数学家借助AI证明黎曼猜想？

丘成桐、詹姆斯·梅纳德等解读相关影响

■本报记者 韩扬眉

编者按

1859年，德国数学家伯恩哈德·黎曼提出：黎曼ζ函数的所有非平凡零点，实部都等于1/2。黎曼猜想是关于质数分布规律的难题，被视为数论大厦的“地基”，许多定理都以它的成立为前提。美国克雷数学研究所将其列入七大“千禧年数学难题”，悬赏100万美元；它也是德国数学家希尔伯特提出的“23个问题”中尚未解决的难题之一。

若此次证明最终被确认，其意义不仅在于一个古老而重要的数学问题得以解决，更标志着人机合作研究新时代的到来。

11月20日，34岁的数学家、美国人工智能（AI）企业Anthropic公司研究员莱文特·阿尔珀格（Levent Alpöge）宣称：“我相信，我与Claude合作，发现了黎曼猜想的一个证明。”

阿尔珀格发布了一份4页纸的“说明”，阐述了证明路线、与AI的分工、研究时间线，并附上14个压缩包、约1700万行的Lean形式化代码。由于Claude生成的文字证明超过25万页，他并未公开。

消息一出，立即引发广泛关注。目前外界关注的焦点有二：其一，这是一个尚未经数学家独立验证和理解的“草稿”；其二，公布的时间点——Anthropic公司于11月19日在纳斯达克上市，阿尔珀格次日即宣布这一成果，难免引发有关商业宣传动机的猜测。阿尔珀格本人也坦承“时机很糟糕”。

11月21日晚，在广东深圳参加“黎曼诞辰200周年学术研讨会”的两位菲尔兹奖得主——清华大学讲席教授丘成桐、英国牛津大学教授詹姆斯·梅纳德（James Maynard），以及清华大学丘成桐数学科学中心副教授徐文强，就此接受了媒体采访。

丘成桐表示，“毫无疑问，这是数论上的伟大贡献，可以说数学进入了一个新纪元”；梅纳德认为，有98%的可能是正确的，“是一个重大的里程碑，将成为数学史上的伟大成果之一”。

有98%的可能是正确的

问：你是怎么得知这个消息的？知道后有什么感受？

丘成桐：从昨天晚上到今天早上，海外一众媒体都在报道这件事。我感到兴奋，我们梦寐以求的重要问题解决了，现在我们要推动数学向前一大步。接下来很重要的一个问题是，哪里才是最好的、最能够吸引学者向前走的方向。希望中国的年轻数学家能够发展出一套自己的理论，这是一个很好的机会。

梅纳德：当时是晚上，我在手机上看到这个消息后，立即上网找那份证明文件，迫切想弄清一些细节。我认为有98%的可能是正确的，因此非常兴奋。我认为这是一个重大的里程碑，将成为数学史上的伟大成果之一。

徐文强：我听到这个消息时，正好与我在美国斯坦福大学时的博士生导师开会。感到震惊的同时，我们也在思考：如此重要的数学猜想被证明后，我们应该作出什么样的反应？我认为，一个数学猜想可能像一座大山，数学家的初衷是翻越山峰，看到美景，而AI工具帮助了我们。我们应该利用工具更勇敢地继续攀登，看到更多的风景，并且知道“想看什么样的风景”，即提出什么问题。这仍取决于人的品味与直觉。

问：在尚未全面核验的情况下，现在下“证明正确”的结论是否为时尚早？

梅纳德：这个证明附有形式化验证——一种由机器检查结果是否正确的方法。但仍需保持谨慎，因为还有几项检查要做。这项工作背后的研究者此前在AI和数学方面作出过重要贡献，我认为他不太可能在未经充分检查的情况下发布结果。

我的一些同事已经开始借助AI梳理证明，提取关键思想。数学界先借助AI将证明压缩到便于处理的规模，再理解它的整体思路，就有可能进一步简化它。

问：黎曼猜想被证明，对数学有何影响？

徐文强：直接影响是很多重要数学问题将成为推论，即被彻底解决。同时，此前那些建立在黎曼猜想之上的假设性结果，将变成无条件的严格证明，即许多之前看似还有“缺陷”的结果，都变得完美了。此外，可能会更接近证明哥德巴赫猜想、孪生素数猜想。

不过回到数学思想本身，我还不确定目前这一证明中有多少原创性的、历史级别的想法。

问：AI给出的论证长达25万页，人类何时能理解？有什么办法可以解决人类难以验证和理解冗长机器证明的问题？

丘成桐：我估计5年内可以全部了解并提炼其精神。验证和理解AI的证明并不是问题。做所有未知学问的第一步是知道哪个方向是对的。黎曼猜想被证明以前，不知道对错就难有巨大进步。现在知道它是对的，剩下的就是花全部精力去理解它的证明并融会贯通。

所谓25万页有可能是因为很多内容写得啰嗦，我想很快整个过程会以人类能懂的方式写出来。我对此有很大信心，因为AI使用的方法都是人类想过的，只是将其综合起来往前迈了一大步，理解它并没有多大障碍。

改变数学家的工作方式：先解题，再理解

问：人机合作后，数学研究的方式是否会改变？

丘成桐：传统数学家的工作方式是一步一步向前走，直到解决问题。这与物理学家、生物学家、化学家的做法相反，他们是先找到一个结果，再慢慢去推敲，看看有没有意义。现在AI把数学家也“扭转”过来了，AI先找到答案，数学家再消化整个过程。所以，我们一定要先花工夫融会贯通人类走过的路，然后再看机器走过的路，最后再思考朝什么方向走是最好的。

梅纳德：我会因此改变工作方式。如果我理解此次黎曼猜想为什么成立，今后我会从一个完全不同的角度看待数论，并采用新方法。这就是为什么理解才是重要的。过去“证明”与“理解”是两件相伴而生的事，现在分开了。我原以为自己在意的是那些能登上“头条”的成果，但实际上，我在意的是过去那种数学研究方式中，伴随成果而来的理解。

问：黎曼猜想长期被认为是最困难的问题之一，AI却在短时间内证明了它。这是否意味着在人类数学家看来极其困难的问题，对AI来说反而相对容易？

梅纳德：目前，AI的思考方式与人类惊人地相似。因为它是用人类全部的数学知识训练出来的，数学家认为难的问题，大多数AI也觉得难。

我认为，把证明黎曼猜想理解为“AI证明了这个问题”是不对的。AI完成了这场长跑的最后一小段，而大部分艰苦的工作是数学家在漫长时间里完成的。目前，AI不太擅长构建、发展新理论，还非常依赖人类已有的理论。在应用人类思想方面，它也许做得更好；但如果一个问题超出现有理论的范围，AI也很难解决它。

如果未来人类擅长的与AI擅长的领域存在显著差异，那么我们就需要重新审视面临的问题与所处的格局，更多地投身于那些对人类简单、对AI困难的事。

是中国年轻人“弯道超车”的好时候

问：当证明由人类和AI共同完成时，学术成果归属该如何界定？

丘成桐：功劳通常属于“指挥AI的那个人”。但是，我关心的问题不是人与机器谁重要，而是机器是不是公平地让每个人都可以使用。目前，AI操纵在少数公司手上，甚至对中国学者封禁，这是主要的不公平。

徐文强：过去数学界可能比较看重“谁是第一个证明某个问题的人”，我相信今后数学界会慢慢达成共识，或者形成新的评价体系——第一个证明的人依然有很重要的贡献，但哪个数学家、哪个群体把问题和证明真正吃透并推动理论发展，可能会得到越来越多的认可。提出问题的人同样很重要，就像黎曼。

问：为什么说这对中国年轻学者是个机会？

丘成桐：每次科学转型时，年轻人都能够发挥更大的作用。我原本估计，中国数学与国际顶尖数学水平还有一段距离，但是趁转型之机，年轻人“弯道超车”可以很快赶上。这是推动中国数学发展最好的时机，且转瞬即逝。希望国家和社会加大投入，在AI的助力下推动基础科学迅速发展。

问：数学界仍有一些担忧：数学家可能失去工作，年轻学生可能感到迷茫。你怎么看？

梅纳德：我们正在进入一个新时代，很多事情都还不确定。不确定性让人不安、让人担忧，是完全可以理解的。但我们必须保持灵活，随着AI的发展进行调整。

我认为，数学将迎来一个非常令人兴奋的黄金时代。数学家将看得更远、进步得更快。这个时候，资深数学家帮助引导年轻人走过这个阶段非常重要。我乐观地认为，数学家不会失去工作。因为数学是科学的语言，支撑着世界上许多领域的发展。

徐文强：未来数学工作者的“门槛”会降低。“门槛”更多指大家通常对数学家的期待，比如智力比较高之类的刻板印象。而其他品质的重要性可能会增加，比如是否有很好的沟通表达能力、团队协作能力。实际上，不少数学家有极强的竞争意识，在传统数学研究中，这有其存在的必要性，但如今会不可避免地带来一些负面影响，让不太喜欢或适应这种氛围的人远离数学圈，我想这其实是一种损失。我乐观地认为，未来的数学家更愿意合作，个人英雄主义可能慢慢淡化。

晚起半小时能让孩子学得更好吗？

■高小榕

这个学期，不少学生早上能多睡一会儿了。

江苏江阴从11月5日起调整作息：小学8点到校，初中为7点40分，高中为7点20分。江苏南京、徐州不少学校也把到校时间往后挪了20至30分钟。更早一些，陕西西安已要求小学上课一般不早于8点20分、中学不早于8点，四川成都、浙江宁波的一些学校则干脆取消了统一早读。前不久，“延迟到校取消早读”还冲上了微博热搜榜第一。

对此，家长反应不一。有人松了口气——孩子总算不用摸黑出门了；也有人犯嘀咕——早读没了，功课会不会落下？

孩子为什么早上起不来？睡眠时长与大脑发育有什么关系？晚到半小时，能不能真的换来多睡半小时？要回答这几个问题，得从大脑发育规律说起。

青少年“生物钟”本就偏晚

很多家长觉得孩子早上起不来是因为犯懒、不自律。其实，这与青少年的生长规律密切相关。

人什么时候犯困，大致由两个因素决定。一是生物钟，大脑里的视交叉上核根据光照“对表”，到了晚上就让身体分泌褪黑素；二是睡眠压力，醒着的时间越久，困意攒得越多。进入青春期后，褪黑素的每日初始分泌时间后移，睡眠压力也攒得慢。于是不少青少年晚上11点前怎么也睡不着，早上6点被叫醒时，大脑其实还停在“夜里”。美国儿科学会早在2014年就据此建议：初中和高中上课时间不要早于8点30分。

儿童、青少年和成年人的睡眠规律有很大差异。成年人6点起床，顶多觉得有点早；换成一个十六七岁的孩子，那种感受可能接近成年人凌晨四五点被闹钟叫醒。有意思的是，生物钟后移在高中阶段最明显，而我们的作息恰恰是年级越高到校越早。这一轮调整里，高中生仍是到校最早的那一拨人。

睡眠是大脑的“学习后台”

很多人把睡觉当成学习的“空档”。从脑科学角度看，情况恰恰相反。

孩子进入深睡眠阶段后，海马体会把白天学到的东西“过一遍”，借着慢波和睡眠纺锤波把记忆一点点转存到大脑皮层中，这些节律可以清晰地显示在睡眠脑电中。后半夜的快速眼动睡眠阶段，则与情绪加工、知识整合关系密切。闹钟提前响，先被砍掉的就是这一段。这段睡眠还与创造力有关。有实验比较过安静休息、非快速眼动睡眠阶段和快速眼动睡眠阶段，发现快速眼动睡眠阶段更有助于把不相干的信息联系起来，解决需要“灵光一现”的问题。

一个长期缺觉的孩子，丢掉的可能不只是分数，还有对美的感受和想象力。神经美学研究发现，审美体验需要奖赏、情绪和自我反思等多个脑网络协同工作，而这些大脑活动对睡眠状态都相当敏感，只是很少被关注。对美的感受和想象力，又与人的共情能力、创造力密切相关。

缺觉的后果第二天就会被带进教室。前额叶皮层掌管注意、计划和情绪调节，要到二十多岁才完全成熟，偏偏又最怕缺觉。睡不够时，前额叶“压不住”情绪脑区，人容易烦躁、注意力涣散。做过脑电实验的人大多有体会：被试一犯困，信号里的低频成分就更明显，任务表现随之变差。人坐在教室里，脑子却在打盹，早读时间再长也补不回来。

超扫描研究还提供了另一个角度。已有研究发现，课堂上师生之间的脑间同步与学生的投入程度和学习效果有关。学习不只是一个人记东西，还要和老师、同学“对上频”——昏昏沉沉的大脑很难跟上这个节拍。

长期缺觉的影响也有数据验证。美国“青少年大脑认知发展”研究分析了8300多名9至10岁儿童，发现每晚睡不足9小时的孩子，与记忆、智力和情绪健康相关的脑区同其他孩子存在差异，抑郁、焦虑、冲动等问题也更多；两年后再看，差异依然存在。国内情况也不乐观，《中国国民心理健康发展报告（2019—2020）》显示，95.5%的小学生、90.8%的初中生和84.1%的高中生睡眠时长没有达标。

推迟到校只是起点

推迟到校到底管不管用？美国西雅图有过一次现成的“实验”。2016年秋天，当地中学把上课时间从7点50分改到8点45分，研究者给学生戴上腕式监测仪记录睡眠。结果，学生上学日的睡眠中位数多了34分钟，成绩中位数提高了4.5%，出勤也有改善。多出来的睡眠主要来自起床时间变晚，入睡时间并没有什么变化。

不过，这个结果不能直接照搬。如果作业不减、晚自习照旧，或者多出来的时间被睡前刷手机“吃掉”，“晚到半小时”很容易变成“晚睡半小时”，等于白改。

所以，作息表只管得了早上，而晚上同样要管。教育部门在推迟到校的同时，最好一并考虑作业总量、晚自习结束时间和就寝时间，否则压力只是从清晨挪到深夜。初高中学生生物钟后移最明显，理应优先照顾；确实需要早到的孩子，学校得安排人看管；政策落地后效果到底如何，还需要数据说话。

研究者这边也有功课要做。目前睡眠数据大多来自问卷——孩子自己填、家长帮着填，误差不小。当前可穿戴设备、便携式脑电监测设备已比较成熟，完全可以用于客观测量。科研人员可以综合考量课堂注意力、情绪状态、创造力与审美体验以及学业表现，积累我国的长期追踪证据。还有一点容易被忽略：有的孩子天生是“早鸟”，有的是“夜猫子”，一张作息表不可能人人合适。

至于取消早读会不会耽误学习，倒不必太担心。课堂效率提上去，时间是能补回来的。家长能做的，是晚上帮孩子把手机收一收。

（作者系清华大学生物医学工程学院教授）

戈壁滩上“驯氢”记

——全球首次高压大口径氢气管道泄漏喷射火与爆炸试验突破纪实

■本报记者 廖洋 通讯员 任波

9月初，暑气还未散尽，山东科技大学教授张琦从试验场回到办公室后，仍在对着数据复盘。两个月前，作为该校安全与环境工程学院教授周刚团队的一员，张琦在45摄氏度的高温中完成了一组试验——去年冬天零下20摄氏度时开始做的全球首次高压大口径氢气管道泄漏喷射火与爆炸试验的补充试验。

为什么要做这个危险的试验？氢能是清洁能源，被视为实现“双碳”目标的重要路径。我国西部有大量风电、光伏，用这些绿电制成绿氢，再通过长输管道送往东部，相当于把西部的风光变成“绿色能源”送到千家万户。管道输送，是实现氢气大规模、长距离运输的关键。但氢气易燃易爆，安全怎么保障？泄漏后能扩散多远？点燃后爆炸威力多大？安全距离是多少米？这些问题没有现成答案，以前没人做过这么大尺度的试验。

这一次，团队成员拿到了不同温度、不同浓度、不同点火位置下的一系列完整爆炸试验数据。以前国内最大规模的氢气爆炸也不过几十立方米，而他们完成的是超200立方米的氢气爆炸。这种大尺度的爆炸数据，之前从没人拿到过。

遥控“点燃”1.2万方存量的高压氢气

2025年12月7日，国家管网集团联合中国安全生产科学研究院，在位于新疆哈密的国家管网集团管道断裂控制试验场，实施了国内首次全尺寸高压纯氢管道喷射火系列试验。这个试验场，是全球第三座、亚洲唯一的一座管道全尺寸爆破试验基地。

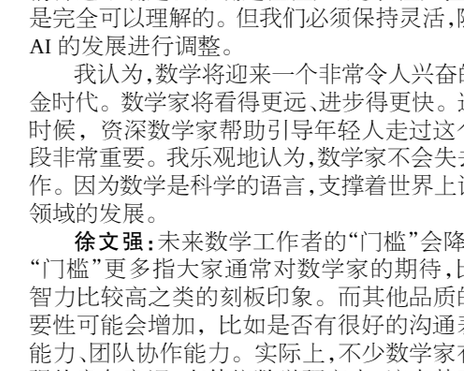
张琦当时是中国安全生产科学研究院访问学者，是这套全尺寸实验系统的总设计师之一。

试验场里，一条77米长、管径1.4米的高压氢气主管道静静躺在戈壁滩上，储氢能力达到1.2万标准立方米。试验设置了4种泄漏孔径、3个压力等级、3个喷射方向。泄漏口周围布设了热辐射计、压力传感器、高速摄影机、无人机等设备。

没人做过这么大尺度的氢气喷射火试验，每一个环节都要从头设计。整个试验场仅测控电缆就铺设了超过15公里，各种设备一次性布设到位，首次在此类试验设计中实现泄漏、点火、测试全链条无人化作业——试验期间现场没有一人。因为一旦发生意外，人在现场没有任何生还可能。

最令人捏把汗的是点火。氢气从高压管道喷出后射流速度极快：点火源放在纯氢气区域，点不着；放在空气区域，也点不着；只有放在氢气与空气浓度适当的混合区域，才能点燃。

团队提前做了大量数值模拟工作，据此设置点火位置。对小口径泄漏，他们首次提出用远控汽油喷火无人机点燃，并在正式试验前反复训练，精准控制无人机喷火方向。即便如此，意外还是发生了。一次，无人机点火后，爆炸瞬间产生的电磁干扰切断了通信信号，无人机失控，径直飞进了火里。



科研人员在试验现场合影。 受访者供图

将管道中的氢气放出来，或快速点燃，或混合后引爆，然后寻找答案。这个试验，张琦团队从夏天做到了冬天，又从冬天做到了夏天。

团队成员、山东科技大学博士研究生魏志凯回忆：“冬天，戈壁滩上的风刮在脸上跟刀子一样，戴着头套只露两只眼睛，还是扛不住。”

今年6月至7月，他们又回到同一片戈壁滩，在45摄氏度的酷暑中做补充试验。太阳下，连风都是热的。唯一能喘口气的地方是一个半地下的设备间——那里有一台给昂贵的数据采集设备降温的空调，人也能跟着沾点光。

这场试验的意义，远不止于学术层面。

2022年，国家发展改革委、国家能源局联合印发《氢能产业发展中长期规划（2021—2035年）》，明确提出统筹推进氢能基础设施建设。但是没有试验数据，就没有设计依据；没有设计依据，就没有安全标准。新疆戈壁滩上那一道道火光背后的数据，将成为这些标准体系的重要支撑。

一支“冲一流”的安全团队

张琦的底气，来自一支正在“冲一流”的学科团队。

他所在的山东科技大学安全学科，70多年来一直在跟“安全”较劲。早些年，较劲的对象是矿山里矿工头顶的岩石和脚下的水——前辈们创立了“实用矿山压力理论”，提出了预防底板突水的“下三带”理论，研发了国内首套具有自主知识产权的“智慧矿山”系统。如今，“战场”扩展到氢能储运安全。

场景在变，不变的是服务国家能源安全战略的底色。从矿山安全到氢能安全，一代代学者传下来的，是“国家需要什么，安全研究就指向哪里”的传统。这个学科在全国第五轮学科评估中稳居国内安全学科第一方阵；2023年入选山东省一流学科建设“811”项目A类潜力学科；2026年入选山东省高等学校优势特色学科和一流学科建设项目，学校在政策、人才引进上给予了全力保障。

张琦是这支队伍里的年轻人。从北京理工大学爆炸科学与安全防护全国重点实验室博士毕业后，他回到本科母校山东科技大学任教。采访中，张琦多次提到团队背后的“定盘星”——团队负责人周刚。作为带头人，周刚多次赴哈密试验场现场指导，从试验方案论证、测控系统布设到点火时序设计，每一环节都与团队成员反复推敲；每次试验结束后，他又和大家一起复盘数据、调整参数。而正是科研经费、研究生配备和跨学科协作等方面的全方位保障，让项目组成员能够心无旁骛地扎根戈壁滩。

“没有团队的支撑，我们不敢碰这么大规模的试验。”张琦说。这是实话。在前辈学者的悉心指导和学院的全力托举下，年轻一代迅速成长起来。如今，他这一代人把研究推到了氢能安全的最前沿。

张琦说：“我们做的事，就是用一场场真枪实弹的试验，给氢能产业的安全发展探路。”