arXiv:2301.13800v1 [math.LO] 31 Jan 2023

# 模型类大小与描述长度之间的单调连接

ReijoJaakkola Antti Kuusisto Miikka Vilander坦佩雷大学 坦佩雷大学 坦佩雷大学 芬兰 赫尔辛基大学 芬兰 芬兰

摘要

本文将模型类的大小与定义它们的公式的最小长度联系起来，也就是与它们的描述复杂度联系起来。我们将研究范围限定为定义域大小为 n 的固定论域模型，并考察相对于能够计数赋值的命题逻辑扩展而言的描述复杂度。这种逻辑称为 GMLU，也可以被视为全可及关系上的 Kripke 模型所对应的分级模态逻辑。GMLU 在表达能力上足以定义赋值多重集；我们还研究其片段 GMLU(d)，它只能计数至整数阈值 d。我们尤其关注 GMLU(d) 等价类的描述复杂度。我们证明，在类型实现构成的偏序集上，等价类按大小排列的序，与按描述复杂度排列的序相同。这也表明，模型类的玻尔兹曼熵与描述复杂度之间存在单调连接。此外，我们刻画了定义域大小 n 与计数阈值 d 之间的关系如何决定是否存在支配类，而支配类实质上是极限概率为一的模型类。为获得这些结果，我们证明了关于 r-关联斯特林数的新估计。作为另一项关键工具，我们还表明，模型类根据其描述复杂度分为两种截然不同的情况。

## 1 引言

本文研究模型类的大小与定义这些类所需公式的最小长度之间的联系。在我们考察的情形中，首先固定一类 $\mathcal{M}$ 共享相同有限大小 $n$论域的模型。随后，我们研究模型类 $M \subseteq \mathcal{M}$ 相对于命题逻辑的一种扩展；该扩展能够对命题赋值进行计数。我们将这种逻辑称为 GMLU，因为它也可以定义为基于具有全关系的 Kripke 模型的分级模态

逻辑。为了获得更精细的结果，我们对 GMLU 加入参数化，并研究其片段 GMLU$_d$，这些片段只能计数至阈值 $d \in \mathbb{Z}_+$。这也使我们能够展示，随着所用逻辑的表达能力逐步增强，公式最小长度与模型类大小之间的关系如何变化。对于模型类 $M \subseteq \mathcal{M}$，相对于 GMLU$_d$ 而言，$M$ 的描述复杂度就是定义 $M$所需的 GMLU$_d$ 公式的最小长度（如果存在这样的公式）。

本文特别关注由 GMLU$_d$ 在$\mathcal{M}$上确定的逻辑等价类 $M$ 的描述复杂度 $C_d$(*M*)。若模型 $\mathfrak{M}$和 $\mathfrak{M}' \in \mathcal{M}$ 满足 GMLU$_d$中的同一组公式，则记作 $\mathfrak{M} \equiv_d \mathfrak{M}'$。注意，$\equiv_d$ 中等价类 $M$ 的 $C_d$(*M*)也可以视为其中每个模型 $\mathfrak{M} \in M$的描述复杂度，因为 GMLU$_d$ 的表达能力恰好足以刻画 $\mathfrak{M}$，使其达到等价关系 $\equiv_d$。从这一角度看，描述复杂度类似于柯尔莫哥洛夫复杂度。柯尔莫哥洛夫复杂度与香农熵之间存在广为人知的联系，例如参见 [11]。[8],[7]中的近期研究展示了一种方法，使我们能够在通过逻辑对关系结构进行分类的情形下，也构想相关结果。具体而言，研究表明，GMLU 的等价类的期望玻尔兹曼熵渐近等于期望描述复杂度（相对于 GMLU）乘以所考虑词汇表的大小。研究还表明，对于 $d = 1$，GMLU$_d$ 中最大的等价类在各个类中具有最大的描述复杂度。本文以这些结果为基础。

首先，作为证明中的一项关键工具，我们将描述复杂度划分为两个不同的类。此划分依据的是：在所描述模型类的模型中，实现不同命题类型的元素数量 $n_i$ $i \in I$；这里，$I$ 只是这些类型的指标集。随后，根据是否存在至少两种不同的类型满足 $n_i = d$ 来确定这一划分。利用这一划分，我们建立了模型类大小与描述复杂度之间的紧密联系。对于每个模型 $\mathfrak{M} \in \mathcal{M}$，令$\overline{n}_{\mathfrak{M}}$ 表示元组 $(n_i)_{i \in I}$，它给出了在 $\mathfrak{M}$中实现命题类型的点的数量 $n_i$。此外，对于大于计数阈值 $d$的数量 $n_i$，我们不再记录，而是在$\overline{n}_{\mathfrak{M}}$中直接填入 $d$。我们在模型上定义偏序集 $(\mathcal{M}, \preceq_\tau)$，其中 $\tau$是词汇表，序$\preceq_\tau$ 基于对元组 $\overline{n}_{\mathfrak{M}}$ 逐坐标比较。序 $\preceq_\tau$ 也直接继承到 $\equiv_d$ 的类上，使得当且仅当对于相应类中的某些（或等价地，所有）模型 $\mathfrak{M}$ 和 $\mathfrak{M}'$，都有 $\mathfrak{M} \preceq_\tau \mathfrak{M}'$ 时，$M \preceq_\tau M'$。我们将证明，对于所有满足以下条件的 $\equiv_d$ 的类 $M$ 和 $M'$：$M$且 $M'$在 $\preceq_\tau$意义下可比较，有

$$|M| < |M'| \;\Leftrightarrow\; C_d(M) < C_d(M').$$

换言之，在 $\preceq_\tau$上，按照大小对模型类进行排序，与按照描述复杂度进行排序完全一致。这是句法与语义之间的紧密联系。作为推论，我们得到模型类的玻尔兹曼熵与描述复杂度之间的对应关系。

接着，我们考察改变定义域大小 $n$ 和计数阈值 $d$ 时， $\equiv_d$ 的类会如何变化。请注意，增大 $d$ 相当于转向表达力越来越强的逻辑。首先我们观察到，对于阈值 $d$ 和 $d' > d$ ，以及由 $\equiv_d$ 和 $\equiv_{d'}$ 给出的模型类分布所对应的香农熵 $H_S(\equiv_d)$ 和 $H_S(\equiv_{d'})$，有

$$H_S(\equiv_d) < H_S(\equiv_{d'}) \text{ when } d' \text{ is at most } n/2, \text{ and}$$
$$H_S(\equiv_d) = H(\equiv_{d'}) \text{ when } d \text{ is at least } n/2.$$

对于期望玻尔兹曼熵 $H_B(\equiv_d)$ 和 $H_B(\equiv_{d'})$，也可得出类似结果，但序相反，即当 $d'$ 至多为 $n/2$ 时， $H_B(\equiv_d) > H_B(\equiv_{d'})$。

为了更好地了解改变 $n$ 和 $d$ 时各类的相对大小，我们证明了类分布的渐近刻画，其中 $n \to \infty$ 和 $d$ 是 $n$ 的函数。我们称 $\equiv_{d(n)}$ 具有支配类，若以趋于 1 的极限概率，大小为 $n$ 的随机模型属于 $\equiv_{d(n)}$ 中大小最大的类。类似地，若以趋于 0 的极限概率，大小为 $n$ 的随机模型属于最大类，则称 $\equiv_{d(n)}$ 中的所有类都消失。于是，当 $n \to \infty$ 时，以下结果成立。

- 若 $d(n) \leq n/2^{|\tau|} - f(n)$，其中 $f(n) = \omega(\sqrt{n})$，则 $\equiv_{d(n)}$ 有一个支配类。
- 若 $d(n) \geq n/2^{|\tau|} - f(n)$，其中 $f(n) = o(\sqrt{n})$，则 $\equiv_{d(n)}$ 没有支配类。
- 若 $d(n) \geq n/2^{|\tau|} + f(n)$，其中 $f(n) = \omega(\sqrt{n})$，则 $\equiv_{d(n)}$中的每个类都是消失的。

这些结果的一个推论是，对于 $d(n) \leq n/t - f(n)$，如果 $f(n) = \omega(\sqrt{n})$，那么以极限概率 1，大小为 $n$ 的两个随机模型无法在 $\mathsf{GMLU}^{d(n)}$ 中区分。最后，我们还给出了 $\equiv_d$ 类分布刻画的一个非渐近变体，其中包括显式界 $d$ ，用于区分 $\equiv_d$ 是否具有多数类的情形。多数类是指在 $\mathcal{M}$ 中包含超过一半模型的类。

关于相关研究，正如前面提到的，熵与柯尔莫哥洛夫复杂度之间存在关联，这一点众所周知。事实上，对于可计算分布，香农熵与柯尔莫哥洛夫复杂度相差至多一个常数。例如， [11], [4], [10]讨论了这一结果。然而，[15] 证明，对于 Rényi 熵和 Tsallis熵，这种一般性关联并不成立。参见

例如， [4], [10], [15] 讨论了 Rényi 熵和 Tsallis熵。就我们所知，逻辑公式长度与熵之间的第一个联系见于 [8], [7], ，其中证明了期望玻尔兹曼熵与描述复杂度渐近等价。

关于其他相关工作，接下来我们将讨论本文采用的证明技巧。为了证明公式大小的界，我们针对 GMLU$_d$逻辑使用公式大小博弈。事实上，标准 Ehrenfeucht-Fra¨ıss´e 博弈和（分级）双模拟博弈的变体都不足以满足需求，因为我们需要处理公式长度，因而需要考虑所有逻辑算子，包括联结词。下文将基于 [8], [7] 中使用的一种类似博弈，构建 GMLU$_d$ 逻辑的公式大小博弈。该博弈以 [5] 中用于标准模态逻辑 ML 的博弈为基础；该博弈被提出并加以发展，用来证明一阶逻辑与 ML 之间存在非初等简洁性差距。Razborov 在 [13], 中提出的第一个公式大小博弈针对命题逻辑。后来，Adler 和 Immerman 为 CTL 定义了该博弈的一个变体 [1]。设计 G MLU$_d$ 的博弈相对直接，且直接基于类似的早期体系，但要运用这些博弈则需要一些非平凡的组合论论证。

除了博弈之外，我们还采用了一系列技术来估计模型类大小和描述复杂度，包括斯特林近似和切尔诺夫界。特别是，为了得到我们的结果，我们证明了关于 $r$-关联斯特林数的新估计，这些估计本身也可能具有独立的研究价值。

简要概括本文，主要目标是阐明描述长度与模型类大小之间关系的整体图景。这也建立了逻辑与熵概念之间的联系。逻辑 GMLU 适用于当前的研究，甚至可以轻松地通过逐步增加 $d$，接触一系列表达力不断增强的逻辑 GMLU$_d$ 。结论部分讨论了将研究推广到更多逻辑的可能性。尽管本文侧重于理论，描述复杂度这一概念也与一系列应用有关。例如，在人工智能可解释性的一些当前活跃研究中，最短长度的规约可以用作较长公式的解释。关于这一主题的研究，参见例如 [2], [6]。

本文的结构安排如下。第 2 节介绍预备知识；第 3 节利用博弈证明描述复杂度的重要下界；第 4 节证明模型类大小与描述复杂度之间的单调连接；第 5 节通过改变 $n$ 和 $d$，探讨类大小分布的相变。第 6 节总结全文。

## 2 预备知识

我们首先定义本文研究的逻辑。设 $\tau$ 是一个有限的命题符号集合。本文始终将 $\tau$ 视为固定不变。分级通用模态逻辑 GMLU$[\tau]$ 的句法定义如下（稍后将解释这些句法选择）：

$$\begin{aligned}\varphi :=& \blacklozenge^{\geq k}\psi \mid \blacksquare^{<k}\psi \mid \blacklozenge^{=k}\psi \mid \blacksquare^{\neq k}\psi \mid \\ & \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \blacklozenge^{\geq k}\varphi \mid \blacksquare^{<k}\varphi \mid \blacklozenge^{=k}\varphi \mid \blacksquare^{\neq k}\varphi \\ \psi :=& p \mid \neg p \mid \psi \wedge \psi \mid \psi \vee \psi\end{aligned}$$

这里有 $p \in \tau$ 和 $k \in \mathbb{N}$。注意，GMLU$[\tau]$ 的公式中，命题符号仅出现在模态算子的作用域内。此外，所有公式均采用否定范式。在本文中，$\neg\varphi$始终表示这样一个公式：其中的 $\neg$ 已被一直下推到文字层面。

设 $\mathfrak{M}$ 为一个论域为 $W$的 Kripke 模型。在本文中，模态逻辑始终具有一元词汇表，因此 Kripke 模型不与二元可及关系相关联。全局分级模态的语义定义如下：$(\mathfrak{M}, w) \vDash \blacklozenge^{\geq k}\varphi \Leftrightarrow$ 存在至少 $d$ 个元素 $v \in W$ 使得 $(\mathfrak{M}, v) \vDash \varphi$，并且 $(\mathfrak{M}, w) \vDash \blacklozenge^{=k}\varphi \Leftrightarrow$ 存在恰好 $d$ 个元素 $v \in W$ 使得 $(\mathfrak{M}, v) \vDash \varphi$。此外，$(\mathfrak{M}, w) \vDash \blacksquare^{<k}\varphi \Leftrightarrow (\mathfrak{M}, w) \vDash \neg\blacklozenge^{\geq k}\neg\varphi$，并且 $(\mathfrak{M}, w) \vDash \blacksquare^{\neq k}\varphi \Leftrightarrow (\mathfrak{M}, w) \vDash \neg\blacklozenge^{=k}\neg\varphi$。布尔联结词 $\neg$、$\wedge$、$\vee$ 的语义按通常方式定义。注意，$\blacklozenge^{\geq k}$ 和 $\blacksquare^{<k}$ 以及 $\blacklozenge^{=k}$ 和 $\blacksquare^{\neq k}$ 互为对偶算子。因此，GMLU 的模态是菱形算子 $\blacklozenge^{\geq k}$、$\blacklozenge^{=k}$ 及其对偶算子$\blacksquare^{<k}$、$\blacksquare^{\neq k}$。直观地说，$\blacksquare^{<k}$（分别地，$\blacksquare^{\neq k}$）表示所有点都满足$\varphi$，但有数量为 $m < k$（分别为 $m \neq k$）的例外。

设 $\mathfrak{M}$ 是基于 $\tau$ 和 $\varphi \in$ GMLU$[\tau]$ 的一个 Kripke 模型。无指点真值关系定义如下：$\mathfrak{M} \vDash \varphi$ 当且仅当 $\mathfrak{M}$，$w \vDash \varphi$ 对所有 $w \in W$ 均成立。由于所有命题符号都出现在全局模态的辖域内，$\mathfrak{M} \vDash \varphi$ 当且仅当存在某个 $w \in W$ 使得 $\mathfrak{M}$，$w \vDash \varphi$。显然，GMLU 公式的真值不依赖于求值点 $w$。这种独立性正是 GMLU 句法如此定义的原因：它保证命题符号处于模态算子的辖域内。对于带指点模型的集合 $M$，我们记 $M \vDash \varphi \Leftrightarrow (\mathfrak{M}, w) \vDash \varphi$，对每个 $(\mathfrak{M}, w) \in M$ 均如此。

基于 $\tau$ 的命题类型 $\pi$ 是一个极大相容的文字集（即命题符号和取反后的命题符号）。因此，$\pi$ 对每个符号 $p \in \tau$ 恰好包含 $p$、$\neg p$ 中的一个。下文将命题类型简称为类型。基于 $\tau$ 的类型数量记为 $t = 2^{|\tau|}$。

公式 $\varphi \in$ GMLU$[\tau]$ 的计数深度记为 depth($\varphi$)，定义如下：

- 深度($\alpha$) $= 0$ 对于任意文字 $\alpha$,

- 深度($\varphi \wedge \psi$) = 深度($\varphi \vee \psi$)=
  最大值(深度($\varphi$), 深度($\psi$)),
- 深度($\blacklozenge^{\geq k}\varphi$) = 深度($\blacksquare^{<k}\varphi$) $= k$,
- 深度($\blacklozenge^{=k}\varphi$) = 深度($\blacksquare^{\neq k}\varphi$) $= k+1$。

我们用 GMLU$_d[\tau]$ 表示 GMLU$[\tau]$ 的计数深度 $d$片段，其中公式的计数深度限制为至多 $d$。本文的结果针对逻辑 GMLU$_d[\tau]$ 表述。注意，深度($\blacklozenge^{=k}\varphi$) $= k+1$，而深度($\blacklozenge^{\geq k}\varphi$) $= k$。我们提供一些直观解释来说明这一选择。首先，$\blacklozenge^{=d-1}\varphi \equiv \blacklozenge^{\geq d-1}\varphi \wedge \neg\blacklozenge^{\geq d}\varphi$，因此我们可以看到，当公式的计数深度限制为 $d$ 时，允许使用的"精确计数"菱形算子 $\blacklozenge^{=k}$ 总是具有 $k \leq d-1$，因而相较于具有 $k \leq d$ 的"阈值计数"菱形算子 $\blacklozenge^{\geq k}$，它们并不增加表达能力。此外，$k+1$ 也对应于在一元一阶逻辑中表达对 $k$ 个元素进行"精确计数"所需的量词数量。

公式 $\varphi \in$ GMLU$[\tau]$ 的大小，记作 size($\varphi$)，定义如下:

- $\mathrm{size}(\alpha) = 1$ for any literal $\alpha$,
- $\mathrm{size}(\varphi \wedge \psi) = \mathrm{size}(\varphi \vee \psi) = \mathrm{size}(\varphi) + \mathrm{size}(\psi) + 1$,
- $\mathrm{size}(\blacklozenge^{\geq k}\varphi) = \mathrm{size}(\blacksquare^{<k}\varphi) = \mathrm{size}(\varphi) + k$,
- $\mathrm{size}(\blacklozenge^{=k}\varphi) = \mathrm{size}(\blacksquare^{\neq k}\varphi) = \mathrm{size}(\phi) + k + 1$.

请注意，所有文字的大小都相同。这是因为我们希望在公式大小方面，将负面信息（即被否定的信息）与正面信息（即未被否定的信息）视为相同。这也解释了我们为何将 GMLU 定义为使公式处于否定范式。

令 $\mathcal{M}$ 为论域固定为 $W = \{1, \ldots, n\}$ 的所有 $\tau$-模型的集合。当 $\mathcal{M}$ 可从上下文中明确得知时，若公式 $\varphi \in$ GMLU$_d$ 对每个 $\mathfrak{M} \in \mathcal{M}$ 都满足 $\mathfrak{M} \vDash \varphi$ 当且仅当 $\mathfrak{M} \in M$，则称该公式定义集合 $M \subseteq \mathcal{M}$。此时，集合 $M$ 称为 GMLU$_d$- 可定义的。GMLU$_d$- 可定义集合 $M$ 的 GMLU $C(M)$ 描述复杂度，是定义 $M$ 的 GMLU $\varphi \in$ GMLU$_d$ 公式的最小大小。

如果模型 $\mathfrak{M}$ 和 $\mathfrak{N}$ 满足完全相同的 GMLU$_d$-公式，我们可以记作 $\mathfrak{M} \equiv_d \mathfrak{N}$。关系 $\equiv_d$ 显然是一个等价关系，并定义了一个自然的相关划分。作为描述复杂度的一个例子，考虑单字母表 $\tau = \{p\}$ 情况下的 GMLU$_1[\tau]$。其中 $p$ 在每个点都为真的模型，在由 $\equiv_1$ 定义的模型划分中构成一个单元素类。该类的描述复杂度为 2，因为定义该类的最小规模公式是 $\blacksquare^{<1}p$。

设 $I := \{1, \ldots, t\}$，并固定一个枚举 $(\pi_i)_{i \in I}$，其中包含命题词汇表 $\tau$ 上的所有类型。现在令 $M$ 为 $\equiv_d$ 的一个等价类

，其中 $\mathcal{M}$ 是论域为 $W = \{1, \ldots, n\}$、大小为 $n$ 的模型集合。对于类型$\pi_i$，该类中的所有模型要么恰好有 $n_i$ 个类型为 $\pi_i$ 的点（其中$n_i < d$），要么 $M$ 中的所有模型都有至少 $d$ 个类型为 $\pi_i$的点。在后一种情况下，我们定义 $n_i := d$。由此，我们可以用 $t$-元组 $\overline{n}$来刻画 $\equiv_d$的各个类。若对于每个 $i \in I$， $t$-元组 $\overline{n} = (n_i)_{i \in I}$ 都满足$n_i \leq d$ ，且要么至少存在一个 $i \in I$使得 $n_i = d$ ，要么 $\sum_{i \in I} n_i = n$，则称其为 $(n, d)$-可容许的。注意，$(n, d)$-可容许元组 $\overline{n}$ 与 $\equiv_d$ 的等价类之间存在一一对应关系。因此，我们可以用 $M_{\overline{n}}$ 表示与 $\overline{n}$对应的类。

给定一个 $(n, d)$-可容许元组 $\overline{n}$ ，以及一个类型 $\pi_i$ ，它在每个模型 $\mathfrak{M} \in M_{\overline{n}}$ 中恰有相同数量 $k$ 的实现点，我们将这个数量 $k$ 记为 $|\pi_i|_{\overline{n}}$。（注意， $k$ 可以大于 $d$。）当上下文明确时，我们通常会省略 $|\pi_i|_{\overline{n}}$ 下标中的元组 $\overline{n}$ 。

沿用 [7], 的定义，我们将类 $M$ 的玻尔兹曼熵定义为 $H_B(M) := \log(|M|)$。正如 [7], 中所述，这一术语源自统计力学；在那里，玻尔兹曼熵通过与某个宏观态相对应的微观态数量来度量该宏观态的随机性。其思想是，较大的宏观态“随机性更高”（或“特异性更低”），因为随机选择命中它的可能性更大。在统计力学中，玻尔兹曼熵的公式为 $k_B \ln \Omega$，其中 $\Omega$ 是微观态的数量，而 $k_B$ 是玻尔兹曼常数。在我们的定义中，我们使用二进制对数（而不使用 $k_B$）。作为一种总体直觉，可以自然地将公式 $\varphi$ （或其定义的类）与宏观态联系起来，而 $\varphi$ 的模型则是相应的微观态。

现在考虑在 $\equiv_d$ 的等价类上定义的如下自然概率分布： $p_{\equiv_d}(M) = |M|/|\mathcal{M}|$对每个类 $M$。我们仍用符号 $\equiv_d$ 表示这个分布（稍微滥用了记号）。我们将分布 $\equiv_d$ 的玻尔兹曼熵定义为 $H_B$ 在分布 $\equiv_d$ 下的期望值，并记作 $H_B(\equiv_d)$。换言之，我们将 $H_B(\equiv_d)$ 定义为 $\sum_{M \in \mathcal{M}/\equiv_d} p_{\equiv_d}(M) H_B(M)$。粗略地说，当 $\equiv_d$ 远离均匀分布时，$H_B(\equiv_d)$ 就较大。

分布 $\equiv_d$ 的玻尔兹曼熵与 $\equiv_d$ 的香农熵 $H_S(\equiv_d)$ 密切相关。我们将香农熵定义为 $-\log(p_{\equiv_d}(M))$ 在分布 $\equiv_d$ 下的期望值。更明确地说，我们将 $H_S(\equiv_d)$ 定义为 $-\sum_{M \in \mathcal{M}/\equiv_d} p_{\equiv_d}(M) \log(p_{\equiv_d}(M))$。与玻尔兹曼熵不同，香农熵通过考察分布的均匀程度来度量 $\equiv_d$ 的随机性。事实上，如果 $\equiv_d$ 包含一个非常大的类，那么它的香农熵就较小，而玻尔兹曼熵则相对较大。下面的结果在更一般的情形下（记号略有不同）由 [7] 给出，它正式确立了这两种熵的概念在本质上互为补充。

命题 2.1. $H_S(\equiv_d) + H_B(\equiv_d) = |\tau| n$

## 3 描述复杂度

本节将探讨划分 $\equiv_d$ 中等价类的 GMLU$_d$ 描述复杂度。我们将运用 GMLU$_d$ 的公式大小博弈。

令 $I = \{1, \dots, t\}$，并令 $(\pi_i)_{i\in I}$ 为命题符号集合 $\tau$ 的类型枚举。令 $d \le n$ 为计数深度。我们考虑由 GMLU$_d$ 对大小为 $n$的模型所诱导的划分。对于每个可容许元组 $\overline{n} = (n_i)_{i\in I}$，都有一个等价类 $M_{\overline{n}}$，其中每个类型 $\pi_i$ 出现 $n_i$ 次；$n_i = d$ 表示类型 $\pi_i$ 至少出现$d$ 次。我们用 $k_d$表示 $d$ 在 $\overline{n}$ 中的出现次数。

这样的类 $M_{\overline{n}}$ 可以通过下面的 GMLU$_d$ 公式定义：

$$\varphi(\overline{n}) := \bigwedge_{n_i<d} \Diamond^{=n_i}\psi(\pi_i) \wedge \bigwedge_{n_i=d} \Diamond^{\ge d}\psi(\pi_i)$$

公式 $\varphi(\overline{n})$ 的大小为 $\sum_{i\in I} n_i + t(2|\tau| + 1) - k_d - 1$。

如果 $n_i = d$ 至多对一个 $i \in I$成立，那么 $M_{\overline{n}}$ 可以由一个更小的公式定义。令 $\pi_j$ 为 $\overline{n}$中 $n_j$ 最大的类型。现在 $M_{\overline{n}}$ 也可以由公式定义为

$$\varphi'(\overline{n}) := \bigwedge_{i\neq j} \Diamond^{=n_i}\psi(\pi_i).$$

为了解释这一点，回想一下，我们将模型限制为大小为 $n$ 的模型。由于所有其他类型都已被精确指定，剩余点唯一可选的类型就是 $\pi_j$。$\varphi'(\overline{n})$ 的大小为 $n - |\pi_j| + (t-1)(2|\tau|+1) - 2$，其中 $|\pi_j|$ 表示类型为 $\pi_j$ 的 $M_{\overline{n}}$ 模型中的点数。

我们定义常数 $c_\tau := t(2|\tau| + 1) - 1$。根据上述公式可知，若 $k_d \ge 2$，则 $C(M_{\overline{n}}) \le \sum_{i\in I} n_i + c_\tau$；若 $k_d \le 1$，则 $C(M_{\overline{n}}) \le n - |\pi_j| + c_\tau$。我们将使用 GMLU$_d$ 的公式大小博弈来证明，这些界在常数 $c_\tau$ 的意义下是最优的。

设 $r_0 \in \mathbb{N}$，并令 $\mathcal{A}_0$、$\mathcal{B}_0$ 为 $\tau$-模型的集合。GMLU$_d$-公式大小博弈 GAME$_d(r_0, \mathcal{A}_0, \mathcal{B}_0)$ 有两名玩家 S 和 D。博弈的局面形式为 $P = (r, \mathcal{A}, \mathcal{B})$，初始局面为 $P_0 = (r_0, \mathcal{A}_0, \mathcal{B}_0)$。在局面 $P$ 中，若 $r = 0$，则 D 获胜。否则，S 可从以下行动中选择：

$p$-步：S 选择一个 $\tau$-文字 $\alpha$。博弈结束。如果 $\mathcal{A} \vDash \alpha$ 且 $\mathcal{B} \vDash \neg\alpha$，则 S 获胜。否则 D 获胜。如果 S 至今尚未走过模态步，则不能执行此步。

$\vee$-步：S 选择 $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$，使得 $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{A}$ 且 $r_1$，并选择 $r_2 \ge 1$，使得 $r_1 + r_2 + 1 = r$。D 选择下一个局面是 $(r_1, \mathcal{A}_1, \mathcal{B})$ 还是 $(r_2, \mathcal{A}_2, \mathcal{B})$。

$\wedge$-步：与 $\vee$-步相同，只是 $\mathcal{A}$和 $\mathcal{B}$ 的角色互换。

$\blacklozenge^{\geq d}$-步：S 选择一个数 $k \in \mathbb{N}$，满足 $k \leq d$ 和 $k < r$。对于每个 $(\mathfrak{M}, w) \in \mathcal{A}$，S 选择 $k$ 个不同的点 $v \in W$。令 $\mathcal{A}'$ 为以这种方式选出的模型 $(\mathfrak{M}, v)$ 的集合。对于每个 $(\mathfrak{M}, w) \in \mathcal{B}$，S 选择 $n-k+1$ 个不同的点 $v \in W$。再次令 $\mathcal{B}'$ 为选出的模型集合。博弈的下一个局面是 $(r-k, \mathcal{A}', \mathcal{B}')$。

$\blacksquare^{<d}$-步：与 $\blacklozenge^{\geq d}$-步相同，只是交换 $\mathcal{A}$ 和 $\mathcal{B}$ 的角色。

$\blacklozenge^{=d}$-步：S 选择一个满足 $k < d$ 和 $k < r$的数 $k \in \mathbb{N}$。对于每个 $(\mathfrak{M}, w) \in \mathcal{A}$，S 选择一个由 $k$ 个互不相同的点组成的集合 $P_{\mathfrak{M},w}$。令 $N_{\mathfrak{M},w} := W \setminus P_{\mathfrak{M},w}$。对于每个 $(\mathfrak{M}, w) \in \mathcal{B}$，S 选择一个由 $k+1$个互不相同的点组成的集合 $P_{\mathfrak{M},w}$，或一个由 $|\mathfrak{M}|-k+1$ 个互不相同的点组成的集合 $N_{\mathfrak{M},w}$。最后，令$\mathcal{A}' := \{(\mathfrak{M}, v) \mid (\mathfrak{M}, w) \in \mathcal{A} \cup \mathcal{B}, v \in P_{\mathfrak{M},w}\}$，并令 $\mathcal{B}' := \{(\mathfrak{M}, v) \mid (\mathfrak{M}, w) \in \mathcal{A} \cup \mathcal{B}, v \in N_{\mathfrak{M},w}\}$。博弈的下一个局面是 $(r-k-1, \mathcal{A}', \mathcal{B}')$。

$\blacksquare^{\neq d}$步：与 $\blacklozenge^{=d}$ 步相同，但交换 $\mathcal{A}$ 和 $\mathcal{B}$ 的角色。

公式大小博弈刻画了能够区分模型类的公式大小。以下定理将这一点形式化：

定理 3.1。以下陈述等价：

1. S 在博弈 $\mathrm{GAME}_d(r, \mathcal{A}, \mathcal{B})$ 中有获胜策略。2. 存在一个 $\varphi \in \mathrm{GMLU}_d[\tau]$，其大小至多为 $r$，使得 $\mathcal{A} \vDash \varphi$ 和$\mathcal{B} \vDash \neg\varphi$。

证明。归纳证明很简单。参见 [5]，其中给出了基本模态逻辑的一个证明版本。□

下面的证明中，注意到如果博弈双方的模型（本质上）相同，那么 D 就有一个简单的获胜策略，这一点会很有用。

引理 3.2。设 $P = (r, \mathcal{A}, \mathcal{B})$ 是博弈 $\mathrm{GAME}_d(r_0, \mathcal{A}_0, \mathcal{B}_0)$ 的一个局面。设同一模型 $\mathfrak{M}$ 存在命题等价版本 $(\mathfrak{M}, w) \in \mathcal{A}$ 以及 $(\mathfrak{M}, v) \in \mathcal{B}$。现在，D 从局面 $P$ 开始拥有获胜策略。

证明。容易看出，同一模型的命题等价版本对会在 S 的任何模态步中保持不变。对于 $\vee$-步和 $\wedge$-步，接下来的两个可能局面中总有一个包含这样一对模型，而 D 的策略就是选择这个局面。□

令 $\overline{n}$ 为一个可容许元组，并假设 $n_1$ 是最大的坐标之一。我们首先定义该博弈的模型集合 $\mathcal{A}_{\overline{n}}$ 和 $\mathcal{B}_{\overline{n}}$。所有模型都具有相同的论域 $W = \{1, \ldots, n\}$。我们记 $\mathrm{supp}(\overline{n}) := \{i \in I \mid n_i > 0\}$。首先按如下方式定义模型 $\mathfrak{M}_0$。对于每个 $i \in I, i \neq 1$，类型 $\pi_i$ 恰好实现 $n_i$ 次。类型 $\pi_1$ 实现 $n - \sum_{i \neq 1} n_i$次。此外，点 1 的类型为 $\pi_1$。直观而言，模型 $\mathfrak{M}_0$ 是类 $M_{\overline{n}}$的一个模型，其中所有未被元组 $\overline{n}$ 固定的点都属于类型$\pi_1$。我们令 $\mathcal{A}_{\overline{n}} := \{(\mathfrak{M}_0, 1)\}$。

对于集合 $\mathcal{B}_{\overline{n}}$，我们为满足某些条件的数对 $(i,j)$ 定义模型 $\mathfrak{M}_{i\to j}$：$\in \mathsf{supp}(\overline{n}) \times \mathsf{supp}(\overline{n})$，如下所示。对于 $i \neq 1$以及任意 $j \in \mathsf{supp}(\overline{n})$，模型 $\mathfrak{M}_{i\to j}$ 由模型 $\mathfrak{M}_0$ 通过将一个类型为 $\pi_i$ 的点改为类型 $\pi_j$ 而得到。如果 $i=1$ 且 $n_j = d$，则模型 $\mathfrak{M}_{i\to j}$ 由 $\mathfrak{M}_0$ 通过将 $|\pi_1|_{\mathfrak{M}_0} - d + 1$ 个类型为 $\pi_1$ 的点 $w \neq 1$ 改为类型 $\pi_j$而得到。如果 $i = 1$ 且 $n_j < d$，则数对 $(i,j)$ 不定义模型。我们令 $\mathcal{B}_{\overline{n}} := \{(\mathfrak{M}_{i\to j}, 1) \mid i,j \in \mathsf{supp}(\overline{n})\}$。

就其元组 $\overline{n}'$而言，模型 $\mathfrak{M}_{i\to j}$ 具有 $n'_i = n_i - 1$ 和$n'_j = n_j + 1$，但若 $n_j = d$，则情况为 $n'_j = n_j = d$。对于所有其他指标 $\ell$，$n'_\ell = n_\ell$。所有模型 $(\mathfrak{M}_0, 1)$ 和 $(\mathfrak{M}_{i\to j}, 1)$ 都是命题等价的，因为它们在点 1 实现了类型 $\pi_1$。

设 $P = (r, \mathcal{A}, \mathcal{B})$ 是博弈 $\mathsf{GAME}_d(r, \mathcal{A}_{\overline{n}}, \mathcal{B}_{\overline{n}})$ 的一个局面。我们定义有向图 $\mathcal{G}(\mathcal{A},\mathcal{B}) := (V, E)$，其中 $V = \mathsf{supp}(\overline{n})$，且 $(i,j) \in E$，当且仅当存在命题等价的 $(\mathfrak{M}_0, w) \in \mathcal{A}$ 和 $(\mathfrak{M}_{i\to j}, v) \in \mathcal{B}$，或者反之（关于 $\mathcal{A}$ 和 $\mathcal{B}$）。若对于每个 $(i,j) \in E$，都有 $i \in C$ 或者 $j \in C$和$n_j < d$都成立，则称集合 $C \subseteq \mathsf{supp}(\overline{n})$ 是$\mathcal{G}(\mathcal{A},\mathcal{B})$ 的一个覆盖。覆盖 $C$ 的代价 $r(C)$ 为

$$r(C) := \sum_{i\in C} n_i.$$

直观地说，覆盖的定义意味着，纳入索引 $i \in I$ 通常会覆盖所有指向 $i$ 和从 $i$ 出发的边。例外情况是，当 $n_i = d$ 时，入边不会被覆盖。

在下面引理的证明中，我们用记号 $\mathsf{Md}(\mathcal{A}) = \{\mathfrak{M} \mid (\mathfrak{M}, w) \in \mathcal{A}, w \in W\}$ 表示至少有一个带指点版本属于带指点模型集合 $\mathcal{A}$ 的模型集合。我们还用 $\mathsf{tp}(A)$ 表示模型 $\mathfrak{M}$ 中点集 $A$ 所实现的命题类型集合；下文将根据点集明确模型 $\mathfrak{M}$。

引理 3.3. 设 $P = (r, \mathcal{A}, \mathcal{B})$ 是博弈 $\mathsf{GAME}_d(r, \mathcal{A}_{\overline{n}}, \mathcal{B}_{\overline{n}})$ 的一个局面，并设

$$R(P) := \min\{r(C) \mid C \text{ is a cover of } \mathcal{G}(\mathcal{A},\mathcal{B})\}.$$

若 $r < R(P)$，则 D 从局面 $P$ 出发有获胜策略。

证明。我们证明，对于 S 的所有可能着法，要么条件 $r < R(P)$ 得以保持，要么 D 因其他原因有获胜策略。由于图 $\mathcal{G}(\mathcal{A},\mathcal{B})$ 的定义关于 $\mathcal{A}$ 和 $\mathcal{B}$ 是对称的，因此我们只需处理每对对偶着法中的一种。

$p$-步：由于 $0 < r < R(P)$，博弈双方都有命题等价模型，因此如果 S 做出任何 $p$-步，D 就会获胜。

$\vee$步：设 $\mathcal{A}_1$、$\mathcal{A}_2 \subseteq \mathcal{A}$和 $r_1$、$r_2 \geq 1$ 分别是 S 的选择，并设 $P_1 = (r_1, \mathcal{A}_1, \mathcal{B})$和 $P_2 = (r_2, \mathcal{A}_2, \mathcal{B})$。对于每个$(i,j) \in E$，存在一对

命题等价模型$(\mathfrak{M}, w) \in \mathcal{A}$ 和$(\mathfrak{M}', v) \in \mathcal{B}$ 作为见证。由于 $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{A}$，每个模型$(\mathfrak{M}, w) \in \mathcal{A}$ 都属于 $\mathcal{A}_1$ 或 $\mathcal{A}_2$。集合 $\mathcal{B}$ 保持不变，因此对于每个$(i, j) \in E$，我们有$(i, j) \in E_1$ 或$(i, j) \in E_2$。现在，若$r_1 \geq R(P_1)$且 $r_2 \geq R(P_2)$，那么局面 $P_1$ 存在一个极小覆盖 $C_1$，满足 $r(C_1) \leq r_1$，对 $P_2$也同样如此。现在集合 $C :$ $= C_1 \cup C_2$ 是局面 $P$ 的一个覆盖，满足 $r(C) \leq r_1 + r_2 < r$，这就产生了矛盾。因此，我们有 $r_1 < R(P_1)$或 $r_2 < R(P_2)$，D 可以通过选择这样一个局面来维持该条件。

$\blacklozenge^{\geq d}$-步：设 $k \leq d$ 是 S 选择的数。接下来的局面是 $P' = (r - k, \mathcal{A}', \mathcal{B}')$。首先注意，如果 $\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{A}) \cap \mathsf{Md}(\mathcal{B})$，那么 S 必须从 $\mathcal{A}$ 中的版本选取 $k$ 个点，并从 $\mathcal{B}$中的版本选取 $n - k + 1$ 个点。因此，下一个局面 $P'$ 将包含命题等价版本 $(\mathfrak{M}_0, w) \in \mathcal{A}'$和 $(\mathfrak{M}_0, v) \in \mathcal{B}'$。根据引理 3.2，这使 D 拥有获胜策略，因此我们假设 $\mathfrak{M}_0$ 只在博弈的一方出现。

情形 $\mathcal{A}$: $\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{A})$。对于每个 $(\mathfrak{M}_0, w) \in \mathcal{A}$，S 选择一个包含 $k$个点的集合 $P_{\mathfrak{M}_0, w}$。令 $P_{\mathfrak{M}_0} := \bigcup_{w \in W} P_{\mathfrak{M}_0, w}$。我们考虑以下情形:

1）我们有 $\sum_{\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i > k$。令 $(i, j) \in E$。由于模型 $\mathfrak{M}_{i \to j}$与模型 $\mathfrak{M}_0$ 的差异仅在于 $n_i' = n_i - 1$ 和 $n_j' \geq n_j$，模型 $\mathfrak{M}_{i \to j}$ 至少有 $k$ 个点，其类型来自 $\mathsf{tp}(P_{\mathfrak{M}_0})$。因此，当 S 选择包含 $n - k + 1$ 个点的集合 $N_{\mathfrak{M}_{i \to j}}$时，其中至少有一个点的类型来自 $\mathsf{tp}(P_{\mathfrak{M}_0})$。这样便维持了这对命题等价模型，并且 $(i, j) \in E'$。

2) 我们有 $\sum_{\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i = k$。令 $(i, j) \in E$。假设 $\pi_i \notin \mathsf{tp}(P_{\mathfrak{M}_0})$ 或$\pi_j \in \mathsf{tp}(P_{\mathfrak{M}_0})$。现在模型 $\mathfrak{M}_{i \to j}$ 至少有 $k$ 个点，其类型来自 $\mathsf{tp}(P_{\mathfrak{M}_0})$，且 $(i, j) \in E'$，如情形 1 所示。

因此假设 $\pi_i \in \mathsf{tp}(P_{\mathfrak{M}_0})$ 且 $\pi_j \notin \mathsf{tp}(P_{\mathfrak{M}_0})$。消除所有这类边属于可接受的最坏情况，因此我们假设 $(i, j) \notin E'$。

总结情形 $\mathcal{A}$，最坏情况是 S 选择一个由类型构成的集合 $\mathsf{tp}(P_{\mathfrak{M}_0})$，其中 $\sum_{\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i = k \leq d$，并消除所有满足 $\pi_i \in \mathsf{tp}(P_{\mathfrak{M}_0})$ 且 $\pi_j \notin \mathsf{tp}(P_{\mathfrak{M}_0})$ 的边。

情形 $\mathcal{B}$: $\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{B})$。对于每个 $(\mathfrak{M}_0, w) \in \mathcal{B}$，S 选择一个由 $N_{\mathfrak{M}_0, w}$ 个点组成的集合。令 $N_{\mathfrak{M}_0} := \bigcup_{w \in W} N_{\mathfrak{M}_0, w}$，并令 $\Pi$ 为 $\mathsf{tp}(N_{\mathfrak{M}_0})$ 的补集。注意，对于每个 $\pi_i \in \Pi$，$n_i < d$，因为 $\mathfrak{M}_0$ 至多只有 $k - 1$ 个点，其类型来自 $\Pi$。

令 $(i, j) \in E$并假设 $\pi_i \in \Pi$或 $\pi_j \notin \Pi$。现在 $\pi_i \notin \mathsf{tp}(N_{\mathfrak{M}_0})$ 或 $\pi_j \in \mathsf{tp}(N_{\mathfrak{M}_0})$，因此模型 $\mathfrak{M}_{i \to j}$ 至少有 $n - k + 1$ 个点，其类型来自 $\mathsf{tp}(N_{\mathfrak{M}_0})$。因此，任意集合 $P_{\mathfrak{M}_{i \to j}, v}$ 个点中都至少包含一个类型来自 $\mathsf{tp}(N_{\mathfrak{M}_0})$ 的点。由此，命题等价模型这一对得以保持，且 $(i, j) \in E'$。

现在假设 $\pi_i \notin \Pi$ 和 $\pi_j \in \Pi$。如果 $\mathfrak{M}_{i \to j}$ 至少有 $n - k + 1$ 个点，其类型来自 $\mathsf{tp}(N_{\mathfrak{M}_0})$，那么 $(i, j)$与上文相同。因此我们假设$\mathfrak{M}_{i \to j}$ 中类型来自 $\mathsf{tp}(N_{\mathfrak{M}_0})$的点少于 $n - k + 1$ 个，这意味着它

至少有 $k$ 个点，其类型来自 $\Pi$。由于 $n_j < d$ 和 $\mathfrak{M}_{i\to j}$ 与$\mathfrak{M}_0$ 的差异仅在于 $n'_i = n_i - 1$ 和 $n'_j = n_j + 1$，唯一剩下的可能是$\mathfrak{M}_{i\to j}$ 恰好有 $k$ 个点，其类型来自 $\Pi$。因此，按照同样的推理，$\mathfrak{M}_0$ 恰好有 $k-1$ 个点，其类型来自 $\Pi$；又由于 $k-1 < d$，我们有$\sum_{i\in\Pi} n_i = k-1$。我们再次将消去这些边的 S 视为最坏情况，并假设 $(i,j) \notin E'$。

总结案例 $\mathcal{B}$，最坏情况是 S 选择一组 $\Pi$ 类型，其中包含 $\sum_{\pi_i\in\Pi} n_i = k-1 < d$，并消除所有具有 $\pi_i \notin \Pi$ 和 $\pi_j \in \Pi$的边。

现在考虑以下局面中的条件 $r - k < R(P')$。根据上述论证，我们可以看出，S 消除从 $\mathcal{G}(\mathcal{A},\mathcal{B})$ 指向 $\mathcal{G}(\mathcal{A}',\mathcal{B}')$ 的边的唯一方法，是在每个模型 $(\mathfrak{M}_0, w)$ 或 $(\mathfrak{M}_{i\to j}, v)$ 中，从 $\mathcal{A}$ 恰好选出所有满足某组 $\Pi$ 类型的点。若$\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{A})$，则有 $\sum_{\pi_i\in\Pi} n_i = k \le d$，S 可以消除从 $\Pi$ 中的类型指向其他类型的所有边。若 $\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{B})$，则有 $\sum_{\pi_i\in\Pi} n_i = k-1$，并且对所有 $\pi_i \in \Pi$都有$n_i < d$。此时，S 可以消除从其他类型指向 $\Pi$中的类型的所有边。在这两种情况下，集合 $C = \{i \in \mathsf{supp}(\overline{\pi}) \mid \pi_i \in \Pi\}$ 覆盖所有被消除的边。$C$ 的代价为 $r(C) = \sum_{i\in C} n_i \le k$。设 $C'$ 是 $P'$ 的一个代价最小的覆盖，则 $r(C') = R(P')$。现在 $C \cup C'$ 是 $P$的一个覆盖，因此$r < R(P) \le R(P') + r(C) \le R(P') + k$。因此 $r - k < R(P')$。

♦$^{=d}$步：令 $k < d$ 为 S 选择的数。下一个局面为 $P' = (r-k-1, \mathcal{A}', \mathcal{B}')$。至于 ♦$^{\ge d}$步，我们可以假设 $\mathfrak{M}_0$ 只出现在博弈的一方。

情形 $\mathcal{A}$：$\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{A})$。对于每个 $(\mathfrak{M}_0, w) \in \mathcal{A}$，S 选择一个 $k$ 点集$P_{\mathfrak{M}_0,\, w}$ 和 $N_{\mathfrak{M}_0,\, w} = W \setminus P_{\mathfrak{M}_0,\, w}$。我们记 $P_{\mathfrak{M}_0} := \bigcup_{w\in W} P_{\mathfrak{M}_0,\, w}$ 和$N_{\mathfrak{M}_0} := \bigcup_{w\in W} N_{\mathfrak{M}_0,\, w}$。我们考虑以下情形：

1) 我们有 $\sum_{\pi_i\in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i > k$。因此，存在命题等价的$w \in P_{\mathfrak{M}_0}$和 $v \in N_{\mathfrak{M}_0}$。这意味着，在下一个局面 $P'$中，博弈双方都有模型 $\mathfrak{M}_0$ 的命题等价版本。根据引理 3.2，D 从 $P'$ 开始有获胜策略。

2) 我们有 $\sum_{\pi_i\in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i = k$。注意，由于 $k < d$，对于$\pi_i \in \mathsf{tp}(P_{\mathfrak{M}_0})$中的所有 $n_i < d$ 也成立。令 $(i,j) \in E$。

假设 $\pi_i, \pi_j \notin \mathsf{tp}(P_{\mathfrak{M}_0})$。现在模型 $\mathfrak{M}_{i\to j}$恰好有 $k$ 个点，其类型来自 $\mathsf{tp}(P_{\mathfrak{M}_0})$。因此，如果 S 选择一个 $k+1$点集 $P_{\mathfrak{M}_{i\to j}}$，那么其中一个点 $v'$ 的类型为 $\pi_\ell \notin \mathsf{tp}(P_{\mathfrak{M}_0})$。根据模型的定义，$n_\ell \ne 0$ 因此模型 $\mathfrak{M}_0$ 中存在一个类型为 $\pi_\ell$的点 $w'$。现在有命题等价的 $(\mathfrak{M}_{i\to j}, v') \in \mathcal{A}'$ 和 $(\mathfrak{M}_0, w') \in \mathcal{B}'$，所以 $(i,j) \in E'$。

类似地，如果 S 为 $\mathfrak{M}_{i\to j}$ 选择一个由 $n-k+1$ 个点组成的集合 $N_{\mathfrak{M}_{i\to j}}$，那么该集合至少包含一个点 $v'$，其类型为 $\pi_\ell \in \mathsf{tp}(P_{\mathfrak{M}_0})$。现在存在一个类型为 $\pi_\ell$的点 $w' \in P_{\mathfrak{M}_0}$。因此，存在命题等价的 $(\mathfrak{M}_0, w') \in \mathcal{A}'$ 和 $(\mathfrak{M}_{i\to j}, v') \in \mathcal{B}'$，所以 $(i,j) \in E'$。

如果 $\pi_i \in \mathsf{tp}(P_{\mathfrak{M}_0})$ 或 $\pi_j \in \mathsf{tp}(P_{\mathfrak{M}_0})$，我们假设 $(i,j) \notin E'$。至于

$\blacklozenge^{\geq d}$-移动，这是我们下界的一个可接受的最坏情况。

综上，情况 $\mathcal{A}$ 中，最坏情况是 S 选择类型集合 $\mathrm{tp}(P_{\mathfrak{M}_0})$，满足 $\sum_{\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})} n_i = k < d$，并消除所有满足 $\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})$ 或 $\pi_j \in \mathrm{tp}(P_{\mathfrak{M}_0})$ 的边 $(i,j) \in E$。

情况 $\mathcal{B}$: $\mathfrak{M}_0 \in \mathrm{Md}(\mathcal{B})$。对于每个 $(\mathfrak{M}_0, w) \in \mathcal{B}$，S 选择一个由 $k+1$ 个点组成的集合 $P_{\mathfrak{M}_0,w}$，或一个由 $n-k+1$ 个点组成的集合 $N_{\mathfrak{M}_0,w}$。共有三种情况：
1) 首先假设存在 $(\mathfrak{M}_0, w)$、$(\mathfrak{M}_0, w') \in \mathcal{B}$，并选定了集合 $P_{\mathfrak{M}_0,w}$ 和 $N_{\mathfrak{M}_0,w'}$。由于 $k+1+n-k+1 > n$，存在 $v \in P_{\mathfrak{M}_0,w} \cap N_{\mathfrak{M}_0,w'}$。在接下来的局面 $P'$ 中，模型 $(\mathfrak{M}_0, v)$ 位于博弈的双方，因此根据引理 3.2，D 从 $P'$ 开始有获胜策略。

2) 因此，假设 S 对每个 $(\mathfrak{M}_0, w) \in \mathcal{B}$ 都选择一个由 $k+1$ 个点组成的集合 $P_{\mathfrak{M}_0,\ w}$，并令 $P_{\mathfrak{M}_0} = \bigcup_{w \in W} P_{\mathfrak{M}_0,\ w}$。令 $(i,j) \in E$，并令 $(\mathfrak{M}_{i\to j}, v) \in \mathcal{A}$ 和 $(\mathfrak{M}_0, w) \in \mathcal{B}$ 为相应的模型。

假设 $\pi_i \notin \mathrm{tp}(P_{\mathfrak{M}_0})$ 或 $\pi_j \in \mathrm{tp}(P_{\mathfrak{M}_0})$。现在，模型 $\mathfrak{M}_{i\to j}$ 至少有 $k+1$ 个点，其类型来自 $\mathrm{tp}(P_{\mathfrak{M}_0})$。因此，大小为 $n-k$ 的集合 $N_{\mathfrak{M}_{i\to j},\ v}$ 至少有一个点，其类型来自 $\mathrm{tp}(P_{\mathfrak{M}_0})$。这样就保持了命题等价的模型对，且 $(i,j) \in E'$。

假设 $\pi_i \in \mathrm{tp}(P_{\mathfrak{M}_0})$ 且 $\pi_j \notin \mathrm{tp}(P_{\mathfrak{M}_0})$。如果 $\mathfrak{M}_{i\to j}$ 至少有 $k+1$ 个点，其类型来自 $\mathrm{tp}(P_{\mathfrak{M}_0})$，那么上述论证仍然成立，且 $(i,j) \in E'$。由于 $k < d$ 和 $\mathfrak{M}_{i\to j}$ 与 $\mathfrak{M}_0$ 的差异仅在于 $n'_i = n_i - 1$ 和 $n'_j \leq n_j + 1$，唯一剩下的可能是，$\mathfrak{M}_{i\to j}$ 恰好有 $k$ 个点，其类型来自 $\mathrm{tp}(P_{\mathfrak{M}_0})$。我们得到 $\sum_{i \in \mathrm{tp}(P_{\mathfrak{M}_0})} n'_i = k$。我们再次假设最坏情况下 $(i,j) \notin E'$。

3）最后，假设 S 对每个 $(\mathfrak{M}_0, w)$ 都选择一个由 $n-k+1$ 个点组成的集合 $N_{\mathfrak{M}_0,\ w}$，并令 $N_{\mathfrak{M}_0} = \bigcup_{w \in W} N_{\mathfrak{M}_0,\ w}$。令 $\Pi$ 为 $\mathrm{tp}(N_{\mathfrak{M}_0})$ 的补集。

假设 $\pi_i \in \Pi$ 或 $\pi_j \notin \Pi$。现在 $\pi_i \notin \mathrm{tp}(N_{\mathfrak{M}_0})$ 或 $\pi_j \in \mathrm{tp}(N_{\mathfrak{M}_0})$，因此模型 $\mathfrak{M}_{i\to j}$ 至少有 $n-k+1$ 个点，其类型来自 $\mathrm{tp}(N_{\mathfrak{M}_0})$。因此，大小为 $k$ 的集合 $P_{\mathfrak{M}_{i\to j},\ v}$ 至少有一个点，其类型来自 $\mathrm{tp}(N_{\mathfrak{M}_0})$。由此，命题等价的模型对得以保持，且 $(i,j) \in E'$。

假设 $\pi_i \notin \Pi$ 且 $\pi_j \in \Pi$。如果 $\mathfrak{M}_{i\to j}$ 至少有 $n-k+1$ 个点，其类型来自 $\mathrm{tp}(N_{\mathfrak{M}_0})$，那么上述论证仍然成立，且 $(i,j) \in E'$。

我们假设 $\mathfrak{M}_{i\to j}$ 中类型属于 $\mathrm{tp}(N_{\mathfrak{M}_0})$ 的点少于 $n-k+1$ 个，这意味着它至少有 $k$ 个点的类型属于 $\Pi$。由于 $k < d$ 和 $\mathfrak{M}_{i\to j}$ 与 $\mathfrak{M}_0$ 仅相差 $n'_i = n_i - 1$ 和 $n'_j \leq n_j + 1$，唯一剩下的可能就是 $\mathfrak{M}_{i\to j}$ 中类型属于 $\Pi$ 的点恰好有 $k$ 个。我们得到 $\sum_{i \in \Pi} n'_i = k$。我们再次假设最坏情况下 $(i,j) \notin E'$。以这种方式消去的边，是从其他类型指向 $\Pi$ 中类型的边。特别地，我们注意到，对于任何以这种方式消去的边 $(i,j)$，$n_j < d$，因为 $\mathfrak{M}_0$ 中类型为 $\pi_j$ 的点至多有 $n-(n-k+1) = k-1$ 个。

总结情况 $\mathcal{B}$，最坏情况是 S 选择一组类型 $\Pi$，并消除所有边（$i$，$j$），满足 $\pi_i \in \Pi$ 或 $\pi_j \notin \Pi$，或者消除所有边

（$i$, $j$），满足 $\pi_i \notin \Pi$ 或 $\pi_j \in \Pi$。在这两种情况下，$\sum_{i\in\Pi} n'_i = k$ 对于模型 $\mathfrak{M}_{i\to j}$，而对于后一种情况，$n_j < d$。

现在我们考虑以下局面中的条件 $r-k-1 < R(P')$。根据上述论证，我们可以看出，S 消除从 $\mathcal{G}(\mathcal{A},\mathcal{B})$ 到 $\mathcal{G}(\mathcal{A}',\mathcal{B}')$ 的边的唯一方法，是在每个模型 $(\mathfrak{M}_0, w)$ 或 $(\mathfrak{M}_{i\to j}, v)$ 中，在 $\mathcal{A}$ 恰好选出满足某个类型集 $\Pi$ 的所有点。若$\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{A})$，则有 $\sum_{\pi_i\in\Pi} n_i = k < d$，S 可以消除通向和来自该类型集 $\Pi$ 的所有边。若 $\mathfrak{M}_0 \in \mathsf{Md}(\mathcal{B})$，则有 $\sum_{\pi_i\in\Pi} n'_i = k < d$，S 可以消除通向或来自该类型集 $\Pi$的所有边。特别地，若 $n_j = d$ 对于 $\pi_j \in \Pi$成立，则只能消除出边。在所有情况下，集合 $C = \{i \in \mathsf{supp}(\overline{n}) \mid \pi_i \in \Pi\}$ 都覆盖了所有被消除的边。$C$ 的代价为 $r(C) = \sum_{i\in C} n_i \le k+1$，因为 $\sum_{i\in C} n_i \le \sum_{i\in C} n'_i + 1$。设 $C'$ 是 $P'$ 的一个覆盖，且代价最小，因此 $r(C') = R(P')$。现在 $C \cup C'$是 $P$的一个覆盖，所以 $r < R(P) \le R(P') + r(C) \le R(P') + k + 1$。因此 $r - k - 1 < R(P')$。 □

现在，只需计算 $M_{\overline{n}}$ 的描述复杂度下界，即求 $\mathcal{G}(\mathcal{A}_{\overline{n}},\mathcal{B}_{\overline{n}})$ 的覆盖的最小代价。

引理 3.4。设 $M_{\overline{n}}$ 是一个类，且至少有两个不同的 $i \in I$ 满足 $n_i = d$。那么 $C(M_{\overline{n}}) \ge \sum n_i.i\in I$

证明。假设 $n_1 = d$。考虑图 $\mathcal{G}(\mathcal{A}_{\overline{n}},\mathcal{B}_{\overline{n}}) = \{(i,j) \in \mathsf{supp}(\overline{n}) \mid i \ne 1$ 或 $n_j = d\}$。该图包含从所有满足 $n_i \ne 0$ 的类型 $\pi_i$ 到彼此的边，但 $\pi_1$ 到满足 $n_j < d$ 的 $\pi_j$ 之间没有边。

首先注意到 $C = \mathsf{supp}(\overline{n})$ 是 $\mathcal{G}(\mathcal{A}_{\overline{n}},\mathcal{B}_{\overline{n}})$ 的一个覆盖，其代价为 $\sum_{i\in I} n_i$。对于 $C' \ne C$，如果 $i \notin C'$ 满足 $i \ne 1$，那么根据覆盖的定义，边 $(i,1) \in E$ 未被覆盖，因为 $i \notin C$ 且 $n_1 = d$。如果 $1 \notin C'$，那么令 $j \in I$ 为除 1 以外满足 $n_j = d$ 的另一个索引。此时边 $(1,d) \in E$ 未被覆盖，因为 $1 \notin C$ 且 $n_j = d$。由此可知，$C$ 是代价最小的覆盖；根据引理 3.3 和定理 3.1，结论成立。

□

引理 3.5. 设 $M_{\overline{n}}$ 是一个类，其中至多有一个 $i \in I$满足 $n_i = d$。设 $\pi_j$是 $M_{\overline{n}}$中实现点最多的类型之一。则 $C(M_{\overline{n}}) \ge \sum_{i\in I\setminus\{j\}} n_i = n - |\pi_j|$。

证明. 我们假设 $n_1 = \max\{n_i \mid i \in I\}$。因此 $\pi_1$ 是实现点最多的类型之一，并且如果 $n_1 < d$，那么对于所有 $i \in I$，$n_i < d$。

我们考虑图 $\mathcal{G}(\mathcal{A}_{\overline{n}},\mathcal{B}_{\overline{n}})$。根据定义可得 $G(\mathcal{A}_{\overline{n}},\mathcal{B}_{\overline{n}}) = \{(i,j) \in \mathsf{supp}(\overline{n}) \mid i \ne 1\}$。该图中，除没有从 1 出发的边外，所有类型 $\pi_i$ with $n_i \ne 0$之间都存在边。

我们首先注意到 $C = \text{supp}(\overline{n}) \setminus \{1\}$ 是 $\mathcal{G}(\mathcal{A}_{\overline{n}}, \mathcal{B}_{\overline{n}})$ 的一个覆盖，其中 $r(C) = \sum_{i \in I \setminus \{1\}} n_i = n - |\pi_1|$。显然，$\text{supp}(\overline{n})$ 是一个代价更高的覆盖。

设 $C' \subset \text{supp}(\overline{n})$。如果存在 $i, j \in \text{supp}(\overline{n})$，且 $i, j \notin C'$，那么边 $(i, j)$ 或 $(j, i)$ 属于 $E$，并且未被覆盖。如果只有一个 $i \in \text{supp}(\overline{n})$，且 $i \notin C'$ 和 $i \neq 1$，那么 $r(C') = \sum_{i \in I \setminus \{j\}} n_i \geq r(C)$，因为 $\pi_1$ 是具有最大 $n_i$的类型之一。我们看到 $C$ 是代价最小的覆盖，因此由引理 3.3 和定理 3.1，结论成立。□

我们将上述引理归纳为以下定理：

定理 3.6。设 $\overline{n}$ 是一个 $(n, d)$-可容许元组，并令 $M_{\overline{n}}$ 为 $\equiv_d$对应的等价类。如果对于至少两个 $i \in I$，有 $n_i = d$，那么 $C(M_{\overline{n}}) \geq \sum_{i \in I} n_i$。否则，$C(M_{\overline{n}}) \geq \sum_{i \in I \setminus \{j\}} n_i = n - |\pi_j|$。

# 4 单调连接

考虑全体 $(n, d)$-可容许元组上的以下严格偏序。

1. $\overline{n} <_s \overline{n}'$ iff $|M_{\overline{n}}| < |M_{\overline{n}'}|$
2. $\overline{n} <_c \overline{n}'$ iff $C(M_{\overline{n}}) < C(M_{\overline{n}'})$

请注意，$|M_{\overline{n}}| = |M_{\overline{n}'}|$和 $C(M_{\overline{n}}) = C(M_{\overline{n}'})$ 均不一定意味着 $\overline{n} = \overline{n}'$，元组 $(1, 2, d)$ 和 $(2, 1, d)$ 就说明了这一点。我们用 $\leq_s$ 和 $\leq_c$ 分别表示在 $<_s$ 和 $<_c$ 上添加自环后得到的偏序。本节的主要目的是证明，存在一个自然且非平凡的偏序，同时包含于 $\leq_s$ 和 $\leq_c$（前提是 $n$ 相对于 $d$ 足够大），由此便建立了模型类大小与其描述复杂度之间的单调连接。为确立这一结果，我们将运用上一节中发展出的复杂工具，以及涉及 $r$-关联斯特林数的进一步组合论论证。

## 4.1 $r$-关联斯特林数

我们将使用 $r$-关联斯特林数来计算给定模型类中的模型数量。给定正整数 $n$、$m$、$r$，满足 $n \geq mr$，我们定义

$$\left\{ {n \atop m} \right\}_{\geq r}$$

表示将 $[n]$ 划分为 $m$个集合、且每个集合的大小至少为 $r$ 的 $[n]$ 的划分数。当 $r = 1$ 时，这些数也称为第二类斯特林数，它们仅表示将 $[n]$ 划分为 $m$个集合 $[3]$ 的划分数。

以下引理在估算模型类的大小时将起关键作用。尽管证明简单而基础，我们仍未能在现有文献中找到这些估算结果。

引理 4.1。设 $n, m, r \in \mathbb{Z}_+$。假设 $n \geq mr$。那么

$$\frac{m^n}{m^{mr}} \leq \left\{ {n \atop m} \right\}_{\geq r} \leq \frac{m^n}{m!}$$

证明。对于上界，注意到

$$\left\{ {n \atop m} \right\}_{\geq r} \leq \left\{ {n \atop m} \right\}_{\geq 1} \leq \frac{m^n}{m!}.$$

最后一个不等式源于这样一个事实： $m^n$ 计数映射 $f\colon [n] \to [m]$的数量，而 $m! \cdot \left\{ {n \atop m} \right\}_{\geq 1}$ 只计数其中为满射的那些 $f$。对于下界，我们利用这样一个事实： $m! \cdot \left\{ {n \atop m} \right\}_{\geq r}$ 计数满足以下性质的映射 $f\colon [n] \to [m]$ ： 对于每个 $1 \leq k \leq m$ ，都有 $|f^{-1}(\{k\})| \geq r$。为了给出这类函数数量的下界，我们计数满足以下性质的映射 $f : [n] \to [m]$ ： 集合 $\{(k-1) \cdot r + 1, \ldots, k \cdot r\}$ （其中 $1 \leq k \leq m$）映射到不同的元素。由于剩余的 $n - mr$ 个元素可以任意映射，这类映射的数量正好是 $m! \cdot m^{n-mr}$。因此

$$m! \cdot m^{n-mr} \leq m! \cdot \left\{ {n \atop m} \right\}_{\geq r},$$

除以 $m!$ 后即可得到所需的下界。 □

我们注意到，如果 $m$ 和 $r$ 远小于 $n$ ——正如它们在我们的应用中那样——那么引理 4.1 的估计（或许出人意料地）相当精确。

引理 4.1 的一个重要推论是下面的引理。再次强调，我们未能在现有文献中找到这种形式的估计。

引理 4.2。设 $n$、 $m$、 $r \in \mathbb{Z}_+$。若 $n \geq mr + 1$，则

$$\left\{ {n \atop m} \right\}_{\geq r} \leq \frac{m^{mr+1}}{m!} \left\{ {n-1 \atop m} \right\}_{\geq r}.$$

证明。由引理 4.1 立即可得。 □

如果 $m$ 和 $r$ 固定，那么引理 4.2 给出了 $r$ 关联斯特林数随 $n$ 增大时的增长率上界。

## 4.2 模型类的组合学

在本小节中，我们利用上述关于 $r$ 关联斯特林数的结果，根据模型类的 $(n, d)$-可容许元组来研究其大小。我们从以下引理开始，该引理给出了模型类大小的一个简单闭式公式。

引理 4.3。设 $\overline{n}$ 是一个 $(n, d)$-可容许元组。设 $i_1, \dots, i_k \in I$ 是满足 $n_{i_\ell} < d$ 的指标。于是有

$$|M_{\overline{n}}| = \binom{n}{n_{i_1}, \dots, n_{i_k}, m} \cdot k_d! \cdot \left\{ {m \atop k_d} \right\}_{\geq d},$$

其中 $m := n - \sum_{i \notin \{i_1, \dots, i_k\}} n_i$ 且 $k_d := t - k$。

证明。 $M_{\overline{n}}$ 中的每个模型都可以按如下方式构造。
(1) 首先，我们从 $\{1, \dots, n\}$ 中选取 $k$ 个子集，其大小分别为 $n_{i_1}, \dots, n_{i_k}$，并规定第 $i_\ell$ 个集合中的每个元素都实现第 $i_\ell$ 种类型。完成这一步的方法数为 $\binom{n}{n_{i_1} \dots ,, n_{i_k}, m}$。
(2) 然后，我们将大小为 $m$ 的剩余子集划分为 $k_d$ 个部分，每部分的大小至少为 $d$。完成这一步的方法数由关联斯特林数 $\left\{ {m \atop k_d} \right\}_{\geq d}$ 给出。

(3) 对每个部分，我们从剩余的类型中选取一个唯一的类型——其数量为 $k_d$ ——并规定该部分中的每个元素都实现与该部分对应的类型。这样做的方法数为 $k_d!$ 。

将上述各因子相乘即可得到结果。 □

以下引理说明，修改可容许元组时，模型类的大小如何变化。

引理 4.4。固定 $d \in \mathbb{Z}_+$ ，并设 $\overline{n}$ 是一个 $(n, d)$-可容许元组。假设相对于 $d$ 和 $|\tau|$， $n$ 足够大。那么对于每个满足 $n_i < d - 1$ 的 $i \in I$，我们有

$$|M_{\overline{n}}| < |M_{\overline{n}'}|,$$

其中 $n'_i = n_i + 1$ 和 $n'_j = n_j$，对每个 $j \neq i$ 均成立。

证明。为简化记号，我们假设 $n_\ell < d$ 当且仅当 $\ell \leq k$。根据引理 4.3，我们需要证明的不等式为

$$\begin{aligned} &\binom{n}{n_1, \dots, n_k, m} \cdot k_d! \cdot \left\{ {m \atop k_d} \right\}_{\geq d} \\ &< \binom{n}{n_1, \dots, n_i + 1, \dots, n_k, m - 1} \cdot k_d! \cdot \left\{ {m-1 \atop k_d} \right\}_{\geq d}, \end{aligned}$$

其中 $m = n - \sum_{\ell=1}^{k} n_\ell$ 且 $k_d = t - k$。注意，由于相对于 $d$而言， $n$ 足够大， $k_d \geq 1$。化简后可得等价不等式

$$\begin{Bmatrix} m \\ k_d \end{Bmatrix}_{\geq d} < \frac{m}{n_i + 1} \begin{Bmatrix} m-1 \\ k_d \end{Bmatrix}_{\geq d},$$

只要……，这便可由引理 4.2 得出

$$\frac{m}{n_i + 1} > \frac{k_d^{k_d d+1}}{k_d!} \Leftrightarrow n > (n_i + 1)\frac{k_d^{k_d d+1}}{k_d!} + \sum_{\ell=1}^{k} n_\ell$$

利用对每个 $1 \leq \ell \leq k$ 都成立的 $n_\ell < d$，以及 $1 \leq k_d \leq t$，即可得到所需结果。

□

引理 4.5。固定 $d \in \mathbb{Z}_+$，并令 $\overline{n}$ 为一个 $(n, d)$-可容许元组。假设相对于 $d$ 和 $|\tau|$， $n$ 足够大。那么对于每个满足 $n_i < d$ 的 $i \in I$，我们有

$$|M_{\overline{n}}| < |M_{\overline{n}'}|,$$

其中，当 $j = i$ 时为 $n'_j = d$，否则为 $n'_j = n_j$。

证明。为简化记号，我们假设 $n_\ell < d$ 当且仅当 $\ell \leq k$。由引理 4.3，我们需要证明的不等式为

$$\begin{aligned} &\binom{n}{n_1, \ldots, n_i, \ldots, n_k, m} \cdot k_d! \cdot \begin{Bmatrix} m \\ k_d \end{Bmatrix}_{\geq d} \\ &< \binom{n}{n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k, m + n_i} \\ &\cdot (k_d + 1)! \cdot \begin{Bmatrix} m + n_i \\ k_d + 1 \end{Bmatrix}_{\geq d}, \end{aligned}$$

其中 $m = n - \sum_{\ell=1}^{k} n_\ell$ 且 $k_d = t - k$。注意，由于相对于 $d$， $n$ 足够大， $k_d \geq 1$。化简后可得等价的不等式

$$\begin{Bmatrix} m \\ k_d \end{Bmatrix}_{\geq d} < \frac{m! n_i!}{(m + n_i)!} \cdot (k_d + 1) \cdot \begin{Bmatrix} m + n_i \\ k_d + 1 \end{Bmatrix}_{\geq d}$$

由引理 4.1可知

$$\begin{Bmatrix} m + n_i \\ k_d + 1 \end{Bmatrix}_{\geq d} \geq \frac{(k_d + 1)^{m+n_i}}{(k_d + 1)^{(k_d+1)d}}$$

and

$$\begin{Bmatrix} m \\ k_d \end{Bmatrix}_{\geq d} \leq \frac{k_d^m}{k_d!}$$

因此，我们只需证明

$$\frac{k_d^m}{k_d!} < \frac{m!n_i!}{(m+n_i)!} \cdot (k_d+1) \cdot \frac{(k_d+1)^{m+n_i}}{(k_d+1)^{(k_d+1)d}}$$

或者等价地，

$$\begin{aligned}\frac{(k_d+1)^{(k_d+1)d}}{(k_d+1)^{n_i+1}k_d!n_i!} &< \frac{m!(k_d+1)^m}{k_d^m(m+n_i)!} \\ &= \frac{(k_d+1)^m}{k_d^m \prod_{j=0}^{n_i-1}(m+n_i-j)} \\ &= \Bigg(\underbrace{\frac{k_d+1}{k_d}}_{>1}\Bigg)^m \Big/ \prod_{j=0}^{n_i-1}(m+n_i-j)\end{aligned}$$

回顾一下 $0 \le n_i < d$ 和 $1 \le k_d \le t$。因此，在上述不等式中，左边是一个常数，所以只需证明右边的公式随着 $n$增大而趋于无穷（回顾 $m = n - \sum_{\ell=1}^{k} n_\ell$，即 $m$就是 $n$减去一个常数）。然而，这显而易见，因为右边的形式为 $f(n)/g(n)$，其中 $f$ 相对于 $n$ 呈指数增长，而 $g$ 相对于 $n$ 仅呈多项式增长。 □

## 4.3 连接大小与描述复杂度

令 $P_{n,d}$ 表示所有 $(n,d)$-可容许元组的集合。我们在 $P_{n,d}$ 上定义一个自然的偏序 $\preceq$ 如下：当且仅当 $n_i \le n_i'$对每个$i \in I$都成立时，$\overline{n} \preceq \overline{n}'$。写作 $\overline{n} \prec \overline{n}'$ 表示 $\overline{n} \preceq \overline{n}'$ 且 $\overline{n} \ne \overline{n}'$。

引理 4.6. 假设 $n$ 相对于 $d$ 和 $|\tau|$足够大。设 $\overline{n}$ 和 $\overline{n}'$ 是 $(n,d)$-可容许元组，且满足 $\overline{n} \prec \overline{n}'$。那么

$$|M_{\overline{n}}| < |M_{\overline{n}'}|$$

证明. 只需证明：若 $\overline{n}'$ 是 $\overline{n}$的直接后继，则$|M_{\overline{n}}| < |M_{\overline{n}'}|$。现在，对每个 $j \ne i$，必定恰好存在一个 $i \in I$使得 $n_i' = n_i + 1$和 $n_j' = n_j$。若 $n_i' < d$，则结论由引理 4.4 得出；若 $n_i' = d$，则结论由引理 4.5 得出。 □

回顾上一节中的常数 $c_\tau = 2^{|\tau|}(2|\tau|+1) - 1$。我们在 $P_{n,d}$ 上定义另一个偏序 $\preceq_\tau$ 如下：当且仅当以下两个条件成立时，$\overline{n} \preceq_\tau \overline{n}'$：

1. 对于每个 $i \in I$，都有 $n_i \le n_i'$。
2. $\overline{n} = \overline{n}'$ 或 $\sum_{i \in I} (n_i' - n_i) > c_\tau$。

粗略地说，$\overline{n} \preceq_\tau \overline{n}'$ 表示：如果这些元组互不相同，那么它们之间相对于序 $\preceq$ 的距离大于 $c_\tau$。同样，写作 $\overline{n} \prec_\tau \overline{n}'$ 表示 $\overline{n} \preceq_\tau \overline{n}'$ 且 $\overline{n} \neq \overline{n}'$。例如，$(1, d) \prec_\tau (2 + c_\tau, d)$，但 $(1, d) \not\preceq_\tau (1 + c_\tau, d)$。

引理 4.7. 假设相对于 $d$ 和 $|\tau|$，$n$ 足够大。设 $\overline{n}$ 和 $\overline{n}'$ 是 $(n, d)$-可容许元组，且 $\overline{n} \prec_\tau \overline{n}'$。那么我们有

$$C(M_{\overline{n}}) < C(M_{\overline{n}'})$$

证明. 首先假设 $d$ 在 $\overline{n}'$ 中至少出现两次。由定理 3.6 可得 $C(M_{\overline{n}'}) \geq \sum_{i \in I} n'_i$。现在根据 $d$ 是否在 $\overline{n}$ 中至少出现两次，分两种情况讨论。首先假设 $d$ 在 $\overline{n}$ 中恰好出现一次，不妨设为 $n_j = d$。那么 $M_{\overline{n}}$ 可由一个大小为

$$c_\tau + \sum_{i \in I - \{j\}} n_i < \sum_{i \in I - \{j\}} n'_i \leq C(M_{\overline{n}'})$$

因此 $C(M_{\overline{n}}) < C(M_{\overline{n}'})$。另一方面，若 $d$ 在 $\overline{n}$ 中至少出现两次，则 $M_{\overline{n}}$ 可由一个大小为

$$c_\tau + \sum_{i \in I} n_i < \sum_{i \in I} n'_i \leq C(M_{\overline{n}'})$$

因此 $C(M_{\overline{n}}) < C(M_{\overline{n}'})$。

接下来假设 $d$ 在 $\overline{n}'$ 中恰好出现一次，不妨设为 $n'_j = d$。由定理 3.6 可得 $C(M_{\overline{n}'}) \geq \sum_{i \in I - \{j\}} n'_i$。由于 $\overline{n} \prec_\tau \overline{n}'$，对于每个 $i \neq j$，都有 $n_i < d$。由于 $d$ 必须在 $\overline{n}$ 中至少出现一次——因为 $\overline{n}$ 是 $(n, d)$-可容许的——所以我们有 $n_j = d$。现在 $M_{\overline{n}}$ 可由一个大小为

$$c_\tau + \sum_{i \in I - \{j\}} n_i < \sum_{i \in I - \{j\}} n'_i \leq C(M_{\overline{n}'})$$

因此 $C(M_{\overline{n}}) < C(M_{\overline{n}'})$。 □

回顾本节开头引入的 $P_{n,d}$ 上的偏序 $\leq_s$ 和 $\leq_c$。下述定理将模型类的大小与其描述复杂度之间的联系形式化。

定理 4.8。若相对于 $d$ 而言，$n$ 足够大，则

$$\preceq_\tau \subseteq \leq_s \cap \leq_c .$$

特别地，如果 $\overline{n}$ 和 $\overline{n}'$ 是两个不同的 $\preceq_\tau$-可比较元组，那么

$$|M_{\overline{n}}| < |M_{\overline{n}'}| \Leftrightarrow C(M_{\overline{n}}) < C(M_{\overline{n}'}).$$

证明。假设 $\overline{n} \prec_\tau \overline{n}'$。引理 4.6 和引理 4.7 保证，如果相对于 $d$而言， $n$ 足够大，那么 $|M_{\overline{n}}| < |M_{\overline{n}'}|$和 $C(M_{\overline{n}}) < C(M_{\overline{n}'})$。因此$\overline{n} \leq_s \overline{n}'$ 和 $\overline{n} \leq_c \overline{n}'$，这就证明了第一个断言。第二个断言直接由第一个断言得出。 □

上述定理的直观含义是，偏序 $\preceq_\tau$同时近似于 $\leq_s$ 和 $\leq_c$。因此， $\preceq_\tau$ 可以被视为 $\leq_s$ 与 $\leq_c$之间一种高度非平凡的单调连接。

我们注意到，上述定理对玻尔兹曼熵同样成立。这是因为序 $\leq_s$ 与基于玻尔兹曼熵的序相同，因为对数是一个递增函数。我们将这两项主张中的后一项表述为一个推论。

推论 4.9。假设相对于 $d$而言， $n$ 足够大。若 $\overline{n}$ 和 $\overline{n}'$是两个不同的、关于 $\preceq_\tau$ 可比较的元组，则

$$H_B(M_{\overline{n}}) < H_B(M_{\overline{n}'}) \Leftrightarrow C(M_{\overline{n}}) < C(M_{\overline{n}'}).$$

# 5 类大小分布的相变

本节中，我们将关注点从单个类转向由 $\equiv_d$给出的整个概率分布。通过允许 $d$ 依赖于 $n$，我们得到定性结果，将 $d$ 的增长率与支配类的出现联系起来；支配类即包含几乎所有模型的类。对于固定的$n$，我们还得到定量结果，说明 $d$ 与$n$ 之间的关系如何决定 $\equiv_d$ 中是否存在一个类，包含大小为 $n$的所有模型中的多数。我们还将指出这些结果对 GMLU$_d$“平均情形”表达能力的影响。

在本节中，我们继续沿用之前的约定：$t = 2^{|\tau|}$ 表示字母表 $\tau$ 的类型数量 $\pi$。我们从以下观察开始，它说明了随着计数深度增加，分布会发生什么变化。

命题 5.1. 假设 $d < d' \leq n/2$。那么

$$H_S(\equiv_d) < H_S(\equiv_{d'}) \text{ and } H_B(\equiv_d) > H_B(\equiv_{d'}).$$

此外，对于每个 $n/2 \leq d \leq d'$，我们都有

$$H_S(\equiv_d) = H_S(\equiv_{d'}) \text{ and } H_B(\equiv_d) = H_B(\equiv_{d'}).$$

证明。根据命题 2.1 $H_S(\equiv_d) + H_B(\equiv_d) = |\tau|n$。因此，只需在玻尔兹曼熵的情形下证明这些结论。我们首先注意到，如果$n/2 \leq d \leq d'$，那么 $H_B(\equiv_d) = H_B(\equiv_{d'})$，因为逻辑 GMLU$_d$ 已经能够将每个结构刻画到同构为止。

那么假设 $d < d' \leq n/2$。考虑类 $M_{\overline{n}}$，其中 $\overline{n} = (0, \ldots, 0, d, d)$。随着深度增加到 $d' > d$，该类至少被划分为两个较小的类，其元组分别为 $(0, \ldots, 0, d, d')$ 和 $(0, \ldots, 0, d', d)$。与此同时，显然没有任何类的大小增加，因此我们看到 $H_B(\equiv_d) > H_B(\equiv_{d'})$。 □

我们借此机会指出前一个结果的一个简单推论。在 [7] 中已经证明，$H_B(\equiv_n) \sim |\tau|n$，我们的意思是 $\lim_{n\to\infty} H_B(\equiv_n)/|\tau|n = 1$。将此结果与命题 5.1 结合起来，可以直接得到如下结论。

推论 5.2. 对于任意计数深度 $d(n)$，有

$$H_B(\equiv_{d(n)}) \sim |\tau|n$$

证明. 不失一般性，假设对每个 $n$，都有 $d(n) \leq n$。由于对任意类 $M$，$\log(|M|) \leq |\tau|n$，因此有 $H_B(\overline{\equiv}_{d(n)}) \leq |\tau|n$。由于对每个 $\varepsilon > 0$，都有 $H_B(\equiv_n) \sim |\tau|n$，所以当 $n$ 足够大时，$H_B(\equiv_n) \geq (1-\varepsilon)|\tau|n$。由于对每个 $\varepsilon > 0$，都有 $H_B(\overline{\equiv}_{d(n)}) \geq H_B(\equiv_n)$，因此只要 $n$ 足够大，就有 $H_B(\overline{\equiv}_{d(n)}) \geq (1-\varepsilon)|\tau|n$。综合这些界即可得到所需结果。 □

因此，从渐近的角度来看，计数深度 $d$ 对 $n$ 的依赖不会影响 $\equiv_d$ 的玻尔兹曼熵。根据命题 2.1，$\equiv_d$ 的香农熵也同样如此。

现在我们进一步分析计数深度对分布的影响。为了表述一些结果，我们将使用标准的渐近记号，在此回顾如下。设 $f$、$g$：$\mathbb{N} \to \mathbb{R}_{>0}$。我们用 $f = o(g)$ 表示 $\lim_{n\to\infty} f(n)/g(n) = 0$。此外，我们用 $f = \omega(g)$ 表示 $\lim_{n\to\infty} f(n)/g(n) = \infty$。

我们首先证明，如果计数深度相对于 $n$ 增长得足够缓慢，那么该分布中就有一个类，包含几乎所有大小为 $n$ 的模型。更正式地说，固定一个函数 $d$：$\mathbb{Z}_+ \to \mathbb{N}$，从而得到一个等价关系序列 $\overline{\equiv}_{d(n)}$。相对于序列 $\overline{\equiv}_{d(n)}$，类序列 $M(n)$ 是一个函数，它为序列中的每个等价关系输出一个单独的类。每个类序列 $M(n)$ 都自然对应于相应的概率序列

$$p_n := p_{\equiv_{d(n)}}(M(n)) = \frac{|M(n)|}{2^{n|\tau|}}.$$

我们称 $\overline{\equiv}_{d(n)}$ 具有支配类，如果存在一个类序列$M(n)$，使得 $p_n \to 1$当 $n \to \infty$时成立。此时，类序列 $M(n)$称为支配 $\overline{\equiv}_{d(n)}$。直观而言，如果大小为 $n$ 的随机模型属于某个类（序列）的极限概率为 1，那么该类（序列）就是支配类（序列）。

我们的证明使用了广为人知的切尔诺夫界。以下引理需要用到下尾估计，而上尾估计将在本节后面使用。

命题 5.3（Chernoff 界 [12]）。设 $X := \sum_{i=1}^{n} X_i$ 是独立的0-1值随机变量之和，其中 $X_i = 1$ 的概率为 $p$，$X_i = 0$ 的概率为 $1 - p$。那么，对于任意 $\delta \geq 0$，都有

$$\textbf{\textit{(Lower tail)}}\ \Pr[X \leq (1-\delta)np] \leq e^{-\delta^2 \frac{np}{2}}$$
$$\textbf{\textit{(Upper tail)}}\ \Pr[X \geq (1+\delta)np] \leq e^{-\delta^2 \frac{np}{2+\delta}}$$

引理 5.4。对于任意 $f(n) = \omega(\sqrt{n})$，若计数深度为 $d(n) \leq n/t - f(n)$，则类序列 $M_{\overline{n}}$（其中 $\overline{n} = (d(n) \ldots, d(n))$）支配$\equiv_{d(n)}$。

证明。我们将证明，对于每个满足 $f(n) = \omega(\sqrt{n})$ 且 $f(n) \leq n/t$ 的 $f(n)$，以极限概率 1，随机模型对每种类型的实现次数都超过 $n/t - f(n)$ 次。给定一种类型 $\pi$，我们令 $X_\pi$ 表示一个随机变量，用于计数 $\pi$ 实现该类型的次数。现在，$X_\pi := \sum_{1 \leq i \leq n} X_{\pi,i}$，其中 $X_{\pi,i}$ 是一个指示随机变量，表示元素 $i$ 实现类型 $\pi$这一事件。由于 $X_{\pi,i}$ 的成功概率为$t^{-1}$，我们有 $\mathrm{E}(X_\pi) = n/t$。

现在，Chernoff 界告诉我们，对于每个 $n$ 以及每个 $0 \leq \delta$，都有

$$\Pr\left[X_\pi \leq (1-\delta)\frac{n}{t}\right] \leq e^{-\delta^2 \frac{n}{2t}}.$$

注意到 $\delta$ 确实可以依赖于 $n$。令 $\delta(n) := \sqrt{2tg(n)/n}$，对于任意$g(n) = \omega(1)$，我们得到

$$e^{-\delta^2 \frac{n}{2t}} = e^{-g(n)} \to 0.$$

此外

$$\delta(n)\frac{n}{t} = \underbrace{\sqrt{2t} \cdot \frac{1}{t}}_{=:C} \cdot \sqrt{g(n)} \cdot \sqrt{n} = C\sqrt{ng(n)}.$$

因此，对于任意函数 $g(n) = \omega(1)$，以高概率，类型 $\pi$ 会被实现超过 $n/t - C\sqrt{ng(n)}$ 次。由于 $\pi$ 是任意的，根据并集界可知，以高概率，每种类型 $\pi$ 都会被实现超过 $n/t - C\sqrt{ng(n)}$ 次。现在，若 $f(n) = \omega(\sqrt{n})$，则令 $g(n) = (f(n)/C\sqrt{n})^2$，可得每种类型都会被实现超过 $n/t - f(n)$ 次。

□

我们证明了，当计数深度足够低时，该分布存在支配类。接下来，我们将证明，当计数深度足够高时，不存在支配类。为此，我们将使用 Robbins [14] 给出的以下斯特林近似不等式形式。

命题 5.5（斯特林近似 [14]）。对于所有 $n \in \mathbb{N}$、 $n > 0$，我们有

$$\sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n e^{\frac{1}{12n}}$$

引理 5.6. 对任意 $f(n) = o(\sqrt{n})$，如果计数深度为 $d(n) \geq n/t - f(n)$，那么 $\equiv_{d(n)}$ 在 $n \to \infty$ 时没有支配类。

证明. 设 $M(n)$ 是一个任意的类序列。我们将证明，存在某个 $n_0$ ，使得对于每个 $n \geq n_0$ ，模型属于类 $M(n)$ 的概率至多为一半。

设 $M_{\overline{n}}$ 是序列 $M(n)$ 中满足 $i, j \leq t$ 且 $n_i \neq n_j$ 的一个类。令 $\overline{n}'$ 由 $\overline{n}$ 得到，方法是在元组中交换数字 $n_i$ 和 $n_j$。由对称性可知 $|M_{\overline{n}'}| = |M_{\overline{n}}|$ ，因此模型属于 $M_{\overline{n}}$ 的概率至多为一半。

剩下要证明的是：当 $n$ 足够大时，序列 $M(n)$ 中元组只重复一个数字的类，其概率至多为一半。设 $d \geq n/t - f(n)$，其中 $f(n) = o(\sqrt{n})$。假设 $d < n/t$，并令 $M = M_{\overline{n}}$，其中 $\overline{n} = (d, \ldots, d)$。注意，这是唯一一个只重复一个数字的 $(n, d)$-可容许元组。我们证明，如果 $n$ 足够大，这个类 $M$ 的概率就小于一半。事实上，我们将证明一个更强的结论：具有元组 $(d(n) \ldots, d(n))$ 的这类序列的极限概率为 0。

上述类 $M$ 的大小由以下总和给出

$$|M| = \sum_{\substack{n_1+\cdots+n_t=n \\ n_i \geq d}} \binom{n}{n_1, \ldots, n_t}.$$

首先注意，当数字$n_1$、..... $n_t$ 相等时，多项式系数最大。利用这一点和斯特林近似，我们得到

$$\begin{aligned}
\binom{n}{n_1, \ldots, n_t} &\leq \binom{n}{\frac{n}{t}, \ldots, \frac{n}{t}} \\
&\leq \frac{\sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n}}}{(\sqrt{2\pi\frac{n}{t}}(\frac{n}{et})^{\frac{n}{t}} e^{\frac{1}{12\frac{n}{t}+1}})^t} \\
&= \frac{\sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n}}}{\sqrt{2\pi\frac{n}{t}}^t (\frac{n}{e})^n \frac{1}{t^n} e^{\frac{t}{12\frac{n}{t}+1}}} \\
&= \sqrt{\frac{t^t}{(2\pi)^{t-1}}} \cdot \frac{1}{\sqrt{n^{t-1}}} \cdot e^{g(n)} \cdot t^n \\
&\leq \sqrt{\frac{t^t}{(2\pi)^{t-1}}} \cdot \frac{1}{\sqrt{n^{t-1}}} \cdot 2^{n|\tau|}
\end{aligned}$$

上面 $e$ 的指数为

$$g(n) = \frac{1}{12n} - \frac{t}{12\frac{n}{t}+1} = \frac{1}{12n} - \frac{t^2}{12n+t}$$
$$= \frac{12n+t-12t^2n}{12n(12n+t)} = \frac{(1-t^2)12n+t}{12n(12n+t)}.$$

显然，对所有正的 $n$， $g(n) < 0$，因此 $e^{g(n)} < 1$，上述估计成立。由插板法，原来的和包含

$$\binom{tf(n)+t-1}{t-1} \leq (2t)^{t-1} \cdot f(n)^{t-1}$$

项，因此最终估计为

$$|M| \leq \sqrt{\frac{t^t}{(2\pi)^{t-1}}} \cdot (2t)^{t-1} \cdot \frac{f(n)^{t-1}}{\sqrt{n^{t-1}}} \cdot 2^{n|\tau|}$$

回想一下，总共有 $2^{n|\tau|}$ 个大小为 $n$ 的模型，且 $f(n) = o(\sqrt{n})$。我们得到以下极限：

$$\lim_{n\to\infty} \frac{\sqrt{\frac{t^t}{(2\pi)^{t-1}}} \cdot (2t)^{t-1} \cdot \frac{f(n)^{t-1}}{\sqrt{n^{t-1}}} \cdot 2^{n|\tau|}}{2^{n|\tau|}} = 0.$$

我们看到，类序列 $M_{\overline{n}}$ 满足 $\overline{n} = (d, \ldots, d)$，其极限概率为 0。因此，对于这样一个单独的类，只要 $n$ 足够大，其概率肯定至多为一半。

现在假设 $d \geq n/t$，并考虑类序列 $M_{\overline{n}}$，其中 $\overline{n} = (n/t, \ldots, n/t)$。这仍然是唯一一个只重复一个数字的 $(n, d)$-可容许元组。这些类的大小由 $\binom{n}{\frac{n}{t}, \ldots, \frac{n}{t}}$ 给出，并且很容易从上文看出，该序列的极限概率也为 0。因此，当 $n$ 足够大时，这样一个单独的类的概率小于一半。 □

直观来说，如果大小为 $n$ 的随机模型属于某个类（序列）的极限概率为零，那么这个类（序列）就是消失的。为正式定义这一概念，固定一个序列 $\equiv_{d(n)}$。我们称 $\equiv_{d(n)}$ 中的所有类在 $n \to \infty$ 时都是消失的，如果对于所有满足 $\equiv_{d(n)}$ 的类序列 $M(n)$，都有 $p_n \to 0$，当 $n \to \infty$ 时。现在我们证明，如果计数深度足够高，那么分布序列中的所有类都是消失的。

引理 5.7. 对于任意满足 $f(n) = \omega(\sqrt{n})$ 的 $f(n)$，如果计数深度为 $d(n) \geq n/t + f(n)$，那么 $\equiv_{d(n)}$ 中的所有类在 $n \to \infty$ 时都是消失的。

证明。固定某个函数 $f(n)$，使得 $f(n) = \omega(\sqrt{n})$。我们首先证明，以极限概率为零，随机模型实现每种类型的次数都少于 $n/t + f(n)$。不失一般性，我们可以假设

$f(n) \le n - n/t$。给定一种类型 $\pi$，我们令 $X_\pi$ 表示与引理 5.4 的证明中相同的随机变量。Chernoff 界再次告诉我们，对于每个$n$ 以及每个 $0 < \delta \le 1$，都有

$$\Pr\left[X_\pi \ge (1+\delta)\frac{n}{t}\right] \le e^{-\delta^2 \frac{n}{t(2+\delta)}} \le e^{-\delta^2 \frac{n}{3t}}$$

请注意，$\delta$ 可以依赖于 $n$。令 $\delta(n) := \sqrt{3tg(n)/n}$，对于任意满足 $g(n) = \omega(1)$且 $g(n) \le n/(3t)$的 $g(n)$（以确保 $\delta(n) \le 1$），我们得到

$$e^{-\delta^2 \frac{n}{3t}} = e^{-g(n)} \to 0.$$

此外

$$\delta(n)\frac{n}{t} = \underbrace{\sqrt{3t}\cdot\frac{1}{t}}_{=:C}\cdot\sqrt{g(n)}\cdot\sqrt{n} = C\sqrt{ng(n)}.$$

因此，对于任意满足 $g(n) = \omega(1)$且 $g(n) \le n/(3t)$的函数 $g(n)$，以高概率，类型 $\pi$ 实现的次数少于 $\frac{n}{t} - C\sqrt{ng(n)}$次。由于 $\pi$ 是任意的，根据并集界可知，以高概率，每个类型 $\pi$ 实现的次数都少于 $\frac{n}{t} - C\sqrt{ng(n)}$次。~~令 $g(n) = (f(n)/C\sqrt{n})^2$~~，可得每个类型实现的次数都少于 $n/t - f(n)$次。

现在考虑任意类序列 $M(n)$。我们要证明，对于每个 $\varepsilon > 0$，只要 $n$ 足够大，就有 $p_n < \varepsilon$。考虑一个类 $M(n)$，并令 $\overline{n}$为相应的（$n$，$d(n)$）-可容许元组，即 $M(n)$是类 $M_{\overline{n}}$。如果存在 $i \in I$ 使得 $n_i = d(n) = n/t + f(n)$，那么只要 $n$ 足够大，由前一个结果可知$p_{\equiv_{d(n)}}(M_{\overline{n}}) < \varepsilon$。接下来假设对每个$i \in I$，都有 $n_i < n/t + f(n)$。在这种情况下，类 $M_{\overline{n}}$ 是一个同构类，这意味着随机模型属于 $M_{\overline{n}}$ 的概率就是

$$\binom{n}{n_1, \ldots, n_t} / 2^{n|\tau|},$$

根据引理 5.6 的证明中的计算，该值至多为一个常数乘以$1/\sqrt{n^{t-1}}$。只要 $n$ 足够大，后一个量必然小于 $\varepsilon$。因此，只要 $n$足够大，$p_{\equiv_{d(n)}}(M_{\overline{n}}) < \varepsilon$。

□

我们将上述结果汇总为以下定理：

定理 5.8. 以下关于计数深度 $d(n)$ 的陈述成立，其中$n \to \infty$。

- *If $d(n) \le n/t - f(n)$ where $f(n) = \omega(\sqrt{n})$, then $\equiv_{d(n)}$ has a dominating class.*

- 如果 $d(n) \geq n/t - f(n)$，其中 $f(n) = o(\sqrt{n})$，那么 $\equiv_{d(n)}$ 没有支配类。

- 如果 $d(n) \geq n/t + f(n)$，其中 $f(n) = \omega(\sqrt{n})$，那么 $\equiv_{d(n)}$ 中的每个类都是消失的。

我们指出上述结果的一个推论。当计数深度过低时，就 GMLU$_d$ 可定义性而言，几乎所有模型都属于同一个支配类。相反，如果计数深度足够高，GMLU$_d$ 就能将模型划分为会消失的类。这些观察直接给出以下结果：

推论 5.9。设 $f(n) = \omega(\sqrt{n})$。如果 $d(n) \leq n/t - f(n)$，那么以极限概率 1，两个大小为 $n$的随机模型无法在 GMLU$^{d(n)}$中被区分。如果 $d(n) \geq n/t + f(n)$，那么以极限概率 1，两个大小为 $n$ 的随机模型可以在 GMLU$^{d(n)}$中被区分。

如果一个类包含超过所有模型一半的模型，我们就称其为多数类。下面的定理是定理 5.8 的定量版本。

定理 5.10。设 $n \in \mathbb{Z}_+$。

- 如果 $d \leq n/t - c_1\sqrt{n}$，其中

$$c_1 := \frac{1}{t}\sqrt{2t\ln(2t)},$$

那么，对于大小为 $n$ 的模型，其分布 $\equiv_d$ 有一个多数类。

- 如果 $d \geq n/t - c_2\sqrt{n}$，其中

$$c_2 := \sqrt{\frac{\pi}{2t^3(4t)^{1/(t-1)}}} < c_1$$

那么，对于大小为 $n$ 的模型，其分布 $\equiv_d$ 没有多数类。

证明。首先假设 $d \leq n/t - c_1\sqrt{n}$。令 $\delta = (tc_1)/\sqrt{n}$，此时$\delta \cdot \frac{n}{t} = c_1\sqrt{n}$ 且 $\delta^2 \cdot n/(2t) = (tc_1)^2/(2t)$。类似于引理 5.4 的证明，应用 Chernoff 界和并集界可得，以严格大于 $(1 - te^{-(tc_1)^2/(2t)})$ 的概率，每种类型至少出现 $n/2 - c_1\sqrt{n}$ 次。简单计算可知，后一个概率等于 $1/2$ （因此选择 $c_1$）。

于是考虑 $d \geq n/t - c_2\sqrt{n}$ 的情形。令 $M = M_{\overline{n}}$，其中 $\overline{n} = (d, \ldots, d)$。利用引理 5.6 的证明中给出的估计以及 $f(n) = c_2\sqrt{n}$，我们得到

$$\frac{|M|}{2^{n|\tau|}} \leq \sqrt{\frac{t^t}{(2\pi)^{t-1}}} \cdot (2t)^{t-1} \cdot \frac{(c_2\sqrt{n})^{t-1}}{\sqrt{n^{t-1}}} = 1/2$$

因此， $M$ 不是多数类。与引理 5.6 的证明中相同的推理表明，不存在其他多数类。 □

当然，我们可以得到推论 5.9 的相应定量版本。

推论 5.11。设 $n \in \mathbb{Z}_{+}$。若 $d \leq n/t - c_1\sqrt{n}$，则两个大小为 $n$ 的随机模型能够在 GMLU$_d$ 中被区分的概率小于 $1/4$。若 $d \geq n/t - c_2\sqrt{n}$，则两个大小为 $n$ 的随机模型能够在 GMLU$_d$ 中被区分的概率至少为 $1/2$。

# 6 结论

我们建立了模型类大小与描述复杂度之间有趣的单调连接，并得到了关于熵的相关结果。此外，我们刻画了当定义域大小 $n$ 和表达能力（以计数深度 $d$ 的形式表示）发生变化时，模型类大小的相变。这些结果阐明了类大小与公式长度之间的相互作用。虽然本文聚焦于 GMLU$_d$，但这些结果旨在概述相关现象。因此，一个显而易见的未来研究方向是探究如何将这些结果推广到一阶逻辑的框架中。在那里，可以通过量词深度和变量数量，获得类似于改变计数深度 $d$ 的自然参数化方式。

超越一阶逻辑之后，一个有趣的研究方向是考察 [9] 中引入的、表达能力达到图灵完备的逻辑 CL，即计算逻辑。在这一框架下研究描述复杂度，将使其与柯尔莫哥洛夫复杂度的联系更加紧密。

致谢。Antti Kuusisto 和 Miikka Vilander 的工作得到芬兰科学院项目“通过逻辑解释人工智能”（XAILOG）的资助，资助编号为 345612（Kuusisto）。Antti Kuusisto 的工作还得到芬兰科学院项目“计算逻辑理论”的资助，资助编号为 324435、328987、352419、352420、352419、353027。

# 参考文献

[1] Micah Adler 和 Neil Immerman。公式大小的 n! 下界。《ACM计算逻辑汇刊》，4(3)：296–314，2003。

[2] Pablo Barcel´o、Mika¨el Monet、Jorge P´erez 和 BernardoSubercaseaux。《从计算复杂性的视角看模型可解释性》。载于 Hugo Larochelle、Marc’Aurelio Ranzato、Raia Hadsell、Maria-Florina Balcan 和 Hsuan-Tien Lin 编，《神经信息处理系统进展》

Processing Systems 33：神经信息处理系统年度会议 2020, NeurIPS 2020，2020年12月6日至12日，线上举行，2020年。[3] Ronald L. Graham、Donald E. Knuth 和 Oren Patashnik。《具体数学——计算机科学基础》。Addison-Wesley，1989年。[4] Peter Gr¨unwald 和 Paul M. B. Vit´anyi。香农信息与柯尔莫哥洛夫复杂度。CoRR, cs.IT/0410002，2004年。[5] Lauri Hella 和 Miikka Vilander。模态逻辑与 $\mu$-演算的公式大小博弈。《逻辑与计算杂志》，29(8):1311–1344, 2019年。[6] Reijo Jaakkola、Tomi Janhunen、Antti Kuusisto、Masood Feyzbakhsh Rankooh 和 Miikka Vilander。通过短公式实现可解释性：带实现的命题逻辑案例。载于《（HYDRA 2022）与组合爆炸问题求解算法实验评估 RCRA 研讨会联合论文集》，第3281卷，CEUR研讨会论文集，第64–77页，2022年。[7] Reijo Jaakkola、Antti Kuusisto 和 Miikka Vilander。描述复杂度与熵的关联。arXiv:2209.12564，2022年。[8] Reijo Jaakkola、Antti Kuusisto 和 Miikka Vilander。描述复杂度与熵的关联。载于 Martin Grohe 和 Johann A. Makowsky 编，《计算机科学理论方面研讨会论文集》，即将出版，2023年。[9] Antti Kuusisto。一阶逻辑的一些图灵完备扩展。载于 Adriano Peron 和 Carla Piazza 编，《第五届博弈、自动机、逻辑与形式验证国际研讨会论文集》，GandALF 2014，意大利维罗纳，2014年9月10日至12日，第161卷，EPTCS，第4–17页，2014年。[10] Sik K. Leung-Yan-Cheong 和 Thomas M. Cover。香农熵与柯尔莫哥洛夫复杂度之间的一些等价关系。《IEEE信息论汇刊》，24(3):331–338，1978年。[11] Ming Li 和 Paul M. B. Vit´anyi。《柯尔莫哥洛夫复杂度及其应用导论》，第4版。计算机科学丛书。施普林格，2019年。[12] Michael Mitzenmacher 和 Eli Upfal。《概率与计算：随机化算法与概率分析》。剑桥大学出版社，2005年。

[13] Alexander A. Razborov。《矩阵方法在计算复杂性下界理论中的应用》。Comb.，10(1):81–93，1990。[14] Herbert Robbins。《关于斯特林公式的一点说明》。The American Mathematical Monthly，62(1):26–29，1955。[15]Andreia Teixeira、Armando Matos、Andre Souto 和 Luis Filipe Coelho Antunes。《熵度量与柯尔莫哥洛夫复杂度》。Entropy，13(3):595–611，2011。