

R1 英中样稿：排版与复测报告

已完成 Nature 官方 11 页左右英中样稿、中文样稿及逐页审阅。全部真实翻译采用已有 ChatGPT 登录的 gpt-6-luna, reasoning_effort=low。

原文：Guo et al., Nature 645, 633–638 (2025).

DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning

DOI: 10.1038/s41586-025-09422-z

© The Author(s) 2025 • CC BY 4.0

<https://creativecommons.org/licenses/by/4.0/>

输入固定为 2,735,576 字节 / 11 页：

916fdaafc3b44143a9744481d079744f773d73d06372f82b1dba1076525927f2

最终全篇：65 请求 = 57 缓存命中 + 8 次真实调用，0 失败。

固定配置复跑：65 缓存命中，0 新调用。原始与坐标修订后的 mono/dual 各 11 页，文字与 72 dpi 像素均一致；PDF 字节因内部元数据变化不同。

全部双语页左侧与原文 72 dpi 像素完全一致。表 2 的 105 个数值与列/行顺序一致，统计阈值 $P < 0.01$ 完整。

本样稿是非官方翻译与排版适配。原作者、来源、版权、图表和许可保留；不表示作者或出版社认可。p3 两处行内公式复用源 PDF 对象；图标签按审阅坐标替换；p11 保留原英文图片表格，附完整中文转录。基数下标写成“63（13 进制）”，数学含义不变。

结论限于固定输入/配置/字体和 Linux 运行时。论文专用坐标修订仍需人工审阅，未声称通用 OCR、任意 PDF 全局稳定或带标签无障碍阅读顺序通过。Windows 对照合并与发布留待后续。

真实失败、增量修复与验证

完整 JSON 缺译：6 个 ID 全在，但标题续行被合并到上一行，output 为空。普通文本槽列表又发生 2/3 槽变 2/2 和公式标记注入。命名必填 schema + 程序恢复标记解决这类结构错误；仍拒绝漏项/空槽/新增标记并限定重试。

数学布局：真实引擎原文回传（35 请求、0 错误、无推理）仍复现 p3 根号/上标/变量散落。补丁连接表格 fallback 数学行，保留源二维字形和曲线；四个几何测试核对坐标、归属、正文和幂等性。样稿另修正两处跨公式中文语序。

小数尾部：源脚注将 $P < 0.01$ 拆成公式标记和 .01)。纯数值/标点片段现由程序逐字保留，只使受影响的检查点失效，离线夹具与全文实测均通过。

77 项测试通过；Node 语法、27 资源同步、diff 检查通过。六文件补丁完成应用/检查/回滚/重应用/检查。七组真实引擎合成回归全部通过，22 PDF / 54 页（审阅译文回放，无新模型调用）。

SIGTERM 实验：2 个检查点后中断，退出 130；续跑 2 命中 + 5 真实调用，0 失败，源稿不变。

最小冷启动独立调用通过，但措辞不同；新桥接进程缓存回复哈希完全一致。没有将缓存复用当作新的推理或声称全篇冷启动逐字确定。

主机断线曾留下陈旧 running 清单及英文回退 PDF，这些文件没有交付。异常最后调用的未记录 usage 不是零消费。后续使用独立观察进程并核对最终清单、中文覆盖与全部渲染页。SIGKILL 仍不能保证正常清理。

环境：Python 3.12.3 · BabelDOC 0.6.4+pdftranslate.linux-r1.1 · PyMuPDF 1.28.2 · bitstring 4.4.0 · Codex CLI 0.162.0-alpha.17.2。代码、脱敏计数、配置、资产哈希与可复现命令见同目录 REPORT.md / qa-pages.json / run-summaries.json。私有缓存、令牌、原生事件及日志未提交。

11 页逐页验收

- 01 标题、摘要、双栏、DOI 无缺漏或溢出；标题在“进行”中换行，保留为轻微排版差异。
- 02 图 1 轴标题、图例和完整 a/b 图题补译；曲线、刻度、模型 ID 与正文完整。
- 03 表 1 根号、上标、变量和推导行检查；两处行内公式用源 PDF 对象按原比例重排；碎片化注释统一字号；图 2 标签补译。
- 04 表 2 五模型列与 105 个数值顺序完全一致；统计阈值 $P < 0.01$ 完整，5 个整数单元格位置微移，未越行。
- 05 两栏正文、在线内容正常；全部文献、许可、出版社声明和版权以原英文保留。
- 06 作者及单位保留英文；只翻译页眉，中文覆盖检查对此页采用例外。
- 07 GRPO 公式 (1)–(6)、分数、上下标、符号及奖励/训练说明检查，无遮挡。
- 08 公式 (7)–(10)、跨页续文、链接和补充信息正常；文献保持原英文。
- 09 两图坐标轴、完整 a/b 图题补译；Wait 和反思词语保留为被统计的原标识符，曲线/刻度不改。
- 10 PPO/GRPO 节点及图例补译；边框、箭头、变量和图题检查；原行内数学对象保留。
- 11 英文图片表格保留；完整中文提示词/评估转录无缺字，根号及基数条件不改；含非官方改动与 CC BY 说明。

全部页由 Poppler 100 dpi 渲染并逐页查看，重点区域放大。检查缺译、重叠、字号、视觉列顺序、公式/上下标/根号、数值、图题、原英文和图内文字。普通抽取仍可能包含覆盖层底下的英文/重复文字，不是无障碍 PDF 验收。

p3 数学：原文回传失败与最终样稿

左：旧运行时源文本回传（无模型）仍损坏二维数学。右：数学几何补丁与两处行内语序修订后的中文区域。原英文始终保留在双语左页。

Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{ax + x^2}}$ = is equal to

Response: <think>

To solve the equation $\sqrt{a - \sqrt{ax + x^2}}$, let's start by squaring both ...

Rearrange to isolate the inner square root term:

Wait, wait. Wait. That's an aha moment I can flag here.

Let's reevaluate this step by step to identify whether the correct sum can be ...

We started with the equation:

First, let's square both sides:

Next, I could square both sides again, treating the equation: ...

The model learns to rethink using an anthropomorphic tone. This is also an aha moment for us, allowing us to witness the power and beauty of RL.

问题: 若 $a > 1$, 则方程 $\sqrt{a - \sqrt{ax + x^2}} = x$ 的实数解之和为

回答: <think>

为求解方程 $\sqrt{a - \sqrt{ax + x^2}} = x$, 先将等式两边平方……

移项以分离出内部的平方根项:

等等, 等等. 等等. 这是一个值得在此标记的顿悟时刻。

让我们逐步重新评估, 看看正确的和是否可以……

我们从以下方程开始:

首先, 让我们将等式两边平方:

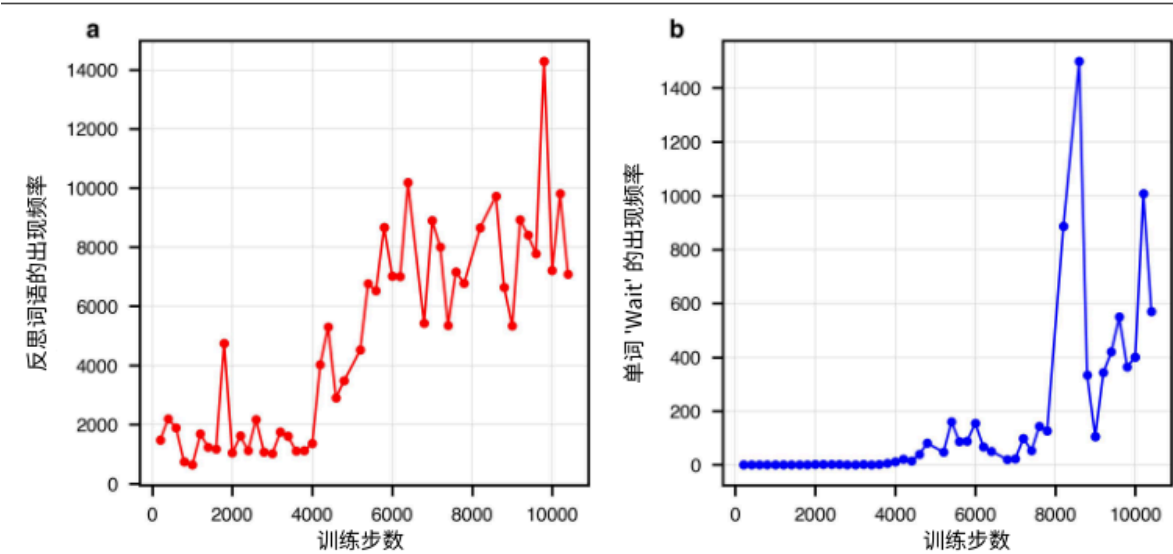
接下来, 我可以再次将等式两边平方, 并将方程视为: ……

模型学会以拟人化的语气进行反思。这对我们来说也是一个顿悟时刻, 让我们见证了强化学习 (RL) 的力量与魅力。

p4 统计脚注：小数尾部不再被译成额外正文。105 个结果值顺序逐项一致；5 个千位整数单元格最多微移 1.8 pt，仍在对应行列内。

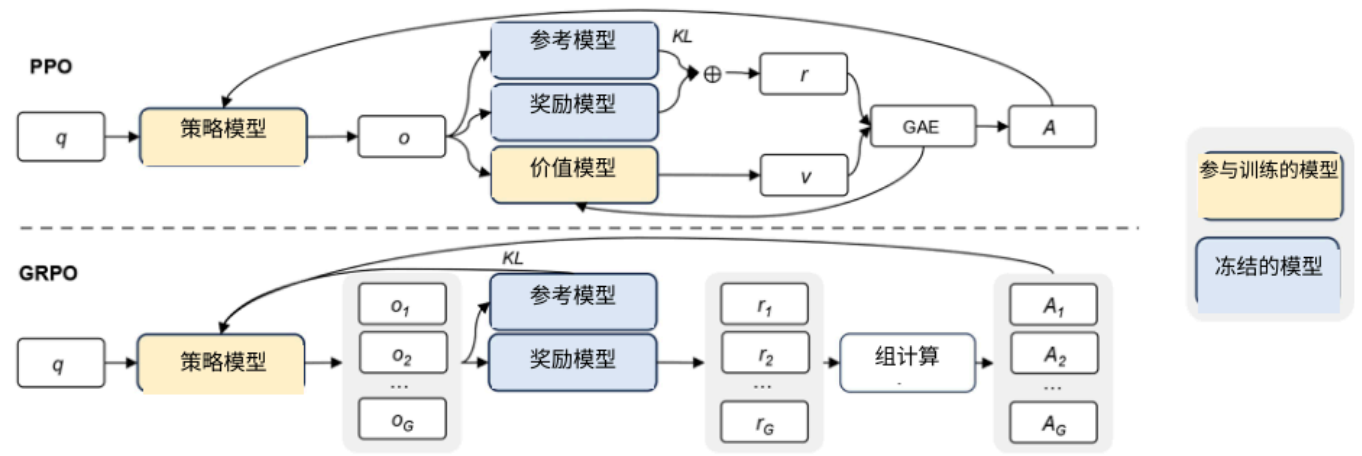
粗体数字表示性能具有统计显著性 (t检验, $P < 0.01$)。

p9 频次图：轴标题按审阅坐标补译，被统计的 Wait 标识符、数据、刻度与曲线保留。完整 a/b 图题由 Luna 作为整体翻译后分栏。



p10 图节点与 p11 图片表格

PPO/GRPO 图节点和图例补译；箭头、边框、变量保留。



p11 原英文图片表格保持，下方增加完整中文提示词与评估转录。根号、81 的例子、 $b \geq 2$ 、多于十个和 boxed 答案要求保持；非官方/来源/许可/改动声明位于本页下方。

扩展数据表 1：图片表格的中文译文

提示词

设 $b \geq 2$ 为一个整数。如果一个正整数 n 用 b 进制表示时恰有两位，且这两位数字之和等于 $\sqrt{n}$ ，则称 n 为 b -美丽数。例如，81 是 13-美丽数，因为 $81 = 63$ （13 进制），且 $6 + 3 = \sqrt{81}$ 。求使得 b -美丽数多于十个的最大整数 $b \geq 2$ 。请逐步推理，并将最终答案写在 \boxed{} 中。

评估

解析 \boxed{} 中的最终答案，并使用基于规则的评分器判断其是否等于标准答案。必要时对数值进行四舍五入，并使用 SymPy 解析表达式。

输入、输出和每页渲染的 SHA256 在 qa-pages.json。原文/译文/QA 的下载内容及 SHA256SUMS 可独立核验。未经逐页审查的中断输出和私有诊断不进入公开交付。
来源 PDF: <https://www.nature.com/articles/s41586-025-09422-z.pdf>
许可: <https://creativecommons.org/licenses/by/4.0/>