最优　　主动粒子的最优导航与机器学习相遇

Mahdi Nasiri[1]，　Hartmut Löwen[2] 和 Benno Liebchen[1] (a)

[1] 凝聚态物理研究所，达姆施塔特工业大学 - *D-64289* 达姆施塔特，德国[2] 杜塞尔多夫海因里希·海涅大学理论物理 *II*：软物质研究所，*D-40225* 杜塞尔多夫，德国

arXiv:2303.05558v1 [cond-mat.soft] 9 Mar 2023

pacs 82.70.Dd – 胶体pacs 05.40.Jc– 布朗运动

摘要 – 如何让“聪明”的主动智能体（如昆虫、微生物或未来的胶体机器人）在复杂环境中灵活行动，从而以最优方式抵达或发现目标（例如气味源、食物或癌细胞），近年来引起了广泛关注。本文概述了从微观到宏观尺度上此类最优导航问题的最新进展，并通过讨论我们面临的一些挑战展望未来。除了举例说明解决最优导航问题的一种基础方法外，本文还重点介绍了采用基于机器学习的方法的研究。即使面对涉及混沌、高维或未知环境的问题，这类基于学习的方法也能发现高效的导航策略；而这些问题几乎无法通过传统解析方法或仿真方法求解。

引言。– 飞机起飞前，
c驾驶员通常会运行一个软件，以规划前往目标
d目的地的路径。该路径旨在代表最佳的折
s中方案，兼顾航行时间、燃油消耗
t导航，以及针对特定环境条件的安全因素，
t例如当前的风况和温度分布，
a以及空域占用。类似的最优导航
p问题也会出现在这样的情形中：一个主动（自推进）智能体能够
c在复杂环境中控制其运动方向或速度，
p这类情形也会出现在机器人和自主
d驾驶汽车。这些智能体通常配备
c摄像头、处理单元和执行器 [1–3], ，使其能够
t在本地感知周围环境，并利用这些信息
t实现目标，例如尽可能快速地跑完赛道，
p或高效地收集废弃物。在这里，摄像头
t通常会生成环境图像，并将其输入
i
送至处理单元，随后处理单元将接收到的信息转化为所需的行动（运动指令），并发送给执行器，由执行器启动运动。在动物界，发展高效导航策略的能力有助于物种生存。例如，为了抵达繁殖地，一些海龟必须在海洋中跋涉数百公里，找到一条高效路径 [4, 5]。同样，昆虫也需要通过复杂的分子环境进行导航，以找到气味源。

(a)电子邮箱： benno.liebchen@pkm.tu-darmstadt.de

由于湍流气流，lar 模式甚至可能呈现混沌 [6–9]。

智能微型游动体——除了宏观尺度的智能体，微生物也能感知来自其环境的信息，并将其用于导航。例如，精子细胞感知卵细胞释放的化学物质浓度梯度的能力 [10,11] 对许多物种的生存至关重要。同样，细菌拥有种类丰富的生化传感器，使其能够测量氧气、营养物质或自诱导物浓度的梯度 [12–14] ，并将其用于导航。

除了生物微型游动体，近二十年来 [15]合成微型游动体也已问世，目前正与生物微型游动体一同在活性物质研究领域中受到研究 [16–35]。合成微型游动体可以由外场 [36–53] （甚至由反馈控制系统 [54–56]）操控，并能通过多种形式的趋性 [57, 58], 对环境作出反应。未来可利用这些特性，帮助它们在血管中穿行，以检测并可能修复突变细胞 [59, 60], 、将药物运送至癌细胞 [61–63] ，或进行显微外科手术 [64]。

尽管宏观尺度上的最优导航问题已被研究数十年，[65,66], 采用的方法包括最优控制理论和动态规划 [67–70],，近来还包括强化学习 [71, 72], ，但在

微观尺度上，相应的探索才刚刚开始。粒子的微小尺寸带来了各种新的挑战：(i) 微型游动体会受到布朗运动（或转向方案中的误差和延迟）引起的显著涨落，因此无法准确预测其导航动作的结果。(ii) 微型游动体会与壁面、障碍物及其他微型游动体发生流体动力学相互作用，这可能从根本上改变以最快速度抵达目标所需的导航策略 [73]。(iii) 环境导致的微型游动体位移速率可能比其自身推进速度（通常为 $\sim \mu m/s$）高出几个数量级，例如在血管中；这与风中的飞机等情况恰恰相反。(iv) 在许多潜在应用中，微型游动体将面对未知环境，只能获取周围环境的局部信息，因此需要可迁移的导航策略。

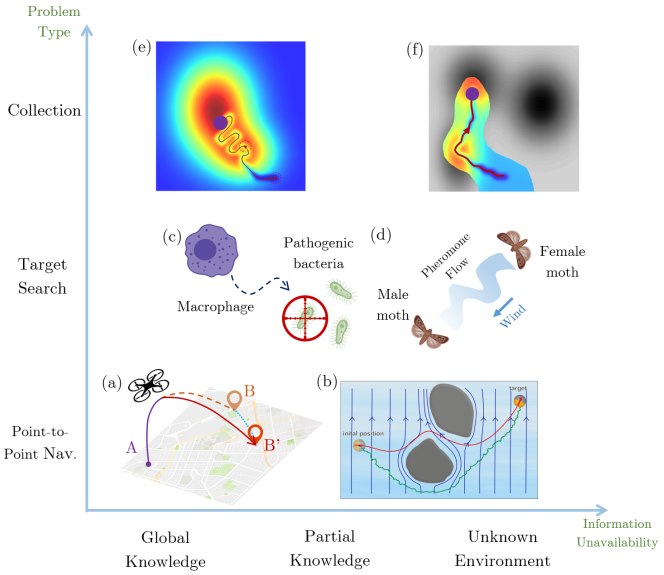

图 1：活性粒子最优导航问题的分类。示例：(a) 确定性环境和 (b) 涨落环境中的最优点对点导航；(c) 捕食者-猎物系统中的目标搜索与捕获；(d) 湍流流场中的气味搜索；在 (e) 已知和 (f) 未知设置下的最优收集与定位。

最优点对点导航。– Ernst Zermelo 于1931年提出：船只在非均匀风场中应如何航行，才能最快到达目标 [74]。泽尔梅洛问题后来与许多领域相关，从活性粒子、类鱼型水下航行器 [75]，到平流层中的无人气球导航 [76]。更一般地，本节探讨：在包含复杂（例如湍流）流场和力场、运动性场（与非均匀光强场中的光驱动微型游动体相关）、黏度场（例如黏趋性情形 [77]）以及复杂障碍物分布的环境中，受约束的自推进智能体应如何操控，才能以最优方式（例如最快、最省或最安全）到达目标。自推进智能体的最优路径（轨迹）及相应导航策略，可通过庞特里亚金原理（用于确定性问题）、哈密顿–雅可比–贝尔曼方程（也适用于随机问题） [69, 78], 、几何方法 [79–81], 、现代优化算法 [82], , 以及（深度）强化学习方法 [83–85], 来确定。尤其是在环境未知（或部分已知），或环境具有高维性和混沌特征、难以获得精确解时，这些方法特别有用（图2和图3）。

精确最优轨迹的初等计算——考虑一个处于与时间无关的二维复杂环境中的过阻尼干性活性粒子（无流体动力学相互作用）。粒子位置的运动方程 $\vec{r}(t) = (x(t)\ y(t))$ 可简洁地写为 [86]

$$\dot{\vec{r}}(t) = v_0(\vec{r})\hat{n}(t) + \vec{f}(\vec{r}) + \sqrt{2D}\vec{\eta}(t). \qquad (1)$$

这里，$f(\vec{r})$表示一般的力场、流场和黏度场，D, $\vec{\eta}$ 表示平移扩散系数以及均值为零、方差为单位值的高斯白噪声。我们先来关注一种理想化情形：智能体可以自由且瞬时地控制其自身推进方向 $\hat{n}(t) = (\cos\psi(t), \sin\psi(t))$，但不能控制其速度 $v_0(\vec{r})$，而速度可能取决于空间位置 [53, 87]。给定起点和终点，目标是找到连接两点的路径 $\vec{r}(t)$（等价地，找到转向角 $\psi(t)$），使活性粒子以最快速度到达目标。为了解决消失噪声 ($D = 0$) 条件下的这一问题，我们现在将航行时间写成路径 $y(x)$ 以及 $y'(x) = dy(x)/dx$ 的泛函，如下所示

$$T\left[y(x), y'(x), x\right] = \int_0^T \mathrm{d}t = \int_{x_A}^{x_B} \frac{\mathrm{d}x}{|\dot{x}(y(x), y'(x), x)|} \qquad (2)$$

并通过求解针对$L\,(y(x)\ y'(x)\ x) = {}_{|\dot{x}\ y\ x}{}^{1}{}_{(()\ y'(x)\ x)}$的欧拉-拉格朗日方程$\mathrm{d}_{\ \mathrm{d}x}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0$（边值问题）使其最小化，其中 $\dot{x}$ 表示 x 方向上的速度。利用（1），可得到拉格朗日函数 [86]

$$L(x, y(x), y'(x)) = \frac{(1 + y'^2)}{\left|f_x + y' f_y \pm \sqrt{v_0^2\,(1 + y'^2) - (f_y - y' f_x)^2}\right|}, \qquad (3)$$

其中 $\vec{f} = (f_x, f_y)$ 和 v_0 可能取决于 x, y。活性粒子与射线光学：欧拉-拉格朗日方程与式（3）结合后，对于恒定的 $\vec{f}$ 和恒定的 v_0，可以很容易求解，得到 $y'(x)$ = const，这表明在任何恒定场中，最短路径也是最快的。也就是说，活性粒子的转向方式恰好能够补偿环境造成的漂移。在分段恒定的环境中，最优轨迹也呈分段恒定，从而得到活性粒子的斯涅尔定律，其中涉及一种广义折射率，该折射率

也可以为负值，如超材料中的光的折射率[88, 89]。（另见 [90]。）利用守恒律（对称性），还可以得到欧拉 - 拉格朗日方程的其他精确解，这表明在复杂环境中，最短路径通常并非最快路径。在旋转流场中，活性粒子有时甚至必须先朝远离目标的方向游动，才能以最快速度到达目标。请注意，若优化其他量，例如沿路径的耗散功率，就会得到不同的拉格朗日函数，因此通常也会得到不同的导航策略 [86]。

流体动力学相互作用——与干燥活性粒子不同，参考文献 [73] 研究了会与壁面和障碍物发生流体动力学相互作用的微型游泳体。一项关键成果表明，微型游泳体所需的最优导航策略，在定性上可能不同于使干燥活性粒子或宏观车辆在相同环境中实现最优轨迹的策略（图 3a）。

强化学习——在非常复杂的环境中（例如混沌或未知情形）通常无法精确求解最优导航问题，但仍可采用基于强化学习的方法 [83]。一些研究采用表格型 Q 学习，对自热泳主动粒子进行高效实时控制（图 2b） [91], ，或训练其在具有无边缘墨西哥帽势的环境中实现最优导航（图 2c） [92]。近期其他研究则采用了行动者—评论家方法 [93]以及深度强化学习[94, 95] ，研究在日益复杂的环境中的最优导航。尤其是文献 [95] 开发了一种基于深度强化学习的方法，用于确定渐近最优轨迹。这里的关键挑战是开发一种能够找到全局最优解，而非某条局部最优路径的方法。应对这一挑战的关键“窍门”是采用基于策略梯度的方法，该方法能够“理解”并直接专注于优化期望总回报（而不是像 Q 学习这样的离策略方法）。为评估该方法，研究人员将其结果与已知的精确最优轨迹进行了比较（图 3）。文献 [96] 将微型游泳体导航问题视为马尔可夫决策问题，并假设已知环境的全局信息，以最小化代价函数的方式求解迷宫（液滴迷宫求解另见 [97] ）。最近，研究人员还训练了智能胶体微型机器人群体，使其能够在迷宫中捕获布朗载荷粒子 [98]。除了迷宫求解，近期一系列研究 [98–100] 还利用（深度）强化学习，探索微型游泳体在未知环境中的导航；在这种环境中，微型游泳体会局部探查障碍物。研究发现，接收局部感觉输入的智能胶体能够利用深度强化学习绕过障碍物到达目标 [99]，并完成复杂的导航和局部定位任务，同时满足时间限制 [100]。

与未知环境中的问题不同，对于混沌（湍流）环境中的最优导航问题，原则上可以使用最优控制理论（庞特里亚金极大值原理）来确定精确最优轨迹。然而，由于在混沌环境中，初始条件的微小变化通常会导致完全不同的终点，因此难以通过打靶法对这一问题进行数值求解；系统地改变初始条件并无助益，而且还需要处理扩展的目标区域。在评估表示精确最优解的方程时遇到的这一困难，使强化学习尤其具有价值。因此，参考文献 [93] 最近针对湍流场中的泽尔梅洛问题，采用了一种基于演员-评论家的强化学习方法（图 2d）。

文献 [94] 则探讨了在深度强化学习方法的输入特征中纳入环境线索（如涡量、流速等）所产生的影响，以及由此形成的、用于湍流中的点对点导航任务的策略。文献 [101] 则采用对抗强化学习方法，训练微型游动体在统计均匀各向同性湍流中实现高效的点对点导航；这些微型游动体的表现优于始终朝目标方向移动这一简单策略。

合成微型游动体的游动方向通常可以通过外场控制 [33], ，而许多生物微型游动体则会通过适当的形状变形自主改变游动方向。与此相应，近期有多项研究采用强化学习来探索可变形智能体的游动机制 [102–105]。相关研究还利用学习方法探究游动体如何变形才能以最快速度游动 [106], 、沿预先确定的路径前进 [107], 、表现出趋化性 [108] ，或实现最优点对点导航 [109–111]。例如，文献 [110] 研究了（三连杆）仿生鱼模型，其输入数据仅包括自身朝向以及与（移动）目标之间的距离。模型由此学会了通用策略，并进一步在训练期间未遇到过的情境中进行探索。

在另一项研究中，研究人员训练智能活性粒子利用潜在的湍流，逃离局部流体陷阱 [112], 并抵达涡量较高的目标区域 [113], ，或朝着可达到的最高高度导航 [114]。最近，配备表格型 Q 学习的智能微型游动体也展现出高效的导航策略（尽管只能获取局部信息），能够在具有各种运动性场的环境中导航 [115]。

搜索并捕获目标。– 另一类最优导航问题关注的是：主动智能体如何移动，才能高效地找到位置（及其动态特性）未知的目标。在这里，我们可以区分——

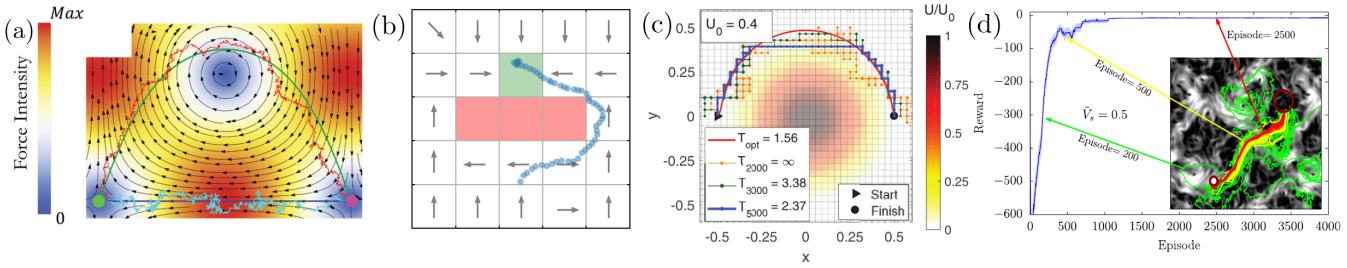

图 2：主动粒子的最优导航。(a) 在布朗噪声存在时活性粒子的轨迹（红色和青色）。粒子调整运动方向，尝试沿着 Taylor-Green 流（箭头和背景颜色）中确定性问题的最优路径（绿色曲线）和最短路径（蓝线）前进 [80]。(b) 在反馈控制胶体实验中，胶体朝目标（绿色单元格）移动时，存在虚拟障碍物（红色单元格）的情况下，学习轨迹（蓝色曲线）及对应策略图（箭头）的快照 [91]。(c) 在墨西哥帽势（背景颜色）中，通过 Q学习进行点对点导航的活性粒子轨迹；训练回合数分别为 2000（黄色）、3000（绿色）和 5000（蓝色），并与精确最优轨迹（红色）进行比较 [92]。(d) 湍流中活性粒子的总奖励随时间的变化（蓝色曲线）及示例轨迹（插图） [93]。更多详情请参见参考文献。

（i）智能体无法从目标处获得任何信息的问题，因此我们将这类目标称为“静默”目标；（ii）目标会发出某些信息的问题，例如以气味的形式发出，并通过扩散或平流传播（图 1d）；以及（iii）智能体知道目标的当前位置（能“看见”猎物），但不了解其动态特性的问题（图 1c）。另一些有趣的例子是，捕食者只能获得有关猎物的间接或部分信息（类型（ii）），例如鲨鱼和鳐鱼通过侧线传感器感知猎物造成的流场 [116,117],，以及趋化性细菌捕食者[118]。请注意，若目标静止不动，则（iii）中的特殊情形对应于上一节讨论的点对点导航。

(i) 寻找“静默”目标：近期，人们基于一项算法研究了（i）类问题。该算法将跃迁路径采样推广至主动布朗粒子动力学，用于在复杂环境中搜索目标 [119,120]。研究表明，对于在圆形约束区域中心搜寻目标的自推进粒子，通过调节自推进速度和特征旋转时间等参数，可以提高搜索效率 [121]。随后，文献 [122] 还研究了环境特征（如空间异质性）对自推进粒子的目标搜索动力学和搜索能力的影响。

（ii）寻找源：人们已深入研究宏观动物在复杂流场中搜索气味源的目标搜索问题 [123–127]。这类流场不仅能让动物之间通过信息素通信，还能比增强分子扩散更有效地将气味传播到远得多的距离 [127]。然而，流场也使得智能体难以根据从其近邻环境中获取的信息预测源的位置。面对稀疏信息时，一种常用的搜索策略是信息觅食法 [128],，它已在昆虫气味搜索问题的背景下得到深入研究。该策略旨在局部最大化信息增益；也就是说，信息觅食智能体本质上沿着（局部）信息增益梯度向上移动，与趋化细菌沿浓度梯度向上移动类似。近期，人们已将信息觅食法与值迭代方法（针对部分可观测马尔可夫决策问题 [129]）[9,130] 以及强化学习方法 [131,132]进行了比较。

(iii) 捕捉以未知方式移动的目标：近期，研究人员使用 Q学习研究了主动布朗粒子的捕食者-猎物问题 [133,134]。文献 [135] 则探讨了在微型游泳者追逐与逃避任务中训练对抗性强化学习智能体所产生的策略。在整个训练过程中，捕食者和猎物制定策略，利用流体动力学相互作用，通过复杂的连续动作和反制动作相互胜过对方。最近，研究还表明，一种深度策略梯度方法能够定性地再现低雷诺数下追逐有限尺寸猎物时的最优捕食路径 [136]。

翱翔：另一类相关问题并不要求目标一定是局部化的，并且已通过强化学习方法进行探索。这类问题涉及鸟类、无人机或（其他）滑翔机的翱翔，也就是这些智能体如何在复杂地形中导航，以发现并利用热气流[137,138]。有趣的是，人们至今仍不清楚鸟类是如何做到这一点的 [138]。

收集问题。– 未来的微型游泳者应如何移动，才能高效收集以未知方式分布的目标，例如毒素或微塑料？这一问题与机器人学中的区域清扫任务密切相关 [139–141],，但在活性物质文献中尚未得到

充分研究（图 1）。
现有研究大多聚焦于随机搜索策略。一些有影响力的研究报告称，在信天翁 [142]，海洋捕食者 [143]，大黄蜂和T细胞 [144]的觅食过程中，观察到了莱维游走（步长从厚尾分布中随机抽取）这种策略。文献 [145] 探讨了利用莱维游走活性粒子收集（不可再生的）稀疏目标。在均匀环境中，人们认为扩散运动与弹道运动的某种组合是最优的；然而，该研究发现，随着障碍物的存在使环境日益复杂，扩散性更强的策略往往能带来更高的目标收集率。类似地，文献 [146]研究了搜索单个目标的跑动-翻滚行走者。

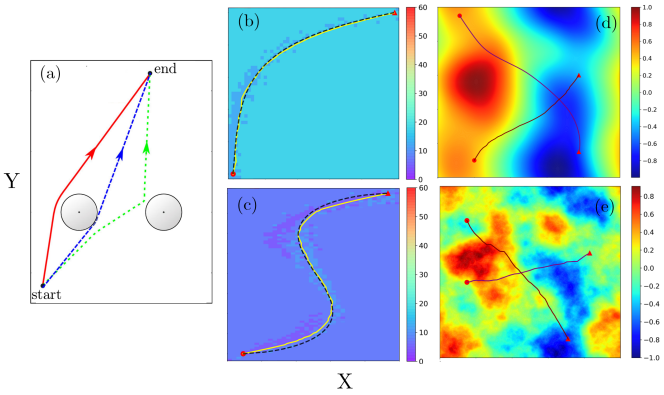

图 3：（a）流体动力学会使活性粒子为最快抵达目标而需采取的导航策略在定性上发生变化：干燥活性粒子（蓝色）以及源偶极强度为 $\sigma = -15$ （绿色）和 $\sigma = 7.5$ （红色）的源偶极微游动体与障碍物（灰色圆盘）发生流体动力学相互作用时的最优轨迹 [73]。（b-e）在给定起点和终点（红点）之间，活性粒子在有线性力场（b）和水平管流（c）时的精确最优轨迹（虚线），与机器学习得到的轨迹（黄色线） [86] 进行比较 [95]。背景颜色表示学习得到的策略图，即偏好的离散化转向“方向” ψ（$\psi \in \{0, .., 59\} \cong [0, 2\pi)$）。（d、e）线条表示高斯随机势中的学习轨迹（背景颜色） [95]。更多详情请参阅参考文献。

开放问题、挑战与展望。
机器学习得到的结果有多好？——在学习某个导航问题（或其他复杂控制问题）的解时，最终得到的轨迹或导航策略与真正最优解有多接近，往往仍不清楚。当然，奖励的收敛并不保证最优性：奖励可能收敛到任意一个局部最优 [83]；即使收敛到全局最优，也往往不清楚这个最优解是否能代表（渐近意义上的）物理最优解，还是仅在给定的奖励定义、离散化、超参数选择和所选学习算法范围内最优。因此，一项重要的开放挑战是开发能够复现精确结果、并可据此进一步推进的强化学习方法 [95]。

涨落——虽然我们预计，涨落会使抵达目标所需的导航策略在定性上发生变化（图 1b），从而实现最优抵达（图 2a） [79]，，例如悬崖行走问题中的情况 [147]，，但在其他问题中，涨落对所需导航策略的影响很小。因此，系统地理解并制定判据，明确涨落在何种情况下会使所需导航策略发生显著的定量变化，甚至定性变化，将十分重要。对于强化学习算法的开发而言，涨落会使环境变得部分可观测，因而智能体所做决策（动作）的结果无法被准确预测。在某些问题中，这种不可预测性会影响训练鲁棒性，这凸显了开发新型深度强化学习方法的必要性（例如信赖域方法 [148, 149]），以便在易变环境中保持稳健学习。

可迁移性与未知环境——微生物需要制定策略，以便在从未遇到过、且可能随时间不可预测地变化的环境中导航并寻找食物；而迄今为止，许多关于活性粒子的导航问题都依赖于固定（或具有确定性动态）的环境。一项早期研究探讨了基于深度强化学习的“胶体机器人”在未知环境中的最优导航 [99]。未来若要开发强大的方法来确定可迁移导航策略，很可能需要借助基于模型的强化学习方法 [150, 151]，更重要的是需要世界模型 [152]。在世界模型中，智能体致力于学习环境的模型（表征）（在此也可以理解为学习系统的物理学），并利用这一环境表征来规划未来的行动。

机器学习的最新进展——目前，我们正见证着新型强化学习方法的开发迅速发展。因此，一些最强大的方法尚未应用于活性物质及相关最优导航问题，也就不足为奇了。近年来，一项极具前景的研究方向已应用于国际象棋、围棋和将棋等知名棋类游戏，即引入融入规划的强化学习算法，例如 AlphaGo 和 AlphaZero [153–155]。我们认为，鉴于这些方法的规划阶段十分稳健（这得益于其内置的蒙特卡洛树搜索，可探索未来可能出现的结果），它们非常适用于那些要求所学习策略具有高度准确性且能够确信其最优性的任务。

交叉相互作用与通信规则——近期，若干研究着重探讨运动智能体如何运用强化学习技术，基于简单的底层原则和激励设计，学习集体行为 [156]，例如群集 [157, 158]。另一条相关研究方向

关注多智能体强化学习 [159] 在微型游动体问题中的应用。可以设想一些复杂任务，例如在环境感知能力较弱的情况下定位癌细胞或收集微塑料（图 1f）。要成功完成这些任务，多个智能微型游动体需要相互协作，并共享它们收集到的宿主环境知识。

参考文献

[1] Smith S. L. et al., *IEEE Trans. Robot.*, 28 (2011) 410.
[2] Levinson J. et al., *Towards fully autonomous driving: Systems and algorithms* 载于 *2011 IEEE Intell. Veh. Symp. (IV)* 2011 163 （IEEE）第 . 页。[3] Galceran E. et al., *Rob. Auton. Syst.*, 61 (2013) 1258.
[4] Lohmann K. J. et al., *Nature*, 428 (2004) 909.
[5] Sequeira A. M., *Curr. Biol.*, 30 (2020) R1054.
[6] Klimontovich Y. L., *Turbulent motion. the structureof chaos* 载于 *Turbulent Motion and the Structure of Chaos* （Springer Dordrecht）1991 年, 第 329 页。[7] Cardé R. T. et al., *J. Chem. Ecol.*, 34 (2008) 854. [8] Bau J. et al., *Integr. Comp. Biol.*, 55 (2015) 461. [9] Rigolli N. et al., *Elife*, 11 (2022) e76989. [10] Spehr M. et al., *Science*, 299 (2003) 2054. [11] Eisenbach M. et al., *Nat. Rev. Mol. Cell Biol.*, 7 (2006) 276. [12] Berg H. C., *E. coli in Motion* （Springer） 2004. [13] Bi S. et al., *Curr. Opin. Microbiol.*, 45 (2018) 22.
[14] Laganenka L. et al., *Nat. Commun.*, 7 (2016) 12984.
[15] Paxton W. F. et al., *J. Am. Chem. Soc.*, 126 (2004) 13424. [16] Elgeti J. et al., *Rep. Prog. Phys.*, 78 (2015) 056601.
[17] Cates M. E. et al., *Annu. Rev. Condens. Matter Phys.* **6** 2015 219 , ( ) .
[18] Bialké J. et al., *J. Non-Cryst. Solids*, 407 (2015) 367.
[19] Menzel A. M., *Phys. Rep.*, 554 (2015) 1. [20] Klapp S. H., *Curr. Opin. Colloid Interface Sci.*, 21 (2016) 76. [21] Ebbens S., *Curr. Opin. Colloid Interface Sci.*, 21 (2016) 14.
[22] Hagan M. F. et al., *Curr. Opin. Cell. Biol.*, 38 (2016) 74.
[23] Maass C. C. et al., *Annu. Rev. Condens. Matter Phys*, 7 2016 171( ) . [24] Patteson A. E. et al., *Curr. Opin. Colloid Interface Sci.* **21** 2016 86 , ( ) .
[25] Zöttl A. et al., *J. Phys. Cond. Mater*, 28 (2016) 253001. [26] Katuri J. et al., *Acc. Chem. Res.*, 50 (2016) 2.
[27] Moran J. L. et al., *Annu. Rev. Fluid Dyn.*, 49 (2017) 511.
[28] Mallory S. A. et al., *Annu. Rev. Phys. Chem.*, 69 (2018) 59. [29] Liebchen B. et al., *Phys. Rev. E*, 98 (2018) 052610. [30] Jülicher F. et al., *Rep. Prog. Phys.*, 81 (2018) 076601. [31] Bär M. et al., *Annu. Rev. Condens. Matter Phys.*, 11 (2020) 441. [32] Agostinelli D. et al., *Math. Eng*, 2 (2020) 230. [33] Bechinger C. et al., *Rev. Mod. Phys.*, 88 (2016) 045006. [34] Hecht L. et al., *arXiv:2102.13007*, (2021) .
[35] Liebchen B. et al., *J. Phys. Cond. Matter*, 34 (2021年) 083002. [36] Chen J. et al., *AIP Adv.*, 4 (2014) 31325.
[37] Mano T. et al., *Proc. Natl. Acad. Sci.*, 114 (2017) E2580.
[38] Palacci J. et al., *Science*, 339 (2013) 936.
[39] Sánchez S. et al., *Angewandte Chemie International Edition* **54** 2015 1414 , ( ) .
[40] Dai B. et al., *Nat. Nanotechnol.*, 11 (2016) 1087.
[41] Dong R. et al., *ACS Nano*, 10 (2016) 839.
[42] Xuan M. et al., *J. Am. Chem. Soc.*, 138 (2016) 6492.
[43] Driscoll M. et al., *Nat. Phys.*, 13 (2017) 375.
[44] Martinez-Pedrero F. et al., *Small*, 13 (2017) 1603449. [45] Nedev S. et al., *ACS Photonics*, 2 (2015) 491.
[46] Zong Y. et al., *ACS Nano*, 9 (2015) 10844.
[47] Moyses H. et al., *Soft Matter*, 12 (2016) 6357.
[48] Liu J. et al., *Nanoscale*, 8 (2016) 19894. [49] Liu J. et al., *Micromachines*, 9 (2018) 232. [50] Mousavi S. M. et al., *Soft Matter*, 15 (2019) 5748. [51] Demirörs A. F. et al., *Soft Matter*, 14 (2018) 4741. [52] Lozano C. et al., *Soft Matter*, 15 (2019) 5185. [53] Jahanshahi S. et al., *Commun. Phys.*, 3 (2020) 127. [54] Khadka U. et al., *Nat. Commun.*, 9 (2018) 3864. [55] Fernandez-Rodriguez M. A. et al., *Nature Commun.* **11** 2020 4223 , ( ) .
[56] Lavergne F. A. et al., *Science*, 364 (2019) 70.
[57] Liebchen B. et al., *Acc. Chem. Res.*, 51 (2018) 2982.
[58] Tsang A. C. et al., *Adv. Intell. Syst.*, 2 (2020) 1900137. [59] Harari Y. N., *Homo Deus: A brief history of tomorrow* (Harper) 2016.
[60] Bunea A.-I. et al., *Micromachines*, 11 (2020) 1048.
[61] Patra D. et al., *Nanoscale*, 5 (2013) 1273.
[62] Ceylan H. et al., *ACS Nano*, 13 (2019) 3353.
[63] Alapan Y. et al., *Sci. Robot.*, 3 (2018) eaar4423.
[64] Vyskocil J. et al., *ACS Nano*, 14 (2020) 8247.
[65] Hart P. E. et al., *IEEE Trans. Syst. Sci. Cybern.*, 4 (1968) 100. [66] Gasparetto A. et al., *Path Planning and Trajectory Planning Algorithms: A General Overview* (Springer International Publishing, Cham) 2015 p. 3.
[67] Martin C. F. et al., *Math. Comput. Model.*, 33 (2001) 237.
[68] Techy L. et al., *J. Guid. Control Dyn.*, 32 (2009) 1736.
[69] Lewis F. L. et al., *Optimal and robust estimation: with an introduction to stochastic control theory* (CRC press, USA) 2017.
[70] Mohamed A. et al., *Int. J. Veh. Perform.*, 4 (2018) 1.
[71] Zhang B. et al., *J. Intell. Robot. Syst.*, 77 (2015) 391.
[72] Xin J. et al., *Application of deep reinforcement learning in mobile robot path planning* in proc. of *2017 Chinese Automation Congress (CAC)* 2017 7112 (IEEE) p. . [73] Daddi-Moussa-Ider A. et al., *Commun. Phys.*, 4 (2021年) 15. [74] Zermelo E., *ZAMM-Z. fur Angew. Math. Mech.*, 11 (1931年) 114.
[75] Yu J. et al., *J. Bionic Eng.*, 15 (2018) 579.
[76] Bellemare M. G. et al., *Nature*, 588 (2020) 77.
[77] Liebchen B. et al., *Phys. Rev. Lett.*, 120 (2018) 208002.

[78] Kirk D. E., *Optimal control theory: an introduction* (Courier Corporation, Massachusetts) 2004.
[79] Piro L. et al., *Phys. Rev. Res.*, 3 (2021) 023125.
[80] Piro L. et al., *New J. Phys.*, 24 (2022) 093037.
[81] Piro L. et al., *Front. Phys.*, 10 (2022) .
[82] Panda M. et al., *Int. J. Autom. Comput.*, 17 (2020) 321.
[83] Sutton R. S. et al., *Reinforcement Learning: An Introduction* 2018 (麻省理工学院出版社, Massachusetts) .[84] Arulkumaran K. et al., *arXiv:1708.05866*, (2017) .
[85] François-Lavet V. et al. *Found. Trends Mach. Learn.* **11** 2018 219 , , ( ) .[86] Liebchen B. et al., *EPL*, 127 (2019) 34003.
[87] Lozano C. et al., *Nature Commun.*, 7 (2016) 12828.
[88] Shelby R. A. et al., *Science*, 292 (2001) 77.
[89] Smith D. R. et al., *Science*, 305 (2004) 788.
[90] Ross T. D. et al., *ACS nano*, 16 (2022) 16191.
[91] Muiños-Landin S. et al., *Sci. Robot.*, 6 (2021) eabd9285.[92] Schneider E. et al., *EPL*, 127 (2019) 64003.
[93] Biferale L. et al., *Chaos*, (2019) 103138.
[94] Gunnarson P. et al., *Nat. Commun.*, 12 (2021) 7143.
[95] Nasiri M. et al., *New J. Phys.*, 24 (2022) 073042.
[96] Yang Y. et al., *ACS Nano*, 12 (2018) 10712.
[97] Jin C. et al., *Proc. Natl. Acad. Sci.*, 114 (2017) 5089.
[98] Xu K. et al., *Adv. Intell. Syst.*, 3 (2021) 2100115.
[99] Yang Y. et al., *Adv. Intell. Syst.*, 2 (2020) 1900106.
[100] Yang Y. et al., *Adv. Theory Simul.*, 3 (2020) 2000034.
[101] Alageshan J. K. et al., *Phys. Rev. E*, 101 (2020) 043110.[102] Tsang A. C. H. et al., *Phys. Rev. Fluids*, 5 (2020) 074101.[103] Liu Y. et al., *Phys. Fluids*, 33 (2021) 062007.[104] Qiu J. et al., *Sci. China: Phys. Mech. Astron.*, 63 (2020) 1.[105] Behrens M. R. et al., *Adv. Intell. Syst.*, 4 (2022) 2200023.[106] Jiao Y. et al., *Phys. Rev. Fluids*, 6 (2021) 050505.[107] Zou Z. et al., *Commun. Phys.*, 5 (2022) .
[108] Hartl B. et al., *Proc. Natl. Acad. Sci. U.S.A.*, 118 (2021) .[109] Amoudruz L. et al., *Adv. Intell. Syst.*, 4 (2022) 2100183.[110] Zhu Y. et al., *Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci.* 2022 09544062221079693 , , ( ) .
[111] Zhu Y. et al., *Front. Phys.*, (2022) 237.
[112] Gustavsson K. et al., *Eur. Phys. J. E*, 40 (2017) 110.
[113] Colabrese S. et al., *Phys. Rev. Fluids*, 3 (2018) 084301.[114] Colabrese S. et al., *Phys. Rev. Lett.*, 118 (2017) 158004.[115] Monderkamp P. A. et al., *arXiv:2209.03047*, (2022) .[116] Kalmijn A. J., *J. Exp. Biol.*, 55 (1971) 371.[117] Gardiner J. M. et al., *Flow sensing in sharks: lateral line contributions to navigation and prey capture* in *Flow sensing in air and water* (Springer Berlin, 海德堡) 2014 p. 127.
[118] Chet I. et al., *J. Bacteriol.*, 106 (1971) 863.
[119] Zanovello L. et al., *Phys. Rev. Lett.*, 126 (2021) 018001.[120] Zanovello L. et al., *J. Chem. Phys.*, 155 (2021) 084901.[121] Wang J. et al., *Chem. Phys.*, 144 (2016) 204702.[122] Wang J. et al., *Soft Matter*, 13 (2017) 758.
[123] Celani A. et al., *Phys. Rev. X*, 4 (2014) 041015.
[124] Baker K. L. et al., *J. Neurosci.*, 38 (2018) 9383.
[125] Durve M. et al., *Phys. Rev. E*, 102 (2020) 012402.
[126] Nguyen D. M. T. et al., *Proc. Natl. Acad. Sci. U.S.A.*, 118 2021 2011916118 ( ) e .[127] Reddy G. et al., *Annu. Rev. Condens. Matter Phys.*, 13 2022 191 ( ) .
[128] Vergassola M. et al., *Nature*, 445 (2007) 406.
[129] Spaan M. T. et al., *J. Artif. Intell. Res.*, 24 (2005) 195.
[130] Heinonen R. A. et al., *arXiv:2207.04277*, (2022) .
[131] Loisy A. et al., *Proc. R. Soc. A*, 478 (2022) 20220118.
[132] Reddy G. et al., *Proc. Natl. Acad. Sci. U.S.A.*, 119 (2022) e2107431118.[133] La H. M. et al., *IEEE Trans. Control Syst. Technol.*, 23 2014 52 ( ) .
[134] Gerhard M. et al., *Phys. Rev. E*, 104 (2021) 054614.
[135] Borra F. et al., *Phys. Rev. Fluids*, 7 (2022) 023103.
[136] Zhu G. et al., *J. Fluid Mech.*, 944 (2022) A3.
[137] Reddy G. et al., *Proc. Natl. Acad. Sci. U.S.A.*, 113 (2016) E4877.[138] Reddy G. et al., *Nature*, 562 (2018) 236.
[139] Ahmadi M. et al., *Continuous area sweeping: A task definition and initial approach* 载于 *'05. Proceed-* (IEEE) 第 *ings. ICAR* 2005 316 页.[140] Ahmadi M. et al., *A multi-robot system for continuous area sweeping tasks* *Proceedings IEEE ICRA* 载于 (IEEE) 2006年第1724页.[141] Shah R. et al., *Deep r-learning for continual area sweeping* *IEEE/RSJ (IROS)* 2020 载于 (IEEE) 第5542页.[142] Viswanathan G. M. et al., *Nature*, 381 (1996) 413.[143] Sims D. W. et al., *Nature*, 451 (2008) 1098.
[144] Zaburdaev V. et al., *Rev. Mod. Phys.*, 87 (2015) 483.
[145] Volpe G. et al., *Proc. Natl. Acad. Sci. U.S.A.*, 114 (2017) 11350.[146] Rupprecht J.-F. et al., *Phys. Rev. E*, 94 (2016) 012117.[147] Van Seijen H. et al., *A theoretical and empirical analysis of expected sarsa* 载于 *IEEE Symp. Adapt. Dyn. Program. Reinf.* 2009 177 (IEEE) 第 页.[148] Schulman J. et al., *Trust region policy optimization* 载于 *Proc. Int. Conf. Mach. Learn.* (机器学习研究论文集) 2015年第1889页.[149] Schulman J. et al., *arXiv:1707.06347*, (2017) .
[150] Kaiser L. et al., *arXiv:1903.00374*, (2019) .
[151] Moerland T. M. et al., *arXiv:2006.16712*, (2020) .
[152] Schrittwieser J. et al., *Nature*, 588 (2020) 604.
[153] Silver D. et al., *Nature*, 529 (2016) 484.
[154] Silver D. et al., *Nature*, 550 (2017) 354.
[155] Silver D. et al., *Science*, 362 (2018) 1140.
[156] Charlesworth H. J. et al., *Proc. Natl. Acad. Sci.*, 116 2019 15362 ( ) .[157] Durve M. et al., *Phys. Rev. E*, 102 (2020) 012601.[158] Hornischer H. et al., *Phys. Rev. Res.*, 4 (2022) 023037.[159] Zhang K. et al., *arXiv:1911.10635*, (2019) .