

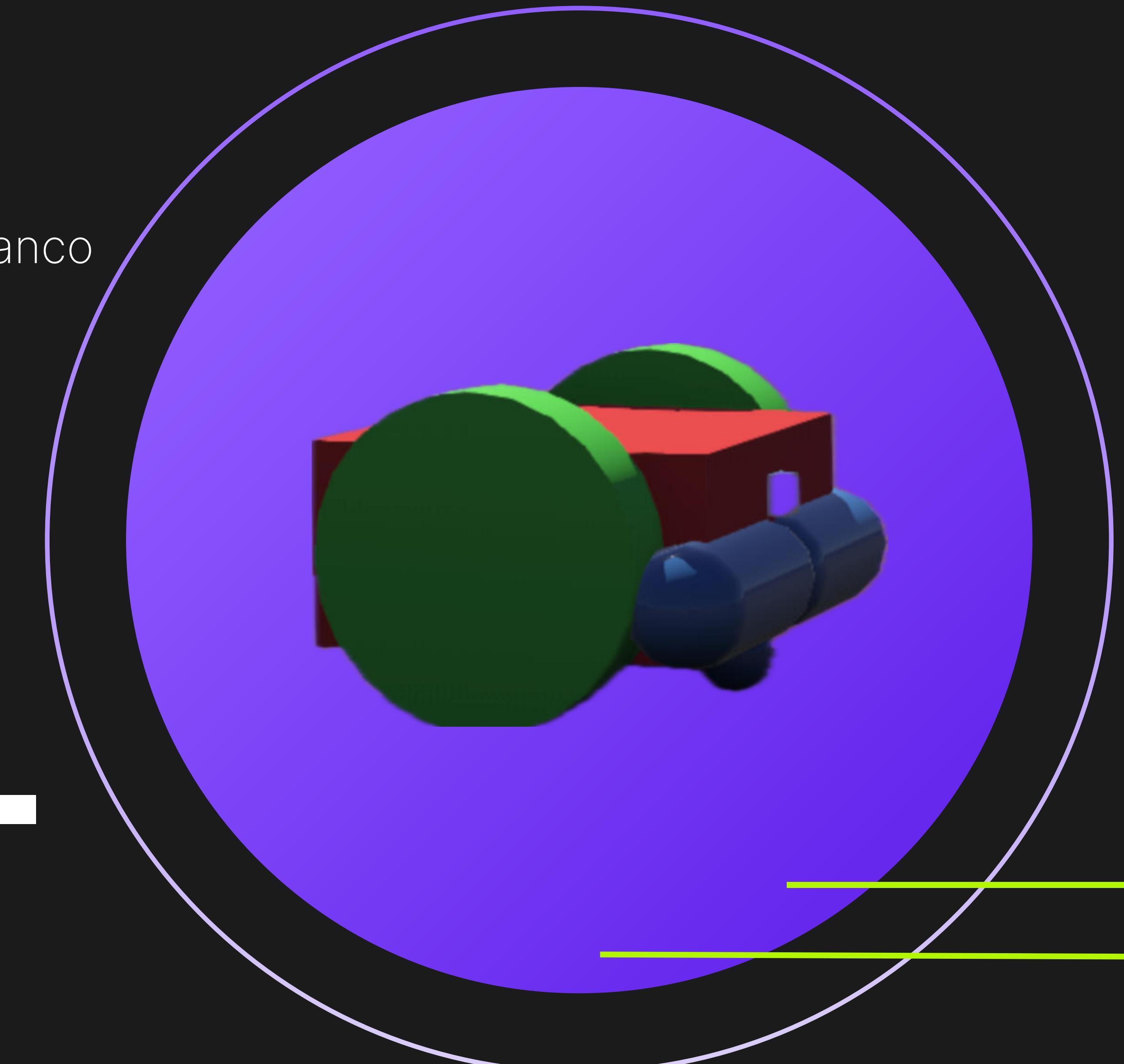
REINFORCEMENT LEARNING - Prof. Capobianco

...

A LOWCOST SIM2REAL



SAPIENZA
UNIVERSITÀ DI ROMA



INDEX

01 GOAL

02 SIM. ENV

03 SIM ROBOT

04 DETAILS

05 STACK

05 REAL ENV

06 REAL ROBOT

07 ALL TOGETHER

07 RESULT

08 IMPROVEMENT

09 CONCLUSION



GOAL

RL IS AMAZING.

Everybody knows it.

GOAL

RL IS AMAZING.

BUT, *does it works in real world?*



GOAL



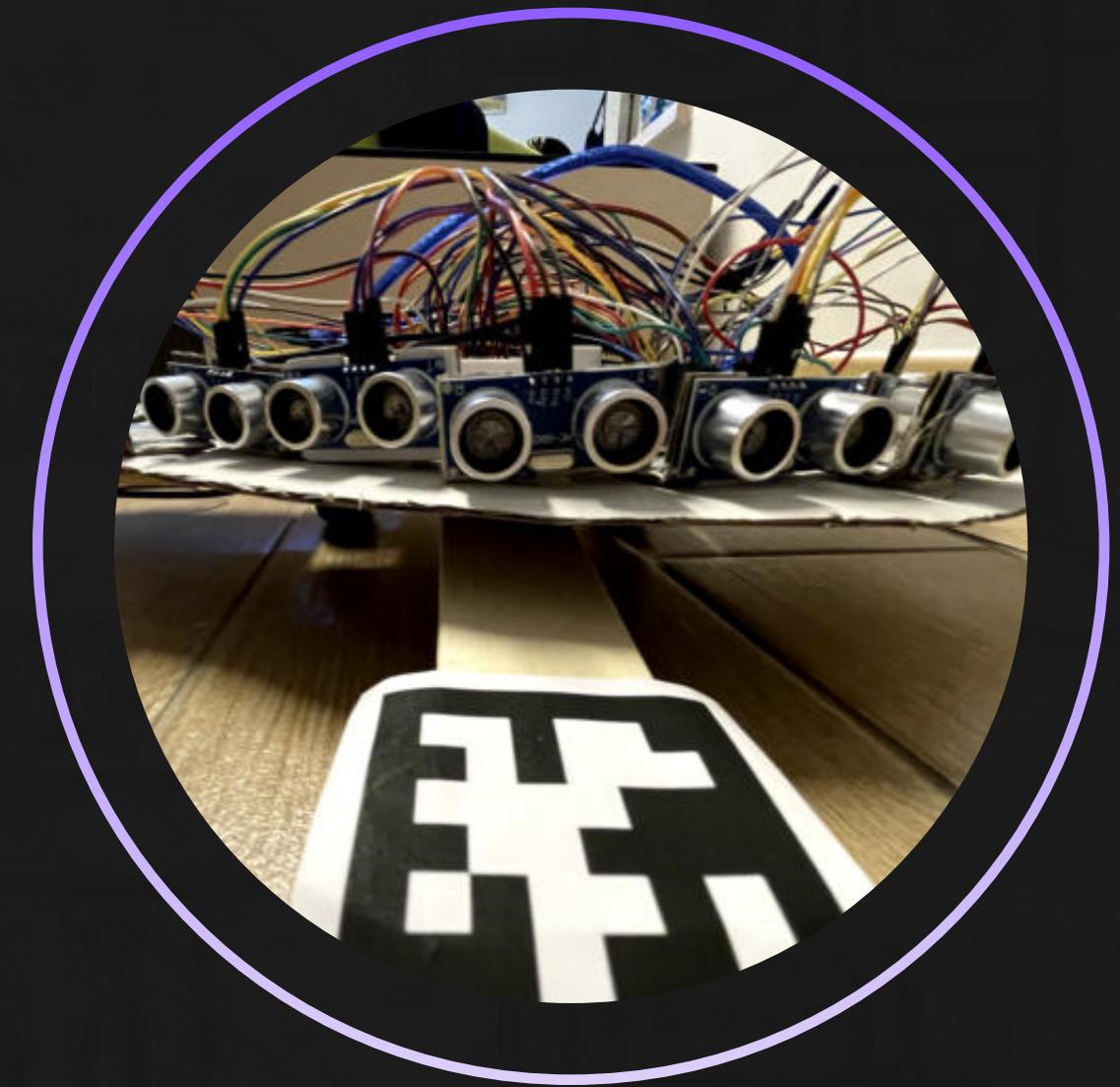
Simulation



GOAL



Simulation



Real

GOAL



Simulation

Real

GOAL

SIM2REAL

PHYSICS & ENV

Physic in simulation not equal to the real world physic, match the environment requires special places.

MECHANICS

Match the simulated mechanics is really a challenge.

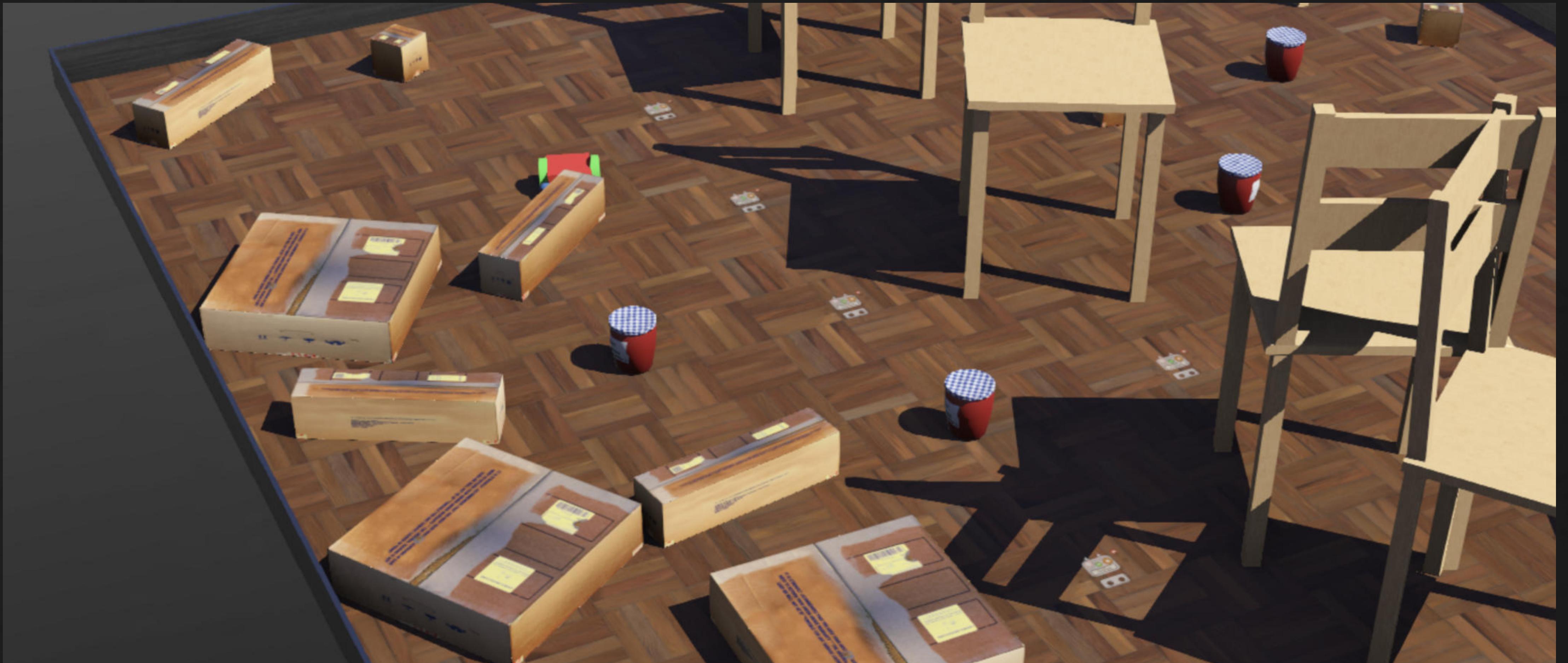
ELECTRONICS

Gap between the abilities of simulated sensor with respect to the real ones.

SOFTWARE STACK

A simple robot may require a lot of interactions between different platforms, libraries and protocols.

SIMULATED ENV



SIMULATED ENV



TARGET

The goal is reach a target
from a given starting position.

AVOID

Avoid obstacle while reading sensors.

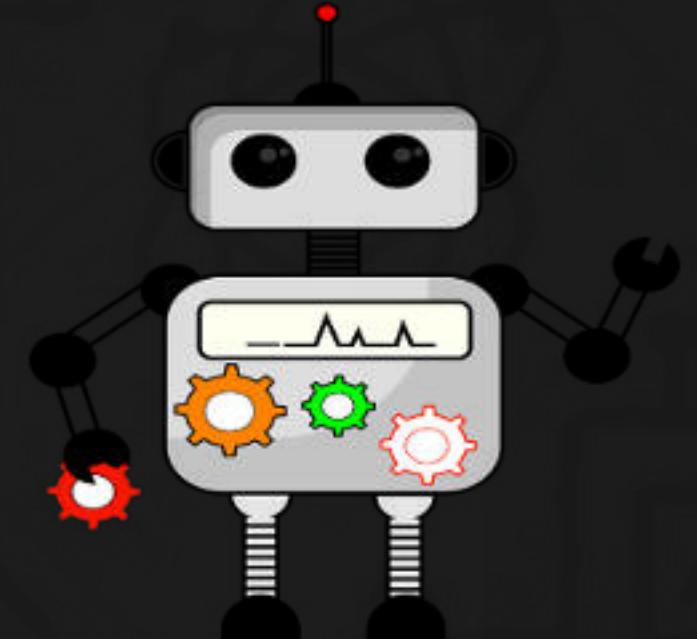
SIMULATED ENV



Tools



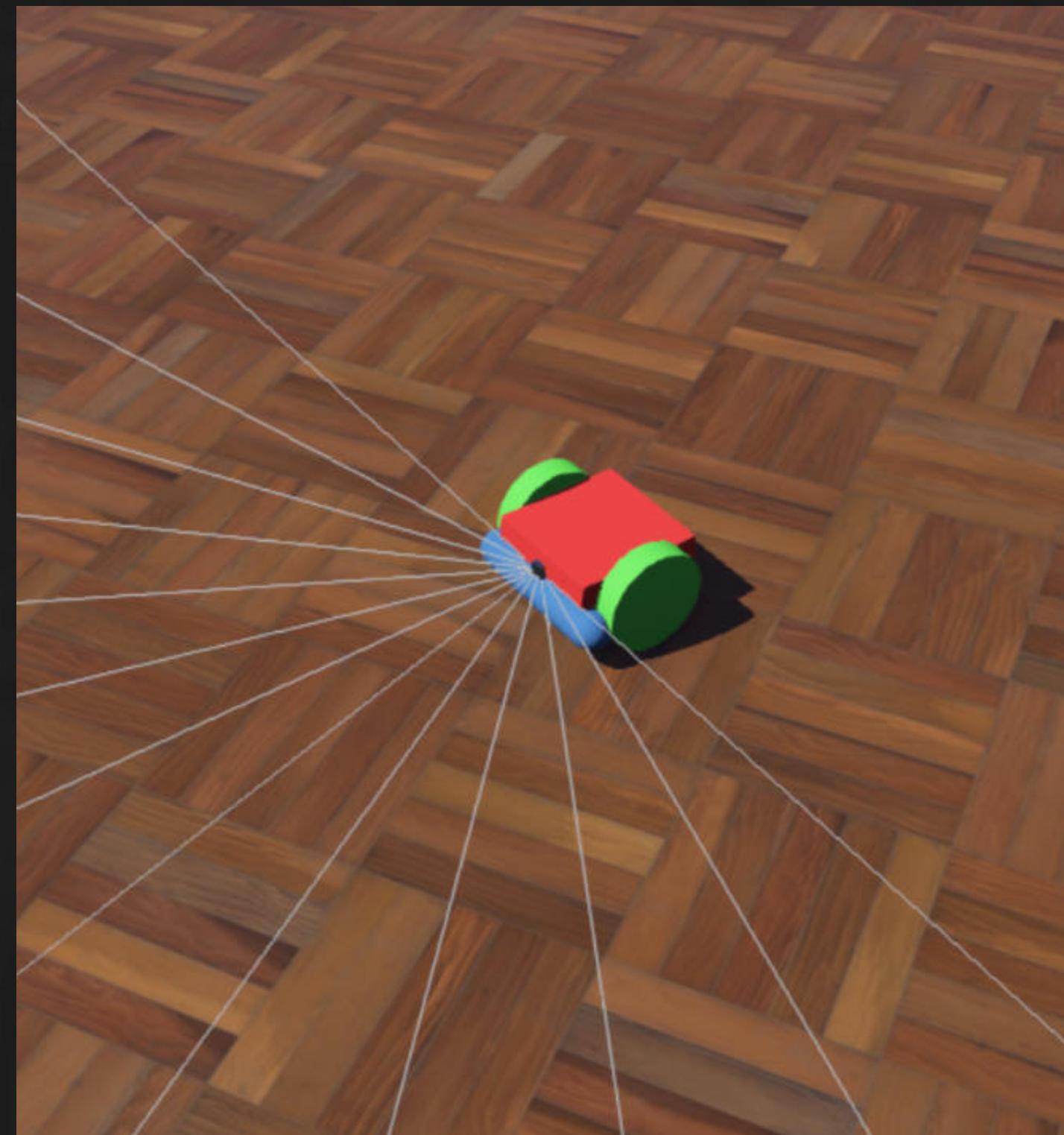
Webots simulator



DeepWorlds

A middleware between Gym and Webots

SIMULATED ROBOT



DIFFERENTIAL DRIVE

A simple differential drive even without a caster wheel.



SENSORS

13 Proximity sensors
2 Long-shape touch sensors

DETAILS - ALGORITHM

PPO with Action Masking

Some **basic rules** can be encoded as
the **default** robot behavior.

DETAILS - ALGORITHM

PPO with Action Masking

Reduce the “search space” of the network:

given the sensors value, a mask is applied on the policy network to avoid the execution of non-sense action.

DETAILS - ALGORITHM

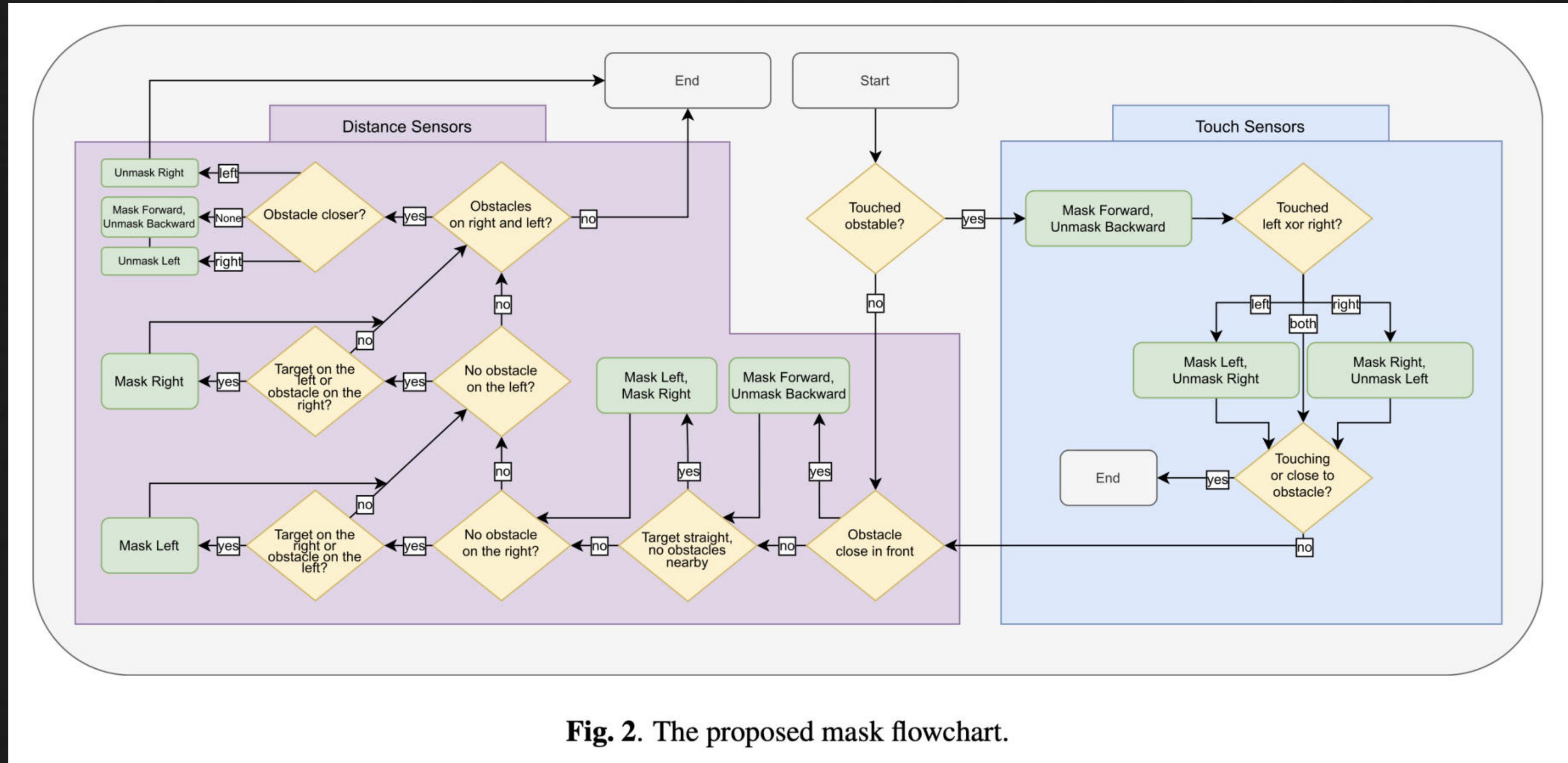


Fig. 2. The proposed mask flowchart

DETAILS - OBS



48 DIM TENSOR

24-dim tensor Prev. timestep

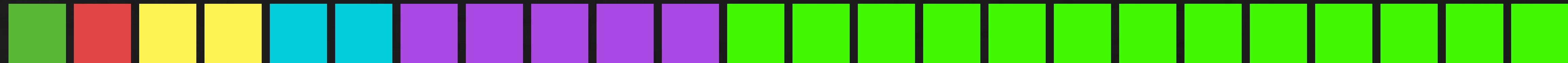
24-dim tensor current. timestep

DETAILS - OBS



48 DIM TENSOR

24-dim tensor Prev. timestep



One Hot Previous Action

13 Distance sensors

Left and right touch sensor

Left and right motor speed

Angle from robot to target

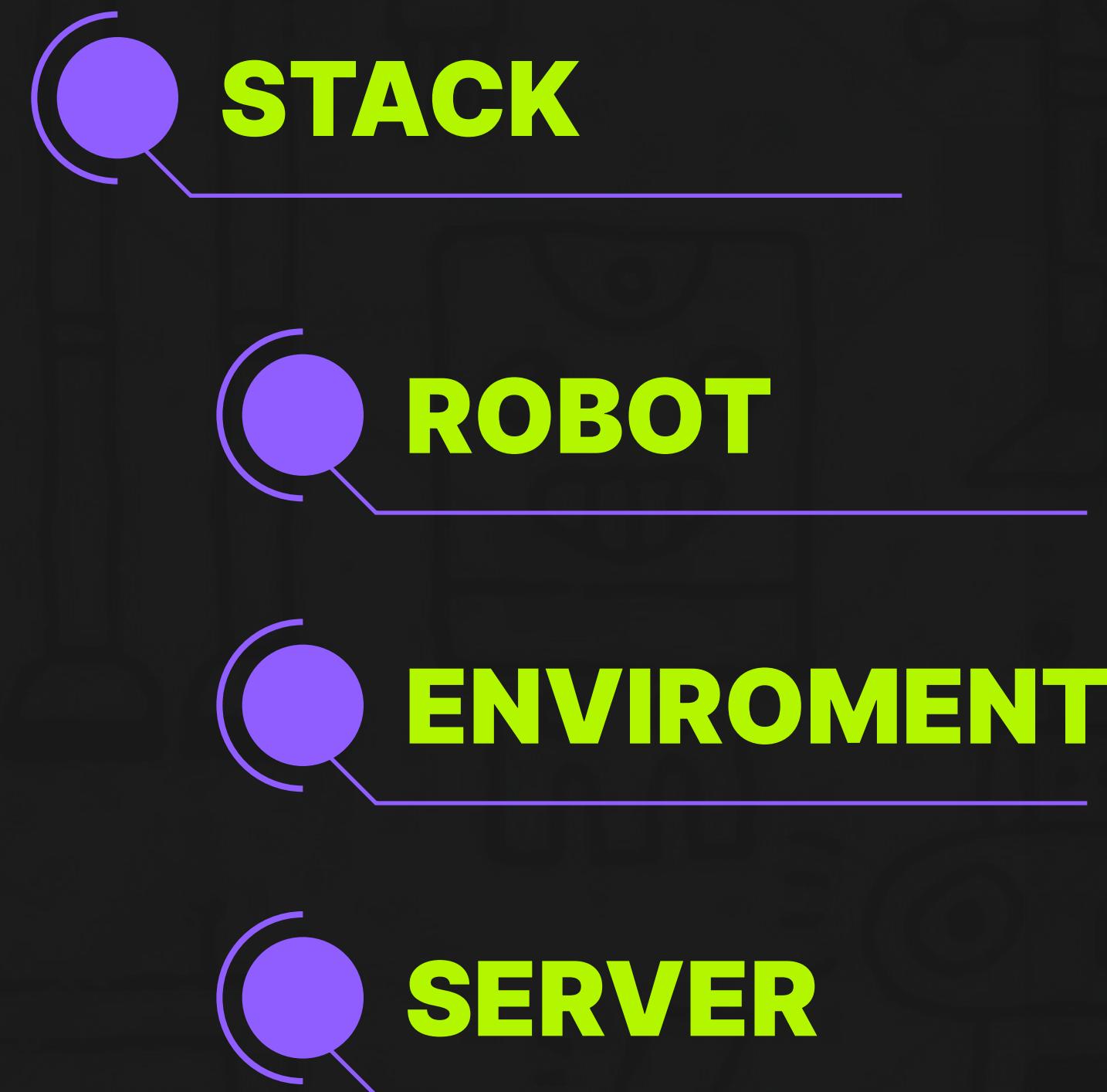
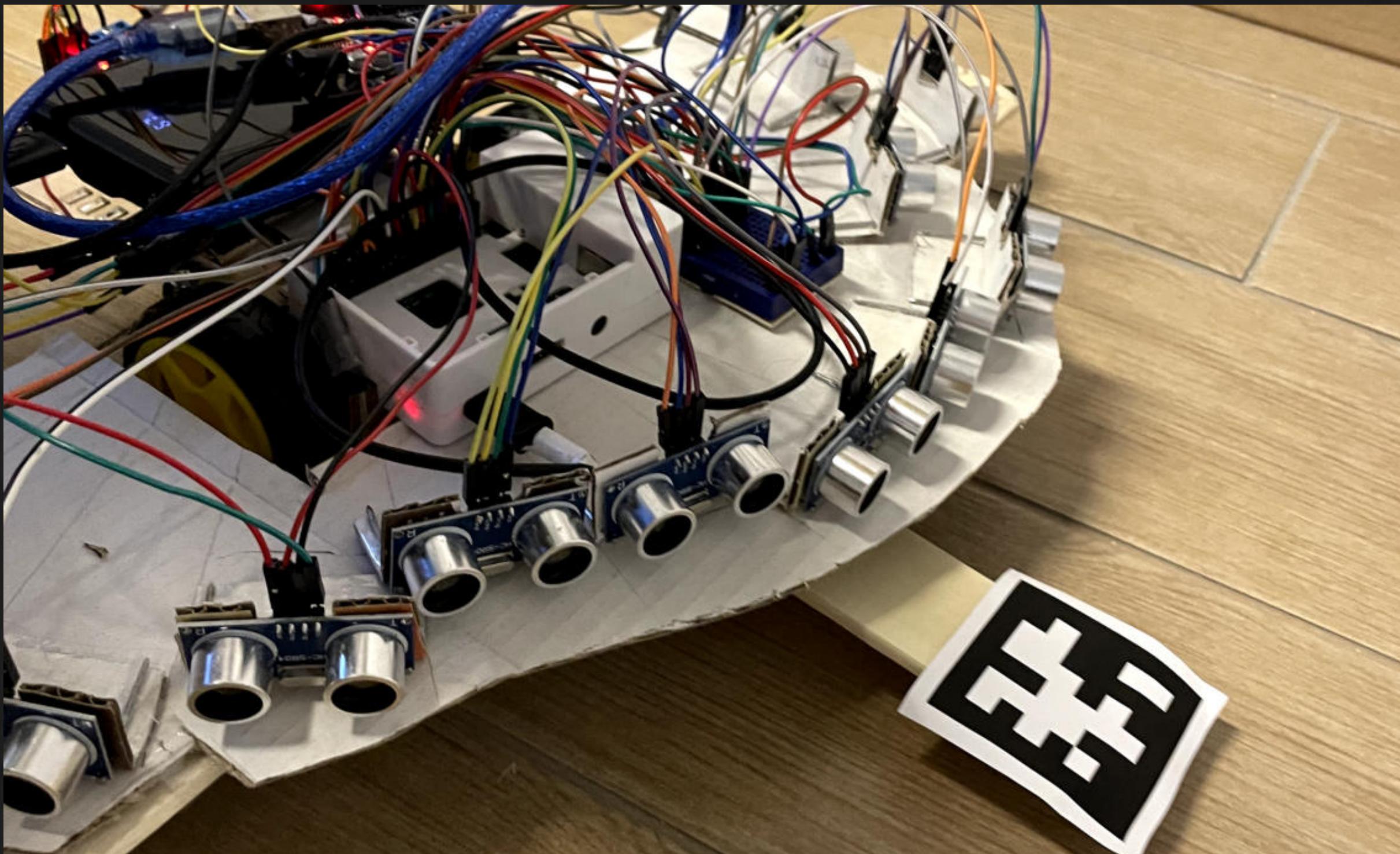
Distance from the target

DETAILS - REWARD

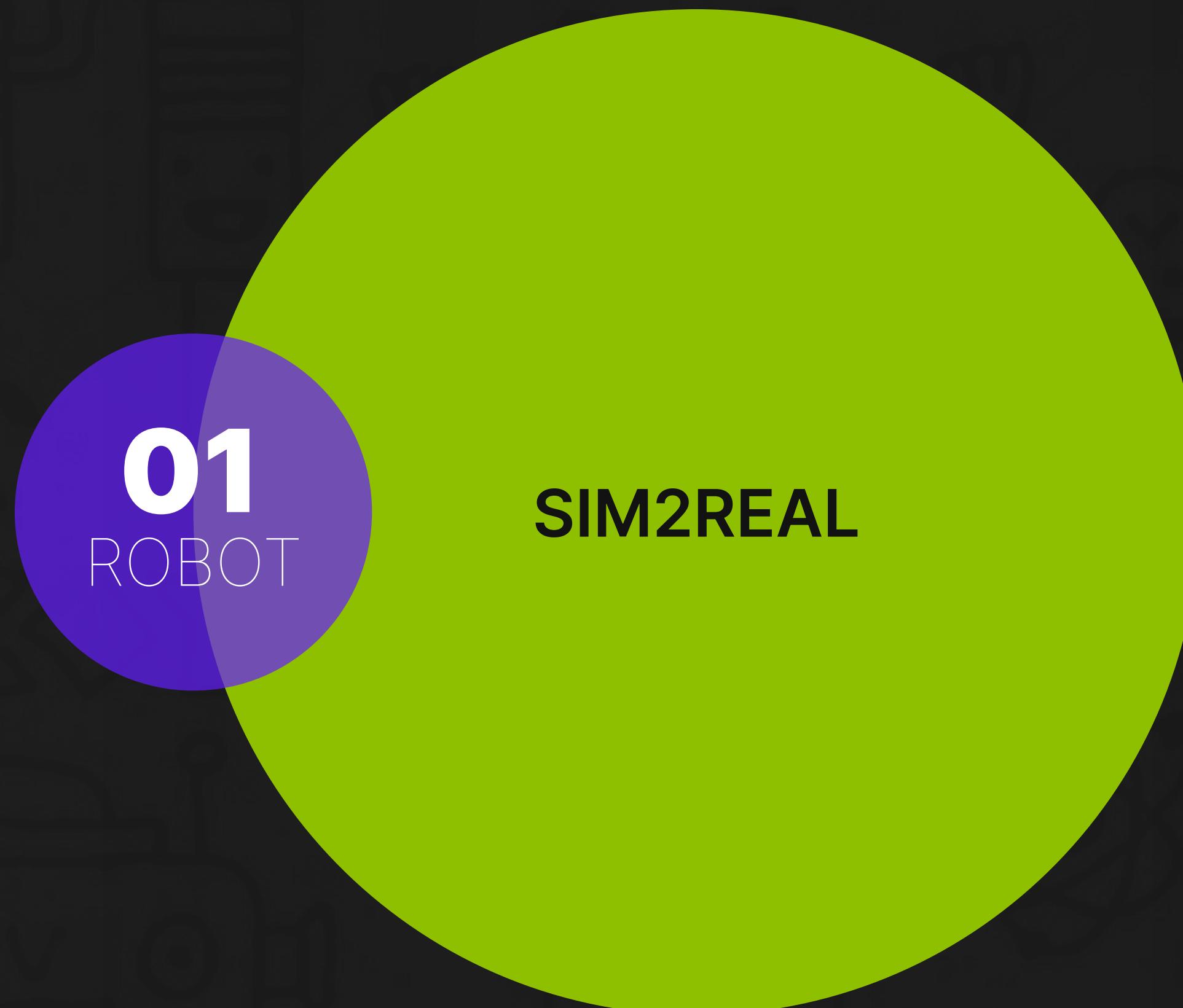
$$R = w_{dr} dr + w_{ar} ar + w_{dsr} dsr + w_{rtr} rtr + w_{cr} cr$$

- dr is the distance to target reward
- ar is the angle to target reward
- dsr is the distance sensors reward
- rtr is the reach target reward
- cr is the collision reward
- w_{dr} , w_{ar} , w_{dsr} , w_{rtr} , and w_{cr} are the weights assigned to each reward component

REAL WORLD



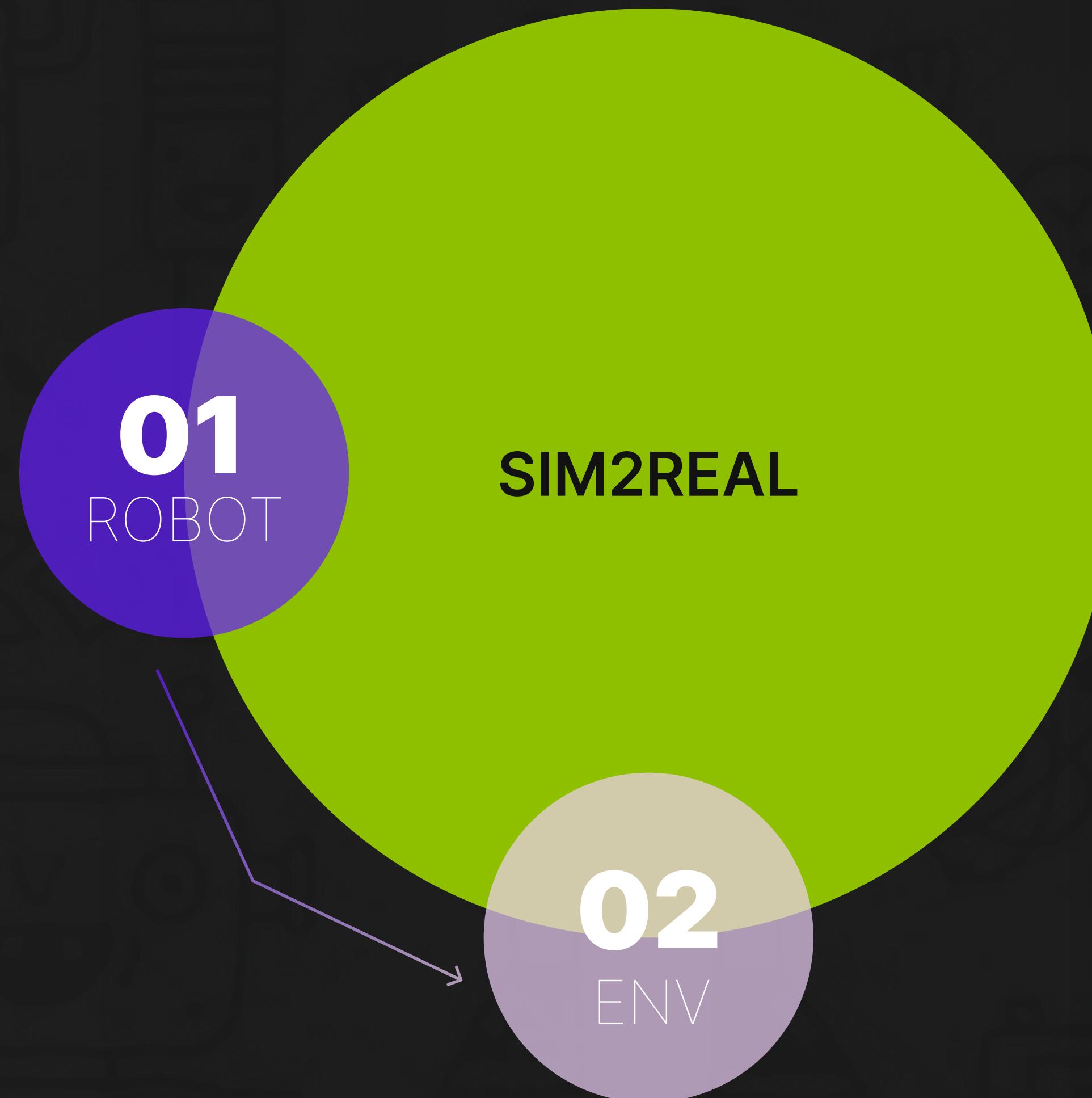
STACK



01 ROBOT

Physical Robot differential drive

STACK



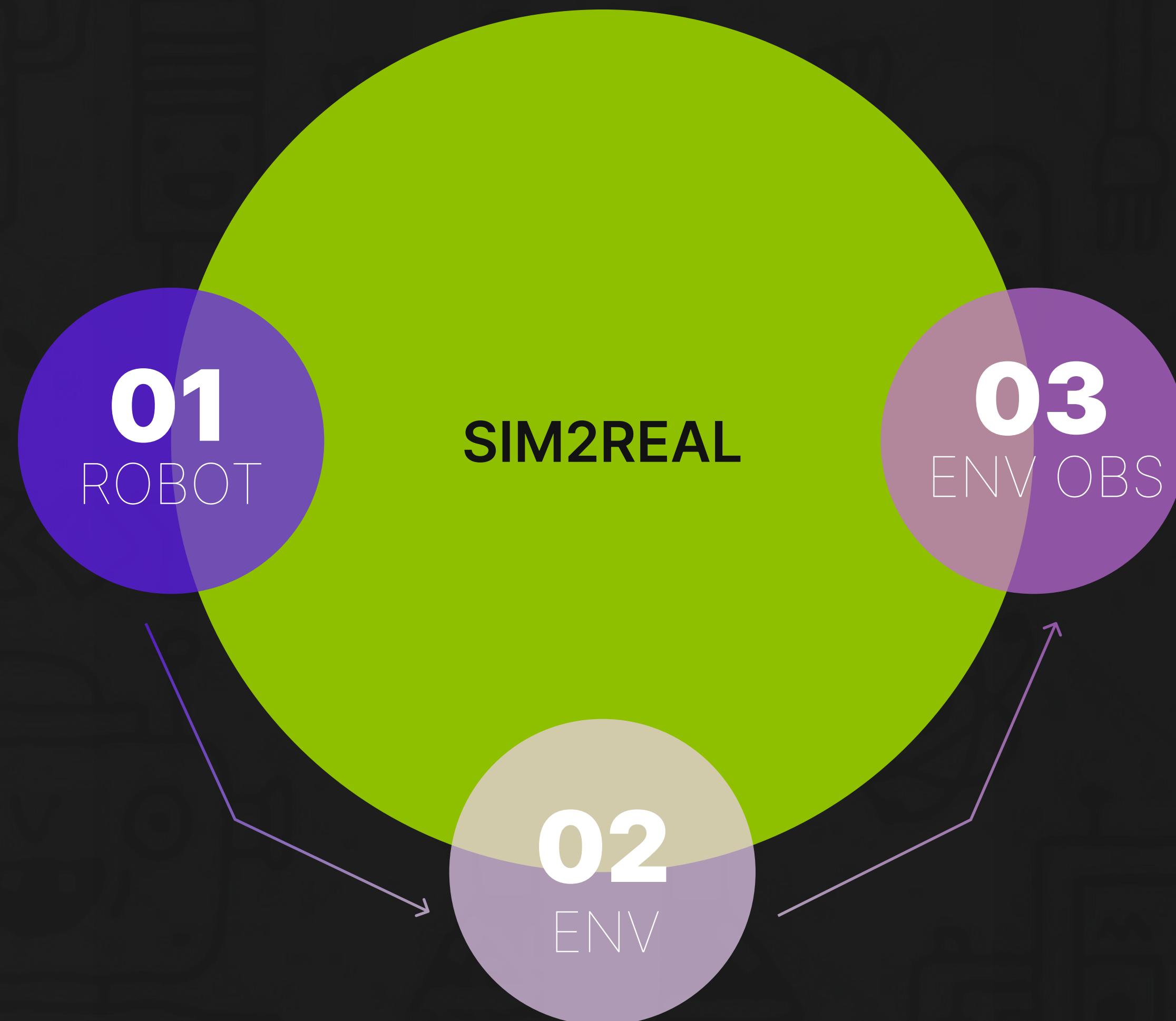
01 ROBOT

Physical Robot differential drive

02 ENV

Wall defined area in which the robot can operate

STACK



01 ROBOT

Physical Robot differential drive

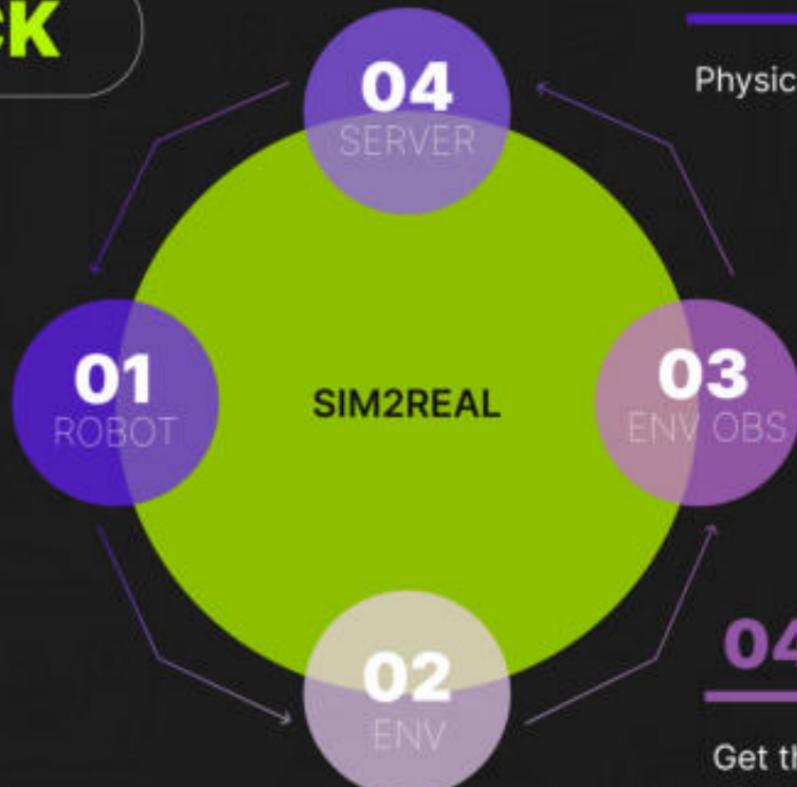
02 ENV

Wall defined area in which the robot can operate

03 ENV OBS

Information such us distance and angle between target and robot

STACK



01 ROBOT

Physical Robot differential drive

02 ENV

Wall defined area in which the robot can operate

03 ENV OBS

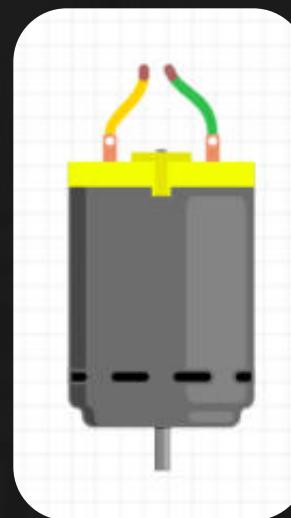
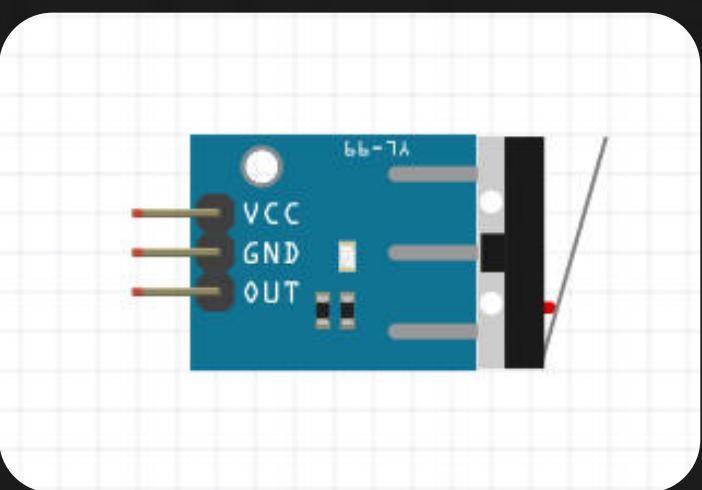
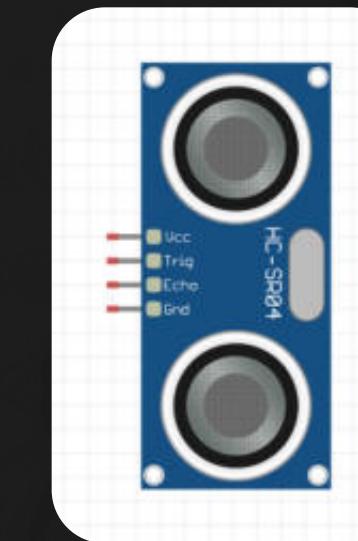
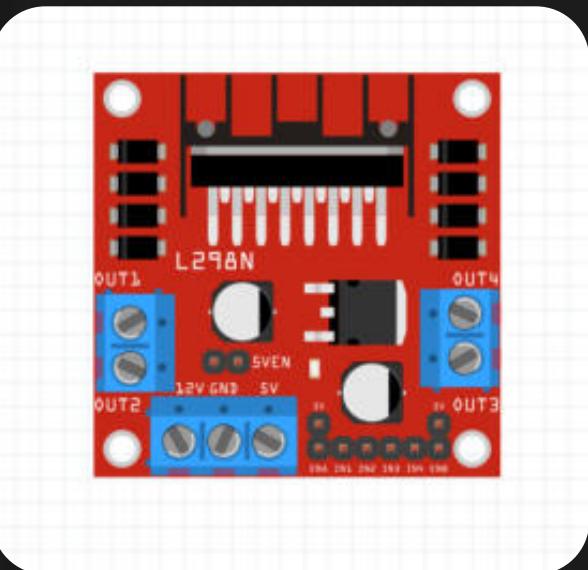
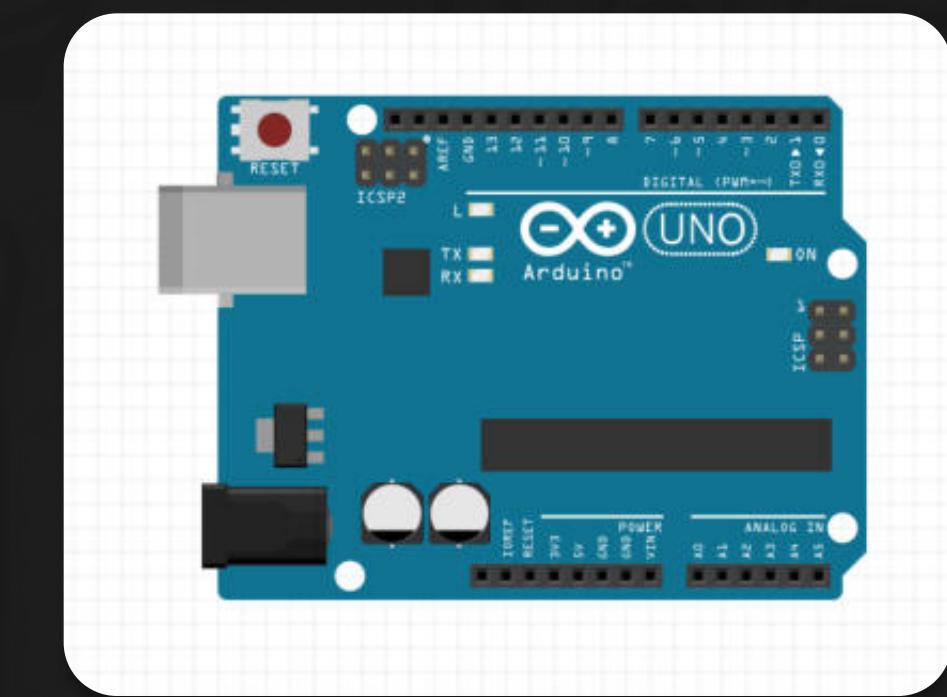
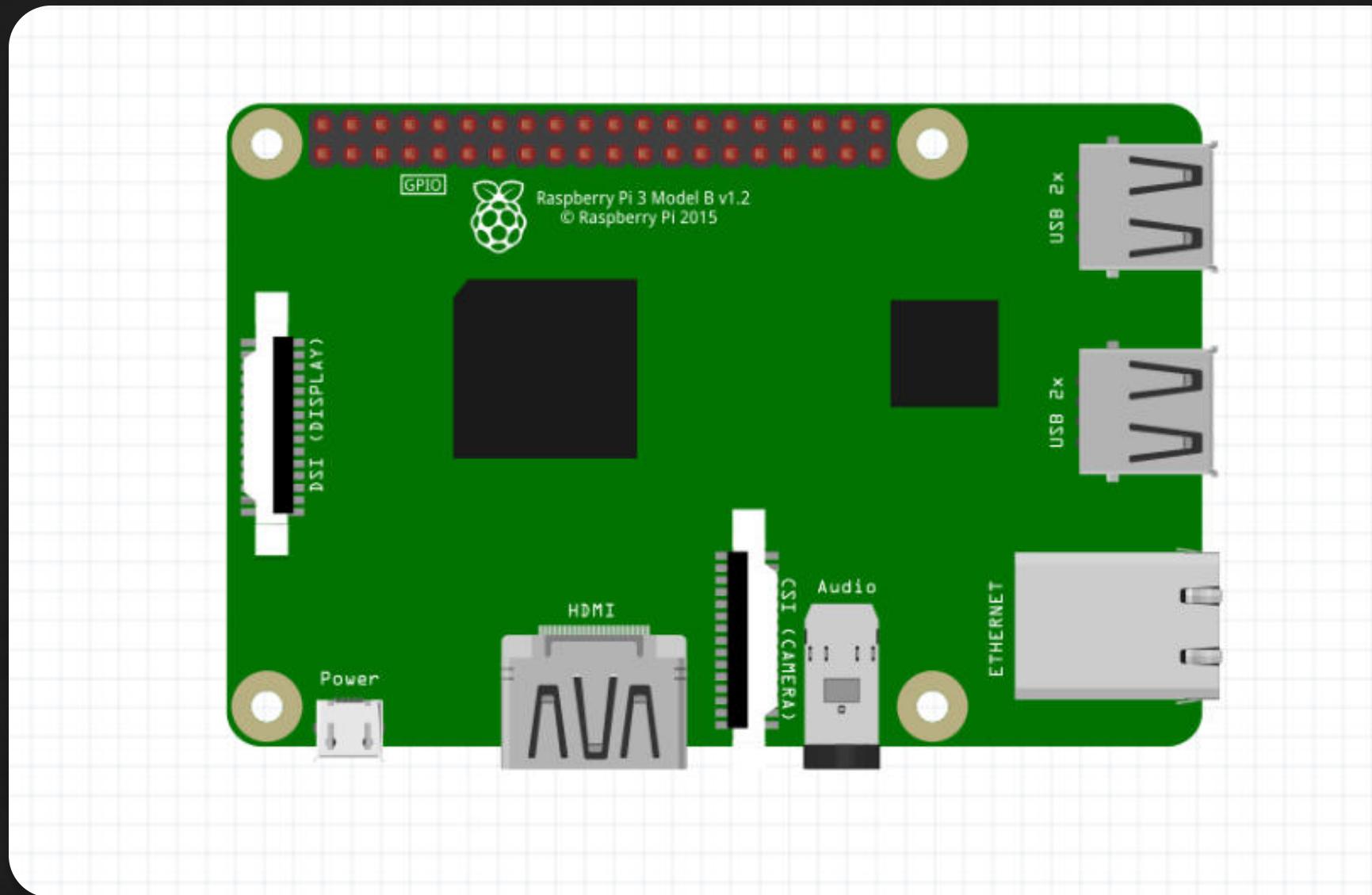
Information such us distance and angle between target and robot

04 SERVER

Get the obsevation, run the policy and give the action

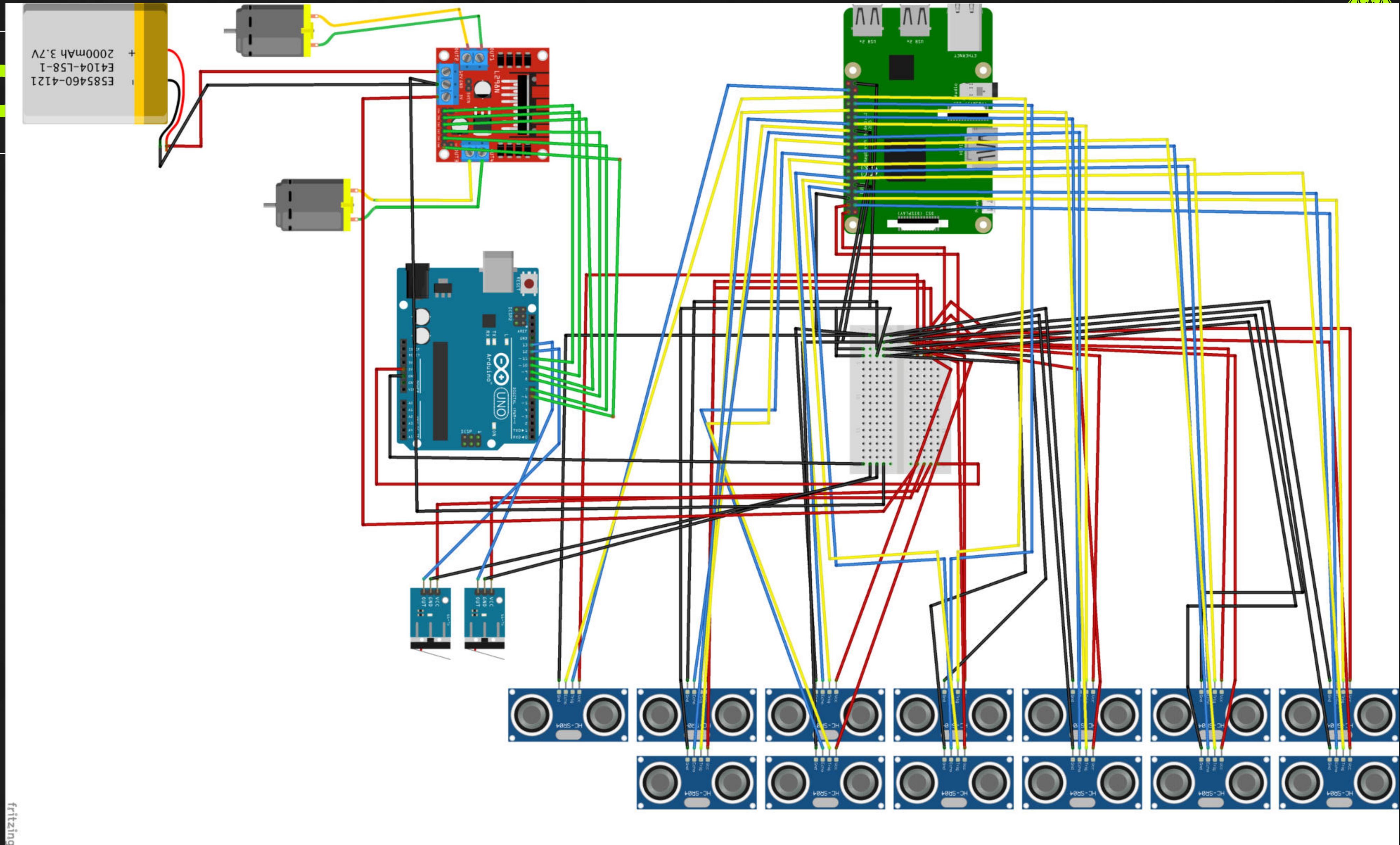


ROBOT





RC



fritzing





RC

121
1-1
3.7V

WHAT THE

MESS!

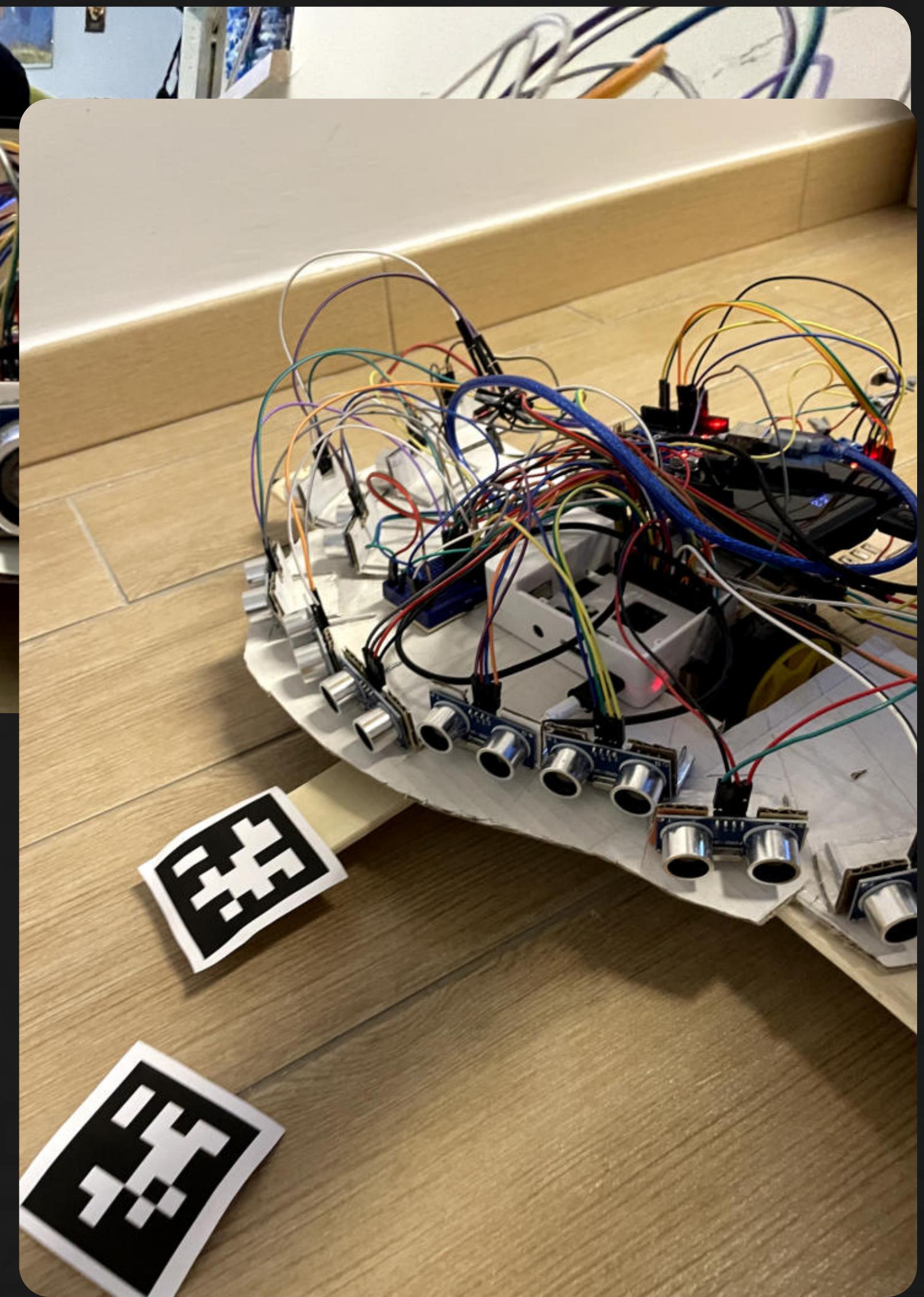
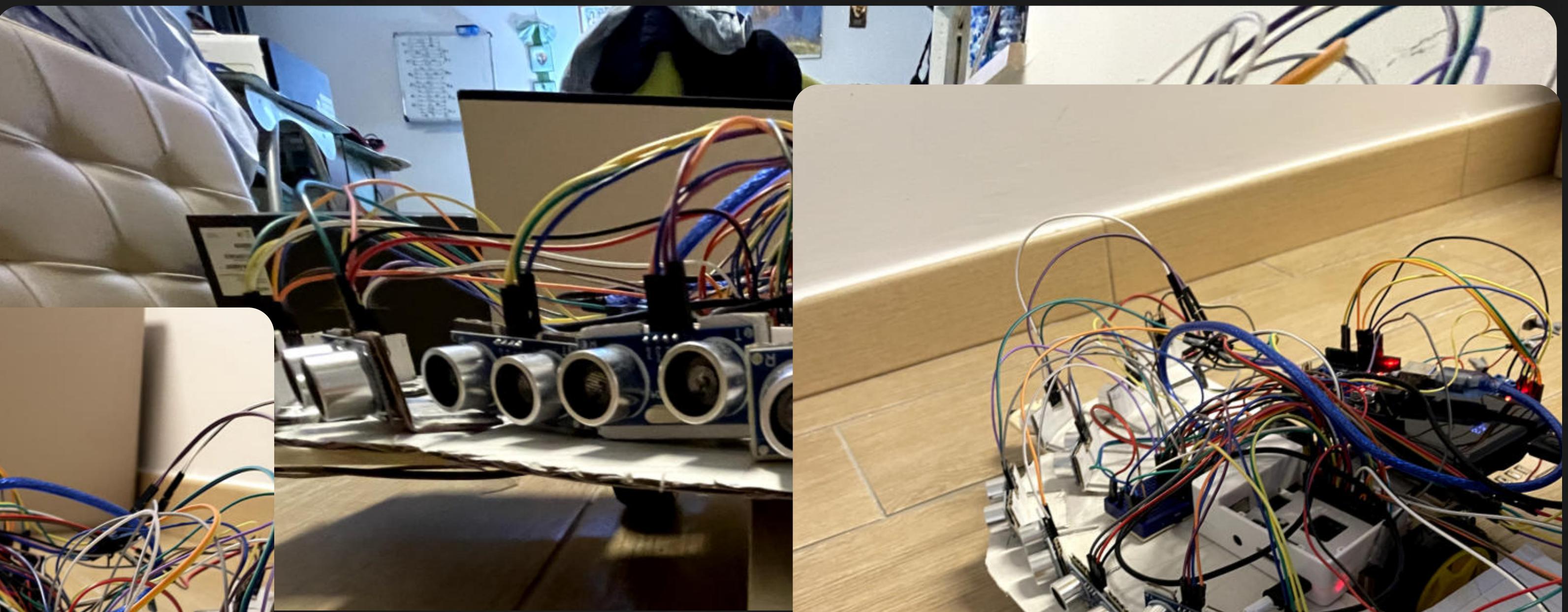
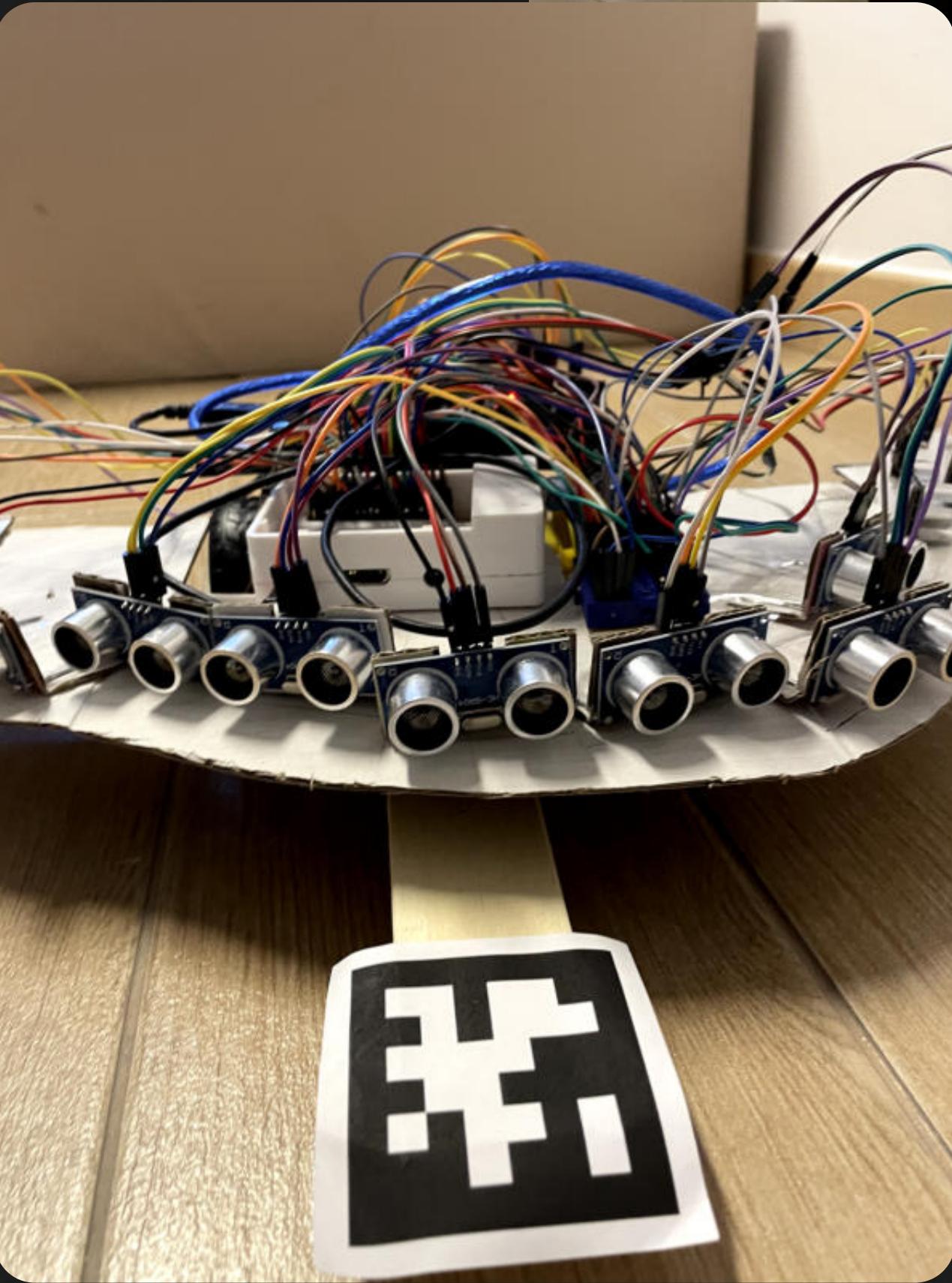
makeameme.org

fritzing





ROBOT

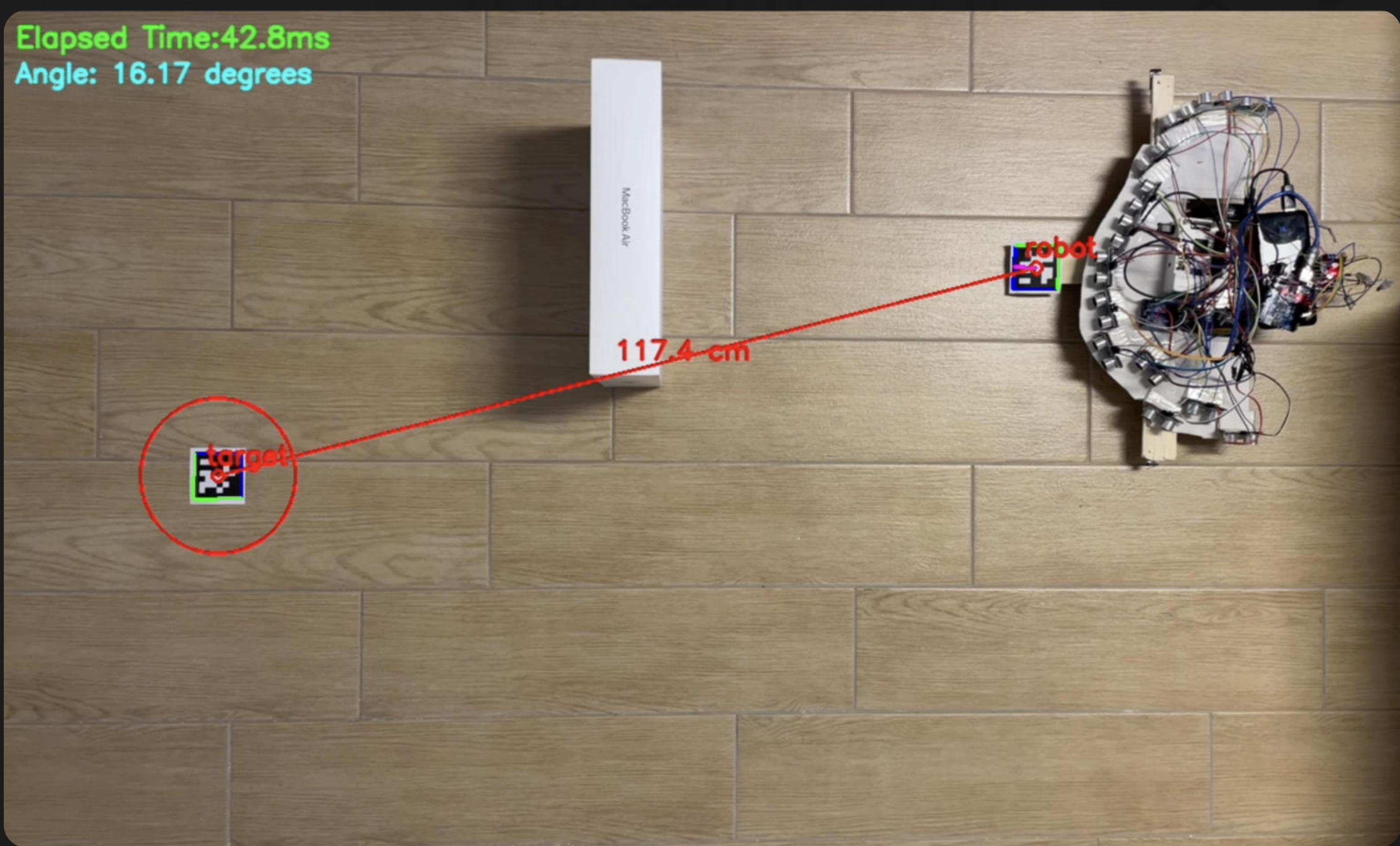




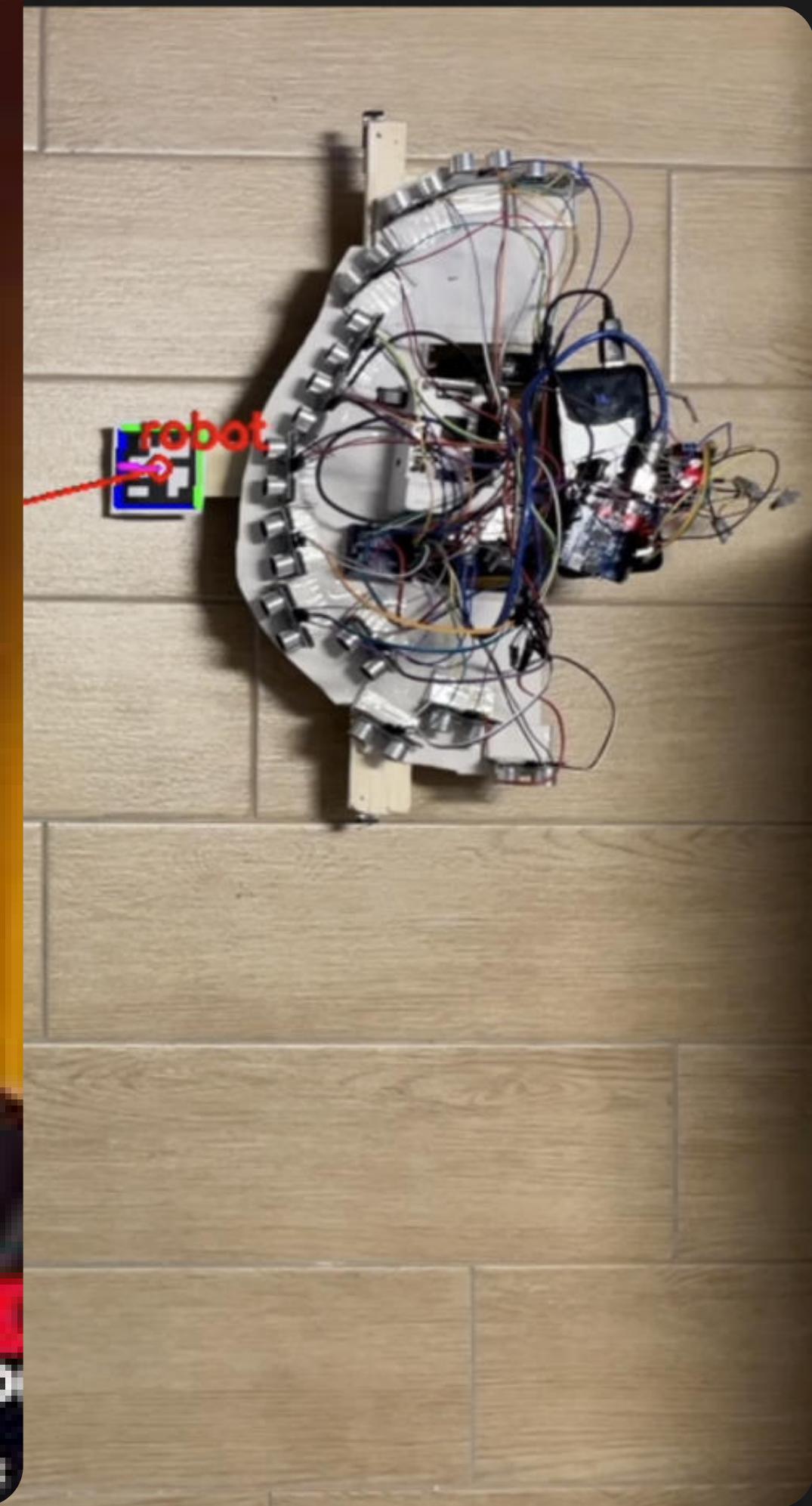
ENV



ENV OBS

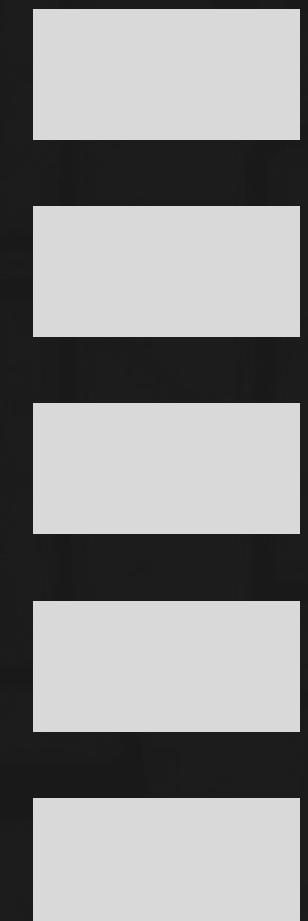
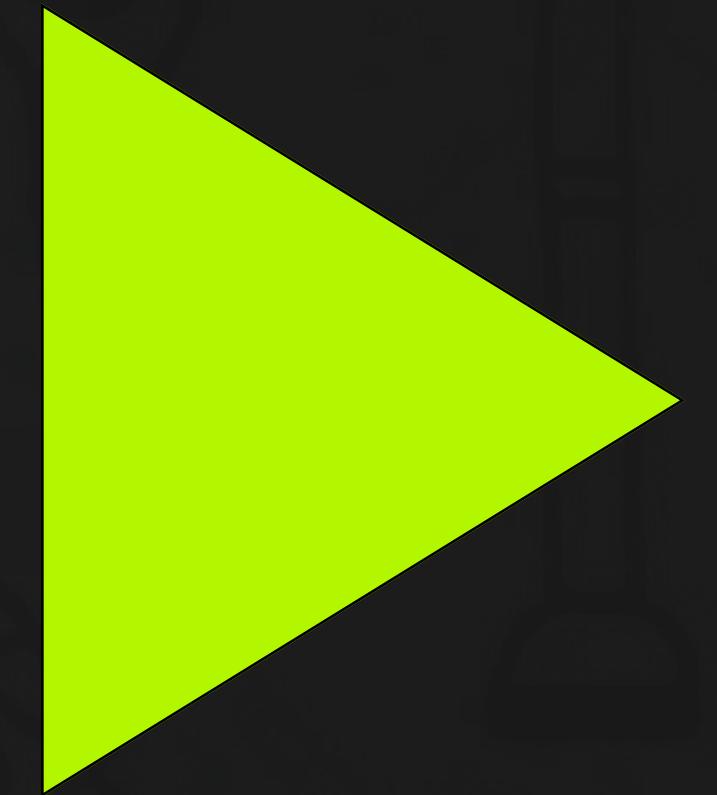


ENV OBS



SERVER

OBSERVATION



ACTION

ALL TOGETHER

01 ROBOT

Serial Handshake Raspberry/Arduino

Read distance sensors

Read Touch sensors

Send Info to the server

ALL TOGETHER

01 ROBOT

Serial Handshake Raspberry/Arduino
Read distance sensors from Raspberry
Read Touch sensors from Arduino
Send Info to the server

04 SERVER

Open Stream with continuity and ask the env observation to get the complete state for inference.
Sends back the response to robot.

ALL TOGETHER

01 ROBOT

Serial Handshake Raspberry/Arduino
Read distance sensors from Raspberry
Read Touch sensors from Arduino
Send Info to the server

04 SERVER

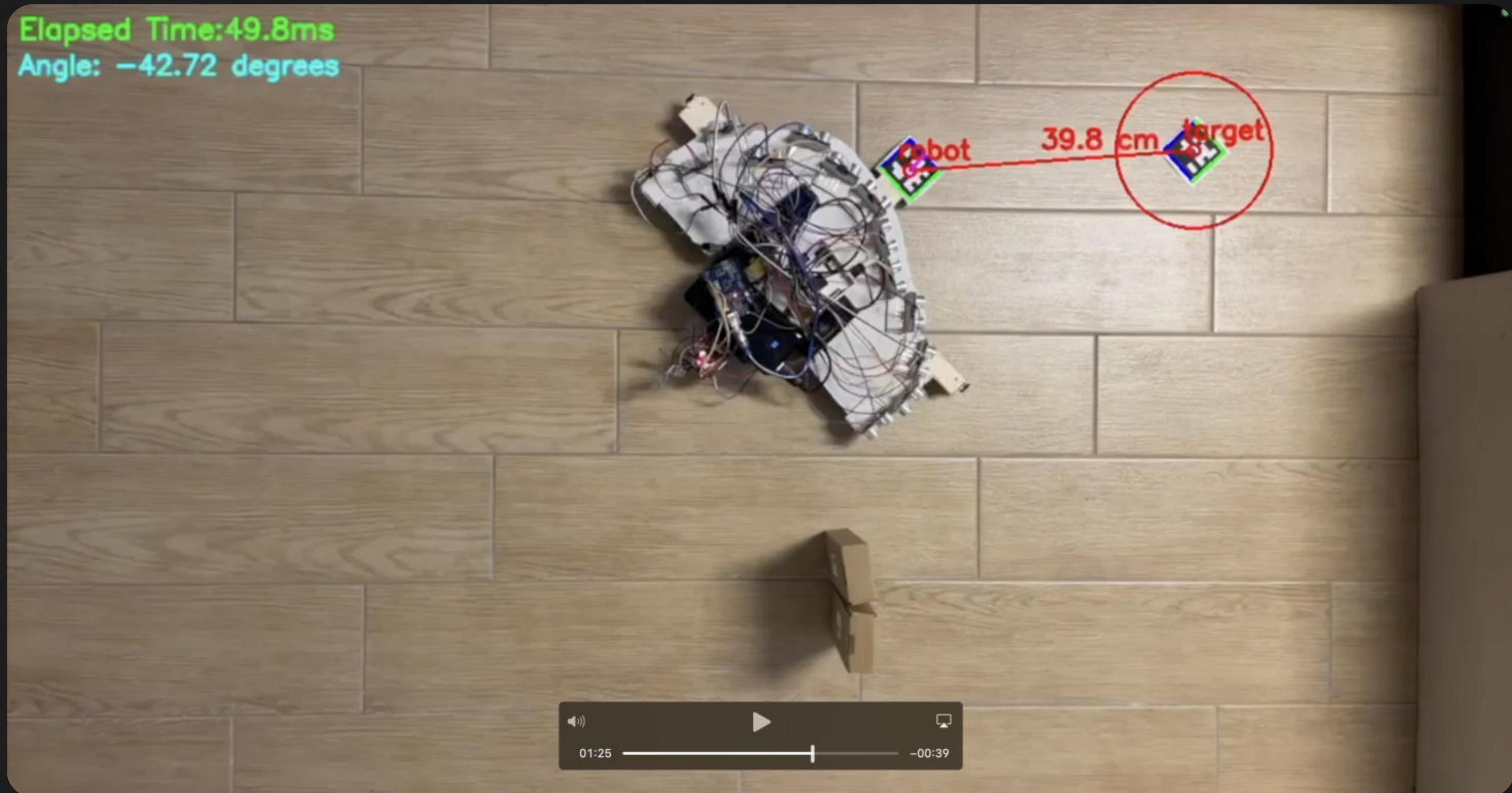
Open Stream with continuity and gets the observation to get the complete state for inference.
Sends back the response to robot

03 ENV OBS

Read distance and angle between target and robot using pupil_apriltags.



RESULT





IMPROV.

Basically Everything

Hardware

Electronics

Sim2Real stack

Communication protocol

CONCLUSION

With not appropriate material a lot can go wrong.

We can consider the experience gained on this project really valuable for our careers.



END



SAPIENZA
UNIVERSITÀ DI ROMA

...

QUESTIONS?

COMITO - RUCCI

Reinforcement Learning, Prof Rob. Capobianco

