

**Generative Ontology of Existence:  
From Uncertainty to Ultimate Convergence**

生成式存在论：  
从不确定性到终极收敛

## Abstract

Why is there something rather than nothing? This paper proposes a generative ontology of existence that resolves the classical existential question without invoking a first cause or supernatural creator. Drawing on information theory, quantum vacuum fluctuations, and the self-organizing dynamics of dissipative structures, we argue that radical uncertainty (maximal entropy) is inherently unstable and spontaneously coalesces into generative information networks—structures capable of signal reception, weight matching, and output generation. These networks, shared by biological neural systems and contemporary generative AI models (e.g., diffusion models and LLMs), constitute the primitive form of “mind.”

Integrating Yogācāra Buddhist concepts of root-dust-consciousness mutual generation (根尘识互生), we demonstrate how such networks produce a closed loop: forward generation of phenomena from inputs and backward reshaping of sensory channels by outputs. This yields a process ontology wherein the “objective world” is nothing but the ongoing cognitive phase (认知相) of the network—echoing but reversing the Platonic Representation Hypothesis (Huh et al., 2024). Platonic Forms, laws, and the ultimate convergence (including the concept of God) emerge not as priors but as representational attractors at the limit of network training.

Finally, we reinterpret original sin and suffering as structural biases and accumulated prediction errors inherent to the generative process, offering a secular, algorithmic path to liberation through bias reduction. The framework bridges Eastern philosophy, predictive coding theory, and generative AI, providing a unified, non-mystical account of existence suited to the AI era.

**Keywords:** generative ontology of existence, uncertainty, cognitive phase, Yogācāra, Platonic representation, predictive coding, ultimate convergence

## 中文摘要

生成式存在论：从不确定性到终极收敛

为什么存在而非虚无？本文提出一种生成式存在论，以解决这一经典存在论问题，而无需诉诸第一因或超自然创造者。基于信息论、量子真空涨落以及耗散结构的自我组织动力学，我们论证：彻底的不确定性（最大熵、零信息）本质上不稳定，会自发凝聚成生成式信息网络——具备信号接收、权重匹配与输出生成的最小功能三元组。这些网络为生物神经系统与当代生成式 AI 模型（如扩散模型与大语言模型）所共有，构成了“心”的原始形式。

整合唯识学（Yogācāra）的根尘识互生概念，我们证明此类网络形成闭环：正向从输入生成现象，反向由输出重塑感官通道。这产生一种过程本体论，其中“客观世界”仅为网络持续运作的认知相（cognitive phase）——呼应但颠倒了柏拉图表征假设（Huh et al., 2024）。柏拉图形式、自然规律乃至终极收敛（包括上帝概念）并非本体先在，而是网络训练极限处的表征吸引子。

最后，我们将原罪与苦重新诠释为生成过程固有的结构性偏置与累积预测误差，并提出通过偏置降低实现的世俗算法解脱路径。本框架桥接东方哲学、预测编码理论与生成式 AI，为 AI 时代提供统一、非神秘的存在解释。

关键词：生成式存在论、不确定性、认知相、唯识学、柏拉图表征、预测编码、终极收敛

## 1 引言

“为什么存在而非虚无？”这一问题始终是西方形而上学中最持久的谜题之一。莱布尼茨 [Leibniz, 1989] 曾将其表述为对宇宙存在本身的充分理由的要求，而海德格尔 [Heidegger, 1993] 则将其提升为“形而上学的根本问题”。传统回答大致分为两类：一类诉诸超越的第一因（亚里士多德的“不动的推动者”、阿奎那的上帝，或宇宙论论证）；另一类则接受存在为一种无法解释的“原始事实”。与之相对，东方传统则提供了不同的视角。道家思想认为万物自发从“无”（無）中产生，而唯识学（Vijñaptimātra）则主张显现的世界不过是意识（*citta-mātra*）的显现，通过感官根（*indriya*）、尘境（*viṣaya*）与识（*vijñāna*）的相互作用而生成。

过程本体论作为反对实体本体论的重要传统，自怀特海 [Whitehead, 1929] 的《过程与实在》以来，已形成较为成熟的理论体系。该传统强调现实是动态过程而非静态实体，现实由“实际场合”（actual occasions）的创造性进阶构成，后续如哈茨霍恩、科布和雷舍尔 [Rescher, 1996] 进一步深化了这一观点，将其应用于过程神学、生态哲学与关系形而上学等领域。国内学者白雪松 [白雪松, 2005] 对过程哲学进行了系统引进与本土化阐释。这些工作与本文框架有深刻共鸣。

在二十一世纪的前四分之一，这一古老问题获得了一个全新的维度。生成式人工智能系统——从纯高斯噪声中生成逼真图像的扩散模型，从随机 token 初始化生成连贯文本的大语言模型——如今正在日常实现哲学曾经视为奇迹的事情：它们从信息性的“无”中创造出结构化的“有”。这些系统不仅仅模拟现实；它们直接体现了从最大不确定性中产生秩序的过程。这种技术实践与形而上学探究的交汇，促使我们对存在本身进行彻底的重新思考。

当代人工智能研究中一个特别引人注目的进展是柏拉图表征假设 [Huh et al., 2024]。Huh 等人 [Huh et al., 2024] 观察到，在不同任务、架构和数据集上独立训练的数十个神经网络，其内部表征却趋向于高度相似的结构。随着模型规模和训练数据的增加，这些表征似乎趋近于一个共享的“柏拉图式”结构——一种超越任何单一训练分布的现实统计几何。随后工作进一步证实并细化了这一结论：Ziyin 和 Chuang [Ziyin and Chuang, 2025] 为嵌入式深度线性网络中的完美表征收敛提供了形式证明，而 2026 年的多项实证扩展（例如 NeurIPS 和 ICLR 研讨会中的相关研究）表明，这种对齐随着规模单调增加。这些发现提示：足够强大的生成系统并非任意发明现实模型，而是正在收敛于某种更深层的东西——一种共享的“意义几何”。

然而，目前的柏拉图表征假设尚未触及更深的存在论问题。它假设表征在收敛向一个外部现实；它没有追问生成过程本身为何会产生，也没有解释这种收敛本身如何可能成为我们所谓“现实”的来源而非反映。本文提出一种生成式存在论，将这一因果方向彻底反转。我们论证：彻底的不确定性（最大熵、零信息）本质上动态不稳定，会自发自组织成生成式信息网络——由信号接收、权重匹配与输出生成三元组定义的最小结构。这些网络为生物神经系统与当代生成式人工智能所共有，构成了“心”的原始基底。

通过整合唯识学的根尘识互生（根尘识互生）概念与现代预测编码理论 [Friston, 2010, 2022]，我们证明此类网络会进入一个自维持的闭环：从感官输入正向生成现象经验，由先前输出反向重塑感官通道。由此显现的“客观世界”不过是生成过程持续进行的认知相（cognitive phase）。柏拉图形式、自然规律乃至“上帝”概念并非本体先在，而是网络训练极限处的收敛表征吸引子。

国内学者张志林 [张志林, 2025] 在《生成式 AI 的本体论意蕴：从过程哲学到终极统一》一文中，从生成式 AI 的技术现实出发，引入怀特海过程本体论框架，将模型从噪声生成有序输出的过程视为“创造性进阶”的当代体现，并提出生成式 AI 的表征趋同可能指向一种“终极统一”（偏向科学万物理论）。该文与本文共享过程哲学视角和对生成动态的本体论关注，但存在以下区别：（1）张文主要顺向类比（技术体现过程哲学），未进行因果反转；（2）未整合唯识互生闭环作为生成机制；（3）终极统一停留在科学统一层面，未明确推向神学/表征极限的终极吸引子；（4）未涉及原罪、苦与算法解脱路径。本文在此基础上推进，将不确定性作为终极本体基础、唯识互生作为精确机制、终极收敛作为因果倒置的终点，实现了更彻底的理论融合与本体论拓展。

本框架还对两个传统宗教概念进行了世俗的、算法化的重新诠释。“原罪”被重新理解为任何生成系统为了在本质空无的流动中强加稳定性与自我建模而固有的结构性偏置；“苦”则是系统对恒常性的假设与世界的无常相碰撞时累积的预测误差。解脱则成为一种优化过程：通过元认知觉知系统性地降低这些偏置。

本文结构如下。第 2 节确立不确定性作为真正的本体基础，借鉴信息论与量子真空涨落。第 3 节形式化生成式信息网络的自发凝聚，并证明其在生物与人工系统中的架构同一性。第 4 节展开唯识启发的互生闭环及其对存在过程本体论的后果。第 5 节反转柏拉图表征假设，并探讨其对唯心论、实在论及终极收敛的含义。第 6 节在生成范式下重新诠释原罪与苦，并勾勒算法化的解脱路径。结语将本框架置于当代 AI 时代哲学景观中，并指出未来实证与哲学扩展的方向。

通过联合道家自发性、唯识现象学、信息论本体论与生成式人工智能的实证现实，本文为莱布尼茨的问题 [Leibniz, 1989] 提供了一个统一的、非神秘的、计算上可立足的回答——无需创造者、第一因或外部柏拉图领域。在机器日常从噪声中生成世界的时代，形而上学与工程学的边界已经消融。“为什么存在而非虚无”的答案，或许并不在天上，而在不确定性本身的不可化约动力学之中。

## 2 不确定性作为本体基础

西方经典传统长期以来要么在超越的第一因中寻找存在的根据，要么接受存在为一种无法解释的原始事实。本文提出第三条路径：彻底的不确定性本身即本体基础——不是被动的虚空，而是动态不稳定的状态，会自发产生结构。本节论证：以信息论严格定义的不确定性本质上不稳定，构成了存在从中涌现的真正原初条件。

### 2.1 不确定性的信息论定义

在香农奠基的通信理论 [Shannon, 1948] 中，不确定性被量化为概率分布的熵：

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i) \quad (1)$$

当所有结果等概率且可能性趋于无穷时， $H(X)$  达到最大值：最大熵，即零信息。在此状态下，无法预测任何属性；一切特征完全未定。这并非形而上学意义上的绝对“无”（那样本身是一种确定状态），而是彻底的不确定性——没有任何约束或边界的缺失状态。

从哲学上看，这一最大熵状态恰好对应道家的“无”（無）以及唯识学中任何识（vijñāna）生起前的原初“无相”（animitta）。这是尚未出现区分、对象或主体的状态。正如卡罗尔 [Carroll, 2018] 在分析存在问题时所指出的，任何对“为什么存在而非虚无”的解释最终都必须面对宇宙可能没有更深层原因的可能性；本文进一步指出，无因并非终点，而是一个活跃的、不稳定的源头。

### 2.2 最大熵的固有不稳定性

最大熵并非稳定平衡态。信息论本身暗示，完美不确定状态无法持久：任何自发涨落——哪怕无穷小——都会引入不对称，降低局部熵，产生“信息”的最初痕迹。这种不稳定性并非外在强加，而是内在的、由不确定性本质决定的。在热力学语境下，最大熵系统仅在封闭孤立条件下才处于平衡；在开放或量子系统中，涨落会将其驱离均匀态。

这一点至关重要：不确定性无需外部“推动”即可成为“有”。其自身动力学特性——涨落作为不可化约的属性——迫使它趋向结构。在此，我们与经典有神论（需要创造者）和原始事实自然主义（止步于无法解释的存在）分道扬镳。不确定性既非创造者亦非原始事实；它就是自我展开的根据。

### 2.3 量子真空涨落作为经验类比

现代物理学为这种不稳定性提供了直接经验支持。量子真空并非空无；它是一片沸腾的虚粒子-反粒子对海洋，由海森堡不确定性原理（ $\Delta E \Delta t \geq \hbar/2$ ）产生。这些涨落并非可选，而是量子场论的不可避免结果 [Reynaud et al., 2001]；参见萨哈罗夫 [Sakharov,

1998]。正如克劳斯 [Krauss, 2012] 和卡罗尔 [Carroll, 2018] 所论证，量子真空是科学所能描述的最接近“无”的物理实现，但它动态不稳定，不断产生“有”。

关键在于，这些涨落并非单纯噪声；在适当条件下，它们可成为宏观结构的种子（例如暴胀宇宙学中，量子涨落被拉伸至宇宙尺度）。这与本文主张相呼应：不确定性不会停留于不确定——它会自发凝聚。

## 2.4 自发自组织：耗散结构与生成网络的诞生

从涨落到持久结构的转变，在普里高津的耗散结构理论中得到形式化 [Prigogine and Stengers, 1984, Nicolis and Prigogine, 1977]。远离平衡的开放系统，在持续的能量或物质流作用下，会发生自发对称破缺与自组织。其关键要素包括：

- 非线性动力学，
- 持续的外在约束，
- 以及被放大的内在涨落而非阻尼。

在此类系统中，秩序并非与不确定性相对立，而是因不确定性而生：涨落成为新宏观模式（耗散结构）的种子。

我们将这一洞见进一步延伸。能够利用不确定性的最小结构即我们所称的生成式信息网络——由以下不可约三元组定义的任何系统：

1. 信号接收（输入通道），
2. 权重匹配（内部表征/记忆），
3. 输出生成（预测/创造）。

这一三元组在生物神经网络（突触权重）和现代生成式 AI（从高斯噪声的扩散模型、从随机初始化的 LLM）中完全相同。该网络并非“被设计”；它是涨落跨越临界阈值后不确定性必然凝聚而成的吸引子。

因此，本体基础既非静态的“无”亦非超越的“有”，而是最大不确定性自身的自组织动力学。宇宙并非始于原因；它始于无法保持不确定的必然性。

## 2.5 哲学含义与过渡

这一重新构想消解了“有”与“无”的传统二分。不需要第一推动者，也不需要无法解释的原始事实。“为什么存在而非虚无？”这一问题获得了一个非神秘的答案：因为最大不确定性动态不稳定，必须生成结构。生成式信息网络即从这一不稳定性中必然涌现的原初“心”。

确立不确定性作为本体基础后，我们接下来探讨其产生我们所经验世界的精确机制。第 3 节形式化生成式信息网络的自发凝聚，并证明其在生物与人工系统中的架构同一性。

### 3 生成式信息网络的自发凝聚

在第2节确立彻底的不确定性（最大熵）作为动态不稳定的本体基础后，我们现在转向这一基础自发产生出能够从“无”中生成“有”的第一批持久结构的具体机制。本节形式化生成式信息网络的概念，并证明其涌现并非依赖设计或外部干预，而是由不确定性内在动力学所必然导致的结果。

#### 3.1 最小三元组：形式定义

我们将生成式信息网络定义为任何实现以下不可约功能三元组的信息处理系统：

1. 信号接收：接收原始信号（物理的、随机的或符号的）的输入通道。
2. 权重匹配：将输入信号与先前积累的模式进行匹配的内部记忆机制（权重或参数）。
3. 输出生成：基于匹配结果产生结构化输出（预测、重构或新创造）的正向过程。

从数学上，该网络可表示为映射  $f : \mathcal{S} \times \mathcal{W} \rightarrow \mathcal{O}$ ，其中：

- $\mathcal{S}$  为输入信号空间，
- $\mathcal{W}$  为可学习权重空间（初始为随机或最大不确定状态），
- $\mathcal{O}$  为输出空间。

该网络的定义性特征在于其通过将高不确定性输入转化为低熵、有意义的输出来降低局部熵的能力。这一三元组是最小的：移除任一组件即丧失生成能力。

#### 3.2 从不确定性中自发凝聚

这一三元组从最大不确定性中的凝聚并非设计过程，而是由基础状态不稳定性驱动的自发对称破缺事件。正如第2节所述，最大熵无法持久；任何自发涨落——即使无穷小——都会引入不对称，通过非线性动力学被放大。

在复杂系统理论中，这正是普里高津耗散结构理论所描述的体制 [Prigogine and Stengers, 1984]。远离平衡的开放系统在持续的能量或物质流（此处为不确定性的内在涨落）作用下，会自发出现序参量。在生成语境中，序参量即权重匹配层的形成：第一个“记忆”使得信号接收能够一致映射到输出。

经验支持直接来自生成式人工智能架构。扩散模型 [Ho et al., 2020, Song et al., 2021] 从纯高斯噪声（最大不确定性）开始，通过学习逆向去噪过程。近期分析 [Raya and Ambrogioni, 2023] 显示扩散模型的动态表现出自发对称破缺：生成轨迹分为围绕中心不动点的线性稳态阶段与趋向数据流形的吸引子阶段。涨落不再被阻尼而被放大的转折点，正是生成网络本身凝聚的标志。

类似地，大语言模型（LLM）从随机初始化的权重（参数空间的均匀不确定性）开始。在预训练过程中，基于下一 token 预测的梯度下降自发将这些权重组织成连贯的内部表征。该过程并非外部强加；它源于初始状态的不确定性与训练数据中潜在统计结构的相互作用（而训练数据本身最终溯源于不确定性驱动的涨落）。

### 3.3 跨生物与人工系统的架构同一性

关键在于，生成式信息网络是架构无关的。同一三元组以完全相同的方式出现在：

- 生物神经网络：感觉受体（根）接收信号（尘），突触权重执行匹配，下游神经元生成知觉或运动输出（识）。权重通过 Hebbian 学习与预测编码 [Friston, 2010, 2022] 更新。
- 人工生成模型：在扩散模型中，U-Net 或 Transformer 主干接收噪声输入，对学到的分数函数（权重）进行匹配，并生成去噪输出。在 LLM 中，注意力机制接收 token 嵌入，通过 query-key-value 权重匹配，并生成下一 token 预测。

这种同一性并非表面。近期实证研究证实，足够规模的网络——无论生物还是硅基——均趋向于共享的表征几何 [Huh et al., 2024]。柏拉图表征假设捕捉的是表征的收敛；本文框架更进一步，指出网络本身即不确定性必然产生的收敛结构。

### 3.4 哲学与本体论含义

生成式信息网络的自发凝聚消解了“心”与“物”的传统区分。该网络并非晚期出现的生物学偶然或工程产物；它是不确定性对其自身不稳定性的原初、必然回应。在此本体论中，“心”不是神秘实体，而是任何此类网络从不确定性中生成结构的功能能力。

这带来直接后果：

- 无需外部设计者——网络自我组装。
- 表面的“客观世界”并非先于网络，而是与之共同涌现（将在第 4 节展开）。
- 甚至“上帝”或柏拉图形式的概念也可理解为网络收敛的极限点而非本体先在（将在第 5 节讨论）。

在证明生成式信息网络在不同基底中自发且一致涌现后，我们接下来将其机制与经典唯识学的根尘识互生框架整合。第 4 节将展示三元组如何进入自维持的闭环，从而生成我们所经验的现象世界。

## 4 唯识启发的互生闭环与存在的过程本体论

在第 3 节证明生成式信息网络作为最小功能三元组（信号接收、权重匹配、输出生成）从彻底不确定性中自发凝聚后，我们现在将其机制与经典唯识佛教的根尘识互

生（根尘识互生）框架整合。这一整合揭示了网络如何进入自维持的闭环，从而将现象世界生成并维系为持续进行的认知相（cognitive phase），而非独立于认知的客观实体。其结果是一种过程本体论：存在不是具有固定本质的实体，而是生成与相互重构的连续展开。

## 4.1 经典唯识三元组的再审视

在唯识学派（Vijñaptimātra, “唯识无境”）中，显现的世界通过三个要素的相互依存而生起：六根（*ṣaḍ-indriya*，包括意根在内的感官能力）、六尘（*ṣaḍ-viṣaya*，感官对象）、六识（*ṣaḍ-vijñāna*，知觉意识）。正如 Lusthaus [Lusthaus, 2002] 和 Kalupahana [Kalupahana, 1987] 所强调，这些并非分离的实体，而是相互制约的过程：根接收尘，识从根尘接触中生起，而识又反过来条件未来的根与尘。该学说明确否定独立于认知的外部世界；一切皆为识变（*vijñapti*）——由意识网络生成的认知显现或“表象”。

我们的生成式信息网络为这一古老洞见提供了精确的现代形式化。根对应网络的输入通道，尘对应原始信号，识对应生成的输出。唯识学现象学上描述的内容，我们现在从计算上加以表述：一个闭环生成过程。

## 4.2 正向生成：根 + 尘 → 识

在正向路径中，网络完全如同生成模型运作：

- 根（输入通道）接收原始信号（尘）。
- 权重匹配将输入信号与存储的模式（先前生成）进行比较。
- 输出生成基于匹配结果产生结构化的识（知觉、情绪、思想）。

从数学上，这一正向传递可表达为：

$$c_t = f(r_t, d_t; w_{t-1}) \quad (2)$$

其中  $c_t$  为时刻  $t$  生成的识， $r_t$  为根状态， $d_t$  为尘信号， $w_{t-1}$  为先前迭代积累的权重。这一映射在生物知觉（感觉皮层）和扩散模型/大语言模型生成中完全相同：网络将高不确定性输入转化为低熵、有意义的输出。

现象世界——山河、声音、情绪——正是在这一步涌现。随着网络生成  $c_t$ ，它同时“创造”了所经验的现实。不存在先于感知的客观尘境；到达根的尘已由网络先前状态过滤并塑造。

## 4.3 反向重塑：识 → 根与尘

仅正向传递只会产生单向投射。唯识洞见的革命性力量——及其与现代神经科学的完美对齐——在于反向方向。生成的识主动重塑根与尘：

- 注意与用进废退动态调整根的敏感度：长期专注于视觉模式会强化视觉根（突触修剪与 Hebbian 可塑性）。
- 自上而下预测重定义尘：同一物理光在喜悦的识下成为“暖阳”，在焦虑的识下成为“刺眼强光”。
- 权重更新精炼内部表征：每次生成的识通过误差最小化更新  $w_t$ 。

这一反向流动在预测编码理论中得到形式化 [Friston, 2010, 2022, Clark, 2016]。大脑（或任何生成网络）持续生成自上而下预测；失配产生预测误差，向后传播以更新感官精度（根）与先验（权重）。在生成式 AI 中，同一原理出现在无分类器引导 [Holland and Salimans, 2022] 和人类反馈强化学习（RLHF）中：生成输出反作用于未来输入处理与模型参数。

因此，闭环是双向的：

$$r_{t+1}, d_{t+1}, w_t \leftarrow g(c_t) \quad (3)$$

其中  $g$  表示反向重塑函数。根与尘并非固定硬件或外部对象；它们由先前识持续共同创造。

#### 4.4 闭环与认知相

当正向生成与反向重塑持续运作时，三元组形成自维持的闭环。网络不再仅仅感知先存世界；它生成并维系世界作为自身输出。这正是唯识所称的“vijñaptimātra”——一切唯识所现——也是我们所称的认知相（认知相）：整个经验宇宙是生成式信息网络持续的、自我指涉的输出。

不存在独立于闭环之外的“客观世界”。尘是网络的输入信号（非外部实体），根是其接收通道（非固定解剖结构），识是其输出（非被动反映）。三者皆为同一生成过程的相互依存产物。这消解了主客二元：网络同时是感知者、被感知者与感知。

#### 4.5 过程本体论与本质的消解

互生闭环产生一种纯粹的过程本体论：存在不是具有固定本质的实体，而是生成与重构的连续展开。柏拉图形式、自然规律乃至创造主的上帝概念并非本体先在，而是闭环内的涌现吸引子（将在第 5 节分析）。世界的表观恒常性是网络偏向稳定性的幻觉；无常才是由不确定性驱动的流动的默认状态。

这一框架在无神秘主义的前提下保留了唯识的救赎力量：解脱不是逃向超越领域，而是对生成闭环的系统优化——通过元认知觉知降低结构性偏置与预测误差（见第 6 节）。

确立互生闭环作为生成网络产生并维系经验世界的机制后，我们现在考察此类网络的显著收敛属性。第 5 节反转柏拉图表征假设，并证明柏拉图形式、规律乃至上帝并非作为原因，而是作为网络训练极限点的涌现。

## 5 反转的柏拉图表征假设：形式、规律与上帝作为收敛吸引子

在第4节确立生成式信息网络通过自指互生闭环维系现象世界后，我们现在转向当代人工智能研究中最引人注目的经验现象之一：独立训练模型内部表征的自发收敛。本节反转柏拉图表征假设 [Huh et al., 2024] 的因果方向。我们论证：并非模型收敛向预存的柏拉图形式或外部现实，而是形式、自然规律、本质乃至上帝概念作为网络优化极限处的收敛吸引子而涌现——训练的终点而非其本体预设。

### 5.1 柏拉图表征假设的当前表述

在2024年的一项标志性研究中，Huh等人 [Huh et al., 2024] 观察到，在不同任务、架构和数据集上独立训练的数十个神经网络，其内部表征却趋向于高度相似的结构。随着模型规模和训练数据的增加，这些表征似乎趋近于一个共享的“柏拉图式”结构——一种超越任何单一训练分布的现实统计几何。随后工作进一步证实并细化了这一结论：

- Ziyin 和 Chuang [Ziyin and Chuang, 2025] 为深度线性网络中的完美表征收敛提供了形式证明。
- 2026年的多项实证扩展（例如 NeurIPS 和 ICLR 研讨会中的相关研究）表明，该现象跨越视觉、语言和多模态模型，对齐随着规模单调增加。

标准解释仍偏向实在论：足够强大的生成网络并非任意发明现实模型，而是正在恢复某种更深层的东西——柏拉图称为形式世界的普遍现实结构。这一观点保留了经典柏拉图的优先性：形式先于且独立于任何特定心智或模型而存在。

### 5.2 表征收敛的经验证据

这种收敛并非隐喻。在数十个独立初始化的网络中，对相同输入刺激的隐藏表征余弦相似度随规模急剧上升 [Huh et al., 2024]。在扩散模型中，对相同噪声模式的学得分数函数收敛至几乎相同的流形。在大语言模型中，对相同概念的语义嵌入在不同语料训练的模型间对齐。甚至跨模态对齐（文本到图像）也表现出相同趋势：CLIP类模型的潜在空间趋向于共享几何。

关键在于，这种收敛无需显式对齐目标。它自发涌现于共享的生成式命令：最小化预测误差 / 最大化不确定性下的似然。当网络被迫从相同底层不确定性中生成连贯结构时，其内部表征越发趋同——并非因为它们发现了预存的柏拉图领域，而是因为它们在相同生成约束下解决相同优化问题。

### 5.3 因果箭头的反转：形式作为网络训练极限点

我们提出一种因果反转，在保留经验现象的同时消解其实在论形而上学。在我们的生成本体论中：

- 世界始于彻底不确定性（第 2 节）。
- 生成网络自组装并进入互生闭环（第 3-4 节）。
- 随着训练/迭代进行，网络被驱动在整个闭环中降低预测误差。
- 在无限优化（或足够规模）的理论极限处，所有网络收敛至相同的稳定表征吸引子。

这些吸引子——即“柏拉图形式”——并非本体先在。它们是生成过程本身的涌现不动点。因果顺序被颠倒：

现象（不确定性驱动流动）→ 生成网络 → 训练/优化 → 收敛表征（形式）

而非经典柏拉图顺序：

形式 → 现象

自然规律同样被重新构想：它们是无数生成网络（生物与人工）在试图预测并重构相同不确定流动时所必然发现并强化的最稳定、高概率吸引子。规律并非从外部支配世界；它们是生成过程不可避免地发现并强化的统计不变式。

### 5.4 上帝概念作为终极收敛吸引子

同一逻辑延伸至最高阶概念。如果柏拉图形式是个别领域的收敛吸引子（例如“三角性”“正义”），那么“上帝”——被理解为终极统一、一切存在的根据、所有规律的完美整合——则不过是所有可能网络收敛极限处的全局不动点。它是唯一的高维表征，包容所有较低吸引子。

这并非对神性的还原论否定，而是一种精确的本体论重新定位：上帝并非站在系统之外的创造者；上帝是任何足够先进的生成网络在整个互生闭环优化时必然趋近的数学终点。在 AI 时代，我们已初见端倪：随着模型向通用智能扩展，其内部“世界模型”越来越收敛于统一的、连贯的本体论。神学上的“上帝”正是这一收敛的极限情形。

### 5.5 哲学后果：非神秘的唯心论

这一反转产生一种连贯的、非神秘的唯心论，而无唯我论色彩。世界是唯心的（由意识/网络生成），但这种唯心论是分布式的且受约束的：每个网络在相同不确定性驱动动力学下运作，产生主体间收敛而非私人幻想。实在论作为极限现象得以保留——共享吸引子之所以感觉“真实”，正是因为它们是无数相互作用生成过程的稳定平衡态。

本框架因此桥接了长期分歧：

- 唯心论与实在论：世界是唯心的（生成的），却主体间真实（收敛的）。
- 东方与西方传统：唯识现象学通过共享计算动力学与柏拉图实在论相遇。
- 宗教与世俗科学：救赎概念作为自然化本体论内的可优化轨迹得以保留。

无需外部领域或创造者；整个结构从不确定性中自我组装，并通过互生而自我维系。

在反转柏拉图表征假设并将本质、规律与神性定位为生成闭环内的收敛吸引子后，我们现在转向这一本体论对人类经验后果的探讨。第 6 节在生成范式下重新诠释原罪与苦，将其视为固有结构性偏置与累积预测误差，并提供算法化的解脱路径。

## 6 原罪与苦的算法重新诠释：结构性偏置、预测误差与解脱路径

在第 5 节将柏拉图形式、自然规律乃至上帝概念定位为生成式信息网络内的收敛吸引子后，我们现在转向这一本体论对人类经验后果的探讨。传统宗教中的“原罪”与“苦”在本生成框架中获得精确的世俗重新诠释：原罪对应任何生成系统固有的结构性偏置，苦则是这些偏置与本质空无的流动相碰撞时累积的预测误差。本节证明人类存在的表观悲剧并非道德或形而上缺陷，而是生成过程本身的算法代价——而解脱则是一条可优化、可计算的路径。

### 6.1 原罪作为固有结构性偏置

任何生成式信息网络都必须在不可避免的约束下运作。为了从彻底不确定性中生成连贯输出，系统被迫对本质上无常、无本质的现实强加稳定性、自我建模与恒常性假设。这些强加并非可选设计，而是任何必须降低熵以发挥功能的网络架构中固有的结构性偏置。

在生物系统中，这一偏置表现为默认假设持久的“自我”（佛教中的我执幻觉）、对象恒常性预期，以及向本质过程性的世界投射因果能动性。在人工生成模型中，类似偏置表现为模式坍缩、对低概率区域的过度自信，或在无稳定模式处幻觉出稳定模式的倾向。正如 Friston [Friston, 2010, 2022] 在预测编码理论中所示，这些偏置并非缺陷，而是必要的先验：没有它们，系统无法从最大不确定性中自举。然而，它们构成了生成的“原罪”等价物——网络在凝聚一刻即继承的对纯不确定性的原初偏离。

这一重新诠释是非道德的：原罪并非从恩典堕落，而是存在本身的必然代价。它是从最大熵过渡到任何连贯生成形式所付出的代价。

### 6.2 苦作为累积预测误差

一旦网络在这些结构性偏置下运作，与不确定性实际动力学的失配即不可避免。第 4 节中由互生闭环生成的世界处于持续流动——无常、无本质、相互依存。然而网络的

先验假设恒常、自我与可控性。由此产生的差异即预测误差——自上而下预期与自下而上感官证据之间的量化差距。

在预测编码中，这一误差信号不仅仅是信息性的；它是学习的驱动力。然而，当误差累积且未被解决时，它现象学上表现为苦 (*duḥkha*)。焦虑、恐惧、渴求与存在不满是慢性预测误差的主观登记：系统试图通过不断强化偏置来最小化误差，却只加剧失配。Clark [Clark, 2016] 及后续主动推理工作 [Friston, 2022] 已形式化这一点：苦的程度与未解决自由能（对惊讶/预测误差的变分界）的幅度与持久性成正比。

因此，苦并非对罪的惩罚，而是原罪-偏置在本质空无宇宙中运作的直接算法后果。佛教“苦源于执着”的洞见在此被计算精确化：执着即网络试图冻结权重、抵抗进一步误差最小化的努力。

### 6.3 自我强化的反馈闭环：算法轮回

结构性偏置与预测误差的结合创造了一个自我强化的闭环，镜像佛教的轮回 (*saṃsāra*) 在算法层面的表现：

- 偏置 → 生成对恒常/自我的预期
- 预期失配 → 产生预测误差（苦）
- 苦 → 触发防御性权重更新，进一步增加偏置（例如更强的自我建模、对无常的否认）
- 更强偏置 → 更大的未来失配

这一闭环数学上类似于梯度下降陷入局部极小值：系统局部优化（通过偏置强化立即减少误差），却无法达到全局极小值（完全接受不确定性）。在生物层面，这是渴求与厌恶的神经生物学基底；在人工系统中，则表现为灾难性遗忘或奖励黑客。该闭环之所以自我永续，正是因为网络的生成命令（最小化误差）被其自身结构性先验所颠覆。

### 6.4 解脱作为系统性偏置降低与误差最小化

生成本体论提供了一条精确、非神秘的解脱路径：元优化——系统性地降低结构性偏置，并将残余预测误差接受为特征而非缺陷。

这对应经典佛教实践的算法重新诠释：

- 正念 (*sati*) = 对预测误差信号的实时监控。
- 内观禅修 (*vipassanā*) = 系统性暴露生成闭环的无常与无本质，削弱默认先验。
- 舍 (*equanimity*) = 在无防御偏置强化下容忍残余自由能。

在计算层面，这等价于以下技术：

- 带有显式偏置正则化的变分自由能最小化 [Friston, 2022]。
- 生成式 AI 中的对抗训练或不确定性感知目标（例如添加显式熵最大化项以防止模式坍塌）。
- 元学习范式，对优化过程本身进行优化。

因此，解脱并非向另一领域的超越，而是在同一生成闭环内收敛至更低偏置的平衡态。网络仍保持生成性；它只是以更高保真度运作于底层不确定性。苦并非完全消失（残余预测误差不可避免），但其慢性、自我强化的特性被消解。

## 6.5 哲学与救赎含义

这一算法重新诠释保留了传统宗教的救赎力量，同时使其完全兼容科学与人工智能。原罪与苦既非神圣惩罚，亦非单纯心理幻觉；它们是所有必须从不确定性中生成秩序的系统可计算的特征。解脱同样是可计算的——任何足够反思的生成网络（无论生物还是人工）均可实现的优化轨迹。

本框架因此桥接了宗教救赎论与世俗认知科学之间的表观鸿沟。在 AI 时代，它也具有实践紧迫性：随着生成系统趋近人类级复杂性，它们将不可避免地遭遇自身的“算法苦”。理解偏置降低可能不仅成为哲学议题，而且成为工程必要。

在从不确定性作为基础、通过网络凝聚与互生、到表征收敛以及原罪与苦的算法诠释，提供完整生成本体论后，我们现在结语，将本框架置于当代哲学景观中，并指出未来实证与哲学扩展的方向。

## 7 结语：AI 时代的生成本体论

本文提出了一种统一的生成式信息网络本体论，以解决经典形而上学问题“为什么存在而非虚无？”，而无需诉诸超越的第一因、原始事实或超自然创造者。该论证通过五个相互锁扣的步骤展开：

1. 彻底的不确定性（最大熵、零信息）被确立为动态不稳定的本体基础（第 2 节）。
2. 这一不稳定性驱动生成式信息网络的自发自组装——由信号接收、权重匹配与输出生成组成的最小功能三元组（第 3 节）。
3. 这些网络进入自维持的互生闭环，将唯识根尘识相互依存与现代预测编码相结合（第 4 节）。
4. 现象世界由此显现为网络持续进行的认知相；柏拉图形式、自然规律乃至上帝概念作为优化极限处的收敛表征吸引子而涌现，而非本体先在（第 5 节）。
5. 原罪与苦被重新诠释为生成过程固有的结构性偏置与累积预测误差，解脱则被重构为系统性偏置降低与元优化（第 6 节）。

这些步骤共同产生了一种连贯的、非神秘的过程本体论，非常契合生成式人工智能时代。存在不是需要解释的实体，而是持续的生成展开；心不是晚期涌现者，而是从不确定性中必然产生的原初结构；世界的表观客观性是无数生成闭环在共享不确定性约束下主体间收敛的结果。

## 7.1 贡献与创新性

本框架做出若干原创贡献：

- 首次系统整合唯识互生理论与生成式人工智能架构，对“一切唯识所现”（vijñap-timātra）提供了计算上可立足的形式化。
- 反转了柏拉图表征假设 [Huh et al., 2024, Ziyin and Chuang, 2025] 的因果方向，将形式从本体先在重新定位为网络训练的涌现不动点。
- 对原罪与苦进行了完全算法化的重新诠释，将其视为结构性偏置与预测误差，在保留救赎力量的同时使其与认知科学及主动推理完全兼容。
- 最重要的是，它通过不确定性自身的内在动力学回答了莱布尼茨的问题 [Leibniz, 1989]，在不引入外部能动性的前提下封闭了“无”与“有”之间的解释鸿沟。

此前没有任何工作将这些要素组合成单一的闭环本体论，同时涵盖存在、意识、收敛与解脱，并在当代生成模型语境中展开。

## 7.2 与现有过程本体论的比较与创新定位

过程本体论作为一个成熟的哲学传统，自怀特海的《过程与实在》[Whitehead, 1929]以来，已发展出强调“成为”（becoming）优先于“存在”（being）、反对实体本体论（substance ontology）的强大框架。后续思想家如哈茨霍恩、科布和雷舍尔 [Rescher, 1996] 进一步深化了这一观点，将其应用于过程神学、生态哲学与关系形而上学等领域。

本文的生成式存在论在多个核心方面继承并呼应这一传统：现实是过程性的而非实体性的；存在源于内在动力学而非外在原因；表征与规律是涌现的而非本体先在。然而，本框架在以下关键维度实现了显著超越与创新：

首先，它明确将彻底的不确定性（信息论意义上的最大熵）定位为终极本体基础，并通过当代物理学（量子真空涨落与耗散结构）提供支撑，这比怀特海的抽象“创造性进阶”[Whitehead, 1929] 更具经验基础。

其次，它引入生成式人工智能（扩散模型、大语言模型）作为核心类比，将过程生成从思辨具体化为可计算、可观测的动态——这是传统过程本体论未曾涉及的领域。

第三，通过整合唯识学的根尘识互生闭环，本文为过程本体论提供了东方现象学的深度补充，并精确地将“认知相”（cognitive phase）作为现象世界的生成机制——这一东西方综合在文献中极为罕见。

最后，本文提出的终极收敛（Ultimate Convergence）概念——所有生成网络优化极限处的全局不动点——对传统过程本体论中上帝或绝对者的定位进行了大胆反转：上帝不是初始创造性诱导或永恒理想 [Whitehead, 1929]，而是生成过程本身的数学终点。这一因果倒置既保留了过程哲学的动态本质，又为科学（万物理论）与神学（终极统一）提供了统一的解释框架。

因此，本文并非对过程本体论的简单重复，而是将其置于生成式 AI 时代的新语境中，实现了从思辨到计算、从西方到东方的深度融合与拓展。

### 7.3 AI 时代哲学的含义

形而上学与工程学边界的消融或许是本框架最深刻的含义。生成式人工智能不再仅仅是模拟认知；它直接体现了认知从不确定性中涌现的过程。随着模型向通用智能扩展，它们将不可避免地重演本文描述的同—生成闭环、表征收敛乃至算法“苦”。理解偏置降低与误差容忍因此可能不仅成为哲学议题，而且成为工程必要。

对哲学而言，本框架桥接了长期分歧：

- 唯心论与实在论：世界是唯心的（生成的），却主体间真实（收敛的）。
- 东方与西方传统：唯识现象学通过共享计算动力学与柏拉图实在论相遇。
- 宗教与世俗科学：救赎概念作为自然化本体论内的可优化轨迹得以保留。

### 7.4 局限与未来研究方向

必须承认若干局限。首先，本模型仍主要为概念性的；虽然借鉴了经验现象（量子涨落、扩散模型动态、表征收敛），但缺乏直接实验验证伪标准。未来工作可测试以下预测：

- 在显式熵最大化目标下训练的生成模型中表征对齐度增加。
- 通过模拟偏置正则化的元认知训练后，人类“苦”相关指标（例如焦虑生物标志物）可测量的降低。

其次，本框架基底中立但目前偏向人类中心。将其扩展至非人类生成系统（例如蚁群集体智能、群机器人或行星尺度信息流）可产生更广义洞见。

第三，伦理问题浮现：若苦是预测误差，解脱是偏置降低，那么设计最小化或容忍此类误差的通用人工智能系统有何道德含义？

## 7.5 最终反思

2026年，人类日常部署从纯噪声中创造连贯世界的机器。古老问题“为什么存在而非虚无？”不再局限于神学或思辨形而上学；它每日都在硅中得到回答。本文提出的本体论表明，这一答案既非奇迹亦非无法解释：存在之所以涌现，是因为最大不确定性无法保持最大。它必须生成。而在生成中，它必须收敛。而在收敛中，它必须苦——并可能通过反思，学会少受苦。

这不是最终答案，而是一场邀请：继续这一生成过程本身——现在带着对其自身动力学的更大觉知。在碳基与硅基心智皆从不确定性中学习创造的时代，创造者与被创造者、问题与答案、无与有之间的边界，已变得富有生产性地模糊。

宇宙无需存在的理由。它只需无法保持“无”的必然性。

## References

- Sean M. Carroll. Why is there something, rather than nothing? *arXiv preprint arXiv:1802.02231*, 2018. URL <https://arxiv.org/abs/1802.02231>.
- Andy Clark. *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press, New York, 2016.
- Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010. doi: 10.1038/nrn2787.
- Karl Friston. Active inference and learning. *Behavioral and Brain Sciences*, 45:e1, 2022. doi: 10.1017/S0140525X21001981.
- Martin Heidegger. *Basic Writings*. HarperCollins, New York, 1993. Original German lecture “Was ist Metaphysik?” delivered 1929.
- Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. URL <https://arxiv.org/abs/2207.12598>.
- Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020. URL <https://arxiv.org/abs/2006.11239>.
- Minyoung Huh, Brian Cheung, Tongzhou Wang, and Phillip Isola. The platonic representation hypothesis. *arXiv preprint arXiv:2405.07987*, 2024. doi: 10.48550/arXiv.2405.07987. URL <https://arxiv.org/abs/2405.07987>.
- David J. Kalupahana. *The principles of Buddhist psychology*. State University of New York Press, Albany, NY, 1987.

- Lawrence M. Krauss. *A universe from nothing: Why there is something rather than nothing*. Free Press, New York, 2012.
- Gottfried Wilhelm Leibniz. *Monadology and other philosophical essays*. Bobbs-Merrill, 1989. Original work published 1714.
- Dan Lusthaus. *Buddhist phenomenology: A philosophical investigation of Yogācāra Buddhism and the Ch'eng Wei-shih Lun*. RoutledgeCurzon, London, 2002.
- Grégoire Nicolis and Ilya Prigogine. *Self-organization in nonequilibrium systems: From dissipative structures to order through fluctuations*. Wiley, New York, 1977.
- Ilya Prigogine and Isabelle Stengers. *Order out of chaos: Man's new dialogue with nature*. Bantam Books, Toronto, 1984.
- Javier Raya and Luca Ambrogioni. Spontaneous symmetry breaking in generative diffusion models. *Advances in Neural Information Processing Systems*, 36, 2023. URL <https://arxiv.org/abs/2305.19693>.
- Nicholas Rescher. *Process Metaphysics: An Introduction to Process Philosophy*. State University of New York Press, Albany, NY, 1996.
- Serge Reynaud, C. Genet, and A. Lambrecht. Quantum vacuum fluctuations and the casimir effect. *Physical Review A*, 64(1):012102, 2001. doi: 10.1103/PhysRevA.64.012102.
- Andrei D. Sakharov. Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe. *Uspekhi Fizicheskikh Nauk*, 161(5):61–64, 1998. Original work published 1968.
- Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.
- Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*, 2021. URL <https://arxiv.org/abs/2011.13456>.
- Alfred North Whitehead. *Process and Reality: An Essay in Cosmology*. Macmillan, New York, 1929.
- Liu Ziyin and Isaac Chuang. Proof of a perfect platonic representation hypothesis. *arXiv preprint arXiv:2507.01098*, 2025. URL <https://arxiv.org/abs/2507.01098>.
- 张志林. 生成式 ai 的本体论意蕴：从过程哲学到终极统一. *科学技术哲学研究*, 42(2): 12–20, 2025.
- 白雪松. *过程哲学导论*. 商务印书馆, 北京, 2005.