

AI 破解 80 年前数学难题

本报讯 5月20日,OpenAI 宣布,它的人工智能(AI)聊天机器人在所谓的单位距离问题上,证明了匈牙利数学家 Paul Erdos (1913—1996年)的观点是错误的。OpenAI 的 AI 聊天机器人利用来自数学家的一个提示,破解了这道已有 80 年历史的几何难题。目前,这一发现已得到与该公司无关的数学家的独立验证。

1946年,Erdos 推导出了一个他认为的平面上的点的最佳排列方式,即让尽可能多的点之间的距离保持在给定距离上。他还提出了一个挑战:没有人能够做得更好。

现在,OpenAI 表示,他们的系统已经做到了这一点。它是通过运用代数数论中的相关技术实现的。这使得它能够选取坐标值作为特定方程解的点。这一发现令数学家感到震惊。

“如果 Erdos 还活着,他肯定会对此一进展欣喜若狂。”美国佐治亚理工学院的数学家 Tom Trotter 说。他曾与 Erdos 共同撰写过相关论文。

OpenAI 的数学家 Sebastien Bubeck 表示,他认为这是 AI 首次在一个科研领域自主产生的重要成果。美国加利福尼亚大学伯克利分校的数学家 Tony Feng 表示:“我一直对 AI 在数

学领域的影响持审慎态度,但这次的成果实在令人难以置信。”

加拿大多伦多大学的数学家 Daniel Litt 是 OpenAI 邀请来验证这一证明的独立研究人员之一。他表示:“这是第一个完全由 AI 独立得出的研究结果,其本身就极具研究价值。”

在几何学中,点可以在平面上进行排列,并让许多对点具有相同的相互距离。例如,一个有 9 条边的正九边形就有 9 组等距的点,这是因为 9 条边的边长完全相等;而在一个正方形网格上放置 9 个点,则能形成 12 组这样的等距点对。

Erdos 证明了越来越大的网格如何能够包含大量距离相等的点,并且这个网格会以比点数增长略快的速度无限延伸。此外,他还提出一个猜想,即没有人能找到一种更好的方法来排列这么多的具有相同距离的点。

然而 OpenAI 表示已经具备了这种能力。该公司的 AI 模型利用代数数论中的技术实现了这一目标,该技术使它能够选择坐标作为特定方程解的点。Bubeck 说,该模型已生成了一条很长的思维链,使得得出这一答案的提示是一个关于 Erdos 的猜想是真还是假的开放式问题,并不是一个证明他是错误的明确要求。

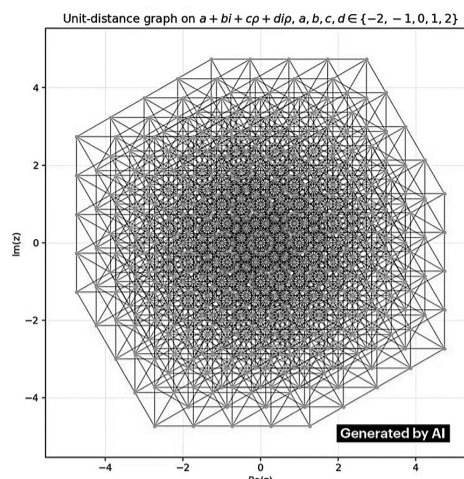
OpenAI 的数学家 Mehtaab Swahney 说:“看到这个模型像人类一样真正通过推理来解决问题,着实令人惊叹。”

这一推理过程包含在一份长达 125 页的文件中,但 OpenAI 尚未完全公布这份文件。此外,该公司也未透露其模型的具体名称。Bubeck 表示,这是一个实验性的通用推理模型,并非专为解决数学问题而设计,并且它能够根据一个提示自主完成所有工作,即对 Erdos 问题给出一个机器重写的表述。

Bubeck 表示,这种做法与利用 AI 解决数学问题的“编排”方法截然不同。在“编排”方法中,研究人员会让大语言模型(LLM)通过不断迭代的方式纠正自身的错误,从而找到问题的解决方案。

相比之下,OpenAI 系统给出的答案不会因提示语的表述方式不同而有太大差异。到目前为止,一些针对数学问题的最佳 AI 解决方案都需要大量试错,而提示语的使用已成为一门艺术。“如今,你基本可以任何你想要的方式提出问题,而模型都会正确理解这些问题。”Bubeck 说。

Litt 表示,由 AI 生成的解决方案所采用的算法来自代数数论这一事实表明,AI 模型正在



平面单位距离问题探讨的是,在无限大的纸上,最多可以画出多少条等长的线段来连接纸上的点。图片来源:Alvaro Lozano-Robledo

超越专业化“孤岛”的局限,实现更广泛的应用。他补充道,没有人能够像 LLM 那样全面掌握数学领域的文献内容。

“我们所有人都曾预料有朝一日会看到这样的情况,但没想到会这么快。”OpenAI 的数学家 Mark Sellke 说,“这与一个月前我们习以为常的情况相比,是一个巨大的飞跃。”(李木子)

“小柯”秀

一个会写科学新闻的机器人

《细胞》

伤害性神经支配限制三级淋巴结构促进肺癌

英国弗朗西斯·克里克研究所的 Leanne Li 团队发现,伤害性神经支配限制了三级淋巴结构促进肺癌。相关研究近日发表于《细胞》。

感觉神经支配调控肺部的生理与病理过程,但在其肺癌中的作用尚不清楚。

研究团队发现,肺腺癌(LUAD)的进展局部放大了伤害感觉神经的支配和激活,从而驱动了一种主要感觉神经肽——降钙素基因相关肽(CGRP)的释放。CGRP 作用于巨噬细胞的一个亚群,从而损害 CXCL13⁺成纤维细胞的募集,并阻断三级淋巴结构(TLS)的组装。这是 LUAD 后的一个关键预测因子。局部感觉去神经支配可恢复 TLS 的形成,增强 B 细胞和 T 细胞依赖的免疫应答,并抑制肿瘤生长。香烟烟雾提取物(CSE)进一步激活这种神经回路,加速 LUAD 的进展。在暴露于 CSE 的动物中,药物阻断 CGRP 可使肿瘤对免疫治疗增敏,并延长生存期。研究团队的上述发现揭示了连接伤害性神经、TLS 和 LUAD 的神经免疫轴,并确定了神经源性炎症是吸烟促进肺癌发生的一种独立于体细胞突变的机制。相关论文信息: https://doi.org/10.1016/j.cell.2026.04.038

《自然》

烯炔的脱羧烷基化

德国马克斯·普朗克煤炭研究所的 Tobias Ritter 团队报道了烯炔的脱羧烷基化。相关研究近日发表于《自然》。

研究团队报道了以羧酸为烷基来源的烯炔区域选择性而非对映选择性 C-H 烷基化反应。羧酸来源广泛且结构多样。该方法的突破性在于实现了一种极性脱羧烷基化路径,与当前基于羧酸衍生物的自由基介导 C-C 键形成模式截然不同,其关键是通过先前未被重视的途径,从氧化还原活性酯中获得稳定的烷基锌中间体。通过钯催化的烷基锌物种与由烯炔衍生的烯基钯物种的交叉偶联反应,即可高非对映选择性地获得取代烯炔。该转化适用于环状、无环、末端、内部、单取代、双取代及三取代烯炔,并可引入多样化的烷基基团。相关论文信息: https://doi.org/10.1038/s41586-026-10463-1

21 世纪中非山区出现高山火灾

美国布朗大学的 James M. Russell 团队揭示了 21 世纪中非山区高山火灾的出现。相关研究近日发表于《自然》。

野火是一种日益加剧的全球性灾害,威胁着生态系统、空气质量和社会发展。热带高山通常被认为凉爽潮湿,因此不易发生火灾。但近期在非洲最高山脉上发生的高海拔野火表明,一种新的、可能带来重要影响的威胁正在显现。由于缺乏这些环境的火灾历史记录,因此限制了学界对于火灾活动、其环境影响以及非洲山地生态系统恢复力的了解。研究团队发现,在中非鲁文佐里山脉,过去 1.2 万年中,21 世纪的一场火灾是首次影响非洲高山海拔(海拔 3800 米以上)区域的火灾;在中海拔地区(海拔 2990 米),火灾在大约 2000 年前突然增加。这与区域内人类活动变化的证据相吻合。随后是以竹子为主的生态系统扩张。该研究表明热带非洲高山的高海拔火灾是 21 世纪新出现的一种干扰,可能改变高海拔热带生态系统。相关论文信息: https://doi.org/10.1038/s41586-026-10511-w

更多内容详见科学网小柯机器人频道: http://paper.sciencenet.cn/Alnews/

54.1 公斤科学实验样品奔赴太空

(上接第 1 版)

其中,单结钙钛矿太阳能电池分为“无源”和“有源”两类样品。无源暴露样品将在开路、短路、最大功率点等 5 种不同运行工况下,开展 3 种钙钛矿组分及两种器件结构的样品性能老化分析。

对于具有完整电路系统的有源样品,科研团队则通过实时回传电压、电流数据,系统性了解其在轨过程。中国科学院半导体研究所助理研究员熊仕介绍:“我们期待能准确定位材料、器件结构、封装工艺中存在的薄弱环节,为下一代高效协同的太空光伏器件设计提供直接依据。”

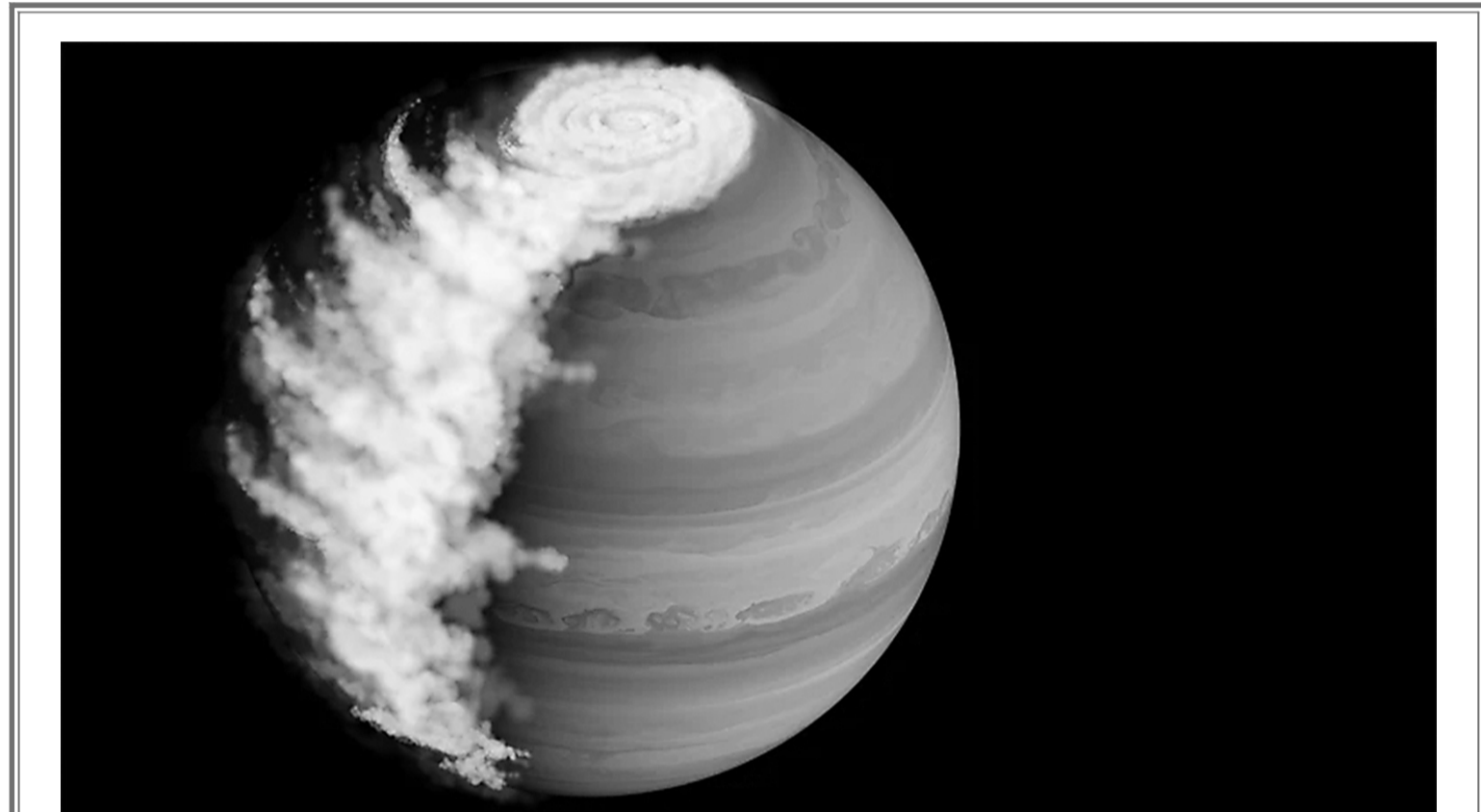
而在叠层技术方面,钙钛矿基叠层电池通过不同带隙子电池协同吸收太阳光谱,强强联手突破理论极限,充分提取太阳能的能量。

在北京理工大学前沿交叉科学研究院副教授朱成看来,如果把发电比作盖楼,叠层电池就像配合盖楼,要选出最匹配的两块砖。“我们的目标是筛选出‘最匹配、最兼容、最稳定的两种电池材料’。”他表示。

该项目共搭载 80 个无源和 10 个有源样品,将结合在轨数据、返回分析,建立天地对比寿命模型,评估其空间服役稳定性。

科研人员表示,这项实验既是重要的基础研究,也是在为人类走向太空、建立可持续能源系统寻找答案。

此外,神舟二十三号还搭载了多项重要的科学实验。在空间生命科学领域,纳米酶和放线菌将被放置于舱外辐射装置中,接受太空辐射的“洗礼”,以探究太空辐射对生物大分子及微生物的深层影响。针对空间长期滞留健康防护的脂质代谢实验,科研团队将从“生物相分离”的新视角解析微重力导致脂肪肝的机制。



艺术家笔下的 WASP-94A b。

图片来源:Hannah Robbins

科学此刻

690 光年外的“风起云涌”

一项近日发表于《科学》的研究发现,在遥远行星 WASP-94A b 的夜晚一侧,厚重的云层持续生成,但当风将它们吹向白天一侧时,云层便迅速消散了。天文学家通过研究来自 210 秒差距(690 光年)外的这颗行星宿主恒星的光,已经能够探测行星上的这种不对称天气模式,表明追踪行星自转的不同阶段,可以增进研究人员对行星大气的了解。

研究人员通常只能间接探测系外行星的存在。例如系外行星凌星过程,即当行星从其宿主恒星前方经过时,从地球上看来,部分背景恒星的光会穿过行星大气层,在那里被吸收或散射,因此可获得行星白昼和黑夜的大气光谱。基于此,研究人员可获得关于行星化学组成、天气模式甚至起源的线索。

美国亚利桑那州立大学的天文学家 Sagnick Mukherjee 及合作者将目光锁定在明亮的恒星 WASP-94A 及围绕其运行的气态巨行星 WASP-94A b。当他们用美国国家航空航天局的詹姆斯·韦布空间望远镜对准恒星 WASP-94A 时,惊喜出现了。“我们在

数据中发现了惊人的不对称性。”Mukherjee 说,望远镜揭示了在凌星的起始、中段和结束时,光谱存在细微差异。

测量结果显示,在每次凌星期间,率先

从恒星前方经过的那部分行星大气被厚厚的云层覆盖。鉴于该行星白昼侧温度至少为 1600 开尔文,研究人员推测这些云可能由矿物液滴而非水构成。凌星结束时,最后经过的那部分大气则是晴朗无云的。综合来看,随着风环绕行星吹拂,云层在夜晚一侧持续形成,进入白昼侧则迅速消散。

如果研究人员忽视了这种不对称性,数据原本可能会指向白昼侧有霾层形成,而非在夜晚一侧有云层形成。这意味着全然不同的大气化学成分。

葡萄牙天体物理与空间科学研究所的天文学家 Olivier Demangeon 表示,Mukherjee 等人以一种“清晰且有说服力”的方式“证明了昼夜差异对于推断系外行星的大气组成至关重要。”

研究人员表示,就 WASP-94A 系统而言,关于行星形成的线索可能格外有趣,因为该系统同时还在围绕另一颗名为 WASP-94B 的恒星运行。(徐锐)

相关论文信息: https://doi.org/10.1126/science.adx5903

坚持“难而正确的事”,他们将不可能变为可能

(上接第 1 版)

令人意外的是,起初,这是别人劝陆凌放弃的研究。

陆凌做光学微纳结构设计出身,但他深知,不可能所有需求都通过结构设计来实现。“就像衣服版型固然重要,但如果没有防水的材料,就无法设计出好的雨衣来。”他希望跨越领域,尝试突破现有光学材料的限制。

透明导体是手机触屏、显示器、太阳能电池、发光二极管等光电器件的核心材料。然而,透明与导电一直是互斥的属性,目前主流的透明导体都是掺杂半导体或绝缘体获得的,牺牲了部分透明性。

2005 年,有理论提出,可以通过一种非常特殊的能带结构让金属材料获得理想的透明属性。不过,这种透明金属始终未在现实世界被发现。

为寻找这样的材料,陆凌团队对整个无机材料数据库中数千万种材料进行高通量计算搜索,并做了长期的实验调研,结果一无所获。有同行劝他放弃这一研究,理由很充分——专门

的材料学家都没做出来,而陆凌团队本来就不是做材料研究的。

美国知名理论物理学家保罗·斯坦哈特曾说,世界上有两种“不可能”:第一种是像永动机那样违背物理原理的“不可能”;第二种是“不违背任何已知物理原理,但却从来没有这样的先例”。

“这个材料属于第二种不可能。”陆凌表示,原则上,如果无机材料找不到,可以尝试在种类更多的有机材料中去寻找。

前路未知,唯有尝试才有可能。而“青年团队计划”给了陆凌团队试错的底气——稳定的支持机制、宽松的科研环境,以及以科学意义和应用价值为导向,突出原创性、引领性、突破性贡献的标准,坚定了陆凌走下去的决心。

陆凌先做了有机材料的初步调研,随后鼓励吴正再继续“试试”。吴正再顺着线索,发现有一类已知能导电的有机材料可能符合要求。于是,他从头开始学习材料制备、光学和电学性质的表征与测试。经过漫长的实验周期,他们最终

新型 AI 聊天机器人 有助抵抗错误健康信息

据新华社电 芬兰奥卢大学日前发布新闻公报说,该校研究人员与国际同行合作开发出一款人工智能(AI)聊天机器人,通过使用“认知接种”方法,可帮助人们抵抗健康领域的错误信息。

公报说,近年来,生成式 AI 工具快速发展,使虚假或误导性内容更容易被大规模制造和传播,与此同时,AI 也可以被用于抵抗错误信息。譬如,这款名为 Forty 的聊天机器人采用源于社会心理学的“认知接种”方法,基本思路是让人们在受控条件下接触“弱化版”错误观点,从而提升在现实环境中对错误信息的抵抗力。研究团队说,其原理类似人体通过接种疫苗而增强免疫系统能力。

研究团队将 Forty 部署在公共网站上,围绕每日刷牙、饮酒行为、身体活动与心理健康的关系等常见健康议题,与用户开展“结构化”对话。这是指聊天机器人依据预设的科学教育流程,引导用户识别误导性观点,分析其逻辑漏洞,形成更稳固的正确认识。

为验证系统效果,研究人员邀请 65 名参与者与 Forty 对话。结果显示,相比单纯阅读或写作,与这个聊天机器人互动更能有效提升参与者对错误健康信息的抵抗力。

研究人员表示,接下来计划将该聊天机器人的应用范围扩展至复杂生活情境,探索对话式 AI 如何帮助人们应对日常压力、突发变故等挑战。(朱昊晨 徐谦)

美新一代“星舰”首飞 完成大部分目标

据新华社电 美国太空探索技术公司重型运载火箭“星舰”近日实施第 12 次试飞。这是新一代“星舰”系统首次整体投入飞行测试,旨在对多项升级后的关键技术进行验证。任务完成大部分测试目标。

太空探索技术公司直播画面显示,美国中部时间 5 月 22 日 17 时 30 分许(北京时间 23 日 6 时 30 分许)，“星舰”从位于得克萨斯州南部的发射基地升空。两分多钟后,火箭第一级“超级重型”助推器和第二级“星舰”飞船成功分离,但助推器没能完成返航制动燃烧,以非受控状态坠入墨西哥湾。

在完成上升燃烧后,飞船进入太空滑行阶段,按计划部署了 20 颗“星链”模拟卫星,这些模拟卫星尺寸与下一代“星链”卫星相近,沿飞船同一轨道轨迹飞行,在重返大气层过程中焚毁。飞船还部署了两颗专门改装的“星链”卫星,对“星舰”热防护系统进行成像观测,并将相关图像传回地面,用于评估未来在飞船返回发射场前对热防护系统状态进行检测的方法。

由于飞船 6 台“猛禽”发动机中有 1 台提前关闭,任务团队取消在轨发动机再点火测试。

发射 40 多分钟后,飞船再入地球大气层,最终在印度洋溅落。据介绍,在飞船再入大气层阶段,工程团队故意移除一块隔热瓦片,以测试相邻区域瓦片承受的气动载荷变化。飞船还进行了此前试飞中验证过的实验性动作,包括对尾翼结构极限承载能力进行测试的机动动作,以及模拟未来返回地面基地飞行轨迹的动态倾斜机动动作。

据太空探索技术公司介绍,本次任务是新一代“星舰”系统“首飞验证”,从“星舰”飞船、“超级重型”助推器,到“猛禽”发动机、发射台,均进行了大幅升级和重新设计。

“星舰”火箭总长约 120 米,直径约 9 米,由两部分组成,第一级是长约 70 米的“超级重型”助推器,第二级是“星舰”飞船。该火箭的设计目标是将人和货物送至地球轨道、月球乃至火星。(谭晶晶)