How Many Legs Do I Have?
Non-Simple Roles in Number Restrictions Revisited

Yevgeny Kazakov, Ulrike Sattler, and Evgeny Zolin

School of Computer Science, The University of Manchester, UK
{ykazakov,sattler,ezolin}@manchester.ac.uk

Abstract. The Description Logics underpinning OWL impose a well-known syntactic restriction in order to preserve decidability: they do not allow to use non-simple roles—that is, transitive roles or their super-roles—in number restrictions. When modeling composite objects, for example in bio-medical ontologies, this restriction can pose problems.

Therefore, we take a closer look at the problem of counting over non-simple roles. On the one hand, we sharpen the known undecidability results and demonstrate that: (i) for DLs with inverse roles, counting over non-simple roles leads to undecidability even when there is only one role in the language; (ii) for DLs without inverses, two transitive and an arbitrary role are sufficient for undecidability. On the other hand, we demonstrate that counting over non-simple roles does not compromise decidability in the absence of inverse roles provided that certain restrictions on role inclusion axioms are satisfied.

1 Introduction

Recently, Description Logics (DLs) [1] have attracted increasing attention, partially due to their usage as logical underpinning of ontology languages such as OIL, DAML+OIL, and OWL[1] [5]. All these languages are based on DLs of the $\mathcal{SHQ}$ family, which are decidable fragments of first order logic and close relatives of modal logics. In DLs, unary predicates/propositional variables are usually called concepts, binary predicates/modal parameters are called roles and, in a nutshell, $\mathcal{SHQ}$ extends $\mathcal{ALC}$ (a notational variant of multi-modal K) with transitivity and role inclusion axioms and with *number restrictions*: these are concepts of the form $(\leqslant n\,R.C)$ for n a non-negative integer, R a role, and C a possibly complex concept. Number restrictions are heavily used to define concepts, e.g., the following expression makes use of standard DL notation to define the concept Human as featherless bipeds:

$$\texttt{Human} = \texttt{Mammal} \sqcap \forall\texttt{hasPart}.\neg\texttt{Feather} \sqcap (\geqslant 2\,\texttt{hasPart.Leg}) \sqcap (\leqslant 2\,\texttt{hasPart.Leg})$$

We find numerous more convincing yet less readable such applications of number restrictions in bio-informatics and medical applications, e.g., they are used to restrict the number of certain components of proteins [8].

---

[1] OWL comes in three flavours, OWL Lite, OWL DL, and OWL Full. Here, we are only concerned with the first two.

Other heavily used features are the above mentioned *transitivity* and *role inclusion* axioms. They allow to express, e.g., that `hasPart` must be interpreted as a transitive relation (which is closely related to the modal logic K4) and that `hasComponent` implies `hasPart`.

Now ontology design and maintenance is a non-trivial task, especially since ontologies can be quite large: e.g., SNOMED and the National Cancer Institute ontology have over 300,000 resp. 17,000 defined concepts. In order to check for consistency and compute the (implicit) concept hierarchy w.r.t. the subsumption relationship, ontology editors make use of DL reasoners[2] which implement decision procedures for concept satisfiability and subsumption w.r.t. DL axioms. For this to be possible, i.e., for these reasoning problems to be decidable for $\mathcal{SHQ}$, their designer had to impose a syntactic restriction: in number restrictions, one can neither use transitive roles nor super-roles of transitive roles, i.e., number restrictions can only be used on *simple roles*. For example, if we want to make use of our definition of `Human`, we have to either refrain from making `hasPart` a transitive role or use, e.g., a (non-transitive) subrole such as `hasComp` of `hasPart` in its number restrictions. Both options are sub-optimal since they result in the loss of other, useful consequences. For the first option, e.g., we could add the following definition of `HumanBird` without causing a (useful) inconsistency:

$$\texttt{HumanBird} = \texttt{Human} \sqcap \exists\texttt{hasPart}.(\texttt{Wing} \sqcap \exists\texttt{hasPart}.\texttt{Feather}).$$

For the second option, e.g., we could add the following definition of `3LHuman` without causing an inconsistency (please note that here we use twice the sub-role `hasComp` of `hasPart` and only once `hasPart`):

$$\texttt{3LHuman} = \texttt{Human} \sqcap \exists\texttt{hasComp}.(\texttt{Leg} \sqcap \texttt{Left}) \sqcap \exists\texttt{hasComp}.(\texttt{Leg} \sqcap \texttt{Right} \sqcap \neg\texttt{Left})$$
$$\sqcap \exists\texttt{hasPart}.(\texttt{Leg} \sqcap \neg\texttt{Right} \sqcap \neg\texttt{Left}).$$

In [6], it is shown that satisfiability of concepts in $\mathcal{SHQ}$ (even in its sublogic $\mathcal{SHN}$) is undecidable if non-simple roles (i.e., transitive roles or their super-roles) are used in number restrictions. In this paper, we explore this area more thoroughly with the goal of finding a more expressive but still decidable DL where we can use non-simple roles in number restrictions. Our contributions are two-fold: on the one hand, we sharpen the above undecidability result and demonstrate that: (i) for DLs such as $\mathcal{SHIN}$ (which extends $\mathcal{SHN}$ with inverse roles), counting over non-simple roles leads to undecidability even with only one role in the language; (ii) for DLs without inverses such as $\mathcal{SHN}$, two transitive and a third role are sufficient for undecidability. On the other hand, we demonstrate that, in the absence of inverse roles, counting over non-simple roles does not compromise decidability provided that they satisfy certain other restrictions regarding role inclusion axioms. Roughly speaking, as long as any two transitive roles are either completely unrelated w.r.t. inclusion or one of them implies the other, we can use them in number restrictions without losing decidability. We believe that the latter result will turn out to be useful in practice since it allows, for example, to capture a transitive role `hasPart` alongside other, possibly transitive roles such as `hasComp` or `hasSegment` *and* to use them all in number restrictions—as long as any two of these transitive roles are related by a (bi)-implication.

---

[2] See `http://www.cs.man.ac.uk/~sattler/reasoners.html` for a list.

2 Preliminaries and Known Results

The vocabulary of a DL consists of disjoint infinite sets of *concept names* CN, *role names* RN, and *individual names* IN. A *role* is an expression of the form r or r^-, where r is a role name. For convenience, we introduce a syntactic operator defined on roles: $\mathsf{Inv}(R) := r^-$, if R is a role name r; and $\mathsf{Inv}(R) := r$, if $R = r^-$ for some role name r. Finally, we use $\mathsf{Card}(M)$ for the cardinality of a set M.

Definition 1 (RBox). An *RBox* $\mathcal{R}$ is a finite collection of *transitivity axioms* of the form $\mathsf{Tr}(R)$ and *role inclusion axioms* of the form $R \sqsubseteq S$, where R, S are roles.

An *interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ consists of a non-empty set $\Delta^{\mathcal{I}}$, its *domain*, and an interpretation function $\cdot^{\mathcal{I}}$ that maps each role name $r \in \mathsf{RN}$ to a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$; $\mathcal{I}$ is *finite* if the domain of $\mathcal{I}$ is finite. We define $(r^-)^{\mathcal{I}} := \{\langle x, y \rangle \mid \langle y, x \rangle \in r^{\mathcal{I}}\}$. We define whether $\mathcal{I}$ *satisfies* an axiom α, written $\mathcal{I} \models \alpha$ as follows: $\mathcal{I} \models \mathsf{Tr}(R)$ iff $R^{\mathcal{I}}$ is transitive, and $\mathcal{I} \models R \sqsubseteq S$ iff $R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$. An interpretation satisfying all axioms in $\mathcal{R}$ is called a *model* of $\mathcal{R}$. An RBox $\mathcal{R}$ *entails* an axiom α, written $\mathcal{R} \models \alpha$, if all models of $\mathcal{R}$ satisfy α.

The deductive *closure* $[\mathcal{R}]$ of $\mathcal{R}$ is the minimal set that contains $\mathcal{R}$ and axioms $R \sqsubseteq R$, for all roles R in $\mathcal{R}$, and that is closed under the following rules:

$$\frac{R \sqsubseteq S \quad S \sqsubseteq T}{R \sqsubseteq T} \qquad \frac{R \sqsubseteq S}{\mathsf{Inv}(R) \sqsubseteq \mathsf{Inv}(S)} \qquad \frac{T \sqsubseteq S \sqsubseteq T \quad \mathsf{Tr}(T)}{\mathsf{Tr}(S)} \qquad \frac{\mathsf{Tr}(T)}{\mathsf{Tr}(\mathsf{Inv}(T))}$$

We write $\mathcal{R} \vdash \alpha$ as an alternative notation for $\alpha \in [\mathcal{R}]$, where α is an RBox axiom.

Definition 2. The set of concepts in DL $\mathcal{ALCIQ}$ is defined by the grammar:

$$\mathbf{C} ::= \bot \mid A \mid \neg C \mid C \sqcap D \mid \exists R.C \mid \leqslant n\, S.C,$$

where $A \in \mathsf{CN}$, $C, D \in \mathbf{C}$, R and S are roles, and n is a non-negative integer.

The interpretation function $\cdot^{\mathcal{I}}$ maps, additionally, each concept name $C \in \mathsf{CN}$ to a subset $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and $\cdot^{\mathcal{I}}$ is extended to complex concepts inductively as follows:

$$\bot^{\mathcal{I}} = \varnothing, \ (\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}, \ (C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}},$$
$$(\exists R.C)^{\mathcal{I}} = \{\, e \in \Delta^{\mathcal{I}} \mid \text{ there exists } d \in C^{\mathcal{I}} \text{ such that } \langle e, d \rangle \in R^{\mathcal{I}}\,\},$$
$$(\leqslant n\, S.C)^{\mathcal{I}} = \{\, e \in \Delta^{\mathcal{I}} \mid \mathsf{Card}(\{d \in C^{\mathcal{I}} \mid \langle e, d \rangle \in S^{\mathcal{I}}\}) \leqslant n\,\}.$$

For C and D $\mathcal{ALCIQ}$ concepts, $C \sqsubseteq D$ is a *general concept inclusion* (GCI), and a finite set of GCIs is called a *TBox*. An interpretation $\mathcal{I}$ satisfies a GCI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$. An interpretation is a *model* of a TBox if it satisfies all its axioms. If a interpretation $\mathcal{I}$ is a model of an RBox $\mathcal{R}$ and a TBox $\mathcal{T}$, then we say that $\mathcal{I}$ is a *model* of $\langle \mathcal{R}, \mathcal{T} \rangle$, or $\langle \mathcal{R}, \mathcal{T} \rangle$ is *satisfiable*. A concept C is *satisfiable w.r.t.* $\langle \mathcal{R}, \mathcal{T} \rangle$ if there exists a model $\mathcal{I}$ of $\langle \mathcal{R}, \mathcal{T} \rangle$ such that $C^{\mathcal{I}} \neq \varnothing$.

As usual, the concept expressions $\top, C_1 \sqcup C_2, \forall R.C$ and $\geqslant n\, S.C$ are assumed to be abbreviations for $\neg \bot, \neg(\neg C_1 \sqcap \neg C_2), \neg(\exists R.\neg C)$ and $\neg(\leqslant (n-1)\, S.\neg C)$ respectively. Concepts of $\mathcal{ALCIQ}$ that do not use number restrictions $(\leqslant n\, R.C)$, or inverse roles, or both, will be called $\mathcal{ALCI}$-, $\mathcal{ALCQ}$-, and $\mathcal{ALC}$ concepts, resp. The letter $\mathcal{N}$ in the name of a DL indicates that this DL supports only number restrictions of the form $(\leqslant n\, R.\top)$.

Please note that, so far, we have introduced RBoxes and $\mathcal{ALCIQ}$ TBoxes separately, i.e., we did not put them into a single logic, which is slightly unusual. Recall that in [6] a role S is called *simple w.r.t.* $\mathcal{R}$ if there is no transitive subrole of S in $\mathcal{R}$. Traditionally, the DL that allows for

- an RBox without inverse roles and an $\mathcal{ALCQ}$ TBox where all roles in number restrictions are simple is called $\mathcal{SHQ}$, and
- an RBox and an $\mathcal{ALCIQ}$ TBox where all roles in number restrictions are simple is called $\mathcal{SHIQ}$.

For $\mathcal{SHIQ}$ and related DLs, roles in number restrictions are restricted to simple ones to ensure decidability of concept satisfiability w.r.t. a TBox and an RBox: in $\mathcal{SHN}$ (and hence $\mathcal{SHIQ}$), non-simple roles in number restrictions lead to the undecidability of the satisfiability problem [6]. Our aim is to find conditions under which we can relax or even get rid of this restriction to simple roles in number restrictions while preserving decidability. This aim can be achieved by extending the notion of a simple role in such a way that it covers, besides roles that are usually called simple, also some transitive roles or their super-roles. In this paper, we focus on a sub-problem, namely, we are looking for conditions on an RBox under which one can use all its roles in number restrictions and still have a decidable logic. Therefore, we introduce the following notion.

Definition 3. Let $\mathcal{L}$ be a logic between $\mathcal{ALC}$ and $\mathcal{ALCIQ}$ and $\mathcal{R}$ an RBox. The problem of $\mathcal{L}(\mathcal{R})$-satisfiability is to determine, given an $\mathcal{L}$-concept C and an $\mathcal{L}$-TBox $\mathcal{T}$, whether C is satisfiable w.r.t. $\langle \mathcal{R}, \mathcal{T} \rangle$. We say that an RBox $\mathcal{R}$ is $\mathcal{L}$-*safe* (or *safe for* $\mathcal{L}$) if $\mathcal{L}(\mathcal{R})$-satisfiability is decidable, and $\mathcal{L}$-*unsafe* otherwise.

Any RBox is $\mathcal{ALCI}$-safe because (i) neither $\mathcal{ALCI}$ nor $\mathcal{SHI}$ support number restrictions, and (ii) since a concept C and a TBox $\mathcal{T}$ are $\mathcal{ALCI}(\mathcal{R})$-satisfiable iff C is satisfiable w.r.t. $\langle \mathcal{R}, \mathcal{T} \rangle$, we have that $\mathcal{ALCI}(\mathcal{R})$ satisfiability can be viewed as the standard $\mathcal{SHI}$ satisfiability problem which is known to be decidable [6]. With a similar argument, any RBox $\mathcal{R}$ without transitivity axioms is $\mathcal{ALCIQ}$-safe because (i) all roles are simple in this case, and (ii) $\mathcal{ALCIQ}(\mathcal{R})$-satisfiability can be viewed as the standard $\mathcal{SHIQ}$ satisfiability problem which is known to be decidable [6]. There are numerous other restrictions on the syntax that could possibly lead to decidability, for example to use only number restrictions of the form $(\leqslant 1\, R)$.

At the same time, we know from [6] that the following RBox Star_4 (with eight roles, of which four are transitive) is $\mathcal{ALCN}$-unsafe:

$$\mathsf{Star}_4 = \{\, s_i \sqsubseteq t_{ij},\, r_j \sqsubseteq t_{ij},\, \mathsf{Tr}(t_{ij}) \mid 0 \leqslant i, j \leqslant 1 \,\}.$$

In what follows, we show that

- there is a large class of RBoxes involving role inclusions and transitivity axioms that are $\mathcal{ALCQ}$-safe (Theorem 4),
- there exists an $\mathcal{ALCIN}$-unsafe RBox with only one transitive role (Theorem 2),
- there exist $\mathcal{ALCN}$-unsafe RBoxes involving only three roles (Theorem 1).

Many proofs in this paper are rather sketchy or completely omitted; all the details however can be found in the accompanying technical report [7].
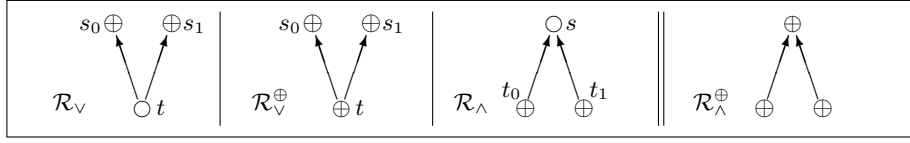
Fig. 1. The first three RBoxes are unsafe (Theorem 1) and for the last one the problem is open.

3 Undecidability Results

Here we show that three roles are sufficient for building an unsafe RBox for $\mathcal{ALCQ}$, whereas for $\mathcal{ALCIQ}$, even one role is sufficient for that. In order to provide a geometric intuition for our results, we depict RBoxes (without inverse roles) as directed graphs whose nodes are non-transitive ($\bigcirc$) and transitive ($\oplus$) roles and arrows represent implications between roles. Our plan is as follows. First, in Theorem 1 we show that the RBoxes $\mathcal{R}_\vee$, $\mathcal{R}_\vee^\oplus$, and $\mathcal{R}_\wedge$ shown in Fig. 1 are $\mathcal{ALCQ}$-unsafe; we give only a sketch of the proof to illustrate the idea. Our conjecture is that the fourth RBox $\mathcal{R}_\wedge^\oplus$ in Fig. 1 is $\mathcal{ALCQ}$-safe. Next, we formulate a more general result (Theorem 2) that the RBoxes depictured in Fig. 3 are also $\mathcal{ALCQ}$-unsafe (its fully detailed proof is given in the accompanying technical report [7]). Finally, in Theorem 3 we demonstrate that the RBox $\{\mathsf{Tr}(r)\}$ is $\mathcal{ALCIQ}$-unsafe. We obtain the undecidability results by reduction from the undecidable *domino problem* (see, e.g., [3]).

Definition 4 (Domino). A *domino system* is a triple $\mathcal{D} = \langle D, H, V \rangle$, where $D = \{d_1, \ldots, d_n\}$ is a finite set of *tile types* and $H, V \subseteq D \times D$ are horizontal and vertical *matching* relations. We say that $\mathcal{D}$ *tiles* $\mathbb{N} \times \mathbb{N}$ if there exists a $\mathcal{D}$-tiling, i.e., a mapping $\tau \colon \mathbb{N} \times \mathbb{N} \to D$ such that, for all $i, j \in \mathbb{N}$, the following *compatibility* conditions hold: $\langle \tau(i,j), \tau(i+1,j) \rangle \in H$ and $\langle \tau(i,j), \tau(i,j+1) \rangle \in V$. The *domino problem* is to check, given a domino system $\mathcal{D}$, whether $\mathcal{D}$ tiles $\mathbb{N} \times \mathbb{N}$.

Our proofs follow the usual pattern: in order to show $\mathcal{L}$-unsafety of some RBox $\mathcal{R}$, we first build an $\mathcal{L}$-TBox $\mathcal{T}_{\mathsf{grid}}$ that, together with $\mathcal{R}$ "encodes", the $\mathbb{N} \times \mathbb{N}$ grid. Then, given a domino system $\mathcal{D}$, we build (efficiently) an $\mathcal{ALC}$-TBox $\mathcal{T}_\mathcal{D}$ that "tiles" the grid and "ensures" the compatibility conditions. Finally, we prove that $\mathcal{D}$ tiles $\mathbb{N} \times \mathbb{N}$ iff some concept (usually a concept name) C is satisfiable w.r.t. $\mathcal{R} \cup \mathcal{T}_{\mathsf{grid}} \cup \mathcal{T}_\mathcal{D}$. We give fine-grained formulations of results by indicating, as a subscript to the name of a logic, the maximal number n occurring in number restrictions ($\leqslant n\, R.C$) in the proof.

Theorem 1. *The RBoxes $\mathcal{R}_\wedge$, $\mathcal{R}_\vee$, and $\mathcal{R}_\vee^\oplus$ shown in Fig. 1 are unsafe for $\mathcal{ALCN}$; more precisely, they are unsafe for $\mathcal{ALCN}_9$ and $\mathcal{ALCQ}_1$.*

Proof (sketch for $\mathcal{R}_\wedge$). We use 16 concept names A_{ij}, $0 \leqslant i, j \leqslant 3$, place them on an $\mathbb{N} \times \mathbb{N}$ grid (by repeating a $[0,3] \times [0,3]$ pattern periodically) and link them with t_0- and t_1-edges as shown in Fig. 2a. We will refer to edges in this grid as $\langle A, r, B \rangle$, where $A, B \in \{A_{ij} \mid 0 \leqslant i, j \leqslant 3\}$ and $r \in \{t_0, t_1\}$. Having this picture in mind, we add the following axioms (a)–(c), (d) to an $\mathcal{ALCQ}$-TBox $\mathcal{T}_{\mathsf{grid}}^\mathcal{Q}$ and axioms (a)–(c), (d') to an $\mathcal{ALCN}$-TBox $\mathcal{T}_{\mathsf{grid}}^\mathcal{N}$, where $i, j, k, \ell \in \{0, 1, 2, 3\}$ and we denote $i \oplus j := (i+j) \bmod 4$.
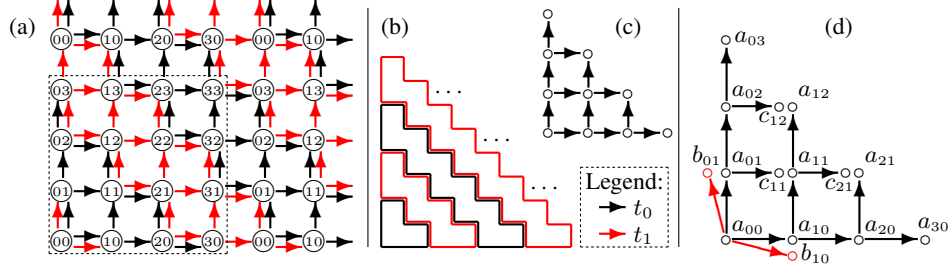
Fig. 2. A grid for Theorem 1: (a) A detailed view of the grid. (b) A grid at a glance. (c) Accessibility relation for R (and similarly for S). (d) A pre-grid for $\mathcal{ALCN}$.

(a) $A_{ij} \sqcap A_{k\ell} \sqsubseteq \bot$, for all $\langle i,j \rangle \neq \langle k,\ell \rangle$, i.e., concepts A_{ij} are pairwise disjoint;
(b) $A \sqsubseteq \exists r.B$ for each edge $\langle A, r, B \rangle$, where $r \in \{t_0, t_1\}$;
(c) $A \sqsubseteq \, \leqslant 1 \, s.B$ for each double edge $\langle A, t_0, B \rangle$ and $\langle A, t_1, B \rangle$;
(d) $A_{ij} \sqsubseteq \, \leqslant 1 \, s.A_{i \oplus 1, j \oplus 1}$ for all $0 \leqslant i,j \leqslant 3$;
(d') $A_{ij} \sqsubseteq \, \leqslant 9 \, s$ for all $0 \leqslant i,j \leqslant 3$ such that $i+j$ is even.

For instance, we have axioms $A_{10} \sqsubseteq \exists t_0.A_{11}$ from (b), $A_{11} \sqsubseteq \, \leqslant 1 \, s.A_{12}$ from (c), $A_{32} \sqsubseteq \, \leqslant 1 \, s.A_{03}$ from (d) in $\mathcal{T}_{\text{grid}}^{\mathcal{Q}}$, and $A_{13} \sqsubseteq \, \leqslant 9 \, s$ from (d') in $\mathcal{T}_{\text{grid}}^{\mathcal{N}}$.

Next, given a domino system $\mathcal{D} = \langle D, H, V \rangle$ with $D = \{d_1, \ldots, d_n\}$, we introduce fresh concept names $D_1, \ldots, D_n$ and add the following $\mathcal{ALC}$-axioms to a TBox $\mathcal{T}_{\mathcal{D}}$:

(e) $\top \sqsubseteq D_1 \sqcup \ldots \sqcup D_n$;
(f) $D_k \sqcap D_\ell \sqsubseteq \bot$, for all $1 \leqslant k < \ell \leqslant n$;
(g) $A \sqcap D_k \sqsubseteq \forall r. \big(B \to \bigsqcup_{\ell: \langle d_k, d_\ell \rangle \in H} D_\ell \big)$ for each horizontal edge $\langle A, r, B \rangle$;
(h) $A \sqcap D_k \sqsubseteq \forall r. \big(B \to \bigsqcup_{\ell: \langle d_k, d_\ell \rangle \in V} D_\ell \big)$ for each vertical edge $\langle A, r, B \rangle$.

Now, for $\mathcal{X} \in \{\mathcal{Q}, \mathcal{N}\}$, we set $\mathcal{K}_{\mathcal{D}}^{\mathcal{X}} := \mathcal{R}_\wedge \cup \mathcal{T}_{\text{grid}}^{\mathcal{X}} \cup \mathcal{T}_{\mathcal{D}}$ and prove the following lemma.

Lemma 1.1 (For $\mathcal{R}_\wedge$). *The concept A_{00} is satisfiable w.r.t. $\mathcal{K}_{\mathcal{D}}^{\mathcal{X}}$ iff $\mathcal{D}$ tiles $\mathbb{N} \times \mathbb{N}$.*

($\Leftarrow$) Given a tiling $\tau \colon \mathbb{N} \times \mathbb{N} \to D$, we build a model $\mathcal{I}$ as follows: set $\Delta^{\mathcal{I}} := \mathbb{N} \times \mathbb{N}$, interpret concepts A_{ij} exactly as in Fig. 2a; roles t_0, t_1 as the transitive closures of the relations depicted by arrows in Fig. 2a; let $s^{\mathcal{I}} := t_0^{\mathcal{I}} \cup t_1^{\mathcal{I}}$; and set $\langle i, j \rangle \in D_k^{\mathcal{I}}$ iff $\tau(i,j) = d_k$. Then $A_{00} \neq \varnothing$, as $\langle 0, 0 \rangle \in A_{00}^{\mathcal{I}}$.

It remains to check that $\mathcal{I} \models \mathcal{K}_{\mathcal{D}}^{\mathcal{X}}$. Clearly, $\mathcal{I} \models \mathcal{R}_\wedge$; here it is important that $\mathcal{R}_\wedge$ has no transitivity axiom for s, as the relation $s^{\mathcal{I}}$ is not transitive. To show that $\mathcal{I} \models \mathcal{T}_{\text{grid}}^{\mathcal{X}}$, one needs the following observation: any element in $\mathcal{I}$ has at most 9 s-successors, which belong to 9 pairwise disjoint sets $A_{ij}^{\mathcal{I}}$. Finally, it is straightforward to show that $\mathcal{I} \models \mathcal{T}_{\mathcal{D}}$.

($\Rightarrow$) Suppose that $\mathcal{I} \models \mathcal{K}_{\mathcal{D}}^{\mathcal{X}}$ and $A_{00}^{\mathcal{I}}$ is nonempty, say $a_{00} \in A_{00}^{\mathcal{I}}$. Then we show that there exist (not necessarily distinct) elements $a_{ij} \in \Delta^{\mathcal{I}}$, for all $i, j \in \mathbb{N}$, that are linked with t_0- and t_1-edges as in Fig. 2a. After that, we "read-off" a $\mathcal{D}$-tiling of $\mathbb{N} \times \mathbb{N}$.

Since $\mathcal{I}$ satisfies axiom (b), we can find in $\mathcal{I}$ a "pre-grid" with the root a_{00}, i.e., elements a_{ij}, b_{ij}, c_{ij} linked with t_0- and t_1-edges as shown in Fig. 2d. Axioms (c) allow to "glue" double edges, i.e., entail that $b_{10} = a_{10}$ and $b_{01} = a_{01}$. Now, with the help of axioms (d), we "glue" cells, i.e., infer that $c_{11} = a_{11}$, $c_{12} = a_{12}$, and $c_{21} = a_{21}$. Thus,
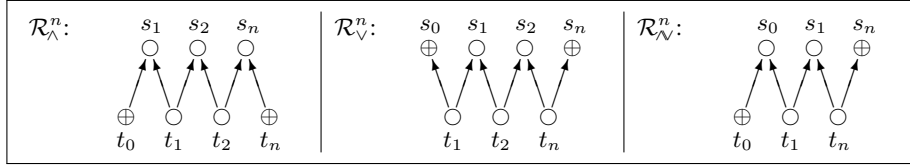
Fig. 3. The chain RBoxes $\mathcal{R}_\wedge^n$, $\mathcal{R}_\vee^n$, $\mathcal{R}_\mathcal{N}^n$, for $n \geqslant 1$, are $\mathcal{ALCN}$-unsafe.

using axioms from $\mathcal{T}_{\text{grid}}^\mathcal{Q}$, we have found in $\mathcal{I}$ a structure depicted in Fig. 2c. In $\mathcal{T}_{\text{grid}}^\mathcal{N}$, axiom (d') applied to the element a_{00}, together with the fact that all $A_{ij}^\mathcal{I}$ are disjoint, entail all the above equalities and hence ensure the existence of the same structure Fig. 2c in $\mathcal{I}$. After that, we repeat the same argument, starting at the root a_{ij} with $i + j = 2$, then with $j + i = 4$ and so on.

Once we have built all the elements a_{ij}, for $i, j \in \mathbb{N}$ we define $\tau \colon \mathbb{N} \times \mathbb{N} \to D$ by setting $\tau(i, j) := d_k$ iff $a_{ij} \in D_k^\mathcal{I}$. By (e) and (f), τ is well-defined and total; and the compatibility conditions easily follow from (g) and (h). Thus τ is indeed a $\mathcal{D}$-tiling of $\mathbb{N} \times \mathbb{N}$. This completes the proof of Lemma 1.1 and hence of Theorem 1. $\dashv$

In order to generalise Theorem 1, we introduce the following RBoxes, depicted in Fig. 3, where $n \geqslant 1$ (observe that $\mathcal{R}_\wedge^1$ and $\mathcal{R}_\vee^1$ correspond to $\mathcal{R}_\wedge$ and $\mathcal{R}_\vee$, resp.):

$$\mathcal{R}_\wedge^n := \{\,\mathsf{Tr}(t_0), \mathsf{Tr}(t_n)\,\} \cup \{\,t_{k-1} \sqsubseteq s_k,\, s_k \sqsupseteq t_k \mid 1 \leqslant k \leqslant n\,\},$$
$$\mathcal{R}_\vee^n := \{\,\mathsf{Tr}(s_0), \mathsf{Tr}(s_n)\,\} \cup \{\,s_{k-1} \sqsupseteq t_k,\, t_k \sqsubseteq s_k \mid 1 \leqslant k \leqslant n\,\},$$
$$\mathcal{R}_\mathcal{N}^n := \{\,\mathsf{Tr}(t_0), \mathsf{Tr}(s_n)\,\} \cup \{\,s_{k-1} \sqsupseteq t_k,\, t_\ell \sqsubseteq s_\ell \mid 1 \leqslant k \leqslant n,\, 0 \leqslant \ell \leqslant n\,\}.$$

Theorem 2. *The RBoxes $\mathcal{R}_\wedge^n$, $\mathcal{R}_\vee^n$, $\mathcal{R}_\mathcal{N}^n$, with $n \geqslant 1$, are $\mathcal{ALCN}$-unsafe.*

Theorem 3. *The RBox $\mathcal{R} := \{\mathsf{Tr}(r)\}$ is unsafe for $\mathcal{ALCIN}$ (more precisely, for $\mathcal{ALCIN}_8$ and $\mathcal{ALCIQ}_1$), even for TBoxes that involve a single role name r.*

Proof. By reduction from the undecidable domino problem for $\mathbb{Z} \times \mathbb{Z}$, which is formulated analogously to Def. 4. Take 16 concept names A_{ij}, $0 \leqslant i, j \leqslant 3$. Place them on the $\mathbb{Z} \times \mathbb{Z}$ grid (by repeating a $[0, 3] \times [0, 3]$ pattern periodically) and link them with r-edges in accordance with Fig. 4a. Now, having this picture in mind (we refer to its edges as $\langle A, r, B \rangle$, where A, B are concept names), we add the following axioms (a)–(c) to an $\mathcal{ALCIQ}$-TBox $\mathcal{T}_{\text{grid}}^\mathcal{Q}$ and axioms (a) and (d) to an $\mathcal{ALCIN}$-TBox $\mathcal{T}_{\text{grid}}^\mathcal{N}$:

(a) All 16 concept names A_{ij}, $0 \leqslant i, j \leqslant 3$, are pairwise disjoint;
(b) $A \sqsubseteq \exists r.B$ and $B \sqsubseteq \exists r^-.A$, for each edge $\langle A, r, B \rangle$;
(c) $A_{ij} \sqsubseteq\, \leqslant 1\, r.A_{k\ell}$ and $A_{k\ell} \sqsubseteq\, \leqslant 1\, r^-.A_{ij}$, for all even i, j and odd k, ℓ;
(d) $A_{ij} \sqsubseteq\, \leqslant 8\, r$ and $A_{k\ell} \sqsubseteq\, \leqslant 8\, r^-$, for all even i, j and odd k, ℓ.

Given a domino system $\mathcal{D}$, we build a $\mathcal{ALC}$-TBox $\mathcal{T}_\mathcal{D}$: axioms (e) and (f) are the same as in the proof of Theorem 1, whereas (g) is the following (and (h) is analogous):

(g) $A \sqcap D_k \sqsubseteq \forall r.\bigl(B \to \bigsqcup_{\ell \colon \langle d_k, d_\ell \rangle \in H} D_\ell\bigr)$ for each right-going edge $\langle A, r, B \rangle$;
    $A \sqcap D_\ell \sqsubseteq \forall r.\bigl(B \to \bigsqcup_{k \colon \langle d_k, d_\ell \rangle \in H} D_k\bigr)$ for each left-going edge $\langle A, r, B \rangle$.
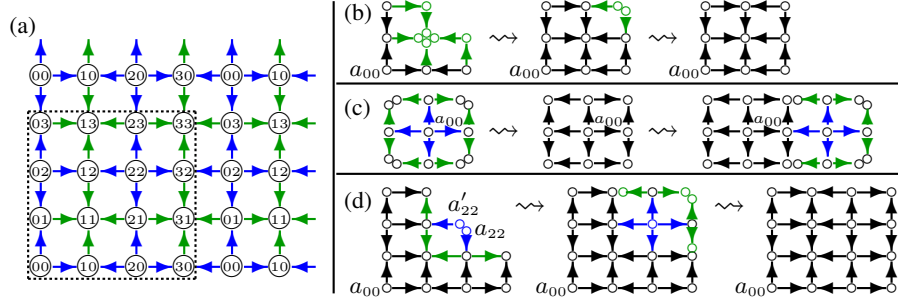
Fig. 4. (a) A grid for Theorem 3 (coloured for decoration only, as we have only 1 role). (b) A pre-grid for $\mathcal{ALCIQ}$. (c) Building a horisontal axis in $\mathcal{ALCIN}$. (d) Building new cells in $\mathcal{ALCIN}$.

Finally, for each $\mathcal{X} \in \{\mathcal{Q}, \mathcal{N}\}$, we set $\mathcal{K}_\mathcal{D}^\mathcal{X} := \mathcal{R} + \mathcal{T}_{\text{grid}}^\mathcal{X} + \mathcal{T}_\mathcal{D}$. It remains to prove

Lemma 3.1. *The concept A_0 is satisfiable w.r.t. $\mathcal{K}_\mathcal{D}^\mathcal{X}$ iff $\mathcal{D}$ tiles $\mathbb{Z} \times \mathbb{Z}$.*

The '$\Leftarrow$' part is proved as in Theorem 1. To prove '$\Rightarrow$', suppose that $\mathcal{I} \models \mathcal{K}_\mathcal{D}^\mathcal{X}$ and $a_{00} \in A_{00}^\mathcal{I}$. Then we show that in $\mathcal{I}$ there are (not necessarily distinct) elements a_{ij}, $i, j \in \mathbb{Z}$, linked via r-edges as shown in Fig. 4a, and then build a $\mathcal{D}$-tiling of $\mathbb{Z} \times \mathbb{Z}$. The steps of building elements a_{ij} are illustrated in Fig. 4(b–d); we omit the details, which can be found in the technical report [7]. $\dashv$

4 Internalization of RBoxes in TBoxes Using Extended Roles

In order to study safety of RBoxes for different DLs, it is somewhat inconvenient to work separately with RBoxes and TBoxes. Therefore, in this section, we demonstrate how RBoxes can be internalized into TBoxes, provided additional role constructors—role unions and transitive closure operator—can be used. We also demonstrate that it is sufficient to focus only on TBoxes of some simple form. The results of this section can be applied to any logic $\mathcal{L}$ between $\mathcal{ALC}$ and $\mathcal{ALCIQ}$.

Definition 5. We say that an $\mathcal{L}$-TBox $\mathcal{T}$ is in a *simple form* if all axioms in $\mathcal{T}$ have the following forms, where $A_{(i)}$, $B_{(j)}$ are concept names, m, n integers, and S a role:

$$\prod A_i \sqcap \prod \neg B_j \sqsubseteq \bot \tag{1}$$

$$A \sqsubseteq \geqslant n\, S.B \tag{2}$$

$$A \sqsubseteq \leqslant m\, S.B \tag{3}$$

Lemma 1 (Simplification of $\mathcal{L}$-TBoxes). *Given an $\mathcal{L}$-TBox $\mathcal{T}$, one can construct in polynomial time an $\mathcal{L}$-TBox $\mathcal{T}_{sf}$ in simple form such that, for every RBox $\mathcal{R}$, $\langle \mathcal{T}, \mathcal{R} \rangle$ is (finitely) satisfiable iff $\langle \mathcal{R}, \mathcal{T}_{sf} \rangle$ is (finitely) satisfiable.*

Definition 6. The set of *extended roles* $\mathbf{R}^{\sqcup,+}$ is defined by the following grammar:

$$\mathbf{R}^{\sqcup,+} ::= R \mid \rho_1 \sqcup \rho_2 \mid \rho^+, \quad \text{where } R \text{ is a role and } \rho_{(i)} \in \mathbf{R}^{\sqcup,+}.$$

The additional role constructors are interpreted as follows: $(\rho_1 \sqcup \rho_2)^\mathcal{I} = \rho_1^\mathcal{I} \cup \rho_2^\mathcal{I}$, $(\rho^+)^\mathcal{I} = (\rho^\mathcal{I})^+$, where $(\cdot) \cup (\cdot)$ and $(\cdot)^+$ are usual operators of union and transitive closure on binary relations. Concepts of $\mathcal{L}(\sqcup, +)$ are defined as for $\mathcal{L}$ except that extended roles can be used in place of roles.

Our goal is to demonstrate that every RBox can be internalized in a simple $\mathcal{L}$-TBox producing an $\mathcal{L}(\sqcup, +)$-TBox of a certain simple form:

Definition 7 (Simple $\mathcal{L}(\sqcup, +)$-TBox). We say that an $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{T}$ is *simple* if every axiom from $\mathcal{T}$ is either of the form (1), (2), or:

$$A \sqsubseteq \, \leqslant m \, (\bigsqcup u_i^+ \sqcup v).B \tag{4}$$

where $A_{(i)}$, $B_{(j)}$ are concept names, m, n integers, and u_i and v are disjunctions of roles: $u_i, v = \bigsqcup R_i$. For a simple TBox $\mathcal{T}$, we denote by $K(\mathcal{T})$ the number of axioms of type (4) in $\mathcal{T}$, by $N(\mathcal{T})$ and $M(\mathcal{T})$ the sum of all numbers n, resp. m, over all axioms of type (2), resp. (4), by $C(\mathcal{T})$ the number of concept names in $\mathcal{T}$.

Definition 8 ($\mathcal{R}$-extension). Given an RBox $\mathcal{R}$, an *extension* of a role S in $\mathcal{R}$ (or the *$\mathcal{R}$-extension* of S, for short) is an extended role $\mathcal{R}(S) \in \mathbf{R}^{\sqcup, +}$ defined as follows:

– If S is transitive in $\mathcal{R}$ then $\mathcal{R}(S) := (\bigsqcup S_i)^+$, where $\{S_i\}$ is the set of all subroles of S in $\mathcal{R}$ (including S itself);
– If S is not transitive, then $\mathcal{R}(S) := \bigsqcup \mathcal{R}(T_i) \sqcup \bigsqcup S_j$, where $\{T_i\}$ is exactly the set of all maximal transitive subroles of S, and $\{S_j\}$ is the set of all subroles of S.

Definition 9 (Internalization of an RBox in an $\mathcal{L}$-TBox).
Let $\mathcal{R}$ be an RBox and $\mathcal{T}$ be a simple $\mathcal{L}$-TBox. The internalization of $\mathcal{R}$ in $\mathcal{T}$ is a simple $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{R}(\mathcal{T}) := \{\mathcal{R}(\alpha) \mid \alpha \in \mathcal{T}\}$, where:

– $\mathcal{R}(\alpha) := \alpha$ if α is of the form (1) or (2), and
– $\mathcal{R}(\alpha) := A \sqsubseteq \, \leqslant m \, (\mathcal{R}(S)).B$ if $\alpha = A \sqsubseteq \, \leqslant m \, S.B$ is of the form (3).

Lemma 2. *Let $\mathcal{R}$ be an RBox and $\mathcal{T}$ a simple $\mathcal{L}$-TBox. Then $\langle \mathcal{R}, \mathcal{T} \rangle$ is satisfiable iff $\mathcal{R}(\mathcal{T})$ is satisfiable.*

5 Decidability Results

As we have demonstrated in Theorem 3, an RBox consisting of just one transitivity axiom is already unsafe for $\mathcal{ALCIN}$. Hence, there is a little room left for non-trivial safe RBoxes for $\mathcal{ALCIN}$. In contrast, the undecidability results in Section 3 for $\mathcal{ALCN}$ require a certain interaction between several transitive roles. This poses a question about safety of those RBoxes that do not fit such a pattern. In this section, we investigate this question and define a relatively large class of so-called admissible RBoxes that, as we will prove, are safe for $\mathcal{ALCQ}$. Since we focus on $\mathcal{ALCQ}$, within this section we assume that there are no inverse roles in RBoxes.

Definition 10. For a TBox $\mathcal{T}$, RBox $\mathcal{R}$, or an axiom α, let $\mathsf{RN}(\mathcal{T})$, $\mathsf{RN}(\mathcal{T})$, $\mathsf{RN}(\alpha)$ denote the set of role names that occur in $\mathcal{T}$, $\mathcal{R}$, α, respectively.

An RBox $\mathcal{R}$ is *strongly admissible* if, for every two transitive roles $T_1, T_2 \in \mathsf{RN}(\mathcal{R})$, we have $\mathcal{R} \vdash T_1 \sqsubseteq T_2$ or $\mathcal{R} \vdash T_2 \sqsubseteq T_1$. An RBox $\mathcal{R}$ is *admissible* if $\mathcal{R} = \bigcup \mathcal{R}_i$ where (1) each $\mathcal{R}_i$ is strongly admissible and (2) $\mathsf{RN}(\mathcal{R}_i) \cap \mathsf{RN}(\mathcal{R}_j) = \varnothing$ for all $i \neq j$.

In the remainder of this section, we prove the following Theorem:

Theorem 4. *Every admissible RBox is $\mathcal{ALCQ}$-safe.*

Note 1. For $\mathcal{R} = \{\mathsf{Tr}(r)\}$, this result corresponds to the decidability of the graded variant of the modal logic **K4** (called **GrK4**), which has already been addressed in [4]. However, we found that the proof in that paper is incorrect (see [7] for details). Therefore, here we re-establish decidability of **GrK4** as a special case of Theorem 4.

First of all, we demonstrate that, for the purpose of proving safety, it is sufficient to focus only on strongly admissible RBoxes.

Lemma 3 (Modularity). *Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be RBoxes with $\mathsf{RN}(\mathcal{R}_1) \cap \mathsf{RN}(\mathcal{R}_2) = \varnothing$ and $\mathcal{L}$ is between $\mathcal{ALC}$ and $\mathcal{ALCIQ}$. Then $\mathcal{R}_1 \cup \mathcal{R}_2$ is $\mathcal{L}$-safe iff $\mathcal{R}_1$ and $\mathcal{R}_2$ are $\mathcal{L}$-safe.*

Proof. The '$\Rightarrow$' part of the lemma is obvious. The '$\Leftarrow$' part of the lemma follows from the results about fusions of DLs from [2]. See [7] for details. $\dashv$

Corollary 1. *Let $\mathcal{L}$ be a logic between $\mathcal{ALC}$ and $\mathcal{ALCIQ}$. Then every admissible RBox is $\mathcal{L}$-safe provided every strongly admissible RBox is $\mathcal{L}$-safe.*

In order to prove that every strongly admissible RBox $\mathcal{R}$ is safe, it is sufficient to show that the problem of satisfiability of a pair $\langle \mathcal{R}, \mathcal{T} \rangle$, with $\mathcal{T}$ an $\mathcal{L}$-TBox, is decidable. Indeed, a concept C is satisfiable w.r.t. $\mathcal{T}$ and $\mathcal{R}$ iff the pair $\langle \mathcal{R}, \mathcal{T} \cup \{\top \sqsubseteq \exists R.C\} \rangle$ is satisfiable, where R is a fresh role. To this end, we first simplify the TBox $\mathcal{T}$ using Proposition 1 and then internalize RBox $\mathcal{R}$ using Definition 9, which will result in some $\mathcal{L}(\sqcup, +)$-TBox of a restricted form, which we call admissible. We then demonstrate that satisfiability of admissible $\mathcal{L}(\sqcup, +)$-TBoxes is decidable.

In what follows, for convenience, we often identify an extended role $u = \bigsqcup R_i$ with the set $\bigcup\{R_i\}$. Using this convention, we can write $r \in u$ or $u \subseteq u'$ for disjunction of roles u and u', as well as $u^{\mathcal{I}}$ for sets of roles u.

Definition 11. A simple $\mathcal{L}(\sqcup, +)$ TBox $\mathcal{T}$ is *admissible* if (i) all axioms of form (4) are of the forms (5) and (6) below, and (ii) for every two axioms $A_1 \sqsubseteq \leqslant m_1 (u_1^+ \sqcup v_1).B_1$ and $A_2 \sqsubseteq \leqslant m_2 (u_2^+ \sqcup v_2).B_2$ of form (6), we have that either $u_1 \subseteq u_2$, or $u_2 \subseteq u_1$.

$$A \sqsubseteq \leqslant m\,(v).B \tag{5}$$
$$A \sqsubseteq \leqslant m\,(u^+ \sqcup v).B \tag{6}$$

In other words, a simple $\mathcal{L}(\sqcup, +)$-TBox is admissible if in every axiom of form (4) there is at most one occurrence of a transitively closed disjunction of roles.

Lemma 4. *Let $\mathcal{T}$ be a simple $\mathcal{L}$-TBox and $\mathcal{R}$ a strongly admissible RBox. Then $\mathcal{R}(\mathcal{T})$ is a simple admissible $\mathcal{L}(\sqcup, +)$-TBox.*

The condition (ii) from Definition 11 can be alternatively formulated as follows:

Proposition 1. *Let $\mathcal{T}$ be a simple admissible $\mathcal{L}(\sqcup, +)$-TBox. Then all roles in $\mathcal{T}$ can be ordered as $r_1, \ldots, r_n$ in such a way that for every axiom $A \sqsubseteq \leqslant m\,(u^+ \sqcup v).B$ of type (6) and every $1 \leqslant i \leqslant j \leqslant n$, we have that $r_j \in u$ implies $r_i \in u$.*
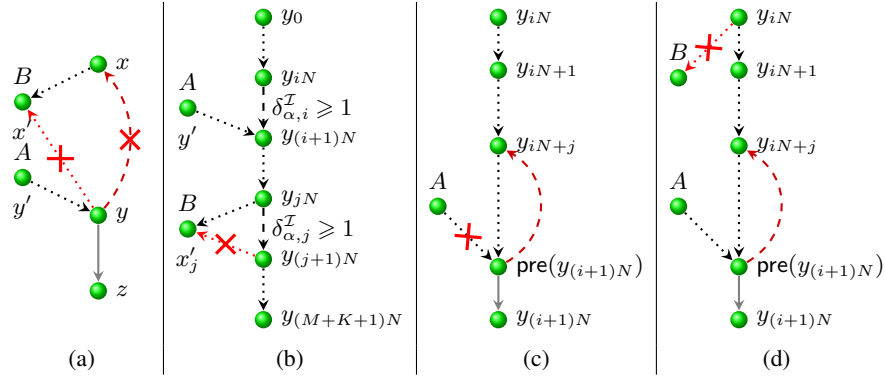
Fig. 5. Looping long chains in a model back

We prove that satisfiability of simple admissible $\mathcal{L}(\sqcup, +)$-TBoxes is decidable by demonstrating the finite model property (FMP) for such TBoxes. The key property that will guarantee FMP is that, in every model of a simple admissible TBox, it is possible to "loop back" every sufficiently long chain of elements connected via roles. This idea is reminiscent to blocking conditions in tableau decision procedures for modal and description logics [6]. The next lemma states that every model of a simple $\mathcal{L}(\sqcup, +)$ TBox can be reduced to a model with bounded branching degree by removing edges that are not "required" by axioms of type (2).

Definition 12. *Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be an $\mathcal{L}$-interpretation. A branching degree of an element $x \in \Delta^{\mathcal{I}}$ in $\mathcal{I}$ is $\deg(\mathcal{I}, x) = \mathsf{Card}\{y \mid \langle x, y \rangle \in r^{\mathcal{I}} \text{ for some } r\}$. A branching degree of $\mathcal{I}$ is $\deg(\mathcal{I}) = \max\{\deg(\mathcal{I}, x) \mid x \in \Delta^{\mathcal{I}}\}$.*

Lemma 5. *Any satisfiable simple $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{T}$ has a model $\mathcal{I}$ with $\deg(\mathcal{I}) \leqslant N(\mathcal{T})$.*

Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be an interpretation. For each axiom α of type (6) in $\mathcal{T}$, we introduce a function $\delta_\alpha^{\mathcal{I}}(x, y)$ defined on elements of $\Delta^{\mathcal{I}}$ as follows:

$$\delta_\alpha^{\mathcal{I}}(x, y) = \begin{cases} \mathsf{Card}\{x' \mid x' \in B^{\mathcal{I}}, \ \langle x, x' \rangle \in (u^+)^{\mathcal{I}}, \ \langle y, x' \rangle \notin (u^+)^{\mathcal{I}}\} \\ \qquad \text{if there exists } y' \in A^{\mathcal{I}} \text{ with } \langle y', y \rangle \in (u^+)^{\mathcal{I}} \\ 0 \qquad \text{otherwise} \end{cases}$$

In other words, if y has a u^+ predecessor in which A holds, $\delta_\alpha^{\mathcal{I}}(x, y)$ equals to the number of elements in which B holds and that are reachable via u^+ from x but not from y (see Fig. 5a). The value of $\delta_\alpha^{\mathcal{I}}(x, y)$ intuitively indicates the number of new u^+ successors of y that might appear and potentially violate the axiom α (at the points, where A holds), if x becomes reachable from y via u^+.

Definition 13. *Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be an interpretation. For an element $x \in \Delta^{\mathcal{I}}$, let $\mathsf{CN}^{\mathcal{I}}(x) := \{A \in \mathsf{CN} \mid x \in A^{\mathcal{I}}\}$ denote the set of concept names that hold at x in $\mathcal{I}$.*

Given a simple admissible $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{T}$, an interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, and $x, y, z \in \Delta^{\mathcal{I}}$, we say that x can foster z for y in $\mathcal{I}$ (w.r.t. $\mathcal{T}$) if (i) $\mathsf{CN}^{\mathcal{I}}(z) = \mathsf{CN}^{\mathcal{I}}(x)$,

(ii) $\langle y, x \rangle \in r^{\mathcal{I}}$ for no atomic role r, and (iii) for every axiom α of type (6) in $\mathcal{T}$, if $\langle y, z \rangle \in r^{\mathcal{I}}$ for some role $r \in u_\alpha = u$, then $\delta_\alpha^{\mathcal{I}}(x, y) = 0$.

Lemma 6 (Model Transformation). *Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be a model of a simple admissible $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{T}$ and x, y, z elements of $\Delta^{\mathcal{I}}$ such that x can foster z for y in $\mathcal{I}$ w.r.t. $\mathcal{T}$. Let $\mathcal{J} = \mathsf{swap}(\mathcal{I}, x, y, z)$ be obtained from $\mathcal{I}$ by setting $A^{\mathcal{J}} := A^{\mathcal{I}}$ and, $r^{\mathcal{J}} := r^{\mathcal{I}} \setminus \{\langle y, z \rangle\} \cup \{\langle y, x \rangle\}$ if $\langle y, z \rangle \in r^{\mathcal{I}}$, and $r^{\mathcal{J}} := r^{\mathcal{I}}$ otherwise, for every concept name A and role name r. Then $\mathcal{J}$ is a model of $\mathcal{T}$.*

Our main lemma states that, in every model of simple admissible $\mathcal{L}(\sqcup, +)$-TBox $\mathcal{T}$, every sufficiently long chain $x_0, \ldots, x_p$ of elements connected with roles contains two elements x_i and x_j with $i < j$ such that x_i can foster x_j for the predecessor x_{j-1} of x_j w.r.t. $\mathcal{T}$. Thus, every sufficiently long chain can be "looped back" using the transformation described in Lemma 6.

Lemma 7 (Main Lemma). *Let $\mathcal{T}$ be a simple admissible $\mathcal{L}(\sqcup, +)$-TBox and $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ a model for $\mathcal{T}$ with $\deg(\mathcal{I}) \leqslant N$. Let $r_1, \ldots, r_n$ be all the role names in $\mathcal{T}$ enumerated according to Proposition 1, k an integer with $1 \leqslant k \leqslant n$, and $x_0, \ldots, x_p$ a sequence of distinct elements in $\Delta^{\mathcal{I}}$ such that, for every $i \geqslant 1$, there exists $\ell \leqslant k$ such that $\langle x_{i-1}, x_i \rangle \in r_\ell^{\mathcal{I}}$. Then there exist i and j with $1 \leqslant i < j \leqslant p$ such that x_i can foster x_j for x_{j-1}, provided that $p \geqslant p_k := ((M + K + 1)N \cdot 2^C + 1)^k$, where $M = M(\mathcal{T})$, $K = K(\mathcal{T})$, and $C = C(\mathcal{T})$ as defined in Definition 7.*

Before proving Lemma 7, we demonstrate the following auxiliary property. For convenience, if x is an element of the sequence $x_0, \ldots, x_p$, i.e., $x = x_i$ for some i, then its predecessor *in this sequence* will be denoted by $\mathsf{pre}(x) := x_{i-1}$.

Lemma 8 (Auxiliary Lemma). *Let a TBox $\mathcal{T}$, a model $\mathcal{I}$, and a sequence $x_0, \ldots, x_p$ be as in Lemma 7. Let $y_0, \ldots, y_q$ be a sub-sequence in $x_1, \ldots, x_p$ such that (i) $q \geqslant (M + K + 1)N$, (ii) $\mathsf{CN}^{\mathcal{I}}(y_0) = \cdots = \mathsf{CN}^{\mathcal{I}}(y_q)$, (iii) $\langle \mathsf{pre}(y_i), y_i \rangle \notin r_\ell^{\mathcal{I}}$ for all $0 \leqslant i \leqslant q$ and $\ell < k$. Then for some $0 \leqslant i < j \leqslant q$, y_i can foster y_j for $\mathsf{pre}(y_j)$.*

Proof. Let U_k be the set of axioms $\alpha = (A \sqsubseteq \leqslant m\,(u^+ \sqcup v).B) \in \mathcal{T}$ of type (6) such that $r_k \in u$. Take any axiom $\alpha \in U_k$ and consider a sequence of values $\delta_{\alpha,i}^{\mathcal{I}} := \delta_\alpha^{\mathcal{I}}(y_{iN}, y_{(i+1)N})$ for $0 \leqslant i \leqslant M + K$ (see Fig. 5b). We claim that at most $m + 1$ of values $\delta_{\alpha,i}^{\mathcal{I}}$ are positive.

Indeed, for the first i with $\delta_{\alpha,i}^{\mathcal{I}} \geqslant 1$, by definition of $\delta_\alpha^{\mathcal{I}}(x, y)$, there exists $y' \in A^{\mathcal{I}}$ with $\langle y', y_{(i+1)N} \rangle \in (u^+)^{\mathcal{I}}$. For all subsequent $j > i$ with $d_\alpha^j \geqslant 1$, there exists an element x_j' such that $\langle y_{jN}, x_j' \rangle \in (u^+)^{\mathcal{I}}$, but $\langle y_{(j+1)N}, x_j' \rangle \notin (u^+)^{\mathcal{I}}$. In particular, all such x_j are distinct for different j. Note that since $r_k \in u$, by Proposition 1, $r_\ell \in u$ for all $\ell \leqslant k$. Hence $\langle y', x_j' \rangle \in (u^+)^{\mathcal{I}}$. Since $\mathcal{I}$ is a model of α, the number of such different j can be at most m.

Hence, the number of different i such that, for some $\alpha \in U_k$, $\delta_{\alpha,i}^{\mathcal{I}} \geqslant 1$, is at most $\sum_{\alpha \in U_k} (m_\alpha + 1) \leqslant M + K$. Since $q \geqslant (M + K + 1)N$, there exists at least one i such that $\delta_{\alpha,i}^{\mathcal{I}} = 0$ for all $\alpha \in U_k$. For every $\alpha \in U_k$, there are two cases possible: either (1) there exists no $y' \in A^{\mathcal{I}}$ such that $\langle y', y_{(i+1)N} \rangle \in (u^+)^{\mathcal{I}}$ (see Fig. 5c), or (2) such a

y' exists, but there exists no $x' \in B^{\mathcal{I}}$ with $\langle y_{iN}, x' \rangle \in (u^+)^{\mathcal{I}}$ (see Fig. 5d). Hence, in particular, $\delta_\alpha^{\mathcal{I}}(y_{iN+j}, \mathsf{pre}(y_{(i+1)N})) = 0$ for all $j < N$ and all $\alpha \in U_k$.

Since $\deg(\mathcal{I}) \leqslant N$ and $\langle \mathsf{pre}(y_{(i+1)N}), y_{(i+1)N} \rangle \in r_\ell^{\mathcal{I}}$ for $\ell \leqslant k$, there exists $j < N$ such that $\langle \mathsf{pre}(y_{(i+1)N}), y_{iN+j} \rangle \notin r^{\mathcal{I}}$, for every r. Since, by condition (iii), we have $\langle \mathsf{pre}(y_{(i+1)N}), y_{(i+1)N} \rangle \notin r_l^{\mathcal{I}}$ for each $\ell < k$, we have $\delta_\alpha^{\mathcal{I}}(y_{iN+j}, \mathsf{pre}(y_{(i+1)N})) = 0$ for each axiom α of type (6) such that $\langle \mathsf{pre}(y_{(i+1)N}), y_{(i+1)N} \rangle \in r_\ell^{\mathcal{I}}$ and $r_\ell \in u$. Indeed, those are exactly $\alpha \in U_k$, because $r_\ell \in u$ implies $r_k \in u$ for every $\ell \geqslant k$ by Proposition 1. Hence, by Definition 13, y_{iN+j} can foster $y_{(i+1)N}$ for $\mathsf{pre}(y_{(i+1)N})$. $\dashv$

Proof (of Lemma 7). We prove the lemma by induction on k, using Lemma 8 both in induction base and induction step. Denote $L := (M + K + 1)N$ for short.

Induction base: For $k = 1$, we have a sequence of elements $x_0, \ldots, x_p \in \Delta^{\mathcal{I}}$ with $p \geqslant p_1 := L \cdot 2^C + 1$ such that $\langle x_{i-1}, x_i \rangle \in r_1^{\mathcal{I}}$, for all $1 \leqslant i \leqslant p$. We claim that there exists a subsequence $y_0, \ldots, y_q$ in $x_1, \ldots, x_p$ with $q \geqslant L$ such that $\mathsf{CN}^{\mathcal{I}}(y_0) = \cdots = \mathsf{CN}^{\mathcal{I}}(y_q)$. Indeed, otherwise, since the number of different values of $\mathsf{CN}^{\mathcal{I}}(x)$ is bounded by 2^C, and the number of elements x in $x_1, \ldots, x_p$ with the same value of $\mathsf{CN}^{\mathcal{I}}(x)$ is at most L, the total number of elements in $x_1, \ldots, x_p$ cannot exceed $L \cdot 2^C$, which contradicts to $p \geqslant p_1$. Now, Lemma 8 can be applied to the sequence $y_0, \ldots, y_q$, since there are no roles r_ℓ with $\ell < k = 1$. By Lemma 8 there exist elements y_i and y_j in this sequence with $0 \leqslant i < j \leqslant q$, such that y_i can foster y_j for $\mathsf{pre}(y_j)$.

Induction Step: Assume that the lemma holds for $k - 1$. Two cases are possible:

(A) There exists a sub-sequence of consecutive elements $x_i, x_{i+1}, \ldots, x_{i+p_{k-1}}$ with $p_{k-1} = (L \cdot 2^C + 1)^{k-1}$ and for each j with $1 \leqslant j \leqslant p_{k-1}$, there exists $\ell \leqslant k - 1$ such that $\langle x_{i+j-1}, x_{i+j} \rangle \in r_\ell^{\mathcal{I}}$. In this case the lemma holds by the induction hypothesis.

(B) Otherwise, in every sequence $x_{ip_{k-1}}, x_{ip_{k-1}+1} \ldots, x_{(i+1)p_{k-1}}$ of consecutive elements with $0 \leqslant i \leqslant p_1 - 1 = L \cdot 2^C$, there exists an element $x_i' = x_{ip_{k-1}+j}$, with $1 \leqslant j \leqslant p_{k-1}$, such that $\langle \mathsf{pre}(x_i'), x_i' \rangle \notin r_\ell^{\mathcal{I}}$ for all $\ell \leqslant k - 1$. By applying a combinatorial argument as in the induction base, from the sequence $x_0', \ldots, x_{p_1-1}'$ of $p_1 = L \cdot 2^C$ distinct elements one can select a subsequence $y_0, \ldots, y_q$ with $q \geqslant L$ such that $\mathsf{CN}^{\mathcal{I}}(y_0) = \cdots = \mathsf{CN}^{\mathcal{I}}(y_q)$. Hence the claim of the lemma follows from Lemma 8 applied to the sequence $y_0, \ldots, y_q$. $\dashv$

Theorem 5. *An admissible $\mathcal{ALCQ}(\sqcup, +)$-TBox $\mathcal{T}$ is satisfiable iff $\mathcal{T}$ has a finite model.*

Proof. The "if" direction is trivial. To prove the "only if" part, we use the following argument: given any model $\mathcal{I}$ of $\mathcal{T}$, we build a finite model $\mathcal{J}$ of $\mathcal{T}$ by "looping back" all sufficiently long paths from some element x_0 using Lemma 6 and Lemma 7 until none of them left, and then removing elements that became disconnected from x_0 after this transformation. For a detailed proof, see the technical report [7]. $\dashv$

Now it is time to harvest our decidability results (see [7] for all proofs).

Theorem 6. *Let $\mathcal{T}$ be an $\mathcal{ALCQ}$-TBox and $\mathcal{R}$ a strongly admissible RBox. Then $\langle \mathcal{R}, \mathcal{T} \rangle$ is satisfiable iff it has a finite model.*

Corollary 2 (Theorem 4). *Every admissible RBox is safe for $\mathcal{ALCQ}$.*

Corollary 3. *Every satisfiable $\mathbf{GrK4}$-formula has a finite model.*
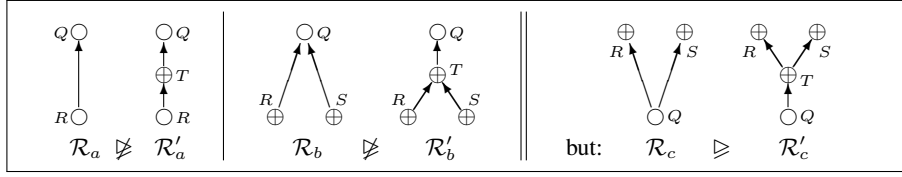
Fig. 6. For $i \in \{a, b\}$, we have $\mathcal{R}_i \sqsubseteq \mathcal{R}'_i$ and $\mathcal{R}_i \not\triangleright \mathcal{R}'_i$; whereas $\mathcal{R}_c \triangleright \mathcal{R}'_c$.

6 Extending RBoxes

In Section 5 we have described a rather large class of $\mathcal{ALCQ}$-safe RBoxes. However, so far, only few RBoxes were shown to be unsafe for $\mathcal{ALCN}$ and $\mathcal{ALCIN}$ in Section 3. In this section we are concerned with a question whether every RBox "containing" any of the patterns described in Section 3 is necessarily unsafe? Or, in general, what happens to the (un)safety of an RBox when the RBox are extended?

It is clear that adding axioms may turn a safe RBox into unsafe and vice versa: an $\mathcal{ALCN}$-safe RBox $\{\mathsf{Tr}(r)\}$ can be extended to an $\mathcal{ALCN}$-unsafe RBox $\mathcal{R}_\wedge$ from Theorem 1; adding to $\mathcal{R}_\wedge$ an inclusion between its incomparable transitive roles yields an $\mathcal{ALCN}$-safe RBox by Theorem 4. So it is not sufficient for an RBox $\mathcal{R}'$ to be unsafe if it contains some unsafe RBox $\mathcal{R}$. The question now is: what additional property an extension $\mathcal{R}'$ of $\mathcal{R}$ should fulfill so that unsafety of $\mathcal{R}$ can be transferred to $\mathcal{R}'$. In this section we demonstrate that it is sufficient to require that $\mathcal{R}'$ is *semantically conservative* over $\mathcal{R}$.

Definition 14. Let $\mathcal{R}$ and $\mathcal{R}'$ be two RBoxes. We say that $\mathcal{R}'$ *is semantically conservative over* $\mathcal{R}$ (notation: $\mathcal{R} \triangleright \mathcal{R}'$), if every model $\mathcal{I}$ of $\mathcal{R}$ can be expanded to a model $\mathcal{I}'$ of $\mathcal{R}'$ by interpreting new role names from $\mathsf{RN}(\mathcal{R}') \setminus \mathsf{RN}(\mathcal{R})$. If, additionally, we have $[\mathcal{R}] \subseteq [\mathcal{R}']$, then $\mathcal{R}'$ is called a *semantic conservative extension of $\mathcal{R}$.*

Example 1. Consider RBoxes depicted in Fig. 6. We have $\mathcal{R}_c \triangleright \mathcal{R}'_c$. Indeed, given any model $\mathcal{I} \models \mathcal{R}_c$, we define $\mathcal{I}'$ by setting $R^{\mathcal{I}'} := R^{\mathcal{I}}$, $S^{\mathcal{I}'} := S^{\mathcal{I}}$, and $T^{\mathcal{I}'} := R^{\mathcal{I}} \cap S^{\mathcal{I}}$. Then $\mathcal{I}' \models \mathcal{R}'_c$; in particular, $T^{\mathcal{I}'}$ is transitive as the intersection of transitive relations.

At the same time, $\mathcal{R}_a \not\triangleright \mathcal{R}'_a$, since one can easily construct two non-transitive relations on some set that have no transitive relations between them. Furthermore, $\mathcal{R}_b \not\triangleright \mathcal{R}'_b$; indeed, take $\Delta^{\mathcal{I}} = \{0, 1, 2\}$ and set $R^{\mathcal{I}} := \{\langle 0, 1\rangle\}$, $S^{\mathcal{I}} := \{\langle 1, 2\rangle\}$, and $Q^{\mathcal{I}} := R^{\mathcal{I}} \cup S^{\mathcal{I}}$. Then we cannot interpret the transitive role T to satisfy $\mathcal{R}'_b$.

Theorem 7 (Preservation of Unsafety under Conservative Extensions of RBoxes).
If $\mathcal{R}'$ is a conservative extension of $\mathcal{R}$ and $\mathcal{R}$ is $\mathcal{L}$-unsafe, then $\mathcal{R}'$ is $\mathcal{L}$-unsafe.

Let $\mathcal{F} := \{\mathcal{R}_\vee^\oplus\} \cup \{\mathcal{R}_\wedge^n, \mathcal{R}_\vee^n, \mathcal{R}_\mathbb{N}^n \mid n \geqslant 1\}$ be the family of RBoxes (see. Fig. 3) that we have shown to be $\mathcal{ALCQ}$-unsafe in Sect. 3. As a consequence of Theorems 2, 3, and 7, we obtain the following result:

Corollary 4. *(1) Any $\mathcal{R}'$ that is a conservative extension of $\{\mathsf{Tr}(r)\}$ is $\mathcal{ALCIQ}$-unsafe; (2) Any $\mathcal{R}'$ that is a conservative extension of some RBox $\mathcal{R} \in \mathcal{F}$ is $\mathcal{ALCQ}$-unsafe.*

It turns out, surprisingly, that properties (1) and (2) in Corollary 4 can be checked in polynomial in the size of $\mathcal{R}'$ (see [7] for details). We conjecture that Corollary 4 describes *all* RBoxes (modulo role renaming) that are unsafe for $\mathcal{ALCIQ}$ and $\mathcal{ALCQ}$.

7 Conclusions and Future Work

Driven by applications, we have looked more closely at the effect of non-simple roles in number restrictions on the decidability of standard DL reasoning problems. We have shown that, in the absence of inverse roles, the restriction imposed by $\mathcal{SHQ}$ to non-simple roles in number restrictions can be relaxed substantially and that, in the presence of inverse roles, this restriction turns out to be crucial for decidability.

These results raise numerous further questions. Firstly, given a DL $\mathcal{L}$, can we formulate necessary and sufficient conditions for an RBox to be $\mathcal{L}$-safe? Secondly, for an interesting class of $\mathcal{L}$-safe RBoxes $\mathcal{R}$, what is the computational complexity of deciding $\mathcal{L}(\mathcal{R})$-satisfiability? And can these decision procedures be implemented and used in practice? Thirdly, in the approach taken here, we allow all roles to occur in number restrictions. Given an $\mathcal{L}$-unsafe RBox $\mathcal{R}$, can we extend the notion of simple roles to regain decidability of $\mathcal{L}(\mathcal{R})$? And how applicable would this be in practice? Finally, in the presence of inverse roles, can we restrict the usage of inverse roles in TBoxes so as to re-gain decidability? For example, would disallowing number restrictions on inverse roles whilst allowing number restrictions on transitive role names help? For the list of other interesting open problems, see the accompanying technical report [7].

References

1. F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, 2003.
2. F. Baader, C. Lutz, H. Sturm, and F. Wolter. Fusions of Description Logics and Abstract Description Systems, Journal of Artificial Intelligence Research (JAIR), 16, 2002, pp. 1–58.
3. E. Börger, E. Grädel, Y. Gurevich. *The Classical Decision Problem*. Perspectives in Mathematical Logic. Springer-Verlag, 1997.
4. C. Cerrato. Decidability by Filtration for Graded Modal Logics (Graded Modalities V). *Studia Logica*, 53:61–74, 1994.
5. I. Horrocks, P. F. Patel-Schneider, and F. van Harmelen. From $\mathcal{SHIQ}$ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.
6. I. Horrocks, U. Sattler, and S. Tobies. Practical reasoning for very expressive description logics. *Logic Journal of the IGPL*, 8(3):239–263, 2000.
7. Y. Kazakov, U. Sattler, and E. Zolin. Is Your RBox Safe? Technical report. The University of Manchester, 2007. Available at `http://www.cs.man.ac.uk/˜ezolin/pub/`
8. K. Wolstencroft, A. Brass, I. Horrocks, P. Lord, U. Sattler, D. Turi, R. Stevens. A Little Semantic Web goes a long way in biology. In *Proc. of the 4th ISWC*, pp. 786–800, 2005.