Towards Elimination of Second-Order Quantifiers in the Separated Fragment

Marco Voigt✉

Max Planck Institute for Informatics and Saarbrücken Graduate School of Computer Science, Saarland Informatics Campus, Saarbrücken, Germany
mvoigt@mpi-inf.mpg.de

Abstract. It is a classical result that the monadic fragment of second-order logic admits elimination of second-order quantifiers. Recently, the separated fragment (SF) of first-order logic has been introduced. SF generalizes the monadic first-order fragment without equality, while preserving decidability of the satisfiability problem. Therefore, it is a natural question to ask whether SF also admits elimination of second-order quantifiers. Interestingly, already Ackermann answered this question in the negative as far as full SF with unrestricted occurrences of second-order quantifiers is concerned. However, with appropriate restrictions on the syntax of a second-order version of SF, one could hope to define a substantial extension of the monadic fragment that admits second-order quantifier elimination. The present note is about preliminary results of ongoing research in this direction. As a first positive result a restricted second-order version of SF is defined that admits the elimination of at least one existential second-order quantifier. The elimination of existential second-order quantifiers from a monadic sentence without equality constitutes a special case of the methods presented here.

Keywords: Second-order quantifier elimination · separated fragment · monadic fragment

1 Introduction

It is a classical result that the monadic fragment of second-order logic admits elimination of second-order quantifiers. This was discovered by Löwenheim [6], Skolem [7], and Behmann [2].

Recently, the *separated fragment (SF)* of first-order logic has been introduced [8]. It constitutes a syntactic generalization of well-known first-order fragments: the Bernays–Schönfinkel–Ramsey fragment—the class of relational $\exists^*\forall^*$ sentences—and the monadic first-order fragment without equality—the class of relational sentences over predicate symbols of arity at most one. The satisfiability problem for SF sentences (*SF-Sat*) is decidable, but computationally very hard: SF-Sat is k-NExpTime-hard for every positive integer k [10]. In other words, SF-Sat is non-elementary. The definition of SF is based on restricting the syntax of first-order sentences in prenex normal form. However, neither the

In: P. Koopmann, S. Rudolph, R. Schmidt, C. Wernhard (eds.): *SOQE 2017 – Proceedings of the Workshop on Second-Order Quantifier Elimination and Related Topics, Dresden, Germany, December 6–8, 2017*, published at http://ceur-ws.org.

arity of predicate symbols nor the shape of quantifier prefixes is restricted. The defining principle for SF sentences is that universally and existentially quantified variables do not occur together in atoms. Leading existential quantifiers are exempt from this rule. The sentence $\forall x_1 \exists y_1 \forall x_2 \exists y_2 .\, R(x_1, x_2) \leftrightarrow Q(y_1, y_2)$ is an exemplary SF sentence.

As SF generalizes the monadic first-order fragment without equality while retaining a decidable satisfiability problem, it is natural to ask whether a second-order version of SF admits elimination of second-order quantifiers. Interestingly, already Ackermann gave a negative answer to this question. In an article from 1935 [1], Ackermann argued that the quantifier $\exists P$ in the following formula cannot be eliminated: $\exists P.\ P(x) \wedge \neg P(y) \wedge \forall uv.\ \neg P(u) \vee P(v) \vee \neg N(u, v)$. The only atom in this formula that could potentially break the separateness condition is $N(u, v)$. But since both variables u and v are universally quantified, universal variables are separated from existential variables and the sentence is in SF.

Although Ackermann's observation seems to be discouraging, it only means that there is, apparently, no straight-forward way of extending the quantifier-elimination techniques that work for second-order monadic logic to the separated fragment. The purpose of this note is to present certain syntactic restrictions that allow the elimination of existentially quantified unary predicate symbols in separated formulas. The presented results are of a preliminary character and are not yet fully developed. They provide only a first hint at some directions that might be worth following in future work.

In Section 2 we present the used notation and some basic results. A definition of the separated fragment is given in Section 3. The main result is developed in Section 4 and concisely formulated in Theorem 8. Finally, Section 5 concludes with a discussion of the results and future directions.

2 Notation and Preliminaries

We consider second-order logic formulas with equality. We call a formula *relational* if it contains neither function nor constant symbols. In all formulas, if not explicitly stated otherwise, we tacitly assume that no variable occurs freely and bound at the same time and that no variable is bound by two different occurrences of quantifiers. For convenience, we sometimes identify tuples $\bar{\mathbf{x}}$ of variables with the set containing all the variables that occur in $\bar{\mathbf{x}}$. By $\mathrm{vars}(\varphi)$ we denote the set of all variables occurring in φ.

The symbol $\models$ denotes the *is-a-model-of* relation as well as semantic entailment of formulas, i.e. $\varphi \models \psi$ holds whenever for every structure $\mathcal{A}$ and every variable assignment β, $\mathcal{A}, \beta \models \varphi$ entails $\mathcal{A}, \beta \models \psi$. The symbol $\models\!\!\!\mid$ denotes *semantic equivalence* of formulas, i.e. $\varphi \models\!\!\!\mid \psi$ holds whenever $\varphi \models \psi$ and $\psi \models \varphi$.

The following are standard lemmas that we simply add for completeness.

Lemma 1 (Miniscoping). *Let φ, ψ, χ be formulas, and assume that x and y do not occur freely in χ. We have the following equivalences, where $\circ \in \{\wedge, \vee\}$:*

(i) $\exists y.(\varphi \vee \psi) \models\!\!\!\mid (\exists y.\varphi) \vee (\exists y.\psi)$		*(ii)* $\forall x.(\varphi \wedge \psi) \models\!\!\!\mid (\forall x.\varphi) \wedge (\forall x.\psi)$	
(iii) $\exists y.(\varphi \circ \chi) \models\!\!\!\mid (\exists y.\varphi) \circ \chi$		*(iv)* $\forall x.(\varphi \circ \chi) \models\!\!\!\mid (\forall x.\varphi) \circ \chi$	

Lemma 2. *Let $\psi[t]$ be some second-order formula in which the term t occurs. Let x be some first-order variable that does not occur in $\psi[t]$. Then, $\psi[t]$ is semantically equivalent to $\forall x.\, x = t \rightarrow \psi[x]$, where $\psi[x]$ is derived from $\psi[t]$ by replacing every occurrence of t with the variable x.*

3 The Separated Fragment

Consider a second-order formula φ. We say that two disjoint sets of first-order variables X and Y are *separated in* φ if and only if for every atom A in φ we have $\text{vars}(A) \cap X = \emptyset$ or $\text{vars}(A) \cap Y = \emptyset$.

The following definition of the separated fragment is a slightly simplified version of the fragment investigated in [8] and in [10]. In contrast to the original, we do not consider constant symbols here.

Definition 3 (Separated fragment (SF)). *The* separated fragment (SF) *of first-order logic consists of all relational first-order sentences with equality that are of the form $\exists \bar{\mathbf{z}}\, \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n \exists \bar{\mathbf{y}}_n.\, \psi$, in which ψ is quantifier free, and in which the two sets $\bar{\mathbf{x}}_1 \cup \ldots \cup \bar{\mathbf{x}}_n$ and $\bar{\mathbf{y}}_1 \cup \ldots \cup \bar{\mathbf{y}}_n$ are separated. The tuples $\bar{\mathbf{z}}$ and $\bar{\mathbf{y}}_n$ may be empty, i.e. the quantifier prefix does not have to start with an existential quantifier and it does not have to end with an existential quantifier either.*

Notice that the variables in $\bar{\mathbf{z}}$ are not subject to any restriction concerning their occurrences.

It is not hard to see that SF generalizes the Bernays–Schönfinkel–Ramsey fragment (relational $\exists^*\forall^*$ prenex formulas with equality) and the monadic first-order fragment without equality (see [8], Theorem 9). The reason why certain monadic sentences with equality do not belong to SF is that, although SF sentences may contain equality, non-separated equations are not allowed in SF. For example, the sentence $\exists y \forall x.\, x = y$ belongs to SF whereas $\forall x \exists y.\, x = y$ does not.

As already mentioned in the introduction, it is known that the satisfiability problem for SF sentences (*SF-Sat*) is decidable and non-elementary [8, 10]. These results rely on an equivalence-preserving transformation from SF into the Bernays–Schönfinkel–Ramsey fragment (BSR): for every SF sentence there is an equivalent sentence in the BSR fragment. This transformation will be the starting point for showing that quantifier elimination is possible for a certain extension of the separated fragment with second-order quantifiers.

Lemma 4. *Let $\varphi := \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n \exists \bar{\mathbf{y}}_n.\, \psi(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{z}})$ be a relational first-order formula in which ψ is quantifier free and the sets $\bar{\mathbf{x}} := \bar{\mathbf{x}}_1 \cup \ldots \cup \bar{\mathbf{x}}_n$ and $\bar{\mathbf{y}} := \bar{\mathbf{y}}_1 \cup \ldots \cup \bar{\mathbf{y}}_n$ are separated. Moreover, we assume that every variable occurring in the quantifier prefix and in $\bar{\mathbf{z}}$ also occurs in the matrix ψ.*

Let $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1} \subseteq \bar{\mathbf{x}}$ and $\widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2} \subseteq \bar{\mathbf{y}}$ be partitions of the sets $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, respectively, such that the $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1}, \widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2}$ are nonempty, pairwise disjoint, and pairwise separated in φ. Then, φ is equivalent to a finite disjunction of formulas of the form

$$\Big(\bigwedge_k \forall \widetilde{\mathbf{x}}''_k.\ \bigvee_\ell K_{k\ell}(\widetilde{\mathbf{x}}''_k, \bar{\mathbf{z}})\Big) \ \wedge\ \Big(\bigwedge_i \exists \widetilde{\mathbf{y}}''_i.\ \bigwedge_j L_{ij}(\widetilde{\mathbf{y}}''_i, \bar{\mathbf{z}})\Big)\ ,$$

where the $K_{k\ell}$ and the L_{ij} are literals whose atoms are renamed variants of atoms that occur in φ. Moreover, any two sets $\widetilde{\mathbf{x}}''_{k_1}, \widetilde{\mathbf{x}}''_{k_2}$ with $k_1 \neq k_2$, $\widetilde{\mathbf{y}}''_{i_1}, \widetilde{\mathbf{y}}''_{i_2}$ with $i_1 \neq i_2$, and $\widetilde{\mathbf{x}}''_k, \widetilde{\mathbf{y}}''_i$ are separated in the resulting formula.

Proof. The proof is an adaptation of the proof of Lemma 12 in [10]. For convenience, we pretend that $\bar{\mathbf{z}}$ is empty. The argument works for nonempty $\bar{\mathbf{z}}$ as well. We will make use of the following auxiliary lemma:

Claim I (Lemma 11 in [10]):
Let I and J_i, $i \in I$, be sets that are finite, nonempty, and pairwise disjoint. The elements of these sets serve as indices. Let

$$\exists \bar{\mathbf{v}}. \bigwedge_{i \in I} \Big(\chi_i(\bar{\mathbf{u}}) \vee \bigvee_{k \in J_i} \eta_k(\bar{\mathbf{v}}, \bar{\mathbf{u}})\Big)$$

be some first-order formula where the χ_i and the η_k denote arbitrary subformulas that we treat as indivisible units in what follows. We say that $f : I \to \bigcup_{i \in I} J_i$ is a *selector* if for every $i \in I$ we have $f(i) \in J_i$. We denote the set of all selectors of this form by $\mathcal{F}$.
Then, the above formula is equivalent to

$$\bigwedge_{\substack{S \subseteq I \\ S \neq \emptyset}} \Big(\bigvee_{i \in S} \chi_i(\bar{\mathbf{u}})\Big) \vee \bigvee_{f \in \mathcal{F}} \Big(\exists \bar{\mathbf{v}}. \bigwedge_{i \in S} \eta_{f(i)}(\bar{\mathbf{v}}, \bar{\mathbf{u}})\Big)\ .$$

$\Diamond$

We transform φ into an equivalent CNF formula of the form

$$\forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n \exists \bar{\mathbf{y}}_n. \bigwedge_{i \in I} \Big(\chi_i(\bar{\mathbf{x}}) \vee \bigvee_{k \in J_i} L_k(\bar{\mathbf{y}})\Big)$$

where I and the J_i are finite, pairwise disjoint sets of indices, the subformulas χ_i are disjunctions of literals, and the L_k are literals. By Claim I, we can construct an equivalent formula of the form

$$\varphi' := \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n. \bigwedge_{\substack{S \subseteq I \\ S \neq \emptyset}} \Big(\bigvee_{i \in S} \chi_i(\bar{\mathbf{x}})\Big) \vee \bigvee_{f \in \mathcal{F}} \Big(\exists \bar{\mathbf{y}}_n. \bigwedge_{i \in S} \eta_{f(i)}(\bar{\mathbf{y}})\Big)$$

where $\mathcal{F}$ is the set of all selectors over the index sets J_i, $i \in I$. Applying miniscoping (Lemma 1), we move inward the universal quantifier block $\forall \bar{\mathbf{x}}_n$ and thus obtain

$$\varphi'' := \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \exists \bar{\mathbf{y}}_{n-1}. \bigwedge_{\substack{S \subseteq I \\ S \neq \emptyset}} \Big(\forall \bar{\mathbf{x}}_n. \bigvee_{i \in S} \chi_i(\bar{\mathbf{x}})\Big) \vee \bigvee_{f \in \mathcal{F}} \Big(\exists \bar{\mathbf{y}}_n. \bigwedge_{i \in S} \eta_{f(i)}(\bar{\mathbf{y}})\Big)\ .$$

We now iterate these two steps in an alternating fashion until all quantifier blocks have been moved inwards in the described way. The constituents of the result $\varphi^{(3)} := \bigwedge_q \Big(\chi_q^{(3)} \vee \bigvee_p \eta_{qp}^{(3)}\Big)$ of this process have the form

$$\chi_q^{(3)} = \forall \bar{\mathbf{x}}_1. \bigvee_{\ell_1} \forall \bar{\mathbf{x}}_2. \bigvee_{\ell_2} \Big(\dots \Big(\bigvee_{\ell_{n-1}} \forall \bar{\mathbf{x}}_n. \bigvee_{i \in S_{\ell_1,\dots,\ell_{n-1}}} \chi_i(\bar{\mathbf{x}}) \Big) \dots \Big)$$

where the $S_{\ell_1,\dots,\ell_n-1}$ are certain subsets of I and the χ_i are still disjunctions of literals, and

$$\eta_{qp}^{(3)} = \exists \bar{\mathbf{y}}_1. \bigwedge_{\ell_1} \exists \bar{\mathbf{y}}_2. \bigwedge_{\ell_2} \Big(\dots \Big(\bigwedge_{\ell_{n-1}} \exists \bar{\mathbf{y}}_n. \bigwedge_{k \in J_{\ell_1,\dots,\ell_{n-1}}} L_k(\bar{\mathbf{y}}) \Big) \dots \Big)$$

where the $J_{\ell_1,\dots,\ell_{n-1}}$ are certain subsets of $\bigcup_{i \in I} J_i$.

By definition of the sets $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_{m_1}$, which are pairwise separated in the $\chi_q^{(3)}$, we can rewrite every $\chi_q^{(3)}$ into the following form by regrouping the inner disjuncts:

$$\chi_q^{(4)} = \forall \bar{\mathbf{x}}_1. \bigvee_{\ell_1} \forall \bar{\mathbf{x}}_2. \bigvee_{\ell_2} \Big(\dots \Big(\bigvee_{\ell_{n-1}} \forall \bar{\mathbf{x}}_n. \bigvee_{i'=1,\dots,m_1} \chi'_{\bar{\ell} i'}(\widetilde{\mathbf{x}}_{i'}) \Big) \dots \Big)$$

where the $\chi'_{\bar{\ell} i'}$ are (possibly empty) disjunctions of literals. Analogously, we rewrite every $\eta_{qp}^{(3)}$ into the form

$$\eta_{qp}^{(4)} = \exists \bar{\mathbf{y}}_1. \bigwedge_{\ell_1} \exists \bar{\mathbf{y}}_2. \bigwedge_{\ell_2} \Big(\dots \Big(\bigwedge_{\ell_{n-1}} \exists \bar{\mathbf{y}}_n. \bigwedge_{j'=1,\dots,m_2} \eta'_{\bar{\ell} j'}(\widetilde{\mathbf{y}}_j) \Big) \dots \Big)$$

where the $\eta'_{\bar{\ell} j'}$ are (possibly empty) conjunctions of literals.

We then observe the following equivalences, starting from $\chi_q^{(4)}$:

$$\begin{aligned}
&\forall \bar{\mathbf{x}}_1. \bigvee_{\ell_1} \forall \bar{\mathbf{x}}_2. \bigvee_{\ell_2} \Big(\dots \Big(\bigvee_{\ell_{n-1}} \forall \bar{\mathbf{x}}_n. \bigvee_{i'=1,\dots,m_1} \chi'_{\bar{\ell} i'}(\widetilde{\mathbf{x}}_{i'}) \Big) \dots \Big) \\
&\models\!\!\!\!\!\!\;=\; \forall \bar{\mathbf{x}}_1. \bigvee_{\ell_1} \forall \bar{\mathbf{x}}_2. \bigvee_{\ell_2} \Big(\dots \Big(\bigvee_{\ell_{n-1}} \bigvee_{i'=1,\dots,m_1} \forall (\bar{\mathbf{x}}_n \cap \widetilde{\mathbf{x}}_{i'}).\ \chi'_{\bar{\ell} i'}(\widetilde{\mathbf{x}}_{i'}) \Big) \dots \Big) \\
&\models\!\!\!\!\!\!\;=\; \forall \bar{\mathbf{x}}_1. \bigvee_{\ell_1} \forall \bar{\mathbf{x}}_2. \bigvee_{\ell_2} \Big(\dots \Big(\bigvee_{i'=1,\dots,m_1} \bigvee_{\ell'_{n-1}} \forall (\bar{\mathbf{x}}_n \cap \widetilde{\mathbf{x}}_{i'}).\ \chi'_{\bar{\ell}' i'}(\widetilde{\mathbf{x}}_{i'}) \Big) \dots \Big) \\
&\qquad \vdots \\
&\models\!\!\!\!\!\!\;=\; \bigvee_{i'=1,\dots,m_1} \forall (\bar{\mathbf{x}}_1 \cap \widetilde{\mathbf{x}}_{i'}). \bigvee_{\ell'_1} \forall (\bar{\mathbf{x}}_2 \cap \widetilde{\mathbf{x}}_{i'}). \bigvee_{\ell'_2} \Big(\dots \Big(\bigvee_{\ell'_{n-1}} \forall (\bar{\mathbf{x}}_n \cap \widetilde{\mathbf{x}}_{i'}).\ \chi'_{\bar{\ell}' i'}(\widetilde{\mathbf{x}}_{i'}) \Big) \dots \Big) \\
&\models\!\!\!\!\!\!\;=\; \bigvee_{i'=1,\dots,m_1} \forall \widetilde{\mathbf{x}}'_{i'}.\ \chi''_{i'}(\widetilde{\mathbf{x}}'_{i'}) \ ,
\end{aligned}$$

where the $\chi''_{i'}$ are disjunctions of literals. Before moving universal quantifiers outwards in the last step of the above transformation, bound variables are renamed such that all quantifiers bind pairwise distinct variables. Analogously, we have

$$\eta^{(4)}_{qp} \models\mid \bigwedge_{j'=1,\ldots,m_2} \exists \widetilde{\mathbf{y}}'_{j'}.\, \eta''_{j'}(\widetilde{\mathbf{y}}'_{i'}) \ ,$$

where the $\eta''_{j'}$ are conjunctions of literals.

Consequently, we have rewritten $\varphi^{(3)} = \bigwedge_q \Big(\chi^{(3)}_q \vee \bigvee_p \eta^{(3)}_{qp} \Big)$ into an equivalent formula $\varphi^{(4)}$ of the form

$$\varphi^{(4)} = \bigwedge_q \Big(\Big(\bigvee_{i'=1,\ldots,m_1} \forall \widetilde{\mathbf{x}}'_{i'}.\, \chi''_{qi'}(\widetilde{\mathbf{x}}'_{i'}) \Big) \vee \Big(\bigvee_p \bigwedge_{j'=1,\ldots,m_2} \exists \widetilde{\mathbf{y}}'_{j'}.\, \eta''_{qpj'}(\widetilde{\mathbf{y}}'_{i'}) \Big) \Big) \ .$$

After renaming bound variables again such that all quantifiers bind pairwise distinct variables, we transform $\varphi^{(4)}$ into an equivalent formula that is a disjunction of formulas of the form

$$\bigwedge_k \Big(\forall \widetilde{\mathbf{x}}''_k.\, \bigvee_\ell K_{k\ell}(\widetilde{\mathbf{x}}''_k) \Big) \ \wedge \ \bigwedge_i \Big(\exists \widetilde{\mathbf{y}}''_i.\, \bigwedge_j L_{ij}(\widetilde{\mathbf{y}}''_i) \Big) \ .$$

□

The just proven lemma will provide the syntactic transformations necessary to eliminate second-order quantifiers that occur in a separated formula under certain conditions.

Example 5. We have already mentioned the following SF sentence in the introduction: $\varphi := \forall x_1 \exists y_1 \forall x_2 \exists y_2.\, R(x_1, x_2) \leftrightarrow Q(y_1, y_2)$. As indicated by Lemma 4, nested alternating quantifiers can be transformed away. An intermediate result of this process is

$$\begin{aligned} \forall x_1 \exists y_1.\, \Big(\big(\forall x_2.\, R(x_1, x_2) \big) \vee \big(\exists y_2.\, \neg Q(y_1, y_2) \big) \Big) \\ \wedge \Big(\big(\forall x_2.\, \neg R(x_1, x_2) \big) \vee \big(\exists y_2.\, Q(y_1, y_2) \big) \Big) \ . \end{aligned}$$

Continuing the transformation process, we eventually obtain

$$\begin{aligned} & \Big(\exists y_1 y_2 y_3.\, Q(y_1, y_2) \wedge \neg Q(y_1, y_3) \Big) \\ & \vee \Big(\big(\forall x_1 x_2.\, R(x_1, x_2) \big) \wedge \big(\exists y_1 y_2.\, Q(y_1, y_2) \big) \Big) \\ & \vee \Big(\big(\forall x_1 x_2.\, \neg R(x_1, x_2) \big) \wedge \big(\exists y_1 y_2.\, \neg Q(y_1, y_2) \big) \Big) \\ & \vee \Big(\big(\forall x_1 x_2 x_3.\, R(x_1, x_2) \vee \neg R(x_1, x_3) \big) \wedge \big(\exists y_1 y_2.\, Q(y_1, y_2) \big) \wedge \big(\exists y_3 y_4.\, \neg Q(y_3, y_4) \big) \Big) \ , \end{aligned}$$

which is equivalent to φ but does not contain any quantifier alternation.

4 Elimination of Second-Order Quantifiers

In this section we formulate syntactic restrictions that enable the elimination of second-order quantifiers over unary predicates from sentences that belong to the separated fragment. The notion of separation of sets of variables in a formula plays a central role in our criterion. However, this time it is not only of interest that universal variables are separated from existential variables. It is rather of importance that within each set of non-separated variables there is at most one that occurs as the argument of the predicate symbol that is bound by the quantifier we intend to eliminate.

Lemma 6. *Let* $\varphi := \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n \exists \bar{\mathbf{y}}_n . \psi(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{z}})$ *be a relational first-order formula in which* ψ *is quantifier free and the sets* $\bar{\mathbf{x}} := \bar{\mathbf{x}}_1 \cup \ldots \cup \bar{\mathbf{x}}_n$ *and* $\bar{\mathbf{y}} := \bar{\mathbf{y}}_1 \cup \ldots \cup \bar{\mathbf{y}}_n$ *are separated. We assume that every variable occurring in the quantifier prefix and in* $\bar{\mathbf{z}}$ *also occurs in the matrix* ψ*.*

Let $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1} \subseteq \bar{\mathbf{x}}$ *and* $\widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2} \subseteq \bar{\mathbf{y}}$ *be partitions of the sets* $\bar{\mathbf{x}}$ *and* $\bar{\mathbf{y}}$*, respectively, such that the* $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1}, \widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2}$ *are nonempty, pairwise disjoint, and pairwise separated in* φ*. Let* P *be a unary predicate symbol satisfying the following conditions:*

(1) For every set $\widetilde{\mathbf{x}}_i$*,* $1 \leq i \leq m_1$*, there is at most one variable* $x_i^* \in \widetilde{\mathbf{x}}_i$ *for which* φ *contains atoms* $P(x_i^*)$*.*
(2) For every set $\widetilde{\mathbf{y}}_i$*,* $1 \leq i \leq m_2$*, there is at most one variable* $y_i^* \in \widetilde{\mathbf{y}}_i$ *for which* φ *contains atoms* $P(y_i^*)$*.*

Then $\exists P. \varphi$ *is equivalent to a finite disjunction of formulas of the form*

$$\begin{aligned}
\theta(\bar{\mathbf{z}}) \;\wedge\; \exists P. \; & \bigwedge_{k_1} \Big(\forall \widetilde{\mathbf{x}}'_{k_1} . \, \chi_{k_1}(\widetilde{\mathbf{x}}'_{k_1}, \bar{\mathbf{z}}) \vee P(x^*_{k_1}) \Big) \wedge \bigwedge_{k_2} \Big(\forall \widetilde{\mathbf{x}}'_{k_2} . \, \chi'_{k_2}(\widetilde{\mathbf{x}}'_{k_2}, \bar{\mathbf{z}}) \vee \neg P(x^*_{k_2}) \Big) \\
& \wedge \bigwedge_{i_1} \Big(\exists \widetilde{\mathbf{y}}'_{i_1} . \, \eta_{i_1}(\widetilde{\mathbf{y}}'_{i_1}, \bar{\mathbf{z}}) \wedge P(y^*_{i_1}) \Big) \wedge \bigwedge_{i_2} \Big(\exists \widetilde{\mathbf{y}}'_{i_2} . \, \eta'_{i_2}(\widetilde{\mathbf{y}}'_{i_2}, \bar{\mathbf{z}}) \wedge \neg P(y^*_{i_2}) \Big) \\
& \wedge \bigwedge_{\ell_1} P(z^*_{\ell_1}) \wedge \bigwedge_{\ell_2} \neg P(z^*_{\ell_2}) \;,
\end{aligned}$$

where (a) the χ_{k_1} *and the* χ'_{k_2} *are disjunctions of literals and the* η_{i_1} *and the* η'_{i_2} *are conjunctions of literals, (b) all the atoms in* θ *and in the* χ_{k_1}*,* χ'_{k_2}*,* η_{i_1}*, and* η_{i_2} *are renamed variants of atoms that occur in* φ *and do not contain the predicate symbol* P*, and (c) the variables* $z^*_{\ell_1}$*,* $z^*_{\ell_2}$ *are pairwise distinct and stem from* $\bar{\mathbf{z}}$*.*

Proof. By Lemma 4, we know that φ can be rewritten to an equivalent formula that is a finite disjunction of formulas in which no universal quantifier lies within the scope of an existential quantifier and vice versa. We apply this transformation to φ and obtain a formula as described in Lemma 4. In the next step, we isolate atoms that exclusively contain variables from $\bar{\mathbf{z}}$, narrow the scopes of first-order quantifiers so that these atoms are not within their scopes anymore, and transform the resulting formulas into a formula φ' that is a disjunction of formulas of the form

$$\Big(\bigwedge_k \forall\widetilde{\mathbf{x}}'_k.\ \bigvee_\ell K_{k\ell}(\widetilde{\mathbf{x}}'_k, \bar{\mathbf{z}})\Big) \ \wedge\ \Big(\bigwedge_i \exists\widetilde{\mathbf{y}}'_i.\ \bigwedge_j L_{ij}(\widetilde{\mathbf{y}}'_i, \bar{\mathbf{z}})\Big) \ \wedge\ \bigwedge_r M_r(\bar{\mathbf{z}})\ ,$$

where the $K_{k\ell}$ and the L_{ij} are literals whose atoms are renamed variants of atoms that occur in φ and contain at least one variable from some $\widetilde{\mathbf{x}}'_k$ or $\widetilde{\mathbf{y}}'_i$. The M_r are literals whose atoms occur in φ and contain exclusively variables from $\bar{\mathbf{z}}$. Moreover, any two sets $\widetilde{\mathbf{x}}'_{k_1}, \widetilde{\mathbf{x}}'_{k_2}$ with $k_1 \neq k_2$, $\widetilde{\mathbf{y}}'_{i_1}, \widetilde{\mathbf{y}}'_{i_2}$ with $i_1 \neq i_2$, and $\widetilde{\mathbf{x}}'_k, \widetilde{\mathbf{y}}'_i$ are separated in φ'. By inspection of the transformations performed in the proof of Lemma 4, we observe that Conditions (1) and (2) are preserved such that they also apply to the sets $\widetilde{\mathbf{x}}'_k$ and $\widetilde{\mathbf{y}}'_i$ with respect to variables x^*_k and y^*_i, respectively.

This enables us to regroup the disjunctions and conjunctions in the constituents of φ' such that each of these disjuncts has the form

$$\begin{aligned} &\bigwedge_{k'}\Big(\forall\widetilde{\mathbf{x}}'_{k'}.\Big(\bigvee_{\ell'} K_{k'\ell'}(\widetilde{\mathbf{x}}'_{k'}, \bar{\mathbf{z}})\Big) \vee [\neg]P(x^*_{k'})\Big)\\ \wedge\ &\bigwedge_{i'}\Big(\exists\widetilde{\mathbf{y}}'_{i'}.\Big(\bigwedge_{j'} L_{i'j'}(\widetilde{\mathbf{y}}'_{i'}, \bar{\mathbf{z}})\Big) \wedge [\neg]P(y^*_{i'})\Big)\\ \wedge\ &\Big(\bigwedge_{r'} M_{r'}(\bar{\mathbf{z}})\Big) \wedge \bigwedge_q [\neg]P(z^*_q)\ , \end{aligned}$$

where the literals $K_{k'\ell'}$, $L_{i'j'}$, and $M_{r'}$ do not contain the predicate symbol P. The variables z^*_q stem from $\bar{\mathbf{z}}$. Moreover, we replace disjuncts (conjuncts) which contain two literals $P(v)$ and $\neg P(v)$ with the logical constant `true` (`false`). Having this, it only remains to regroup conjuncts and distribute the existential quantifier $\exists P$ over the topmost disjunction, in order to obtain the formula advertised in the lemma. □

The formula resulting from the lemma gives us the right starting point for the elimination of the second-order quantifier $\exists P$ from a formula. Before we elaborate on this, we present the lemma that we shall employ for elimination.

Lemma 7 (Basic elimination lemma, see [11] and [2]). *Let P be a unary predicate symbol and let χ, η be first-order formulas in which P does not occur. Then, $\exists P.\ \big(\forall x.\, \chi \vee P(x)\big) \wedge \big(\forall x.\, \eta \vee \neg P(x)\big)$ is semantically equivalent to $\forall x.\, \chi \vee \eta$.*

Consider a formula $\varphi := \forall\bar{\mathbf{x}}_1\exists\bar{\mathbf{y}}_1 \ldots \forall\bar{\mathbf{x}}_n\exists\bar{\mathbf{y}}_n.\psi(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{z}})$ as described in Lemma 6. Moreover, let there be sets $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1}$ and $\widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2}$ and a unary predicate symbol P as described in the lemma. Then, Lemma 6 stipulates the existence of a formula equivalent to φ that is a disjunction of formulas of the form

$$\begin{aligned} \theta(\bar{\mathbf{z}})\ \wedge\ \exists P.\ &\bigwedge_{k_1}\Big(\forall\widetilde{\mathbf{x}}'_{k_1}.\ \chi_{k_1}(\widetilde{\mathbf{x}}'_{k_1}, \bar{\mathbf{z}}) \vee P(x^*_{k_1})\Big) \wedge \bigwedge_{k_2}\Big(\forall\widetilde{\mathbf{x}}'_{k_2}.\ \chi'_{k_2}(\widetilde{\mathbf{x}}'_{k_2}, \bar{\mathbf{z}}) \vee \neg P(x^*_{k_2})\Big)\\ \wedge &\bigwedge_{i_1}\Big(\exists\widetilde{\mathbf{y}}'_{i_1}.\ \eta_{i_1}(\widetilde{\mathbf{y}}'_{i_1}, \bar{\mathbf{z}}) \wedge P(y^*_{i_1})\Big) \wedge \bigwedge_{i_2}\Big(\exists\widetilde{\mathbf{y}}'_{i_2}.\ \eta'_{i_2}(\widetilde{\mathbf{y}}'_{i_2}, \bar{\mathbf{z}}) \wedge \neg P(y^*_{i_2})\Big)\\ \wedge &\bigwedge_{\ell_1} P(z^*_{\ell_1}) \wedge \bigwedge_{\ell_2} \neg P(z^*_{\ell_2})\ , \end{aligned}$$

in which we can eliminate the quantifier $\exists P$ as follows. The shape of the above formula is very similar to what Behmann called "Eliminationshauptform" in [2] (see [11] for a modern exposition of Behmann's results related to quantifier elimination). With the next two transformation steps we come closer to the syntactic shape of the "Eliminationshauptform". First, we narrow the scope of the first-order quantifiers that do not bind variables x_k^* or y_i^*.

$$\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists P.\ & \bigwedge_{k_1}\Big(\forall x_{k_1}^*.\ \underbrace{\big(\forall(\widetilde{\mathbf{x}}'_{k_1}\setminus\{x_{k_1}^*\}).\ \chi_{k_1}(\widetilde{\mathbf{x}}'_{k_1},\bar{\mathbf{z}})\big)}_{=:\ \chi_{k_1}^*} \vee P(x_{k_1}^*)\Big) \\
\wedge \ & \bigwedge_{k_2}\Big(\forall x_{k_2}^*.\ \underbrace{\big(\forall(\widetilde{\mathbf{x}}'_{k_2}\setminus\{x_{k_2}^*\}).\ \chi_{k_2}(\widetilde{\mathbf{x}}'_{k_2},\bar{\mathbf{z}})\big)}_{=:\ \chi_{k_2}^*} \vee \neg P(x_{k_2}^*)\Big) \\
\wedge \ & \bigwedge_{i_1}\Big(\exists y_{i_1}^*.\ \underbrace{\big(\exists(\widetilde{\mathbf{y}}'_{i_1}\setminus\{y_{i_1}^*\}).\ \eta'_{i_1}(\widetilde{\mathbf{y}}'_{i_1},\bar{\mathbf{z}})\big)}_{=:\ \eta_{i_1}^*} \wedge P(y_{i_1}^*)\Big) \\
\wedge \ & \bigwedge_{i_2}\Big(\exists y_{i_2}^*.\ \underbrace{\big(\exists(\widetilde{\mathbf{y}}'_{i_2}\setminus\{y_{i_2}^*\}).\ \eta'_{i_2}(\widetilde{\mathbf{y}}'_{i_2},\bar{\mathbf{z}})\big)}_{=:\ \eta_{i_2}^*} \wedge \neg P(y_{i_2}^*)\Big) \\
\wedge & \bigwedge_{\ell_1} P(z_{\ell_1}^*) \wedge \bigwedge_{\ell_2} \neg P(z_{\ell_2}^*)
\end{aligned}$$

Next, we treat the subformulas χ_k^* and η_i^* as indivisible units, move universal quantifiers outwards that occur in different conjuncts (and merge them while doing so), pull first-order existential quantifiers outwards (without merging them), and rename the variables that are bound by the moved quantifiers. Moreover, we reorder the conjunctions in the scope of the quantifier blocks $\exists\bar{\mathbf{u}}$ and $\exists\bar{\mathbf{v}}$.

$$\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists P.\ & \Big(\forall x.\ \underbrace{\Big(\bigwedge_{k_1}\chi_{k_1}^*\big[x_{k_1}^*/x\big]\Big)}_{=:\ \chi_1^*(x,\bar{\mathbf{z}})} \vee P(x)\Big) \\
\wedge \ & \Big(\forall x.\ \underbrace{\Big(\bigwedge_{k_2}\chi_{k_2}^*\big[x_{k_2}^*/x\big]\Big)}_{=:\ \chi_2^*(x,\bar{\mathbf{z}})} \vee \neg P(x)\Big) \\
\wedge \ & \Big(\exists\bar{\mathbf{u}}.\ \underbrace{\Big(\bigwedge_{i_1}\eta_{i_1}^*\big[y_{i_1}^*/u_{i_1}\big]\Big)}_{=:\ \eta_1^*(\bar{\mathbf{u}},\bar{\mathbf{z}})} \wedge \bigwedge_{i_1} P(u_{i_1})\Big) \\
\wedge \ & \Big(\exists\bar{\mathbf{v}}.\ \underbrace{\Big(\bigwedge_{i_2}\eta_{i_2}^*\big[y_{i_2}^*/v_{i_2}\big]\Big)}_{=:\ \eta_2^*(\bar{\mathbf{v}},\bar{\mathbf{z}})} \wedge \bigwedge_{i_2} \neg P(v_{i_2})\Big) \\
\wedge \ & \Big(\bigwedge_{\ell_1} P(z_{\ell_1}^*) \wedge \bigwedge_{\ell_2} \neg P(z_{\ell_2}^*)\Big)
\end{aligned}$$

In what follows we treat the χ_1^*, χ_2^* and η_1^*, η_2^* as indivisible units. One more step remains to establish a kind of "Eliminationshauptform". We move the quantifier blocks $\exists\bar{\mathbf{u}}$ and $\exists\bar{\mathbf{v}}$ outwards over the $\exists P$, reorder the conjuncts within the scope of $\exists P$, and narrow the scope of $\exists P$ such that it does not contain the η_1^*, η_2^* anymore. Moreover, we make use of Lemma 2 and turn the literals $P(u_{i_1})$ into subformulas $\forall x.\, x = u_{i_1} \to P(x)$. We proceed analogously with the literals $\neg P(v_{i_2})$, $P(z^*_{\ell_1})$, and $\neg P(z^*_{\ell_2})$.

$$
\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists\bar{\mathbf{u}}\bar{\mathbf{v}}.\ & \eta_1^*(\bar{\mathbf{u}},\bar{\mathbf{z}}) \wedge \eta_2^*(\bar{\mathbf{v}},\bar{\mathbf{z}}) \\
& \wedge \exists P.\ \Big(\forall x.\ \chi_1^*(x,\bar{\mathbf{z}}) \vee P(x)\Big) \ \wedge \ \Big(\forall x.\ \chi_2^*(x,\bar{\mathbf{z}}) \vee \neg P(x)\Big) \\
& \qquad \wedge \ \Big(\forall x.\ \bigwedge_{i_1}\big(x = u_{i_1} \to P(x)\big)\Big) \ \wedge \ \Big(\forall x.\ \bigwedge_{i_2}\big(x = v_{i_2} \to \neg P(x)\big)\Big) \\
& \qquad \wedge \ \Big(\forall x.\ \bigwedge_{\ell_1}\big(x = z^*_{\ell_1} \to P(x)\big)\Big) \ \wedge \ \Big(\forall x.\ \bigwedge_{\ell_2}\big(x = z^*_{\ell_2} \to \neg P(x)\big)\Big)
\end{aligned}
$$

At this point, the subformula staring with $\exists P$ is in "Eliminationshauptform". After converting the implications into disjunctions and factoring out the $[\neg]P(x)$, we arrive at a formula from which the second-order quantifier $\exists P$ can be eliminated immediately via the basic elimination lemma.

$$
\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists\bar{\mathbf{u}}\bar{\mathbf{v}}.\ & \eta_1^*(\bar{\mathbf{u}},\bar{\mathbf{z}}) \wedge \eta_2^*(\bar{\mathbf{v}},\bar{\mathbf{z}}) \\
& \wedge \exists P.\ \Big(\forall x.\ \big(\chi_1^*(x,\bar{\mathbf{z}}) \wedge \bigwedge_{i_1} x \neq u_{i_1} \wedge \bigwedge_{\ell_1} x \neq z^*_{\ell_1}\big) \vee P(x)\Big) \\
& \qquad \wedge \ \Big(\forall x.\ \big(\chi_2^*(x,\bar{\mathbf{z}}) \wedge \bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2}\big) \vee \neg P(x)\Big)
\end{aligned}
$$

Using Lemma 7, we eliminate the quantifier $\exists P$ and obtain the following result.

$$
\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists\bar{\mathbf{u}}\bar{\mathbf{v}}.\ & \eta_1^*(\bar{\mathbf{u}},\bar{\mathbf{z}}) \wedge \eta_2^*(\bar{\mathbf{v}},\bar{\mathbf{z}}) \\
& \wedge \forall x.\ \Big(\big(\chi_1^*(x,\bar{\mathbf{z}}) \wedge \bigwedge_{i_1} x \neq u_{i_1} \wedge \bigwedge_{\ell_1} x \neq z^*_{\ell_1}\big) \\
& \qquad \vee \big(\chi_2^*(x,\bar{\mathbf{z}}) \wedge \bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2}\big)\Big)
\end{aligned}
$$

In order to convert this result into a somewhat nicer form, we proceed as described in the proof of Lemma 19 in [11]. In particular, we remove the disequations $x \neq y$, where x is a universally quantified variable. To this end, we first distribute disjunction over conjunction within the scope of the quantifier $\forall x$.

$$
\begin{aligned}
\theta(\bar{\mathbf{z}}) \ \wedge \ \exists\bar{\mathbf{u}}\bar{\mathbf{v}}.\ & \eta_1^*(\bar{\mathbf{u}},\bar{\mathbf{z}}) \wedge \eta_2^*(\bar{\mathbf{v}},\bar{\mathbf{z}}) \\
& \wedge \ \forall x.\ \Big(\chi_1^*(x,\bar{\mathbf{z}}) \vee \chi_2^*(x,\bar{\mathbf{z}})\Big) \\
& \qquad \wedge \ \Big(\big(\bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2}\big) \vee \chi_1^*(x,\bar{\mathbf{z}})\Big)
\end{aligned}
$$

$$\wedge\ \Big((\bigwedge_{i_1} x \neq u_{i_1} \wedge \bigwedge_{\ell_1} x \neq z^*_{\ell_1}) \vee \chi^*_2(x, \bar{\mathbf{z}}) \Big)$$
$$\wedge\ \Big((\bigwedge_{i_1} x \neq u_{i_1} \wedge \bigwedge_{\ell_1} x \neq z^*_{\ell_1}) \vee (\bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2}) \Big)$$

Next, we factor the subformulas χ^*_1, χ^*_2, and $\bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2}$ into the conjunctions with which they are disjunctively connected. Moreover, we turn the resulting disjunctions into implications.

$$\begin{aligned}
\theta(\bar{\mathbf{z}})\ \wedge\ \exists \bar{\mathbf{u}}\bar{\mathbf{v}}.\ &\eta^*_1(\bar{\mathbf{u}}, \bar{\mathbf{z}}) \wedge \eta^*_2(\bar{\mathbf{v}}, \bar{\mathbf{z}}) \\
&\wedge\ \forall x.\ \Big(\chi^*_1(x, \bar{\mathbf{z}}) \vee \chi^*_2(x, \bar{\mathbf{z}}) \Big) \\
&\qquad \wedge\ \Big(\bigwedge_{i_2} \big(x = v_{i_2} \to \chi^*_1(x, \bar{\mathbf{z}})\big) \wedge \bigwedge_{\ell_2} \big(x = z^*_{\ell_2} \to \chi^*_1(x, \bar{\mathbf{z}})\big) \Big) \\
&\qquad \wedge\ \Big(\bigwedge_{i_1} \big(x = u_{i_1} \to \chi^*_2(x, \bar{\mathbf{z}})\big) \wedge \bigwedge_{\ell_1} \big(x = z^*_{\ell_1} \to \chi^*_2(x, \bar{\mathbf{z}})\big) \Big) \\
&\qquad \wedge\ \Big(\bigwedge_{i_1} \big(x \neq u_{i_1} \to (\bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2})\big) \Big) \\
&\qquad \wedge\ \Big(\bigwedge_{\ell_1} \big(x \neq z^*_{\ell_1} \to (\bigwedge_{i_2} x \neq v_{i_2} \wedge \bigwedge_{\ell_2} x \neq z^*_{\ell_2})\big) \Big)
\end{aligned}$$

Finally, we apply Lemma 2 in a reverse fashion to remove the universal variable x from some of the subformulas.

$$\begin{aligned}
&\theta(\bar{\mathbf{z}}) \wedge \big(\forall x.\ \chi^*_1(x, \bar{\mathbf{z}}) \vee \chi^*_2(x, \bar{\mathbf{z}})\big) \\
&\wedge \exists \bar{\mathbf{u}}\bar{\mathbf{v}}.\ \eta^*_1(\bar{\mathbf{u}}, \bar{\mathbf{z}}) \wedge \eta^*_2(\bar{\mathbf{v}}, \bar{\mathbf{z}}) \\
&\qquad \wedge\ \bigwedge_{i_2} \chi^*_1(v_{i_2}, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{\ell_2} \chi^*_1(z^*_{\ell_2}, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{i_1} \chi^*_2(u_{i_1}, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{\ell_1} \chi^*_2(z^*_{\ell_1}, \bar{\mathbf{z}}) \\
&\qquad \wedge\ \bigwedge_{i_1}\bigwedge_{i_2} u_{i_1} \neq v_{i_2}\ \wedge\ \bigwedge_{i_1}\bigwedge_{\ell_2} u_{i_1} \neq z^*_{\ell_2}\ \wedge\ \bigwedge_{\ell_1}\bigwedge_{i_2} z^*_{\ell_1} \neq v_{i_2}\ \wedge\ \bigwedge_{\ell_1}\bigwedge_{\ell_2} z^*_{\ell_1} \neq z^*_{\ell_2}
\end{aligned}$$

Consequently, we get the following result.

Theorem 8. *Let* $\varphi := \forall \bar{\mathbf{x}}_1 \exists \bar{\mathbf{y}}_1 \ldots \forall \bar{\mathbf{x}}_n \exists \bar{\mathbf{y}}_n.\ \psi(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{z}})$ *be a relational first-order formula in which* ψ *is quantifier free and the sets* $\bar{\mathbf{x}} := \bar{\mathbf{x}}_1 \cup \ldots \cup \bar{\mathbf{x}}_n$ *and* $\bar{\mathbf{y}} := \bar{\mathbf{y}}_1 \cup \ldots \cup \bar{\mathbf{y}}_n$ *are separated. We assume that every variable occurring in the quantifier prefix and in* $\bar{\mathbf{z}}$ *also occurs in the matrix* ψ.

Let $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1} \subseteq \bar{\mathbf{x}}$ *and* $\widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2} \subseteq \bar{\mathbf{y}}$ *be partitions of the sets* $\bar{\mathbf{x}}$ *and* $\bar{\mathbf{y}}$*, respectively, such that the* $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_{m_1}, \widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_{m_2}$ *are nonempty, pairwise disjoint, and pairwise separated in* φ*. Let* P *be a unary predicate symbol satisfying the following conditions:*

(1) For every set $\widetilde{\mathbf{x}}_i$*,* $1 \leq i \leq m_1$*, there is at most one variable* $x^*_i \in \widetilde{\mathbf{x}}_i$ *for which* φ *contains atoms* $P(x^*_i)$*.*

(2) For every set $\widetilde{\mathbf{y}}_i$, $1 \leq i \leq m_2$, there is at most one variable $y_i^* \in \widetilde{\mathbf{y}}_i$ for which φ contains atoms $P(y_i^*)$.

Then $\exists P.\, \varphi$ is equivalent to some first-order formula φ' that is a finite disjunction of formulas of the form

$$\begin{aligned}
&\theta(\bar{\mathbf{z}}) \wedge \big(\forall x.\, \chi_1^*(x, \bar{\mathbf{z}}) \vee \chi_2^*(x, \bar{\mathbf{z}})\big) \\
&\wedge\ \exists \bar{\mathbf{u}}\bar{\mathbf{v}}.\, \eta_1^*(\bar{\mathbf{u}}, \bar{\mathbf{z}}) \wedge \eta_2^*(\bar{\mathbf{v}}, \bar{\mathbf{z}}) \\
&\qquad \wedge\ \bigwedge_{i_2} \chi_1^*(v_{i_2}, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{\ell_2} \chi_1^*(z_{\ell_2}^*, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{i_1} \chi_2^*(u_{i_1}, \bar{\mathbf{z}})\ \wedge\ \bigwedge_{\ell_1} \chi_2^*(z_{\ell_1}^*, \bar{\mathbf{z}}) \\
&\qquad \wedge\ \bigwedge_{i_1}\bigwedge_{i_2} u_{i_1} \neq v_{i_2}\ \wedge\ \bigwedge_{i_1}\bigwedge_{\ell_2} u_{i_1} \neq z_{\ell_2}^*\ \wedge\ \bigwedge_{\ell_1}\bigwedge_{i_2} z_{\ell_1}^* \neq v_{i_2}\ \wedge\ \bigwedge_{\ell_1}\bigwedge_{\ell_2} z_{\ell_1}^* \neq z_{\ell_2}^*
\end{aligned}$$

where all free predicate symbols and all free first-order variables also occur freely in $\exists P.\, \varphi$. Moreover, all the u_{i_1} are variables from $\bar{\mathbf{u}}$, the v_{i_2} are from $\bar{\mathbf{v}}$, and the $z_{\ell_1}^*$ and $z_{\ell_2}^*$ are certain free variables from $\bar{\mathbf{z}}$.

Example 9. Consider the sentence $\varphi :=\ \exists P.\, \forall x_1 \exists y \forall x_2.\, R(x_1, x_2) \leftrightarrow P(y)$. We transform it into the equivalent sentence

$$\begin{aligned}
\exists P.\ &\Big(\forall x_1 x_2 x_3.\, R(x_1, x_2) \vee \neg R(x_1, x_3)\Big) \\
&\wedge \Big(\big(\forall x_1 x_2.\, R(x_1, x_2)\big) \vee \big(\exists y.\, \neg P(y)\big)\Big) \\
&\wedge \Big(\big(\forall x_1 x_2.\, \neg R(x_1, x_2)\big) \vee \big(\exists y.\, P(y)\big)\Big)\ .
\end{aligned}$$

For the sake of simplicity, we narrow the scope of $\exists P$ so that it only stretches over the last two conjuncts, which we thereafter transform into a disjunction of conjunctions. This yields

$$\begin{aligned}
&\Big(\forall x_1 x_2 x_3.\, R(x_1, x_2) \vee \neg R(x_1, x_3)\Big) \\
&\wedge\ \Big(\exists P.\ \Big(\big(\forall x_1 x_2.\, R(x_1, x_2)\big) \wedge \big(\exists y.\, P(y)\big)\Big) \\
&\qquad\quad \vee \Big(\big(\forall x_1 x_2.\, \neg R(x_1, x_2)\big) \wedge \big(\exists y.\, \neg P(y)\big)\Big) \\
&\qquad\quad \vee \Big(\big(\exists y.\, P(y)\big) \wedge \big(\exists y.\, \neg P(y)\big)\Big)\Big)\ .
\end{aligned}$$

Since we can distribute the quantifier $\exists P$ over disjunction, it is enough to eliminate $\exists P$ in the following three formulas:

(1) $\exists P.\, \exists y.\, P(y)$
 $\models\!\!\!\!\dashv\quad \exists y.\, \exists P.\, \forall x.\, \big(x \neq y \vee P(x)\big) \wedge \big(\texttt{true} \vee \neg P(x)\big)$
 $\models\!\!\!\!\dashv\quad \exists y \forall x.\, x \neq y \vee \texttt{true}$
 $\models\!\!\!\!\dashv\quad \texttt{true}$

(2) $\exists P.\, \exists y.\, \neg P(y)$
 $\models\!\!\!\!\dashv\quad \exists y.\, \exists P.\, \forall x.\, \big(x \neq y \vee \neg P(x)\big) \wedge \big(\texttt{true} \vee P(x)\big)$
 $\models\!\!\!\!\dashv\quad \exists y \forall x.\, x \neq y \vee \texttt{true}$
 $\models\!\!\!\!\dashv\quad \texttt{true}$

(3) $\exists P. \big(\exists y. P(x)\big) \wedge \big(\exists y. \neg P(x)\big)$
$\models\!\mid\quad \exists y_1 y_2. \exists P. \big(\forall x. x \neq y_1 \vee P(x)\big) \wedge \big(\forall x. x \neq y_2 \vee \neg P(x)\big)$
$\models\!\mid\quad \exists y_1 y_2 \forall x. x \neq y_1 \vee x \neq y_2$
$\models\!\mid\quad \exists y_1 y_2. y_1 \neq y_2$

Hence, φ is semantically equivalent to

$$\big(\forall x_1 x_2 x_3. R(x_1, x_2) \vee \neg R(x_1, x_3)\big)$$
$$\wedge \Big(\big(\forall x_1 x_2. R(x_1, x_2)\big) \vee \big(\forall x_1 x_2. \neg R(x_1, x_2)\big) \vee \big(\exists y_1 y_2. y_1 \neq y_2\big)\Big) .$$

Several remarks regrading the shape of the resulting formulas in Theorem 8 are in order. (a) Although the elimination of $\exists P$ potentially introduces new (dis)equations, these only involve existentially quantified and free variables. This means, the separation conditions are not violated by these newly introduced equations. Hence, the introduction of these atoms in one elimination step does not pose an obstacle to the iterated elimination of multiple existential second-order quantifiers. (b) As the subformulas $\chi_1^*(v_{i_2}, \bar{\mathbf{z}})$ may contain universal quantifiers $\forall w$ and atoms $R(\ldots w \ldots v_{i_2} \ldots)$, the separateness condition regarding universally and existentially quantified variables might be violated when introducing the subformulas $\chi_1^*(v_{i_2}, \bar{\mathbf{z}})$ and, similarly, the subformulas $\chi_2^*(u_{i_1}, \bar{\mathbf{z}})$. (c) Perhaps more severely, the introduction of atoms $R(\ldots w \ldots v_{i_2} \ldots)$ may create a connection between sets $\widetilde{\mathbf{x}}_k$ and $\widetilde{\mathbf{y}}_i$, if $w \in \widetilde{\mathbf{x}}_k$ and $v_{i_2} \in \widetilde{\mathbf{y}}_i$. Then, the sets $\widetilde{\mathbf{x}}_k$ and $\widetilde{\mathbf{y}}_i$ are not separated anymore in formulas that contain the new atom. Similar effects might affect pairs $\widetilde{\mathbf{x}}_k, \widetilde{\mathbf{x}}_{k'}$ and $\widetilde{\mathbf{y}}_i, \widetilde{\mathbf{y}}_{i'}$. Hence, if we were to predict whether elimination of both second-order quantifiers in a formula $\exists Q \exists P. \varphi$ is possible using the methods outlined above, we would need to predict which sets of variables will be separated in the formula that results from eliminating $\exists P$.

The above observations seem to make it hard to formulate a version of Theorem 8 that clearly facilitates iterative elimination of multiple quantifiers. On the other hand, it might be worthwhile to base the theorem on a generalization of the separated fragment that still has a decidable satisfiability problem. The *generalized Bernays–Schönfinkel–Ramsey fragment (GBSR)* is described in [9]. In GBSR sentences universally and existentially quantified variables may occur together in atoms under certain restrictions. Roughly speaking, if the existential variable is quantified outside the scope of the quantifier binding the universal variable, the two may occur jointly in atoms. Every GBSR sentence can be transformed into an equivalent sentence that is a finite disjunction of formulas of the form $\exists \bar{\mathbf{y}}. \bigwedge_i \forall \bar{\mathbf{x}}_i. \bigvee_j L_{ij}(\bar{\mathbf{y}}, \bar{\mathbf{x}}_i)$ where the L_{ij} are literals. Hence, Observation (b) might cause fewer troubles in the GBSR setting.

Another interesting aspect is that the symmetry regarding the two Conditions (1) and (2) in Theorem 8 is perhaps more restrictive than necessary. It seems that Condition (2) is obsolete, as the resulting formula in Lemma 6 could be generalized in such a way that the restriction imposed by (2) is not satisfied but second-order quantifiers can still be eliminated.

Altogether, it is subject to future investigations whether Theorem 8 can be enhanced to facilitate iterative elimination of multiple quantifiers.

5 Discussion

We have developed a preliminary result regarding the elimination of second-order quantifiers in a logic fragment that extends the monadic first-order fragment without equality and the Bernays–Schönfinkel–Ramsey fragment.

Notice that the elimination of $\exists P$ from a formula $\exists P.\, \varphi$, where φ is a monadic first-order formula *without equality*, constitutes a special case of the method shown in the present note. The reason is that, if every atom contains at most one variable, then variables cannot occur jointly in atoms. Hence, given a monadic first-order formula φ without equality, any two singleton sets $\{x\}, \{y\}$ of variables are separated in φ. Consequently, any formula $\exists P.\, \varphi$ with monadic φ satisfies the prerequisites of Theorem 8, if we choose the sets $\widetilde{\mathbf{x}}_k$ and $\widetilde{\mathbf{y}}_i$ to be singleton sets covering all the variables that occur bound in φ.

The presented result can only be a first step towards the formulation of a novel fragment of second-order logic that (a) extends the monadic second-order fragment, (b) is based on the concept of separateness of certain variables at the atomic level, (c) admits elimination of second-order quantifiers, also in an iterated fashion. The discussion following Theorem 8 already makes clear that a lot remains to be done, in order to achieve this goal. Furthermore, there seems to be no good reason to confine ourselves to the elimination of quantifiers over unary predicates, but aim for higher arities as well. Moreover, (b) can be weakened by taking boolean structure into account instead of only concentrating on the atoms in a given formula. For example, the formula $\exists P. \forall xy.\, P(x) \wedge \big(P(y) \vee R(x, y)\big)$ does not satisfy the prerequisites of Theorem 8, as $\{x\}$ and $\{y\}$ are not separated and the set $\{x, y\}$ contains two variables that occur as arguments of P. However, the theorem can be applied to the equivalent formula $\exists P. \forall x_1 x_2 y.\, P(x_1) \wedge \big(P(y) \vee R(x_2, y)\big)$, as the sets $\{x_1\}$ and $\{x_2, y\}$ are separated and x_2 does not occur as argument of P. As a third possible improvement, equations between universal and existential variables should be allowed in a less restrictive way than they are in the present note. To this end, some of the methods that are used to handle equations during quantifier elimination in the monadic second-order fragment might be applicable in the more general setting as well.

In the present note we concentrate on transforming the input formulas syntactically until the basic elimination lemma (Lemma 7) is applicable. In future work, it is of course advisable to also try other known approaches, such as the ones described in [4], e.g. the SCAN algorithm, the DLS* algorithm, hierarchical theorem proving, or variations thereof. The unmodifed DLS algorithm, as presented in [4], fails on the logic fragment described in Theorem 8 in the present note. In particular, the preprocessing phase is not always able to transform the input into the required form, although this is possible in principle. This is already true for monadic sentences such as $\varphi := \exists P.\, \forall x \exists y.\, \big(\neg P(x) \vee P(y)\big) \wedge \big(P(x) \vee \neg P(y)\big)$, which is equivalent to $\exists P.\, \forall x \exists y.\, P(x) \leftrightarrow P(y)$. Conradie gave a necessary and sufficient condition on the syntax of formulas in which DLS can successfully eliminate an existential second-order quantifier [3]. It turns out that the occurrences of P in φ violate Conradie's condition in many ways. (Every occurrence of P is in *malignant* conjunctions and disjunctions and inside a $\forall\exists$-*scope*.) Nonetheless, it is not hard to see that there is a first-order formula that is equivalent to φ, namely `true`. A slight modification of the DLS preprocessing step in the spirit of Claim I, used in the proof of Lemma 4, might already solve this particular issue.

Acknowledgement The present author is indebted to the anonymous reviewers for their constructive criticism and valuable suggestions.

References

1. Ackermann, W.: Untersuchungen über das Eliminationsproblem der mathematischen Logik. Mathematische Annalen 110, 390–413 (1935)
2. Behmann, H.: Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem. Mathematische Annalen 86(3–4), 163–229 (1922)
3. Conradie, W.: On the strength and scope of DLS. Journal of Applied Non-Classical Logics 16(3-4), 279–296 (2006)
4. Gabbay, D., Schmidt, R., Szałas, A.: Second-Order Quantifier Elimination: Foundations, Computational Aspects and Applications. College Publications (2008)
5. van Heijenoort, J.: From Frege to Gödel – A Source Book in Mathematical Logic, 1879–1931. Harvard University Press (2002)
6. Löwenheim, L.: Über Möglichkeiten im Relativkalkül. Mathematische Annalen 76, 447–470 (1915), an English translation can be found in [5].
7. Skolem, T.: Untersuchungen über die Axiome des Klassenkalküls und über Produktations- und Summationsprobleme welche gewisse Klassen von Aussagen betreffen. Videnskapsselskapets Skrifter I. Mat.-Nat. Klasse (3) (1919)
8. Sturm, T., Voigt, M., Weidenbach, C.: Deciding First-Order Satisfiability when Universal and Existential Variables are Separated. In: Logic in Computer Science (LICS'16). pp. 86–95 (2016)
9. Voigt, M.: On Generalizing Decidable Standard Prefix Classes of First-Order Logic Submitted, a preprint version is available under arXiv:1706.03949 [cs.LO].
10. Voigt, M.: A Fine-Grained Hierarchy of Hard Problems in the Separated Fragment. In: Logic in Computer Science (LICS'17). pp. 1–12 (2017)
11. Wernhard, C.: Heinrich Behmann's Contributions to Second-Order Quantifier Elimination from the View of Computational Logic. Tech. rep., TU Dresden (2015)