课程笔记

萨顿谈 AI 的苦涩教训：从历史规律到基础模型的未来

cnfjlhj & Codex

2026 年 7 月 20 日

演讲者：Richard S. Sutton（强化学习奠基人）
视频作者/频道：居鲁仕（Bilibili UP 主）
发布日期：2026 年 7 月（WAIC 2026）
视频时长：约 13 分 42 秒
视频链接：`https://www.bilibili.com/video/BV1VHKA67Eit`

目录

1 引言：谁在讲，讲了什么

2019 年，Richard S. Sutton 发表了一篇仅有两页纸的短文《**The Bitter Lesson**》（苦涩的教训）。这篇短文在 AI 社区引发了巨大反响，成为被反复引用的”微型经典”。Sutton 是强化学习领域的奠基人之一，他的 TD-learning 算法、策略梯度方法、以及与 Barto 合著的经典教材《Reinforcement Learning: An Introduction》深刻塑造了整个领域。

在 WAIC 2026 的这场演讲中，Sutton 亲自回顾了《苦涩的教训》的核心论点，并将其延伸到一个当下最紧迫的问题：**这条历史规律是否同样适用于大语言模型和基础模型？**

背景知识：Richard Sutton 与强化学习

Richard S. Sutton 是加拿大计算机科学家，阿尔伯塔大学教授，DeepMind 杰出研究科学家。他被誉为”强化学习之父”，其主要贡献包括：

- Temporal Difference (TD) Learning —强化学习的核心算法框架
- Policy Gradient Theorem —策略梯度方法的数学基础
- Dyna 架构—将规划与学习统一的框架

他的研究哲学一贯强调：**通用方法 + 算力扩展**胜过 **人类专家知识 + 精巧设计**，这一哲学正是《苦涩的教训》的核心。

本讲核心问题

大语言模型和基础模型是否也会被”苦涩的教训”所支配？即：它们目前依赖人类知识的路线，长期来看是否会被基于算力扩展和搜索学习的方法所取代？

1.1 本章小结

本节介绍了 Sutton 的学术地位和《苦涩的教训》的背景，明确了本场演讲的核心问题——将历史规律投射到当下的基础模型之上。带着这个问题，我们进入 Sutton 对苦涩教训原文的逐段解读。

2 苦涩的教训：原文与解读

2.1 历史观察：研究者反复犯的同一个错

Sutton 首先回顾了《苦涩的教训》的发表历程：最初于 2018 年以演讲形式呈现，”让在场的不少科学家感到不快”（severely upset some of the scientists present），随后在 2019 年正式发表并在 Twitter 上广泛传播，成为 AI 圈的一种”微型传奇”（mini legend）。

文章的核心论点基于一个反复出现的历史模式：

苦涩的教训——核心论断（Sutton 原文释义）

历史观察：AI 研究者反复尝试将人类知识构建到智能体中。
短期效果：这种做法在短期内总是有效的，也让研究者本人感到满意。
长期效果：从长期来看，这种做法会**停滞**（plateaus）甚至**抑制**进一步进展。
突破的来源：真正的突破性进展最终总是通过**对立的方法**实现——基于扩展计算量的搜索与学习。

Sutton 特别强调：这就是为什么这个教训是” 苦涩” 的——因为研究者自己偏好的方法虽然一度有效，但最终被其他方法所**取代**（superseded）。研究者投入了大量心血的领域知识和精巧设计，在算力扩展的浪潮面前不堪一击。

2.2 四个历史案例

Sutton 在演讲中列举了四个关键的历史案例来支撑这一论点：

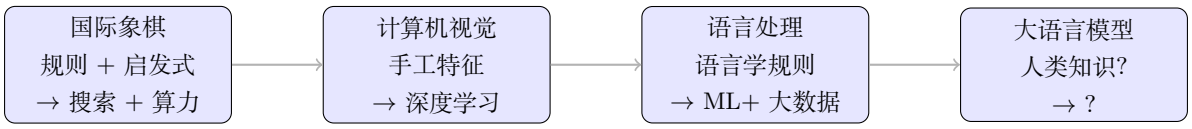

图 1: 苦涩的教训：四个领域中” 人类知识 → 算力扩展” 的反复模式。前三个已成历史定论，第四个正是 Sutton 本场演讲要探讨的问题。

1. **计算机国际象棋**：早期方法依赖人类棋手的知识——开局库、残局库、评估函数中的启发式规则。最终被纯粹的搜索 + 算力方法（如 AlphaZero 的自我对弈）所取代。
2. **计算机视觉**：从手工设计特征（SIFT、HOG、Haar-like features）到端到端深度学习，卷积神经网络通过大规模数据 + 算力直接从像素中学习表示。
3. **语言处理**：最早的 NLP 方法基于语言学规则和专家知识，被基于统计的方法（n-gram、隐马尔可夫模型）所取代，而统计方法最终又被大规模机器学习所取代。
4. **大语言模型**：这是 Sutton 提出的新问题——当前的大语言模型是否也在重复同样的模式?

常见误解：这不是物理定律

Sutton 明确指出：苦涩的教训是一条**历史观察**（historical lesson），**不是物理或科学定律**（not a law of physics or science）。它不**必然**再次发生。然而，在过去 70 年的 AI 研究中，这一模式已经重复了太多次，我们不得不对它保持警惕。

2.3 二十六字的浓缩版

在演讲中，Sutton 提到他在 X（Twitter）上用一条推文将《苦涩的教训》浓缩为 **26 个词**：

苦涩的教训——26 词精简版

“Don’t be distracted by human knowledge as AI has been historically. Instead focus on methods that will scale with computation, methods that create knowledge, not try to take advantage of existing human knowledge so that you can scale with search and learning.”
不要像 AI 历史上的做法那样被人类知识分散注意力。相反，应专注于能随算力扩展的方法——那些**创造**知识的方法，而非试图利用现有人类知识的方法，这样你才能随搜索与学习而扩展。

这段话值得逐句拆解：

- **“Don’t be distracted by human knowledge”** —不要被人类知识**分心**。Sutton 用的词是”distracted”，暗示人类知识看似有用，实际上是一种**注意力陷阱**——它让你偏离了真正能带来长期突破的方向。

- **"methods that will scale with computation"** —选择能随算力**扩展**的方法。关键词是"scale":当算力增长 10 倍、100 倍时，方法本身应该自动变强，而不需要人为重新设计。
- **"methods that create knowledge"** vs **"take advantage of existing human knowledge"** —这是全段最深刻的对比：好的方法**创造**新知识，坏的方法**利用**已有的人类知识。前者是开放式探索，后者是封闭式编码。

2.4 本章小结

Sutton 的核心论点可以压缩为一句话：**在 AI 研究中，利用人类知识的做法短期有效但长期被取代，而基于算力扩展的搜索与学习才是真正带来突破的路线**。这一模式在象棋、视觉、语言处理中反复上演。接下来，Sutton 将这条历史规律直接投射到当下最热的基础模型之上。

3 基础模型：苦涩的教训适用吗？

3.1 Sutton 的判断：既是也不是

这是本场演讲最关键的问题，也是 Sutton 最直接地回应时代关切的部分。他的回答是：**既适用，也不适用**（briefly yes and no）。

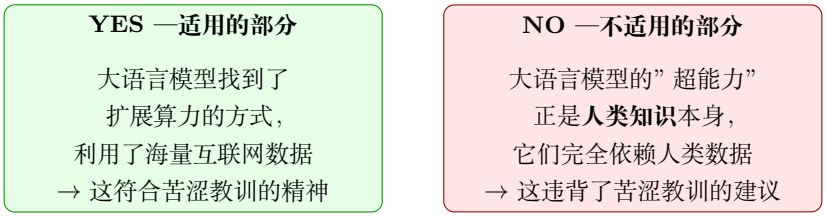

图 2: Sutton 对" 苦涩的教训是否适用于基础模型" 的二分回答。

适用的部分（YES）：大语言模型确实找到了一种大幅扩展算力的方式。它们利用了当今可用的大规模计算资源和互联网上的海量数据集。可以说，将大语言模型带到今天这一步，本身就体现了对" 苦涩的教训" 的认知——找到了一种可扩展的方法。

不适用的部分（NO）：同样明显的是，大语言模型和基础模型**建立在人类知识之上**。它们的" 超能力"（superpower）就是人类知识、人类数据。而这恰恰是苦涩的教训所警告的——不要依赖人类知识。

3.2 Sutton 的预判：被取代是大概率事件

Sutton 在演讲中给出了明确的立场：

Sutton 的核心预判

"我认为**很可能**（likely），它们目前正受到制约，未来也会受到制约，而且我怀疑**在长期它们终将被取代**——正如历史上反复发生的那样——因为它们完全依赖现有人类知识，而非试图创造或超越人类知识。"

这是一个相当大胆的判断。要知道，发表这番话的场合是 WAIC 2026——在大语言模型和基础模型如日中天的时代。Sutton 的逻辑链条如下：

1. 苦涩的教训说：依赖人类知识的方法，长期被基于算力扩展的搜索/学习所取代。
2. 大语言模型的全部力量来自人类知识（互联网文本数据）。
3. 因此，大语言模型在本质上属于”被苦涩的教训警告”的那一类方法。
4. 因此，大语言模型在长期很可能被取代。

需要辨析的关键张力

Sutton 的判断有一个微妙之处：大语言模型**确实**利用了算力扩展（这是符合苦涩教训的部分），但它们扩展的是**对人类知识的压缩和检索**，而非**对新知识的创造**。苦涩的教训推崇的不仅是”扩展”，更是”通过搜索和学习**创造**知识”。大语言模型在这一点上存在根本性缺失。

3.3 本章小结

Sutton 的回答是”既是也不是”：大语言模型在算力扩展方面符合苦涩教训的精神，但其核心依赖人类知识这一事实，使其很可能在长期被取代。这一判断将我们引向一个更根本的问题：基础模型**到底是什么**？只有回答了这个本体论问题，才能判断它们的终局。

4 大世界视角：为什么全知是不可能的

4.1 基础模型到底是什么？

为了回答基础模型在极限情况下的命运，Sutton 认为我们必须先回答一个更根本的问题：**基础模型到底是什么？**

他指出，”基础模型”（foundation model）已经变成了一个被滥用的流行语（buzzword）——人们在使用这个词时，往往并不清楚自己到底在指什么。而要判断基础模型的未来，必须先搞清楚它的本质。

Sutton 给出了他自己的定义：

Sutton 对基础模型的定义

基础模型是所有已确立的人类知识的容器（containers of all agreed human knowledge）。
它们**不是**像科学家那样会创造新知识的实体，它们渴望成为的是**包罗万象的知识容器**——存储所有已被人类确立、达成共识的知识。

注意这个定义中的关键限定词：**”已确立的、达成共识的”**（agreed, established）。Sutton 随即指出，如果基础模型的目标是包含**所有**知识——包括未知的、不确定的、尚未发现的知识——那是不可能的。但如果是包含**已确立的、达成共识的**人类知识，那是可能的。

4.2 大世界视角：世界远大于任何心智

接下来，Sutton 展开了本场演讲中最具哲学深度的部分——**大世界视角**（Big World Perspective）。他用几页幻灯片来阐释这个核心思想。

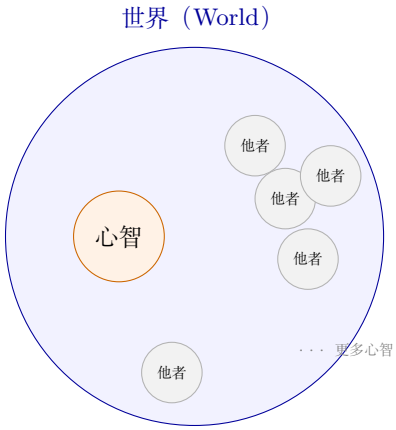

图 3: 大世界视角示意图：世界远大于任何单一心智。世界不仅包含物理环境，还包含**许多其他心智**——每个心智都有自己独特的知识和体验。

Sutton 首先纠正了一个直觉性的错误图景：人们可能以为世界和心智是”大致可比的”（roughly comparable），就像一个画面中同时画着世界和一个代表心智的小人，两者尺寸相仿。但正确的图景应该是：**世界比任何心智都大得多、复杂得多**。

为什么？因为心智只是一**个**心智，而世界包含了**许多其他心智**。世界不仅由物理环境构成，更由无数其他有自己独特知识和体验的智能体构成。

4.3 一个直觉化的例子：你无法完全了解另一个人

Sutton 用一个非常贴近生活的例子来让这个抽象观点落地：

Sutton 的直觉例子

“即使我刚到这里，我对在座的各位也已有了一个心智模型——因为你们的心理状态对我很重要，我正在试图把一些想法从我的脑中传递给你们。然而，我**显然不可能**拥有关于你们心理状态的完整知识。这只是一个基本事实的一个具体实例——世界比任何心智都复杂得多。”

这个例子的精妙之处在于：它把一个看似抽象的哲学命题变成了每个人都能感同身受的日常经验。你永远无法完全”读取”另一个人的内心——这不是技术限制，而是**本体论事实**。

4.4 推论：广度必然意味着浅度

从”世界远大于任何心智”这一前提出发，Sutton 推导出一个关键推论：

广度与深度的不可兼得

任何基础模型（或任何心智）只有在接受浅薄性（superficiality）的前提下，才能声称拥有通用性（generality）。
它不可能同时声称对世界拥有**深入而完整**的知识。这**不应该**是我们的目标。

换句话说，通用性和深度是**此消彼长**的。一个试图覆盖一切知识的系统，注定只能停留在表面；而一个在某个领域拥有深度知识的系统，注定无法覆盖一切。这不是工程问题，而是信息论和认识论层面的根本限制。

4.5 本章小结

大世界视角揭示了一个根本性的事实：世界比任何心智都大得多、复杂得多，因此没有任何心智能包含对世界的完整知识。基础模型可以追求成为”已确立人类知识的容器”，但不能追求”全知”。通用性必须以浅薄性为代价。这一认识将我们引向 Sutton 对基础模型未来角色的定位。

5 美丽的真理：谦逊的力量

5.1 为什么叫”美丽的真理”？

Sutton 将大世界视角称为一个**美丽的真理**（beautiful truth）。这乍听矛盾——”你不可能全知”怎么可能是”美丽的”？Sutton 从两个角度阐释了它的美：

1. **它让我们保持谦逊**：承认我们的局限性，帮助我们变得谦卑。我们不会幻想自己能理解一切。
2. **它解除了”超级心智”的压迫**：如果没有人能理解一切，那么我们就不必被”会比你更了解你自己、比你更了解你朋友”的超级心智的想法所压迫。

超级心智同样受限

Sutton 明确指出：超级心智**并非全能**。它受到与你我**定性上相同**（qualitatively the same）的限制。我们每个人都将是自己生活、自己世界角落的**专家**——超级心智也不例外。

这是一个深刻的观点：大世界视角的限制不是人类专属的弱点，而是**任何有限心智**的本质属性。无论你有多少参数、多少算力，只要你是一个**有限**的心智，你就无法包含无限复杂的世界。超级心智也会有自己擅长和不擅长的领域。

5.2 对错误野心的纠正

然而，”美丽的真理”对于”错误的野心”（false ambition）来说是个问题。什么错误的野心？就是基础模型试图成为**全知**的野心——试图涵盖所有知识、回答所有问题的野心。

需要丢弃的错误希望

Sutton 要求我们**丢弃**（discard）全知的错误希望，**接受**（accept）基础模型一个**更小但仍有意义**的角色。这不是放弃，而是对基础模型的**重新定位**——从一个不切实际的目标（全知）转向一个可实现的、有价值的目标。

5.3 本章小结

大世界视角是一个”美丽的真理”：它让我们谦逊，也解除了超级心智压迫的恐惧。但同时也要求我们放弃基础模型全知的错误野心，转而接受一个更现实、更有意义的定位。这就引出了 Sutton 对 AI 未来格局的构想。

6 AI 的未来格局：三类心智

6.1 不是一种 AI，而是多种

Sutton 认为，未来不会只有一种 AI，而是会有**多种不同类型**的 AI 共存。他将这些 AI 分为三类，形成一个从” 完整心智” 到” 专用工具” 的光谱，基础模型位于中间。

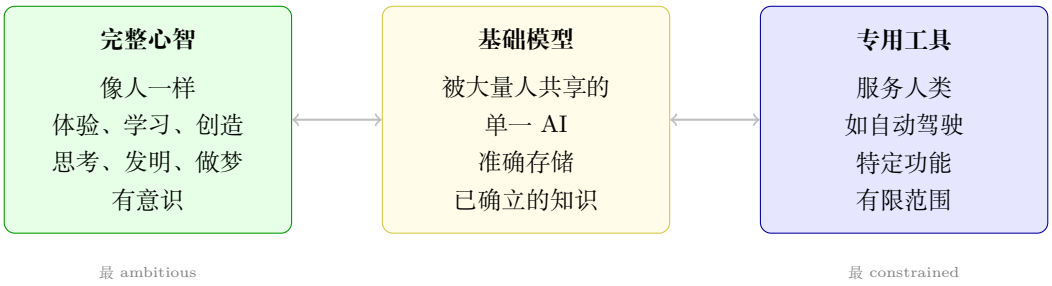

图 4: Sutton 构想的 AI 未来格局：三类心智/智能体在功能光谱上的分布。

6.2 第一类：完整心智——像人一样的 AI

Sutton 首先描述了他认为最具野心的 AI 类型：

完整心智：Sutton 的终极愿景

Sutton 的个人野心是创造**像人一样的完整心智**——能够体验、学习、创造、思考、发明、做梦，甚至具有意识的 AI。他相信这**终将实现**。
但关键判断是：**这些不会是基础模型**。完整心智会像人一样，在其生命周期中持续学习，持续对自己接触到的世界角落变得专业化。

注意几个要点：

- **“这些不会是基础模型”** —Sutton 明确将完整心智和基础模型区分为不同的东西。基础模型不是通往完整心智的路径，而是另一种类型的 AI。
- **“它们也不会是超级的”** —完整心智不会全知。它们像人一样，有自己的专业领域和局限。
- **“持续学习”** —完整心智的关键特征是**终身学习**（continual learning），在生命周期中不断适应。这与当前大语言模型” 训练一次、固化使用” 的模式根本不同。

6.3 第二类：专用工具——服务人类的简单 AI

光谱的另一端是专用工具型 AI：

- 例如**自动驾驶汽车**：做特定的事来辅助人类。
- 它们不一定复杂。
- 它们有**指定的范围**（designated scope）——在明确的边界内工作。

这类 AI 不需要通用性，只需要在指定领域内可靠地工作。

6.4 第三类：基础模型——准确的知识存储

位于光谱中间的，是基础模型。Sutton 对基础模型的重新定位如下：

基础模型的正确定位

基础模型是**被大量人共享的单一 AI**，其角色是成为**准确的、关于当前已达成共识的知识存储**（accurate stores of current agreed-on knowledge）。
这是比当前基础模型**更高的**野心——因为当前的模型经常**幻觉**（hallucinating），被要求做出超出其数据范围的事。我们应该接受” 成为已确立知识的准确代表” 这一野心。

Sutton 对当前基础模型的批评非常精准：

1. **幻觉问题**：当前模型经常生成虚假信息。
2. **越界问题**：我们经常要求模型做出超出其数据范围的事——这是**问题**，不应该是目标。
3. **正确的目标**：成为” 已确立知识的准确代表”——这本身就是一个**伟大且有意义**的目标。

6.5 超级百科全书：基础模型的理想形态

Sutton 用一个生动的比喻来描述基础模型的理想形态：**超级百科全书**（super encyclopedias）。

基础模型

基础模型的角色应该是：

- 让已确立的知识**易于获取**
- 以**定制化**的方式呈现——适应不同理解水平的用户
- 成为一种” 超级百科全书”

Sutton 认为，**如果我们能真正实现这一点，这将是一项伟大、有意义且重要的成就**。

这个定位的核心转变是：从” 试图回答一切问题” 到” 准确呈现已确立的知识”。前者是一个不可能的目标（因为大世界视角告诉我们全知是不可能的），后者是一个可实现的、有价值的目标。

6.6 本章小结

Sutton 构想了 AI 未来的三类格局：完整心智（像人一样终身学习）、专用工具（指定范围内服务人类）、基础模型（准确的知识存储/超级百科全书）。基础模型不是通往完整心智的路径，而是一种独立的、有价值的 AI 类型——前提是放弃全知幻想，接受” 准确存储已确立知识” 的定位。

7 总结与延伸

7.1 Sutton 的结语

Sutton 以简洁有力的话语结束了演讲：

Sutton 的结语

"我认为这就是基础模型的未来。我只是希望我们专注于它，拥抱这个未来——**作为当前知识准确存储的 AI**，我认为这就是未来。谢谢大家。"

7.2 核心论点回顾

回顾整场演讲，Sutton 的论证可以提炼为以下逻辑链：

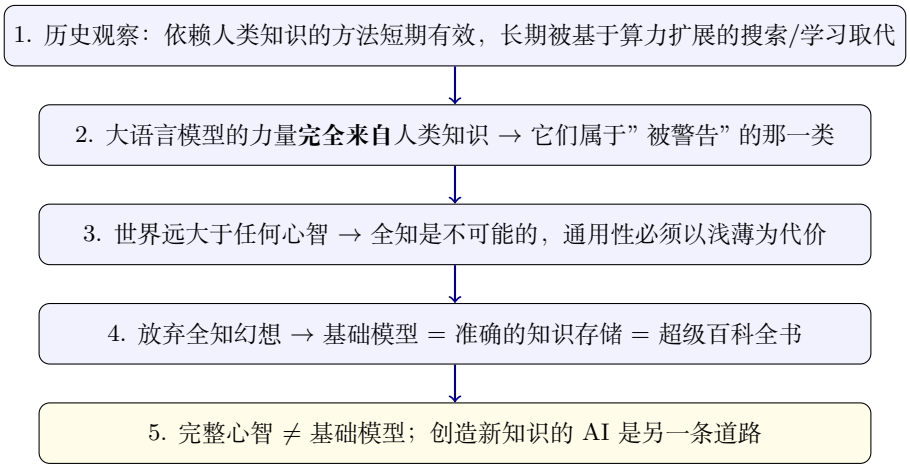

图 5: Sutton 演讲的完整论证逻辑链。

7.3 深度综合：三个层面的启示

7.3.1 认识论层面：知识的两种范式

Sutton 暗中区分了两种根本不同的知识范式：

1. **知识编码范式**（Knowledge Encoding）：将人类已有知识编码进系统。这是" 利用已有知识" 的路线。代表：专家系统、手工特征、以及当前的大语言模型（将互联网上的人类文本编码进参数）。

2. **知识创造范式**（Knowledge Creation）：通过搜索和学习让系统**自己发现**新知识。这是" 创造知识" 的路线。代表：AlphaZero（自我对弈发现棋艺）、强化学习智能体（通过环境交互发现策略）。

苦涩的教训本质上是说：知识编码范式短期有效但长期被知识创造范式取代。Sutton 对大语言模型的判断正是基于这一区分——大语言模型是知识编码范式的巅峰代表，因此长期可能被取代。

7.3.2 技术层面：持续学习 vs 固化训练

Sutton 提到完整心智会" 在其生命周期中持续学习"（continue to learn during their lives）。这一点触及了当前大语言模型的一个根本局限：

- 当前大语言模型是**训练一次、固化使用**的模式——训练完成后，模型参数不再变化，所有" 知识" 都被压缩在固定的权重中。

- 完整心智需要**持续学习**（continual learning）——在与世界交互的过程中不断更新自己的知识和策略。
- 这在技术上是一个巨大的挑战：当前模型在训练后的”持续学习”容易导致**灾难性遗忘**（catastrophic forgetting）——学习新知识的同时丢失旧知识。

Sutton 作为强化学习的奠基人，他心目中的”完整心智”几乎必然是一个**强化学习智能体**——通过与环境的持续交互来学习和适应。这与当前以”预训练 + 微调”为核心范式的大语言模型形成了鲜明对比。

7.3.3 战略层面：研究者应该怎么做？

对于 AI 研究者，Sutton 的演讲传递了一个清晰的战略信号：

对研究者的战略建议

1. **如果你在做基础模型**：不要追求”全知”——接受”准确存储已确立知识”的定位，解决幻觉问题，做好知识检索和定制化呈现。
2. **如果你在追求通用人工智能（AGI）**：基础模型不是通往 AGI 的路径——你需要的是能**持续学习**、**创造新知识**、通过与世界交互而进化的系统。
3. **对于所有人**：不要被人类知识分心。选择能随算力扩展的方法，选择能创造知识而非利用已有知识的方法。

7.4 延伸思考与开放问题

1. **大语言模型真的只是”知识编码”吗？**当前一些前沿研究正在尝试让大语言模型进行”搜索”（如 OpenAI 的 o1 系列使用的思维链搜索和自我对弈验证）。这是否意味着大语言模型正在向”知识创造”范式靠拢？如果是，Sutton 的判断可能需要修正。
2. **”创造知识”的边界在哪里？** Sutton 将 AlphaZero 式的自我对弈视为”创造知识”，但 AlphaZero 的”知识”仍然局限于一个规则明确的封闭世界（棋盘游戏）。在开放世界中，”创造知识”意味着什么？如何评估？
3. **持续学习的工程可行性**：Sutton 呼唤的”完整心智”需要持续学习，但当前技术在这方面进展有限。这是否意味着通往”完整心智”的路径比 Sutton 暗示的更加漫长？
4. **基础模型的”超级百科全书”定位是否过于保守？**如果基础模型被定位为”已确立知识的准确存储”，那么谁来判断什么知识是”已确立”的？在知识快速演进的领域，这种定位是否会让基础模型始终滞后于前沿？

7.5 拓展阅读

- **Sutton 原文**：Richard S. Sutton, “The Bitter Lesson”, 2019. `cs.toronto.edu/~hinton/.../BitterLesson.pdf`
- **Sutton 26 词推文**：X/Twitter, 2026, 总结苦涩教训的核心要义。
- **Rich Sutton 个人主页**：`incompleteideas.net` —包含其全部论文和研究哲学阐述。

- **强化学习经典教材**：Sutton & Barto, *Reinforcement Learning: An Introduction* (2nd ed., MIT Press, 2018) —理解 Sutton 思想体系的必读。