arXiv:2303.17578v1 [cs.LG] 30 Mar 2023

在线可 学习与部分概念类的消歧

S

Tsun-Ming Cheung* Hamed Hatami † Pooya Hatami ‡ Kaave Hosseini §

摘要

在近期的一篇文章中，Alon、Hanneke、Holzman 和 Moran（FOCS ’21）提出了一个统一框架，用于研究部分概念类的可学习性。他们研究的一个核心问题是：部分概念类的可学习性是否总能从其扩展为某个全概念类后的可学习性中继承而来。

他们证明，PAC 学习并非如此，但将这一问题留给了更强的在线可学习性概念。

我们通过构造一类可在线学习的部分概念来解决这一问题，但将其扩展为全概念类后，该全概念类无法在线学习（甚至无法进行 PAC 学习）。

1 引言

在许多实际的学习问题中，之所以能够处理学习任务，是因为我们只需预测满足特定性质的数据点的标签。在二元分类问题的设定下，我们通常不必学习一个全概念 $h: \mathcal{X} \to \{0,1\}$，而是满足于学习它的部分版本 $h: \mathcal{X} \to \{0,1,\star\}$，其中 $h(x) = \star$ 表示将 0 和 1 都视为可接受的预测。这种允许未指定预测的松弛方式，使更广泛的学习任务变得可处理。

例如，可以通过观察某人以往的投票模式，来预测其是否赞成或反对各种政治立场。此人可能对某些政治态度没有明确的看法，因此根据其以往经历，可能无法预测其在这些问题上的投票。然而，如果在缺乏坚定信念的情况下允许“赞成”和“反对”都作为可接受的预测，那么这一学习任务或许就能够实现。

对这一现象进行了深入研究的一个例子是大间隔半空间学习。在这个问题中，定义域是任意欧几里得空间中某个有界区域内的点集，概念则是半空间：根据每个点是否属于该半空间，将其映射为 1 或 0。众所周知，当底层欧几里得空间的维数很高时，学习一个半空间需要大量样本。然而，在大间隔设定下，我们只需在点到定义超平面的距离至少达到某个间隔时，正确预测该点的标签。Rosenblatt 提出的经典感知机算法等标准学习算法 [Ros58], 表明，这种对学习

*麦吉尔大学计算机科学学院。电子邮箱：tsun.ming.cheung@mail.mcgill.ca。

†麦吉尔大学计算机科学学院。电子邮箱：hatami@cs.mcgill.ca。获加拿大自然科学与工程研究委员会资助。

‡俄亥俄州立大学计算机科学与工程系。pooyahat@gmail.com。获美国国家科学基金会项目 CCF-1947546§罗切斯特大学计算机科学系。电子邮箱：kaave.hosseini@rochester.edu

要求的放宽使得即使在高维欧几里得空间中，这个问题也变得易于处理。受此类例子的启发，Alon、Hanneke、Holzman 和 Moran [AHHM22] 开创了对部分概念类可学习性的系统研究 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$。他们聚焦于概率近似正确（PAC）学习和在线学习这两种框架。部分概念类的 PAC 可学习性定义见 [AHHM22]。我们在定义 1.4 中定义在线可学习性。

PAC 学习是一个优雅的理论框架，其特征由 Vapnik–Chervonenkis（VC）维数这一组合参数刻画。PAC 学习基本定理指出，全二元概念类可进行 PAC 学习，当且仅当其 VC 维数有限。类似地，全概念类的在线可学习性由一个称为 Littlestone 维数（LD）的组合参数刻画。我们分别在定义 2.2 和定义 2.3 中正式定义 VC 维数和 Littlestone 维数。Alon、Hanneke、Holzman 和 Moran [AHHM22] 证明，这些关于 PAC 学习和在线可学习性的刻画可推广到部分概念类的情形。

定理 1.1（[AHHM22，定理 1 和 15]）。设 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$为一个部分概念类。

- $\mathbb{H}$ 可进行 PAC 学习，当且仅当 VC($\mathbb{H}$) $< \infty$。
- $\mathbb{H}$ 当且仅当 LD($\mathbb{H}$) $< \infty$时，具有在线可学习性。

根据 VC 维和 Littlestone 维数的定义，对于每个部分概念类$\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$，都有 VC($\mathbb{H}$) $\leq$ LD($\mathbb{H}$)。特别地，在线可学习性总是意味着 PAC 可学习性。

在 [AHHM22] 中研究的核心问题之一是：部分概念类的可学习性是否总能继承自某个全概念类的可学习性。为了明确这一问题，我们需要定义部分概念类的消歧这一概念。正式定义将在第 2.2 节给出；我们可以先将部分类的强消歧理解为：为该类中的每个部分概念，将每个 $\star$ 指定为 1 或 0。当 $\mathcal{X}$ 为无限集时，考虑我们将在定义 2.5 中给出的较弱的消歧概念更为自然。当 $\mathcal{X}$ 为有限集时，消歧和强消歧这两个概念相同。

考虑在 PAC 学习或在线学习中学习部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$的问题。如果部分概念类 $\mathbb{H}$ 具有一个可进行 PAC 学习的消歧 $\overline{\mathbb{H}} \subseteq \{0,1\}^{\mathcal{X}}$，那么 $\mathbb{H}$ 是可进行 PAC 学习的。这是由 VC($\mathbb{H}$) $\leq$ VC($\overline{\mathbb{H}}$)得到的，或者只需在 $\mathbb{H}$上运行 $\overline{\mathbb{H}}$ 的 PAC 学习算法即可。类似地，如果 $\mathbb{H}$ 的一个消歧 $\overline{\mathbb{H}}$ 是可进行在线学习的，那么$\mathbb{H}$ 是可进行在线学习的。

每个部分概念类的可学习性，是否都继承自某个全概念类的消歧的可学习性?

问题 1.2（非正式 [AHHM22]）。每个可学习的部分类是否都有一个可学习的消歧?

借助定理 1.1, [AHHM22] 的 VC 维数刻画，已证明对于 PAC 学习，问题 1.2 的答案是否定的。

定理 1.3（[AHHM22，定理 11]）。对于每个 $n \in \mathbb{N}$，都存在一个部分概念类$\mathbb{H}_n \subseteq \{0,1,\star\}^{[n]}$，其 VC 维数为 VC($\mathbb{H}_n$) $= 1$，使得对 $\mathbb{H}_n$的任意消歧 $\overline{\mathbb{H}}$，VC($\overline{\mathbb{H}}$) $\geq (\log\ n)^{1-o(1)}$。此外，对于 $\mathcal{X} = \mathbb{N}$，存在 $\mathbb{H}_\infty \subseteq \{0,1,\star\}^{\mathcal{X}}$，其 VC 维数为 VC($\mathbb{H}_\infty$) $= 1$，且对于 $\mathbb{H}_\infty$的每种消歧 $\overline{\mathbb{H}}$，VC($\overline{\mathbb{H}}$) $= \infty$。

尽管定理 1.3 针对 PAC 学习的情形，对问题 1.2 给出了强有力的否定回答，但在线学习情形下的问题仍未解决。粗略来说，该问题将 $\mathbb{H}$ 上的有界 VC 维假设加强为有界 Littlestone 维数（LD）；后者关系到 $\mathbb{H}$的在线可学习性。

[AHHM22] 中的作者还提出了第二个开放问题：用多项式增长的假设取代 VC 维数有界的假设。该假设比 Littlestone 维数有界更弱，但比 VC 维数有界更强。

正如下文所述，我们的主要结果解决了这两个开放问题。

在线可学习性。在线学习按连续的轮次进行：在第 t 轮，学习者会收到一个实例 $x_t \in \mathcal{X}$，并需要预测其标签。预测标签后，正确标签 $y_t \in \{0,1\}$ 会揭示给学习者。请注意，即使对于部分概念类，我们也要求正确标签为 0 或 1。学习者的目标是在这一过程中尽可能少地犯预测错误。我们假设真实标签始终是可实现的，即存在一个部分概念 $h \in \mathbb{H}$，满足 $h(x_i) = y_i$，对于所有 $i = 1, \ldots, t$ 均成立。

定义 1.4（在线可学习性）。如果存在一个错误界 $m := m(\mathbb{H}) \in \mathbb{N}$，使得对于每个 $T \in \mathbb{N}$，都存在一个学习算法，对每个可实现序列 $(x_i, y_i)_{i=1,\ldots,T}$，至多犯 m 次错误，则称部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$ 在线可学习。

全概念类的在线可学习性等价于 Littlestone 维数有界。在定理 1.1 中，Alon、Hanneke、Holzman 和 Moran [AHHM22] 证明，同样的等价关系也适用于部分类。他们提出了问题 1.2 的如下表述。

如果一个部分类是在线可学习的，那么它是否存在一个在线可学习的消歧?

更确切地说，他们提出了以下问题：

问题 1.5（[AHHM22]）。设 $\mathbb{H}$ 是一个满足 LD($\mathbb{H}$) $< \infty$的部分类。是否存在 $\mathbb{H}$ 的一个消歧 $\overline{\mathbb{H}}$，使得 LD($\overline{\mathbb{H}}$) $< \infty$? 是否存在一个满足 VC($\overline{\mathbb{H}}$) $< \infty$ 的消歧?

我们对问题 1.5 给出否定答案：

定理 1.6（主定理）。对于每个 $n \in \mathbb{N}$，都存在一个部分概念类 $\mathbb{H}_n \subseteq \{0,1,\star\}^{[n]}$，满足 LD($\mathbb{H}_n$) ≤ 2，且其每个消歧 $\overline{\mathbb{H}}$ 都满足 LD($\overline{\mathbb{H}}$) $\geq$ VC($\overline{\mathbb{H}}$) $= \Omega(\log \log n)$。因此，对于 $\mathcal{X} = \mathbb{N}$，存在 $\mathbb{H}_\infty \subseteq \{0,1,\star\}^{\mathcal{X}}$，满足 LD($\mathbb{H}_\infty$) ≤ 2，且对于其每个消歧 $\overline{\mathbb{H}}$，都有 LD($\overline{\mathbb{H}}$) $\geq$ VC($\overline{\mathbb{H}}$) $= \infty$。

多项式增长。证明全概念类 $\mathbb{H} \subseteq \{0,1\}^n$ 的 VC 维数具有超常数下界的一种通用策略，是证明该类的规模超多项式增长。定理 1.3 和定理 1.6 采用的就是这种方法。对于 VC 维数为 d 的全概念类 $\mathbb{H} \subseteq \{0,1\}^n$，有 $2^d \leq |\mathbb{H}| \leq O(n^d)$：根据 VC 维数的定义，下界显而易见；上界则是著名的 Sauer-Shelah-Perles（SSP）引理的推论。

定理 1。 7（Sauer-Shelah-Perles 引理 [Sau72]）。设 $\mathbb{H} \subseteq \{0,1\}^n$ 且 VC($\mathbb{H}$) $= d$。那么

$$|\mathbb{H}| \leq \binom{n}{\leq d} := \sum_{i=0}^{d} \binom{n}{i} = O(n^d).$$

对于部分概念类，SSP 引理的直接类比并不成立： [AHHM22] 证明了存在 $\mathbb{H} \subseteq \{0,1,\star\}^{[n]}$，其 VC($\mathbb{H}$) $= 1$，使得每种消歧 $\overline{\mathbb{H}}$的大小都为$|\overline{\mathbb{H}}| \geq n^{\Omega(\log n)}$。结合全类的 SSP 引理，这一结果立即推出定理 1.3。

值得注意的是，在 Littlestone 维数有界这一更强假设下，原 SSP 引理中的多项式增长性质仍然成立。

定理 1.8（[AHHM22]）。每个部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{[n]}$，只要其 Littlestone 维数 $\mathbb{H} \leq d$，就存在一个消歧 $\overline{\mathbb{H}}$，使得 $|\overline{\mathbb{H}}| \leq O(n^d)$。

我们称部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$具有参数为 $d \in \mathbb{N}$ 的多项式增长性质，如果对于每个有限的 $\mathcal{X}' \subseteq \mathcal{X}$，都存在 $\mathbb{H}|_{\mathcal{X}'}$ 的一个消歧 $\overline{\mathbb{H}|_{\mathcal{X}'}}$，其大小至多为 $O(|\mathcal{X}'|^d)$。注意，根据定理 1.8，每个 Littlestone 维数为 d 的部分概念类都具有参数为 d 的多项式增长性质。

Alon、Hanneke、Holzman 和 Moran 提出了以下问题:

问题 1.9（[AHHM22]）。设 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$ 是一个具有多项式增长性质的部分概念类。是否存在 $\mathbb{H}$ 的一个消歧 $\overline{\mathbb{H}}$，使得 VC($\overline{\mathbb{H}}$) $< \infty$?

注意，对 $\mathbb{H}$ 的消歧或其限制直接应用 SSP 引理，无法（以否定方式）解决问题 1.9。不过，定理 1.6 与定理 1.8 结合起来，同样可以否定问题 1.9。

定理 1.10. 对于每个 $n \in \mathbb{N}$，都存在一个具有参数 2 的多项式增长的 $\mathbb{H} \subseteq \{0,1,\star\}^{[n]}$，使得 $\mathbb{H}$ 的每种消歧 $\overline{\mathbb{H}}$ 都满足 VC($\overline{\mathbb{H}}$) $= \Omega(\log \log n)$。

因此，对于 $\mathcal{X} = \mathbb{N}$，存在一个具有参数 2 的多项式增长的 $\mathbb{H}_\infty \subseteq \{0,1,\star\}^{\mathcal{X}}$，使得 $\mathbb{H}_\infty$ 的每种消歧 $\overline{\mathbb{H}_\infty}$ 都满足 VC($\overline{\mathbb{H}_\infty}$) $= \infty$。

Alon-Saks-Seymour 问题。 [AHHM22] 中定理 1.3 的证明依赖于 G¨o¨os [G¨o¨o15] 的突破性结果及其后续改进 [BBDG+22] ，这些成果使 Alon、Saks 和 Seymour 提出的“二分团划分数与色数”问题获得了几乎最优的超多项式界。图 G 的二分团划分数记为 bp(G)，它是能够划分 G 的边集的完全二部图（二分团）的最小数量。Alon、Saks 和 Seymour 猜想，二分团划分数为 k 的图的色数至多为 $k+1$。Huang 和 Sudakov 在 [HS12] 中通过证明这两个参数之间存在超线性差距，否定了 Alon-Saks-Seymour 猜想。后来，G¨o¨os [G¨o¨o15] 取得突破，证明了二者之间存在超多项式分离。

我们的主要结果定理 1.6 同样建立在前述图构造之上。然而，与以往工作不同，我们的定理要求顶点数量有一个合理的上界。由于这些构造源自一系列复杂的归约，涉及查询复杂度、通信复杂度和图论 [BLT14, G¨o¨o15, GLM+16, BBDG+22]，因此有必要仔细审查这些构造，以确保满足所需参数。我们在第 3.3 节中给出一系列重新编排且部分简化的构造，从而证明以下定理。

定理 1.11（对 Alon-Saks-Seymour 猜想的小规模反驳）。存在一个图G，其有 $2^{\Theta(k^4 \log^3 k)}$ 个顶点，且存在大小为 $2^{O(k \log^4 k)}$ 的二分团划分，但其色数至少为 $2^{\Omega(k^2)}$。

定理 1.11 本质上归功于 [BBDG+22]。我们对该定理的贡献是，为 G 的大小给出了一个明确且经过优化的界。

标准最优算法。定理 1.6 给出了一个部分类的例子，其 Littlestone 维数为 ≤ 2，且其任意一种消歧的 VC 维数均为 Ω(log log n)。能否改进 Ω(log log n) 这一 下界尚不明确。特别是，一个值得探讨的问题是：Littlestone 维数至多为 2 的部类，其任意一种消歧的 VC 维数是否都为O(log log n)。一种自然的思路是利用标准最优算法（SOA）来获得这样的上界。

SOA 是 Littlestone [Lit88] 提出的一种在线学习算法，可以学习 Littlestone 维数有界的概念类。Alon、Hanneke、Holzman 和 Moran 在定理 1.8 的证明中指出，将 SOA 应用于 Littlestone 维数为 d 的部分概念类 $\mathbb{H}$，可得到大小为 $|\overline{\mathbb{H}}| \leq O(n^d)$ 的消歧，因此其 VC 维数为 $O(d\log n)$。这表明，定理 1.6 关于消歧的 VC 维数的下界无法改进到超过 O(log n)。因此，自然会问：能否为基于 SOA 的消歧得到更优的 VC 维数上界?

我们通过构造一族 Littlestone 维数为 d 的部分概念类 $\mathbb{H}$，对这个问题给出了否定答案；对于这族概念类，SOA 算法得到的消歧具有 Ω($d$log(n/d)) 的 VC 维数。

定理 **1 12** $d \leq n$ $\mathbb{H}_{n\,d} \subseteq$ . . 对 $\{0,1,\star\}^{[n]}$ $d \leq$ $\mathbb{H}_{n\,d} \leq d+1$ $\mathbb{H}_{n\,d}$ 于每个自然数，都存在一个部分概念类，使得其 LD(·)，且 · 的 SOA 消歧具有 Ω($d$log(n/d)) 的 VC 维数。

2 预备知识与背景

对于正整数 k，我们记 $[k] := \{1,\ldots,k\}$。我们约定，$\{0,1\}^0$ 或$\{0,1,\star\}^0$ 仅包含空字符串，并将其记为 ()。

我们采用标准的计算机科学渐近记号，例如大 O 记号，并使用渐近波浪号记号来隐藏多对数因子。

2.1 VC 维数与 Littlestone 维数

设 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$为一个部分概念类。当定义域 $\mathcal{X}$有限时，我们有时将$\mathbb{H}$ 视为一个部分矩阵 $\mathrm{M}_{\mathcal{X}\times\mathbb{H}}$，其中每一行对应一个点 $x \in \mathcal{X}$，每一列对应一个概念 $h \in \mathbb{H}$，且矩阵项定义为 M(x,h) $= h(x)$。

接下来，我们定义部分类的 VC 维数和 Littlestone 维数，它们推广了全类中这些概念的定义。如 [AHHM22], 所示，部分类的 VC 维数和 Littlestone 维数分别刻画 PAC 可学习性和在线可学习性。

定义 2.1（打散集）。如果对于每个模式 $y \in \{0,1\}^n$，都存在 $h \in \mathbb{H}$，使得对于所有 $i \in [n]$，$h(x_i) = y_i$，则有限点集 $C = \{x_1,\ldots,x_n\} \subseteq \mathcal{X}$ 被部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$ 打散。

定义 2.2（VC 维数）。部分类 $\mathbb{H}$ 的 VC 维数记为 VC($\mathbb{H}$)，它是满足以下条件的最大 d：存在一个大小为 d 的 $\mathcal{X}$ 子集，该子集被 $\mathbb{H}$ 打散。如果不存在这样的最大 d，则定义 VC($\mathbb{H}$) $= \infty$。

将 $\mathbb{H}$ 视为矩阵，其 VC 维数是满足以下条件的最大 d：对应的部分矩阵 $\mathrm{M}_{\mathcal{X}\times\mathbb{H}}$ 包含一个维度为 $d \times 2^d$的零一子矩阵，其中各列枚举所有 d位零一模式。

Littlestone 维数是通过决策树的打散来定义的，而不是通过集合。考虑一棵高度为 d 的完整二叉决策树，其中每个非叶 v 都标记有一个元素$x_v \in \mathcal{X}$。我们用字符串 $v \in \bigcup_{k=0}^{d}\{0,1\}^k$ 来标识树中的每个节点，该字符串对应从根节点到该节点的路径。也就是说，根节点是空字符串，它的两个子节点分别对应 $\{0,1\}$中的两个元素；更一般地，节点 $\vec{v} \in \{0,1\}^k$ 的两个子节点是 $\{0,1\}^{k+1}$ 中的字符串 $\vec{v}0$ 和$\vec{v}1$ 。

我们称这样的树被部分概念类 $\mathbb{H}$ 打散，当且仅当对于每个叶节点 $y \in \{0,1\}^d$，都存在 $h \in \mathbb{H}$，使得对于每个 $i \in [d]$，$h(x_{y[<i]}) = y_i$，其中 $y[<i]$ 是 y 的前（$i-1$）位。换言之，将 h 输入决策树后，会到达叶节点 y。

定义 2.3（Littlestone 维数）。部分概念类 $\mathbb{H}$的 Littlestone 维数记作 LD($\mathbb{H}$)，定义为满足以下条件的最大 d：存在一棵高度为 d 的满二叉决策树，其节点标记为 $\mathcal{X}$，且被 $\mathbb{H}$打散。如果不存在这样的最大 d，则定义 LD($\mathbb{H}$) $= \infty$。

概念类 $\mathbb{H}$ 的对偶是将点与概念的角色互换后得到的概念类。具体而言，$\mathbb{H} \in \{0,1,\star\}^{\mathcal{X}}$的对偶类记作 $\mathbb{H}^\top$，它由函数 f_x 组成：对于每个 $x \in \mathcal{X}$，都有 $\mathbb{H} \to \{0,1,\star\}$，其定义为对每个 $h \in \mathbb{H}$，$f_x(h) = h(x)$。当 $\mathcal{X}$有限时，取对偶相当于转置概念类的矩阵。对偶类的 VC 维数与原类的 VC 维数满足以下不等式

$$\mathrm{VC}(\mathbb{H}^\top) \le 2^{\mathrm{VC}(\mathbb{H})+1} - 1$$

（参见 [Mat02]），这给出了原类 VC 维数的一个下界。

2.2 消歧

我们首先正式定义强消歧和消歧。如前所述，当定义域 $\mathcal{X}$ 为有限集时，这两个概念是一致的。

定义 2.4（强消歧）。部分概念类 $\mathbb{H} \subseteq\{0,1,\star\}^{\mathcal{X}}$的一个强消歧是一个全概念类 $\overline{\mathbb{H}} \subseteq \{0,1\}^{\mathcal{X}}$，使得对于每个 $h \in \mathbb{H}$，都存在一个 $\bar{h} \in \overline{\mathbb{H}}$，它在点 $h^{-1}(\{0,1\})$ 上与 h 一致。

定义 2.5（消歧）。部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$的一个消歧是一个全概念类 $\overline{\mathbb{H}} \subseteq \{0,1\}^{\mathcal{X}}$，使得对于每个 $h \in \mathbb{H}$ 以及每个有限的 $S \subseteq h^{-1}(\{0,1\})$，都存在 $h \in \overline{\mathbb{H}}$，它在 S 上与 h 一致。

学习算法通常可以通过将算法的预测赋给未指定的值，来对部分概念类进行消歧。与我们的工作相关的是 Littlestone 的标准最优算法所进行的消歧。[AHHM22] 中指出，该算法能够对 Littlestone 维数有界的部分类进行“高效”消歧。接下来我们将介绍这种消歧。

考虑一个定义域可数的部分概念类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$，其定义域为 $\mathcal{X}$，并且 x_1、x_2、..... 是 $\mathcal{X}$ 的一个排序。给定 $\vec{b} \in \{0,1,\star\}^k$，令 $\mathbb{H}|_{\vec{b}}$ 为概念 h 的集合，其中对于每个 $i \in [k]$，$h(x_i) = b_i$。为方便起见，我们将 $\mathbb{H}|_{()} = \mathbb{H}$ 视为同一对象。对于该算法，我们采用约定 LD($\emptyset$) $= -1$。

SOA 通过迭代得到消歧，并为 $\mathbb{H}$ 中的每个 $\star$ 赋予一个 $0/1$ 值：对于每个 $k \in \mathbb{N}$，考虑每个 $\vec{b} \in \{0,1\}^{k-1}$ 对应的 $\mathbb{H}|_{\vec{b}}$。选取使 LD($\mathbb{H}|_{\vec{b}c}$) 最大的 $c \in \{0,1\}$；若出现并列，则优先选择 $c = 0$，并对每个 $h \in \mathbb{H}|_{\vec{b}\star}$ 将 c 赋给 $h(x_k) = \star$。

我们用记号 $\overline{\mathbb{H}}^{\mathrm{SOA}}$ 表示部分概念类 $\mathbb{H}$ 的 SOA 消歧。如前所述，对于 Littlestone 维数为 d 的部分类，定理 1.8 给出了关于 $\left|\overline{\mathbb{H}}^{\mathrm{SOA}}\right|$ 的上界 $\binom{n}{\leq d} = O(n^d)$。该定理由 SOA 在线学习的错误界推出，这依赖于一项关键性质：只要 $\mathbb{H}|_{\vec{b}} \neq \emptyset$，就至少有一种 $c \in \{0,1\}$ 的选择满足 LD($\mathbb{H}|_{\vec{bc}}$) $\leq$ LD($\mathbb{H}|_{\vec{b}}$) -1。

3 证明

本节将给出定理1.6、1.10、1.11和1.12的证明。

3.1 定理1.6和1.10的证明

如前所述，定理1.10是定理1.6和定理1.8的直接推论。我们将重点证明定理1.6。

设 $G = (V, E)$是定理1.11给出的图，其包含 $|V| = n = 2^{\Theta(k^4 \text{对数}^3\, k)}$ 个顶点，并具有大小为 $m = 2^{O(k \log^4 k)}$的二分团划分。我们将用 G 构造一个部分概念类$\mathbb{G} \subseteq \{0,1,\star\}^V$。这个构造恰好是 [AHHM22] 在其定理1.6的证明中所用部分概念类的对偶。

令 $\{B_1, \ldots, B_m\}$ 为 G 的边的大小为 m 的二分团划分。我们为每个二分团固定一个定向 $B_i = L_i \times R_i$。定义 $\mathbb{G} \subseteq \{0,1,\star\}^V$ 如下。对于每个 $i \in [m]$，将一个概念 $h_i \colon V \to \{0,1,\star\}$ 与二分团 B_i 关联，其定义为

$$h_i(v) = \begin{cases} 0 & \text{if } v \in L_i \\ 1 & \text{if } v \in R_i \\ \star & \text{otherwise} \end{cases}.$$

我们首先注意到，该概念类的 Littlestone 维数至多为 2。

断言 3.1. LD($\mathbb{G}$) ≤ 2。

证明。我们证明，将 $\mathbb{G}$ 视为矩阵时，它不包含 $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ 作为子矩阵；然后证明，存在这个子矩阵是 Littlestone 维数大于 2 的必要条件。

如果 $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ 作为子矩阵出现在 $\mathbb{G}$ 中，那么存在 $i \neq j$ 和 $u \neq v \in V(G)$，使得$h_i(v) = h_j(v) = 1$，且 $h_i(u) = h_j(u) = 0$。然而，这意味着 $v \in R_i \cap R_j$ 和 $u \in L_i \cap L_j$，进而推出边 $\{u, v\}$ 同时被 B_i 和 B_j 覆盖，这与每条边恰好被覆盖一次的假设矛盾。

另一方面，对于利特尔斯通维数大于 2 的类 $\mathbb{H} \subseteq \{0,1,\star\}^{\mathcal{X}}$，存在一棵被 $\mathcal{X}$ 标签的高度为 3 的满二叉树打散。特别地，存在 h、$h' \in \mathbb{H}$ 以及点 $x_{()}$、x_1、x_{10}，使得

$$\begin{aligned} h(x_{()}) = 1, \quad h(x_1) = 0, \quad h(x_{10}) = 0, \\ h'(x_{()}) = 1, \quad h'(x_1) = 0, \quad h'(x_{10}) = 1. \end{aligned}$$

这意味着，限制在列 $\{x_{()}, x_1\}$ 和行 $\{h, h'\}$ 上的子矩阵为 $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$。由此得出 LD($\mathbb{G}$) ≤ 2。

□

定理 1.6 的证明。考虑上面的部分概念类 $\mathbb{G} \subseteq \{0,1,\star\}^V$。根据断言 3.1，我们有 $\mathrm{LD}(\mathbb{G}) \leq 2$。我们证明，对于 $\mathbb{G}$的任意消歧 $\overline{\mathbb{G}}$，都有 $\mathrm{VC}(\overline{\mathbb{G}}) \geq \Omega(\log\log n)$。这里的论证与定理 1.3 的证明类似。

考虑 $\overline{\mathbb{G}}$ $\mathbb{G}$ 的一种消歧。注意，如果 $\overline{\mathbb{G}}$ 中两列 u 和 v 相同，那么 u 和 v 之间没有边；否则，某个 h_i 会对 u 和 v 中的一个赋值 0，而对另一个赋值 1。因此，如果两列 u 和 v 相同，我们可以给对应的顶点赋予相同的颜色。由此，$\overline{\mathbb{G}}$ 中不同列的数量至少为色数 $\chi(G)$ $\geq 2^{\Omega(k^2)}$。根据 SSP 引理（定理 1.7），如果 $\mathrm{VC}(\overline{\mathbb{G}}^\top) \leq d$，那么 $\overline{\mathbb{G}}$ 至多有 $O(m^d)$ 个不同的列。因此，

$$2^{\Omega(k^2)} \leq O(m^d).$$

代入 $m = 2^{\tilde{O}(k)}$ 可得 $d = \tilde{\Omega}(k)$。最后，

$$\mathrm{VC}(\overline{\mathbb{G}}) \geq \Omega(\log \mathrm{VC}(\overline{\mathbb{G}}^\top)) \geq \Omega(\log k) \geq \Omega(\log\log n).$$

至此，定理 1.6 第一部分的证明完成。

对于第二部分，我们采用 AHHM22 的定理 11] 证明中的相同构造。设 $\mathbb{H}_\infty$ 是由 $n \in \mathbb{N}$ 上 $\mathbb{H}_n$ 的若干个不相交副本构成的并；每个副本都支撑在一个定义域 $\mathcal{X}_n$ 上，这些定义域彼此不相交，且与 $\mathbb{H}_n$ 也不相交； $\mathbb{H}_n$ 的部分概念在其定义域之外扩展为 $\star$。由于 $\mathbb{H}_\infty$ 的任一消歧 $\mathbb{H}$ 都同时消歧所有 $\mathbb{H}_n$，Sauer-Shelah-Perles 引理意味着 $\mathrm{VC}(\mathbb{H})$ 必须是无限的。

□

3.2 通过 SOA 算法进行消歧（定理 1.12）

本节旨在证明定理 1.12。

定理 1.12 的证明。我们通过证明以下命题来完成证明：对于任意 $r, d \in \mathbb{N}$，存在一个定义在 $[n]$ 上的部分概念类 $\mathbb{H}_{r,d}$，其中 $n = d(2^r + r)$，使得 $d \leq \mathrm{LD}(\mathbb{H}_{r,d}) \leq d+1$，且 SOA 消歧的 VC 维数为 $\geq dr$，并且至少有 2^{dr} 个不同的行。对于 n 的其他情形，只需平凡地扩展定义域即可。

对于任意 $r, d \in \mathbb{N}$，定义

$$\mathcal{F}_{r,d} = \{F \subseteq [d2^r] : |F| = d\}.$$

注意 $|\mathcal{F}_{r,d}| = \binom{d2^r}{d} \geq 2^{dr}$。我们将 $\mathcal{F}_{r,d}$ 中的集合依次记为 $F_1, \ldots, F_{\binom{d2^r}{d}}$ 按自然顺序排列。

接下来，我们定义定义域为 $[d(2^r + r)]$ 的部分概念类 $\mathbb{H}_{r,\,d}$。该类由部分概念 $h_{i,\,j}$ 构成，其中 $i \in [\binom{d2^r}{d}]$ 和 $j \in [dr]$ 定义如下：

$$h_{i,j}(x) = \begin{cases} 1 & \text{if } x \in F_i \\ 0 & \text{if } x \in [d2^r] \setminus F_i \\ \beta(i,j) & \text{if } x = d2^r + j \\ \star & \text{otherwise} \end{cases},$$

其中，$\beta(i,j)$ 表示 i 的 dr 位二进制表示中的第 j 位（若 $i \in [2^{dr}]$），否则 $\beta(i,j) = \star$。

我们首先证明 $d \le \mathrm{LD}(\mathbb{H}_{r,d}) \le d+1$。注意，存在一组 2^d 个指标 $I \subseteq [d2^r]$，它们

$$\{F_i \cap [d] : i \in I\} = \mathcal{P}([d]),$$

因此，$[d]$ 可以被 $\{h_{i,1}$ 打散：$i \in I\}$，从而 $\mathrm{LD}(\mathbb{H}_{r,\ d}) \ge \mathrm{VC}(\mathbb{H}_{r,\ d}) \ge d$。另一方面，注意到对于任意 $f \in \mathbb{H}_{r,\ d}$，$|f^{-1}(1)| \le d+1$，这意味着 $\mathrm{LD}(\mathbb{H}_{r,\ d}) \le d+1$。

接下来，我们考虑 SOA 消歧。我们声称，$\{d2^r+1, \ldots, d(2^r+r)\}$ 可以被 $\{h_{i,1} : i \in [2^{dr}]\}$打散。不存在针对 $x \in [d2^r]$的消歧。对于 $x > d2^r$，注意到对于任意$\vec{b} \in \{0,1\}^{x-1}$，要么$\mathbb{H}_{r,\ d}|_{\vec{b}} = \emptyset$，要么

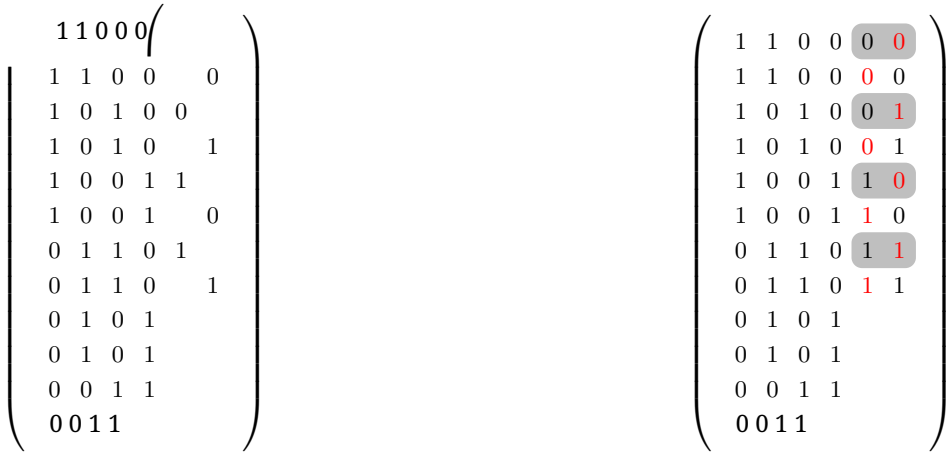
$$\mathbb{H}_{r,d}|_{\vec{b}} = \{h_{i,j} : j \in [dr]\},$$

其中 $i \in [d2^r]$ 使得 $F_i = \{k \in [d2^r] : \ b_k = 1\}$。我们重点讨论后一种情况，并将范围限制为 $i \in [2^{dr}]$。恰有一个 $c \in \{0,1\}$ 使得$\mathbb{H}_{r,\ d}|_{\vec{b}c} \neq \emptyset$，即 $c = \beta(i, x-d2^r)$，且在这种情况下$\mathbb{H}_{r,\ d}|_{\vec{b}c} = \{h_{i,c}\}$。这迫使算法通过设置 $f(x) = h_{i,\ c}(x) = \beta(i, x-d2^r)$，对每个满足 $\vec{b} \in \{0,1\}^{x-1}$的函数 f进行消歧。这样，每个$h_{i,\ j}$最终都会被消歧为同一个全函数：

$$\overline{h_{i,j}}(x) = \begin{cases} 1 & \text{if } x \in F_i \\ 0 & \text{if } x \in [d2^r] \setminus F_i \\ \beta(i, x-d2^r) & \text{if } x > d2^r \end{cases}.$$

特别地，对于每个 $i \in [2^{dr}]$，比特串 $(h_{i,\ 1}(d2^r+1), \ldots, h_{i,\ 1}(d2^r+dr))$ 是 i 的 dr 位二进制表示。这为 $\mathrm{VC}(\mathbb{H}_{r,\ d}^{\mathrm{SOA}}) \ge dr$ 提供了一个见证。 □

作为示例，我们在图 1 中给出了 SOA 消歧的一些关键步骤的矩阵表示，如下所示。



(a) $\mathbb{H}_{1,\ 2}$ 的矩阵表示：用星号填充空格　　全部为空　　(b) $\overline{\mathbb{H}_{1,\ 2}}^{\mathrm{SOA}}$：阴影条目表示打散发生的位置

图1：$\mathbb{H}_{1,\ 2}$ 及其 SOA 消歧

3.3 对 Alon-Saks-Seymour 猜想的小规模反驳（定理 1.11）

本节将详细介绍定理 1.11 的构造。起点是构造一个由 [BBDG+22] 在查询复杂度中提出的布尔函数。随后，该布尔函数经过若干次归约，被转换为一个图，具体如下所述。

我们先介绍一些与证书复杂度概念相关的基本定义。设$f\colon \{0,1\}^n \to \{0,1\}$是一个布尔函数。对于 $b \in \{0,1\}$和输入 $x \in f^{-1}(b)$，若部分输入$\rho \in \{0,1,\star\}^n$与 ρ一致，且对于每个与 ρ一致的 $x' \in \{0,1\}^n$，都有 $f(x') = b$，则称$\rho \in \{0,1,\star\}^n$为 b-证书。 ρ的大小是 ρ中非$\star$条目的数量。定义 $\mathrm{C}_b(f,x)$ 为 x的 b-证书的最小大小。 f的 b-证书复杂度记为 $\mathrm{C}_b(f)$，它是对所有 $x \in f^{-1}(b)$ 的 $\mathrm{C}_b(f,x)$ 的最大值。

无歧义　　b-证书复杂度 f，记作 $\mathrm{UC}_b(f)$，是满足以下条件的最小 k

1. 每个输入 $x \in f^{-1}(b)$ 都有一个大小至多为 k 的 b-证书 ρ_x；
2. 对于每个 $x \neq y \in f^{-1}(b)$，都有 $\rho_x \neq \rho_y$。

[BBDG+22] 的主要结果是 UC_1 与 C_0 之间的以下分离。

定理 3.2（[BBDG+22，定理 1]）。存在一个函数 $f\colon \{0,1\}^{12n^4\log^2 n} \to \{0,1\}$，使得 $\mathrm{UC}_1(f) = O(n\log^3 n)$ 且 $\mathrm{C}_0(f) = \Omega(n^2)$。

构造的下一步是将分离证书复杂度 UC_1 和 C_0 的函数转化为一个通信问题。这通过“提升”技巧实现：给定一个函数 $f\colon \{0,1\}^n \to \{0,1\}$ 和一个“构造块”函数 $g\colon \{0,1\}^k \times \{0,1\}^k \to \{0,1\}$，我们将 $f \circ g^n\colon \{0,1\}^{nk} \times \{0,1\}^{nk} \to \{0,1\}$ 定义为

$$f \circ g^n([x_1,\dots,x_n],[y_1,\dots,y_n]) = f(g(x_1,y_1),\dots,g(x_n,y_n)).$$

对于通信问题 $f\colon \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ 和 $b \in \{0,1\}$，令 $\mathrm{Cov}_b(f)$ 表示覆盖 f 的所有 b 项所需的 b- 单色矩形的最小数量。我们用 $\mathrm{UCov}_b(f)$ 表示划分 f 的所有 b 项所需的 b- 单色矩形的最小数量。下面的定理建立了通信复杂度参数与证书复杂度参数之间的联系。

定理 3.3（[GLM+16，定理 33]）。存在一个构造块 $g\colon \{0,1\}^k \times \{0,1\}^k \to \{0,1\}$ 满足$k = \Omega(\log n)$，使得对于每个 $f\colon \{0,1\}^n \to \{0,1\}$，我们都有

$$\log \mathrm{Cov}_b(f \circ g^n) = \Omega(k\,\mathrm{C}_b(f)).$$

注意，对于每个 $b \in \{0,1\}$，都有 $\log \mathrm{UCov}_b(f \circ g^n) \le 2k\,\mathrm{UC}_b(f)$。结合定理 3.3，我们可以将定理 3.2 中 UC_1 与 C_0 的分离“提升”为 UCov_1 与 Cov_0的分离。

推论 3.4。存在一个函数 $f\colon \{0,1\}^{O(n^4\log^3 n)} \times \{0,1\}^{O(n^4\log^3 n)} \to \{0,1\}$，使得

$$\log \mathrm{Cov}_0(f) = \Omega(n^2) \qquad and \qquad \log \mathrm{UCov}_1(f) = n\log^4 n.$$

接下来，我们将展示如何把这些通信参数转换为完全二部图划分数和色数这两个图参数。

引理 3.5. 设 $h\colon \{0,1\}^t \times \{0,1\}^t \to \{0,1\}$ 是一个布尔函数，满足 $\mathrm{Cov}_0(h) = c$ 且 $\mathrm{UCov}_1(h) = m$。存在一个图 $G = (V, E)$，其顶点数至多为 2^{2t}，且 $\mathrm{bp}(G) \le m^2$ 和 $\chi(G) \ge \sqrt{c}$。

证明。定义图 G with $V := h^{-1}(0)$ 如下。若两个顶点 (x, y)、$(x', y') \in V$ 在 G 中相邻，当且仅当 $h(x, y') = 1$ 或 $h(x', y) = 1$。根据构造，若 $\{(x_1, y_1) \dots, (x_\ell, y_\ell)\} \subseteq V$ 是一个独立集，那么 $\{x_1, \dots, x_\ell\} \times \{y_1, \dots, y_\ell\}$ 就是 h 的一个 0-单色矩形。因此，G 的每个使用 $\chi(G)$ 种颜色的正常顶点着色，都对应于 h 的一个 0-覆盖，其中包含 $\chi(G)$ 个 0-单色矩形。因此，$\chi(G) \ge c$。

接下来，我们证明存在一个小的二分团集合，使得 E 的每条边都被这些二分团覆盖至少一次且至多两次。令 $h^{-1}(1) = \bigcup_{i=1}^{m} (A_i \times B_i)$ 为将 $h^{-1}(1)$ 划分为 m 个 1-单色矩形的一种划分。注意，每个 1-单色矩形 $A_i \times B_i$ 都对应于 G 中的一个二分团 $Q_i := S_i^- \times S_i^+$，其中

$$S_i^- := \{(x, y) \in V(G) : \ x \in A_i\} \text{ and } S_i^+ = \{(x, y) \in V(G) : \ y \in B_i\}.$$

注意，G 中的每条边 $\{(x, y), (x', y')\}$ 至少被 $Q_1, \dots, Q_m$ 覆盖一次，至多被覆盖两次；后一种情况发生在 $h(x, y') = h(x', y) = 1$ 时。

因此，我们构造出了一个至多有 2^{2t} 个顶点的图 G，使得 $\chi(G) \ge c$，并且至多有 m 个二分团，其中 G 中的每条边都出现在至少一个、至多两个二分团中。

定义 H_2 为 G 的子图，它由 $Q_1, \dots, Q_m$ 中恰好被两个二分团覆盖的所有边组成。对每个 $i, j \in [m]$，定义 $Q_{ij} = (S_i^- \cap S_j^+) \times (S_i^+ \cap S_j^-)$。注意，每个 Q_{ij} 都是 H_2 的一个二分团，而且 H_2 的每条边恰好出现在一个 Q_{ij} 中。因此，H_2 的完全二部图划分数至多为 m^2。现在，若 $\chi(H_2) \ge \sqrt{c}$，我们就得到所需的图 H_2。否则，假设 $\chi(H_2) < \sqrt{c}$，并考虑 H_2 的一个使用 $\sqrt{c}$ 种颜色、颜色类为 $V_1, \dots, V_{\sqrt{c}}$ 的正常顶点着色。由于 $\chi(G) \ge c$，必存在 i，使得 G 在 V_i 上的诱导子图（记为 $G[V_i]$）满足 $\chi(G[V_i]) \ge \sqrt{c}$。由于 V_i 是 H_2 的一个独立集，因此将二分团 $Q_1, \dots, Q_m$ 限制到 V_i 后，构成了 $G[V_i]$ 的一个二分团划分。 □

引理 3.5 和推论 3.4 共同推出定理 1.11。

注 3.6。除了为图的规模提供有效界之外，引理 3.5 还简化了先前工作 [BBDG+22、Göö15、BLT14、Yan91] 中为实现二分团划分与色数之间的超多项式分离而采用的原始归约链。下面我们将简要介绍原始证明，并强调其中的差异。

(i) 与我们对定理 1.11 的证明类似，该归约链始于推论 3.4 提供的函数 f，使得

$$\log \mathrm{Cov}_0(f) = \Omega(n^2) \qquad \textit{and} \qquad \log \mathrm{UCov}_1(f) = n \log^4 n.$$

(ii) Yannakasis [Yan91]（另见 [Göö15，图 1]）展示了如何使用 f 在 $\mathrm{UCov}_1(f) = 2^{O(n \log^4 n)}$ 个顶点上构造一个图 F，使得 F 的每个团-稳定集分离器的大小至少为 $\mathrm{Cov}_0(f) = 2^{\Omega(n^2)}$。这里，团-稳定集分离器是 F 中的一组割，使得对于 F 中的每一对不相交的团 C 和稳定集 I（C，I），该集合中都存在一个割（A，B），满足 $C \subseteq A$ 和 $I \subseteq B$。

(iii) Bousquet 等人， [BLT14，引理 23] 展示了如何使用 F 构造一个新的图 G，其所谓的有向完全二部图打包数至多为 $2^{n\log^4 n}$，且色数为

$$\chi(G) \geq 2^{\Omega(n^2)}。$$

(iv) 随后，通过 BLT 中的最后一步归约，将图 G 转化为另一张图 H 中二分团划分数与色数之间的分离。

上述归约链不适用于我们的应用，因为步骤 (iii) 中的图 G为 F的每一对团 C与稳定集 I（即 (C, I)）都设置了一个顶点，因此无法有效地限制图 G的顶点数。我们对定理 1.11 的证明绕过了步骤 (ii)，采用更直接的方法构造一张规模较小的图 G，使其具有与步骤 (iii) 中的图 G 类似的性质。

4 结论

仍有一些自然问题尚未得到解答。第一个问题是：在更强的假设 $\mathrm{LD}(\mathbb{H}) = 1$ 下，是否存在一个类似的定理 1.6 示例 $\mathbb{H}$？

问题 4.1。设 $\mathbb{H}$ 是一个满足 $\mathrm{LD}(\mathbb{H}) = 1$的部分类。是否存在一个由全类 $\overline{\mathbb{H}}$ 对 $\mathbb{H}$进行的消歧，使得 $\mathrm{LD}(\overline{\mathbb{H}}) < \infty$？是否存在满足 $\mathrm{VC}(\overline{\mathbb{H}}) < \infty$的消歧？

定理 1.10 表明，对于部分类而言，具有多项式增长并不足以保证 PAC 可学习性。一个自然的定理修正版，是采用限制更严格的线性增长假设。

问题 4.2. 设 $\mathbb{H} \subseteq \{0, 1, \star\}^{\mathcal{X}}$ 具有参数为 1 的多项式增长。是否存在 $\mathbb{H}$ 的一个消歧 $\overline{\mathbb{H}}$，使得 $\mathrm{VC}(\overline{\mathbb{H}}) < \infty$？

另一个问题是，能否将定理 1.6 中的下界 $\Omega(\log\log n)$ 提高到 $\Omega(\log n)$。

问题 4.3. 能否将定理 1.6 中的下界提高到 $\mathrm{VC}(\overline{\mathbb{H}}) \geq \Omega(\log n)$？

禁止组合模式。证明概念类的 VC 维数上界的一种自然方法，是证明该类不包含某个特定的组合模式。例如， [AHHM22] 中定理 1.3 的构造利用了这样一个事实：该概念类（视为矩阵）不包含组合模式 $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ 和 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$；任何满足 $\mathrm{VC}(\mathbb{H}) \geq 2$ 的概念类 $\mathbb{H}$ 都包含这些模式。类似地，定理 1.6 中的对偶构造禁止了模式 $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$，而任何满足 $\mathrm{LD}(\mathbb{H}) \geq 3$ 的概念类 $\mathbb{H}$ 都必然包含该模式。

问题 4.4。假设 $\mathbb{H} \subseteq \{0, 1, \star\}^{[n]}$ 不包含模式 $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$。每个消歧 $\overline{\mathbb{H}}$ 是否都满足 $\mathrm{VC维}(\mathbb{H}) = O(1)$？

致谢 我们感谢 Mika G¨o¨os 澄清了 [BLT14、G¨o¨o15、GLM+16、BBDG+22]中的归约。

参考文献

[AHHM22] Noga Alon、Steve Hanneke、Ron Holzman 和 Shay Moran。部分概念类的 PAC 可学习性理论。载于 2021 年 IEEE 第 62 届计算机科学基础年会（FOCS），第 658–671 页。IEEE，2022 年。[BBDG+22] Kaspars Balodis、Shalev Ben-David、Mika G¨o¨os、Siddhartha Jain 和 Robin Kothari。无歧义析取范式与 Alon-Saks-Seymour。载于 2021 年 IEEE 第 62 届计算机科学基础年会——FOCS 2021，第 116–124 页。IEEE 计算机学会，加利福尼亚州洛斯阿拉米托斯， [2022] ©2022 年。[BLT14] N. Bousquet、A. Lagoutte 和 S. Thomass´e。团与独立集。European J. Combin.，40:73–92，2014 年。[GLM+16] Mika G¨o¨os、Shachar Lovett、Raghu Meka、Thomas Watson 和 David Zuckerman。矩形是非负侏儒函数。SIAM J. Comput.，45(5):1835–1869，2016 年。[G¨o¨o15] Mika G¨o¨os。团与独立集的下界。载于 2015 年 IEEE 第 56 届计算机科学基础年会——FOCS 2015，第 1066–1076 页。IEEE 计算机学会，加利福尼亚州洛斯阿拉米托斯，2015 年。[HS12] Hao Huang 和 Benny Sudakov。Alon-Saks-Seymour 猜想及相关问题的一个反例。《组合数学》，32(2):205–219，2012 年。[Lit88] Nick Littlestone。无关属性大量存在时的快速学习：一种新的线性阈值算法。《机器学习》，2(4):285–318，1988 年。[Mat02] Jiˇr´ı Matouˇsek，编。《离散几何讲义》。Springer 纽约，2002 年。[Ros58] Frank Rosenblatt。感知机：用于大脑中信息存储和组织的概率模型。《心理学评论》，65 6:386–408，1958 年。[Sau72] Norbert Sauer。论集合族的密度。《组合理论杂志》A 辑，13(1):145–147，1972 年。[Yan91] Mihalis Yannakakis。用线性规划表达组合优化问题。J. Comput. System Sci.，43(3):441–466，1991 年。