
Bcent框架：让机器人告别“木偶式执行”，走向“感知-认知-行为”的闭环协同

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/39932.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

Bcent框架：让机器人告别“木偶式执行”，走向“感知-认知-行为”的闭环协同。 本篇论文将机器人等智能体建模为由感知体（P-body）、认知体（C-body）和行为体（B-body）组成的闭环系统，强调主动感知、认知学习与行为控制的协同。该机制使机器人能够在交互过程中不断修正自己的内部模型，提升感知与行为协调能力，其有效性在 D’Kitty 的导航任务中得到了验证。

清华大学孙富春教授团队联合中国人民大学、汉堡大学研究人员，在中国工程院院刊Engineering发表题为Robot Cognitive Learning by Considering Physical Properties（融合物理属性的机器人认知学习）的研究论文，提出面向机器人认知学习的Bcent（Brain – Body Co-developmENT）框架。论文的核心不是把机器人简单理解为观察到动作的单向映射，而是把感知、认知与行为重新组织为一个持续交互、知识持续更新的闭环系统。

长期以来，许多机器人系统更像是在执行一条预先写好的控制链：传感器给出感知输入，策略模块通过执行器输出动作，系统是否能够适应更复杂的环境，很大程度取决于设计者事先写进去了多少规则。这样的组织方式在局部任务中可以奏效，但一旦环境变化、观测受扰或任务链条拉长，系统就容易暴露出局部有效、整体不足的问题。

Bcent的意义，正是在于它不再满足于让机器人照着做，而是试图回答机器人如何在感知、规划和执行之间形成更稳定的信息回路。如图1所示，论文将机器人等智能体表述为由P-body、C-body、B-body组成的三体协同系统：P-body负责获取并调整观察，C-body负责生成和更新中长期子目标，B-body负责面向子目标执行动作并把结果反馈回来。这样的结构，使机器人从线性流程转向闭环协同。

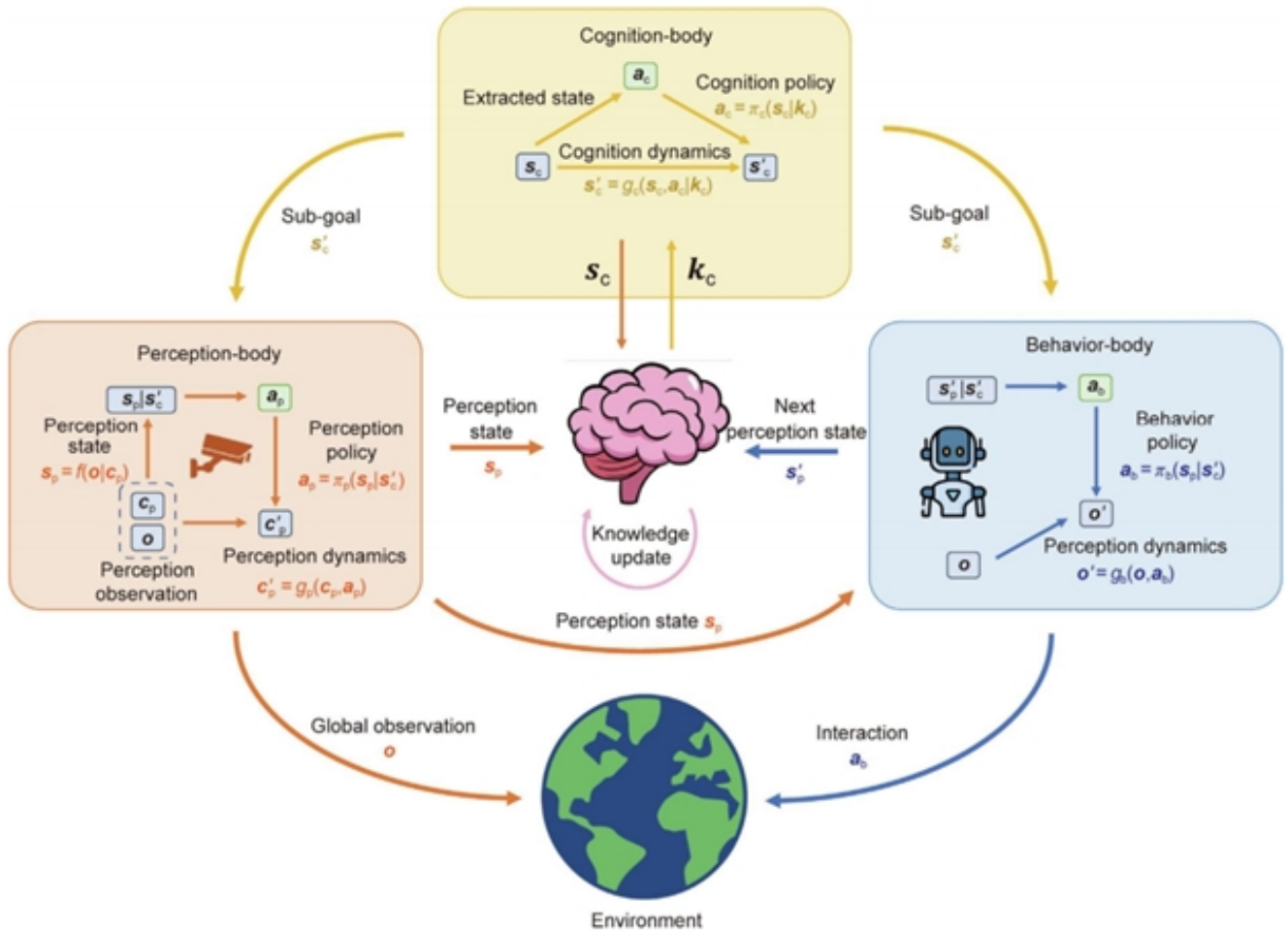


图1.Bcent框架的感知、认知、行为三体协同架构。

Bcent框架的系统论构建

在学术层面，Bcent框架可以被理解为对传统感知-决策-执行线性链条的一次系统化重构。论文并不是简单增加一个新模块，而是重新界定系统内部各部分的职责：P-body对环境进行主动观察并输出低维状态，C-body在抽象状态空间中进行知识利用和知识更新，B-body则在子目标引导下完成与环境的具体交互。

这种重构的关键，在于引入了认知体这一中间层。它既不是单纯的控制器，也不是孤立的知识库存储器，而是把历史经验、当前状态和后续子目标联结起来的调度中枢。于是，机器人不再只是看到什么就做什么，而是能够围绕任务目标逐步组织感知和行为。

如图2所示，论文把知识库设计为可初始化、可调用、可更新的结构：既可以来自专家示范或人工编码，也可以在后续交互中不断修正。换句话说，Bcent真正想建立的，不是一条更长的控制链，而是一种更合理的系统组织方式。

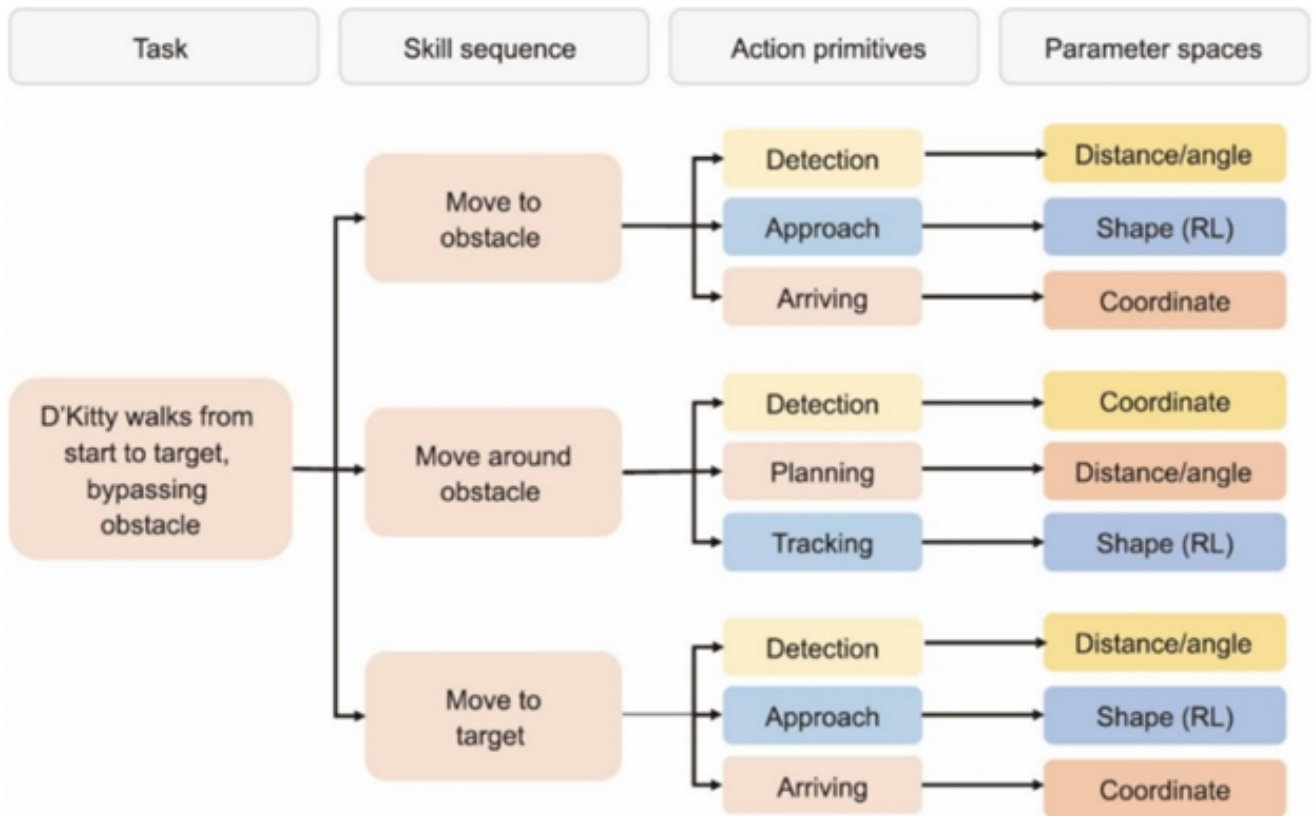


图2.Bcent框架中知识库的分类设计与构建流程。

三体协同的反馈控制律

框架的核心创新，体现在三体之间形成了明确的闭环交互关系，而不仅仅是数据在模块之间的单向传递。论文中，P-body、C-body、B-body都有各自的状态、策略与动力学，并通过子目标和状态残差彼此耦合。图3给出的，正是这种知识应用与部署的整体流程。

主动感知逻辑：在Bcent中，P-body不再局限于被动接收传感器输入，而是会根据C-body提供的子目标主动调整观测配置。以实验中的可移动相机为例，它可以改变位置和视角，以便更稳定地跟踪D' Kitty的状态。这一设计使感知不再只是前端输入，而成为服务于任务推进的主动过程。

知识利用与更新：C-body一方面基于当前状态和知识库生成子目标序列，引导后续感知与行为；另一方面又会利用真实交互带回的状态偏差，持续修正自身的认知模型。论文强调，C-body的时间尺度可以比P-body、B-body更长，因此它更适合承担中长期规划和知识更新的角色。

任务分解与行为学习：在B-body一侧，原本复杂的长程任务被拆解为多个由子目标驱动的短程子任务。这样一来，行为策略不需要在一开始就直接面对完整问题，而是能够在更明确的阶段性目标下逐步学习控制动作。这种认知给方向、行为去落实的分工，使整体训练更有结构。

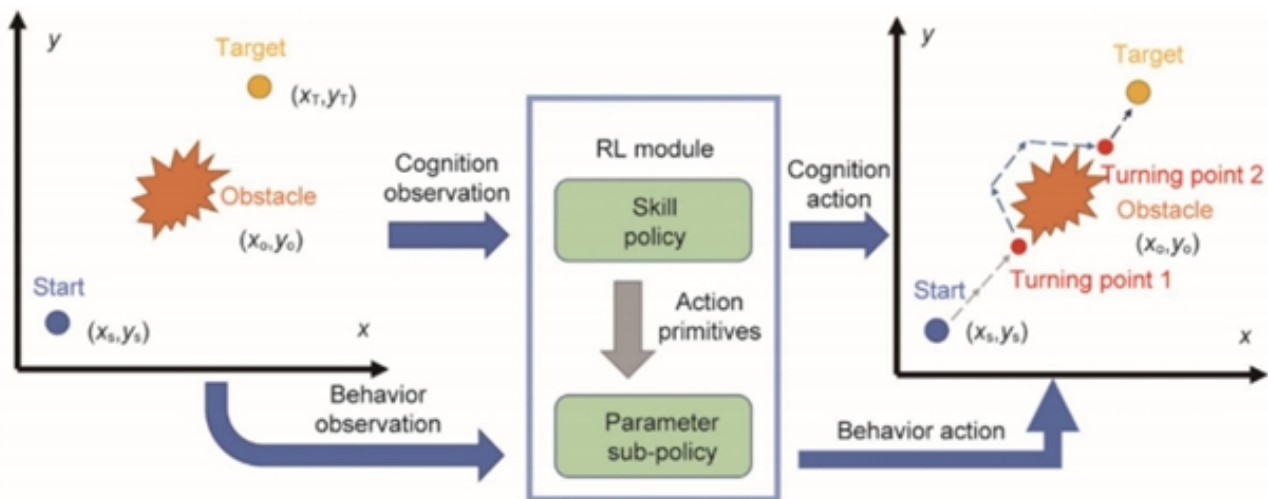


图3.知识的应用与部署。

长程任务的推理与决策

为了让机器人不仅能做出即时反应，还能面向更长时域组织行为，Bcent在C-body中显式建模了认知动力学。论文的做法不是把规划理解为抽象口号，而是让C-body在低维状态空间中对后续状态进行预测，并据此生成一串可执行的子目标。

图4是D' Kitty的运动与导航任务实现示意。在这一任务中，机器人需要从起点移动到目标点，同时处理障碍物、身体姿态和运动控制等问题。C-body的作用，就是根据当前感知状态和知识模型，对后续路径作出前瞻性的组织，再把阶段性目标分发给P-body与B-body。

因此，Bcent的长程决策并不意味着机器人已经具备了泛化到一切场景的完备推理能力，而是说明在当前任务中，系统可以通过子目标序列把全局任务拆解为一连串更容易执行和更新的局部过程。这一步，是从即时控制走向层级规划的重要组织方式。

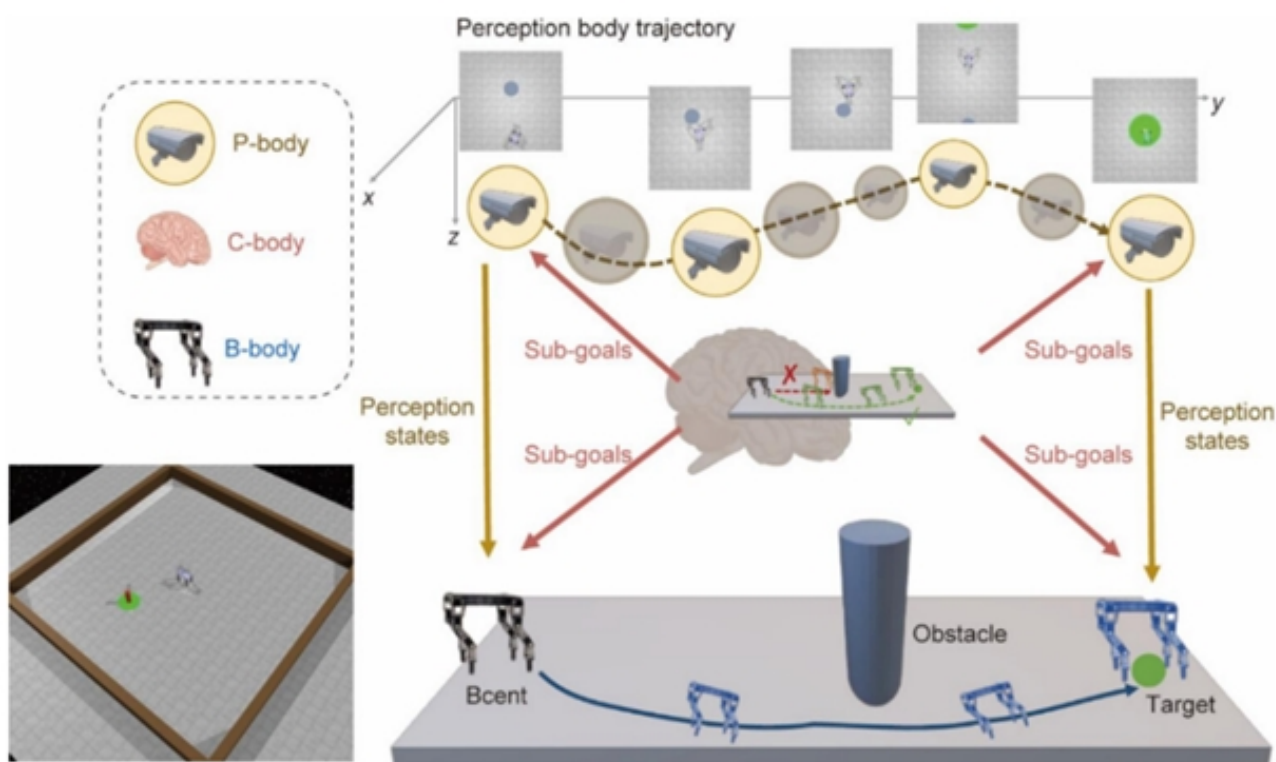


图4.DKitty运动与导航任务的实现。

DKitty机器人的实证表现

论文选择四足机器人DKitty的运动与导航任务作为验证场景：上方相机作为P-body，负责主动感知与目标跟踪；规划模块作为C-body，负责给出子目标轨迹；四足机器人本体作为B-body，负责站立、行走与避障。这一任务天然需要感知、规划与控制三者紧密协同，因此很适合检验Bcent的闭环设计。

从感知实验看，完整方法Policy + sub-goal在表1中取得了最高的mAP-0.5 指标（0.597），明显优于固定相机、随机移动相机以及不带子目标引导的对照设置。这意味着C-body提供的子目标不仅服务于后续行为控制，也实质性改善了主动感知阶段对机器人状态的跟踪质量。

图5展示了C-body与B-body的网络结构关系：C-body预测下一阶段子目标，B-body在环境中执行动作并在时间间隔 Δt 后反馈到达状态，后者再成为C-body更新认知模型的监督信号。这个结构说明，论文想强调的不是单个策略网络有多强，而是不同身体之间如何通过反馈构成可持续更新的闭环。

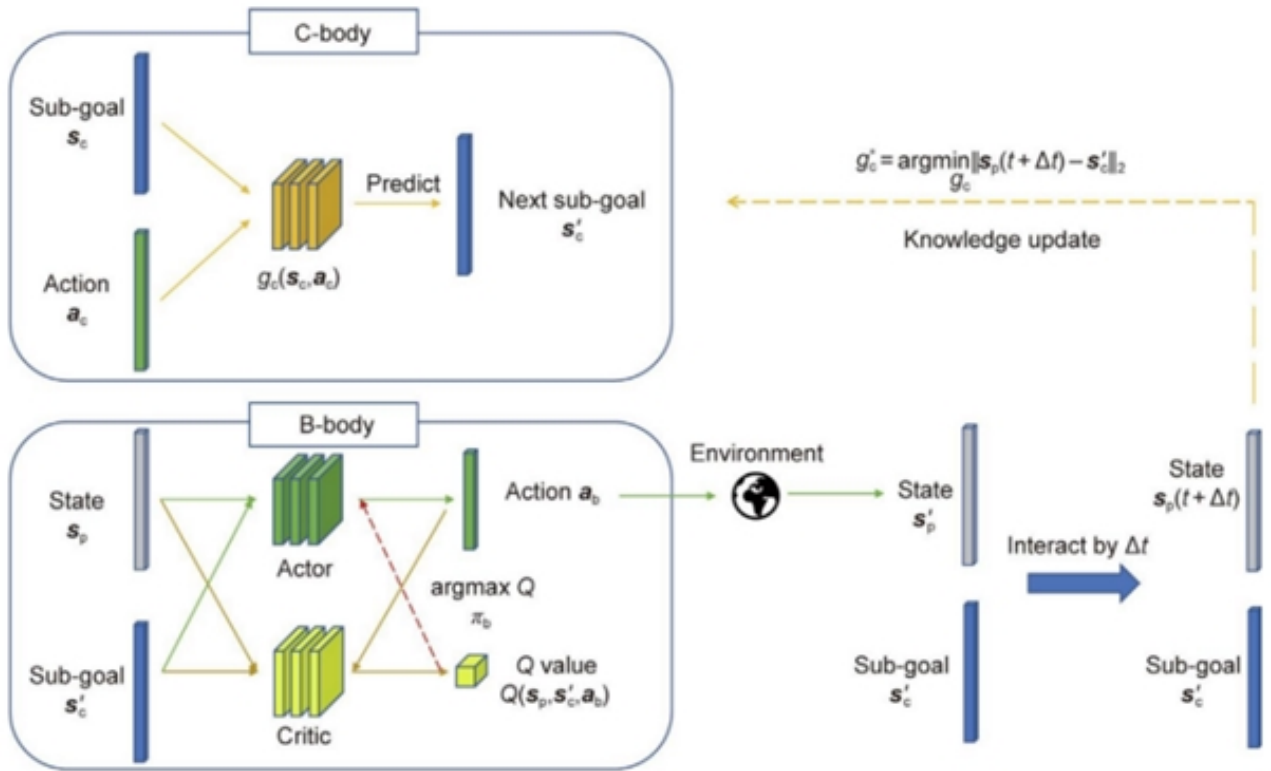


图5.C-body与B-body的网络架构。

在运动与导航实验中，Bcent框架在奖励曲线和成功率上均优于PPO、SAC、TDM等基线方法。更重要的是，当系统使用由P-body预测得到的状态而非地面真实状态时，整体性能虽然有所下降，但仍然显著优于基线。这表明，该框架并不依赖理想化的完美感知，而是在更接近真实应用条件下依然保持了较好的协同效果。

图6所示的奖励曲线，直观展示了这一点：与仅依赖单一行为学习的对照方法相比，PCB-body（P-body、C-body、B-body）在训练过程中获得了更稳定、更高的回报水平。这也从实验层面支撑了论文的核心判断—机器人能力的提升，不只取决于更强的单点模块，也取决于系统内部是否建立了有效的信息回路。

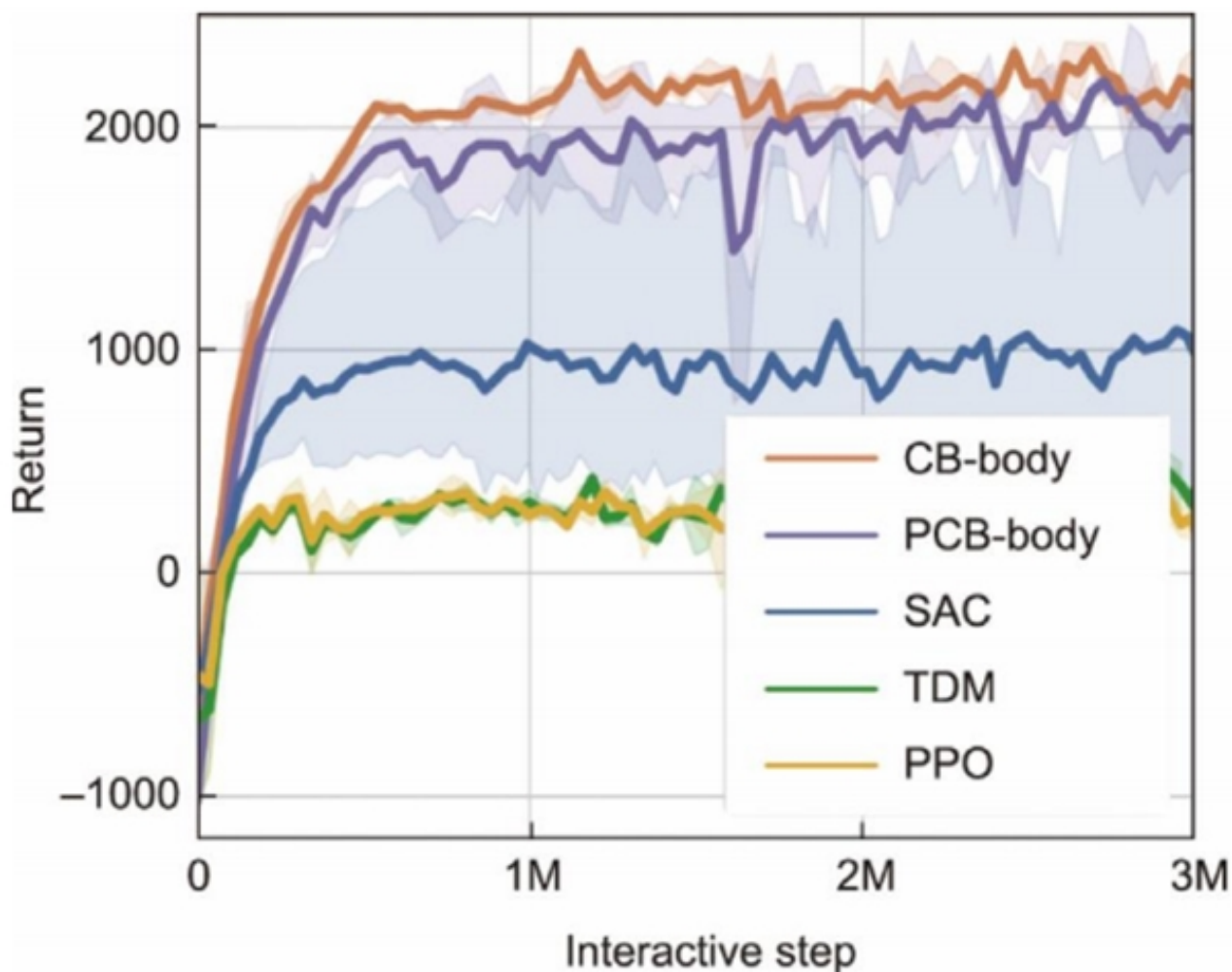


图6.运动与导航任务的奖励（Reward）曲线对比。

总结

这项工作的价值，并不在于宣称机器人已经理解了全部物理世界，而在于提出了一种更合理的机器人系统组织方式：把主动感知、认知规划、知识更新和行为控制放在同一个闭环中，让机器人能够在交互过程中不断修正自己的内部模型，并以子目标为桥梁协调不同层级的决策。

对于具身智能研究而言，这一思路具有启发性。它提醒我们，机器人能力的提升，未必只依赖更大的模型或更多的数据，也可能来自系统结构本身的重构——尤其是在感知、认知和行为之间建立真正有效的反馈回路。Bcent提供的，正是这样一种从单向执行走向闭环协同的方法视角。

从传播角度看，这篇论文最值得强调的，不是把机器人描绘成拥有灵魂的新物种，而是更准确地指出：机器人也需要一种更接近真实任务需求的内部组织框架。在这个框架里，机器人不再只是被动接收输入、机械输出动作，而是在感知、认知与行为的相互作用中逐步形成更稳健的任务能力。这也是Bcent对具身智能研究最有分量的贡献。（来源：EngineeringJournals微信公众号）

相关论文信息：<https://www.sciencedirect.com/science/article/pii/S2095809924006441?via%3Dihub>

特别声明：本文转载仅仅是出于传播信息的需要，并不意味着代表本网站观点或证实其内容的真实性；如其他媒体、网站或个人从本网站转载使用，须保留本网站注明的“来源”，并自负版权等法律责任；作者如果不希望被转载或者联系转载稿费事宜，请与我们联系。

作者：孙富春等 来源：《工程》

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发