Normalisations of Existential Rules: Not so Innocuous!

David Carral, Lucas Larroque, Marie-Laure Mugnier, Michaël Thomazo

Normalisations of Existential Rules: Not so Innocuous!

David Carral[1] , **Lucas Larroque**[2*] , **Marie-Laure Mugnier**[1] , **Michaël Thomazo**[3]
[1]LIRMM, Inria, University of Montpellier, CNRS, Montpellier, France
[2]DI ENS, ENS, CNRS, PSL University, Paris, France
[3]Inria, DI ENS, ENS, CNRS, PSL University, Paris, France
{david.carral, michael.thomazo}@inria.fr, lucas.larroque@ens.psl.eu, mugnier@lirmm.fr

Abstract

Existential rules are an expressive knowledge representation language mainly developed to query data. In the literature, they are often supposed to be in some normal form that simplifies technical developments. For instance, a common assumption is that rule heads are atomic, i.e., restricted to a single atom. Such assumptions are considered to be made without loss of generality as long as all sets of rules can be normalised while preserving entailment. However, an important question is whether the properties that ensure the decidability of reasoning are preserved as well. We provide a systematic study of the impact of these procedures on the different chase variants with respect to chase (non-)termination and FO-rewritability. This also leads us to study open problems related to chase termination of independent interest.

1 Introduction

Existential rules are an expressive knowledge representation language mainly developed to query data (Baget et al. 2009; Calì, Gottlob, and Lukasiewicz 2009). Such rules are an extension of first-order function-free Horn rules (like those of Datalog) with existentially quantified variables in the rule heads, which allows to infer the existence of unknown individuals.

Querying a knowledge base (KB) $\mathcal{K} = \langle \mathcal{R}, F \rangle$, where $\mathcal{R}$ is a set of existential rules and F a set of facts, consists in computing all the answers to queries that are logically entailed from $\mathcal{K}$. Two main techniques have been developed, particularly in the context of the fundamental (Boolean) conjunctive queries. The *chase* is a bottom-up process that expands F by rule applications from $\mathcal{R}$ towards a fixpoint. It produces a *universal* model of $\mathcal{K}$, i.e., a model of $\mathcal{K}$ that homomorphically maps to all models of $\mathcal{K}$, which is therefore sufficient to decide query entailment. *Query rewriting* is a dual technique, which consists in rewriting a query q with the rules in $\mathcal{R}$ into a query q' such that q is entailed by $\mathcal{K}$ if and only if q' is entailed by F solely.

Conjunctive query answering being undecidable for existential rules (Beeri and Vardi 1981), both the chase and query rewriting may not terminate. There is however a wide range of rule subclasses defined by syntactic restrictions that ensure *chase termination* on any set of facts (see, e.g., various acyclicity notions in (Grau et al. 2013)) or the existence, for any conjunctive query, of a (finite) rewriting into a first-order query, a property referred to as *FO-rewritability* (Calvanese et al. 2007).

In the literature, existential rules are often supposed to be in some normal form that simplifies technical developments. For instance, a common assumption is that rule heads are atomic, i.e., restricted to a single atom. On the one hand, the use of single-head rules greatly simplifies the presentation of theoretical arguments (e.g., (Calì, Gottlob, and Pieris 2012)). On the other hand, this restriction may also simplify implementations; e.g., the optimisation procedure presented in (Tsamoura et al. 2021) exploits single-head rules to clearly establish the provenance of each fact computed during the chase. Moreover, after normalisation, we can apply existing methods to effectively determine if the chase terminates for an input single-head existential rule set if this set is linear (Leclère et al. 2019) or guarded (Gogacz, Marcinkowski, and Pieris 2020). Normal form assumptions are often made without loss of generality as long as all sets of rules can be normalised while preserving all interesting entailments. However, an important question is whether the properties that ensure the decidability of reasoning are preserved as well. In particular, what is the impact of common normalisation procedures on fundamental properties like chase termination or FO-rewritability?

In fact, the chase is a family of algorithms, which differ from each other in their termination properties. Here, we consider the four main chase variants, namely: the *oblivious* chase (Calì, Gottlob, and Kifer 2008a), the *semi-oblivious* (aka skolem) chase (Marnette 2009), the *restricted* (aka standard) chase (Fagin et al. 2003) and the *core* chase (Deutsch, Nash, and Remmel 2008). As the core chase has the inconvenience of being non-monotonic (i.e., the produced set of facts does not grow monotonically), we actually study a monotonic variant that behaves similarly regarding termination, namely the *equivalent* chase (Rocher 2016). The ability of a chase variant to halt on a given KB is directly related to its power of reducing logical redundancies introduced by rules. The oblivious chase blindly performs all possible rule applications, while the equivalent chase terminates exactly when the KB admits a finite universal model. The other variants lie between these two extremes. For practical efficiency reasons, the most implemented variant is the restricted chase. However, it is the only variant sensitive to

[*]Our work started when Lucas was intern at LIRMM-Inria.

	$\mathbb{O}$	$\mathbb{SO}$	$\exists$-$\mathbb{R}$	$\forall$-$\mathbb{R}$	$\exists$-$\mathbb{DF}$-$\mathbb{R}$	$\forall$-$\mathbb{DF}$-$\mathbb{R}$	$\mathbb{E}$
Single-Piece	$=$	$+$	$\neq$	$\neq$	$\neq$	$\neq$	$=$
One-Way	$=$	$=$	$-$	$-$	$-$	$-$	$-$
Two-Way	$=$	$=$	$+$	$-$	$=$	$=$	$=$

Table 1: Impact on chase termination. Chase variants are denoted as follows: $\mathbb{O}$: Oblivious; $\mathbb{SO}$: semi-oblivious; ($\mathbb{DF}$-)$\mathbb{R}$: (Datalog-first) restricted; $\mathbb{E}$: equivalent.

the order of rule applications: for a given KB, there may be sequences of rule applications that terminate, while others do not. We study a natural strategy, called *Datalog-first restricted chase*, which prioritises Datalog rules (whose head does not include existential quantifiers) thus achieving termination in many real-world cases (Carral, Dragoste, and Krötzsch 2017). Moreover, experiments have shown that it is indeed a very efficient strategy (Urbani et al. 2018).

On the other hand, we consider two well-known normalisation procedures of a set of rules: *single-piece*-decomposition, which breaks rule heads into subsets called pieces and outputs a logically equivalent rule set (Baget et al. 2011; König et al. 2013); and *atomic*-decomposition, which requires to introduce fresh predicates and outputs a set of atomic-head rules that form a conservative extension of the original set, hence preserve entailment (Calì, Gottlob, and Kifer 2008b; Calì, Gottlob, and Pieris 2012).

Contributions. We provide a systematic study of the impact of these procedures on the different chase variants with respect to chase (non-)termination and FO-rewritability. This also leads us to solve some open problems related to chase termination, which are of independent interest.

Although the relationships between most chase variants with respect to chase termination are well understood (Grahne and Onet 2018), the question remained open regarding the restricted chase and its Datalog-first version. Unexpectedly, we found that Datalog-first strategies are not always optimal: we exhibit a rule set $\mathcal{R}$ such that the restricted chase has a terminating sequence on any KB $\langle\mathcal{R}, F\rangle$ but there is a KB $\langle\mathcal{R}, F\rangle$ on which no Datalog-first strategy terminates (Section 3).

While it appears that none of the considered decompositions influences FO-rewritability, the situation is very different concerning chase termination, as summarized in Table 1. Note that we distinguish between two behaviors for restricted chases: $\exists$ means that at least one chase sequence terminates on any KB ("sometimes-termination") and $\forall$ that all sequences terminate on any KB ("termination"). Single-piece-decomposition (Section 4) has no impact on the oblivious and equivalent chases (noted =), a positive impact on the semi-oblivious chase (noted +), and an erratic impact on (Datalog-first-)restricted chase (noted $\neq$). The standard atomic-decomposition (Section 5), called *one-way* in the table, has a negative impact on all chase variants, except for the (semi-)oblivious ones. Looking for a well-behaved atomic-decomposition procedure, we study a new one, named *two-way* (Section 6). A salient property of this decomposition is that it preserves the existence of a finite universal model. As shown in the table, two-way behaves better than one-way: it preserves (sometimes-)termination of the Datalog-first restricted chase and may even improve the sometimes-termination of the restricted chase. However, the negative impact on the termination of the restricted chase remains. These findings led us to an intriguing question: does a computable normalisation procedure exist that produces atomic-head rules and exactly preserves the termination of the restricted chase? We show that the answer is negative by a complexity argument (Section 7). More specifically, we study the decidability status of the following problem: Given a KB $\mathcal{K} = \langle\mathcal{R}, F\rangle$, does the restricted chase terminate on $\mathcal{K}$? We show that the associated membership problem is at least at the second level of the arithmetical hierarchy (precisely Π^0_2-hard) when there is no restriction on $\mathcal{R}$, while it is recursively enumerable (in Σ^0_1) when $\mathcal{R}$ is a set of atomic-head rules. Since $\Sigma^0_1 \subsetneq \Pi^0_2$, we obtain the negative answer to our question.

The complete proofs for all of the results in this paper can be found on an arXiv submission with the same name.

2 Preliminaries

First-Order Logic (FOL) We define $\mathtt{Preds}$, $\mathtt{Cons}$, and $\mathtt{Vars}$ to be mutually disjoint, countably infinite sets of *predicates*, *constants*, and *variables*, respectively. Every $P \in \mathtt{Preds}$ has an *arity* $\mathsf{ar}(P) \geq 0$. Let $\mathtt{Terms} = \mathtt{Cons} \cup \mathtt{Vars}$ be the set of *terms*. We write lists $t_1, \ldots, t_n$ of terms as $\vec{t}$ and often treat them as sets. For a formula or set thereof U, let $\mathtt{Preds}(U)$, $\mathtt{Cons}(U)$, $\mathtt{Vars}(U)$, and $\mathtt{Terms}(U)$ be the sets of all predicates, constants, variables, and terms that occur in U, respectively.

An *atom* is a FOL formula $P(\vec{t})$ with P a $|\vec{t}|$-ary predicate and $\vec{t} \in \mathtt{Terms}$. For a formula U, we write $U[\vec{x}]$ to indicate that $\vec{x}$ is the set of all free variables that occur in U.

Definition 1. An *(existential) rule* R is a FOL formula

$$\forall\vec{x}\forall\vec{y}.\big(B[\vec{x},\vec{y}] \to \exists\vec{z}.H[\vec{x},\vec{z}]\big) \quad (1)$$

where $\vec{x}$, $\vec{y}$, and $\vec{z}$ are pairwise disjoint lists of variables; and B and H are (finite) non-empty conjunctions of atoms, called the *body* and the *head* of R, respectively. The set $\vec{x}$ is the *frontier* of R. If $\vec{z}$ is empty, then R is a *Datalog* rule.

Next, we often denote a rule such as R above by $B \to H$ or $B \to \exists\vec{z}.H$, omitting all or some quantifiers.

A *factbase* F is an existentially closed (finite) conjunction of atoms. A *Boolean conjunctive query* (BCQ) has the same form as a factbase, and we often identify both notions.A *knowledge base* (KB) $\mathcal{K}$ is a tuple $\langle\mathcal{R}, F\rangle$ with $\mathcal{R}$ a rule set and F a factbase. We often identify rule bodies, rule heads, and factbases with (finite) sets of atoms.

Given atom sets F and F', a *homomorphism* π from F to F' is a function with domain $\mathtt{Vars}(F)$ such that $\pi(F) \subseteq F'$; π is an *isomorphism* from F to F' if additionally, π is injective and π^{-1} is a homomorphism from F to F'. A homomorphism π from F to F' is a *retraction* if π is the identity over $\mathtt{Vars}(F) \cap \mathtt{Vars}(F')$ (next, we often use this notion with $F' \subseteq F$).

We identify logical interpretations with atom sets. An atom set F *satisfies* a rule $R = B \to H$ if, for every homomorphism π from B to F, there is an extension $\hat{\pi}$ of π with $\hat{\pi}(H) \subseteq F$; equivalently, F is a *model* of R. An atom set $\mathcal{M}$ is a *model of a factbase* F if there is a homomorphism from F to $\mathcal{M}$, and it is a *model of a KB* $\langle \mathcal{R}, F\rangle$ if it is a model of F and satisfies all rules in $\mathcal{R}$. Given KBs or atom sets A and B, A *entails* B, written $A \models B$, if every model of A is a model of B; A and B are *equivalent* if $A \models B$ and $B \models A$. Given atom sets F and F', it is known that $F \models F'$ iff there is a homomorphism from F' to F.

Definition 2. A model $\mathcal{M}$ of a KB $\mathcal{K}$ is *universal* if there is a homomorphism from $\mathcal{M}$ to every model of $\mathcal{K}$.

Every KB $\mathcal{K}$ admits some (possibly infinite) universal model. Hence, $\mathcal{K} \models Q$ for any BCQ Q iff there is a homomorphism from a universal model of $\mathcal{K}$ to Q. The *BCQ entailment* problem takes as input a KB $\mathcal{K}$ and a BCQ Q and asks if $\mathcal{K} \models Q$; it is undecidable (Beeri and Vardi 1981).

Next, we will consider transformations of rule sets that introduce fresh predicates. To specify the relationships between a rule set and its decomposition, we will rely on the notion of conservative extension:

Definition 3 (Conservative extension)**.** Let $\mathcal{R}$ and $\mathcal{R}'$ be two rule sets such that $\texttt{Preds}(\mathcal{R}) \subseteq \texttt{Preds}(\mathcal{R}')$. The set $\mathcal{R}'$ is a *conservative extension* of the set $\mathcal{R}$ if (1) the restriction of any model of $\mathcal{R}'$ to the predicates in $\texttt{Preds}(\mathcal{R})$ is a model of $\mathcal{R}$, and (2) any model $\mathcal{M}$ of $\mathcal{R}$ can be extended to a model $\mathcal{M}'$ of $\mathcal{R}'$ that has the same domain (i.e., $\texttt{Terms}(\mathcal{M})$ = $\texttt{Terms}(\mathcal{M}')$) and agrees with $\mathcal{M}$ on the interpretation of the predicates in $\texttt{Preds}(\mathcal{R})$ (i.e., they have the same atoms with predicates in $\texttt{Preds}(\mathcal{R})$).

When $\mathcal{R}'$ is a conservative extension of $\mathcal{R}$, for any factbase F the KBs $\langle \mathcal{R}, F\rangle$ and $\langle \mathcal{R}', F\rangle$ entail the same (closed) formulas on $\texttt{Preds}(\mathcal{R})$, in particular BCQs.

The chase The chase is a family of procedures that repeatedly apply rules to a factbase until a fixpoint is reached. We formally define such procedures before stating their correctness with respect to factbase entailment in Proposition 7.

Definition 4 (Triggers and derivations)**.** Given a fact set F, a *trigger* t on F is a tuple $\langle R, \pi\rangle$ with $R = B \to \exists\vec{z}.H$ a rule and π a homomorphism from B to F. Let $\mathsf{support}(t) = \pi(B)$ and $\mathsf{output}(t) = \pi^R(H)$, where π^R is the extension of π that maps every variable $z \in \vec{z}$ to the fresh variable z_t that is unique for z and t. A *derivation* from a KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$ is a sequence $\mathcal{D} = (\emptyset, F_0), (t_1, F_1), \ldots$ such that:

1. Every F_i in $\mathcal{D}$ is a factbase; moreover, $F_0 = F$.
2. Every t_i in $\mathcal{D}$ is a trigger $\langle R, \pi\rangle$ on F_{i-1} such that $R \in \mathcal{R}$, $\mathsf{output}(t_i) \not\subseteq F_{i-1}$, and $F_i = F_{i-1} \cup \mathsf{output}(t_i)$.

The *result* of $\mathcal{D}$, written $\mathsf{res}(\mathcal{D})$, is the union of all the factbases in $\mathcal{D}$. Let $\mathsf{triggers}(\mathcal{D})$ be the set of all triggers in $\mathcal{D}$ and $\mathsf{length}(\mathcal{D}) = |\mathsf{triggers}(\mathcal{D})|$ be the length of $\mathcal{D}$.

Different chase variants build specific derivations according to different criteria of trigger applicability. Below, the letters $\mathbb{O}$, $\mathbb{SO}$, $\mathbb{R}$, and $\mathbb{E}$ respectively refer to so-called oblivious, semi-oblivious, restricted, and equivalent[1] variants.

Definition 5 (Applicability)**.** A trigger $t = \langle R, \pi\rangle$ on a factbase F is (i) $\mathbb{O}$-*applicable* on F if $\mathsf{output}(t) \not\subseteq F$, (ii) $\mathbb{SO}$-*applicable* on F if $\mathsf{output}(t') \not\subseteq F$ for every trigger $t' = (R, \pi')$ with $\pi(x) = \pi'(x)$ for all $x \in \mathsf{fr}(R)$, (iii) $\mathbb{R}$-*applicable* on F if there is no retraction from $F \cup \mathsf{output}(t)$ to F, and (iv) $\mathbb{E}$-*applicable* on F if there is no homomorphism from $F \cup \mathsf{output}(t)$ to F.

Example 1. *Consider the KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$ with $\mathcal{R} = \{R = P(x, y) \to \exists z.P(y, z) \land P(z, y)\}$ and $F = \{P(a, b)\}$ with a and b some constants. The trigger $t_1 = (R, \pi_1)$ with $\pi_1 = \{x \mapsto a, y \mapsto b\}$ is $\mathbb{X}$-applicable on $F_0 = F$ (for any $\mathbb{X}$), and $\mathsf{output}(t_1) = \{P(b, z_{t_1}), P(z_{t_1}, b)\}$. There are two new triggers on $F_1 = F \cup \mathsf{output}(t_1)$, both $\mathbb{O}$- and $\mathbb{SO}$-applicable, but neither $\mathbb{R}$- nor $\mathbb{E}$-applicable. For instance, consider $t_2 = (R, \pi_2)$ with $\pi_2 = \{x \mapsto b, y \mapsto z_{t_1}\}$ and $\mathsf{output}(t_2) = \{P(z_{t_1}, z_{t_2}), P(z_{t_2}, z_{t_1})\}$: there is a retraction from $F_1 \cup \mathsf{output}(t_2)$ to F_1, which maps z_{t_2} to b.*

Definition 6 (($\mathbb{DF}$-)$\mathbb{X}$-Chase)**.** For an $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{R}, \mathbb{E}\}$, an $\mathbb{X}$-*derivation from a KB* $\mathcal{K} = \langle \mathcal{R}, F\rangle$ is a derivation $\mathcal{D}$ such that every trigger $t_i \in \mathsf{triggers}(\mathcal{D})$ is $\mathbb{X}$-applicable on F_i; $\mathcal{D}$ is a $\mathbb{DF}$-$\mathbb{X}$-*derivation* if it gives priority to Datalog rules: for any $t_i = \langle R, \pi\rangle \in \mathsf{triggers}(\mathcal{D})$, if R is a non-Datalog rule, then F_{i-1} satisfies every Datalog rule in $\mathcal{R}$. A ($\mathbb{DF}$-)$\mathbb{X}$-derivation $\mathcal{D}$ is *fair* if for every F_i occurring in $\mathcal{D}$ and trigger t $\mathbb{X}$-applicable on F_i, there is some $j > i$ such that t is not $\mathbb{X}$-applicable on F_j. A ($\mathbb{DF}$-)$\mathbb{X}$-derivation is *terminating* if it is fair and finite.

The result of any fair $\mathbb{X}$-derivation is a universal model of the KB, for $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{R}\}$, and has a retraction to a universal model for $\mathbb{X} = \mathbb{E}$. Therefore, we obtain:

Proposition 7. *Consider a BCQ Q, a KB $\mathcal{K}$, and some fair $\mathbb{X}$-derivation $\mathcal{D}$ from $\mathcal{K}$ where $\mathbb{X} \in \{\mathbb{Y}, \mathbb{DF}\text{-}\mathbb{Y}\}$ and $\mathbb{Y} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{R}, \mathbb{E}\}$. Then, $\mathcal{K} \models Q$ iff $\mathsf{res}(\mathcal{D}) \models Q$.*

Decidable Classes of Rule Sets We now define classes of rule sets that ensure the decidability of BCQ entailment, based either on chase termination or on query rewritability.

Definition 8 (Chase-Terminating Sets)**.** For a $\mathbb{Y} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{R}, \mathbb{E}\}$ and an $\mathbb{X} \in \{\mathbb{Y}, \mathbb{DF}\text{-}\mathbb{Y}\}$, let $CT^{\mathbb{X}}_{\forall\forall}$ (resp. $CT^{\mathbb{X}}_{\forall\exists}$) be the set of all rule sets $\mathcal{R}$ such that every (resp. some) fair $\mathbb{X}$-derivation from every KB $\langle \mathcal{R}, F\rangle$ is finite.

When $\mathcal{R} \in CT^{\mathbb{X}}_{\forall\forall}$ (resp. $CT^{\mathbb{X}}_{\forall\exists}$), $\mathcal{R}$ ensures the *termination* (resp. *sometimes termination*) of the $\mathbb{X}$-chase.

Example 2. *Consider the KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$ from Example 1. All fair $\mathbb{O}$- or $\mathbb{SO}$-derivations from $\mathcal{K}$ are infinite. The only one fair $\mathbb{R}$-derivation (resp. $\mathbb{E}$-derivation) from $\mathcal{K}$ is $\mathcal{D} = (\emptyset, F_0), (t_1, F_1)$. Any fair $\mathbb{R}$-derivation from a KB with $\mathcal{R}$ is finite and hence, $\mathcal{R} \in CT^{\mathbb{R}}_{\forall\forall}$ (and $\mathcal{R} \in CT^{\mathbb{E}}_{\forall\forall}$).*

[1]The equivalent chase behaves as the better-known core chase regarding termination: it halts exactly when the KB has a finite universal model. The difference lies in the fact that the core chase computes a minimal universal model (i.e., a core). The equivalent chase has the advantage of being monotonic ($\forall i, F_i \subseteq F_{i+1}$).

Definition 9 (FO-rewritability)**.** A rule set $\mathcal{R}$ is *FO-rewritable* if for any BCQ Q, there is a (finite) BCQ set $\{Q_1, \ldots, Q_n\}$ such that, for every factbase F, $\langle \mathcal{R}, F\rangle \models Q$ iff $F \models Q_i$ for some $1 \leq i \leq n$.

In our proofs, we rely on a property equivalent to FO-rewritability: the bounded derivation depth property, which has the advantage of being based on (a breadth-first version of) the chase (Calì, Gottlob, and Lukasiewicz 2009). See (Gottlob et al. 2014) about the equivalence between both properties.

Definition 10 (BDDP)**.** For a rule R and a factbase F, let $R(F) \supseteq F$ be the minimal factbase that includes $\mathsf{output}(t)$ for every trigger t with R. For a rule set $\mathcal{R}$, let $\mathcal{R}(F) = \bigcup_{R \in \mathcal{R}} R(F)$. For a KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$, let $\mathsf{Ch}_0(\mathcal{K}) = F$ and $\mathsf{Ch}_i(\mathcal{K}) = \mathcal{R}(\mathsf{Ch}_{i-1}(\mathcal{K}))$ for every $i \geq 1$.

A rule set $\mathcal{R}$ has the *bounded derivation depth property (BDDP)* if, for any BCQ Q, there is some $k \geq 0$ such that, for every factbase F, $\langle \mathcal{R}, F\rangle \models Q$ iff $\mathsf{Ch}_k(\langle \mathcal{R}, F\rangle) \models Q$.

Normalisation Procedures Finally, we formally define normalisation procedures and their impact on the above properties. A *normalisation procedure* is a function f that maps rule sets to rule sets (complying with a certain shape) such that for any rule set $\mathcal{R}$, $\langle \mathcal{R}, F\rangle \models Q$ iff $\langle f(\mathcal{R}), F\rangle \models Q$ for any factbase F and BCQ Q on $\texttt{Preds}(\mathcal{R})$.

Definition 11. Consider some $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{R}, \mathbb{DF}\text{-}\mathbb{R}, \mathbb{E}\}$. Then, a normalisation procedure f:

- *Preserves termination* of the $\mathbb{X}$-chase if $f(CT^{\mathbb{X}}_{\forall\forall}) \subseteq CT^{\mathbb{X}}_{\forall\forall}$; it preserves *sometimes-termination* of the $\mathbb{X}$-chase if $f(CT^{\mathbb{X}}_{\forall\exists}) \subseteq CT^{\mathbb{X}}_{\forall\exists}$.
- *Preserves non-termination* of the $\mathbb{X}$-chase if $f(\overline{CT^{\mathbb{X}}_{\forall\forall}}) \subseteq \overline{CT^{\mathbb{X}}_{\forall\forall}}$. Otherwise, f *may gain termination.*
- *Preserves rewritability* if it maps FO-rewritable rule sets to FO-rewritable rule sets.

3 Generality of Chase-Terminating Rule Sets

One of our goals is to study normalisation procedures that preserve membership over the sets of chase-terminating rule sets from Definition 8. To be systematic, we clarify the equality and strict-subset relations between these sets in Theorems 12 and 13, respectively. Grahne and Onet already proved most of the claims in these theorems (see Theorem 4.5, Propositions 4.6 and 4.7, and Corollary 4.8 in (Grahne and Onet 2018)); we reprove some of them again to be self-contained. However, note that all results regarding Datalog-first chase variants are our own contribution.

Theorem 12. *For every $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}, \mathbb{E}\}$, we have that*

$$CT^{\mathbb{X}}_{\forall\forall} = CT^{\mathbb{X}}_{\forall\exists} = CT^{\mathbb{DF}\text{-}\mathbb{X}}_{\forall\forall} = CT^{\mathbb{DF}\text{-}\mathbb{X}}_{\forall\exists}$$

Sketch. To show that the theorem holds if $\mathbb{X} = \mathbb{O}$ (resp. $\mathbb{X} = \mathbb{SO}$), it suffices to prove that all fair $\mathbb{X}$-derivations from an input KB $\mathcal{K}$ produce the same result (resp. same result up to isomorphism); see forthcoming Lemma 16.

All fair $\mathbb{E}$-derivations from an input KB $\mathcal{K}$ are finite iff $\mathcal{K}$ admits a finite universal model (Rocher 2016). Hence, the theorem holds if $\mathbb{X} = \mathbb{E}$. □

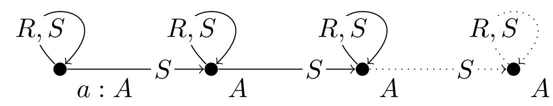

Figure 1: The only result of the $\mathbb{DF}$-$\mathbb{R}$-chase from the KB $\mathcal{K} = \langle \mathcal{R}, \{A(a)\}\rangle$ introduced in the proof of Theorem 13

The equalities in Theorem 12 simplify our work: for instance, if a function preserves termination of the oblivious chase, then we know that it also preserves sometimes-termination of this variant. Alas, the remaining sets of chase-terminating rule sets are not equal:

Theorem 13. *The following hold:*

$$CT^{\mathbb{O}}_{\forall\forall} \subset CT^{\mathbb{SO}}_{\forall\forall} \subset CT^{\mathbb{R}}_{\forall\forall} \subset CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$$
$$\subset CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists} \subset CT^{\mathbb{R}}_{\forall\exists} \subset CT^{\mathbb{E}}_{\forall\forall}$$

Sketch. The subset inclusions follow by definition; we present some rule sets to show that these are strict:

$$\{P(x,y) \to \exists z.P(x,z)\} \in CT^{\mathbb{SO}}_{\forall\forall} \setminus CT^{\mathbb{O}}_{\forall\forall}$$

$$\{P(x,y) \to \exists z.P(y,z) \wedge P(z,y)\} \in CT^{\mathbb{R}}_{\forall\forall} \setminus CT^{\mathbb{SO}}_{\forall\forall}$$

$$\{P(x,y) \to \exists z.P(y,z),$$
$$P(x,y) \to P(y,x)\} \in CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall} \setminus CT^{\mathbb{R}}_{\forall\forall}$$

$$\{P(x,y) \to \exists z.P(y,z) \wedge P(z,y),$$
$$P(x,y) \to \exists z.P(y,z)\} \in CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists} \setminus CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$$

$$\{P(x,y) \to \exists z.P(y,z),$$
$$P(x,y) \wedge P(y,z) \to P(y,x)\} \in CT^{\mathbb{E}}_{\forall\forall} \setminus CT^{\mathbb{R}}_{\forall\exists}$$

Moreover, the rule set $\mathcal{R} = \{(2\text{–}6)\}$ is in $CT^{\mathbb{R}}_{\forall\exists} \setminus CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$:

$$A(x) \to R(x,x) \tag{2}$$
$$R(x,y) \wedge S(y,z) \to S(x,x) \tag{3}$$
$$A(x) \wedge S(x,y) \to A(y) \tag{4}$$
$$A(x) \to \exists z.R(x,z) \tag{5}$$
$$R(x,y) \to \exists z.S(y,z) \tag{6}$$

To show that $\mathcal{R} \notin CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$ we prove that the KB $\mathcal{K} = \langle \mathcal{R}, \{A(a)\}\rangle$ does not admit terminating $\mathbb{DF}$-$\mathbb{R}$-derivations. Specifically, all fair $\mathbb{DF}$-$\mathbb{R}$-derivations from $\mathcal{K}$ yield the same result, which is depicted in Figure 1. Rule (2) is applied first, then the following pattern is repeated: apply rule (6) followed by Datalog rules (3), (4) and (2). Rule (5) is never applicable since priority is given to rule (2). To show that $\mathcal{R} \in CT^{\mathbb{R}}_{\forall\exists}$ we verify that every KB of the form $\langle \mathcal{R}, F\rangle$ admits a terminating $\mathbb{R}$-derivation. We can produce such a derivation by exhaustively applying the rules in $\mathcal{R}$ in the following order:[2] first, apply rules (4), (5) and (6); then, apply (3); finally, apply (2). □

Our main achievement is showing that $CT^{\mathbb{R}}_{\forall\exists} \setminus CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$ is non-empty; thus proving that Datalog-first strategies are not necessarily the most terminating for the restricted chase.

[2] To understand why this strategy results in a terminating $\mathbb{R}$-derivation, we suggest to first try it on $F = \{A(a)\}$, which yields $\{A(a), R(a,z_1), S(z_1,z_2), S(a,a), R(a,a)\}$.

4 Single-Piece Decomposition

The single-piece decomposition (piece-decomposition in short) is a procedure that splits a rule $R = B \to \exists z.H$ into several rules $R_1, \ldots, R_n$ that have the same body as R, and whose head is a subset of H that (directly or indirectly) shares some existential variable in H.

Definition 14. The *piece graph* of a rule $R = B \to \exists\vec{z}.H$ is the graph whose vertices are the atoms in H, and with an edge between a and a' if $\vec{z} \cap \mathtt{Vars}(a) \cap \mathtt{Vars}(a')$ is non-empty. A *(rule) piece* of R is the conjunction of atoms corresponding to a (connected) component of its piece graph.

The *piece-decomposition* of a rule $R = B \to \exists\vec{z}.H$ is the rule set $\mathsf{sp}(R) = \{B \to \exists\vec{v}.H' \mid H' \text{ is a piece of } R\}$. For a rule set $\mathcal{R}$, let $\mathsf{sp}(\mathcal{R}) = \bigcup_{R\in\mathcal{R}} \mathsf{sp}(R)$.

Example 3. *Consider the rule* (7) *and its single-piece decomposition* $\mathsf{sp}((7)) = \{(8\text{–}10)\}$*:*

$$R(x,y) \to \exists z,u.P(x,z) \wedge A(z) \wedge A(u) \wedge P(x,y) \quad (7)$$
$$R(x,y) \to \exists z.P(x,z) \wedge A(z) \quad (8)$$
$$R(x,y) \to \exists u.A(u) \quad (9)$$
$$R(x,y) \to P(x,y) \quad (10)$$

Piece-decomposition is indeed a normalisation procedure, since it preserves logical equivalence:

Proposition 15. *A rule set $\mathcal{R}$ is equivalent to the set* $\mathsf{sp}(\mathcal{R})$.

The following lemma is later applied to show that the piece-decomposition preserves termination of the oblivious and semi-oblivious chase in Theorem 18:

Lemma 16. *Consider some fair $\mathbb{X}$-derivations $\mathcal{D}$ and $\mathcal{D}'$ from a KB $\mathcal{K}$. If $\mathbb{X} = \mathbb{O}$, then $\mathsf{res}(\mathcal{D}) = \mathsf{res}(\mathcal{D}')$. If $\mathbb{X} = \mathbb{SO}$, then $\mathsf{res}(\mathcal{D})$ is isomorphic to $\mathsf{res}(\mathcal{D}')$.*

Definition 17. Given some $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}\}$ and a KB $\mathcal{K}$, let $\mathsf{Ch}_{\mathbb{X}}(\mathcal{K})$ be some (arbitrarily chosen) atom set that is isomorphic to the result of all fair $\mathbb{X}$-derivations from $\mathcal{K}$.

Theorem 18. *The piece-decomposition preserves the termination of the $\mathbb{O}$-chase and $\mathbb{SO}$-chase.*

Sketch. Consider some $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}\}$ and some $\mathbb{X}$-derivation $\mathcal{D}$ from a KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$. We can show via induction on $\mathcal{D}$ that there is an injective homomorphism from $\mathsf{Ch}_{\mathbb{X}}(\langle \mathsf{sp}(\mathcal{R}), F\rangle)$ to $\mathsf{Ch}_{\mathbb{X}}(\mathcal{K})$. Therefore, finiteness of $\mathsf{Ch}_{\mathbb{X}}(\mathcal{K})$ implies finiteness of $\mathsf{Ch}_{\mathbb{X}}(\langle \mathsf{sp}(\mathcal{R}), F\rangle)$. □

The piece-decomposition does not preserve the termination of any restricted chase variant. The reason is that it allows for intertwining the application of split rules that come from different original rules, resulting in new application strategies that may lead to non-termination.

Theorem 19. *The piece-decomposition does not preserve termination of the $\mathbb{R}$- or the $\mathbb{DF}$-$\mathbb{R}$-chase.*

Sketch. Consider the rule set $\mathcal{R} = \{(11), (12)\}$ and its piece-decomposition $\mathsf{sp}(\mathcal{R}) = \{(12\text{–}14)\}$:

$$P(x,y) \to P(y,y) \wedge A(y) \quad (11) \qquad P(x,y) \to P(y,y) \quad (13)$$
$$A(x) \to \exists z.P(x,z) \quad (12) \qquad P(x,y) \to A(y) \quad (14)$$

The set $\mathcal{R}$ is in $CT^{\mathbb{R}}_{\forall\forall}$ because triggers with (12) are not $\mathbb{R}$-applicable to the output of triggers with (11). The set $\mathsf{sp}(\mathcal{R})$ is not in $CT^{\mathbb{R}}_{\forall\forall}$ because the KB $\langle \mathsf{sp}(\mathcal{R}), \{A(a)\}\rangle$ admits the following non-terminating $\mathbb{R}$-derivation $(\emptyset, F_0), (t_1, F_1), \ldots$:

$$F_0 = \{A(a)\}, \qquad F_4 = \{P(z_1,z_1)\} \cup F_3,$$
$$F_1 = \{P(a,z_1)\} \cup F_0, \qquad F_5 = \{A(z_2)\} \cup F_4,$$
$$F_2 = \{A(z_1)\} \cup F_1, \qquad F_6 = \{P(z_2,z_3)\} \cup F_5,$$
$$F_3 = \{P(z_1,z_2)\} \cup F_2, \qquad \ldots$$

This derivation is built by first applying rule (12) (leading to F_1), then indefinitely repeating the sequence of rule applications (14), (12), and (13). In contrast, the only fair $\mathbb{R}$-derivation with $\mathcal{R}$ would apply (12) then (11), leading to $\{A(a), P(a,z_1), P(z_1,z_1), A(z_1)\}$.

To get a similar behavior with the $\mathbb{DF}$-$\mathbb{R}$ chase, we introduce "dummy" existential variables in rules (11) and (12), so that their piece-decomposition has no Datalog rules:

$$P(x,y,v) \to \exists u,w.P(y,y,u) \wedge A(y,w) \quad (15)$$
$$A(x,v) \to \exists z,u.P(x,z,u) \quad (16)$$

Applying analogous arguments we can show that $\mathcal{R}' = \{(15), (16)\}$ is in $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$ and that $\mathsf{sp}(\mathcal{R}')$ is not. □

Initially, we believed that the piece-decomposition would preserve sometimes-termination of the $\mathbb{R}$-chase. Our intuition was that, given a terminating $\mathbb{R}$-derivation from a KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$, we could replicate this derivation from $\langle \mathsf{sp}(\mathcal{R}), F\rangle$ by applying the split rules in $\mathsf{sp}(\mathcal{R})$ piece by piece. Surprisingly, this is not always possible:

Theorem 20. *The piece-decomposition does not preserve the sometimes-termination of the $\mathbb{R}$-chase.*

Sketch. The following set $\mathcal{R} = \{(17\text{–}22)\}$ is in $CT^{\mathbb{R}}_{\forall\exists}$ and its piece-decomposition $\mathsf{sp}(\mathcal{R}) = \{(18\text{–}24)\}$ is not. The set $\mathcal{R}$ is adapted from $\{(2\text{–}6)\}$ (proof of Th. 13). Note that (17) is split into two equivalent rules (23) and (24). To show that $\mathsf{sp}(\mathcal{R}) \notin CT^{\mathbb{R}}_{\forall\exists}$, we start again from $\{A(a)\}$. Again, some R-atom is created and leads to apply other rules. With $\mathcal{R}$, applying (17) then (18) creates an atom of form $R(y_1, z_1)$, while with $\mathsf{sp}(\mathcal{R})$, applying (23) then (18) creates an atom $R(y_1, y_1)$. This loop leads to non-termination.

$$A(x) \to \exists y,z.U(x,y) \wedge H(y,x) \wedge U(x,z) \wedge H(z,x) \quad (17)$$
$$U(x,y) \wedge U(x,z) \to R(y,z) \quad (18)$$
$$U(x,z) \wedge R(y,z) \to \exists v.R(z,v) \quad (19)$$
$$R(x,y) \wedge R(y,z) \to \exists v.S(z,v) \quad (20)$$
$$R(x,y) \wedge S(y,z) \to S(x,x) \quad (21)$$
$$A(x) \wedge U(x,y) \wedge S(y,z) \to \exists v.H(z,v) \wedge A(v) \quad (22)$$
$$A(x) \to \exists y.U(x,y) \wedge H(y,x) \quad (23)$$
$$A(x) \to \exists z.U(x,z) \wedge H(z,x) \quad (24)$$

□

Theorem 21. *The piece-decomposition does not preserve the sometimes-termination of the $\mathbb{DF}$-$\mathbb{R}$-chase.*

Sketch. The set $\mathcal{R} = \{(25–29)\}$ is in $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$ while $\mathsf{sp}(\mathcal{R}) = \{(26–31)\}$ is not. Note that the only difference with $\{(2–6)\}$ is the atom $H(x, y)$ in the first rule, making it non-Datalog, which prevents its early application.

$$A(x) \rightarrow \exists y.R(x, x) \wedge H(x, y) \quad (25)$$
$$R(x, y) \wedge S(y, z) \rightarrow S(x, x) \quad (26)$$
$$A(x) \wedge S(x, y) \rightarrow A(y) \quad (27)$$
$$A(x) \rightarrow \exists y.R(x, y) \quad (28) \qquad A(x) \rightarrow R(x, x) \quad (30)$$
$$R(x, y) \rightarrow \exists z.S(y, z) \quad (29) \qquad A(x) \rightarrow \exists y.H(x, y) \quad (31)$$

□

Since piece-decomposition preserves logical equivalence (Proposition 15), one directly obtains that it preserves termination of the equivalent chase. Indeed, $\langle \mathcal{R}, F\rangle$ admits a finite universal model iff $\langle \mathsf{sp}(\mathcal{R}), F\rangle$ admits one:

Theorem 22. *The piece-decomposition preserves the termination of the $\mathbb{E}$-chase.*

The piece-decomposition may gain termination:

Theorem 23. *The piece-decomposition may gain termination (and sometimes-termination) of the $\mathbb{SO}$-, the $\mathbb{R}$-, and the $\mathbb{DF}$-$\mathbb{R}$-chase but not of the $\mathbb{O}$- and $\mathbb{E}$-chase.*

Sketch. Consider the set $\mathcal{R} = \{P(x, y) \rightarrow \exists z.P(x, z) \wedge R(x, y)\}$, which is not in $CT^{\mathbb{SO}}_{\forall\forall}$, $CT^{\mathbb{R}}_{\forall\exists}$, $CT^{\mathbb{R}}_{\forall\forall}$, $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$, or $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$. However, $\mathsf{sp}(\mathcal{R})$ is in all of these sets.

Concerning the $\mathbb{O}$-chase, we show via induction that $\mathsf{Ch}_{\mathbb{O}}(\langle \mathcal{R}, F\rangle)$ and $\mathsf{Ch}_{\mathbb{O}}(\langle \mathsf{sp}(\mathcal{R}), F\rangle)$ are isomorphic for any rule set $\mathcal{R}$ and factbase F (where $\mathsf{Ch}_{\mathbb{O}}(\cdot)$ is the function from Definition 17, which maps a KB to its only $\mathbb{O}$-chase result). Hence, all $\mathbb{O}$-derivations from $\langle \mathcal{R}, F\rangle$ are terminating iff all $\mathbb{O}$-derivations from $\langle \mathsf{sp}(\mathcal{R}), F\rangle$ are terminating. Concerning the $\mathbb{E}$-chase, we rely again on Proposition 15. □

To show that the piece-decomposition preserves FO-rewritability we show that it preserves the BDDP property.

Theorem 24. *A rule set $\mathcal{R}$ is BDDP iff $\mathsf{sp}(\mathcal{R})$ is BDDP.*

Sketch. We can prove by induction that, for any KBs $\langle \mathcal{R}, F\rangle$ and any $i \geq 1$, the factbases $\mathsf{Ch}_i(\langle \mathcal{R}, F\rangle)$ and $\mathsf{Ch}_i(\langle \mathsf{sp}(\mathcal{R}), F\rangle)$ are isomorphic. □

5 One-Way Atomic Decomposition

Piece-decomposition may not produce atomic-head rules; a useful restriction considered in many contexts. The following procedure is classically used to produce such rules:

Definition 25. The *one-way atomic decomposition* of a rule $R = B[\vec{x}, \vec{y}] \rightarrow \exists \vec{z}.H[\vec{x}, \vec{z}]$ is the rule set $\mathit{1ad}(R)$ that contains the rule $B \rightarrow \exists \vec{z}.X_R(\vec{x}, \vec{z})$ and, for each atom $P(\vec{t}) \in H$, the rule $X_R(\vec{x}, \vec{z}) \rightarrow P(\vec{t})$, where X_R is a fresh predicate unique for R, of arity $|\vec{x}| + |\vec{z}|$. Given a rule set $\mathcal{R}$, let $\mathit{1ad}(\mathcal{R}) = \bigcup_{R\in\mathcal{R}} \mathit{1ad}(R)$.

Example 4. *Consider the rule* (32) *and its one-way atomic decomposition* $\mathit{1ad}((32)) = \{(33–35)\}$*:*

$$R(x, y) \rightarrow \exists z.P(x, z) \wedge S(x, y, z) \quad (32)$$
$$R(x, y) \rightarrow \exists z.X_{(32)}(x, y, z) \quad (33)$$
$$X_{(32)}(x, y, z) \rightarrow P(x, z) \quad (34)$$
$$X_{(32)}(x, y, z) \rightarrow S(x, y, z) \quad (35)$$

Note that piece-decomposition would not decompose rule (32)*, i.e.,* $\mathsf{sp}((32)) = \{(32)\}$

Strictly speaking, $\mathcal{R}$ and $\mathit{1ad}(\mathcal{R})$ cannot be logically equivalent because they are built on different sets of predicates; however, it is straighforward to check that $\mathit{1ad}(\mathcal{R})$ is a conservative extension of $\mathcal{R}$. Therefore, one-way atomic decomposition is indeed a normalisation procedure.

The following is a corollary of forthcoming Theorem 34:

Theorem 26. *The one-way atomic decomposition preserves termination of the $\mathbb{O}$- and $\mathbb{SO}$-chase.*

An interesting phenomenon occurs with the one-way atomic decomposition: the notions of $\mathbb{SO}$-applicability on $\mathcal{R}$ and $\mathbb{R}$-applicability on $\mathit{1ad}(\mathcal{R})$ coincide:

Lemma 27. *Consider a KB $\langle \mathcal{R}, F\rangle$ and a finite derivation $\mathcal{D} = (\emptyset, F_0), (t_1, F_1), \ldots, (t_n, F_n)$ from $\langle \mathit{1ad}(\mathcal{R}), F\rangle$. Then, a trigger t with a rule in $\mathit{1ad}(\mathcal{R})$ is $\mathbb{R}$-applicable on F_n iff it is $\mathbb{SO}$-applicable on F_n.*

Proof. ($\Rightarrow$): from Definition 5. ($\Leftarrow$): Let $t = (R, \pi)$ with $R \in \mathit{1ad}(\mathcal{R})$. If R is Datalog, the notions of $\mathbb{SO}$- and $\mathbb{R}$-applicability coincide for every factbase. Otherwise, R is of the form $B[\vec{x}, \vec{y}] \rightarrow \exists \vec{z}.X_{R'}(\vec{x}, \vec{z})$ where $X_{R'} \notin \mathtt{Preds}(\mathcal{R})$ and $R' \in \mathcal{R}$ is of the form $B[\vec{x}, \vec{y}] \rightarrow \exists \vec{z}.H[\vec{x}, \vec{z}]$. If t is $\mathbb{SO}$-applicable on F_n, then, for every trigger $t' = (R, \pi')$ with $\pi(\vec{x}) = \pi'(\vec{x})$, it holds that $\mathtt{output}(t') \not\subseteq F_n$. By Definition 25, R is the only rule in $\mathit{1ad}(\mathcal{R})$ with $X_{R'}$ in its head, hence $\pi''(X_{R'}(\vec{x}, \vec{z})) \notin F_n$ for every extension π'' of π. Hence, t is $\mathbb{R}$-applicable on F_n. □

Intuitively, Lemma 27 implies that, after applying the one-way atomic decomposition, $\mathbb{R}$-applicability becomes as loose and unrestrictive as $\mathbb{SO}$-applicability. Therefore:

Theorem 28. *The one-way atomic decomposition does not preserve termination nor sometimes-termination of the $\mathbb{R}$- or the $\mathbb{DF}$-$\mathbb{R}$-chase.*

Proof. By Theorem 13, there is a rule set $\mathcal{R} \notin CT^{\mathbb{SO}}_{\forall\forall}$ that is in $CT^{\mathbb{R}}_{\forall\forall}$, $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$, $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$, and $CT^{\mathbb{R}}_{\forall\exists}$.[3] By Theorem 26, $\mathit{1ad}(\mathcal{R})$ is not in $CT^{\mathbb{SO}}_{\forall\forall}$; that is, there is some KB $\mathcal{K}$ of the form $\langle \mathit{1ad}(\mathcal{R}), F\rangle$ that does not admit any terminating $\mathbb{SO}$-derivation. By Lemma 27, every terminating $\mathbb{R}$-derivation from $\mathcal{K}$ is also a terminating $\mathbb{SO}$-derivation from $\mathcal{K}$. Therefore, $\mathcal{K}$ does not admit any terminating $\mathbb{R}$-derivation, hence $\mathit{1ad}(\mathcal{R})$ is not in $CT^{\mathbb{R}}_{\forall\forall}$, $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$, $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists}$, or $CT^{\mathbb{R}}_{\forall\exists}$. □

Theorem 29. *The one-way atomic decomposition does not preserve the termination of the $\mathbb{E}$-chase.*

[3]Such rule set in given in the proof of Theorem 13; see also $\{(36)\}$ in the proof of Theorem 29.

Proof. Consider the rule set $\mathcal{R} = \{(36)\}$ (see also Example 1) and its decomposition $1ad(\mathcal{R}) = \{(37–39)\}$:

$$P(x, y) \to \exists z.\ P(y, z) \wedge P(z, y) \quad (36)$$
$$P(x, y) \to \exists z. X_{(36)}(y, z) \quad (37)$$
$$X_{(36)}(y, z) \to P(y, z) \quad (38)$$
$$X_{(36)}(y, z) \to P(z, y) \quad (39)$$

The rule set $\mathcal{R}$ is in $CT^{\mathbb{E}}_{\forall\forall}$ since every $\mathbb{E}$-derivation from a KB $\langle \mathcal{R}, F\rangle$ yields a finite result, which is a subset of $F \cup \{P(a, z_a), P(z_a, a) \mid a \in \texttt{Terms}(F), z_a \notin \texttt{Terms}(F)\}$.

The rule set $1ad(\mathcal{R})$ is not in $CT^{\mathbb{E}}_{\forall\forall}$ since the KB $\mathcal{K} = \langle 1ad(\mathcal{R}), \{P(a, b)\}\rangle$ has no terminating $\mathbb{E}$-derivation. In fact, all fair $\mathbb{E}$-derivations from $\mathcal{K}$ yield the same result:

$$\{P(a, b), X_{(36)}(b, z_1), P(b, z_1), P(z_1, b)\} \cup$$
$$\{X_{(36)}(z_i, z_{i+1}), P(z_i, z_{i+1}), P(z_{i+1}, z_i) \mid i \geq 1\}$$

□

Again, to show that the one-way decomposition preserves FO-rewritability, we show that it preserves BDDP.

Theorem 30. *A rule set $\mathcal{R}$ is BDDP iff 1ad($\mathcal{R}$) is BDDP.*

Sketch. ($\Rightarrow$): For factbases F restricted to the original vocabulary Σ, we prove that $\mathsf{Ch}_i(F, \mathcal{R}) = \mathsf{Ch}_{2i}(F, 1ad(\mathcal{R}))_{|\Sigma}$. Dealing with arbitrary factbases is tackled by a weakening of this correspondence. ($\Leftarrow$): If $\mathcal{R}$ is not BDDP, there are Q and $\{F_i\}_{i\in\mathbb{N}}$ such that for all i, $F_i, \mathcal{R} \models Q$ and $\mathsf{Ch}_i(F_i, \mathcal{R}) \not\models Q$. Since for all F_i on Σ, $\mathsf{Ch}_i(F_i, \mathcal{R}) = \mathsf{Ch}_{2i}(F_i, 1ad(\mathcal{R}))_{|\Sigma}$, it holds that $\mathsf{Ch}_{2i}(F_i, 1ad(\mathcal{R})) \not\models Q$, hence $1ad(\mathcal{R})$ is not BDDP. □

6 Two-Way Atomic Decomposition

Despite the fact that it produces a conservative extension of the original rule set, the one-way atomic decomposition does not preserve the existence of a finite universal model; hence, it does not preserve equivalent chase termination.

Example 5. *As in the proof of Theorem 29, consider $\mathcal{R} = \{(36)\}$, its decomposition $1ad(\mathcal{R}) = \{(37–39)\}$, and the factbase $F = \{P(a, b)\}$. Then, $\mathcal{U} = \{P(a, b), P(b, z_1), P(z_1, b)\}$ is a finite universal model for $\langle \mathcal{R}, F\rangle$ that cannot be extended (keeping the same domain) into is a universal model of $\langle 1ad(\mathcal{R}), F\rangle$. Indeed, the set*

$$\{P(a, b), P(b, z_1), P(z_1, b), X_{(36)}(b, z_1), X_{(36)}(z_1, b)\}$$

is the smallest extension of $\mathcal{U}$ that is a model for $\langle 1ad(\mathcal{R}), F\rangle$, but it is not universal.

Hence, we define a notion similar to that of conservative extension, but whose purpose is to guarantee the preservation of the equivalent chase termination.

Definition 31. Let $\mathcal{R}$ and $\mathcal{R}'$ be two rule sets such that $\texttt{Preds}(\mathcal{R}) \subseteq \texttt{Preds}(\mathcal{R}')$. The set $\mathcal{R}'$ is a *universal-conservative extension* of the set $\mathcal{R}$ if, for any factbase F with $\texttt{Preds}(F) \subseteq \texttt{Preds}(\mathcal{R})$,

1. The restriction of any universal model of $\langle \mathcal{R}', F\rangle$ to the predicates in $\texttt{Preds}(\mathcal{R})$ is a universal model of $\langle \mathcal{R}, F\rangle$.
2. Any universal model $\mathcal{M}$ of $\langle \mathcal{R}, F\rangle$ can be extended to a universal model of $\langle \mathcal{R}', F\rangle$ that has the same domain and agrees with $\mathcal{M}$ on the interpretation of $\texttt{Preds}(\mathcal{R})$.

We now introduce a normalisation procedure that produces universal-conservative extensions:

Definition 32. The *two-way atomic decomposition* of a rule $R = B[\vec{x}, \vec{y}] \to \exists\vec{z}.H[\vec{x}, \vec{z}]$ is the rule set $2ad(R) = 1ad(R) \cup \{H[\vec{x}, \vec{z}] \to X_R(\vec{x}, \vec{z})\}$, with X_R the fresh predicate in $1ad(R)$. For a rule set $\mathcal{R}$, we let $2ad(\mathcal{R}) = \bigcup_{R\in\mathcal{R}} 2ad(R)$.

Example 6. *Consider again the rule* (32) *from Example 4. Then its two-way atomic decomposition is $2ad((32)) = 1ad((32)) \cup \{(40)\}$:*

$$P(x, z) \wedge S(x, y, z) \to X_{(32)}(x, y, z) \quad (40)$$

We establish that this new decomposition is indeed a normalisation procedure that has the desired property.

Proposition 33. *The rule set $2ad(\mathcal{R})$ is a conservative extension and a universal-conservative extension of $\mathcal{R}$.*

We can now focus our interest again on chase termination. Both atomic decompositions behave like the single-piece decomposition (Theorem 18) regarding the oblivious and the semi-oblivious chase:

Theorem 34. *Both atomic decompositions preserve the termination of the $\mathbb{O}$-chase and the $\mathbb{SO}$-chase.*

Sketch. Consider $\mathbb{X} \in \{\mathbb{O}, \mathbb{SO}\}$ and $\mathcal{K} = \langle \mathcal{R}, F\rangle$ a KB. First note that for these $\mathbb{X}$-chases, applying a rule cannot prevent the application of another rule. Hence, since $1ad(\mathcal{R}) \subseteq 2ad(\mathcal{R})$, it is sufficient to prove the result for $2ad(\mathcal{R})$. The proof is similar to that of Theorem 18: we show by induction on an arbitrary derivation $\mathcal{D}$ from $\mathcal{K}$ that there is an injective homomorphism from $\mathsf{Ch}_{\mathbb{X}}(\langle 2ad(\mathcal{R}), F\rangle)$ restricted to the predicates in $\texttt{Preds}(\mathcal{R})$ to $\mathsf{Ch}_{\mathbb{X}}(\mathcal{K})$, which leads to a similar conclusion. □

The behavior of the restricted chase is again less easily characterized, as we will see in the next results.

Theorem 35. *The two-way atomic decomposition preserves sometimes-termination of the $\mathbb{R}$-chase; it may also gain termination of this chase variant.*

Sketch. To prove preservation, consider a KB $\mathcal{K} = \langle \mathcal{R}, F\rangle$ such that $\mathcal{R} \in CT^{\mathbb{R}}_{\forall\exists}$. Then, there is a terminating $\mathbb{R}$-derivation $\mathcal{D}$ from $\mathcal{K}$. We can then show by induction that if a trigger $t = (R, \pi)$ with $R = B \to H$ is applied at some step, the trigger $t' = (B \to X_R, \pi)$ is applicable at the same step, then the triggers $t_i = (X_R \to H_i, \pi^R)$ also are, and that applying t' and all the t_i successively yields the same result (when restricted to the predicates in $\mathcal{R}$) as applying t. This shows that we can replicate a terminating derivation, and thus that the sometimes-termination is preserved.

We now present an example where we gain termination. Consider the rule set $\mathcal{R} = \{(41–45)\}$:

$$A(x) \to \exists y, z\ R(x, x, x) \wedge R(x, y, z) \quad (41)$$
$$R(x, y, z) \to \exists t.R(x, x, t) \quad (42)$$

$$R(x,x,y) \to \exists z.S(x,y,z) \quad (43)$$
$$R(x,x,y) \wedge S(x,y,z) \to S(x,x,x) \quad (44)$$
$$A(x) \wedge S(x,x,y) \to A(y) \quad (45)$$

There is no terminating $\mathbb{R}$-derivation on the KB $\langle \mathcal{R}, \{A(a)\} \rangle$ but there is one on $\langle 2ad(\mathcal{R}), F \rangle$ for any F. $\square$

Theorem 36. *The two-way atomic decomposition does not preserve the termination of the $\mathbb{R}$-chase.*

Proof. Consider the rule set $\mathcal{R} = \{(36)\}$ introduced in the proof of Theorem 29 and $2ad(\mathcal{R}) = \{(37\text{–}39), (46)\}$:

$$P(x,y) \to \exists z.\ P(y,z) \wedge P(z,y) \quad (36)$$
$$P(x,y) \to \exists z.X_{(36)}(y,z) \quad (37)$$
$$X_{(36)}(y,z) \to P(y,z) \quad (38)$$
$$X_{(36)}(y,z) \to P(z,y) \quad (39)$$
$$P(y,z) \wedge P(z,y) \to X_{(36)}(y,z) \quad (46)$$

The $\mathbb{R}$-chase yields the same result as the $\mathbb{E}$-chase on $\mathcal{R}$, so $\mathcal{R} \in CT^{\mathbb{R}}_{\forall\forall}$. We then construct an infinite derivation from $\langle 2ad, P(a,b) \rangle$. First, apply (37), and (38). Then, repeat the following pattern: (37), (38) (on the new variable), then (39) and (46) (on the variables of the previous loop). Applying (37) again before applying (46) yields an infinite chain. $\square$

Again, the $\mathbb{R}$-chase is not well-behaved with respect to atomic decomposition. However, the $\mathbb{DF}$-$\mathbb{R}$-chase behaves exactly as desired regarding the two-way atomic decomposition. In fact, we can show an even stronger result: any $\mathbb{DF}$-$\mathbb{R}$-derivation from a KB $\langle \mathcal{R}, F \rangle$ can be replicated by a $\mathbb{DF}$-$\mathbb{R}$-derivation from $\langle 2ad(\mathcal{R}), F \rangle$, and conversely.

Theorem 37. *The 2-way atomic decomposition has no impact on the (sometimes-)termination of the $\mathbb{DF}$-$\mathbb{R}$-chase; i.e., $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall} = 2ad(CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall})$ and $CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists} = 2ad(CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\exists})$.*

Sketch. One can prove that any fair $\mathbb{DF}$-$\mathbb{R}$-derivation $\mathcal{D}$ from $\langle \mathcal{R}, F \rangle$ with $\mathcal{R} \in CT^{\mathbb{DF}\text{-}\mathbb{R}}_{\forall\forall}$ can be replicated to yield a fair $\mathbb{DF}$-$\mathbb{R}$-derivation $\mathcal{D}'$ from $\langle 2ad(\mathcal{R}), F \rangle$ such that $\mathcal{D}'$ is finite if and only if $\mathcal{D}$ is. The reciprocal is also true. $\square$

The following result follows from Proposition 33

Theorem 38. *: The two-way atomic decomposition preserves the termination of the $\mathbb{E}$-chase.*

The single-piece decomposition may gain termination for some chase variants (Theorem 23); we are interested to know if the same can happen with atomic decompositions. Unfortunately, there is no way for a non-terminating rule set to gain termination, as stated next:

Proposition 39. *If a chase variant does not terminate on a rule set $\mathcal{R}$, it does not terminate on 1ad($\mathcal{R}$) and 2ad($\mathcal{R}$).*

Sketch. For each trigger with a rule R in the original infinite fair derivation, one can consider the corresponding triggers with *1ad*(R) or *2ad*(R), and thus produce an infinite fair derivation. $\square$

Regarding FO-rewritability, the two-way atomic decomposition behaves similarly to the one-way atomic decomposition (which can be proven similarly, see Theorem 30).

Theorem 40. *A rule set $\mathcal{R}$ is BDDP iff 2ad($\mathcal{R}$) is BDDP.*

7 No Normalisation for the Restricted Chase

Normalisation procedures studied so far do not maintain the status of the termination of the $\mathbb{R}$-chase. This raises the question of the existence of such a procedure. We show here that no computable function can map rule sets to sets of rules having atomic head while preserving termination and non-termination of the $\mathbb{R}$-chase. To do that, we show that with atomic-head rules, the class of rule sets $CT^{\mathbb{R}}_{F\forall}$ for which every fair $\mathbb{R}$-derivation from $\langle \mathcal{R}, F \rangle$ is finite is a recursively enumerable set. With arbitrary rules, we show it is hard for Π^0_2, the second level of the arithmetic hierarchy (Rogers 1987). A complete problem for Π^0_2 is to decide whether a given Turing machine halts on every input word; it remains complete when inputs are restricted to words on a unary alphabet.

Proposition 41. *For any factbase F, the subset of $CT^{\mathbb{R}}_{F\forall}$ composed of sets of atomic-head rules is recognizable.*

Sketch. With atomic-head rules, it is known that the existence of an infinite fair restricted derivation is equivalent to the existence of an infinite restricted derivation (Gogacz, Marcinkowski, and Pieris 2020). Using König's lemma, one can show that the chase terminates iff there exists a k such that any fair $\mathbb{R}$-derivation is of length at most k. $\square$

Proposition 42. *There exists a factbase F such that $CT^{\mathbb{R}}_{F\forall}$ is Π^0_2-hard.*[4]

Sketch. Given a Turing machine (TM) M whose input alphabet is unary, we build a KB $\mathcal{K} = \langle \mathcal{R}_w \cup \mathcal{R}_M, F \rangle$ s.t. every fair $\mathbb{R}$-derivation from $\mathcal{K}$ is finite iff M halts on every input. Regardless the chase variant, simulating a TM with a rule set such that the chase terminates whenever the TM halts is classical; we reuse the rule set $\mathcal{R}_M$ provided in (Bourgaux et al. 2021), which we recall in Figure 4 for self-containment. We show that we can assume wlog that all the rules of $\mathcal{R}_M$ are applied after all the rules of $\mathcal{R}_w$ (listed in Figure 3). The set $\mathcal{R}_w$ is used to generate from F arbitrarily large input tape representations in a terminating way. To ensure that any fair $\mathbb{R}$-derivation from $\langle \mathcal{R}_w, F \rangle$ terminates, we reuse the emergency brake technique from (Krötzsch, Marx, and Rudolph 2019), which allows one to stop the derivation at any desired length. The representation of an input word of length j is a set of atoms of the shape $\{\mathtt{Nxt}(c^j_i, c^j_{i+1}), \mathtt{S}_1(c^j_i) \mid 0 \leq i < j\} \cup \{\mathtt{S}_\_(c^j_j), \mathtt{Frst}(c^j_0), \mathtt{End}(c^j_j)\}$. As detailed below, the factbase F contains the representation of the input words of length 0 and 1 (Item 1), atoms used as seeds to build larger words (Item 2) and atoms that initialize the emergency brake (Items 3 and 4):

1. $\mathtt{Frst}(c^1_0), \mathtt{S}_1(c^1_0), \mathtt{Nxt}(c^1_0, c^1_1), \mathtt{End}(c^1_1), \mathtt{S}_\_(c^1_1)$, $\mathtt{Frst}(c^0_0), \mathtt{End}(c^0_0), \mathtt{S}_\_(c^0_0)$
2. $\mathtt{Int}(a), \mathtt{NF}(a, nf_1), \mathtt{R}(nf_1), \mathtt{NF}(nf_1, b), \mathtt{D}(nf_1, b)$

[4] Note that this contradicts the first item of Theorem 5.1 in (Grahne and Onet 2018). However, no proof is given for that statement, which is uncorrectly attributed to (Deutsch, Nash, and Remmel 2008).

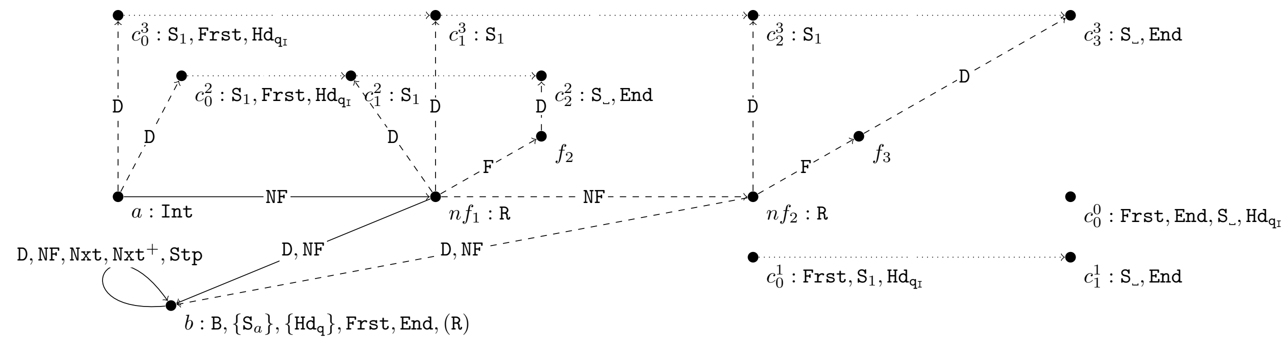

Figure 2: The effect of $\mathcal{R}_w$ on F: Dashed atoms are $\mathtt{D}, \mathtt{F}$ or $\mathtt{NF}$ atoms generated by the chase; $\mathtt{Nxt}$ atoms used by $\mathcal{R}_M$ are dotted.

$$\mathtt{B}(b) \wedge \mathtt{NF}(z,x) \wedge \mathtt{R}(x) \rightarrow \exists y.\mathtt{NF}(x,y) \wedge \mathtt{R}(y) \wedge \mathtt{D}(y,b) \wedge \mathtt{NF}(y,b) \tag{47}$$

$$\mathtt{B}(b) \rightarrow \mathtt{R}(b) \tag{48}$$

$$\mathtt{NF}(x,y) \rightarrow \exists z.\mathtt{F}(y,z) \tag{49}$$

$$\mathtt{F}(x,y) \rightarrow \exists z.\mathtt{D}(y,z) \wedge \mathtt{End}(z) \wedge \mathtt{S}_{\_}(z) \tag{50}$$

$$\mathtt{NF}(t,x) \wedge \mathtt{NF}(x,y) \wedge \mathtt{D}(y,z) \rightarrow \exists u.\mathtt{Nxt}(u,z) \wedge \mathtt{D}(x,u) \wedge \mathtt{S}_1(u) \tag{51}$$

$$\mathtt{NF}(t,x) \wedge \mathtt{F}(x,y) \wedge \mathtt{D}(y,z) \rightarrow \exists u.\mathtt{Nxt}(u,z) \wedge \mathtt{D}(x,u) \wedge \mathtt{S}_1(u) \tag{52}$$

$$\mathtt{Int}(x) \wedge \mathtt{NF}(x,y) \wedge \mathtt{D}(y,z) \rightarrow \exists u.\mathtt{Nxt}(u,z) \wedge \mathtt{D}(x,u) \wedge \mathtt{S}_1(u) \wedge \mathtt{Frst}(u) \tag{53}$$

$$\mathtt{Frst}(x) \rightarrow \mathtt{Hd}_{\mathtt{q}_\mathtt{I}}(x) \tag{54}$$

Figure 3: Rules $\mathcal{R}_w$ to create the initial tapes

3. $\mathtt{B}(b), \mathtt{F}(b,b), \mathtt{NF}(b,b), \mathtt{D}(b,b), \mathtt{Nxt}(b,b), \mathtt{Lst}(b), \mathtt{Frst}(b)$
4. $\mathtt{Hd}_\mathtt{s}(b), \mathtt{S}_l(b), \mathtt{End}(b), \mathtt{Stp}(b,b), \mathtt{Nxt}^+(b,b)$

The chase works as follows: after generating a *non-final* (NF) chain with Rule (47), the *brake* (B) is made *real* (R) by Rule (48), which prevents any extension of the non-final chain through restricted rule applications. A *final* (F) element is added after each non-final element by Rule (49), and from each final element a tape is created, by traversing the chain, marking as *done* (D) processed elements, thanks to Rules (50)-(53). Figure 2 depicts the result of any $\mathbb{R}$-chase derivation from $\langle \mathcal{R}_w, F \rangle$ in which $\mathtt{R}(b)$ has been derived after exactly one application of Rule (47). Rule (54) sets the initial state on the first cell. □

As it is known recursively enumerable sets are a strict subsets of Π_2^0 (Rogers 1987), the following theorem follows.

Theorem 43. *No computable function f exists that maps rule sets to rule sets having atomic-head rules such that $\mathcal{R} \in CT_{\forall\forall}^{\mathbb{R}}$ if and only if $f(\mathcal{R}) \in CT_{\forall\forall}^{\mathbb{R}}$.*

$$\begin{aligned}
\mathtt{Nxt}(x,y) &\rightarrow \mathtt{Nxt}^+(x,y)\\
\mathtt{Nxt}^+(x,y) \wedge \mathtt{Nxt}^+(y,z) &\rightarrow \mathtt{Nxt}^+(x,z)\\
\mathtt{Nxt}(x,y) \wedge \mathtt{Stp}(x,z) \wedge \mathtt{Stp}(y,w) &\rightarrow \mathtt{Nxt}(z,w)\\
\mathtt{End}(x) \wedge \mathtt{Stp}(x,z) &\rightarrow \exists v.\mathtt{Nxt}(z,v) \wedge \mathtt{S}_{_}(v) \wedge \mathtt{End}(v)\\
\mathtt{Hd}_\mathtt{q}(x) \wedge \mathtt{Nxt}^+(x,y) \wedge \mathtt{S}_c(y) &\rightarrow \exists z.\mathtt{Stp}(y,z) \wedge \mathtt{S}_c(z)\\
\mathtt{Hd}_\mathtt{q}(x) \wedge \mathtt{Nxt}^+(y,x) \wedge \mathtt{S}_c(y) &\rightarrow \exists z.\mathtt{Stp}(y,z) \wedge \mathtt{S}_c(z)\\
\mathtt{Hd}_\mathtt{q}(x) \wedge \mathtt{S}_a(x) &\rightarrow \exists z.\mathtt{Stp}(x,z) \wedge \mathtt{S}_b(z)\\
\mathtt{Hd}_\mathtt{q}(x) \wedge \mathtt{S}_a(x) \wedge \mathtt{Stp}(x,z) \wedge \mathtt{Nxt}(z,w) &\rightarrow \mathtt{Hd}_\mathtt{r}(w)\\
\mathtt{Hd}_\mathtt{q}(x) \wedge \mathtt{S}_a(x) \wedge \mathtt{Stp}(x,z) \wedge \mathtt{Nxt}(w,z) &\rightarrow \mathtt{Hd}_\mathtt{r}(w)
\end{aligned}$$

Figure 4: Rules $\mathcal{R}_M$ for the Turing Machine simulation: the last three rules are instantiated w.r.t.the transition function of M.

This applies in particular to normalisation procedures producing rule sets with atomic-head rules.

8 Conclusion

As shown in this paper, normalisation procedures do have an impact, sometimes unexpected, on chase termination. This is particularly true regarding the restricted chase, which is the most relevant in practice but also the most difficult to control. We extend the understanding of its behavior by three results. We show that the Datatog-first strategy is in fact not always the most terminating, which goes against a common belief. We introduce a new atomic-decomposition (two-way), which behaves nicely, in particular regarding the Datalog-first restricted chase, but still has a negative impact on the restricted chase termination. This leads to us to show a more fundamental decidability result, which implies that no computable atomic-decomposition exists that exactly preserves the termination of the restricted chase (i.e., termination and non-termination). Note however that our result does not rule out the existence of a computable normalisation procedure into atomic-head rules that would improve the termination of the restricted chase, although this seems unlikely. Future work includes investigating normalisation procedures for first-order logical formulas, to translate these into the existential rule framework.

Acknowledgements

This work was partly supported by the ANR project CQFD (ANR-18-CE23-0003).

References

Baget, J.; Leclère, M.; Mugnier, M.; and Salvat, E. 2009. Extending decidable cases for rules with existential variables. In *IJCAI*, 677–682.

Baget, J.; Leclère, M.; Mugnier, M.; and Salvat, E. 2011. On rules with existential variables: Walking the decidability line. *Artif. Intell.* 175(9-10):1620–1654.

Beeri, C., and Vardi, M. Y. 1981. The implication problem for data dependencies. In *International Colloquium on Automata, Languages, and Programming*, 73–85. Springer.

Bourgaux, C.; Carral, D.; Krötzsch, M.; Rudolph, S.; and Thomazo, M. 2021. Capturing homomorphism-closed decidable queries with existential rules. In Bienvenu, M.; Lakemeyer, G.; and Erdem, E., eds., *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021, Online event, November 3-12, 2021*, 141–150.

Calì, A.; Gottlob, G.; and Kifer, M. 2008a. Taming the infinite chase: Query answering under expressive relational constraints. In *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning*, KR'08, 70–80. AAAI Press.

Calì, A.; Gottlob, G.; and Kifer, M. 2008b. Taming the infinite chase: Query answering under expressive relational constraints. In Brewka, G., and Lang, J., eds., *Principles of Knowledge Representation and Reasoning: Proc. of the Eleventh International Conference, KR 2008, Sydney, Australia, September 16-19, 2008*, 70–80. AAAI Press.

Calì, A.; Gottlob, G.; and Lukasiewicz, T. 2009. A general datalog-based framework for tractable query answering over ontologies. In *PODS*, 77–86. ACM.

Calì, A.; Gottlob, G.; and Pieris, A. 2012. Towards more expressive ontology languages: The query answering problem. *Artif. Intell.* 193:87–128.

Calvanese, D.; De Giacomo, G.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2007. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *Journal of Automated Reasoning (JAR)* 39(3):385–429.

Carral, D.; Dragoste, I.; and Krötzsch, M. 2017. Restricted chase (non)termination for existential rules with disjunctions. In Sierra, C., ed., *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, 922–928. ijcai.org.

Deutsch, A.; Nash, A.; and Remmel, J. B. 2008. The chase revisited. In Lenzerini, M., and Lembo, D., eds., *Proc. of the 27th ACM SIGMOD-SIGACT-SIGART Symp. on Principles of Database Systems, PODS 2008*, 149–158. ACM.

Fagin, R.; Kolaitis, P. G.; Miller, R. J.; and Popa, L. 2003. Data exchange: Semantics and query answering. In Calvanese, D.; Lenzerini, M.; and Motwani, R., eds., *Database Theory - ICDT 2003, 9th International Conference, Siena, Italy, January 8-10, 2003, Proc.*, volume 2572 of *Lecture Notes in Computer Science*, 207–224. Springer.

Gogacz, T.; Marcinkowski, J.; and Pieris, A. 2020. All-instances restricted chase termination. In Suciu, D.; Tao, Y.; and Wei, Z., eds., *Proceedings of the 39th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2020, Portland, OR, USA, June 14-19, 2020*, 245–258. ACM.

Gottlob, G.; Kikot, S.; Kontchakov, R.; Podolskii, V. V.; Schwentick, T.; and Zakharyaschev, M. 2014. The price of query rewriting in ontology-based data access. *Artif. Intell.* 213:42–59.

Grahne, G., and Onet, A. 2018. Anatomy of the chase. *Fundam. Informaticae* 157(3):221–270.

Grau, B. C.; Horrocks, I.; Krötzsch, M.; Kupke, C.; Magka, D.; Motik, B.; and Wang, Z. 2013. Acyclicity notions for existential rules and their application to query answering in ontologies. *J. Artif. Intell. Res.* 47:741–808.

König, M.; Leclère, M.; Mugnier, M.; and Thomazo, M. 2013. Sound, complete and minimal ucq-rewriting for existential rules. *CoRR* abs/1311.3198.

Krötzsch, M.; Marx, M.; and Rudolph, S. 2019. The power of the terminating chase (invited talk). In Barceló, P., and Calautti, M., eds., *22nd International Conference on Database Theory, ICDT 2019, March 26-28, 2019, Lisbon, Portugal*, volume 127 of *LIPIcs*, 3:1–3:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Leclère, M.; Mugnier, M.; Thomazo, M.; and Ulliana, F. 2019. A single approach to decide chase termination on linear existential rules. In Barceló, P., and Calautti, M., eds., *Proceedings of the 22nd International Conference on Database Theory, ICDT 2019, Lisbon, Portugal*, volume 127 of *LIPIcs*, 18:1–18:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Marnette, B. 2009. Generalized schema-mappings: from termination to tractability. In Paredaens, J., and Su, J., eds., *Proc. of the Twenty-Eigth ACM SIGMOD-SIGACT-SIGART Symp. on Principles of Database Systems, PODS 2009, June 19 - July 1, 2009, Providence, Rhode Island, USA*, 13–22. ACM.

Rocher, S. 2016. *Querying Existential Rule Knowledge Bases: Decidability and Complexity.* Ph.D. Dissertation, University of Montpellier, France.

Rogers, H. 1987. *Theory of Recursive Functions and Effective Computability*. Cambridge, MA, USA: MIT Press.

Tsamoura, E.; Carral, D.; Malizia, E.; and Urbani, J. 2021. Materializing knowledge bases via trigger graphs. *Proc. VLDB Endow.* 14(6):943–956.

Urbani, J.; Krötzsch, M.; Jacobs, C. J. H.; Dragoste, I.; and Carral, D. 2018. Efficient model construction for horn logic with vlog - system description. In *IJCAR*, volume 10900 of *Lecture Notes in Computer Science*, 680–688. Springer.