Presburger Concept Cardinality Constraints in Very Expressive Description Logics

Allegro sexagenarioso ma non ritardando

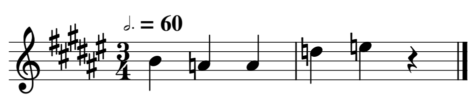
"Johann" Sebastian Rudolph[0000−0002−1609−2080]

Computational Logic Group, TU Dresden, Germany
sebastian.rudolph@tu-dresden.de

Abstract. Quantitative information plays an increasingly important role in knowledge representation. To this end, many formalisms have been proposed that enrich traditional KR formalisms by counting and some sort of arithmetics. Baader and Ecke (2017) propose an extension of the description logic $\mathcal{ALC}$ by axioms which express correspondences between the cardinalities of concepts by means of Presburger arithmetics. This paper extends their results, enhancing the expressivity of the underlying logic as well as the constraints while preserving complexities. It also widens the scope of investigation from finite models to the classical semantics where infinite models are allowed. We also provide first results on query entailment in such logics. As opposed to prior work, our results are established by polynomially encoding the cardinality constraints in the underlying logic.

Keywords: F♯ Major · B Major.

Prelude: Dedication



This work is inspired by Franz. It was him who introduced me to the idea of extending the quantitative capabilities in description logics using Presburger arithmetic constraints. Quite fittingly, this happened during a colloquial discussion at a workshop of our DFG PhD training group QuantLA – *Quantitative Logics and Automata*. Already back then, I was musing that it should be possible to encode certain cardinality constraints in expressive description logics, using some auxiliary vocabulary.
Thankfully, this festschrift provided the most suitable occasion, motivation, and pressure[1] to explore that option thoroughly, arguably with quite decent results. I'd like to dedicate this "composition" to Franz and thank him for bringing up this interesting topic and, generally, for plenty of inspiring discussions, occupational support, and pleasant collaboration.

[1] I hereby sincerely apologize to Anni, Carsten, Cesare, Frank, and Uli for pushing the submission deadline.

1 Introduction: Toward Quantitative Description Logics

Enriching knowledge representation formalisms with features for counting and basic arithmetic regarding domain individuals is a worthwhile endeavor. So far, mainstream description logics provide only very limited support in this respect: *(qualified) number restrictions* allow for enforcing concrete upper and lower bounds on the number of an individual's role neighbors. These limited capabilities fall short of some basic practical knowledge representation requirements, such as expressing statistical information [21]. As an example, assume that on the occasion of a distinguished scientist's 60th birthday, fellow researchers group together to produce a festschrift consisting of distinct contributed papers. The publisher requires the festschrift to have not less than 140 and not more than 800 pages. This could intuitively be expressed by a statement like

$$140 \leq |\texttt{Page}| \leq 800,$$

assuming Page denotes the class of all the festschrift's pages. Let's say, the editors manage to recruit a total of 73 authors:

$$|\texttt{Author}| = 73.$$

They assume that the average number of contributors per paper is between 2 and 3:

$$2 \cdot |\texttt{Paper}| \leq |\texttt{Author}| \leq 3 \cdot |\texttt{Paper}|.$$

They impose the condition that each paper must have at least 10 and at most 40 pages:

$$\texttt{Paper} \sqsubseteq \geqslant 10\ \texttt{OnPage}.\top \sqcap \leqslant 40\ \texttt{OnPage}.\top.$$

They also notice that just one author (the "outlier") contributes to two papers, whereas the others contribute to one.

$$\texttt{Author} \sqcap \neg\{\texttt{outlier}\} \sqsubseteq {=}1\ \texttt{Contributes}.\top \qquad \{\texttt{outlier}\} \sqsubseteq {=}2\ \texttt{Contributes}.\top$$

Some background knowledge for our domain needs to be specified: roughly speaking, authors are precisely contributors, papers are precisely "contributees" and they are precisely the things occurring on pages and on nothing but pages.

$$\texttt{Author} \equiv \exists\texttt{Contributes}.\top \quad \texttt{Paper} \equiv \exists\texttt{Contributes}^{-}.\top$$
$$\texttt{Paper} \equiv \exists\texttt{OnPage}.\top \quad \texttt{Page} \equiv \exists\texttt{OnPage}^{-}.\top$$

As an important last ingredient, it needs to be specified that no two distinct papers can occur on the same page (i.e., OnPage is inverse functional):

$$\top \sqsubseteq \leqslant 1\ \texttt{OnPage}^{-}.\top.$$

If a reasoner supporting all the modeling features in this specification existed, the editors could now find out by a satisfiability check if the planned festschrift can be published under the given assumptions (which is the case). Also, by removing the first statement and checking for its entailment instead, they could find out if the publisher's space constraints are guaranteed to be met in view of the given information (which is not the case). Alas, currently, axioms alike the first three statements are not supported by mainstream description logics.

In a line of recent work, Franz and others have addressed the shortcomings in description logics on the quantitative side. For instance, extending results from [12], Franz proposed $\mathcal{ALCSCC}$ [1], an extension of the basic description logic $\mathcal{ALC}$ by constraints expressed in the quantifier-free fragment of Boolean Algebra with Presburger Arithmetic (QFBAPA) [19] over role successors. The described constraints are *local*, as they always refer to an individual under consideration, as opposed to *global* constraints, which range over the full domain and compare cardinalities of concepts. The latter were introduced in [5], giving rise to the notions of $\mathcal{ALC}$ extended cardinality boxes (ECboxes) and – striving for more favorable complexity results – their "light version" $\mathcal{ALC}$ restricted cardinality boxes (RCboxes). As a natural next step, [2] introduced and investigated $\mathcal{ALCSCC}$ ECboxes and RCboxes, enabling both local and global cardinality constraints in a joint formalism. Pushing the envelope further, [3] showed that local and global constraints can be tightly integrated leading to the starkly more expressive logic $\mathcal{ALCSCC}^{++}$, for which ECbox consistency checking is still NExpTime-complete. On the downside, conjunctive query entailment becomes undecidable in this logic. Moreover, as an (albeit massively calculation-enhanced) version of plain $\mathcal{ALC}$, $\mathcal{ALCSCC}^{++}$ is lacking basic modeling features that are normally taken for granted in description logics. Most notably, it does not feature role inverses, which are crucial to draw level with popular logics from other families, such as two-variable logics.

Decidability and complexity results for the logics discussed above were established via the solution of large systems of (in)equalities as well as elaborate constructions and transformations of models. We show that, if we limit our attention to global cardinality constraints, we can use an alternative, reduction-based approach, and expand the existing results simultaneously in three directions:

- *Incorporation of role inverses.* As stated earlier, the existing results are for description logics without the feature of role inverses. In fact, it is notoriously difficult to incorporate this feature into the (in)equality-system-based machinery hitherto used. Interestingly, the method proposed in this paper not only allows for incorporating role inverses, it actually does require their presence in the logic.
- *Relaxation of restrictions on RCboxes.* As mentioned above, RCboxes were introduced as a light version of ECboxes in order to obtain more favorable complexity results. We show that some of the restrictions made can be relaxed without endangering this complexity gain. This actually motivates us to introduce *extended restricted cardinality boxes* (ERCboxes) as low-complexity, high-expressivity middle-ground between ECboxes and RCboxes.
- *Reasoning over finite and arbitrary models.* Previous results confine themselves to a finite-model setting, which is arguably the right choice for practical modeling tasks in concrete scenarios where arithmetic is applied. However, the traditional semantics of description logics allows for infinite models. Hence we extend the scope of our investigations to also include the case of arbitrary models. Next to an appropriate extension of the underlying arithmetic (as described in the beginning of Section 2) this raises some deeper model-theoretic concerns, which we address in Section 5.

2 Ostinato: Preliminaries

Numbers. We recall that $\mathbb{N}$ denotes the set of natural numbers (including 0). Throughout this paper, whenever natural numbers occur in some expression, we assume binary encoding. We let $\mathbb{N}^\infty = \mathbb{N} \cup \{\infty\}$. Basic arithmetic and comparison operations are extended from $\mathbb{N}$ to $\mathbb{N}^\infty$ in the straightforward way, in particular adding anything to ∞ yields ∞, $0 \cdot \infty = 0$, and $n \cdot \infty = \infty$ for every $n \geq 1$. For $n \in \mathbb{N}^\infty$, we let $[n] = \{i \mid i < n\}$, in particular $[\infty] = \mathbb{N}$. For some set S, we let $|S|$ denote the number of elements of S if it is finite and ∞ otherwise.

Description Logics. We give the definition of the extremely expressive description logic $\mathcal{SROIQB}$ which is obtained from the well-known description logic $\mathcal{SROIQ}$ [16] by allowing arbitrary Boolean constructors on simple roles. We assume that the reader is familiar with description logics [4,6,26].
The description logics considered in this paper are based on four disjoint sets of *individual names* $\mathbf{N_I}$, *concept names* $\mathbf{N_C}$, *simple role names* $\mathbf{N_R^s}$, and *non-simple role names* $\mathbf{N_R^n}$ (containing the *universal role* $\overline{\top} \in \mathbf{N_R}$). Furthermore, we let $\mathbf{N_R} := \mathbf{N_R^s} \cup \mathbf{N_R^n}$.

Definition 1 (syntax of $\mathcal{SROIQB}$). *A $\mathcal{SROIQB}$ Rbox for $\mathbf{N_R}$ is based on a set $\mathbf{R}$ of* atomic roles *defined as $\mathbf{R} := \mathbf{N_R} \cup \{R^- \mid R \in \mathbf{N_R}\}$, where we set* $\mathrm{Inv}(R) := R^-$ *and* $\mathrm{Inv}(R^-) := R$ *to simplify notation. In turn, we distinguish* simple atomic roles $\mathbf{R}^\mathrm{s} := \mathbf{N_R^s} \cup \mathrm{Inv}(\mathbf{N_R^s})$ *and* non-simple roles $\mathbf{R}^\mathrm{n} := \mathbf{N_R^n} \cup \mathrm{Inv}(\mathbf{N_R^n})$.
The set of simple roles $\mathbf{B}$ *is defined as follows:*

$$\mathbf{B} ::= \mathbf{N_R^s} \mid \neg\mathbf{B} \mid \mathbf{B} \cap \mathbf{B} \mid \mathbf{B} \cup \mathbf{B} \mid \mathbf{B} \setminus \mathbf{B}.$$

Moreover, a simple role will be called safe*, if it does not contain $\neg$.*
A generalized role inclusion axiom *(RIA) is a statement of the form $S \sqsubseteq R$ with simple roles S and R, or of the form*

$$S_1 \circ \ldots \circ S_n \sqsubseteq R$$

where each S_i is a (simple or non-simple) role, and where R is a non-simple atomic role, none of them being $\overline{\top}$. A set of such RIAs will be called a generalized role hierarchy. *A role hierarchy is* regular *if there is a strict partial order $\prec$ on the non-simple roles $\mathbf{R}^\mathrm{n}$ such that*

- $S \prec R$ *iff* $\mathrm{Inv}(S) \prec R$*, and*
- *every RIA is of one of the forms*

$$R\circ R \sqsubseteq R \quad R^- \sqsubseteq R \quad S_1\circ\ldots\circ S_n \sqsubseteq R \quad R\circ S_1\circ\ldots\circ S_n \sqsubseteq R \quad S_1\circ\ldots\circ S_n\circ R \sqsubseteq R$$

  such that $R \in \mathbf{N_R}$ is a (non-inverse) role name, and $S_i \prec R$ for $i = 1, \ldots, n$ whenever S_i is non-simple.

A $\mathcal{SROIQB}$ Rbox *is a regular role hierarchy.*[2]

[2] The original definition of $\mathcal{SROIQ}$ Rboxes also features explicit axioms expressing role reflexivity, asymmetry, and role disjointness. However, in the presence of (safe) Boolean role constructors, these can be expressed, so we omit them here.

Table 1: Semantics of $\mathcal{SROIQB}$ role and concept constructors for interpretation $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$.

Name	Syntax	Semantics
inverse role	R^-	$\{(x,y) \in \Delta^\mathcal{I} \times \Delta^\mathcal{I} \mid (y,x) \in R^\mathcal{I}\}$
universal role	$\overline{\top}$	$\Delta^\mathcal{I} \times \Delta^\mathcal{I}$
role negation	$\neg S$	$\{(x,y) \in \Delta^\mathcal{I} \times \Delta^\mathcal{I} \mid (x,y) \notin R^\mathcal{I}\}$
role conjunction	$S \cap R$	$S^\mathcal{I} \cap R^\mathcal{I}$
role disjunction	$S \cup R$	$S^\mathcal{I} \cup R^\mathcal{I}$
role difference	$S \setminus R$	$S^\mathcal{I} \setminus R^\mathcal{I}$
top	$\top$	$\Delta^\mathcal{I}$
bottom	$\bot$	$\emptyset$
negation	$\neg C$	$\Delta^\mathcal{I} \setminus C^\mathcal{I}$
conjunction	$C \sqcap D$	$C^\mathcal{I} \cap D^\mathcal{I}$
disjunction	$C \sqcup D$	$C^\mathcal{I} \cup D^\mathcal{I}$
nominals	$\{\mathsf{a}\}$	$\{\mathsf{a}^\mathcal{I}\}$
univ. restriction	$\forall R.C$	$\{x \in \Delta^\mathcal{I} \mid (x,y) \in R^\mathcal{I} \text{ implies } y \in C^\mathcal{I}\}$
exist. restriction	$\exists R.C$	$\{x \in \Delta^\mathcal{I} \mid \text{for some } y \in \Delta^\mathcal{I}, (x,y) \in R^\mathcal{I} \text{ and } y \in C^\mathcal{I}\}$
Self concept	$\exists S.\mathsf{Self}$	$\{x \in \Delta^\mathcal{I} \mid (x,x) \in S^\mathcal{I}\}$
qualified number	$\leqslant n\, S.C$	$\{x \in \Delta^\mathcal{I} \mid \lvert\{y \in \Delta^\mathcal{I} \mid (x,y) \in S^\mathcal{I} \text{ and } y \in C^\mathcal{I}\}\rvert \leq n\}$
restriction	$\geqslant n\, S.C$	$\{x \in \Delta^\mathcal{I} \mid \lvert\{y \in \Delta^\mathcal{I} \mid (x,y) \in S^\mathcal{I} \text{ and } y \in C^\mathcal{I}\}\rvert \geq n\}$

Given a $\mathcal{SROIQB}$ Rbox $\mathcal{R}$, the set of concept expressions *(short: concepts)* $\mathbf{C}$ *is inductively defined as follows:*

- $\mathbf{N_C} \subseteq \mathbf{C}$, $\top \in \mathbf{C}$, $\bot \in \mathbf{C}$,
- *for $C, D \in \mathbf{C}$ concepts, $R \in \mathbf{B} \cup \mathbf{R}^\mathrm{n}$ a (simple or non-simple) role, $S \in \mathbf{B}$ a simple role, $\mathsf{a} \in \mathbf{N_I}$, and $n \in \mathbb{N}$ a non-negative integer, the expressions $\neg C$, $C \sqcap D$, $C \sqcup D$, $\{\mathsf{a}\}$, $\forall R.C$, $\exists R.C$, $\exists S.\mathsf{Self}$, $\leqslant n\, S.C$, and $\geqslant n\, S.C$ are also concepts.*

Throughout this paper, the symbols C, D will be used to denote concepts. A $\mathcal{SROIQB}$ Tbox *is a set of* general concept inclusion axioms *(GCIs) of the form $C \sqsubseteq D$. We use $C \equiv D$ as a shorthand for $C \sqsubseteq D$, $D \sqsubseteq C$.*
An individual assertion *can have any of the following forms: $C(\mathsf{a})$, $R(\mathsf{a},\mathsf{b})$, $\neg S(\mathsf{a},\mathsf{b})$, $\mathsf{a} \approx \mathsf{b}$, $\mathsf{a} \not\approx \mathsf{b}$, with $\mathsf{a},\mathsf{b} \in \mathbf{N_I}$ individual names, $C \in \mathbf{C}$ a concept, and $R, S \in \mathbf{B} \cup \mathbf{R}^\mathrm{n}$ roles with S simple. A $\mathcal{SROIQB}$* Abox *is a set of individual assertions.*
A $\mathcal{SROIQB}$ knowledge base *$\mathcal{K}$ is a triple $(\mathcal{A}, \mathcal{T}, \mathcal{R})$ where $\mathcal{R}$ is a regular Rbox while $\mathcal{A}$ and $\mathcal{T}$ are an Abox and a Tbox for $\mathcal{R}$, respectively. We use the term* axiom *to uniformly refer to any single statement contained in $\mathcal{A}$, $\mathcal{T}$, or $\mathcal{R}$.*

We further provide the semantics of $\mathcal{SROIQB}$ knowledge bases.

Definition 2 (semantics of $\mathcal{SROIQB}$). *An interpretation $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ consists of a set $\Delta^\mathcal{I}$ called* domain *together with a function $\cdot^\mathcal{I}$ mapping individual names to elements of $\Delta^\mathcal{I}$, concept names to subsets of $\Delta^\mathcal{I}$, and role names to subsets of $\Delta^\mathcal{I} \times \Delta^\mathcal{I}$.*
The function $\cdot^\mathcal{I}$ is inductively extended to roles and concepts as shown in Table 1. An interpretation $\mathcal{I}$ satisfies *an axiom φ (written: $\mathcal{I} \models \varphi$) if the respective condition is satisfied:*

- $\mathcal{I} \models S \sqsubseteq R$ *if* $S^\mathcal{I} \subseteq R^\mathcal{I}$,
- $\mathcal{I} \models S_1 \circ \ldots \circ S_n \sqsubseteq R$ *if* $S_1^\mathcal{I} \circ \ldots \circ S_n^\mathcal{I} \sqsubseteq R^\mathcal{I}$ *($\circ$ being overloaded to denote the standard composition of binary relations here),*
- $\mathcal{I} \models C \sqsubseteq D$ *if* $C^\mathcal{I} \subseteq D^\mathcal{I}$,
- $\mathcal{I} \models C(\mathtt{a})$ *if* $\mathtt{a}^\mathcal{I} \in C^\mathcal{I}$,
- $\mathcal{I} \models R(\mathtt{a}, \mathtt{b})$ *if* $(\mathtt{a}^\mathcal{I}, \mathtt{b}^\mathcal{I}) \in R^\mathcal{I}$,
- $\mathcal{I} \models \neg S(\mathtt{a}, \mathtt{b})$ *if* $(\mathtt{a}^\mathcal{I}, \mathtt{b}^\mathcal{I}) \notin S^\mathcal{I}$,
- $\mathcal{I} \models \mathtt{a} \approx \mathtt{b}$ *if* $\mathtt{a}^\mathcal{I} = \mathtt{b}^\mathcal{I}$,
- $\mathcal{I} \models \mathtt{a} \not\approx \mathtt{b}$ *if* $\mathtt{a}^\mathcal{I} \neq \mathtt{b}^\mathcal{I}$.

An interpretation $\mathcal{I}$ satisfies *a knowledge base* $\mathcal{K}$ *(we then also say that* $\mathcal{I}$ *is a* model *of* $\mathcal{K}$ *and write* $\mathcal{I} \models \mathcal{K}$*) if it satisfies all axioms of* $\mathcal{K}$*. A knowledge base* $\mathcal{K}$ *is* (finitely) satisfiable *if it has a (finite) model. Two knowledge bases are* equivalent *if they have exactly the same models. They are* (finitely) equisatisfiable *if either both are (finitely) unsatisfiable or both are (finitely) satisfiable.*

The description logic $\mathcal{SHOIQB}$ is obtained from $\mathcal{SROIQB}$ by discarding the universal role $\overline{\top}$ as well as the Self concept and allowing only RIAs of the form $R \sqsubseteq S$ or $R \circ R \sqsubseteq R$. If we also disallow $R \circ R \sqsubseteq R$, we obtain $\mathcal{ALCHOIQB}$. For any of these three logics, replacing $\mathcal{B}$ in the name by b disallows role negation (but preserves role difference) while removing $\mathcal{B}$ entirely also disallows role conjunction, disjunction and difference. Dropping $\mathcal{O}$ from any description logic's name disables nominal concepts $\{o\}$, while dropping $\mathcal{I}$ disables role inverses $\cdot^-$, and dropping $\mathcal{H}$ disables RIAs of the form $R \sqsubseteq S$. For any description logic $\mathcal{L}$ that does not feature the Self concept (the universal role $\overline{\top}$), we denote by $\mathcal{L}^{\mathsf{Self}}$ (by $\mathcal{L}_{\overline{\top}}$) the logic with this feature added.

Queries. In queries, we use *variables* from a countably infinite set $\mathbf{V}$. A Boolean *positive two-way regular path query* (P2RPQ) is a formula $\exists \boldsymbol{x}.\varphi$, where φ is a positive Boolean expression (i.e., one using only $\wedge$ and $\vee$) over atoms of the form $C(t)$ or $T(s,t)$, where s and t are elements of $\boldsymbol{x} \cup \mathbf{N_I}$, C is a concept, and T is a *regular role expression* from $\mathbf{T}$, defined by

$$\mathbf{T} ::= \mathbf{R} \mid \mathbf{T} \cup \mathbf{T} \mid \mathbf{T} \circ \mathbf{T} \mid \mathbf{T}^* \mid id(\mathbf{C}).$$

If q does not use disjunction and all T are simple roles, it is called a *conjunctive query* (CQ). A *variable assignment* π for $\mathcal{I}$ is a mapping $\mathbf{V} \to \Delta^\mathcal{I}$. For $x \in \mathbf{V}$, we set $x^{\mathcal{I},\pi} := \pi(x)$; for $\mathtt{c} \in \mathbf{N_I}$, we set $\mathtt{c}^{\mathcal{I},\pi} := \mathtt{c}^\mathcal{I}$. $T(s,t)$ evaluates to true under π and $\mathcal{I}$ if $(s^{\mathcal{I},\pi}, t^{\mathcal{I},\pi}) \in T^\mathcal{I}$, with $T^\mathcal{I}$ obtained as detailed in Table 2. $C(t)$ evaluates to true under π and $\mathcal{I}$ if $t^{\mathcal{I},\pi} \in C^\mathcal{I}$. A P2RPQ $q = \exists \boldsymbol{x}.\varphi$ is *satisfied* by $\mathcal{I}$ (written: $\mathcal{I} \models q$) if there is a variable assignment π (called *match*) such that φ evaluates to true under $\mathcal{I}$ and π. A P2RPQ q is *(finitely) entailed* from a KB $\mathcal{K}$ if every (finite) model of $\mathcal{K}$ satisfies q.

3 Subject: Extending Knowledge Bases by Presburger-Style Concept Cardinality Constraints

In this section, we introduce extended cardinality boxes (and several restricted versions thereof) as means for expressing quantitative global knowledge.

Table 2: Semantics of regular role expressions for interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$.

Name	Syntax	Semantics
union	$T_1 \cup T_2$	$T_1^{\mathcal{I}} \cup T_2^{\mathcal{I}}$
concatenation	$T_1 \circ T_2$	$T_1^{\mathcal{I}} \circ T_2^{\mathcal{I}}$
Kleene star	T^*	$\bigcup_{i \geq 0} (T^{\mathcal{I}})^i$
concept test	$id(C)$	$\{(x,x) \mid x \in C^{\mathcal{I}}\}$

Definition 3 (concept cardinality constraint, ECbox, RCbox, ERCbox). *A* concept cardinality constraint *(short: constraint)* $\mathfrak{c}$ *is an expression of the form*

$$n_0 + n_1|\mathtt{A}_1| + \ldots + n_k|\mathtt{A}_k| \leq m_0 + m_1|\mathtt{B}_1| + \ldots + m_\ell|\mathtt{B}_\ell|, \tag{1}$$

where $\mathtt{A}_1, \ldots, \mathtt{A}_k, \mathtt{B}_1, \ldots, \mathtt{B}_\ell$ *are concept names and all* n_i *and* m_i *are natural numbers. A concept cardinality constraint is* restricted *if* $n_0 = m_0 = 0$ *and* semi-restricted *if* $m_0 = 0$*. An* extended cardinality box *(ECbox) is a positive Boolean combination of concept cardinality constraints. A* restricted cardinality box *(RCbox) is a conjunction of restricted cardinality constraints. An* extended restricted cardinality box *(ERCbox) is a positive Boolean combination of semi-restricted cardinality constraints.*
Satisfaction of a concept cardinality constraint $\mathfrak{c}$ *by an interpretation* $\mathcal{I}$ *(written as* $\mathcal{I} \models \mathfrak{c}$*) is verified as follows: every expression* $|\mathtt{A}|$ *is mapped to* $|\mathtt{A}^{\mathcal{I}}|$*. The constraint is evaluated in the straightforward way over* $\mathbb{N}^\infty$*. Satisfaction of constraints is then lifted to satisfaction of ECboxes in the obvious manner.*

Definition 4 (EKB, ERKB, RKB). *For some description logic* $\mathcal{L}$*, an* extended $\mathcal{L}$ knowledge base *($\mathcal{L}$ EKB) is a quadruple* $(\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$ *where* $(\mathcal{A}, \mathcal{T}, \mathcal{R})$ *is an* $\mathcal{L}$ *knowledge base and* $\mathcal{E}$ *is an ECbox. An EKB is a* restricted knowledge base *(RKB) if* $\mathcal{E}$ *is an RCbox. It is an* extended restricted knowledge base *(ERKB) if* $\mathcal{E}$ *is an ERCbox.*

Obviously, RKBs (RCboxes) are properly subsumed by ERKBs (ERCboxes) which in turn are properly subsumed by EKBs (ECboxes). One general insight of this paper is that upper complexity bounds persist when generalizing the previously defined RCboxes to the newly defined, more expressive ERCboxes.
Our syntactic formulation of ECboxes is somewhat more restrictive than that in prior work [5], but we will show that the differences are immaterial. Using our more restricted form allows for a more uniform presentation of our results.
First, the original work allows expressions $|C|$ for arbitrary concept descriptions C. We note that our definition does not restrict expressivity since general Tboxes allow for axioms $\mathtt{A} \equiv C$, so complex concept expressions in cardinality constraints can be replaced by fresh concept names and defined in the Tbox. Resorting to plain concept names in constraints allows us to consider cardinality boxes uniformly independently from the used description logic.
Second, instead of positive weighted sums of concept cardinalities as left and right hand sides, the original work allows for arbitrary functions built from integers z and expressions of the form $|\mathtt{A}|$ using functions $+$ (binary) and $z\cdot$ (unary). It is, however,

easy to see that each comparison on such more liberal expressions can be polynomially translated into an equivalent comparison of positive weighted sums.
Third, the original work allows for extended cardinality constraints using other modes of comparison than just "$\leq$": $\alpha = \beta$, $\alpha < \beta$, or n dvd α. However, all these constraints can be rewritten into (combinations of) constraints only using "$\leq$" as follows: $\alpha = \beta$ is replaced by $(\alpha \leq \beta) \wedge (\beta \leq \alpha)$, $\alpha < \beta$ by $\alpha{+}1 \leq \beta$,[3] and n dvd α can be rewritten into $(n|\mathtt{A}| \leq \alpha) \wedge (\alpha \leq n|\mathtt{A}|)$ for some fresh concept name A.
Fourth, the original work defined ECboxes as arbitrary (not just positive) Boolean combinations of constraints. However, this work only considered finite models. We note that under this assumption, negated constraints can be rewritten into negation-free (combinations of) constraints in the following way: replace $\neg(\alpha \leq \beta)$ by $\beta{+}1 \leq \alpha$, $\neg(\alpha < \beta)$ by $\beta \leq \alpha$, $\neg(\alpha = \beta)$ by $(\alpha{+}1 \leq \beta) \vee (\beta{+}1 \leq \alpha)$, and $\neg(n \text{ dvd } \alpha)$ by $(n|\mathtt{A}| + |\mathtt{B}| \leq \alpha) \wedge (\alpha \leq n|\mathtt{A}| + |\mathtt{B}|) \wedge (1 \leq |\mathtt{B}|) \wedge (|\mathtt{B}| + 1 \leq n)$ for fresh concept names A and B. Note again, that this rewriting is not equivalent for infinite models.[4]

As observed before, ECboxes allow to express nominal concepts by enforcing that a concept must have cardinality exactly one. However, this is not possible with RCboxes nor ERCboxes.

4 Exposition: Statement of Results

With the notion of ECboxes, RCboxes and ERCboxes in place, we can now formally state results that can be derived from prior work before giving an outlook on the results established in this paper. We first note some results that can be obtained as easy consequences of previous publications.

- Finite satisfiability of Abox-free $\mathcal{SHQ}b$ RKBs is in ExpTime. For $\mathcal{ALCHQ}b$, this is an immediate consequence of earlier work on $\mathcal{ALCSCC}$ RCboxes [2]. Adding transitivity is possible since it can be handled via the classical "box pushing" approach [33,29].
- Finite satisfiability of $\mathcal{SHOQB}$ EKBs is in NExpTime. This follows from [3] together with the observation that the logic $\mathcal{ALCSCC}^{++}$ considered there allows to express qualified number restrictions, nominals (hence also Aboxes), and arbitrary Boolean role expressions. Again, transitivity can be dealt with via "box pushing".
- Finite CQ entailment over Abox-free $\mathcal{ALCHQ}b$ RKBs is in 2ExpTime, as immediate consequence of the corresponding result for $\mathcal{ALCSCC}$ RCboxes in [3].

At the core of our method is the insight that expressive description logics in and of themselves hold enough expressive means to simulate ECboxes without noteworthy blow-up. We will show that:

[3] For this, we have to postulate $\infty < \infty$, which is debatable, but could be justified by the fact that there is an injective, non-surjective mapping between any two countably infinite sets.

[4] In fact, the constraint expression $(1 + |\mathtt{A}| = |\mathtt{A}|) \wedge \neg(|\mathtt{A}| = |\mathtt{B}|)$ would enforce finiteness of the extension of B, which is not axiomatizable in first order logic, neither finitely nor infinitely. For good reasons (see Section 5) we define ECboxes in a way that a first-order axiomatization is still possible.

(A) ERCboxes can be succinctly simulated in any description logic that can express $\mathcal{ALCIQ}_{\overline{\top}}$ GCIs and
(B) ECboxes can be succinctly simulated in any description logic that can express $\mathcal{ALCOIQ}$ GCIs.

This "simulation", made formally precise in Section 6, is sufficiently authentic for both satisfiability checking and query entailment. Consequently, we are able to significantly strengthen the aforementioned results as follows (further detailed in Section 7):

- Satisfiability and finite satisfiability of $\mathcal{SHIQ}b_{\overline{\top}}^{\mathsf{Self}}$ ERKBs is ExpTime-complete.
- Satisfiability and finite satisfiability of $\mathcal{SHOIQB}^{\mathsf{Self}}$ EKBs is NExpTime-complete.
- Entailment of P2RPQs as well as finite entailment of CQs from $\mathcal{ALCHIQ}b_{\overline{\top}}^{\mathsf{Self}}$ ERKBs are 2ExpTime-complete.
- Entailment of unions of conjunctive queries from $\mathcal{ALCHOIQ}b$ EKBs is decidable and coN2ExpTime-hard.

Yet, before going into the details of our translation, we have to take care of a nuisance, arising from counting in the presence of infinity.

5 Interlude: Countability

Dealing with infinity can be tricky [7,9,10,27]. In the general case, allowing for infinite models might require us to account for the presence of several distinct infinite cardinalities. In the realms of first-order logic, the Löwenheim-Skolem Theorem [31] ensures that it suffices to consider models of countable cardinality, in which only one type of infinity can occur.
In our setting, however, expressibility in first-order logic cannot be easily taken for granted. In fact, even the very simple RCbox $(|\mathtt{A}| \leq |\mathtt{B}|) \wedge (|\mathtt{B}| \leq |\mathtt{A}|)$, stating that A and B contain the same number of individuals, cannot be expressed using a first-order sentence (as can be shown by an easy argument using Ehrenfeucht-Fraïssé games).
We manage to resolve the issue by showing that ECboxes can be expressed by countable (but possibly infinite) first-order theories, noting that Löwenheim-Skolem still applies in this case.

Lemma 1 *Let $\mathcal{E}$ be an ECbox. Then there exists a countable first-order theory $\Phi_{\mathcal{E}}$ logically equivalent to $\mathcal{E}$.*

Proof. We construct $\Phi_{\mathcal{E}}$ from $\mathcal{E}$. For convenience, we introduce some notation: Given a concept name A and a number $n \in \mathbb{N}$, we let $\boldsymbol{f}(|\mathtt{A}| \geq n)$ denote the first-order sentence

$$\exists x_1, \dots x_n. \bigwedge_{1 \leq i \leq n} \mathtt{A}(x_i) \ \wedge \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j,$$

and we let $\boldsymbol{f}(|\mathtt{A}| \leq n)$ denote the first-order sentence

$$\forall x_0, x_1, \dots x_n. \Big(\bigwedge_{0 \leq i \leq n} \mathtt{A}(x_i) \Big) \rightarrow \bigvee_{0 \leq i < j \leq n} x_i = x_j.$$

Note that the first-order sentences have precisely the intended meaning. Now, considering some cardinality constraint $\mathfrak{c}$ of the form

$$n_0 + n_1|\mathtt{A}_1| + \ldots + n_k|\mathtt{A}_k| \leq m_0 + m_1|\mathtt{B}_1| + \ldots + m_\ell|\mathtt{B}_\ell|,$$

we let $Bad_\mathfrak{c}$ denote the (most likely infinite) set of first-order sentences

$$\Big\{ \bigwedge_{1\leq i\leq k} \boldsymbol{f}(|\mathtt{A}_i| \geq a_i) \wedge \bigwedge_{1\leq i\leq \ell} \boldsymbol{f}(|\mathtt{B}_i| \leq b_i) \;\Big|\; n_0 + \sum_{1\leq i\leq k} n_i a_i > m_0 + \sum_{1\leq i\leq \ell} m_i b_i \Big\},$$

and note that $\mathcal{I} \models \mathfrak{c}$ if and only if $\mathcal{I} \not\models \varphi$ holds for all $\varphi \in Bad_\mathfrak{c}$. Next, let $\mathfrak{C}$ denote the set of all constraints occurring in $\mathcal{E}$. We let $\mathfrak{Bad}$ consist of all sets $\mathfrak{D} \subseteq \mathfrak{C}$ for which the Boolean expression obtained from $\mathcal{E}$ by replacing all $\mathfrak{c} \in \mathfrak{D}$ with **false** and all $\mathfrak{c} \in \mathfrak{C} \setminus \mathfrak{D}$ with **true** evaluates to **false**. Finally, we let $\Phi_\mathcal{E}$ consist of all sentences $\neg(\varphi_1 \wedge \ldots \wedge \varphi_m)$ for which there is some $\{\mathfrak{c}_1, \ldots, \mathfrak{c}_m\} \in \mathfrak{Bad}$ such that $\varphi_i \in Bad_{\mathfrak{c}_i}$ for $1 \leq i \leq m$. ◠

As planned, we can now use this insight to make sure that even in the presence of ECboxes, we can restrict our attention to countable models, as long as the rest of the knowledge base is expressible in first-order logic.

Theorem 2 *Let $\mathcal{L}$ be a description logic such that any $\mathcal{L}$ knowledge base $(\mathcal{A}, \mathcal{T}, \mathcal{R})$ is equivalent to a countable first-order logic theory $\Psi_{(\mathcal{A},\mathcal{T},\mathcal{R})}$.*

1. *Every satisfiable $\mathcal{L}$ EKB has a countable model.*
2. *An $\mathcal{L}$ EKB $\mathcal{K}$ entails a P2RPQ q iff q is satisfied by all countable models of $\mathcal{K}$.*

Proof. 1. This is actually a special case of the case below: pick $q = \exists x.\bot(x)$.

2. The "only if" direction is trivial. For the "if" direction, first observe that any Boolean P2RPQ q can be expressed as a possibly infinite disjunction $\bigvee_{q' \in Q_q} q'$ of (finite) Boolean CQs. Let $\mathcal{K} = (\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$. Toward a contradiction, suppose $\mathcal{K} \not\models q$, i.e., there is a model $\mathcal{I}$ (of arbitrary cardinality) such that $\mathcal{I} \models \mathcal{K}$ but $\mathcal{I} \not\models q$. Then, by Lemma 1, we know that $\mathcal{I}$ is a model of the countable first-order theory $\Psi_{(\mathcal{A},\mathcal{T},\mathcal{R})} \cup \Phi_\mathcal{E} \cup \{\neg q' \mid q' \in Q_q\}$. Now we can apply the Löwenheim-Skolem Theorem downward and obtain that there must be a countable model $\mathcal{J}$ of this theory as well. By construction, $\mathcal{J}$ is a countable model of $\mathcal{K}$ but does not satisfy q, a contradiction. ◠

6 Development: Eliminating Cardinality Boxes

The basic underlying idea of our method is to model satisfaction of cardinality constraints by performing the necessary calculations and comparisons "physically" inside the model, using the domain elements for tallying. However, in the case of finite interpretations, it might happen that evaluating the cardinality constraints produces numbers that are greater than the number of domain elements (note that this danger is material, since expressive description logics allow for enforcing restricted domain sizes). Hence, we somehow have to make sure that our models are allowed to contain enough domain elements.

6.1 Shift: Making Space through Relativization

To this end, we employ a folklore technique called *relativization*, through which a (possibly domain-restricting) knowledge base $\mathcal{K}$ is transformed into a knowledge base $\mathcal{K}^{\wr}$ that allows for models with arbitrary domain sizes (by means of admitting "silent" or "non-active" domain elements, which do not participate in any relation), but every model of $\mathcal{K}^{\wr}$ "contains" a model of $\mathcal{K}$ in a formally defined way, so $\mathcal{K}^{\wr}$ is an authentic replacement of $\mathcal{K}$ when it comes to satisfiability testing or querying. In the context of description logics, similar techniques have been applied in [15,18].

Definition 5 (relativization). *We let $\top_{\text{new}}$ be a fresh concept name. The function $\cdot^{\wr}$ mapping concepts to concepts is recursively defined as follows:*

$$\begin{aligned}
\mathtt{A}^{\wr} &= \mathtt{A} & \{\mathtt{a}\}^{\wr} &= \{\mathtt{a}\}\\
\top^{\wr} &= \top_{\text{new}} & (\forall R.C)^{\wr} &= \top_{\text{new}} \sqcap \forall R.(\neg\top_{\text{new}} \sqcup C^{\wr})\\
\bot^{\wr} &= \bot & (\exists R.C)^{\wr} &= \exists R.C^{\wr}\\
(\neg C)^{\wr} &= \top_{\text{new}} \sqcap \neg C^{\wr} & ({\geqslant}nS.C)^{\wr} &= {\geqslant}nS.C^{\wr}\\
(C_1 \sqcap C_2)^{\wr} &= C_1^{\wr} \sqcap C_2^{\wr} & ({\leqslant}nS.C)^{\wr} &= {\leqslant}nS.C^{\wr}\\
(C_1 \sqcup C_2)^{\wr} &= C_1^{\wr} \sqcup C_2^{\wr} & (\exists S.\mathsf{Self})^{\wr} &= \exists S.\mathsf{Self}
\end{aligned}$$

Given a P2RPQ q, we let $q^{\wr}$ denote the query obtained by replacing each concept C in q by $C^{\wr}$. Moreover, we extend $\cdot^{\wr}$ to EKBs $\mathcal{K} = (\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$ by letting $\mathcal{K}^{\wr} = (\mathcal{A}', \mathcal{T}', \mathcal{R}, \mathcal{E})$, where $\mathcal{A}'$ contains

- $\top_{\text{new}}(\mathtt{a})$ *for every* $\mathtt{a} \in \mathbf{N_I}$ *occurring in* $\mathcal{K}$,
- *for every assertion* $C(\mathtt{a})$ *from* $\mathcal{A}$ *the assertion* $C^{\wr}(\mathtt{a})$, *and*
- *all assertions of the form* $\mathtt{a} \approx \mathtt{b}$, $\mathtt{a} \not\approx \mathtt{b}$, $\neg S(\mathtt{a}, \mathtt{b})$, *and* $R(\mathtt{a}, \mathtt{b})$ *from* $\mathcal{A}$,

while $\mathcal{T}'$ contains

- $\mathtt{A} \sqsubseteq \top_{\text{new}}$ *for every* $\mathtt{A} \in \mathbf{N_C}$ *occurring in* $\mathcal{K}$,
- $\exists\mathtt{P}.\top \sqsubseteq \top_{\text{new}}$ *and* $\top \sqsubseteq \forall\mathtt{P}.\top_{\text{new}}$ *for every* $\mathtt{P} \in \mathbf{N_R} \setminus \{\overline{\top}\}$ *occurring in* $\mathcal{K}$, *as well as*
- *for every GCI* $C_1 \sqsubseteq C_2$ *from* $\mathcal{T}$ *the GCI* $C_1^{\wr} \sqsubseteq C_2^{\wr}$.

It is now not too hard to establish the following lemma, explicating the formerly claimed very close connection between the models of $\mathcal{K}$ and $\mathcal{K}^{\wr}$.

Lemma 3 (relativization: model synchronicity) *Let $\mathcal{K} = (\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$ be an EKB and let $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ be in interpretation with $\top_{\text{new}}^{\mathcal{J}} \neq \emptyset$. Then $\mathcal{J}$ is a (finite) model of $\mathcal{K}^{\wr}$ if and only if there exists a (finite) set $\Delta_{\blacksquare}$ and a (finite) model $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ of $\mathcal{K}$ such that $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}} \cup \Delta_{\blacksquare}$ and $\cdot^{\mathcal{J}} = \cdot^{\mathcal{I}} \cup \{\top_{\text{new}} \mapsto \Delta^{\mathcal{I}}\}$.*

Proof. (Sketch.) The not immediate cases are a direct consequence of the correspondence $C^{\mathcal{I}} = C^{\wr\mathcal{J}}$ which is proven by induction over the structure of C.

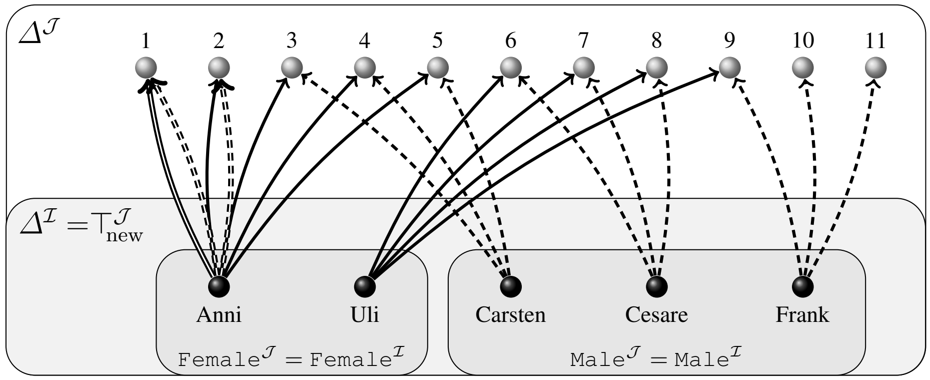


Fig. 1: Illustration of construction enforcing constraints.

6.2 Episode: Illustrative Example

We now know how to ensure that a model can contain enough elements for counting and hence are able to avoid “out of memory errors” in the course of our model-internal computation. Next, we describe in detail how to express concept cardinality constraints and consequently ECboxes polynomially with in-house means of expressive description logics, requiring just some extra vocabulary.

We first describe the core idea behind our modeling by means of an easy example. Assume, we would like to implement the constraint

$$1 + 4 \cdot |\texttt{Female}| \leq 2 + 3 \cdot |\texttt{Male}|$$

on the set of this volume’s editors, which, of course, can be simplified by subtracting one on both sides, but we will not do so for the sake of the example. Assume that, by means of relativization, we have already ensured that as many as needed “silent elements” can be present in a model. In order to ensure that the constraint is satisfied, we proceed as follows (aiming at a setting as displayed in Fig. 1, where the silent elements are in the top line while the “proper elements” can be found in the bottom line): we introduce several types of left-hand-side roles (denoted $\texttt{LHS}_{\ldots}$, depicted by solid arrows in the figure) and right-hand-side roles (denoted $\texttt{RHS}_{\ldots}$, depicted by dashed arrows in the figure) and we make sure that every individual in `Female` has (at least) four outgoing (single-line) left-hand-side roles, while every individual in `Male` has (at most) three outgoing (single-line) right-hand-side roles and any individual not in `Male` has no such outgoing roles whatsoever. Also, to account for the left- and right-hand-side constant terms, we pick one volunteering domain element, say Anni, as the source of one (double-line) left-hand-side role and of two (double-line) right-hand side roles. Then, we make sure that every domain element may receive at most one left-hand-side role. Under these circumstances, the “$\leq$” condition can be enforced by requiring that any element receiving a left-hand-side role must also be receiving a right-hand-side role. This (somewhat simplified) example will hopefully elucidate the modeling presented in the following.

6.3 Modulation: General Construction and Proof

After explaining the underlying ideas of our construction, we now provide the general definition of the technique of eliminating ECboxes from EKBs yielding plain KBs. We introduce the applied transformation and formally show that it has the announced properties.

Definition 6 (knowledge base transformation, $\mathcal{K}^{\mathfrak{tr}}$). *Let $\mathfrak{c}$ be the following cardinality constraint:*

$$n_0 + n_1|\mathtt{A}_1| + \ldots + n_k|\mathtt{A}_k| \leq m_0 + m_1|\mathtt{B}_1| + \ldots + m_\ell|\mathtt{B}_\ell|.$$

Then the Tbox $\mathcal{T}_\mathfrak{c}$ contains the following axioms (with fresh role names $\mathtt{AUX}$, $\mathtt{RHS}^i_\mathfrak{c}$, $\mathtt{LHS}^i_\mathfrak{c}$ *and – if needed – a fresh individual name* $\mathtt{o}$*):*

$$\top \sqsubseteq \exists\mathtt{AUX}.\big(\top_{\text{new}} \sqcap {\geqslant} n_0 \mathtt{LHS}^0_\mathfrak{c}.\top\big) \quad (2)$$

$$\mathtt{A}_i \sqsubseteq {\geqslant} n_i \mathtt{LHS}^i_\mathfrak{c}.\top \qquad \textit{for} \quad 1 \leq i \leq k \quad (3)$$

$$\exists {\mathtt{LHS}^i_\mathfrak{c}}^-.\top \sqcap \exists {\mathtt{LHS}^j_\mathfrak{c}}^-.\top \sqsubseteq \bot \qquad \textit{for} \quad 0 \leq i < j \leq k \quad (4)$$

$$\top \sqsubseteq {\leqslant} 1.{\mathtt{LHS}^i_\mathfrak{c}}^-.\top \qquad \textit{for} \quad 0 \leq i \leq k \quad (5)$$

$$\exists \mathtt{RHS}^0_\mathfrak{c}.\top \sqsubseteq \bot \qquad \textit{in case} \quad m_0 = 0 \quad (6)$$

$$\exists \mathtt{RHS}^0_\mathfrak{c}.\top \sqsubseteq \{\mathtt{o}\} \qquad \textit{in case} \quad m_0 > 0 \quad (7)$$

$$\exists \mathtt{RHS}^i_\mathfrak{c}.\top \sqsubseteq \mathtt{B}_i \qquad \textit{for} \quad 1 \leq i \leq \ell \quad (8)$$

$$\top \sqsubseteq {\leqslant} m_i.\mathtt{RHS}^i_\mathfrak{c} \qquad \textit{for} \quad 0 \leq i \leq \ell \quad (9)$$

$$\mathtt{Cstrt}_\mathfrak{c} \sqcap \bigsqcup_{0 \leq i \leq k} \exists {\mathtt{LHS}^i_\mathfrak{c}}^-.\top \sqsubseteq \bigsqcup_{0 \leq i \leq \ell} \exists {\mathtt{RHS}^i_\mathfrak{c}}^-.\top \quad (10)$$

Given an ECbox $\mathcal{E}$, let $C_\mathcal{E}$ be the concept expression obtained from $\mathcal{E}$ by replacing every $\mathfrak{c}$ in $\mathcal{E}$ by $\mathtt{Cstrt}_\mathfrak{c}$, every $\wedge$ by $\sqcap$, and every $\vee$ by $\sqcup$. Let $\mathsf{Sync}(\mathtt{Cstrt}_\mathfrak{c})$ denote $\{\exists\overline{\top}.\mathtt{Cstrt}_\mathfrak{c} \sqsubseteq \mathtt{Cstrt}_\mathfrak{c}\}$ if the underlying description logic supports the universal role $\overline{\top}$ and $\{\top \sqsubseteq \exists\mathtt{AUX}^-.\{\mathtt{o}\},\ \exists\mathtt{AUX}.\mathtt{Cstrt}_\mathfrak{c} \sqsubseteq \forall\mathtt{AUX}.\mathtt{Cstrt}_\mathfrak{c}\}$ otherwise.
Then, we let

$$\mathcal{T}_\mathcal{E} = \{\top \sqsubseteq C_\mathcal{E}\} \cup \bigcup_{\mathfrak{c} \textit{ from } \mathcal{E}} \mathcal{T}_\mathfrak{c} \cup \mathsf{Sync}(\mathtt{Cstrt}_\mathfrak{c}). \quad (11)$$

Finally, for an EKB $\mathcal{K} = (\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$ with $\mathcal{K}^{\natural} = (\mathcal{A}', \mathcal{T}', \mathcal{R}, \mathcal{E})$ let $\mathcal{K}^{\mathfrak{tr}} = (\mathcal{A}', \mathcal{T}' \cup \mathcal{T}_\mathcal{E}, \mathcal{R}, \emptyset)$ be the corresponding transformed *KB.*

The following observations are immediate from the construction of $\mathcal{K}^{\mathfrak{tr}}$.

Lemma 4 (syntactic properties of $\mathcal{K}^{\mathfrak{tr}}$) *For any EKB $\mathcal{K}$ in some description logic $\mathcal{L}$:*

1. *$\mathcal{K}^{\mathfrak{tr}}$ can be computed from $\mathcal{K}$ in polynomial time.*
2. *If $\mathcal{L}$ subsumes $\mathcal{ALCIQ}_{\overline{\top}}$ and $\mathcal{K}$ is an ERKB, then $\mathcal{K}^{\mathfrak{tr}}$ is a (plain) $\mathcal{L}$ KB.*
3. *If $\mathcal{L}$ subsumes $\mathcal{ALCOIQ}$, then $\mathcal{K}^{\mathfrak{tr}}$ is a (plain) $\mathcal{L}$ KB.*

Next we prove a rather close relationship between the models of $\mathcal{K}$ and $\mathcal{K}^{\mathfrak{tr}}$.

Lemma 5 (semantic properties of $\mathcal{K}^{\mathsf{tr}}$) *For an EKB $\mathcal{K}$ in some description logic $\mathcal{L}$, the following hold:*

1. *For every (finite) model $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ of $\mathcal{K}^{\mathsf{tr}}$, the interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ with $\Delta^{\mathcal{I}} = \top_{\text{new}}^{\mathcal{J}}$ and $\cdot^{\mathcal{I}}$ the appropriate restriction of $\cdot^{\mathcal{J}}$ is a (finite) model of $\mathcal{K}$.*
2. *For every (finite) countable model $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ of $\mathcal{K}$ there is a (finite) countable model $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ of $\mathcal{K}^{\mathsf{tr}}$ such that with $\Delta^{\mathcal{I}} = \top_{\text{new}}^{\mathcal{J}}$ and $\cdot^{\mathcal{I}}$ is the appropriate restriction of $\cdot^{\mathcal{J}}$.*

Proof. Let $\mathcal{K} = (\mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{E})$. We show the two parts consecutively.

1. To show the first part, assume a (finite) model $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ of $\mathcal{K}^{\mathsf{tr}}$. We now show that $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ is a model of $\mathcal{K}$. For all Abox, Tbox, and Rbox axioms, satisfaction follows from Lemma 3. Now consider $\mathcal{E}$. We pick an arbitrary $\delta \in \Delta^{\mathcal{J}}$ and let $\mathfrak{C} = \{\mathfrak{c} \mid \delta \in \mathtt{Cstrt}_{\mathfrak{c}}^{\mathcal{J}}\}$. Due to the axiom $\top \sqsubseteq C_{\mathcal{E}}$, we know that $\mathfrak{C}$ is such that simultaneous satisfaction of all $\mathfrak{c} \in \mathfrak{C}$ implies satisfaction of $\mathcal{E}$. Hence we proceed to prove that this is indeed the case. First note that the axioms $\mathsf{Sync}(\mathtt{Cstrt}_{\mathfrak{c}})$ ensure $\mathtt{Cstrt}_{\mathfrak{c}}^{\mathcal{J}} = \Delta^{\mathcal{J}}$ for every $\mathfrak{c} \in \mathfrak{C}$. Furthermore, for any $\mathfrak{c} \in \mathfrak{C}$ of the form

$$n_0 + n_1|\mathtt{A}_1| + \ldots + n_k|\mathtt{A}_k| \leq m_0 + m_1|\mathtt{B}_1| + \ldots + m_\ell|\mathtt{B}_\ell|,$$

satisfaction of the $\mathfrak{c}$ by $\mathcal{J}$ follows from the following three inequalities:

$$n_0 + n_1|\mathtt{A}_1^{\mathcal{J}}| + \ldots + n_k|\mathtt{A}_k^{\mathcal{J}}| \leq \sum_{i=0}^{k} |\{\delta \mid (\delta', \delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}|, \qquad (\#)$$

$$\sum_{i=0}^{k} |\{\delta \mid (\delta', \delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}| \leq \sum_{i=0}^{\ell} |\{\delta \mid (\delta', \delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}|, \qquad (\%)$$

$$\sum_{i=0}^{\ell} |\{\delta \mid (\delta', \delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}| \leq m_0 + m_1|\mathtt{B}_1^{\mathcal{J}}| + \ldots + m_\ell|\mathtt{B}_\ell^{\mathcal{J}}|. \qquad (\phi)$$

We will now consecutively show each of these statements.

(#) On one hand, the axioms $\top \sqsubseteq \leqslant 1.\mathtt{LHS}_{\mathfrak{c}}^{i\,-}.\top$ make sure that

$$|\{\delta \mid (\delta', \delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}| = |\mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}|.$$

On the other hand, whenever $1 \leq i \leq k$, the axiom $\mathtt{A}_i \sqsubseteq \geqslant n_i \mathtt{LHS}_{\mathfrak{c}}^{i}.\top$ ensures

$$|\mathtt{A}_i^{\mathcal{J}}| \cdot n_i \leq |\mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}|,$$

while for $i = 0$, the axiom $\top \sqsubseteq \exists \mathtt{AUX}.\geqslant n_0 \mathtt{LHS}_{\mathfrak{c}}^{0}.\top$ enforces that there must be some $\delta \in (\geqslant n_0 \mathtt{LHS}_{\mathfrak{c}}^{0}.\top)^{\mathcal{J}}$ and consequently

$$n_0 \leq |\mathtt{LHS}_{\mathfrak{c}}^{0\,\mathcal{J}}|.$$

Putting everything together, we obtain

$$n_0 + \sum_{i=1}^{k} n_i|\mathtt{A}_i^{\mathcal{J}}| \leq \sum_{i=0}^{k} |\mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}| = \sum_{i=0}^{k} |\{\delta \mid (\delta', \delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}|$$

as claimed.

(𝄋) First note that the axioms of the form $\exists\mathtt{LHS}_{\mathfrak{c}}^{i\,-}.\top \sqcap \exists\mathtt{LHS}_{\mathfrak{c}}^{j\,-}.\top \sqsubseteq \bot$ ensure

$$|\{\delta \mid (\delta',\delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le k\}| = \sum_{i=0}^{k} |\{\delta \mid (\delta',\delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}|.$$

Also, the axiom $\mathtt{Cstrt}_{\mathfrak{c}} \sqcap \bigsqcup_{0\le i\le k} \exists\mathtt{LHS}_{\mathfrak{c}}^{i\,-}.\top \sqsubseteq \bigsqcup_{0\le i\le \ell} \exists\mathtt{RHS}_{\mathfrak{c}}^{i\,-}.\top$ enforces

$$\{\delta \mid (\delta',\delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le k\} \subseteq \{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le \ell\}$$

(remembering that $\mathtt{Cstrt}_{\mathfrak{c}}^{\mathcal{J}} = \Delta^{\mathcal{J}}$) and consequently

$$|\{\delta \mid (\delta',\delta) \in \mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le k\}| \le |\{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le \ell\}|.$$

It remains to note that

$$|\{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}, 0 \le i \le \ell\}| \le \sum_{i=0}^{\ell} |\{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}|$$

holds unconditionally, so putting the established correspondences together shows our claim.

(𝄌) First note that, the axiom $\exists\mathtt{RHS}_{\mathfrak{c}}^{0}.\top \sqsubseteq \{\mathtt{o}\}$ (or, alternatively, $\exists\mathtt{RHS}_{\mathfrak{c}}^{0}.\top \sqsubseteq \bot$ in case $m_0 = 0$) ensures $\delta = \mathtt{o}^{\mathcal{J}}$ whenever we find $(\delta,\delta') \in \mathtt{RHS}_{\mathfrak{c}}^{0\,\mathcal{J}}$. But therefrom, using the axiom $\top \sqsubseteq \leqslant m_0.\mathtt{RHS}_{\mathfrak{c}}^{0}$, we can derive

$$|\mathtt{RHS}_{\mathfrak{c}}^{0\,\mathcal{J}}| \le m_0.$$

Likewise, for $1 \le i \le \ell$, we obtain $\delta' \in \mathtt{B}_i^{\mathcal{J}}$ for each $(\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}$ due to the axiom $\exists\mathtt{RHS}_{\mathfrak{c}}^{i}.\top \sqsubseteq \mathtt{B}_i$. Yet, for every such δ', at most m_i distinct corresponding δ can exist due to the axiom $\top \sqsubseteq \leqslant m_i.\mathtt{RHS}_{\mathfrak{c}}^{i}$ and hence

$$|\mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}| \le |\mathtt{B}_i^{\mathcal{J}}| \cdot m_i.$$

Moreover, as projecting will never increase the size of a set, we obtain

$$|\{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}| \le |\mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}|.$$

Yet then, combining these statements yields

$$\sum_{i=0}^{\ell} |\{\delta \mid (\delta',\delta) \in \mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}\}| \le \sum_{i=0}^{\ell} |\mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}}| \le m_0 + \sum_{i=1}^{\ell} m_i |\mathtt{B}_i^{\mathcal{J}}|$$

as claimed.

2. Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be a (finite) countable model of $\mathcal{K}$. We first give a construction for $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ and then show modelhood for $\mathcal{K}^{tr}$. Let $n_{\max} \in \mathbb{N}^{\infty}$ be the largest value obtained when evaluating all the left and right hand sides of all the constraints in $\mathcal{E}$. Then let $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}} \cup [n_{\max}]$. Note that $\Delta^{\mathcal{J}}$ is finite, whenever $\Delta^{\mathcal{I}}$ is. We let $\cdot^{\mathcal{J}}$ coincide with $\cdot^{\mathcal{I}}$ for all individual names, concept names and role names from $\mathcal{K}$. It remains to define the fresh auxiliary vocabulary of $\mathcal{K}^{tr}$. To this end, let $\mathfrak{C}$ be the set of cardinality constraints occurring in $\mathcal{E}$ which are satisfied in $\mathcal{I}$. Now, pick one δ' from $\Delta^{\mathcal{I}}$ and let

- $\mathtt{o}^{\mathcal{J}} = \delta'$,
- $\mathtt{AUX}^{\mathcal{J}} = \Delta^{\mathcal{J}} \times \{\delta'\}$,
- $\top_{\text{new}}^{\mathcal{J}} = \Delta^{\mathcal{I}}$,
- $\mathtt{Cstrt}_{\mathfrak{c}}^{\mathcal{I}} = \Delta^{\mathcal{J}}$ whenever $\mathfrak{c} \in \mathfrak{C}$ and $\mathtt{Cstrt}_{\mathfrak{c}}^{\mathcal{I}} = \emptyset$ otherwise,
- For any $\mathfrak{c}$ of the form $n_0 + n_1|\mathtt{A}_1| + \ldots + n_k|\mathtt{A}_k| \leq m_0 + m_1|\mathtt{B}_1| + \ldots + m_\ell|\mathtt{B}_\ell|$, let

$$\mathfrak{L}_{\mathfrak{c}} = \{0\}\times[n_0] \cup \{1\}\times[n_1]\times\mathtt{A}_1^{\mathcal{J}} \cup \ldots \cup \{k\}\times[n_k]\times\mathtt{A}_k^{\mathcal{J}} \text{ and}$$
$$\mathfrak{R}_{\mathfrak{c}} = \{0\}\times[m_0] \cup \{1\}\times[m_1]\times\mathtt{B}_1^{\mathcal{J}} \cup \ldots \cup \{\ell\}\times[m_\ell]\times\mathtt{B}_\ell^{\mathcal{J}}.$$

  Then we let $\cdot^{\mathfrak{c}\sharp} : \mathfrak{L}_{\mathfrak{c}} \to \big[|\mathfrak{L}_{\mathfrak{c}}|\big]$ and $\cdot^{\mathfrak{c}\flat} : \mathfrak{R}_{\mathfrak{c}} \to \big[|\mathfrak{R}_{\mathfrak{c}}|\big]$ be bijective enumeration functions for $\mathfrak{L}_{\mathfrak{c}}$ and $\mathfrak{R}_{\mathfrak{c}}$. Now we let
  - $\mathtt{LHS}_{\mathfrak{c}}^{0\,\mathcal{J}} = \{(\delta', (0,j)^{\mathfrak{c}\sharp}) \mid j \in [n_0]\}$,
  - $\mathtt{LHS}_{\mathfrak{c}}^{i\,\mathcal{J}} = \{(\delta, (i,j,\delta)^{\mathfrak{c}\sharp}) \mid \delta \in \mathtt{A}_i^{\mathcal{J}},\ j \in [n_i]\}$, for $1 \leq i \leq k$,
  - $\mathtt{RHS}_{\mathfrak{c}}^{0\,\mathcal{J}} = \{(\delta', (0,j)^{\mathfrak{c}\flat}) \mid j \in [m_0]\}$, and
  - $\mathtt{RHS}_{\mathfrak{c}}^{i\,\mathcal{J}} = \{(\delta, (i,j,\delta)^{\mathfrak{c}\flat}) \mid \delta \in \mathtt{B}_i^{\mathcal{J}},\ j \in [m_i]\}$, for $1 \leq i \leq \ell$.

It is now straightforward to check that by construction, $\mathcal{J}$ satisfies all axioms from $\mathcal{T}_{\mathcal{E}}$. As far as the relativized Abox and Tbox axioms and the unchanged Rbox axioms from $\mathcal{K}^{\natural}$ are concerned, their satisfaction follows from Lemma 3. ◠

With these model correspondences in place, we can now establish the results regarding preservation of satisfiability and query entailment as well as their complexities.

Theorem 6 (eliminability of ECboxes) *Let $\mathcal{K}$ be an EKB in some (finitely or at least countably) first-order expressible description logic $\mathcal{L}$. Then the following hold:*

1. *$\mathcal{K}$ and $\mathcal{K}^{\natural r}$ are (finitely) equisatisfiable.*
2. *Given a P2RPQ q, $\mathcal{K}$ (finitely) entails q exactly if $\mathcal{K}^{\natural r}$ (finitely) entails $q^{\natural}$.*
3. *If $\mathcal{L}$ subsumes $\mathcal{ALCIQ}_{\overline{\top}}$, then the complexities of (finite) satisfiability and (finite) CQ or P2RPQ entailment for $\mathcal{L}$ ERKBs coincide with those of plain $\mathcal{L}$ KBs.*
4. *If $\mathcal{L}$ subsumes $\mathcal{ALCOIQ}$, then the complexities of (finite) satisfiability and (finite) CQ or P2RPQ entailment for $\mathcal{L}$ EKBs coincide with those of plain $\mathcal{L}$ KBs.*

Proof.

1. On one hand, given a (finite) model of $\mathcal{K}$, Theorem 2 makes sure that we can assume it is countable and thus, Item 2 of Lemma 5 provides us with a (finite) model of $\mathcal{K}^{\natural r}$. On the other hand, given a (finite) model of $\mathcal{K}^{\natural r}$, we can invoke Item 1 of Lemma 5 to obtain a (finite) model of $\mathcal{K}$.
2. We show the equivalent statement that $\mathcal{K}$ does not (finitely) entail q exactly if $\mathcal{K}^{\natural r}$ does not (finitely) entail $q^{\natural}$. Consider a (finite) $\mathcal{I}$ with $\mathcal{I} \models \mathcal{K}$ but $\mathcal{I} \not\models q$. Theorem 2 allows us to assume that $\mathcal{I}$ is countable. Then Item 2 of Lemma 5 ensures that there is a model of $\mathcal{K}^{\natural r}$ which by construction does not satisfy $q^{\natural}$. Vice versa, consider a (finite) $\mathcal{J}$ with $\mathcal{J} \models \mathcal{K}^{\natural r}$ but $\mathcal{J} \not\models q^{\natural}$. Then Item 1 of Lemma 5 provides us with a (finite) model of $\mathcal{K}$ not satisfying q by construction.
3. This follows from the two previous items and Lemma 4, Items 1 and 2.
4. This follows from the two previous items and Lemma 4, Items 1 and 3. ◠

We note that this theorem does not only hold for CQs and P2RPQs, but it easily extends to all query formalisms where non-satisfaction can be expressed via countable first-order theories. Among others, this includes all Datalog queries [28].

7 Recapitulation: Results

Theorem 6 can now be put to use by harvesting a number of findings from known results. We will go through the results announced in Section 4 and discuss their provenance and possible further ramifications.

- Satisfiability and finite satisfiability of $\mathcal{SHIQ}b_{\top}^{\mathsf{Self}}$ ERKBs is ExpTime-complete. Noting that transitivity can be equisatisfiably removed via box-pushing (along the lines of [33,29]), yielding $\mathcal{ALCHIQ}b_{\top}^{\mathsf{Self}}$, ExpTime-completeness for the latter can be obtained via minor extensions of [33] for arbitrary models and [20] for finite models. Both also follow from the corresponding result for $\mathcal{GC}^2$, the guarded two-variable fragment as defined by Pratt-Hartmann [23]. Using these results, the application of standard techniques [11,17,29] allow to establish 2ExpTime-completeness for finite and arbitrary satisfiability of $\mathcal{SRIQ}b$ ERKBs.
- Satisfiability and finite satisfiability of $\mathcal{SHOIQB}^{\mathsf{Self}}$ EKBs is NExpTime-complete. Again, as laid out in [29], transitivity can be removed preserving satisfiability (yielding $\mathcal{ALCHOIQB}^{\mathsf{Self}}$), which is just a syntatic variant of $\mathcal{C}^2$, the two-variable fragment of first-order logic, for which the respective complexity results were established by Pratt-Hartmann [22]. Based on these findings, N2ExpTime-completeness of finite and arbitrary satisfiability of $\mathcal{SROIQB}$ EKBs is a rather direct consequence [17,29].
- P2RPQ entailment as well as finite CQ entailment from $\mathcal{ALCHIQ}b_{\top}^{\mathsf{Self}}$ ERKBs are 2ExpTime-complete. Note that $\mathcal{ALCHIQ}b_{\top}^{\mathsf{Self}}$ is a syntactic variant of $\mathcal{GC}^2$, therefore 2ExpTime-completeness of finite entailment of CQs follows from [24], while P2RPQ entailment is a consequence of [8].
- Entailment of unions of CQs from $\mathcal{ALCHOIQ}b$ EKBs is decidable and coN2ExpTime-hard. This is a consequence of the respective results for plain $\mathcal{ALCHOIQ}b$ KBs [27,14].

8 Coda: Conclusion

Inspired by previous work on quantitative extensions of $\mathcal{ALC}$ driven by Franz [1,5,2,3], we investigated the possibility of extending the expressivity of the underlying logic in the presence of global cardinality constraints. Using a novel idea of simulating the cardinality information via modeling features readily available in mainstream description logics, we were able to show that significant complexity-neutral extensions are possible. Moreover, we laid the formal foundations for adequately dealing with models of infinite domain size.

There are plenty of avenues for future work. We reiterate, that the logics considered here are tailored toward "global counting", whereas "local counting" (that is Presburger constraints over individuals' role successors) is not supported. For example, $\mathcal{ALCSCC}$ [1]

would allow us to express that some course is gender-balanced if and only if it has as many female participants as it has male ones. While this is beyond the capabilities of any of the logics considered here, the presence of inverses and nominals allows us to at least enforce that a concrete given course tcs is indeed gender-balanced, using the following Tbox and ERCbox statements:

$$\texttt{MalInC} \equiv \texttt{Male} \sqcap \exists\texttt{hasParticipant}^{-}.\{\texttt{tcs}\} \tag{12}$$

$$\texttt{FemInC} \equiv \texttt{Female} \sqcap \exists\texttt{hasParticipant}^{-}.\{\texttt{tcs}\} \tag{13}$$

$$(|\texttt{MalInC}| \leq |\texttt{FemInC}|) \wedge (|\texttt{FemInC}| \leq |\texttt{MalInC}|) \tag{14}$$

In fact, we can even go one step further and express that tcs is gender-balenced exactly if $\texttt{GendBal}(\texttt{tcs})$ holds as follows:

$$\texttt{MalInC} \equiv \texttt{Male} \sqcap \exists\texttt{hasParticipant}^{-}.\{\texttt{tcs}\} \tag{15}$$

$$\texttt{FemInC} \equiv \texttt{Female} \sqcap \exists\texttt{hasParticipant}^{-}.\{\texttt{tcs}\} \tag{16}$$

$$\texttt{BalTCS} \equiv \texttt{GendBal} \sqcap \{\texttt{tcs}\} \tag{17}$$

$$\begin{aligned}&\big(|\texttt{BalC}| {\leq} 0 \wedge (|\texttt{MalInC}| {+} 1 {\leq} |\texttt{FemInC}| \vee |\texttt{FemInC}| {+} 1 {\leq} |\texttt{MalInC}|)\big)\\ &\quad \vee \big(1 {\leq} |\texttt{BalC}|) \wedge (|\texttt{MalInC}| {\leq} |\texttt{FemInC}| \wedge |\texttt{FemInC}| {\leq} |\texttt{MalInC}|)\big)\end{aligned} \tag{18}$$

A more thorough investigation about which local counting features can be realized by global ones using advanced description logic modeling features is clearly an interesting starting point for future work.

On another note, in the case of reasoning with arbitrary models, it would be very handy from a modeler's perspective to have a way of expressing that a concept may have only finitely many elements. As mentioned before, with such statements, we leave the realms of first-order logic for good. However, for instance, an inspection of Pratt-Hartmann's work on the two-variable fragment of first-order logic with counting strongly suggests that such "finiteness constraints" can be accommodated at no additional complexity cost [22,25].

Finally, the reduction presented in this paper could potentially turn out to be of practical value, since it allows to express elaborate quantitative information by means of standardized ontology languages, which are supported by existing, highly optimized reasoning engines [13,30,32]. This having said, this proposal would only work for reasoning under the classical (i.e., arbitrary-model) semantics and, admittedly, it is also rather questionable if existing reasoners would cope well with large values in qualified number restrictions. Yet, conversely, this work might motivate developers of reasoning engines to come up with better implementations as to support statistical and other quantitative modeling.

Acknowledgements

This work (and, in particular, the conception of ERCboxes) was supported by the European Research Council (ERC) Consolidator Grant 771779 *A Grand Unified Theory of Decidability in Logic-Based Knowledge Representation* (DeciGUT). The author would also like to thank the two anonymous reviewers as well as Andreas Ecke and Bartosz Bednarczyk for their comments.

References

1. Baader, F.: A new description logic with set constraints and cardinality constraints on role successors. In: Dixon, C., Finger, M. (eds.) Proc. 11th Int. Symposium on Frontiers of Combining Systems (FroCoS'17). LNCS, vol. 10483, pp. 43–59. Springer (2017)
2. Baader, F.: Expressive cardinality constraints on $\mathcal{ALCSCC}$ concepts. In: Proc. 34th ACM/SIGAPP Symposium On Applied Computing (SAC'19). ACM (2019)
3. Baader, F., Bednarczyk, B., Rudolph, S.: Satisfiability checking and conjunctive query answering in description logics with global and local cardinality constraints. In: Proc. 32nd Int. Workshop on Description Logics (DL'19). CEUR Workshop Proceedings, CEUR-WS.org (2019), submitted
4. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, second edn. (2007)
5. Baader, F., Ecke, A.: Extending the description logic $\mathcal{ALC}$ with more expressive cardinality constraints on concepts. In: 3rd Global Conf. on Artificial Intelligence (GCAI'17). EPiC Series in Computing, vol. 50, pp. 6–19. EasyChair (2017)
6. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
7. Bach, J.S.: Canon circularis per tonos. In: The Musical Offering, BWV, vol. 1079. Leipzig (1747)
8. Bednarczyk, B., Rudolph, S.: Worst-case optimal querying of very expressive description logics with path expressions and succinct counting. In: Proc. 32nd Int. Workshop on Description Logics (DL'19). CEUR Workshop Proceedings, CEUR-WS.org (2019), submitted
9. Cantor, G.: Beiträge zur Begründung der transfiniten Mengenlehre. Mathematische Annalen **46**(4), 481–512 (1895)
10. Cantor, G.: Beiträge zur Begründung der transfiniten Mengenlehre. Mathematische Annalen **49**(2), 207–246 (1897)
11. Demri, S., Nivelle, H.: Deciding regular grammar logics with converse through first-order logic. J. of Logic, Language and Information **14**(3), 289–329 (2005)
12. Demri, S., Lugiez, D.: Complexity of modal logics with presburger constraints. Journal of Applied Logic **8**(3), 233–252 (2010)
13. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z.: HermiT: an OWL 2 reasoner. Journal of Automated Reasoning **53**(3), 245–269 (2014)
14. Glimm, B., Kazakov, Y., Lutz, C.: Status QIO: an update. In: Rosati, R., Rudolph, S., Zakharyaschev, M. (eds.) Proc. 24th Int. Workshop on Description Logics (DL'11). CEUR Workshop Proceedings, vol. 745. CEUR-WS.org (2011)
15. Glimm, B., Rudolph, S., Völker, J.: Integrated metamodeling and diagnosis in OWL 2. In: Patel-Schneider, P.F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J.Z., Horrocks, I., Glimm, B. (eds.) Proc. 9th Int. Semantic Web Conf. (ISWC'10), Part I. LNCS, vol. 6496, pp. 257–272. Springer (2010)
16. Horrocks, I., Kutz, O., Sattler, U.: The even more irresistible $\mathcal{SROIQ}$. In: Doherty, P., Mylopoulos, J., Welty, C.A. (eds.) Proc. 10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'06). pp. 57–67. AAAI Press (2006)
17. Kazakov, Y.: $\mathcal{RIQ}$ and $\mathcal{SROIQ}$ are harder than $\mathcal{SHOIQ}$. In: Brewka, G., Lang, J. (eds.) Proc. 11th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'08). pp. 274–284. AAAI Press (2008)
18. Krötzsch, M., Rudolph, S.: Nominal schemas in description logics: Complexities clarified. In: Baral, C., De Giacomo, G., Eiter, T. (eds.) Proc. 14th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'14). pp. 308–317. AAAI Press (2014)

19. Kuncak, V., Rinard, M.C.: Towards efficient satisfiability checking for Boolean algebra with Presburger arithmetic. In: Pfenning, F. (ed.) Proc. of the 21st Int. Conf. on Automated Deduction (CADE'07). LNCS, vol. 4603, pp. 215–230. Springer (2007)
20. Lutz, C., Sattler, U., Tendera, L.: The complexity of finite model reasoning in description logics. Inf. Comput. **199**(1-2), 132–171 (2005)
21. Peñaloza, R., Potyka, N.: Towards statistical reasoning in description logics over finite domains. In: Moral, S., Pivert, O. (eds.) Proc. of the 11th Int. Conf. on Scalable Uncertainty Management (SUM'17). LNCS, vol. 10564. Springer (2017)
22. Pratt-Hartmann, I.: Complexity of the two-variable fragment with counting quantifiers. J. of Logic, Language and Information **14**, 369–395 (2005)
23. Pratt-Hartmann, I.: Complexity of the guarded two-variable fragment with counting quantifiers. J. Log. Comput. **17**(1), 133–155 (2007)
24. Pratt-Hartmann, I.: Data-complexity of the two-variable fragment with counting quantifiers. Inf. Comput. **207**(8), 867–888 (2009)
25. Pratt-Hartmann, I.: Personal communication. March 19, 2019 (2019)
26. Rudolph, S.: Foundations of description logics. In: Polleres, A., d'Amato, C., Arenas, M., Handschuh, S., Kroner, P., Ossowski, S., Patel-Schneider, P.F. (eds.) Reasoning Web. Semantic Technologies for the Web of Data – 7th International Summer School 2011, LNCS, vol. 6848, pp. 76–136. Springer (2011)
27. Rudolph, S., Glimm, B.: Nominals, inverses, counting, and conjunctive queries or: Why infinity is your friend! J. of Artificial Intelligence Research **39**, 429–481 (2010)
28. Rudolph, S., Krötzsch, M.: Flag & check: Data access with monadically defined queries. In: Hull, R., Fan, W. (eds.) Proc. 32nd Symposium on Principles of Database Systems (PODS'13). pp. 151–162. ACM (2013)
29. Rudolph, S., Krötzsch, M., Hitzler, P.: Cheap Boolean role constructors for description logics. In: Hölldobler, S., Lutz, C., Wansing, H. (eds.) Proc. 11th European Conf. on Logics in Artificial Intelligence (JELIA'08). LNAI, vol. 5293, pp. 362–374. Springer (2008)
30. Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y.: Pellet: A practical OWL-DL reasoner. J. of Web Semantics **5**(2), 51–53 (2007)
31. Skolem, T.: Über einige Grundlagenfragen der Mathematik. Skrifter utgitt av det Norske Videnskaps-Akademi i Oslo, I. Matematisk-naturvidenskabelig Klasse **7**, 1–49 (1929)
32. Steigmiller, A., Liebig, T., Glimm, B.: Konclude: System description. Journal of Web Semantics **27**, 78–85 (2014)
33. Tobies, S.: Complexity Results and Practical Algorithms for Logics in Knowledge Representation. Ph.D. thesis, RWTH Aachen, Germany (2001)