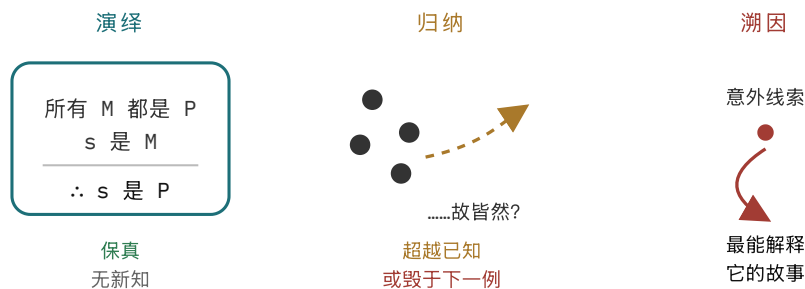


# 逻辑与有效推理

从已知通往未知有三条路，只有一条真正保险。



三种推理引擎——保证自左至右递减

演绎把你锁在安全的房间里；归纳与溯因送你出门——代价是确定性。

一个陌生人走进伦敦的诊室。不过片刻，夏洛克·福尔摩斯便断言：此人是一名刚从阿富汗归来的退役军医。华生愕然。肤色黝黑，手腕却苍白——那是在海外晒出的，并非海滨日光。手臂僵直，是战伤；面容憔悴，显露出艰辛与热病。福尔摩斯称之为「演绎」，这个词已随他流传了一个多世纪。但他用错了词。福尔摩斯所施展的——也正是令他不朽的技艺——并非演绎。它更谦逊、更冒险，也更具创造力。

这个误称是再好不过的入口，因为今日全篇围绕一个几乎人人混淆的区分：推理不止一种，而它们给出的保证并不相同。有些推理天衣无缝——只要承认前提，结论便无处可逃；另一些则丰硕却可错——它们越过证据，可能被明日的意

外推翻。把前者当后者、或把后者当前者，是人类错误中惊人比例的根源。所以我们要把界线画清楚。

前两日，我们都在推理的外围徘徊。在第1日，我们追问真信念何以成为知识——并撞上阿格里帕三难：任何理由要么无穷追问，要么陷入循环论证，要么在某处武断止步。到第2日，休谟的归纳问题表明，再多观察也无法证明一条普遍定律，因此波普尔告诉我们应去证伪而非证实。今日我们打开引擎本身。先前两个谜题其实都关乎两种特定推理模式的边界；如今我们为三者命名，看逻辑如何蜕变为数学，再跟随它抵达本课程最奇异的前沿——以零容错的精确度核查证明的机器。今日最明亮的线索，是计算。



佩吉特笔下的福尔摩斯成了「演绎」的公众形象；然而那些著名的诊断式跳跃，通常先是溯因：线索在前，最佳解释在后。

—— 模型

## 三种引擎，三种担保

如果今日只能记住一件事，就记住这个三分法。推理不是一种活动，而是三种——按承诺的大小排列。

演绎是保真引擎。结论早已蕴涵在前提之中；有效的演绎只是把它展开。承认所有人都会死，且苏格拉底是人，你便无法回避「苏格拉底会死」——否认它便是自相矛盾。这种安全的代价，是演绎具有**非扩展性**：它从不向你揭示关于世界的真正新知，只是重新排列你已拥有的东西。数学是演绎艺术被推至极限的形态，这也正是数学家何以如此笃定——以及他们的确定性为何永远无法回答关于这个宇宙的任何一个问题。

归纳是概括引擎。你见过太阳升起一万次，便推断它明日仍将升起。直到1697年以前，所有人记录下的天鹅都是白色，于是「所有天鹅皆白」看似牢不可破。归纳是**扩展性**的：它增添内容，将已有的案例推向未知。正因如此，它不保真。这是第2日休谟埋下的炸弹，至今仍在滴答作响：任何有限次的观察，都无法在逻辑上担保下一次。归纳是经验知识真正成长的方式，但它不提供逻辑保证。

溯因是解释引擎——也是多数人从未学过名称的那一种。你遇到一个令人惊讶的事实，于是寻找一个假设：若它为真，惊讶便会消散。美国博学家查尔斯·桑德斯·皮尔士（Charles Sanders Peirce, 1839-1914）将它单独提出，视为唯一真正具有创造性的模式：它不是检验或展开旧观念，而是生成新观念。「科学的每一块进步之板，最初都是由溯因独自铺就的。」他写道。演绎与归纳处理你已有的假设；溯因则回答假设最初从何而来。

回到福尔摩斯。黝黑肤色、僵直手臂、憔悴面容——这些都是令人惊讶的事实；福尔摩斯一跃而至最能同时解释它们的假设：一名从炎热战场归来的战伤军医。但请注意，这一跃并无担保。此人也可能是个演员，夏天在摩洛哥度假，打网球时扭伤了肩膀。福尔摩斯的结论是最佳解释，而非唯一解释——这正是溯因的标志——而非演绎。（这一点在[第4日](#)还会回来，届时我们将用概率把「最佳解释」变得精确。）

### 一个伟大误称的形态

福尔摩斯并不孤单。我们说医生「诊断」——那是溯因，从症状推理到最可能产生它们的疾病。听发动机的技工、勘查犯罪现场的侦探、盯着反常读数的科学家：他们都在溯因，都在跃向那个能把奇怪变得平常的解释。甚至你读到的这句话也依赖它——你推断这些文字背后有一颗心智，因为这是它们有序排列的最佳解释，而非某个定理强迫你如此推断。溯因是我们游于其中的水；我们只是很少叫出它的名字。

### —— 人人都容易混淆的区分

## 有效并不等于为真

在演绎引擎内部，住着整个逻辑中最易被误解的观念；厘清它，比背诵一打谬误更能切中要害。那就是有效性与健全性的区别。

一个论证是有效的，当且仅当它的形式保证：前提为真时，结论必为真。有效性是形状的性质，而非内容的性质——它只问论证的骨架，不问骨架里装了什么。《互联网哲学百科全书》表述得干净利落：一个论证有效，「当且仅当它的形式使得前提为真而结论为假成为不可能」。而健全性要求更多——一个论证是健全的，仅当它既有效，且所有前提实际为真。

真正容易令人失足的是这一点：一个有效论证完全可能导出一个荒诞的假结论。请看：

*所有鸟都会飞。企鹅是鸟。因此，企鹅会飞。*

形式毫无瑕疵——「所有 M 都是 P；s 是 M；因此 s 是 P」，正是「苏格拉底会死」那一例所套用的模子。若前提为真，结论就不得不跟随。所以这个论证完全有效。但它也显然不可靠，因为第一个前提是假的：并非所有鸟都会飞。有效性只认证管道的结构；健全性还要追问管中流淌的是否为清水。一个有效却不健全的论证，就像一条做工完美的管道，输送的却是污水。

这可不是钻牛角尖。它是归谬法——数学中最锋利的工具之一——背后的工作原理：要证明某前提为假，就先假设它，有效地推理到一个你已经知道为假的结论，于是假结论便逆流而上，反证前提为假。整个技巧恰恰依赖一个有效论证故意产出假结论。有效性是舟，真理是货；学会分别追踪二者，你读任何论证时都会少一层迷雾。

—— 当形式破裂时

## 藏在每个「如果」里的两种谬误

若有效形式是安全路径，谬误便是伪装成同一路径的陷阱。其中最危险的一批藏在条件推理——「若 P，则 Q」形式的命题——之中，因为无效式与有效式往往只有一步之遥。

两个有效招式是老朋友。肯定前件式：若 P 则 Q；P 真；故 Q。否定后件式：若 P 则 Q；Q 假；故 P 假。两者滴水不漏。现在轮到它们那对危险的孪生冒牌货登场。

肯定后件的推法是：若 P 则 Q；Q 真；故 P。它抓错了箭头方向。「若某人住在圣迭戈，他就住在加利福尼亚。Joe 住在加利福尼亚。因此 Joe 住在圣迭戈。」但加利福尼亚很大；Joe 完全可能在萨克拉门托。结论可能为真，这正是该谬误如此诱人的原因——它有时碰巧命中正确答案——而一个通过有缺陷的论证到达的真结论，正是第 1 日那个盖梯尔陷阱穿上了逻辑学家的外衣。

否定前件是它的镜像：若 P 则 Q；P 假；故 Q 假。「如果下雨，地面会湿。没下雨。所以地面不湿。」但别忘了洒水器、爆裂的水管、打翻的水桶。排除一个原因，并不等于排除结果本身；同一结果完全可以有多条来路。

一个教学经典例子能把结构刻进记忆：若一只动物是狗，它就有四条腿。这只动物有四条腿。因此它是狗。猫、马，甚至桌子都会抗议。这种荒谬正是关键——它与圣迭戈例子共用同一种破碎形式，只是把荒诞感放大，让齿轮滑脱清晰可见。（欧仁·尤内斯库在他的戏剧《犀牛》中整整一幕都建立在这个谬误之上：一位逻辑学家庄严地证明，一只有四条腿的猫必定是狗。）

这些是形式谬误——骨架断裂。它们的近亲，非形式谬误，缺陷不在形式而在内容：post hoc ergo propter hoc（公鸡打鸣，太阳升起，因此公鸡召唤了黎明）、人身攻击、悄悄偷换词义的歧义。形式谬误靠检查骨架便可识破；非形式谬误则要读清文字实际在做什么。

## 条件论证形式表

形式	模式	判定	理由
肯定前 件式	若 P 则 Q; P; 故 Q	有效	肯定充分条件，结论便逃不掉。
否定后 件式	若 P 则 Q; 非- Q; 故非-P	有效	若 Q 必随 P 而来，则 Q 不在场便可排除 P。
肯定后 件	若 P 则 Q; Q; 故 P	无效	Q 可能有别的原因：Joe 可以住在加利福尼亚，却不住在圣迭戈。
否定前 件	若 P 则 Q; 非- P; 故非-Q	无效	排除一个充分原因，不等于排除 Q 的所有来路：洒水器仍可打湿地面。

—— 脉络

## 逻辑如何变成数学

你正在使用的这套机制有着深远的历史，并最终转向一个出人意料的方向：在二十三个世纪里，对好论证的研究慢慢变成了一门代数的分支。这个故事有四座里程碑。

亚里士多德（公元前 4 世纪）在《前分析篇》中建立了第一个形式系统。他的天才在于以字母充当占位符——「所有 A 是 B」——从而研究脱离内容的论证形式。这是词项逻辑：它处理「人」「会死」这类词项之间的关系。中世纪逻辑学家以助记名兴致勃勃地编录有效的三段论式——Barbara、Celarent、Darii。这些名字不是人名，而是密码：元音标记命题类型，A 表示「所有 S 都是 P」，E 表示「没有 S 是 P」，I 表示「有些 S 是 P」，O 表示「有些 S 不是 P」。因此 Barbara 是 AAA，Celarent 是 EAE，Darii 是 AII；例如 Barbara 意味着：所有 M 都是 P；所有 S 都是 M；所以所有 S 都是 P。近两千年间，这就是逻辑。

斯多葛学派，尤其是克律西波斯（约公元前 279–206 年），建立了第二条与之平行的逻辑，历史几乎让它失传。亚里士多德处理词项，斯多葛学派则用我们日常仍在使用的联结词处理整个命题：如果……那么、并且、或者、并非。克律西波斯列出五条「不可证明式」——基本推理图式，第一条（「若第一，则第二；但第一；故第二」）正是肯定前件式。这便是命题逻辑，也是每一块计算机芯片内部逻辑的远古源头。斯多葛学派很可能已经对联结词有了真值函数的理解——通过组成部分的真假判断整体的真假——这比后人重新发现早了两千年。20 世纪逻辑学家扬·武卡谢维奇曾令学者惊讶地主张，斯多葛逻辑并非亚里士多德的穷亲戚，而是「同等级的成就」。随后它被掩埋多年，亚里士多德独尊——这提醒我们，思想史并非一场整齐的接力赛。

乔治·布尔把两个传统推上了新轨道。1854 年，他在《思维规律的研究》中做了一件大胆的事：把逻辑推理当作计算。令 1 为全域，0 为空无；乘法即「且」，加法即「或」。骤然之间，有效推理的规律看上去如同代数定律。「我们不应再把逻辑与形而上学相联系，」布尔宣称，「而应把逻辑与数学相联系。」他的书销量平平，同代人也大惑不解。直到几十年后，1937 年克劳德·香农注意到布尔的二值代数精确描述了电路开关，布尔代数才成为数字逻辑名副其实的基础。你此刻用来阅读这段文字的设备中，每一个 AND 门都是克律西波斯的一句话在硅中的实现。

戈特洛布·弗雷格完成了自亚里士多德以来最大的跳跃。他那薄薄一卷、令人生畏的《概念文字》（Begriffsschrift, 1879）引入了量词——形式的「所有」（ $\forall$ ）与「存在」（ $\exists$ ）——以及谓词逻辑。亚里士多德的词项逻辑会被「马皆动物，故马头皆动物头」这类论证难倒；弗雷格的机制不仅能处理它，而且远不止此——它把命题解析为以个体为变元的函数。它常被称为符号逻辑史上最伟大

的一部著作。但悲剧性的尾声随之而来：弗雷格梦想把全部算术还原为纯粹逻辑，就在第二卷即将付梓之际，年轻的伯特兰·罗素寄来一封信，里面藏着一个悖论——所有不包含自身的集合构成的集合：它是否包含自身？无论回答「是」或「否」，都会自相矛盾。弗雷格宏大的基础工程由此崩裂。但他的逻辑在废墟中幸存，成为我们今天仍在讲授的现代符号逻辑。（那个悖论的幽灵，以及它所暗示的边界，将在第 28 日重新浮现；届时哥德尔将证明，没有任何形式系统能满足数学家曾怀有的全部希望。）

—— 辩论

## 逻辑是发现还是发明？

这里有一个听起来像沙龙游戏、实则直抵根本的问题。那些基岩般的定律——同一律（A 是 A）、矛盾律（A 与非-A 不能同真）、排中律（A 或非-A，没有第三种）——看似无可回避。但它们究竟栖居何处？是实在的特征，即使心智不存在也编织在宇宙之中？是思维的特征，任何思考者都无法逃避的语法？还是人类的约定——真实且具约束力，但终究是被选择出来的，犹如象棋规则？

### 逻辑实在论

被发现

定律是客观的、独立于心智的世界结构。我们并不立法规定矛盾律，正如我们并不立法规定素数——我们只是发现它。逻辑是从实在中读出的。

### 心理主义

思维规律

定律描述心智必须如何运作——实为心理学的一个分支。弗雷格与胡塞尔猛烈抨击这一点：逻辑真理是精确且先验的，而心理学是经验且模糊的。

### 约定主义

被发明

定律是我们因有用而采纳的约定——一旦选定便具约束力，但并非由宇宙降下。奇怪的是，尽管它与道德

### 可修正性

经验的？

奎因与普特南提出了激进的想法：即便逻辑也可能因经验理由而被修正——量子力学可能把我们推向非经

反实在论渊源甚深，这个立场却很少有充分发展的版本。

典逻辑，恰如相对论曾把我们推向非欧几何。

最后一个方框，正把问题引向今日的前沿。历史上大部分时间里，「思维规律」似乎不可触碰——质疑它们仿佛锯断自己正坐着的树枝。但二十世纪产生了严谨且可运作的替代逻辑，它们悄然放弃某条神圣定律，却依旧运转。一旦你看到这些替代逻辑确实能承担实际工作，那个宏大的形而上学问题便会软化成一个更实际、也更耐人寻味的问题：不是「哪种逻辑为真？」而是「对这项工作来说，哪种逻辑才是合适的工具？」下面就来看这些替代者。

—— 前沿 · 2026

## 三条活跃前沿，以及一道炒作过滤网

本课程每日都以前沿研究收尾，每条主张都标明了它能承受多少重量。逻辑的前沿出奇地具体：它运行在真实计算机上，核查真实证明，并且最近与人工智能发生碰撞——这要求我们擦亮眼睛。

前沿 01 [已确立]

### 故意打破规则的那些逻辑

「经典」逻辑并非唯一一致的选项；它只是更广阔的逻辑图景中一个已站稳脚跟的位置，每一套替代逻辑都放弃了大多数人以为不可动摇的某条定律。

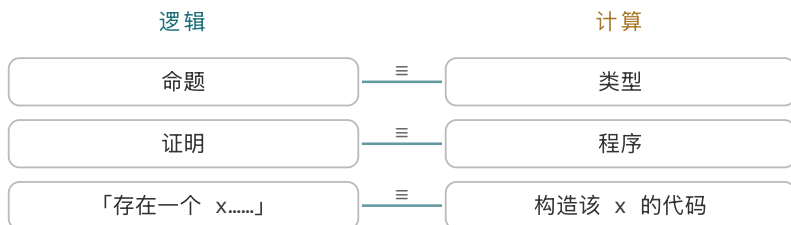
直觉主义逻辑放弃了排中律。它由 L. E. J. 布劳威尔开创，1920-30 年代由阿伦德·海廷形式化，坚持一条陈述只有在你能构造出其证明时才算为真。你不能凭空断言「A 或非-A」——你必须证明其中一边。一个例子能尖锐地说明动机：排中律会让你轻松断言，对任何计算机程序，「它停机或不停机」——然而（我们将在第 27 日看到）不存在判定停机的通用方法，因此没有构造支撑这一断言。直觉主义说：那就别断言。这听起来像是哲学洁癖，直到你发现经由一种美得值得专设一框的对应关系，它竟通向计算机科学的心脏。

次协调逻辑放弃了爆炸原则。在经典逻辑中，单个矛盾是灾难性的：从「P 且非-P」你可以推出任何东西（原则 *ex contradictione quodlibet*，「由矛盾可得任意结论」）——一处不一致，整座系统便付之一炬。次协调逻辑拒绝这一点，让你即便某些矛盾潜入，仍能合理推理——这对大型数据库、法典，以及任何局部不一致却仍有整体价值的信息集都很有用。更强硬的哲学表亲，**双面真理论**——格雷厄姆·普里斯特认为有些矛盾实际为真，如说谎者语句「这句话是假的」——则争议大得多。务必区分二者：你可以采纳次协调逻辑（关于爆炸原则的技术选择）而不成为双面真理论者（关于真实矛盾的本体论主张）。前者是工具；后者是世界观。

模糊逻辑完全放弃了二值限制。1965 年，卢菲特·扎德让真值在 0 到 1 的连续区间滑动，以刻画模糊性——「水是温的」为 0.7 真——建立在 1920 年代扬·武卡谢维奇的多值逻辑之上。它运行在控制系统与家电中。而模态逻辑——关于必然与可能（ $\square$  与  $\diamond$ ）的逻辑——以及经过精心选择的时态逻辑，支撑着硬件与软件的形式验证：某些具体片段既能表达有用性质，又足够克制，可以保留模型检查所需的可判定性。这些不是博物馆中的藏品，而是现代技术世界实际运转的逻辑。

#### 桥梁 · 命题即类型

直觉主义逻辑之所以重要的根本原因，是柯里-霍华德对应：在适当的形式系统中，命题对应类型，证明对应程序。证明一个定理，可以被看作构造一个居于相应类型之中的程序式对象——反过来也一样。



这就是为什么下面若干证明助手建立在类型论基础之上——也是为何逻辑与计算，我们五条线索之二，并非彼此相邻的邻域，而是同一片疆域的两面视图。（将在第 27-29 日继续展开。）

## 前沿 02 [已确立]

### 零容错的证明：证明助手的崛起

亚里士多德的梦想是一条紧密到无人能够怀疑的推理链。二十三个世纪后，这个梦想有了软件化身。证明助手是一种程序，其中每一步逻辑都必须通过机器核查；没有任何一步能凭权威、直觉或一句「显然」而蒙混过关。主流系统包括 Lean（现为 Lean 4）、Rocq（原名为 Coq，2025 年更名）、Agda 与 Isabelle/HOL。Lean、Rocq 与 Agda 属于类型论家族；Isabelle/HOL 则建立在经典高阶逻辑之上。目标相同，基础不同。

Lean 的社区共建数学库 `mathlib`，是世界上最大的统一数学形式化库之一：超过 278,000 条定理与 132,000 个定义——2026 年 6 月统计时如此，且仍在增长——覆盖了某著名「形式化这些」挑战清单上 100 个问题中的 84 个。这不是玩具。看看它已验证的成果：

#### 2022 · 已完成

液体张量实验。2020 年 12 月，菲尔兹奖得主彼得·朔尔策向全世界发出挑战，要求验证其「凝聚数学」中一个他本人都不太确定的定理。约翰·科默兰与亚当·托帕兹带领的团队在 Lean 中完成了验证，于 2022 年 7 月 14 日完成。一位一线数学家借助机器，获得对一份复杂到人类审稿人难以安心核查的 proof 的信心——这正是关键所在。

#### 2023 · 三周内完成

多项式弗雷曼-鲁萨猜想。蒂姆·高尔斯、本·格林、弗雷迪·曼纳斯与陶哲轩发表这一加性组合学结果的证明数日后，陶哲轩启动了一个 Lean 项目来形式化它——三周后便宣布依赖图「被一片可爱的绿色完全覆盖」。形式化几乎与研究同步推进。

#### 2024-25 · 已完成

等式理论项目。陶哲轩的合作实验（2024 年 9 月启动）旨在判定 4,694 条代数定律之间的蕴涵关系——若把每条定律对自身的平凡蕴涵也算入，共有 22,033,636 个有序对；若只数非平凡图边，则为 22,028,942 条——

结合人类证明、自动证明器、AI 与 Lean 验证，50 余位贡献者，在 200 多天內完成了工作。这是一种大规模协作、机器核查数学的新范式。

2024-2029 · 进行中

费马大定理。凯文·巴扎德由 EPSRC 资助的项目（2024 年 4 月启动，伦敦帝国理工学院）旨在形式化 FLT——并非怀尔斯的原始证明，而是一条现代路线。巴扎德「谨慎乐观」地认为自己能把它归约到 1980 年代已知的结果，但坦率承认整个项目「至少需要 5 年」。尚未完成——最诚实也准确的说法是：它仍在进行中，也是那 100 个挑战问题中尚未闭合的最后一项。

这种确定性已从纯数学延伸到生命所依赖的系统。Rocq 中被证明正确的 C 编译器 CompCert；一项著名的编译器查错研究耗费约六个 CPU 年，试图诱使它生成错误代码，却一无所获——「我们测试过的唯一一个 Csmith 无法找到错误代码的编译器」——同时在 GCC 与 LLVM 中找出了大量 bug。seL4 是第一个在 Isabelle/HOL 中拥有完整机器检查功能性正确性证明的操作系统微内核：在其明确列出的假设下，C 实现细化了形式规格，因此整类崩溃与不安全行为不是靠希望避免，而是被定理排除。这些不是普通承诺，而是关于软件的有条件定理。这就是逻辑的机械化所能做的事——而且它已确立。

前沿 03 [已确立] [争议/炒作]

## 当 AI 遇见证明核查器

最新、也最被喧嚣包围的前沿，是机器学习与形式证明的碰撞——此处正是前沿校准器该上场的时候，因为标题与事实之间已经出现了漂移。

先看真正的里程碑。2024 年 7 月，DeepMind 的 AlphaProof 与 AlphaGeometry 2 联手，在国际数学奥林匹克（IMO）6 道题中解出 4 道，获得 28 分——位居银牌档顶端，仅比 29 分的金牌线低 1 分。它甚至攻克了令人畏惧的第 6 题，这道题在约 600 名人类参赛者中只有 5 人完整解出。该方法于 2025 年 11 月 12 日在线发表于 Nature，正式版本于 2026 年刊出。真正把它同

聊天机器人式空谈区分开的关键设计事实是：AlphaProof 在 Lean 内部工作。它把约一百万道自然语言问题自动形式化为约 8000 万条 Lean 陈述，然后以 AlphaZero 风格的循环训练自己，其中每一步都由 Lean 核查。用 DeepMind 的话说，「无需担心幻觉」——因为一个幻觉步骤根本无法编译。神经网络提供创造性搜索，证明助手提供真值基准。这种结合真实且重要。[已确立]

2025 年 7 月，门槛再次抬高：DeepMind（Gemini「Deep Think」模型）与 OpenAI 都报告了金牌分数——6 题中解出 5 题，35 分——而且引人注目的是，它们在时限内以端到端自然语言完成，而非在 Lean 内部完成。DeepMind 的结果由 IMO 官方认证；OpenAI 的结果是内部评分。确实令人印象深刻。但也正是在这里，第 1 日练出的校准直觉该派上用场：

- 「金牌」是一个分数，不是加冕。这些是竞赛题——数学中狭窄、限时、已知存在简短答案的一角。它们不是开放的研究问题，而且据官方 2025 年结果，仍有 26 名人类参赛者得分超过两个 AI 系统。
- 离开 Lean 是一种取舍，不是无代价的升级。2024 年的银牌是形式验证的——由机器保证正确。2025 年的自然语言金牌是人工评分的，意味着我们重新依赖可能藏有细微漏洞的散文。更通用，却更不确定。别让「金牌胜过银牌」的叙事掩盖了认识论根基的转移。
- 它昂贵且狭窄。每道困难的 2024 年题需要两到三天的计算，而且题目还需先被人工形式化为 Lean 陈述。这还称不上通用数学智能。

最需要明确否定的说法是：AI 尚未「解决数学」，也没有使数学家变得多余。[争议/炒作] 没有任何 AI 独立证明过一个著名的开放猜想并被接受为里程碑。关于定理证明代理找到小型 Lean 证明、或帮助完成狭窄形式化任务的报道虽然有趣，但仍早期、范围有限，也还不能替代被数学共同体接受的研究数学；它们应归入[线索]，留待日后审视，而非大肆宣扬。真正的革命比标题更安静、也更持久：一条延续 2300 年的标准——证明是一条无人能怀疑的链——终于交由机器以零容错执行，而 AI 正在学习沿这些严苛轨道搜索。（我们将在第 138-145 日深入追寻这一主题。）

## 关于虚构来源的注记

本课程的前沿校准器有一条必须明说的规则：凡是指向未来日期预印本编号的引用，一律剔除。这一领域的搜索结果中，充斥着信誓旦旦引用尚不存在论文的条目。以上每个里程碑都可追溯到真实、有日期、可核实的原始来源——已发表的 Nature 论文、官方竞赛结果、具名研究者本人的公开宣布。当一条关于 AI 与数学的声明无法这样追溯时，正确反应不是兴奋，而是怀疑。

## —— 开放问题

# 真正尚未解决的是什么

二十三个世纪之后，有效推理的研究依然留有真正未决的问题：

- 是否只有一种真逻辑，还是有许多种？当直觉主义逻辑、次协调逻辑与模糊逻辑都能切实派上用场，「正确逻辑」便渐渐显得更像工具选择，而非宇宙事实——但多元论者与一元论者仍真正地各执一词。
- 发现还是发明？逻辑定律是从实在中读出、嵌入任何可能心智，还是由约定采纳？经验物理学能否如普特南所想迫使我们修正？
- 溯因究竟是什么？「最佳解释推理」是真正第三种模式，还是换了外衣的归纳？甚至皮尔士本人是否将其理解为最佳解释推理（而非仅仅生成假设），学者之间亦有争议。
- 机械化证明能否改变数学本身？若一个结果为真，却只有计算机核查过证明，有没有人真正理解它？一个已验证却不透明的证明，与一个能带来洞见的人类证明，价值是否相同？
- 以及将萦绕 AI 单元的问题：当一台机器输出一个真实且得到充分支持的定理时，它是否知道任何东西——还是它只是第 1 日那个终极盖梯尔案例的翻版——因与理解毫无关系的理由而恰巧正确？（第 138-145 日。）

◆ 用三句话概括今日

核心观点

推理有三种引擎、三种担保——演绎保真却不扩展内容，归纳概括却可能被下一例打破，溯因跃向最佳解释——而在演绎内部，有效性（形式成立）与健全性（形式成立且前提为真）是完全不同的两件事。

最佳类比

夏洛克·福尔摩斯的「演绎」其实是溯因——对线索的最佳解释，而非保证结论——而一个有效却不健全的论证，是一条接合严密却输送污水的管道。

当下争议

逻辑是发现还是发明（以及是否只有一种真逻辑，还是一套工具），如今被一条真实的前沿所激化：Lean 等证明助手以零容错验证前沿数学，AI 已达奖牌水准——但并未真正「解决数学」。

---

今日线索 › 计算（柯里-霍华德：证明对应程序；硅芯片中的布尔代数；证明助手）· 信息（形式化使证明内容可被机器核查）· 涌现（大规模协作证明判定约 2200 万个蕴涵关系）——也将演绎与归纳衔接到[第 1 日](#)与[第 2 日](#)。

明日 → 第 04 日

## 概率成为扩展的逻辑

今日负责扩展却可能失手的引擎是归纳，而溯因留给我们一个任务：判断哪种解释最佳。明日，我们为二者加上数字刻度。概率原来并非与逻辑分离的学科，而是部分信念的自然延伸——蒙提霍尔问题将展示我们的直觉能错得多离谱，而贝叶斯定理又如何纠正它们。带上今日对天衣无缝与只是看似合理的区分：你即将学习如何演算「看似合理」。

---

—— 来源

## 来源与延伸阅读

1. "Validity and Soundness." Internet Encyclopedia of Philosophy (accessed 2026). [iep.utm.edu/val-snd](http://iep.utm.edu/val-snd) --基于形式的有效性与健康性区分。
2. "Deductive and Inductive Arguments." Internet Encyclopedia of Philosophy. [iep.utm.edu/ded-ind](http://iep.utm.edu/ded-ind) --保真推理与扩展性推理之分。
3. Douven, I. "Abduction." Stanford Encyclopedia of Philosophy (rev. 2021). [plato.stanford.edu/entries/abduction](http://plato.stanford.edu/entries/abduction) --皮尔士、最佳解释推理，以及关于溯因究竟是什么的学术争论。
4. "Aristotle's Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/aristotle-logic](http://plato.stanford.edu/entries/aristotle-logic) --三段论、《前分析篇》与词项逻辑。
5. Bobzien, S. "Ancient Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/logic-ancient](http://plato.stanford.edu/entries/logic-ancient) --克律西波斯、斯多葛不可证明式与命题逻辑；武卡谢维奇的重新评估。
6. Boole, G. (1854). An Investigation of the Laws of Thought. London: Walton & Maberly. See "George Boole, The Laws of Thought," PhilPapers. [philpapers.org/rec/BOOTLO-4](http://philpapers.org/rec/BOOTLO-4) --逻辑作为代数；「逻辑与数学」。

7. "Origins of Boolean Algebra in the Logic of Classes." Mathematical Association of America (Convergence). [old.maa.org](http://old.maa.org) --布尔、文恩、皮尔士，以及经香农（1937）通往数字逻辑之路。
8. "Frege's Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/frege-logic](http://plato.stanford.edu/entries/frege-logic) --《概念文字》（1879）、量词、谓词逻辑与罗素悖论。
9. "Intuitionistic Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/logic-intuitionistic](http://plato.stanford.edu/entries/logic-intuitionistic) --布劳威尔、海廷、对排中律的拒斥、BHK 解释。
10. Priest, G., Berto, F. & Weber, Z. "Dialetheism" and "Paraconsistent Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/dialetheism](http://plato.stanford.edu/entries/dialetheism) --爆炸原则、次协调性与双面真理论、Logic of Paradox。
11. "Fuzzy logic." Wikipedia (accessed 2026). [en.wikipedia.org/wiki/Fuzzy\\_logic](http://en.wikipedia.org/wiki/Fuzzy_logic) --扎德（1965）、 $[0,1]$  上的真值、多值 / 武卡谢维奇根源。
12. Garson, J. "Modal Logic." Stanford Encyclopedia of Philosophy. [plato.stanford.edu/entries/logic-modal](http://plato.stanford.edu/entries/logic-modal) --必然 / 可能与计算机科学及验证应用。
13. "Curry-Howard correspondence." Wikipedia (accessed 2026). [en.wikipedia.org/wiki/Curry-Howard\\_correspondence](http://en.wikipedia.org/wiki/Curry-Howard_correspondence) --命题即类型，证明即程序。
14. "Mathlib statistics." Lean community (accessed June 2026). [leanprover-community.github.io/mathlib\\_stats.html](http://leanprover-community.github.io/mathlib_stats.html) --当前定理与定义数量。
15. "100 theorems in Lean." Lean community (accessed June 2026). [leanprover-community.github.io/100.html](http://leanprover-community.github.io/100.html) --Wiedijk 的 100 个定理基准中已有 84 个在 Lean 中形式化。
16. Commelin, J. & Topaz, A. et al. "Liquid Tensor Experiment." Lean community blog (completion 14 July 2022); Scholze's original challenge (Dec 2020). [leanprover-community.github.io](http://leanprover-community.github.io) --机器核查一位菲尔兹奖得主自己都不太确定的证明。
17. Tao, T. "Formalizing the proof of PFR in Lean4." [terrytao.wordpress.com](http://terrytao.wordpress.com) (Nov 2023). Gowers, Green, Manners & Tao, "On a conjecture of Marton," *Annals of Mathematics* (2025). [terrytao.wordpress.com](http://terrytao.wordpress.com)
18. Tao, T. et al. "The Equational Theories Project." Project announced Sept 2024; retrospective paper Dec 2025 (arXiv:2512.07087). [teorth.github.io/equational\\_theories](http://teorth.github.io/equational_theories) --22,033,636 个含自蕴涵的有序对；22,028,942 条非平凡图边；50 余位贡献者，Lean 验证。

19. Buzzard, K. "Fermat's Last Theorem project." Lean community blog (launch 30 April 2024); EPSRC grant EP/Y022904/1 (2024–2029), Imperial College London. [leanprover-community.github.io](https://leanprover-community.github.io) --进行中; 「至少需要 5 年」。
20. Leroy, X. et al. "CompCert" – a formally verified C compiler. Yang, Chen, Eide & Regehr, "Finding and Understanding Bugs in C Compilers," PLDI (2011). [compcert.org](https://compcert.org) --约六个 CPU 年未找到错误代码。
21. Klein, G. et al. (2009). "seL4: Formal Verification of an OS Kernel." SOSP '09. [sel4.systems](https://sel4.systems) --首个操作系统微内核功能性正确性的机器检查证明 (Isabelle/HOL)。
22. "AI achieves silver-medal standard solving International Mathematical Olympiad problems." Google DeepMind blog (25 July 2024). [deepmind.google](https://deepmind.google) --AlphaProof + AlphaGeometry 2; 28 分; 在 Lean 中工作。
23. Hubert, T., Mehta, R., Sartran, L. et al. (2026). "Olympiad-level formal mathematical reasoning with reinforcement learning." Nature 651: 607–613. doi:10.1038/s41586-025-09833-y. [nature.com/articles/s41586-025-09833-y](https://nature.com/articles/s41586-025-09833-y) --AlphaProof 方法论文; 2025 年 11 月 12 日在线发表, 2026 年 3 月 13 日正式出版; 约 8000 万道 Lean 问题。
24. "Advanced version of Gemini with Deep Think officially achieves gold-medal standard at the IMO." Google DeepMind blog (July 2025). [deepmind.google](https://deepmind.google) --35/42, 官方认证; 时限内的自然语言证明。
25. "66th IMO 2025." International Mathematical Olympiad. [imo-official.org/editions/2025](https://imo-official.org/editions/2025) 和 [individual results](https://imo-official.org/editions/2025/individual-results) --630 名参赛者; 金牌线 35; 人类分数分布。
26. "Our First Proof submissions." OpenAI (2026). [openai.com/index/first-proof-submissions](https://openai.com/index/first-proof-submissions) --OpenAI 对其 2025 年 7 月 IMO 金牌级结果的后续总结, 35/42 分。
27. "Philosophy of logic" & "Logical realism." Wikipedia / Stanford Encyclopedia of Philosophy (accessed 2026). [plato.stanford.edu/entries/logical-pluralism](https://plato.stanford.edu/entries/logical-pluralism) --实在论、约定主义、奎因 / 普特南关于修正逻辑、逻辑多元论。

第 03 日完 · 尚有 177 次深入