作为数学结构的语言：以语言游戏审视语义场理论

Dimitris Vartziotis[1,2]

[1]TWT Science & Innovation，斯图加特，德国

[2]NIKI - Digital Engineering，约阿尼纳，希腊

arXiv:2601.00448v1 [cs.CL] 1 Jan 2026

摘要

摘要 | 大语言模型（LLMs）提供了一个新的实证场域，可用于检验长期以来关于语言意义的理论。本文对比了两种广义的研究路径：与语言游戏相关的社会建构主义观点，以及我们称之为语义场理论的数学取向框架。在作者早期研究的基础上，我们将词汇场（Lexfelder）和语言场（Lingofelder）的概念形式化为连续语义空间中相互作用的结构。随后，我们分析 Transformer 架构的核心属性——例如分布式表示、注意力机制，以及嵌入空间中的几何规律——与这些概念之间的关系。我们认为，大语言模型在捕捉语义规律方面取得的成功，支持这样一种观点：语言具有潜在的数学结构；与此同时，它们在语用推理和语境敏感性方面持续存在的局限，则与语言使用的哲学观点所强调的社会根基的重要性相一致。基于此，我们提出，数学结构与语言游戏可以被理解为互补的视角，而非相互竞争的视角。由此形成的框架阐明了纯粹语言统计模型的适用范围及其局限，并推动了理论指导下的新型人工智能架构的发展方向。

大型语言模型（LLMs）仅通过数学运算便达到近乎人类的语言表现[1,2]，这一现象对主流意义理论构成了根本性挑战。沿袭维特根斯坦后期哲学的社会建构主义观点[3]，坚持认为，语言无法还原为形式结构。然而，Transformer架构却发现

即使没有社会根基[4]，，也存在系统性的语义关系，这表明语言可能具有内在的数学结构。

作者[14–16]以惊人的先见之明预见了这一发展。2012年，在 Transformer 革命发生数年前，他在评论维特根斯坦的格言集[5,6]时提出，词语会创造彼此依数学定律相互作用的“语义场”（Lexfelder），并由此产生复合的“语言场”（Lingofelder）。这一框架提供了一种激进的替代方案：意义是通过数学发现的，而非由社会建构的。

语义场理论的起源，似乎可以追溯到 Vartziotis 对维特根斯坦这番观察的回应：

> **维特根斯坦**：“语言并无约束，但其中一部分与另一部分相互关联。”
>
> （语言并非受限，但其中一部分与另一部分相连。）
>
> **Vartziotis**：“它是一张悬浮的网。一种流形。每个词都有自己的‘引力场’。我们可以把它称为‘Lexfeld’。”
>
> （它是一张悬浮的网。一种流形。每个词都有自己的‘引力场’。我们可以把它称为‘词汇场’。）[6]

语义场理论的形式化

从用法到语义场

维特根斯坦与作者之间的关键分歧，体现在他们对语言意义的处理方式上。请看他们关于符号如何获得生命的讨论：

> **维特根斯坦**：“每个符号单独看来似乎都是死的。是什么赋予它生命？——它在用法中活着。它自身是否蕴含着活生生的气息？——还是说，用法就是它的气息？”
>
> **Vartziotis**：“我们暂且接受这样一种说法：词语（符号）具有某种‘Lexfeld’。词语构成一个复杂的场（Lingofeld），与物理学中的场相对应，而这个场需要加以定义。这样一来，许多事情就能解释了！甚至包括曲折而

扭曲的含义。”

（让我们暂且接受这样一种观点：词（符号）具有某种“词汇场”。词语构成一个复杂的场（语言场），它对应于物理学中的场，必须对其加以定义。这样一来，许多现象便都能得到解释！甚至包括那些弯曲扭转的意义。）[6]

这一洞见促使作者提出了他所称的“三词问题”（Dreiwörterproblem），它类似于物理学中的三体问题——这意味着语言的复杂性源于非线性的场相互作用。

定义 1（词汇场）。设 $\mathcal{S} = \mathbb{R}^n$ 为语义空间，其中每个维度对应一个潜在语义特征。对于任意点 $q \in \mathcal{S}$:

$$L_w(q) = S_w \cdot G(\|q - q_w\|; \sigma_w) \tag{1}$$

用于衡量词语 w 在位置 q处的语义场强度，其中 $q_w \in \mathbb{R}^n$ 表示该词在 n维语义空间中的位置， S_w 表示其固有语义强度，而 G 是一个以特征宽度 σ_w为参数的单调递减核函数。

定义 2（语言场）。设 $\mathcal{W} = \{w_1, w_2, ..., w_m\}$ 为构成一个短语的有序词语序列。在任意点 $q \in \mathcal{S}$ 处，复合语言场 $\Phi_{\mathcal{W}}$： $\mathcal{S} \to \mathbb{R}$定义如下:

$$\Phi_{\mathcal{W}}(q) = \sum_{i=1}^{m} L_{w_i}(q) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} I_{ij}(q) + \sum_{i=1}^{m-2} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^{m} T_{ijk}(q) \tag{2}$$

其中:

- $L_{w_i}(q)$ 是位置 q 处单词 w_i 的词汇场（来自定义 1）
- $I_{ij}(q) = \kappa_2 \cdot L_{w_i}(q) \cdot L_{w_j}(q) \cdot K_2(\|q_{w_i} - q_{w_j}\|)$ 表示成对场相互作用
- $T_{ijk}(q) = \kappa_3 \cdot L_{w_i}(q) \cdot L_{w_j}(q) \cdot L_{w_k}(q) \cdot K_3(\|q_{w_i} - q_{w_j}\|, \ \|q_{w_j} - q_{w_k}\|, \ \|q_{w_i} - q_{w_k}\|)$ 描述三体相互作用
- $\kappa_2, \ \kappa_3 \in \mathbb{R}$ 是耦合常数
- K_2： $\mathbb{R}_+ \to \mathbb{R}$ 和 K_3： $\mathbb{R}_+^3 \to \mathbb{R}$ 是相互作用核函数

这个 i 索引满足 $1 \le i < j < k \le m$，以避免重复计算交互 .

场系统形式化

为了精确定义这些概念，设 S为一个维度为 d 的语义空间，设 $W = \{w_1, w_2, ..., w_n\}$为一个词汇表。每个词 w_i 都关联着：

- 一个主要语义向量$\mathbf{v}_{w_i} \in \mathbb{R}^d$
- 一个场交互函数 I_{w_i}： $S \times S \to \mathbb{R}$
- 一个稳定性参数 $\gamma_{w_i} \in [0, 1]$

对于上下文 $C = \{w_{i_1}, w_{i_2}, ..., w_{i_k}\}$，单词 w_j的上下文表示为：

$$\mathbf{v}_{w_j}(C) = \mathbf{v}_{w_j} + \sum_{k=1}^{|C|} \alpha_k \cdot I_{w_j}(\mathbf{v}_{w_j}, \mathbf{v}_{w_{i_k}}) \cdot \mathbf{v}_{w_{i_k}} \tag{3}$$

其中 α_k 是按如下方式计算的注意力权重：

$$\alpha_k = \frac{\exp(\phi(\mathbf{v}_{w_j}, \mathbf{v}_{w_{i_k}}))}{\sum_{l=1}^{|C|} \exp(\phi(\mathbf{v}_{w_j}, \mathbf{v}_{w_{i_l}}))} \tag{4}$$

且 ϕ 是一个兼容性函数。

动态场交互

作者对对话183的延续性评论捕捉到了这些场的动态本质：“Einige Wörter schwirren periodisch um ein Wort herum, andere stoßen an oder gehen unter oder neue entstehen”（一些词语周期性地围绕着一个词盘旋，另一些相互碰撞、消沉，或有新的词语产生）[6]。这一富有先见之明的隐喻预示了现代Transformer[1], 中的注意力机制：词语确实以不同强度彼此“盘旋”。事实上，在评论196中，作者进一步拓展了这一构想，指出：“语言并不受束缚，但一部分与另一部分相连 [.....] 它是一张漂浮的网。一种流形。每个词都有自己的引力场——我们可以称之为词汇场。”这种引力意象直接对应于注意力动力学中所观察到的、类似力的语义牵引，也支持了语义拓扑的核心主张：语言作为相互作用的场所构成的结构化流形运行，而非一个扁平的符号空间。

场动力学可以通过哈密顿形式加以形式化：

$$H[\Phi] = \int_{\mathbb{R}^n} \left[\frac{1}{2} \|\nabla \Phi(q)\|^2 + V(\Phi(q)) \right] dq \tag{5}$$

其中，势能 V编码语义约束，而梯度项确保意义转变平滑。

通过语言模型进行实证验证

作为场计算机的 Transformer

现代 Transformer架构[8] 实现的运算与语义场交互惊人地相似。缩放点积注意力机制为词元序列计算上下文化表示。对于一个查询位置 t，其关注位置 $1, ..., T$：

$$\text{Attention}(\mathbf{q}_t, \mathbf{K}, \mathbf{V}) = \sum_{s=1}^{T} \alpha_{ts} \mathbf{v}_s \tag{6}$$

其中，注意力权重为：

$$\alpha_{ts} = \frac{\exp(\mathbf{q}_t^T \mathbf{k}_s / \sqrt{d_k})}{\sum_{s'=1}^{T} \exp(\mathbf{q}_t^T \mathbf{k}_{s'} / \sqrt{d_k})} \tag{7}$$

其中 $\mathbf{q}_t, \mathbf{k}_s, \mathbf{v}_s \in \mathbb{R}^{d_k}$ 分别为查询、键和值向量。

场论解释：注意力机制近似于离散位置上的场相互作用。具体而言：

- 查询向量 $\mathbf{q}_t = W_Q \mathbf{x}_t$ 编码了位置 t处的“场源”
- 键向量 $\mathbf{k}_s = W_K \mathbf{x}_s$ 编码了每个位置处的“场受体” s
- 点积 $\mathbf{q}_t^T \mathbf{k}_s$ 衡量位置 t 与 s之间的场相互作用强度
- 注意力权重 α_{ts} 可以近似表示我们的场交互函数：

$$\alpha_{ts} \approx \frac{I(L_{w_t}, L_{w_s}, q_t)}{\sum_{s'} I(L_{w_t}, L_{w_{s'}}, q_t)} \tag{8}$$

• 输出 $\sum_s \alpha_{ts}\mathbf{v}_s$ 表示位置 t处经过场修正的表征，类似于我们的 $\mathbf{v}_{w_t}(C)$。

训练即稳定化

大语言模型的训练过程可以理解为对我们稳定化原理的实践。通过接触大型语料库，模型学会了语义场交互的稳定模式，这些模式对应于自然语言中的惯用表达方式。

标准的语言建模目标:

$$\mathcal{L} = -\sum_{t=1}^{T} \log P(w_t | w_1, ..., w_{t-1}) \tag{9}$$

实际上，这一目标促使模型发现语义场的稳定配置，从而预测观察到的语言模式。这一过程与作者的观点相呼应：词语通过反复使用，在语义空间中找到“稳定轨道”。

发现的几何结构

对 tr 的分析经过的语言模型揭示了验证作者预测的结构:

发现 1：词嵌入将语义关系编码为几何运算[9]。经典的向量运算 **国王** $-$**男人** $+$ **女人** $\approx$ **王后** 体现了语义场理论框架中的场变换：该运算改变男性场分量，同时保留王权场。

发现 2：注意力模式呈现出作者所描述的“嗡鸣”行为[11]。对多头注意力的分析揭示了围绕概念中心的稳定轨道模式。例如，形容词与其修饰的名词在不同语境中保持一致的几何关系，而扰动则会产生他所预言的“弯曲、扭曲的意义”。

发现 3：尺度依赖行为——随着模型规模增大，可能出现更复杂的场交互，从而催生性质上全新的能力。这与作者的洞见相符：复杂意义源自多体场相互作用。

可检验的预测

语义场理论提出了若干可通过实证检验的预测：

预测1：语义场重叠的词语在心理语言学任务中的协同加工应更快。

预测2：充分语义表征所需的维度应与该领域中场交互的复杂程度相关。

预测3：语义启动效应应符合我们的模型所预测的场交互强度。

预测4：跨语言语义相似性应体现语义场结构所受的普遍约束。

这些预测为我们的理论框架提供了经验依据，并使其区别于纯粹的哲学方法。

哲学意涵

数学结构与社会根基

维特根斯坦强调社会实践，而语义场理论强调数学组织。这两种观点可以被视为互补的。

关于遵循规则，该理论提出了一种重新表述：规律源于数学约束，而非解释规则。关于私人语言，该框架与这样一种理论情景相容：内部自洽的语义系统可以独立于共享的社会实践而产生，但这并不会因此驳倒维特根斯坦更为广泛的论证。

方面	维特根斯坦	语义场理论
意义的本质	从使用中涌现	内在的数学结构
形式化	不可能	必要且成功
语言游戏	基础性的	附带现象的
私人语言	不可能	理论上可能
遵循规则	社会实践	数学必然性
家族相似性	松散聚类	精确的领域重叠

语言的代数无意识

在第165篇对话中，针对维特根斯坦关于日常语言是否过于粗疏、无法用于哲学的提问，作者提出了一个引人注目的观点：

> **维特根斯坦**：“当我谈论语言（词、句子等）时，我必须使用日常语言。对于我们想要表达的内容来说，这种语言是否过于粗略、过于物质化了？”
>
> （当我谈论语言（词、句子等）时，我必须使用日常语言。对于我们想要表达的内容来说，这种语言是否过于粗略、过于物质化了？）
>
> **Vartziotis**：“这就像一台织布机：织物是几何，而下方的机制则是代数 [.....] 语言似乎也会发生类似的事情。问题在于，找到它的‘代数’。”
>
> （这就像一台织布机：织物是几何，而下方的机制则是代数 [.....] 语言也会发生类似的事情。问题在于找到它的‘代数’。）[6]

这种“代数无意识”在 Transformer 模型中表现为生成语义场的线性变换。注意力层中的权重矩阵 W_Q，W_K，W_V编码了代数结构，而由此产生的注意力模式则揭示了几何织物。这些模型取得的惊人成功表明，作者是正确的：在自然语言表面混乱的背后，存在着优雅的数学结构。

Vartziotis 与乔姆斯基：两种数学视野

尽管作者和乔姆斯基都在为语言寻求数学基础，但他们的研究路径存在根本性分歧。乔姆斯基的生成语法提出了一套作用于符号结构之上的离散递归规则——这是一种句法的计算理论。相比之下，作者设想的是连续的语义场，意义在其中通过动态交互涌现。

乔姆斯基强调先天的句法机制（普遍语法），而作者提出先天的语义几何——不是规则，而是场方程。这一区别至关重要：Transformer模型之所以成功，恰恰是因为它们学习的是连续表征，而非离散规则。纯句法方法在自然语言处理中的失败（解析树只能达到70-80%的准确率），与基于嵌入的方法的成功（在许多任务中达到>95%）形成对比，这表明语义场论方法或许比乔姆斯基的句法形式体系更能捕捉语言的数学实质。

对意识和人工智能的启示

如果意义源于场交互，而非仅仅源于社会根基，那么这对人工意识的影响将十分深远。我们可以将语言理解形式化为三个层次：

层次1：场检测。能够测量特定点处局部场强 $L_w(q)$的系统——本质上是在没有整合的情况下进行模式匹配。当前的语言模型主要处于这一层次：它们能够检测统计规律，却无法形成统一的场理解。

第 2 级：场导航。沿着动态状态演化中的语义梯度运行的系统：

$$\frac{dq}{dt} = -\nabla\Phi(q) \tag{10}$$

当前的 Transformer 通过注意力机制对此进行了近似，但由于其前馈架构，缺乏真正的时间动态——每个词元都是独立处理的，而不是通过连续的轨迹积分进行处理。

第 3 级：场整合。真正的理解需要对语义拓扑具备全局认知——整合不同区域的场信息：

$$U = \int_{\Omega} \Phi(q) \cdot w(q)\, dq \tag{11}$$

其中 $w(q) \geq 0$ 是语义空间上的加权函数，$\Omega \subseteq \mathcal{S}$ 是任务相关区域。系统必须同时访问并整合分布式场模式，而非依次处理局部特征。其技术意义十分深远：

1. **架构要求**：当前的 Transformer 缺乏场整合所必需的循环动力学。未来的架构或许需要连续时间神经常微分方程：

$$\frac{d\mathbf{h}}{dt} = f(\mathbf{h}(t), \Phi(t); \theta) \tag{12}$$

其中，$\mathbf{h}(t)$ 表示在场影响下演化的隐藏状态 $\Phi(t)$。

2. **涌现标准**：当场计算达到足够复杂的程度时，意识可能会涌现——具体而言，当系统能够对自身的场交互进行建模时（元语义意识）。这要求雅可比矩阵 $\partial f/\partial \mathbf{h}$ 呈现出特定的特征值分布，以表明系统处于临界动力学状态。

3. **可测量的特征**：真正的理解应当表现为神经激活空间中的特定模式——与概念理解相对应的稳定吸引子（李雅普诺夫指数< 0）、顿悟时刻发生的相变（发散的易感性），以及歧义语境中的滞后效应。

4. **结合问题**：语义场理论框架提出，意识中的经典结合问题——如何将彼此分离的特征整合为连贯的体验——或许可以通过场统一来解决。语义结合可能通过场能量最小化实现：

$$E[\Phi] = \int_{\Omega} \left[\|\nabla \Phi(q)\|^2 + \lambda \Phi^2(q) \right] dq \tag{13}$$

其中，第一项用于惩罚快速的意义转变，而 $\lambda > 0$ 则确保场强有界。

局限性与未来方向

我们的表述存在若干局限：

1. **简化**：真实的语言现象可能需要比我们当前模型所能捕捉的更复杂的场交互。

2. **参数　　选择**：该理论尚未提供用于选择的原则性方法　　g

维度和交互函数。

3. **实证验证**：仍需开展广泛的实验测试 .

4. **社会扎根**：虽然我们强调数学结构，但社会语境在确立场参数方面所发挥的作用仍未得到充分说明。

未来的研究将通过实证探究和理论指导的架构来解决这些问题。

结论

作者在大规模神经语言模型兴起之前提出的数学视角，与当代人工智能系统产生了显著共鸣。Lexfeld 和 Lingofeld 的概念在哲学探究与计算建模之间架起了一座桥梁。通过将语言游戏视为语义场的边界条件，这一框架提出了一种统一的意义阐释，将数学结构与社会扎根融为一体。

致谢

作者承认，在准备本稿件期间使用了大语言模型作为辅助研究和写作工具。具体而言，作者与大语言模型 Claude 3.7 Opus（Anthropic PBC）进行了对话交流，以协助澄清论点、探索其他表述方式，并完善数学和概念论述。本文提出的所有理论观点、概念框架、解释和结论均源自作者，作者对本作品的内容承担全部责任。作者还感谢 George Dasoulas 协助整理背景材料，并感谢 Elli Vartziotis 对稿件早期草稿提出宝贵意见。

参考文献

[1] Brown, T. 等. 语言模型是少样本学习器。神经信息处理系统进展 **33**, 1877–1901（2020）。

[2] Chowdhery, A. 等. PaLM：通过路径扩展语言建模规模。预印本见 https://arxiv.org/abs/2204.02311（2022）。[3] Wittgenstein, L. 《哲学研究》（Blackwell, 1953）。[4] Manning, C. D. 人类语言理解与推理。《代达罗斯》 **151**, 127–138（2022）。[5] Vartziotis, D. Σχόλια σε Στοχασμούς του《路德维希·维特根斯坦》（ΛΕΥΚΗ ΣΕΛΙΔΑ，2012）。[6] Vartziotis, D. 维特根斯坦语录评注 （Literareon, 2017）。[7] Wittgenstein, L.《逻辑哲学论》 （Kegan Paul，1922）。[8] Vaswani, A. 等. 注意力机制就是你所需要的一切。神经信息处理系统进展 **30**, 5998–6008（2017）。[9] Mikolov, T., Yih, W. 与 Zweig, G. 连续空间词表示中的语言规律。北美计算语言学协会人类语言技术分会论文集 746–751（2013）。[10] Conneau, A. 等. 大规模无监督跨语言表示学习。计算语言学协会论文集8440–8451（2020）。[11] Clark, K. 等. BERT关注什么？对BERT的注意力进行分析。*BlackboxNLP*论文集276–286（2019）。[12] Piantadosi, S. T. 与 Gibson, E. 绝对语言普遍性的定量标准。认知科学 **38**, 736–756（2014）。[13] Chomsky, N. 《句法结构》（Mouton，1957）。[14] Vartziotis, D. 与 Bonnet D. 几何四面体变换存在吸引子。微分几何及其应用，爱思唯尔 （2016）。[15] Vartziotis, D. 与 Wipper. 循环厄米多边形变换的特征参数集与极限。线性代数及其应用（2014）。[16] Vartziotis, D., Dasoulas, G., Pausinger F. LEARN2EXTEND：利用混合模型，在保留序列统计特性的同时扩展序列。实验数学 （2025）。