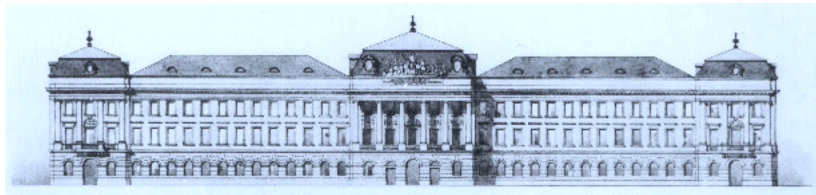
I N F S Y S
RESEARCH
R E P O R T



INSTITUT FÜR INFORMATIONSSYSTEME

ARBEITSBEREICH WISSENSBASIERTE SYSTEME

QUERY REWRITING FOR HORN-$\mathcal{SHIQ}$ PLUS RULES

Thomas Eiter   Magdalena Ortiz   Mantas Šimkus
TrungKien Tran   Guohui Xiao

INFSYS RESEARCH REPORT 1843-12-04
APRIL 2012

Institut für Informationssysteme
AB Wissensbasierte Systeme
Technische Universität Wien
Favoritenstraße 9-11
A-1040 Wien, Austria
Tel: +43-1-58801-18405
Fax: +43-1-58801-18493
sek@kr.tuwien.ac.at
www.kr.tuwien.ac.at

TU
WIEN

INFSYS RESEARCH REPORT
INFSYS RESEARCH REPORT 1843-12-04, APRIL 2012

QUERY REWRITING FOR HORN-$\mathcal{SHIQ}$ PLUS RULES

Thomas Eiter[1] Magdalena Ortiz[1] Mantas Šimkus[1]
TrungKien Tran [2] Guohui Xiao[1]

Abstract. Query answering over Description Logic (DL) ontologies has become a vibrant field of research. Efficient realizations often exploit database technology and rewrite a given query to an equivalent SQL or Datalog query over a database associated with the ontology. This approach has been intensively studied for conjunctive query answering in the $\mathcal{DL}$-*Lite* and $\mathcal{EL}$ families, but is much less explored for more expressive DLs and queries. We present a rewriting-based algorithm for conjunctive query answering over Horn-$\mathcal{SHIQ}$ ontologies, possibly extended with recursive rules under limited recursion as in $\mathcal{DL}$+*log*. This setting not only subsumes both $\mathcal{DL}$-*Lite* and $\mathcal{EL}$, but also yields an algorithm for answering (limited) recursive queries over Horn-$\mathcal{SHIQ}$ ontologies (an undecidable problem for full recursive queries). A prototype implementation shows its potential for applications, as experiments exhibit efficient query answering over full Horn-$\mathcal{SHIQ}$ ontologies and benign downscaling to $\mathcal{DL}$-*Lite*, where it is competitive with comparable state of the art systems.

[1]Institut für Informationssysteme, Technische Universität Wien, 1040 Wien, Austria. Email: (eiter,ortiz,xiao)@kr.tuwien.ac.at simkus@dbai.tuwien.ac.at
[2]STARLab, Vrije Universiteit Brussel. Email: truntran@vub.ac.be

Acknowledgements: This work was supported by the Austrian Science Fund (FWF) grants P20840 and T515.

$\text{village}(\textit{Carichi})$	$\text{hasHDI}(\textit{Carichi}, \textit{low})$
$\text{state}(\textit{Chihuahua})$	$\text{hasHDI}(\textit{Mexico}, \textit{high})$
$\text{country}(\textit{Mexico})$	$\text{hasHDI}(\textit{Islamabad}, \textit{high})$
$\text{capital}(\textit{Islamabad})$	$\text{hasHDI}(\textit{Brasilia}, \textit{high})$
$\text{country}(\textit{Pakistan})$	$\text{isLocatedIn}(\textit{Carichi}, \textit{Chihuahua})$
$\text{capital}(\textit{Brasilia})$	$\text{isLocatedIn}(\textit{Chihuahua}, \textit{Mexico})$
$\text{country}(\textit{Brazil})$	$\text{isLocatedIn}(\textit{Islamabad}, \textit{Pakistan})$
	$\text{isLocatedIn}(\textit{Brasilia}, \textit{Brazil})$

(a) $\textit{trans}(\text{isLocatedIn})$ (c) $\text{Country} \sqsubseteq \exists \text{hasCapital}.\text{capital}$

(b) $\text{hasCapital} \sqsubseteq \text{isLocatedIn}^-$ (d) $\text{Country} \sqsubseteq \forall \text{hasCapital}.\text{city}$

(e) $\text{Country} \sqsubseteq \leqslant 1\ \text{isLocatedIn}^-.\text{capital}$

$$(q_1) \quad \text{disadvantagedTerritory}(x, y) \leftarrow \text{hasHDI}(x, \textit{low}), \text{isLocatedIn}(x, y), \text{country}(y), \text{hasHDI}(y, \textit{high})$$

$$(q_2) \quad \text{hasDevelopedCapital}(x) \leftarrow \text{country}(x), \text{hasCapital}(x, y), \text{city}(y), \text{hasHDI}(y, \textit{high})$$

Table 1: An example ontology and queries

1 Introduction

Description Logics (DLs) are the primary tool for representing and reasoning about knowledge given by an *ontology*. They are mostly fragments of first-order logic with a clear-cut semantics, convenient syntax and decidable reasoning, performed by quite efficient algorithms. This has led to important applications of DLs in areas like Ontology Based Data Access (OBDA), Data Integration and the Semantic Web, where the OWL standard is heavily based on DLs.

An important reasoning task in DLs is query answering similar as in databases, where a database-style query is evaluated over an ontology, viewing it as an enriched database.

Example 1. Consider the following sociopolitical ontology. The Human Development Index (HDI) of certain territories T, whose value V may be low, medium or high (as in the UN Development Programme) is given by facts $\text{hasHDI}(T, V)$. Further facts classify territories as cities, countries, etc. and relate their locations. The facts are shown in the two left columns of Table 1. The axioms (a)–(e) on the right hand side provide a terminology (in DL syntax) stating that: (a) the isLocatedIn relation is transitive; (b) the capital of a territory is located in that territory; (c) every country has a capital; (d) only cities can be capitals; and (e) only one capital can be located in each country. The query q_1 can be used to retrieve disadvantaged territories that lie in countries with high HDI but have a low HDI themselves. Observe that if we evaluate q_1 over the database (i.e., the facts), it returns no answer: indeed, Mexico is the only country with high HDI, and there is no fact $\text{islocatedIn}(X, \textit{Mexico})$ such that territory X has low HDI. However, if we evaluate q_1 over the full ontology, we can infer from axiom (a) that Carichi is located in Mexico, and return $(\textit{Carichi}, \textit{Mexico})$ as an answer. The query q_2, which retrieves countries whose capital city has a high HDI, would also have an empty answer over the database, but from the axioms (b)–(e) we can infer that Brasilia is the capital of Brazil and Islamabad the capital of Pakistan, and return both countries as an answer to the query.

To supply this reasoning service, a number of challenges must be faced. *Conjunctive queries* (CQs) have typically much higher complexity than standard reasoning in a DL, and recursive DATALOG queries are undecidable even in very weak DLs, including the ones considered here [Levy and Rousset, 1998]. For reasoning with large instance data, *translating* queries into database query languages has proved to be efficient. Calvanese et al. (2007b) introduced a *query rewriting* technique for the $\mathcal{DL}$-*Lite* family of DLs, where the terminological information is incorporated into the query in such a way that it can be straight evaluated over the database facts. For example, a rewriting of query q_1 in Table 1 should include, among other queries,

$$\mathrm{disadvantagedTerritory}(x, y) \leftarrow \mathrm{hasHDI}(x, \mathit{low}), \mathrm{country}(y), \\ \mathrm{hasCapital}(y, x), \mathrm{hasHDI}(y, \mathit{high}),$$

which adds all tuples (x, y) to the query answer that can be inferred using axiom (b). Such rewriting approaches have been developed for answering CQs in DLs of the $\mathcal{DL}$-*Lite* family, and to a lesser extent for $\mathcal{EL}$, but they are practically unexplored for more expressive DLs and queries (see Related Work for details).

In this paper we present a rewriting-based method for query answering over ontologies in Horn-$\mathcal{SHIQ}$ (the disjunction-free fragment of $\mathcal{SHIQ}$), which extends the members of the $\mathcal{DL}$-*Lite* and the $\mathcal{EL}$ families. $\mathcal{DL}$-*Lite* and $\mathcal{EL}$ are prominent DLs which underlie the OWL 2 QL and the OWL 2 EL profiles, respectively. They offer different expressiveness while allowing for tractable reasoning. For example, axiom (b) is allowed in most DLs of the $\mathcal{DL}$-*Lite* family but not in $\mathcal{EL}$, while (c) is allowed in $\mathcal{EL}$ but not in $\mathcal{DL}$-*Lite*. Axioms (a), (d) and (e) are not expressible in either of them, but they are expressible in Horn-$\mathcal{SHIQ}$. Despite the increase in expressivity, reasoning in Horn-$\mathcal{SHIQ}$ is still tractable in data complexity.

In this paper, we make the following contributions:

- We provide a practical algorithm for rewriting queries over Horn-$\mathcal{SHIQ}$ ontologies. It first applies a special resolution calculus, and then rewrites the query w.r.t. the saturated TBox into a DATALOG program ready for evaluation over any ABox. It runs in polynomial time in data complexity, and is worst-case optimal.

- It can handle CQs and the more general *weakly DL-safe* DATALOG queries in the style of $\mathcal{DL}$+*log* [Rosati, 2006], where only existentially quantified variables may be bound to 'anonymous' domain elements implied by axioms.

- The algorithm supports transitive roles, which are considered relevant in practice [Sattler, 2000], although challenging for query answering (Glimm et al. 2006, Eiter et al. 2009). It simultaneously allows for full existential quantification, inverse roles, and number restrictions, covering and extending the OWL2 profiles QL, EL and RL.

- A prototype implementation for CQ answering (without transitive roles) shows that our approach behaves well in practice. In experiments it worked efficiently and it scaled down nicely to $\mathcal{DL}$-*Lite*, where it is competitive with state of the art query rewriting systems.

2 Description Logic Horn-$\mathcal{SHIQ}$

As usual, we assume countably infinite sets $\mathsf{N_C} \supset \{\top, \bot\}$ and $\mathsf{N_R}$ of *atomic concepts* and *role names*, respectively. $\mathsf{N_R} \cup \{r^- \mid r \in \mathsf{N_R}\}$ is the set of *roles*. If $r \in \mathsf{N_R}$, then $\mathsf{inv}(r) = r^-$ and $\mathsf{inv}(r^-) = r$. *Concepts* are inductively defined as follows: (a) each $A \in \mathsf{N_C}$ is a concept, and (b) if C, D are concepts and r is a role, then $C \sqcap D$, $C \sqcup D$, $\neg C$, $\forall r.C$, $\exists r.C$, $\geqslant n\, r.C$ and $\leqslant n\, r.C$, for $n \geq 1$, are concepts. An expression $C \sqsubseteq D$, where C, D are concepts, is a *general concept inclusion axiom (GCI)*. An expression $r \sqsubseteq s$, where r, s are roles, is a *role inclusion (RI)*. A *transitivity axiom* is an expression $trans(r)$, where r is a role. A TBox $\mathcal{T}$ is a finite set of GCIs, RIs and transitivity axioms. We let $\sqsubseteq^*_\mathcal{T}$ denote the reflexive transitive closure of $\{(r, s) \mid r \sqsubseteq s \in \mathcal{T}$ or $\mathsf{inv}(r) \sqsubseteq \mathsf{inv}(s) \in \mathcal{T}\}$. A role s is *transitive* in $\mathcal{T}$ if $trans(s) \in \mathcal{T}$ or $trans(s^-) \in \mathcal{T}$. A role s is *simple* in $\mathcal{T}$ if there is no transitive r in $\mathcal{T}$ s.t. $r \sqsubseteq^*_\mathcal{T} s$. $\mathcal{T}$ is a $\mathcal{SHIQ}$ terminology if roles in concepts of the form $\geqslant n\, r.C$ and $\leqslant n\, r.C$ are simple. The semantics for TBoxes is given by *interpretations* $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$. We write $\mathcal{I} \models \mathcal{T}$ if $\mathcal{I}$ is a *model* of $\mathcal{T}$. See [Baader *et al.*, 2007] for more details.

A TBox $\mathcal{T}$ is a Horn-$\mathcal{SHIQ}$ TBox (in normalized form), if each GCI in $\mathcal{T}$ takes one the following forms:

(F1) $A_1 \sqcap \ldots \sqcap A_n \sqsubseteq B$, (F3) $A_1 \sqsubseteq \forall r.B$,

(F2) $A_1 \sqsubseteq \exists r.B$, (F4) $A_1 \sqsubseteq \leqslant 1\, r.B$,

where $A_1, \ldots, A_n, B$ are concept names and r is a role. Axioms (F1) are called *existential*. W.l.o.g. we treat here only Horn-$\mathcal{SHIQ}$ TBoxes in normalized form; our results generalize to full Horn-$\mathcal{SHIQ}$ by means of TBox *normalization*; see e.g. [Kazakov, 2009; Krötzsch *et al.*, 2007] for a definition and normalization procedures.

An Horn-$\mathcal{ALCHIQ}$ TBox is a Horn-$\mathcal{SHIQ}$ TBox with no transitivity axioms. Horn-$\mathcal{ALCHIQ}^\sqcap$ TBoxes are obtained by allowing *role conjunction* $r_1 \sqcap r_2$, where r_1, r_2 are roles and in any interpretation $\mathcal{I}$, $(r_1 \sqcap r_2)^\mathcal{I} = r_1^\mathcal{I} \cap r_2^\mathcal{I}$ (we use it for a similar purpose as Glimm et al. (2008)). We let $\mathsf{inv}(r_1 \sqcap r_2) = \mathsf{inv}(r_1) \sqcap \mathsf{inv}(r_2)$ and assume w.l.o.g. that for each role inclusions $r \sqsubseteq s$ of an Horn-$\mathcal{ALCHIQ}^\sqcap$ TBox $\mathcal{T}$, (i) $\mathsf{inv}(r) \sqsubseteq \mathsf{inv}(s) \in \mathcal{T}$, and (ii) $s \in \{p, p^-\}$ for a role name p. For a set W and a concept or role conjunction $\Gamma = \gamma_1 \sqcap \ldots \sqcap \gamma_m$, we write $\Gamma \subseteq W$ for $\{\gamma_1, \ldots, \gamma_m\} \subseteq W$.

3 Ontologies and Knowledge Bases

Following [Levy and Rousset, 1998] we now define *knowledge bases (KBs)*. Let $\mathsf{N_I}$, $\mathsf{N_V}$ and $\mathsf{N_D}$ be countable infinite sets of *constants* (or, *individuals*), *variables* and DATALOG *relations*, respectively; we assume these sets as well as $\mathsf{N_C}$ and $\mathsf{N_R}$ are all mutually disjoint. Each $\sigma \in \mathsf{N_D}$ has an associated non-negative integer *arity*. An *atom* is an expression $p(\vec{t})$, where $\vec{t} \in (\mathsf{N_I})^n \cup (\mathsf{N_V})^n$, and (i) $p \in \mathsf{N_C}$ and $n = 1$, (ii) $p \in \mathsf{N_R}$ and $n = 2$, or (iii) $p \in \mathsf{N_D}$ and n is the arity of p. If $\vec{t} \in (\mathsf{N_I})^n$, then $p(\vec{t})$ is *ground*. Ground atoms $A(a)$ and $r(a, b)$, where $A \in \mathsf{N_C}$ and r is a role, are *concept* and *role assertions*, respectively. An ABox $\mathcal{A}$ is a finite set of ground atoms. A rule ρ is an expression of the form

$$h(\vec{u}) \leftarrow p_1(\vec{v_1}), \ldots, p_m(\vec{v_m}), \quad (1)$$

where $h(\vec{u})$ is an atom (the *head*), $\{p_1(\vec{v_1}), \ldots, p_m(\vec{v_m})\}$ are also atoms (the *body* atoms, denoted $body(\rho)$), and $\vec{u}, \vec{v_1}, \ldots, \vec{v_m}$ are tuples of variables. The variables in $\vec{u}$ are *distinguished*. A *KB* is a tuple $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$, where $\mathcal{T}$ is a TBox, $\mathcal{A}$ is an ABox, and $\mathcal{P}$ is a set of rules (a *program*).

The semantics for a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$ is given by extending an interpretation $\mathcal{I}$ to symbols in $\mathsf{N_I} \cup \mathsf{N_D}$. For any $c \in \mathsf{N_I}$ and $p \in \mathsf{N_D}$ of arity n, we have $c^\mathcal{I} \in \Delta^\mathcal{I}$ and $p^\mathcal{I} \subseteq (\Delta^\mathcal{I})^n$. A *match* for a rule ρ of the form (1) in $\mathcal{I}$ is a mapping from variables in ρ to elements in $\Delta^\mathcal{I}$ such that $\pi(\vec{t}) \in p^\mathcal{I}$ for each body atom $p(\vec{t})$ of ρ. We define:

(a) $\mathcal{I} \models \rho$ if $\pi(\vec{u}) \in h^\mathcal{I}$ for every match π for ρ in $\mathcal{I}$,

(b) $\mathcal{I} \models \mathcal{P}$ if $\mathcal{I} \models \rho$ for each $\rho \in \mathcal{P}$,

(c) $\mathcal{I} \models \mathcal{A}$ if $(\vec{c})^\mathcal{I} \in p^\mathcal{I}$ for all $p(\vec{c}) \in \mathcal{A}$,

(d) $\mathcal{I} \models \mathcal{K}$ if $\mathcal{I} \models \mathcal{T}$, $\mathcal{I} \models \mathcal{A}$ and $\mathcal{I} \models \mathcal{P}$.

Finally, given a ground atom $p(\vec{c})$, $\mathcal{K} \models p(\vec{c})$ if $(\vec{c})^\mathcal{I} \in p^\mathcal{I}$ for all models $\mathcal{I}$ of $\mathcal{K}$. We recall *weak DL-safety* [Rosati, 2006]. A KB $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$ is weakly DL-safe if each rule $\rho \in \mathcal{P}$ satisfies the next condition: every distinguished variable x of ρ occurs in some body atom $p(\vec{t})$ of ρ such that $p \in \mathsf{N_D}$. We make the *Unique Name Assumption (UNA)*.

A KB $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \emptyset)$ is an *ontology* (we will use $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ for brevity). A *conjunctive query (CQ)* q over $\mathcal{O}$ is a rule of the form (1) such that h does not occur in $\mathcal{O}$. The *answer* to q over $\mathcal{O}$ is $ans(\mathcal{O}, q) = \{\vec{c} \in \mathsf{N_I}^{|\vec{u}|} \mid (\mathcal{T}, \mathcal{A}, \{q\}) \models h(\vec{c})\}$. Note that $\vec{c} \in ans(\mathcal{O}, q)$ iff in any model $\mathcal{I}$ of $\mathcal{O}$ there exists a match π for q such that $\pi(\vec{u}) = (\vec{c})^\mathcal{I}$.

Note that, for a KB $\mathcal{K} = (\emptyset, \mathcal{A}, \mathcal{P})$, $\mathcal{A} \cup \mathcal{P}$ is an ordinary DATALOG program with constraints (cf. [Dantsin *et al.*, 2001]). By *models* of DATALOG programs, we mean Herbrand models, and we recall that a consistent $\mathcal{A} \cup \mathcal{P}$ has a unique least (Herbrand) model $MM(\mathcal{A} \cup \mathcal{P})$.

We will also consider programs $\mathcal{P}$ containing atoms $r^-(x, y)$, $r \in \mathsf{N_R}$, with the semantics given by the semantics of $\mathcal{P}'$ obtained by replacing in $\mathcal{P}$ each $r^-(x, y)$ by $r(y, x)$.

3.1 Elimination of Transitivity

It is handy to eliminate transitivity axioms from $\mathcal{SHIQ}$ TBoxes (see, e.g., Hustadt et al. (2007)). We use the transformation from [Kazakov, 2009], which ensures the resulting TBox is in normal form.

Definition 1. Let $\mathcal{T}^*$ be the Horn-$\mathcal{ALCHIQ}$ TBox obtained from a Horn-$\mathcal{SHIQ}$ TBox $\mathcal{T}$ by (i) adding for every $A \sqsubseteq \forall s.B \in \mathcal{T}$ and every transitive role r with $r \sqsubseteq^*_\mathcal{T} s$, the axioms $A \sqsubseteq \forall r.B^r$, $B^r \sqsubseteq \forall r.B^r$ and $B^r \sqsubseteq B$, where B^r is a fresh concept name; and (ii) removing all transitivity axioms.

The transformation does not preserve answers to CQs where non-simple roles occur. However, we can relax the notion of match and then use the translated TBox for answering arbitrary CQs.

Definition 2. Let $\mathcal{T}$ be a Horn-$\mathcal{SHIQ}$ TBox. A $\mathcal{T}$*-match* for a query q in an interpretation $\mathcal{I}$ is a mapping π from variables of q to elements in $\Delta^\mathcal{I}$ that satisfies the following:

(a) If $\alpha = p(\vec{t})$ is a body atom in q, where $p \in \mathsf{N_C} \cup \mathsf{N_D}$ or p is a simple role in $\mathcal{T}$, then $\pi(\vec{t}) \in p^\mathcal{I}$.

$$\frac{M \sqsubseteq \exists S.N \sqcap N' \quad N \sqsubseteq A}{M \sqsubseteq \exists S.N \sqcap N' \sqcap A}\ \mathbf{R}^c_\sqsubseteq \qquad \frac{M \sqsubseteq \exists S \sqcap S'.N \quad S \sqsubseteq r}{M \sqsubseteq \exists S \sqcap S' \sqcap r.N}\ \mathbf{R}^r_\sqsubseteq \qquad \frac{M \sqsubseteq \exists S.N \sqcap \bot}{M \sqsubseteq \bot}\ \mathbf{R}_\bot$$

$$\frac{M \sqsubseteq \exists S \sqcap r.N \quad A \sqsubseteq \forall r.B}{M \sqcap A \sqsubseteq \exists S \sqcap r.N \sqcap B}\ \mathbf{R}_\forall \qquad \frac{M \sqsubseteq \exists S \sqcap \mathsf{inv}(r).N \sqcap A \quad A \sqsubseteq \forall r.B}{M \sqsubseteq B}\ \mathbf{R}^-_\forall$$

$$\frac{M \sqsubseteq \exists S \sqcap r.N \sqcap B \quad A \sqsubseteq\ \leqslant 1\, r.B \quad M' \sqsubseteq \exists S' \sqcap r.N' \sqcap B}{M \sqcap M' \sqcap A \sqsubseteq \exists S \sqcap S' \sqcap r.N \sqcap N' \sqcap B}\ \mathbf{R}_\leq$$

$$\frac{M \sqsubseteq \exists S \sqcap \mathsf{inv}(r).N_1 \sqcap N_2 \sqcap A \quad A \sqsubseteq\ \leqslant 1\, r.B \quad N_1 \sqcap A \sqsubseteq \exists S' \sqcap r.N' \sqcap B \sqcap C}{M \sqcap B \sqsubseteq C \quad M \sqcap B \sqsubseteq \exists S \sqcap \mathsf{inv}(S' \sqcap r).N_1 \sqcap N_2 \sqcap A}\ \mathbf{R}^-_\leq$$

Table 2: Inference rules. $M^{(\prime)}$, $N^{(\prime)}$, (resp., $S^{(\prime)}$) are conjunctions of atomic concepts (roles); A, B are atomic concepts

(b) If $\alpha = s(x, y)$ with s non-simple, then there exist a transitive $r \sqsubseteq^*_\mathcal{T} s$ and $d_1 \in \Delta^\mathcal{I}, \ldots, d_k \in \Delta^\mathcal{I}$ such that $d_1 = \pi(x)$, $d_k = \pi(y)$, and $(d_i, d_{i+1}) \in r^\mathcal{T}$ for all each $1 \leq i < k$; we call this sequence $d_1 \in \Delta^\mathcal{I}, \ldots, d_k \in \Delta^\mathcal{I}$ an *r-path from $\pi(x)$ to $\pi(y)$*.

The set $ans^\mathcal{T}(\mathcal{O}, q)$ is defined as $ans(\mathcal{O}, q)$ but using $\mathcal{T}$-matches instead of matches. The next characaterization follows from known techniques, see e.g. [Eiter *et al.*, 2012b] for a similar result.

Proposition 1. *For any Horn-$\mathcal{SHIQ}$ ontology $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ and CQ q, we have $ans(\mathcal{O}, q) = ans^\mathcal{T}((\mathcal{T}^*, \mathcal{A}), q)$.*

4 Canonical Models

A stepping stone for the tailored query answering methods for Horn DLs and languages like DATALOG$^\pm$ is the *canonical model property* [Eiter *et al.*, 2008b; Ortiz *et al.*, 2011; Calì *et al.*, 2009]. In particular, for a consistent Horn-$\mathcal{ALCHIQ}^\sqcap$ ontology $\mathcal{O} = (\mathcal{T}, \mathcal{A})$, there exists a model $\mathcal{I}$ of $\mathcal{O}$ that can be homomorphically embedded into any other model $\mathcal{I}'$ of $\mathcal{O}$. We show that such an $\mathcal{I}$ can be built in three steps:

(1) Close $\mathcal{T}$ under specially tailored inferences rules.
(2) Close $\mathcal{A}$ under all but existential axioms of $\mathcal{T}$.
(3) Extend $\mathcal{A}$ by "applying" the existential axioms of $\mathcal{T}$.

For Step (1), we tailor from the inference rules in [Kazakov, 2009; Ortiz *et al.*, 2010] a calculus to support model building for Horn-$\mathcal{ALCHIQ}^\sqcap$ ontologies.

Definition 3. Given a Horn-$\mathcal{ALCHIQ}^\sqcap$ TBox $\mathcal{T}$, $\Xi(\mathcal{T})$ is the TBox obtained from $\mathcal{T}$ by exhaustively applying the inference rules in Table 2.

For Step (2), we use DATALOG rules that express the semantics of GCIs, ignoring existential axioms.

$B(y) \leftarrow A(x), r(x,y)$ for each $A \sqsubseteq \forall r.B \in \mathcal{T}$
$B(x) \leftarrow A_1(x), \ldots, A_n(x)$ for all $A_1 \sqcap \ldots \sqcap A_n \sqsubseteq B \in \Xi(\mathcal{T})$
$r(x,z) \leftarrow r(x,y), r(y,z)$ for all transitive roles r in $\mathcal{T}$
$r(x,y) \leftarrow r_1(x,y), \ldots, r_n(x,y)$ for all $r_1 \sqcap \ldots \sqcap r_n \sqsubseteq r \in \mathcal{T}$
$\bot(x) \leftarrow A(x), r(x,y_1), r(x,y_2), B(y_1), B(y_2), y_1 \neq y_2$ for each $A \sqsubseteq\ \leqslant 1\, r.B \in \mathcal{T}$
$\Gamma \leftarrow A(x), A_1(x), \ldots, A_n(x), r(x,y), B(y)$ for all $A_1 \sqcap \ldots \sqcap A_n \sqsubseteq \exists r_1 \sqcap \ldots \sqcap r_m.B_1 \sqcap \ldots \sqcap B_k$ and $A \sqsubseteq\ \leqslant 1\, r.B$ of $\Xi(\mathcal{T})$ such that $r{=}r_i$ and $B{=}B_j$ for some i, j with $\Gamma \in \{B_1(y), \ldots, B_k(y), r_1(x,y), \ldots, r_k(x,y)\}$

Table 3: (Completion rules) DATALOG program $\mathsf{cr}(\mathcal{T})$.

Definition 4. Given a Horn-$\mathcal{ALCHIQ}^{\sqcap}$ TBox $\mathcal{T}$, $\mathsf{cr}(\mathcal{T})$ is the DATALOG program described in Table 3.

Given a consistent Horn-$\mathcal{ALCHIQ}^{\sqcap}$ ontology $\mathcal{O} = (\mathcal{T}, \mathcal{A})$, the least model $\mathcal{J}$ of the DATALOG program $\mathsf{cr}(\mathcal{T}) \cup \mathcal{A}$ is almost a canonical model of $\mathcal{O}$; however, existential axioms may be violated. We deal with this in Step (3), by extending $\mathcal{J}$ with new domain elements as required by axioms $A \sqsubseteq \exists r.N$ in $\Xi(\mathcal{T})$, akin to database *chase* [Maier and Mendelzon, 1979] where fresh values and tuples are introduced to satisfy the given dependencies.

Definition 5. Let $\mathcal{T}$ be a Horn-$\mathcal{ALCHIQ}^{\sqcap}$ TBox and $\mathcal{I}$ an interpretation. A GCI $M \sqsubseteq \exists S.N$ is *applicable at* $e \in \Delta^{\mathcal{I}}$ if

(a) $e \in M^{\mathcal{I}}$,

(b) there is no $e' \in \Delta^{\mathcal{I}}$ with $(e, e') \in S^{\mathcal{I}}$ and $e' \in N^{\mathcal{I}}$,

(c) there is no axiom $M' \sqsubseteq \exists S'.N' \in \mathcal{T}$ such that $e \in (M')^{\mathcal{I}}$, $S \subseteq S'$, $N \subseteq N'$, and $S \subset S'$ or $N \subset N'$.

An interpretation $\mathcal{J}$ obtained from $\mathcal{I}$ by an *application* of an applicable axiom $M \sqsubseteq \exists S.N$ at $e \in \Delta^{\mathcal{I}}$ is defined as:

- $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}} \cup \{d\}$ with d a new element not present in $\Delta^{\mathcal{I}}$ (we call d a *successor* of e),
- For each atomic $A \in \mathsf{N_C}$ and $o \in \Delta^{\mathcal{J}}$, we have $o \in A^{\mathcal{J}}$ if (a) $o \in \Delta^{\mathcal{I}}$ and $o \in A^{\mathcal{I}}$; or (b) $o = d$ and $A \in N$.
- For each role name r and $o, o' \in \Delta^{\mathcal{J}}$, we have $(o, o') \in r^{\mathcal{J}}$ if (a) $o, o' \in \Delta^{\mathcal{I}}$ and $(o, o') \in r^{\mathcal{I}}$; or (b) $(o, o') = (e, d)$ and $r \in S$; or (c) $(o, o') = (d, e)$ and $\mathsf{inv}(r) \in S$.

We denote by *chase*$(\mathcal{I}, \mathcal{T})$ a possibly infinite interpretation obtained from $\mathcal{I}$ by applying the existential axioms in $\mathcal{T}$. We require the application to *fair*: the application of an applicable axiom can not be infinitely postponed.

We note that *chase*$(\mathcal{I},\mathcal{T})$ is unique up to renaming of domain elements. As usual in DLs, it can be seen as a 'forest': application of existential axioms simply attaches 'trees' to a possibly arbitrarily shaped $\mathcal{I}$. The following statement can be shown similarly as in [Ortiz *et al.*, 2011].

Proposition 2. *Let $\mathcal{O}=(\mathcal{T},\mathcal{A})$ be a Horn-$\mathcal{ALCHIQ}^{\sqcap}$ ontology. Then $\mathcal{O}$ is consistent iff $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$ consistent. Moreover, if $\mathcal{O}$ is consistent, then*

(a) chase$(MM(\mathcal{A}\cup\mathsf{cr}(\mathcal{T})),\Xi(\mathcal{T}))$ is a model of $\mathcal{O}$, and

(b) chase$(MM(\mathcal{A}\cup\mathsf{cr}(\mathcal{T})),\Xi(\mathcal{T}))$ can be homomorphically embedded into any model of $\mathcal{O}$.

Proof. Let $\mathcal{O}=(\mathcal{T},\mathcal{A})$ be a Horn-$\mathcal{ALCHIQ}^{\sqcap}$ ontology.

Suppose $\mathcal{O}$ is consistent and $\mathcal{J}$ is a model of $\mathcal{O}$. We first show that $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$ is consistent, and afterwards show (a) and (b). Due to the UNA, we can w.l.o.g. assume that $a^{\mathcal{J}}=a$ for each constant $a\in\mathsf{N_I}$. A model of $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$ can built by simply restricting the domain of $\mathcal{J}$ to constants. Let $\mathcal{J}'$ be the interpretation such that

- $\Delta^{\mathcal{J}'}=\mathsf{N_I}$;

- $A^{\mathcal{J}'}=A^{\mathcal{J}}\cap\Delta^{\mathcal{J}'}$ and $r^{\mathcal{J}'}=r^{\mathcal{J}}\cap\Delta^{\mathcal{J}'}\times\Delta^{\mathcal{J}'}$, for all concepts names A and role names r.

$\mathcal{J}'$ is a model of $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$ because $\mathcal{J}$ is a model of $\mathcal{T}$ and since all axioms in $\Xi(\mathcal{T})$ are logical consequences of $\mathcal{T}$.

Assume the least model $\mathcal{I}_{\mathcal{A}}$ of $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$, which exists due to consistency $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$. Let $\mathcal{I}_{\mathcal{O}}=$ *chase*$(\mathcal{I}_{\mathcal{A}},\Xi(\mathcal{T}))$. We show next that $\mathcal{I}_{\mathcal{O}}$ is a model of $\mathcal{O}$, i.e. show (a). To show the statement we need some book-keeping when building $\mathcal{I}_{\mathcal{O}}$. We assume $\Delta^{\mathcal{I}_{\mathcal{A}}}=\mathsf{N_I}$ and prescribe the naming of fresh domain elements introduced during the chase procedure. In particular, if d is a successor of e according to Definition 5, then d is an expression of the form $e\cdot n$, where n is a integer. We show that $\mathcal{I}_{\mathcal{O}}$ satisfies each statement in $\mathcal{O}$:

(1) For assertions $A(a)\in\mathcal{A}$ and $r(a,b)\in\mathcal{A}$, we have $a^{\mathcal{I}_{\mathcal{A}}}\in A^{\mathcal{I}_{\mathcal{A}}}$ and $(a^{\mathcal{I}_{\mathcal{A}}},b^{\mathcal{I}_{\mathcal{A}}})\in r^{\mathcal{I}_{\mathcal{A}}}$ because $\mathcal{I}_{\mathcal{A}}$ is a model of $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$. We have $a^{\mathcal{I}_{\mathcal{O}}}\in A^{\mathcal{I}_{\mathcal{O}}}$ and $(a^{\mathcal{I}_{\mathcal{O}}},b^{\mathcal{I}_{\mathcal{O}}})\in r^{\mathcal{I}_{\mathcal{O}}}$ because $\mathcal{I}_{\mathcal{O}}$ is an extension $\mathcal{I}_{\mathcal{A}}$ by construction.

(2) Assume an axiom $M\sqsubseteq B\in\Xi(\mathcal{T})$, where M is a conjunction of atomic concepts, and also assume a domain element $e\in M^{\mathcal{I}_{\mathcal{O}}}$. Note that $\mathcal{T}\subseteq\Xi(\mathcal{T})$. If $e\in\mathsf{N_I}$, then $e\in B^{\mathcal{I}_{\mathcal{O}}}$ since $\mathcal{I}_{\mathcal{A}}$ is a model of $\mathcal{A}\cup\mathsf{cr}(\mathcal{T})$. Assume $e=w\cdot n$. We know e is a successor of w introduced in $\mathcal{I}_{\mathcal{O}}$ by an application of some $M'\sqsubseteq\exists S.N\in\Xi(\mathcal{T})$. By the construction of $\mathcal{I}_{\mathcal{O}}$, e satisfies exactly the atomic concepts in N. It remains to see that $B\in N$. This follows from the inference rule ($\mathbf{R}^{c}_{\sqsubseteq}$). Indeed, if $B\notin N$, then we can apply ($\mathbf{R}^{c}_{\sqsubseteq}$) to obtain the axiom $M'\sqsubseteq\exists S.N\sqcap B\in\Xi(\mathcal{T})$. This makes $M'\sqsubseteq\exists S.N\in\Xi(\mathcal{T})$ inapplicable at e due a violation of (c) in Definition 5.

(3) To show that existential axioms are satisfied, first take an arbitrary domain element $e\in A^{\mathcal{I}_{\mathcal{O}}}$. We say $M\sqsubseteq\exists S.N\in\Xi(\mathcal{T})$ is *relevant for* e if there is no axiom $M'\sqsubseteq\exists S'.N'\in\mathcal{T}$ such that $e\in(M')^{\mathcal{I}}$, $S\subseteq S'$, $N\subseteq N'$, and $S\subset S'$ or $N\subset N'$. To prove that $\mathcal{I}_{\mathcal{O}}$ satisfies each

existential axiom of $\Xi(\mathcal{T})$, it suffices to show that $\mathcal{I}_\mathcal{O}$ satisfies each existential axiom that is relevant for e. To this end, assume $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ relevant for e. Suppose $e \in M^{\mathcal{I}_\mathcal{O}}$ and $e \notin (\exists S.N)^{\mathcal{I}_\mathcal{O}}$. Then $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ is applicable in $\mathcal{I}_\mathcal{O}$ at e according to Definition 5. This leads to a contradiction: by the fairness of chase, the axiom $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ must be applied and thus $e \in (\exists S.N)^{\mathcal{I}_\mathcal{O}}$.

(4) Assume an axiom $A \sqsubseteq \forall r.B \in \mathcal{T}$ and a domain element $e \in A^{\mathcal{I}_\mathcal{O}}$. Suppose there is $e' \in \Delta^{\mathcal{I}_\mathcal{O}}$ such that $(e, e') \in r^{\mathcal{I}_\mathcal{O}}$ and $e' \notin B^{\mathcal{I}_\mathcal{O}}$. Due to the definition of $\mathcal{I}_\mathcal{O}$, we have 3 possible cases:

(i) $e, e' \in \mathsf{N_I}$ and $(e, e') \in r^{\mathcal{I}_\mathcal{A}}$. We have that $e' \in B^{\mathcal{I}_\mathcal{O}}$ because $\mathcal{I}_\mathcal{A}$ is a model of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$ by assumption.

(ii) $e' = e \cdot n$ for some integer n, where e' was introduced by applying some axiom $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$. Note that, by the construction of $\mathcal{I}_\mathcal{O}$, we must have $e \in M^{\mathcal{I}_\mathcal{O}}$ and $r \in S$. From the inference rule ($\mathbf{R}_\forall$) we know that $M \sqcap A \sqsubseteq \exists S.N \sqcap B \in \Xi(\mathcal{T})$. We know that $e \in (M \sqcap A)^{\mathcal{I}_\mathcal{O}}$. Then due to maximality of $M \sqsubseteq \exists S.N$ at e, we have $N \sqcap B = N$, i.e. $B \in N$. By the construction of $\mathcal{I}_\mathcal{O}$, $e' \in B^{\mathcal{I}_\mathcal{O}}$.

(iii) $e = e' \cdot n$ for some integer n, where e was introduced by applying some axiom $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$. By the construction of $\mathcal{I}_\mathcal{O}$, we have $r^- \in S$ and $A \in N$. Then by the inference rule ($\mathbf{R}_\forall^-$), we have $M \sqsubseteq B \in \Xi(\mathcal{T})$. We have already shown above that $\mathcal{I}_\mathcal{O}$ satisfies $M \sqsubseteq B$. Since $e' \in M^{\mathcal{I}_\mathcal{O}}$ by the construction of $\mathcal{I}_\mathcal{O}$, we have $e' \in A^{\mathcal{I}_\mathcal{O}}$.

(5) Assume a role inclusion $S \sqsubseteq r \in \Xi(\mathcal{T})$ and a pair $(e, e') \in S^{\mathcal{I}_\mathcal{O}}$. Due to the definition of $\mathcal{I}_\mathcal{O}$, we have 2 possible cases:

(i) $e, e' \in \mathsf{N_I}$. Then $(e, e') \in r^{\mathcal{I}_\mathcal{O}}$ because $\mathcal{I}_\mathcal{A}$ is a model of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$ by assumption.

(ii) $e' = e \cdot n$ for some integer n, where e' was introduced by applying some axiom $M \sqsubseteq \exists S'.N \in \Xi(\mathcal{T})$ with $S \subseteq S'$. We know from the inference rule ($\mathbf{R}_\sqsubseteq^r$) that $M \sqsubseteq \exists S' \sqcap r.N \in \Xi(\mathcal{T})$. Due to maximality of $M \sqsubseteq \exists S'.N$, we must have $S' \sqcap r = S'$, which implies $(e, e') \in r^{\mathcal{I}_\mathcal{O}}$.

(iii) $e = e' \cdot n$ for some integer n, where e was introduced by applying some axiom $M \sqsubseteq \exists S'.N \in \Xi(\mathcal{T})$ with $S^- \subseteq S'$. Note that $S^- \sqsubseteq r^- \in \mathcal{T}$ (see preliminaries). We know from the inference rule ($\mathbf{R}_\sqsubseteq^r$) that $M \sqsubseteq \exists S' \sqcap r^-.N \in \Xi(\mathcal{T})$. Again, due to maximality of $M \sqsubseteq \exists S'.N$, we must have $S' \sqcap r^- = S'$, which implies $(e', e) \in (r^-)^{\mathcal{I}_\mathcal{O}}$ and $(e, e') \in r^{\mathcal{I}_\mathcal{O}}$.

(6) Assume an axiom $A \sqsubseteq\ \leqslant 1\, r.B \in \mathcal{T}$ and a domain element $e \in A^{\mathcal{I}_\mathcal{O}}$. Suppose there is $e_1, e_2 \in \Delta^{\mathcal{I}_\mathcal{O}}$ such that $e_1 \neq e_2$, $\{(e, e_1), (e, e_2)\} \subseteq r^{\mathcal{I}_\mathcal{O}}$ and $\{e_1, e_2\} \subseteq B^{\mathcal{I}_\mathcal{O}}$. We have the following possible cases:

(i) $\{e_1, e_2\} \subseteq \mathsf{N_I}$. Then by the construction of $\mathcal{I}_\mathcal{O}$ we must have $e \in \mathsf{N_I}$. We arrive at a contradiction to the assumption that $\mathcal{I}_\mathcal{A}$ is a model of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$; the constraint representing $A \sqsubseteq\ \leqslant 1\, r.B \in \mathcal{T}$ must be violated.

(ii) $e_1, e \in \mathsf{N_I}$ and e_2 is of the form $e_2 = e \cdot n$ for some integer. Assume e_2 was introduced by applying an applicable axiom $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ at e. Note we have $e \in M^{\mathcal{I}_\mathcal{O}}$. By a rule of the last type in Table 3, we have that $e_1 \in N^{\mathcal{I}_\mathcal{A}}$ and $(e, e_1) \in S^{\mathcal{I}_\mathcal{A}}$. This shows that $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ was never applicable at e. Contradiction.

(iii) $e_2, e \in \mathsf{N_I}$ and e_1 is of the form $e_1 = e \cdot n$ for some integer. Symmetric to the above.

(iv) e_1, e_2 are of the form $e_1 = e \cdot n$ and $e_2 = e \cdot n'$. Suppose e_1, e_2 where introduced by applying axioms $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ and $M' \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ at e. Then by the construction of $\mathcal{I}_\mathcal{O}$ we have $r \in S$, $r \in S'$, $B \in N$ and $B \in N'$. Then by the inference rule ($\mathbf{R}_\leq$), we have $M \sqcap M' \sqcap A \sqsubseteq \exists S \sqcap S'.N \sqcap N' \in \Xi(\mathcal{T})$. Since $e \in (M \sqcap M' \sqcap A)^{\mathcal{I}_\mathcal{O}}$, we have a violation of applicability of $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ and $M' \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ at e, i.e. they are not maximal.

(v) $e = e_1 \cdot n$ and $e_2 = e \cdot n'$ obtained by applying some axioms $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ and $M' \sqsubseteq \exists S'.N' \in \Xi(\mathcal{T})$ at e_1 and e, respectively. By the construction of $\mathcal{I}_\mathcal{O}$, we have have $A \in N$, $r^- \in S$, $r \in S'$ and $B \in N'$. Then by the inference rule ($\mathbf{R}_\leq^-$), we have $M \sqcap B \sqsubseteq C \in \Xi(\mathcal{T})$ for all $C \in N'$ and also $M \sqcap B \sqsubseteq \exists S \sqcap (S')^-.N \in \Xi(\mathcal{T})$. Since $e_1 \in (M \sqcap B)^{\mathcal{I}_\mathcal{O}}$, we have $(S^-)^- \subset S$ by the maximality of $M \sqsubseteq \exists S.N$. Due to point (2) in this proof, we also have $e_1 \in C^{\mathcal{I}_\mathcal{O}}$ for all $C \in N'$. This shows that $M' \sqsubseteq \exists S'.N' \in \Xi(\mathcal{T})$ was not applicable at e, i.e. maximality violated.

(7) It remains to see that $\perp^{\mathcal{I}_\mathcal{O}} = \emptyset$. First note that $\mathsf{N_I} \cap \perp^{\mathcal{I}_\mathcal{O}} = \emptyset$ because $\mathcal{I}_\mathcal{A}$ is a model of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$. Thus it suffices to prove the following statement: if $e \cdot n \in \perp^{\mathcal{I}_\mathcal{O}}$, then also $e \in \perp^{\mathcal{I}_\mathcal{O}}$. Assume some $e \cdot n \in \perp^{\mathcal{I}_\mathcal{O}}$. Suppose $e \cdot n$ was introduced by applying an axiom $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$. Then by the definition of $\mathcal{I}_\mathcal{O}$, $\perp \in N$. By the inference rule ($\mathbf{R}_\perp$), we have $M \sqsubseteq \perp \in \Xi(\mathcal{T})$. Since $e \in M^{\mathcal{I}_\mathcal{O}}$, by point (2) in this proof we have $e \in \perp^{\mathcal{I}_\mathcal{O}}$.

It remains to see (b), i.e. that $\mathcal{I}_\mathcal{O}$ can be homomorphically embedded into any model $\mathcal{I}$ of $\mathcal{O}$. A homomorphism h from $\mathcal{I}_\mathcal{O}$ to $\mathcal{I}$ can be inductively defined as follows:

- $h(e) = e^\mathcal{I}$ for all $e \in \mathsf{N_I} \cap \Delta^{\mathcal{I}_\mathcal{O}}$. It is straightforward to see that $e_1 \in A^{\mathcal{I}_\mathcal{O}}$ and $(e_1, e_2) \in r^{\mathcal{I}_\mathcal{O}}$ imply $e_1 \in A^\mathcal{I}$ and $(e_1, e_2) \in r^\mathcal{I}$ for all $e_1, e_2 \in \mathsf{N_I}$, concepts A and roles r.

- Assume $e \cdot n \in \Delta^{\mathcal{I}_\mathcal{O}}$ was introduced in $\mathcal{I}_\mathcal{O}$ by an application of $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$. Note that $e \in M^{\mathcal{I}_\mathcal{O}}$. It suffices to define $h(e \cdot n) = e'$ where $e' \in \Delta^\mathcal{I}$ is an element such that $S \subseteq \{r \mid (h(e), e') \in r^\mathcal{I}\}$ and $N \subseteq \{A \mid e' \in A^\mathcal{I}\}$. Note that such e' exists. Indeed, by the induction hypothesis, $h(e) \in M^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\Xi(\mathcal{T})$, we must have $h(e) \in (\exists S.N)^\mathcal{I}$.

It remains to show that consistency of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$ implies consistency of $\mathcal{O}$. Assume $\mathcal{O}$ is inconsistent and suppose $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$ is consistent. Then there exists the least model $\mathcal{I}_\mathcal{A}$ of $\mathcal{A} \cup \mathsf{cr}(\mathcal{T})$, and thus $\mathcal{I}_\mathcal{O} = \mathit{chase}(MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T})), \Xi(\mathcal{T}))$ is defined. As we shown in (a), $\mathcal{I}_\mathcal{O} \models \mathcal{O}$. Contradiction. $\square$

In database terms, this means that checking consistency of $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ reduces to evaluating the (plain) DATALOG query $\mathsf{cr}(\mathcal{T})$ over the database $\mathcal{A}$. Note that $\Xi(\mathcal{T})$ can be computed in exponential time in size of $\mathcal{T}$: the calculus only infers axioms of the form $M \sqsubseteq B$ and $M \sqsubseteq \exists S.N$, where M, N are conjunctions of atomic concepts, B is atomic and S is a conjunction of roles. The number of such axiom is single exponential in the size of $\mathcal{T}$.

5 Rewriting Rules and Programs

The following is immediate from Propositions 1 and 2:

Theorem 3. *Let $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ be a Horn-$\mathcal{SHIQ}$ ontology. Then $\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*)$ is consistent iff $\mathcal{O}$ is consistent. Moreover, if $\mathcal{O}$ is consistent, then $ans(\mathcal{O}, q) = ans^{\mathcal{T}}(\mathcal{I}_{\mathcal{O}}, q)$ for every CQ q, where $\mathcal{I}_{\mathcal{O}} = chase(MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*), \Xi(\mathcal{T}^*))$.*

Computing $ans^{\mathcal{T}}(\mathcal{I}_{\mathcal{O}}, q)$ is still tricky because $\mathcal{I}_{\mathcal{O}}$ can be infinite. Hence we rewrite q into a set Q of CQs such that $ans^{\mathcal{T}}(\mathcal{I}_{\mathcal{O}}, q) = \bigcup_{q' \in Q} ans(MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*), q')$. That is, we only need to evaluate the queries in Q over the DATALOG program $\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*)$. Since this can be easily done directly in DATALOG, we have an algorithm for answering q over $\mathcal{O}$, which we later generalize to KBs.

5.1 Rewriting rules with simple roles only

We will first present a simplified version of our rewriting algorithm that rewrites a rule ρ assuming that r is a simple role for all atoms of the form $r(x, y)$ that occur in its body. This version can be explained more easily, and it will allow us to give a better explanation of the general algorithm.

The intuition is the following. Suppose that ρ has a non-distinguished variable x, and that there is some match π in $\mathcal{I}_{\mathcal{O}}$ such that $\pi(x)$ is an object in the 'tree part' introduced by the chase procedure and it has no descendant in the image of π, that is, $\pi(x)$ it is a leaf in the forest shaped image of ρ under π. Then for all atoms $r(y, x)$ of ρ, the "neighbor" variable y must mapped to the parent of $\pi(x)$. A rewrite step makes a choice of such an x, and employs an existential axiom from $\Xi(\mathcal{T})$ to 'clip off' x, eliminating all query atoms that mention it. By repeating this procedure, we can clip off all variables matched in the tree part and obtain a rule that has a match in $MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T}))$.

The one-step clipping off of a variable works as follows. For a CQ ρ and a Horn-$\mathcal{ALCHIQ}^{\sqcap}$ TBox $\mathcal{T}$, we write $\rho \rightarrow_{\mathcal{T}} \rho'$ if ρ' can be obtained from ρ in the following steps:

(S1) Select in ρ an arbitrary non-distinguished variable x such that there are no atoms of the form $r(x, x)$ in ρ.

(S2) Replace each role atom $r(x, y)$ in ρ, where y is arbitrary, by the atom $\mathsf{inv}(r)(y, x)$.

(S3) Let $V_p = \{y \mid \exists r : r(y, x) \in body(\rho)\}$, and select some $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T})$ such that

   (a) $\{r \mid r(y, x) \in body(\rho) \wedge y \in V_p\} \subseteq S$, and

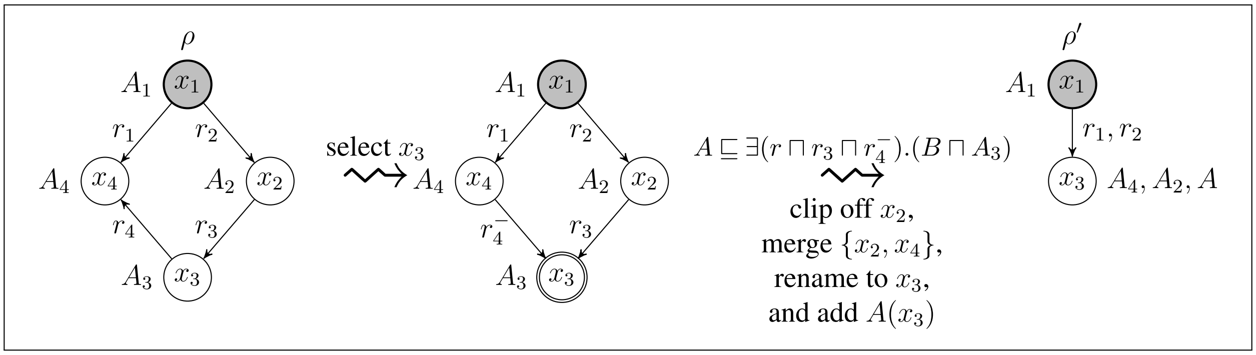


Figure 1: Example 2, query rewriting with only simple roles

(b) $\{A \mid A(x) \in body(\rho)\} \subseteq N$.

(S4) Drop from ρ each atom containing x.

(S5) Rename each $y \in V_p$ of ρ by x.

(S6) Add the atoms $\{A(x) \mid A \in M\}$ to $body(\rho)$.

We illustrate the rewriting step with two examples:

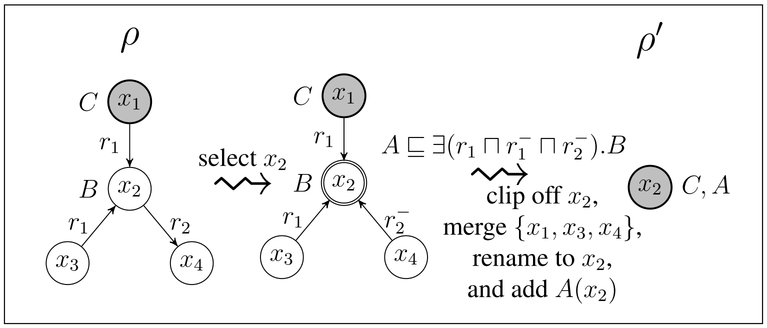
Example 2. Let $\rho : q(x_1) \leftarrow A_1(x_1), r_2(x_1, x_2), A_2(x_2), r_3(x_2, x_3), A_3(x_3), r_1(x_1, x_4), A_4(x_4), r_4(x_3, x_4)$ in Figure 1, and assume that $A \sqsubseteq \exists(r \sqcap r_3 \sqcap r_4^-).(B \sqcap A_3)$ is in $\Xi(\mathcal{T})$ and that all roles are simple. We choose the variable x_3, replace $r_4(x_3, x_4)$ by $r_4^-(x_4, x_3)$ in step (S2), and get $V_p = \{x_2, x_4\}$. Intuitively, if $\pi(x_3)$ is a leaf in a tree-shaped match π, then x_2 and x_4 must both be mapped to the parent of $\pi(x_3)$. Since the GCI $A \sqsubseteq \exists(r \sqcap r_3 \sqcap r_4^-).(B \sqcap A_3)$ in $\Xi(\mathcal{T})$ satisfies (S3.a,b), we can drop the atoms containing x_3 from ρ, and perform (S5) and (S6) to obtain the rewritten query $\rho' : q(x_1) \leftarrow A_1(x_1), r_1(x_1, x_3), r_2(x_1, x_3), A_4(x_3), A_2(x_3), A(x_3)$.

Example 3. In this example, illustrated in Figure 2a, we again assume that all roles are simple. Let $\rho : q(x_1) \leftarrow C(x_1), B(x_2), r_1(x_1, x_2), r_1(x_3, x_2), r_2(x_2, x_4)$, and assume $A \sqsubseteq \exists(r_1 \sqcap r_1^- \sqcap r_2^-).B \in \Xi(\mathcal{T}^*)$. In (S1) we select the non-distinguished variable x_2. Next, in (S2), we replace $r_2(x_2, x_4)$ by $r_2^-(x_4, x_2)$. Since all roles are simple, we do nothing in (S3). In (S4) we choose $V_\ell = \{x_2\}$ and $V_p = \{x_1, x_3, x_4\}$, and in (S5), $A \sqsubseteq \exists(r_1 \sqcap r_1^- \sqcap r_2^-).B$. Then we clip off x_2 in (S6), merge all variables in V_p and rename them to x_2 in (S7), and add $A(x_2)$ in (S8), to obtain $\rho' : q(x_2) \leftarrow C(x_2), A(x_2)$.

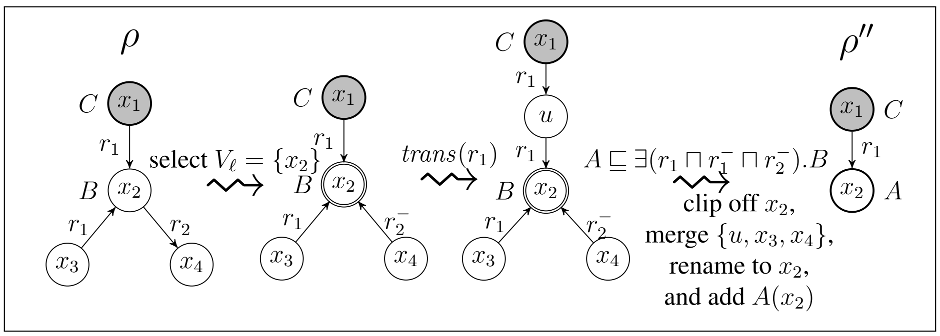
5.2 Rewriting arbitrary rules

Now we present the rewriting algorithm for the general case, and show that it is sound and complete.

As above, suppose that ρ has a match and $\pi(x)$ is a leaf of its forest shaped image, for some variable x. The most significant difference in the presence of non-simple roles is that if the query



(a) Example 3: Query rewriting with only simple roles



(b) Example 4: Query rewriting with non-simple roles

Figure 2: Examples of query rewriting

has an atom $r(y, x)$ and r is non-simple, then $\pi(y)$ is not necessarily the parent p of $\pi(x)$. Instead, $\pi(y)$ can be an ancestor of p, or $\pi(y) = \pi(x)$ may hold. Hence, instead of just x, we guess a set of distinguished variables V_ℓ that are mapped together at some leaf node $\pi(x)$. Then we guess a subset of the neighbor variables whose match is higher up in the tree, and for them we introduce an 'intermediate' variable u that can be matched at the parent p. In this way we can forget about the variables that are matched to ancestors of p, and assume that all the neighbours V_p of the variables in V_ℓ are matched at p. We can then proceed similarly as above and clip off all variables in V_ℓ using an axiom from $M \sqsubseteq \exists S.N$ that ensures the existence of a match for them. This axiom must now also ensure that $\pi(x)$ is an r-successor of itself for every atom $r(x, y)$ such that $x, y \in V_\ell$. This is verified by the new condition (S5c), which relies on the fact that a node e is an r-successor of itself in $\mathcal{I}_\mathcal{O}$ iff both $e, e' \in s^{\mathcal{I}_\mathcal{O}}$ and $e', e \in s^{\mathcal{I}_\mathcal{O}}$ hold for some transitive $s \sqsubseteq^*_\mathcal{T} r$, where e' is either the parent or a child of e in $\mathcal{I}_\mathcal{O}$.

Definition 6. For a rule ρ and a Horn-$\mathcal{SHIQ}$ TBox $\mathcal{T}$, we write $\rho \rightarrow_\mathcal{T} \rho'$ if ρ' is obtained from ρ by the following steps:

(S1) Select an arbitrary non-empty set V_ℓ of non-distinguished variables in ρ.

(S2) Replace each role atom $r(x, y)$ in ρ, where $x \in V_\ell$ and $y \notin V_\ell$ is arbitrary, by the atom $\mathsf{inv}(r)(y, x)$.

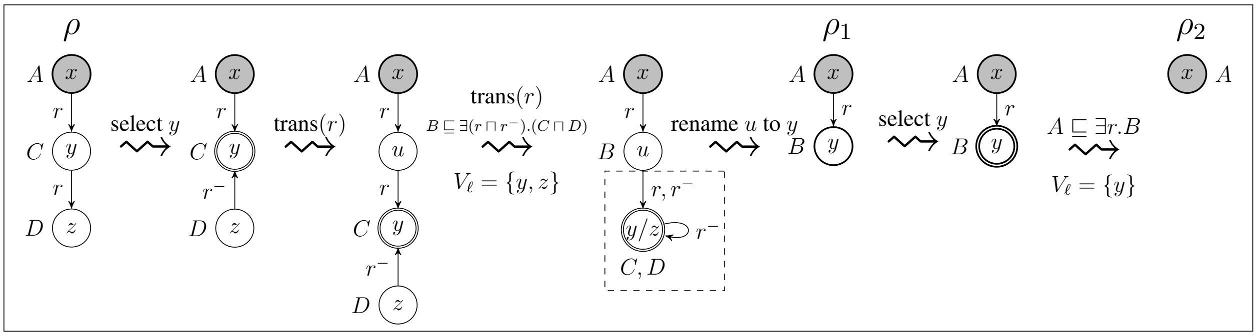


Figure 3: Example 5

(S3) For each atom $\alpha = s(y,x)$ in ρ, where where $x \in V_\ell$, $y \notin V_\ell$ is arbitrary and s is non-simple, either leave α untouched or replace it by two atoms $r(y,u), r(u,x)$, where u is a fresh variable and r is a transitive role with $r \sqsubseteq^*_\mathcal{T} s$.

(S4) Let $V_p = \{y \mid \exists r : r(y,x) \in \mathit{body}(\rho) \wedge x \in V_\ell \wedge y \notin V_\ell\}$.

(S5) Select some $M \sqsubseteq \exists S.N \in \Xi(\mathcal{T}^*)$ such that

(a) $\{r \mid r(y,x) \in \mathit{body}(\rho) \wedge x \in V_\ell \wedge y \in V_p\} \subseteq S$,

(b) $\{A \mid A(x) \in \mathit{body}(\rho) \wedge x \in V_\ell\} \subseteq N$, and

(c) for each atom $r(x,y)$ in $\mathit{body}(\rho)$ with $x, y \in V_\ell$ there is a transitive $s \sqsubseteq^*_\mathcal{T} r$ such that

i. $\{s, s^-\} \subseteq S$, or

ii. there is an axiom $M' \sqsubseteq \exists S'.N' \in \Xi(\mathcal{T}^*)$ such that $M' \subseteq N$ and $\{s, s^-\} \subseteq S'$.

(S6) Drop each atom from ρ containing a variable from V_ℓ.

(S7) Select some $x \in V_\ell$ and rename each $y \in V_p$ of ρ by x.

(S8) Add the atoms $\{A(x) \mid A \in M\}$ to ρ.

We write $\rho \rightarrow^*_\mathcal{T} \rho'$ if ρ' can be obtained from ρ by finitely many rewrite iterations. We let $\mathsf{rew}_\mathcal{T}(\rho) = \{\rho' \mid \rho \rightarrow^*_\mathcal{T} \rho'\}$. For a set $\mathcal{P}$ of rules, $\mathsf{rew}_\mathcal{T}(\mathcal{P}) = \bigcup_{\rho \in \mathcal{P}} \mathsf{rew}_\mathcal{T}(\rho)$.

Example 4 (ctd)**.** Recall $\rho : q(x_1) \leftarrow C(x_1), B(x_2), r_1(x_1,x_2), r_1(x_3,x_2), r_2(x_2,x_4)$ and $A \sqsubseteq \exists(r_1 \sqcap r_1^- \sqcap r_2^-).B \in \Xi(\mathcal{T}^*)$ from Example 3, but now assume that $\mathit{trans}(r_1) \in \mathcal{T}$. As shown in Figure 2b, in (S1) we choose $V_\ell = \{x_2\}$, and in (S3) we choose to replace $r_1(x_1,x_2)$ with $r_1(x_1,u), r_1(u,x_2)$. In (S4) we get $V_p = \{u, x_3, x_4\}$. Then we proceed similarly as above to obtain $\rho'' : q(x_1) \leftarrow C(x_1), r_1(x_1,x_2), A(x_2)$.

Example 5. Assume $\mathcal{T} = \{r \sqsubseteq r^-,\ \mathit{trans}(r),\ A \sqsubseteq \exists r.B,\ B \sqsubseteq \exists r.C,\ C \sqsubseteq D\}$. Let $\rho : q(X) \leftarrow A(x), r(x,y), C(y), D(z), r(y,z)$. By saturation rules $\mathbf{R}^c_\sqsubseteq$ and $\mathbf{R}^r_\sqsubseteq$, we have $B \sqsubseteq \exists(r \sqcap r^-).(C \sqcap D) \in \Xi(\mathcal{T})$.

In the first round, in (S1) we select y. In (S2), we replace $r(y,z)$ by $r^-(z,y)$. In (S3), as r is transitive, we replace $r(x,z)$ by $r(x,u)$ and $r(u,y)$. In (S4), we choose $V_\ell = \{y,z\}$, $V_p = \{u\}$. In (S5), we choose $B \sqsubseteq \exists(r \sqcap r^-).(C \sqcap D) \in \Xi(\mathcal{T})$, which satisfies (S5.a), (S5.b), and (S5.c1). In (S6), we drop atoms containing y or z from $body(\rho)$. In (S7), we rename u to y. Finally in (S8), we add $B(y)$ to the body and get $\rho_1 : q(x) \leftarrow A(x), r(x,y), B(y)$.

In the second round, we select y in (S1), $V_\ell = \{y\}, V_p = \{x\}$ in (S3), and $A \sqsubseteq \exists r.B$ in (S5). Following the similar steps, we get another rewritten rule $\rho_2 : q(X) \leftarrow A(x)$.

The following is crucial:

Theorem 4. *Assume a consistent Horn-$\mathcal{SHIQ}$ ontology $\mathcal{O} = (\mathcal{T},\mathcal{A})$ and a conjunctive query q. Then $ans(\mathcal{O},q) = \bigcup_{q' \in \mathsf{rew}_\mathcal{T}(q)} ans(MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*)), q')$.*

Proof. Let $\mathcal{I}_\mathcal{O} = chase(\mathcal{J}, \Xi(\mathcal{T}^*))$, where $\mathcal{J} = MM(\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*))$. It suffices to show $ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q) = ans(\mathcal{J}, \mathsf{rew}_\mathcal{T}(q))$.

We first show $ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q) \supseteq ans(\mathcal{J}, \mathsf{rew}_\mathcal{T}(q))$. Suppose $h(\vec{x})$ is the head atom of q. Assume a tuple $\vec{u} \in ans(\mathcal{J}, \mathsf{rew}_\mathcal{T}(q))$. Then there is a query $q' \in \mathsf{rew}_\mathcal{T}(q)$ and a match $\pi_{q'}$ for q' in $\mathcal{J}$ such that $\vec{u} = \pi_{q'}(\vec{x})$. By the construction of $\mathcal{I}_\mathcal{O}$, we also have $\vec{u} \in ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q')$. If $q' = q$, then we are done. Suppose $q' \neq q$. Then there is $n > 0$ such that $q_0 \to_\mathcal{T} q_1, \cdots, q_{n-1} \to_\mathcal{T} q_n$ with $q_0 = q$ and $q_n = q'$. Thus to prove the claim it suffices to show that $\vec{u} \in ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q_i)$ implies $\vec{u} \in ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q_{i-1})$, where $0 < i \leq n$.

Suppose π_{q_i} is a match for q_i in $\mathcal{I}_\mathcal{O}$ with $\vec{u} = \pi_{q'}(\vec{x})$, i.e. $\vec{u} \in ans^\mathcal{T}(\mathcal{I}_\mathcal{O}, q_i)$. Let V_ℓ be the set chosen in (S1), let $x \in V_\ell$ be the variable chosen in (S7), and let $M \sqsubseteq \exists S.N$ be the axiom chosen in (S5). Moreover, let $d = \pi_{q_i}(x)$. Due to step (S8) in the rewriting and the fact that $\mathcal{I}_\mathcal{O}$ is a model of $\mathcal{O}$, we have $d \in (\exists S.N)^{\mathcal{I}_\mathcal{O}}$. Then there is $d' \in \Delta^{\mathcal{I}_\mathcal{O}}$ such that $(d,d') \in S^{\mathcal{I}_\mathcal{O}}$ and $d' \in N^{\mathcal{I}_\mathcal{O}}$. Define the mapping $\pi_{q_{i-1}}$ for the variables of q_{i-1} as follows: (a) $\pi_{q_{i-1}}(z) = d'$ for all variables $z \in V_\ell$, (b) $\pi_{q_{i-1}}(u) = d$ for all variables $u \in V_p$, and (c) $\pi_{q_{i-1}}(z) = \pi_{q_i}(z)$ for the remaining variables z. Then $\pi_{q_{i-1}}$ is a match for q_{i-1} in $\mathcal{I}_\mathcal{O}$. To see this, assume an atom α in q_{i-1}. We show that $\pi_{q_{i-1}}$ makes α true in $\mathcal{I}_\mathcal{O}$. There can be two possibilities:

(i) α has an occurrence of a variable from V_ℓ. In this case we have 3 more possibilities:

  (a) α is a unary atom of the form $\alpha = A(z)$. Then $z \in V_\ell$ and $\pi_{q_{i-1}}(z) = d'$ by construction of $\pi_{q_{i-1}}$. As noted above, $d' \in N^{\mathcal{I}_\mathcal{O}}$. By (S5.b) we have $A \in N$.

  (b) α is a binary atom $\alpha = r(y,x)$, where $y \in V_p$. We know $\pi_{q_{i-1}}(y) = d$ and $\pi_{q_{i-1}}(x) = d'$. As noted above, $(d,d') \in S^{\mathcal{I}_\mathcal{O}}$. By (S5.b) we have $r \in S$.

  (c) α is a binary atom $\alpha = r(y,x)$, where $y \in V_\ell$. We know $\pi_{q_{i-1}}(y) = \pi_{q_{i-1}}(x) = d'$. By (S5.c), there is a transitive $s \sqsubseteq^*_\mathcal{T} r$ such that $\{s, s^-\} \subseteq S$, or there is an axiom $M' \sqsubseteq \exists S'.N' \in \Xi(\mathcal{T}^*)$ such that $M' \subseteq N$ and $\{s, s^-\} \subseteq S'$. Since $\mathcal{I}_\mathcal{O}$ is a model of $\mathcal{O}$, we have $(d',d') \in r^{\mathcal{I}_\mathcal{O}}$.

(ii) α does not have an occurrence of a variable from V_ℓ. We distinguish the following cases:

  (a) α has no variables from V_p. Then $\alpha \in body(q_i)$ and the claim follows from (c) in the definition of $\pi_{q_{i-1}}$.

(b) α is a unary atom $\alpha = A(u)$ with $u \in V_p$, which was replaced by $A(x)$ in (S7). By construction of $\pi_{q_{i-1}}$ we have $\pi_{q_{i-1}}(u) = d$. As $A(x) \in body(q_i)$, we have that $\pi_{q_i}(x) = d$ implies $d \in A^{\mathcal{I}_{\mathcal{O}}}$ as desired.

(c) α is a binary atom $\alpha = r(u,z)$ with $u \in V_p$ and $z \notin V_p$, which was replaced by $r(x,z)$ in (S7). Since π_{q_i} is a match for q_i in $\mathcal{I}_{\mathcal{O}}$ and $r(x,z) \in body(q_i)$, π_{q_i} satisfies $r(x,z)$. By construction of $\pi_{q_{i-1}}$ we have $\pi_{q_{i-1}}(u) = \pi_{q_i}(x) = d$ and $\pi_{q_{i-1}}(z) = \pi_{q_i}(z)$, hence π_{q_i} satisfies $r(u,z)$.

(d) the cases $\alpha = r(z,u)$ with either $u \in V_p$ and $z \notin V_p$, or $\{z,u\} \subseteq V_p$, are both analogous to the previous one.

We show $ans^{\mathcal{T}}(\mathcal{I}_{\mathcal{O}}, q) \subseteq ans(\mathcal{J}, \mathsf{rew}_{\mathcal{T}}(q))$. To show this we need some book-keeping when chasing $\mathcal{J}$ w.r.t. $\Xi(\mathcal{T}^*)$. We prescribe the naming of fresh domain elements introduced during the chase procedure. In particular, if d is a successor of e according to Definition 5, then d is an expression of the form $e \cdot n$, where n is a integer. For $d \in \Delta^{\mathcal{J}}$, let $|d| = 0$. For the elements $w \cdot n \in \Delta^{\mathcal{I}_{\mathcal{O}}}$, let $|w \cdot n| = |w| + 1$.

Suppose $h(\vec{x})$ is the head atom of q. Assume a tuple $\vec{u} \in ans^{\mathcal{T}}(\mathcal{I}_{\mathcal{O}}, q)$. By definition, there is match π_q for q in $\mathcal{I}_{\mathcal{O}}$ such that $\vec{u} = \pi_q(\vec{x})$. We have to show that there exists $q' \in \mathsf{rew}_{\mathcal{T}}(q)$ and a match $\pi_{q'}$ for q' in $\mathcal{J}$ such that $\vec{u} = \pi_{q'}(\vec{x})$. For any match π' in $\mathcal{I}_{\mathcal{O}}$, let

$$deg(\pi') = \sum_{y \in rng(\pi')} |\pi'(y)|.$$

Then, given that $q \in \mathsf{rew}_{\mathcal{T}}(q)$, to prove the claim it suffices to prove the following statement: if $q_1 \in \mathsf{rew}_{\mathcal{T}}(q)$ has a match π_{q_1} for q_1 in $\mathcal{I}_{\mathcal{O}}$ such that $\vec{u} = \pi_{q_1}(\vec{x})$ and $deg(\pi_{q_1}) > 0$, then there exists $q_2 \in \mathsf{rew}_{\mathcal{T}}(q)$ that has a match π_{q_2} for q_2 in $\mathcal{I}_{\mathcal{O}}$ such that $\vec{u} = \pi_{q_2}(\vec{x})$ and $deg(\pi_{q_2}) < deg(\pi_{q_1})$.

Assume $q_1 \in \mathsf{rew}_{\mathcal{T}}(q)$ as above. Since $deg(\pi_{q_1}) > 0$ by assumption, there must exists a variable x of q_1 such that $\pi_{q_1}(x) \notin \mathsf{N_I}$. Take such an x for which there is no variable x' of q_1 with $\pi_{q_1}(x)$ being a prefix of $\pi_{q_1}(x')$. That is, there is no variable x' of q_1 with $\pi_{q_1}(x') = \pi_{q_1}(x) \cdot w$ for some w. Intuitively, the image of π_{q_1} induces a subforest in $\mathcal{I}_{\mathcal{O}}$; the variable x is mapped into a leaf node in this forest.

Let $d_x = \pi_{q_1}(x)$, and d_p be the parent element of d_x, i.e. $d_x = d_p \cdot n$ for some integer n. We know from the construction of $\mathcal{I}_{\mathcal{O}}$ that d_x was introduced by an application of an axiom $ax = M \sqsubseteq \exists S.N \in \Xi(\mathcal{T}^*)$ such that $d_p \in M^{\mathcal{I}_{\mathcal{O}}}$. We take a query q_2 obtained from q_1 as follows:

- For Step (S1) choose $V_\ell = \{y \in var(q_1) \mid \pi_{q_1}(y) = d_x\}$ (note that since $d_x \notin \mathsf{N_I}$, all such y are non-distinguished).

- For Step (S3) let $\Gamma = \{s(y,x) \in q_1 \mid x \in V_\ell \wedge \pi_{q_1}(y) \neq d_x \wedge \pi_{q_1}(y) \neq d_p\}$ be the set of atoms we choose to rewrite. Note that due to the selection of the atoms in Γ and since π_{q_1} is a $\mathcal{T}$-match for q_1, by definition of $\mathcal{T}$-matches, for every atom $s(y,x) \in \Gamma$ there exists a transitive role r_s with $r_s \sqsubseteq^*_{\mathcal{T}} s$ such that there is an r_s-path from $\pi_{q_1}(y)$ to $\pi_{q_1}(x)$. Using this role r_s, we rewrite $s(y,x)$ into $r_s(y,u), r_s(u,x)$. Observe that, since d_p is the parent of d_x in $\mathcal{I}_{\mathcal{O}}$ and $\pi_1(y) \neq d_p$, then d_p is in the r_s-path from $\pi_{q_1}(y)$ to $\pi_{q_1}(x)$ and the following holds:

($\dagger$) there is an r_s-path from $\pi_{q_1}(y)$ to d_p, and $(d_p, d_x) \in {r_s}^{\mathcal{I}_\mathcal{O}}$.

Observe also that if $\Gamma \neq \emptyset$, then in Step (S4) we get that $u \in V_p$.

- For Step (S5), choose ax given above. To see that (S5.a) holds, take any $r(y, x)$ where $x \in V_\ell$ and $y \in V_p$. We have to show $r \in S$, and we have two cases:

  i. $y \in var(q_1)$ and $\pi_{q_1}(y) = d_p$. Since π_{q_1} is a $\mathcal{T}$-match for q_1 and d_x is a successor of d_p, we must have $(\pi_{q_1}(y), \pi_{q_1}(x)) \in r^{\mathcal{I}_\mathcal{O}}$. Then due to the construction of $\mathcal{I}_\mathcal{O}$, $r \in S$.

  ii. if $y = u$ is the fresh variable introduced in Step (S3), then r is the role r_s chosen above and by ($\dagger$) we have $(d_p, d_x) \in r^{\mathcal{I}_\mathcal{O}}$, which implies $r \in S$ due to the construction of $\mathcal{I}_\mathcal{O}$.

  To see that (S5.b) holds, take any $A(z)$ where $z \in V_\ell$. We have to show $A \in N$. Since π_{q_1} is a $\mathcal{T}$-match for q_1, we have $\pi_{q_1}(z) \in A^{\mathcal{I}_\mathcal{O}}$. Since $\pi_{q_1}(z) = d_x$, by construction of $\mathcal{I}_\mathcal{O}$ we have $A \in N$.

  Finally, we check that (S5.c) holds. Take an atom $r(x, y)$ in q_1 such that $x, y \in V_\ell$. Since $\pi_{q_1}(z) = \pi_{q_1}(x) = d_x$ and π_{q_1} is a $\mathcal{T}$-match, we have a "self-loop" from d_x to itself, that is, there is a transitive $s \sqsubseteq^*_\mathcal{T} r$ and an s-path from d_x to d_x. This path must pass through a domain element $d \neq d_x$, an in particular it muss pass a d that is either the parent d_x or some child of d_x. Due to the construction of $\mathcal{I}_\mathcal{O}$, (i.) is satisfied in the former case and (ii.) is satisfied in the latter case.

Finally, a match π_{q_2} for q_2 in $\mathcal{I}_\mathcal{O}$ such that $\vec{u} = \pi_{q_2}(\vec{x})$ and $deg(\pi_{q_2}) < deg(\pi_{q_1})$ is obtained from π_{q_1} by setting (a) $\pi_{q_2}(z) = \pi_{q_1}(z)$ for all z of q_2 with $z \neq x$, and (b) $\pi_{q_2}(x) = d_p$. It is easy to check that π_{q_2} is a $\mathcal{T}$-match for q_2 because q_2 is intuitively a subquery of q_1. Observe that $vars(q_2) \subseteq vars(q_1)$ because any new variable introduced in Step (S3) is eliminated in Step (S7). Hence, $deg(\pi_{q_2}) < deg(\pi_{q_1})$ follows from the fact that (i) $|\pi_{q_2}(z)| = |\pi_{q_1}(z)|$ for all z of q_2 with $z \neq x$, and (ii) $|\pi_{q_2}(x)| = |\pi_{q_1}(x)| - 1$.

□

By the above reduction, we can answer q over $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ by posing $\mathsf{rew}_\mathcal{T}(q)$ over the DATALOG program $\mathcal{A} \cup \mathsf{cr}(\mathcal{T}^*)$. The method also applies to KBs $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$, where $\mathcal{T}$ is in Horn-$\mathcal{SHIQ}$ and $\mathcal{P}$ is weakly DL-safe. The ground atomic consequences of $\mathcal{K}$ can be collected by fixed-point computation: until no new consequences are derived, pose rules in $\mathcal{P}$ as CQs over $(\mathcal{T}, \mathcal{A})$ and put the obtained answers into $\mathcal{A}$. If we employ the rewriting in Definition 6, this computation can achieved using a plain DATALOG program.

Theorem 5. *For a ground atom α over a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$ where $\mathcal{T}$ is a Horn-$\mathcal{SHIQ}$ TBox and $\mathcal{P}$ is weakly DL-safe, we have $(\mathcal{T}, \mathcal{A}, \mathcal{P}) \models \alpha$ iff $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_\mathcal{T}(\mathcal{P}) \cup \mathcal{A} \models \alpha$.*

Proof. First of all, let $\mathcal{K}_1 \models^\mathcal{T} \alpha_1$ be defined as $\mathcal{K}_1 \models \alpha_1$ but using the notion of a $\mathcal{T}$-match instead of a (plain) match. Since $(\mathcal{T}, \mathcal{A}, \mathcal{P}) \models \alpha$ iff $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^\mathcal{T} \alpha$, it suffices to show $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^\mathcal{T} \alpha$ iff $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_\mathcal{T}(\mathcal{P}) \cup \mathcal{A} \models \alpha$.

Let $\mathcal{P}' = \mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_\mathcal{T}(\mathcal{P})$.

For the "if" direction, the interesting case is where $(\mathcal{T}^*, \mathcal{A})$ is consistent. Note first that $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^{\mathcal{T}} \alpha'$ for all $\alpha' \in \mathcal{A}$. Hence, intuitively, it suffices to show that the rules of $\mathcal{P}'$ applied on $\mathcal{A}$ derive consequences of $(\mathcal{T}^*, \mathcal{A}, \mathcal{P})$. In particular, assume a rule

$$r = h(\vec{u}) \leftarrow b_1(\vec{v_1}), \ldots, b_m(\vec{v_m})$$

in $\mathcal{P}'$ and take a mapping $\pi : vars(r) \to \mathsf{N_I}$. To prove the claim it suffices to show that $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^{\mathcal{T}} b_1(\pi(\vec{v_1})), \ldots, (\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^{\mathcal{T}} b_m(\pi(\vec{v_m}))$ implies $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^{\mathcal{T}} h(\pi(\vec{u}))$.

The statement is straightforward if r is a rule in $\mathsf{cr}(\mathcal{T}^*)$, because $\mathsf{cr}(\mathcal{T}^*)$ encodes a subset of $\Xi(\mathcal{T}^*)$, which contains only logical consequences of $\mathcal{T}^*$.

Suppose $r \in \mathsf{rew}_{\mathcal{T}}(r')$, for some rule $r' \in \mathcal{P}$. Let $\mathcal{K}' = (\mathcal{T}^*, \mathcal{A}', \mathcal{P})$, where

$$\mathcal{A}' = \mathcal{A} \cup \{b_1(\pi(\vec{v_1})), \ldots, b_m(\pi(\vec{v_m})\}.$$

By applying Theorem 4, we get $h(\pi(\vec{u})) \in ans((\mathcal{T}^*, \mathcal{A}'), r')$. Hence, $\mathcal{K}' \models^{\mathcal{T}} h(\pi(\vec{u}))$. Since $\mathcal{K}' \equiv (\mathcal{T}^*, \mathcal{A}, \mathcal{P})$ due to the induction hypothesis, we also get $(\mathcal{T}^*, \mathcal{A}, \mathcal{P}) \models^{\mathcal{T}} h(\pi(\vec{u}))$.

We prove the "only if" direction. The only interesting case is where $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A}$ is consistent. In this case, it suffices to show the existence of a model $\mathcal{I}$ of $(\mathcal{T}^*, \mathcal{A}, \mathcal{P})$ such that $\mathcal{I} \not\models \alpha$ for all α such that $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A} \not\models \alpha$. Let $\mathcal{A}'$ be the set of all ground α such that $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A} \models \alpha$. We let $\mathcal{I} = chase(\mathcal{A}', \Xi(\mathcal{T}^*))$. Since the chase procedure does change the gound atoms that are entailed, $\mathcal{I} \not\models \alpha$ for all α such that $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A} \not\models \alpha$. It only remains to see that

(a) $\mathcal{I} \models (\mathcal{T}^*, \mathcal{A})$. Due to consistency of $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A}$, we also have that $\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}'$ is consistent. Due to Theorem 2, it suffices to show that $\mathcal{A}' = MM(\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}')$. Trivially, $\mathcal{A}' \subseteq MM(\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}')$. For $\mathcal{A}' \supseteq MM(\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}')$, assume there is $\beta \in MM(\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}')$ with $\beta \notin \mathcal{A}'$. Then β is derived via a rule $r \in \mathsf{cr}(\mathcal{T}^*)$ using some match π in $\mathcal{A}'$. Then it must be the case that $\beta \in \mathcal{A}'$ because by construction $\mathcal{A}'$ is closed under the rules in $\mathsf{cr}(\mathcal{T}^*)$.

(b) $\mathcal{I} \models \mathcal{P}$. Assume a rule $r \in \mathcal{P}$ with a mapping π from variables of r to $\Delta^{\mathcal{I}}$ such that $\mathcal{I} \models b(\pi(\vec{v}))$ for each body atom $b(\vec{v})$ of r. We have to show that $\mathcal{I} \models h(\pi(\vec{u}))$, where $h(\vec{u})$ is the head of r. Due to weak DL-safety of $\mathcal{P}$, $\pi(x) \in \mathsf{N_I}$ for each variable x in $\vec{u}$. In other words, π is an ordinary match for a conjunctive query. In particular, $\pi(\vec{v}) \in ans((\mathcal{T}^*, \mathcal{A}'), r)$ since $\mathcal{A}' = MM(\mathsf{cr}(\mathcal{T}^*) \cup \mathcal{A}')$. Then due to Theorem 4, we have a match π' for some $r' \in \mathsf{rew}_{\mathcal{T}}(\mathcal{P})$ in $\mathcal{A}'$. Since $\mathcal{A}'$ is closed under rules in $\mathsf{rew}_{\mathcal{T}}(\mathcal{P})$, we have $h(\pi(\vec{u})) \in \mathcal{A}'$ and thus $\mathcal{I} \models h(\pi(\vec{u}))$.

□

The algorithm obtained by the above reduction is worst-case optimal in terms of combined and data complexity.

Theorem 6. *For a ground atom α over a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$ where $\mathcal{T}$ is a Horn-$\mathcal{SHIQ}$ TBox and $\mathcal{P}$ is weakly DL-safe, checking $(\mathcal{T}, \mathcal{A}, \mathcal{P}) \models \alpha$ is* EXPTIME*-complete in general, and* PTIME*-complete when only the size of $\mathcal{A}$ is counted (i.e. in data complexity).*

Proof. By Theorem 5, checking $(\mathcal{T}, \mathcal{A}, \mathcal{P}) \models \alpha$ is equivalent to deciding $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A} \models \alpha$. We analyze the computational cost of the latter check.

We first recall that $\Xi(\mathcal{T}^*)$ can be computed in exponential time in size of $\mathcal{T}$ and is independent from $\mathcal{A}$: the calculus in Table 2 only infers axioms of the form $M \sqsubseteq B$ and $M \sqsubseteq \exists S.N$, where M, N are conjunctions of atomic concepts, B is atomic and S is a conjunction of roles. The number of such axiom is single exponential in the size of $\mathcal{T}$.

Observe that $\mathsf{rew}_{\mathcal{T}}(\mathcal{P})$ is finite and computable in time exponential in the size of $\mathcal{T}$ and $\mathcal{P}$: rules in $\mathsf{rew}_{\mathcal{T}}(\rho)$, where $\rho \in \mathcal{P}$, use only relation names and variables that occur in ρ and $\mathcal{T}$ (fresh variables introduced in (S3) are eliminated in (S6) and (S7)). Hence, the size of each rule resulting from a rewrite step is of size polynomial in the size of $\mathcal{T}$ and $\mathcal{P}$, and thus the number of rules in $\mathsf{rew}_{\mathcal{T}}(\mathcal{P})$ is at most exponential in the size of $\mathcal{T}$ and $\mathcal{P}$. The size of $\mathsf{rew}_{\mathcal{T}}(\mathcal{P})$ is constant when data complexity is considered.

Furthermore, the *grounding* of $\mathsf{cr}(\mathcal{T}^*) \cup \mathsf{rew}_{\mathcal{T}}(\mathcal{P}) \cup \mathcal{A}$ is exponential in the size of $\mathcal{K}$, but polynomial for fixed $\mathcal{T}$ and $\mathcal{P}$. By the complexity of DATALOG, it follows that the algorithm resulting from Theorem 5 is exponential in combined but polynomial in data complexity.

The above complexity result is worst-case optimal, and applies already to plain conjunctive queries [Eiter *et al.*, 2008b]. □

6 Related Work and Conclusion

Since Calvanese *et al.* (2007b) introduced query rewriting in their seminal work on $\mathcal{DL}$-*Lite*, many query rewriting techniques have been developed and implemented, e.g. (Perez-Urbina et al. 2009, Rosati and Almatelli 2010, Chortaras et al. 2011, Gottlob et al. 2011), usually aiming at an optimized rewriting size. Some of them also go beyond $\mathcal{DL}$-*Lite*; e.g. Perez-Urbina et al. cover $\mathcal{ELHI}$, while Gottlob et al. consider DATALOG$^{\pm}$. Most approaches rewrite a query into a (union of) CQs; Rosati and Almatelli generate a non-recursive Datalog program, while Perez-Urbina et al. produce a CQ for $\mathcal{DL}$-*Lite* and a (recursive) Datalog program for DLs of the $\mathcal{EL}$ family. Our approach rewrites a CQ into a union of CQs, but generates (possibly recursive) DATALOG rules to capture the TBox.

Our technique resembles Rosati's [2007] for CQs in $\mathcal{EL}$, which replaces query atoms by existential concepts, then applies some TBox saturation and translates the rewritten queries and the TBox into Datalog. The main difference is that in Rosati's technique the rewriting takes place *before* TBox saturation, resulting in an algorithm that is best-case exponential in the size of the query. This is avoided in our approach since a rewrite step occurs only if the saturated TBox has an applicable existential axiom. Another comparable technique is the *combined approach* of Lutz et al. [2009]. In order to do query answering in $\mathcal{EL}$ with off-the-shelf RDBMSs, the authors expand the data in the ABox 'materializing' a part of the canonical model that can be used for query answering after some query rewritings. Viewing our approach as a variation of the combined approach suggests an alternative query evaluation technique: we can first close the ABox under the rules in $\mathsf{cr}(\mathcal{T})$, and then evaluate the rewritten query $\mathsf{rew}(q)$ over the closed ABox.

Rewriting approaches for more expressive DLs are less common. The most notable exception is Hustadt et al.'s translation of $\mathcal{SHIQ}$ terminologies into disjunctive DATALOG [Hustadt *et al.*, 2007], which is implemented in the KAON2 reasoner. The latter can be used to answer queries over arbitrary ABoxes, but supports only instance queries. An extension to CQs (without transitive roles) is given in [Hustadt *et al.*, 2004], but it is not implemented. To our knowledge, also the extension of the rewriting in [Pérez-Urbina *et al.*, 2009] to nominals remains to be implemented [Pérez-Urbina *et al.*, 2010]. In [Ortiz *et al.*, 2010] a DATALOG rewriting is used to establish complexity bounds of standard reasoning in the Horn fragments of $\mathcal{SHOIQ}$ and $\mathcal{SROIQ}$, but it does not cover CQs.

To our knowledge, CQ answering for Horn-$\mathcal{SHIQ}$ and beyond has not been implemented before. Respective algorithms for full $\mathcal{SHIQ}$ were first given in [Glimm *et al.*, 2008] and (Calvanese et al. 2007a), and for Horn-$\mathcal{SHIQ}$ in [Eiter *et al.*, 2008b]. They are all of theoretical interest (to prove complexity results) but not suited for practical implementation, due to prohibitive sources of complexity.

7 Conclusion

We presented a rewriting-based algorithm for answering CQs over Horn-$\mathcal{SHIQ}$ ontologies. Our prototype implementation shows potential for practical applications, and further optimizations will improve it. Future versions of CLIPPER will support transitive roles and queries formulated in weakly DL-safe DATALOG, for which the theoretic foundations have been already developed.

As an interesting application, we mention that our method allows to improve reasoning with *DL-programs*, which loosely couple rules and ontologies [Eiter *et al.*, 2008a]. To avoid the overhead caused by the interaction of a rule reasoner and an ontology reasoner of traditional methods, the *inline evaluation* framework translates ontologies into rules [Heymans *et al.*, 2010; Eiter *et al.*, 2012a]. The techniques of this paper can be faithfully integrated into the inline evaluation framework to efficiently evaluate DL-programs involving Horn-$\mathcal{SHIQ}$ ontologies.

References

[Baader *et al.*, 2007] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, second edition, 2007.

[Calì *et al.*, 2009] Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz. Datalog±: a unified approach to ontologies and integrity constraints. In *Proceedings of the 12th International Conference on Database Theory*, ICDT '09, pages 14–30, New York, NY, USA, 2009. ACM.

[Calvanese *et al.*, 2007a] Diego Calvanese, Thomas Eiter, and Magdalena Ortiz. Answering regular path queries in expressive description logics: An automata-theoretic approach. In *AAAI*, pages 391–396. AAAI Press, 2007.

[Calvanese *et al.*, 2007b] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. Autom. Reasoning*, 39(3):385–429, 2007.

[Chortaras *et al.*, 2011] Alexandros Chortaras, Despoina Trivela, and Giorgos Stamou. Optimized query rewriting for OWL 2 QL. In *Proceedings of the 23rd International Conference On Automated Deduction*, CADE'11, pages 192–206, Berlin, Heidelberg, 2011. Springer-Verlag.

[Dantsin *et al.*, 2001] Evgeny Dantsin, Thomas Eiter, Georg Gottlob, and Andrei Voronkov. Complexity and expressive power of logic programming. *ACM Computing Surveys*, 33(3):374–425, 2001.

[Eiter *et al.*, 2008a] T. Eiter, G. Ianni, T. Lukasiewicz, R. Schindlauer, and H. Tompits. Combining answer set programming with description logics for the Semantic Web. *Artificial Intelligence*, 172(12-13):1495–1539, 2008.

[Eiter *et al.*, 2008b] Thomas Eiter, Georg Gottlob, Magdalena Ortiz, and Mantas Simkus. Query answering in the description logic Horn-$\mathcal{SHIQ}$. In Steffen Hölldobler, Carsten Lutz, and Heinrich Wansing, editors, *JELIA*, volume 5293 of *LNCS*, pages 166–179. Springer, 2008.

[Eiter *et al.*, 2009] Thomas Eiter, Carsten Lutz, Magdalena Ortiz, and Mantas Simkus. Query answering in description logics with transitive roles. In Craig Boutilier, editor, *IJCAI*, pages 759–764, 2009.

[Eiter *et al.*, 2012a] Thomas Eiter, Thomas Krennwallner, Patrik Schneider, and Guohui Xiao. Uniform evaluation of nonmonotonic DL-programs. In *FoIKS'12*, volume 7153 of *LNCS*, pages 1–22. Springer, March 2012.

[Eiter *et al.*, 2012b] Thomas Eiter, Magdalena Ortiz, and Mantas Simkus. Conjunctive query answering in the description logic SH using knots. *J. Comput. Syst. Sci.*, 78(1):47–85, 2012.

[Glimm *et al.*, 2006] Birte Glimm, Ian Horrocks, and Ulrike Sattler. Conjunctive query answering for description logics with transitive roles. In Bijan Parsia, Ulrike Sattler, and David Toman, editors, *Description Logics*, volume 189 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2006.

[Glimm *et al.*, 2008] Birte Glimm, Carsten Lutz, Ian Horrocks, and Ulrike Sattler. Conjunctive query answering for the description logic SHIQ. *J. Artif. Intell. Res. (JAIR)*, 31:157–204, 2008.

[Gottlob *et al.*, 2011] G. Gottlob, G. Orsi, and A. Pieris. Ontological queries: Rewriting and optimization. In *(ICDE), IEEE 27th International Conference on Data Engineering 2011*, pages 2 –13, april 2011.

[Heymans *et al.*, 2010] S. Heymans, T. Eiter, and G. Xiao. Tractable reasoning with DL-programs over Datalog-rewritable description logics. In *ECAI'10*, pages 35–40. IOS Press, 2010.

[Hustadt *et al.*, 2004] Ullrich Hustadt, Boris Motik, and Ulrike Sattler. A decomposition rule for decision procedures by resolution-based calculi. In *In: Proc. 11th Int. Conf. on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR*, pages 21–35. Springer, 2004.

[Hustadt *et al.*, 2007] Ullrich Hustadt, Boris Motik, and Ulrike Sattler. Reasoning in description logics by a reduction to disjunctive datalog. *J. Autom. Reasoning*, 39(3):351–384, 2007.

[Kazakov, 2009] Yevgeny Kazakov. Consequence-driven reasoning for Horn SHIQ ontologies. In Craig Boutilier, editor, *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, pages 2040–2045, 2009.

[Krötzsch *et al.*, 2007] Markus Krötzsch, Sebastian Rudolph, and Pascal Hitzler. Complexity boundaries for Horn description logics. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI'07)*, pages 452–457. AAAI Press, 2007.

[Levy and Rousset, 1998] Alon Y. Levy and Marie-Christine Rousset. Combining Horn rules and description logics in CARIN. *Artif. Intell.*, 104:165–209, September 1998.

[Lutz *et al.*, 2009] C. Lutz, D. Toman, and F. Wolter. Conjunctive query answering in the description logic el using a relational database system. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence IJCAI09*. AAAI Press, 2009.

[Maier and Mendelzon, 1979] David Maier and Albert Mendelzon. Testing implications of data dependencies. *ACM Transactions on Database Systems*, 4:455–469, 1979.

[Ortiz *et al.*, 2010] Magdalena Ortiz, Sebastian Rudolph, and Mantas Simkus. Worst-case optimal reasoning for the Horn-DL fragments of OWL 1 and 2. In Fangzhen Lin, Ulrike Sattler, and Miroslaw Truszczynski, editors, *KR*. AAAI Press, 2010.

[Ortiz *et al.*, 2011] Magdalena Ortiz, Sebastian Rudolph, and Mantas Simkus. Query answering in the Horn fragments of the description logics SHOIQ and SROIQ. In Toby Walsh, editor, *IJCAI*, pages 1039–1044. IJCAI/AAAI, 2011.

[Pérez-Urbina *et al.*, 2009] Héctor Pérez-Urbina, Boris Motik, and Ian Horrocks. A comparison of query rewriting techniques for DL-Lite. In Bernardo Cuenca Grau, Ian Horrocks, Boris Motik, and Ulrike Sattler, editors, *Description Logics*, volume 477 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2009.

[Pérez-Urbina *et al.*, 2010] H. Pérez-Urbina, B. Motik, and I. Horrocks. Tractable query answering and rewriting under description logic constraints. *J. Applied Logic*, 8(2):186–209, 2010.

[Rosati and Almatelli, 2010] Riccardo Rosati and Alessandro Almatelli. Improving query answering over DL-Lite ontologies. In *Proceedings of the Twelfth International Conference on Principles of Knowledge Representation and Reasoning (KR 2010)*, 2010.

[Rosati, 2006] Riccardo Rosati. DL+log: Tight integration of description logics and disjunctive datalog. In *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning (KR 2006)*, pages 68–78, 2006.

[Rosati, 2007] Riccardo Rosati. On conjunctive query answering in EL. In *Proceedings of the 2007 International Workshop on Description Logic (DL 2007)*. CEUR Electronic Workshop Proceedings, 2007.

[Sattler, 2000] Ulrike Sattler. Description logics for the representation of aggregated objects. In *Proceedings of the 14th European Conference on Artificial Intelligence. IOS*, pages 239–243. Press, 2000.