
名称无损花香：大型语言模型的范畴同伦论^{*}

A PREPRINT

Sridhar Mahadevan

Adobe Research and University of Massachusetts, Amherst
smahadev@adobe.com, mahadeva@umass.edu

August 15, 2025

ABSTRACT

自然语言充满了表面上不同但意思相同的陈述，例如“Charles Darwin wrote”和“Charles Darwin is the author of”。在这些情况下，大型语言模型（LLM）应该产生相同的下一个词的概率，但通常并非如此。已经探索了经验性的变通方法，例如使用 k -NN 的句子相似性估计来生成平滑的估计。在本文中，我们更抽象地解决这个问题，引入了用于 LLM 的范畴同伦框架。我们引入了一个 LLM 马尔可夫范畴，以表示由 LLM 生成的语言中的概率分布，其中句子的概率，例如“Charles Darwin wrote”，由马尔可夫范畴中的一个箭头定义。然而，这种方法遇到了困难，因为语言充满了等价的重述，每个重述在 LLM 马尔可夫范畴中生成一个非同构的箭头。为了解决这一基本问题，我们使用范畴同伦技术来捕捉 LLM 马尔可夫范畴中的“弱等价性”。我们详细概述了从高等代数 K 理论到模型范畴，将过去半个世纪发展出的强大的理论成果应用于 LLM 的范畴同伦方法。

Keywords Large Language Models · Category Theory · Homotopy Theory · AI · Machine Learning

^{*}Draft under review.

Contents

1 介绍

近年来，大型语言模型（LLMs）基于 Transformer 模型 [?] 和相关的结构化状态空间序列模型 [Gu et al., 2022] 的成功，彻底改变了人工智能。LLMs 通过使用海量数据 [Bommasani et al., 2022] 构建大型基础模型，使得近期实现人工通用智能（AGI）[Morris et al., 2023] 成为可能。最近的进展，例如 DeepSeek-R1 [DeepSeek-AI et al., 2025] 表明，通过使用基于强化学习 [?] 的试错搜索来构建问题，可以大幅降低构建这些模型的计算成本。然而，这些系统到底在学习什么仍然不清楚。此外，关于 LLMs 的现有理论结果，一方面表明它们无法识别 Dyck 语言或计算奇偶函数 [Hahn, 2020, Merrill et al., 2022]，另一方面又表明它们是序列上的通用函数逼近器 [?]

我们的方法受到语言范畴模型的启发 [Asudeh and Giorgolo, 2020, Bradley et al., 2022, Coecke, 2019]，但這些方法未解决语义同伦的问题：何时语言中两个看似不同的句子“表示”相同的意思？例如，Bradley et al. [2022] 为 LLM 定义了一个句法范畴（见定义 ??）和一个语义范畴（见定义 ??），其中对象是句子的片段 x ，态射 $x \rightarrow y$ 代表通过 y 可能扩展片段 x 的方式。可能的补全集合被表示为单位区间 $[0, 1]$ 的一个富加性单体预序范畴中的对象，其中逻辑连接词如蕴含被定义为伴随函子。Merrill et al. [2024] 调查了是否可以从 LLM 中的共现模式中学习到蕴含。这些研究都没有关注同伦。在关于使用对称单体范畴建模语言的庞大作品中量子自然语言处理 [Coecke, 2020, Coecke et al., 2010]，将自然语言转换成用于量子计算的电路 [Coecke and Kissinger, 2017]。有一个名为 lambeq [Kartsaklis et al., 2021] 的 Python 包。量子 NLP 工作尚未解决同伦问题。他们的方法基于由 Lambek 前群文法定义的语言组合理论，这是一种由部分有序幺半群定义的范畴类型 [Lambek, 2006]。

LLMs 将句子前缀编码为固定大小的表示，并使用此编码预测文本中的下一个单词。但对于自然语言的释义来说，这种下一个标记的预测是具有挑战性的：例如，“Charles Dickens is the author of” 应生成与 “Charles Dickens wrote” 相同的下一个标记分布。Khandelwal et al. [2020] 使用最近邻方法来提高 LLM 在此类下一个标记预测问题上的性能。我们的工作提出了一种严格的同伦理论来分析相似的文本片段。我们的同伦框架也与关于 LLM 模型蒸馏的研究 [Hsieh et al., 2023] 相关，但这些方法通常依赖于监督训练来构建较小的特定任务模型，而我们提出的是对模型参数所在的低维流形的内部分析。最后，我们的工作与利用商业 LLM 实现中的对数概率或对数几率信息来揭示内部架构细节 [Finlayson et al., 2024] 的研究相关，这些研究通过寻找线性低维矩阵流形来实现，其中 logit 分数所处的位置，但我们的方法使用了更强大的非线性同伦方法。

为了阐述主要思想，我们提出使用马尔科夫范畴 [Fritz, 2020] 的框架来建模大语言模型，这是一种对称单模范畴，每个对象上都有共模的“复制-删除”结构。这种马尔科夫范畴最近被证明能够进行广泛而复杂的概率和统计推理，并且具有一种与替代的“汇编语言”测度理论基础相比，使用更为简便的图示基础。自然语言短语上的概率分布在马尔科夫范畴中被编码为箭头。

$$f : \psi \rightarrow \text{Charles Darwin wrote}$$

然而，一个即时的困境是，仅靠一个 LLM 马尔科夫类别本身无法识别该释义

$$g : \mu \rightarrow \text{Charles Darwin is the author of}$$

应被视为与上一个箭头“同构”。换句话说，在 LLM 马尔科夫范畴中，存在许多在语法上看似不同但在语义上相同的对象。我们通过将范畴论的同伦技术引入 LLMs，特别是使用范畴论技术对它们进行建模来解决这一基本问题。

2 LLMs 中的同伦

同伦意味着决定两个任意对象在某种意义上是否等价。例如，如果存在一个连续映射 $h : [0, 1] \times X \rightarrow \mathbb{R}$ ，使得 $h(0, x) = f(x)$ 和 $h(1, x) = g(x)$ ，那么两个实值函数 $f, g : X \rightarrow \mathbb{R}$ 被认为是同伦。如果你想象绘制 f 和 g 的图，那么脑海中的图像就是平滑地扰动 f 的图，使其与 g 的图对齐。我们希望在捕捉“语法”和“语义”的 LLM 范畴上定义同伦（定义 ?? 到定义 ??）。Quillen [1967] 提出了模型类别作为在任何范畴中进行抽象同伦的一种方法，我们将在第 7.3 节进行回顾。我们在第 7 节中证明了 LLMs 定义了模型范畴。

同伦的概念最初在代数拓扑中出现，用来描述何时两个空间实质上是相同的，例如，一个咖啡杯如同一个甜甜圈，因为它们都有一个孔，并且它们在拓扑上与“两个圆的楔和”完全不同，后者看起来像一个八字形。通常，提升图 [Gavrilovich, 2017] 用于定义同伦（参见图 1）。? 使用提升图来在关系数据库中制定查询。在那个应用中， $X = \text{我在飞} \dots$ 类似于一个像 SQL 这样的语言中的数据库查询， Y 是我们想要从中获得答案的数据库（它可能包含查询的所有可能完成），而提升代表查询的潜在答案。

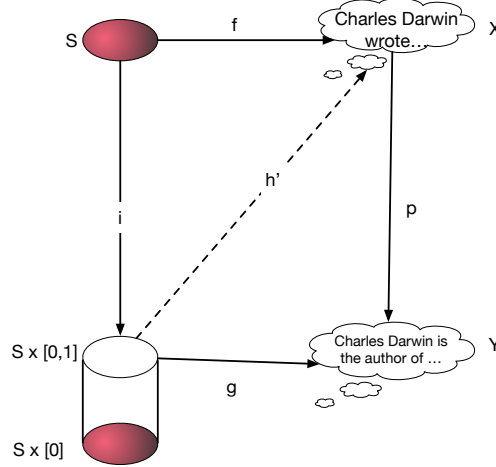


Figure 1: 用于 LLM 的提升图：任何这样的交换图如果它的复合态射为 h' ，则有一个“提升” $g \circ i = p \circ f$ 。应用于 LLM，它询问文本句子之间的相似性，即一个片段是否是另一个片段的等效改述，是否可以在拓扑“意义”上解释为当圆在单位区间路径 $[0, 1]$ 上移动时如何适应圆柱。

3 大语言模型的范畴论

我们引入了范畴理论的框架 MacLane [1971]，强调其在建模大型语言模型（LLMs）中的应用。简单来说，一个范畴 C 就是一个对象 $c, d \in C$ （任意复杂）的集合和对于每一对对象 $c, d \in C$ 的一个箭头 $C(c, d)$ 的集合。什么构成一个对象是可以商榷的：在 LLMs 的情况下，正如我们将在下面看到的，自然的定义对象的方法是其定义为“tokens”（大致理解为自然语言中的单词）。自然地，单词可以连接成句子，这引导我们使用对称幺半范畴 Fong and Spivak [2018]。在我们介绍这些之前，我们回顾一些将 Transformer 模型建模为范畴的简单方法，然后回顾 [Bradley et al., 2022] 利用富范畴来建模 LLMs 的下一个词分布的工作。在下一部分，我们将正式介绍 LLM 马尔可夫范畴，这将是我们在本文中的主要框架。

Definition 1. category 是一个图 $\mathcal{G}$ ，并带有两个附加的函数： $\text{id} : O \rightarrow A$ ，将每个对象 $c \in C$ 映射为一个箭头 id_c ；以及 $\circ : A \times_O A \rightarrow A$ ，将每对可合成的态射 $\langle f, g \rangle$ 映射为它们的复合态射 $g \circ f$ 。

3.1 Transformer 中的注意力机制

Transformer 结构的核心是注意力分数的计算。对某些类型的注意力分数的详细回顾在 [Chaudhari et al., 2021] 中进行了。在忽略许多细节的情况下，整体思路是计算嵌入在欧几里得空间中的两个标记之间的某种“点积”。我们提供了一些方法的简要回顾，然后定义了一类 Transformer 模型。特别值得注意的是，Transformer 本质上计算置换等变函数 [?]。为了将这种基本对称的注意力分数计算扩展到非对称的自然语言应用，输入以某种方式加权，在原始 Transformer 模型中使用正弦函数 [?]，后来在更一般的相对位置编码框架中进行了研究 [?]。注意力模型的总体结构给出为

$$A(q, K, v) = \sum_i p(a(k_i, q)) \times v_i$$

，其中 K 和 V 是键值对， q 是查询， a 是对齐函数， p 是分布函数。某些示例对齐函数在 Table 1 中给出。

3.2 作为一个类别的 Transformer

为了帮助具体化主论文中更抽象的呈现，我们定义一个 Transformer 模型的范畴。很容易将我们下面的范畴模型扩展到生成式 AI 系统的其他构建模块，包括结构化状态空间序列模型 [Gu et al., 2022] 或图像扩散模型 [?]。与所有生成式 AI 系统一样，Transformer 模型的基本结构是由模块化组件组成的组合结构，每个组件在向量空间 $\mathbb{R}^{d \times n}$ 上计算一个置换等变函数，该向量空间由 n 长度的标记序列组成，每个标记嵌入在一个维度为 d 的空间中。我们可以定义一个可交换的图表来展示置换等变性质，如下所示。

首先，按照 [Fong et al., 2019]，我们可以通常将类型为 (n_1, n_2) 的神经网络层定义为子集 $C \subseteq [n_1] \times [n_2]$ ，其中 $n_1, n_2 \in \mathbb{N}$ 是自然数，以及 $[n] = \{1, \dots, n\}$ 。这些数字 n_1 和 n_2 用于定义每一层的输入和输出数量， C 是一组连接， $(i, j) \in C$ 表示在网络图中节点 i 与节点 j 相连。为每一层定义激活函数 $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ 是直接的，但基本上每个网络层定义了一个参数化函数 $I : \mathbb{R}^{|C|+n_2} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2}$ ，其中 $|C|$ 定义了每个网络边的边权重， $\mathbb{R}^{n_2}$ 系数编码了各个单元的偏置。

Function	Equation
similarity	$a(k_i, q) = \text{sim}(k_i, q)$
dot product	$a(k_i, q) = q^T k_i$
scaled dot product	$a(k_i, q) = \frac{q^T k_i}{\sqrt{d_k}}$
general	$a(k_i, q) = q^T W k_i$
biased general	$a(k_i, q) = k_i(Wq + b)$
activated general	$a(k_i, q) = \text{act}(q^T W k_i + b)$
generalized kernel	$a(k_i, q) = \phi(q)^T \phi(k_i)$

Table 1: 一些对齐函数 [Chaudhari et al., 2021]。

我们可以将这些专门化到 Transformer 模型中，特别是，要注意 Transformer 模型计算出专门类型的置换等变函数，如下方交换图所定义的（实际中，Transformer 模型添加了一个外部位置编码，例如 [?] 中定义的或后来的变种，例如相对位置嵌入 [?]，我们将在下面讨论）。

$$\begin{array}{ccccc}
 X & \xrightarrow{f} & Y & \xrightarrow{g} & Z \\
 \downarrow P & & \downarrow P & & \downarrow P \\
 XP & \xrightarrow{f} & YP & \xrightarrow{g} & ZP
 \end{array}$$

在上述交换图中，顶点是对象，箭头是定义 Transformer 块动作的态射。这里， $X \in \mathbb{R}^{d \times n}$ 是一个长度为 n 的 token 序列，维度为 d 。 P 是一个排列矩阵。由 Transformer 块计算的函数 f 满足 $f(XP) = f(X)P$ 的条件。该性质在上述图中通过设置 $Y = f(X)P$ 来定义，该设置可以通过两种方式实现：首先将输入通过矩阵 P 进行排列，然后应用 f ，或者通过

让我们更详细地理解 Transformer 模型的排列等变性质。虽然在文献中有许多不同的变体，但我们的 Transformer 模型符号是基于 [?]。基于下面描述的方法，这些变体可以很直接地适应我们的范畴框架。Transformer 模型本质上是组合的，这使得它们特别方便用范畴理论进行建模。

Definition 2. 一个 **变压器** 块是一个序列到序列的函数映射 $\mathbb{R}^{d \times n} \rightarrow \mathbb{R}^{d \times n}$ 。通常有两个层：一个自注意力层和一个逐字令牌的前馈层。我们假设令牌被嵌入到维度为 d 的空间中。具体来说，我们将 Transformer 块的输入 $X \in \mathbb{R}^{d \times n}$ 建模为 n 长度的 d 维度令牌序列，其中每个块计算定义为 $t^{h,m,r} : \mathbb{R}^{d \times n} : \mathbb{R}^{d \times n}$ 的以下函数：

$$\begin{aligned}
 \text{Attn}(X) &= X + \sum_{i=1}^h W_O^i W_V^i X \cdot \sigma[W_K^i X]^T W_Q^i X \\
 \text{FF}(X) &= \text{Attn}(X) + W_2 \cdot \text{ReLU}(W_1 \cdot \text{Attn}(X) + b_1 \mathbf{1}_n^T),
 \end{aligned}$$

其中 $W_O^i \in \mathbb{R}^{d \times n}$ 、 $W_K^i, W_Q^i, W_V^i \in \mathbb{R}^{d \times n}$ 、 $W_2 \in \mathbb{R}^{d \times r}$ 、 $W_1 \in \mathbb{R}^{r \times d}$ 和 $b_1 \in \mathbb{R}^r$ 。Transformer 块的输出为 $\text{FF}(X)$ 。按照惯例，“头”的数量是 h ，每个“头”的大小 m 是注意层的主要参数，而“隐藏”前馈层的大小为 r 。

Transformer 模型以输入对象 $X \in \mathbb{R}^{d \times n}$ 为输入，这些对象表示 n 长度的令牌序列在 d 维空间中，并作为表示排列等变函数的态射 $f : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^{d \times n}$ ，使得对于任何排列矩阵 $f(XP) = f(X)P$ 。? 展示了上述定义的 Transformer 模型所计算的实际函数是一个排列等变映射。

具体而言，我们定义了一个变压器的范畴 $\mathcal{C}_T$ ，其中对象是向量 $x \in \mathbb{R}^{d \times n}$ ，表示长度为 n 的 d 维符号序列，而可组合的箭头是由变压器块 $t^{h,m,r}$ 的组成构成的排列等变函数 $\mathcal{T}^{h,m,r}$ ，每个块有 h 个大小为 m 的头部，以及一个包含 r 个隐藏节点的前馈层。在一个范畴中，对象通过箭头或态射相互作用。在变压器模型的范畴 $\mathcal{C}_T$ 中，态射是等变映射 f ，通过这些映射，一个变压器模型块可以与另一个块组合。

为了克服仅限于排列等变函数的限制，在 Transformer 的实现中通常包括一个相对位置嵌入函数（RPE）[?]。

现在，我们根据下一个标记分布定义 LLM 的类别，并引入一个基于由 Khandelwal et al. [2020] 提出的 k -NN LLM 模型的混合类别。为了定义 LLM 的更“语义”的类别，我们使用丰富的余前层：这些是 Yoneda

嵌入，其中一个标记 x 被映射到所有可能的补全集 $\mathcal{L}(x, -)$ ——在单位区间 $[0, 1]$ 上丰富的反预层。Yoneda 嵌入通常将一个对象 x 映射到一个协变的集合值功能 $\mathcal{L}(x, -)$ ，但在特殊情况下，例如对于 LLM，这个函子本身是一个丰富类别的对象，例如单位区间 $[0, 1]$ 上的对称单态预序。我们可以将 LLM 类别 $\mathcal{L}$ 的 Yoneda 嵌入视为定义句子意义的“语义”类别。最后，让我们基于 Khandelwal et al. [2020] 定义的 k -NN 大型语言模型定义一种新的类别类型（详情请参见第 3.3 节）。Khandelwal et al. [2020] 表明， k -NN LLM 由于更准确地模拟等价短语（如 查尔斯·狄更斯写到 和 查尔斯·狄更斯是作者）的下一个标记分布，而比常规 LLM 具有更高的困惑度得分。

3.3 使用最近邻语言模型的合成标记概率

Khandelwal et al. [2020] 提出了 k -NN 最近邻语言模型，它结合了由每个训练示例 $\langle c_i, w_i \rangle$ 的键-查询值对 $\langle k_i, v_i \rangle$ 构建的合成概率，其中 c_i 是一个上下文短语，例如 查尔斯·狄更斯写道，而 $f(\cdot)$ 是一个函数，它将上下文 c 映射为由预训练的 Transformer 模型计算的固定长度向量表示。例如， $f(c)$ 可以将 c 映射为由某个自注意层（我们在上一节中描述过）输出的固定长度向量。数据存储 $\langle \mathcal{K}, \mathcal{V} \rangle$ 是从 $\mathcal{D}$ 中所有训练示例构建的所有键值对的集合：

$$\langle \mathcal{K}, \mathcal{V} \rangle = \{(f(c_i), w_i) | (c_i, w_i) \in \mathcal{D}\}$$

在测试时，由 LLM 学习的语言模型生成输出分布 $p_{LM}(y|x)$ ，并使用 $f(x)$ 查询数据存储 $\langle \mathcal{K}, \mathcal{V} \rangle$ ，利用某个距离函数 $d(\cdot, \cdot)$ 检索其 k 最近邻 $\mathcal{N}$ 。然后使用其负距离的 softmax 在邻居上计算一个合成概率：

$$p_{KNN}(y|x) \propto \sum_{(k_i, v_i) \in \mathcal{N}} \mathbf{1}_{y=v_i} \exp(-d(k_i, f(x)))$$

然后对于 k -NN LLM 给出最终分布

$$p(y|x) = \lambda p_{KNN}(y|x) + (1 - \lambda) p_{LM}(y|x)$$

因此，我们可以以这种方式为 k -NN LLM 定义一个新的 LLM 类别，称为 $\mathcal{L}_{knn}$ ，下一词的概率是学习概率与最近邻预测线性插值的结合。Khandelwal et al. [2020] 显示了在大型数据集上的一系列实证结果，表明这种插值概率提高了 LLM 具有更高困惑度评分的预测能力。

3.4 端和余端的演算

我们可以定义自然变换空间，这是两个函子 $F, G : \mathcal{L}_s \Rightarrow \mathcal{L}_t$ 之间的变换，这些函子在 LLMs 的“语法”范畴和它的“语义”范畴之间进行映射，其方式是用 (co) 端的微积分来描述 [Loregian, 2021]。楔形物范畴由一组由双函子 $F : \mathcal{C}^{op} \times \mathcal{C} \rightarrow \mathcal{D}$ 构成的对象组成，以及称为双自然变换（对角自然变换的缩写）的每对双函子 F, G 之间的箭头集合定义。我们将在下面看到，楔形范畴中的初始和终端对象对应于由 Yoneda 首次表述的一个美丽概念，称为共同终点或结束 [?]。Loregian [2021] 对共端微积分有一个极好的处理方法，我们将在下面使用它。

Definition 3. 给定一对双函子 $F, G : \mathcal{C}^{op} \times \mathcal{C} \rightarrow \mathcal{D}$ ，定义自然双变换如下：

$$\begin{array}{ccc}
 & F(c', c) & \\
 \swarrow F(f, c) & & \searrow F(c', f) \\
 F(c, c) & & F(c', c') \\
 \downarrow G(c, c) & & \downarrow G(c', c') \\
 G(c, c) & & G(c', c') \\
 \swarrow G(c, f) & & \searrow G(f, c') \\
 & G(c, c') &
 \end{array}$$

正如 [Loregian, 2021] 观察到的那样，就像一个自然变换通过填补它们对终端范畴上态射 Ff 和 Fg 作用之间的空缺来在两个常规函子 F 和 G 之间进行插值一样，一个双自然变换“填补了”上面六边形顶部和六边形底部之间的空隙。

我们可以通过它将所有东西映射到的对象来定义一个常数双函子 $\Delta_d : \mathcal{C}^{op} \times \mathcal{C} \rightarrow \mathcal{D}$ ，即输入的对象对 $(c, c') \rightarrow d$ 都映射到对象 $d \in \mathcal{D}$ ，并且两个输入态射 $(f, f') \rightarrow \mathbf{1}_d$ 都映射到 d 上的恒等态射。我们现在可以定义楔形和边的相邻边。

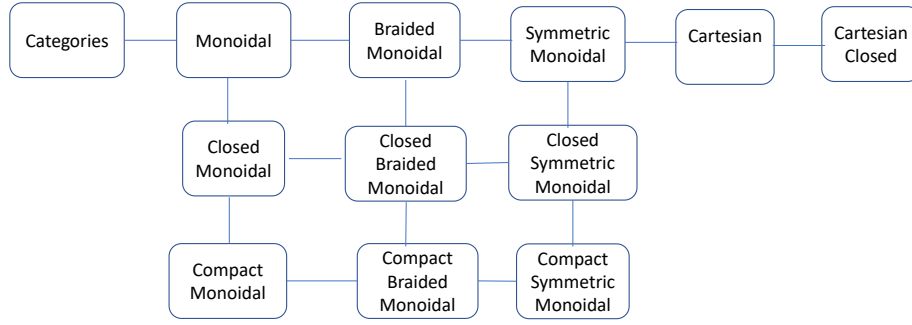


Figure 2: 么半范畴的结构。

Definition 4. 对于双函子 $F: C^{\text{op}} \times C \Rightarrow \mathcal{D}$ 的 **楔形** 是一个从对象 $d \in \mathcal{D}$ 上的常函子到 F 的自然变换 $\Delta_d \rightarrow F$ 。同样地，我们可以通过自然变换 $P \Rightarrow \Delta_d$ 来为双函子 F 定义一个 **边缘联合**。

我们现在可以定义一个类楔形体，其中的每个对象都是楔子，对于箭头，我们选择使下图可交换的共域范畴中的箭头。

有了这些定义，我们可以再次用始对象和终对象来定义通用性质。

值得注意的是，可以正式证明概率定义了类别 [Avery, 2016] 的终端，而流行的降维方法如 UMAP [McInnes et al., 2018] 所实现的数据集的拓扑嵌入，对应于联合项 [MacLane, 1971]。这些联系表明楔形和反楔形的类别在全面理解 LLMs 中具有重要的规范意义。

3.5 增强的 LLM 类别

接下来让我们为 LLM 定义一些增强范畴的性质。通常，我们对 $\mathcal{V}$ 是闭合的情形感兴趣，从而它可以在其自身上进行增强，此时上述关系简化为 $C(x, y) \leq \mathcal{V}(fx, fy)$ 。

Definition 5. 如果 C 是一个在交换么半序 $\mathcal{V}$ 上富集的范畴，那么 **丰富余前层** 是一个满足 $C(x, y) \leq \mathcal{V}(fx, fy)$ 的函数 $f: C \rightarrow \mathcal{V}$ ，对范畴 C 中的所有对象 x 和 y 。对于 LLMs 的范畴， y 将表示文本片段 x 的扩展，而 $\mathcal{V}(fx, fy)$ 定义扩展的概率。

我们实际上可以在富集的 LLM 范畴上定义一个拓扑 [Mac Lane and Moerdijk, 1992]，因为它具有内部指数对象，正如每个拓扑都包含这样的对象。我们可以使用以下定义（其中富集端的概念在第 3.4 节中有更详细的定义 - 另见 [Loregian, 2021]）。对于 $\mathcal{V}$ 是单位区间上的对称单模序的特殊情况，这个富集的余末端变得简单得多。内部同构对象 $[x, y] \in \mathcal{V}$ 定义为截断除法，上述表达式大大简化为以下：这些定义和结果本质上可以被视为广义度量空间中度量 Yoneda 引理的一个特例，其中单位区间 $[0, 1]$ 通过 $-\log$ 映射映射到 $[0, \infty]$ 中的非负距离。

4 单体 LLM 范畴

在本节中，我们提出了一个基于称为马尔可夫范畴 [Fritz, 2020] 的对称单体范畴表示来分析 LLM 的范畴框架。我们首先简要解释对称单体范畴，然后更正式地介绍马尔可夫范畴。我们将我们的方法与先前有关 LLM 范畴表示的工作进行比较 [Bradley et al., 2022]。

4.1 对称单（对称么）范畴

因果性的范畴模型 [Fong, 2012, Fritz and Klingler, 2023, Jacobs et al., 2018, Mahadevan, 2023] 通常是在对称单范畴上定义的，下面我们简要回顾一下 [MacLane, 1971, Richter, 2020]（见图 2）。

Definition 6. 一个 **单积范畴** 是一个类别 C ，并且具有函子 $\otimes: C \times C \rightarrow C$ ， C 的单位对象 e ，以及定义为

$$\begin{aligned} \alpha_{C_1, C_2, C_3}: C_1 \otimes (C_2 \otimes C_3) &\cong (C_1 \otimes C_2) \otimes C_3, \\ \lambda_C: e \otimes C &\cong C, \\ \rho: C \otimes e &\cong C, \end{aligned}$$

的自然同构 α, λ, ρ 。

自然同构必须满足被称为“五边形”和“三角形”图的相干性条件 [MacLane, 1971]。在 [MacLane, 1971] 中显示的一个重要结果是，这些相干性条件保证所有格式良好的图都必须交换。有许多自然的幺半范畴例子，最简单的是有限集的范畴，在 [Fritz, 2020] 中被称为有限集，其中每个对象 C 是一个集合，张量积 $\otimes$ 是集合的笛卡尔积，函数作为箭头作用。确定性因果模型可以在范畴有限集中表述。其他例子包括以关系作为态射的集合范畴，以及希尔伯特空间的范畴 [Heunen and Vicary, 2019]。马尔可夫范畴 [Fritz and Klingler, 2023] 是幺半范畴，其中单位元 e 也是终对象，意味着每个对象 X 都有一个唯一的“删除”态射 $d_e : X \rightarrow e$ 关联。[Fritz, 2020] 显示它们构成了概率和统计推理的统一基础。

Definition 7. 一个 **对称单部范畴** 是一个单控范畴 $(C, \otimes, e, \alpha, \lambda, \rho)$ ，并带有自然同构

$$\tau_{C_1, C_2} : C_1 \otimes C_2 \cong C_2 \otimes C_1, \text{ for all objects } C_1, C_2$$

，其中 τ 满足额外条件：对于所有对象 C_1, C_2 $\tau_{C_2, C_1} \circ \tau_{C_1, C_2} \cong 1_{C_1 \otimes C_2}$ ，以及对于所有对象 C ， $\rho_C = \lambda_C \circ \tau_{C, e} : C \otimes e \cong C$ 。

需要一个额外的六边形公理来确保 τ 自然同构与 α 兼容。在马尔可夫范畴中， τ 算子被称为“交换” [Fritz, 2020]。在人工智能中大多数感兴趣的情况下，对称幺半范畴是在一些便利的基础范畴丰富的上，包括向量空间，或预序例如单位区间 $[0, 1]$ ，其中如果且仅当 $a \leq b$ ，则从 $a \rightarrow b$ 存在唯一的态射。5

Definition 8. 一个 **V-丰富的范畴** 由一个常规类别 C 组成，使得对于 C 中的每一对对象 x 和 y ，其态射 $C(x, y) \in \mathcal{V}$ ，通常被称为一个 $\mathcal{V}$ -hom 对象。当 $(\text{calV}, \leq, \otimes, 1)$ 是一个交换的幺半序时，我们有以下条件

- $1 \leq C(x, x)$
- $C(y, z) \otimes C(x, y) \leq C(x, z)$

4.2 LLM 马尔可夫范畴

在本节中，我们介绍了通过马克洛夫范畴 [Fritz, 2020] 建模 LLMs 的框架，这些范畴被提出作为因果推理、概率和统计的统一分类模型。它们是对称单oidal 范畴，这在上文中已进行回顾，结合在每个对象上的共单oidal 结构。重要的是，马克洛夫范畴是半笛卡尔的，因为它们不使用统一的复制，而是包含一个由确定态射定义的笛卡尔子范畴（见下文）。我们简要回顾了马克洛夫范畴，省略的重大附加细节可以在 [Cho and Jacobs, 2019, Fritz, 2020, Fritz and Klingler, 2023] 中找到。方程是用弦图的图示语言书写的，这可以显示为一种形式语言，相当于写下代数方程 [Selinger, 2010]。

我们可以简单地将由 LLM 计算的下一个 token 的分布建模为马尔可夫范畴中的态射。我们在下面更正式地定义 LLM 马尔可夫范畴。

Definition 9. 一个 **LLM Markov 范畴** C_{LLM} [Fritz, 2020] 是一个对称的单部范畴，其中每个对象 $X \in C_{\text{LLM}}$ 定义一个 LLM 标记，并配备有由共乘 $\text{copy}_X : X \rightarrow X \otimes X$ 和伴随单元 $\text{del}_X : X \rightarrow I$ 给出的交换余半群结构，在字符串图中表示为

$$\text{copy}_X = \begin{array}{c} \text{---} \cup \\ \text{---} \end{array} \quad \text{del}_X = \begin{array}{c} \bullet \\ | \end{array}$$

(1)

，并满足交换余半群方程，

$$\begin{array}{c} \text{---} \cup \\ \text{---} \end{array} \bullet = \begin{array}{c} \bullet \\ | \end{array} \cup \text{---}$$

(2)

$$\bullet \cup \text{---} = \text{---} \cup \bullet = \begin{array}{c} \text{---} \cup \\ \text{---} \end{array} \quad \begin{array}{c} \cup \\ \bullet \end{array} = \begin{array}{c} \cup \\ \bullet \end{array}$$

(3)

以及与单部结构的兼容性,

$$\begin{array}{c} \bullet \\ | \\ X \otimes Y \end{array} = \begin{array}{c} \bullet \\ | \\ X \end{array} \begin{array}{c} \bullet \\ | \\ Y \end{array}$$

$$\begin{array}{c} X \otimes Y \quad X \otimes Y \\ \diagdown \quad \diagup \\ \bullet \\ | \\ X \otimes Y \end{array} = \begin{array}{c} X \quad Y \quad X \quad Y \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \\ \bullet \quad \bullet \\ | \quad | \\ X \quad Y \end{array}$$

(4)

和 del 的自然性, 这意味着

$$\begin{array}{c} \bullet \\ | \\ \boxed{f} \\ | \\ \bullet \end{array} = \begin{array}{c} \bullet \\ | \\ \bullet \end{array}$$

(5)

对于每个态射 f 。

让我们简要解释这些定义。 $\text{del}_X : X \rightarrow I$ 本质上像是在概率分布上积分, 总是会得到 1。因此, I , 即张量积的单位, 是仿射 CDU 和马尔可夫范畴中的终端对象。贝叶斯法则变成了一个 解体规则 [Cho and Jacobs, 2019], 这个只在具有条件的马尔可夫范畴中可用 (即, 在那里可以范畴地细化 $P(y|x)$ 条件分布)。注意, 在定义为实数上的可测函数的随机变量的连续情况下, 必须非常小心地定义条件 [Halmos, 1974]。 copy_X 过程是统一的和确定性的, 意思是如果你对两个变量 $X \otimes Y$ 进行张量积然后复制得到的对象, 那与先将 X 复制到 $X \otimes X$ 并将 Y 复制到 $Y \otimes Y$, 然后进行张量积以及交换操作 (见方程 9) 完全相同。只有 del_X 以“统一”的方式发挥作用, 意思是如果你使用某个函数 f 处理变量 X 然后删除 $f(X)$ (意思是对其进行边缘化), 那相当于简单地删除 X 。然而, copy_X 并不是这样定义的, 我们在下面讨论这种微妙性, 因为它对于理解为什么马尔可夫范畴是半笛卡尔的很重要。为了将它们转换成拓扑斯, 我们需要结果是笛卡尔封闭的, 这也是为什么我们需要使用宫田引理构建前幺范畴来保证获得一个拓扑斯。

4.3 马尔可夫范畴中的笛卡尔结构

我们现在讨论 Markov 范畴中的一个 Cartesian 范畴子类别, 这个子类别涉及均匀的 copy_X 和 del_X 态射。Markov 范畴的一个基本性质是它们是半笛卡尔的, 因为单位对象也是终端对象。但是, 对于这些复制和删除算子的建模方式, 会出现一些细微问题, 如下所述。

Definition 10. 一个对称单 *oidal* 范畴 $\mathcal{C}$ 是 **笛卡尔**, 如果张量积 $\otimes$ 是范畴积。

如果 $\mathcal{C}$ 和 $\mathcal{D}$ 是对称单部范畴, 那么函子 $F : \mathcal{C} \rightarrow \mathcal{D}$ 是单体的, 如果张量积最多被保持为一致的自然同构。 F 是严格单体的, 如果所有的单体结构都被精确地保持, 包括 $\otimes$ 、单位对象 I 、对称性、结合性和单位自然同构。用 **MON** 表示以严格函子为箭头的对称单部范畴的范畴。让我们回顾一下由 Heunen and Vicary [2019] 提供的基本定义, 这将进一步阐明仿射 CDU 和马尔可夫范畴中的笛卡尔结构。

Definition 11. 在所有对称单 *oidal* 范畴的环境范畴 **MON** 中, *comonoid* 的子范畴 **共同** 被定义为任意特定范畴 $\mathcal{C}$ 的“余代数”对象 $(X, \text{copy}_X, \text{del}_X)$ 的集合, 其中 X 位于 $\mathcal{C}$, 并且箭头被定义为从 $(X, \text{copy}_X, \text{del}_X)$ 到 $(Y, \text{copy}_Y, \text{del}_Y)$ 以一致方式作用的 *comonoid* 同态, 也就是说, 如果 $f : X \rightarrow Y$ 是 $\mathcal{C}$ 中的任何态射, 那么:

$$\begin{aligned} (f \otimes f) \circ \text{copy}_X &= \text{copy}_Y \circ f \\ \text{del}_Y \circ f &= \text{del}_X \end{aligned}$$

Heunen and Vicary [2019] 在范畴 **coMON** 中定义了“均匀复制和删除”的过程, 我们现在将其与马尔可夫范畴联系起来。值得强调的一个微妙区别是, 在定义 9 中, 在马尔可夫范畴中, 只有 del_X 是“均匀的”, 而不是 copy_X , 根据 Heunen and Vicary [2019] 所定义。这种区别可以通过适当地修改与复制相关的 PROP 映射的定义, 在像马尔可夫范畴一样的半笛卡尔的 cPROP 范畴中建模。

Definition 12. 对称单轴范畴 C 允许 **均匀删除**，如果对于余么半对象的子范畴 C_{coMON} 中的所有对象，有一个自然变换 $e_X: X \xrightarrow{e_X} I$ ，如公式 9 所示。

这个条件被 Cho and Jacobs [2019] 称为箭头 e_X 上的因果关系条件。基本上，它表示如果你处理某些对象然后丢弃它，这等同于在不处理的情况下丢弃它。

Theorem 1. [Heunen and Vicary, 2019] 一个对称单积范畴 C 如果且仅当 I 是终端对象时，具有统一删除。

正如在 [Fritz, 2020] 中所指出的，这一性质适用于马尔可夫范畴，并且在 [Heunen and Vicary, 2019] 中给出了一个简单的图示证明。

Definition 13. 如果一个对称的单核范畴 C 存在 **均匀复制**，则存在一个自然变换 $copy_X: X \rightarrow X \otimes X$ 满足方程 9 和方程 9。

我们现在可以陈述一个在 [Heunen and Vicary, 2019] 中证明的重要结果（定理 4.28），该结果与 Fox [1976] 早些时候展示的更一般的结果相关。

Theorem 2. 对称单体范畴 C 满足以下等价条件。

- 类别 C 是 **笛卡尔**，张量积 $\otimes$ 由范畴积给出，张量单位由终端对象给出。
- 对称单 *oidal* 范畴 C 具有 **均匀复制和删除**，并且方程 9 成立。

如 Fritz [2020] 所述，并非所有 Markov 范畴都是笛卡尔的，因为它们 **拷贝** X 不是一致的，而只有 **删除** X 是。例如，考虑范畴 **FinStoch**，其中一个联合分布由态射 $\psi: I \rightarrow X \otimes Y$ 指定。在这种情况下，边缘分布可以构成复合态射

$$\begin{aligned} I &\xrightarrow{\psi} X \otimes Y \xrightarrow{\text{del}_Y} X \\ I &\xrightarrow{\psi} X \otimes Y \xrightarrow{\text{del}_X} Y \end{aligned}$$

。但是，要求这情况下 $\otimes$ 是范畴乘积意味着上述复合定义的边缘分布必须与联合分布一一对应。

4.4 单斜范畴的量子自然语言处理

我们简要回顾了先前在使用对称么半群范畴 [Coecke, 2020, Coecke et al., 2010] 建模语言结合量子计算 [Coecke and Kissinger, 2017] 方面的工作。这种方法已经在称为 **Lambeck** [Kartsaklis et al., 2021] 的 **DiscoPY** Python 包中实现。与我们的工作相比，其显著区别在于其工作中没有解决范畴同伦问题。该方法基于由 Lambek 预群语法定义的语言的组合理论，即由部分序么群 [Lambek, 2006] 定义的一种范畴。

在 Python 包 **lambeck** 中，自然语言句子使用统计 CCG 解析器 [?] 进行解析。其中，CCG 表示组合范畴语法 [Lewis and Steedman, 2014]。给定一个长度为 N 的句子 x ，CCG 树 τ 的概率计算为超标记（或类别） c_i 的概率的乘积，它们本质上是分解模型：

$$P(\tau|x) = \prod_{i \in [1, N]} P_{\text{tag}}(c_i|x)$$

有一种高效的解析方法，称为 A^* 解析 [Klein and Manning, 2003]，可以用于实现 CCG 的解析。**lambeck** 将 CCG 解析树转换为字符串图 [Selinger, 2010]，这是一种图示化的方法，用于定义句子的组成结构。经过一些预处理后，字符串图被转换为量子电路。这项工作中有很多有趣的想法，但进行更详细的讨论超出了本文的范围。主要要点是，我们接下来讨论的同伦问题没有在这里解决。一个有趣的未来研究问题是，将模型类别同伦框架与如 **lambeck** 这样的软件包结合起来。

5 范畴同伦论

在本节介绍模型类别 [Quillen, 1967, Hovey, 2020] 的框架之前，我们需要提供一些关于同伦理论 [Quillen, 1967] 的背景。为了重申我们最初的问题，如果我们将句子片段视为 LLM 马尔可夫范畴中的对象，我们立即遇到的一个问题是，其中许多对象应该被视为相同，即便它们不是同构的。同伦在拓扑学的研究中出现，作为一种手段来理解何时两个形状大致相同（即在同伦下）。我们通过提升图来解释捕获这一概念的想法。

5.1 范畴中的提升问题

提升问题在数学的众多领域中提供了定义基本概念优雅方法 [Gavrilovich, 2017]。例如，单射和满射函数的概念、拓扑中的分离概念以及许多其他基本构造都可以表述为提升问题的解决方案。使用 SQL 等语言在关系数据库中的数据库查询可以形式化为提升问题 [?]。提升问题定义了将结构分解为更简单部分并重新组合的方法。我们希望更抽象地理解 LLM，了解提升图及其在模型范畴和单纯集合中的应用。

Definition 14. 令 C 为一个范畴。在 C 中，**lifting problem** 是一个在 C 中的交换图 σ 。

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & & \downarrow p \\ B & \xrightarrow{v} & Y \end{array}$$

Definition 15. 令 C 为一个范畴。**solution to a lifting problem** 在 C 中是一个满足以下图示中所指示的 $p \circ h = v$ 和 $h \circ f = \mu$ 的在 C 中的态射 $h: B \rightarrow X$ 。

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ B & \xrightarrow{v} & Y \end{array}$$

Definition 16. 令 C 为一个范畴。如果在 C 中给定了两个态射 $f: A \rightarrow B$ 和 $p: X \rightarrow Y$ ，我们说 f 相对于 p 具有 **left lifting property**，或者说 p 相对于 f 具有 **right lifting property**，如果对于每一对满足方程 $p \circ \mu = v \circ f$ 的态射 $\mu: A \rightarrow X$ 和 $v: B \rightarrow Y$ ，下图中所示的关联提升问题成立。

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ B & \xrightarrow{v} & Y \end{array}$$

承认由映射 $h: B \rightarrow X$ 给出的一个满足 $p \circ h = v$ 和 $h \circ f = \mu$ 的解。

“分数范畴” [Gabriel et al., 1967] 旨在形式化这样一个概念，即在 LLM 范畴中存在弱等价——即意义相同的句子——需要在构造的同伦范畴中变为同构。将这些弱等价转化为另一个同伦范畴中的同构，这个问题可通过以下的普遍性质 [Borceux, 1994] 形式化：

Definition 17. 考虑一个 LLM 范畴 $\mathcal{L}$ 和一个箭头类 Σ 。当存在一个范畴 $\mathcal{L}(\Sigma^{-1})$ 和一个函子 $\phi: \mathcal{L} \rightarrow \mathcal{L}(\Sigma^{-1})$ 具有以下普遍性质时，称 **category of fractions** $\mathcal{L}(\Sigma^{-1})$ 存在：

1. $\forall f \in \Sigma, \phi(f)$ 是一个同构。
2. 如果 $\mathcal{L}'$ 是一个 LLM 范畴，并且 $F: \mathcal{L} \rightarrow \mathcal{L}'$ 是一个函子，使得对于所有态射 $f \in \Sigma$ ， $F(f)$ 是一个同构，那么存在一个唯一的函子 $G: \mathcal{L}(\Sigma^{-1}) \rightarrow \mathcal{L}'$ 使得 $G \circ \phi = F$ 。

6 LLM 范畴中的同伦

Definition 18. 设 X 和 Y 是 LLM 马尔可夫范畴中的两个对象，并且假设给出了一个从 f_0 到 f_1 的态射对 $f_0, f_1: X \rightarrow Y$ 。一个从 f_0 到 f_1 的 **homotopy** 是一个满足 $f_0 = h|_{0 \times X}$ 和 $f_1 = h|_{1 \times X}$ 的态射 $h: \Delta^1 \times X \rightarrow Y$ 。

我们现在介绍一种正式的方法来定义由 LLM 使用一个范畴的神经定义的拓扑空间的代数不变量。如 [Segal, 1968] 中所示，一个范畴的神经是一个范畴作为单纯对象的全嵌入。在单纯集合的意义上，神经就是 n 单纯形的集合，它由长度为 $n \geq 0$ 的可组合箭头对定义。

我们现在可以使用抽象同伦以一种更抽象的方式定义自然语言句子的等价类。在 LLM 范畴 $\mathcal{L}$ 中，如果两个对象 C 和 C' 通过如图 ?? 所示的锯齿形结构相关联，则它们在同一个等价类 $\mathcal{E}$ 中。这是一类路径分量上的结果的特殊情况（见 [Richter, 2020]）。

这种 LLM 范畴 $\mathcal{L}$ 与其拓扑实现 $\mathcal{BL}$ 之间的关系为我们提供了在图 1 中所需的 LLM 的同伦特征。我们可以如下定义一个 LLM 范畴 $\mathcal{L}$ 的高阶同伦群。

根据定义 ?? 或定义 ??，我们知道一个 LLM 范畴 $\mathcal{L}$ 是在一个对称单积范畴上丰富的，这为其同伦结构的研究提供了大量已知的结果。

该定理的证明基于简单的图表追踪，并依赖于（小）对称单积范畴的标准结果。

6.1 LLM 代数 K 理论

最后，我们基于上述结果利用 Grayson-Quillen K-理论 [Quillen, 1973, Grayson, 1976] 来定义 LLM 同伦。

Definition 19. 设 $(\mathcal{L}, \otimes, e, \tau)$ 是一个 LLM Markov 范畴。记 $\mathcal{L}^{-1}\mathcal{L}$ 为范畴，其对象是 $\mathcal{L}$ 的对象对。 $\mathcal{L}^{-1}\mathcal{L}$ 中从 (C_1, D_1) 到 (C_2, D_2) 的态射是态射对的等价类

$$(f : C_1 \otimes E \rightarrow C_2, g : D_1 \otimes E \rightarrow D_2)$$

其中 E 是 $\mathcal{L}$ 的一个对象。这种态射对等价于

$$(f' : C'_1 \otimes E' \rightarrow C_2, g' : D_1 \otimes E' \rightarrow D_2)$$

如果存在一个同构 $h \in \mathcal{L}(E, E')$ ，使得以下图表交换：

$$\begin{array}{ccc} (C_1 \otimes E, D_1 \otimes E) & \xrightarrow{(1_{C_1} \otimes h, 1_{D_1} \times h)} & (C_1 \otimes E', D_1 \otimes E') \\ & \searrow (f, g) \quad \swarrow (f', g') & \\ & (C_2, D_2) & \end{array}$$

范畴 $\mathcal{L}^{-1}\mathcal{L}$ 称为 $\mathcal{L}$ 的 **Grayson-Quillen** 构造。

Theorem 3. 范畴 $\mathcal{L}^{-1}\mathcal{L}$ 也是对称的单积范畴，并且存在一个弛单积函子 $j : \mathcal{L} \rightarrow \mathcal{L}^{-1}\mathcal{L}$ ，而 $\pi_0(\mathcal{BL}^{-1}\mathcal{L})$ 是一个阿贝尔群。

证明： 该证明直接来源于 [Richter, 2020] 中的一般性结果，引理 13.3.2。 $\square$

Definition 20. 与 LLM 类别 $\mathcal{L}$ 相关联的 **K-理论** 空间是分类空间 $\mathcal{KL} = \mathcal{BL}^{-1}\mathcal{L}$ ，其中 $\mathcal{L}$ 的 n^{th} K-群是其 n^{th} 同伦群 $\pi_n \mathcal{BL}^{-1}\mathcal{L}$ 。具体来说，基本群 $\pi_0(\mathcal{BL}^{-1}\mathcal{L})$ 是由路径分量 $\pi_0(\mathcal{L})$ 诱导的阿贝尔么半群的 Grothendieck 群完备化。

现在让我们将该过程与 LLM 等价类的 0^{th} 同伦群连接起来。

Theorem 4. LLM 类别 $\mathcal{L}$ 的高代数 K 理论定义为

$$K_0(\mathcal{L}) = \pi_0(\mathcal{KL}) \simeq G_0(\pi_0(\mathcal{L}))$$

证明： 该证明直接来源于在任何对称单弱范畴中都成立的更一般性结果（见 [Richter, 2020] 中的引理 13.3.4）。 $\square$

6.2 LLM 群体

群胚是其态射可逆的范畴。我们可以根据其诱导群胚的分类空间来表征 LLM 等价类。

Definition 21. 定义 LLM 范畴为范畴 $\mathcal{L}_G$ ，其对象定义为范畴中连通路径的等价类，其可逆态射对应于将一个等价类映射回自身的可逆边。

Theorem 5. LLM 范畴群的分类空间被定义为

$$\mathcal{BL}_G = \bigsqcup_i \mathcal{BL}_G^i$$

，其中不相交的索引 i 跨越等价类。

证明： 该证明直接来源于 [Richter, 2020] 中关于群胚分类空间的一般性结果。 $\square$

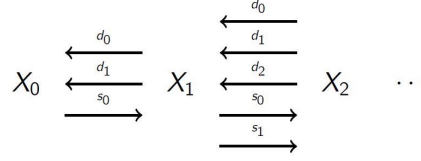


Figure 3: LLM 的单纯集合 [May, 1992] 表示是由一系列分级的集合组成，其中 0-单形是一个令牌集，1-单形是一组相邻令牌对，通常情况下， n -单形是一组长度为 n 的令牌序列。面映射和退化运算符在 n -单形及其较低或较高维度的单形之间映射。

7 将 LLM 视为模型范畴

为了在一个 LLM Markov 范畴上定义一个模型范畴，我们需要使用神经函子 [Segal, 1968] 将其转换为一个单纯集合。在此过程中有许多复杂的问题需要解决，这些问题超出了本文的介绍范围。首先，应用于一个范畴的标准神经函子产生一个单纯集合（直观上，每个 n -单形是所有 n -长度可组合态射的集合），但这种特殊类型的单纯集合是一个拟范畴 [Joyal, 2002]。这样的拟范畴仅允许填充内角，一般来说并不构成模型范畴。然而，这是我们分析的一个关键，当我们在一个 LLM Markov 范畴中有对象导致群结构时，因为它们本质上是对彼此的旁述，应反映相同的下一令牌分布，那么神经函子确实会产生一个模型范畴 [Riehl and Verity, 2019]。对于置换范畴，神经函子的对称幺半范畴的表现也是良好的 [Richter, 2020]。这里我们略过了一些细节，但它们对本文的主要目标并不重要。

7.1 单纯集合

本节的核心目标是展示 LLMs 定义了模型范畴（定理 6）。为了展示这一结果，我们需要引入单纯形集的范畴工具 [May, 1992]，这是一种拓扑空间的组合表示，它们是 CW 复形 [Munkres, 1984]。单纯形集推广了有向图，并构成高阶范畴理论的基础 [Boardman and Vogt, 1973, Joyal, 2002, Lurie, 2009]。

范畴 Δ 拥有非空序数 $[n] = \{0, 1, \dots, n\}$ 作为对象，且有序映射 $[m] \rightarrow [n]$ 作为箭头。 Δ 中的一个重要性质是任何箭头可以分解为一个单射和满射映射的复合，其中每个都可以分解为一个基本单射的序列 $\delta_i : [n] \rightarrow [n+1]$ ，被称为忽略 $i \in [n]$ 的 coface 映射，和一个基本满射的序列 $\sigma_i : [n] \rightarrow [n-1]$ ，被称为重复 $i \in [n]$ 的 co-degeneracy 映射。基本单形 $\Delta([n])$ 是全部映射到 $[n]$ 中的预层，即表现函子 $\Delta(-, [n])$ 。由 Yoneda 引理 [MacLane, 1971] 可以确认到， n -单形 $x \in X_n$ 可以通过对应的映射 $\Delta[n] \rightarrow X$ 来识别。 Δ 中的每个态射 $f : [n] \rightarrow [m]$ 会函子性地映射到 S 中的映射 $\Delta[m] \rightarrow \Delta[n]$ 。

在范畴 Δ 中，任何态射都可以定义为共同退化和余面算子的序列，其中共面算子 $\delta_i : [n-1] \rightarrow [n], 0 \leq i \leq n$ 定义为：

$$\delta_i(j) = \begin{cases} j, & \text{for } 0 \leq j \leq i-1 \\ j+1 & \text{for } i \leq j \leq n-1 \end{cases}$$

类似地，共退化算符 $\sigma_j : [n+1] \rightarrow [n]$ 被定义为

$$\sigma_j(k) = \begin{cases} j, & \text{for } 0 \leq k \leq j \\ k-1 & \text{for } j < k \leq n+1 \end{cases}$$

注意，在逆变映射下，余面映射变成面映射，余退化映射变成退化映射。也就是说，对于任何单纯对象（或集合） X_n ，我们有 $X(\delta_i) := d_i : X_n \rightarrow X_{n-1}$ ，同样地，还有 $X(\sigma_j) := s_j : X_{n-1} \rightarrow X_n$ 。

这些箭头的组合定义了某些众所周知的性质 [May, 1992, Richter, 2020]：

$$\begin{aligned} \delta_j \circ \delta_i &= \delta_i \circ \delta_{j-1}, \quad i < j \\ \sigma_j \circ \sigma_i &= \sigma_i \circ \sigma_{j+1}, \quad i \leq j \\ \sigma_j \circ \delta_i(j) &= \begin{cases} \sigma_i \circ \sigma_{j+1}, & \text{for } i < j \\ 1_{[n]} & \text{for } i = j, j+1 \\ \sigma_{i-1} \circ \sigma_j, & \text{for } i > j+1 \end{cases} \end{aligned}$$

Definition 22. 单纯形对象 是一个反变函子 $X : \Delta^{op} \rightarrow C$ ，其中 Δ 是上面介绍的序数范畴，而 C 是任意对象和箭头的范畴。

单纯集是从 Δ^{op} 到 **Sets** 范畴的反变函子。我们正在将这一概念推广到任何范畴，包括在第 ?? 节中定义的 LLM 范畴。每个 n 单纯形 $X[n]$ 实际上是范畴 C 的对象的集合，并且其态射是从 Δ 的弱序保留态射中映射而来的。

Example 1. 在范畴 C 中的单纯形对象 X 的“顶点”是 C 中的对象 X_0 ，“边” X_1 是它的箭头 $f: C_i \rightarrow C_j$ ，其中 C_i 和 C_j 是 C 中的对象。注意， X_0 是一个反变函子 $X: [0] \rightarrow C$ ，由于 $[0]$ 只有一个对象，该函子的效果是挑选出 C 中的对象。单纯形对象 $X_1: [1] \rightarrow C$ 。给定任意这样的箭头，面运算符 $d_0 f = C_j$ 和 $d_1 f = C_i$ 可以恢复每个箭头的起始和目标。此外，给定范畴 C 中的一个对象 X ，我们可以将退化运算符 $s_0 X$ 视为它的恒等态射 $1_X: X \rightarrow X$ 。

Example 2. 给定一个范畴 C ，我们可以定义一个 n -单纯形 X_n ，它是 C 中的一个单纯形对象，具有 the sequence: 。

$$X_n = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

面操作符 d_0 作用于 X_n 得到序列

$$d_0 X_n = C_1 \xrightarrow{f_2} C_2 \xrightarrow{f_3} \dots \xrightarrow{f_n} C_n$$

，其中对象 C_0 以及与之相关的态射 f_0 被“删除”。

Example 3. 给定一个范畴 C ，以及一个范畴 X_n 中单纯对象的 n -单纯形 X_n ，面算子 d_n 作用于 X_n 产生序列

$$d_n X_n = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} C_{n-1}$$

，其中对象 C_n 以及进入它的态射 f_n 被“删除”。

Example 4. 给定一个范畴 C ，以及一个简约对象的 n -单形 X_n ，面算子 $d_i, 0 < i < n$ 应用于 X_n 时得到序列

$$d_i X_n = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots C_{i-1} \xrightarrow{f_{i+1} \circ f_i} C_{i+1} \dots \xrightarrow{f_n} C_n$$

，其中对象 C_i 被“删除”，且态射 f_i 与态射 f_{i+1} 复合。

Example 5. 给定一个范畴 C 和一个定义在该范畴上的单纯对象的 n -单纯形 X_n ，退化算子 $s_i, 0 \leq i \leq n$ 应用到 X_n 得到序列

$$s_i X_n = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots C_i \xrightarrow{1_{C_i}} C_i \xrightarrow{f_{i+1}} C_{i+1} \dots \xrightarrow{f_n} C_n$$

，其中对象 C_i 通过插入其恒等态射 1_{C_i} 来“重复”。

7.2 Kan 复形

为了证明 LLMs 定义了模型范畴，我们使用提升图引入了 Kan 复形。

Definition 23. standard simplex Δ^n 是由所有从 $[m]$ 映射出来的集合定义的单纯集合：

$$([m] \in \Delta) \mapsto \mathbf{Hom}_\Delta([m], [n])$$

根据惯例， $\Delta^{-1} := \emptyset$ 。标准的 0 简单形 Δ^0 将每个 $[n] \in \Delta^{op}$ 映射到单元集合 $\{\bullet\}$ 。

Definition 24. 令 X 表示一个单纯对象，其中 X_n 是其 n^{th} 单形。如果对于每个整数 $n \geq 0$ ，给定一个子集 $Y_n \subseteq X_n$ ，使得应用于 Y_n 的面和退化映射

$$d_i: X_n \rightarrow X_{n-1} \quad s_i: X_n \rightarrow X_{n+1}$$

导致

$$d_i: Y_n \rightarrow Y_{n-1} \quad s_i: Y_n \rightarrow Y_{n+1}$$

，则集合 $\{Y_n\}_{n \geq 0}$ 定义了一个 **simplicial subset** $Y_\bullet \subseteq X_\bullet$ 。

Definition 25. boundary 是一个单纯形集 $(\partial \Delta^n): \Delta^{op} \rightarrow \mathbf{Set}$ ，定义为

$$(\partial \Delta^n)([m]) = \{\alpha \in \mathbf{Hom}_\Delta([m], [n]) : \alpha \text{ is not surjective}\}$$

注意，边界 $\partial \Delta^n$ 是标准 n -单纯形 Δ^n 的一个单纯子集。

Definition 26. Horn $\Lambda_i^n: \Delta^{op} \rightarrow \mathbf{Set}$ 被定义为

$$(\Lambda_i^n)([m]) = \{\alpha \in \mathbf{Hom}_\Delta([m], [n]) : [n] \not\subseteq \alpha([m]) \cup \{i\}\}$$

直观地说，角 Λ_i^n 可以被视为从 n -单纯形 Δ^n 中移除其内部及其 i 顶点相对的结果所形成的单纯子集。

Definition 27. 设 $f: X \rightarrow S$ 为范畴 C 中单纯对象的一个态射。我们称 f 为 **Kan 纤维化**，如果对于每个 $n > 0$ 和每个 $0 \leq i \leq n$ ，所有上提问题，

$$\begin{array}{ccc}
\Lambda_i^n & \xrightarrow{\sigma_0} & X \\
\downarrow \sigma & \nearrow \sigma & \downarrow f \\
\Delta^n & \xrightarrow{\bar{\sigma}} & S
\end{array}$$

都有解。更确切地说，对于单纯集合 $\sigma_0: \Lambda_i^n \rightarrow X$ 的每个映射和每个拓展了 $f \circ \sigma_0$ 的 n -单形，我们可以将 σ_0 扩展为满足 $f \circ \sigma = \bar{\sigma}$ 的 n -单形 $\sigma: \Delta^n \rightarrow X$ 。

Example 6. 给定范畴 C 中的单纯对象 X ，一个作为 *Kan* 纤维化的投影映射 $X \rightarrow \Delta^0$ 被称为 **Kan complex**。

Example 7. 在范畴 C 中，任何单纯对象之间的同构都是 *Kan* 纤维化。

Example 8. 范畴中的 *Kan* 纤维化集合在收缩下是封闭的。

7.3 模型类别

一个优雅的抽象同伦框架由 Quillen [1967] 开发，称为模型类别，它基于将箭头分类为三种类型的态射：一组共纤维化（例如，图 1 左侧的“单射”箭头 i ），一组纤维化（图 1 右侧的满射箭头 p ），以及一组弱等价，它们必须满足以下性质。我们的一个主要结果（见定理 6）是 LLMs 定义了模型范畴。

Definition 28. 一个模型类别 M 是一个具备三类态射的范畴：弱等价性、上合并映射和纤维化。一个既是（余）纤维化也是弱等价的映射被称为非循环（共）纤维化。一个具有这三类态射的模型范畴 M 必须满足以下五个条件：

1. **MC1**: 类别 M 具有所有小极限和余极限。
2. **MC2**: 如果 f 和 g 是 M 中的任意两个映射，并且 $g \circ f$ 是定义好的，并且如果这些映射中的任意两个—— f 、 g 和 $g \circ f$ ——是弱等价，那么第三个也会是。
3. **MC3**: 如果 f 和 g 是 M 中的映射，并且 f 是 g 的收缩，并且 g 是弱等价、纤维化或协纤维化，则 f 也是。
4. **MC4**: 给定一个如下面定义的提升图，如果 f 是协纤维化且 p 是非循环纤维化，或者 i 是非循环协纤维化且 p 是纤维化，则点箭头 h 存在。

$$\begin{array}{ccc}
A & \xrightarrow{\mu} & X \\
\downarrow f & \nearrow h & \downarrow p \\
B & \xrightarrow{v} & Y
\end{array}$$

5. **MC5**: 任何在 M 中的映射 g 可以用两种方式分解： $g = q \circ i$ ，其中 i 是协纤维化且 q 是非循环纤维化，或者 $g = p \circ j$ ，其中 j 是非循环协纤维化且 p 是纤维化。

Definition 29. 设 C 为任意一个包含弱等价子范畴 $\mathcal{W}$ 的范畴。“自由范畴” $F(C, \mathcal{W}^{-1})$ 是通过在 C 中添加箭头 $\mathcal{W}$ 的反向箭头而定义的，它与 C 具有相同的对象，其态射是可合成态射 $(f_1, \dots, f_n)$ 的字符串，其中 f_i 要么是 C 的一个箭头，要么是箭头 w_i 在 $\mathcal{W}$ 中的反向 w_i^{-1} 。同伦范畴 HoC 被定义为“自由”范畴通过对所有对象 x 的关系 $1_x = (1_x)$ 、对所有可合成态射 $f, g \in C$ 的关系 $(f, g) = (g \circ f)$ 、以及对所有 $w \in \mathcal{W}$ 的 $1_{dom_w} = (w^{-1}, w)$ 进行商构造的结果。

模型范畴的详细概述在许多教材中给出 [Hovey, 2020, May and Ponto, 2012, Quillen, 1967]。

7.4 LLM 模型类别

我们论文的一个关键结果是 LLM 马尔可夫范畴定义了模型范畴。在概述证明之前，我们引入两个基本定义。

Definition 30. 设 C 和 C' 是 LLM Markov 范畴 $\mathcal{L}$ 中的一对对象。我们说 C 是 C' 的 **a retract**，若存在映射 $i: C \rightarrow C'$ 和 $r: C' \rightarrow C$ 使得 $r \circ i = id_C$ 。

Definition 31. 设 $\mathcal{L}$ 为一个 LLM Markov 范畴。我们称态射 $f: C \rightarrow D$ 为 **retract of another morphism** $f': C \rightarrow D$ ，如果它是 f' 的一个收缩态，从函子范畴 $\mathbf{Hom}([1], \mathcal{L})$ 的对象来看。回顾一下， $[1] = \{0, 1\}$ 有一个非平凡态射 $0 \rightarrow 1$ 。 $\mathcal{L}$ 的态射集合 T 是 **closed under retracts**，如果对于 $\mathcal{L}$ 的每一对态射，如果 f 是 f' 的收缩态，并且 f' 属于 T ，则 f 也属于 T 。

Theorem 6. LLM 马尔可夫范畴定义模型范畴。

证明大纲：证明是辛普里希亚集合定义模型范畴这一结果的一个特例 [Quillen, 1967, Hovey, 2020, May and Ponto, 2012]：

- *Fibrations are Kan complexes* : Kan 复形是那些满足所有角包含物提升性质的复形。使用神经函子在 LLM 马尔可夫范畴上构造的单纯集合是 Kan 复形，因为每个 n -单纯形是一个 n -长度的序列。通过应用面算子，LLM 的任何 $n-1$ -单纯形 X_{n-1} 都是 LLM 的 n -单纯形 X_n 的一个可缩影（见例 2-4）。
- *Cofibrations are monomorphisms of simplicial sets* : 如果我们在 LLM 单纯集合之间的映射定义为其对应 n -单纯形（这些单纯形是可合成态射的集合）之间的组成映射，那么 LLM 模型范畴中的上纤维是单射映射。
- *Weak equivalences are associated with topological embeddings of simplicial sets* : 两个 LLM 单纯集合间的弱等价是它们拓扑实现之间的相应弱等价（我们在第 ?? 节回顾 LLM 的拓扑嵌入）。

□

8 笛卡尔拓扑空间

在本节中，我们讨论拓扑空间的“良好”类别，显示明显的分类导致了一个对称单体结构，其中张量积 $\otimes$ 被定义为拓扑空间的乘积，但它不是一个笛卡尔范畴。限制在 CW 复形（或紧生成的弱 Hausdorff 空间 [Munkres, 1984]）的子类别中，可以定义一个合适的笛卡尔结构。这一回顾将激发使用从 LLM 类别通过单纯集合构造拓扑结构的动机，这在主文中已进行了讨论。

8.1 拓扑空间中的笛卡尔结构

我们在本文中的目标是构建自然语言句子的等价类的拓扑表示，这些句子在一个大型语言模型中表示。为此，我们需要了解一些关于拓扑空间类别的基本知识 [Borceux, 1994]，以及为何在主要论文中描述的单纯形集构造能够导致一个良好的拓扑结构。

Definition 32. 范畴 **顶部** 定义了所有拓扑空间的范畴，其中对象是拓扑空间，箭头是从一个空间到另一个空间的连续函数。

范畴 **顶部** 具有自然的对称单oidal 结构，具体定义见以下结果（参见 [Borceux, 1994]，第 7 章）。

Theorem 7. 类别 **顶点** 上有一个自然的闭对称单元素结构，其中

- 两个空间 X 和 Y 的张量积 $X \otimes Y$ 的底层集合必须是底层集合的笛卡尔积 $X \times Y$ 。
- 构成内在“同态对象”函数空间 $[X, Y]$ 的集合必然是从 X 到 Y 的连续映射的集合 $\mathbf{Top}(X, Y)$ 。

然而，这种自然的幺半结构不具有正确的拓扑性质，因为以下结果：

Theorem 8. 范畴 **顶部** 不是 *Cartesian* 封闭的。

和以前一样，我们要求笛卡尔结构必须具备构造指数对象的能力。

Definition 33. 只有当积函子 $- \times X : \mathbf{Top} \rightarrow \mathbf{Top}$ 有右伴随时，拓扑空间 X 才是 **可指数化**。

幸运的是，可以构造这样一个拓扑空间的子范畴。

Definition 34. 一个拓扑空间 X 是局部紧致的，当且仅当其中每个点的邻域都包含该点的一个紧邻域。

这种限制为构建可指数化的对象带来了令人满意的结果。

Theorem 9. 局部紧空间是可指数的。对于函子 $- \times X$ 的右伴随由将 Y 映射到连续函数集 $\mathbf{Top}(X, Y)$ 的函子给出，并以紧开拓扑进行拓扑化。

$\mathbf{Top}(X, Y)$ 上的紧开拓扑定义为

$$\langle K, U \rangle = \{f \in \mathbf{Top}(X, Y) | f(K) \subseteq U\}$$

其中 K 是 X 的紧子集， U 是 Y 的开子集。

拓扑空间上的理想笛卡尔结构由紧致豪斯多夫空间和紧致开映射的范畴给出。

Theorem 10. 紧致豪斯多夫空间和紧致连续映射的范畴 **CGHaus** 是笛卡尔闭的。

最后，我们注意到主文中描述的单纯集合构造得到了 CW 复形（紧生成的弱豪斯多夫），其定义如下：

Definition 35. 如果任何紧致 Hausdorff 域 X 的连续映射的像在 X 中是闭的，则空间 X 是 **弱 Hausdorff**。

这些性质在理解 LLM 范畴的分类空间的拓扑结构时会很有用，尤其是它们导向称为 H -空间的结合和乘法空间。

8.2 拓扑嵌入作为末端

我们现在引入一个视角，即单纯对象的拓扑嵌入也可以被解释为一个余端 [Loregian, 2021, MacLane, 1971]（我们在第 3.4 节回顾（余）端演算）：

$$F : \Delta^0 \times \Delta \rightarrow \mathbf{Top}$$

其中 F 作为从 Δ 到 **Sets** 的函子作用，映射 $[n] \mapsto X_n$ ，而协变地作为从 Δ 到拓扑空间范畴 **Top** 的函子映射 $[m] \mapsto \Delta^m$ 。该余端定义了一个单纯对象 X 的拓扑嵌入，其中 X_n 表示具有相关概率 p 的长度为 n 的可组合态射。给定该单纯对象，我们现在可以构建它的一个余端对象作为拓扑实现

$$\int^n (X_n) \cdot \Delta n$$

，其中 $X : \Delta^{op} \rightarrow \mathbf{C}$ 是由从单纯范畴 Δ 到单纯对象范畴的反变函子定义的单纯对象，而 $\Delta : |\Delta| \rightarrow \mathbf{Top}$ 是一个函数，从单纯范畴 Δ 的拓扑 n -单纯形实现到拓扑空间 **Top**。在此，符号 $(X_n) \cdot \Delta n$ 意味着“共幂”，代表拓扑 n -单纯形的 X_n 份不交并集 $\sqcup_{X_n} \Delta$ 。

9 LLM 模糊单纯形对象

现在我们研究在第 ?? 节的 LLM 范畴上定义的“模糊”单纯集合 [McInnes et al., 2018]。单纯对象可以使用最初由 Milnor [1957] 提出的构造嵌入到拓扑空间中。

Definition 36. 一个单纯形对象的 **几何实现** $|X|$ 在一个定义为拓扑空间的 LLM 类别 $\mathcal{L}$ 中的 X ，其中 X_n 是指假设具有离散拓扑的 n -单形（即， X_n 的所有子集都是开放集），并且 Δ^n 表示拓扑 n -单形：²

$$\Delta^n = \{(p_0, \dots, p_n) \in \mathbb{R}^{n+1} \mid 0 \leq p_i \leq 1, \sum_i p_i = 1\}$$

。空间 $\Delta^n, n \geq 0$ 可以被看作是具有以下退化和面映射的余筒形拓扑空间：

$$\delta_i(t_0, \dots, t_n) = (t_0, \dots, t_{i-1}, 0, t_i, \dots, t_n) \text{ for } 0 \leq i \leq n$$

$$\sigma_j(t_0, \dots, t_n) = (t_0, \dots, t_j + t_{j+1}, \dots, t_n) \text{ for } 0 \leq i \leq n$$

。注意 $\delta_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$ ，而 $\sigma_j : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ 。上面定义商空间的等价关系由以下给出：

$$(d_i(x), (t_0, \dots, t_n)) \sim (x, \delta_i(t_0, \dots, t_n))$$

$$(s_j(x), (t_0, \dots, t_n)) \sim (x, \sigma_j(t_0, \dots, t_n))$$

。

10 总结与未来工作

为了从范畴上捕捉语言中充斥着表面上不同但意义相同的陈述这一事实，例如“查尔斯·达尔文写了”和“查尔斯·达尔文是作者”这类语句，我们需要在范畴语言模型中引入抽象同伦理论的思想。我们定义了一个 LLM 马尔可夫范畴，以表示由 LLM 生成的语言中的概率分布，其中语句的概率，例如“查尔斯·达尔文写了”由马尔可夫范畴中的箭头定义。为解决将语言释义所表示的非同构对象视为本质等同这一基本问题，我们使用范畴同伦技术来捕捉 LLM 马尔可夫范畴中的“弱等价”。我们提供了关于将范畴同伦应用于 LLM 的详细概述，从高代数 K 理论到模型范畴，这些建立在过去半个世纪发展起来的强大理论结果之上。

在本文中，我们主要关注同伦，但也可以在模糊单纯形集上构建一个链同调结构，该结构将一个 LLM 映射到一系列阿贝尔群 [?]。在未来的工作中，我们希望将这些理论工具应用于经验性 LLM 研究 [Khandelwal et al., 2020] 中的语言释义尝试，目标是深入理解实际部署的 LLM 系统。

我们的论文提供了对 LLM 同伦结构的抽象理论分析。它并不是为了实际演示如何解决下一个词分布的平滑问题，以反映表面上不同语句之间的语义相似性。这类实证研究已经在文献中有所报道，比如 k -NN LLMs [Khandelwal et al., 2020]。范畴同伦理论的本质是利用升降图解 [Gavrilovich, 2017] 的概念捕捉在语义相似性背后的通用性质。它并不直接解决计算问题，将这类见解与实践更紧密地联系起来是我们的长期目标之一。

²在“模糊”单纯集合中，每个对象 $(X[n], p)$ 都嵌入到一个“扩展”的空间 $\{(p_0, \dots, p_n) \mid \sum_i p_i = -\log p\}$ 中。

References

- Ash Asudeh and Gianluca Giorgolo. Enriched Meanings: Natural Language Semantics with Category Theory . Oxford University Press, 09 2020. ISBN 9780198847854. doi: 10.1093/oso/9780198847854.001.0001. URL <https://doi.org/10.1093/oso/9780198847854.001.0001>.
- Tom Avery. Codensity and the giry monad. *Journal of Pure and Applied Algebra* , 220(3):1229–1251, March 2016. ISSN 0022-4049. doi: 10.1016/j.jpaa.2015.08.017. URL <http://dx.doi.org/10.1016/j.jpaa.2015.08.017>.
- M. Boardman and Rainer Vogt. Homotopy invariant algebraic structures on topological spaces . Springer, Berlin, 1973.
- Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri Chatterji, Annie Chen, Kathleen Creel, Jared Quincy Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshthe Khani, Omar Khattab, Pang Wei Koh, Mark Krass, Ranjay Krishna, Rohith Kudipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Ben Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, Julian Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Rob Reich, Hongyu Ren, Frieda Rong, Yusuf Roohani, Camilo Ruiz, Jack Ryan, Christopher Ré, Dorsa Sadigh, Shiori Sagawa, Keshav Santhanam, Andy Shih, Krishnan Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. On the opportunities and risks of foundation models, 2022.
- Francis Borceux. Handbook of Categorical Algebra , volume 2 of Encyclopedia of Mathematics and its Applications . Cambridge University Press, 1994. doi: 10.1017/CBO9780511525858.
- TD. Bradley, J. Terilla, and Y. Vlassopoulos. An enriched category theory of language: From syntax to semantics. *La Matematica* , 1:551–580, 2022.
- Sneha Chaudhari, Varun Mithal, Gungor Polatkan, and Rohan Ramanath. An attentive survey of attention models, 2021. URL <https://arxiv.org/abs/1904.02874>.
- Kenta Cho and Bart Jacobs. Disintegration and bayesian inversion via string diagrams. *Mathematical Structures in Computer Science* , 29(7):938–971, March 2019. ISSN 1469-8072. doi: 10.1017/s0960129518000488. URL <http://dx.doi.org/10.1017/S0960129518000488>.
- Bob Coecke. The mathematics of text structure. *ArXiv* , abs/1904.03478, 2019. URL <https://api.semanticscholar.org/CorpusID:102351346>.
- Bob Coecke. The mathematics of text structure, 2020.
- Bob Coecke and Aleks Kissinger. Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning . Cambridge University Press, 2017.
- Bob Coecke, Mehrnoosh Sadrzadeh, and Stephen Clark. Mathematical foundations for a compositional distributional model of meaning, 2010. URL <https://arxiv.org/abs/1003.4394>.
- DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, Bing Xue, Bingxuan Wang, Bochao Wu, Bei Feng, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, Damai Dai, Deli Chen, Dongjie Ji, Erhang Li, Fangyun Lin, Fucong Dai, Fuli Luo, Guangbo Hao, Guanting Chen, Guowei Li, H. Zhang, Han Bao, Hanwei Xu, Haocheng Wang, Honghui Ding, Huajian Xin, Huazuo Gao, Hui Qu, Hui Li, Jianzhong Guo, Jia Shi Li, Jiawei Wang, Jingchang Chen, Jingyang Yuan, Junjie Qiu, Junlong Li, J. L. Cai, Jiaqi Ni, Jian Liang, Jin Chen, Kai Dong, Kai Hu, Kaige Gao, Kang Guan, Kexin Huang, Kuai Yu, Lean Wang, Lecong Zhang, Liang Zhao, Litong Wang, Liyue Zhang, Lei Xu, Leyi Xia, Mingchuan Zhang, Minghua Zhang, Minghui Tang, Meng Li, Miaojun Wang, Mingming Li, Ning Tian, Panpan Huang, Peng Zhang, Qiancheng Wang, Qinyu Chen, Qiusi Du, Ruiqi Ge, Ruisong Zhang, Ruizhe Pan, Runji Wang, R. J.

- Chen, R. L. Jin, Ruyi Chen, Shanghao Lu, Shangyan Zhou, Shanhuang Chen, Shengfeng Ye, Shiyu Wang, Shuiping Yu, Shunfeng Zhou, Shuting Pan, S. S. Li, Shuang Zhou, Shaoqing Wu, Shengfeng Ye, Tao Yun, Tian Pei, Tianyu Sun, T. Wang, Wangding Zeng, Wanxia Zhao, Wen Liu, Wenfeng Liang, Wenjun Gao, Wenqin Yu, Wentao Zhang, W. L. Xiao, Wei An, Xiaodong Liu, Xiaohan Wang, Xiaokang Chen, Xiaotao Nie, Xin Cheng, Xin Liu, Xin Xie, Xingchao Liu, Xinyu Yang, Xinyuan Li, Xuecheng Su, Xuheng Lin, X. Q. Li, Xiangyue Jin, Xiaojin Shen, Xiaosha Chen, Xiaowen Sun, Xiaoxiang Wang, Xinnan Song, Xinyi Zhou, Xianzu Wang, Xinxia Shan, Y. K. Li, Y. Q. Wang, Y. X. Wei, Yang Zhang, Yanhong Xu, Yao Li, Yao Zhao, Yaofeng Sun, Yaohui Wang, Yi Yu, Yichao Zhang, Yifan Shi, Yiliang Xiong, Ying He, Yishi Piao, Yisong Wang, Yixuan Tan, Yiyang Ma, Yiyuan Liu, Yongqiang Guo, Yuan Ou, Yuduan Wang, Yue Gong, Yuheng Zou, Yujia He, Yunfan Xiong, Yuxiang Luo, Yuxiang You, Yuxuan Liu, Yuyang Zhou, Y. X. Zhu, Yanhong Xu, Yanping Huang, Yaohui Li, Yi Zheng, Yuchen Zhu, Yunxian Ma, Ying Tang, Yukun Zha, Yuting Yan, Z. Z. Ren, Zehui Ren, Zhangli Sha, Zhe Fu, Zhean Xu, Zhenda Xie, Zhengyan Zhang, Zhewen Hao, Zhicheng Ma, Zhigang Yan, Zhiyu Wu, Zihui Gu, Zijia Zhu, Zijun Liu, Zilin Li, Ziwei Xie, Ziyang Song, Zizheng Pan, Zhen Huang, Zhipeng Xu, Zhongyu Zhang, and Zhen Zhang. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning, 2025. URL <https://arxiv.org/abs/2501.12948>.
- Matthew Finlayson, Xiang Ren, and Swabha Swayamdipta. Logits of API-Protected LLMs leak proprietary information, 2024. URL <https://arxiv.org/abs/2403.09539>.
- Brendan Fong. Causal theories: A categorical perspective on bayesian networks. Master’s thesis, Oxford University, 2012.
- Brendan Fong and David I Spivak. Seven Sketches in Compositionality: An Invitation to Applied Category Theory . Cambridge University Press, 2018.
- Brendan Fong, David I. Spivak, and Rémy Tuyéras. Backprop as functor: A compositional perspective on supervised learning. In 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019 , pages 1–13. IEEE, 2019. doi: 10.1109/LICS.2019.8785665. URL <https://doi.org/10.1109/LICS.2019.8785665>.
- James Fox. Universal Coalgebras . Dissertation, McGill University, 1976.
- Tobias Fritz. A synthetic approach to markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics* , 370:107239, August 2020. ISSN 0001-8708. doi: 10.1016/j.aim.2020.107239. URL <http://dx.doi.org/10.1016/j.aim.2020.107239>.
- Tobias Fritz and Andreas Klingler. The d-separation criterion in categorical probability. *Journal of Machine Learning Research* , 24(46):1–49, 2023. URL <http://jmlr.org/papers/v24/22-0916.html>.
- P. Gabriel, P. Gabriel, and M. Zisman. Calculus of Fractions and Homotopy Theory . Calculus of Fractions and Homotopy Theory. Springer-Verlag, 1967. ISBN 9780387037776. URL <https://books.google.com/books?id=UEQZAQAIAAJ>.
- Misha Gavrillovich. The unreasonable power of the lifting property in elementary mathematics, 2017. URL <https://arxiv.org/abs/1707.06615>.
- Daniel Grayson. Higher algebraic k-theory: Ii. In Michael R. Stein, editor, *Algebraic K-Theory* , pages 217–240, Berlin, Heidelberg, 1976. Springer Berlin Heidelberg. ISBN 978-3-540-37964-5.
- Albert Gu, Karan Goel, and Christopher Ré. Efficiently modeling long sequences with structured state spaces. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022* . OpenReview.net, 2022. URL <https://openreview.net/forum?id=uYLFoz1v1AC>.
- Michael Hahn. Theoretical limitations of self-attention in neural sequence models. *Trans. Assoc. Comput. Linguistics* , 8:156–171, 2020. doi: 10.1162/TACL_A_00306. URL https://doi.org/10.1162/tacl_a_00306.
- Paul Halmos. *Measure Theory* . Springer, 1974.
- Chris Heunen and Jamie Vicary. *Categories for Quantum Theory: An Introduction* . Oxford University Press, November 2019. doi: 10.1093/oso/9780198739623.001.0001.
- Mark Hovey. Model categories. *Foundations of Stable Homotopy Theory* , 2020. URL <https://api.semanticscholar.org/CorpusID:241323754>.
- Cheng-Yu Hsieh, Chun-Liang Li, Chih-Kuan Yeh, Hootan Nakhost, Yasuhisa Fujii, Alexander Ratner, Ranjay Krishna, Chen-Yu Lee, and Tomas Pfister. Distilling step-by-step! outperforming larger language models with less training data and smaller model sizes, 2023. URL <https://arxiv.org/abs/2305.02301>.
- Bart Jacobs, Aleks Kissinger, and Fabio Zanasi. Causal inference by string diagram surgery, 2018. URL <https://arxiv.org/abs/1811.08338>.

- A. Joyal. Quasi-categories and kan complexes. *Journal of Pure and Applied Algebra* , 175(1):207–222, 2002. ISSN 0022-4049. doi: [https://doi.org/10.1016/S0022-4049\(02\)00135-4](https://doi.org/10.1016/S0022-4049(02)00135-4). URL <https://www.sciencedirect.com/science/article/pii/S0022404902001354>. Special Volume celebrating the 70th birthday of Professor Max Kelly.
- Dimitri Kartsaklis, Ian Fan, Richie Yeung, Anna Pearson, Robin Lorenz, Alexis Toumi, Giovanni de Felice, Konstantinos Meichanetzidis, Stephen Clark, and Bob Coecke. lambeq: An efficient high-level python library for quantum nlp, 2021. URL <https://arxiv.org/abs/2110.04236>.
- Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. Generalization through memorization: Nearest neighbor language models, 2020. URL <https://arxiv.org/abs/1911.00172>.
- Dan Klein and Christopher D. Manning. A* parsing: Fast exact Viterbi parse selection. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics* , pages 119–126, 2003. URL <https://aclanthology.org/N03-1016/>.
- J. Lambek. Pregroups and natural language processing. *The Mathematical Intelligencer* , 28:41–48, 2006.
- Mike Lewis and Mark Steedman. A* CCG parsing with a supertag-factored model. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* , pages 990–1000, Doha, Qatar, October 2014. Association for Computational Linguistics. doi: 10.3115/v1/D14-1107. URL <https://aclanthology.org/D14-1107/>.
- Fosco Loregian. (Co)end Calculus . London Mathematical Society Lecture Note Series. Cambridge University Press, 2021. doi: 10.1017/9781108778657.
- Jacob Lurie. Higher Topos Theory . *Annals of mathematics studies*. Princeton University Press, Princeton, NJ, 2009. URL <https://cds.cern.ch/record/1315170>.
- Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic a First Introduction to Topos Theory* . Springer New York, New York, NY, 1992. ISBN 9781461209270 1461209277. URL <http://link.springer.com/book/10.1007/978-1-4612-0927-0>.
- Saunders MacLane. *Categories for the Working Mathematician* . Springer-Verlag, New York, 1971. Graduate Texts in Mathematics, Vol. 5.
- Sridhar Mahadevan. Universal causality. *Entropy* , 25(4):574, 2023. doi: 10.3390/E25040574. URL <https://doi.org/10.3390/e25040574>.
- J.P. May. *Simplicial Objects in Algebraic Topology* . University of Chicago Press, 1992.
- J.P. May and K. Ponto. *More Concise Algebraic Topology: Localization, Completion, and Model Categories* . Chicago Lectures in Mathematics. University of Chicago Press, 2012. ISBN 9780226511788. URL <https://books.google.com/books?id=SHhmXUPskFwC>.
- Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2018. URL <https://arxiv.org/abs/1802.03426>.
- William Merrill, Ashish Sabharwal, and Noah A. Smith. Saturated transformers are constant-depth threshold circuits, 2022. URL <https://arxiv.org/abs/2106.16213>.
- William Merrill, Zhaofeng Wu, Norihito Naka, Yoon Kim, and Tal Linzen. Can you learn semantics through next-word prediction? the case of entailment, 2024. URL <https://arxiv.org/abs/2402.13956>.
- John Milnor. The geometric realization of a semi-simplicial complex. *The Annals of Mathematics* , 65(2): 357–362, 1957.
- Meredith Ringel Morris, Jascha Sohl-Dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet, and Shane Legg, editors. Levels of AGI for Operationalizing Progress on the Path to AGI , 2023. Original arXiv title in November 2023 was "Levels of AGI": Operationalizing Progress on the Path to AGI. Final title for publication as a position paper at ICML 2024 is: Levels of AGI for Operationalizing Progress on the Path to AGI.
- James R. Munkres. *Elements of algebraic topology* . Addison-Wesley, 1984. ISBN 978-0-201-04586-4.
- Daniel Quillen. Higher algebraic k-theory: I. In H. Bass, editor, *Higher K-Theories* , pages 85–147, Berlin, Heidelberg, 1973. Springer Berlin Heidelberg. ISBN 978-3-540-37767-2.
- Daniel G. Quillen. *Homotopical algebra* . Springer, 1967.
- B. Richter. *From Categories to Homotopy Theory* . Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2020. ISBN 9781108479622. URL <https://books.google.com/books?id=pnzUDwAAQBAJ>.