arXiv:2601.18663v1 [math.LO] 26 Jan 2026

PAC学习的一致可计算性

VASCO BRATTKA [a] 和 GUILLAUME CHIRACHE [b]

[a]联邦国防军大学计算机科学学院，慕尼黑，85577 诺伊比贝格，德国；以及南非开普敦大学数学与应用数学系，朗德博斯 7700*e-mail address*：Vasco.Brattka@cca-net.de

[b] ´巴黎综合理工学院，91128 帕莱索，法国*e-mail address*：guillaume.chirache.2022@polytechnique.org

Abstract. 我们利用 Weihrauch 复杂度研究 PAC学习的一致可计算性。我们重点考察闭概念类，它们分别以正信息、负信息或完整信息表示。在诸多结果中，我们证明，基于正信息的适当 PAC 学习等价于贝尔空间上的极限运算；而基于正信息的非适当 PAC 学习则与弱柯尼希引理密切相关，并且当我们拥有关于可接受假设的一些负信息时，它甚至与该引理等价。如果允许任意假设，那么基于正信息的非适当 PAC 学习仍属于有限型DNC范围，这意味着它是非确定性可计算的，但无法通过概率算法实现。这些结果也可视为对统计学习基本定理之构造性程度的分类。如果将 VC 维数的上界作为额外输入信息提供，上述结果也都成立。我们还研究了如下问题：如果未给出 VC 维数，而只是保证其有限，或者概念类以负信息或完整信息表示，这些结果会受到怎样的影响。最后，我们还对 VC 维数运算本身的复杂度进行了分类；这一问题本身也具有独立研究价值。结果表明，对于正信息或完整信息，该运算等价于二进制排序问题；对于负信息，则等价于排序跳跃。这一分类也使我们能够得出有关 PAC 可学习性之博雷尔复杂度的结论。

关键词：可计算分析、Weihrauch 复杂度和机器学习。

1. Introduction

近年来，研究界对 PAC 学习的可计算性性质越来越感兴趣，已有多位作者研究 PAC 学习的可计算变体，相关研究始于 Agarwal、Ananthakrishnan、Ben-David、Lechner 和 Urner 的合作工作 [2, 1, 45, 27, 30, 3, 36]。Soloveichik 早已证明了这一方向上的一项初步结果 [44]。Kattermann 和 Krapp 最近的工作提供了对这些结果的良好综述 [36]。在上述大多数研究中，PAC 学习都是从非统一视角加以考察，重点在于单个概念类及其学习器的可计算性性质。我们则采取不同立场，从统一视角研究 PAC 学习，并运用可计算分析 [49, 17] 和 Weihrauch

复杂度 [15]。可计算分析中的概念已应用于学习理论领域的其他研究，例如 [24, 26, 31, 22, 23, 8]。尤其值得一提的是，Ackerman、Asilis、Di、Freer 和 Tristan [1] 已使用 Weihrauch 可约性分析连续情形下的 PAC 学习。

使用 Weihrauch 复杂度的统一方法，使我们能够同时对以下三个问题给出简洁而精确的答案。这些问题大致分别来自逻辑、可计算性理论和理论计算机科学的视角：

（1）统计学习基本定理的非构造性有多强？（2）对于以下问题，其可计算性理论复杂度是什么：给定一个具有有限 VC 维数的概念类，为其寻找一个学习器。（3）PAC学习具有什么计算性质：它是可计算的、lim 可计算的、非确定性可计算的，还是其他形式？

为了描述我们的结果，我们首先回顾计算学习理论中的一些基本定义。有关细节，请读者参阅标准参考文献，例如 [41, 38]。

用 $\mathcal{S} := (\mathbb{N} \times \{0,1\})^*$ 表示由样本组成的集合，其中样本是由有序对（x， y）构成的有限字 $\in \mathbb{N} \times \{0,1\}$。一个假设是一个函数 $h\colon \mathbb{N} \to \{0,1\}$，而一个学习器是一个函数 $A : \mathcal{S} \to 2^{\mathbb{N}}$，它为每个样本生成一个假设。我们也交替地将 $2^{\mathbb{N}}$ 中的点视为二进制序列，或视为 $\mathbb{N}$ 的子集，并通过其特征函数来识别。用 $\mathcal{L}$ 表示所有学习器的集合。一个假设类是一个集合 $\mathcal{H} \subseteq 2^{\mathbb{N}}$。有时我们使用两个假设类 $\mathcal{C} \subseteq \mathcal{H} \subseteq 2^{\mathbb{N}}$，此时也将其中较小的一个 $\mathcal{C}$ 称为概念类。PAC学习的目标是：对于每个合适的概念类 $\mathcal{C} \subseteq \mathcal{H}$，找到一个学习器 $A\colon \mathcal{S} \to \mathcal{H}$，使该学习器能够很好地适应该概念类。学习器很好地适应概念类的含义，是用概率来表达的；粗略地说，目标是让学习器仅以较小的概率犯下较大的错误（PAC学习代表“可能近似正确”学习）。

我们使用概率分布 $\mathcal{D}\colon \ \mathbb{N} \times \{0,1\} \to \mathbb{R}$，它们只是取非负值且满足$\sum_{(n,i)\in\mathbb{N}\times\{0,1\}} \mathcal{D}(n, i) = 1$的函数。我们用 $\mathcal{P}$ 表示此类概率分布的集合。对于概率分布 $\mathcal{D}$ 和假设 h，我们用真实误差（风险）表示 h 根据 $\mathcal{D}$ 所具有的

$$L_{\mathcal{D}}(h) := \mathbb{P}_{(x,y)\sim\mathcal{D}}[h(x) \neq y].$$

这里， $\mathbb{P}_{(x,y)\sim\mathcal{D}}$表示所有按照 $\mathcal{D}$分布的$(x, y) \in \mathbb{N} \times \{0,1\}$ 上的概率。同样，我们用 $\mathbb{P}_{S\sim\mathcal{D}^n}$ 表示所有按照 $\mathcal{D}$独立同分布的长度为 n 的样本 $S \in \mathcal{S}$ 上的概率。

现在，我们可以定义学习器 PAC 学习一个假设类是什么意思。我们使用康托尔配对函数，其中 $\langle n, k\rangle := \frac{1}{2}(n + k)(n + k + 1) + k$对所有$n, k \in \mathbb{N}$都成立。

定义 1（PAC 学习）。设 $\varnothing \neq \mathcal{C} \subseteq \mathcal{H} \subseteq 2^{\mathbb{N}}$ 为假设类， $A : \mathcal{S} \to \mathcal{H}$ 为学习器， $m\colon \mathbb{N} \to \mathbb{N}$ 为我们称为样本复杂度的函数。若对于所有 $i,\ j \in \mathbb{N}$ 和 $n \geq m\langle i,\ j\rangle$，以及每个定义在 $\mathbb{N} \times \{0,1\}$ 上的概率分布 $\mathcal{D}$，下述条件均成立，则称 A*PAC 学习* $\mathcal{C}$*相对于* $\mathcal{H}$，*在* m *的条件下成立*：

$$\mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \leq \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h) + 2^{-i}\right] > 1 - 2^{-j}. \tag{1.1}$$

若 $\mathcal{C} = \mathcal{H}$，则称 A （恰当地）*PAC 学习* $\mathcal{C}$，且 m。若 $\mathcal{C} \subseteq \mathcal{H}$ 且 $\mathcal{H} = 2^{\mathbb{N}}$，则称 A *非恰当地 PAC 学习* $\mathcal{C}$，且 m。

参数 $\varepsilon = 2^{-i}$ 有时称为准确度，而 $\delta = 2^{-j}$ 称为置信度。粗略地说，该条件表示：对于所有足够大规模的样本 S $n \geq m\langle i,\ j\rangle$以高置信度，学习器产生的假设 A（S） $\in \mathcal{H}$ 的真实误差应以高准确度接近概念类 $\mathcal{C}$ 中可达到的最优真实误差$\inf_{h\in\mathcal{C}} L_{\mathcal{D}}$（$h$）（两处“高”分别指较大的 i 和 j）。非适当 PAC 学习也称为表示无关学习 [41,备注 3.2]。不过，我们会采用更简短的“非适当学习”这一术语。

定义 2 (PAC 可学习性)。 设 $\varnothing \neq \mathcal{C} \subseteq \mathcal{H} \subseteq 2^{\mathbb{N}}$ 为假设类。若存在一个学习器 $A\colon \mathcal{S} \to \mathcal{H}$ 以及一个样本复杂度$m : \mathbb{N} \to \mathbb{N}$，使得 A 相对于 $\mathcal{H}$ 以 mPAC 学习 $\mathcal{C}$，则称$\mathcal{C}$ 相对于 $\mathcal{H}$ *PAC* 可学习。若 $\mathcal{C} = \mathcal{H}$，则简称 $\mathcal{C}$ （适当）*PAC* 可学习。若 $\mathcal{C} \subseteq \mathcal{H}$ 且 $\mathcal{H} = 2^{\mathbb{N}}$，则简称 $\mathcal{C}$ 非适当 *PAC* 可学习，

若无特别说明，PAC 可学习性总是指适当 PAC 可学习性。一个显而易见的问题是，是否存在 PAC 可学习的概念类 $\mathcal{C}$，如果存在，又该如何刻画它们。统计学习基本定理对此给出了完整的答案。该定理由 Blumer、Ehrenfeucht、Haussler 和 Warmuth 证明 [4, 定理 2。1]（另见 [41, 定理 6。7]），其工作建立在 Vapnik 和 Chervonenkis 早期研究的基础上 [47]。基本定理指出，$\mathcal{C}$ 是 PAC 可学习的，当且仅当其 VC 维数有限。

我们回顾一下*VC* 维数（也称为*Vapnik Chervonenkis* 维数）的定义。首先，对于每个概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$ 和 $A \in 2^{\mathbb{N}}$，我们定义 $\mathcal{C} \cap A := \{B \cap A\colon B \in \mathcal{C}\}$，并称 A 被 $\mathcal{C}$ 打散，如果 $|\mathcal{C} \cap A| = 2^{|A|}$，其中 $|X|$ 表示集合 X 的基数。使用这一术语，我们可以定义

$$\mathrm{VCdim}(\mathcal{C}) := \sup\{|A| \in \mathbb{N} : A \subseteq \mathbb{N} \text{ is finite and shattered by } \mathcal{C}\} \in \mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$$

对于每个非空类 $\mathcal{C}$。如果 $\mathcal{C} = \varnothing$，那么上确界是在空集上取值，自然的取值应为 -1。然而，为了保持在自然数范围内，我们只好人为地定义 $\mathrm{VCdim}(\varnothing) := 0$。

对于可 PAC 学习的类 $\mathcal{C}$，存在 PAC 学习器的通用类。若学习器 $A\colon \mathcal{S} \to \mathcal{C}$ 满足以下条件，则称其为类 $\mathcal{C}$ 的经验风险最小化器

$$A(S) \in \mathrm{argmin}_{h\in\mathcal{C}} L_S(h) := \{h \in \mathcal{C} : L_S(h) = \min\{L_S(h) : h \in \mathcal{C}\}\}$$

对于所有 $S \in \mathcal{S}$。这里，我们用以下符号表示

$$L_S(h) := \frac{|\{i \in \{1, ..., n\} : h(x_i) \neq y_i\}|}{n}$$

样本 $S = ((x_1, y_1) ..., (x_n, y_n)) \in \mathcal{S}$对于假设$h$ 的所谓 经验风险 $\mathbb{N} \to \{0, 1\}$。

现在，我们可以总结统计学习基本定理的一个与我们相关的版本（参见 [4, 定理 2.1] 和 [41, 定理 6.7 和 6.8]）。

定理 3（统计学习基本定理）。令 $\varnothing \neq \mathcal{C} \subseteq 2^{\mathbb{N}}$。那么

$$\mathcal{C} \text{ is PAC learnable} \iff \mathrm{VCdim}(\mathcal{C}) \text{ is finite.}$$

事实上，还可以作出以下额外的定量表述[1] ，其中常数是固定的c, $C > 0$ 且与 $\mathcal{C}$ 无关：

(1) 若 VCdim($\mathcal{C}$) $\leq d$，则每个经验风险最小化器 $A\colon \mathcal{S} \to \mathcal{C}$ 都能通过样本复杂度 m 小于或等于 $m_d\langle i, j\rangle := c \cdot 2^i \cdot (d \cdot i + j)$，对 $\mathcal{C}$ 进行 PAC 学习。(2) 若 $d \leq$ VCdim($\mathcal{C}$)，则不存在任何学习函数 $A\colon \mathcal{S} \to \mathcal{H}$，能够对任意 $\mathcal{H} \supseteq \mathcal{C}$ 进行 PAC 学习；相对于 $\mathcal{H}$，其样本复杂度 m 小于 $M_d\langle i, j\rangle := C \cdot 2^i \cdot (d + j)$。

不失一般性，我们可以假设 $c \in \mathbb{N}$，并且 $C \in \mathbb{R}$ 是可计算的。这意味着 $m_d\colon \mathbb{N} \to \mathbb{N}$ 和 $M_d\colon \mathbb{N} \to \mathbb{R}$ 都是可计算函数（在 d 中亦如此）。

由于我们想要分析统计学习基本定理的一致可计算性理论内涵，因此需要考察其作为一个 $\forall\exists$–陈述的逻辑形式。对于最有趣的方向“$\Longleftarrow$”，我们基本上得到了适当 PAC 学习情形下的如下逻辑形式：

$$(\forall \mathcal{C})(\mathrm{VCdim}(\mathcal{C}) < \infty \Longrightarrow (\exists (A, m) \in \mathcal{L} \times \mathbb{N}^{\mathbb{N}})(A \text{ PAC learns } \mathcal{C} \text{ with } m))$$

为了分析这一陈述的 Weihrauch 复杂度，我们可以直接将相应的计算问题解释为一个多值映射 PPAC，它以$\mathcal{C}$ 为输入，并输出一个可能的数对 (A, m)。然而，这立即引出了若干问题：

(A) 我们如何表示 $\mathcal{C}$，以及这里允许使用什么样的集合？(B) 将 d =VC 维数($\mathcal{C}$)的一个上界作为额外的输入信息提供，与仅承诺 $\mathcal{C}$ 的 VC 维数有限（或具有某个固定的有限值）相比，是否会产生区别？(C) 我们是否需要在输出端给出样本复杂度 m，还是说没有它 A 也有用？

关于问题 (C)，很容易看出，从实践角度来说，没有相应样本复杂度 m 的学习器 A 几乎毫无用处。稍后我们会在第 6 节末尾简要再谈这个问题，并在那里进行一个案例研究；但我们不会系统地研究所有不将样本复杂度作为输出的 PAC 学习变体。

关于问题 (A)，我们选择将研究范围限制在拓扑闭概念类 $\mathcal{C}$。这一选择至少有两个理由。首先，一个集合与其闭包的 VC 维数相同，因此闭概念类是一般概念类的自然代表。将所有对象限制为闭集，不会使 PAC 可学习性方面的任何实质内容有所损失。这一点尤其适用于非适当情形，在这种情形下，我们不要求学习器生成的假设属于给定的概念类 $\mathcal{C}$。其次，闭集类的基数为连续统，可以使用可计算性理论和可计算分析中的标准技术来处理。事实上，在可计算分析中，表示闭子集 $2^{\mathbb{N}}$ 的空间 $\mathcal{A}$ 有三种成熟的标准方式 [20, 34]。根据分别使用完全信息、正信息或负信息，我们将这些空间记为 $\mathcal{A}$、 $\mathcal{A}_+$ 和 $\mathcal{A}_-$。下一节将给出更精确的定义。

总的来说，我们将研究 PAC 学习的三种不同的一般版本。除了通常的适当 PAC 学习和非适当 PAC 学习之外，事实证明，考虑

[1]该定理给出的界通常用 $\varepsilon = 2^{-i}$ 和 $\delta = 2^{-j}$表示，此时它们写作 经过简单计算即可得到我们所采用的定理版本中的界（常数 c 和 C的取值作了适当调整）。

$$m_d(\varepsilon, \delta) = c \cdot \frac{d \cdot \log(\frac{1}{\varepsilon}) + \log(\frac{1}{\delta})}{\varepsilon} \text{ and } M_d(\varepsilon, \delta) = C \cdot \frac{d + \log(\frac{1}{\delta})}{\varepsilon}.$$

另一种 PAC 学习变体也很有用，我们称之为 相对 *PAC* 学习；在这种学习中，除了概念类 $\mathcal{C}$，我们还提供更大的假设类 $\mathcal{H} \supseteq \mathcal{C}$，允许从中选取结果。在这种情况下，我们考虑由正信息给出的类 $\mathcal{C}$，并以否定方式描述允许的假设，因为这样可以排除不需要的假设。

按照 Weihrauch 复杂度中的惯例，[15], 我们使用部分多值映射 $f :\subseteq X \rightrightarrows Y$来表示数学问题。其思想是，dom($f$) 是问题的实例集合，而 $f(x) \subseteq Y$则是对于特定实例$x \in$ dom(f) 的问题的解集合。这样，我们便可如下表示 PAC 学习问题的三种变体。

定义 4（PAC 学习问题）。我们定义如下内容。

（1）相对 PAC 学习问题

$$\begin{aligned} \mathsf{RPAC} :\subseteq \mathcal{A}_+ \times \mathcal{A}_- \times \mathbb{N} &\rightrightarrows \mathcal{L} \times \mathbb{N}^{\mathbb{N}} \\ (\mathcal{C}, \mathcal{H}, d) &\mapsto \{(A, m) : A \text{ PAC learns } \mathcal{C} \text{ relative to } \mathcal{H} \text{ with } m\} \end{aligned}$$

其 dom(RPAC) $= \{(\mathcal{C},\ \mathcal{H},\ d) : \mathrm{VCdim}(\mathcal{C}) \leq d$ 且 $\varnothing \neq \mathcal{C} \subseteq \mathcal{H}\}$。

（2）适当的*PAC* 学习问题

$$\begin{aligned} \mathsf{PPAC} :\subseteq \mathcal{A} \times \mathbb{N} &\rightrightarrows \mathcal{L} \times \mathbb{N}^{\mathbb{N}} \\ (\mathcal{C}, d) &\mapsto \{(A, m) : A \text{ properly PAC learns } \mathcal{C} \text{ with } m\} \end{aligned}$$

其定义域为 dom(PPAC) $= \{(\mathcal{C},\ d)$: VC 维数($\mathcal{C}$) $\leq d$ 和 $\mathcal{C} \neq \varnothing\}$。

（3）非适当 PAC 学习问题

$$\begin{aligned} \mathsf{IPAC} :\subseteq \mathcal{A} \times \mathbb{N} &\rightrightarrows \mathcal{L} \times \mathbb{N}^{\mathbb{N}} \\ (\mathcal{C}, d) &\mapsto \{(A, m) : A \text{ improperly PAC learns } \mathcal{C} \text{ with } m\} \end{aligned}$$

其定义域为 dom(IPAC) $= \{(\mathcal{C},\ d)$: VC 维数($\mathcal{C}$) $\leq d$ 和 $\mathcal{C} \neq \varnothing\}$。

如果分别用 $\mathcal{A}_+$ 或 $\mathcal{A}_-$替换空间 $\mathcal{A}$，我们也使用类似的记号 $\mathsf{PPAC}^{\pm}$ 和 $\mathsf{IPAC}^{\pm}$，并在上标中使用 $+$ 或 $-$。

关于问题（B），我们将看到，给出 VC 维数的上界在某些情况下会产生影响，而在另一些情况下则不会。为了能够准确表述这些差异，我们为上述所有学习问题引入变体：固定 VC 维数的上界，且不再将其作为输入信息提供。

定义 5（有界 PAC 学习问题）。对于所有 $d \in \mathbb{N}_\infty$，我们按照定义 4 定义 $\mathsf{RPAC}_{<d}$、$\mathsf{PPAC}_{<d}$ 和 $\mathsf{IPAC}_{<d}$，但不提供输入 $d \in \mathbb{N}$，并使用以下定义域:

(1) 相对 PAC 学习$_{<d}$的 dom $= \{(\mathcal{C},\ \mathcal{H})$: VC 维数($\mathcal{C}$) $< d$ 且 $\varnothing \neq \mathcal{C} \subseteq \mathcal{H}\}$。(2) $\mathsf{PPAC}_{<d}$的 dom $= \{\mathcal{C}$: VC 维数($\mathcal{C}$) $< d$ 且 $\mathcal{C} \neq \varnothing\}$。

(3) $\mathsf{IPAC}_{<d}$的 dom $= \{\mathcal{C}$: VC 维数($\mathcal{C}$) $< d$ 且 $\mathcal{C} \neq \varnothing\}$。

我们采用类似的记法 $\mathsf{PPAC}^{\pm}_{<d}$ 和 $\mathsf{IPAC}^{\pm}_{<d}$，并使用上标 $+$ 或 $-$，表示将空间 $\mathcal{A}$分别替换为 $\mathcal{A}_+$ 或 $\mathcal{A}_-$。

对于适当和非适当 PAC 学习，我们可以在图 1 的表格中总结部分结果。对于每种情况，表中给出的可计算性概念都是相应类型概念下的最优上界。

输入	适当 PAC 学习	非适当 PAC 学习
$\mathcal{C} \in \mathcal{A}$, $d >$ VC 维数($\mathcal{C}$)	可计算	可计算
$\mathcal{C} \in \mathcal{A}_+$, $d >$ VC 维数($\mathcal{C}$)	极限可计算	非确定性可计算
$\mathcal{C} \in \mathcal{A}_-, d >$ VC 维数($\mathcal{C}$)	极限可计算	可计算
$\mathcal{C} \in \mathcal{A}$	有限心智变更可计算	有限心智变更可计算
$\mathcal{C} \in \mathcal{A}_+$	极限可计算	极限可计算
$\mathcal{C} \in \mathcal{A}_-$	双重极限可计算	双重极限可计算

图 1 ：PPAC、PPAC$<\infty$ 和 IPAC、IPAC$<\infty$ 的可计算性属性，分别对应 .

然而，使用 Weihrauch 复杂度，我们可以得到一幅比图 1 表格所呈现的图景精细得多的图景。我们还可以回答有关相应类别完备性的问题，并给出更精确的界（另见图 5 中的示意图）。例如，我们还可以推导出如下结论：从正信息进行、具有给定上界 IPAC$^+$ 的非适当 PAC 学习并非对非确定性计算完备，但它仍然不存在概率算法；尤其是，按照 [13] 的意义，它不是拉斯维加斯可计算的。这是因为存在一些可计算实例，使得其所有解都具有 PA图灵度。图 2 的示意图展示了一些 PAC 学习的不可计算版本（红色），以及 Weihrauch 复杂度中的若干基准问题（蓝色，所有问题将在下文定义）。

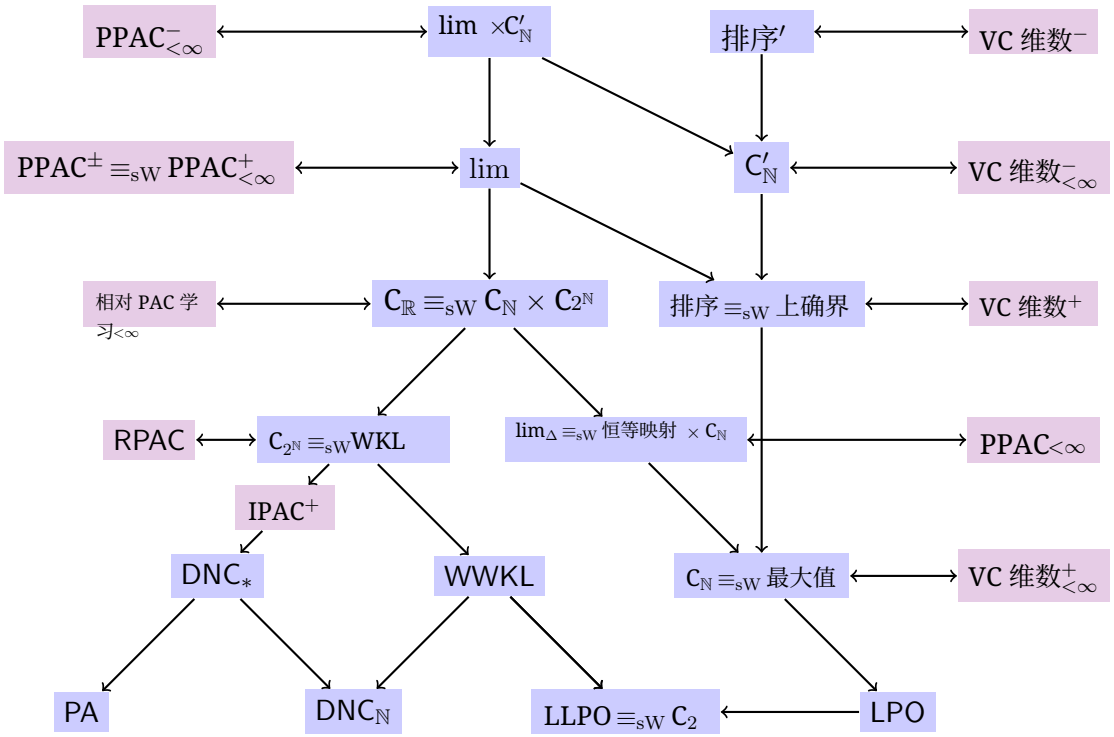

图2：强 Weihrauch 格中的 PAC 学习。

在接下来的第 2 节中，我们将给出本研究所需的可计算分析和 Weihrauch 复杂度的定义。第 3 节中，我们将研究 VC 维数的分类；这一问题本身也值得关注。将 VC 维数限制在具有有限 VC 维数的集合上进行计算，其复杂度将在我们对 PAC 学习问题的分类中发挥作用，尤其是在输入未提供 VC 维数上界的情形下。我们的结果还直接导出了 PAC 可学习性的博雷尔复杂度分类。接着，在第 4 节中，我们将研究相对 PAC 学习，并证明它等价于弱 K″onig 引理 WKL。第 5 节中，我们将研究适当 PAC 学习，并证明无论概念类是以正向还是反向方式表示，它都等价于贝尔空间上的极限运算。最后，在第 6 节中，我们将研究基于正信息的非适当 PAC 学习，并证明它严格位于问题 PA（计算皮亚诺算术的完备扩张）与 WKL（弱 K″onig 引理）之间。这一分类使我们能够得出上文已提及的结论。

事实证明，在研究非适当 PAC 学习时，使用 DelleRose、Kozachinskiy、Rojas 和 Steifer [27] 引入的见证函数 这一概念非常有用。

定义 6（见证）。设 $\mathcal{C} \subseteq 2^{\mathbb{N}}$ 为一个概念类。若函数 $f\colon \mathbb{N}^{k+1} \to \{0,1\}^{k+1}$ 满足以下条件，则称其为 k–见证，或简称为 $\mathcal{C}$ 的见证函数：

$$f(x_1, ..., x_{k+1}) \neq (h(x_1), ..., h(x_{k+1}))$$

对于所有的 $x_1 < ... < x_{k+1} \in \mathbb{N}$ 和 $h \in \mathcal{C}$。

利用这一定义，我们可以将 VC 维数的概念刻画为

$$\mathrm{VCdim}(\mathcal{C}) = \min\{k \in \mathbb{N} : \mathcal{C} \text{ admits a } k\text{–witness}\},$$

约定 $\min(\varnothing) = \infty$。上述作者使用 k–见证的概念来定义有效 *VC* 维数

$$\mathrm{eVCdim}(\mathcal{C}) = \min\{k \in \mathbb{N} : \mathcal{C} \text{ admits a computable } k\text{–witness}\}.$$

借助这一概念，作者证明了统计学习基本定理的以下非统一可计算版本；Sterkenburg [45]，此前的研究为此奠定了基础，他隐含地证明了这一等价关系的一个方向。

定理 7（Delle Rose、Kozachinskiy、Rojas 和 Steifer，2023年）。概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$ 存在一个非适当可计算学习器 $A : \mathcal{S} \to 2^{\mathbb{N}}$（具有可计算的样本复杂度），当且仅当其有效 VC 维数有限。

我们从第 6 节开始，分析这一结果证明的统一计算内容，这将帮助我们确定 IPAC^+ 的复杂度。

最后，在第 7 节的结论中，我们回顾研究结果，并提出若干悬而未决的问题。

2. Preliminaries on Computable Analysis and Weihrauch Complexity

我们介绍一些可计算分析和 Weihrauch 复杂度中的概念，并请读者参阅 [17, 49] 以了解更多细节。我们采用基于表示的可计算分析方法，并回顾一下，空间 X 的表示是一个满射的部分映射 $\delta_X :\subseteq \mathbb{N}^{\mathbb{N}} \to X$。在这种情况下，$(X, \delta_X)$ 称为表示空间。若有两个表示空间 (X, δ_X) 和 (Y, δ_Y)，则部分函数 $F :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ 称为某个部分多值函数 $f :\subseteq X \rightrightarrows Y$ 的实现器，如果

$$\delta_Y F(p) \in f\delta_X(p)$$

对所有 $p \in \mathrm{dom}(f\delta_X)$ 均成立。在这种情况下，我们也写作 $F \vdash f$。表示空间上的任意部分多值函数 $f :\subseteq X \rightrightarrows Y$ 都称为问题；当且仅当一个问题具有可计算实现器时，才称其为可计算的。其他概念，例如连续性、有限心智变更可计算性或极限可计算性，也可以类似地针对问题定义。

一个可计算度量空间 (X, d, α) 是满足以下条件的空间：$d : X \times X \to \mathbb{R}$ 是 X 上的度量，并且存在一个稠密序列 $\alpha : \mathbb{N} \to X$，使得 $d \circ (\alpha \times \alpha) : \mathbb{N}^2 \to \mathbb{R}$ 是可计算的。这类空间由其柯西表示 δ_X 表示，该表示可定义为

$$\delta_X(p) = x :\iff \lim_{n\to\infty} \alpha p(n) = x \text{ and } (\forall j)(\forall i > j)\ d(\alpha p(i), \alpha p(j)) < 2^{-j}.$$

实数 $\mathbb{R}$ 表示为一个可计算度量空间，并配有有理数的标准枚举 $q : \mathbb{N} \to \mathbb{Q}$。有时我们也使用实数的单侧表示。例如，$\mathbb{R}_< = \mathbb{R} \cup \{\infty\}$ 表示为

$$\delta_{\overline{\mathbb{R}}_<}(p) := \sup_{n\in\mathbb{N}} q_{p(n)}$$

并且用 $\mathbb{R}_<$ 表示该空间在不包含无穷大的实数上的限制。空间 $\mathbb{R}_>$ 以对称的方式定义。对于每个表示空间 X，还有序列空间 $X^{\mathbb{N}}$ 以及其上的有限字空间 X^* 的规范表示，其中 X。

用 $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ 表示余自然数空间，其表示为 $\delta_{\mathbb{N}_\infty}(0^n 1p) := n$ ，其中 $n \in \mathbb{N}$ ，且 $p \in \mathbb{N}^{\mathbb{N}}$ 和 $\delta_{\mathbb{N}_\infty}(\widehat{0}) = \infty$。该空间与 $\mathbb{R}$的紧子空间 $\{0\} \cup \{2^{-n} : n \in \mathbb{N}\}$ 可计算同构。

对于一个可计算度量空间（X， d， α），我们用 $B(x, \varepsilon) := \{y \in X : d(x, y) < \varepsilon\}$表示以 $x \in X$ 为中心、半径为 $\varepsilon > 0$ 的开球，并通过 $B_{\langle n,\, k\rangle} := B(\alpha(n)\ q_k)$ 得到开有理球的一个编号，其中 $\langle n, k\rangle \in \mathbb{N}$ 表示自然数 n， $k \in \mathbb{N}$的康托配对。对于康托空间 $2^{\mathbb{N}}$，我们通常考虑有限字的一个双射可计算枚举 $\mathsf{w} : \mathbb{N} \to \{0,1\}^*$， $i \mapsto \mathsf{w}_i$。在这种情况下，我们也使用开球 $B_n := \mathsf{w}_n 2^{\mathbb{N}}$。

我们分别用 $\mathcal{O}(X)$ 和 $\mathcal{A}(X)$表示 X的开子集空间和闭子集空间。事实上，针对正信息、负信息和完整信息，我们需要三种不同版本的闭子集空间，分别记为 $\mathcal{A}_+(X)$、 $\mathcal{A}_-(X)$ 和 $\mathcal{A}(X)$。这些空间被定义为表示空间。我们将 $\mathrm{range}(p) - 1 := \{n \in \mathbb{N} : n + 1 \in \mathrm{range}(p)\}$ 写作 $p \in \mathbb{N}^{\mathbb{N}}$，以便允许将常零序列作为空集的名称。

定义 8（超空间表示）。设 X 是一个可计算度量空间。我们定义以下表示，其中 p， $q \in \mathbb{N}^{\mathbb{N}}$ 且闭 $A \subseteq X$：

(1) $\delta_{\mathcal{O}(X)}$ 由 $\delta_{\mathcal{O}(X)}(p)$ 表示 $\mathcal{O}(X) := \bigcup_{n \in \mathrm{range}(p)-1} B_n$。(2) $\delta_{\mathcal{A}_-}$由 $\delta_{\mathcal{A}_-}(X)(p)$ 表示 $\mathcal{A}_-(X) := X \setminus \delta_{\mathcal{O}(X)}(p)$。(3) $\delta_{\mathcal{A}_+}$由 $\delta_{\mathcal{A}_+}(X)(p)$ 表示 $\mathcal{A}_+(X) = A :\iff \mathrm{range}(p) - 1 = \{n \in \mathbb{N} : B_n \cap A \neq \varnothing\}$。(4) $\delta_{\mathcal{A}(X)}$ 由 $\delta_{\mathcal{A}(X)}\langle p, q\rangle = A$ 表示 $\mathcal{A}(X)$: $\iff \delta_{\mathcal{A}_+}(X)(p) = \delta_{\mathcal{A}_-}(X)(q) = A$。

这里 $\langle . \rangle$ 表示贝尔空间 $\mathbb{N}^{\mathbb{N}}$ 上的某个标准配对函数（我们用这个记号表示有序对，也表示序列的配对）。上述超空间的表示在可计算分析中被广泛使用 [20, 49, 17]。

表示空间 (X, δ) 中的点 $x \in X$ 称为可计算的，如果它有一个可计算名称，也就是说，存在一个可计算名称 $p \in \mathbb{N}^{\mathbb{N}}$，使得 $\delta(p) = x$。表示 δ 还在 X上诱导出一种拓扑，即其终拓扑；这是 X 上使 δ 连续的最大拓扑。对于引入的超空间，我们得到以下结果（见 [20]）：

(1) 空间 $\mathcal{A}_+(X)$ 中的可计算点称为可计算可枚举闭集。空间 $\mathcal{A}_+(X)$上的诱导拓扑是下 *Fell* 拓扑。(2) 空间 $\mathcal{A}_-(X)$ 中的可计算点称为余递归可枚举闭。空间 $\mathcal{A}_-(X)$上的诱导拓扑是上 *Fell* 拓扑。(3) 空间 $\mathcal{A}(X)$中的可计算点称为可计算闭。空间 $\mathcal{A}(X)$上的诱导拓扑是 *Fell* 拓扑。(4) 空间 $\mathcal{O}(X)$ 中的可计算点称为可计算可枚举开。空间 $\mathcal{O}(X)$上的诱导拓扑是 *Scott* 拓扑。

这种关于可计算集的术语推广了 c.e.、co-c.e. 和可计算集 $A \subseteq \mathbb{N}$ 的经典概念，因为它们分别正是 $\mathcal{A}_+(\mathbb{N})$、 $\mathcal{A}_-(\mathbb{N})$ 和 $\mathcal{A}(\mathbb{N})$ 中的可计算点。由于我们主要会在康托空间 $X = 2^{\mathbb{N}}$ 中使用这些记号，在这一特殊情形下，我们也简写为 $\mathcal{A}_+$、 $\mathcal{A}_-$ 和 $\mathcal{A}$。

有时，使用空间 $\mathcal{A}_+(X)$ 的另一种刻画对我们很有用。具体而言，如果 X 是一个完备度量空间，那么我们也可以通过枚举稠密子集 [20]来编码闭子集 $A \subseteq X$。我们用 A 表示子集 $A \subseteq X$的闭包。

命题 9（闭子集的序列表示）。 设 X 是一个可计算度量空间。那么

$$S : X^{\mathbb{N}} \to \mathcal{A}_+(X), (x_n)_{n\in\mathbb{N}} \mapsto \overline{\{x_n : n \in \mathbb{N}\}}$$

是可计算的；此外，如果 X 是完备的，那么 S 存在一个多值可计算逆，其定义域为所有非空的 $A \in \mathcal{A}_+(X)$。

众所周知，具有完整信息的康托空间的闭子集可以与修剪树 [20] 等同起来。我们用 Tr 表示由其特征函数表示的二叉树空间。

命题 10（修剪树）。 映射

$$T : \mathcal{A} \to \mathrm{Tr}, A \mapsto T_A := \{w \in \{0,1\}^* : w2^{\mathbb{N}} \cap A \neq \varnothing\}$$

是一个可计算嵌入，即它是可计算的、单射，并且具有可计算的部分逆。

我们注意到，如果 $[T] := \{p \in 2^{\mathbb{N}} : (\forall n \in \mathbb{N})\ p|_n \in T\}$ 表示树 $T \in \mathrm{Tr}$ 的无限路径集合，那么对所有 $A \in \mathcal{A}$，$[T_A] = A$。

在此我们简要讨论，与考虑假设类 $\mathcal{H} \in \mathcal{A}_+$ 相对应的有效性概念，和文献中使用的其他概念有何关联。Agarwal 等人 [2] 引入了所谓的 RER 假设类 $\mathcal{H}_e = \{\varphi_i : i \in \mathrm{W}_e\}$，其中$\varphi$ 是部分可计算函数 $\varphi_i :\subseteq \mathbb{N} \to \mathbb{N}$ 的标准 G¨odel 编号，而 $\mathrm{W}_e := \mathrm{dom}(\varphi_e)$ 是可计算可枚举集$W_e \subseteq \mathbb{N}$的导出标准编号。这里，仅当所有满足 $i \in \mathrm{W}_e$的函数 φ_i 均为全函数时，才定义 $\mathcal{H}_e$。DR 假设类的相关概念则通过指标的可判定集而非枚举来定义。根据命题 9，我们的可计算可枚举闭集假设类 $\mathcal{H}$ 恰好是 RER 类的闭包。然而，谈到一致可计算性时，使用 G¨odel 编号就不那么自然了，因为函数 $f :\subseteq \mathbb{N} \to \mathcal{A}_+, e \mapsto \mathcal{H}_e$是可计算的，但它没有可计算左逆，即使将定义域限制在单元素集合 $\mathcal{H}_e$上也不行。这是因为，为给定的可计算序列寻找 G¨odel 数本身就是一个非平凡的计算问题（其 Weihrauch 复杂度已在 [8]中得到研究）。换言之，从统一视角考察 PAC 学习时，使用 G¨odel 数会人为地模糊其复杂度，这正是我们选择更适合这一目的的方法的原因。Akbari 和 Harrison-Trainor [3] 已经指出，在 PAC 学习的语境下使用 G¨odel 数可能不够自然；他们提出的解决方案是考虑相对化 RER 类$\mathcal{H}_{\langle p,e\rangle} = \{\varphi_i^p : i \in \mathrm{W}_e^p\}$，其中所有内容都相对于某个预言机 $p \in \mathbb{N}^{\mathbb{N}}$进行相对化。同样，只有当类中仅包含全函数时，才定义 $\mathcal{H}_{\langle p,e\rangle}$。虽然在[3] 中，这些概念是在一个锥上考虑的，但我们基本上直接研究这类的闭包。我们将把这一结论表述得更精确；它是命题 9 的推论。

推论 11（相对化 RER 类）。 函数 $F :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathcal{A}_+, \langle p,\ e\rangle \to \mathcal{H}_{\langle p,\ e\rangle}$是可计算的、满射的，并且具有可计算左逆。

如果需要，本推论可以帮助读者将我们关于空间 $\mathcal{A}_+$ 的结果转换到 RER 类的框架中。在我们的许多构造中，这些类已经是闭的相对化 RER 类（在这些情况下无需再取闭包）。

现在我们开始介绍 Weihrauch 复杂度中的概念（综述见 [15] ）。这里，$\mathrm{id} : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ 表示贝尔空间上的恒等映射。

定义 12（Weihrauch 可约性）。设 $f :\subseteq X \rightrightarrows Y$和 $g :\subseteq Z \rightrightarrows W$为问题。我们称

(1) 如果存在可计算函数H，$K :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$，使得只要 $G \vdash g$成立，就有 $H\langle$恒等映射 $GK\rangle \vdash f$，则称 f可 Weihrauch 归约到 g，记作 $f \leq_{\mathrm{W}} g$。(2) 如果存在可计算的H，$K :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$，使得只要 $G \vdash g$成立，就有 $HGK \vdash f$，则称 f强 Weihrauch 可约到 g，记作 $f \leq_{\mathrm{sW}} g$。

我们用 $\equiv_{\mathrm{W}}$ 和 $\equiv_{\mathrm{sW}}$ 表示分别由预序$\leq_{\mathrm{W}}$ 和 $\leq_{\mathrm{sW}}$诱导的等价关系。相应的等价类分别称为*Weihrauch*度和强 *Weihrauch* 度。此外，还有一种纯拓扑版本的（强）Weihrauch 可约性，分别记作 $\leq^*_{\mathrm{W}}$ 和 $\leq^*_{\mathrm{sW}}$。在这种情况下，只要求 H，K连续。我们的所有结果对于连续 Weihrauch 可归约性同样成立。图 3 展示了这种可约性关系。

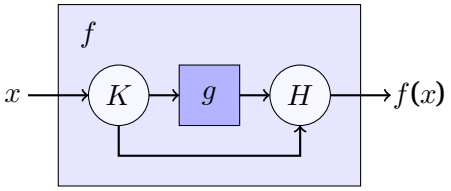

图 3：Weihrauch 可约性 $f \leq_{\mathrm{W}} g$。

一类可用于校准其他问题的重要问题是选择问题。可计算度量空间 X 的选择问题 C_X是指在负表示闭集 $A \in \mathcal{A}_-(X)$中寻找一个点的问题。

定义 13（选择问题）。可计算度量空间 X 的选择问题定义为

$$\mathsf{C}_X :\subseteq \mathcal{A}_-(X) \rightrightarrows X, A \mapsto A$$

其定义域为 $\mathrm{dom}(\mathsf{C}_X) := \{A \in \mathcal{A}_-(X) : A \neq \varnothing\}$。

我们将使用选择问题 C_2, $\mathsf{C}_{\mathbb{N}}$、$\mathsf{C}_{2^{\mathbb{N}}}$ 和 $\mathsf{C}_{\mathbb{R}}$。对于双点空间$2 = \{0, 1\}$ 的选择也称为 LLPO，即较弱的有限全知原理。有限全知原理LPO：$\mathbb{N}^{\mathbb{N}} \to \{0, 1\}$ 是单元素集$\{\widehat{0}\}$ 的特征函数，其对应的序列为常零序列 $\widehat{0}$。一般而言，$\widehat{n} \in \mathbb{N}^{\mathbb{N}}$ 表示取值为 $n \in \mathbb{N}$的常值序列。

我们定义若干后续将用到的基准问题。我们通常将数 $n \in \mathbb{N}$与 $X = \{0, ..., n - 1\}$等同。以下定义中的极限是相对于 $\mathbb{N}^{\mathbb{N}}$上的积拓扑而言的。我们使用相对 PA 度的关系$\ll$，即对于 p，$q \in \mathbb{N}^{\mathbb{N}}$ 表示每个无限二进制 p可计算树都有一条 q可计算路径。在这种情况下，我们称 q 具有相对于 p的*PA* 度。这一术语源于这样一个事实：对于可计算的 p，这对应于寻找皮亚诺算术的一个完备扩张的复杂度。

定义14（问题）。我们考虑以下适用于所有 $X \in \mathbb{N} \cup \{\mathbb{N}\}$的问题：
(1) $\lim :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$，$\langle p_0, p_1, p_2, ...\rangle \mapsto \lim_{i\to\infty} p_i$称为极限问题。

(2) $\lim_\Delta :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ 是 lim 在最终常值序列（p_i）上的限制。(3) WKL $:\subseteq \mathrm{Tr} \rightrightarrows 2^{\mathbb{N}}$, $T \mapsto [T]$ 称为弱柯尼希引理。(4) WWKL 是 WKL 在具有 $[T]$ 正测度的树 T 上的限制，称为弱弱柯尼希引理。(5) SORT : $2^{\mathbb{N}} \to 2^{\mathbb{N}}, p \mapsto \widehat{0}$ 若 $p \in 2^{\mathbb{N}}$ 包含无限多个零，而 $p \mapsto 0^n\widehat{1}$ 若 p恰好包含 n 个零，则称为排序问题。(6) sup : $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}_\infty, p \mapsto \sup\{p(i) : i \in \mathbb{N}\}$ 称为上确界问题。(7) max $:\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}, p \mapsto \max\{p(i) : i \in \mathbb{N}\}$ 称为最大值问题。(8) B $:\subseteq \mathbb{N}^{\mathbb{N}} \rightrightarrows \mathbb{N}$, $p \mapsto \{n \in \mathbb{N} : (\forall k \in \mathbb{N})\ p(k) \leq n\}$ 称为有界性问题。(9) $\mathrm{ACC}_X :\subseteq \mathcal{A}_-(X) \rightrightarrows X$ 是 C_X 在集合 $A \in \mathcal{A}_-(X)$ 上的限制，其中 $|X \setminus A| \leq 1$，称为全选或余唯一选择问题。(10) DNC_X : $\mathbb{N}^{\mathbb{N}} \to X^{\mathbb{N}}, p \mapsto \{q : (\forall n)\ \varphi_n^p(n) \neq q(n)\}$ 称为对角非可计算问题。(11) PA : $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$, $p \mapsto \{q : p \ll q\}$ 是皮亚诺算术度问题。所有问题均取其自然定义域，例如 lim 定义于收敛序列，WKL 定义于无限树，max以及 B 定义于有界序列等。

这基本上可由 [16, 18] 中的结果推出，给定的问题按图2所示排序（其中箭头 $f \leftarrow g$ 表示 $f \leq_{\mathrm{sW}} g$）。

我们还需要若干在 Weihrauch 复杂度中常用的问题算子。表示空间的 跳跃 (X, δ) 是表示空间 $(X, \delta \circ \lim)$，我们也将其记作 X'。

定义 15（问题上的运算）. 设 $f :\subseteq X \rightrightarrows Y$, $g :\subseteq Z \rightrightarrows W$ 是问题。我们定义以下运算:

(1) 并行化 $\widehat{f} :\subseteq X^{\mathbb{N}} \rightrightarrows Y^{\mathbb{N}}, (x_n)_{n\in\mathbb{N}} \mapsto \mathrm{X}_{n\in\mathbb{N}} f(x_n)$。(2) 乘积 $f \times g$: $\subseteq X \times Z \rightrightarrows Y \times W, (x, z) \mapsto f(x) \times g(z)$。(3) 复合乘积 $f * g :=$ 最大值 $\leq_{\mathrm{W}} \{F \circ G : F \leq_{\mathrm{W}} f$ 和 $G \leq_{\mathrm{W}} g\}$。(4) 跳跃 $f' :\subseteq X' \rightrightarrows Y, x \mapsto f(x)$。

我们注意到，除跳跃只关于强 Weihrauch 可约性具有单调性外，所有这些运算关于 Weihrauch 可约性都是单调的，因此都可以扩展到 Weihrauch 度 [14]。直观地说，复合乘积刻画了先使用 g 、再（经过某种中间计算后）使用 f所能得到的最强问题。复合乘积由 [14]引入。[19]证明了相应的最大值总是存在的。严格来说，$f * g$ 只定义了一个普通的 Weihrauch 度，但该度有一个特定实例，它关于强 Weihrauch 可约性是最大的，并且还使我们能够表述涉及复合乘积的强 Weihrauch 归约（详情见 [19, 15]）。

我们将需要弱 K″önig 引理 WKL 的度的若干刻画，这些刻画实质上已在 [12, 9, 16]中得到证明。我们回顾一下，$\mathcal{L} = (2^{\mathbb{N}})^{\mathcal{S}}$ 表示一类与康托空间 $2^{\mathbb{N}}$ 在可计算同构意义下相同的学习器，其同构通过对$\mathbb{N} \times \mathcal{S}$的某种标准编码实现。

Pro　　位置 16（弱 K″onig 引理）。 $\mathrm{WKL} \equiv_{\mathrm{sW}} \mathsf{C}_{2^{\mathbb{N}}} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathcal{L}} \equiv_{\mathrm{sW}} \mathsf{C}_2 \equiv_{\mathrm{sW}} \mathrm{DNC}_2$.

自然数上的选择 $\mathsf{C}_{\mathbb{N}}$ 可按如下方式刻画 [11, 9, 15].

命题 17（自然数上的选择）。 $\mathrm{B} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} \equiv_{\mathrm{sW}}$ 最大值 $\equiv_{\mathrm{W}} \lim_\Delta \equiv_{\mathrm{sW}}$ 恒等映射 $\times \mathsf{C}_{\mathbb{N}}$和 $\mathrm{B} <_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}} <_{\mathrm{sW}} \lim_\Delta$。

欧几里得空间上的选择 $\mathsf{C}_{\mathbb{R}}$ 可以写成 $\mathsf{C}_{\mathbb{N}}$ 和 $\mathsf{C}_{2^{\mathbb{N}}}$的乘积的不同形式，即 $\mathsf{C}_{\mathbb{R}} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}} \times \mathsf{C}_{2^{\mathbb{N}}} \equiv_{\mathrm{sW}} \mathsf{C}_{2^{\mathbb{N}}} * \mathsf{C}_{\mathbb{N}}$，这一点在 [9, 14] 中已基本得到证明。尽管 $\mathsf{B} <_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}}$，我们仍得到 $\mathsf{B} \times \mathsf{C}_{2^{\mathbb{N}}} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}} \times \mathsf{C}_{2^{\mathbb{N}}}$，因为输入可以通过 $\mathsf{C}_{2^{\mathbb{N}}}$ 复制（如 $\mathrm{id} \times \mathsf{C}_{2^{\mathbb{N}}} \leq_{\mathrm{sW}} \mathsf{C}_{2^{\mathbb{N}}}$）。

命题 18（欧几里得空间上的选择）。 $\mathsf{C}_{\mathbb{R}} \equiv_{\mathrm{sW}} \mathsf{C}_{2^{\mathbb{N}}} * \mathsf{C}_{\mathbb{N}} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}} \times \mathsf{C}_{2^{\mathbb{N}}} \equiv_{\mathrm{sW}} \mathsf{B} \times \mathsf{C}_{2^{\mathbb{N}}}$。

以下刻画广为人知，且已在 [11, 6]中得到证明。

命题 19（极限）。 $\lim \equiv_{\mathrm{sW}} \mathsf{LPO} \equiv_{\mathrm{sW}} \lim * \mathsf{C}_{\mathbb{N}}$。

以下刻画最初由 Higuchi 和 Kihara 证明 [31, 命题 7。9]（另见 [16, 定理 5。2]）。我们注意到，ACC_n 也称为 LLPO_n。

命题20（对角不可计算性）。 $\mathrm{DNC}_X \equiv_{\mathrm{sW}}$反链选择$_X$ 适用于所有至少包含两个元素的 $X \in \mathbb{N} \cup \{\mathbb{N}\}$。

我们引入 DNC_n 的一个新变体，它与非适当 PAC 学习的分析相关，并将其非正式地定义为 $\mathrm{DNC}_* := \sqcap_{k\geq 2} \mathrm{DNC}_k$。其思路是，它对应于逻辑问题 $(\exists k \geq 2)\ \mathrm{DNC}_k$。由于我们没有一般性地定义问题序列上的运算 $\sqcap$[2], ，因此在此临时定义问题 DNC_*。

定义 21（有限DNC）。 我们将 $\mathrm{DNC}_* :\subseteq \mathsf{X}_{k=2}^{\infty} \mathcal{A}_{-}(k)^{\mathbb{N}} \rightrightarrows \mathbb{N}^{\mathbb{N}}$ 定义为

$$\mathrm{DNC}_*(((A_{n,k})_{n\in\mathbb{N}})_{k\geq 2}) := \{\langle k, p\rangle : (\forall n)\ p(n) \in A_{n,k}\}$$

其中 $|k \setminus A_{n,\ k}| \leq 1$ 对所有 $n,\ k$ 都成立。我们称之为有限*DNC*问题。

换言之，DNC_* 的实例是序列 $(B_k)_{k\geq 2}$，其中每个 $B_k = (A_{n,\ k})_{n\in\mathbb{N}}$ 都是反链选择$_k$ 的一个实例，而解是满足 $p \in$ 反链选择$_k(B_k)$ 的一对 $(k,\ p)$。我们将使用问题 DNC_* 来区分非适当 PAC 学习与 $\mathrm{DNC}_{\mathbb{N}}$、皮亚诺算术和弱弱柯尼希引理。这之所以可行，是因为我们有以下结果。

命题22（有限DNC）。 $\mathrm{PA} <_{\mathrm{W}} \mathrm{DNC}_*$，$\mathrm{DNC}_{\mathbb{N}} <_{\mathrm{W}} \mathrm{DNC}_*$，且对所有 $n \in \mathbb{N}$，$\mathrm{DNC}_* \not\leq_{\mathrm{W}} \mathrm{WWKL}^{(n)}$。

证明。由命题20可得 $\mathrm{DNC}_{\mathbb{N}} \equiv_{\mathrm{sW}} \mathrm{ACC}_{\mathbb{N}} \leq_{\mathrm{sW}} \mathrm{DNC}_*$。后者显然成立，因为 $f :\ \mathcal{A}_{-}(\mathbb{N}) \times \mathbb{N} \to \mathcal{A}_{-}(k)$，$(A, k) \mapsto A \cap k$ 可在 k 中一致地计算。归约 $\mathrm{PA} \leq_{\mathrm{sW}} \mathrm{DNC}_k = \mathrm{ACC}_k$ 众所周知，也容易看出，因为 $p \ll q$ 成立，当且仅当每个q 都计算一个函数 $f : \mathbb{N} \to k$，该函数相对于 p 是对角不可计算的。这对 k 一致成立，因此我们得到 $\mathrm{PA} \leq_{\mathrm{sW}} \mathrm{DNC}_*$。由推论7[16, 可知，对所有 $n \in \mathbb{N}$，$\mathrm{PA} \not\leq_{\mathrm{W}} \mathrm{WWKL}^{(n)}$。因此，我们得到对所有 $n \in \mathbb{N}$，$\mathrm{DNC}_* \not\leq_{\mathrm{W}} \mathrm{WWKL}^{(n)}$。由推论7 [16, 可知，$\mathrm{DNC}_{\mathbb{N}} \leq_{\mathrm{sW}} \mathrm{WWKL}$。因此，$\mathrm{DNC}_* \not\leq_{\mathrm{W}} \mathrm{DNC}_{\mathbb{N}}$。另一方面，$\mathrm{DNC}_{\mathbb{N}} \not\leq_{\mathrm{W}} \mathrm{PA}$ 也已为人所知（甚至 $\mathrm{ACC}_{\mathbb{N}} \leq_{\mathrm{W}} \mathrm{PA}$ 都不成立，因为 PA 在 [16] 的意义下稠密实现），因此可推出 $\mathrm{DNC}_* \not\leq_{\mathrm{W}} \mathrm{PA}$。 □

这一结果表明，问题 DNC_* 与图2中展示的其他著名问题有关。

[2]对于问题序列 $(f_k)_{k\in\mathbb{N}}$，我们可以定义 $\sqcap_{k\in\mathbb{N}} f_k$（见 [13]），使其成为关于连续 Weihrauch 可约性的 f_i 的下确界。然而，对于可计算 Weihrauch 可归约性而言，自然问题不存在非平凡的下确界 [32, 推论 3.18]。

众所周知，极限映射也等价于将康托空间中某个闭子集的负信息转换为正信息，反之亦然。我们将需要这一结果的以下方向 [10, 命题 4.2、4.5]。

命题 23（闭子集信息的转换）。 两个恒等函数 $\mathrm{id}:\mathcal{A}_+\to\mathcal{A}_-, A\mapsto A$ 和 $\mathrm{id}:\mathcal{A}_-\to\mathcal{A}_+$, $A\mapsto A$ 都强 Weihrauch 可约于 lim。

在推论 51 中，我们将加强这一结果。最后提及，$\mathsf{C}_{\mathbb{N}}$、$\mathsf{C}_{2^{\mathbb{N}}}\equiv_{\mathrm{sW}}$、WKL、弱弱柯尼希引理和 lim 刻画了一些自然问题类，这些问题类此前均已被独立研究过 [9, 13]。

定理 24（可计算问题类）。 设 f 是一个问题。那么：
（1） $f\leq_{\mathrm{W}}\mathsf{C}_{\mathbb{N}}\iff f$ 可通过有限次思维转换计算。（2） $f\leq_{\mathrm{W}}$ WKL $\iff f$ 可非确定性计算。（3） $f\leq_{\mathrm{W}}$弱弱柯尼希引理 $\iff f$ 可拉斯维加斯计算。（4） $f\leq_{\mathrm{W}}\lim\iff f$ 极限可计算。

这里我们不单独定义所涉及的可计算性概念，但定理 24 中的刻画也可以视为一种定义。我们只指出，有限心智变化计算和极限计算常用于算法学习理论的语境中 [29]。

3. VC Dimension

VC 维数的概念由 Vapnik 和 Chervonenkis [47] 提出。由于它有许多应用，甚至超出了机器学习理论的范畴，因此对其复杂度进行精确分类本身就具有独立的研究价值。在我们对 PAC 学习的研究中，主要需要关于有限 VC 维数集合的结果。

我们将证明，利用完整信息计算闭子集$\mathcal{C}\in\mathcal{A}$ 的 VC 维数，或利用正信息计算闭子集 $\mathcal{C}\in\mathcal{A}_+$ 的 VC 维数，其复杂度与二进制排序问题相同。对于利用负信息的闭子集 $\mathcal{C}\in\mathcal{A}_-$，该问题等价于二进制排序问题的跳跃。

定义 25（VC 维数问题）。
（1）用 VC 维数 表示： $\mathcal{A}\to\mathbb{N}_\infty$ 我们将确定闭集的 VC 维数这一问题称为 VC 维数问题，正如引言中所定义的那样。（2）用 VC 维数$_{<\infty}$: $\subseteq\mathcal{A}\to\mathbb{N}$ 我们将确定限制在有限 VC 维数闭集上的 VC 维数这一问题称为 VC 维数问题。（3）用有界 VC 维数: $\subseteq\mathcal{A}\rightrightarrows\mathbb{N}$, $\mathcal{C}\mapsto\{d\in\mathbb{N}:\text{VC 维数}(\mathcal{C})\leq d\}$ 我们将确定有限 VC 维数闭集的 VC 维数上界这一问题称为有界 VC 维数问题。
如果这些问题分别定义在 $\mathcal{A}_+$ 或 $\mathcal{A}_-$ 上，而不是定义在 $\mathcal{A}$上，我们就在每个问题上添加上标 $+$ 或 $-$。

现在我们将二进制排序问题 SORT 用作基准问题。该问题最初由 Neumann 和 Pauly [39], 提出并研究，尽管 Weihrauch [48] 早已以 EC_1 之名研究过同一个普通 Weihrauch 等价类（等价性见 [7] ）。我们用 SORT$_{<\infty}$ 表示限制在仅含有限个零的二进制序列上的排序问题。作为准备，我们证明二进制排序问题与上确界问题实际上是等价的。

引理 26. 排序 $\equiv_{\mathrm{sW}}$ 上确界且排序$_{<\infty} \equiv_{\mathrm{sW}}$ 最大值 $\equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}}$。

证明。 我们证明第一个等价关系。

排序 $\leq_{\mathrm{sW}}$ 上确界。给定一个序列 $p \in 2^{\mathbb{N}}$，通过设定$q(0) = 1 - p(0)$，生成一个序列 $q \in \mathbb{N}^{\mathbb{N}}$。对于 $i \in \mathbb{N}$，若 $p(i+1) = 1$，则令 $q(i+1) = q(i)$；若 $q(i+1) = q(i) + 1$，则令 $p(i+1) = 0$。这样，$q(i)$ 就是 $p|_{i+1}$ 中零的个数。然后，对 q 应用上确界，得到零个数的一个$\delta_{\mathbb{N}_\infty}$–名称，即 $\mathsf{SORT}(p)$。

上确界 $\leq_{\mathrm{sW}}$ 排序。给定一个序列 $p \in \mathbb{N}^{\mathbb{N}}$，先写入 $p(0)$ 个零，并设 $\sup := p(0)$。然后，对于每个读到的符号 $p(i)$ $(i \geq 1)$，写入一个 1；若 $p(i) > \sup$，则写入 $p(i) - \sup$ 个零，并随后更新 $\sup := p(i)$。当 p 无界时，所得序列 q 含有无穷多个零；否则，含有 $\sup_i p(i)$ 个零。因此，对 q 应用排序，即可生成 $\mathsf{SORT}(p)$ 的一个名称。

排序$_{<\infty} \equiv_{\mathrm{sW}}$ 最大值的等价性分别由上述相同的构造在仅含有限个零的序列和有界序列上的限制得到。只需使用以下额外事实：$\mathsf{id} : \mathbb{N} \hookrightarrow \mathbb{N}_\infty$ 是一个可计算嵌入。最大值 $\equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}}$ 的等价性是众所周知的，见 [15, 定理 11.7.13]。 □

现在我们可以陈述关于 VC 维数的主要结果。用 $w \cdot p$ 表示一个字 $w \in \mathbb{N}^*$ 与一个序列 $p \in \mathbb{N}^{\mathbb{N}}$ （或另一个字 $p \in \mathbb{N}^*$）的连接。用 $h|_k$表示函数 $h :\subseteq \mathbb{N} \to \mathbb{N}$ 到 $k = \{0, ..., k-1\}$的限制。

定理27 （VC 维数）。 VC 维数 $\equiv_{\mathrm{sW}}$ VC 维数$^+ \equiv_{\mathrm{sW}}$ 排序。

证明。 根据引理 26，只需证明

$$\mathsf{SORT} \leq_{\mathrm{sW}} \mathrm{VCdim} \leq_{\mathrm{sW}} \mathrm{VCdim}^+ \leq_{\mathrm{sW}} \sup .$$

第二个归约显然成立。我们证明另外两个归约。

(1) VC 维数$^+ \leq_{\mathrm{sW}}$ 上确界。给定 $\mathcal{C} \in \mathcal{A}_+$，根据命题 9，我们可以假设存在一个序列 $(h_n)_{n\in\mathbb{N}}$ 属于 $2^{\mathbb{N}}$，使得 $\mathcal{C} = \{h_n : n \in \mathbb{N}\}$。我们定义 $\mathcal{C}_k := \{h_n|_k \cdot \widehat{0} : n \leq k\}$。于是得到

$$\mathrm{VCdim}(\mathcal{C}) = \sup_{k\in\mathbb{N}} \mathrm{VCdim}(\mathcal{C}_k).$$

这是成立的，因为任何被 $\mathcal{C}$ 打散的集合 $A \subseteq \mathbb{N}$，必定也被某个长度为 $\max(A) + 1$ 的 h_n 的某个前缀打散。$\mathcal{C}_k$ 的 VC 维数很容易计算，因为只需考虑长度为 k 的有限前缀 $k+1$。

(2) 排序 $\leq_{\mathrm{sW}}$ VC 维数。给定一个字符串 $p \in 2^{\mathbb{N}}$，定义如下子集

$$\mathcal{C} = \{h : \mathbb{N} \to \{0,1\} : (\forall i \in \mathbb{N})(p(i) = 1 \implies h(i) = 0)\} \in \mathcal{A}.$$

注意，$\mathcal{C}$ 的 VC 维数是函数 h 在 $\mathcal{C}$ 中可以取 0 和 1 两个值的指标数量，根据定义，这恰好等于 p 中零的数量。

要从 p 计算 $\mathcal{C} \in \mathcal{A}$，我们从 $i = 0$ 开始，遍历所有自然数 $i \in \mathbb{N}$。对于每个 i，长度为 $i+1$ 的可扩展前缀是满足约束 $a_k = 0$ 的如下形式：$a_0 \dots a_i \in \{0,1\}^{i+1}$，其中对所有满足 $p(k) = 1$的 $k \leq i$ 均成立。读取 $p|_{i+1}$，识别这些前缀并将相关的球写入正信息，然后将长度为 $i+1$ 的所有其他前缀写入负信息。

然后，应用 VC 维数来求出 $\mathcal{C}$ 在 $\mathbb{N}_\infty$中的 VC 维数。这是 p中零的数量的一个 $\delta_{\mathbb{N}_\infty}$–名称，恰好就是 $\mathsf{SORT}(p)$。 □

基本相同的证明也表明

$$\mathsf{SORT}_{<\infty} \leq_{\mathrm{sW}} \mathrm{VCdim}_{<\infty} \leq_{\mathrm{sW}} \mathrm{VCdim}^{+}_{<\infty} \leq_{\mathrm{sW}} \max .$$

因此，借助引理 26，我们得到以下结果。

推论 28. VC 维数$_{<\infty} \equiv_{\mathrm{sW}}$ VC 维数$^{+}_{<\infty} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{N}}$ 且有界 VC 维数 $\equiv_{\mathrm{sW}}$ 有界 VC 维数$^{+} \equiv_{\mathrm{sW}} \mathrm{B}$。

第二条等价链可由 VC 维数$_{<\infty} \equiv_{\mathrm{sW}}$ VC 维数$^{+}_{<\infty} \equiv_{\mathrm{sW}}$ 最大值推出，只需使用相同的归约函数，因为 B 只计算最大值的一个上界。

我们还可以从定理27的证明得出，VC 维数映射作为实值函数是下半可计算的。下面的推论将对此作更形式化的表述。

推论 29. VC 维数，作为类型为 VC 维数：$\mathcal{A}_{+} \to \mathbb{R}_{<}$ 的映射，是可计算的。这对于用 $\mathcal{A}$ 代替 $\mathcal{A}_{+}$ 的情形也成立。

值得注意的是，加入负信息并不会降低计算 VC 维数的 Weihrauch 复杂度。直观而言，负信息无关紧要，可以这样理解：这些信息能帮助我们知道，在某个深度范围内，所考虑的假设类没有打散某个子集。但我们无法据此对 VC 维数得出任何结论，因为同样大小的子集完全可能在该类更深的层次上被打散。

现在我们来说明，如果只包含负信息，情况会如何。与先前的结果一致，在使用这些信息之前，先将其转换为正信息，才是利用它们的最佳方式。

定理 30（带负信息的 VC 维数）。$\mathrm{VCdim}^{-} \equiv_{\mathrm{sW}} \mathrm{VCdim}' \equiv_{\mathrm{sW}}$ 排序′。

Proof. $\mathrm{VCdim}^{-} \leq_{\mathrm{sW}}$ VCd $\mathrm{im}' \leq_{\mathrm{sW}}$ SO $^{\mathrm{RT}'}$。Thi s follows from Theorem 27 with Proposition 23。

排序′ $\leq_{\mathrm{sW}}$ VC 维数$^{-}$。设 $(p_n) \in (\mathbb{N}^{\mathbb{N}})^{\mathbb{N}}$ 是一个收敛到 $p \in 2^{\mathbb{N}}$ 的序列，并令$\ell \in \mathbb{N}_{\infty}$ 表示 p 中零的个数。我们要计算 VC 维数为 ℓ 的闭集 $\mathcal{C}$ 的负信息。我们进行构造，使得对于任意 $h \in \mathcal{C}$ 和 $i \in \mathbb{N}$，都有 $h(i) = 0$，但某些 i 例外，我们将其称为开放路径；在这些路径中，$h(i)$ 可以取 0 和 1 两个值。我们希望开放路径的数量为 ℓ，从而得到 VC 维数 $(\mathcal{C}) = \ell$。

为此，我们跟踪一个名为 open paths 的变量，其中包含 $\mathbb{N} \times \mathbb{N}$ 的元素列表。对于列表中的 $(i,\ k) \in \mathbb{N} \times \mathbb{N}$，$i$ 表示一个我们保持开放的索引，而 k 表示 (p_n) 中与 i 关联的零的位置。我们还跟踪一个变量 N，初始值为 0，它是在描述 $\mathcal{C}$ 时我们处理过的最新索引。

现在，构造 $q \in \mathbb{N}^{\mathbb{N}}$ 使得

$$\mathcal{C} = 2^{\mathbb{N}} \setminus \bigcup_{m+1 \in \mathrm{range}(q)} \mathsf{w}_m 2^{\mathbb{N}} = \delta_{\mathcal{A}_-}(q), \tag{3.1}$$

其中 $\mathsf{w} : \mathbb{N} \to \{0,1\}^*$ 是我们的可计算双射编号。我们重复以下步骤：
(1) 读取$p_{\mathsf{N}|_{\mathsf{N}+1}}$。(2) 对于任意满足 $p_{\mathsf{N}}(k) = 0$ 的 $k \leq$ 自然数N，且不存在使 (i,k) 属于开放路径的 $i \in \mathbb{N}$，为 k添加一条开放路径，也就是将 (N k) 添加到开放路径中,- 然后递增 N。(3) 对于任意满足 $p_{\mathsf{N}}(k) \neq 0$ 的 $k \leq$ 自然数N,$^{-}$ 若存在某个 $i \in \mathbb{N}$ 使 (i,k) 属于开放路径（即，k 被“错误地保持开放”），则将 (i,k) 从开放路径中移除。(4) 将 $\bar{\mathsf{N}}$ 加一，以关闭另一条路径。

(5) 有效地关闭截至此时尚未保持开放的所有路径，也就是将所有$m \in \mathbb{N}$写入 q，使得 $|\mathsf{w}_{m-1}| \leq$ 自然数 $\mathsf{N}+1$ **并且**存在 $i \leq$ 自然数N，满足 $\mathsf{w}_{m-1}(i) = 1$，但不存在满足条件的 k，使得 (i,k) 属于开放路径。

对于任意 $k \in \mathbb{N}$，满足条件的 $i \in \mathbb{N}$ 要么不存在，要么恰有一个，使得 $(i,\ k)$ 无限期地保留在开放路径中；当且仅当 $p(k) = 0$ 时，这样的 $i \in \mathbb{N}$ 存在。事实上，若 $p(k) = 0$ 成立，则存在一个最小的 $N \in \mathbb{N}$，使得对所有 $n \geq N$，$p_n(k) = 0$ 均成立；当读到 $p_N(k)$ 时，k 会被加入，此后再也不会被移除。若 $p(k) = 1$，则加入开放路径中的任何 k 最终都会被移除。

借助自然数，这一构造还确保不会有任何 i 同时以不同的 k 作为配对 $(i,\ k)$ 留在开放路径中。因此，恰有 $\ell \in \mathbb{N}_\infty$ 个值 $i \in \mathbb{N}$，使得对于某个 $k \in \mathbb{N}$，$(i,\ k)$ 无限期地保留在开放路径中。令 $i_1 < i_2 < \ldots$ 为这些值。取 $j_1 := i_1$ 和 $j_{k+1} := i_{k+1} - i_k - 1$，我们得到 $\mathcal{C}$ 具有如下形式

- $\mathcal{C} = 0^{j_1} \cdot \{0,1\} \cdot 0^{j_2} \cdot \{0,1\} \cdot \ldots \cdot 0^{j_\ell} \cdot \{0,1\} \cdot \widehat{0}$ 若 ℓ 是有限的，或者
- $\mathcal{C} = 0^{j_1} \cdot \{0,1\} \cdot 0^{j_2} \cdot \{0,1\} \cdot 0^{j_3} \cdot \{0,1\} \cdot \ldots$，若 $\ell = \infty$。

特别地，VC 维数$(\mathcal{C}) = \ell$。

因此，一旦生成 q，就可以将其输入 VC 维数$^-$ 的实现器，从而生成 SORT′结果的名称。 □

将范围限制为有限 VC 维数的集合以及仅含有限多个零的序列后，结合推论 28 和引理 26，同样的考虑可得以下结论。

推论 31. VC 维数$^-_{<\infty} \equiv_{\mathrm{sW}} \mathsf{C}'_{\mathbb{N}}$ 且有界 VC 维数$^- \equiv_{\mathrm{sW}} \mathsf{B}'$。

我们也可以根据所得结果，轻易推出关于 PAC 可学习概念类这一类的博雷尔复杂度的结论。这一结果对应于 Schaefer [40, 定理 4.1] 获得的算术层级分类（另见 Calvert [21]）。由于这对我们而言只是一个附带结果，因此此处不正式讨论和引入博雷尔层级。关于表示空间中博雷尔复杂度的解释，请读者参见 [5, 10]。

定理 32（PAC 可学习性的博雷尔复杂度）。在有效博雷尔层级中，$\mathcal{PL}$ PAC 可学习的闭概念类在 $\mathcal{A}$ 和 $\mathcal{A}_+$中是 Σ^0_2–完全的，并且在 $\mathcal{A}_-$ 中是 Σ^0_3–完全的。

证明。根据统计学习基本定理定理 3，我们有

$$\mathcal{PL} := \{\mathcal{C} \in \mathcal{A} : \mathcal{C} \text{ PAC learnable}\} = \{\mathcal{C} \in \mathcal{A} : \mathrm{VCdim}(\mathcal{C}) < \infty \text{ and } \mathcal{C} \neq \varnothing\}.$$

函数 $f\colon\ \mathcal{A} \to 2^{\mathbb{N}}$，它将每个概念类 $\mathcal{C} \in \mathcal{A}$ 映射到 $f(\mathcal{C})$：若 $n =$VCdim$(\mathcal{C}) < \infty$（或 $\mathcal{C} = \varnothing$），则取值为 $= 0^n\widehat{1}$；若 $n = \infty$，则取值为 $f(\mathcal{C}) = \widehat{0}$。根据定理27，该函数与 SORT 强 Weihrauch等价，因此可归约到 lim。这意味着 f 是 Σ^0_2–可测的 [5]。因此，$f^{-1}(2^{\mathbb{N}} \setminus \{\widehat{0}\})$ 是一个有效的 Σ^0_2–集，$\mathcal{PL} = f^{-1}(2^{\mathbb{N}} \setminus \{\widehat{0}\}) \setminus \{\varnothing\}$ 也同样如此，因为 $\{\varnothing\}$ 在 $\mathcal{A}$中是余递归可枚举闭的。将 f 视为类型为 $f\colon \mathcal{A}_+ \to 2^{\mathbb{N}}$ 的函数，可类似地证明 $\mathcal{PL}$ 也是 $\mathcal{A}_+$中的有效 Σ^0_2–集，因为 $\{\varnothing\}$ 在 $\mathcal{A}_+$中也余递归可枚举闭。若将 f 视为类型为 $f\colon \mathcal{A}_- \to 2^{\mathbb{N}}$的函数，则由定理 30可得，$f$ 与 SORT′ 强Weihrauch等价，因此可归约到 lim′；这意味着 $\mathcal{PL}$ 是有效 Σ^0_3–可测的。此处用到的是，$\{\varnothing\}$ 在 $\mathcal{A}_-$中是可计算可枚举开集，这归因于康托空间 $2^{\mathbb{N}}$的紧致性。

还需证明完备性结论。众所周知，包含所有仅有有限个数字 1 的二进制序列的类 $\mathcal{F} \subseteq 2^{\mathbb{N}}$，在（有效）博雷尔层级 [37, 中是 Σ^0_2–完备的，见练习 23.1]。我们考虑可计算函数 $g\colon\ 2^{\mathbb{N}} \to \mathcal{A}$，其满足

$$g : 2^{\mathbb{N}} \to \mathcal{A}, p \mapsto \{0,1\}^{p(0)} \cdot \{0,1\}^{p(1)} \cdot \{0,1\}^{p(2)} \cdot \ldots,$$

其中我们假设 $\{0,1\}^0=\{0\}$。于是 $g(p)\neq\varnothing$ 并且

$$p\in\mathcal{F}\iff \mathrm{VCdim}(g(p))<\infty\iff g(p)\in\mathcal{PL}$$

对所有 $p\in 2^{\mathbb{N}}$均成立。因此， g 给出了一个有效的 Wadge 归约，见证了 $\mathcal{PL}$ 在 $\mathcal{A}$中具有 Σ^0_2–完备性。这也意味着它在 $\mathcal{A}_+$ 中具有 Σ^0_2–完备性。

我们仍需证明 $\mathcal{PL}$ 在 $\mathcal{A}_-$中的 Σ^0_3–完备性。众所周知，集合$\mathcal{G}:=\{p\in 2^{\mathbb{N}}:(\forall^\infty k)(\exists n)\ p\langle n,k\rangle=1\}$ 是（有效地） Σ^0_3–完备的，见 [37, 练习 23.2]。我们证明存在一个可计算函数 $h:2^{\mathbb{N}}\to\mathcal{A}_-$，使得 $p\in\mathcal{G}\iff h(p)\in\mathcal{PL}$，并且所有 $h(p)$ 都非空。给定 $p\in 2^{\mathbb{N}}$，我们按如下方式在 $2^{\mathbb{N}}$中计算序列 $(p_n)_{n\in\mathbb{N}}$。令 $p_n(k)=0$，如果 $p\langle i,k\rangle=0$，对 $i=0,...,n$ 均如此；如果 $p\langle m,k\rangle=1$，则切换为 $p_n(k)=1$，对所有 $n\geq m$ 均如此。因此，序列 $(p_n)_{n\in\mathbb{N}}$ 收敛，并且其极限在满足存在 $n\in\mathbb{N}$ 使得 $p\langle n,k\rangle=1$的那些位置 k 上包含 1。现在，正如定理 30 的证明中那样，我们为这个收敛序列 $(p_n)_{n\in\mathbb{N}}$ 计算一个非空概念类 $\mathcal{C}\in\mathcal{A}_-$，并令 $h(p):=\mathcal{C}$。于是我们得到

$$\begin{aligned} p\in\mathcal{G} &\iff (\forall^\infty k)(\exists n)\ p\langle n,k\rangle=1\iff(\forall^\infty k)\lim_{n\to\infty}p_n(k)=1\\ &\iff \mathrm{VCdim}(h(p))<\infty\iff h(p)\in\mathcal{PL}.\end{aligned}$$

证明至此完成。 □

我们注意到， $\mathcal{A}$是一个完全可度量可分空间，而我们得到的博雷尔结构是标准的 Effros 博雷尔结构 [37, 10]。空间 $\mathcal{A}_-$ 和 $\mathcal{A}_+$ 是拟波兰空间，此类空间上的博雷尔复杂度也可以直接定义 [25] （其方式与我们通过实现器所采用的方法等价）。

有效博雷尔完备性也可用于刻画判定 PAC 可学习性的特征函数 $\chi_{\mathcal{PL}}:\ \mathcal{A}\to\{0,1\}$ 的 Weihrauch 复杂度。若分别使用 $\mathcal{A}_+$ 或 $\mathcal{A}_-$ 代替 $\mathcal{A}$，我们记作 $\chi^+_{\mathcal{PL}}$ 或 $\chi^-_{\mathcal{PL}}$。

C推论 33 （PAC 可学习性）。 $\chi_{\mathcal{PL}}\equiv_{\mathrm{sW}}\chi^+_{\mathcal{PL}}\equiv_{\mathrm{sW}}\mathrm{LPO}'$ 以及 $\chi^-_{\mathcal{PL}}\equiv_{\mathrm{sW}}\mathrm{LPO}''$。

证明。 由定理 32 的证明，我们可以提取归约 $\chi_{\mathcal{F}}\leq_{\mathrm{sW}}\chi_{\mathcal{PL}}\leq_{\mathrm{sW}}\mathrm{LPO}'$（$\chi^+_{\mathcal{PL}}$同样如此）以及 $\chi_{\mathcal{G}}\leq_{\mathrm{sW}}\chi^-_{\mathcal{PL}}\leq_{\mathrm{sW}}\mathrm{LPO}''$。归约 $\mathrm{LPO}'\leq_{\mathrm{sW}}\chi_{\mathcal{F}}$和 $\mathrm{LPO}''\leq_{\mathrm{sW}}\chi_{\mathcal{G}}$成立，这是因为 $\mathrm{LPO}^{(n)}=\mathrm{LPO}\circ\lim^{(n-1)}$ 是某个有效 Π^0_{n+1}–集的特征函数，而$n\geq 1$ （由于 lim 是有效 Σ^0_2–可测的，例如由 [5, 命题 9。1] 可知）。 □

4. Relative PAC Learning

本节证明：只要概念类 $\mathcal{C}\in\mathcal{A}_+$ 由正信息给出、假设类 $\mathcal{H}\in\mathcal{A}_-$ 由负信息给出，并且将 $\mathcal{C}$ 的 VC 维数上界作为额外的输入信息提供，相对 PAC 学习在可计算意义下等价于弱 K″ önig 引理。为了证明关于相对 PAC 学习的上述结果，我们需要进行一些准备工作。

引理 34 （概率分布集合）。 概率分布的集合 $\mathcal{P}\ \mathcal{D}:\ \mathbb{N}\times\{0,1\}\to\mathbb{R}$ 是 $\mathbb{R}^{\mathbb{N}\times\{0,1\}}$ 的可计算可枚举闭子集。

证明。 为简单起见，我们使用可计算编码 $f:\ \mathbb{N}\to\mathbb{N}\times\{0,1\}$ 其中 $f(2n)=(n,\ 0)$，以及 $f(2n+1)=(n,\ 1)$，从而可以将 $\mathbb{N}\times\{0,1\}$ 与 $\mathbb{N}$ 对应起来。现在，给定一个非空有理区间的有限序列 $(a_i,\ b_i)\subseteq\mathbb{R}$ 其中 $a_i,\ ...\ b_i\in\mathbb{Q}$，对于 $i=0,\ ...\ k$，我们需要

识别基本开集 $(\mathbf{X}_{i=0}^{k}(a_i,\ b_i))\ \times \mathbb{R}^{\mathbb{N}}$ 是否包含一个概率分布$\mathcal{D}$： $\mathbb{N} \to \mathbb{R}$。然而，当且仅当

$$\sum_{i=0}^{k} a_i < 1 \text{ and } (\forall i = 0, ..., k)\ b_i \geq 0.$$

这甚至是一种可判定性质。也很容易看出，概率分布集合是一个闭子集，因此整体上是可计算可枚举闭集。 □

L引理 35 (风险的可计算性)。 以下函数是可计算的 e:
(1) $L\colon \mathcal{P} \times 2^{\mathbb{N}} \to \mathbb{R},\ (\mathcal{D},\ h) \mapsto L_{\mathcal{D}}(h)$。(2) $\inf\ L\colon \subseteq \mathcal{P} \times \mathcal{A}_+ \to \mathbb{R}_>,\ (\mathcal{D},\ \mathcal{C}) \mapsto \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h)$。这里 $\inf L$ 仅对非空的 C有定义。

证明。 如果 $\mathcal{D}$ 是一个概率分布，那么对于每个值 $n \in \mathbb{N}$，我们都可以确定一个$k \in \mathbb{N}$，使得 $\sum_{x\leq k,\ y\in\{0,1\}} \mathcal{D}(x,\ y)\ > 1 - 2^{-n}$。这使我们能够计算

$$L_{\mathcal{D}}(h) = \sum_{(x,y)\notin \mathrm{graph}(h)} \mathcal{D}(x, y)$$

对于每个 $h \in 2^{\mathbb{N}}$，都能达到任意指定的精度 2^{-n}，这就证明了 (1)。对于 (2)，我们注意到，根据命题 9，非空的 $\mathcal{C} \in \mathcal{A}_+$ 可以由一个序列 $(h_n)_{n\in\mathbb{N}}$ 给出，该序列满足$\mathcal{C} = \{h_n\colon\ n \in \mathbb{N}\}$。于是我们得到

$$\inf L(\mathcal{D}, \mathcal{C}) = \inf_{n\in\mathbb{N}} L_{\mathcal{D}}(h_n),$$

根据 (1)，它可以作为一个数在 $\mathbb{R}_>$ 中计算。 □

接下来，我们证明相对 PAC 学习可以归约到弱柯尼希引理。我们利用 $\mathsf{WKL} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathcal{L}}$ 这一事实，并证明，给定 $\mathcal{C}$、 $\mathcal{H}$ 以及 $\mathcal{C}$ 的 VC 维数上界，我们可以识别出一个足够大的非学习器类。

命题 36. 相对 PAC 学习 $\leq_{\mathrm{sW}} \mathsf{WKL}$。

证明。 给定 $\mathcal{C} \in \mathcal{A}_+$、 $\mathcal{H} \in \mathcal{A}_-$ 和 $d \in \mathbb{N}$，且 VC维$(\mathcal{C})\ \leq d$。那么，我们可以计算出 $\mathcal{C}$ 的样本复杂度，其形式为

$$m_d\langle i, j\rangle = c \cdot 2^i(d \cdot i + j)$$

其中 $c \in \mathbb{N}$ 是某个适当的通用常数。因此，根据基本定理 3，我们知道存在一个学习器 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$，其值域 $\mathrm{range}(A)\ \subseteq \mathcal{H}$ （甚至可以取值域 $\mathrm{range}(A)\ \subseteq \mathcal{C}$ 的学习器，但我们不想使用这一条件），它以样本复杂度 m_d PAC 学习 $\mathcal{C}$；也就是说，对于每个 $i,\ j \in \mathbb{N}$ 和 $n \geq m_d\langle i,\ j\rangle$ 以及每个 $\mathcal{D} \in \mathcal{P}$，学习条件 (1.1) 都成立。

为了证明我们可以使用 $\mathsf{C}_{\mathcal{L}} \equiv_{\mathrm{sW}} \mathsf{WKL}$ 找到这样的学习器，并给定 $\mathcal{C},\ \mathcal{H}$ 和d，只需证明我们可以可计算地识别出函数 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$ 不是这样的学习器。但这等价于

(1) $(\exists S \in \mathcal{S})\ A(S)\ \notin \mathcal{H}$ 或 (2) $(\exists \mathcal{D} \in \mathcal{P})(\exists i,\ j \in \mathbb{N})(\exists n \geq m_d\langle i,\ j\rangle)\ \mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S))\right.$ $\left.> \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h)\ + 2^{-i}\right] > 2^{-j}$。

第一个条件显然是相对于 A 和 $\mathcal{H}$ 的可计算可枚举开条件，因为 $\mathcal{H} \in \mathcal{A}_-$ 是根据负信息给出的。对于第二个条件，我们注意到，根据引理 34， $\mathcal{P}$ 概率分布的集合是可计算可枚举闭集。因此，如果量词之后的条件是可计算可枚举开的，那么对 $\mathcal{D}$ 进行存在量化会导出一个可计算可枚举开条件。为此，我们注意到

根据引理 35，在给定 A、$\mathcal{D}$ 和 S 的情况下，我们可以计算 $L_{\mathcal{D}}(A(S)) \in \mathbb{R}$，并且在给定 $\mathcal{C}$ 和 $\mathcal{D}$ 的情况下，我们可以计算 $\inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h) \in \mathbb{R}_{>}$。因此，条件

$$U = \left\{ S \in \mathcal{S} : L_{\mathcal{D}}(A(S)) > \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h) + 2^{-i} \right\}$$

在其参数 $\mathcal{D}, \mathcal{C}$ 和 A 上是可计算可枚举的。这意味着我们可以计算概率

$$\mathbb{P}_{S\sim\mathcal{D}^n}[S \in U] = \sum_{((x_1,y_1),\ldots,(x_n,y_n))\in U} \prod_{i=1}^{n} \mathcal{D}(x_i, y_i) \in \mathbb{R}_{<}$$

关于 $n \in \mathbb{N}$ 一致地，因为我们可以将所有能够确认的正实例加总起来。这足以使条件 $\mathbb{P}_{S\sim\mathcal{D}^n}[S \in U] > 2^{-j}$ 在其所有参数上都是可计算可枚举开的。因此，借助 $\mathsf{C}_{\mathcal{L}} \equiv_{\mathrm{sW}} \mathsf{WKL}$，我们可以找到一个学习器 A，使其相对于 $\mathcal{H}$，以样本复杂度 m_d 对 $\mathcal{C}$ 进行 PAC 学习。这使我们能够将对 (A, m_d) 的描述作为输出信息。 □

$\widehat{\ }$在证明的逆向部分，我们利用了这样一个事实：$\mathsf{WKL} \equiv_{\mathrm{sW}} \mathsf{C}_2$，并且对于 C_2 的任意$\widehat{\ }$实例 $(B_k)_{k\in\mathbb{N}}$，我们都能计算出一个概念类 $\mathcal{C} \in \mathcal{A}_+$，其 VC 维数为 $\mathrm{VCdim}(\mathcal{C}) \leq 1$，以及一个假设类 $\mathcal{H} \in \mathcal{A}_-$，使得 $\mathcal{C} \subseteq \mathcal{H}$，并且任意学习器 $A : \mathcal{S} \to \mathcal{H}$ 若能相对于 $\mathcal{H}$ 对 $\mathcal{C}$ 进行 PAC 学习，就能让我们为每个 $k \in \mathbb{N}$ 重构答案 $i \in \mathsf{C}_2(B_k)$。

命题 37. $\mathsf{WKL} \leq_{\mathrm{sW}}$ 相对 PAC 学习。

证明。给定一个集合序列 $B = (B_k)_{k\in\mathbb{N}}$，其中 $B_k \in \mathcal{A}_-(\{0,1\})$ 满足对所有 $B_k \neq \varnothing$，我们按如下方式计算 $k \in \mathbb{N}$:

$\widehat{\ }$(1) 将 0 加入 $\mathcal{C}_B$。(2) 若 $1 \notin B_k$，则将 $h_{k,0} := 0^{k+1}1\widehat{0}$ 加入 $\mathcal{C}_B$。(3) 若 $0 \notin B_k$，则将 $h_{k,1} := 0^{k+1}11\widehat{0}$ 加入 $\mathcal{C}_B$。

也就是说

$$\mathcal{C}_B = \{\widehat{0}\} \cup \{h_{k,b} : 1 - b \notin B_k\}.$$

并且我们按如下方式计算 $\mathcal{H}_B \in \mathcal{A}_-$，其中 $\mathcal{C}_B \subseteq \mathcal{H}_B$：

(1) 从 $\mathcal{H}_B$ 开始，其中包含所有可能属于 $\mathcal{C}_B$ 的元素。(2) 若 $1 \notin B_k$，则从 $\mathcal{H}_B$ 中移除 $h_{k,1}$。(3) 若 $0 \notin B_k$，则从 $\mathcal{H}_B$ 中移除 $h_{k,0}$。

也就是说

$$\mathcal{H}_B = \{\widehat{0}\} \cup \{h_{k,b} : k \in \mathbb{N}, b \in \{0,1\}\} \setminus \{h_{k,b} : b \notin B_k\}.$$

由于对所有 $k \in \mathbb{N}$ 都有 $B_k \neq \varnothing$，因此可以保证 $\mathcal{C}_B \subseteq \mathcal{H}_B$；这是因为 $\mathcal{H}_B$ 的构造始于一个包含所有可能属于 $\mathcal{C}_B$ 的对象的集合，并且只移除那些确定不会被加入 $\mathcal{C}_B$ 的路径。我们注意到，$\mathcal{C}_B$ 的 VC 维数小于或等于 1。现在假设 $(A, m) \in \mathsf{RPAC}(\mathcal{C}_B, \mathcal{H}_B, 1)$。为了确定 $\mathsf{C}_2(B_k)$，我们需要找到 $b \in B_k$。学习器 A 必须遵守学习条件 (1.1)，特别是对于狄拉克分布 $\mathcal{D}$，其中 $\mathcal{D}(k+1, 1) = 1$。若 $B_k = \{b\}$ 满足 $b \in \{0,1\}$，则 $L_{\mathcal{D}}(h) = 0$ 仅当 $h = h_{k,b} \in \mathcal{C}_B$ 时成立，而对于所有其他 $h \in \mathcal{H}_B \setminus \{h_{k,b}\}$，$L_{\mathcal{D}}(h) = 1$。因此，对于大小为 $n \geq m\langle 2$ 的样本 $S = ((k+1, 1), (k+1, 1) \ldots, (k+1, 1)) \in \mathcal{S}$，为了满足条件，学习器必然必须返回 $A(S) = h_{k,b}$，且满足 $h_{k,b}(k+2) = b \in B_k$

$$\mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \leq \inf_{h\in\mathcal{C}_B} L_{\mathcal{D}}(h) + \frac{1}{4} \right] > \frac{3}{4}. \tag{4.1}$$

若 $B_k = \{0,1\}$，则对于某个 $h \in \mathcal{H}_B$，我们也有 $A(S) = h$；无论如何，$h(k+2) \in B_k$。
⌢这就证明了归约 $\mathsf{WKL} \equiv_{\mathrm{sW}} \mathsf{C}_2 \leq_{\mathrm{sW}} \mathsf{RPAC}$。 □

If we modify RPAC to $\mathsf{RPAC}^{\mathrm{fix}} :\subseteq \mathcal{A}_+ \times \mathcal{A}_- \times \mathbb{N} \rightrightarrows \mathcal{L}$ with

$$\mathsf{RPAC}^{\mathrm{fix}}(\mathcal{C}, \mathcal{H}, d) := \{A : A \text{ PAC learns } \mathcal{C} \text{ relatively to } \mathcal{H} \text{ with } m_d\}$$

与 RPAC 具有相同的定义域，也就是说，如果学习器始终必须使用固定的样本复杂度 m_d，那么上述所有证明仍然成立。因此，我们证明了以下结论。对于任意固定的有限 VC 维数 ≥ 1，这些证明同样成立。

定理 38（相对 PAC 学习）。$\mathsf{RPAC} \equiv_{\mathrm{sW}} \mathsf{RPAC}_{<d} \equiv_{\mathrm{sW}} \mathsf{RPAC}^{\mathrm{fix}} \equiv_{\mathrm{sW}} \mathsf{WKL}$，且$1 < d < \infty$。

由定理 38，我们可以得到各种正面和负面推论。例如，根据克利尼树（即没有可计算路径的可计算二叉树）的存在性，我们可以得出：存在包含于余可计算可枚举闭假设类中的可计算可枚举闭概念类，但它们不允许具有可计算样本复杂度的可计算学习器。然而，在推论 73 中，我们将表述一个更强的负面结果。

另一方面，低基定理导致了以下推论（见 [9]）。回顾一下，如果学习器 $A : \mathcal{S} \to \mathcal{H}$ 被视为康托空间 $\mathcal{L}$ 中的一个点，则称其为低的，当且仅当它的图灵跳跃图灵可归约于停机问题。这意味着，从其预言机关于跳跃的能力不强于可计算学习器这一意义上说，该学习器几乎是可计算的。

推论 39（相对 PAC 学习器的低度性）。对于每个具有有限 VC 维数 d 的可计算可枚举闭概念类 $\mathcal{C}$，以及每个包含 $\mathcal{C}$ 的余递归可枚举闭超类 $\mathcal{H}$，都存在一个低度学习器 $A\colon \mathcal{S} \to \mathcal{H}$，它能相对于 $\mathcal{H}$ 对 $\mathcal{C}$ 进行 PAC 学习，且具有可计算样本复杂度 m_d。

最后，已知 [9, 定理 7.2]，WKL 在 [50]意义下是非确定性可计算问题的完全问题。因此我们得到以下结论。

推论 40（非确定性相对 PAC 学习）。相对PAC学习问题 RPAC 是非确定性可计算问题类的完全问题。

本节最后，我们研究这样一个问题：如果不将 VC 维数的上界作为额外输入信息提供，而仅将 VC 维数有限作为前提条件，相对 PAC 学习的复杂度会受到怎样的影响。也就是说，我们要研究如下的 $\mathsf{RPAC}_{<\infty}$。我们证明，它实际上等价于欧几里得选择 $\mathsf{C}_{\mathbb{R}}$。

定理 41（无维数的相对 PAC 学习）。相对 PAC 学习$_{<\infty} \equiv_{\mathrm{sW}}$ 闭包引理$_{\mathbb{R}}$。

证明。根据推论 28，计算一个具有有限 VC 维数的类$\mathcal{C} \subseteq 2^{\mathbb{N}}$ 的 VC 维数$^{+}_{<\infty}$ 问题等价于闭包引理$_{\mathbb{N}}$。因此，结合定理 38 以及命题 16 和 18，我们得到

$$\mathsf{RPAC}_{<\infty} \leq_{\mathrm{sW}} \mathsf{RPAC} * \mathrm{VCdim}^{+}_{<\infty} \leq_{\mathrm{sW}} \mathsf{WKL} * \mathsf{C}_{\mathbb{N}} \equiv_{\mathrm{sW}} \mathsf{C}_{\mathbb{R}}.$$

⌢根据命题 18，只需证明 $\mathsf{B} \times \mathsf{C}_2 \leq_{\mathrm{sW}}$ 相对 PAC 学习$_{<\infty}$ 即可完成证明。我们采用与命题 37 证明中类似的构造。给定 $p \in \mathrm{dom}(\mathsf{B})$ 和⌒$B = (B_k)_{k\in\mathbb{N}} \in \mathrm{dom}(\mathsf{C}_2)$，我们计算集合 $\mathcal{C} = \mathcal{C}_B \cup \mathcal{C}_p \in \mathcal{A}_+$ 和 $\mathcal{H} = \mathcal{H}_B \cup \mathcal{H}_p \in \mathcal{A}_-$，其中 $\mathcal{C}_B$ 和 $\mathcal{H}_B$ 正是命题 37 证明中的集合。并且，令 $d_p :=$最大值$\{p(n) : n \in \mathbb{N}\}$，我们定义（其中 $\cdot$ 表示连接）：

- $\mathcal{C}_p := 1 \cdot \{0,1\}^{d_p} \cdot \widehat{0}$,
- $\mathcal{H}_p := 1 \cdot 2^{\mathbb{N}}$.

注意到 d_p = VC 维数($\mathcal{C}_p$)。给定 (A, m) $\in$ 相对 PAC 学习$_{<\infty}$($\mathcal{C}$, $\mathcal{H}$)，我们可以完全按照命题 37 证明中的方法，确定所有 $k \in \mathbb{N}$ 对应的 $\mathsf{C}_2(B_k)$ 的值。此外，由d = VC 维数($\mathcal{C}$)，根据基本定理 3 可得，每个 $n \in \mathbb{N}$ 都满足

$$n \geq \frac{m\langle 0,0\rangle}{C} \geq \frac{M_d\langle 0,0\rangle}{C} = d = \mathrm{VCdim}(\mathcal{C}_B \cup \mathcal{C}_p) \geq \mathrm{VCdim}(\mathcal{C}_p) = d_p,$$

满足 $n \in$B(p)。由于我们可以不失一般性地假设常数 $C \in \mathbb{R}$是可计算的，因此给定 m，我们可以确定相应的 n。归约的证明至此完成。

□

5. 适当 PAC 学习

本节研究适当 PAC 学习。如果我们将适当 PAC 学习设置（$\mathcal{C} = \mathcal{H}$）视为相对设置的一个特例，那么我们所考虑的假设类由完整信息描述。不过，考虑假设类仅由正信息描述的情形也很有意义。我们还知道，如果一个假设类是适当PAC可学习的，那么对于 $\mathcal{C}$，任何经验风险最小化器都能$\mathcal{C}$恰当 PAC 学习。因此，定义一个经验风险最小化问题也很有意义。

定义 42（经验风险最小化器）。我们将 ERM：$\mathcal{A} \setminus \{\varnothing\} \rightrightarrows \mathcal{L}$ 定义为

$$\mathsf{ERM}(\mathcal{C}) := \{A \in \mathcal{L} : A \text{ is an empirical risk minimizer for } \mathcal{C}\}.$$

再次，我们通过将 $\mathcal{A}$替换为 $\mathcal{A}_+$ 或 $\mathcal{A}_-$，分别定义 ERM^+、ERM^-。

我们首先证明，对于概念类具有完整信息时，适当 PAC学习是可计算的，无论 VC 维数的上界是否被明确给出，还是仅承诺其具有某个特定的有限值。

定理 43（具有完整信息和维数的适当 PAC 学习）。我们得到，对于 $1 < d < \infty$，$\mathsf{ERM} \equiv_{\mathrm{sW}} \mathsf{PPAC} \equiv_{\mathrm{sW}} \mathsf{PPAC}_{<d} \equiv_{\mathrm{sW}} \mathsf{id}$ 是可计算的。

证明。我们断言，可以得到归约链

$$\mathrm{id} \leq_{\mathrm{sW}} \mathsf{PPAC}_{<2} \leq_{\mathrm{sW}} \mathsf{PPAC}_{<d} \leq_{\mathrm{sW}} \mathsf{PPAC} \leq_{\mathrm{sW}} \mathsf{ERM} \leq_{\mathrm{sW}} \mathrm{id} \tag{5.1}$$

对所有 $1 < d < \infty$ 均成立。第二个和第三个归约显然成立。第一个归约可以利用可计算函数 $2^{\mathbb{N}} \to \mathcal{A}$，$p \mapsto \{p\}$ 以及 $\mathsf{id} \equiv_{\mathrm{sW}} \mathsf{id}_{2^{\mathbb{N}}}$这一事实得到。我们还需要证明其他归约。

我们首先证明 ERM 是可计算的，这将推出最后一个归约。由于我们拥有关于假设类 $\mathcal{C} \in \mathcal{A}$ 的完整信息，根据命题 10，对于任意样本 $S = ((x_1, y_1), \ldots, (x_n, y_n)) \in \mathcal{S}$，我们可以读取该信息，直到确切知道长度为 $x_{\max} + 1$ 的哪些前缀是可扩展的，其中 $x_{\max} :=$ 最大值 $\{x_1, \ldots, x_n\}$。然后，我们可以轻松计算这些可扩展前缀的经验风险，并利用正信息将其中最优的前缀扩展为假设 $h \in \mathcal{C}$。这将产生一个经验风险最小化器 A，从而解决问题 ERM。

现在我们证明 $\mathsf{PPAC} \leq_{\mathrm{sW}} \mathsf{ERM}$。根据定理 3，计算得到的经验风险最小化器能够恰当 PAC 学习 $\mathcal{C}$，且具有 $m_{\mathrm{VCdim}(\mathcal{C})}$。根据输入信息，我们知道 VC 维数($\mathcal{C}$)的上界为 d，因此只需返回 (A, m_d)即可解决问题 PPAC。这

已经证明了 PPAC $\leq_{\mathrm{W}}$ ERM。为了得到强归约，我们需要通过问题 ERM 将假设类的上界传递进去。不过，这可以通过可计算函数来实现

$$\mathcal{A} \times \mathbb{N} \to \mathcal{A}, (\mathcal{C}, d) \mapsto \mathcal{C}_d := 0^d \cdot 1 \cdot \mathcal{C}.$$

我们注意到，VC 维数($\mathcal{C}_d$) = VC 维数($\mathcal{C}$)。如果 $A_d \in$ 经验风险最小化($\mathcal{C}_d$)，那么我们可以得到某个学习器$A \in$经验风险最小化($\mathcal{C}$)，其由下式隐式定义：

$$0^d \cdot 1 \cdot A((x_1, y_1), ..., (x_n, y_n)) = A_d((x_1 + d + 1, y_1), ..., (x_n + d + 1, y_n)).$$

值 d 本身可以从 range(A_d) 中的任意值恢复。 □

现在我们考虑这样一种概念类：具有完整信息，并且唯一的承诺是该类的 VC 维数是有限的。

定理44 （具有完整信息且无维度的适当PAC学习）。 我们有 $\mathrm{PPAC}_{<\infty} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$。

证明。 我们有 $\mathrm{PPAC}_{<\infty} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$，因为我们可以返回 ($A$, m_d)，其中 A 按照前一个证明中的方式计算，而 d =VC 维数($\mathcal{C}$)则利用 $\mathcal{C}$与 $\mathsf{C}_{\mathbb{N}}$的正信息计算。

对于另一个方向，我们证明 $\mathsf{B} \leq_{\mathrm{W}} \mathrm{PPAC}_{<\infty}$，这已足以推出 $\mathsf{B} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$。给定 $p \in \mathrm{dom}(\mathsf{B})$，我们写出假设类 $\mathcal{C} \in \mathcal{A}$ 的完整信息，使得 VC 维数($\mathcal{C}$) = 最大值(p)：

$$\mathcal{C} := 0^{n_1} \cdot \{0, 1\} \cdot 0^{n_2} \cdot \{0, 1\} \cdot ... \cdot 0^{n_{\max(p)}} \cdot \{0, 1\} \cdot \widehat{0},$$

其中 n_1, n_2, ... $n_{\max(p)} \in \mathbb{N}$ 根据需要取足够大。在其他方面，也就是说，当我们在 p 中遇到更高的值时，就选择增大 VC 维数。这是可行的，因为在构造 $\mathcal{C}$的过程中，该集合在任意时刻都只确定到某个深度，因此我们可以在稍后允许取值 0 和 1。

然后，我们得到 $\mathrm{PPAC}_{<\infty}$($\mathcal{C}$) 的一个输出 (A, m)。与定理 41 的证明类似，设$n \in \mathbb{N}$ 满足 $n \geq {}^{m\langle 0}{}_{\cdot}{}_{C}{}^{0\rangle}$。于是~~$n \geq$~~ VC 维数($\mathcal{C}$)，因此 $n \in \mathsf{B}(p)$。 □

我们注意到，该结论无法加强为强等价，因为对于某些输入，$\mathrm{PPAC}_{<\infty}$只能产生不可计算输出，而$\mathsf{C}_{\mathbb{N}}$ 总是产生可计算输出。不过，利用可计算映射

$$\mathcal{A} \times 2^{\mathbb{N}} \to \mathcal{A}, (\mathcal{C}, p) \mapsto \{\langle q, p\rangle : q \in \mathcal{C}\}$$

结合前一个证明中的构造，我们得到以下推论，它表明 $\mathrm{PPAC}_{<\infty}$ 的强等价类恰好是 $\mathsf{C}_{\mathbb{N}}$的柱化；这个度还包含在离散拓扑 [9]下 $\mathbb{N}^{\mathbb{N}}$ 的极限 $\lim_{\Delta}$。

推论 45. $\mathrm{PPAC}_{<\infty} \equiv_{\mathrm{sW}}$ 恒等映射 $\times \mathsf{C}_{\mathbb{N}} \equiv_{\mathrm{sW}} \lim_{\Delta}$。

现在我们研究仅具有关于概念类的正信息或负信息的适当 PAC 学习，结果表明它等价于极限映射。完整信息下适当 PAC 学习的跳跃也同样如此。后一种情形也可视为在概念类上提供了更弱的极限信息。

定理46 （具有较弱信息的适当 PAC 学习）。 我们得到
(1) $\mathrm{ERM}^{+} \equiv_{\mathrm{sW}} \mathrm{PPAC}^{+}_{<d} \equiv_{\mathrm{sW}} \mathrm{PPAC}^{+} \equiv_{\mathrm{sW}} \lim$ 对于 $1 < d \leq \infty$。
(2) $\mathrm{ERM}^{-} \equiv_{\mathrm{sW}} \mathrm{PPAC}^{-}_{<d} \equiv_{\mathrm{sW}} \mathrm{PPAC}^{-} \equiv_{\mathrm{sW}} \lim$ 对于 $1 < d < \infty$。

(3) $\mathrm{ERM}' \equiv_{\mathrm{sW}} \mathrm{PPAC}'_{<d} \equiv_{\mathrm{sW}} \mathrm{PPAC}' \equiv_{\mathrm{sW}} \lim$，当 $1 < d < \infty$ 时。

证明。(3) 是定理 43 的直接推论，因为 $\mathrm{id}' \equiv_{\mathrm{sW}} \lim$。(1) 和 (2) 对于 $1 < d < \infty$ 可以通过与 (5.1) 中类似的归约链来证明，即我们有

$$\lim \leq_{\mathrm{sW}} \mathrm{PPAC}^{\pm}_{<2} \leq_{\mathrm{sW}} \mathrm{PPAC}^{\pm}_{<d} \leq_{\mathrm{sW}} \mathrm{PPAC}^{\pm} \leq_{\mathrm{sW}} \mathrm{ERM}^{\pm} \leq_{\mathrm{sW}} \lim,$$

其中，“±”的含义是：我们可以在所有地方都采用“+”，或者在所有地方都采用“–”。第二个和第三个归约显然仍然成立。根据 (3) 和命题 23，最后提到的归约成立，因为 $\mathrm{ERM}^{\pm} \leq_{\mathrm{sW}} \mathrm{ERM}' \equiv_{\mathrm{sW}} \lim$。$\mathrm{PPAC}^{\pm} \leq_{\mathrm{sW}} \mathrm{ERM}^{\pm}$ 的归约可以按照定理 43 的证明来证明，注意到其中使用的函数在仅有负信息或仅有正信息的情况下也是可计算的。

$$\mathcal{A}_{\pm} \times \mathbb{N} \to \mathcal{A}_{\pm}, (\mathcal{C}, d) \mapsto \mathcal{C}_d := 0^d \cdot 1 \cdot \mathcal{C}$$

我们仍需证明 $\lim \leq_{\mathrm{sW}} \mathrm{PPAC}^{\pm}_{<2}$，为此使用 $\lim \equiv_{\mathrm{sW}} \widehat{\ }\mathrm{LPO}$，这由命题 19 得出。$\widehat{\ }$对于 LPO 的给定输入 $p = \langle p_0, p_1, \ldots \rangle$，我们可以计算出概念类

- $\mathcal{C}^+ = \{\widehat{0}\} \cup \{0^k 1\widehat{0} : (\exists i \in \mathbb{N})\ p_k(i) = 1\} \in \mathcal{A}_+$,
- $\mathcal{C}^- = \{\widehat{0}\} \cup \{0^k 1\widehat{0} : (\forall i \in \mathbb{N})\ p_k(i) = 0\} \in \mathcal{A}_-$.

关于 $\mathcal{C}^+$ 的正信息可通过在 p_k 中遇到 1 时添加假设 $h_k = 0^k 1\widehat{0}$来计算；关于 $\mathcal{C}^-$ 的负信息可通过在 p_k中遇到 1 时从 $\{\widehat{0}\} \cup \{0^k 1\widehat{0}$ 中移除 $h_k : \ k \in \mathbb{N}\}$ 来计算。注意，我们有 $\mathrm{VCdim}(\mathcal{C}^{\pm}) = 1$。

现在，我们得到一个学习器 A 和一个样本复杂度 m，适用于具有 $\mathrm{PPAC}^{\pm}_{<2}(\mathcal{C}^{\pm})$ 的 $\mathcal{C}^{\pm}$。利用m，我们计算出样本量 $n \geq m\langle 2, 2\rangle$，使得准确度误差和置信误差均为$1\,4$。于是，对于大小为 n的样本 $S = ((k, 1) \ldots, (k, 1))$，学习器 A必须满足式（4.1）中的学习条件。因此，考虑狄拉克分布 $\mathcal{D}_{(k,1)} = 1$，我们得到 $h = A(S)$满足 $h(k) = 1$，当$\mathcal{C}^+$ 时当且仅当 p_k 中存在一个 1，而当 $\mathcal{C}^-$时当且仅当 p_k 中不存在 1。在这两种情况下，我们都可以从 $h = A(S)$中恢复关于 $\widehat{\ }\mathrm{LPO}(p)$的信息，由此证明 $\widehat{\ }\mathrm{LPO} \leq_{\mathrm{sW}} \mathrm{PPAC}^{\pm}_{<2}$。

对于情形 $d = \infty$，还需证明 (1)。不过，我们现在可以按照定理 41 的证明进行，并使用 $\mathsf{C}_{\mathbb{N}}$ 来计算 VC 维数（见推论 28）。由命题 19 可得

$$\mathrm{PPAC}^{+}_{<2} \leq_{\mathrm{sW}} \mathrm{PPAC}^{+}_{<\infty} \leq_{\mathrm{sW}} \mathrm{PPAC}^{+} * \mathrm{VCdim}^{+}_{<\infty} \leq_{\mathrm{sW}} \lim * \mathsf{C}_{\mathbb{N}} \equiv_{\mathrm{sW}} \lim,$$

结合情形 $d = 2$ 下的 (1)，即可完成证明。 □

我们注意到，将 lim 归约到适当 PAC 学习问题，只需在单个输入 $m\langle 2,\ 2\rangle$上计算样本复杂度 m。因此，全部不可计算性都包含在学习器中，由此得到以下非一致结果，它强化了已知结果：存在 VC 维数为 1 且不允许可计算学习器的 RER 概念类；该结果由 Agarwal 等人证明。[2, 定理 9]。

推论 47（停机问题图灵度的适当学习器）。存在一个可计算可枚举闭（以及另一个余递归可枚举闭）概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$，其 VC 维数为 1，使得每个能够适当学习 $\mathcal{C}$ 的学习器 $A : \mathcal{S} \to \mathcal{C}$ 必然具有停机问题的图灵度。

根据定理 43，这样的概念类 $\mathcal{C}$ 不可能是可计算闭的（即同时是可计算可枚举闭集和余可计算可枚举闭集）。而停机问题的图灵度给出了所能达到的最大复杂度。我们称一个闭概念类极限可计算，

如果它在空间 $\mathcal{A}'$ 中是可计算的。这个类的集合既包括可计算可枚举闭概念类，也包括余可计算可枚举闭概念类。

推论 48（极限可计算风险最小化器）。 每个有限 VC 维数的极限可计算闭概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$ 都存在一个极限可计算的经验风险最小化器A: $\mathcal{S} \to \mathcal{C}$。

下面的引理说明 PAC 学习问题具有幂等性，也就是说，我们可以用一个实例解决两个实例。

引理 49（PAC 学习的幂等性）。 给定两个非空概念类 $\mathcal{C}_i \subseteq 2^{\mathbb{N}}$，其 VC 维数为 $\mathrm{VCdim}(\mathcal{C}_i) = d_i$，其中 $i \in \{0,1\}$，概念类 $\mathcal{C} = \langle \mathcal{C}_0, \mathcal{C}_1 \rangle$ 满足 $\mathrm{VCdim}(\mathcal{C}) = d_0 + d_1$。给定一个学习器 $A : \mathcal{S} \to 2^{\mathbb{N}}$，用于 $\mathcal{C}$，其样本复杂度为 m，我们可以计算出学习器 $A_i : \mathcal{S} \to 2^{\mathbb{N}}$，用于 $\mathcal{C}_i$，且样本复杂度相同，为 m。此外，如果$A : \mathcal{S} \to \mathcal{C}$ 是 $\mathcal{C}$ 的适当学习器，那么学习器 $A_i : \mathcal{S} \to \mathcal{C}_i$ 就是 $\mathcal{C}_i$ 的适当学习器，其中 $i \in \{0,1\}$。

证明。类 $\mathcal{C} = \langle \mathcal{C}_0, \mathcal{C}_1 \rangle = \{\langle p, q \rangle : \ p \in \mathcal{C}_0 \text{和 } q \in \mathcal{C}_1\}$ 定义如下：来自 $\mathcal{C}_0$ 的假设填入偶数位置，来自 $\mathcal{C}_1$ 的假设填入奇数位置。设 $B \subseteq \mathbb{N}$是一个被 $\mathcal{C}$打散的有限集合。则 $\mathcal{C}_i$ 打散 $B_i := \{n \in \mathbb{N} : 2n + i \in B\}$，其中 $i \in \{0,1\}$且 $|B| = |B_0| + |B_1|$，从而证明 $\mathrm{VCdim}(\mathcal{C}) \leq d_0 + d_1$。反过来，如果 $\mathcal{C}_i$ 针对 $i \in \{0,1\}$打散有限集合$B_i \subseteq \mathbb{N}$，那么 $\mathcal{C}$ 打散不相交集合 $2B_0$ 和 $2B_1 + 1$ 的并集 B，其中 $|B| = |B_0| + |B_1|$，由此得到反向不等式。这就证明了 $\mathrm{VCdim}(\mathcal{C}) = d_0 + d_1$。

现在令 $A : \mathcal{S} \to 2^{\mathbb{N}}$ 是 $\mathcal{C}$ 的一个学习器，其样本复杂度为 m。然后我们定义学习器 $A_i : \mathcal{S} \to 2^{\mathbb{N}}$，如下所示:

$$A_i((x_1, y_1), ..., (x_n, y_n)) := \mathrm{pr}_{i+1} A((2x_1 + i, y_1), ..., (2x_n + i, y_n))$$

对所有样本 $S = ((x_1, y_1) ..., (x_n, y_n)) \in \mathcal{S}$，其中 $\mathrm{pr}_1 : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}, \langle p, q \rangle \mapsto p$ 表示对偶数位置的投影，$\mathrm{pr}_2 : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}, \langle p, q \rangle \mapsto q$ 表示对奇数位置的投影。由于学习器 A 必须对 $\mathcal{C}$ 和所有分布 $\mathcal{D}$满足学习条件 (1.1)，因此对于那些只在偶数位置或只在奇数位置上赋予全部权重的分布，它也满足该条件。但这意味着 A_i 对 $\mathcal{C}_i$ 以及所有分布都满足学习条件，且样本复杂度相同，为 m。为此，我们注意到，如果 $\mathcal{D}_i$ 是某个分布，那么就有一个对应的分布 $\mathcal{D}$，它与 $\mathcal{D}_i$ 赋予相同的权重，但只在偶数位置或奇数位置上赋权，具体取决于 $i = 0$ 还是 $i = 1$，由此得到 $\inf_{h \in \mathcal{C}_i} L_{\mathcal{D}_i}(h) = \inf_{h \in \mathcal{C}} L_{\mathcal{D}}(h)$。显然， $\mathcal{L} \to \mathcal{L} \times \mathcal{L}, A \mapsto (A_0, A_1)$ 是可计算的。

显然，如果 $\mathrm{range}(A) \subseteq \mathcal{C}$，则上述定义的学习器 A_i 满足 $\mathrm{range}(A_i) \subseteq \mathcal{C}_i$。 □

我们注意到，在定理46的（2）和（3）中不能加入情形 $d = \infty$，因为在这种情况下我们会得到不同的复杂度。

命题 50。我们得到 $\mathsf{PPAC}^{-}_{<\infty} \equiv_{\mathrm{sW}} \mathsf{PPAC}'_{<\infty} \equiv_{\mathrm{sW}} \lim \times \mathsf{C}'_{\mathbb{N}}$。

证明。 由于 $\mathsf{B} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$，我们得到 $\mathsf{id} \times \mathsf{B} \equiv_{\mathrm{sW}} \mathsf{id} \times \mathsf{C}_{\mathbb{N}}$，因此 $\lim \times \mathsf{B}' \equiv_{\mathrm{sW}} \lim \times \mathsf{C}'_{\mathbb{N}}$。所以只需证明

$$\lim \times \mathsf{B}' \leq_{\mathrm{sW}} \mathsf{PPAC}^{-}_{<\infty} \leq_{\mathrm{sW}} \mathsf{PPAC}'_{<\infty} \equiv_{\mathrm{sW}} \lim \times \mathsf{C}'_{\mathbb{N}}.$$

第二个归约及等价性由命题 19 和推论 45 得出。对于另一个归约，我们使用推论 31，它蕴含 $\mathsf{B}' \leq_{\mathrm{sW}} \mathsf{PPAC}^{-}_{<\infty}$，因为我们可以按照定理 41 的证明，利用样本复杂度 m 和常数 C，通过 $n \geq \underline{\quad}^{m\langle 0}{}_{C}{}^{0\rangle}$ $\geq \mathrm{VCdim}(\mathcal{C})$ 计算出 VC 维数的一个上界。另一方面，定理46

蕴含 $\lim \leq_{\mathrm{sW}} \mathrm{PPAC}^{-} \leq_{\mathrm{sW}} \mathrm{PPAC}^{-}_{<\infty}$。归约 $\lim \times \mathrm{B}' \leq_{\mathrm{sW}} \mathrm{PPAC}^{-}_{<\infty}$ 随之成立，因为引理 49 蕴含 $\mathrm{PPAC}^{-}_{<\infty} \times \mathrm{PPAC}^{-}_{<\infty} \leq_{\mathrm{sW}} \mathrm{PPAC}^{-}_{<\infty}$，这是因为 $\mathcal{A}_- \times \mathcal{A}_- \to \mathcal{A}_-$，$(\mathcal{C}_0,\ \mathcal{C}_1) \mapsto \langle \mathcal{C}_0,\ \mathcal{C}_1 \rangle$ 是可计算的。□

定理 46 的证明中的构造还证明了，将正信息转换为负信息，反之亦然，与极限运算等价，即使将范围限制在 VC 维数为一的闭集上也是如此。这是因为，只需将该证明中的 $\mathcal{C}^+$ 转移到 $\widehat{\ }$，或者反过来，就能计算 LPO。下一个推论中的另一方向由命题 23 得出。

推论 51. 恒等映射 $\mathrm{id}:\ \mathcal{A}_+ \to \mathcal{A}_-$ 和 $\mathrm{id}:\ \mathcal{A}_- \to \mathcal{A}_+$ 与 lim 强Weihrauch等价，即使将范围限制在 VC 维数为一的闭集上也是如此。

6. 非适当 PAC 学习

本节要分析非适当 PAC 学习问题及其变体。由于非适当 PAC 学习就是相对于 $2^{\mathbb{N}}$ 的 PAC 学习，我们可以直接得到

$$\mathsf{IPAC}^{+} \leq_{\mathrm{sW}} \mathsf{RPAC} \equiv_{\mathrm{sW}} \mathsf{WKL}$$

利用定理 38。但上界 WKL 是否紧致？IPAC^{+} 是否与之等价？我们将看到事实并非如此，但结果表明 IPAC^{+} 属于有限型DNC范围；我们将对此作出精确定义，并指出它带来的若干有趣推论。

为了开展我们的研究，分析 Delle Rose 等人给出的基本定理的可计算版本——定理 7——的证明，有助于我们理解非适当 PAC 学习。因此，我们的首要目标是提取他们证明中的一致性内容。

首先，我们需要回顾无免费午餐定理 [41, 定理 5。1], 该定理是机器学习中的基本定理之一，大致指出每个学习器都会在某项任务上遭遇严重失败。Agarwal、Ananthakrishnan、Ben-David、Lechner 和 Urner 推导出了该定理的一个可计算版本 [2, 引理 19]，借此可以针对每个学习器和样本量 n，计算出某个大小为 $k \geq 2n$ 的有限集合上的一个特定均匀分布，使得该学习器在此分布下遭遇严重失败。然而，他们在定理表述中对学习器 A 所作的可计算性假设是多余的；他们的证明实际上表明了以下形式上更强的一致性结果。

定理52（无免费午餐定理）。给定任意学习器 $A:\ \mathcal{S} \to 2^{\mathbb{N}}$ 和 $n \in \mathbb{N}$，以及自然数 $x_1 < \ldots < x_k$（对于某个 $k \geq 2n$），我们可以可计算地找到一个函数$g:\ \{x_1, \ldots, x_k\} \to \{0,1\}$，使得

$$\mathbb{P}_{S \sim \mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \geq \frac{1}{8}\right] \geq \frac{1}{7}$$

对于 {上的均匀分布 $\mathcal{D}$，$(x_1,\ g(x_1))$ ...，$(x_k,\ g(x_k))$}。

现在，我们基本上可以原样沿用 Delle Rose 等人 [27,定理 11] 以及 Sterkenburg [45, 引理 9] 中的证明，从而得到以下结果，它揭示了定理 7 证明中的统一性内容。为确保完备性，我们将简要回顾他们的证明，以确认所有结论都统一成立。记 $\mathcal{W}_k$ 为函数集合 $f:\ \mathbb{N}^k \to \{0,1\}^k$，记 $\mathcal{W} := \bigsqcup_{k \in \mathbb{N}} \mathcal{W}_k$ 为见证函数空间，其中 $\bigsqcup$ 表示余积。该空间的表示方式如下：名称 $\langle k, p \rangle$ 附带显式值 k 以及对 f的描述 $p \in \mathbb{N}^{\mathbb{N}}$。

定理 53（见证与不适当学习器）。我们得到如下结果：
(1) 给定某个非空概念类 $\mathcal{C}$（该概念类未给出）的见证 $f \in \mathcal{W}$，我们可以计算地找到一个学习器 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$ 和一个样本复杂度 $m \in \mathbb{N}^{\mathbb{N}}$，使得 A 以 m 不适当地学习 $\mathcal{C}$。
(2) 给定一个学习器 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$ 和一个样本复杂度 $m \in \mathbb{N}^{\mathbb{N}}$，使得 A 以样本复杂度 m 不适当地学习某个非空概念类 $\mathcal{C}$（该概念类未给出），我们可以计算地找到 $\mathcal{C}$ 的一个见证 $f \in \mathcal{W}$。

证明。(1) 我们概述 Delle Rose 等人 [27, 定理 11] 中的证明，以说明它如何给出所需的一致可计算解。给定某个非空概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$（该概念类未给出）的见证 $f \in \mathcal{W}$，我们有 $k \in \mathbb{N}$，以及见证函数 $f\colon \mathbb{N}^{k+1} \to \{0,1\}^{k+1}$。特别地，我们知道 VCdim($\mathcal{C}$) $\leq k$。遵循 Delle Rose 等人的做法，我们称假设 $h\colon \mathbb{N} \to \{0,1\}$ 对 良好，如果 supp(h) $:= h^{-1}\{1\}$ 是有限的，并且对所有 $x_1 < \ldots < x_{k+1} <$ max supp(h)，我们得到

$$f(x_1, \ldots, x_{k+1}) \neq (h(x_1), \ldots, h(x_{k+1})).$$

用 $\mathcal{H}_f$ 表示所有对 f良好的假设所构成的类。目标是计算一对 (A, m)，使得 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$ 是一个学习器，能够相对于$\mathcal{H}$ 学习 $\mathcal{C} := \mathcal{C} \cup \mathcal{H}_f$，其样本复杂度为 $m\colon \mathbb{N} \to \mathbb{N}$。正如 Delle Rose 等人所证明的，我们得到 VCdim($\mathcal{H}$) $\leq k+1$。给定 f后，我们可以为 $\mathcal{H}_f$计算一个经验风险最小化器$A\colon \mathcal{S} \to 2^{\mathbb{N}}$，因为给定 $S = ((x_1, y_1), \ldots, (x_n, y_n)) \in \mathcal{S}$，且其 $M =$最大值为 $\{x_1, \ldots, x_n\}$，我们只需遍历所有有限多个良好序列 h，满足 supp(h) $\subseteq \{0, \ldots, M\}$，从而找到误差 $L_S(h)$ 最小的序列。正如 Delle Rose 等人所证明的，这个 h 也是 $\mathcal{H}$ 的经验风险最小化器，因此根据定理 3，它是一个学习器，能够学习$\mathcal{C} \subseteq \mathcal{H}$，其样本复杂度为 m_{k+1}。因此，输入 $f \in \mathcal{W}$后，我们可以返回 (A, m_{k+1}) 作为结果。

(2) 现在，我们简要概述 Sterkenburg [45, 引理 9] 给出的证明，以说明它如何导出所需的均匀可计算解。给定一个学习器 $A : \mathcal{S} \to 2^{\mathbb{N}}$，它在 $\mathcal{L}$ 中学习一个非空概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$（该概念类未给定），其样本复杂度为 $m\colon \mathbb{N} \to \mathbb{N}$。我们有，对于 $n = m\langle 4, 3\rangle$，$\mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \leq \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h) + \frac{1}{16}\right] > 1 - \frac{1}{8}$，因此特别地，

$$\mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \geq \inf_{h\in\mathcal{C}} L_{\mathcal{D}}(h) + \frac{1}{8}\right] < \frac{1}{7} \tag{6.1}$$

对于每个概率分布 $\mathcal{D}$均成立。对于 $k = 2n$ 以及所有自然数 $x_1 < \ldots < x_k$，根据无免费午餐定理52，我们可以找到一个函数 $g\colon \{x_1, \ldots, x_k\} \to \{0,1\}$，使得对于 {上的均匀分布 $\mathcal{D}$，x_1，g(x_1)) ...,(x_k，g(x_k))}，我们有

$$\mathbb{P}_{S\sim\mathcal{D}^n}\left[L_{\mathcal{D}}(A(S)) \geq \frac{1}{8}\right] \geq \frac{1}{7},$$

这意味着对于每个 $h \in \mathcal{C}$，graph(g) $\not\subseteq$graph(h)。特别地，如果我们定义

$$f(x_1, \ldots, x_k) := (g(x_1), \ldots, g(x_k)).$$

这就完成了证明。 □

注意，如果我们从某个数对 (A, m) 开始，先应用 (2)，再应用 (1)，那么证明会得到一个新的数对 (A', m')，其中 m' 的形式为 $m' = m_{k+1}$，其中某个 $k \in \mathbb{N}$。因此，m' 总是可计算的。

推论54（可计算样本复杂度）。给定某个学习器 $A:\mathcal{S}\to 2^{\mathbb{N}}$ 和某个 $m\colon \mathbb{N}^{\mathbb{N}}\to\mathbb{N}^{\mathbb{N}}$，使得 A 能以样本复杂度 m 学习某个非空概念类 $\mathcal{C}\subseteq 2^{\mathbb{N}}$（该概念类未给定），我们可以通过可计算方法找到另一个学习器 $A'\colon \mathcal{S}\to 2^{\mathbb{N}}$ 以及一个可计算样本复杂度 $m'\colon \mathbb{N}\to\mathbb{N}$，使得 A' 能以样本复杂度 m' 学习 $\mathcal{C}$。

从这个意义上说，我们总能获得某种可计算的样本复杂度及与之对应的学习器。顺便指出，定理 53 中由 Delle Rose 等人和 Sterkenburg 给出的统一表述不仅蕴含定理 7，还蕴含下面的相对化版本；该版本也推广了经典统计学习基本定理的定性部分，即定理 3。

推论55（相对化统计学习基本定理）。对于非空概念类 $\mathcal{C}\subseteq 2^{\mathbb{N}}$ 和一个图灵度$\mathbf{a}$，以下命题等价：$\mathcal{C}$ 具有一个非适当学习器 $A\colon \mathcal{S}\to 2^{\mathbb{N}}$ 该学习器可从$\mathbf{a}$ 计算得到（具有可计算的样本复杂度m），当且仅当 $\mathcal{C}$ 具有一个可从$\mathbf{a}$计算得到的见证函数。

定理 53 告诉我们，对于给定了某个 VC 维数上界的概念类，计算其见证的难度等同于计算该类的非适当学习器及相应的样本复杂度。为了更形式化地表达这一点，我们引入计算见证这一问题。

定义 56（见证问题）。我们考虑见证问题 $\mathsf{WIT}:\subseteq \mathcal{A}\times\mathbb{N}\rightrightarrows\mathcal{W}$，其定义如下：

$$\mathsf{WIT}(\mathcal{C},d):=\{f:\mathbb{N}^{k+1}\to\{0,1\}^{k+1}: f \text{ is a witness function for } \mathcal{C}\}$$

且 $\mathrm{dom}(\mathsf{WIT})=\{(\mathcal{C},\ d)\colon \mathrm{VCdim}(\mathcal{C})\le d \text{ 且 } \mathcal{C}\neq\varnothing\}$。如果将空间 $\mathcal{A}$ 分别替换为 $\mathcal{A}_+$ 或 $\mathcal{A}_-$，我们也使用带上标 $+$ 或 $-$的类似记号。

我们立即得到定理 53 的以下推论。

推论 57（不适当学习与见证）。我们得到以下等价关系：$\mathsf{IPAC}\equiv_{\mathrm{sW}}\mathsf{WIT}$、$\mathsf{IPAC}^+\equiv_{\mathrm{sW}}\mathsf{WIT}^+$，以及 $\mathsf{IPAC}^-\equiv_{\mathrm{sW}}\mathsf{WIT}^-$。

在许多情况下，研究见证问题比研究非适当 PAC 学习问题本身更容易。例如，我们可以直接得到以下结论。

命题 58（计算见证）。WIT^- 是可计算的。

证明。给定概念类 $\mathcal{C}\in\mathcal{A}_-$，且 $d\ge\mathrm{VCdim}(\mathcal{C})$，我们知道，对于 $k=d+1$，存在一个 $(k-1)$–见证函数 $f\colon \mathbb{N}^k\to\{0,1\}^k$，用于 $\mathcal{C}$。给定自然数 $x_1<\ldots<x_k$，我们考虑 $M:=\max\{x_1,\ldots,x_k\}$ 以及 $W_b:=\{w\in\{0,1\}^{M+1}:(w(x_1)\ldots,w(x_k))\neq b\}$，对每个 $b\in\{0,1\}^k$ 都成立。由于 $\mathcal{C}\in\mathcal{A}_-$ 是紧集，给定 $(x_1,\ldots,x_k)$ 和 b，该性质

$$\mathcal{C}\subseteq\bigcup_{w\in W_b} w2^{\mathbb{N}}$$

根据 [20, 定理 4.9 和 4.10]，在所有参数下都是可计算可枚举的。由于存在一个 $(k-1)$–见证函数，该性质必定对有限多个 $b\in\{0,1\}^k$ 中的某一个成立。因此，我们可以找到一个使其成立的 $b\in\{0,1\}^k$，并令 $f(x_1,\ldots,x_k):=b$，从而定义一个见证函数。由此可得到一个算法，在给定 $\mathcal{C}\in\mathcal{A}_-$ 和 d 的情况下计算 f。□

借助推论 28 和推论 31，以下推论可直接得出。

推论5 9. IPAC 和 IPAC^- 可计算，$\mathrm{IPAC}_{<\infty} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$ 和 $\mathsf{B}' \leq_{\mathrm{sW}} \mathrm{IPAC}^-_{<\infty} \leq_{\mathrm{W}} \mathsf{C}'_{\mathbb{N}}$。

命题 58 的证明中的算法还表明，如果已知精确 VC 维数，我们总能找到一个见证函数。因此，我们得到如下结论。

推论60 （余递归可枚举闭类的非适当 PAC 可学习性）。 每个非空的余递归可枚举闭概念类 $\mathcal{C}$ 都满足 VC 维数($\mathcal{C}$) =扩展VC维数($\mathcal{C}$)。特别地，每个非空的余递归可枚举闭概念类 $\mathcal{C}$ 存在一个可计算的非适当学习器（其样本复杂度可计算），当且仅当它具有有限 VC 维数。

现在我们想研究基于正信息的非适当 PAC 学习 IPAC^+。为了让读者对相关内容有所了解，我们在图4中更细致地展示了图2所示 Weihrauch 格的相关部分。更多细节见[16]。

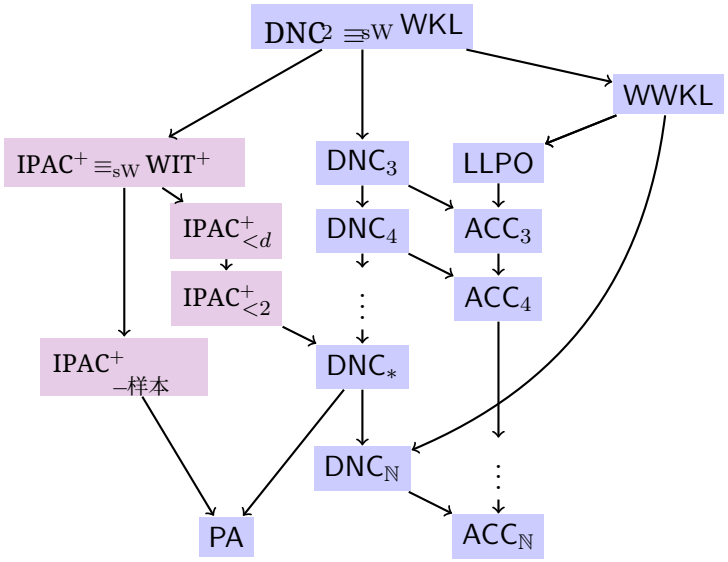

图4：Weihrauch格中DNC的邻域。

我们已经提到，$\mathrm{IPAC}^+ \leq_{\mathrm{W}} \mathrm{WKL}$ 可由定理 38 推出。根据以下观察，还可以得到另一种独立的证明。

命题 61 （计算见证集）。 该映射

$$F : \mathcal{A}_+ \to \mathcal{A}_-(\mathcal{W}), \mathcal{C} \mapsto \{f \in \mathcal{W} : f \text{ is a witness function for } \mathcal{C}\}$$

是良定义且可计算的。特别地，给定概念类的见证集是 $\mathcal{W}$ 的一个闭子集。

Proof. 对于 $\mathcal{C} \in \mathcal{A}_+$ 和 f，可立即得到这一结果： $\mathbb{N}^k \to \{0,1\}^k \in \mathcal{W}$

$$f \notin F(\mathcal{C}) \iff (\exists h \in \mathcal{C})(\exists x_1 < \ldots < x_k)\ f(x_1, \ldots, x_k) = (h(x_1), \ldots, h(x_k)).$$

对集合 $\mathcal{C} \in \mathcal{A}_+$ 进行存在量化，且谓词为可计算可枚举开谓词时，其结果相对于$\mathcal{C}$是可计算可枚举开的，由此即可得到所需结论。 □

注意，见证函数的集合可能为空。但如果将$\mathcal{C}$ 的 VC 维数作为输入信息给出，那么我们知道存在某个 k，使得 $\mathcal{W}_k$ 必定包含一个适用于 $\mathcal{C}$ 的见证函数，因此 $\mathsf{IPAC}^+ \equiv_{\mathrm{sW}} \mathsf{WIT}^+ \leq_{\mathrm{sW}}$ WKL 成立（这里我们隐含地使用了 $\mathcal{W}_k$ 与 $2^{\mathbb{N}}$ 可计算同构，因此 $\mathcal{C}_{\mathcal{W}_k} \equiv_{\mathrm{sW}} \mathcal{C}_{2^{\mathbb{N}}} \equiv_{\mathrm{sW}} \mathsf{WKL}$）。

接下来，我们通过考察见证函数，尝试将 IPAC^+ 与 WKL 区分开来。事实上，由于 $\mathsf{WKL} \equiv_{\mathrm{sW}} \widehat{\quad}\mathsf{LLPO}$，为此只需证明 $\mathsf{LLPO} \not\leq_{\mathrm{W}} \mathsf{WIT}^+$。我们证明一个在两方面都更强的结果。首先，我们甚至考虑一般的 $n \geq 2$ 下的 ACC_n（注意，$\mathsf{ACC}_2 \equiv_{\mathrm{sW}} \mathsf{LLPO}$）；其次，我们给出一种拓扑分离。这里，$\leq^*_{\mathrm{W}}$ 表示 Weihrauch 可约性的拓扑版本，其中函数 H 和 K 只需连续即可。

命题 62. 反链选择$_n \not\leq^*_{\mathrm{W}} \mathsf{WIT}^+$ 对所有 $n \geq 2$。

证明。反设对于某个 $n \geq 2$，$\mathsf{ACC}_n \leq^*_{\mathrm{W}} \mathsf{WIT}^+$。那么存在连续的 H、K，使得只要 G 是 WIT^+ 的实现器，$H\langle \mathrm{id}\ GK\rangle$ 就是 ACC_n 的实现器。现在我们考虑集合 $A = \{0,...,n-1\} \in \mathcal{A}_-(\{0,...,n-1\})$ 的一个名称 $p \in \mathbb{N}^{\mathbb{N}}$。该集合是 ACC_n 的合法输入，因此 $K(p)$ 必须是某个概念类 $\mathcal{C} \in \mathcal{A}_+$ 以及其 VC 维数上界 $d \geq \mathsf{VC}$ 维数$(\mathcal{C})$ 的名称，其中 $\mathcal{C}$ 是该概念类。WIT^+ 的实现器可以生成 $\mathcal{C}$ 的任意见证 $f \in \mathcal{W}$ 的名称。我们只考虑 $k = 2^n(d+1)$ 的情形，即 $f \in \mathcal{W}_k$。集合 $\mathcal{C}$ 的可能见证构成闭集 $W \subseteq \mathcal{W}_k$，因为只需检查见证的有限前缀即可确认它是否失败，见命题 61。由于 W 是紧空间 $(\{0,1\}^k)^{\mathbb{N}^k}$ 的闭子集，它本身也是紧的。我们可以假设 $\mathcal{W}$ 的表示方式使函数通过某种自然编码被双射表示。因此，我们可以将 W 与其名称集合等同起来。于是，H 在紧集 $\{p\} \times W$ 上一致连续，并且存在某个有限前缀 $w \sqsubseteq p$，使得对于每个 $f \in W$，H 都足以让其在 $\{0,...,n-1\}$ 中输出一个自然数。此外，根据 K 的连续性，我们可以假设 w 足够长，从而使 K 在输出端已给出 VC 维数的上界 d。对于每个 $i \in \{0,...,n-1\}$，都存在 w 的一个扩张 $p_i \in \mathbb{N}^{\mathbb{N}}$，它是集合 $= \{0,...,n-1\} \setminus \{i\}$: A_i 的名称。根据 w 的选择，这些 p_i 会被映射到概念类 $\mathcal{C}_i \in \mathcal{A}_+$ 的名称 $K(p_i)$，且满足 VC 维数$(\mathcal{C}_i) \leq d$。因此，集合 $= \mathcal{C} \cup \bigcup_{i=0}^{n-1} \mathcal{C}_i : \mathcal{C}'$ 是 $n+1$ 个集合的并集，每个集合的 VC 维数均为 $\leq d$；由并集的 VC 维数公式可知，VC 维数$(\mathcal{C}') \leq 2^n(d+1) - 1$。[3]因此，存在 $\mathcal{C}'$ 的一个见证函数 $f \in \mathcal{W}_k$，使得 $k = 2^n(d+1)$。由于 $\mathcal{C} \subseteq \mathcal{C}'$，这个见证函数也满足 $f \in W$，因此 H 会将 p 与 f 的一个名称一起映射到某个数 $i \in \{0,...,n-1\}$。由于 $\mathcal{C}_i \subseteq \mathcal{C}$，函数 f 也是 $\mathcal{C}_i$ 的一个见证函数，因此，当输入 $\mathcal{C}_i$ 和 d 的名称时，它可以是 WIT^+ 的实现器的一个可能输出。根据 w 的选择，当输入 p_i 和 f 时，H 也会得到同一个数 i。然而，由于 $i \notin A_i$，这个结果是错误的。因此，该归约不成立，矛盾！

□

证明中的论证无法推广到可数情形，因为 $\mathsf{ACC}_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{WIT}^+$。事实上，我们证明甚至 DNC_* 也可归约于 WIT^+，从而得到一个更强的结果。

命题 63。$\mathsf{DNC}_* \leq_{\mathrm{sW}} \mathsf{WIT}^+$。

证明。给定一个集合序列 $A = ((A_{n,m})_{n\in\mathbb{N}})_{m\geq 2}$ 作为 DNC_* 的输入数据，我们知道每个 $A_{n,m}$中至多缺少一个 $i \in \{0,...,m-1\} = m$。目标是找到某个数对 $(m,p) \in \mathbb{N} \times \mathbb{N}^{\mathbb{N}}$，使得 $(\forall n)\ p(n) \in A_{n,m}$。这个目标必须借助

[3]对于概念类 $\mathcal{H}_i$，由 $\mathsf{VCdim}(\mathcal{H}_i) \leq d$ 可得到 $\mathsf{VCdim}(\mathcal{H}_1 \cup \mathcal{H}_2) \leq 2d+1$，其归纳地得到 $\mathsf{VCdim}(\bigcup_{i=1}^{n+1} \mathcal{H}_i) \leq 2^n(d+1)-1$，尽管已有更好的估计结果 [41, 练习 6.11]。

适当的概念类 $\mathcal{C} \in \mathcal{A}_+$ 所对应的 $\mathsf{WIT}^+(\mathcal{C}, d)$ 来实现，同时还需要一个界 $d \geq \mathsf{VCdim}(\mathcal{C})$，该界必须根据输入数据计算得出。事实上，我们只会尝试寻找形如 $(2^k, p)$ 的数对，其中 $(\forall n)\ p(n) \in A_{n,2^k}$，且集合的索引为 2 的幂。这是因为见证问题会产生类型为 $f : \mathbb{N}^k \to \{0,1\}^k$ 的函数。这类函数恰好有 2^k 个可能的输出，我们用 $\mathsf{bin}_k(i) \in \{0,1\}^k$ 表示按 $i \in \{0, ..., 2^k - 1\}$ 的长度为 k 的二进制表示编号时，第 $(i+1)$ 个长度为 k 的二进制字。事实上，我们可以计算

$$\mathcal{C} = \{\widehat{0}\} \cup \{0^{2^{\langle n,k\rangle}} \cdot \mathsf{bin}_k(i) \cdot \widehat{0} : i \notin A_{n,2^k}, i < 2^k, k \geq 1, n \in \mathbb{N}\} \in \mathcal{A}_+$$

来自集合序列 A。我们采用满足$\langle n,k\rangle \geq k$的标准康托配对定义。因此，前缀 $0^{2^{\langle n,k\rangle}}$，再加上每个集合 $A_{n,2^k}$ 中至多缺少一个$i < 2^k$ 这一事实，确保了对于不同的值 $\langle n,k\rangle$，两个假设中形如 $\mathsf{bin}_k(i)$ 的两个块不会重叠，因此 $\mathsf{VCdim}(\mathcal{C}) \leq 1$。给定 $\mathcal{C}$，并将 VC 维数的上界 1 作为输入传给 WIT^+，我们得到某个见证函数$f \in \mathsf{WIT}^+(\mathcal{C}, 1)$，其类型为 $f : \mathbb{N}^k \to \{0,1\}^k$，其中某个 $k \geq 1$。现在，对每个 $n \in \mathbb{N}$，我们确定某个值 $p(n) \in \{0, ..., 2^k - 1\}$，使得

$$f(2^{\langle n,k\rangle}, 2^{\langle n,k\rangle} + 1, ..., 2^{\langle n,k\rangle} + k - 1) = \mathsf{bin}_k(p(n))$$

并且由于 f是 $\mathcal{C}$的见证函数，我们可以得出 $(\forall n)\ p(n) \in A_{n,2^k}$。因此，$\langle 2^k, p\rangle$是 $\mathsf{DNC}_*(A)$ 的一个合适解，这就证明了该归约。 □

由于该归约得到一个 VC 维数为 1 的概念类，我们立即得到以下推论。

推论64. $\mathsf{DNC}_* \leq_{\mathrm{sW}} \mathsf{IPAC}^+_{<2} \leq_{\mathrm{sW}} \mathsf{IPAC}^+_{<d} \leq_{\mathrm{sW}} \mathsf{IPAC}^+ <_{\mathrm{W}} \mathsf{WKL}$ 对所有 $0 < d < \infty$ 均成立。

由于 $\mathsf{DNC}_2 \equiv_{\mathrm{sW}} \mathsf{WKL}$，我们特别有

$$\mathsf{DNC}_* \leq_{\mathrm{sW}} \mathsf{IPAC}^+ <_{\mathrm{W}} \mathsf{DNC}_2. \tag{6.2}$$

这就是我们所说的 IPAC^+ 属于有限型*DNC*范围的含义，并且这意味着，在若干方面，IPAC^+ 的表现类似于一个与 DNC_n 相似（但并不等价）的问题，适用于某个 $n \in \mathbb{N}$。

由于皮亚诺算术和 $\mathsf{DNC}_{\mathbb{N}}$ 是 DNC_*的两个（事实上不可比较的）下界，我们得到若干重要结论。以下结论由命题22得出。

推论65. 皮亚诺算术 $<_{\mathrm{W}} \mathsf{IPAC}^+_{<2} \not\leq_{\mathrm{W}}$ 弱弱柯尼希引理$^{(n)}$ 对所有 $n \in \mathbb{N}$。

命题22的证明隐含地使用了 Jokusch 和 Soare 的一项经典结果 [35,推论5。4], 该结果表明，PA 度的集合是零测度的。从某种意义上说，这排除了从正表示概念类进行非适当 PAC 学习时存在有意义的概率算法的任何可能性。具体而言，这类概念类的非适当 PAC 学习不是一个概率问题，也不是 [13] 意义下的拉斯维加斯可计算问题。

推论66。 基于正信息的非适当 PAC 学习 IPAC^+ 不是拉斯维加斯可计算的（即便将范围限制在 VC 维数为 1 的概念类上也不行）。

根据 [16, 命题6，6] 已知 $\mathsf{WKL} \leq_{\mathrm{W}} \mathsf{C}'_{\mathbb{N}} * \mathsf{PA}$。因此，$\mathsf{C}'_{\mathbb{N}}$ 也为 IPAC^+ 与 WKL 相距多远提供了一个上界。

推论 67. $\mathsf{WKL} \leq_{\mathrm{sW}} \mathsf{C}'_{\mathbb{N}} * \mathsf{IPAC}^+$.

我们还可以得出一些关于问题 $\mathrm{IPAC}^+_{<\infty}$ 的结论。与命题 50 的证明类似，我们可以使用引理 49 得到一个包含 $\mathrm{IPAC}^+_{<\infty}$ 的区间。

推论 68. $\mathrm{DNC}_* \times \mathsf{C}_{\mathbb{N}} \leq_{\mathrm{W}} \mathrm{IPAC}^+ \times \mathsf{C}_{\mathbb{N}} \leq_{\mathrm{W}} \mathrm{IPAC}^+_{<\infty} \leq_{\mathrm{W}} \mathrm{IPAC}^+ * \mathsf{C}_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{R}}$。

最后，我们想简要讨论引言中提出的第三个问题（C），即如果从 PAC 学习问题的输出端省去样本复杂度，会发生什么。我们只考虑 IPAC^+，因为这是我们所考虑的最弱的自然非可计算 PAC 学习问题。我们将 IPAC^+_- 样本记为 IPAC^+ 的一个版本，其中去掉了输出中的样本复杂度。IPAC^+_- 样本 比 IPAC^+弱多少？如果输出只是一个没有任何样本复杂度保证的学习器，那么该学习器在小规模样本上可以做任何事情，因此，对于定义域中的任意输入（$\mathcal{C}$，d），输出集合 IPAC^+_- 样本($\mathcal{C}$，d) 都是 $\mathcal{L}$ 的稠密子集。因此，IPAC^+_- 样本($\mathcal{C}$，d) 在 [16] 的意义下稠密实现；由 [16, 命题 4.3]，我们得到以下结论。

推论 69. 反链选择$_{\mathbb{N}} \not\leq_{\mathrm{W}} \mathrm{IPAC}^+_-$ 样本。

如果去掉输出端的样本复杂度，那么这同样适用于我们考虑过的所有其他 PAC 学习问题。从这个意义上说，它们会变得非常弱，因为在一个明确的意义上，无法从中提取任何离散信息。这与一种直觉相符：从实际角度来看，没有样本复杂度的学习器毫无用处。

积极的一面是，我们可以从定理 53 的证明中提取一些信息。在该证明中，根据数对 (A，m) 计算见证函数时，样本复杂度 m 只需在单个输入上进行计算，即可确定 $n = m\langle 4, 3\rangle$。如果还给出了概念类 $\mathcal{C} \in \mathcal{A}_+$，那么我们可以直接猜测数字 n，并使用命题 61 验证见证函数 $f\colon \mathbb{N}^k \to \{0,1\}^k$。如果见证函数未通过验证，我们就用 $n+1$ 替换 n，然后重新开始。由此得到以下结果。

命题 70. 给定一个非空且 VC 维数有限的概念类 $\mathcal{C} \in \mathcal{A}_+$，以及一个学习器 $A\colon \mathcal{S} \to 2^{\mathbb{N}}$，它可以不适当地学习 $\mathcal{C}$，我们可以通过有限心智变化计算确定 $\mathcal{C}$ 的一个见证函数$f \in \mathcal{W}$。因此，我们也可以通过有限心智变化计算找到另一个学习器 $A'\colon \mathcal{S} \to 2^{\mathbb{N}}$，以及一个（可计算的）样本复杂度 m，使得 A' 能以不适当的方式学习 $\mathcal{C}$，且满足 m。

由此得到以下推论。

推论 71. $\mathrm{IPAC}^+ \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} * \mathrm{IPAC}^+_-$ 样本。

结合推论 67、$\mathsf{C}'_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} * \mathsf{C}'_{\mathbb{N}}$ 这一事实（例如可由 [43,定理 7.2]推出），以及 $\mathrm{PA} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} * f \implies \mathrm{PA} \leq_{\mathrm{W}} f$这一事实（由于有限心智变化计算只能产生图灵可归约于相应实例的解，因此该事实成立），我们得到以下结论。

推论 72。 $\mathrm{PA} \leq_{\mathrm{W}} \mathrm{IPAC}^+_-$ 样本 和 $\mathrm{WKL} \leq_{\mathrm{W}} \mathsf{C}'_{\mathbb{N}} * \mathrm{IPAC}^+_-$ 样本。

从某种意义上说，这一结果再次证实，PAC学习不可计算性的关键来源是学习器，而非样本复杂度。事实上，如果我们将样本复杂度 $m \in \mathbb{N}^{\mathbb{N}}$ 替换为自然数 $m\langle 4, 3\rangle \in \mathbb{N}$，我们的所有结果仍然成立，因为样本复杂度

在我们的任何证明中，都从未用于大于 $\langle 4,\ 3\rangle$ 的输入（并且由于我们可以假设$m\langle 4,\ 3\rangle \geq m\langle 2,\ 2\rangle$，其中 $\langle 2,\ 2\rangle$ 是对 m 求值的唯一另一个输入）。

我们还可以提出一个推论，用于讨论属于可计算可枚举闭概念类的学习器的复杂度。我们的结果强化了 Sterkenburg [45, 定理 3]所证明的结果，即存在 VC维数1 的 RER 概念类 $\mathcal{C}$ 不允许可计算学习器。这一结果是命题 63 和定理 53 的证明的推论。由于我们只需要利用单个输入上的样本复杂度$m\langle 4,\ 3\rangle$ 在该证明中，全部不可计算性都包含在学习器中，因此我们再次无需在此提及样本复杂度。

推论 73（PA 度的非适当学习器）。 存在一个 VC维数1 的可计算可枚举闭概念类 $\mathcal{C} \subseteq 2^{\mathbb{N}}$，使得 $\mathcal{C}$ 的每个见证函数以及每个学习器 $A : \mathcal{S} \to 2^{\mathbb{N}}$（其非适当地学习 $\mathcal{C}$）都具有 PA 度。该类 $\mathcal{C}$ 必然具有无限有效 VC 维数。

另一个方向的正面结果见推论 39，该推论同样适用于非适当学习器（即相对于 $\mathcal{H} = 2^{\mathbb{N}}$ 的学习器）。

7. Conclusions

本文研究了以下问题的可计算性方面：给定 一个 VC 维数具有上界的概念类，寻找 一个针对该概念类且具有相应样本复杂度的学习器。我们强调，我们是从统一视角考察这一问题的，这意味着我们关注的是寻找学习器的复杂度（以及相应的样本复杂度）。此前已有的非一致结果则更关注可计算学习器的存在性，以及能够保证这类学习器存在的条件。我们的一些结果表明，可以将见证函数视为概念类的额外输入信息。它们包含的信息多于单纯的 VC 维数上界。图5中的图示从这一视角总结了我们的结果。

该图展示了足以获得经验风险最小化器，或分别获得具有相应样本复杂度的适当或非适当 PAC 学习器的计算成本。从概念类的某些类型信息出发，可以通过在图中选取成本效率最高的路径来计算这些成本，其中所有成本都需要与相应 Weihrauch 度的复合乘积 $*$ 进行复合。根据我们的结果，这通常会为相应问题给出相当紧的上界。总体假设是，我们处理的是具有有限 VC 维数的非空闭概念类 $C \subseteq 2^{\mathbb{N}}$。

对逆向数学 [42, 33, 28]感兴趣的读者可能会问，我们的结果如何转化为二阶算术中的分类。虽然某些转换是可行的 [46]，但并不存在可以让我们自动转换这些结果的直接元定理。逆向数学是一种证明论方法，旨在寻找足以（或许也必要）证明二阶算术中某些结果的公理系统。

据我们所知，统计学习基本定理尚未在逆向数学中得到分析。在许多情况下，

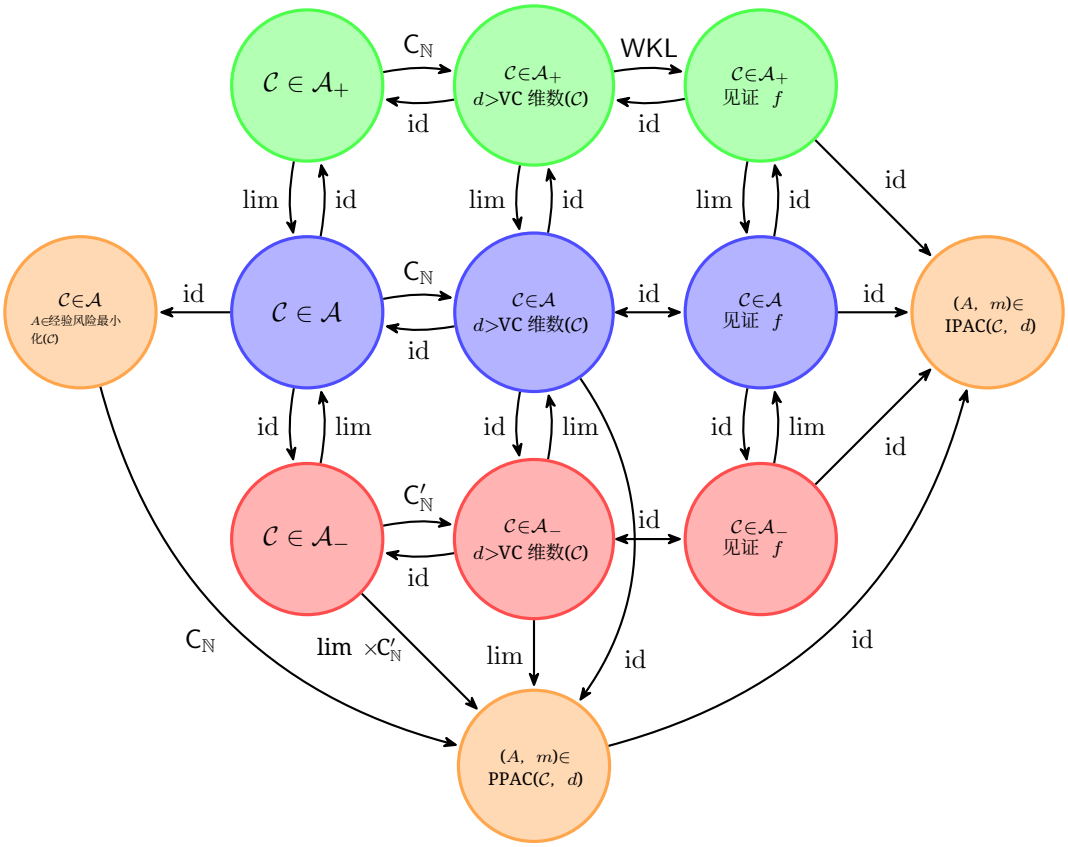

图5: 有限 VC 维数闭概念类的 PAC 学习成本 $\mathcal{C} \subseteq 2^{\mathbb{N}}$ n.

Weihrauch 度逆向数学系统

lim	ACA_0
WKL	WKL_0^*
WWKL	弱弱柯尼希引理*
$\mathsf{DNC}_{\mathbb{N}}$	DNR
$\mathsf{C}_{\mathbb{N}}$	$\mathbf{I\Sigma^0_1}$
$\mathsf{C}'_{\mathbb{N}}$	$\mathbf{I\Sigma^0_2}$

图6: Weihrauch 复杂度与逆向数学。

结果最终体现为资源敏感的形式，以及逆向数学中相应分类的统一版本。在这些情况下，图6中表格的对照表给出了 Weihrauch度与逆向数学体系之间的对应关系。[4]

在我们的研究中，有一个尚未回答的问题：从正信息进行非适当 PAC 学习，是否取决于概念类的 VC 维数。

问题 74. 对于所有 $d > 1$，$\mathrm{IPAC}^+_{<d} <_{\mathrm{W}} \mathrm{IPAC}^+_{<d+1}$ 是否成立?

因此，我们未能区分图4示意图中的所有问题。我们甚至没有证明，对于 $n > 2$，IPAC^+ 不能归约到 DNC_n 。

[4]该系统 ms WKL_0^* 和弱弱柯尼希引理$^*_0$ 指的是不含 Σ^0_1归纳法$\mathbf{I\Sigma}^0_1$的这些公理系统的版本。

Q问题 75。是否存在 $d > 1$ 和 $n > 2$，使得 $\mathrm{IPAC}^+_{<d} \leq_{\mathrm{W}} \mathrm{DNC}_n$ 成立 ？

尽管显然 $\mathrm{IPAC}^+ \leq_{\mathrm{W}} \mathrm{DNC}_2$ 且 $\mathrm{DNC}_n \not\leq_{\mathrm{W}} \mathrm{IPAC}^+$，但我们尚无任何结果能够帮助回答上述问题。

致谢。 第一作者获得了南非国家研究基金会（NRF）的资助，资助编号为 151597。第二作者获得了欧盟伊拉斯谟计划的资助。我们还感谢亚历山大·冯·洪堡基金会提供的支持。我们谨向 Emmanuel Rauzy 致谢，感谢他参与了富有帮助的讨论。

References

[1] N. Ackerman、J. Asilis、J. Di、C. Freer 和 J.-B. Tristan。连续特征的可计算PAC学习。载于 *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science*，LICS '22，美国纽约州纽约，2022年。计算机协会。[2] S. Agarwal、N. Ananthakrishnan、S. Ben-David、T. Lechner 和 R. Urner。论可计算学习器的可学习性。载于 A. Kontorovich 和 G. Neu 编，*Proceedings of the 31st International Conference on Algorithmic Learning Theory* 117 *Proceedings of Machine Learning Research* 48 60，第 卷，页码 –。PMLR，2020年2月8日至11日。[3] S. Akbari 和 M. Harrison-Trainor。自然假设类的可计算学习。载于 N. Haghtalab 和 A. Moitra 编，*Proceedings of Thirty Eighth Conference on Learning Theory*，*Proceedings of Machine Learning Research*第291卷，页码2–21。PMLR，2025年6月30日至7月4日。[4] A. Blumer、A. Ehrenfeucht、D. Haussler 和 M. K. Warmuth。可学习性与 Vapnik-Chervonenkis 维数。*Journal of the Association for Computing Machinery*，36(4):929–965，1989年。[5] V. Brattka。有效博雷尔可测性与函数可约性。*Mathematical Logic Quarterly*，51 1 19( ): –44，2005年。[6] V. Brattka。图灵跳跃与极限之间的伽罗瓦联络。*Logical Methods in Computer Science*，14 3 13 1 37，2018年（ :): – Aug. 。[7] V. Brattka。Weihrauch 复杂度与哈根可计算分析学派。载于 B. L¨owe 和 D. Sarikaya 编，*60 Jahre DVMLG*，*Tributes*第48卷，页码13–44。College Publications，伦敦，2022年。[8] V. Brattka。论学习程序的复杂度。载于 G. Della Vedova、B. Dundua、S. Lempp 和 F. Manea 编，*Unity of Logic and Computation*，*Lecture Notes in Computer Science*第13967卷，页码166–177，沙姆，2023年。施普林格。第19届欧洲可计算性会议。[9] V. Brattka、M. de Brecht 和 A. Pauly。闭选择与统一低基定理。*Annals of Pure and Applied Logic* 163 986 1008, 2012 , : –。[10] V. Brattka 和 G. Gherardi。可计算度量空间上拓扑运算的博雷尔复杂度。*Journal of Logic and Computation*，19(1):45–76，2009年。[11] V. Brattka 和 G. Gherardi。可计算分析中的有效选择与有界性原理。*The Bulletin of Symbolic Logic* 17 1 73 117, 2011 , ( ): –。[12] V. Brattka 和 G. Gherardi。Weihrauch度、全知原理与弱可计算性。*Journal of Symbolic Logic* 76 1 143 176, 2011 , ( ): –。[13] V. Brattka、G. Gherardi 和 R. H¨olzl。概率 *Information and Computation* 242 249 286, 2015 可计算性与选择。, : –。[14] V. Brattka、G. Gherardi 和 A. Marcone。博尔查诺-魏尔斯特拉斯定理是弱柯尼希引理的跳跃。*Annals of Pure and Applied Logic*，163:623–655，2012年。[15] V. Brattka、G. Gherardi 和 A. Pauly。可计算分析中的 Weihrauch 复杂度。载于 V. Brattka 和 P. Hertling 编，*Handbook of Computability and Complexity in Analysis*，《可计算性理论与应用》，页码367–417。施普林格，沙姆，2021年。[16] V. Brattka、M. Hendtlass 和 A. P. Kreuzer。论可计算性理论的统一计算内容。*Theory of Computing Systems*，61(4):1376–1426，2017年。[17] V. Brattka 和 P. Hertling 编。*Handbook of Computability and Complexity in Analysis*，《可计算性理论与应用》，沙姆，2021年。施普林格。

[18] V. Brattka、R. H¨olzl 和 R. Kuyper。蒙特卡罗可计算性。载于 H. Vollmer 和 B. Vall´ee 编, *34th Symposium on Theoretical Aspects of Computer Science (STACS 2017)* 66 *Leibniz* , 第 卷, *International Proceedings in Informatics (LIPIcs)* 17 1 17 14, 2017 页码 : – : 德国达格斯图尔, 。Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik。[19] V. Brattka 和 A. Pauly。论 Weihrauch度的代数结构。 *Logical Methods in Computer Science* 14 4 4 1 36, 2018 , (:): –。

[20] V. Brattka 和 G. Presser。度量空间子集上的可计算性。 *Theoretical Computer Science*，305 43 76，2003年: –。[21] W. Calvert。PAC学习、VC 维数与算术层级。 *Archive for Mathematical Logic*，54 7 8 871 883，2015年（ -): –。[22] V. Cipriani。 *Many Problems Different Frameworks Classification of Problems in Computable Analysis* , ,。博士 *and Algorithmic Learning Theory* 2023 论文，乌迪内大学，。[23] T. Crook、J. Morgan、A. Pauly 和 M. Roggenbach。机器学习（经验证）的可计算性视角。载于 A. Madeira 和 M. A. Martins 编， *Recent Trends in Algebraic Development Techniques*，页码63–80，沙姆，2023年。施普林格·自然瑞士。[24] M. de Brecht。 *Topological and Algebraic Aspects of Algorithmic Learning Theory*。博士论文，京都大学信息学研究生院，2010年。[25] M. de Brecht。拟波兰空间。 *Annals of Pure and Applied Logic*，164(3):356–381，2013年。[26] M. de Brecht 和 A. Yamamoto。概念空间的拓扑性质（完整版）。 *Information and Computation* 208 4 327 340, 2010 , ( ): –。

[27] V. Delle Rose、A. Kozachinskiy、C. Rojas 和 T. Steifer。寻找见证或打散：可计算PAC学习的全貌。载于 G. Neu 和 L. Rosasco 编, *Proceedings of Thirty Sixth Conference on Learning Theory* 195 *Proceedings of Machine Learning Research* 511 524 , 第 卷, 页码 –。PMLR，2023年7月12日至15日。[28] D. D. Dzhafarov 和 C. Mummert。 *Reverse Mathematics*。《可计算性理论与应用》。施普林格，2022年。[29] M. E. Gold。极限递归。 *The Journal of Symbolic Logic*，30(1):28–48，1965年。[30] P. Gourdeau、L. Tosca 和 R. Urner。论鲁棒PAC学习的可计算性。载于 S. Agrawal 和 A. Roth 编， *Proceedings of Thirty Seventh Conference on Learning Theory*, *Proceedings of Machine Learning Research*第247卷，页码2092–2121。PMLR，2024年6月30日至7月3日。[31] K. Higuchi 和 T. Kihara。Muchnik度 I：不连续性、可学习性与构造主义。*Annals of Pure and Applied Logic*，165(5):1058–1114，2014年。[32] K. Higuchi 和 A. Pauly。Weihrauch 可约性的度结构。 *Logical Methods in Computer Science* 9 2 2 02, 17, 2013 , ( ): :。

[33] D. R. Hirschfeldt。 *Slicing the Truth: On the Computable and Reverse Mathematics of Combinatorial Principles* 28 *Lecture Notes Series Institute for Mathematical Sciences National University* , 第 卷, ,。*of Singapore* 2015 世界科学出版社，新加坡，。[34] Z. Iljazovi´c 和 T. Kihara。度量空间子集的可计算性。载于 V. Brattka 和 P. Hertling 编， *Handbook of Computability and Complexity in Analysis*，《可计算性理论与应用》，页码29–69。施普林格，沙姆，2021年。[35] C. G. Jockusch, Jr. 和 R. I. Soare。 Π^0_1 类与理论 *Transactions of the American Mathematical Society* 173 33 56, 1972 的度。，: –。[36] D. Kattermann 和 L. S. Krapp。递归可枚举表示的类与统计学习基本定理的可计算版本。arXiv 2511.02644，2025年。[37] A. S. Kechris。 *Classical Descriptive Set Theory*， *Graduate Texts in Mathematics*第156卷。施普林格，柏林，1995年。[38] M. Mohri、A. Rostamizadeh 和 A. Talwalkar。 *Foundations of machine learning*。自适应计算与机器学习。麻省理工学院出版社，马萨诸塞州剑桥，第二版，2018年。[39] E. Neumann 和 A. Pauly。代数计算模型的拓扑视角。 *Journal of Complexity*，44 1（补编 C）：–22，2018年。[40] M. Schaefer。判定 Vapnik–ˇCervonenkis 维数是 Σ^p_3 完全的。， *Journal of Computer and System Sciences* 58 1 177 182, 1999 ( ): –。[41] S. Shalev-Shwartz 和 S. Ben-David。 *Understanding Machine Learning: From Theory to Algorithms*。剑桥大学出版社，纽约，2014年。

[42] S. G. Simpson。 *Subsystems of Second Order Arithmetic*。数学逻辑展望丛书。施普林格，柏林，1999年。
[43] G. Sold
、 *Annals of Pure and* a 和 M. *Applied Logic* 174 7 103270, 2023
Valenti。问题一阶部分的代数性质。， （）：。[44] D. Soloveichik。任意可计算分类器的统计学习。arXiv 0806.3537，2008年。[45] T. F. Sterkenburg。论可计算学习器下可学习性的刻画。载于 P.-L. Loh 和 M. Raginsky（编）， *Proceedings of Thirty Fifth Conference on Learning Theory*, *Proceedings of Machine Learning Research*第 178 卷，第 3365–3379 页。PMLR，2022年7月2日至5日。[46] P. Uftring。包含 E-PA$^{\omega}$ + QF-AC^{0,0} 的系统中 Weihrauch 可约性的刻画。*The Journal of Symbolic Logic*, 86(1):224–261，2021年。[47] V. N. Vapnik 和 A. Y. Chervonenkis。事件相对频率一致收敛到其概率。 *Theory of Probability and Its Applications*，16(2):264–280，1971年。[48] K. Weihrauch。实数表示之间某些转换器的不连续度。技术报告 TR-92-050，国际计算机科学研究所，伯克利，1992年7月。[49] K. Weihrauch。 *Computable Analysis*。施普林格，柏林，2000年。[50] M. Ziegler。实数超计算与连续性。 *Theory of Computing Systems*，41(1):177–206，2007年。

Appendix: A result on effective VC dimension

在 [36, 问题 6.1] Kattermann 和 Krapp 曾提出，是否存在一个 PAC 可学习类 $\mathcal{C}$，它有一个可计算的非适当学习器和一个不可计算的假设 $h \in \mathcal{C}$。本附录旨在肯定地回答这一问题（使用定理 7）。

命题 76. 存在一个假设的可计算序列 $(h_i)_{i\in\mathbb{N}}$： $\mathbb{N} \to \{0,1\}$，该序列收敛到一个不可计算的、具有停机问题的图灵度的 h： $\mathbb{N} \to \{0,1\}$，并且 $\mathcal{H} = \{h\} \cup \{h_i : i \in \mathbb{N}\}$ 构成一个有效 VC 维数为 1 的假设类。

证明. 令 $r : \mathbb{N} \to \mathbb{N}$ 为停机问题 $K = \mathrm{range}(r)$ 的一个可计算枚举。令 $R_i := \{r(0), ..., r(i)\}$, $m_i :=$ 最大值(R_i)，且 $v_{i,j} \in \{0,1\}^*$ 定义如下：

$$v_{i,j} := \begin{cases} 1 \cdot 0 & \text{if } j \in R_i \\ 0 \cdot 1 & \text{otherwise} \end{cases}$$

对于所有 $i \in \mathbb{N}$ 和 $j \le m_i$。我们定义一个形如以下形式的可计算序列 $(h_i)_{i\in\mathbb{N}}$：

$$h_i := u_i \cdot \widehat{0} \text{ with } u_i := v_{i,0} \cdot 0^{k_{i,0}} \cdot v_{i,1} \cdot 0^{k_{i,1}} \cdot ... v_{i,m_i-1} \cdot 0^{k_{i,m_i-1}} \cdot v_{i,m_i}$$

其偶数值为 $k_{i,j} \in \mathbb{N}$，这些值通过关于 $i \in \mathbb{N}$的归纳构造来定义。第 $i = 0$ 阶段的归纳从 $k_{0,j} := 0$（对所有 $j < m_0$）开始。在归纳步骤 $i \to i+1$中，假设我们已经定义了 $h_0, ..., h_i$，相应的字为 $u_0, ..., u_i$，偶数值为 $k_{0,j}, ...k_{i,j}$。我们需要定义 h_{i+1}。若 $r(i+1) \in R_i$，则令 $h_{i+1} := h_i$，因此对于所有 $j < m_{i+1} = m_i$，有 $k_{i+1,j} := k_{i,j}$。现在假设 $r(i+1) \notin R_i$，我们需要区分两种子情况。若 $r(i+1) > m_i$，则对于所有满足 $m_i \le j < m_{i+1}$ 的 j，令 $k_{i+1,j} := 0$，并对 $j < m_i$令 $k_{i+1,j} := k_{i,j}$。有趣的情况是 $\iota := r(i+1) < m_i$。在这种情况下，对于所有满足 $\iota \le j < m_i$的 j，令 $k_{i+1,j} := 0$；对 $j < \iota - 1$令$k_{i+1,j} := k_{i,j}$，并为 $k_{i+1,\iota-1}$ 选取一个必为偶数的值，使得

$$|u_i| = |v_{i+1,0} \cdot 0^{k_{i+1,0}} \cdot ... \cdot v_{i+1,\iota-1} \cdot 0^{k_{i+1,\iota-1}}|.$$

换句话说，u_{i+1} 与 u_i 有一个公共前缀，其中包含块$v_{i+1,\,\iota-1}$，但$v_{i+1,\,\iota} \neq v_{i,\,\iota}$。在公共块 $v_{i+1,\,\iota-1}$ 之后，字 u_{i+1} 用零填充，直到长度达到u_i，然后继续接上$v_{i+1,\,\iota}$。由于 K 是一个无限集合，因此可以保证序列$(|u_i|)_{i\in\mathbb{N}}$ 无界且单调递增。

这个构造确保，如果某个位置 $n < |u_i|$上的比特从 h_i 变为 h_{i+1}，那么它只能从1变为0，而绝不会从0变为1。如果位置$n < |u_i|$ 上的比特从1变为0，那么 h_i 在满足 $n < k < |u_i|$ 的位置 k 上的所有其他比特也都会被设为0。新的比特1只能在位置 $n \ge |u_i|$引入。这个性质使我们能够如下计算见证 w： $\mathbb{N}^2 \to \{0,1\}^2$，其中 $\{h_i : i \in \mathbb{N}\}$。给定两个自然数 $n < k$，我们首先确定某个满足 $|u_m| > k$ 的 $m \in \mathbb{N}$，然后设置

$$w(n,k) := \begin{cases} (0,1) & \text{if } (\exists i \le m)\ (h_i(n), h_i(k)) = (1,1) \\ (1,1) & \text{otherwise} \end{cases}.$$

We e解释在两种情况下，为什么这样选择见证函数的取值是可行的：
(1) 如果存在 $i \le m$，使得 $(h_i(n)\ h_i(k)) = (1, 1)$，那么，根据上述归纳构造，位置 n 处 h_i 中的比特 1 之后不能被 $h_j(n) = 0$ 替换，而不同时将位置 k 处的比特 1 替换为$h_j(k) = 0$。同样，也不可能存在更早的假设 h_j，使得 $j < i$ $h_j(n) = 0$ 且 $h_j(k) = 1$，因为这将推出 $|u_j| > k$，从而位置 n 处的 0 不可能被改为$h_i(n) = 1$。因此，对于所有 $i \in \mathbb{N}$，均有 $(h_i(n)\ h_i(k)) \neq (0, 1)$。

(2) 如果对于所有 $i \leq m$，$(h_i(n)\ h_i(k)) \neq (1,\ 1)$，那么我们断言这对所有 $i \in \mathbb{N}$都成立。这是因为，如果 $h_m(n) = 0$ 或 $h_m(k) = 0$，则相应的比特之后不能被设为 1。由于模式 (1，1) 此前也未出现过，因此我们可以得出，对于所有 $i \in \mathbb{N}$，$(h_i(n)\ h_i(k)) \neq (1,\ 1)$。

总之，我们描述了见证函数 w 的一种计算方法：$\mathbb{N}^2 \to \{0,1\}^2$ for$\{h_i : i \in \mathbb{N}\}$，因此也描述了集合 $\{h_i : i \in \mathbb{N}\}$ 的闭包 $\mathcal{H} = \{h\} \cup \{h_i : i \in \mathbb{N}\}$ 的一种计算方法。假设 h 属于该闭包，因为构造保证序列 $(h_i)_{i\in\mathbb{N}}$ 收敛到 h，其图灵度与停机问题相同，因为 h 的形式为

$$h = v_0 \cdot 0^{2k_0} \cdot v_1 \cdot 0^{2k_1} \cdot v_2 \cdot 0^{2k_2} ...,$$

其中 $v_j = 1 \cdot 0 \iff j \in K$，否则为 $v_j = 0 \cdot 1$。 □