Partially Finite Model Reasoning in Description Logics

Tomasz Gogacz[1], **Filip Murlak**[1], **Marcin Przybyłko**[1], **Alexandra Rogova**[1] and **Michał Skrzypczak**[1]
[1]University of Warsaw
{t.gogacz, fmurlak, m.przybylko, rogova, mskrzypczak}@mimuw.edu.pl

Abstract

Aiming to harmonise finite and infinite model reasoning, we initiate the study of *partially finite models*, where the reasoning task comes with a formula that specifies a part of the model that must be finite. We focus on the problem of *partially finite query entailment* in description logics (DLs): given a knowledge base (KB), a query, and a distinguished concept, decide whether the query holds in all models of the KB that interpret the distinguished concept as a finite set. To break the ground, we work with the DL $\mathcal{S}$, an extension of the basic DL $\mathcal{ALC}$ with transitive roles, which is one of the simplest cases where finite and infinite query entailment diverge. Generalising previous results on the finite and infinite cases, we show that also partially finite entailment of conjunctive queries is in 2-ExpTime for $\mathcal{S}$. The solution involves sophisticated infinite model surgery and goes far beyond combining the arguments for the two special cases. As a direct application, we show how the problem of query containment in the presence of closed predicates can be solved by reduction to partially finite query entailment.

1 Introduction

Query answering is one of the fundamental tasks in knowledge-base related reasoning, where a query language – often, *conjunctive queries* – is enhanced by an ontology – a set of inference rules. The ontology augments the query by facilitating the unification of heterogeneous data, allowing reconstruction of incomplete data, or allowing domain knowledge injection (Calì *et al.* 2012; Calvanese *et al.* 2009).

Classically, in ontology-mediated data access one asks for so-called *certain answers*, that is, answers valid in every, possibly infinite, model of the knowledge base. This *infinite model reasoning* mode has been the subject of a rich and intense study for many ontology languages, including existential rules and description logics, e.g. (Calì *et al.* 2009; Calvanese *et al.* 2012; Pérez-Urbina *et al.* 2008; Amendola *et al.* 2018b). In the alternative mode of *finite model reasoning*, inspired by applications in database theory, only finite models are considered; see (Murlak 2023) for an overview.

For many ontology and query languages the above modes of query entailment coincide, which is known as *finite controllability*. Some languages, however, are not finitely controllable even with respect to conjunctive queries: two well-known examples among DLs are $\mathcal{S}$ and $\mathcal{ALCIF}$. For an in-depth study see, e.g., (Amendola *et al.* 2018a; Bednarczyk and Kieronski 2022; Figueira *et al.* 2020; Gogacz and Marcinkowski 2017).

In this work we explore a hybrid mode of *partially finite model reasoning* (Gogacz *et al.* 2020a), which allows only models that interpret a specified concept F as a finite set of individuals. This is a common generalisation of both classical reasoning modes that allows applying different modes in different parts of the knowledge base. It could be useful in federated systems comprising both classical databases and knowledge-based components or in the context of partial materialisation strategies for the infinite reasoning mode (Qin *et al.* 2021). More generally, it is a powerful modelling tool, allowing the knowledge engineer to decide which parts of the model should be finite and which can be infinite. The distinction might simply reflect reality: for instance, while the future of a reactive system is infinite, its past is finite. But it might also be a pragmatic choice: even if a concept is always finite in reality, correct axiomatisation might be harder when finiteness is taken into account. Under the partially finite reasoning mode, the knowledge engineer can decide which finite parts of the world should be modelled accurately as finite, and for which the cruder unrestricted reasoning mode should be used, allowing for simpler axiomatisation at the cost of possibly missing inferences.

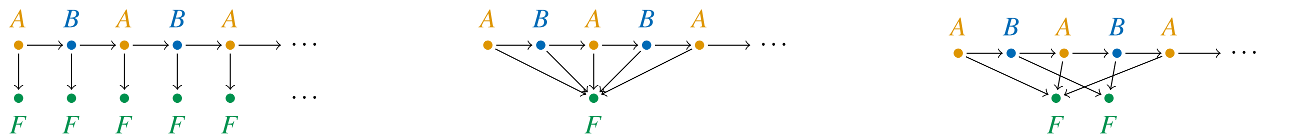
Challenges. Consider the knowledge base consisting of the ABox $\{A(a)\}$ and the TBox

$$\{A \sqsubseteq \exists r.B,\ B \sqsubseteq \exists r.A,\ A \sqcup B \sqsubseteq \exists r.F\}\,.$$

In every model of this knowledge base, there is an infinite chain of alternating A-, and B-labelled nodes, each linked to an F-labelled node. A natural infinite model for this TBox is represented in Figure 1a. In the partially finite reasoning mode, however, we might be required to ensure that there are only finitely many F-labelled nodes. There are several ways to make the model conform to this new requirement. The most obvious choice, shown in Figure 1b, is to merge all F-labelled nodes into one and redirect all edges to this single representative. This, however, might not always be a good solution. Consider the query

$$\exists x \exists y \exists z\ r(x,y) \wedge r(x,z) \wedge r(y,z)\,.$$

which checks if some node and its successor share a common successor. The model in Figure 1a shows that this query



(a) Unrestricted (infinitely many F nodes). (b) Partially finite (single F node). (c) Partially finite (two F nodes).

Figure 1: Unrestricted and partially finite interpretations.

is not entailed over infinite models. But the model in Figure 1b does satisfy the query. Is there one that does not? Instead of merging all F-nodes in the model from Figure 1a into one, separately merge those connected to an A-labelled node and those connected to a B-labelled node, as shown in Figure 1c. The resulting model does not satisfy the query. Hence, we can conclude that the query is not entailed.

The partially finite model in Figure 1c is relatively simple, but it is tailored to a specific query – unlike the unrestricted model in Figure 1a, which can serve as a universal countermodel for all conjunctive queries that are not entailed. For

$$\exists x \exists y \exists z \exists w \; r(x,y) \wedge r(y,z) \wedge r(x,w) \wedge r(z,w)$$

we would need a different partially finite countermodel. And indeed, no partially finite model can serve as a countermodel for all conjunctive queries that are not entailed.

Contribution. We begin the investigation of partially finite entailment by exploring this problem for the logic $\mathcal{S}$, a description logic that extends $\mathcal{ALC}$ with role transitivity. As a relatively small and syntactically simple logic without finite controllability, it is a prime candidate for this initial study. As our main contribution, we show that for $\mathcal{S}$ partially finite entailment of conjunctive queries is 2-ExpTime-complete, matching the complexities of both finite (Ibáñez-García *et al.* 2026) and infinite entailment (Eiter *et al.* 2009; Glimm *et al.* 2008; Glimm *et al.* 2007). Additionally, to illustrate the usefulness of partially finite entailment, we apply it to solve query containment in the presence of closed predicates.

We start the presentation with the necessary preliminaries in Section 2. In Section 3, we describe a solution for a single transitive role. In Section 4, we develop a model-theoretic tool that allows us to solve the general problem in Section 5. In Section 6, we discuss the connection to query containment with closed predicates. Finally, we conclude with a discussion of potential future work in Section 7. An appendix with full proofs can be found in the full version of the paper (Gogacz *et al.* 2026).

Related Work. Partially finite variants of entailment and satisfiability were introduced by Gogacz *et al.* (2020a) in the context of ontology focusing, but only satisfiability (dubbed MIXED-SAT) was studied. It was shown to be ExpTime-complete for $\mathcal{ALCHIF}$ and NExpTime-complete for $\mathcal{ALCHOIF}$. Improved complexity bounds were obtained for logics from the DL-Lite family.

In the 1980's and 1990's, a line of work investigated *circumscription* (McCarthy 1980; Bonatti *et al.* 2009), and *negation as failure rules* (Gelfond and Przymusinska 1986; Donini *et al.* 2002; Katsuno 1990) as alternatives to the big divide of Closed vs Open World Assumptions. In a *circumscribed knowledge base*, a set of predicates (or concepts) must be minimal (but not necessarily finite), whereas in a language with *negation as failure rules*, a set of predicates can be marked to be interpreted as false unless provably true. Gelfond *et al.* (1989), show that these two approaches can be made equivalent under some natural constraints. More recently, combinations of the classical DL open-world semantics and closed-world rules have been considered (Lifschitz 1991; Motik and Rosati 2010; Knorr *et al.* 2011): unlike partially finite entailment, they suffer from increased complexity or even undecidability.

Partially finite entailment can be seen as a relaxation of entailment with *closed predicates*, where considered models might extend the interpretation of a closed predicate but must keep it finite. The complexity of this problem has been studied in the context of description logics, ranging from DL-Lite and $\mathcal{EL}$ (Lutz *et al.* 2019; Andresel *et al.* 2020; Ngo *et al.* 2016), to very expressive DLs like $\mathcal{ALCHI}$ (Bajraktari *et al.* 2018; Lutz *et al.* 2019) and $\mathcal{ALCHQIO}$ (Ahmetaj *et al.* 2016; Lukumbuzya *et al.* 2020), as well as existential rules (Gogacz *et al.* 2020b; Marconi and Rosati 2025; Benedikt and Bourhis 2018; Bienvenu and Bourhis 2019).

2 Preliminaries

We fix disjoint infinite countable sets $\mathtt{N_C}$ of concept names, $\mathtt{N_R}$ of role names, $\mathtt{N_I}$ of individuals, and $\mathtt{N_V}$ of variables.

The Description Logic $\mathcal{S}$. We consider the description logic $\mathcal{S}$ which extends $\mathcal{ALC}$ with role transitivity. We assume that $\mathtt{N_R}$ is split into the sets $\mathtt{N_R^t}$ and $\mathtt{N_R^{nt}}$ of *transitive* and *non-transitive role names*. $\mathcal{S}$ *concepts*, usually denoted by C or D, are then defined just like for $\mathcal{ALC}$ by the grammar

$$C, D := \bot \mid \top \mid A \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists r.C \mid \forall r.C$$

where A ranges over concept names from $\mathtt{N_C}$ and r ranges over role names from $\mathtt{N_R}$. A *concept inclusion* is a formula of the shape $C \sqsubseteq D$ where C and D are $\mathcal{S}$ concepts. A TBox is a finite set of concept inclusions.

Without loss of generality we can assume that TBoxes are normalised and only use concept inclusion of the forms

$$\textstyle\bigsqcap_i A_i \sqsubseteq \bigsqcup_j B_j\,, \qquad A \sqsubseteq \forall r.B\,, \qquad A \sqsubseteq \exists r.B\,,$$

with empty conjunction and disjunction treated as $\top$ and $\bot$ (Gutiérrez-Basulto *et al.* 2023).

An ABox is a finite set of concept assertions of the form $A(a)$ and role assertions of the form $r(a,b)$ where A is a concept name, r is a role name and $\{a,b\}$ are individuals.

Interpretations. An interpretation $\mathcal{I}$ consists of a (possibly infinite) domain $\Delta^{\mathcal{I}}$ and a function $\cdot^{\mathcal{I}}$ that maps each concept name A to $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and each role name r to $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. We extend $\cdot^{\mathcal{I}}$ to $\mathcal{S}$ concepts as follows:

$$\bot^{\mathcal{I}} = \emptyset, \quad \top^{\mathcal{I}} = \Delta^{\mathcal{I}}, \quad (\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}},$$
$$(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}, \quad (C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}},$$
$$(\exists r.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \exists b.\ (a,b) \in r^{\mathcal{I}} \wedge b \in C^{\mathcal{I}}\},$$
$$(\forall r.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \forall b.\ (a,b) \in r^{\mathcal{I}} \implies b \in C^{\mathcal{I}}\}.$$

We call a role $r \subseteq \Delta \times \Delta$ *transitive* if for all $a,b,c \in \Delta$, $(a,b) \in r$ and $(b,c) \in r$ implies $(a,c) \in r$. The *transitive closure* of $r \subseteq \Delta \times \Delta$ (with respect to Δ), written as r^* is the least transitive $r' \subseteq \Delta \times \Delta$ such that $r \subseteq r'$. An interpretation $\mathcal{I}$ is *transitive* if $t^{\mathcal{I}}$ is transitive for all $t \in \mathtt{N}_{\mathtt{R}}^{\mathtt{t}}$. The *transitive closure* of an interpretation $\mathcal{I}$, written as $\mathcal{I}^*$ is the interpretation such that $\Delta^{\mathcal{I}^*} = \Delta^{\mathcal{I}}$, $A^{\mathcal{I}^*} = A^{\mathcal{I}}$, $r^{\mathcal{I}^*} = r^{\mathcal{I}}$, and $t^{\mathcal{I}^*} = \left(t^{\mathcal{I}}\right)^*$ for all $A \in \mathtt{N}_{\mathtt{C}}$, $r \in \mathtt{N}_{\mathtt{R}}^{\mathtt{nt}}$, and $t \in \mathtt{N}_{\mathtt{R}}^{\mathtt{t}}$.

An interpretation $\mathcal{I}$ *satisfies*: a concept inclusion $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$; a concept assertion $A(a)$ if $a \in A^{\mathcal{I}}$; a role assertion $r(a,b)$ if $(a,b) \in r^{\mathcal{I}}$; a TBox $\mathcal{T}$, written as $\mathcal{I} \models \mathcal{T}$, if it is transitive and satisfies each concept inclusion in $\mathcal{T}$; and an ABox $\mathcal{A}$, written as $\mathcal{I} \models \mathcal{A}$, if it satisfies each assertion in $\mathcal{A}$. (Note that we adopt the Standard Name Assumption; in particular, $\Delta^{\mathcal{I}}$ includes all individuals used in $\mathcal{A}$.)

An interpretation $\mathcal{I}$ is a *subinterpretation* of an interpretation $\mathcal{J}$, written as $\mathcal{I} \subseteq \mathcal{J}$, if $\Delta^{\mathcal{I}} \subseteq \Delta^{\mathcal{J}}$ and $A^{\mathcal{I}} \subseteq A^{\mathcal{J}}$, $r^{\mathcal{I}} \subseteq r^{\mathcal{J}}$ for all $A \in \mathtt{N}_{\mathtt{C}}$, and $r \in \mathtt{N}_{\mathtt{R}}$. A subinterpretation $\mathcal{I}$ of $\mathcal{J}$ is *induced* by $\Delta \subseteq \Delta^{\mathcal{J}}$, written as $\mathcal{I} = \mathcal{J} \restriction \Delta$, if $\Delta^{\mathcal{I}} = \Delta$, $A^{\mathcal{I}} = A^{\mathcal{J}} \cap \Delta$ and $r^{\mathcal{I}} = r^{\mathcal{J}} \cap \Delta \times \Delta$ for all $A \in \mathtt{N}_{\mathtt{C}}$, and $r \in \mathtt{N}_{\mathtt{R}}$. For $\Sigma \subseteq \mathtt{N}_{\mathtt{R}}$, an interpretation $\mathcal{I}$ is *over* Σ if $r^{\mathcal{I}} = \emptyset$ for all $r \notin \Sigma$. If $\Sigma = \{r\}$ we say simply that $\mathcal{I}$ is over r. A concept name A *occurs* in $\mathcal{I}$ if $A^{\mathcal{I}} \neq \emptyset$. We consider only interpretations in which finitely many concept names occur.

A homomorphism h from $\mathcal{I}$ to $\mathcal{J}$, written as $h : \mathcal{I} \to \mathcal{J}$ is a function from $\Delta^{\mathcal{I}}$ to $\Delta^{\mathcal{J}}$ such that $e \in A^{\mathcal{I}}$ implies $h(e) \in A^{\mathcal{J}}$, $(e,e') \in r^{\mathcal{I}}$ implies $(h(e),h(e')) \in r^{\mathcal{J}}$, and $h(a) = a$ for all $e,e' \in \Delta^{\mathcal{I}}$, $A \in \mathtt{N}_{\mathtt{C}}$, $r \in \mathtt{N}_{\mathtt{R}}$, $a \in \mathtt{N}_{\mathtt{I}}$. By $\mathrm{rg}(h) \subseteq \Delta^{\mathcal{J}}$ we denote the set of values of h, i.e., $\{h(e) \mid e \in \Delta^{\mathcal{I}}\}$.

An interpretation $\mathcal{I}$ can be seen as directed multigraph $G_{\mathcal{I}}$ whose nodes are the elements of $\mathcal{I}$ and edges are obtained as the disjoint union of the interpretations of role names in $\mathcal{I}$. Nodes of $G_{\mathcal{I}}$ are labelled with sets of concept names and edges are labelled with role names (note that parallel edges have different labels). Throughout the paper we apply the standard graph-theoretic terminology directly to interpretations. In particular, we will speak of reachability and *strongly-connected components* (SCCs) of interpretations, which are maximal sets of elements in which every element is reachable from every other element.

Knowledge Bases. A knowledge base (KB) $\mathcal{K}$ is a pair $(\mathcal{T},\mathcal{A})$ composed of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. An interpretation $\mathcal{I}$ is a *model* of a KB $\mathcal{K} = (\mathcal{T},\mathcal{A})$, written as $\mathcal{I} \models \mathcal{K}$, if $\mathcal{I} \models \mathcal{T}$, $\mathcal{I} \models \mathcal{A}$, and $\mathcal{I}$ is transitive.

We write $\|\mathcal{K}\|$ for the total size of concept inclusions, transitivity declarations, and assertions in $\mathcal{K}$. By $\mathtt{CN}(\mathcal{K})$, $\mathtt{Rol}(\mathcal{K})$, and $\mathtt{Ind}(\mathcal{K})$, we denote the sets of all concept names, role names and individuals that appear in $\mathcal{K}$. $\mathcal{K}$ is *over* $\Sigma \subseteq \mathtt{N}_{\mathtt{R}}$ if $\mathtt{Rol}(\mathcal{K}) \subseteq \Sigma$. We extend this notation and terminology in the natural way to TBoxes and ABoxes.

A *unary type* is a set of concept names. A unary type over $\Gamma \subseteq \mathtt{N}_{\mathtt{C}}$ is a set of concept names from Γ. For a KB $\mathcal{K}$, we write $\mathtt{Tp}(\mathcal{K})$ for the set of all unary types over $\mathtt{CN}(\mathcal{K})$. We write $\mathtt{tp}^{\mathcal{I}}(e)$ for the *type of element e in interpretation $\mathcal{I}$*, defined as $\mathtt{tp}^{\mathcal{I}}(e) = \{A \in \mathtt{N}_{\mathtt{C}} \mid e \in A^{\mathcal{I}}\}$.

Conjunctive Queries. We consider *conjunctive queries* (CQs), i.e. queries of the form $\exists \bar{x}.\ q_1(\bar{y}_1) \wedge q_2(\bar{y}_2) \wedge \ldots \wedge q_n(\bar{y}_n)$ where $\bar{x}$ is a tuple of variables, $\bar{y}_1, \bar{y}_2, \ldots \bar{y}_n$ are tuples of variables from $\bar{x}$, and each q_i is either a unary atom of the form $A(x)$ for some $A \in \mathtt{N}_{\mathtt{C}}$ and $x \in \mathtt{N}_{\mathtt{V}}$ or a binary atom of the form $r(x,x')$ for some $r \in \mathtt{N}_{\mathtt{R}}$ and $x,x' \in \mathtt{N}_{\mathtt{V}}$.

A *match* of a CQ q in an interpretation $\mathcal{I}$ is a function π that maps each variable in q to an element of $\Delta^{\mathcal{I}}$ such that $\pi(x) \in A^{\mathcal{I}}$ for each atom in q of the shape $A(x)$ and $(\pi(x),\pi(x')) \in r^{\mathcal{I}}$ for each atom in q of the shape $r(x,x')$. We say that an interpretation $\mathcal{I}$ satisfies a CQ q, written as $\mathcal{I} \models q$, if there is a match of q in $\mathcal{I}$.

A *union of conjunctive queries* (UCQ) is a finite set Q of CQs. An interpretation $\mathcal{I}$ satisfies Q, written as $\mathcal{I} \models Q$, if $\mathcal{I} \models q$ for some $q \in Q$. It is well known that if $\mathcal{I} \models Q$ and there is a homomorphism from $\mathcal{I}$ to $\mathcal{J}$, then $\mathcal{J} \models Q$.

A query Q is *over* $\Sigma \subseteq \mathtt{N}_{\mathtt{R}}$ if uses role names only from Σ.

Partially Finite Entailment. A KB $\mathcal{K}$ *entails* a query Q, written as $\mathcal{K} \models Q$, if every model of $\mathcal{K}$ satisfies Q. A *counter-model* is a model of $\mathcal{K}$ that does not satisfy Q. Similarly, $\mathcal{K}$ *finitely entails* Q, written as $\mathcal{K} \models_{\mathit{fin}} Q$, if every *finite* model of $\mathcal{K}$ satisfies Q. We focus on a common generalisation of these problems, where the implication must hold for models where a specified concept is finite. Given a concept F, we say that $\mathcal{K}$ *F-entails* Q, written $\mathcal{K} \models_F Q$, if every model $\mathcal{I}$ of $\mathcal{K}$ with $F^{\mathcal{I}}$ finite, satisfies Q. For simplicity, we assume F is a concept name.

PARTIALLY FINITE ENTAILMENT	
Input:	A KB $\mathcal{K}$, a concept name F, and a query Q
Output:	True iff $\mathcal{K} \models_F Q$

Throughout the paper we assume that F is the distinguished concept to be interpreted as a finite set. We call an element e of interpretation $\mathcal{I}$ *critical* if $e \in F^{\mathcal{I}}$.

3 Single Transitive Role

We first deal with the case of a single transitive role t: throughout the section we assume that the interpretations of all remaining role names are empty. To make this subcase easier to use in the general solution, we work with rooted and labelled interpretations.

An *X-labelled interpretation* $(\mathcal{I},\lambda)$ is an interpretation $\mathcal{I}$ along with a labelling function $\lambda : \Delta^{\mathcal{I}} \to X$ that assigns to each element of $\mathcal{I}$ a label from X. The set X will vary, depending on the context. A homomorphism $h\colon (\mathcal{I},\lambda) \to (\mathcal{I}',\lambda')$ is a homomorphism $h\colon \mathcal{I} \to \mathcal{I}'$ such that $\lambda(a) = \lambda'(h(a))$ for all $a \in \Delta^{\mathcal{I}}$.

A *rooted interpretation* $(\mathcal{I}, a)$ is an interpretation $\mathcal{I}$ along with a distinguished element $a \in \Delta^{\mathcal{I}}$ called the *root*. A *homomorphism* $h\colon (\mathcal{I}, a) \to (\mathcal{I}', a')$ is a homomorphism $h\colon \mathcal{I} \to \mathcal{I}'$ such that $h(a) = a'$.

A *rooted X-labelled interpretation* is $(\mathcal{I}, a, \lambda)$ where $(\mathcal{I}, a)$ is a rooted interpretation and $(\mathcal{I}, \lambda)$ is an X-labelled interpretation. A homomorphism $h\colon (\mathcal{I}, a, \lambda) \to (\mathcal{I}', a', \lambda')$ is a homomorphism $h\colon \mathcal{I} \to \mathcal{I}'$ such that both $h\colon (\mathcal{I}, a) \to (\mathcal{I}', a')$ and $h\colon (\mathcal{I}, \lambda) \to (\mathcal{I}', \lambda')$ are homomorphisms.

This additional labelling in X will be used inside our inductive bottom-up construction to store the already-simplified parts of our interpretation as external values (labels), allowing us to focus on the currently modified part.

Intuitively, we would like to replace a given (rooted X-labelled) interpretation with a simpler, more regular one, while preserving certain properties. Eventually, we will care about satisfying a given KB and not satisfying a given query. For now, we use an abstract sufficient condition, formulated purely in terms of interpretations.

For an interpretation $\mathcal{I}$, an element $u \in \Delta^{\mathcal{I}}$, and a transitive role name t, we write $\mathtt{rch}_t^{\mathcal{I}}(u)$ for the set of concept names reachable from u, i.e. all A such that $A^{\mathcal{I}} \setminus \{u\}$ contains a node t-reachable from u.

A homomorphism $h\colon \mathcal{I} \to \mathcal{I}'$ is *t-strong* if

$$\mathtt{tp}^{\mathcal{I}}(a) = \mathtt{tp}^{\mathcal{I}'}(h(a)) \quad \text{and} \quad \mathtt{rch}_t^{\mathcal{I}}(a) = \mathtt{rch}_t^{\mathcal{I}'}(h(a))$$

for all $a \in \Delta^{\mathcal{I}}$. Clearly, the composition of two t-strong homomorphisms is a t-strong homomorphism.

Lemma 1. *Consider a TBox $\mathcal{T}$ over t, a UCQ Q, and transitive interpretations $\mathcal{I}$ and $\mathcal{I}'$ such that there is a t-strong homomorphism $h\colon \mathcal{I} \to \mathcal{I}'$.*

- *If $\mathcal{I}' \models \mathcal{T}$ and $\mathcal{I}' \not\models Q$, then $\mathcal{I} \models \mathcal{T}$ and $\mathcal{I} \not\models Q$.*
- *If h is surjective and $\mathcal{I} \models \mathcal{T}$, then $\mathcal{I}' \models \mathcal{T}$.*

The notion of t-strong homomorphism lifts naturally to rooted and X-labelled interpretations: we call a homomorphism $h\colon (\mathcal{I}, a, \lambda) \to (\mathcal{I}', a', \lambda')$ t-strong if the underlying homomorphism from $\mathcal{I}$ to $\mathcal{I}'$ is t-strong.

Our goal is the following: for a transitive rooted X-labelled interpretation $(\mathcal{I}, a, \lambda)$ over t, construct a rooted X-labelled interpretation over t that has a small finite representation and maps to $(\mathcal{I}, a, \lambda)$ via a t-strong homomorphism.

Next, we make this goal more concrete by explaining how we represent infinite interpretations using finite ones.

3.1 Quasi-Unravelling

We modify the standard unravelling to keep the set of critical elements finite: the *quasi-unravelling* procedure builds a tree-like model by making a separate copy of an element for each path leading to this element, except for critical elements, of which only one copy is kept (see Figure 2).

Let $\mathcal{I}$ be an interpretation over a single transitive role name t. We refer to SCCs of $\mathcal{I}$ as *clusters*. Notice that since we have a single transitive role, a cluster is either a single element (with or without a loop) or a clique. We call a cluster in $\mathcal{I}$ *critical* if it contains a critical element. We write $\Delta_F^{\mathcal{I}}$ for the union of all critical clusters in interpretation $\mathcal{I}$. For an element $a \in \Delta^{\mathcal{I}}$, we let $K_a^{\mathcal{I}}$ be the cluster in $\mathcal{I}$ that contains a if this cluster is critical, and $K_a^{\mathcal{I}} = \{a\}$ otherwise.

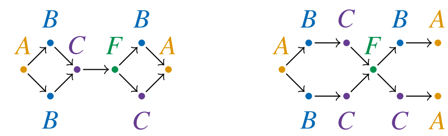


Figure 2: An interpretation (left) and its quasi-unravelling (right).

Definition 2 (Quasi-unravelling)**.** *The quasi-unravelling of a rooted interpretation $(\mathcal{I}, a)$ over a single transitive role t is the rooted interpretation $(\widetilde{\mathcal{I}}, a)$ over t obtained as follows. The domain of $\widetilde{\mathcal{I}}$ is the union of $\Delta_F^{\mathcal{I}} \cup \{a\}$ and the set of pairs of the form (K, p) where K is either $K_a^{\mathcal{I}}$ or a critical cluster of $\mathcal{I}$, and p is a non-empty path in $\mathcal{I}$ that begins in a successor of an element of K and visits elements from $\Delta^{\mathcal{I}} \setminus (\Delta_F^{\mathcal{I}} \cup \{a\})$ only. For a concept name A, we let*

$$A^{\widetilde{\mathcal{I}}} = \Big(A^{\mathcal{I}} \cap \big(\Delta_F^{\mathcal{I}} \cup \{a\}\big)\Big) \cup \big\{(K, pu) \in \Delta^{\widetilde{\mathcal{I}}} \bigm| u \in A^{\mathcal{I}}\big\}.$$

Finally, $t^{\widetilde{\mathcal{I}}}$ is the transitive closure of

$$\begin{gathered}\Big(t^{\mathcal{I}} \cap \big((\Delta_F^{\mathcal{I}} \cup \{a\}) \times \Delta_F^{\mathcal{I}}\big)\Big) \cup \\ \cup \big\{\big(u, (K, p)\big) \bigm| (K, p) \in \Delta^{\widetilde{\mathcal{I}}}, u \in K\big\} \cup \\ \cup \big\{\big((K, p), (K, pu)\big) \bigm| (K, p), (K, pu) \in \Delta^{\widetilde{\mathcal{I}}}\big\} \cup \\ \cup \big\{\big((K, pu), v\big) \bigm| (K, pu) \in \Delta^{\widetilde{\mathcal{I}}}, (u, v) \in t^{\mathcal{I}}, v \in \Delta_F^{\mathcal{I}}\big\}.\end{gathered}$$

The reader might be familiar with a different style of defining unravellings where copies of elements from $(\mathcal{I}, a)$ are added successively to $(\widetilde{\mathcal{I}}, a)$. In our definition, elements from $\Delta_F^{\mathcal{I}} \cup \{a\}$ are copies of themselves, and each (K, p) is a copy of the last element in path p.

The quasi-unravelling $(\widetilde{\mathcal{I}}, a, \widetilde{\lambda})$ of a rooted X-labelled interpretation $(\mathcal{I}, a, \lambda)$ is the quasi-unravelling $(\widetilde{\mathcal{I}}, a)$ of $(\mathcal{I}, a)$ with the labelling function $\widetilde{\lambda}$ defined as $\widetilde{\lambda}(u) = \lambda(u)$ for $u \in \Delta_F^{\mathcal{I}} \cup \{a\}$ and $\widetilde{\lambda}((K, pu)) = \lambda(u)$ for $(K, pu) \in \Delta^{\widetilde{\mathcal{I}}} \setminus (\Delta_F^{\mathcal{I}} \cup \{u\})$.

Quasi-unravelling preserves the satisfaction of TBox constraints and conjunctive queries. Indeed, mapping each copy of an element of $\mathcal{I}$ to its original gives a t-strong homomorphism from $\widetilde{\mathcal{I}}$ to $\mathcal{I}^*$.

Lemma 3. *If $(\widetilde{\mathcal{I}}, a, \widetilde{\lambda})$ is the quasi-unravelling of $(\mathcal{I}, a, \lambda)$ where $\mathcal{I}$ is an interpretation over t then there is a t-strong homomorphism $h\colon (\widetilde{\mathcal{I}}, a, \widetilde{\lambda}) \to (\mathcal{I}^*, a, \lambda)$. Moreover, if each element in $\mathcal{I}$ is reachable from a, we can additionally require h to be surjective.*

From Lemmas 1 and 3 we get the following.

Lemma 4. *Consider an interpretation $\mathcal{I}$ over t and an element a of $\mathcal{I}$ from which every other element of $\mathcal{I}$ is reachable. Let $(\widetilde{\mathcal{I}}, a)$ be the quasi-unravelling of $(\mathcal{I}, a)$. Then*

- *for every TBox $\mathcal{T}$ over t, $\widetilde{\mathcal{I}} \models \mathcal{T}$ iff $\mathcal{I}^* \models \mathcal{T}$;*
- *for every UCQ Q over t, $\widetilde{\mathcal{I}} \models Q$ implies $\mathcal{I}^* \models Q$.*

Lemma 4 shows that $(\mathcal{I},a)$ is a good representation of $(\widetilde{\mathcal{I}},a)$, but the second property is not sufficient to allow deciding if $\widetilde{\mathcal{I}} \models Q$ by just examining $\mathcal{I}$. Later we will show how to pick $(\mathcal{I},a)$ so that $\widetilde{\mathcal{I}} \models Q$ iff $\mathcal{I} \models Q$. This will involve interpretations of a special shape, described next.

3.2 Elementary Interpretations

Elementary interpretations are defined hierarchically, with three levels of increasing complexity, illustrated in Figure 3.

Three basic building blocks form the first level (Figure 3a). We call an interpretation $\mathcal{I}$ over t a *singleton* if the associated graph $G_{\mathcal{I}}$ is a single isolated node; a *cycle* if $G_{\mathcal{I}}$ is a simple cycle; and a *critical cluster* if $G_{\mathcal{I}}$ consists of a single SCC and $\mathcal{I}$ contains a critical element. Note that a singleton interpretation is a critical cluster as long as its unique element is critical. We lift this terminology to rooted interpretations naturally.

Loop trees, defined below, constitute the second level (Figure 3b). They are similar to *cactuses* defined e.g. in (Danecki 1984). The operation of *attaching* an interpretation $(\mathcal{J},v)$ to an element u' of an interpretation $(\mathcal{I},u)$ results in interpretation $(\mathcal{I}',u)$ where $\mathcal{I}'$ is $\mathcal{I}\cup\mathcal{J}$ with an additional edge from u' to v; note that $\Delta^{\mathcal{I}}$ and $\Delta^{\mathcal{J}}$ need not be disjoint.

Definition 5. *We define* loop trees *inductively as follows.*

- *Every rooted singleton interpretation is a loop tree.*
- *If $(\mathcal{I},u)$ is a singleton or a cycle and has no critical elements, and $(\mathcal{I}_1,u_1),\dots,(\mathcal{I}_n,u_n)$ are loop trees that share no non-critical elements with each other and with $(\mathcal{I},u)$, then the interpretation obtained from $(\mathcal{I},u)$ by attaching $(\mathcal{I}_1,u_1),\dots,(\mathcal{I}_n,u_n)$ to u is a loop tree.*

Finally, we define elementary interpretations as trees built from critical clusters and loop trees (Figure 3c).

Definition 6. *We define* elementary interpretations *inductively as follows.*

- *Every rooted critical cluster and every loop tree is an elementary interpretation.*
- *If $(\mathcal{I},u)$ is a loop tree and $(\mathcal{I}_1,u_1),\dots,(\mathcal{I}_n,u_n)$ are pairwise disjoint elementary interpretations such that u_i is the only element shared by $(\mathcal{I},u)$ and $(\mathcal{I}_i,u_i)$, and it is critical in both interpretations, then $(\mathcal{I}\cup\mathcal{I}_1\cup\dots\cup\mathcal{I}_n,u)$ is elementary.*
- *If $(\mathcal{I},u)$ is a critical cluster, $(\mathcal{J},v)$ is an elementary interpretation without cycles visiting v, and the only element shared by $(\mathcal{I},u)$ and $(\mathcal{J},v)$ is v, then $(\mathcal{I}\cup\mathcal{J},u)$ is an elementary interpretation.*

Elementary interpretations will be used as finite representations of potential countermodels, via quasi-unravelling. Not every countermodel has such a representation, but, as we show next, each can be turned into one that does.

3.3 Regularisation of Interpretations

We are now ready to realize the goal of this section: given an interpretation $\mathcal{I}$ with finitely many critical elements, we find a small elementary interpretation $\mathcal{E}$ with a t-strong homomorphism from the quasi-unravelling of $\mathcal{E}$ to $\mathcal{I}$.

Theorem 7. *For every transitive rooted X-labelled interpretation $(\mathcal{I},a,\iota)$ over a transitive role t with finitely many critical elements there is an elementary rooted X-labelled interpretation $(\mathcal{E},a,\varepsilon)$ over t of size at most $(\ell+1)^{(\ell+1)^2}$ for $\ell=|\mathtt{CN}(\mathcal{I})|$ whose quasi-unravelling $(\widetilde{\mathcal{E}},a,\widetilde{\varepsilon})$ maps to $(\mathcal{I},a,\iota)$ via a t-strong homomorphism.*

To prove this theorem, we first observe that the length of simple paths that visit critical elements can be bounded. We call a path *simple* if it never visits the same element twice, and *critical* if it visits only critical elements.

Lemma 8. *Consider a transitive interpretation $\mathcal{I}$ with finitely many critical elements and let $b\in\Delta^{\mathcal{I}}$. Then there is an induced subinterpretation $\mathcal{J}$ of $\mathcal{I}$ such that $b\in\Delta^{\mathcal{J}}$, $\mathtt{rch}_t^{\mathcal{J}}(u)=\mathtt{rch}_t^{\mathcal{I}}(u)$ for all $u\in\Delta^{\mathcal{J}}$, and every simple critical path in $\mathcal{J}$ that does not visit b has length at most $|\mathtt{CN}(\mathcal{I})|$.*

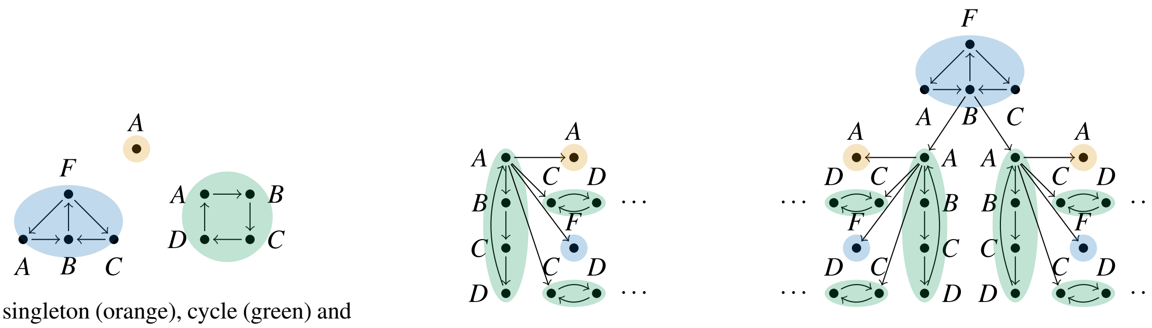
Proof of Theorem 7. The *critical depth* of a rooted X-labelled interpretation $(\mathcal{I},a,\iota)$ is the maximal length of a simple critical path in $\mathcal{I}$ that does not visit the root cluster $K_a^{\mathcal{I}}$. If $\mathcal{I}$ contains finitely many critical elements, then the critical depth of $(\mathcal{I},a,\iota)$ is finite. By Lemma 8, we can assume that the critical depth of $(\mathcal{I},a,\iota)$ is at most ℓ. We can also assume that all elements of $\mathcal{I}$ are reachable from a. By induction on d, we will construct for $(\mathcal{I},a,\iota)$ of critical depth d a suitable elementary $(\mathcal{E},a,\varepsilon)$ of size $(\ell+1)^{(\ell+1)\cdot(d+1)}$.

Consider $(\mathcal{I},a,\iota)$ of critical depth d. Let $(\widetilde{\mathcal{I}},a,\widetilde{\iota})$ be the quasi-unravelling of $(\mathcal{I},a,\iota)$. Note that the critical depth of $(\widetilde{\mathcal{I}},a)$ is d as well.

We call a cluster K in $\widetilde{\mathcal{I}}$ *non-root* if $K\neq K_a^{\widetilde{\mathcal{I}}}$. Because all elements in $\mathcal{I}$ are reachable from a, all critical clusters in $\widetilde{\mathcal{I}}$ are reachable from $K_a^{\widetilde{\mathcal{I}}}$. We call a non-root critical cluster *minimal* if it is not reachable from any other non-root critical cluster. Let $K_1,K_2,\dots,K_m$ be the minimal non-root critical clusters in $\widetilde{\mathcal{I}}$ (if $d=0$, there are none). For each $i\in\{1,2,\dots,m\}$ pick a critical element $c_i\in K_i$ and let $C=\{c_1,\dots,c_m\}$.

For $c\in C$, let $(\widetilde{\mathcal{I}}_c,\widetilde{\iota}_c)$ be the X-labelled interpretation obtained by restricting $(\widetilde{\mathcal{I}},\widetilde{\iota})$ to the domain consisting of c and all elements reachable from c. Clearly, the critical depth of $(\widetilde{\mathcal{I}}_c,c,\widetilde{\iota}_c)$ is at most $d-1$. Let $(\mathcal{E}_c,c,\varepsilon_c)$ be the elementary interpretation of size at most $(\ell+1)^{(\ell+1)\cdot d}$ obtained from $(\widetilde{\mathcal{I}}_c,c,\widetilde{\iota}_c)$ by the induction hypothesis. Without loss of generality we assume that all elements of $\mathcal{E}_c$ except c are fresh; that is, they do not occur in any other interpretations.

It remains to deal with the subinterpretation of $(\widetilde{\mathcal{I}},\widetilde{\iota})$ induced by the set $\Delta^{\widetilde{\mathcal{I}}}\setminus\bigcup_{c\in C}\Delta^{\widetilde{\mathcal{I}}_c}$ of elements not reachable from C. By the definition of quasi-unravelling, this interpretation is the transitive closure of the union of the restriction $(\mathcal{I}_a,\iota_a)$ of $(\mathcal{I},\iota)$ to $K_a^{\mathcal{I}}$ and the restriction $(\widetilde{\mathcal{I}}_S,\widetilde{\iota}_S)$ of $(\widetilde{\mathcal{I}},\widetilde{\iota})$ to the set S consisting of a and all elements of the form $(K_a^{\mathcal{I}},p)$ from $\Delta^{\widetilde{\mathcal{I}}}$. Note that all elements of the latter form are non-critical. Hence, $S\setminus\{a\}$ contains no critical elements. Moreover, $\widetilde{\mathcal{I}}_S$ is the transitive closure of a tree with root a. We think of S as if it were this tree and speak of



(a) A singleton (orange), cycle (green) and cluster (blue).

(b) A loop tree.

(c) An elementary interpretation.

Figure 3: The three levels of elementary interpretations. Critical elements indicated with concept name F.

nodes, leaves, subtrees, children, and descendants in S. We write $(\widetilde{\mathcal{I}}_{S\cup C}, \widetilde{\iota}_{S\cup C})$ for the restriction of $(\widetilde{\mathcal{I}}, \widetilde{\iota})$ to $S \cup C$.

In the reminder of the proof we construct a loop tree whose unravelling maps homomorphically to $(\widetilde{\mathcal{I}}_{S\cup C}, \widetilde{\iota}_{S\cup C})$, and then combine it with $(\mathcal{E}_c, c, \varepsilon_c)$ for $c \in C$ and a bounded-size cluster corresponding to $(\mathcal{I}_a, \iota_a)$ to obtain $(\mathcal{E}, a, \varepsilon)$.

We build the loop tree by induction. We define the *rank* of a node v in S as the size of the set $\mathtt{tp}^{\widetilde{\mathcal{I}}}(v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v)$ of concept names that occur in v or elements of $\widetilde{\mathcal{I}}$ (not necessarily in S) reachable from v. For each descendant v' of v, $\mathtt{tp}^{\widetilde{\mathcal{I}}}(v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v) \supseteq \mathtt{tp}^{\widetilde{\mathcal{I}}}(v') \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v')$, so ranks never increase as we go down S. We produce a loop tree for each subtree of S by induction on the rank of the subtree's root v. The easy case is when the loop tree for v can be built from a descendant v' of v and loop trees already constructed for descendants of v' of lower rank. If this is impossible, we identify an infinite sequence of descendants of v that can be folded into a cycle that can be extended to a loop tree for v by attaching loop trees already constructed for nodes of lower rank. The hard part is to ensure that the unravelling of the whole loop tree (not just the cycle) maps to $(\widetilde{\mathcal{I}}_{S\cup C}, \widetilde{\iota}_{S\cup C})$.

The actual claim we prove includes additional conditions that facilitate the inductive proof and ensure compatibility with $(\mathcal{E}_c, c, \varepsilon_c)$. We say that a homomorphism h *fixes* C if $h(c) = c$ for each $c \in \mathrm{dom}(h) \cap C$.

Claim 9. *For each $v \in S \setminus \{a\}$ of rank q there is a loop tree $(\mathcal{L}_v, u_v, \lambda_v)$ of size at most $(\ell+1)^q$ and a homomorphism*

$$h_v : \big(\widetilde{\mathcal{L}}_v, u_v, \widetilde{\lambda}_v\big) \to \big(\widetilde{\mathcal{I}}_{S\cup C}, v', \widetilde{\iota}_{S\cup C}\big)$$

for some v' such that

- *v' is an element of S reachable from v and*

$$\mathtt{tp}^{\widetilde{\mathcal{I}}}(v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v) = \mathtt{tp}^{\widetilde{\mathcal{I}}}(v') \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v')\,; \tag{1}$$

- *h_v extends to a t-strong homomorphism fixing C*

$$\widehat{h}_v : \big(\widetilde{\mathcal{E}}_v, u_v, \widetilde{\varepsilon}_v\big) \to \big(\widetilde{\mathcal{I}}, v', \widetilde{\iota}\big)$$

where $\big(\mathcal{E}_v, \varepsilon_v\big)$ is the union of $\big(\mathcal{L}_v, \lambda_v\big)$ and $\big(\mathcal{E}_c, \varepsilon_c\big)$ for all $c \in C$ such that $c \in \Delta^{\mathcal{L}_v}$.

Proof of Claim 9. To make the construction of $\big(\mathcal{L}_v, u_v, \lambda_v\big)$ more uniform, for $c \in C$, we let $(\mathcal{L}_c, u_c, \lambda_c)$ be the restriction of $(\widetilde{\mathcal{I}}, c, \widetilde{\iota})$ to $\{c\}$ with all edges removed. Note that $(\mathcal{L}_c, u_c, \lambda_c)$ satisfies the conditions in the statement of the claim with $c' = c$ and the identity function for h_c.

Consider first a node $v \in S \setminus \{a\}$ of rank 0. We define $(\mathcal{L}_v, u_v, \lambda_v)$ as the restriction of $(\widetilde{\mathcal{I}}, v, \widetilde{\iota})$ to $\{v\}$. We let $v' = v$ and take the identity function for h_v. The condition in the first bullet holds trivially. For the second bullet, $\Delta^{\mathcal{L}_v} = \{v\}$ and $v \notin C$, so $(\mathcal{E}_v, u_v, \varepsilon_v) = (\mathcal{L}_v, u_v, \lambda_v)$ and h_v itself is the required t-strong homomorphism from $(\widetilde{\mathcal{E}}_v, u_v, \widetilde{\varepsilon}_v)$ to $\big(\widetilde{\mathcal{I}}, v', \widetilde{\iota}\big)$ because $\mathtt{tp}^{\widetilde{\mathcal{E}}_v}(u_v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{E}}_v}(u_v) = \emptyset = \mathtt{tp}^{\widetilde{\mathcal{I}}}(v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v)$.

Suppose now that we have shown the claim for nodes of rank at most $q-1$ and consider a node $v \in S \setminus \{a\}$ of rank $q > 0$. Let S_v be the subtree of S rooted at v and let R_v be the subset of S_v consisting of all nodes of rank q. Note that R_v forms a *prefix* of S_v: if a node from S belongs to R_v, so do all its ancestors in S. We will think of R_v as a tree, too. Note also that every $v' \in R_v$ satisfies condition (1). We have two cases, depending on whether R_v contains a leaf or not.

Case 1. R_v contains a leaf v'; that is, all descendants of v' in S have lower rank. Pick a minimal subset W of elements of $S \cup C$ reachable from v' in $\widetilde{\mathcal{I}}$ such that

$$\mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v') = \bigcup_{w \in W} \mathtt{tp}^{\widetilde{\mathcal{I}}}(w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(w)\,. \tag{2}$$

Each element $w \in W$ is an element of $S \setminus \{a\}$ of rank at most $q-1$ or an element of C. By the induction hypothesis and the initial step of the proof of the claim, $(\mathcal{L}_w, u_w, \lambda_w)$ exists for all $w \in W$. We let $u_v = v'$ and define $(\mathcal{L}_v, \lambda_v)$ as follows. We restrict $(\widetilde{\mathcal{I}}, \widetilde{\iota})$ to $\{v'\}$ and for each $w \in W$ we attach to v' a *quasi-fresh* copy of the interpretation $(\mathcal{L}_w, u_w, \lambda_w)$, where each non-critical element is replaced by a fresh element. Because $|W| \le q \le \ell$ and $|\Delta^{\mathcal{L}_w}| \le (\ell+1)^{q-1}$ for all $w \in W$,

$$|\Delta^{\mathcal{L}_v}| \le 1 + q \cdot (\ell+1)^{q-1} \le (\ell+1)^q\,.$$

A suitable h_v: $\big(\widetilde{\mathcal{L}}_v, u_v, \widetilde{\lambda}_v\big) \to \big(\widetilde{\mathcal{I}}_{S\cup C}, v', \widetilde{\iota}_{S\cup C}\big)$ can be obtained by setting $h_v(u_v) = v'$ and combining homomorphisms h_w for $w \in W$ as follows. Consider a path p in the quasi-fresh copy of $\mathcal{L}_w$ used in $\mathcal{L}_v$, beginning in the copy of u_w. There is a corresponding path p' in $\mathcal{L}_w$, beginning in

u_w. We let $h_v(u_v p) = h_w(p')$. For every $c \in C \cap \Delta^{\widetilde{\mathcal{L}}_v}$ we let $h_v(c) = c$. It is routine to check that h_v is a homomorphism and that it extends to a suitable t-strong homomorphism fixing C. Let us only verify that $\mathtt{rch}_t^{\widetilde{\mathcal{E}}_v}(u_v) = \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v')$. Recall that $u_v = v'$. By construction,

$$\mathtt{rch}_t^{\widetilde{\mathcal{E}}_v}(v') = \bigcup_{w \in W} \mathtt{tp}^{\widetilde{\mathcal{E}}_w}(u_w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{E}}_w}(u_w)\,.$$

By the induction hypothesis, for each $w \in W$, using the existence of a t-strong homomorphism from $(\widetilde{\mathcal{E}}_w, u_w, \widetilde{\varepsilon}_w)$ to $(\widetilde{\mathcal{I}}, w', \widetilde{\iota})$, followed by (1) for w, we get

$$\begin{aligned}\mathtt{tp}^{\widetilde{\mathcal{E}}_w}(u_w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{E}}_w}(u_w) = \mathtt{tp}^{\widetilde{\mathcal{I}}}(w') \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(w') = \\ = \mathtt{tp}^{\widetilde{\mathcal{I}}}(w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(w)\,.\end{aligned}$$

We conclude by applying (2).

Case 2. R_v does not contain a leaf; that is, each node in R_v has a child in R_v. We need a similar, but stronger property for concept names. For a subset U of S, we call a concept name A *dense in* U if each element in U has a proper descendant in U that belongs to $A^{\widetilde{\mathcal{I}}}$. We construct a subtree R'_v of R_v such that each concept name occurring in R'_v is dense in R'_v.

Let $\mathtt{tp}^{\widetilde{\mathcal{I}}}(v) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v) = \{A_1, \ldots, A_n\}$. We process A_i one by one and eliminate those that are not dense by passing to a smaller subtree of R_v, also without leaves. Let R_v^i be the subtree of R_v obtained after processing $A_1, \ldots, A_i$, with $R_v^0 = R_v$. If A_{i+1} is dense in R_v^i or does not occur in R_v^i, we let $R_v^{i+1} = R_v^i$. If A_{i+1} occurs in R_v^i but is not dense in R_v^i, there is an element u of R_v^i without proper descendants in $A_{i+1}^{\widetilde{\mathcal{I}}} \cap R_v^i$. Since R_v^i has no leaves, node u has a child u' in R_v^i. We define R_v^{i+1} as the the subtree of R_v^i rooted at u'. Note that R_v^{i+1} is non-empty and has no leaves. Moreover, A_{i+1} does not occur in R_v^{i+1}, each concept name that does not occur in R_v^i does not occur in R_v^{i+1} either, and each concept name dense in R_v^i is also dense in R_v^{i+1}. We let $R'_v = R_v^n$.

Let $B_1, \ldots, B_k$ be the concept names that occur in R'_v. Because all these concept names are dense in R'_v, we can find an infinite sequence

$$v_1^1, v_2^1, \ldots, v_k^1,\; v_1^2, v_2^2, \ldots, v_k^2,\; \ldots,\; v_1^i, v_2^i, \ldots, v_k^i,\; \ldots$$

of elements from R'_v that form a subsequence of a branch in R'_v ensuring that $v_j^i \in B_j^{\widetilde{\mathcal{I}}}$ for all i and j. Next, for each i, we pick a minimal subset W_i of elements of $(S \setminus R'_v) \cup C$ reachable from v_1^i in $\widetilde{\mathcal{I}}$ such that

$$\mathtt{rch}_t^{\widetilde{\mathcal{I}}}(v_1^i) \setminus \{B_1, \ldots, B_k\} = \bigcup_{w \in W_i} \mathtt{tp}^{\widetilde{\mathcal{I}}}(w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(w)\,.$$

Each element of W_i is an element of $S \setminus \{a\}$ of rank at most $q-1$ or an element of C. By the induction hypothesis and the initial step of the proof of the claim, for each $w \in \bigcup_{i=1}^{\infty} W_i$, $(\mathcal{L}_w, u_w, \lambda_w)$ is already defined and is an X-labelled loop tree of size at most $(\ell+1)^{q-1}$ using only concept names from $\mathtt{CN}(\widetilde{\mathcal{I}})$ and critical elements from C. Clearly, there are only finitely many such interpretations, up to an isomorphism fixing C. In consequence, there is an infinite sequence $\ell_1 < \ell_2 < \ldots$ such that

- the sets $\{(\mathcal{L}_w, u_w, \lambda_w) \mid w \in W_{\ell_i}\}$ are equal for all i (up to an isomorphism fixing C),
- the tuples $(\mathtt{tp}^{\widetilde{\mathcal{I}}}(v_1^{\ell_i}), \ldots, \mathtt{tp}^{\widetilde{\mathcal{I}}}(v_k^{\ell_i}))$ are equal for all i, and
- the tuples $(\iota(v_1^{\ell_i}), \ldots, \iota(v_k^{\ell_i}))$ are equal for all i.

Let $u_v = v_1^{\ell_1}$ and construct $(\mathcal{L}_v, \lambda_v)$ by taking the restriction of $(\widetilde{\mathcal{I}}, \widetilde{\iota})$ to $\{v_1^{\ell_1}, v_2^{\ell_1}, \ldots, v_k^{\ell_1}\}$ with all edges dropped, adding edges $(v_1^{\ell_1}, v_2^{\ell_1}), \ldots, (v_{k-1}^{\ell_1}, v_k^{\ell_1}), (v_k^{\ell_1}, v_1^{\ell_1})$, and attaching to $v_1^{\ell_1}$ a quasi-fresh copy of $(\mathcal{L}_w, u_w, \lambda_w)$ for each $w \in W_{\ell_1}$. Then,

$$|\Delta^{\mathcal{L}_v}| \le k + (q-k)\cdot(\ell+1)^{q-1} \le (\ell+1)^q\,.$$

Let us now define $h_v \colon (\widetilde{\mathcal{L}}_v, u_v, \widetilde{\lambda}_v) \to (\widetilde{\mathcal{I}}_{S\cup C}, v', \widetilde{\iota}_{S\cup C})$ for $v' = v_1^{\ell_1}$. For $c \in C \cap \Delta^{\widetilde{\mathcal{L}}_v}$, let $h_v(c) = c$. The elements of $\widetilde{\mathcal{L}}_v$ obtained by unravelling the cycle $v_1^{\ell_1}, v_2^{\ell_1}, \ldots, v_k^{\ell_1}$ in $\mathcal{L}_v$ are mapped to the elements $v_1^{\ell_1}, v_2^{\ell_1}, \ldots, v_k^{\ell_1}, v_1^{\ell_2}, v_2^{\ell_2}, \ldots, v_k^{\ell_2}, \ldots$ in $\widetilde{\mathcal{I}}$: for $i \ge 1$ and $j \le k$ we let

$$\underbrace{\left(v_1^{\ell_1} v_2^{\ell_1} \cdots v_k^{\ell_1}\right) \cdots \left(v_1^{\ell_1} v_2^{\ell_1} \cdots v_k^{\ell_1}\right)}_{i-1} v_1^{\ell_1} v_2^{\ell_1} \cdots v_j^{\ell_1} \;\mapsto\; v_j^{\ell_i}\,.$$

For the remaining elements we rely on h_w for $w \in W_{\ell_1}$. For every $w \in W_{\ell_1}$ and $i \ge 1$ there is $w^{(i)} \in W_{\ell_i}$ such that $(\mathcal{L}_w, u_w, \lambda_w)$ is isomorphic to $(\mathcal{L}_{w^{(i)}}, u_{w^{(i)}}, \lambda_{w^{(i)}})$ via an isomorphism $\eta_{w,i}$ fixing C. Consider a path p in the quasi-fresh copy of $(\mathcal{L}_w, u_w, \lambda_w)$ used in $(\mathcal{L}_v, u_v, \lambda_v)$, starting from the copy of u_w. Using the isomorphism $\eta_{w,i}$, we find a corresponding path $p^{(i)}$ in $(\mathcal{L}_{w^{(i)}}, u_{w^{(i)}}, \lambda_{w^{(i)}})$, starting from $u_{w^{(i)}}$. We let

$$\underbrace{\left(v_1^{\ell_1} v_2^{\ell_1} \cdots v_k^{\ell_1}\right) \cdots \left(v_1^{\ell_1} v_2^{\ell_1} \cdots v_k^{\ell_1}\right)}_{i-1} v_1^{\ell_1} p \;\mapsto\; h_{w^{(i)}}(p^{(i)})\,.$$

The element v' is reachable from v and satisfies (1), because $v' \in R'_v \subseteq R_v$. Verifying that h_v extends to a suitable t-strong homomorphism is straightforward, based on the properties of $v_1^{\ell_1}, v_2^{\ell_1}, \ldots, v_k^{\ell_1}, v_1^{\ell_2}, v_2^{\ell_2}, \ldots, v_k^{\ell_2}, \ldots$ and the inductive hypothesis. This completes the proof of Claim 9. $\square$

With $(\mathcal{L}_v, u_v, \lambda_v)$ at hand for all $v \in S \setminus \{a\}$, we can define the elementary interpretation we are looking for. Let W be a minimal subset of $S \cup C$ such that

$$\mathtt{rch}_t^{\widetilde{\mathcal{I}}}(a) \setminus \mathtt{rch}_t^{\mathcal{I}_a}(a) = \bigcup_{w \in W} \mathtt{tp}^{\widetilde{\mathcal{I}}}(w) \cup \mathtt{rch}_t^{\widetilde{\mathcal{I}}}(w)\,.$$

Let V be a minimal subset of $K_a^{\mathcal{I}}$ such that $a \in V$ and $\mathtt{rch}_t^{\mathcal{I}_a}(a) = \bigcup_{v \in V} \mathtt{tp}^{\mathcal{I}_a}(v)$. Let $(\mathcal{I}'_a, \iota'_a)$ be the restriction of $(\mathcal{I}_a, \iota_a)$ to V. Let $(\mathcal{L}_a, \lambda_a)$ the X-labelled interpretations obtained from $(\mathcal{I}'_a, \iota_a)$ by attaching to a a quasi-fresh copy of $(\mathcal{L}_w, u_w, \lambda_w)$ for all $w \in W$. Then,

$$\begin{aligned}|\Delta^{\mathcal{L}_a}| \le 1 + |\mathtt{rch}_t^{\mathcal{I}_a}(a)| + \left(\ell - |\mathtt{rch}_t^{\mathcal{I}_a}(a)|\right)\cdot(\ell+1)^{\ell} \le \\ \le (\ell+1)^{\ell+1}\,.\end{aligned}$$

As in Case 1 in the claim, we can construct a homomorphism $h_a : (\widetilde{\mathcal{L}}_a, a, \widetilde{\lambda}_a) \to (\widetilde{\mathcal{I}}_{S\cup C}, a, \widetilde{\iota}_{S\cup C})$ that extends to a t-strong homomorphism $\widehat{h}_a : (\widetilde{\mathcal{E}}_a, a, \widetilde{\varepsilon}_a) \to (\widetilde{\mathcal{I}}, a, \widetilde{\iota})$ where $(\mathcal{E}_a, \varepsilon_a)$ is the union of $(\mathcal{L}_a, \lambda_a)$ and $(\mathcal{E}_c, \varepsilon_c)$ for all $c \in C$ such that $c \in \Delta^{\mathcal{L}_a}$. Note that

$$|\Delta^{\mathcal{E}_a}| \leq (\ell+1)^{\ell+1} \cdot (\ell+1)^{(\ell+1)\cdot d} \leq (\ell+1)^{(\ell+1)(d+1)}.$$

Hence, $(\mathcal{E}_a, a, \varepsilon_a)$ is the interpretation $(\mathcal{E}, a, \varepsilon)$ we seek. $\square$

4 Partially Finite Coloured Blocking

This section provides a variant of the *coloured blocking* theorem (Gogacz *et al.* 2018; Gogacz *et al.* 2019), suitable for partially finite models (over multiple roles). The original theorem offers a method to turn an infinite countermodel into a finite one, under certain assumptions justified by the *a priori* existence of a finite countermodel. Here, we only know that a partially finite countermodel exists, which does not give equally strong assumptions. In what follows we describe an alternative construction that preserves countermodels and ensures partial finiteness when applied to an interpretation satisfying suitable weaker assumptions.

Quotients. Let $\mathcal{I}$ be an interpretation and $\sim$ an equivalence relation on $\Delta^{\mathcal{I}}$. We write $[x]_\sim$ for the equivalence class $\{y \in \Delta^{\mathcal{I}} \mid y \sim x\}$ of $x \in \Delta^{\mathcal{I}}$ and $\mathcal{I}/\sim$ for the *quotient interpretation* where

- $\Delta^{(\mathcal{I}/\sim)} = \{[x]_\sim \mid x \in \Delta^{\mathcal{I}}\}$,
- $A^{(\mathcal{I}/\sim)} = \{[x]_\sim \mid x \in A^{\mathcal{I}}\}$ for all concept names A,
- $r^{(\mathcal{I}/\sim)} = \{([x]_\sim, [y]_\sim) \mid (x,y) \in r^{\mathcal{I}}\}$ for all role names r.

The quotient construction naturally induces the *quotient homomorphism* $h_\sim : \mathcal{I} \to \mathcal{I}/\sim$ defined as $h_\sim(x) = [x]_\sim$.

Distances. An interpretation $\mathcal{I}$ can be viewed as an undirected graph with an edge between $x \in \Delta^{\mathcal{I}}$ and $y \in \Delta^{\mathcal{I}}$ whenever (x,y) or (y,x) belongs to $r^{\mathcal{I}}$ for some $r \in \mathbb{N}_{\mathrm{R}}$. By an *undirected path* in $\mathcal{I}$ we mean a path in this undirected graph. The *distance in* $\mathcal{I}$ is the function $d_{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \to \mathbb{N} \cup \{\infty\}$ where $d_{\mathcal{I}}(x,y)$ is the minimal length of an undirected path in $\mathcal{I}$ between x and y, or ∞ if there is no such path. Naturally, $d_{\mathcal{I}}$ satisfies the triangle inequality: $d_{\mathcal{I}}(x,z) \leq d_{\mathcal{I}}(x,y) + d_{\mathcal{I}}(y,z)$ for all $x,y,z \in \Delta^{\mathcal{I}}$. In the case of a single transitive role, the distance is always 0, 1, or ∞. But already for two transitive roles $d_{\mathcal{I}}$ can take all possible values.

For $k \geq 0$ we define the *k-neighbourhood* of an element $x \in \Delta^{\mathcal{I}}$ as the set $N_k(x) = \{y \in \Delta^{\mathcal{I}} \mid d_{\mathcal{I}}(x,y) \leq k\}$. We will often be interested in the induced subinterpretation $\mathcal{I} \upharpoonright N_k(x)$, which might be infinite.

Coloured Interpretations. Let us fix a set C of *colours*. A *coloured interpretation* is an interpretation $\mathcal{I}$ along with a function $\kappa : \Delta \to C$ for some $\Delta \subseteq \Delta^{\mathcal{I}}$ that assigns colours to some elements of $\mathcal{I}$. If $\mathcal{I}$ is coloured by κ and $\mathcal{J}$ is coloured by θ then a homomorphism $h : \mathcal{I} \to \mathcal{J}$ *preserves colours* if for every $x \in \Delta^{\mathcal{I}}$ we have: $x \in \mathrm{dom}(\kappa) \Leftrightarrow h(x) \in \mathrm{dom}(\theta)$ and if $x \in \mathrm{dom}(\kappa)$ then $\kappa(x) = \theta(h(x))$. Two interpretations $\mathcal{I}$ and $\mathcal{J}$ coloured by κ and θ are *homomorphically equivalent* if there is a homomorphism from $\mathcal{I}$ to $\mathcal{J}$ and a homomorphism from $\mathcal{J}$ to $\mathcal{I}$, both of them preserving colours. Homomorphic equivalence is reflexive, symmetric, and transitive, but it does not guarantee an isomorphism between the interpretations.

k-Neighbourhood Equivalence. The original coloured blocking relies on the finite number of isomorphism types of k-neighbourhoods. As we cannot assume it here, we use the weaker notion of homomorphism equivalence.

Consider $k \geq 0$ and an interpretation $\mathcal{I}$ coloured by κ. For each $x \in \Delta^{\mathcal{I}}$, the interpretation $\mathcal{I} \upharpoonright N_k(x)$ is naturally coloured by $\kappa \upharpoonright (N_k(x) \cap \mathrm{dom}(\kappa))$. We define *k-neighbourhood equivalence* $\sim_k$ as the equivalence relation on $\Delta^{\mathcal{I}}$ such that $x \sim_k y$ iff $x = y$ or $x,y \in \mathrm{dom}(\kappa)$, $\kappa(x) = \kappa(y)$, and the coloured interpretations $\mathcal{I} \upharpoonright N_k(x)$ and $\mathcal{I} \upharpoonright N_k(y)$ are homomorphically equivalent via homomorphisms that maps x to y and vice versa. Note that $\sim_k$ merges only coloured elements.

For $k \geq 0$, we say that an interpretation $\mathcal{I}$ coloured by κ is *k-sparse* if for every $x \in \Delta^{\mathcal{I}}$ and every two $y, y' \in N_k(x) \cap \mathrm{dom}(\kappa)$, if $y \neq y'$ then $\kappa(y) \neq \kappa(y')$.

Theorem 10. *Let $K = 2 \cdot k^3$ for $k \geq 0$. Consider a K-sparse interpretation $\mathcal{I}$ coloured by κ and a finite interpretation $\mathcal{Q}$ with $|\Delta^{\mathcal{Q}}| \leq k$. Then $\mathcal{Q}$ maps homomorphically into $\mathcal{I}$ iff $\mathcal{Q}$ maps homomorphically into $\mathcal{I}/\sim_K$.*

The right-to-left implication is clear as the quotient homomorphism $h_{\sim_K}$ maps $\mathcal{I}$ to $\mathcal{I}/\sim_K$. For the converse, consider a homomorphism $h : \mathcal{Q} \to \mathcal{I}/\sim_K$. We need to lift it to a homomorphism from $\mathcal{Q}$ to $\mathcal{I}$. We begin by dividing $\mathcal{Q}$ into finitely many *pieces*. Initially, each piece contains at most two coloured elements. For such pieces we can lift the corresponding restriction of h to a homomorphism to $\mathcal{I}$, using $h_{\sim_K}$. Then, we iteratively glue pieces into larger ones, while preserving the existence of homomorphisms to $\mathcal{I}$. This is possible because $\sim_K$-equivalent elements of $\mathcal{I}$ have homomorphically equivalent neighbourhoods. Ultimately all pieces are glued together into a single piece $\mathcal{Q}$ with a corresponding homomorphism into $\mathcal{I}$.

In the following section we apply Theorem 10 to a suitably coloured interpretation of a special shape and prove that the resulting quotient interpretation is partially finite.

5 Multiple Roles Allowed

In this section we show how to solve the entailment problem in the general case. The strategy is familiar: we characterize non-entailment in terms of structurally simple witnesses and show that the existence of such a witness is decidable.

A rooted interpretation $(\mathcal{I}, a)$ is *piecewise single-role* if it is the union of a family of pairwise disjoint rooted interpretations called *pieces*, that are either singleton interpretations or single-role interpretations over a transitive role, called *transitive pieces*, with some additional edges between elements from different pieces. Specifically, the pieces can be arranged into a potentially infinite tree such that

- the root piece is a singleton interpretation with root a;
- for every two adjacent pieces in the tree, $\mathcal{I}$ contains exactly one edge from an element in the parent piece to the root of the child piece, and there are no other edges in $\mathcal{I}$ between elements from different pieces.

To disambiguate decompositions into pieces, we assume that, for every transitive role name t, if a piece contains a t-edge or its root has an incoming t-edge, then there are no t-edges originating in this piece and leading to other pieces.

An interpretation $(\mathcal{G},a)$ is *piecewise elementary* if it is piecewise single-role and each non-singleton piece is a transitive piece. The unravelling of a piecewise single-role interpretation $(\mathcal{G},a)$ is a piecewise single-role interpretation $(\widetilde{\mathcal{G}},a)$ obtained as follows. Starting from the (singleton) root piece of $(\mathcal{G},a)$, repeat the following exhaustively: for each newly added element e' that is a copy of element e from $\mathcal{G}$ (or e itself), for each piece $(\mathcal{J},b)$ of $(\mathcal{G},a)$ such that there is an r-edge from e to b, add a fresh copy of the quasi-unravelling of $(\mathcal{J},b)$, with an r-edge from e' to the root. Finally, close all transitive roles by transitivity.

Theorem 11. *For $\mathcal{K}=(\mathcal{T},\{A(a)\})$, a concept name F, and a UCQ Q, if $\mathcal{K}\not\models_F Q$ then there is a piecewise elementary interpretation $(\mathcal{G},a)$ of piece size at most $(\ell+1)^{(\ell+1)^2}$ for $\ell=|\mathtt{CN}(\mathcal{K})|$ and degree at most $\|\mathcal{T}\|$ whose quasi-unravelling $(\widetilde{\mathcal{G}},a)$ satisfies $\widetilde{\mathcal{G}}\models\mathcal{K}$ and $\widetilde{\mathcal{G}}\not\models Q$.*

We build $(\mathcal{G},a)$ starting from the unravelling $(\mathcal{J},a)$ of any countermodel. Using Theorem 7, for each maximal connected component $(\mathcal{C},c)$ of $\mathcal{J}$ over a transitive role t we can produce a suitable elementary interpretation $(\mathcal{E},c)$ whose quasi-unravelling maps to $(\mathcal{C},c)$ via a t-strong homomorphism. In order to be able to construct $(\mathcal{G},a)$ from these elementary interpretations, we need them to be compatible: if $\mathcal{J}$ contains components $(\mathcal{C}_1,c_1)$ and $(\mathcal{C}_2,c_2)$ over a transitive role s such that some copies of the same element from $(\mathcal{E},c)$ are mapped to c_1 and c_2, then $(\mathcal{E}_1,c_1)$ and $(\mathcal{E}_2,c_2)$ must be identical. This can be ensured using X-labellings for a suitable X, but it requires picking the elementary interpretations bottom up. How can we do it in an infinite tree? We define initial fragments $(\mathcal{J}_k,a)$ of $(\mathcal{J},a)$ of depth k. For each $(\mathcal{J}_k,a)$ we build a finite piecewise elementary interpretation $(\mathcal{G}_k,a)$ bottom up. Then we use König's lemma to extract a single infinite interpretation $(\mathcal{G},a)$.

A characterization of non-entailment could be obtained by proving the converse of Theorem 11, but the condition $\widetilde{\mathcal{G}}\not\models Q$ makes it hard to decide the existence of such witnesses. Instead, we replace this condition with a weaker one, that is easier to check and still sufficient to prove the converse.

We use a blow-up operation, which can also be seen as partial quasi-unravelling. Given a piecewise elementary interpretation $(\mathcal{G},a)$ and a positive integer n, we define $(\mathcal{G}_n,a)$ as follows. We process $\mathcal{G}$ top-down, piece by piece. In each piece we blow-up the cycles in loop trees, also in the top-down order: we unravel each cycle to a cycle $2n+1$ times longer. Each copy e' of element e from the original cycle gets its own fresh copy of the subinterpretation of $\mathcal{G}$ induced by elements reachable from e in $\mathcal{G}$ without visiting any other elements on the cycle. When all the cycles are blown-up, we partially close the resulting interpretation by transitivity: we add all edges from the transitive closure except those between elements from the same blown-up cycle. Finally, in each blown-up cycle we transitively close the first n of the $2n+1$ copies of the original cycle.

Because partial quasi-unravelling followed by full quasi-unravelling is essentially the same as just full quasi-unravelling, we easily get the following.

Lemma 12. *For every piecewise elementary interpretation $(\mathcal{G},a)$ and positive integer n, the quasi-unravelling $(\widetilde{\mathcal{G}_n},a)$ of $(\mathcal{G}_n,a)$ maps homomorphically into the quasi-unravelling $(\widetilde{\mathcal{G}},a)$ of $(\mathcal{G},a)$.*

By Lemma 12, Theorem 11 holds with $\widetilde{\mathcal{G}}\not\models Q$ replaced with $\widetilde{\mathcal{G}_n}\not\models Q$ for any n. We can also prove the converse.

Lemma 13. *For $\mathcal{K}=(\mathcal{T},\{A(a)\})$, a concept name F, and a UCQ Q, if there is a piecewise elementary interpretation $(\mathcal{G},a)$ of bounded piece size and bounded degree such that $\widetilde{\mathcal{G}}\models\mathcal{K}$ and $\widetilde{\mathcal{G}_n}\not\models Q$, then $\mathcal{K}\not\models_F Q$.*

A partially finite countermodel is obtained by applying Theorem 10 to $\widetilde{\mathcal{G}}_n$ with suitably coloured critical elements to ensure sparsity. As quasi-unravelling does not duplicate critical elements, we can sparsely colour $\mathcal{G}_n$ instead. This can be done because $\mathcal{G}_n$ has bounded degree.

Theorem 11, Lemma 12, and Lemma 13 together establish an equivalent criterion for non-entailment. But how does replacing $\widetilde{\mathcal{G}}\not\models Q$ with $\widetilde{\mathcal{G}}_n\not\models Q$ help with decidability? The construction ensures that $\widetilde{\mathcal{G}_n}$ and $\mathcal{G}_n$ satisfy the same conjunctive queries of size at most n.

Lemma 14. *For a piecewise elementary interpretation $(\mathcal{G},a)$, a positive integer n, and a conjunctive query Q with at most n variables, $\mathcal{G}_n\models Q$ iff $\widetilde{\mathcal{G}_n}\models Q$.*

The left-to-right implication follows from the absence of short cycles in $\mathcal{G}_n$. The converse would be a standard property of unravellings if not for the transitive closure in the definition of quasi-unravelling. We emulate the missing edges using the partial transitive closure in the definition of $\mathcal{G}_n$.

By Lemma 14, we can replace $\widetilde{\mathcal{G}}_n\not\models Q$ in our criterion with $\mathcal{G}_n\not\models Q$ where n is the maximal number of variables in a CQ from Q. The latter condition can be checked on the fly during a type-elimination procedure, which can be used to check the existence of a suitable witness.

Lemma 15. *Given a KB $\mathcal{K}=(\mathcal{T},\{A(a)\})$, a concept name F, and a UCQ Q, one can decide in 2ExpTime whether there is a piecewise elementary interpretation $(\mathcal{G},a)$ of piece size at most $(\ell+1)^{(\ell+1)^2}$ for $\ell=|\mathtt{CN}(\mathcal{K})|$ and degree at most $\|\mathcal{T}\|$ such that $(\widetilde{\mathcal{G}},a)\models\mathcal{K}$ and $\mathcal{G}_n\not\models Q$.*

Putting together all the ingredients, along with the standard method for simplifying the ABox (see e.g. (Ibáñez-García *et al.* 2026)), we obtain our main result.

Theorem 16. *Partially finite entailment of conjunctive queries for $\mathcal{S}$ is 2ExpTime-complete.*

The lower bound holds already for finite and unrestricted entailment (Ibáñez-García *et al.* 2026).

6 An Application to Closed Predicates

As our last contribution, we present a result illustrating that partially finite reasoning occurs naturally in problems related to closed predicates. In the problem of *query containment in the presence of closed predicates*, we are given two Boolean CQs q_1, q_2, a TBox $\mathcal{T}$, and a set $\mathcal{F}$ of concept names, and we have to decide whether $q_1 \subseteq_{\mathcal{T},\mathcal{F}} q_2$, that is, whether for every ABox $\mathcal{A}$ and every interpretation $\mathcal{I}$ such that $\mathcal{I} \models (\mathcal{T},\mathcal{A})$ and $A^{\mathcal{I}} = \{a \mid A(a) \in \mathcal{A}\}$ for each $A \in \mathcal{F}$, if $\mathcal{I} \models q_1$ then $\mathcal{I} \models q_2$. Notice that in the formulation of the problem we essentially quantify universally over partially finite interpretations $\mathcal{I}$ with F defined as the union of predicates from $\mathcal{F}$. And indeed, this problem can be solved via a Turing reduction to partially finite entailment.

Theorem 17. *With an oracle for partially finite entailment, one can decide query non-containment in the presence of closed predicates in nondeterministic polynomial time.*

Proof. Let q_1, q_2 be conjunctive queries, $\mathcal{T}$ a TBox, and $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ a set of concept names. We will show that $q_1 \subsetneq_{\mathcal{T},\mathcal{F}} q_2$ iff $(\mathcal{T},\mathcal{A}_{q_1'}) \not\models_F q_2$ for $F = F_1 \sqcup F_2 \sqcup \dots \sqcup F_n$ and some ABox $\mathcal{A}_{q_1'}$ obtained from a homomorphic image q_1' of q_1 by interpreting variables as individuals and atoms as assertions. This suffices, because the algorithm for query non-containment in the presence of closed predicates can then simply guess the correct homomorphic image of q_1 and test partially finite entailment.

Suppose that $(\mathcal{T},\mathcal{A}_{q_1'}) \not\models_F q_2$ for some $\mathcal{A}_{q_1'}$ as above. Then, there is an interpretation $\mathcal{I}$ such that $\mathcal{I} \models (\mathcal{T},\mathcal{A}_{q_1'})$, $\mathcal{I} \not\models q_2$, and $F^{\mathcal{I}}$ is finite. Consequently, $q_1 \not\subseteq_{\mathcal{T},\mathcal{F}} q_2$ as witnessed by $\mathcal{A} = \bigcup_{i=1}^{n} \{F_i(a) \mid a \in F_i^{\mathcal{I}}\}$ and $\mathcal{I}$.

Conversely, suppose that $q_1 \not\subseteq_{\mathcal{T},\mathcal{F}} q_2$. Then there is an ABox $\mathcal{A}$ and an interpretation $\mathcal{I}$ such that $\mathcal{I} \models (\mathcal{T},\mathcal{A})$, $F_i^{\mathcal{I}} = \{a \mid F_i(a) \in \mathcal{A}\}$ for each i, $\mathcal{I} \models q_1$, and $\mathcal{I} \not\models q_2$. Clearly, $F^{\mathcal{I}}$ is finite. Some homomorphic image q_1' of q_1 maps to $\mathcal{I}$ via an injective homomorphism h. Assuming that in $\mathcal{A}_{q_1'}$ variables of q_i' are replaced with their images via h, $\mathcal{I} \models (\mathcal{T},\mathcal{A}_{q_1'})$. Hence, $\mathcal{I}$ witnesses that $(\mathcal{T},\mathcal{A}_{q_1'}) \not\models_F q_2$. □

In the above argument we need to check all homomorphic images of query q_1 because we adopt the Standard Names Assumption. Had we not done so, we could use a single ABox that could collapse to either image of the query and the algorithm would work in deterministic polynomial time.

7 Conclusion

With this paper we aim to initiate a systematic study of partially finite reasoning, where the reasoning tasks are performed under the assumption that some, explicitly defined, parts of the interpretations are necessarily finite. We have shown that in the case of query entailment this mode of reasoning, subsuming both query entailment and finite query entailment, does not come with an increased cost, retaining its 2-ExpTime complexity. Natural next steps include inspecting other logics without finite controllability, like $\mathcal{ALCIF}$, as well as other reasoning tasks, such as query containment, enumeration, or direct access.

Acknowledgements

This work was supported by Poland's NCN grants 2018/30/E/ST6/00042 (Gogacz, Murlak, Przybyłko, Rogova) and 2024/55/B/ST6/00318 (Skrzypczak).

AI Declaration

The authors have not employed any Generative AI tools.

References

Shqiponja Ahmetaj, Magdalena Ortiz, and Mantas Simkus. Polynomial Datalog rewritings for ontology mediated queries with closed predicates. In Reinhard Pichler and Altigran Soares da Silva, editors, *Proceedings of the 10th Alberto Mendelzon International Workshop on Foundations of Data Management, Panama City, Panama, May 8-10, 2016*, volume 1644 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2016.

Giovanni Amendola, Nicola Leone, and Marco Manna. Finite controllability of conjunctive query answering with existential rules: Two steps forward. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 5189–5193. ijcai.org, 2018.

Giovanni Amendola, Nicola Leone, Marco Manna, and Pierfrancesco Veltri. Enhancing existential rules by closed-world variables. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 1676–1682. ijcai.org, 2018.

Medina Andresel, Magdalena Ortiz, and Mantas Simkus. Query rewriting for ontology-mediated conditional answers. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 2734–2741. AAAI Press, 2020.

Labinot Bajraktari, Magdalena Ortiz, and Mantas Simkus. Combining rules and ontologies into clopen knowledge bases. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 1728–1735. AAAI Press, 2018.

Bartosz Bednarczyk and Emanuel Kieronski. Finite entailment of local queries in the Z family of description logics. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 5487–5494. AAAI Press, 2022.

Michael Benedikt and Pierre Bourhis. PSPACE hardness of mixed world query answering for atomic queries under guarded tgds. Technical report, Technical report, University of Oxford, 2018.

Meghyn Bienvenu and Pierre Bourhis. Mixed-world reasoning with existential rules under active-domain semantics.

In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 1558–1565. ijcai.org, 2019.

Piero A. Bonatti, Carsten Lutz, and Frank Wolter. The complexity of circumscription in DLs. *J. Artif. Intell. Res.*, 35:717–773, 2009.

Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz. Datalog extensions for tractable query answering over ontologies. In Roberto De Virgilio, Fausto Giunchiglia, and Letizia Tanca, editors, *Semantic Web Information Management - A Model-Based Perspective*, pages 249–279. Springer, 2009.

Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Ontologies and databases: The DL-Lite approach. In Sergio Tessaris, Enrico Franconi, Thomas Eiter, Claudio Gutierrez, Siegfried Handschuh, Marie-Christine Rousset, and Renate A. Schmidt, editors, *Reasoning Web. Semantic Technologies for Information Systems, 5th International Summer School 2009, Brixen-Bressanone, Italy, August 30 - September 4, 2009, Tutorial Lectures*, volume 5689 of *Lecture Notes in Computer Science*, pages 255–356. Springer, 2009.

Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati. View-based query answering in description logics: Semantics and complexity. *J. Comput. Syst. Sci.*, 78(1):26–46, 2012.

Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz. A general Datalog-based framework for tractable query answering over ontologies. *Journal of Web Semantics*, 14:57–83, 2012. Special Issue on Dealing with the Messiness of the Web of Data.

Ryszard Danecki. Propositional dynamic logic with strong loop predicate. In *International Symposium on Mathematical Foundations of Computer Science*, pages 573–581. Springer, 1984.

Francesco M. Donini, Daniele Nardi, and Riccardo Rosati. Description logics of minimal knowledge and negation as failure. *ACM Trans. Comput. Log.*, 3(2):177–225, 2002.

Thomas Eiter, Carsten Lutz, Magdalena Ortiz, and Mantas Šimkus. Query answering in description logics with transitive roles. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, IJCAI'09, page 759–764, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc.

Diego Figueira, Santiago Figueira, and Edwin Pin Baque. Finite controllability for ontology-mediated query answering of CRPQ. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 381–391, 2020.

Michael Gelfond and Halina Przymusinska. Negation as failure: Careful closure procedure. *Artif. Intell.*, 30(3):273–288, December 1986.

Michael Gelfond, Halina Przymusinska, and Teodor Przymusinski. On the relationship between circumscription and negation as failure. *Artificial Intelligence*, 38(1):75–94, 1989.

Birte Glimm, Ian Horrocks, and Ulrike Sattler. Conjunctive query entailment for SHOQ. In Diego Calvanese, Enrico Franconi, Volker Haarslev, Domenico Lembo, Boris Motik, Anni-Yasmin Turhan, and Sergio Tessaris, editors, *Proceedings of the 2007 International Workshop on Description Logics (DL2007), Brixen-Bressanone, near Bozen-Bolzano, Italy, 8-10 June, 2007*, volume 250 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.

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

Tomasz Gogacz and Jerzy Marcinkowski. Converging to the chase – a tool for finite controllability. *Journal of Computer and System Sciences*, 83(1):180–206, 2017.

Tomasz Gogacz, Yazmín Angélica Ibáñez-García, and Filip Murlak. Finite query answering in expressive description logics with transitive roles. In *International Conference on Principles of Knowledge Representation and Reasoning*, 2018.

Tomasz Gogacz, Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Jean Christoph Jung, and Filip Murlak. On finite and unrestricted query entailment beyond SQ with number restrictions on transitive roles. In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 1719–1725. ijcai.org, 2019.

Tomasz Gogacz, Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Filip Murlak, Magdalena Ortiz, and Mantas Simkus. Ontology focusing: Knowledge-enriched databases on demand. In Giuseppe De Giacomo, Alejandro Catalá, Bistra Dilkina, Michela Milano, Senén Barro, Alberto Bugarín, and Jérôme Lang, editors, *ECAI 2020 - 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain, August 29 - September 8, 2020 - Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020)*, volume 325 of *Frontiers in Artificial Intelligence and Applications*, pages 745–752. IOS Press, 2020.

Tomasz Gogacz, Sanja Lukumbuzya, Magdalena Ortiz, and Mantas Simkus. Datalog rewritability and data complexity of ALCHOIF with closed predicates. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 434–444, 2020.

Tomasz Gogacz, Filip Murlak, Marcin Przybyłko, Alexandra Rogova, and Michał Skrzypczak. Partially finite model reasoning in description logics (extended version). *CoRR*, abs/2604.25549, 2026.

Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Jean Christoph Jung, and Filip Murlak. Answering regular path queries mediated by unrestricted SQ ontologies. *Artif. Intell.*, 314:103808, 2023.

Yazmín Ibáñez-García, Jean Christoph Jung, Vincent Michielini, and Filip Murlak. Revisiting conjunctive query entailment for S. In Sven Koenig, Chad Jenkins, and Matthew E. Taylor, editors, *Fortieth AAAI Conference on Artificial Intelligence, Thirty-Eighth Conference on Innovative Applications of Artificial Intelligence, Sixteenth Symposium on Educational Advances in Artificial Intelligence, AAAI 2026, Singapore, January 20-27, 2026*, pages 19152–19159. AAAI Press, 2026.

H. Katsuno. Closed world assumptions having precedence in predicates. *New Gener. Comput.*, 8(3):185–209, 1990.

Matthias Knorr, José Júlio Alferes, and Pascal Hitzler. Local closed world reasoning with description logics under the well-founded semantics. *Artif. Intell.*, 175(9-10):1528–1554, 2011.

Vladimir Lifschitz. Nonmonotonic databases and epistemic queries. In John Mylopoulos and Raymond Reiter, editors, *Proceedings of the 12th International Joint Conference on Artificial Intelligence. Sydney, Australia, August 24-30, 1991*, pages 381–386. Morgan Kaufmann, 1991.

Sanja Lukumbuzya, Magdalena Ortiz, and Mantas Simkus. Resilient logic programs: Answer set programs challenged by ontologies. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 2917–2924. AAAI Press, 2020.

Carsten Lutz, Inanç Seylan, and Frank Wolter. The data complexity of ontology-mediated queries with closed predicates. *Log. Methods Comput. Sci.*, 15(3), 2019.

Lorenzo Marconi and Riccardo Rosati. Consistent query answering over existential rules with open and closed predicates. In Toby Walsh, Julie Shah, and Zico Kolter, editors, *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, pages 15083–15091. AAAI Press, 2025.

John McCarthy. Circumscription—a form of non-monotonic reasoning. *Artificial intelligence*, 13(1-2):27–39, 1980.

Boris Motik and Riccardo Rosati. Reconciling description logics and rules. *J. ACM*, 57(5):30:1–30:62, 2010.

Filip Murlak. Finite-model reasoning for graph queries and description logics. In Marco Console and Boris Konev, editors, *Reasoning Web. Declarative Artificial Intelligence: Knowledge, Rules, Logic - 19th International Summer School 2023 Oslo, Norway, September 21-24, 2023, Tutorial Lectures*, volume 15400 of *Lecture Notes in Computer Science*, pages 23–53. Springer, 2023.

Nhung Ngo, Magdalena Ortiz, and Mantas Simkus. Closed predicates in description logics: Results on combined complexity. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016*, pages 237–246. AAAI Press, 2016.

Héctor Pérez-Urbina, Boris Motik, and Ian Horrocks. Rewriting conjunctive queries over description logic knowledge bases. In Klaus-Dieter Schewe and Bernhard Thalheim, editors, *Semantics in Data and Knowledge Bases, Third International Workshop, SDKB 2008, Nantes, France, March 29, 2008, Revised Selected Papers*, volume 4925 of *Lecture Notes in Computer Science*, pages 199–214. Springer, 2008.

Xiaoyu Qin, Xiaowang Zhang, Muhammad Qasim Yasin, Shujun Wang, Zhiyong Feng, and Guohui Xiao. SUMA: A partial materialization-based scalable query answering in OWL 2 DL. *Data Sci. Eng.*, 6(2):229–245, 2021.