自指与相对不完备性：从逻辑障碍到困难性假设

JieLuo Ke Xu

复杂关键软件环境国家重点实验室，北京航空航天大学，北京100083，中国{luojie，kexu}@buaa.edu.cn

2026年11月25日

arXiv:2611.32095v1 [cs.LO] 24 Nov 2026

摘要

我们定义了相对不完备性，并利用它为哥德尔不完备性与有限 SAT 中局部观察的局限性提供了一个统一框架。如果与某一背景相容的两个实现对目标问题给出相反答案，则称该背景相对于该目标问题是不完备的。哥德尔对角化产生一个语句，使其真值在同一理论的两个模型中不同；而依赖解的阻塞则产生两个在已观察到的子句上相同、但在可满足性方面不同的公式。这两种构造都实现了真与接受之间的共同对角关系，迫使可靠的部分判定器对目标问题无法作出决定。

对于一个可满足公式，在不改变其变量支撑集的情况下替换一个未观察子句，当且仅当删除该子句后剩余的所有解在其支撑集上都一致时，才能使该公式变得不可满足。如果原公式具有唯一解，则该条件等价于该子句具有冗余性。更一般地，对一个可满足公式的观测，只有在其包含一个位置，而该公式在该位置上不同于每个不可满足补全时，这一观测才能确切确定可满足性。对于唯一可满足公式，每个充分观察都必须包含的位置，恰好就是那些冗余位置。我们还将解阻断算术化，并构造一个指涉自身阻塞数据的带标签不动点。最后，我们使用自指的 K-SAT 来说明，相对不完备性如何为表述定量困难性假设提供逻辑基础。

关键词： 自指；相对不完备性；局部观察；算术化；解阻断。

1 引言

何时已有的事实足以回答一个问题？如果两个完备情形与所有这些事实都一致，却给出相反的答案，那么该答案就不可能仅由这些事实推出。本文研究的问题是：一种构造如何保证，对于指定类别中的每个背景，都能产生这种失败。将标准自然数中为真的所有语句作为公理，得到完备理论 Th(ℕ)。它不是可计算枚举的。哥德尔定理则适用于

具有足够算术强度的有效公理化理论，并满足适当的一致性或可靠性假设。在有限情形中，完整输入同样决定可满足性，但下文的阻塞构造使每一种允许的局部观测都与相反答案相容。两种情形的限制有所不同：一种限制有效公理化，另一种限制对输入部分内容的访问。共同的要求是在指定类别的背景中始终构造出相反的实现。

哥德尔构造使一个语句断言其自身在指定理论中不可证 [G31]；唯一解阻塞使公式的唯一满足赋值决定一个排除该赋值的约束

，[FLL+26]。在各自的假设下，这两种构造都使保留的背景无法判定一个目标断言：该理论无法判定哥德尔句子，而局部观察无法判定完整公式的可满足性。对于每个符合条件的理论，语句都是由该理论的证明谓词构造的。对于固定母公式的每个允许的观察，都可以替换一个未观察子句，同时反转可满足性。我们要探讨的是，这些构造如何实现语义真值与接受之间的共同关系，以及何时能够通过保留观察的局部修改来实现有限实现。

我们采用元数学视角，将理论、观察背景以及它们支撑正确判断的能力作为研究对象。问题在于：即使每个完整实现都有一个确定的有限目标值，从指定背景中仍能推出哪些结论。通过确定给定信息体系留下哪些问题未决，相对不完备性为 formulates 困难性假设提供了逻辑起点。

1.1 一个有限例子

例 1.1。假设一个公式在变量 x、y 上有四个带索引的子句，每个子句都使用这两个变量。前三个子句已经被观察到：

$$Q = (x \vee y) \wedge (x \vee \neg y) \wedge (\neg x \vee y).$$

它们分别排除了 (0、0)、(0、1) 和 (1、0)，因而留下唯一的赋值 (1、1)。第四个子句尚未被观察到。两个可能的完整公式是

$$F = Q \wedge (x \vee y), \qquad F^{-} = Q \wedge (\neg x \vee \neg y).$$

第一个保留了解 (1、1)；第二个排除了它，因此不可满足。二者具有相同的前三个子句，并且在每个位置上具有相同的变量支撑集。

观察已经确定了唯一可能的解，但尚未确定未观察子句是否允许该解。观察事实的任何后果都无法区分F 和 F^{-}，因为两个完整公式都满足这些事实。读取第四个子句即可解决这个问题。

替换项选自母公式自身的解：(1，1) 确定了子句 $\neg x \vee \neg y$。原来的第四个子句与第一个重复，因此删除它不会产生新的解。这两点说明了为什么替换会逆转可满足性。其未观测的位置则解释了为什么现有信息得以保留。

我们将这种无法确定的情形称为相对不完备性。同一定义既适用于模型对某个语句意见不一的算术理论，也适用于相容完备化对可满足性意见不一的局部 SAT 观测。有些部分观测确实能够判定可满足性：只要看到两个单位子句 x 和 $\neg x$，就足以确定其不可满足性，无论其余子句是什么。我们的任务是找出这样的构造：对于一个明确指定的类别中的每一种观测，都能保证存在相反的完备化。一个含糊的观测可以说明这种机制；一般性结果则必须证明全称量词成立。

1.2 背景及相关工作

算术方面的论述始于哥德尔 [G31]。Li [L14，第 5.1 节和第 5.3] 节阐述了如何从命题层面的表达式 $X \leftrightarrow B(X)$ 过渡到关于语句数值编码的公式，并解释了为何不动点构造是一种有限代换过程，而非循环定义。Boolos、Burgess 和 Jeffrey [BBJ07] 则提供了下文所用的可表性、完备性与不完备性的逻辑背景。

Lawvere [L69] 通过求值与不动点统一了对角论证。Yanofsky [Y03]给出了集合论表述及诸多应用。在其初等形式中，

构造从 e 开始： $A \times A \to Y$ 以及一个无不动点映射 α ： $Y \to Y$，并构造$a \mapsto \alpha(e(a,\ a))$。所得函数无法用一列 e 表示。该表述适用于有限集和无限集。我们给出一个明确的求值映射，将这两种构造置于这一集合论模式中。随后，有限情形的问题转向其实现：哪些解结构能够通过改变一个子句位置，同时保持指定观测不变，实现语义反转?

第二条研究路线探讨算术句子必须如何做，才能将某种性质归于自身。Halbach 和 Visser [HV14] 考察了算术自指中的这一问题。Grabmayr、Halbach 和 Ye [GHY23] 研究了对角算子和编码的约束，这些约束将范式性的自指与偶然的对角句子区分开来。这些分析与第5节中的带标签的对角构造相关；在该构造中，对句子编码的引用被用于提取固定阻塞数据。

Popescu 和 Traytel [PT19] 提出了哥德尔不完备定理的抽象条件，并在 Isabelle/HOL 中对其进行了验证。他们对句法、可表性和语义假设的分离说明了为什么共同的对角步骤必须辅以确立其实现的条件。在此处，有限实现由编辑限制和解集的投影所支配。

解阻塞构造源自自指的 K-SAT [FLL+26]。我们采用其中的确定性要素，并证明这里所需的蕴含关系。提供合适的对数宽度母公式的概率分析仍见于该工作。我们的论证适用于每个满足所述唯一性、冗余性和观测假设的母公式。

1.3 结果与结构

共同的证明步骤是引理 2.6：其真值是其接受指示函数的补的对象，无法由可靠的部分判定器判定。第 3 节使用哥德尔不动点表明，所述类别中的每个理论 T 相对于其哥德尔句子而言都是不完备的 G_T。第 4 节使用解阻断表明，只要局部观测理论 $\Gamma_S(F)$遗漏了一个可以进行阻塞的位置，它相对于可满足性断言 A_H 就是不完备的。 S

有限结果精确确定了何时可以实现这一构造。定理 4.1 刻画了通过保留解集的单元素投影实现的单子句阻塞。定理 4.3表明，这是通过一次未观测的、保持支撑集的编辑实现对角线反转的充要条件。对于唯一解，该条件化简为冗余性。定理 4.6通过与不可满足完备化的差异集的交集刻画了所有充分观测。推论 4.7 确定了每个充分观测都必须包含的位置。

量词至关重要。每个可靠的可计算枚举扩张 T 皮亚诺算术都有一个独立句子 G_T。对于每个固定的唯一可满足 F，凡是让某个冗余位置不可见的观测，都有一个不可满足补全，因此，同一个可满足性断言仍独立于每一种此类观测理论。算术目标随理论而变；有限目标保持不变，而伴随对象可能变化。

第 6 节借助 Li 对形式语言与其元语言环境的区分，阐明自我识别障碍。这一区分也将算法设计视为一种数学活动，并将对问题复杂性的研究视为元数学的。相对不完备性通过指出所有受限于相同信息的方法所共有的逻辑障碍，为这一区分赋予了具体形式。随后，我们用自指 K-SAT 说明，可证明的信息障碍如何促成定量困难性假设。

第2 节提供语义框架和通用证明步骤。第5 节将两种构造都与集合论对角模式联系起来，用算术表达解阻断，并构造一个从自身编码中读取阻塞数据的自指句子。该

初等对角障碍继承自既有框架；有限结果具体说明了其在观测和编辑约束下的实现。带标签的算术句子有着不同的用途：它引用自身存储的阻塞数据。

2 上下文与相对不完备性

2.1 上下文与实现

这个例子区分了三个对象：保留下来的事实、与这些事实相容的完全公式，以及可满足性问题。我们将它们分别称为上下文、其实现和目标。对于算术理论，上下文就是该理论，其实现就是模型，而目标就是某个语句的真值。以下定义仅记录这些共有特征。

定义 2.1。 一个语义上下文系统是一个元组

$$(\mathcal{X}, \mathcal{I}, \mathcal{C}, P),$$

其中，$\mathcal{X}$ 是实现的集合，$\mathcal{I}$ 是上下文的集合，$\mathcal{C}(I) \subseteq \mathcal{X}$ 是 I的实现集，且$P\colon \mathcal{X} \to \{0, 1\}$ 是一个目标谓词。若$\mathcal{C}(I) \neq \varnothing$，则称上下文可容许。若上下文分离 P，则

$$\mathcal{C}(I) \cap P^{-1}(0) \neq \varnothing, \qquad \mathcal{C}(I) \cap P^{-1}(1) \neq \varnothing. \tag{1}$$

见证（1）的一对称为关于 I的语义分离。

在例 中，1.1， I 记录前三个子句和全部四个支撑集，$\mathcal{C}(I)$ 由可能的第四个子句的符号模式组成，而 P 询问由此得到的公式是否可满足。公式 F 和 F^{-} 证明了（1）。

有两种实用的表述方式。对于观察映射 $V : \mathcal{X} \to \mathcal{Y}$，取上下文$y \in V(\mathcal{X})$ 和 $\mathcal{C}(y) = V^{-1}(y)$。集合 $V^{-1}(y) = \{x \in \mathcal{X} : V(x) = y\}$ 由所有与观测 y相容的实现构成；按标准术语，它是纤维 V 在 y上的原像。在有限公式的设定中，其中的元素是与已观察到的子句相容的完全公式，我们称之为相容完备化。对于逻辑语言，则取句子集作为上下文，取模型作为实现。在算术应用中，论域为 $\mathbb{N}$ 的可数模型就足够：完备性和向下 Löwenheim–Skolem 定理为皮亚诺算术的一致扩展的模型提供了这类表述。它们的算术运算不必是 $\mathbb{N}$上的标准运算。

命题 2.2 (由上下文决定)。 对于可容许上下文 I，当且仅当 P在 $\mathcal{C}(I)$上为常数时，其信息才能恰好决定 P。在观测表示中，当且仅当存在某个函数 $h : V(\mathcal{X}) \to \{0, 1\}$，使得 $P = h \circ V$，也就是说，对每个可达的观测，所有相容实现都具有相同的目标值。

证明。 第一项断言成立，是因为非空集合 $P(\mathcal{C}(I))$ 恰好在不同时包含 0 和 1 时才是单元素集合。对于第二项，分解使得 P 在每个集合 $V^{-1}(y)$ 上为常数。反过来，如果 P 在每个这样的集合上为常数，就将 $h(y)$ 定义为 $V^{-1}(y)$ 上的该共同取值。 □

命题 2.2 涉及一整类决策规则。如果与同一观测相容的两个实现具有相反的目标值，那么任何仅依赖于 V 的规则都不可能在整个 $\mathcal{X}$上正确。反过来，存在分解 $P = h \circ V$意味着语义充分性； h 的可计算性和资源需求则是有待分析的进一步属性。对于有限观测空间， h 可以用有限表表示，但其构造、存储以及跨不同实例规模的统一计算还会带来进一步的资源问题。一个完整公式决定了其可满足性；高效地确定该值则是另一个问题。

从相同背景中推导出更多事实，也无法排除这样一对实现中的任何一个：每个有效推论都必须对二者同时成立。为正式表述这一点，假设上下文是定义在 $\mathcal{X}$ 上、由条件构成的集合。记

$$\mathrm{Con}(I) = \{\varphi : \text{every } x \in \mathcal{C}(I) \text{ satisfies } \varphi\}.$$

命题 2.3 (在推论下的保持性)。 若 $I \subseteq J \subseteq \mathrm{Con}(I)$，则 $\mathcal{C}(J) = \mathcal{C}(I)$。特别地， I 上的每一种语义分离，在加入其任意一组有效推论后仍然成立。

证明。 包含关系 $I \subseteq J$ 给出 $\mathcal{C}(J) \subseteq \mathcal{C}(I)$。 I 的每一种实现都满足J的每个成员，因此反向包含关系也成立。 □

因此，一个上下文及其有效推论具有相同的实现。可靠推理无法排除原有背景的任何实现，尽管它可以使某个隐含推论明确可用。这就是演绎在此处不增加信息的语义意义。推导隐含推论是一项计算任务，其代价可以单独研究。我们对理论和观测记录同样采用保持性。

2.2 相对不完备性

定义 2.4。 若可容许上下文 I 允许对 P进行语义分离，则称其相对于 P 不完备。对于相容的经典理论 Γ 和语句 A，若 Γ 相对于 A不完备，则称之为

$$\Gamma \nvdash A \qquad \text{and} \qquad \Gamma \nvdash \neg A.$$

命题 2.5 （语义与演绎表述）。 在经典命题逻辑或一阶逻辑中，一致理论 Γ 相对于 A 不完备，当且仅当它有满足 A的模型，也有满足 $\neg A$的模型。

证明。 可靠性意味着这两个模型都与证明任一备选命题相矛盾。反过来，若两个备选命题都不可证，则由演绎定理， $\Gamma \cup \{A\}$和 $\Gamma \cup \{\neg A\}$都一致。完备性保证它们存在模型。 □

这一术语涉及理论相对于某个目标所包含的信息。底层演绎演算的完备性保证了它与相反模型之间的等价关系。描述中的某个真子部分可能已经足以确定特定目标；相对不完备性的特征是存在相反实现。

2.3 在一类背景下的保证

针对单个上下文的分离与覆盖整类上下文的保证是不同的断言。为比较后者，令$\mathcal{B}$ 对语义上下文系统中允许的上下文 I_b 和目标 P_b 进行索引，并允许目标依赖于 b。所需的陈述是

$$\forall b \in \mathcal{B}\ \exists X_b^+, X_b^- \in \mathcal{C}(I_b): \quad P_b(X_b^+) = 1, \qquad P_b(X_b^-) = 0. \tag{2}$$

在算术应用中， $b = T$ ，该构造给出了 $P_b = P_{G_T}$。在有限应用中， $b = S$，上下文记录观察到的符号，而 P_b 始终是可满足性。公式 F 为每个 S提供相同的正实现；只有负伴随对象需要变化。

方程（2）记录了针对一类背景所作保证的证明义务。引用机制规定了该保证如何得到实现。在当前应用中需要分析的是：对某个理论的证明谓词或某个公式自身的解进行引用，如何提供目标实现或对立实现，以及为什么该构造对每个合格背景都有效。接下来的两节将分别确立这些保证。

2.4 真与接受

部分求值器可能接受、拒绝某个问题，或对其置于未决状态。其论域是可用记录，而不一定是对象本身。令 $V\colon \mathcal{X} \to \mathcal{I}$ 且 $P\colon \mathcal{X} \to \{0, 1\}$。映射$D\colon \mathcal{I} \to \{1, 0, \bot\}$ 若对每个 x 都成立，则称其为可靠的，

$$D(V(x)) = 1 \implies P(x) = 1, \qquad D(V(x)) = 0 \implies P(x) = 0.$$

记为 $a_D(I) = \mathbf{1}[D(I) = 1]$。因此，拒绝和未作决定的接受指示函数均为零。对 D不作任何有效性假设。

L引理 2.6 (真与接受相反)。 若 D 具有可靠性，且对象 x_D 满足 s

$$P(x_D) = 1 - a_D(V(x_D)), \tag{3}$$

则 $D(V(x_D)) = \bot$ 且 $P(x_D) = 1$。

证明。 接受将意味着 $P(x_D) = 1$，由可靠性可知 $P(x_D) = 0$，由（3）可知。因此不可能接受，而（3）给出 $P(x_D) = 1$。现在可靠性排除了拒绝。 □

我们称（3）为反接受关系。它的实现就是构造问题；该引理提供了共同的排除论证。任意具有可靠性的求值器都可能返回 $\bot$，即使信息已经确定了答案。若要从未作决定推导出语义分离，就需要一种能够记录背景所有推论的机制。

对于相容的经典理论 Γ，定义其关于句子的演绎状态如下：

$$D_\Gamma(A) = \begin{cases} 1, & \Gamma \vdash A, \\ 0, & \Gamma \vdash \neg A, \\ \bot, & \text{otherwise.} \end{cases} \tag{4}$$

命题 2.5 指出， $D_\Gamma(A) = \bot$ 当且仅当背景理论存在使 A 具有两种真值的模型。这里用到的是演绎演算的完备性，而不是理论的完备性。在引理 2.6 的算术应用中，研究对象是句子， V 是恒等映射，而P 是标准真值；相反模型随后由命题 2.5 得出。在有限应用中，研究对象是完整公式， V 是子句观测，而 P 是可满足性。

2.5 两种指称形式

这两种构造使用以下指称形式。

编码引用始于一个表达式的编码以及一个定义在编码上的谓词。通过对角构造闭合该引用，得到 G ，使得 $T \vdash G \leftrightarrow \psi(\ulcorner G \urcorner)$。构造本身，包括所选用的编码和代入操作，都是数据的一部分。

依赖见证的反馈始于两类对象：源 $x \in \mathcal{D}$ 和见证 $w \in \mathcal{W}$，二者由 $E(x, w)$ 关联。如果 x 有唯一见证，则将其记为 $\sigma(x)$。然后将修改指定为

$$x \longmapsto \sigma(x) \longmapsto B(\sigma(x), \theta) \longmapsto K(x, \sigma(x), \theta), \tag{5}$$

其中， θ 记录了修改所允许的位置或支撑。该操作与该源的见证相关联，而不是与独立选取的见证相关联。反馈构造必须分别确定所有见证会发生什么，以及保留了哪些上下文。

对于解阻断， （5）是 [FLL+26] 中称为自指的语义依赖：唯一的满足赋值确定一个约束，并将其插回其母公式中。我们采用这种特定的依赖关系，而不是将对输入性质的任意依赖作为自指的定义。

对于每一种机制，证明都必须确立否定操作的效果以及上下文的保持。这些步骤导出了（1）。

3 哥德尔式自指与相对不完备性 s

3.1 编码之前的思路

在第 2 节的框架中，算术理论 T 是上下文， T 的模型是其实现，而目标是在这些模型中判定一个语句的真。要确立相对不完备性，我们需要一个语句，使其在同一 T的不同模型中真值各异。对于下文所讨论类别中的每个理论，哥德尔构造都会产生这样一个目标。

考虑一个旨在表达“该语句在 T 中没有证明”的语句。假设 T能够验证具体的证明，并且该语句已被正确构造。如果T 证明了它，那么这个具体证明也会让 T 确立它拥有一个证明。该语句断言与此相反的内容，因此一致性排除了这样的证明。于是，该语句陈述了一个关于不存在证明的真实事实。如果 T 在标准自然数中是可靠的，它也不能证明该语句的否定。

在这成为一个数学论证之前，还剩下两项任务。算术必须表达证明是什么，而短语“该语句”必须被替换为对正在构造的语句的精确指称。编码使表达式和证明能够作为自然数来处理。对角引理通过一次明确的代换提供了这一指称。我们先完成这一构造，然后证明独立性，最后利用完备性得到所保留理论的相反模型。可靠性的假设意味着， T 的每个定理在标准自然数中都为真；它强于一致性，足以支持此处所采用的论证。

3.2 编码与不动点构造

固定一种算术表达式的标准有效编码。如果 G 是一个语句， $\ulcorner G\urcorner$ 表示其编码；当它出现在算术公式中时，则表示该编码对应的数词。将数词代入公式是原始递归的。令 diag(u) 表示这样得到的编码：将 u 的数词代入由 u编码的公式的指定自由变量中；对于格式错误的输入，则规定一个固定的默认值。其图具有算术表示 Diag(u, v)，其全定义性和函数性均可在皮亚诺算术中证明。

引理 3.1（对角引理）。 对于每个具有一个自由变量的算术公式 $\psi(z)$ ，存在一个语句 G ，使得

$$\mathsf{PA} \vdash G \leftrightarrow \psi(\ulcorner G\urcorner). \tag{6}$$

证明。 令 $A(u)$ 为 $\exists v(\mathrm{Diag}(u,\ v) \wedge \psi(v))$，令 $a = \ulcorner A(u)\urcorner$，并设 $G = A(\bar{a})$。根据代换的定义，$\mathrm{diag}(a) = \ulcorner G\urcorner$。皮亚诺算术验证了这一具体计算及输出的唯一性，因此它证明了 $A(\bar{a}) \leftrightarrow \psi(\ulcorner G\urcorner)$。 □

这是通常的对角构造；参见 [L14，第5节。3] 以及 [BBJ07]。公式A 先写出，然后再代入它的编码。并不假定任何语句在构造出来之前就已存在。

设 T是皮亚诺算术的一个可计算可枚举扩张。可以使用有限推导记录和公理的枚举阶段证书 T，从而得到一个原始递归证明检查关系。其存在闭包给出标准谓词 $\mathrm{Prov}_T(z)$。将引理 3.1 应用于$\neg\,\mathrm{Prov}_T(z)$ 可得

$$T \vdash G_T \leftrightarrow \neg\,\mathrm{Prov}_T(\ulcorner G_T\urcorner). \tag{7}$$

（7）中的关系是 T中的可证性，而不是任意模型中的真。这一区分对于其一致解释至关重要。

3.3 固定理论上的相对不完备性

一个独立句子为该理论留下两个一致扩展：一个通过添加该句子得到，另一个通过添加其否定得到。完备性将这些扩展转化为模型。因此，这些相反的可能性是共同理论的模型；它们并不表示标准自然数中的真值发生了变化。

定理 3.2（作为相对不完备性的独立性）。 设 T 是一个扩展皮亚诺算术的一致理论，且 G 是一个语句。以下条件等价：

1. $T \nvdash G$ 和 $T \nvdash \neg G$；
2. 存在 T 的可数模型 M_+、M_-，满足

$$\mathcal{M}_+ \models G, \qquad \mathcal{M}_- \models \neg G.$$

因此， G 独立于 T 当且仅当 T 相对于 G不完备。在命题 *2.5*的语义表述中，这两个模型见证了上下文 T 上的分离，其中所用谓词为 $P_G(\mathcal{M}) = \mathbf{1}[\mathcal{M} \models G]$。

证明。若 $T \nvdash \neg G$，则演绎定理意味着 $T \cup \{G\}$ 是一致的。类似地，当 $T \nvdash G$时， $T \cup \{\neg G\}$ 是一致的。由完备性和向下 Löwenheim–Skolem 定理可得这些理论的可数模型。反过来，根据一阶演绎规则的可靠性，如果这两个模型存在，那么 T 不能推出 G 或 $\neg G$ 。 □

定理 3.3（哥德尔相对不完备性）。 设 $\mathcal{T}$ 是皮亚诺算术 PA 在标准自然数中可靠的可计算枚举扩张类。对于每个 $T \in \mathcal{T}$，构造 （7）给出一个句子 G_T，使得 T 相对于 G_T是不完备的。因此

$$\forall T \in \mathcal{T}\ \exists G_T : \quad T \nvdash G_T \quad and \quad T \nvdash \neg G_T. \tag{8}$$

证明。 使用 D_T （4）中的构造，以句子为对象，并令 $P(A) = \mathbf{1}[\mathbb{N} \models A]$。 T的可靠性使该求值器可靠。标准证明谓词满足 $\mathbb{N} \models \mathrm{Prov}_T(\ulcorner A \urcorner)$ 当且仅当 $T \vdash A$。由于（7）在 $\mathbb{N}$中成立，

$$P(G_T) = 1 - \mathbf{1}[T \vdash G_T] = 1 - a_{D_T}(G_T).$$

引理 2.6 给出 $D_T(G_T) = \bot$ 以及 $\mathbb{N} \models G_T$。定理 3.2 提供了 T的两个模型，在这些模型中，G_T 具有相反的真值。它们证明了对于该目标， T 的相对不完备性。 □

理论 T 在所述类中是任意的。该构造并非只是找出一个被某个选定理论遗漏的句子：它利用每个理论自身的证明谓词，构造出一个该理论无法判定的句子。这就是（8）中的普遍保证。对于此处使用的特定哥德尔句子，可靠性假设是一个方便的充分条件。仅假设一致性时，则可以改用罗瑟句子 [R36]。当前构造保留（7），从而使该指称及其否定保持显式。

定理3.3 中的两个模型都满足 T的每个定理。它们并非初等等价：G_T本身就能区分二者。在满足 $\neg G_T$的模型中，谓词$\mathrm{Prov}_T(\ulcorner G_T \urcorner)$ 成立，尽管不存在标准的有限证明。这样的模型是非标准模型，可能包含非标准的所谓证明编码。它的存在并不意味着已经找到了标准证明。

该构造给出了一个独立语句，并通过完备性得到了取值相反的模型。它并未指定一种有效变换，能将 T 的任意模型转变为一个使 G_T取值相反的模型。分离是关于某个共同理论的存在性陈述。第4 节将同一判据应用于把已观察到的子句作为保留的信息。在那里，两种实现是相容完备化，而将它们区分开的目标是可满足性。

4 有限解阻断与相对不完备性

我们首先刻画在保留观测的同时，何时可以替换一个未观察子句，使可满足公式变为不可满足公式。然后，我们利用这一构造确立相对不完备性，并准确判定哪些观测足以决定可满足性。

4.1 单子句阻断的精确条件

例 1.1使用了一个唯一的全局解。更一般地，如果多个解在某个子句涉及的变量上都一致，那么该子句可以一次排除许多解。相关解是删除原子句后剩余的解。我们从这一精确条件开始。

固定变量 $x_1, \dots, x_N$和带索引支撑集 $H = (J_1, \dots, J_M)$，每个支撑集均非空。每个子句恰好使用其支撑集中的每个变量一次，且可取任一符号。带索引公式 $F = (C_1, \dots, C_M)$可以包含重复子句。用 $\mathrm{Sol}(F)$ 表示其解集，$D_d = C_d$，并用 $Q_d = F - D_d$ 表示删除对应的带索引出现项。当满足以下条件时，该位置是冗余的：

$$\mathrm{Sol}(Q_d) = \mathrm{Sol}(F). \tag{9}$$

对于 $t \in \{0,1\}^J$，定义

$$B_{t,J} = \bigvee_{i \in J} \ell_i, \qquad \ell_i = \begin{cases} x_i, & t_i = 0, \\ \neg x_i, & t_i = 1. \end{cases} \tag{10}$$

对于完整赋值 s, $B_{s,J}$是 $B_{s|_J,J}$ 的缩写。其使假集合恰为 $\{a : a|_J = t\}$。等价地，对于布尔整数变量，

$$B_{t,J}(x) = 1 \iff \sum_{i \in J} ((1 - t_i)x_i + t_i(1 - x_i)) \geq 1. \tag{11}$$

定理 4.1（投影判据）。设 F 可满足，且 $t \in \{0,1\}^{J_d}$。以下条件等价：

1. $Q_d \wedge B_{t,\,J_d}$ 不可满足；2. $\pi_{J_d}(\mathrm{Sol}(Q_d)) = \{t\}$；3. d 在 F 中是冗余的，且 $\pi_{J_d}(\mathrm{Sol}(F)) = \{t\}$。

特别地，当且仅当保留的解集在 J_d上的投影是单元素投影时，d 处的保持支撑的替换才会使 F 不可满足。

证明。替换后的解集为

$$\mathrm{Sol}(Q_d \wedge B_{t,J_d}) = \{a \in \mathrm{Sol}(Q_d) : a|_{J_d} \neq t\}.$$

由于 F 可满足，Q_d 也可满足，从而证明前两个条件等价。第三个条件由冗余性推出第二个条件。反过来，在第二个条件成立时，取 F的一个解。它在 t上的限制为 t，因此满足原子句 C_d。Q_d的每个解都有该限制，因而满足 C_d。因此 d 是冗余的，解集及其投影均相同。J_d 上的每个子句都是 $B_{t,\,J_d}$，且对应唯一一个 t，由此证明最后一个断言。

□

将阻断位置的集合定义为

$$\mathcal{B}(F) = \{d : |\pi_{J_d}(\mathrm{Sol}(F - D_d))| = 1\}. \tag{12}$$

对于每个 $d \in \mathcal{B}(F)$，令 t_d为唯一投影，并令 $F^{(d)}$用 B_{t_d,J_d}替换 C_d。该替换保持带索引的无向关联图不变，包括每个子句宽度。它与 F 的差异恰好在于 d，因为 F 是可满足的，而 $F^{(d)}$ 不是。

推论 4.2 （唯一解阻塞）。 如果 $\mathrm{Sol}(F) = \{s\}$，那么 $B(F)$ 恰好是 $D(F)$ 中的冗余位置集合。对于每个这样的冗余位置，

$$F^{(d)} = (F - D_d) \wedge B_{s,J_d}$$

与 F 具有相同带索引支撑的公式是不可满足的。

证明。E 的投影$\{s\}$ 是单元素集合，因此定理 4.1 将该判据归约为冗余性 y. □

这恢复了 [FLL+26] 中的确定性阻塞步骤，并证明了其在单位置、保持支撑的替换类中的逆命题。全局唯一性是充分条件，但并非必要条件：对于变量 x、y、z，公式 $(x) \wedge (y) \wedge (x \vee y)$ 有两个解，它们在 z上存在差异；将其第三个子句替换为 $\neg x \vee \neg y$ 可排除这两个解。

4.2 编辑限制下的对角线逆转

令 $\mathcal{X}_H$ 表示具有带索引支撑 H 的所有公式，并设 $P(G) = \mathbf{1}[\mathrm{Sol}(G) \neq \varnothing]$。令 $V_S(G)$ 记录位置 S 处的子句以及公共支撑系统。将 $\mathcal{N}_S(F)$ 记作 $\mathcal{X}_H$ 中可由 F 得到的公式，即至多更改一个位于 S之外的位置（包括 F 本身）所得的公式。其所有成员都具有观测 $V_S(F)$。

定理 4.3 (受限对角线实现)。 对于可满足的 F，令 H 为所有函数h： $V_S(X_H) \to \{0,1\}$ 的集合。以下命题等价：

1. 存在映射 C： $H \to N_S(F)$，使得对每个 h，

$$P(C(h)) = 1 - h(V_S(C(h))). \tag{13}$$

2. $N_S(F)$ 包含一个不可满足公式。

3. $B(F) \not\subseteq S$.

当这些条件成立时，一个固定的 $d \in B(F) \setminus S$ 给出了如下构造

$$C(h) = \begin{cases} F^{(d)}, & h(V_S(F)) = 1, \\ F, & h(V_S(F)) = 0. \end{cases} \tag{14}$$

证明。将第一个条件应用于常值函数 $h \equiv 1$，即可得到第二个条件。$\mathrm{N}_S(F)$ 中的一个不可满足成员与可满足公式 F 恰好在一个位置 $d \notin S$ 上不同。由定理 4.1 可得 $d \in \mathrm{B}(F)$，从而证明第三个条件。在第三个条件下，$P(F) = 1$，$P(F^{(d)}) = 0$，且 (14) 的两个分支都保持 $V_S(F)$。这些恒等式证明了第一个条件。 □

定理中的函数 h 是任意的二元预测，并不假定它们正确。对于可靠的部分判定器 D，将其接受指示函数 $h = a_D$ 代入（14）。由此得到的对象满足（3）；引理 2.6 给出 D（V_S（F））$= \perp$。一旦确定了不可判定性，选择器就会返回 F。因此，结论针对的是原始固定的母公式，尽管证明中使用了条件选择。这一构造只断言其存在性，并未对找到 d、t_d或判定求值器的状态所需的代价作出限制。

推论 4.4。 每个满足 B（F）$\not\subseteq S$ 的观测 S，都存在一个不可满足补全与可满足的 F一致。如果 F 是唯一可满足的，这就涵盖了省略任意冗余位置的所有 S，因而也涵盖了所有 $|S| < |D$（F）$|$。

证明。 选择 $d \in \mathcal{B}$（F）$\setminus S$并使用 $F^{(d)}$。 □

定理 4.3 中的必要性涉及单位置替换。覆盖$\mathcal{B}(F)$ 的观察仍可能存在在多个位置上不同的不可满足补全。

4.3 局部观察下的相对不完备性

观测理论是上下文，相容完备式是其实现，而目标是完整公式的可满足性。需要区分两类布尔变量。变量 x_i 描述合取范式的一个候选解。下文引入的变量描述合取范式中出现了哪些文字。观察者知道其中一些输入符号，并询问完整输入是否存在一个解。将这个问题编码为一个命题断言，就能用通常的逻辑后承准确表达观测所确定的内容。

固定一个定义在 $[N]$ 上的带索引支撑系统 $H=(J_1,\ \ldots\ldots,\ J_M)$，其中每个 J_j 均非空。对于每个$i \in J_j$，引入一个命题变量 $z_{j,\ i}$：取值 0 表示文字 x_i，取值 1 表示$\neg x_i$。因此，赋值 z 表示一个支撑集为 H 的完整合取范式 F_z。

对于固定的赋值 $a \in \{0,\ 1\}^N$，令

$$L_{a_i}(z_{j,i}) = \begin{cases} z_{j,i}, & a_i = 0, \\ \neg z_{j,i}, & a_i = 1, \end{cases} \qquad E_{H,a}(z) = \bigwedge_{j=1}^{M} \bigvee_{i \in J_j} L_{a_i}(z_{j,i}).$$

Thu当且仅当 a 满足 F_z 时，s $E_{H,\ a}(z)$ 成立。定义有限命题断言

$$A_H(z) = \bigvee_{a \in \{0,1\}^N} E_{H,a}(z). \tag{15}$$

其真值就是 F_z的可满足性。该析取式给出一种语义编码；无需对其简洁性作任何限制。

用 z^F 表示 F 的符号编码。对于 $S \subseteq [M]$，令

$$\Gamma_S(F) = \{z_{j,i} : j \in S,\ z^F_{j,i} = 1\} \cup \{\neg z_{j,i} : j \in S,\ z^F_{j,i} = 0\}, \tag{16}$$

其中 i 遍历 J_j。其模型恰好是这样的公式编码：其支撑集 H在观测位置上与 F 一致。这些公理描述的是输入公式，而不是将其子句作为对赋值 x的约束。支撑系统是固定且公开的。观察者并未得到未知完整公式唯一可满足的承诺；这样的承诺本身就已经确定了 A_H。

定理 4.5 (局部观测下的相对不完备性**)。** 对于支撑集为H的可满足 F，每个满足 $B(F) \not\subseteq S$ 的 S 都满足

$$\Gamma_S(F) \nvdash A_H, \qquad \Gamma_S(F) \nvdash \neg A_H. \tag{17}$$

将观测事实的全部逻辑后承加入后，这些陈述仍然成立。完整描述 $\Gamma_{[M]}(F)$ 证明了 A_H。对于唯一可满足的 F，该假设恰好是 $D(F) \not\subseteq S$。

证明。对于任意可实现的观测 y，令 Γ_y 为其单位符号事实，并令 $D(y)$表示它们在 A_H上的演绎状态。命题演绎的可靠性使得 D 对于可满足性是可靠的。定理 4.3 后的选择器和引理 2.6 强制 $D(V_S(F)) = \bot$，从而得到（17）。明确地，对于 $d \in \mathrm{B}(F) \setminus S$，$z^F$和 $z^{F^{(d)}}$是 $\Gamma_S(F)$ 的相反模型。命题 2.3 保证它们在逻辑后承下保持不变。完整描述只有赋值 z^F，它满足 A_H，因此命题完备性证明了完整信息结论。推论 4.2 给出了唯一解的特例。

□

4.4 全相反补全与不可或缺的位置

第 4.3 节基于单个未观测编辑，给出了相对不完备性的一个充分条件。为了准确刻画局部观测何时能够决定可满足性，我们考虑所有相容的不可满足补全。对于可满足的 $F \in \mathcal{X}_H$，定义

$$\Delta(F,G) = \{j : C_j(F) \neq C_j(G)\}, \qquad \mathcal{E}_F = \{\Delta(F,G) : G \in \mathcal{X}_H,\ \mathrm{Sol}(G) = \varnothing\}.$$

这些差异集非空。该族本身可能为空；此时，所有支撑为 H 的公式都是可满足的。

定理 4.6 (充分观测)。 对于可满足的 F，观测 S 证明 A_H 当且仅当它与 E_F 的每个成员都有交集：

$$\Gamma_S(F) \vdash A_H \quad \Longleftrightarrow \quad \forall E \in \mathcal{E}_F,\ S \cap E \neq \varnothing. \tag{18}$$

否则 A_H 独立于 Γ_S（F）。当 E_F 为空时，右侧条件是空泛成立的 y.

证明。 一个不可满足的 G 恰好在 $S \cap \Delta(F, G) = \varnothing$ 时与观测一致。因此，右侧表示观测到的符号事实的每个模型都编码了一个可满足公式。命题逻辑的可靠性与完备性给出（18）。如果仍有一个不相交的差异集，其公式和 F 便提供相反模型。由于 F 总是满足该观测，$\Gamma_S(F)$ 绝不会证明 $\neg A_H$。 □

称 S 在 F 处充分，若其满足（18）；称 d 在 F 处不可或缺，若每个充分的 S都包含 d。

Cor推论 4.7 (不可或缺的位置)。 对于可满足的 F，以下条件等价 ：

$$d \in \mathcal{B}(F), \qquad \{d\} \in \mathcal{E}_F, \qquad d \text{ is indispensable at } F.$$

对于唯一可满足的 F，这些正是冗余位置。

证明。 第一个等价关系是定理 4.1。单元素差异集迫使每个充分观察都包含 d。反之，若 $\{d\} \notin \mathcal{E}_F$，则集合 $[M] \setminus \{d\}$ 与每个差异集相交，因为没有差异集为空或包含于$\{d\}$中。因此，在没有 d的情况下它仍然是充分的。应用推论 4.2即可得到唯一性。 □

冗余性涉及子句对解集的影响；不可或缺性涉及对输入符号的知晓。删除冗余子句不会丢失对解的任何约束，但其未观测位置可能包含一个排除所有这些解的子句。

例4.8。 设 $F = (x) \wedge (y) \wedge (x \vee y)$。它的唯一解是 (1，1)，唯一的冗余位置是3。仅观测位置3是不充分的：$(\neg x) \wedge (\neg y) \wedge (x \vee y)$ 是一个不可满足补全。包含极小差异集恰好是 {3} 和 {1，2}。事实上，两个单位子句确定了一个赋值，而二元子句当且仅当排除该赋值时会导致不可满足性。如果位置3发生变化，差异集包含{3}；如果位置3不变，则两个单位子句都必须变为负，得到 {1，2}。两个极小集都会出现。因此，S 当且仅当 $3 \in S$ 和 $S \cap \{1，2\} \neq \varnothing$ 时才是充分的。

该示例将不可或缺的信息与充分信息区分开来。它还说明了为什么单一位置判据无法刻画相对不完备性的所有来源。

5 对角实现与算术化

前述结果都涉及反接受关系，但实现方式各不相同。我们先明确它与集合论对角模式之间的联系。然后，我们描述算术中的阻塞，并将这一描述与导致独立性的构造区分开来。

5.1 求值映射的实现

基本的 Lawvere–Yanofsky 模式取 $e\colon A\times A\to\{0,1\}$，并构造 $g(a)=1-e(a,a)$。对于每个 b，$g(b)$ 与 $e(b,b)$ 的值都不同；因此，g 并非任何一列 $e(-,b)$ [L69, Y03]。

命题 5.1（反接受的求值映射）。设非空集合 A 指标接受谓词 $\mathrm{Acc}_a\colon X\to\{0,1\}$ 以及满足 x_a $P(x_a)=1-\mathrm{Acc}_a(x_a)$ 的对象。则

$$e(a,b)=\mathrm{Acc}_b(x_a),\qquad g(a)=P(x_a)$$

实例化二元对角模式，并且 g 不由 e的任何列表示。若这些谓词是可靠的部分求值器的接受指示函数，则每个对应的求值器都会将 x_a 留为未决定。

证明。由假设可得 $g(a)=1-e(a,a)$。在 b 处求值会排除列 b。最后一个断言由引理2.6得出。□

对于算术，指标化扩展 皮亚诺算术的可靠可计算枚举表示 T_a，取$x_a=G_{T_a}$，并令 $\mathrm{Acc}_b(A)=\mathbf{1}[T_b\vdash A]$。标准真值为 P。定理 3.3 通过对角代入验证所需关系；可靠表示的指标集不必是有效可枚举的。

对于定理 4.3 中固定的 F、S，将 V_S 的部分求值器族 D_a 编入指标，并令$x_a=C(a_{D_a})$。令 $\mathrm{Acc}_b(G)=a_{D_b}(V_S(G))$。该选择器实现相同的关系。所有 x_a 都具有观测 $y=V_S(F)$，因此在此实例中，$e(a,b)=a_{D_b}(y)$独立于 a。每一列都是常值的。这一精确嵌入捕捉了与同一观测相容的相反补全之间的共同对角关系。解阻断使该关系在子句编辑约束下获得了具体实现。

事实上，对于固定观测 y 且不存在编辑限制时，相容完备式能够取得两个目标值，当且仅当每个二元预测器 h 都能被某个相容完备式反驳 x_h，其中 $P(x_h)=1-h(y)$。一个方向选择相反的值；另一个方向则使用常值预测器 0 和 1。定理 4.3 补充了该等价关系所遗漏的结构性问题：何时可以通过在固定支撑集上替换一个未观察子句来实现这一选择？答案是投影判据，而推论 4.7 确定了其观测后果。其贡献在于对集合论模式进行受约束的有限实现，并给出了将允许的编辑与解结构联系起来的精确判据。

5.2 两项保证及其机制

定理 5.2 (相对不完备性的两种实现)。以下构造实现 (3)并确立 (2)中的保证。

1. 对于每个可靠的可计算枚举扩张 T，对角代入构造G_T，使得 $D_T(G_T)=\bot$。理论 T相对于 G_T是不完备的；后者的真值在 T的两个模型中取相反值。*2.* 固定一个可满足的 F 和一个满足 $B(F)\not\subseteq S$的观测 S。应用于演绎状态接受指示函数的选择器 *(14)* 表明，$\Gamma_S(F)$ 相对于 A_H是不完备的。对于唯一可满足的 F，这涵盖了每个省略冗余位置的 S。

两种分离在演绎闭包下均得以保持。在第二种情况下，当且仅当满足关于 $B(F)$ 的既定条件时，才存在至多使用一次未观测编辑的保持支持的实现。

证明。定理 3.3 给出了 T 为 G_T赋予相反的真值。定理 4.3和 4.5 给出了相反的补全，分别由一个可满足公式及其不可满足伴随公式表示。命题 2.5 指出，这些配对分别是在各自上下文中相对不完备性的见证。命题2.3给出了在推论下的保持性。□

在算术中，该语句指涉其自身的证明关系。在有限阻塞中，解集决定一个排除子句；另行选择源公式还是其伴随公式，会使求值器的接受与之相反。后一选择将依赖于解的修改与常见的对角论证联系起来；而编码的自指则由下文带标记的算术构造另行提供。表1记录了这一区别。

组件	算术实现	有限实现
内部依赖	语句自身的证明谓词	源的解投影
构造	对角代入	阻塞与条件选择
保留的背景	固定理论 T	固定观测 $V_S(F)$
目标	G_T，随 T 变化	固定的可满足性断言 A_H
实现条件	编码与可靠的有效理论	未见处的单元素投影
相反模型	$T + G_T$ 和 $T + \neg G_T$的模型	位置 母体和伴随对象的符号编码 伴随对象

表1：共同反接受论证的两种实现。

5.3 用于解阻断的算术谓词

阻塞断言表明，母公式有一个解，移除选定子句后仍保留其解，而替换项排除了该解。每个条件都可以通过有限过程进行检查。对该过程进行编码，就能将这一描述转化为算术陈述，同时将母体与伴随对象保持为相互区别的对象。

选择一种对带索引合取范式、支撑集和赋值进行原始递归编码的方法。声明的变量数 $n(f)$ 和子句数 $m(f)$ 都是有效公式编码 f 的组成部分。一个数 $a < 2^{n(f)}$ 通过其二进制位对赋值进行编码。令 $\mathsf{Sat}(f, a)$ 表示： f 是有效合取范式编码， a 属于其值域，并且编码赋值满足每个子句。验证、位提取和求值都是原始递归的。

我们在皮亚诺算术的保守定义扩张中使用这些函数的符号。利用它们的图公式可以将这些符号消去。特别地，幂运算和下文的有限序列运算具有可由皮亚诺算术证明的全定义性。这些谓词取自其直接的有限算法，因此，它们的基本正确性恒等式可以在皮亚诺算术中通过归纳法得到证明。

定义

$$\mathrm{Unique}(f, a) \;:\Longleftrightarrow\; \mathrm{Sat}(f, a) \wedge \forall b < 2^{n(f)} \big(\mathrm{Sat}(f, b) \to b = a\big), \tag{19}$$

$$\mathrm{Red}(f, d) \;:\Longleftrightarrow\; 1 \le d \le m(f) \wedge \forall b < 2^{n(f)} \big(\mathrm{Sat}(f, b) \leftrightarrow \mathrm{Sat}(\mathrm{Del}(f, d), b)\big). \tag{20}$$

删除保留所声明的变量全集。令 $\mathrm{Rep}(f、\ d、\ a、\ J)$ 表示通过将位置 d 替换为由 a 和 J指定的子句（10）而得到的编码。无效输入将获得固定的默认输出。记 $\mathrm{Adm}(f、\ d、\ J)$ 为以下原始递归条件： f是一个在互异变量上具有非空子句的带索引合取范式， d 是一个有效位置，并且 $J =\mathrm{supp}(D_d)$。如果要求精确宽度，还应加入相应检查。

定义 5.3。 算术阻塞关系为

$$\begin{aligned} \mathrm{Block}(f, g) \;:\Longleftrightarrow\; \exists a, d, J\, [\; & \mathrm{Adm}(f, d, J) \wedge \mathrm{Unique}(f, a) \\ & \wedge \mathrm{Red}(f, d) \wedge g = \mathrm{Rep}(f, d, a, J)\,]. \end{aligned} \tag{21}$$

定义不可满足谓词(g) := 有效合取范式谓词(g) $\wedge \forall b < 2^{n(g)} \neg\, \mathsf{Sat}(g,\ b)$。

命题 5.4（算术正确性）。 对于刚刚指定的编码，

$$\mathsf{PA} \vdash \forall f\, \forall g\, \big(\mathrm{Block}(f,g) \to \mathrm{Unsat}(g)\big).$$

皮亚诺算术还证明：对于每个阻断对，所声明的变量数与子句数分别相等，且带索引的无向关联图相等。

证明。在皮亚诺算术中进行推导，并取 a、d、J 作为 (21) 的见证。若 $\mathsf{Sat}(g,b)$ 成立，则替换正确性蕴含 $\mathsf{Sat}(\mathsf{Del}(f,d)\ b)$。冗余性给出 $\mathsf{Sat}(f,b)$，而唯一性给出 $b = a$。阻塞子句的正确性表明，该子句的任何满足赋值都至少在 J 中的一个位置上不同于 a。这与 $b = a$ 矛盾。 g 的有效性及其所声明参数的保持性由可容许性条件推出。图恒等式成立，是因为替换保留了位置和支撑。每一步要么是恒等式，要么是关于原始递归运算的有界全称蕴含；这些结论均由皮亚诺算术中相应的有限列表归纳加以验证。 □

因此，“唯一解被禁止”这一非正式断言具有精确的算术解释。它将母代码 f 与伴随代码 g 联系起来。对于 f，该赋值是唯一的；它不是 g 的满足赋值。现在，我们利用这些谓词构造一个算术句子，断言编码在该句子自身中的阻塞数据的有效性。

5.4 用于解阻断的自指句子

算术句子 G_p 的自然语言解读如下：

> 我自身编码中携带的数据指定了一个有效的唯一解阻塞构造。

句法标签存储这些数据，解析函数读取该标签。对角代入使该解析函数的参数成为完整句子本身的编码。所得不动点在编码层面表达了自指；其数学内容仍然是关于所存储数据的断言。

关系（21）本身可以置于一个对角句子中。令 p 编码一个三元组 $(f,\ d,\ J)$，并定义

$$R(p) := \mathrm{Triple}(p) \wedge \mathrm{Adm}(f,d,J) \wedge \mathrm{Red}(f,d) \wedge \exists a\ \mathrm{Unique}(f,a), \tag{22}$$

其中各分量通过固定投影函数获得。伴随代码于是为 $\mathrm{Rep}(f,\ d,\ a,\ J)$；唯一的 a 确定了它，而命题 5.4 证明了其不可满足性。因此， R 断言该阻塞构造的有效性：母体具有唯一解，而伴随对象将其排除。

选择一种句法约定，使公式可以以标签 $\bar{p} = \bar{p}$ 作为左合取项开头。令 $\mathsf{Payload}(z)$ 从此前导标签中提取数词，并对其他表达式使用默认值。这是一个原始递归解析函数。尽管该句法标签在逻辑上是冗余的，我们仍将其保留在书面公式中。

命题 5.5（带标签的对角嵌入）。 对于每个自然数 p，都可以有效地构造一个算术句子 G_p，使得

$$\mathsf{PA} \vdash G_p \leftrightarrow R\big(\mathrm{Payload}(\ulcorner G_p \urcorner)\big), \tag{23}$$
$$\mathsf{PA} \vdash \mathrm{Payload}(\ulcorner G_p \urcorner) = \bar{p}, \tag{24}$$
$$\mathsf{PA} \vdash G_p \leftrightarrow R(\bar{p}). \tag{25}$$

证明。 利用第3节的编码，构造

$$A_p(u) := (\bar{p} = \bar{p}) \wedge \exists v\big(\mathrm{Diag}(u,v) \wedge R(\mathrm{Payload}(v))\big).$$

令 $a_p = \ulcorner A_p(u) \urcorner$ 并令 $G_p = A_p(\bar{a}_p)$。对角函数将 a_p 映射为 $\ulcorner G_p \urcorner$。其具体计算和函数性，以及真标签方程，给出（23）。代换

不会改变前导标签，因此直接解析得到（24）。代换这个等式得到（25）。所有缩写都可以在编码前展开；按照这一约定，前导标签会被保留。

□

方程（23）–（25）具有直接的解读。第一条表明， G_p 断言了从其自身编码中读取的数据具有阻塞性质。第二条将这些数据确定为 p。因此，第三条将该断言归约为 $R(\bar{p})$。对于每个固定的有效阻塞数据，有限验证证明了 $R(\bar{p})$ 在皮亚诺算术中成立，因而也证明了 G_p。

独立性取决于对角构造中所使用的谓词和引用机制。在（7）中，应用于该语句编码的谓词否定了在 T中存在该语句自身证明的可能性。假设存在一个证明，那么它恰好提供了该语句所否定其存在的对象。在（23）中，谓词则读取一个固定数据，并检查其阻塞性质。对于有效数据， G_p 的证明恰好验证了它所断言的阻塞性质。同样的带标签构造适用于以任意算术谓词替代 R的情形；其内容取决于该谓词。这就是为什么除了不动点等价 [HV14、GHY23]之外，还必须明确指定引用机制。

定理 4.5 中证明的有限独立性属于另一对对象：观测理论 $\Gamma_S(F)$ 和目标 A_H。其产生的原因在于，观测理论既有一个可满足补全，也有一个不可满足补全。通过完整的阻塞数据进行验证，便确立了这一构造；而观察者的背景 $\Gamma_S(F)$ 仍然同时允许这两种补全。证明指定的信息无法确定答案，与这种无法确定性是相容的。

因此，所展示的两个自指关系具有不同的作用。方程（3）涉及接受，并推动共同的非决策论证。方程（23）涉及从语句的编码中提取数据，并记录有限构造。观察定理将 A_H 用作其目标，而带标签的语句则在编码层面描述了该构造。

6 困难性假设的逻辑基础

前述结果表明，理论或局部观察可能使某个目标处于未决状态。现在我们要探讨，这类信息约束如何指导计算困难性假设的表述。

6.1 元语言环境与自我识别障碍

在「盲人摸象」这个成语背后的寓言中，每个观察者都只能触摸到大象的不同部位，并试图据此推断大象的整体形状。假设观察者只能触摸某个指定部位。两个物体在该部位摸起来可能完全相同，但整体形状却不同。仅凭获准触摸的部位做出的任何判断，对这两个物体都必须相同。为了确立这一限制，我们将观测方法本身作为研究对象，并比较那些能够产生相同观测结果的物体。

李的环境原则为区分这两种推理提供了一个框架 Li's principle of environments providesa framework for distinguishing these two kinds ofreasoning [L14, 第 10.1–10.2 节，原则 10。1]。对象语言提供句法，模型对其作出解释，而元语言环境则是我们定义语言及其模型并证明相关结果的地方。算术推导依据给定理论的规则进行。证明某个理论无法判定一个语句，则需要考察该理论的公理和可证性关系。同样，SAT 判定规则根据已观察到的子句进行推理，而关于其局限性的证明则考察与这些子句相容的完整公式。

该自我识别障碍由 Fang 等人讨论 [FLL+26，第 6 节。1] 当系统使用自身的方法来探究这些方法本身的局限时，就会出现这种情况。只要探究局限于底层背景，它就会继承该背景的局限。寓言中的观察者或许能够认识到，允许进行的触摸使整体形状

保持开放。如果观察者使用同样的触摸来检验某个特定猜测是否正确，这一检验仍然无法区分这两种形状。元语言分析通过比较两种可能性，并展示观测留下了哪些未解决之处，揭示了这一盲点。

命题 2.2 精确刻画了根本的信息极限。可容许背景无法唯一确定二元目标，当且仅当它允许两个相容实现，而二者的目标值相反。只要分析、代数和几何方法的判断都仅依赖于同一背景，它们便都面临这一极限。假设某条仅限于该背景的规则在两个实现上都正确，那么面对相同的信息，它就必须给出不同的答案。这一矛盾表明，可靠的部分规则必须在该背景下对目标保持未决。

在此处研究的两种情形中，自指为这一极限提供了具体见证。哥德尔构造利用理论自身的可证性谓词，得到一个独立句子（定理 3.3）。随后，该理论的两个模型对这个语句的判断相左。在有限构造中，母公式的唯一解确定了阻塞子句。对于每个未能观测到冗余子句位置的观测，在该位置进行替换，都会保留观测结果，同时将一个可满足公式变为不可满足公式（定理 4.5）。这一对实现揭示了所有依据该观测进行判断的规则共同面临的局限。关于一整类方法之局限的论述，也正是复杂性问题的起点。

6.2 两个层面的算法与复杂性

算法是一种数学方法，其定义方式使其能够机械地执行。要建立无条件上界，只需构造一个算法，并证明它能在规定的资源界内正确地解决问题。一般下界涉及指定计算模型所允许的每一个正确算法。它属于元数学的范畴，因为它以整个算法类为研究对象。由于算法是机械化的数学方法，研究算法的局限性，也就是研究数学以机械化形式呈现时的局限性。

完全通过所研究方法开展的探究，会继承这些方法自身的盲点。如果

	算术理论	**有限 SAT** **观察**	**一般算法** **复杂性**
内部方法	使用……推导语句 该理论的公理和 规则。	根据……判断可满足性 规定的观测。	运行算法于♀♀♀ 指定的计算 模型中运行算法。
元层次 问题	哪些句子处于 理论能力所不及的范围 无法判定?	哪些观察使 可满足性仍无法确定, 对于每一条局限于该 观测信息?	有哪些资源限制 适用于整个 模型允许的算法类 吗?
……的基础 分析	两个理论模型 对目标 语句。	两个公式在 观测上相同，但意见不一 关于可满足性。	一个已识别的结构性 障碍促使我们提出一项 关于某一指定算法类的资源声明 算法类 算法。
此处状态	哥德尔构造 给出一个独立的 语句，在……的 假设下 定理 3.3。	解阻断给出 相反补全 在以下假设下 定理 4.5。	SRK-ESH 表述了一个 最坏情况时间假设 针对指定的承诺 族。

表 2：内部方法与元数学问题。

证明某种不可能性需要作出这些方法无法作出的区分，那么数学若在这些方法内部运作，就无法无条件地确立这种不可能性。这就是第6.1节所讨论的自我识别障碍。

有限 SAT 场景提供了一个具体例子。假设某个程序只能接收指定的子句观测，却被要求核验实际输入的答案“可满足”。如果一个可满足公式和一个不可满足公式给出相同的观测，那么该程序在两种情况下掌握的信息相同，尽管所提出的答案只对其中一种情况正确。在元数学层面，对这两个相容公式进行比较，就能确立一项所有受限于该观测的程序都共有的局限。

表 2比较了各框架内方法的作用，以及我们可以就框架本身提出的问题。

前两个场景展示了元数学结果如何揭示一项局限，而所有受某项明确限制的方法都共同受制于这一局限。第三个场景引出了第 6.3 节探讨的问题。这类结果如何帮助我们提出关于整个算法类的资源限制假设?

6.3 从相对不完备性到困难性假设

证明 P $\neq$ NP 长期以来一直是计算复杂性领域的一项核心挑战。首先，如上所述，算法设计是在数学层面进行的，而对算法局限的研究则是在元数学层面进行的。其次，尽管数十年来人们不断努力，在不对空间另作限制的通用确定性模型中，SAT 和许多其他典型 NP 完全问题的无条件超线性最坏情况时间下界仍难以确立。这两点使我们认为，P $\neq$ NP 很可能独立于 ZFC 这样的标准数学理论。因此，传统的探究顺序或许已经颠倒。一个长期以来以证明为目标的猜想，本身可以作为工作假设，用以发展计算困难性理论。这一转变使得如何选择工作假设本身也成为一个方法论问题。[1] P $\neq$ NP 作为出发点有多合适？更一般地说，我们应依据什么来选择困难性假设?

支持 P $\neq$ NP 的两个常见依据是多项式时间归约和算法设计经验。归约表明，对于任何一个 NP 完全问题，若存在多项式时间算法，就能为 NP 中的所有问题构造多项式时间算法。持续未能找到这样的算法，为这一猜想提供了经验支持。这里所寻求的逻辑基础，应当能够确定某个 NP 完全问题的一个结构特征，而这一特征会对每一种候选多项式时间算法构成共同障碍。归约和算法设计方面的论证并未指出这一障碍究竟是什么。

指数时间假设（ETH）和 3SUM 假设的发展史说明了这些证据的局限性 [IP01、VW18]。ETH 认为，判定 3-SAT（涉及 n 个变量）在最坏情况下需要指数时间。OpenAI 通过给出一种确定性算法，计算固定宽度 k-合取范式公式的满足赋值数量，并在允许多项式因子的情况下达到 $2^{o(n)}$ 时间，从而反驳了这一假设。检验该数量是否为正，即可在相同的时间界内判定3-SAT [OA26，定理 1。1]。3SUM 假设排除了对任意固定 $O(n^{2-\varepsilon})$都适用的 $\varepsilon > 0$-时间算法。Alman 和 Vassilevska Williams 针对整数有多项式界的情形，给出了一个确定性的 $O(n^{1.9992})$-时间算法，从而反驳了该假设 [AVW26]。在这两种情况下，新算法都跨越了此前的研究使研究者有理由猜想的时间阈值。

困难性假设的元数学基础，始于所有受明确限制的方法共同面临的结构性障碍。相对不完备性提供了这样

[1]人工智能的快速发展可能正在开启工业规模的数学证明生成时代。随着越来越多的证明发现与验证工作能够实现自动化，提出富有成果的猜想并确定哪些假设有充分依据，可能会日益成为数学和理论计算机科学中的核心任务。

一种障碍。在算术中，同一理论的两个模型可能对某个语句意见相左。在有限情形下，两个完整公式可以在所有已观察到的子句上相同，却具有不同的可满足性。任何局限于规定信息的方法，在这两种情况下都只能基于相同依据作出判断。既然正确答案不同，这些信息就无法决定目标。

Fang 等人提出了一项用于困难性假设的双重证据标准 [FLL+26，第 7 节。2]按本文采用的术语，一个有充分依据的困难性假设应同时依托数学证据和元数学证据。在数学层面，对已知算法的分析会展示它们实现了哪些加速，以及其技术在哪些方面有所不足。在元数学层面，逻辑论证会指出所有受明确限制的方法共同面临的结构性障碍，并解释该障碍与所猜想的资源界有何关联。两类证据都应针对同一个实例族、计算模型和资源规模。

自指 K-SAT 穷举搜索假设（SRK-ESH）展示了这两类证据如何为一项具体的资源声明提供依据。如果对唯一可满足母公式的观测遗漏了一个冗余子句位置，那么在该位置阻塞唯一解，就会得到一个与该观测一致的不可满足伴随公式。因此，相对不完备性揭示了一个区别：承诺族上的每个正确求解器都必须作出判断，而规定的局部信息却无法消解这一点。在算法方面，Fang 等人考察的策略表明，对于同一承诺族，即使给出完整公式，也尚未证实指数中存在固定比例节省。以 2^N 赋值测试的穷举搜索基准$0 < \delta < 1$为参照，SRK-ESH 断言，对于每个固定的$O^*(2^{(1-\delta)N})$，不存在一致确定性算法能在最坏情况时间 N 内判定该承诺族，其中 O^* 表示变量数， [FLL+26，第 7 节。2]。

7 结论

相对不完备性为哥德尔不完备性与有限 SAT 中局部观察的局限性提供了一个共同刻画。在算术中，对角化产生了一个语句，使同一理论的两个模型对此意见不一致。在有限 SAT 中，解阻断产生了两个在已观察子句上相同、但在可满足性上不同的公式。真与接受之间的共同关系解释了为什么可靠的部分判定器必须对相应问题保持未决。对于一个可满足公式，当删除后剩余的解在被替换子句的支撑集上全部一致时，一次保持支撑的替换恰好可以使其变得不可满足。对于唯一可满足母公式，符合条件的位置恰好是其冗余位置，而这些位置正是每次充分的子句观测都必须包含的位置。求值映射将这两种构造都与既有的集合论对角模式联系起来。算术阻塞关系和带标签不动点在编码层面描述了这一有限构造。

背景及其演绎闭包拥有相同的模型，因此，可靠推理会保留与其相容的任何相反实现。这一信息视角与 Chaitin 对形式系统的分析有关 [C74]。可用信息所支撑的判断本身就是一个可以在元语言层面证明的主张。在有限情形中，相对不完备性揭示了哪些预先规定的观察会使可满足性保持未决。这一阻碍促成了 SRK-ESH，其定量主张涉及判定指定承诺族所需的最坏情况时间。

参考文献

[AVW26] Josh Alman 和 Virginia Vassilevska Williams。通过稀疏非均衡图中的三角形，实现真正次二次时间的 3SUM 和真正次三次时间的 APSP。arXiv:2610.06783，2026 年。

[BBJ07] George S. Boolos、John P. Burgess 和 Richard C. Jeffrey。《可计算性与逻辑》。第 5 版。剑桥大学出版社，2007 年。[C74] Gregory J. Chaitin。形式系统的信息论限制。《ACM会刊》，21(3):403–424，1974。doi:10.1145/321832.321839。[FLL+26] Wen Fang、Xianxian Li、Jun Liu、Jie Luo、Yongxin Tong 和 Ke Xu。自指 K- 可满足性问题与哥德尔不完备定理的有限类比。arXiv:2607.01671v3，2026 年。https://arxiv.org/abs/2607.01671v3.[G31] 库尔特·哥德尔。关于《数学原理》及相关系统中形式上不可判定的命题 I。《数学与物理月刊》，38:173–198，1931。doi:10.1007/BF01700692。[GHY23] Balthasar Grabmayr、Volker Halbach 和 Lingyuan Ye。元数学中的自指类型。《哲学逻辑杂志》，52:1005–1052，2023。doi:10.1007/s10992-022-09696-y。[HV14] Volker Halbach 和 Albert Visser。算术中的自指（一）。《符号逻辑评论》，7(4):671–691，2014。doi:10.1017/S1755020314000288。[IP01] Russell Impagliazzo 和 Ramamohan Paturi。论 k- 可满足性问题的复杂性。《计算机与系统科学杂志》，62(2):367–375，2001。doi:10.1006/jcss.2000.1727。[L69] F. William Lawvere。对角论证与笛卡尔闭范畴。载于《范畴论、同调论及其应用 II》，《数学讲义》92，第 134–145 页。施普林格，1969 年。附作者评注重印于《范畴论及其应用重印文集》，15:1–13，2006 年。https://www.tac.mta.ca/tac/reprints/articles/15/tr15abs.html.[L14] Wei Li。《数理逻辑：信息科学基础》。第二次修订版。《计算机科学与应用逻辑进展》25。施普林格巴塞尔，2014 年。doi:10.1007/978-3- 0348-0862-0。[OA26] OpenAI。《指数时间假设是错误的》。预印本，2026 年 10 月 17 日。修订于 2026 年 10 月 25 日。修订版预印本。[PT19] Andrei Popescu 和 Dmitriy Traytel。哥德尔不完备定理的形式化验证抽象论述。载于 Pascal Fontaine 编，《自动演绎—CADE 27》，《数学讲义》11716，第 442–461 页。施普林格，2019 年。doi:10.1007/978-3-030- 29436-6 26。_[R36] Barkley Rosser。哥德尔和丘奇定理的若干推广。《符号逻辑杂志》，1(3):87–91，1936。doi:10.2307/2269028。[VW18] Virginia Vassilevska Williams。算法与复杂性中的若干细粒度问题。载于《国际数学家大会论文集》（ICM 2018），第 3431–3472 页，2018 年。doi:10.1142/9789813272880_0188.[Y03] Noson S. Yanofsky。处理自指悖论、不完备性与不动点的统一方法。《符号逻辑通报》，9(3):362–386，2003 年。 https://arxiv.org/abs/math/0305282。