具身智能-08

刘华平

2025年4月9日

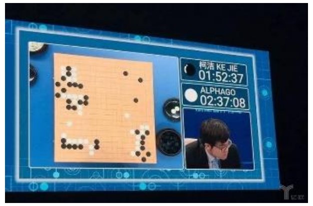
课程内容安排

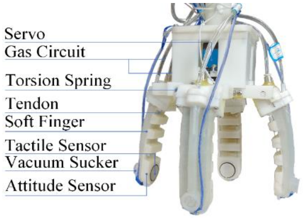
课次	周次	上课内容	软件
1	1	绪论	
2	2	深度学习	
3	3	强化学习1	Gym, Mujuco
4	4	强化学习2	Gym, Mujuco
	5	作业准备	
5	6	自监督与持续学习	
	7	开题	Powerpoint
6	8	形态智能	Gym, Mujuco
7	9	视觉导航：VLN	AI2THOR
	10	主动感知：VSN，EQA	AI2THOR
8	11	五一放假	
9	12	具身学习	AI2THOR
10	13	多体智能	AI2THOR
11	14	面向具身智能的AIGC	AI2THOR
	15-16	成果准备与展示	Powerpoint

具身智能的体系

➢ 狭义与广义的具身智能

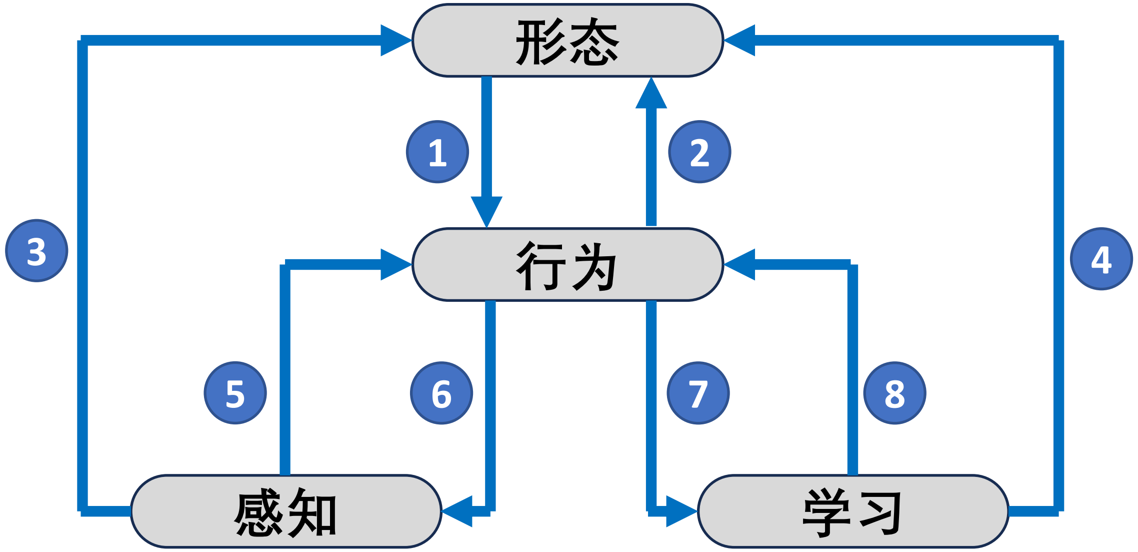
离 身 智 能　　　　具 身 智 能

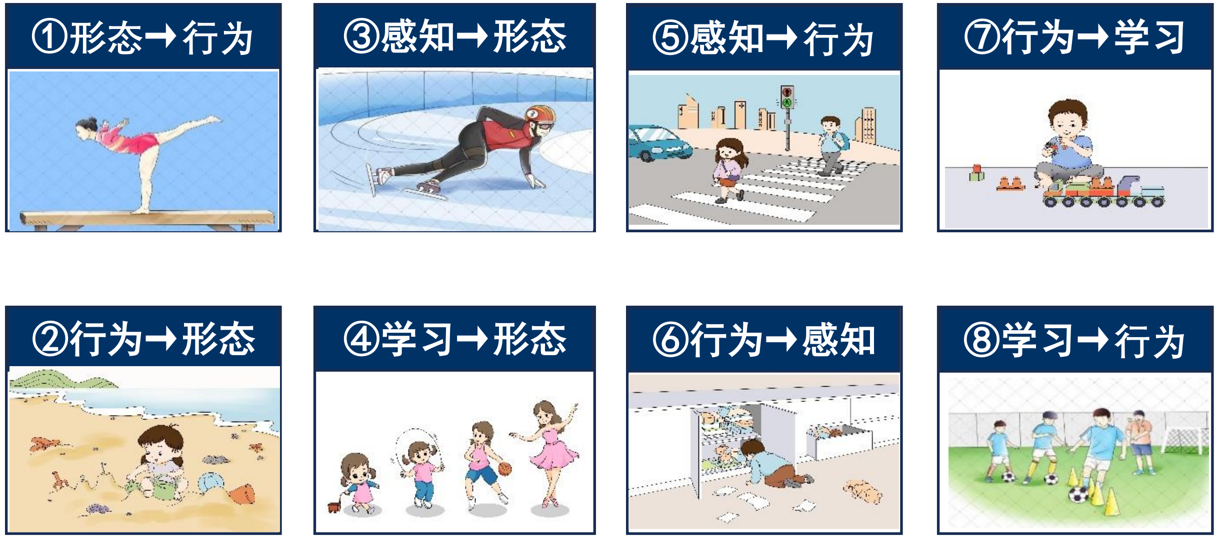




具身智能的体系

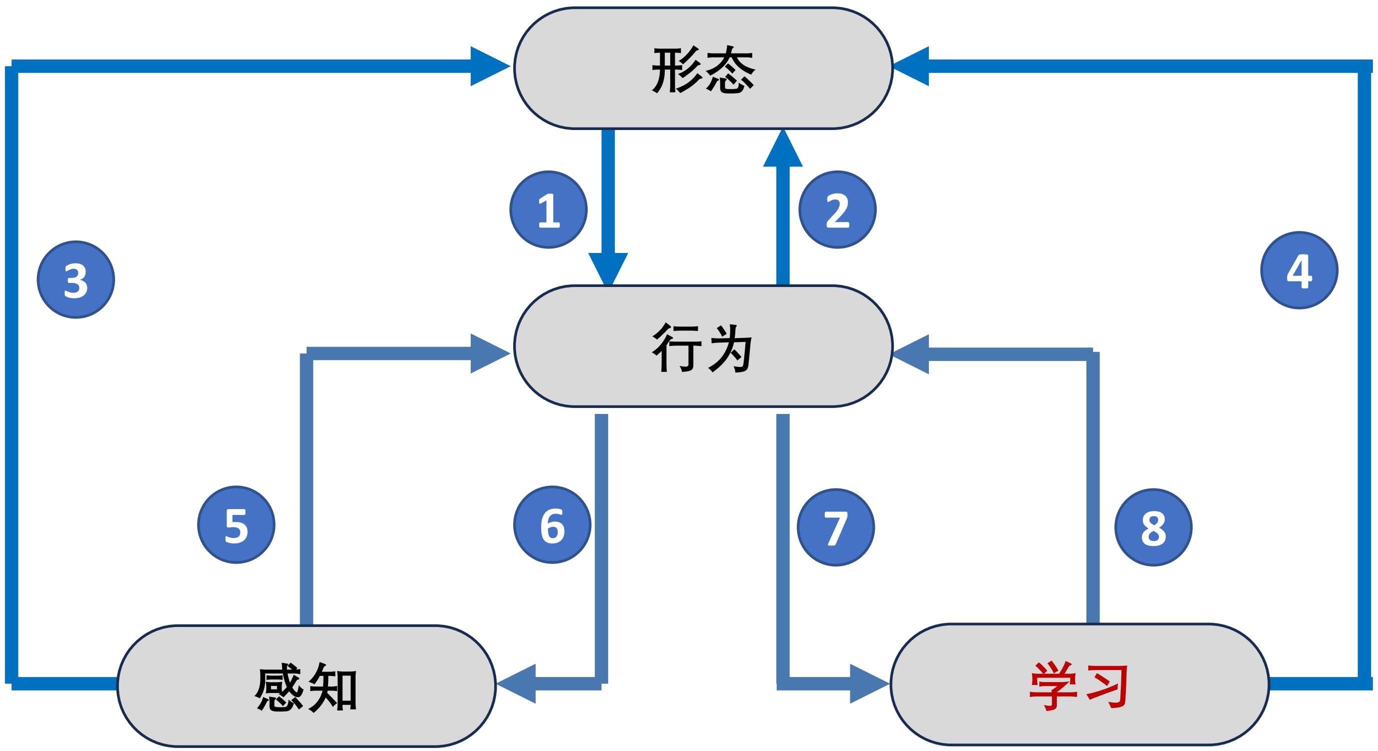
➢ 具身智能的体系结构



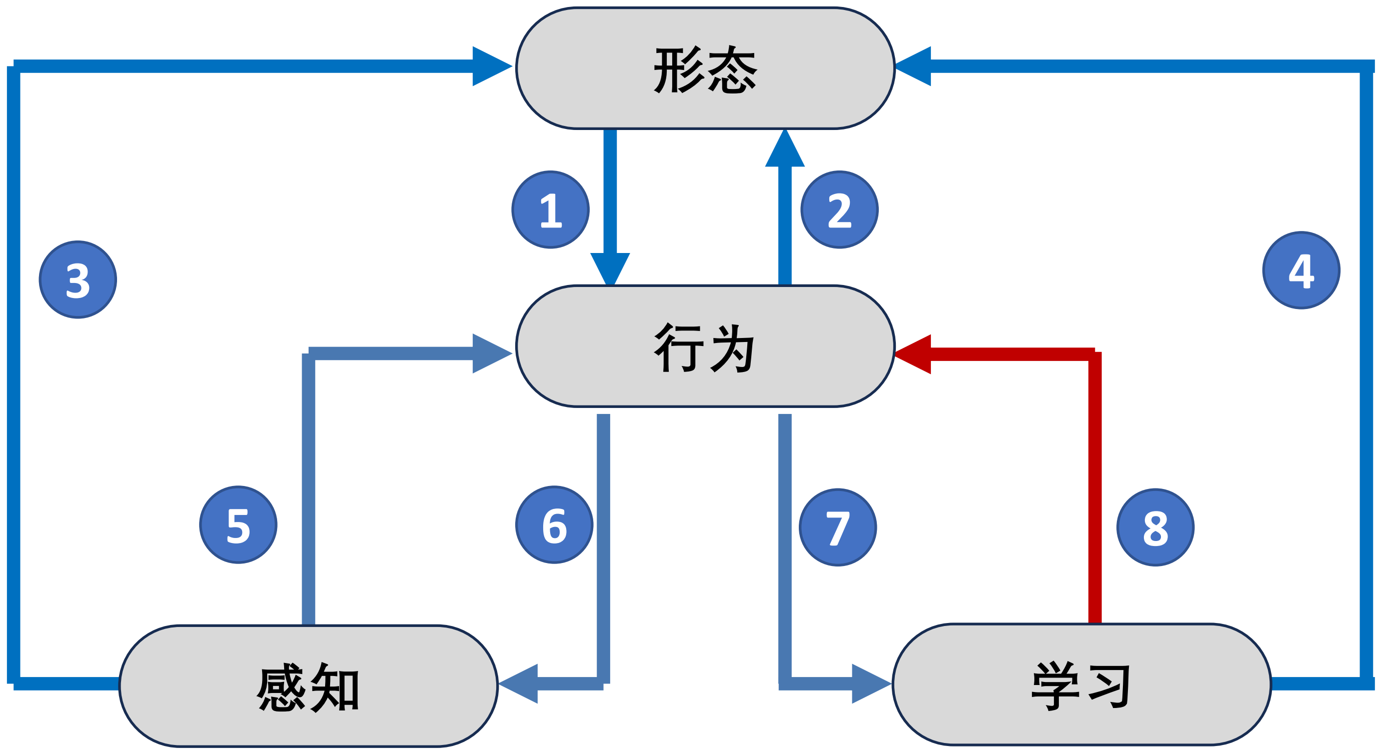


① 基于形态的行为生成	③ 基于感知的形态变换	⑤ 基于感知的行为生成	⑦ 基于行为的自主学习
② 基于行为的形态控制	④ 基于学习的形态优化	⑥ 基于行为的主动感知	⑧ 基于学习的行为优化

关键技术



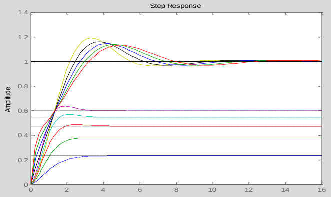
关键技术

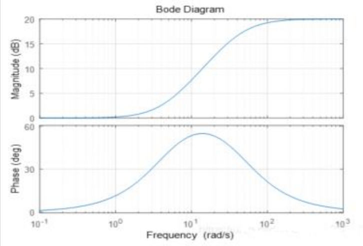


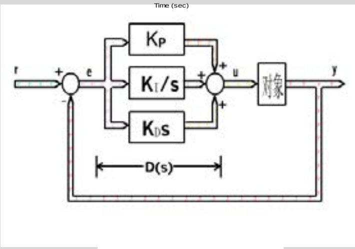
0 学习→行为：强化学习

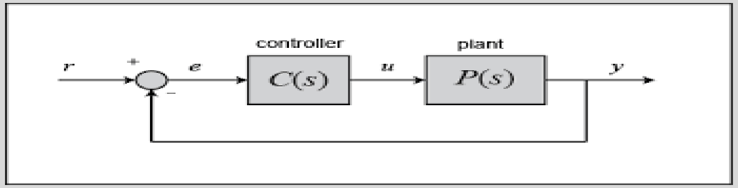
经典控制：经验

$$G_c(s)=\frac{s+\frac{1}{T}}{s+\frac{1}{\alpha T}}\quad(\alpha<1)$$

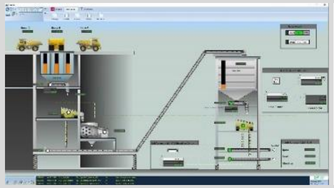


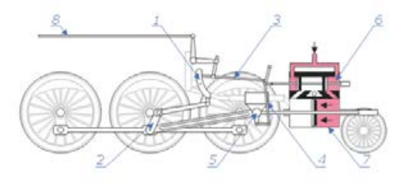






$$u_t=K_P e_t+K_I\int e_t dt+K_D\frac{de_t}{dt}$$





现代控制：模型

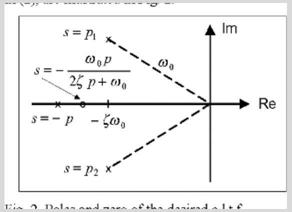
state equations

$$\begin{cases}\dot{x}_1=f_1(x_1,x_2,\ldots,x_n,u_1,\ldots,u_m)\\ \vdots\\ \dot{x}_n=f_n(x_1,x_2,\ldots,x_n,u_1,\ldots,u_m)\end{cases}$$

output equations

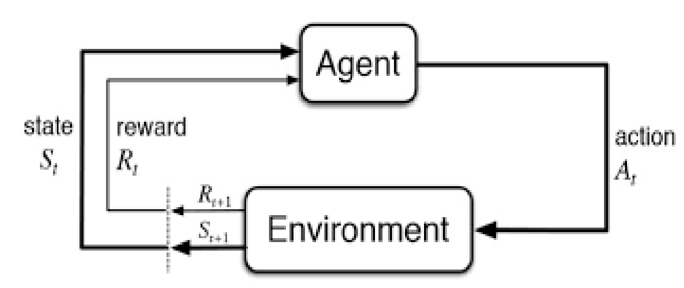
$$\begin{cases}y_1=h_1(x_1,x_2,\ldots,x_n,u_1,\ldots,u_m)\\ \vdots\\ y_p=h_p(x_1,x_2,\ldots,x_n,u_1,\ldots,u_m)\end{cases}$$

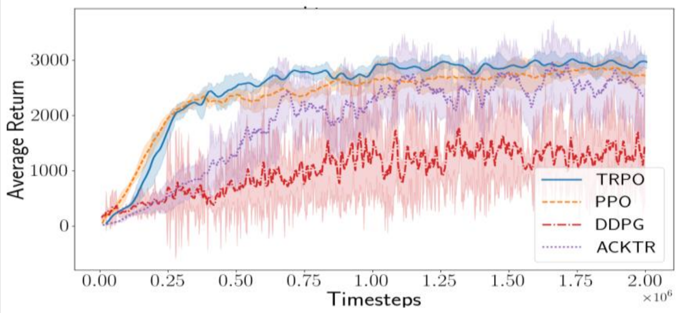
$$\begin{cases}\dot{x}_t=Ax_t+Bu_t\\ y_t=Cx_t+Du_t\end{cases}$$



$$u_t=-Kx_t$$

智能控制：学习

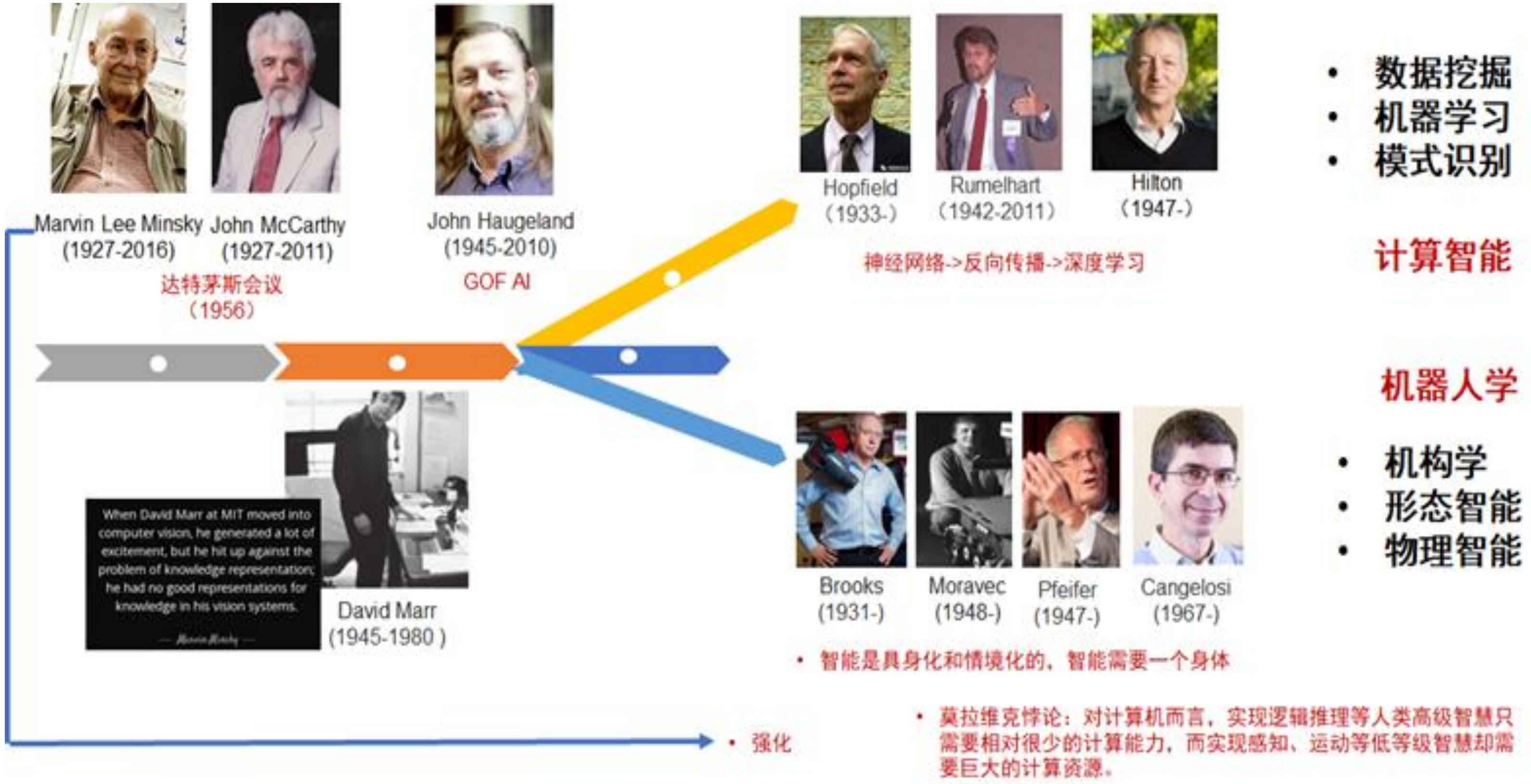




$$\pi_\theta(a_t\mid s_t)$$

0 学习→行为：强化学习

➢ 起源



人工神经网络
Artificial Neural Networks
Walter Pitts, 1943

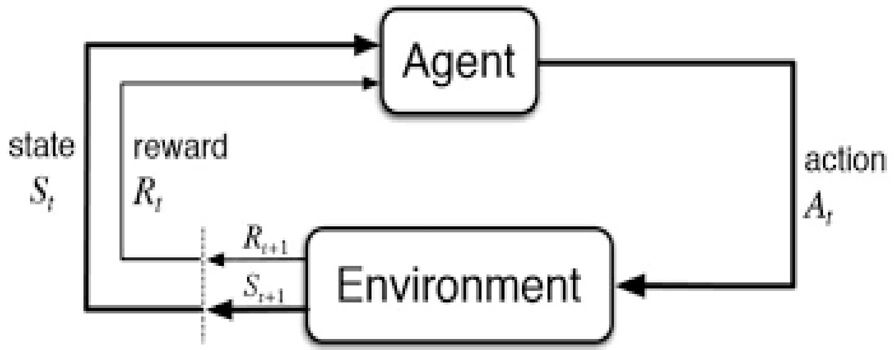
深度学习
Deep Learning
Geoffrey Hinton, 2006

深度强化学习
DQN
DeepMind, 2015

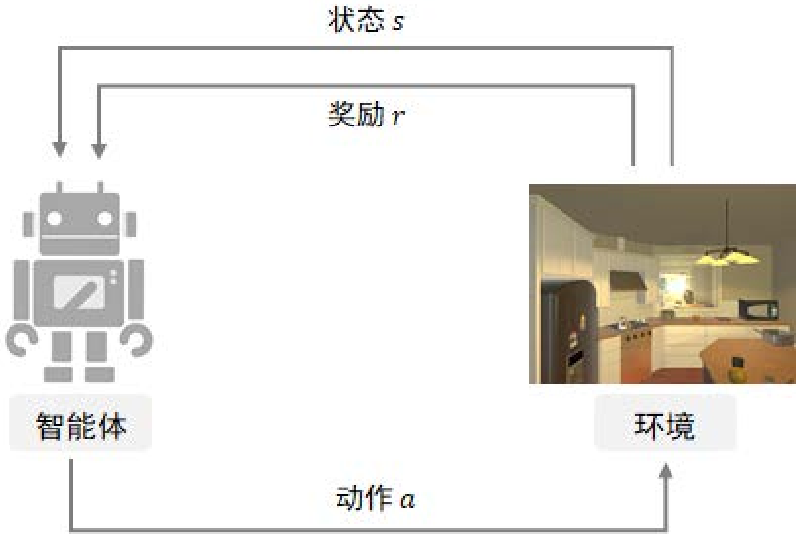
最优控制与动态规划
Markov Decision Process
Richard Bellman, 1957

时序差分与最优控制
Q-learning
Chris Watkins, 1989

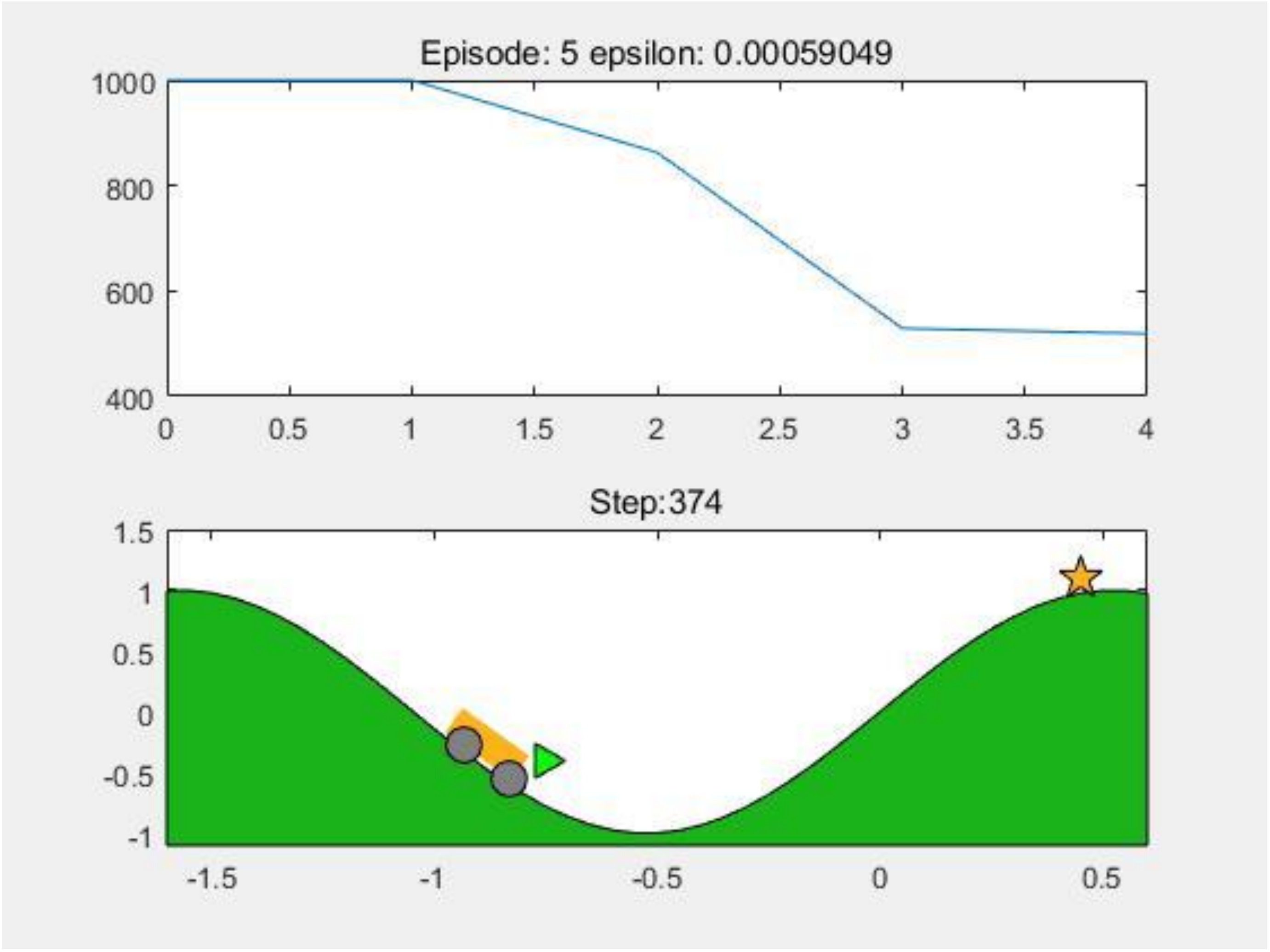
0 学习➔行为：强化学习



Agent
state
S_t
reward
R_t
action
A_t
R_{t+1}
S_{t+1}
Environment



状态 s
奖励 r
智能体
环境
动作 a



Episode: 5 epsilon: 0.00059049
1000
800
600
400
0
0.5
1
1.5
2
2.5
3
3.5
4
Step:374
1.5
1
0.5
0
-0.5
-1
-1.5
-1
-0.5
0
0.5

0 学习→行为：强化学习

➢ 两类基本算法

$$\max_{\theta} \quad \mathrm{E}[\sum_{t=0}^{H} R(s_t)|\pi_\theta]$$

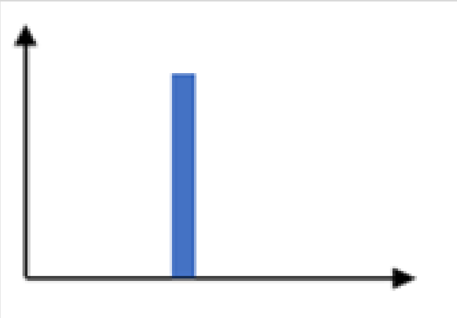
Q-学习算法

$$Q(s_t,a_t)=r_{t+1}+\gamma r_{t+2}+\gamma^2 r_{t+3}+\gamma^3 r_{t+4}+\cdots$$

$$Q(s_t,a_t)=r_{t+1}+\gamma Q(s_{t+1},a_{t+1})$$

$$Q(s_t,a_t)\leftarrow Q(s_t,a_t)+\alpha\left[r+\gamma \max_a Q(s_{t+1},a)-Q(s_t,a_t)\right]$$

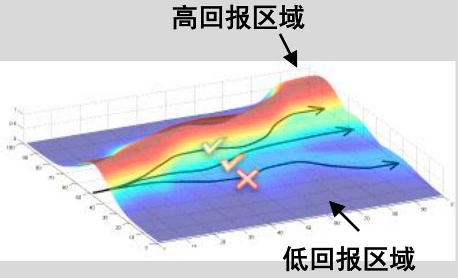
$$\pi^*(s)=\arg\max_{a\in\mathcal{A}} Q^*(s,a)$$



发展：DQN

策略梯度算法

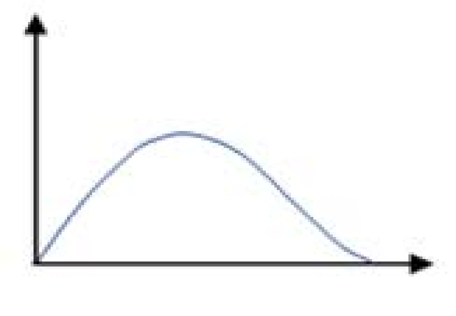
$$J(\pi)=\mathbb{E}_{\tau\sim p_\pi(\tau)}\left[\sum_{t=0}^{\infty}\gamma^t r(s_t,a_t,s_{t+1})\right]$$



$$\nabla_\theta J(\pi_\theta)\approx\frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{H^{(n)}}R(\tau^{(n)})\nabla_\theta \log\pi_\theta(a_t^{(n)}\mid s_t^{(n)})$$

$$\theta\leftarrow\theta+\eta\nabla_\theta J$$

$$\pi_\theta(a\mid s)$$



发展：REINFORCE

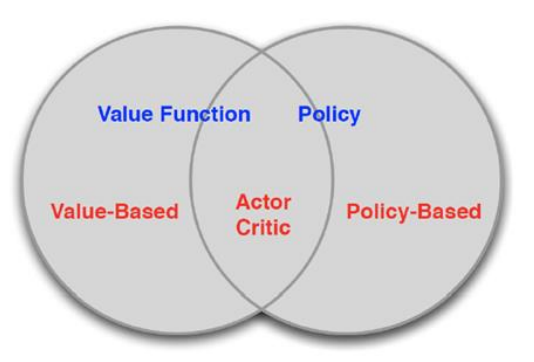
0 学习→行为：强化学习

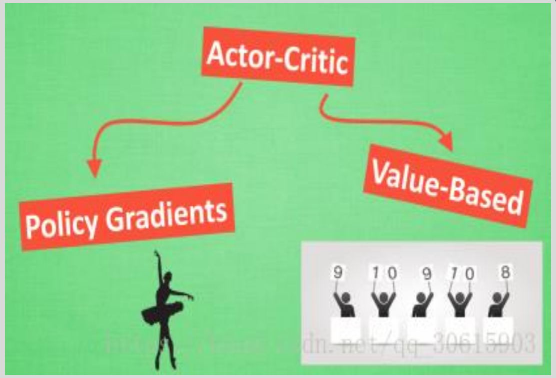
➢ 两类典型算法

Actor-Critic

$$\nabla_\theta J(\pi_\theta) = \mathbb{E}_{s\sim p_0(s)}\mathbb{E}_{a\sim\pi_\theta(\cdot|s)}\left(Q^{\pi_\theta}(s,a)\cdot\nabla_\theta \log \pi_\theta(a\,|\,s)\right)$$

$$\hat{Q}_\omega(s,a) \approx Q^{\pi_\theta}(s,a) \qquad \nabla_\theta J(\pi_\theta) = \hat{Q}_\omega(s_t,a_t)\cdot\nabla_\theta \log \pi_\theta(a_t\,|\,s_t)$$





PPO

$$A^{\pi_\theta}(s_t,a_t) = Q^{\pi_\theta}(s_t,a_t) - V^{\pi_\theta}(s_t)$$

$$\mathbb{E}_{\tau\sim p_{\pi_\theta}(\tau)}\left(\sum_{t=0}^{\infty}\gamma^t A^{\pi_{\theta\text{old}}}(s_t,a_t)\right)$$

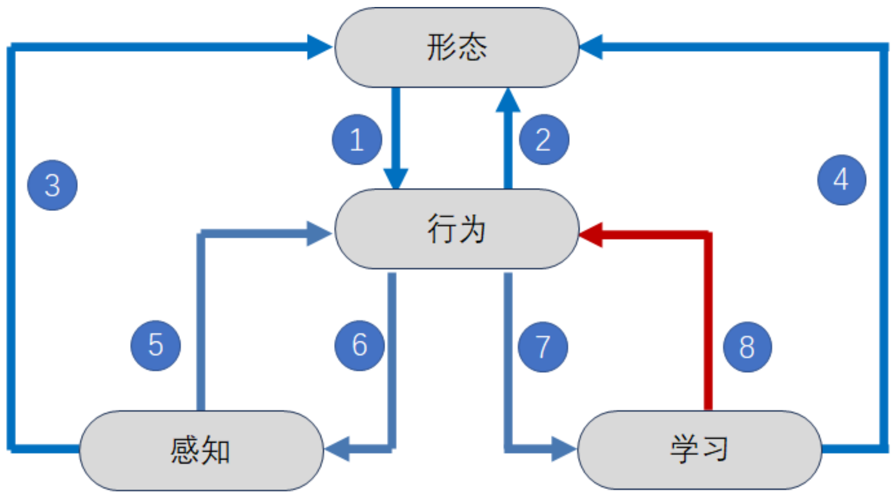
$$L_{\pi_{\theta\text{old}}}(\pi_\theta) = \mathbb{E}_{\tau\sim p_{\pi_{\theta\text{old}}}(\tau)}\left(\sum_{t=0}^{\infty}\gamma^t \rho_t(\theta) A^{\pi_{\theta\text{old}}}(s_t,a_t)\right)$$

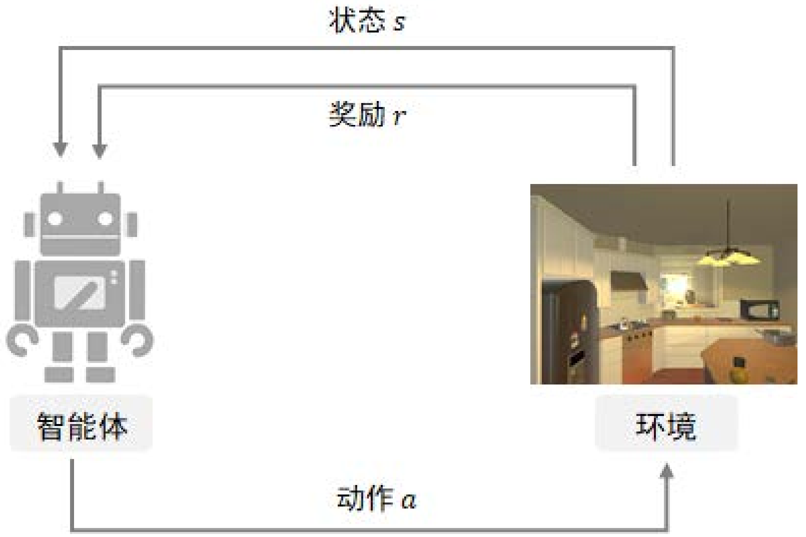
$$L_{\pi_{\theta\text{old}}}(\pi_\theta) = \mathbb{E}_{\tau\sim p_{\pi_{\theta\text{old}}}(\tau)}\left(\sum_{t=0}^{\infty}\gamma^t \min\left\{(\rho_t(\theta) A^{\pi_{\theta\text{old}}}(s_t,a_t), \text{clip}(\rho_t(\theta),1-\varepsilon,1+\varepsilon)A^{\pi_{\theta\text{old}}}(s_t,a_t)\right\}\right)$$

$$\text{clip}(\rho_t(\theta),1-\varepsilon,1+\varepsilon) = \begin{cases} \rho_t(\theta) & 1-\varepsilon \le \rho_t(\theta) \le 1+\varepsilon \\ 1-\varepsilon & \rho_t(\theta) < 1-\varepsilon \\ 1+\varepsilon & \rho_t(\theta) > 1+\varepsilon \end{cases}$$

0 学习→行为：强化学习

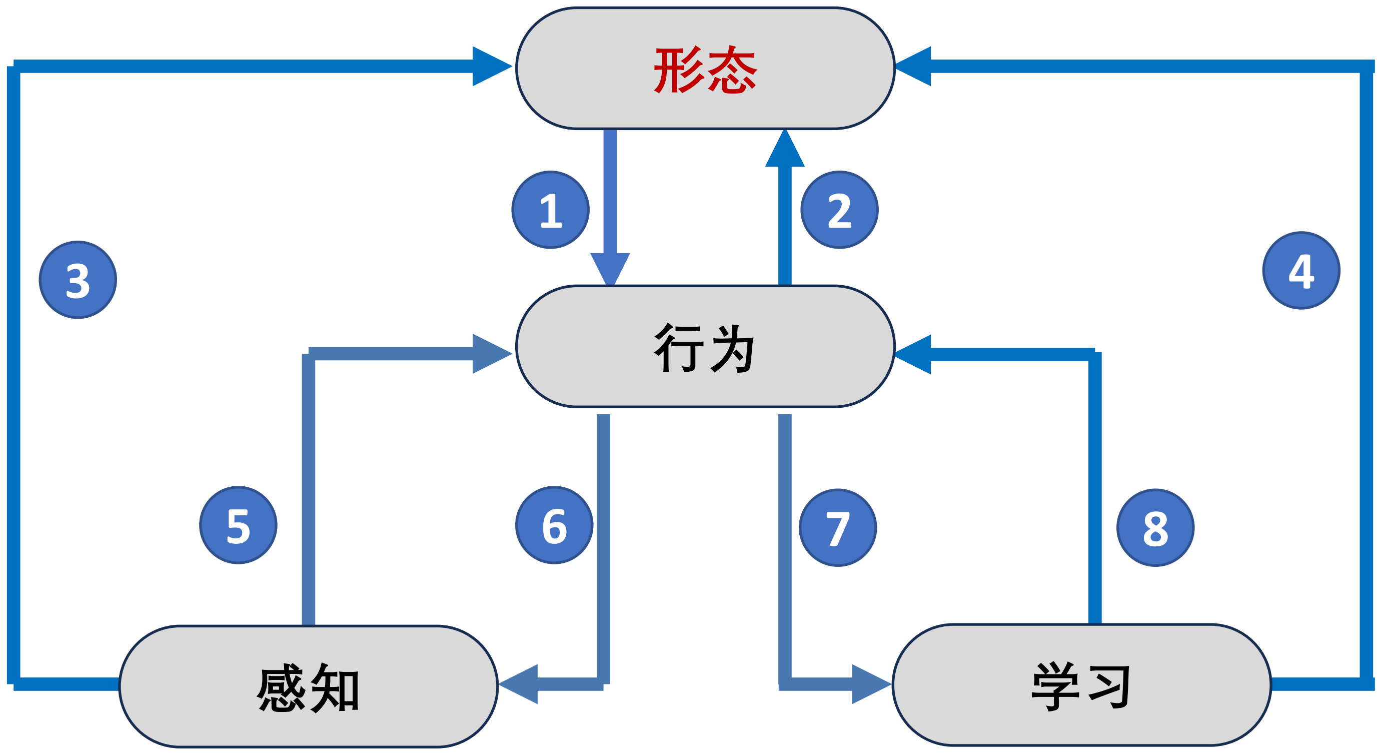
➢ 小结





$$\pi_{\theta}(a_t \mid s_t) \Rightarrow \boxed{\text{A2C, A3C, SAC PPO}} \Rightarrow r = ?$$

- 模仿学习、层次化强化学习、离线强化学习，… …


形态
行为
感知
学习
1
2
3
4
5
6
7
8

形态

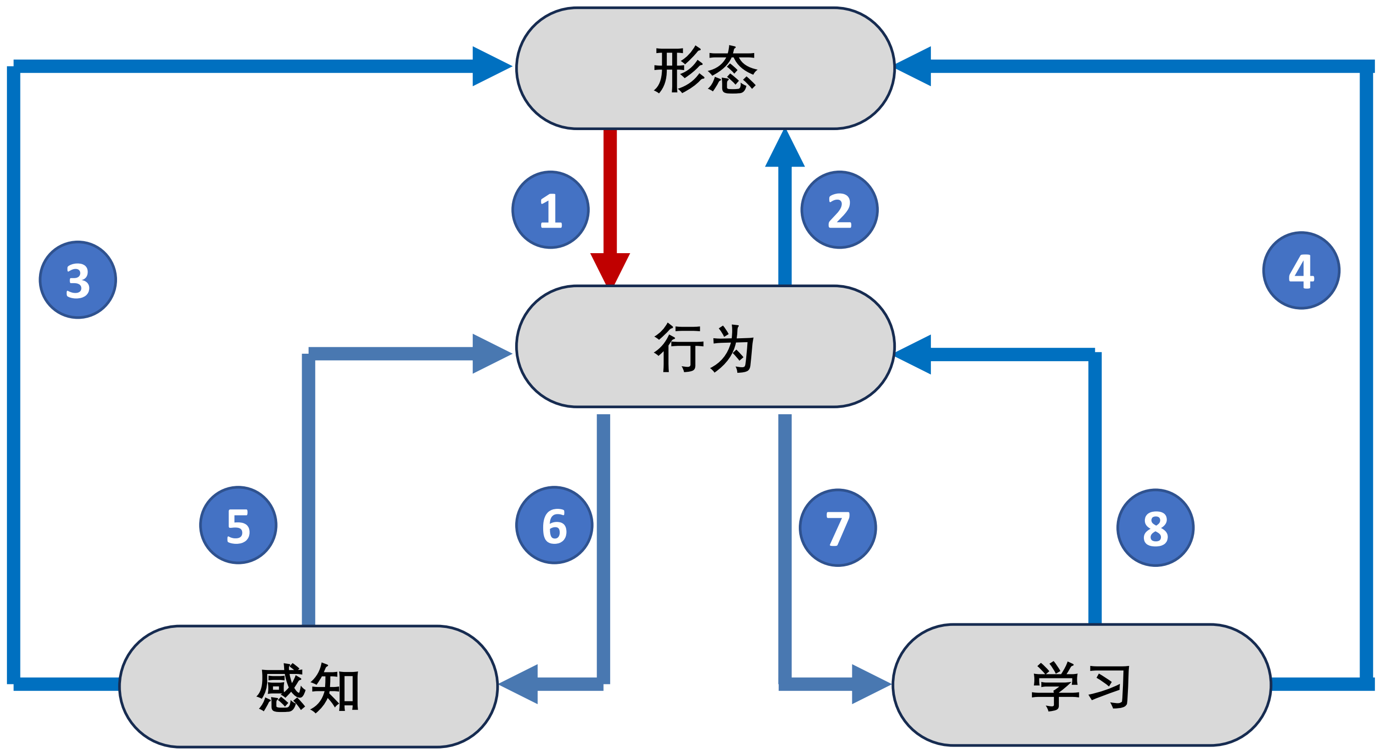
- 例子

Strandbeest（风力仿生兽）

形态

- 例子

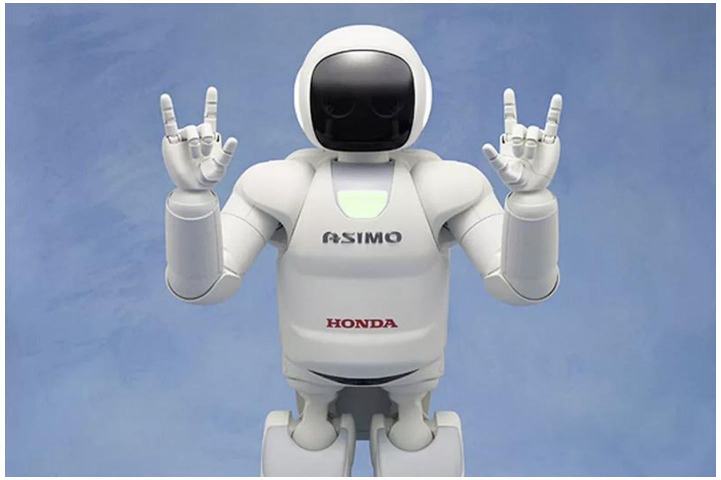
1 形态➝行为



1 形态➞行为

➢ 边界

由于形态计算方面的研究与仿生机器人的研究关系非常密切，二者之间的关系会引起混淆。事实上，形态计算更关心的是利用形态来产生行为，而并非从形态上简单地模仿某些生物。很多仿生型机器人，只是形体上模仿了动物的身体，这种模仿可能能获得一些自由度上的突破，例如腿式机器人相比轮式能爬楼，但在行为控制方面，并没有充分利用形态自身的优势，而且仍需要设计复杂的控制器。本田的Asimo机器人，尽管已经充分接近人类的外形，但其每一个关节都需要用特定的算法来控制。这些情形都不属于形态计算之列。

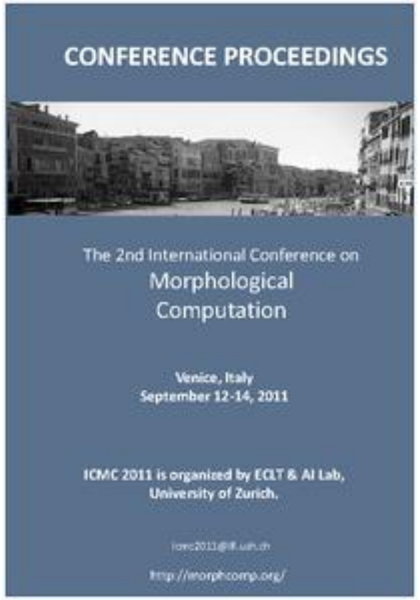


1 形态➔行为

➢ 边界

In 2007, a workshop at the first International Conference on Morphological Computing in Venice, Italy, led by Norman Packard, informally defined morphological computing as any process that

- (a) serves for a computational purpose
- (b) has clearly assignable input and output states
- (c) is programmable, where ‘programmable’ is understood in the broad sense that a programmer can vary the behavior of the system by varying a set of parameters.”

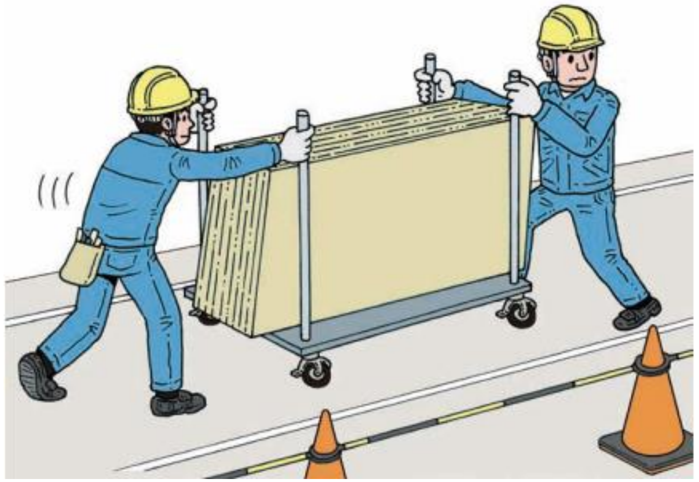


This definition implies that conventional digital computing systems are instances of morphological computing. However, many researchers require that at least part of the computation should rely in a significant way on the physical dynamics of the system implementing the computation, which should be entirely defined by the “program.” By this definition, the execution of a conventional computer program that completely specifies the transformation of an input to an output cannot be classed as morphological computation.

1 形态→行为

➢ 边界

Moving from robotics to physiology, we can use similar concepts to interpret the behavior of natural systems. For example, consider the way one changes one's posture when carrying a heavy load. Though we have no strict proof, we interpret the change as a reshaping of the body's attractor landscape. The human body is a dynamical system that is optimized for walking under normal conditions. A heavy load changes the properties of the system. To readjust we alter the positions of some of our joints and change the tension on some of our muscles. This seems to be a form of morphological control.



1 形态➔行为

➢ 边界

Another example of the interplay between physiological dynamics, control by the brain, and technology is downhill skiing. Skiers continuously adjust their posture to the terrain in order to gain stability against perturbations (without losing the ability to steer). At the same time, the choice of the mechanical properties of skis and boots influences the dynamical behavior of the “skier,” considered as an integrated system. We hypothesize that these changes represent forms of morphological control. For example, certain types of skis (carving skis) are generally regarded as easier to use than the skis used by athletes. In other words, they reduce the complexity of the control problem posed by downhill skiing. However, they also offer less precise control. This explains why racers choose
different, harder-to-use models.

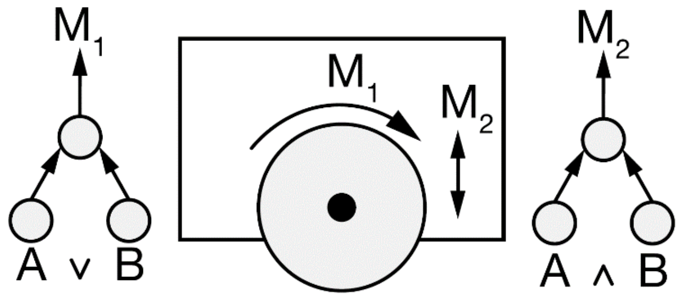
1 形态➞行为

➢ 形态计算

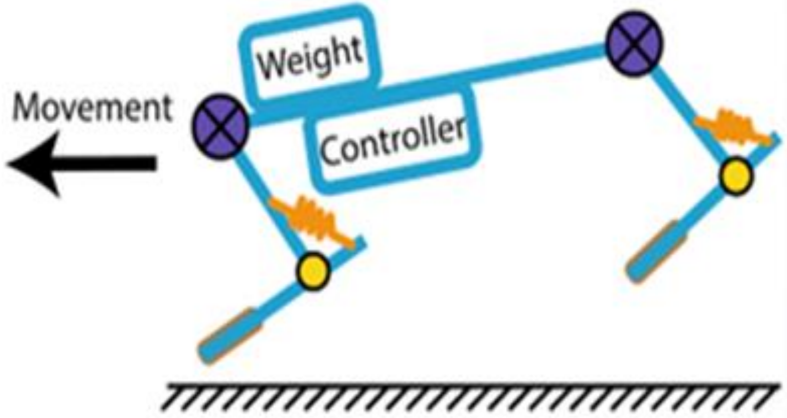
- 用较少的控制！

1 形态→行为

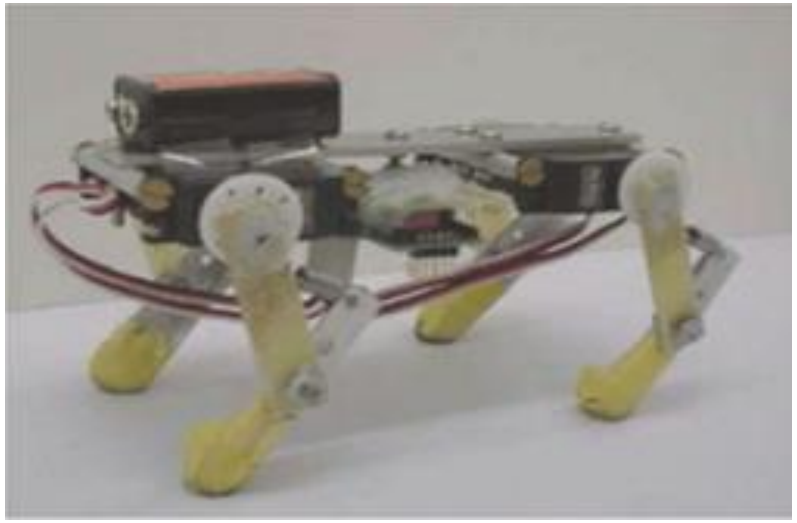
➢ 基本原理：XOR机器人



A	B	M_1	M_2	moving
0	0	0	0	false
0	1	1	0	true
1	0	1	0	true
1	1	1	1	false

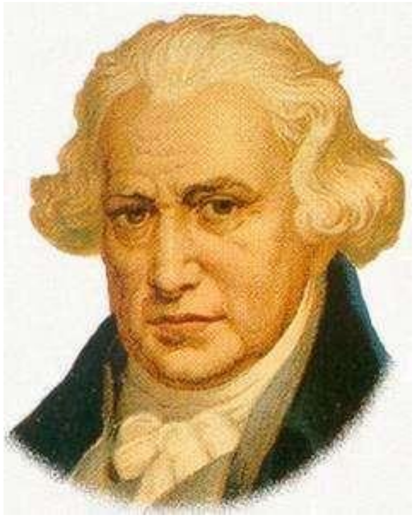


- **M_1 = A OR B**
- **M_2 = A AND B**



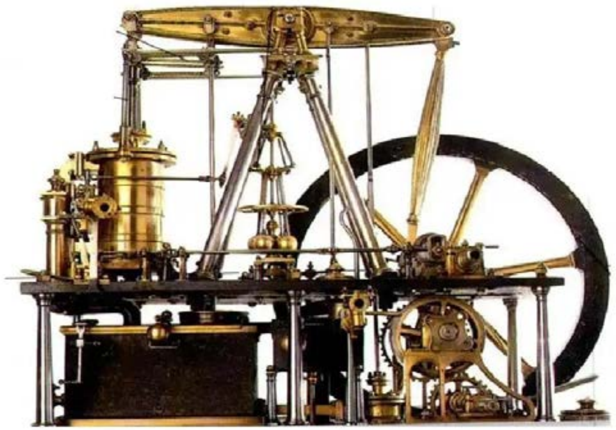
1 形态→行为

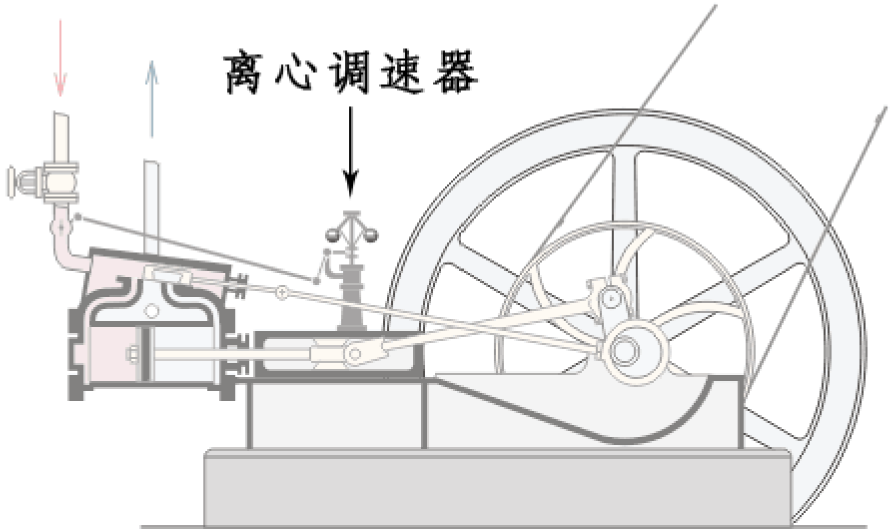
➢ 离心调速器

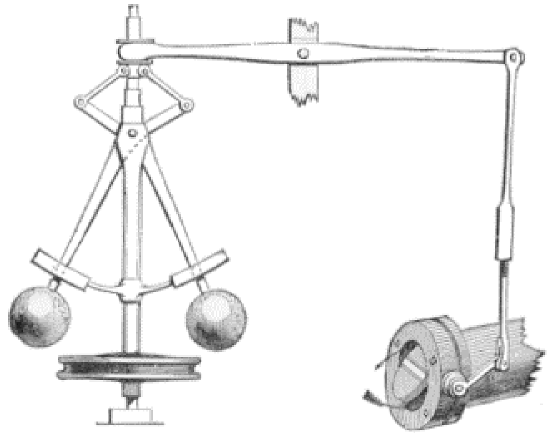


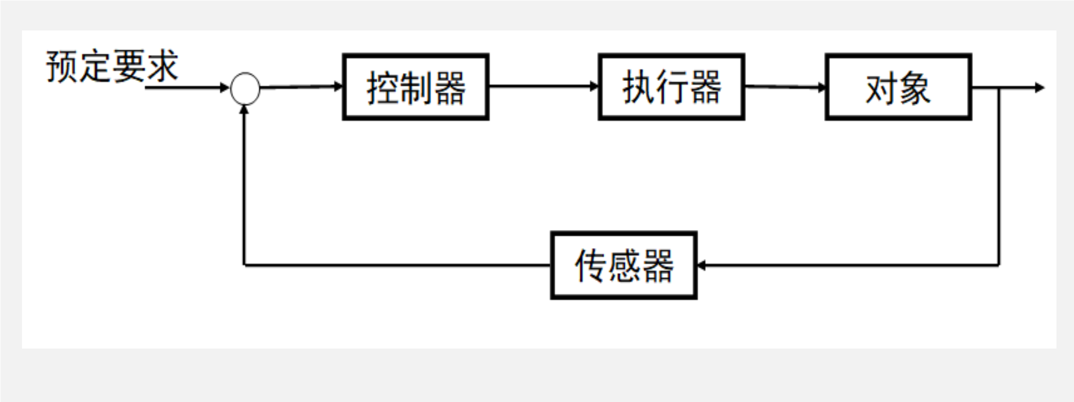
1736-1819

瓦特改良的蒸汽机



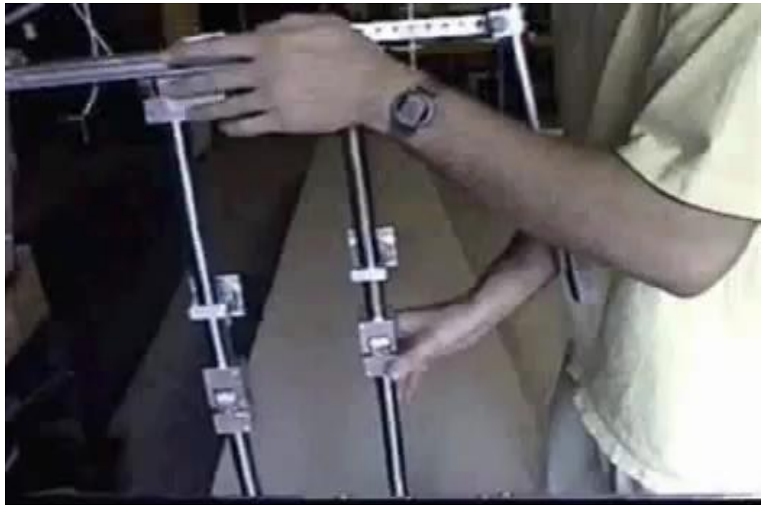


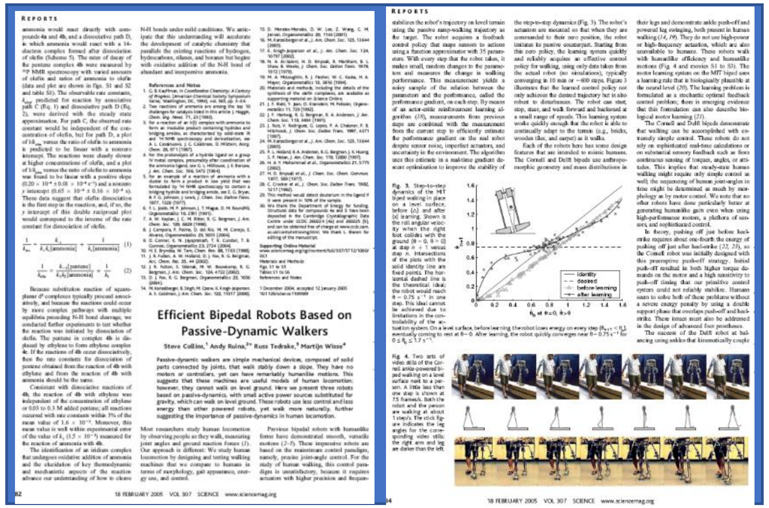




1 形态→行为

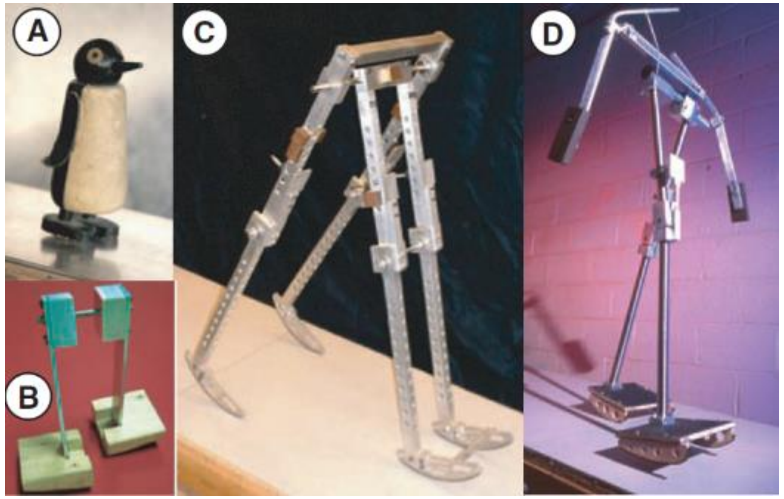
➢ 被动行走机器人（Passive Walking Robot）

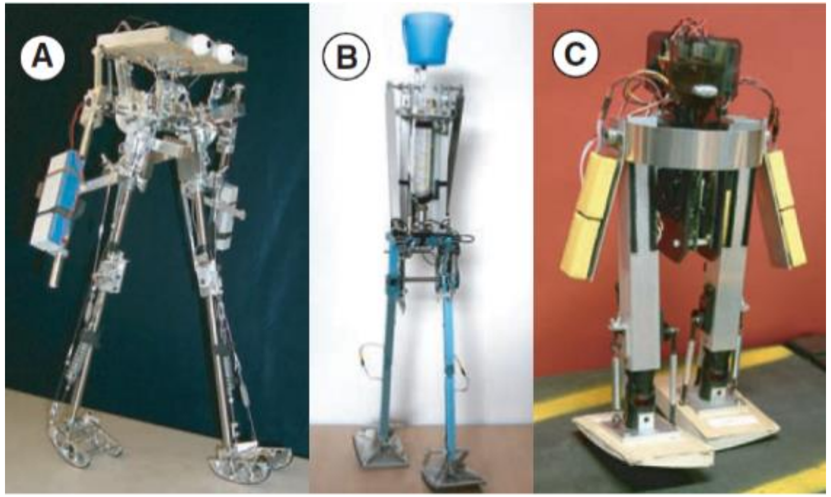



REPORTS

Efficient Bipedal Robots Based on Passive-Dynamic Walkers

Steve Collins, Andy Ruina, Russ Tedrake, Martijn Wisse

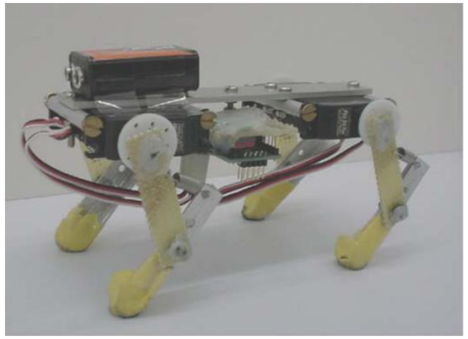


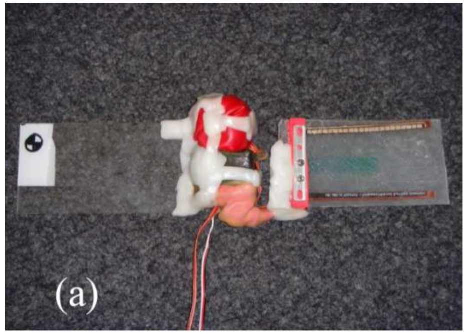


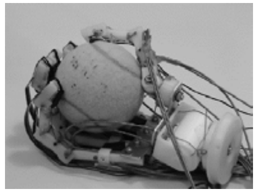
1 形态→行为

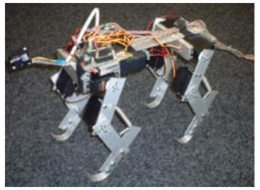
➢ 行为控制

- Pfeifer在Science上撰文，对具身形态的研究工作建立了一个统一框架。并指出，对机器人形态动力学与环境交互的充分利用可以有助于能量的节省，以及自然运动的产生，而且也可以显著降低对控制器与信息处理计算的要求。这也是“形态计算”这一术语的内在核心。







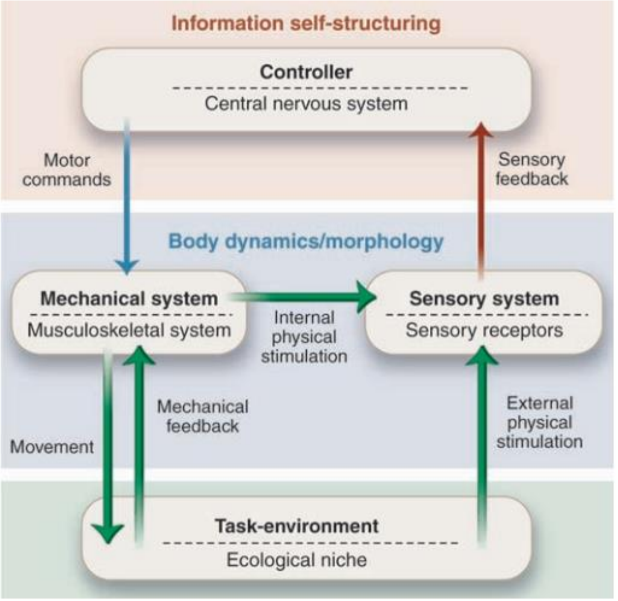


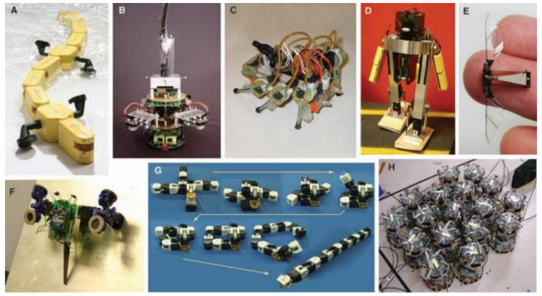
Robotics

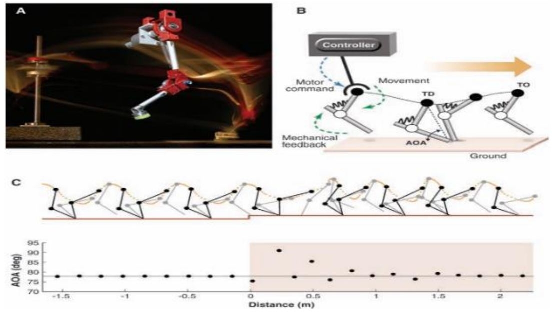
REVIEW

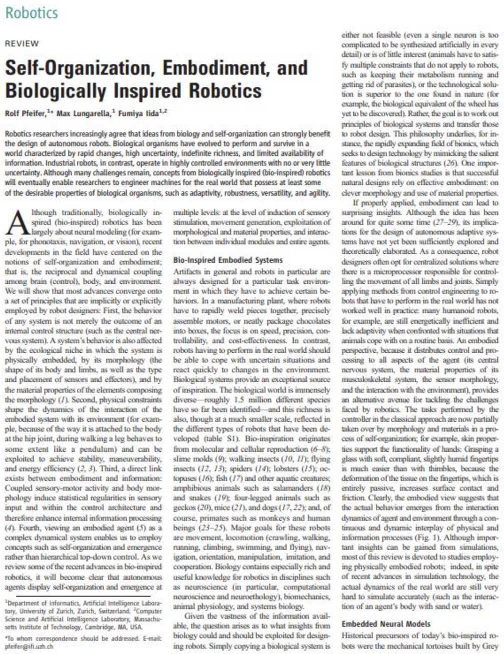
Self-Organization, Embodiment, and Biologically Inspired Robotics

Rolf Pfeifer,[1*] Max Lungarella,[1] Fumiya Iida[1,2]









This paper presents a review of embodied intelligence. where the authors define it as "the computational approach to the design and understanding of intelligent behavior in embodied and situated agents through the consideration of the strict coupling between the agent and its environment, mediated by the constraints of the agent's own body, perceptual and motor system, and brain". Their review is mainly based on a proposed architecture that links morphology, action, perception and learning shown in the right figure of Figure 2. Embodied intelligence is indeed an important concept a fresh review can be useful for the relevant communities. Nevertheless, the reviewer really cannot accept the paper based on the current state with the following comments:

(1) The core of the review presented in this paper is based on the proposed architecture shown in in the right figure of Figure 2. However, embodied intelligence is not a new concept and there are already fundamental frameworks explaining it, and they are widely accepted by the community. A classical one is of course the architecture/framework explained in a well cited work in [1], which is further refined in [2]. Please explain what are the differences between the proposed architecture and the one explained in [1][2], and why the authors find it necessary to expand the framework.

[1] Rolf Pfeifer et al., "Self-organization, embodiment, and biologically inspired robotics", Science 318(5853): 1088-1093, 2007.

[2] Rolf Pfeifer et al., "Cognition from the bottom up: on biological inspiration, body morphology and soft materials", Trends Cogn Sci, 18(8):404-13, 2014.

[3] Daniel E. Koditschek., "What is robotics? Why do we need it and how can we get it", Annu. Rev. Control Robot. Auton. Syst. 2021.

[4] Keyan Ghazi-Zahedi et al., "Evaluating morphological computation in muscle and DC-motor driven models of hopping movements", Front. Robot. AI, vol. 3, 2016.

[5] Shiv A. Katiyar et al., "Energy harvesting for robots with adaptive morphology", Soft Robotics 10(2): 365-379, 2023.

[6] Stephen Coyle et al., "Bio-inspired soft robotics: Material selection, actuation, and design", Extreme Mechanics Letters 22:51-59, 2018.

Reply: Thank you very much for your insightful comments. Indeed, we totally agree that the embodied intelligence framework proposed in [1-2] is a very classic and

Embodied Intelligence: A Synergy of Morphology, Action, Perception and Learning

HUAPING LIU, Tsinghua University, Beijing, China
DI GUO, Beijing University of Posts and Telecommunications, Beijing, China
ANGELO CANGELOSI, The University of Manchester, Manchester, United Kingdom of Great Britain and Northern Ireland

Embodied intelligence emphasizes that the intelligence is affected by the tight coupling of brain, body, and environment. It is continuously and dynamically generated through the process of information perception and physical interaction with the environment. During the past years, the research scope of embodied intelligence has also been expanding and has attracted great attention from various communities. At the same time, a huge number of works relevant to embodied intelligence have been proposed, especially in recent years. In this article, we present a comprehensive survey of embodied intelligence from the perspective that it is a synergy of morphology, action, perception, and learning, providing a thorough summary and categorization of existing studies. Specifically, as embodied intelligence is a synergy of all these components rather than themselves alone, we mainly focus on the connections across these four components (morphology, action, perception, and learning) and identify areas where future research can benefit from their intrinsic connections.

CCS Concepts: • **Computer systems organization → Robotic autonomy;**

Additional Key Words and Phrases: Embodied intelligence, morphology, action, perception, learning

ACM Reference Format:
Huaping Liu, Di Guo, and Angelo Cangelosi. 2025. Embodied Intelligence: A Synergy of Morphology, Action, Perception and Learning. *ACM Comput. Surv.* 57, 7, Article 186 (March 2025), 36 pages. https://doi.org/10.1145/3717059

1 Introduction

Embodied intelligence is the computational approach to the design and understanding of intelligent behavior in embodied and situated agents through the consideration of the strict coupling between the agent and its environment, mediated by the constraints of the agent's own body, perceptual and motor system, and brain [23]. Embodied intelligence emphasizes that the intelligence is affected by the tight coupling of brain, body, and environment. It is continuously and dynamically generated

This work was supported by the National Natural Science Fund for Key International Collaboration under grant 62120106005 and the National Natural Science Foundation of China under grant 62273054. Cangelosi's work was in part supported by the EPSRC CRADLE project (EP/X02489X/1), the ERC Advanced project "eTALK" (Selected by the Horizon ERC, funded by UKRI) and the US AFOSR/EOARD project CASPER++.

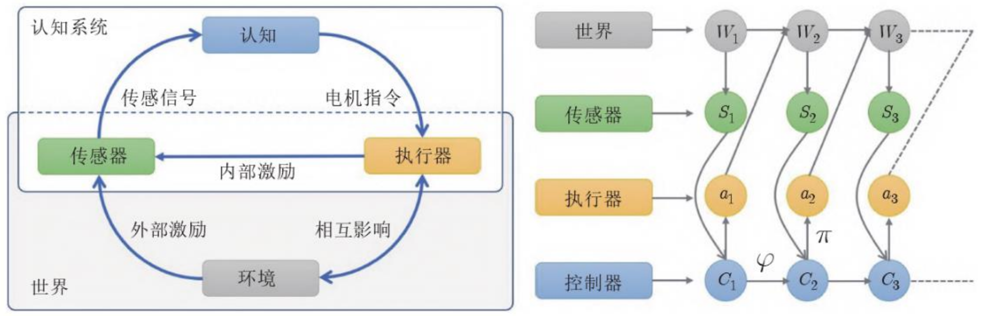
Authors' Contact Information: Huaping Liu, Tsinghua University, Beijing, China; e-mail: hpliu@tsinghua.edu.cn; Di Guo, Beijing University of Posts and Telecommunications, Beijing, Beijing, China; e-mail: guodi.gd@gmail.com; Angelo Cangelosi, The University of Manchester, Manchester, United Kingdom of Great Britain and Northern Ireland; e-mail: angelo.cangelosi@manchester.ac.uk.

ACM 0360-0300/2025/03-ART186
https://doi.org/10.1145/3717059

1 形态→行为

➢ 理论模型

信息论模型



$$\beta(s \mid w) : W \to \Delta_S$$

$$\varphi(c' \mid s, c) : S \times C \to \Delta_C$$

$$\pi(a \mid c) : C \to \Delta_A$$

$$\alpha(w' \mid w, a) : W \times A \to \Delta_W$$

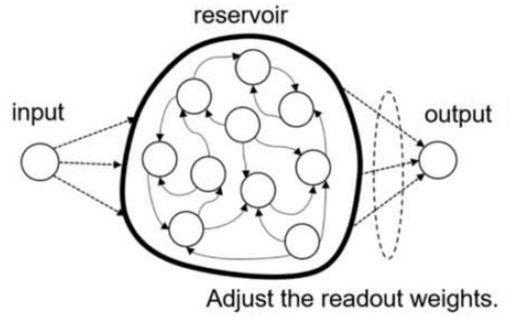
$$\pi(a \mid s) : S \to \Delta_A$$

$$KL(\alpha \,||\, \tilde{\alpha}) = \sum_{w', w, a} p(w', w, a) \log_2 \frac{\alpha(w' \mid w, a)}{\tilde{\alpha}(w' \mid a)}$$

$$\max_{\pi(a|s)} \ \mathbb{E}\big(Q(s,a)\big) - \text{Complexity}$$

通过抑制控制器的复杂度来强迫形态发挥作用

储备池模型

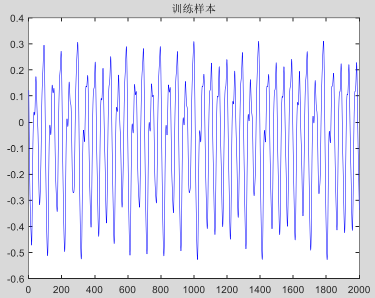


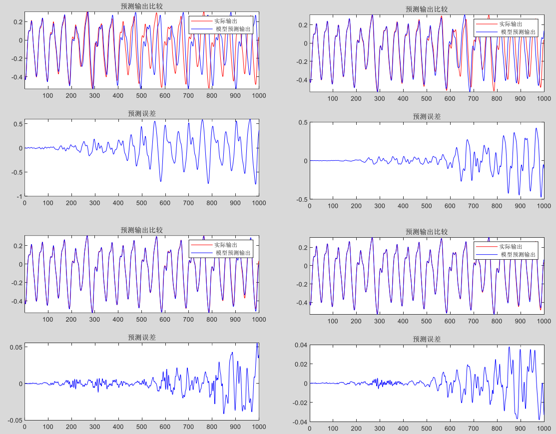
$$h_t = \sigma(W_x \cdot x_t + W_h \cdot h_{t-1})$$

$$y_t = W_y \cdot h_t$$

$$W_y = YH^T(HH^T + \lambda I)^{-1}$$

$$\dot{y}(t) = -0.1y(t) + \frac{0.2y(t-16.8)}{1 + y^{10}(t-16.8)}$$

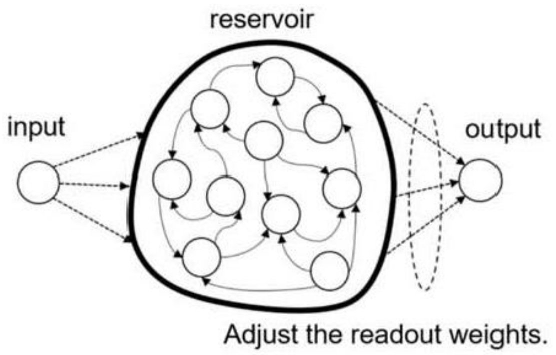


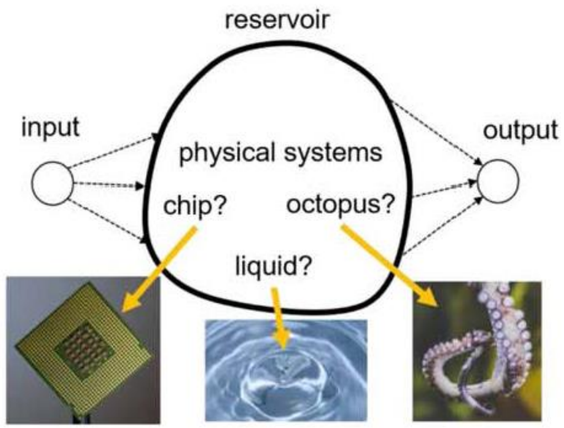


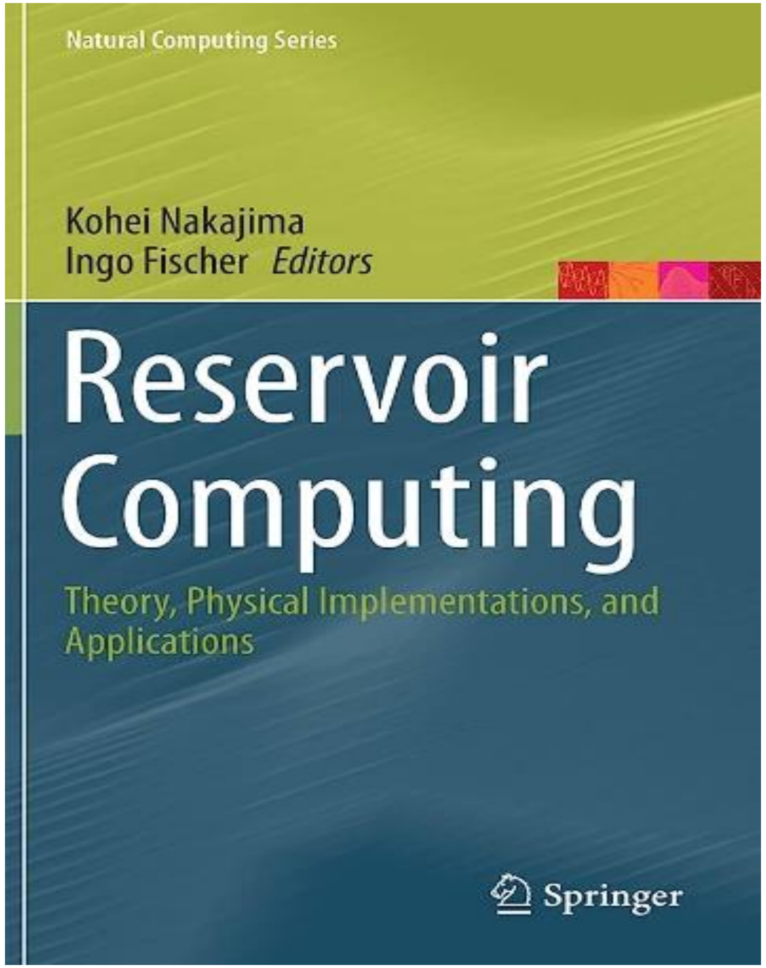
利用具身智能体的身体动力学来实现储备池

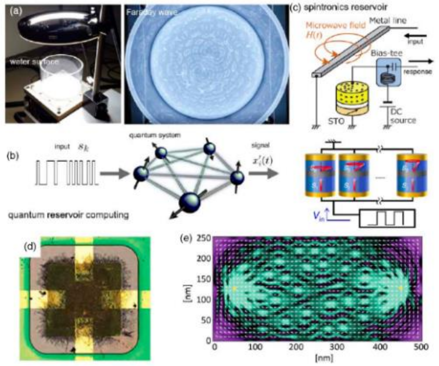
1 形态➞行为

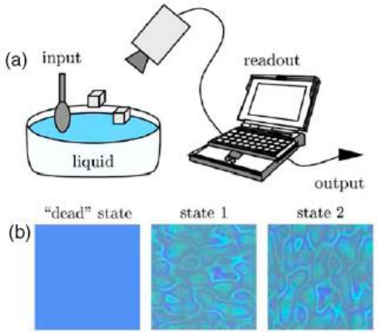
➢ 物理储备池模型

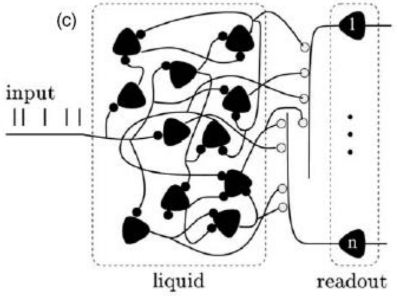






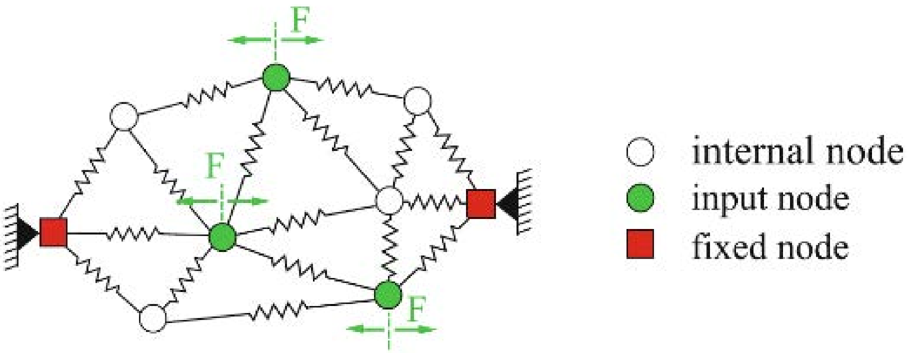






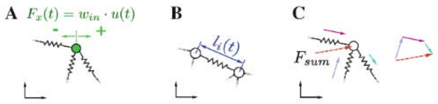
1 形态→行为

➢ 物理储备池模型

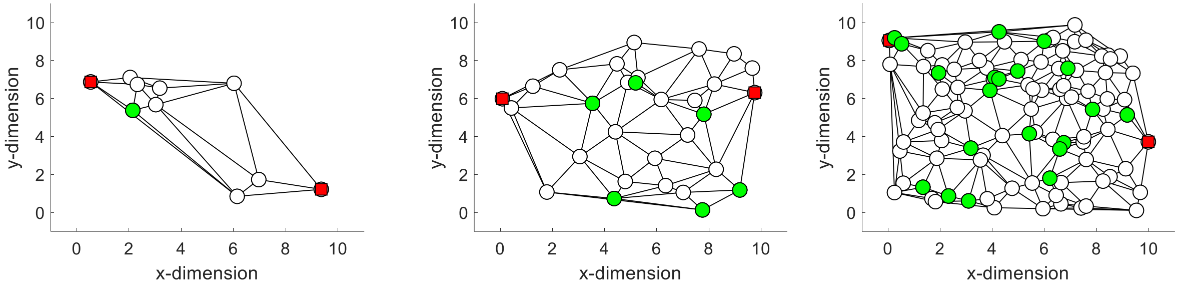


$$y_t = 0.3y_{t-1} + 0.05y_{t-1} \cdot \sum_{t'=t-10}^{t-1} y_{t'} + 1.5u_{t-10}u_{t-1} + 0.1$$

$$u_t = 0.1\sin(4.22\pi t)\cdot\sin(7.46\pi t)\cdot\sin(8.66\pi t)$$

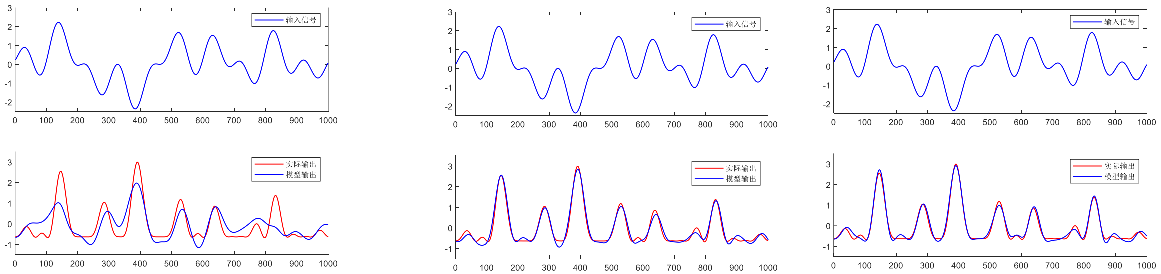


$$\begin{cases} \dot{p}_i(t) = v_i(t) \\ \dot{v}_i(t) = -\kappa(p_i(t)) - \upsilon(v_i(t)) + u_i(t) \end{cases}$$



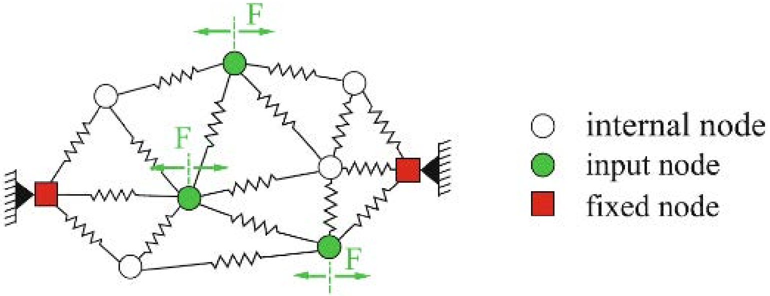
$$\kappa(p) = k_3 p^3 + k_1 p \qquad \upsilon(v) = d_3 v^3 + d_1 v$$

$$y(t) = \sum_{i=1}^{L} w_i l_i(t)$$

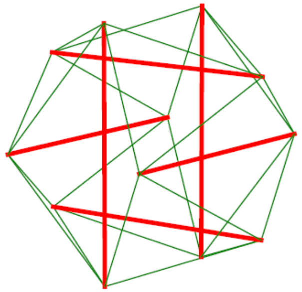


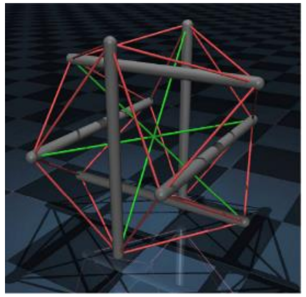
1 形态➞行为

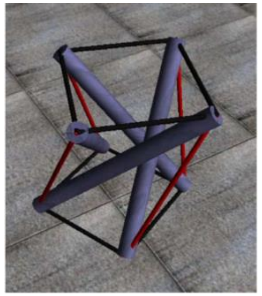
➢ 物理储备池模型：张拉整体（ Tensegrity ）机器人

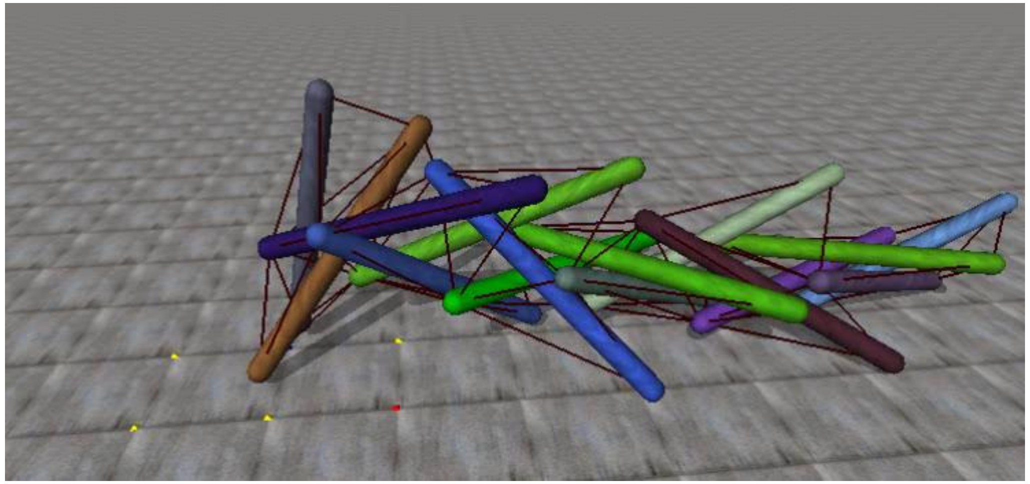


质块改成刚性棍



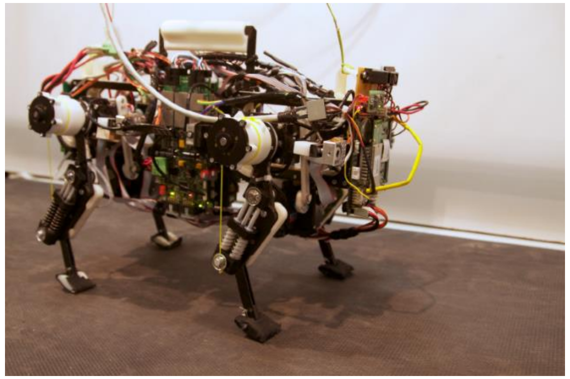


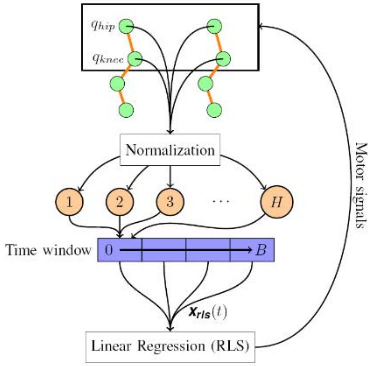




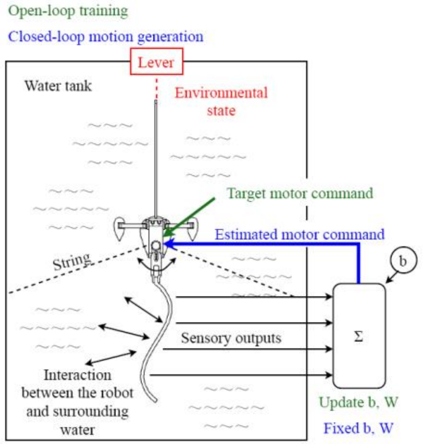
1 形态→行为

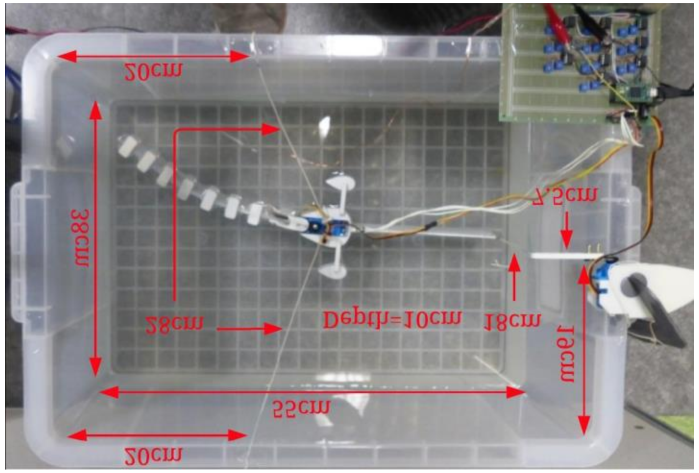
➢ 物理储备池：例子



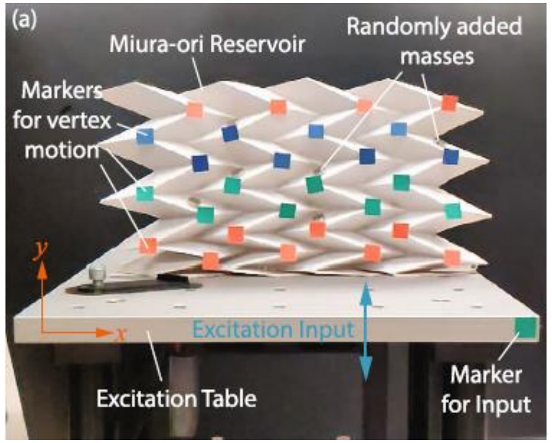


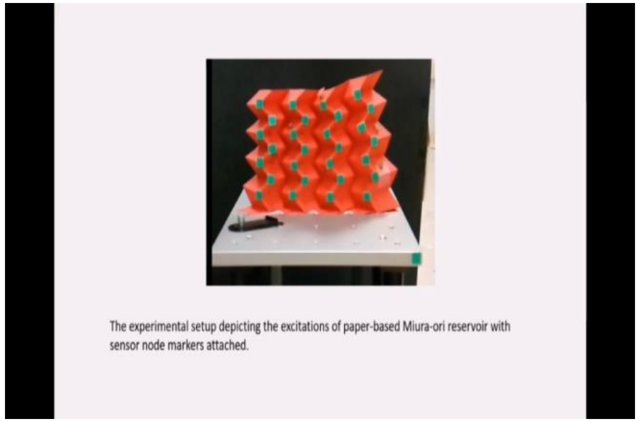
- Degrave J, Caluwaerts K, Dambre J, et al. Developing an embodied gait on a compliant quadrupedal robot, IROS, 2015



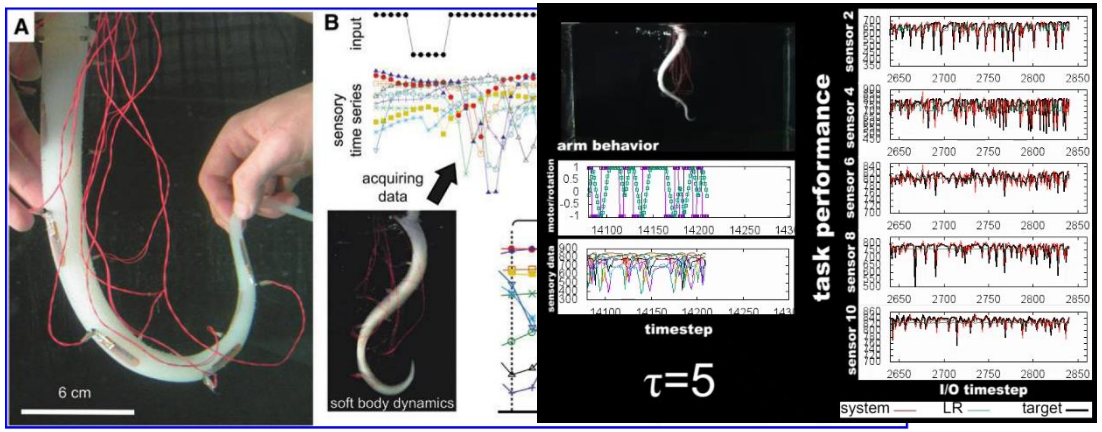


- Horii Y, Inoue K, Nishikawa S, et al. Physical reservoir computing in a soft swimming robot, ALIFE 2021.





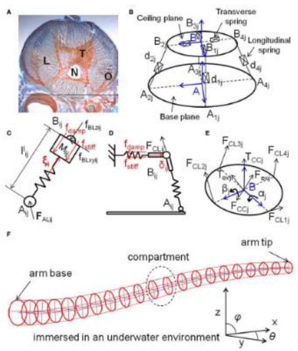
- Bhovad P, Li S. Physical reservoir computing with origami and its application to robotic crawling, Scientific Reports, 2021,

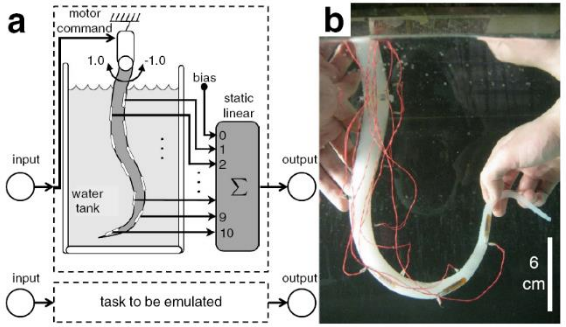


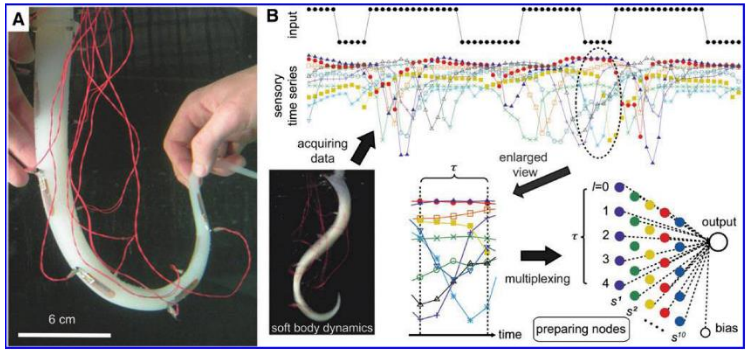
- Nakajima K, Hauser H, Li T, et al. Information processing via physical soft body[J]. Scientific reports, 2015

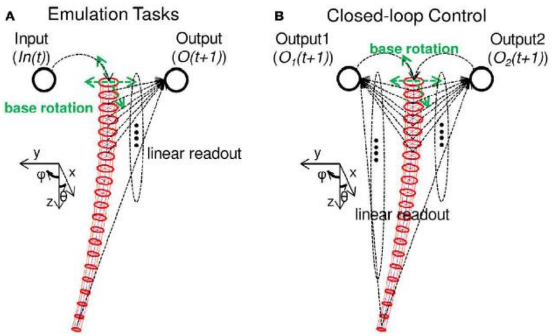
1 形态➞行为

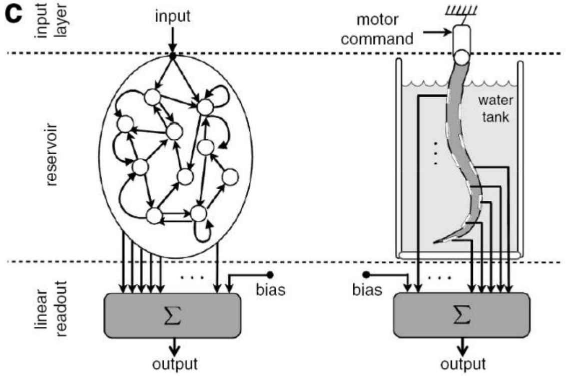
➢ 物理储备池：章鱼软体









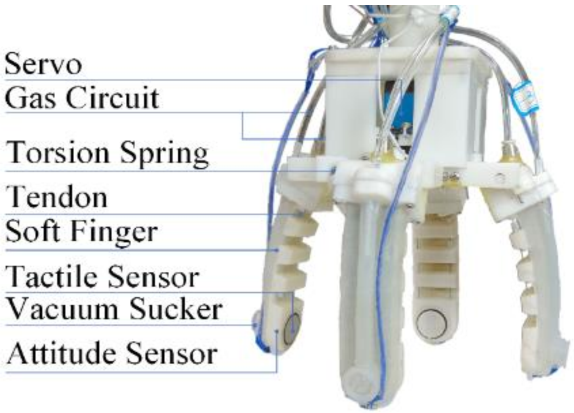


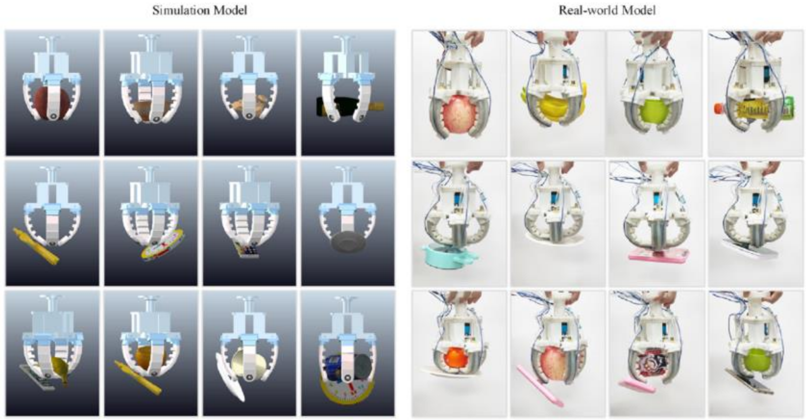
章鱼启发

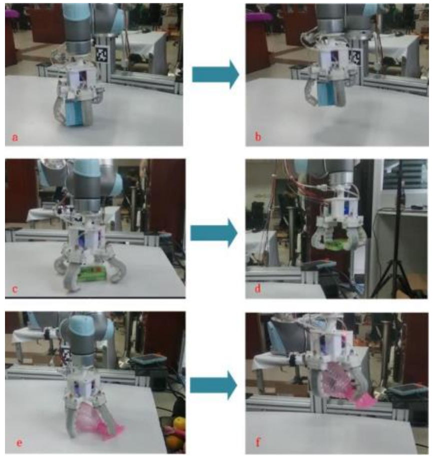
$$O(t+1)=\sum_{i=0}^{10} w_i s_i(t)$$

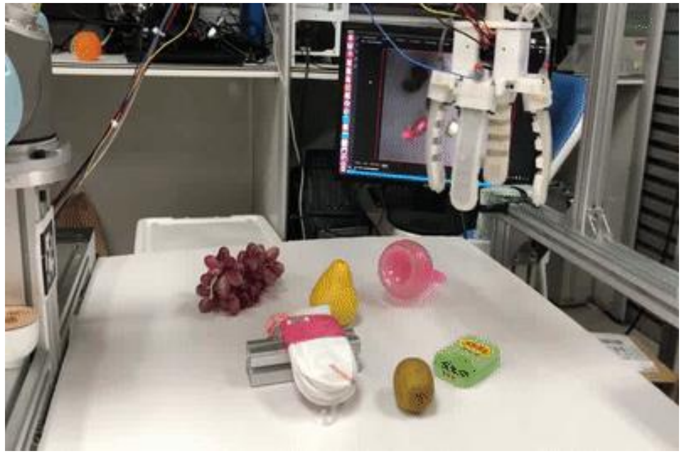
1 形态→行为

➢ 物理储备池：软体机器人





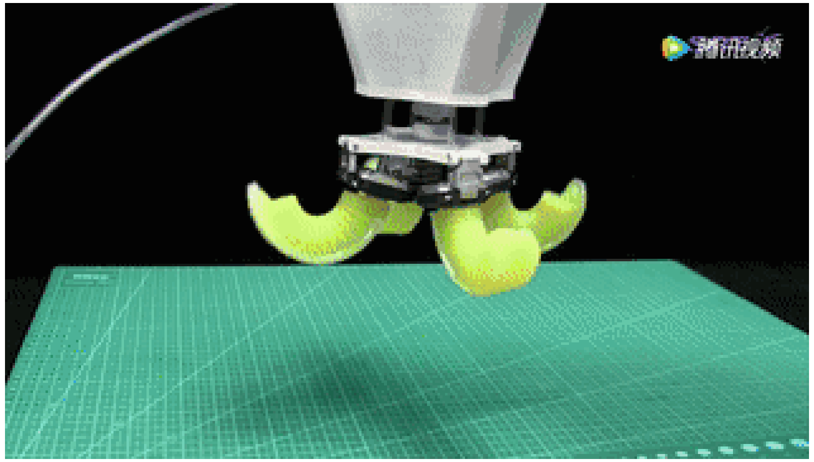


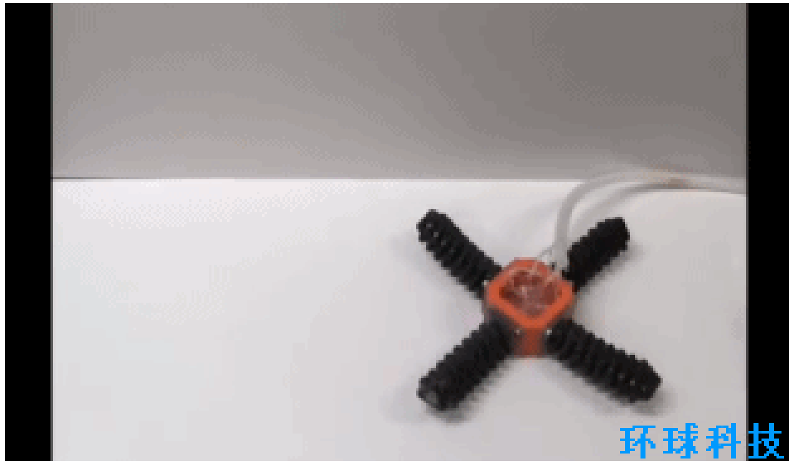


- Multimode grasping soft gripper achieved by layer jamming structure and tendon-driven mechanism. Soft Robotics, 2021
- Hybrid robotic grasping with a soft multimodal gripper and a deep multistage learning scheme, T-RO, 2022

1 形态→行为

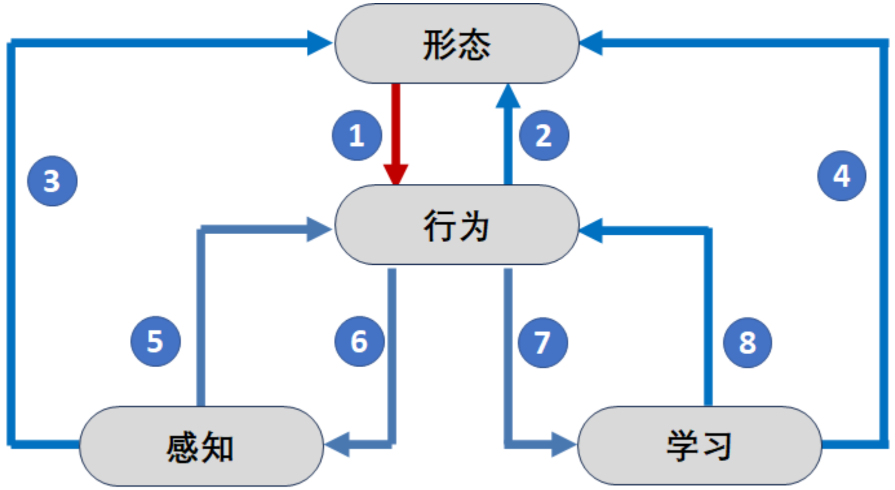
➢ 物理储备池：软体机器人

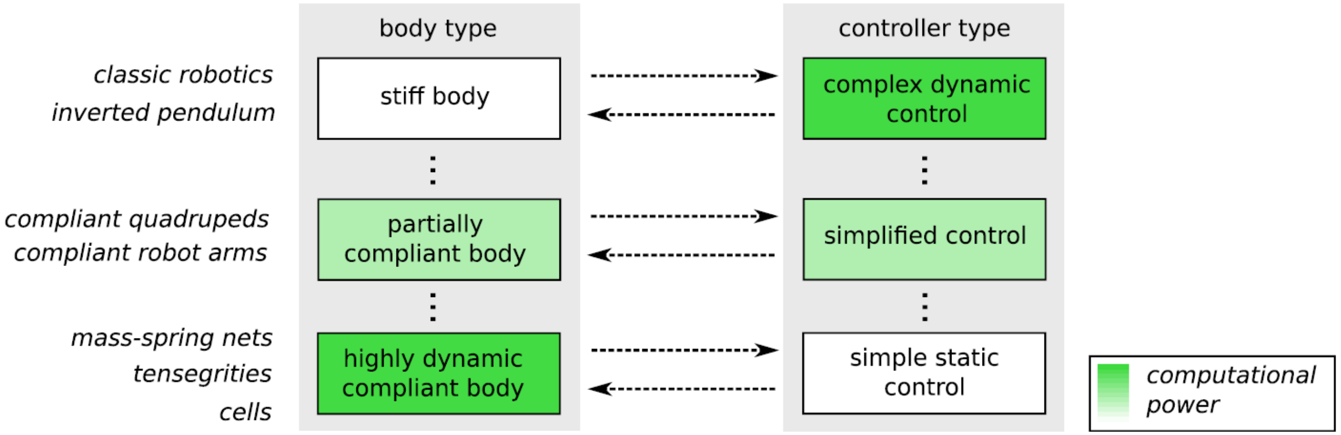


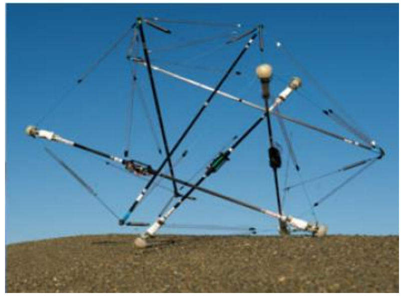


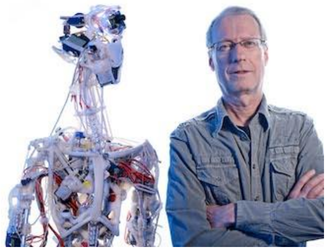
1 形态→行为：形态计算

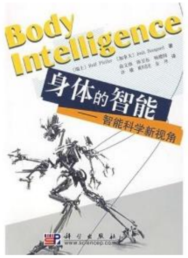
➢ 小结

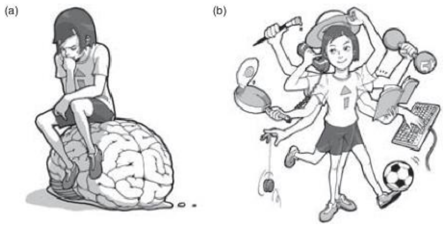




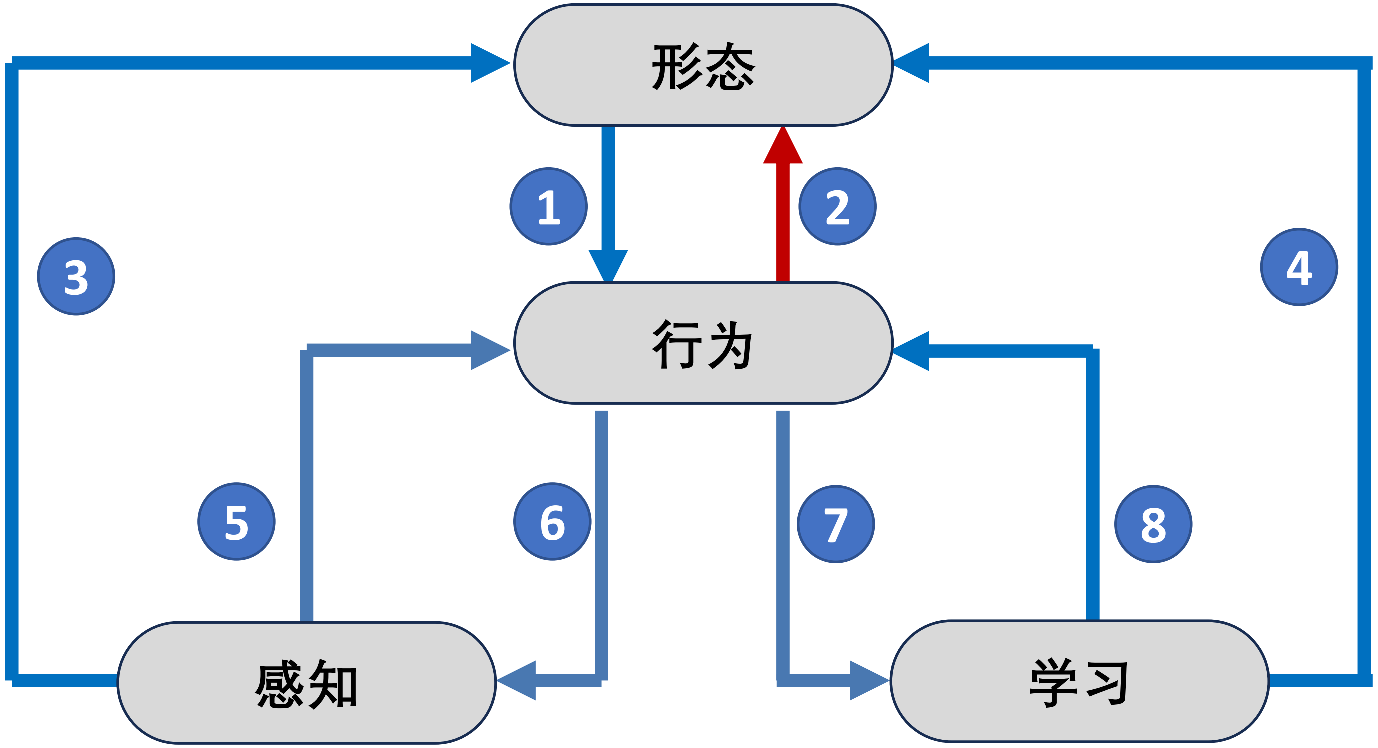








2 行为➞形态：形态控制



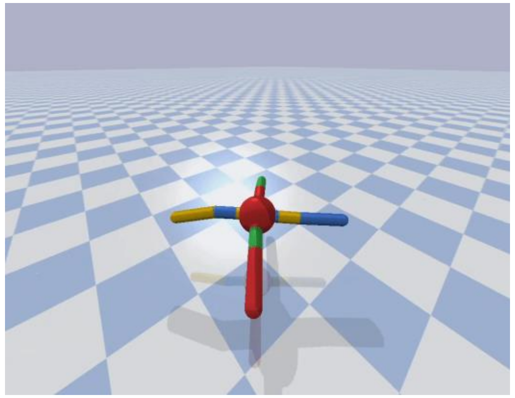
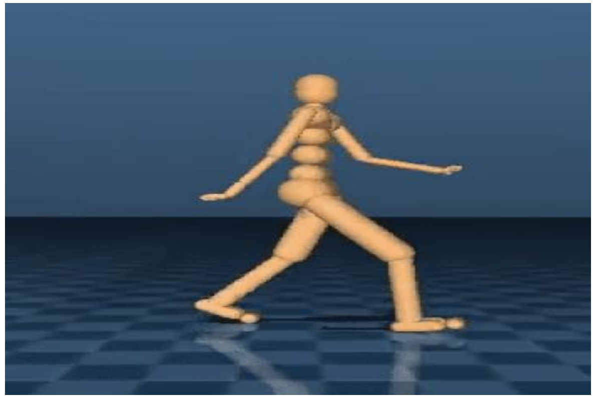
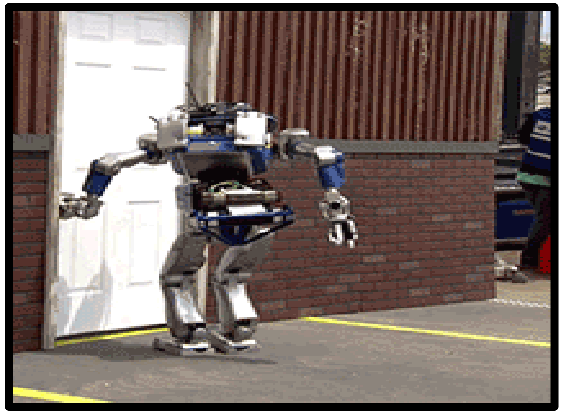
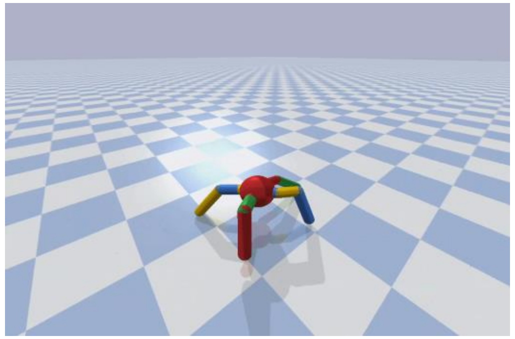
2 行为→形态：形态控制

- 形态约束

2 行为→形态：形态控制

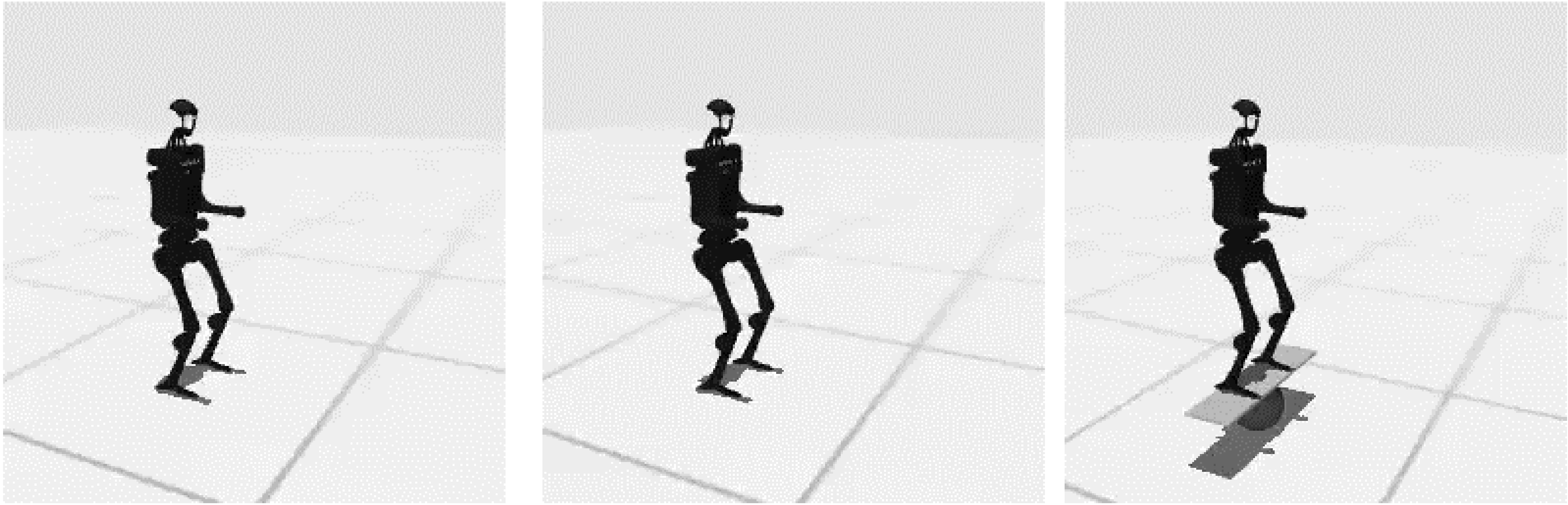
➢ 强化学习的维数灾

- 机器人自由度多，优化学习困难
- 机器人形态复杂，难以利用
- 在不同形态之间的迁移

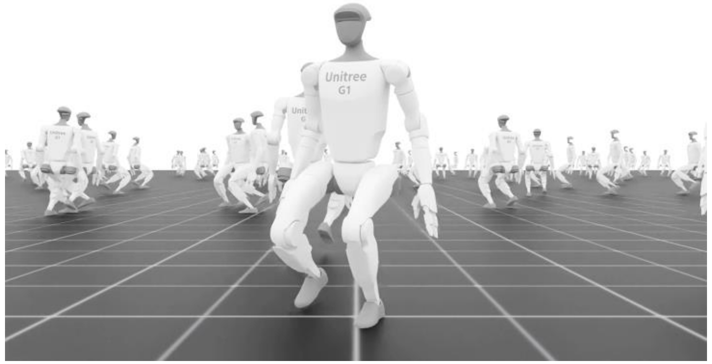
2 行为→形态：形态控制

➢ 强化学习的维数灾



2 行为→形态：形态控制

➢ 强化学习的维数灾



Iteration 0

Iteration 1

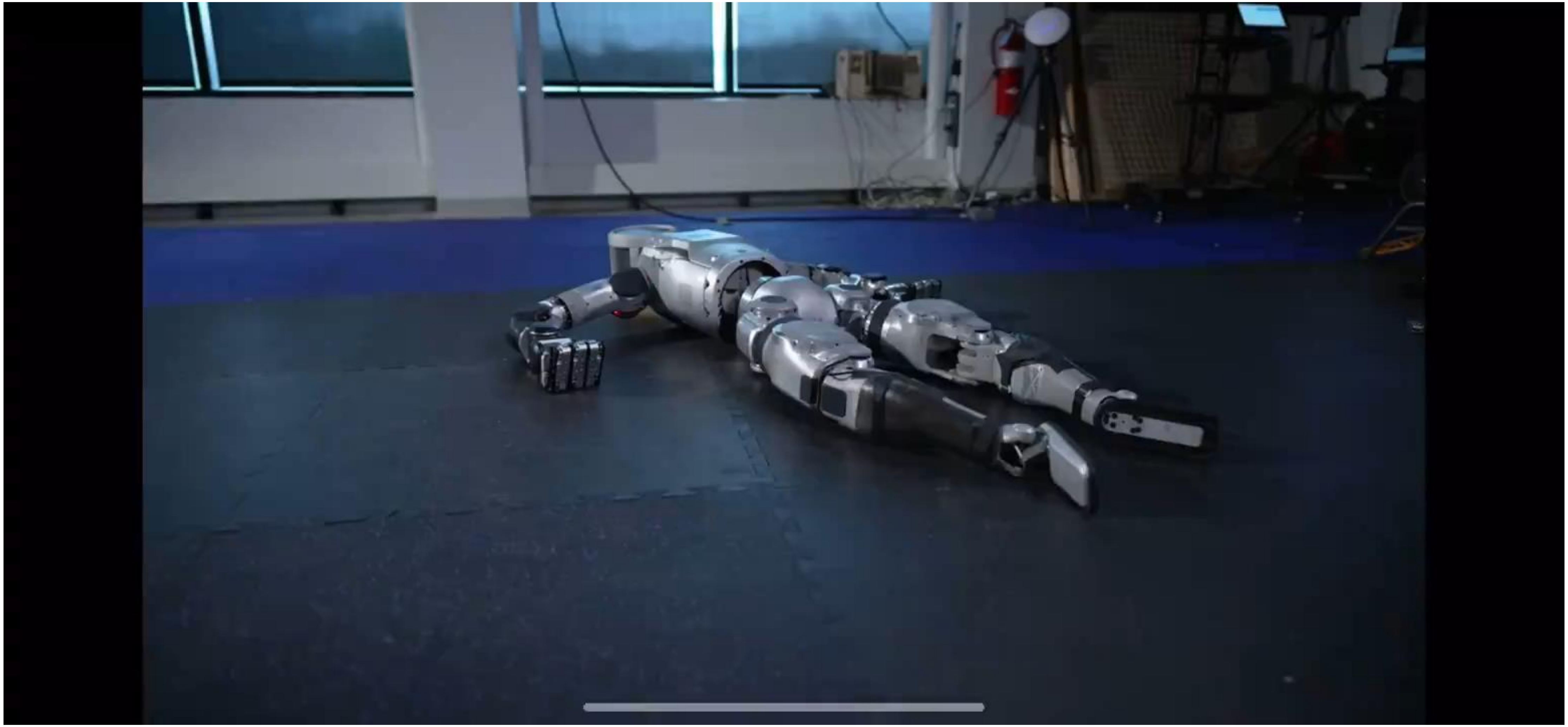
Iteration 2

Iteration 3

Iteration 4

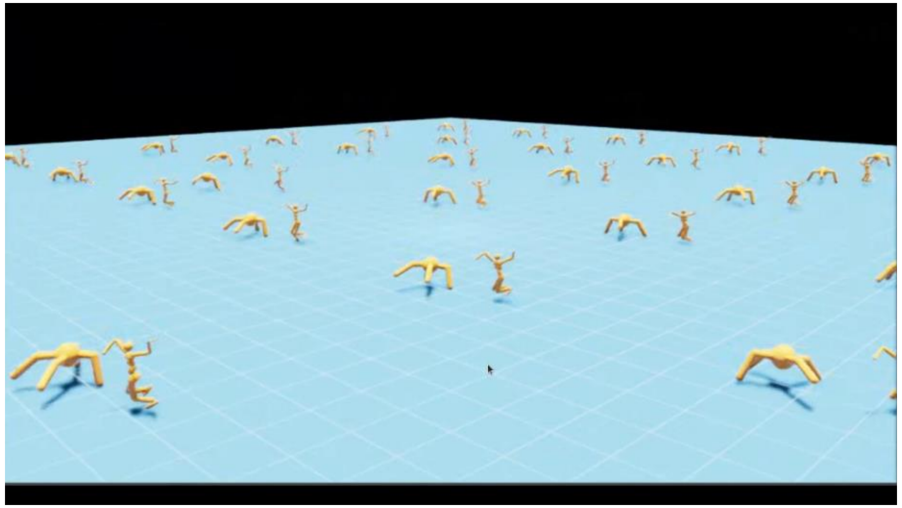
2 行为→形态：形态控制

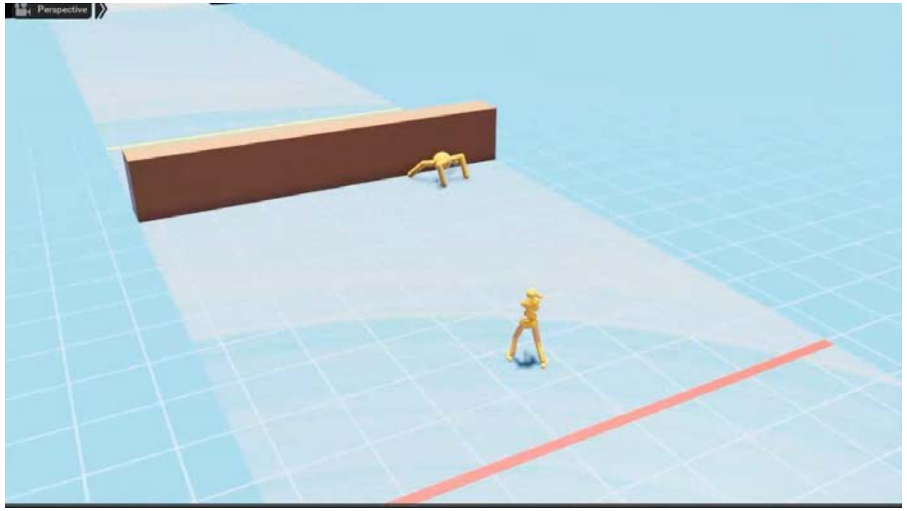
➢ 强化学习的维数灾

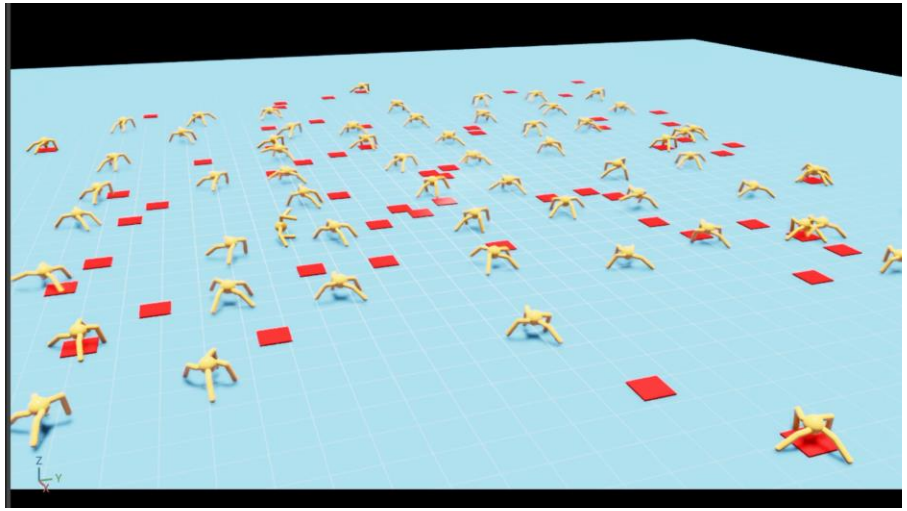


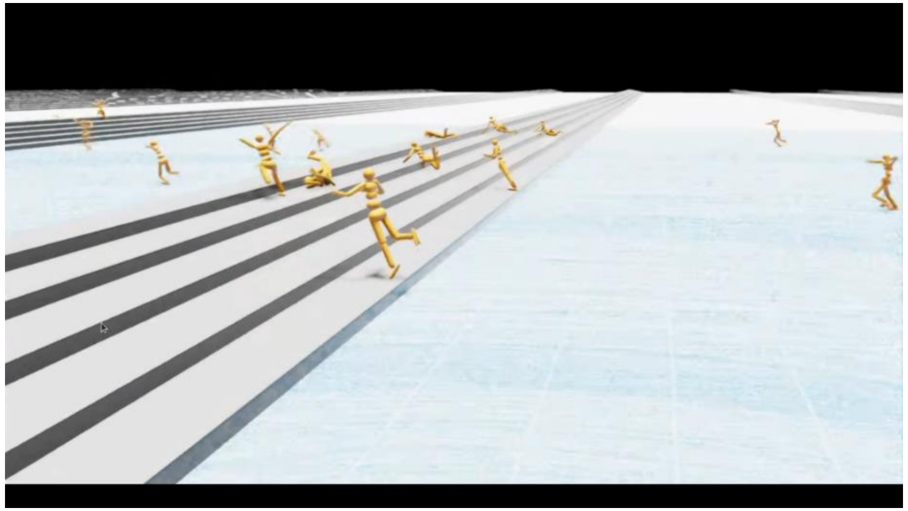
2 行为→形态：形态控制

➢ 强化学习的维数灾



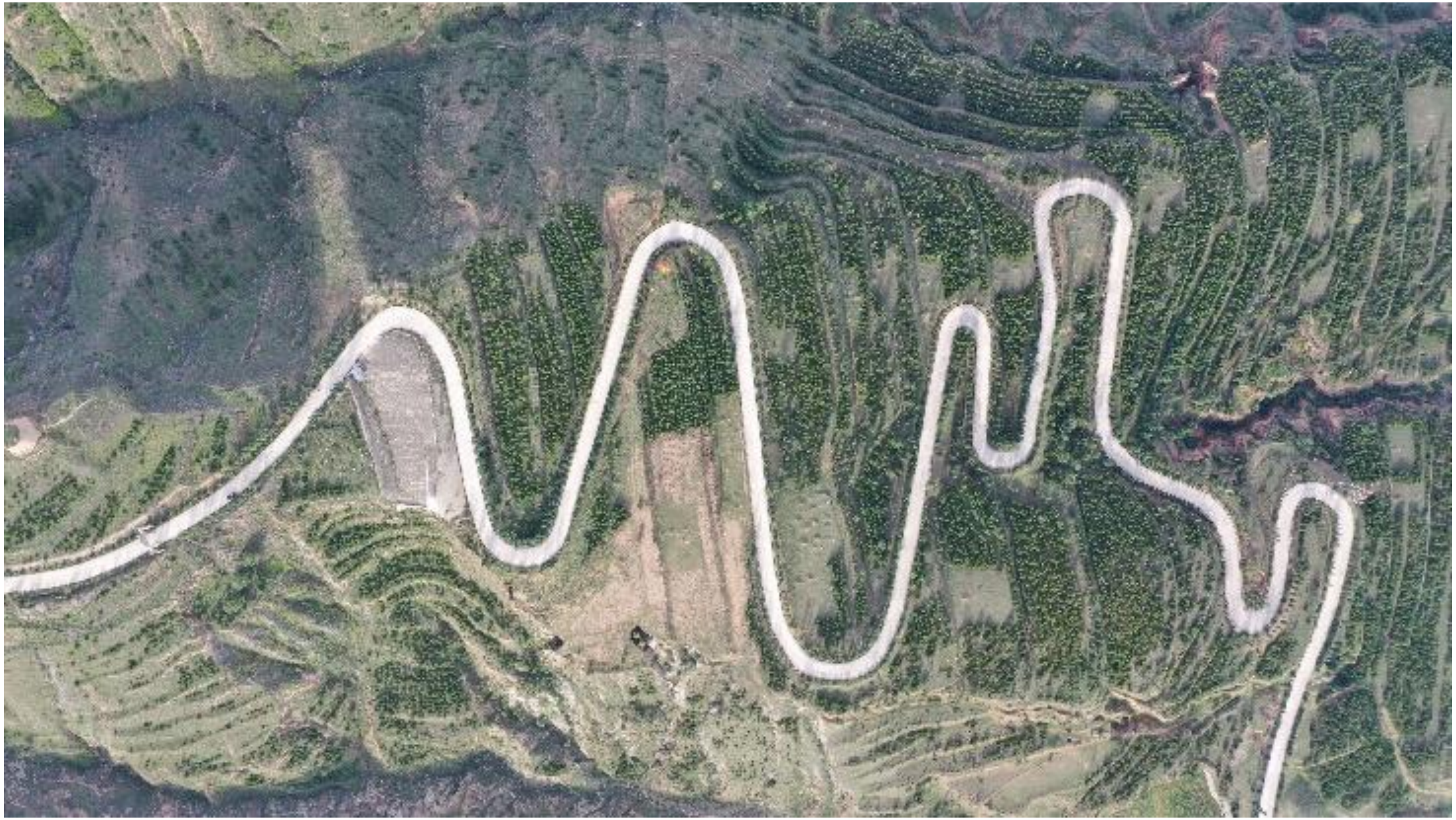






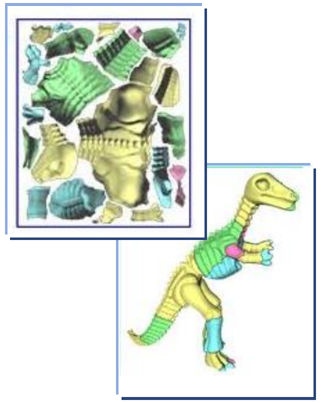
2 行为→形态：形态控制

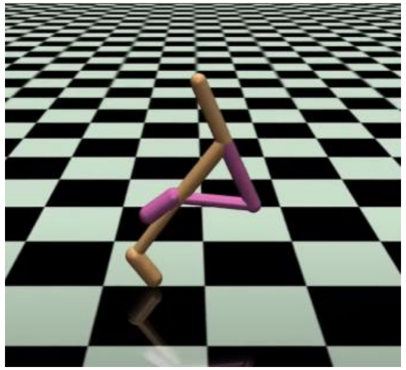
➢ 结构

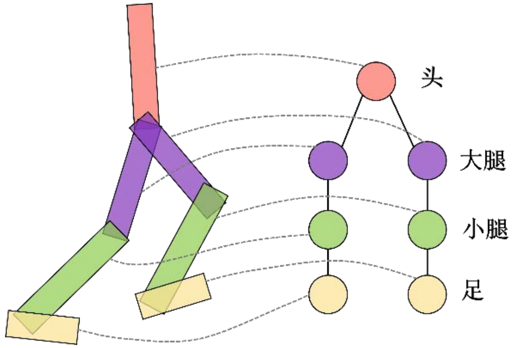


2 行为→形态：形态控制

➢ 引入结构性约束







$$\pi_\theta(a_t \mid s_t)$$

图结构表示

$$D = (V, E, Y)$$

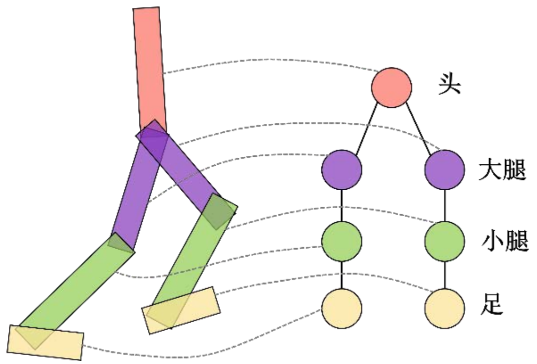
$N_{out}(v)$ 所有以v作为头节点的有向边所指向的尾节点的集合。这一集合反映的是节点v的“输出节点”。

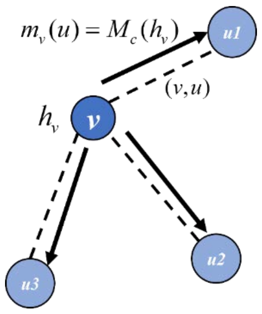
$N_{in}(v)$ 所有以v作为尾节点的边的头结点的集合。这一集合反映的是节点v的“输入节点”。

$$Y_V(v_i) \in \{1, 2, ..., |Y_V|\} \qquad Y_E(\{v_i, v_j\}) \in \{1, 2, ..., |Y_E|\}$$

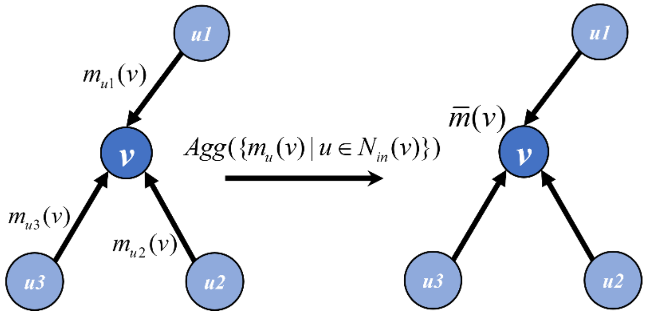
2 行为→形态：形态控制

➢ 形态的表示：图神经网络





$$h_v^{(0)}=\Phi_1(o_v)$$



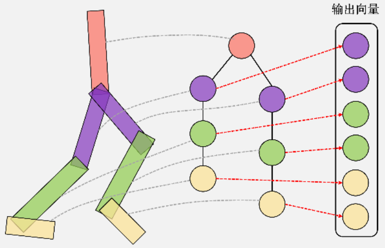
$$s_t=\underset{v\in V}{\mathrm{Concat}}\{h_v^{(0)}\}$$

- 消息的扩散 $\quad m_v^{(k)}(u)=\Phi_2(h_v^{(k)})$
- 消息的汇聚 $\quad \bar{m}_v^{(k)}=\dfrac{1}{|N_{in}(v)|}\sum_{u\in N_{in}(v)} m_u^{(k)}(v)$
- 状态的更新 $\quad h_v^{(k+1)}\leftarrow \Phi_3(h_v^{(k)},\bar{m}_v^{(k)})$

重复 K 次

$$s_t=\underset{v\in V}{\mathrm{Concat}}\{h_v^{(K)}\}$$

输出模型



$$\pi_\theta^{(v)}(a_t^{(v)}\mid s_t)$$

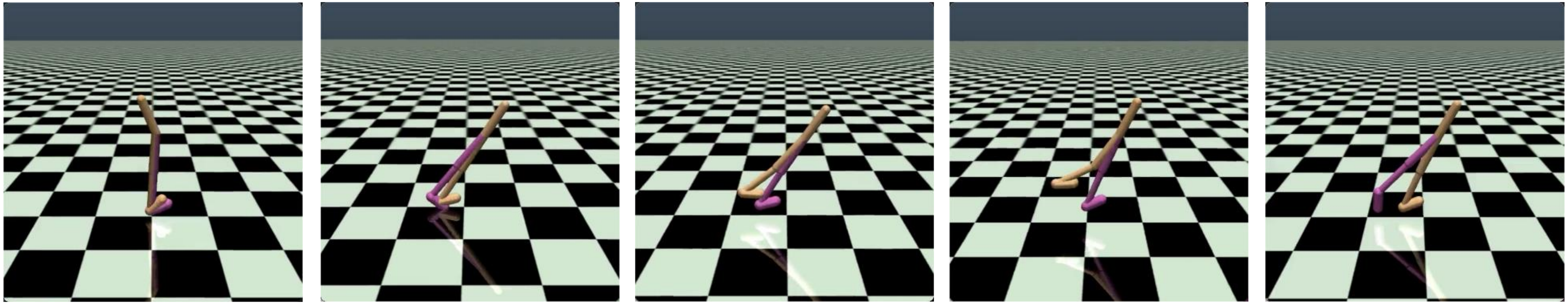
$$\pi_\theta(a_t\mid s_t)=\prod_{v\in O}\pi_\theta^{(v)}(a_t^{(v)}\mid s_t)$$

$$\mu_v=\Phi_O(h_v^{(K)})$$

$$\pi_\theta^{(v)}(a_t^{(v)}\mid s_t)=\frac{1}{\sqrt{2\pi\sigma_v}}e^{(a_t^{(v)}-\mu_v)^2/(2\sigma_v^2)}$$

2 行为→形态：形态控制

➢ 奖励函数设计

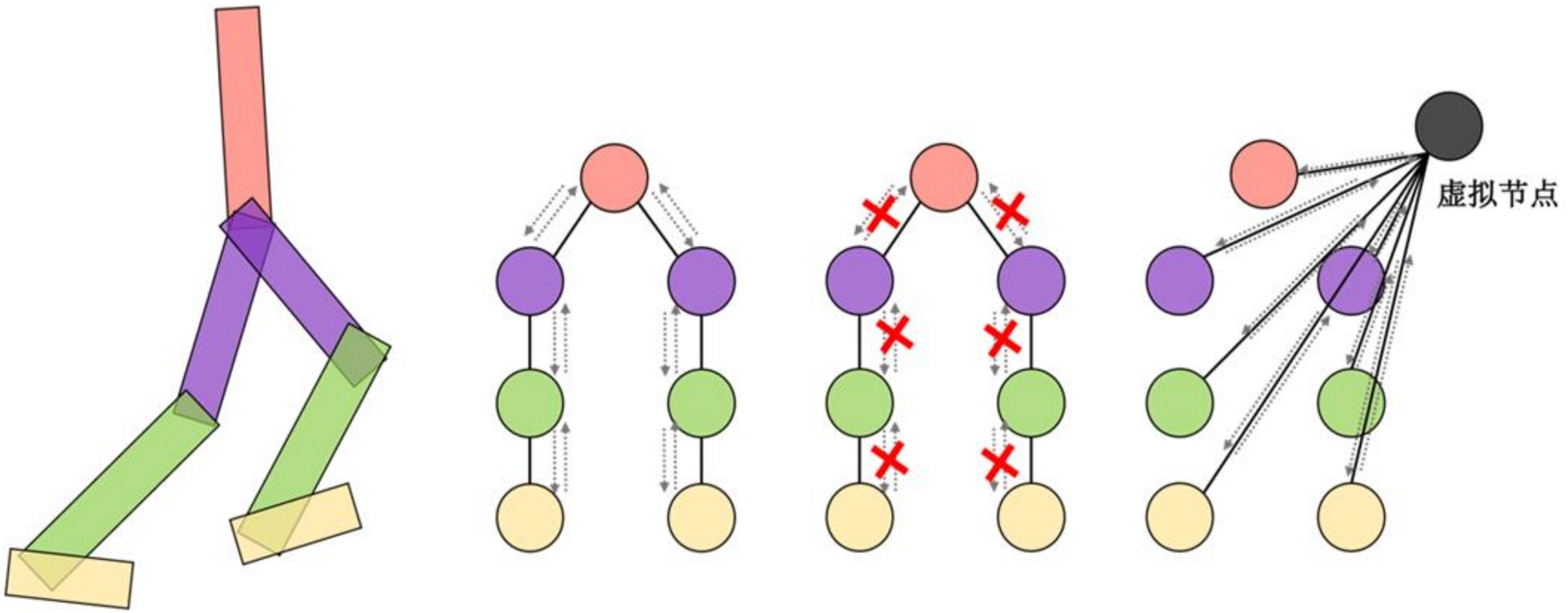


- 生存奖励 $r_{alive} = constant$
- 前进奖励 $r_{move} = (x_{after} - x_{before}) / dt$
- 控制奖励 $r_{control} = -\eta \cdot \sum_{i=1}^{|V|} a_i^2$

$$r = r_{alive} + r_{move} + r_{control}$$

2 行为→形态：形态控制

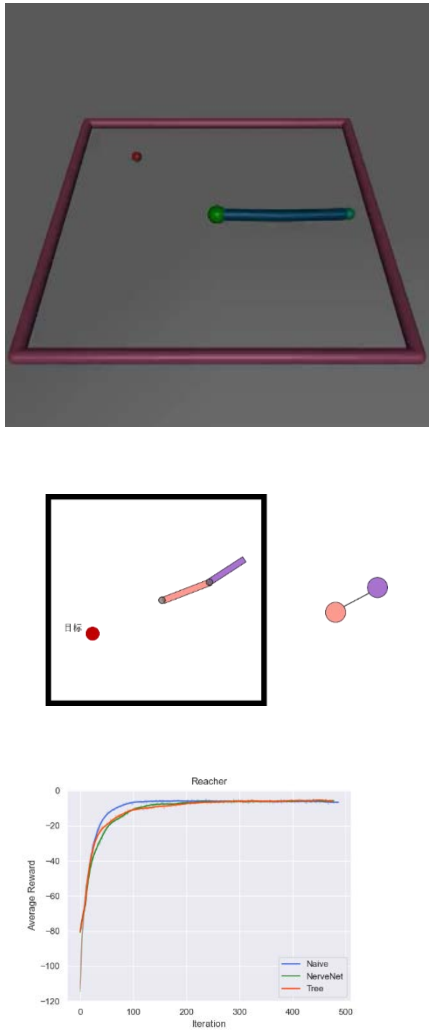
➢ 退化情形



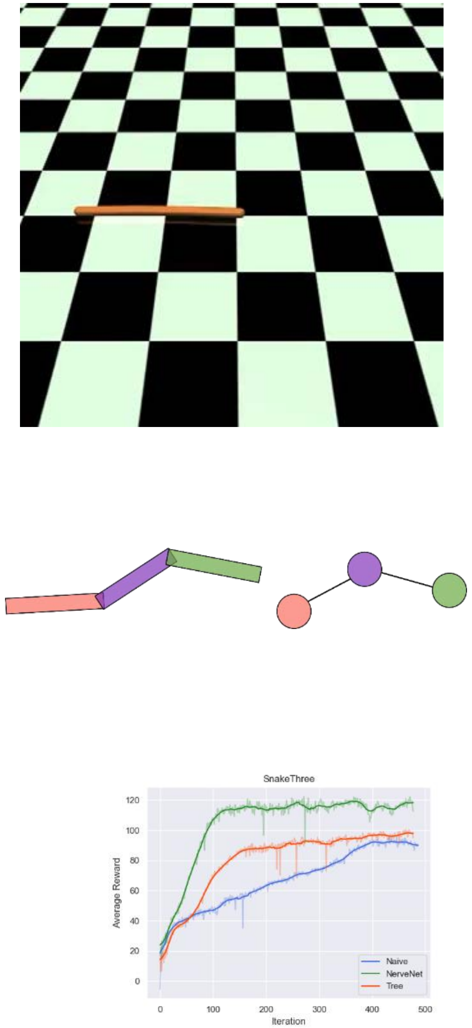
2 行为→形态：形态控制

➢ 例子：形态控制

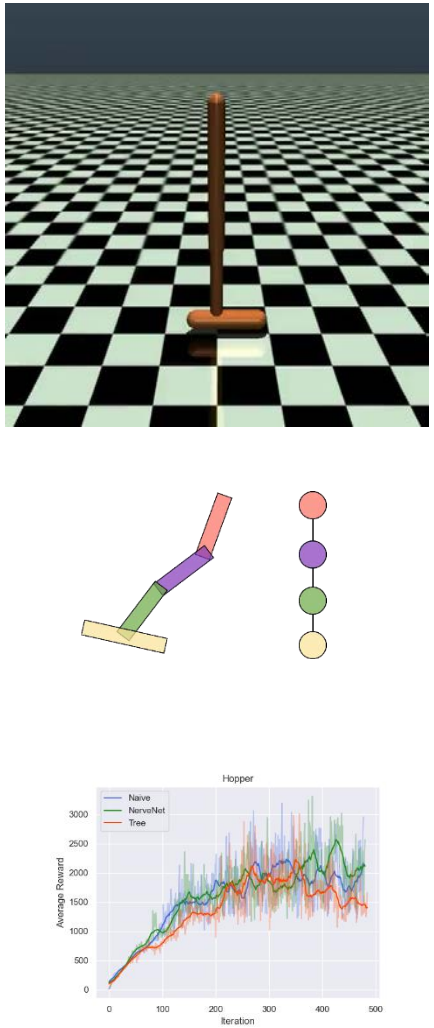
Reacher



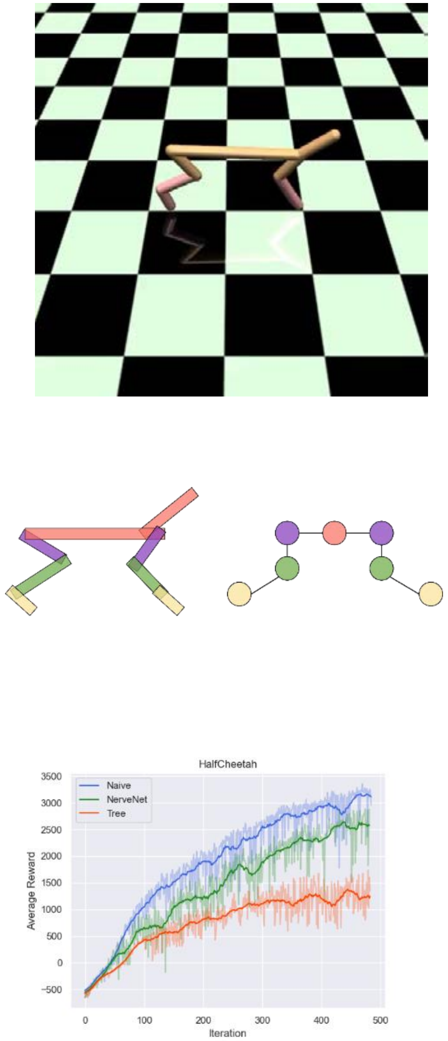
SnakeThree



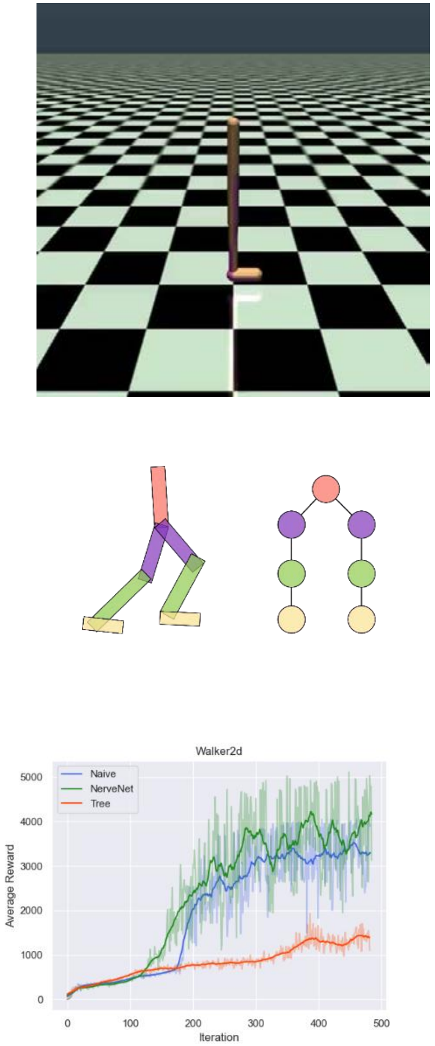
Hopper



HalfCheetah



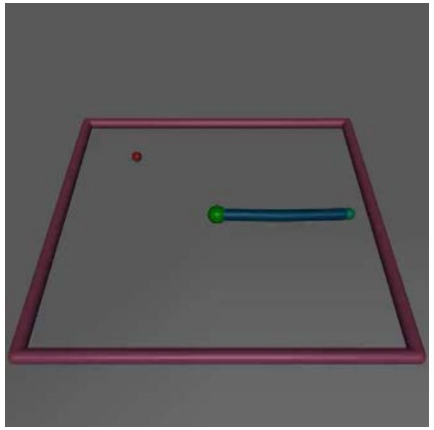
Walker2D

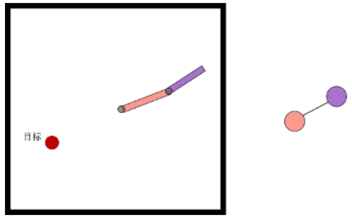


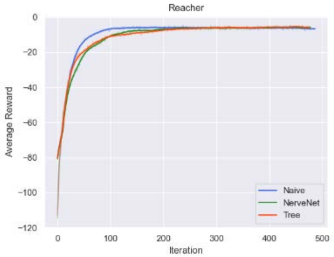
2 行为➞形态：形态控制

➢ 例子：形态控制

Reacher





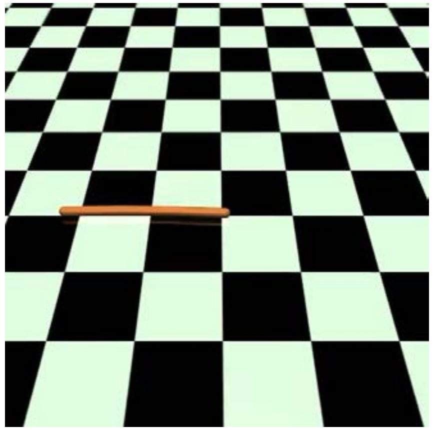


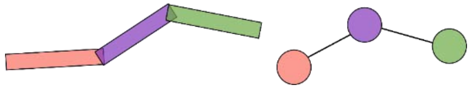
- Reacher是一个具备两个旋转关节的机械臂，任务目标是控制智能体的末端执行器接近一个在随机位置产生的目标。
- 组成奖励函数的两项均为惩罚项。实验结果的奖励值会随着训练的进行逐步增大并趋近于零。
- 由于Reacher环境中智能体结构相对简单，结构化策略的形态控制难以展现出其对形态结构理解的优势。

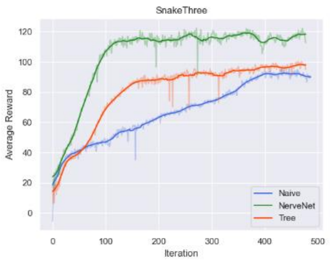
2 行为→形态：形态控制

➢ 例子：形态控制

SnakeThree





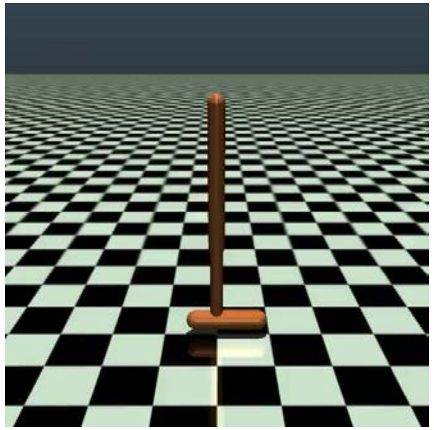


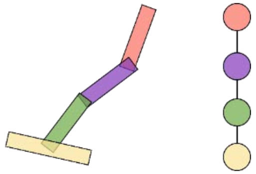
- SnakeThree是由三个刚体肢节首尾相连而成的蛇形智能体，每个关节连接两个肢节，在平面上初始化并开始运动，通过有规律地扭曲身体实现蛇行运动。
- 使用结构化策略的形态控制的NerveNet展现出了更好的效果。

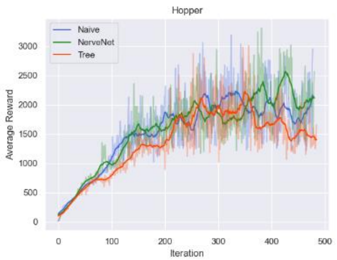
2 行为→形态：形态控制

➢ 例子：形态控制

Hopper





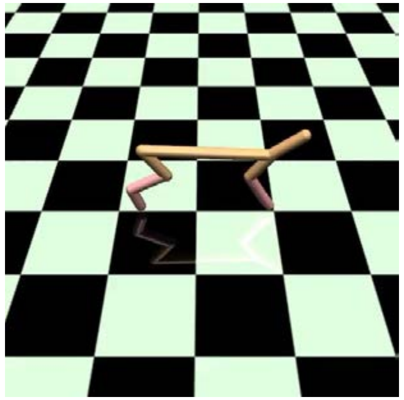


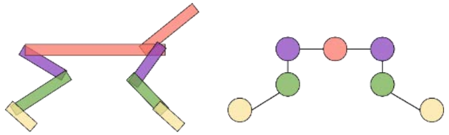
- Hopper是一个二维的单腿形体，由四个主要部分组成：顶部的躯干、中间的大腿、底部的小腿、以及支撑整个身体所依靠的足，任务目标是通过对连接身体四个部分的三个关节施加力矩，使身体向前移动。
- 采用结构化策略下的形态控制方法NerveNet没有体现出强势的竞争力，说明三种网络信息传递的拓扑结构复杂程度区别较小，NerveNet并没有将结构信息的优势充分体现出来。
- 视频中Hopper智能体首先处于直立状态，随后躯干前倾做出向前跳跃的趋势，并使用足作为主要的发力机构，落地时使用足首先着地并调整躯干做缓冲，使整个动作看起来更加协调。

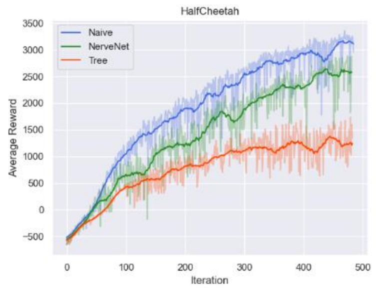
2 行为→形态：形态控制

➢ 例子：形态控制

HalfCheetah





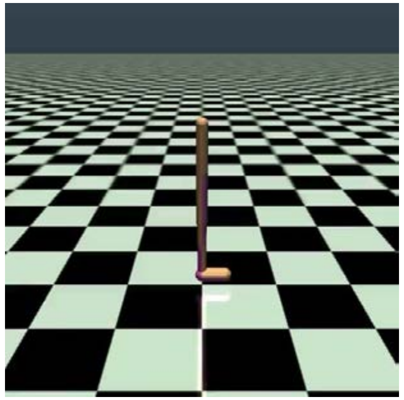


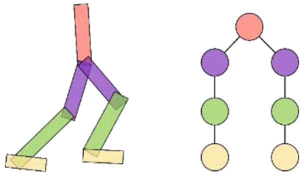
- HalfCheetah是一个二维机器人，由9个肢节和8个关节相连，任务目标是在关节上施加力矩，使身体尽可能快地向前跑。
- NerveNet并未在该任务环境中占据优势，收敛速度与收敛后的平均奖励低于不采用任何信息传递的Naive。除此以外，Tree网络全面落后于具有正确结构信息的NerveNet与无结构信息的Naive。这一点上也侧面证实了，在部分任务中神经网络考虑智能体形态结构时具有一定的局限性。
- 视频中，智能体从静止开始，后腿发力向前跳跃，前腿着地进行缓冲，并重复这一动作，快速向前奔跑跳跃。

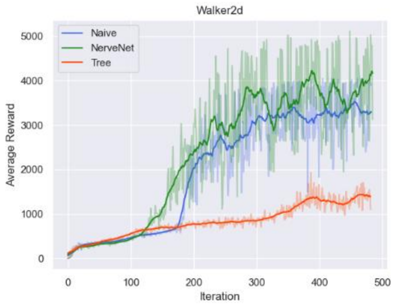
2 行为→形态：形态控制

➢ 例子：形态控制

Walker2D





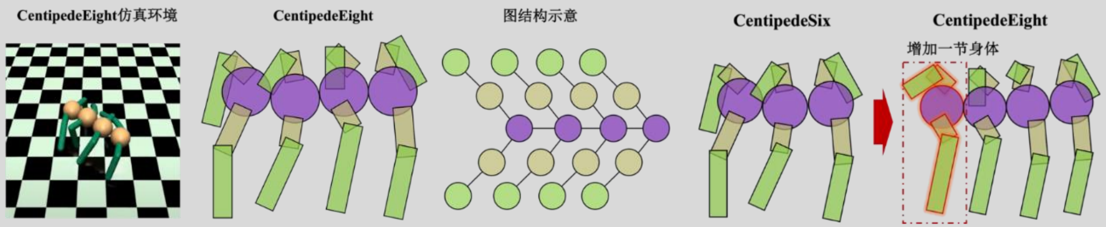


- Walker2D是一种二维的类人型智能体，任务目标为控制两组脚、小腿和大腿相协调，使智能体尽可能快而稳得向前方向移动。
- 使用NerveNet的结构化策略控制与无结构信息的Naive网络效果相似，而有错误结构信息的Tree结果不尽如人意。
- 视频中，智能体从直立状态开始，缓慢迈开双腿向前行走，随后双腿交替发力蹬地前进。

2 行为→形态：形态控制

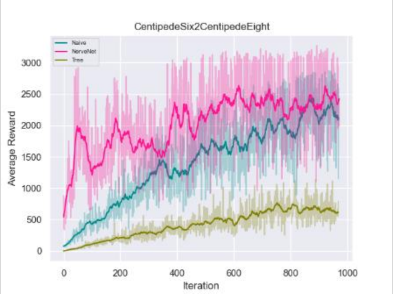
➢ 例子：迁移能力

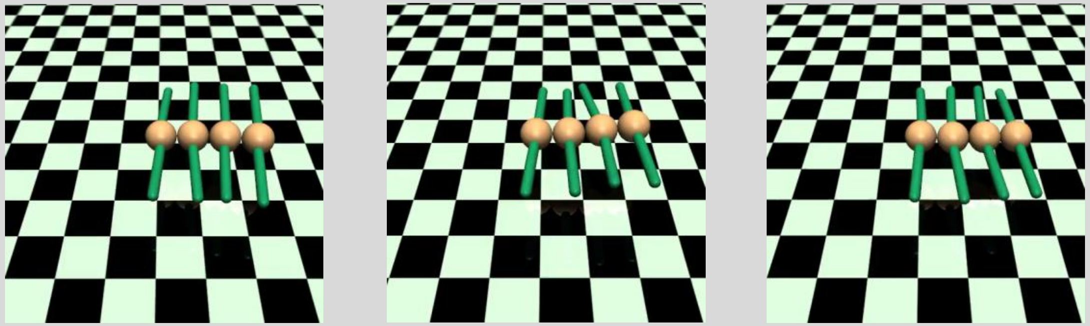
形态迁移



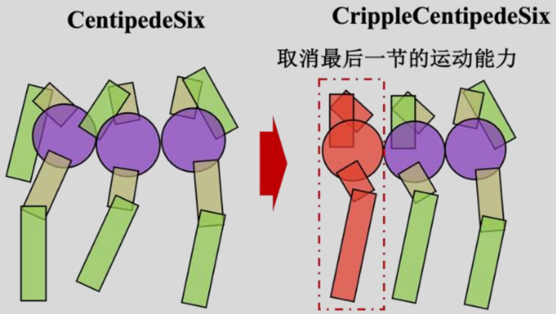
$$r = r_{alive} + r_{move} + r_{control} + r_{contact}$$

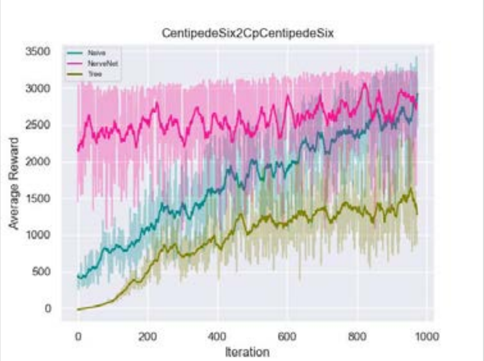
$$r_{contact} = -\eta_{cnt} \cdot \sum_{j \in SmallFoot} frc_j^2$$

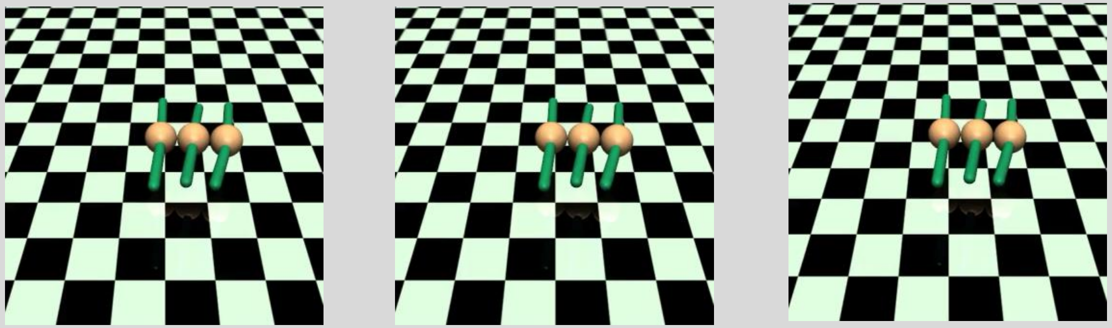




故障迁移



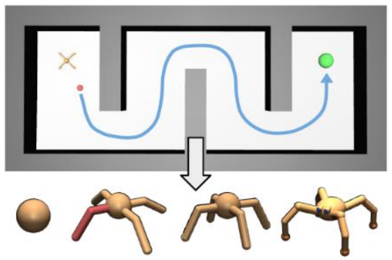


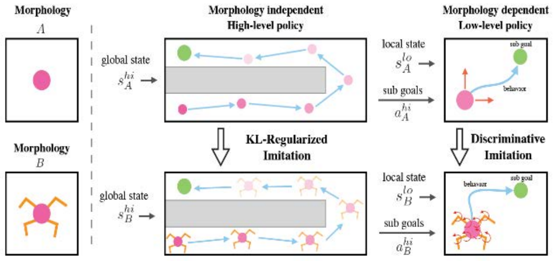


2 行为→形态：形态控制

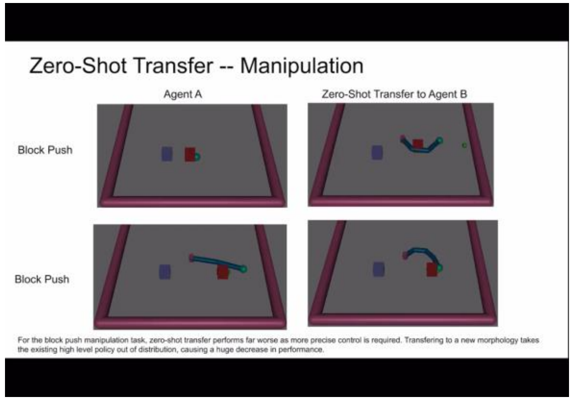
➢ 前沿研究

• 形态迁移



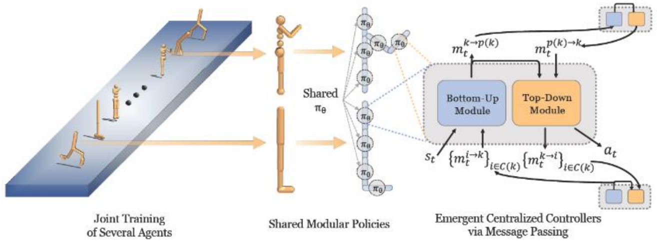


存在的挑战：如果不同形态的智能体的底层策略显著不同，则高层策略的迁移也很难成功。为此，作者引入互信息来极小化形态与底层行为之间的差异，从而实现不同形态智能体底层策略的对齐。



$$\min_{\theta_B^{lo}} I(\text{morphology}; \text{behavior})$$

• 统一控制



$$\max_{\theta} \mathbb{E}_{\pi_\theta} \sum_{n=1}^{N} \sum_{t=0}^{\infty} \left[\gamma^t r_t^n \left(\{s_{t+1}^k\}_{k=1}^{K_n}, \{a_t^k\}_{k=1}^{K_n} \right) \right]$$

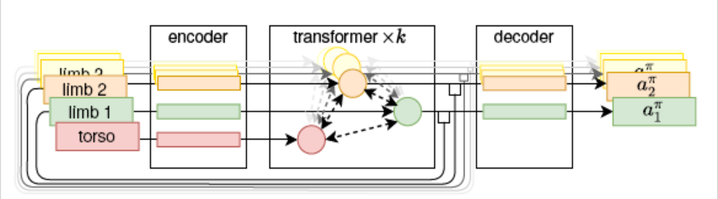
利用树形结构提出了不同形态智能体的统一控制方法。其核心在于将智能体的形态分解为不同的独立模块，针对独立的模块可以独立设计控制策略，这样就可以针对不同形态实现统一的控制。为了解决独立控制器之间难以协调的问题，论文研究了不同模块之间的通讯机制。

- Hejna D, Pinto L, Abbeel P. Hierarchically decoupled imitation for morphological transfer, ICML 2020
- Huang W, Mordatch I, Pathak D. One policy to control them all: Shared modular policies for agent-agnostic control, ICML, 2020

2 行为→形态：形态控制

➢ 前沿研究：Transformer

指出形态信息对应基于GNN的形态控制作用并不明显，且容易受到oversmoothing的影响。他们进一步提出了自注意力机制，允许节点直接直接通讯，并忽略了形态信息。然而，节点的位置信息对于自注意力机制并未被充分揭示。



AMORPHEUS

将形态（主要是position）信息嵌入transformer模型，用于异构形态的联合策略学习。解决了over-smoothing问题。

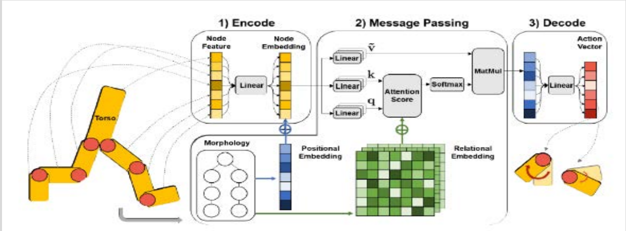
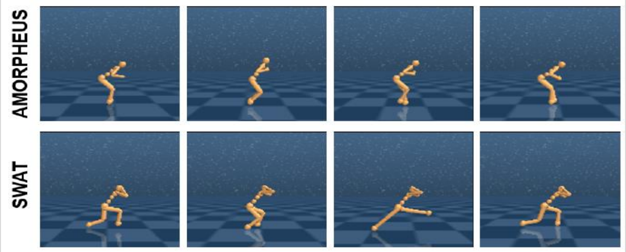
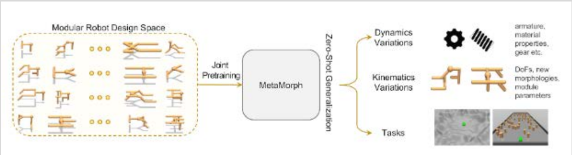
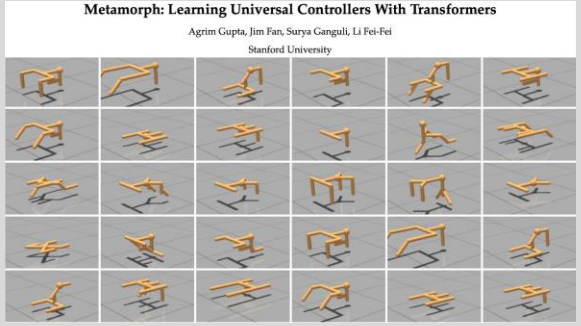


Figure 3: The overview of SWAT.



把机器人的形态(morphology)看做是一种输入transformer的模态，进而学习一个通用策略同时解决大量不同机器人的控制问题。为形态预训练提供了新的思路。

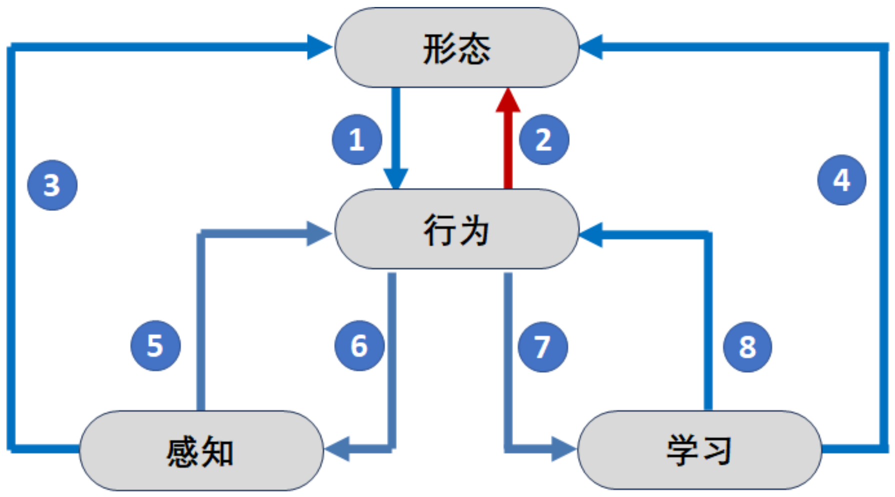


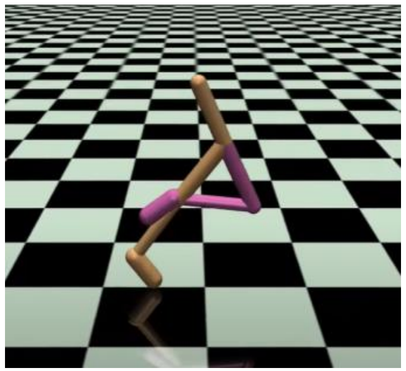


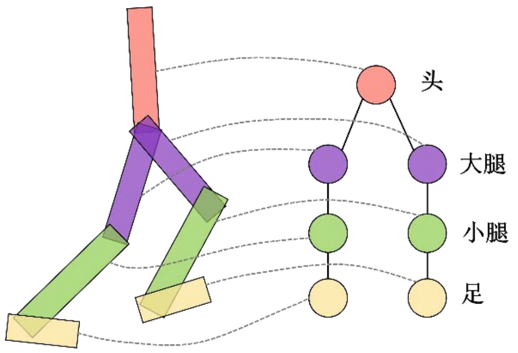
- Kurin V, Igl M, Rocktäschel T, et al. My body is a cage: the role of morphology in graph-based incompatible control, ICLR,2021
- Hong S, Yoon D, Kim K E. Structure-Aware Transformer Policy for Inhomogeneous Multi-Task Reinforcement Learning, ICLR. 2022.
- Gupta A, Fan L, Ganguli S, et al. MetaMorph: Learning Universal Controllers with Transformers[J]. arXiv preprint arXiv:2203.11931, 2022.

2 行为→形态：形态控制

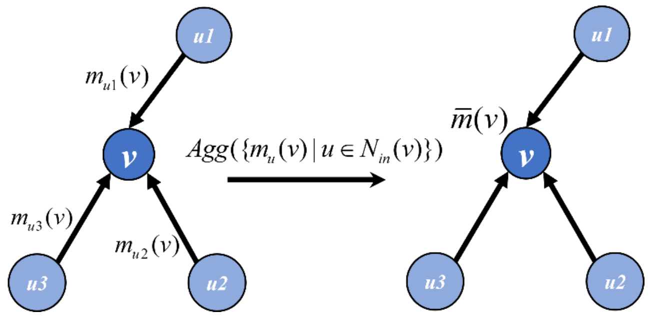
➢ 小结

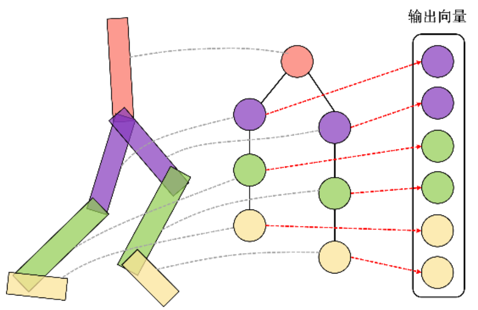






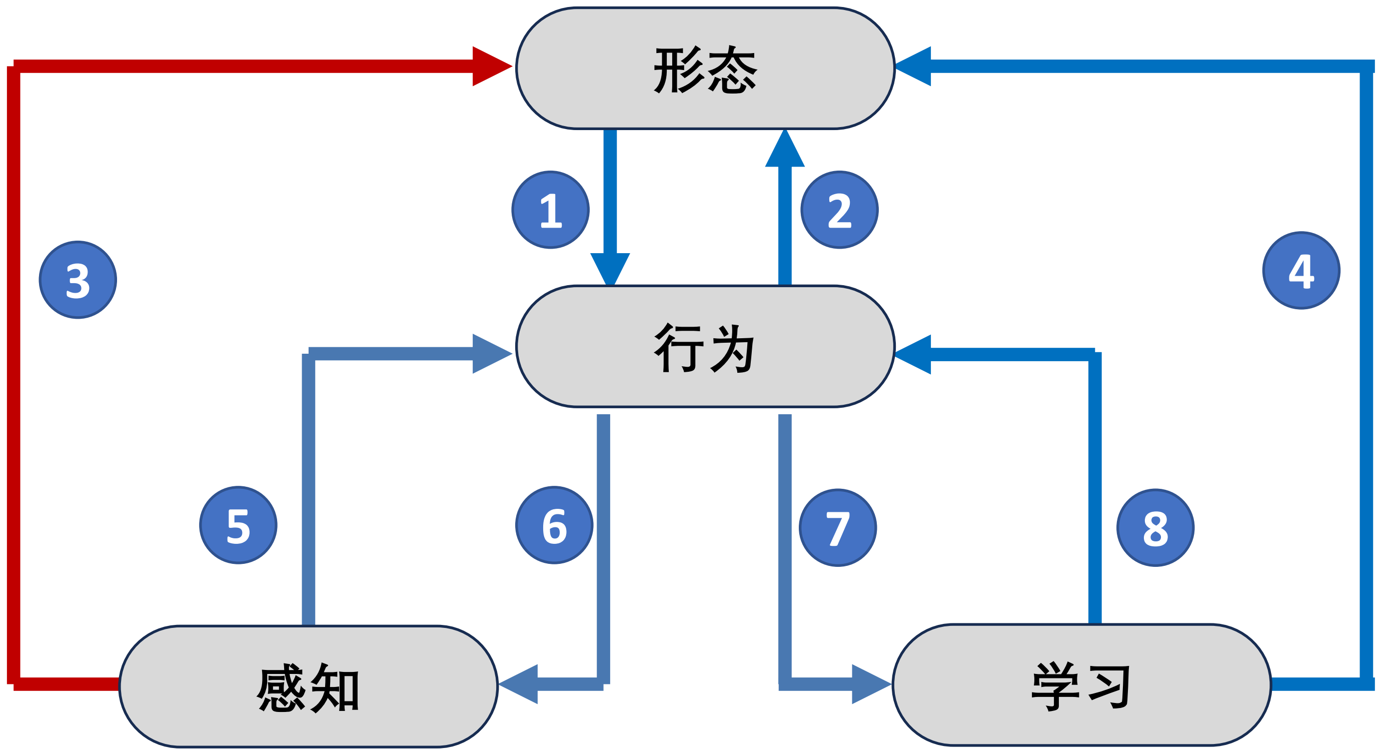
$$\pi_\theta(a_t \mid s_t)$$





- 图神经网络、Transformer，……

3 感知→形态：形态变换



3 感知➞形态：形态变换

➢ 地形适应的形态变换

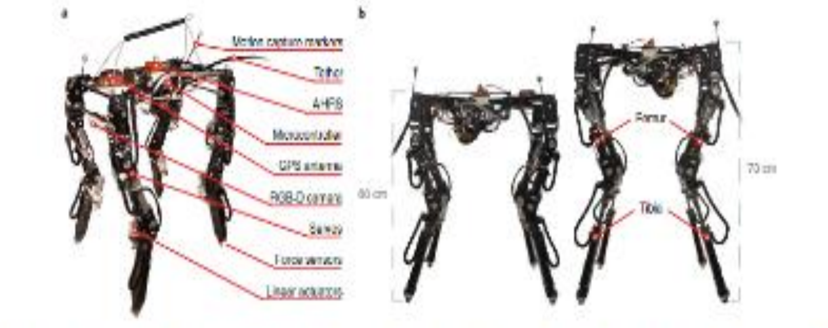
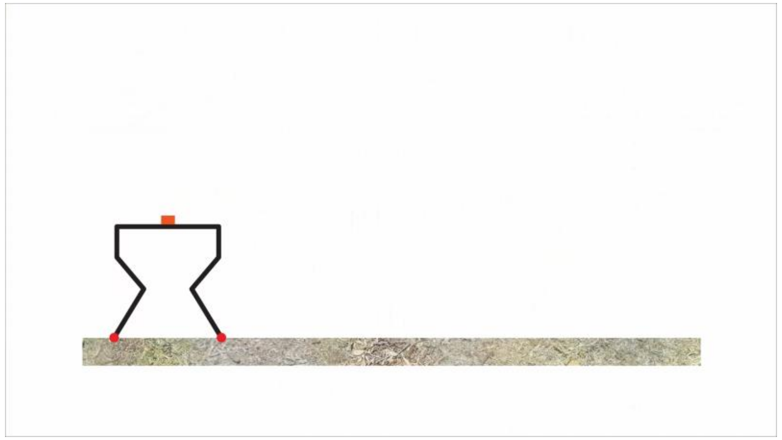
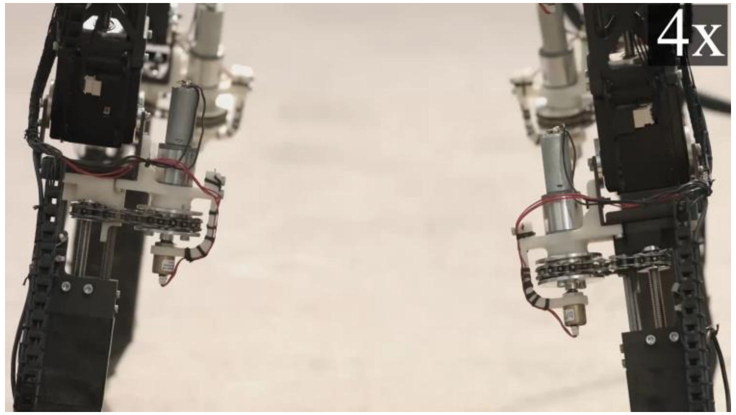
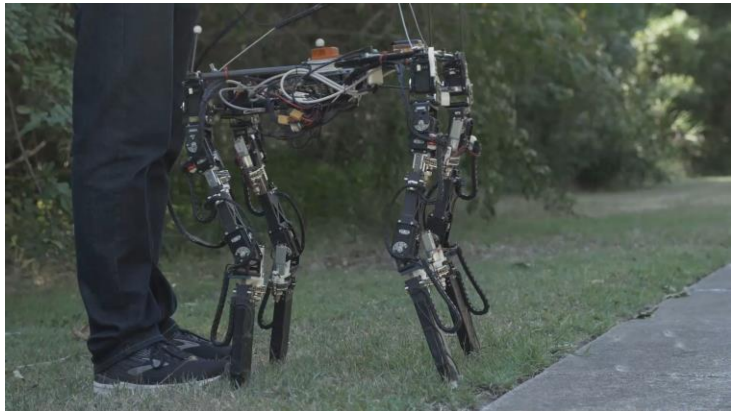


Fig. 1 | The morphologically adaptive robot used in this study. a, An overview of the main components of the robot. b, The robot with the shortest (left) and longest (right) leg configuration. AHS, attitude and heading reference system; GPS, global positioning system; RGB-D, red, green, blue and depth.





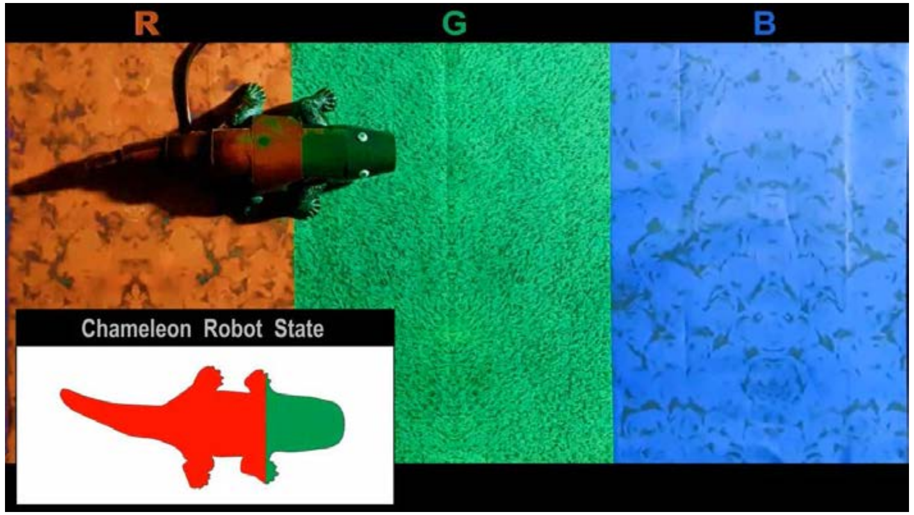


- Nygaard T F, Martin C P, Torresen J, et al. Real-world embodied AI through a morphologically adaptive quadruped robot[J]. Nature Machine Intelligence, 2021, 3(5): 410-419.

3 感知→形态：形态变换

➢ 仿变色龙软体机器人

将人造变色龙皮肤应用于软体机器人上，并结合颜色传感器和反馈控制系统，从而使得这种设备级的自适应人造伪装技术能够检测到背景环境的颜色，并令变色龙软体机器人上实现实时背景颜色匹配。



- Kim H, Choi J, Kim K K, et al. Biomimetic chameleon soft robot with artificial crypsis and disruptive coloration skin[J]. Nature communications, 2021, 12(1): 1-11.

3 感知➞形态：形态变换

➢ 变形机器人

nature communications

Article

https://doi.org/10.1038/s41467-023-39018-y

Multi-Modal Mobility Morphobot (M4) with appendage repurposing for locomotion plasticity enhancement

Received: 4 January 2023

Accepted: 22 May 2023

Published online: 27 June 2023

Eric Sihite[1], Arash Kalantari[2], Reza Nemovi[1], Alireza Ramezani[3] & Morteza Gharib[1]

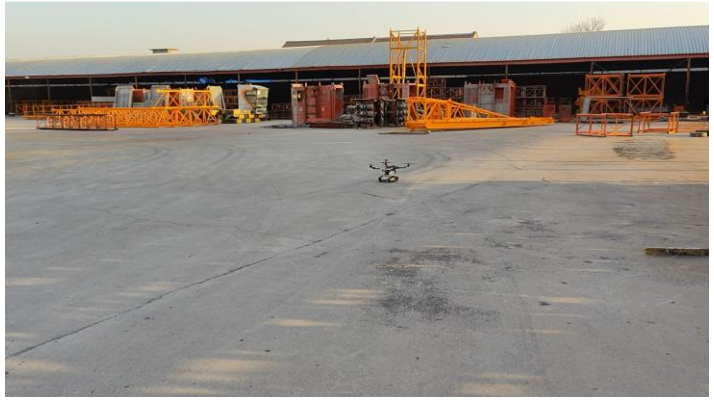
论文链接：https://www.nature.com/articles/s41467-023-39018-y

3 感知→形态：形态变换

➢ 两栖飞行汽车

I代飞车（无人）

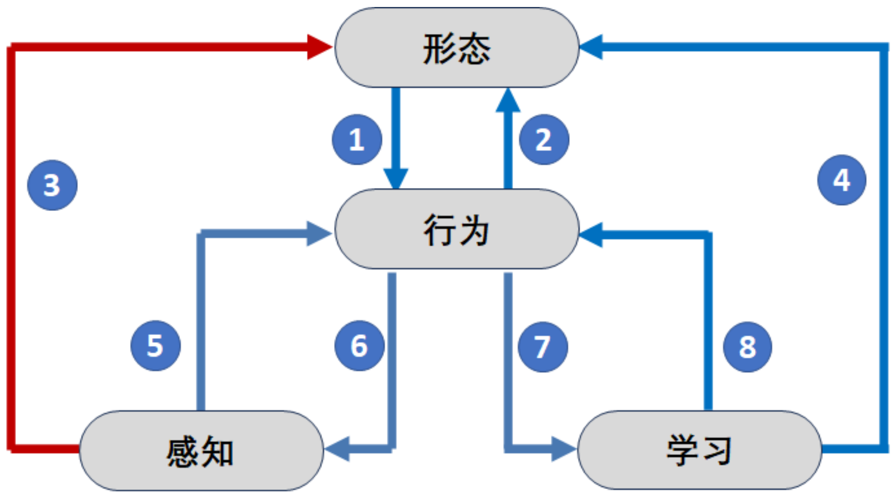
II代飞车（可折叠）

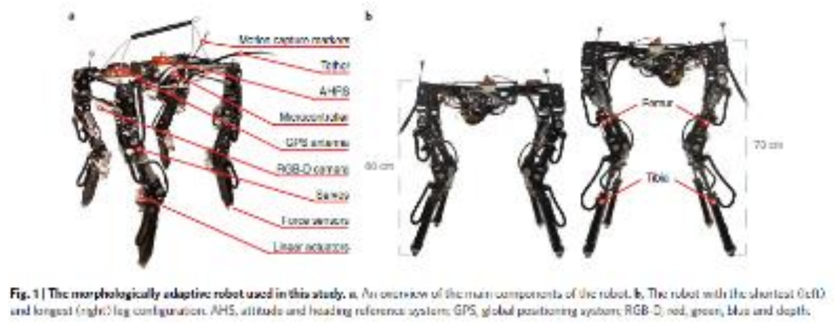


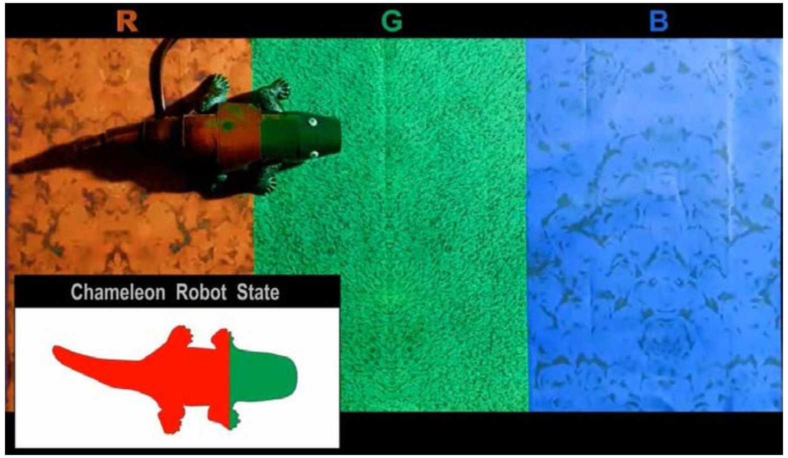
III代飞车（载人）

3 感知→形态：形态变换

➢ 小结

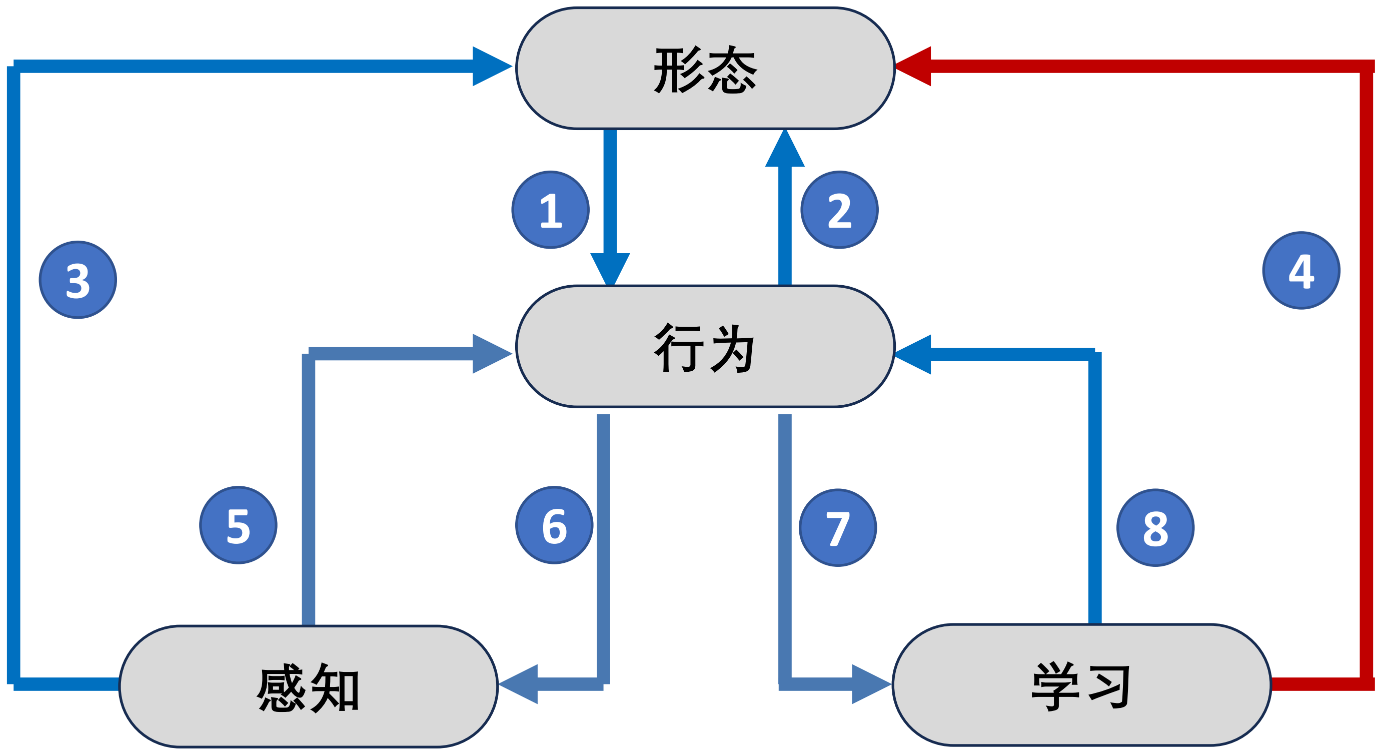






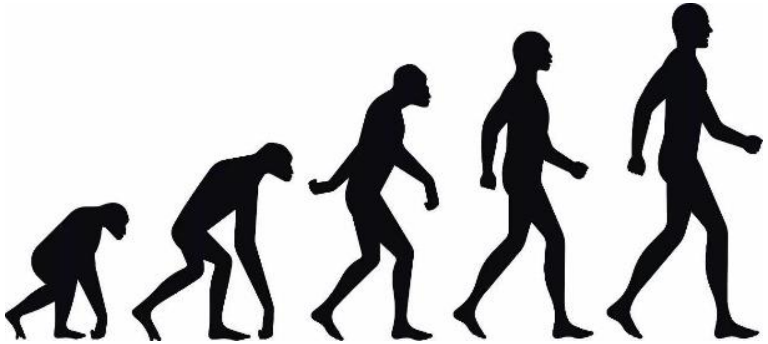
- 材料、机构，……

4 学习→形态：形态生成



4 学习→形态：形态生成

➢ 脑-体协同进化



$$(D^*,\pi^*)=\underset{D,\pi}{\arg\max}\,J(D,\pi)$$

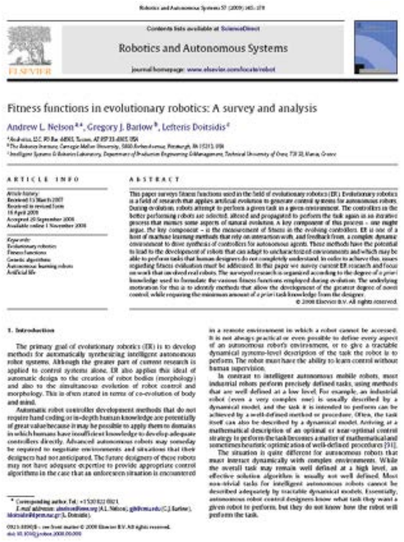
$$\pi^*=\underset{\pi}{\arg\max}\,J(\bar{D},\pi)$$

4 学习➔形态：形态生成

➢ 脑-体协同进化

- 利用学习思想实现脑-体协同进化的思想在具身认知与具身智能研究的早期就得到了充分的重视[Pfeifer et al, 2006a]，有时也称为进化机器人（evoluiontary robots）、人工生命（artificial life）。区别于早期进化机器人学方面的研究主要侧重于利用进化学习算法优化设计特定形态机器人的控制策略，而并不能影响机器人的形态

- 人类更擅长设计物理系统，而非智能控制系统


Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Fitness functions in evolutionary robotics: A survey and analysis

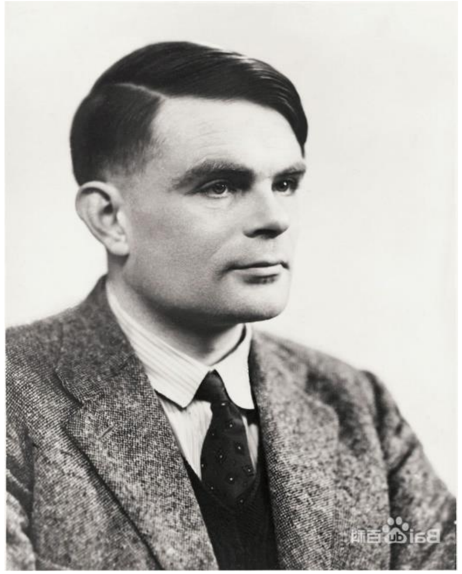
Andrew L. Nelson, Gregory J. Barlow, Lefteris Doitsidis

ARTICLE INFO

ABSTRACT

4 学习→形态：形态生成

➢ 脑-体协同进化



“拥有一个身体就是拥有一个通用的装置、拥有一个涵盖所有类型的知觉展开图式。” 阿兰·图灵（Alan Turing）在论文《计算机器与智能》（Computing Machinery and Intelligence）中，提出了一种能借助传感器与环境互动并自行学习的人工智能，而这正是如今“具身智能”的最初构想。

- “We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child.”

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game." It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A." The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

Now suppose X is actually A, then A must answer. It is A's object in the game to try and cause C to make the wrong identification. His answer might therefore be:

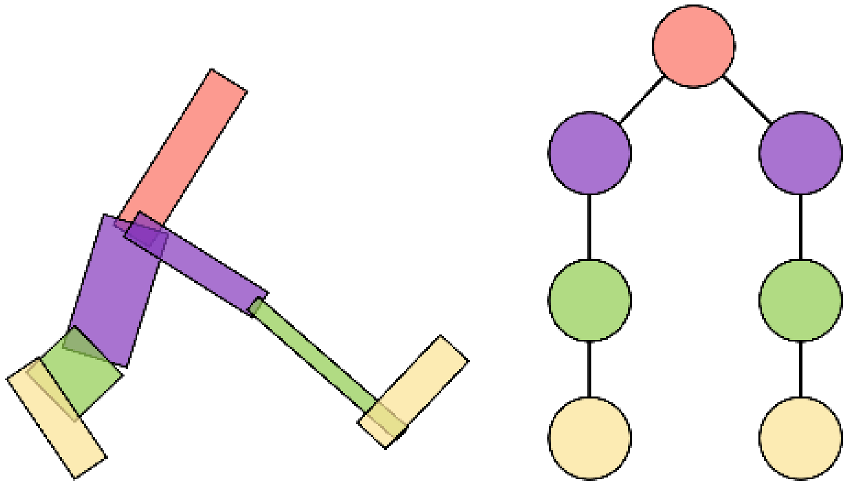
"My hair is shingled, and the longest strands are about nine inches long."

In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. Alternatively the question and answers can be repeated by an intermediary. The object of the game for the third player (B) is to help the interrogator. The best strategy for her is probably to give truthful answers. She can add such things as "I am the woman, don't listen to him!" to her answers, but it will avail nothing as the man can make similar remarks.

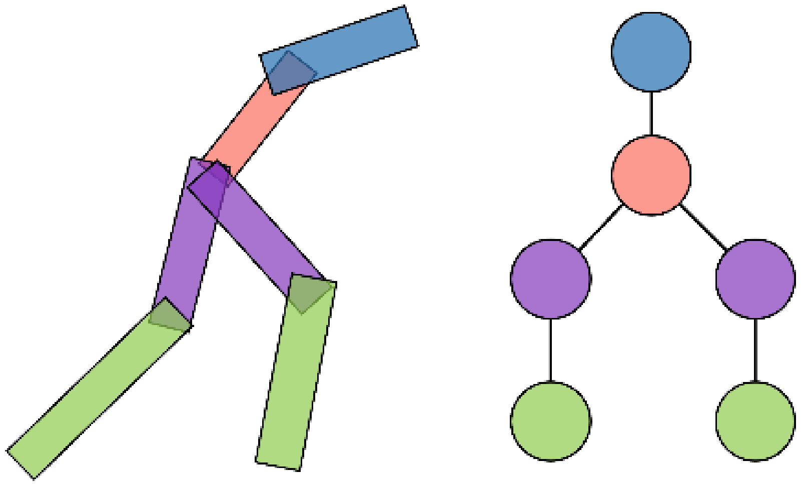
We now ask the question, "What will happen when a machine takes the part of A in this game?" Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, "Can machines think?"

4 学习→形态：形态生成

➢ 脑-体协同进化



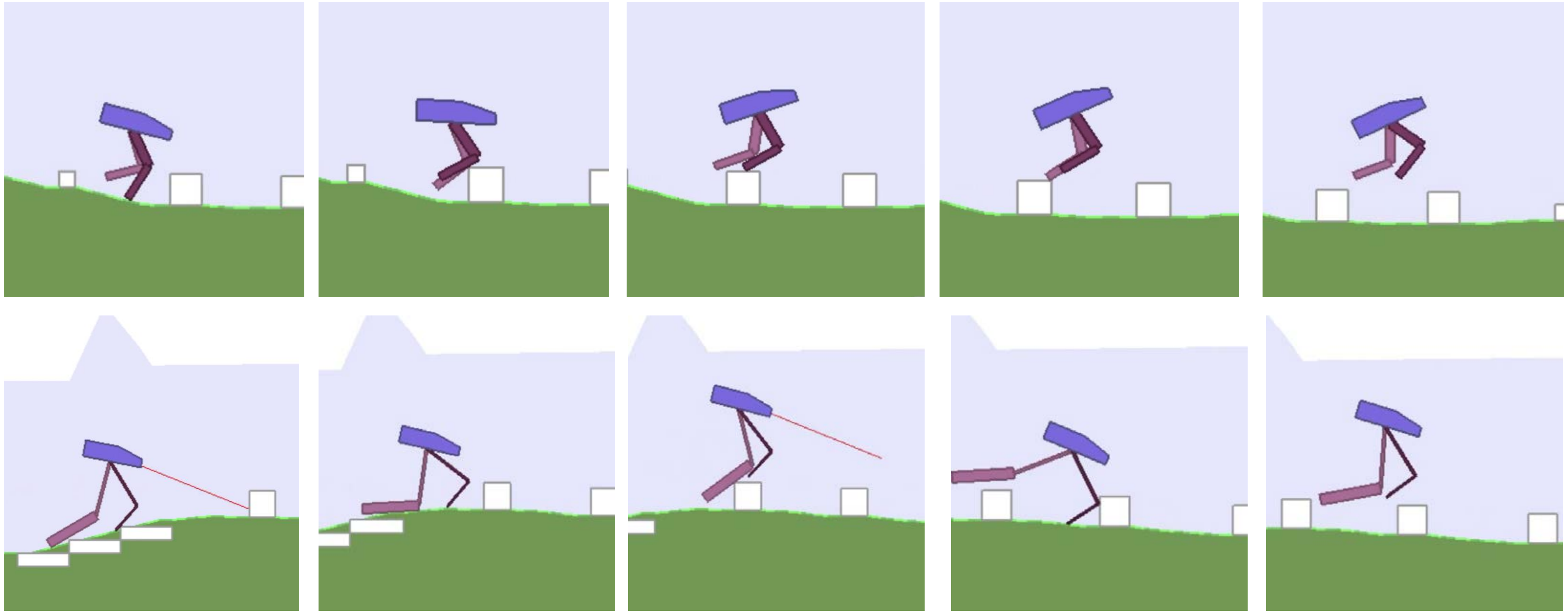
形态参数



形态结构

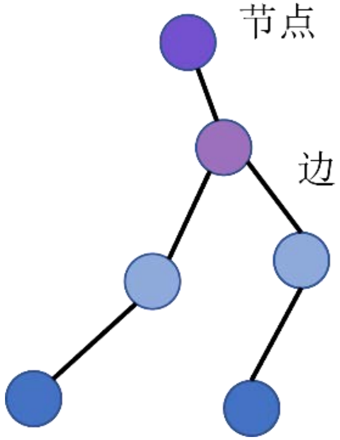
4 学习➝形态：形态生成

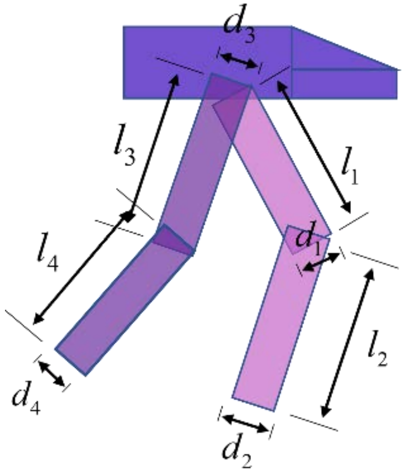
➢ 形态参数的优化



4 学习→形态：形态生成

➢ 形态参数的优化





$$\{l_1, d_1, l_2, d_2, l_3, d_3, l_4, d_4\}$$

$$J(\theta_w) = \int R(w) p_{\theta_w}(w) dw$$

形态参数对应的奖励　　形态参数　　形态参数对应的分布

$$\theta_w^* = \arg\max_{\theta_w} \int R(w) p_{\theta_w}(w) dw$$

$$\nabla_{\theta_w} J(\theta_w) = \int_w R(w) \nabla_{\theta_w} \log p_{\theta_w}(w) dw$$

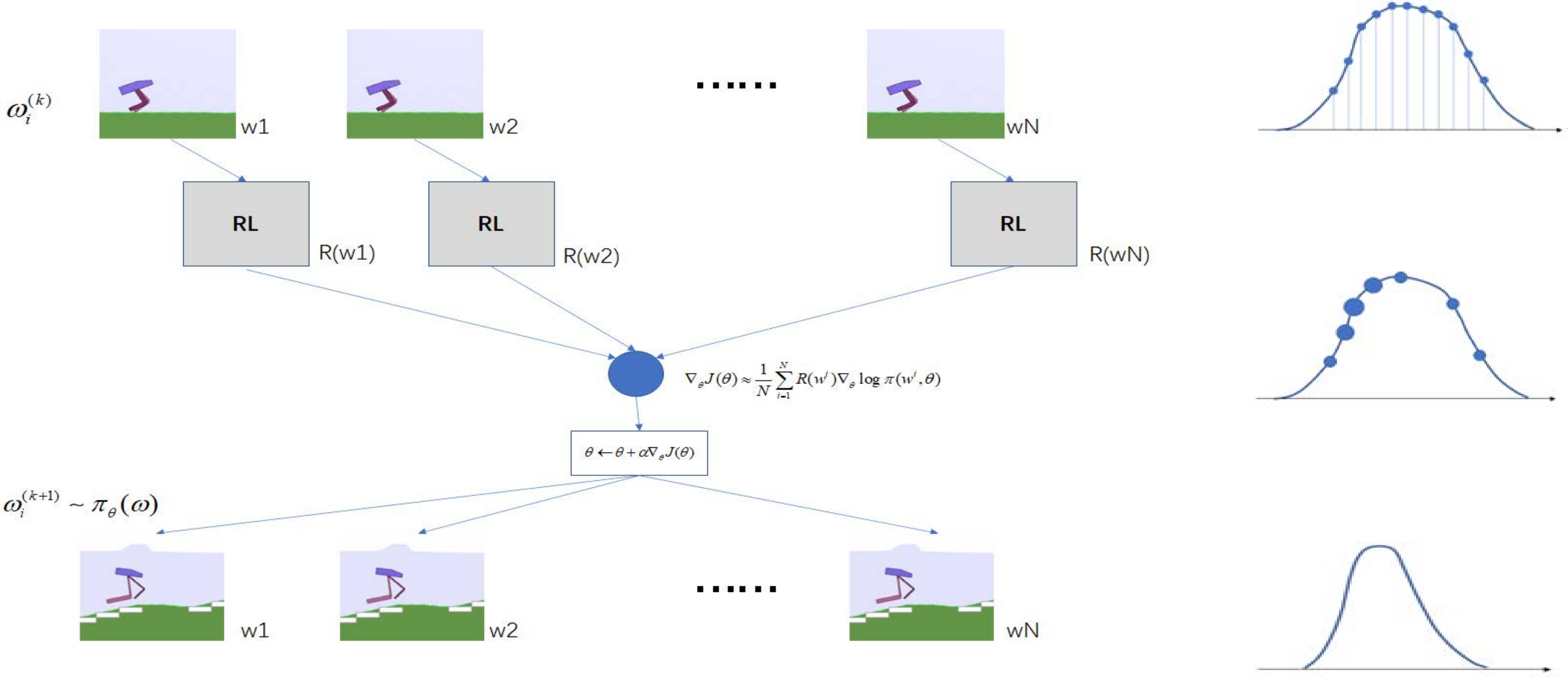
$$\theta_w \leftarrow \theta_w + \eta \nabla_{\theta_w} J(\theta_w)$$

$$w^{(i)} \sim p_{\theta_w}(w)$$

$$\nabla_{\theta_w} J(\theta_w) \approx \frac{1}{N} \sum_{i=1}^{N} R(w^i) \nabla_{\theta_w} \log p_{\theta_w}(w^i)$$

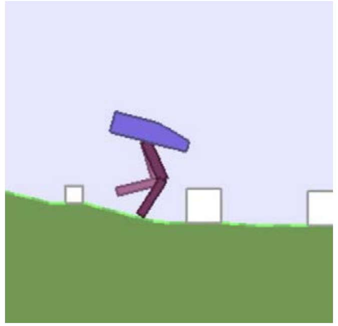
4 学习➞形态：形态生成

➢ 形态参数的优化



4 学习➞形态：形态生成

➢ 形态参数的优化



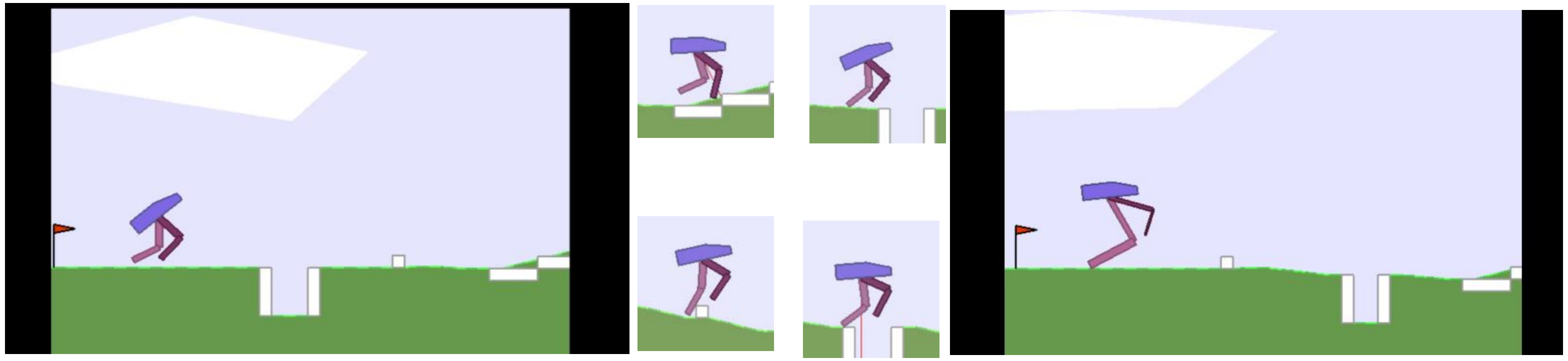
$$r_{alive} = -100$$

$$r_{move} = \eta_{mv} \cdot (x_{after} - x_{before})$$

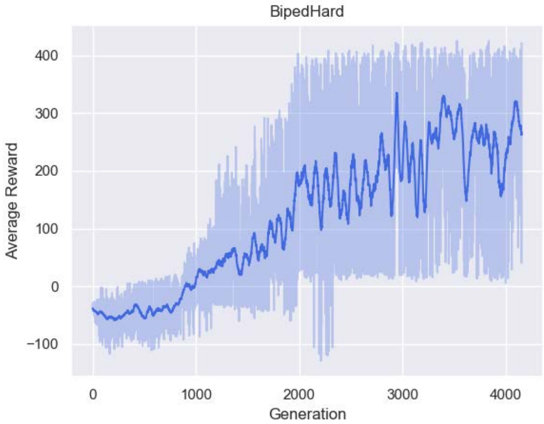
$$r_{balance} = -\eta_{balance} \cdot |\alpha|$$

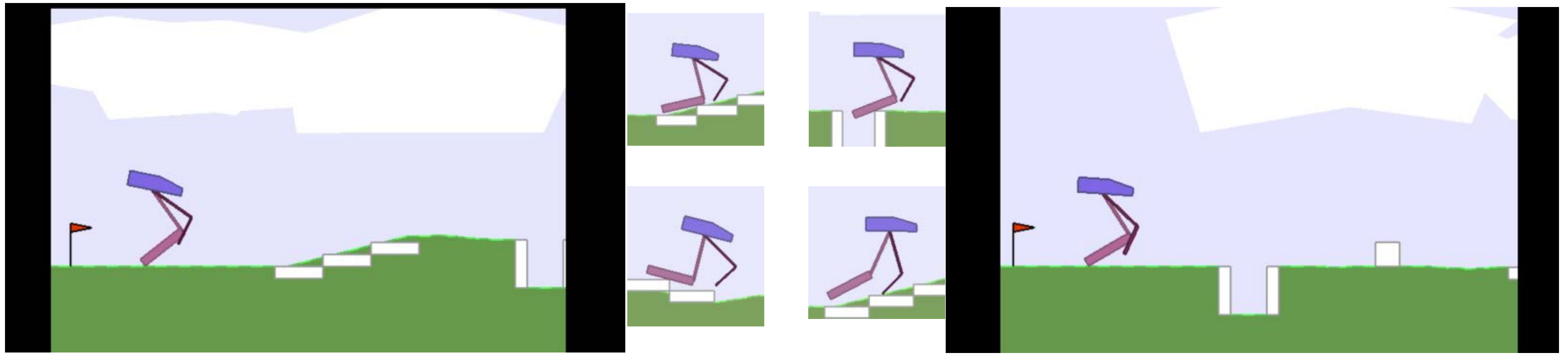
$$r_{control} = -\eta_{ctl} \cdot T_0 \cdot \sum_{i=1}^{|V|} a_i^2$$

$$r = r_{alive} + r_{move} + r_{balance} + r_{control}$$



Reward equals to around 200

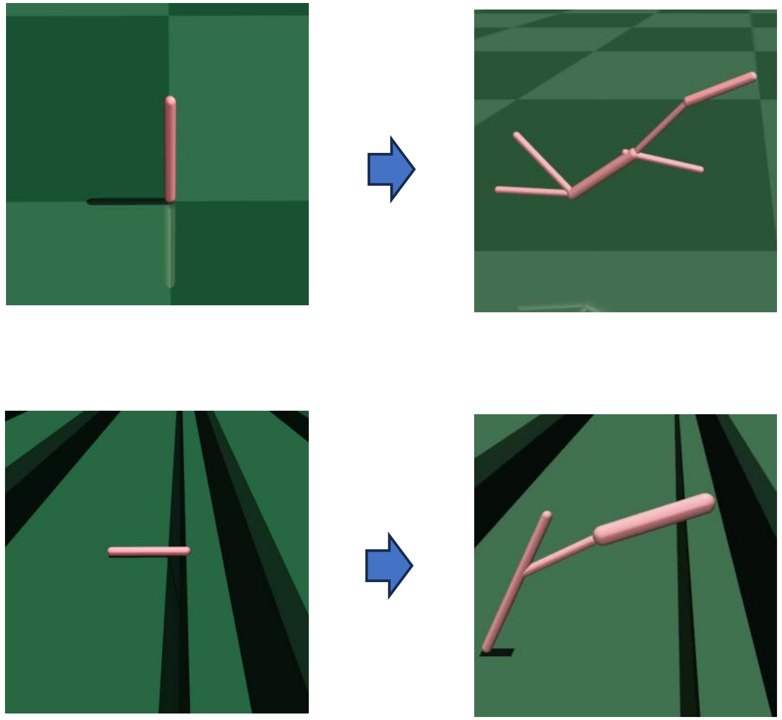




Reward equals to around 300

4 学习➞形态：形态生成

➢ 结构-参数的联合优化



$$D_t = (V_t, E_t, Z_t)$$

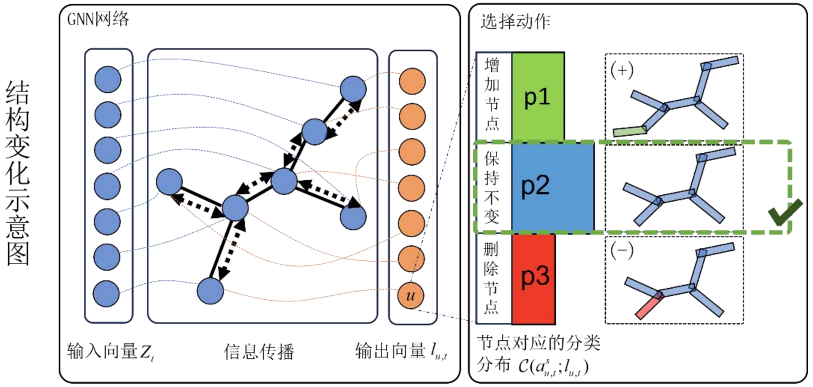
$$s_t = (s_t^e, D_t, \Phi_t) \qquad a_t = (a_t^d, a_t^e)$$

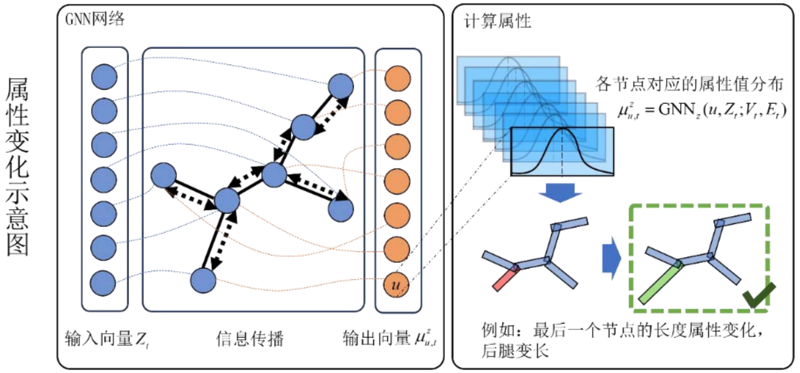
$$T(s_{t+1}^e, D_{t+1}, \Phi_{t+1} \mid s_t^e, D_t, \Phi_t, a_t^d, a_t^e)$$

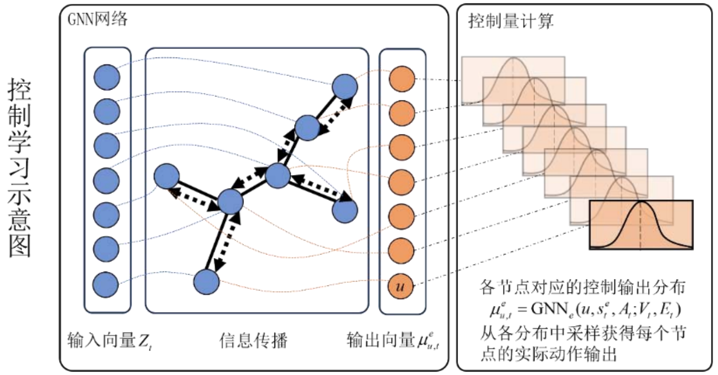
$$\pi_\theta(a_t^d, a_t^e \mid s_t^e, D_t, \Phi_t) = \begin{cases} \pi_{\theta_1}^d(a_t^d \mid D_t, \Phi_t) & \Phi_t = \text{形态变化} \\ \pi_{\theta_2}^e(a_t^e \mid s_t^e, D_t, \Phi_t) & \Phi_t = \text{形态控制} \end{cases}$$

4 学习→形态：形态生成

➤ 结构-参数的联合优化



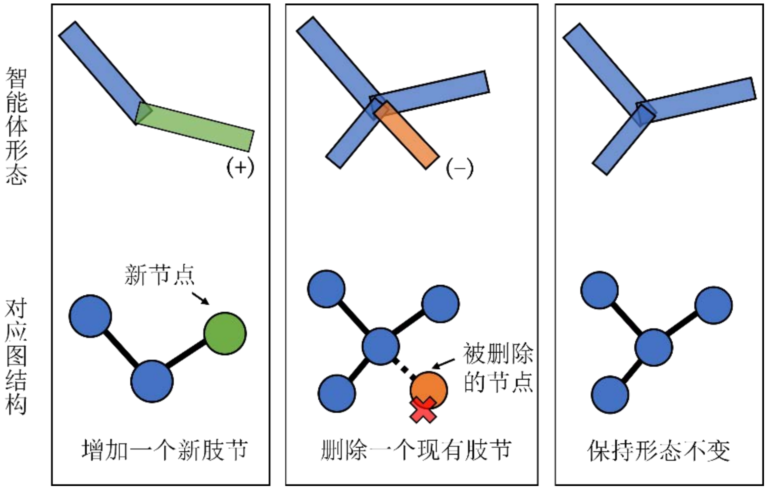


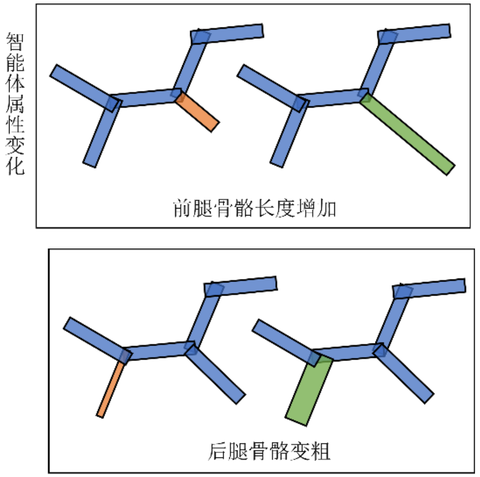


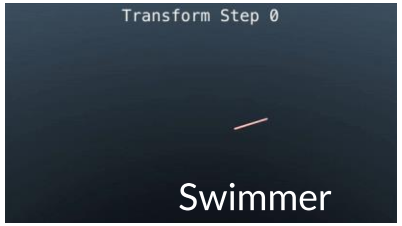
$$\pi_\theta^s(a_t^s \mid D_t,\Phi_t)=\prod_{u\in V_t}\pi_\theta^s(a_{u,t}^s \mid D_t,\Phi_t)$$

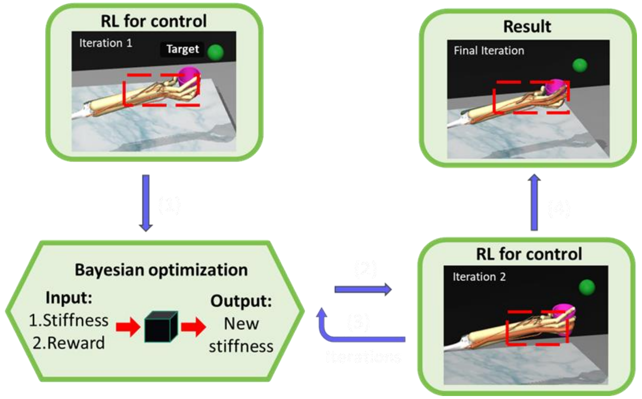
$$\pi_\theta^z(a_t^z \mid D_t,\Phi_t)=\prod_{u\in V_t}\pi_\theta^z(a_{u,t}^z \mid D_t,\Phi_t)$$

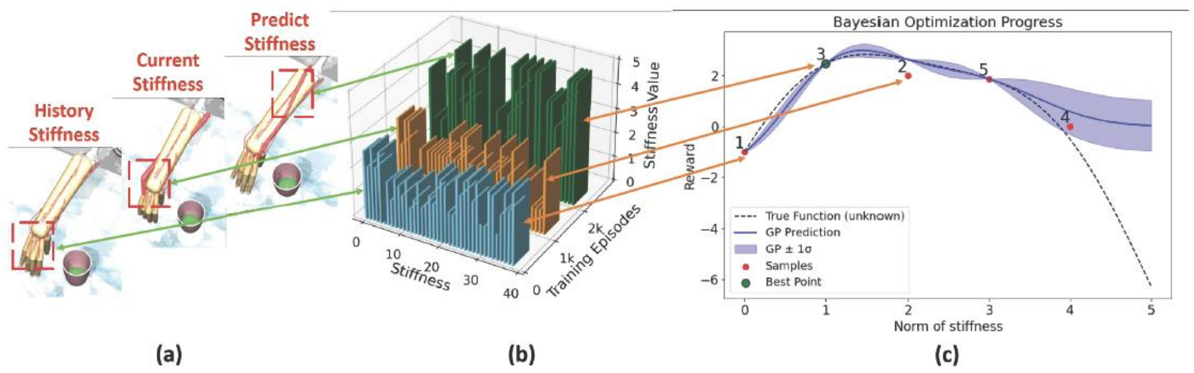
$$\pi_\theta^e(a_t^e \mid s_t^e,D_t,\Phi_t)=\prod_{u\in V_t}\pi_\theta^e(a_{u,t}^e \mid s_t^e,D_t,\Phi_t)$$

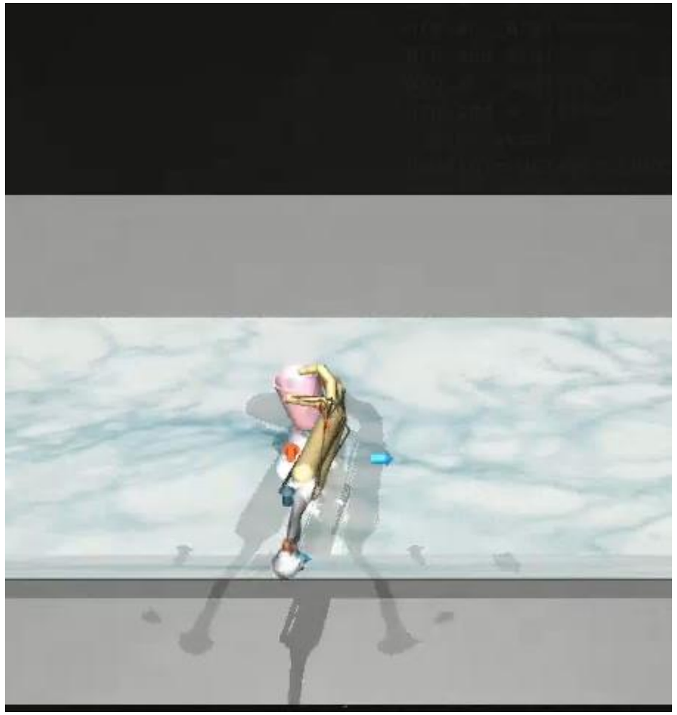






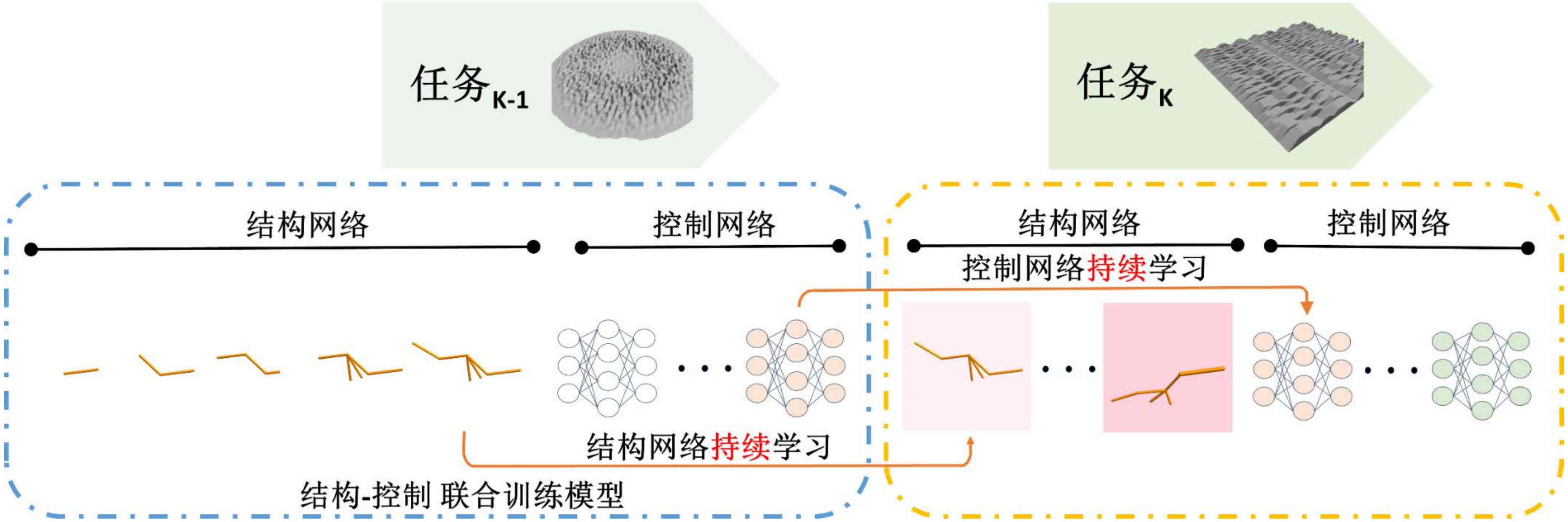

RL for control
Iteration 1
Target
Result
Final Iteration
Bayesian optimization
Input:
1.Stiffness
2.Reward
Output:
New
stiffness
RL for control
Iteration 2

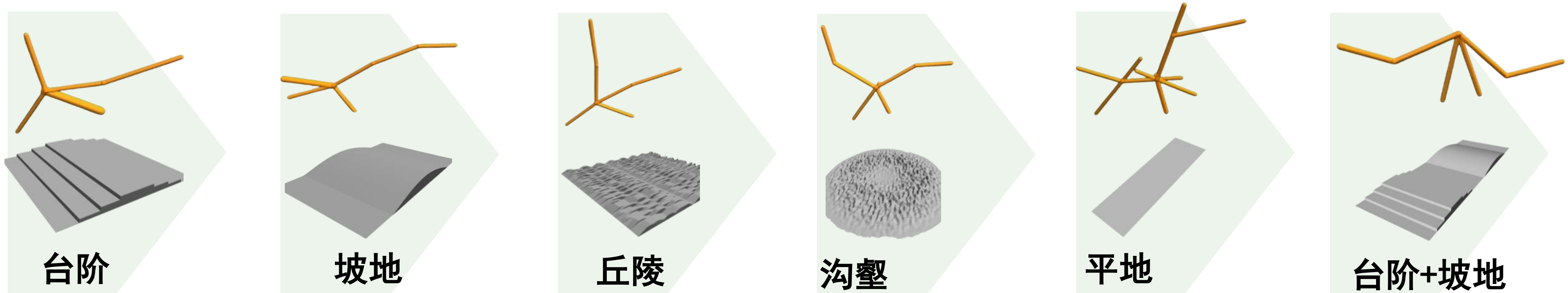

History
Stiffness
Current
Stiffness
Predict
Stiffness
Stiffness Value
Stiffness
Training Episodes
Bayesian Optimization Progress
Reward
Norm of stiffness
True Function (unknown)
GP Prediction
GP ± 1σ
Samples
Best Point
(a)
(b)
(c)



4 学习→形态：形态生成

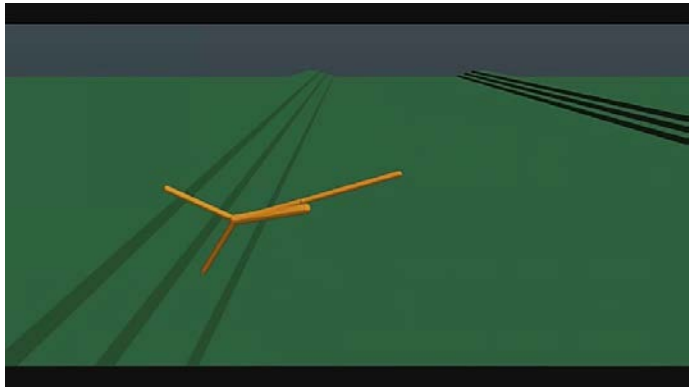
➢ 结构-参数的联合持续优化



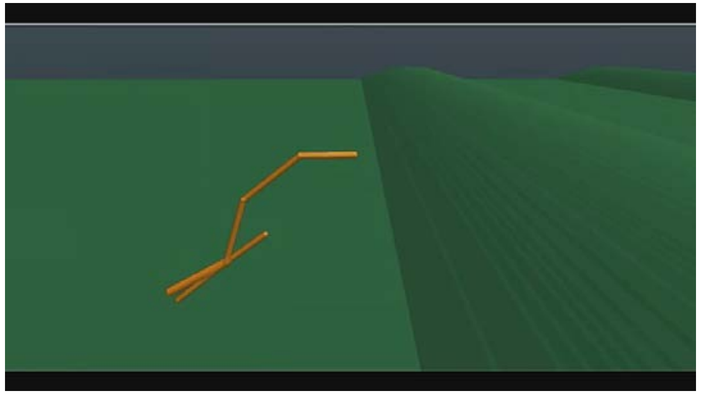


4 学习→形态：形态生成

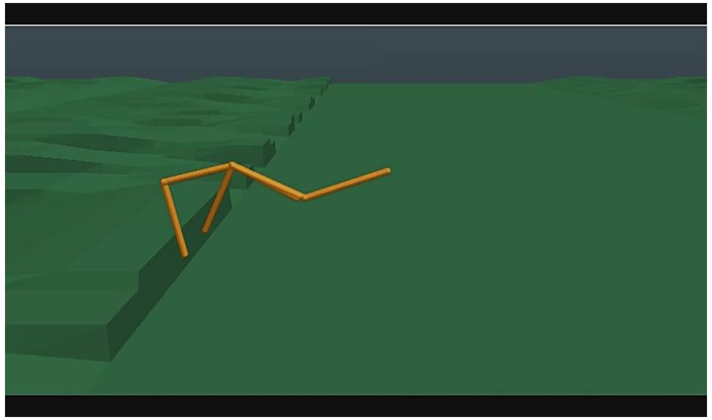
➢ 结构-参数的联合**持续**优化



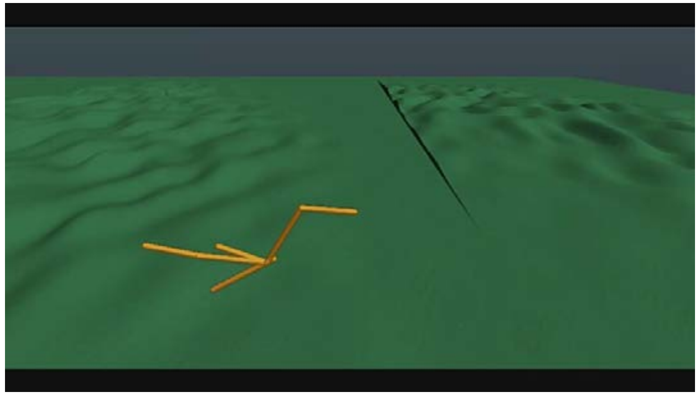
任务$_1$ 台阶



任务$_2$ 坡地

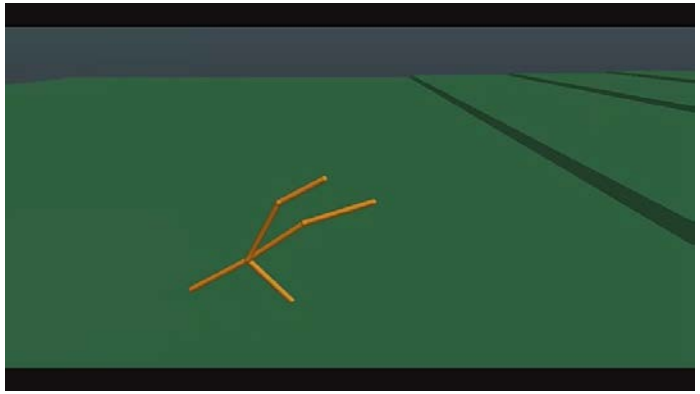


任务$_3$ 丘陵



任务$_4$ 沟壑

任务$_5$ 平地

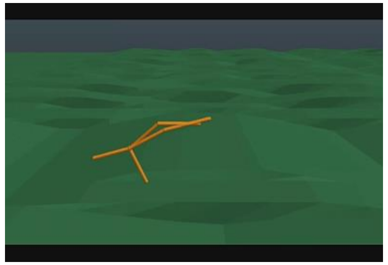


任务$_6$ 台阶+坡地

4 学习➜形态：形态生成

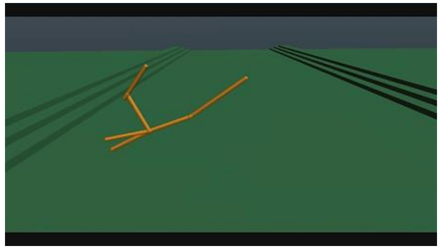
➢ 多环境**持续**优化过程中的智能体形态变化过程

任务顺序

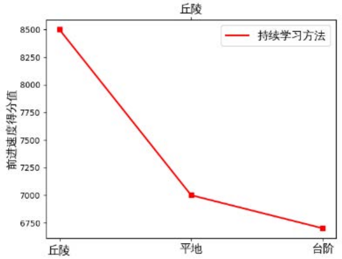


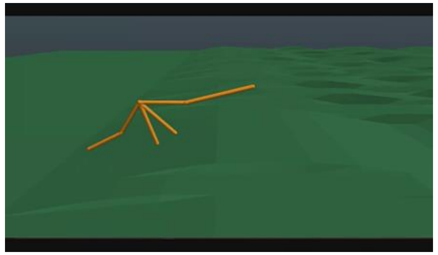
任务$_1$ 丘陵

任务$_2$ 平地

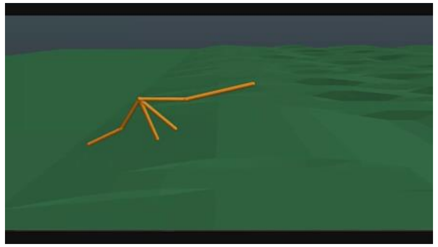


任务$_3$ 坡地

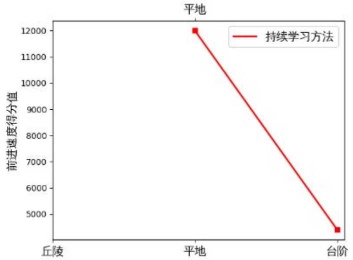




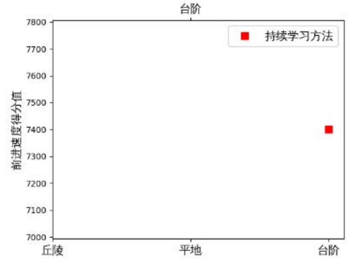
测丘陵



测丘陵



测平地

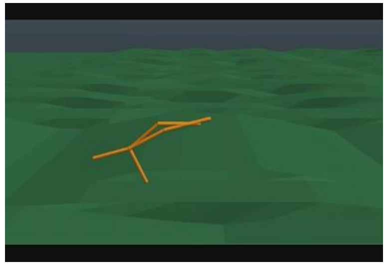


- 为了适应平地任务，智能体演化出了更加健壮的双腿；
- 而在台阶任务中，双腿则无需如此强壮。

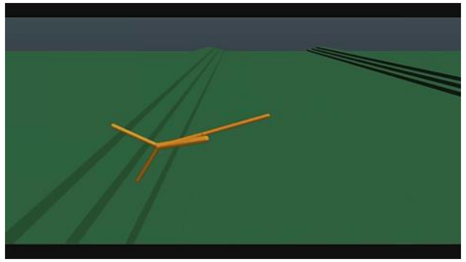
4 学习➜形态：形态生成

- 多环境**持续**优化过程中的智能体形态变化过程

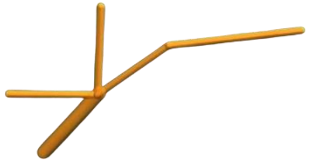
任务顺序

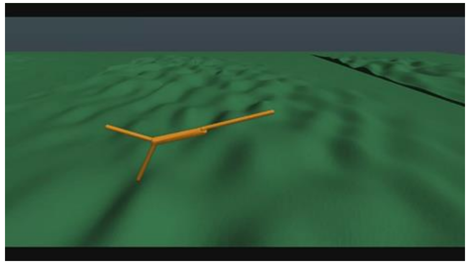


任务$_1$ 丘陵

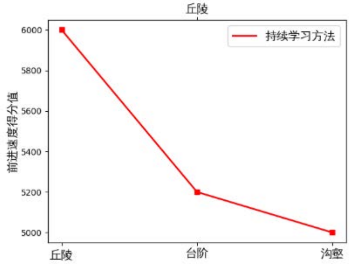


任务$_2$ 台阶

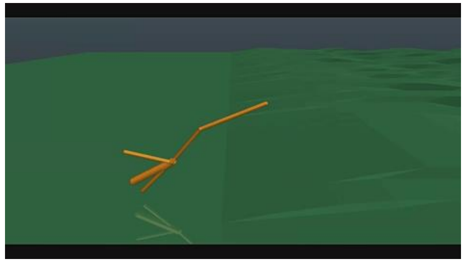




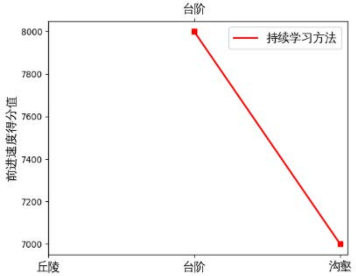
任务$_3$ 沟壑

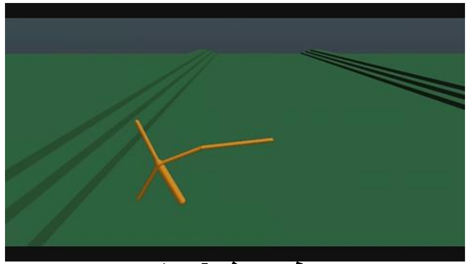


测丘陵

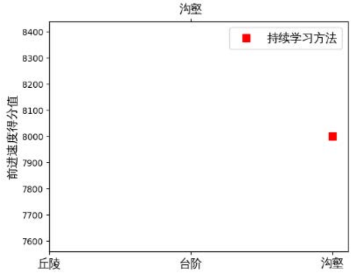


测丘陵





测台阶

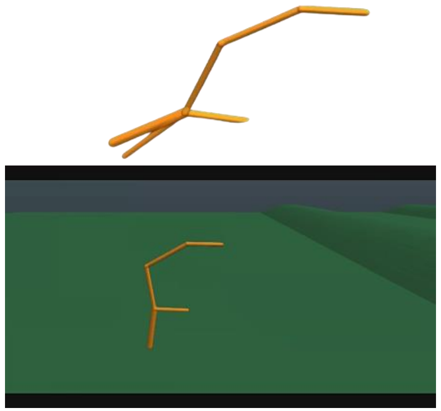


- 为了适应台阶任务，智能体退化了尾部功能；
- 而在沟壑任务中，它却增强了一条前腿的功能。

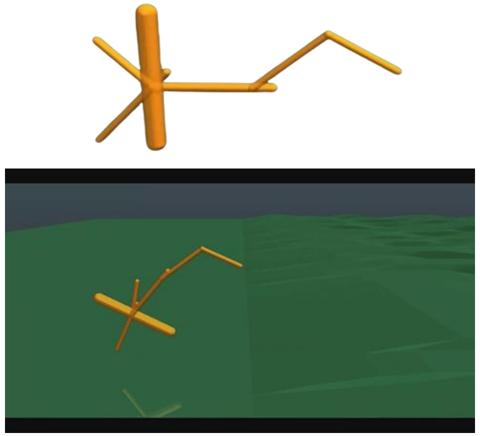
4 学习➜形态：形态生成

➢ 多环境**持续**优化过程中的智能体形态变化过程对应视频

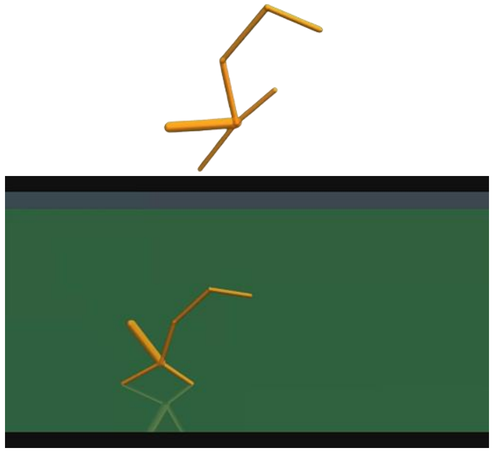
任务顺序



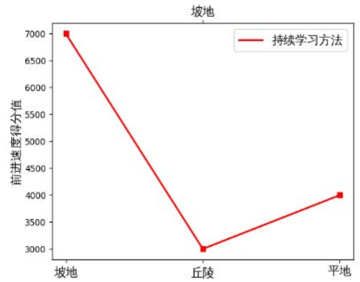
任务$_1$ 坡地



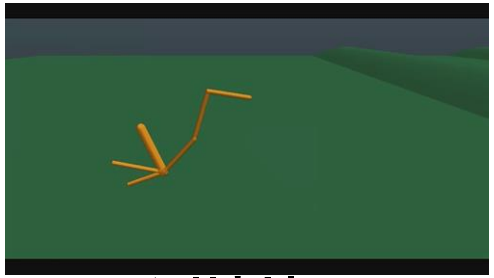
任务$_2$ 丘陵



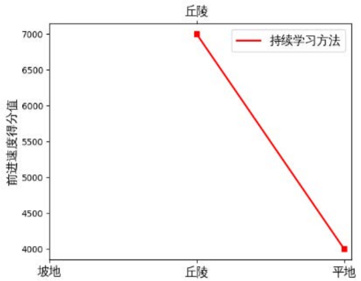
任务$_3$ 平地



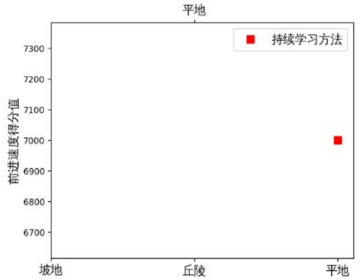
测坡地



测坡地



测丘陵



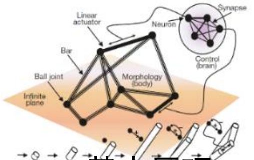
- 在丘陵任务中，智能体对其中一条前腿进行了强化并形成了对称结构；
- 而在平地环境中，则强化了其尾部以增强平衡性。

4 学习→形态：形态生成

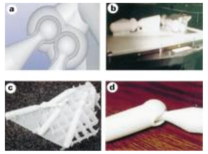
➢ 前沿研究

- 利用进化学习框架实现了虚拟环境下形态与控制器的协同优化。将三维刚性机器人的结构表示为有向图基因表示并利用图上的进化算法来优化机器人形态的设计
- 建立了可以自动生成三维虚拟生物的系统，并在物理世界中展现了不同的竞争策略
- 使用有向图作为描述形态和行为的遗传语言，定义了一个无限个可能结果的多维空间

- 利用可变长度圆柱形部件构建可物理可实现的进化机器人。
- 利用行走能力作为适应度函数的评估，在仿真环境中大约经历300-600代迭代即可。利用商业化的快速成型技术转换为物理系统。

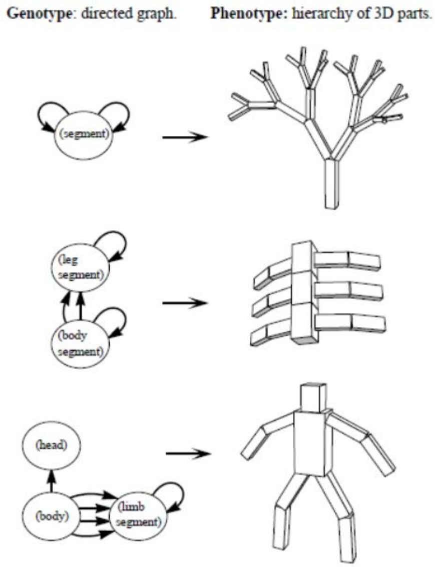


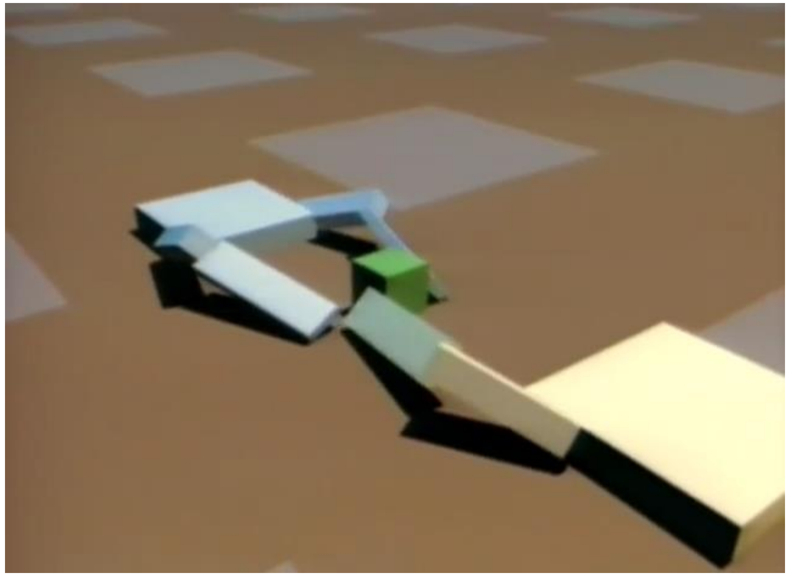
基本原理

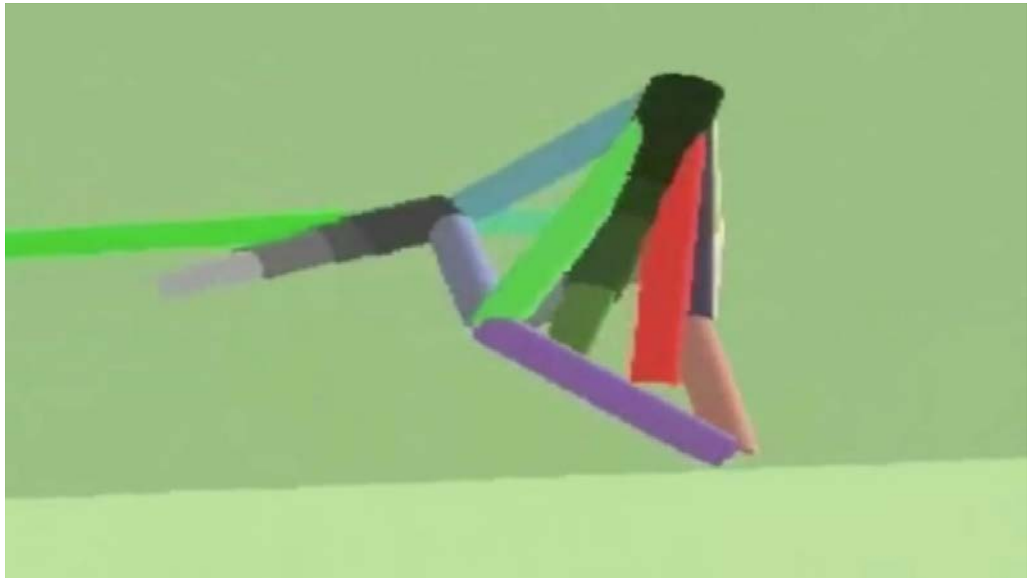


成型

三组形态







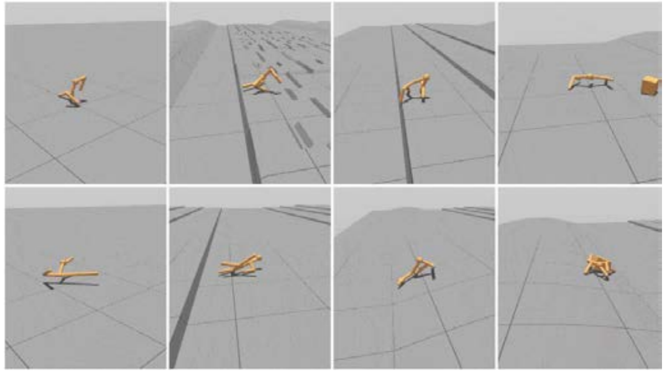
- Sims, Karl. "Evolving 3D morphology and behavior by competition." *Artificial life* 1.4 (1994): 353-372.

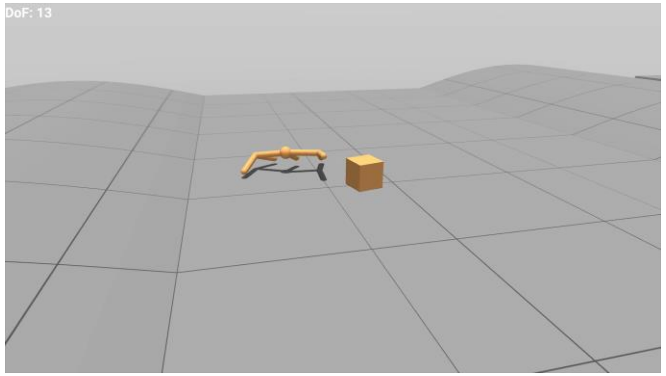
- Lipson H, Pollack J B. Automatic design and manufacture of robotic lifeforms[J]. Nature, 2000, 406(6799): 974-978.

4 学习➞形态：形态生成

➢ 前沿研究

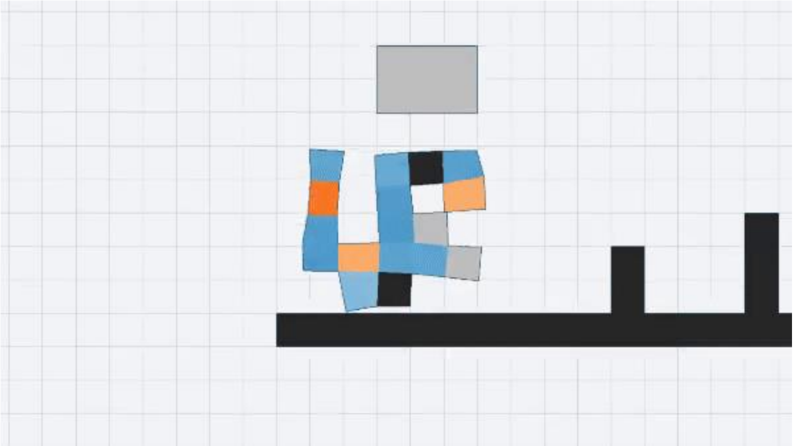
- “进化游乐场”的环境，探索人工智能中具身智能的发展及其与环境的关系，以及在计算机实验中的学习应用。
- 通过「形态学习」（morphological learning）证明了进化生物学中的「鲍德温效应」。在进化生物学中，鲍德温效应提出，在进化过程的早期世代一生中最学会的行为将逐渐成为本能，甚至可能遗传给后代。在过去的6亿年里，进化带来了无数形态的美：从古老的两侧对称的昆虫到各种各样的动物形态。这项研究不仅提出了一种新的计算框架，即深度进化强化学习（DERL），而且通过形态学习首次证明了达尔文-鲍德温效应。形态学习对于自然界中的动物进化至关重，现已在我们创建的 AI 智能体中展现





- Gupta A, Savarese S, Ganguli S, et al. Embodied intelligence via learning and evolution[J]. Nature communications, 2021, 12(1): 5721.

- 在不依赖人提供任何初始形态的前提下，算法可以自主根据任务需要，进化出适合任务的身体结构和动作，并且不断通过进化自己，让任务完成得越来越好。在上面案例中，当机器人学会“进化”自己的身体之后，灰色矩形块被扔得越来越远。
- Evolution Gym 专门为软体机器人而开发，涵盖 30 多个不同的任务环境，包括跑步、上台阶、攀爬、搬运物体等。Evolution Gym 中的机器人看起来像是柔软、可移动的俄罗斯方块，整体呈网格状结构，由许多个“细胞”作为基本单元组成，其中包括可以自由形变的软体细胞、坚硬的刚体细胞、以及可以主动收缩或扩张的致动器细胞。这种灵活的形态，使得机器人可以自由“进化”其形状，最终在不同地形上完成一系列运动和操纵物体等任务。只模拟了二维软体机器人



- Evolution Gym: A Large-Scale Benchmark for Evolving Soft Robots，NIPS，2021

4 学习➞形态：形态生成

➢ 前沿研究

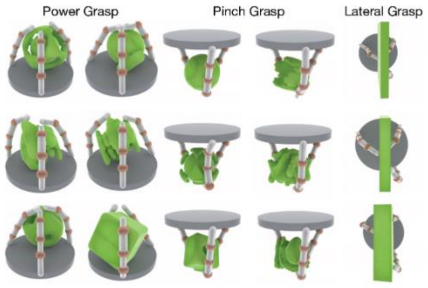
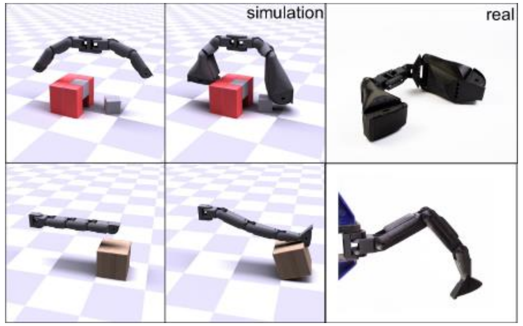
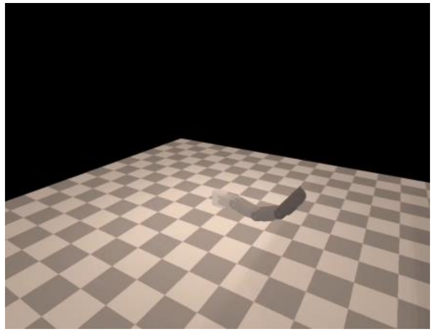
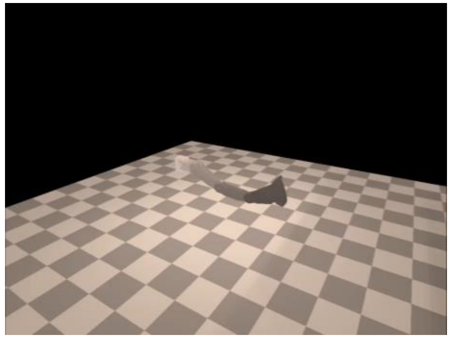
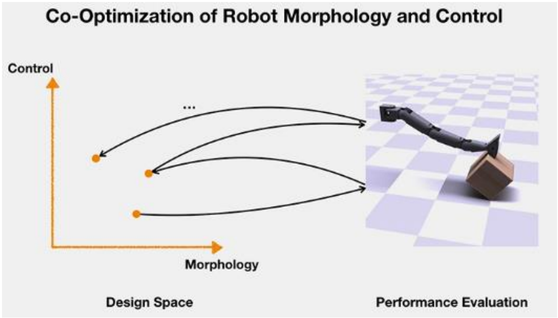


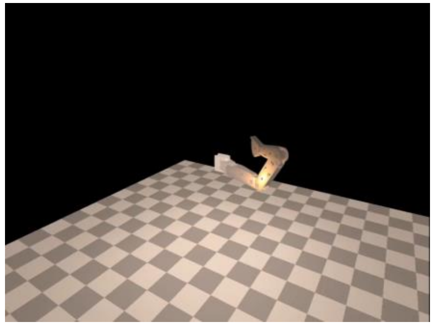
Fig. 1: **Emergent hand morphology and behavior.** We optimize hand design and control on grasping a diverse set of objects, aiming at maximizing the success rate and minimizing the hand complexity. Here are examples of the emergent hand design and grasping behavior from our model in three types of grasps. Our optimized hand is both effective and cost-efficient in its design.

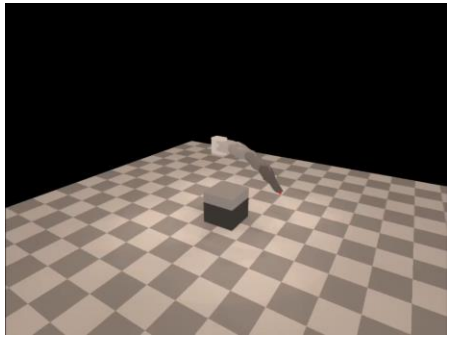








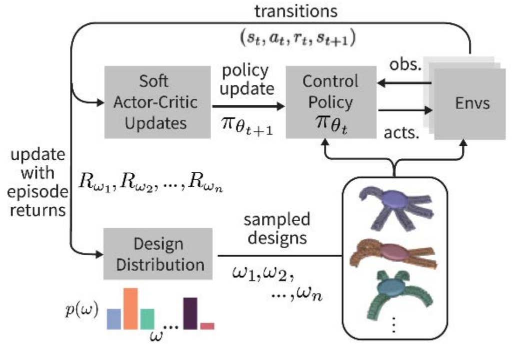




- Pan X, Garg A, Anandkumar A, et al. Emergent hand morphology and control from optimizing robust grasps of diverse objects[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 7540-7547.
- Jie Xu, Tao Chen, Lara Zlokapa, Michael Foshey, Wojciech Matusik, Shinjiro Sueda, Pulkit Agrawal，An End-to-End Differentiable Framework for Contact-Aware Robot Design，RSS 2021

4 学习➞形态：形态生成

➢ 前沿研究



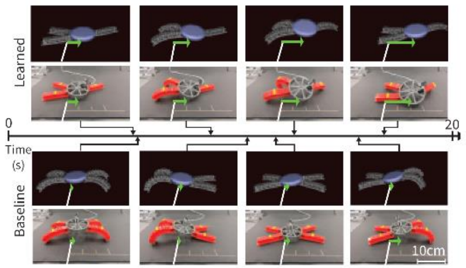
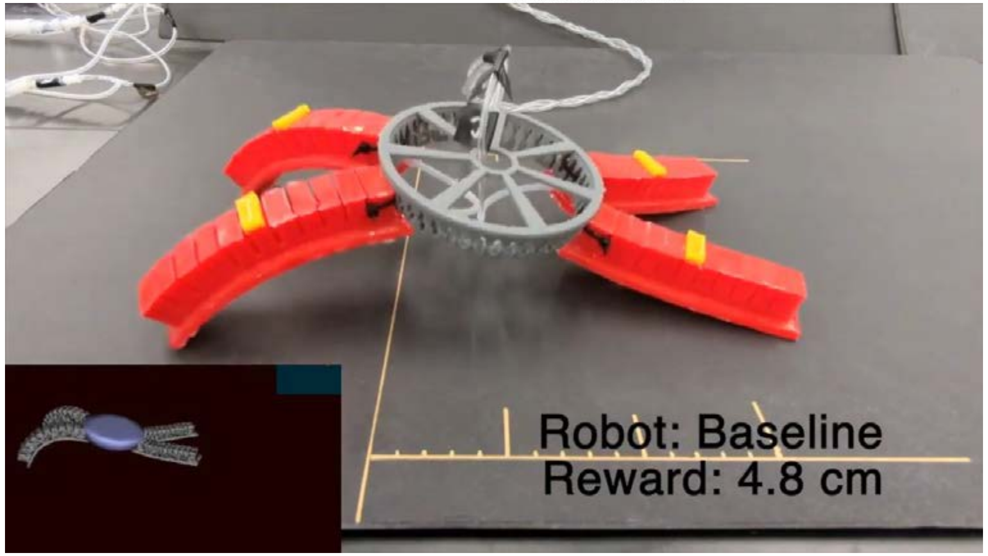
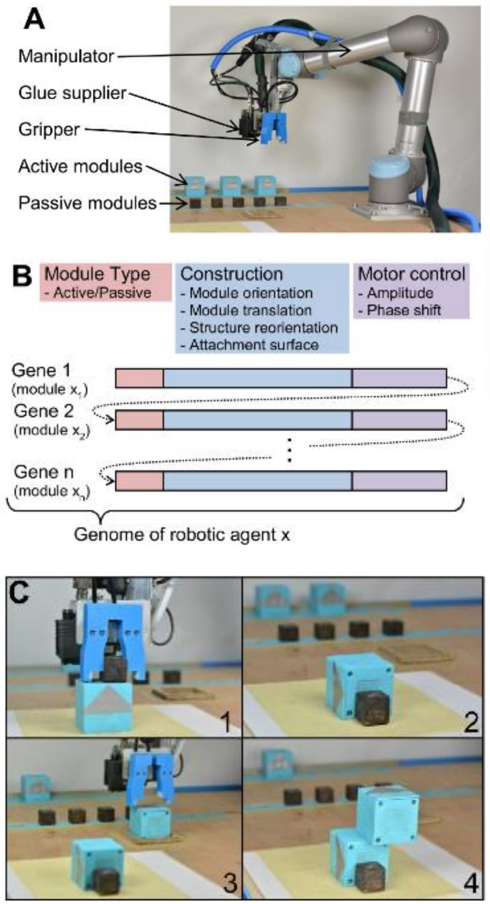
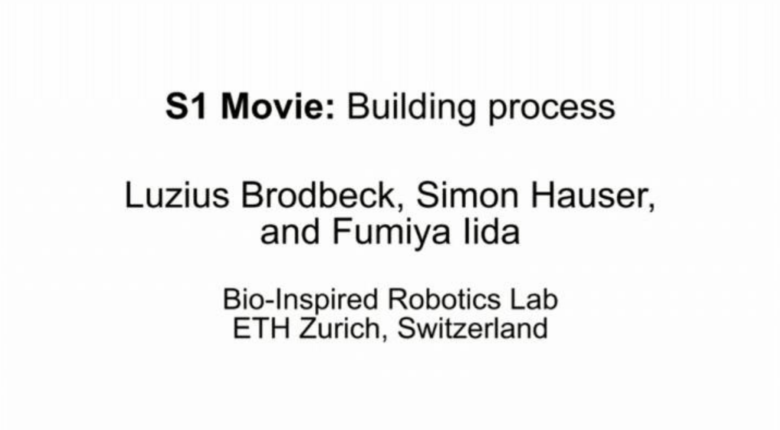
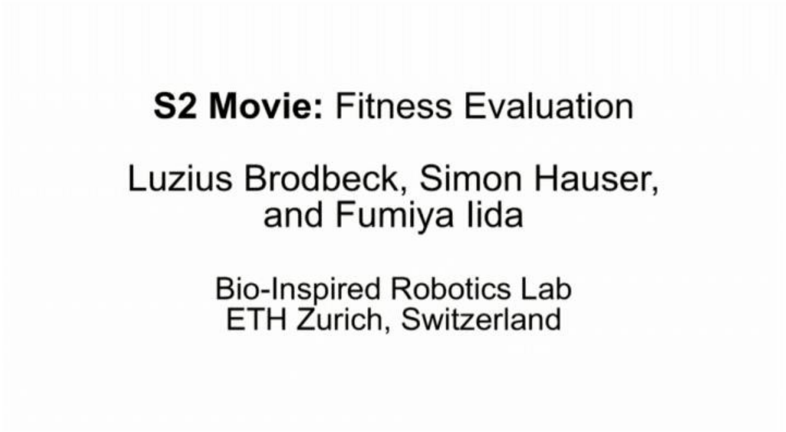


Fig. 1. Our framework jointly learns the design and control of crawling soft robots (top) that outperform an expert-designed baseline (bottom). While trained exclusively in simulation, our learned robots are capable of zero-shot sim-to-real transfer, with the optimal design moving more than 2× faster than the baseline in the real world.





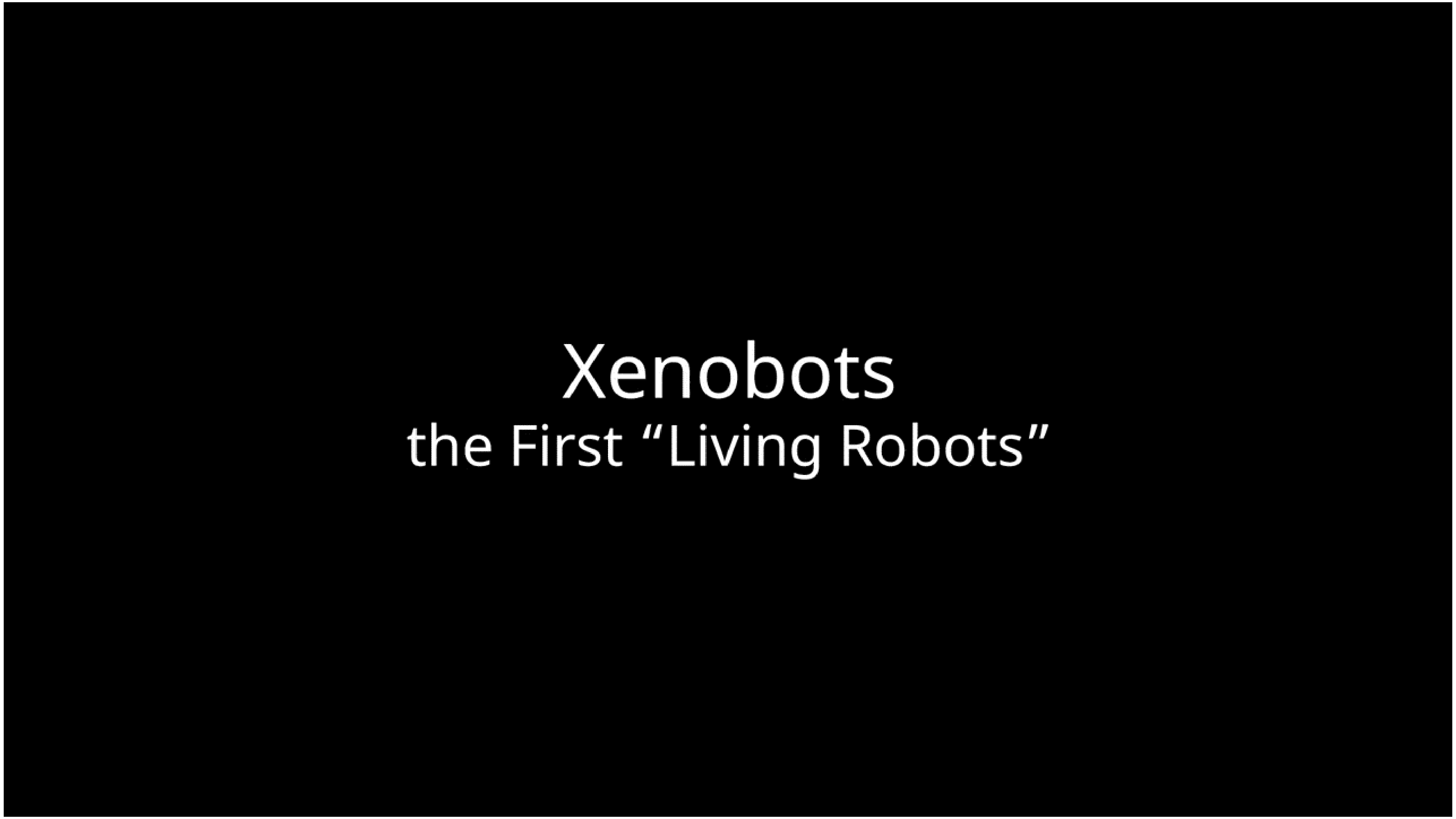




- Schaff C, Sedal A, Walter M R. Soft Robots Learn to Crawl: Jointly Optimizing Design and Control with Sim-to-Real Transfer[J]. arXiv preprint arXiv:2202.04575, 2022.
- Brodbeck, L., Hauser, S., Iida, F.: Morphological evolution of physical robots through model-free phenotype development. PLoS ONE 10(6) (2015) 1{17

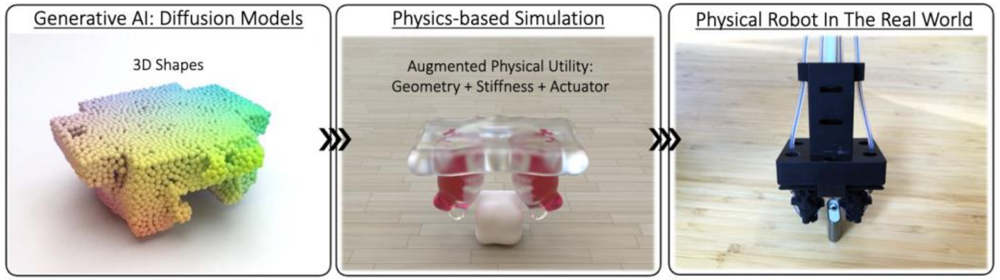
4 学习➝形态：形态生成

- **前沿研究**



4 学习➞形态：形态生成

➢ 前沿研究：DiffuseBot



Algorithm 2 Sampling: Diffusion As Co-design

Initialize: initial sample $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
while within maximal number of diffusion steps $t \geq 0$ **do**
  Perform regular per-step diffusion update: $\mathbf{x}_t \Leftarrow \mathbf{x}_{t+1}$.
  if perform co-design **then**
    while within K steps **do**
      Run update in (6) and (7) to overwrite $\mathbf{x}_t$.
    end while
  end if
end while

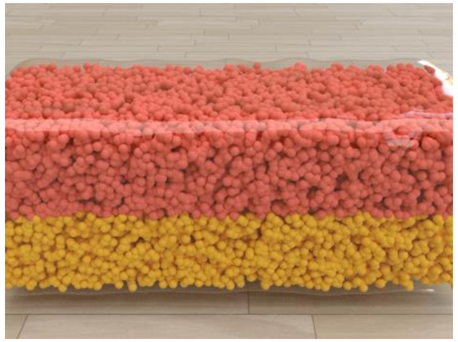
Balancing

Landing

Crawling

Hurdling

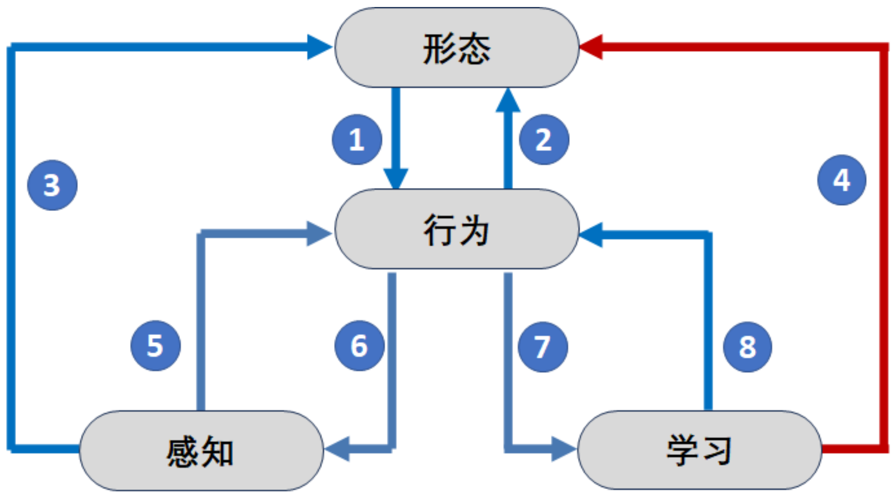
Gripping

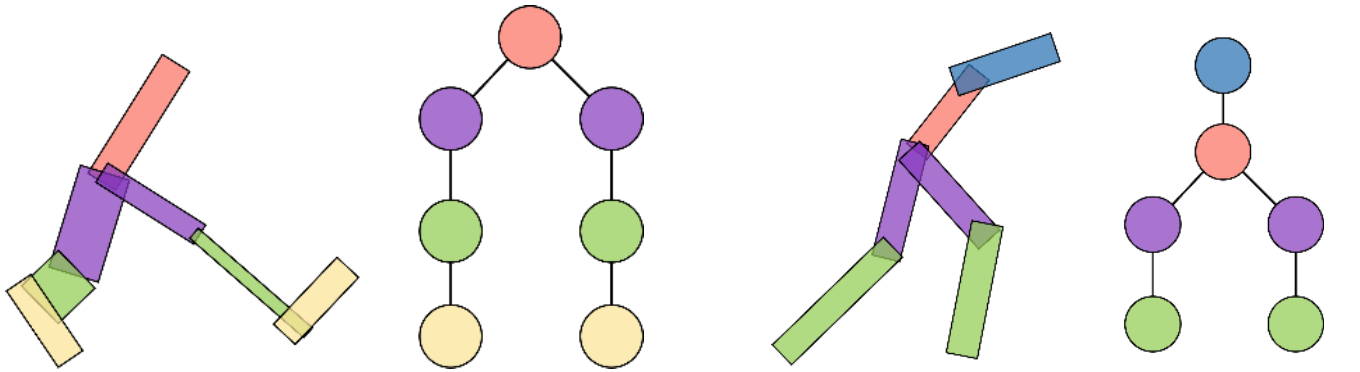


Box Moving

4 学习→形态：形态生成

➢ 小结





$$(D^*,\pi^*)=\arg\max_{D,\pi} J(D,\pi)$$

$$J(\theta_w)=\int R(w)p_{\theta_w}(w)dw$$

$$\theta_w^*=\arg\max_{\theta_w}\int R(w)p_{\theta_w}(w)dw$$

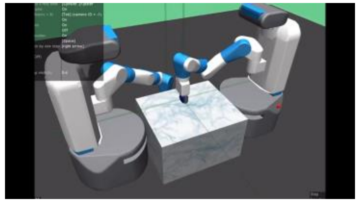
$$\nabla_{\theta_w}J(\theta_w)\approx\frac{1}{N}\sum_{i=1}^{N}R(w^i)\nabla_{\theta_w}\log p_{\theta_w}(w^i)$$

- 遗传强化学习，……

总结

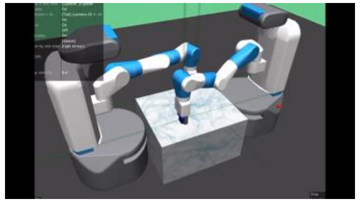
➢ 多体具身交互学习

基于竞争、对抗的多体交互学习方法



- “Compete”:

The protagonist competes with the adversary to get the block first and then put it to the goal position.

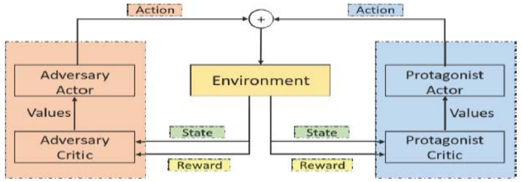


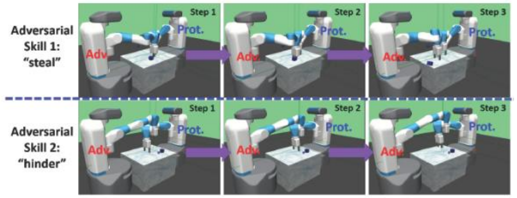
- “Push”:

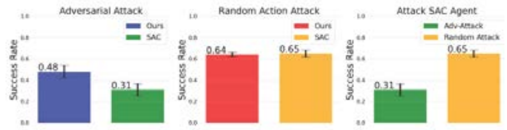
The protagonist pushes the adversary robot arm away to get the block and accomplish the task.

对抗行为学习

- 通过两个智能体之间的行为对抗（“偷走”与“抢回”），得到了鲁棒性更强的智能体行为策略。



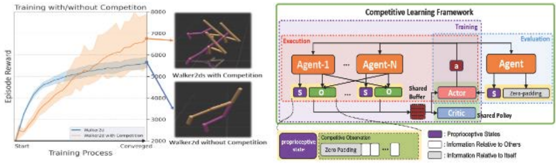


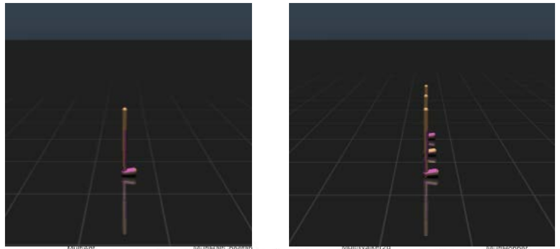


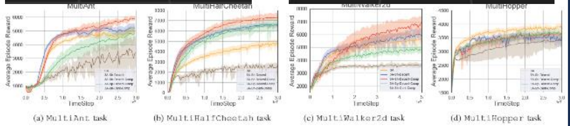
- Adversarial Skill Learning for Robust Manipulation, **ICRA**, 2021

竞争行为学习

- 通过在智能体独立训练过程中引入竞争信息，得到了比单一训练更好的结果，揭示了“竞争促进学习”的机制。



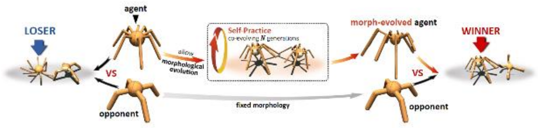


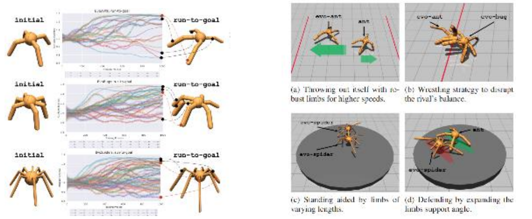


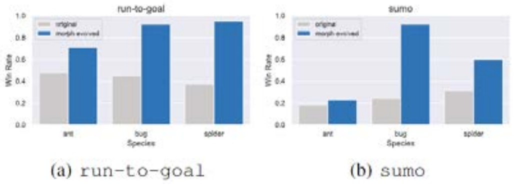
- Stimulate the Potential of Robots via Competition, **ICRA**, 2024
- Adversarial decision making against intelligent targets in cooperative multi-agent systems, **IEEE T-CDS**, 2023

对抗形态学习

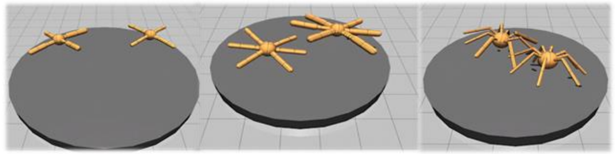
- 模拟人类身体发育机理，探索了智能体在对抗过程中可改变形态的方法，为形态优化提供了新的解决思路。



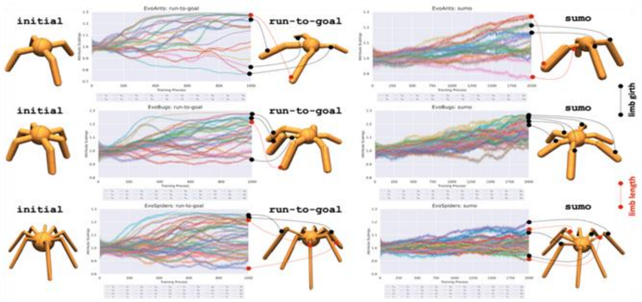




- CompetEvo: Towards Morphological Evolution from Competition, **IJCAI**, 2024

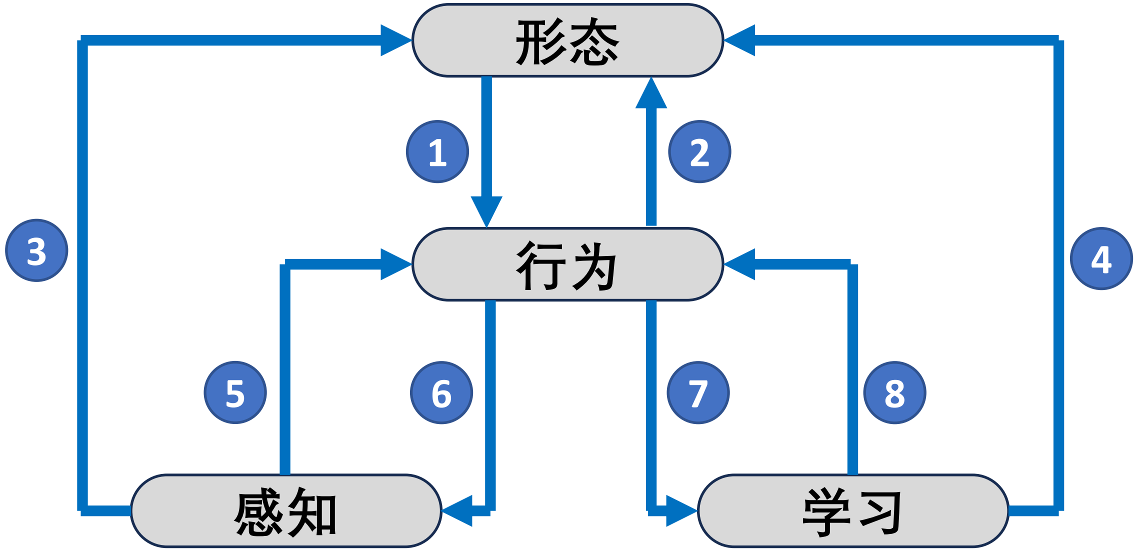


ant-evoant　bug-evobug　spider-evospider



总结

➢ 具身智能的体系结构

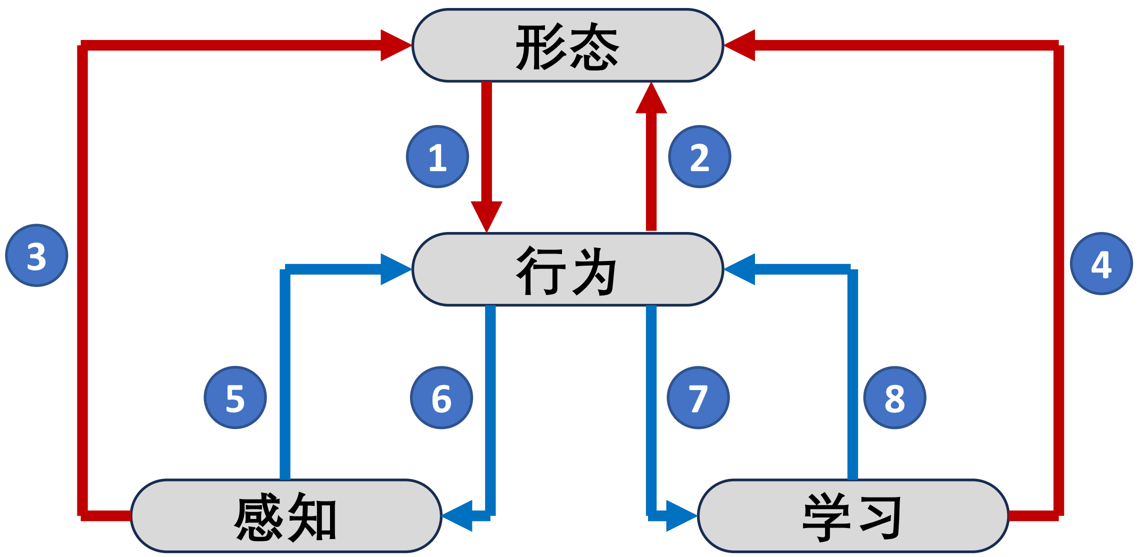


① 基于形态的行为生成 ③ 基于感知的形态变换 ⑤ 基于感知的行为生成 ⑦ 基于行为的自主学习

② 基于行为的形态控制 ④ 基于学习的形态优化 ⑥ 基于行为的主动感知 ⑧ 基于学习的行为优化

总结

➢ 具身智能的体系结构



① 基于形态的行为生成

② 基于行为的形态控制

③ 基于感知的形态变换

④ 基于学习的形态优化

①形态→行为

②行为→形态

③感知→形态

④学习→形态

第 49 卷 第 6 期　　自 动 化 学 报　　Vol. 49, No. 6

2023 年 6 月　　ACTA AUTOMATICA SINICA　　June, 2023

基于形态的具身智能研究: 历史回顾与前沿进展

刘华平[1]　郭 迪[2]　孙富春[1]　张新钰[3]

摘　要　具身智能强调智能受脑、身体与环境协同影响, 更侧重关注智能体与环境的"交互". 因此, 在具身智能的研究中, 智能体的物理形态与感知、学习、控制的关系起到至关重要的作用. 当前, 具身智能综合吸收了机构学领域关于形态、结构, 机器学习领域关于感知、学习, 以及机器人领域关于行为、控制等的相关研究成果, 形成了相对完整、独立并仍在蓬勃发展的学科分支. 但是, 目前尚无文献完整地梳理基于形态的具身智能研究进展. 本文从这个角度出发, 重点围绕基于形态计算的行为生成、基于学习的形态控制, 以及基于学习的形态优化这三方面总结重要的研究进展, 凝炼相关的科学问题, 并总结未来的发展方向, 可为具身智能的研究提供参考.

关键词　具身智能, 形态智能, 形态计算, 形态控制, 形态–控制协同优化

引用格式　刘华平, 郭迪, 孙富春, 张新钰. 基于形态的具身智能研究: 历史回顾与前沿进展. 自动化学报, 2023, 49(6): 1131−1154

DOI　10.16383/j.aas.c220564

Morphology-based Embodied Intelligence: Historical Retrospect and Research Progress

LIU Hua-Ping[1]　GUO Di[2]　SUN Fu-Chun[1]　ZHANG Xin-Yu[3]

Abstract　Embodied intelligence emphasizes that the intelligence is influenced by the interaction among brain, body and environment. It is more focused on the interaction between the agent and environment. Therefore, the relationship between the physical morphology and perception, learning, and control of the intelligent agent plays a vital role in the research of embodied intelligence. Currently, embodied intelligence comprehensively utilizes the research results from the communities of mechanism, machine learning, and robotics to form a new ever-growing branch. However, there is still no complete survey to summarize the research progress of the morphology-based embodied intelligence. In this paper, we focus on the aspects of morphology computation based behavior generation, learning based morphology control, and learning based morphology optimization to summarize important research progress, scientific issues and future development directions, which can provide reference for the embodied intelligence research.

Key words　Embodied intelligence, morphology intelligence, morphology computation, morphology control, morphology-control collaboration optimization

Citation　Liu Hua-Ping, Guo Di, Sun Fu-Chun, Zhang Xin-Yu. Morphology-based embodied intelligence: Historical retrospect and research progress. *Acta Automatica Sinica*, 2023, 49(6): 1131−1154

现代人工智能起源于上世纪五十年代的达特茅斯会议. 在此之后的一段时期内, 对人工智能的研究主要限于符号处理范式 (也被称为符号主义). 然而, 符号主义的局限性很快在实际应用中暴露出来, 并催动了联结主义的发展, 形成了包括多层感知机、前向神经网络、循环神经网络, 直至今日风靡学术界与产业界的深度神经网络等多种方法. 这种用人工神经网络模拟认知过程的方法在适应、泛化与学习方面的确取得了很大的进展, 但并未真正解决智能体与真实物理世界交互的难题, 在可解释性、鲁棒性等方面也面临很大的挑战[1]. 事实上, 关于符号主义与联结主义所存在的问题的讨论在上世纪七、八十年代就引起了极大的关注. 汉斯-莫拉维克 (Hans Moravec)、罗德尼-布鲁克斯 (Rodney Brooks) 与马文·明斯基 (Marvin Minsky) 等提出的"莫拉维克悖论" (一般通俗地表述为: 要让电脑如成人般地下棋是相对容易的, 但是要让电脑有如一岁小孩般的感知和行动能力却是相当困难甚至是不可能的) 就体现了很多学者的担忧. 针对这一问题, 明斯基从行为学习的角度提出了"强化学习"的概念. 布鲁

收稿日期 2022-07-09　录用日期 2023-01-16

Manuscript received July 9, 2022; accepted January 16, 2023

国家自然科学基金 (62025304, 62273054) 资助

Supported by National Natural Science Foundation of China (62025304, 62273054)

本文责任编委 金耀初

Recommended by Associate Editor JIN Yao-Chu

1. 清华大学计算机科学与技术系 北京 100084　2. 北京邮电大学人工智能学院 北京 100876　3. 清华大学车辆与运载学院 北京 100084

1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084　2. School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876　3. School of Vehicle and Mobility, Tsinghua University, Beijing 100084