

# 数学物理奥赛评测超金牌线

## 人类平均分仅得0.01的数学题,被一款上海“小”模型做对了

上海人工智能实验室(上海AI实验室)联合团队近日公布了一项面向奥赛级科学推理的新成果:他们提出一套通专融合训练方案,并基于30B-A3B架构构建出科学推理模型SU-01。在完全不调用外部工具、不执行代码、不依赖专用符号求解器的纯自然语言推理条件下,这一模型在数学和物理双学科奥赛评测中均达到金牌水平。

在第66届国际数学奥林匹克(IMO 2025)和2026年美国数学奥林匹克(USAMO 2026)中,SU-01均取得35分,超过金牌线。其中,USAMO 2026第三题拿到满分,而该题所有人类选手的平均分仅为0.01分,最高分不到5分。在2024年和2025年国际物理奥林匹克(IPhO)评测中,模型得分同样超越金牌线。

这一表现背后,是上海AI实验室“通专融合”理念的直接验证:奥赛级科学推理不一定需要超大规模模型,也不必为数学、物理等学科分别搭建繁琐的专用系统。通过统一的训练目标、奖励设计及“生成-验证-修正”推理机制,一般尺寸的模型也能在数学证明、物理推导等高难度任务中形成可复用的专家级推理能力。

研究团队将这一过程设计为“先塑造行为,再扩展能力”的后训练闭环,分三步来递进实现——

第一步是行为塑造。利用约33.8万条高质量解题轨迹进行反向困惑度课程监督微调,让通用模型学会如何组织证明、检查假设和修复漏洞,把“严谨证明”的行为范式植入模型。

第二步是能力强化。经过200步两阶段

强化学习:第一阶段提升直接求解能力,第二阶段引入证明级奖励模型,使模型不仅“答对”,更学会给出完整、可检验的推理过程。

第三步是推理扩展。面对奥赛级难题时,模型启动多轮“生成-验证-修正”循环,将训练阶段学到的自验证与自修正行为延伸至长程证明搜索中。

评测结果直观反映了这套方法的有效性。在证明质量基准ProofBench上,SU-01直接生成得分为57.6%,经推理扩展后提升至70.2%,显著优于同尺寸模型,并接近Gemini-3.1-Pro等前沿模型的表现。更值得关注的是其长程推理能力:在USAMO 2026的解题过程中,模型单次生成证明的中位长度达到10.6万个词元,修正阶段也长达8.3万个词元。这意

味着一个30B量级的通用模型,能够持续进行超过十万吨量级的有效推理,把计算资源集中用于构建逻辑、定位漏洞和完善论证。

在与人类选手的直接对比中,SU-01同样经受住了“地狱难度”的考验。USAMO 2026人类选手平均分8.59分,中位数仅6分;第三题平均分更是低至0.01分,无一人超过5分。SU-01正是在这道题上拿到满分,说明它的推理能力并非靠简单题目拉高分,而是真正具备攻克超高难度证明题的实力。

研究团队认为,科学发现是对智能的终极考验,也是验证通专融合的舞台。当AI能够像科学家一样进行严谨、长程且可自我验证的思考时,就向“AGI for Science”的目标更近了一步。 本报记者 邵阳

# 瘦小上海姑娘 为爱增重十余斤

## 107天前为陌生儿童捐干细胞,昨天再捐淋巴细胞

昨天,上海大学2024级社会学专业硕士研究生陈诗琪,用一种特殊的方式传递爱的力量——她成为上海市第815例造血干细胞捐献者,将为远方一个陌生的儿童送上生的希望。

值得一提的是,107天前,她刚为这个孩子完成造血干细胞捐献。5月初,得知孩子需要移植淋巴细胞以巩固其免疫系统,陈诗琪又义无反顾地走进复旦大学附属华山医院造血干细胞采集中心,成为上海第13例二次淋巴细胞捐献者。

“像是在平淡的青春里,再次接住了那颗命运落下的果子。”别人眼中的“壮举”,陈诗琪说起来却格外平静。她说,自己是个随心而行的人。每次过生日许愿时,她最常对自己说的就是:“希望我平和且坚定。”

2022年,陈诗琪在一次献血后主动申请加入了中华骨髓库。人海茫茫,匹配成功的概率并不大,等待的过程也可能漫长,当时她并没有多想。没想到,4年后,她就听到了生命的召唤。她说,她找不到不去的理由。“这个世界上,刚好有一个具体的人需要帮助,而自己

刚好有这个能力,那就不要犹豫,坚定地选择去帮一把吧!”陈诗琪说。为了让身体达到造血干细胞的捐献标准,这个苗条的姑娘努力将体重从90多斤,慢慢增重到了100多斤。

然而,短短百余天后,听说她要进行第二次捐献时,家人难免担忧。陈诗琪并没有急躁,她一边耐心地与红十字会、医护人员沟通,一边查阅详细的捐赠指南,科学了解整个采集过程。经过诚恳耐心的沟通,父母终于理解并支持她所做的决定。

生活里的陈诗琪,是一个充满好奇心、步履不停的女孩。上课之外,她喜欢去上海各处体验新鲜事物:今天去练击剑,明天参加不同领域的志愿活动。在不断点亮“上海地图”的过程中,她努力拓宽生活的边界。

她说,自己学的是社会学。无论是独自骑行在上海街头静静欣赏风景,还是在捐献中与红十字会、医护人员的真诚沟通,抑或是与远方素未谋面的小朋友血脉相连,这种与社会建立真实联结的过程,都让她有一种无法言说的感动。和同龄人一样,她偶尔也会



陈诗琪是上海第十三例二次淋巴细胞捐献者

对未来产生迷茫,需要在琐碎中不断寻找内心的平衡。这次捐献,恰好成为她找寻生命意义、确认自身价值的一个重要契机。

“每次想到自己切切实实帮助到了别人,我还是会感到开心与欣慰。”陈诗琪说。

记者从上海大学了解到,自2001年起,上海大学率先在沪上高校启动造血干细胞捐献志愿者招募,引领青年学子传递生命希望。上大法学院经济法系2004届本科生孙晓蕾成为全国第一位捐献造血干细胞的在校大学生。“愿病人早日康复,因你重获新生。让我们携手,把这份大爱延续下去。”孙晓蕾为师妹送上祝福。截至目前,上海大学已累计完成14例成功捐赠,用青春力量诠释责任与大爱。 本报记者 陆梓华 通讯员 余思序

本报讯(记者 陆梓华)以“积微成著,筑基未来”为主题,2026第二十三届上海教育博览会明天在上海展览中心开幕。本届教博会将继续采用线上、线下相结合的方式,其中,5月22日至24日举行线下展。线上展同步上线,持续至6月底,观众可通过搜索“上海教育博览会”小程序浏览展会。

线下展期间,将有120多家单位在上海展览中心集中展示,分为市级展区、长三角教育协同发展展区、高等教育展区、区域教育展区、职业教育展区、特色办学主题展区、协同育人实践展区。现场还将举办高质量教育主题交流分享会、职业体验、学生展演等形式多样的特色活动。为践行“健康第一”教育理念,本届教博会还特邀上海各区花样跳绳优秀代表队到场展演,开展充满青春活力的一分钟单摇跳绳交流活动。

本届线上展特设“上海教育发展成就”展示专区,观众可浏览各参展单位的特色展示内容,包括学校教育特色、创新实践成果、教师优秀教学案例、学生素质教育成果视频等,打造线上互动交流平台。

# 第二十三届上海教博会明天开幕 百多家单位集中亮相

INTERNATIONAL DAY FOR BIODIVERSITY 2026  
Acting locally for global impact  
护一方生灵 泽万物共荣  
2026.5.22  
中国·上海·崇明

2026 上海崇明生态岛国际论坛  
Shanghai Chongming Eco-Island International Forum  
Decade of Ecological Commitment Joint Stride for a Green Future  
十年共守生态初心 携手迈向绿色未来  
2026.5.22-5.24  
中国·上海·崇明