1 Introduction

Answering conjunctive queries (CQs) over ontologies is an important reasoning task with many applications in knowledge representation and data management. A flurry of research efforts have significantly improved our understanding of this problem, and led to concrete solutions for many ontologies based either on description logics (DLs) [Calvanese et al., 2007; Stefanoni et al., 2014; Calvanese et al., 2014; Bienvenu et al., 2016] or on existential rules [Baget et al., 2011; Cuenca Grau et al., 2013; Cali et al., 2013].

Nevertheless, the problem remains very challenging in theory and in practice. For the popular OWL Web Ontology Language (based on DLs), it is still unknown if the problem is decidable [Rudolph and Glimm, 2010]. For existential rules, which are syntactically closer to CQs, query answering is a basic reasoning task, but it is also undecidable. At the same time, existential rules are too weak to capture the non-deterministic features of ontology languages like OWL.

In this work, we study existential rules with disjunction, which generalise many expressive ontology languages. Our results apply to DLs (and OWL) and many rule languages.

Example 1. The following rules capture basic part-whole relationships (meronymy) and disjunctive information.

\[
\begin{align*}
\text{Bicycle}(x) & \rightarrow \exists v. \text{hasPart}(x,v) \land \text{Wheel}(v) \\
\text{Wheel}(x) & \rightarrow \text{SpokeWheel}(x) \lor \text{DiscWheel}(x) \\
\text{SpokeWheel}(x) & \rightarrow \exists w. \text{partOf}(x,w) \land \text{Bicycle}(w) \\
\text{hasPart}(x, y) & \rightarrow \text{partOf}(y, x) \\
\text{partOf}(x, y) & \rightarrow \text{hasPart}(y, x)
\end{align*}
\]

Similar modelling can be found, e.g., in many medical ontologies. This particular example can also be expressed in OWL.

The oblivious chase provides a sound and complete reasoning algorithm for this logic [Bourhis et al., 2016], computing rule entailments in a bottom-up fashion and branching on disjunctive choices. Unfortunately, the chase often fails to terminate. In Example 1, each bicycle introduces a new wheel (1), which could be a spoke wheel (2), which may introduce a new bicycle (3) – an infinite chain of bikes and wheels. To avoid this, we define a restricted chase that tries to reuse existing elements before introducing new ones.

However, the restricted chase may still fail to terminate, and determining this is undecidable in general [Beeri and Vardi, 1981]. We therefore develop new acyclicity criteria that can ensure termination in the spirit of previous approaches for rules without disjunctions [Marnette, 2009; Krötzsch and Rudolph, 2011; Cuenca Grau et al., 2013]. Existing notions can actually be applied to disjunctive rules by replacing \( \lor \) with \( \land \) for testing acyclicity. However, this only works for an oblivious chase, where non-termination is preserved if rules are logically strengthened. Our restricted chase is more challenging, since additional entailments can lead to termination. Indeed, the restricted chase terminates for Example 1 since rule (3) never needs to be applied to a wheel that was newly introduced by rule (1). Without the rules (4) and (5), this would not be true, and an infinite chase would be required. Only Carral et al. [2016] seem to have studied this setting so far, but their results are specific to the non-disjunctive DL Horn-\( \mathcal{SRQ} \) and do not apply to existential rules.

Acyclicity is sufficient but not necessary for termination. In experiments, we are often left with a significant amount of non-acyclic ontologies of which we cannot say anything. We therefore also develop a cyclicity notion to detect non-termination of the restricted chase. To the best of our knowledge, this is the first proposal for such a criterion. It allows us to show that a majority of practical ontologies is such that the chase will either terminate over all possible sets of input facts, or will necessarily be infinite over at least some inputs.

In summary, our main contributions are:

- we propose restricted joint acyclicity as a simple criterion for restricted chase termination on disjunctive rules,
- we design more general criteria by extending model-faithful acyclicity and model-summarising acyclicity,
we characterise the complexity of query answering and of checking acyclicity for each of the new notions,

- we introduce the first criterion for checking non-termination of the restricted chase,
- we empirically evaluate our results on a large corpus of real-world ontologies.

Experiments suggest that our new notions can significantly improve over the state of the art, and that, moreover, a large subset of non-acyclic ontologies are indeed cyclic. This extended report includes an appendix with additional details on proofs omitted from the conference version.

2 Preliminaries

We consider a standard first-order signature based on mutually disjoint, countably infinite sets of constants $C$, function symbols $F$, variables $V$, and predicates $P$. Every function symbol or predicate $f$ has an arity $ar(f) \geq 0$. Terms are built from $C$, $V$, and $F$ as usual. We write lists of variables $(x_1, \ldots, x_n)$ as $\mathbf{x}$, and similarly for term lists $t$. We treat such lists as sets when order is irrelevant. We will use function symbols only to represent auxiliary terms introduced during reasoning (via skolemisation; see below). An atom is a formula $p(t)$ with $p \in P$ and $ar(p) = |t|$. A term or formula is ground if it contains no variables. A fact is a ground atom. A disjunctive existential rule, or simply rule, is a formula

$$\forall x, y, (B[x, y] \rightarrow \bigvee_{i=1}^{n} \exists v_i.H_i(x, v_i)) \quad (6)$$

where $n \geq 1$, and where $B$ (the body) and $H_i$ (the heads) are non-empty conjunctions of atoms that contain only variables from $x \cup y$ and $x \cup v_i$, respectively, and no constants or function symbols. We omit the universal quantifiers when writing rules. The variables $x$ are called frontier variables. A rule is deterministic if $n = 1$ and nondeterministic otherwise; it is generating if it contains an existential variable and non-generating otherwise. When convenient, we treat conjunctions, such as $B$ and $H_i$, as sets of atoms.

We consider finite sets of rules $R$, where we assume without loss of generality that each variable in $R$ occurs only in the scope of a single quantifier within a single rule (i). An instance $I$ is a finite set of function-free facts. A program is a pair $\langle R, I \rangle$ with $R$ a rule set and $I$ an instance.

A Boolean conjunctive query (BCQ) is a formula $\exists v.Q[v]$, where $Q$ is a conjunction of function-free, constant-free atoms using only variables from $v$. A program $\langle R, I \rangle$ entails a BCQ $\exists v.Q[v]$ if $\bigwedge R \land \bigwedge I \models \exists v.Q[v]$ under standard first-order semantics. It is well known that arbitrary conjunctive query answering can be reduced to BCQ entailment.

A (ground) substitution $\sigma$ is a partial function from variables to (ground) terms. We denote finite substitutions as $[x_1/t_1, \ldots, x_n/t_n]$ or $[x/t]$, and we set $\sigma(x) = x$ if $\sigma$ is undefined for $x$. Given a term or formula $F$, we write $F\sigma$ for the expression obtained by concurrently replacing all unbound occurrences of variables $x$ in $F$ by $\sigma(x)$.

The skolemisation $sk(\rho)$ of a rule $\rho$ as in (6) is the formula $\forall x, y, (B[x, y] \rightarrow \bigvee_{i=1}^{n} sk(H_i)[x])$ where $sk(H_i)$ is obtained from $H_i$ by replacing each variable $v \in v_i$ by the term $f_v(x)$, where $f_v$ is a fresh skolem function symbol specific to $v$ (which, by (i), occurs in only one quantifier).

The Restricted Chase We consider a restricted version of the disjunctive chase [Bourhis et al., 2016], where rules are only applicable if their heads are not satisfied by existing facts. Moreover, we impose an order of rule applications that defers the use of generating rules.

Definition 1. Consider a rule $\rho$ of form (6), a ground substitution $\sigma$ defined exactly on the variables $x \cup y$ from $\rho$, and a set of facts $F$. Then $(\rho, \sigma)$ is applicable to $F$ if (a) $F \models B\sigma$, and (b) $F \not\models \bigvee_{i=1}^{n} \exists v_i.H_i\sigma$. In this case, the result of applying $(\rho, \sigma)$ to $F$ is the set $\{F \cup sk(H_1)\sigma, \ldots, F \cup sk(H_n)\sigma\}$, consisting of all sets of facts obtained from $F$ by adding the skolemised, $\sigma$-instantiated atoms of some head of $\rho$.

Note that condition (a) is the same as $B\sigma \subseteq F$, while (b) states that there is no extension $\sigma$ of $\sigma$ to the variables $v_i$, such that $H_i\sigma \subseteq F_i$, for any $i \in \{1, \ldots, n\}$. The chase is the result of a possibly infinite process of recursive rule applications:

Definition 2. A chase tree of $\langle R, I \rangle$ is a (possibly infinite) tree; each node is labelled by a set of facts, such that:

1. the root is labelled with $I$,
2. if a node labelled $F$ has $n$ children labelled $F_1, \ldots, F_n$, then there is a rule $\rho \in R$ and substitution $\sigma$ such that $\{F_1, \ldots, F_n\}$ is the result of applying $(\rho, \sigma)$ to $F$,
3. if a node $\alpha$ is labelled with $F$ and $(\rho, \sigma)$ is applicable to $F$, then each path starting from $\alpha$ contains a node in which $(\rho, \sigma)$ is no longer applicable due Definition 1 (b),
4. generating rules are only applied in a node where no non-generating rule of $R$ is applicable.

The result of a restricted chase is the (possibly infinite) set of all (possibly infinite) sets of facts that are obtained as the union of all set of facts along some path.

Condition (3) ensures fair, exhaustive application, while (4) improves the rule application strategy to reduce the amount of applicable generating rules. Bourhis et al. [2016] omit (4) and Definition 1 (b), but restricted disjunctive chase algorithms were considered before, e.g., by Deutsch and Tannen [2002].

Example 2. Consider the Example 1, where we use first letters to abbreviate predicates from now on, and the instance $\{B(c)\}$. We obtain a finite chase tree with two leaves $F_1 = \{B(c), hP(c, f_v(c)), W(f_v(c)), pO(f_v(c), c), SW(f_v(c))\}$ and $F_2 = \{B(c), hP(c, f_v(c)), W(f_v(c)), pO(f_v(c), c), DW(f_v(c))\}$. In particular, rule (3) is not applicable to $F_1$, since the facts $pO(f_v(c), c)$ and $B(c)$ already satisfy the head of this rule for the substitution $[x/c]$.

In general, the chase tree and even its result is not unique, since the order of rule applications may matter, but we get the following consequence of well-known results:

Fact 1. A program $\langle R, I \rangle$ entails a BCQ $\exists v.Q$ iff $F \models \exists v.Q$ holds for all sets of facts $F$ in the result of an (arbitrary) restricted chase.

If the chase terminates, the chase tree is finite, and the result is the set of all (finite) leaf labels. In this case, Fact 1 leads to a decision procedure for BCQ entailment. Unfortunately, chase termination is undecidable even for deterministic rules [Beeri and Vardi, 1981]. We therefore study sound but incomplete tests for restricted chase termination.
3 Restricted Joint Acyclicity

We first consider a limited but easy-to-check condition to ensure chase termination. As noted in the introduction, we could apply existing criteria for the deterministic skolem chase, but the result is often unsatisfactory:

Example 3. On the skolemisation of the rules in Example 1, the oblivious chase may produce an infinite set of facts \( \{pO(c, f_{x}(c)), SW(f_{x}(c)), pO(f_{x}(c), e_{1}), SW(f_{x}(c)), hP(f_{x}(c), e_{1}), hP(f_{x}(c), e_{2}), hP(f_{x}(c), e_{1}), f_{x}(c), f_{x}(c), W(f_{x}(c)), f_{x}(c)\} \).

To address such cases, we extend the notion of joint acyclicity (JA), proposed by Krotzsch and Rudolph [2011] for deterministic rules. JA proceeds in two steps: (1) for each existential variable \( v \), we compute a set \( \Omega_v \) of predicate positions, to which values generated for \( v \) might propagate; (2) we build a dependency graph to show if the value generated for a variable \( v \) may participate in the generation of a new value for a variable \( w \). A rule set is JA if this graph has no cycles. To adapt this idea to the restriction chase, we treat \( \lor \) as \( \land \) for (1) and add a new blocking criterion for (2). For the next definition, recall assumption (1) from Section 2.

Definition 3. A position in a predicate \( p \) is a pair \( (p, i) \) with \( i \in \{1, \ldots, ar(p)\} \). A term \( t \) occurs at position \( (p, i) \) in a formula \( \phi \) if \( \phi \) contains an atom \( p(s) \) with \( s_i = t \). Given a rule \( \rho \) and variable \( z \), a body position (head position) of \( z \) is a position in the body (head) of \( \rho \) at which \( z \) occurs.

For a rule set \( \mathcal{R} \) and an existential variable \( v \) of \( \mathcal{R} \), a set \( \Omega_v \) of positions is defined recursively: (i) \( \Omega_v \) contains all head positions of \( v \) in a rule of \( \mathcal{R} \); (ii) for every universally quantified variable \( x \) in \( \mathcal{R} \), if \( \Omega_v \) contains every body position of \( x \), then \( \Omega_v \) also contains every head position of \( x \).

Example 4. For the rules of Example 1, we obtain \( \Omega_v = \{hP(2), \langle W, 1 \rangle, \langle SW, 1 \rangle, \langle DW, 1 \rangle, \langle pO, 1 \rangle\} \) and \( \Omega_w = \{\langle pO, 2 \rangle, \langle B, 1 \rangle, \langle hP, 1 \rangle\} \).

For a set of facts \( F \) and set of rules \( \mathcal{R} \), let \( \mathcal{R}_{dng}(F) \) be the set of facts obtained from \( F \) by exhaustive application of all deterministic, non-generating rules of \( \mathcal{R} \). For an existentially quantified variable \( v \), let \( H_v \) denote the (unique) head conjunction \( v \) occurs in, and let \( B_v \) be the body of the corresponding rule.

Definition 4. The restricted dependency graph of a rule set \( \mathcal{R} \) has the existentially quantified variables \( v \) of \( \mathcal{R} \) as its nodes, and an edge \( v \rightarrow w \) if \( w \) occurs in a rule \( \rho_w \) of form (6) with a frontier variable \( x \in x \) such that

(\( a \)) all body positions of \( x \) occur in \( \Omega_v \), and

\( b \)) for the set \( F = \{B_w \cup H_z[v/x] \cup B_u\} \), where \( \sigma \) replaces all variables \( z \) by distinct constants \( c_z \), we have \( \mathcal{R}_{dng}(F) \models \langle \exists w, H_x \rangle \sigma \).

\( \mathcal{R} \) is restricted jointly acyclic (RJA) if its restricted dependency graph has no cycles.

Example 5. We consider Example 1 and rename variables by subscripts of rule numbers. The existential dependency graph has nodes \( v \) and \( w \). For the potential edge \( v \rightarrow w \), condition (a) of Definition 4 is satisfied for rule (3) and variable \( x_3 \), since \( SW(1) \in \Omega_v \) (see Example 4). However, the set \( F \) of condition (b) is \( \{SW(c_2), P(c_1), SW(c_3)\} \) and the application of deterministic, non-generating rules (4) and (5) yields \( \mathcal{R}_{dng}(F) = F \cup \{pO(c_3, c_1)\} \), which satisfies rule (3). Hence, there is no edge \( v \rightarrow w \). A similar argument rules out \( w \rightarrow v \). For \( v \rightarrow v \) and \( w \rightarrow w \), condition (a) is not satisfied. The graph therefore has no edges and the rule set is RJA.

RJA rules lead to a finite chase tree. The following property is the essence of this claim.

Lemma 2. If the restricted chase tree of \( \langle \mathcal{R}, \mathcal{I} \rangle \) contains a term \( f_{x}(t) \) where \( t_i = f_{x}(s) \) for some \( i \in \{1, \ldots, |t|\} \), then there is an edge \( v \rightarrow w \) in the restricted dependency graph.

Proof. We denote rules etc. as in Definition 4, and assume that \( f_{x}(t) \) was derived by applying \( \langle \rho_w, \theta \rangle \). The sets \( \Omega_v \) estimate the possible positions of \( f_{x}(i) \)-terms, hence the applicability of \( \rho_w \) to \( f_{x}(s) \) implies Definition 4 (a). Moreover, let \( F_w \) be the set of facts as it was when \( \rho_w \) was applied in the chase tree. There is a homomorphism \( h : \mathcal{F} \rightarrow F_w \) with \( h(B_w \sigma) = B_w \theta \) and \( h(B_s \sigma) \) the premise of the application of \( \rho_w \) that produced \( f_{x}(s) \). By our chase strategy, \( \mathcal{R}_{dng}(F_w) = F_w \). Hence, if \( \mathcal{R}_{dng}(F \models \langle H_w \sigma \rangle \), then \( F_w \models \langle h(H_x \sigma) = H_w \theta \rangle \), such that \( \langle \rho_w, \theta \rangle \) would not be applicable. Since it is, we get \( \mathcal{R}_{dng}(F) \not\models \langle H_w \sigma \rangle \) as claimed.

Theorem 3. Deciding BCQ entailment for programs \( \langle \mathcal{R}, \mathcal{I} \rangle \) where \( \mathcal{R} \) is RJA is coN2EXPTime-complete, even if the arity of predicates is bounded.

Proof sketch. Membership follows since BCQ non-entailment can be shown by finding a model of \( \langle \mathcal{R}, \mathcal{I} \rangle \) that does not satisfy the query. For this we may nondeterministically guess a branch of the chase tree. The maximal nesting depth of function terms in the chase tree is bounded by the number of existentially quantified variables in \( \mathcal{R} \), since a greater depth can only be achieved by repeating a function symbol, which would make the restricted dependency graph cyclic by Lemma 2. The maximal number of terms of linear depth is doubly exponential, so there are double exponentially many possible ground facts overall. A set of facts of this size can be computed in 2EXPTime.

Hardness is established by modifying the construction of a 2EXPTime Turing machine given for deterministic, weakly acyclic rules by Cali et al. [2010]. The construction yields a grid of doubly exponential size, using predicates of arity \( \leq 3 \). Using disjunction in rules, it is not hard to simulate a nondeterministic Turing machine in the same way.

Theorem 4. Deciding if \( \mathcal{R} \) is RJA is EXPTime-complete, coNP-complete if the arity of predicates is bounded, and P-complete if the number of variables per rule is bounded.

Proof. Sets \( \Omega_v \) can be computed in polynomial time, and there are only polynomially many possible edges and body variables \( x \) as in Definition 4 to be considered. However, Definition 4 (b) corresponds to the EXPTime-complete Datalog reasoning task of checking non-entailment of a set of facts [Dantsin et al., 2001]. The task becomes coNP-complete
for predicates of bounded arity: hardness follows from hardness of conjunctive query entailment (rule bodies are CQs); membership follows since there are only polynomially many ground facts over this signature, hence the derivation of any such fact can be represented as a polynomial directed acyclic graph with (true) facts as nodes and edges connecting premises with conclusions, such that fact entailment can be checked by guessing this graph and verifying each rule application. The task becomes P-complete for bounded numbers of variables, since grounding (instantiation of rules with constants) polynomially reduces fact entailment to propositional Horn logic entailment. In all cases, detecting cycles in the (polynomial) dependency graph is possible in P.

The complexity of Theorem 4 is dominated by the reasoning for condition (b) in Definition 4. In practice, this reasoning task is usually fairly simple, since only a very small number of facts is given, and most rules can be ignored.

4 Restricted Model-Faithful Acyclicity

While RJA is fairly easy to check, it is not sufficient to capture all realistic cases. We therefore develop more general, though also more complex criteria.

Example 6. We extend Example 1 with the following rules:

\[
\begin{align*}
\text{SpokeWheel}(x_0) & \rightarrow \exists u \text{.hasPart}(x_0, u) \wedge \text{Spoke}(u) \\
\text{Spoke}(x_7) & \rightarrow \exists z \text{.partOf}(x_7, z) \wedge \text{Bicycle}(z) \\
\text{hasPart}(x_9, y_9) & \wedge \text{hasPart}(y_8, z_8) \rightarrow \text{hasPart}(x_9, z_8)
\end{align*}
\]

The resulting rule set still leads to a finite restricted chase for all instances, but it is not RJA. Indeed, the restricted dependency graph contains a cycle \( v \rightarrow u \rightarrow z \rightarrow v \). For example, when considering \( z \rightarrow v \), the set \( \mathcal{F} \) in Definition 4 (b) is \( \{B(c_1), pO(c_7, c_1), S(c_7)\} \), which cannot entail \( W(c_1) \).

For cases as in Example 6, we extend the notion of model-faithful acyclicity (MFA) [Cuenca Grau et al., 2013]. To determine if a set of deterministic rules \( \mathcal{R} \) is MFA, one computes the chase on \( (sk(\mathcal{R}), I_\mathcal{R}) \), where \( I_\mathcal{R} \) is the critical instance, which contains all possible ground facts based on predicates of \( \mathcal{R} \) and the single constant symbol \( * \). \( \mathcal{R} \) is MFA if this chase terminates without introducing a cyclic term \( f(t) \), which is such that \( f \) occurs in the terms \( t \). Alternatively, a cyclic term must appear after at most doubly exponentially many steps. Deciding MFA indeed is 2ExpTime-complete.

MFA uses the fact that the chase terminates on every instance if it terminates on the critical instance. This is not true for the restricted chase, as no rule is applicable in the presence of the critical instance. We therefore consider a relaxed condition of applicability, which, in the spirit of Definition 4 (b), determines a rule’s applicability from a smaller set of facts not including the whole critical instance. The body of the rule (instance) that is to be applied can always be assumed as given. Further facts can be obtained from this body’s skolem terms, since each skolem function is introduced by one specific rule:

Definition 5. For a rule set \( \mathcal{R} \) and a ground term \( t = f_a(s) \) using skolem functions from \( sk(\mathcal{R}) \), the set \( \mathcal{F}_i \) contains all ground facts involved in the derivation of facts containing \( t \):

(1) Let \( B[x_1, y] \rightarrow \bigvee_i \exists v_i. H_i[x_1, v_i] \) be the unique rule that contains \( v \) in head disjunct \( H_i \), and consider the substitution \( \theta = [x/s, y/e] \), where \( e \) is a list of fresh constant symbols not used elsewhere. Then \( B \theta \cup sk(H_0) \theta \subseteq \mathcal{F}_i \).

(2) For every functional term \( s_j \in s \), we have \( \mathcal{F}_{s_j} \subseteq \mathcal{F}_i \).

Example 7. For the rules from Examples 6 and 1, and term \( t = f_a(f_a(b)) \), we have \( \mathcal{F}_t = \{SW(f_a(b)), hP(f_a(b), t), S(t)\} \cup \mathcal{F}_{f_a(b)} \) with \( \mathcal{F}_{f_a(b)} = \{B(b), hP(f_a(b), f_a(b)), W(f_a(b))\} \).

The next example illustrates another difficulty: even if we only take a rule’s body into account to check its applicability, we might get much fewer derivations on the critical instance than on other instances.

Example 8. Consider the rule \( \rho : p(x, y) \rightarrow \exists v.p(v, v) \vee \exists w.p(y, w) \). On the critical instance, one could try to apply \( \rho \) with substitution \( [x/\star, y/\star] \). But already the instantiated body \( p(\star, \star) \) prevents the application of \( \rho \), since \( p(\star, \star) \models \exists v.p(v, v) \). The restricted chase terminates immediately. Yet, on the instance \( \{p(a, b)\} \), one can compute an infinite set of facts \( \{p(a, b), p(b, f_a(b)), p(f_a(b), f_a(f_a(b)))\} \), so the restricted chase is not finite in general.

To handle this issue, we rename distinct occurrences of \( \star \).

Definition 6. Consider a rule \( \rho : B[x_1, y] \rightarrow \bigvee_i \exists v_i.H_i[x_1, v_i] \), and a ground substitution \( \sigma \) defined exactly on \( x \cup y \). Let \( \rho' \) be such that, for all \( x \in \mathcal{V} \), \( \rho'(x) \) is a function with each occurrence of a constant renamed so that no constant occurs more than once in the image of \( \rho' \). The set \( B_{\rho, \sigma} \) is the union of \( B_{\rho'} \) and each of the sets \( \mathcal{F}_i \) for which there is a skolem term \( t \) in \( B_{\rho'} \). We say that \( (\rho, \sigma) \) is blocked if \( \mathcal{R}_{\mathcal{R}}(B_{\rho, \sigma}) \models \bigvee_i \exists v_i.H_i \sigma' \).

Example 9. Consider the rules of Example 6, and especially rule \( p(8) \) under substitution \( \{x_7/t\} \) with \( t = f_a(f_a(b)) \). Intuitively speaking, \( t \) represents a spoke that was introduced as part of wheel \( f_a(b) \), which in turn is part of bicycle \( b \). We want to show that \( \rho \) does not need to be applied to introduce another bicycle \( f_a(f_a(b)) \). We do not need to rename any constants here, so \( B_{\rho, \sigma} = \{S(t)\} \cup \mathcal{F}_i \) as in Example 7. While \( B_{\rho, \sigma} \) does not satisfy the head of \( \rho \) yet, we get \( \mathcal{R}_{\mathcal{R}}(B_{\rho, \sigma}) = B_{\rho, \sigma} \cup \{hP(b, t), pO(t, f_a(b)), pO(f_a(b), f_a(f_a(b)))\} \). Therefore \( \mathcal{R}_{\mathcal{R}}(B_{\rho, \sigma}) \models \exists z.pO(t, z) \wedge B(z) \) and \( (\rho, \sigma) \) is blocked as expected.

Our adaptation of MFA conducts a (deterministic) chase on the critical instance, but applies rules only if not blocked.

Definition 7. For a rule set \( \mathcal{R} \), \( \text{RMFA}(\mathcal{R}) \) is the least set of facts for which \( I_\mathcal{R} \subseteq \text{RMFA}(\mathcal{R}) \) and, whenever \( \rho : B \rightarrow \bigvee_i \exists v_i.H_i \) is a rule in \( \mathcal{R} \), and \( \sigma \) is such that \( B_\sigma \subseteq \text{RMFA}(\mathcal{R}) \) and \( (\rho, \sigma) \) is not blocked, then \( sk(H_1) \sigma \cup \ldots \cup sk(H_n) \sigma \subseteq \text{RMFA}(\mathcal{R}) \). \( \mathcal{R} \) is restricted model-faithfully acyclic (RMFA) if \( \text{RMFA}(\mathcal{R}) \) does not contain a cyclic term.

Example 10. The rules of Examples 1 and 6 together are RMFA, as one can easily check along the lines of Example 9.

Theorem 5. Deciding if \( \mathcal{R} \) is RMFA is 2ExpTime-complete even if the arity of predicates or the number of variables per rule is bounded. It is ExpTime-complete if each rule contains at most one frontier variable.
Proof sketch. Membership in 2Exptime follows as in Theorem 3 by bounding the possible ground skolem terms. Similarly, rules with one frontier variable lead to unary skolem functions, which can form only exponentially many terms.

2Exptime-hardness can be shown as for the case of MFA, where it was done by reduction from the 2Exptime-hard problem of BCQ entailment checking for weakly acyclic (WA) rules [Cuenca Grau et al., 2013]. The hardness proof for WA in turn is based on a direct Turing machine construction using predicates of bounded arity and rules with a bounded number of variables [Cali et al., 2010]. One may verify that no rule application is ever blocked in this particular construction of Definition 8. The recursive case (2) for subterms becomes irrelevant).

As nullary "skolem" function symbols in Definition 5 (the explicit during the chase instead of using cyclic terms. Definition 6, any chase derives at most doubly exponentially many terms, and σ preserves non-cyclicity. Hardness follows by Theorem 3, as any RJA rule set can be shown to be RMFA (cf. the relation of JA and MFA [Cuenca Grau et al., 2013]).

Theorem 5 motivates the search for a simpler test that still extends RJA. We can achieve this by adapting model-summarising acyclicity (MSA) to our setting [Cuenca Grau et al., 2013]. This criterion resembles MFA in that a chase on the critical instance is conducted to discover cycles. However, instead of using skolem terms, existential variables now are replaced by fresh constants, and cycles are tracked explicitly during the chase instead of using cyclic terms. Definition 6 can remain unchanged if we treat the fresh constants as nullary "skolem" function symbols in Definition 5 (the recursive case (2) for subterms becomes irrelevant).

Definition 8. For a rule set R, let S be a binary predicate not used in R, and let θ be a substitution that maps each existentially quantified variable v in R to a unique fresh constant cv. RMSA(R) is the least set of facts for which B(R) ⊆ RMSA(R) and, whenever ρ : B[x, y] → ∃v1.H1[x, v1] is a rule in R, and σ is such that Bσ ⊆ RMSA(R) and ⟨ρ, σ⟩ is not blocked, then H1θσ ∪ . . . ∪ Hnθσ ∪ {S(xσ, vθ) | x ∈ x, v ∈ v1, 1 ≤ i ≤ n} ⊆ RMSA(R).

R is restricted model-summarising acyclic (RMSA) if RMSA(R) does not contain a directed cycle of S-relations.

BCQ answering remains as hard as for RJA and RMFA, but recognising RMSA is only as hard as for RJA. The proof is similar to the proofs of Theorems 4 and 5.

Theorem 8. Deciding if R is RMSA is Exptime-complete, and P-complete if the number of variables per rule is fixed.

Example 11. RMSA cannot capture Example 6, but it generalises RJA. Consider the set R of the rules (1), (2), (7), and

\[ \text{hasPart}(x, y) \land \text{BicycleChain}(y) \rightarrow \text{Bicycle}(x). \] (10)

\( R \) is RMSA since rule (10) is never applicable to fresh constants (which do not have parts that are bicycle chains). However, \( R \) is not RJA since (10) leads to \( \{ \text{Bicycle}, 1 \} \in \Omega. \)

5 Proving Nontermination

Even if rules are not acyclic by any of our criteria, they might still have a finite chase. In this section, we introduce a complementary criterion that is sufficient (but not necessary) to show that the chase will be infinite. By combining this with our previous acyclicity notions, we hope to decide the question of chase termination for most practical ontologies.

Sufficient conditions for nontermination can also look for cycles, e.g., by detecting cyclic terms in a chase as done for MFA. The critical instance cannot be used here, since it overestimates what can really be derived repeatedly. However, a cyclic term \( f_v(\ldots f_v(t) \ldots) \) might indicate nontermination if it was derived in a chase that started from nothing but the facts \( B_v \cup sk(H_v) \), where we replace each variable \( z \) by a fresh constant \( c_z \) (\( B_v \) and \( H_v \) are the unique body and head for \( v \) as in Section 3). Indeed, if this happens, then each application of the rule of \( v \) creates a set of facts that eventually enables another application of the same rule that will produce a skolem term of increased nesting depth. This discussion suggests a criterion for showing nontermination of the skolem chase:

Definition 9. Consider a set \( R \) of deterministic rules, and a rule \( \rho : B[x, y] \rightarrow \exists w.H(x, v) \) in \( R \). The set \( Z_\rho \) is obtained from \( B \cup sk(H) \) by replacing all occurrences of every variable \( z \) with a fresh constant \( c_z \). Let \( F_\rho \) be the set of facts obtained by exhaustive application of rules from \( sk(R) \) to \( Z_\rho \) under the condition that no rule is applied to facts with a cyclic term. Then \( R \) has a \( p \)-cycle if there is a variable \( v \in v \) such that \( F_\rho \) contains a cyclic term \( f_w(\ldots f_v(t) \ldots) \). \( R \) is model-faithful cyclic (MFC) if it has a \( p \)-cycle for a generating rule \( \rho \) in \( R \).

Excluding facts with cyclic terms in the computation of \( F_\rho \) is necessary to ensure termination, since otherwise that chase might be infinite although none of the cyclic terms have the form \( f_w(\ldots f_v(t) \ldots) \) with \( v \in v \). With this restriction, however, MFC can be checked in 2Exptime.

Example 12. The rules in Example 1 are MFC if we replace \( \lor \) by \( \land \) to make them deterministic. Let \( \rho \) be rule (1). Then \( Z_\rho = \{ B(c), hP(c, f_v(c)), W(f_v(c)) \} \) and \( F_\rho \) contains the facts \( SW(f_v(c)), po(f_v(c), f_w(f_v(c))), B(f_w(f_v(c))), \) and \( hP(f_w(f_v(c)), f_v(f_w(f_v(c)))) \), which has a cyclic term.
Using the argumentation given before Definition 9, we can establish the following result.

**Theorem 9.** If $\mathcal{R}$ is MFC, then there is an instance $\mathcal{I}$ for which the chase of $(sk(\mathcal{R}), \mathcal{I})$ is infinite.

Unfortunately, the approach of performing a chase on minimal sets of facts fails when the application of rules can be prevented by the presence of additional facts. We therefore perform an additional applicability check for the restricted chase. Recall that $\mathcal{T}_\mathcal{R}$ denotes the critical instance for $\mathcal{R}$.

**Definition 10.** Consider a set $\mathcal{R}$ of deterministic rules, a rule $\rho : B[x, y] \rightarrow \exists v.H[x, v] \in \mathcal{R}$, and a ground substitution $\sigma$ on $x \cup y$. Let $\sigma^*$ be such that $\sigma^*(z) = \sigma(z)$ with all constants replaced by $*$ for all $z \in V$. Let $\mathcal{R}^*$ be $\mathcal{R}$ with existential quantifiers omitted and existential variables replaced by $*$, and let $\rho^* \in \mathcal{R}^*$ be the rule obtained from $\rho$.

We define $\mathcal{U}_{\rho, \sigma, \mathcal{R}} = \mathcal{T}_\mathcal{R} \cup B\sigma^* \cup \bigcup_{t \in (x \cup y)\sigma^*} \mathcal{F}_t$. The set $\mathcal{U}_{\rho, \sigma, \mathcal{R}}$ is obtained by exhaustive application of rules from $\mathcal{R}^*$ to $\mathcal{I}_\rho$, with the exception of the rule $\rho^*$ under substitution $\sigma^*$. Then $(\rho, \sigma)$ is unblockable for $\mathcal{R}$ if $\mathcal{U}_{\rho, \sigma, \mathcal{R}} \not\models \exists v.H\sigma^*$.

Intuitively speaking, $\mathcal{U}_{\rho, \sigma, \mathcal{R}}$ represents a gross overestimation of what might be derivable in a situation where we would like to apply a rule $\rho$. Note that, for $(\rho, \sigma)$ to be unblockable, the body $B\sigma$ must contain a functional term.

**Lemma 10.** Consider the chase tree for a program $(\mathcal{R}, \mathcal{I})$, let $\mathcal{F}$ be the label of some node $n$ in this tree, and let $\mathcal{R}_\lambda$ be $\mathcal{R}$ with $\lor$ replaced by $\land$. If $\rho : B[x, z] \rightarrow \exists v.H[x, v] \in \mathcal{R}$ is a deterministic rule in $\mathcal{R}$, $\mathcal{F} \models B\sigma$, and $(\rho, \sigma)$ is unblockable for $\mathcal{R}_\lambda$, then $(\rho, \sigma)$ is applied in a node above or below $n$.

**Proof sketch.** For a contradiction, suppose that the preconditions hold but $(\rho, \sigma)$ is not applied. Since applications are fair, Definition 1 (b) is violated in all leaf nodes below $n$. For any such leaf $n'$ labelled $\mathcal{F}'$, we map $\mathcal{F}'$ to $\mathcal{U}_{\rho, \sigma, \mathcal{R}}$ as follows: every term $t$ in $B\sigma^*$ (including subterms) is mapped to itself; all other terms are mapped to $*$. One can show by induction over the chase of $\mathcal{F}'$ that this is a homomorphism. Since $(\rho, \sigma)$ is not applicable to $\mathcal{F}'$, we find $\mathcal{U}_{\rho, \sigma, \mathcal{R}} \not\models \exists v.H\sigma^*$.

**Definition 11.** For a set $\mathcal{R}$ of rules and a rule $\rho \in \mathcal{R}$, we define $\mathcal{I}_\rho$ and $\mathcal{F}_\rho$ as in Definition 9, but with the additional restriction that the computation of $\mathcal{F}_\rho$ uses only applications of deterministic rules that are unblockable for $\mathcal{R}_\lambda$, obtained from $\mathcal{R}$ by replacing $\lor$ with $\land$.

As before $\mathcal{R}$ has a restricted $\rho$-cycle if there is a variable $v \in V$ such that $\mathcal{F}_\rho$ contains a cyclic term $f_0(\ldots \ f_1(\ldots t \ldots ))$. $\mathcal{R}$ is restricted model-faithful cyclic (RMFC) if it has a $\rho$-cycle for some generating rule $\rho \in \mathcal{R}$.

The following result is obtained by combining the arguments for Theorem 9 with the insights from Lemma 10.

**Theorem 11.** If $\mathcal{R}$ is RMFC, then there is an instance $\mathcal{I}$ for which the restricted chase of $(\mathcal{R}, \mathcal{I})$ is infinite.

## 6 Evaluation

To evaluate the effectiveness of our criteria, we have used MOWL.Corp, a large corpus of real-world OWL ontologies [Matentzoglu and Parsia, 2014; Matentzoglu et al., 2013], which we transformed into rules. To this end, we first normalised ontologies by structural decomposition of complex axioms, and then rewrote axioms into first-order logic to obtain rules. We refer to Cuenca Grau et al. [2013] for details on this standard process; our normal forms are as in their Table 1 but with an added form $A_1 \cap \ldots \cap A_n \subseteq B_1 \cup \ldots \cup B_m$.

We excluded ontologies with nominals (oneOf) and at-most-restrictions (maxCardinality) since they require equality reasoning. There are well-known techniques for this [Cuenca Grau et al., 2013], but they are not our focus. We then considered all ontologies with up to 1,000 existential quantifiers after normalisation, leading to a set of 1,576 ontologies.

We have implemented tests for RMSA, RMFA, RMFC, MSA, and MFA using RDFox [Motik et al., 2014] as a rule engine. The creation of new terms during the chase and our blocking conditions are implemented on top of RDFox. For MSA and MFA, we replaced $\lor$ by $\land$ in all ontologies. We treat $\top$ (universal class) and $\bot$ (empty class) as regular unary predicates, and we modify our tests to ensure that all elements are always in $\top$; likewise for the universal and the empty property. Special treatment of $\bot$ is not needed, as inconsistencies are expected and should be ignored when chasing from the critical instance. We have not implemented RJA since we found the more general RMSA to perform well.

Table 1 shows our results for ontologies without (top) and with (bottom) disjunctions, grouped by their number of existential axioms ($\#\exists$). The column “#” gives the number of ontologies per group; “open” counts ontologies that are neither RMFA nor RMFC. We can see that RMSA performs better than MFA, while (R)MFA hardly improves over (R)MSA. Using MFA, chase termination remains open for 602 ontologies overall (38.2%). The combination of RMFA and RMFC reduces this number to 250 (15.8%). As expected, many ontologies are indeed cyclic, but there are also an additional 85 that are acyclic (14.1% of the formerly open ones). In the deterministic case, our notions perform rather well and allow us to characterize 96.3% ontologies as acyclic or cyclic.

### Table 1: Experimental results

<table>
<thead>
<tr>
<th>#\exists</th>
<th># MSA</th>
<th>MFA</th>
<th>RMSA</th>
<th>RMFA</th>
<th>RMFC</th>
<th>open</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–5</td>
<td>443</td>
<td>293</td>
<td>293</td>
<td>314</td>
<td>314</td>
<td>127</td>
</tr>
<tr>
<td>5–69</td>
<td>368</td>
<td>243</td>
<td>243</td>
<td>272</td>
<td>272</td>
<td>72</td>
</tr>
<tr>
<td>70–1K</td>
<td>409</td>
<td>348</td>
<td>348</td>
<td>350</td>
<td>350</td>
<td>40</td>
</tr>
<tr>
<td>1–1K</td>
<td>1220</td>
<td>884</td>
<td>884</td>
<td>936</td>
<td>936</td>
<td>239</td>
</tr>
</tbody>
</table>

## 7 Conclusion

To the best of our knowledge, this is the first systematic study of termination of the restricted chase on existential rules (with disjunctions) and the first ever approach to restricted chase nontermination. We have shown our criteria to be theoretically and empirically more general than previous notions, deciding termination for 84.2% of the tested ontologies.
Our work motivates and enables further research on chase-based reasoning procedures for ontologies. Many tableau-based OWL reasoners already implement chase-like algorithms that could be a starting point. In an early test with a modified version of the reasoner HermiT [Motik et al., 2009], we have already answered conjunctive queries over an acyclic ontology with tens of thousands of facts. We believe this is a highly promising direction in description logics ontologies and existential rules alike.

References


A Proofs for Section 3

**Theorem 3.** Deciding BCQ entailment for programs \((\mathcal{R}, \mathcal{I})\) where \(\mathcal{R}\) is RJA is \text{coN2EXPTIME}-complete, even if the arity of predicates is bounded.

**Proof.** Towards membership, we determine the maximal number of ground (skolem) terms and corresponding facts that may occur in the chase. Let \(n\) be the number of skolem functions in \(sk(\mathcal{R})\), and let \(m\) be the maximal arity of such functions. The maximal nesting depth of ground terms in the chase is \(n\), since a greater depth can only be achieved by repeating a function symbol, which would make the restricted dependency graph cyclic by Lemma 2. Ground skolem terms then correspond to trees of depth at most \(n\), fan-out at most \(m\), and with leaves from the set \(C_\mathcal{I}\) of constants in \(\mathcal{I}\). Such trees have at most \(m^n\) leaves, and at most \(n \cdot m^n\) nodes in total. As each node is assigned a constant or function symbol, there are at most \(T = (|C_\mathcal{I}| + n)^{n \cdot m^n}\) trees, and hence ground skolem terms, overall.

Now if \((\mathcal{R}, \mathcal{I})\) contains \(k\) different predicate symbols of arity at most \(\ell\), then the maximal number of ground facts based on \(T\) terms is \(A = kT^{\ell^r} = k(|C_\mathcal{I}| + n)^{\ell^{n \cdot m^n}}. A\) is therefore double exponential in (measures that depend linearly on) the size of \((\mathcal{R}, \mathcal{I})\).

Membership then follows since BCQ non-entailment can be shown by finding a model of \((\mathcal{R}, \mathcal{I})\) that does not satisfy the query. For this we may non-deterministically guess a branch of a chase tree which, by Definition 2, may at most have depth \(A\).

For hardness, we reduce the word problem of double-exponentially time-bounded non-deterministic Turing machines (TMs) to BCQ non-entailment. Consider a N2EXPTIME Turing Machine (TM) \(M\). We simulate the computation of \(M\) on an input string \(I\) by constructing a program \((\mathcal{R}, \mathcal{I})\) such that \((\mathcal{R}, \mathcal{I})\) does not entail some nullary predicate \text{Reject} if and only if \(M\) accepts \(I\).

To address computation steps and tape cells, we recall a construction by Cali et al. to (deterministically) construct a chain of double exponentially many elements. Let \(\mathcal{I} = \{r_0(0), r_0(1), succ_0(0,1), min_0(0), max_0(1)\}\). For each \(i \in \{0, \ldots, n-1\}\), with \(n\) the length of the input \(I\), we add the following rules:

\[
\begin{align*}
    r_i(x) \land r_i(y) &\rightarrow \exists z s_i(x, y, z) \\
    s_i(x, y, z) &\rightarrow r_{i+1}(z) \\
    s_i(x, y, z) \land s_i(x, y', z') \land succ_i(y, y') &\rightarrow succ_{i+1}(z, z') \\
    s_i(x, y, z) \land s_i(x', y', z') \land max_i(y) \land min_i(y') &\land succ_i(x, x') \rightarrow succ_{i+1}(z, z') \\
    min_i(x) \land s_i(x, x, y) &\rightarrow min_{i+1}(y) \\
    max_i(x) \land s_i(x, x, y) &\rightarrow max_{i+1}(y)
\end{align*}
\]

It can be shown, by induction on \(i\), that in any path of any chase tree of \((\mathcal{R}, \mathcal{I})\), the relation \(r_n\) contains \(2^{2^n}\) elements, which are linearly ordered by \(succ_n\).

The remaining TM simulation follows standard constructions (cf. [Dantsin et al., 2001]), using elements of the \(r_n\) chain to refer to specific time points and tape cells when encoding a run of the TM. Non-deterministic transitions are captured using rules with disjunction. Assuming that the state of \(M\) at step \(s\) is captured with facts \(State_q(s)\) for all states \(Q\), we can complete the simulation by adding rules:

\[
State_q(s) \land \max_n(s) \rightarrow \text{Reject}
\]

for all non-accepting states \(q\) of \(M\). We can assume without loss of generality that \(M\) runs for the maximum double-exponential number of steps on all rejecting runs, so that the query \text{Reject} is entailed if and only if there are no accepting runs.

As observed by Cali et al. [2010], the resulting rules are weakly acyclic (WA), and in particular this holds for the rules used to construct \(r_n\). In consequence, the rules are also RJA, since RJA generalises joint acyclicity which in turn generalises WA.

B Proofs for Section 4

**Theorem 5.** Deciding if \(\mathcal{R}\) is RMFA is \text{2EXPTIME}-complete even if the arity of predicates or the number of variables per rule is bounded. It is \text{EXPTIME}-complete if each rule contains at most one frontier variable.

**Proof.** Membership is shown as for Theorem 3, since we obtain the same upper bounds for the number of ground skolem terms that are not cyclic. The chase is deterministic in this case, leading to the claimed complexity. Note that the problem of deciding whether a rule and a substitution are blocked is in \text{EXPTIME}, since it only requires only reasoning with deterministic non-generating rules (corresponding to Datalog [Dantsin et al., 2001]). Performing blocking checks therefore does not worsen the overall complexity of the RMFA check.

Rules with at most one frontier variable lead to unary skolem functions. In consequence, only exponentially many terms and atoms may occur in the chase of programs with rule sets satisfying such restriction. Indeed, the size of the frontier limits the arity of skolem functions, such that ground skolem terms are now trees of fan-out 1, i.e., words, with at most \(n\) function symbols or constants at each node. Therefore, deciding RMFA membership of a one frontier variable rule sets is in \text{EXPTIME}. As before, blocking checks do not worsen this complexity.

\text{2EXPTIME}-Hardness can be shown as for the case of MFA, where it was done by reduction from the \text{2EXPTIME}-hard problem of BCQ entailment checking for weakly acyclic (WA) rules [Cuenca Grau et al., 2013]. The hardness proof for WA in turn is based on a direct Turing machine construction using predicates of bounded arity and rules with a bounded number of variables [Cali et al., 2010]. One may verify that no rule application is ever blocked in this particular construction and thus, we can use the same argument for our purposes. \text{EXPTIME}-hardness can be obtained using a known proof for the case of MFA, which sports the same complexity for small frontiers [Cuenca Grau et al., 2013]. In particular, the construction in their Lemma 59 works even when using blocking because, as in the previous case, no rule application is ever blocked.
Lemma 6. Let $\pi^*$ be the function that maps a term $t$ to the term obtained from $t$ by replacing all constants with $\ast$. Then, for every term $t$ that occurs in any restricted chase of $(R, I)$ for any instance $I$, the term $\pi^*(t)$ occurs in $\text{RMFA}(R)$.

Proof. Consider any set of facts $F$ that is the label of some node in the restricted chase, and which contains a term $t$. We show that $\pi^*$ is a homomorphism from $F$ to $\text{RMFA}(R)$, which implies that $\pi^*(t)$ occurs in $\text{RMFA}(R)$.

We proceed by induction along the rule applications that have produced $F$. The base case is clear since $\pi^*(t)$ is a homomorphism from $I$ to $\text{RMFA}(R)$. For the induction step, assume that the result of applying rule $\rho : B \rightarrow \bigvee_i \exists v_i.H_i$ under substitution $\sigma$ to a previous set of facts $F'$, and in particular that $F = F' \cup \text{sk}(H_k)\sigma'$ for some index $k$. By the induction hypothesis, $\pi^*$ is a homomorphism from $F'$ to $\text{RMFA}(R)$. Therefore $\pi^*(\rho, \sigma) \subseteq \text{RMFA}(R)$.

Let $\sigma'$ be such that $\sigma'(x) = \pi^*(\sigma(x))$ for all $x \in V$. We show that $(\rho, \sigma')$ is not blocked. Indeed, consider the set $B_{\rho, \sigma'}$ from Definition 6, and let $\sigma'$ be the substitution from this definition, i.e., $\sigma'(x) = \sigma'(x)$ with each occurrence of a constant renamed. Then $B_{\rho, \sigma'}$ can be obtained from $B_{\rho, \sigma}$ by replacing each of the renamed constants $c$ that occurs at a certain position in $B_{\rho, \sigma'}$ by the constant $\eta(c)$ that occurs at the same position in $B_{\sigma}$. We can extend this mapping $\eta$ to terms and atoms that contain constants from $B_{\rho, \sigma}$. We claim that $\eta$ can be further extended to a homomorphism from $B_{\rho, \sigma'}$ to $F'$. Indeed, $\eta(B_{\rho, \sigma'}) = B_{\sigma'} \subseteq F'$ by definition. This also implies that each skolem (sub)term $f_\ell(s)$ in $B_{\sigma'}$ has a corresponding term $\eta(f_\ell(s)) = f_\ell(\eta(s))$ that occurs in $F'$. This can only be if the rule $\rho' : B' \rightarrow \bigvee_i \exists v_i.H'_i$ that introduces $f_\ell(\eta(s))$ has been applied in $F'$, which requires that $(B' \cup \text{sk}(H'_i))\theta \subseteq F'$ for the head $H'_i$ that contains $v$ and some substitution $\theta$ that maps the frontier $\ell'$ of $\rho'$ to $\eta(s)$. The set $F_{f_\ell(s)}$ considered in Definition 6 and defined in Definition 5 may contain additional constants $c_v$ introduced for non-frontier variables $v$; we extend $\eta$ to map each such constant $c_v$ to $\theta$. This yields $\eta(f_\ell(s)) = (B' \cup \text{sk}(H'_i))\theta \subseteq F'$, and finishes our construction of the homomorphism $\eta : B_{\rho, \sigma'} \rightarrow F'$.

Now suppose for a contradiction that $(\rho, \sigma, c)$ is blocked, i.e., $R_{\text{Exp}}(B_{\rho, \sigma}) = \bigvee_i \exists v_i.H'_i.\sigma'$. Clearly, $\eta(B_{\rho, \sigma}) \subseteq F'$ implies $\eta(R_{\text{Exp}}(\eta(B_{\rho, \sigma}))) \subseteq F'$, so we obtain $F' \models \eta(\bigvee_i \exists v_i.H'_i.\sigma') = \bigvee_i \exists v_i.H_i.\sigma$. Hence, $(\rho, \sigma, c)$ does not satisfy Definition 1 (b) for $F'$, contradicting our assumptions.

We therefore find that $(\rho, \sigma, c)$ is not blocked. Combining this with our earlier observations that $\pi^*(B_{\rho, \sigma}) = B_{\sigma} \subseteq \text{RMFA}(R)$, we can apply Definition 7 to obtain $\text{sk}(H_k)\sigma' = \pi^*(\text{sk}(H_k))\sigma \subseteq \text{RMFA}(R)$. This completes the proof.

Theorem 7. Deciding BCQ entailment for programs $(R, I)$ where $R$ is $\text{RMFA}$ is $\text{coN2ExpTIME}$-complete, even if the arity of predicates is bounded.

Proof. Membership is shown as for Theorem 3. Indeed, by Lemma 6, cyclic terms are not derived in the chase, since $\pi^*$ maps non-cyclic terms to non-cyclic terms. We can apply the same counting argument as for Theorem 3 to establish a doubly exponential upper bound for the number of facts that can be derived in any branch of the restricted chase.

Hardness follows by Theorem 3, as any RJA rule set can be shown to be RMFA. This proof is analogous to the proof for the relation of JA and MFA [Cuenca Grau et al., 2013], so we do not repeat it here.

Theorem 8. Deciding if $R$ is RMSA is $\text{ExpTIME}$-complete, and $\text{P}$-complete if the number of variables per rule is bounded.

Proof. Membership follows since the number of derivable ground facts in $\text{RMSA}(R)$ is exponential (in the arity of predicate symbols). The number of constants used in computing $\text{RMSA}(R)$ is the number of existentially quantified variables in $R$ plus one (for the constant $\ast$), and therefore linear in the size of $R$. Checking whether a rule is blocked likewise corresponds to a Datalog-like computation, leading to an overall $\text{ExpTIME}$ upper bound.

When the number of variables per rule is bounded by a constant, then the grounding of the rules (obtained by uniformly replacing each variable by some constant) yields a polynomially large propositional Horn theory, for which reasoning is in $\text{P}$ [Dantsin et al., 2001]. The same applies to the blocking check.

For hardness, we reduce the $\text{ExpTIME}$-hard fact entailment problem of Datalog [Dantsin et al., 2001] to RMSA checking. Let $R$ be a set of Datalog rules (i.e., non-generating, deterministic rules), let $I$ be an instance, and let $A(c)$ be a ground fact for which we want to check if $R \cup I \models A(c)$. Without loss of generality, we can assume that each rule body in $R$ forms a connected structure. This can be achieved by introducing a fresh binary predicate $U$, adding a fact $U(a, b)$ for all constants $a$ and $b$ in $I$, and extending each rule body with an atom $U(x, y)$ for all pairs of variables $x$ and $y$ in this body. Clearly, this is a polynomial reduction.

We define a set $R'$ of existential rules using a fresh binary predicate $R$ and a list of variables $w$ consisting of one variable $w_n$ for every constant $\ast$ in $I$. $R'$ then contains all rules of $R$ and one additional rule

$$R(x, y) \wedge A(y) \rightarrow \exists w.R(y, w_n) \land \bigwedge_{p(c_1, \ldots, c_n) \in I} p(c_1, \ldots, c_n)$$

Hardness of RMSA checking follows from the fact that $R \cup I \models A(c)$ if and only if the set $R'$ is not RMSA. Indeed, (12) applies to the critical instance once. This application creates a copy of $I$ with new constants. If $A(c)$ can be derived from $I$, then a similar derivation is possible from this copy of $I$, leading to another application of rule (12). This process can be repeated so that the chase does indeed not terminate; in particular $R'$ is not RMSA in this case. The assumption that rule bodies are connected is important to ensure that each copy of $I$ is used in isolation when applying rules.

We can make an analogous argument to show that the RMSA check is $\text{P}$-hard when the number of variables per rule is bounded. However, (12) does not have a bounded number of variables (unless we also impose a bound on the number of constants). Therefore, we now use the fact that entailment in propositional Horn logic is already $\text{P}$-hard. Let $H$ be a propositional Horn theory consisting (without loss of generality) of
facts $p$ and binary rules $p \land q \rightarrow r$. For each propositional letter $p$, we introduce a unary predicate $T_p$, and each Horn rule $p \land q \rightarrow r$ is rewritten to $T_p(x) \land T_q(x) \rightarrow T_r(x)$. Rule (12) is then replaced by a rule

$$R(x, y) \land T_q(y) \rightarrow \exists w. R(y, w) \land \bigwedge_{p \in \mathcal{H}} T_p(w)$$

Then $\mathcal{H} \models q$ if and only if the resulting set of rules is not RMSA. \hfill \square

C Proofs for Section 5

Theorem 9. If $\mathcal{R}$ is MFC, then there is an instance $I$ for which the chase of $(\langle \mathcal{R} \rangle, I)$ is infinite.

Proof. We can show the claim using the critical instance $I = I_R$. Assume that $\mathcal{R}$ is MFC and that $\rho : B \rightarrow \exists w. H$ is a rule in $\mathcal{R}$ that satisfies the conditions of Definition 9. We show by induction that the chase of $(\langle \mathcal{R} \rangle, I)$ contains $n$ distinct applications of $\rho$, for every $n$, and is therefore infinite.

The base case $n = 1$ follows since all rules are applicable at least once to $I = I_R$.

For the inductive step, assume that $\rho$ was applied $n$ times during the chase, with the $n$th application using substitution $\sigma$. Consider the set $I_\rho$ from Definition 9, and let $\pi$ be a function that maps each constant $c_\pi$ in $I_\rho$ to a variable $z$ in $\rho$ to $\pi$. We extend $\pi$ to skolem terms and to ground facts as usual. Then $\pi$ is a homomorphism from $I_\rho$ to the set $\mathcal{B}_\pi \cup \mathcal{H}(\mathcal{R})$, which in turn is a subset of the chase of $(\langle \mathcal{R} \rangle, I)$.

Since rule derivations are preserved under homomorphisms, we find that $\pi$ extends to a homomorphism from $F_\rho$ to the chase of $(\langle \mathcal{R} \rangle, I)$. Therefore, since $F_\rho$ contains a cyclic term $f_v(\ldots f_v(t) \ldots)$ for some variable $v \in \mathcal{V}$, we find that $\pi(f_v(\ldots f_v(t) \ldots)) = f_v(\ldots f_v(\pi(t)) \ldots)$ occurs in the chase. Since $I_\rho$ contains no cyclic terms, the computation of $F_\rho$ involved an application of rule $\rho$ (the only rule that may introduce an $f_v$-term) to facts derived from $I_\rho$. This additional $(n + 1)$th application of $\rho$ must also occur in the chase for $\pi(f_v(\ldots f_v(t) \ldots))$ to occur there. \hfill \square

Lemma 10. Consider a chase tree for a program $(\langle \mathcal{R}, I \rangle)$, let $F$ be the label of some node $n$ in this tree, and let $\mathcal{R}_n$ be $\mathcal{R}$ with $\forall$ replaced by $\land$. If $\rho : B \rightarrow \exists w. H$ is a deterministic rule in $\mathcal{R}$, $F \models B \sigma$, and $(\rho, \sigma)$ is unblockable for $\mathcal{R}_n$, then $(\rho, \sigma)$ is applied in a node above or below $n$.

Proof. For a contradiction, suppose that the preconditions hold but $(\rho, \sigma)$ is not applied. Since applications are fair, Definition 1 (b) is violated in all leaf nodes below $n$, which are saturated under the application of non-generating rules (condition (4) of Definition 2). For any such leaf $n'$ labelled $F'$, we map $F'$ to $U_{\rho, \sigma, \mathcal{R}}$ using a function $\pi$. For every term or subterm $t$ that occurs in $B \sigma$, we define $\pi(t) = t$; for all other terms we define $\pi(t) = \ast$. We extend $\pi$ to facts and sets of facts as usual.

We show that $\pi$ is a homomorphism from $F'$ to $U_{\rho, \sigma, \mathcal{R}}$. We proceed by induction over the chase step in which each fact was derived. The base case is the starting instance $I$, for which we find that $\pi(I) \subseteq I_\mathcal{R} \subseteq I_{\rho, \sigma, \mathcal{R}} \subseteq U_{\rho, \sigma, \mathcal{R}}$.

For the induction step, assume that $\pi$ is a homomorphism from some intermediate set of facts $F_1$ to $U_{\rho, \sigma, \mathcal{R}}$, and that the next set of facts on the chase path is $F_2$, obtained by applying the skolemised version of rule $\rho' : B' \rightarrow \exists w. H'_1$ under substitution $\sigma'$. That is $F_2 = F_1 \cup \{sk(H'_1 \sigma')\}$ for some index $k$. We consider two cases.

(1) If $sk(H'_1 \sigma')$ introduces a new (ground) skolem term $t$ that occurs as a (sub)term in $B \sigma^*$, then $F_2 \subseteq U_{\rho, \sigma, \mathcal{R}}$. We claim that $\pi(sk(H'_k \sigma')) \subseteq F_2$. Indeed, for the frontier $x'$ of $\rho'$, $t$ has the form $f(x', \sigma')$, and all (sub)terms of $x'$ therefore also occur in $B \sigma^*$. We therefore obtain $\pi(x' \sigma') = x' \sigma'$. In other words, $\pi$ is the identity on all frontier variables of $\rho'$, and therefore $\pi(sk(H_k \sigma')) = sk(H_k \sigma')$. By the definitions, we get $sk(H'_k \sigma') \subseteq F_2$, so $\pi(sk(H'_k \sigma')) \subseteq F_2 \subseteq U_{\rho, \sigma, \mathcal{R}}$ as required.

(2) If $sk(H'_k \sigma')$ contains no new skolem term that occurs in $B \sigma$, then $\pi(sk(H'_k \sigma'))$ is contained in the facts that are derived by applying the rule $\rho'' \in \mathcal{R}^*$ (obtained from $\rho$ by replacing $\forall$ with $\land$, and existential variables with $\ast$) to the facts $\pi(sk(B \sigma')$). By the induction hypothesis, the latter are contained in $U_{\rho, \sigma, \mathcal{R}}$, so that $\rho''$ is indeed applied. This shows the claim and completes the induction.

Since $(\rho, \sigma)$ is not applicable to $F'$, we find that $F' \models \exists w. H \sigma$. Together with $\pi(F') \subseteq U_{\rho, \sigma, \mathcal{R}}$ we obtain $U_{\rho, \sigma, \mathcal{R}} \models \pi(\exists w. H \sigma) = \exists w. H \sigma^*$. Therefore, $(\rho, \sigma)$ is not unblockable -- contradiction. \hfill \square

Theorem 11. If $\mathcal{R}$ is RMFC, then there is an instance $I$ for which the restricted chase of $(\langle \mathcal{R}, I \rangle)$ is infinite.

Proof. We can use the same inductive argument as in the proof of Theorem 9 to show that the restricted chase over the instance $I = I_\mathcal{R}$ as in Definition 9 contains infinitely many applications of a rule $\rho$. In fact, since RMFC only considers deterministic rules for detecting cycles, all paths of the restricted chase must be infinite in this case. The fact that $I_\mathcal{R}$ contains skolem terms does not concern us, since we could simply replace them with fresh constants to get a corresponding infinite chase from an instance that has no function symbols.

The argument for the base case is the same as for the induction step, since we start in both cases with a set of facts that corresponds to a situation where $\rho$ has just been applied. We can argue that another application of $\rho$ is possible just as we did for the proof of Theorem 9, where we additionally note that, by Lemma 10, it is certain that every unblockable (deterministic) rule will eventually be applied. This ensures that the derivation of cyclic terms as required in the induction can indeed be found in the restricted chase. \hfill \square