arXiv:2603.09558v1 [cs.DB] 10 Mar 2026

禁止团：迈向 BDD/FC 猜想的下一步

Lucas Larroque*Inria, DI ENS, ENS, CNRS, PSL University*
lucas.larroque@inria.fr

Piotr Ostropolski-Nalewaja
University of Wrocław
postropolski@cs.uni.wroc.pl

米夏¨埃尔·托马佐*Inria, DI ENS, ENS, CNRS, PSL University*
michael.thomazo@inria.fr

摘要

本文探讨存在规则领域中的一个基本未决问题：有界推导深度规则集（bdd ⇒fc）的有限可控性猜想。我们通过证明，由 bdd 规则集生成的通用模型不可能包含任意大的竞赛图（任意大的有向团），却不蕴含环查询，从而朝着积极解决这一猜想迈进一步，$\exists x\ E(x, x)$。这一简洁而优雅的结果缩小了 (bdd ⇒fc) 猜想的潜在反例范围。

1 引言

过去十年间，数据库与知识表示领域之间的交流合作成果丰硕，旨在促进遗留数据库的集成与查询。其总体思路是允许用户使用逻辑形式系统来编写查询，并通过一阶公式约束其中谓词的含义。这引出了一个核心推理问题，称为基于本体的查询回答 (OBQA)：输入一个数据库 $\mathcal{I}$、一个规则集 $\mathcal{R}$ 和一个布尔合取查询 q，判断是否有 $\mathcal{I}, \mathcal{R} \models q$，其中 $\models$ 表示一阶逻辑的蕴涵关系。换句话说，是否任意一个 $\mathcal{I}$ 和 $\mathcal{R}$ 的模型（有限或无限）都是 q 的模型?

规则通常表示为存在规则[1]，它们是形如$\forall \bar{x},\ \bar{y}\ B(\bar{x},\ \bar{y}) \rightarrow \exists \bar{z}\ H(\bar{y},\ \bar{z})$ 的公式，其中 B 和 H 是原子的合取，分别称为规则的规则体 和规则头 。众所周知，如果不对规则施加其他约束，OBQA 是一个不可判定问题。

OBQA 的可判定性因此，人们投入了大量研究，致力于为 $\mathcal{R}$ 设计条件，以确保在 $\mathcal{R}$ 下布尔合取查询回答的可判定性，有时甚至确保其可处理性。理解这些条件的一项重要工具是**追赶** [24]。简而言之，追赶是一种算法，它添加新鲜项和新原子，以确保规则体的每个映射都能扩展为规则头的映射。这个可能无限的过程所产生的结果，是 $\mathcal{I}$ 和 $\mathcal{R}$的一个特定模型，它具有泛性质 [11]：它可以同态映射到 $\mathcal{I}$ 和 $\mathcal{R}$的任意模型。当追赶是有限的（例如，规则集满足某种无环性 [13, 23, 10, 16, 8] 时便是如此），它就能为查询回答提供一个判定过程。即使追赶不是有限的，只要它是一个树宽有界的结构，也同样如此——守卫规则 [5] 通常就属于这种情况。

最后，对我们的工作而言极为重要的是 UCQ-可重写性。若规则集 $\mathcal{R}$ 满足以下条件，则称其具有 UCQ-可重写性：合取查询（CQs）始终可以被重写。具体而言，对于每个合取查询 q，都存在一个合取查询的并（UCQ） $Q_{\mathcal{R}}$，使得对于任意数据库 $\mathcal{I}$：

$$\mathcal{I}, \mathcal{R} \models q \iff \mathcal{I} \models Q_{\mathcal{R}}.$$

鉴于追赶可能生成无限结构，规则集的 UCQ-可重写性保证了 OBQA 的可判定性，因为可以计算出重写结果，随后利用它查询数据库。UCQ-可重写性成为广泛研究的主题并不令人意外，在过去几十年中，这促成了众多子类的引入，包括：- 线性理论，其规则体中至多包含一个原子 [6]；守卫有界推导深度理论，它推广了线性理论 [3, 9]；黏性理论，其中对连接施加了严格限制 [7]。

UCQ-可重写性也以许多不同的名称被独立研究，这凸显了这一性质的根本性；其中值得注意的有有限合一集类（fus）和有界推导深度性质（bdd）。有限合一集类与反向链接过程的概念密切相关，该过程本质上是一种“反向追踪”。在这种方法中，为了回答蕴含问题，首先从一个查询开始，并尝试通过反向应用规则，从数据库中推导出一个子结构。bdd类将在下一节中正式介绍，其起源要早得多，可追溯至20世纪80年代。它源自Datalog程序有界性的经典概念 [15]。

[1]存在规则也被称为 *tuple-generating dependencies* [1] *conceptual* 、Datalog 以及文献中的其他名称。 *graph rules* [28] ± [19] ∀∃-*rules* [2]

有限数据库中的OBQA然而，数据库是有限结构；若认为 $\mathcal{I}$ 是对现实世界的部分描述，而 $\mathcal{R}$是应当在有限现实世界中成立的约束，那么一阶逻辑所提供的经典蕴涵概念便与我们所期望的蕴涵概念不符。事实上，我们更关心的是："在任何包含 $\mathcal{I}$ 且满足 $\mathcal{R}$的有限数据库中，q是否都成立？"这种语义显然与前一种语义相关，因为在非受限语义下， $\mathcal{I}$ 和$\mathcal{R}$ 的结论在有限语义下也成立。然而，反过来却不成立，以下典型示例便是明证。

例 1. 考虑 $\mathcal{I} = \{\mathrm{E}(a, b)\}$，以及包含以下两条规则的 $\mathcal{R}$ $\forall x \forall y\, \mathrm{E}(x, y) \rightarrow \exists z\, \mathrm{E}(y, z)$ 和 $\forall x \forall y \forall z\, \mathrm{E}(x, y) \wedge \mathrm{E}(y, z) \rightarrow \mathrm{E}(x, z)$。查询 $\exists x\, \mathrm{E}(x, x)$ 在非受限语义下不是 $\mathcal{I}, \mathcal{R}$ 的结论，追赶可作为见证。然而，在任何有限模型中都必然存在一个环；由于第二条规则强制了传递性，因此也必然存在一个自环。

有限可控性 有限推理为用户提供了一个用于判定非蕴涵的半判定过程（枚举有限解释，并检查是否存在一个不是 Q 模型的解释），但蕴涵不再是半可判定的，因为有限全模型可能不存在。Rosati [27], 早已提出一个有趣的问题：给定规则集 $\mathcal{R}$，对于任意数据库 $\mathcal{I}$ 和任意查询 Q，非受限语义与有限语义是否一致？若是，则称规则集 $\mathcal{R}$ 具有有限可控性 （fc）。注意，只要规则集具有有限可控性，查询回答便是可判定的。事实上，用于半判定一阶蕴涵的经典数学工具仍然可用，而有限结构可以作为非蕴涵的见证。Rosati 证明，包含依赖这一类受到严格限制的存在规则具有有限可控性 [27]。这一结果后来从两个方面得到了推广：推广到守卫规则 [4]，也推广到黏性规则 [18]。注意，这两个类都以不同方式推广了包含依赖：守卫规则具有有界树宽性质，而黏性规则具有有界推导深度性质。

Gogacz 和 Marcinkowski [17]提出了一个广为人知的猜想：具有有界推导深度（bdd）性质的规则集具有有限可控性（bdd ⇒fc）。此外，他们证明了该猜想在一种特定的受限情形下成立：bdd 规则集定义在二元签名上，且其规则头仅包含一个原子 [17]。然而，这一结果适用范围极其有限，而解决一般形式下该猜想的进展甚微。有限可控性是否适用于所有 bdd 规则集，仍是一个重要的未解决问题。

主要贡献 我们的主要贡献是证明 bdd 规则集具有某种模型论性质。我们用 $\textsc{Tournaments}_{\mathrm{E}}$ 表示以下查询：对于所有整数 k，存在一个集合 $\{x_1, \ldots, x_k\}$，由给定实例中的元素组成，使得 $\mathrm{E}(x_i, x_j)$ 或 $\mathrm{E}(x_j, x_i)$ 对所有 $i \neq j$ 都在该实例中成立；并用$\textsc{Loop}_{\mathrm{E}}$ 表示查询 $\exists x\ \mathrm{E}(x, x)$。那么，以下是我们的主要结果：

定理 1. 对于每个 有界推导深度 规则集 $\mathcal{R}$ 和每个实例 $\mathcal{I}$，我们都有：

$$(\mathcal{I},\mathcal{R}) \models \textsc{Tournaments}_{\mathrm{E}} \quad \Rightarrow \quad (\mathcal{I},\mathcal{R}) \models \textsc{Loop}_{\mathrm{E}}. \qquad (♙)$$

除了揭示 bdd 规则集的全称模型的结构性质外，这一结果也为证明（bdd ⇒fc）猜想迈出了重要一步。具体而言，它通过排除最自然的反例，缩小了该猜想潜在反例的范围。下面简要解释其原因。

首先注意，任何蕴含 $\textsc{Tournaments}_{\mathrm{E}}$ 且不蕴含 $\textsc{Loop}_{\mathrm{E}}$ 的模型都是无限的，因为在 $\textsc{Tournaments}_{\mathrm{E}}$ 中，项 x_1、.....、x_k 必须彼此不同，才能使 $\textsc{Loop}_{\mathrm{E}}$ 不被蕴含。因此，在有限情形下，$\textsc{Tournaments}_{\mathrm{E}}$ 确实蕴含 $\textsc{Loop}_{\mathrm{E}}$；而任何在（不受限制的）情形下不满足这一蕴含的 bdd 规则集，都将推翻（bdd ⇒fc）。

不过请注意，例 1 并不构成这样的反例，因为其规则集并非 bdd：要蕴含 E(a、b)，传递性规则至少必须应用与 a 和 b 之间的距离相同的次数，而这一次数无法独立于数据库的大小而得到界定。在该例的基础上，可以尝试使用 bdd 规则来模仿这种行为，即用 $\forall x\forall x'\forall y\forall y'$ E(x、x') ∧ E(y、y') → E(x、y') 替代传递性规则；该规则会蕴含传递性规则。经过这一修改，规则集终于成为 bdd，且追赶蕴含$\textsc{Tournaments}_{\mathrm{E}}$。然而，正如性质 (p)所预期的那样，一旦蕴含 $\exists x\exists y$ E(x、y)，这一新规则就会触发 $\exists x$ E(x、x) 的蕴含。

虽然定理 1 并不蕴含（有界推导深度 ⇒fc）猜想，因为原则上可能存在另一种类型的反例，但我们确信，后续各节所呈现的洞见以及精心整理的工具集，是在一般情形下解决（有界推导深度 ⇒fc）猜想这一努力中的一个重要里程碑。

论文结构 我们首先给出必要的预备知识，然后在第 3 节陈述主要结果。该证明分两个主要步骤进行。第一步（第 4 节）涉及一系列规则集“手术”。尽管其中一些已广为人知，但我们所需要的性质是针对当前具体情况量身定制的，因而需要单独的形式化证明。第二步（第5节）主要由一个复杂的模型论论证构成，用以表明性质 (p) 对变换后的规则成立。

2 预备知识

我们假定读者熟悉一阶逻辑以及图论中的基本概念。本节仅对这些主题作简要回顾。

2.1 一阶逻辑

我们定义 Preds 和 Vars 分别为谓词与变量的可数无穷集，且二者互不相交。我们为每个 P ∈ Preds 关联

一个元数 ar(P) ≥ 0。对于每个 $n \geq 0$，所有 n元谓词的集合也是可数无穷集。签名是谓词的集合。我们将变量的元组 $(x_1, \ldots, x_n)$ 记作 $\bar{x}$，并且常常将这些元组视为集合。变量元组的支撑是其中出现的变量所组成的集合。变量元组$\bar{y} = (y_1, \ldots, y_m)$与另一个元组 $\bar{x} = (x_1, \ldots, x_n)$兼容，当且仅当 $m = n$，并且对于所有 i，$j \leq n$，只要 $x_i = x_j$，就有 $y_i = y_j$。此外，若对于所有i，要么 $y_i = x_i$，要么存在某个 j使得 $y_i = y_j = x_j$，则称 $\bar{y}$ 是 $\bar{x}$的特化。任意两个元组 $\bar{x} = (x_1, \ldots, x_n)$ 和 $\bar{t} = (t_1, \ldots, t_n)$，只要 $\bar{t}$与$\bar{x}$ 兼容，就定义一个替换 $[\bar{x} \mapsto \bar{t}]$：$x_i \mapsto t_i$。

谓词 P 上的**原子**，或称 P-原子，是形如 P($\bar{x}$) 的一阶公式，其中 $\bar{x}$ 是变量元组，P 是 $|\bar{x}|$ 元谓词。对于零元谓词 P，我们用 P 简写零元原子 P()。签名$\mathbb{S}$ 上的**实例** 是由 $\mathbb{S}$中的谓词上的原子构成的集合。为简化表述，我们假设所有实例都包含一个零元事实 $\top$。原子或实例的**活动域** 是其中出现的变量集合。原子（或其集合）的活动域记作 adom(S)。两个实例 $\mathcal{I}_1$ 和 $\mathcal{I}_2$ 的**不交并** $\mathcal{I}_1 \dot{\cup} \mathcal{I}_2$ 为 $\mathcal{I}_1 \cup \sigma(\mathcal{I}_2)$，其中 σ 是一个函数，它将$\mathcal{I}_2$ 中的变量映射为新变量，这些新变量不会出现在 $\mathcal{I}_1$中。对于一阶公式 F 和变量元组 $\bar{x}$，我们用 $F(\bar{x})$ 表示 $\bar{x}$ 是 F中所有自由变量构成的集合。

存在规则一个（存在）规则 ρ 是如下形式的一阶逻辑公式：

$$\forall \bar{x}, \bar{y}\ B(\bar{x}, \bar{y}) \to \exists \bar{z}\ H(\bar{y}, \bar{z})$$

其中，B 和 H 是非空的原子合取。我们将 $\bar{y}$ 称为 ρ的前沿，并分别将 $B(\bar{x}, \bar{y})$ 和 $\exists \bar{z}\ H(\bar{y}, \bar{z})$称为 ρ的规则体和规则头。如果某规则不含存在变量，即 $\bar{z}$ 为空，则该规则是$Datalog$ 规则。给定规则 ρ，集合 body(ρ)、head(ρ) 和 fr(ρ)分别表示 ρ的规则体、规则头和前沿。我们将 body(ρ)（分别为 head(ρ)）中的元素称为 ρ的体原子（分别为头部原子）。书写规则时省略 $\forall$量词。如果规则集中的任意规则所含的原子均属于签名 $\mathbb{S}$中的某个谓词，则称该规则集基于签名 $\mathbb{S}$ 。

同态 A 替换 π 是一个从 Vars 到 Vars 的函数。对于一个原子 P($\bar{t}$)，令 π(P($\bar{t}$)) 表示将 P($\bar{t}$) 中每个变量 x 的所有出现都替换为 $\pi(x)$（如果后者有定义）后得到的原子。从一个实例 $\mathcal{A}$ 到一个实例 $\mathcal{B}$ 的同态 π是一个替换，使得 $\pi(\mathcal{A}) \subseteq \mathcal{B}$。从 $\mathcal{A}$ 到 $\mathcal{B}$ 的同构 π 是一个从$\mathcal{A}$ 到 $\mathcal{B}$的单射且满射同态。若存在从 $\mathcal{A}$ 到 $\mathcal{B}$ 的同态，以及从 $\mathcal{B}$ 到 $\mathcal{A}$的同态，则两个原子集 $\mathcal{A}$ 和 $\mathcal{B}$ 是同态等价 的。我们用 $\leftrightarrow$ 符号表示同态等价。

查询A 合取查询(CQ) $q(\bar{x})$ 是一个二元组，其中一阶公式的形式为 $\exists \bar{z}\ B(\bar{x}, \bar{z})$，$B$ 是非空的原子的合取，而元组$\bar{x}$ 由答案变量组成，其支撑是该公式的自由变量集合。A

若 CQ 不含答案变量，则它是布尔的。合取查询的并 (UCQ) $Q(\bar{x})$ 是一组 CQ 的析取，我们常将其视为 CQ 的集合$\{ q_1(\bar{x}_1) \ldots, q_n(\bar{x}_n) \}$，以及一个答案变量元组 $\bar{x}$，使得对所有 i，$\bar{x}_i$ 都是 $\bar{x}$的特化。给定一个 UCQ $Q(\bar{x}, \bar{y})$ 和一个与 $\bar{x}$兼容的变量元组 $\bar{x}'$，我们用 $Q(\bar{x}', \bar{y})$表示 UCQ $Q[\bar{x} \mapsto \bar{x}'](\bar{x}', \bar{y})$。在标准一阶语义下，若对于一个与 $\bar{x}$ 兼容的、由 adom(I) 中元素构成的给定元组 $\bar{t}$，存在从 $q_k(\bar{x}_k[\bar{x} \mapsto \bar{t}])$ 到 $\mathcal{I}$ 的同态，其中某个 $1 \leq k \leq n$满足条件，则实例 $\mathcal{I}$蕴含 UCQ $Q(\bar{x}) = \{ q_1(\bar{x}_1) \ldots, q_n(\bar{x}_n) \}$。此外，若该同态是单射，则 $\mathcal{I}$ 单射蕴含 Q。我们分别用 $\mathcal{I} \models Q(\bar{t})$ 和 $\mathcal{I} \models_{\text{inj}} Q(\bar{t})$表示这些事实。对于规则集 $\mathcal{R}$、实例 $\mathcal{I}$和 UCQ Q，我们写作 $\langle \mathcal{R}, \mathcal{I} \rangle \models Q(\bar{t})$，表示 $\mathcal{R} \cup \mathcal{I}$在非受限模型上蕴含 $Q(\bar{x})$，其中 $\bar{t}$；而 $\langle \mathcal{R}, \mathcal{I} \rangle \models_{\text{inj}} Q(\bar{t})$ 表示 $\mathcal{R} \cup \mathcal{I}$在非受限模型上单射蕴含 $Q(\bar{x})$，其中 $\bar{t}$。

2.2 追赶

给定规则集 $\mathcal{R}$，某个实例 $\mathcal{I}$ 上的一个（$\mathcal{R}$-）触发器 是由规则$\rho \in \mathcal{R}$ 与从规则体(ρ)到 $\mathcal{I}$ 的某个子集的同态 h 所构成的对 $\langle \rho, \mathrm{h} \rangle$。触发器 $\tau = \langle \rho, \mathrm{h} \rangle$ 在 $\mathcal{I}$ 上的输出 是实例输出(τ) = h′(规则头(ρ))，其中 h′ 是通过将规则头(ρ)的存在量化变量映射到新变量来扩展 h 的同态。我们用触发器($\mathcal{I}, \mathcal{R}$)表示 $\mathcal{I}$ 与 $\mathcal{R}$ 的全部触发器的集合。（遗忘式）追赶 [21, 6] 的第 k 步记为 $Ch_k(\mathcal{I}, \mathcal{R})$，并定义为

$$Ch_0(\mathcal{I}, \mathcal{R}) = \mathcal{I} \qquad Ch_{n+1}(\mathcal{I}, \mathcal{R}) \;=\; Ch_n(\mathcal{I}, \mathcal{R}) \cup \bigcup_{\tau \in T_n} \mathsf{output}(\tau) \qquad (\text{for } n \in \mathbb{N}_{>0})$$

其中 T_n =触发器($Ch_n(\mathcal{I}, \mathcal{R})$ $\mathcal{R}$)\触发器($Ch_{n-1}(\mathcal{I}, \mathcal{R})$ $\mathcal{R}$)。从 $\langle \mathcal{I}$、$\mathcal{R} \rangle$出发的追赶结果记为 $Ch(\mathcal{I}, \mathcal{R})$，是所有$n$上 $Ch_n(\mathcal{I}, \mathcal{R})$ 的并集。众所周知，$Ch(\mathcal{I}, \mathcal{R})$ 是 $\mathcal{I}$ 与 $\mathcal{R}$ 的通用模型。为便于阅读，我们将 $Ch_n(\{\top\}, \mathcal{R})$ 记为 $Ch_n(\mathcal{R})$，将 $Ch(\{\top\}, \mathcal{R})$ 记为 $Ch(\mathcal{R})$。

给定一个实例 $\mathcal{I}$ 和一个规则集 $\mathcal{R}$，追赶项 是活动论域 adom($Ch(\mathcal{I}, \mathcal{R})$) \adom($\mathcal{I}$) 的一个元素。因此，每个追赶项都是通过应用某个触发器创建的。由某个触发器 $\langle \rho, \mathrm{h} \rangle$ 创建的追赶项 t 的前沿 是集合 h(fr(ρ))。

2.3 UCQ-可重写性

定义 2. 若对于所有合取查询 $q(\bar{x})$，都存在一个合取查询的并$Q(\bar{x})$，使得对于所有与 $\bar{x}$兼容的实例 $\mathcal{I}$ 以及由 adom($\mathcal{I}$) 中元素组成的元组 $\bar{t}$，都有

$$\langle \mathcal{I}, \mathcal{R} \rangle \models q(\bar{t}) \iff \mathcal{I} \models Q(\bar{t}).$$

合取查询的并 $Q(\bar{x})$ 是 $q(\bar{x})$ 的（UCQ-）重写，记作 rew($q(\bar{x})$ $\mathcal{R}$)。

请注意，rew($q(\bar{x})$ $\mathcal{R}$) 并非唯一；不过，我们可以取其中最小的一个，它在变量的双射重命名下是唯一确定的 [22]。此外，合取查询的并重写的概念可以轻松地从合取查询推广到合取查询的并。

定义 3. 若对于每个合取查询 q，都存在某个 $k \geq 0$，使得对于所有实例 $\mathcal{I}$，$\langle \mathcal{I}, \mathcal{R} \rangle \models q$ 当且仅当 $Ch_k(\mathcal{I}, \mathcal{R}) \models q$，则称一个规则集具有有界推导深度（*bdd*）。我们用 bdd 表示具有 bdd 的规则集类别。对于给定的 q，我们将上述最小的 k 称为 q 的 bdd-常数，并记为 bdd(q, $\mathcal{R}$)。

有趣的是，这些概念是一致的。有关此结果的完整证明，请参见 [20]。

命题 4. 一个规则集是可合取查询的并重写的，当且仅当它具有有界推导深度。

引理 5. 设 $\mathcal{R}_1$ 和 $\mathcal{R}_2$ 都具有有界推导深度。若对于每个实例 $\mathcal{I}$，都有$Ch(Ch(\mathcal{I}, \mathcal{R}_1), \mathcal{R}_2)$ 与 $Ch(\mathcal{I}, \mathcal{R}_1 \cup \mathcal{R}_2)$ 同态等价，那么 $\mathcal{R}_1 \cup \mathcal{R}_2$ 就具有有界推导深度。

证明。设 Q 是合取查询 q 针对 $\mathcal{R}_2$ 的一个重写，接着设 Q' 是 Q 针对 $\mathcal{R}_1$的一个重写。注意，Q' 是 q 针对 $\mathcal{R}_1 \cup \mathcal{R}_2$ 的一个重写，因为$Ch(Ch(\mathcal{I}, \mathcal{R}_1), \mathcal{R}_2) \leftrightarrow Ch(\mathcal{I}, \mathcal{R}_1 \cup \mathcal{R}_2)$。 □

之后我们将需要单射重写，也就是说，我们希望将定理 2 改述如下：

定义 2（改述版）。若对所有合取查询 $q(\bar{x})$，都存在一个合取查询的并 $Q(\bar{x})$，使得对于所有实例 $\mathcal{I}$ 以及活动论域 adom($\mathcal{I}$)中与 $\bar{x}$相容的元素元组 $\bar{t}$，都有：

$$\langle \mathcal{I}, \mathcal{R} \rangle \models q(\bar{t}) \iff \mathcal{I} \models_{\text{inj}} Q(\bar{t})$$

合取查询的并 $Q(\bar{x})$ 是 $q(\bar{x})$ 的 单射重写，记作 rew^{inj}($q(\bar{x})$, $\mathcal{R}$)。

由以下内容可知，上述定义确实是一个等价定义：

命题 6. 对于每个合取查询的并 Q，都存在另一个合取查询的并 Q_{inj}，使得对于每个实例 $\mathcal{I}$ 和元组 $\bar{a}$：

$$\mathcal{I} \models Q(\bar{a}) \iff \exists q \in Q_{\text{inj}} \quad \mathcal{I} \models_{\text{inj}} q(\bar{a}) \iff \mathcal{I} \models Q_{\text{inj}}(\bar{a}).$$

证明。我们将证明 Q 为合取查询的情形；证明很容易推广到合取查询的并的情形。设 q 为一个合取查询，$\bar{x}$ 为其变量组成的元组。注意，每当 q 通过同态 h 映射到某个实例 $\mathcal{I}$ 时，都存在 $\bar{x}$ 的一个特化 $\bar{y}$，以及一个单射同态 h'，使得 $[\bar{x} \mapsto \bar{y}] \circ h' = h$。那么，所需的合取查询的并只需取 $\bar{x}$ 的所有特化 $\bar{y}$ 对应的 $q[\bar{x} \mapsto \bar{y}]$ 的析取即可。要得到第二个等价关系，注意上述构造是幂等的。

□

2.4 其他有用概念

图论概念 一个（有向）图 是二元组 $\langle V, E \rangle$，由集合 V 的顶点以及 $V \times V$ 的子集 E 组成，其元素称为边。图的大小 是其顶点的数量。我们常说一条边$\langle v, w \rangle \in E$ 从 v 到 w。一个自环 是从节点到其自身的边。一个

子图 $\langle V', E' \rangle$ 是图 $\langle V, E \rangle$ 的一个图，使得 $V' \subseteq V$ 且 $E \subseteq E'$。$\langle V, E \rangle$ 的子图 $\langle V', E' \rangle$ 是导出的，如果对所有 $v, w \in V'$，若 $\langle v, w \rangle \in E$，则 $\langle v, w \rangle \in E'$。两个顶点 s 与 t 之间的路径是顶点序列 $v_0, \ldots, v_n$，使得 $v_0 = s, v_n = t$，且对所有 $i < n, \langle v_i, v_{i+1} \rangle \in E$。环是从节点到其自身的路径。若图不含任何环，则该图是无环的。竞赛图是这样的图：对所有不同的顶点 v, w，我们有 $\langle v, w \rangle \in E$ 或[2] $\langle w, v \rangle \in E$。图 $\langle V, E \rangle$ 的边的 k-着色是从 E 到 $\{1, \ldots, k\}$ 的函数。[3] 给定图 G 及其边的 k-着色 C，G 的子图 $\langle V', E' \rangle$ 对某个 $i \leq k$ 以 i 着色，如果对所有 $e \in E$，我们有 $C(e) = i$。以下是有向图的拉姆齐定理的一种改述。

定理 7. *对于任意整数 $s_1, \ldots, s_k \geq 1$，存在某个整数 $R(s_1, \ldots, s_k)$，使得对于任意大小至少为 $R(s_1, \ldots, s_k)$ 的竞赛图 T，以及对 T 的边的任意 k 着色，都存在 T 的一个子图，它是一个大小为 s_i 、以某种 i 着色的竞赛图。*

多重集 集合 D 上的多重集是一个从 D 到 $\mathbb{N}$ 的函数 M。若 $\{x \in D \mid M(x) > 0\}$ 有限，则该多重集是有限的。有限多重集 M 的大小为 $|M| = \sum_{x \in D} M(x)$。空多重集为 $\emptyset_\mathsf{m}$：$x \mapsto 0$。给定由 D 中元素组成的有限列表 $[x_1, \ldots, x_n]$，我们将多重集 $x \mapsto |\{i \mid x_i = x\}|$ 记为 $\{x_1, \ldots, x_n\}_\mathsf{m}$。给定一个从某个有限集合 E 到 D 的函数 f，我们将多重集 $x \mapsto |f^{-1}(x)|$ 记为 $\{f(x) \mid x \in E\}_\mathsf{m}$。两个多重集 M 和 N 的并集为多重集 $M \cup_\mathsf{m} N$：$x \mapsto M(x) + N(x)$，它们的交集为 $M \cap_\mathsf{m} N$：$x \mapsto \min(M(x) N(x))$，差集为 $M \setminus_\mathsf{m} N$：$x \mapsto \max(M(x) - N(x), 0)$。

考虑 $<$上的严格序。多重集 M 上的最大值在有定义时为 $\max_\mathsf{m}(M) = \max\{x \in D \mid M(x) > 0\}$，否则无定义。多重集上的严格字典序 $<_{\text{lex}}$ 归纳定义如下：对于所有非空多重集 M、$\emptyset_\mathsf{m} <_{\text{lex}} M$，以及给定的两个非空多重集 M和 N，当且仅当 $\max_\mathsf{m}(M) < \max_\mathsf{m}(N)$，或者 $\max_\mathsf{m}(M) = \max_\mathsf{m}(N)$ 且 $(M \setminus_\mathsf{m} \{\max_\mathsf{m}(M)\}_\mathsf{m}) <_{\text{lex}} (N \setminus_\mathsf{m} \{\max_\mathsf{m}(N)\}_\mathsf{m})$ 时，有 $M <_{\text{lex}} N$。字典序 $\leq_{\text{lex}}$ 定义为：当且仅当 $M <_{\text{lex}} N$ 或 $M = N$（按通常定义）时，有 $M \leq_{\text{lex}} N$。

引理 8。 *如果 $<$ 在 D上是良基的，那么对于所有 k，序 $<_{\text{lex}}$ 在 D 上大小至多为 k 的多重集集合上是良基的。*

证明。我们对 k进行归纳。由于大小为 0 的唯一多重集是 $\emptyset_\mathsf{m}$，因此 $k = 0$ 的情形显然成立。接下来，考虑由 D上的大小至多为 k 的多重集构成的集合 $\mathbb{S}_k$。由于 $\mathbb{S}_k$ 中的所有元素都是有限的，它们都有最大值。由于$<$ 在 D上是良基的，集合 $\{\max_\mathsf{m}(M) \mid M \in \mathbb{S}_k\}$ 存在最小值 m。考虑集合 $\mathbb{S}_{k-1} = \{M \mid M \cup_\mathsf{m} \{m\}_\mathsf{m} \in \mathbb{S}_k \wedge \max_\mathsf{m}(M \cup_\mathsf{m} \{m\}_\mathsf{m}) = m\}$。由于该集合中的所有多重集大小都至多为 $k-1$，根据归纳假设，它存在最小值 $M'_{\min}$。接着，我们断言 $M_{\min} = M'_{\min} \cup_\mathsf{m} \{m\}_\mathsf{m}$是

[2]注意，这里的“或”是包容性的，与竞赛图的通常定义相反。[3]注意，我们并不要求相邻顶点具有不同的颜色。

$\mathbb{S}_k$的最小值。考虑 $\mathbb{S}_k$中的一个元素 M。若 $M \setminus_{\mathsf{m}} \{m\}_{\mathsf{m}} \notin \mathbb{S}_{k-1}$，则根据 $\mathbb{S}_{k-1}$的定义，$\mathsf{max}_{\mathsf{m}}(M) > m = \mathsf{max}_{\mathsf{m}}(M_{\min})$，因此 $M_{\min} \leq_{\text{lex}} M$。否则，若 $M \setminus_{\mathsf{m}} \{m\}_{\mathsf{m}} \in \mathbb{S}_{k-1}$，则$\mathsf{max}_{\mathsf{m}}(M) = m = \mathsf{max}_{\mathsf{m}}(M_{\min})$，且$M_{\min} \setminus_{\mathsf{m}} \{m\}_{\mathsf{m}} = M'_{\text{最小值}} \leq_{\text{lex}} M \setminus_{\mathsf{m}} \{m\}_{\mathsf{m}}$，因为 M'_{min}是 $\mathbb{S}_{k-1}$的最小值。因此， $M_{\min} \leq_{\text{lex}} M$，证明完毕。

□

3 主要结果

定义 9. 给定规则集 $\mathcal{R}$、实例 $\mathcal{I}$ 以及二元谓词 E，我们用 $\textsc{Tournaments}_{\mathsf{E}}$ 表示如下查询：[4] 对所有整数 k，存在一个 E-竞赛图，其大小为 k，且位于给定实例中。

定义 10. 我们定义 E-环查询 $\exists x\, \mathsf{E}(x,\ x)$，并用 $\textsc{Loop}_{\mathsf{E}}$ 表示它。

在本文的其余部分，我们固定一个二元谓词 E，以同时用于 $\textsc{Tournaments}_{\mathsf{E}}$ 和 $\textsc{Loop}_{\mathsf{E}}$。

定理 1（重述）。对每个签名 $\mathbb{S}$、每个 *UCQ* 可重写规则集 $\mathcal{R}$（定义于 $\mathbb{S}$ 之上），以及每个实例 $\mathcal{I}$，我们有：

$$Ch(\mathcal{I},\mathcal{R}) \models \textsc{Tournaments}_{\mathsf{E}} \quad \Rightarrow \quad Ch(\mathcal{I},\mathcal{R}) \models \textsc{Loop}_{\mathsf{E}}. \qquad (\clubsuit)$$

在本定理以及具有类似结构的陈述中，我们将 (p) 称为结论，而其前面的部分是假设。我们将用本文余下的篇幅证明上述主要结果。

4 化简主定理

本节的目标是将本文主定理中的假设简化为一个更易于控制的假设：所考虑的实例始终为$\{\top\}$，且规则集满足若干性质。定理28体现了这一总体目标；正如我们将在本节中证明的那样，定理28等价于定理 1。这里的难点在于证明定理28推出定理 1，因为反方向显然成立。证明分若干步进行，每一步都证明它与某个中间命题（定理11、18和23）等价，并逐步加强假设。本节所用的工具包括若干复杂程度不一的规则集变换。

证明的所有部分（关于定理 1 与定理28等价性的证明）都将遵循同一个公式，我们会在讨论下一步时一并说明。

[4]为了使 $\models$ 对 $\textsc{Tournaments}_{\mathsf{E}}$ 有意义，我们可以通过以下二阶逻辑公式定义该查询：$\forall \mathtt{X}$ 有限，$\exists \mathtt{Y}\ \forall x, y\ \exists x', y'\ \ \mathtt{X}(x) \wedge \mathtt{X}(y) \ \rightarrow\ \mathtt{Y}(x, x') \wedge \mathtt{Y}(y, y') \wedge \mathtt{E}(x', y')$. X 的有限性通过如下方式表达：从 X 到 X 的所有函数都是单射，当且仅当它们是满射。

4.1 将实例编码到规则集中

我们从第一个中间命题开始，将定理 1 中的实例设为 $\{\top\}$。

引理 11. 对于每个签名 $\mathbb{S}$，以及每个定义在 $\mathbb{S}$ 上的 *UCQ* 可重写规则集 $\mathcal{R}$，我们有：

$$Ch(\mathcal{R}) \models \textsc{Tournaments}_{\mathtt{E}} \quad \Rightarrow \quad Ch(\mathcal{R}) \models \textsc{Loop}_{\mathtt{E}}. \qquad (\Leftarrow)$$

本步骤的目标是证明定理 11 与定理 1 等价。由于我们限制了潜在实例的集合，证明的 $(\Rightarrow)$ 方向是平凡的，即定理 1 蕴含定理 11。

在本证明步骤及后续步骤中，我们将始终用逆否法证明 $(\Leftarrow)$ 方向。因此，假设签名 $\mathbb{S}$、规则集 $\mathcal{R}$ 与实例 $\mathcal{I}$ 构成定理 1 的一个反例，现在我们将构造 $\mathcal{R}'$，使得 $\mathcal{R}'$ 与 $\mathbb{S}$ 构成定理 11 的一个反例。

规则集手术

反例的构造将始终包含某种规则集手术。

定义 12。 对于任意实例 $\mathcal{J}$，我们用 $\top \to \mathcal{J}$ 表示以下规则

$$\top \quad \to \quad \exists f(\mathsf{adom}(\mathcal{J})) \quad \bigwedge_{\mathtt{A}(\bar{t}) \in \mathcal{J}} \mathtt{A}(f(\bar{t}))$$

其中，f 是将项双射重命名为新变量所得的结果。

观察 13。 设 ρ 是任意规则，$\top$ 是其规则体，则对于任意规则集 $\mathcal{S}$ 和任意实例 $\mathcal{J}$，有：

$$Ch(\mathcal{J}, \mathcal{S} \cup \{\rho\}) \leftrightarrow Ch(Ch(\mathcal{J}, \{\rho\}),\, \mathcal{S}).$$

证明。 注意，ρ 总是会触发，且只会触发一次。 □

我们只需令 $\mathcal{R}' = \mathcal{R} \cup \{\top \to \mathcal{I}\}$。构造出 $\mathcal{R}'$ 后，现在我们将证明，新构造的规则集 $\mathcal{R}'$ 与原始的 $\mathbb{S}$ 构成定理 11 的一个反例。这始终分为两部分：证明假设成立——在此情形下即证明 $\mathcal{R}'$ 可 UCQ 重写——以及证明结论不成立——在此情形下即证明：

$$Ch(\mathcal{R}') \models \textsc{Tournaments}_{\mathtt{E}} \ \wedge \ Ch(\mathcal{R}') \not\models \textsc{Loop}_{\mathtt{E}}$$

手术的主要性质

与接下来的步骤一样，我们在此介绍手术的主要性质。这些小节旨在以模块化且可复用的方式阐述证明的主要基本组成部分。正如引言中所述，其中一些技术在学术界早已广为人知。然而，这些技术往往缺少

恰当的表述或正式的陈述。值得注意的是，大多数技术在 UCQ-可重写性方面的性质远未广为人知，而这些性质本身也是值得学术界单独关注的结果。最后，我们指出，通常会陈述两项重要性质——一项用于“证伪结论”部分，另一项用于“证明假设”部分。

观察 14. 对于每一对实例 $\mathcal{J}$ 和 $\mathcal{J}'$ 以及规则集 $\mathcal{S}$，我们有 $Ch(\mathcal{J} \cup \mathcal{J}', \mathcal{S})$ 与 $Ch(\mathcal{J}, \mathcal{S} \cup \{\top \to \mathcal{J}'\})$同态等价。

证明. 考虑以下等价序列:

$$\begin{aligned} Ch(\mathcal{J} \cup \mathcal{J}', \mathcal{S}) &\leftrightarrow Ch(\, Ch(\mathcal{J}, \{\top \to \mathcal{J}'\}),\, \mathcal{S}\,) && \text{(Definition of } \{\top \to \mathcal{J}'\}) \\ &\leftrightarrow Ch(\mathcal{J}, \mathcal{S} \cup \{\top \to \mathcal{J}'\}) && \text{(From Theorem 13)} \quad \square \end{aligned}$$

以下内容使我们能够概括手术的第一个性质:

推论 15. 对于任意实例 $\mathcal{J}$ 和任意规则集 $\mathcal{S}$，我们有:

$$Ch(\mathcal{J}, \mathcal{S}) \;\leftrightarrow\; Ch(\{\top\}, \mathcal{S} \cup \{\top \to \mathcal{J}\}).$$

并且以下是它的第二个性质:

观察 16. 给定一个可 *UCQ* 重写的规则集 $\mathcal{S}$ 和一个实例$\mathcal{J}$，我们有

$\mathcal{S} \cup \{\top \to \mathcal{J}\}$ 可 *UCQ* 重写。

证明。 由定理 5 和 13 可知， $\mathcal{S} \cup \{\top \to \mathcal{J}\}$ 可 UCQ 重写。

□

证伪结论：（N）

引理 17。 $Ch(\mathcal{R}') \models \textsc{Tournaments}_{\mathtt{E}} \wedge Ch(\mathcal{R}') \not\models \textsc{Loop}_{\mathtt{E}}$。

证明。 由以下内容得出:

- $Ch(\mathcal{I}, \mathcal{R}) \leftrightarrow Ch(\mathcal{R}')$（定理 15）
- $Ch(\mathcal{I}, \mathcal{R}) \models \textsc{Tournaments}_{\mathtt{E}} \wedge Ch(\mathcal{I}, \mathcal{R}) \not\models \textsc{Loop}_{\mathtt{E}}$(来自假设)。

□

证明定理 11 的假设

在这种情况下，我们必须证明 $\mathcal{R}'$ 是可 UCQ 重写的。根据定理 16，这可以由 $\mathcal{R}$ 是可 UCQ 重写的推出。

4.2 归约为二元签名

引理 18. 对于每个二元签名 $\mathbb{S}$，以及每个基于 $\mathbb{S}$ 的 *UCQ* 可重写规则集$\mathcal{R}$，我们有：

$$Ch(\mathcal{R}) \models \textsc{Tournaments}_{\mathtt{E}} \quad \Rightarrow \quad Ch(\mathcal{R}) \models \textsc{Loop}_{\mathtt{E}} \qquad (\text{🔔})$$

现在我们证明定理 11 与定理 18 等价。(⇒) 方向是平凡的，我们通过对位法证明 (⇐)。设 $\mathbb{S}$ 为一个签名，$\mathcal{R}$ 为一个规则集，二者共同构成定理 11 的一个反例。

再化

给定一个元数为 ar(A) > 2 的关系符号 A，我们将 reify(A) 定义为如下二元谓词集合 $\{\mathrm{A}_1, \ldots, \mathrm{A}_{\mathsf{ar}(\mathtt{A})}\}$。给定一个划分为至多二元谓词与高元谓词的有限签名 $\mathbb{S} = \mathbb{S}_{\leq 2} \uplus \mathbb{S}_{\geq 3}$，我们将 $\mathbb{S}$ 的再化版本 定义为二元签名 reify($\mathbb{S}$) $= \mathbb{S}_{\leq 2} \uplus \bigcup_{\mathtt{A} \in \mathbb{S}_{\geq 3}}$ reify(A). 给定一个原子 $\alpha = \mathrm{A}(x_1, \ldots, x_n)$，我们将 reify($\alpha$) 定义为 $\{\mathrm{A}_i(x_i, x_\alpha) \mid 1 < i \leq n\}$，对于 $n > 2$，并令 reify 在元数小于三的原子上为恒等映射。我们以自然的方式将 reify 提升到实例、规则和查询上。

我们证明，$\mathcal{R}' =$ reify($\mathcal{R}$) 和 $\mathbb{S}' =$ reify($\mathbb{S}$) 构成定理 18 的一个反例。

再化的主要性质

我们从 Feller 等人那里得到再化的第一个性质。[14]（再化的性质 (ii)：）

引理 19。对于任意实例 $\mathcal{J}$ 和任意规则集 $\mathcal{S}$，我们有

$$Ch(\mathsf{reify}(\mathcal{J}), \mathsf{reify}(\mathcal{S})) \leftrightarrow \mathsf{reify}(Ch(\mathcal{J}, \mathcal{S})).$$

据我们所知，再化与 UCQ-可重写性之间的关系此前尚未得到研究：

引理 20。对于任意可 *UCQ* 重写的规则集 $\mathcal{S}$，我们有 $\mathsf{reify}(\mathcal{S})$ 也是可 *UCQ* 重写的。

证明。给定一个元数大于二、签名为 $\mathcal{S}$ 的符号 A，令 $\rho_{\mathtt{A}}$ be 为一条规则

$$\mathtt{A}(x_1, \ldots, x_n) \quad \rightarrow \quad \exists z \bigwedge_{\mathtt{A}_i \in \mathsf{reify}(\mathtt{A})} \mathtt{A}_i(x_i, z)$$

即 $\rho_{\mathtt{A}}$ 是一条规则，它只是将 A 投影 到其重化变体。令 $\mathcal{S}'$ be 为 $\mathcal{S}$，并为 $\mathcal{S}$ 的所有相关签名 A 添加规则 $\rho_{\mathtt{A}}$。注意，$\mathcal{S}'$ 是可 UCQ 重写的——因为 $\mathcal{S}$ 是可 UCQ 重写的，而且新增规则无法递归触发，也无法触发来自 $\mathcal{S}$的规则。

对于任意实例 $\mathcal{J}$，将 $Ch(\mathcal{J}, \mathcal{S}')$ 限制到再化签名 reify(Ξ)，所得结果与 reify($Ch(\mathcal{J}, \mathcal{R})$) 同构。根据定理 19，它也与$Ch$(reify($\mathcal{J}$),reify($\mathcal{R}$)) 同构。由此可知，若取任意一个关于 reify(Ξ) 的合取查询 Q，并令 Φ 为其相对于 $\mathcal{S}'$ 的重写，则 reify(Φ) 就是 Q 相对于 reify($\mathcal{S}$) 的重写 —— 这意味着 reify($\mathcal{S}$) 是可 UCQ 重写的。 □

证伪结论：(B)

利用定理 19 可知，reify($Ch(\mathcal{R})$) 与$Ch(\mathcal{R}')$ 同态等价；因此，若结论 (N) 不成立，则 (B) 也不成立。

证明定理 18 的假设

由于 $\mathbb{S}'$ 是二元的，且我们有定理 20，因此我们已证明定理 18 的假设对于 $\mathcal{R}'$ 和签名 $\mathbb{S}'$ 成立，证明至此结束。

4.3 精简规则头

现在我们引入两个定义，只要追赶过程中仅应用非 Datalog 规则，它们就能为我们提供一种方便的结构：为原子赋予方向，并使每个项都由给定谓词的原子唯一确定一个前驱。

定义 21（前向存在）。 规则 ρ 是前向存在的，当且仅当 对于它的每个头部原子 $A(x,\ y)$，都有 x 是边界变量，且 y 是存在变量。规则集是前向存在的 当且仅当 其中每条非 Datalog 规则都是前向存在的。

定义 22（谓词唯一）。 二元签名 $\mathbb{S}$ 上的规则集 $\mathcal{R}$ 是谓词唯一的，当且仅当对于 $\mathcal{R}$ 中的每个非 Datalog 规则 ρ，每个谓词 $\mathtt{E} \in \mathbb{S}$ 在 ρ 的规则头中最多出现一次。

请注意，上述条件并不等同于单规则头规则集，因为

$$\mathtt{A}(x), \mathtt{B}(y) \ \rightarrow \ \exists z \ \ \mathtt{D}(x,z), \mathtt{E}(y,z)$$

是一个谓词唯一的前向存在规则。

引理 23。 对于每个二元签名 $\mathbb{S}$ 以及每个可 *UCQ* 重写、前向存在且谓词唯一的规则集 $\mathcal{R}$ （其定义在 $\mathbb{S}$ 上），我们有：

$$Ch(\mathcal{R}) \models \textsc{Tournaments}_{\mathtt{E}} \quad \Rightarrow \quad Ch(\mathcal{R}) \models \textsc{Loop}_{\mathtt{E}}. \qquad (♖)$$

现在我们证明定理 18 与定理 23 等价。(⇒) 方向显然成立；对于 (⇐)，我们通过逆否命题来证明。设 $\mathbb{S}$ 是一个签名，$\mathcal{R}$ 是一个规则集，二者共同构成了定理 18 的一个反例。

规则集手术

设 A 和 B（带有指标）是新引入的关系符号。给定一条规则

$$\rho = B(\bar{x}, \bar{y}) \ \rightarrow \ \exists \bar{z} \ \ H(\bar{y}, \bar{z})$$

我们定义：

- 存在规则$\rho_{\text{初始}}$ 为

$$B(\bar{x},\bar{y}) \;\;\rightarrow\;\; \exists w \;\; \mathtt{A}_0^{\rho}(w) \wedge \bigwedge_{y\in\bar{y}} \mathtt{A}_{y,w}^{\rho}(y,w);$$

- 存在规则 $\rho_{\exists}$ 为

$$\mathtt{A}_0^{\rho}(w) \wedge \bigwedge_{y\in\bar{y}} \mathtt{A}_{y,w}^{\rho}(y,w) \;\;\rightarrow\;\; \exists\bar{z} \bigwedge_{y'\in\bar{y}\cup\{w\}} \bigwedge_{z\in\bar{z}} \mathtt{B}_{y',z}^{\rho}(y',z);$$

- Datalog 规则 ρ_{DL} 为

$$\bigwedge_{y'\in\bar{y}\cup\{w\}} \bigwedge_{z\in\bar{z}} \mathtt{B}_{y',z}^{\rho}(y',z) \;\;\rightarrow\;\; H(\bar{y},\bar{z}).$$

给定规则集 $\mathcal{S}$ 令 $\bigtriangledown(\mathcal{S})$ 为包含 ρ_{init}、$\rho_{\exists}$ 以及 ρ_{DL} 中所有规则的规则集。现在令 $\mathcal{R}' = \bigtriangledown(\mathcal{R})$，并令 $\mathbb{S}'$ 包含所有出现在 $\mathcal{R}'$中的谓词。我们证明 $\mathcal{R}'$ 和 $\mathbb{S}'$ 构成了定理 23 的一个反例。

精简的主要性质

由于 $\bigtriangledown(\mathcal{S})$ 的规则只是为追赶引入中间步骤——请注意，规则之间没有产生不必要交互的余地——我们得到如下结论：

引理 24. 对于每个规则签名 $\mathbb{S}$、规则集 $\mathcal{S}$ 以及关于 $\mathbb{S}$的实例 $\mathcal{J}$，当限制在 $\mathbb{S}$ 上时，$Ch(\mathcal{J},\mathcal{S})$ 与 $Ch(\mathcal{J},\bigtriangledown(\mathcal{S}))$ 同态等价。

证明。 证明可在 A 节中找到。 □

引理 25。 给定规则集 $\mathcal{S}$，我们有 $\bigtriangledown(\mathcal{S})$ 是前向存在且谓词唯一的。此外，如果 $\mathcal{S}$ 是可 *UCQ* 重写的，那么 $\bigtriangledown(\mathcal{S})$ 也是可 *UCQ* 重写的。

证明。 证明可在 B 节中找到 □

证伪结论：(R)

由定理 24 可知，将 $\{\top\}$ 和 $\mathcal{R}$ 的追赶限制在签名 $\mathbb{S}$ 上时，与将 $\{\top\}$ 和$\mathcal{R}'$ 的追赶限制在该签名上时所得的结果同态等价。因此，如果结论 (B) 不成立，那么 (R) 也不成立。

证明定理 23 的假设

鉴于 $\mathcal{R}$ 可 UCQ 重写，根据定理25，我们有 $\mathcal{R}'$ 满足要求。

4.4 主体重写

本节将使用受 [25]启发的 [26],中的快速性。

定义26 (快速)。 规则集 $\mathcal{R}'$ 是快速的，当且仅当 对于每个实例 $\mathcal{I}$ 和 $Ch(\mathcal{I},\mathcal{R}')$ 的每个原子 β，若 β 的所有边界项都出现在 adom($\mathcal{I}$) 中，则$\beta \in Ch_1(\mathcal{I},\mathcal{R}')$。

定义27 (Regal)。 二元签名上的规则集 $\mathcal{R}$ 是*Regal*，当且仅当 它可 UCQ 重写、快速、前向存在且谓词唯一。

现在我们可以陈述定理28，这正是我们在第5节中证明的结论。

定理28。 对于每个二元签名 $\mathbb{S}$ 以及 $\mathbb{S}$ 上的每个 *Regal* 规则集 $\mathcal{R}$，我们有：

$$Ch(\mathcal{R}) \models \textsc{Tournaments}_{\mathtt{E}} \quad \Rightarrow \quad Ch(\mathcal{R}) \models \textsc{Loop}_{\mathtt{E}}. \qquad (\bowtie)$$

现在我们证明定理 23 与定理28 等价。($\Rightarrow$) 方向显然成立，我们通过逆否命题证明 ($\Leftarrow$)。设 $\mathbb{S}$ 是一个签名， $\mathcal{R}$ 是一个规则集，它们共同构成定理 23 的一个反例。

规则集手术

接下来 [26]考虑如下定义：

定义29 （主体重写）。 给定规则集 $\mathcal{S}$ 以及形式为 $B(\bar{x},\bar{y}) \rightarrow \exists \bar{z}\ H(\bar{y},\bar{z})$ 的存在规则$\rho \in \mathcal{S}$，令 rew($\rho,\mathcal{S}$) 为如下规则集：

$$\Big\{\ q(\bar{x}',\bar{y}') \ \rightarrow\ \exists \bar{z}\ H(\bar{y}',\bar{z}) \ \Big|\ \exists \bar{x}'\ q(\bar{x}',\bar{y}') \in \mathtt{rew}(\exists \bar{x}\ B(\bar{x},\bar{y}),\mathcal{S})\ \Big\}.$$

最后，令 rew($\mathcal{S}$) $= \mathcal{S} \cup \bigcup_{\rho\in\mathcal{S}}$ rew($\rho,\mathcal{S}$)。

令 $\mathcal{R}'$ 为 rew($\mathcal{R}$)。现在我们证明， $\mathcal{R}'$ 和 $\mathbb{S}$构成定理28的一个反例。

主体重写的主要性质

我们将以下内容（引理 42, [26]）作为手术的第一个性质：

引理 30 ([26])。 对于任意有界推导深度规则集 $\mathcal{S}$和任意实例 $\mathcal{J}$，都有 $Ch(\mathcal{J}、\mathcal{S}) \leftrightarrow Ch(\mathcal{J}、\mathtt{rew}(\mathcal{S}))$。

作为引理 30 的一个简单结论，我们有以下结果。

引理 31. `rew` 规则集手术保留规则集的可 *UCQ* 重写、谓词唯一性和前向存在性。

证明. 设 $\mathcal{S}$ 为一个满足以下条件的规则集:

- 可 *UCQ* 重写：由于 $\mathcal{S}$ 和 rew($\mathcal{S}$) 的追赶在同态意义下等价（定理 30），因此，针对 $\mathcal{S}$ 的重写可以直接用作针对 rew($\mathcal{S}$) 的重写。

- 谓词唯一或前向存在：注意，与 $\mathcal{S}$ 相比，rew($\mathcal{S}$) 中规则的规则头保持不变——除了可能对前沿变量进行合并。因此，由于 $\mathcal{S}$ 具有谓词唯一性和前向存在性，rew($\mathcal{S}$) 也具有这两种性质。

□

此外，我们采用以下结论（引理 28, [26]）

引理 32 ([26])。 对于任意 *UCQ* 可重写规则集 $\mathcal{S}$，$\texttt{rew}(\mathcal{S})$ 是快速的。

4.4.1 规则性的主要性质

本节列出我们在下一节中会用到的规则性规则集的所有相关性质。设 $\mathcal{S}$ 为一个规则性规则集，$\mathcal{S}^{DL}$ 和 $\mathcal{S}^{\exists}$ 分别为其中 Datalog 规则和非 Datalog 规则的子集。首先注意到，由于 $\mathcal{S}$ 的快速性，可以仅在 Ch($\mathcal{S}^{\exists}$) 之上应用 Datalog 规则来得到 Ch($\mathcal{S}$)。

引理 33. $Ch(\mathcal{S})$ 与 $Ch(Ch(\mathcal{S}^{\exists})\ \mathcal{S}^{DL})$ 同态等价。

证明。从 Ch(Ch($\mathcal{S}^{\exists}$) $\mathcal{S}^{DL}$) 到 Ch($\mathcal{S}$) 的同态显然存在，因为存在从 Ch($\mathcal{S}^{\exists}$) 到 Ch($\mathcal{S}$) 的同态（因为 $\mathcal{S}^{\exists} \subseteq \mathcal{S}$），且$\mathcal{S}^{DL} \subseteq \mathcal{S}$。反过来，我们通过对 i 归纳，证明存在一个 h_i （若 $i > 0$，则扩展 h_{i-1}），将 Ch_i($\mathcal{S}$) 的项映射到 Ch_i($\mathcal{S}^{\exists}$) 的项，使得 h_i是从 Ch_i($\mathcal{S}$) 到 Ch(Ch_i($\mathcal{S}^{\exists}$) $\mathcal{S}^{DL}$) 的同态。对于 $i = 0$，结论显然成立。若结论对 $i \geq 0$成立，则考虑一个属于 Ch_{i+1}($\mathcal{S}$) 但不属于 Ch_i($\mathcal{S}$) 的项 t。它必定是由某个非 Datalog 规则 ρ 创建的，该规则的规则体通过 h 映射到 Ch_i($\mathcal{S}$)。因此，$\mathrm{h}_i \circ$h 是从 ρ 的规则体到 Ch(Ch_i($\mathcal{S}^{\exists}$) $\mathcal{S}^{DL}$) 的同态，并且所创建原子的所有前沿项都属于 Ch_i($\mathcal{S}^{\exists}$)。由于 $\mathcal{S}^{DL} \cup \mathcal{S}^{\exists}$具有快速性，这些原子可以通过一次规则 ρ的应用创建；该规则必定属于 $\mathcal{S}^{\exists}$，因为它是非 Datalog 规则。我们通过将 t映射到应用 ρ所创建的对应项，扩展 h_i。对于 Datalog 规则应用，设 $\langle \rho, \mathrm{h} \rangle$是在步骤 $i + 1$应用的一个触发器。 π(body(ρ)) $\subseteq Ch_i$($\mathcal{S}$)，因此 $\langle \rho, \mathrm{h}_i \circ \pi \rangle$ 是一个可应用于Ch(Ch_i($\mathcal{S}^{\exists}$) $\mathcal{S}^{DL}$) 的触发器，其结果因而属于 Ch(Ch_i($\mathcal{S}^{\exists}$) $\mathcal{S}^{DL}$)。

□

然而，Ch($\mathcal{S}^{\exists}$) 的结构比 Ch($\mathcal{S}$) 好得多，我们很快就会利用这一点。为了描述这种结构，我们为每个项关联一个时间戳。

定义34。 给定 Ch($\mathcal{S}^{\exists}$) 的一个项 t，我们将 t 的时间戳 TS(t)定义为满足 $t \in$adom(Ch_i($\mathcal{S}^{\exists}$)) 的最小整数 i。给定 Ch($\mathcal{S}^{\exists}$) 的一个项集T，我们将 TS$_{\mathsf{m}}$(T) 定义为多重集 $\{$ TS(t) $\mid t \in T \}_{\mathsf{m}}$。

鉴于 $\mathcal{S}$ 的签名是二元的，我们可以将实例（以及查询）视为有向图。

观察 35。 $Ch(\mathcal{S}^{\exists})$ 是一个有向无环图。

证明。由于 $\mathcal{S}$ 是前向存在的（且数据库不包含任何项），对于 $Ch(\mathcal{S}^{\exists})$ 中的所有原子 A(t, u)，我们都有 TS(t) $<$ TS(u)。因此，如果某个项 t 是某个环的一部分，则可得到 TS(t) <TS(t)，这与矛盾相冲突。 □

证伪结论：（）

根据定理 30，$\{\top\}$ 和 $\mathcal{R}$ 的追赶与 $\{\top\}$和 $\mathcal{R}'$ 的追赶是同态等价的。因此，如果结论 (B) 不成立，那么 (R) 也不成立。

证明定理 28 的假设

已知 rew 保留规则集的可 UCQ 重写、谓词唯一和前向存在性质（定理 31），且 $\mathcal{R}'$ 是快速的（定理 32），我们知道 $\mathcal{R}'$ 是 Regal。

5 证明定理28定理28

固定一个规则性规则集 $\mathcal{R}$ 。令 $\mathcal{R}^{\exists}$ 定义其非 Datalog 部分，并令 $\mathcal{R}^{DL}$表示 $\mathcal{R}$ 的 Datalog 规则集。我们将通过反证法证明定理28。假设 $\textsc{Tournaments}_{\textsc{E}}$ 在 $Ch(\mathcal{R}$)中成立，但 $\textsc{Loop}_{\textsc{E}}$ 不成立。根据定理 33， $\textsc{Tournaments}_{\textsc{E}}$ 在 $Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$ 中也成立，但 $\textsc{Loop}_{\textsc{E}}$ 不成立。对于 $n \in \mathbb{N}$，令 K_n为 $Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$ 中一个大小为 n 的 E-锦标赛。

证明思路如下。由于 $\mathcal{R}$ 具有规则性，对于 K_n 的每个 E(s, t)，E 相对于 $\mathcal{R}$ 的单射重写 Q 在 $Ch(\mathcal{R}^{\exists})$ 中对 s 和 t 成立。在第 5.1节中，我们将证明，任意大的竞赛图都可以由 Q 的一个特定形状的析取项见证（称为谷查询）。最后，在第 5.2 节中，我们将证明，单个谷查询在不产生自环的情况下只能生成大小为 3 的团，从而完成证明。

5.1 山谷查询

令 Q 为针对 $\mathcal{R}$ 对 E(x, y) 的单射重写。根据定理 6，且由于 $\mathcal{R}$ 可 UCQ 重写，该重写存在。

定义 36。 给定 s 和 t，使得 E(s, t) $\in Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$，我们将 $Ch(\mathcal{R}^{\exists})$ 中 E(s, t) 的见证 集合定义为 $\mathcal{W}(s, t) = \{\, q \in Q \mid Ch(\mathcal{R}^{\exists}) \models_{\text{inj}} q(s, t) \,\}$。

观察 37。 对于所有满足E(s, t) $\in Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$的 s 和 t，集合$\mathcal{W}(s, t)$ 非空。

证明。 由定理 6 得出，并且事实是 Q 是针对 $\mathcal{R}$ 对 E(x, y) 的单射重写，其中包含 $\mathcal{R}^{DL}$。 □

$Ch(\mathcal{R}^{\exists})$ 的无环性使我们能够在 $\mathcal{W}(s,t)$ 中找到具有特殊形状的查询。

定义 38. 给定一个作为有向无环图的实例 $\mathcal{I}$（或合取查询），我们定义项 $\mathcal{I}$ 上的严格偏序 $<_{\mathcal{I}}$ 如下：给定 $\mathcal{I}$ 中的两个项 s、t，当且仅当存在一条从 s 到 t 的有向路径时，$s <_{\mathcal{I}} t$ 成立。我们将定义为的自反闭包。我们用 和 分别作为)和 $\leq_{Ch(\mathcal{R}^{\exists})}$ 的简写。

定义 39. 谷查询 $q(x, y)$ 是一个二元合取查询，它是一个有向无环图，且其仅有的 $<_q$-极大变量是 x 和 y。

引理 40 (去峰论证)。对于所有满足 $\mathtt{E}(s,t) \in Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$ 的 s 和 t，集合 $\mathcal{W}(s,t)$ 包含一个谷查询。

证明。反设结论不成立。考虑一个非谷查询 $q \in \mathcal{W}(s, t)$，以及一个从 q 到 $Ch(\mathcal{R}^{\exists})$ 的单射同态 h，使得 $\mathsf{h}(x) = s$、$\mathsf{h}(y) = t$，并且 $\mathsf{TS_m}(\mathsf{h}(q))$ 在所有此类查询和同态中是 $\leq_{\text{lex}}$-极小的。根据定理 37 以及定理 8，鉴于 $\leq_{\text{lex}}$ 在大小至多为 $\max\{\,|q'| \mid q' \in Q\,\}$ 的多重集上是良基的，这样的配对确实存在。

接下来，考虑变量 $z \in \exists\mathsf{vars}(q)$，它是 $\leq_q$-极大的。这样的变量存在，因为 q 不是谷查询。由于 h 是单射，$\mathsf{h}(z)$ 在 $\mathsf{h}(q)$ 中没有任何出边。记 Z 为 q 中包含变量 z 的原子集合，并令 $\langle \rho,\ \pi \rangle$ 为创建 $\mathsf{h}(z)$ 的触发器。由于 $\mathcal{R}$ 是前向存在量化的，且 z 是 $\leq_q$-极大的，因此 $\mathsf{h}(Z) \subseteq \pi(\mathsf{head}(\rho))$。于是我们定义 $\mathcal{I} = \mathsf{h}(q) \setminus \mathsf{h}(Z) \cup \pi(\mathsf{body}(\rho))$。注意，根据此定义，$\mathcal{I} \subseteq Ch(\mathcal{R}^{\exists})$。

触发器 $\langle \rho,\ \pi \rangle$ 可应用于 $\mathcal{I}$，并产生 $\pi(\mathsf{head}(\rho))$，其中包含 $\mathsf{h}(Z)$（至多相差 $\mathsf{h}(z)$ 的重命名）。因此，由于 z 在 q 中是存在量化的，我们有 $Ch(\mathcal{I},\ \mathcal{R}) \models q(s,\ t)$。于是，$Ch(\mathcal{I},\ \mathcal{R}) \models \mathsf{E}(s,\ t)$；又由于 $Q(x,\ y)$ 是 $\mathsf{E}(x,\ y)$ 的单射重写，因此可推出，对于某个 $q' \in Q$，$\mathcal{I} \models_{\text{inj}} q'(s,\ t)$。考虑见证这一事实的单射同态 h'。

首先注意到，由于 $h'(q') \subseteq \mathcal{I}$，我们有 $\mathsf{TS_m}(h'(q')) \leq_{\text{lex}} \mathsf{TS_m}(\mathcal{I})$。接着，由于 $\mathsf{h}(z)$ 是由触发器 $\langle \rho,\ \pi \rangle$ 创建的，根据追赶的定义，对于所有满足 $x \in \pi(\mathsf{body}(\rho))$ 的情况，我们有 $\mathsf{TS}(x) <\mathsf{TS}(\mathsf{h}(z))$。因此，$\mathsf{TS_m}(I) <_{\text{lex}} \mathsf{TS_m}(\mathsf{h}(q))$，这意味着 $\mathsf{TS_m}(h'(q')) <_{\text{lex}} \mathsf{TS_m}(\mathsf{h}(q))$，与 $\langle q,\ h \rangle$ 的极小性矛盾。 □

命题41。如果 $Ch(\mathcal{R}\ )$ 中存在任意规模的 E-竞赛图，则存在由 $Ch(\mathcal{R}^{\exists})$ 上单个谷查询定义的任意规模竞赛图。

证明。首先，如前所述，根据定理33，如果 $Ch(\mathcal{R}\ )$ 中存在任意规模的 E-竞赛图，那么 $Ch(Ch(\mathcal{R}^{\exists}\ \mathcal{R}^{DL})$ 中也存在任意规模的 E-竞赛图。接下来我们运用拉姆齐定理。

对于 $Ch(Ch(\mathcal{R}^{\exists})\ \mathcal{R}^{DL})$ 中给定的 E-锦标赛，我们用 $\mathcal{W}(s,t)$ 中任意一个谷查询为每条边 $E(s,t)$ 着色。这是定理40的结果。拉姆齐定理保证，对于任意 n，如果原竞赛图的规模大于 $R(n,\ldots,n)$（参数个数与 Q 中的元素个数相同），那么就存在一个规模为 n 的竞赛图，它由单个谷查询定义。 □

5.2 由单个谷查询定义的锦标赛

现在我们来分析由单个谷查询定义的团。这里的主要工具是： $Ch(\mathcal{R}^\exists)$ 中谷查询的像完全由其答案变量的像决定。谷查询之所以具有这一强性质，是因为 $\mathcal{R}$ 同时满足谓词唯一性和前向存在性。

引理 42。 设 $q(x, \bar{y})$ 是一个合取查询，使得 $y <_q x$ 对于每个 $y \in \bar{y}$，那么集合

$$\{ \langle s, \bar{t} \rangle \in \mathsf{adom}(Ch(\mathcal{R}^\exists_\bowtie)) \times \mathsf{adom}(Ch(\mathcal{R}^\exists_\bowtie))^{|\bar{y}|} \quad | \quad Ch(\mathcal{R}^\exists_\bowtie) \models q(s, \bar{t}) \}$$

是一个函数。

证明。我们首先证明 $|\bar{y}| = 1$ （且 y 是 $\bar{y}$ 中唯一变量）时的情形，并且 $q(x, y)$ 是一条从 y 到 x的路径。我们根据 $q(x, y)$ 中从 y 到 x 的路径长度进行归纳，证明对于所有 $s \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，至多存在一个 $t \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，使得 $Ch(\mathcal{R}^\exists) \models q(s, t)$。当该路径长度为 0 时，结论显然成立，因此假设其长度至少为 1。于是，$q(x, y) = q'(x, z) \wedge \mathsf{A}(y, z)$，其中 A 是 $\mathbb{S}$ 中的一个谓词，而 $q'(x, z)$ 是长度为 $|q(x, y)| - 1$ 的路径查询。因此，根据归纳假设，对于所有 $s \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，至多存在一个 $u \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，使得 $Ch(\mathcal{R}^\exists) \models q'(s, u)$。

接下来，假设存在一些项 s、 t 和 t'，使得 $Ch(\mathcal{R}^\exists) \models q(s, t) \wedge q(s, t')$。那么，存在唯一的 u，使得 $Ch(\mathcal{R}^\exists) \models q'(s, u)$，因此原子 $\mathsf{A}(t, u)$ 和 $\mathsf{A}(t', u)$ 都出现在 $Ch(\mathcal{R}^\exists)$ 中。令 $\langle \rho, \mathsf{h} \rangle$ 为在 $Ch(\mathcal{R}^\exists)$ 中引入 u 的触发器。由于 $\mathcal{R}^\exists$ 是前向存在的， t 和 t' 都是 ρ 的边界项的像。然而，由于 $\mathcal{R}^\exists$ 具有边界函数性，$\mathsf{head}(\rho)$ 中至多有一个 A 原子。因此，必有 $t = t'$，由此归纳法得证。

完整结论随即由这一情形推出。事实上，对于所有 $y \in \bar{y}$，考虑 $q(x, \bar{y})$ 中从 y 到 x 的一条路径。这条路径定义了一个路径查询 $q'_y(x, y)$，因此对于所有 $s \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，至多存在一个 $t \in \mathsf{adom}(Ch(\mathcal{R}^\exists))$，使得 $Ch(\mathcal{R}^\exists) \models q'_y(s, t)$。由此，利用如下事实即可得到完整结论：对于所有 s 和 $\bar{t}$， $Ch(\mathcal{R}^\exists) \models q(s, \bar{t})$ 蕴含 $Ch(\mathcal{R}^\exists) \models q'_y(s, t)$。 □

命题43. 如果一个谷查询定义了一个大小为4的E-竞赛图，那么它也定义了一个E-自环。

证明. 设 q 是一个谷查询。

- 如果 $q(x, y)$不连通，那么 $q(x, y) = q_1(x) \wedge q_2(y) \wedge q_3$，其中 q_i 和 q_j的变量集合不相交，条件是 $i \neq j$。由于 k_1 与k_2之间存在一条边，且二者具有对称性，我们不妨设$Ch(\mathcal{R}^\exists) \models q_1(k_1) \wedge q_2(k_2)$。 k_3与 k_4之间也存在一条边，因此我们再次不妨设 $Ch(\mathcal{R}^\exists) \models q_1(k_3) \wedge q_2(k_4)$。最后， k_1 与 k_3之间存在一条边，因此要么 $Ch(\mathcal{R}^\exists) \models q_2(k_1)$，要么 $Ch(\mathcal{R}^\exists) \models q_2(k_3)$。在这两种情况下， $Ch(\mathcal{R}^\exists) \models \exists u\ q_1(u) \wedge q_2(u) \wedge q_3$，所以$Ch(\mathcal{R}^\exists) \models \exists u\ \mathsf{E}(u, u)$，从而定义了一个 E-环。因此，在其他所有情况下，我们都假设 q 是连通的。

- 如果 $q(x,y)$包含一个单独的 $\leq_q$-极大顶点，例如 x （因为 y 是 $\leq_q$-极大的情形具有对称性）。因此， $y < x$，所以根据定理 42, 集合 $\{\langle s,t\rangle \in$ 活动论域$(Ch(\mathcal{R}^\exists))^2 \mid Ch(\mathcal{R}^\exists) \models q(s,t)\}$ 是一个函数。因此， K 中的每个顶点的出度至多为一，这与 K 是一个大小为 4 的竞赛图这一事实相矛盾。

- 如果 x 和 y 在 $q(x,y)$ 中都是 $\leq_q$-极大的，那么令 $\bar{v}$ 为 q 中小于 x 和 y （针对 $\leq_q$）的存在变量组成的元组。于是， $q(x,y) = \exists \bar{v}\ q_x(x,\bar{v}) \wedge q_y(\bar{v},y)$, 其中 q_x （分别是 q_y）包含所有小于 x （分别是 y）的变量。根据定理42, 集合 $\{\langle x,\bar{v}\rangle \mid Ch(\mathcal{R}^\exists) \models q_x(x,\bar{v})\}$ 和 $\{\langle y,\bar{v}\rangle \mid Ch(\mathcal{R}^\exists) \models q_y(\bar{v},y)\}$ 都是函数，我们分别记为 f_x 和 f_y。由于 K 是一个含有4个元素的竞赛图，除元素重命名外，它包含以下集合： $\{\mathrm{E}(k_1,k_2), \mathrm{E}(k_1,k_3), \mathrm{E}(k_2,k_3)\}$。

  接下来我们证明 $Ch(\mathcal{R}) \not\models \mathrm{E}(k_2,k_2)$。首先，注意到如果 $q(s,t) \in Ch(\mathcal{R}^\exists)$, 那么 $f_x(s) = f_y(t)$。因此，我们得到以下三个等式： $f_x(k_1) = f_y(k_2)$、 $f_x(k_1) = f_y(k_3)$ 和 $f_x(k_2) = f_y(k_3)$。于是，复合这些等式可得 $f_x(k_2) = f_y(k_2)$，因此 $Ch(\mathcal{R}^\exists) \models q_x(k_2, f_x(k_2)) \wedge q_y(f_x(k_2)\ k_2)$。由此， $Ch(\mathcal{R}^\exists) \models q(k_2,k_2)$，且 $Ch(\mathcal{R}) \not\models \mathrm{E}(k_2,k_2)$，证明完毕。 □

6 讨论

竞赛图的定义

虽然主要结果针对固定关系 E 上的锦标赛，但我们的结果可以直接推广到任何可由二元合取查询并定义的关系。考虑任意二元合取查询并 $Q(x,y) = \bigvee_{i=1}^{k} q_i(x,y)$，其中每个 q_i 都是一个合取查询。只需对每个 i，将以下规则加入所考虑的规则集：

$$q_i(x,y) \to \mathrm{E}(x,y).$$

当 E 是一个新的谓词符号时，添加这一项不会影响该规则集的 UCQ-可重写性。

任意可着色性

这一研究方向的自然下一步，是证明对于 UCQ 可重写规则集而言，不可能定义色数任意高的结构而不蕴含以下 $\textsc{Loop}_{\mathrm{E}}$ 查询：

猜想 44。 对于每个签名 $\mathbb{S}$、每个关于$\mathbb{S}$的 *UCQ* 可重写规则集 $\mathcal{R}$ 以及每个实例 $\mathcal{I}$, 我们都有：

$Ch(\mathcal{I},\ \mathcal{R})|_{\mathrm{E}}$ *cannot be colored with a finite number of colors* $\Rightarrow Ch(\mathcal{I},\ \mathcal{R}) \models \textsc{Loop}_{\mathrm{E}}$.

我们相信，如果这一猜想得以证明，将构成 UCQ 可重写规则集的一项优雅的模型论性质，并代表着朝证明（有界推导深度 $\Rightarrow$fc）猜想迈出的重要一步。下面具体说明。与本文给出的结果类似，如果定理44成立，就能排除反例空间中的很大一部分：任何无法用有限数量的颜色着色的结构，都不能同态嵌入到一个不蕴含 $\textsc{Loop}_{\mathrm{E}}$ 的有限结构中。

我们希望，本文中发展出的工具能为定理44的证明奠定基础。不过，我们注意到，这样的证明并不能直接通过扩展第5节中的技术得到。目前的证明旨在说明，存在一个由单个谷查询见证的四元竞赛图。然而，我们知道，存在色数任意高且不含四团的结构：

定理45 (埃尔″什 [12])。 存在围长和色数都任意高的图。

锦标赛规模界

另一个独立但相关的研究方向，是探讨以下问题：

问题46。 给定一个可 *UCQ* 重写规则集 $\mathcal{R}$ ，使得 $Ch(\{\top\},\mathcal{R}) \not\models \textsc{Loop}_{\mathrm{E}}$，最大的 n 是多少，使得 $Ch(\mathcal{I},\mathcal{R})$ 容许一个规模为n的竞赛图？

细心的读者会注意到，前文给出的证明可以推出形式为 $N(4,\ldots,4)$（含有 $|Q|$ 个参数）的竞赛图最大规模的一个上界——如果追赶中存在一个至少达到此规模的竞赛图，就可以运用第 5.2 节的工具证明 $\textsc{Loop}_{\mathrm{E}}$ 可推出。

我们将定理46留作一个开放问题，因为它将关于存在规则的多个重要问题联系在一起，尤其是在 UCQ-可重写片段中。

致谢

Piotr Ostropolski-Nalewaja 获得了波兰国家科学中心（NCN）项目 2022/45/B/ST6/00457 的资助。

参考文献

[1] S. Abiteboul、R. Hull 和 V. Vianu。《数据库基础》。Addison-Wesley, 1995 年。

[2] J.-F. Baget, M. Lecl`ere, M.-L. Mugnier 和 E. Salvat. 关于含存在变量的规则：在可判定性边界上探索。《人工智能》，175(9):1620–1654，2011。

[3] P. Barcel´o, G. Berger, C. Lutz 和 A. Pieris. 前沿守卫本体中介查询的一阶可改写性。载于第*27*届国际人工智能联合会议（*IJCAI 2018*），*2018*年*7*月*13*日至*19*日，瑞典斯德哥尔摩。，第1707–1713页，2018年。[4] V. B´ar´any, G. Gottlob 和 M. Otto. 守卫片段查询。载于*2010*年第*25*届*IEEE*计算机科学逻辑年会，第1–10页，2010年。[5] A. Cal`ı, G. Gottlob 和 M. Kifer. 驯服无限追赶：在表达力强的关系约束下进行查询回答。载于第十一届知识表示与推理原理国际会议论文集，KR'08，第70–80页。AAAI出版社，2008年。[6]A. Cal`ı, G. Gottlob 和 M. Kifer. 驯服无限追赶：在表达力强的关系约束下进行查询回答。人工智能研究杂志，48(1):115–174，2013年10月。[7] A. Cal`ı, G. Gottlob 和 A. Pieris. 本体查询的高级处理。*VLDB*基金会论文集，3，2010年。[8] D. Carral, I. Dragoste 和 M. Kr¨otzsch. 含析取的存在规则的受限追赶（不）终止性。载于 C. Sierra 编，第*26*届国际人工智能联合会议（*IJCAI 2017*），澳大利亚，第922–928页。ijcai.org，2017年。[9] C. Civili 和 R. Rosati. 关于二元守卫存在规则上合取查询的一阶可改写性（扩展摘要）。2015年7月。[10] B. Cuenca Grau, I. Horrocks, M. Kr¨otzsch, C. Kupke, D. Magka, B. Motik 和 Z. Wang. 存在规则的无环性概念及其在本体查询回答中的应用。《人工智能研究杂志》（*JAIR*），47:741–808，2013年。[11] A. Deutsch, A. Nash 和 J. Remmel. 重新审视追赶。2008年。[12] P. Erd¨os. 图论与概率。加拿大数学杂志，11:34–38，1959年。[13] R. Fagin, P. G. Kolaitis, R. J. Miller 和 L. Popa. 数据交换：语义与查询回答。计算机科学讲义（包括人工智能讲义和生物信息学讲义子系列），2572，2003年。[14] T. Feller, T. S. Lyon, P. Ostropolski-Nalewaja 和 S. Rudolph. 存在规则的有限团宽集合：迈向可判定且高表达力查询的一般判据。载于 F. Geerts 和 B. Vandevoort 编，第*26*届数据库理论国际会议（*ICDT 2023*），第255卷，莱布尼茨国际信息学论文集（*LIPIcs*），第18:1–18:18页，德国达格施图尔，2023年。Schloss Dagstuhl – Leibniz-信息学中心。

[15] H. Gaifman、H. Mairson、Y. Sagiv与M. Y. Vardi。数据库逻辑程序的不可判定优化问题。《*ACM*期刊》，40(3):683–713，1993年7月。[16] L. Gerlach与D. Carral。析取Skolem追赶的一般无环性与环性概念。载于B. Williams、Y. Chen与J. Neville编，第三十七届*AAAI*人工智能会议论文集，*2023*，美国，第6372–6379页。AAAI出版社，2023年6月。[17] T. Gogacz与J. Marcinkowski。论BDD/FC猜想。载于第*32*届*ACM SIGMOD-SIGACT-SIGAI*数据库系统原理研讨会论文集，PODS '13，第127–138页，美国纽约州纽约市，2013年。计算机协会。[18] T. Gogacz与J. Marcinkowski。收敛于追赶——一种有限可控性工具。《计算机与系统科学杂志》，83(1):180–206，2017年。[19] G. Gottlob。Datalog+/-：本体与完整性约束的统一方法。载于V. D. Antonellis、S. Castano、B. Catania与G. Guerrini编，第*17*届意大利高级数据库系统研讨会论文集（*SEBD'09*），第5–6页。Edizioni Seneca，2009年。[20] G. Gottlob、S. Kikot、R. Kontchakov、V. Podolskii、T. Schwentick与M. Zakharyaschev。基于本体的数据访问中查询重写的代价。《人工智能》，213:42–59，2014年。[21] D. Johnson与A. Klug。在函数依赖与包含依赖下检验合取查询的包含性。《计算机与系统科学杂志》，28(1):167–189，1984年。[22] M. K¨onig、M. Lecl`ere、M. Mugnier与M. Thomazo。面向存在规则的可靠、完备且最小的并集合取查询重写。《语义网》，6(5):451–475，2015年。[23] M. Kr¨otzsch与S. Rudolph。通过结合无环性与守卫性扩展可判定存在规则。2011年。[24] D. Maier、A. O. Mendelzon与Y. Sagiv。检验数据依赖的蕴涵。《*ACM*数据库系统汇刊》（*TODS*），4，1979年。[25] P. Ostropolski-Nalewaja、J. Marcinkowski、D. Carral与S. Rudolph。通往查询可重写性前沿的旅程。载于L. Libkin与P. Bar- cel´o编，*PODS '22*：国际数据管理会议，美国宾夕法尼亚州费城，*2022*年*6*月*12*日至*17*日，第359–367页。美国计算机协会，2022年。[26] P. Ostropolski-Nalewaja与S. Rudolph。通向表达性查询的黏性路径：存在规则下导航查询的可判定性。载于第*21*届知识表示与推理原理国际会议论文集，第574–584页，2024年8月。

[27] R. Rosati。论开放世界假设下数据库中合取查询回答的有限可控性。计算机与系统科学杂志，77(3):572– 594，2011年。[28] E. Salvat与M.-L. Mugnier。图规则的可靠且完备的正向和反向链接。载于P. W. Eklund、G. Ellis与G. Mann编，第*4*届国际概念结构会议论文集（*ICCS'96*），计算机科学讲义第1115卷，第248–262页。施普林格，1996年。

引理 24 的证明

令 $\triangledown_{\text{init}}(\mathcal{S})$ 为 $\triangledown(\mathcal{S})$ 的子集，其中仅包含形如 ρ_{init} 的规则；令 $\triangledown_{\exists}(\mathcal{S})$ 和 $\triangledown_{\text{DL}}(\mathcal{S})$ 以类似方式定义。

观察 47。对于任意规则集 $\mathcal{S}$，$\triangledown(\mathcal{S})$ 的签名可以划分为三个互不相交的签名 $\mathbb{S}$、$\mathbb{S}_{\text{A}}$ 和 $\mathbb{S}_{\text{B}}$，使得

- $\mathbb{S}$ 是 $\triangledown_{\text{DL}}(\mathcal{S})$ 中规则头以及 $\triangledown_{\text{init}}(\mathcal{S})$ 中规则体的签名，
- $\mathbb{S}_{\text{A}}$ 是 $\triangledown_{\text{init}}(\mathcal{S})$ 中规则头以及 $\triangledown_{\exists}(\mathcal{S})$ 中规则体的签名，
- $\mathbb{S}_{\text{B}}$ 是 $\triangledown_{\exists}(\mathcal{S})$ 中规则头以及 $\triangledown_{\text{DL}}(\mathcal{S})$ 中规则体的签名，

引理 48。对于每个规则集 $\mathcal{S}$、每个实例 $\mathcal{J}$ 和每个整数 i，都存在一个同态 $\sigma_i^{\triangledown}$，将 $Ch_i(\mathcal{J},\mathcal{S})$ 映射到 $Ch_{3i}(\mathcal{J},\triangledown(\mathcal{S}))$。

证明。我们对 i 使用归纳法。对于基例，有 $Ch_0(\mathcal{J},\mathcal{S}) = Ch_0(\mathcal{J},\triangledown(\mathcal{S})) = \mathcal{J}$，因此 $\sigma_0^{\triangledown}$ 是恒等映射。现在，在归纳步骤中，假设已构造出 $\sigma_i^{\triangledown}$，并考虑 $Ch_i(\mathcal{J},\mathcal{S})$ 上的一个触发器 $\tau = \langle \rho, \mathsf{h} \rangle$。于是，$\sigma_i^{\triangledown}(\mathsf{h}(\text{规则体}(\rho_{\text{init}}))) = \sigma_i^{\triangledown}(\mathsf{h}(\text{规则体}(\rho)))$（因为这些规则体相等），因此触发器 $\langle \rho_{\text{init}}, \sigma_i^{\triangledown} \circ \mathsf{h} \rangle$ 是 $Ch_{3i}(\mathcal{J},\triangledown(\mathcal{S}))$ 上的一个触发器。由此，触发器 $\langle \rho_{\exists}, \sigma_i^{\triangledown} \circ \mathsf{h} \rangle$ 是 $Ch_{3i+1}(\mathcal{J},\triangledown(\mathcal{S}))$ 上的一个触发器，因为 $\rho_{\exists}$ 的规则体就是 ρ_{init} 的规则头。对于步骤 $3i+2$ 中的 $\langle \rho_{\text{DL}}, \sigma_i^{\triangledown} \circ \mathsf{h} \rangle$，同理可证。因此，$\sigma_i^{\triangledown}(\text{输出}(\tau))$ 与 $\text{输出}(\langle \rho_{\text{DL}}, \sigma_i^{\triangledown} \circ \mathsf{h} \rangle)$ 之间存在一个同构，所以我们可以利用这些同构扩展 $\sigma_i^{\triangledown}$，处理在 $Ch_{i+1}(\mathcal{J},\mathcal{S})$ 中步骤 $i+1$ 应用的所有触发器，并构造出一个同态 $\sigma_{i+1}^{\triangledown}$，将 $Ch_{i+1}(\mathcal{J},\mathcal{S})$ 映射到 $Ch_{3i+3}(\mathcal{J},\triangledown(\mathcal{S}))$。□

我们用 $\mathbb{S}$ 表示 $\mathcal{S}$ 的签名，并对每个实例 $\mathcal{J}$，定义 $\mathcal{J}$ 在签名 $\mathbb{S}$ 上的限制 $\mathcal{J}_{|\mathbb{S}}$。

观察 49。对于每个实例 $\mathcal{J}$、规则集 $\mathcal{S}$、规则 $\rho \in \mathcal{S}$ 和同态 h

- 若 $\langle \rho_{\exists}, \mathsf{h} \rangle$ 是一个触发器，且 $\mathsf{h}(\text{规则体}(\rho_{\exists}))$ 中的某个原子被映射到 $Ch_1(\mathcal{J},\triangledown_{\text{初始}}(\mathcal{S})) \setminus \mathcal{J}$，那么存在某个触发器 $\langle \rho_{\text{初始}}, \mathsf{h}' \rangle$，对于 $\mathcal{J}$，使得 $\text{输出}(\langle \rho_{\text{初始}}, \mathsf{h}' \rangle) = \mathsf{h}(\text{规则体}(\rho_{\exists}))$。
- 若 $\langle \rho_{\text{DL}}, \mathsf{h} \rangle$ 是一个触发器，且 $\mathsf{h}(\text{规则体}(\rho_{\text{DL}}))$ 中的某个原子被映射到 $Ch_1(\mathcal{J},\triangledown_{\exists}(\mathcal{S})) \setminus \mathcal{J}$，那么存在某个触发器 $\langle \rho_{\exists}, \mathsf{h}' \rangle$，对于 $\mathcal{J}$，使得 $\text{输出}(\langle \rho_{\exists}, \mathsf{h}' \rangle) = \mathsf{h}(\text{规则体}(\rho_{\text{DL}}))$。

证明。请注意，这些规则的规则体中的所有原子都包含一个项，而该项是前一条规则规则头中的存在变量。因此，必须通过规则应用生成这个变量。由于对于所有 $\rho \in \mathcal{S}$，ρ_{init} 的规则头都是 $\rho_{\exists}$ 的规则体，且 $\rho_{\exists}$ 和 ρ_{DL} 也同样如此，我们便得到了结论。□

引理 50。对于每个规则集 $\mathcal{S}$ 和实例 $\mathcal{J}$，设 $\mathcal{J}' = Ch_1(Ch_1(\mathcal{J}, \triangledown_\exists(\mathcal{S}))\ \triangledown_{\mathrm{DL}}(\mathcal{S}))|_{\mathbb{S}}$。对于所有 i，存在一个同态 $\pi_i^\triangledown$，从 $Ch_{3i}(\mathcal{J}, \triangledown(\mathcal{S}))$ 到

$$\mathcal{R}_i = Ch_1(Ch_1(Ch_i(\mathcal{J}', \mathcal{S}), \triangledown_{\mathrm{init}}(\mathcal{S})), \triangledown_\exists(\mathcal{S})) \cup Ch_1((\mathcal{J} \setminus \mathcal{J}_{|\mathbb{S}}), \triangledown_\exists(\mathcal{S}))$$

证明。我们通过对 i 进行归纳法来构造 $\pi_i^\triangledown$。对于基例，$Ch_0(\mathcal{J}, \triangledown(\mathcal{S})) = \mathcal{J}$ 包含在 $\mathcal{R}_0$ 中，因为 $\mathcal{J}_{|\mathbb{S}} \subseteq \mathcal{J}'$，所以 $\pi_0^\triangledown$ 是恒等映射。接下来，我们通过假设 $\pi_i^\triangledown$ 已构造来讨论归纳情形。在步骤 $3i+1$、$3i+2$ 或 $3i+3$，可以应用在 $Ch_{3i+3}(\mathcal{J}, \triangledown(\mathcal{S})) \setminus Ch_{3i}(\mathcal{J}, \triangledown(\mathcal{S}))$ 中生成原子的触发器，并使用 $\triangledown_{\mathrm{init}}(\mathcal{S})$、$\triangledown_\exists(\mathcal{S})$ 或 $\triangledown_{\mathrm{DL}}(\mathcal{S})$ 中的一条规则。设 $\tau = \langle r, \mathrm{h} \rangle$ 是这样的触发器。为了构造 $\pi_{i+1}^\triangledown$，我们展示在这九种情形下，如何将 $\pi_i^\triangledown(\mathrm{output}(\tau))$ 映射到 $\mathcal{R}_{i+1}$。

步骤 $3i+1$ 和 $r = \rho_{\mathrm{init}}$： 这些规则的规则体定义在 $\mathbb{S}$ 上，这意味着 $\langle r, \pi_i^\triangledown \circ \mathrm{h} \rangle$ 是一个作用于 $Ch_i(\mathcal{J}', \mathcal{S})$ 上的触发器。因此，$\pi_i^\triangledown(\mathrm{output}(\langle r, \mathrm{h} \rangle))$ 映射到 $\mathrm{output}(\langle r, \pi_i^\triangledown \circ \mathrm{h} \rangle) \subseteq Ch_1(Ch_i(\mathcal{J}', \mathcal{S})\ \triangledown_{\mathrm{init}}(\mathcal{S})) \subseteq \mathcal{R}_i$。

步骤 $3i+1$ 和 $r = \rho_\exists$：这些规则的规则体位于 $\mathbb{S}_{\mathtt{A}}$ 上，因此只能映射到 $\mathcal{J} \setminus \mathcal{J}_{|\mathbb{S}}$ 中的原子，或映射到 $Ch_1(Ch_i(\mathcal{J}', \mathcal{S})\ \triangledown_{\mathrm{init}}(\mathcal{S}))$ 中的原子。如果 $\pi_i^\triangledown(\mathrm{h}(\mathrm{body}(\rho_\exists)))$ 包含在 $\mathcal{J}$ 中，那么 $\pi_i^\triangledown(\mathrm{output}(\langle \rho_\exists, \mathrm{h} \rangle))$ 可以映射到 $Ch_1((\mathcal{J} \setminus \mathcal{J}_{|\mathbb{S}})\ \triangledown_\exists(\mathcal{S}))$。否则，根据定理 49，存在某个触发器 $\langle \rho_{\mathrm{init}}, \mathrm{h}' \rangle$ 使得 $\pi_i^\triangledown(\mathrm{h}(\mathrm{body}(\rho_\exists))) = \mathrm{output}(\langle \rho_{\mathrm{init}}, \mathrm{h}' \rangle)$。因此，$\pi_i^\triangledown(\mathrm{output}(\langle \rho_\exists, \mathrm{h} \rangle))$ 可以映射到 $\mathrm{output}(\langle \rho_\exists, \pi_i^\triangledown \circ \mathrm{h} \rangle) \subseteq \mathcal{R}_i$。

步骤 $3i+1$ 和 $r = \rho_{\mathrm{DL}}$：如果 $\pi_i^\triangledown(\mathrm{h}(\mathrm{body}(\rho_{\mathrm{DL}}))) \subseteq \mathcal{J} \setminus \mathcal{J}_{|\mathbb{S}}$，那么 $\pi_i^\triangledown(\mathrm{output}(\langle \rho_{\mathrm{DL}}, \mathrm{h} \rangle))$ 可以映射到 $\mathrm{output}(\langle \rho_{\mathrm{DL}}, \pi_i^\triangledown \circ \mathrm{h} \rangle) \subseteq \mathcal{J}' \subseteq \mathcal{R}_i$，因为 $\triangledown_{\mathrm{DL}}(\mathcal{S})$ 中的规则只会生成 $\mathbb{S}$ 上的原子。否则，通过两次应用定理 49，存在一些触发器 $\langle \rho_\exists、\ \mathrm{h}' \rangle$ over $Ch_1(Ch_i(\mathcal{J}', \mathcal{S})\ \triangledown_{\mathrm{init}}(\mathcal{S}))$ 和 $\langle \rho_{\mathrm{init}}、\ \mathrm{h}'' \rangle$over $Ch_i(\mathcal{J}', \mathcal{S})$，使得 $\mathrm{output}(\langle \rho_{\mathrm{init}}, \mathrm{h}'' \rangle) = \mathrm{h}'(\mathrm{body}(\rho_\exists))$，且 $\mathrm{output}(\langle \rho_\exists, \mathrm{h}' \rangle) = \pi_i^\triangledown(\mathrm{h}(\mathrm{body}(\rho_{\mathrm{DL}})))$。因此，触发器 $\langle \rho、\mathrm{h}'' \rangle$ 是 $Ch_i(\mathcal{J}', \mathcal{S})$ 上的一个触发器，使得 $\mathrm{output}(\langle \rho, \mathrm{h}'' \rangle)$ 和 $\mathrm{output}(\langle \rho_{\mathrm{DL}}, \pi_i^\triangledown \circ \mathrm{h} \rangle)$ 同构。因此，我们可以扩展这一同构，将集合 $\pi_i^\triangledown(\mathrm{output}(\langle \rho_{\mathrm{DL}}, \mathrm{h} \rangle))$ 映射到 $\mathrm{output}(\langle \rho, \mathrm{h}'' \rangle) \subseteq Ch_{i+1}(\mathcal{J}', \mathcal{S}) \subseteq \mathcal{R}_{i+1}$。

步骤 $3i+2$ 和 $r = \rho_{\mathrm{init}}$： $\mathrm{h}(\mathrm{body}(\rho_{\mathrm{init}}))$ 的某个原子必须通过 $\triangledown_{\mathrm{DL}}(\mathcal{S})$ 中的一条规则，在步骤 $3i+1$ 映射到新创建的原子上。对于所有这样的触发器 $\langle \rho'_{\mathrm{DL}}, \mathrm{h}' \rangle$，根据前一种情况，我们可以将 $\pi_i^\triangledown(\mathrm{output}(\langle \rho'_{\mathrm{DL}}, \mathrm{h}' \rangle))$ 映射到 $Ch_{i+1}(\mathcal{J}', \mathcal{S})$ 中。因此，根据定理 49，我们可以将 $\pi_i^\triangledown(\mathrm{h}(\mathrm{body}(\rho_{\mathrm{init}})))$ 完全映射到 $Ch_{i+1}(\mathcal{J}', \mathcal{S})$ 中，使得 $\langle \rho_{\mathrm{init}}, \pi_i^\triangledown \circ \mathrm{h} \rangle$ 是其触发器，从而 $\pi_i^\triangledown(\mathrm{output}(\langle \rho_{\mathrm{init}}, \mathrm{h} \rangle))$ 可以映射到 $\mathrm{output}(\langle \rho_{\mathrm{init}}, \pi_i^\triangledown \circ \mathrm{h} \rangle) \subseteq \mathcal{R}_{i+1}$。

步骤 $3i+2$ 和 $r \in \{\rho_\exists, \rho_{\mathrm{DL}}\}$： 这些情况的处理方式与步骤 $3i+1$ 完全相同。

步骤 $3i+3$ 和 $r \in \{\rho_{\mathrm{init}}, \rho_{\mathrm{DL}}\}$： 这些情况的处理方式与步骤 $3i+2$ 完全相同。

步骤 $3i+3$ 和 $r = \rho_\exists$：h(body($\rho_\exists$)) 中的某个原子必须通过 $\triangledown_{\text{init}}(\mathcal{S})$ 中的一条规则，在步骤 $3i+2$ 映射到新创建的原子。对于所有此类触发器$\langle\, \rho'_{\text{init}}, \mathsf{h}' \,\rangle$，根据前述情况，我们可以将 $\pi_i^{\triangledown}$(output($\langle\, \rho'_{\text{init}}, \mathsf{h}' \,\rangle$)) 映射到 $Ch_1(Ch_{i+1}(\mathcal{J}', \mathcal{S})\ \triangledown_{\text{init}}(\mathcal{S}))$。因此，根据定理 49，我们可以将 $\pi_i^{\triangledown}$(h(body($\rho_\exists$))) 完全映射到 $Ch_1(Ch_{i+1}(\mathcal{J}', \mathcal{S})\ \triangledown_{\text{init}}(\mathcal{S}))$，使得 $\langle\, \rho_\exists, \pi_i^{\triangledown} \circ \mathsf{h} \,\rangle$ 是其触发器，从而 $\pi_i^{\triangledown}$(output($\langle\, \rho_\exists, \mathsf{h} \,\rangle$)) 可以被映射到 output($\langle\, \rho_\exists, \pi_i^{\triangledown} \circ \mathsf{h} \,\rangle$) $\subseteq \mathcal{R}_{i+1}$。 □

引理 24 (重述)。对于每个规则签名 $\mathbb{S}$、规则集 $\mathcal{S}$ 以及其上的实例 $\mathcal{J}$，当限制在 $\mathcal{S}$ 的签名上时，$Ch(\mathcal{J}, \mathcal{S})$ 与 $Ch(\mathcal{J}, \triangledown(\mathcal{S}))$同态等价。

证明。注意到

$$\big(Ch_1(Ch_1(Ch_i(\mathcal{J}', \mathcal{S}), \triangledown_{\text{init}}(\mathcal{S})), \triangledown_\exists(\mathcal{S})) \cup (\mathcal{J} \setminus \mathcal{J}_{|\mathbb{S}})\big)_{|\mathbb{S}} = Ch_i(\mathcal{J}', \mathcal{S})$$

并且当 $\mathcal{J}$ 是关于 $\mathbb{S}$ 时，$\mathcal{J}' = \mathcal{J}$。 □

B 引理 25 的证明

引理 25 (重述)。给定规则集 $\mathcal{S}$，我们有 $\triangledown(\mathcal{S})$ 是前向存在且谓词唯一的。此外，如果 $\mathcal{S}$ 可 *UCQ* 重写，那么 $\triangledown(\mathcal{S})$ 也可 *UCQ* 重写。

根据 $\triangledown(\mathcal{S})$ 的定义可知，它是前向存在且谓词唯一的。现在我们证明它也可 UCQ 重写。令 $\triangledown_\exists(\mathcal{S})$ 为 $\{\, \rho_\exists \mid \rho \in \mathcal{S} \,\}$，且 $\triangledown_{\text{DL}}(\mathcal{S})$ 为 $\{\, \rho_{\text{DL}} \mid \rho \in \mathcal{S} \,\}$。对于 $\rho \in \mathcal{S}$，定义规则 ρ_{full} 如下：

$$\rho_{\text{full}} \quad = \quad \phi(\bar{x}, \bar{y}) \ \rightarrow\ \exists \bar{z}, w\ \ \psi(\bar{y}, \bar{z}) \ \ \wedge\ \ \mathtt{A}_0^\rho(w) \wedge \bigwedge_{y \in \bar{y}} \mathtt{A}^\rho_{y,w}(y, w) \ \ \wedge \ \ \bigwedge_{z \in \bar{z}} \mathtt{B}^\rho_{y',z}(y', z)$$

并令 $\mathcal{S}_{\text{full}}$ 为集合 $\{\, \rho_{\text{full}} \mid \rho \in \mathcal{S} \,\}$ 和 $\mathcal{S}' = \mathcal{S}_{\text{full}} \cup \triangledown_\exists(\mathcal{S}) \cup \triangledown_{\text{DL}}(\mathcal{S})$。显然，对于任意实例 $\mathcal{I}$：

$$Ch(\mathcal{I}, \mathcal{S}') \leftrightarrow Ch(\mathcal{I}, \triangledown(\mathcal{S})).$$

由上述内容可知，关于 $\mathcal{S}'$ 和 $\triangledown(\mathcal{S})$ 的重写是等价的——因此，$\mathcal{S}'$ 可 UCQ 重写当且仅当 $\triangledown(\mathcal{S})$ 可 UCQ 重写——所以只需证明 $\mathcal{S}'$ 可 UCQ 重写。现在，观察以下内容：

$$Ch(\mathcal{I}, \mathcal{S}') \leftrightarrow Ch(\ \ Ch(\mathcal{I},\ \triangledown_\exists(\mathcal{S}) \cup \triangledown_{\text{DL}}(\mathcal{S})),\ \ \mathcal{S}_{\text{full}}).$$

我们认为这一点成立，因为 $\triangledown_\exists(\mathcal{S}) \cup \triangledown_{\text{DL}}(\mathcal{S})$ 是非递归的，而且来自 $\mathcal{S}_{\text{full}}$ 的规则应用不会触发来自 $\triangledown_\exists(\mathcal{S}) \cup \triangledown_{\text{DL}}(\mathcal{S})$ 的规则——这些规则的规则头已分别包含在 $\mathcal{S}_{\text{full}}$ 的规则头中。

根据上述等价关系，我们可以认为，如果 $\mathcal{S}_{\text{full}}$ 是可 UCQ 重写的，那么 $\mathcal{S}'$ 也是。理由是：给定一个 CQ q 及其在 $\mathcal{S}_{\text{full}}$ 下的有界深度常数 c，其在 $\mathcal{S}'$ 下的有界深度常数不大于 $c+2$，因为 $\bigtriangledown_{\exists}(\mathcal{S}) \cup \bigtriangledown_{\text{DL}}(\mathcal{S})$ 是非递归的，而且 $\bigtriangledown_{\exists}(\mathcal{S}) \cup \bigtriangledown_{\text{DL}}(\mathcal{S})$ 下的所有查询的有界深度常数至多为 2。

因此，只要证明 $\mathcal{S}_{\text{full}}$ 是可 UCQ 重写的，我们就完成了证明。令 $\mathbb{S}'$ 包含出现在 $\mathcal{S}_{\text{full}}$中的新谓词 A 或 B（带下标）。注意， $\mathbb{S}'$ 中的谓词不会在 $\mathcal{S}_{\text{full}}$ 中递归使用，并且将范围限制为 $\mathbb{S}$ 中的谓词后，规则集 $\mathcal{S}_{\text{full}}$ 和 $\mathcal{S}$ 完全相同。因此， $\mathcal{S}_{\text{full}}$ 和 $\mathcal{S}$ 的推导相同——由此我们得出结论： $\mathcal{S}_{\text{full}}$确实是可 UCQ 重写的，因为 $\mathcal{S}$ 是可 UCQ 重写的。 □