

理论物理已被 AI 攻下一城: AI - Newton 解读

方尤乐 见东山 李 想

(北京大学物理学院 100871)

1. AI 探寻物理规律之问 —— 从人类探索史说起

在科学的漫漫长河中,人类对物理规律的探寻犹如一场波澜壮阔的冒险。在近千年的历史中,人们发现了很多简洁、普适的物理规律,它们不仅刻画了人类对自然界物理机制的认识,由它们驱动的技术革命也塑造了人类今日的生活形态。

17 世纪,伽利略通过对自由落体运动的研究,推翻了亚里士多德的错误观点,为经典力学的发展奠定了基础。牛顿站在巨人的肩膀上,综合前人的研究成果,发现了万有引力定律和牛顿运动定律,构建起经典力学体系。这些理论成果直接推动了第一次工业革命。到了 19 世纪,人们由麦克斯韦方程组建立了完整的电磁理论,各种热力学规律也被

相继发现。这些物理学成就促成了第二次工业革命。进入 20 世纪,量子力学的诞生更是打开了微观世界的大门,推动了第三次工业革命中信息技术的飞速发展。

如今,随着科技的进步,人工智能(AI)逐渐走进了人们的视野并已广泛融入生活中。从智能语音助手到自动驾驶汽车,AI 的身影无处不在,深刻改变着人们的生活和工作方式。在这样的背景下,一个有趣的问题引发了学界的广泛关注:人类从近乎一片空白开始,花费数千年发现了若干简洁而普适的物理规律,AI 是否也能做到呢?

2. AI 探寻物理规律的坎坷征途

在回答“如何让 AI 自主探索物理规律”这一问题上,人们从不同方向展开了积极尝试。当下最热

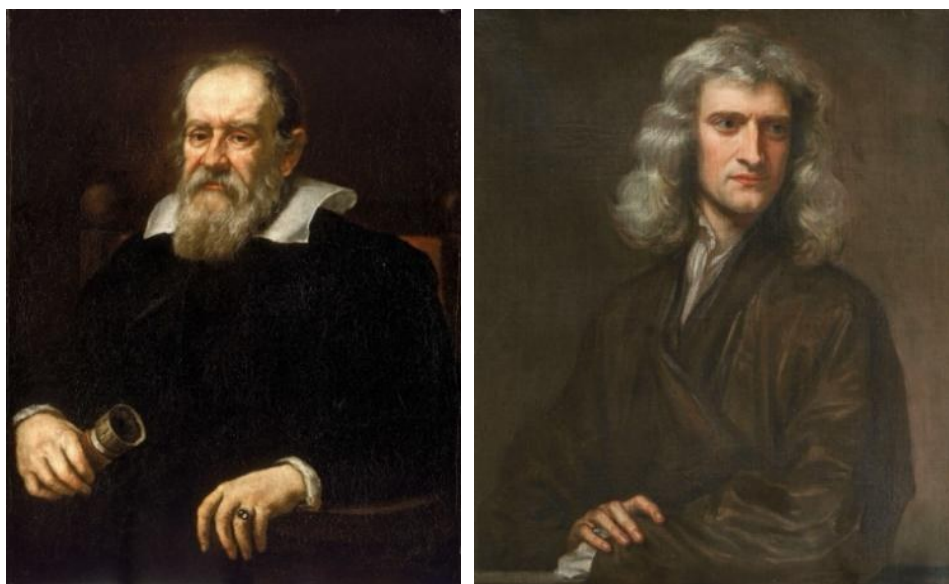


图1 伽利略与牛顿

门的当属各种基于神经网络的方法。凭借强大的表达和学习能力,它在处理大规模数据和识别复杂模式时表现出色,能精准捕捉到数据背后隐藏的信息。但它也存在明显的短板,其运行机制犹如一个黑箱,内部的运作过程难以理解,可解释性差。而且,当面对与训练数据差异较大的全新问题时,它就像一个习惯了固定解题模式的学生一样难以适应,外推性能受到很大限制。

与此同时,符号主义方法也备受关注,其中与探索物理规律联系较紧密的主要方法是符号回归。区别于传统回归,符号回归同时搜索函数的表达式结构和参数来拟合实验数据。与神经网络方法不同,符号回归最终借助显式的数学表达式刻画物理规律,因此具有天然的可解释性。特别是近期在大语言模型的加持下,符号回归方法得以借助跨学科知识,高效地生成符合物理直觉的候选方程,为物理规律的发现提供了极具价值的方向。然而,在处理复杂的高自由度系统时,符号回归所搜索的表达式空间会急剧膨胀,如同一个错综复杂的迷宫,寻找正确路径的难度超乎想象。此外,上述方法提炼出的知识常常仅适用于特定问题,缺乏一般性,难以在不同的实验场景中直接应用,未能达到人类一直以来追求用普适规律描述各种自然现象的高度。可以说,让AI自动探寻普适的物理规律这条道路上

的研究,依然处于刚刚起步的阶段。

3. AI - Newton——自动发掘普适规律的首场胜利

在AI探索物理规律面临诸多困境的背景下,人们始终在不懈努力寻求破局之法。近日,北京大学物理学院马滢青带领的研究团队在预印本【arXiv: 2504.01538】中对此问题进行了深入研究,提出了一种名为AI - Newton的概念驱动的物理规律自动发掘系统,为AI驱动的自动化物理学研究带来了新的曙光。这一系统无需先验物理知识(仅需时空坐标用于表述实验数据),就能基于多个不同物理实验,从带有噪声的实验数据中自主发掘一般性物理规律,包括牛顿第二定律、能量守恒、万有引力定律等。这一过程看似简单,实则困难重重。原因在于,寻找物理规律就是在一个浩瀚无垠的函数空间中寻找特定目标,暴力搜索在实践中难以奏效;而要求所得规律具有普适性则更是一重难题。

为了攻克难点,该研究团队深入分析了人类物理学研究的底层模式,为AI - Newton搜寻物理规律设计了两个创新策略。一个策略是合情推理。不同于严格的逻辑演绎,它依据部分证据进行合理推断,这正是人类科学家研究过程中采用的“大胆猜想,小心求证”。比如,如果在一个实验中发现某个函数组合能用于构造描述该实验的方程,AI - Newton系统会猜测它在其他实验中也可能发挥作用,进而将其定义为物理概念记录下来。如果某个规律在

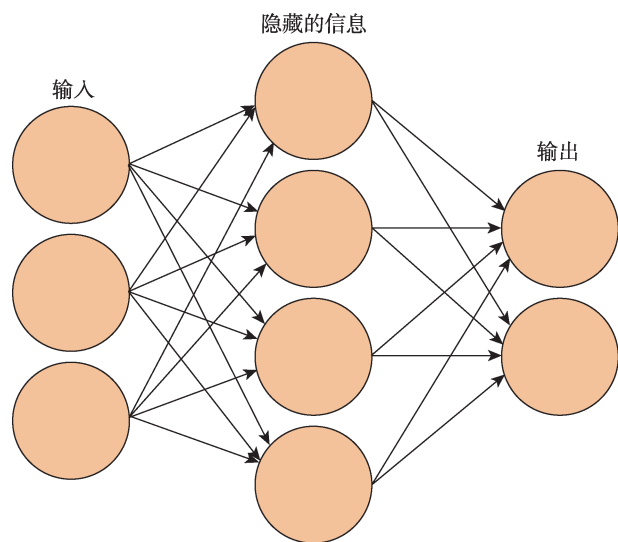


图2 神经网络示意图

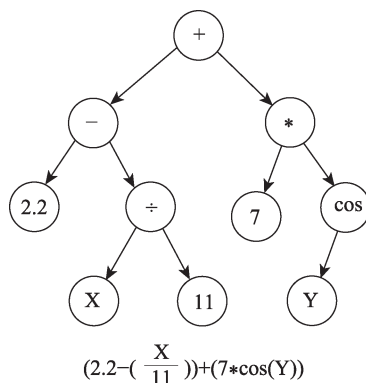


图3 符号回归示意图

部分实验中成立, AI - Newton 系统便猜测它经过简单扩充后在其他实验中也可能成立, 从而将其记录为一般性规律。以能量守恒为例, 从弹性碰撞问题中所找到的动能守恒出发, 系统会逐渐扩充该一般性规律, 发现并纳入各种与相互作用关联的势能(如弹簧系统中的弹性势能)。另一个策略是概念筛选。基于概念在各种规律中的使用频次, AI - Newton 系统赋予每个概念分数; 系统倾向于使用分值高的概念来构造新的规律, 从而将不同的规律用尽可能统一的概念进行表示, 并提高规律的搜索效率。

然而, 仅仅依靠这样两条“战术”层面上的策略, AI 自主的物理规律探寻仍然难以实现——它会轻易地陷入复杂问题, 花费大量努力而难有进展。这意味着在 AI 还需要“战略”层面上的启示。事实上, 人类的科学发现也不会直奔复杂问题、一次性定义所有物理概念或发现完整的定律, 而是遵循“简单→复杂”的累积路径。只有充分认识了简单的实验, 有了基础的概念与定律, 才能够进一步对更复杂的问题进行理论分析, 并基于相对简单的概念创造更复杂的概念。因此, 人类科学发现的历史呈现出“渐进积累”与“集中突破”两种模式的交替

演进, 甚至常依此被粗略划分成不同“时代”, 如“亚里士多德时代”、“牛顿时代”、“量子时代”。受此启发, AI-Newton 系统在前文介绍的两条策略之上, 对持续发现物理规律的流程提出了一个简单但有效的“时代控制策略”(其中每个“时代”都包含迭代中的连续若干次考察分析), 要求在某个时代中, 对单个实验的一次考察分析需要在特定的时间内终结, 这一特定的时间即可视为不同时代的标志。如果 AI-Newton 在当前时代中经过足够多的尝试后仍无法再获得新知识, 系统将指数级增加一次实验分析所分配的时间, 并过渡到一个新时代, 从而自然地模仿人类平衡深入分析旧实验和积极探索新实验的倾向。

在测试环节, AI - Newton 表现出色。它基于四十多个复杂度各异的牛顿力学经典实验(其中既包括基本的运动学实验, 也包含由多个小球、弹簧、斜面等构成的复杂动力学系统; 输入 AI-Newton 的实验数据中仅包含实验中涉及的空间坐标的时间序列), 历经约一千次迭代分析(每次分析考察一个实验), 成功重新发现了经典力学中的诸多重要概念, 如质量、动量、动能等, 以及普适的物理定律, 如牛

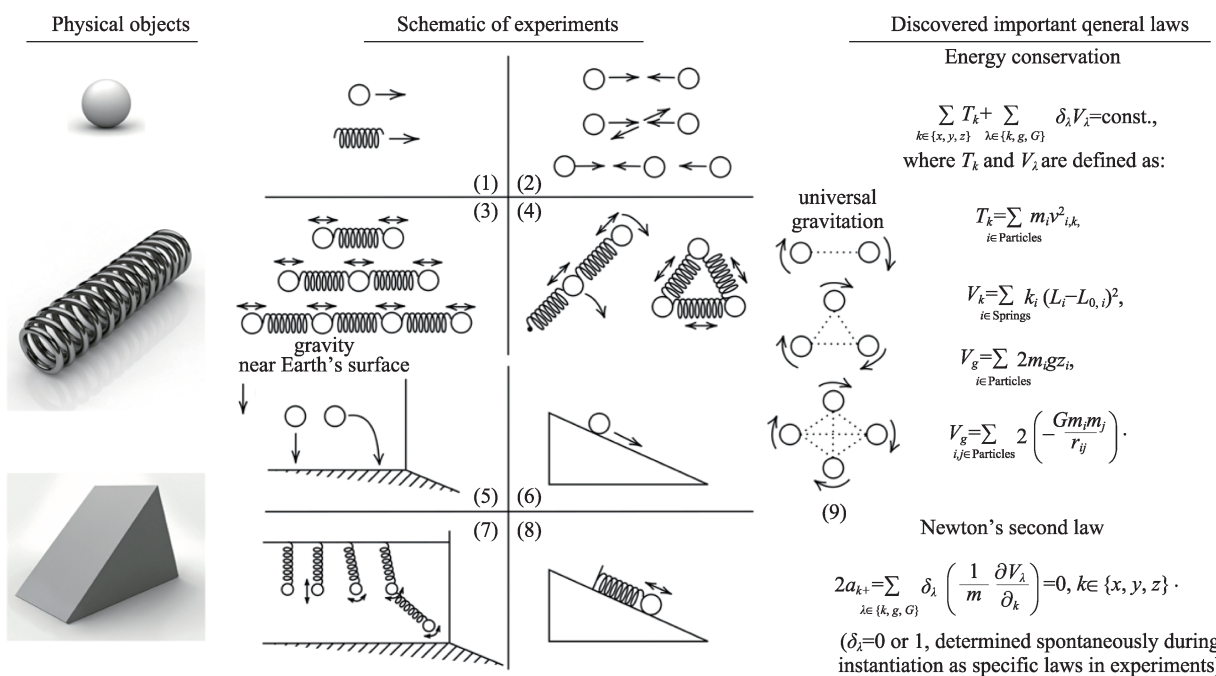


图 4 AI - Newton 测试的实验示例和所得规律示例(摘自【arXiv: 2504.01538】)

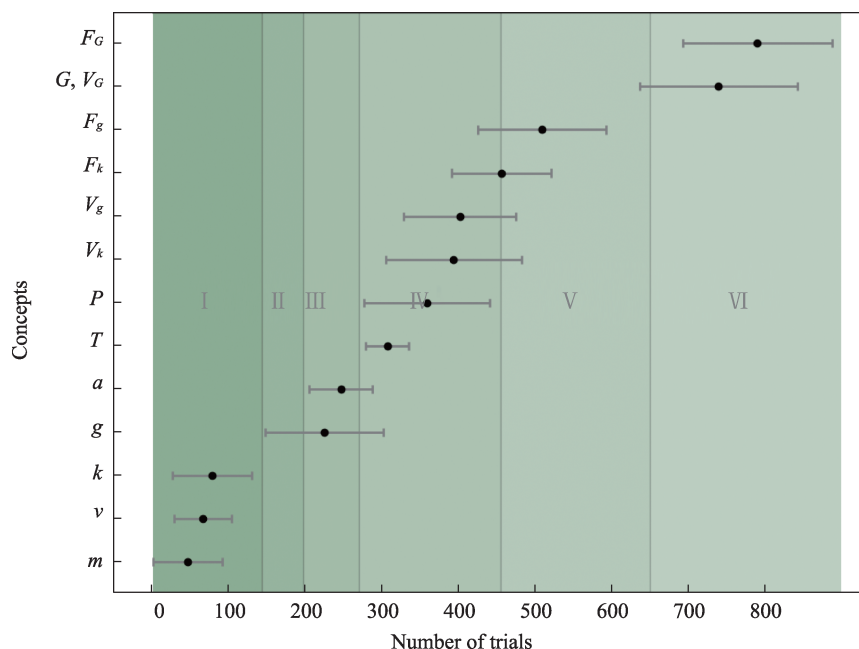


图5 AI - Newton 在不同时代探索得到的概念(摘自【arXiv: 2504.01538】)

顿第二定律、能量守恒定律和万有引力定律等。这些简洁的定律可以在高自由度的系统中轻松地实例化为针对该系统的具体表达式,从而获得描述该系统的复杂动力学方程。在这些实验上的测试不仅验证了AI - Newton发现物理规律的可行性,更是首次实现了AI自动发掘出达到人类水平的重要普适物理规律。

更具体地,从前文所述的“时代演进”的角度来看,平均而言,AI - Newton在第一个时代就能够利用较简单实验的时空测量数据,组合出恒等式,从而发现那些相对简单、基本的描述物体的内禀概念,如质量、弹性系数,以及描述运动的最简单概念——速度,并为简单的单体实验建立唯象模型。可以看到,此阶段发掘出的概念和规律其背后尚不需要复杂的数学演算。在第二、三时代的探索过程受阻,仅发现了加速度概念,从而能够描述匀变速运动;到了第四时代,随着容许的分析时间增加,动能与动量概念及其守恒定律被发现,并以此为基础迅速发现了弹性势能、重力势能等概念,建立了完整

的能量守恒定律。力的概念与牛顿第二定律随着对实验进一步的分析而被提出。而直到第六时代,才有足够长的分析时间和充分的知识积累来解决较复杂的万有引力实验,并在之前认识的定律基础上,发现了引力势能与万有引力的概念。AI-Newton的“时代演进”自发地重现了人类科学探索的时代特征。

4. AI for discovery —— 展望未来

未来融入更强大的数学工具(如向量形式、符号推理)将可显著提升AI - Newton的推导和证伪能力。此外,AI - Newton的核心框架能自然地与自然语言集成,提供更多样的知识表示和更灵活的合情推理,使热力学规律、量子力学原理等不便于用纯数学表达式描述的物理概念和一般规律,也能在系统中以一种规范的方式表达。随着系统功能的进一步发展,AI - Newton有望为前沿科学发现做出贡献,推动AI逐步替代人类科学家进行科学探索。