

QCon
全球软件开发大会

如何在浏览器内构建一个具身智能平台

Tim Qian



📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AI Ops
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AICon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AICon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AICon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AICon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LM Ops
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

自我介绍



Tim Qian

github.com/timqian

独立开发者, vibe coder

兴趣: 开源、人工智能、区块链、机器人

开发了 star-history.com, bambot.org, chart.xkcd, puzi.io, openprompt.co 等产品





BamBot 的故事

AI 热潮之前，我曾经鼓捣过的机器人



● 宇数机器人, Aloha, AI 热潮



2024 年 5 月
¥99000



2024 年 1 月
\$32000

从 SO-100 机械臂 到 BamBot

Standard Open SO-100 & SO-101 Arms



2024年10月
~\$100 机械臂

<https://github.com/huggingface/lerobot>



Tim Qian @Tim_Qian · Mar 18

Introducing BamBot, an open source, low cost (~\$300) humanoid robot.

Inspired by @LeRobotHF, so-100 arm and lekiwi



58

292

1.4K

188K

2025年3月
~\$300 机器人

<https://github.com/timqian/bambot>

如何控制它?

lerobot-bambot Public

Forked from [huggingface/lerobot](#)

Modified lerobot for bambot

Python 17 2 Apache License 2.0 Updated on Mar 25

1. Fork lerobot

2. 改代码和文档

Table of contents

- [A. Source the parts](#)
- [B. Install software Pi](#)
- [C. Setup LeRobot laptop/pc](#)
- [D. Assemble the arms](#)
- [E. Calibrate](#)
- [F. Teleoperate](#)
- [G. Record a dataset](#)
- [H. Visualize a dataset](#)
- [I. Replay an episode](#)
- [J. Train a policy](#)
- [K. Evaluate your policy](#)



Tip

If you have any questions or need help, please reach out on [Discord](#) in the channel [#mobile-so-100-arm](#).

A. Source the parts

Follow this [README](#). It contains the bill of materials, with a link to source the parts, as well as the instructions to 3D print the parts, and advice if it's your first time printing or if you don't own a 3D printer.

Before assembling, you will first need to configure your motors. To this end, we provide a nice script, so let's first install LeRobot. After configuration, we will also guide you through assembly



第一个买家



Tim Qian @Tim_Qian · Apr 4

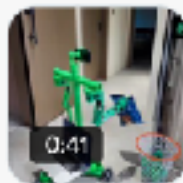


Sold the first bambot! I have to say that bambot v0.0.1 is not in a suitable state for sale, lot of places need to be improved. But do you think it's possible that, in the not too distant future, everyone will own a bot like this one?

—just as they own a phone today



Tim Qian @Tim_Qian · Mar 18



Introducing Bambot, an open source, low cost (~\$300) humanoid robot.

Inspired by @LeRobotHF, so-100 arm and Iekiwi



9



4



81



6.2K





一个麻烦

没有相关技术背景
很难上手让机器人自己干活



目前主流的两种机器人控制平台

ROS (Robot Operating System)

1 定义消息接口 (使用标准消息)

ROS 提供了标准的速度控制消息:

```
builtin
geometry_msgs/Twist
```

复制代码

2 发布控制命令 (Python)

```
python
import rospy
from geometry_msgs.msg import Twist

rospy.init_node('robot_controller')
pub = rospy.Publisher('/cmd_vel', Twist, queue_size=10)
rate = rospy.Rate(10)

while not rospy.is_shutdown():
    cmd = Twist()
    cmd.linear.x = 0.2 # 前进速度 (m/s)
    cmd.angular.z = 0.6 # 转动速度 (rad/s)
    pub.publish(cmd)
    rate.sleep()
```

复制代码

这个节点会不断向 /cmd_vel 发布速度 0.2 m/s 的命令。

3 硬件控制节点 (接收并下发出电机)

```
python
import rospy
from geometry_msgs.msg import Twist
import serial
```

复制代码

LeRobot

这是最小运行例子 (仿真版)

```
python
from lerobot import load_dataset, train, load_policy, ControlEnv

# 1. 准备数据
ds = load_dataset("lerobot/datasets/pusht")

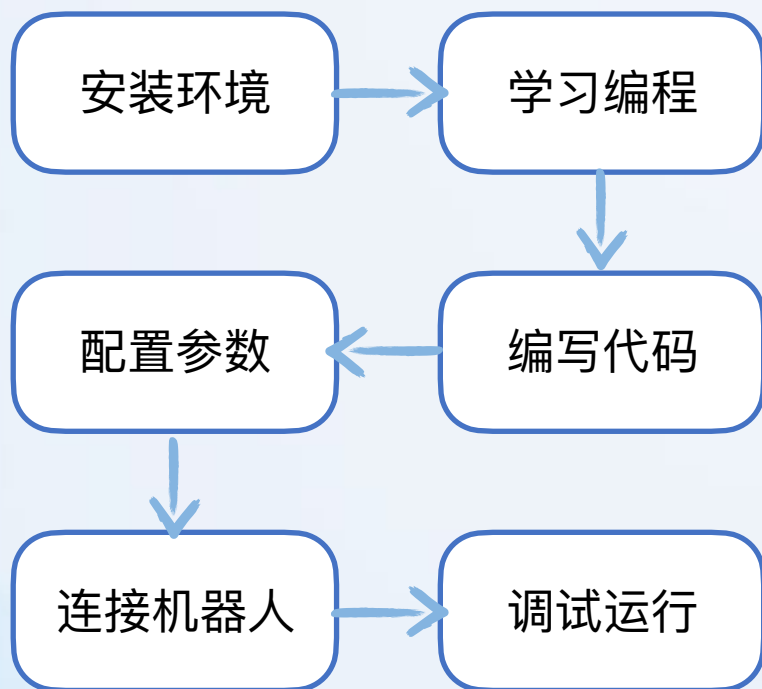
# 2. 训练
train.run(dataset="lerobot/datasets/pusht", total_steps=2000)

# 3. 推理/部署
env = ControlEnv("pusht")
policy = load_policy("runs/pusht/model.pt")

obs = env.reset()
for _ in range(500):
    action = policy(obs)
    obs, _, done, _ = env.step(action)
    env.render()
```

复制代码

传统方式

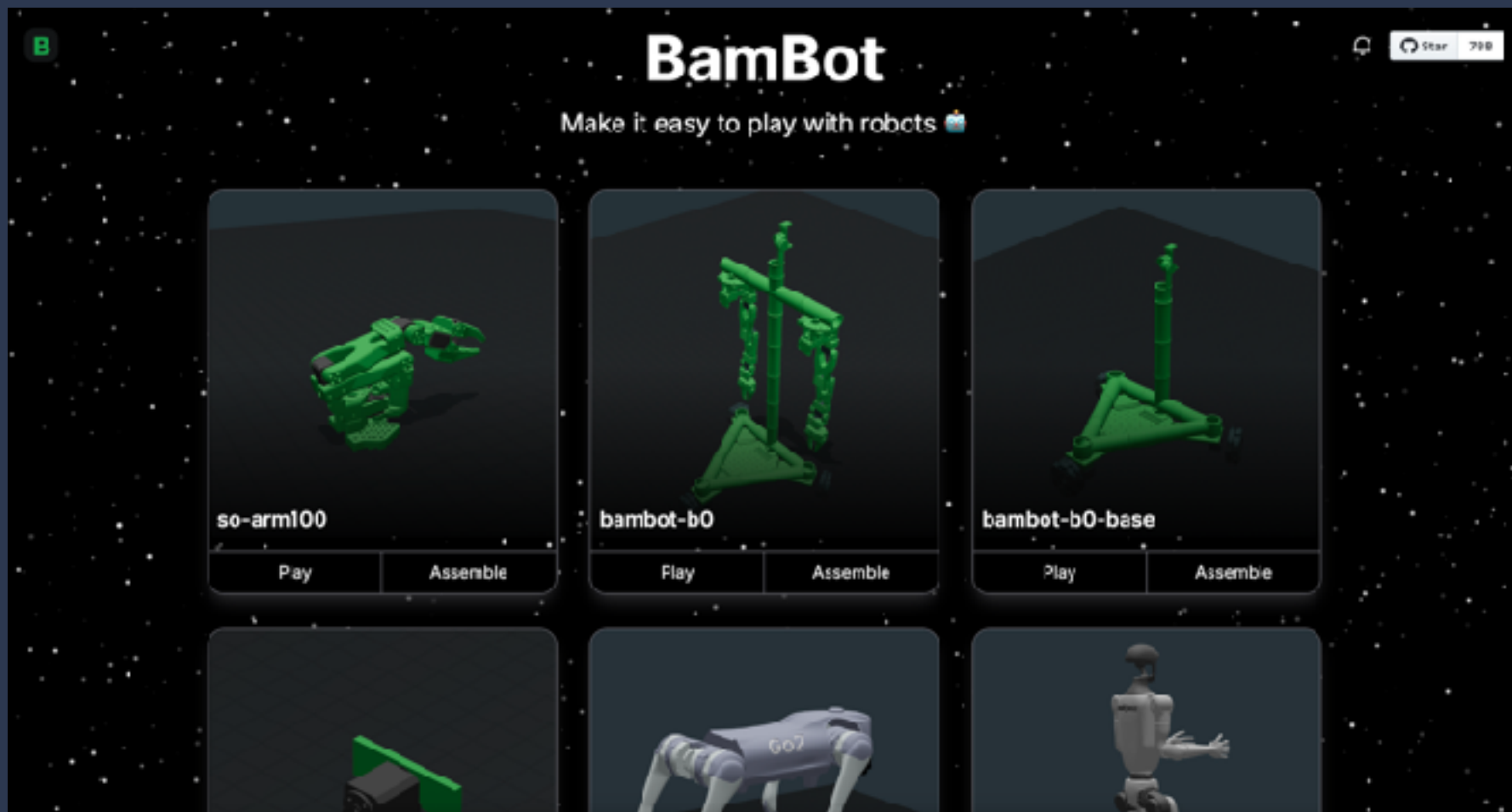


有没有可能



02 用户体验设计

● 着陆页：用户选择想要操控的机器人



● 机器人仿真控制页面

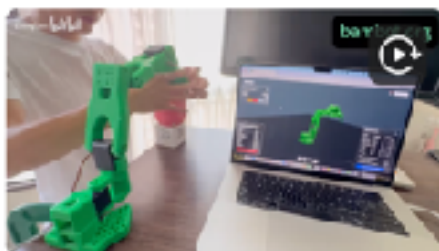


以 APP 的形式支持不同的功能



通过 bambot.org 遥操作虚拟 so-100 机械臂

遥操作



在网站上轻松录制和回放 so-101 机械臂的动作

动作录制和回放



在网站上愉快玩耍宇树机器人

数据采集

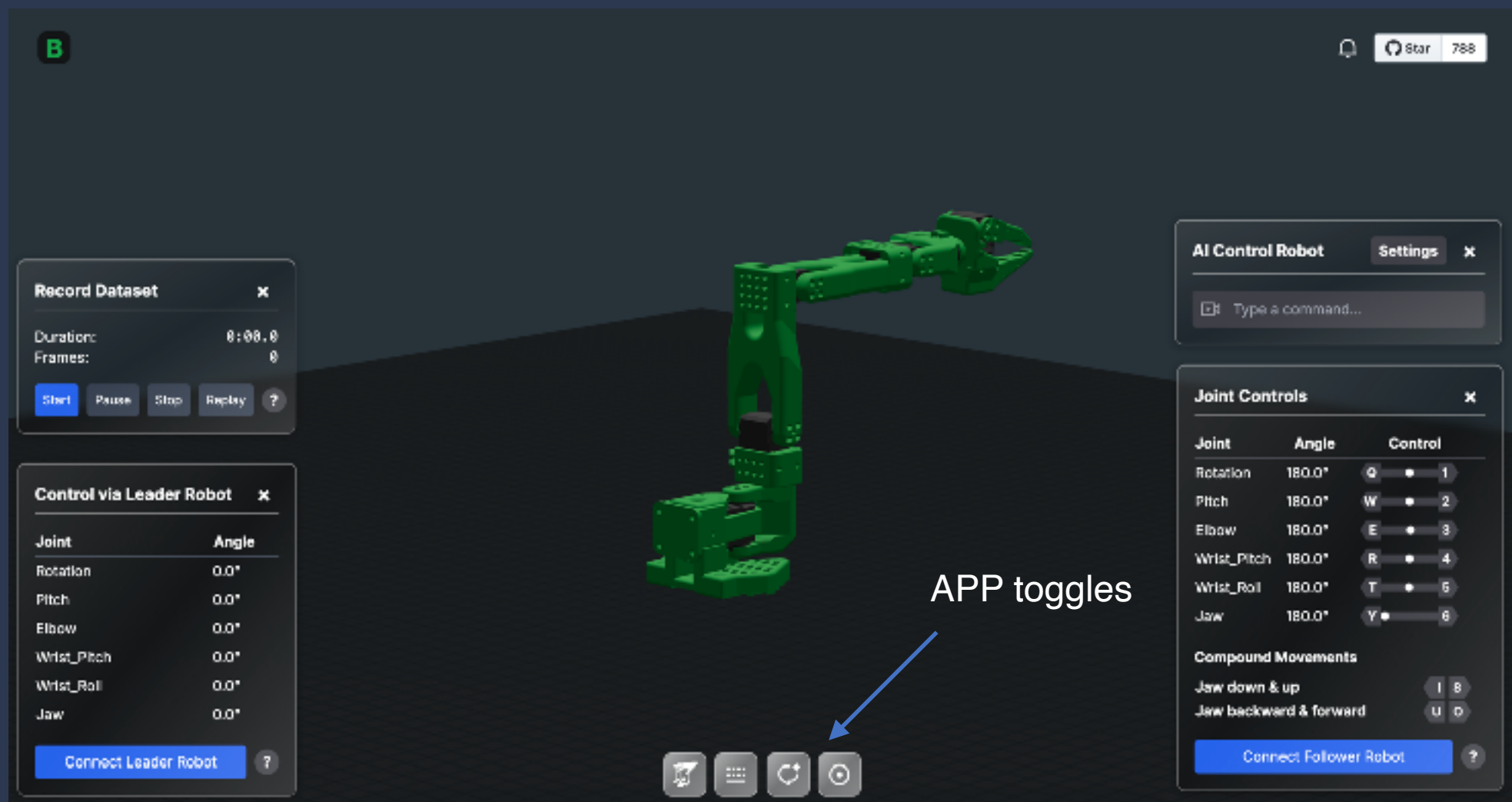


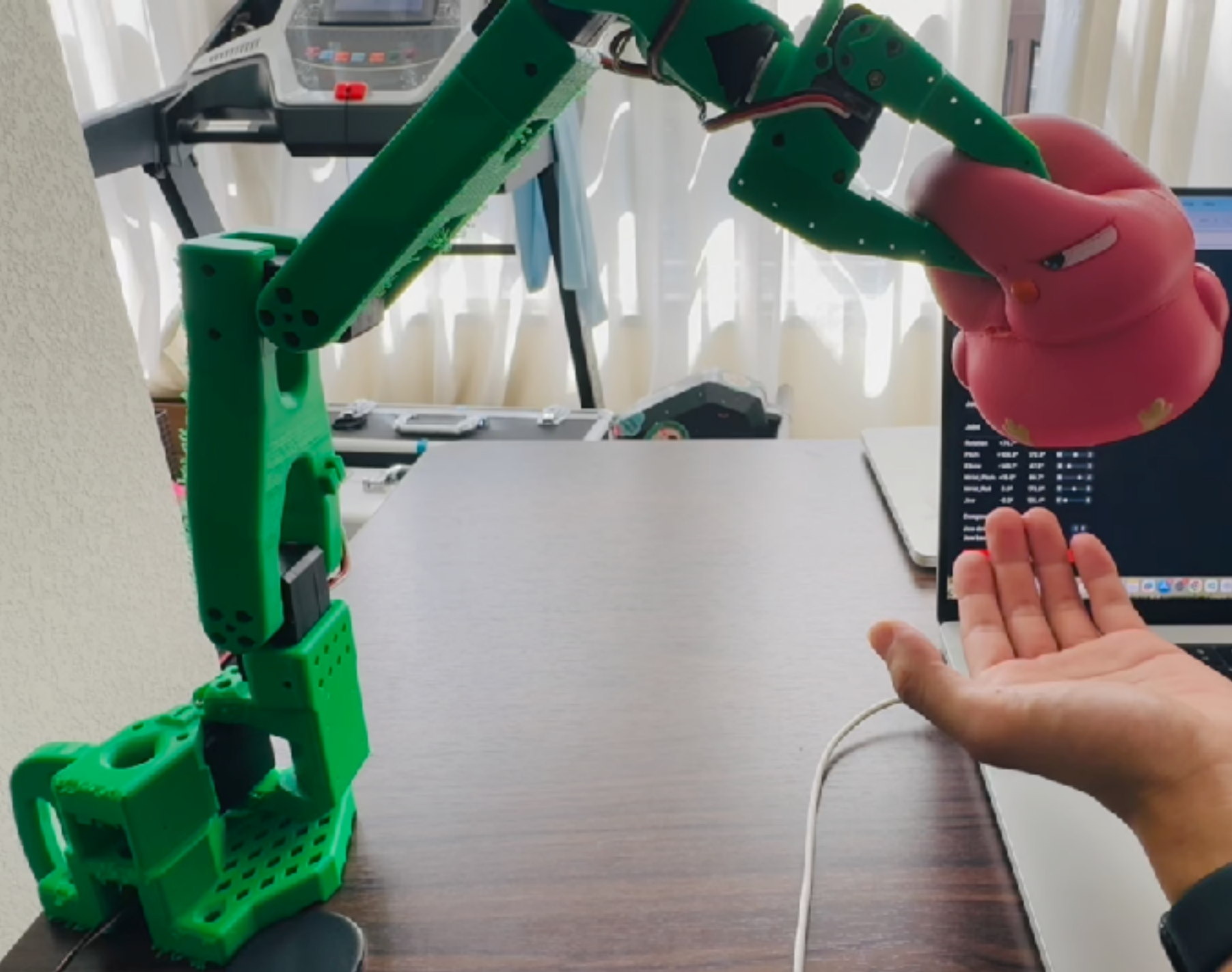
实现了大语言模型控制 so-arm100 机械臂，欢迎大家试玩

大模型控制机器人

<https://space.bilibili.com/11206127>

● 机器人仿真控制页面 – APPs





03

如何构建这个平台

3.1 在浏览器里 3D 渲染和仿真



描述机器人的两种文件格式

URDF

(Unified Robot Description Format)

- 基于 xml
- ROS (Robot Operating System) 生态系统的标准格式
- 主要用于机器人可视化、运动学和简单动力学仿真

MJCF

(MuJoCo Modeling Language)

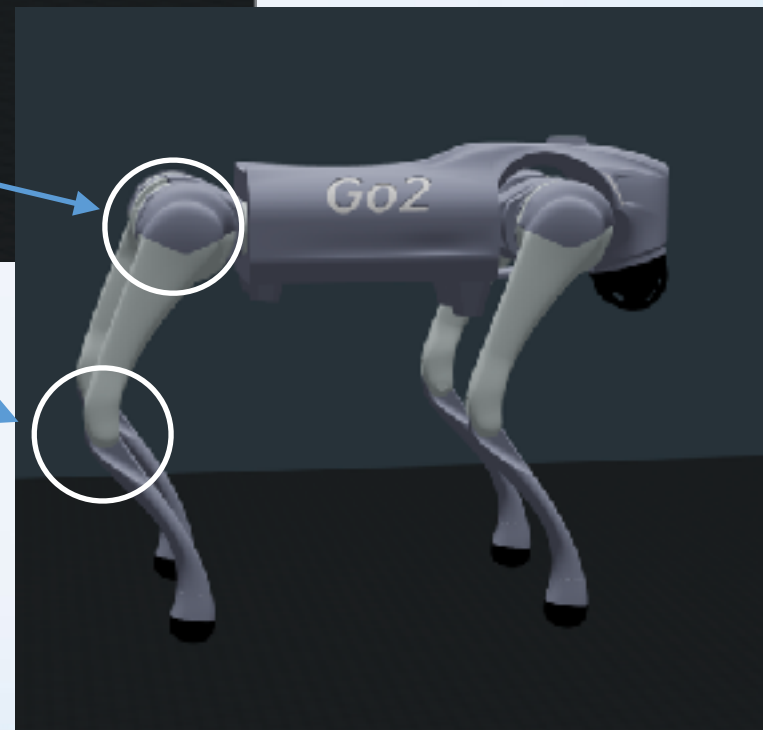
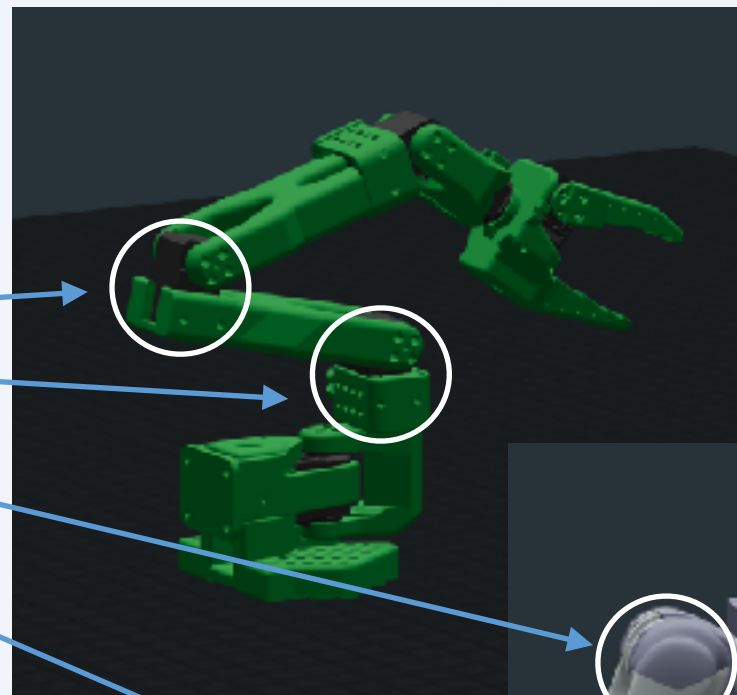
- 基于 xml
- MuJoCo物理引擎的原生格式
- 更强大的物理仿真能力（接触、摩擦、软体等）

收集不同机器人 3d 模型: https://github.com/robot-descriptions/robot_descriptions.py

● 机器人的两个核心概念

Joint (关节)

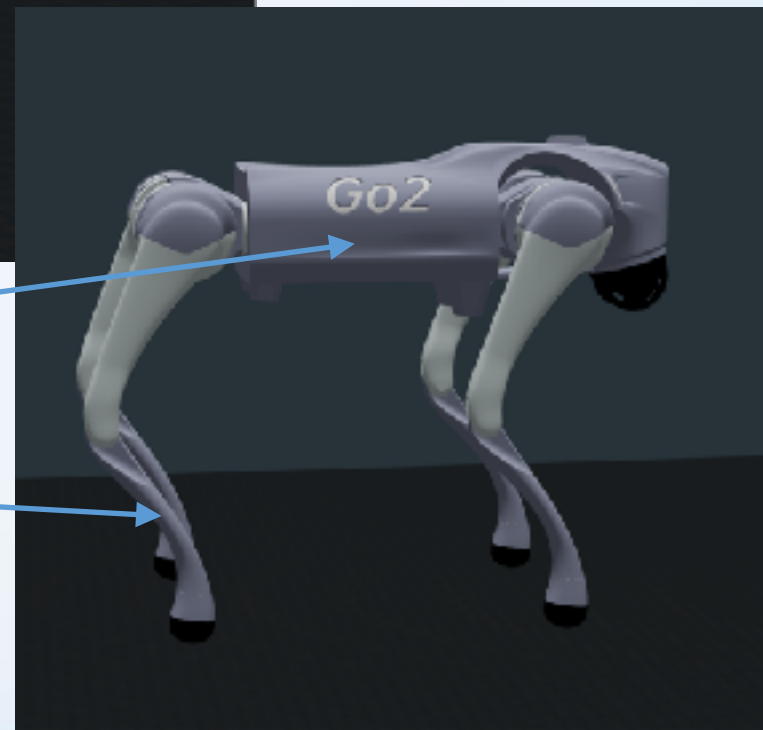
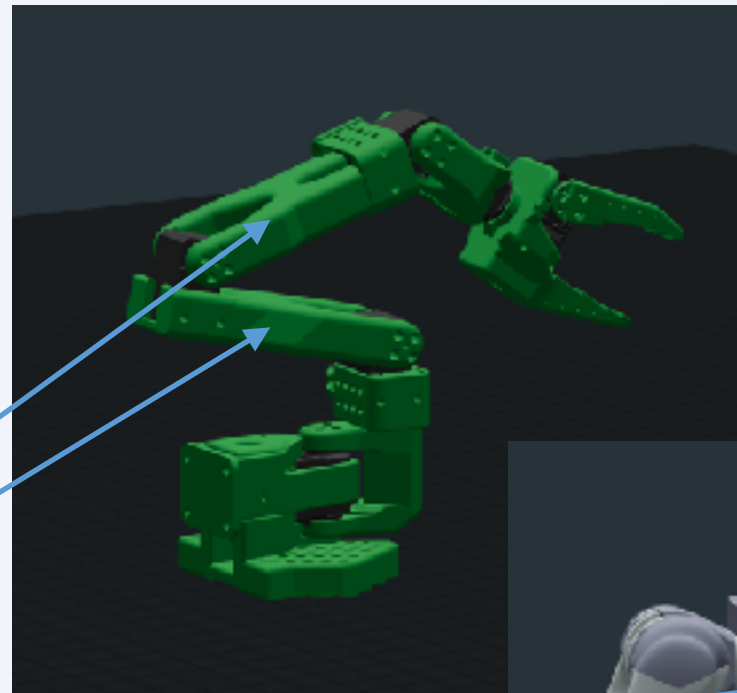
Link (连杆/刚体)



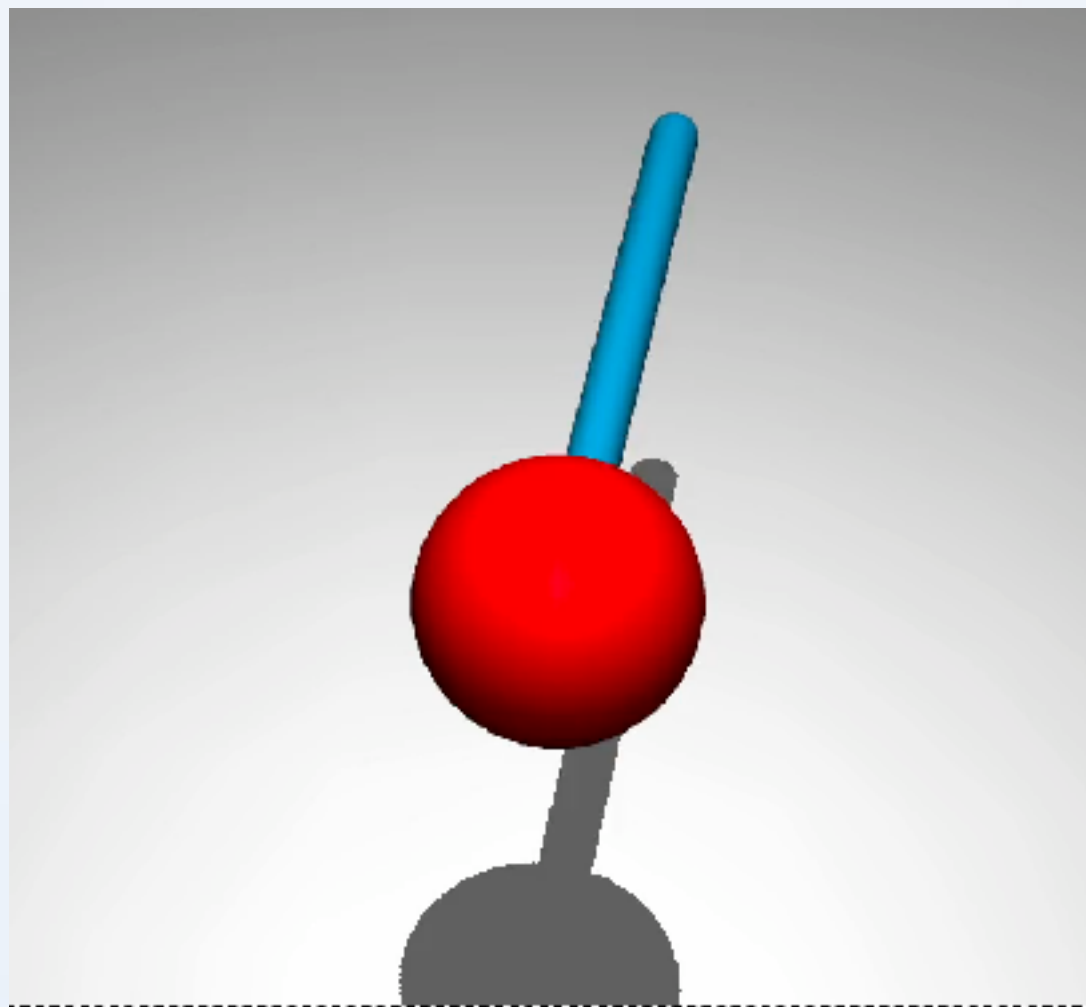
● 机器人的两个核心概念

Joint (关节)

Link (连杆/刚体)



例子





例子

URDF

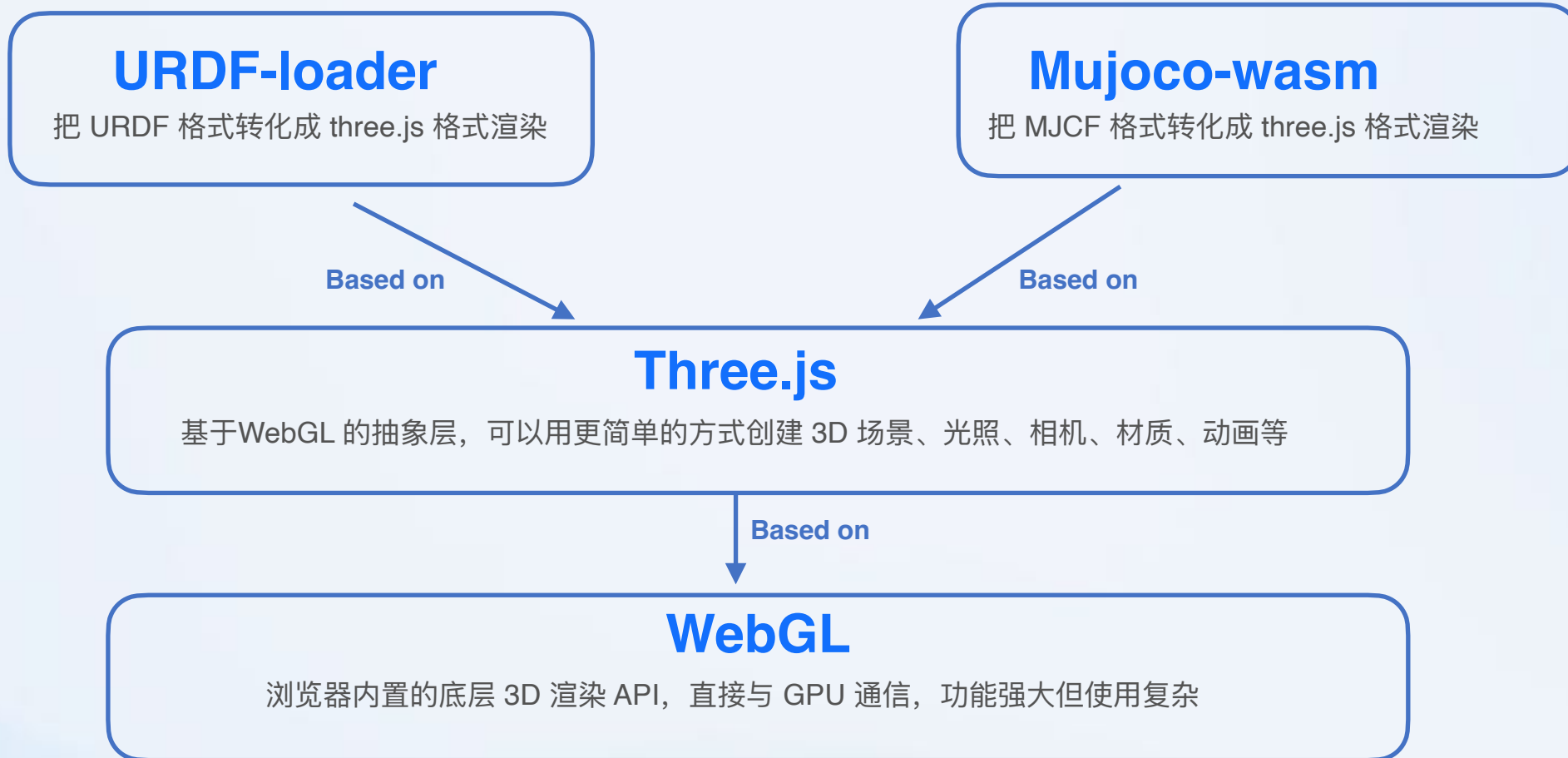
```
model.xml
1 <?xml version="1.0"?>
2 <robot name="ball_pendulum">
3
4   <!-- Link 1: 小球 -->
5   <link name="ball">
6     <visual><geometry><sphere radius="0.1"/></geometry></visual>
7   </link>
8
9   <!-- Link 2: 杆子 -->
10  <link name="rod">
11    <visual><geometry><box size="0.5 0.04 0.04"/></geometry></visual>
12  </link>
13
14  <!-- Joint: 绕 Z 轴旋转 -->
15  <joint name="joint1" type="continuous">
16    <parent link="ball"/>
17    <child link="rod"/>
18    <axis xyz="0 0 1"/>
19  </joint>
20 </robot>
```

MJCF

```
1 <mujoco model="ball_pendulum">
2
3   <worldbody>
4     <!-- 固定的小球 -->
5     <body name="ball" pos="0 0 0.5">
6       <geom type="sphere" size="0.1" rgba="1 0 0 1"/>
7
8     <!-- 可以绕球转的杆子 -->
9     <body name="rod" pos="0 0 0">
10      <!-- 关节: 绕 Z 轴旋转 -->
11      <joint name="joint1" type="hinge" axis="0 0 1"/>
12
13      <!-- 杆子 (从球心向外延伸) -->
14      <geom type="capsule" fromto="0 0 0 0.5 0 0"
15        size="0.02" rgba="0.2 0.6 0.8 1"/>
16    </body>
17  </worldbody>
18 </mujoco>
```



如何在浏览器里面渲染



<https://github.com/mrdoob/three.js>

https://github.com/zalo/mujoco_wasm

<https://github.com/gkjohnson/urdf-loaders>

BamBot

Record Dataset

Duration: 0:12.1

Frames: 320/606

Start

Pause

Stop

Stop Replay

?



Joint Controls

Joint	Angle	Control
FL_hip_joint	0.0°	Q 1
FL_thigh_joint	24.0°	W 2
FL_calf_joint	-48.0°	E 3
FR_hip_joint	0.0°	R 4
FR_thigh_joint	-78.2°	T 5
FR_calf_joint	-84.3°	Y 6
RL_hip_joint	0.0°	Z A
RL_thigh_joint	24.0°	X S
RL_calf_joint	-48.0°	C D
RR_hip_joint	0.0°	V F
RR_thigh_joint	24.0°	B G
RR_calf_joint	-48.0°	N H

Connect Follower Robot

?

3.2

连接各种机器人和传感器

浏览器的能力

Web serial API

通过串行端口与设备通信



连接舵机驱动板、ESP32、Arduino 等开发板，读取传感器数据

连接电机/舵机

Web HID API

与 HID (Human Interface Device) 设备交互



游戏手柄、操纵杆、键盘、鼠标
(特殊功能 3D 空间鼠标)

连接控制器

Web USB API

访问 USB 设备



USB 摄像头，3D 打印机等

连接相机

3.3 多种 AI 模型适配

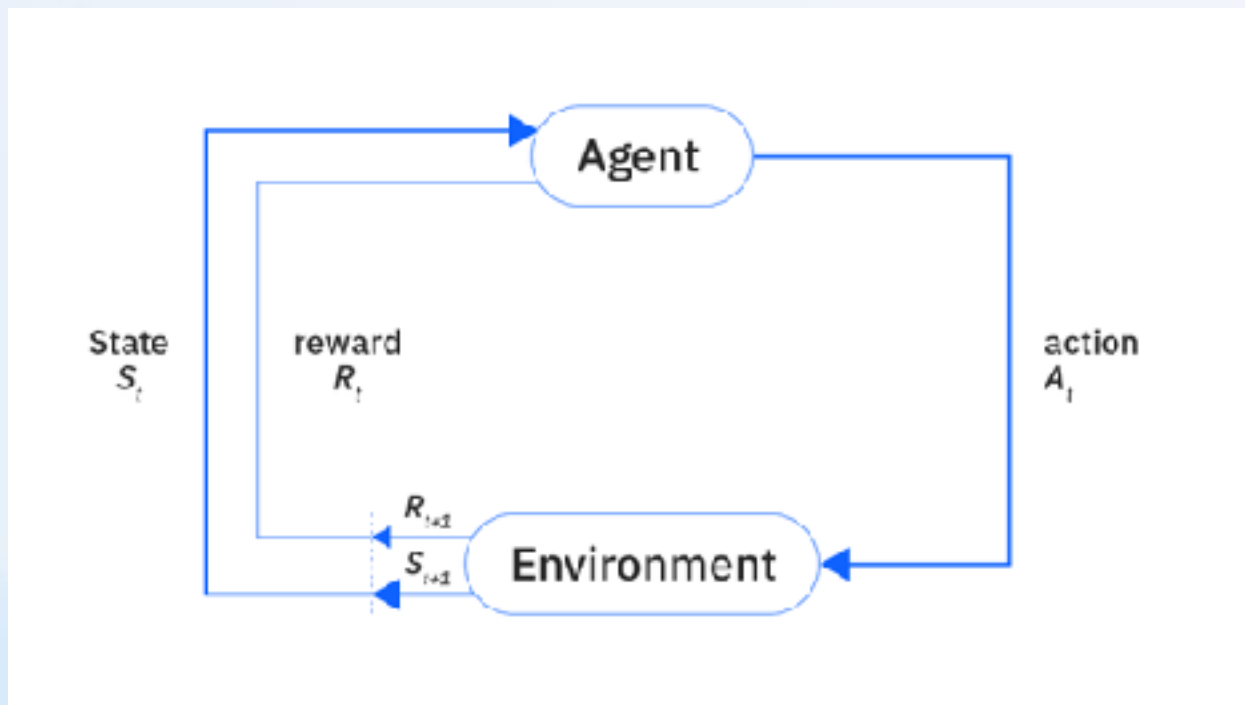
RL

ACT/Diffusion policies

LLM

VLA

Reinforcement Learning (强化学习)

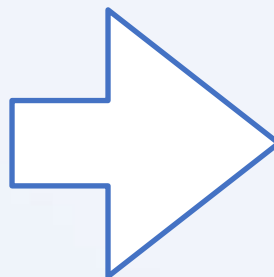
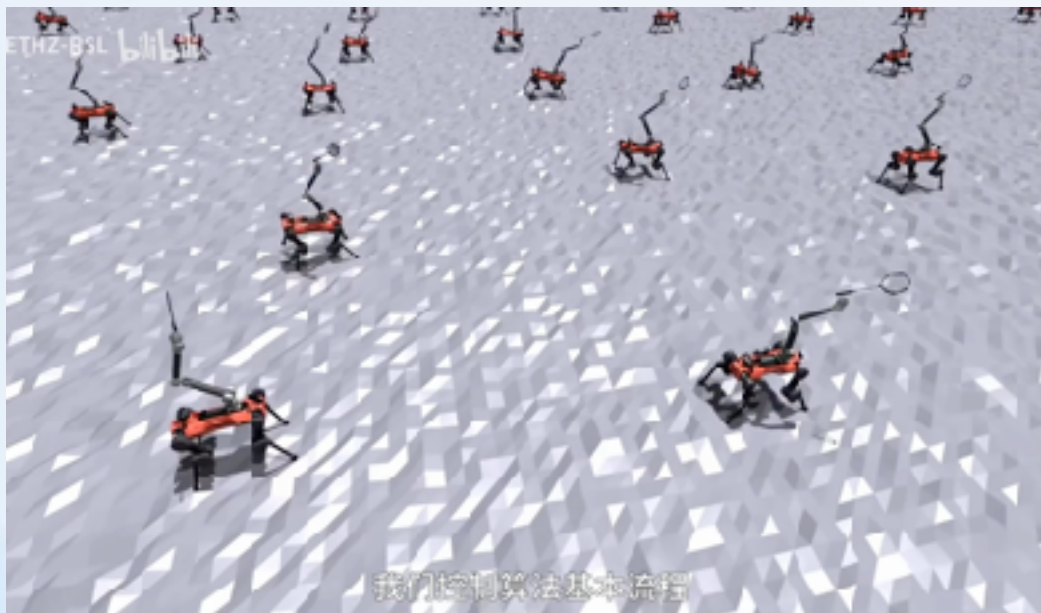


试错 + 奖励反馈循环:

1. 机器人观察当前状态(如关节角度、传感器数据)
2. 根据策略选择动作(如移动、抓取)
3. 环境返回新状态和奖励信号(完成任务得正奖励,失败得负奖励)
4. 机器人根据奖励调整策略,重复这个过程

例子 – 机器人学打羽毛球

训练模型



真机部署



2025 年 3 月 - ETH Zurich (苏黎世联邦理工学院)

<https://www.science.org/doi/epdf/10.1126/scirobotics.adu3922>

Reinforcement Learning (强化学习)

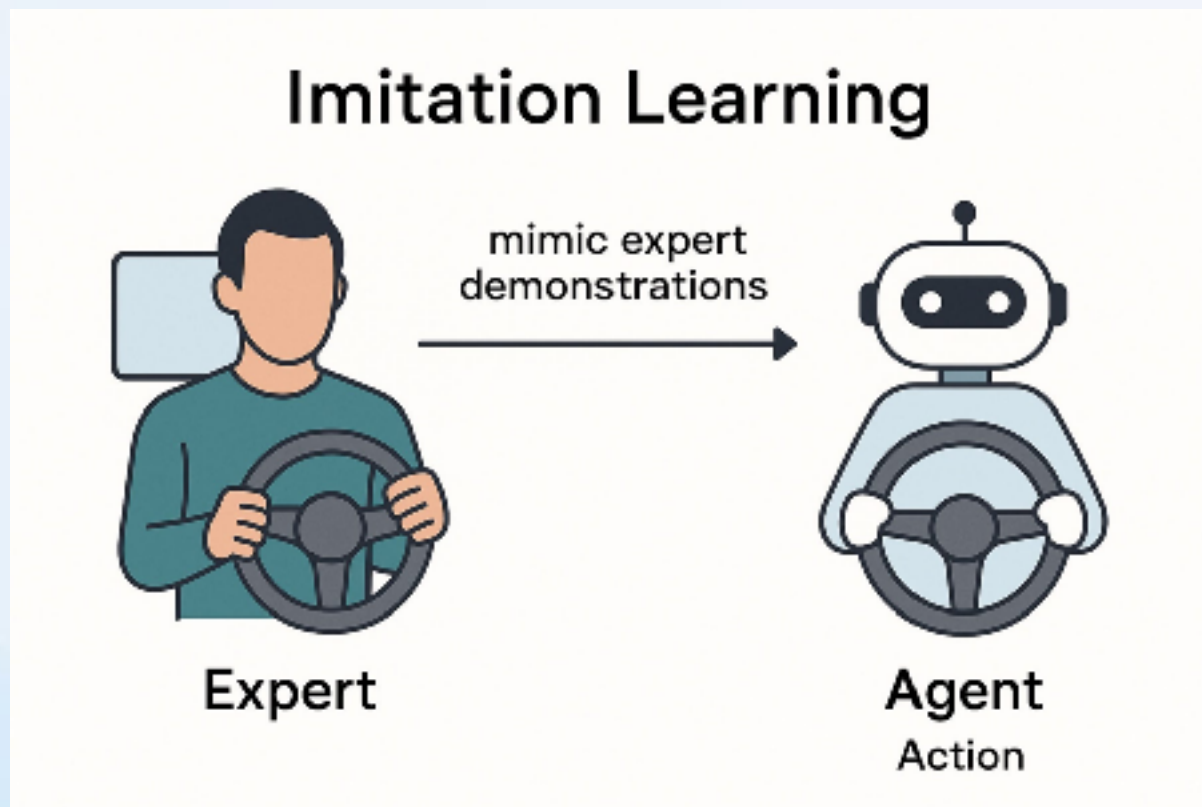
主要优势

- 可以自主学习，无需示范
- 潜力超越人类：理论上可以找到最优策略

主要局限

- 迁移学习困难

ACT / Diffusion Policies (模仿学习)



使用场景

- 难以定义奖励函数：任务目标难以量化
 - 依赖专家示范：需要高质量的演示数据
- 典型应用领域
- 自动驾驶
 - 机器人操作：抓取、装配等复杂任务
 - 人机交互行为模仿

例子 – Aloha / Mobile Aloha

示例教学



部署模型



2024 年 1 月 Stanford (斯坦福)
<https://mobile-aloha.github.io/>



ACT / Diffusion Policies (模仿学习)

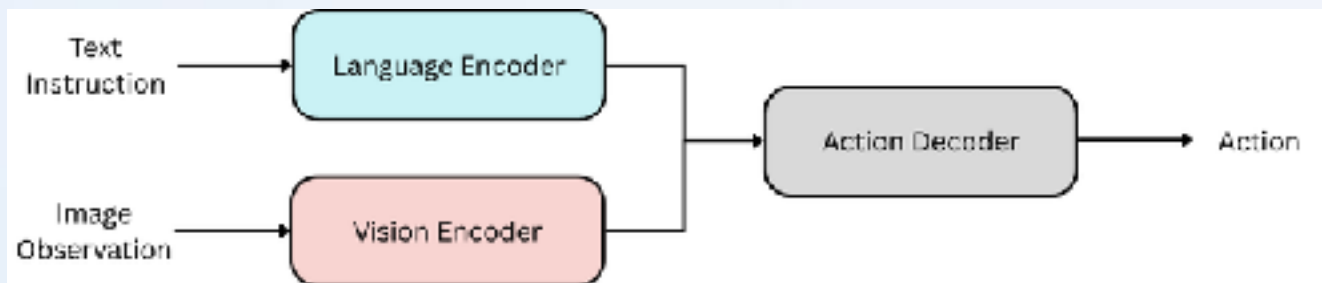
主要优势

- 学习效率高：直接从专家经验学习，避免大量试错
- 适合复杂任务：无需精确建模奖励函数

主要局限

- 依赖数据质量：专家示范的质量直接影响学习效果
- 难以超越示范：模型表现上限受限于专家水平

VLA (Vision-Language-Action)



端到端的机器人学习架构
将视觉感知、语言理解和动作执行
统一在一个模型中

目标是模型的通用化

基本架构

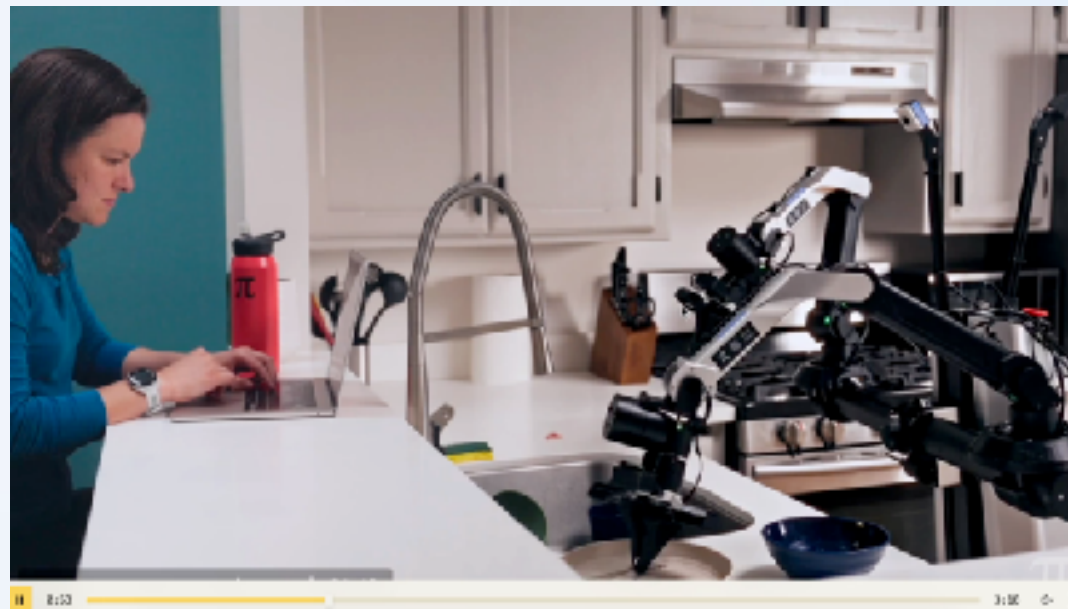
多模态输入融合

- 视觉编码器: 处理相机图像,提取视觉特征(通常使用预训练的视觉Transformer如ViT)
- 语言编码器: 处理自然语言指令,将其编码为语义表示(通常使用预训练的语言模型)
- 融合机制: 将视觉和语言特征进行交互融合,建立指令与场景的对应关系

动作生成

- 融合后的特征通过解码器生成机器人的动作序列
- 动作通常表示为关节角度、末端执行器位置,或离散/连续的控制信号

例子：pi-0.5



2025 年 4月 Physical Intelligence

<https://www.physicalintelligence.company/blog/pi05>



VLA (Vision–Language–Action)

典型应用领域

- 通用机器人操作：基于语言指令的多样化任务
- 具身智能：家庭服务、仓储等机器人

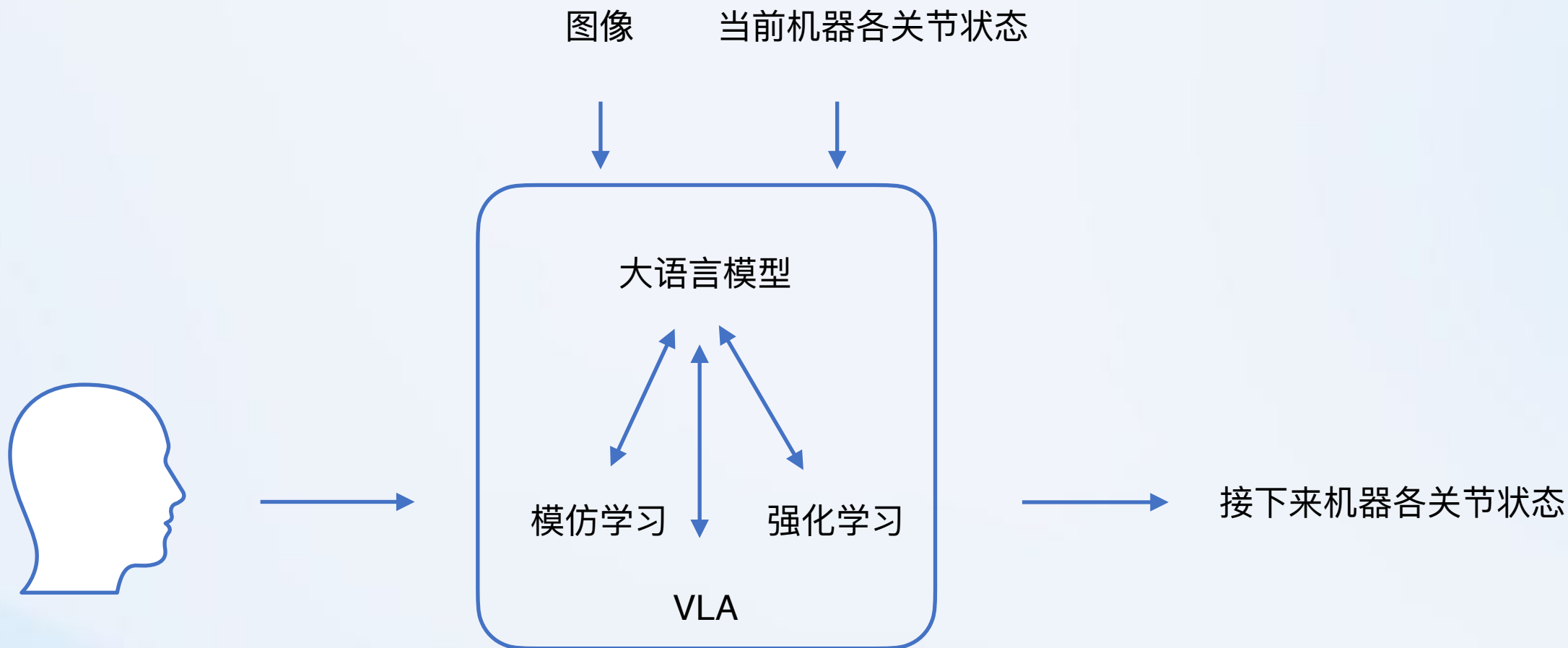
主要优势

- 泛化能力强：可执行未见过的任务指令
- 自然交互：通过自然语言指定任务
- 知识迁移：预训练知识可迁移到动作空间

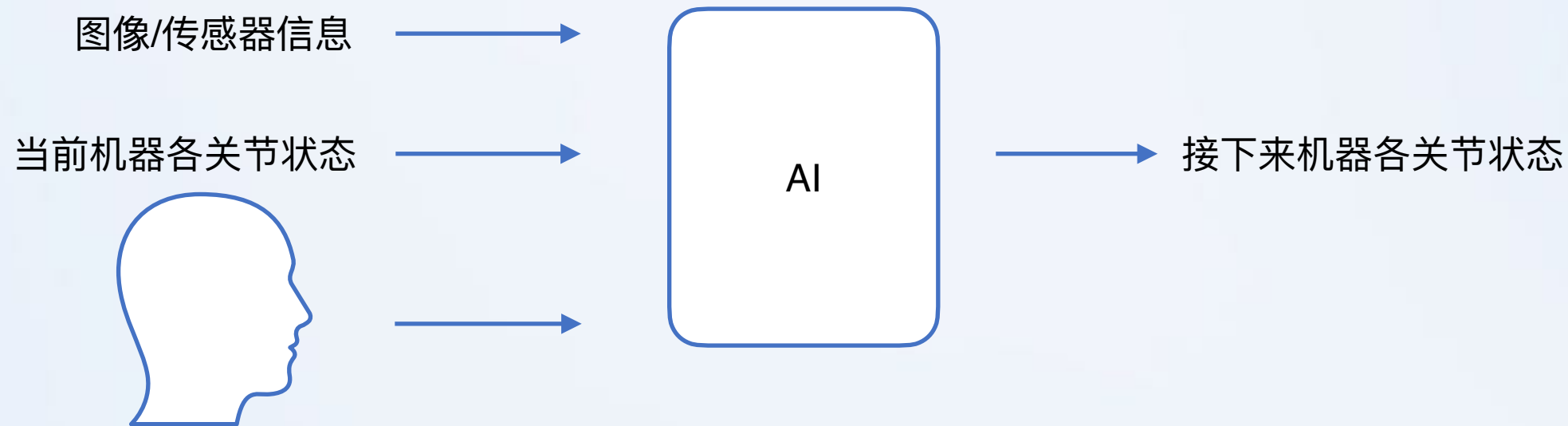
主要局限

- 计算资源需求大：训练和推理成本高
- 实时性挑战：推理速度影响控制频率
- 精细操作能力有限：高精度任务表现不如专用模型

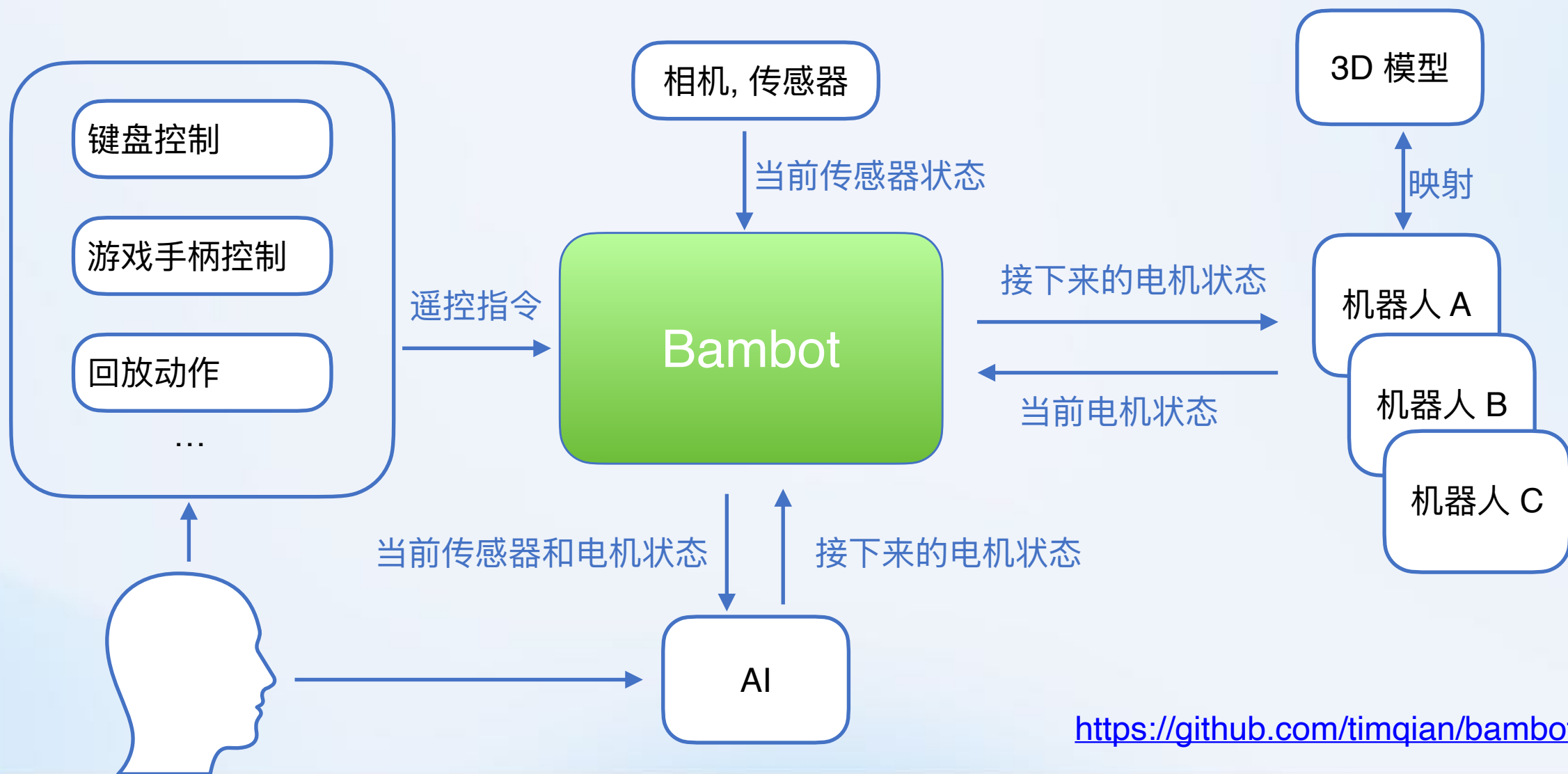
混合模型



模型的共性



BamBot 架构



<https://github.com/timqian/bambot>

两朵乌云

无法直接在浏览器运行 AI 模型，需要通过 API 与模型交互

基于 URDF-loader，只有 3D 渲染，没有物理仿真，无法用于模型的训练和仿真

Bambot

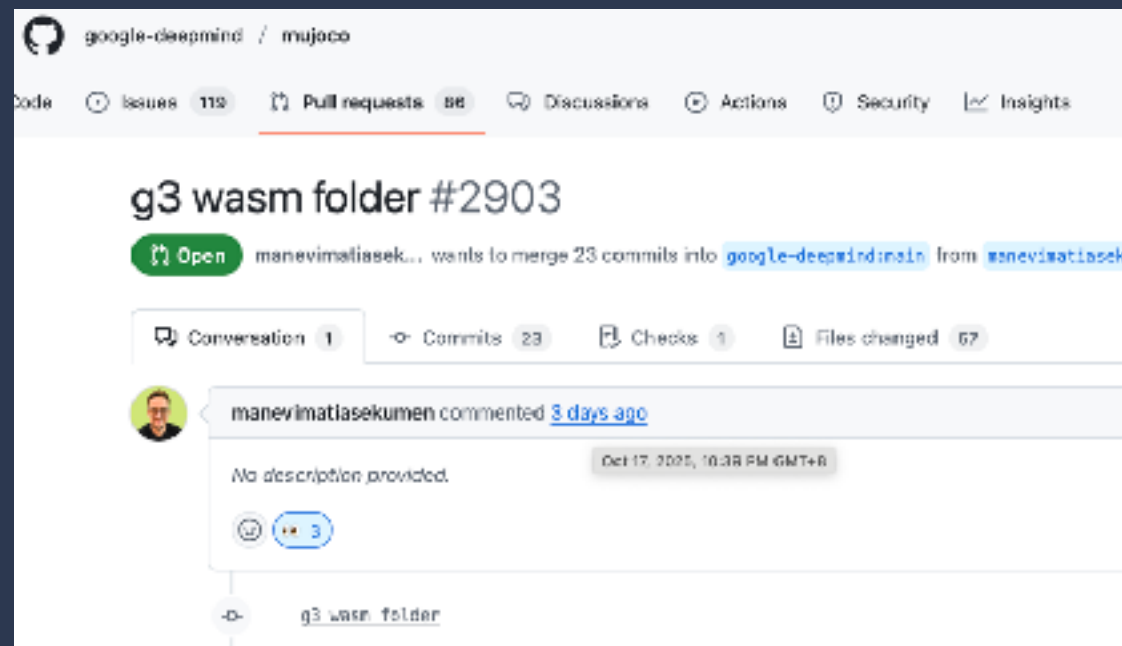
04 最新进展

浏览器直接跑 AI 模型：ONNX Runtime



2025 年 10月 Tatsuki Tsujimoto
<https://github.com/ttkjtjmt/muwanx>

Mujoco in the browser (官方支持)



- Issue: <https://github.com/google-deepmind/mujoco/issues/2585#issuecomment-3213407529>
- PR: <https://github.com/google-deepmind/mujoco/pull/2903>

不远的未来

1. 无需代码就可以通过示教的方式教机器人工作
2. RL 模型仍需在浏览器之外训练
3. VLA 模型不断进步，兑现承诺

BamBot 的未来 – 开源项目的商业化挑战

1. 制作更好的，更便宜的机器卖给普通用户？
2. 聚焦于软件，与机器人厂商合作接入更多机器？
3. 与有机器人自动化需求的企业合作？

📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AI Ops
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AICon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AICon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AICon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AICon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LM Ops
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

THANKS

<https://github.com/timqian/bambot>

