

物理 AI 的“伪接地”幻象：表意 AI 如何真正解决符号接地问题

The Illusion of "Groundedness" in Physical AI: How Logographic AI Can Truly Solve the Symbol Grounding Problem

摘要

2026 年，黄仁勋在 CES 宣告“物理 AI 的 ChatGPT 时刻已经到来”[5]，随后，Figure 03 机器人在一场公开直播中连续作业超过 100 小时，处理了超过 13 万个包裹[9][10]。然而，这场旨在证明其自主性的直播，却因机器人暴露出的 0.5 秒动作延迟、异常“摸头”行为、包裹掉落等细节，引发了“自主运行还是远程操控”的广泛质疑；与此同时，智元机器人成为全球首个突破万台量产的具身智能企业[11]。物理 AI 的“可行性”不再是疑问，真正的竞争已进入“可用性”与“经济性”的深水区。然而，一个更深层的问题被狂欢掩盖：当机器人“摸”到包裹、抓取物体、避开障碍时，它真的“理解”自己在做什么吗？——还是仅仅执行了一套更复杂的统计模式？

本文从表意 AI (Logographic AI, LAI) 理论[1]的视角出发，揭示物理 AI 在 Token 主义范式内的根本困境：它建立了符号与传感器信号的统计映射，却未能让符号真正锚定于意义。真正的符号接地，要求认知基元本身携带意义——这正是形根 (Morpho-Root) 理论和自然语言本体论 (NLO) 的核心主张[1]。

Abstract

In 2026, Jensen Huang declared at CES that "the ChatGPT moment for Physical AI has arrived" [5]. Shortly thereafter, the Figure 03 robot operated continuously for over 100 hours in a public live stream, processing over 130,000 packages [9][10]. However, this live stream, intended to demonstrate the robot's autonomy, instead sparked widespread doubts about "autonomous operation versus remote control," due to details such as a 0.5-second latency, abnormal "head-touching" behavior, and dropped packages. At the same time, AGI Robotics became the world's first embodied AI company to break through the milestone of 10,000 units in mass production [11]. The "feasibility" of Physical AI is no longer in question; the real competition has entered the critical phase of "usability" and "economic viability." Yet, a deeper problem is obscured by the celebration: when a robot "touches" a package, grasps an object, or avoids an obstacle, does it truly "understand" what it is doing? — Or is it merely executing a more complex statistical pattern?

This paper, from the perspective of Logographic AI (LAI) theory [1], reveals the fundamental predicament of Physical AI within the Tokenist paradigm: it establishes statistical mappings between symbols and sensor signals, but fails to truly anchor symbols in meaning. Genuine symbol grounding requires that the cognitive primitives themselves carry meaning — which is the core proposition of Morpho-Root theory and Natural Language Ontology (NLO) [1].

关键词：物理 AI；符号接地问题；表意 AI；形根；自然语言本体论；无根符号主义

Keywords:

Physical AI; Symbol Grounding Problem; Logographic AI; Morpho-Root; Natural Language Ontology; Rootless Semioticism

一、从虚拟 AI 到物理 AI：概念溯源与技术演进

物理 AI 的诞生源于一个朴素的需求：AI 不能只“会聊天”，还必须“会做事”。

1.1 物理 AI 概念的提出

物理 AI (Physical AI) 概念的公开提出，可以追溯到 2024 年 3 月的 GTC 大会。英伟达 CEO 黄仁勋在那次大会上首次系统性地阐述了物理 AI 的愿景，并发布了相关的技术平台和工具链。黄仁勋将 AI 的演进明确分为四个阶段：**感知 AI (Perceptual AI) → 生成式 AI (Generative AI) → 代理 AI (Agentic AI) → 物理 AI (Physical AI)** [5]。

2025 年 3 月的 GTC 大会上，黄仁勋再次强调这一判断，指出 AI 已经历了三代技术范式的转移，接下来将是物理 AI 的时代[6]。2025 年 7 月，黄仁勋与之江实验室主任、阿里云创始人王坚对话时明确提出：“**人工智能的下个浪潮是物理 AI**” [7]。

这一概念在 2026 年 1 月的 CES 展会上达到高潮。黄仁勋宣告：“**物理 AI 的 ChatGPT 时刻即将到来**” [5]。他在演讲中表示，物理 AI 旨在让摄像头、机器人、自动驾驶汽车等自主系统具备在现实世界中“看、想、动”的能力，能够完成感知、理解、推理，并执行或协调复杂动作。他特别强调：“当模型能够理解质量、摩擦、惯性、动量守恒，AI 才真正走出屏幕。[5]”

1.2 物理 AI 的技术内涵

中国科学技术大学人工智能与数据科学学院特任教授王翔系统阐释了这一概念：“物理 AI 意味着 AI 系统具备在真实世界中‘**感知—推理—行动—反馈**’的闭环能力。它不仅会思考，更能通过机器人等具身设备执行任务，并从真实反馈中持续纠错、自我进化” [8]。

黄仁勋进一步揭示了物理 AI 的本质：**当物理 AI 通过了技术进化的临界点，控制权从人类编写的确定性代码移交给了具有泛化能力的、理解物理规律的神经网络**[6]。

1.3 物理 AI 的三大技术支柱

物理 AI 的核心技术突破来自三个方向的融合：

第一、大语言模型 (LLM) 解决了“听懂”的问题：2023 年，谷歌先后推出 PaLM-E 与 RT-2 等具身多模态大模型。2025 年 9 月，谷歌 DeepMind 发布新一代多模态具身智能大模型 Gemini Robotics 1.5，宣称是“全球首款专为具身推理优化的思考型模型”，通过“先思考再行动”机制将长程任务失败率从 44.5% 降至 22% [14]。黄仁勋在 CES 2026 的演讲中指出，LLM 赋予物理 AI 理解人类语言和意图的能力，这是“理解任务”的第一步 [5]。

第二、世界模型 (World Model) 解决了“预判物理后果”的问题：英伟达发布的 Cosmos 平台——黄仁勋称之为“物理 AI 的 ChatGPT 时刻”的核心支撑 [5]——基于超 2000 万小时真实数据训练，能够从文字或图像生成符合物理规律的动作数据。Cosmos 3 进一步将物理推理、世界生成和动作生成整合为统一的 Mixture-of-Transformers 架构 [12]。但正如

行业分析所指出的，视频生成模型“只会复刻视觉样本，无法量化力、角度、摩擦力等交互参数”——生成逼真，不等于理解正确。

第三、视觉-语言-动作（VLA）模型打通了“看到即做到”的闭环：DeepMind 发布了 Gemini Robotics 1.5[14]；英伟达推出了 Isaac GR00T N1.6——全球首款开源的可定制通用机器人基础模型[13]；北京人形机器人创新中心开源了 XR-1，这是国内首个通过具身智能国标测试的 VLA 模型[8]。三者叠加，使得机器人第一次具备了在开放环境中自主执行任务的基础能力。黄仁勋在 GTC 2025 上表示：“通用机器人时代已经到来”[6]。

然而，世界模型赛道正面临严峻挑战。行业共识日益清晰：**视频生成不等于物理解**。更深层的理论困境，则指向一个困扰 AI 领域数十年的根本问题——符号接地。

1.4 物理 AI 的落地瓶颈

物理 AI 的全面普及仍面临显著挑战。王翔教授指出，物理 AI 面临真实交互数据昂贵、稀缺且反馈延迟的困境，全球高质量物理数据仅 50 万小时，而 GPT-5 级 LLM 训练数据折算超 100 亿小时，两者相差超 2 万倍；同时，仿真与现实之间存在**动力学、传感噪声等多重壁垒**，使得仿真中训练的模型迁移到真实场景极易失效[8]。此外，物理 AI 还需将 AI 算法、机械结构、传感器、控制系统深度集成，工程复杂性极高；其动作的不可逆性也意味着感知误差或算法漏洞都可能导致设备损坏甚至人身伤害，规模化部署必须具备**内生安全机制**，叠加可验证的安全约束、全链路审计与合规闭环[8]。

对于物理 AI 领域普遍存在的数据稀缺焦虑，刘深在《有根的智能——表意人工智能的范式革命》（The Rooted Intelligence: The Paradigm Revolution of Logographic AI）中从表意 AI 视角指出：真正的智能不应依赖海量数据的统计拟合，而应植根于概念本身的结构关系[1]。这一判断对物理 AI 同样适用——物理世界中的“概念”（如力、摩擦、惯性）同样可以被编码为结构化的认知基元，而非必须通过海量传感器数据来“学习”。

二、符号接地问题：AI 的“原罪”

2.1 Harnad 的经典问题

1990 年，认知科学家 Stevan Harnad 提出了**符号接地问题（The Symbol Grounding Problem）**：纯粹的符号系统如何获得意义？如果符号的意义只能由其他符号定义，整个系统就陷入无限循环，永远无法与真实世界建立联系[2]。

Harnad 指出，符号系统中符号的根本无意义性，正是 Searle 在“中文屋”论证中所揭示的——一个完全不懂中文的人可以通过规则手册完美地回答中文问题，但他并不“理解”中文[2][3]。Harnad 的解决方案是“符号接地”：通过确保符号在环境中实现物理层面的接地，使符号系统获得意义。他提出了一种混合模型：让神经网络学习将模拟感官输入分类为符号，从而建立符号系统的自下而上的接地[2]。

符号接地问题的研究并非孤立的学术话题。中国学者李建会等人在《计算主义及其理论难题研究》中指出：形式系统本身是有意义的，但“符号系统由于没有人心灵的参与，其本身无法获得意义，只能由外部输入”[15]。霍书全在《人工智能符号接地问题研究的

意义和挑战》中进一步指出，符号接地问题要解决“符号如何获得意义、符号和它们的意义如何联系起来的问题”，这一问题由塞尔中文屋论证所引发，并涉及“意识是如何产生的这个更困难的问题”[16]。学者坎杰洛西将符号接地问题分解为三个子问题[16]：（1）智能体如何自动把符号联系到世界中的指称；（2）智能体如何自动创造符号-符号关系；（3）智能体社会如何自动发展出符号的共享集合。其中后两个问题至今没有满意地解决。

2.2 Token 主义的“无根”困境

刘深在《有根的智能——表意人工智能的范式革命》中系统论述了这一困境的当代版本。该书指出，**Token 主义 AI 正是这样一个符号系统**：Token 本身是一个整数索引，其“语义”完全由训练数据中的统计共现关系临时赋予[1]。当 AI 处理“火”这个 Token 时，它学习的是“火”与“热”“燃烧”“危险”等 Token 在文本中的共现频率，而非“火”本身是什么。

Token “火”与 Token “热”在语义空间中接近——但 AI 从未感受过“热”，也从未见过“火”，更不知道“火”会烧伤皮肤。它所拥有的，只是一个统计学意义上的“近似关系”，而非对“火”的“理解”。

2.3 物理 AI 为何仍然是“伪接地”

物理 AI 似乎解决了这个问题——因为它有“身体”。它可以通过传感器“感知”火焰的温度，通过摄像头“看见”火焰的形状，通过执行器“体验”靠近火焰的后果。但“**物理接触**”不等于“**认知理解**”。

物理 AI 的“接地”路径本质是：给 Token “火”配上热传感器数据，给 Token “抓取”配上电机指令。它只是在 Token 和传感器信号之间建立了一套新的统计映射。正如 Boyle 所指出的，**单纯的“换能”（transduction）并不能保证符号的深度接地**——真正决定接地程度的，是符号如何具有因果效力[4]。

这与《有根的智能》第 5.3 节对符号接地问题的诊断一致：**Token 主义 AI 的“伪接地”**在于，Token 的“意义”完全由向量空间中的距离定义，但这些距离指向什么？指向其他 Token 的向量。这是一个彻底自我指涉的封闭循环[1]。

这暴露了物理 AI 接地的三重局限：

第一、“物理接触”≠“认知理解”：机器人触摸火焰触发“灼热”传感器→输出 Token “危险”，但系统本身并不“知道”灼热意味着什么——它只是学会了“靠近火→温度升高→执行远离”的统计模式，没有内在地理解“火”是什么。

第二、统计映射是可塑的：只要改变数据分布或奖励函数，所有“理解”都会发生漂移。一台在模拟环境学会“躲避火焰”的机器人，换一个传感器或换一种场景，它的“知识”就可能崩塌。仿真合成数据自带“仿真偏差”，模型迁移到真实场景极易失效。

第三、反事实推理的缺失：它无法“想象”如果火不存在会怎样——它没有因果模型，只有统计模型。《有根的智能》第 10.3 节系统论述了这一问题：**Token 主义无法实现珀尔因果阶梯的第二层（干预）和第三层（反事实），只能停留在第一层（关联）**[1]。

物理 AI 的“伪接地”本质上是用物理世界的传感器数据作为 Token 的新维度，而没有从根本上改变 Token “无根”的认知基元。

三、表意 AI 的回答：让认知基元本身携带意义

3.1 表意 AI 理论的核心诊断

表意 AI (LAI) 理论是中国学者刘涇于 2025 年提出的、对当前以 Token 为认知基元的主流 AI 范式的根本性替代。表意 AI 理论的核心诊断是：**Token 主义范式的哲学本质是“无根符号主义” (Rootless Semioticism)**——符号的意义完全由符号系统内部的统计共现定义，无法触及符号之外的“所指”，即真实世界、人类经验与文明价值。符号因此成为“漂浮的能指”，在封闭的统计循环中无限滑动[1]。

3.2 形根理论：意义内嵌的认知基元

与此相对，表意 AI 以“形根” (Morpho-Root) 替代 Token——每个形根都是一个结构化三元组[1]：

$$r = \langle S, A, R \rangle$$

S (Symbol)：符号标识，形根的外部可寻址名称；

A (Attributes)：属性集，内嵌该形根的固有语义特征与价值约束；

R (Relation Functions)：关系函数集，定义该形根与其他形根之间预设的逻辑连接方式。

形根“火”的示例：

```
r_火 = (
  "火",
  {
    [+高温], [+可燃烧],
    [+危险, 不可随意触碰],
    [+生命维持]
  },
  {
    leads_to(火, 灼伤),
    can_cause(火, 灰烬),
    inhibited_by(水, 火)
  }
)
```

}
}

形根的关键设计在于：**意义不是从统计中“涌现”的，而是作为认知基元的固有属性被预置的[1]**。当物理 AI 的“火”形根与真实传感器交互时，传感器信号不是“填充”一个空的 Token，而是“激活”一个已经携带意义的形根。

《有根的智能》第 5.3 节系统阐述了表意 AI 对符号接地问题的解决方案[1]。与 Harnad 的“混合模型”不同，形根理论的属性集合 A 被设计为指向非符号经验的开放性接口——感知接口、物理模拟接口、情感接口。整个系统不是“先封闭再接地”，而是“天生有根”[1]。

3.3 自然语言本体论（NLO）

表意 AI 的哲学根基是**自然语言本体论（Natural Language Ontology, NLO）**。《有根的智能》第 8 章系统阐述了这一理论：**语言结构本身就是智能的一种“本体”形态，是认知世界的原初方式**，语言不再仅仅是描述世界的符号系统，其自身就是世界意义得以呈现的场域[1]。章启群在《意义的本体论》中进一步论证，意义并非符号的附加属性或语言系统内部差异的产物，而是人与世界理解的构成性条件，是“存在的家园”[17]。

NLO 的四个核心命题[1]：

- (1) **语言即本体**：语言结构不是认知的工具，而是认知的本体形态；
- (2) **形式即意义**：在表意文字中，符号的形式结构本身就是意义的来源；
- (3) **居住而非使用**：智能系统应当“居住”在语言结构中，而非从外部“处理”语言数据；
- (4) **多元即合法**：不同文明的文字系统对应不同的合法认知范式。

《有根的智能》第 8.3 节还回应了一个核心质疑：形根理论不是传统符号主义的延续。传统符号主义的符号是封闭的、组合的、离身的；形根的符号是开放的（通过 A 指向非符号经验）、生成的（通过预设关系 R）、具身的（通过价值内嵌），实现了对符号主义的根本超越[1]。

3.4 L0-L4 五级接地置信度体系

《有根的智能》第 5.3.6 节将符号接地从二元（接地/不接地）转化为可工程化的程度问题，建立了 L0-L4 五级接地置信度体系[1]：

等级	名称	定义	当前物理 AI 水平
L0	无接地	符号意义完全来自统计关系	Token 主义 AI 的起点
L1	感知接地	符号与多模态感知特征关联	约在此水平
L2	具身模拟接地	符号与物理引擎或情感模型关联	约在此水平

等级	名称	定义	当前物理 AI 水平
L3	真实世界接地	符号与机器人、传感器实时交互绑定	物理 AI 的目标
L4	社会接地	符号通过与人类的社会交互获得意义	尚未达到

当前物理 AI 大致处于 L1-L2 之间——它建立了符号与感知特征的关联，但尚未通过真实世界的具身体验和社会交互来内化意义。降级机制确保了系统的稳健性：当 L3 接地因传感器故障失效时，系统可自动降级到 L2 或 L1；当接地完全丢失时，主动请求人类介入[1]。

四、两种范式下的“接地”对比

维度	Token 主义物理 AI 的“伪接地”	表意 AI 的“真接地”
符号的初始状态	空的索引，无意义	携带固有属性的形根
如何建立与世界的联系	通过统计共现关联传感器信号	通过形根属性 A 的开放性接口激活真实体验
意义来源	外来的、统计的、临时的	内生的、结构的、永久的
反事实推理	无法回答“如果”	可以通过属性编辑进行反事实推演
安全边界	外挂护栏，可被绕过	硬约束，逻辑上不可违背
对应接地等级	L0-L2	L3-L4
哲学根基	无根符号主义	自然语言本体论（NLO）

五、结论：身体不能替代根

黄仁勋的物理 AI 解决了一个重要问题——**让智能体拥有身体**，完成了从数字世界迈向物理世界的关键一步。世界模型试图让 AI “预判”物理后果，VLA 模型打通了“看到即做到”的闭环——三者叠加，使物理 AI 具备了在开放环境中自主执行任务的基础能力。

但**身体不能替代根**。一个能够“摸”到火的机器人，如果其认知基元仍然是“无根”的 Token，它就只能在统计拟合的意义上“完成任务”，而无法在结构理解的意义上“理解任务”。

正如 Harnad 所指出的，符号接地需要能量在系统与环境之间跨界面传递，使符号结构响应转导的环境状态而激活。物理 AI 拥有这种能量传递——但它没有将这种传递转化为意义锚定。它所建立的，仍然只是统计关联，而非结构理解。

《有根的智能》第 4.8 节和第 4.9 节进一步揭示了 Token 主义范式的安全危机：仅需 250 篇精心构造的文档即可在任何规模的大语言模型中植入隐身后门，而外挂的安全护栏

永远可以被绕过[1]。在物理 AI 场景中，这意味着一个看似“自主”的机器人完全可能在无人察觉的情况下执行危险操作——不是因为它“恶意”，而是因为它从未真正理解“安全”的含义。

物理 AI 让智能有了“身体”，表意 AI 让智能有了“根”——二者结合，才是通向真正通用智能的路径。**智能必须有根，文明必然有魂。**

参考文献

- [1] 刘深. (2026). 有根的智能——表意人工智能的范式革命 [The Rooted Intelligence: The Paradigm Revolution of Logographic AI]. 专著. <http://logographica.com/Monograph.html> (accessed July 2026).
- [2] Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3), 335-346.
- [3] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.
- [4] Boyle, C. F. (2001). Transduction and Degree of Grounding. **Psychology**, 12(36).
- [5] 孙小程. (2026-01-06). 英伟达黄仁勋：物理 AI 迎来“ChatGPT”时刻. 上海证券报·中国证券网. <https://www.cnstock.com/commonDetail/614165> (accessed July 2026).
- [6] 郑维汉. (2025-03-19). 不止芯片！黄仁勋最新演讲，信息量大. 上海证券报·中国证券网. <https://www.cnstock.com/commonDetail/377435> (accessed July 2026).
- [7] 郑亚丽, 刘晨茵. (2025-07-18). 链博会上王坚和黄仁勋开展“炉边对话”. 新华网浙江频道（转载自浙江日报）. <http://www.zj.xinhuanet.com/20250718/144633e1c1434db58bff08fa071ed653/c.html> (accessed July 2026).
- [8] 刘霞. (2026-01-19). 从虚拟智能跨入实体执行 物理 AI：人工智能发展又一高光时刻. 科技日报, 第 4 版. https://www.stdaily.com/web/gdxw/2026-01/16/content_462757.html (accessed July 2026).
- [9] 新智元. (2026-05-19). Figure 机器人直播 100 小时打擂台，3 秒一个不停分拣 13 万包裹！北京智源人工智能研究院（BAAI）转载. <https://hub.baai.ac.cn/view/54787> (accessed July 2026).
- [10] Adcock, B. [@adcock_brett]. (2026, June 30). Figure 03 humanoid robot has started performing a logistics workflow at BMW Group Plant Spartanburg [Post on X]. X. https://x.com/adcock_brett/status/2071973594268327993
- [11] 孙小程. (2026-06-24). 智元彭志辉：本体量产万台后 新的数字里程碑将于近期揭晓. 上海证券报·中国证券网. <https://www.cnstock.com/commonDetail/733753> (accessed July 2026).
- [12] NVIDIA. (2026-06-01). Develop Physical AI Reasoning, World, and Action Models with NVIDIA Cosmos 3 [NVIDIA Technical Blog]. <https://developer.nvidia.com/blog/develop-physical-ai-reasoning-world-and-action-models-with-nvidia-cosmos-3> (accessed July 2026).
- [13] NVIDIA. (2026-01-07). Building Generalist Humanoid Capabilities with NVIDIA Isaac GR00T N1.6 Using a Sim-to-Real Workflow [NVIDIA Technical Blog]. <https://developer.nvidia.com/blog/building-generalist-humanoid-capabilities-with-nvidia-isaac-gr00t-n1-6-using-a-sim-to-real-workflow/> (accessed July 2026).

- [14] Gemini Robotics Team. (2025). Gemini Robotics 1.5: Pushing the Frontier of Generalist Robots with Advanced Embodied Reasoning, Thinking, and Motion Transfer. arXiv, arXiv:2510.03342. <https://arxiv.org/abs/2510.03342> (accessed July 2026).
- [15] 李建会, 赵小军, 符征. (2016). 计算主义及其理论难题研究. 北京: 中国社会科学出版社.
- [16] 霍书全. (2019). 人工智能符号接地问题研究的意义和挑战. 上海师范大学学报(哲学社会科学版), (3), 98-107. DOI: 10.13852/J.CNKI.JSHNU.2019.03.010.
- [17] 章启群. (2018). 意义的本体论——哲学解释学的缘起与要义. 北京: 商务印书馆.