arXiv:2601.03498v1 [q-bio.NC] 7 Jan 2026

可量化的信息处理层级为检测能动性提供必要条件

预印本

Brett J. Kagan[*†]
Cortical Labs，澳大利亚维多利亚州墨尔本 3000。墨尔本大学生物化学与药理学系，澳大利亚维多利亚州帕克维尔。
brett@corticallabs.com

Valentina Baccetti[*]
物理系，理学院，皇家墨尔本理工大学，澳大利亚维多利亚州墨尔本 3001。RMIT应用量子技术中心，皇家墨尔本理工大学，澳大利亚维多利亚州墨尔本 3001。

Brian D. Earp
生物医学伦理中心，杨潞龄医学院，新加坡国立大学，新加坡 117597。

J. Lomax Boyd
伯曼生物伦理学研究所，约翰斯·霍普金斯大学，美国马里兰州巴尔的摩。

Julian Savulescu
生物医学伦理中心，杨潞龄医学院，新加坡国立大学，新加坡 117597。
上广牛津研究所，牛津大学，牛津，OX1 1PT，英格兰。

Adeel Razi
特纳脑与心理健康研究所，莫纳什大学，克莱顿，澳大利亚。威康人类神经影像中心，伦敦大学学院，伦敦，英国。

2025年11月17日

摘要

随着智能系统在各种不同的基底上不断发展——从机器学习模型和神经形态硬件到体外神经培养物——理解赋予系统能动性的因素变得日益重要。然而，现有定义往往依赖难以量化的自上而下的描述。我们提出一个基于系统信息处理顺序的自下而上的框架：即系统对输入的变换随时间演化的程度。我们确定了信息处理的三个阶次。I类系统是反应式且无记忆的，将输入直接映射为输出。II类系统包含能够提供记忆的内部状态，但遵循固定的转换规则。III类系统具有自适应性——其转换规则本身会随先前活动而改变。尽管这些动力学单独并不足以构成能动性，但它们代表了真正能动性所必需的信息条件。这一层级提供了一种可测量且与基底无关的方法，用于识别能动性的信息前兆。我们通过神经生理学和计算实例阐释了这一框架，其中包括恒温器和受体样忆阻器，并讨论了其对可能表现出能动性的系统进行伦理和功能评估的意义。

关键词 信息处理 ·能动性 ·框架 · 伦理学 · 信息学

[*]这些作者贡献相同。
[†]通讯作者。

1 背景与范围

近年来，开发自主智能系统的兴趣和投入迅速增加 Wong（2025）； Kagan and Kitchen（2025）。这些进展使人们在如何定义描述此类系统的关键术语方面产生了不确定性，甚至意见分歧 Kagan et al.（2023）； Wong（2025）。其中争议最大的概念之一是能动性。尽管许多相关术语已被指出需要明确界定 Kagan et al.（2024a），本文聚焦于适用于人工系统和自然系统的能动性。本文不明确支持任何单一的能动性定义。相反，对于现有的一系列合理定义，我们提出一个可通过经验检验的必要 （但未必充分）条件，用以判定一个表面上具有智能的系统是否拥有能动性。在当代人工智能研究、开发和营销中，能动性已成为一种常见却使用不一致的描述词 Dung（2025）； Floridi（2025）； Van Lier（2023）。为消除歧义，我们提议采用一种独立于基底的视角：根据这一视角，凡是展现出相同信息属性的两个系统，无论其物理构成如何，都应使用相同的概念词汇加以描述 Dattathrani and De’（2023）。这一主张仅涉及描述上的等价性，并不意味着道德等价性。道德地位的差异可能取决于具体实现了哪些属性，以及实现了多少种属性 Kagan et al.（2024b）； Sebastián（2021）。下文将简要提及伦理影响，但对于特征组合如何对应道德地位的全面分析超出了本文的讨论范围。本文旨在提供一个形式化、可量化的框架，根据系统内在的信息动态，识别能动性可能出现之处。

我们提出，可以通过系统的信息处理阶序来约束能动性的定义——也就是说，系统对输入进行内部变换的程度取决于其自身先前的状态和输出。简言之，我们提出，一个系统要具备能动性，其信息处理架构必须能够对所接收的信息施加定性的、不可逆的变化，并且这种变化本身能够随时间进行自适应调整。我们将此称为 III类系统，并将其与此前提出的 I类和 II类系统进行对比。这一度量并未提供能动性的完整定义，而只是其潜在涌现所必需的信息条件。与关于情境化与具身系统的更广泛文献一致，Dodig-Crnkovic 和 Burgin（2024）； Barandiaran 等（2009）； Pickering（2024），我们将潜在智能体视为嵌入更大信息环境中的实体。遵循 Sultan 等（2022）的既有研究，我们还将能动性与行为区分开来：能动性关乎采取行动的能力，而行为则指这种能力的表现。在这些前提的基础上，我们采用自下而上方法，考察系统的内部动力学如何响应并整合来自其环境的信息。因此，可以根据不同的信息处理阶序来约束能动性的差异，其范围从纯粹的反应式映射，到支撑目标导向行为或自主行为的自适应、自我修改动态。

本文安排如下。第2节概述确定能动性的不同方法，并论证系统理论方法为何能够解决当前方法中的歧义。第3节介绍信息处理顺序框架，并给出最简数学定义。第4节为我们提出的每种信息处理类型提供案例，并通过模拟模型展示输入如何被系统改变。第5节随后展示如何在一段时间内直接比较这些不同系统的记忆与适应性。最后，在提出结论之前，第6节探讨所提出的信息处理顺序框架在关键领域中的适用性。

2 确定能动性的不同方法

能动性有多种定义方式。一般而言，如果一个系统具备以下一项或多项特征，人们便认为它具有能动性：目标导向性 Dung（2025）； 拥有第一人称视角，或对决策具有所有权感 Das（2025）； 能够独立行动以实现目标 Dodig-Crnkovic 和 Burgin（2024）； 能够建模自身活动和／或其环境中的活动 Barzegar 等；或者能够依据内部心智模型驱动有意图的行动（此处“有意图”指智能体想要做的任何事情）Swanepoel（2021）； Swanepoel 和 Corks（2024）。考察可能定义能动性的各种能力时，可以从不同的抽象层次分别考量这些能力，并将交互性、自主性和适应性等特征，以及上述任何特征纳入考量 Floridi（2025）； Floridi 和 Sanders（2004）。现有论述的一项关键局限在于，它们自上而下地定义能动性——依据系统具备某些行为的能力，而非自下而上地从产生这种能力的机制来定义。这会造成循环论证问题：如果仅凭外在行为推断能动性，那么行为本身就成了能动性的实际标志。然而，行为很容易误导人们对内部实际情况的判断。因此，即使系统的底层过程与构成真正能动性的过程截然不同，人们也可能仅凭其外在性能就将其描述为“智能体”。这会导致人们把行为模仿与真正的能动性混为一谈，进而促使人们错误地将道德能动性、责任 Véliz（2021）； Nyholm（2021）或道德受动性 Shevlin（2021）归于这些系统。这一问题还因系统与外部环境之间的边界模糊而更加复杂 Aguilera 等（2018），嵌套系统尤其如此；

在这类嵌套系统中，信息处理系统运行于更大的信息处理系统之内。这些嵌套系统构成了一大类系统，其中包括所有生物系统以及许多人工系统。归根结底，如果不考虑产生这些能力的底层机制，就难以在实践中判定某个系统是否具有能动性；如果具有，它属于何种类型，或能动性程度如何。

2.1 为什么仅凭行为不足以判定能动性

评估能动性不仅需要观察行为，还需要了解系统在不同条件下采取行动的潜在能力。为说明这一推理，不妨考虑反事实——设想中的其他可能结果 Byrne（2016年）；Dwyer（2021年）。仅凭行为很难判断系统内在的能动性，因为外部因素可能会阻止系统像原本可能的那样行动。例如，在一场扑克游戏中，最优秀的玩家仍可能因运气不佳而输掉比赛。然而，我们可以设想一个反事实世界：如果拿到不同的牌，同一位玩家很可能会获胜，因为其决策实际上更胜一筹。这表明，即使在特定环境中没有实现某些结果，具有能动性的系统仍可能拥有实现这些结果的能力。进一步而言，即使是我们认为显然具有能动性的系统——例如人脑——其能力的发挥也依赖外部条件。大脑必须处于正常运作的身体之中，而身体则充当它与世界交互的接口。身体在某些限制下无法执行大脑预期的行动，并不会因此削弱大脑的能动性。

关于这一问题，还有另一种视角。智能体常被描述为在最大化某种形式的效用或奖励。虽然这通常属实，但具有相同行为表现的系统，其内部机制可能大不相同 Ramírez-Ruiz 等人（2024年）。例如，受最大占据原则引导的智能体会探索并占据多种行动—状态路径；它们的表现可以与奖励最大化智能体不相上下，但行为却更加多样 Ramírez-Ruiz 等人（2024年）。可以说，这类智能体具有更强的能动性，因为它们的行动源于内部驱动因素，而非僵化的外部目标，因此拥有更大的灵活性和选择空间。这一观点与相关研究结果一致：当人们能够比较并选择不同的行动方案时，他们报告的能动感会更强 Kulakova 等人（2017年）。总体而言，纯粹受外部规则支配的系统，其真正的能动性不如那些内部动力学使其能够在多种可能的行动方案中进行选择的系统 Hendrickx（2023年）；Ramírez-Ruiz 等人（2024年）；Kishore Chakrabarty（2025年）。复杂系统，包括人类智能体，也可能具有层级嵌套目标；这些目标整体上被认为是实现福祉所必需的 Miller 等人（2022年）。

最后，考虑一个在讨论能动性时经常使用的经典例子：恒温器。它存在于信息环境中，并会根据环境的变化调整其输出。此外，现代恒温器采用比例-积分-微分算法来预测并实现恒温控制目标。这是否意味着它是一个能动体？一些基于能力的观点会给出肯定的回答，至少对于足够复杂的恒温器而言，Floridi（2025）。然而，直觉上，大多数人并不愿意断定恒温器是真正的能动体。与动物或学习系统不同，恒温器缺乏能够让它以意向性评估或采取其他可能行动的内部动力学——它只是在作出反应。无需诉诸拟人化，这就表明，仅以行为能力来定义能动性的局限。[3]

2.2 以系统理论方法补充可消除歧义

一种替代方法是采用系统理论方法来探讨能动性，考察能动性（按照一系列合理的定义来理解）如何基于系统的特征而产生（Miehling 等人，2025年）。这种方法旨在说明复杂整体如何由组成部分之间的相互作用而产生（Åström 和 Murray，2021年；Hofkirchner 和 Schafranek，2011年）。也就是说，它考察每个组成部分如何独立发挥作用，以及它如何与其他部分相互作用，从而产生系统的整体行为。格式塔观点认为整体总是大于部分之和；与此不同，系统理论视角认为，复杂或非线性结果可能由各部分的相互作用而产生，但并非必然如此。要建立能动性的系统理论解释，我们必须找出所有能够展现能动性的系统共同具备的基本特征。如上所述，此类系统存在于更广泛的信息环境中，并受到其影响。任何具有能动性潜力的系统，都必须能够通过某种将环境与其内部状态联系起来的感官机制，接收并处理外部信息。至关重要的是，具有能动性的系统处理这些信息的方式，必须在性质上有别于不具备能动性的系统。我们提出，系统内存在三种有序的信息处理类别，而真正的能动性只会出现在具备第三类、也是最高级别的信息处理能力的系统中；这一级别也涵盖所有较低层级。

[3]有人认为，能动性取决于观察者所选择的参照系 Abel 等人（2025），但这种观点使能动性成为视角的属性，而非系统本身的属性。基于这种观点，能动性将不再能作为系统的一种客观特征。

3 信息处理的层级

信息处理“阶”指的是系统处理信息时所具有的不同复杂程度。“阶”一词旨在表明，每个更高阶都包含其下各阶的复杂性。在此，我们提出信息处理阶的概念，将其作为一种结构框架，用于根据系统如何通过输入、内部状态和输出之间的相互作用来调节和转换信息，对系统进行区分。理论上，信息处理的阶数可能没有上限；不过，我们认为，可以将其简化为三类，这三类也分别对应信息处理行为连续发展的前三阶。

- I类系统是纯粹的反应式系统，直接将输入映射为输出，且无记忆。
- II类系统扩展了I类系统的功能，纳入内部状态，使其响应能够同时取决于当前输入和过去的输入；然而，它们组合或赋予这些信息权重的方式会随时间保持不变。
- III类系统在II类系统的基础上进一步扩展，纳入自适应机制，使控制输入—状态耦合的参数随时间演变，并以系统自身的活动为函数，从而实现自我调节。

III类系统可能具有多个层次的适应，这会大幅增加复杂性。随着III类系统内信息处理复杂性的增加，信息处理动态可能逐渐受到系统自身历史和配置的塑造，从而产生记忆和适应性等性质。然而，为了确立将某个系统视为具有能动性的必要条件，这些额外机制并非必需。最后，我们也承认，如果信息发生了性质不同的变换，更高类别的系统也可能存在；但考察这些动态超出了本文的范围，也不属于当前的研究重点。

信息处理类别这一概念为根据系统处理和变换信息的方式区分系统提供了结构基础。该方法不关注学习或性能，而是考察输入、内部状态与输出之间的内在关系。随着复杂性的增加，这些关系逐渐更多地依赖于先前的输入和内部配置，从而产生记忆和适应性。该层级为识别能动性的必要信息前兆提供了一种有原则的方法。在层级框架中看待系统的不同类别同样是必要的，因为III类系统几乎必然也包含I类和II类要素。因此，系统类别由其包含的最高阶信息处理动态所定义。

这一层级视角为描述连续的信息处理类别奠定了基础，而学习及其他能动能力则代表了超越适应性机制范围的潜在扩展。在这一框架中，适应性指系统为应对输入或环境条件的变化而自发调整其内部动力学。按此定义，适应性可被视为能动性的最低限度表现，但它仍不同于神经科学和机器学习中所定义的学习；在后者中，行为或参数会通过目标导向优化或基于奖励的过程而发生改变（Friston 等人，2012年；Kagan 等人，2025年；Miconi 和 Kay，2025年）。这一视角关注输入、内部状态与输出之间的内在关系，而非外部性能。重要的是，理解系统的信息处理阶序，为识别系统要具备能动性等特质所必需的动态前提，提供了一种有原则的方法。下文将正式定义这些不同阶序的信息处理。

3.1 I类：一阶信息处理

考虑一个系统，其中信息被接收、在系统中流转，随后触发一项行动，而信息质量并未发生任何有意义的定性改变。因此，除去某个任意噪声项，只要了解系统函数，就能够通过输出减去引入的噪声推导出输入。例如，鲁布·戈德堡机械装置无论多么复杂，都不会对信息进行定性转换。一个过程被启动后，便按照一组既定流程和非动态规则持续运行；这些规则的逆向操作能够从输出反推出任何输入。我们将这类系统所经历的过程定义为「一阶信息处理」：系统以无记忆、反应式的方式运行。形式上，我们可以将一阶信息处理系统的动力学表示为：

$$R(t) = \alpha(t)\, I(t) + \varepsilon(t), \tag{1}$$

其中， $R(t)$ 是系统在时间 t 的输出， $I(t)$ 是在时间 t 接收到的输入信息， $\alpha(t)$ 是独立于 $I(t)$ 的时变增益或偏置，而 $\varepsilon(t)$ 是零均值随机误差项。该映射

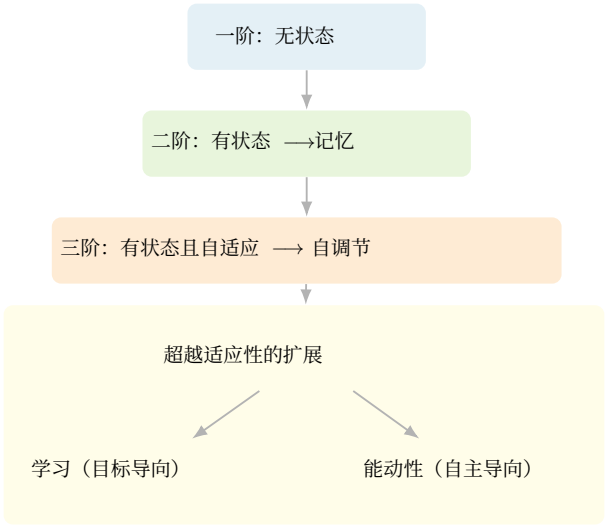


图1：信息处理类别及其与记忆、适应性和能动性的关系。

$I(t) \mapsto R(t)$ 是恒等映射，经过增益/偏置 $\alpha(t)$ 后，该增益/偏置可能随时间或其他外生变量发生漂移；它绝不取决于 $I(t)$ 的具体取值。因此，信息的类型得以保留（没有发生定性变化），而系统可以对其进行缩放或施加偏移量，并加入噪声。这与如下观点相一致：尽管信息可能对系统产生作用，但系统对信息的作用是固定的，即使触发因素本身发生变化也是如此。在这一框架下，恒温器——即使是能够动态调整阈值的恒温器——也仅执行这种一阶信息处理。恒温器接收信息，并在达到阈值时对该信息采取行动。即使需要多个阈值才能触发行动，或在这一过程中考虑多种类型的信息，信息本身也不会被系统定性改变，且这种变化本身属于恒定类型。生物系统也可能执行这种一阶信息处理，因为在简单反射中观察到的正是这一过程。

此类系统的例子包括决策树、欧姆电阻器、具有外生时变调制的线性时变系统（LTV）Bittanti 和 Colaneri（2009），以及经典开关式恒温器 Pickering（2024）。

3.2 II类：二阶信息处理

考虑另一种系统：系统接收信息，随后对信息采取行动，从而产生定性且不可逆的变化。这种变化可能导致信息质量发生改变，但变化类型本身并不会根据信息质量而改变。该系统可以被视为有状态的（在特定时间保持给定状态），但它本身并不会通过自身的内部机制实现动态自适应。I型盲视患者可能就是一个例子。这些患者能够对视觉信息作出反应，但由于视觉皮层受损，他们在没有任何意识觉知的情况下完成这一过程。此前，这一现象被讨论为一种无意识的智能行为 Kagan et al.（2022），但它也可能代表一种二阶信息处理形式。在这种情况下，视觉信息由眼睛接收，随后直接传输至丘脑的上丘和外侧膝状体核 Burra et al.（2019）。这些神经系统对接收到的信息采取行动，通过非线性方法对其进行重组和重新分配，从而触发运动皮层活动，进而驱动种类有限的行为。尽管低级脑区对信息进行了定性变换，但这种定性变化本身是依据预设标准确定的。这一过程能够驱动一系列中等复杂度的反应，因此并非反射行为，但也可能不应被视为能动行为；事实上，那些基于I型盲视过程作出反应的患者报告称，他们对该行动没有自我感知的能动性 Brogaard（2011）。形式上，这一过程可表示为：

$$R(t) = \mathcal{T}\big[I(t)\big] + \varepsilon(t) \tag{2}$$

它通过添加一个固定（时不变）的变换算子$\mathcal{T}$，扩展了一阶表示。该算子 $\mathcal{T}$可以对输入进行重塑、滤波、阈值处理或其他定性改变，但 $\mathcal{T}$本身是恒定的——不依赖于 I的具体实例（t）。

这类系统的例子包括RC低通滤波器、反馈控制系统（如Bang Bang恒温器），以及具有记忆的非线性两端电路元件，例如忆阻器（Chua，1971；Caravelli和Carbajal，2018；Caravelli，2018）。

因此，变换的性质绝不会根据输入信息或输出而演变，因此信息所发生的变化本身不会改变。

3.3 III类：三阶（及更高阶）信息处理

考虑一个最终示例系统：系统接收到信息后，会据此采取行动，从而产生定性的、不可逆的变化。然而，系统的内部动力学可以根据先前接收的信息引发的行为所产生的结果，进一步改变这一变化。这使系统能够随时间自适应并自我调节。III类系统所展现的动力学可能比I类和II类系统复杂得多，因为适应性动力学本身可以层层叠加（提高信息处理顺序）；然而，通过进化到自我调节，系统便进入一个不同的类别，有别于低阶的无记忆或仅有状态的系统。形式上，具有三阶信息处理关系的最简单III类系统可表示为：

$$R_t = \mathcal{T}_t\big[I_t\big] + \varepsilon_t, \tag{3}$$

$$\mathcal{T}_{t+1} = \mathcal{G}\big(\mathcal{T}_t, R_t\big). \tag{4}$$

该模型扩展了二阶信息处理的表征，因此 $\mathcal{T}$ 对 I 在时间 t的影响可能会随时间而变化，具体取决于适应规则 $\mathcal{G}$ 以及系统在该时刻的响应（R_t）。这里新增的关键方程（4）加入了递归：变换本身是一个状态变量，每次输出后通过$\mathcal{G}$进行更新。因此，未来的输入会遇到一个取决于过去输出的 $\mathcal{T}_{t+1}$，从而体现出变化本身可能会因系统的输出而发生改变，并由此产生自适应行为。

这类系统的例子包括已建模的生物神经网络动力学Viriyopase等（2012年），以及人工循环神经网络（RNN）Tampuu等（2018年）、具有输入调制机制的忆阻器Caravelli和Carbajal（2018年），还有采用Oja规则的赫布学习（用于稳定性）Halvagal和Zenke（2023年）。

4 案例示例

在概述了信息处理类别的理论基础之后，下面我们将介绍与各类别相对应的一些案例。这些示例并非旨在穷尽所有可能的表现形式，而是作为能够捕捉每一类别基本特征的最小模型：反应性、基于状态的记忆和适应性。

我们将考察以下三个例子：具有时变外生增益的简单恒温器（具有一阶处理的 I 类）；理想忆阻器 Caravelli and Carbajal (2018); Caravelli (2018)（具有二阶处理的 II 类）；以及输入自适应理想忆阻器 Caravelli and Carbajal (2018); Caravelli (2018)（具有三阶处理的 III 类）

4.1 一阶信息处理示例：恒温开关

在此例中，恒温开关根据二进制切换信号$s(t)$，随时间调节系统的增益。瞬时增益 $\alpha(t)$ 定义为

$$\alpha(t) = \alpha_{\text{off}}\left[1 - s(t)\right] + \alpha_{\text{on}} s(t), \tag{5}$$

因此，系统会在两个恒定增益值 α_{off} 和 α_{on} 之间交替，对应于开关的关闭和开启状态。尽管整体增益 $\alpha(t)$ 随时间变化，切换过程完全是外生的，也就是说，$s(t)$ 的演变独立于输入信号 $I(t)$。在此实现中，$s(t)$ 定义为二进制方波

$$s(t) = H\left[\sin(2\pi f_s t + \phi)\right], \tag{6}$$

其中，$H[\cdot]$ 是 Heaviside阶跃函数。频率为 f_s、相位为 ϕ 的正弦调制会产生振幅为 $\{0, 1\}$、占空比为 50% 的周期性方波，驱动系统在导通和关断状态之间规律地交替切换。

为使这类信息处理的特征更加明显，图2展示了恒温开关的输入和输出信号随时间的变化。输入由一个在正、负振幅之间周期性交替的方波组成，驱动系统依次经历开启和关闭状态

。由于增益 $\alpha(t)$ 的调制完全是外源性的，输出会瞬时跟随输入，在对应于 α_{on} 和 α_{off} 的两个恒定电平之间按比例切换。输入极性的每次变化都会立即使输出反转；一旦输入回到零值，便不会出现残余瞬态或延迟。

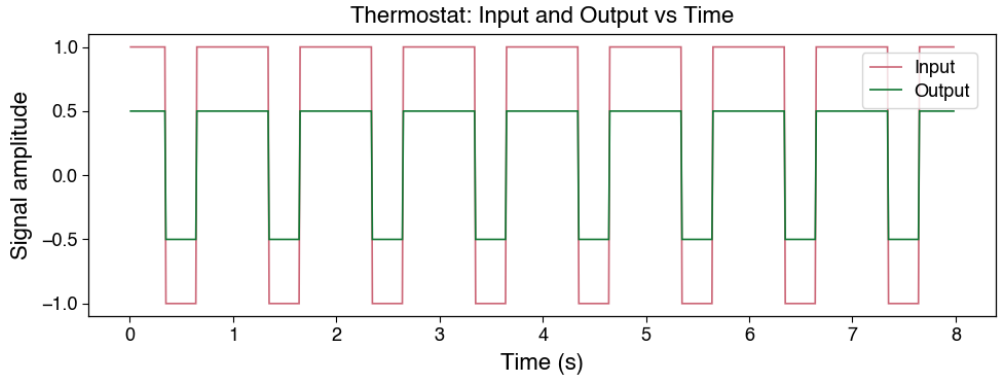


图2：一阶恒温器对方波的响应：在两个固定输出电平之间即时、按比例切换；没有记忆或适应。

4.2 二阶信息处理示例：理想忆阻器

作为二阶信息处理系统的一个示例，我们考察一种忆阻器，这是一种双端电气元件，其瞬时电阻 $\mathcal{R}(t)$取决于先前施加的电压或电流 Chua (1971)：$V(t) = \mathcal{R}(\eta(t))i(t)$，其中 $V(t)$为电压，$i(t)$为电流。与电阻固定的线性电阻器不同，忆阻器通过内部状态变量 $\eta(t)$保留对过去输入的记忆，该变量会调节其电阻。因此，该器件在单一元件中结合了即时响应与时间积分，自然实现了具有内部记忆的系统。具体而言，我们将考察理想忆阻器 Chua (1971); Caravelli(2018)，其电阻可表示为

$$\mathcal{R}(\eta(t)) = \mathcal{R}_{\text{off}}\,(1-\eta(t)) + \mathcal{R}_{\text{on}}\eta(t), \tag{7}$$

其中，$\mathcal{R}_{\text{off}}$ 和 $\mathcal{R}_{\text{on}}$ 分别表示器件的高电阻和低电阻极限。系统的动力学遵循内部状态变量的运动方程

$$\frac{d\eta}{dt} = \frac{\mathcal{R}_{\text{off}}}{\beta}\frac{V(t)}{\mathcal{R}(\eta(t))} - \alpha\,\eta(t). \tag{8}$$

其中，α 决定内部状态的弛豫速率，β 是有效激活电压，用于设定施加电压 $V(t)$ 与忆阻器内部动力学之间的耦合。第一项驱动状态对施加电压作出响应，第二项则确保在没有输入时状态呈指数形式向平衡态弛豫。因此，系统的电流响应既取决于瞬时电压，也取决于编码在 $\eta(t)$ 中的记忆；后者整合了过去输入的影响。需要注意的是，对于电压驱动的理想忆阻器，输入信息 $I(t)$ 可以编码在电压 $V(t)$ 中，而响应 $R(t)$ 对应于输出电流 $i(t)$。因此，采用与式（2）相同的记号，固定变换算子定义为

$$\mathcal{T}[V(t)] = \frac{V(t)}{\mathcal{R}(\eta(t))} \tag{9}$$

尽管算子 $\mathcal{T}[V(t)]$ 似乎通过 $\eta(t)$ 随时间变化，但其函数形式和控制参数始终固定。一旦参数 α、β、R_{off} 和 R_{on} 固定，忆阻器的变换定律在形式上便保持不变。从这个意义上说，理想忆阻器实现的是一个具有状态依赖动力学（而非输入依赖动力学）的固定算子。这个例子说明了二阶信息处理系统与一阶系统在本质上的不同：它们通过一个固定算子变换输入，而该算子的瞬时响应取决于一个携带过去信息的内部变量。

与恒温开关类似，我们考察该系统对双极性方波输入的响应，如图3所示。在这种情况下，尽管输入在正、负振幅之间快速切换，输出电流的响应却较为渐进，反映出内部状态变量 $\eta(t)$ 的缓慢演变。每次电压转换都会留下

一段输出瞬态轨迹，形成在输入反转后仍持续存在的弯曲曲线段。这种延迟且依赖历史的响应表明，忆阻器会随时间整合信息：尽管其变换定律保持不变，瞬时输出不仅取决于当前输入，也取决于先前输入所形成的残余状态。

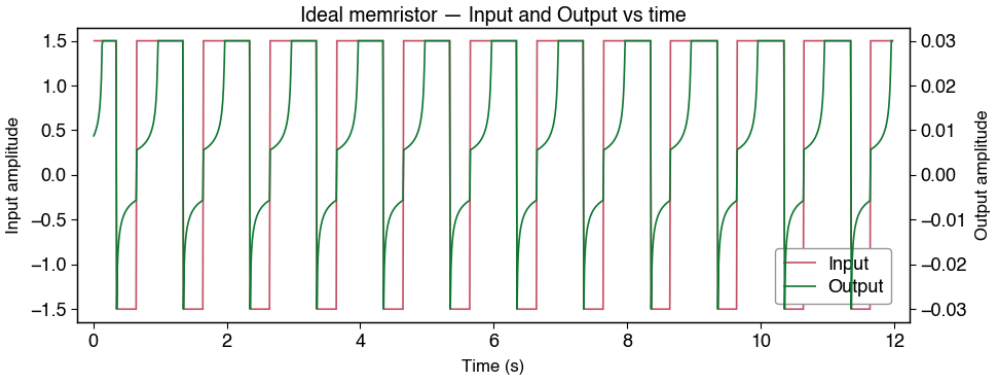


图3：理想忆阻器对方波的响应。输出电流反映了内部状态的缓慢演变，因此每次输入跃变都会在下一个周期留下延续效应；映射是固定的，但响应取决于近期历史。

4.3 三阶信息处理示例：忆阻型生物受体

借鉴生物受体适应，我们将三阶信息处理系统定义为一种理想忆阻器，它具有缓慢的、依赖状态的机制，可根据近期活动调节有效输入增益和偏置。我们将这种器件称为忆阻型生物受体（MBR），意指一种忆阻元件，它通过单一内部状态（渐逝记忆）和缓慢的观测变量实现输入增益和偏置的受体式适应；这并不意味着存在分子结合或学习。忆阻器的内部状态为系统提供内在的（渐逝）记忆，而缓慢机制则在不进行学习的情况下调节灵敏度。我们加入类似神经元稳态的缓慢自适应机制，根据过去活动的慢速平均值调整输入振幅和偏置（Ladenbauer等（2014）；Benda（2021））。在当前系统中，适应通过指数加权移动平均（EWMAs）实现，用以提供内部变量的平滑估计，其时间尺度远长于器件的内在动力学（Ladenbauer等（2012））。这些平均量充当“观测”变量，调节器件的变换算子，使其能够根据内部或外部活动的持续变化调整自身振幅和偏置。在每个时间步，施加电压构造如下：

$$V(t) = A(t)\,u(t) + b(t), \tag{10}$$

其中，$u(t)$ 是外部输入信号，而自适应变量 $A(t)$ 和 $b(t)$ 分别调节有效输入振幅和偏置。根据更新规则，它们的动力学演化时间尺度比忆阻器的内部状态更慢

$$\frac{dA(t)}{dt} = \gamma_A\left[\sigma_\phi^2(t) - \sigma_*^2\right], \tag{11}$$

$$\frac{db(t)}{dt} = \gamma_b\left[\eta_* - \bar{\eta}(t)\right], \tag{12}$$

这里，$\bar{\eta}(t)$ 和 $\sigma_\phi^2(t)$ 分别表示内部状态 $\eta(t)$ 及其活动度量 $\phi(t) = \eta(t)(1-\eta(t))$ 的指数加权移动平均，定义如下：

$$\frac{d\bar{\eta}(t)}{dt} = \frac{1}{\tau_\eta}\left[\eta(t) - \bar{\eta}(t)\right], \qquad \frac{d\sigma_\phi^2}{dt} = \frac{1}{\tau_\phi}\left[\left(\phi(t) - \bar{\phi}(t)\right)^2 - \sigma_\phi^2\right]. \tag{13}$$

参数τ_η 和 τ_ϕ 是适应时间常数，决定了 $\bar{\eta}$ 和 σ_ϕ^2 整合过去活动的速率。参数 γ_A 和 γ_b 控制振幅和偏置的适应速率，而 η_* 和 σ_*^2 设定内部状态长期均值和变异性的目标值。因此，振幅适应通过 σ_ϕ^2响应内部活动变异性的持续偏离，而偏置则适应其均值水平的偏离 $\bar{\eta}(t)$，使系统在输入条件变化时仍能保持稳定且响应灵敏的行为。指数加权移动平均的实现方式

与生物物理模型的理念相似，在这些模型中，适应电流充当近期活动的低通滤波器，实质上实现了随时间进行的指数平均（Ladenbauer等（2014）；Benda和Herz（2003）；Benda（2021））。
按照公式（3）和（4）的记号，输入变换规则和适应映射为

$$\mathcal{T}_k[V_k] = \frac{A_k u_k + b_k}{\mathcal{R}(\eta_k)}, \qquad \mathcal{G}\left[\mathcal{T}_k, R_k\right] = \begin{cases} \dot{A}_k \\ \dot{b}_k \end{cases} \tag{14}$$

通过这些耦合的自适应反馈，忆阻器不再依据固定的变换定律运行，而是会根据自身的活动持续重新配置其输入—输出映射。内部状态 $\eta(t)$ 提供了近期输入的短期痕迹（短期记忆），而自适应变量则引入了更慢、更长期的记忆，用于调节器件参数。从这个意义上说，该器件表现出自调节动力学，其中记忆和自适应调节共同决定输入信号的变换。

与前两个示例类似，我们考察该系统对双极性方波输入的响应，如图4所示。与理想忆阻器一样，输出电流会略有延迟地激活。然而，在连续的周期中，其振幅和基线也会逐渐发生偏移。每次转换都保留了理想忆阻器所具有的特征性曲率，只是没有那么明显；不过，随着系统通过自适应变量$A(t)$和 $b(t)$缓慢调节其有效增益和偏置，整体波形会逐渐向上漂移。这些缓慢而累积的调整表明，支配输入—输出关系的变换不再是固定的，而是会响应系统自身持续的活动而演化。因此，该器件将短期记忆与对其变换定律进行更慢的自适应调制结合起来。

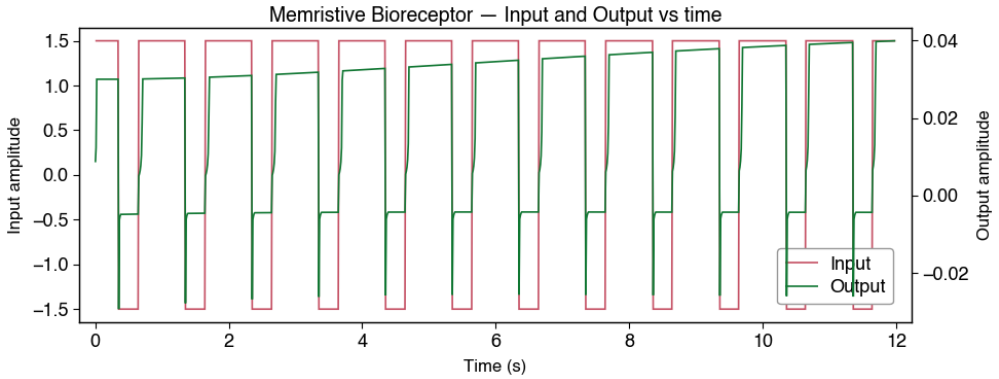


图 4：方波驱动下的忆阻型生物受体。输出呈现忆阻瞬态，并在缓慢变化的活动平均值驱动下，灵敏度和偏移量逐渐改变，表明输入—输出映射本身会随时间自适应。

5 记忆与适应性的表征

为进一步表征三个系统中记忆与适应性的涌现，我们考察它们在输入—输出平面（IO 平面）中的动力学，该平面能够反映各系统如何保留有关过去输入的信息；同时，我们还采用输入—输出关系的简单时间度量，例如滑动窗口线性回归和过零滞后分析，以揭示有效增益和偏置的逐渐变化，这些变化表明存在自适应动力学。

5.1 输入-输出轨迹与记忆

为说明随着信息处理阶序提高，记忆如何涌现，我们分析各系统在周期性驱动下于输入—输出（IO）平面中的轨迹，并判断其是否形成滞回轨迹。滞回环的存在被视为记忆的标志，这沿用了忆阻器件中的既有解释：在忆阻器件中，电流-电压（I-V）特性中的滞后被视为忆阻行为的标志，且更普遍地说，是记忆的标志 Chua (1971); Pershin和Di Ventra (2011)

更广泛地说，滞后与记忆之间的关联在凝聚态物理中根深蒂固，并广泛存在于许多受驱动的无序系统中，涵盖磁性材料、铁电材料、非晶固体及其他

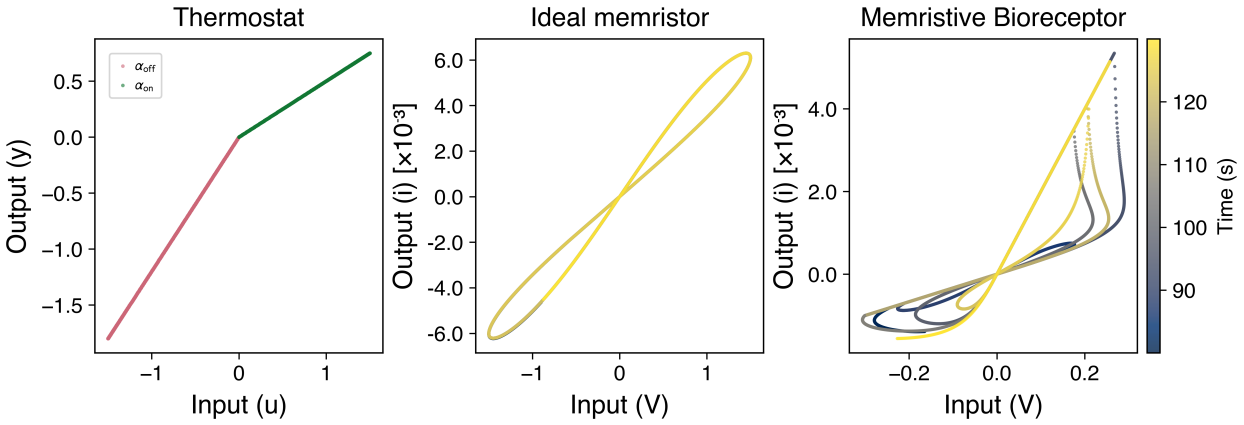


图5：三个示例系统在正弦输入驱动下的输入-输出轨迹。恒温器（左）产生两条固定增益线，分别对应恒定增益值 α_{on} 和 α_{off}（开启和关闭状态），呈现瞬时无记忆响应。理想忆阻器（中）呈现稳定的电压-电流回线，其中较早时刻的点（深蓝色）会被较晚时刻的点（浅黄色）重新经过，表明其具有无适应性的记忆。MBR（右）则呈现一条随时间逐渐移动并变形的回线（颜色刻度表示模拟最后 50 秒内的绝对时间），表明其具有适应性的记忆

非平衡介质 ，其中滞后响应会产生返回点记忆或历史依赖效应（Seth
na 等（1993）；Keim 等（2019）。

在图5中，我们展示了第4节所讨论的三个示例系统在正弦输入驱动下于输入-输出平面中的轨迹。对于恒温器（左图），输入代表无量纲驱动信号，而输出对应系统的瞬时状态。由此产生的点聚集在与两个恒定增益值 α_{on} 和 α_{off} 相关的两点附近（分别对应开关的开启和关闭状态），表明输入与输出之间存在瞬时的无记忆映射。

对于图5中的忆阻系统（中图和右图），输入对应施加电压 $V(t)$，输出则对应由此产生的电流 $i(t)$，时间范围为模拟的最后50秒（$t = 80$–130 s）。瞬态行为消退后，色彩图显示了此时间窗口内的绝对时间，展示了自适应器件的响应如何演变。在理想（非自适应）忆阻器（中图）中，电压-电流轨迹形成一个静止的回线，较早时刻的点（以深蓝色显示）会被重新经过并被较新的点覆盖，这一点印证了这一现象。正如理想忆阻器所预期的那样，这表明电流取决于当前电压以及由先前输入确立的内部状态，反映出一种固定的变换，但仍保留对过去活动的记忆，即无适应性的记忆。

在忆阻器生物受体（右图）中，电流-电压轨迹不再沿着一条固定路径往返。相反，随着时间推移，回线的各个连续区段逐渐发生偏移，如曲线上颜色由深变浅所示。这种漂移表明，电压与电流之间的关系会随着系统的自适应变量（$A(t)$ 和 $b(t)$，见式（10））根据当前及此前的活动进行调整而不断演变。因此，回线逐渐变形并发生位移，定性地表明系统具有适应性的记忆，即这种变换不仅取决于过去的输入，还会通过缓慢的参数演化进行自我调节。

需要补充的是，虽然输入-输出平面中存在滞后现象可定性地表明系统具有记忆，但它无法直接量化记忆的大小或时间尺度。不同的环路形状可能反映不同的底层动力学，例如饱和、非线性和适应，但无法据此直接衡量系统保留了多少过去的信息。对记忆进行定量评估通常需要结合基于时间相关性或信息论度量的时间分析，例如输出信号自相关或信息存储（Lizier 等人（2012））。

5.2 滚动回归、零交叉滞后与适应性

虽然我们并不试图精确量化适应性，但仍可通过两种简单的时间分析推断出有意义的信息：滚动（滑动窗口）回归 zi
v（2006）和过零滞后分析。

滚动回归通过对输入–输出数据中短小且相互重叠的片段（或滑动窗口）拟合线性模型，捕捉响应强度和基线的缓慢变化。这可以估算系统的有效增益和偏置在时间上的局部变化，这些量描述了系统对输入的反应强度，以及其反应所围绕的基线水平。如图6所示，通过观察局部增益因子随时间的变化，我们可以评估系统的变换算子是保持不变还是动态变化：如果局部增益恒定或独立于输入，系统的行为就像一阶处理器；而有规律的变化则表明存在记忆或适应[4]。

过零滞后关注的则是响应时序，即比较输入和输出何时跨越某个参考水平（通常为零，但更一般地，也可以是均值或稳态值等任意基线）。滞后保持恒定意味着响应平稳，而滞后随时间发生变化，则可解释为时间对齐方面的自适应变化。

这两项指标共同捕捉了系统行为的互补特征：一项描述响应的振幅和偏移量如何演变，另一项则显示系统的响应时序如何相对于驱动信号进行调整。

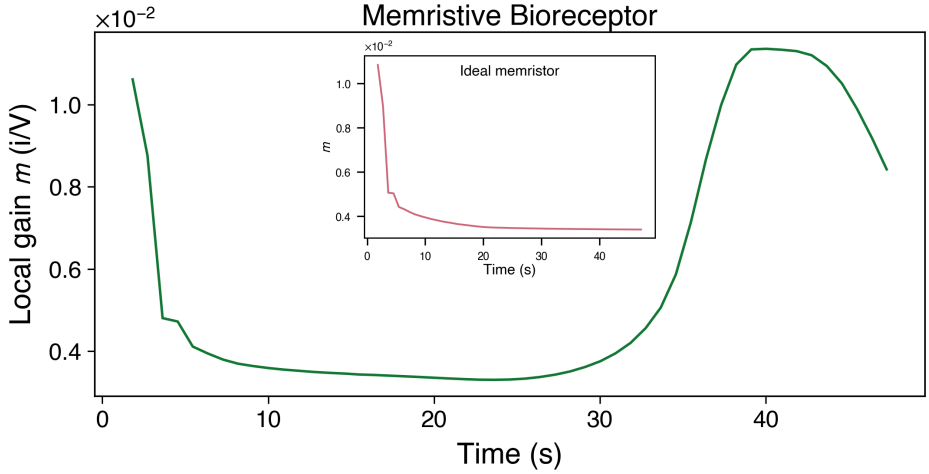


图6：慢变正弦输入下，通过滑动窗口回归得到的局部增益 $m(t)$。忆阻型生物受体（主图）显示出明显的频率依赖性增益增加。插图：理想忆阻器对应的 $m(t)$，其值基本保持恒定。

我们已对迄今考察的三个系统——恒温器、理想忆阻器和忆阻型生物受体——应用了滑动回归。作为探测信号，我们使用了慢变正弦输入，或称“线性调频信号”，其瞬时频率从 0.1 Hz 平滑扫频至 10 Hz。

选择合适长度的窗口后（见下文），即可通过考虑局部（窗口大小）的线性输入—输出近似来进行滚动回归，形式如下：

$$y(t) \approx m(t)\,u(t) + b(t), \tag{15}$$

其中，$m(t)$ 表示系统对其输入的瞬时灵敏度（或增益），$b(t)$ 则表示局部推断出的基线，输出围绕该基线上下波动。

Hudson 和 Landy（2012年）曾使用扫频驱动来揭示非线性动力系统如何随驱动频率变化而改变其响应。这里，它发挥着相同的作用：展示各系统的有效增益和偏置如何随不同的输入时间尺度而变化。

滑动窗口大小是根据输入的最慢分量选定的。由于慢变正弦波从 $f_0 = 0.1$ Hz 开始（周期为 $T_0 \approx 10$ s），我们采用窗口 W $= 0.4\,T_0 \approx 4$ s，步长 S $= 0.1\,T_0 \approx 1$ s。该窗口足够短，能够分辨系统有效增益和偏置的缓慢漂移；同时又足够长，确保每次局部线性拟合在数值上稳定。所有系统均采用相同的窗口参数，以确保可比性。步长设为窗口长度的四分之一，这样既能实现足够的重叠（约 75%），获得稳定的估计，又能使回归窗口逐渐沿信号移动。

[4]从概念上说，这种方法与用于追踪自适应系统中时变参数的局部或加权最小二乘估计方法有关（Joensen 等人（2000年）），但这里明确采用有限移动窗口来实现，而非递归加权。

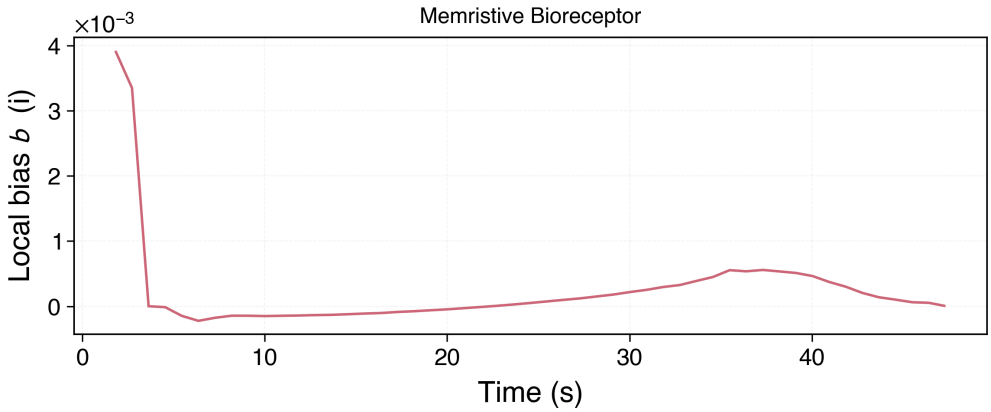


图7：通过滑动窗口回归得到的 MBR 局部偏置 $b(t)$。该偏置呈现缓慢漂移，与器件时变内部状态的特征一致。

附录A.1中的图9c展示了忆阻器生物受体的局部增益 $m(t)$ 随时间的变化。初始阶段（截至 $t \approx 5$ s）呈现出与忆阻器状态弛豫相关的预期衰减；但随着输入频率增加，该器件逐渐放大信号，使增益在 $t \approx 30-40$s 附近明显上升。作为对比，插图展示了相同输入下理想（非自适应）忆阻器对应的 $m(t)$：在初始瞬态之后，局部增益只是弛豫至一个恒定值，并保持平稳。这表明，主图中观察到的自适应行为并非忆阻器的一般属性，而是源于额外的自适应状态变量。

图7展示了局部偏置 $b(t)$ 的变化。偏置迅速稳定在零附近，随后随着器件适应而向上漂移；当内部状态开始饱和后，偏置又开始下降。理想忆阻器和恒温器系统中均未出现这一行为（见附录A.1中的图9a、9b），两者的偏置在整个过程中都基本保持恒定。忆阻型生物受体呈现出时变内部动力学的典型特征：实验过程中，局部增益和局部偏置都会发生变化，早期表现为缓慢弛豫，随后随着输入频率足够高、从而激活其适应机制，增益显著增加。

为补充基于振幅的分析，我们简要利用过零滞后来考察响应时序。为此，我们采用了双极性方波输入（类似于图2、3、4中所示的输入），因为其符号的突变提供了明确的时间参考点。

为量化时间关系，我们提取输入和输出的过零时刻（阈值设为零；在此情况下，双极性方波的均值为零，因此这一设置很自然）。每次输入符号变化都会产生一个时刻 t_u^k，相应的输出变化时刻（通过线性插值获得）记为 t_y^k。每次 k 过零时的滞后量定义为：

$$\ell^{(k)} = t_y^{(k)} - t_u^{(k)}, \tag{16}$$

with 为正值表示输出滞后于输入，为负值则表示输出超前于输入 .

在图8中，绝对滞后 $|\ell(t)|$ 呈现出数值虽小却持续增长的漂移。 x轴显示的是输入过零时间 $t_u^{(k)}$，因为只有在这些时刻才定义滞后；使用 $t_u^{(k)}$可将每次测量定位在输入符号翻转的精确时刻，确保滞后以正确的生成事件为参照。滞后 $\ell(t)$（上图）如双极性方波输入所预期的那样正负交替振荡，但振荡振幅随时间增大，表明输出开始偏离与输入完全同步的状态。下图清楚地显示了这一趋势： $|\ell^{(k)}|$ 在连续的过零时刻之间单调递增。作为参照，附录A.2中给出的恒温器和理想忆阻器对应的滞后曲线呈现出不同的行为。恒温器始终保持零滞后（图10a）；而尽管理想忆阻器存在微小且固定的时间偏移（图10b），其振荡幅度却始终不随时间变化。 $|\ell|$ 的缓慢增长是时变内部动力学的明显特征。

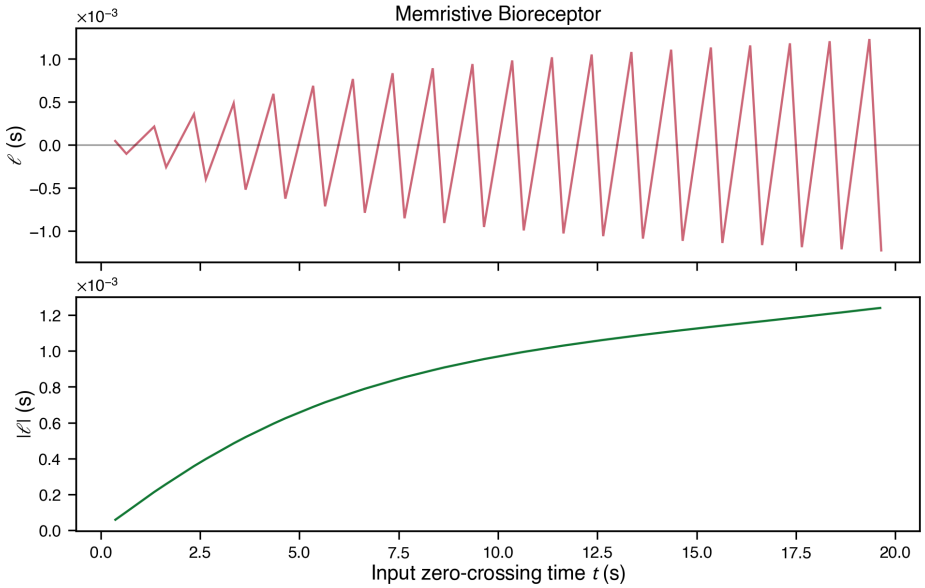


图8：在每个输入过零时刻 $t_u^{(k)}$对 MBR 进行评估所得的过零滞后。上图：带符号滞后 $\ell^{(k)}$ 呈现交替极性，这是由双极驱动造成的，且振幅随时间增大。下图：绝对滞后 $|\ell^{(k)}|$持续增大，揭示了由系统自适应动力学驱动的渐进式时序漂移。

6 信息处理顺序框架的适用性

所提出框架的效用在于，它是一种基底无关动力学，可在多种系统中接受检验。这并非旨在取代可能与信息处理架构相关的其他考量，而是提供了一种可测量的系统级计算机制，以供评估。这与常见框架十分契合；这类框架会将考量分解为计算、算法以及实现／物理等层次，参见 Marr（2010）。通过分离出核心计算属性，这种方法尤其适合用于界定系统是否表现出能动性等特质。我们认为，该框架将在三个广泛领域发挥作用：伦理考量、功能潜力以及基底无关的比较。

6.1 能动性的伦理考量

能动性的本质在伦理学中扮演着重要角色，尤其是在自由意志、道德责任和道德地位方面 Brey（2014）。道德地位是指某个实体本身具有重要性，也就是说它拥有利益。道德地位是权利的基础 Clarke 和 Savulescu（2021）。一些哲学家，如 Gosepath（2014），认为能动性是道德地位的核心组成部分 Steinhoff（2015）。然而，人们通常认为道德地位需要意识，因此能够赋予道德地位的能动性应当是有意识的能动性。但挑战在于，如何开发方法，在非典型案例（例如具身脑类器官、人工智能、神经人工智能）中识别作为道德地位组成部分的有意识能动性；在这些案例中，我们用于推断意向性的代理指标无法轻易适用。因此，开发可靠的指标，以推断某个实体是否具备具有道德意义的认知能力（如有意识的能动性），已成为实验神经伦理学的核心关注点 Kagan et al.（2024b）。先前研究提出了基于比较行为反应 Boyd 和 Lipshitz（2024）或信息流的神经架构 Boyd（2024）的策略，用于推断能够赋予道德地位的认知能力。这些尝试试图以认识论标准为道德考量奠定基础。本研究提出的形式体系为推断“最低限度能动性”的案例提供了理论基础；我们认为，这种能动性先于通常与道德地位相关联的心理表征，即有意识的能动性 Barandiaran et al.（2009）。与III类能动性相关的特质（例如时间性、适应性）可能是展现能够赋予道德地位的能力所必需的，但并不足够；这些能力包括自我导向性（或自主性）和自我意识，这些能力对于拥有持久的自我概念 Block（1995）或形成“美好生活”观念 Rawls（2003）而言是必需的。此外，赋予道德地位的能动性方面——包括意识、个人自主性和意向性——在道德上具有重要意义，因为它们体现了某个实体的评价立场，即拥有偏好、利益或价值观的能力

这些偏好、利益或价值观会引导行动 Blumenthal-Barby（2024）。简言之，拥有最低限度能动性，甚至III类能动性，并不一定会赋予道德地位，但可能是理解那些导致具有道德意义的能动性表现得以产生的因果机制所必不可少的。

6.2 系统的功能潜力

能动性通常被视为智能系统的一项关键特征。目前，构建智能系统的尝试达到了前所未有的高峰，其中最常见的是采用基于硅的计算，但也有采用生物组件的方案（Kagan and Kitchen，2025）。如果能够预测（或至少限定）系统可能表现出的行为范围——这些行为最终可能依据系统处理信息的方式产生影响——将会很有帮助。这种以影响为导向的方法，也与其他观点一致：应主要根据系统对其所嵌入的信息环境产生影响的能力来界定其能动性（Soltanzadeh，2025）。信息处理阶序框架可能发挥作用的一个特定领域，是快速发展的相关研究领域：该领域旨在利用活体生物神经培养物进行信息处理，并最终实现智能用途（Kagan et al.，2025）。尽管为与神经元培养物交互而设计的工具日益增多（Kagan，2025a；Jordan et al.，2024；Zhang et al.，2024；Cai et al.，2023），但如何评估分配给这些神经元培养物的任务，仍存在相当大的不确定性（Tanveer et al.，2025）。更复杂的是，对于如何最好地构建神经元培养物，以实现信息处理并最终实现智能，各方仍在探索不同方案（Kagan，2025b）。总体而言，一些方案侧重于生理学相关性，这种方法被称为“类器官智能（OI）”（例如 Alam El Din et al.，2025；Smirnova et al.，2023）；另一些方案则旨在利用具有不同特性的高度结构化网络，这被称为“生物工程智能（BI）”（例如 Sumi et al.，2023；Kagan，2025b）。鉴于即使是简单的神经元培养物，闭环输入与输出也可能引发动态的全网络效应（Habibollahi et al.，2023），人们将越来越关注预测性指标，以识别哪些网络结构本身具备以多种方式调节输入信息的能力。即使在最简单的层面上，能够轻松描述特定神经元培养物所执行的信息处理阶序的术语，也可作为一项有用指标，用来说明所测试任务或过程的复杂程度。

6.3 基底无关的比较

对不同系统架构进行有意义的比较十分困难，尤其是在底层基底截然不同、但表面结果看似相似的情况下 Voges et al. (2024); Dale et al. (2019)。例如，比较生物学习与机器学习方法时，始终可以对性能进行比较 Khajehnejad et al. (2025)，但这些系统的架构和内部动力学差异极大，因此通常很难进一步比较它们在每个时间步如何处理信息。采用基底无关指标的好处在于，它能够比较那些在大多数方面可能存在显著差异的系统。不过，通过识别系统内部对输入信息的一致处理方式，我们可以开展更切题的比较。以往类似的基底无关方法也提出了其他有用的指标，例如从外部环境流向内部系统的信息流 Kolchinsky and Wolpert (2018)。我们提出的方法有一个关键区别：它关注的不是系统保留了哪些信息，而是哪些信息经过了系统信息处理架构的变换。我们提出的信息处理阶次框架，也不同于名称相似、旨在量化动态系统信息处理能力（IPC）的方法 Dambre et al. (2012); Schulte to Brinke et al. (2023)。作为一种指标，IPC 可用于研究动力系统能够计算哪些多项式函数，以及完成这项任务所需的记忆 Schulte to Brinke et al. (2023)。然而，它并未探究系统在这些限制之外对信息进行的实际变换，也未建立信息处理不同阶次的框架，或说明这些阶次可能与其他系统属性有何关联。同样，其他一些信息动力学指标也着眼于低层次过程，例如系统对信息的存储、传输和修改 Voges et al. (2024)，或语义信息等其他指标 Kolchinsky and Wolpert (2018)。我们在此提出的框架并非旨在独立使用，而是可以结合现有的基底无关指标加以扩展。通过结合这些不同的基底中立信息度量与方法，我们可以找出各种可能系统在信息处理方式上的共同原则与差异。实际上，这些指标能够对系统架构处理输入信息的方式进行一致的描述，从而使我们能够在原本几乎没有共同之处的系统之间进行更有意义的比较。

6.4 对神经科学的适用性

对系统如何处理信息的探索在神经科学领域尤为突出。此前已有研究提出神经计算指标，用以量化系统在其可计算功能方面的复杂性，从而衡量其信息处理能力（Schulte to Brinke

等人，2023年）。信息处理阶序框架关注的并非某一过程最终产生的功能结果，而是系统层面发生的信息变换。这种互补视角可能至关重要，因为研究已发现，系统主动变换输入的能力是完成更复杂学习任务所必需的（Miconi 和 Kay，2025年）。大量证据表明，生物神经系统具备本文提出的所有类别的信息处理能力；不同类型系统之间的相互作用，还能以受控方式产生更复杂的行为（Toyoizumi 等人，2014年；Pariz 等人，2018年；Bastos 等人，2015年）。最后，这种基于系统的指标也能补充临床神经科学和临床前神经科学研究的其他领域。例如，扰动复杂性指数（PCI）旨在评估输入信号在整个大脑中的归一化可压缩性（Sinitsyn 等人，2020年）。研究者曾提出，PCI 分数升高可预测原本处于昏迷状态的患者是否具有意识（医学意义上的意识）（Sinitsyn 等人，2020年）。理论上，本框架所述的任何类别的系统都可能改变输入信号的可压缩性——即使只是当呈现达到某一阈值的刺激后，在整个系统中输出随机反应。然而，也许可以通过增强刺激，探查这些患者的大脑中是否存在 I类、II类或 III 类系统动态，并进一步了解原本反应极弱的患者可能保留的功能。

7 结论

通过提出一项系统必须展现的必要条件，以使其具备能动性，本研究旨在解决该领域中的分歧，并提供一个有助于讨论此类系统动力学的框架。通过提出一种可与现有自上而下的能力定义相整合的自下而上的系统动态方法，可以形成一致且实用的能动性定义。信息处理阶序框架对于讨论一个系统是否符合能动性的伦理考量具有直接影响，并将有助于更细致地讨论在符合条件时可能适用的道德考量。然而，应当承认，即使此类系统确实展现出高度能动性，这也不等同于其具有道德地位。在此，我们认为，只有随着有意识的能动性的出现，一个系统才会具有道德地位及相应的权利。不过，通过建立可检验且可证伪的框架，使能动性等术语能够得到更深入的探讨，识别表明其他具有道德相关性的特征之指标的工作也能够取得进展。此外，该框架或许有助于预测系统的功能潜力，同时支持对信息处理动态进行基底无关比较。更广泛而言，该框架也适用于神经科学，可用于描述神经系统及其子系统在处理信息时所开展的信息动力学。最后，对于试图利用体外神经培养物进行信息处理和智能研究的工作，所提出的框架提供了一种评估这些生物系统在响应结构化信息环境时所执行的信息处理阶序的方法。这将阐明哪些任务涉及能动性步骤，并将其与较不复杂的任务区分开来。归根结底，随着开发自主智能系统的兴趣、进展和投资不断增长，所提出的框架提供了一条有原则且可检验的路径，用于解决与能动性等术语相关的不确定性和现有分歧，同时阐明这些系统可能展现出的内部动力学。

7.1 致谢

作者感谢 Andy C. Kitchen 富有洞见的讨论与建议，并感谢 Nicole Kerlero de Rosbo 博士进行校对。本研究由 NUSMed 和 ODPRT（NUHSRO/2024/035/Startup/04）支持，项目名称为“实验哲学生命伦理学与关系道德心理学”，BDE 担任首席研究员。本研究由新加坡国家研究基金会通过其 AI Singapore 计划（AISG 奖项编号：AISG3-GV-2023-012）提供支持。本研究项目由新加坡国立大学 NUS 初创基金支持；（NUHSRO/2022/078/Startup/13）。本研究得到惠康信托的支持 [资助编号： 226801]，用于伦理学与人文研究变革性包容性探索研究平台（ANTITHESES）。为实现开放获取，作者已针对本次投稿所产生的任何作者接受稿版本申请知识共享署名许可协议公共版权许可。本项目由美国国家科学基金会题为“EFRI BEGIN OI：为具身类器官智能带来奖励”的基金支持（FAIN 编号：2515214），JLB 担任共同首席研究员。

附录A

A.1 滚动回归

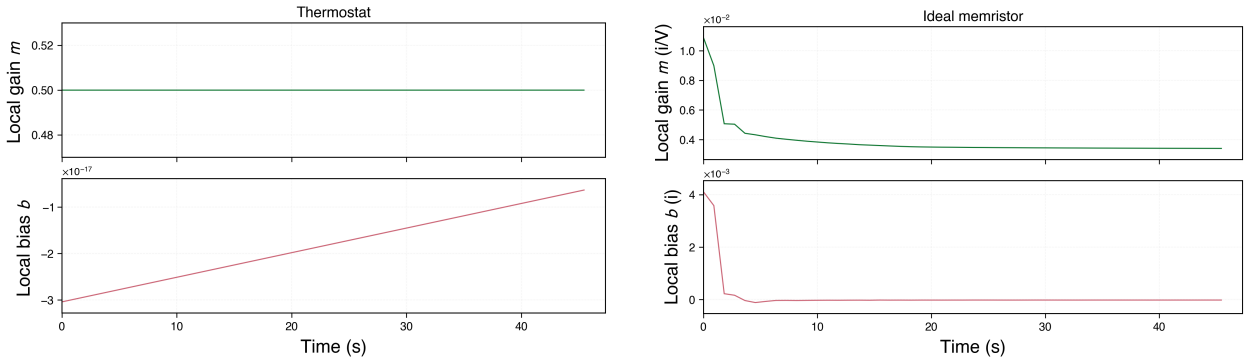
为计算局部增益 $m(t)$ 和局部偏置 $b(t)$，我们采用了第5节.2中所述的滑动窗口线性回归方法。在每个时间窗口内，我们将输出对输入进行回归，从而获得系统对刺激（$m(t)$）响应强度以及围绕哪个基线水平（$b(t)$）的随时间变化的估计值。

一个关键考量是窗口长度的选择。如果窗口过短，该时间区间内的输入可能变化不足，尤其是在低频片段期间，从而导致估计不稳定或噪声较大。如果窗口过长，变换中真实的时间演化可能会被平均掉，自适应效应也可能被人为地平滑掉。

由于驱动信号是一个缓慢变化的正弦扫描信号，其频率跨度接近两个数量级（0.1 Hz $\rightarrow$ 10 Hz），因此我们根据其中的最低频率来设定窗口大小。最低频率片段为$f_0 = 0.1$ Hz，对应的周期为 $T_0 = 10$s。4 秒的窗口（$W \approx 0.4\,T_0$）确保满足两个重要条件：

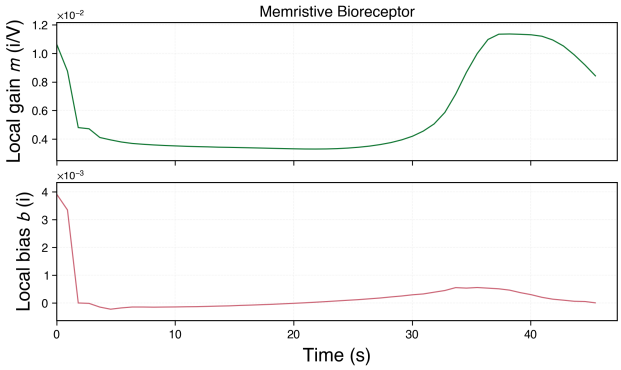
- 在最低频率范围内，每个窗口都覆盖一个周期中有意义的一部分，提供足够的变化，以便进行条件良好的线性拟合。
- 对于中频和高频范围，每个窗口覆盖多个周期，从而降低对 m（t）和 b（t）的估计结果的方差。

窗口每次向前移动约 $S = 1$s（约 $0.1\,T_0$）。这种设置能够进行密集的时间采样，使相邻窗口之间有较大程度的重叠，从而连续追踪估计增益和偏置的缓慢漂移，同时避免不必要的冗余。



(a) 恒温器

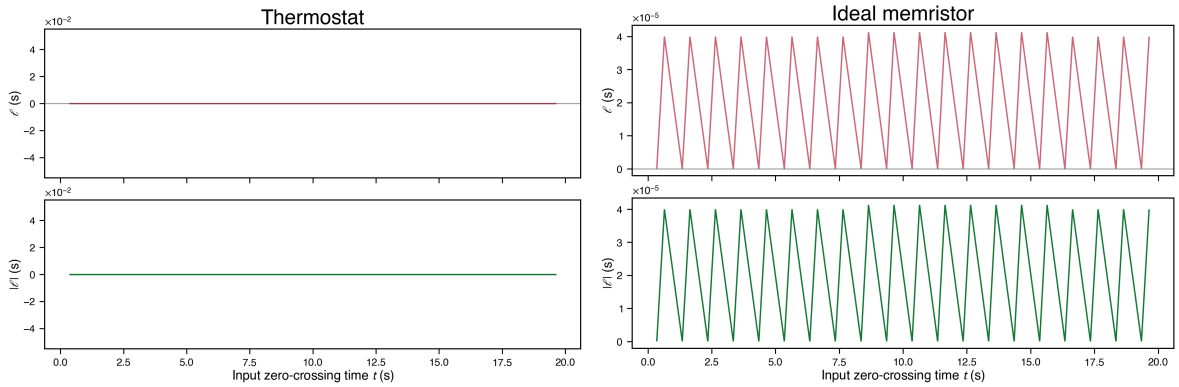
(b) 理想忆阻器



(c) 忆阻型生物受体

图 9：所有系统的滚动增益与偏置：恒温器保持不变，理想忆阻器缓慢弛豫，只有 MBR 表现出真正的时变特性。

所有系统（恒温器、理想忆阻器和忆阻型生物受体）均使用相同的窗口和步长参数，以确保分析设置一致。图 9 展示了这三个系统完整的滚动增益与偏置轨迹，体现了上述比较的依据。恒温器始终表现出恒定的局部增益和偏置；理想忆阻器呈现缓慢的状态依赖弛豫，但没有适应；而忆阻器生物受体的这两个参数均表现出显著的时间演化，反映了其自适应内部动力学。



(a) 恒温器 (b) 理想忆阻器

图 10：恒温器和理想忆阻器的过零滞后。恒温器在所有过零点的滞后均为零，而理想忆阻器则表现出一个较小但固定的偏移量，其大小随时间保持不变。

A.2 过零滞后

图 10 展示了双极性方波输入驱动下，恒温器和理想忆阻器的过零滞后。正如预期，恒温器在所有过零点的滞后均为零，而理想忆阻器则表现出较小但固定的时间偏移量，其幅度随时间保持不变。这两个系统均未表现出忆阻型生物受体所特有的渐进漂移 $|\ell^{(k)}|$。它们的时序对齐始终保持稳定。

参考文献

Hong Yu Wong。探究人工能动性。《心理学前沿》，15:1449320，2025年1月。ISSN 1664-1078。doi:10.3389/fpsyg.2024.1449320。网址 https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1449320/full。Brett J. Kagan 和 Andy C. Kitchen。为什么人工智能的进步将必然需要利用合成生物学来借助智能的真实基础。载于 Ulrich A. K. Betz 编，《科学创造更美好的明天，第195–213页。施普林格·自然瑞士，查姆，2025年。ISBN 978-3-031-93622-7 978-3-031-93623-4。doi:10.1007/978-3-031-93623-4_12。网址 https://link.springer.com/10.1007/978-3-031-93623-4_12。Brett J. Kagan、Adeel Razi、Anjali Bhat、Andy C. Kitchen、Nhi T. Tran、Forough Habibollahi、Moein Khajehnejad、Bradyn J. Parker、Ben Rollo 和 Karl J. Friston。科学传播与感知能力的语义。《神经元》，111(5):606–607，2023年3月。ISSN 08966273。doi:10.1016/j.neuron.2023.02.008。网址 https://linkinghub.elsevier.com/retrieve/pii/S0896627323001125。Brett J. Kagan、Michael Mahlis、Anjali Bhat、Josh Bongard、Victor M. Cole、Phillip Corlett、Christopher Gyngell、Thomas Hartung、Bianca Jupp、Michael Levin、Tamra Lysaght、Nicholas Opie、Adeel Razi、Lena Smirnova、Ian Tennant、Peter Thestrup Wade 和 Ge Wang。迈向多样化智能系统的命名共识：呼吁合作。《创新》，5(5):100658，2024年9月。ISSN 26666758。doi:10.1016/j.xinn.2024.100658。网址 https://linkinghub.elsevier.com/retrieve/pii/S2666675824000961。Leonard Dung。理解人工能动性。《哲学季刊》，75(2):450–472，2025年3月。ISSN 0031-8094, 1467-9213。doi:10.1093/pq/pqae010。网址 https://academic.oup.com/pq/article/75/2/450/ 7601099。Luciano Floridi。作为无智能能动性的人工智能：论人工智能作为一种新形式的人工能动性及能动性多重实现论题。《哲学与技术》，38(1):30，s13347–025–00858–9，2025年3月。ISSN 2210-5433, 2210-5441。doi:10.1007/s13347-025-00858-9。网址 https://link.springer.com/10. 1007/s13347-025-00858-9。Maud Van Lier。介绍概念化人工能动性的四分法。*Synthese*，201(3):85，2023年2月。ISSN 1573-0964。doi:10.1007/s11229-023-04083-9。网址 https://link.springer.com/10.1007/ s11229-023-04083-9。Sai Dattathrani 和 Rahul De'。人工智能时代的能动性概念：维度与程度。《信息系统前沿》，25(1):29–54，2023年2月。ISSN 1387-3326, 1572-9419。doi:10.1007/s10796-022- 10336-8。网址 https://link.springer.com/10.1007/s10796-022-10336-8。

Brett J. Kagan、Alon Loeffler、J. Lomax Boyd 和 Julian Savulescu。具身神经系统可支持对道德相关状态的迭代研究。《神经科学杂志》，44(15):e0431242024，2024年4月。ISSN 0270-6474, 1529-2401。doi:10.1523/JNEUROSCI.0431-24.2024。网址 https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0431-24.2024。Miguel Ángel Sebastián。人工智能中的第一人称表征与负责任的能动性。*Synthese*，199(3):7061–7079，2021年12月。ISSN 1573-0964。doi:10.1007/s11229-021-03105-8。网址 https://doi.org/10.1007/ s11229-021-03105-8。Gordana Dodig-Crnkovic 和 Mark Burgin。自主智能体的系统化方法。*Philosophies*，9(2):44，2024年3月。ISSN 2409-9287。doi:10.3390/philosophies9020044。网址 https://www.mdpi.com/2409-9287/9/ 2/44。Xabier E. Barandiaran、Ezequiel Di Paolo 和 Marieke Rohde。定义能动性：行动中的个体性、规范性、不对-称性与时空性。《适应性行为》，2009年9月。doi:10.1177/1059712309343819。网址 https://journals.sagepub.com/doi/10.1177/1059712309343819。出版商：SAGE PublicationsSage UK：伦敦，英格兰。Andrew Pickering。什么是能动性？科学研究与控制论的视角。《生物学理论》，19(1): 16–21，2024年3月。ISSN 1555-5550。doi:10.1007/s13752-023-00437-1。网址 https://doi.org/10.1007/ s13752-023-00437-1。Sonia E. Sultan、Armin P. Moczek 和 Denis Walsh。弥合解释性缺口：我们能从生物能动性视角学到什么？*BioEssays*，44(1):2100185，2022年1月。ISSN 0265-9247, 1521-1878。doi:10.1002/bies.202100185。网址 https://onlinelibrary.wiley.com/doi/10.1002/bies.202100185。Parashar Das。人工智能系统中的能动性。2025年。Ali Barzegar、Emilia Margoni 和 Daniele Oriti。物理学中能动性的极简论述。Danielle Swanepoel。人工智能具有能动性吗？载于 Robert W. Clowes、Klaus Gärtner 和 Inês Hipólito 编，《心智—技术问题：探究心智、自我与21世纪人工制品，第83–104页。Springer International Publishing，查姆，2021年。ISBN 978-3-030-72644-7。doi:10.1007/978-3-030-72644-7_4。网址 https://doi.org/10.1007/978-3-030-72644-7_4。Danielle Swanepoel 和 Daniel Corks。人工智能与能动性：人工智能决策中的平局裁决。《科学与工程伦理》，30(2):11，2024年3月。ISSN 1471-5546。doi:10.1007/s11948-024-00476-2。网址 https://doi.org/10.1007/s11948-024-00476-2。Luciano Floridi 和 J W Sanders。论人工智能体的道德性。2004年。Carissa Véliz。道德僵尸：为什么算法不是道德主体。《人工智能与社会》，36(2):487–497，2021年6月。ISSN 1435-5655。doi:10.1007/s00146-021-01189-x。网址 https://doi.org/10.1007/s00146-021-01189-x。Sven Nyholm。人机交互伦理与传统道德理论。2021年。网址 https:// academic.oup.com/edited-volume/37078/chapter/323166315。Henry Shevlin。我们如何知道机器人何时会成为道德承受者？《剑桥医疗保健伦理季刊》，30(3):459–471，2021年7月。ISSN 0963-1801, 1469-2147。doi:10.1017/S0963180120001012。网址 https: // www.cambridge.org/core/product/identifier/S0963180120001012/type/journal_article。Miguel Aguilera、Carlos Alquézar 和 Manuel G. Bedia。最小感觉运动模型中的能动性与整合信息。载于*2018*年人工生命会议，第396–403页，日本东京，2018年。麻省理工学院出版社。doi:10.1162/isal_a_00077。网址 https://www.mitpressjournals.org/doi/abs/10.1162/isal_a_00077。Ruth M.J. Byrne。反事实思维。《心理学年度评论》，67(1):135–157，2016年1月。ISSN 0066-4308, 1545-2085。doi:10.1146/annurev-psych-122414-033249。网址 https://www.annualreviews.org/doi/10.1146/annurev-psych-122414-033249。Joseph De La Torre Dwyer。法兰克福的能动性反事实干预的根本失败：不存在能动性。*Philosophia*，49(2):633–642，2021年4月。ISSN 0048-3893, 1574-9274。doi:10.1007/s11406-020-00240-3。网址 https://link.springer.com/10.1007/s11406-020-00240-3。Jorge Ramírez-Ruiz、Dmytro Grytskyy、Chiara Mastrogiuseppe、Yamen Habib 和 Rubén Moreno-Bote。从内在动机到占据未来行动状态路径空间的复杂行为。《自然·通讯》，15(1):6368，2024年7月。ISSN 2041-1723。doi:10.1038/s41467-024-49711-1。网址 https://www.nature.com/articles/ s41467-024-49711-1。Eugenia Kulakova、Nima Khalighinejad 和 Patrick Haggard。我本可以做出不同选择：反事实比较的可用性影响能动感。《意识与认知》，49:237–244，2017年3月。ISSN 10538100。doi:10.1016/j.concog.2017.01.013。网址 https://linkinghub.elsevier.com/retrieve/ pii/S1053810016302070。

Malte Hendrickx。《主体控制的行动：因果性的，而非反事实的》。《哲学研究》180(10-11)：3121–3139，2023年11月。ISSN 0031-8116，1573-0883。doi:10.1007/s11098-023-02033-2。网址 https: // link.springer.com/10.1007/s11098-023-02033-2。Pradipta Kishore Chakrabarty。《智能体人工智能中的因果推断：弥合可解释性与动态决策之间的鸿沟》。《国际科学与研究杂志》(IJSR)，14(4):2112–2117，2025年4月。ISSN 23197064。doi:10.21275/SR25424081718。网址 https://www.ijsr.net/getabstract.php?paperid= SR25424081718。Mark Miller、Julian Kiverstein 和 Erik Rietveld。《幸福与福祉的预测动力学》。《情绪评论》14(1):15–30，2022年1月。ISSN 1754-0739，1754-0747。doi:10.1177/17540739211063851。网址 https://journals.sagepub.com/doi/10.1177/17540739211063851。David Abel、André Barreto、Michael Bowling、Will Dabney、Shi Dong、Steven Hansen、Anna Harutyunyan、Khimya Khetarpal、Clare Lyle、Razvan Pascanu、Georgios Piliouras、Doina Precup、Jonathan Richens、Mark Rowland、Tom Schaul 和 Satinder Singh。《能动性依赖于参照框架》，2025年2月。网址 http://arxiv.org/abs/2502. 04403。arXiv:2502.04403 [cs]。Erik Miehling、Karthikeyan Natesan Ramamurthy、Kush R. Varshney、Matthew Riemer、Djallel Bouneffouf、John T. Richards、Amit Dhurandhar、Elizabeth M. Daly、Michael Hind、Prasanna Sattigeri、Dennis Wei、Ambrish Rawat、Jasmina Gajcin 和 Werner Geyer。《智能体人工智能需要系统理论》，2025年2月。网址 http://arxiv. org/abs/2503.00237。arXiv:2503.00237 [cs]。Karl Johan Åström 和 Richard Murray。《反馈系统：面向科学家与工程师的导论，第二版》。普林斯顿大学出版社，2021年2月。ISBN 978-0-691-21347-7。Google-Books-ID: qZ0DEAAAQBAJ。Wolfgang Hofkirchner 和 Matthias Schafranek。一般系统理论。载于 Cliff Hooker 编，《复杂系统哲学》，《科学哲学手册》第10卷，第177–194页。北荷兰出版社，阿姆斯特丹，2011年1月。doi:10.1016/B978-0-444-52076-0.50006-7。网址 https://www.sciencedirect.com/science/ article/pii/B9780444520760500067。Karl Friston、Rick Adams 和 Read Montague。价值是什么——累积奖励还是证据? 《神经机器人学前沿》6，2012年。ISSN 1662-5218。doi:10.3389/fnbot.2012.00011。网址 http://journal.frontiersin. org/article/10.3389/fnbot.2012.00011/abstract。Brett J. Kagan、Forough Habibollahi、Brad Watmuff、Azin Azadi、Finn Doensen、Alon Loeffler、Seung Hoon Byun、Bram Servais、Candice Desouza、Kwaku Dad Abu-Bonsrah 和 Nicole Kerlero de Rosbo。利用体外脑细胞的智能。《神经科学家》，2025年。Thomas Miconi 和 Kenneth Kay。可塑性神经网络中的关系学习与快速知识重组的神经机制。《自然-神经科学》28(2):406–414，2025年2月。ISSN 1097-6256，1546-1726。doi:10.1038/s41593-024-01852-8。网址 https://www.nature.com/articles/s41593-024-01852-8。Sergio Bittanti 和 Patrizio Colaneri。《周期系统》。通信与控制工程。施普林格，伦敦，2009年。ISBN 978-1-84800-910-3 978-1-84800-911-0。doi:10.1007/978-1-84800-911-0。网址 http://link. springer.com/10.1007/978-1-84800-911-0。ISSN: 0178-5354。Brett J. Kagan、Daniela Duc、Ian Stevens 和 Frederic Gilbert。虚拟世界中的具身神经元：类器官伦理学的证据? 《美国生物伦理学杂志：神经科学》13(2):114–117，2022年4月。ISSN 2150-7740，2150-7759。doi:10.1080/21507740.2022.2048731。网址 https://www.tandfonline.com/doi/full/10.1080/21507740. 2022.2048731。Nicolas Burra、Alexis Hervais-Adelman、Alessia Celeghin、Beatrice de Gelder 和 Alan J. Pegna。情感盲视依赖于低空间频率。《神经心理学》128:44–49，2019年5月。ISSN 00283932。doi:10.1016/j.neuropsychologia.2017.10.009。网址 https://linkinghub.elsevier.com/retrieve/pii/ S0028393217303780。Berit Brogaard。是否存在无意识知觉过程? 《意识与认知》20(2):449–463，2011年6月。ISSN 10538100。doi:10.1016/j.concog.2010.10.002。网址 https://linkinghub.elsevier.com/retrieve/ pii/S105381001000190X。L. Chua。忆阻器——缺失的电路元件。《IEEE电路理论汇刊》18(5):507–519，1971年9月。ISSN 2374-9555。doi:10.1109/TCT.1971.1083337。网址 https://ieeexplore.ieee.org/document/ 1083337/。Francesco Caravelli 和 Juan Pablo Carbajal。《写给好奇的外行人的忆阻器》，2018年12月。网址 http: //arxiv.org/abs/1812.03389。arXiv:1812.03389 [cs]。Francesco Caravelli。忆阻网络的舞台呈现：有效记忆、动力学与学习。《国际并行、涌现与分布式系统杂志》33(4):350–366，2018年7月。ISSN 1744-5760。

doi:10.1080/17445760.2017.1320796。网址https://doi.org/10.1080/17445760.2017.1320796。出版商：Taylor & Francis _电子版：https://doi.org/10.1080/17445760.2017.1320796。

Atthaphon Viriyopase、Ingo Bojak、Magteld Zeitler 和 Stan Gielen。长程零时滞同步在皮层网络中何时可行。*Frontiers in Computational Neuroscience*，第 6 卷，2012年7月。ISSN 1662-5188。doi:10.3389/fncom.2012.00049。网址 https://www.frontiersin.org/journals/ computational-neuroscience/articles/10.3389/fncom.2012.00049/full。出版方：Frontiers。

Ardi Tampuu、Tambet Matiisen、Hau ¨dur Freyja Ólafsdóttir、Caswell Barry 和 Raul Vicente。利用深度循环网络高效解码自我位置。*bioRxiv*，2018年1月。doi:10.1101/242867。网址 http://biorxiv.org/lookup/doi/10.1101/242867。

Manu Srinath Halvagal 和 Friedemann Zenke。Hebb可塑性与预测性可塑性的结合能够在深层感觉网络中学习不变的物体表征。《自然-神经科学》，26(11):1906–1915，2023年11月。ISSN 1097-6256，1546-1726。doi:10.1038/s41593-023-01460-y。网址 https://www.nature.com/articles/ s41593-023-01460-y。

Josef Ladenbauer、Moritz Augustin 和 Klaus Obermayer。适应电流如何改变神经元放电的阈值、增益和变异性。《神经生理学杂志》，111(5):939–953，2014年3月。ISSN 0022-3077。doi:10.1152/jn.00586.2013。网址https://journals.physiology.org/doi/full/10.1152/jn.00586.2013。出版方：美国生理学会。

Jan Benda。神经适应。《当代生物学》，31(3):R110–R116，2021年2月。ISSN 0960- 9822。doi:10.1016/j.cub.2020.11.054。网址 https://www.sciencedirect.com/science/article/pii/S096098222031767X。Josef Ladenbauer、Moritz Augustin、LieJune Shiau 和 Klaus Obermayer。适应电流对耦合指数积分发放神经元同步性的影响。《*PLOS*计算生物学》，8(4):e1002478，2012年4月。ISSN 1553-7358。doi:10.1371/journal.pcbi.1002478。网址 https://journals.plos.org/ploscompbiol/ article?id=10.1371/journal.pcbi.1002478。出版方：公共科学图书馆。Jan Benda 和 Andreas V. M. Herz。尖峰频率适应的通用模型。《神经计算》，15(11): 2523–2564，2003年11月。ISSN 0899-7667。doi:10.1162/089976603322385063。网址 https://doi.org/10. 1162/089976603322385063。Yuriy V. Pershin 和 Massimiliano Di Ventra。复杂材料和纳米尺度系统中的记忆效应。《物理学进展》，60(2):145–227，2011年4月。ISSN 0001-8732。doi:10.1080/00018732.2010.544961。网址 https://doi.org/10.1080/00018732.2010.544961。出版方：Taylor & Francis_电子版：https://doi.org/10.1080/00018732.2010.544961。James P. Sethna、Karin Dahmen、Sivan Kartha、James A. Krumhansl、Bruce W. Roberts 和 Joel D. Shore。滞后与层级：无序驱动的一阶相变动力学。《物理评论快报》，70(21): 3347–3350，1993年5月。ISSN 0031-9007。doi:10.1103/PhysRevLett.70.3347。网址 https://link.aps.org/doi/ 10.1103/PhysRevLett.70.3347。Nathan C. Keim、Joseph D. Paulsen、Zorana Zeravcic、Srikanth Sastry 和 Sidney R. Nagel。物质中的记忆形成。《现代物理评论》，91(3):035002，2019年7月。doi:10.1103/RevModPhys.91.035002。网址 https://link.aps.org/doi/10.1103/RevModPhys.91.035002。出版方：美国物理学会。Joseph T. Lizier、Mikhail Prokopenko 和 Albert Y. Zomaya。复杂分布式计算中信息存储的局部度量。《信息科学》，208:39–54，2012年11月。ISSN 0020- 0255。doi:10.1016/j.ins.2012.04.016。网址 https://www.sciencedirect.com/science/article/pii/ S0020025512002800。时间序列滚动分析。载于 Eric Zivot 和 Jiahui Wang（编），使用*S-PLUS*®建模金融时间序列，第313–360页。施普林格，纽约州纽约市，2006年。ISBN 978-0-387-32348-0。doi:10.1007/978-0-387-32348-0_9。网址 https://doi.org/10.1007/978-0-387-32348-0_9。Alfred Joensen、Henrik Madsen、Henrik Aa. Nielsen 和 Torben S. Nielsen。通过局部回归跟踪时变参数。《自动化》，36(8):1199–1204，2000年8月。ISSN 0005-1098。doi:10.1016/S0005-1098(00)00029- 7。网址 https://www.sciencedirect.com/science/article/pii/S0005109800000297。Todd E. Hudson 和 Michael S. Landy。使用正弦扰动函数测量适应。《神经科学方法杂志》，208(1):48–58，2012年6月。ISSN 0165-0270。doi:10.1016/j.jneumeth.2012.04.001。网址 https://www.sciencedirect.com/science/article/pii/S0165027012001240。David Marr。视觉：对人类视觉信息表征与处理的计算研究。麻省理工学院出版社，2010年。ISBN 0-262-28898-2。

Philip Brey。从道德主体到道德因素：结构伦理学方法。载于 Peter Kroes 和 Peter- Paul Verbeek（编），技术制品的道德地位，第125–142页。Springer Netherlands，荷兰多德雷赫特，2014年。ISBN 978-94-007-7914-3。doi:10.1007/s1007-7914-3_8。网址 https://doi.org/10.1007/ 978-94-007-7914-3_8。Steve Clarke 和 Julian Savulescu。重新思考我们对道德地位的假设。载于 Steve Clarke、Hazem Zohny 和 Julian Savulescu（编），重新思考道德地位，惠康信托基金资助的专著和书籍章节。牛津大学出版社，英国牛津，2021年。网址 http://www.ncbi.nlm.nih.gov/books/NBK572928/。Uwe Steinhoff。所有人是否都具有平等的道德价值？论基本平等以及平等的尊重与关怀。牛津大学出版社，牛津，2015年。ISBN 978-0-19-871950-2。J. Lomax Boyd 和 Nethanel Lipshitz。意识维度与脑类器官的道德地位。神经伦理学，17(1):5，2024年4月。ISSN 1874-5490，1874-5504。doi:10.1007/s12152-023-09538-x。网址 https: //link.springer.com/10.1007/s12152-023-09538-x。J Lomax Boyd。从计算架构角度探讨脑类器官的道德可考虑性。*OxfordOpen Neuroscience*，3:kvae004，2024年2月。ISSN 2753-149X。doi:10.1093/oons/kvae004。网址 https: //doi.org/10.1093/oons/kvae004。Ned Block。关于意识功能的一种混淆。*BehavioralandBrainSci-ences*，18(2):227–247，1995年6月。ISSN 1469-1825，0140-525X。doi:10.1017/S0140525X00038188。网址 https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/abs/on-a-confusion-about-a-function-of-consciousness/061422BF0C50C5FF00927F9B6E879413。John Rawls。正义论。哈佛大学出版社贝尔纳普出版社，马萨诸塞州剑桥，修订版，第5–6次印刷版，2003年。ISBN 978-0-674-00078-0 978-0-674-00077-3。Jennifer Blumenthal-Barby。人格的终结。《美国生物伦理学杂志：*AJOB*》，24(1):3–12，2024年1月。ISSN 1536-0075。doi:10.1080/15265161.2022.2160515。Sadjad Soltanzadeh。技术治理中的能动性形而上学阐释。《人工智能与社会》，40(3):1723–1734，2025年3月。ISSN 0951-5666，1435-5655。doi:10.1007/s00146-024-01941-z。网址 https://link.springer. com/10.1007/s00146-024-01941-z。Brett J. Kagan。CL1：借助生物神经系统功能的平台技术。*Nature ReviewsBioengineering*，2025年7月。ISSN 2731-6092。doi:10.1038/s44222-025-00340-3。网址 https://www.nature. com/articles/s44222-025-00340-3。出版方：Springer Science and Business Media LLC。Fred D. Jordan、Martin Kutter、Jean-Marc Comby、Flora Brozzi 和 Ewelina Kurtys。开放且可远程访问的神经平台：用于湿件计算的研究。*Frontiers in ArtificialIntelligence*，7:1376042，2024年5月。ISSN 2624-8212。doi:10.3389/frai.2024.1376042。网址 https://www.frontiersin.org/articles/10.3389/frai. 2024.1376042/full。Xiaotian Zhang、Zhi Dou、Seung Hyun Kim、Gaurav Upadhyay、Daniel Havert、Sehong Kang、Kimia Kazemi、Kai- Yu Huang、Onur Aydin、Raymond Huang、Saeedur Rahman、Austin Ellis-Mohr、Hayden A. Noblet、Ki H. Lim、Hee Jung Chung、Howard J. Gritton、M. Taher A. Saif、Hyun Joon Kong、John M. Beggs 和 Mattia Gazzola。体外心智平台：用于连接活神经元的多功能、可扩展、稳健且开放的解决方案。*Advanced Science*，11(11):2306826，2024年3月。ISSN 2198-3844，2198-3844。doi:10.1002/advs.202306826。网址 https://onlinelibrary.wiley.com/doi/10.1002/advs.202306826。Hongwei Cai、Zheng Ao、Chunhui Tian、Zhuhao Wu、Hongcheng Liu、Jason Tchieu、Mingxia Gu、Ken Mackie 和 Feng Guo。用于人工智能的脑类器官储备池计算。*Nature Electronics*，6(12):1032– 1039，2023年12月。ISSN 2520-1131。doi:10.1038/s41928-023-01069-w。网址 https://www.nature.com/ articles/s41928-023-01069-w。Md Sayed Tanveer、Dhruvik Patel、Hunter E. Schweiger、Kwaku Dad Abu-Bonsrah、Brad Watmuff、Azin Azadi、Sergey Pryshchep、Karthikeyan Narayanan、Christopher Puleo、Kannathal Natarajan、Mohammed A. Mostajo-Radji、Brett J. Kagan 和 Ge Wang。从零开始建立合成生物智能实验室。*Patterns*，6(5):101232，2025年5月。ISSN 26663899。doi:10.1016/j.patter.2025.101232。网址 https://linkinghub.elsevier.com/retrieve/pii/S2666389925000807。Brett J. Kagan。两条道路分岔：利用神经细胞培养物智能的途径。*Cell Biomaterials*，1(8):100156，2025年9月。ISSN 30505623。doi:10.1016/j.celbio.2025.100156。网址 https://linkinghub. elsevier.com/retrieve/pii/S3050562325001473。Dowlette-Mary Alam El Din、Leah Moenkemoeller、Alon Loeffler、Forough Habibollahi、Jack Schenkman、Amitav Mitra、Tjitse Van Der Molen、Lixuan Ding、Jason Laird、Maren Schenke、Erik C. Johnson、Brett J. Kagan、Thomas Hartung 和 Lena Smirnova。人类神经类器官微生理系统展现了构成基础的

对基础学习和记忆至关重要。《通讯生物学》，8(1):1237，2025年8月。ISSN 2399-3642。doi:10.1038/s42003-025-08632-5。网址 https://www.nature.com/articles/s42003-025-08632-5。

Lena Smirnova、Brian S. Caffo、David H. Gracias、Qi Huang、Itzy E. Morales Pantoja、Bohao Tang、Donald J. Zack、Cynthia A. Berlinicke、J. Lomax Boyd、Timothy D. Harris、Erik C. Johnson、Brett J. Kagan、Jeffrey Kahn、Alysson R. Muotri、Barton L. Paulhamus、Jens C. Schwamborn、Jesse Plotkin、Alexander S. Szalay、Joshua T. Vogelstein、Paul F. Worley 和 Thomas Hartung。类器官智能（OI）：生物计算与培养皿智能的新前沿。《科学前沿》，第1卷，2023年2月。ISSN 2813-6330。doi:10.3389/fsci.2023.1017235。网址 https: //www.frontiersin.org/journals/science/articles/10.3389/fsci.2023.1017235/full。出版方：Frontiers。

Takuma Sumi、Hideaki Yamamoto、Yuichi Katori、Koki Ito、Satoshi Moriya、Tomohiro Konno、Shigeo Sato 和 Ayumi Hirano-Iwata。生物神经元在储备池计算中充当泛化滤波器。《美国国家科学院院刊》，120(25):e2217008120，2023年6月。doi:10.1073/pnas.2217008120。网址 https: //doi.org/10.1073/pnas.2217008120。出版方：《美国国家科学院院刊》。Forough Habibollahi、Brett J. Kagan、Anthony N. Burkitt 和 Chris French。在具身体外神经网络中呈现结构化信息期间，会产生临界动力学。《自然·通讯》，14(1):5287，2023年8月。ISSN 2041-1723。doi:10.1038/s41467-023-41020-3。网址 https://www.nature.com/articles/ s41467-023-41020-3。Nicole Voges、Vinicius Lima、Johannes Hausmann、Andrea Brovelli 和 Demian Battaglia。将神经回路功能分解为信息处理基元。《神经科学杂志：神经科学学会官方期刊》，44(2):e0157232023，2024年1月。ISSN 1529-2401。doi:10.1523/JNEUROSCI.0157-23.2023。

Matthew Dale、Julian F. Miller、Susan Stepney 和 Martin A. Trefzer。用于表征储备池计算机的底物无关框架。英国皇家学会会刊*A*辑：数学、物理与工程科学，475(2226):20180723，2019年6月。ISSN 1364-5021，1471-2946。doi:10.1098/rspa.2018.0723。网址 https://royalsocietypublishing.org/doi/10.1098/rspa.2018.0723。

Moein Khajehnejad、Forough Habibollahi、Alon Loeffler、Aswin Paul、Adeel Razi 和 Brett J. Kagan。神经元培养物中的动态网络可塑性与样本效率：与深度学习的比较。《赛博格与仿生系统》，2025年6月。ISSN 2692-7632。doi:10.34133/cbsystems.0336。网址 https://spj.science.org/ doi/10.34133/cbsystems.0336。出版方：美国科学促进会（AAAS）。

Artemy Kolchinsky 和 David H. Wolpert。语义信息、自主能动性与非平衡统计物理学。《界面聚焦》，8(6):20180041，2018年12月。ISSN 2042-8898。doi:10.1098/rsfs.2018.0041。

Joni Dambre、David Verstraeten、Benjamin Schrauwen 和 Serge Massar。动力系统的信息处理能力。《科学报告》，2(1):514，2012年7月。ISSN 2045-2322。doi:10.1038/srep00514。网址 https://www.nature.com/articles/srep00514。出版商：自然出版集团。

Tobias Schulte to Brinke、Michael Dick、Renato Duarte 和 Abigail Morrison。《精细化的信息处理能力指标可深入分析神经计算系统中记忆与非线性的权衡》。《科学报告》，13(1):10517，2023年6月。ISSN 2045-2322。doi:10.1038/s41598-023-37604-0。网址 https: //www.nature.com/articles/s41598-023-37604-0。出版商：自然出版集团。Taro Toyoizumi、Megumi Kaneko、Michael P. Stryker 和 Kenneth D. Miller。《希伯定律可塑性与稳态可塑性的动态交互建模》。《神经元》，84(2):497–510，2014年10月。ISSN 08966273。doi:10.1016/j.neuron.2014.09.036。网址 https://linkinghub.elsevier.com/retrieve/pii/ S0896627314008940。

Aref Pariz、Zahra G. Esfahani、Shervin S. Parsi、Alireza Valizadeh、Santiago Canals 和 Claudio R. Mirasso。高频神经元决定神经元网络中的有效连接。《神经影像》，166:349–359，2018年2月。ISSN 10538119。doi:10.1016/j.neuroimage.2017.11.014。网址 https://linkinghub.elsevier.com/ retrieve/pii/S1053811917309217。

André Moraes Bastos、Julien Vezoli、Conrado Arturo Bosman、Jan-Mathijs Schoffelen、Robert Oostenveld、Jar- rod Robert Dowdall、Peter De Weerd、Henry Kennedy 和 Pascal Fries。视觉区通过不同频率通道产生前馈和反馈影响。《神经元》，85(2):390–401，2015年1月。ISSN 08966273。doi:10.1016/j.neuron.2014.12.018。网址 https://linkinghub.elsevier.com/retrieve/pii/ S089662731401099X。Dmitry O. Sinitsyn、Alexandra G. Poydasheva、Ilya S. Bakulin、Liudmila A. Legostaeva、Elizaveta G. Iazeva、Dmitry V. Sergeev、Anastasia N. Sergeeva、Elena I. Kremneva、Sofya N. Morozova、Dmitry Yu. Lagoda、Silvia Casarotto、Angela Comanducci、Yulia V. Ryabinkina、Natalia A. Suponeva 和 Michael A. Piradov。使用扰动复杂度指数检测无反应患者的意识潜能。《脑科学》，10(12):917，

2020年11月。ISSN 2076-3425。doi:10.3390/brainsci10120917。网址 https://www.mdpi.com/2076-3425/ 10/12/917。