arXiv:2304.12109v1 [cs.LO] 24 Apr 2023

伪随机有限模型

Jan Dreier[1] 和 Jamie Tucker-Foltz[2]

[1]维也纳技术大学。 dreier@ac.tuwien.ac.at
[2]哈佛大学。 jtuckerfoltz@gmail.com

2023年4月25日

摘要

我们从逻辑可定义性的角度研究伪随机性和伪随机生成器。在常规去随机化和有限模型论相关结果的基础上，我们证明，可以在多项式时间内确定性地构造图和关系结构，使其对于一阶逻辑或最小不动点逻辑的任意句子而言，在统计上都无法与随机结构区分。这引出了这样一个问题：能否通过从熵更低的简单结构出发的逻辑转导来实现这类构造？换言之，逻辑公式能否成为伪随机生成器？我们完整刻画了对于一阶逻辑、不动点逻辑以及带奇偶性的不动点逻辑，何时可以实现这一点；对于带奇偶性的一阶逻辑，我们则给出了部分结果和猜想。

1 引言

决策问题的最坏情形可处理性，在有限模型论领域有一个优雅的对应概念：问题的难度不再依据解决问题所需的计算资源来衡量，而是依据定义问题所需的逻辑构造来衡量。从 Fagin 的开创性工作 [13] 到近期“刻画多项式时间的探索” [16], ，人们已经用逻辑描述了决策问题的各类复杂性类 [20]。近期的研究也开始将描述复杂性拓展到近似领域 [3, 28]。然而，从逻辑视角探讨平均情形计算问题的研究仍然寥寥无几。

密码学的一个基础猜想是，密码学安全的伪随机生成器（PRG）存在：这类确定性多项式时间算法将随机输入映射为更长的伪随机输出，使得任何多项式时间算法都无法区分其输出与真正的随机性。这可以看作生成器与对手之间的一场博弈：生成器构造伪随机输出，而对手以概率1 2 看到的要么是生成器的输出，要么是真正的随机字符串，并且必须以显著高于随机猜测的概率区分二者。伪随机生成器的存在等价于单向函数的存在；后者是密码学领域一个标准且广受认可的猜想（见第2.4节）。本文研究逻辑公式充当生成器或对手的能力。我们考察的逻辑包括一阶逻辑（FO），以及它的两个研究充分的扩展：加入计算最小不动点（LFP）和奇偶性运算符的逻辑（FO[⊕]和 LFP[⊕]）。

我们不再考察随机字符串，而是通过随机图和关系结构来衡量随机性。这使我们能够借鉴关于这些逻辑的现有结果，并研究一些引人入胜的

无序结构特有的问题，也就是说，这类结构没有一个总是被解释为全序的二元关系符号。对于不同的关系签名，我们考察每个元组以概率 1 2 出现在某个关系中的结构。对于单个对称二元关系，这便得到例如经典的 Erd¨os-Renyi 随机图 $G(n, 1/2)$。逻辑在此类结构上的表达能力已有充分研究：一阶逻辑 [14]和最小不动点逻辑 [5] 的零一律 φ指出，对于任意句子，随着论域 n 规模增大，满足该句子的概率会以指数速度收敛到零或一，具体取决于 φ。一阶逻辑[$\oplus$]也存在这一现象的一个变体（由 [22], 在第5节中详细回顾）：对于随机图，满足该句子的概率会收敛到两个极限（分别对应偶数个 n 和奇数个 n）。另一方面，最小不动点逻辑[$\oplus$]能够以高概率 [17],规范化随机结构，这意味着（通过 Immerman-Vardi 定理 [19,29]），就随机输入而言，最小不动点逻辑[$\oplus$] 的能力与多项式时间计算相当，因此显然不可能存在零一律。

本文研究的结构并非随机，而是伪随机的。我们在第 2 节提出了这一概念的形式化定义，其与私钥密码学中的标准定义十分相似。[1] 零一律表明，一阶逻辑和最小不动点逻辑应当存在多项式时间伪随机生成器：要生成能够欺骗基于某个公式的对手的伪随机性 φ，只需构造一个结构，使其接受 φ 当且仅当极限概率为一而非零——而几乎所有结构都满足这一性质。然而，复杂性理论中充斥着许多类似任务（称为“大海捞针” [2, 第 21]章），它们最终都被证明相当困难。

1.1 存在闭图的确定性构造

我们在第 3 节提出的首个主要结果，是一种确定性的多项式时间构造方法，可以构造出一阶逻辑和最小不动点逻辑无法与真正随机的结构区分的图及其他结构（定理 12）。为此，我们构造由所谓的扩展公理（定义见第 3 节）描述的模式。这些模式在随机结构中以很高的概率出现，并使一阶逻辑和最小不动点逻辑句子满足零一收敛。

对于图而言，这可以理解为一种无限构造的有限类比：拉多图 是满足所有扩张公理的唯一无限图（在同构意义下）。从无限 Erd¨os-Renyi 图 $G(\infty, 1/2)$ 中抽样，以概率 1 得到拉多图。现已有若干显式构造[2] ，并且可以在多项式时间内计算它们的有限前缀。然而，这些前缀不 满足扩张公理，因此即使面对一阶逻辑对手，也无法提供伪随机结构。

事实证明，以确定性方式构造拉多图的有限类比物是一项更具挑战性的任务，这些类比物通常也称为存在闭图。满足扩展公理的显式图已经被发现，但它们是基于数学结构特殊性质的精巧构造，例如佩利图 [4], 二进制矩阵 [6],和仿射设计 [9]；有关这些构造的概述，请参见 Bonato [7] 。这些构造都无法明显推广到具有多个任意元数的非对称关系的更复杂有限模型。[3] 为此，我们利用去随机化领域的工具，尤其是

[1]据我们所知，本文首次研究逻辑可定义性背景下的密码学概念。已有关于“伪随机图”的文献，但该术语通常指满足图的某些有用统计性质的图，例如子图密度和谱隙。这些图在我们的语境下通常并不具有伪随机性。关于综述，参见 Krivelevich 和 Sudakov [23]。[2]例如，使用顶点集 $\mathbb{Z}_{\geq 1}$，当 $x < y$ 且y 的二进制表示的 x^{th} 位为 1 时，连接 x 和 y [1,27]。据我们所知，存在一个例外：Even-Zohar、Farber 和 Mead [12] 通过巧妙应用范德蒙德行列式，将 Paley 图构造扩展到了超图。然而，如何进一步扩展这一思想，使其适用于任意关系签名，仍不清楚。

Θ \ φ	一阶逻辑	最小不动点逻辑	最小不动点逻辑[⊕]
FO	✗	✗	✗
LFP	✗	✗	✗
最小不动点逻辑[⊕]	✓	✓	⟺单向函数存在

表 1：从具有一个二元关系的关系签名到具有一个三元关系的关系签名的伪随机生成器是否存在，取决于转导 Θ和对手 φ 所允许采用的逻辑。对于右下角的条目，伪随机生成器存在等价于私钥密码学中的标准猜想：单向函数存在（也等价于猜想 3）。所有可能的签名对的一般分类见表 2。

完美哈希函数和通用集，构建了一个更稳健且灵活的框架。我们的方法与先前的方法大不相同，并且能够轻松处理图以及任意的关系结构。

1.2 逻辑与逻辑

正如我们将对手限制为在给定逻辑中可定义一样，我们也可以通过将生成器限制在某种逻辑中来“拉平竞争环境”。为了定义通过逻辑从输入生成输出的含义，我们考察转导的概念 [10] （在第2.2节中有形式化定义）。在本文中，转导是由逻辑公式组成的元组，它将具有某个签名 σ 的所有结构映射为具有相同论域但不同签名 τ的结构。例如，考虑一个输入签名 σ，其中包含一个二元关系，以及一个输出签名 τ，其中包含一个三元关系 T。那么，从 σ到 τ的转导 Θ由单个 σ-公式 $\theta(x,y,z)$组成；该公式将每个输入 σ-结构 $\mathbb{A}$关联到 τ-结构 $\Theta(\mathbb{A})$，其中当且仅当在 $\mathbb{A}$中 $\theta(x,y,z)$成立时， $T(x,y,z)$成立。如果将 Θ应用于一个随机的 σ-结构会产生一个伪随机的 τ-结构，我们就称 Θ是从 σ到 τ的伪随机生成器。因此，在我们的例子中， Θ 接收一个具有 n^2 比特熵的二元σ-结构，并且必须生成一个具有 n^3 比特熵的伪随机输出（使对手无法将其与三元 τ-结构区分开来）。我们认为这是PRG的一个恰当的模型论类比。

我们在第4节证明了第二个主要结果：根据我们要求生成器和对手所属的逻辑，考察 {一阶逻辑、最小不动点逻辑、最小不动点逻辑[⊕]}中所有可能的逻辑组合，以及输入和输出签名，可以完整刻画这类伪随机生成器何时存在。表1 列出了特殊情形下的结果，其中 σ 仅包含一个二元关系，而 τ 仅包含一个三元关系。

请注意，在这种模型论语境下，熵具有不同的“形状”。考虑从具有单个三元关系的结构到具有三个二元关系的结构的伪随机生成器。通常，将包含 n^3 个独立随机比特的输入重塑并截断为包含 $3n^2$ 个比特的输出是微不足道的，但逻辑是否也能完成这项任务并不明确。我们提出了一种“熵”的度量，它完全决定了哪些签名之间存在伪随机生成器。出人意料的是，这一概念取决于所考察的逻辑组合，并且并不总能在签名上定义一个全序。

1.3 带奇偶性的一级逻辑

值得注意的是，我们迄今考察的逻辑中，没有一种能提供一个对来自同一逻辑的对手无条件安全的伪随机生成器：一阶逻辑和最小不动点逻辑过于弱小，而 LFP[⊕] 又过于

强。为了寻找一个有趣的折中方案，我们转而考虑一阶逻辑[⊕]。它远不足以表达多项式时间，却仍能表达有趣的性质，例如“存在奇数个三角形”，而该性质在随机图中以极限概率 1 2成立。我们考虑一阶逻辑[⊕]-转导，它通过添加每条超边，将一个普通图转化为一个 t-超图，当且仅当存在奇数个顶点与以下每个顶点都相邻时，才添加该超边： $\{v_1, \ldots, v_t\}$ v_1、……、 v_t。我们在第 5 节证明，这种转导是一个对一阶逻辑和最小不动点逻辑安全的伪随机生成器。具体而言，我们证明，经过转导得到的超图以高概率满足相应的超图扩张公理。因此，我们的转导将一个具有$\binom{n}{2}$ 比特熵的随机图转化为一个超图，而一阶逻辑或最小不动点逻辑的对手无法将其与具有 t比特熵的真实 $\binom{n}{t}$ -超图区分开来。我们猜想，这种特定的转导对一阶逻辑[⊕]也安全，因此一阶逻辑[⊕]存在一个对自身安全的伪随机生成器。

2 预备知识

对于任意非负整数 n，令 $[n] := \{1, 2, 3, \ldots, n\}$。$\log(n)$ 表示以2为底的对数。我们称函数 $f(n)$ 是 可忽略的，如果对于任意多项式函数 p，当 n 足够大时，对所有 $|f(n)| \leq 1/p(n)$ 均成立。我们用 $f(n) = \mathrm{negl}(n)$ 表示这一点。对于任意有限集 S，我们用 $s \sim S$ 表示从 S 中均匀抽样元素 s；用 $s \sim \mathcal{D}$ 表示按照在 S 上的分布 $\mathcal{D}$ 进行抽样。对于任意非负整数 k， $\binom{S}{k}$ 是所有 S 的 k 元子集构成的集合。

2.1 模型与逻辑

在本文中，我们研究两类随机结构。第一类是 关系结构，它们是相对于 关系签名 $\sigma = \langle R_1, R_2, \ldots, R_t \rangle$ 定义的，其中每个关系符号 R_i 都具有相应的 元数 a_i。一个 σ- 结构 $\mathbb{A}$ 是一个有限论域 A，以及一列关系 $\langle R_1^{\mathbb{A}}, R_2^{\mathbb{A}}, \ldots, R_t^{\mathbb{A}} \rangle$；其中每个 $R_i^{\mathbb{A}}$ 都是论域 A 上的 a_i 元关系，即 A^{a_i} 的子集。我们将 $\mathbb{A}$ 与其论域 A 等同，因此例如写作 $a \in \mathbb{A}$ 而不是 $a \in A$。我们用 $\mathrm{Str}[\sigma]$ 表示所有 σ-结构构成的集合，用 $\mathrm{Str}[\sigma, n]$ 表示所有具有 n 个元素的 σ- 结构构成的集合。从 $\mathrm{Str}[\sigma, n]$ 中进行均匀抽样会得到一个 随机 σ-结构，其中论域中的每个有序 a_i 元组（元素可以重复）独立地以 1 2 的概率属于关系 R_i。

其次， t-超图是关系结构，由一个对称的 t 元超边关系构成，其中每个 t 元组都由 t 个互异元素组成。2-超图就是一类图。Erd¨os-Renyi 随机图 $G(n, 1/2)$ 定义为所有 n 顶点图上的均匀分布，也就是说，每条边都以 $1/2$的概率独立出现。类似地，$G_t(n, 1/2)$表示所有 n 元 t超图上的均匀分布，其中每组由 t 个互异顶点构成的集合都以 $1/2$ 的概率独立形成一条超边。

对于任意图类或关系结构类 $\mathcal{P}$ 以及逻辑 L，如果存在一条 L 中的句子 φ，使得 $\mathcal{P}$ 恰好由所有满足 φ 的图/结构组成，我们就称 $\mathcal{P}$ 在 L 中是可定义的。我们假设读者熟悉一阶逻辑（一阶逻辑）。我们将带有最小不动点算子的一阶逻辑扩展记作最小不动点逻辑。我们不必详细介绍最小不动点逻辑，只需指出它具有以下性质。*Immerman-Vardi* 定理指出，如果存在一个关系符号，其解释是论域上的全序，那么最小不动点逻辑的表达能力等价于多项式时间计算：

定理 1 (Immerman [19], Vardi [29])。 设 σ 是一个包含二元关系符号$\leq$的签名， $\mathcal{P}$ 是 σ-结构的一个性质，并且对于任意具有性质 $\mathcal{P}$的 σ-结构 $\mathbb{A}$， $\leq^{\mathbb{A}}$ 都是其论域上的全序。那么，当且仅当 $\mathcal{P}$ 可用最小不动点逻辑定义时，它可在多项式时间内判定。

通常，将有限结构上的运算视为算法会更容易理解，同时默认该算法计算出的性质可以用给定的逻辑表达。本文中，我们通常不会详细说明这种转换的具体方式。Immerman-Vardi 定理保证，只要我们以某种方式在论域上定义了一个全序，就可以将此顺序叠加到结构上，并将定理 1 应用于这个“隐式”结构。这样，我们就能成功地将算法中所有多项式时间的步骤转换为最小不动点逻辑。

2.2 转导

转导是一种从一个模型定义另一个模型的方法，两者的签名可能不同。在关系结构的语境下，我们如下定义转导。设 σ 和 τ 是签名，其中 $\tau = \langle R_1, R_2, \ldots, R_t \rangle$，且 R_i 的元数为 a_i。那么，从 σ 到 τ 的转导（也称为在 σ 中对 τ 的解释）是一列 σ-公式 $\Theta = (\theta_1, \theta_2, \ldots, \theta_t)$，其中每个 θ_i 都有 a_i 个自由变量 $x_1, x_2, \ldots, x_{a_i}$。为简化记法，我们将 $\Theta : \mathsf{Str}[\sigma] \to \mathsf{Str}[\tau]$ 视为一个函数，将 σ-结构映射为 τ-结构。给定任意一个 σ-结构 $\mathbb{A}$，我们定义 $\Theta(\mathbb{A})$ 为一个与 $\mathbb{A}$ 具有相同论域的 τ-结构；当且仅当 $\mathbb{A} \models \theta_i(x_1, x_2, \ldots, x_{a_i})$ 成立时，关系 $R_i^{\Theta(\mathbb{A})}$ 在元组 $(x_1, x_2, \ldots, x_{a_i})$ 上成立。在图和超图的语境下，我们用 $\theta(G)$ 表示基于公式 θ 创建单个关系的转导。

2.3 奇偶算子

我们可以为任意逻辑 L 添加一个 奇偶性 算子，从而形成新的逻辑 $L[\oplus]$。表达式 $\oplus y\ \varphi$ 询问论域 y 中是否有奇数个元素满足子公式 φ。

该算子显著增强了随机结构上的最小不动点逻辑 LFP 的表达能力。特别是，它使得我们能够以概率 $1 - \mathrm{negl}(n)$ [17] 在论域上定义一个全序。在附录 B 中，我们证明了这一结果的以下变体：它表明，我们几乎总能使用一个元数为 $k \geq 2$ 的关系，同时对论域进行排序并提取 $\Omega(n^k)$ 个独立随机比特。随后，Immerman-Vardi 定理使我们能够对这些比特应用多项式时间计算。

引理 2。设 σ 是一个关系签名，其中包含一个元数为 $k \geq 2$ 的关系符号 R；设 τ 是一个关系签名，其中包含元数为 2 的符号“$\leq$”和元数为 k 的符号 Q。存在一个 最小不动点逻辑 $[\oplus]$ 转导 Γ：结构类 $[\sigma] \to$ 结构类 $[\tau]$，使得对于随机 σ-结构 $\mathbb{A}$，以概率 $1 - \mathrm{negl}(n)$，

(i) $\leq^{\Gamma(\mathbb{A})}$ 在论域上定义一个全序，并且

(ii) $\left|Q^{\Gamma(\mathbb{A})}\right| = \Omega(n^k)$.

此外，对于 k-元组 $(x_1, x_2, \ldots, x_k) \in Q^{\Gamma(\mathbb{A})}$，即使在对 $\leq^{\Gamma(\mathbb{A})}$ 和 $Q^{\Gamma(\mathbb{A})}$ 的任意实现进行条件化之后，$(x_1, x_2, \ldots, x_k) \in R^{\mathbb{A}}$ 均为 $\frac{1}{2}$，并且在 $(x_1, x_2, \ldots, x_k) \in Q^{\Gamma(\mathbb{A})}$ 之间两两独立。 –

2.4 普通伪随机性与密码学

关于本文所使用的密码学概念，读者可参阅 Boneh 和 Shoup [8] 合著的教材，其中有全面的讨论。这里我们只作简要概述。

一个伪随机生成器是函数集合 $f_n : \{0,1\}^n \to \{0,1\}^{\ell(n)}$，满足对于任意多项式规模电路集合 $A_n : \{0,1\}^{\ell(n)} \to \{0,1\}$,

$$\left| \Pr_{s \sim \{0,1\}^n} [A_n(f_n(s)) = 1] - \Pr_{t \sim \{0,1\}^{\ell(n)}} [A_n(t) = 1] \right| = \mathrm{negl}(n).$$

我们也可以将函数集合 A_n 视为一台能够访问“建议”位的单个图灵机（为每个 n 存储一个电路），该图灵机在多项式时间内运行（以模拟该电路）。“建议”位可以以任意（不一定是可计算的）方式依赖于 n 。

人们相信，存在伪随机生成器，可以将 n 个随机位扩展为 $\ell(n) > n$ 个伪随机位。

猜想 3。 存在一个多项式时间伪随机生成器 f_n： $\{0,1\}^n \to \{0,1\}^{n+1}$。

已知这等价于密码学中的一系列其他猜想，包括:

- 单向函数（OWF）的存在性。这类函数易于计算，但给定随机输入的输出后，难以求出其逆。
- 存在一种计算安全的私钥加密方案，可供持有预先分发的短密钥的双方使用。
- 即使面对选择明文或选择密文攻击，也能提供更强的安全性保证；在这类攻击中，我们假设对手可以通过查询来访问对其他消息的加密或解密操作。

这种等价关系中最难的一部分，源于 H˚astad、Impagliazzo、Levin 和 Luby 的著名成果 [18]，，该成果展示了如何从任意单向函数构造伪随机生成器。由于单向函数是更基础、更简单的对象，猜想3 通常被称为单向函数猜想。该猜想意味着 P $\neq$NP。

请注意，在这些密码学对象各自的定义中，我们必须始终允许对手是非一致的，否则某些归约就无法顺利进行。例如，一种经典归约称为长度扩展定理[4] ，它将一个伪随机生成器 f_n : $\{0,1\}^n \to \{0,1\}^{n+1}$ 迭代多项式 $p(n)$ 次，从而生成一个伪随机生成器 g_n :$\{0,1\}^n \to \{0,1\}^{p(n)}$。每次迭代后，我们输出一位，并将剩余的 n 位送回 f_n ，用于下一次迭代，直到输出 $p(n)$ 位。安全性的证明使用混合论证，该论证考察 $p(n)$ 位字符串上的概率分布 H_0， H_1， H_2，……， $H_{p(n)}$；其中，在 H_i中，前 i 位均匀随机选取，其余位则按照定义中的方式，通过迭代 f_n 生成。只需证明，对于每个 i，没有对手能以非可忽略的优势区分从 H_i 抽取的字符串与从 H_{i+1} 抽取的字符串；由此可知，没有对手能区分 H_0 与 $H_{p(n)}$，这意味着 g_n 是安全的。这是因为，任何能够区分 H_i 与 H_{i+1} 的对手，都可以转换成一个破坏 f_n安全性的对手。然而，该算法需要随机性，而目前唯一已知的去随机化方法，是将随机位的一次良好取值编码为非一致建议。

定理17的一种情形的证明中，我们需要对长度扩展构造作出一种新颖的改造，使其能够实现为最小不动点逻辑$[\oplus]$-转导。

2.5 模型论伪随机性

考虑一组概率分布 $(\mathcal{D}_n)^{n\in\mathbb{Z}_{\geq 1}}$，其中 $\mathcal{D}_n$ 遍历结构类 $\mathrm{Str}[\tau, n]$。现在，对于逻辑 L ，如果对每个 τ-句子 $\varphi \in L$,

$$\left| \Pr_{\mathbb{A}\sim\mathcal{D}_n}[\mathbb{A} \models \varphi] - \Pr_{\mathbb{A}\sim \mathrm{Str}[\tau,n]}[\mathbb{A} \models \varphi] \right| = \mathrm{negl}(n).$$

[4]参见 Boneh 和 Shoup [8, 定理 3.3]。

普通伪随机生成器的概念中，对手可以访问多项式数量的非均匀建议位；而这个模型论定义则完全是均匀的，也没有提及建议。几乎所有结果都不需要建议（具体而言，除了定理 17 的第 (i)部分外，所有结果都是如此），这可以视为该模型的一项优势。不过，为了将我们的结果与密码学基础猜想联系起来，我们也会让逻辑对手能够访问建议。这是通过互不相交的多项式大小建议结构 $(\mathbb{X}_n)^{n\in\mathbb{Z}_{\geq 1}}$ 来实现的；这些结构可以非均匀地依赖于 n。

对于一个 τ-结构 $\mathbb{A}$ 和一个 ρ-结构 $\mathbb{X}_n$，我们用 $(\mathbb{A}, \mathbb{X}_n)$ 表示通过将这两个结构进行不交并[5]而得到的 $(\tau \cup \rho)$-结构。若一个分布 $(\mathcal{D}_n)^{n\in\mathbb{Z}_{\geq 1}}$ 对于逻辑 L是伪随机的，则对于每个建议签名 ρ、每个 ρ-结构序列 $(\mathbb{X}_n)^{n\in\mathbb{Z}_{\geq 1}}$（其中 $\mathbb{X}_n$ 的论域大小为 n）以及每个 $(\tau \cup \rho)$-句子 $\varphi \in L$,

$$\left| \Pr_{\mathbb{A}\sim\mathcal{D}_n} [(\mathbb{A}, \mathbb{X}_n) \models \varphi] - \Pr_{\mathbb{A}\sim \mathrm{Str}[\tau,n]} [(\mathbb{A}, \mathbb{X}_n) \models \varphi] \right| = \mathrm{negl}(n).$$

最后，对于任意一对逻辑 L_1 和 L_2，从 σ 到 τ 的 (L_1, L_2)-伪随机生成器 是一个 L_1 转导 Θ： $\mathrm{Str}[\sigma] \to \mathrm{Str}[\tau]$，使得通过对$\mathbb{C}_n \sim \mathrm{Str}[\sigma, n]$ 进行抽样并输出 $\Theta(\mathbb{C}_n)$ 所得到的分布族对于 L_2 是伪随机的。

3 最小不动点逻辑和扩展公理的伪随机结构

关于一阶逻辑或最小不动点逻辑中的对手，构造伪随机结构的一项核心性质是所谓的扩展公理。目前，我们先为图定义这些公理，之后再给出关系结构更复杂的定义。

定义 4。图 G 满足 k-扩展公理 EA_k，如果对于所有大小为 k的集合 $S \subseteq V(G)$ 以及所有 $T \subseteq S$，都存在 $v \notin S$（称为 (S, T) 的扩展顶点），它与 T中的每个顶点都相邻，并与 $S \setminus T$中的每个顶点都不相邻。

非正式地说， k-扩展公理指出，在 Ehren- feucht–Fra¨ıss´e 一阶逻辑（有限或无限）的博弈中，任意一组 棋子都可以通过增加一枚棋子，以所有可能的方式扩展。[6] 因此，如果两个结构都满足 k-扩展公理，那么无论棋子位于何处，Duplicator 总能找到应对之策。换言之，对于任意给定的最小不动点逻辑句子 φ，都存在 k，使得对于任意一对图 G_1， G_2 只要二者都满足 k-扩展公理 EA_k，就有 $G_1 \models \varphi \Leftrightarrow G_2 \models \varphi$。虽然这是一个标准论证，但为求完整，我们将在引理 11 中证明此结论（即使存在额外的建议结构亦然）。

一个重要的观察是，随机图以高概率满足扩展公理。正如我们在引理10中所展示的，对于每个正整数 k,

$$\Pr_{G\sim G(n,1/2)} \left[G \not\models \mathrm{EA}_k \right] = \mathrm{negl}(n).$$

FO [14] 和最小不动点逻辑 [5] 的零一律随即成立：任何一个量词秩为 k 的句子，要么在所有满足 k扩张公理的图中成立，要么在其中都不成立，因此在

[5]（$\mathbb{A}$， $\mathbb{X}_n$）的论域是不交并，由 $\mathbb{A}$ 和 $\mathbb{X}_n$ 的论域组成； $\tau \cup \rho$ 中的每个关系都解释为这两个结构中相应关系的不交并，其中缺失的关系视为空关系。[6]虽然理解本文并不需要这些内容，但如需进一步了解卵石游戏的背景，读者可参阅 Libkin [24] 。

几乎所有或几乎没有图中成立。特别是，无限随机图 $G(\infty, 1/2)$（也称为拉多图）以概率 1 满足所有 $k \in \mathbb{Z}_{\geq 1}$ 的 k扩张公理。

在本节中，我们给出一种多项式时间确定性构造，可构造满足 $k =$的 n 元图和关系结构，使其满足 k-扩展公理；这意味着一阶逻辑和最小不动点逻辑的对手无法将它们与真正随机的图和结构区分开来。

通用集与完美哈希函数 以下定义提供了一种自然的方法：仅使用对数数量的元素，从一个大小为线性的集合中"扩展"大小为 k 的集合。

定义 5。设 $\mathcal{F}$ 是一个从 $[n]$ 到 $[k]$的函数族。若对于每个大小为 k 的集合 $S \subseteq [n]$，都存在 $f \in \mathcal{F}$ 使得 $\{f(s) \mid s \in S\} = [k]$，则称 $\mathcal{F}$ 为一个 (n,k)-完美哈希函数族。若对于 $[n]$ 的每个大小为 k 的子集 $S \subseteq [n]$，集合族 $\{S \cap U\colon\ U \in \mathcal{U}\}$ 包含 S 的所有 2^k 子集，则称由 $[n]$ 的子集构成的族 $\mathcal{U}$ 为一个 (n,k)-通用集。

小规模的完美哈希函数和通用集都存在，并且可以高效地构造。我们使用以下经典结果。

定理 6（Naor、Shulman 和 Srinivasan [26]）。存在一个常数 c 使得，对于 n，$k \geq 1$，可以构造一个（n，k）-完美哈希函数族和一个（n，k）-通用集，其大小为$2^{ck}\log(n)$，所需时间为 $2^{ck}n\log(n)$。

锦标赛 A锦标赛 是无向完全图的一种边定向。这意味着它是一个有向图，对于每一对顶点 a，b 要么包含从 a 到b 的有向边，要么包含从 b 到 a 的有向边，但不会同时包含两者。通过穷举，我们找到了一个大小为 $\log(\Theta(n)$的锦标赛图，它为每个大小为 k 的集合预留一个指向该集合中所有元素的顶点，并利用它来分配扩展顶点，如图1所示。

引理 7。令 k 为一个非负整数。存在一个大小为 2^{3k} 的锦标赛 F，使得对于每个大小为 k的 $S \subseteq V(F)$，都存在一个顶点，该顶点通过有向边指向 S中的每个顶点。

证明。考虑一个随机锦标赛 F，其上有 2^{3k} 个顶点，其中每条边的方向都独立且均匀地随机选取。我们使用概率方法，通过证明存在非零概率使得 F 具有所述性质，来证明该命题。固定一个集合 $S \subseteq V(F)$，其大小为 k。一个固定顶点的边均指向 S 中每个顶点的概率恰好为 2^{-k}。不存在某个顶点的边均指向 S 中每个顶点的概率（利用附录A中的观察24）至多为

$$(1-2^{-k})^{2^{3k}-k} \leq e^{-(2^{3k}-k)/2^k}.$$

根据并集界，对于某个大小为 k的 S 个顶点集合，不存在某个顶点的边均指向 S 中每个顶点的概率至多为

$$\binom{2^{3k}}{k} \cdot e^{-(2^{3k}-k)/2^k}.$$

基本微积分表明，对于所有 k，该项都小于 1。这意味着 F 具有所述性质的概率为非零概率。

□

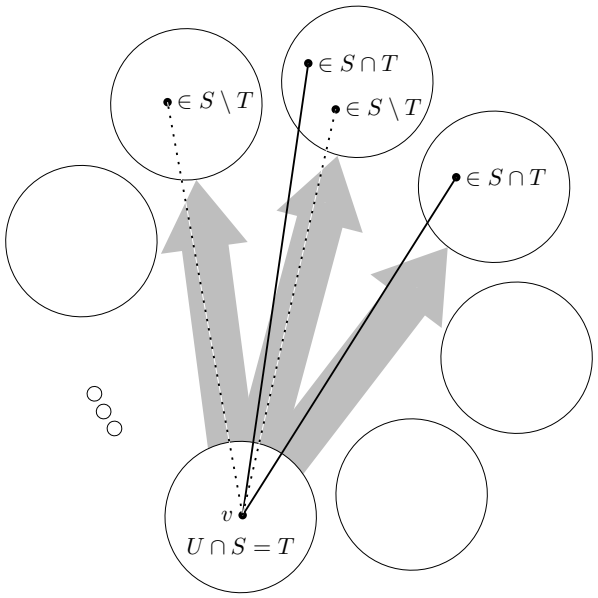


图 1：我们构造的有限拉多图。要为集合 S 和 T 找到一个扩展顶点，首先选择一个锦标赛节点，使其有有向边（粗灰色箭头）指向所有包含 S 部分的锦标赛节点。在这个锦标赛节点内，从一个包含 $U \cap S = T$ 的 $(n,\ k)$-通用集中选择一个集合 U。存在一个满足 $N(v) \cap S = T$ 的扩展顶点 v。

3.1 图

We p首先介绍我们针对图的构造，之后介绍针对关系结构的构造 .

定理 8。存在一个常数 c，使得对于每个 n，都可以在时间 $O(n^3)$ 内构造一个满足扩张公理 EA_k 的 n 顶点图，对所有 $k \leq \lfloor \log(\log(n))/c \rfloor$ 均成立。

由于扩张公理以高概率成立，并且决定 LFP 句子的真值，因此立即可知，LFP 无法区分真正的随机图 $G(n,\ 1/2)$ 与定理8的伪随机输出。我们将在定理 12 中对此进行形式化。

定理 *8* 的证明。 我们证明下面这一命题，它可以推出该定理：存在一个常数c，使得对于每个 n，$k \in \mathbb{Z}_{\geq 1}$ 满足 $2^{k^c} \leq n$时，都可以在 $O(2^{2^{ck}} n^2)$ 时间内构造一个满足 EA 的 n-顶点图$_k$。

顶多有 $2^{2^{3k \cdot 2}}$ 个具有 2^{3k} 个顶点的锦标赛。通过枚举所有这些锦标赛，根据引理 7，我们找到一个锦标赛 F，其顶点集为 $[2^{3k}]$，满足：对于每个 $S \subseteq [2^{3k}]$ 大小为 k的 $j \in [2^{3k}]$，都存在一个顶点 S，使其向 S中的每个顶点都有一条有向边。

我们将构造一个顶点集为 $[n]$ 的图 G，满足扩展公理 EA_k。首先利用定理 6 构造一个 $(n,\ k)$-通用集 $\mathcal{U}$。接下来，我们需要将顶点集 $[n]$ 划分为 2^{3k} 个部分 P_1, ..., $P_{2^{3k}}$，每个部分的大小至少为 $|\mathcal{U}|$（这些部分将对应于锦标赛中的顶点）。下面说明，可以选择参数 c，使得对所有 $n \geq 2^{k^c}$ 都能做到这一点。设 c' 为定理 6 中的常数，从而可以将我们的 $(n,\ k)$-通用集的大小界定为 $2^{c'k} \log(n)$。我们可以选择 c，使得对于每个 $k \in \mathbb{Z}_{\geq 1}$,

$$2^{3k} \cdot 2^{c'k} \cdot 2 \leq 2^{k^c}/k^c.$$

因此，对于每个 $n \in \mathbb{Z}_{\geq 1}$，$2^{k^c} \leq n$ 成立 $2^{k^c}/k^c \leq n/\log(n)$，从而

$$2^{3k} \cdot 2^{c'k} \log(n) \cdot 2 \leq n.$$

这意味着我们可以将顶点划分为若干部分 P_1、……、$P_{2^{3k}}$，每部分的大小至少为

$$\lfloor 2^{c'k} \log(n) \cdot 2 \rfloor \geq |\mathcal{U}|.$$

接下来，我们要选择一个函数 $\textsc{pattern}\colon [n] \to \mathcal{U}$。由于我们有 $|P_j| \geq |\mathcal{U}|$，因此可以保证

$$\{\textsc{pattern}(v) \mid v \in P_j\} = \mathcal{U}$$

对所有 $j \in [2^{3k}]$ 都成立。我们进一步用 $\mathsf{index}(v)$ 表示满足 $v \in P_j$ 的索引 j。为了构造图 G，现在对所有的点对 $\{u, v\} \in \binom{[n]}{2}$ 执行以下操作：

> 不妨通过重新命名，假设锦标赛 F 中的弧从 $\mathsf{index}(v)$ 指向 $\mathsf{index}(u)$。当且仅当 $u \in \textsc{pattern}(v)$ 时，向 G 添加边 $\{u, v\}$。

正确性 设 $S \subseteq [n]$，且令 $T \subseteq S$。我们利用引理 7 构造了锦标赛 F，使得可以选取一个顶点 $j \in [2^{3k}]$，它有有向边指向 $\{\mathsf{index}(s) \mid s \in S\}$ 中的所有顶点。由于 $\mathcal{U}$ 是一个 (n,k)-通用集，存在 $U \in \mathcal{U}$ 使得 $S \cap U = T$。由于 $\textsc{pattern}$ 在 P_j 上的限制被选为满射，我们可以选取 $v \in P_j$，使得 $\textsc{pattern}(v) = U$。由于 F 没有自环，$v \notin S$。我们选取了 v，使得对于所有 $u \in S$，$\mathsf{index}(v)$ 在 F 中有一条指向 $\mathsf{index}(u)$ 的有向边。因此，v 与 $u \in S$ 相邻，当且仅当 $u \in \textsc{pattern}(v) = T$。

运行时间 构造 F 所需时间至多为 $2^{2^{3k \cdot 2}} \cdot k^2$。根据定理 6，我们可以在 n, k)-通用集所需的时间 $2^{c'k} n \log(n$内构造出该集合。构造的其余部分所需时间为$O(|\mathcal{U}| \cdot n^2)$。通过适当缩放 c，我们得到运行时间 $O(2^{2^{ck}} n^2)$。 □

3.2 关系结构

对于一般签名 k，σ_k -扩张公理 EA 的定义 σ 变得更加复杂，因此我们的有限拉多构造也更加晦涩，尽管最终使用的论证与图的情形相同。给定集合 S，我们找到一个锦标赛节点，其方向指向包含 S 的锦标赛节点。现在，我们不再使用通用集，而是从一族 f (n, k)-完美哈希函数中选取一个函数，该函数在 S 上是双射。我们的构造添加一个扩张元素，使其对所有满足 S 上 f 为双射的集合都有效。关于 k-扩张公理 EA_k^σ（定义25）的定义以及定理 9的证明，见附录C。

定理 9。 设 σ 是一个关系签名。存在一个常数 c，使得对于每个 n，都可以在时间 $O(n^c)$ 内构造一个满足扩张公理EA_k^σ （对于所有 $k \leq \lfloor \log(\log(n))/c \rfloor$）的 n 元 σ 结构。

虽然我们没有证明这一点，但定理 9 很容易修改为也能生成伪随机超图。正如我们将在下一节看到的，与图的情形一样，扩张公理决定了一阶逻辑和最小不动点逻辑公式的真值，并且以高概率成立。

3.3 伪随机图与结构

现在让我们讨论为什么定理 8和定理 9的构造能为一阶逻辑和最小不动点逻辑生成伪随机结构。首先请注意，随机图、超图和关系结构都很容易满足相应的扩张公理。

引理 10。 对于每个签名 σ 、每个 k 和每个 t,

$$\Pr_{G\sim G(1/2,n)}\left[G \not\models \mathrm{EA}_k\right] = \mathrm{negl}(n),$$
$$\Pr_{G\sim G_t(1/2,n)}\left[G \not\models \mathrm{EA}_k^t\right] = \mathrm{negl}(n),$$
$$\Pr_{\mathbb{A}\sim \mathrm{Str}[\sigma,n]}\left[\mathbb{A} \not\models \mathrm{EA}_k^\sigma\right] = \mathrm{negl}(n).$$

这是一个众所周知的结果 [11, 第 4.1 节和第 4。2], 节，但为完整起见，我们在附录D 中给出证明。

接下来，我们通过卵石游戏（例如参见 [25, 引理 12。7]）论证扩展公理决定一阶逻辑或最小不动点逻辑句子的真值。这个证明同样直接明了，并收录于附录 E中。

引理 11。 考虑一对具有相同签名的图/超图/结构 $\mathbb{A}_1$、 $\mathbb{A}_2$ ， 它们都满足相应的 k扩张公理。于是，对于每个建议结构 $\mathbb{X}$，以及每个量词秩至多为 k 且签名匹配的最小不动点逻辑 句子 φ ，有 ($\mathbb{A}_1$、 $\mathbb{X}$) $\models \varphi \Leftrightarrow (\mathbb{A}_2$、 $\mathbb{X}) \models \varphi$。

现在，引理 10 和引理 11 连同定理 8 （针对图）以及定理 9（针对结构）立即推出本节的主要定理。

定理 12. 存在确定性的多项式时间算法，可以构造对于一阶逻辑 和最小不动点逻辑而言是伪随机的图或关系结构。

4 伪随机转导

现在我们转而探讨转导能否充当伪随机生成器。具体而言，我们要问：对于哪些逻辑 L_1 和 L_2 ，以及哪些关系签名 σ 和 τ ，存在从 σ 到 τ的 (L_1, L_2)-伪随机转导，也就是说，存在一种能够抵御 L_2的 L_1-转导。本节将对所有逻辑对 $L_1, L_2 \in$ {一阶逻辑,最小不动点逻辑,最小不动点逻辑[$\oplus$]}以及所有关系签名给出完整分类。

显然，这类转导是否存在取决于 σ 和 τ中关系的数量及其元数。例如，如果 τ 的元数多重集是 σ元数多重集的子集，我们就可以将这些关系符号与如下形式的公式简单地一一对应

$$\theta_i(x_1, x_2, \ldots, x_{a_i}) = R_j(x_1, x_2, \ldots, x_{a_i})$$

并忽略来自 σ的额外关系符号。这一转导在任何逻辑中都是可定义的，并将产生一个均匀随机的 τ-结构，因此对于任何逻辑而言也都是伪随机的。更有趣的情况是，目标 τ-结构本质上比输入σ-结构具有更多“熵”，例如从单个二元关系转变为单个三元关系。我们针对这种情况的结果总结于表 1中。

在其他情况下，如何衡量熵并不显而易见。例如，是否存在一个从具有一个三元关系的签名到具有

Θ \ φ	FO	LFP	最小不动点逻辑[$\oplus$]
FO	$\sigma \succeq_S \tau$	$\sigma \succeq_S \tau$	$\sigma \succeq_S \tau$
LFP	$\sigma \succeq_S \tau$	$\sigma \succeq_S \tau$	$\sigma \succeq_S \tau$
最小不动点逻辑[$\oplus$]	$\sigma \succeq_L \tau$ 或 σ 非全一元	$\sigma \succeq_L \tau$ 或 σ 非全一元	$\sigma \succeq_L \tau$ 或 (σ 并非全一元且存在单向函数)

表 2：关于签名的条件 σ 以及 τ 从L_1、 L_2到 σ 的 τ伪随机生成器存在的条件。行对应生成器逻辑 L_1 ，列对应对手逻辑 L_2。这是对表1的推广。

十个二元关系的签名的伪随机转导？有趣的是，我们发现，衡量熵的“正确”方式取决于转导逻辑 L_1；在这一特定情况下，FO 和 LFP 的答案是否定的，而 LFP[$\oplus$] 的答案是肯定的。

设 $\sigma = \langle R_1, R_2, \ldots, R_t \rangle$ 和 $\tau = \langle R'_1, R'_2, \ldots, R'_{t'} \rangle$ 是签名，其中 R_i 的元数为 a_i，而R'_i 的元数为 a'_i。我们将证明，对于 LFP[$\oplus$] ，恰当的熵度量是以下“字典序” $\succeq_L$。当且仅当 $\sigma \succeq_L \tau$ 的元数组 (a_1， a_2，……， a_t) 按降序排列后，按（非严格）字典序大于 τ对应的元组 (a'_1， a'_2，……， $a'_{t'}$) 时，我们称 $\sigma \succeq_L \tau$。等价地，

$$\sigma \succeq_L \tau \iff \sum_{i=1}^{t} (t+t')^{a_i} \geq \sum_{i=1}^{t'} (t+t')^{a'_i}.$$

正如预期， $\succeq_L$ 是一个全序，与我们对随机结构中熵的直观理解相符：为一个 a元关系赋予 n^a的熵，而这一熵渐近地大于任意由 $b < a$构关系组成的组合的熵。

当转导逻辑 L_1 是 FO 或 LFP 时，情况会变得更加复杂。在这里，正确的熵度量将是由下式定义的序 $\succeq_S$：

$$\sigma \succeq_S \tau \iff \text{for all } k \in \mathbb{Z}_{\geq 1}, \sum_{i=1}^{t} T(a_i, k) \geq \sum_{i=1}^{t'} T(a'_i, k),$$

其中

$$T(a,k) := \sum_{j=0}^{k} (-1)^{k-j} j^a \binom{k}{j}$$

计算从大小为 a 的集合到大小为 k [15, 的集合]（6.19）的满射数量。注意， $\succeq_S$是 $\succeq_L$的细化，而不是全序。也就是说，从 $\succeq_L$ 的意义上具有更多熵，是存在伪随机生成器的必要条件，但不是充分条件。直观地说，这是因为 FO 和 LFP 的能力不足以将输入熵的某些“形状”转换为所需的输出形状。

我们在表2中总结了伪随机转导的存在性如何取决于签名的逻辑和熵，由此表1 可作为特殊情形推导出来。

4.1 LFP 的不可能性

现在我们陈述并　概述以下结果的证明，并填写表格2的上三分之二　　.

定理 13。 设 $\sigma = \langle R_1, R_2, \ldots, R_t \rangle$ 和 $\tau = \langle R'_1, R'_2, \ldots, R'_{t'} \rangle$ 为关系签名，其中R_i 的元数为 a_i ，R'_i 的元数为 a'_i。以下陈述等价：

(i) 存在一个从 σ 到 τ 的 (最小不动点逻辑，一阶逻辑)伪随机生成器。

(ii) 存在一个无量词一阶逻辑转导 Θ : 结构类$[\sigma] \to$ 结构类$[\tau]$ ，使得对于所有 n，将 Θ 应用于 $\mathbb{A} \sim$ 结构$[\sigma, n]$ 所得到的 结构类$[\tau, n]$ 上的分布，在统计上与分布 $\mathbb{B} \sim$ 结构类$[\tau, n]$相同。

(iii) $\sigma \succeq_S \tau$。

要理解这为何补全了表 2中的一阶逻辑和最小不动点逻辑两行，首先假设 $\sigma \succeq_S \tau$。那么根据 (ii)，存在一个一阶逻辑转导（因而也是一个最小不动点逻辑转导） Θ ，它生成的 τ-结构在统计上与随机 τ-结构相同。因此，即便计算能力不受限制的对手也无法区分这两种情况，所以 Θ 能够抵御任何对手逻辑，从而使最上面的两行均得到肯定结果。另一方面，如果我们没有 $\sigma \succeq_S \tau$，那么根据 (i)，即便由最小不动点逻辑定义的生成器，也无法骗过一阶逻辑中的每个对手；较弱的生成器和较强的对手也同样如此，因此最上面的两行均得到否定结果。

这个定理可以视为关于最小不动点逻辑的不可能性定理。 (i)与 (ii) 等价，这意味着不存在“新颖的”（最小不动点逻辑，一阶逻辑）伪随机生成器：任何可由最小不动点逻辑定义的转导，只要它对较弱的一阶逻辑也安全，就可以替换为一种极其简单的替代转导，其中所有公式都是一阶逻辑公式，且不含量词。而这种简单的转导“安全性已达极致”：其输出分布不仅在计算上与真正的随机分布不可区分，而且就是真正的随机分布。因此，在此情境下，最小不动点逻辑的所有额外表达能力对于生成伪随机结构都无关紧要。

正如我们先前所指出的，在这些逻辑中用于刻画伪随机生成器的熵的概念 $\succeq_S$ 只是一个偏序。例如，如果 σ 由一个二元关系构成，而τ 由两个一元关系构成，那么定理 13 意味着不存在从 σ 到 τ 或从 τ 到 σ的伪随机生成器，因为

$$\sum_{i=1}^{t} T(a_i, 2) = 2 > 0 = \sum_{i=1}^{t'} T(a'_i, 2)$$

but

$$\sum_{i=1}^{t} T(a_i, 1) = 1 < 2 = \sum_{i=1}^{t'} T(a'_i, 1).$$

完整证明篇幅很长且技术性很强，因此推迟至附录F。这里我们概述全部 3 个情形的论证。

（ii） $\implies$ （i）：这显然成立，因为具有 Θ 中所述性质的转导（ii）是任意L_1， L_2-伪随机生成器，其中 L_1 和 L_2 满足 L_1 扩展无量词一阶逻辑。

(iii) $\implies$ (ii)：考虑以下示例。假设我们要构造一个转导，将由三个元数分别为 $a_1 = 3$、 $a_2 = 1$、$a_3 = 1$ 的关系组成的输入签名 σ 转换为输出签名 τ ，后者也包含三个关系，但元数为 $a'_1 = a'_2 = a'_3 = 2$。如果 σ 有一个二元关系，我们就可以直接用它填充 τ 中的一个关系；但由于没有二元关系，我们必须使用三元关系的以下特化形式： $R_1(x, x, y)$、 $R_1(x, y, x)$ 和 $R_1(y, x, x)$。对于 $x \neq y$，这些元组将彼此独立地属于关系 R_1 ，因此可以分别用来填充 τ 中的三个二元关系。然而，这种处理只涵盖了关系中的“二元部分”，也就是涉及两个不同变量的部分。为了确定哪些关系 $R'_i(x, x)$

可以保留，我们只能使用一次 $R_1(x,\ x,\ x)$，因此另外两个必须使用一元关系 $R_2(x)$ 和 $R_3(x)$。于是，将二元部分和一元部分分开处理后，我们得到一个无量词的一阶逻辑转导 $\Theta = (\theta_1,\ \theta_2,\ \theta_3)$，它能将均匀随机的 σ 结构转换为均匀随机的 τ 结构：

$$\begin{aligned}
\theta_1(x_1, x_2) &:= R_1(x_1, x_1, x_2) \\
\theta_2(x_1, x_2) &:= \big((x_1 \neq x_2) \wedge R_1(x_1, x_2, x_1)\big) \vee \big((x_1 = x_2) \wedge R_2(x_1)\big) \\
\theta_3(x_1, x_2) &:= \big((x_1 \neq x_2) \wedge R_1(x_1, x_2, x_2)\big) \vee \big((x_1 = x_2) \wedge R_3(x_1)\big).
\end{aligned}$$

更一般地，$\sigma \succeq_S \tau$ 这一条件相当于检查：对于 k每一种可能的不同变量个数，σ 中是否有足够多的关系的 k元部分，能够覆盖 τ中关系的 k元部分。

$\neg$(iii) $\Longrightarrow$ $\neg$(i)：这是论证中最困难的部分，我们必须在此处运用零一律来证明：无论最小不动点逻辑转导 $\Theta = (\theta_1, \theta_2, \ldots, \theta_{t'})$ 做什么，我们都可以构造一个一阶逻辑句子 φ，使其输出与真正随机的 τ-结构区分开来。为此，我们必须对每个公式 θ_i 进行处理，将其改写成零一律适用的形式。例如，假设其中一个公式是

$$\theta_i(x_1, x_2, x_3) = \forall x_4\ R_1(x_3, x_4).$$

通过区分 是否 $x_4 \in \{x_1, x_2, x_3\}$，我们首先将这个句子等价地改写为 s

$$\theta_i(x_1, x_2, x_3) = R_1(x_3, x_1) \wedge R_1(x_3, x_2) \wedge R_1(x_3, x_3) \wedge \forall x_4 \notin \{x_1, x_2, x_3\}\ R_1(x_3, x_4).$$

然后我们注意到，$R_1(x_3, \cdot)$ 可以看作是在移除 x_1、x_2和 x_3后得到的结构上的一个随机一元关系。因此，根据应用于该结构的零一律，我们知道这个句子的最后部分

$$\forall x_4 \notin \{x_1, x_2, x_3\}\ R_1(x_3, x_4),$$

可以替换为 TRUE 或 FALSE（此处为 FALSE），所得公式以概率 $1 - \mathrm{negl}(n)$ 等价。这样，我们可以从公式中移除所有量词。对多项式数量的可忽略误差项应用并集界，可得所得无量词公式 θ_i 以概率 $1 - \mathrm{negl}(n)$ 等价。 –

最后，我们令 k 为一个不满足条件 $\sigma \succeq_S \tau$ 的正整数，并选取一个较大（但为常数）的正整数 c。我们写出一个句子 φ，用于检验以下条件：“对于一个由 c 个互异论域元素组成的 c-元组（y_1、y_2、…、y_c）在仅使用来自 $\{y_1, y_2, \ldots, y_c\}$ 的至多 k 个互异变量时，相对于关系可能呈现的每一种行为，都确实存在一个以该方式呈现的 c-元组。”显然，真正随机的 τ-结构以概率 $1 - \mathrm{negl}(n)$ 满足 φ。另一方面，我们证明，对于足够大的 c，$\Theta = (\theta_1$、θ_2、…、$\theta_{t'})$ 生成满足 φ 的结构的概率为 0。此步骤由以下引理严格证明；该引理的证明见附录 G，之后也会用于证明当 σ 仅含一元关系时，最小不动点逻辑 LFP$[\oplus]$ 的不可能性。

定义14。给定整数 $1 \leq k \leq c$、具有元数 a_1、a_2、……、a_t的关系签名 $\sigma = \langle R_1, R_2, \ldots, R_t \rangle$、一个 σ-结构 $\mathbb{A}$ 以及互异元素 y_1、y_2、……、$y_c \in \mathbb{A}$，(c、k)-型(y_1、y_2、……、y_c) 在 $\mathbb{A}$ 中是满足以下条件的子集

$$T \subseteq \bigcup_{i=1}^{t} \bigcup_{\substack{S \subseteq [c] \\ |S| \leq k}} \{(i, S, g) \mid g : [a_i] \to S \text{ is surjective}\}$$

使得 T（通过对 $(y_1、y_2、\ldots\ldots、y_c)$ 进行索引）准确描述了 $\mathbb{A}$ 中哪些关系成立于包含 $(y_1、y_2、\ldots\ldots、y_c)$ 中至多 k 个互异元素的元组上。换言之，$(i、S、g) \in T$表示 $(y_{g(1)}、y_{g(2)}、\ldots\ldots、y_{g(a_i)}) \in R_i^{\mathbb{A}}$。

引理*15*。设 σ 和 τ 是关系签名，且某个正整数 k 违反了以下条件：$\sigma \succeq_S \tau$。设 $f\colon \mathit{Str}[\sigma] \to \mathit{Str}[\tau]$ 是一个函数，使得对于任意正整数 $c \geq k$ 和任意 σ-结构 $\mathbb{A}$，如果任意一对 c-元组 $(y_1, y_2, \ldots, y_c)$、$(y'_1, y'_2, \ldots, y'_c) \in \mathbb{A}^c$ 在 $\mathbb{A}$ 中具有相同的 (c,k)-类型，那么它们在 $f(\mathbb{A})$ 中的 (c,k)-类型也相同。那么，对均匀随机的 σ-结构应用 f 所得到的分布不是对一阶逻辑伪随机的。

我们可以将这个引理应用于 $f := \overline{\Theta}$，因为它是无量词的，由此可得 Θ 不是伪随机生成器。由于原始转导 Θ 与 Θ 的差异可以忽略不计，因此可推出 Θ 也不是伪随机生成器。

4.2 LFP 的可能性$[\oplus]$

加入奇偶算子后，我们可以规避零一律，因此伪随机转换是可行的。在至少存在一个非一元关系的情况下，我们可以将输入结构规范化，从而利用定理12中的多项式时间构造，为一阶逻辑和最小不动点逻辑定义伪随机输出结构。

推论 16。对于 $L_2 \in \{$一阶逻辑、最小不动点逻辑$\}$，存在一个 (最小不动点逻辑$[\oplus]$, L_2)-伪随机生成器，将 σ 映射到 τ，当且仅当 σ 至少具有一个非一元关系，或者 σ 和 τ 都只有一元关系，且 σ 包含的关系数量不少于 τ。

证明。如果 σ 至少具有一个非一元关系，那么根据引理 $1 - \mathrm{negl}(n)$ 的概率可以将整个输入结构规范化（即定义一个全序）。假设这种情况发生，根据 Immerman-Vardi 定理（定理1），我们拥有多项式时间计算的全部能力，因此可以应用定理 12中的确定性算法。

如果 σ 和 τ 都只有一元关系，且 σ 的数量不少于 τ，那么只需将它们一一对应并忽略多余的部分，显然就存在一个伪随机转导。

如果这两个条件都不成立，就意味着 $\sigma \succeq_S \tau$ 在 $k = 1$上不满足。设 $\Theta : \mathrm{Str}[\sigma] \to \mathrm{Str}[\tau]$ 是任意一个最小不动点逻辑$[\oplus]$ 转导。对于任意整数 c，如果两个 c-元组在一个 σ-结构 $\mathbb{A}$中具有相同的（c, 1）-类型，这意味着对于每个 $1 \leq i \leq c$，两个元组中对应的 i^{th} 元素上的一元关系都相同。由于 σ 只有一元关系，因此存在一个自同构 $\mathbb{A}$ 将一个元组映射到另一个元组。由于 Θ 是一个转导，这意味着在 $\Theta(\mathbb{A})$中存在一个类似的自同构，由此可知这两个元组在 $\Theta(\mathbb{A})$中具有相同的（c, 1）-类型。因此，引理15适用，所以 Θ 不是一阶逻辑或最小不动点逻辑的伪随机转导。 □

鉴于这一结论，表2 中唯一尚未讨论的格子是右下角的格子，即两种逻辑都是最小不动点逻辑时$[\oplus]$。以下结果完全刻画了何时存在（最小不动点逻辑$[\oplus]$，最小不动点逻辑$[\oplus]$）-伪随机生成器，这取决于 σ 和 τ 的比较熵（这次用 $\succeq_L$而非 $\succeq_S$来度量）。

定理 17。设 σ 和 τ 为关系签名。

(*i*) 若 $\sigma \prec_L \tau$和 σ 包含至少一个非一元关系，则存在从（最小不动点逻辑$[\oplus]$，最小不动点逻辑$[\oplus]$）-伪随机生成器，从 σ 到 τ 当且仅当存在长度递增的普通伪随机生成器（即猜想*3* 成立）。

(*ii*) 若 $\sigma \succeq_L \tau$，那么就存在一个从 σ 到 τ 的(最小不动点逻辑$[\oplus]$，最小不动点逻辑$[\oplus]$)-伪随机生成器 .

(iii) 否则，不存在从 σ 到 τ的 $(\mathrm{LFP}[\oplus], \mathrm{LFP}[\oplus])$-伪随机生成器。

最困难也最有趣的陈述是 (i)，其证明涉及在字符串与关系结构之间来回转换，并且这种转换在一阶逻辑$[\oplus]$中是可定义的。事实证明，证明的反向过程相对简单：从普通伪随机生成器构造伪随机转导。可是，正向过程则面临多重挑战。要从伪随机转导 Θ构造普通伪随机生成器，基本思路是利用随机输入位构造一个结构 $\mathbb{A}$，模拟对它应用 Θ，然后读取其中的关系，从而生成比起始时更多的随机位。为证明安全性，我们必须论证：任何能够区分随机字符串与伪随机字符串的普通对手，都可以转化为能够区分随机结构与伪随机结构的逻辑对手。要定义这样的逻辑对手，读取关系中的随机位这一后处理步骤必须能够用一阶逻辑$[\oplus]$ 以序不变的方式定义，也就是说，不能依赖论域元素原本任意指定的次序。当 τ 包含一个额外的k元关系时，对于 $k \geq 2$，我们可以利用引理 2 完成后处理步骤。因此，唯一剩下的情况是：对于所有$k \geq 2$， τ 中的 k元关系数量都与 σ 相同，但一元关系更多。这是迄今为止最具挑战性的问题；我们通过将密码学中的长度扩展定理（第 2.4 节回顾）以一种新颖的方式改造并应用于我们的领域来解决它。具体来说，我们编写一个一阶逻辑$[\oplus]$ 公式，迭代 Θ 多项式数量次，并在每次迭代中从 τ 的额外一元关系中提取一个随机奇偶位，如图 3 所示。

我们唯一使用奇偶算子的地方是应用引理 2 （多次）时。若将 $\mathrm{LFP}[\oplus]$替换为受到广泛研究的带计数的不动点逻辑 （FPC），证明过程完全相同。

定理 *17* 的证明。 (i)：我们定义

$$p(n) := \sum_{i=1}^{t} n^{a_i}, \qquad\qquad p'(n) := \sum_{i=1}^{t'} n^{a'_i}.$$

对于反向方向，假设普通伪随机生成器存在。令 R 为 σ中的一个元数至少为 2 的关系，并令 $f\colon \{0,1\}^n \to \{0,1\}^{p'(n)}$ 为一个伪随机生成器。考虑由以下算法描述的 $\mathrm{LFP}[\oplus]$ 转导。输入 $\mathbb{A}$ 后，我们首先应用引理 2 中的转导，使用 R 以概率 $1-\mathrm{negl}(n)$ 获得论域 $\leq$ 上的一个全序，以及至少 $\Omega(n^2)$ 位来自 R 的独立随机性，这些随机位位于关系 Q所描述的位置。随后，我们根据$\leq$，在前 n 位随机性上模拟 f （通过定理 1），从而获得 $p'(n)$ 位伪随机性；然后，我们再次按照全序 $\leq$，用这些伪随机位填充 τ中的所有关系。显然，这一切都可以通过 $\mathrm{LFP}[\oplus]$ 转导来完成。

为导出矛盾，假设某个 $\mathrm{LFP}[\oplus]$ 对手 φ 能够破坏Θ的安全性。然后考虑算法 A：该算法在输入 $t \in \{0,1\}^{p'(n)}$ 时，使用t 确定每个关系，从而生成一个 τ-结构 $\mathbb{B}$，随后计算 φ，并且当且仅当 $\mathbb{B} \models \varphi$ 时输出 1。显然，当 t 是一个均匀随机字符串时， A 输出 1 的概率与 $\mathbb{B} \models \varphi$ 对一个随机 $\mathbb{B} \sim \mathrm{Str}[\tau, n]$输出 1 的概率相同。另一方面，对于均匀随机的 $s \sim \{0,1\}^n$，将 $t = f(s)$ 作为输入，等价于反复应用引理 2，直至成功对结构进行排序并生成 n 个具有独立随机性的有序比特，然后将 f 应用于所得结果，如图 2 所示。对于一个能够成功排序的随机结构 $\mathbb{A}$， A 随后当且仅当 $\Theta(\mathbb{A}) \models \varphi$ 时输出 1。因此，在 $t = f(s)$ 的情况下， A 输出 1 的概率，恰好等于对于一个随机 $\mathbb{A} \sim \mathrm{Str}[\sigma, n]$， $\Theta(\mathbb{A}) \models \varphi$ 的概率，但以 $\mathbb{A}$已成功排序这一事件为条件；我们将其记为抽样 $\mathbb{A} \sim \mathrm{Str}_{\leq}[\sigma, n]$。根据引理 2，这仅仅

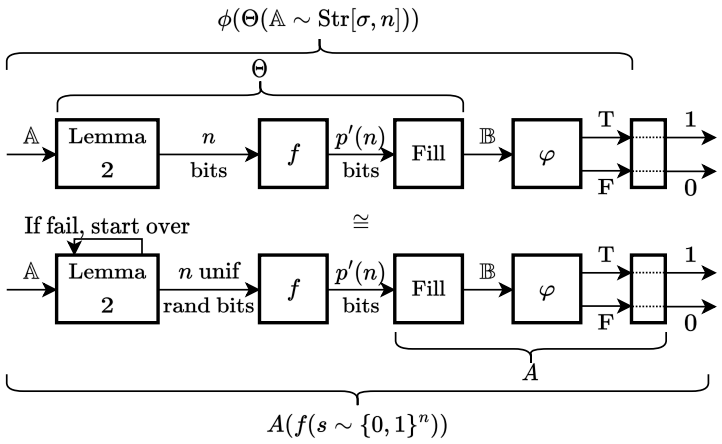


图 2：伴随定理 17 (i)证明的逆向部分的示意图，其中≅ 表示 $\{0,1\}$ 上的输出概率分布仅有可忽略的差异，因为引理2 以可忽略概率失效。

与无条件的 $\Pr_{\mathbb{A}\sim\mathrm{Str}[\sigma,n]}[\Theta(\mathbb{A}) \models \varphi]$ 相差一个可忽略的量。根据三角不等式，我们有

$$\begin{aligned}\left|\Pr_{t\sim\{0,1\}^{p'(n)}}[A(t)=1] - \Pr_{s\sim\{0,1\}^n}[A(f(s))=1]\right| &= \left|\Pr_{\mathbb{B}\sim\mathrm{Str}[\tau,n]}[\mathbb{B}\models\varphi] - \Pr_{\mathbb{A}\sim\mathrm{Str}_{\le}[\sigma,n]}[\Theta(\mathbb{A})\models\varphi]\right| \\ &\ge \left|\Pr_{\mathbb{B}\sim\mathrm{Str}[\tau,n]}[\mathbb{B}\models\varphi] - \Pr_{\mathbb{A}\sim\mathrm{Str}_{\le}[\sigma,n]}[\Theta(\mathbb{A})\models\varphi]\right| \\ &\quad - \left|\Pr_{\mathbb{A}\sim\mathrm{Str}_{\le}[\sigma,n]}[\Theta(\mathbb{A})\models\varphi] - \Pr_{\mathbb{A}\sim\mathrm{Str}[\sigma,n]}[\Theta(\mathbb{A})\models\varphi]\right|,\end{aligned}$$

该概率不可忽略，因为第一项不可忽略，而第二项可忽略。因此， A 破坏了 f的安全性，这与假设矛盾。

对于正向部分，假设 Θ： $\mathrm{Str}[\sigma] \to \mathrm{Str}[\tau]$ 是一个（LFP[⊕]、LFP[⊕]）-伪随机生成器。为了定义一个普通伪随机生成器，需要考虑两种情况。令k 为使得 σ 和 τ 具有不同数量的元数为 k的关系的最大元数，并首先考虑 $k \ge 2$的情况。令 $g(n) = \Omega(n^k)$ 为引理 2中的函数，并令

$$h(n) := \sum_{k'=k}^{\infty} \left|\{i \mid a_i = k'\}\right| \cdot n^{k'}.$$

根据 k 的定义方式以及 $\sigma \prec_L \tau$ 这一事实，我们知道，对于所有 $k' \ge k$， $n^{k'}$ 的系数在 $h(n)$ 和 $p'(n)$ 中相同，而 n^k 系数在 $h(n)$ 中严格更大。因此，当 n 足够大时， $p'(n) \le h(n) - n^k + g(n)$，因为 $g(n) = \Omega(n^k)$ 支配了来自 τ 中元数为 $< k$ 的关系的所有其余次数 $< k$ 项。因此，只需定义一个伪随机生成器 $f\colon \{0,1\}^{p(n)} \to \{0,1\}^{h(n)-n^k+g(n)}$。在输入 $s \in \{0,1\}^{p(n)}$ 时，我们使用所有随机位来生成一个随机的 σ-结构 $\mathbb{A}$。随后运行 Θ，得到一个 τ-结构 $\mathbb{B} = \Theta(\mathbb{A})$，并在该结构上使用引理 2，选取 τ 中任意一个 k-元关系 R。以 $1 - \mathrm{negl}(n)$ 的概率，我们得到论域上的一个全序，并可以从 R 中提取并输出 $g(n)$ 个有序随机比特，同时从元数至少为 k 的其他关系中进一步提取 $h(n) - n^k$ 个有序随机比特。（如果排序步骤失败，我们就输出任意字符串。）

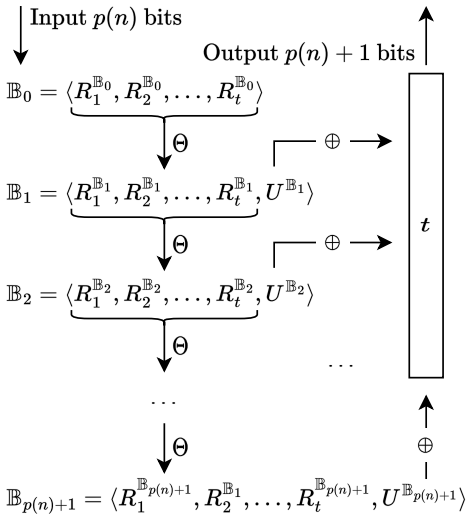


图 3：使用伪随机转导 Θ的普通伪随机生成器f长度扩展构造：$\{0,1\}^{p(n)} \to \{0,1\}^{p(n)+1}$。

假设为了导出矛盾，某个对手 A 能够破坏 f 的安全性。那么我们通过以下算法描述一个 $\mathsf{LFP}[\oplus]$ 对手 φ，以破坏 Θ。给定一个 τ-结构 $\mathbb{B}$，我们对该结构进行排序，并按照 f 的最后一步提取 $h(n) = n^k + g(n)$ 个随机位；当且仅当 A 在所得字符串上输出 1 时，返回 true。显然，根据定理 1，这可以在 $\mathsf{LFP}[\oplus]$ 中实现。如果 $\mathbb{B}$ 是从结构类 $\mathsf{Str}[\tau, n]$ 中随机采样的，那么以 $1 - \mathsf{negl}(n)$ 的概率，输入给 A 的是一个均匀随机的 $p(n)$-比特字符串。另一方面，测试 $\mathbb{B} = \Theta(\mathbb{A}) \models \varphi$ 是否属于 $\mathbb{A} \sim \mathsf{Str}[\sigma, n]$，等价于生成一个随机的 $p(n)$-比特字符串，用它填充 $\mathbb{A}$ 中的关系，应用 Θ，应用引理 2 提取 $h(n) = n^k + g(n)$ 个随机位，最后将所得字符串输入 A。执行这些步骤，恰好等价于在均匀随机的 $p(n)$ 比特字符串上运行 f，然后将结果输入 A。因此，由于 A 破坏了 f 的安全性，根据正向方向中的类似论证可知，φ 破坏了 Θ 的安全性，这与我们的假设矛盾。

最后，考虑 $k = 1$的情形。换言之，σ 和 τ 在每个大于 1 的元数上具有相同数量的关系，但 τ 的元数为 1 的关系严格多于 σ。不失一般性地，假设 σ 和 τ 具有相同的关系符号，但 τ 恰好多出一个一元关系符号 U（如果多出一个以上的一元关系，则忽略多出的部分）。定义 $f\colon \{0,1\}^{p(n)} \to \{0,1\}^{p(n)+1}$ 如下。我们使用输入比特生成一个随机σ-结构 $\mathbb{B}_0$，然后反复应用 Θ $p(n)$ 次，并忽略 U 关系。令 $\mathbb{B}_i$ 表示经过 i 次迭代后得到的中间 τ-结构。随后输出一个字符串 t，其中t 的第 i^{th} 位等于 $\left|U^{\mathbb{B}_i}\right|$ 模 2，即在第 i^{th} 次迭代的输出中，U 成立的元素数量的奇偶性（见图 3）。

为了证明 f是安全的，我们使用混合论证。对于每个 $i \in \{0, 1, 2, 3, \ldots, p(n) + 1\}$，令$H_i$ 表示定义在 $(p(n) + 1)$ 位字符串上的分布，其中前 i 位是均匀随机选取的，而剩余位则从一个随机的 σ-结构开始，按照 f的定义迭代 Θ 来确定。假设为了导出矛盾，存在某个对手 A 能够攻破 f的安全性。也就是说，

$$\left| \Pr_{s \sim \{0,1\}^{p(n)}}[A(f(s)) = 1] - \Pr_{t \sim \{0,1\}^{p(n)+1}}[A(t) = 1] \right|$$

是不可忽略的。由于 H_0 等价于前一种分布，而 $H_{p(n)+1}$ 等价于后一种分布，根据三角不等式可知，对于某个 $i^* \in \{0,1,2,3,\ldots,p\ (n)\ \}$，

$$\left| \Pr_{y \sim H_{i^*}} [A(y) = 1] - \Pr_{y \sim H_{i^*+1}} [A(y) = 1] \right|$$

是不可忽略的（因为任意 $p(n)$ 个可忽略函数之和仍然是可忽略的）。我们将利用这一点定义一个最小不动点逻辑$[\oplus]$ 对手 φ 来攻破 Θ。首先考虑以下随机算法 A'。给定一个 τ结构 $\mathbb{B}$，我们均匀采样 $r \sim \{0,1\}^{i^*}$，然后从 $\mathbb{B}$开始，迭代 Θ、 $p(n) - i^*$ 次，以从 U 关系中提取 $p(n) + 1 - i^*$ 个奇偶校验位，就像 f 的定义中那样（我们也包含 $U^{\mathbb{B}}$ 的初始奇偶性），并将结果追加到 r中。随后，我们在所得的$p(n)+1$ 位字符串上运行 A，并输出其结果。注意，当输入 $\mathbb{B}$ 是一个随机τ结构时，提供给 A 的输入分布恰好是 H_{i^*+1}；而当对于$\mathbb{A} \sim$ 结构类 $\mathsf{Str}[\sigma, n]$有 $\mathbb{B} = \Theta(\mathbb{A})$ 时，提供给 A 的输入分布恰好是 H_{i^*}。因此，我们有

$$\left| \Pr_{\substack{\mathbb{A} \sim \mathrm{Str}[\sigma,n] \\ r \sim \{0,1\}^{i^*}}} [A'(\Theta(\mathbb{A})) = 1] - \Pr_{\substack{\mathbb{B} \sim \mathrm{Str}[\tau,n] \\ r \sim \{0,1\}^{i^*}}} [A'(\mathbb{B}) = 1] \right|$$

并非可忽略。剩下的工作是将 A' 转换为一个 $\mathsf{LFP}[\oplus]$ 句子 φ，这样便会与我们关于 Θ 是安全的假设相矛盾。主要困难在于， A' 是一个随机算法。虽然 $\mathsf{LFP}[\oplus]$-转导无法为 r 生成随机位，但我们可以将一个良好的实现 r^* 编码到建议结构中并使用它。所谓良好的实现，是指满足以下条件的样本 $r \in \{0,1\}^{i^*}$：

$$\begin{aligned} &\left| \Pr_{\mathbb{A} \sim \mathrm{Str}[\sigma,n]} [A'(\Theta(\mathbb{A}), r^*) = 1] - \Pr_{\mathbb{B} \sim \mathrm{Str}[\tau,n]} [A'(\mathbb{B}, r^*) = 1] \right| \\ &\geq \left| \Pr_{\substack{\mathbb{A} \sim \mathrm{Str}[\sigma,n] \\ r \sim \{0,1\}^{i^*}}} [A'(\Theta(\mathbb{A})) = 1] - \Pr_{\substack{\mathbb{B} \sim \mathrm{Str}[\tau,n] \\ r \sim \{0,1\}^{i^*}}} [A'(\mathbb{B}) = 1] \right|, \end{aligned}$$

根据平均原理，这样的样本必然存在。我们还将一个全序编码到建议结构中，以便使用归纳定义的关系存储每次迭代的输出 Θ。因此，我们可以将 A' 转换为一个 $\mathsf{LFP}[\oplus]$句子 φ，从而完成情形 (i)的证明。

(ii)：如果 $(a'_1, a'_2, \ldots, a'_{t'})$ 是 $(a_1, a_2, \ldots, a_t)$ 的一个排列，那么显然可以通过按照该排列将关系彼此映射，得到一个 $(\mathsf{LFP}[\oplus], \mathsf{LFP}[\oplus])$-伪随机生成器。否则，设 k 是满足以下条件的最大元数： σ 和 τ 中元数为k的关系数量不同。如果 $k = 1$，那么同样很容易构造一个从 σ 到 τ 的伪随机生成器：只需按元数匹配关系，并忽略 σ 中多出的一个或多个一元关系。如果 $k \geq 2$，那么较低元数部分可能无法匹配（也就是说，当 $\sigma \succeq_L \tau$ 但 $\sigma \not\succeq_S \tau$ 时）。此时，我们再次援引引理 2，使用一个 k元关系以概率 $1 - \mathsf{negl}(n)$ 对论域进行排序，然后依次用 Q 定义的 $\Omega(n^k)$ 个随机位填充较低元数的关系。由于这以概率 $1 - \mathsf{negl}(n)$ 产生一个均匀随机的 τ结构，所以所得分布显然对于任意逻辑都是伪随机的。

(iii)：在此情况下， σ 仅包含一元关系，而 $\sigma \prec_L \tau$，这意味着 $\sigma \not\succeq_S \tau$ 因为 $\succeq_S$ 是 $\succeq_L$的细化。根据推论16，甚至不存在从 σ到 τ的（$\mathsf{LFP}[\oplus]$，FO）伪随机生成器，因此显然也不可能存在（$\mathsf{LFP}[\oplus]$，$\mathrm{LFP}[\oplus]$）伪随机生成器。

□

5 带奇偶性的一阶逻辑

转导 $\theta^t(x_1, \ldots\ldots, x_t) := \oplus y \bigwedge_{i=1}^{t} E(y, x_i)$ 要判断是否存在奇数个顶点 y，使其与来自 $x_1, \ldots\ldots, x_t$ 的所有顶点都相邻。它从一个普通图生成一个 t-超图。我们通过证明以下扩张公理以高概率均成立，论证 FO 和最小不动点逻辑无法区分 $\theta^t(G(n, 1/2))$ 与真正的随机t-超图 $G_t(n, 1/2)$。

定义 18。 t-超图满足扩张公理 EA_k^t，当且仅当对于所有大小为 k 的集合 S 以及所有 $T \subseteq \binom{S}{t-1}$，存在 $v \notin S$，使得对于每个 $\{s_1, \ldots, s_{t-1}\} \in \binom{S}{t-1}$，当且仅当 $\{s_1, \ldots, s_{t-1}\} \in T$ 时，我们都有一条超边 $\{s_1, \ldots, s_{t-1}, v\}$。

引理11 指出，即使对于超图，相应的扩张公理也决定了 LFP 句子的真值，即使存在一个建议结构；由引理 10可知，随机超图以高概率满足扩张公理：对于每个 k 和 t,

$$\Pr_{G \sim G_t(1/2,n)}\left[G \not\models \mathrm{EA}_k^t\right] = \mathrm{negl}(n).$$

作为本节的主要结果，对 $G(n, 1/2)$ 中某些集合的奇偶性的概率分析将揭示，在经过转导的图 $\theta^t(G(n, 1/2))$ 中同样成立。

引理 19。 对于每个 k 和 t,

$$\Pr_{G \sim G(n,1/2)}\left[\theta^t(G) \not\models \mathrm{EA}_k^t\right] = \mathrm{negl}(n).$$

我们在下一小节中证明这个引理。于是，下面的定理立即由引理 10、11 和 19 得出。

定理 20。 对于每个 t, θ^t 是一个从图到 t-超图的 （一阶逻辑$[\oplus]$，最小不动点逻辑）伪随机生成器。

如果能找到一种逻辑，使其允许构造能够抵抗同一逻辑中的对手的伪随机生成器，那将非常有趣。我们认为 FO$[\oplus]$ 是这一目标的候选逻辑。目前的证明仅表明其能够抵抗 FO 或最小不动点逻辑对手，因为扩展公理不足以确定 FO$[\oplus]$ 句子的真值。然而，Kolaitis 和 Kopparty [22] 以类似的方式对 FO$[\oplus]$ 给出了一个强有力的结果：要确定任意 FO$[\oplus]$句子的真值，只需知道某些“子图频率”，也就是在某个大小范围内若干指定子图出现次数的奇偶性。如果这一结论能够推广到超图，那么，为了欺骗 FO$[\oplus]$，只需转导出一个能够模拟这些子图频率的 t-超图。因此，可以将 [22] 理解为以下猜想的证据。

猜想 21。 对于每个 t, θ^t 都是从图到 t-超图的 （一阶逻辑$[\oplus]$，一阶逻辑$[\oplus]$）-伪随机生成器。

5.1 随机图中的奇偶性概率

还需证明，以高概率， $\theta^t(G(n, 1/2))$满足扩展公理。粗略地说，在下面的引理中， $|C| = t - 1$ 对应于 θ^t 的输入变量，也就是说，“与 $\{v\} \cup C$ 中每个顶点都相邻的顶点数量”的奇偶性将决定 $\{v\} \cup C$ 是否构成超边。

引理 22。 设 $k \in \mathbb{Z}_{\geq 1}$。在 $G(n,1/2)$ 中，以 $1-\mathrm{negl}(n)$ 的概率，对于每个大小为 k 的顶点集 $S \subseteq [n]$ 以及每个 $\mathcal{C} \subseteq \{C \mid \emptyset \neq C \subseteq S\}$，都存在 $v \notin S$，使得对于所有 $\emptyset \neq C \subseteq S$，$C \in \mathcal{C}$ 当且仅当与 $\{v\} \cup C$ 中每个顶点都相邻的顶点数量为奇数。

我们通过以下类似命题，借助合适向量空间上的精细基构造，推导出引理 22 ；我们先证明该命题。

引理 23. 设 $k \in \mathbb{Z}_{\geq 1}$。在 $G(n,1/2)$ 中，以 $1-\mathrm{negl}(n)$ 的概率，对于每个大小为 k 的顶点集 $S \subseteq [n]$ 以及每个 $\mathcal{B} \subseteq \{B \mid \emptyset \neq B \subseteq S\}$，都存在 $v \notin S$，使得对于所有 $\emptyset \neq B \subseteq S$，当且仅当与 $\{v\} \cup B$ 中每个顶点都相邻、且与 $S \setminus B$ 中每个顶点都不相邻的顶点数量为奇数时，$B \in \mathcal{B}$。

证明。 $G(n,1/2)$ 的分布与以下随机过程相同：按照任意顺序逐一揭示元组 $\{u,v\} \in \binom{[n]}{2}$ 之间的所有邻接关系；每个元组以概率 $1/2$ 被揭示为相邻或不相邻。

令 S 为一个由 k 个顶点组成的集合，并且 $\mathcal{B} \subseteq \{B \mid \emptyset \neq B \subseteq S\}$。我们首先揭示所有与 S 相 incident 的邻接关系。揭示出的边根据顶点与 S 的连接方式，将图的顶点划分为 2^k 个集合。对于 $B \subseteq S$，我们用 $V_B \subseteq [n]$ 表示在 S 中邻域恰好为 B 的顶点。

令 $B \subseteq S$。由于每条边都以概率 $1/2$ 独立出现，位于 S 外的一个顶点属于 V_B 的概率为 2^{-k}。由于我们只需确定错误概率的渐近行为，因此可以始终假设 n 相较于 k 足够大。我们可以假设 $k \leq n/2$，这意味着 $n-k \geq n/2$，因此该集合中的顶点期望数量为

$$E\big[|V_B \setminus S|\big] = (n-k)\cdot 2^{-k} \geq n/2^{k+1}.$$

位于 S 外的顶点属于 V_B 的事件彼此独立。Chernoff 界指出，独立二元随机变量之和与其期望值显著偏离的概率至多为指数级小。应用 Chernoff 界可得

$$\begin{aligned}\Pr\Big[|V_B \setminus S| \leq n/2^{k+2}\Big] &\leq \Pr\Big[|V_B \setminus S| \leq E[|V_B \setminus S|]/2\Big] \\ &\leq e^{-E[|V_B \setminus S|]/12} \\ &\leq e^{-n/2^{k+5}}.\end{aligned}$$

根据并集界，$E\big[|V_B \setminus S|\big] \leq n/2^{k+3}$ 对任意 $B \subseteq A$ 成立的概率至多为

$$2^k \cdot e^{-n/2^{k+5}}. \tag{1}$$

因此，我们暂且假设，以某种方式揭示了与 S 关联的边，使得对于所有 $B \subseteq S$，$|V_B \setminus S| \geq n/2^{k+2}$。当 n 相对于 k 足够大时，这意味着 $V_B \setminus S$ 非空。这样，我们就可以为每个 $\emptyset \neq B \subseteq S$ 选取一个“切换”顶点 $v_B \in V_B \setminus S$。

考虑一个顶点 $v \in V_\emptyset \setminus S$。我们已经揭示了 v 与 S之间的边。接下来揭示 v 与 $V \setminus \big(V_\emptyset \cup \{v_B \mid B \subseteq S\}\big)$之间剩余的边。我们定义一个“奇偶模式”$\mathcal{B}(v) \subseteq \{B \mid \emptyset \neq B \subseteq S\}$，其中 $B \in \mathcal{B}(v)$ 当且仅当 v 与 V_B中的奇数个顶点相邻。随着我们逐一揭示v 与满足 $\emptyset \neq B \subseteq S$的顶点 v_B 之间的边，奇偶模式 $\mathcal{B}(v)$ 会随时间变化。当揭示通向 v_B 的边时，$B \in \mathcal{B}(v)$ 是否成立这一事实会以 $1/2$的概率翻转，而奇偶模式的所有其他值保持不变。由于所有边彼此独立，这保证了之后，对于 $B \neq \emptyset$，$B \in \mathcal{B}(v)$ 是否成立这些事件彼此独立，且发生的概率为 $1/2$。由于 v 与 $V \setminus V_\emptyset$ 之间的所有边都已揭示，之后揭示更多边时，$\mathcal{B}(v)$ 将不再变化。

事件 $\mathcal{B}(v) = \mathcal{B}$ 发生的概率恰好为 2^{-2^k-1}，这等价于我们的目标：对于所有 $\emptyset \neq B \subseteq S$，$B \in \mathcal{B}$ 当且仅当与 $\{v\} \cup B$ 中每个顶点相邻、且与 $S \setminus B$ 中每个顶点都不相邻的顶点数为奇数。对于所有其他 $v \in V_\emptyset \setminus S$，我们也以同样的方式继续揭示边并分析奇偶模式 $\mathcal{B}(v)$。这 $|V_\emptyset \setminus S|$ 个事件 $\mathcal{B}(v) = \mathcal{B}$ 彼此独立。因此，没有顶点 $v \in V_\emptyset$ 呈现奇偶模式 $\mathcal{B}(v) = \mathcal{B}$ 的概率等于（第一个不等式使用观察24，第二个不等式使用 $|V_\emptyset \setminus S| \geq n/2^{k+2}$）

$$\left(1 - 2^{-2^k+1}\right)^{|V_\emptyset \setminus S|} \leq e^{-|V_\emptyset \setminus S|/2^{2^k-1}} \leq e^{-n/\left(2^{k+2} \cdot 2^{2^k-1}\right)} = e^{n/2^{2^k+k+1}}. \tag{2}$$

对于任意给定的 S 和 $\mathcal{B}$，如果存在一个顶点实现该奇偶模式，则所声称的陈述成立。对于 S 和 $\mathcal{B}$，至多有 $n^k \cdot 2^{2^k}$ 种选择；对于每种选择，错误概率均由（1）和（2）之和界定。因此，根据并集界，总错误概率至多为

$$\left(n^k \cdot 2^{2^k}\right) \cdot \left(2^k \cdot e^{-n/2^{k+5}} + e^{-n/2^{2^k+k+1}}\right).$$

由于 k 是固定的，这是关于 n 的可忽略函数。 □

现在我们继续证明引理22 和19，从而完成定理20的证明。

引理*22*的证明。固定 $S \subseteq [n]$。对于每个 $v \in [n]$，令 $\mathcal{B}(v)$ 为 $\{B \mid \emptyset \neq B \subseteq S\}$的子集，使得对于所有 $\emptyset \neq B \subseteq S$，当且仅当满足 $B \in \mathcal{B}(v)$ 时，与 $\{v\} \cup B$ 中的每个顶点都相邻且与 $S \setminus B$ 中的每个顶点都不相邻的顶点数为奇数。类似地，令 $\mathcal{C}(v)$ 为 $\{C \mid \emptyset \neq C \subseteq S\}$ 的子集，使得对于所有 $\emptyset \neq C \subseteq S$，当且仅当满足 $C \in \mathcal{C}(v)$ 时，与 $\{v\} \cup C$ 中的每个顶点都相邻的顶点数为奇数。

对于某个顶点 v 和集合 C，令 X 表示与 $\{v\} \cup C$ 中每个顶点都相邻的顶点。我们可以根据顶点与 $S \setminus C$ 的邻接关系，对 X 中的顶点进行划分。现在，当且仅当我们将 X 划分为奇数个大小为奇数的部分时，X 的大小才为奇数。换言之，

$$C \in \mathcal{C}(v) \iff \left\{B \mid C \subseteq B \subseteq S, B \in \mathcal{B}(v)\right\} \text{ has odd size.} \tag{3}$$

同样可以证明，对于顶点 v 和集合 B，有

$$\mathcal{B}(v) = \{B\} \implies \mathcal{C}(v) = \{C \mid B \subseteq C \subseteq S\}. \tag{4}$$

接下来，我们考虑带有对称差运算 $\oplus$ 的向量空间 $2^{\{B|\emptyset \neq B \subseteq S\}}$。例如，对于该向量空间中的向量 $\{B_1, B_2\}$ 和 $\{B_2, B_3\}$（其中 $\emptyset \neq B_1, B_2, B_3 \subseteq S$），有 $\{B_1, B_2\} \oplus \{B_2, B_3\} = \{B_1, B_3\}$。我们断言，这些集合

$$\left\{\{B \mid C \subseteq B \subseteq S\} \mid \emptyset \neq C \subseteq S\right\}$$

构成该向量空间的一组基。为证明这一点，我们按 i 递减的顺序迭代构造所有满足 $\emptyset \subseteq I \subseteq S$，$|I| = i$ 的单位向量 $\{I\}$，使其成为我们所述基的线性组合。对于 $i > |S|$，无需证明。为了生成一个满足 $|I| = i - 1$ 的单位向量 $\{I\}$，我们从向量$\{B \mid I \subseteq B \subseteq S\}$ 开始，并使用 $\oplus$ 运算以及已经构造出的向量 $\{B\}$，消去大小至少为 i 的所有多余项 $I \subsetneq B \subseteq S$。

接下来，我们将论证，对于所有的 v、v_1、$v_2 \in [n]$，

$$\mathcal{B}(v) = \mathcal{B}(v_1) \oplus \mathcal{B}(v_2) \implies \mathcal{C}(v) = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2). \tag{5}$$

假设 $\mathcal{B}(v) = \mathcal{B}(v_1) \oplus \mathcal{B}(v_2)$。对于集合 C，我们注意到

$$\begin{aligned} C \in C(v) &\iff \{B \mid C \subseteq B \subseteq A, B \in (\mathcal{B}(v_1) \oplus \mathcal{B}(v_2))\} \text{ has odd size} \qquad \text{(from (3))} \\ &\iff \{B \mid C \subseteq B \subseteq A, B \in \mathcal{B}(v_1)\} \text{ and } \{B \mid C \subseteq B \subseteq A, B \in \mathcal{B}(v_2)\} \\ &\qquad \text{have sizes of different parity} \\ &\iff (C \in C(v_1) \text{ and } C \notin C(v_2)) \text{ or } (C \notin C(v_1) \text{ and } C \in C(v_2)). \end{aligned}$$

这证明了（5）。

根据引理23，我们可以假设，对于每个

$$\mathcal{B} \subseteq \{B \mid \emptyset \neq B \subseteq S\},$$

存在 $v \in [n]$，使得 $\mathcal{B}(v) = \mathcal{B}$。特别地，对于每个 $\emptyset \neq B \subseteq S$，存在 $v \in [n]$，使得 $\mathcal{B}(v) = \{B\}$；根据（4），这意味着 $\mathcal{C}(v) = \{C \mid B \subseteq C \subseteq S\}$。因此，根据上面的断言，$\{\mathcal{C}(v) \mid v \in [n]\}$ 是我们向量空间的一个基。为了最终证明这一引理的断言，任取一个 $\mathcal{C} \subseteq \{C \mid \emptyset \neq C \subseteq S\}$。由于 $\{\mathcal{C}(v) \mid v \in [n]\}$ 是我们向量空间的一个基，我们可以将 $\mathcal{C}$ 表示为 $\mathcal{C} = \mathcal{C}(v_1) \oplus \ldots \oplus \mathcal{C}(v_t)$，其中顶点为 v_1、……、$v_t \in [n]$。本段开头对引理 23 的调用意味着存在一个顶点 v，使得 $\mathcal{B}(v) = \mathcal{B}(v_1) \oplus \ldots \oplus X(v_t)$。反复应用（5）即可推出 $\mathcal{C}(v) = \mathcal{C}(v_1) \oplus \ldots \oplus C(v_t)$。因此，$\mathcal{C}(v) = \mathcal{C}$。 □

引理 19 的证明。根据引理 22，以 n 为参数，错误概率可忽略：对于每个大小为 k 的顶点集 $S \subseteq [n]$ 以及每个 $\mathcal{C} \subseteq \{C \mid \emptyset \neq C \subseteq S\}$，存在 $v \notin S$，使得对于所有 $\emptyset \neq C \subseteq S$，$C \in \mathcal{C}$ 当且仅当与 $\{v\} \cup C$ 中每个顶点都相邻的顶点数为奇数。因此，特别地，对于每个 $T \subseteq \binom{S}{t-1}$，存在 $v \notin S$，使得对于每个 $\{s_1, \ldots, s_{t-1}\} \in \binom{S}{t-1}$，$\{s_1, \ldots, s_{t-1}\} \in T$ 当且仅当与 $\{s_1, \ldots, s_{t-1}, v\}$ 中每个顶点都相邻的顶点数为奇数。后者等价于 $\{s_1, \ldots, s_{t-1}, v\}$ 是 $\theta^t(G)$ 中的一条边。因此，以可忽略的错误概率，$G \models \mathrm{EA}^t_k$。 □

6 结论

本文定义了伪随机生成器的模型论类比。我们全面分类了哪些由一阶逻辑、最小不动点逻辑以及最小不动点逻辑[⊕] 生成的伪随机生成器与对手逻辑能够存在，并确定性地构造出能够欺骗一阶逻辑和最小不动点逻辑对手的结构，即使在任意关系签名的范围内也是如此。最后，我们提出了一个一阶逻辑[⊕]转导的潜在候选，它可充当一个能够抵御一阶逻辑[⊕]自身的伪随机生成器。从概念层面而言，这项工作为伪随机性提供了新颖的视角，并为密码学领域的描述复杂性开辟了新的方向；在这一领域，平均情形分析至关重要。

致谢

我们非常感谢 Boaz Barak 对早期草稿提出的宝贵建议和反馈。

本材料部分基于美国国家科学基金会研究生研究奖学金计划在资助编号 DGE1745303 项下提供支持的工作。本文所表达的任何意见、研究结果以及结论或建议均属于作者观点，不一定反映美国国家科学基金会的立场。

参考文献

[1] Wilhelm Ackermann。一般集合论的一致性。数学年刊，114(1)：305–315，1937。[2] Sanjeev Arora 和 Boaz Barak。计算复杂性：一种现代方法。剑桥大学出版社，2009。[3] Albert Atserias 和 Anuj Dawar。可定义的不可近似性：给 Duplicator 带来的新挑战。*J. Log. Comput.*，29(8)：1185–1210，2019。[4] Andreas Blass、Geoffrey Exoo 和 Frank Harary。佩利图满足所有一阶邻接公理。图论杂志，5(4)：435–439，1981。[5] Andreas Blass、Yuri Gurevich 和 Dexter Kozen。带不动点算子的逻辑的零一律。*Inf. Control.*，67(1-3)：70–90，1985。[6] Andreas Blass 和 Benjamin Rossman。具有扩展性质的显式图。*Bull.EATCS*，86：166–175，2005。[7] Anthony Bonato。对 n-扩展图的探索。*Contributions to Discrete Mathematics*，4(1)，2009。[8] Dan Boneh 和 Victor Shoup。应用密码学研究生课程。2015。[9] Peter J. Cameron 和 Dudley Stark。具有 n-扩展性质的强正则图的丰富构造。*Electron. J. Comb.*，9(1)，2002。[10] Bruno Courcelle 和 Joost Engelfriet。图结构与单子二阶逻辑：一种语言理论方法，第 138 卷。剑桥大学出版社，2012。[11] Heinz-Dieter Ebbinghaus 和 J¨org Flum。有限模型论。施普林格·柏林·海德堡，1995。[12] Chaim Even-Zohar、Michael Farber 和 Lewis Mead。充足单纯复形。欧洲数学杂志，8(1)：1–32，2022。[13] Ronald Fagin。广义一阶谱与多项式时间可识别集合。计算复杂性，7：43–73，1974。[14] Ronald Fagin。有限模型上的概率。*J. Symb. Log.*，41(1)：50–58，1976。[15] Ronald L. Graham、Donald E. Knuth 和 Oren Patashnik。具体数学：计算机科学的基础。艾迪生-韦斯利，1989。[16] Martin Grohe。寻找一种刻画 PTIME 的逻辑。载于第二十三届 *IEEE* 计算机科学逻辑年度研讨会论文集，*LICS 2008，2008* 年 *6* 月 *24–27* 日，美国宾夕法尼亚州匹兹堡，第 267–271 页。IEEE 计算机学会，2008。[17] Lauri Hella、Phokion G. Kolaitis 和 Kerkko Luosto。有限模型论中逻辑的几乎处处等价。*Bull. Symb. Log.*，2(4)：422–443，1996。[18] Johan H ˚Astad、Russell Impagliazzo、Leonid A. Levin 和 Michael Luby。由任意单向函数构造伪随机生成器。*SIAM Journal on Computing*，28(4)：1364–1396，1999。

[19] Neil Immerman。在多项式时间内可计算的关系查询。信息与控制，68(1-3):86– 104，1986。[20] Neil Immerman。描述复杂性。计算机科学研究生教材。施普林格，1999。[21] Richard M. Karp。图的典范编号算法的概率分析。载于组合学与数学其他领域之间的关系（*Ohio State Univ.*，哥伦布，俄亥俄州，*1978* 年纯数学研讨会论文集），《纯数学研讨会论文集》，XXXIV，365–378 页。美国数学学会，普罗维登斯，罗得岛州，1979。[22] Phokion G. Kolaitis 和 Swastik Kopparty。随机图与奇偶量词。*J.ACM*，60(5)，2013 年 10 月。[23] Michael Krivelevich 和 Benny Sudakov。伪随机图。载于更多集合、图与数，199–262 页。施普林格，2006。[24] Leonid Libkin。有限模型论要素。《理论计算机科学文本》。EATCS 丛书。施普林格，2004。[25] Leonid Libkin。有限模型论要素。施普林格，2004。[26] Moni Naor、Leonard J Schulman 和 Aravind Srinivasan。分割器与近最优去随机化。载于*IEEE* 第 *36* 届年度计算机科学基础会议论文集，182–191 页。IEEE，1995。[27] R. Rado。通用图与通用函数。*Acta Arithmetica*，9(4):331–340，1964。[28] Jamie Tucker-Foltz。带计数的定点逻辑中唯一博弈的不可近似性。载于第 *36* 届年度 *ACM/IEEE* 计算机科学逻辑研讨会，*LICS 2021*，意大利罗马，*2021* 年 *6* 月 *29* 日至 *7* 月 *2* 日，1–13 页。IEEE，2021。[29] Moshe Y. Vardi。关系查询语言的复杂性（扩展摘要）。载于第 *14* 届年度 *ACM* 计算理论研讨会论文集，*1982* 年 *5* 月 *5* 日至 *7* 日，美国加利福尼亚州旧金山，137–146 页，1982。

A 错误概率

为完整起见，我们回顾一个有助于界定错误概率的简单事实。

观察24. 由于 $1-x \leq e^x$ 对所有 x都成立，因此我们可以得出，对所有 a, b，

$$\left(1-\frac{1}{a}\right)^b \leq e^{-b/a}.$$

对于 $a=2^k$ 和 $b=n$，例如，我们可以对在 k 次抛硬币中未能得到 k 次正面的概率作出如下界定，其中实验允许重复 n 次:

$$(1-2^{-k})^n \leq e^{-n/2^k},$$

当 n 相较于 k足够大时，这个值会越来越小。

B 引理 2 的证明2

引理 2. 设 σ 是一个关系签名，其中包含一个元数为 $k \geq 2$的关系符号 R；设 τ是一个关系签名，其中包含元数为 *2* 的符号“$\leq$”以及元数为 k的符号 Q。存在一个 最小不动点逻辑$[\oplus]$转导 Γ : 结构类$[\sigma] \to$ 结构类$[\tau]$，使得对于随机 σ-结构$\mathbb{A}$，其概率为 $1-\operatorname{negl}(n)$，

(i) $\leq^{\Gamma(\mathbb{A})}$ 在论域上定义一个全序，并且

(ii) $\left|Q^{\Gamma(\mathbb{A})}\right| = \Omega(n^k)$.

此外，始终有如下情况：对于 k-元组 $(x_1, x_2, \ldots, x_k) \in Q^{\Gamma(\mathbb{A})}$，即使在以 $\leq^{\Gamma(\mathbb{A})}$ 和 $Q^{\Gamma(\mathbb{A})}$的任意实现为条件化之后，$(x_1, x_2, \ldots, x_k) \in R^{\mathbb{A}}$ 的概率也都为 $1\,2$，并且在 $(x_1, x_2, \ldots, x_k) \in Q^{\Gamma(\mathbb{A})}$之间两两独立。

证明。考虑以下算法。令 S 为论域中使 $R(x, x, \ldots, x)$ 成立的元素 x 的集合，并令 S为其补集。根据 Chernoff 界，以概率1 –negl(n)， S 和S 各自都至少包含 $\frac{n}{3}$ 个元素。我们在 S 上定义图 G，并在 S 上定义图G：当且仅当 $R(x, y, y, \ldots, y)$ 和 $R(y, x, x, \ldots, x)$ 中恰有一个成立时，相应顶点集中的两个不同顶点 x 和 y 之间才有一条边。注意，这两个图始终都是均匀随机图。随后，我们将 Karp 的算法 [21] 应用于 G 和 G，以计算顶点集 S 和S上的规范全序。由于 S 是可定义的，我们可以将这两个全序串接起来，使 S 中的所有元素都排在 S中的所有元素之前。Hella、Kolaitis 和 Luosto [17] 证明了 Karp算法可以实现为一个最小不动点逻辑$[\oplus]$ 公式，而其他所有步骤显然都可以在一阶逻辑中完成。这就给出了一个最小不动点逻辑$[\oplus]$ 定义来定义 $\leq$。随后，我们定义关系 Q 在所有元组 $(x, y_1, y_2, \ldots, y_{k-1})$ 上成立，其中 $x \in S$ 且每个 $y_i \in S$。

在随机图上运行时，Karp算法未能找到规范顺序的概率是顶点数的可忽略函数。因此，假设每个顶点集的大小至少为 $\frac{n}{3}$，则两个图中至少有一个未能成功排序的概率至多为 negl$(\frac{n}{3})$ + negl$(\frac{n}{3})$ =negl(n)。此外，同样假设每个顶点集的大小至少为 $\frac{n}{3}$， Q将至少包含

$$\left(\frac{n}{3}\right) \cdot \left(\frac{n}{3}\right)^{k-1} = \frac{1}{3^k} \cdot n^k = \Omega(n^k)$$

k元组。最后，独立性条件源于这样一个事实：规范化算法只考虑图 G 和G，因此从不检查 Q中的任何元组是否属于 R。

□

C 关系拉多结构

我们将扩展公理的定义以及有限拉多图的确定性构造推广到关系结构。

定义 25。 固定一个签名 σ。我们称一个 k-原子类型 是由元组 $(R,(i_1,\ldots,i_a))$ 组成的集合，其中$R \in \sigma$ 是元数为 a 的关系符号，而 $i_1,\ldots,i_a \in \{0,\ldots,k\}$ 是满足 $0 \in \{i_1,\ldots,i_a\}$的指标。我们用 types(σ,k) 表示所有可能的 k-原子类型的集合。如果 σ 中的关系的元数为$a_1,\ldots,a_t$，则 $|\text{types}(\sigma,k)| \leq 2^{\sum_{i=1}^{t}(k+1)^{a_i}}$。

在一个 σ-结构 $\mathbb{A}$中，我们称顶点 $v \in \mathbb{A}$ 满足 元组 $(v_1,\ldots,v_k)$ 上给定的 k-原子类型 $\in \mathbb{A}^k$，如果对于上述所有元组 $(R,(i_1,\ldots,i_a))$，都有 $(v_{i_1},\ldots,v_{i_a}) \in R^{\mathbb{A}}$ 当且仅当 $R(i_1,\ldots,i_a)$ 包含于该原子类型中。

我们称 $\mathbb{A}$ 满足扩张公理 EA_s^σ，如果对于所有两两不同的 $v_1,\ldots,v_k \in \mathbb{A}$ 以及所有k-原子类型，都存在 $v_0 \neq v_1,\ldots,v_k$，使其在 $(v_1,\ldots,v_k)$ 上满足给定类型。

有限 Rado 结构的构造与定理 8的思路相同，但可能的原子类型数量更多，使得构造过程更加晦涩。结合引理 10 和引理11，下述定理蕴含（见定理 12），最小不动点逻辑无法区分随机 σ结构与下述定理的输出。

定理 9。 设 σ 是一个关系签名。存在一个常数 c，使得对每个 n，都可以在时间 $O(n^c)$ 内构造一个 n元素的 σ结构，该结构满足扩张公理EA_k^σ，其中所有$k \leq \lfloor \log(\log(n))/c \rfloor$均成立。

证明。 我们改为证明以下断言：存在常数 c，使得对于每个 n，$k \in \mathbb{Z}_{\geq 1}$ 满足$2^{k^c} \leq n$，都可以在时间 $O(2^{2^{ck}}n^c)$ 内构造一个含 n 个元素的 σ-结构，使其满足扩展公理 EA_k^σ。由此即可推出定理的结论。

我们再次使用引理 7 和穷举搜索，找出一个锦标赛 F，其顶点集为$[2^{3k}]$，使得对于每个大小为 $S \subseteq [2^{3k}]$ 的 k，都存在一个顶点 $j \in [2^{3k}]$，它向 S中的每个顶点都有有向边。

我们将构造一个论域为 $[n]$ 的结构 $\mathbb{A}$，使其满足扩展公理 EA_k^σ。首先，利用定理 6 构造一个 (n,k)-完美哈希函数族 $\mathcal{F}$。接下来，我们需要将论域 $[n]$ 划分为 2^{3k} 个部分 P_1、……、$P_{2^{3k}}$，使每个部分的大小至少为 $|\mathcal{F}| \cdot k! \cdot |\text{类型}(\sigma,k)|$。下面说明可以选择参数 c，使得对于所有 $n \geq 2^{k^c}$，这都是可行的。设 c' 为定理 6 中的常数，这样我们可以将 (n,k)-完美哈希函数族的大小上界设为 $2^{c'k}\log(n)$。设元数属于 σ 的关系为 a_1、……、a_t。那么，k-原子类型的数量至多为 $|\text{类型}(\sigma,k)| \leq 2^{\sum_{i=1}^{t}(k+1)^{a_i}}$。我们可以选择 c，使得对于每个$k \in \mathbb{Z}_{\geq 1}$,

$$2^{3k} \cdot k! \cdot 2^{c'k} \cdot |\text{types}(\sigma,k)| \cdot 2 \leq 2^{k^c}/k^c.$$

因此，对于每个 $n \in \mathbb{Z}_{\geq 1}$，当 $2^{k^c} \leq n$ 成立时，$2^{k^c}/k^c \leq n/\log(n)$，从而

$$2^{3k} \cdot k! \cdot 2^{c'k} \log(n) \cdot |\text{types}(\sigma,k)| \cdot 2 \leq n.$$

This 这意味着，我们可以将论域划分为若干部分 P_1、……、$P_{2^{3k}}$，每部分的大小至少为st

$$\lfloor 2^{c'k} k! \cdot \log(n) \cdot |\text{types}(\sigma,k)| \cdot 2 \rfloor \geq k! \cdot |\mathcal{F}| \cdot |\text{types}(\sigma,k)|.$$

为方便起见，我们定义一个扩展函数集 $\mathcal{F}^*$，其中对每个函数$f \in \mathcal{F}$和每个排列 $\pi\colon [k] \to [k]$，都包含函数 $\pi \circ f$。注意，$|\mathcal{F}^*| \leq |\mathcal{F}| \cdot k!$ 此外，对于所有互异元素 v_1，……，v_k，存在 $f \in \mathcal{F}^*$，使得对所有 $l \in [k]$，都有 $f(v_l) = l$。

我们将各部分 P_1，……，$P_{2^{3k}}$设得足够大，以便可以选取一个函数 $\text{PATTERN}: [n] \to \mathcal{F}^* \times$类型$(\sigma, k)$，使得对每个 $j \in [2^{3k}]$，函数 PATTERN 限制在 P_j 上时都是满射。我们进一步用 index(v) 表示满足 $v \in P_j$ 的索引 j。为了构造我们的结构$\mathbb{A}$，现在我们对每个 $R \in \sigma$ 或元数为 a 且由（不一定互异的）元素组成的序列 u_1，……，$u_a \in \mathbb{A}$ 执行以下操作：

> 令 $v_0 \in \{u_1, \ldots, u_a\}$ 为唯一元素，使得在我们的锦标赛 F 中，从索引(v_0) 到 { 中所有顶点都存在有向边：索引$(u_1) \ldots,$索引$(u_a)\} \setminus \{$索引$(v_0)\}$，并令 $\text{PATTERN}(v_0) = (f, \tau)$。当且仅当存在 $(R, (i_1, \ldots, i_a)) \in \tau$ 满足以下条件时，将元组 $(u_1, \ldots, u_a)$ 加入 $R^{\mathbb{A}}$：
>
> $$\begin{cases} u_l = v_0 & \text{for all } l \text{ with } i_l = 0, \\ f(u_l) = i_l & \text{for all } l \text{ with } i_l \neq 0. \end{cases}$$

正确性 令 $v_1, \ldots, v_k \in \mathbb{A}$ 两两不同，并令 τ 为一个 k-原子类型。我们使用引理 7 构造了锦标赛 F，使得可以选取 $j \in [2^{3k}]$，令其到索引$(v_1) \ldots,$索引(v_k) 中所有顶点都有一条有向边。集合 $\mathcal{F}^*$ 的构造保证存在 $f \in \mathcal{F}^*$，使得对所有 l，$f(v_l) = l$。由于将 PATTERN 限制在 P_j 上时选为满射，我们也可以选取 $v_0 \in P_j$，使得 $\text{PATTERN}(v_0) = (f, \tau)$。由于 F 没有自环，$v_0 \neq v_1, \ldots, v_k$。

现在考虑一个元组 $(R,$ $(i_1,$ ……，$i_a))$，其中 $R \in \sigma$ 是一个元数为 a 的关系符号，而 i_1，…… $i_a \in \{0, \ldots, k\}$ 是满足 $0 \in \{i_1, \ldots, i_a\}$ 的索引。为了证明 $\mathbb{A}$ 满足扩展公理 EA_k^σ，我们必须证明 $(v_{i_1},$ ……，$v_{i_a}) \in R^{\mathbb{A}}$ 当且仅当 $R(i_1,$ ……，$i_a) \in \tau$。注意，$v_0 \in \{v_{i_1}, \ldots, v_{i_a}\}$ 是唯一的顶点，使得在 F 中，有一条有向边从 index(v_0) 指向 { 中由 index(u_1)……，index$(u_a)\} \setminus \{$index$(v_0)\}$ 索引的所有顶点。此外，$\text{PATTERN}(v_0) = (f,$ $\tau)$。另外，请注意，f 的选择满足

$$\begin{cases} v_{i_l} = v_0 & \text{for all } l \text{ with } i_l = 0, \\ f(v_{i_l}) = i_l & \text{for all } l \text{ with } i_l \neq 0. \end{cases}$$

因此，按照我们的构造，$(v_{i_1},$ ……，$v_{i_a}) \in R^{\mathbb{A}}$ 当且仅当 $(R,$ $(i_1,$ ……，$i_a)) \in \tau$，从而满足扩展公理。

运行时间 构造 F 所需时间至多为 $2^{2^{3k \cdot 2}} \cdot k^2$。根据定理 6，我们可以在 $2^{c'k} n \log(n)$ 时间内构造 (n, k)-通用族。构造的剩余部分所需时间为$O(|$类型$(\sigma, k)| \cdot n^a$，其中 a 是 σ中各关系的最大元数。通过适当重新缩放 c 使得 $c \geq a$，可得到运行时间 $O(2^{2^{ck}} n^c)$。

□

D 引理 10的证明

引理 10。 对于每个签名 σ 以及任意 k、t，

$$\Pr_{G \sim G(1/2,n)}\left[G \not\models \text{EA}_k\right] = \text{negl}(n),$$
$$\Pr_{G \sim G_t(1/2,n)}\left[G \not\models \text{EA}_k^t\right] = \text{negl}(n),$$
$$\Pr_{\mathbb{A} \sim \text{Str}[\sigma,n]}\left[\mathbb{A} \not\models \text{EA}_k^\sigma\right] = \text{negl}(n).$$

证明。 我们仅对关系结构证明该命题，因为图和超图的证明与此类似且更简单。固定一个签名 $\sigma = \langle R_1, \ldots, R_t \rangle$，其中包含元数为 a_1，……，a_t 的关系。

于是， k-类型在 σ 上至多指定 $\sum_{i=1}^{t}(k+1)^{a_i}$ 个连接。固定一个元组 v_1，……， $v_k \in \mathbb{A}^k$ <style id='22}以及一个原子 k-类型。顶点 $v \notin S$ 满足 $\sum_{i=1}^{t}(k+1)^{a_i}$-类型所指定的每个连接的概率均为 $1/2$，且彼此独立。对于 $n-k$ 共有 v种选择。利用观察 24，没有顶点满足这些要求的概率为

$$\left(1-2^{-\sum_{i=1}^{t}(k+1)^{a_i}}\right)^{n-k} \leq e^{(n-k)/\sum_{i=1}^{t}(k+1)^{a_i}}.$$

对于 n^k 有 v_1、……、 v_k 的选择，原子 k-类型至多有 $2^{\sum_{i=1}^{t}(k+1)^{a_i}}$ 种。因此，并集界给出了所有这些选项下失败概率的上界，至多为

$$n^k \cdot 2^{\sum_{i=1}^{t}(k+1)^{a_i}} \cdot e^{(n-k)/\sum_{i=1}^{t}(k+1)^{a_i}}.$$

由于 σ 和 k 是固定的，该函数关于 n是可忽略的。 □

E 引理 11 的证明11

引理 11。 考虑一对具有相同签名的图/超图/结构 $\mathbb{A}_1, \mathbb{A}_2$ ， 它们都满足相应的 k-扩张公理。于是，对于每个建议结构 $\mathbb{X}$，以及每个量词秩至多为 k 且签名匹配的最小不动点逻辑 句子 φ ， 都有 $(\mathbb{A}_1, \mathbb{X}) \models \varphi \Leftrightarrow (\mathbb{A}_2, \mathbb{X}) \models \varphi$。

证明。我们通过证明 Duplicator 在具有 k 个棋子的无穷卵石博弈 [25,定义 11.4]中获胜，来说明两个结构满足直到量词秩 k 为止的相同 LFP 句子；该博弈在 $(\mathbb{A}_1, \mathbb{X})$ 和 $(\mathbb{A}_2, \mathbb{X})$ 上进行。假设我们有棋子 p_1、……、 p_k，将其放置在结构 $(\mathbb{A}_1, \mathbb{X})$ 上，并有棋子 q_1、……、 q_k，将其放置在 $(\mathbb{A}_2, \mathbb{X})$ 上；我们需要为 Duplicator 提出一个获胜策略，使其始终保持两个（有序的） k-元组棋子 p_1、……、 p_k和 q_1、……、 q_k 分别描述 $(\mathbb{A}_1, \mathbb{X})$ 与 $(\mathbb{A}_2, \mathbb{X})$ 之间的部分同构。如果 Spoiler 将棋子 p_i 放在 $(\mathbb{A}_1, \mathbb{X})$ 的一个 $\mathbb{X}$-元素上，那么 Duplicator 就将 q_i 放在 $(\mathbb{A}_2, \mathbb{X})$ 中完全相同的元素上。如果 Spoiler 将 p_i放在 $(\mathbb{A}_1, \mathbb{X})$ 的一个 $\mathbb{A}_1$-元素上，那么 Duplicator 观察（在关系结构的情况下） p_i 在 $\mathbb{A}_1$ 中位于 p_1、……、 p_k 中那些落在 $\mathbb{A}_1$上的棋子子集上所满足的$\leq k$-原子类型，并利用 k-扩张公理，将 q_i 放置到 $\mathbb{A}_2$ 中的一个元素上，使其在已放置的棋子上满足相同的 $\leq k$-原子类型，且这些棋子位于 $\mathbb{A}_2$中。结构的部分同构性由两个 k-原子类型相同这一点推出。因此，Duplicator 可以无限地遵循该策略，并且不会输掉博弈。图和超图的情形完全类似。 □

F 定理 13 的证明

我们需要下面这个引理，它将最小不动点逻辑的零一律扩展到了含自由变量的公式。

引理 26 （含自由变量的零一律）。 对于每个 最小不动点逻辑 σ-公式 $\varphi(\bar{x})$ ， 都存在一个无量词的 一阶逻辑 σ-公式 $\varphi(\bar{x})$ ， 使得当 $\overline{n}$ 趋于无穷大时，

$$\Pr_{\mathbb{A}\sim \mathrm{Str}[\sigma,n]}\left[\mathbb{A} \models \varphi(\bar{u}) \ \textit{iff}\ \mathbb{A} \models \overline{\varphi}(\bar{u}) \ \textit{for every}\ |\bar{x}|\textit{-tuple}\ \bar{u}\ \textit{over}\ [n]\right] \geq 1-\mathrm{negl}(n).$$

证明。 设 φ 是一个最小不动点逻辑公式，其自由变量为 x_1、 x_2、……、 x_ℓ，其形式为 $\varphi = \exists x_{\ell+1}\, \varphi^*$，其中 $x_{\ell+1}$ 是一个新变量，而 φ^* 是公式的其余部分。令 $\varphi^*[x_i/x_{\ell+1}]$ 表示该公式

将 φ^* 中所有变量 $x_{\ell+1}$ 的出现替换为 x_i 后得到。于是，我们可以将 φ 等价地改写为

$$\varphi = \left(\bigvee_{i\in[\ell]} \varphi^*[x_i/x_{\ell+1}] \right) \vee \exists x_{\ell+1} \left(\left(\bigwedge_{i\in[\ell]} x_{\ell+1} \neq x_i \right) \wedge \varphi^* \right). \tag{6}$$

形式为 $x_i = x_{l+1}$ 且 $x_i \neq x_{l+1}$ 属于 φ^* 的原子公式可以分别替换为 FALSE 和 TRUE。对每个量词重复这一过程，我们重写 φ，使得每个量词只对与 x_1、x_2、.....、x_ℓ 不同的变量进行量化，并且自由变量与量化变量之间不存在任何（其他）比较关系。

接下来，我们将关系中共同出现的自由变量和量化变量分离开来。令 σ' 为扩展签名，其中对于每个 a元关系 R_p、每个 $a' \leq a$ 以及变量 $\bar{x}$ 中的每个 a'元组 $\bar{y}$，我们添加一个新的（$a - a'$）元关系符号 $R_i^{\bar{y}}$。通过将每个原子公式 $R_i(\bar{y}\bar{z})$替换为 $R_i^{\bar{y}}(\bar{z})$，我们从φ 得到一个新公式 φ'；其中，该原子公式中的变量被划分为两部分，使得 $\bar{y}$和 $\bar{z}$分别包含 φ的自由变量和约束变量。由于在 $\mathrm{Str}[\sigma, n]$ 和 $\mathrm{Str}[\sigma', n]$中，所有关系都以二分之一的概率彼此独立地分布，因此对于每个元组 $\bar{u}$

$$\Pr_{\mathbb{A}\sim\mathrm{Str}[\sigma,n]}[\mathbb{A} \models \varphi(\bar{u})] \quad = \quad \Pr_{\mathbb{A}'\sim\mathrm{Str}[\sigma',n]}[\mathbb{A}' \models \varphi'(\bar{u})].$$

注意，φ' 中的变量不会出现在任何原子公式中。接下来，我们将 φ' 改写为析取范式并进行划分

$$\varphi(\bar{x}) = \bigvee_{i=1}^{s} \xi_i(\bar{x}) \wedge \psi_i(\bar{x}),$$

使得每个公式 $\xi_i(\bar{x})$ 都是无量词的，并且每个公式 $\psi_i(\bar{x})$ 只对不同于 $\bar{x}$ 的变量进行量化，且不包含依赖于来自 $\bar{x}$ 的变量的关系。因此，我们可以将 ψ_i 视为某个 σ' 结构中的句子，其中元素 $\bar{x}$ 已被移除。换言之，对于一个$|\bar{x}|$元组 $\bar{u}$

$$\Pr_{\mathbb{A}\sim\mathrm{Str}[\sigma,n]}[\mathbb{A} \models \varphi(\bar{u})] \quad = \quad \Pr_{\mathbb{A}'\sim\mathrm{Str}[\sigma',n]} \Pr_{\mathbb{A}''\sim\mathrm{Str}[\sigma'',n-|\bar{x}|]} \left[\bigvee_{i=1}^{s} \mathbb{A}' \models \xi_i(\bar{u}) \wedge \mathbb{A}'' \models \psi_i \right].$$

将每个 ψ_i 视为一个句子后，Blass、Gurevich 和 Kozen 的 LFP零一律 [5]（或引理 10 与引理 11 的组合）意味着，ψ_i 要么以 negl(n) 的概率成立，要么 $1 - \mathrm{negl}(n)$。相应地，令 ψ_i 为 TRUE 或 FALSE。于是，对于固定的 $\bar{u}$ 和 i，ψ_i 与 $\psi_i(\bar{u})$ 以 negl(n) 的概率不同。我们需要界定存在某个元组 $\bar{u}$ 和某个 i，使得 ψ_i 与 $\psi_i(\bar{u})$ 不同的概率。元组 $\bar{u}$ 只有多项式数量，而 i 只有常数数量，因此根据并集界，以及多项式数量的可忽略函数之和仍为可忽略函数这一事实，该陈述成立。 □

现在我们可以证明 LFP 的不可能性结果。

定理 13。设 $\sigma = \langle R_1, R_2, \ldots, R_t \rangle$ 和 $\tau = \langle R'_1, R'_2, \ldots, R'_{t'} \rangle$ 为关系签名，其中 R_i 的元数为 a_i，R'_i 的元数为 a'_i。以下陈述等价：

（*i*）存在一个从 σ 到 τ 的 （LFP，一阶逻辑）伪随机生成器。

(ii) 存在一个无量词一阶逻辑转导 Θ：结构类$[\sigma] \to$ 结构类$[\tau]$，使得对于所有 n，将 Θ 应用于 $\mathbb{A} \sim$结构类$[\sigma, n]$ 后在 结构类$[\tau, n]$ 上得到的分布，与分布 $\mathbb{B} \sim$ 结构类$[\tau, n]$ 统计上相同。

(iii) $\sigma \succeq_S \tau$。

证明。
(ii) $\Longrightarrow$(i)：显然。
(iii) $\Longrightarrow$ (ii)：对于正整数 a 和 k，令 $\mathsf{eq}(a,k)$ 表示 $[a]$ 上恰有 k 个等价类的等价关系集合，并令

$$\mathcal{S}_k := \{(i, \mathrm{eq}\,([a_i], k)) \mid i \in [t]\}, \qquad \mathcal{S}'_k := \{(i', \mathrm{eq}\,([a'_{i'}], k)) \mid i' \in [t']\}.$$

注意，对于任意 a 和 k，$T(a,k) = |\mathsf{eq}\,(a,k)| \cdot k!$，因为从一个集合 S 到一个大小为 k 的集合的满射函数，可以理解为先选取一个等价关系，以确定 S 中的哪些元素会映射到同一个元素，然后再选取一个从等价类到目标集合的双射（等价类的数量必须恰好为 k）。因此，假设 $\sigma \succeq_S \tau$，我们知道，对于每个正整数 k,

$$|\mathcal{S}_k| = \sum_{i=1}^{t} \frac{T(a_i,k)}{k!} \geq \sum_{i=1}^{t'} \frac{T(a'_i,k)}{k!} = \left|\mathcal{S}'_k\right|.$$

因此，存在注入 $f_k\colon \mathcal{S}'_k \to \mathcal{S}_k$。

我们通过以下算法描述一个转导 $\Theta = (\theta_1, \theta_2, \ldots, \theta_{t'}) : \mathsf{Str}[\sigma] \to \mathsf{Str}[\tau]$。假设我们希望确定某个索引 $1 \leq i' \leq t'$对应的 $\theta_{i'}(x_1, x_2, \ldots, {}^{x_{a'_{i'}}})$ 的真值。我们首先检查这些 x 变量之间的相等关系，从而得到索引集 $[a'_{i'}]$上的一个等价关系 $\sim'$。它将有若干个 k 等价类。接着令 $(i, \sim) := f_k(i', \sim')$。最后，返回 $R_i(y_1, y_2, \ldots, y_{a_i})$ 的真值，其中 y_j 指的是满足以下条件的变量 $x_{j'}$：通过测试 $(y_1, y_2, \ldots, y_j)$ 之间的相等关系所确定的等价关系为 $\sim$，且 y_j 属于 ℓ^{th} 等价类，该等价类包含 $\sim$，其中 ℓ 满足 $x_{j'}$ 属于 ℓ^{th} 等价类，该等价类包含 $\sim'$。注意，这里的所有“计算”都是根据相等性比较，确定如何将变量路由到 σ 中的某个关系；这完全可以通过一个较大（但有限）的无量词一阶逻辑句子纯粹以语法方式实现。

这一构造的要点是，对于每个 i 以及每个 $x_1, x_2, \ldots\ldots, {}^{x_{a_{i'}}}$，$\theta_i(x_1, x_2, \ldots\ldots, x_{a_i})$ 的真值取决于对输入 σ-结构$\mathbb{A}$中某个关系的独立赋值。通过检查是否存在两个元组 $(x_1, x_2, \ldots\ldots, {}^{x_{a_{i'}}})$ 和 $(x_1, x_2, \ldots\ldots, {}^{x_{a_{j'}}})$ 被映射到 $\mathbb{A}$ 中相同的关系赋值，即可验证这一点。为此请注意，由于所有变量都被使用，若 $\{x_1, x_2, \ldots, x_{a_{i'}}\} \neq \{x_1, x_2, \ldots, x_{a_{j'}}\}$，则这些元组会被映射到 $\mathbb{A}$ 中不同的赋值。否则，若它们包含相同的 k 互异元素集合，但等价关系类型不同，则它们会被映射到 σ 中不同的关系赋值，或不同的等价关系类型，因为 f_k 是单射。而如果它们具有相同的等价关系类型，那么当且仅当首次出现的顺序相同，也就是说，这些元组完全相同，它们才会被映射到相同的赋值。因此，根据独立性，Θ 会根据均匀随机的 σ-结构生成均匀随机的 τ 结构。

¬(iii) $\Longrightarrow$ ¬(i)：给定 $\Theta = (\theta_1, \theta_2, \ldots, \theta_{t'})$，其中每个 θ_i 都是一个具有 a'_i 个自由变量的 σ-公式。将引理 26 应用于 Θ中的每个公式，我们得到一个无量词的一阶逻辑转导 $\Theta := (\theta_1, \theta_2, \ldots, \theta_{t'})$，使得对于 $\mathbb{A} \sim$结构类 $[\sigma, n]$，结构 $\Theta(\mathbb{A})$ 与$\Theta(\mathbb{A})$以概率 $1 - \mathsf{negl}(n)$ 等价。

由于 $\overline{\Theta}$ <style id='4}>是无量词的，因此对于任意 k，它都满足引理15 的假设。于是，由于我们假设 $\sigma \not\succeq_S \tau$，引理15 告诉我们Θ 不是伪随机生成器，这意味着存在某个句子 φ 能够以不可忽略的优势将其输出与均匀随机的 τ 结构区分开来。因此，Θ 也不可能是伪随机生成器，因为三角不等式以及非可忽略函数加上可忽略函数仍为非可忽略函数这一事实意味着 φ 破坏了 Θ 的安全性。 □

G 引理15的证明

引理*15*。设 σ 和 τ 是关系签名，并且某个正整数 k 违反了条件 $\sigma \succeq_S \tau$。令 $f\colon \mathit{Str}[\sigma] \to \mathit{Str}[\tau]$ 为一个函数，使得对于任意正整数 $c \geq k$以及任意 σ-结构 $\mathbb{A}$，如果任意一对 c-元组（y_1、y_2、…、y_c）、（y'_1、y'_2、…、y'_c）的（c、k）-类型 $\in \mathbb{A}^c$在 $\mathbb{A}$ 中相同，那么它们的（c、k）-类型在 f（$\mathbb{A}$）中也相同。则将 f 应用于均匀随机的 σ-结构所得到的分布对一阶逻辑并非伪随机。

证明。令 k 为满足以下条件的正整数

$$\sum_{i=1}^{t} T(a_i, k) < \sum_{i=1}^{t'} T(a'_i, k). \tag{7}$$

令 $\mathcal{T}_c$ 为 (c,k)-类型的集合。对于任意 c和任意 $T \in \mathcal{T}_c$，令 $\varphi_T(y_1, y_2, \ldots, y_c)$ 为一阶逻辑公式，用于检验 $(y_1, y_2, \ldots, y_c)$ 是否两两不同且具有 (c,k)-类型 T。我们声称存在一个 c，使得对于所有充分大的 n，一阶逻辑句子

$$\varphi_c := \bigwedge_{T \in \mathcal{T}_c} \exists y_1, y_2, \ldots, y_c \varphi_T(y_1, y_2, \ldots, y_c)$$

在 $\mathbb{A} \sim$结构类中区分 $\mathbb{B} \sim \mathrm{Str}[\tau]$ 与 $f(\mathbb{A})$。

显然，当 $\mathbb{B} \sim$结构类$[\tau]$时，$\mathbb{B} \models \varphi_c$ 的 $1-$概率为 $\mathrm{negl}(n)$；这是观察结果 24 的直接推论。另一方面，我们将证明，对于足够大的 c，不存在 $\mathbb{A} \in$结构类$[\sigma, n]$使得 $f(\mathbb{A}) \models \varphi_c$。

我们通过计算以下数量来继续证明：N_σ 以及 N_τ 可能的（c，k）-类型的数量，即给定的 c-元组在任意 σ-结构或 τ-结构中分别可能具有的类型。令 $\mathcal{P}$（R）表示R的幂集，从定义14中的表达式出发，我们有

$$\begin{aligned} N_\sigma &= \left| \mathcal{P}\left(\bigcup_{i=1}^{t} \bigcup_{\substack{S \subseteq [c] \\ |S| \leq k}} \{(i, S, g) : \mid g : [a_i] \to S \text{ is surjective}\} \right) \right| \\ &= \left| \mathcal{P}\left(\bigcup_{k'=1}^{k} \bigcup_{i=1}^{t} \bigcup_{\substack{S \subseteq [c] \\ |S| = k'}} \{(i, S, g) : \mid g : [a_i] \to S \text{ is surjective}\} \right) \right| \\ &= 2^{\sum_{k'=1}^{k} \sum_{i=1}^{t} T(a_i, k') \binom{c}{k'}}. \end{aligned}$$

根据关于 f 的假设，对于任意固定的结构 $\mathbb{A}$，给定的任意 c-元组在 $f(\mathbb{A})$ 中只能具有 N_σ种可能的 $(c,\ k)$-类型。然而，存在

$$N_\tau = 2^{\sum_{k'=1}^{k} \sum_{i=1}^{t} T(a'_i, k') \binom{c}{k'}}$$

共计（c，k）-类型。对于足够大的 c，在对 k' 求和时，$\binom{c}{k'}$-项占主导地位；对于 N_σ，该项的系数为

$$\sum_{i=1}^{t} T(a_i, k),$$

而对于 N_τ，系数为

$$\sum_{i=1}^{t} T(a_i', k).$$

因此，由方程（7）可知 $N_\sigma < N_\tau$。因此，无论 $\mathbb{A}$ 是什么，都必定存在某种k-类型，没有任何 c-元组在 f（$\mathbb{A}$）中具有这种类型，因此 f（$\mathbb{A}$）$\not\models \varphi_c$。 □