斯蒂芬·沃尔夫勒姆的计算宇宙：一次深度访谈的全方位解读

前言：一场通往计算宇宙的智力探险

一份为未来科学家准备的指南

欢迎你，未来的科学家。你即将踏上一段非同寻常的智力旅程。这份报告将完整呈现英国物理学家、计算机科学家、企业家斯蒂芬·沃尔夫勒姆（Stephen Wolfram）在伦敦数学科学研究所的一场深度访谈。但这远不止是一篇翻译稿。它的核心目标是为你——一位站在知识前沿、对世界充满好奇心的高三学生——提供一把钥匙，去开启理解21世纪科学思想的一扇大门。

斯蒂芬·沃尔夫勒姆是一位当代罕见的思想家和“建造者” 。他不仅提出了颠覆性的科学理论，还亲手创造了实现这些理论的强大工具，如全球数百万科学家和工程师使用的计算软件 Mathematica 和知识引擎 Wolfram|Alpha 。他的人生轨迹本身就是一部传奇：15岁发表第一篇科学论文，20岁获得加州理工学院理论物理学博士学位，21岁成为当时最年轻的麦克阿瑟“天才奖”得主 。然而，他早早地离开了传统的学术轨道，选择了一条更独立、更具挑战性的道路，致力于探索一个宏大而统一的理念：**我们宇宙的本质是计算**。

这份报告将逐字逐句地翻译沃尔夫勒姆的访谈内容，并在每一段译文之后，为你提供一段长度相当甚至更长的深度解读。这些解读旨在：

1. **搭建知识的桥梁**：解释访谈中出现的关键概念，如“元胞自动机”、“计算不可约性”和“Ruliad”，将它们与你已有的物理、数学和生物学知识联系起来。
2. **揭示思想的脉络**：展示沃尔夫勒姆的论证逻辑，看他如何从一个简单的想法出发，逐步构建出一个能够解释物理定律、生命演化甚至人类意识的宏伟框架。
3. **提供批判性的视角**：沃尔夫勒姆的理论极具争议，报告也会客观地呈现科学界对他的批评和质疑，让你理解科学进步并非一帆风顺，而是在激烈的思想碰撞中前行。

核心思想：一种新科学

沃尔夫勒姆工作的核心论点，集中体现在他2002年出版的巨著《一种新科学》（*A New Kind of Science*）中。他认为，几个世纪以来，科学主要依赖复杂的数学方程式来描述世界。然而，自然界中无处不在的复杂性——从湍流的形态到生物的多样性——往往无法用简洁的公式来概括。沃尔夫勒姆提出，宇宙的底层逻辑可能并非复杂的数学，而是**简单的计算规则** 。就像几个简单的音符可以组合成宏伟的交响乐一样，宇宙通过在时空中不断迭代执行这些简单规则，涌现出了我们所见的一切复杂现象。

这趟旅程将挑战你的许多既有观念。它要求你像沃尔夫勒姆一样，敢于回到最基本的问题，重新审视那些被认为是理所当然的“基石”。这不仅是学习知识，更是学习一种全新的思考方式——**计算思维**。准备好了吗？让我们一同进入斯蒂芬·沃尔夫勒姆的计算宇宙。

第一部分：历史的重量与探寻根基

本部分涵盖了访谈的开篇，聚焦于沃尔夫勒姆如何看待科学史，以及历史在他的当代发现中扮演的关键角色。他认为，历史并非尘封的故纸堆，而是蕴藏着未来突破线索的战略地图。

1.1 开场：背景与影响

【访谈原文翻译】

>> 托马斯·弗林克：真是高朋满座。哇哦。我是托马斯·弗林克，伦敦数学科学研究所的主任。我们是英国第一家独立的物理学和数学研究机构。我们非常高兴能请到斯蒂芬·沃尔夫勒姆先生。这是我们第一次见面。斯蒂芬虽然是英国人，但却是从美国远道而来。

>> 斯蒂芬·沃尔夫勒姆：我是在美国长大的，现在这里是我的家。

>> 托马斯·弗林克：事实上，当我在德克萨斯州长大的时候，有两位科学家对我产生了巨大的影响。第一位是罗杰·彭罗斯，我在十几岁时读了他的书，他让我对物理学感到兴奋。第二位影响我的人，是在我作为加州理工学院本科生时遇到的斯蒂芬·沃尔夫勒姆，他让我对研究产生了热情。特别是他关于元胞自动机的工作，我作为本科生发表的第一篇论文就是一个关于地貌学中模式形成的元胞自动机模型。所以，能请你来这里真是太好了。

斯蒂芬是一位物理学家、计算机科学家、人工智能开发者、数学家、科学传播者和科技创始人。他在十几岁离开牛津时，就已经发表了10篇理论物理学论文。他去了加州理工学院，在那里获得了博士学位，我想大概是在20岁时，成为麦克阿瑟天才奖最年轻的获得者。之后他加入了加州理工学院的教职团队。大约就在那个时候，斯蒂芬开始发展一些想法，这些想法后来演变成了 Mathematica——你们中大多数人可能都知道，Mathematica 是我们都用来进行符号操作和处理方程的工具，后来它发展成了一门庞大的编程语言。

几年后，他去了普林斯顿高等研究院，我想也正是在那个时候，斯蒂芬开始真正对由简单规则产生的复杂行为感到兴奋。如今，斯蒂芬是 Wolfram Research 公司的创始人，这家公司是 Mathematica、Wolfram|Alpha、Wolfram Language 以及最近的物理学项目和元数学项目的幕后推手。他在至少十几个科学技术领域都是先驱。所以，能有机会采访他，我们感到非常荣幸。我们将采访他一个小时，然后进行大约30到40分钟的问答环节。与我同在的还有阿南约·巴塔查里亚，他是我们的首席科学作家。阿南曾在《自然》杂志担任编辑，之后在《经济学人》担任科学作家，最近还写了约翰·冯·诺依曼的权威传记。那么，让我们开始吧。

>> 阿南约·巴塔查里亚：斯蒂芬，欢迎来到伦敦数学科学研究所。我们是英国唯一独立的理论物理研究所，虽然我们很自豪能身处法拉第过去的房间里，但我们独立于皇家学会。所以有时会有些混淆，但我们确实独立于皇家学会。斯蒂芬，我们一直在痴迷地听你的播客等等。作为冯·诺依曼的传记作者，或者说他最新的传记作者，让我印象深刻的是你对科学史的热情。你似乎对此非常感兴趣。所以我想问你，你认为科学史应该或实际上如何在当代发现中发挥作用？

【深度解读】

这段开场白不仅仅是礼节性的寒暄，它为整个访谈设定了重要的基调和背景。

首先，**对话的场所极具象征意义**。伦敦数学科学研究所（LIMS）被强调为“英国第一家独立的物理学和数学研究机构”，并且坐落在“法拉第过去的房间里”，但又“独立于皇家学会”。这暗示了一种精神：既继承了以法拉第为代表的伟大科学传统，又寻求一种独立于主流建制（如皇家学会）的、更灵活、更具开创性的研究模式。这恰好与沃尔夫勒姆本人的职业生涯相呼应——他大部分的突破性工作都是在他自己创立的公司和研究所里，而非传统大学体制内完成的 。

其次，**主持人的背景揭示了沃尔夫勒姆思想的深远影响**。主任托马斯·弗林克坦言，他本科时代的第一篇科研论文，就是受到了沃尔夫勒姆关于“元胞自动机”（Cellular Automata）工作的启发。这并非偶然，它直接点出了沃尔夫勒姆科学思想的核心工具和方法论——**通过简单的计算模型来理解复杂世界的模式形成**。元胞自动机这个概念将在访谈后半部分被反复提及，这里提前埋下了伏笔，让听众知道这是一种能够应用于具体科学问题（如地貌学）的强大思想。

最后，**另一位主持人的身份至关重要**。阿南约·巴塔查里亚是约翰·冯·诺依曼（John von Neumann）的传记作者 。冯·诺依曼是20世纪最伟大的数学家和科学家之一，他不仅是计算机架构的奠基人，也是元胞自动机理论的共同开创者 。然而，沃尔夫勒姆对冯·诺依曼的看法颇具争议，他认为冯·诺依曼虽然开创了这个领域，但并未真正洞察到其最核心的奥秘——即极简规则能自发产生无限复杂性 。因此，当冯·诺依曼的传记作者向沃尔夫勒姆提问时，这不仅仅是一个普通的采访，而是一场跨越时空的思想对话，隐含着对科学遗产、学术传承和思想突破的深刻探讨。这个问题“科学史如何在当代发现中发挥作用？”也因此变得格外尖锐和深刻。

1.2 科学史：一张通往未来的藏宝图

【访谈原文翻译】

>> 斯蒂芬·沃尔夫勒姆：你知道，当我认为自己想出了一些新颖且与众不同的东西时，我首先要做的事情之一就是通过回顾人们以前为什么有不同的看法，来确保我对自己所谈论的东西有充分的了解。例如，最近我一直在研究热力学第二定律。我从12岁起就对热力学第二定律感兴趣，从那时起就一直试图弄明白它的原理。我想我终于搞懂了，然后我就会想：为什么以前的人没有发现这一点？所以我决定，我最好回过头去，真正理解热力学第二定律的全部历史，追溯到19世纪20年代等等。在梳理了那段历史之后，我现在知道了，为什么在那么长的时间里，人们没有看到那些我现在认为显而易见的东西。

所以对我来说，这是一种自我定位的方式，你知道，我如何确定我现在所做的事情是有意义的？我在物理学领域最近做的很多事情，都打破了人们长期以来的观念。比如，其中一个问题是：空间是离散的还是连续的？自古以来人们就一直在思考这个问题。我的意思是，在古代，有德谟克利特的原子论，也有赫拉克利特的一切皆流的思想。直到19世纪末，任何事物是离散的还是连续的，都还不清楚。而在19世纪末，很可能就在这座建筑里，人们发现了关于物质分子结构的东西。

我直到最近才知道，在20世纪初，大多数著名的物理学家都会说，是的，空间也是离散的。他们非常努力地尝试建立离散的空间模型。但他们无法使其与相对论保持一致，所以到了20世纪30年代，他们就放弃了。关于这方面的文献很少，因为他们放弃了，没有

写下来。最终，我在20世纪90年代，以及大约五六年前，终于弄清楚了空间如何可以是离散的，而这解锁了各种各样不同的东西。但是，了解这段历史对我来说非常重要，让我感觉，你知道，我可能真的知道自己在说什么，因为这正是100年前人们在做的事情，他们只是卡住了。

>> 阿南约·巴塔查里亚：当我研究《来自未来的人》那本书以及与物理学家交谈时，我感觉没有多少人真正有同样的想法。我的意思是，你会得到一种相当陈词滥调的看法，比如爱因斯坦独自坐在他的专利局里，突然间就发现了相对论，没有任何……你知道……前期的铺垫，谁是洛伦兹？谁是庞加莱？诸如此类的。

>> 斯蒂芬·沃尔夫勒姆：你必须记住，那个时代的专利局就是当时的硅谷。我的意思是，在伯尔尼专利局工作，就像在一家硅谷公司工作一样。所以那并不完全是……你知道……被扔到外面的黑暗中。

>> 阿南约·"巴塔查里亚：是的。是的。绝对是。但你是否认为，对于当代物理学或数学来说，这是一个问题——人们似乎对自己的根源知之甚少，也许没有认真对待历史？

【深度解读】

沃尔夫勒姆在这里揭示了他独特的研究方法论：**将科学史作为一种战略工具，而非仅仅是背景知识**。对他而言，历史研究的目的不是为了瞻仰过去的辉煌，而是为了精准地定位当下最有价值的突破口。

他的逻辑链条是这样的：

1. **发现一个新想法**：当他认为自己对某个基本问题（如热力学第二定律）有了新的理解时。
2. **战略性回溯历史**：他会系统地研究该问题的整个历史，从起源（19世纪20年代）到发展。
3. **寻找“卡点”（Sticking Point）**：他的研究重点不是看前人做对了什么，而是**“为什么人们没有看到那些我现在认为显而易见的东西？”** 他在寻找的是历史上的思想瓶颈或技术障碍。
4. **定位突破口**：一旦找到“卡点”，他就知道了自己的新理论需要解决的核心矛盾。例如，关于“空间是离散的还是连续的”这个问题，他发现20世纪初的物理学巨匠们（如爱因斯坦和海森堡）普遍相信空间是离散的，但他们最终放弃了，因为无法将离散空间模型与爱因斯坦的相对论兼容 。这个“卡点”——**离散性与相对论的矛盾**——就成了沃尔夫勒姆物理学项目要攻克的核心堡垒。

这种方法论与传统科学教育中呈现的“英雄式”科学发现叙事形成了鲜明对比。传统的叙事常常将爱因斯坦描绘成一个横空出世的天才，独自在专利局里颠覆了整个物理学。而沃尔夫勒姆则提醒我们，爱因斯坦身处的伯尔尼专利局是当时的技术创新中心，相当于“那个时代的硅谷”，他并非与世隔绝。更重要的是，任何科学突破都建立在洛伦兹、庞加莱等前人的工作之上。

沃尔夫勒姆的观点对你这样的学习者有着深刻的启示：

- **知识的“考古学”**：当你学习一个理论时，不仅要学习它的结论，更要去了解它是在怎样的历史背景下、为了解决什么问题而提出的。那些被放弃的理论、失败的实验、未解的悖论，往往比成功的理论更能揭示该领域的深层结构和未来的可能性。

- **自信的来源**：沃尔夫勒姆说，了解历史让他“感觉自己可能真的知道在说什么”。这种自信并非源于盲目自大，而是建立在对历史脉络的清晰认知之上。他知道自己的工作不是凭空想象，而是在回应一个被搁置了近一个世纪的重大科学问题。这是一种深刻的、有根基的学术自信。

总而言之，沃尔夫勒姆将科学史从一门“人文学科”转变为一种强大的“科学研究工具”。他像一位侦探，在历史的故纸堆中寻找线索，定位那些被遗忘的战场，然后带着现代的工具（强大的计算机和新的计算思想）重返战场，试图一举攻克前人未能逾越的难关。

1.3 科学的制度化与创新的土壤

【访谈原文翻译】

>> 斯蒂芬·沃尔夫勒姆：看，问题在于，当一个科学领域形成时，通常会有一些方法论上的进步，然后有五年、十年，也许二十年，是采摘“低垂果实”的时期。大量的发现涌现，领域经历一个非常快速的成长期。然后，事情就慢下来了，在接下来的一百年里，你知道，它会是缓慢增长。这个领域变得非常制度化，会有一种力量抵制新思想等等。我认为……你知道……那种认为应该有……你知道……可以产生新事物的地方，而这些地方通常是……你知道……在现有事物的整个制度化结构之外的。我的意思是，在任何……你知道，我很幸运地参与了许多在我研究它们时还是新兴的领域。而且，你知道，新兴领域的增长速度远大于那些发展成熟、制度化的领域。

>> 阿南约·巴塔查里亚：既然我们谈到了历史和影响，你能告诉我们，当你年轻时，刚开始涉足粒子物理学等领域时，有哪些科学家影响或启发了你？

>> 斯蒂芬·沃尔夫勒姆：你知道，这类事情的问题在于，最终你总会真的见到那些人。你会发现他们并不像……嗯……这……你知道，我想我更多地是被一种集体氛围所激励，那就是在20世纪70年代，当我为粒子物理学感到兴奋时，每周都有新进展发生。这有点像今天的机器学习。你知道，每周都有进步。那是非常鼓舞人心的。我当时认识了那个时代粒子物理学领域的大多数泰斗。嗯……我想说……你知道……迪克·费曼（Dick Feynman）是我相当熟悉的一个人。嗯……他……我想说……我在某些事情上同意他，但在其他事情上则不然。但他有一个……你知道……他是一个非常热衷于探究事物本质的人，这一点我也非常喜欢。我的感觉是，这就是你取得进步的方式，通过真正理解事物的根基。

你刚才问到领域的发展以及人们关注什么。我注意到的一件事是，你知道，有些领域的创始人，他们总是对基础是否真的正确有点不确定，并且非常乐意去探索这些基础。然后你往下走，比如说，五个学术代际之后，到那个时候，从事该领域研究的人会说：基础？我们知道基础是正确的。它们是……你知道……50年前就建立起来了，我们不需要再看基础了。而就我而言，我在许多事情上取得的进展，往往是通过回头再次攻击这些基础，因为已经有50年或100年没人看过它们了。

【深度解读】

在这一段中，沃尔夫勒姆提出了一个关于科学领域发展的“生命周期模型”，并借此解释了为什么他选择了一条非传统的科研道路，以及为什么他如此强调“攻击基础”。

科学领域的生命周期模型可以概括为两个阶段：

1. **爆发期（“采摘低垂果实”）**：当一个新的方法论或工具出现时（例如，量子力学的建立、计算机的发明），会开启一个短暂而辉煌的时期。在这个阶段，有大量的、相对容易的发现等待着人们去摘取。沃尔夫勒姆将上世纪70年代的粒子物理学和今天的机器学习作为例子，这两个领域在特定时期都呈现出“每周都有新进展”的爆发式增长。
2. **停滞与制度化期**：当“低垂的果实”被采摘完毕后，该领域的发展速度会显著放缓。此时，它会变得高度“制度化”——形成固定的研究范式、学术等级、经费评审机制和教育体系。这种制度化的结构虽然有助于知识的传承和深化，但其内在的保守性往往会“抵制新思想”，使得颠覆性的创新变得异常困难。

基于这个模型，沃尔夫勒姆揭示了他的核心策略：**避开成熟领域，攻击被遗忘的基础**。他认为，在一个高度制度化的领域里，学术传承就像一场“传话游戏”。领域的创始人（第一代）对理论基础是抱有审慎和怀疑态度的，因为他们是亲手奠基的人。但传到第五代学者时，这些基础已经被奉为金科玉律，变成了教科书里不容置疑的“事实”。他们会说：“我们不需要再看基础了。”

这恰恰为像沃尔夫勒姆这样的“局外人”创造了巨大的机会。当一个领域的基础已经“50年或100年没人看过”时，意味着：

- **知识的代差**：这50-100年间，数学、计算机科学等相关领域可能已经发展出了全新的工具和观念。用这些新工具去重新审视旧的基础，很可能会发现前人无法看到的东西。
- **思维的僵化**：领域内的学者由于长期在同一个范式内思考，可能已经形成了思维定势，对基础的某些根本性缺陷视而不见。
- **高杠杆效应**：对基础的任何一点微小修正，都可能引发整个领域的连锁反应，产生巨大的影响力。这是一种“高杠杆”的创新策略。

他提到对理查德·费曼（Richard Feynman）的欣赏，也并非全盘接受，而是有选择性的。他欣赏的是费曼那种“探究事物本质”的精神，这与他“攻击基础”的策略不谋而合。这表明，他所推崇的并非某个特定的科学家偶像，而是一种敢于质疑、直面根本的科学精神。

对于立志于科学研究的学生来说，这里的教训是深刻的：不要仅仅满足于在一个既定框架内解决问题。要时刻保持对学科基础的批判性审视，并主动学习那些来自“领域之外”的新工具和新思想。真正的重大突破，往往发生在那些被主流学界认为是“早已解决”或“不值得研究”的地方。

第二部分：生命与物理的计算之心

这是访谈的核心部分。沃尔夫勒姆在这里将他最深刻、最具颠覆性的思想串联起来，试图用一个统一的计算框架来解释两个看似毫不相关的宏大主题：**生命的演化**和**物理学的基本定律（特别是热力学第二定律）**。他将揭示，在这两者背后，都隐藏着同一个根本性的原理。

2.1 生物学的新根基：从元胞自动机到演化论

【访谈原文翻译】

>> 托马斯·弗林克：我能……我能……我能提一个我认为需要被攻击的基础吗？嗯，我想到的是能够学习的自我复制机器。也就是生命，但我更愿意不用“生物学”这个词，因为我希望保持更广泛的概括性。你知道，你第一次写信给我们（或者说给我），是在我们完成了那份包含23个数学挑战的清单之后，那份清单就挂在外面的大厅里。我们做这个是因为（英国新的科研机构）ARIA正在创立。它有点像英国版的DARPA，其中一些挑战就是关于……你知道……关于生命的。它是如何运作的？一种生命的热力学模型。物理学似乎对生命毫无解释。你知道，如果你把……说到基础……突变、遗传和选择放进一个盒子里摇晃，并不会发生什么了不起的事情，至少我是这么认为的。我当然没见过什么值得大书特书的结果。那么，到底是怎么回事？为什么……为什么生命的基础……你知道……如此……你知道……薄弱？

>> 斯蒂芬·沃尔夫勒姆：你知道，碰巧在过去几周里，我一直在研究的正是这个问题。

>> 托马斯·弗林克：啊，那真是时机正好。

>> 斯蒂芬·沃尔夫勒姆：我……我大概在几周后，会最终把我一直在研究的东西发布出来，但我可以先跟你说说。所以，嗯……好吧……这个问题的一种思考方式是，我们有一个生命系统。回到法拉第等人的时代，人们会说我们是由“原生质”（protoplasm）构成的。那原生质到底是什么鬼东西？你知道，我们是由液体构成的吗？比如说，液体是什么？液体就是一群分子在随机地碰撞，偶尔发生反应等等。但我们可能并不是真正的液体，因为我们从分子生物学中知道，存在着所有这些主动运输过程，它们在移动分子，并精心组织……你知道……这些大分子如何排列等等。所以，我们真正的形态是某种**宏观上被精心编排的分子集合体**（bulk orchestrated collection of molecules）。

我们能有一个关于“宏观编排”的理论吗？问题就在于此。嗯……你知道……我们有统计物理学理论，我们说，如果你把足够多的气体分子放进一个盒子里，热力学第二定律很可能会主导，分子们的构型将被我们模拟为大致随机的，然后我们可以从中得出各种结论，比如气体定律和流体动力学等等。但在生命这个例子中，分子们不是随机碰撞，而是被精心编排的，我们能建立一个关于“宏观编排”的理论吗？

我思考这个问题已经好几年了，几周前终于取得了一点突破。所以，事情大概是这样的，我以前没有谈论过这个，所以可能会有点粗糙。嗯……好的，它始于我去年想明白的一件事，这件事我在20世纪80年代就试图弄明白但失败了。就是之前提到的元胞自动机系统，约翰·冯·诺依曼等人研究过的。你知道，它们有非常简单的规则。它们只是一个细胞阵列，规则仅仅是说，这个细胞的颜色由它上一行细胞的颜色决定，比如说。而最大的……最大的惊喜是，即使规则可以很简单，系统的行为却可以非常复杂。

好的，这我在80年代就知道了。然后我想，让我用这个来作为生物学的模型吧。让我说……你知道……我有这些规则，但它们就像是基因组规则，是基因组。然后系统的行为，也就是元胞自动机所做的事情，就像是表型（phenotype）。让我尝试去突变这些规则，看看我能对表型做什么。我在80年代尝试过，但没有成功。

去年，我带着一个新的直觉回到了这个问题上。这个新的直觉是**机器学习**，以及机器学习确实有效这个事实。以及从2011年左右开始的那个惊喜，即，如果你足够猛烈地“敲打”一个神经网络，它就会学习，这在当时对任何人来说都不是显而易见的。我的意思

是，我们……你知道……它能成功这个事实本身就是一个巨大的惊喜。所以我想，让我们试着更猛烈地“敲打”这个系统，看看它是否能演化。

结果它确实能。你发现会发生这样的事情：我用的最简单的适应度函数（fitness function）是，让一个模式存活尽可能长的时间，但又不是无限长。然后你发现，它会发现这些非常精巧的形式。它会做出一些发现，比如，哦，如果你让这个“奔跑者”（runner）这样走，然后再那样走，那么你就能活得更久一点。然后它会基于这个发现继续构建。你知道，再为基因组的突变掷一次骰子，它会做出一个不同的发现，然后以一种不同的方式去构建。这非常奇妙地像化石记录。你知道，比如说，它发现了如何制造……我不知道……爪子之类的东西，然后它会制造出越来越精巧的爪子来做事。

所以这就是我去年意识到的事情，就是……你得到了……生物学在做什么？它在利用计算的这种能力来制造复杂的东西，并且它是在适应这些适应度函数的情况下这样做的。而其中……呃……我发现一件既有趣又令人震惊的事情是，你会得到这些非常精巧的模式，这些问题的解决方案，是你作为人类永远也猜不到的。它们不是你能够设计出来的任何东西。它们只是作为“不可约计算”（irreducible computation）的某些特征被发现，而这些特征恰好能起作用。

我最近尝试做的一件事是，把这些模式中的一些喂给一个大语言模型（LLM），然后说：“写一段描述。” 它写出了一段描述。它说，你知道，这里有……它会编造一些名字，给这些不同的小结构命名等等。你读起来，感觉就像一本生物学教科书。这有点……你知道……它真的是这样。就好像在生物学中，我们被抛给了这样一个……你知道……随机计算恰好起作用的例子。顺便说一下，我认为这与机器学习的工作方式非常相似，但那是另一个故事了。

无论如何，回到那个问题，关于……你知道……什么是生物物质，那到底是什么？我意识到的是，你看到的是，你有一些适应度函数，对吧？比如“尽可能活得久”等等。然后你会得到一些关于如何实现该适应度的解决方案。好的，然后我思考了下面这件事。有没有可能……哦，我应该先解释另一件事。

【深度解读】

这一段是理解沃尔夫勒姆生物学思想的关键，他在这里提出了一个全新的、基于计算的生命演化模型。为了理解它，我们需要拆解他的论证步骤。

第一步：重新定义生命——“宏观编排的分子集合体”

访谈者提出的问题非常深刻：“物理学似乎对生命毫无解释”。传统的物理学，如统计力学，擅长描述大量分子的**随机**行为（如盒子里的气体），并从中得出宏观定律（如气体定律）。但生命的核心特征恰恰是**非随机性**。从分子生物学我们知道，细胞内的分子运动是被高度组织和“精心编排”的，存在着主动运输、精确折叠等过程。沃尔夫勒姆将生命体定义为“宏观上被精心编排的分子集合体”，从而将问题的核心从“生命是什么物质”转向“**生命是如何实现这种宏观编排的**”。他试图为这种“编排”本身建立一个理论。

第二步：引入模型——元胞自动机（CA）

为了研究“编排”的原理，沃尔夫勒姆选择了他最熟悉的工具——元胞自动机。这里，我们需要为这个概念建立一个清晰的图像。

- **什么是元胞自动机？** 想象一下一排无限长的格子（一维CA），每个格子要么是黑色要么是白色（两种状态）。时间以离散的“步”或“代”推进。在每一步，每个格子的下一代颜色，都由一个**简单固定的规则**根据它自己以及它左右两个邻居**当前**的颜色来决定 。例如，一个规则可能是：“如果左邻居是黑色，你自己是白色，右邻居是白色，那么你下一代就变成黑色。” 这样的规则组合起来，可以覆盖所有8种可能的邻居颜色组合。
- **核心惊喜**：沃尔夫勒姆在80年代的重大发现是，即使规则本身极其简单，从一个简单的初始状态（比如只有一个黑格）开始演化，其产生的模式却可以**极其复杂和不可预测** 。有些规则产生简单的重复图案，而另一些（如著名的“规则30”）则产生看起来完全随机的复杂结构。
- **生物学类比**：沃尔夫勒姆将这个模型类比于生物学：
  - **CA规则 ↔ 基因组（Genotype）**：一套简单的、决定细胞行为的指令。
  - **CA演化出的模式 ↔ 表型（Phenotype）**：由基因组指导发育成的复杂生物体。

第三步：80年代的失败与今日的突破——“敲打”系统

沃尔夫勒姆坦言，他在80年代尝试用这个模型模拟生物演化失败了。他当时的做法是随机“突变”CA的规则（基因组），然后观察表型的变化，但没有得到有意义的“演化”。

突破来自于一个全新的直觉——**机器学习的成功**。2011年后，深度学习的实践表明，只要有足够大的模型、足够多的数据和足够强的计算能力（可以理解为“猛烈地敲打”），神经网络就能“学会”非常复杂的任务。这启发了他用类似的方法来“训练”他的元胞自动机模型。他设定了一个简单的“适应度函数”（fitness function），这在演化理论中相当于环境的选择压力。他用的例子是：“**让模式存活尽可能长的时间，但又不是无限长。**”

第四步：惊人的结果——计算的“化石记录”

当他用这个适应度函数来筛选成千上万个随机突变的CA规则时，系统“演化”出了令人惊叹的结果。它不再是无意义的模式，而是会自发地“发现”各种精巧的结构和策略来延长自己的“生命”。比如，系统可能会发现一种像“滑翔机”一样的结构，然后通过突变不断优化它，让它走得更远、更稳定。这个过程，沃尔夫-勒姆形容“**非常奇妙地像化石记录**”。就像生物演化史上，一旦“发明”了眼睛或翅膀这种成功的结构，就会在其基础上不断进行修饰和改良。

更令人震惊的是，这些演化出的“解决方案”是人类设计师永远无法想到的。它们是计算宇宙中固有的、“非人”的智慧。当他让大语言模型（LLM）去描述这些模式时，LLM竟然编造出了一套听起来很专业的术语来描述这些结构的互动，读起来“**就像一本生物学教科书**”。

这引出了一个深刻的结论：生物学所展现的复杂性和精巧性，可能并非源于某种神秘的“生命力”或“智能设计”，而是**计算本身固有的创造力**。生命，在沃尔夫勒姆看来，就是利用了计算宇宙中那些能够满足环境适应度要求的、被偶然“发现”的复杂计算过程。它是一种“碰巧起作用的随机计算”。

2.2 计算不可约性与热力学第二定律的新解释

【访谈原文翻译】

>> 斯蒂芬·沃尔夫勒姆：生物演化为什么会起作用？换句话说，为什么我们能够成功地……你知道……适应，变得越来越适应环境等等？我认为答案需要几个步骤，也许我不会把所有步骤都讲完，但基本上要点是……好吧，实际上我应该先回头说点别的，那就是热力学第二定律为什么会起作用。我提到我最近研究了这个问题。

好的，有一个观察结果是，非常简单的规则可以产生非常复杂的行为。这个观察结果的推论之一，是一种我称之为“**计算不可约性**”（computational irreducibility）的现象。你可能会想，如果你有了某个东西的规则，那你就大功告成了，你就知道了关于系统行为的一切。从某种意义上说，确实如此，因为你可以运行这些规则，看看会发生什么。但问题是，你知道，你可能需要运行十亿步规则。问题在于，你能不能跳过这些步骤，直接说：“我知道会发生什么，我有一个答案的公式。” 对吧？我的论点是，对于计算系统来说，通常你无法做到这一点。它们通常是**计算上不可约的**。了解它们将要发生什么的唯一方法，就是去运行它们，看看会发生什么。好的。

所以这是一个基本的现象，它导致了我们随处可见的这种表观上的复杂性等等。好的。热力学第二定律。它的想法是，你有一个可能非常简单的初始条件。分子在盒子里四处碰撞。原则上，你总是可以逆转你得到的构型，回到你最初的状态。但实际上，你无法做到这一点。我认为你无法做到的原因，是因为**你，作为这个过程的观察者，在计算上是弱小的，相比于系统实际力学过程中发生的不可约计算而言**。

>> 主持人：你的意思是，我们无法找出那个复杂的构型，如果时间倒流，它会回到那个有序的、特殊的构型？

>> 斯蒂芬·沃尔夫勒姆：不，我们看到的是这些构型中的一个，而对我们来说，它“哦，它来自一个简单的东西”这一点并不明显，因为我们必须运行所有的计算才能弄明白。由于计算不可约性，我们无法做到这一点。所以在某种意义上，**第二定律是我们作为计算上受限的观察者，相对于底层动力学的计算能力而言，所导致的一个结果**。

这很有趣，回到……我又一次被这个房间提醒，回到了19世纪60年代。好的，当时热力学第二定律还很年轻，像开尔文勋爵这样的人到处宣扬，它将导致宇宙的“热寂”。过一段时间后，整个宇宙都会衰败，我们最终只会剩下……你知道……随机的热量，一切都会变成那样。那么，这幅图景错在哪里？

关键在于，对于像我们今天的观察者来说，看起来事物会趋于随机化。但是……但是，由今天的我们这些分子构成的未来，在未来将会是某种……你知道……不同的东西。如果我们是那个未来的更复杂的观察者，我们就会知道：“哦，那些分子……你知道……在万亿年后，是2025年发生的这场对话的结果。” 只是因为我们是……你知道……作为计算上受限的观察者，我们只会说：“哦，那只是随机的热量。” 所以……所以，你知道……宇宙的热寂只对像我们这样的观察者存在。热力学第二定律只对像我们这样的观察者存在。

事实证明，这种**底层计算不可约性与像我们这样的观察者的计算弱小性之间的相互作用**，原来是我认为我在过去几年里发现的最令人兴奋的事情之一，那就是**物理定律的来源**。广义相对论、量子力学和热力学第二定律，都是这种相互作用的结果。

> 几个月前的一个惊喜是，我意识到**生物演化是同一个故事**。也就是说，你有所有这些可能发生的突变，你在问，哪些突变是个好主意？嗯，有一些适应度函数定义了哪些是好主意。如果那个适应度函数在计算上极其复杂，它会挑选出某个奇怪的分支。但正是因为它是一个**计算上弱小的适应度函数**，生物演化才能起作用。
>
> 举个例子，假设每个生物体一出生就必须会做微积分。基本上，你不会有任何生物体能存活下来，因为这件事太难了，太精细了，无法让生物体做到。但因为它所要做的只是……你知道……闲逛，然后制造更多的生物体等等，这是一个足够粗糙的事情，以至于……有机体发育的底层计算不可约性足够强大，使得那种弱目标是可以实现的。

【深度解读】

这是整场访谈的理论核心。沃尔夫勒姆在这里正式提出了他最重要的两个概念——“计算不可约性”和“计算等价性原理”（尽管他没有明确说出后者），并用它们来统一地解释物理学的根基（热力学第二定律）和生物学的根基（演化论）。

第一步：定义“计算不可约性”（Computational Irreducibility, CI）

这个概念是理解沃尔夫勒姆整个思想体系的基石。

- **定义**：对于一个由确定性规则驱动的系统，如果我们无法找到一个“捷径”（比如一个数学公式）来预测它在未来的状态，而只能一步一步地模拟它的演化过程，那么这个系统就是计算不可约的 。
- **类比**：想象一下烤蛋糕。食谱（规则）和原料（初始条件）都是确定的。但你无法通过看一眼食谱和原料就直接“计算”出蛋糕最终的味道和质地。你必须严格按照步骤——混合、搅拌、烘烤——走完整个过程，才能知道结果。这个烘焙过程本身就是“不可约的” 。同样，对于一个计算不可约的系统，**时间的流逝本身就是一种不可简化的计算过程** 。
- **来源**：CI 是“简单规则能产生复杂行为”这一现象的直接推论。正是因为行为如此复杂和不可预测，我们才找不到捷径。

第二步：用CI重新解释热力学第二定律

这是沃尔夫勒姆对物理学的一个颠覆性贡献。

特征	传统观点 (玻尔兹曼)	沃尔夫勒姆的计算观点
随机性的来源	**统计概率**：一个系统（如气体）的无序状态数量远远多于有序状态，因此系统“几乎总是”会演化到更无序的状态 。	**内在生成**：系统遵循简单的确定性规则，但其行为是计算不可约的，从而**内在**地生成了看起来随机的复杂模式 。
熵的本质	**微观状态的度量**：熵是对系统微观无序程度或我们所缺失的信息的度量。	**观察者与系统间的计算差异**：熵是我们作为**计算能力有限的观察者**，在面对一个计算不可约的系统时，所必然感知到的现象 。
观察者的	**被动**：观察者的“粗粒化”视角帮助定义了宏观状态，但不是定律的根本	**核心与主动**：**观察者的计算局限性是第二定律存在的原因**。我们因为无法“看穿”系统复杂的计算过程，所以只能将其感知

特征	传统观点 (玻尔兹曼)	沃尔夫勒姆的计算观点
角色	原因。	为无序的“热”。
可逆性悖论	**一个历史难题**：既然微观定律是可逆的，为什么宏观世界的时间之箭是单向的？	**轻松化解**：底层的规则确实是可逆的。但要想从一个复杂的最终状态“反向计算”回简单的初始状态，这个过程同样是计算不可约的，对于我们这样的有限观察者来说，这在实践上是不可能的。

这个新解释最惊人的推论是：**宇宙的“热寂”是相对的**。对于我们这些“计算上弱小”的观察者来说，宇宙的最终状态看起来是一片均匀、无序的随机热量。但对于一个计算能力足够强大的“超级观察者”而言，这片“热量”中仍然包含了宇宙全部的历史信息和结构。它能“看穿”这种表面的随机，识别出这些分子的构型正是我们今天这场对话在亿万年后演化的结果。因此，**热力学第二定律不是宇宙的终极宿命，而是我们观察视角的一种必然产物**。

第三步：将CI框架应用于生物演化

沃尔夫勒姆接着指出，生物演化遵循着**完全相同的逻辑**。

- **强大的底层计算**：一个生物体的发育过程（从受精卵到成体）是一个极其复杂的、计算不可约的过程。它有能力产生无穷无尽的复杂形态和功能。
- **弱小的“观察者”**：这里的“观察者”就是**环境的适应度函数**。环境对生物提出的要求通常是计算上“简单”或“粗糙”的，比如“找到食物”、“躲避天敌”、“成功繁殖”。
- **演化得以可能**：正是因为环境的目标是“弱小的”，生物体强大的、不可约的计算能力才有了足够的“创作空间”去探索和发现各种能够达成这些简单目标的复杂解决方案。如果环境的要求是计算上极其复杂的（沃尔夫勒姆的玩笑例子：“一出生就必须会微积分”），那么任何随机突变都几乎不可能恰好满足这个精细的要求，演化就会失败。

统一的图景

至此，一个宏大而统一的图景浮现出来：无论是物理学中的时间之箭（熵增），还是生物学中的演化之箭（复杂性增加），其根源都在于**一个计算不可约的强大底层系统**与一个**计算上相对弱小的采样/观察过程**之间的相互作用。这不仅解释了这两个领域的核心问题，沃尔夫勒姆更宣称，广义相对论和量子力学也源于同样的相互作用。这是一个试图统一科学主要分支的、雄心勃勃的纲领。

2.3 人工生命与“规则系综”

【访谈原文翻译】

>> 主持人：那么人工生命呢？我的意思是，它算是失败了吗？或者说，我们在人工生命中看到什么成果了吗？我知道这个领域已经持续很长时间了……

>> 斯蒂芬·沃尔夫勒姆：我不认为那是……

>> 斯蒂芬·沃尔夫勒姆：嗯，只是为了完成刚才的思路……所以生物演化之所以能起作用，是因为由环境定义的适应度函数，在计算上比生物体能做的底层力学要简单。好

> 的。那么问题来了，关于生物体正在做的底层的东西，我们能说些什么呢？
>
> 我意识到的是，你可以定义许多不同的适应度函数，但只要你知道关于它们的一件事——即它们在**计算上是简单的**——这就告诉了你一些信息，关于哪些底层规则会是在经过适应性演化以满足某个计算上简单的适应度函数后，能够成功的规则。
>
> 所以，在传统统计力学中，有“系综”（ensembles）这样的概念，即一个系统所有可能构型的集合，然后你说，我们看到的行为是这个系统所有可能构型中的典型行为。所以我有一个东西，我称之为“**规则系综**”（rule ensemble），它指的是那些与“为了一个计算上简单的目的而演化出来”这一事实相一致的规则集合。
>
> 所以我的论点是，生物学……你可以这样思考生物学的理论：不是通过了解剪接体（splicesome）或其他任何东西的所有具体机制细节，而仅仅是通过了解这一个事实——**它们是为了一个计算上简单的目的而演化出来的**。而这个事实，会反过来影响到底层正在发生的实际微观动力学。
>
> 于是你开始发现各种各样的事情。我称之为“**类机械行为**”（mechanoidal behavior）。这种行为表现为，在小尺度上似乎存在着某种机制在起作用，并且这种机制的出现具有某种必然性。

【深度解读】

在这一小节，沃尔夫勒姆将他关于生物演化的思想推向了一个更具预测性的理论框架。他试图回答一个更深层次的问题：既然我们知道了生物演化的**原理**（弱适应度函数筛选强计算过程），那么我们能否预测生物体内部会演化出**什么样的微观机制**？

为了做到这一点，他引入了一个类比于统计物理学的强大概念：“**规则系综**”（Rule Ensemble）。

第一步：理解传统物理学中的“系综”

首先，我们需要理解他所类比的“系综”（Ensemble）在物理学中是什么意思。在统计力学中，当我们研究一个宏观系统（比如一杯水）时，我们不可能知道每个水分子的精确位置和速度。于是，物理学家们想出了一个聪明的办法：他们想象一个由无数个这杯水的“副本”组成的集合，这些副本在宏观性质（如温度、压强）上完全相同，但在微观上（每个分子的具体状态）则涵盖了所有可能性。这个巨大的虚拟集合就叫做“系综”。然后，物理学家们假设，我们现实中观察到的那杯水的宏观行为（比如它的沸点），就等于这个系综中所有副本行为的**平均值或典型行为**。这是一种用统计方法绕过微观复杂性的强大工具。

第二步：构建生物学的“规则系综”

沃尔夫勒姆将这个思想从“状态的集合”（系综）推广到了“**规则的集合**”（规则系综）。他的逻辑是：

1. 宇宙中存在着无数种可能的简单计算规则（就像元胞自动机的规则库）。
2. 生命并不是从这些规则中随机挑选一个来用。相反，生命是通过演化被“塑造”出来的。
3. 塑造它的核心原则我们已经知道了：它必须能够成功地适应一个**计算上简单**的环境目标（适应度函数）。

4. 因此，我们可以定义一个特殊的规则集合，这个集合包含了所有那些“**擅长于在简单目标下演化出复杂解决方案**”的规则。这个集合，就是沃尔夫勒姆所说的“规则系综”。

第三步：理论的预测力——“类机械行为”的必然性

这个“规则系综”理论的强大之处在于它的预测能力。沃尔夫勒姆宣称，我们不再需要去研究地球上生物演化的所有偶然细节（比如剪接体这个具体的分子机器是如何出现的），我们只需要知道一个更高层次的抽象事实：**地球上的所有生物机制，都必然是从这个“规则系综”中抽取的样本**。

既然它们都来自同一个特定的规则集合，那么它们就应该会共享某些普遍的、底层的特征。沃尔夫勒姆将这些共同特征称为“**类机械行为**”（Mechanoidal Behavior）。他预测，在生命的微观尺度上，我们会发现大量看起来像是精心设计的“小机器”或“小机制”在运作。这种“机械性”并非偶然，而是源于那些能够成功适应简单目标的规则所固有的性质。

这个理论框架的雄心是巨大的。它试图将生物学从一门很大程度上依赖于描述和历史叙事的学科，转变为一门像物理学一样，可以从第一性原理出发进行预测的硬科学。它认为，生命的许多核心特征，比如模块化、鲁棒性、以及各种分子机器的存在，可能都不是地球生命史的偶然产物，而是任何一个遵循“计算上简单的演化目标”的系统所必然会涌现的普遍规律。这为寻找地外生命，甚至理解人工生命和人工智能的未来发展，提供了一个全新的理论视角。

第三部分：人工智能、数学与知识的版图

在这一部分，访谈转向了当前最热门的话题之一：人工智能（AI）。沃尔夫勒姆清晰地剖析了他眼中AI的两种不同范式，并借此引申出他对数学本质的独特看法。他认为，数学并非客观真理的抽象王国，而更像是一种人类特有的、带有艺术性的创造活动。

3.1 两种AI辅助发现的模式

【访谈原文翻译】

>> 主持人：嗯……让我把你拉回到你最近在做的那些实验。这大概是用你的……呃……一维元胞自动机做的吧？

>> 斯蒂芬·沃尔夫勒姆：碰巧是。是的。

>> 主持人：是的。是的。

>> 斯蒂芬·沃尔夫勒姆：我也用二维的、图灵机之类的试过。没什么区别。

>> 主持人：我们对AI辅助发现非常感兴趣。你进来的时候见到了杨教授。他……他在数学和理论物理领域做AI辅助发现。还有你见过的伊夫·肯尼。他……他探索可积模型的AI辅助发现。嗯……所以你从第一天起就参与其中了，对吧？你能给我们简要地讲讲你是如何……

>> 斯蒂芬·沃尔夫勒姆：嗯，我认为自己过去40年都活在“AI之梦”里。好吧。什么是AI之梦？在我看来，AI之梦就是，你知道，我有一件想做的事，让我尽可能自动地完成

它。所以我一直在构建工具，让它达到这样一个程度，无论是工具还是……我想说……管理结构，比如公司之类的，都能让从“我有一个想法”到“这个想法被实现”的过程尽可能高效。

所以……你知道……我构建的工具不是神经网络。它们……你知道……我构建的主要工具是Wolfram Language，其核心思想是提供一种方式……一种符号表示，用来计算性地思考世界。就像……你知道……500年前的数学符号让谈论数学类的事情变得流畅一样，我们的目标是让谈论计算类的事情变得流畅。

而这对我来说，最强大的方法论就是使用这一整套工具，然后向外望向**计算宇宙**（computational universe），在数以万亿计的程序中进行搜索，看看……你知道……外面有什么。而外面的东西非常令人惊讶。现在，这非常不同。你知道，原始计算宇宙所揭示的东西非常令人惊讶，非常**非人类**，非常……嗯……违背我们的直觉。

神经网络和那种传统则有点不同。因为……你知道……大语言模型（LLM）在做的事情，可以这么说，在第一近似下，是**碾碎人类的知识**。当然，你可以用比这更多的东西来训练它们。但在第一近似下，它就是碾碎人类知识，然后给我们反馈这种经过“人性化”处理的东西。

举个例子，如果你想发现数学定理，你知道，我研究过一个问题，就是数学的长期极限是什么。你知道，在数学文献中，有三四百万个定理被写了下来。当我们拥有一万亿个或一千万亿个定理时会发生什么？数学的连续极限是什么？当我们……当我们推导出所有定理时，关于那个结构，其实有很多可以说的。但问题是，当你开始……你知道……枚举可能的定理时，你枚举出的大多数定理，虽然是完全有效的，但它们是那种典型的**人类只会耸耸肩的定理**。你知道，就像……我们为什么要在乎这个？你知道，重要的是，这些定理中哪些是人类在乎的。

而我认为，当前这一代的LLM等等……我还没能让这个方法成功，但原则上，可以想象，那是一个它们能派上用场的地方，就是说：“嘿，外面有所有这些你可以抽象地找到的东西，但哪些东西可能是我们人类真正在乎的？”

【深度解读】

沃尔夫勒姆在这里清晰地划分了两种截然不同的人工智能（AI）范式。这不仅是对当前AI技术格局的深刻洞察，也揭示了他自己40年来工作的核心哲学。

范式一：以LLM为代表的“人类知识研磨机”

- **核心原理**：这种AI（如GPT-4）通过学习海量的人类创造的数据（文本、图片、代码）来工作。它的本质是一个强大的**模式匹配和统计推断**引擎。
- **工作方式**：它将人类的语言和知识“碾碎”成高维向量空间中的数学表示，学习其中的关联和结构，然后再根据这些学习到的模式生成新的、看起来符合人类逻辑和创造力的内容。
- **产出特征**：它的产出本质上是**对人类已有知识的内插、外推和重组**。它非常擅长模仿人类的风格、总结人类的观点、在人类已有的概念框架内进行推理。它是一个卓越的“人类助理”。
- **局限性**：它的“视野”被其训练数据所限制。它很难凭空创造出完全超越人类现有概念框架的东西。它探索的是“人类知识”这个已知的岛屿。

范式二：沃尔夫勒姆的“计算宇宙探险家”

- **核心原理**：这种AI范式并非学习人类数据，而是直接**探索所有可能计算规则的抽象空间**。
- **工作方式**：它通过系统性地枚举和运行简单的程序（如元胞自动机、图灵机等），观察它们从最简单的初始条件出发会产生什么样的行为。这是一种不带任何偏见的、纯粹的经验性探索。
- **产出特征**：它的发现往往是“**非人类的**”、“**违背直觉的**”。它能揭示出宇宙固有的、不依赖于人类观察和理解的计算结构。例如，元胞自动机“规则30”产生的复杂性，就是人类在探索之前完全无法想象的。
- **局限性**：它发现的大多数东西对人类来说可能是“**无意义的**”或“**无法理解的**”。沃尔夫勒姆将其描述为“人类只会耸耸肩的定理”。这些发现在数学上是正确的，但在人类构建的知识体系中找不到位置。

为了更清晰地理解这两种范式的区别，我们可以构建一个表格：

表1：两种AI辅助发现的模式对比

特征	LLM驱动的AI (人类知识研磨机)	沃尔夫勒姆的计算探索 (计算宇宙探险家)
核心方法	从现有数据中学习统计模式	枚举和运行简单程序以发现其内在行为
数据来源	人类知识的总和（文本、图像、代码等）	所有可能计算规则的抽象空间（“Ruliad”）
发现的性质	**以人类为中心**：在人类知识框架内发现新的联系和模式。	**“外星”或“非人类”**：发现独立于人类认知的、宇宙固有的计算结构。
角色定位	强大的**“助理”**，帮助人类整理、加速和扩展已有的知识。	纯粹的**“发现者”**，揭示全新的、可能难以理解的现象。
面临的挑战	如何超越训练数据的偏见和局限？	如何从海量的“无意义”发现中，筛选出对人类有价值和可理解的部分？

沃尔夫勒姆最后提出的一个有趣观点是，这两种AI范式或许可以结合。他的“计算宇宙探险家”可以发现无数“外星”般的真理，而LLM这种“人类知识研磨机”则可以凭借其对人类兴趣和价值观的理解，帮助我们从中筛选出那些“**我们人类可能真正在乎的**”部分。这为未来AI与科学的结合提供了一个富有想象力的图景：一个负责探索未知，一个负责翻译和诠释。

3.2 数学：一种人类的艺术创作

【访谈原文翻译】

>> 主持人：那么，这是否就是人类数学家的角色？你还认为人类数学家有存在的价值吗？因为……你知道……现在只要你上社交媒体，你可能不上，这很明智，但如果你碰

巧上了，年轻的数学家们都非常担心数学会因为AI而终结。

>> 斯蒂芬·沃尔夫勒姆：我认为这涉及到数学在根本上是什么的问题。好的，所以……你知道……在19世纪末出现的观点是，数学是这种形式化的东西，是可以机械推导的。希尔伯特有这个想法，我们只需写下数学的公理，然后用这台机器去研磨，就能得出所有正确的数学定理。哥德尔定理在某种程度上摧毁了这个想法。但事实是，哥德尔定理摧毁了这个想法，而大多数在职的数学家却不关心哥德尔定理，这件事本身就说明了一些问题。

换句话说，这种仅仅在公理层面，一步一步向前推进的数学，并不是数学家通常所做的事情。这里有一个类比，实际上与我们刚才谈到的热力学有关。你知道，在热力学中，有分子动力学这个层面，有所有这些分子在四处碰撞。我们可以说……你知道……为了知道气体的行为，我们必须去观察每个分子的行为。这是一种可能性。但事实是，我们知道我们不必这样做。我们知道我们可以在流体力学的层面上进行研究，只讨论……你知道……速度场之类的东西，而不是每个分子具体在做什么。

问题是，数学真正在做什么？数学是在这个公理层面上运作的东西吗？还是说，人类所做的数学，是某种……更像是在流体力学这个层面上运作的东西？我认为，一件有趣的事情是，更高层次的数学是可能的。它可能不是……流体力学可能是不可能的。可能情况是，了解流体将如何运动的唯一方法，是下降到底层，计算出每个分子的运动。但事实上，高层次的流体力学是可能的。同样的事情也适用于我们的物理模型，适用于时空结构等等。

在数学中，**高层次数学之所以可能，是另一种这样的现象**。举个例子，如果你拿毕达哥拉斯定理（勾股定理）来说，在一个典型的证明辅助系统中，如果你想在公理的基础上（比如集合论公理）证明毕达哥拉斯定理，你会得到一个包含……你知道……7000个引理之类的东西，一个巨大而复杂的烂摊子，对吧？但事实是，没人在乎。他们只想讨论毕达哥拉斯定理本身。顺便说一句，有很多不同的包含7000个引理的集合，可以不同地证明毕达哥拉斯定理。

但是，数学家们在他们通常实践数学时，是在这种**人性化的、流体力学的层面**上运作，而不是在这种公理层面上。而且我认为……你知道……嗯……我想说……你知道……自动化定理证明是“让我们自动发现数学”的早期迭代。我相信，我25年前发现的一个结果，至今仍然是通过自动化定理证明在数学中找到的、而事先不被认为是正确的唯一例子。

我当时对布尔代数最简单的公理系统感兴趣。人们花了大约50年的时间，试图……你知道……精简出最简单的东西，你知道，$x \vee y = y \vee x$等等等等。布尔代数最简单的基础是什么？我当时猜测可能有一个非常简单的。于是我开始搜索，然后我找到了这个小小的东西，它是一个单一的公理，里面有六个“与非”（NAND）操作，它能推导出整个布尔代数。我是通过使用自动化定理证明找到它的。我应该说，今年早些时候，我再次尝试用AI来理解那个证明。我和其他所有人一样，完全失败了。

>> 主持人：但你是通过枚举得到这个的。

>> 斯蒂芬·沃尔夫勒姆：完全枚举所有……

>> 斯蒂芬·沃尔夫勒姆：我是通过枚举找到了这个**待证的定理**。然后我用自动化定理证明来**证明**了它。好的。

>> 主持人：那个证明大约有120个引理长。

>> 斯蒂芬·沃尔夫勒姆：而且它**无法理解**。对。我的意思是，我尝试过……你知道……我用最好的AI工具和可能一点点小聪明，做出了相当认真的努力，试图解码这个证明到底在搞什么鬼。但它就像……它就是……你知道……你可以验证这些步骤。验证步骤很容易，但它无法理解这是什么。它没有路标。它不像你……你知道……它不像证明进行到一个定理，然后你说："哦，这是高斯证明过的东西。" 它存在于那个……元数学空间的"**外星世界**"里。

>> 主持人：什么……对不起，我没跟上。那么，这告诉我们人类数学家的角色是什么呢？这是……我们是……

>> 斯蒂芬·沃尔夫勒姆：嗯，我认为**人类的数学更像是一种艺术性的努力**（artistic endeavor）。我的意思是，数学是……你知道……你选择做什么，以及你在一个人性化的层面上做它，可以这么说。当你谈论证明时……你知道……那个自动化……用现代语言来说，25年前为这个东西找到的等式证明，它真的是一个有用的证明吗？它不是一个能向我这个人类**解释**它为什么为真的证明。它是一个**验证**它为真的证明。嗯……也许我可以基于这个事实继续构建，但它不是一个……阐述任何东西的证明。

我认为……你知道……就……你知道……元数学空间是无限的，非常无限，在数学空间里你可以走向许多不同的方向，你可以证明许多不同种类的定理，你可以发明许多数学领域。**你选择发明哪些，这是一个非常人性化的事情**。对吧？这就像问，对于人类语言，我们有……典型语言中有5万个词汇。我们发明了某些词，因为我们发现它们对于我们倾向于思考的那些事情很有用。你可以想象发明数百万个其他的词，实际上……你知道……AI在内部已经这样做了……你知道……数百万个……实际上是事物的"词汇"。但我们……你知道……我们用我们的人类文明，填充了这个"概念空间"的某些部分。

你知道，如果你开始生成……这里有一个有趣的实验。你用一个生成式AI图像系统，你说："生成一张猫的图片。" 它会照做。然后你说："让我改变嵌入空间中的数字，制造一个不完全是猫的东西。" 它离猫有点远。有一个"猫岛"，那里的图片或多或少看起来像猫。离开"猫岛"，你就进入了我称之为"**概念间空间**"（interconcept space）的地方。在这个空间里，有图片，你可以看着它们，但它们不是我们有词汇来描述的东西。它们不是我们用人类概念填充过的地方。而这实际上有点令人谦卑，即使是一个非常简单的生成式AI系统，概念间空间所占的比例也绝对是压倒性的。就像……你知道……只有 10^{600} 分之一左右的东西，是我们给过名字的。

数学也是一样。有一个巨大的概念间空间，里面充满了我们人类从未决定要去关心的东西。而那是一个……你知道……那是一个人类的选择，我们选择走向哪个方向，就像在人类语言中发明词汇一样。

【深度解读】

面对“AI是否会终结数学”的焦虑，沃尔夫勒姆给出了一个深刻而富有诗意的回答：**人类数学的本质并非机械的定理证明，而是一种艺术性的、人性化的创造活动**。他通过几个层次的论证来支撑这个观点。

第一层：数学的两个层面——“分子动力学” vs “流体力学”

沃尔夫勒姆首先区分了两种不同层次的数学：

- **公理层面（“分子动力学”）**：这是数学最底层的逻辑结构。从少数几个公理（如集合论公理）出发，通过严格的逻辑推理，一步一步地构建出整个数学大厦。这个过程是机械的、可以被计算机执行的。然而，其产物往往是极其冗长和“无法理解”的。他举了两个例子：
  1. 用自动化定理证明系统从基本公理证明勾股定理，会产生一个包含7000个引理的“烂摊子”。
  2. 他自己发现的、仅用一个公理就能定义布尔代数的证明，包含了120个引理，即使借助最强的AI也无法理解其“思想”。这个证明存在于一个“外星世界”，它能**验证**（verify）真理，但无法向人类**解释**（explain）真理。
- **人类层面（“流体力学”）**：这是绝大多数数学家实际工作的层面。他们并不直接与底层的公理打交道，而是使用更高层次的概念、直觉和审美判断（如“优雅”、“深刻”）。他们像是在研究流体的宏观行为，而不是追踪每一个水分子的轨迹。他们关注的是那些能够被人类心智所理解、能够讲述一个连贯“故事”的数学结构。

这个类比揭示了为什么哥德尔定理（它指出了公理层面的局限性）对大多数数学家的日常工作影响不大，因为他们根本就不在那个层面上操作。

第二层：数学作为一种“选择”的艺术

基于上述区分，沃尔夫勒姆的核心论点浮出水面：在无限广阔的、所有可能定理组成的“元数学空间”中，人类数学家所做的工作，本质上是一种**选择**。

- **概念间空间（Interconcept Space）**：他用生成式AI的例子生动地说明了这个概念。在AI生成的图像空间中，只有极小的一部分（他估计是 10^{600} 分之一）对应着我们人类有命名、有概念的东西（如“猫”、“狗”、“桌子”）。绝大部分空间是“概念间的”，充满了我们无法识别、无法命名的“怪物”。
- **数学的“概念岛屿”**：数学也是如此。所有可能被证明的定理构成了一个汪洋大海，而人类数学家们在过去的几千年里，只是在这片大海中选择并建立了一些我们认为有意义、有关联、有美感的“概念岛屿”（如数论、几何、拓扑）。
- **选择即创造**：这种选择的过程，就像在语言中发明新词汇一样，是一种**人性化的创造活动**。我们选择研究那些与我们心智结构相容、与我们从物理世界中获得的直觉相符、能够相互连接形成优美理论的数学领域。因此，数学与其说是一种对柏拉图式客观真理的“发现”，不如说是一种人类心智在抽象可能性空间中的“**艺术创作**”。

对“AI终结数学”的回应

在这个框架下，“AI终结数学”的担忧就显得多余了。

- AI（特别是自动化定理证明器）非常擅长在“公理层面”进行探索，它们可能会发现无数人类无法理解的、“外星”般的定理。
- 但是，**决定哪些定理是“有趣的”，哪些数学方向是“值得探索的”**，这个充满审美和价值判断的过程，仍然是人类数学家的核心角色。
- AI不会终结数学，但它可能会改变数学家的工作方式。AI可以成为一个强大的“外星探险家”，带回来自“概念间空间”的奇珍异宝，而人类数学家的任务，则是去理解、诠释这些发现，并将它们编织进人类的知识挂毯中。

最终，沃尔夫勒姆将数学家从一个“真理的矿工”重新定位为一个“**意义的建筑师**”。这不仅为人类在AI时代的角色提供了辩护，也为数学这门古老学科赋予了更深的人文主义色彩。

3.3 物理学、数学与可理解的宇宙

【访谈原文翻译】

>> 主持人：我注意到你……你给我看了一篇你写的关于物理学如何为数学提供信息的文章。

>> 斯蒂芬·沃尔夫勒姆：是的，你对此有一些看法。

>> 主持人：嗯，是的。因为……因为有一个问题，好吧，首先要理解什么是科学？好的。你知道，自然世界做它该做的事，它在进行所有这些计算等等。而我们，另一方面，我们的心智只能做整个自然世界所能做的计算的一小部分。所以我看待科学的方式……你知道……什么是科学？我认为它就像一座桥梁，连接着世界上实际发生的事情和我们能用我们的心智理解的叙事。所以，**科学是从自然世界中提取出那一小段能装进人类心智的叙事线索的方式**。

所以……你知道……当你问，嗯，什么……你知道……能装进人类心智的是什么，有些事情我们可以谈论，因为它们……它们符合我们心智的结构。就数学是实现这一目标的载体而言，这更多地说明了我们迄今为止已经建立起思考方式的那些事物的类型，可以这么说。换句话说，物理学本身并非内在地是数学的。而就我们认为我们现在实际知道物理学的“机器代码”是什么样子而言，我们可以说，这就是机器代码的样子。在那机器代码之上，有很多东西……有很多不同的特定……好的，所以从这个机器代码出发，宇宙中实际发生着这种不可约的计算。我们从中取出一些小的切片，一些**可约的**（reducible）切片，那些我们能理解的。所以当我们……

>> 主持人：什么是“可计算上可约的口袋”？

>> 斯蒂芬·沃尔夫勒姆：是的。是的。所以我的意思是，这个想法有点……

>> 主持人：是的。对。

>> 斯蒂芬·沃尔夫勒姆：所以我的意思是，每当你有一个这样的计算不可约过程，要想详细了解各处会发生什么，是需要这种不可约量的计算工作的。但是，在任何计算不可约的系统内部，实际上**不可避免地会存在计算可约性的口袋**（pockets of computational

reducibility)，那些你实际上可以言之有物的地方。事实上，这样的口袋必须有无限多个。

>> 主持人：你的意思是，比如素数，我们没有有效的方法来得到所有的素数，但我们或许能说一些关于素数分布的规律。那是一个……

>> 斯蒂芬·沃尔夫勒姆：不可约的口袋。

>> 主持人：可约的口袋。

>> 斯蒂芬·沃尔夫勒姆：对，那是一个可约的口袋。而存在无限多个可约口袋这个事实，就是为什么我们永远不会耗尽发明。我们总是能在物理世界中找到一些小地方，在那里我们可以取得进展。在数学中，我们也永远不会耗尽……可以建立的定理等等，那些我们可以取得进展的地方。

>> 主持人：那么这与……我应该解释一下，几个月前我写了一篇文章，关于为什么物理学和物理直觉似乎能引出非常有趣的数学。另一个问题，那个反过来的问题，已经被讨论了几十年了。你知道，为什么数学在给予我们物理学、在描述物理学、在成为物理学的语言方面如此出色？那个问题已经被讨论得烂了。但我把那个问题反过来看，而那个问题也至少在一些科学圈子里被讨论过，但我没在更广泛的媒体上看到过讨论。所以……嗯……你是说这与我们的物理直觉有关吗？呃……那一点点……是它催生了我们觉得有趣的数学问题吗？

>> 斯蒂芬·沃尔夫勒姆：对。我的意思是，所以第一个问题是，你知道，物理世界做它该做的事，我们只感知到它的某些方面，对吧？这是第一件事。所以有很多事情在发生……你知道……无论是微小的量子涨落还是别的什么，我们只是举手投降说：“这对我们来说看起来是随机的。” 嗯……那……你知道……但有些事情我们可以做出陈述。

比如，牛顿真的很幸运，他研究的是刚体力学，他研究的是……你知道……我们有……他没有去做，比如说，流体力学。如果他不是试图用……你知道……固体物体的运动来建立力学定律，而是试图建立流体力学定律，那就会包含流体湍流之类的东西，他就不可能提出任何简单的、可约的关于世界如何运作的陈述。你知道，但他最终得到了世界的一个切片，一个可以做出相当简单的……叙事性陈述的切片，对吧？

我认为，这就是……你知道……在某种意义上，为什么……你知道……嗯……数学……就数学捕捉了某种……你知道……可约性而言，我们所做的数学捕捉了某种……可约性，那是我们可以用来理解世界上发生的事情的切片的东西。它不是世界上发生的全部事情。它是世界上发生的事情中，我们有能力理解的那一部分，因此我们用它来构建技术，比如说。

【深度解读】

这一段对话触及了一个古老而深刻的科学哲学问题，即“数学在物理学中不可思议的有效性”。但沃尔夫勒姆和访谈者巧妙地将其反转，探讨了“为什么物理学能激发有趣的数学”。沃尔夫勒姆的回答，是他整个计算世界观的逻辑延伸，为这个问题提供了一个全新的、统一的解释。

第一步：重新定义“科学”——在不可约的海洋中寻找可约的岛屿

沃尔夫勒姆首先给出了一个他自己对“科学”的定义，这个定义是理解后续所有论证的关键。

- **宇宙的现实**：宇宙本身在进行着海量的、计算不可约的复杂计算。这是世界的本来面目，一个充满复杂性和不可预测性的“计算海洋”。
- **人类心智的局限**：我们的大脑，作为宇宙的一部分，其计算能力是极其有限的。我们无法处理和理解宇宙的全部复杂性。
- **科学的使命**：科学，就是在这片不可约的计算海洋中，寻找并打捞出那些为数不多的、**计算上可约的**（computationally reducible）“小岛”。这些“可约的口袋”（pockets of reducibility）是宇宙中那些行为简单、有规律、可预测的部分。科学的本质，就是“**从自然世界中提取出那一小段能装进人类心智的叙事线索**”。

这个定义有一个深刻的推论：**科学之所以可能，不是因为宇宙本身是简单的，而是因为在极其复杂的宇宙中，必然存在着无限多个我们可以理解的、简单的“口袋”**。这保证了科学探索永无止境。

第二步：解释“数学的有效性”——一个选择效应

现在，我们可以回答那个经典问题了：“为什么数学能如此完美地描述物理世界？” 沃尔夫勒姆的答案是，这并非什么宇宙的奇迹，而是一个**双重选择效应**的结果。

1. **物理学的选择**：物理学家（如牛顿）并非研究了宇宙的全部。他们之所以能成功地建立起简洁的物理定律（如牛顿运动定律），是因为他们幸运地选择了那些**恰好是计算可约的**物理现象来研究（如刚体力学、天体轨道）。如果牛顿一开始就试图为计算不可约的现象（如流体湍流）建立简单定律，他将会一败涂地。我们称之为“物理定律”的东西，只是对宇宙中那些罕见的可约“口袋”的描述。
2. **数学的选择**：正如上一节所讨论的，人类数学家也并非研究了所有可能的数学结构。他们“艺术性地”选择了那些符合人类直觉、结构优美、易于理解的数学分支来发展。这些数学分支，其内在特性恰恰就是“**可约性**”和“**结构性**”。

将这两点结合起来，结论就显而易见了：**我们用我们选择发展的、擅长描述简单结构的数学，去描述我们选择研究的、宇宙中那些恰好行为简单的部分，发现它们彼此匹配，这并不奇怪。** 这就像我们特意制造了一把方形的钥匙（数学），然后在一大堆锁孔中找到了一个方形的锁孔（物理学的可约部分），然后惊叹于钥匙和锁孔的完美匹配。

第三步：回答反向问题——“物理学为何能激发数学”

有了以上基础，回答这个反向问题就变得水到渠成。 物理世界是我们接触和感知现实的主要来源。当我们观察物理世界时，我们的直觉会自然而然地被那些“可约的口袋”所吸引，因为只有这些部分是我们可以理解和把握的。 因此，物理直觉引导我们去关注那些具有简单、重复、对称等特性的现象。而这些特性，恰恰是人类数学家认为“有趣”、“深刻”和“值得研究”的数学问题的源泉。物理学就像一个向导，它帮助我们在无限的抽象可能性中，指明了那些既存在于现实世界，又能与我们心智结构产生共鸣的数学方向。它为数学这门“艺术创作”提供了最丰富的灵感和素材。

总而言之，沃尔夫勒姆认为，物理学和数学之间看似“不可思议”的和谐关系，源于它们都是人类有限心智在面对无限复杂宇宙时，共同聚焦于“计算可约性”这一狭窄窗口的必然结果。它们是同一枚硬币的两面，共同构成了我们能够理解的那个“科学世界”。

3.4 发现的机器与科学的未来

【访谈原文翻译】

>> 斯蒂芬·沃尔夫勒姆：当我们建造技术时，大多数时候，至少在现代之前，我们最终……实际上……

>> 主持人：我能……我能……说到牛顿，我能插句话吗？是的。

>> 斯蒂芬·沃尔夫勒姆：嗯……我其实想先说一件关于技术的事，这是一个有趣的事实。你知道，后工业革命时代的人们期望能理解技术，他们期望能看到齿轮，看到它是如何工作的等等。在那之前，你得到一头驴，驴会做你想让它做的事，你并不会真的去问驴是如何工作的。我们现在又回到了更接近那种情况的境地，可以这么说。而且……嗯……我们以前也经历过，但无论如何，请讲。

>> 主持人：所以，我夏天在为英国政府写一份关于AI辅助发现的基金申请。其中一件事，这是在我们知道要见斯蒂芬之前，我们写道：你知道，有趣的是，我们发现物理学和数学的方式与牛顿当年非常相似。你知道，还是纸和笔。也许有两件事改变了这一点，一个是计算，易于获取的计算，它给了我们更多数据；另一个是符号操作，也就是Mathematica。嗯……但你知道，也许这两样东西各自使我们的科学发现速度翻了一番，但看起来……而且，确实，当你观察科学发现实践中是如何完成的，它主要还是由大学完成的。你知道，我们……我们这个组织感兴趣的事情之一是，你知道，我们如何……我们如何能加速发现？

我的意思是，如果你看看……如果你看看……你知道……职业……运动员在过去75年里，基本上从业余走向了专业，不断推动人类成就的极限。你……你已经……我感觉你已经从学术角度、从科技领导者……你知道……科技企业家的角度，研究了多种关于如何做发现的思考方式，而且我认为你在纯科学和技术之间来回切换。如果你要创造一台“**发现机器**”（discovery machine），它现在会是什么样子？

>> 斯蒂芬·沃尔夫勒姆：是的。我的意思是，看，我很幸运，或者说这实际上是计划好的，我一直在技术开发和基础科学之间交替进行。这是一个非常有用的循环，因为，你知道，基础科学向你展示了你在技术上能做什么，而技术为你构建了让你能做基础科学的工具。但我必须说，在发现方面，对我来说，**影响发现数量的绝对主导因素是战略**（strategy），即你试图发现什么？你问的是什么问题？这远比如何做的具体方法重要。

现在，你知道，我碰巧花了很多时间试图优化如何做的具体方法，而且我觉得我们在这方面已经取得了很大进展。现在，问题……你知道……所以这是经常出错的地方。例如，在大学里，它们是无管理组织。你知道，你是个教授，从来没人告诉你该做什么。如果他们这么做了，你会非常不高兴。嗯……但事实是，人们……你知道……花费数十年……你知道……他们只是在重复他们论文里做过的事情，但结果并不太好等等。

我记得多年前，我曾在贝尔实验室做顾问，那里有一个……稍微更好一点，更……你知道……他们有一种使命感，那种……你知道……至少我看到的部分不是那么强硬，但仍然存在一种“我们试图实现什么”的流向。所以我认为……你知道……有一些方面，关于你如何做出重大的……你知道……你必须……**你所问问题的战略，我认为是最重要的**

事情之一，这也是为什么我喜欢你的那个问题清单，因为我觉得它们相当不错。嗯……而且我认为……嗯……你知道……不知何g……

我之前提到过，关于领域的根基……你知道……人们通常避免研究领域的根基，因为他们说：“哦，那100年前就建立了，没什么可做的了。” 事实是，如果你能在领域的根基上取得进展，那将是**极高杠杆**的。我的意思是，那是……而且通常你能取得进展，不仅仅是因为很久没人看过它了，还因为现在存在的方法与以前存在的方法真的非常不同。

但是，你知道，就科学的制度结构之类的事情而言……我的意思是，其中一件事是，你知道，科学……你知道……人们应该小心自己对科学的期望，因为……你知道……你有一个领域，它很小，很年轻，充满创业精神，很多事情在发生。然后它成功了，然后它变得庞大，你建立了这些巨大的制度结构，然后……然后新事物就很难发生了，因为人们会说：“哦，那不符合……我们正在思考的路线，你知道……那不可能获得资助等等等等。” 所以这很复杂。

我确实思考过……当然……关于基础科学，以及应该如何做基础科学，应该如何为基础科学买单，以及为什么要为基础科学买单等等。我很幸运，我做的理论科学并不那么昂贵，我一直……你知道……用我科技生涯的资源来支持它。现在我们有一个独立的沃尔夫勒姆研究所，资源多一些，但……你知道……这个问题……我的意思是，总有一个问题，为什么社会要支持基础科学？这是一个……一个根本性的问题。

我会说……我注意到，每当某个实体对基础科学如何被使用拥有垄断地位时，它就有理由支持基础科学。例如，你知道，鼎盛时期的贝尔实验室……你知道……各种各样的创新……贝尔实验室将是主要的……主要的受益者。美国政府，作为经济体，有很多种事情，如果取得了进展，美国政府将会受益。你知道，这要……我知道，在我们世界的小角落里，在各种数学计算中，我们做基础研究，因为我们很清楚我们拥有利用它的分销渠道。这使得……投资于纯基础研究在那些领域是理性的。但是……你知道……更普遍地，至少对我来说，如何让这件事运作起来，远没有那么清晰。

说到牛顿，我发现一个有趣的思想实验是，如果你是1687年的老艾萨克，你发明了微积分，你确信在接下来的几百年里，微积分将产生数万亿美元的价值。你能做什么来……你知道……你是否创造……你知道……“微积分币”之类的东西来……你知道……试图……嗯……做些什么来加速那个价值的实现。我不确定。

【深度解读】

面对“如何加速科学发现”以及如何构建一台“发现机器”的宏大问题，沃尔夫勒姆的回答出人意料地将重点从“工具”转向了“**战略**”。他认为，在科学发现中，**问对问题远比拥有先进的工具更重要**。

科学发现的核心驱动力：战略 > 方法

沃尔夫勒姆指出，尽管他本人花费了半生精力来打造最高效的科学工具（如Mathematica和Wolfram Language），但他深知，这些工具本身并不能保证重大发现的产生。真正的瓶颈在于**战略选择**。

- **问题的选择**：“你试图发现什么？你问的是什么问题？”这被他视为“绝对主导因素”。一个平庸的问题，即使有最强大的工具，也只能产生平庸的答案。而一个深刻的、根本性的问题，则有可能开启一个全新的领域。
- **大学的“无管理”困境**：他尖锐地批评了现代大学的科研模式。教授们拥有极大的自由度，但这种“无管理”的状态也导致了战略上的漂移。许多科研人员倾向于在自己博士论文的框架内进行修补和延续，而不是去挑战更宏大、更具风险的根本性问题。
- **贝尔实验室的启示**：他以鼎盛时期的贝尔实验室为例，说明一种带有“使命感”和战略方向的科研环境，可能比完全自由放任的环境更能催生重大创新。

高杠杆战略：攻击被遗忘的基础

沃尔夫勒姆再次强调了他之前提到的“攻击基础”的策略，并将其定义为一种“**极高杠杆**”（incredibly high leverage）的科学投资。

- **为什么杠杆高？** 因为对一个领域的基础进行任何微小的修正或颠覆，其影响都会沿着知识树向上扩散，可能导致整个领域的重构。
- **为什么可行？** 因为这些基础往往被忽视了几十年甚至上百年，而在这期间，我们已经拥有了全新的思想工具（如计算思维）和技术工具（如强大的计算机模拟）。用新武器去攻击旧战场，胜算大大增加。

基础科学的经济学困境

沃尔夫勒姆也坦诚地指出了支持这种高风险、高回报的基础科学研究的现实困境：“**为什么社会要支持基础科学？**”

- **垄断与回报**：他提出了一个敏锐的观察——只有当一个实体能够“垄断”基础科学成果所带来的应用时，它才有足够的动力去进行纯粹的基础研究。例如：
  - **贝尔实验室**：作为曾经的电信巨头，它知道任何通信相关的基础物理学突破，最终都会让它受益。
  - **美国政府**：作为国家经济的管理者，它知道许多基础科学的进步最终会转化为国家整体的经济和军事实力。
  - **Wolfram Research**：在他自己的公司里，因为拥有将数学和计算新思想转化为产品的“分销渠道”（Mathematica），所以投资于相关的基础研究是“理性的”。
- **牛顿的“微积分币”思想实验**：这个思想实验非常巧妙地揭示了基础科学价值实现的长期性和分散性。牛顿发明微积分时，不可能预见到它将在几百年后支撑起整个现代工程和金融体系。他无法通过发行“微积分币”来将未来的巨大价值在当下变现，从而资助自己的研究。这形象地说明了为什么纯粹的基础科学很难在常规的市场经济中获得足够的投资。

总而言之，沃尔夫勒姆心中的“发现机器”，其核心部件不是更快的计算机或更聪明的AI，而是一个能够**持续提出深刻、根本性问题的战略引擎**。他认为，当前的科学体制（特别是大学）在这个核心部件上存在缺陷。而要修复它，则需要解决一个更深层次的经济学和社会学问题：如何建立一种机制，来激励和资助那些回报最高、但风险也最大、周期也最长的“攻击基础”型研究。

第四部分：宏大问题：上帝、自由意志与未来

在访谈的快速问答环节，沃尔夫勒姆被要求用极短的时间回答一些人类思想史上最宏大、最古老的哲学问题。他毫不犹豫地将自己统一的计算框架应用到这些问题上，给出了一系列简洁而极具颠覆性的答案。这部分内容集中展现了他的世界观如何从科学延伸到哲学的广阔领域。

4.1 上帝存在吗？

【访谈原文翻译】

>> 主持人：斯蒂芬·沃尔夫勒姆。

>> 斯蒂芬·沃尔夫勒姆：是的。

>> 主持人：当今科学有什么问题？

>> 斯蒂芬·沃尔夫勒姆：它太庞大了。从事科学的人太多了，这导致它过于制度化，使得创新事物的发生变得太困难。

>> 主持人：你如何解决它？我的意思是，这是同一个问题。没关系，慢慢来。

>> 斯蒂芬·沃尔夫勒姆：我不知道。我的意思是，我真的……我不知道。我认为，它得到解决的地方，是那些因新方法论等而产生的新生事物的“萌芽”之处。而且我认为……你知道……如何最好地支持这一点，是一个非常有趣的问题。我的意思是，你们正在做的事情是一个……一个有趣的想法。你知道，我们正在尝试做的事情，实际上是一个有点类似的想法。这在全球生态系统中如何运作，我还不知道。

>> 主持人：下一个。

>> 主持人：嘿，我真的成功地……

>> 主持人：爱你。你……你做得非常出色。所以，这里有一个非常简单的问题。上帝存在吗？

>> 斯蒂芬·沃尔夫勒姆：这是一个有趣的问题。所以，你知道，我们的物理学项目得出的结论之一是……这真的会很有挑战性。好的。嗯……我有一个……我有一个……好的。当我还是个孩子的时候，我要……我要在这个问题上花点时间。当我还是个孩子的时候，你知道，我在英国长大，那里有……你知道……国教等等，而且……嗯……我总是……你知道……反对所有那些东西。我记得……你知道……关于是否有灵魂的问题。我当时想，如果灵魂是真实的东西，它必须有质量。灵魂的质量是多少？当然，这是个错误的问题，因为现在我们对计算有了更好的理解，我们意识到，存在一种心智中的概念，它不是一个有质量的物理实体。所以我被那个观察适当地上了一课。

好的。所以在我们的物理学项目中，一个大问题是……有点像……我们为什么得到了这个宇宙，而不是另一个？我意识到的是，有这样一个“**Ruliad**”的概念，即所有可能计算的纠缠极限。事实证明，似乎……有点像……Ruliad，这个……所有可能的计算，这

个抽象定义的东西，在某种意义上……有点像……是所有可能存在的一切。而像我们这样的观察者，**不可避免地**会以我们看待物理学的方式来感知Ruliad。

好的。这和上帝有什么关系？好的。所以问题之一是，宇宙为什么存在？是否……你知道……宇宙之外必须有某个东西使它存在？这个Ruliad的整个想法，**让宇宙的存在成为一种必然**。它没有让**我们**的存在成为必然。而这实际上与所有这些关于生物学的问题密切相关。我快要……做到了。但是……呃……你知道……我认为，关于是否需要在宇宙之外有某个东西作为宇宙的原因的这个问题……

>> 主持人：是的。

>> 斯蒂芬·沃尔夫勒姆：我现在开始理解，有一种方式可以看到为什么宇宙必须作为一个必然的事物而存在。宇宙中存在像我们这样的观察者，则没有那么必然。那是一件……要证明存在像我们这样的观察者是可能的，是一个科学问题。然后我们时间到了。

>> 主持人：我们确实时间到了。

>> 斯蒂芬·沃尔夫勒姆：也许我们稍后可以回到上帝是否存在的问题上。嗯……是的。

【深度解读】

沃尔夫勒姆对“上帝是否存在”这个终极问题的回答，是他整个物理学和哲学思想的顶点。他没有简单地回答“是”或“否”，而是重新定义了问题的框架，用一个全新的、极其抽象的概念——“**Ruliad**”——来取代传统上帝的角色。

第一步：理解“Ruliad”是什么

这是沃尔夫勒姆思想体系中最宏大、也最难理解的概念。我们可以分层来理解它：

- **基础是计算**：宇宙万物皆为计算。
- **超越单个宇宙**：我们通常认为的宇宙，是遵循一套特定物理规则的计算过程。但沃尔夫勒姆问，为什么是这套规则，而不是别的规则？
- **所有规则的集合**：Ruliad就是对这个问题的终极回答。它不是一个特定的计算，而是**所有可能的计算规则，以所有可能的方式，从所有可能的初始条件开始运行，所产生的所有结果的集合**。它是一个包罗万象的、抽象的数学对象，是“所有计算可能性”的极限和总和 。
- **一个唯一的、必然的存在**：由于Ruliad包含了**所有**可能性，所以它本身的存在不依赖于任何外部选择或创造。它不是被“设计”出来的，它的存在是一种逻辑上的**必然**。它就是所有可能性的集合体，它不可能不是它所是的样子。

第二步：Ruliad如何取代上帝？

传统神学中，“上帝”通常扮演两个核心角色：

1. **第一因（First Cause）**：回答“为什么世界会存在？”、“谁创造了宇宙？”。
2. **设计者（Designer）**：回答“为什么宇宙的规律是这个样子，而不是别的样子？”。

沃尔夫勒姆的Ruliad概念，同时消解了这两个问题：

- **宇宙的存在是必然的**：对于“为什么有物而非无物？”的问题，沃尔夫勒姆的答案是：因为“物”（所有可能计算的总和，即Ruliad）的存在是逻辑上必然的，它不需要一个外部的“创造者”。宇宙的存在成了一个**无需解释的、必然的事实**。
- **我们的物理定律也是必然的（对于我们而言）**：对于“为什么物理定律是这样？”的问题，他的答案更进一步。他认为，我们所感知的物理定律（相对论、量子力学等），并非Ruliad中一套被“特殊选中”的规则，而是**任何像我们这样的观察者在观察Ruliad时，所必然会感知到的宏观规律**。我们的心智结构、我们对时间连续性的信念、我们的计算局限性，共同“过滤”了Ruliad无限的复杂性，最终呈现出的就是我们所熟知的物理学。

第三步：我们存在的“偶然性”

然而，这个理论有一个非常重要的区分：

- **宇宙（Ruliad）的存在是必然的**。
- **我们的存在是偶然的**。

Ruliad的必然性，并不能保证在它的某个角落里，会演化出像人类这样能够进行自我意识和科学探索的观察者。证明“像我们这样的观察者的存在是可能的”，沃尔夫勒姆认为，这是一个具体的、需要通过类似他生物演化模型来研究的“**科学问题**”，而不是一个哲学上的必然。

结论：一个计算的、非人格化的“上帝”

最终，沃尔夫勒姆的回答可以被理解为一种深刻的**计算泛神论或自然神论**。如果非要有一个“上帝”，那么这个“上帝”就是Ruliad本身——一个抽象的、必然存在的、包罗万象的计算结构。但这个“上帝”没有意识，没有人格，不进行干预。它就是存在本身。这个观点消除了对一个超自然造物主的需求，但同时为宇宙的结构和存在本身注入了一种新的、令人敬畏的数学和逻辑之美。

4.2 我们有自由意志吗？

> **【访谈原文翻译】**
>
> **>> 主持人**：所以，我们有自由意志吗？
>
> **>> 斯蒂芬·沃尔夫勒姆**：好的。关于这个问题我已经思考了很多。我认为，我之前提到的“计算不可约性”这个现象，正是解开自由意志如何运作这个结的关键。在计算不可约性中，存在一个确定的底层规则。但是，要想知道将会发生什么，你却无法绕开让它实际运行、观察其结果的过程。
>
> 所以，对我们来说，没有自由意志的定义性特征将是：你可以从我们外部观察，然后说：“我知道他接下来要说什么了。” 那将是……你知道……“我知道那只飞蛾会做什么，它会撞向窗户扑向灯光。” 它可以说是没有自由意志的。
>
> **自由意志，在某种操作层面上，就是你无法预测这个东西将要做什么**。它的意志并非明显地由这些底层规则所决定。而我认为，这正是计算不可约性的故事。

>> 主持人：是的，答案是肯定的。

>> 主持人：我们有自由意志吗？

>> 斯蒂芬·沃尔夫勒姆：嗯，从这个意义上说，你无法通过比仅仅观察我们行动更有效的方法来预测我们将要做什么。

>> 主持人：是的。

>> 斯蒂芬·沃尔夫勒姆：但是要问是否……从我们外部来看，可以这么说，如果你有一个计算上任意复杂的东西，它能预测将会发生什么吗？那么我认为答案是肯定的。所以这是一个有点微妙的……

>> 主持人：我想这取决于你如何定义自由意志。

>> 斯蒂芬·沃尔夫勒姆：对。

【深度解读】

自由意志问题是哲学史上最古老的辩论之一，它通常被视为**决定论**（Determinism）与**自由**（Freedom）之间不可调和的矛盾。

- **决定论的观点**：宇宙中的每一个事件，包括人类的思想和行为，都是由先前的因果链条完全决定的。如果物理定律是确定的，那么从宇宙大爆炸的那一刻起，你今天会想什么、做什么，都早已注定。
- **自由意志的观点**：人类拥有真正的选择能力，我们是自己行为的最终发起者，不受制于物理因果的必然性。

沃尔夫勒姆利用“**计算不可约性**”（CI）这一概念，巧妙地绕过了这个二元对立，提供了一个相容论（Compatibilism）的现代版本。

沃尔夫勒姆的解决方案：区分“因果决定”与“实践预测”

他的核心论点是，我们必须严格区分两个概念：

1. **因果上的决定性（Causal Determination）**：我们的思想和行为，作为大脑中物理和化学过程的结果，完全遵循底层的物理定律。从这个意义上说，我们是**被决定的**。
2. **实践上的可预测性（Practical Predictability）**：我们是否能**在事前**通过某种计算捷径，准确地预测出这些思想和行为的结果？

沃尔夫勒姆指出，由于我们大脑的运作过程是**计算不可约的**，所以对于第二点，答案是“**否**”。即使我们的行为是完全由底层规则决定的，但要预测这些行为，唯一的办法就是完整地模拟大脑中每一个神经元的每一次放电——这个过程的复杂度和耗时，与我们实际思考和行动所花费的时间是相当的。**不存在预测的“捷径”**。

自由意志的操作性定义

基于此，沃尔夫勒姆为“自由意志”提供了一个全新的、**操作性**的定义： **自由意志，就是我们的行为对于任何外部观察者（包括我们自己）而言，在实践上是不可预测的。**

- **为什么我们感觉自己是自由的？** 因为我们无法“计算”出自己下一秒会想什么。我们只能通过“活在其中”来体验这个思考的过程。这种内在的、对自身未来状态的不可知性，就是我们体验到的“自由”。
- **为什么我们认为他人是自由的？** 因为我们无法通过观察一个人的当前状态，就用一个简单的公式算出他十分钟后会说什么话。我们只能等待，让他“运行”自己的思考过程，然后观察结果。

他用扑灯的飞蛾作为反例。我们可以相当准确地预测飞蛾的行为，因为驱动它行为的算法是计算上“可约的”、简单的。因此我们说飞蛾没有自由意志。而人类大脑的复杂性，使其行为是计算不可约的，因此我们拥有自由意志。

一个微妙的结论

沃尔夫勒姆最后补充了一个重要的限定：这种自由是相对于“计算能力与我们相当”的观察者而言的。如果存在一个计算能力**无限**的“超级观察者”（类似于拉普拉斯妖），它或许能够完整地模拟我们的每一个原子，从而预测我们的行为。但对于宇宙中任何计算能力有限的实体来说，我们的意志是“自由”的。

这个观点巧妙地保留了物理学的决定论基础，同时捍卫了人类自由体验的真实性。它将自由意志从一个神秘的、超自然的属性，转变为一个根植于计算复杂性理论的、可被科学理解的现象。自由不再是“摆脱因果”，而是“**超越预测**”。

4.3 人工智能的未来：AGI的迷思与泡沫的现实

【访谈原文翻译】

>> 主持人：我们是否正处于一场AI市场泡沫之中？

>> 斯蒂芬·沃尔夫勒姆：完全换个话题。

>> 斯蒂芬·沃尔夫勒姆：答案显然是肯定的。好的。我的意思是，你知道，嗯……AI当然是有用的东西。它是一个……你知道……就像这种……你知道……被发现的野生动物。我们刚刚发现了马，或者别的什么，现在我们可以……你知道……现在我们正在建造……你知道……马车之类的东西。很多价值将来自于马车，而不是马本身。我想……我想……嗯……你知道……在……嗯……当市场变得如此泡沫化时，情况很复杂。你知道，那常常会为一些事后的填充创造机会，而这些填充物最终会非常有价值。你知道，即使最初投资的东西本身不是……你知道……只是……引发人们前来投入大量资金的旗帜。

>> 主持人：对，就像互联网泡沫一样。

>> 斯蒂芬·沃尔夫勒姆：我的意思是，它……你知道……有点……是的。

>> 主持人：对。对。嗯，好的。在AI泡沫这个话题上，通用人工智能（AGI）还有多远？

>> 主持人：而且……而且……你害怕它吗？

>> 斯蒂芬·沃尔夫勒姆：我本来要……

>> 主持人：你害怕它吗？我本来要单独提问的。这是一个额外问题。

>> 斯蒂芬·沃尔夫勒姆：不，我的意思是，我认为这有点……一个无意义的东西，因为你说，嗯，它到底意味着什么？嗯，它意味着像人一样智能的东西。然后你说，嗯，你知道，在我从事计算和类似AI领域的这段时间里，有很多次人们都说，当AI能做XYZ时，比如做符号积分，那么我们就会知道它们是智能的。很快，它们就能做符号积分了，然后每个人都说："这只是工程"，事实也确实如此。

你知道，这是一个反复出现的事情。而你真正说它……你知道……"它做的就像我们一样，但它是一台机器"的时候，是……你知道……我们做的每件事都只有我们能做。你基本上必须制造一个……你知道……人类的完整复制品，可以这么说，才能得到……所有的一切。所以我认为……我不认为它真的……我认为这是一个不怎么说得通的概念。

而且我认为……嗯……不，那……我害怕它吗？不，因为我不认为这是一个说得通的概念。我的意思是，我认为……你知道……这将会很有趣，因为……你知道……人们思考……你知道……外星智能，而我总是想，嗯，那……你知道……天气也有它自己的想法。在我对计算如何运作的看法中，天气在进行各种各样的计算，与我们大脑中的计算没有太大不同，只是它恰好与我们思考事物的方式不一致。

我的猜测是，人们开始接受……你知道……自然世界充满智能这个事实的时间，大概会和他们开始接受……你知道……存在……你知道……在不同地方有这些不同形式的智能，它们或多或少地与我们人类的运作方式一致这个事实的时间差不多。

【深度解读】

沃尔夫勒姆对当前AI热潮的看法，体现了他一贯的、基于第一性原理的深刻洞察力。他既承认AI的巨大价值，又对其中最流行的概念——通用人工智能（AGI）——提出了根本性的批判。

关于AI泡沫：价值在"马车"，而非"马"

他对"AI市场是否存在泡沫"的回答是"显然是的"。但他紧接着用了一个非常精妙的比喻来阐明他的观点：**AI模型本身是"马"，而真正的长期价值在于围绕它建造的"马车"**。

- **"马"**：指的是底层的、核心的AI技术，比如大型语言模型本身。发现"马"是一次性的、革命性的事件，它带来了新的动力。
- **"马车"**：指的是利用这种新动力所构建的无数具体应用、产品、工作流程和商业模式。这部分是渐进的、多样化的，并且是价值创造的主要来源。

这个比喻的历史参照是互联网泡沫。在泡沫时期，人们疯狂投资于"互联网"这个概念本身（相当于"马"），但许多公司都失败了。然而，泡沫破裂后，建立在互联网基础设施之上的、真正解决实际问题的应用（如电子商务、社交媒体，相当于"马车"）蓬勃发展，创造了巨大的、持久的价值。沃尔夫勒姆预测，当前的AI热潮也会遵循类似的路径。

对通用人工智能（AGI）的批判：一个"无意义的概念"

沃尔夫勒姆对AGI的看法尤其具有颠覆性。他认为AGI是一个“**移动的球门**”，一个定义不清、因此也无法实现的伪概念。

- **历史的教训**：他指出，历史上，人们曾认为许多任务是“智能”的标志，如符号积分、下棋等。然而，一旦计算机掌握了这些技能，人们便不再称之为“智能”，而将其重新归类为“**只是工程**”。这个过程会无限重复下去。
- **“智能”的定义问题**：AG-I的目标是达到“人类水平的智能”。但“人类水平”是一个包罗万象、模糊不清的概念。要完全复制人类的所有能力，你最终需要制造一个人类的完整复制品，但这已经脱离了AI的范畴。
- **真正的“外星智能”早已存在**：这是他最深刻的洞见。我们不必等待未来某个时刻AGI的降临才能遇到“非人智能”。根据他的计算世界观，**智能（即复杂的计算过程）在宇宙中是无处不在的**。
    - **天气的“智能”**：他举例说，“天气也有它自己的想法”。一个飓风系统内部进行的流体力学计算，其复杂性可能远超人脑。它只是不具备我们所说的“目标”或“意识”，其计算过程与人类的认知目标不“对齐”（aligned）。 .
    - **计算宇宙的智能**：更广泛地说，他研究的元胞自动机等简单程序，已经展现了极其复杂的、能解决问题（如素数判断）的“智能”行为。

因此，沃尔夫勒姆认为，我们对AI的恐惧和期待，很大程度上源于一种**人类中心主义**的偏见，即认为只有“像我们一样”的智能才是真正的智能。他预测，未来我们将逐渐认识到，宇宙本身就是一个充满各种形式、各种层次的“智能”的地方。AI的发展，其真正的意义不在于创造一个我们的复制品（AGI），而在于它将迫使我们拓宽对“智能”本身的定义，并学会与我们周围早已存在的、各种“非人类”的计算过程进行交流和共存。这是一种将AI问题从工程学和伦理学，提升到宇宙学和存在论层面的宏大视角。

4.4 冯·诺依曼 vs 格罗滕迪克：科学遗产的两种范式

【访谈原文翻译】

>> 主持人：好的。这是我的最后一个问题。嗯，亚历山大·格罗滕迪克（Alexander Grothendieck）还是约翰·冯·诺依曼（John von Neumann）？选一个。

>> 斯蒂芬·沃尔夫勒姆：为了什么目的？

>> 主持人：嗯，去见个面。

>> 斯蒂芬·沃尔夫勒姆：嗯，随你怎么解释。作为一种影响，作为一个……一个你被吸引的人。

>> 斯蒂芬·沃尔夫勒姆：我的意思是……

>> 主持人：那是什么？

>> 主持人：我想到的是那个……

> **>> 斯蒂芬·沃尔夫勒姆**：是的。是的。是的。不，我的意思是，我对约翰·冯·诺依曼的了解远多于对亚历山大·格罗滕迪克的了解。所以我……我会说……嗯……你知道……我一直认为冯·诺依曼是一个问了很多非常有趣问题的人，但有时没有选择深入到他本可以达到的深度。嗯……而且……你知道……问了正确的问题，但没有深入挖掘并真正找到根本性的东西。
>
> 格罗滕迪克则有点……也许是相反，他深入挖掘，嗯……你知道……问了最深刻的问题，但并不总是清楚为什么要问那些问题。我的意思是……你知道……其中一件事，无穷大广群（infinity groupoid），我以前从未想过我会发现它有任何用处，但它实际上与我提到的那个Ruliad对象密切相关。所以，嗯……但我……我会说……嗯……你知道，如果有人能把冯·诺依曼研究过的所有东西都拿来，然后说：“这是个好问题。现在，让我们用一种根本性的方式来回答它。” 对吧？我的意思是，我认为那将会是……
>
> 比如量子力学，你知道，他对量子力学的表述，我认为是……嗯……你知道……在某些方面，比如量子计算，它把人们引向了非常错误的方向。因为人们……你知道……在冯·诺依曼的量子力学表述中，它就像是……然后你进行一次测量，“砰”，它就发生了。事情在物理世界中并不是那样运作的，测量实际上是一个过程。你知道，事实是，当你想象一台量子计算机时，有所有这些计算的线程在发生，但要知道答案是什么，你必须把这些计算的线程编织在一起。但冯·诺依曼告诉我们，它只是一个测量算符。所以没人……你知道……人们没有倾向于去拆解它，然后说：“把那些量子计算的线程编织在一起，算出答案到底需要多大的努力？” 我认为……我认为那是一个他有点把人们引向错误方向的地方。但无论如何……

【深度解读】

这个问题看似是一个简单的二选一，但实际上它触及了关于科学风格、学术遗产和思想深度的核心问题。沃尔夫勒姆的回答，不仅揭示了他对这两位20世纪数学巨匠的评价，更重要的是，它清晰地勾勒出了他为自己定位的学术生态位。

为了理解他的回答，我们首先需要知道这两位科学家的风格差异：

- **约翰·冯·诺依曼（John von Neumann）**：一位百科全书式的天才，其研究领域横跨纯粹数学、应用数学、物理学、计算机科学和经济学。他的标志性特点是**开创性**和**广度**。他总能敏锐地抓住一个新领域的核心问题，并为其建立坚实的数学框架（如博弈论、元胞自动机、量子力学的数学基础）。他是一位伟大的“**问题提出者**”和“**领域开创者**”。
- **亚历山大·格罗滕迪克（Alexander Grothendieck）**：一位风格截然不同的数学家，被认为是20世纪最伟大的代数几何学家。他的特点是**极致的抽象**和**深度**。他不喜欢解决孤立的问题，而是致力于建立庞大、普适的理论框架（如概形理论），使得原来的难题在他的新框架下变得“显然”。他是一位伟大的“**理论建造者**”和“**深度挖掘者**”。

沃尔夫勒姆对两人的评价，恰好反映了这种风格差异：

- **对冯·诺依曼的评价**：“问了非常有趣的问题，但有时没有选择深入到他本可以达到的深度。”他认为冯·诺依曼经常在为某个领域打下基础后，就转向下一个新领域，而没有将最初的问题追问到底。他以冯·诺依曼对量子力学的“测量”处理为例，认为这种过于简洁的数学形式

（测量算符），掩盖了“测量”作为一个复杂物理过程的内在机制，从而在某种程度上“误导”了后来的量子计算研究。

- **对格罗滕迪克的评价**：“深入挖掘……问了最深刻的问题，但并不总是清楚为什么要问那些问题。” 他承认格罗滕迪克理论的极端深刻性，但暗示其有时会为了抽象而抽象，脱离了更广泛的科学问题背景。

沃尔夫勒姆的自我定位：两者的综合

通过这番评价，沃尔夫勒姆实际上是在描绘他自己理想中的科学范式，一种试图**结合冯·诺依曼的广度与格罗滕迪克的深度**的模式。他推崇冯·诺依曼那种“问对大问题”的能力，但他不满足于仅仅建立框架，他要像格罗滕迪克一样，深入到最根本的层面，找到问题的最终答案。他的整个科学生涯，可以看作是这样一个过程：“如果有人能把冯·诺依曼研究过的所有东西都拿来，然后说：‘这是个好问题。现在，让我们用一种根本性的方式来回答它。’”

争议与视角

必须指出，沃尔夫勒姆对自己与冯·诺依曼关系的定位，在科学界是存在争议的。一些批评者认为，沃尔夫勒姆在强调自己贡献的同时，不公平地贬低了冯·诺依曼和约翰·康威（“生命游戏”的发明者）等先驱的洞察力 。他们认为，“简单规则产生复杂性”的核心思想，在康威的“生命游戏”中已经得到了淋漓尽致的体现，而沃尔夫勒姆只是对其进行了系统性的研究和推广 。

表2：元胞自动机历史中的关键人物

科学家	关键贡献	沃尔夫勒姆的观点 / 外部评论
约翰·冯·诺依曼	元胞自动机（CA）的共同发明者。用它来探索自我复制的逻辑。设计了一个复杂的“通用构造器” 。	**沃尔夫勒姆的观点**：“他绝对没有看到”简单规则可以创造复杂性。他的工作过于复杂和工程化 。
约翰·康威	“生命游戏”的发明者，这个著名的CA展示了从极简规则中涌现出的复杂类生命行为。	**沃尔夫勒姆的观点**：“和冯·诺依曼一样” 。
斯蒂芬·沃尔夫勒姆	对最简单的（“初等”）一维CA进行了系统性研究。发现了“规则30”的复杂性。提出了“计算等价性原理”，主张简单规则产生复杂性是一种宇宙的普遍现象。	**外部评论**：批评者认为沃尔夫勒姆淡化了冯·诺依曼和康威的基础性工作，并指出“简单规则产生复杂性”的思想在康威的工作中已广为人知 。

对于你这样的学习者而言，这段对话提供了一个宝贵的案例，让你看到科学史是如何被后人书写和解读的。它揭示了科学进步中不同角色之间的张力：是最初提出一个“游戏般”想法的人更重要（康威），还是将这个想法提升为一套普适性科学哲学和系统性研究纲领的人贡献更大（沃尔夫勒姆的追求）？这没有简单的答案，但思考这个问题本身，就是理解科学如何发展的关键一步。

结论：你在Ruliad中的位置

穿越这场与斯蒂芬·沃尔夫勒姆的深度思想对话，我们仿佛进行了一次穿越其宏大计算宇宙的旅行。从科学史的战略价值，到生命与物理的统一计算基础，再到对AI、数学和人类终极问题的独特解答，一个连贯而深刻的世界观呈现在我们面前。

沃尔夫勒姆范式的核心

我们可以将沃尔夫勒姆构建的思想大厦总结为几个核心支柱：

1. **计算的普适性**：宇宙的本质不是由连续的数学方程式所描述，而是由离散的、简单的计算规则在时空中不断迭代所驱动。从星系的形成到意识的运作，万物皆为计算。
2. **复杂性的起源**：我们世界中无处不在的复杂、丰富甚至看似随机的现象，都源于这些底层简单规则的重复执行。这是他“一种新科学”的核心论点。
3. **计算不可约性**：许多这样的计算过程是“不可约的”，意味着我们无法找到预测其行为的捷径。时间的流逝，在根本上就是这种不可简化的计算过程。这一概念是他解释热力学第二定律和自由意志的关键。
4. **观察者的中心地位**：我们所感知的物理定律（如相对论、量子力学），并非宇宙唯一的、绝对的真理，而是像我们这样“计算能力有限”且“相信时间是连续的”观察者，在观察包罗万象的计算可能性总和——**Ruliad**——时，所必然得出的宏观结论。物理学，在某种意义上，是我们自身存在方式的反映。

给未来探索者的邀请

沃尔夫勒姆的理论，无论其最终的对错，都为我们提供了一套前所未有的强大思想工具和一种全新的世界观。对于站在科学门槛上的你，这次访谈的意义远不止于了解一位特立独行的科学家的奇思妙想。它更像是一份邀请，邀请你：

- **学习计算思维**：无论你未来从事哪个领域，掌握计算的语言和思想（例如通过学习Wolfram Language）都将是核心竞争力。它让你不仅能使用现成的模型，更能创造全新的模型来理解世界。
- **敢于叩问基础**：不要将教科书中的公理和定律视为理所当然。像沃尔夫勒姆一样，去追问那些最根本的“为什么”，去探索那些被前人因技术或思想局限而放弃的领域。那里是最高杠杆的创新之地。
- **拥抱“非人”的智慧**：未来的科学发现，可能越来越多地来自于AI对“计算宇宙”的探索。这些发现可能最初看起来是“外星的”、“无法理解的”。学会欣赏、解读并利用这种非人类的智慧，将是下一代科学家的必备技能。

你，作为一个独特的观察者，存在于Ruliad这个无限的可能性空间之中。你所学的每一个概念，你所做的每一次思考，都是在Ruliad中开辟一条属于你自己的路径。沃尔夫勒姆的工作告诉我们，这片空间中还有无垠的未知领域等待探索，还有无限多的“可约性口袋”等待被发现。

科学的未来，将属于那些不仅能站在巨人肩膀上，更敢于质疑巨人脚下基石的人。这趟旅程的终点，正是你未来探索的起点。祝你在这场伟大的智力探险中，发现属于你自己的新大陆。