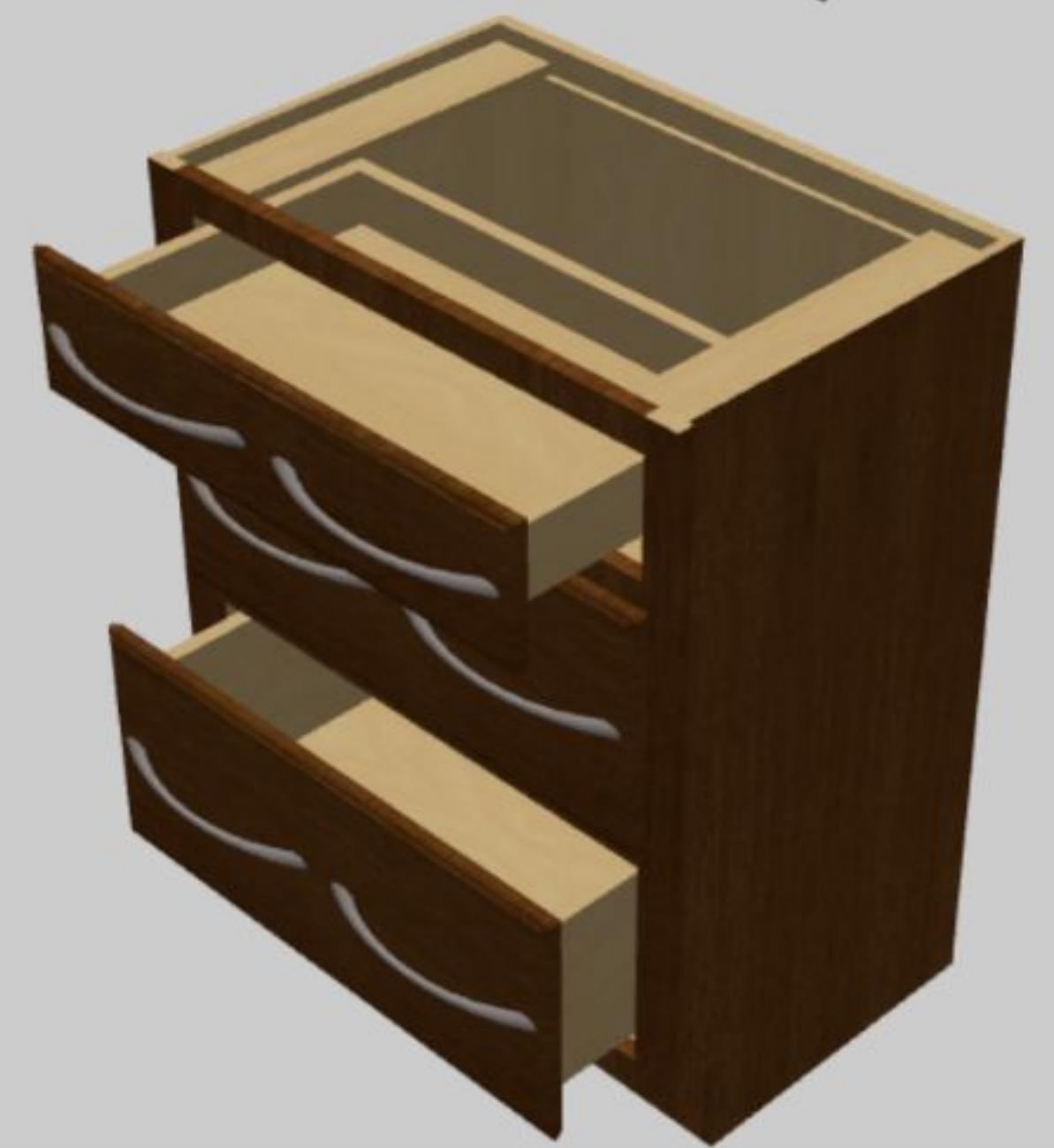
Shape cleaning
(ShapeNet, 3D Warehouse)

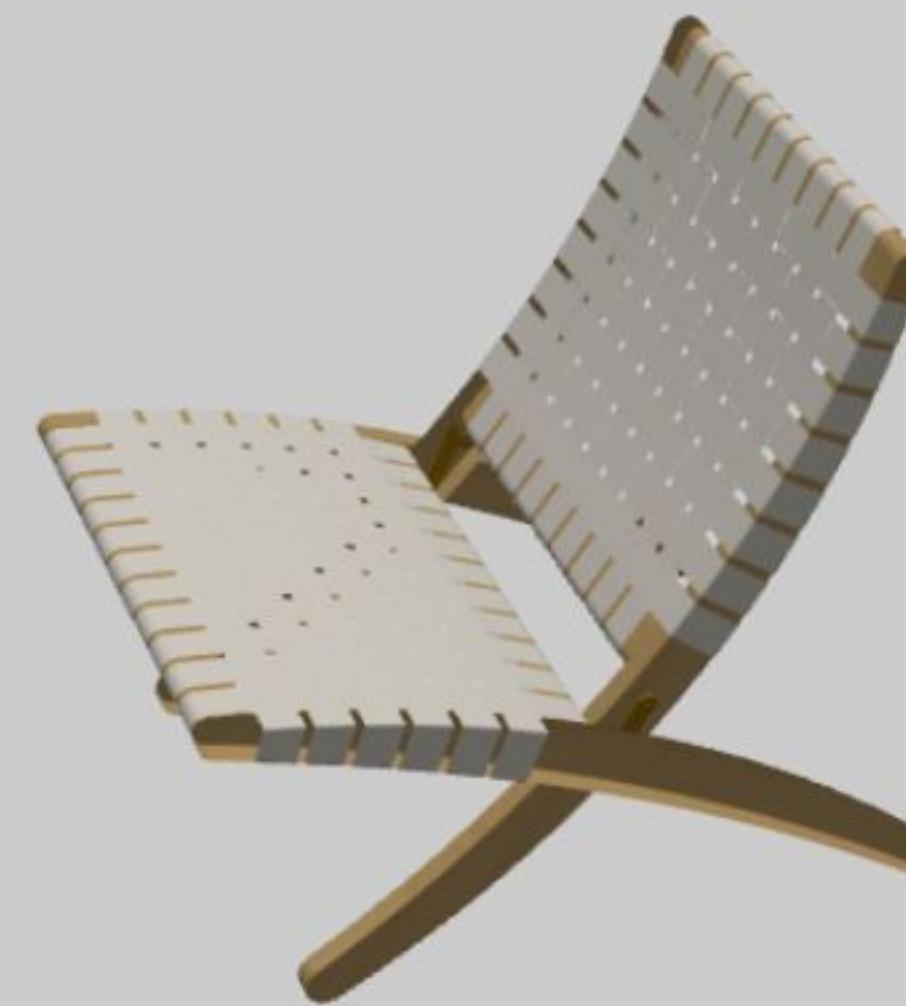
A screenshot of the Part Hierarchy (AOG) interface. On the left, the 'Part Hierarchy' panel shows a tree structure of a chair's components. In the center, the 'Q & A' panel displays a question about the 'Back Frame' and its sub-components like 'Back Frame Vertical Bar'. On the right, the '3D Preview' panel shows a 3D model of a chair with colored parts (purple, red, green) corresponding to the annotations. Below these are '3D Interface', 'Model Screenshots', and 'Part Definition' sections.

Part segmentation
(PartNet)

A screenshot of the Motion annotation (PartNet-Mobility) interface. It shows a 'Part Tree' on the left with checked items: 'FoldingChair', 'seat', and 'base_body'. The main area contains 'Question 1 : Subtree Mobility' with instructions and a 'Motion Tree' diagram. The diagram shows a folding chair with various parts highlighted in color (green, blue, red) to indicate their mobility levels relative to each other.

Motion annotation
(PartNet-Mobility)







SAPIEN Asset

SAPIEN Asset

PartNet-Mobility Dataset

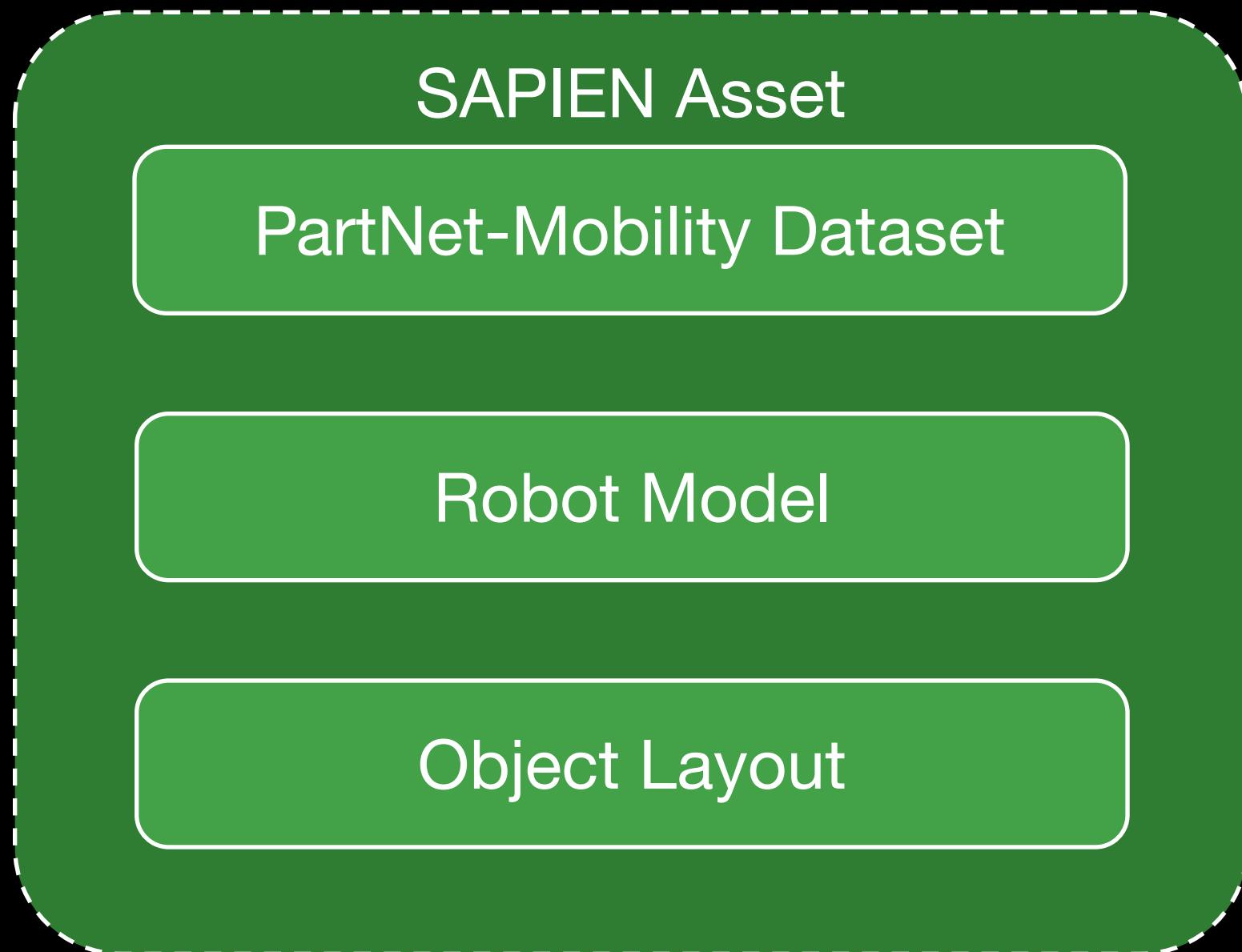
SAPIEN Asset

SAPIEN Asset

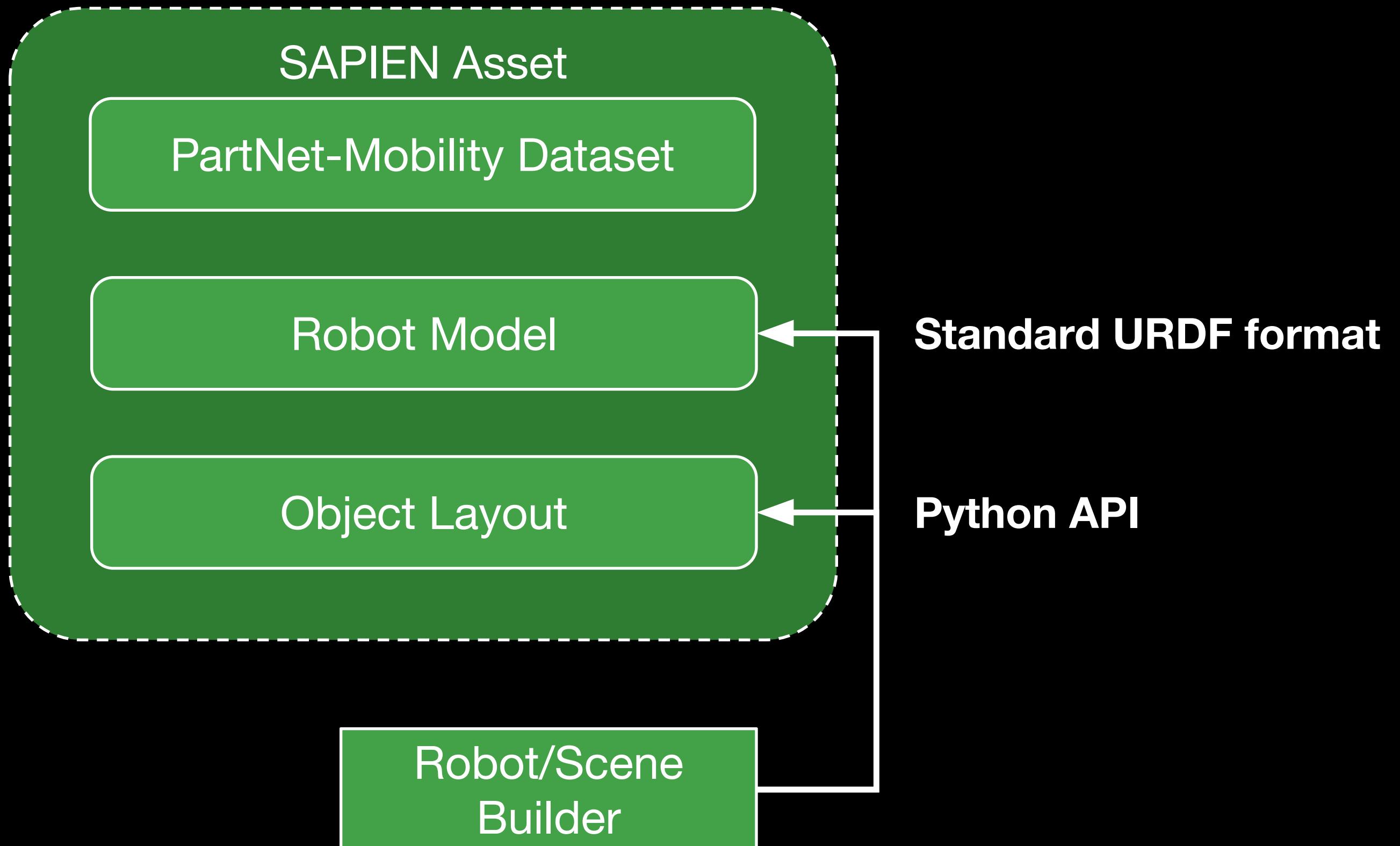
PartNet-Mobility Dataset

Robot Model

SAPIEN Asset

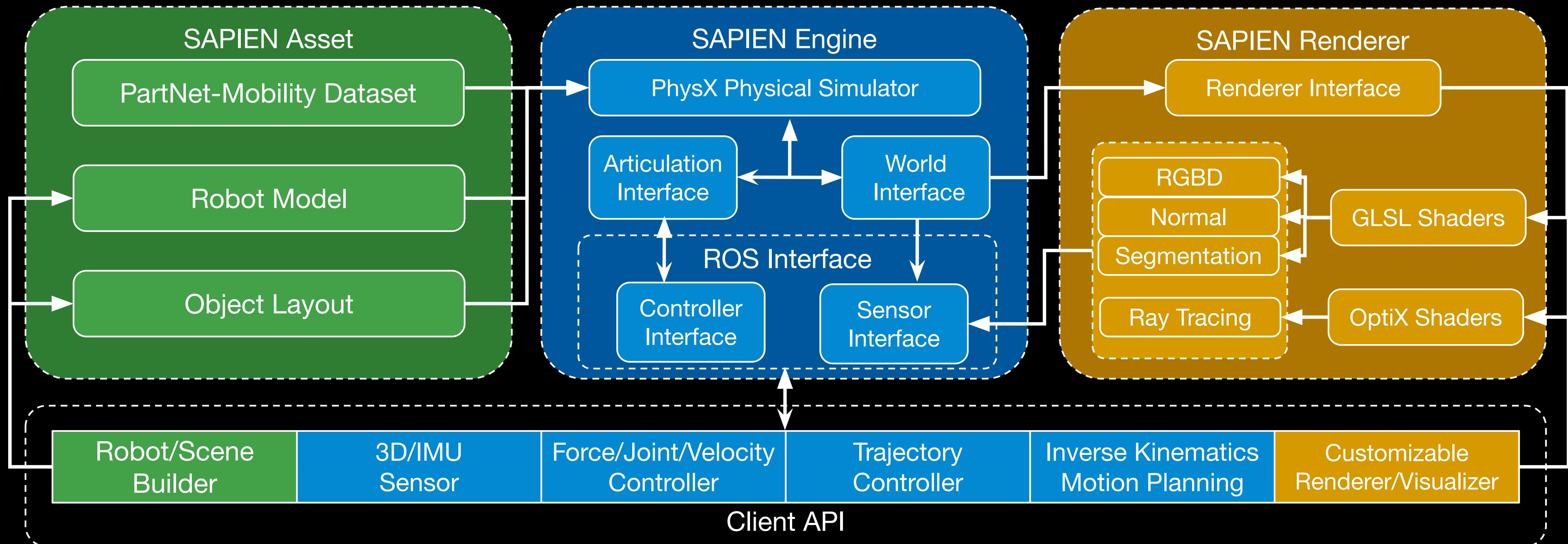


SAPIEN Asset



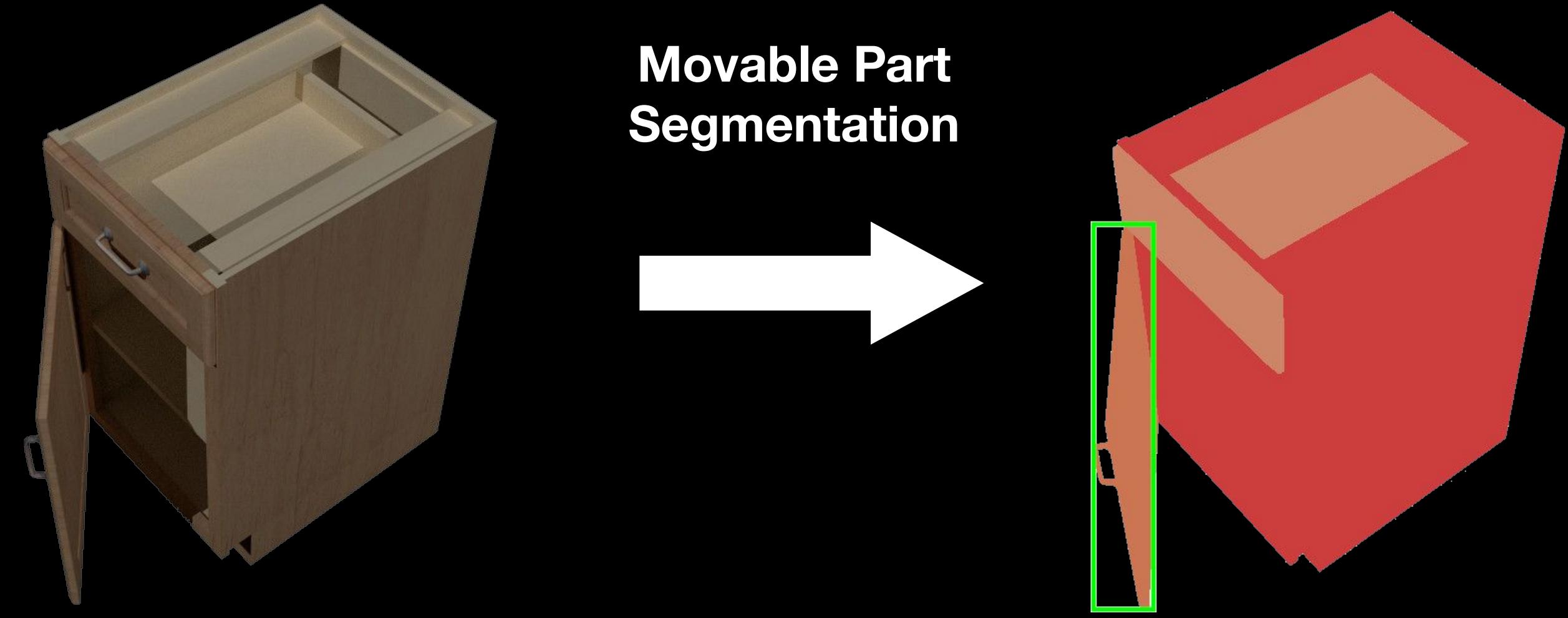
```
<robot name="panda" xmlns:xacro="http://www.ros.org/wiki/xacro">
  <link name="panda_link0">
    <visual>
      <geometry>
        <mesh filename="franka_description/meshes/visual/link0.dae"/>
      </geometry>
    </visual>
    <collision>
      <geometry>
        <mesh filename="franka_description/meshes/collision/link0.stl"/>
      </geometry>
    </collision>
  </link>
  <link name="panda_link1">
    <visual>
      <geometry>
        <mesh filename="franka_description/meshes/visual/link1.dae"/>
      </geometry>
    </visual>
    <collision>
      <geometry>
        <mesh filename="franka_description/meshes/collision/link1.stl"/>
      </geometry>
    </collision>
  </link>
  <joint name="panda_joint1" type="continuous">
    <safety_controller k_position="100.0" k_velocity="40.0" soft_lower_limit="-2.8973" soft_upper_limit="2.8973"/>
    <origin rpy="0 0 0" xyz="0 0 0.333"/>
    <parent link="panda_link0"/>
    <child link="panda_link1"/>
    <axis xyz="0 0 1"/>
    <limit effort="87" lower="-2.8973" upper="2.8973" velocity="2.1750"/>
  </joint>
  <link name="panda_link2">
    <visual>
      <geometry>
        <mesh filename="franka_description/meshes/visual/link2.dae"/>
      </geometry>
    </visual>
```

SAPIEN



Task Demonstrations

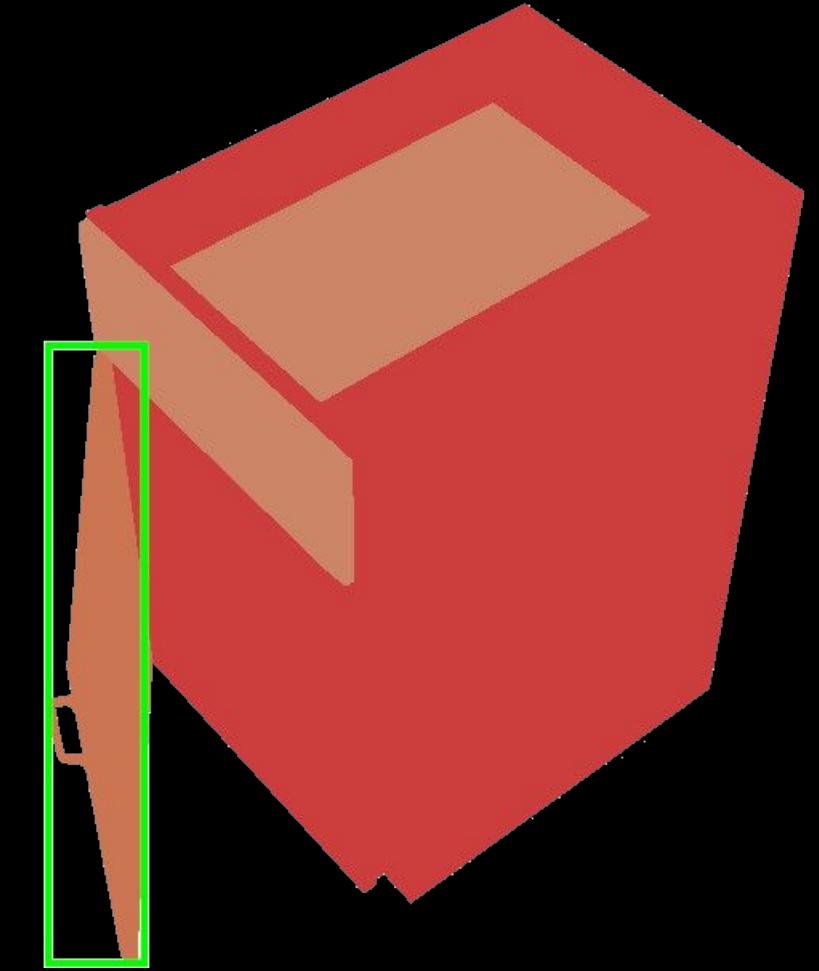
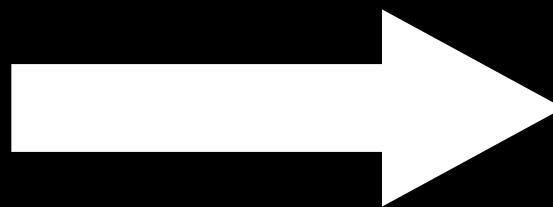
Task Demonstrations



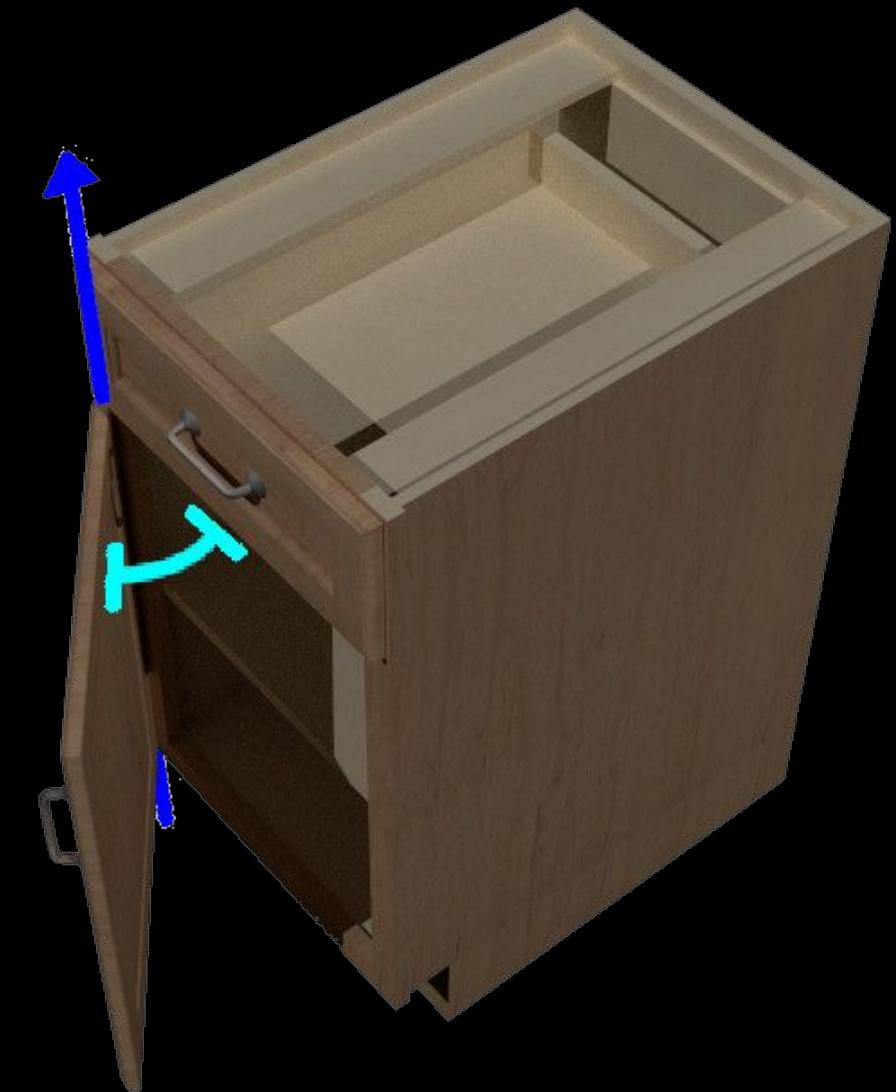
Task Demonstrations



Movable Part Segmentation



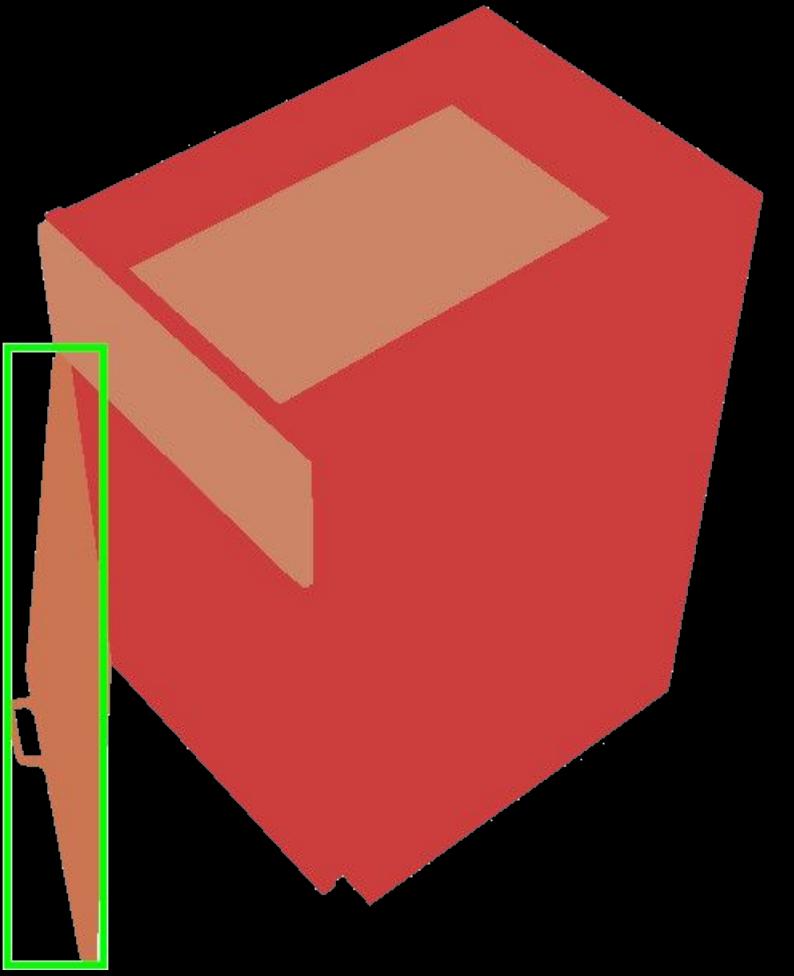
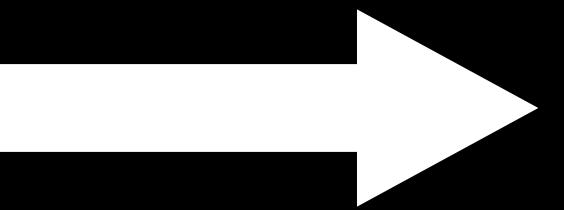
Motion Parameter Estimation



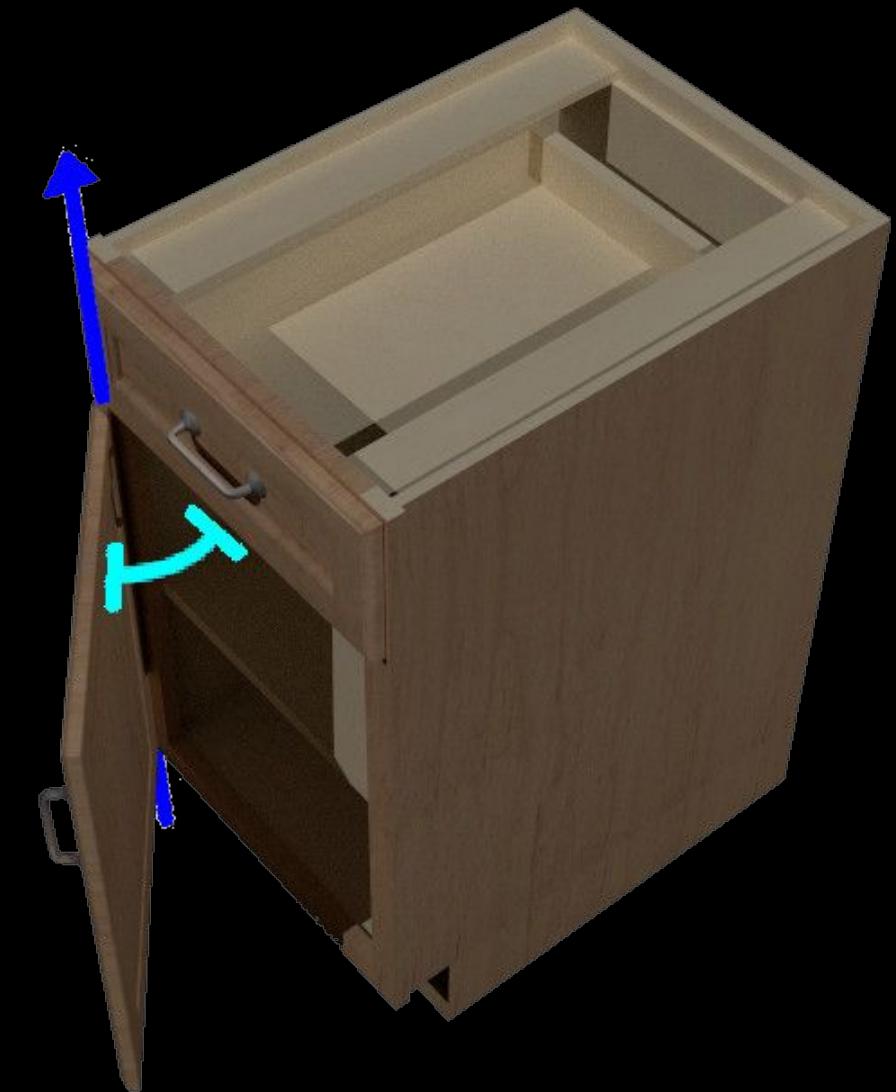
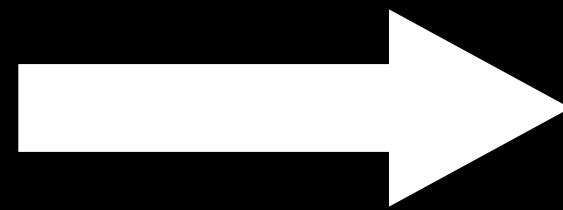
Task Demonstrations



Movable Part Segmentation



Motion Parameter Estimation

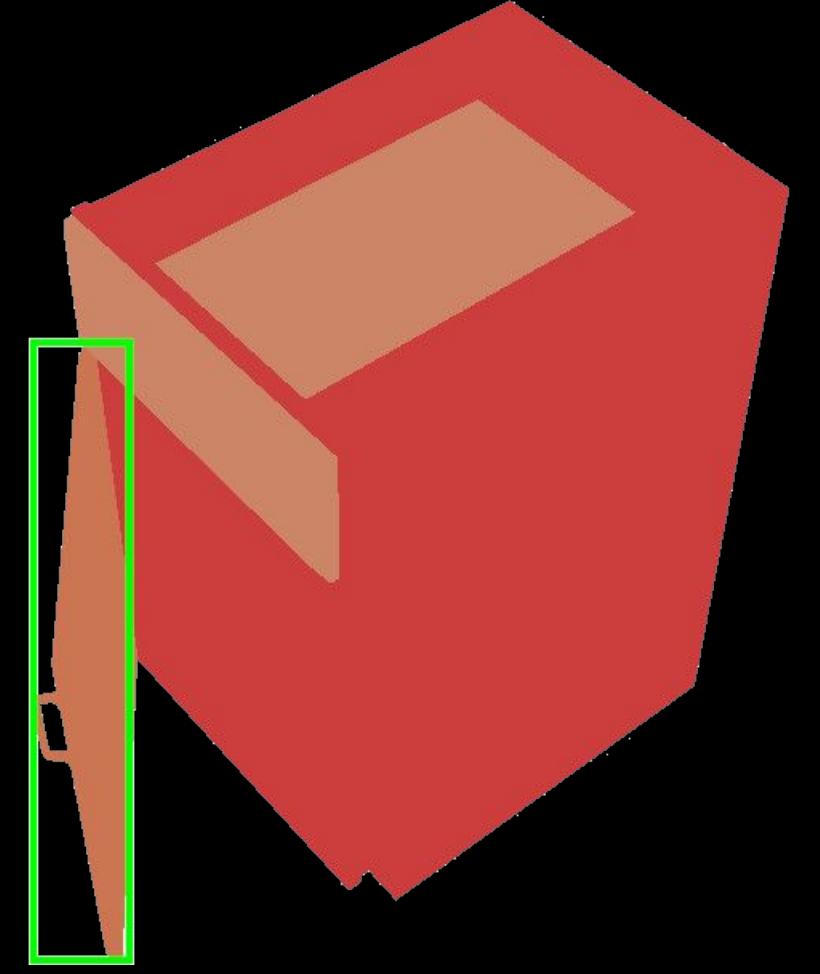
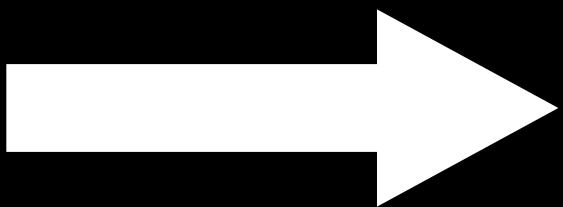


Part Manipulation

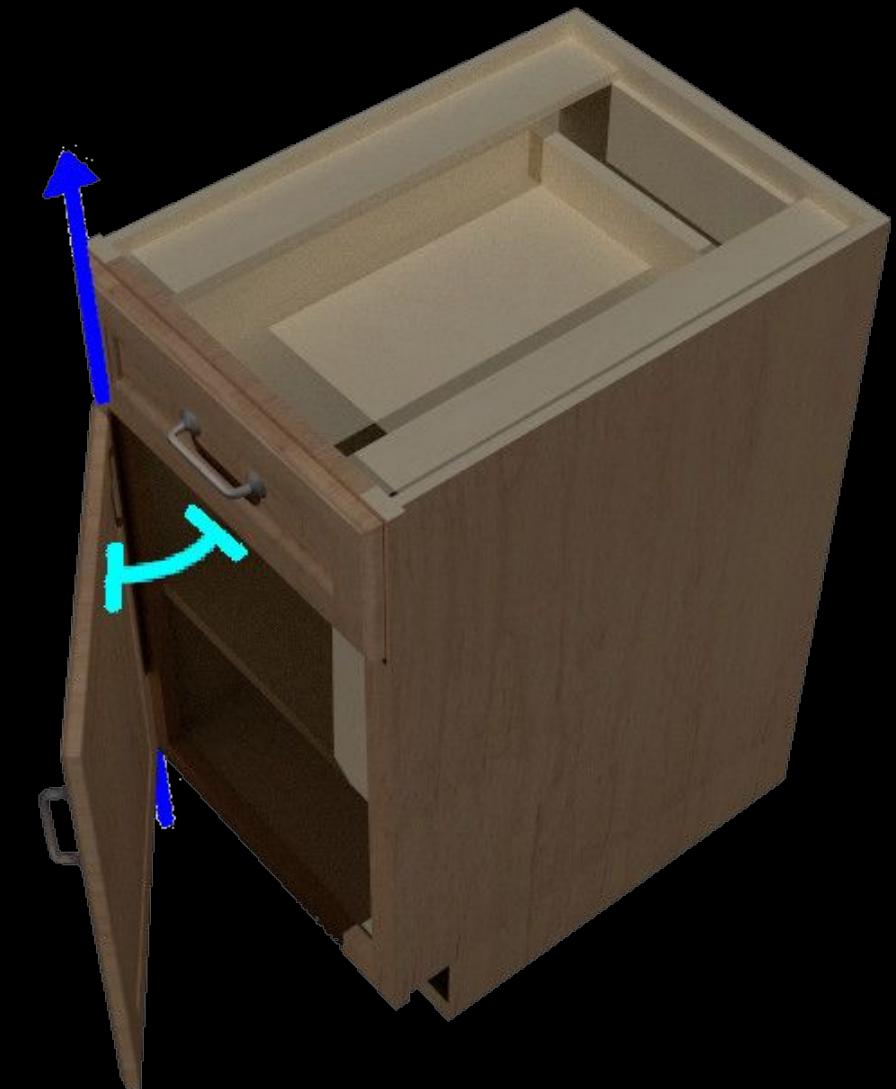
Task Demonstrations



Movable Part Segmentation



Motion Parameter Estimation



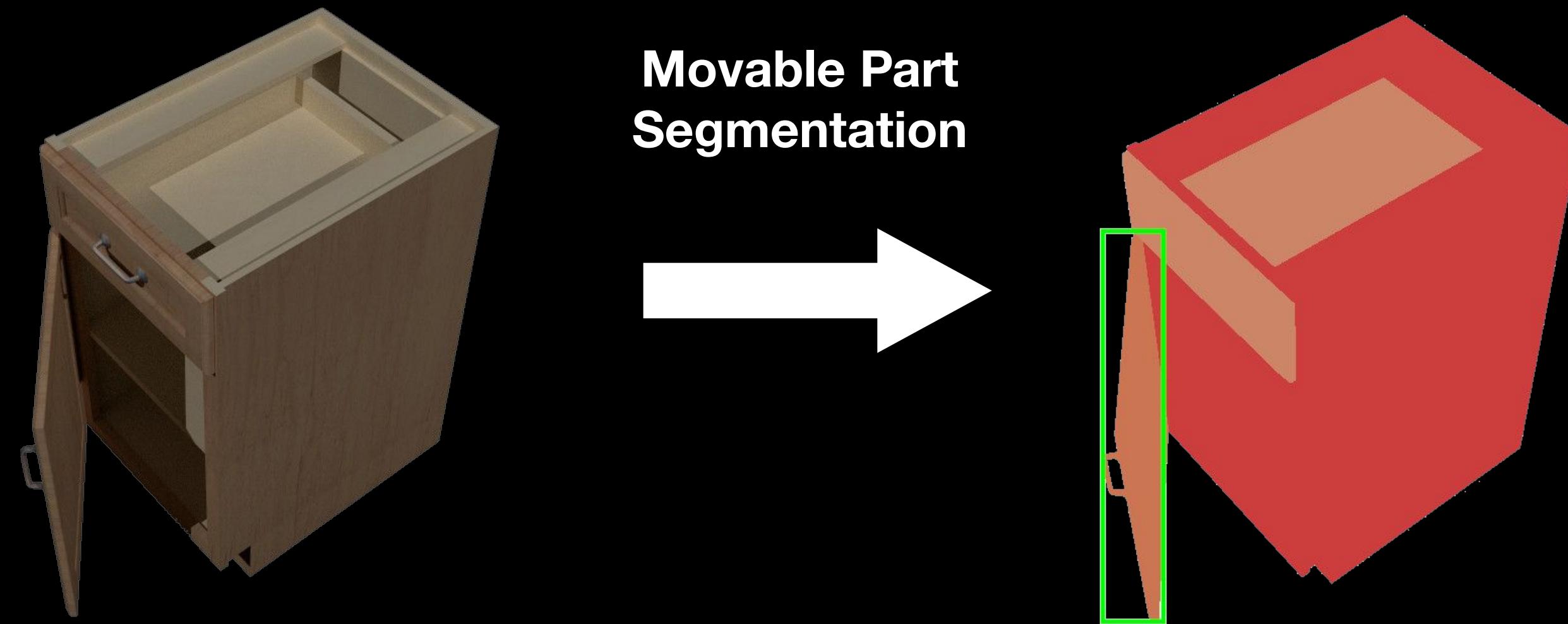
Long-horizon Planning



Part Manipulation



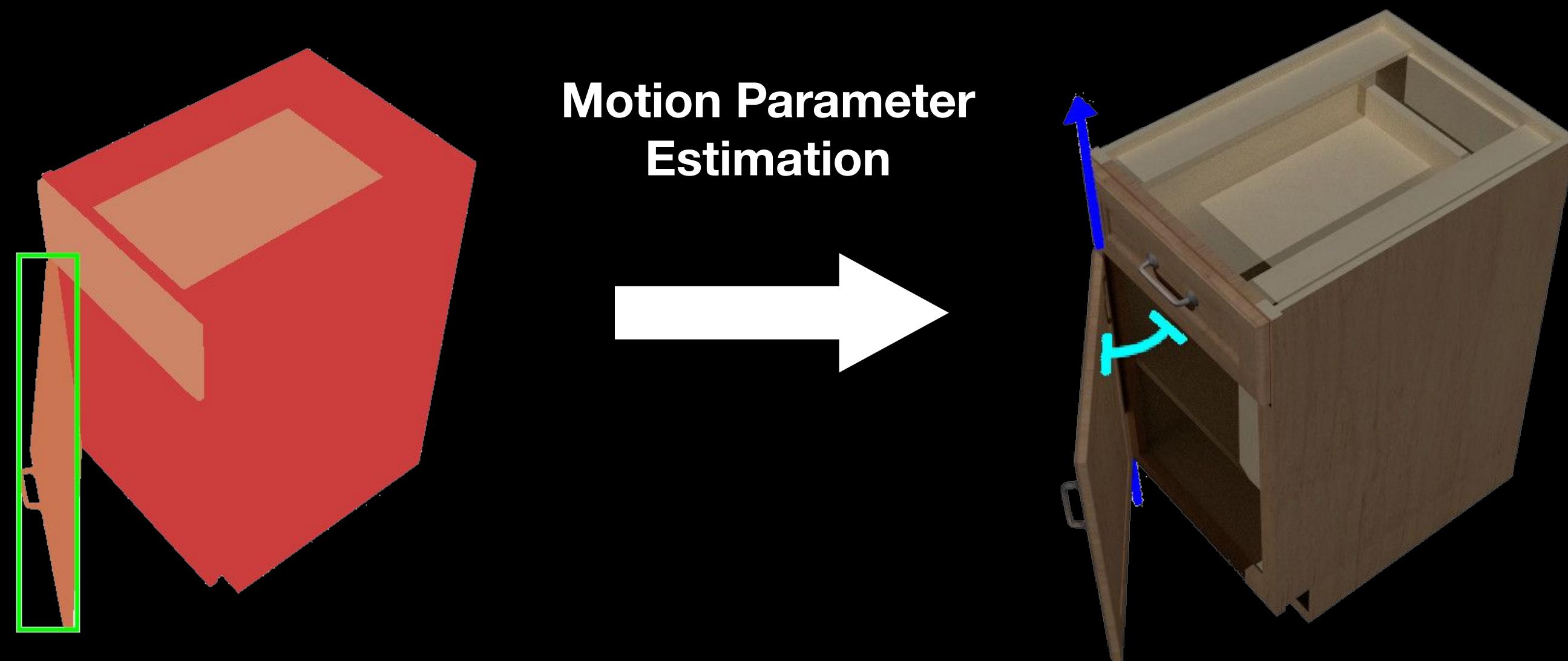
Task Demonstrations



**Movable Part
Segmentation**

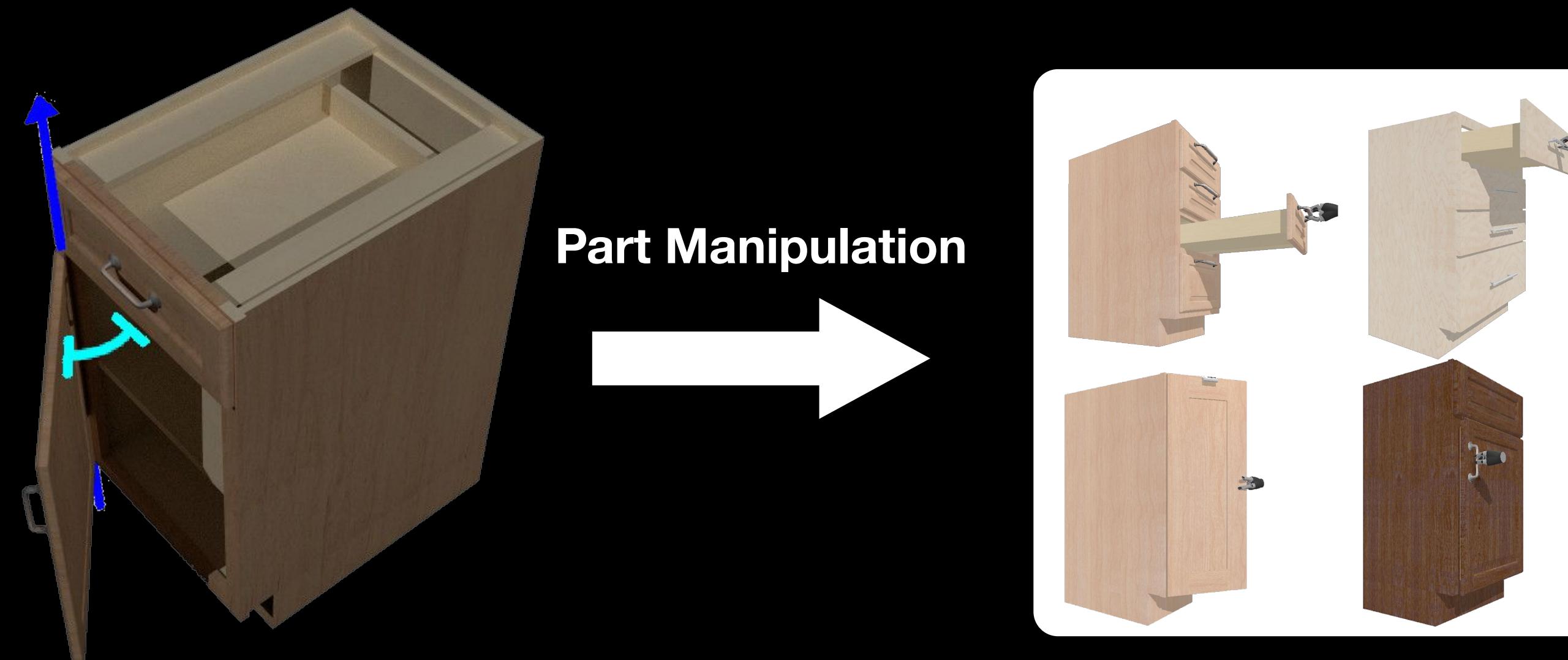
- Standard vision problem: detection and segmentation.

Task Demonstrations



- **New vision task**
- **Important for control tasks**
 - Given the motion parameters, we can use control methods to manipulate the parts

Task Demonstrations



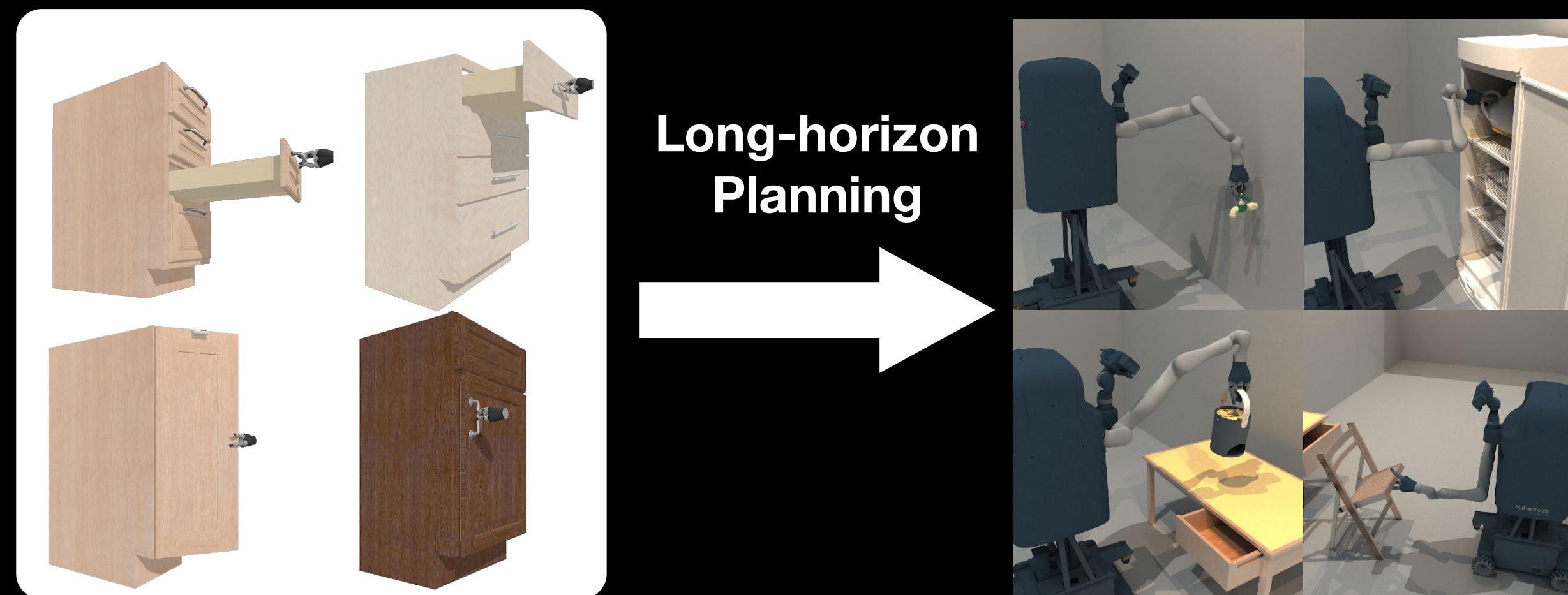
- **Manipulation and control**
 - Reinforcement learning
 - Imitation learning
 - robotics control

Task Demonstrations



- **Manipulation and control**
 - Reinforcement learning
 - Imitation learning
 - robotics control

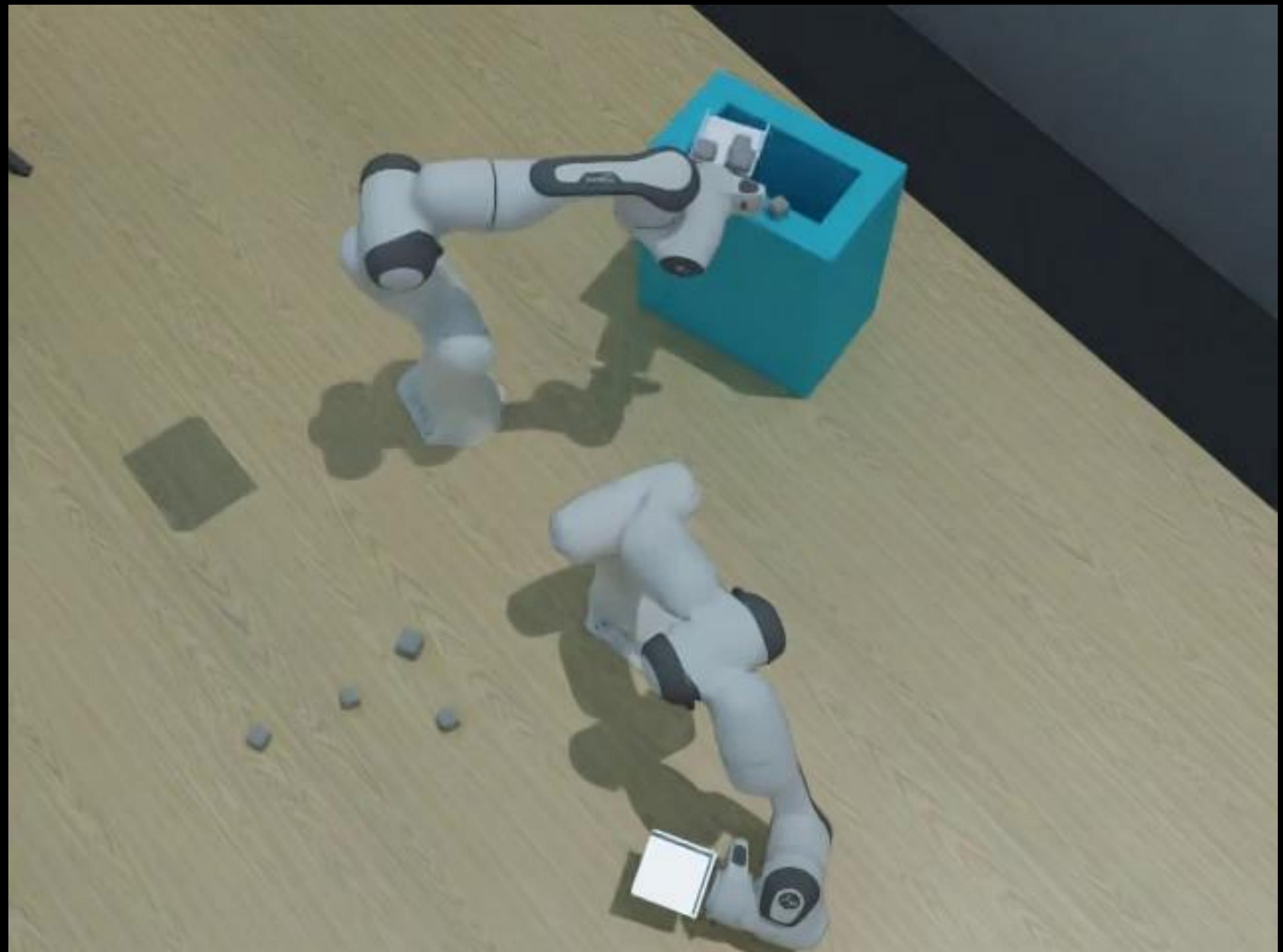
Task Demonstrations



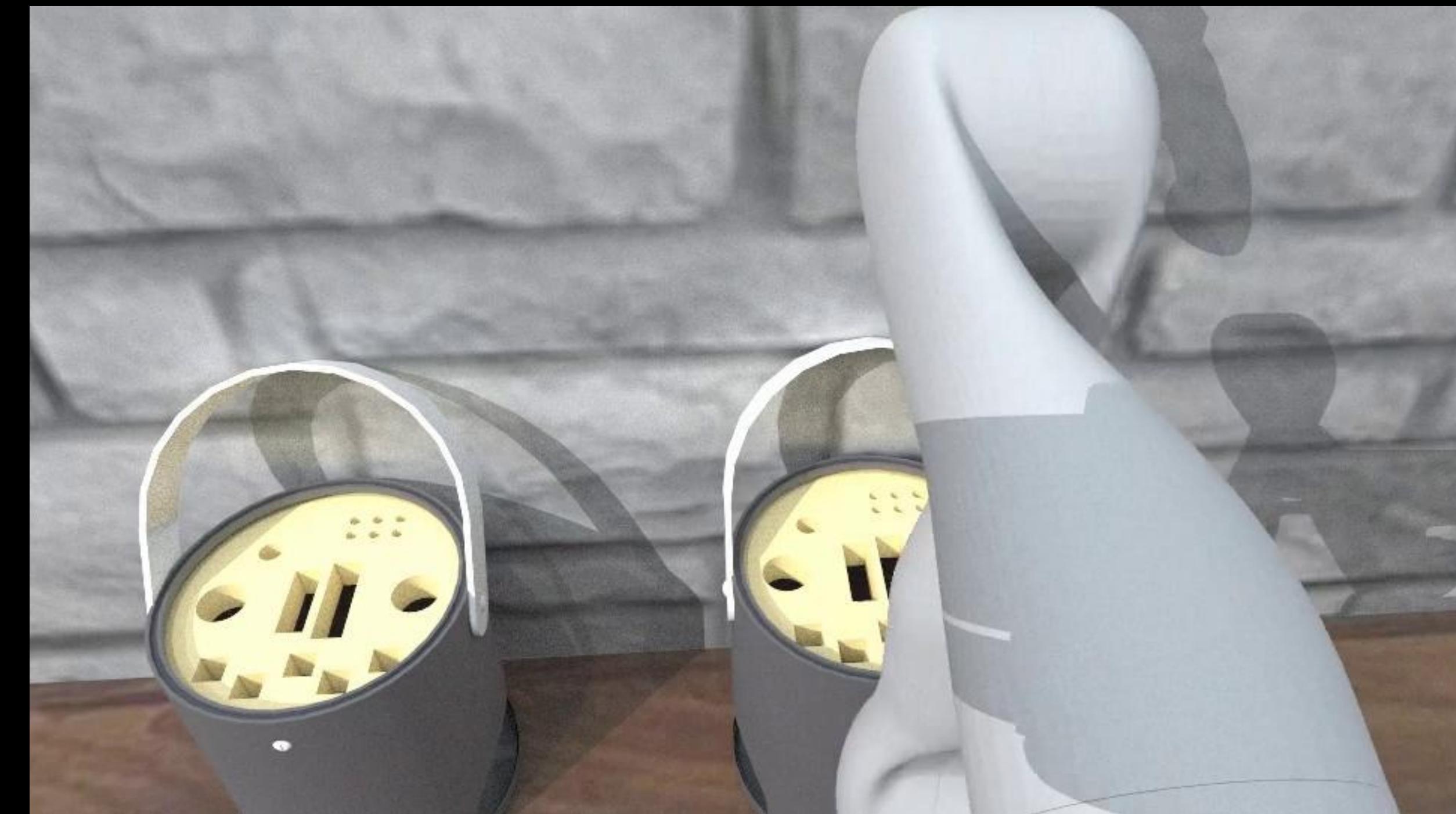
- **Planning**
 - Achieve meaningful tasks
 - “Home assistant”

Future Plans/Applications

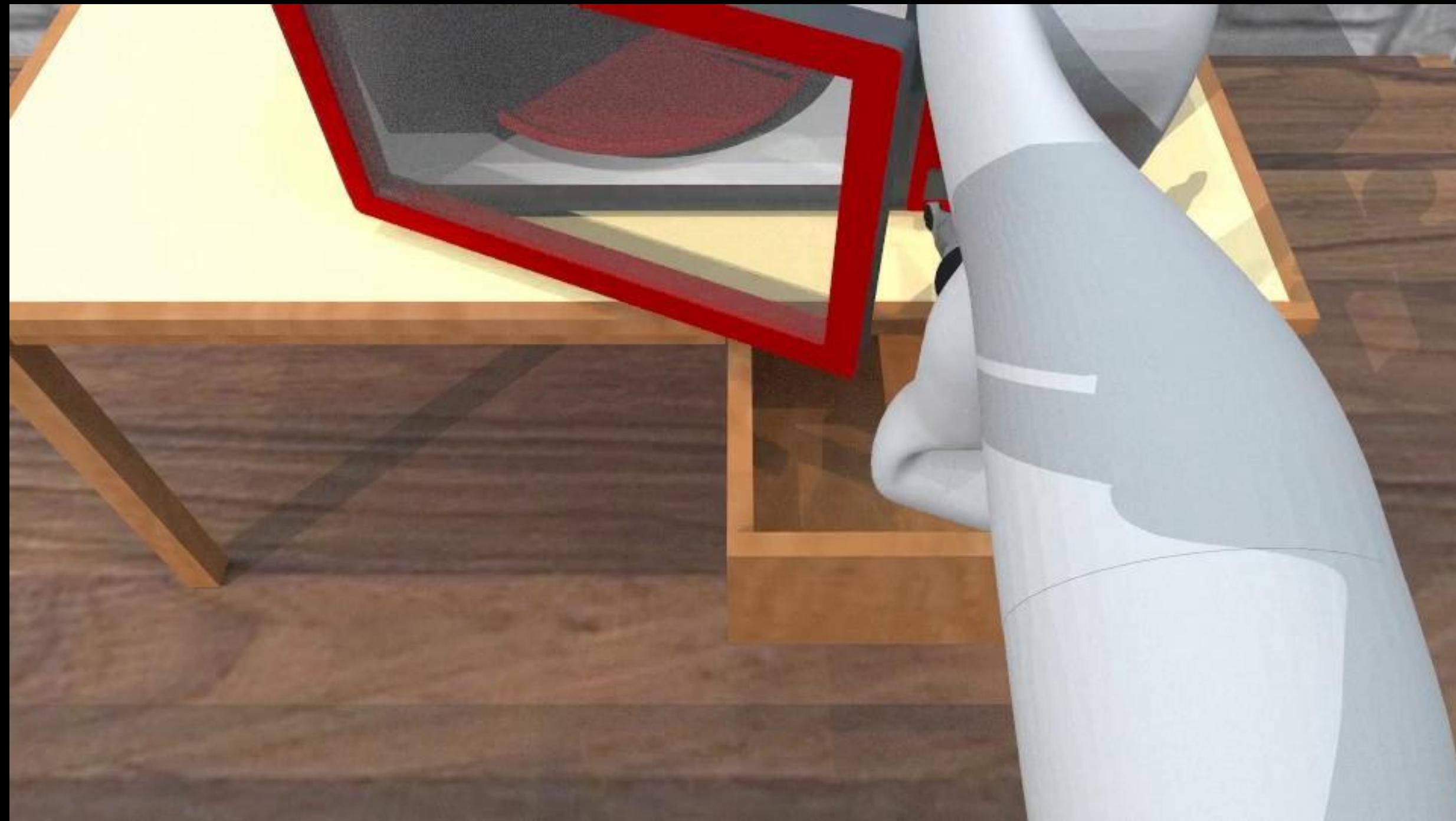
- Benchmark (SAPIEN Challenge)
 - Vision tasks
 - Manipulation tasks
- Education Platform
 - Control Algorithms
 - Robot Learning



Demos

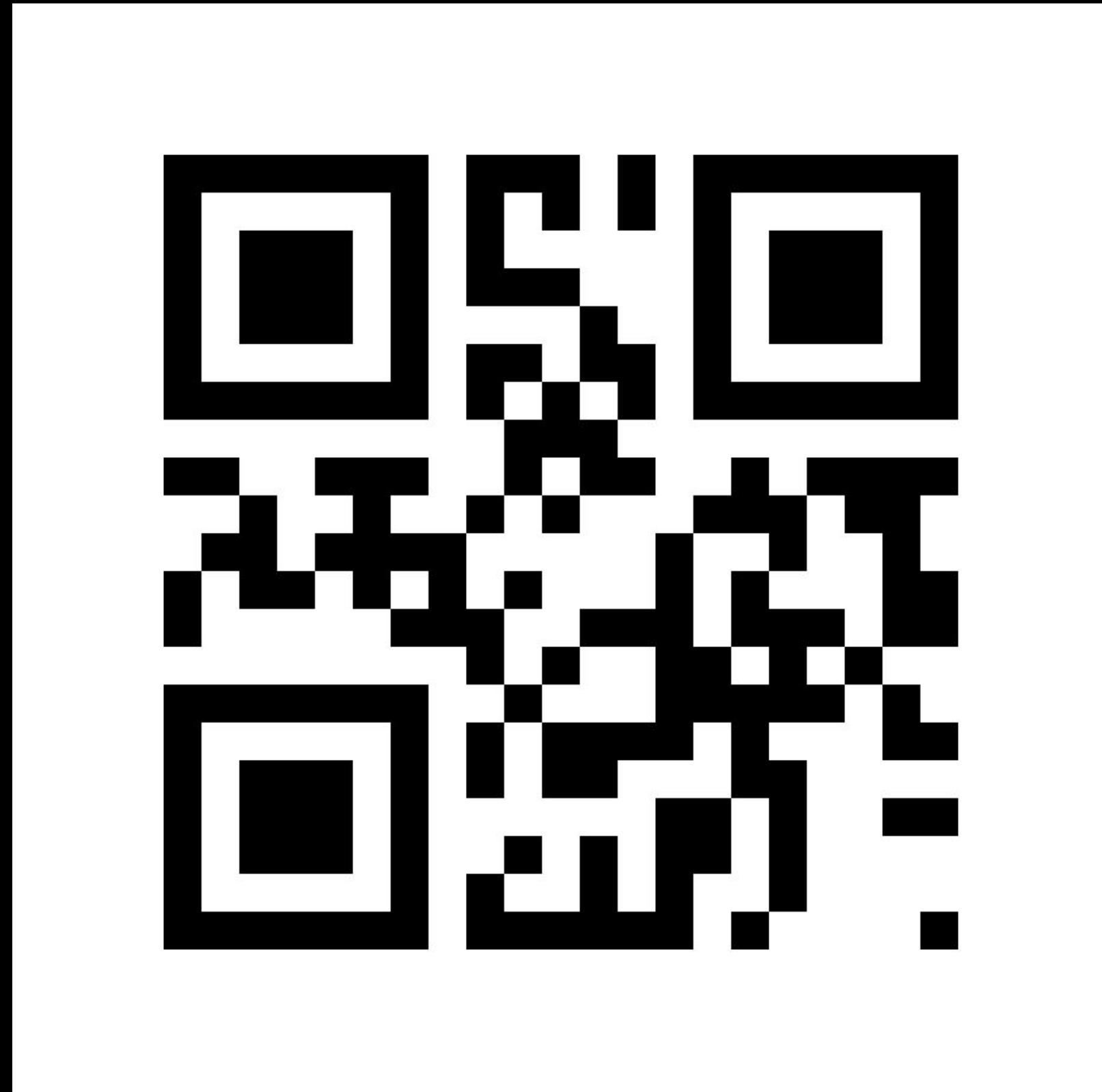


Demos



Summary

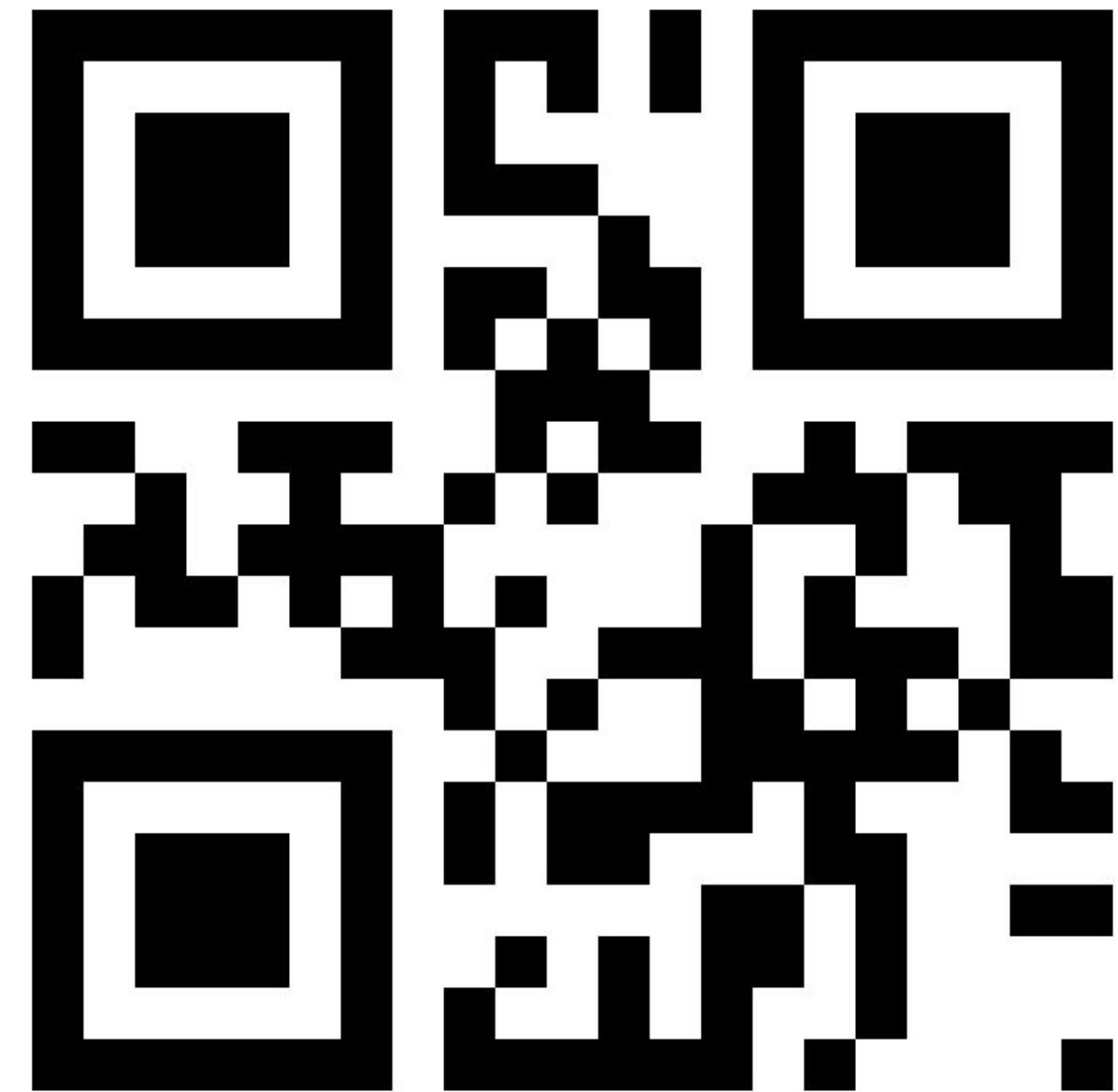
- SAPIEN provides simulation and rendering for vision and manipulation.
- SAPIEN hosts a large-scale dataset for articulated objects.
- SAPIEN is open source.
- pip install sapien



<https://sapien.ucsd.edu>

Thank you

- Questions?



<https://sapien.ucsd.edu>