# The Combined Approach to Query Answering in Horn-$\mathcal{ALCHOIQ}$

**David Carral** and **Irina Dragoste** and **Markus Krötzsch**
Center for Advancing Electronics Dresden (cfaed), TU Dresden, Germany
firstname.lastname@tu-dresden.de

## Abstract

Combined approaches have become a successful technique for solving conjunctive query (CQ) answering over description logics (DL) ontologies. Nevertheless, existing approaches are restricted to tractable DL languages. In this work, we extend the combined method to the more expressive DL Horn-$\mathcal{ALCHOIQ}$—a language for which CQ answering is ExpTime-complete—in order to develop an efficient and scalable CQ answering procedure which is worst-case optimal for Horn-$\mathcal{ALCHOIQ}$ and $\mathcal{ELHO}$ ontologies. We implement and study the feasibility of our algorithm, and compare its performance to the DL reasoner Konclude.

## Introduction

Answering conjunctive queries (CQs) over Description Logics (DL) ontologies is an important reasoning task with many applications in knowledge representation and data management. Intensive research efforts in recent years have significantly improved our understanding of this problem, and led to efficient algorithms and implementations for many DL languages (Calvanese et al. 2007; Bienvenu et al. 2016). *Query rewriting* was an important step towards widespread practical implementation in legacy databases, but it is limited to DLs of sub-polynomial data complexity. This limitation was overcome by the so-called *combined approach*, which answers CQs in two steps:

1. *Materialisation*: data is augmented to build a query-independent interpretation, which may not be a model but is complete for CQ answering.
2. *Filtration*: queries are evaluated over this interpretation and unsound answers are discarded in a filtration step.

This approach has made CQ answering feasible for many further DLs (Kontchakov et al. 2010; 2011; Lutz et al. 2013; Stefanoni, Motik, and Horrocks 2013; Feier et al. 2015).

However, for many expressive DL languages, the problem remains challenging in theory and in practice. All previously cited works on query rewriting and combined approaches for DL are restricted to tractable languages. For $\mathcal{SROIQ}$—the DL underpinning the OWL Web Ontology Language—, it is unknown if the problem is decidable at all (Rudolph and Glimm 2010). Concrete complexity bounds are known for fragments of this logic, e.g., for Horn-$\mathcal{SROIQ}$ (Ortiz, Rudolph, and Simkus 2010), but these results did not give rise to practical implementations yet. Indeed, common methods for showing decidability often lead to algorithms that always run in worst-case complexity.

We address this limitation by proposing a new combined approach to CQ answering over ontologies of the DL Horn-$\mathcal{ALCHOIQ}$ (Krötzsch, Rudolph, and Hitzler 2013), for which CQ answering is ExpTime-complete (Ortiz, Rudolph, and Simkus 2011). Our procedure generalises previous works on lightweight, tractable DLs (Kontchakov et al. 2010; 2011; Lutz et al. 2013; Stefanoni, Motik, and Horrocks 2013; Feier et al. 2015) to this expressive language, but at the same time exhibits worst-case optimal, pay-as-you-go behaviour. In particular, our algorithm runs in exponential time for Horn-$\mathcal{ALCHOIQ}$ and in non-deterministic polynomial time for $\mathcal{ELHO}$ (Baader, Brandt, and Lutz 2005).

The pay-as-you-go behaviour is embodied in our materialisation step, which extends ideas on consequence-based reasoning (Kazakov 2009; Simančík, Kazakov, and Horrocks 2011; Bate et al. 2016) to a DL that combines nominals, at-most quantifiers, and inverse roles. The filtration step then adapts a technique of Feier et al. (2015), which is comparatively lightweight.

To show practical feasibility, we have implemented the materialisation step of our approach using a Datalog reasoner as a blackbox system. Even without filtration, this suffices to solve standard DL reasoning tasks such as satisfiability and assertion retrieval, and we evaluate performance on expressive real-world ontologies. For these ontologies, our implementation requires only moderately sized Datalog programs, and can often outperform *Konclude* (Steigmiller, Liebig, and Glimm 2014)—considered the leading DL reasoner (Parsia et al. 2017). We interpret this as an indication that pay-as-you-go behaviour does indeed occur in practice. A complete CQ answering implementation based on our approach therefore seems feasible.

In summary, our main contributions are:

- We present the first combined approach applicable to a non-tractable DL fragment, generalising combined approaches by Kontchakov et al. (2010; 2011), Lutz et al. (2013), Stefanoni et al. (2013), and Feier et al. (2015) to significantly more expressive logics.

| | | |
|---|---|---|
| (1) | $\bigsqcap_{i=1}^{n} C_i \sqsubseteq D$ | $\bigcap_{i=1}^{n} C_i^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ |
| (2) | $C \sqsubseteq \exists R.D$ | $C^{\mathcal{I}} \subseteq \{\delta \mid \exists \gamma \in D^{\mathcal{I}}.\langle \delta, \gamma \rangle \in R^{\mathcal{I}}\}$ |
| (3) | $\exists R.C \sqsubseteq D$ | $\{\delta \mid \exists \gamma \in C^{\mathcal{I}}.\langle \delta, \gamma \rangle \in R^{\mathcal{I}}\} \subseteq D^{\mathcal{I}}$ |
| (4) | $C \sqsubseteq \leqslant 1\, R.D$ | $C^{\mathcal{I}} \subseteq \{\epsilon \mid \forall \delta, \gamma \in D^{\mathcal{I}}.(\langle \epsilon, \delta \rangle \in R^{\mathcal{I}} \wedge \langle \epsilon, \gamma \rangle \in R^{\mathcal{I}}) \rightarrow \delta = \gamma\}$ |
| (5) | $C \sqsubseteq \{a\}$ | $C^{\mathcal{I}} \subseteq \{a^{\mathcal{I}}\}$ |
| (6) | $R \sqsubseteq S$ | $R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$ |
| (7) | $C(a)$ | $a^{\mathcal{I}} \in C^{\mathcal{I}}$ |
| (8) | $R(a,b)$ | $\langle a^{\mathcal{I}}, b^{\mathcal{I}} \rangle \in R^{\mathcal{I}}$ |

Figure 1: Syntax (left) and semantics (right) of Horn-$\mathcal{ALCHOIQ}$ axioms in normal form, where $C_{(i)}, D \in \mathbf{C}$, $R, S \in \mathbf{R}$, and $a, b \in \mathbf{I}$

- We show that our approach is worst-case optimal for standard reasoning and CQ answering for the DLs Horn-$\mathcal{ALCHOIQ}$ and $\mathcal{ELHO}$, in terms of both data and combined complexity.
- We develop an efficient implementation to solve standard reasoning tasks over Horn-$\mathcal{ALCHOIQ}$ ontologies.
- We conduct an empirical evaluation with four data intensive ontologies (two real-world and two benchmarks) that shows performance gains over the DL reasoner Konclude.

This is a report including an appendix with additional details on proofs omitted from the conference version.

## Preliminaries

We consider logical theories based on *finite* signatures consisting of mutually disjoint sets $\mathbf{C}$ of *concepts* (unary predicates), $\mathbf{R}$ of *roles* (binary predicates), and $\mathbf{I}$ of *individuals* (constants). We require $\bot, \top \in \mathbf{C}$.

**Description Logics** Additionally, a Horn-$\mathcal{ALCHOIQ}$ signature has inverse roles: there is a bijective and irreflexive function $\cdot^{-} : \mathbf{R} \rightarrow \mathbf{R}$ with $R^{--} = R$ for all $R \in \mathbf{R}$. Without loss of generality (Krötzsch, Rudolph, and Hitzler 2013), we define Horn-$\mathcal{ALCHOIQ}$ using a restricted set of normalised axioms over some signature, which we introduce on the left hand side of Figure 1. Axioms of the form (7) or (8) are also referred to as *class* and *role assertions*, respectively, or simply *assertions*.

A *(Horn-$\mathcal{ALCHOIQ}$) ontology* is a finite set of axioms. For an ontology $\mathcal{O}$, let $\sqsubseteq_{\mathcal{O}}^{*}$ be the minimal transitive, reflexive relation defined over roles such that $R \sqsubseteq_{\mathcal{O}}^{*} S$ and $R^{-} \sqsubseteq_{\mathcal{O}}^{*} S^{-}$ for all $R \sqsubseteq S \in \mathcal{O}$.

We define the semantics of ontologies using *interpretations*. An *interpretation* $\mathcal{I}$ is a pair $\langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ with $\Delta^{\mathcal{I}}$ a set of *domain elements*, and $\cdot^{\mathcal{I}}$ a function mapping individuals to domain elements, concepts to subsets of $\Delta^{\mathcal{I}}$, and roles to binary relations over $\Delta^{\mathcal{I}}$, such that $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$, $\bot^{\mathcal{I}} = \emptyset$, and $(R^{-})^{\mathcal{I}} = (R^{\mathcal{I}})^{-}$ for all $R \in \mathbf{R}$. We extend the mapping defined by an interpretation $\mathcal{I}$ to concept/role conjunctions, and nominals in the standard way. That is, $(\sqcap_{i=1}^{n} C_i)^{\mathcal{I}} = \cap_{i=1}^{n} C_i^{\mathcal{I}}$, $(\sqcap_{i=1}^{n} R_i)^{\mathcal{I}} = \cap_{i=1}^{n} R_i^{\mathcal{I}}$, and $\{a\}^{\mathcal{I}} = \{a^{\mathcal{I}}\}$ with $C_i \in \mathbf{C}$ and $R_i \in \mathbf{R}$ for all $i \in \{1, \ldots, n\}$, and $a \in \mathbf{I}$.

An interpretation $\mathcal{I}$ *satisfies* (*entails*) an axiom $\alpha$, written $\mathcal{I} \models \alpha$, if the corresponding condition in Figure 1 holds. Interpretation $\mathcal{I}$ *satisfies* (is a *model* of) an ontology $\mathcal{O}$, written $\mathcal{I} \models \mathcal{O}$, if it satisfies all axioms in $\mathcal{O}$. If there is such an interpretation, we say that $\mathcal{O}$ is *satisfiable*. The ontology $\mathcal{O}$ *entails* an axiom $\alpha$, written $\mathcal{O} \models \alpha$, if $\mathcal{I} \models \alpha$ for all $\mathcal{I} \models \mathcal{O}$.

**Datalog** Consider a countably infinite set $\mathbf{V}$ of variables disjoint from $\mathbf{C}$, $\mathbf{R}$, and $\mathbf{I}$. The set of *terms* is $\mathbf{T} = \mathbf{V} \cup \mathbf{I}$. Lists of terms $t_1, \ldots, t_n$ are abbreviated as $\vec{t}$. An *atom* is a formula of the form $C(t)$ or $R(t, u)$ with $C \in \mathbf{C}$, $R \in \mathbf{R}$, and $t, u \in \mathbf{T}$. A *(Datalog) rule* is a formula of the form

$$B_1 \wedge \ldots \wedge B_n \rightarrow H_1 \wedge \ldots \wedge H_m,$$

with $n \geq 0$ *body atoms* $B_i$ and $m \geq 1$ *head atoms* $H_j$, and where each variable in the head must also occur in the body. A *fact* is a variable-free atom, i.e., a ground rule with $n = 0$ and $m = 1$. A *substitution* is a partial function $\sigma : \mathbf{V} \rightarrow \mathbf{I}$. For a formula $\varphi$, $\varphi\sigma$ is obtained from $\varphi$ by replacing all free variables $x$ in the domain of $\sigma$ with $\sigma(x)$.

We define the semantics of rules via the *chase* procedure.

**Definition 1.** *A* chase sequence *of a rule set $\mathcal{R}$ is a maximal sequence $\mathcal{F}^0, \ldots, \mathcal{F}^n$ of sets of facts, such that $\mathcal{F}^0 = \emptyset$, and for all $i \in \{1, \ldots, n\}$ there is a rule $\rho = \beta \rightarrow \eta \in \mathcal{R}$ and a substitution $\sigma$ with (1) $\beta\sigma \subseteq \mathcal{F}^{i-1}$, (2) $\eta\sigma \not\subseteq \mathcal{F}^{i-1}$, and (3) $\mathcal{F}^i = \mathcal{F}^{i-1} \cup \eta\sigma$. The* chase of $\mathcal{R}$*, denoted with $\mathcal{R}^{\infty}$, is the final element of any (arbitrarily chosen) chase sequence.*

The set $\mathcal{R}^{\infty}$ is unique for a rule set $\mathcal{R}$ irrespectively of the chosen chase sequence, and coincides with the least Herbrand model of $\mathcal{R}$. A set of facts $\mathcal{F}$ *entails* an assertion $\alpha$, written $\mathcal{F} \models \alpha$, if $\alpha \in \mathcal{F}$. The set $\mathcal{R}$ *entails* an assertion $\alpha$, written $\mathcal{R} \models \alpha$, if $\mathcal{R}_{\mathcal{O}}^{\infty} \models \alpha$.

**Conjunctive Queries** A *conjunctive query* (CQ) is a formula of the form $\exists \vec{x}.\beta$ where $\beta$ (the *body*) is a conjunction of atoms. A *Boolean CQ* (BCQ) is a CQ where all the variables are existentially quantified. Without loss of generality, we restrict ourselves to the task of solving BCQ entailment and consider only queries that do not contain individuals.

A *(variable) assignment* $Z$ for an interpretation $\mathcal{I}$ is a function $Z : \mathbf{V} \rightarrow \Delta^{\mathcal{I}}$. An interpretation $\mathcal{I}$ *entails* a BCQ $q = \exists \vec{x}.\beta$, written $\mathcal{I} \models q$, if there is an assignment $Z$ for $\mathcal{I}$ with $\langle Z(\vec{y}) \rangle \in P^{\mathcal{I}}$ for all $P(\vec{y}) \in \beta$. An ontology $\mathcal{O}$ *entails* a BCQ $q$, written $\mathcal{O} \models q$, if $\mathcal{I} \models q$ for all $\mathcal{I} \models \mathcal{O}$. Similarly, a set of facts $\mathcal{F}$ *entails* $q$, written $\mathcal{F} \models q$, if there is a substitution $\sigma$ with $\beta\sigma \subseteq \mathcal{F}$. A rule set $\mathcal{R}$ *entails* $q$, written $\mathcal{R} \models q$, if the chase of $\mathcal{R}$ entails $q$. Note that, in all of the above cases, our definition of entailment coincides with that of first-order logic.

## Materialisation Phase

We now present the materialisation step of our combined approach, which leads to a query-independent set of facts that can be exploited to solve BCQ entailment. Moreover,

$$\mathcal{R}_{\mathsf{Top}} = \{C(x) \to \top(x) \mid C \in \mathbf{C}\} \cup \{\mathbb{R}(x,y) \to \top(x) \wedge \top(y) \mid \mathbb{R} \in \mathbf{R}^{\sqcap}\}$$

$$\mathcal{R}_{\mathsf{Role}} = \{\mathbb{R}(x,y) \wedge \mathsf{N}(y) \to \mathbb{R}^{-}(y,x) \mid \mathbb{R} \in \mathbf{R}^{\sqcap}\} \cup \{\mathbb{R}(x,y) \to R(x,y) \mid \mathbb{R} \in \mathbf{R}^{\sqcap}, R \in \mathbb{R}\}$$

$$\mathcal{R}_{\mathsf{Nm}} = \{\mathsf{N}(a), \top(a) \mid a \in \mathbf{I}\}$$

$$\begin{aligned}\mathcal{R}_{\mathsf{Eq}} = \{&x \approx y \to y \approx x, x \approx y \wedge y \approx z \to x \approx z\} \cup \\ \{&C(x) \wedge x \approx y \to C(y) \mid C \in \mathbf{C}^{+}\} \cup \\ \{&\mathbb{R}(x,y) \wedge x \approx z \to \mathbb{R}(z,y) \mid \mathbb{R} \in \mathbf{R}^{\sqcap}\} \cup \\ \{&\mathbb{R}(x,y) \wedge y \approx z \to \mathbb{R}(x,z) \mid \mathbb{R} \in \mathbf{R}^{\sqcap}\}\end{aligned}$$

Figure 2: Auxiliary rule sets $\mathcal{R}_{\mathsf{Top}}$, $\mathcal{R}_{\mathsf{Eq}}$, $\mathcal{R}_{\mathsf{Role}}$, and $\mathcal{R}_{\mathsf{Nm}}$

we show that this set of facts can be directly used to decide satisfiability and assertion entailment.

We obtain this set of facts as the chase of a rule set $\mathcal{R}_{\mathcal{O}}$, which is defined over an extended signature (for the remainder of the paper, let $\mathcal{O}$ be a fixed ontology defined over some signature $\langle \mathbf{C}, \mathbf{R}, \mathbf{I} \rangle$). We let $\mathbf{C}^{\sqcap} = \{\sqcap_{i=1}^{n} C_i \mid C_1, \ldots, C_n \in \mathbf{C}, n \geq 1\}$. In the following, $\mathbb{C}$ and $\mathbb{D}$ are always used to denote elements of $\mathbf{C}^{\sqcap}$. Likewise, we consider role conjunctions $\mathbf{R}^{\sqcap} = \{\sqcap_{i=1}^{n} R_i \mid R_1, \ldots, R_n \in \mathbf{R}, n \geq 1\}$, denoted by $\mathbb{R}$ and $\mathbb{S}$. We tacitly identify elements of $\mathbf{C}^{\sqcap}$ and $\mathbf{R}^{\sqcap}$ with the sets of the concepts and roles that they contain, and use expressions such as $R \in \mathbb{R}$ and $\mathbb{C} \subseteq \mathbb{D}$ with this intention. For roles, we further define $\mathbb{R}^{-} = \sqcap_{R \in \mathbb{R}} R^{-}$.

The signature of the rule set $\mathcal{R}_{\mathcal{O}}$ is $\langle \mathbf{C}^{+}, \mathbf{R}^{+}, \mathbf{I}^{+} \rangle$, where $\mathbf{C}^{+} = \mathbf{C} \uplus \{\mathsf{N}\}$ with $\mathsf{N}$ a unary predicate, $\mathbf{R}^{+} = \mathbf{R}^{\sqcap} \uplus \{\approx\}$ with $\approx$ a binary predicate, and $\mathbf{I}^{+} = \mathbf{I} \uplus \{t_{\mathbb{C}} \mid \mathbb{C} \in \mathbf{C}^{\sqcap}\}$ with new constants of form $t_{\mathbb{C}}$. We assume $t_{\mathbb{C}} = t_{\mathbb{D}}$ if $\mathbb{C} = \mathbb{D}$ when these are considered as sets. Note that, $\mathbf{C} \subseteq \mathbf{C}^{\sqcap}$ and $\mathbf{R} \subseteq \mathbf{R}^{\sqcap}$ by definition.

We use the (finite) chase $\mathcal{R}_{\mathcal{O}}^{\infty}$ to represent (potentially infinite) models $\mathcal{I}$ of $\mathcal{O}$, where constants $a \in \mathbf{I}$ represent the named elements of the domain $\Delta^{\mathcal{I}}$ and the constants $t_{\mathbb{C}}$ represent (possibly many) anonymous domain elements in $\mathbb{C}^{\mathcal{I}}$. The special predicate $\mathsf{N}$ classifies representatives that behave like named individuals, i.e., the derivation of $\mathsf{N}(t_{\mathbb{C}})$ during the computation of the chase implies that $t_{\mathbb{C}}$ represents the unique element in $\mathbb{C}^{\mathcal{I}}$ for any model $\mathcal{I}$. Note that this may be the case if, e.g., $C \sqsubseteq \{a\} \in \mathcal{O}$ for some $C \in \mathbb{C}$.

The occurrence of fact $C(a)$ during the computation of the chase indicates that all elements represented by $a$ are in $C^{\mathcal{I}}$ for all $\mathcal{I} \models \mathcal{O}$. The occurrence of a fact $\mathbb{R}(a,b)$ indicates that all elements represented by $a$ are connected to some element represented by the second constant by all of the roles $R \in \mathbb{R}$. It is important to distinguish such joint connections that exist in the model from incidental co-occurrences that are an artefact of the re-use of representatives $t_{\mathbb{C}}$. The special predicate $\approx$ represents equality, which we model explicitly only between constants in $\mathsf{N}$. A fact of the form $a \approx b$ indicates that the classes of elements represented by the constants $a$ and $b$ are indeed the same in all models. The intuitions discussed in the previous paragraphs are formally introduced by the claims in the proof of Lemma 1.

(1) $\bigwedge_{i=1}^{n} C_i(x) \to D(x)$

(2) $C(x) \to R(x, t_D) \wedge D(t_D)$

(3) For all $\mathbb{R} \in \mathbf{R}^{\sqcap}$ with $R \in \mathbb{R}$, and all $\mathbb{C} \in \mathbf{C}^{\sqcap}$:

(3.1) $R(x,y) \wedge C(y) \to D(x)$

(3.2) $C(x) \wedge \mathbb{R}^{-}(x, t_{\mathbb{C}}) \to \mathbb{R}^{-}(x, t_{\mathbb{C} \sqcap D}) \wedge \bigwedge_{X \in \mathbb{C} \sqcap D} X(t_{\mathbb{C} \sqcap D})$

(5) $C(x) \to a \approx x$

(6) For $\mathbb{S} = \{T \mid R \sqsubseteq_{\mathcal{O}}^{*} T\}$:

(6.1) $R(x,y) \to \mathbb{S}(x,y)$

(6.2) $R^{-}(x,y) \to \mathbb{S}^{-}(x,y)$

(7) $C(a)$

(8) $R(a,b)$

Figure 3: Rule sets for axiom types (1)–(3) and (5)–(8) of Figure 1 (see Figure 4 for type (4))

The set of rules $\mathcal{R}_{\mathcal{O}}$ is defined as a combination of the auxiliary rules in Figure 2, and the axiom-specific rules in Figures 3 and 4, both of which we explain below.

**Definition 2.** *For each axiom $\alpha$ of one of the types introduced in Figure 1, let $\mathcal{R}_{\alpha}$ denote the corresponding rule set defined in Figures 3 and 4. For an ontology $\mathcal{O}$, the rule set $\mathcal{R}_{\mathcal{O}}$ is defined as:*

$$\mathcal{R}_{\mathcal{O}} = \bigcup\nolimits_{\alpha \in \mathcal{O}} \mathcal{R}_{\alpha} \cup \mathcal{R}_{\mathit{Top}} \cup \mathcal{R}_{\mathit{Role}} \cup \mathcal{R}_{\mathit{Nm}} \cup \mathcal{R}_{\mathit{Eq}}$$

Since we consider a finite signature, $\mathcal{R}_{\mathcal{O}}$ is finite, but exponential (due to the exponentially many roles $\mathbb{R} \in \mathbf{R}^{\sqcap}$ and constants $t_{\mathbb{C}} \in \mathbf{I}^{+}$). The rules of Figure 2 axiomatise the intended semantics of $\top$, role conjunction, $\mathsf{N}$, and $\approx$. The first part of $\mathcal{R}_{\mathsf{Role}}$ expresses the semantics of inverse roles, which are part of the signature, but do not have a built-in semantics in Datalog. We only invert roles that connect to constants in $\mathsf{N}$. The second part of $\mathcal{R}_{\mathsf{Role}}$ recovers individual roles from role conjunctions. $\mathcal{R}_{\mathsf{Nm}}$ comprises basic facts for the named individuals. Finally, $\mathcal{R}_{\mathsf{Eq}}$ are a standard equality theory, which could be omitted if $\approx$ is defined with a special semantics, as in some Datalog engines.

In Figure 3, rules (1), (2), (3.2), (5), (7), and (8) are basically direct translations of the corresponding DL axioms into first-order implications. In (6.1) and (6.2), we use a role conjunction $\mathbb{S}$ that represents the upward closure of $R$ in the role hierarchy. The necessary reflexive transitive closure can be computed in polynomial time. Note that rules (3) and (4) are instantiated for any $\mathbb{R}$ (and $\mathbb{S}$) that contain $R$.

Rule (3.2) applies domain axioms along inverse relations that lead to representative $t_{\mathbb{C}}$ by initialising a new individual $t_{\mathbb{C} \sqcap D}$ to which any other information (roles in $\mathbb{R}^{-}$ and arbitrary concepts $X$) are copied. This inverse case is not needed for individuals in $\mathbf{I}$, since they can be treated with (3.1) after flipping the direction of the inverse predicate using $\mathcal{R}_{\mathsf{Role}}$. Rules (4) handle functional roles $R$ as follows: (4.1) is similar to the usual first-order translation of functionality, written

(4) For all $\mathbb{R}, \mathbb{S} \in \mathbf{R}^{\sqcap}$ with $R \in \mathbb{R}$ and $R \in \mathbb{S}$, and all $\mathbb{C}, \mathbb{D} \in \mathbf{C}^{\sqcap}$:

(4.1) $D(y) \wedge R^{-}(y,x) \wedge C(x) \wedge R(x,z) \wedge D(z) \wedge \mathsf{N}(z) \to y \approx z$

(4.2) $C(x) \wedge \mathbb{R}(x, t_{\mathbb{C}}) \wedge D(t_{\mathbb{C}}) \wedge \mathbb{S}(x, t_{\mathbb{D}}) \wedge D(t_{\mathbb{D}}) \to (\mathbb{R} \sqcap \mathbb{S})(x, t_{\mathbb{C} \sqcap \mathbb{D}}) \wedge \bigwedge_{X \in \mathbb{C} \sqcap \mathbb{D}} X(t_{\mathbb{C} \sqcap \mathbb{D}})$

(4.3) $D(y) \wedge \mathbb{R}^{-}(y,x) \wedge C(x) \wedge \mathbb{S}(x, t_{\mathbb{C}}) \wedge D(t_{\mathbb{C}}) \to \bigwedge_{X \in \mathbb{C}} X(y) \wedge (\mathbb{R}^{-} \sqcap \mathbb{S}^{-})(y,x)$

(4.4) $D(y) \wedge R^{-}(y,x) \wedge C(x) \wedge \mathsf{N}(x) \to \mathsf{N}(y)$

Figure 4: Rule sets for axiom type (4) of Figure 1

using one inverted occurrence of $R$ and restricted to cases where the target individual of $R$ is in class $\mathsf{N}$; (4.2) merges two anonymous individuals into a new anonymous individual; (4.3) folds the properties of an anonymous individual back into its grandparent; and (4.4) propagates the property of being named along inverse functional roles.

**Example 1.** *Consider the following ontology:*

| | | | |
|---|---|---|---|
| $A \sqsubseteq \exists R.B$ | (1) | $B \sqsubseteq \exists S.C$ | (5) |
| $B \sqsubseteq \exists V.D$ | (2) | $C \sqsubseteq \{c\}$ | (6) |
| $\top \sqsubseteq \leqslant 1\, S.\top$ | (3) | $A(a)$ | (7) |
| $V \sqsubseteq S$ | (4) | $V(b,c)$ | (8) |

*Among other things, these axioms entail $D(c)$. Indeed, we can derive this fact with the following chase sequence. For each inference, we give the applied rule before the colon, with subscripts indicating the ontology axiom it originated from, and the facts that it was applied to after the colon.*

| | | |
|---|---|---|
| $R(a,t_B), B(t_B)$ | $(2)_{(1)} : (7)$ | (9) |
| $V(t_B,t_D), D(t_D)$ | $(2)_{(2)} : (9)$ | (10) |
| $S(t_B,t_C), C(t_C)$ | $(2)_{(5)} : (9)$ | (11) |
| $(V \sqcap S)(t_B,t_D)$ | $(6.1)_{(4)} : (10)$ | (12) |
| $(V \sqcap S)(t_B,t_{C\sqcap D})$, $C(t_{C\sqcap D}), D(t_{C\sqcap D})$ | $(4.2)_{(3)} : (11), (12)$ | (13) |
| $V(t_B,t_{C\sqcap D})$ | $\mathcal{R}_{\mathit{Role}} : (13)$ | (14) |
| $c \approx t_{C\sqcap D}$ | $(5)_{(6)} : (13)$ | (15) |
| $t_{C\sqcap D} \approx c$ | $\mathcal{R}_{\mathit{Eq}} : (15)$ | (16) |
| $D(c)$ | $\mathcal{R}_{\mathit{Eq}} : (16)$ | (17) |

*Additional inferences are possible, but one cannot obtain all inferences that one might expect in DL. For example, $R^{-}(t_B,a)$ is not entailed, since we do not have $\mathsf{N}(t_B)$. However, if we add to the ontology an additional axiom*

$$\top \sqsubseteq \leqslant 1\, V^{-}.\top \tag{18}$$

*then we can further compute $B(b)$ as follows:*

| | | |
|---|---|---|
| $V^{-}(c,b)$ | $\mathcal{R}_{\mathit{Role}} : (8), \mathcal{R}_{\mathit{Nm}}$ | (19) |
| $V^{-}(t_{C\sqcap D},b)$ | $\mathcal{R}_{\mathit{Eq}} : (19), (15)$ | (20) |
| $t_B \approx b$ | $(4.1)_{(18)} : (14), (20), \mathcal{R}_{\mathit{Nm}}$ | (21) |
| $B(b)$ | $\mathcal{R}_{\mathit{Eq}} : (9), (21)$ | (22) |

*In this case, all expected conclusions can be obtained since all auxiliary individuals can be inferred to be in $\mathsf{N}$.*

The rules of $\mathcal{R}_{\mathcal{O}}$ entail enough relevant inferences to be used to decide standard reasoning tasks over Horn-$\mathcal{ALCHOIQ}$ ontologies.

**Theorem 1.** *$\mathcal{O}$ is satisfiable iff $\mathcal{R}_{\mathcal{O}} \not\models \exists x.\bot(x)$. If $\mathcal{O}$ is satisfiable, then it entails an assertion $\alpha$ iff $\mathcal{R}_{\mathcal{O}} \models \alpha$.*

Theorem 1 follows from Lemmas 1, 3, and 4, shown in the next section. Before this, we first discuss the complexity of our approach. The rule set $\mathcal{R}_{\mathcal{O}}$ is exponential in the size of $\mathcal{O}$, due to the exponential number of roles $\mathbb{R} \in \mathbf{R}^{+}$ and of individuals of the form $t_{\mathbb{C}}$ in $\mathbf{I}^{+}$. However, the Datalog rules in $\mathcal{R}_{\mathcal{O}}$ contain at most three variables, making their propositional logic grounding polynomial in the size of $\mathcal{R}_{\mathcal{O}}$. Since propositional Horn logic entailment can be decided in polynomial time, this already yields a worst-case optimal ExpTime algorithm for Horn-$\mathcal{ALCHOIQ}$.

We can also achieve worst-case optimal reasoning for simpler DL fragments. A practically relevant case is $\mathcal{ELHO}$, the fragment Horn-$\mathcal{ALCHOIQ}$ that does not contain "at most" quantifiers or inverse roles:

**Definition 3.** *An ontology $\mathcal{O}$ is $\mathcal{ELHO}$ if (i) $\mathcal{O}$ does not contain axioms of the form (4) and, (ii) for every role $R \in \mathbf{R}$, $\mathcal{O}$ contains only one of $R$ and $R^{-}$, but not both.*

Without "at most" restrictions, we can omit all rules of type (4). Without inverses, we can further discard rules (3.2) and (6.2). The remaining rules of Figure 3 only introduce facts about constants of the form $t_C$ with $C$ a single concept, and about roles $\mathbb{R}$ where $\mathbb{R}$ is either a single role or the $\sqsubseteq^{*}_{\mathcal{O}}$ closure of such a role as in (6). Therefore, only polynomially many signature symbols are required, and we can restrict the rules in Figure 2 to these symbols as well. The resulting polynomial set of Datalog rules with at most three variables can be evaluated in polynomial time.

We can sum up our results as follows. Recall that *data complexity* is measured with respect to the number of (normalised) assertions in an ontology, while *combined complexity* refers to the ontology (and its signature) as a whole.

**Theorem 2.** *The approach of Theorem 1 decides consistency and assertion entailment for Horn-$\mathcal{ALCHOIQ}$ in polynomial time for data complexity, and in exponential time for combined complexity. When restricting to rules in the relevant signature for $\mathcal{ELHO}$, the algorithm runs in polynomial time for combined and data complexity.*

In practice, the better algorithm for $\mathcal{ELHO}$ can be used as a basis for a pay-as-you-go algorithm for Horn-$\mathcal{ALCHOIQ}$, which adds rules on demand during materialisation. In this optimised procedure, we start with only those rules whose premises use only the signature of the given ontology (but whose conclusions may use further symbols). Whenever a newly derived fact uses additional signature symbols, rules with premises that also use this symbol are added. Adding rules during materialisation is not difficult in Datalog, and is supported by existing engines.

## Correctness of the Materialisation

We now establish the correctness of our approach, where we give proof sketches that show some of the most important cases. The full case analysis is found in the appendix.

The next lemma establishes soundness. This is not an obvious property, since our approach represents potentially infinite models with a finite materialisation that re-uses the same constant symbols as representatives for many domain elements. This could result in unsound inferences, as indeed it happens for BCQ entailment, where an additional filtration step is thus needed. For computing assertions, however, the method is sound:

**Lemma 1.** *If $\mathcal{R}_\mathcal{O}$ entails an assertion $\alpha$, then $\mathcal{O} \models \alpha$.*

*Proof sketch.* For all $a \in \mathbf{I}^+$, let $\mathsf{Ex}(a) = \{a\}$ if $a \in \mathbf{I}$, and $\mathsf{Ex}(a) = \mathbb{C}$ if $a$ is of the form $t_\mathbb{C}$.

Let $\mathcal{F}^0, \ldots, \mathcal{F}^n$ be some fixed chase sequence of $\mathcal{R}_\mathcal{O}$. Then, we show via induction on $i \in \{0, \ldots, n\}$ that the following claims hold for any model $\mathcal{I}$ of $\mathcal{O}$:

(a) If $C(a) \in \mathcal{F}^i$ with $C \in \mathbf{C}$, then $\mathsf{Ex}(a)^\mathcal{I} \subseteq C^\mathcal{I}$.

(b) If $\mathbb{R}(a, b) \in \mathcal{F}^i$, then for all $\delta \in \mathsf{Ex}(a)^\mathcal{I}$ there is some $\gamma \in \mathsf{Ex}(b)^\mathcal{I}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^\mathcal{I}$.

(c) If $a \approx b \in \mathcal{F}^i$, then $\mathsf{Ex}(a)^\mathcal{I} = \mathsf{Ex}(b)^\mathcal{I}$.

(d) If $a \in \mathbf{I}^+$ occurs in $\mathcal{F}^i$, then $\mathsf{Ex}(a)^\mathcal{I} \neq \emptyset$.

(e) If $\mathsf{N}(a) \in \mathcal{F}^i$, then $|\mathsf{Ex}(a)^\mathcal{I}| = 1$.

The lemma follows from (a) and (b); the rest of the claims are included to structure our induction argument.

For the base case $i = 0$, all the claims trivially hold, since $\mathcal{F}^0 = \emptyset$. For the induction step, consider $i \in \{1, \ldots, n\}$, and assume that all the claims hold for $i - 1$ (IH). We show that the claims remain true by distinguishing the following cases based on the type of the rule $\rho = \beta \to \eta \in \mathcal{R}_\mathcal{O}$ and the substitution $\sigma$ such that $\beta\sigma \subseteq \mathcal{F}^{i-1}$ and $\mathcal{F}^i = \mathcal{F}^{i-1} \cup \eta\sigma$. Note that each type of rule needs to establish only the claims whose premise might be affected by its application. In this proof sketch, we only consider five cases.

Let $\rho$ be of the form (1). Then, $\{C_j(\sigma(x)) \mid 1 \leq j \leq m\} \subseteq \mathcal{F}^{i-1}$ and hence, $\mathsf{Ex}(\sigma(x))^\mathcal{I} \subseteq \cap_{j=1}^m C_j^\mathcal{I}$ by (IH.a). Since $\sqcap_{j=1}^m C_j \sqsubseteq D \in \mathcal{O}$, $\mathsf{Ex}(\sigma(x))^\mathcal{I} \subseteq D^\mathcal{I}$ and (a) holds.

Let $\rho$ be of the form (3.2). Then, $C(\sigma(x)), \mathbb{R}^-(\sigma(x), t_\mathbb{C}) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R}$. By definition, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^\mathcal{I} = \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} \cap D^\mathcal{I}$. Hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^\mathcal{I} \subseteq \bigcap_{X \in \mathbb{C}} X^\mathcal{I} \cap D^\mathcal{I}$ and (a) holds. By (IH.a) and (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(x))^\mathcal{I} \subseteq C^\mathcal{I}$ there is some $\gamma \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I}$ with $\langle \delta, \gamma \rangle \in (\mathbb{R}^-)^\mathcal{I}$. Since $\exists R.C \sqsubseteq D \in \mathcal{O}$, $\gamma \in D^\mathcal{I}$ and (b) holds. By (IH.d), $\mathsf{Ex}(\sigma(x))^\mathcal{I} \neq \emptyset$ and hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^\mathcal{I} \neq \emptyset$ and (d) holds.

Let $\rho$ be of the form (4.3). Then, $D(\sigma(y))$, $\mathbb{R}^-(\sigma(y), \sigma(x)), C(\sigma(x)), \mathbb{S}(\sigma(x), t_\mathbb{C}), D(t_\mathbb{C}) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R}$ and $R \in \mathbb{S}$. By (IH.a) and (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(y))^\mathcal{I} \subseteq D^\mathcal{I}$, (1) there is some $\epsilon \in \mathsf{Ex}(\sigma(x))^\mathcal{I} \subseteq C^\mathcal{I}$ with $\langle \delta, \epsilon \rangle \in (\mathbb{R}^-)^\mathcal{I}$, and (2) there is some $\gamma \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} \subseteq D^\mathcal{I}$ with $\langle \epsilon, \gamma \rangle \in \mathbb{S}^\mathcal{I}$. Since $C \sqsubseteq\ \leqslant 1\, R.D \in \mathcal{O}$, $\delta \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} = \mathbb{C}^\mathcal{I}$ and (a) holds. Moreover, $\langle \delta, \epsilon \rangle \in (\mathbb{R}^- \sqcap \mathbb{S}^-)^\mathcal{I}$ and (b) holds.

Let $\rho$ be of the form (4.4). Then, $D(\sigma(y))$, $R^-(\sigma(y), \sigma(x)), C(\sigma(x)), \mathsf{N}(\sigma(x)) \in \mathcal{F}^{i-1}$. By (IH.a), (IH.b), (IH.d), and (IH.e), (1) there is a single element $\epsilon \in \mathsf{Ex}(\sigma(x))^\mathcal{I} \subseteq C^\mathcal{I}$, (2) $\langle \epsilon, \delta \rangle \in R^\mathcal{I}$ for all $\delta \in \mathsf{Ex}(\sigma(y))^\mathcal{I} \subseteq D^\mathcal{I}$, and (3) $\mathsf{Ex}(\sigma(y))^\mathcal{I} \neq \emptyset$. Since $C \sqsubseteq\ \leqslant 1\, R.D \in \mathcal{O}$, $|\mathsf{Ex}(\sigma(y))^\mathcal{I}| = 1$ and (e) holds.

Let $\rho$ be of the form (Eq.5). Then, $\mathbb{R}(\sigma(x), \sigma(y)), \sigma(y) \approx \sigma(z) \in \mathcal{F}^{i-1}$. By (IH.b) and (IH.c), for all $\delta \in \mathsf{Ex}(\sigma(x))^\mathcal{I}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^\mathcal{I} = \mathsf{Ex}(\sigma(z))^\mathcal{I}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^\mathcal{I}$. Hence, (b) holds. $\square$

To show completeness, we use the chase of $\mathcal{R}_\mathcal{O}$ to define a structure $\mathcal{U}_\mathcal{O}$, which is a universal model of $\mathcal{O}$—i.e., a model that can be homomorphically embedded into any other model of $\mathcal{O}$. We then show that, for any assertion $\alpha$, $\mathcal{R}_\mathcal{O} \not\models \alpha$ implies $\mathcal{U}_\mathcal{O} \not\models \alpha$. In turn, this implies $\mathcal{O} \not\models \alpha$ and our approach is indeed complete for assertion retrieval. We first define the domain $\Delta^{\mathcal{U}_\mathcal{O}}$ by "unravelling" $\mathcal{R}_\mathcal{O}^\infty$.

**Definition 4.** *The set $\Delta$ and the function $\mathsf{i} : \Delta \to \mathbf{I}^+$ are defined recursively:*

*(i) For all $a \in \mathbf{I}^+$, we have $[a] \in \Delta$ for the equivalence class $[a] = \{b \mid a \approx b \in \mathcal{R}_\mathcal{O}^\infty\} \cup \{a\}$. Let $\mathsf{i}([a])$ be an arbitrary but fixed representative $\mathsf{i}([a]) \in [a]$.*

*(ii) For all $\delta \in \Delta$, $\mathbb{R} \in \mathbf{R}^\sqcap$, and $\mathbb{C} \in \mathbf{C}^\sqcap$, we have $\delta.d_{\mathbb{R},\mathbb{C}} \in \Delta$. Let $\mathsf{i}(\delta.d_{\mathbb{R},\mathbb{C}}) = t_\mathbb{C}$.*

Note that, for an individual $a \in \mathbf{I}$, possibly $\mathsf{i}([a]) \neq a$.

**Definition 5.** *We recursively define the domain $\Delta^{\mathcal{U}_\mathcal{O}} \subseteq \Delta$:*

1. *For all $a \in \mathbf{I}^+$ with $\mathsf{N}(a) \in \mathcal{R}_\mathcal{O}^\infty$, we have $[a] \in \Delta^{\mathcal{U}_\mathcal{O}}$.*
2. *For all $\delta \in \Delta^{\mathcal{U}_\mathcal{O}}$, we have $\delta.d_{\mathbb{R},\mathbb{C}} \in \Delta^{\mathcal{U}_\mathcal{O}}$ if*
   - *2.a $\mathbb{R}(\mathsf{i}(\delta), t_\mathbb{C}) \in \mathcal{R}_\mathcal{O}^\infty$, and*
   - *2.b for all $\mathbb{S}(\mathsf{i}(\delta), t_\mathbb{D}) \in \mathcal{R}_\mathcal{O}^\infty$, $\mathbb{R} \cup \mathbb{C} \not\subseteq \mathbb{S} \cup \mathbb{D}$, and*
   - *2.c for all $a \in \mathbf{I}^+$ with $\mathsf{N}(a) \in \mathcal{R}_\mathcal{O}^\infty$; there is some $R \in \mathbb{R}$ with $R(\mathsf{i}(\delta), a) \notin \mathcal{R}_\mathcal{O}^\infty$, or some $C \in \mathbf{C}$ with $C(t_\mathbb{C}) \in \mathcal{R}_\mathcal{O}^\infty$ and $C(a) \notin \mathcal{R}_\mathcal{O}^\infty$, and*
   - *2.d if $\delta$ is of the form $\gamma.d_{\mathbb{S},\mathbb{D}}$; $\mathbb{R}^- \not\subseteq \mathbb{S}$, or there is some $C \in \mathbf{C}$ with $C(t_\mathbb{C}) \in \mathcal{R}_\mathcal{O}^\infty$ and $C(\mathsf{i}(\gamma)) \notin \mathcal{R}_\mathcal{O}^\infty$.*

Conditions (2b-2d) ensure that, in some cases, an element is not introduced in $\Delta^{\mathcal{U}_\mathcal{O}}$ if the corresponding existential restriction is already satisfied by another element. These restrictions become relevant when showing that axioms of the form (4) are satisfied by $\mathcal{U}_\mathcal{O}$.

**Definition 6.** *Let $\Delta^{\mathcal{U}_\mathcal{O}}$ be as in Definition 5. The interpretation function $\cdot^{\mathcal{U}_\mathcal{O}}$ is defined by setting:*

1. *For all $a \in \mathbf{I}$, we have $a^{\mathcal{U}_\mathcal{O}} = [a]$.*
2. *For all $C \in \mathbf{C}$ and $\delta \in \Delta^{\mathcal{U}_\mathcal{O}}$, we have $\delta \in C^{\mathcal{U}_\mathcal{O}}$ if $C(\mathsf{i}(\delta)) \in \mathcal{R}_\mathcal{O}^\infty$.*
3. *For all $R \in \mathbf{R}$ and $\delta, \gamma \in \Delta^{\mathcal{U}_\mathcal{O}}$, we have $\langle \delta, \gamma \rangle \in R^{\mathcal{U}_\mathcal{O}}$ if*
   - *3.a $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$, or*
   - *3.b $R^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)), \mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_\mathcal{O}^\infty$, or*
   - *3.c $\gamma$ is of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ with $R \in \mathbb{R}$, or*
   - *3.d $\delta$ is of the form $\gamma.d_{\mathbb{R}^-,\mathbb{C}}$ with $R \in \mathbb{R}$.*

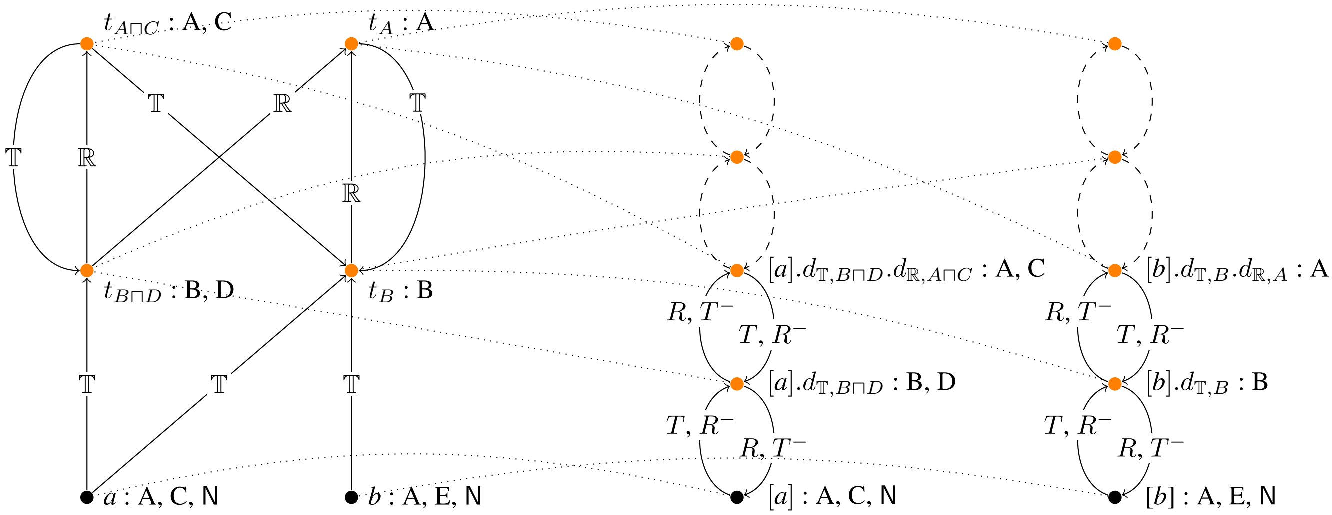


Figure 5: Representation of $\mathcal{R}_{\mathcal{O}}^{\infty}$ (left, inferences of rules $\mathcal{R}_{\mathsf{Top}}$ and $\mathcal{R}_{\mathsf{Roles}}$ not shown) and $\mathcal{U}_{\mathcal{O}}$ (right), where $\mathbb{T} = T \sqcap R^-$ and $\mathbb{R} = R \sqcap T^-$; dotted lines indicate how individuals in $\mathcal{U}_{\mathcal{O}}$ are represented in $\mathcal{R}_{\mathcal{O}}^{\infty}$

The relation between the chase of $\mathcal{R}_{\mathcal{O}}$ and the universal model $\mathcal{U}_{\mathcal{O}}$ is illustrated by the following example.

**Example 2.** *Let $\mathcal{O}$ be the ontology containing the axioms:*

$$
\begin{array}{cc}
A \sqsubseteq \exists T.B & B \sqsubseteq \exists R.A \\
\exists T^-.C \sqsubseteq D & \exists R^-.D \sqsubseteq C \\
T^- \sqsubseteq R & R^- \sqsubseteq T \\
A(a) \quad C(a) & A(b) \quad E(b)
\end{array}
$$

*Figure 5 illustrates the structure of $\mathcal{R}_{\mathcal{O}}^{\infty}$ (left) and $\mathcal{U}_{\mathcal{O}}$ (right). In the chase of $\mathcal{R}_{\mathcal{O}}$, only four constants of the form $t_{\mathbb{C}} \in \mathbf{I}^+$ are reused to "satisfy" all existential restrictions.*

*As previously discussed, a constant $t_{\mathbb{C}}$ intuitively represents all domain elements in the domain $\Delta^{\mathcal{I}}$ of some model $\mathcal{I}$ of $\mathcal{O}$ which belong to the interpretation of the intersection of all concepts in $\mathbb{C}$, i.e., all $\delta \in \Delta^{\mathcal{I}}$ with $\delta \in \cap_{C \in \mathbb{C}} C^{\mathcal{I}}$. Its properties are those shared by all individuals that are members of such concept interpretations in $\mathcal{I}$. The dotted arrows in Figure 5 link each element $\gamma \in \Delta^{\mathcal{U}_{\mathcal{O}}}$ to its corresponding representative $\mathsf{i}(\gamma)$ in the chase of $\mathcal{R}_{\mathcal{O}}$.*

**Lemma 2.** *If $\mathcal{R}_{\mathcal{O}} \not\models \exists x.\bot(x)$, then $\mathcal{U}_{\mathcal{O}}$ is an interpretation.*

*Proof.* The necessary conditions follow from Definition 6. We have $\bot^{\mathcal{U}_{\mathcal{O}}} = \emptyset$ due to (2) and the precondition of the lemma; $\top^{\mathcal{U}_{\mathcal{O}}} = \Delta^{\mathcal{U}_{\mathcal{O}}}$ due to $\mathcal{R}_{\mathsf{Top}} \cup \mathcal{R}_{\mathsf{Nm}} \subseteq \mathcal{R}_{\mathcal{O}}$ and (2); and $(R^-)^{\mathcal{U}_{\mathcal{O}}} = (R^{\mathcal{U}_{\mathcal{O}}})^-$ for all $R \in \mathbf{R}$ due to (3). $\square$

We strengthen the previous result and show that, if $\mathcal{O}$ is satisfiable, then $\mathcal{U}_{\mathcal{O}}$ is a model of $\mathcal{O}$.

**Lemma 3.** *The following are equivalent: (1) $\mathcal{O}$ is satisfiable, (2) $\mathcal{R}_{\mathcal{O}} \not\models \exists x.\bot(x)$, and (3) $\mathcal{U}_{\mathcal{O}} \models \mathcal{O}$.*

*Proof sketch.* $(1) \Rightarrow (2)$: If $\mathcal{R}_{\mathcal{O}} \models \exists x.\bot(x)$, then $\mathcal{O}$ does not admit any model by (a) and (d) from Lemma 1.

$(2) \Rightarrow (3)$: By (2) and Lemma 2, $\mathcal{U}_{\mathcal{O}}$ is an interpretation. It remains to show that $\mathcal{U}_{\mathcal{O}}$ satisfies every axiom $\alpha \in \mathcal{O}$, which is done by a lengthy analysis of cases included in the appendix.

$(3) \Rightarrow (1)$: By the definition of satisfiability. $\square$

**Lemma 4.** *Consider an assertion $\alpha$. If $\mathcal{R}_{\mathcal{O}} \not\models \exists x.\bot(x)$ and $\mathcal{R}_{\mathcal{O}} \not\models \alpha$, then $\mathcal{O} \not\models \alpha$.*

*Proof.* $\mathcal{U}_{\mathcal{O}}$ is a model of $\mathcal{O}$ by Lemma 3. By Definition 6, $\mathcal{R}_{\mathcal{O}} \not\models \alpha$ implies $\mathcal{U}_{\mathcal{O}} \not\models \alpha$ and hence, $\mathcal{O} \not\models \alpha$. $\square$

## Filtration Phase

The Datalog model $\mathcal{R}_{\mathcal{O}}^{\infty}$ cannot be used directly for solving BCQ entailment over $\mathcal{O}$ as this set of facts entails BCQs which are not entailed by $\mathcal{O}$.

**Example 3.** *Consider the ontology $\mathcal{O}$ from Example 2, and the BCQ $q = \exists \vec{x}.E(x) \wedge T(x,y) \wedge R(y,z) \wedge C(z)$. Even though $\mathcal{R}_{\mathcal{O}}^{\infty} \models q$, $\mathcal{U}_{\mathcal{O}} \not\models q$ and hence $\mathcal{O} \not\models q$.*

The BCQ from Example 3 contains a fork—a confluent chain of binary atoms—which finds a match in $\mathcal{R}_{\mathcal{O}}^{\infty}$ but not in all models of $\mathcal{O}$. There are other types of unsound matches that can be entailed by $\mathcal{R}_{\mathcal{O}}^{\infty}$, e.g., cycles.

**Example 4.** *For the ontology $\mathcal{O} = \{A \sqsubseteq \exists R.A\}$ and BCQ $q = \exists x.R(x,x)$, we find $\mathcal{R}_{\mathcal{O}}^{\infty} \models q$ whereas $\mathcal{U}_{\mathcal{O}} \not\models q$.*

To avoid unsound answers, we adapt a technique by Feier et al. (2015): first we expand $\mathcal{R}_{\mathcal{O}}^{\infty}$ into the set of facts $\mathcal{C}_{\mathcal{O}}$ on which we compute query matches; then we employ a filtration method to determine which of these matches correspond to a match in all models. We specify $\mathcal{C}_{\mathcal{O}}$ as a set of facts over the signature of $\mathcal{R}_{\mathcal{O}}$, where we create several copies of the form $t^i_{\mathbb{R},\mathbb{C}}$ with $i \in \{0,1,2\}$ and $\mathbb{R} \in \mathbf{R}^{\sqcap}$ for each "anonymous" individual $t_{\mathbb{C}}$ with $\mathsf{N}(t_{\mathbb{C}}) \notin \mathcal{R}_{\mathcal{O}}^{\infty}$. Moreover, for every role $R \in \mathbf{R}$, we introduce an auxiliary role $R_{\triangleright}$ to record the original order of elements in $\mathcal{R}_{\mathcal{O}}^{\infty}$.

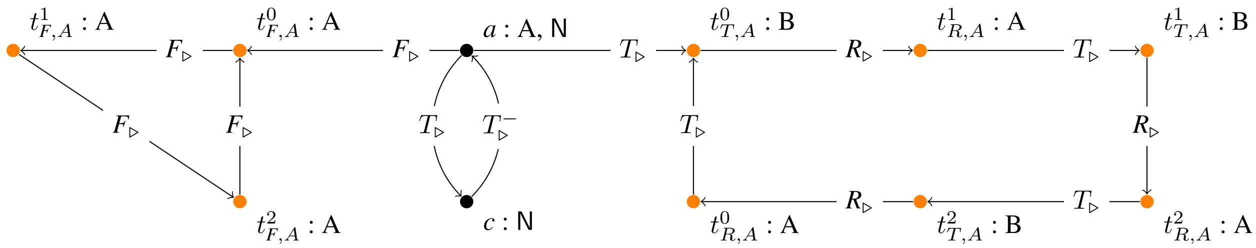

Figure 6: Representation of $\mathcal{C}_\mathcal{O}$

**Definition 7.** *Let $\mathbf{I}^\times$ be the set of all individuals of the form $t^i_{\mathbb{R},\mathbb{C}}$, such that $\mathbb{R} \in \mathbf{R}^\sqcap$, $\mathbb{C} \in \mathbf{C}^\sqcap$, and $i \in \{0, 1, 2\}$. Let $\preceq$ be some reflexive total order on the set of all pairs $\langle \mathbb{R}, \mathbb{C} \rangle$.*

*Then, we define $\mathcal{C}_\mathcal{O}$ as the minimal set of facts over individuals from $\mathbf{I}^+ \cup \mathbf{I}^\times$ such that:*

- *If $\mathsf{N}(a), R(a, b), \mathsf{N}(b) \in \mathcal{R}^\infty_\mathcal{O}$, then $R_\triangleright(a, b), R^-_\triangleright(b, a) \in \mathcal{C}_\mathcal{O}$, where $a, b \in \mathbf{I}^+$ can also be of form $t_\mathbb{C}$.*
- *If $\mathsf{N}(a), \mathbb{R}(a, t_\mathbb{C}) \in \mathcal{R}^\infty_\mathcal{O}$ and $\mathsf{N}(t_\mathbb{C}) \notin \mathcal{R}^\infty_\mathcal{O}$, then $\{R_\triangleright(a, t^0_{\mathbb{R},\mathbb{C}}) \mid R \in \mathbb{R}\} \subseteq \mathcal{C}_\mathcal{O}$.*
- *If $t^i_{\mathbb{R},\mathbb{C}}$ is in $\mathcal{C}_\mathcal{O}$, $\mathbb{S}(t_\mathbb{C}, t_\mathbb{D}) \in \mathcal{R}^\infty_\mathcal{O}$ and $\mathsf{N}(t_\mathbb{D}) \notin \mathcal{R}^\infty_\mathcal{O}$, then $\{S_\triangleright(t^i_{\mathbb{R},\mathbb{C}}, t^j_{\mathbb{S},\mathbb{D}}) \mid S \in \mathbb{S}\} \subseteq \mathcal{C}_\mathcal{O}$ where*
  - *$j = (i + 1) \mod 2$ if $\langle \mathbb{R}, \mathbb{C} \rangle \preceq \langle \mathbb{S}, \mathbb{D} \rangle$, and*
  - *$j = i$ otherwise.*
- *If $t^i_{\mathbb{S},\mathbb{C}}$ occurs in $\mathcal{C}_\mathcal{O}$ and $R(t_\mathbb{C}, a), \mathsf{N}(a) \in \mathcal{R}^\infty_\mathcal{O}$, then $R_\triangleright(t^i_{\mathbb{S},\mathbb{C}}, a)$, $R^-_\triangleright(a, t^i_{\mathbb{S},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$.*
- *If $a \in \mathbf{I}^+$ is in $\mathcal{C}_\mathcal{O}$ and $C(a) \in \mathcal{R}^\infty_\mathcal{O}$, then $C(a) \in \mathcal{C}_\mathcal{O}$.*
- *If $t^i_{\mathbb{R},\mathbb{C}}$ is in $\mathcal{C}_\mathcal{O}$ and $C(t_\mathbb{C}) \in \mathcal{R}^\infty_\mathcal{O}$, then $C(t^i_{\mathbb{R},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$.*
- *If $R_\triangleright(a, b) \in \mathcal{C}_\mathcal{O}$, then $R(a, b), R^-(b, a) \in \mathcal{C}_\mathcal{O}$.*

**Example 5.** *Let $\mathcal{O}$ be the following ontology:*

$$A \sqsubseteq \exists T.B \quad B \sqsubseteq \exists R.A \quad A \sqsubseteq \exists F.A \quad A(b) \quad T(a, c)$$

*Assuming $\langle T, B \rangle \preceq \langle R, A \rangle$, the set of facts $\mathcal{C}_\mathcal{O}$ is as represented in Figure 6.*

As with $\mathcal{R}^\infty_\mathcal{O}$, we cannot directly use the set $\mathcal{C}_\mathcal{O}$ to solve BCQ entailment over $\mathcal{O}$.

**Example 6.** *Let $\mathcal{O}$ be the ontology from Example 5. There are BCQs, such as $q = \exists x, y, z.F(x, y) \wedge F(y, z) \wedge F(z, x)$, with $\mathcal{C}_\mathcal{O} \models q$ and $\mathcal{O} \not\models q$. Note how $q$ matches a cycle in $\mathcal{C}_\mathcal{O}$ that may not occur in models of $\mathcal{O}$, such as $\mathcal{U}_\mathcal{O}$. Also, there are BCQs containing forks such as $\exists x, y.F(x, y) \wedge T(z, y)$, entailed by $\mathcal{C}_\mathcal{O}$ but not by $\mathcal{O}$.*

Alas, not all cyclic matches entailed by $\mathcal{C}_\mathcal{O}$ are spurious! This is easy to see if we consider cycles consisting only of named individuals: the ontology $\{R(a, b), S(b, c), V(c, a)\}$, e.g., entails the BCQ $\exists x, y, z.R(x, y) \wedge S(y, z) \wedge V(z, w)$. Interestingly, cyclic BCQs might also have correct matches that involve anonymous individuals.

**Example 7.** *Consider the following ontology $\mathcal{O}$:*

$$A \sqsubseteq \exists R.B \quad B \sqsubseteq \exists S.C \quad C \sqsubseteq \{c\} \quad V(c, a) \quad A(a)$$

*The cyclic BCQ $q = \exists x, y, z.R(x, y) \wedge S(y, z) \wedge V(z, x)$ has a match in $\mathcal{C}_\mathcal{O}$ where $y$ maps to an anonymous individual. This match is correct and $\mathcal{O} \models q$.*

To filter unsound BCQ answers we introduce the notion of a valid match.

**Definition 8.** *Consider an ontology $\mathcal{O}$, a BCQ $q = \exists \vec{x}.\beta$, and a substitution $\sigma$ such that $\beta\sigma \subseteq \mathcal{C}_\mathcal{O}$. Then, $G_{q,\sigma}$ is the minimal directed graph (DG) such that,*

1. *for each $x \in \vec{x}$, there is a vertex $v(x)$ in $G_{q,\sigma}$; and*
2. *for all $x, y \in \vec{x}$, $v(x) \to v(y) \in G_{q,\sigma}$ if there is some $R \in \mathbf{R}$ such that (i) $R(x, y) \in \beta$ or $R^-(y, x) \in \beta$, (ii) $R_\triangleright(x, y)\sigma \in \mathcal{C}_\mathcal{O}$, and (iii) $R^-_\triangleright(y, x)\sigma \notin \mathcal{C}_\mathcal{O}$.*

*Moreover, $F_{q,\sigma}$ is the graph that results from exhaustively applying the following rule to $G_{q,\sigma}$: if there are some $x, y, z, w \in \vec{x}$ with $\sigma(x) = \sigma(y)$, $v(z) = v(w)$, and $\{v(x) \to v(z), v(y) \to v(w)\} \subseteq G_{q,\sigma}$; then $v(x) = v(y)$.*

*The substitution $\sigma$ is a* valid match *for $\mathcal{O}$ and $q$ if $F_{q,\sigma}$ is a rooted directed forest.*

If $F_{q,\sigma}$ is a forest, then no node $v(x)$ mapped to an anonymous individual can be reached from two different nodes in $F_{q,\sigma}$, thus preventing spurious answers due to forks in the query. By requiring $F_{q,\sigma}$ to be acyclic, we also prevent unsound answers that may be due to cycles in the query being mapped to cycles of anonymous individuals in $\mathcal{C}_\mathcal{O}$ that are no instances of the special predicate $\mathsf{N}$. Cyclic structures over such anonymous individuals may never occur in the universal model $\mathcal{U}_\mathcal{O}$ of an Horn-$\mathcal{ALCHOIQ}$ ontology $\mathcal{O}$ and hence, such queries may not be entailed.

**Theorem 3.** *A consistent ontology $\mathcal{O}$ entails a BCQ $q$ iff there is some valid match $\sigma$ for $\mathcal{O}$ and $q$.*

Note that we can check for inconsistency using Theorem 1, in which case all BCQs are entailed.

Our proof of Theorem 3 in the appendix proceeds in several steps. For soundness, we first show that we can correctly read off certain entailments from $\mathcal{C}_\mathcal{O}$. As in the proof of Lemma 1, we define a characteristic class expression $\mathsf{Ex}(a)$ for each $a$ of $\mathcal{C}_\mathcal{O}$: if $a \in \mathbf{I}$, then $\mathsf{Ex}(a) = \{a\}$; else if $a = t_\mathbb{C}$ or $a = t^i_{\mathbb{R},\mathbb{C}}$, then $\mathsf{Ex}(a) = \mathbb{C}$. One can then show similar semantic correspondences as for Lemma 1. Soundness is then established by considering a valid match $\sigma$ on $\mathcal{C}_\mathcal{O}$, and by constructing, for an arbitrary model $\mathcal{I}$ of $\mathcal{O}$, a corresponding match of the BCQ. Completeness is established by an inverse construction, turning a BCQ match on $\mathcal{U}_\mathcal{O}$ into a valid match on $\mathcal{C}_\mathcal{O}$. In both cases, we proceed inductively by defining matches for an increasing sequence of queries $q_1, \ldots, q_n$ where $q_n = q$ is the BCQ under consideration.

As in the case of assertion entailment, our approach yields a worst-case optimal algorithm for both Horn-$\mathcal{ALCHOIQ}$ and $\mathcal{ELHO}$:

**Theorem 4.** *The approach of Theorem 3 decides BCQ entailment for Horn-$\mathcal{ALCHOIQ}$ in exponential time for combined complexity, and in polynomial time for data complexity. When restricting to the relevant signature of $\mathcal{ELHO}$ for materialisation as in Theorem 2, the algorithm runs in non-deterministic polynomial time for combined complexity.*

*Proof.* We first consider the case of Horn-$\mathcal{ALCHOIQ}$. The argument for proving Theorem 2—grounding an exponential rule set $\mathcal{R}_{\mathcal{O}}$, and computing its propositional entailments in linear time—shows that the chase $\mathcal{R}_{\mathcal{O}}^{\infty}$ is of exponential size (and of polynomial size with respect to the number of assertions). The extension of $\mathcal{R}_{\mathcal{O}}^{\infty}$ to $\mathcal{C}_{\mathcal{O}}$ is possible in polynomial time in the size of $\mathcal{R}_{\mathcal{O}}^{\infty}$. The number of possible matches $\sigma$ of a BCQ $q$ on $\mathcal{C}_{\mathcal{O}}$ is exponential in the size of $q$ and polynomial in the size of $\mathcal{C}_{\mathcal{O}}$. For each $\sigma$, we can check in polynomial time in $\mathcal{C}_{\mathcal{O}}$ whether it is a valid match. The overall procedure therefore runs in exponential time in the size of $\mathcal{O}$ and $q$, and in polynomial time in the number of assertions in $\mathcal{O}$.

In the case of $\mathcal{ELHO}$, as argued for Theorem 2, $\mathcal{R}_{\mathcal{O}}^{\infty}$ is of polynomial size and can be computed in polynomial time by restricting to the signature relevant for $\mathcal{ELHO}$. $\mathcal{C}_{\mathcal{O}}$ contains only individuals $t^{i}_{\mathbb{R},\mathbb{C}}$ for cases where $\mathbb{R}(c, t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$, so that their overall number is polynomial in the size of $\mathcal{R}_{\mathcal{O}}^{\infty}$. The same therefore holds for the size of $\mathcal{C}_{\mathcal{O}}$. A query match $\sigma$ can be guessed non-deterministically in polynomial time, and it is again polynomial to verify that it is valid. □

The previous results are worst-case optimal: BCQ entailment over Horn-$\mathcal{ALCHOIQ}$ ontologies is ExpTime-hard (and P-hard for data complexity) since this is true even for standard reasoning in this DL (Krötzsch, Rudolph, and Hitzler 2013); NP-hardness of BCQ entailment for $\mathcal{EL}$ ontologies follows from the fact that BCQ entailment over a set of assertions is already NP-hard.

## Proof of Concept

We evaluate a prototype implementation of the materialisation phase, which we consider the performance-critical part of our algorithm. In contrast, our filtration phase uses a polynomial algorithm, which is computationally similar to the filtration in other combined approaches that have already been shown to be efficient in practical cases (Feier et al. 2015). Since materialisation decides fact entailment for Horn-$\mathcal{ALCHOIQ}$ (Theorem 1), we can meaningfully compare performance against standard DL reasoners.

Our prototype implementation uses the RDFox Datalog engine (SVN version 2776) for computing the chase (Motik et al. 2014), and implements the optimised materialisation where we add new rules on demand during the computation of the chase, as discussed after Theorem 2. We further modify the process by adding facts $\top(c)$ and (if applicable) $\mathsf{N}(c)$ directly when loading or creating new individuals, thus omitting rules $\mathcal{R}_{\mathsf{Top}}$ and facts $\mathcal{R}_{\mathsf{Nom}}$. We also omit $\mathcal{R}_{\mathsf{Eq}}$, since we rely on the built-in equality reasoning support of RDFox instead (Motik et al. 2015). Some of the considered ontologies contain (in)equality assertions of the form $a \not\approx b$ and $a \approx b$. To deal with these, we simply rely on the built-in equality reasoning of RDFox and include an extra rule $x \not\approx x \to \bot(x)$ to detect inconsistencies entailed by inequality assertions.

We use our implementation to solve assertion retrieval, i.e., the reasoning task that consists in computing all class and role assertions that are entailed by a given ontology. We compared performance with that of Konclude (v0.6.2), a leading DL reasoner (Steigmiller, Liebig, and Glimm 2014), which we used as a command-line client on local input files.

Since query answering is most relevant in data-intensive applications, we use ontologies with large sets of assertions (ABoxes). We considered two standard benchmarks, *LUBM* (Guo et al. 2007) and *UOBM* (Ma et al. 2006); and two real-world ontologies from the bio-domain, *Reactome* and *Uniprot*, which were used in the evaluation of PAGOdA (Zhou et al. 2015). We have normalised these ontologies and removed axioms not expressible in Horn-$\mathcal{ALCHOIQ}$, such as role chains or disjunctions. The resulting ontologies contained 108 (LUBM), 254 (UOBM), 481 (Reactome), and 317 (Uniprot) terminological axioms, respectively. None of these ontologies belonged to a known tractable fragment of Horn-$\mathcal{ALCHOIQ}$. For each ontology, we consider ABoxes of various sizes, generated by using the size parameter for the benchmarks (LUBM, UOBM), and by sampling the real-world ABoxes (Reactome, Uniprot) using the method by Zhou et al. (2015). All ontology files and own software used in the evaluation can be found from an online repository.[1]

Our test system is a commodity laptop (16GB RAM, 500GB SSD, CPU i7-8550U/4 cores/1.8GHz, Windows 10). We configured the operating system to allow up to 28GB of virtual memory. We have measured wall-clock times spent during reasoning, ignoring the time required for parsing and loading. Konclude reports detailed times, while for RDFox we have measured the time from within our prototype.

Figure 7 shows the results. Note the logarithmic scale in the case of Reactome. Konclude ran out of memory for the two largest of the Uniprot samples, hence no times are reported there. Detailed measurement results can also be found in our evaluation repository. For LUBM, we started materialisation with 215 rules and computed the chase. Based on the results, another six rules were added and the chase was started again, without any further rules needed this time. Likewise, UOBM and Reactome required the chase to run four times, and Uniprot five times. The total number of rules added for each ontology was 6 (LUBM), 19 (UOBM), 14 (Reactome), and 59 (Uniprot). Considering the exponential number of rules that might be required in the worst case, these figures are quite moderate.

The performance results show that our prototype can already achieve competitive performance. It is about three times faster than Konclude on LUBM, and about the same on UOBM. For Reactome, Konclude has an initial performance advantage but slows down exponentially as ABoxes increase. We see a similar picture for Uniprot, with Konclude running out of memory for the larger datasets, while our prototype continues to scale approximately linearly.

[1] `https://github.com/irina-dragoste/combined-approach-horn_alchoiq`

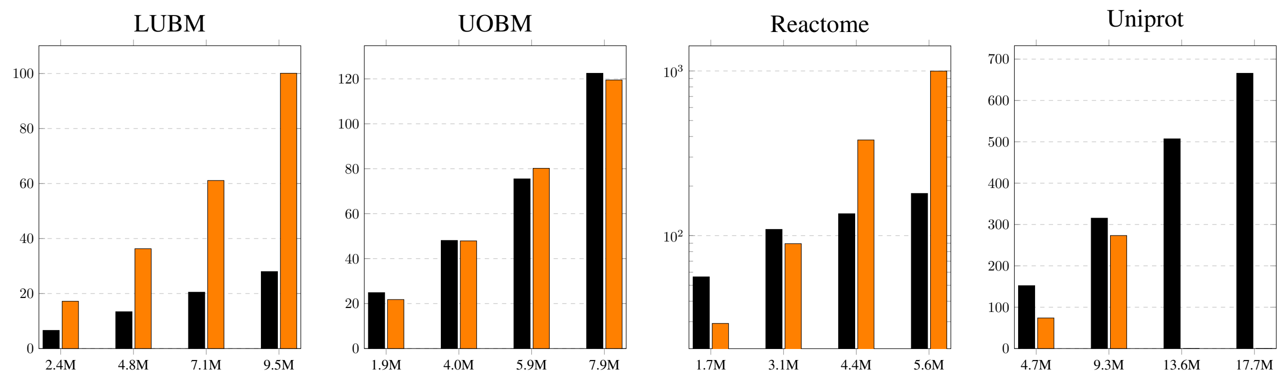


Figure 7: Reasoning times for ABox materialisation for our implementation (black) and Konclude (orange), measuring time in seconds in the $y$-axis and the number of assertions in the $x$-axis

The scalability advantage of RDFox is not unexpected, since DL reasoners are not optimised for ontologies with a large number of assertions. Konclude supports DL ontologies beyond Horn-$\mathcal{ALCHOIQ}$, hence may not take full advantage of optimisations possible for our case. On the other hand, Konclude supports only class assertion retrieval, whereas our implementation also computes the entailed role assertions. Nevertheless, the tasks solved by the two systems are similar enough to use Konclude as a meaningful baseline in a feasibility study. Moreover, there are no dedicated Horn-$\mathcal{ALCHOIQ}$ reasoners, so Konclude still is indicative of the best possible performance available to practitioners today.

## Related Work

Combined approaches have been developed for $\mathcal{ELHO}^r_\bot$ (Stefanoni, Motik, and Horrocks 2013)—a DL that extends our $\mathcal{ELHO}$ with role ranges—, DL-Lite (Kontchakov et al. 2010), DL-Lite$_\mathcal{R}$ (Lutz et al. 2013)—DL-Lite with role hierarchies—, and RSA ontologies (Feier et al. 2015)—a tractable class of Horn ontologies that extends $\mathcal{ELHO}$ and DL-Lite$_\mathcal{R}$ (Carral et al. 2014). Our approach solves BCQ entailment over the more expressive class of Horn-$\mathcal{ALCHOIQ}$ ontologies, though it does not always achieve the same worst-case complexity on specific fragments.

The translation of DL reasoning problems to Datalog has also been explored previously, for logics from $\mathcal{EL}$ (Krötzsch 2011) to Horn-$\mathcal{SROIQ}$ (Ortiz, Rudolph, and Simkus 2010). Interestingly, the latter approach by Ortiz et al. has also been used to establish upper bounds for deciding BCQ entailment in such expressive DLs (Ortiz, Rudolph, and Simkus 2011), which indirectly also yields a Datalog-based procedure for Horn-$\mathcal{ALCHOIQ}$. However, Ortiz et al. use polynomial Datalog programs with predicates of polynomial arity, while we obtain exponential Datalog programs with predicates of fixed arity. This allows us to use subsets of rules to obtain lower complexities for tractable fragments, and to use existing Datalog engines (which are not prepared to handle predicates with arities $> 100$).

Eiter et al. gave a method for rewriting Horn-$\mathcal{SHIQ}$ ontologies and CQs to Datalog with fixed predicate arities (2012). A crucial difference to all many other works in DL query answering is that the unique name assumption is made. The approach also differs from ours in that it supports transitive roles but no nominals.

Our materialisation phase shares some similarities with consequence-based reasoning procedures (Kazakov 2009; Simančík, Kazakov, and Horrocks 2011; Simančík, Motik, and Horrocks 2014; Bate et al. 2016). Such approaches make use of "types"—akin to our individuals of the form $t_\mathbb{C}$—, which represent certain combinations of features in arbitrary models. To the best of our knowledge, no such procedure supports DLs with nominals, "at most" quantifiers, and inverse roles—a combination of constructors that is known to be difficult to deal with (Rudolph and Glimm 2010). Moreover, consequence-based reasoning remains a method for standard reasoning, which does not address CQ answering.

## Conclusions

To the best of our knowledge, we have presented the first combined approach for conjunctive query answering over the expressive DL Horn-$\mathcal{ALCHOIQ}$. We combine two powerful methods—consequence-based reasoning and filtration—both of which were shown to enable practically feasible implementations on their own. Indeed, consequence-based reasoning features provable pay-as-you-go behaviour that leads to faster runtimes on simpler ontologies (see Theorem 2), while filtration can take advantage of delegating most work to the query engines of highly optimised databases, which makes this part rather scalable in practice (Feier et al. 2015).

We have also provided empirical evidence for the practical applicability of our method, even on a prototypical implementation that uses a standard Datalog engine without tight integration. A fully integrated system for DL reasoning and query answering therefore seems to be feasible, and indeed would be a promising direction for further research. In particular, we would like to explore the use of our rule engine VLog (Urbani, Jacobs, and Krötzsch 2016), which promises advantages in terms of memory usage. VLog does not yet support equality reasoning natively, but we conjecture that explicit equality reasoning plays only a minor role for overall performance in our encoding.

# References


Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the $\mathcal{EL}$ envelope. In Kaelbling, L., and Saffiotti, A., eds., *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI'05)*, 364–369. Professional Book Center.

Bate, A.; Motik, B.; Cuenca Grau, B.; Simančík, F.; and Horrocks, I. 2016. Extending consequence-based reasoning to $\mathcal{SRIQ}$. In Baral, C.; Delgrande, J. P.; and Wolter, F., eds., *Proc. 15th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'16)*, 187–196. AAAI Press.

Bienvenu, M.; Hansen, P.; Lutz, C.; and Wolter, F. 2016. First order-rewritability and containment of conjunctive queries in Horn description logics. In Kambhampati, S., ed., *Proc. 25th Int. Joint Conf. on Artificial Intelligence (IJCAI'16)*, 965–971. IJCAI/AAAI Press.

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

Carral, D.; Feier, C.; Cuenca Grau, B.; Hitzler, P.; and Horrocks, I. 2014. Pushing the boundaries of tractable ontology reasoning. In *Proc. 13th Int. Semantic Web Conf. (ISWC'14), Part II*, volume 8797 of *LNCS*, 148–163. Springer.

Carral, D.; Dragoste, I.; and Krötzsch, M. 2018. The combined approach to query answering in Horn-$\mathcal{ALCHOIQ}$. Technical report with full appendix; available at `https://kbs.inf.tu-dresden.de/HornALCHOIQ2018`.

Eiter, T.; Ortiz, M.; Simkus, M.; Tran, T.; and Xiao, G. 2012. Query rewriting for Horn-$\mathcal{SHIQ}$ plus rules. In Hoffmann, J., and Selman, B., eds., *Proc. 26th AAAI Conf. on Artificial Intelligence (AAAI'12)*. AAAI Press.

Feier, C.; Carral, D.; Stefanoni, G.; Grau, B. C.; and Horrocks, I. 2015. The combined approach to query answering beyond the OWL 2 profiles. In Yang, Q., and Wooldridge, M., eds., *Proc. 24th Int. Joint Conf. on Artificial Intelligence (IJCAI'15)*, 2971–2977. AAAI Press.

Guo, Y.; Qasem, A.; Pan, Z.; and Heflin, J. 2007. A requirements driven framework for benchmarking semantic web knowledge base systems. *IEEE Trans. Knowl. Data Eng.* 19(2):297–309.

Kazakov, Y. 2009. Consequence-driven reasoning for Horn $\mathcal{SHIQ}$ ontologies. In Boutilier, C., ed., *Proc. 21st Int. Joint Conf. on AI (IJCAI'09)*, 2040–2045. IJCAI.

Kontchakov, R.; Lutz, C.; Toman, D.; Wolter, F.; and Zakharyaschev, M. 2010. The combined approach to query answering in DL-Lite. In Lin et al. (2010), 247–257.

Kontchakov, R.; Lutz, C.; Toman, D.; Wolter, F.; and Zakharyaschev, M. 2011. The combined approach to ontology-based data access. In Walsh (2011), 2656–2661.

Krötzsch, M.; Rudolph, S.; and Hitzler, P. 2013. Complexities of Horn dls. *ACM Trans. Comput. Logic* 14(1):2:1–2:36.

Krötzsch, M. 2011. Efficient rule-based inferencing for OWL EL. In Walsh (2011), 2668–2673.

Lin, F.; Sattler, U.; and Truszczynski, M., eds. 2010. *Proc. 12th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'10)*. AAAI Press.

Lutz, C.; Seylan, I.; Toman, D.; and Wolter, F. 2013. The combined approach to OBDA: Taming role hierarchies using filters. In Alani, H.; Kagal, L.; Fokoue, A.; Groth, P. T.; Biemann, C.; Parreira, J. X.; Aroyo, L.; Noy, N. F.; Welty, C.; and Janowicz, K., eds., *Proc. 12th Int. Semantic Web Conf. (ISWC'13)*, volume 8218 of *LNCS*, 314–330. Springer.

Ma, L.; Yang, Y.; Qiu, Z.; Xie, G. T.; Pan, Y.; and Liu, S. 2006. Towards a complete OWL ontology benchmark. In *The 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, June 11-14, 2006, Proceedings*, 125–139.

Motik, B.; Nenov, Y.; Piro, R.; Horrocks, I.; and Olteanu, D. 2014. Parallel materialisation of Datalog programs in centralised, main-memory RDF systems. In *AAAI*.

Motik, B.; Nenov, Y.; Piro, R.; and Horrocks, I. 2015. Combining rewriting and incremental materialisation maintenance for datalog programs with equality. In Yang, Q., and Wooldridge, M., eds., *Proceedings of the 24th International Joint Conference on AI, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, 3127–3133. AAAI Press.

Ortiz, M.; Rudolph, S.; and Simkus, M. 2010. Worst-case optimal reasoning for the Horn-DL fragments of OWL 1 and 2. In Lin et al. (2010), 269–279.

Ortiz, M.; Rudolph, S.; and Simkus, M. 2011. Query answering in the Horn fragments of the description logics $\mathcal{SHOIQ}$ and $\mathcal{SROIQ}$. In Walsh (2011), 1039–1044.

Parsia, B.; Matentzoglu, N.; Gonçalves, R. S.; Glimm, B.; and Steigmiller, A. 2017. The OWL reasoner evaluation (ORE) 2015 competition report. *J. of Autom. Reasoning* 59(4):455–482.

Rudolph, S., and Glimm, B. 2010. Nominals, inverses, counting, and conjunctive queries or: Why infinity is your friend! *J. of Artif. Intell. Res.* 39:429–481.

Simančík, F.; Kazakov, Y.; and Horrocks, I. 2011. Consequence-based reasoning beyond Horn ontologies. In Walsh (2011), 1093–1098.

Simančík, F.; Motik, B.; and Horrocks, I. 2014. Consequence-based and fixed-parameter tractable reasoning in description logics. *Artif. Intell.* 209:29–77.

Stefanoni, G.; Motik, B.; and Horrocks, I. 2013. Introducing nominals to the combined query answering approaches for $\mathcal{EL}$. In *Proc. 27th AAAI Conf. on Artificial Intelligence (AAAI'13)*. AAAI Press.

Steigmiller, A.; Liebig, T.; and Glimm, B. 2014. Konclude: System description. *J. of Web Semantics* 27:78–85.

Urbani, J.; Jacobs, C.; and Krötzsch, M. 2016. Column-oriented Datalog materialization for large knowledge graphs. In *Proc. 30th AAAI Conf. on Artificial Intelligence (AAAI'16)*, 258–264. AAAI Press.

Walsh, T., ed. 2011. *Proc. 22nd Int. Joint Conf. on Artificial Intelligence (IJCAI'11)*. AAAI Press/IJCAI.

Zhou, Y.; Cuenca Grau, B.; Nenov, Y.; Kaminski, M.; and Horrocks, I. 2015. PAGOdA: Pay-as-you-go ontology query answering using a Datalog reasoner. *J. of Artif. Intell. Res.* 54:309–367.

## Proof of Lemma 1

We denote the various types of rule in $\mathcal{R}_{\mathsf{Top}}$, $\mathcal{R}_{\mathsf{Role}}$, $\mathcal{R}_{\mathsf{Nm}}$, and $\mathcal{R}_{\mathsf{Eq}}$ by indexing them in the order given. E.g., ($\mathsf{Top}$.2) refers to the rule $\mathbb{R}(x,y) \to \top(x) \wedge \top(y)$, and ($\mathsf{Eq}$.3) refers to $C(x) \wedge x \approx y \to C(y)$.

**Lemma 1.** *If $\mathcal{R}_{\mathcal{O}}$ entails an assertion $\alpha$, then $\mathcal{O} \models \alpha$.*

*Proof.* For all $a \in \mathbf{I}^+$, let $\mathsf{Ex}(a) = \{a\}$ if $a \in \mathbf{I}$, and $\mathsf{Ex}(a) = \mathbb{C}$ if $a$ is of the form $t_{\mathbb{C}}$.

Given a chase sequence $\mathcal{F}^0, \dots, \mathcal{F}^n$ of $\mathcal{R}_{\mathcal{O}}$, we show via induction on $i \in \{0, \dots, n\}$ that the following claims hold for any model $\mathcal{I}$ of $\mathcal{O}$:

(a) If $C(a) \in \mathcal{F}^i$ with $C \in \mathbf{C}$, then $\mathsf{Ex}(a)^{\mathcal{I}} \subseteq C^{\mathcal{I}}$.

(b) If $\mathbb{R}(a,b) \in \mathcal{F}^i$, then for all $\delta \in \mathsf{Ex}(a)^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(b)^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^{\mathcal{I}}$.

(c) If $a \approx b \in \mathcal{F}^i$, then $\mathsf{Ex}(a)^{\mathcal{I}} = \mathsf{Ex}(b)^{\mathcal{I}}$.

(d) If $a \in \mathbf{I}^+$ occurs in $\mathcal{F}^i$, then $\mathsf{Ex}(a)^{\mathcal{I}} \neq \emptyset$.

(e) If $\mathsf{N}(a) \in \mathcal{F}^i$, then $|\mathsf{Ex}(a)^{\mathcal{I}}| = 1$.

The lemma follows from (a) and (b); the rest of the claims are included to structure our induction argument.

For the base case $i = 0$, all the claims trivially hold, since $\mathcal{F}^0 = \emptyset$. For the induction step, consider $i \in \{1, \dots, n\}$, and assume that all the claims hold for $i-1$ (IH). We show that the claims remain true by distinguishing all possible cases based on the type of the rule $\rho = \beta \to \eta \in \mathcal{R}_{\mathcal{O}}$ and the substitution $\sigma$ such that $\beta\sigma \subseteq \mathcal{F}^{i-1}$ and $\mathcal{F}^i = \mathcal{F}^{i-1} \cup \eta\sigma$. Note that each type of rule needs to establish only the claims whose premise might be affected by its application.

(1) Since $\beta\sigma \subseteq \mathcal{F}^{i-1}$, $\{C_j(\sigma(x)) \mid 1 \leq j \leq m\} \subseteq \mathcal{F}^{i-1}$ and hence, $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq \bigcap_{j=1}^m C_j^{\mathcal{I}}$ by (IH.a). Since $\bigsqcap_{j=1}^m C_j \sqsubseteq D \in \mathcal{O}$, $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and (a) holds.

(2) Then, $C(\sigma(x)) \in \mathcal{F}^{i-1}$.

(a) By definition, $\mathsf{Ex}(t_D)^{\mathcal{I}} = D^{\mathcal{I}}$.

(b) By (IH.a), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$. Since $C \sqsubseteq \exists R.D \in \mathcal{O}$, for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ there is some $\gamma \in D^{\mathcal{I}} = \mathsf{Ex}(t_D)^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in R^{\mathcal{I}}$.

(d) By (IH.d), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \neq \emptyset$ and hence, $\mathsf{Ex}(t_D)^{\mathcal{I}} \neq \emptyset$.

(3.1) Since $\rho$ is applicable, $R(\sigma(x), \sigma(y))$, $C(\sigma(y)) \in \mathcal{F}^{i-1}$. By (IH.a) and (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in R^{\mathcal{I}}$. Since $\exists R.C \sqsubseteq D \in \mathcal{O}$, $\delta \in D^{\mathcal{I}}$ and (a) holds.

(3.2) Then, $C(\sigma(x)), \mathbb{R}^-(\sigma(x), t_{\mathbb{C}}) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R}$.

(a) By definition, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^{\mathcal{I}} = \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} \cap D^{\mathcal{I}}$. Hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^{\mathcal{I}} \subseteq \bigcap_{X \in \mathbb{C}} X^{\mathcal{I}} \cap D^{\mathcal{I}}$.

(b) By (IH.a) and (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in (\mathbb{R}^-)^{\mathcal{I}}$. Since $\exists R.C \sqsubseteq D \in \mathcal{O}$, $\gamma \in D^{\mathcal{I}}$ .

(d) By (IH.d), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \neq \emptyset$ and hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap D})^{\mathcal{I}} \neq \emptyset$.

(4.1) Then, $D(\sigma(y)), R^-(\sigma(y), \sigma(x)), C(\sigma(x)), R(\sigma(x), \sigma(z)), D(\sigma(z)), \mathsf{N}(\sigma(z)) \in \mathcal{F}^{i-1}$. By (IH.a), (IH.b), (IH.d), and (IH.e), (1) $\mathsf{Ex}(\sigma(y))^{\mathcal{I}} \neq \emptyset$, (2) for all $\delta \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ there is some $\epsilon \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$ with $\langle \epsilon, \delta \rangle \in R^{\mathcal{I}}$, and (3) $\langle \epsilon, \gamma \rangle \in R^{\mathcal{I}}$ with $\gamma \in \mathsf{Ex}(\sigma(z))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ the only element in $\mathsf{Ex}(\sigma(z))^{\mathcal{I}}$. Since $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$, $\delta = \gamma$. Hence, $\mathsf{Ex}(\sigma(y))^{\mathcal{I}} = \mathsf{Ex}(\sigma(z))^{\mathcal{I}}$ and (c) holds.

(4.2) Then, $C(\sigma(x)), \mathbb{R}(\sigma(x), t_{\mathbb{C}}), D(t_{\mathbb{C}}), \mathbb{S}(\sigma(x), t_{\mathbb{D}}), D(t_{\mathbb{D}}) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R} \cap \mathbb{S}$.

(a) $\mathsf{Ex}(t_{\mathbb{C} \sqcap \mathbb{D}})^{\mathcal{I}} = \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} \cap \mathsf{Ex}(t_{\mathbb{D}})^{\mathcal{I}}$. Hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap \mathbb{D}})^{\mathcal{I}} \subseteq \bigcap_{X \in \mathbb{C} \sqcap \mathbb{D}} X^{\mathcal{I}}$.

(b) By (IH.a) and (IH.b), for all $\epsilon \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$, there is some $\delta \in \mathsf{Ex}(\sigma(t_{\mathbb{C}}))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ with $\langle \epsilon, \delta \rangle \in \mathbb{R}^{\mathcal{I}}$, and there is some $\gamma \in \mathsf{Ex}(t_{\mathbb{D}})^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ with $\langle \epsilon, \gamma \rangle \in \mathbb{S}^{\mathcal{I}}$. Since $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$, $\delta = \gamma$ and $\delta \in \mathsf{Ex}(t_{\mathbb{C} \sqcap \mathbb{D}})^{\mathcal{I}}$ with $\langle \epsilon, \delta \rangle \in (\mathbb{R} \sqcap \mathbb{S})^{\mathcal{I}}$.

(d) By (IH.d), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \neq \emptyset$ and hence, $\mathsf{Ex}(t_{\mathbb{C} \sqcap \mathbb{D}})^{\mathcal{I}} \neq \emptyset$.

(4.3) Then, $D(\sigma(y)), \mathbb{R}^-(\sigma(y), \sigma(x)), C(\sigma(x)), \mathbb{S}(\sigma(x), t_{\mathbb{C}}), D(t_{\mathbb{C}}) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R} \cap \mathbb{S}$.

(a) By (IH.a) and (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, (1) there is some $\epsilon \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$ with $\langle \delta, \epsilon \rangle \in (\mathbb{R}^-)^{\mathcal{I}}$, and (2) there is some $\gamma \in \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ with $\langle \epsilon, \gamma \rangle \in \mathbb{S}^{\mathcal{I}}$. Since $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$, $\delta \in \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} = \mathbb{C}^{\mathcal{I}}$.

(b) Moreover, $\langle \delta, \epsilon \rangle \in (\mathbb{R}^- \sqcap \mathbb{S}^-)^{\mathcal{I}}$.

(4.4) Then, $D(\sigma(y)), R^-(\sigma(y), \sigma(x)), C(\sigma(x)), \mathsf{N}(\sigma(x)) \in \mathcal{F}^{i-1}$. By (IH.a), (IH.b), (IH.d), and (IH.e), (1) there is a single element $\epsilon \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$, (2) $\langle \epsilon, \delta \rangle \in R^{\mathcal{I}}$ for all $\delta \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, and (3) $\mathsf{Ex}(\sigma(y))^{\mathcal{I}} \neq \emptyset$. Since $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$, $|\mathsf{Ex}(\sigma(y))^{\mathcal{I}}| = 1$ and (e) holds.

(5) Then, $C(\sigma(x)) \in \mathcal{F}^{i-1}$. By (IH.a) and (IH.d), there is some $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$. Since $C \sqsubseteq \{a\} \in \mathcal{O}$, $\delta = a^{\mathcal{I}}$ for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$. Hence, $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} = \mathsf{Ex}(a)^{\mathcal{I}} = \{a^{\mathcal{I}}\}$ and (c) holds.

(6.1) Then, $R(\sigma(x), \sigma(y)) \in \mathcal{F}^{i-1}$. By (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in R^{\mathcal{I}}$. Then, $\langle \delta, \gamma \rangle \in \mathbb{S}^{\mathcal{I}}$ with $\mathbb{S} = \{S \mid R \sqsubseteq^*_{\mathcal{O}} S\}$ and (b) holds.

(6.2) Analogous to the previous case.

(7) By definition, $\mathsf{Ex}(a)^{\mathcal{I}} = \{a^{\mathcal{I}}\}$ and (d) holds. Since $C(a) \in \mathcal{O}$, $a^{\mathcal{I}} \in C^{\mathcal{I}}$ and (a) holds.

(8) By definition, $\mathsf{Ex}(a)^{\mathcal{I}} = \{a^{\mathcal{I}}\}$ and $\mathsf{Ex}(b)^{\mathcal{I}} = \{b^{\mathcal{I}}\}$, and (d) holds. Since $R(a, b) \in \mathcal{O}$, $\langle a^{\mathcal{I}}, b^{\mathcal{I}} \rangle \in R^{\mathcal{I}}$ and hence, (b) holds.

(Top.1) Since $\mathcal{I}$ is an interpretation, $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq \top^{\mathcal{I}}$ and (a) holds.

(Top.2) Analogous to the previous case.

(Role.1) Then, $\mathbb{R}(\sigma(x), \sigma(y)), \mathsf{N}(\sigma(y)) \in \mathcal{F}^{i-1}$. By (IH.b), (IH.d) and (IH.e), there is at least one element $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ such that $\langle \delta, \gamma \rangle \in \mathbb{R}^{\mathcal{I}}$ with $\gamma$ the only element in $\mathsf{Ex}(\sigma(y))^{\mathcal{I}}$. Therefore, $\langle \gamma, \delta \rangle \in \mathbb{R}^{-\mathcal{I}}$ and (b) holds.

(Role.2) Then, $\mathbb{R}(\sigma(x), \sigma(y)) \in \mathcal{F}^{i-1}$ with $R \in \mathbb{R}$. By (IH.b), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^{\mathcal{I}}$, and (b) holds.

(Nm.1) By definition, $\mathsf{Ex}(a)^{\mathcal{I}} = \{a^{\mathcal{I}}\}$. Hence, $|\mathsf{Ex}(a)^{\mathcal{I}}| = 1$ and (d) and (e) hold.

(Nm.2) Since $\mathcal{I}$ is an interpretation, $\mathsf{Ex}(a)^{\mathcal{I}} \subseteq \top^{\mathcal{I}}$ and (a) holds. By definition, $\mathsf{Ex}(a)^{\mathcal{I}} = \{a^{\mathcal{I}}\}$ and hence, (d) holds.

(Eq.1) Then, $\sigma(x) \approx \sigma(y) \in \mathcal{F}^{i-1}$. By (IH.c), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} = \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$ and (c) holds.

(Eq.2) Then, $\sigma(x) \approx \sigma(y)$, $\sigma(y) \approx \sigma(z) \in \mathcal{F}^{i-1}$. By (IH.c), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} = \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$ and $\mathsf{Ex}(\sigma(y))^{\mathcal{I}} = \mathsf{Ex}(\sigma(z))^{\mathcal{I}}$, and (c) holds.

(Eq.3) Then, $C(\sigma(x)), \sigma(x) \approx \sigma(y) \in \mathcal{F}^{i-1}$. By (IH.c), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} = \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$. We study the following cases:

- Let $C \in \mathbf{C}$. By (IH.a), $\mathsf{Ex}(\sigma(x))^{\mathcal{I}} \subseteq C^{\mathcal{I}}$ and (a) holds.
- Let $C = \mathsf{N}$. By (IH.e), $|\mathsf{Ex}(\sigma(x))^{\mathcal{I}}| = 1$ and (e) holds.

(Eq.4) Then, $\mathbb{R}(\sigma(x), \sigma(y)), \sigma(x) \approx \sigma(z) \in \mathcal{F}^{i-1}$. By (IH.b) and (IH.c), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}} = \mathsf{Ex}(\sigma(z))^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^{\mathcal{I}}$. Hence, (b) holds.

(Eq.5) Then, $\mathbb{R}(\sigma(x), \sigma(y)), \sigma(y) \approx \sigma(z) \in \mathcal{F}^{i-1}$. By (IH.b) and (IH.c), for all $\delta \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(\sigma(y))^{\mathcal{I}} = \mathsf{Ex}(\sigma(z))^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in \mathbb{R}^{\mathcal{I}}$. Hence, (b) holds.

$\square$

## Proof of Lemma 10

In this section, we show Lemma 10, which can be directly used to solve the second implication in the proof of Lemma 3.

**Lemma 5.** *If $C(c) \in \mathcal{R}^\infty_{\mathcal{O}}$ for all $C \in \mathbb{E}$ for some $c \in \mathbf{I}^+$ and $\mathbb{E} \in \mathbf{C}^\sqcap$, then:*

*(a) If $D(t_{\mathbb{E}}) \in \mathcal{R}^\infty_{\mathcal{O}}$ with $D \in \mathbf{C}$, then $D(c) \in \mathcal{R}^\infty_{\mathcal{O}}$.*

(b) *If* $\mathbb{R}(t_{\mathbb{E}}, b) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ *for some* $b \in \mathbf{I}^{+}$*, then* $\mathbb{R}(c, b) \in \mathcal{R}^{\infty}_{\mathcal{O}}$*.*

(c) *If* $t_{\mathbb{E}} \approx b \in \mathcal{R}^{\infty}_{\mathcal{O}}$ *(resp.* $b \approx t_{\mathbb{E}} \in \mathcal{R}^{\infty}_{\mathcal{O}}$*) for some* $b \in \mathbf{I}^{+}$*, then* $c \approx b \in \mathcal{R}^{\infty}_{\mathcal{O}}$ *(resp.* $b \approx c \in \mathcal{R}^{\infty}_{\mathcal{O}}$*).*

(d) *If* $\mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$*, then* $t_{\mathbb{E}} \approx c \in \mathcal{R}^{\infty}_{\mathcal{O}}$*.*

*Proof.* We show (a)-(d) via induction. Namely, we verify that, for all rules $\rho = \beta \to \eta \in \mathcal{R}_{\mathcal{O}}$ and all substitutions $\sigma$ with $\beta\sigma \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$, claims (a)-(d) hold for all the facts in $\eta\sigma$; provided that (IH) these claims hold for all the facts in $\beta\sigma$. We show that the this implication holds by distinguishing all possible cases based on the type of the rule $\rho$ and the range of the substitution $\sigma$.

(1) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $\{C_i(t_{\mathbb{E}}) \mid 1 \leq i \leq n\} \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$ and $\{C_i(c) \mid 1 \leq i \leq n\} \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$. Hence, $D(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (a) holds.

(2) We study two cases:
- Let $t_{\mathbb{E}} = t_D$. Then, $\mathbb{E} = D$, and (a) holds since $D(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by the premise of the lemma.
- Let $\sigma(x) = t_{\mathbb{E}}$. Then, $C(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $C(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $R(c, t_D) \in \mathcal{R}^{\infty}_{\mathcal{O}}$, and (b) holds.

(3.1) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $R(t_{\mathbb{E}}, \sigma(y)), C(\sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $R(c, \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $D(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$, and (a) holds.

(3.2) We study two cases:
- Let $t_{\mathbb{E}} = t_{\mathbb{C} \sqcap D}$. Then, $\mathbb{E} = \mathbb{C} \sqcap D$, and (a) holds since $\{X(c) \mid X \in \mathbb{C} \sqcap D\} \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$ by the premise of the lemma.
- Let $\sigma(x) = t_{\mathbb{E}}$. Then, $C(t_{\mathbb{E}}), \mathbb{R}^{-}(t_{\mathbb{E}}, t_{\mathbb{C}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $C(c), \mathbb{R}^{-}(c, t_{\mathbb{C}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $\mathbb{R}^{-}(c, t_{\mathbb{C} \sqcap D})$ and (b) holds.

(4.1) We study two cases:
- Let $t_{\mathbb{E}} = \sigma(y)$. Then, $D(t_{\mathbb{E}}), R^{-}(t_{\mathbb{E}}, \sigma(x)), C(\sigma(x)), R(\sigma(x), \sigma(z)), D(\sigma(z)), \mathsf{N}(\sigma(z)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $D(c), R^{-}(c, \sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $c \approx \sigma(z) \in \mathcal{R}^{\infty}_{\mathcal{O}}$
- Let $t_{\mathbb{E}} = \sigma(z)$. Then, $\mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $t_{\mathbb{E}} \approx c \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $\sigma(y) \approx c \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by (4.1) and ($\mathsf{Eq}$.2).

In either case, (c) holds.

(4.2) We study two cases:
- Let $t_{\mathbb{E}} = t_{\mathbb{C} \sqcap \mathbb{D}}$. Then, $\mathbb{E} = \mathbb{C} \sqcap \mathbb{D}$, and (a) holds since $\{X(c) \mid X \in \mathbb{C} \sqcap \mathbb{D}\} \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$ by the premise of the lemma.
- Let $\sigma(x) = t_{\mathbb{E}}$. Then, $C(t_{\mathbb{E}}), \mathbb{R}(t_{\mathbb{E}}, t_{\mathbb{C}}), D(t_{\mathbb{C}}), \mathbb{S}(t_{\mathbb{E}}, t_{\mathbb{D}}), D(t_{\mathbb{D}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $C(c), \mathbb{R}(c, t_{\mathbb{C}}), \mathbb{S}(c, t_{\mathbb{D}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $(\mathbb{R} \sqcap \mathbb{S})(c, t_{\mathbb{C} \sqcap \mathbb{D}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (b) holds.

(4.3) Let $\sigma(y) = t_{\mathbb{E}}$. Then, $D(t_{\mathbb{E}}), \mathbb{R}^{-}(t_{\mathbb{E}}, \sigma(x)), C(\sigma(x)), \mathbb{S}(\sigma(x), t_{\mathbb{C}}), D(t_{\mathbb{C}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $D(c), \mathbb{R}^{-}(c, \sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $\{X(c) \mid X \in \mathbb{C}\} \subseteq \mathcal{R}^{\infty}_{\mathcal{O}}$ and (a) holds. Furthermore, $(\mathbb{R}^{-} \sqcap \mathbb{S}^{-})(c, \sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (b) holds.

(4.4) Let $\sigma(y) = t_{\mathbb{E}}$. Then, $D(t_{\mathbb{E}}), R^{-}(t_{\mathbb{E}}, \sigma(x)), C(\sigma(x)), \mathsf{N}(\sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $R(\sigma(x), t_{\mathbb{E}})$, $\mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by (4.4) and ($\mathsf{Role}$.1). Moreover, $D(c), R^{-}(c, \sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $c \approx t_{\mathbb{E}} \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by (4.1), and (d) holds.

(5) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $C(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $C(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $c \approx a \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (c) holds.

(6.1) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $R(t_{\mathbb{E}}, \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $R(c, \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $\mathbb{S}(c, \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (b) holds.

(6.2) Analogous to the previous case.

($\mathsf{Top}$.1) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $C(t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $C(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by IH. Hence, $\top(c) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and (a) holds.

($\mathsf{Top}$.2) Analogous to the previous case

(Role.1) Let $\sigma(y) = t_{\mathbb{E}}$. Then, $\mathbb{R}(\sigma(x), t_{\mathbb{E}}), \mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $t_{\mathbb{E}} \approx c \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Moreover, $\mathbb{R}(\sigma(x), c), \mathsf{N}(c) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by (Eq.3) and (Eq.5). Hence, $\mathbb{R}^{-}(c, \sigma(x)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (b) holds.

(Role.2) Let $\sigma(x) = t_{\mathbb{E}}$. Then, $\mathbb{R}(t_{\mathbb{E}}, \sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\mathbb{R}(c, \sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $R(c, \sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (b) holds.

(Eq.1) We consider two cases:
- Let $t_{\mathbb{E}} = \sigma(x)$. Then, $t_{\mathbb{E}} \approx \sigma(y) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $c \approx \sigma(y) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $\sigma(y) \approx c \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (c) holds.
- Let $t_{\mathbb{E}} = \sigma(y)$. Analogous to the previous case.

(Eq.2) We consider two cases:
- Let $t_{\mathbb{E}} = \sigma(x)$. Then, $t_{\mathbb{E}} \approx \sigma(y), \sigma(y) \approx \sigma(z) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $c \approx \sigma(y) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $c \approx \sigma(z) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (c) holds.
- Let $t_{\mathbb{E}} = \sigma(z)$. Analogous to the previous case.

(Eq.3) Let $t_{\mathbb{E}} = \sigma(y)$. Then, $C(\sigma(x)), \sigma(x) \approx t_{\mathbb{E}} \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\sigma(x) \approx c \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $C(c) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (a) holds.

(Eq.4) Let $t_{\mathbb{E}} = \sigma(z)$. Then, $\mathbb{R}(\sigma(x), \sigma(y)), \sigma(x) \approx t_{\mathbb{E}} \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\sigma(x) \approx c \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $\mathbb{R}(c, \sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and (b) holds.

(Eq.5) Let $t_{\mathbb{E}} = \sigma(x)$. Then, $\mathbb{R}(t_{\mathbb{E}}, \sigma(y)), \sigma(y) \approx \sigma(z) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\mathbb{R}(c, \sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by IH. Hence, $\mathbb{R}(c, \sigma(z))$ and (b) holds.

□

**Lemma 6.** *If* $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ *for some* $\delta, \gamma \in \Delta^{\mathcal{U}_{\mathcal{O}}}$*, then* $\delta = \gamma$.

*Proof.* The lemma follows if $\mathsf{N}(\mathsf{i}(\delta)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Note that, this implies that both $\delta$ and $\gamma$ are equivalence classes with $\mathsf{i}(\delta) \in \gamma$ and $\mathsf{i}(\gamma) \in \delta$ by Definition 5 and hence, $\delta = \gamma$. Therefore, the lemma follows if $a \approx b \in \mathcal{R}_{\mathcal{O}}^{\infty}$ implies $\mathsf{N}(a), \mathsf{N}(b) \in \mathcal{R}_{\mathcal{O}}^{\infty}$—an implication proven via induction in the following paragraph.

We verify that, for all rules $\rho = \beta \to x \approx y \in \mathcal{R}_{\mathcal{O}}$ and all substitutions $\sigma$ with $\beta\sigma \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$, the previous implication holds for $\sigma(x) \approx \sigma(y)$, provided that (IH) it holds for the facts in $\beta\sigma$. We show this in the following case by case analysis in which we consider the four types of rules of the form $\beta \to x \approx y$ which may occur in $\mathcal{R}_{\mathcal{O}}$.

(Eq.1) Then, $\sigma(y) \approx \sigma(x) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By IH, $\mathsf{N}(\sigma(y)), \mathsf{N}((\sigma(x))) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(Eq.2) Then, $\sigma(x) \approx \sigma(y), \sigma(y) \approx \sigma(z) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ . By IH, $\mathsf{N}(\sigma(x)), \mathsf{N}(\sigma(z)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(4.1) Then, $D(\sigma(y)), R^{-}(\sigma(y), \sigma(x)), C(\sigma(x)), R(\sigma(x), \sigma(z)), D(\sigma(z)), \mathsf{N}(\sigma(z)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), (Eq.1), and (Eq.3), $\sigma(y) \approx \sigma(z), \sigma(z) \approx \sigma(y), \mathsf{N}(\sigma(y)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(5) Then, $C(\sigma(x)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (Nm.1), (5), and (Eq.3), $\mathsf{N}(a), a \approx \mathsf{i}(\sigma(x)), \mathsf{N}(\sigma(x)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

□

**Lemma 7.** *If* $\mathbb{V}(t_{\mathbb{E}}, b) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ *and* $\mathsf{N}(t_{\mathbb{E}}) \notin \mathcal{R}_{\mathcal{O}}^{\infty}$ *with* $t_{\mathbb{E}} \in \mathbf{I}^{+} \setminus \mathbf{I}$ *and* $b \in \mathbf{I}$*, then* $\mathbb{V}(t_{\mathbb{E}}, t_{\mathbb{D}})$*,* $t_{\mathbb{D}} \approx b \in \mathcal{R}_{\mathcal{O}}^{\infty}$ *for some* $t_{\mathbb{D}} \in \mathbf{I}^{+} \setminus \mathbf{I}$.

*Proof.* We show the lemma via induction. More precisely, we verify that, for all rules $\rho = \beta \to \eta \in \mathcal{R}_{\mathcal{O}}$ and all substitutions $\sigma$ such that $\beta\sigma \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\eta\sigma = \mathbb{V}(t_{\mathbb{E}}, b)$, the lemma holds the fact $\eta\sigma$, provided that (IH) the lemma holds for every fact in $\beta\sigma$. We show that the this implication holds by distinguishing all possible cases based on the substitution $\sigma$ and the type of the rule $\rho$.

(4.3) Let $\sigma(y) = t_{\mathbb{E}}$, $\sigma(x) = b$, and $\mathbb{R}^{-} \sqcap \mathbb{S}^{-} = \mathbb{V}$. Then, $D(t_{\mathbb{E}}), \mathbb{R}^{-}(t_{\mathbb{E}}, b), C(b), \mathbb{S}(b, t_{\mathbb{C}}), D(t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By IH, $t_{\mathbb{D}} \approx b, \mathbb{R}^{-}(t_{\mathbb{E}}, t_{\mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ for some $t_{\mathbb{D}} \in \mathbf{I}^{+} \setminus \mathbf{I}$. By (Eq.3) and (Eq.4), $C(t_{\mathbb{D}}), \mathbb{S}(t_{\mathbb{D}}, t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.3), $\mathbb{R}^{-} \sqcap \mathbb{S}^{-}(t_{\mathbb{E}}, t_{\mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(6.1) Let $\sigma(x) = t_{\mathbb{E}}$, $\sigma(y) = b$, and $\mathbb{S} = \mathbb{V}$. Then, $R(t_{\mathbb{E}}, b) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By IH, $R(t_{\mathbb{E}}, t_{\mathbb{D}}), t_{\mathbb{D}} \approx b \in \mathcal{R}_{\mathcal{O}}^{\infty}$ for some $t_{\mathbb{D}} \in \mathbf{I}^{+} \setminus \mathbf{I}$. By (6.1), $\mathbb{S}(t_{\mathbb{E}}, t_{\mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(6.2) Analogous to the previous case.

(Role.1) Let $\sigma(y) = t_{\mathbb{E}}$, $\sigma(x) = b$, and $\mathbb{R}^- = \mathbb{V}$. Then, $\mathbb{R}(b, t_{\mathbb{E}}), \mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(Role.2) Analogous to case (6.1).

(Eq.4) Let $\sigma(z) = t_{\mathbb{E}}$, $\sigma(y) = b$, and $\mathbb{R} = \mathbb{V}$. Then, $\mathbb{V}(\sigma(x), b), \sigma(x) \approx t_{\mathbb{E}} \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By Lemma 6, $\mathsf{N}(t_{\mathbb{E}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

(Eq.5) Let $\sigma(x) = t_{\mathbb{E}}$, $\sigma(z) = b$, and $\mathbb{R} = \mathbb{V}$. Then, $\mathbb{V}(t_{\mathbb{E}}, \sigma(y)), \sigma(y) \approx b \in \mathcal{R}_{\mathcal{O}}^{\infty}$. We assume that $\sigma(y) \in \mathbf{I}$ as otherwise the lemma trivially holds. By IH, $\mathbb{V}(t_{\mathbb{E}}, t_{\mathbb{D}}), t_{\mathbb{D}} \approx \sigma(y) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ for some $t_{\mathbb{D}} \in \mathbf{I}^+ \setminus \mathbf{I}$. By (Eq.2), $t_{\mathbb{D}} \approx b \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

□

**Lemma 8.** *A domain element of the form $\delta.d_{\mathbb{S},\mathbb{C}} \in \Delta$ may not occur in $\Delta^{\mathcal{U}_{\mathcal{O}}}$ if*

1. *there is an axiom of the form $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$ with $R \in \mathbb{S}$, and*
2. *$C(\mathsf{i}(\delta)), R(\mathsf{i}(\delta), a), D(a), \mathsf{N}(a), D(t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ for some $a \in \mathbf{I}^+$.*

*Proof.* Suppose that $\delta.d_{\mathbb{S},\mathbb{C}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$ for a contradiction. Then, $\mathbb{S}(\mathsf{i}(\delta), t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 5 and $\{X(a) \mid X \in \mathbb{C}\} \cup \{R^-(a, \mathsf{i}(\delta))\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$ by (Role.1) and (4.3). By Lemma 5, $\{X(a) \mid X(t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$. Therefore, by Definition 5, $\delta.t_{\mathbb{S},\mathbb{C}} \notin \Delta^{\mathcal{U}_{\mathcal{O}}}$ if $\{S(\mathsf{i}(\delta), a) \mid S \in \mathbb{S}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$, which follows irrespectively of whether $\mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

- Let $\mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.3) and (Role.1), $(R^- \sqcap \mathbb{S}^-)(a, \mathsf{i}(\delta)), (R \sqcap \mathbb{S})(\mathsf{i}(\delta), a) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (Role.2), $\{S(\mathsf{i}(\delta), a) \mid S \in \mathbb{S}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$.
- Let $\mathsf{N}(\mathsf{i}(\delta)) \notin \mathcal{R}_{\mathcal{O}}^{\infty}$. Then, $\mathsf{i}(\delta)$ is of the form $t_{\mathbb{E}}$ by (Nm.1). Hence, by Lemma 7, there is some element $t_{\mathbb{D}} \in \mathbf{I}^+ \setminus \mathbf{I}$ such that $R(\mathsf{i}(\delta), t_{\mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$, and $t_{\mathbb{D}} = a$ or $t_{\mathbb{D}} \approx a \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.2), $\{(R \sqcap \mathbb{S})(\mathsf{i}(\delta), t_{\mathbb{C}\sqcap\mathbb{D}})\} \cup \{X(t_{\mathbb{C}\sqcap\mathbb{D}}) \mid X \in \mathbb{C} \sqcap \mathbb{D}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$. Hence, $t_{\mathbb{C}\sqcap\mathbb{D}} \approx a \in \mathcal{R}_{\mathcal{O}}^{\infty}$ or $t_{\mathbb{C}\sqcap\mathbb{D}} = a$ by (c) from Lemma 5. By (Eq.5) and (Role.2), $\{(R \sqcap \mathbb{S})(\mathsf{i}(\delta), a)\} \cup \{S(\mathsf{i}(\delta), a) \mid S \in \mathbb{S}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$.

□

**Lemma 9.** *A domain element of the form $\delta.d_{\mathbb{R},\mathbb{C}}.d_{\mathbb{S},\mathbb{D}} \in \Delta$ may not occur in $\Delta^{\mathcal{U}_{\mathcal{O}}}$ if*

1. *there is an axiom of the form $C \sqsubseteq \leqslant 1\, R.D \in \mathcal{O}$ with $R \in \mathbb{R}^-$ and $R \in \mathbb{S}$, and*
2. *$D(\mathsf{i}(\delta)), C(t_{\mathbb{C}}), D(\mathsf{i}(t_{\mathbb{D}})) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.*

*Proof.* Suppose that $\delta.d_{\mathbb{R},\mathbb{C}}.d_{\mathbb{S},\mathbb{D}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$ for a contradiction. Then, $\mathbb{R}(\mathsf{i}(\delta), t_{\mathbb{C}}))$, $\mathbb{S}(t_{\mathbb{C}}, t_{\mathbb{D}})) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 5. By (4.3) and (1) from Lemma 5, $\{\mathbb{R} \sqcap \mathbb{S}^-(\mathsf{i}(\delta), t_{\mathbb{C}})\} \cup \{X(\mathsf{i}(\delta)) \mid X(t_{\mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}\} \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Since $\delta.d_{\mathbb{R},\mathbb{C}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$, $\mathbb{R} = \mathbb{R} \sqcap \mathbb{S}^-$ by Definition 5. Therefore, $\delta.d_{\mathbb{R},\mathbb{C}}.d_{\mathbb{S},\mathbb{D}} \notin \mathcal{U}_{\mathcal{O}}$. □

**Lemma 10.** *If $\mathcal{R}_{\mathcal{O}} \not\models \exists x.\bot(x)$, then $\mathcal{U}_{\mathcal{O}} \models \mathcal{O}$.*

*Proof.* By Lemma 2, $\mathcal{U}_{\mathcal{O}}$ is an interpretation for $\mathcal{O}$. Therefore, to show that $\mathcal{U}_{\mathcal{O}}$ is a model of $\mathcal{O}$, it suffices to check that $\mathcal{U}_{\mathcal{O}}$ satisfies every axiom in $\mathcal{O}$. We verify this in the following case by case analysis, in which we consider all the different types of axioms (1)–(8) introduced in Figure 1.

(1) Let $\delta \in \bigcap_{i=1}^{n} C_i^{\mathcal{U}_{\mathcal{O}}}$. By Definition 6, $\{C_i(\mathsf{i}(\delta)) \mid 1 \leq in\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$. Hence, $D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by (1). By Definition 6, $\delta \in D^{\mathcal{U}_{\mathcal{O}}}$.

(3) Let $\langle \delta, \gamma \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$, and $\gamma \in C^{\mathcal{U}_{\mathcal{O}}}$. Then, $C(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 6. Also by Definition 6, some of the following cases hold.

- $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (3.1), $D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.
- $R^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)), \mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (Role.1) and (3.1), $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.
- $\gamma$ is of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ with $R \in \mathbb{R}$. Then, $\mathbb{R}(\mathsf{i}(\delta), \mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 5. By (Role.2) and (3.1), $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$.

- $\delta$ is of the form $\gamma.d_{\mathbb{R}^-,\mathbb{C}}$ with $R \in \mathbb{R}^-$. Then, $\mathbb{R}^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)) = \mathbb{R}^-(\mathsf{i}(\gamma), t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 5. Suppose for a contradiction that $D \notin \mathbb{C}$. Then, $\mathbb{R}^-(\mathsf{i}(\gamma), t_{\mathbb{C} \sqcap D}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\{X(t_{\mathbb{C} \sqcap D}) \mid X(t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}\} \subseteq \mathcal{R}_{\mathcal{O}}^{\infty}$ by (3.2) and Lemma 5. Therefore, $\delta = \gamma.d_{\mathbb{R}^-,\mathbb{C}} \notin \Delta^{\mathcal{U}_{\mathcal{O}}}$ by Definition 5. Since $\delta \in \Delta^{\mathcal{U}_{\mathcal{O}}}$, $D \in \mathbb{C}$. Hence, $D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by (2), (3.2), and (4.2).

  In either case, $D(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and hence, $\delta \in D^{\mathcal{U}_{\mathcal{O}}}$ by Definition 6.

(2) Let $\delta \in C^{\mathcal{U}_{\mathcal{O}}}$. Then, $C(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $R(\mathsf{i}(\delta), t_D), D(t_D) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by (2). Then, by Definition 5, some of the following cases hold.
- There is some $\delta.d_{\mathbb{R},\mathbb{C}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$ with $R \in \mathbb{R}$ and $D \in \mathbb{C}$. By Definition 5, $D(t_{\mathbb{C}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Therefore, $\langle \delta, \delta.d_{\mathbb{R},\mathbb{C}} \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$ and $\delta.d_{\mathbb{R},\mathbb{C}} \in D^{\mathcal{U}_{\mathcal{O}}}$ by Definition 6.
- $\delta$ is of the form $\gamma.d_{\mathbb{R}^-,\mathbb{C}}$, $R \in \mathbb{R}^-$, and $D(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Then, $\langle \delta, \gamma \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$ and $\gamma \in D^{\mathcal{U}_{\mathcal{O}}}$ by Definition 6.
- There is some $\gamma \in \Delta^{\mathcal{U}_{\mathcal{O}}}$ with $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), D(\mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Then, $\langle \delta, \gamma \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$ and $\gamma \in D^{\mathcal{U}_{\mathcal{O}}}$ by Definition 6.

(4) Let $\epsilon \in C^{\mathcal{U}_{\mathcal{O}}}$, and $\delta, \gamma \in D^{\mathcal{U}_{\mathcal{O}}}$ with $\langle \epsilon, \delta \rangle, \langle \epsilon, \gamma \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$. By Definition 6, $C(\mathsf{i}(\epsilon)), D(\mathsf{i}(\delta)), D(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$, and some of the following cases hold.
- $R(\mathsf{i}(\epsilon), \mathsf{i}(\delta)), \mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By ($\mathsf{Role}$.1), $R^-(\mathsf{i}(\delta), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Again, by Definition 6, some of the following cases hold.
  - $R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $R^-((\gamma), \mathsf{i}(\epsilon)), \mathsf{N}(\mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $\gamma$ is of the form $\epsilon.d_{\mathbb{S},\mathbb{C}}$ with $R \in \mathbb{S}$. By Lemma 8, this case results in a contradiction.
  - $\epsilon$ is of the form $\gamma.d_{\mathbb{S},\mathbb{C}}$ with $R^- \in \mathbb{S}$. Then, $\mathbb{S}(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by Definition 5. By ($\mathsf{Role}$.2), $R^-(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
- $R^-(\mathsf{i}(\delta), \mathsf{i}(\epsilon)), \mathsf{N}(\mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.4) and ($\mathsf{Role}$.1), $\mathsf{N}(\mathsf{i}(\delta)), R(\mathsf{i}(\epsilon), \mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Again, by Definition 6, some of the following cases hold.
  - $R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $R^-(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)), \mathsf{N}(\mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\gamma) \approx \mathsf{i}(\delta) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $\gamma$ is of the form $\epsilon.d_{\mathbb{S},\mathbb{C}}$. By Lemma 8, this case results in a contradiction.
  - $\epsilon$ is of the form $\gamma.d_{\mathbb{S},\mathbb{C}}$ with $R^- \in \mathbb{S}$. By Definition 5, $\mathbb{S}(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By ($\mathsf{Role}$.2) and (4.1), $R^-(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)), \mathsf{i}(\gamma) \approx \mathsf{i}(\delta) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Then, $\delta = \gamma$ by Lemma 6.
- $\delta$ is of the form $\epsilon.d_{\mathbb{R},\mathbb{C}}$ with $R \in \mathbb{R}$. By Definition 5, $\mathbb{R}(\mathsf{i}(\epsilon), \mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Again, by Definition 6, some of the following cases hold.
  - $R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By Lemma 8, this case results in a contradiction.
  - $R^-(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)), \mathsf{N}(\mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.4) and ($\mathsf{Role}$.1), $\mathsf{N}(\mathsf{i}(\gamma)), R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By Lemma 8, this case results in a contradiction.
  - $\gamma$ is of the form $\epsilon.d_{\mathbb{S},\mathbb{D}}$ with $R \in \mathbb{S}$. By Definition 5, $\mathbb{S}(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By the Definition of $\mathsf{i}(\cdot)$, $\mathsf{i}(\delta) = t_{\mathbb{C}}$ and $\mathsf{i}(\gamma) = t_{\mathbb{D}}$. By (4.2), $(\mathbb{R} \sqcap \mathbb{S})(\mathsf{i}(\epsilon), t_{\mathbb{C} \sqcap \mathbb{D}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (2b) from Definition 5, $\mathbb{R} \cup \mathbb{C} \not\subset \mathbb{R} \cup \mathbb{S} \cup \mathbb{C} \cup \mathbb{D} \not\subset$ and $\mathbb{S} \cup \mathbb{D} \not\subset \mathbb{R} \cup \mathbb{S} \cup \mathbb{C} \cup \mathbb{D}$. Therefore, $\mathbb{R} = \mathbb{S}$, $\mathbb{C} = \mathbb{D}$, and $\delta = \gamma$.
  - $\epsilon$ is of the form $\gamma.d_{\mathbb{S},\mathbb{D}}$ with $R^- \in \mathbb{S}$. By Lemma 9, this case results in a contradiction.
- $\epsilon$ is of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ with $R^- \in \mathbb{R}$. By Definition 5, $\mathbb{R}(\mathsf{i}(\delta), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By ($\mathsf{Role}$.2), $R^-(\mathsf{i}(\delta), \mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. Again, by Definition 6, some of the following cases hold.
  - $R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $R^-(\mathsf{i}(\gamma), \mathsf{i}(\epsilon)), \mathsf{N}(\mathsf{i}(\epsilon)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.4) and ($\mathsf{Role}$.1), $R(\mathsf{i}(\epsilon), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By (4.1), $\mathsf{i}(\delta) \approx \mathsf{i}(\gamma) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and $\delta = \gamma$ by Lemma 6.
  - $\gamma$ is of the form $\epsilon.d_{\mathbb{S},\mathbb{D}}$ with $R \in \mathbb{S}$. By Lemma 9, this case results in a contradiction.
  - $\epsilon$ is of the form $\gamma.d_{\mathbb{S},\mathbb{D}}$. Then, $\delta = \gamma$.

(5) Let $\delta \in C^{\mathcal{U}_{\mathcal{O}}}$. By Definition 6, $C(\mathsf{i}(\delta)) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and by (5), $\mathsf{i}(\delta) \approx a \in \mathcal{R}_{\mathcal{O}}^{\infty}$. By Definitions 5 and 6, $\mathsf{i}(a^{\mathcal{U}_{\mathcal{O}}}) \approx a \in \mathcal{R}_{\mathcal{O}}^{\infty}$ and hence, $\mathsf{i}(\delta) \approx \mathsf{i}(a^{\mathcal{U}_{\mathcal{O}}}) \in \mathcal{R}_{\mathcal{O}}^{\infty}$ by ($\mathsf{Eq}$.2) and ($\mathsf{Eq}$.3). Then, $\delta = \gamma$ by Lemma 6.

(6) Let $\langle \delta, \gamma \rangle \in R^{\mathcal{U}_\mathcal{O}}$. By Definition 6, some of the following cases holds.
- $R(\mathsf{i}(\delta), \mathsf{i}(\gamma)), \mathsf{N}(\mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$. By (6.1), $\mathbb{R}(\mathsf{i}(\delta), \mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$ with $\mathbb{S} = \{V \mid R \sqsubseteq_\mathcal{O}^* V\}$. By ($\mathsf{Role}$.2), $S(\mathsf{i}(\delta), \mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$.
- $R^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)), \mathsf{N}(\mathsf{i}(\delta)) \in \mathcal{R}_\mathcal{O}^\infty$. By (6.2), $\mathbb{R}^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)) \in \mathcal{R}_\mathcal{O}^\infty$ with $\mathbb{S} = \{S \mid V \sqsubseteq_\mathcal{O}^* V\}$. By ($\mathsf{Role}$.2), $S^-(\mathsf{i}(\gamma), \mathsf{i}(\delta)) \in \mathcal{R}_\mathcal{O}^\infty$.
- $\gamma$ is of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ and $R \in \mathbb{R}$. By Definition 5, $\mathbb{R}(\mathsf{i}(\delta), \mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$. Using an inductive argument on any chase sequence of $\mathcal{R}_\mathcal{O}$, we can show that, for all $V \in \mathbb{R}$, $\mathbb{R}$ contains every role $W \sqsupseteq_\mathcal{O}^* V$. Hence, $S \in \mathbb{R}$ and ($\mathsf{Role}$.2), $S(\mathsf{i}(\delta), \mathsf{i}(\gamma)) \in \mathcal{R}_\mathcal{O}^\infty$.
- $\delta$ is of the form $\gamma.d_{\mathbb{R}^-,\mathbb{C}}$ with $R \in \mathbb{R}^-$. Analogous to the previous case.

In either case, $\langle \delta, \gamma \rangle \in S^{\mathcal{U}_\mathcal{O}}$ by Definition 6.

(7) By (7), $C(a) \in \mathcal{R}_\mathcal{O}^\infty$. By Definitions 5 and 6, $\mathsf{i}(a^{\mathcal{U}_\mathcal{O}}) \approx a \in \mathcal{R}_\mathcal{O}^\infty$. By ($\mathsf{Eq}$.2) and ($\mathsf{Eq}$.4), $C(\mathsf{i}(a^{\mathcal{U}_\mathcal{O}})) \in \mathcal{R}_\mathcal{O}^\infty$. Hence, $a^{\mathcal{U}_\mathcal{O}} \in C^{\mathcal{U}_\mathcal{O}}$ by Definition 6.

(8) By (8), $R(a, b) \in \mathcal{R}_\mathcal{O}^\infty$. By Definitions 5 and 6, $\mathsf{i}(a^{\mathcal{U}_\mathcal{O}}) \approx a, \mathsf{i}(b^{\mathcal{U}_\mathcal{O}}) \approx b \in \mathcal{R}_\mathcal{O}^\infty$. Since $\mathcal{R}_{\mathsf{Eq}} \subseteq \mathcal{R}_\mathcal{O}$, $R(\mathsf{i}(a^{\mathcal{U}_\mathcal{O}}), \mathsf{i}(b^{\mathcal{U}_\mathcal{O}})) \in \mathcal{R}_\mathcal{O}^\infty$. By ($\mathsf{Nom}$.1), $\mathsf{N}(b) \in \mathcal{R}_\mathcal{O}^\infty$ and therefore, $\mathsf{N}(\mathsf{i}(b^{\mathcal{U}_\mathcal{O}})) \in \mathcal{R}_\mathcal{O}^\infty$ by ($\mathsf{Eq}$.2) and ($\mathsf{Eq}$.4). By Definition 6, $\langle a^{\mathcal{U}_\mathcal{O}}, b^{\mathcal{U}_\mathcal{O}} \rangle \in R^{\mathcal{U}_\mathcal{O}}$.

□

## Proof of Lemma 12

We start by extending the function $\mathsf{Ex}(\cdot)$ introduced in the proof of Lemma 1 to $\mathbf{I}^\times$ and showing a preliminary result, later used in the proof of Lemma 12.

**Definition 9.** *For all $a \in \mathbf{I}^\times$, let $\mathsf{Ex}(a) = \mathbb{C}$ if $a$ is of the form $t_\mathbb{C}$ or $t_{\mathbb{R},\mathbb{C}}^i$, and $\mathsf{Ex}(a) = \{a\}$ otherwise (i.e., if $a \in \mathbf{I}$).*

**Lemma 11.** *For any model $\mathcal{I}$ of $\mathcal{O}$, we have that:*
1. *If $C(a) \in \mathcal{C}_\mathcal{O}$ with $C \in \mathbf{C}$, then $\mathsf{Ex}(a)^\mathcal{I} \subseteq C^\mathcal{I}$.*
2. *If $a, b \in \mathbf{I}^\times$ occur $\mathcal{C}_\mathcal{O}$, for every $\delta \in \mathsf{Ex}(a)^\mathcal{I}$ there is some $\gamma \in \mathsf{Ex}(b)^\mathcal{I}$ with $\langle \delta, \gamma \rangle \in R^\mathcal{I}$ for all $R_\triangleright(a, b) \in \mathcal{C}_\mathcal{O}$.*
3. *If some $a \in \mathbf{I}^\times$ occurs in $\mathcal{C}_\mathcal{O}$, then $\mathsf{Ex}(a)^\mathcal{I} \neq \emptyset$.*
4. *If $\mathsf{N}(a) \in \mathcal{C}_\mathcal{O}$, then $|\mathsf{Ex}(a)^\mathcal{I}| = 1$.*

*Proof.* In the remainder of the argument, we use (a)-(d) to refer to the corresponding claims from the proof of Lemma 1.
1. Let $C(a) \in \mathcal{C}_\mathcal{O}$. By Definition 7, if $a$ is of the form $t_{\mathbb{R},\mathbb{C}}^i$, then $C(t_\mathbb{C}) \in \mathcal{R}_\mathcal{O}^\infty$. Otherwise, $C(a) \in \mathcal{R}_\mathcal{O}^\infty$. Either way, $\mathsf{Ex}(a)^\mathcal{I} \subseteq C^\mathcal{I}$ by (a).
2. If there is not any $R \in \mathbf{R}$ with $R_\triangleright(a, b) \notin \mathcal{C}_\mathcal{O}$, then the claim holds since $\mathsf{Ex}(b)^\mathcal{I} \neq \emptyset$ by (3). To show that (2) also holds otherwise, we study the following cases:
   - Let $\mathsf{N}(a), \mathsf{N}(b) \in \mathcal{C}_\mathcal{O}$. By Definition 7 and ($\mathsf{Role}$.1), for each $R_\triangleright(a, b) \in \mathcal{C}_\mathcal{O}$, there is some $\mathbb{R}(a, b) \in \mathcal{R}_\mathcal{O}^\infty$ with $R \in \mathbb{R}$. By (b) and (e), for all $\mathbb{R}(a, b) \in \mathcal{R}_\mathcal{O}^\infty$, we have that $\langle \delta, \gamma \rangle \in \mathbb{R}^\mathcal{I}$ with $\delta$ and $\gamma$ the only elements in $\mathsf{Ex}(a)^\mathcal{I}$ and $\mathsf{Ex}(b)^\mathcal{I}$, respectively.
   - $\mathsf{N}(a) \in \mathcal{C}_\mathcal{O}$ and $\mathsf{N}(b) \notin \mathcal{C}_\mathcal{O}$. Then, $b$ is of the form $t_{\mathbb{R},\mathbb{C}}^i$ by Definition 7, and there is a single element $\delta \in \mathsf{Ex}(a)^\mathcal{I}$ by (e). We consider two different cases:
   - Let $\mathbb{R}(a, t_\mathbb{C}) \in \mathcal{R}_\mathcal{O}^\infty$. By Definition 7, for every $R_\triangleright(a, t_{\mathbb{R},\mathbb{C}}^i) \in \mathcal{C}_\mathcal{O}$ we have that either $R \in \mathbb{R}$, or there is some $\mathbb{S}(t_\mathbb{C}, a) \in \mathcal{R}_\mathcal{O}^\infty$ with $R \in \mathbb{S}^-$. By (b), $\langle \delta, \gamma \rangle \in \mathbb{R}^\mathcal{I}$ for some $\gamma \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} = \mathsf{Ex}(t_{\mathbb{R},\mathbb{C}}^i)^\mathcal{I}$. Moreover, for all $\epsilon \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} = \mathsf{Ex}(t_{\mathbb{R},\mathbb{C}}^i)^\mathcal{I}$, $\langle \epsilon, \delta \rangle \in \mathbb{S}^\mathcal{I}$ for all $\mathbb{S}(t_\mathbb{C}, a) \in \mathcal{R}_\mathcal{O}^\infty$ and hence, $\langle \delta, \gamma \rangle \in (\mathbb{S}^-)^\mathcal{I}$ for all $\mathbb{S}(t_\mathbb{C}, a) \in \mathcal{R}_\mathcal{O}^\infty$.
   - Let $\mathbb{R}(a, t_\mathbb{C}) \notin \mathcal{R}_\mathcal{O}^\infty$. By Definition 7, for every $R_\triangleright(a, t_{\mathbb{R},\mathbb{C}}^i) \in \mathcal{C}_\mathcal{O}$ there is some $\mathbb{S}(t_\mathbb{C}, a) \in \mathcal{R}_\mathcal{O}^\infty$ with $R \in \mathbb{S}^-$. By (b), for all $\gamma \in \mathsf{Ex}(t_\mathbb{C})^\mathcal{I} = \mathsf{Ex}(t_{\mathbb{R},\mathbb{C}}^i)^\mathcal{I}$, we have that $\langle \delta, \gamma \rangle \in (\mathbb{S}^-)^\mathcal{I}$ for all $\mathbb{S}(t_\mathbb{C}, a) \in \mathcal{R}_\mathcal{O}^\infty$.

- Let $\mathsf{N}(a) \notin \mathcal{R}^{\infty}_{\mathcal{O}}$, and $\mathsf{N}(b) \in \mathcal{R}^{\infty}_{\mathcal{O}}$. By Definition 7, $a$ is of the form $t^{i}_{\mathbb{R},\mathbb{C}}$. Moreover, for every $R_{\triangleright}(t^{i}_{\mathbb{R},\mathbb{C}}, b) \in \mathcal{C}_{\mathcal{O}}$ there is some $\mathbb{S}(t_{\mathbb{C}}, b) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ with $R \in \mathbb{S}$. By (b) and (e), for all $\delta \in \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} = \mathsf{Ex}(t^{i}_{\mathbb{R},\mathbb{C}})^{\mathcal{I}}$ and all $\mathbb{S}(t_{\mathbb{C}}, b) \in \mathcal{R}^{\infty}_{\mathcal{O}}$, $\langle \delta, \gamma \rangle \in \mathbb{S}^{\mathcal{I}}$ with $\gamma$ the only element in $\mathsf{Ex}(b)^{\mathcal{I}}$.
- Let $\mathsf{N}(a) \notin \mathcal{C}_{\mathcal{O}}$ and $\mathsf{N}(b) \notin \mathcal{C}_{\mathcal{O}}$. By Definition 7, $a$ is of the form $t^{i}_{\mathbb{R},\mathbb{C}}$, and $b$ is of the form $t^{j}_{\mathbb{S},\mathbb{D}}$. Moreover, $\mathbb{S}(t_{\mathbb{C}}, t_{\mathbb{D}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$, and $R \in \mathbb{S}$ for all $R_{\triangleright}(a, b) \in \mathcal{C}_{\mathcal{O}}$. By (b), for all $\delta \in \mathsf{Ex}(b)^{\mathcal{I}} = \mathsf{Ex}(t_{\mathbb{C}})^{\mathcal{I}} = \mathsf{Ex}(t^{i}_{\mathbb{R},\mathbb{C}})^{\mathcal{I}}$ there is some $\gamma \in \mathsf{Ex}(b)^{\mathcal{I}} = \mathsf{Ex}(t_{\mathbb{D}})^{\mathcal{I}} = \mathsf{Ex}(t^{j}_{\mathbb{S},\mathbb{D}})^{\mathcal{I}}$ with $\langle \delta, \gamma \rangle \in \mathbb{S}^{\mathcal{I}}$.

3. Let $a \in \mathbf{I}^{\times}$ be some individual occurring in $\mathcal{C}_{\mathcal{O}}$. By Definition 7, if $a$ is of the form $t^{i}_{\mathbb{R},\mathbb{C}}$, then $t_{\mathbb{C}}$ occurs in $\mathcal{R}^{\infty}_{\mathcal{O}}$. Otherwise, $a$ occurs in $\mathcal{R}^{\infty}_{\mathcal{O}}$. Either way, $\mathsf{Ex}(a)^{\mathcal{I}} \neq \emptyset$ by (d).
4. Let $\mathsf{N}(a) \in \mathcal{C}_{\mathcal{O}}$. By Definition 7, if $a$ is of the form $t^{i}_{\mathbb{R},\mathbb{C}}$, then $\mathsf{N}(t_{\mathbb{C}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$. Otherwise, $\mathsf{N}(a) \in \mathcal{R}^{\infty}_{\mathcal{O}}$. Either way, $|\mathsf{Ex}(a)^{\mathcal{I}}| = 1$ by (e).

$\square$

We proceed to show that our BCQ entailment algorithm is sound.

**Lemma 12.** *If there is a valid match for an ontology $\mathcal{O}$ and a query $q$, then $\mathcal{O} \models q$.*

*Proof.* We assume that there is a valid match $\sigma$ for $\mathcal{O}$ and $q = \exists \vec{x}.\beta$ as otherwise the lemma trivially holds. By Definition 8, $\sigma$ is a substitution such that the graph $F_{q,\sigma}$ is a forest, and $\beta\sigma \subseteq \mathcal{C}_{\mathcal{O}}$.

To prove the lemma, we show that, $\mathcal{I} \models q'$ for any (arbitrarily chosen) model $\mathcal{I}$ of $\mathcal{O}$ with $q' = \exists \vec{x}.\beta'$ a normalised query such that $q' \models q$. This normalised query $q'$ is obtained by exhaustively applying the following transformation rules to $q$:

- If $v(x) = v(y)$ and $x \neq y$ for some $x, y \in \vec{x}$, replace all occurrences of $x$ in $q$ with $y$.
- If $R(x, y) \in \beta$, then add $R^{-}(y, x) \in \beta$.

Since we only merge two variables $x, y \in \vec{x}$ if $v(x) = v(y)$, the graph $F_{q',\sigma}$ is a forest. Moreover, since $v(x) = v(y)$ implies $\sigma(x) = \sigma(y)$ by Definition 8, and $R(a, b) \in \mathcal{C}_{\mathcal{O}}$ iff $R^{-}(b, a) \in \mathcal{C}_{\mathcal{O}}$ by Definition 7; $\beta'\sigma \subseteq \mathcal{C}_{\mathcal{O}}$. By the definition of $q'$, if $v(y) \to v(x), v(z) \to v(x) \in F_{q',\sigma}$ for some $x, y, z \in \vec{x}$, then we can show by contradiction that $y = z$. To identify the unique predecessor variable we introduce the partial function $\mathsf{pr}$: For $x \in \vec{x}$, $\mathsf{pr}(x) = y$ if there is some $y$ with $v(y) \to v(x) \in F_{q',\sigma}$, and $\mathsf{pr}(x)$ is undefined otherwise.

To show $\mathcal{I} \models q'$, we define a sequence $q_1, \ldots, q_n$ of BCQs with $q_n = q'$ and then verify via induction that $\mathcal{I} \models q_i$ for all $i \in \{1, \ldots, n\}$. Let $x_1, \ldots, x_n$ be some sequence containing all the variables in $\vec{x}$ such that $v(x_j) \to v(x_i) \notin F_{q',\sigma}$ for all $1 \leq i < j \leq n$ (note that, there is at least one such sequence since $F_{q',\sigma}$ is a forest). Let $q_1 = \exists x_1.\beta_1, \ldots, q_n = \exists x_1, \ldots, x_n.\beta_n$ be the sequence of BCQs such that, for all $i \in \{1, \ldots, n\}$, $\beta_i$ is the conjunction containing an atom $\alpha$ in the body of $q'$ if all the variables occurring in $\alpha$ also occur in the sequence $x_1, \ldots, x_i$.

To show that $\mathcal{I} \models q_i$ for all $i \in \{1, \ldots, n\}$ we introduce an assignment $Z$ for $\mathcal{I}$ such that, for all $x \in \vec{x}$, $Z(x)$ is a domain element satisfying all of the following:

1. If $\mathsf{N}(\sigma(x)) \in \mathcal{C}_{\mathcal{O}}$, then $Z(x)$ is the only element in $\mathsf{Ex}(\sigma(x_i))^{\mathcal{I}}$. Note that, $|\mathsf{Ex}(\sigma(x))^{\mathcal{I}}| = 1$ by (4) from Lemma 11.
2. If $\mathsf{N}(\sigma(x)) \notin \mathcal{C}_{\mathcal{O}}$ and $v(x)$ does not have a predecessor in $F_{q,\sigma}$, then $Z(x)$ is some (arbitrarily chosen) element in $\mathsf{Ex}(\sigma(x))^{\mathcal{I}}$. Note that, $|\mathsf{Ex}(\sigma(x))^{\mathcal{I}}| \neq \emptyset$ by (3) from Lemma 11.
3. If $\mathsf{N}(\sigma(x)) \notin \mathcal{C}_{\mathcal{O}}$ and $v(x)$ has a predecessor in $F_{q,\sigma}$, then $Z(x)$ is some domain element with $\langle Z(\mathsf{pr}(x)), Z(x) \rangle \in R^{\mathcal{I}}$ for all $R(\mathsf{pr}(x), x) \in \beta_i$ and $Z(x) \in \mathsf{Ex}(\sigma(x))^{\mathcal{I}}$.

Note that, the assignment $Z$ is well defined because $F_{q',\sigma}$ is acyclic.

We show via induction that there must be a domain element $Z(x)$ satisfying the conditions established in case (3), as this is not immediately obvious. I.e., we show that there is a domain element such as $Z(x)$ assuming that (IH) *there is a domain element* $Z(\mathsf{pr}(x))$. Since $\beta'\sigma \subseteq \mathcal{C}_\mathcal{O}$, $R(\sigma(\mathsf{pr}(x)), \sigma(x)) \in \mathcal{C}_\mathcal{O}$ for all $R(\mathsf{pr}(x), x) \in \beta_i$. Hence, $R_\triangleright(\sigma(\mathsf{pr}(x)), \sigma(x)) \in \mathcal{C}_\mathcal{O}$ or $R^-_\triangleright(\sigma(x), \sigma(\mathsf{pr}(x))) \in \mathcal{C}_\mathcal{O}$ by Definition 7. Suppose for a contradiction that $S_\triangleright(\sigma(\mathsf{pr}(x)), \sigma(x)) \notin \mathcal{C}_\mathcal{O}$ for some $S(\mathsf{pr}(x), x) \in \beta_i$. Then, $S^-_\triangleright(\sigma(x), \sigma(\mathsf{pr}(x))) \in \mathcal{C}_\mathcal{O}$, $v(x) \to v(\mathsf{pr}(x)) \in F_{q',\sigma}$ by Definition 8, and the graph $F_{q',\sigma}$ contains a cycle. Theorefore, we conclude that, $R_\triangleright(\sigma(\mathsf{pr}(x)), \sigma(x)) \in \mathcal{C}_\mathcal{O}$ for all $R(\mathsf{pr}(x), x) \in \beta_i$. By the definition of $Z$, $Z(\mathsf{pr}(x)) \in \mathsf{Ex}(\sigma(\mathsf{pr}(x)))^\mathcal{I}$ and by (2) from Lemma 11, there is some $\gamma \in \mathsf{Ex}(\sigma(x))^\mathcal{I}$ with $\langle Z(\mathsf{pr}(x)), \gamma\rangle \in R^\mathcal{I}$ for all $R(\mathsf{pr}(x), x) \in \beta_i$.

Given some $i = 1, \ldots, n$, we show via induction that $\mathcal{I}, Z \models q_i$, i.e., we verify that $\langle Z(\vec{y})\rangle \in P^\mathcal{I}$ for all $P(\vec{y}) \in \beta_i$. If $i \geq 2$ we assume that (IH) $\mathcal{I}, Z \models q_{i-1}$.

- Let $P(\vec{y}) \in \beta_i$ with $y \neq x_i$ for all $y \in \vec{y}$. Then, $\langle Z(\vec{y})\rangle \in P^\mathcal{I}$ by IH. Note that, $\beta_1$ may not contain atoms of this form and hence, in this case we do not need to apply the IH.
- Let $C(x_i) \in \beta_i$. Since $\beta'\sigma \subseteq \mathcal{C}_\mathcal{O}$, $C(\sigma(x_i)) \in \mathcal{C}_\mathcal{O}$. By (1) from Lemma 11, $\mathsf{Ex}(\sigma(x_i))^\mathcal{I} \subseteq C^\mathcal{I}$. Since $Z(x_i) \in \mathsf{Ex}(\sigma(x_i))^\mathcal{I}$, $Z(x_i) \in C^\mathcal{I}$.
- Let $R(y, x_i) \in \beta_i$. We study the following different cases:
  - Let $v(y) \to \mathsf{pr}(x_i) \in F_{q',\sigma}$. Then, $\mathsf{pr}(x) = y$, and $\langle Z(y), Z(x_i)\rangle \in R^\mathcal{I}$ by the definition of $Z$.
  - $v(y) \to \mathsf{pr}(x_i) \notin F_{q',\sigma}$. Since $R(\sigma(y), \sigma(x_i)) \in \mathcal{C}_\mathcal{O}$, $R_\triangleright(\sigma(y), \sigma(x_i)) \in \mathcal{C}_\mathcal{O}$ or $R^-_\triangleright(\sigma(x_i), \sigma(y)) \in \mathcal{C}_\mathcal{O}$ by Definition 7. Moreover, $v(x_i) \to v(y) \notin F_{q,\sigma}$ by the definition of $q_i$. Therefore, $R_\triangleright(\sigma(x_i), \sigma(y)), R^-_\triangleright(\sigma(y), \sigma(x_i)) \in \mathcal{C}_\mathcal{O}$ by Definition 8. By Definition 7, we either have that $\sigma(y) \in \mathbf{I}^+$ or $\sigma(x_i) \in \mathbf{I}^+$, and hence, $\mathsf{N}(\sigma(y)) \in \mathcal{C}_\mathcal{O}$ or $\mathsf{N}(\sigma(x_i)) \in \mathcal{C}_\mathcal{O}$ . We consider two cases:
    - ∗ Let $\mathsf{N}(x_i) \in \mathcal{C}_\mathcal{O}$. By the definition of $Z$, $Z(x_i)$ is the only element $\mathsf{Ex}(\sigma(x_i))^\mathcal{I}$ and $Z(y) \in \mathsf{Ex}(\sigma(y))^\mathcal{I}$. By (2) from Lemma 11, $\langle Z(y), Z(x_i)\rangle \in R^\mathcal{I}$.
    - ∗ Let $\mathsf{N}(y) \in \mathcal{C}_\mathcal{O}$. By the definition of $Z$, $Z(x_i) \in \mathsf{Ex}(\sigma(x_i))^\mathcal{I}$ and $Z(y)$ is the only element $\mathsf{Ex}(\sigma(y))^\mathcal{I}$. By (2) from Lemma 11, $\langle Z(y), Z(x_i)\rangle \in R^\mathcal{I}$.
- Let $R(x_i, y) \in \beta_i$. Then, $R^-(y, x_i) \in \beta_i$ by the definition of the query $q'$ and, as shown in the previous case, $\langle Z(y), Z(x_i)\rangle \in (R^-)^\mathcal{I}$. Therefore, $\langle Z(x_i), Z(y)\rangle \in R^\mathcal{I}$.

□

## Proof of Lemma 13

**Lemma 13.** *If there are not any valid matches for* $\mathcal{O}$ *and* $q$*, then* $\mathcal{O} \not\models q$*.*

*Proof.* We show the following claim, from which the lemma follows: $(*)$ If $\mathcal{U}_\mathcal{O}$ entails $q = \exists\vec{x}.\beta$, then there is a valid match for $\mathcal{O}$ and $q$. By $(*)$, if there are not any valid matches for $\mathcal{O}$ and $q$, then $\mathcal{U}_\mathcal{O}$ does not entail $q$. In turn, this implies that $\mathcal{O}$ does not entail $q$ since $\mathcal{U}_\mathcal{O}$ is a model of $\mathcal{O}$ by Lemma 3.

To show $(*)$, we prove that, if there is some assignment $Z$ with $Z(\vec{y}) \in P^{\mathcal{U}_\mathcal{O}}$ for all $P(\vec{y}) \in \beta$, one can construct a valid match $\sigma$ for $q$ and $\mathcal{O}$. Let $\sigma$ be the substitution mapping the variables in $\vec{x}$ to individuals in $\mathbf{I}^\times$ occurring in $\mathcal{C}_\mathcal{O}$ such that, for all $x \in \vec{x}$, all of the following hold:

- Let $Z(x)$ be an equivalence class. Then, $\sigma(x) = \mathsf{i}(Z(x))$ with $\mathsf{i}(\cdot)$ the function introduced in Definition 4.
- Let $Z(x)$ be some element of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ such that (1) $Z(y) \neq \delta$ for all $y \in \vec{x}$, or (2) $Z(y) = \delta$ for some $y \in \vec{x}$ with $Z(y)$ an equivalence class. Then, $\sigma(x) = t^0_{\mathbb{R},\mathbb{C}}$.
- Let $Z(x)$ be an element of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ such that $Z(y) = \delta$ for some $y \in \vec{x}$ with $\delta$ an element of the form $\gamma.d_{\mathbb{S},\mathbb{D}}$. Then, $\sigma(y)$ is of the form $t^i_{\mathbb{S},\mathbb{D}}$ by the definition of $\sigma$; and $\sigma(x) = t^{(i+1) \mod 2}_{\mathbb{R},\mathbb{C}}$ if $\langle\mathbb{S},\mathbb{D}\rangle \preceq \langle\mathbb{R},\mathbb{C}\rangle$, and $\sigma(x) = t^i_{\mathbb{R},\mathbb{C}}$ otherwise.

Note that, $\sigma$ is well defined since $\sigma(x) = \sigma(y)$ for all $y, z \in \vec{x}$ with $Z(y) = Z(z)$.

We show claim $(*)$ via induction; namely, in the following paragraph, we introduce a sequence of queries $q_1, \ldots, q_n$ with $q_n \equiv q$ and then show via induction that, for all $i \in \{1, \ldots, n\}$, $\sigma$ is a valid match for $q_i$.

Let $\mathsf{d} : \Delta \to \mathbb{N}$ such that, for all $\delta \in \Delta$, $\mathsf{d}(\delta) = 1$ if $\delta$ is an equivalence class, and $\mathsf{d}(\delta) = \mathsf{d}(\gamma) + 1$ if $\delta$ is of the form $\gamma.d_{\mathbb{R},\mathbb{C}}$. We extend $\mathsf{d}$ to atoms such that, given some $P(x_1, \ldots, x_n) \in \beta$, $\mathsf{d}(P(x_1, \ldots, x_n)) = \mathit{max}(\mathsf{d}(Z(x_1)), \ldots, \mathsf{d}(Z(x_n)))$. Let $q_0 = \exists \vec{x}.\beta_0$ with $\beta_0$ the empty conjunction and let $\alpha'_1, \ldots, \alpha'_n$ be some sequence containing all the atoms in $\beta$ such that $\mathsf{d}(\alpha'_i) \geq \mathsf{d}(\alpha'_j)$ for all $1 \leq i < j \leq n$. Moreover, let $\alpha_1, \ldots, \alpha_n$ be the sequence such that, for all $i \in \{1, \ldots, n\}$, $\alpha_i = R^-(y, x)$ if $\alpha_i = R(x, y)$ with $\mathsf{d}(Z(x)) > \mathsf{d}(Z(y))$, and $\alpha_i = \alpha'_i$ otherwise. Then, for all $i \in \{1, \ldots, n\}$, $q_i = \exists \vec{x}.\beta_i$ with $\beta_i = \alpha_1 \wedge \ldots \wedge \alpha_i$.

Before our inductive argument, we state and prove some auxiliary results.

**Lemma 14.** *If $v(x) \to v(y) \in G_{\sigma,q_i}$ for some $i \in \{0, \ldots, n\}$, then $\mathsf{N}(\sigma(y)) \notin \mathcal{C}_{\mathcal{O}}$.*

*Proof.* Suppose for a contradiction that $\mathsf{N}(\sigma(y)) \in \mathcal{C}_{\mathcal{O}}$. Since $v(x) \to v(y) \in G_{\sigma,q_i}$, for some $R \in \mathbf{R}$, (1) $R(x, y) \in \beta$ or $R^-(y, x) \in \beta$, (2) $R_{\triangleright}(\sigma(x), \sigma(y)) \in \mathcal{C}_{\mathcal{O}}$, and (3) $R^-_{\triangleright}(\sigma(y), \sigma(x)) \notin \mathcal{C}_{\mathcal{O}}$. We show that assuming $\mathsf{N}(\sigma(y)) \in \mathcal{C}_{\mathcal{O}}$ results in a contradiction is implied irrespectively of whether $\mathsf{N}(\sigma(x)) \in \mathcal{C}_{\mathcal{O}}$:

- $\mathsf{N}(\sigma(x)) \in \mathcal{C}_{\mathcal{O}}$. Since $R_{\triangleright}(\sigma(x), \sigma(y)) \in \mathcal{C}_{\mathcal{O}}$, $R(\sigma(x), \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ or $R^-(\sigma(y), \sigma(x)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$. By Definition 7, $R^-_{\triangleright}(\sigma(y), \sigma(x)) \in \mathcal{C}_{\mathcal{O}}$.
- $\mathsf{N}(\sigma(x)) \notin \mathcal{C}_{\mathcal{O}}$. Then, $\sigma(x)$ is of the form $t^i_{\mathbb{R},\mathbb{C}}$. Since $R_{\triangleright}(\sigma(x), \sigma(y)) \in \mathcal{C}_{\mathcal{O}}$, $R(t_{\mathbb{C}}, \sigma(y)) \in \mathcal{R}^{\infty}_{\mathcal{O}}$. By Definition 7, $R^-_{\triangleright}(\sigma(y), \sigma(x)) \in \mathcal{C}_{\mathcal{O}}$.

□

**Lemma 15.** *If $v(x) \to v(y) \in G_{\sigma,q_i}$ for some $i \in \{0, \ldots, n\}$, then $Z(y)$ is of the form $Z(x).d_{\mathbb{R},\mathbb{C}}$.*

*Proof.* By Lemma 14, $\mathsf{N}(\sigma(y)) \notin \mathcal{C}_{\mathcal{O}}$. Hence $Z(y)$ is of the form $\delta.d_{\mathbb{S},\mathbb{D}}$ by Definition 5, and $\sigma(y) = t^i_{\mathbb{S},\mathbb{D}}$ for some $i \in \{0, 1, 2\}$ by definition of $\sigma$. Since $v(x) \to v(y) \in G_{\sigma,q_i}$, there is some $R \in \mathbf{R}$, (1) $R(x, y) \in \beta$ or $R^-(y, x) \in \beta$, (2) $R_{\triangleright}(\sigma(x), t^i_{\mathbb{S},\mathbb{D}}) \in \mathcal{C}_{\mathcal{O}}$, and (3) $R^-_{\triangleright}(t^i_{\mathbb{S},\mathbb{D}}, \sigma(x)) \notin \mathcal{C}_{\mathcal{O}}$. Since $\langle Z(x), Z(y) \rangle \in R^{\mathcal{U}_{\mathcal{O}}}$, one of the following cases must hold by Definition 6.

1. There is some $\mathbb{V} \in \mathbf{R}^{\sqcap}$ such that $\mathbb{V}(t_{\mathbb{D}}, \mathsf{i}(Z(x))), \mathsf{N}(\mathsf{i}(Z(x))) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ and $R^- \in \mathbb{V}$. Then, $\sigma(x) = \mathsf{i}(Z(x))$, and $R^-_{\triangleright}(t^i_{\mathbb{S},\mathbb{D}}, \mathsf{i}(Z(x))) \in \mathcal{C}_{\mathcal{O}}$ by Definition 7. Hence, this case entails a contradiction (cf. (4.c)).
2. $Z(y)$ is of the form $Z(x).d_{\mathbb{S},\mathbb{D}}$ and $R \in \mathbb{S}$. Since $Z(x).d_{\mathbb{S},\mathbb{D}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$, $\mathbb{S}(\mathsf{i}(Z(x)), \mathsf{i}(Z(y))) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by Definition 5 and the lemma does hold in this case.
3. $Z(x)$ is of the form $Z(y).d_{\mathbb{V},\mathbb{E}}$ for some $\mathbb{V} \in \mathbf{R}$ with $R^- \in \mathbb{V}$. Then, $Z(x).d_{\mathbb{V},\mathbb{E}} = \delta.d_{\mathbb{S},\mathbb{D}}.d_{\mathbb{V},\mathbb{E}} \in \Delta^{\mathcal{U}_{\mathcal{O}}}$, and hence, $\mathbb{V}(t_{\mathbb{D}}, t_{\mathbb{E}}) \in \mathcal{R}^{\infty}_{\mathcal{O}}$ by Definition 5. Moreover, $\sigma(x)$ is of the form $t^j_{\mathbb{V},\mathbb{E}}$ by the definition of $\sigma$. Two possible cases arise:
   - $\langle \mathbb{V}, \mathbb{E} \rangle \preceq \langle \mathbb{S}, \mathbb{D} \rangle$. Then, $R^-_{\triangleright}(t^j_{\mathbb{S},\mathbb{D}}, t^i_{\mathbb{V},\mathbb{E}}) \in \mathcal{C}_{\mathcal{O}}$ with $j = i \mod 2$.
   - $\langle \mathbb{S}, \mathbb{D} \rangle \not\preceq \langle \mathbb{V}, \mathbb{E} \rangle$. Then, $R^-_{\triangleright}(t^j_{\mathbb{S},\mathbb{D}}, t^i_{\mathbb{V},\mathbb{E}}) \in \mathcal{C}_{\mathcal{O}}$ with $j = i$.

   In either case, we have that $R_{\triangleright}(t^j_{\mathbb{V},\mathbb{E}}, t^i_{\mathbb{S},\mathbb{D}}), R^-_{\triangleright}(t^i_{\mathbb{S},\mathbb{D}}, t^j_{\mathbb{V},\mathbb{E}}) \in \mathcal{C}_{\mathcal{O}}$, which contradicts Definition 7.

Since (2) is the only case that does not entail a contradiction, the lemma follows.

□

**Lemma 16.** *If $v(x) \to v(y) \in F_{\sigma,q_i}$ for some $i \in \{0, \ldots, n\}$, then $Z(y)$ is of the form $Z(x).d_{\mathbb{R},\mathbb{C}}$.*

*Proof.* If $v(x) \to v(y) \in F_{\sigma,q_i}$, then there are some sequences of vertices $v(x_1),\ldots,v(x_n)$ and $v(y_2),\ldots,v(y_n)$ in $G_{\sigma,q_i}$ such that,

- for all $2 \leq i \leq n$, $v(x_{i-1}) \to v(x_i) \in G_{\sigma,q_i}$ and $\sigma(x_i) = \sigma(y_i)$;
- for all $3 \leq i \leq n$, $v(y_{i-1}) \to v(y_i) \in G_{\sigma,q_i}$; and
- $v(x) = v(x_1)$, $v(y) = v(y_2)$, and $v(x_n) = v(y_n)$.

Since $v(x_{n-1}) \to v(x_n) \in G_{\sigma,q_i}$, we conclude that $Z(x_n)$ is of the form $Z(x_{n-1}).d_{\mathbb{R},\mathbb{C}}$ and $\mathbb{R}(\mathsf{i}(Z(x_{n-1})),\mathsf{i}(Z(x_n))) \in \mathcal{R}^\infty_\mathcal{O}$ by Lemma 15. Moreover, since $v(y_{n-1}) \to v(x_n) \in G_{\sigma,q_i}$, $Z(x_n)$ is of the form $Z(y_{n-1}).d_{\mathbb{S},\mathbb{D}}$ and $\mathbb{S}(\mathsf{i}(Z(y_{n-1})),\mathsf{i}(Z(x_n))) \in \mathcal{R}^\infty_\mathcal{O}$ again by Lemma 15. Therefore, $Z(x_{n-1}) = Z(y_{n-1})$. Using an analogous argument, we can conclude that $Z(x_i) = Z(y_i)$ for all $2 \leq i \leq n-2$. We can make such an argument as we have that $v(x_i) = v(y_i)$ for all $2 \leq i \leq n$ by the definition of $F_{\sigma,q_i}$. Finally, by Lemma 15, $Z(y)$ is of the form $Z(x).d_{\mathbb{R},\mathbb{C}}$ and $\mathbb{R}(\mathsf{i}(Z(x)),\mathsf{i}(Z(y))) \in \mathcal{R}^\infty_\mathcal{O}$. ☐

We show via induction that, for all $i \in \{0,\ldots,n\}$, $\sigma$ is a valid match for $\mathcal{O}$ and $q_i$. The base case is trivial, since $q_0$ is a query with an empty body. To show the induction step we check that, for all $i \in \{1,\ldots,n\}$, $\sigma$ is a valid match for $q_i = \exists\vec{x}.\beta_i = \exists\vec{x}.\beta_{i-1} \wedge \alpha_i$ provided (IH) *$\sigma$ is a valid match for $q_{i-1} = \exists\vec{x}.\beta_{i-1}$*. In the following case by case analysis we show that the induction step holds, irrespectively of whether $\alpha_i$ is a unary or binary atom, and of which type of elements occur in the range of $Z$. For every case, we have show that (a) $\beta_i\sigma \subseteq \mathcal{C}_\mathcal{O}$ and (b) $F_{q_i,\sigma}$ is a forest. Since $F_{q_{i-1},\sigma}$ is a forest and $\beta_{i-1}\sigma \subseteq \mathcal{C}_\mathcal{O}$ by IH, (a) follows if $\alpha_i\sigma \in \mathcal{C}_\mathcal{O}$ and (b) follows if $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.

1. *Let $\alpha_i = C(z)$ with $Z(z)$ an equivalence class.*
   a. Then, $\sigma(z) = \mathsf{i}(Z(z))$. Since $Z(z) \in C^{\mathcal{U}_\mathcal{O}}$, $C(\mathsf{i}(Z(z))) \in \mathcal{R}^\infty_\mathcal{O}$ by Definition 6. Hence, $C(\mathsf{i}(Z(z))) \in \mathcal{C}_\mathcal{O}$ by Definition 7.
   b. In this case, $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.
2. *Let $\alpha_i = C(z)$ with $Z(z)$ a domain element of the form $\delta.d_{\mathbb{R},\mathbb{C}}$.*
   a. Then, $\sigma(z)$ is of the form $t^j_{\mathbb{R},\mathbb{C}}$. Since $Z(z) \in C^{\mathcal{U}_\mathcal{O}}$, $C(t_\mathbb{C}) \in \mathcal{R}^\infty_\mathcal{O}$ by Definition 6. Hence, $C(t^j_{\mathbb{R},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$ by Definition 7.
   b. In this case, $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.
3. *Let $\alpha_i$ be of the form $R(x,y)$ such that $Z(x)$ and $Z(y)$ are equivalence classes.*
   a. Then, $\sigma(x) = \mathsf{i}(Z(x))$ and $\sigma(y) = \mathsf{i}(Z(y))$, and $\mathsf{N}(\mathsf{i}(Z(x))), \mathsf{N}(\mathsf{i}(Z(y))) \in \mathcal{R}^\infty_\mathcal{O}$. Since $\langle Z(x), Z(y)\rangle \in R^{\mathcal{U}_\mathcal{O}}$, there is some $\mathbb{R} \subseteq \mathbf{R}$ with $R \in \mathbb{R}$, and $\mathbb{R}(\mathsf{i}(Z(x)),\mathsf{i}(Z(y))) \in \mathcal{R}^\infty_\mathcal{O}$ or $\mathbb{R}^-(\mathsf{i}(Z(y)),\mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$. Hence, $R(\mathsf{i}(Z(x)),\mathsf{i}(Z(y))) \in \mathcal{C}_\mathcal{O}$.
   b. By Lemma 16, $v(x) \to v(y) \notin F_{q,\sigma}$. Hence, $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.
4. *Let $\alpha_i = R(x,y)$, $Z(y)$ is a domain element of the form $Z(x).d_{\mathbb{R},\mathbb{C}}$, and $R \in \mathbb{R}$.*
   a. We consider two possible cases:
      - $Z(x)$ is an equivalence class. Then, $\sigma(x) = \mathsf{i}(Z(x))$ and $\sigma(y) = t^0_{\mathbb{R},\mathbb{C}}$. Since $Z(y) \in \Delta^{\mathcal{U}_\mathcal{O}}$, $\mathbb{R}(\mathsf{i}(Z(x)), t_\mathbb{C}) \in \mathcal{R}^\infty_\mathcal{O}$ and hence, $R(\mathsf{i}(Z(x)), t^0_{\mathbb{R},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$.
      - $Z(x)$ is of the form $\gamma.d_{\mathbb{S},\mathbb{D}}$. Then, $\sigma(x)$ is of the form $t^i_{\mathbb{S},\mathbb{D}}$ and $\sigma(y)$ is of the form $t^j_{\mathbb{R},\mathbb{C}}$. Since $Z(y) \in \Delta^{\mathcal{U}_\mathcal{O}}$, $\mathbb{R}(t_\mathbb{D}, t_\mathbb{C}) \in \mathcal{R}^\infty_\mathcal{O}$. By the definition of $\sigma$, $j = (i+1) \bmod 2$ if $\langle\mathbb{S},\mathbb{D}\rangle \preceq \langle\mathbb{R},\mathbb{C}\rangle$, and $j = i$ if $\langle\mathbb{S},\mathbb{D}\rangle \not\preceq \langle\mathbb{R},\mathbb{C}\rangle$. In either case, we have that $\mathbb{R}(t^i_{\mathbb{S},\mathbb{D}}, t^j_{\mathbb{R},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$
   b. Let us assume that $F_{\sigma,q_i} \neq F_{\sigma,q_{i-1}}$ as otherwise $\sigma$ is a valid match for $q_i$ and $\mathcal{O}$ by IH. By the definition of $q_i$ and Lemma 16, for all $z \in \vec{x}$, $v(z) \to v(x) \notin F_{\sigma,q_i}$. If this was not the case, then $\beta_i$ would contain a binary atom of the form $\alpha = S(z,x)$ with $\mathsf{d}(\alpha) < \mathsf{d}(R(x,y))$—the occurrence of such an atom in $\beta_i$ would contradict the definition of $q_i$. Therefore, $F_{\sigma,q_i} = \{v(x) \to v(y)\} \cup F_{\sigma,q_{i-1}}$ is not a rooted forest if and only if there is some vertex $v(z)$ with $v(z) \to v(y)$ and $v(z) \neq v(x)$. Suppose for a contradiction that such a vertex $v(z)$ exists. By Lemma 16, $Z(y)$ is of the form

$Z(z).d_{\mathbb{S},\mathbb{D}}$. Hence, $\mathbb{R} = \mathbb{S}$, $\mathbb{C} = \mathbb{D}$, and $Z(z) = Z(x)$; and $\sigma(v) = \sigma(x)$ by the definition of $\sigma$. By the definition of $F_{\sigma,q_i}$, $v(v) = v(x)$.

5. *Let $\alpha_i = R(x,y)$, $Z(y)$ is a domain element of the form $Z(x).d_{\mathbb{R},\mathbb{C}}$, and $R \notin \mathbb{R}$.*

(a) Since $\langle Z(x), Z(y)\rangle \in R^{\mathcal{U}_\mathcal{O}}$ and $R \notin \mathbb{R}$, $Z(x)$ is an equivalence class (and hence, $\mathsf{N}(\mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$), and there is some $\mathbb{S}(t_\mathbb{C}, \mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$ with $R^- \in \mathbb{S}$. Then, $\sigma(x) = \mathsf{i}(Z(x))$ and $\sigma(y)$ is of the form $t^j_{\mathbb{R},\mathbb{C}}$. By Definition 7, $R(\mathsf{i}(Z(x)), t^j_{\mathbb{R},\mathbb{C}}) \in \mathcal{C}_\mathcal{O}$.

(b) Since $\mathbb{S}(t_\mathbb{C}, \mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$ with $R^- \in \mathbb{S}$, $R^-(t^j_{\mathbb{R},\mathbb{C}}, \mathsf{i}(Z(x))) \in \mathcal{C}_\mathcal{O}$. Therefore, either $v(x) \to v(y) \in F_{q_{i-1},\sigma}$ or $v(x) \to v(y) \notin F_{q_i,\sigma}$. Either way, $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.

6. *Let $\alpha_i = R(x,y)$ with $Z(y)$ a domain element of the form $\delta.d_{\mathbb{R},\mathbb{C}}$ and $Z(x) \neq \delta$.*

(a) Then, $Z(x)$ is an equivalence class, $\sigma(x) = \mathsf{i}(Z(x))$, and $\sigma(y)$ is of the form $t^j_{\mathbb{R},\mathbb{C}}$. Moreover, $\mathsf{N}(\mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$, and there is some $\mathbb{S}(t_\mathbb{C}, \mathsf{i}(Z(x))) \in \mathcal{R}^\infty_\mathcal{O}$ with $R^- \in \mathbb{S}$. Therefore, $R(t^j_{\mathbb{R},\mathbb{C}}, \mathsf{i}(Z(x))) \in \mathcal{C}_\mathcal{O}$.

(b) By Lemma 16, $v(x) \to v(y) \notin F_{q,\sigma}$. Hence, $F_{q_{i-1},\sigma} = F_{q_i,\sigma}$.

By the definition of $q_i$, $\mathsf{d}(x) \leq \mathsf{d}(y)$ for all $R(x,y) \in \beta_i$, and hence, we do not need to consider binary atoms in the above case by case analysis for which this is not the case. □