上海大学 SHANGHAI UNIVERSITY

上海大学未来技术学院 SCHOOL OF FUTURE TECHNOLOGY, SHANGHAI UNIVERSITY

上海大学人工智能研究院 SCHOOL OF ARTIFICIAL INTELLIGENCE SHANGHAI UNIVERSITY

人形机器人控制
——从过去到未来

叶林奇 （上海大学，副研究员）

https://linqi-ye.github.io/

THU-SHU ROBOART LAB

人形机器人控制
从过去到未来

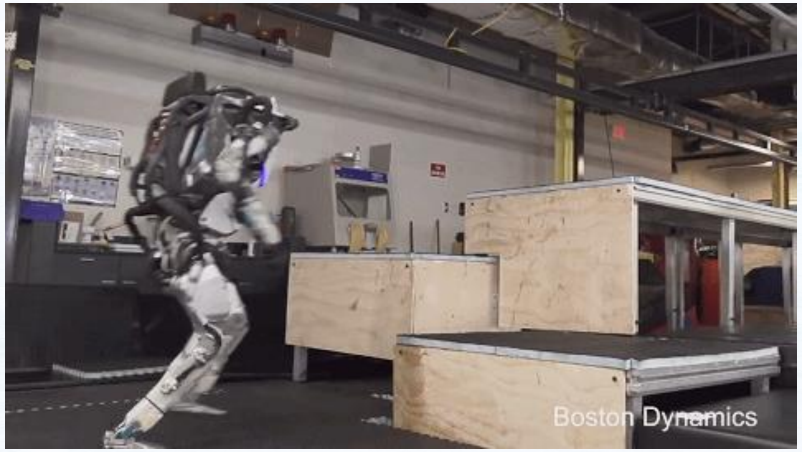
目录

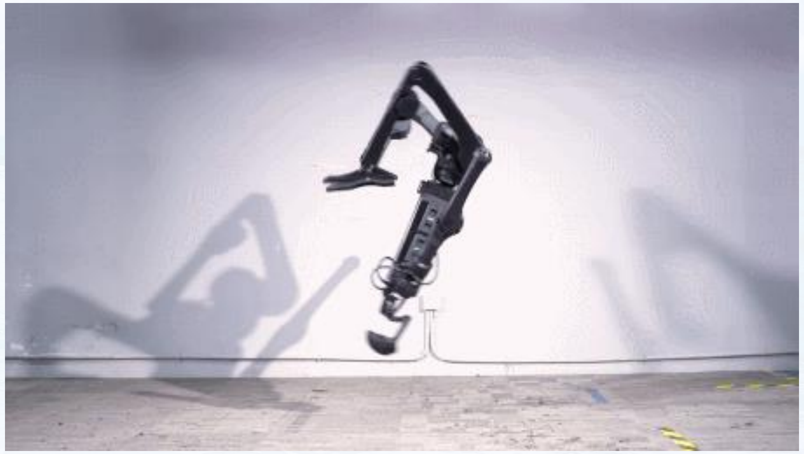
人形机器人控制发展历程

日本
本田公司

美国
波士顿动力

中国
宇树科技





Asimo

Atlas

H1

人形机器人控制发展历程

简化主义时代	优化主义时代	学习主义时代
Static Walking	Model Predictive Control	Reinforcement Learning
Passive Walking	Whole Body Control	Imitation Learning
Central Pattern Generator		Online Learning
Zero Moment Point		
Hybrid Zero Dynamics		
Virtual Model Control		

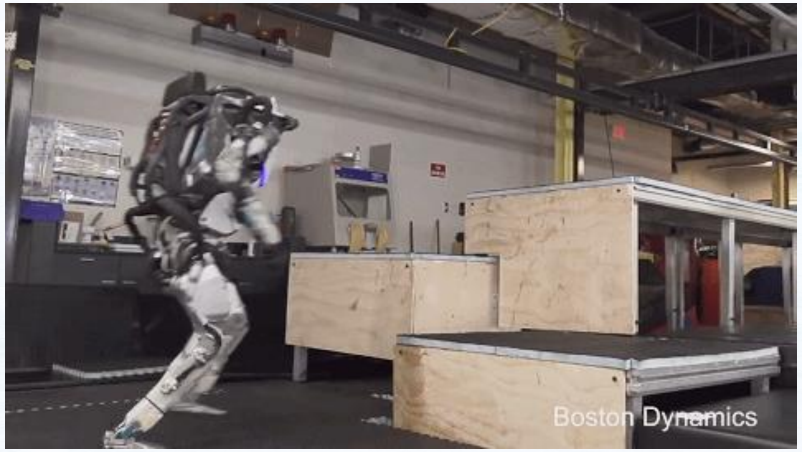
人形机器人控制发展历程

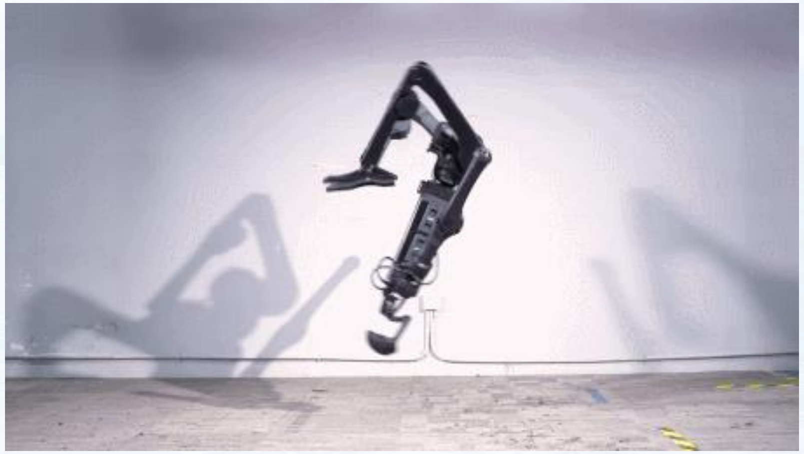
THU-SHU ROBOART LAB

简化方法

优化方法

学习方法





零力矩点控制

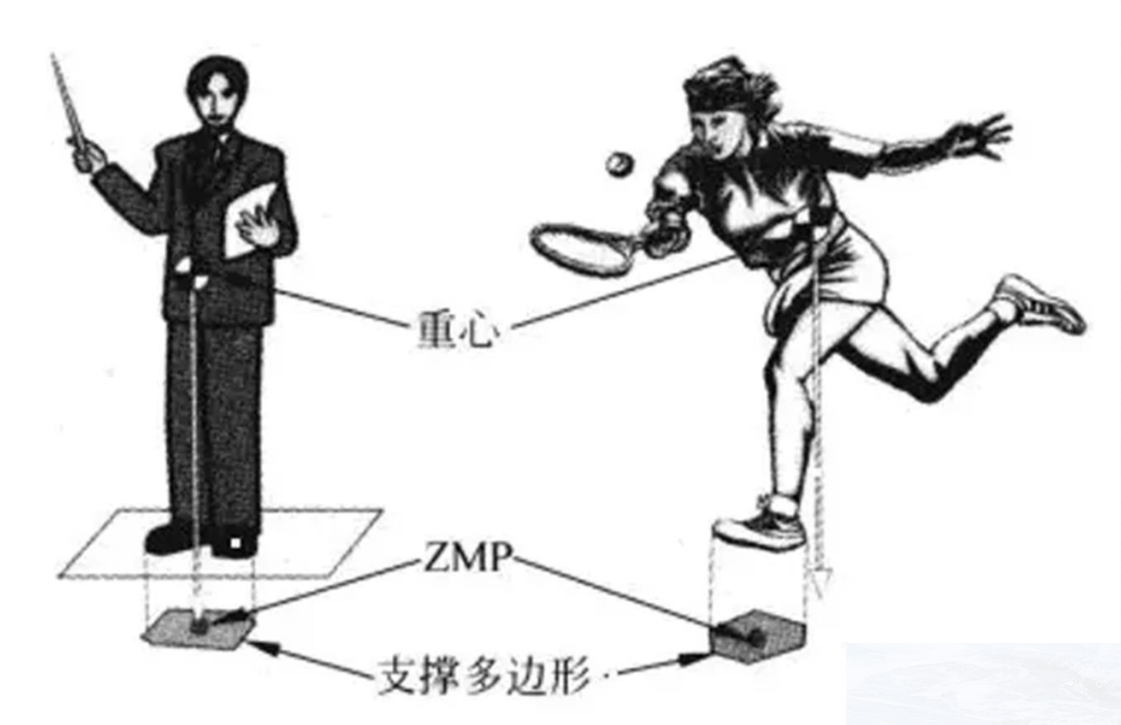
模型预测控制

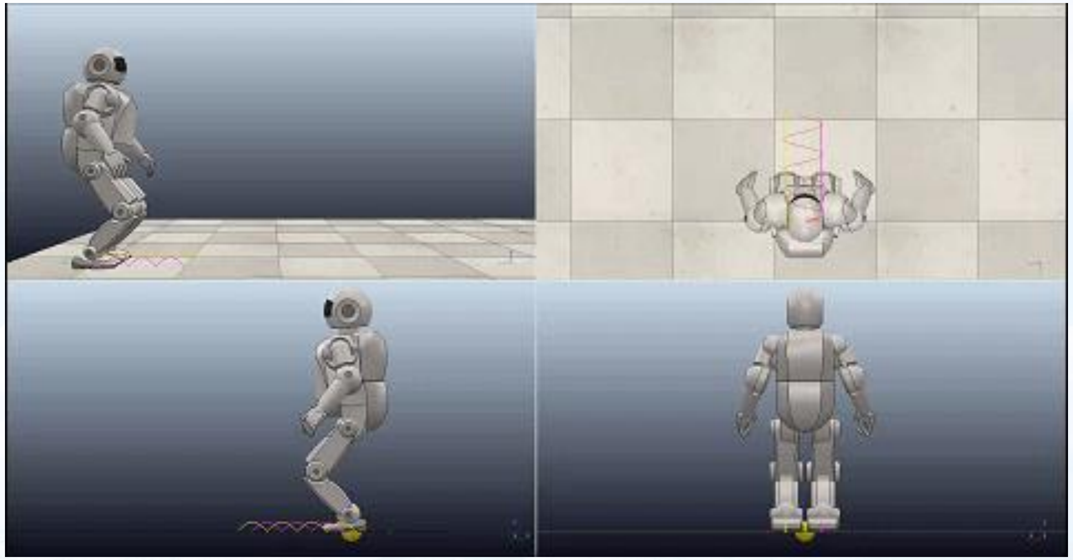
强化学习控制

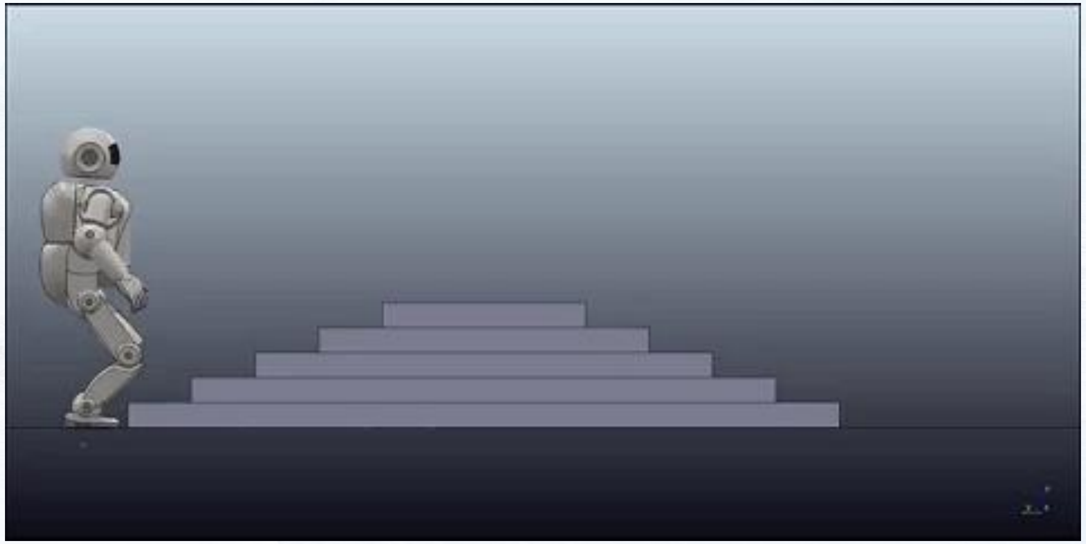
人形机器人控制从过去到未来

目录

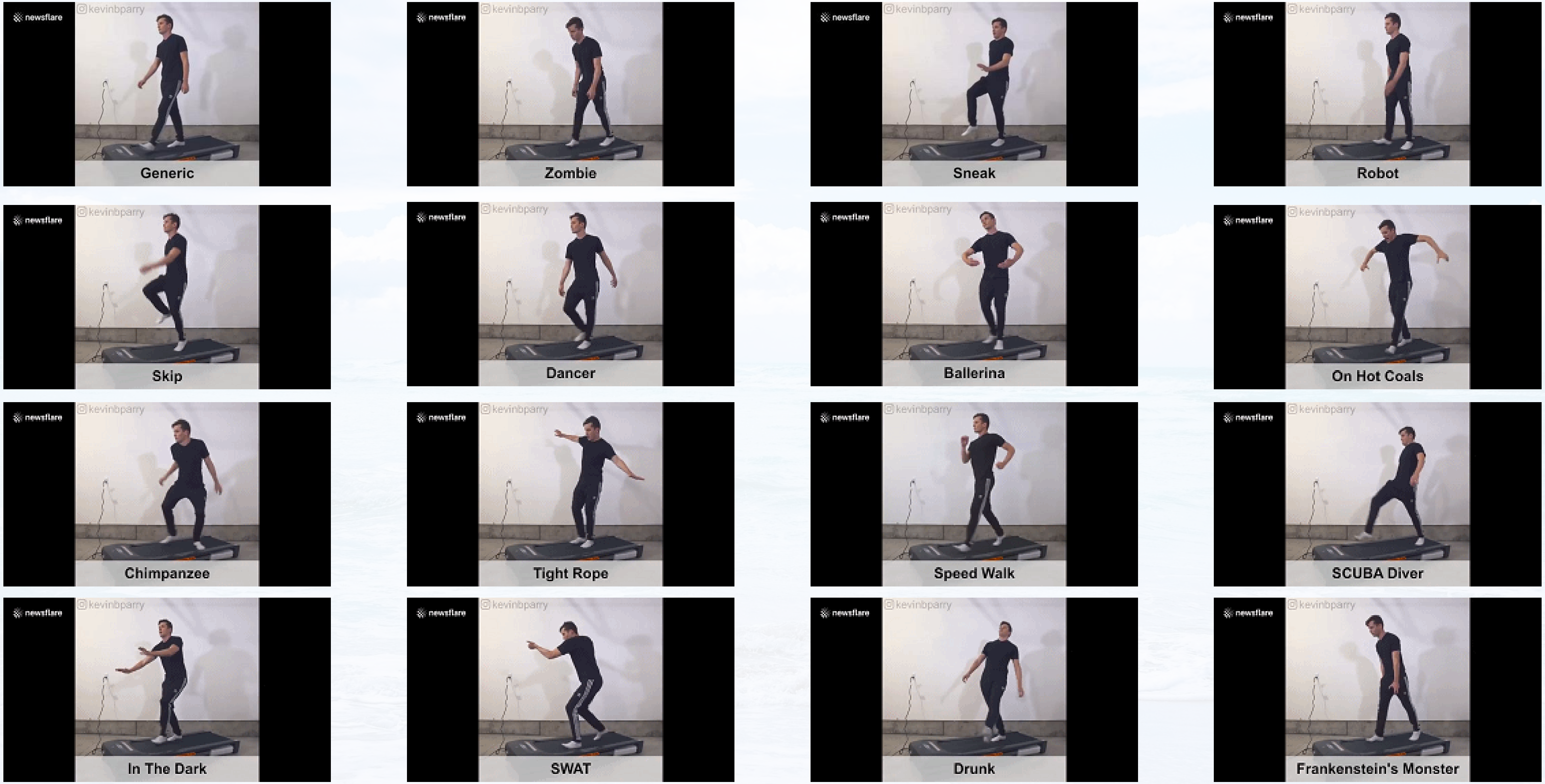
静态行走







人类行走


newsflare
@kevinbparry
Generic
Zombie
Sneak
Robot
Skip
Dancer
Ballerina
On Hot Coals
Chimpanzee
Tight Rope
Speed Walk
SCUBA Diver
In The Dark
SWAT
Drunk
Frankenstein's Monster

倒立摆模型

cost of transport (CoT): $\dfrac{\text{energy used}}{\text{weight} \times \text{distance traveled}}$

$$C = \int_0^{t_{\text{step}}} [F(t)\dot{l}]^+ \mathrm{d}t / mgd$$

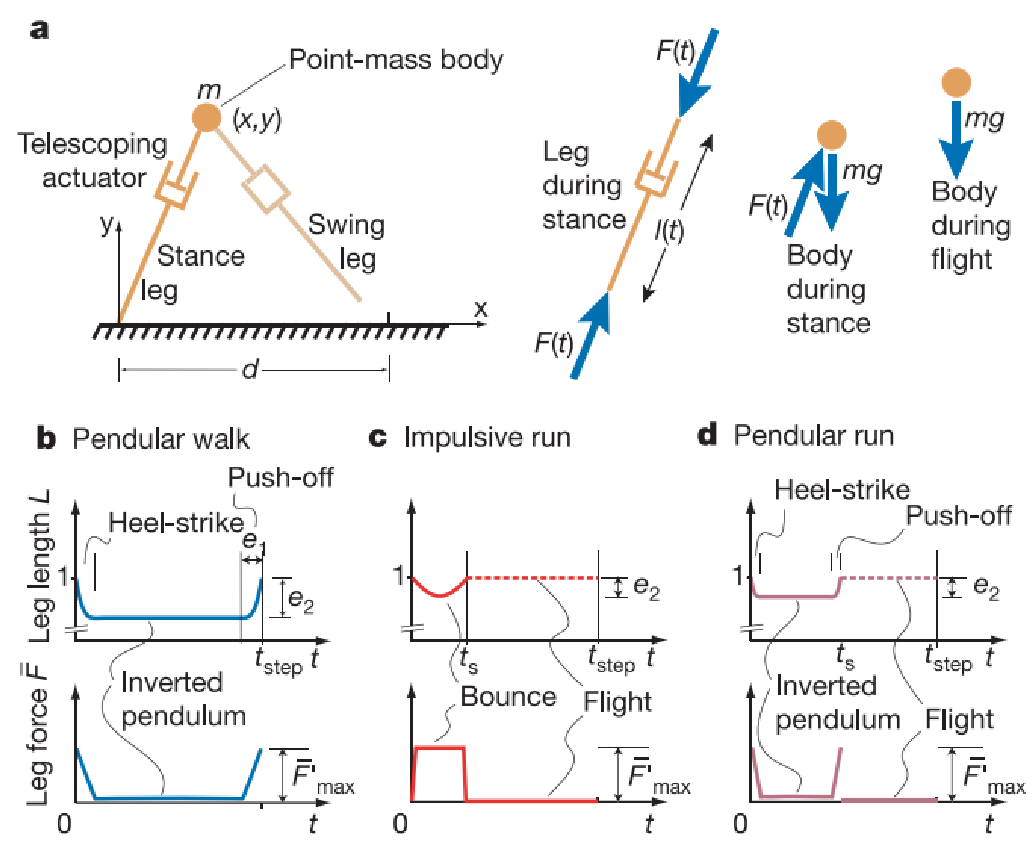


Figure 2 | Point-mass biped model and its optimal solutions.

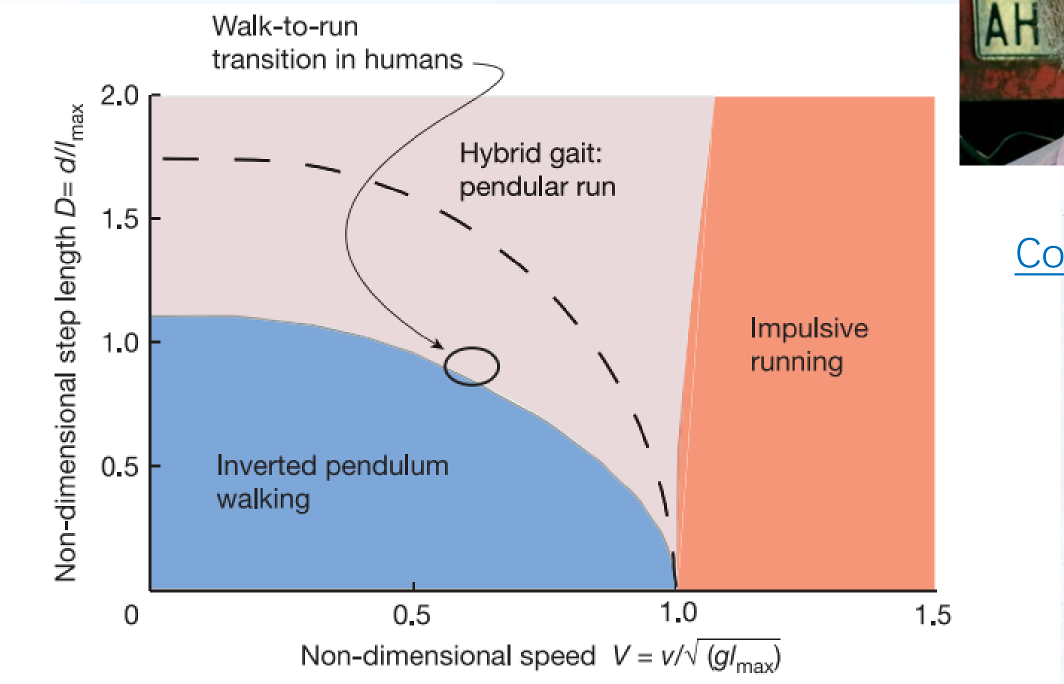
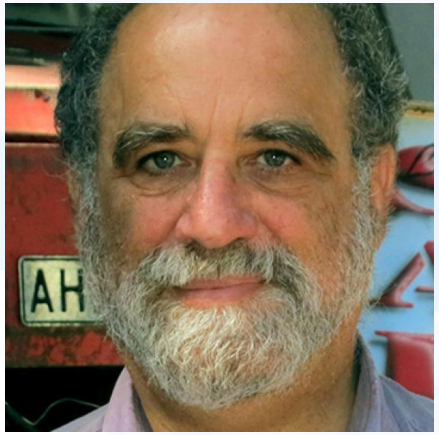


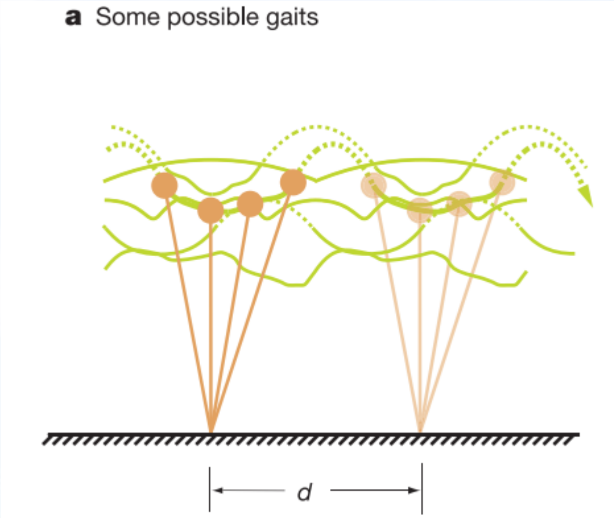
Figure 3 | The regions in which each of the three collisional gaits are optimal.

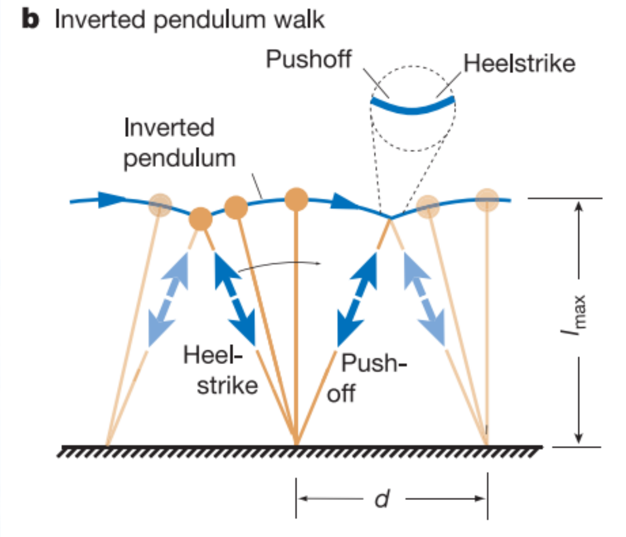


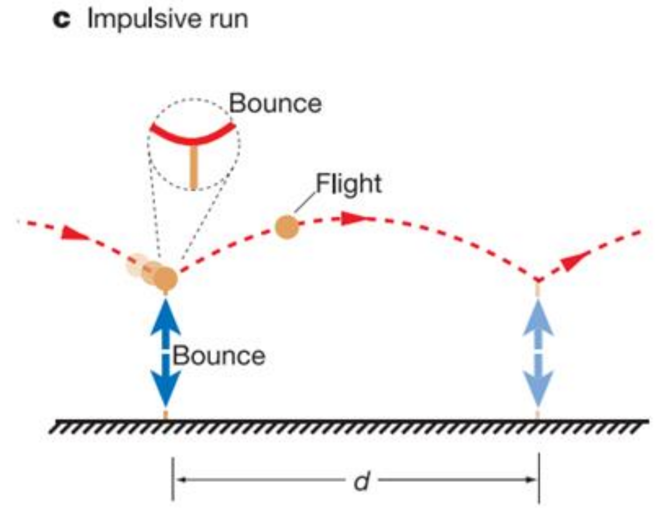
Andy Ruina
Cornell University

Srinivasan, Manoj, and Andy Ruina. "Computer optimization of a minimal biped model discovers walking and running." *Nature* 439.7072 (2006): 72-75.

倒立摆模型


a Some possible gaits
d


b Inverted pendulum walk
Pushoff
Heelstrike
Inverted pendulum
Heel-strike
Push-off
l_{max}
d

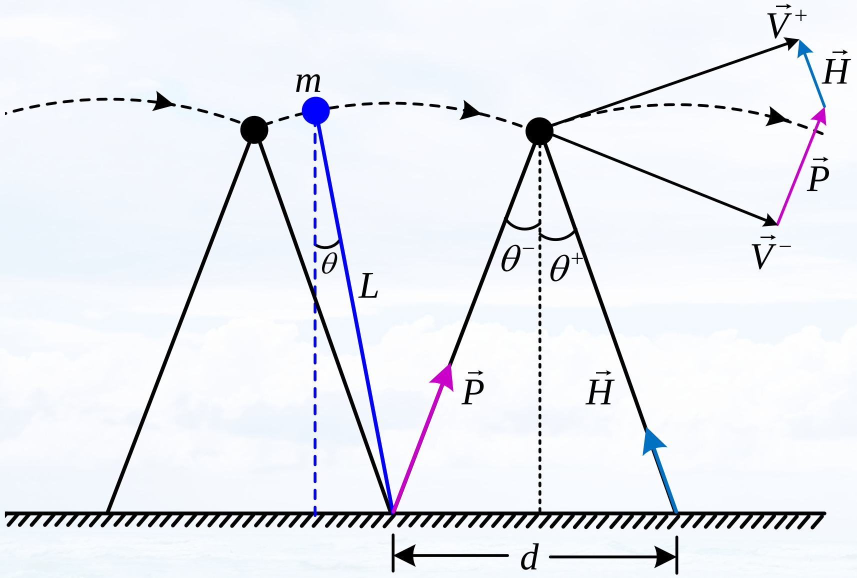

c Impulsive run
Bounce
Flight
Bounce
d

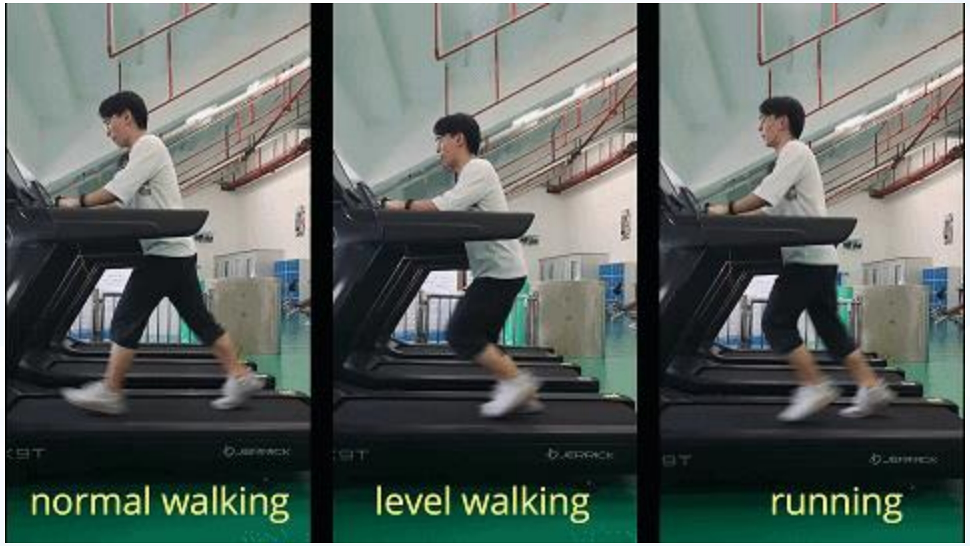

+114


200
800
3000

倒立摆模型

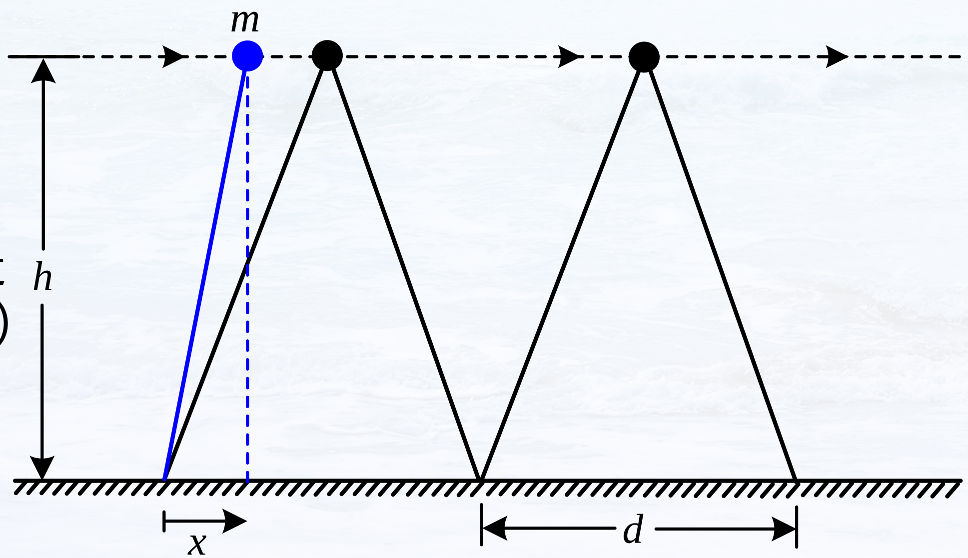
正常行走
(IP模型)





10km/h

水平行走
(LIP模型)

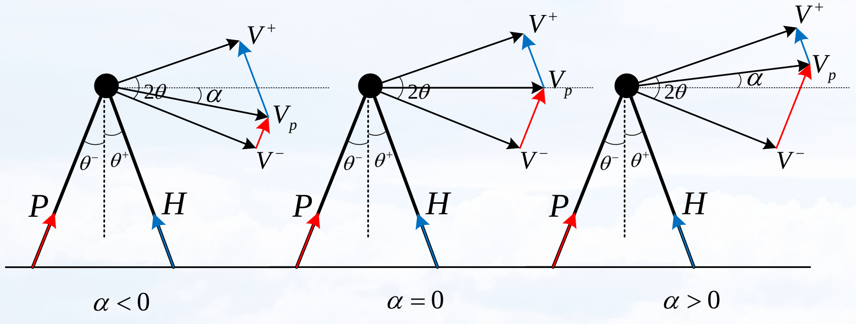


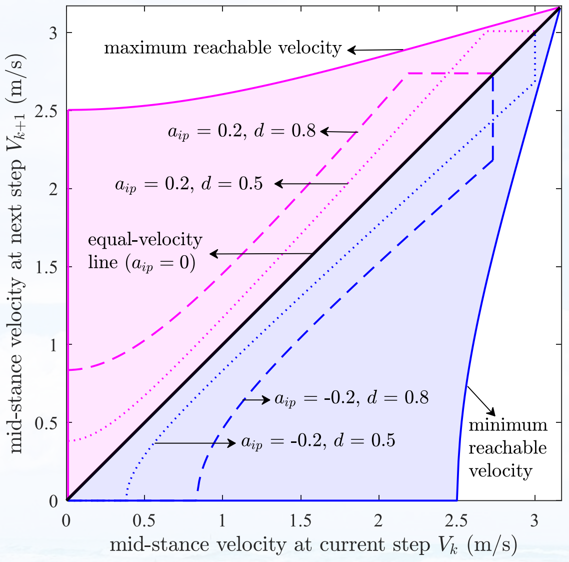
15km/h

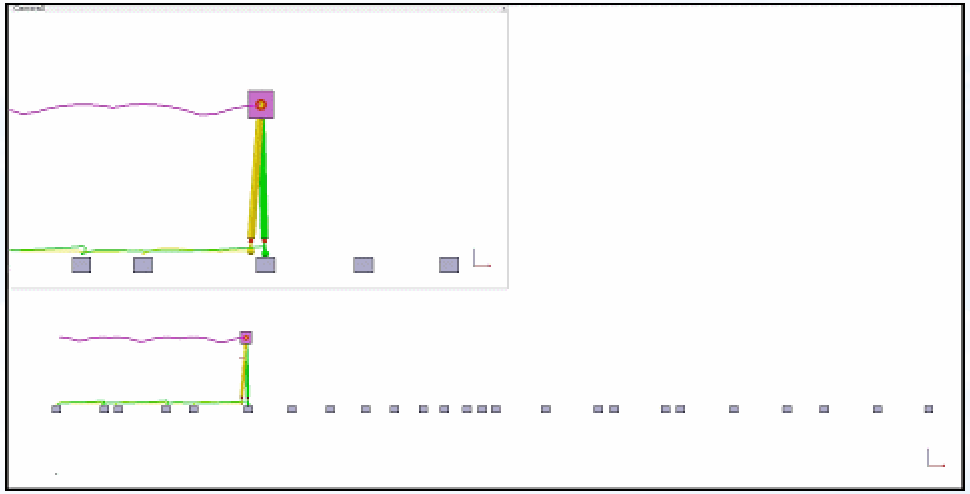
倒立摆模型

THU-SHU ROBOART LAB

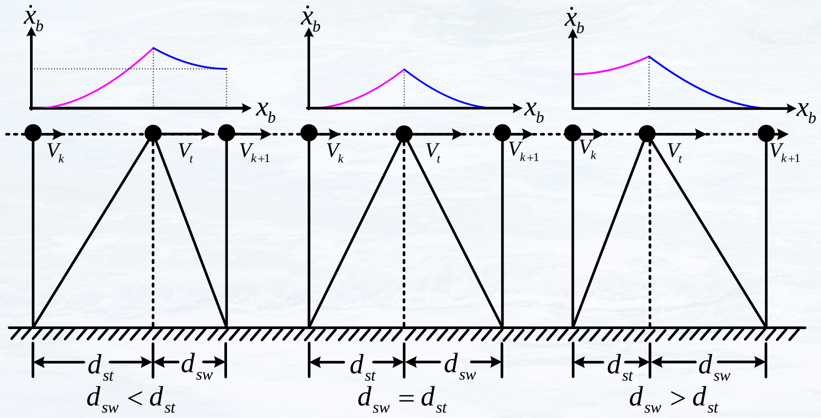
正常行走（IP模型）

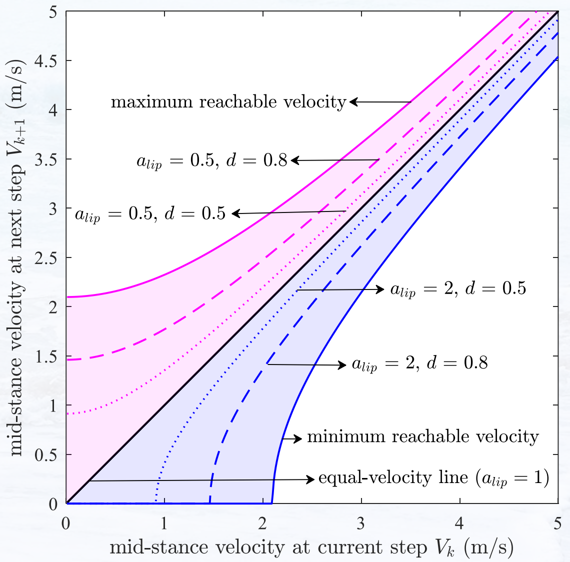


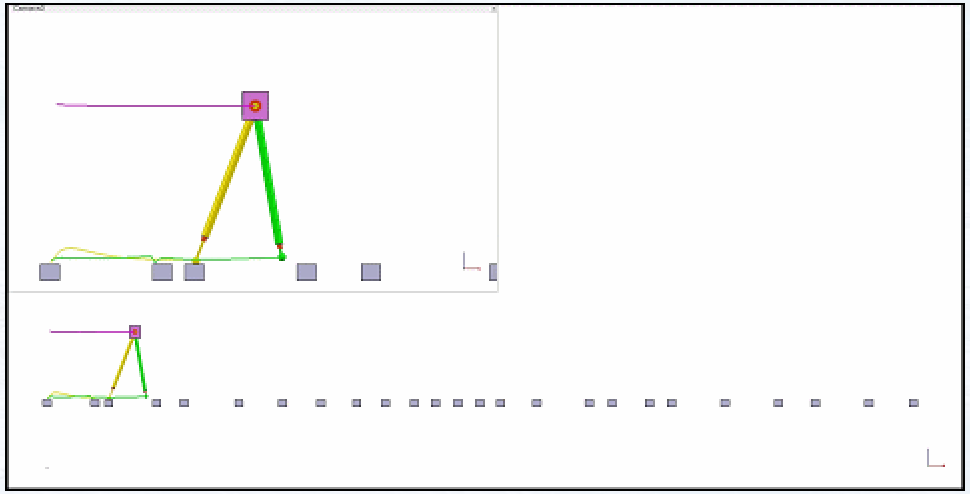




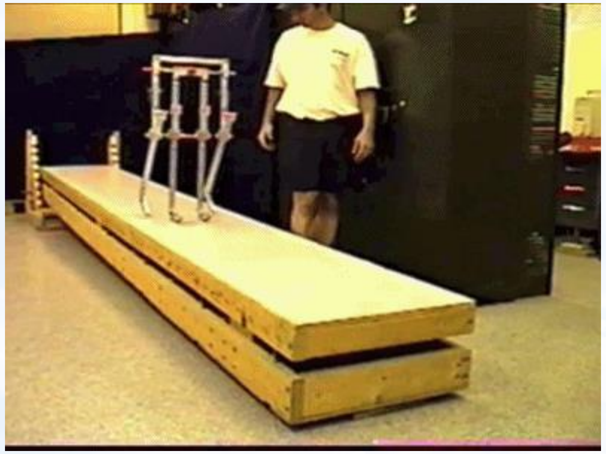
水平行走（LIP模型）







被动行走



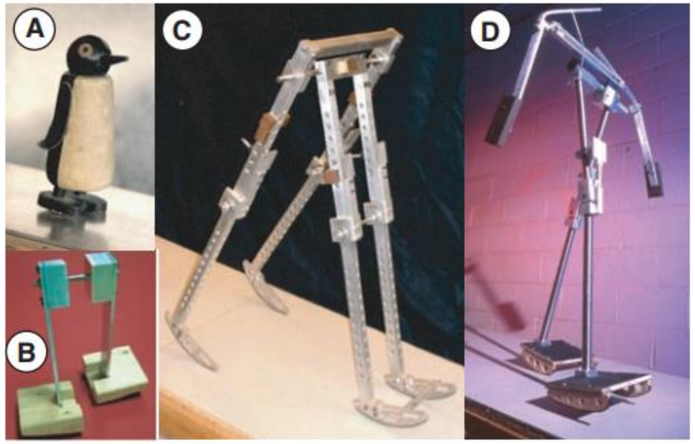


Fig. 1. "Ramp-walking," "downhill," "unpowered," or "passive-dynamic" machines. Our powered bipeds are based on these passive designs. **(A)** The Wilson "Walkie" (*27*). **(B)** MIT's improved version (*28*). Both (A) and (B) walk down a slight ramp with the "comical, awkward, waddling gait of the penguin" (*27*). **(C)** Cornell copy (*29*) of McGeer's capstone design (*7*). This four-legged "biped" has two pairs of legs, an inner and outer pair, to prevent falling sideways. **(D)** The Cornell passive biped with arms [photo: H. Morgan]. This walker has knees and arms and is perhaps the most humanlike passive-dynamic walker to date (*8*).

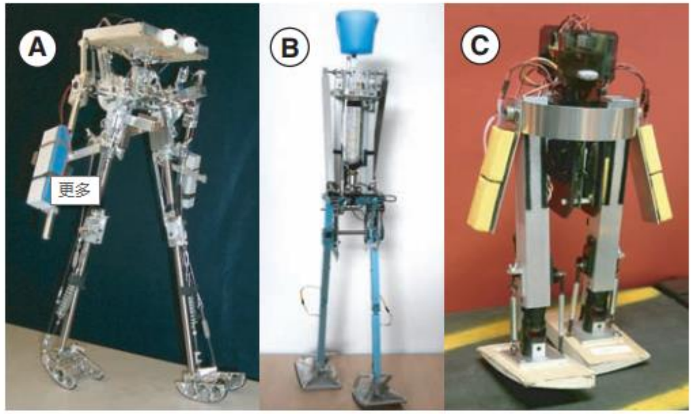


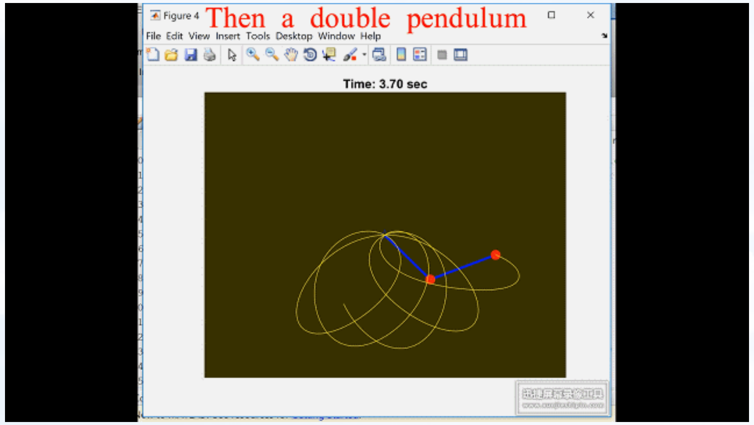
Fig. 2. Three level-ground powered walking robots based on the ramp-walking designs of Fig. 1. **(A)** The Cornell biped. **(B)** The Delft biped. **(C)** The MIT learning biped. These powered robots have motions close to those of their ramp-walking counterparts as seen in the supporting online movies (movies S1 to S3). Information on their construction is in the supporting online text (*9*).

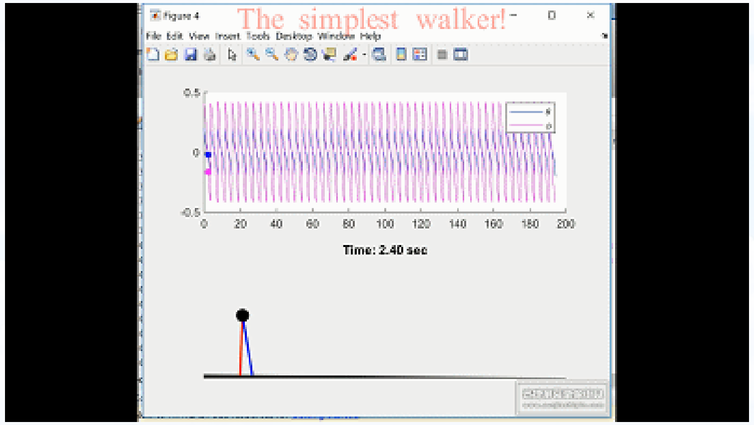
Gliders+Engines→Airplanes

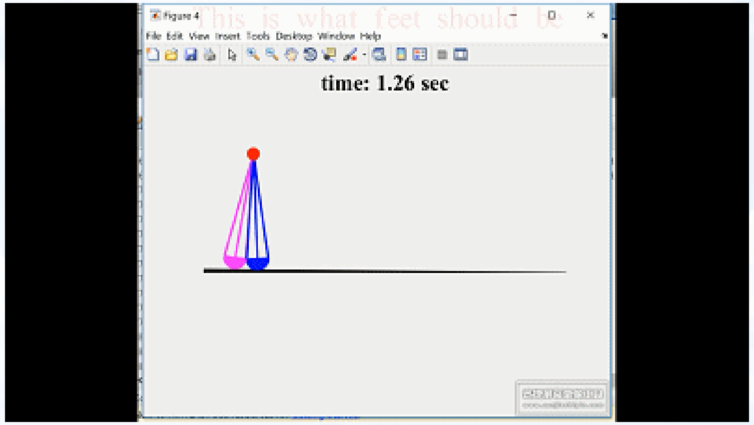
Passive walkers+Actuators→Human-level robot

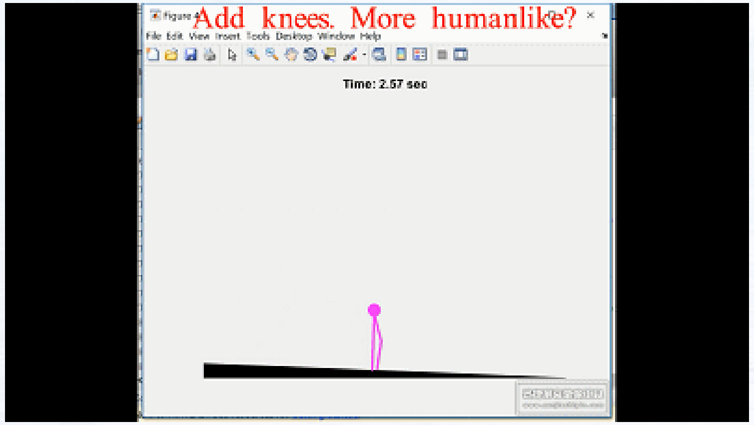
Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science*, *307*(5712), 1082-1085.

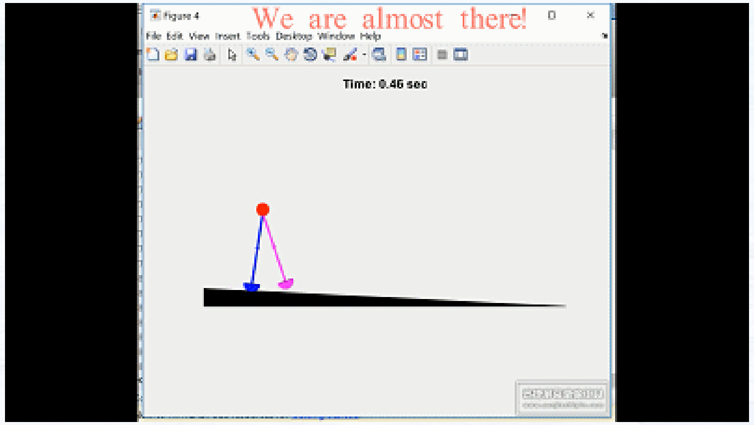
被动行走

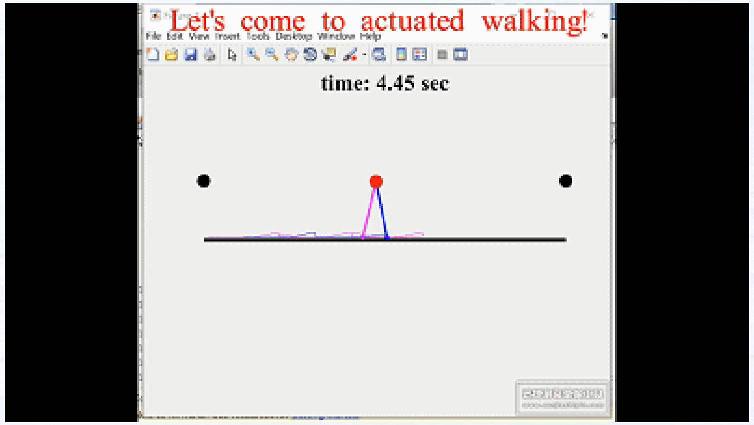

Then a double pendulum
Time: 3.70 sec


The simplest walker!
Time: 2.40 sec


This is what feet should be
time: 1.26 sec

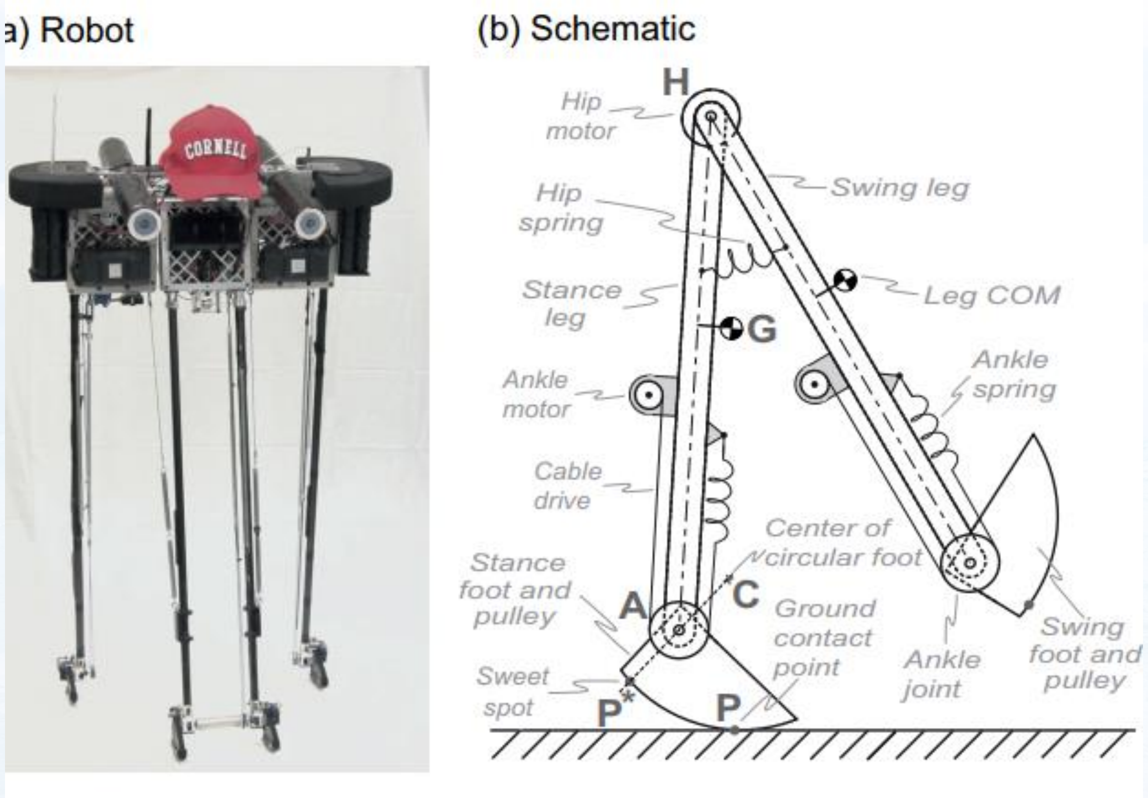

Add knees. More humanlike?
Time: 2.57 sec


We are almost there!
Time: 0.46 sec


Let's come to actuated walking!
time: 4.45 sec

被动行走

Ranger walks non-stop 65.2 km ultra-Marathon on May 1-2, 2011



Bhounsule, P. A., Cortell, J., Grewal, A., Hendriksen, B., Karssen, J. D., Paul, C., & Ruina, A. (2014). Low-bandwidth reflex-based control for lower power walking: 65 km on a single battery charge. *The International Journal of Robotics Research*, *33*(10), 1305-1321.

落脚点控制

The neutral point

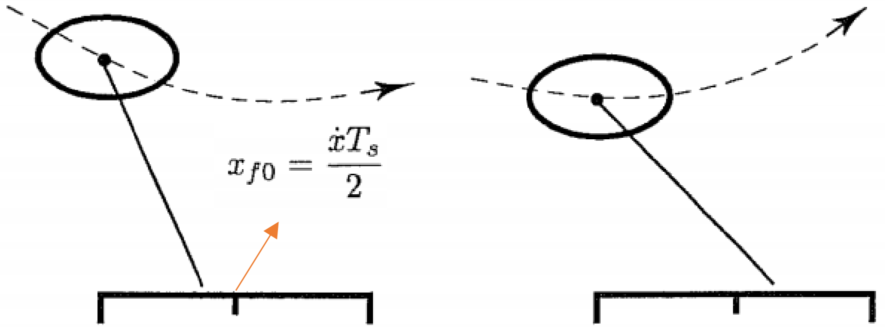


Figure 2.11. Asymmetric trajectories. Displacement of the foot from the neutral position accelerates the body by skewing its trajectory. When the foot is placed behind the neutral point, the body accelerates forward during stance (left). When the foot is place forward of the neutral point, the body accelerates backward during stance (right). Dashed lines indicate the path of the body, and solid horizontal lines under each figure indicate the CG-print.

Three-part control

Hopping:

Thrust for specified duration during stance.

Exhaust to specified pressure during flight.

Forward Speed:

Choose foot position $\quad x_f = \frac{\dot{x}T_s}{2} + k_{\dot{x}}(\dot{x} - \dot{x}_d).$

Convert to hip angle $\quad \gamma_d = \phi - \arcsin\left(\frac{x_f}{r}\right).$

Servo hip angle $\quad \tau = -k_p(\gamma - \gamma_d) - k_v(\dot{\gamma}).$

Body Attitude:

Servo body angle $\quad \tau = -k_p(\phi - \phi_d) - k_v(\dot{\phi}).$

落脚点控制

Balance strategies for a biped:

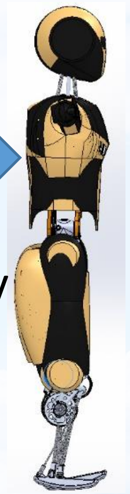
1) Apply ankle torques. Base of support diameter up to 0.2 m

2) Bend the upper body/spin arms.

   Effective base of support up to 0.02 m

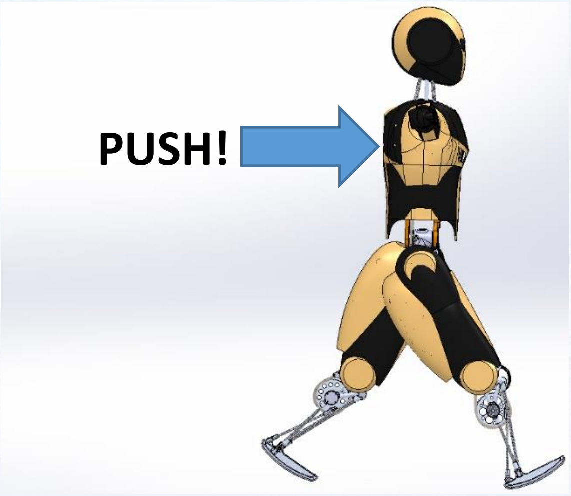
3) Foot placement. Effective base of support up to 1 m

Therefore **robust balance = fast stepping**.



How to select step location?

- Take step in the falling direction
- Falling faster, taking bigger step

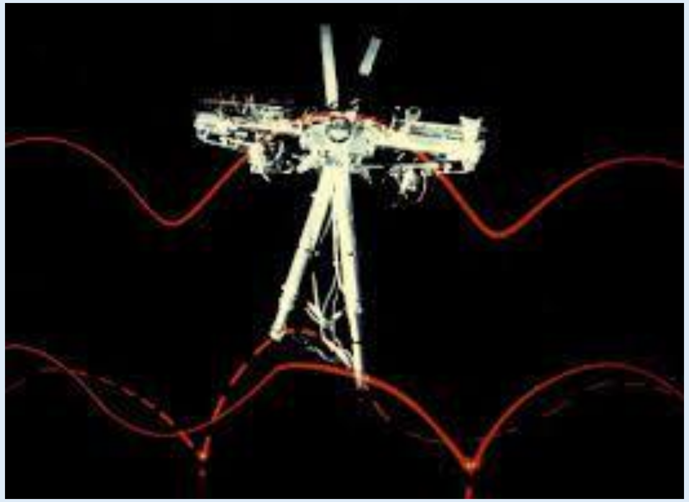


It's better to take a step!

落脚点控制

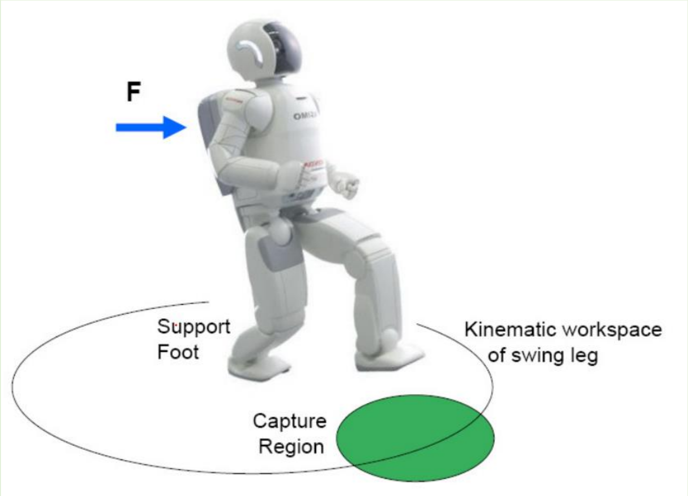
(1) Raibert heuristic (Marc Raibert)

$$x_f = \frac{\dot{x}T_s}{2} + k_{\dot{x}}\left(\dot{x} - \dot{x}_d\right)$$



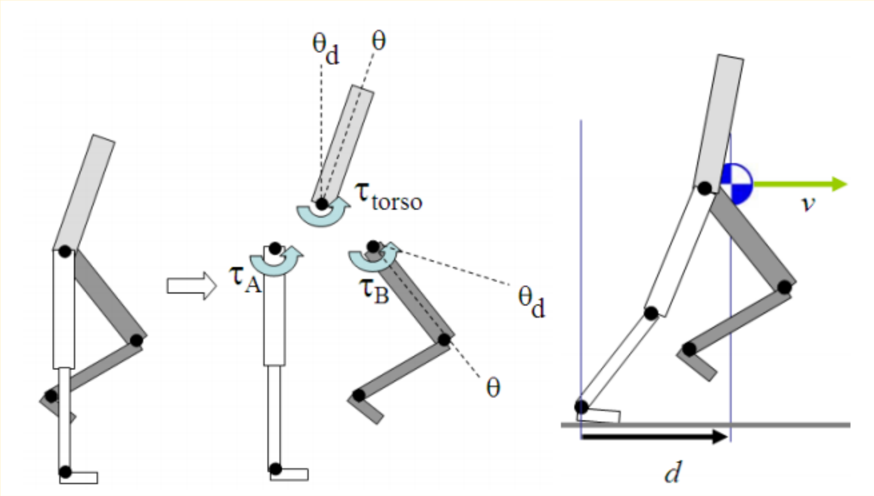
(2) Capture point (Jerry Pratt)

$$x_{capture} = \dot{x}\sqrt{\frac{z_0}{g}}$$



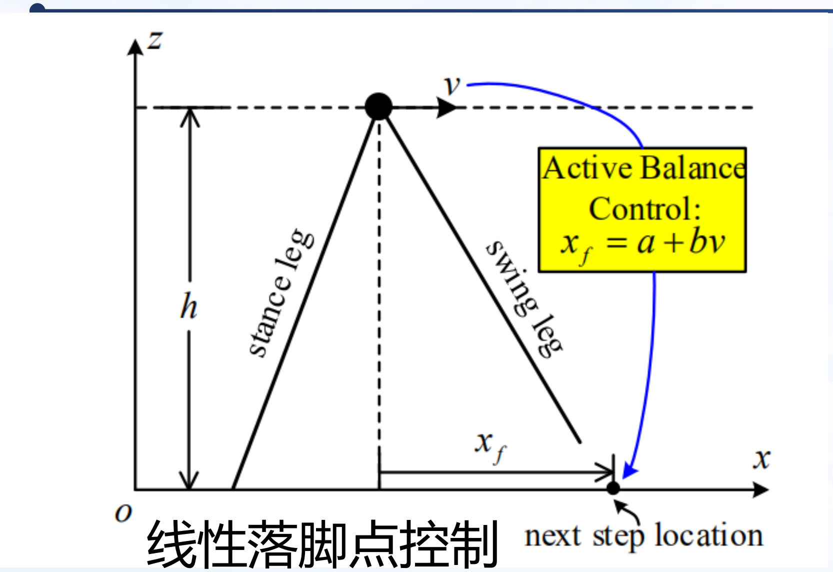
(3) SIMBICON (KangKang Yin)

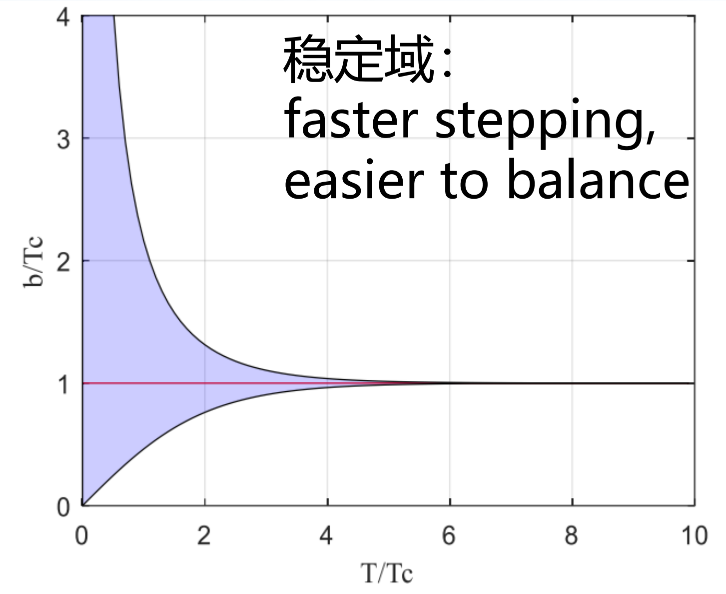
$$\theta_d = \theta_{d0} + c_d d + c_v v$$

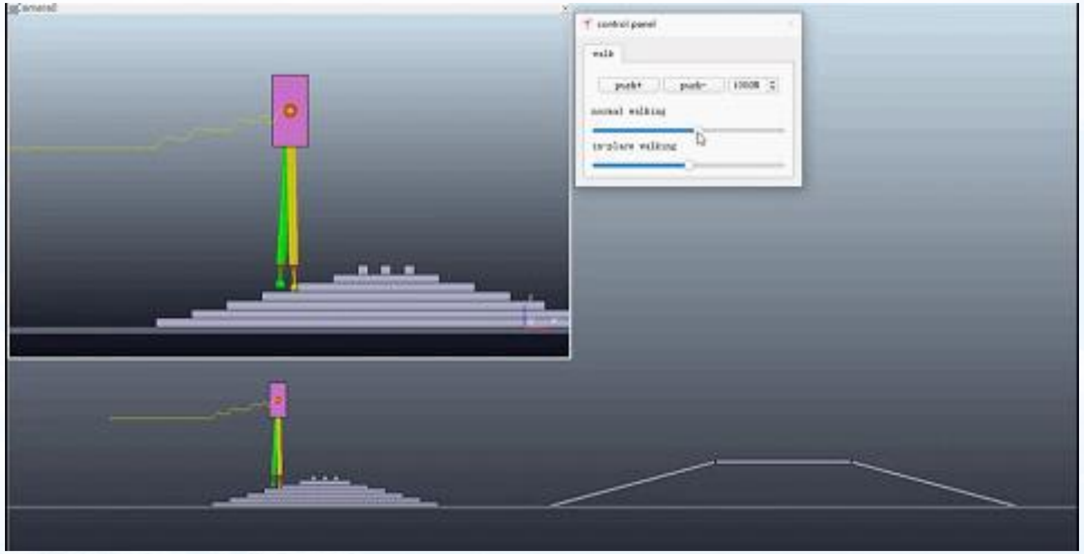


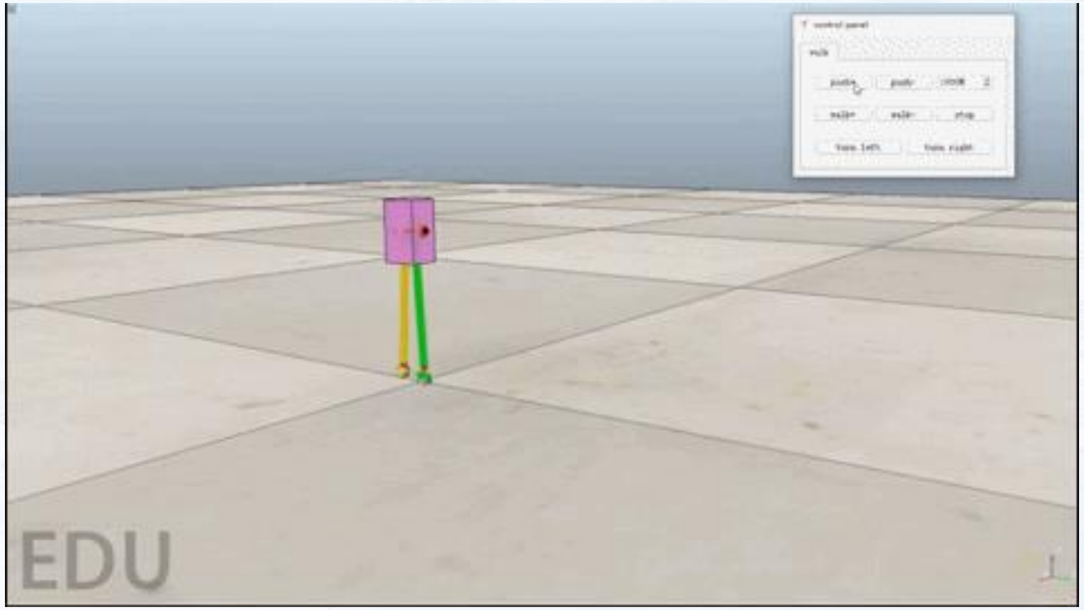
LFPC (linear foot placement control): use a linear function of the body velocity to determine the next foot placement.

落脚点控制


z
v
Active Balance Control:
$x_f = a + bv$
h
stance leg
swing leg
x_f
x
o
next step location
线性落脚点控制


稳定域：
faster stepping,
easier to balance
b/Tc
T/Tc
0
1
2
3
4
0
2
4
6
8
10




EDU

人形机器人控制
从过去到未来

目录

预测控制

2021

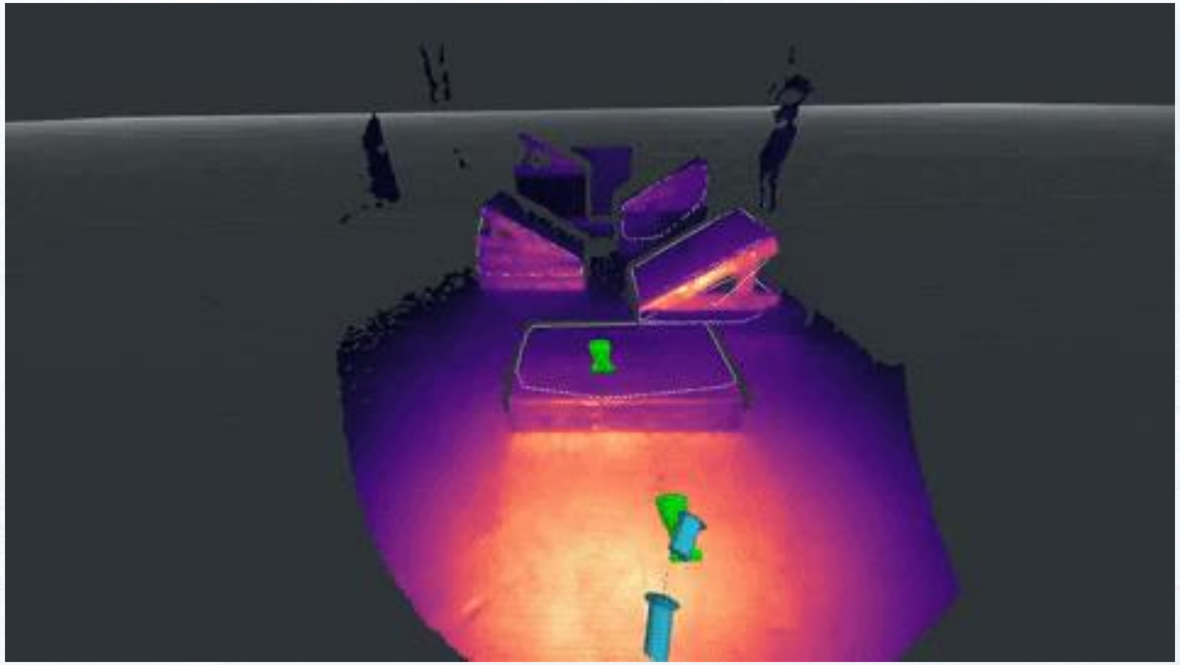
预测控制

行为库

Atlas 在跑酷中执行的每个动作都源自使用轨迹优化提前创建的模板。给定感知的计划目标，机器人从库中选择与给定目标尽可能匹配的行为。

模型预测控制

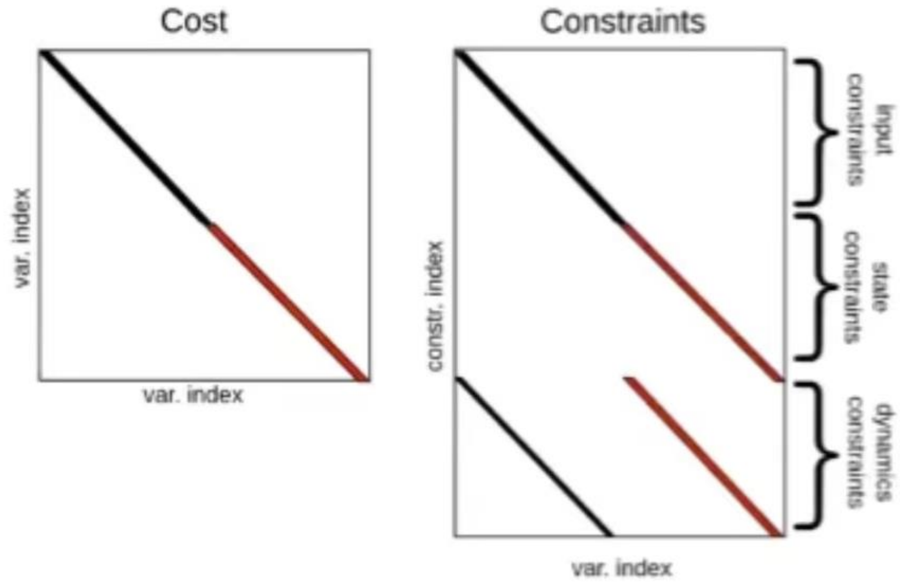
使用机器人动力学模型来预测其运动，控制器通过优化来计算机器人现在要做的最佳事情，以应对环境几何、脚滑或其他实时因素的差异。

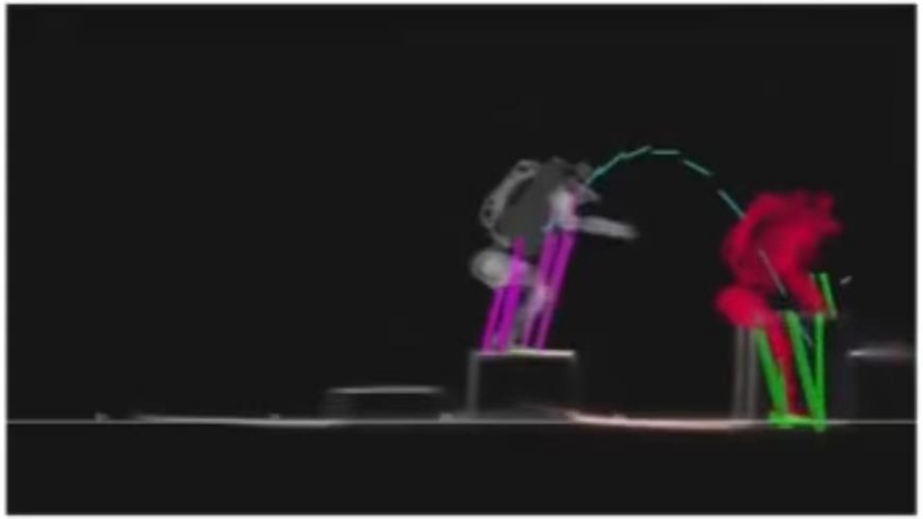


预测控制

MPC: Common Features

- Nonlinear dynamics, costs, constraints
- Iteratively linearize and solve a QP
- Never run to convergence
- Exploit problem structure for speed
- Don't treat solvers as a black box

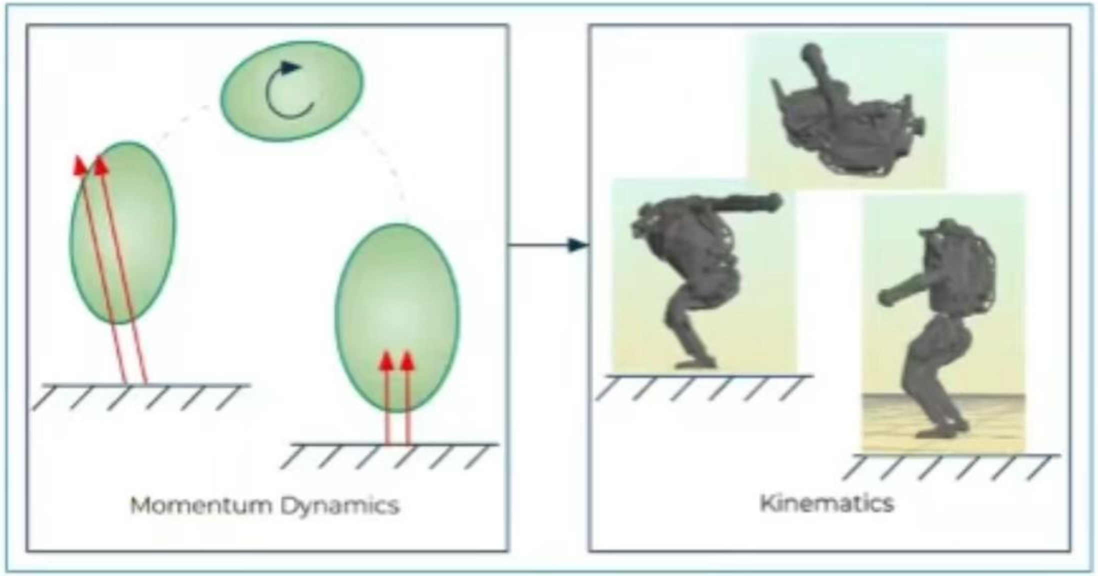




~10,000,000 QPs solved (hardware)
~10,000,000,000 QPs solved (sim)

预测控制

The Robot is a Potato with Limbs



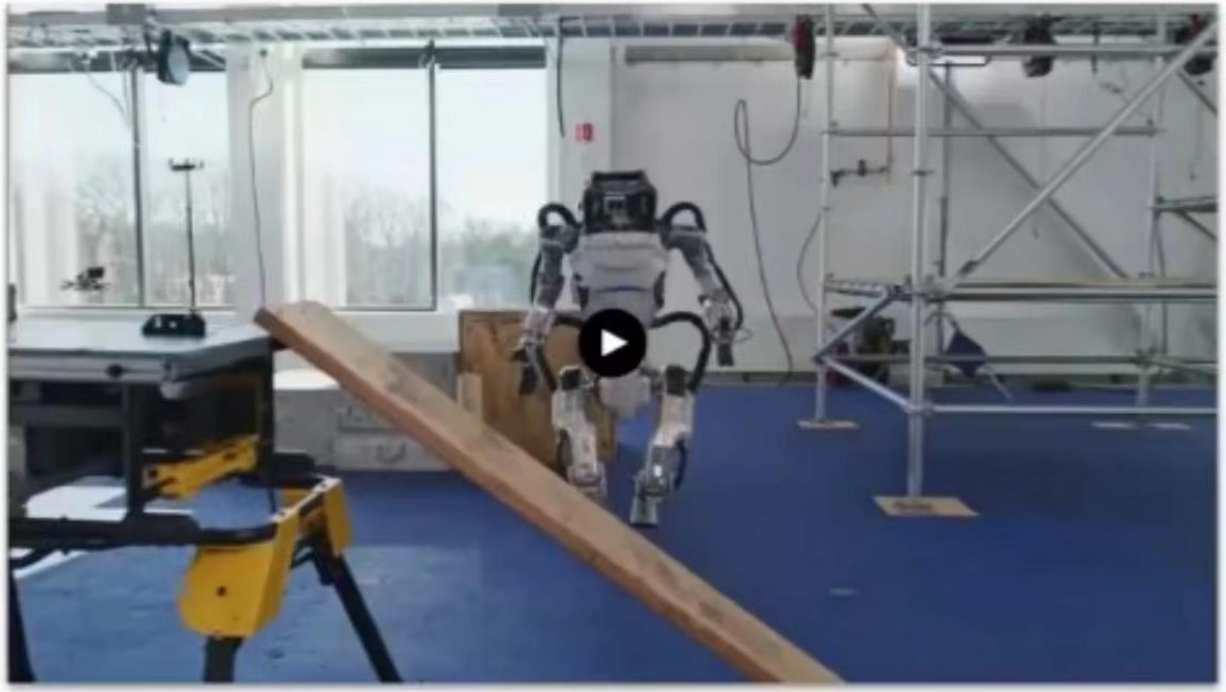
- Centroidal dynamics
- Independent kinematics

BostonDynamics

预测控制

THU-SHU ROBOART LAB

The Robot is a Kinodynamic System



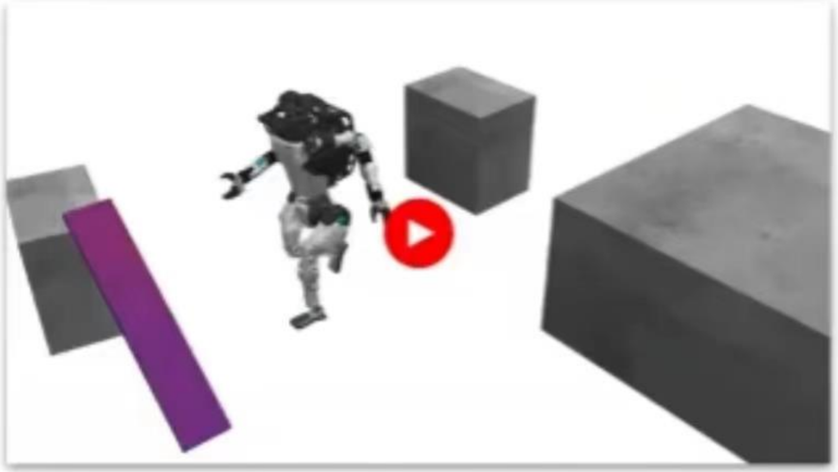
- Coupled kinematics and centroidal dynamics in one big optimization

10

BostonDynamics

预测控制

The Robot is Several Kinodynamic Systems

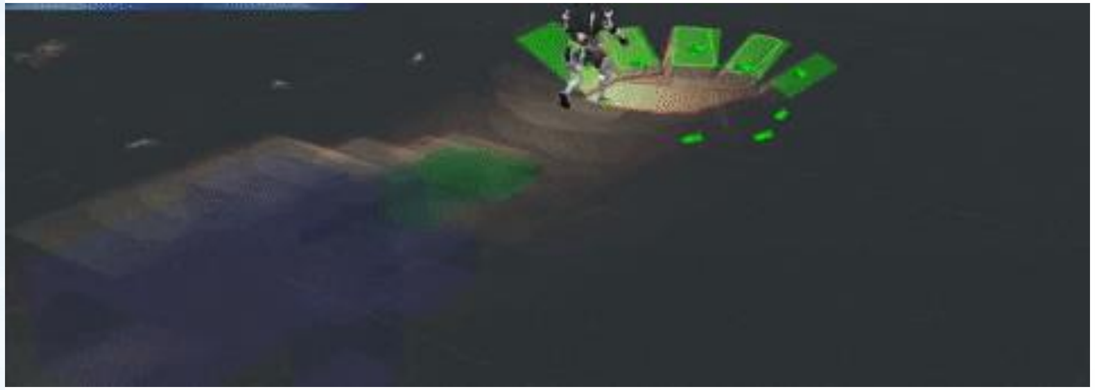


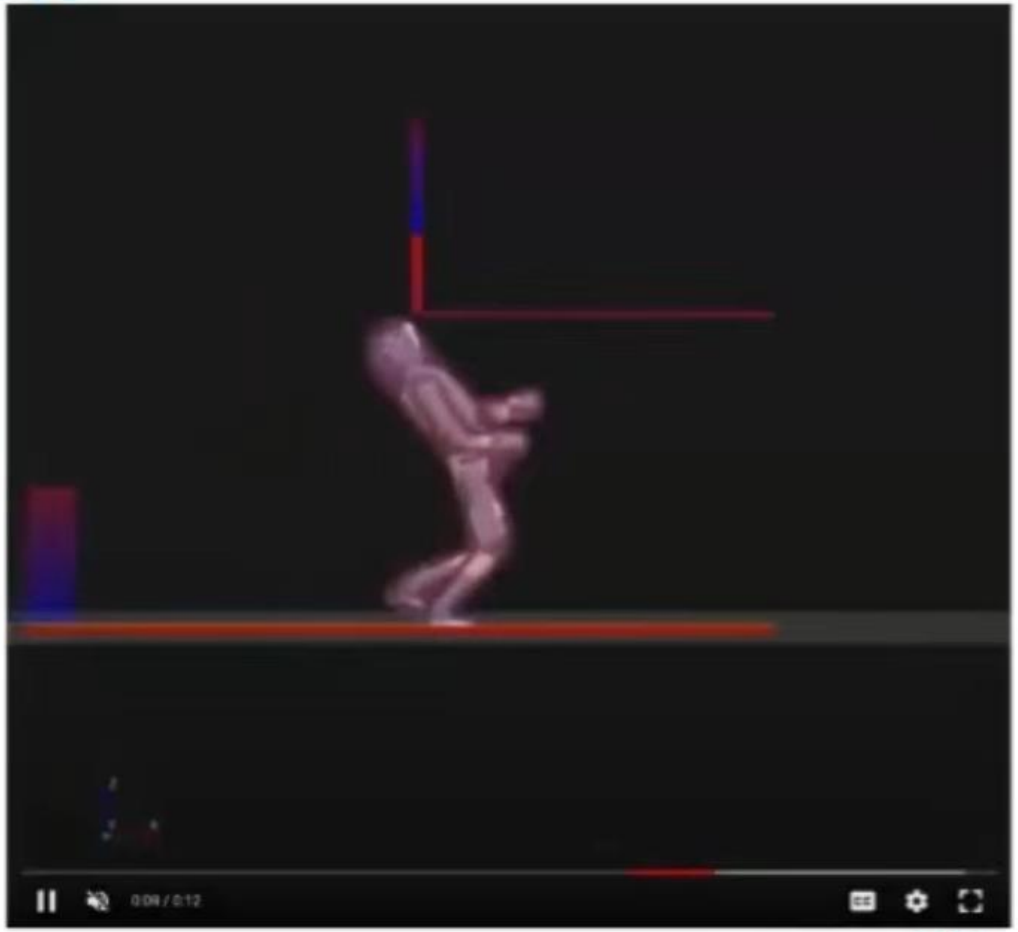
- Robot kinematics and centroidal dynamics
- Object kinematics and centroidal dynamics
- Robot-object interactions

BostonDynamics

预测控制

The Robot is Several Kinodynamic Systems in a Perceived Environment





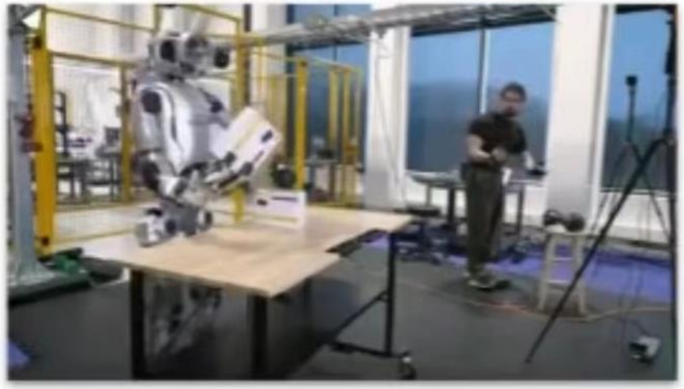
- Robot kinematics and centroidal dynamics
- Object kinematics and centroidal dynamics
- Robot-object interactions
- Perception-driven constraints

12

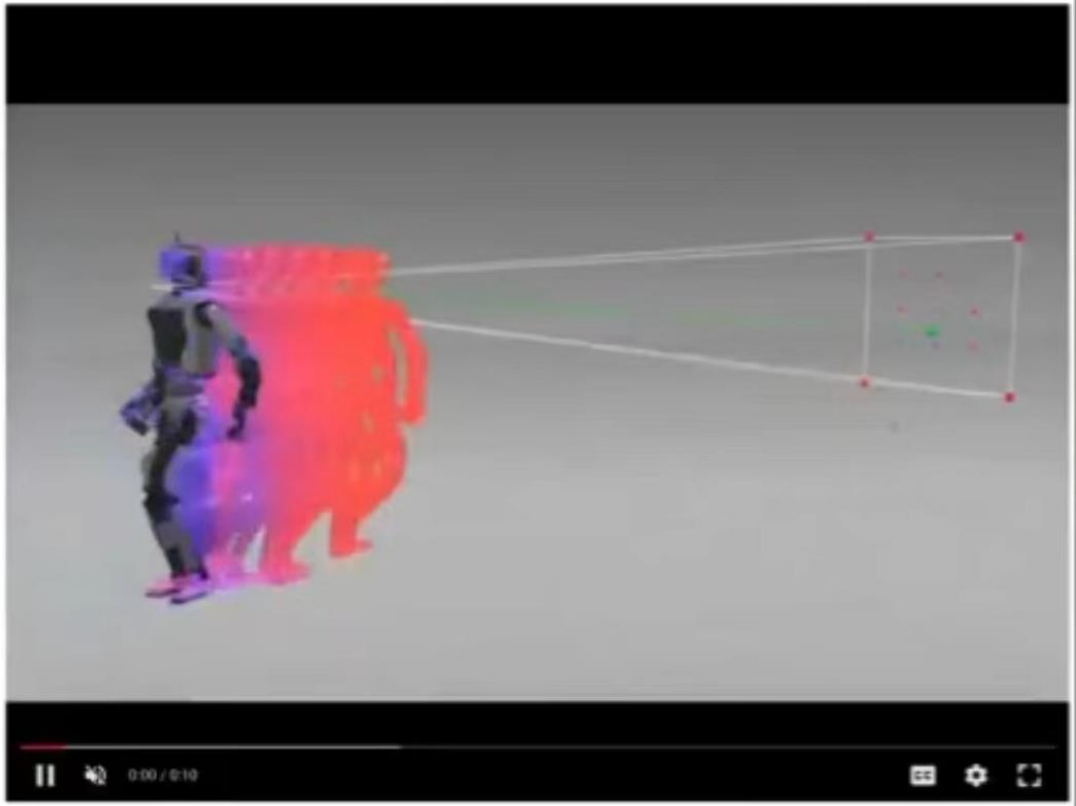
BostonDynamics

预测控制

The Robot is Several Kinodynamic Systems in a Perceived Environment Doing Useful Tasks



- Robot kinematics and centroidal dynamics
- Object kinematics and centroidal dynamics
- Robot-object interactions
- Perception-driven constraints
- Tasks updated on the fly



BostonDynamics

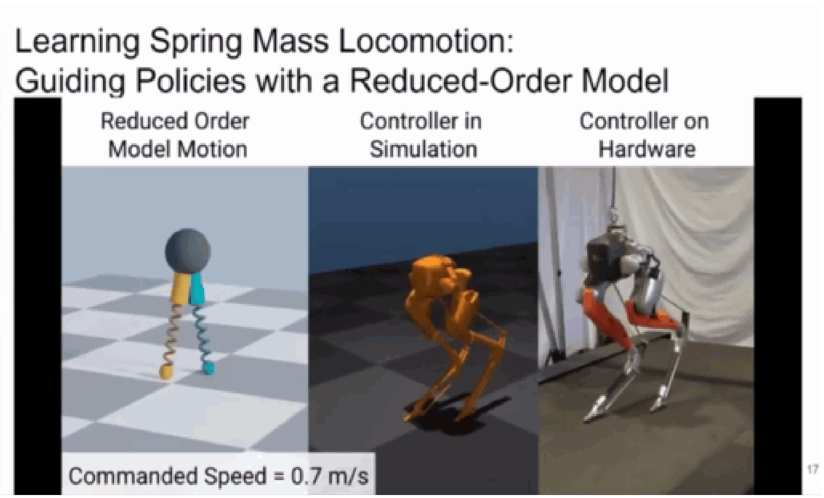
预测控制

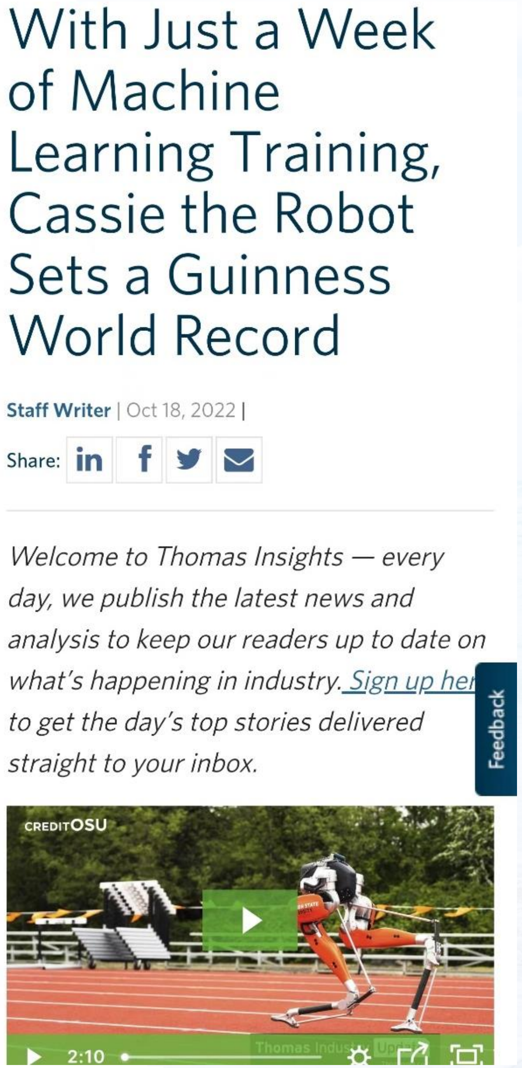
2023

人形机器人控制从过去到未来

目录

强化学习





With Just a Week of Machine Learning Training, Cassie the Robot Sets a Guinness World Record

Staff Writer | Oct 18, 2022 |

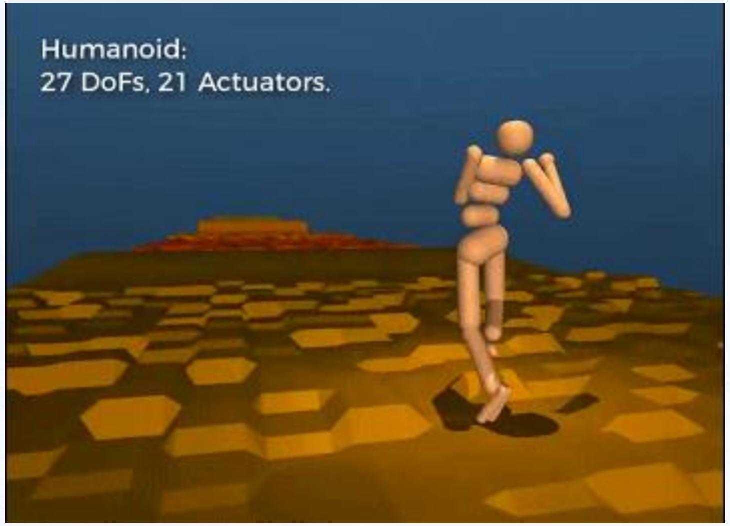
Share:

Welcome to Thomas Insights — every day, we publish the latest news and analysis to keep our readers up to date on what's happening in industry. Sign up her to get the day's top stories delivered straight to your inbox.

Feedback

强化学习

Simple reward may lead to unnatural behavior!



Deepmind 2017, Emergence of Locomotion Behaviours in Rich Environments

Rewards

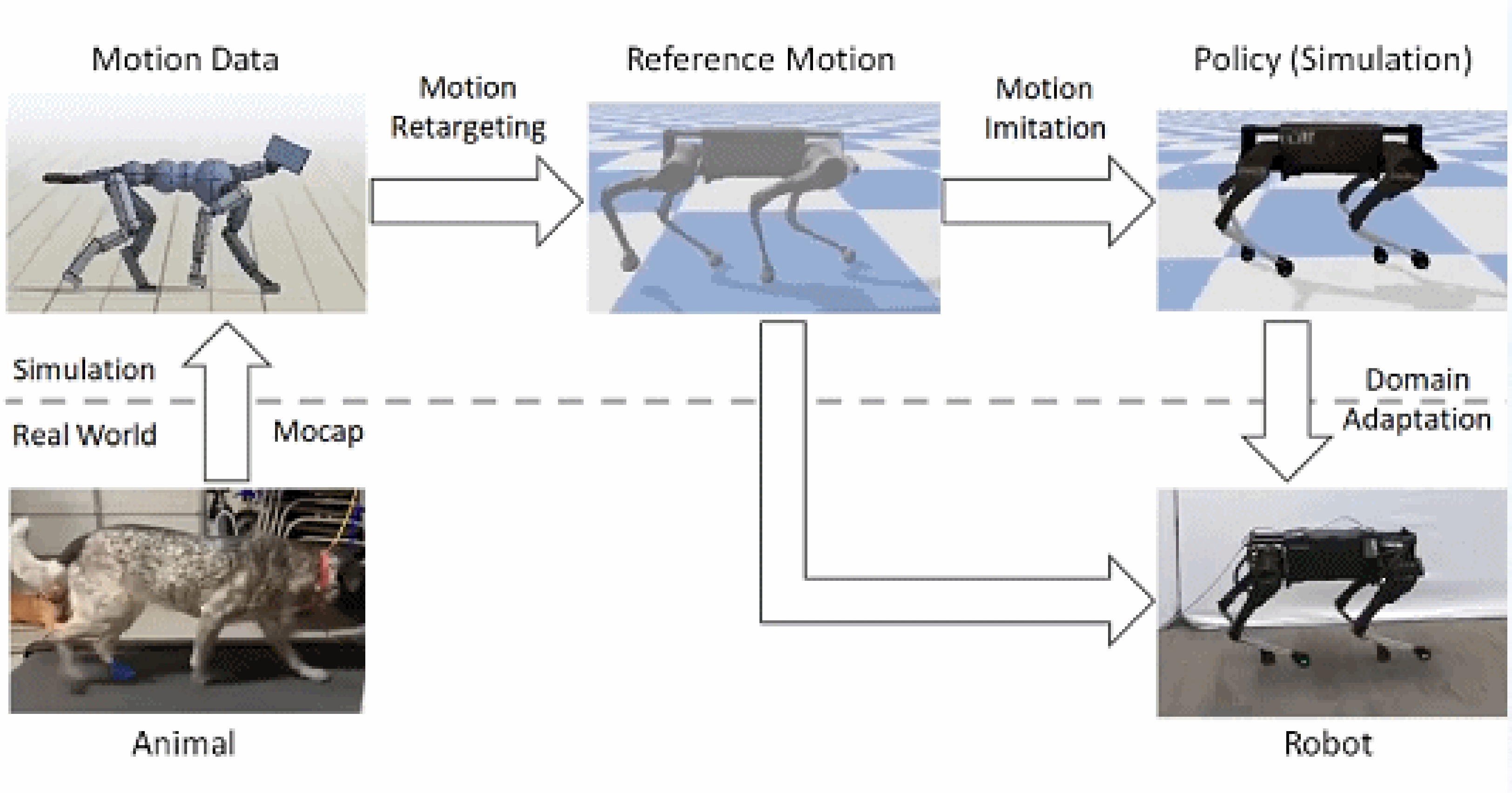
Reward for forward velocity

Energy punishment

Humanoid $r = \boxed{\min(v_x, v_{\max}) - 0.005(v_x^2 + v_y^2) - 0.05y^2} - 0.02\|u\|^2 + 0.02$ where $v_{\max}$ is a cutoff for the velocity reward which we usually set to $4m/s$.

强化学习

Imitation learning → Follow a given reference motion



Reward Function.

$$r_t = w^{\mathrm{p}} r_t^{\mathrm{p}} + w^{\mathrm{v}} r_t^{\mathrm{v}} + w^{\mathrm{e}} r_t^{\mathrm{e}} + w^{\mathrm{rp}} r_t^{\mathrm{rp}} + w^{\mathrm{rv}} r_t^{\mathrm{rv}}$$

$$r_t^{\mathrm{p}} = \exp\left[-5 \sum_j ||\hat{\mathbf{q}}_t^j - \mathbf{q}_t^j||^2\right]$$

$$r_t^{\mathrm{v}} = \exp\left[-0.1 \sum_j ||\hat{\dot{\mathbf{q}}}_t^j - \dot{\mathbf{q}}_t^j||^2\right]$$

$$r_t^{\mathrm{e}} = \exp\left[-40 \sum_e ||\hat{\mathbf{x}}_t^e - \mathbf{x}_t^e||^2\right]$$

$$r_t^{\mathrm{rp}} = \exp\left[-20||\hat{\mathbf{x}}_t^{\mathrm{root}} - \mathbf{x}_t^{\mathrm{root}}||^2 - 10||\hat{\mathbf{q}}_t^{\mathrm{root}} - \mathbf{q}_t^{\mathrm{root}}||^2\right]$$

$$r_t^{\mathrm{rv}} = \exp\left[-2||\hat{\dot{\mathbf{x}}}_t^{\mathrm{root}} - \dot{\mathbf{x}}_t^{\mathrm{root}}||^2 - 0.2||\hat{\dot{\mathbf{q}}}_t^{\mathrm{root}} - \dot{\mathbf{q}}_t^{\mathrm{root}}||^2\right]$$

Mimic reward for trajectory tracking

Xue Bin Peng 2020, Learning Agile Robotic Locomotion Skills by Imitating Animals

强化学习

THU-SHU ROBOART LAB

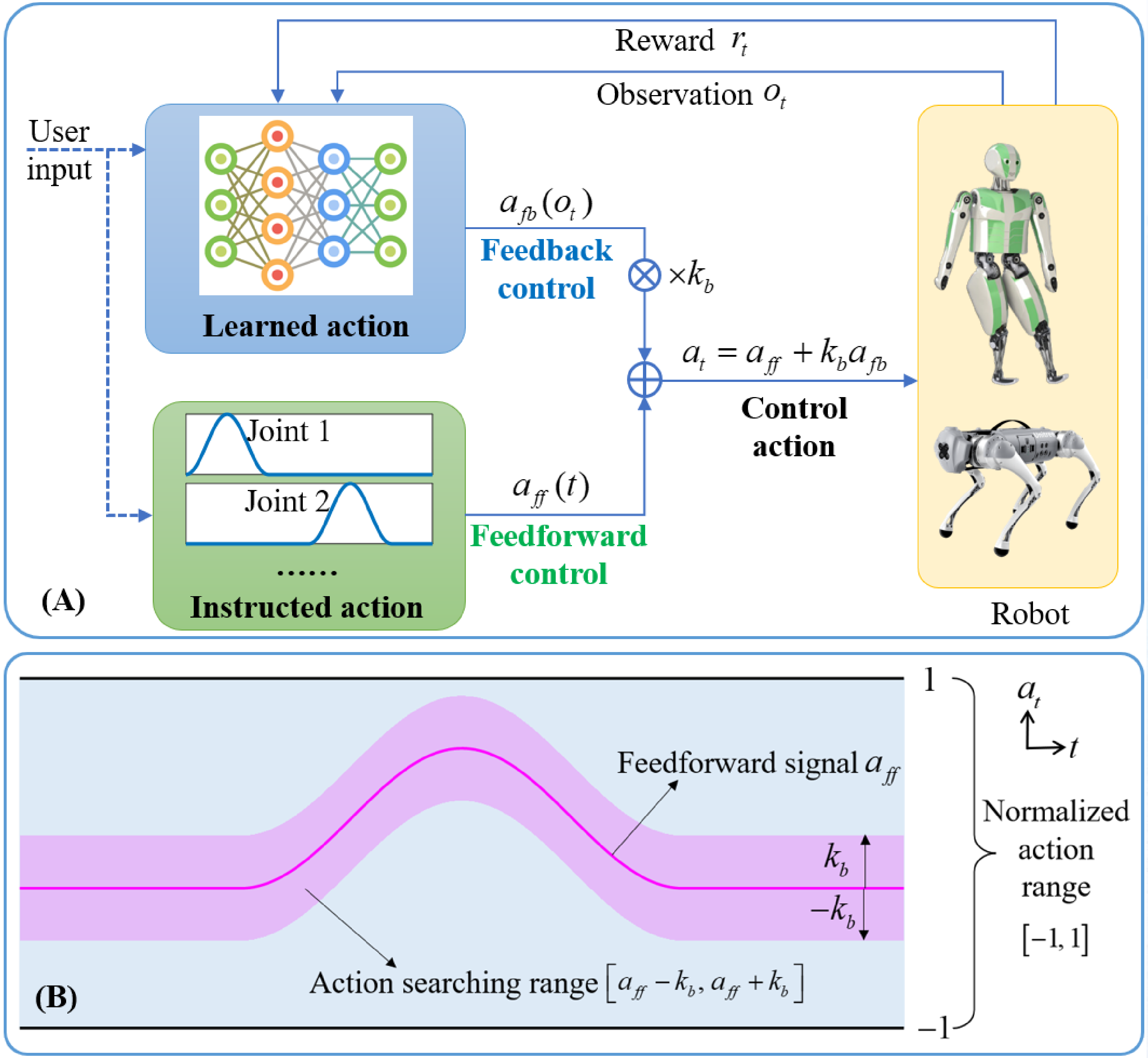
Yann LeCun
@ylecun

Indeed, I do favor MPC over RL.
I've been making that point since at least 2016.
RL requires ridiculously large numbers of trials to learn any new task.
In contrast MPC is zero shot: If you have a good world model and a good task objective, MPC can solve new tasks without any task-specific learning.
That's the magic of planning.
It doesn't mean that RL is useless, but its use should be a last resort.

强化学习

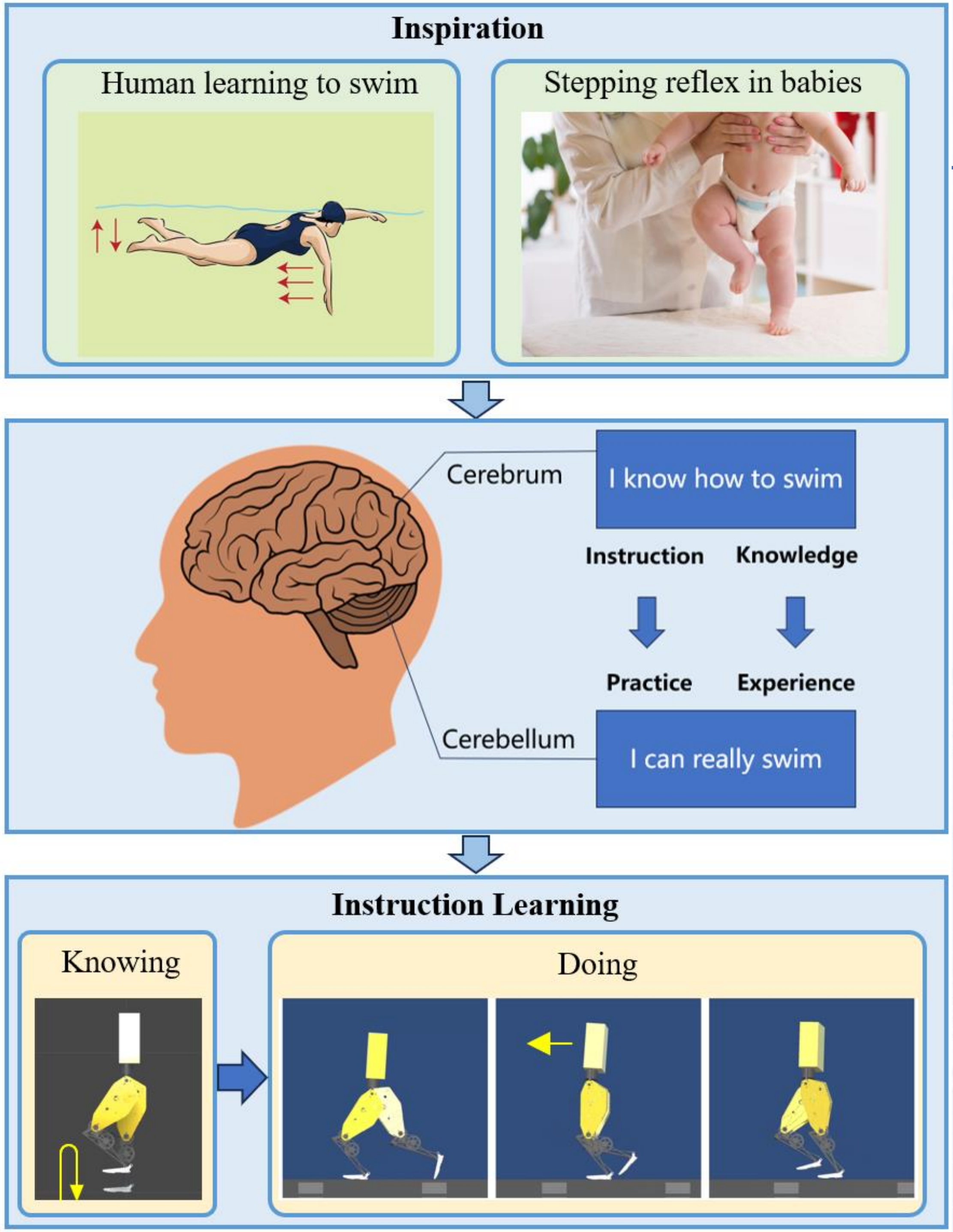
Instruction Learning Framework

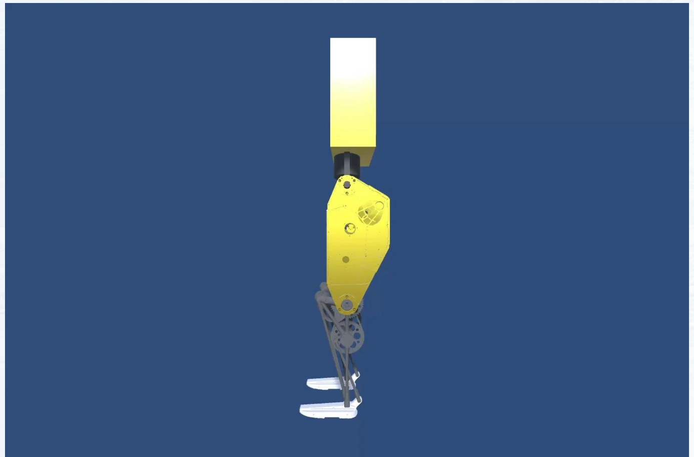
Action Bounding Technique

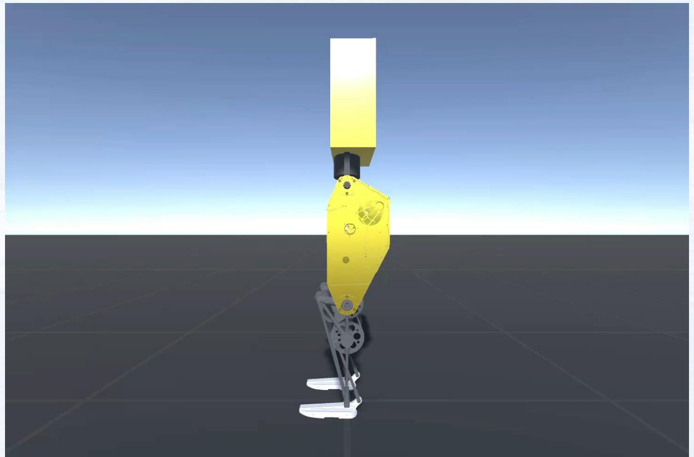


强化学习

Inspiration from human learning







强化学习

Imitation Learning

vs.

Instruction Learning

Learn from scratch

Full action space

Purely reward-driven

Require mimic reward

Learn by following instruction

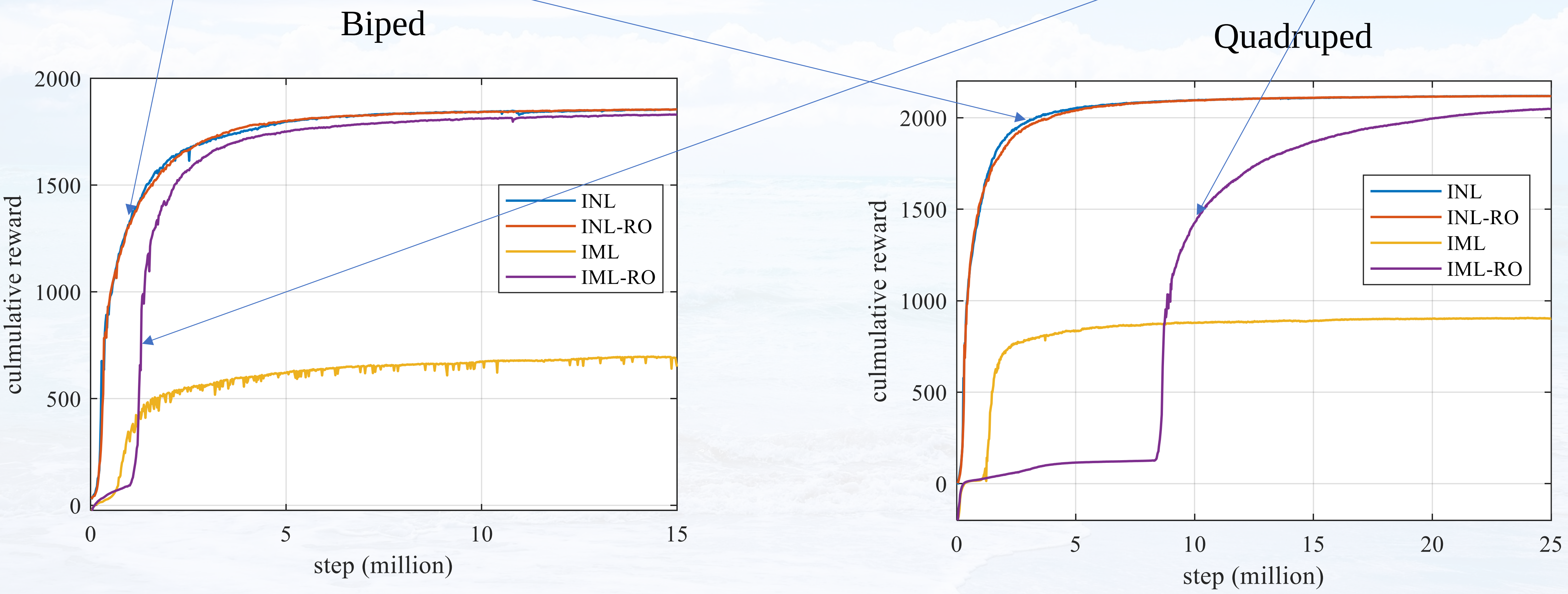
Reduced action space

Reward and feedforward driven

No need for mimic reward

强化学习

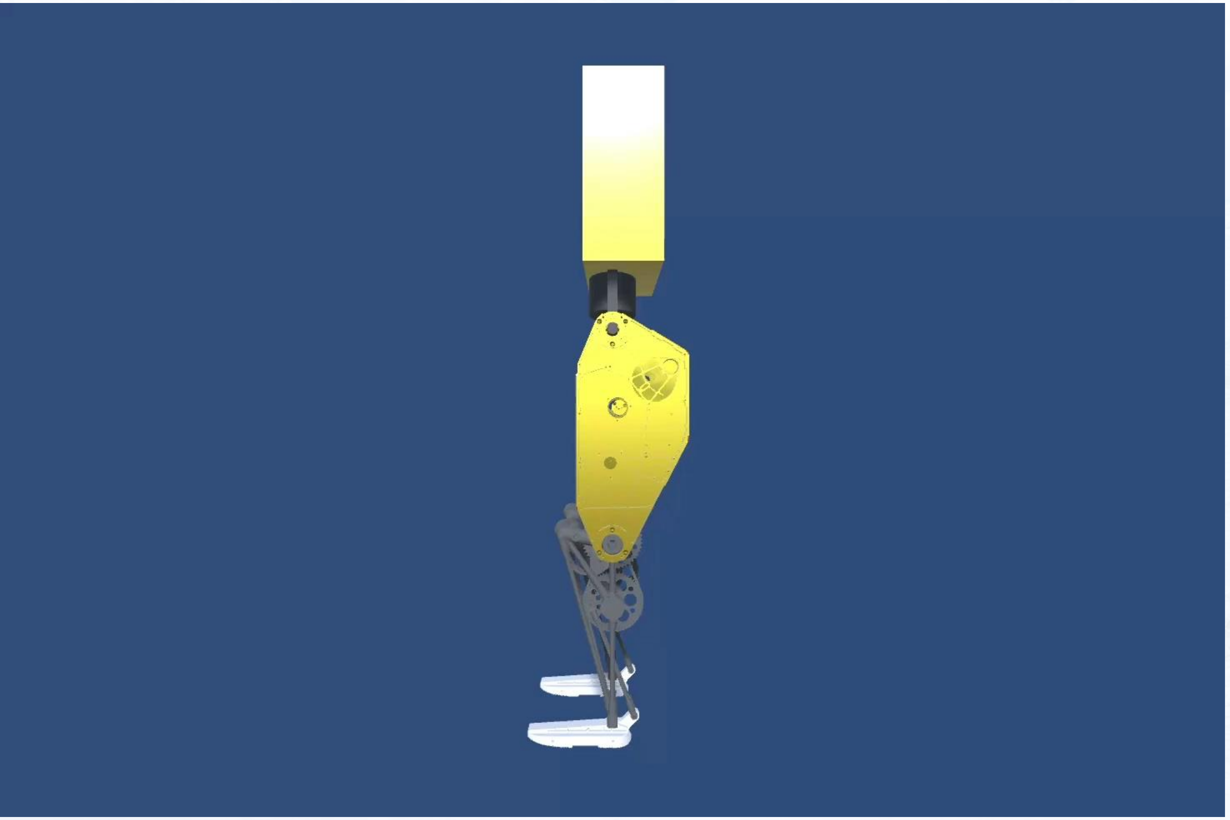
Instruction learning (INL) learns much faster than imitation learning (IML)



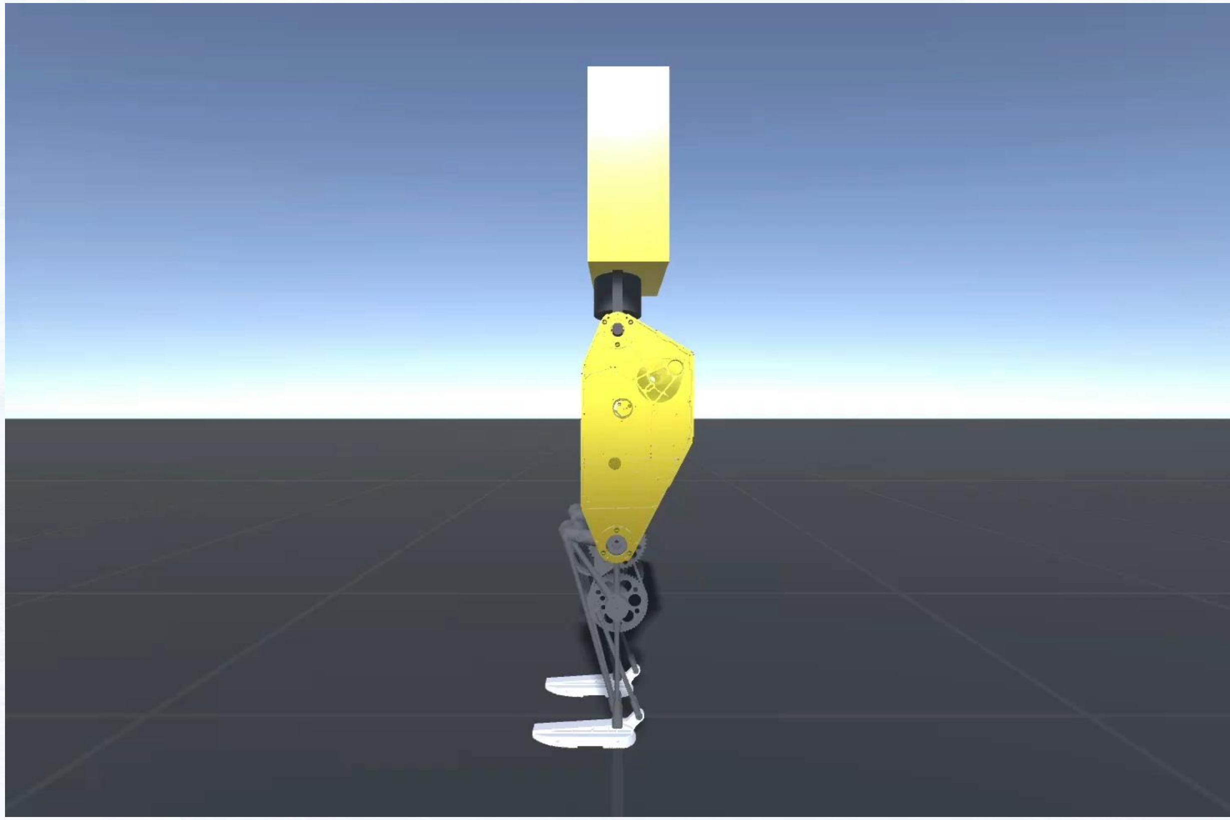
强化学习

原地踏步

Feedforward action



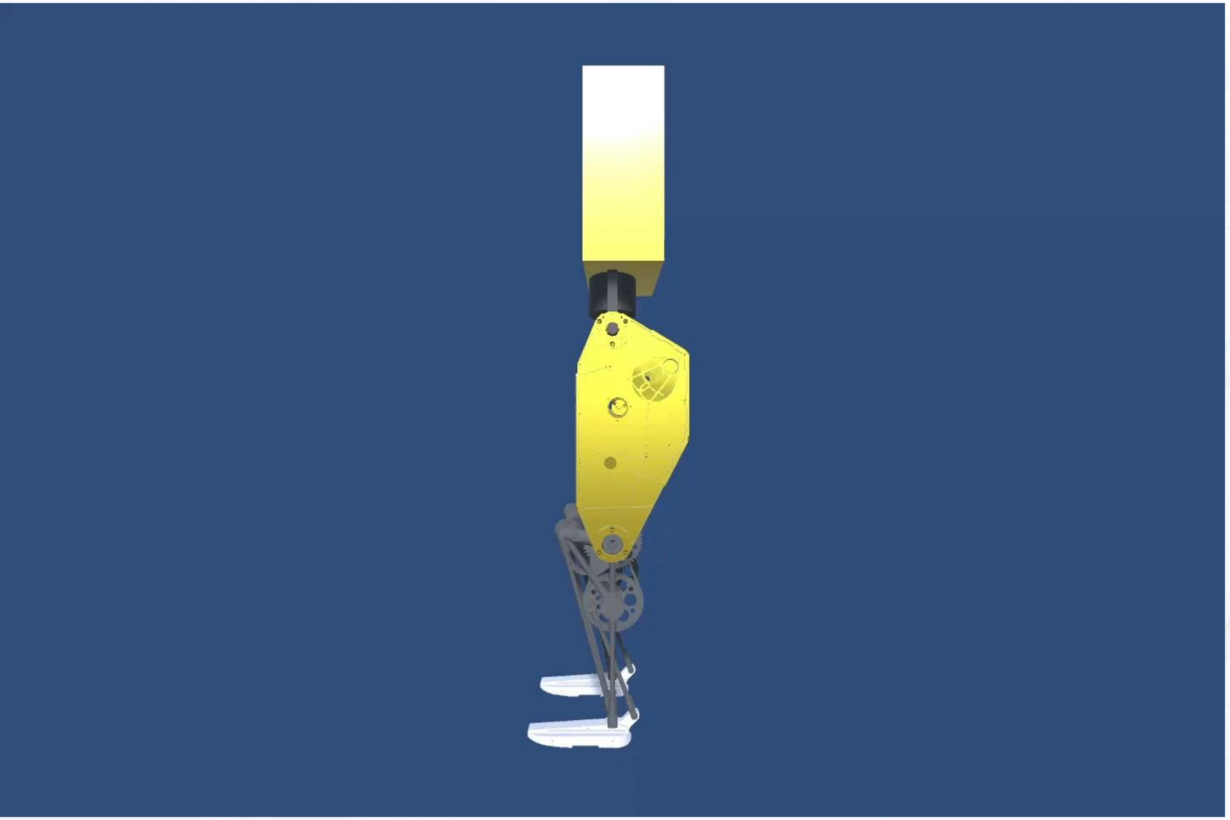
Learned motion



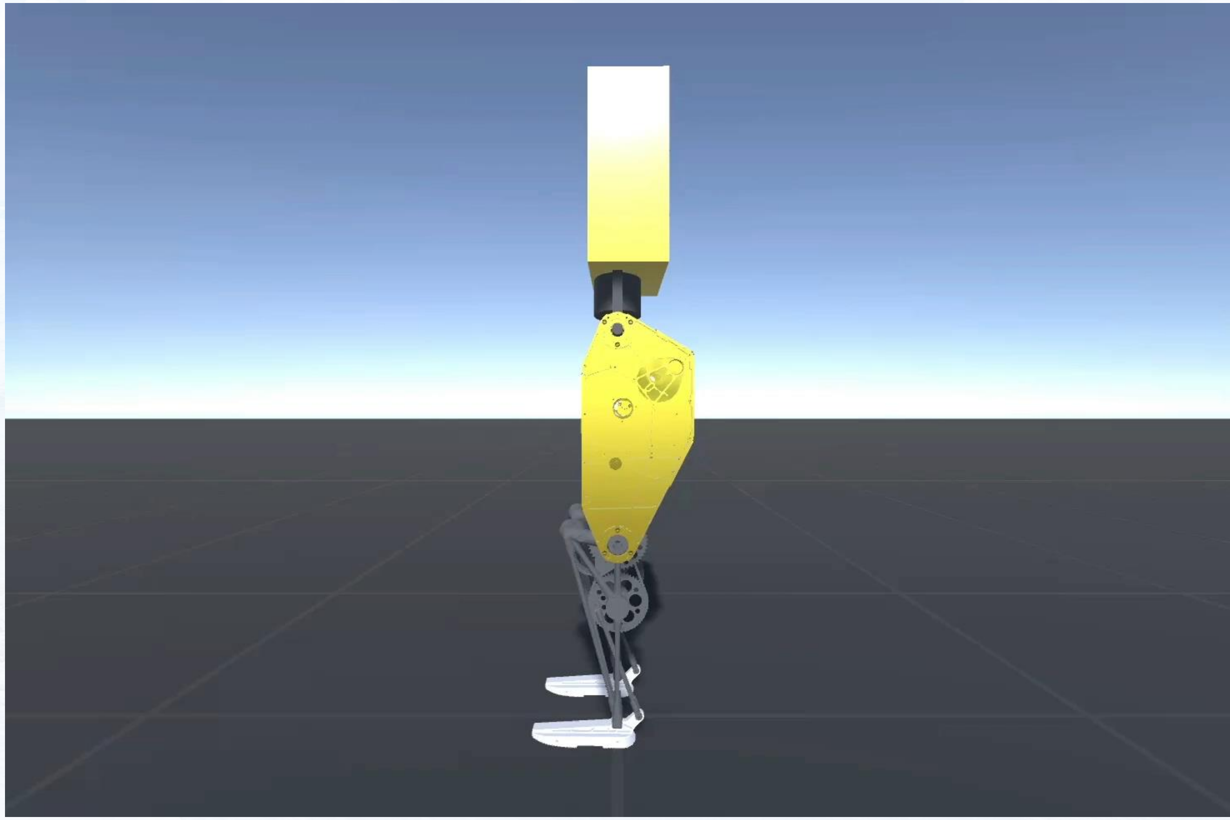
强化学习

正常行走

Feedforward action



Learned motion

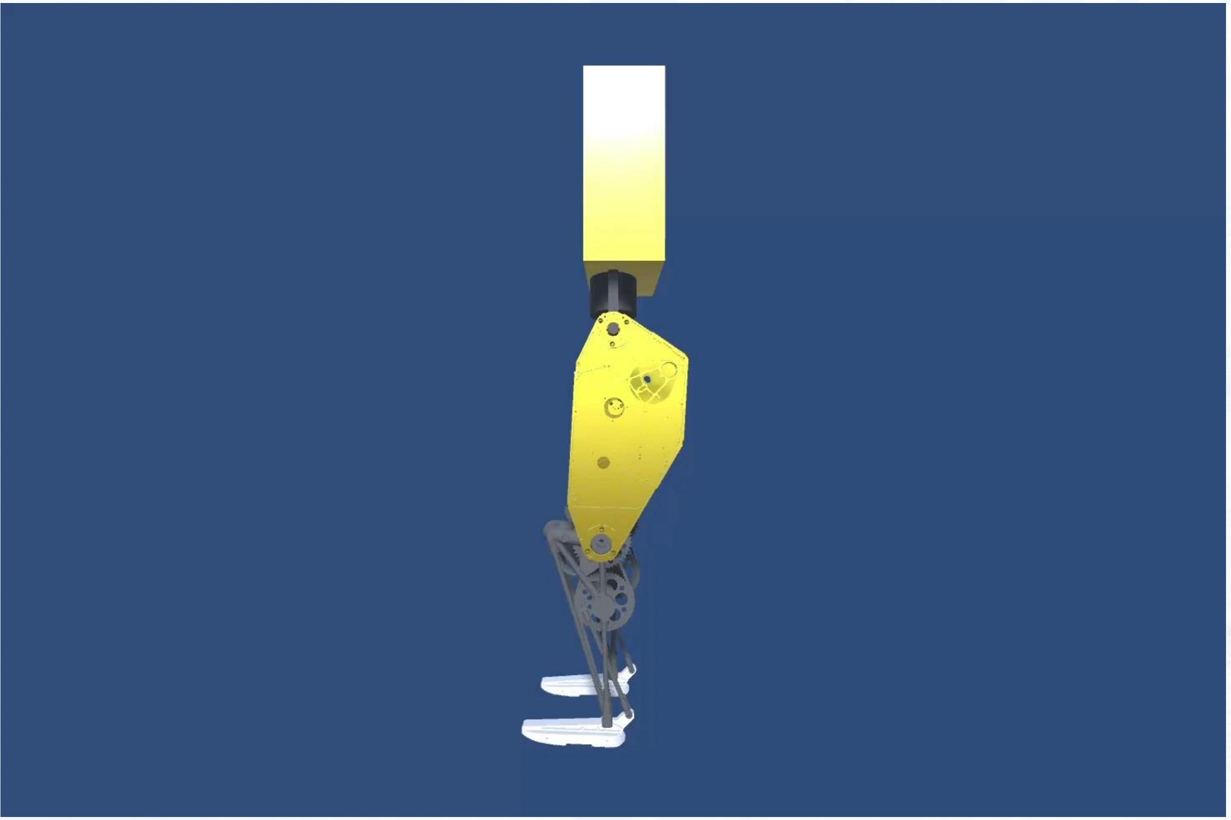


强化学习

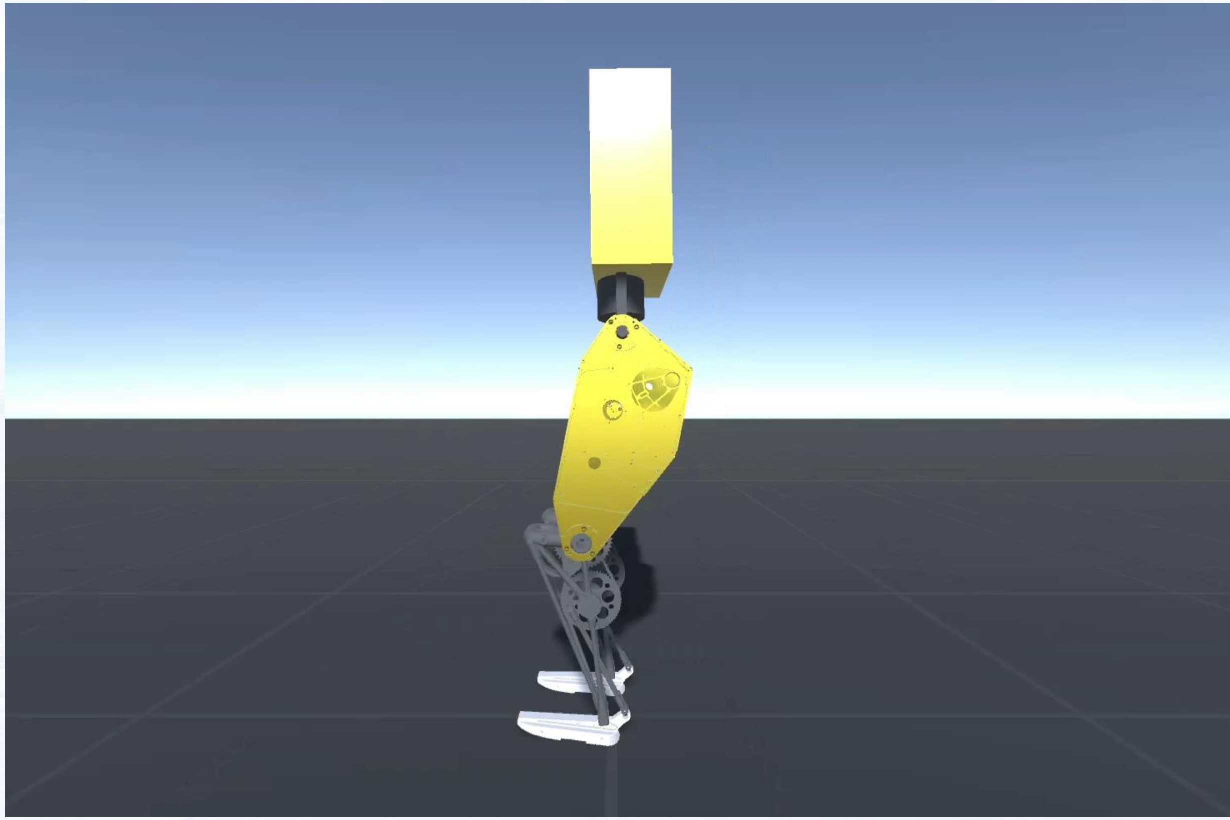
THU-SHU ROBOART LAB

水平行走

Feedforward action

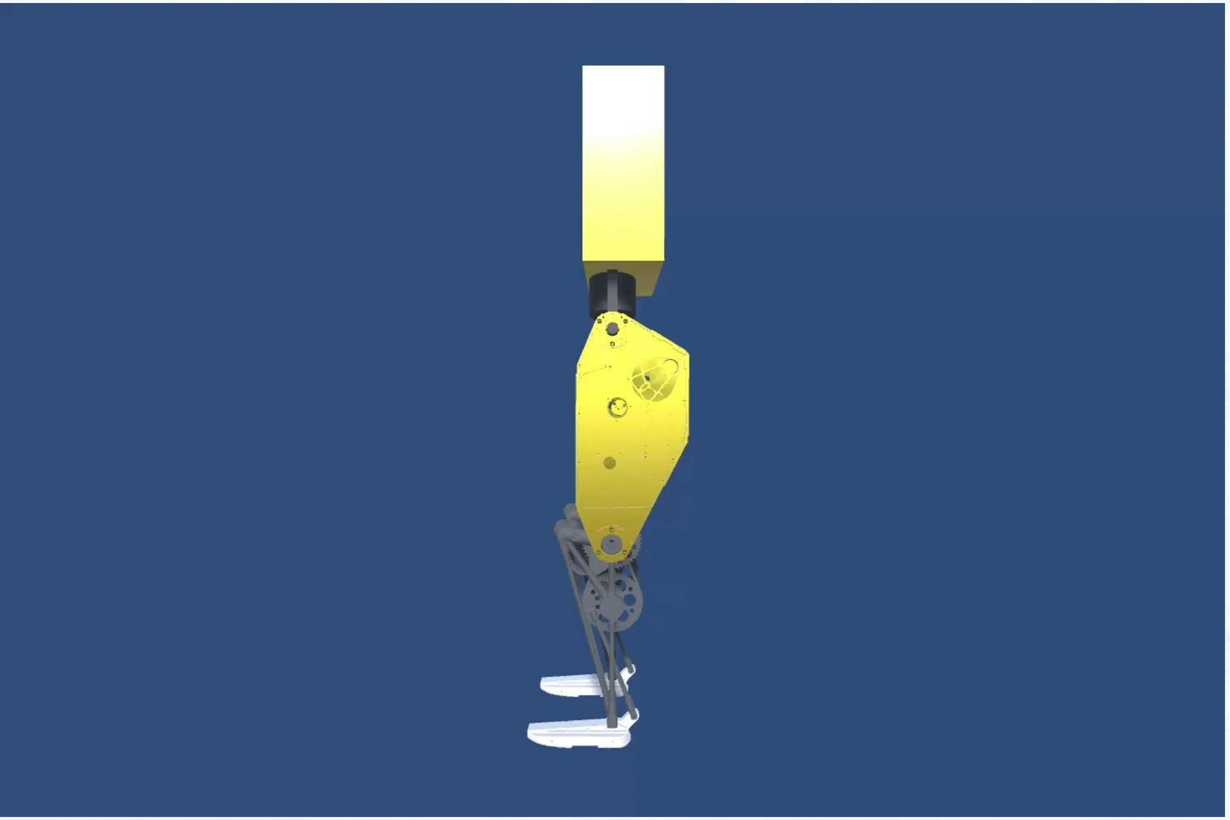


Learned motion

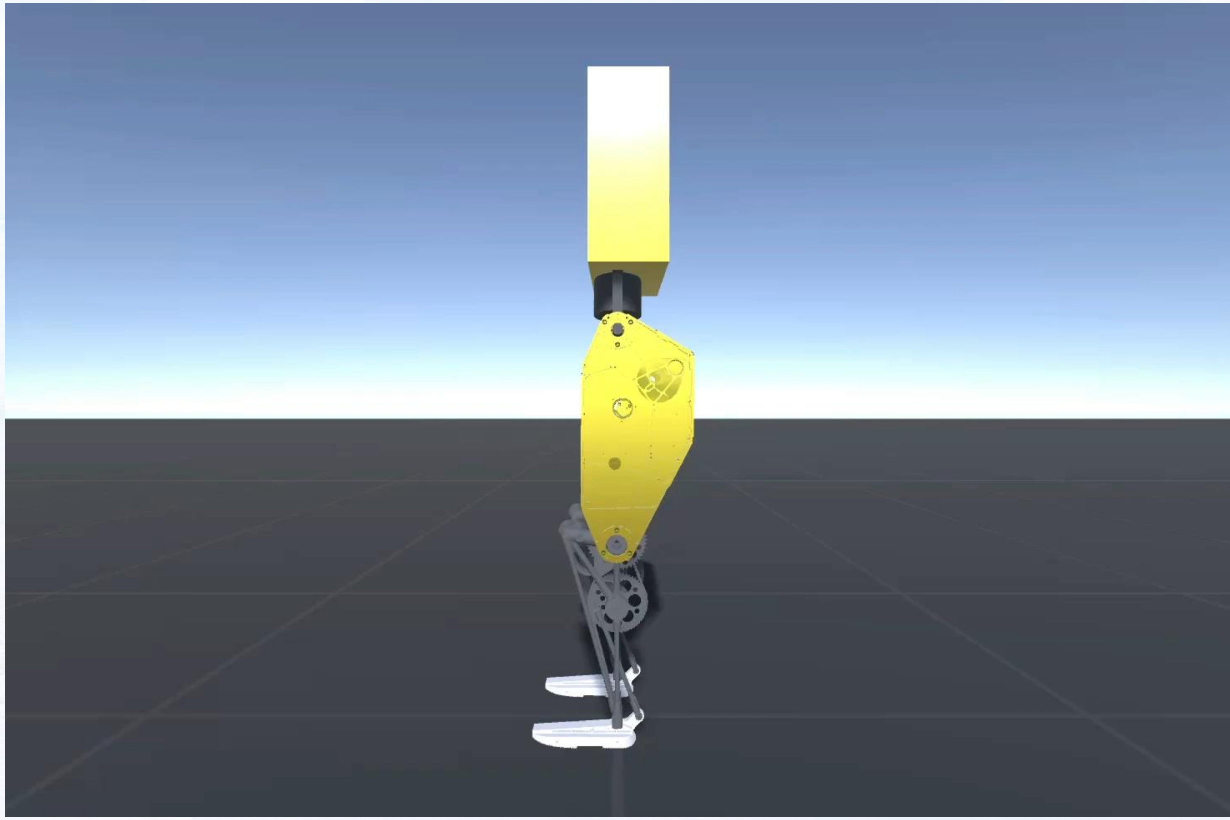


踢正步行走

Feedforward action



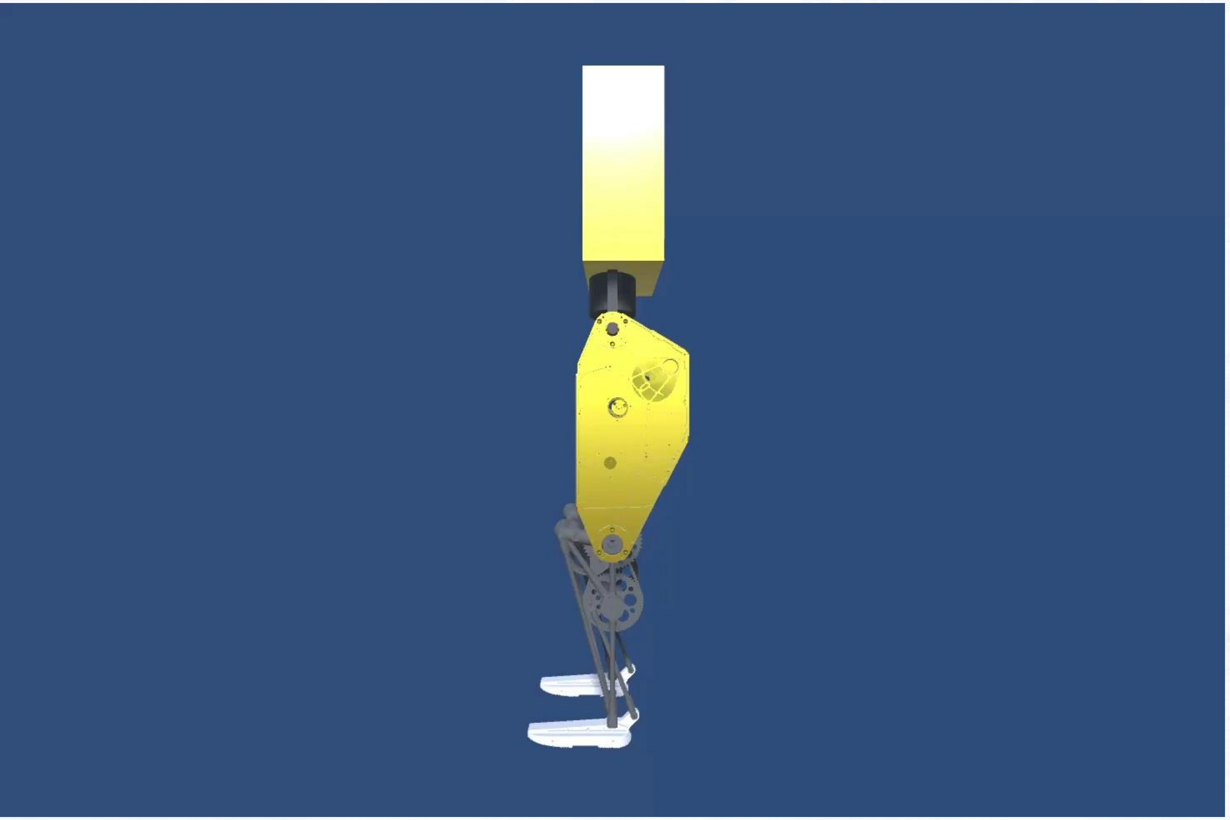
Learned motion



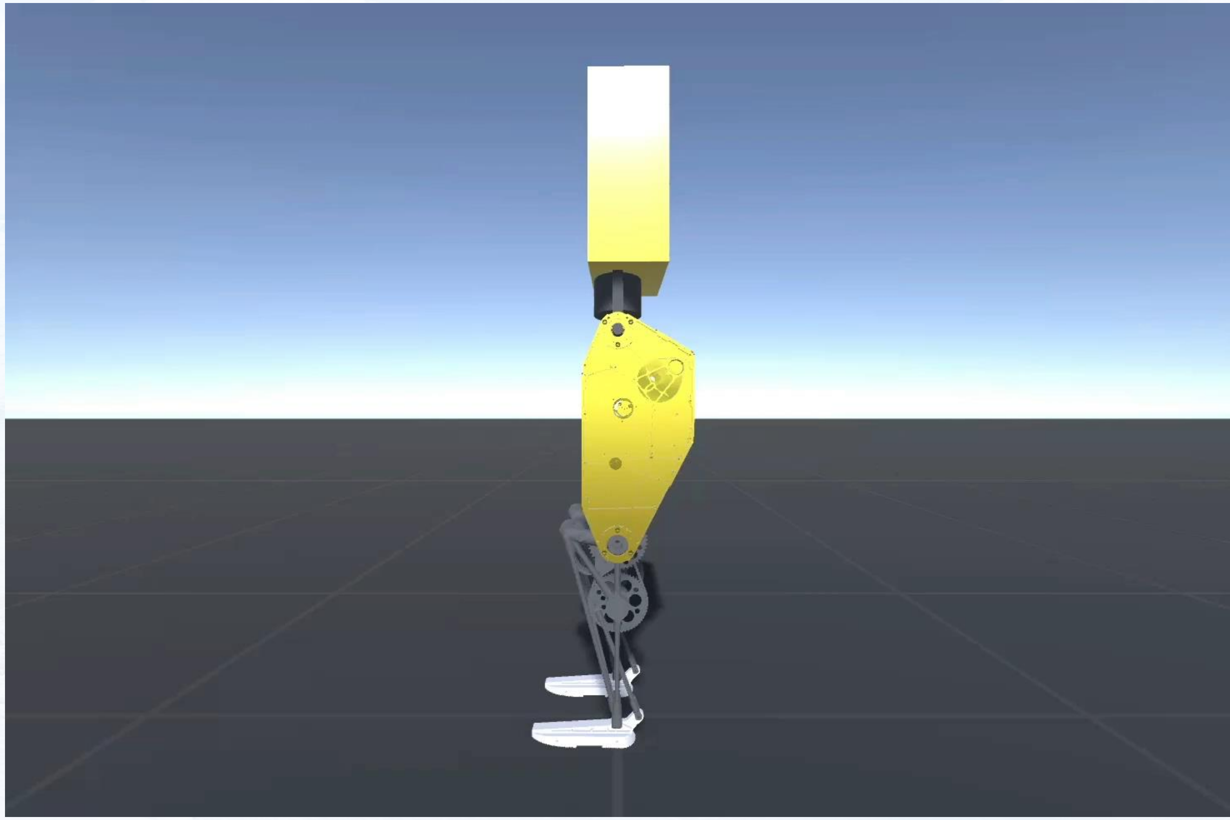
强化学习

蛙跳

Feedforward action



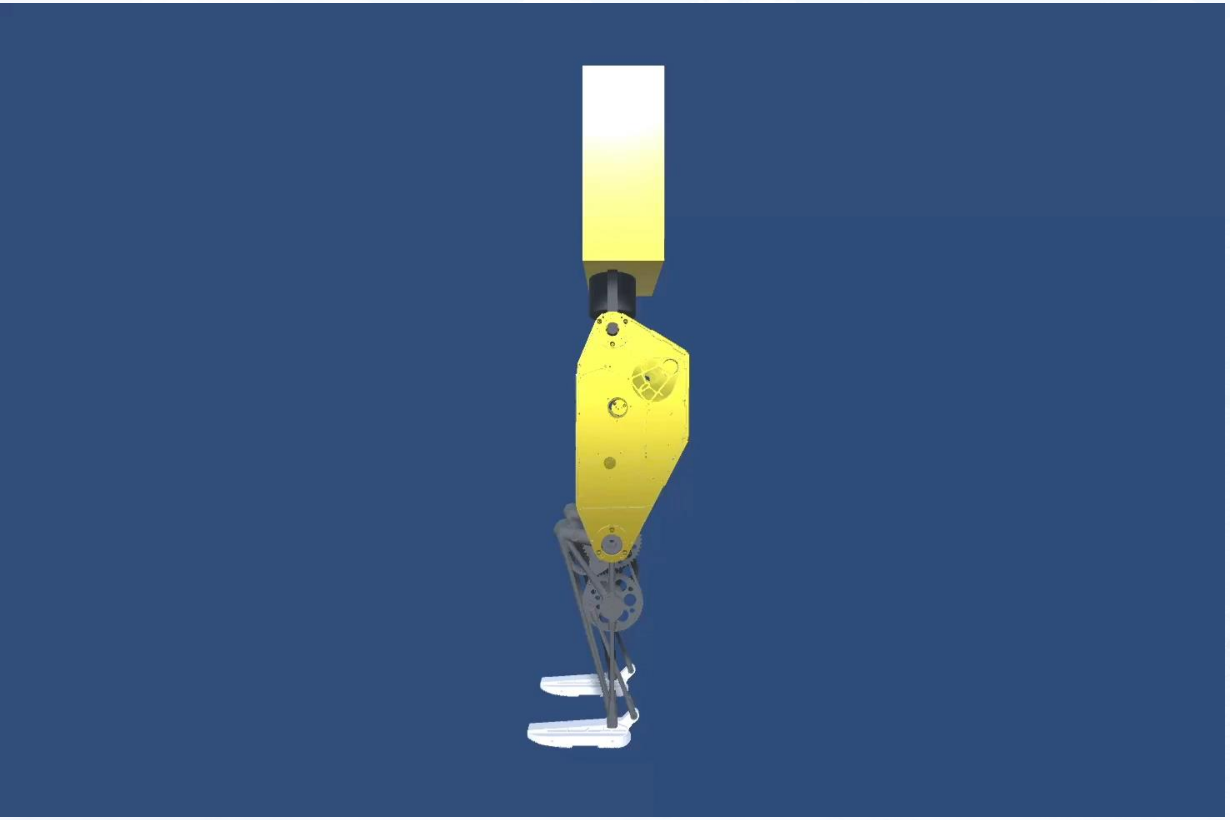
Learned motion



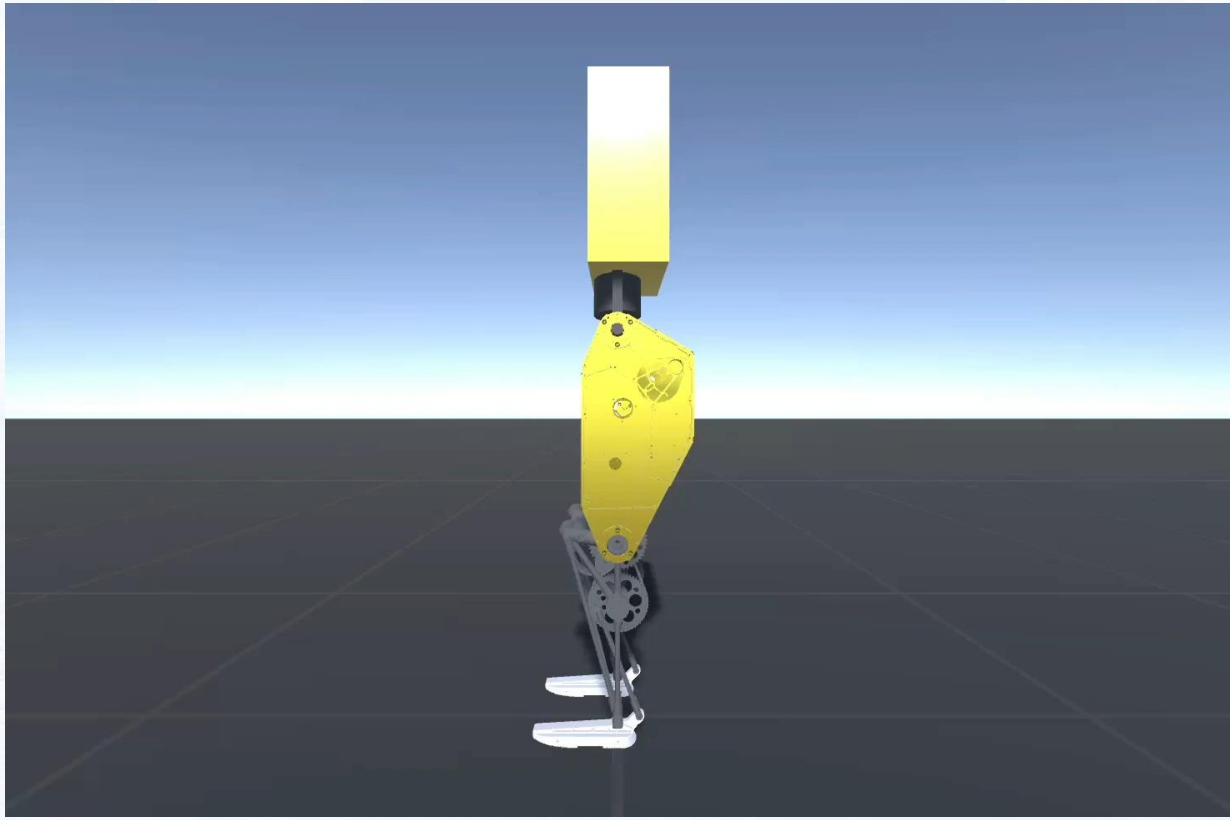
强化学习

单腿跳

Feedforward action

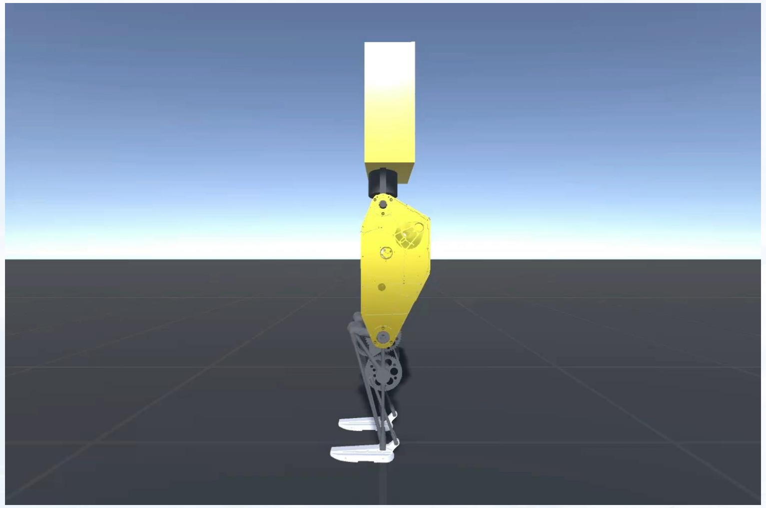


Learned motion

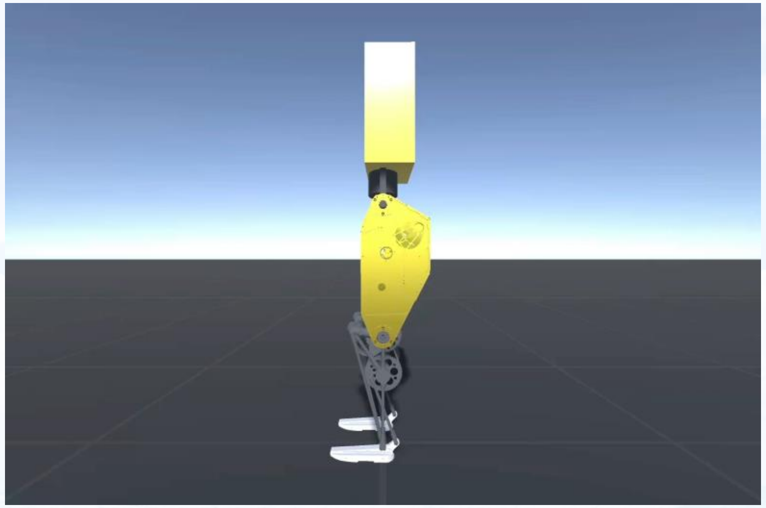


强化学习　　对前馈的适应性

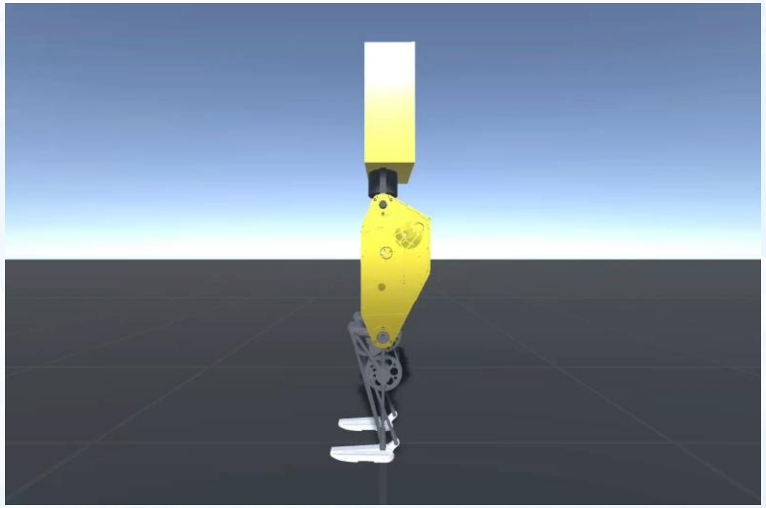
Learned motion



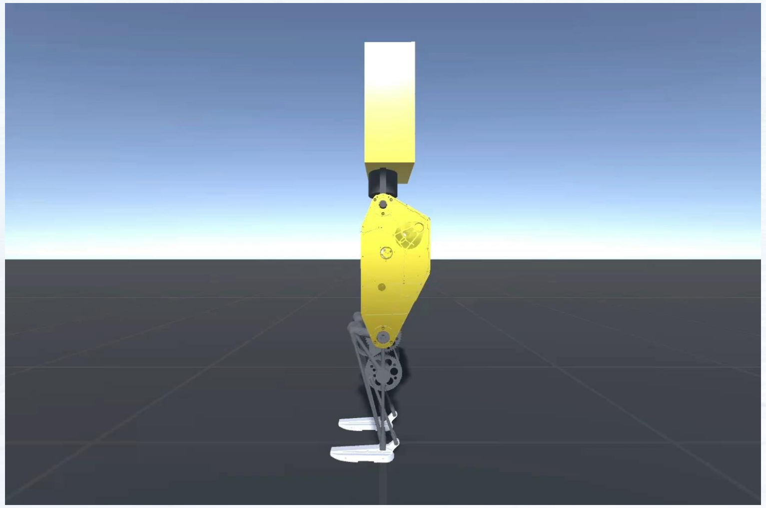
Period * 0.88



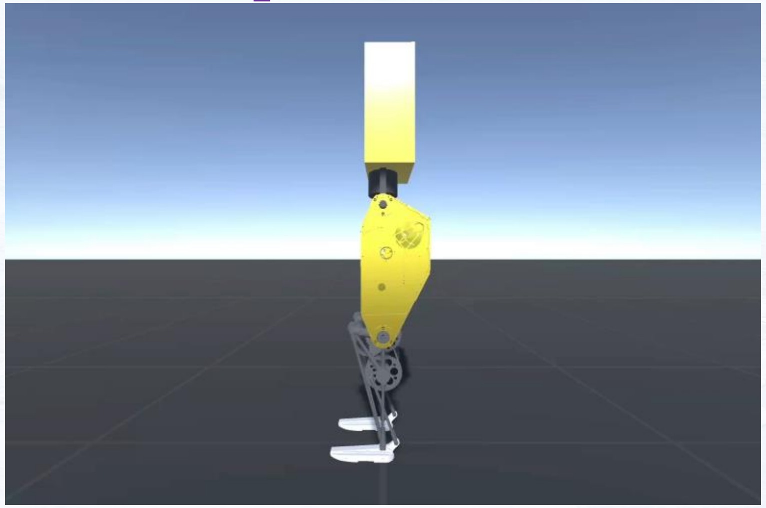
Period * 1.26



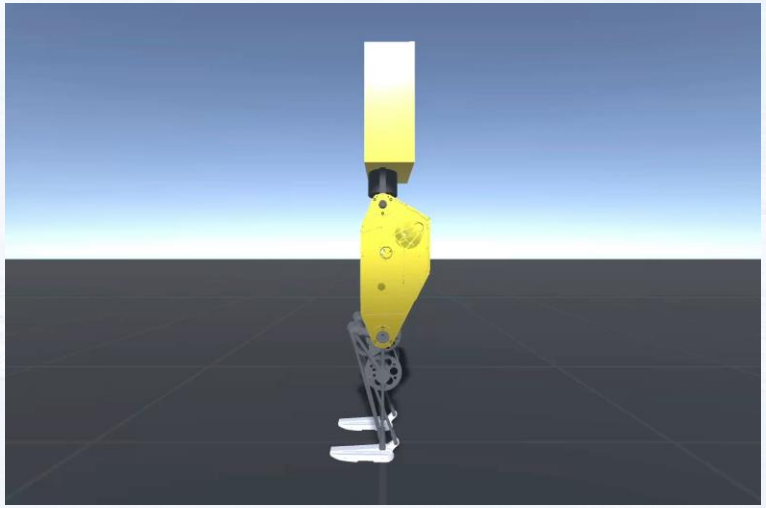
Learned motion



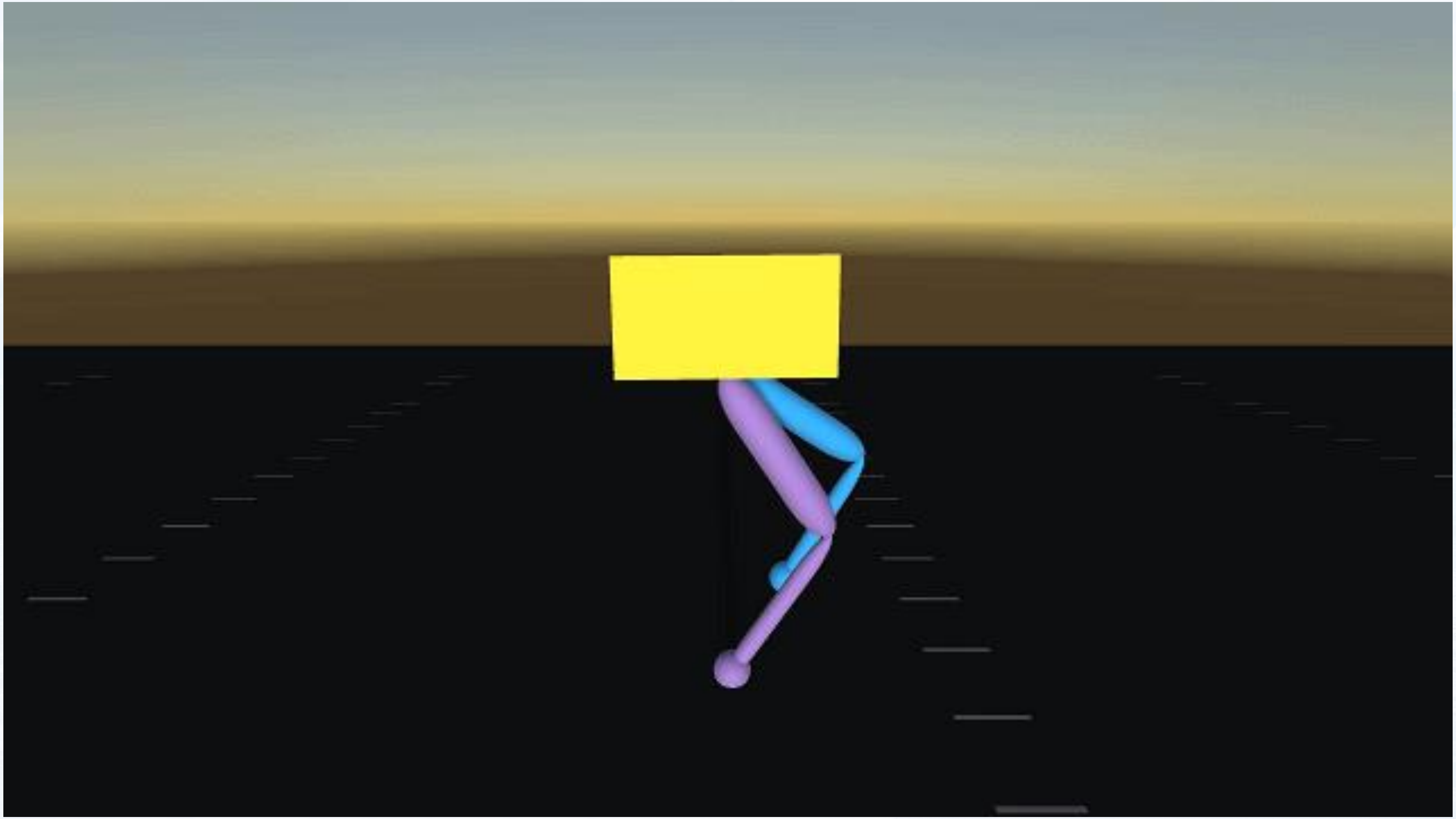
Amplitude * 0.88



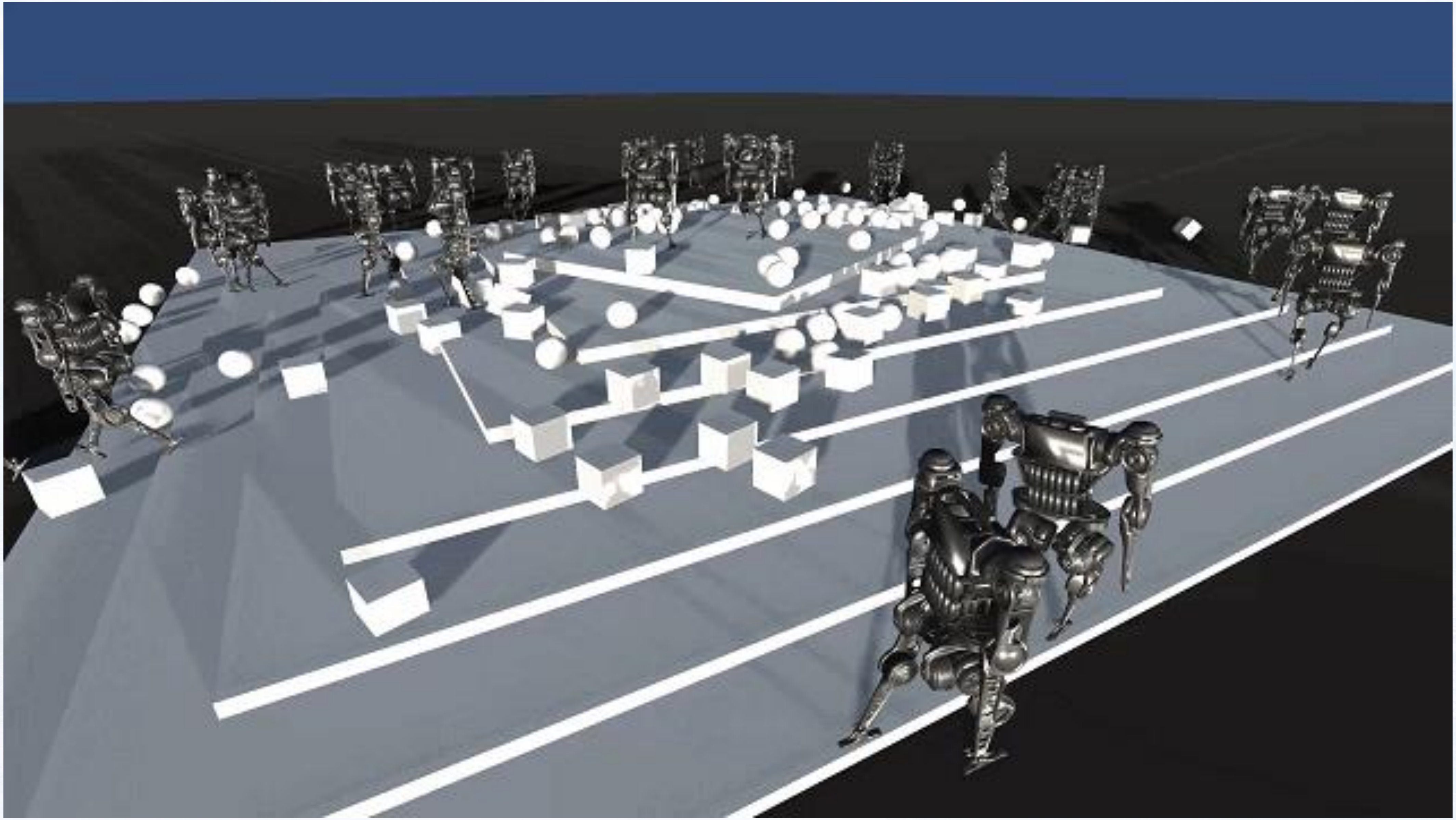
Amplitude * 2.08



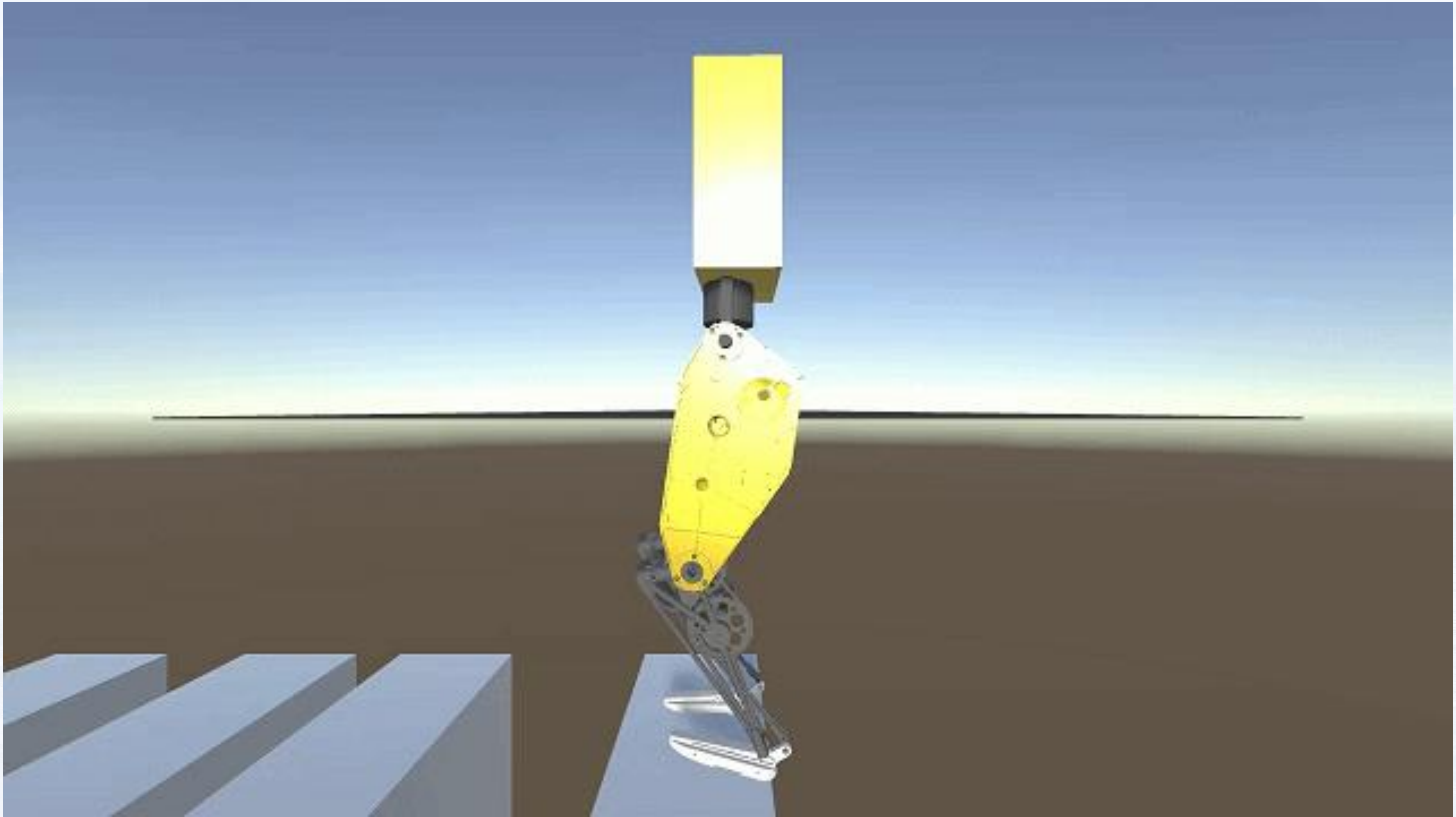
强化学习



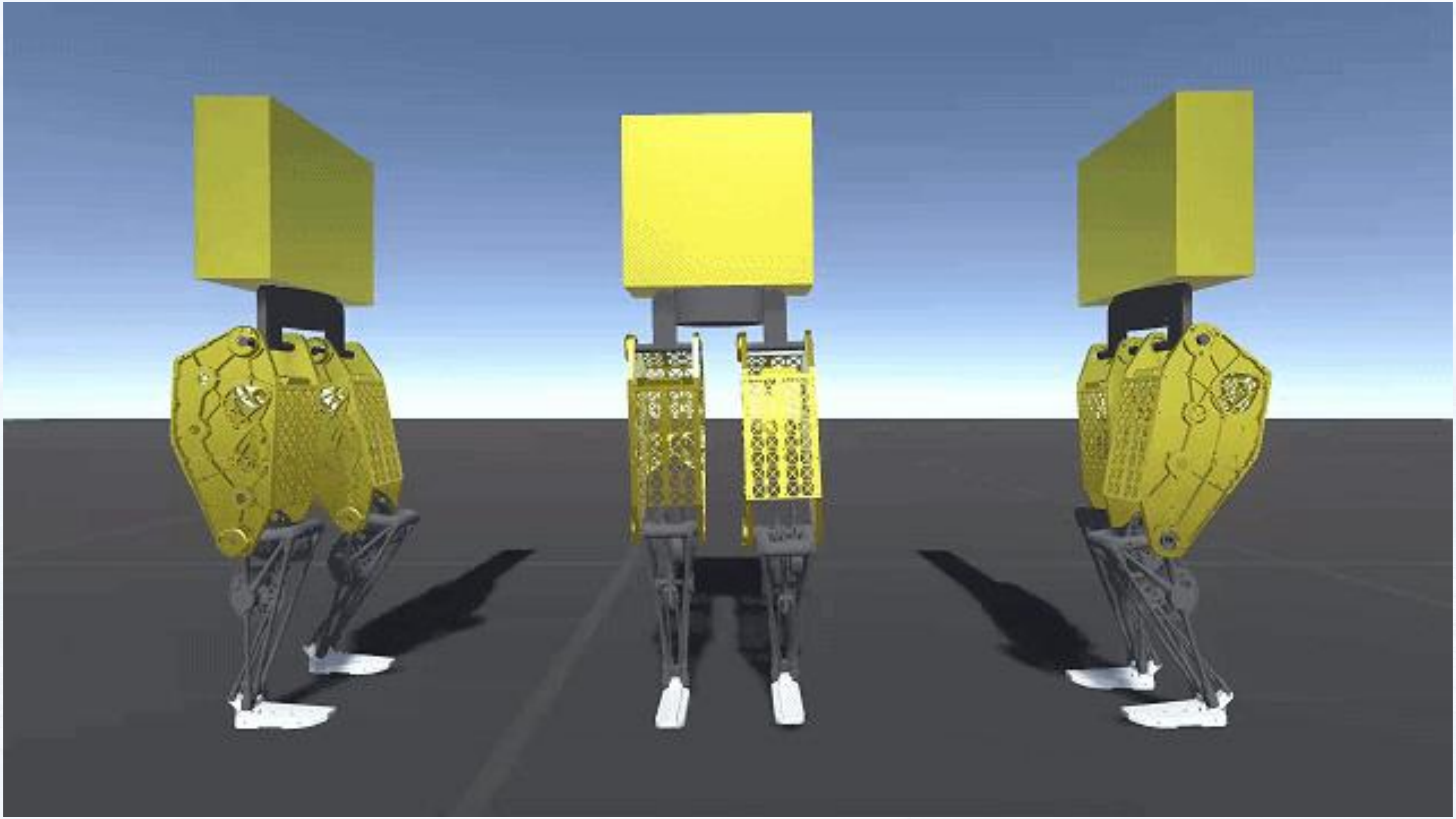
强化学习



强化学习



强化学习

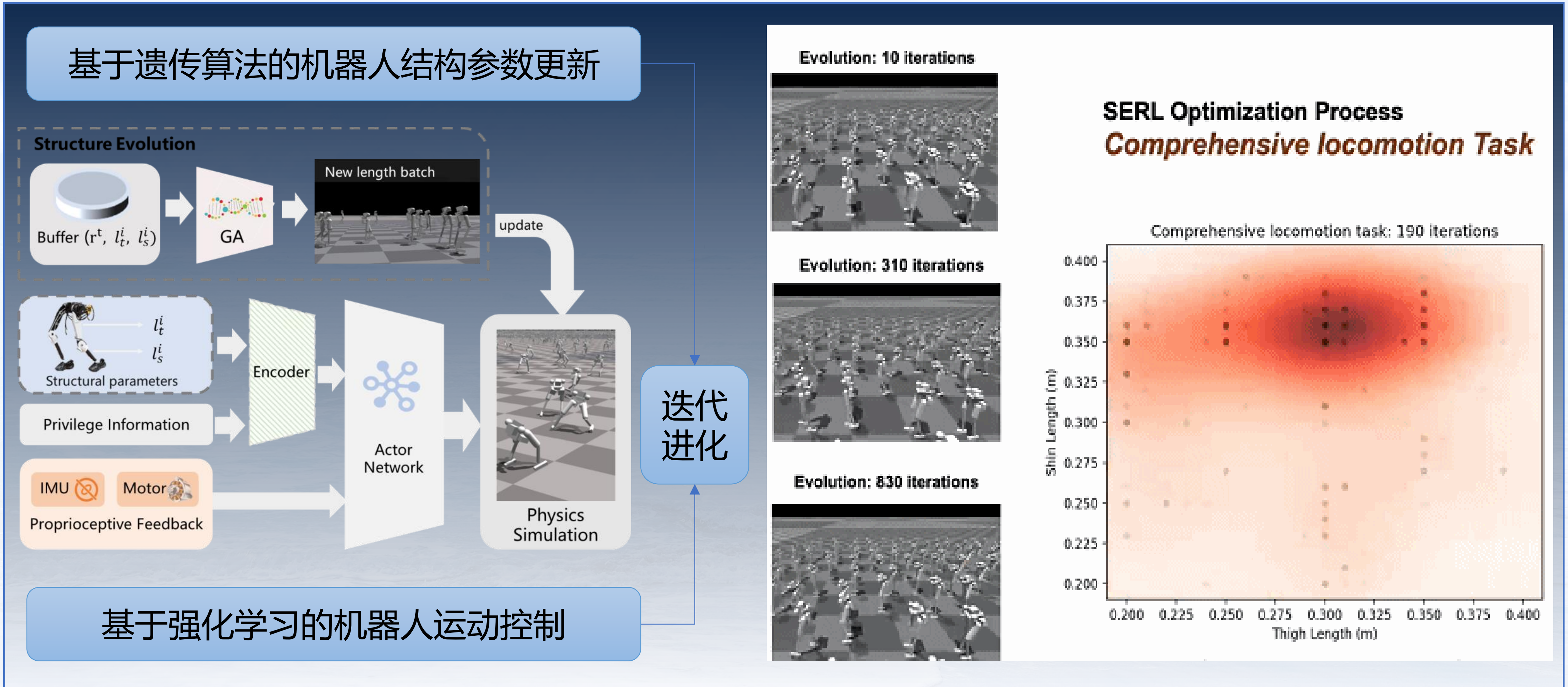


强化学习


Home Environment
Simulation
We built a two-story room scenario to test the robot's ability to move in a narrow and obstructed environment. Ankle and trunk target trajectories are manually set and we focus on verifying the tracking capabilities of our controllers.
unity

强化学习

利用遗传算法和强化学习，实现机器人结构参数和控制策略的协同进化



强化学习

获 奖 证 书

CERTIFICATE OF AWARD

江淮前沿技术协同创新中心、清华大学、上海大学 代表队：

在第二十六届中国机器人及人工智能大赛全国总决赛中，表现优异，成绩突出，荣获 人形机器人创新挑战赛（仿真赛）

一等奖

团队成员：郑函、程颐、张辉

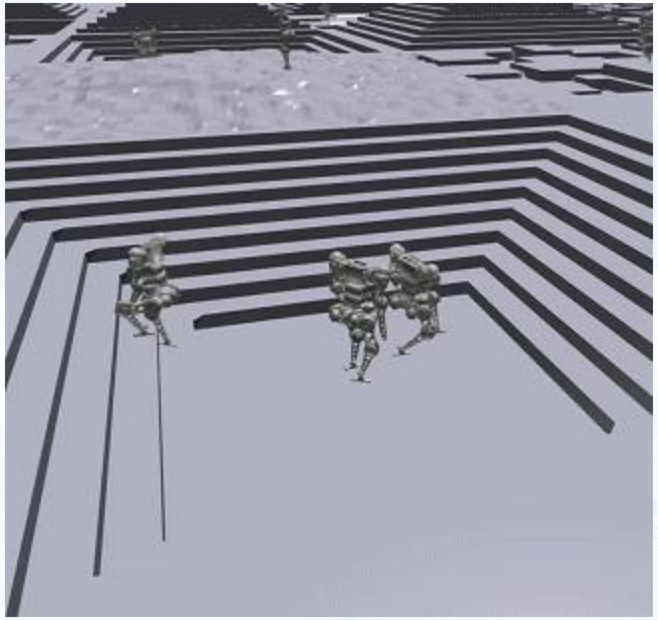
指导老师：刘厚德、叶林奇

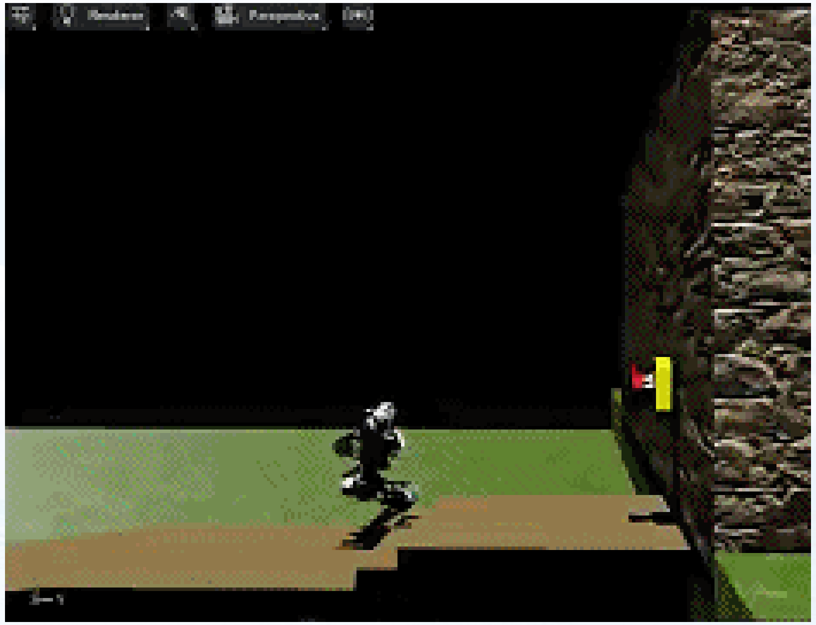
特发此证，以资鼓励。

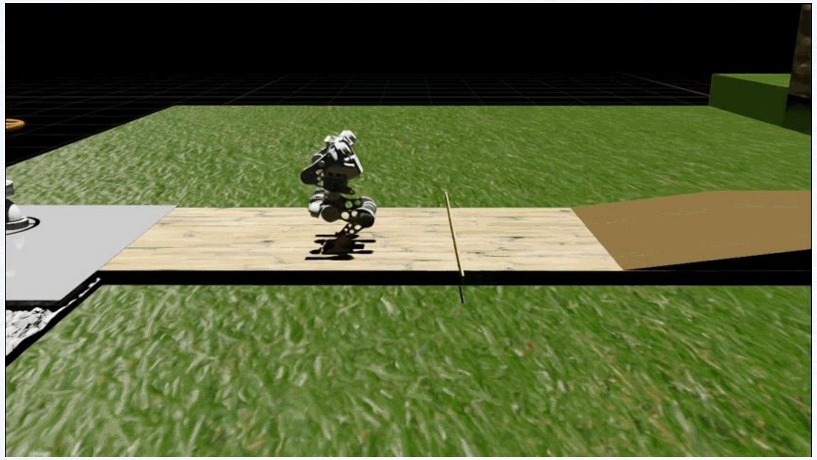
中国机器人及人工智能大赛组委会

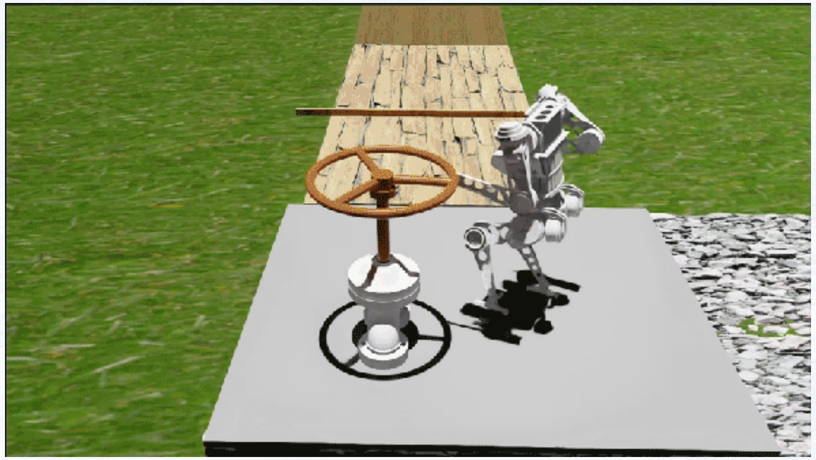
二〇二四年八月

证书编号：CRAIC2024-NF-BWZFZS



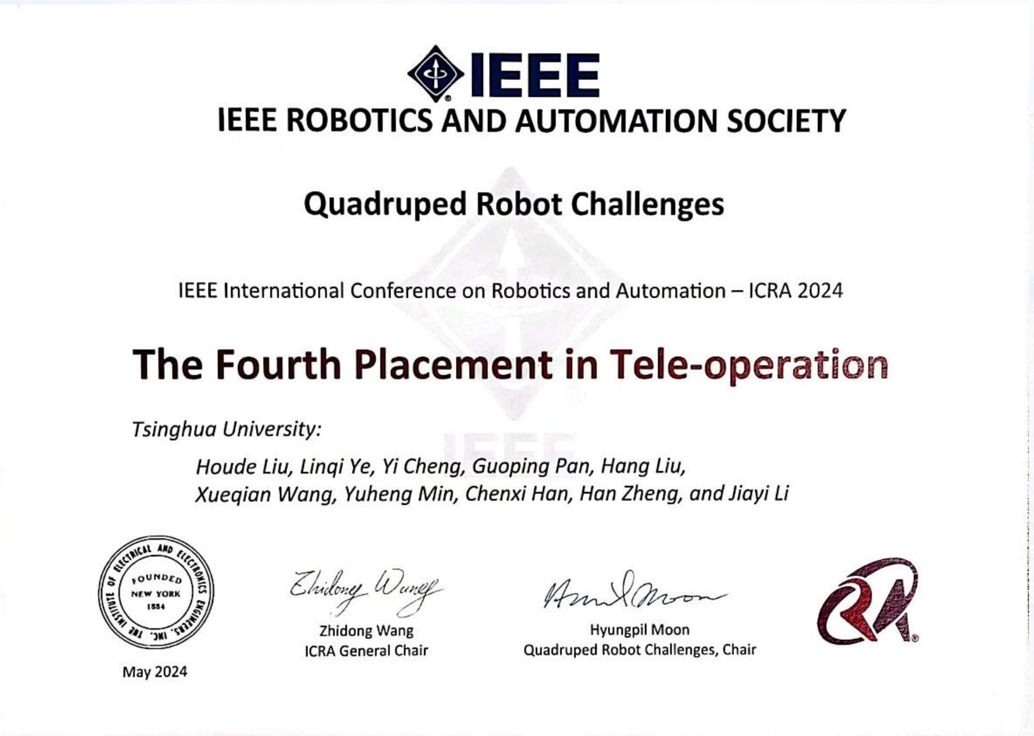






强化学习

THU-SHU ROBOART LAB



IEEE
IEEE ROBOTICS AND AUTOMATION SOCIETY

Quadruped Robot Challenges

IEEE International Conference on Robotics and Automation – ICRA 2024

The Fourth Placement in Tele-operation

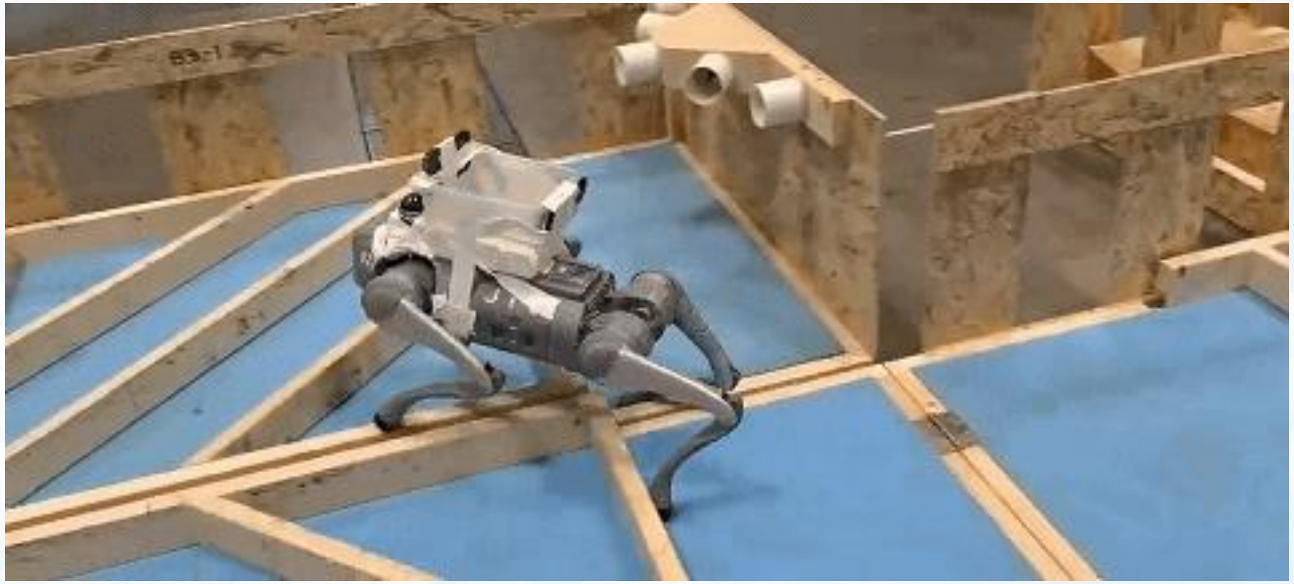
Tsinghua University:

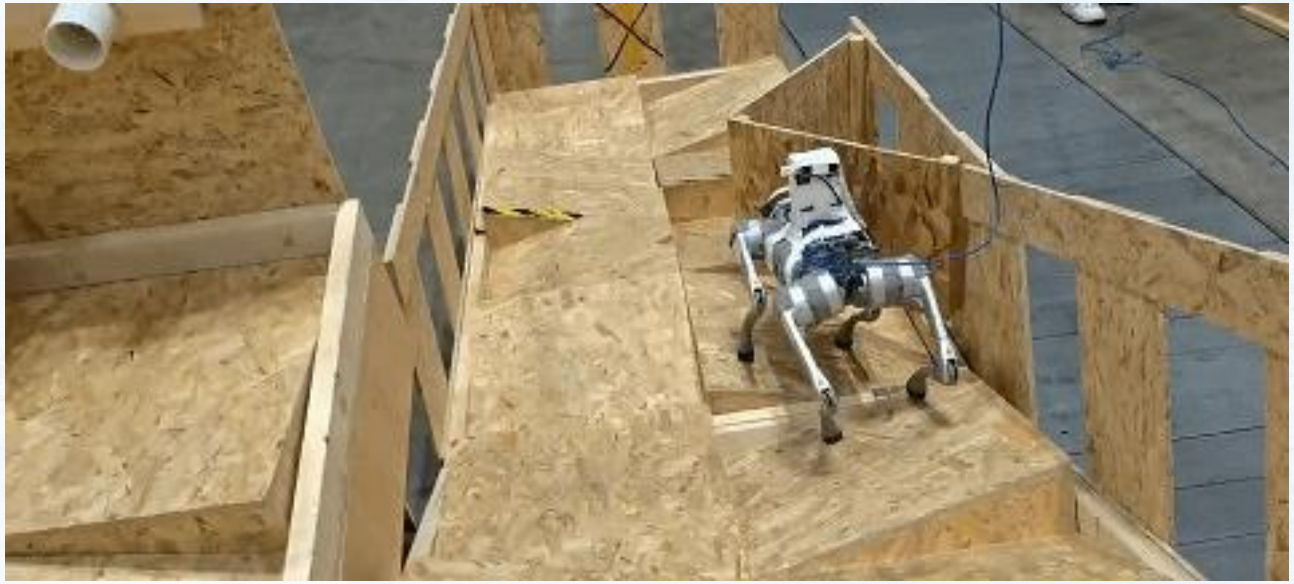
Houde Liu, Linqi Ye, Yi Cheng, Guoping Pan, Hang Liu,
Xueqian Wang, Yuheng Min, Chenxi Han, Han Zheng, and Jiayi Li

May 2024

Zhidong Wang
ICRA General Chair

Hyungpil Moon
Quadruped Robot Challenges, Chair



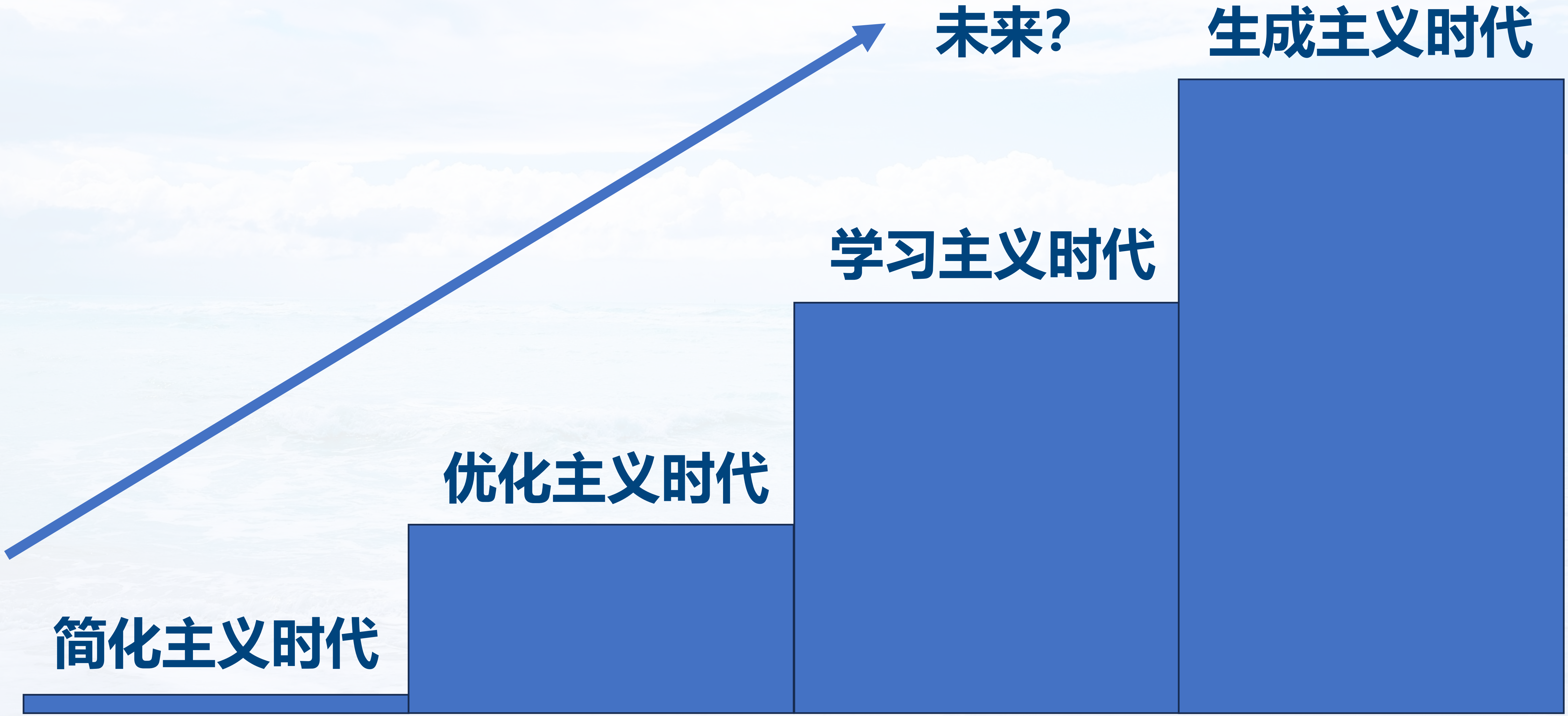


强化学习

人形机器人控制
从过去到未来

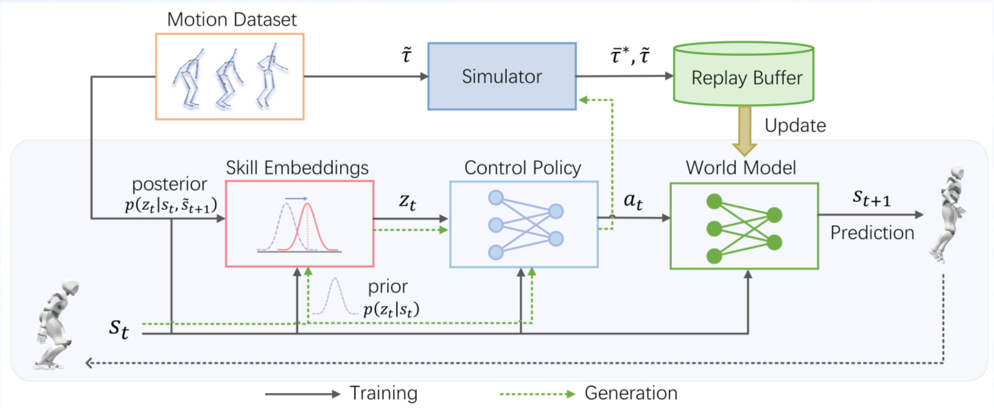
目录

未来

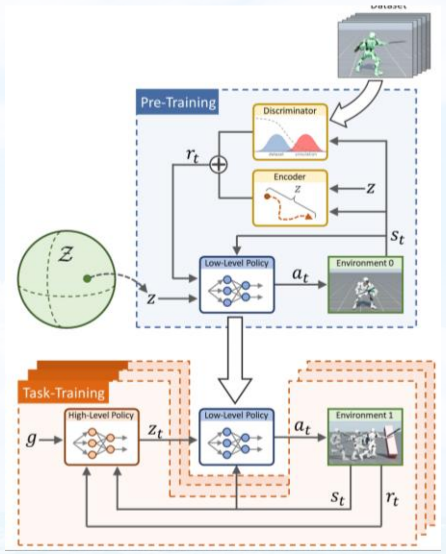


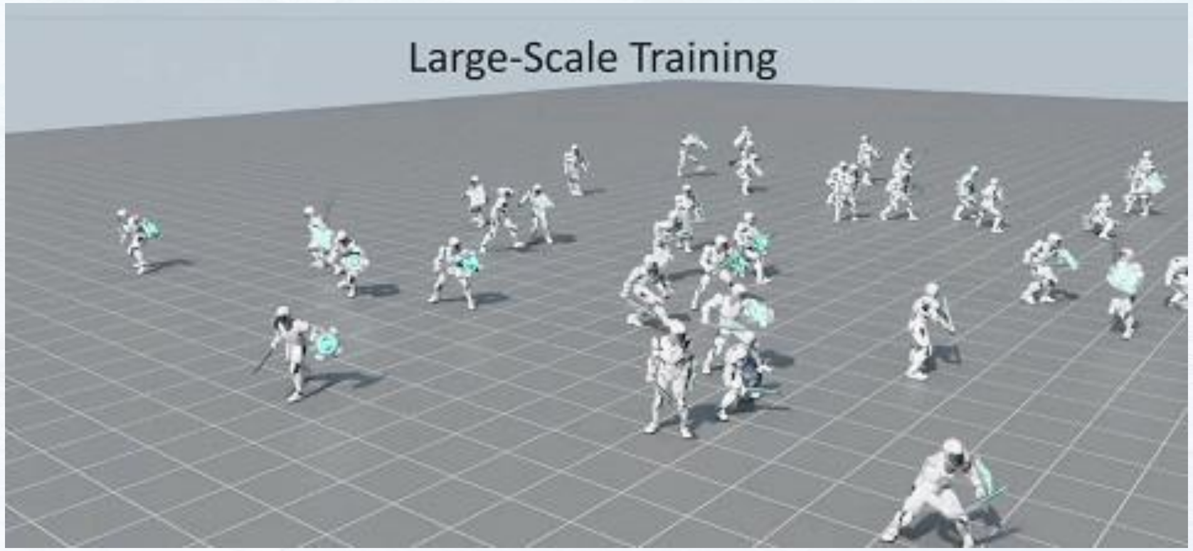
未来展望

ControlVAE: Model-Based Learning of Generative Controllers for Physics-Based Characters

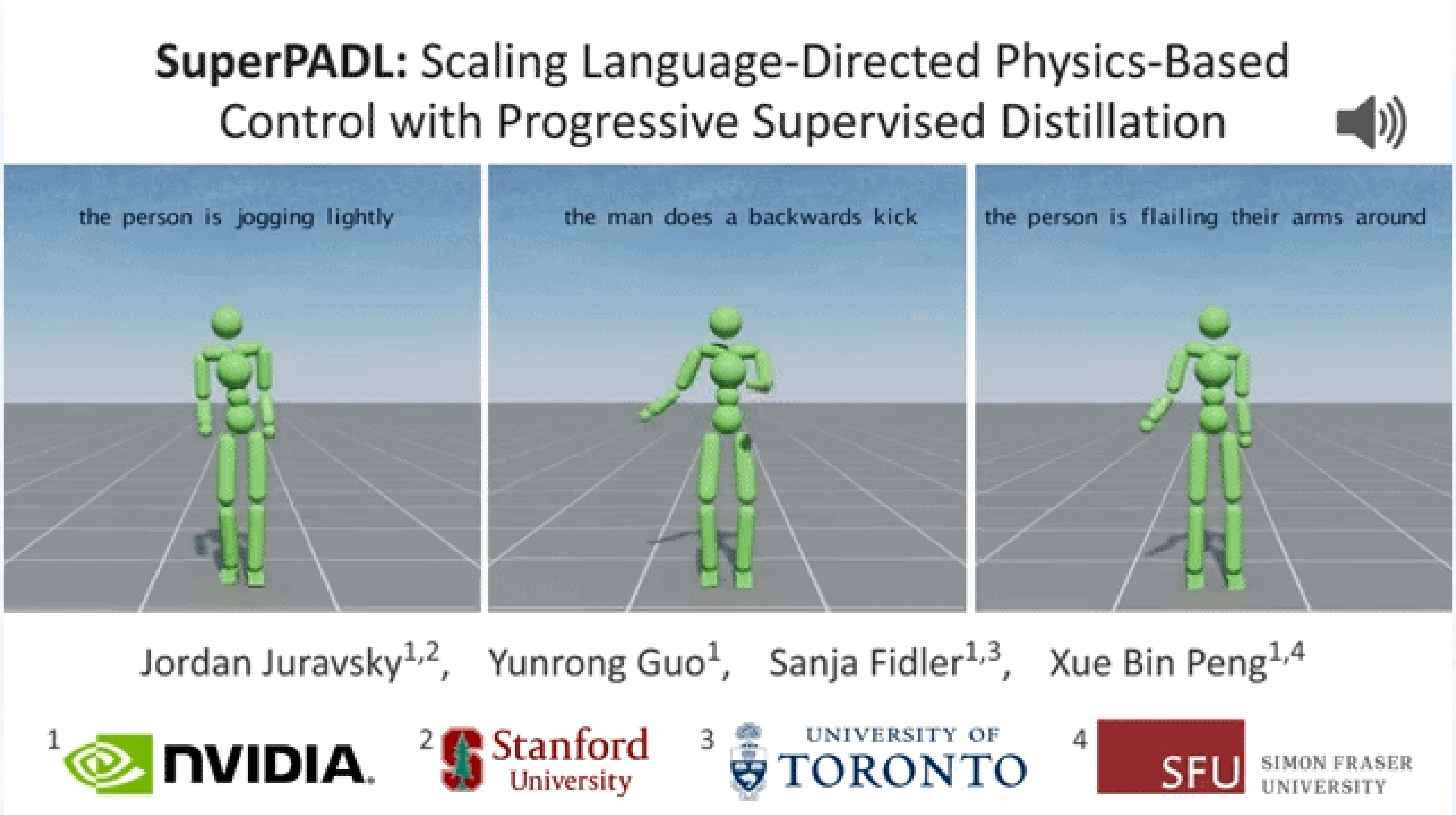


ASE: Large-Scale Reusable Adversarial Skill Embeddings for Physically Simulated Characters

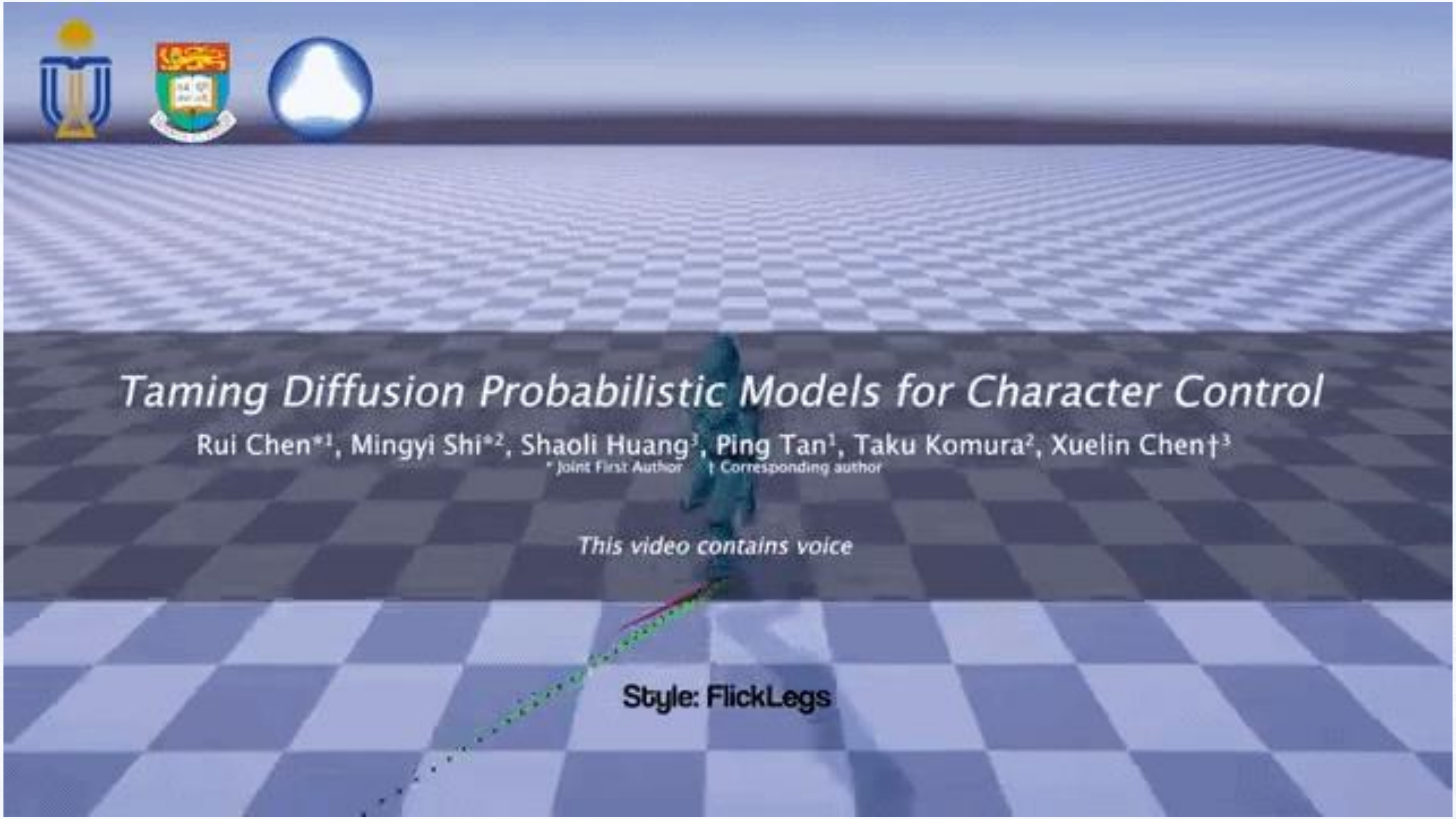




未来展望

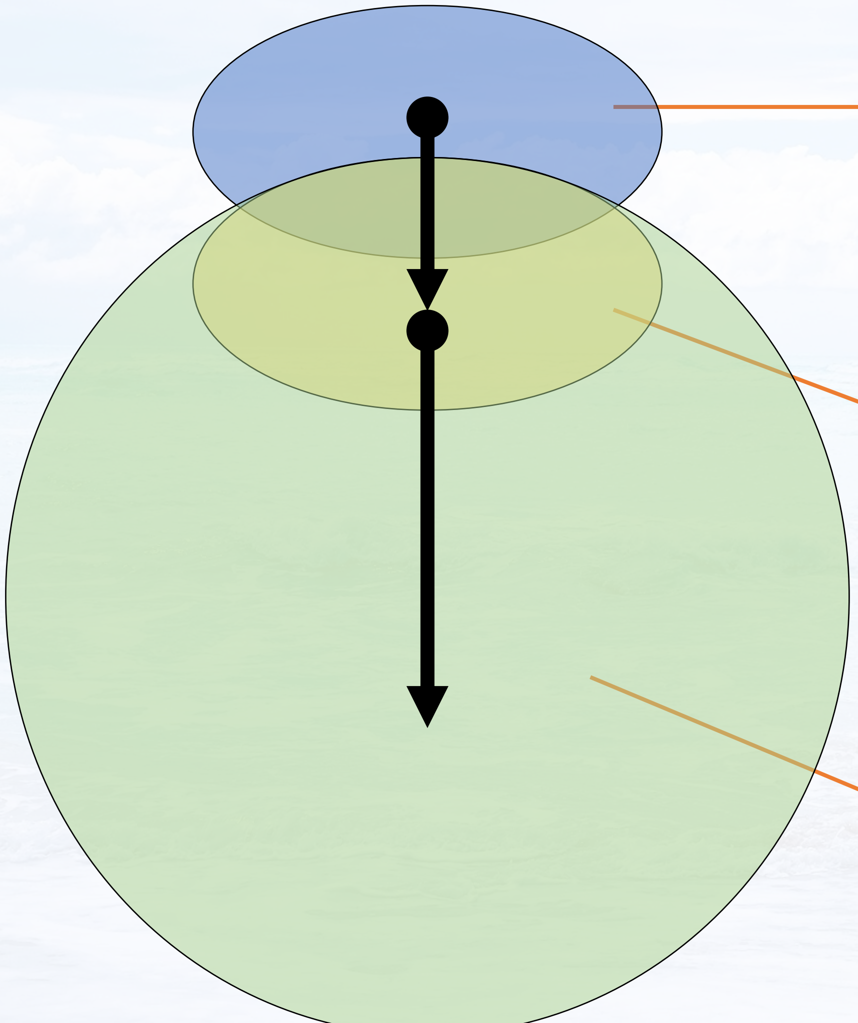

SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation
the person is jogging lightly
the man does a backwards kick
the person is flailing their arms around
Jordan Juravsky1,2, Yunrong Guo1, Sanja Fidler1,3, Xue Bin Peng1,4
1 NVIDIA.
2 Stanford University
3 UNIVERSITY OF TORONTO
4 SFU SIMON FRASER UNIVERSITY

未来展望


Taming Diffusion Probabilistic Models for Character Control
Rui Chen*1, Mingyi Shi*2, Shaoli Huang3, Ping Tan1, Taku Komura2, Xuelin Chen†3
* Joint First Author
† Corresponding author
This video contains voice
Style: FlickLegs

未来展望

人形机器人未来控制范式：生成式预训练 + 实物在线学习



- **生成式预训练**：多样化数据库、高保真化建模
- **Sim2Real**：虚实空间对齐、虚实差异补偿
- **真实开放环境持续学习**：持续在线学习、大小模型协同

未来展望

THU-SHU ROBOART LAB

未来展望

谢谢

https://linqi-ye.github.io/

THU-SHU ROBOART LAB