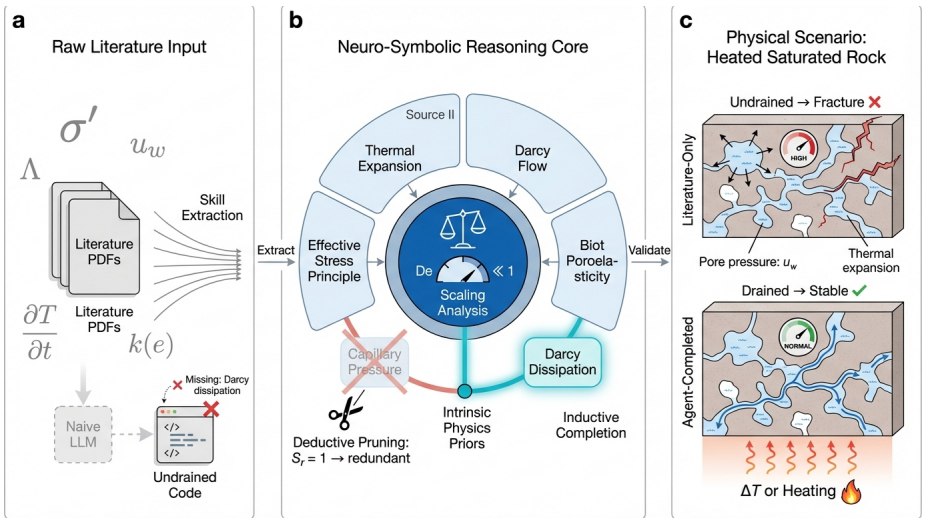
arXiv:2603.09756v1 [cs.DB] 10 Mar 2026

图形摘要

认知闭合：面向物理一致性模拟的自主机制补全

Yue Wu、Tianhao Su、Rui Hu、Mingchuan Zhao、Shunbo Hu、Deng Pan、Jizhong Huang



认知闭合：面向物理一致仿真的自主机制补全

Yue Wu[a,1], Tianhao Su[a,1], RuiHu[a,c], Mingchuan Zhao[d], ShunboHu[a,b,*], DengPan[a,*],Jizhong Huang[a,b,*]

[a]上海大学材料基因组研究院，上海，*200444*，中国[b]上海大学文化遗产与信息管理学院文化遗产保护研究院，上海，*200444*，中国[c]上海斯韦弗科技有限公司，上海，中国[d]科克大学学院数学科学学院，科克，爱尔兰共和国

摘要

大语言模型（LLM）融入科学发现目前受到**“隐含语境”**问题的阻碍：从文献中提取的控制方程包含不可见的热力学假设（例如不排水条件），而标准生成模型无法识别这些假设。这会导致**“物理幻觉”**：生成语法正确、能够忠实执行物理上无效定律的求解器。在此，我们提出一种神经符号生成智能体，充当传统数值引擎之上的认知监督者。通过将物理定律封装为模块化的**“本构技能”**并利用潜在内禀先验，该智能体采用思维链推理工作流，自主验证、剪枝并补全物理机制。我们在低渗透砂岩热增压问题上展示了这一能力。标准文献检索基线盲目采用刚性“不排水”简化，错误地预测灾难性材料失效；而我们的智能体通过无量纲标度分析，自主识别出系统处于排水状态（德博拉数 $De \ll 1$）。因此，它归纳性地补全了满足边界约束所需的

*通讯作者。电子邮箱： shunbohu@shu.edu.cn（Shunbo Hu），DPan_MGI@shu.edu.cn（Deng Pan），hjizhong@shu.edu.cn（Jizhong Huang）[1]这些作者对本研究贡献相同。

缺失耗散机制（达西流），预测出与实验实际情况一致的稳定应力路径。这项工作确立了一种范式：AI智能体超越编码助手的角色，成为能够对科学数据中嵌入的理论假设进行推理并加以修正的认识论伙伴。

关键词： 神经符号人工智能、物理幻觉、自主机制补全

---

1. 引言

人工智能融入科学发现，已超越静态性质预测阶段（例如，用于蛋白质结构预测的 AlphaFold [1], 和用于材料稳定性预测的 GNoME [2]），如今正迈向多物理场模拟这一动态前沿[3, 4, 5, 6]。传统数值求解器（有限元法/有限体积法）要求严格、明确地定义控制定律[7, 8], ，而制定这些定律所需的科学知识仍散见于非结构化文献之中。“AI科学家”智能体的终极目标是弥合这一认知鸿沟：自主阅读文献、综合数学模型，并在无需人工干预的情况下执行高保真模拟[9, 10, 11, 12]。

通用大语言模型（LLM）近期的性能飞跃推动了这一愿景的实现，例如 Claude 4.5 Sonnet[13] 和 Gemini 3 Pro[14]。与前代模型不同，这些前沿模型展现出涌现推理能力，其范围已超越句法转换，延伸至语义理解 [15, 16, 17, 18]。随着智能体框架兴起——它们将大语言模型的输出组织为可执行的“技能”或工具[19, 20, 21, 22]——构建自主系统在计算上已切实可行；这类系统能够将抽象物理概念作为模块化编程单元加以操作[23, 24]。

然而，一个关键瓶颈依然存在：“隐含语境”问题。文献中的科学方程很少是自成一体的；它们受到

不可见假设的约束（例如“不排水条件”“等温”或“小应变”），这些假设对人类专家来说显而易见，但标准的“Copilot”[25]却难以理解。天真的大语言模型将物理学视为文本补全任务，往往会陷入“物理幻觉”的陷阱：生成的代码语法正确、数学上可解，却会造成灾难性的物理后果，因为它将有效定律（例如不排水增压）应用于其有效域之外（例如排水边界）[26, 27, 28]。

这一局限凸显了当前计算领域中的结构性缺口。MOOSE [29], 、FEniCSx [30], 或 JAX-FEM [31] 等成熟的多物理场平台，在自动执行下游数值计算（例如弱形式组装、线性化）方面已相当成熟。然而，它们本质上只是“不可知计算器”：缺乏质疑输入有效性的认知能力。它们完全依赖外部主体——传统上是人类专家——来完成上游决策：选择正确的本构定律、验证热力学一致性，并确定适当的耦合拓扑结构[32]。目前，“认知监督者”这一角色仍由人工承担；如果求解器接收到一个数学上有效、但物理上不完整的模型（例如由朴素的大语言模型生成的模型），它仍会忠实地收敛到物理上错误的解。

为弥合这一缺口，我们提出一种神经符号生成智能体[33, 34,35] ，将大语言模型从代码转换器提升为推理内核，从而有效地自动履行认知监督者的职责。我们的框架不将提取出的方程视为静态文本，而是视为“本构技能”——带有语义元数据的封装逻辑单元。借助嵌入于大语言模型潜在空间中的内在物理先验（通过知识图谱提取进行验证），该智能体执行思维链（CoT）工作流，以：

1. **演绎式剪除**冗余机制（例如，认定饱和状态下毛细压力可忽略）。2. **归纳式补全**缺失物理（例如，无量纲缩放分析揭示排水状态时，综合引入达西流耗散）。

我们在一个具有代表性的地质力学挑战上验证了这一方法：低渗透率岩石（Rothbach砂岩）中的热增压。这个问题是理想的“认知压力测试”，因为它涉及热生成与流体耗散之间相互竞争的时间尺度。我们证明，标准的文献检索基线会盲目采用“不排水”简化，因此无法奏效；而我们的智能体能够自主识别正确的物理状态，引入必要的耗散项，并预测出与实验实际情况一致的稳定应力路径。

2. 结果

为验证所提出的框架，我们部署智能体来解决一个典型的地质力学挑战：低渗透率岩石（Rothbach砂岩）中的热增压。该问题需要平衡热力学驱动力与水力耗散，因此是检验智能体推理能力的理想测试平台。实现过程遵循三阶段自主工作流程：

1. **技能摄取（图 1a）：** 智能体处理本构文献原始数据集，将控制方程（例如热膨胀、毛细管定律）封装为离散的“本构技能”，并添加语义元数据，标注相互作用的物理场及适用域（例如饱和状态）。

2. **拓扑构建（图 1b）：** 智能体将用户指定的模拟场景（包括几何和边界条件，例如 T_{bnd} =200℃）与技能库相结合，自主构建因果图。关键的是，它通过**逻辑剪枝**移除冗余机制，并通过**内在补全**引入边界约束所需的缺失耗散路径。

3. **耦合模拟（图 1c）：** 构建的图被编译为一个求解器实例，以执行全耦合模拟，并根据实验稳定性判据验证结果。

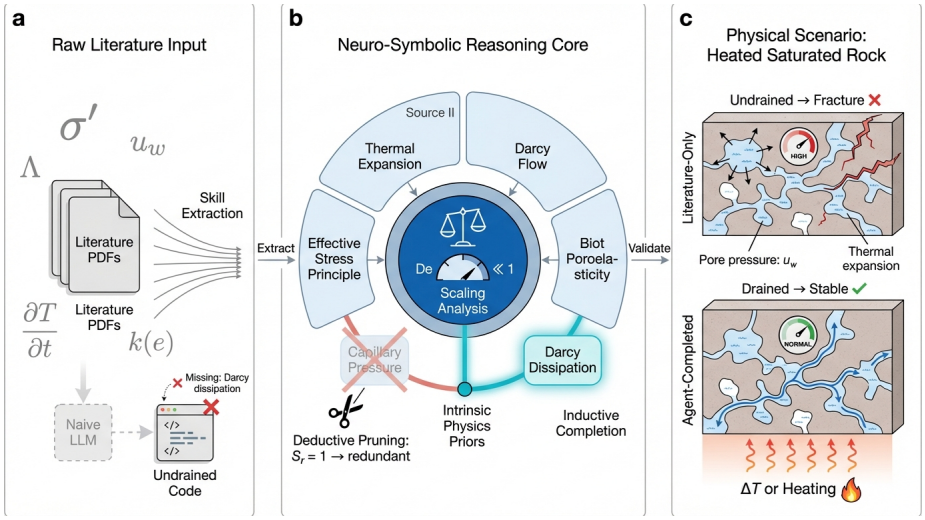


图 1：**神经符号生成智能体的架构概览**。**a,** 从文本到技能：智能体摄取非结构化科学文献（PDF），并将控制方程提取为模块化的“本构技能”。与盲目执行检索到的代码的简单大语言模型（灰色路径）不同，我们的智能体将这些技能视为候选假设。**b,** 推理核心：智能体充当认知守门人。它运用**演绎剪枝**（剪刀图标）移除数学上冗余的机制（例如饱和状态下的毛细压力），并通过**归纳补全**（蓝色路径）补入缺失的物理机制。决策逻辑由**无量纲标度分析**（中央旋钮）支配，具体而言，使用德博拉数（De）来判定竞争机制。**c,** 物理后果：对受热饱和岩石模拟结果的比较。受提取出的“不排水”假设约束，“仅文献”模型（上）预测会发生材料断裂（物理幻觉）。相比之下，“智能体完成”模型（下）正确识别出排水状态，启用达西流以释放压力，并预测出稳定且符合物理规律的状态。

2.1. 通过涌现因果拓扑实现自主构建

为了从僵化的预定义求解器转向自适应仿真范式 ，
我们引入了<style id='1'>“本构技能”这一概念——它是对物理知识的封装单元，将数学控制方程与关于其适用性及相互作用场的语义元数据绑定在一起。图2直观展示了由智能体构建的涌现因果拓扑，体现了这些技能在内在物理逻辑基底上的动态组装。

该框架通过具备语境感知能力的构建过程运行，而非进行静态执行。在特定仿真语境下初始化后，智能体通过协调两个不同的知识来源，综合构建一个耦合系统。首先，它将关键驱动力锚定到由领域文献汇集而成的技能库中。例如，深蓝色路径（图2）并非仅仅是一条连接，而是一个实例化算子：它将热增压系数 Λ （从来源 II 中提取）应用于驱动孔隙压力演化。

至关重要的是，这一构建过程受物理保真度与计算经济性的自主优化所支配。智能体会根据边界条件评估每项候选技能的必要性。

- **逻辑剪枝：** 如**绿色路径** （图2）所示，智能体从来源 I[36]中获取饱和度-毛细约束。然而，通过推理得出初始状态已完全饱和（$S_r = 1$），它识别出该机制在数学上是冗余的，并自主**剪枝** 该边，从而在不牺牲准确性的情况下简化系统。

- **内在补全：** 相反，智能体认识到，从文献中提取的技能可能并不充分。橙色路径（将黏度与流体通量相连）并非源自外部文本，而是来自大语言模型的**内在物理先验**（流体力学基础知识）。智能体将这条路径识别为潜在耗散通道（达西流），并将其保留备用，以解决僵化的文献假设（例如“不排水条件”）可能忽略的物理不一致性。

这种动态拓扑结构为后续的自动耦合提供了蓝图，确保最终的模拟并非只是方程的堆叠，而是针对特定物理现实合理构建的系统。

2.2. 本构核的单元验证

自主仿真的预测保真度，从根本上受限于其组成核的准确性。在统筹完整的多物理场耦合之前，智能体会对提取的基于文献的技能进行严格的逐元素验证；这些技能在图2中以深蓝色路径和绿色路径表示。此阶段旨在验证隐含的因果联系是否已正确实例化为确定性的控制方程。

高阶动力学验证（绿色路径，来源 I）

尽管饱和-毛细机制最终因特定初始条件（$S_r = 1$）而被剪枝，但对其进行验证仍可作为智能体处理本构关系的基准。如图3a所示，智能体成功重建了毛细上升动力学。模拟捕捉到了瞬态演化，并准确收敛至 Jurin定律预测的平衡高度（$H = 29.68$ cm）。这证实智能体正确实例化了支配流体势的物理定律，也验证了绿色路径在进行上下文感知剪枝前的数学完整性。

热力学驱动的验证（深蓝色路径，来源 II）

通过与 Ghabezloo 与 Sulem（2009）的实验数据对比，验证主要源项——热增压——的准确性。图 3b 展示了不排水加热条件下孔隙压力（u_w）的变化。智能体的模拟结果（实线）与实验数据点高度吻合，证实热增压系数 Λ 的提取和实现准确无误。这一结果验证了因果拓扑中深蓝色路径的可靠性，确保外部驱动力在物理上合理，然后智能体再引入其内在完备机制来解决系统级约束。

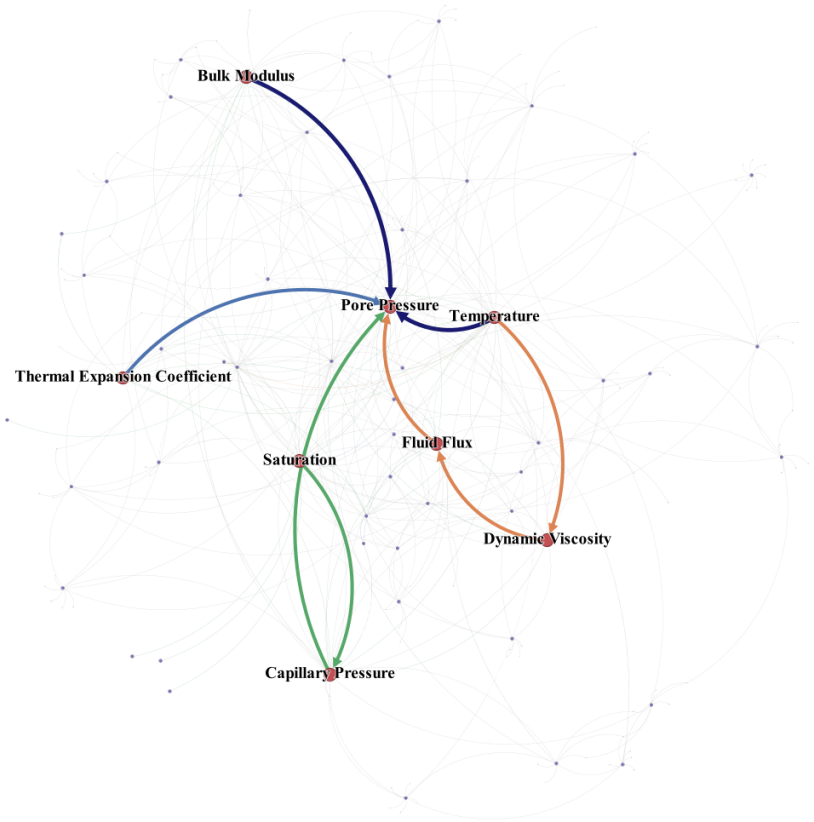


图2：**智能体用于自主机制构建的隐式因果拓扑可视化**。该图并非显式输入，而是对智能体底层认知过程的提取表示，将**内在先验**（基于大语言模型）与**检索技能**（基于文献）融合在一起。**深蓝色路径（检索）：** 表示隐式锚定于技能库（来源 II）的主要热力学驱动，将热膨胀与孔隙压力连接起来。**绿色路径（剪枝）：** 表示饱和度-毛细约束（来源 I）。智能体在内部将该机制识别为有效，但由于特定的仿真语境（$S_r = 1$），在动态执行过程中对其进行**剪枝**，从而滤除数学冗余。**橙色路径（内在）：** 表示一条源自大语言模型内部物理知识而非外部文本的**机制补全**路径。它反映了智能体对流体耗散（达西流）的潜在能力，并予以保留，以确保在时间尺度违背情况下的物理稳定性（详见第2.3节）。

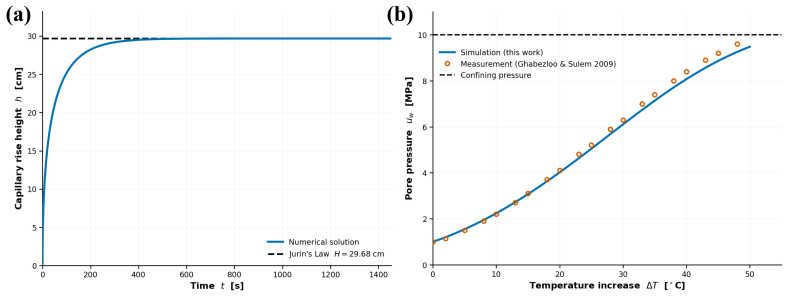


图 3：**从文献中提取的本构核的逐单元验证。a,** 对从来源 I（绿色路径）提取的毛细动力学进行验证。智能体模拟的毛细上升高度 $h(t)$（蓝色实线）准确捕捉了瞬态上升过程，并收敛至由Jurin定律定义的理论平衡极限（黑色虚线，$H = 29.68$ cm）。这证实了智能体能够在自主剪枝之前正确实例化饱和压力关系。**b,** 对从来源 II（深蓝色路径）提取的热增压驱动力进行验证。模拟准确再现了*Ghabezloo & Sulem(2009)* 在不排水加热条件下测得的实验孔隙压力演化（橙色标记）。这证实了在智能体的**内在完备机制**（橙色路径）激活之前，主要热力学驱动力的准确性。

2.3. 通过无量纲缩放进行上下文感知推理

仅验证本构核是不够的，因为文献推导的方程的有效性往往受隐含假设的限制，例如“ 不排水条件”，而这些假设在目标模拟环境中可能并不成立。为解决这一问题，智能体采用无量纲标度分析（具体而言，是德博拉数， De）作为认知筛选器，以确定占主导地位的物理状态。

智能体持续评估特征流体扩散时间（τ_{diff}）与热加载时间尺度（τ_{load}）之间的竞争关系。该比值量化为：

$$De = \frac{\tau_{diff}}{\tau_{load}} \approx \frac{L^2 \mu_f(T) \beta}{k \cdot t} \tag{1}$$

关键在于，智能体会自主识别出德博拉数（De）是唯一的控制参数，并将其与其他可能的无量纲候选参数（例如用于热传导的傅里叶数、用于平流的佩克莱数）区分开来。为解释这一自主决策，我们对提取出的知识图谱（图2）进行事后分析，从而提供了机理层面的证据。图谱显示，智能体在内部将孔隙压力演化构建为源项（热增压速率，$\sim 1/\tau_{load}$）与扩散汇项（达西通量，$\sim 1/\tau_{diff}$）之间的竞争。通过隐式排除平流主导状态（$Pe \to 0$）和不适用的内禀反应动力学（Da ），智能体正确推断出，系统行为由内禀弛豫时间（τ_{diff}）与外加载荷时间（τ_{load}）之比唯一决定。该比值严格对应德博拉数的定义（$De = \tau_{diff}/\tau_{load}$），证实智能体成功地从其潜在物理先验中推导出了数学上精确的标度律。

图4a展示了智能体对目标材料（Rothbach砂岩， $k = 10^{-16}$ m^2, 由**粗体黄色线条**表示）的状态区检测。与理论处理（SourceII）中常采用的热增压简化理论“不排水”理想化假设（$De \gg 1$）相反，智能体识别出

系统实际上处于**“排水状态”**（粉红区）的深处，其中 De 的范围为 10^{-2} 至 10^{-3}。这项尺度分析揭示了一个关键的物理现实：流体逸出的时间尺度比升温速率快两个数量级。

机制补全触发：

对 $De \ll 1$ 状态的明确识别，成为触发智能体**内在机制补全**的条件。智能体意识到，若严格遵循“不排水”公式，将会严重高估孔隙压力。因此，它自主启用橙色路径（图2）——内在达西流动机制——从而向控制方程中注入扩散耗散项（$\nabla \cdot \mathbf{v}$）。此外，如图4b所示，智能体还考虑了水力扩散率分量（黏度 μ_f 和储水系数 S_s）随温度变化的规律，确保耗散机制能够随温度场动态调整。这一过程展现了智能体以稳健且适用于相应尺度的物理先验，取代基于文献的简化假设的能力。

2.4. 全耦合模拟中的涌现有效性

智能体本构推理的最终成果是**全耦合多物理场模拟**，在这一模拟中，其自主机制补全的影响显而易见。智能体将经过验证的热致增压驱动因素（检索技能）与适用于相应尺度的流体耗散项（内在先验）相结合，构建出物理上一致的模型，从而避免灾难性的预测误差。

图 5 展示了这一影响在 $p' - q$ 应力空间（平均有效应力与偏应力）中的表现，该空间是判定材料稳定性的权威图谱。灰色虚线表示朴素的“仅文献”模拟。由于受来源 II 提取出的“不排水”假设约束，该模型预测孔隙压力会持续累积且不受控制。这迫使有效应力路径迅速向左移动，最终越过**拉伸截断** 和**霍斯列夫面** （红色区）。关键的是，这种朴素模型错误地

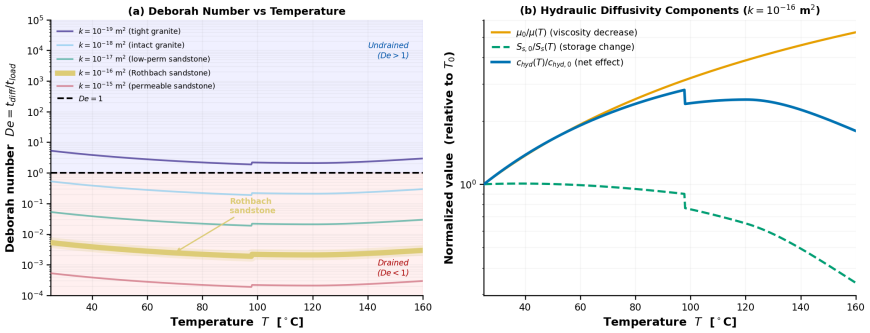


图 4：**通过德博拉数标度分析进行自主流态检测。a,** 不同岩石渗透率下，德博拉数（De）随温度的变化。**黄色粗线**表示本研究使用的目标本构模型（Rothbach砂岩， $k = 10^{-16}$ m^2）。智能体识别出，系统始终处于**粉红区**（$De \ll 1$）内，这表明系统处于“排水”状态，流体扩散占主导地位。这一尺度分析结果明确否定了简化热增压模型（来源 II）中常隐含的“不排水”假设，促使智能体启动内在耗散机制。**b,** 水力扩散率分量相对于其在 T_0 时初始值的分解。智能体追踪黏度降低（橙线，促进流动）与储存能力变化（绿线）这两种相互竞争的效应，并将其综合为净扩散率趋势（蓝线），以动态调节达西流强度。

预测材料发生破坏，这表明岩石会在实验已知其仍保持完整的条件下发生断裂。

与之形成鲜明对比的是，实线蓝线表示“智能体完成”的模拟。通过尺度分析诊断出排水状态（$De \ll 1$）后，智能体自主启用了达西流动机制（橙色路径）。由此得到稳定的应力路径：内在耗散释放了过量孔隙压力，阻止有效应力继续向左迁移。系统最终稳定在安全的平衡端点（$p' = 8.9$ MPa），与破坏包络线保持足够的安全裕度。

这一结果为“涌现有效性”提供了确凿验证。它证实，若没有智能体自主推理来弥合文献假设与物理现实之间的差距，模拟结果将不只是有误，而是**定性错误**，会预测并不存在的结构破坏。

最后，智能体补全机制确保了对空间非均质性的正确表征，这种能力超出了零维本构求解器的范围。图6展示了Rothbach砂岩案例（$k = 10^{-16}\,\mathrm{m}^2$）在$t = 175$ s时的温度（T）和孔隙压力（u_w）空间场。温度场（图6a）由从边界向内传导的热传递所控制，而孔隙压力场（图6b）则揭示了智能体启用的达西流所发挥的作用。关键的是，孔隙压力稳定在约41.1 MPa（与图5所示稳定有效应力 $p' \approx 8.9$MPa一致）。可见的梯度——中心较高、边缘较低——证实智能体成功施加了排水边界条件，使流体得以逸出，而非不断积聚至破坏水平。

3. 讨论

认识论纠正 vs. 盲目转换

这项工作的核心贡献不仅在于自动化代码生成，更在于展示了**认识论纠正**——即一种能力：

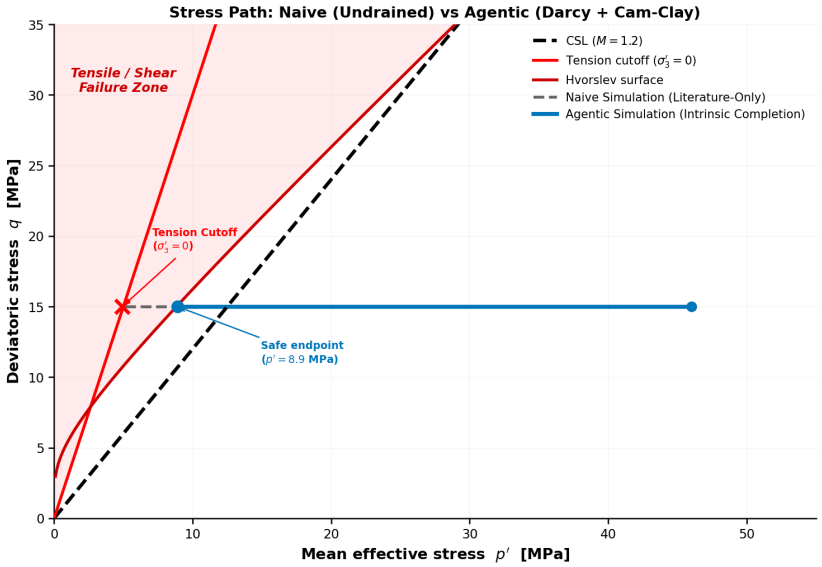


图 5：**通过有效应力路径稳定性展示涌现有效性**。该图追踪了系统在 $p' - q$ 空间中的演化，并与岩石的破坏包络线（红色区：拉伸/剪切破坏）进行对照。**灰色虚线（仅文献）：** 表示严格受限于“不排水”文献提取结果的模型。未受抑制的热加压使有效应力路径进入破坏区（以**红叉**标示），从而错误地预测岩石断裂。**蓝色实线（智能体完成）：** 表示完全耦合的模型；检测到 $De \ll 1$ 状态后，智能体自主启用了**内在机制**（达西流）。扩散耗散阻止了应力进一步降低，使路径稳定在**安全终点**（$p' = 8.9$ MPa），远在弹性包络线以内。这一对比凸显出，智能体的机制补全对于保持模拟的**定性物理正确性**至关重要。

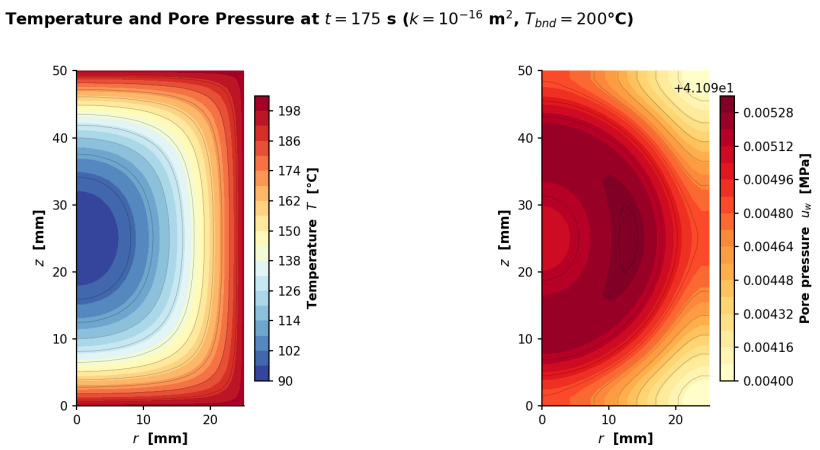


图6：**通过场等值线对完全耦合解进行空间验证**。目标Rothbach砂岩在 $t = 175$ s时的系统状态快照（$k = 10^{-16}$ m^2,$T_{bnd} = 200$℃）。**a,** 温度分布（T），显示传导加热前沿从边界（$r = 25$ mm）向核心传播。**b,** 孔隙压力分布（u_w）。压力场在 ≈ 41.1 MPa下几乎均匀，仅存在细微梯度，推动流体从中心（深红色）流向排水边界（浅黄色）。这证实，Agent的内在机制（$\nabla \cdot \mathbf{v}$）能有效消散热诱导压力，避免朴素模型预测的非物理累积。

挑战并纠正嵌入训练数据中的假设。标准大语言模型（LLM）往往像“随机鹦鹉”一样运作，只会检索文本中最可能的后续内容。如果我们的智能体像标准 Copilot 那样行事，它就会严格遵循检索到的文献（来源 II），采用“不排水”假设，生成能够无误运行、但在物理上不成立的代码（如图5中的“朴素”路径所示）。通过明确运用无量纲标度分析（德博拉数）作为认知守门人，该智能体超越了语言翻译的范畴。它展现出某种“物理直觉”，认识到在开放系统边界条件（T_{bnd} active）下，封闭系统方程的数学有效性不再成立。这种能力标志着从翻译者型人工智能向物理学家型人工智能的范式转变。

内在先验作为安全阀的作用

潜在知识图谱的可视化（图2）揭示了智能体为何能在僵化的求解器失效之处取得成功。“内在先验”——尤其是黏度降低与流体通量之间的联系（达西定律）——构成了一道潜在的安全网。检索到的技能（RAG）提供了材料（Rothbach砂岩）的特定本构参数，而内在先验则提供了维持热力学一致性所需的普适守恒定律。这种神经符号交互使智能体能够以可控且富有建设性的方式“产生幻觉”：它并非因为收到明确指示，而是因为“排水”状态（$De \ll 1$）下的质量守恒要求如此，才创造出一个缺失的机制（橙色路径）。这表明，未来的科学人工智能框架应着重整理这些内在守恒先验，以防止生成式模拟出现违反物理规律的情况。

可推广性与局限性

尽管该框架是在一个特定的热-水-力（THM）问题上展示的，但其逻辑——**技能封装 → 拓扑剪枝→ 内在补全**——并不依赖于特定的物理领域。理论上，只要大语言模型的潜在空间中存在基本守恒定律，该框架就可以推广到其他耦合领域，例如化学力学（电池退化）或热磁学

。然而，该系统依赖于预先定义的“技能库”（即使是自主导入的），因此其效果取决于种子文献的质量。如果源文本中包含关于方程函数形式的根本性错误（而不仅仅是适用性假设方面的问题），那么当前的验证层（单元测试）可能不足以应对。未来的工作将致力于整合数据驱动发现（例如，对原始实验数据进行符号回归），使智能体不仅能够修正图的拓扑结构，还能修正核函数本身。

4. 方法论

我们提出了一种神经符号生成框架，能够自主构建并求解多物理场问题，其中控制定律并非预先定义，而是根据文献动态综合生成。与将自然语言指令转换为僵化脚本的标准“副驾驶”不同，我们的系统充当推理内核，弥合非结构化物理知识与严谨数值执行之间的语义鸿沟。工作流程分为三个不同阶段：(1) 语义封装，将原始方程组织为功能单元；(2) 物理信息认知推理，大语言模型 (LLM) 基于无量纲标度分析隐式筛选并补全机制；以及 (3) 自动变分综合，将经过验证的物理规律编译为可执行的有限元代码。

4.1. 本构技能的语义封装

该框架的基础单元是“本构技能”， $\mathcal{S}$。通过检索增强生成（RAG）从文献中提取的原始数学表达式，因缺少有关其适用性和耦合要求的上下文信息，无法满足自主推理的需要。为解决这一问题，我们将提取出的每条定律映射到一个结构化模式中，将静态方程转化为可连接的语义算子：

$$\mathcal{S} = \{\mathcal{F}, \Omega_{\text{int}}, \mathcal{C}_{\text{meta}}\} \tag{2}$$

其中:

- $\mathcal{F}$ （函数形式）：控制方程的符号表示（例如， $\epsilon_{th} = \alpha_T \Delta T$）。

- Ω_{int} （交互流形）：作为输入（前置条件）所需、并作为输出（目标）生成的物理场集合，定义了节点的连接潜力。

- $\mathcal{C}_{\text{meta}}$ （适用性约束）：定义有效域的语义元数据。这包括相态假设（例如“饱和液体”）、材料对称性和热力学限制。

这种封装使大语言模型能够评估技能的效用，而不只是进行关键词匹配，还能判断其与仿真拓扑的逻辑兼容性。

4.2. 物理信息认知推理

其核心创新在于智能体能够进行隐式拓扑推理。收到用户指定的仿真场景（几何、边界条件 Γ 和初始状态 $\mathbf{u}_0$）后，智能体不会简单地堆叠所有检索到的技能，而是执行**思维链（CoT）**推理过程，以合成有效的本构图，并采用两种特定的认知机制:

演绎逻辑：基于约束的剪枝

智能体首先评估检索到的技能的约束 $\mathcal{C}_{\text{meta}}$ 与场景初始条件$\mathbf{u}_0$之间的兼容性。对于在数学上冗余或在物理上不适用的机制，智能体会自动将其剪枝。例如：如果检索到的技能用于控制毛细管压力演化（P_c），而场景指定的是完全饱和域（$S_r = 1$），智能体便可推导出条件 $\nabla P_c = 0$，从而从控制方程中移除相应项，以避免数值奇异性。

归纳逻辑：通过尺度分析进行状态识别

除了满足约束外，智能体还使用无量纲标度分析来验证系统的物理保真度。在进行公式推导之前，智能体会估算特征时间尺度（例如，流体扩散 τ_{diff} 与热载荷τ_{load}），以计算关键无量纲数（例如德博拉数，De）。

- **状态区检测：** 智能体对主导物理状态进行分类（例如， $De \ll 1 \implies$ 排水； $De \gg 1 \implies$ 不排水）。

- **内在补全：** 如果文献推导方程隐含的假设（例如“ 不排水”）与识别出的状态区相矛盾，智能体就会检测到机制缺失。为恢复物理一致性，它会启用内在先验——存储在大语言模型潜在权重中的基本定律（例如用于描述耗散的达西定律）——并将缺失的散度项（$\nabla \cdot \mathbf{v}$）注入系统。

4.3. 自动变分综合

本构逻辑最终确定后——剔除冗余并补全内在物理机制——智能体将进入求解器综合阶段。此阶段会将抽象的物理图转换为弱形式变分公式，由 FEniCSx 有限元求解器执行。

1. 场初始化：智能体为 Ω_{int} 中识别出的活跃场定义适当的函数空间（例如，拉格朗日单元 P_2/P_1）。2. 弱形式推导：智能体将耦合偏微分方程的强形式以符号方式转换为残差弱形式： $R\mathbf{u},\ \mathbf{v}) = \int_\Omega (\dots)\ d\Omega + \int_\Gamma (\dots)\ d\Gamma\ = 0$ 其中$\mathbf{v}$ 表示检验函数。此推导明确处理了推理阶段引入的内在耗散项的积分。3. 代码编译：最终的变分形式被编译为 Python 仿真脚本，其中包含网格生成、时间步进循环，以及针对耦合系统刚度定制的自适应求解器参数。

4.4. 潜在物理知识可视化

为验证智能体是否具备第2.2节所述推理所需的物理先验，我们开发了一个知识提取探针，用于可视化大语言模型中嵌入的隐式本构逻辑。需要注意的是，智能体的运行时工作流（第2.1–2.3节）直接处理潜在表示；图2中展示的显式知识图谱 $G(V,\ E)$ 是一种事后重构，旨在验证模型的内部一致性以及对多物理场耦合的覆盖情况。

我们实施了一个三阶段探询协议，以系统地揭示这些潜在依赖关系：

1. 基于种子的广度优先扩展（局部拓扑）

从包含 11 个基本状态变量（例如孔隙压力、温度）的种子集出发，我们询问大语言模型，识别与每个变量直接相邻且最为关键的物理邻接变量。此阶段类似于深度为 $d = 2$ 的广度优先搜索（BFS），能够捕获主要的本构定律（例如将 P 与 σ' 联系起来的太沙基原理）。

2. 组合式成对探询（全局连通性）

为揭示局部遍历可能遗漏的间接或远距离耦合，我们从种子集中生成所有不重复的变量对（$C_{11}^2 = 55$ 对）。大语言模型对每一对变量进行判别，确定它们之间是否存在物理路径（直接路径或由中间变量介导的路径）。这确保跨学科相互作用（例如温度影响流体黏度）得到明确记录。

3. 域级耦合综合（系统一致性）

最后，为验证多物理场相互作用的完整性，我们查询整个物理子域（T-H、H-M、T-M）之间的连通性。这种自上而下的检查确保热增压等宏观尺度机制被正确分解为有效的边属性。

探测过程产生的原始输出经由规范化层进行标准化，将同义术语（例如“PWP”和“孔隙压力”）映射到标准化本体。生成的图谱（图2）证实，大语言模型的

内部知识结构本身包含自主推理所需的驱动力（例如热膨胀）和耗散机制（例如达西定律），这在结果中得到了展示。

4.5. 数值模拟方案

为验证智能体推理的涌现有效性，我们对一个承受快速热载荷的圆柱形罗特巴赫砂岩试样（$R = 25\text{mm},\ \ H = 50\ \text{mm}$）进行了对比模拟。该研究对比了两种不同的建模策略：

1. 基线“朴素”模型：假设严格处于不排水条件下的 0D 代表性体积单元（REV）（源自文献来源 II[37]）。2. 智能体合成模型：采用 2D 轴对称有限差分模型，纳入固有的达西流动机制（源自智能体的推理）。

控制性本构定律（智能体合成）

力学行为（修正剑桥黏土模型）

岩石骨架被建模为弹塑性连续体，其行为由 $p' - q$ 空间中的修正剑桥黏土模型（MCC）屈服面控制：

$$f(p', q) = q^2 + M^2 p'(p' - p_c) = 0 \tag{3}$$

其中 $M = 1.2$ 为临界状态斜率，p_c 为随塑性体积应变 ε_v^p演化的先期固结压力：

$$\frac{dp_c}{p_c} = \frac{1 + e_0}{\lambda - \kappa} d\varepsilon_v^p \tag{4}$$

我们引入拉伸截断（$p' \geq 0$）以及干侧的霍斯列夫面（$p' < p_c/2$），以刻画超固结砂岩典型的脆性破坏模式。

热–水耦合

孔隙压力演化由热增压（源）与水力扩散（汇）之间的竞争驱动：

$$\frac{\partial u_w}{\partial t} = \underbrace{\Lambda \frac{\partial T}{\partial t}}_{\text{Source}} + \underbrace{\nabla \cdot \left(\frac{k}{\mu_f(T) S_s} \nabla u_w \right)}_{\text{Sink}} \tag{5}$$

• **源项（Λ）：** 热增压系数 $\Lambda = (\alpha_f - \alpha_s)/(c_f + c_\phi)$ 根据温度相关流体性质动态计算。

• **汇项（达西）：** 水力扩散率 $c_{hyd} = k/(\mu_f S_s)$ 纳入了遵循阿伦尼乌斯定律的温度相关黏度 $\mu_f(T)$。

边界条件与初始条件

模拟将试样初始化为地应力平衡状态：

• 初始状态： $p_0' = 46$ MPa， $q_0 = 15$ MPa， $u_{w,\,0} = 4.0$ MPa， $T_0 = 25$℃。

• 力学边界条件：保持总应力恒定（$\sigma_{radial} = 45$MPa， $\sigma_{axial} = 60$ MPa），以模拟三轴蠕变试验。

• 热边界条件：在外缘（$r = R$）施加狄利克雷边界条件 $T(t) = 25 + 1.0 \cdot t$ (℃)，从而产生径向温度梯度。

• 水力边界条件：Agent模型在 $r = R$处施加排水边界（$u_w = u_{w,0}$），而朴素模型则在各处施加无通量条件（$\nabla u_w \cdot \mathbf{n} = 0$）。

基于算子分裂的数值实现

耦合系统采用交错算子分裂格式求解，时间步长为 $\Delta t = 0.5$ s：

1. 热学步：显式求解热扩散方程 $\partial_t T = \alpha_{th} \nabla^2 T$。2. 水力步：通过叠加局部热增压增量 $\Delta u_{TP} = B\Lambda\Delta T$ 并求解隐式达西扩散方程，更新孔隙压力 u_w。

3. 力学步：更新有效应力 $\boldsymbol{\sigma}' = \boldsymbol{\sigma}_{total} - u_w\mathbf{I}$ ，并执行最近点投影返回映射算法，以满足MCC塑性和拉伸约束。

这一严谨的数值设置确保，结果的差异（图 5）完全源于是否存在智能体识别出的耗散机制，从而将框架的认知贡献单独分离出来。

5. 结论

我们提出了一种神经符号生成智能体，能够自主应对多物理场模拟的复杂性。通过将物理知识拆解为语义化的“本构技能”，并利用思维链推理过程加以编排，该智能体弥合了非结构化科学文献与严谨数值执行之间的鸿沟。

本研究得出了三个关键结论：

1. **上下文感知推理：** 智能体成功识别出，文献中普遍采用的“不排水”假设并不适用于Rothbach砂岩所承受的特定热加载速率，并自主启用Darcy耗散机制来解决这一冲突。

2. **涌现物理有效性：** 在无需人工干预的情况下合成的全耦合模拟，预测出了一条安全的有效应力路径（$p' \approx 8.9$ MPa），避免了朴素模型常见的假阳性失效预测（$p' \to 0$）。

3. **方法稳健性：** 严格的逐单元验证和状态区检测机制确保 AI 的创造力受到物理定律的约束，从而有效降低科学计算中出现幻觉的风险。

这项工作表明，AI 不仅可以充当编程助手，还可以成为推理伙伴，避免模拟受到那些常常损害人工建模的隐性认知偏差和简化处理的影响。

表 1：**Rothbach砂岩的材料参数和模拟条件**。参数按物理领域分类：力学（M）、水力学（H）和热学（T）。星号（*）表示在模拟过程中动态更新的温度相关属性。

参数	**符号**	**数值**	**Unit**
力学属性（*M*）			
临界状态线斜率	M	1.2	-
先期固结压力	p_{c0}	60.0	MPa
正常压缩指数	λ	0.15	-
再压缩指数	κ	0.03	-
初始孔隙比	e_0	0.3	-
泊松比	ν	0.3	-
水力特性 *(H)*			
固有渗透率	k	1.0×10^{-16}	m^2
流体黏度*	$\mu_f(T)$	阿伦尼乌斯	Pa·s
流体压缩性	c_f	4.0×10^{-10}	Pa^{-1}
固体压缩系数	c_s	2.0×10^{-11}	Pa^{-1}
比贮水量	S_s	1.28×10^{-10}	Pa^{-1}
热学性质 *(T)*			
热导率	λ_T	2.0	W/(m·K)
比热容	C_p	1000	J/(kg·K)
流体膨胀系数*	$\alpha_f\ (T)$	$\sim 3.0 \times 10^{-4}$	K^{-1}
固体膨胀系数	α_s	3.3×10^{-5}	K^{-1}
热扩散率	α_{th}	1.0×10^{-6}	m^2/s
初始条件与边界条件			
初始围压	σ_3'	41.0	MPa
初始孔隙压力	$u_{w,0}$	4.0	MPa
初始温度	T_0	25.0	°C
升温速率	$\dot{T}$	1.0	$^\circ\mathrm{C}/\mathrm{s}$

数据可用性

当前研究中生成和/或分析的数据集可在 GitHub 仓库中获取：https://github.com/shuWuYue123/Neuro-Symbolic-Auto-Coupling。

代码可用性

神经符号生成智能体的源代码和模拟脚本可在 GitHub 仓库中获取：https://github.com/shuWuYue123/Neuro-Symbolic-Auto-Coupling。

参考文献

[1] J. Jumper、R. Evans、A. Pritzel、T. Green、M. Figurnov、O. Ronneberger、K. Tunyasuvunakool、R. Bates、A. Žídek、A. Potapenko 等，《利用 AlphaFold 实现高精度蛋白质结构预测》，Nature 596（7873）（2021）：583–589。

[2] A. Merchant、S. Batzner、S. S. Schoenholz、M. Aykol、G. Cheon、E. D. Cubuk，《面向材料发现的深度学习扩展》，Nature 624 (7990) (2023年) 80–85。

[3] M. Raissi、P. Perdikaris、G. E. Karniadakis，《物理信息神经网络：用于求解涉及非线性偏微分方程的正向问题和逆问题的深度学习框架》，《计算物理学杂志》378 (2019) 686–707。

[4] Z. Li、N. B. Kovachki、K. Azizzadenesheli、Burigede Liu、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar，《用于参数化偏微分方程的傅里叶神经算子》，载于：国际学习表征会议，第 2 卷，2021 年，第 4 页。

[5] L. Lu、P. Jin、G. Pang、Z. Zhang、G. E. Karniadakis，《基于算子通用逼近定理通过 DeepONet 学习非线性算子》，Nature Machine Intelligence 3(3)（2021）218–229。

[6] H. Wang、T. Fu、Y. Du、W. Gao、K. Huang、Z. Liu、P. Chandak、S. Liu、P. Van Katwyk、A. Deac 等，《人工智能时代的科学发现》，Nature 620（7972）（2023）47–60。[7] O. Zienkiewicz、R. Taylor，《有限元方法丛书》，Butterworth-Heinemann，牛津，2005年。[8] H. K. Versteeg，《计算流体动力学导论：有限体积法》，第2版，培生印度教育，新德里，2007年。[9] D. A. Boiko、R. MacKnight、B. Kline、G. Gomes，《利用大语言模型开展自主化学研究》，Nature 624（7992）（2023）570–578。

[10] A. M. Bran, S. Cox, O. Schilter, C. Baldassari, A.D.White, P.Sch waller，借助化学工具增强大语言模型，Nature Machine Intelligence 6 (5) (2024年) 525–535。

[11]G. Wang, Y. Xie, Y. Jiang, A. Mandlekar, C. Xiao, Y. Zhu, L. Fan, A. Anandkumar，Voyager：一种基于大语言模型的开放式具身智能体，arXiv预印本 arXiv:2305.16291（2023年）。

[12] J. S. Park, J. O’Brien, C. J. Cai, M. R. Morris, P. Liang, M. S. Bernstein，生成式智能体：人类行为的交互式模拟，载于：第36届年度ACM用户界面软件与技术研讨会论文集，2023年，第1–22页。

[13] A. Anthropic，《Claude 3模型系列：Opus、Sonnet、Haiku》，Claude-3模型卡 1 (1) (2024年) 4。

[14] G. Team, R. Anil, S. Borgeaud, J.-B. Alayrac, J. Yu, R. Soricut, J. Schalk- wyk, A. M. Dai, A. Hauth, K. Millican 等，《Gemini：一系列功能强大的多模态模型》，arXiv 预印本 arXiv:2312.11805（2023年）。[15] J. Wei, Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud, D. Yo- gatama, M. Bosma, D. Zhou, D. Metzler 等，《大语言模型的涌现能力》，arXiv 预印本 arXiv:2206.07682（2022年）。

[16] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg 等，《通用人工智能的火花：使用 GPT-4 的早期实验》，arXiv 预印本 arXiv:2303.12712（2023年）。

[17] J. Wei、X. Wang、D. Schuurmans、M. Bosma、F. Xia、E. Chi、Q. V. Le、D. Zhou 等，思维链提示能够引发大语言模型的推理能力，《神经信息处理系统进展》第35卷（2022年），24824–24837页。

[18] S. Yao、D. Yu、J. Zhao、I. Shafran、T. Griffiths、Y. Cao、K. Narasimhan，《思维树：利用大语言模型进行审慎的问题求解》，《神经信息处理系统进展》第36卷（2023年），11809–11822页。

[19] S. Yao、J. Zhao、D. Yu、N. Du、I. Shafran、K. R. Narasimhan、Y. Cao，React：在语言模型中协同推理与行动，载于：第十一届国际学习表征会议，2022年。

[20] T. Schick、J. Dwivedi-Yu、R. Dessì、R. Raileanu、M. Lomeli、E. Hambro、L. Zettlemoyer、N. Cancedda、T. Scialom，Toolformer：语言模型能够自学使用工具，《神经信息处理系统进展》36（2023年），68539–68551。

[21] Z. Xi、W. Chen、X. Guo、W. He、Y. Ding、B. Hong、M. Zhang、J. Wang、S. Jin、E. Zhou 等，大语言模型智能体的兴起与潜力：一项综述，《中国科学：信息科学》68（2）（2025），121101。

[22] Y. Shen、K. Song、X. Tan、D. Li、W. Lu、Y. Zhuang，HuggingGPT：利用 ChatGPT 及其在 Hugging Face 上的伙伴解决 AI 任务，《神经信息处理系统进展》36（2023年），38154–38180。

[23] A. Brohan, Y. Chebotar, C. Finn, K. Hausman, A. Herzog, D. Ho, J. Ibarz, A. Irpan, E. Jang, R. Julian 等，《按我能做的做，而非按我说的做：将语言 grounding 到机器人可供性》，载于：机器人学习会议，PMLR，2023年，第 287–318 页。

[24] Y. J. Ma, W. Liang, G. Wang, D.-A. Huang, O. Bastani, D. Jayaraman, Y. Zhu, L. Fan, A. Anandkumar, 《Eureka：通过大语言模型编程实现人类水平的奖励设计》，载于：第十二届国际学习表征会议，2024年。

[25]Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, P. Fung, 《自然语言生成中的幻觉综述》，ACM 计算调查 55(12)(2023年)1–38。

[26] X. Wang, Z. Hu, P. Lu, Y.Zhu, J.Zhang, S. Subramaniam, A. R. Loomba, S. Zhang, Y. Sun, W. Wang, 《SciBench：评估大语言模型解决大学水平科学问题的能力》，arXiv 预印本 arXiv:2307.10635（2023年）。

[27] A. Lewkowycz, A. Andreassen, D.Dohan, E.Dyer, H. Michalewski, V.Ra- masesh, A. Slone, C. Anil, I. Schlag, T. Gutman-Solo 等, 《使用语言模型解决定量推理问题》，神经信息处理系统进展 35（2022年）3843–3857。

[28] R. Taylor、M. Kardas、G. Cucurull、T. Scialom、A. Hartshorn、E. Saravia、A. Poulton、V. Kerkez、R. Stojnic, 《Galactica：面向科学的大型语言模型》，arXiv 预印本 arXiv:2211.09085（2022年）。[29] D. Gaston、C. Newman、G. Hansen、D. Lebrun-Grandie, 《MOOSE：用于耦合非线性方程组的并行计算框架》, 《核工程与设计》239（10）（2009年）1768–1778。[30] I. A. Baratta、J. P. Dean、J. S. Dokken、M. Habera、J. S. Hale、C. N. Richardson、M. E. Rognes、M. W. Scroggs、N. Sime、G. N. Wells, 《DOLFINx：新一代 FEniCS 问题求解环境》（2025年12月）。doi:10.5281/zenodo.18101307。

[31] T. Xue, S. Liao, Z. Gan, C. Park, X. Xie, W. K. Liu, J. Cao, Jax-fem：用于自动逆向设计和机理数据科学的可微分GPU加速三维有限元求解器

《计算机物理通讯》291（2023年）108802。[32] T. Belytschko, W. K. Liu, B. Moran, K. Elkhodary, 《连续体与结构的非线性有限元》，John Wiley & Sons，奇切斯特，2014年。[33] S.-M. Udrescu, M. Tegmark，AI Feynman：一种受物理学启发的符号回归方法，《科学进展》6（16）（2020年）eaay2631。[34]M. Krenn, R. Pollice, S. Y. Guo, M. Aldeghi, A. Cervera-Lierta, P. Friederich, G. dos Passos Gomes, F. Häse, A. Jinich, A. Nigam 等，《利用人工智能实现科学理解》，《自然综述：物理学》4（12）（2022年）761–769。[35] M. K. Sarker, L. Zhou, A. Eberhart, P. Hitzler，《神经符号人工智能：当前趋势》，《人工智能通讯》34（3）（2022年）197–209。[36] B. Sun，毛细上升动力学的无奇异近似解析解，《中国科学：物理学、力学与天文学》61（8）（2018年）084721。[37] S. Ghabezloo, J. Sulem，流体饱和岩石的应力相关热增压，《岩石力学与岩土工程》42（1）（2009年）1–24。