The Bernays-Schönfinkel-Ramsey Class of Separation Logic with Uninterpreted Predicates

Mnacho Echenim, Radu Iosif, Nicolas Peltier

The Bernays-Schönfinkel-Ramsey Class of Separation Logic with Uninterpreted Predicates

MNACHO ECHENIM, Univ. Grenoble Alpes, CNRS, LIG
RADU IOSIF, Univ. Grenoble Alpes, CNRS, Verimag
NICOLAS PELTIER, Univ. Grenoble Alpes, CNRS, LIG

This paper investigates the satisfiability problem for Separation Logic with k record fields, with unrestricted nesting of separating conjunctions and implications. It focuses on prenex formulæ with a quantifier prefix in the language $\exists^*\forall^*$, that contain uninterpreted (heap-independent) predicate symbols. In analogy with first-order logic, we call this fragment *Bernays-Schönfinkel-Ramsey Separation Logic* [$\mathsf{BSR}(\mathsf{SL}^k)$]. In contrast with existing work on Separation Logic, in which the universe of possible locations is assumed to be infinite, we consider both finite and infinite universes in the present paper. We show that, unlike in first-order logic, the (in)finite satisfiability problem is undecidable for $\mathsf{BSR}(\mathsf{SL}^k)$. Then we define two non-trivial subsets thereof, for which the finite and infinite satisfiability problems are PSPACE-complete, respectively, assuming that the maximum arity of the uninterpreted predicate symbols does not depend on the input. These fragments are defined by controlling the polarity of the occurrences of separating implications, as well as the occurrences of universally quantified variables within their scope. These decidability results have natural applications in program verification, as they allow to automatically prove lemmas that occur in e.g. entailment checking between inductively defined predicates and validity checking of Hoare triples expressing partial correctness conditions.

Additional Key Words and Phrases: Separation logic, Bernays-Schönfinkel-Ramsey class, decision procedures, complexity, PSPACE-completeness

ACM Reference Format:
Mnacho Echenim, Radu Iosif, and Nicolas Peltier. 2019. The Bernays-Schönfinkel-Ramsey Class of Separation Logic with Uninterpreted Predicates. *ACM Trans. Comput. Logic* 1, 1 (December 2019), 46 pages.

1 INTRODUCTION

Separation Logic [14, 20] (SL) is a logical framework used in program verification to describe properties of the dynamically allocated memory, such as topologies of data structures (lists, trees), (un)reachability between pointers, etc. In a nutshell (formal definitions are given below), given an integer $k \geq 1$, the logic SL^k is obtained from the first-order theory of a finite partial function $h : \mathsf{U} \rightharpoonup \mathsf{U}^k$ called a *heap*, by adding two non-classical connectives:

1. the *separating conjunction* $\phi_1 * \phi_2$, that asserts the existence of a split of the heap into disjoint heaps satisfying ϕ_1 and ϕ_2 respectively, and
2. the *separating implication*, or *magic wand* $\phi_1 \mathrel{-\!\!*} \phi_2$, stating that each extension of the heap by a disjoint heap satisfying ϕ_1 must satisfy ϕ_2.

Authors' addresses: Mnacho Echenim, Mnacho.Echenim@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, CNRS, LIG, , Grenoble, France, 38000; Radu Iosif, Radu.Iosif@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, CNRS, Verimag, Grenoble, France, 38000; Nicolas Peltier, Nicolas.Peltier@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, CNRS, LIG, , Grenoble, France, 38000.

1529-3785/2019/12-ART $15.00
https://doi.org/

Intuitively, the set U denotes the universe of possible of memory locations (cells) and k is the number of record fields in each memory cell. The separating connectives $*$ and $-\!\!*$ may be used to express dynamic transformations of the heap. As such, they allow for concise definitions of program semantics, via weakest precondition calculi [14] and easy-to-write specifications of recursive linked data structures (e.g. singly- and doubly-linked lists, trees with linked leaves and parent pointers, etc.), when inductive definitions are added [20].

Investigating the decidability and complexity of the satisfiability problem for fragments of SL is thus of great theoretical and practical interest. In contrast to first-order logic for which the decision problem has been thoroughly investigated (see, e.g., [3]), only a few results are known for SL. The earliest such results show undecidability of SL^k and the PSPACE-completeness of its quantifier-free fragment, for any $k \geq 2$ [7]. These results have been subsequently refined, by showing undecidability of SL^1, even if only two quantified variables are allowed [8]. Decidability of SL^1 is shown for the fragment without the magic wand connective, but the complexity lower bound is not elementary recursive. This lower bound drops if at most one quantified variable is allowed, in which case SL^1 is PSPACE-complete. Extending SL^1 with higher-order inductive predicates, such as reachability, leads to undecidability in the presence of the magic wand and becomes PSPACE-complete if the magic wand is not allowed [9].

A salient feature of SL is the ability of describing recursive data structures by means of inductive definitions. The axioms defining such interpreted predicates use a very restricted fragment of SL, consisting of atoms (equalities, disequalities and single cell descriptions) joined with separating conjunctions, called the *symbolic heap* fragment. Since negation does not occur within symbolic heaps, one must consider the satisfiability and entailment problems separately. For instance, satisfiability of a symbolic heap is EXPTIME-complete, in general, and NP-complete if the maximum arity of the predicates is a constant, not part of the input [5]. On the other hand, entailment between symbolic heaps is undecidable in general, and becomes elementary recursive under certain conditions guaranteing that the treewidth of each model is bounded by the size of the inductive definition [13]. In particular, the problem is EXPTIME-hard [1] and the more restricted problem of the validity of entailments of the form $P(x_1, \ldots, x_n) \models Q(x_1, \ldots, x_n)$ has been recently shown to belong to 2EXPTIME [15].

In this paper, we consider prenex SL formulæ with a quantifier prefix in the language $\exists^*\forall^*$, possibly containing heap-independent uninterpreted[1] predicate symbols. In analogy with the Bernays-Schönfinkel-Ramsey fragment of first-order logic with $\exists^*\forall^*$ quantifier prefix, equality and uninterpreted predicates and without function symbols of arity greater than 0 [BSR(FO)] [18], we call this fragment *Bernays-Schönfinkel-Ramsey* SL [$\mathsf{BSR}(\mathsf{SL}^k)$].

As far as we are aware, all existing work on SL assumes that the universe U is countably infinite. This assumption is not necessarily realistic in practice since the available memory is usually finite, although the bound depends on the hardware and is not known in advance. However, reasoning about pointer-manipulating programs under the finite memory assumption proves to be harder than under the assumption that memory is infinite, when the bound on the memory size is not known à priori. In particular, the frame rule of classical Separation Logic [**?**], which is a crucial enabler of local reasoning, breaks, in general, for programs that allocate memory, because, intuitively, adding frames is not possible unless enough free memory is available. Nevertheless, restricted versions of the frame rule still hold, with additional side conditions on the structure of the programs and/or the context to which it is applied. A thorough investigation of the soundness of the frame rule for bounded memory domains is, however, out of the scope of this paper and considered as future work.

In this paper we consider the satisfiability problem for $\mathsf{BSR}(\mathsf{SL}^k)$, with $k \geq 2$, in both cases of finite and infinite universe, referred to as finite and infinite satisfiability, respectively. We show that both problems are undecidable (unlike in the BSR fragment of first-order logic) and that they become PSPACE-complete under some additional restrictions, related to the occurrences of the magic wand and universal variables, namely:

[1]By "uninterpreted" we mean that the interpretation of such predicate symbols is not fixed by a theory or by inductive definitions.

I. The infinite satisfiability problem is PSPACE-complete if the positive occurrences of $-\!\!*$ (i.e., the occurrences of $-\!\!*$ that are in the scope of an even number of negations) contain no universal variables.
II. The finite satisfiability problem is PSPACE-complete if there is no positive occurrence of $-\!\!*$ (i.e., $-\!\!*$ only occurs in the scope of an odd number of negations). This additional restriction stems from the fact that, actually, the finite satisfiability problem becomes undecidable even for only one positive occurrence of a $-\!\!*$ with no variable within its scope.

These results establish sharp decidability frontiers within $\mathsf{BSR}(\mathsf{SL}^k)$. In both cases, we assume that the arity of the uninterpreted predicate symbols is bounded by a constant (the satisfiability problem is already NEXPTIME-complete for BSR first-order formulæ with unbounded predicate arity [17]). In contrast, the number k of record fields is not bounded and may be part of the input. Reasoning on finite domains is more difficult than on infinite ones, due to the possibility of asserting cardinality constraints on unallocated cells, which explains that the latter condition is more restrictive than the former one. However, the finite universe hypothesis is especially useful when dealing with bounded memory issues, for instance checking that the execution of a program satisfies its postcondition, provided that there are sufficiently many available memory cells.

Theory-parameterized versions of $\mathsf{BSR}(\mathsf{SL}^k)$ have been shown to be undecidable in [19], e.g. when integer linear arithmetic is used to reason about locations, and claimed to be PSPACE-complete for countably infinite and finite unbounded location sorts, with no relation other than equality. In the present paper, we show that this claim is wrong, and draw a precise chart of decidability for both infinite and finite satisfiability of $\mathsf{BSR}(\mathsf{SL}^k)$, for $k \geq 2$. To complete the picture, the entire prenex fragment of SL^1 has been recently shown decidable but not elementary recursive, whereas the fragment $\mathsf{BSR}(\mathsf{SL}^1)$ is PSPACE-complete [10].

Undecidability is shown by reduction from BSR first-order formulæ with two monadic function symbols, for which satisfiability is known to be undecidable [3]. To establish the decidability results, we first show that every quantifier-free SL formula can be transformed into an equivalent boolean combination of formulæ of some specific patterns, called *test formulæ*. This result is interesting in itself, since it provides a precise and intuitive characterization of the expressive power of SL: it shows that separating connectives can be confined to a small set of test formulæ. Such expressive completeness results were already known for infinite universes (see, e.g., [16]), but our transformation algorithm also provides insights on the form of the obtained formulæ, especially on the polarity of occurrences of some test formulæ, which turns out to be useful latter on in the remainder of the paper. Further, we extend the expressive completeness result to finite universes, with additional test formulæ asserting cardinality constraints on unallocated cells.

One advantage of the translation to test formulæ is that the latter can be straightforwardly translated into first-order formulæ, by encoding the heap as a $(k+1)$-ary predicate. Note that another translation of quantifier-free SL^k into first-order logic with equality has been described in [6]. There, the small model property of quantifier-free SL^k [7] is used to bound the number of first-order variables to be considered and the separating connectives are interpreted as first-order quantifiers. The result is an equisatisfiable first-order formula. This translation scheme cannot be, however, directly applied to $\mathsf{BSR}(\mathsf{SL}^k)$, which does not have a small model property, being moreover undecidable.

We focus first on the infinite satisfiability problem and show that, if the above condition (I) is satisfied, then the obtained first-order formulæ are in the BSR(FO) class. The infinite satisfiability problem for BSR(SL) is thus reduced to the satisfiability problem for BSR(FO), with some additional constraints on the cardinality of the interpretation: the universe must be infinite, and the heap must be finite. We show that these constraints may be handled by relying on an existing characterization of the models of BSR(FO) formulæ with infinitely countable universe [12].

For the finite satisfiability class, the decidability proof is more involved, as the obtained first-order formulæ are not in BSR(FO), even if the above condition (II) is satisfied. However, this problem can be overcome by focusing

on some class of structures satisfying additional properties ensuring that a reduction to BSR(FO) is feasible. Note that in this case, the cardinality constraints on the universe and heap are straightforward to handle, as the BSR(FO) class is finitely controllable (i.e., every satisfiable BSR(FO) formula has a finite model).

The above transformation algorithm does not by itself provide an efficient decision procedure, as the size of the obtained boolean combination of test formulæ is exponential w.r.t. that of the initial (BSR) formula. The PSPACE upper bound thus relies on a careful analysis of the maximal size of the test formulæ. The analysis reveals that, although the boolean combination of test formulæ is of exponential size, its so-called *minterms* (i.e., the conjunctions in its disjunctive normal form) are of polynomial size and can be enumerated in polynomial space. The above algorithms can thus be refined to run in polynomial space.

This paper is an extended and thoroughly revised version of the conference paper [11]. The latter paper only handles SL formulæ with no uninterpreted predicate symbols. The addition of uninterpreted predicate symbols has a limited impact on the transformation of SL formulæ into boolean combinations of test formulæ. Indeed, since these predicates do not depend on the heap the corresponding atoms can be easily shifted outside of the separated connectives. However, non trivial adaptations are required in the satisfiability tests, since the presence of uninterpreted predicates makes it much more difficult to ensure that the considered formula has a model of the expected cardinality (finite or infinite).

Applications

Let us sketch two applications of our results to program verification. The first application is building proofs of validity for the entailments between inductively defined predicates in SL. The second application is proving the validity of Hoare triples with SL as base logic.

Checking Entailment between Inductively Defined Predicates. In contrast to other approaches [5**?**], our logic does not allow for inductively defined predicates (the predicates we consider are independent of the heap). Still, our results, embedded in inductive proof procedures, could prove useful to check entailment between formulæ containing such predicates. Consider for instance the following inductive definitions, describing a list segment with strictly increasing data fields and a possibly cyclic list segment, with no restrictions on the data, respectively:

$$\widehat{\mathsf{ls}}(x,y,d) \leftarrow \mathsf{emp} \wedge x \approx y \ \vee \ \exists z \exists e \,.\, d \prec e \wedge (x \mapsto (d,z) * \widehat{\mathsf{ls}}(z,y,e)) \qquad \text{sorted list segment from } x \text{ to } y$$

$$\mathsf{ls}(x,y) \leftarrow \mathsf{emp} \wedge x \approx y \ \vee \ \exists u \exists f \,.\, x \mapsto (f,u) * \mathsf{ls}(u,y) \qquad \text{unrestricted list segment from } x \text{ to } y$$

Intuitively, a list segment is either empty, in which case the head and the tail coincide [$\mathsf{emp} \wedge x \approx y$], or it contains at least one element. We denote by $x \mapsto (d,z)$ the fact that x is the only allocated memory location, which moreover points to a pair (d,z), where d is a data field and z is a pointer field. When writing $x \mapsto (d,z) * \mathsf{ls}(z,y,e)$ we mean that $x \mapsto (d,z)$ and $\widehat{\mathsf{ls}}(z,y,e)$ must hold over disjoint parts of the heap. The constraint $d \prec e$, in the inductive definition of $\widehat{\mathsf{ls}}$, captures the fact that the list is strictly increasing, $\prec$ being an uninterpreted predicate symbol that satisfies the transitivity and anti-symmetry axioms below:

$$\forall a \forall b \forall c \,.\, a \prec b \wedge b \prec c \rightarrow a \prec c \qquad\qquad \forall a \forall b \,.\, a \prec b \wedge b \prec a \rightarrow a \approx b$$

Now consider a fragment of the inductive proof showing that any sorted list segment is also a list segment:

$$\dfrac{\dfrac{\widehat{\mathsf{ls}}(z,y,e) \vdash \mathsf{ls}(z,y)}{d \prec e \wedge x \mapsto (d,z) * \widehat{\mathsf{ls}}(z,y,e) \vdash \exists u \exists f \,.\, x \mapsto (f,u) * \mathsf{ls}(u,y) \vee \mathsf{emp} \wedge x \approx y}}{\widehat{\mathsf{ls}}(x,y,d) \vdash \mathsf{ls}(x,y)} \quad \begin{array}{l} d \prec e \wedge x \mapsto (d,z) \models \exists u \exists f \,.\, x \mapsto (f,u) \\ \text{by instantiation } u \leftarrow z, f \leftarrow d \end{array}$$

The bottom inference rule introduces one of the two cases produced by unfolding the inductive definitions on both sides of the sequent[2]. Note that the quantifications $\exists z, e$ on the left-hand side have been omitted because they can be eliminated by using the standard $\exists$-left rule of the sequent calculus (if z and e are fresh variables). The second inference rule is a reduction of the sequent obtained by unfolding, to a sequent matching the initial one (by renaming z to x and e to d), and allows to close this branch of the proof by an inductive argument, based on the principle of infinite descent [4]. The simplification applied by the second inference above relies on the validity of the entailment $d \prec e \wedge x \mapsto (d, z) \models \exists u \exists f \,.\, x \mapsto (f, u)$, which reduces to the (un)satisfiability of the formula $d \prec e \wedge x \mapsto (d, z) \wedge \forall u \forall f \,.\, \neg x \mapsto (f, u)$. The latter falls into the $\mathsf{BSR}(\mathsf{SL}^2)$ fragment. A consequence of the results in this paper is that, if the inductive rules contain no occurrence of $\mathrel{-\!\!*}$ and $\forall$, then there exist algorithms for solving the above entailment problem in both finite and infinite universes, in the presence of uninterpreted predicates. The only requirement is that the axiomatization of these predicates can be done using BSR(FO), i.e., that the interpretation of these predicates does not depend on the heap.

Checking Inductive Invariants with Universal Quantifiers. Purely universal SL formulæ are also useful to express pre- or post-conditions asserting "local" constraints on the shape of the data structures manipulated by a program. For instance, the atomic proposition $x \mapsto (p, n, d)$ states that the value of the heap at x is the triple (p, n, d), where n (resp. p) is the location of the next (resp. previous) cell in the list and d is a data value. Moreover, $x \mapsto (p, n, d)$ holds if and only if there is no location, other than x, in the domain of the heap. With this in mind, the following formula describes a well-formed doubly-linked sorted list:

$$\forall x_1, x_2, x_3, x_4, x_5, y_1, y_2 \,.\, x_1 \mapsto (x_2, x_3, y_1) * x_2 \mapsto (x_4, x_5, y_2) * \top \rightarrow x_5 \approx x_1 \wedge y_1 \prec y_2 \tag{1}$$

Such constraints cannot be expressed by using inductively defined predicates for which the entailment problem is known to be decidable[3], which shows the practical relevance of the considered fragment. The separating implication (magic wand) seldom occurs in such shape constraints. However, it is useful to describe the dynamic transformations of the data structures, as in the following Hoare-style axiom, giving the weakest precondition of a universal formula $\forall \mathbf{u} \,.\, \psi$ with respect to redirecting the i-th record field of x to z [14]:

$$\{\mathsf{x} \mapsto (y_1, \ldots, y_k) * [\mathsf{x} \mapsto (y_1, \ldots, y_{i-1}, \mathsf{z}, \ldots, y_k) \mathrel{-\!\!*} \forall \mathbf{u} \,.\, \psi]\}\ \mathsf{x.i} := \mathsf{z}\ \{\forall \mathbf{u} \,.\, \psi\}$$

For example, the Hoare-style axiom for the weakest precondition of the universal formula $\forall \mathbf{u} \,.\, \psi$ when redirecting the 'next' field in a doubly-linked list is

$$\{\mathsf{x} \mapsto (p, n, d) * [\mathsf{x} \mapsto (p, \mathsf{z}, d) \mathrel{-\!\!*} \forall \mathbf{u} \,.\, \psi]\}\ \mathsf{x.next} := \mathsf{z}\ \{\forall \mathbf{u} \,.\, \psi\}.$$

Intuitively, the formula $\mathsf{x} \mapsto (p, n, d) * [\mathsf{x} \mapsto (p, \mathsf{z}, d) \mathrel{-\!\!*} \forall \mathbf{u} \,.\, \psi]$ holds when the heap can be separated into disjoint parts, one in which cell x is allocated, and one that, when extended with a heap in which the 'next' field of x is mapped to z, satisfies $\forall \mathbf{u} \,.\, \psi$. The universal formula $\forall \mathbf{u} \,.\, \psi$ could be the doubly-linked list invariant (1) for instance.

In the general case, the precondition for the redirection of the i-th record field of x to z is equivalent to $\forall \mathbf{u} \,.\, \mathsf{x} \mapsto (y_1, \ldots, y_k) * [\mathsf{x} \mapsto (y_1, \ldots, y_{i-1}, \mathsf{z}, \ldots, y_k) \mathrel{-\!\!*} \psi]$ because, although hoisting universal quantifiers outside of the separating conjunction is unsound in general, this is possible here due to the special form of the left-hand side $\mathsf{x} \mapsto (y_1, \ldots, y_{i-1}, \mathsf{z}, \ldots, y_k)$ which unambiguously defines a single heap cell.

Checking entailment between two universal formulæ boils down to checking the satisfiability of a $\mathsf{BSR}(\mathsf{SL}^k)$ formula, which can be done thanks to the decidability results in our paper. In particular, checking that $\forall \mathbf{u} \,.\, \psi$ is an invariant of the program statement x.i := z amounts to checking that the formula $\forall u \,.\, \psi \wedge \exists \mathbf{u} \,.\, \neg[\mathsf{x} \mapsto$

[2]The second case $\mathsf{emp} \wedge x \approx y \vdash \exists u \exists f \,.\, x \mapsto (f, u) * \mathsf{ls}(u, y) \vee \mathsf{emp} \wedge x \approx y$ is trivial and omitted for clarity.

[3]This is due to the fact that some of the edges, for instance those pointing to list values, may be "dangling". In other words, this structure does not fulfill the so called *establishment* condition of [13].

$(y_1, \ldots, y_k) * (\mathsf{x} \mapsto (y_1, \ldots, y_{i-1}, \mathsf{z}, \ldots, y_k) \mathrel{-\!\!*} \psi)]$ is unsatisfiable. Because the magic wand occurs negated, this formula falls into a decidable class defined in the present paper, for both finite and infinite satisfiability.

Roadmap

The remainder of the paper is structured as follows. In Section 2, usual notions and results are briefly reviewed and the definition of the logic SL^k is provided. In Section 4 a set of formula patterns, called *test formulæ*, is introduced, and it is shown that these patterns can be expressed in first-order logic. In Section 5, an algorithm is described to transform every SL^k formula into an equivalent boolean combination of test formulæ. The output formula is of exponential size, however, we show that the conjunctions of literals occurring in its disjunctive normal form are of polynomial size and may be enumerated in polynomial space. In Section 6, the $\mathsf{BSR}(\mathsf{SL}^k)$ class is investigated and (un)decidability and complexity results are established based on the previous transformation algorithms. Section 7 briefly concludes the paper.

2 PRELIMINARIES

2.1 First Order Logic

Syntax. We denote by $\mathbb{Z}$ and $\mathbb{N}$ the sets of integer and natural numbers, respectively. Let $\mathbb{Z}_\infty = \mathbb{Z} \cup \{\infty\}$ and $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$, where for each $n \in \mathbb{Z}$ we have $n + \infty = \infty$ and $n < \infty$. For any countable set S, we denote by $||S|| \in \mathbb{N}_\infty$ the cardinality of S.

Let U be a sort symbol denoting a *universe* sort and let B be the usual boolean sort. We consider a countably infinite set Var of variables of sort U, ranged over by x, y, z, and a countably infinite set $\mathcal{F}$ of function symbols. Each function symbol $f \in \mathcal{F}$ has a sort $\sigma(f) \in \{\mathsf{U}, \mathsf{B}\}$. A function symbol f takes $\#(f) \geq 0$ arguments of sort U. If $\#(f) = 0$ we call f a *constant* and if $\#(f) = 1$ we say that f is *monadic*. If $\sigma(f) = \mathsf{B}$, f is called a *predicate*. First-order (FO) terms t and formulæ φ are defined by the following grammar:

$$t \;:=\; x \mid f(t_1, \ldots, t_{\#(f)}) \qquad\qquad \varphi \;:=\; \bot \mid \top \mid t_1 \approx t_2 \mid q(t_1, \ldots, t_{\#(q)}) \mid \varphi \wedge \varphi \mid \neg\varphi \mid \exists x \,.\, \varphi$$

where $x \in \mathsf{Var}$, $f, q \in \mathcal{F}$, $\sigma(f) = \mathsf{U}$ and $\sigma(q) = \mathsf{B}$. The logical symbols $\bot$ and $\top$ denote the boolean constants false and true, respectively. As usual, $f(t_1, \ldots, t_n)$ is simply written f if $n = 0$. We write $\varphi_1 \vee \varphi_2$ for $\neg(\neg\varphi_1 \wedge \neg\varphi_2)$, $\varphi_1 \rightarrow \varphi_2$ for $\neg\varphi_1 \vee \varphi_2$, $\varphi_1 \leftrightarrow \varphi_2$ for $\varphi_1 \rightarrow \varphi_2 \wedge \varphi_2 \rightarrow \varphi_1$ and $\forall x \,.\, \varphi$ for $\neg\exists x \,.\, \neg\varphi$. The *size* of a formula φ, denoted as $\mathsf{size}(\varphi)$, is the number of occurrences of symbols in it.

We denote by $\mathsf{Var}(\varphi)$ the set of variables that occur free in φ, i.e. not in the scope of a quantifier, by $\mathcal{F}(\phi)$ the set of function symbols occurring in ϕ, by $\mathcal{P}(\phi)$ the set of predicate symbols in $\mathcal{F}(\phi)$ and by $\mathsf{Const}(\phi)$ the set of constants of sort U in ϕ.

A vector of variables will often be denoted by $\mathbf{x}, \mathbf{y} \ldots$, and x_i will denote the i-th component of $\mathbf{x}$. An equation $\mathbf{x} \approx \mathbf{y}$ with $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ denotes the formula $\bigwedge_{i=1}^n x_i \approx y_i$.

Semantics. First-order formulæ are interpreted over FO-*structures*[4] $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$, where $\mathfrak{U}$ is a nonempty countable set, called a *universe*, the elements of which are called *locations*; $\mathfrak{s} : \mathsf{Var} \rightharpoonup \mathfrak{U}$ is a partial mapping of variables to elements of $\mathfrak{U}$, called a *store* and $\mathcal{I}$ interprets each function symbol f by a function $f^{\mathcal{I}} : \mathfrak{U}^{\#(f)} \rightarrow \mathfrak{U}$ if $\sigma(f) = U$ or by a relation $f^{\mathcal{I}} \subseteq \mathfrak{U}^{\#(f)}$ if $\sigma(f) = \mathsf{B}$. A structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is *finite* when $||\mathfrak{U}|| \in \mathbb{N}$ and *infinite* otherwise.

By writing $\mathcal{S} \models \varphi$, for a structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$, we mean that $\mathsf{Var}(\varphi) \subseteq \mathsf{dom}(\mathfrak{s})$ and φ is true when interpreted in $\mathcal{S}$. This relation is defined recursively on the structure of φ, as usual. When $\mathcal{S} \models \varphi$, we say that $\mathcal{S}$ is a *model* of φ. A formula is [finitely] *satisfiable* when it has a [finite] model. Given two formulæ φ_1 and φ_2, we say that φ_1 *entails* φ_2 (written $\varphi_1 \models \varphi_2$) when every model of φ_1 is a model of φ_2, and that φ_1 and φ_2 are *equivalent* (written $\varphi_1 \equiv \varphi_2$) when $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \varphi_1 \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \varphi_2$, for every structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$. For any store $\mathfrak{s}$ on $\mathfrak{U}$, variables $x_1, \ldots, x_n$ and

[4]These will simply be called structures, when no confusion arises.

elements $\ell_1, \ldots, \ell_n \in \mathfrak{U}$, we denote by $\mathfrak{s}[x_1 \leftarrow \ell_1, \ldots, x_n \leftarrow \ell_n]$ the store that coincides with $\mathfrak{s}$ on every variable not in $\{x_1, \ldots, x_n\}$ and maps x_i to ℓ_i, for all $i = 1, \ldots, n$. We also call $\mathfrak{s}[x_1 \leftarrow \ell_1, \ldots, x_n \leftarrow \ell_n]$ an *extension* of $\mathfrak{s}$. If $\mathbf{y} = (y_1, \ldots, y_n)$ is a vector of variables, and $\mathfrak{s}$ is a store, then $\mathfrak{s}(\mathbf{y})$ denotes the vector $(\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_n))$.

BSR(FO) Formulæ. The *Bernays-Schönfinkel-Ramsey* fragment of FO [BSR(FO)] is the set of formulæ of the form $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_m \,.\, \varphi$, where φ is a quantifier-free formula and all function symbols $f \in \mathcal{F}(\varphi)$ of arity $\#(f) > 0$ have sort $\sigma(f) = \mathsf{B}$. For simplicity we often restrict ourselves to BSR(FO) formulæ containing no existential quantification. This is without any loss of generality, since $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_m \,.\, \varphi$ is satisfiable if and only if $\forall y_1 \ldots \forall y_m \,.\, \varphi$ is satisfiable.

Definition 2.1. Consider the structures $\mathcal{S} \stackrel{\text{def}}{=} (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ and $\mathcal{S}' \stackrel{\text{def}}{=} (\mathfrak{U}', \mathfrak{s}', \mathcal{I}')$. The structure $\mathcal{S}'$ is called a *restriction* of $\mathcal{S}$ to $\mathfrak{U}'$ if $\mathfrak{U}' \subseteq \mathfrak{U}$, $\mathfrak{s}'(x) = \mathfrak{s}(x)$ for every $x \in \mathrm{dom}(\mathfrak{s})$, $q^{\mathcal{I}'} = q^{\mathcal{I}} \cap \mathfrak{U}'^{\#(q)}$ for every predicate symbol q and $f^{\mathcal{I}'} = f^{\mathcal{I}}$ for every function symbol f.

The following proposition states a well-known property of BSR(FO):

Proposition 2.2. *Let φ be a formula in* BSR(FO) *with no existential quantifier and let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be a model of φ. If $\mathfrak{U}'$ is a nonempty subset of $\mathfrak{U}$ such that $\{\mathfrak{s}(x) \mid x \in \mathrm{Var}(\varphi)\} \cup \{c^{\mathcal{I}} \mid c \in \mathrm{Const}(\varphi)\} \subseteq \mathfrak{U}'$ and $\mathcal{S}' \stackrel{\text{def}}{=} (\mathfrak{U}', \mathfrak{s}', \mathcal{I}')$ is a restriction of $\mathcal{S}$ to $\mathfrak{U}'$, then $\mathcal{S}'$ is a model of φ. As a consequence, if φ is satisfiable, then it admits a model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $||\mathfrak{U}|| \leq \max(1, ||\mathrm{Var}(\varphi)|| + ||\mathrm{Const}(\varphi)||)$.*

Proof. See for instance [12, Theorem 3]. □

The decidability of BSR(FO) is a consequence of the above small model property. It is known that the satisfiability problem for this class is NEXPTIME-complete [**?**]. The condition requiring the absence of function symbols of sort U in BSR(FO) is justified by the fact that undecidability occurs as soon as two monadic function symbols are allowed. Let $\mathrm{BSR}^2(\mathrm{FO})$ be the extension of BSR(FO) consisting of the formulæ $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_m \,.\, \varphi$, where φ is a quantifier-free formula in which at most two monadic function symbols occur.

Proposition 2.3. *The satisfiability problem is undecidable for* $\mathrm{BSR}^2(\mathrm{FO})$*, even if only one universal quantifier and no predicates are allowed.*

Proof. See [3, Theorem 4.1.8]. □

2.2 Separation Logic

Syntax Let k be a strictly positive integer. The logic SL^k is the set of formulæ generated by the grammar:

$$\varphi \quad := \quad \bot \mid \top \mid \mathsf{emp} \mid x \approx y \mid x \mapsto (y_1, \ldots, y_k) \mid q(x_1, \ldots, x_{\#(q)}) \mid \varphi \wedge \varphi \mid \neg\varphi \mid \varphi * \varphi \mid \varphi \mathrel{-\!\!*} \varphi \mid \exists x \,.\, \varphi$$

where $x, y, y_1, \ldots, y_k, x_1, \ldots, x_{\#(q)} \in \mathrm{Var}$, $q \in \mathcal{F}$ and $\sigma(q) = \mathsf{B}$. The connectives $*$ and $\mathrel{-\!\!*}$ are respectively called the *separating conjunction* and *separating implication* (or *magic wand*). The *size* $\mathrm{size}(\varphi)$ and set of *free variables* $\mathrm{Var}(\varphi)$ of an SL^k formula φ are defined as for first-order formulæ, as well as the formulæ $\varphi_1 \vee \varphi_2$, $\varphi_1 \rightarrow \varphi_2$, $\varphi_1 \leftrightarrow \varphi_2$ and $\forall x \,.\, \varphi$. Moreover, we write $\varphi_1 \mathrel{-\!\!\circ} \varphi_2$ for $\neg(\varphi_1 \mathrel{-\!\!*} \neg\varphi_2)$ and call the symbol $\mathrel{-\!\!\circ}$ *septraction* in the following. Throughout the paper, we assume that the arity of the predicate symbols occurring in the SL^k formulæ is bounded by a constant, whereas k is not necessarily bounded.

Definition 2.4. Given a SL^k formula ϕ and a subformula ψ of ϕ, we say that ψ *occurs at polarity* $p \in \{-1, 0, 1\}$ iff one of the following holds:

(1) $\phi = \psi$ and $p = 1$,
(2) $\phi = \neg\phi_1$ and ψ occurs at polarity $-p$ in ϕ_1,
(3) $\phi = \phi_1 \wedge \phi_2$ or $\phi = \phi_1 * \phi_2$, and ψ occurs at polarity p in ϕ_i, for some $i = 1, 2$,
(4) $\phi = \phi_1 \mathrel{-\!\!*} \phi_2$ and either ψ is a subformula of ϕ_1 and $p = 0$, or ψ occurs at polarity p in ϕ_2, or

(5) $\phi = \exists x \,.\, \phi_1$ and ψ occurs at polarity p in ϕ_1.

A polarity of $1, 0$ or -1 is also referred to as positive, neutral or negative, respectively.

Note that our notion of polarity is slightly different than the usual one, because the antecedent of a separating implication is of neutral polarity while the antecedent of an implication is usually of negative polarity. This is meant to strengthen upcoming decidability results (see Remark 3.4).

Semantics SL^k formulæ are interpreted over SL-*structures* $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$, where $\mathfrak{U}$, $\mathfrak{s}$ and $\mathcal{I}$ are defined as for FO and $\mathfrak{h} : \mathfrak{U} \rightharpoonup_{fin} \mathfrak{U}^k$ is a finite partial mapping of $\mathfrak{U}$ to k-tuples of elements of $\mathfrak{U}$, called a *heap*. As for FO, a structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ is *finite* when $||\mathfrak{U}|| \in \mathbb{N}$ and *infinite* otherwise. We denote by $\mathrm{dom}(\mathfrak{h})$ the domain of the heap $\mathfrak{h}$ and by $||\mathfrak{h}|| \in \mathbb{N}$ the cardinality of $\mathrm{dom}(\mathfrak{h})$. A location $\ell \in \mathfrak{U}$ (resp. a variable x) is *allocated* in $\mathcal{S}$ if $\ell \in \mathrm{dom}(\mathfrak{h})$ (resp. if $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$). Two heaps $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are *disjoint* iff $\mathrm{dom}(\mathfrak{h}_1) \cap \mathrm{dom}(\mathfrak{h}_2) = \emptyset$, in which case $\mathfrak{h}_1 \uplus \mathfrak{h}_2$ denotes their union. $\mathfrak{h}'$ is an *extension* of $\mathfrak{h}$ iff $\mathfrak{h}' = \mathfrak{h} \uplus \mathfrak{h}''$, for some heap $\mathfrak{h}''$. The relation $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \varphi$ is defined recursively on the structure of φ, as follows:

$$
\begin{array}{lcl}
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \top & \Leftrightarrow & \text{always} \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \bot & \Leftrightarrow & \text{never} \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models x \approx y & \Leftrightarrow & \mathfrak{s}(x) = \mathfrak{s}(y) \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models q(x_1, \ldots, x_{\#(q)}) & \Leftrightarrow & (\mathfrak{s}(x_1), \ldots, \mathfrak{s}(x_{\#(q)})) \in q^{\mathcal{I}} \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{emp} & \Leftrightarrow & \mathfrak{h} = \emptyset \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models x \mapsto (y_1, \ldots, y_k) & \Leftrightarrow & \mathfrak{h} = \{\langle \mathfrak{s}(x), (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k)) \rangle\} \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \varphi_1 \wedge \varphi_2 & \Leftrightarrow & (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \varphi_i, \text{for all } i = 1, 2 \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\varphi & \Leftrightarrow & (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \not\models \varphi \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \exists x \,.\, \varphi_1 & \Leftrightarrow & \text{there exists } u \in \mathfrak{U} \text{ such that} (\mathfrak{U}, \mathfrak{s}[x \leftarrow u], \mathcal{I}, \mathfrak{h}) \models \varphi_1 \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \varphi_1 * \varphi_2 & \Leftrightarrow & \text{there exist disjoint heaps } h_1, h_2 \text{ such that } h = h_1 \uplus h_2 \\
 & & \text{and } (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models \varphi_i, \text{ for } i = 1, 2 \\
(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \varphi_1 \mathrel{-\!\!*} \varphi_2 & \Leftrightarrow & \text{for all heaps } \mathfrak{h}' \text{ disjoint from } \mathfrak{h} \text{ such that } (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \varphi_1, \\
 & & \text{we have } (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models \varphi_2
\end{array}
$$

Satisfiability, entailment and equivalence are defined for SL^k as for FO formulæ. We write $\phi \equiv^{fin} \psi$ (resp. $\phi \equiv^{inf} \psi$) if ϕ has the same truth value as ψ in all finite (resp. infinite) structures.

Remark 2.5. *The cardinality of the universe has a deep impact on the semantics of* SL *formulæ. For instance, the formula* $\phi = \neg\mathsf{emp} \mathrel{-\!\!*} \bot$ *states that no nonempty heap disjoint from the current heap exists, which is always false in an infinite universe (since every heap is finite) but is true in a finite universe where all elements are allocated.* ■

3 THE BSR(SL^k) CLASS

In this section, we give the definition of the Bernays-Schönfinkel-Ramsey fragment of SL^k and provide a brief summary of the results proved in this paper.

Definition 3.1. The Bernays-Schönfinkel-Ramsey fragment of SL^k, denoted by $\mathrm{BSR}(\mathsf{SL}^k)$, is the set of formulæ of the form $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_m \,.\, \phi$, where ϕ is a quantifier-free SL^k formula.

Note that, since there is no function symbol of sort U in SL^k, there is no restriction, other than the form of the quantifier prefix, defining $\mathrm{BSR}(\mathsf{SL}^k)$. As for FO, we will often restrict ourselves to $\mathrm{BSR}(\mathsf{SL}^k)$ formulæ containing no existential quantifier. As satisfiability is not decidable for $\mathrm{BSR}(\mathsf{SL}^k)$ (see Theorem 3.3 below), we define two fragments of $\mathrm{BSR}(\mathsf{SL}^k)$ for which finite and infinite satisfiability are respectively decidable. The definition is based on the polarity (see Definition 2.4) of the occurrences of the symbol $\mathrel{-\!\!*}$ and on the universal variables occurring within their scope.

Definition 3.2. Given an integer $k \geq 1$, we define:

(1) $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ as the set of formulæ $\forall y_1 \ldots \forall y_m \,.\, \phi$ such that for all $i \in [1, m]$ and all formulæ $\psi_1 \mathrel{-\!\!*} \psi_2$ occurring at polarity 1 in ϕ, we have $y_i \notin \mathrm{Var}(\psi_1) \cup \mathrm{Var}(\psi_2)$,

(2) $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$ as the set of formulæ $\forall y_1 \ldots \forall y_m \,.\, \phi$ such that no formula $\psi_1 \mathrel{-\!\!*} \psi_2$ occurs at polarity 1 in ϕ.

Note that $\mathsf{BSR}^{fin}(\mathsf{SL}^k) \subsetneq \mathsf{BSR}^{inf}(\mathsf{SL}^k) \subsetneq \mathsf{BSR}(\mathsf{SL}^k)$, for any $k \geq 1$. We know state the main results of the paper.

Theorem 3.3. *The satisfiability problem is undecidable for* $\mathsf{BSR}(\mathsf{SL}^k)$. *The infinite satisfiability problem for* $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ *and the finite satisfiability problem for* $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$ *are both PSPACE-complete.*

The remainder of the paper is devoted the proof of Theorem 3.3 (see Theorems 6.1, 6.11 and 6.20).

Remark 3.4. *Because the polarity of the antecedent of a separating implication is neutral, Definition 3.2 imposes no constraint on the occurrences of separating implications at the* left *of an occurrence of* $\mathrel{-\!\!*}$. ■

4 TEST FORMULÆ FOR SL^k

4.1 Definition and Basic Properties

We define a small set of SL^k patterns of formulæ, possibly parameterized by a positive integer, called *test formulæ*. These patterns capture properties related to allocation, points-to relations in the heap and cardinality constraints.

Definition 4.1. The following patterns are called *test formulæ*:

$$x \hookrightarrow \mathbf{y} \stackrel{\text{def}}{=} x \mapsto (y_1, \ldots, y_k) * \top \qquad |U| \geq n \stackrel{\text{def}}{=} \top \mathrel{-\!\!\circ} |h| \geq n$$

$$\mathrm{alloc}(x) \stackrel{\text{def}}{=} x \mapsto \underbrace{(x, \ldots, x)}_{k \text{ times}} \mathrel{-\!\!*} \bot \qquad |h| \geq |U| - n \stackrel{\text{def}}{=} |h| \geq n + 1 \mathrel{-\!\!*} \bot$$

$$x \approx y \qquad q(x_1, \ldots, x_{\#(q)}) \qquad |h| \geq m \stackrel{\text{def}}{=} \begin{cases} |h| \geq m - 1 * \neg\mathrm{emp}, & \text{if } 0 < m < \infty \\ \top, & \text{if } m = 0 \\ \bot, & \text{if } m = \infty \end{cases}$$

where $x, y \in \mathrm{Var}$, $q \in \mathcal{F}$, $\sigma(q) = \mathsf{B}$, $x_1, \ldots, x_{\#(q)}, y_1, \ldots, y_k \in \mathrm{Var}$, $n \in \mathbb{N}$ and $m \in \mathbb{N}_\infty$.

If ϕ is a test formula of the form $t \geq s$ then the formula $\neg\phi$ will often be denoted by $t < s$. For a set of variables $X \subseteq \mathrm{Var}$, let $\mathrm{alloc}(X) \stackrel{\text{def}}{=} \bigwedge_{x \in X} \mathrm{alloc}(x)$ and $\mathrm{nalloc}(X) \stackrel{\text{def}}{=} \bigwedge_{x \in X} \neg\mathrm{alloc}(x)$. The trivial test formulæ $|h| \geq 0$ and $|h| \geq \infty$ are introduced for reasons that will become clear in Section 5. The semantics of test formulæ is very natural: $x \hookrightarrow \mathbf{y}$ means that x points to vector $\mathbf{y}$, $\mathrm{alloc}(x)$ means that x is allocated, and the arithmetic expressions are interpreted as usual, where $|h|$ and $|U|$ respectively denote the number of allocated cells and the number of locations (possibly ∞). Formally:

Proposition 4.2. *Given an* SL*-structure* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$, *the following equivalences hold, for all variables* $x, y_1, \ldots, y_k \in \mathrm{Var}$ *and integers* $n \in \mathbb{N}$:

$$(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models x \hookrightarrow \mathbf{y} \;\Leftrightarrow\; \mathfrak{h}(\mathfrak{s}(x)) = \mathfrak{s}(\mathbf{y}) \qquad (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq |U| - n \;\Leftrightarrow\; ||\mathfrak{h}|| \geq ||\mathfrak{U}|| - n$$

$$(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| \geq n \;\Leftrightarrow\; ||\mathfrak{U}|| \geq n \qquad (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq n \;\Leftrightarrow\; ||\mathfrak{h}|| \geq n$$

$$(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathrm{alloc}(x) \;\Leftrightarrow\; \mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$$

Proof. Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be an SL-structure. We establish each statement separately.

- **$\mathcal{S} \models x \hookrightarrow \mathbf{y} \Leftrightarrow \mathfrak{h}(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$.** Assume that $\mathcal{S} \models x \hookrightarrow \mathbf{y}$. Then by definition, there exist disjoint heaps $\mathfrak{h}_1, \mathfrak{h}_2$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models x \mapsto \mathbf{y}$, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_2) \models \top$ and $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$. Thus $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}_1) \subseteq \mathrm{dom}(\mathfrak{h})$ and $\mathfrak{h}(\mathfrak{s}(x)) = \mathfrak{h}_1(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$. Conversely, assume $\mathfrak{h}(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$. Then $\mathfrak{h}$ is of the form $\mathfrak{h}_1 \uplus \mathfrak{h}_2$, where $\mathfrak{h}_1$ is the restriction of $\mathfrak{h}$ to $\{\mathfrak{s}(x)\}$ and $\mathfrak{h}_2$ is the restriction of $\mathfrak{h}$ to $\mathfrak{U} \setminus \{\mathfrak{s}(x)\}$. By definition, $\mathfrak{h}_1 = \langle \mathfrak{s}(x), (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k)) \rangle$, hence $\mathfrak{h}_1 \models x \mapsto \mathbf{y}$. Furthermore, $\mathfrak{h}_2 \models \top$. Thus $\mathcal{S} \models x \hookrightarrow \mathbf{y}$.

- $\mathcal{S} \models \mathbf{alloc}(x) \Leftrightarrow \mathfrak{s}(x) \in \mathbf{dom}(\mathfrak{h})$. Assume that $\mathcal{S} \models \mathrm{alloc}(x)$. This means that there is no heap $\mathfrak{h}'$ disjoint from $\mathfrak{h}$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models x \mapsto (x, \ldots, x)$. If $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$, then the heap $\mathfrak{h}'$ defined as $\mathfrak{h}' = \langle \mathfrak{s}(x), (\mathfrak{s}(x), \ldots, \mathfrak{s}(x)) \rangle$ is disjoint from $\mathfrak{h}$ and we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models x \mapsto (x, \ldots, x)$. Thus $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$. Conversely, assume $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$. By definition, for any heap $\mathfrak{h}'$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models x \mapsto (x, \ldots, x)$ we have $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}')$, hence $\mathfrak{h}' \cap \mathfrak{h} \neq \emptyset$. Thus $\mathcal{S} \models \mathrm{alloc}(x)$.
- $\mathcal{S} \models \mathbf{|h| \geq n} \Leftrightarrow \mathbf{||\mathfrak{h}|| \geq n}$. Assume that $\mathcal{S} \models |h| \geq n$. Then since $\mathfrak{h}$ has a finite domain, it is clear that $||\mathfrak{h}|| \geq n$ if $n = 0$ and that no such structure exists if $n = \infty$. When $n \geq 1$, we prove the result by induction on n. By definition, $\mathcal{S} \models |h| \geq n - 1 * \neg\mathrm{emp}$, hence there exist disjoint heaps $\mathfrak{h}_1, \mathfrak{h}_2$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models |h| \geq n - 1$, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_2) \models \neg\mathrm{emp}$ and $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$. By the induction hypothesis $||\mathfrak{h}_1|| \geq n - 1$ and by definition, $||\mathfrak{h}_2|| \geq 1$, so that $||\mathfrak{h}_1 \uplus \mathfrak{h}_2|| \geq n$. Conversely, assume that $||\mathfrak{h}|| \geq n$. Since $\mathfrak{h}$ is finite, this entails that $n \neq \infty$. If $n = 0$ then $\mathcal{S} \models |h| \geq n$ always holds. Otherwise, we prove the result by induction on n. Consider $\ell \in \mathrm{dom}(\mathfrak{h})$ and let $\mathfrak{h}_1$ and $\mathfrak{h}_2$ respectively denote the restrictions of $\mathfrak{h}$ to $\mathfrak{U} \setminus \{\ell\}$ and to $\{\ell\}$, so that $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$. Since $||\mathfrak{h}_1|| \geq n - 1$, by the induction hypothesis $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models |h| \geq n - 1$, and since $\mathrm{dom}(\mathfrak{h}_2) \neq \emptyset$, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_2) \models \neg\mathrm{emp}$. Thus $\mathcal{S} \models |h| \geq n$.
- $\mathcal{S} \models \mathbf{|U| \geq n} \Leftrightarrow \mathbf{||\mathfrak{U}|| \geq n}$. Assume that $\mathcal{S} \models |U| \geq n$. Then there exists a heap $\mathfrak{h}_1$ disjoint from $\mathfrak{h}$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}_1) \models |h| \geq n$. This entails that $||\mathfrak{h} \uplus \mathfrak{h}_1|| \geq n$ and since $\mathrm{dom}(\mathfrak{h} \uplus \mathfrak{h}_1) \subseteq \mathfrak{U}$, necessarily, $||\mathfrak{U}|| \geq n$. Conversely, if $||\mathfrak{U}|| \geq n$, then there exists a set $L \subseteq \mathfrak{U}$ such that $\mathrm{dom}(\mathfrak{h}) \cap L = \emptyset$ and $||L|| = n - ||\mathfrak{h}||$. Let $\mathfrak{h}'$ be any heap of domain L. Then $\mathfrak{h}$ and $\mathfrak{h}'$ are disjoint and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models |h| \geq n$, which proves that $\mathcal{S} \models |U| \geq n$.
- $\mathcal{S} \models \mathbf{|h| \geq |U| - n} \Leftrightarrow \mathbf{||\mathfrak{h}|| \geq ||\mathfrak{U}|| - n}$.. Assume that $\mathcal{S} \models |h| \geq |U| - n$. By definition, this entails that there is no heap disjoint from $\mathfrak{h}$ with a domain of cardinality at least $n + 1$. In particular, if $L = \mathfrak{U} \setminus \mathrm{dom}(\mathfrak{h})$, and $\mathfrak{h}'$ is any heap of domain L, then $\mathrm{dom}(\mathfrak{h}) \cap \mathrm{dom}(\mathfrak{h}') = \emptyset$, hence $||\mathfrak{h}'|| \leq n$. Since $||\mathfrak{U}|| = ||\mathfrak{h}|| + ||\mathfrak{h}'||$, we deduce that $||\mathfrak{h}|| \geq ||\mathfrak{U}|| - n$. Conversely, if $||\mathfrak{h}|| \geq ||\mathfrak{U}|| - n$ then $||\mathfrak{U} \setminus \mathrm{dom}(\mathfrak{h})|| \leq n$, hence there is no heap disjoint from $\mathfrak{h}$ with a domain of cardinality at least $n + 1$, so that $\mathcal{S} \models |h| \geq |U| - n$.

□

Not all atoms of SL^k are test formulæ, for instance $x \mapsto \mathbf{y}$ and emp are not test formulæ. However, by Proposition 4.2, we have the equivalences $x \mapsto \mathbf{y} \equiv x \hookrightarrow \mathbf{y} \wedge \neg|h| \geq 2$ and $\mathrm{emp} \equiv \neg|h| \geq 1$. Note that, for any $n \in \mathbb{N}$, the test formulæ $|U| \geq n$ and $|h| \geq |U| - n$ are trivially true and false respectively, if the universe is infinite.

4.2 A Generalization of Test Formulæ

For technical convenience, we extend the previous patterns to express more general cardinality constraints. For every $n \in \mathbb{N}$, we denote by $|U| \simeq n$ (resp., $|h| \simeq n$) the formula $|U| \geq n \wedge |U| < n + 1$ (resp., $|h| \geq n \wedge |h| < n + 1$). Similarly, $|h| \simeq |U| - n$ denotes either $|h| \geq |U| - n \wedge |h| < |U| - (n - 1)$ (if $n > 0$) or $|h| \geq |U| - 0$ (if $n = 0$). We then extend the notation $|h| \geq t$ to the case where t is a linear function of $|U|$, with coefficients in $\mathbb{Z}$.

Definition 4.3. Given integers $\alpha, \beta \in \mathbb{Z}$, where $\alpha \notin \{0, 1\}$, let

$$|h| \geq \alpha \cdot |U| + \beta \stackrel{\text{def}}{=} \begin{cases} \bot & \text{if } \alpha > 1,\ \beta > 0 \\ \top & \text{if } \alpha, \beta < 0 \\ |U| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil \wedge \bigwedge_{1 \leq n \leq \left\lfloor \frac{-\beta}{\alpha-1} \right\rfloor} (|U| \simeq n \rightarrow |h| \geq \alpha \cdot n + \beta) & \text{if } \alpha > 1,\ \beta \leq 0 \\ \bigwedge_{1 \leq n < \left\lfloor \frac{-\beta}{\alpha} \right\rfloor} (|U| \simeq n \rightarrow |h| \geq \alpha \cdot n + \beta) & \text{if } \alpha < 0,\ \beta \geq 0 \end{cases}$$

If $\alpha = 0$ and $\beta < 0$ then $|h| \geq \alpha.|U| + \beta \stackrel{\text{def}}{=} \top$. If $\alpha = 1$ and $\beta > 0$ then $|h| \geq \alpha.|U| + \beta \stackrel{\text{def}}{=} \bot$.

Note that the cases $\alpha = 0, \beta \geq 0$ and $\alpha = 1, \beta \leq 0$ are already covered by Definition 4.1. The following proposition states that the semantics of these formulæ is as expected.

Proposition 4.4. *Given an* SL*-structure* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$*, we have* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot |U| + \beta$ *iff* $||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$*, for all* $\alpha, \beta \in \mathbb{Z}$*,* $\alpha \notin \{0, 1\}$*.*

Proof. We distinguish the four cases below:

- If $\alpha > 1$ and $\beta > 0$ then $||\mathfrak{U}|| \geq ||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$ never holds.
- If $\alpha < 0$ and $\beta < 0$ then $||\mathfrak{h}|| \geq 0 \geq \alpha \cdot ||\mathfrak{U}|| + \beta$, always holds.
- If $\alpha > 1$ and $\beta \leq 0$, assume first that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot |U| + \beta$. Then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$, thus $1 \leq ||\mathfrak{U}|| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$ by Proposition 4.2. If $||\mathfrak{U}|| > \left\lfloor \frac{-\beta}{\alpha-1} \right\rfloor$ then $||\mathfrak{U}|| \geq \left\lfloor \frac{-\beta}{\alpha-1} \right\rfloor + 1 = \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$, which contradicts $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$, by Proposition 4.2. Otherwise, we have $||\mathfrak{U}|| = n$, with $1 \leq n \leq \left\lfloor \frac{-\beta}{\alpha-1} \right\rfloor$. In this case $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot n + \beta$, which implies $||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$, by Proposition 4.2. Conversely, assume that $||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$. Since necessarily $||\mathfrak{U}|| \geq ||\mathfrak{h}||$, we obtain $||\mathfrak{U}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$, i.e., $||\mathfrak{U}|| > \alpha \cdot ||\mathfrak{U}|| + \beta - 1$ and thus $||\mathfrak{U}|| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$, so that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| < \left\lceil \frac{1-\beta}{\alpha-1} \right\rceil$. Moreover, if $n = ||U||$ then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot n + \beta$ by Proposition 4.2.
- If $\alpha < 0$ and $\beta \geq 0$, assume first that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot |U| + \beta$. If, moreover, $||\mathfrak{U}|| \geq \frac{-\beta}{\alpha}$, then $\alpha \cdot ||\mathfrak{U}|| + \beta \leq 0$, thus $||\mathfrak{h}|| \geq 0 \geq \alpha \cdot ||\mathfrak{U}|| + \beta$ holds. Otherwise, $1 \leq ||\mathfrak{U}|| < \left\lfloor \frac{-\beta}{\alpha} \right\rfloor$ and if $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| \simeq n$, for some $1 \leq n < \left\lfloor \frac{-\beta}{\alpha} \right\rfloor$, then we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot n + \beta$, thus $||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$, by Proposition 4.2. Conversely, assume that $||\mathfrak{h}|| \geq \alpha \cdot ||\mathfrak{U}|| + \beta$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |U| \simeq n$, for some integer $1 \leq n < \left\lfloor \frac{-\beta}{\alpha} \right\rfloor$. By Proposition 4.2, we have $||\mathfrak{U}|| = n$ and $||\mathfrak{h}|| \geq \alpha \cdot n + \beta$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| \geq \alpha \cdot |U| + \beta$.

□

4.3 From Test formulæ to FO

The introduction of test formulæ (Definition 4.1) is motivated by the reduction of the (in)finite satisfiability problem for quantified boolean combinations thereof to the same problem for FO. The reduction is based on a straightforward encoding of the heap as a $(k+1)$-ary predicate symbol, however it is devised below in such a way that the obtained formula is in the BSR class, if possible. To this purpose, we also use a monadic predicate symbol encoding the domain of the heap and boolean constants encoding cardinality constraints. We thus introduce several special (pairwise distinct) function symbols: a $(k+1)$-ary predicate $\mathfrak{p}$, a monadic predicate $\mathfrak{d}$, boolean constants $\mathfrak{a}_n$, $\mathfrak{b}_n$ and $\mathfrak{c}_n$, and the following constants of sort $\mathfrak{U}$: $\mathfrak{u}_n$, $\mathfrak{u}_n^i$, $\mathfrak{v}_n$, $\mathfrak{w}_n$ and ξ_x^i, for $n \geq 0$, $i \in [1, k]$ and $x \in \mathrm{Var}$. The symbol $\mathfrak{p}$ will encode the heap, $\mathfrak{d}$ will encode the domain of the heap, the constants $\mathfrak{a}_n$, $\mathfrak{b}_n$ and $\mathfrak{c}_n$ encode the constraints over the number of (allocated or unallocated) locations, and $\mathfrak{u}_n$, $\mathfrak{u}_n^i$, $\mathfrak{v}_n$, $\mathfrak{w}_n$ and ξ_x^i are interpreted as pairwise distinct elements of the universe, used to express such constraints in FO.

Given a quantified boolean combination of test formulæ ϕ not containing the above symbols, the FO formula $\tau(\phi)$ is defined by induction on the structure of ϕ:

$$\begin{array}{rcl@{\qquad}rcl}
\tau(|h| \geq n) & \stackrel{\text{def}}{=} & \mathfrak{a}_n & \tau(|U| \geq n) & \stackrel{\text{def}}{=} & \mathfrak{b}_n \\
\tau(|h| \geq |U| - n) & \stackrel{\text{def}}{=} & \neg \mathfrak{c}_{n+1} & \tau(\neg \phi_1) & \stackrel{\text{def}}{=} & \neg \tau(\phi_1) \\
\tau(x \hookrightarrow \mathbf{y}) & \stackrel{\text{def}}{=} & \mathfrak{p}(x, y_1, \ldots, y_k) & \tau(\mathrm{alloc}(x)) & \stackrel{\text{def}}{=} & \mathfrak{d}(x) \\
\tau(\phi_1 \wedge \phi_2) & \stackrel{\text{def}}{=} & \tau(\phi_1) \wedge \tau(\phi_2) & \tau(\exists x \,.\, \phi_1) & \stackrel{\text{def}}{=} & \exists x \,.\, \tau(\phi_1) \\
\tau(q(x_1, \ldots, x_{\#(q)})) & \stackrel{\text{def}}{=} & q(x_1, \ldots, x_{\#(q)}) & \tau(x \approx y) & \stackrel{\text{def}}{=} & x \approx y
\end{array}$$

The special symbols are related by the following axioms:

$$\begin{array}{ll@{\qquad}l}
(\mathit{Heap}) & \forall x \forall \mathbf{y} \forall \mathbf{y}' \,.\, \mathfrak{p}(x, \mathbf{y}) \wedge \mathfrak{p}(x, \mathbf{y}') \rightarrow \mathbf{y} \approx \mathbf{y}' & \\
(\mathit{Dom}_x) & \mathfrak{d}(x) \rightarrow \mathfrak{p}(x, \xi_x^1, \ldots, \xi_x^k) & \text{for each } x \in \mathrm{Var}(\phi)
\end{array}$$

$$(A_0)\quad \mathfrak{a}_0 \qquad (A_n)\quad \left\{ \begin{array}{l} \mathfrak{a}_n \rightarrow (\mathfrak{a}_{n-1} \wedge \mathfrak{p}(\mathfrak{u}_n, \mathfrak{u}_n^1, \ldots, \mathfrak{u}_n^k) \wedge \bigwedge_{i=1}^{n-1} \neg \mathfrak{u}_i \approx \mathfrak{u}_n) \\ \wedge\ \forall x \forall \mathbf{y} \,.\, \neg \mathfrak{a}_n \wedge \mathfrak{p}(x, \mathbf{y}) \rightarrow \bigvee_{i=1}^{n-1} x \approx \mathfrak{u}_i \end{array} \right\}$$

$$(B_0)\quad \mathfrak{b}_0 \qquad (B_n)\quad \left\{ \begin{array}{l} \mathfrak{b}_n \rightarrow (\mathfrak{b}_{n-1} \wedge \bigwedge_{i=1}^{n-1} \neg \mathfrak{v}_i \approx \mathfrak{v}_n) \\ \wedge\ \forall x \,.\, \neg \mathfrak{b}_n \rightarrow \bigvee_{i=1}^{n-1} x \approx \mathfrak{v}_i \end{array} \right\}$$

$$(C_0)\quad \mathfrak{c}_0 \qquad (C_n)\quad \forall \mathbf{y} \,.\, \mathfrak{c}_n \rightarrow (\mathfrak{c}_{n-1} \wedge \neg \mathfrak{p}(\mathfrak{w}_n, \mathbf{y}) \wedge \bigwedge_{i=1}^{n-1} \neg \mathfrak{w}_n \approx \mathfrak{w}_i)$$

Intuitively, $\mathfrak{p}$ encodes the heap in the following sense. If $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$ then there exists a heap $\mathfrak{h}$ on $\mathfrak{U}$ such that $\mathbf{y} = \mathfrak{h}(x) \Leftrightarrow (x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$. The constant $\mathfrak{a}_n$ (resp. $\mathfrak{b}_n$) is true if there are at least n cells in the domain of the heap (resp. in the universe), namely $\mathfrak{u}_1, \ldots, \mathfrak{u}_n$ (resp. $\mathfrak{v}_1, \ldots, \mathfrak{v}_n$). If $\mathfrak{c}_n$ is true, then there are at least n locations $\mathfrak{w}_1, \ldots, \mathfrak{w}_n$ outside of the domain of the heap (i.e., n unallocated locations), but the converse does not hold. Indeed, the axioms C_n do not state the equivalence of $\mathfrak{c}_n$ with the existence of at least n free locations, because such an equivalence cannot be expressed in BSR(FO)[5]. Similarly, the axiom *Dom* states that if x is allocated then $\mathfrak{d}(x)$ holds, but the converse is true only for $x \in \mathrm{Var}(\phi)$ (as stated by the axiom Dom_x). Again, adding the implication $\forall x \,.\, \mathfrak{d}(x) \rightarrow \exists y_1, \ldots, y_k \,.\, \mathfrak{p}(x, y_1, \ldots, y_k)$ would result in a formula that is not in BSR(FO). Instead, we only assert finitely many (skolemized) instances of the latter formula, for every free variable x, which is sufficient for our purpose. As a consequence, the transformation preserves sat-equivalence only if the formulæ $|h| \geq |U| - n$ or $\mathrm{alloc}(x)$ with $x \notin \mathrm{Var}(\phi)$ occur only at negative polarity (see Lemma 4.9, Point 2).

Definition 4.5. Given a structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$ and a heap $\mathfrak{h}$ on $\mathfrak{U}$, if $x = \mathfrak{h}(\mathbf{y}) \Leftrightarrow (x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$, then we say that $\mathfrak{h}$ is *associated with* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$. An element $x \in \mathfrak{U}$ is *allocated in* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ (resp. *points to* $\mathbf{y}$ *in* $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$) if there exists $\mathbf{y} \in \mathfrak{U}^k$ such that $(x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$ (resp. if $(x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$).

Definition 4.6. For a quantified boolean combination of test formulæ ϕ, we let $\mathcal{N}(\phi)$ be the maximum integer n occurring in a test formula θ of the form $|h| \geq n$, $|U| \geq n$, or $|h| \geq |U| - n$ from ϕ and define $\mathcal{A}(\phi) \stackrel{\text{def}}{=} \mathit{Heap} \wedge \bigwedge_{i=0}^{\mathcal{N}(\phi)} A_i \wedge \bigwedge_{i=0}^{\mathcal{N}(\phi)} B_i \wedge \bigwedge_{i=0}^{\mathcal{N}(\phi)+1} C_i \wedge \mathit{Dom} \wedge \bigwedge_{x \in \mathrm{Var}(\phi)} \mathit{Dom}_x$ as the conjunction of axioms related to ϕ.

Example 4.7. Let ϕ be the SL^1 formula: $x \hookrightarrow y \wedge |h| \geq 2 \wedge |h| < |U|$. Then $\tau(\phi) = \mathfrak{p}(x, y) \wedge \mathfrak{a}_2 \wedge \mathfrak{c}_1$, and $\mathcal{A}(\phi)$ contains, among others, the following formulæ[6]:

$$\forall x, y, z \,.\, \mathfrak{p}(x, y) \wedge \mathfrak{p}(x, z) \rightarrow y \approx z$$
$$\mathfrak{a}_0 \wedge (\mathfrak{a}_1 \rightarrow \mathfrak{a}_0 \wedge \mathfrak{p}(\mathfrak{u}_1, \mathfrak{u}_1^1)) \wedge (\mathfrak{a}_2 \rightarrow \mathfrak{a}_1 \wedge \mathfrak{p}(\mathfrak{u}_2, \mathfrak{u}_2^1) \wedge \neg \mathfrak{u}_1 \approx \mathfrak{u}_2)$$
$$\mathfrak{c}_0 \wedge (\forall y \,.\, \mathfrak{c}_1 \rightarrow \mathfrak{c}_0 \wedge \neg \mathfrak{p}(\mathfrak{w}_1, y))$$

The formula $\tau(\phi) \wedge \mathcal{A}(\phi)$ states that $\mathfrak{p}(x, y)$ holds, that $\mathfrak{p}$ is a partial function and that there exist at least two distinct allocated elements (namely $\mathfrak{u}_1$ and $\mathfrak{u}_2$) and one unallocated element ($\mathfrak{w}_1$).

Let ϕ' be the SL^2 formula $\mathrm{alloc}(u) \wedge \forall y \,.\, \neg u \approx y \rightarrow \neg \mathrm{alloc}(y)$. Then $\tau(\phi') = \mathfrak{d}(u) \wedge (\forall y \,.\, \neg u \approx y \rightarrow \neg \mathfrak{d}(y))$, where the relevant axioms in $\mathcal{A}(\phi')$ are:

$$\forall x, y_1, y_2, z_1, z_2 \,.\, \mathfrak{p}(x, y_1, y_2) \wedge \mathfrak{p}(x, z_1, z_2) \rightarrow y_1 \approx z_1 \wedge y_2 \approx z_2$$
$$\forall x, y_1, y_2 \,.\, \mathfrak{p}(x, y_1, y_2) \rightarrow \mathfrak{d}(x)$$
$$\mathfrak{d}(u) \rightarrow \mathfrak{p}(u, \xi_u^1, \xi_u^2)$$

■

The relationship between ϕ and $\tau(\phi)$ is stated below.

[5]The converse of C_n: $\forall x \,.\, (\neg \mathfrak{c}_n \wedge \forall \mathbf{y} \,.\, \neg \mathfrak{p}(x, \mathbf{y})) \rightarrow \bigvee_{i=1}^{n-1} x \approx \mathfrak{w}_i$ is not in BSR(FO).

[6]For simplicity, only the relevant axioms are given.

Definition 4.8. A formula ϕ is BSR-*compatible* if: (i) each test formula $|h| \geq |U| - n$ in ϕ occurs at a negative polarity (ii) if a formula $\mathrm{alloc}(x)$ occurs at positive polarity in ϕ, then $x \in \mathrm{Var}(\phi)$.

Lemma 4.9. *Let ϕ be a quantified boolean combination of test formulæ. The following hold, for any universe $\mathfrak{U}$ and any store $\mathfrak{s}$:*

(1) if $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \phi$, for a heap $\mathfrak{h}$, then $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\phi) \wedge \mathcal{A}(\phi)$ for an interpretation $\mathcal{J}$ coinciding with $\mathcal{I}$ on every symbol not occurring in $\mathcal{A}(\phi)$ and such that $\mathfrak{h}$ is associated with $(\mathfrak{U}, \mathfrak{s}, \mathcal{J})$;

(2) if ϕ is BSR-*compatible and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\phi) \wedge \mathcal{A}(\phi)$ for an interpretation $\mathcal{I}$ such that $||\mathfrak{p}^{\mathcal{I}}|| \in \mathbb{N}$, then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \phi$, where $\mathfrak{h}$ denotes the heap associated with $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$.*

Proof. (1) Let $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a model of ϕ. Let $\mathcal{J}$ be an interpretation coinciding with $\mathcal{I}$ on every symbol occurring in ϕ, and extended to the symbols $\mathfrak{p}, \mathfrak{a}_i, \mathfrak{b}_i, \mathfrak{c}_j, \mathfrak{u}_i, \mathfrak{v}_i, \mathfrak{w}_i$, for $i \in [0, \mathcal{N}(\phi)]$ and $j \in [0, \mathcal{N}(\phi)+1]$, as follows: for all $\ell_0, \ldots, \ell_k \in \mathfrak{U}$ we set $(\ell_0, \ldots, \ell_k) \in \mathfrak{p}^{\mathcal{J}}$ iff $\mathfrak{h}(\ell_0) = (\ell_1, \ldots, \ell_k)$ and $\mathfrak{d}^{\mathcal{I}} = \mathrm{dom}(\mathfrak{h})$. The interpretation of the boolean constants is defined below:

$$\mathfrak{a}_i^{\mathcal{J}} \stackrel{\text{def}}{=} \begin{cases} \top & \text{if } 0 \leq i \leq \min(||\mathfrak{h}||, \mathcal{N}(\phi)) \\ \bot & \text{if } i > \min(||\mathfrak{h}||, \mathcal{N}(\phi)) \end{cases}$$
$$\mathfrak{b}_i^{\mathcal{J}} \stackrel{\text{def}}{=} \begin{cases} \top & \text{if } 0 \leq i \leq \min(||\mathfrak{U}||, \mathcal{N}(\phi)) \\ \bot & \text{if } i > \min(||\mathfrak{U}||, \mathcal{N}(\phi)) \end{cases}$$
$$\mathfrak{c}_i^{\mathcal{J}} \stackrel{\text{def}}{=} \begin{cases} \top & \text{if } 0 \leq i \leq \min(||\mathfrak{U}|| - ||\mathfrak{h}||, \mathcal{N}(\phi) + 1) \\ \bot & \text{if } i > \min(||\mathfrak{U}|| - ||\mathfrak{h}||, \mathcal{N}(\phi) + 1) \end{cases}$$

The constants of sort U are interpreted as locations, as follows:

- $\mathfrak{u}_1^{\mathcal{J}}, \ldots, \mathfrak{u}_{\min(||\mathfrak{h}||, \mathcal{N}(\phi))}^{\mathcal{J}}$ are pairwise distinct locations in $\mathrm{dom}(\mathfrak{h})$ and $\mathfrak{u}_n^i$ is the i-th component of the vector referred to by $\mathfrak{u}_n$.
- $\mathfrak{v}_1^{\mathcal{J}}, \ldots, \mathfrak{v}_{\min(||\mathfrak{U}||, \mathcal{N}(\phi))}^{\mathcal{J}}$ are pairwise distinct locations in $\mathfrak{U}$.
- $\mathfrak{w}_1^{\mathcal{J}}, \ldots, \mathfrak{w}_{\min(||\mathfrak{U}|| - ||\mathfrak{h}||, \mathcal{N}(\phi)+1)}^{\mathcal{J}}$ are pairwise distinct locations in $\mathfrak{U} \setminus \mathrm{dom}(\mathfrak{h})$.

The other symbols are interpreted arbitrarily. It is straightforward to check that $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \mathcal{A}(\phi)$. We prove that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\psi)$ for every subformula ψ of ϕ by induction on the structure of ϕ:

- $\psi = x \approx y$: We have $\tau(\psi) = \psi$. Further, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi \Leftrightarrow \mathfrak{s}(x) = \mathfrak{s}(y) \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \psi$.
- $\psi = q(x_1, \ldots, x_{\#(q)})$: We have $\tau(\psi) = \psi$. Moreover, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi \Leftrightarrow (\mathfrak{s}(x_1), \ldots, \mathfrak{s}(x_{\#(q)})) \in q^{\mathcal{I}}$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \psi \Leftrightarrow (\mathfrak{s}(x_1), \ldots, \mathfrak{s}(x_{\#(q)})) \in q^{\mathcal{J}}$. Because $\mathcal{I}$ and $\mathcal{J}$ coincide on every symbol occurring in ψ, $q^{\mathcal{I}} = q^{\mathcal{J}}$. Thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \psi$.
- $\psi = |h| \geq n$: $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $||\mathfrak{h}|| \geq n$ by Proposition 4.2. Since $n \leq \mathcal{N}(\psi)$, we have $||\mathfrak{h}|| \geq n \Leftrightarrow n \leq \min(||\mathfrak{h}||, \mathcal{N}(\psi)) \Leftrightarrow \mathfrak{a}_n^{\mathcal{J}} = \top \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\psi)$.
- $\psi = |U| \geq n$: $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $||\mathfrak{U}|| \geq n$, by Proposition 4.2. Since $n \leq \mathcal{N}(\psi)$, we have $||\mathfrak{U}|| \geq n \Leftrightarrow n \leq \min(||\mathfrak{U}||, \mathcal{N}(\psi)) \Leftrightarrow \mathfrak{b}_n^{\mathcal{I}} = \top \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\psi)$.
- $\psi = |h| \geq |U| - n$: $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $||\mathfrak{h}|| \geq ||\mathfrak{U}|| - n$, by Proposition 4.2, i.e., iff $n + 1 > ||\mathfrak{U}|| - ||\mathfrak{h}||$. Since $n \leq \mathcal{N}(\psi)$, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi \Leftrightarrow n + 1 > \min(||\mathfrak{U}|| - ||\mathfrak{h}||, \mathcal{N}(\psi) + 1) \Leftrightarrow \mathfrak{c}_{n+1}^{\mathcal{I}} = \bot \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \neg c_{n+1} \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\psi)$.
- $\psi = x \hookrightarrow (y_1, \ldots, y_k)$: $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $\mathfrak{h}(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$ iff $(\mathfrak{s}(x), \mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k)) \in \mathfrak{p}^{\mathcal{J}}$ iff $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \mathfrak{p}(x, y_1, \ldots, y_k)$.
- $\psi = \mathrm{alloc}(x)$: $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ iff $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$ iff $\mathfrak{s}(x) \in \mathfrak{d}^{\mathcal{I}}$ iff $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \mathfrak{d}(x)$.
- The cases $\psi = \psi_1 \wedge \psi_2$, $\psi = \neg\psi_1$ and $\psi = \exists x \,.\, \psi_1$ are by the inductive hypothesis, since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi_i \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\psi_i)$, for all $i = 1, 2$.

(2) Let $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be a model of $\tau(\phi) \wedge \mathcal{A}(\phi)$, such that $||\mathfrak{p}^{\mathcal{I}}|| \in \mathbb{N}$. We define a heap $\mathfrak{h}$ as follows: for each $(k+1)$-tuple of locations $\ell_0, \ldots, \ell_k \in \mathfrak{U}$ such that $(\ell_0, \ldots, \ell_k) \in \mathfrak{p}^{\mathcal{I}}$, we set $\mathfrak{h}(\ell_0) \stackrel{\text{def}}{=} (\ell_1, \ldots, \ell_k)$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$

and $||\mathfrak{p}^{\mathcal{I}}|| \in \mathbb{N}$, $\mathfrak{h}$ is a finite partial function. Let ϕ_{nnf} be the negation normal form of ϕ. It is easy to check that $\tau(\phi_{nnf}) \equiv \tau(\phi)$. We prove that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\psi) \Rightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ for every subformula ψ in ϕ_{nnf}:

- $\psi = |h| \geq n$: $\tau(\psi) = \mathfrak{a}_n$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{a}_n \Rightarrow \mathfrak{a}_n^{\mathcal{I}} = \top$. Since $n \leq \mathcal{N}(\psi)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \bigwedge_{i=0}^{\mathcal{N}(\psi)} A_j$, we have $\mathfrak{a}_j^{\mathcal{I}} = \top$ and $\mathfrak{u}_j^{\mathcal{I}} \in \mathrm{dom}(\mathfrak{h})$, for all $j \in [1, n]$. Because $\mathfrak{u}_j^{\mathcal{I}}$ are pairwise distinct, for $j \in [1, n]$, we obtain that $||\mathfrak{h}|| \geq n$, and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ follows, by Proposition 4.2.
- $\psi = |h| < n$: $\tau(\psi) = \neg\mathfrak{a}_n$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \neg\mathfrak{a}_n \Rightarrow \mathfrak{a}_n^{\mathcal{I}} = \bot$. Since $n \leq \mathcal{N}(\psi)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \bigwedge_{i=0}^{\mathcal{N}(\psi)} A_j$, each location $\ell \in \mathrm{dom}(\mathfrak{h})$ must be one of $\mathfrak{u}_1^{\mathcal{I}}, \ldots, \mathfrak{u}_{n-1}^{\mathcal{I}}$, thus $||\mathrm{dom}(\mathfrak{h})|| \leq n-1$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models |h| < n$ follows, by Proposition 4.2.
- $\psi = |U| \geq n$: $\tau(\psi) = \mathfrak{b}_n$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{b}_n \Rightarrow \mathfrak{b}_n^{\mathcal{I}} = \top$. Since $n \leq \mathcal{N}(\psi)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \bigwedge_{i=0}^{\mathcal{N}(\psi)} B_j$, we have $\mathfrak{b}_j^{\mathcal{I}} = \top$, for all $j \in [1, n]$. Because $\mathfrak{v}_j^{\mathcal{I}}$ are pairwise distinct, for all $j \in [1, n]$, we obtain that $||\mathfrak{U}|| \geq n$, and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ follows, by Proposition 4.2.
- $\psi = |U| < n$: $\tau(\psi) = \neg\mathfrak{b}_n$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \neg\mathfrak{b}_n \Rightarrow \mathfrak{b}_n^{\mathcal{I}} = \bot$. Since $n \leq \mathcal{N}(\psi)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \bigwedge_{i=0}^{\mathcal{N}(\psi)} B_j$, we have that each location $\ell \in \mathfrak{U}$ must be one of $\mathfrak{v}_1^{\mathcal{I}}, \ldots, \mathfrak{v}_{n-1}^{\mathcal{I}}$, thus $||\mathfrak{U}|| \leq n-1$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$ follows, by Proposition 4.2.
- $\psi = |h| \geq |U| - n$: this case is impossible because $|h| \geq |U| - n$ must occur at a negative polarity in ψ.
- $\psi = |h| < |U| - n$: $\tau(\psi) = \mathfrak{c}_{n+1}$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{c}_{n+1} \Rightarrow \mathfrak{c}_{n+1} = \top$. Since $n \leq \mathcal{N}(\psi)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \bigwedge_{i=0}^{\mathcal{N}(\psi)+1} C_j$, we obtain that $\mathfrak{w}_j^{\mathcal{I}} \in \mathfrak{U} \setminus \mathrm{dom}(\mathfrak{h})$, for all $j \in [1, n+1]$. Since $\mathfrak{w}_j^{\mathcal{I}}$ are pairwise disjoint, we obtain $||U|| - ||\mathfrak{h}|| \geq n+1$ thus $(\mathfrak{U}, \mathfrak{s}, \mathfrak{h}) \models \psi$ follows, by Proposition 4.2.
- $\psi = \mathrm{alloc}(x)$. Since ψ occurs at positive polarity and ϕ is BSR-compatible, necessarily $x \in \mathrm{Var}(\phi)$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{d}(x)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Dom}_x$, we must have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{p}(x, \xi_x^1, \ldots, \xi_x^k)$, and therefore $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$. Thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$.
- $\psi = \neg\mathrm{alloc}(x)$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathfrak{d}(x)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Dom}$, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \forall y_1, \ldots, y_k \,.\, \neg\mathfrak{p}(x, y_1, \ldots, y_k)$, thus $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$. Hence $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi$.
- $\psi \in \{x \approx y, \neg x \approx y, q(\mathbf{x}), \neg q(\mathbf{x}), x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y}\}$: The equivalence statement $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi \Leftrightarrow (\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \psi$ is proven in the same way as for point (1).
- The cases $\psi = \psi_1 \wedge \psi_2$, $\psi = \psi_1 \vee \psi_2$, $\exists x \,.\, \psi_1$ are by inductive hypothesis.

□

The following proposition states essential syntactic properties of $\tau(\varphi) \wedge \mathcal{A}(\varphi)$.

Proposition 4.10. *Let $\varphi = \forall \mathbf{y} \,.\, \phi$, where ϕ is a boolean combination of test formulæ, with* $\mathrm{Var}(\varphi) = \{x_1, \ldots, x_n\}$. *The formula $\tau(\varphi) \wedge \mathcal{A}(\varphi)$ is a* BSR(FO) *formula with no existential quantifier such that* $||\mathrm{Const}(\tau(\varphi) \wedge \mathcal{A}(\varphi))|| = k \cdot n + (k+6) \cdot \mathcal{N}(\varphi) + 5$ *(where k denotes the number of record fields) and* $\mathrm{Var}(\tau(\varphi) \wedge \mathcal{A}(\varphi)) = \mathrm{Var}(\varphi)$.

Proof. The proof is by a straightforward inspection of $\tau(\varphi)$ and of the axioms in $\mathcal{A}(\varphi)$. There are $k \cdot n$ constants $\xi_{x_i}^j$, $\mathcal{N}(\varphi)+1$ constants $\mathfrak{a}_i, \mathfrak{b}_i$ and $\mathfrak{w}_i$, $\mathcal{N}(\varphi)$ constants $\mathfrak{u}_i, \mathfrak{v}_i$, $\mathcal{N}(\varphi)+2$ constants $\mathfrak{c}_i$, and $k \cdot \mathcal{N}(\varphi)$ constants $\mathfrak{u}_i^j$. □

5 FROM QUANTIFIER-FREE SL^k TO TEST FORMULÆ

This section establishes the expressive completeness result of the paper, namely that any quantifier-free SL^k formula is equivalent, on both finite and infinite models, to a boolean combination of test formulæ. Starting from a quantifier-free SL^k formula φ, we define a set $\mu(\varphi)$ of conjunctions of test formulæ and their negations, called *minterms*, such that $\varphi \equiv \bigvee_{M \in \mu(\varphi)} M$. The definition of $\mu(\varphi)$ depends on the cardinality of the universe (finite or infinite). The number of minterms in $\mu(\varphi)$ is exponential in the size of φ, however, the size of every $M \in \mu(\varphi)$ is bounded by a polynomial in the size of φ and, as we show, checking the membership of a given minterm M in $\mu(\varphi)$ can be done in PSPACE.

5.1 Minterms

We introduce some definitions and notations, and establish basic properties.

Definition 5.1. A *literal* is a test formula or its negation. A *minterm* M is a set of literals, interpreted as the conjunction of its elements, that contains:

- at most one literal of the form $|U| \geq n$;
- at most one literal of the form $|U| < n$;
- exactly one literal $|h| \geq \min_M$, where $\min_M \in \mathbb{N} \cup \{|U| - n \mid n \in \mathbb{N}\}$;
- exactly one literal $|h| < \max_M$, where $\max_M \in \mathbb{N}_\infty \cup \{|U| - n \mid n \in \mathbb{N}\}$.

Definition 5.2. Given a minterm M, we define the sets:

$$M^e \stackrel{\text{def}}{=} M \cap \{x \approx y, \neg x \approx y \mid x, y \in \mathsf{Var}\} \qquad M^a \stackrel{\text{def}}{=} M \cap \{\mathsf{alloc}(x), \neg\mathsf{alloc}(x) \mid x \in \mathsf{Var}\}$$
$$M^u \stackrel{\text{def}}{=} M \cap \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\} \qquad M^p \stackrel{\text{def}}{=} M \cap \{x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y} \mid x, \mathbf{y} \in \mathsf{Var}^{k+1}\}$$
$$M^f \stackrel{\text{def}}{=} M \cap \{q(\mathbf{x}), \neg q(\mathbf{x}) \mid q \in \mathcal{F}, \sigma(q) = \mathsf{B}, \mathbf{x} \in \mathsf{Var}^{\#(q)}\}$$

Thus, $M = M^e \cup M^u \cup M^a \cup M^p \cup M^f \cup \{|h| \geq \min_M, |h| < \max_M\}$, for each minterm M.

Proposition 5.3. *Given a minterm M, for all structures $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ and $\mathcal{S}' = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}')$ we have $\mathcal{S} \models M^e \wedge M^u \wedge M^f \Leftrightarrow \mathcal{S}' \models M^e \wedge M^u \wedge M^f$.*

Proof. This is immediate, since the semantics of the test formulæ in $M^e \cup M^u \cup M^f$ does not depend on the heap. □

Definition 5.4. Given a set of variables $X \subseteq \mathsf{Var}$, a minterm M is (1) *E-complete* for X iff for all $x, y \in X$, exactly one of $x \approx y \in M$, $\neg x \approx y \in M$ holds, and (2) *A-complete* for X iff for each $x \in X$ exactly one of $\mathsf{alloc}(x) \in M$, $\neg\mathsf{alloc}(x) \in M$ holds.

For a literal ℓ, we denote by $\overline{\ell}$ its complement, i.e., $\overline{\theta} \stackrel{\text{def}}{=} \neg\theta$ and $\overline{\neg\theta} \stackrel{\text{def}}{=} \theta$, where θ is a test formula. If T is a set of literals, then we denote by $\mathsf{atoms}(T)$ the set of all test formulæ ϕ such that either ϕ or $\neg\phi$ occurs in T. The equivalence relation $x \approx_T y$ is defined as $T \models x \approx y$ and we write $x \not\approx_T y$ for $T \models \neg x \approx y$. Observe that $x \not\approx_T y$ is not the complement of $x \approx_T y$. For a set X of variables, $|X|_T$ is the number of equivalence classes of $\approx_T$ in X. Two tuples $\mathbf{y}, \mathbf{y}' \in \mathsf{Var}^k$ are *T-distinct* if $y_i \not\approx_T y'_i$, for some $i \in [1, k]$.

Proposition 5.5. *If M is E-complete for $\mathrm{Var}(M)$, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M$ and $X \subseteq \mathrm{Var}(M)$, then $|X|_M = ||\mathfrak{s}(X)||$.*

Proof. This is an immediate consequence of the fact that if $x, x' \in X$, then $\mathfrak{s}(x) = \mathfrak{s}(x')$ if and only if $M \models x \approx x'$. □

Definition 5.6. For a set T of literals, let:

$$\begin{aligned}
\mathsf{av}(T) &\stackrel{\text{def}}{=} \left\{x \in \mathsf{Var} \mid \exists x' \in \mathsf{Var} \,.\, x \approx_T x',\ T \cap \{\mathsf{alloc}(x'), x' \hookrightarrow \mathbf{y} \mid \mathbf{y} \in \mathsf{Var}^k\} \neq \emptyset\right\} \\
\mathsf{nv}(T) &\stackrel{\text{def}}{=} \{x \in \mathsf{Var} \mid \exists x' \in \mathsf{Var} \,.\, x \approx_T x',\ \neg\mathsf{alloc}(x') \in T\} \\
\mathsf{fp}_X(T) &\stackrel{\text{def}}{=} T \cap \{\mathsf{alloc}(x), \neg\mathsf{alloc}(x), x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y} \mid x \in X, \mathbf{y} \in \mathsf{Var}^k\} \\
\#_{\mathrm{a}}(T) &\stackrel{\text{def}}{=} |\mathsf{av}(T)|_T \\
\#_{\mathrm{n}}(X, T) &\stackrel{\text{def}}{=} |X \cap \mathsf{nv}(T)|_T
\end{aligned}$$

For notational convenience, we also let $\mathsf{fp}_a(T) \stackrel{\text{def}}{=} \mathsf{fp}_{\mathsf{av}(T)}(T)$.

Intuitively, $\mathsf{av}(T)$ (resp. $\mathsf{nv}(T)$) is the set of variables that must be (resp. are never) allocated in every (resp. any) model of T. The symbol $\#_{\mathrm{a}}(T)$ represents the number of equivalence classes of $\approx_T$ containing variables allocated in every model of T; $\#_{\mathrm{n}}(X, T)$ represents the number of equivalence classes of $\approx_T$ containing variables from X that are not allocated in any model of T and $\mathsf{fp}_X(T)$ is the *footprint* of T relative to the set $X \subseteq \mathsf{Var}$, i.e. the set of formulæ describing allocation and points-to relations over variables from X. For example, if $T = \{x \approx z, \mathsf{alloc}(x), \neg\mathsf{alloc}(y), \neg z \hookrightarrow \mathbf{y}\}$, then $\mathsf{av}(T) = \{x, z\}$; $\mathsf{nv}(T) = \{y\}$; $\#_{\mathrm{a}}(T) = 1$; $\#_{\mathrm{n}}(\{y\}, T) = 1$; $\mathsf{fp}_a(T) = \{\mathsf{alloc}(x), \neg z \hookrightarrow \mathbf{y}\}$ and $\mathsf{fp}_{\mathsf{nv}(T)}(T) = \{\neg\mathsf{alloc}(y)\}$.

Proposition 5.7. *Given a set T of test formulæ and a structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$, if $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{fp}_a(T)$, then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \mathsf{fp}_a(T)$ for every extension $\mathfrak{h}'$ of $\mathfrak{h}$.*

Proof. Assume that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{fp}_a(T)$ and let $\phi \in \mathsf{fp}_a(T)$. If ϕ is of the form $\neg\mathsf{alloc}(x)$, then since $x \in \mathsf{av}(T)$, necessarily, T contains an atom of the form $\mathsf{alloc}(x')$ or $x' \hookrightarrow \mathbf{y}$, where x' is a variable such that $x' \approx_T x$. In both cases, $\mathsf{fp}_a(T)$ must be unsatisfiable, contradicting the assumption that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{fp}_a(T)$. If ϕ is of the form $x \hookrightarrow \mathbf{y}$, then, since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \phi$, we have $\mathfrak{h}(\mathfrak{s}(x)) = \mathfrak{s}(\mathbf{y})$, thus $\mathfrak{h}'(\mathfrak{s}(x)) = \mathfrak{s}(\mathbf{y})$ (since $\mathfrak{h}'$ is an extension of $\mathfrak{h}$) so that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \phi$. The proof is similar if $\phi = \mathsf{alloc}(x)$. If $\phi = \neg x \hookrightarrow \mathbf{y}$ and T contains an atom of the form $\mathsf{alloc}(x')$ for some variable x' such that $x \approx_T x'$, then $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$ and $\mathfrak{h}(\mathfrak{s}(x)) \neq \mathfrak{s}(\mathbf{y})$. This entails that $\mathfrak{h}'(\mathfrak{s}(x)) \neq \mathfrak{s}(\mathbf{y})$ (since $\mathfrak{h}'$ is an extension of $\mathfrak{h}$) and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \phi$. Otherwise, because $x \in \mathsf{av}(T)$, T must contain an atom of the form $x' \hookrightarrow \mathbf{y}'$ for some variable x' such that $x \approx_T x'$. Thus, $\mathfrak{h}(\mathfrak{s}(x)) = \mathfrak{s}(\mathbf{y}') \neq \mathfrak{s}(\mathbf{y})$, and we deduce that $\mathfrak{h}'(\mathfrak{s}(x)) \neq \mathfrak{s}(\mathbf{y})$. □

Definition 5.8. Given the minterms M_1, M_2, let $\mathsf{npto}(M_1, M_2) \stackrel{\text{def}}{=} (M_1 \cap M_2) \cap \{\neg x \hookrightarrow \mathbf{y} \mid x \notin \mathsf{av}(M_1 \cup M_2), \mathbf{y} \in \mathsf{Var}^k\}$ be the set of negative points-to literals common to M_1 and M_2, involving left-hand side variables not allocated in either M_1 or M_2.

For example, if $M_1 = \{x \hookrightarrow y, \neg y \hookrightarrow z, \neg y \hookrightarrow u, \neg z \hookrightarrow u, |h| \geq 1, |h| < \infty\}$ and $M_2 = \{x \hookrightarrow y, \neg y \hookrightarrow z, \neg z \hookrightarrow u, \mathsf{alloc}(z), |h| \geq 1, |h| < \infty\}$. Then $\mathsf{npto}(M_1, M_2) = \{\neg y \hookrightarrow z\}$. Observe that $M_1 * M_2$ necessarily entails $\mathsf{npto}(M_1, M_2)$, since the assertion $y \hookrightarrow z$ cannot hold in any part of the heap.

We now introduce some conditions that are necessary for a minterm to be satisfiable. The first condition is that the same element cannot point to distinct vectors.

Definition 5.9. Given a minterm M, its *points-to closure* is $\mathsf{pc}(M) \stackrel{\text{def}}{=} \bot$ if there exist literals $x \hookrightarrow \mathbf{y}, x' \hookrightarrow \mathbf{y}' \in M$ such that $x \approx_M x'$ and $\mathbf{y}$, $\mathbf{y}'$ are M-distinct; and $\mathsf{pc}(M) \stackrel{\text{def}}{=} M$, otherwise.

Intuitively, $\mathsf{pc}(M)$ is $\bot$ iff M contradicts the fact that the heap is a partial function. For instance, let $M = \{x \hookrightarrow (y_1, y_2), x' \hookrightarrow (y_1', y_2'), x \approx x', \neg y_1 \approx y_1', |h| \geq 1, |h| < \infty\}$. We have $\mathsf{pc}(M) = \bot$, and it is clear that M is unsatisfiable as the same location cannot point to both (y_1, y_2) and (y_1', y_2'). Note that we do not assert the equality $\mathbf{y} \approx \mathbf{y}'$, instead we only check that it is not falsified. This is sufficient for our purpose because in the following we always assume that the considered minterms are E-complete.

The second condition is that the alloc and point-to literals should be consistent:

Definition 5.10. A minterm M is *footprint-consistent* if for all $x, x' \in \mathsf{Var}$ and $\mathbf{y}, \mathbf{y}' \in \mathsf{Var}^k$, such that $x \approx_M x'$ and $y_i \approx_M y_i'$ for all $i \in [1, k]$, we have (1) if $\mathsf{alloc}(x) \in M$ then $\neg\mathsf{alloc}(x') \notin M$, and (2) if $x \hookrightarrow \mathbf{y} \in M$ then $\{\neg\mathsf{alloc}(x'), \neg x' \hookrightarrow \mathbf{y}'\} \cap M = \emptyset$.

Proposition 5.11. *If M is a footprint-consistent minterm, then $\mathsf{nv}(M) \cap \mathsf{av}(M) = \emptyset$. If, moreover, M is E-complete for $\mathrm{Var}(M)$, then $\mathfrak{s}(X) \cap \mathfrak{s}(\mathsf{av}(M)) = \emptyset$ for each set X disjoint from $\mathsf{av}(M)$ and each model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ of M.*

Proof. Suppose first that $x \in \mathsf{nv}(M) \cap \mathsf{av}(M)$. Then there exist literals $\neg\mathsf{alloc}(x')$ and $\mathsf{alloc}(x'')$ in M such that $x \approx_M x'$ and $x \approx_M x''$, which contradicts the footprint consistency of M. For the second point, suppose

753 that $\ell \in \mathfrak{s}(X) \cap \mathfrak{s}(\mathsf{av}(M))$. Then there exist variables $x \in X$ and $x' \in \mathsf{av}(M)$ such that $\mathfrak{s}(x) = \mathfrak{s}(x') = \ell$. If M
754 is E-complete, either $x \approx x' \in M$ or $\neg x \approx x' \in M$. The first case contradicts $x \notin \mathsf{av}(M)$ and the second case
755 contradicts $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M$. □

756
757 Footprint-consistency is not sufficient for satisfiability. For example, $\{x \hookrightarrow y, x' \hookrightarrow y', \neg y \approx y', |h| < 2\}$ is at
758 the same time footprint-consistent and unsatisfiable, because x and x' point to distinct elements but there is
759 at most one allocated location. We thus introduce additional conditions related to the cardinality of the heap
760 or of the universe. Intuitively, for any minterm M, we define a formula $\mathsf{dc}(M)$ that asserts that $\min_M < \max_M$
761 and that the domain contains enough elements to allocate all cells. Essentially, given a structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$, if
762 $\mathfrak{h}(x)$ is known to be defined and distinct from n pairwise distinct vectors of locations $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then necessarily
763 at least $n+1$ vectors must exist. Since there are $||\mathfrak{U}||^k$ vectors of length k, we must have $||\mathfrak{U}||^k \geq n+1$, hence
764 $||\mathfrak{U}|| \geq \sqrt[k]{n+1}$. For instance, if

765 $$M = \{\neg x \hookrightarrow y_i \mid i \in [1,n]\} \cup \{\mathsf{alloc}(x)\} \cup \{y_i \not\approx y_j \mid i,j \in [1,n], i \neq j\}$$

766
767 then it is clear that M is unsatisfiable if there are less than n locations, since x cannot be allocated in this case.

768 *Definition 5.12.* Given a minterm M, the *domain closure* of M is $\mathsf{dc}(M) \stackrel{\text{def}}{=} \bot$ if either $\min_M = n_1$ and $\max_M = n_2$
769 for some $n_1, n_2 \in \mathbb{Z}$ such that $n_1 \geq n_2$, or $\min_M = |U| - n_1$ and $\max_M = |U| - n_2$, where $n_2 \geq n_1$; and otherwise:

770
771
772
773
$$\begin{array}{rcl} \mathsf{dc}(M) & \stackrel{\text{def}}{=} & M \cup \left\{|U| \geq \left\lceil \sqrt[k]{\max_{x \in \mathsf{av}(M)}(\delta_x(M)+1)} \right\rceil\right\} \\ & & \cup \{|U| \geq n_1 + n_2 + 1 \mid \min_M = n_1, \max_M = |U| - n_2, n_1, n_2 \in \mathbb{N}\} \\ & & \cup \{|U| < n_1 + n_2 \mid \min_M = |U| - n_1, \max_M = n_2, n_1, n_2 \in \mathbb{N}\}, \end{array}$$

774 where $\delta_x(M)$ is the number of pairwise M-distinct tuples $\mathbf{y}$ for which there exists $\neg x' \hookrightarrow \mathbf{y} \in M$ such that
775 $x \approx_M x'$. For any SL-structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$, we denote by $\min_M^{\mathcal{S}}, \max_M^{\mathcal{S}} \in \mathbb{N}_\infty$ the values obtained by replacing
776 $|U|$ with $||\mathfrak{U}||$ in $\min_M$ and $\max_M$, respectively.

777
778 *Example 5.13.* Let $M = \{|h| \geq 0, |h| < \infty, \mathsf{alloc}(y_0)\} \cup \{\neg y_i \approx y_j \mid i,j \in [0,n], i \neq j\} \cup \{y_0 \hookrightarrow y_i \mid i \in [1,n]\}$.
779 Then $y_0 \in \mathsf{av}(M)$, $\delta_x(M) = n$ and $\mathsf{dc}(M) = M \cup \{|U| \geq n+1\}$. This states that all models of M contain at least
780 $n+1$ locations: $y_1, \ldots, y_n$ and the image of y_0 by the heap.
781 Let $M' = \{|h| \geq 1, |h| < |U| - 1\}$. Then $\mathsf{dc}(M') = M' \cup \{|U| \geq 3\}$. All models of M' contain at least 3 locations
782 (one allocated and two non allocated). ■

783 Proposition 5.14. *Given a minterm M, $\min_M^{\mathcal{S}} < \max_M^{\mathcal{S}}$ for every model $\mathcal{S}$ of $\mathsf{dc}(M)^u$.*

784
785 Proof. Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ and $n_1, n_2 \in \mathbb{N}_\infty$. We distinguish the following cases:

786 - If $\min_M = n_1$ and $\max_M = n_2$ then $n_1 < n_2$ must be the case, or else $\mathsf{dc}(M) \equiv \bot$, in contradiction with
787 $\mathcal{S} \models \mathsf{dc}(M)^u$.
788 - If $\min_M = n_1$ and $\max_M = |U| - n_2$ then $|U| \geq n_1 + n_2 + 1 \in \mathsf{dc}(M)$ and since $\mathcal{S} \models \mathsf{dc}(M)^u$, we obtain
789 $n_1 < ||\mathfrak{U}|| - n_2$.
790 - If $\min_M = |U| - n_1$ and $\max_M = n_2$ then $|U| < n_1 + n_2 \in \mathsf{dc}(M)$ and since $\mathcal{S} \models \mathsf{dc}(M^u)$, we obtain
791 $||\mathfrak{U}|| - n_1 < n_2$.
792 - If $\min_M = |U| - n_1$ and $\max_M = |U| - n_2$ then $n_2 < n_1$ must be the case, or else $\mathsf{dc}(M) \equiv \bot$, in contradiction
793 with $\mathcal{S} \models \mathsf{dc}(M)^u$.

794 □

795
796 Proposition 5.15. *For any minterm M, we have $M \equiv \mathsf{pc}(M) \equiv \mathsf{dc}(M)$.*

797 Proof. It is clear that $\mathsf{pc}(M) \models M$ and $\mathsf{dc}(M) \models M$. Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a model of M. If $\mathcal{S} \not\models \mathsf{pc}(M)$ then
798 necessarily $\mathsf{pc}(M) = \bot$ and there exist variables $x, x' \in \mathrm{Var}(M)$ such that $x \hookrightarrow (y_1, \ldots, y_k), x' \hookrightarrow (z_1, \ldots, z_k) \in$
799

$M, x \approx_M x'$ and $(y_1, \ldots, y_k)$ and $(z_1, \ldots, z_k)$ are M-distinct, i.e., there exists $i \in [1, k]$ such that $M \models \neg y_i \approx z_i$. We have $\mathfrak{h}(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$, $\mathfrak{h}(\mathfrak{s}(x')) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$ and $\mathfrak{h}(\mathfrak{s}(x)) = \mathfrak{h}(\mathfrak{s}(x'))$, thus $\mathfrak{s}(y_i) = \mathfrak{s}(z_i)$, for all $i \in [1, k]$, a contradiction. Thus $\mathcal{S} \models \mathsf{pc}(M)$. For a variable $x \in \mathsf{av}(M)$, let $\neg x_1 \hookrightarrow \mathrm{y}_1, \ldots, \neg x_n \hookrightarrow \mathrm{y}_n \in M$ be all literals such that $x_1 \approx_M \ldots \approx_M x_n \approx_M x$ and $\mathrm{y}_i \not\approx_M \mathrm{y}_j$ for all $i \neq j$. Then $\mathfrak{h}(\mathfrak{s}(x)) \in \mathfrak{U}^k \setminus \{\mathfrak{s}(\mathrm{y}_1), \ldots, \mathfrak{s}(\mathrm{y}_n)\}$, thus $||\mathfrak{U}||^k \geq n + 1 = \delta_x(M) + 1$. Since this holds for each $x \in \mathsf{av}(M)$, we have $\mathcal{S} \models |U| \geq \left\lceil \sqrt[k]{\max_{x \in \mathsf{av}(M)}(\delta_x(M) + 1)} \right\rceil$. Furthermore, if $|h| \geq n_1, |h| < |U| - n_2 \in M$ then, since $\mathcal{S} \models M$, $||\mathfrak{U}|| - n_2 > ||\mathfrak{h}|| \geq n_1$, thus $||U|| \geq n_1 + n_2 + 1$ and $\mathcal{S} \models |U| \geq n_1 + n_2 + 1$. Analogously, we obtain $\mathcal{S} \models |U| < n_1 + n_2$ in the case $|h| < n_1, |h| \geq |U| - n_2 \in M$. □

5.2 Eliminating Spatial Connectives

We now show how to eliminate the connectives $*$ and $-\!\!*$, i.e., to transform a formula of the form $\phi_1 * \phi_2$ or $\phi_1 -\!\!* \phi_2$ into an equivalent boolean combination of test formulæ, assuming ϕ_1 and ϕ_2 have already been transformed. We solve this problem by restricting ourselves to the case where ϕ_1 and ϕ_2 are minterms satisfying some additional properties. We first consider the separating conjunction.

Lemma 5.16. *Let M_1, M_2 be two minterms that are footprint-consistent and E-complete for* $\mathrm{Var}(M_1 \cup M_2)$*, with* $\mathsf{atoms}(M_1^p) = \mathsf{atoms}(M_2^p)$*. Then $M_1 * M_2 \equiv \mathsf{elim}_*(M_1, M_2)$, where*

$$\begin{aligned}
\mathsf{elim}_*(M_1, M_2) \stackrel{\text{def}}{=}\ & M_1^e \wedge M_2^e \wedge M_1^f \wedge M_2^f \wedge \mathsf{dc}(M_1)^u \wedge \mathsf{dc}(M_2)^u \wedge && (2)\\
& \bigwedge_{x \in \mathsf{av}(M_1),\ y \in \mathsf{av}(M_2)} \neg x \approx y \wedge \mathsf{fp}_a(M_1) \wedge \mathsf{fp}_a(M_2) \wedge && (3)\\
& \mathsf{nalloc}(\mathsf{nv}(M_1) \cap \mathsf{nv}(M_2)) \wedge \mathsf{npto}(M_1, M_2) \wedge && (4)\\
& |h| \geq \min_{M_1} + \min_{M_2} \ \wedge\ |h| < \max_{M_1} + \max_{M_2} - 1 && (5)\\
& \wedge\ \eta_{12} \wedge \eta_{21} && (6)
\end{aligned}$$

and $\eta_{ij} \stackrel{\text{def}}{=} \bigwedge_{Y \subseteq \mathsf{nv}(M_j) \setminus \mathsf{av}(M_i)} \Big(\mathsf{alloc}(Y) \rightarrow (|h| \geq \#_{\mathrm{a}}(M_i) + |Y|_{M_i} + \min_{M_j} \wedge\ \#_{\mathrm{a}}(M_i) + |Y|_{M_i} < \max_{M_i})\Big)$.

Intuitively, if M_1 and M_2 hold separately, then all heap-independent literals from $M_1 \cup M_2$ must be satisfied (2), the variables allocated in M_1 and M_2 must be pairwise distinct and their footprints, relative to the allocated variables, jointly asserted (3). Moreover, unallocated variables on both sides must not be allocated and common negative points-to literals must be asserted (4). Since the heap satisfying $\mathsf{elim}_*(M_1, M_2)$ is the disjoint union of the heaps for M_1 and M_2, its bounds are the sum of the bounds on both sides (5) and the variables that M_2 never allocates (the set $\mathsf{nv}(M_2)$) may occur allocated in the heap of M_1 and vice versa, thus the constraints η_{12} and η_{21}, respectively (6).

The proof of Lemma 5.16 requires the following result:

Proposition 5.17. *Let M_1, M_2 be two minterms that are footprint-consistent and E-complete for* $\mathrm{Var}(M_1 \cup M_2)$ *and let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a model of* $\mathsf{elim}_*(M_1, M_2)$*. Let L_i, Y_i, A_i be the following sets, for $i = 1, 2$:*

$$\begin{aligned}
L_i &= \{\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}) \mid x \in \mathsf{nv}(M_{3-i}) \setminus \mathsf{av}(M_i)\}\\
Y_i &= \{x \in \mathrm{Var} \mid \mathfrak{s}(x) \in L_i\}\\
A_i &= \{\mathfrak{s}(x) \mid x \in \mathsf{av}(M_i)\}
\end{aligned}$$

Then $L_1 \cap L_2 = \emptyset$, $L_i \cap (A_1 \cup A_2) = \emptyset$ (for $i = 1, 2$) and $\mathcal{S} \models \mathsf{alloc}(Y_1) \wedge \mathsf{alloc}(Y_2)$.

Proof. We have the following results:

- **$L_1 \cap L_2 = \emptyset$.** By contradiction, suppose that there exists $\ell \in L_1 \cap L_2$. Then $\ell = \mathfrak{s}(y_1) = \mathfrak{s}(y_2)$ for some $y_1 \in \mathsf{nv}(M_1)$ and $y_2 \in \mathsf{nv}(M_2)$. Because M_1 is E-complete for $\mathrm{Var}(M_1 \cup M_2)$, exactly one of $y_1 \approx y_2$, $\neg y_1 \approx y_2$ belongs to M_1. But $\neg y_1 \approx y_2 \in M_1$ contradicts $\mathfrak{s}(y_1) = \mathfrak{s}(y_2)$ and $y_1 \approx y_2 \in M_1$ leads to

$y_2 \in \mathsf{nv}(M_1)$. Symmetrically, $y_1 \in \mathsf{nv}(M_2)$, thus $y_1, y_2 \in \mathsf{nalloc}(\mathsf{nv}(M_1) \cap \mathsf{nv}(M_2))$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{nalloc}(\mathsf{nv}(M_1) \cap \mathsf{nv}(M_2))$ by (4), we have $\ell \notin \mathrm{dom}(\mathfrak{h})$, which contradicts with the fact that $L_1 \cup L_2 \subseteq \mathrm{dom}(\mathfrak{h})$, according to the definition of L_1 and L_2.

- $\boldsymbol{L_i \cap (A_1 \cup A_2) = \emptyset}$**.** First, $L_i \cap A_i = \emptyset$ because M_i is E-complete for $\mathrm{Var}(M_1 \cup M_2)$, and by Proposition 5.11. Second, $L_i \cap A_{3-i} = \emptyset$ because M_i is E-complete for $\mathrm{Var}(M_1 \cup M_2)$ and $\mathsf{nv}(M_{3-i}) \cap \mathsf{av}(M_{3-i}) = \emptyset$, by Proposition 5.11.
- $\boldsymbol{\mathcal{S} \models \mathbf{alloc}(Y_1) \wedge \mathbf{alloc}(Y_2)}$**.** this follows immediately from the fact that $L_1 \cup L_2 \subseteq \mathrm{dom}(\mathfrak{h})$ by definition of L_1, L_2.

□

We are now in the position to prove Lemma 5.16:

Proof. Suppose first that $M_1^e \neq M_2^e$. Since M_1 and M_2 are E-complete for $\mathrm{Var}(M_1 \cup M_2)$, there must exist a literal $x \approx y \in M_1^e$ such that $\neg x \approx y \in M_2^e$, or vice versa. In both cases however $M_1 * M_2 \equiv \mathsf{elim}_*(M_1, M_2) \equiv \bot$. Thus we consider from now on that $M_1^e = M_2^e$.

- $\boldsymbol{M_1 * M_2 \models \mathbf{elim}_*(M_1, M_2)}$**.** Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a model of $M_1 * M_2$. Then there exist disjoint heaps $\mathfrak{h}_1$ and $\mathfrak{h}_2$ such that $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models M_i$, for all $i = 1, 2$. Below we show that $\mathcal{S}$ is a model of the formulæ (2), (3), (4), (5) and (6).
  (2) Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models M_i^e \wedge M_i^u \wedge M_i^f$, by Proposition 5.3, we also have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_i^e \wedge M_i^u \wedge M_i^f$, for $i = 1, 2$. By Proposition 5.15, we obtain further that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathrm{dc}(M_i)^u$, for $i = 1, 2$.
  (3) Since $\mathrm{dom}(\mathfrak{h}_1) \cap \mathrm{dom}(\mathfrak{h}_2) = \emptyset$, for every $x \in \mathsf{av}(M_1)$ and $y \in \mathsf{av}(M_2)$, we must have $\mathfrak{s}(x) \neq \mathfrak{s}(y)$, hence $\mathcal{S} \models \neg x \approx y$. Further, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models M_i$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models \mathsf{fp}_a(M_i)$ and, by Proposition 5.7, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{fp}_a(M_i)$, for $i = 1, 2$.
  (4) Consider a variable $x \in \mathsf{nv}(M_1) \cap \mathsf{nv}(M_2)$. Then there exist variables x_1 and x_2 such that $\neg\mathsf{alloc}(x_1) \in M_1$, $x \approx_{M_1} x_1$, $\neg\mathsf{alloc}(x_2) \in M_2$ and $x \approx_{M_2} x_2$. Hence $\mathfrak{s}(x) = \mathfrak{s}(x_1) \notin \mathrm{dom}(\mathfrak{h}_1)$ and $\mathfrak{s}(x) = \mathfrak{s}(x_2) \notin \mathrm{dom}(\mathfrak{h}_2)$, thus $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\mathsf{alloc}(x)$. Since x was chosen arbitrarily, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{nalloc}(\mathsf{nv}(M_1) \cap \mathsf{nv}(M_2))$. Secondly, let $\neg x \hookrightarrow \mathbf{y} \in M_1 \cap M_2$, for some $x \notin \mathsf{av}(M_1 \cup M_2)$. Since $\mathrm{dom}(\mathfrak{h}_1) \cap \mathrm{dom}(\mathfrak{h}_2) = \emptyset$, only the following are possible:
  i. $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}_1)$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models M_1$, we must have $\mathfrak{h}_1(\mathfrak{s}(x)) \neq \mathfrak{s}(\mathbf{y})$. Then $\mathfrak{h}(\mathfrak{s}(x)) \neq \mathfrak{s}(\mathbf{y})$ thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg x \hookrightarrow \mathbf{y}$.
  ii. $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}_2)$ and $\mathfrak{h}_2(x) \neq \mathfrak{s}(\mathbf{y})$ is symmetrical.
  iii. $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}_1) \cup \mathrm{dom}(\mathfrak{h}_2)$, then $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg x \hookrightarrow \mathbf{y}$.
  Since $\neg x \hookrightarrow \mathbf{y} \in \mathsf{npto}(M_1, M_2)$ was chosen arbitrarily, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{npto}(M_1, M_2)$.
  (5) Since $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$, we have $||\mathfrak{h}|| = ||\mathfrak{h}_1|| + ||\mathfrak{h}_2||$, thus the first two constraints are obtained by summing up the constraints $\min^{\mathcal{S}}_{M_i} \leq ||\mathfrak{h}_i|| < \max^{\mathcal{S}}_{M_i}$, for $i = 1, 2$.
  (6) We prove $\mathcal{S} \models \eta_{12}$, the proof for $\mathcal{S} \models \eta_{21}$ being symmetrical. Consider a set $Y \subseteq \mathsf{nv}(M_2) \setminus \mathsf{av}(M_1)$ and suppose that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{alloc}(Y)$. For each $y \in Y$ we must have $\mathfrak{s}(y) \in \mathrm{dom}(\mathfrak{h}_1)$, because $\mathfrak{s}(y) \notin \mathrm{dom}(\mathfrak{h}_2)$ and $\mathfrak{s}(y) \in \mathrm{dom}(\mathfrak{h})$. Moreover, $\mathfrak{s}(Y) \cap \mathfrak{s}(\mathsf{av}(M_1)) = \emptyset$ because $Y \cap \mathsf{av}(M_1) = \emptyset$ and M_1 is E-complete for $\mathrm{Var}(M_1 \cup M_2)$, by Proposition 5.11. Thus $\#_{\mathsf{a}}(M_1) + |Y|_{M_1} \leq ||\mathfrak{h}_1|| < \max^{\mathcal{S}}_{M_1}$ and $||\mathfrak{h}|| = ||\mathfrak{h}_1|| + ||\mathfrak{h}_2|| \geq \#_{\mathsf{a}}(M_1) + |Y|_{M_1} + \min^{\mathcal{S}}_{M_2}$, as required.
- $\boldsymbol{\mathsf{elim}_*(M_1, M_2) \models M_1 * M_2}$**.** Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a model of $\mathsf{elim}_*(M_1, M_2)$. We will find $\mathfrak{h}_1$ and $\mathfrak{h}_2$ such that $\mathfrak{h} = \mathfrak{h}_1 \uplus \mathfrak{h}_2$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models M_i$, for $i = 1, 2$. Since $\mathcal{S} \models \min_{M_1} + \min_{M_2} \leq |h| \wedge |h| < \max_{M_1} + \max_{M_2} - 1$ by (5), we have, by Proposition 4.2:

$$\min^{\mathcal{S}}_{M_1} + \min^{\mathcal{S}}_{M_2} \leq ||\mathfrak{h}|| < \max^{\mathcal{S}}_{M_1} + \max^{\mathcal{S}}_{M_2} - 1 \tag{7}$$

Let us now define the following sets, for $i = 1, 2$:

$$\begin{array}{rcl} L_i & = & \{\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}) \mid x \in \mathsf{nv}(M_{3-i}) \setminus \mathsf{av}(M_i)\} \\ Y_i & = & \{x \in \mathsf{Var} \mid \mathfrak{s}(x) \in L_i\} \\ A_i & = & \{\mathfrak{s}(x) \mid x \in \mathsf{av}(M_i)\} \end{array}$$

By Proposition 5.17, we have $L_1 \cap L_2 = \emptyset$, $L_i \cap (A_1 \cup A_2) = \emptyset$, for $i = 1, 2$ and $\mathcal{S} \models \mathsf{alloc}(Y_1) \wedge \mathsf{alloc}(Y_2)$. Moreover, because $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \eta_{12} \wedge \eta_{21}$, the following hold, for $i = 1, 2$:

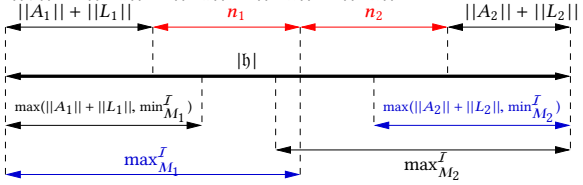
$$||\mathfrak{h}|| \geq ||A_i|| + ||L_i|| + \min^{\mathcal{S}}_{M_{3-i}} \quad (8) \qquad ||A_i|| + ||L_i|| < \max^{\mathcal{I}}_{i} \quad (9)$$

We prove the following relation by distinguishing the cases below:

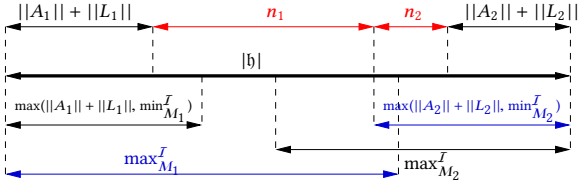
$$\max(\min^{\mathcal{S}}_{M_1}, ||A_1|| + ||L_1||) + \max(\min^{\mathcal{S}}_{M_2}, ||A_2|| + ||L_2||) \leq ||\mathfrak{h}|| \quad (10)$$

(1) if $\min^{\mathcal{S}}_{M_1} \geq ||A_1|| + ||L_1||$ then we have $\min^{\mathcal{S}}_{M_1} + \max(\min^{\mathcal{S}}_{M_2}, ||A_2|| + ||L_2||) \leq ||\mathfrak{h}||$ by (5) and (8). The case $\min^{\mathcal{S}}_{M_2} \geq ||A_2|| + ||L_2||$ is symmetric, and

(2) otherwise, if $\min^{\mathcal{S}}_{M_1} < ||A_1|| + ||L_1||$ and $\min^{\mathcal{S}}_{M_2} < ||A_2|| + ||L_2||$, because $\mathcal{S} \models \bigwedge_{x \in \mathsf{av}(M_1),\ y \in \mathsf{av}(M_2)} \neg x \approx y$, the sets of locations L_1, L_2, A_1 and A_2 are pairwise disjoint and, since $L_1 \cup L_2 \cup A_1 \cup A_2 \subseteq \mathrm{dom}(\mathfrak{h})$, it must be the case that $||\mathfrak{h}|| \geq ||A_1|| + ||L_1|| + ||A_2|| + ||L_2||$.



(a)



(b)

Fig. 1

Furthermore, we have $||\mathfrak{h}|| < \max^{\mathcal{S}}_{M_1} + \max^{\mathcal{S}}_{M_2} - 1$ by (7) and one of the following cases holds (see Fig. 1):

(1) If $\max^{\mathcal{S}}_{M_1} - 1 \leq ||\mathfrak{h}|| - \max(||A_2|| + ||L_2||, \min^{\mathcal{S}}_{M_2})$ then let $n_1 \stackrel{\text{def}}{=} \max^{\mathcal{S}}_{M_1} - ||A_1|| - ||L_1|| - 1$ and $n_2 \stackrel{\text{def}}{=} ||\mathfrak{h}|| - \max^{\mathcal{S}}_{M_1} - ||A_2|| - ||L_2|| + 1$ (Fig. 1 (a)). We have that $n_1 \geq 0$ by (9) and $n_2 \geq 0$ by the hypothesis $\max^{\mathcal{S}}_{M_1} - 1 \leq ||\mathfrak{h}|| - \max(||A_2|| + ||L_2||, \min^{\mathcal{S}}_{M_2})$.

(2) Otherwise, let $n_1 \stackrel{\text{def}}{=} ||\mathfrak{h}|| - ||A_1|| - ||L_1|| - \max^{\mathcal{S}}(||A_2|| + ||L_2||, \min^{\mathcal{S}}_{M_2})$ and $n_2 \stackrel{\text{def}}{=} \max^{\mathcal{S}}(||A_2|| + ||L_2||, \min^{\mathcal{S}}_{M_2}) - ||A_2|| - ||L_2||$ (Fig. 1 (b)). We have $n_1 \geq 0$ by (10) and $n_2 \geq 0$ is immediate.

In both cases, the following holds, for $i = 1, 2$:

$$\min^{\mathcal{S}}_{M_i} \leq ||A_i|| + ||L_i|| + n_i < \max^{\mathcal{S}}_{M_i} \quad (11)$$

We have used the fact that $\min^{\mathcal{S}}_{M_i} < \max^{\mathcal{S}}_{M_i}$, for $i = 1, 2$, which is a consequence of the fact that $\mathcal{S} \models \mathrm{dc}(M_i)^u$, by (2) and Proposition 5.14.

Further, we have that $||\mathfrak{h}|| = \sum_{i=1,2} ||A_i|| + ||L_i|| + n_i$. Moreover, there are exactly $n_1 + n_2$ locations in $\mathrm{dom}(\mathfrak{h}) \setminus (A_1 \cup L_1 \cup A_2 \cup L_2)$, thus we can partition this set into N_1 and N_2 such that $||N_i|| = n_i$ and define $\mathfrak{h}_i$ to be the restriction of $\mathfrak{h}$ to $A_i \cup L_i \cup N_i$, for $i = 1, 2$. It remains to be shown that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_i) \models M_i$, for $i = 1, 2$. Below we do the proof for $i = 1$, the case $i = 2$ being similar.

Clearly, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models M_1^e \wedge M_1^f \wedge M_1^u$, because $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_1^e \wedge M_1^f \wedge \mathrm{dc}(M_1)^u$, by Proposition 5.3. Further, by (11) and Proposition 4.2, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models |h| \geq \min_{M_1} \wedge |h| < \max_{M_1}$. There remains to show that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models M_1^a \wedge M_1^p$.

(M_1^a) Let $\mathrm{alloc}(x) \in M_1^a$ be a literal. Then $x \in \mathrm{av}(M_1)$, thus $\mathfrak{s}(x) \in A_1$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \mathrm{alloc}(x)$ follows, by the definition of $\mathfrak{h}_1$. Dually, let $\neg\mathrm{alloc}(x) \in M_1^a$ be a literal. Then, we have $x \in \mathrm{nv}(M_1)$. We distinguish the following cases:

- If $x \in \mathrm{av}(M_2)$ then $\mathfrak{s}(x) \in A_2$ and since $\mathrm{dom}(\mathfrak{h}_1) \cap A_2 = \emptyset$, we have $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}_1)$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg\mathrm{alloc}(x)$.
- Otherwise, $x \in \mathrm{nv}(M_1) \setminus \mathrm{av}(M_2)$. Again, we distinguish the cases:
  - if $x \in Y_2$ then $\mathfrak{s}(x) \in L_2$ and because $\mathrm{dom}(\mathfrak{h}_1) \cap L_2 = \emptyset$, we obtain $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}_1)$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg\mathrm{alloc}(x)$.
  - otherwise, $x \notin Y_2$, thus $\mathfrak{s}(x) \notin L_2$. But since $x \in \mathrm{nv}(M_1) \setminus \mathrm{av}(M_2)$, by the definition of L_2, it must be the case that $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\mathrm{alloc}(x)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg\mathrm{alloc}(x)$ follows.

(M_1^p) Let $x \hookrightarrow \mathbf{y} \in M_1^p$ be a literal. Then $x \in \mathrm{av}(M_1)$ and $\mathfrak{s}(x) \in A_1$. Moreover, we have $x \hookrightarrow \mathbf{y} \in \mathrm{fp}_a(M_1)$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models x \hookrightarrow \mathbf{y}$, by (3). Since $\mathfrak{h}$ and $\mathfrak{h}_1$ agree on A_1, we also have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models x \hookrightarrow \mathbf{y}$. Dually, let $\neg x \hookrightarrow \mathbf{y} \in M_1^p$. If $x \in \mathrm{av}(M_1)$ then $\neg x \hookrightarrow \mathbf{y} \in \mathrm{fp}_a(M_1)$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg x \hookrightarrow \mathbf{y}$, since $\mathfrak{h}$ and $\mathfrak{h}_1$ agree on A_1. Otherwise, if $x \notin \mathrm{av}(M_1)$, we distinguish the cases:

- if $x \in \mathrm{av}(M_2)$ then $\mathfrak{s}(x) \in A_2$, and since $\mathrm{dom}(\mathfrak{h}_1) \cap A_2 = \emptyset$, we have $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}_1)$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg x \hookrightarrow \mathbf{y}$.
- otherwise, $x \notin \mathrm{av}(M_2)$, and since $\mathrm{atoms}(M_1^p) = \mathrm{atoms}(M_2^p)$, we have $\{x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y}\} \cap M_2 \neq \emptyset$. Since $x \notin \mathrm{av}(M_2)$, the only possibility is $\neg x \hookrightarrow \mathbf{y} \in M_2$, thus $\neg x \hookrightarrow \mathbf{y} \in \mathrm{npto}(M_1, M_2)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg x \hookrightarrow \mathbf{y}$, by (4). Since $\mathfrak{h}$ is an extension of $\mathfrak{h}_1$, we obtain that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}_1) \models \neg x \hookrightarrow \mathbf{y}$ as well.

□

We provide simple examples of application.

Example 5.18. Consider the following minterms:

$$M_1 = E \cup \{|h| \geq 2, |h| < 4, x \hookrightarrow y, \mathrm{alloc}(y), \neg y \hookrightarrow x, \neg z \hookrightarrow z\}$$

$$M_2 = E \cup \{|h| \geq 1, |h| < 2\}$$

with $E = \{\neg x \approx y, \neg y \approx z, \neg x \approx z\}$. Then $M_1 * M_2 \equiv E \cup \{|h| \geq 3, |h| < 5, x \hookrightarrow y, \mathrm{alloc}(y), \neg y \hookrightarrow x\}$.

Let $M_1' = \{|h| \geq 0, |h| < 1, \neg x \approx y\}$ and $M_2' = \{|h| \geq 0, |h| < \infty, \neg x \approx y, \neg\mathrm{alloc}(x)\}$. Then $M_1' * M_2' \equiv \{|h| \geq 0, |h| < \infty, \mathrm{alloc}(x) \rightarrow 1 < 1\} \equiv \{|h| \geq 0, |h| < \infty, \neg\mathrm{alloc}(x)\}$. Indeed, no model of $M_1' * M_2'$ may allocate x since the part of the heap that corresponds to M_1' is empty and $M_2' \models \neg\mathrm{alloc}(x)$. ■

Remark 5.19. *Note that* $\mathrm{elim}_*(M_1, M_2)$ *contains negative occurrences of test formulæ* $\mathrm{alloc}(x)$ *that do not occur in* $M_1 \cup M_2$. *Such occurrences are introduced at Lines 4 and 6, due to the fact that we consider the closure of* $\neg\mathrm{alloc}(x)$ *formulæ w.r.t. all the equalities in* M_1, M_2. *For example, if* $M_1 = \{\neg\mathrm{alloc}(x), x \approx y, |h| \geq 0, |h| < \infty\}$ *and* $M_2 = \{|h| \geq 0, |h| < \infty\}$, *then* $y \in \mathrm{nv}(M_1)$ *and* $\mathrm{alloc}(y)$ *occurs at negative polarity in* $\mathrm{elim}_*(M_1, M_2)$. *This is problematic because upcoming results depend on the fact that the polarity of* $\mathrm{alloc}(x)$ *formulæ is preserved (Lemma 5.29). However, if* $\mathrm{alloc}(x)$ *occurs at a negative polarity in* $\mathrm{elim}_*(M_1, M_2)$, *then there exists a literal* $\neg\mathrm{alloc}(x') \in M_1 \cup M_2$, *such that* $\mathrm{elim}_*(M_1, M_2) \models x \approx x'$, *making the negative occurrence of* $\mathrm{alloc}(x)$ *actually redundant. Consequently, equivalence is preserved when only the test formulæ* $\mathrm{alloc}(x)$ *such that* $\neg\mathrm{alloc}(x) \in M_1 \cup M_2$ *occur at negative polarity in*

$\mathsf{elim}_*(M_1, M_2)$. *This refined version of* $\mathsf{elim}_*(M_1, M_2)$ *is used in the proof of Lemma 5.29. However, taking this observation into account at this point would clutter the definition of* $\mathsf{elim}_*(M_1, M_2)$. ■

Next, we show a similar result for the separating implication. For technical convenience, we translate the septraction $M_1 \mathrel{-\!\!\circ} M_2$, instead of $M_1 \mathrel{-\!\!*} M_2$, as an equivalent boolean combination of test formulæ. This is without loss of generality, because $M_1 \mathrel{-\!\!*} M_2 \equiv \neg(M_1 \mathrel{-\!\!\circ} \neg M_2)$. Unlike with the case of the separating conjuction (Lemma 5.16), here the definition of the boolean combination of test formulæ depends on whether the universe is finite or infinite.

If the complement of some literal $\ell \in \mathsf{fp}_a(M_1)$ belongs to M_2 then no extension by a heap that satisfies ℓ may satisfy $\overline{\ell}$. Therefore, as an additional simplifying assumption, we suppose that $\mathsf{fp}_a(M_1) \cap \overline{M_2} = \emptyset$, so that $M_1 \mathrel{-\!\!\circ} M_2$ is not trivially unsatisfiable.

Lemma 5.20. *Let* M_1 *and* M_2 *be footprint-consistent minterms that are E-complete for* $\mathrm{Var}(M_1 \cup M_2)$*, such that: (a)* M_1 *is A-complete for* $\mathrm{Var}(M_1 \cup M_2)$*, (b)* $\mathsf{atoms}(M_2^a \cup M_2^p) \subseteq \mathsf{atoms}(M_1^a \cup M_1^p)$*, and (c)* $\mathsf{fp}_a(M_1) \cap \overline{M_2} = \emptyset$*. Then,* $M_1 \mathrel{-\!\!\circ} M_2 \equiv^{fin} \mathsf{elim}^{fin}_{\mathrel{-\!\!\circ}}(M_1, M_2)$ *and* $M_1 \mathrel{-\!\!\circ} M_2 \equiv^{inf} \mathsf{elim}^{inf}_{\mathrel{-\!\!\circ}}(M_1, M_2)$*, where:*

$$\mathsf{elim}^{\dagger}_{\mathrel{-\!\!\circ}}(M_1, M_2) \stackrel{\text{def}}{=} \mathsf{pc}(M_1)^e \wedge M_2^e \wedge M_1^f \wedge M_2^f \wedge \mathsf{dc}(M_1)^u \wedge \mathsf{dc}(M_2)^u \wedge \quad (12)$$

$$\mathsf{nalloc}(\mathsf{av}(M_1)) \wedge \mathsf{fp}_{\mathsf{nv}(M_1)}(M_2) \wedge \quad (13)$$

$$|h| \geq \min_{M_2} - \max_{M_1} + 1 \wedge |h| < \max_{M_2} - \min_{M_1} \quad (14)$$

$$\wedge\ \lambda^{\dagger} \quad (15)$$

with $\lambda^{fin} \stackrel{\text{def}}{=} \bigwedge_{Y \subseteq \mathrm{Var}(M_1 \cup M_2)} \mathsf{nalloc}(Y) \rightarrow \left(\begin{array}{l} |h| < |U| - \min_{M_1} - \#_{\mathrm{n}}(Y, M_1) + 1 \\ \wedge\ |U| \geq \min_{M_2} + \#_{\mathrm{n}}(Y, M_1) \end{array} \right)$, *and* $\lambda^{inf} \stackrel{\text{def}}{=} \top$.

Intuitively, a heap satisfies $M_1 \mathrel{-\!\!\circ} M_2$ iff it has an extension, by a disjoint heap satisfying M_1, that satisfies M_2. Thus, $\mathsf{elim}^{\dagger}_{\mathrel{-\!\!\circ}}(M_1, M_2)$ must entail the heap-independent literals of both M_1 and M_2 (12). Next, no variable allocated by M_1 must be allocated by $\mathsf{elim}^{\dagger}_{\mathrel{-\!\!\circ}}(M_1, M_2)$, otherwise no extension by a heap satisfying M_1 is possible and, moreover, the footprint of M_2 relative to the unallocated variables of M_1 must be asserted (13). The heap's cardinality constraints depend on the bounds of M_1 and M_2 (14) and, if Y is a set of variables not allocated in the heap, these variables can be allocated in the extension (15). Actually, this is where the finite universe assumption first comes into play. If the universe is infinite, then there are enough locations outside the heap to be assigned to Y. However, if the universe is finite, then it is necessary to ensure that there are at least $\#_{\mathrm{n}}(Y, M_1)$ free locations to be assigned to Y (15). We now give the proof of Lemma 5.20.

Proof. If $\mathsf{pc}(M_1) = \bot$ then $M_1 \mathrel{-\!\!\circ} M_2 \equiv \mathsf{elim}_{\mathrel{-\!\!\circ}}(M_1, M_2) \equiv \bot$. Also, since M_1 and M_2 are E-complete for $\mathrm{Var}(M_1 \cup M_2)$, if we suppose that $M_1^e \neq M_2^e$ then $M_1 \mathrel{-\!\!\circ} M_2 \equiv \mathsf{elim}_{\mathrel{-\!\!\circ}}(M_1, M_2) \equiv \bot$. From now on, we will assume that $\mathsf{pc}(M_1) = M_1$ and $M_1^e = M_2^e$.

- **$M_1 \mathrel{-\!\!\circ} M_2 \models \mathsf{elim}_{\mathrel{-\!\!\circ}}(M_1, M_2)$.** Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a structure such that $\mathcal{S} \models M_1 \mathrel{-\!\!\circ} M_2$. Then there exists a heap $\mathfrak{h}'$ disjoint from $\mathfrak{h}$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2$. Below we prove that $\mathcal{S}$ is also a model of the formulæ (12), (13), (14) and (15), respectively.
  (12) We have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1^e \wedge M_1^u \wedge M_1^f$, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_1^e \wedge M_1^u \wedge M_1^f$ by Proposition 5.3, and by Proposition 5.15, we deduce that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{pc}(M_1)^e \wedge \mathsf{dc}(M_1)^u \wedge M_1^f$. Analogously, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_2^e \wedge \mathsf{dc}(M_2)^u \wedge M_2^f x$ follows from $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2$ by Propositions 5.3 and 5.15.
  (13) Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1$, also $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \mathsf{alloc}(\mathsf{av}(M_1))$ and since $\mathrm{dom}(\mathfrak{h}') \cap \mathrm{dom}(\mathfrak{h}) = \emptyset$, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{nalloc}(\mathsf{av}(M_1))$. To prove that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{fp}_{\mathsf{nv}(M_1)}(M_2)$, we consider four cases, depending on the form of the literal:

– If $\mathsf{alloc}(x) \in M_2$ and $x \in \mathsf{nv}(M_1)$, then $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}) \cup \mathrm{dom}(\mathfrak{h}')$ and $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}')$, thus $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{alloc}(x)$, by Proposition 4.2.
– The cases $x \hookrightarrow \mathbf{y} \in M_2$ and $x \in \mathsf{nv}(M_1)$ use a similar argument.
– If $\neg\mathsf{alloc}(x) \in M_2$ and $x \in \mathsf{nv}(M_1)$, then $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h} \cup \mathfrak{h}')$, hence $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\mathsf{alloc}(x)$, by Proposition 4.2.
– If $\neg x \hookrightarrow \mathbf{y} \in M_2$ and $x \in \mathsf{nv}(M_1)$ then $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}')$ and either:
  * $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg x \hookrightarrow \mathbf{y}$, by Proposition 4.2, or
  * $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h})$ in which case $\mathfrak{h}' \uplus \mathfrak{h}$ and $\mathfrak{h}$ agree on $\mathfrak{s}(x)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg x \hookrightarrow \mathbf{y}$.

(14) We have $||\mathfrak{h} \uplus \mathfrak{h}'|| = ||\mathfrak{h}|| + ||\mathfrak{h}'||$ and since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2$, we obtain $\min^{\mathcal{S}}_{M_2} \leq ||\mathfrak{h}|| + ||\mathfrak{h}'|| < \max^{\mathcal{S}}_{M_2}$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1$ we also have $\min^{\mathcal{S}}_{M_1} \leq ||\mathfrak{h}'|| < \max^{\mathcal{S}}_{M_1}$, thus $\min^{\mathcal{S}}_{M_1} \leq ||\mathfrak{h}'|| \leq \max^{\mathcal{S}}_{M_1} - 1$, i.e., $-\max^{\mathcal{S}}_{M_1} + 1 \leq -||\mathfrak{h}'|| \leq -\min^{\mathcal{S}}_{M_1}$ so that $\min^{\mathcal{S}}_{M_2} - \max^{\mathcal{S}}_{M_1} + 1 \leq ||\mathfrak{h}|| < \max^{\mathcal{S}}_{M_2} - \min^{\mathcal{S}}_{M_1}$.
(15) Assume that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{nalloc}(Y)$ for a set $Y \subseteq \mathrm{Var}(M_1 \cup M_2)$, which implies that $\mathrm{dom}(\mathfrak{h}) \cap \mathfrak{s}(Y) = \emptyset$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1$, we also have $\mathrm{dom}(\mathfrak{h}') \cap \mathfrak{s}(\mathsf{nv}(M_1)) = \emptyset$. Thus $||\mathfrak{U}|| \geq ||\mathfrak{h}|| + ||\mathfrak{h}'|| + ||\mathfrak{s}(Y \cap \mathsf{nv}(M_1))|| \geq ||\mathfrak{h}|| + \min^{\mathcal{S}}_{M_1} + \#_{\mathrm{n}}(Y, M_1)$, because $||\mathfrak{h}'|| \geq \min^{\mathcal{S}}_{M_1}$ and $||\mathfrak{s}(Y \cap \mathsf{nv}(M_1))|| = |Y \cap \mathsf{nv}(M_1)|_{M_1} = \#_{\mathrm{n}}(Y, M_1)$, by Proposition 5.5, since M_1 is E-complete. Therefore, $||\mathfrak{h}|| \leq ||\mathfrak{U}|| - \min^{\mathcal{S}}_{M_1} - \#_{\mathrm{n}}(Y, M_1)$. Moreover, since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2$, we obtain $|U| \geq ||\mathfrak{h} \uplus \mathfrak{h}'|| + \#_{\mathrm{n}}(Y, M_1) \geq \min^{\mathcal{S}}_{M_2} + \#_{\mathrm{n}}(Y, M_1)$.

- **$\mathsf{elim}_{-\!\circ}(M_1, M_2) \models M_1 \mathrel{-\!\!\circ} M_2$.** Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ be a structure such that $\mathcal{S} \models \mathsf{elim}_{-\!\circ}(M_1, M_2)$. We build a heap $\mathfrak{h}'$ such that $\mathrm{dom}(\mathfrak{h}) \cap \mathrm{dom}(\mathfrak{h}') = \emptyset$, $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2$. First, for each variable $x \in \mathsf{av}(M_1)$ such that $x' \hookrightarrow \mathbf{y} \in M_1^p$ for some variable x' with $x \approx_{M_1} x'$, we add the tuple $(\mathfrak{s}(x), \mathfrak{s}(\mathbf{y}))$ to $\mathfrak{h}'$. Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathrm{pc}(M_1)^e$, for any pair of variables $x \approx_{M_1} x'$ if $x \hookrightarrow \mathbf{y}, x' \hookrightarrow \mathbf{y}' \in M_1$ then $y_i \approx_{M_1} y_i'$, and the result is a functional relation. We define:

$$\begin{array}{rcl} A & = & \{x \in \mathsf{av}(M_1) \mid \forall x' \forall \mathbf{y} \,.\, x \approx_{M_1} x' \rightarrow x' \hookrightarrow \mathbf{y} \notin M_1^p\} \\ V_x & = & \{(\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k)) \in \mathfrak{U}^k \mid x \approx_{M_1} x',\ \neg x' \hookrightarrow \mathbf{y} \in M_1^p\},\ \text{for } x \in \mathsf{av}(M_1) \\ N & = & \{x \in \mathrm{Var}(M_1 \cup M_2) \mid \mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h})\} \end{array}$$

Intuitively, A denotes the set of variables that must be allocated but with no constraint on their image; this set is independent of the interpretation under consideration. The set V_x denotes the set of images the allocated variable x cannot point to, and N denotes the set of variables that are not allocated in $\mathfrak{h}$.
For each $x \in A$ we choose a tuple $(\ell_1, \ldots, \ell_k) \in \mathfrak{U}^k \setminus V_x$ and let $\mathfrak{h}'(\mathfrak{s}(x)) = (\ell_1, \ldots, \ell_k)$. Since M_1 is E-complete, we have $||V_x|| \leq \delta_x(M_1)$ for each $x \in A$, and such a choice is possible because $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathrm{dc}(M_1)^u$, thus $||\mathfrak{U}^k|| \geq \delta_x(M_1) + 1$.
Since $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \mathsf{nalloc}(N)$, if $\mathfrak{U}$ is finite, by (14) it must be the case that:

$$||\mathfrak{h}|| < ||\mathfrak{U}|| - \min^{\mathcal{S}}_{M_1} - \#_{\mathrm{n}}(N, M_1) + 1 \tag{16}$$

$$||\mathfrak{U}|| \geq \min^{\mathcal{S}}_{M_2} + \#_{\mathrm{n}}(N, M_1) \tag{17}$$

Finally, let $L \subseteq \mathfrak{U} \setminus (\mathrm{dom}(\mathfrak{h}) \cup \mathfrak{s}(\mathsf{av}(M_1)) \cup \mathfrak{s}(\mathsf{nv}(M_1)))$ be a finite set of locations of cardinality $||L|| = \max(\min^{\mathcal{S}}_{M_1}, \min^{\mathcal{S}}_{M_2} - ||\mathfrak{h}||) - \#_{\mathrm{a}}(M_1)$. Choosing such a set L is possible, because either $\mathfrak{U}$ is infinite, or $\mathfrak{U}$ is finite, in which case:

$$\begin{array}{rcl} ||\mathfrak{U}|| & \geq & \max(\min^{\mathcal{S}}_{M_1} + ||\mathfrak{h}||, \min^{\mathcal{S}}_{M_2}) + \#_{\mathrm{n}}(N, M_1), \text{ by (16) and (17)} \\ & \geq & \max(\min^{\mathcal{S}}_{M_1}, \min^{\mathcal{S}}_{M_2} - ||\mathfrak{h}||) - \#_{\mathrm{a}}(M_1) + ||\mathfrak{h}|| + \#_{\mathrm{a}}(M_1) + \#_{\mathrm{n}}(N, M_1) \\ & = & ||L|| + ||\mathfrak{h}|| + \#_{\mathrm{a}}(M_1) + \#_{\mathrm{n}}(N, M_1) \\ & \geq & ||L|| + ||\mathrm{dom}(h) \cup \mathfrak{s}(\mathsf{av}(M_1)) \cup \mathfrak{s}(\mathsf{nv}(M_1))|| \end{array}$$

where the last inequality is a consequence of Proposition 5.5. We choose an arbitrary tuple $(\ell_1, \ldots, \ell_k) \in \mathfrak{U}^k$ and let $\mathfrak{h}'(\ell) = (\ell_1, \ldots, \ell_k)$ for all $\ell \in L$. Because $\mathfrak{U}$ is non-empty, such a tuple exists. Consequently, we have $\mathrm{dom}(\mathfrak{h}') = \mathfrak{s}(\mathrm{av}(M_1)) \cup L$ and $\mathrm{dom}(\mathfrak{h}') \cap \mathrm{dom}(\mathfrak{h}) = \emptyset$ because $\mathfrak{s}(\mathrm{av}(M_1)) \cap \mathrm{dom}(\mathfrak{h}) = \emptyset$ by (13) and $L \cap \mathrm{dom}(\mathfrak{h}) = \emptyset$ by construction. We now prove:

- $\boldsymbol{(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1}$. Clearly $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_1^e \wedge M_1^u \wedge M_1^f$ by (12) and Proposition 5.15. To show $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models M_1^a$, observe that $\mathfrak{s}(x) \in \mathrm{dom}(\mathfrak{h}')$ for each $x \in \mathrm{av}(M_1)$, hence for each literal $\mathrm{alloc}(x) \in M_1$ we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \mathrm{alloc}(x)$. Moreover, we have $\mathrm{dom}(\mathfrak{h}') \cap \mathfrak{s}(\mathrm{nv}(M_1)) = (\mathfrak{s}(\mathrm{av}(M_1)) \cup L) \cap \mathfrak{s}(\mathrm{nv}(M_1)) = \emptyset$, because M_1 is footprint consistent and E-complete for $\mathrm{Var}(M_1 \cup M_2)$, by Proposition 5.11. Thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \neg\mathrm{alloc}(x)$ for each literal $\neg\mathrm{alloc}(x) \in M_1^a$. For each literal $x \hookrightarrow \mathbf{y} \in M_1^p$ we have $\mathfrak{h}'(\mathfrak{s}(x)) = (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$ by construction, thus $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models x \hookrightarrow \mathbf{y}$. For each literal $\neg x \hookrightarrow \mathbf{y} \in M_1^p$, we distinguish two cases.
  - If $x \in \mathrm{av}(M_1)$, then $(\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k)) \in V_x$ hence $\mathfrak{h}(\mathfrak{s}(x)) \neq (\mathfrak{s}(y_1), \ldots, \mathfrak{s}(y_k))$ by construction.
  - If $x \notin \mathrm{av}(M_1)$, then since M_1 is A-complete for $\mathrm{Var}(M_1 \cup M_2)$, we have $x \in \mathrm{nv}(M_1)$, thus $\mathfrak{s}(x) \notin \mathrm{dom}(\mathfrak{h}') = \mathfrak{s}(\mathrm{av}(M_1)) \cup L$.

  We finally prove that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models |h| \geq \min_{M_1} \wedge |h| < \max_{M_1}$. Since $\mathrm{dom}(\mathfrak{h}') = \mathfrak{s}(\mathrm{av}(M_1)) \cup L$ and $\mathfrak{s}(\mathrm{av}(M_1)) \cap L = \emptyset$, we have $||\mathfrak{h}'|| = ||\mathfrak{s}(\mathrm{av}(M_1))|| + ||L|| = \max(\min^{\mathcal{S}}_{M_1}, \min^{\mathcal{S}}_{M_2} - ||\mathfrak{h}||)$. If $||\mathfrak{h}'|| = \min^{\mathcal{I}}_{M_1}$ then $||\mathfrak{h}'|| < \max^{\mathcal{I}}_{M_1}$ because $\mathcal{S} \models \mathrm{dc}(M_1)^u$, which implies that $\min^{\mathcal{S}}_{M_1} < \max^{\mathcal{S}}_{M_1}$, by Proposition 5.14. Otherwise $||\mathfrak{h}'|| = \min^{\mathcal{S}}_{M_2} - ||\mathfrak{h}|| \geq \min^{\mathcal{I}}_{M_1}$ and we have by (14) $||\mathfrak{h}|| \geq \min^{\mathcal{I}}_{M_2} - \max^{\mathcal{I}}_{M_1} + 1$, thus $||\mathfrak{h}|| > \min^{\mathcal{I}}_{M_2} - \max^{\mathcal{I}}_{M_1}$, and therefore $||\mathfrak{h}'|| < \max^{\mathcal{I}}_{M_1}$.
- $\boldsymbol{(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}' \uplus \mathfrak{h}) \models M_2}$. We have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}' \uplus \mathfrak{h}) \models M_2^e \wedge M_2^f \wedge M_2^u$ because $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models M_2^e \wedge M_2^f \wedge M_2^u$ and these formulæ do not depend on the heap. Next, for a given variable x, let $\alpha_x \in \{\mathrm{alloc}(x), \neg\mathrm{alloc}(x), x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y} \mid \mathbf{y} \in \mathrm{Var}^k\} \cap M_2$ be a literal and let $\overline{\alpha}_x$ denote its complement. If $x \in \mathrm{nv}(M_1)$ then $\alpha_x \in \mathrm{fp}_{\mathrm{nv}(M_1)}(M_2)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \alpha_x$ by (13). Moreover, because $\mathfrak{h}$ and $\mathfrak{h} \uplus \mathfrak{h}'$ agree on $\mathfrak{s}(\mathrm{nv}(M_1))$, we obtain $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models \alpha_x$. Otherwise $x \notin \mathrm{nv}(M_1)$ hence $x \in \mathrm{av}(M_1)$ because M_1 is A-complete for $\mathrm{Var}(M_1 \cup M_2)$, and since $\alpha_x \in M_2^a \cup M_2^p$ and $\mathrm{atoms}(M_2^a \cup M_2^a) \subseteq \mathrm{atoms}(M_1^a \cup M_1^p)$, we have $\alpha_x \in \mathrm{fp}_a(M_1)$, because the case $\overline{\alpha}_x \in \mathrm{fp}_a(M_1)$ is in contradiction with $\mathrm{fp}_a(M_1) \cap \overline{M_2} = \emptyset$ (condition (c) of the Lemma). But then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}') \models \alpha_x$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models \alpha_x$ follows, by Proposition 5.7. We have thus proved that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h} \uplus \mathfrak{h}') \models M_2^a \cup M_2^p$. We are left with proving that $\min^{\mathcal{S}}_{M_2} \leq ||\mathfrak{h}|| + ||\mathfrak{h}'|| = \max(\min^{\mathcal{I}}_{M_1} + ||\mathfrak{h}||, \min^{\mathcal{S}}_{M_2}) < \max^{\mathcal{S}}_{M_2}$. If $\min^{\mathcal{S}}_{M_1} + ||\mathfrak{h}|| \leq \min^{\mathcal{S}}_{M_2}$ the result follows from the fact that $\mathcal{S} \models \mathrm{dc}(M_2)^u$, which implies $\min^{\mathcal{S}}_{M_2} < \max^{\mathcal{S}}_{M_2}$, by Proposition 5.14. Otherwise, $||\mathfrak{h}|| + ||\mathfrak{h}'|| = \min^{\mathcal{S}}_{M_1} + ||\mathfrak{h}|| > \min^{\mathcal{S}}_{M_2}$ and $||\mathfrak{h}|| + ||\mathfrak{h}'|| < \max^{\mathcal{S}}_{M_2}$ follows from (14).

□

Example 5.21. Let $M_1 = \{\mathrm{alloc}(x), \neg\mathrm{alloc}(y), \neg x \approx y, |h| \geq 1, |h| < 2\}$, $M_2 = \{\neg x \approx y, |h| \geq 3, |h| < \infty, \neg x \hookrightarrow x, \neg y \hookrightarrow y\}$. Then $M_1 \mathrel{-\!\!*} M_2 \equiv^{inf} \{|h| \geq 2, |h| < \infty, \neg\mathrm{alloc}(x), \neg y \hookrightarrow y\}$. ■

5.3 Translating Quantifier-free SL^k into Minterms

We prove next that each quantifier-free SL^k formula is equivalent to a finite disjunction of minterms. Intuitively, these disjunctions are defined by induction on the structure of the formula. The base cases and classical connectives are easy to handle. For formulæ $\psi_1 * \psi_2$ or $\psi_1 \mathrel{-\!\!*} \psi_2$, the transformation is first applied on ψ_1 and ψ_2, then the following equivalences are used to shift $*$ and $\mathrel{-\!\!*}$ innermost in the formula:

$$(\phi_1 \vee \phi_2) * \phi \equiv (\phi_1 * \phi) \vee (\phi_2 * \phi) \qquad (\phi_1 \vee \phi_2) \mathrel{-\!\!*} \phi \equiv (\phi_1 \mathrel{-\!\!*} \phi) \vee (\phi_2 \mathrel{-\!\!*} \phi)$$
$$\phi * (\phi_1 \vee \phi_2) \equiv (\phi * \phi_1) \vee (\phi * \phi_2) \qquad \phi \mathrel{-\!\!*} (\phi_1 \vee \phi_2) \equiv (\phi \mathrel{-\!\!*} \phi_1) \vee (\phi \mathrel{-\!\!*} \phi_2)$$

Afterwards, the operands of $*$ and $-\!\!\circ$ are minterms, and the result is obtained using the equivalences from Lemmas 5.16 and 5.20, respectively (up to a transformation into disjunctive normal form). The only difficulty is that these lemmas impose some additional conditions on the minterms (e.g., being E-complete, or A-complete). However, the conditions are easy to enforce by case splitting, as illustrated by Example 5.22.

Example 5.22. Consider the formula $x \mapsto x \mathrel{-\!\!\circ} y \mapsto y$. It is easy to check that $x \mapsto x \equiv M_1$, where $M_1 = x \hookrightarrow x \wedge |h| \geq 1 \wedge |h| < 2$ and $y \mapsto y \equiv M_2$, where $M_2 = y \hookrightarrow y \wedge |h| \geq 1 \wedge |h| < 2$. To apply Lemma 5.20, we need to ensure that M_1 and M_2 are E-complete, which may be done by adding either $x \approx y$ or $x \not\approx y$ to each minterm. We also have to ensure that M_1 is A-complete, thus for $z \in \{x, y\}$, we add either $\mathrm{alloc}(z)$ or $\neg\mathrm{alloc}(z)$ to M_1. Finally, we must have $\mathrm{atoms}(M_2^a \cup M_2^p) \subseteq \mathrm{atoms}(M_1^a \cup M_1^p)$, thus we add either $y \hookrightarrow y$ or $\neg y \hookrightarrow y$ to M_1. After removing redundancies, we get (among others) the minterms: $M_1' = x \hookrightarrow x \wedge |h| \geq 1 \wedge |h| < 2 \wedge x \approx y$ and $M_2' = y \hookrightarrow y \wedge |h| \geq 1 \wedge |h| < 2 \wedge x \approx y$. Afterwards we compute $\mathrm{elim}^{fin}_{-\!\!\circ}(M_1', M_2') = x \approx y \wedge \neg\mathrm{alloc}(x) \wedge |h| \geq 0 \wedge |h| < 1$. ■

To describe the transformation in a more formal way, we first need to show that the conjunction of two minterms can be written as a disjunction of minterms. To this aim, given minterms M_1 and M_2, we define the sets of constraints $\mathrm{minh}(M_1, M_2)$ and $\mathrm{maxh}(M_1, M_2)$ by taking the conjunction of the lower and upper bounds on the cardinality of the heap and keeping the most restrictive bounds.

Definition 5.23.

$$\mathrm{minh}(M_1, M_2) \stackrel{\text{def}}{=} \begin{cases} \{|h| \geq \max(\min_{M_1}, \min_{M_2})\} \\ \quad \text{if } \min_{M_1}, \min_{M_2} \in \mathbb{N} \\ \left\{ \begin{array}{l} |h| \geq \min_{M_i} \wedge |U| < \min_{M_i} + m + 1, \\ |h| \geq \min_{M_{3-i}} \wedge |U| \geq \min_{M_i} + m + 1 \end{array} \right\} \\ \quad \text{if } \min_{M_i} \in \mathbb{N}, \min_{M_{3-i}} = |U| - m,\ i = 1, 2 \\ \{|h| \geq |U| - \min(m_1, m_2)\} \\ \quad \text{if } \min_{M_i} = |U| - m_i,\ i = 1, 2 \end{cases}$$

$$\mathrm{maxh}(M_1, M_2) \stackrel{\text{def}}{=} \begin{cases} \{|h| < \min(\max_{M_1}, \max_{M_2})\} \\ \quad \text{if } \max_{M_1}, \max_{M_2} \in \mathbb{N}_\infty \\ \{|h| < \max_{M_i}\} \\ \quad \text{if } \max_{M_{3-i}} = \infty, \max_{M_i} = |U| - m,\ i = 1, 2 \\ \left\{ \begin{array}{l} |h| < \max_{M_i} \wedge |U| \geq \max_{M_i} + m, \\ |h| < |U| - m \wedge |U| < \max_{M_i} + m \end{array} \right\} \\ \quad \text{if } \max_{M_i} \in \mathbb{N}, \max_{M_{3-i}} = |U| - m,\ i = 1, 2 \\ \{|h| < |U| - \max(m_1, m_2)\} \\ \quad \text{if } \max_{M_i} = |U| - m_i,\ i = 1, 2 \end{cases}$$

For instance, if $M_1 = \{|h| \geq 2, |h| < |U| - 1\}$ and $M_2 = \{|h| \geq 3, |h| < |U| - 2\}$, then $\mathrm{minh}(M_1, M_2) = \{|h| \geq 3\}$ and $\mathrm{maxh}(M_1, M_2) = \{|h| < |U| - 2\}$. Heterogeneous constraints are merged by performing a case split on the value of $|U|$. For example, if $M_1 = \{|h| \geq |U| - 4\}$ and $M_2 = \{|h| \geq 1\}$, then the first condition prevails if $|U| \geq 5$ yielding: $\mathrm{minh}(M_1, M_2) = \{|h| \geq 1 \wedge |U| < 5, |h| \geq |U| - 4 \wedge |U| \geq 5\}$. The disjunction of minterms equivalent to a conjunction of two minterms is then defined as follows:

Definition 5.24. For any minterms M_1, M_2, let $[M_1, M_2] \stackrel{\text{def}}{=} \left\{ \bigwedge_{i=1,2} M_i^e \wedge M_i^f \wedge M_i^a \wedge M_i^p \wedge M_i^u \wedge \mu \wedge \nu \mid \mu \in \mathrm{minh}(M_1, M_2), \nu \in \mathrm{maxh}(M_1, M_2) \right\}$. We extend this notation recursively to any set of minterms of size $n > 2$: $[M_1, M_2, \ldots, M_n] \stackrel{\text{def}}{=} \bigcup_{M \in [M_1, \ldots, M_{n-1}]} [M, M_n]$.

Proposition 5.25. *Given minterms $M_1, \ldots, M_n$, we have $\bigwedge_{i=1}^n M_i \equiv \bigvee_{M \in [M_1, \ldots, M_n]} M$.*

Proof. We prove the result for $n = 2$, the general result follows by induction. For $n = 2$, this is a consequence of the fact that $|h| \geq \min_{M_1} \wedge |h| \geq \min_{M_2} \equiv \bigvee_{\mu \in \mathrm{minh}(M_1, M_2)} \mu$, and $|h| < \max_{M_1} \wedge |h| < \max_{M_2} \equiv \bigvee_{\nu \in \mathrm{maxh}(M_1, M_2)} \nu$. We prove the first fact in the case where $\min_{M_1} = m_1$ and $\min_{M_2} = |U| - m_2$, the other cases are similar. Consider a structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ such that $\mathcal{S} \models |h| \geq m_1 \wedge |h| \geq |U| - m_2$. Then $||\mathfrak{h}|| \geq m_1$ and $||\mathfrak{h}|| \geq ||\mathfrak{U}|| - m_2$, and we distinguish two cases.

- if $m_1 \geq ||U|| - m_2$, then necessarily $||\mathfrak{U}|| < m_1 + m_2 + 1$, so that $\mathcal{S} \models |h| \geq m_1 \wedge |U| < m_1 + m_2 + 1$.
- otherwise, we have $||U|| \geq m_1 + m_2 + 1$, so that $\mathcal{S} \models |h| \geq |U| - m_2 \wedge |U| \geq m_1 + m_2 + 1$.

Conversely, if $\mathcal{S}$ is a structure such that either $\mathcal{S} \models |h| \geq m_1 \wedge |U| < m_1 + m_2 + 1$ or $\mathcal{S} \models |h| \geq |U| - m_2 \wedge |U| \geq m_1 + m_2 + 1$, then it is straightforward to verify that $\mathcal{S} \models |h| \geq m_1 \wedge |h| \geq |U| - m_2$. □

The following proposition states some properties of the literals occurring in $[M_1, \ldots, M_n]$.

Proposition 5.26. *Given minterms $M_1, \ldots, M_n$ and $M \in [M_1, \ldots, M_n]$, if $\ell \in M$ is a literal then either $\ell \in M_i$, for some $i = 1, \ldots, n$, or $\ell \in \{|U| \geq m_1 + m_2, |U| < m_1 + m_2, |U| \geq m_1 + m_2 + 1, |U| < m_1 + m_2 + 1\}$, where $M_1 \cup \cdots \cup M_n$ contains two literals $\ell_i \in \{|h| \geq m_i, |h| < m_i, |h| \geq |U| - m_i, |h| < |U| - m_i\}$, for $i = 1, 2$.*

Proof. Assume that $n = 2$. If $\ell \notin M_1 \cup M_2$ then by definition of $[M_1, M_2]$, necessarily ℓ occurs in $\mathrm{minh}(M_1, M_2) \cup \mathrm{maxh}(M_1, M_2)$ and the proof is immediate, by definition of these sets. The proof for $n > 2$ goes by induction on n. □

For two sets K, L of literals, a *completion* of K w.r.t. L is a set of literals K' that is minimal with respect to inclusion of sets, such that $K \subseteq K'$ and $\mathrm{atoms}(L) \subseteq \mathrm{atoms}(K)$ (i.e., $K \subseteq K'$ and for every $\ell \in L$, K' contains either ℓ or $\overline{\ell}$). We denote by $(K)^L$ the set of completions of K w.r.t. L.

Proposition 5.27. *If K and L are sets of literals, then $K \equiv \bigvee_{\psi \in (K)^L} \psi$. If further K is a minterm and L contains no literals of the form $|h| \geq t$ or $|h| < t$, then every set $P \in (K)^L$ is a minterm such that $\mathrm{Var}(P) = \mathrm{Var}(K) \cup \mathrm{Var}(L)$, $\min_P = \min_K$ and $\max_P = \max_K$.*

Proof. Immediate, by the definition of $(K)^L$. □

For a literal ℓ, let $[\ell]^{\mathrm{mt}}$ be an equivalent minterm obtained from ℓ by adding the missing lower/upper bounds on the cardinality of the heap, namely $|h| \geq 0$ if $\ell \notin \{|h| \geq n, |h| \geq |U| - n \mid n \in \mathbb{Z}\}$ and $|h| < \infty$ if $\ell \notin \{|h| < n, |h| < |U| - n \mid n \in \mathbb{Z}\}$. We extend this notation to sets (i.e., conjunctions) of literals as $[\ell_1, \ldots, \ell_n]^{\mathrm{mt}} \stackrel{\mathrm{def}}{=} [[\ell_1]^{\mathrm{mt}}, \ldots, [\ell_n]^{\mathrm{mt}}]$. We have $\ell \equiv [\ell]^{\mathrm{mt}}$ for any literal ℓ and $L \equiv \bigvee_{M \in [L]^{\mathrm{mt}}} M$, for any set L of literals. For a boolean combination of literals ϕ, we denote by $(\phi)^{\mathrm{dnf}}$ its disjunctive normal form. We assume from now on that the disjunctive normal form of a formula is canonical and all the conjunctions are incomparable with respect to logical entailment.

Given a formula ϕ in disjunctive normal form $\phi = \bigvee_{i=1}^{n} C_i$, where $C_1, \ldots, C_n$ are conjunctions (represented by sets) of literals, we define $[\phi]^{\mathrm{mt}} \stackrel{\mathrm{def}}{=} \bigcup_{i=1}^{n} [C_i]^{\mathrm{mt}}$. We have $[\phi]^{\mathrm{mt}} \equiv \bigvee_{M \in [\phi]^{\mathrm{mt}}} M$. Further, let $\mathsf{E}(L) \stackrel{\mathrm{def}}{=} \{x \approx y \mid x, y \in \mathrm{Var}(L)\}$ and $\mathsf{A}(L) \stackrel{\mathrm{def}}{=} \{\mathsf{alloc}(x) \mid x \in \mathrm{Var}(L)\}$, for a set L of literals.

For each $\dagger \in \{\mathit{fin}, \mathit{inf}\}$, we define the set of minterms $\mu^{\dagger}(\phi)$ recursively on the structure of ϕ:

$$\mu^{\dagger}(\mathsf{emp}) \stackrel{\mathrm{def}}{=} \{|h| \simeq 0\} \qquad \mu^{\dagger}(x \mapsto \mathbf{y}) \stackrel{\mathrm{def}}{=} \{x \hookrightarrow \mathbf{y} \wedge |h| \simeq 1\}$$

$$\mu^{\dagger}(x \approx y) \stackrel{\mathrm{def}}{=} \{x \approx y \wedge |h| \geq 0 \wedge |h| < \infty\} \qquad \mu^{\dagger}(q(x_1, \ldots, x_{\#(q)}) \stackrel{\mathrm{def}}{=} \left\{q(x_1, \ldots, x_{\#(q)}) \wedge |h| \geq 0 \wedge |h| < \infty\right\}$$

$$\mu^\dagger(\phi_1 \wedge \phi_2) \stackrel{\text{def}}{=} \bigcup_{\substack{M_i \in \mu^\dagger(\phi_i) \\ i=1,2}} [M_1, M_2]$$

$$\mu^\dagger(\neg\phi_1) \stackrel{\text{def}}{=} \bigcup \left\{ \left[\overline{\ell_1}, \ldots, \overline{\ell_n}\right]^{\text{mt}} \middle| \ell_i \in M_i,\ i \in [1,n] \right\}, \text{ where } \mu^\dagger(\phi_1) = \{M_1, \ldots, M_n\}$$

$$\mu^\dagger(\phi_1 * \phi_2) \stackrel{\text{def}}{=} \bigcup_{\substack{M_i \in \mu^\dagger(\phi_i) \\ i=1,2}} \left\{ \left[(\mathsf{elim}_*(P_1, P_2))^{\text{dnf}}\right]^{\text{mt}} \middle| N_j \in (M_j)^{\mathsf{E}(M_1 \cup M_2)}, P_j \in (N_j)^{N_{3-j}^p},\ j = 1,2 \right\}$$

$$\mu^\dagger(\phi_1 \mathrel{-\!\!\circ} \phi_2) \stackrel{\text{def}}{=} \bigcup_{\substack{M_i \in \mu^\dagger(\phi_i) \\ i=1,2}} \left\{ \left[\left(\mathsf{elim}^\dagger_{\mathrel{-\!\!\circ}}(Q_1, N_2)\right)^{\text{dnf}}\right]^{\text{mt}} \middle| N_j \in (M_j)^{\mathsf{E}(M_1 \cup M_2)}, P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}, Q_1 \in (P_1)^{M_2^a \cup M_2^p},\ j = 1,2 \right\}$$

Intuitively, $\mu^\dagger(\phi_1 * \phi_2)$ and $\mu^\dagger(\phi_1 \mathrel{-\!\!\circ} \phi_2)$ are obtained by first recursively computing $\mu^\dagger(\phi_1)$ and $\mu^\dagger(\phi_2)$, then extending the obtained minterms in such a way that the hypotheses of Lemmas 5.16 or 5.20 are satisfied, and finally applying $\mathsf{elim}^\dagger_*$ and $\mathsf{elim}^\dagger_{\mathrel{-\!\!\circ}}$, respectively.

Lemma 5.28. *Given a quantifier-free* SL^k *formula* ϕ*, the following equivalences hold: (1)* $\phi \equiv^{fin} \bigvee_{M \in \mu^{fin}(\phi)} M$*, and (2)* $\phi \equiv^{inf} \bigvee_{M \in \mu^{inf}(\phi)} M$*.*

Proof. We show that $\phi \equiv^{fin} \bigvee_{M \in \mu^{fin}(\phi)} M$ by induction on the structure of ϕ. The fact that $\phi \equiv^{inf} \bigvee_{M \in \mu^{inf}(\phi)} M$ is proved in the same way. The base cases are immediate and the inductive cases are dealt with below:

- If $\phi = \phi_1 \wedge \phi_2$ and $\phi_i \equiv^{fin} \bigvee_{M_i \in \mu^{fin}(\phi_i)} M_i$ for $i = 1, 2$ by the inductive hypothesis and Proposition 5.25, we have:
$$\begin{array}{lll} \phi & \equiv^{fin} & \bigvee_{M_1 \in \mu^{fin}(\phi_1)} M_1 \wedge \bigvee_{M_2 \in \mu^{fin}(\phi_2)} M_2 \\ & \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} M_1 \wedge M_2 \\ & \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{M \in [M_1, M_2]} M \end{array}$$
- If $\phi = \neg\phi_1$, $\mu^{fin}(\phi_1) = \{M_1, \ldots, M_n\}$, $M_i = \{\ell_{i1}, \ldots, \ell_{in_i}\}$ for all $i \in [1,n]$, then since $\phi_1 \equiv^{fin} \bigvee_{i=1}^n \bigwedge_{j=1}^{n_i} \ell_{ij}$ by the inductive hypothesis, we have:
$$\begin{array}{lll} \neg\phi_1 & \equiv^{fin} & \bigwedge_{i=1}^n \bigvee_{j=1}^{n_i} \overline{\ell_{ij}} \\ & \equiv^{fin} & \bigwedge_{i=1}^n \bigvee_{j=1}^{n_i} \left[\overline{\ell_{ij}}\right]^{\text{mt}} \\ & \equiv^{fin} & \bigvee \left\{ \left[\overline{\ell_1}\right]^{\text{mt}} \wedge \ldots \wedge \left[\overline{\ell_n}\right]^{\text{mt}} \middle| \ell_i \in M_i,\ i \in [1,n] \right\} \\ & \equiv^{fin} & \bigvee \left\{ \left[\overline{\ell_1}, \ldots, \overline{\ell_n}\right]^{\text{mt}} \middle| \ell_i \in M_i,\ i \in [1,n] \right\} \end{array}$$
- If $\phi = \phi_1 * \phi_2$ and $\phi_i \equiv^{fin} \bigvee_{M \in \mu^{fin}(\phi_i)} M$ for $i = 1, 2$ by the induction hypothesis, we compute successively[7]:
$$\begin{array}{lll} & & (\phi_1 * \phi_2)\ [\text{distributivity of } * \text{ with } \vee] \\ & \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} M_1 * M_2 \\ & & \left[\text{because } M_i \equiv \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} N_i\right] \\ & \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} N_1 * N_2 \\ & & \left[\text{because } N_i \equiv \bigvee_{P_i \in (N_i)^{N_{3-i}^p}} P_i\right] \\ & \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} \\ & & \bigvee_{P_i \in (N_i)^{N_{3-i}^p}} P_1 * P_2 \end{array}$$

[7] See Definition 5.2 for the definition of N^p.

At this point, observe that N_i, and thus P_i, are E-complete for $\mathrm{Var}(M_1 \cup M_2)$, for $i = 1, 2$. Moreover, $\mathrm{atoms}(P_1^p) = \mathrm{atoms}(P_2^p)$, because $P_i \in (N_i)^{N_{3-i}^p}$, for $i = 1, 2$. We can thus apply Lemma 5.16 and infer that:

$$\begin{array}{rcl} P_1 * P_2 & \equiv & \mathsf{elim}_*(P_1, P_2) \\ & \equiv & (\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}} \\ & \equiv & \bigvee_{M \in \left[(\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}}\right]^{\mathsf{mt}}} M \end{array}$$

- If $\phi = \phi_1 \mathrel{-\!\!\circ} \phi_2$ and $\phi_i \equiv^{fin} \bigvee_{M \in \mu^{fin}(\phi_i)} M$, $i = 1, 2$, by the induction hypothesis, we compute, successively:

$$\begin{array}{rl} & (\phi_1 \mathrel{-\!\!\circ} \phi_2) \text{ [distributivity of } \mathrel{-\!\!\circ} \text{ with } \vee] \\ \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} M_1 \mathrel{-\!\!\circ} M_2 \\ & \left[\text{because } M_i \equiv \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} N_i\right] \\ \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} N_1 \mathrel{-\!\!\circ} N_2 \\ & \left[\text{because } N_1 \equiv \bigvee_{P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}} P_1\right] \\ \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} \\ & \bigvee_{P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}} P_1 \mathrel{-\!\!\circ} N_2 \\ & \left[\text{because } P_1 \equiv \bigvee_{Q_1 \in (P_1)^{N_2^a \cup N_2^p}} Q_1\right] \\ \equiv^{fin} & \bigvee_{M_i \in \mu^{fin}(\phi_i),\ i=1,2} \bigvee_{N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}} \\ & \bigvee_{P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}} \bigvee_{Q_1 \in (P_1)^{N_2^a \cup N_2^p}} Q_1 \mathrel{-\!\!\circ} N_2 \end{array}$$

Observe that N_i and thus P_i are E-complete for $\mathrm{Var}(M_1 \cup M_2)$, for $i = 1, 2$. Moreover, P_1 is A-complete for $\mathrm{Var}(M_1 \cup M_2)$, because $P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}$ and $\mathrm{atoms}(N_2^a \cup N_2^p) \subseteq \mathrm{atoms}(Q_1^a \cup Q_1^p)$, because $Q_1 \in (P_1)^{N_2^a \cup N_2^p}$. Then we can apply Lemma 5.20 and infer that:

$$\begin{array}{rcl} Q_1 \mathrel{-\!\!\circ} N_2 & \equiv^{fin} & \mathsf{elim}^{fin}_{\mathrel{-\!\!\circ}}(Q_1, N_2) \\ & \equiv & \left(\mathsf{elim}^{fin}_{\mathrel{-\!\!\circ}}(Q_1, N_2)\right)^{\mathsf{dnf}} \\ & \equiv & \bigvee_{M \in \left[\left(\mathsf{elim}^{fin}_{\mathrel{-\!\!\circ}}(Q_1, N_2)\right)^{\mathsf{dnf}}\right]^{\mathsf{mt}}} M \end{array}$$

□

As explained in Section 4.3, boolean combinations of minterms can only be transformed into sat-equivalent BSR(FO) formulæ if there is no positive occurrence of test formulæ $|h| \geq |U| - n$ or $\mathsf{alloc}(x)$ (see Definition 4.8 and the second item of Lemma 4.9). Consequently, we relate the polarity of these formulæ in some minterm $M \in \mu^{fin}(\phi) \cup \mu^{inf}(\phi)$ with that of a separating implication within ϕ. The analysis depends on whether the universe is finite or infinite.

Lemma 5.29. *For any quantifier-free* SL^k *formula* ϕ*, the following properties hold:*

(1) For all $M \in \mu^{inf}(\phi)$*, we have* $M \cap \{|h| \geq |U| - n, |h| < |U| - n \mid n \in \mathbb{N}\} = \emptyset$.

(2) If $|h| \geq |U| - n \in M$ *(resp.* $|h| < |U| - n \in M$*) for some minterm* $M \in \mu^{fin}(\phi)$*, then a formula* $\psi_1 \mathrel{-\!\!*} \psi_2$ *occurs at a positive (resp. negative) polarity in* ϕ.

(3) If $\mathsf{alloc}(x) \in M$ *(resp.* $\neg\mathsf{alloc}(x) \in M$*) for some minterm* $M \in \mu^{inf}(\phi)$*, then a formula* $\psi_1 \mathrel{-\!\!*} \psi_2$*, such that* $x \in \mathrm{Var}(\psi_1) \cup \mathrm{Var}(\psi_2)$*, occurs at a positive (resp. negative) polarity in* ϕ.

(4) If $M \cap \{\mathsf{alloc}(x), \neg\mathsf{alloc}(x) \mid x \in \mathrm{Var}\} \neq \emptyset$ *for some minterm* $M \in \mu^{fin}(\phi)$*, then a formula* $\psi_1 \mathrel{-\!\!*} \psi_2$*, such that* $x \in \mathrm{Var}(\psi_1) \cup \mathrm{Var}(\psi_2)$*, occurs in* ϕ *at some polarity* $p \in \{-1, 1\}$*. Moreover,* $\mathsf{alloc}(x)$ *occurs at a polarity* $-p$*, only if* $\mathsf{alloc}(x)$ *is in the scope of a* λ^{fin} *subformula*[8] *of a formula* $\mathsf{elim}^{fin}_{\mathrel{-\!\!\circ}}(M_1, M_2)$ *used to compute* $\bigvee_{M \in \mu^{fin}(\phi)} M$.

[8] See equation (15) in Lemma 5.20.

Proof.

(1) By induction on the structure of ϕ, one shows that no literal from $\{|h| \geq |U| - n, |h| < |U| - n \mid n \in \mathbb{N}\}$ is introduced during the construction of $\mu^{inf}(\phi)$.

(2) Let $\ell \in M \cap \{|h| \geq |U| - n, |h| < |U| - n \mid n \in \mathbb{N}\}$ be a literal. The proof is by induction on the structure of ϕ:

- The cases $\phi = \mathsf{emp}$, $\phi = x \hookrightarrow \mathbf{y}$, $\phi = q(\mathbf{x})$ and $\phi = x \approx y$ are trivial, because $\ell \notin \mu^{fin}(\phi)$.
- $\phi = \phi_1 \wedge \phi_2$: we have $M \in [M_1, M_2]$, for some minterms $M_i \in \mu^{fin}(\phi_i)$, for $i = 1, 2$. By Proposition 5.26, since $\ell \notin \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\}$, we deduce that $\ell \in M_1 \cup M_2$ and the proof follows from the induction hypothesis, since any formula occurring in ϕ_i, $i = 1, 2$, occurs at the same polarity in ϕ.
- $\phi = \neg\phi_1$: assuming $\mu^{fin}(\phi_1) = \{M_1, \ldots, M_m\}$, we have $M \in \left[\overline{\ell_1}, \ldots, \overline{\ell_m}\right]^{\mathsf{mt}}$, for some literals $\ell_i \in M_i$, $i \in [1, m]$. By Proposition 5.26, we deduce that $\ell = \overline{\ell_i}$ for some $i = 1, \ldots, n$, because $\ell \notin \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\}$. By the induction hypothesis, there exists a formula $\psi_1 \mathrel{-\!\!*} \psi_2$ occurring at polarity $p \in \{1, -1\}$ in ϕ_1, where $p = 1$ if $\ell_i = |h| \geq |U| - n$ and $p = -1$ if $\ell_i = |h| < |U| - n$. Then ℓ occurs at polarity $-p$ in M and $\psi_1 \mathrel{-\!\!*} \psi_2$ occurs at polarity $-p$ in ϕ.
- $\phi = \phi_1 * \phi_2$: for $i = 1, 2$, there exist minterms $M_i \in \mu^{fin}(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$ and $P_i \in (N_i)^{N_{3-i}^p}$, such that $M \in \left[(\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}}\right]^{\mathsf{mt}}$. Since by hypothesis $\ell \in \{|h| \geq |U| - n, |h| < |U| - n \mid n \in \mathbb{N}\}$, by Proposition 5.26, this literal is necessarily introduced by $\mathsf{elim}_*(P_1, P_2)$ and, by inspection of $\mathsf{elim}_*(P_1, P_2)$, one of the following must hold:
  - $\ell = |h| \geq \min_{M_1} + \min_{M_2}$, where $\min_{M_1}$ and/or $\min_{M_2}$ is of the form $|U| - n$. By the induction hypothesis ϕ_i contains a formula $\psi_1 \mathrel{-\!\!*} \psi_2$ at polarity 1, for some $i = 1, 2$, and the proof is completed.
  - $\ell = |h| < \max_{M_1} + \max_{M_2} - 1$, where $\max_{M_1}$ and/or $\max_{M_2}$ is of the form $|U| - n$. The proof is similar, with polarity -1.
  - $\ell = |h| \geq \#_{\mathsf{a}}(M_i) + |Y|_{M_i} + \min_{M_j}$, where $\min_{M_j}$ is of the form $|U| - n$. The proof is similar.
- $\phi = \phi_1 \mathrel{-\!\!\circ} \phi_2 = \neg(\phi_1 \mathrel{-\!\!*} \neg\phi_2)$: there exist minterms $M_i \in \mu^{fin}(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$, for $i = 1, 2$, $P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}$ and $Q_1 \in (P_1)^{M_2^a \cup M_2^p}$, such that $M \in \left[\left(\mathsf{elim}_{\mathrel{-\!\!\circ}}^{fin}(Q_1, N_2)\right)^{\mathsf{dnf}}\right]^{\mathsf{mt}}$. By inspection of $\mathsf{elim}_{\mathrel{-\!\!\circ}}^{fin}(Q_1, N_2)$, one of the following cases must occur:
  - $\ell = |h| \geq \min_{M_2} - \max_{M_1} - 1$, where $\min_{M_2}$ is of the form $|U| - n_2$. By the induction hypothesis, ϕ_2 contains a formula $\psi_1 \mathrel{-\!\!*} \psi_2$ at polarity 1, and this formula also occurs at polarity 1 in ϕ, thus the proof is completed. Note that if $\max_{M_1} = |U| - n_1$ then either $\min_{M_2} = |U| - n_2$ and $|h| \geq \min_{M_2} - \max_{M_1} - 1 = |h| \geq n_1 - n_2$, or $\min_{M_2} = n_2 \in \mathbb{N}$ and $|h| \geq \min_{M_2} - \max_{M_1} - 1 = |h| \geq -|U| + (n_1 + n_2) = \bigwedge_{1 \leq n < n_1 + n_2} |U| \simeq n \rightarrow |h| \geq n_1 + n_2 - n$ by Definition 4.3, thus $|h| \geq \min_{M_2} - \max_{M_1} - 1$ contains no literal of the above form.
  - $\ell = |h| < \max_{M_2} - \min_{M_1}$. The proof is similar.
  - $\ell = |h| < |U| - \min_{M_1} - \#_{\mathsf{n}}(Y, M_1) + 1$. In this case since $(\phi_1 \mathrel{-\!\!*} \neg\phi_2)$ occurs at polarity -1 in ϕ, the proof is completed.

(3) Let $\ell \in M \cap \{\mathsf{alloc}(x), \neg\mathsf{alloc}(x) \mid x \in \mathsf{Var}\}$ be a literal occurring in some minterm $M \in \mu^{inf}(\phi)$. The proof is by induction on the structure of ϕ:

- The cases $\phi = \mathsf{emp}$, $\phi = x \hookrightarrow \mathbf{y}$, $\phi = x \approx y$ and $\phi = q(\mathbf{x})$ are trivial, because $\ell \notin \mu^{inf}(\phi)$.
- The cases $\phi = \phi_1 \wedge \phi_2$ and $\phi = \neg\phi_1$ are similar to point (2) of the Lemma.
- $\phi = \phi_1 * \phi_2$: there exist minterms $M_i \in \mu^{inf}(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$ and $P_i \in (N_i)^{N_{3-i}^p}$, such that $M \in \left[(\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}}\right]^{\mathsf{mt}}$, for all $i = 1, 2$. By inspection of $\mathsf{elim}_*(P_1, P_2)$, one of the following cases must occur:
  - $\ell = \neg\mathsf{alloc}(x)$ with $x \in \mathsf{nv}(M_1) \cap \mathsf{nv}(M_2)$. Assuming that the definition of $\mathsf{elim}_*(P_1, P_2)$ is changed according to Remark 5.19, it must be the case that $\neg\mathsf{alloc}(x)$ occurs at a positive polarity in M_1 or M_2. Then, by the induction hypothesis ϕ_i contains a subformula $\psi_1 \mathrel{-\!\!*} \psi_2$ at polarity -1 with $x \in \mathsf{Var}(\psi_1) \cup \mathsf{Var}(\psi_2)$. But then $\psi_1 \mathrel{-\!\!*} \psi_2$ also occurs at polarity -1 in ϕ and the proof is completed.
  - $\ell = \neg\mathsf{alloc}(x)$ with $x \in Y \subseteq \mathsf{nv}(M_j)$. Similar to the previous case.

- $\phi = \phi_1 \multimap \phi_2 = \neg(\phi_1 \mathrel{-\!\!*} \neg\phi_2)$: there exist minterms $M_i \in \mu^{inf}(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$, for $i = 1, 2$, $P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}$ and $Q_1 \in (P_1)^{M_2^a \cup M_2^p}$, such that $M \in \left[\left(\mathsf{elim}^{inf}_{\multimap}(Q_1, N_2)\right)^{\mathsf{dnf}}\right]^{\mathsf{mt}}$. By inspection of $\mathsf{elim}^{inf}_{\multimap}(Q_1, N_2)$, the only possible case is $\ell = \neg\mathsf{alloc}(x)$ with $x \in \mathsf{av}(M_1)$ (Equation (13) in Lemma 5.20), thus $x \in \mathrm{Var}(\phi_1) \cup \mathrm{Var}(\phi_2)$ and $(\phi_1 \mathrel{-\!\!*} \neg\phi_2)$ occurs at polarity -1 in ϕ, which completes the proof.

(4) The proof is similar to point (3). The only difference is that $\mathsf{alloc}(x)$ may occur in the λ^{fin} subformula (Equation (15) in Lemma 5.20) of the $\mathsf{elim}^{fin}_{\multimap}(Q_1, N_2)$, in which case its polarity may be different from that of $\phi_1 \mathrel{-\!\!*} \phi_2$. □

Note that Property 3 in Lemma 5.29 does not hold for $\mu^{fin}(\phi)$:

Example 5.30. Consider a fixed number $n \geq 1$, as well as the following formulæ:

$$\begin{array}{rcl} \phi & \stackrel{\text{def}}{=} & |h| \simeq U - n \\ \psi_1 & \stackrel{\text{def}}{=} & (\neg\mathsf{alloc}(x) \wedge |h| \simeq n) \mathrel{-\!\!*} \bot \\ \psi_2 & \stackrel{\text{def}}{=} & \mathsf{alloc}(x) \end{array}$$

We verify that $\psi_2 \wedge \phi \equiv^{fin} \neg\psi_1 \wedge \phi$:

- If $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi_2 \wedge \phi$, then $\mathfrak{s}(x)$ is allocated in $\mathfrak{h}$ and there are exactly n unallocated cells. Then the heap $\mathfrak{h}'$ whose domain is the set of unallocated cells in $\mathfrak{h}$ is disjoint from $\mathfrak{h}$ and satisfies $\neg\mathsf{alloc}(x) \wedge |h| \simeq n$, which proves that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\psi_1$.
- If $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \neg\psi_1 \wedge \phi$, then there are exactly n unallocated cells in $\mathfrak{U}$, and there exists a heap $\mathfrak{h}'$ disjoint from $\mathfrak{h}$ with n elements in its domain, non of which is $\mathfrak{s}(x)$. Thus, $\mathfrak{s}(x)$ must occur in the domain of $\mathfrak{h}$, and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h}) \models \psi_2$.

However, the polarity of $\mathsf{alloc}(x)$ is positive in ψ_2, whereas x only occurs in the scope of neutral occurrences of $\mathrel{-\!\!*}$ in $\neg\psi_1$. ■

We provide another example illustrating Property 4.

Example 5.31. Let $M_1 = \{|h| \geq 0, |h| < 2, \neg\mathsf{alloc}(x)\}$ and $M_2 = \{|h| \geq 0, |h| < \infty, \neg x \hookrightarrow x\}$. We have $M_1 \multimap M_2 \equiv^{fin} \neg x \approx y \wedge |h| \geq 0 \wedge |h| < |U| \wedge \neg\mathsf{alloc}(x) \rightarrow (|U| \geq 2 \wedge |h| < |U| - 1)$. The last two formulæ are parts of λ^{fin} in Lemma 5.20: $|h| < |U|$ ensures that there exists at least one free location (so that there exists a disjoint heap satisfying M_1), and if x is not allocated, then there must actually exist 2 free locations, since x cannot be allocated in the extension. Observe that $\mathsf{alloc}(x)$ occurs positively in the latter formula (since it is in scope of 2 negations), whereas x only occurs in the scope of negative (or neutral) occurrences of $\mathrel{-\!\!*}$ in $M_1 * M_2$ (i.e., positive occurrences of $\multimap$). This happens because $\mathsf{alloc}(x)$ occurs in λ^{fin}. ■

5.4 Testing Membership in $\mu^{\dagger}(\phi)$ in PSPACE

Given a quantifier-free SL^k formula ϕ, the number of minterms occurring in $\mu^{fin}(\phi)$ (resp. $\mu^{inf}(\phi)$) is exponential in the size of ϕ, in the worst case. Therefore, an optimal decision procedure cannot generate and store these sets explicitly, but rather must enumerate minterms lazily. We show that (i) the size of the minterms in $\mu^{fin}(\phi) \cup \mu^{inf}(\phi)$ is bounded by a polynomial in the size of ϕ, and that (ii) the problem "*given a minterm M, does M occur in $\mu^{fin}(\phi)$ (resp. in $\mu^{inf}(\phi)$)?*" is in PSPACE. To this aim, we define a measure on a quantifier-free formula ϕ, which bounds

the size of the minterms in the sets $\mu^{fin}(\phi)$ and $\mu^{inf}(\phi)$, inductively on the structure of the formulæ:

$$\begin{array}{rcl@{\qquad}rcl}
\mathcal{M}(\top) & \stackrel{\text{def}}{=} & 0 & \mathcal{M}(\bot) & \stackrel{\text{def}}{=} & 0 \\
\mathcal{M}(x \approx y) & \stackrel{\text{def}}{=} & 0 & \mathcal{M}(q(\mathbf{x})) & \stackrel{\text{def}}{=} & 0 \\
\mathcal{M}(\text{emp}) & \stackrel{\text{def}}{=} & 1 & \mathcal{M}(x \mapsto \mathbf{y}) & \stackrel{\text{def}}{=} & 2 \\
\mathcal{M}(\neg\phi_1) & \stackrel{\text{def}}{=} & \mathcal{M}(\phi_1) & \mathcal{M}(\phi_1 \wedge \phi_2) & \stackrel{\text{def}}{=} & \max(\mathcal{M}(\phi_1), \mathcal{M}(\phi_2)) \\
\mathcal{M}(\phi_1 * \phi_2) & \stackrel{\text{def}}{=} & \sum_{i=1}^{2}(\mathcal{M}(\phi_i) + ||\text{Var}(\phi_i)||) & \mathcal{M}(\phi_1 \mathrel{-\!\!*} \phi_2) & \stackrel{\text{def}}{=} & \sum_{i=1}^{2}(\mathcal{M}(\phi_i) + ||\text{Var}(\phi_i)||)
\end{array}$$

The intuition is that $\mathcal{M}(\phi)$ is an upper bound on natural number occurring in the test formulæ in $\mu^{fin}(\phi) \cup \mu^{inf}(\phi)$, when viewed as linear inequalities on $|U|$ and $|h|$. For instance, $\mathcal{M}(\text{emp})$ is 1, because $\text{emp} \equiv |h| < 1$, whereas $\mathcal{M}(x \mapsto \mathbf{y})$ is 2, because $\mathcal{M}(x \mapsto \mathbf{y}) \equiv x \hookrightarrow \mathbf{y} \wedge |h| \geq 1 \wedge |h| < 2$. The extension to the standard connectives is straightforward, but the handling of the separating connectives is more involved: first, the combination of two inequalities may increase the bound (for instance, $|h| \geq 1 * |h| \geq 2 \equiv |h| \geq 3$) and second, the elimination of these connectives yields additional inequalities (see Lemma 5.16 and Lemma 5.20).

Proposition 5.32. *For any $n \in \mathbb{N}$, we have:*

$$\mathcal{M}(|h| \geq n) = \mathcal{M}(|U| \geq n) = n$$
$$\mathcal{M}(|h| \geq |U| - n) = n + 1$$

Proof. By induction on $n \geq 0$. □

Note that, because $|h| < \infty$ is a shorthand for $\top$, we have $\mathcal{M}(|h| < \infty) = 0$.

Definition 5.33. A minterm M is *$\mathcal{M}$-bounded* by a formula ϕ, if for each literal $\ell \in M$, the following hold: (i) $\mathcal{M}(\ell) \leq \mathcal{M}(\phi)$ if $\ell \in \{|h| \geq \min_{M_i}, |h| < \max_{M_i}\}$; (ii) $\mathcal{M}(\ell) \leq 2\mathcal{M}(\phi) + 1$, if $\ell \in \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\}$.

Proposition 5.34. *Given minterms $M_1, \ldots, M_n$ all $\mathcal{M}$-bounded by ϕ, each minterm $M \in [M_1, \ldots, M_n]$ is also $\mathcal{M}$-bounded by ϕ.*

Proof. This is an immediate corollary of Proposition 5.26. □

The following lemma provides the required result:

Lemma 5.35. *Given a quantifier-free SL^k formula ϕ, each minterm $M \in \mu^{fin}(\phi) \cup \mu^{inf}(\phi)$ is $\mathcal{M}$-bounded by ϕ.*

Proof. We prove that each $M \in \mu^{fin}(\phi)$ is $\mathcal{M}$-bounded by ϕ. The proof for $M \in \mu^{inf}(\phi)$ follows from the observation that, because of the definition of $\text{elim}^{inf}_{-\circ}$, for each $M \in \mu^{inf}(\phi)$ there exists $M' \in \mu^{fin}(\phi)$ such that $\mathcal{M}(M) \leq \mathcal{M}(M')$. By induction on the structure of ϕ:

- If $\phi = \text{emp}$ then $\mu^{fin}(\phi) = \{|h| \geq 0 \wedge |h| < 1\}$, $\mathcal{M}(|h| \geq 0) = 0$, $\mathcal{M}(|h| < 1) = \mathcal{M}(|h| \geq 1) = 1$ and $\mathcal{M}(\text{emp}) = 1$, by definition.
- If $\phi = x \mapsto \mathbf{y}$ then $\mu^{fin}(\phi) = \{x \hookrightarrow \mathbf{y} \wedge |h| \geq 1 \wedge |h| < 2\}$, $\mathcal{M}(|h| \geq 1) = 1$, $\mathcal{M}(|h| < 2) = 2$ and $\mathcal{M}(x \mapsto \mathbf{y}) = 2$, by definition.
- If $\phi = q(\mathbf{y})$ with $q \in \mathcal{F}$ then $\mu^{fin}(\phi) = \{q(\mathbf{y}) \wedge |h| \geq 0 \wedge |h| < \infty\}$, $\mathcal{M}(|h| \geq 0) = 0$, $\mathcal{M}(|h| < \infty) = 0$ and $\mathcal{M}(q(\mathbf{y})) = 0$, by definition.
- If $\phi = x \approx y$ then $\mu^{fin}(\phi) = \{x \approx y \wedge |h| \geq 0 \wedge |h| < \infty\}$ and $\mathcal{M}(|h| \geq 0) = \mathcal{M}(|h| < \infty) = 0$, by definition.
- If $\phi = \phi_1 \wedge \phi_2$, let $\ell \in M$ be a literal, where $M \in \mu^{fin}(\phi_1 \wedge \phi_2)$ is a minterm. Then $M \in [M_1, M_2]$, for some minterms $M_i \in \mu^{fin}(\phi_i)$, $i = 1, 2$ and the proof follows from Proposition 5.34, because M_i is $\mathcal{M}$-bounded by ϕ_i and $\mathcal{M}(\phi_i) \leq \mathcal{M}(\phi)$, so that M_i is $\mathcal{M}$-bounded by ϕ, for $i = 1, 2$.

- If $\phi = \neg\phi_1$ assume that $\mu^{fin}(\phi_1) = \{M_1, \ldots, M_m\}$. Let $\ell \in M$ be a literal, where $M \in \mu^{fin}(\neg\phi_1)$ is a minterm. Then $M \in \left[\left[\overline{\ell_1}\right]^{\text{mt}}, \ldots, \left[\overline{\ell_n}\right]^{\text{mt}}\right]$, for some literals $\ell_i \in M_i$, $i \in [1, m]$. By the induction hypothesis, $\overline{\ell_i}$ is $\mathcal{M}$-bounded by ϕ, for every $i \in 1, \ldots, n$, thus the same holds for ℓ_i. Since $\mathcal{M}(|h| \geq 0) = \mathcal{M}(|h| < \infty) = 0$, we deduce that $\left[\overline{\ell_i}\right]^{\text{mt}}$ is $\mathcal{M}$-bounded by ϕ, and the proof follows from Proposition 5.34.
- If $\phi = \phi_1 * \phi_2$, let $\ell \in M$ be a literal, where $M \in \mu^{fin}(\phi_1 * \phi_2)$. Then there exist minterms $M_i \in \mu^{fin}(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$ and $P_i \in (N_i)^{N^P_{3-i}}$, such that $M \in \left[(\text{elim}_*(P_1, P_2))^{\text{dnf}}\right]^{\text{mt}}$, for $i = 1, 2$. First assume that ℓ is of the form $|h| \geq t$ or $|h| < t$. We only consider the case where ℓ occurs in $\text{elim}_*(P_1, P_2)$, the rest of the cases follow from Proposition 5.34. We distinguish the following:
  - ℓ is a subformula of $|h| \geq \min_{P_1} + \min_{P_2} = |h| \geq \min_{M_1} + \min_{M_2}$, because $\min_{P_i} = \min_{M_i}$, for $i = 1, 2$, by Proposition 5.27. By the inductive hypothesis we have $\mathcal{M}(|h| \geq \min_{M_i}) \leq \mathcal{M}(\phi_i)$, for $i = 1, 2$. If $\min_{M_i} \in \mathbb{N}$ for $i = 1, 2$ then $\ell = |h| \geq \min_{M_1} + \min_{M_2}$ and we have:
$$\begin{array}{rcl} \mathcal{M}(\ell) = \mathcal{M}(|h| \geq \min_{M_1} + \min_{M_2}) & = & \mathcal{M}(|h| \geq \min_{M_1}) + \mathcal{M}(|h| \geq \min_{M_2}) \\ & \leq & \mathcal{M}(\phi_1) + \mathcal{M}(\phi_2) \leq \mathcal{M}(\phi). \end{array}$$
    If $\min_{M_i} = |U| - n_i$ and $n_i, \min_{M_{3-i}} \in \mathbb{N}$, then $\ell = |h| \geq \min_{M_1} + \min_{M_2}$ and we obtain:
$$\begin{array}{rcl} \mathcal{M}(\ell) = \mathcal{M}(|h| \geq \min_{M_1} + \min_{M_2}) & = & \mathcal{M}(|h| \geq |U| - (n_i - \min_{M_{3-i}})) \\ & \leq & \mathcal{M}(|h| \geq |U| - n_i) \\ & \leq & \mathcal{M}(\phi_i) \leq \mathcal{M}(\phi). \end{array}$$
    Otherwise, $\min_{M_i} = |U| - n_i$, for $i = 1, 2$, where $n_1, n_2 \in \mathbb{N}$, thus by Definition 4.3:
$$\begin{array}{rcl} |h| \geq \min_{M_1} + \min_{M_2} & = & |h| \geq 2 \cdot |U| - n_1 - n_2 \\ & = & |U| < 1 + n_1 + n_2 \wedge \\ & & \bigwedge_{1 \leq n \leq n_1 + n_2} |U| \simeq n \rightarrow |h| \geq 2n - n_1 - n_2 \end{array}$$
    and
    - either $\ell \in \{|U| \geq n, |U| < n + 1\}$ for some $n \in [1, n_1 + n_2]$, and we have $\mathcal{M}(\ell) \leq n + 1 \leq n_1 + n_2 + 1 \leq 2(\mathcal{M}(\phi_1) + \mathcal{M}(\phi_2)) + 1 = 2\mathcal{M}(\phi) + 1$;
    - or $\ell = |h| \geq 2n - n_1 - n_2$ for some $n \in [1, n_1 + n_2]$, and we have $\mathcal{M}(\ell) = 2n - n_1 - n_2 \leq n_1 + n_2 = \mathcal{M}(\phi_1) + \mathcal{M}(\phi_2) = \mathcal{M}(\phi)$.
  - The proof in the case where ℓ is a subformula of $|h| < \max_{M_1} + \max_{M_2} - 1$ is analogous.
  - $\ell = |h| \geq \#_{\text{a}}(P_i) + |Y|_{P_i} + \min_{P_{3-i}}$, where $Y \subseteq \text{nv}(P_{3-i}) \setminus \text{av}(P_i)$, for some $i = 1, 2$. Because $Y \cap \text{av}(P_i) = \emptyset$, we have (Definition 5.6 and Proposition 5.27): $\#_{\text{a}}(P_i) + |Y|_{P_i} \leq ||\text{Var}(P_i)|| + ||\text{Var}(P_{3-i})|| \leq ||\text{Var}(\phi_1)|| + ||\text{Var}(\phi_2)||$ and thus $\mathcal{M}(\ell) \leq \mathcal{M}(|h| \geq \min_{P_{3-i}}) + ||\text{Var}(\phi_1)|| + ||\text{Var}(\phi_2)|| \leq \mathcal{M}(\phi)$.

  Now assume $\ell \in \{|U| \geq m, |U| < m \mid m \in \mathbb{N}\}$. Then one of the following holds:
  - $\ell \in \text{dc}(P_i)^u$, for some $i = 1, 2$, and we have two cases:
    - $\ell \in \{|U| \geq n_1 + n_2 + 1, |U| < n_1 + n_2\}$, where $\min_{P_i} = \min_{M_i} = n_1$ and $\max_{P_i} = \max_{M_i} = |U| - n_2$. By the induction hypothesis, we have $n_1, n_2 \leq \mathcal{M}(\phi_i)$, thus $\mathcal{M}(\ell) \leq 2\mathcal{M}(\phi_i) + 1 \leq 2\mathcal{M}(\phi) + 1$.
    - $\ell = |U| \geq \left\lceil \sqrt[k]{\max_{x \in \text{av}(M)}(\delta_x(P_i) + 1)} \right\rceil$, in which case either $\text{Var}(M_1) \cup \text{Var}(M_2) = \emptyset$ so that we have $\left\lceil \sqrt[k]{\max_{x \in \text{av}(M)}(\delta_x(P_i) + 1)} \right\rceil = 0$ and the proof is immediate, or we have $\mathcal{M}(\ell) \leq \sqrt[k]{||\text{Var}(M_i)||^k + 1} \leq |\text{Var}(M_i)| + 1 \leq 2\mathcal{M}(\phi) + 1$.
  - $\ell = |U| > n_i + \#_{\text{a}}(P_i) + |Y|_{M_i}$, where $Y \subseteq \text{nv}(M_{3-i}) \setminus \text{av}(M_i)$ and $\max_{M_i} = |U| - n_i$, for some $i = 1, 2$. Because $Y \cap \text{av}(P_i) = \emptyset$, we have $\#_{\text{a}}(P_i) + |Y|_{P_i} \leq ||\text{Var}(P_i)|| + ||\text{Var}(P_{3-i})|| \leq ||\text{Var}(\phi_1)|| + ||\text{Var}(\phi_2)||$ and thus $\mathcal{M}(\ell) \leq \mathcal{M}(\phi_i) + ||\text{Var}(\phi_1)|| + ||\text{Var}(\phi_2)|| \leq 2\mathcal{M}(\phi) + 1$.
- If $\phi = \phi_1 \multimap \phi_2$, consider a literal $\ell \in M$, where $M \in \mu^{fin}(\phi_1 \multimap \phi_2)$. Then there exist minterms $M_i \in \mu^{fin}(\phi_i)$ and $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$, for $i = 1, 2$, and minterms $P_1 \in (N_1)^{\mathsf{A}(M_1 \cup M_2)}$ and $Q_1 \in (P_1)^{M_2^a \cup M_2^p}$, such that

$M \in \left[\left(\mathsf{elim}^{fin}_{-\circ}(Q_1, N_2)\right)^{\mathsf{dnf}}\right]^{\mathsf{mt}}$. We only consider the case where ℓ occurs in $\mathsf{elim}^{fin}_{-\circ}(Q_1, N_2)$, in the remaining cases, the result follows directly from Proposition 5.34. If ℓ is of the form $|h| \geq t$ or $|h| < t$ then either:

- ℓ is a subformula of $|h| \geq \min_{N_2} - \max_{Q_1} - 1 = |h| \geq \min_{M_2} - \max_{M_1} - 1$, because $\min_{N_2} = \min_{M_2}$ and $\max_{Q_1} = \max_{P_1} = \max_{N_1} = \max_{M_1}$ by Proposition 5.27. Then $\min_{M_2} \in \{n_2, |U| - n_2\}$ and $\max_{M_1} \in \{n_1, |U| - n_1\}$ with $n_1, n_2 \in \mathbb{N}_\infty$, and by the induction hypothesis $n_i \leq \mathcal{M}(\phi_i)$. If $\max_{M_1} = n_1$ or $\min_{M_2} \neq n_2$, then by an inspection of the different cases and using Proposition 5.32, we have $\ell = |h| \geq \min_{M_2} - \max_{M_1} + 1$, thus:
$$\mathcal{M}(\ell) = \mathcal{M}(|h| \geq \min_{M_2} - \max_{M_1} + 1) \leq n_1 + n_2 \leq \mathcal{M}(\phi_1) + \mathcal{M}(\phi_2) \leq \mathcal{M}(\phi)$$
Otherwise, $\min_{M_2} = n_2$ and $\max_{M_1} = |U| - n_1$ hence:
  - either $\ell \in \{|U| \geq n, |U| < n + 1\}$, for some $n \in [1, n_1 + n_2 - 1]$ and we have $\mathcal{M}(\ell) \leq n + 1 \leq n_1 + n_2 \leq 2(\mathcal{M}(\phi_1) + \mathcal{M}(\phi_2)) + 1 = \mathcal{M}(\phi)$;
  - or $\ell = |h| \geq n_1 + n_2 - n$, for some $n \in [1, n_1 + n_2 - 1]$ and we have $\mathcal{M}(\ell) = n_1 + n_2 - n \leq n_1 + n_2 - 1 \leq \mathcal{M}(\phi_1) + \mathcal{M}(\phi_2) = \mathcal{M}(\phi)$.
- The case $\ell = |h| < \max_{N_2} - \min_{Q_1}$ is proved in a similar way.
- $\ell = |h| < |U| - \min_{Q_1} - \#_{\mathrm{n}}(Y, Q_1) + 1$, for some $Y \subseteq \mathrm{Var}(Q_1 \cup N_2)$. Because $\mathsf{nv}(Q_1) \subseteq \mathsf{nv}(P_1) \subseteq \mathrm{Var}(\phi_1) \cup \mathrm{Var}(\phi_2)$, we have $\#_{\mathrm{n}}(Y, Q_1) \leq ||\mathrm{Var}(\phi_1)|| + ||\mathrm{Var}(\phi_2)||$. Moreover, $\min_{Q_1} = \min_{M_1}$ by Proposition 5.27. We distinguish the following cases:
  - If $\min_{M_1} \in \mathbb{N}$, we compute:

$$\begin{array}{rcl} \mathcal{M}(\ell) & = & \min_{M_1} + \#_{\mathrm{n}}(Y, Q_1) - 1, \text{ by Proposition 5.32} \\ & \leq & \mathcal{M}(\phi_1) + ||\mathrm{Var}(\phi_1)|| + ||\mathrm{Var}(\phi_2)|| \text{ since } \mathcal{M}(|h| \geq \min_{M_1}) \leq \mathcal{M}(\phi_1), \text{ by the inductive hypothesis.} \\ & \leq & \mathcal{M}(\phi). \end{array}$$

  - Otherwise, $\min_{M_1} = |U| - n_1$, for some $n_1 \in \mathbb{N}$, thus $\ell = |h| < n_1 - \#_{\mathrm{n}}(Y, Q_1) + 1$. By Proposition 5.32, we have $\mathcal{M}(|h| \geq \min_{M_1}) = n_1 + 1$ and $\mathcal{M}(\ell) = n_1 - \#_{\mathrm{n}}(Y, Q_1) + 1$, therefore:
$$\begin{array}{rcl} \mathcal{M}(\ell) & = & \mathcal{M}(|h| \geq \min_{M_1}) - \#_{\mathrm{n}}(Y, Q_1) \\ & \leq & \mathcal{M}(\phi_1) \leq \mathcal{M}(\phi) \end{array}$$

If ℓ is of the form $|U| \geq m$ or $|U| < m$, with $m \in \mathbb{N}$, then either:

- if $\ell \in \mathsf{dc}(Q_1) \cup \mathsf{dc}(N_2)$ the argument is similar to the previous case $\phi = \phi_1 * \phi_2$,
- otherwise, $\ell = |U| \geq \min_{M_2} + \#_{\mathrm{n}}(Y, M_1)$ and either $\min_{M_2} \in \mathbb{N}$, in which case $\mathcal{M}(\ell) = \min_{N_2} + \#_{\mathrm{n}}(Y, Q_1) \leq \mathcal{M}(\phi_2) + ||\mathrm{Var}(\phi_1)|| + ||\mathrm{Var}(\phi_2)|| \leq \mathcal{M}(\phi)$ as in the previous, or $\min_{M_2} = |U| - n_2$, for some $n_2 \in \mathbb{N}$, in which case $\ell \equiv n_2 \geq \#_{\mathrm{n}}(Y, Q_1)$ and $\mathcal{M}(\ell) = 0$.

□

Since $\mathcal{M}(\phi)$ is polynomially bounded by $\mathsf{size}(\phi)$, this entails that it is possible to check whether $M \in \mu^{fin}(\phi)$ (resp. $\mu^{inf}(\phi)$) using space bounded also by a polynomial in $\mathsf{size}(\phi)$.

Proposition 5.36. *Given a quantifier-free* SL^k *formula* ϕ *and a minterm* $M \in \mu^{fin}(\phi) \cup \mu^{inf}(\phi)$*, we have* $\mathsf{size}(M) = O(\mathsf{size}(\phi)^2)$*. As a consequence,* $\mathcal{N}(\bigvee_{M \in \mu^{inf}(\phi)}) = O(\mathsf{size}(\phi)^2)$ *and* $\mathcal{N}(\bigvee_{M \in \mu^{fin}(\phi)}) = O(\mathsf{size}(\phi)^2)$.

Proof. We give the proof for $M \in \mu^{fin}(\phi)$, the case $M \in \mu^{inf}(\phi)$ being similar. Let $\ell \in M$ be a literal. We distinguish the following cases, based on the form of ℓ:

- $\ell \in \{\mathsf{alloc}(x), \neg\mathsf{alloc}(x) \mid x \in \mathrm{Var}\}$: ℓ must occur in ϕ or has been introduced by $\mu^{fin}(.)$, in which case, at most $2 \cdot ||\mathrm{Var}(\phi)||$ such literals are introduced.
- $\ell \in \{x \hookrightarrow \mathbf{y}, \neg x \hookrightarrow \mathbf{y} \mid x \in \mathrm{Var}, \mathbf{y} \in \mathrm{Var}^k\} \cup \{q(\mathbf{z}) \mid q \in \mathcal{F}, \mathbf{z} \in \mathrm{Var}^{\#(q)}\}$: ℓ occurs in ϕ, since $\mu^{fin}(.)$ does not introduce literals of this form.

- $\ell \in \{x \approx y, \neg x \approx y \mid x, y \in \mathsf{Var}\}$: ℓ occurs in ϕ or has been introduced by $\mu^{fin}(.)$, in which case at most $2 \cdot ||\mathsf{Var}(\phi)||^2$ such literals are introduced.
- $\ell \in \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\}$: by Lemma 5.35, $\mathcal{M}(\ell) \leq 2\mathcal{M}(\phi) + 1$, thus $\mathsf{size}(\ell) = O(\mathsf{size}(\phi)^2)$ for each such literal. Furthermore, M contains at most two literals of this form (up to redundancy).
- $\ell \in \{|h| \geq \min_M, |h| < \max_M\}$: by Lemma 5.35, $\mathcal{M}(\ell) \leq \mathcal{M}(\phi)$ and consequently, $\mathsf{size}(\ell) = O(\mathsf{size}(\phi)^2)$ for each such literal. Furthermore, M contains exactly two literals of this form by definition of minterms.

Summing up, we obtain that $\mathsf{size}(M) = O(\mathsf{size}(\phi)^2)$. This second result follows immediately. □

Proposition 5.37. *Let L be a set of literals and ϕ be a boolean combination of literals. The problem of deciding whether $L \in (\phi)^{\mathsf{dnf}}$ is in* $\mathsf{NSPACE}(\mathsf{size}(L) + \mathsf{size}(\phi))$.

Proof. W.l.o.g., we may assume that ϕ is in negation normal form. The algorithm is nondeterministic and proceeds recursively on the structure of ϕ:

- $\phi = \ell$ is a literal: then $(\phi)^{\mathsf{dnf}} = \{\ell\}$ hence it suffices to verify whether $L = \{\ell\}$, using $O(\mathsf{size}(L) + \mathsf{size}(\phi))$ space.
- $\phi = \phi_1 \vee \phi_2$: then $(\phi)^{\mathsf{dnf}} = (\phi_1)^{\mathsf{dnf}} \cup (\phi_2)^{\mathsf{dnf}}$ and we check that one of $L \in (\phi_1)^{\mathsf{dnf}}$ and $L \in (\phi_2)^{\mathsf{dnf}}$ holds. By the induction hypothesis, checking $L \in (\phi_i)^{\mathsf{dnf}}$ can be done using $O(\mathsf{size}(L) + \mathsf{size}(\phi_i))$ space. Since the working space used for $L \in (\phi_1)^{\mathsf{dnf}}$ can be reused for $L \in (\phi_2)^{\mathsf{dnf}}$, the entire check takes $O(\mathsf{size}(L) + \mathsf{size}(\phi))$ space.
- $\phi = \phi_1 \wedge \phi_2$: then $L \in (\phi)^{\mathsf{dnf}} \Leftrightarrow L = L_1 \cup L_2$, with $L_1 \in (\phi_1)^{\mathsf{dnf}}$ and $L_2 \in (\phi_2)^{\mathsf{dnf}}$, thus we guess two subsets L_1 and L_2 with $L_1 \cup L_2 = M$ and check that $L_i \in (\phi_i)^{\mathsf{dnf}}$, using $O(\mathsf{size}(L_i) + \mathsf{size}(\phi_i))$ space, for $i = 1, 2$. Since we must store L_2 during the check $L_1 \in (\phi_1)^{\mathsf{dnf}}$ and the working space can be reused for $L_2 \in (\phi_2)^{\mathsf{dnf}}$, the entire check takes $O(\mathsf{size}(L) + \mathsf{size}(\phi))$ space.

□

Proposition 5.38. *Let L be a set of literals and let M_1, M_2 be minterms. Checking whether $L \in ((\mathsf{elim}_*(M_1, M_2))^{\mathsf{dnf}}$ is in* $\mathsf{NSPACE}(\mathsf{size}(L) + \mathsf{size}(M_1) + \mathsf{size}(M_2))$.

Proof. The algorithm proceeds by induction on the structure of $(\mathsf{elim}_*(M_1, M_2))^{\mathsf{dnf}}$ as in the proof of Proposition 5.37. The only difference concerns the subformulæ η_{ij} (Line 6 in Lemma 5.16) which cannot be constructed explicitly since they are of exponential size. However, η_{ij} is of positive polarity, and to check that $L \in \left(\eta_{ij}\right)^{\mathsf{dnf}}$, it suffices to guess a set of variables $Y \subseteq \mathsf{nv}(M_j) \setminus \mathsf{av}(M_i)$ and check whether:

$$L \in \left(\mathsf{alloc}(Y) \rightarrow (|h| \geq \#_{\mathsf{a}}(M_i) + |Y|_{M_i} + \min_{M_j} \ \wedge \ \#_{\mathsf{a}}(M_i) + |Y|_{M_i} < \max_{M_i})\right)^{\mathsf{dnf}}$$

The size of the above formula is of the order of $O(\mathsf{size}(M_1) + \mathsf{size}(M_2))$, thus $L \in ((\mathsf{elim}_*(M_1, M_2))^{\mathsf{dnf}}$ can be checked in $\mathsf{NSPACE}(\mathsf{size}(L) + \mathsf{size}(M_1) + \mathsf{size}(M_2))$, by Proposition 5.37. □

Proposition 5.39. *Let L be a set of literals and let M_1, M_2 be minterms. The problems whether $L \in \left((\mathsf{elim}^{fin}_{\multimap}(M_1, M_2)\right)^{\mathsf{dnf}}$ and $L \in \left((\mathsf{elim}^{inf}_{\multimap}(M_1, M_2)\right)^{\mathsf{dnf}}$ are both in* $\mathsf{NSPACE}(\mathsf{size}(L) + \mathsf{size}(M_1) + \mathsf{size}(M_2))$.

Proof. The proof is similar to that of Proposition 5.38 (again, the formula $\lambda^{\dagger}$ is exponential, but does not have to be constructed explicitly). □

Proposition 5.40. *Checking whether $M \in [M_1, \ldots, M_n]$, where $M, M_1, \ldots, M_n$ are minterms, $n \geq 2$, is in* $\mathsf{NSPACE}(\mathsf{size}(M) + (\mathsf{size}(M_1) + \ldots + \mathsf{size}(M_n))^2)$.

Proof. The proof is by induction on $n \geq 2$. If $n = 2$ then by definition of $[M_1, M_2]$ it suffices to check that $M = M_1^f \wedge M_1^e \wedge M_1^a \wedge M_1^p \wedge M_1^u \wedge M_2^f \wedge M_2^e \wedge M_2^a \wedge M_2^p \wedge M_2^u \wedge \mu \wedge \nu$ for some $\mu \in \mathsf{minh}(M_1, M_2)$, $\nu \in \mathsf{maxh}(M_1, M_2)$. By definition, the size of each formula in $\mathsf{minh}(M_1, M_2) \cup \mathsf{maxh}(M_1, M_2)$ is of the order of $O(\mathsf{size}(M_1) + \mathsf{size}(M_2))$, thus the algorithm requires $O(\mathsf{size}(M) + \mathsf{size}(M_1) + \mathsf{size}(M_2))$ space.

If $n > 2$, $M \in [M_1, \ldots, M_n] \Leftrightarrow M \in [M', M_n]$, where $M' \in [M_1, \ldots, M_{n-1}]$. By Proposition 5.26, the literals in M' are either literals from $M_1, \ldots, M_{n-1}$ or occur in $\{|U| \geq m_1 + m_2, |U| < m_1 + m_2, |U| \geq m_1 + m_2 + 1, |U| < m_1 + m_2 + 1\}$, where $M_1 \cup \cdots \cup M_{n-1}$ contains two literals ℓ_1 and ℓ_2 and ℓ_i is of the form $|h| \geq m_i, |h| < m_i, |h| \geq |U| - m_i$ or $|h| < |U| - m_i$, for $i = 1, 2$. Thus $\mathsf{size}(M') \leq \sum_{i=1}^{n-1} \mathsf{size}(M_i)$. The nondeterministic algorithm guesses and stores a minterm M_1' of size at most $\sum_{i=1}^{n-1} \mathsf{size}(M_i)$ and checks that $M \in \left[M_1', M_n\right]$ and that $M_1' \in [M_1, \ldots, M_{n-1}]$. According to the base case $n = 2$, the first check takes up $O(\mathsf{size}(M) + \mathsf{size}(M_1') + \mathsf{size}(M_n)) = O(\mathsf{size}(M) + \sum_{i=1}^{n} \mathsf{size}(M_i))$ space, and the second check takes space $O(\mathsf{size}(M_1') + (\sum_{i=1}^{n-1} \mathsf{size}(M_i))^2) = O((\sum_{i=1}^{n} \mathsf{size}(M_i))^2)$, by the induction hypothesis. Because we only need to store M_1' between the two checks, the algorithm takes $O(\mathsf{size}(M) + (\sum_{i=1}^{n} \mathsf{size}(M_i))^2)$ space. □

Proposition 5.41. *Let M be a minterm and let L be a set of literals. The problem of checking whether $M = [L]^{\mathsf{mt}}$ is in* $\mathsf{NSPACE}(\mathsf{size}(M) + (\sum_{\ell \in L} \mathsf{size}(\ell))^2)$.

Proof. By definition, $[L]^{\mathsf{mt}} = [[\ell_1]^{\mathsf{mt}}, \ldots, [\ell_n]^{\mathsf{mt}}]$, with $L = \{\ell_1, \ldots, \ell_n\}$, and each minterm $[\ell_i]^{\mathsf{mt}}$ is of size $O(\mathsf{size}(\ell_i))$, thus the proof follows immediately from Proposition 5.40. □

Lemma 5.42. *Given a minterm M and an SL^k formula ϕ, the problems of checking whether $M \in \mu^{fin}(\phi)$ and $M \in \mu^{inf}(\phi)$ are in PSPACE.*

Proof. We show the existence of a nondeterministic algorithm that decides $M \in \mu^{fin}(\phi)$ in space $O(\mathsf{size}(M) + \mathsf{size}(\phi)^8)$. The PSPACE upper bound is by an application of Savitch's Theorem [21]. We only give the proof for $M \in \mu^{fin}(\phi)$, the proof for $M \in \mu^{inf}(\phi)$ is similar and omitted. By induction on the structure of ϕ, we distinguish the following cases:

- $\phi = \mathsf{emp}$: we check $M = |h| \simeq 0$ in space $O(\mathsf{size}(M) + \mathsf{size}(\phi))$.
- $\phi = x \mapsto \mathsf{y}$: we check $M = \{x \hookrightarrow \mathsf{y} \wedge |h| \simeq 1\}$ in space $O(\mathsf{size}(M) + \mathsf{size}(\phi))$.
- $\phi = q(x_1, \ldots, x_{\#(q)})$: we check $M = \left\{q(x_1, \ldots, x_{\#(q)}) \wedge |h| \geq 0 \wedge |h| < \infty\right\}$ in space $O(\mathsf{size}(M) + \mathsf{size}(\phi))$.
- $\phi = \phi_1 \wedge \phi_2$: $M \in \mu^{fin}(\phi) \Leftrightarrow M \in [M_1, M_2]$ with $M_i \in \mu^{fin}(\phi_i)$, for every $i = 1, 2$. Since, by Proposition 5.36, $\mathsf{size}(M_i) = O(\mathsf{size}(\phi_i)^2) = O(\mathsf{size}(\phi)^2)$, for $i = 1, 2$, it suffices to guess two such minterms M_1 and M_2, check that $M_i \in \mu^{fin}(\phi_i)$, $i = 1, 2$ and that $M \in [M_1, M_2]$. By the induction hypothesis, checking $M_i \in \mu^{fin}(\phi_i)$ requires space $O(\mathsf{size}(M_i) + \mathsf{size}(\phi_i)^8)$, for each $i = 1, 2$, and by the proof of Proposition 5.40 in the case $n = 2$, checking $M \in [M_1, M_2]$ requires space $O(\mathsf{size}(M) + \mathsf{size}(M_1) + \mathsf{size}(M_2)) = O(\mathsf{size}(M) + \mathsf{size}(\phi))$. Since we only need to store M_1 and M_2 between the checks, the entire procedure takes space $O(\mathsf{size}(M) + \mathsf{size}(\phi)^8)$.
- $\phi = \neg\phi_1$: $M \in \mu^{fin}(\phi)$ if and only if $M \in \left[\left[\overline{\ell_1}\right]^{\mathsf{mt}}, \ldots, \left[\overline{\ell_m}\right]^{\mathsf{mt}}\right]$, for some literals $\ell_i \in M_i$, $i \in [1, m]$, where $\mu^{fin}(\phi) = \{M_1, \ldots, M_m\}$. For any $i \in [1, m]$, we distinguish the following cases:
  - if $\ell_i \in \left\{x \hookrightarrow \mathsf{y}, \neg x \hookrightarrow \mathsf{y} \mid x \in \mathsf{Var}, \mathsf{y} \in \mathsf{Var}^k\right\}$ then ℓ_i occurs in ϕ_1, thus there are at most $\mathsf{size}(\phi_1)$ such literals,
  - if $\ell_i \in \{x \approx y, \neg x \approx y \mid x, y \in \mathsf{Var}\}$ then there are at most $2||\mathsf{Var}(\phi)||^2$ such literals,
  - if $\ell_i \in \{|U| \geq n, |U| < n \mid n \in \mathbb{N}\}$, by Lemma 5.35, $\mathcal{M}(\ell_i) \leq 2\mathcal{M}(\phi_1) + 1$, thus there are at most $2\mathcal{M}(\phi_1) + 1 = O(\mathsf{size}(\phi_1))^2$ such literals.

  Summing up, we obtain that $||\{\ell_i \mid i \in [1, m]\}|| = O(\mathsf{size}(\phi_1)^2)$. Thus it suffices to guess a set $\left\{\ell_1', \ldots, \ell_n'\right\}$ of literals and a set of minterms $\left\{M_1', \ldots, M_n'\right\}$ such that $\ell_i' \in M_i'$, where $n = O(\mathsf{size}(\phi_1)^2)$ and $\mathsf{size}(M_i') = O(\mathsf{size}(\phi_1)^2)$, for all $i \in [1, n]$. Then we can check that:
  - $M_i' \in \mu^{fin}(\phi_1)$, which can be done in space $O(\mathsf{size}(M_i') + \mathsf{size}(\phi_1)^8) = O(\mathsf{size}(\phi_1)^2 + \mathsf{size}(\phi_1)^8) = O(\mathsf{size}(\phi_1)^8)$, by the inductive hypothesis,

– $M \in \left[\left[\overline{\ell_1}\right]^{\mathsf{mt}}, \ldots, \left[\overline{\ell_n}\right]^{\mathsf{mt}}\right]$, which can be done in space $O(\mathrm{size}(M) + (n \cdot \mathrm{size}(\phi_1)^2)^2) = O(\mathrm{size}(M) + \mathrm{size}(\phi_1)^8)$, by Proposition 5.40. Observe that this case is the most complex one, and it leads to the exponent 8 in the above inductive invariant.

To ensure that the set $\{\ell_1, \ldots, \ell_m\}$ contains no literal other than $\ell'_1, \ldots, \ell'_n$, we also have to check that every minterm M_j, for $j \in [1, m]$ contains a literal ℓ'_i, for some $i \in [1, n]$. To this aim, we use a nondeterministic algorithm for the complement: we guess a minterm M' $\mathcal{M}$-bounded by ϕ_1, check that $M' \in \mu(\phi_1)$ and that it contains no literal ℓ_i, for $i \in [1, n]$. By the inductive hypothesis, this is possible in space $O(\mathrm{size}(M') + \mathrm{size}(\phi_1)^8) = O(\mathrm{size}(\phi_1)^2 + \mathrm{size}(\phi_1)^8) = O(\phi_1^8)$. Then, checking that every minterm M_j, for $j \in [1, m]$ contains a literal ℓ'_i, for some $i \in [1, n]$ can be done in the same amount of space, using a nondeterministic algorithm, see e.g. [2, Corollary 4.21].

- $\phi = \phi_1 * \phi_2$: $M \in \mu^{\mathit{fin}}(\phi)$ iff there exist minterms $M_i \in \mu(\phi_i)$, $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$ and $P_i \in (N_i)^{\mathcal{N}^p_{3-i}}$, such that $M \in \left[(\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}}\right]^{\mathsf{mt}}$, for $i = 1, 2$. We first guess minterms M_1, M_2 of size $O(\mathrm{size}(\phi_1)^2)$ and $O(\mathrm{size}(\phi_2)^2)$, respectively, check that $M_i \in \mu^{\mathit{fin}}(\phi_i)$, then guess $N_i \in (M_i)^{\mathsf{E}(M_1 \cup M_2)}$ and $P_i \in (N_i)^{\mathcal{N}^p_{3-i}}$, for $i = 1, 2$. This is feasible since by definition each minterm in these sets is of size $O(\mathrm{size}(M_1) + \mathrm{size}(M_2))$. Next, we guess minterms M', M'', of size $O(\mathrm{size}(M_1) + \mathrm{size}(M_2))$ as well, and check that $M' \in (\mathsf{elim}_*(P_1, P_2))^{\mathsf{dnf}}$ in space $O(\mathrm{size}(M') + \mathrm{size}(P_1) + \mathrm{size}(P_2))$, by Proposition 5.40 and $M'' \in [M']^{\mathsf{mt}}$ in space $O(\mathrm{size}(M'') + \mathrm{size}(M')^2)$, by Proposition 5.41.
- $\phi_1 \mathrel{-\!\!\circ} \phi_2$: the proof is similar to the previous case.

□

6 BERNAYS-SCHÖNFINKEL-RAMSEY SL^k

This section contains the results concerning decidability of the (in)finite satisfiability problems within the $\mathsf{BSR}(\mathsf{SL}^k)$ fragment. First, we show that, contrary to $\mathsf{BSR}(\mathsf{FO})$, the satisfiability of $\mathsf{BSR}(\mathsf{SL}^k)$ is undecidable for $k \geq 2$. Second, we carve two nontrivial fragments of $\mathsf{BSR}(\mathsf{SL}^k)$, for which the infinite and finite satisfiability problems are both PSPACE-complete. defined based on restrictions of (i) polarities of the occurrences of the separating implication, and (ii) occurrences of universally quantified variables in the scope of separating implications. These results draw a rather precise chart of decidability within the $\mathsf{BSR}(\mathsf{SL}^k)$ fragment.

6.1 Undecidability of $\mathsf{BSR}(\mathsf{SL}^k)$

Theorem 6.1. *The finite and infinite satisfiability problems are both undecidable for formulæ in* $\mathsf{BSR}(\mathsf{SL}^k)$ *even if the formulæ contain no uninterpreted predicates.*

Proof. Let $\varphi = \forall x \,.\, \phi$ be a formula in $\mathsf{BSR}^2(\mathsf{FO})$, where ϕ is quantifier-free, contains no predicate symbol, one variable x, one constant symbol c and two monadic function symbols f and g of sort U. It is known that the finite satisfiability problem is undecidable for such formulæ, by Proposition 2.3. We reduce this problem to the infinite and finite satisfiability problems for $\mathsf{BSR}(\mathsf{SL}^k)$ formulæ. We proceed by first *flattening* each term in ϕ consisting of nested applications of f and g. The result is an equivalent sentence $\varphi_{\mathit{flat}} = \forall x_1 \ldots \forall x_n \,.\, \phi_{\mathit{flat}}$, in which the only terms are x_i, c, $f(x_i)$, $g(x_i)$, $f(c)$ and $g(c)$, for $i \in [1, n]$. For example, the formula $\forall x \,.\, f(g(x)) \approx c$ is flattened into $\forall x_1 \forall x_2 \,.\, g(x_1) \not\approx x_2 \vee f(x_2) \approx c$. The formal construction is standard and thus omitted. We define the following $\mathsf{BSR}(\mathsf{SL}^2)$ formulæ, for $\dagger \in \{\mathit{fin}, \mathit{inf}\}$:

$$\varphi^{\dagger}_{\mathsf{sl}} \stackrel{\mathsf{def}}{=} \alpha^{\dagger} \wedge x_c \hookrightarrow (y_c, z_c) \wedge \forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n \forall z_1 \ldots \forall z_n \,.\, \bigwedge_{i=1}^{n} (x_i \hookrightarrow (y_i, z_i) \rightarrow \phi_{\mathsf{sl}}) \tag{18}$$

where[9] $\alpha^{fin} \stackrel{\text{def}}{=} |h| \geq |U| - 0$, $\alpha^{inf} \stackrel{\text{def}}{=} \forall x \forall y \forall z \,.\, x \hookrightarrow (y,z) \rightarrow \text{alloc}(y) \wedge \text{alloc}(z)$ and ϕ_{sl} is obtained from ϕ_{flat} by replacing each occurrence of c by x_c, each term $f(c)$ (resp. $g(c)$) by y_c (resp. z_c) and each term $f(x_i)$ (resp. $g(x_i)$) by y_i (resp. z_i). Next, we show that the following statements are equivalent:

(1) φ_{flat} has a finite model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$,

(2) $\varphi_{\text{sl}}^{fin}$ has a finite model $(\mathfrak{U}, \mathfrak{s}', \mathcal{I}, \mathfrak{h})$, and

(3) $\varphi_{\text{sl}}^{inf}$ has an infinite model $(\mathfrak{U}^\infty, \mathfrak{s}', \mathcal{I}, \mathfrak{h})$.

"(1) $\Rightarrow$ (2)" We define the store $\mathfrak{s}' \stackrel{\text{def}}{=} \mathfrak{s}[x_c \leftarrow c^{\mathcal{I}}, y_c \leftarrow f^{\mathcal{I}}(c^{\mathcal{I}}), z_c \leftarrow g^{\mathcal{I}}(c^{\mathcal{I}})]$ and the heap $\mathfrak{h}$ such that $\text{dom}(\mathfrak{h}) = \mathfrak{U}$ and $\mathfrak{h}(\ell) \stackrel{\text{def}}{=} (f^{\mathcal{I}}(\ell), g^{\mathcal{I}}(\ell))$, for all $\ell \in \mathfrak{U}$. By construction, we have $(\mathfrak{U}, \mathfrak{s}', \mathfrak{h}) \models \alpha^{fin} \wedge x_c \hookrightarrow (y_c, z_c)$, because $\text{dom}(\mathfrak{h}) = \mathfrak{U}$ and $\mathfrak{h}(c^{\mathcal{I}}) = (f^{\mathcal{I}}(c^{\mathcal{I}}), g^{\mathcal{I}}(c^{\mathcal{I}}))$. Consider a store $\mathfrak{s}'' \stackrel{\text{def}}{=} \mathfrak{s}'[x_i \leftarrow \ell_i, y_i \leftarrow \ell_i', z_i \leftarrow \ell_i'' \mid i = 1, \ldots, n]$, for an arbitrary set $\{\ell_i, \ell_i', \ell_i'' \mid i \in [1,n]\} \subseteq \mathfrak{U}$ and assume that $(\mathfrak{U}, \mathfrak{s}'', \mathfrak{h}) \models \bigwedge_{i=1}^n x_i \hookrightarrow (y_i, z_i)$. Then by definition of $\mathfrak{h}$, for all $i \in [1,n]$, we have $\ell_i' = f^{\mathcal{I}}(\ell_i)$ and $\ell_i'' = g^{\mathcal{I}}(\ell_i)$; hence, $(\mathfrak{U}, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$. Since ℓ_i, ℓ_i' and ℓ_i'' are arbitrary, for $i \in [1,n]$, this proves that $(\mathfrak{U}, \mathfrak{s}', \mathfrak{h})$ is a finite model of $\varphi_{\text{sl}}^{fin}$.

"(2) $\Rightarrow$ (3)" We define $\mathfrak{U}^\infty \stackrel{\text{def}}{=} \mathfrak{U} \uplus L$, where L is an infinite set of locations not in $\mathfrak{U}$. Clearly $(\mathfrak{U}^\infty, \mathfrak{s}', \mathfrak{h}) \models \alpha^{inf}$, because $x \hookrightarrow (y,z)$ is false for any extension of $\mathfrak{s}'$ with a pair of the form $[x \leftarrow \ell]$, $[y \leftarrow \ell]$ or $[z \leftarrow \ell]$, where $\ell \in L$. Furthermore, the valuation of $x_c \hookrightarrow (y_c, z_c)$ is unchanged between $(\mathfrak{U}, \mathfrak{s}', \mathfrak{h})$ and $(\mathfrak{U}^\infty, \mathfrak{s}', \mathfrak{h})$. Consider a store $\mathfrak{s}'' \stackrel{\text{def}}{=} \mathfrak{s}'[x_i \leftarrow \ell_i, y_i \leftarrow \ell_i', z_i \leftarrow \ell_i'' \mid i = 1, \ldots, n]$, for an arbitrary set $\{\ell_i, \ell_i', \ell_i'' \mid i \in [1,n]\} \subseteq \mathfrak{U}$ and assume that $(\mathfrak{U}, \mathfrak{s}'', \mathfrak{h}) \models \bigwedge_{i=1}^n x_i \hookrightarrow (y_i, z_i)$. Then necessarily, $\left\{\ell_i, \ell_i', \ell_i'' \mid i \in [1,n]\right\} \cap L = \emptyset$. Next, we show that $(\mathfrak{U}, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}} \Leftrightarrow (\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$, by induction on the structure of ϕ_{sl}. Since $(\mathfrak{U}, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$ by the hypothesis, we have $(\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$, thus $(\mathfrak{U}^\infty, \mathfrak{s}, \mathfrak{h}) \models \varphi_{\text{sl}}^{inf}$.

"(3) $\Rightarrow$ (1)" Let $\mathfrak{U} \stackrel{\text{def}}{=} \text{dom}(\mathfrak{h}) \cup \{\ell_1, \ell_2 \mid \exists \ell \in \mathfrak{U}^\infty \,.\, \mathfrak{h}(\ell) = (\ell_1, \ell_2)\}$. Since $\mathfrak{h}$ is finite, so is $\mathfrak{U}$. Let $\mathfrak{s}$ be an arbitrary[10] store on $\mathfrak{U}$ and define $\mathcal{I}$ such that:

- $c^{\mathcal{I}} = \mathfrak{s}'(x_c)$, and,
- for each $\ell \in \mathfrak{U}$, such that $\mathfrak{h}(\ell) = (\ell', \ell'')$, we have $f^{\mathcal{I}}(\ell) = \ell'$ and $g^{\mathcal{I}}(\ell) = \ell''$.

Note that $c^{\mathcal{I}} \in \mathfrak{U}$, because by hypothesis $(\mathfrak{U}^\infty, \mathfrak{s}', \mathfrak{h}) \models x_c \hookrightarrow (y_c, z_c)$, hence $\mathfrak{s}'(x_c) \in \text{dom}(\mathfrak{h})$. Similarly, $f^{\mathcal{I}}(\ell), g^{\mathcal{I}}(\ell) \in \mathfrak{U}$, for each $\ell \in \mathfrak{U}$, by the definition of $\mathfrak{U}$. Moreover, since $(\mathfrak{U}^\infty, \mathfrak{s}', \mathfrak{h}) \models \alpha^{inf}$ we obtain that $f^{\mathcal{I}}$ and $g^{\mathcal{I}}$ are well-defined total functions. For each set $\{\ell_i \mid i = 1, \ldots, n\} \subseteq \mathfrak{U}$, the function $\mathfrak{s}'' = \mathfrak{s}[x_i \leftarrow \ell_i, y_i \leftarrow f^{\mathcal{I}}(\ell_i), z_i \leftarrow g^{\mathcal{I}}(\ell_i) \mid i = 1, \ldots, n]$ is a store on $\mathfrak{U}^\infty$ such that $(\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models x_i \hookrightarrow (y_i, z_i)$ for every $i \in [1,n]$, hence $(\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$. By induction on the structure of ϕ_{flat}, one shows that $(\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}} \Leftrightarrow (\mathfrak{U}, \mathfrak{s}'', \mathcal{I}) \models \phi_{flat}$. Since $(\mathfrak{U}^\infty, \mathfrak{s}'', \mathfrak{h}) \models \phi_{\text{sl}}$, we have $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \phi_{flat}$. □

Note that, by the previous proof, the undecidability result still holds for finite satisfiability if a single occurrence of $-\!\!*$ is allowed, in a ground formula (indeed, we may take $\alpha^{fin} = (|h| \geq |U| - 0) = (\neg\text{emp} -\!\!* \bot)$). For infinite satisfiability one occurrence of $-\!\!*$ is still sufficient, however there must be a universally quantified variable within the scope of $-\!\!*$.

The reductions (18) use positive occurrences of test formulæ $|h| \geq |U| - n$ and $\text{alloc}(x)$, where x is universally quantified. We obtain decidable subsets of $\text{BSR}(\text{SL}^k)$ by devising conditions that are sufficient to discard positive occurrences of such formulæ from $\mu^\dagger(\phi)$, where $\dagger \in \{fin, inf\}$ and $\forall y_1 \ldots \forall y_m \,.\, \phi$ is a $\text{BSR}(\text{SL}^k)$ formula. Note that $\mu^{inf}(\phi)$ contains no formulæ of the form $|h| \geq |U| - n$ (as such test formulæ are trivially false in all infinite structures) which explains why slightly less restrictive conditions are needed for infinite structures. As we shall see (Proposition 6.5), these conditions are sufficient to ensure that the formula $\forall y_1, \ldots, \forall y_m \,.\, \bigvee_{M \in \mu^{inf}(\phi)} M$ is BSR-compatible (but not that $\forall y_1, \ldots, \forall y_m \,.\, \bigvee_{M \in \mu^{fin}(\phi)} M$ is BSR-compatible, see Section 6.2.3 for details).

[9] Note that an equivalent definition of α^{fin} is $\alpha^{fin} \stackrel{\text{def}}{=} \forall x \,.\, \text{alloc}(x)$.

[10] The store is arbitrary because φ contains no free variables.

6.2 Decidability Proofs

6.2.1 Model Checking. We first show that the first-order model checking problem, considering the translation of minterms to FO, is in PSPACE. We first recall the following well-known result, proved for instance in [22].

Proposition 6.2. *Let $\mathcal{S}$ be an* FO*-structure and let ϕ be an* FO *formula. The problem of testing whether $\mathcal{S} \models \phi$ is in PSPACE.*

Proposition 6.2 does not by itself entails the desired result since $||\mu^{\dagger}(\phi)||$ is exponential w.r.t. size(ϕ). We need the following:

Lemma 6.3. *Given a finite* FO*-structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ and an* SL *formula $\forall y_1 \ldots \forall y_m \,.\, \phi$ where ϕ is quantifier-free, the problem $\mathcal{S} \models \tau(\forall y_1 \ldots \forall y_m \,.\, \bigvee_{M \in \mu^{\dagger}(\phi)} M)$ is in PSPACE, for each $\dagger \in \{fin, inf\}$.*

Proof. Since PSPACE is closed under complement (see, e.g., [2, Corollary 4.21]), we consider instead the problem $\mathcal{S} \models \neg\tau(\forall y_1 \ldots \forall y_m \,.\, \bigvee_{M \in \mu^{\dagger}(\phi)} M)$. Because $\tau(.)$ is homomorphic w.r.t. the propositional connectives, we have the equivalences:

$$\begin{array}{rcl} \neg\tau(\forall y_1 \ldots \forall y_m \,.\, \bigvee_{M \in \mu^{\dagger}(\phi)} M) & \equiv & \neg\forall y_1 \ldots \forall y_m \,.\, \tau(\bigvee_{M \in \mu^{\dagger}(\phi)} M) \\ & \equiv & \exists y_1 \ldots \exists y_m \,.\, \neg\tau(\bigvee_{M \in \mu^{\dagger}(\phi)} M) \\ & \equiv & \exists y_1 \ldots \exists y_m \,.\, \neg \bigvee_{M \in \mu^{\dagger}(\phi)} \tau(M) \\ & \equiv & \exists y_1 \ldots \exists y_m \,.\, \bigwedge_{M \in \mu^{\dagger}(\phi)} \neg\tau(M) \end{array}$$

To check that $\mathcal{S} \not\models \tau(\forall y_1 \ldots \forall y_m \,.\, \bigvee_{M \in \mu^{\dagger}(\phi)} M)$, we may thus guess locations $\ell_1, \ldots, \ell_m \in \mathfrak{U}$ and check that $(\mathfrak{U}, \mathfrak{s}[y_1 \leftarrow \ell_1] \ldots [y_m \leftarrow \ell_m], \mathcal{I}) \models \bigwedge_{M \in \mu^{\dagger}(\phi)} \neg\tau(M)$. There remains to prove that the latter test in is PSPACE. To this aim, we consider again the complement problem $(\mathfrak{U}, \mathfrak{s}[y_1 \leftarrow \ell_1] \ldots [y_m \leftarrow \ell_m], \mathcal{I}) \not\models \bigwedge_{M \in \mu^{\dagger}(\phi)} \neg\tau(M)$. We guess a minterm M that is $\mathcal{M}$-bounded by ϕ, then check that $M \in \mu^{\dagger}(\phi)$ and that $(\mathfrak{U}, \mathfrak{s}[y_1 \leftarrow \ell_1] \ldots [y_m \leftarrow \ell_m], \mathcal{I}) \models \tau(M)$. The first check is in PSPACE, by Lemma 5.42. The second check is also in PSPACE, by Proposition 6.2. □

Remark 6.4. *Note that the size of an* FO*-structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is exponential w.r.t. the arity of the symbols in $\mathcal{F}$. In our context, the arity of all symbols is bounded by a constant, except that of the special symbol $\mathfrak{p}$ that encodes the heap. Further, in the following (see for instance the proof of Theorem 6.11), we will only consider structures that satisfy the formula Heap in Definition 4.6 , so that $\mathfrak{p}^{\mathcal{I}}$ is a partial function and $||\mathfrak{p}^{\mathcal{I}}|| \leq ||\mathfrak{U}||$. Hence we may assume that the size of $\mathcal{S}$ is polynomial in $||\mathfrak{U}|| + k + \mathrm{dom}(\mathfrak{s})$.* ■

6.2.2 Infinite Satisfiability ($\mathsf{BSR}^{inf}(\mathsf{SL}^k)$*).* We start by showing decidability, in PSPACE, of the infinite satisfiability problem for the $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ fragment. We first establish the following result:

Proposition 6.5. *Let $\varphi = \forall y_1 \ldots \forall y_m \,.\, \phi$ be a formula in $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$, where ϕ is quantifier-free. The formula $\chi \stackrel{\text{def}}{=} \bigvee_{M \in \mu^{inf}(\phi)} M$ is* BSR*-compatible.*

Proof. By Lemma 5.29 (1), no formula of the form $|h| \geq |U| - i$ occurs positively in χ. Furthermore, if $\mathrm{alloc}(x)$ positively occurs in χ, then it must occur in a minterm in $\mu^{inf}(\phi)$, and by Lemma 5.29 (3), x necessarily occurs in the scope of a positive occurrence of $-\!\!*$, which entails by definition of $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ that $x \notin \{y_1, \ldots, y_n\}$. Consequently, χ is BSR-compatible. □

Proposition 6.5, together with Lemma 5.28, ensures that a reduction from $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ to BSR(FO) is feasible. However, we also have to ensure that the cardinality of the universe is infinite and that the cardinality of the heap is finite, which cannot be expressed in FO. To this aim, we rely on existing results about the cardinality of models of BSR(FO) formulæ. The definition and theorem below are from [12] (they have been slightly adapted to handle formulæ containing free variables).

Definition 6.6. Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be an FO-structure. Let $A \stackrel{\text{def}}{=} \{\mathfrak{s}(x) \mid x \in \text{dom}(\mathfrak{s})\} \cup \{c^{\mathcal{I}} \mid c \in \mathcal{F}, \#(c) = 0, \sigma(c) = \mathsf{U}\}$ and $B \stackrel{\text{def}}{=} \mathfrak{U} \setminus A$. The structure $\mathcal{S}$ is *m-repetitive* if $||B|| \geq m$ and there exists a total order $\prec$ on $\mathfrak{U}$ such that for every $n \leq m$ and strictly increasing sequences $e_1 \prec \cdots \prec e_n$ and $e'_1 \prec \cdots \prec e'_n$ of elements in B, for every predicate symbol $q \in \mathcal{F}$ and every $d_1, \ldots, d_{\#(q)} \in A \cup \{e_1, \ldots, e_n\}$ the following holds:

$$(d_1, \ldots, d_{\#(q)}) \in q^{\mathcal{I}} \Leftrightarrow (d'_1, \ldots, d'_{\#(q)}) \in q^{\mathcal{I}}, \qquad \text{where } d'_i \stackrel{\text{def}}{=} \begin{cases} e'_j \text{ if } d_i = e_j \\ d_i \text{ otherwise} \end{cases}$$

The following theorem, proved in [12], characterizes the existence of an infinite model of a BSR(FO) formula. The intuition is that, due to the above condition, the interpretation of the predicate symbols in an m-repetitive model fulfills some symmetry properties that make it possible to extend this model into an infinite one by adding infinitely many copies of existing elements. Conversely, it is possible to show that every infinite model (actually, every model of sufficiently large cardinality) admits a restriction that is m-repetitive (the proof is based on Ramsey's theorem for hypergraphs [?]).

Theorem 6.7. *Consider a* BSR(FO) *formula φ containing n free variables and constants, no existential quantifier and m distinct universally quantified formulæ. The formula φ has an infinite model if and only if it has an m-repetitive model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $||\mathfrak{U}|| \leq n + m$.*

Proof. See [12, Theorems 4 and 5]. The addition of free variables is not problematic as they can be handled as constants. □

Proposition 6.8. *Testing whether a first-order structure $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is m-repetitive for a given $m \in \mathbb{N}$ is in PSPACE.*

Proof. The algorithm is straightforward: it is clear that A and B can be computed in polynomial time, then it suffices to guess some total order $<$ on $\mathfrak{U}$, to iterate over the increasing sequences $(e_1, \ldots, e_n), (e'_1, \ldots, e'_n) \in B^n$, with $n \leq m$, over the predicate symbols $q \in \mathcal{F}$ and elements $d_1, \ldots, d_{\#(q)} \in A \cup \{e_1, \ldots, e_n\}$, to compute in each case the elements $d'_1, \ldots, d'_{\#(q)}$ according to Definition 6.6 and to check that the equivalence $(d_1, \ldots, d_{\#(q)}) \in q^{\mathcal{I}} \Leftrightarrow (d'_1, \ldots, d'_{\#(q)}) \in q^{\mathcal{I}}$ holds. □

Theorem 6.7 and Proposition 6.8 provide an effective method to decide whether a formula ϕ in BSR(FO) has an infinite model. To ensure that the domain of the predicate $\mathfrak{p}$ encoding the heap is finite we rely on the following definition and result:

Definition 6.9. Let φ be a BSR(FO) formula. We denote by $\varphi_{\mathfrak{p}}$ the formula:

$$\forall x_1, \ldots, x_{k+1} \cdot \bigvee_{i=1}^{k+1} \bigwedge_{z \in \text{Var}(\varphi) \cup \text{Const}(\varphi)} \neg x_i \approx z \rightarrow \neg \mathfrak{p}(x_1, \ldots, x_{k+1})$$

where $x_1, \ldots, x_{k+1}$ are pairwise distinct variables not occurring in φ.

Proposition 6.10. *Let φ be a* BSR(FO) *formula. The two following assertions are equivalent.*

- *φ has an infinite model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $\mathfrak{p}^{\mathcal{I}}$ is finite.*
- *$\varphi \wedge \varphi_{\mathfrak{p}}$ has an infinite model.*

Proof. Assume that φ admits a model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $||\mathfrak{U}|| = \infty$ and $||\mathfrak{p}^{\mathcal{I}}|| \in \mathbb{N}$. Let $A \stackrel{\text{def}}{=} \{\mathfrak{s}(x) \mid x \in \text{Var}(\varphi)\} \cup \{c^{\mathcal{I}} \mid c \in \text{Const}(\varphi)\}$, let B be the set of elements of $\mathfrak{U}$ that do not occur in any vector in $\mathfrak{p}^{\mathcal{I}}$ and $\mathfrak{U}' \stackrel{\text{def}}{=} A \cup B$. Since $\mathfrak{p}^{\mathcal{I}}$ is finite, necessarily B and $\mathfrak{U}'$ are both infinite. By Proposition 2.2, the restriction $(\mathfrak{U}', \mathfrak{s}', \mathcal{I}')$ of $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ to $\mathfrak{U}'$ validates φ, since φ is a BSR(FO) formula and $A \subseteq \mathfrak{U}'$. It is clear that $(\mathfrak{U}', \mathfrak{s}', \mathcal{I}') \models \varphi_{\mathfrak{p}}$, since by

definition, $\mathfrak{U}'^{k+1} \cap \mathfrak{p}^{\mathcal{I}} = A^{k+1} \cap \mathfrak{p}^{\mathcal{I}}$. Conversely, let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be an infinite model of $\varphi \wedge \varphi_{\mathfrak{p}}$. Then by definition of $\varphi_{\mathfrak{p}}$ for every $(\ell_1, \ldots, \ell_{k+1})$ such that $(\ell_1, \ldots, \ell_{k+1}) \in \mathfrak{p}^{\mathcal{I}}$, and for every $i \in [1, k+1]$, either $\ell_i = c^{\mathcal{I}}$ for some $c \in \mathcal{F}(\varphi)$, or $\ell_i = \mathfrak{s}(x)$ for some $x \in \mathrm{Var}(\varphi)$. Since $\mathcal{F}(\varphi)$ and $\mathrm{Var}(\varphi)$ are both finite, $\mathfrak{p}^{\mathcal{I}}$ is also finite. □

Putting all results together, we obtain the first decidability result of this paper:

Theorem 6.11. *The infinite satisfiability problem for* $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ *is PSPACE-complete.*

Proof. PSPACE-hardness is an immediate consequence of the fact that the quantifier-free fragment of SL^k, without the separating implication, but with the separating conjunction and negation, is PSPACE-hard [7, Proposition 5].

To show membership in PSPACE, let $\varphi = \forall y_1 \ldots \forall y_m \,.\, \phi$ be a formula in $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$, where ϕ is quantifier-free and $\mathrm{Var}(\varphi) = \{x_1, \ldots, x_n\}$. Let $\varphi' \stackrel{\text{def}}{=} \forall y_1 \ldots \forall y_m \,.\, \chi$, with $\chi \stackrel{\text{def}}{=} \bigvee_{M \in \mu^{inf}(\phi)} M$ and let $\psi \stackrel{\text{def}}{=} \tau(\varphi') \wedge \mathcal{A}(\varphi')$. By Lemma 5.28, $\varphi \equiv^{inf} \varphi'$. By Proposition 6.5, χ is BSR-compatible and we deduce by Lemma 4.9 that φ' (and hence φ) has an infinite model iff ψ has an infinite model where the interpretation of $\mathfrak{p}$ is finite.

We now show how to solve the latter problem. By Proposition 4.10, ψ is a BSR(FO) formula with no existential variable and contains $k \cdot n + (k+6) \cdot \mathcal{N}(\chi) + 5$ constants. By Proposition 5.36, $\mathcal{N}(\chi) = O(\mathrm{size}(\phi)^2)$, thus we deduce that ψ is a BSR(FO) formula, with $O(k \cdot \mathrm{size}(\varphi)^2)$ constants and free variables. By Proposition 6.10, ψ has an infinite model where the interpretation of $\mathfrak{p}$ is finite iff $\psi \wedge \psi_{\mathfrak{p}}$ has an infinite model. By Theorem 6.7, $\psi \wedge \psi_{\mathfrak{p}}$ has an infinite model iff it has an m-repetitive model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ of cardinality $||\mathfrak{U}|| = O(k \cdot \mathrm{size}(\varphi)^2)$, because $\psi_{\mathfrak{p}}$ is a BSR(FO) formula with no existential variable and contains no constant or free variable other than those in ψ.

The algorithm is then defined as follows. We guess an FO-structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $||\mathfrak{U}|| = O(k \cdot \mathrm{size}(\varphi)^2)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$ (where *Heap* is the formula in Definition 4.6). Note that since k may depend on the input, $\mathfrak{U}^k$ is of exponential size, hence in principle the interpretation of $\mathfrak{p}$ may be exponential. However, since we assume that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$, for every element $x \in \mathfrak{U}$, there is at most one vector $\mathbf{y} \in \mathfrak{U}^k$ such that $(x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$, hence $||\mathfrak{p}^{\mathcal{I}}|| \leq ||\mathfrak{U}||$. To ensure that *Heap* holds, it suffices to guess a subset of $\mathfrak{U}$ (the set of allocated locations), and choose for every element x in this subset one vector $\mathbf{y} \in \mathfrak{U}^k$ such that $(x, \mathbf{y}) \in \mathfrak{p}^{\mathcal{I}}$. Moreover, the arity of each predicate symbol in φ that are different from $\mathfrak{p}$ is bounded by a constant, thus their interpretation is polynomial w.r.t. $\mathfrak{U}$. Then we check that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is m-repetitive and that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\varphi') \wedge \mathcal{A}(\varphi') \wedge \psi_{\mathfrak{p}}$. This test is feasible in PSPACE:

- the problem of testing whether $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is m-repetitive is in PSPACE by Proposition 6.8.
- the problem $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\varphi')$ is in PSPACE by Lemma 6.3,
- the problems $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathcal{A}(\varphi')$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \psi_{\mathfrak{p}}$ are both in PSPACE, by Proposition 6.2.

□

Remark 6.12. *The algorithm given in the proof of Theorem 6.11 is based on guessing some structure of size s, with $s = O(k \cdot \mathrm{size}(\varphi)^2)$. To apply the algorithm one needs of course to know an upper bound of s. Because our aim in the present paper is only to prove the existence of such an algorithm, we do not bother to give this bound explicitly, as this would only hinder readability, and we only state that it exists. However, the bound can easily be extracted from the above proofs, if needed. Similarly, an explicit bound on the size of the minterms in $\mu^{inf}(\phi)$ could be extracted from the proof of Lemma 5.42.* ■

6.2.3 Finite Satisfiability ($\mathsf{BSR}^{fin}(\mathsf{SL}^k)$*).* We now prove that finite satisfiability is PSPACE-complete for the class $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$, defined as the set of formulæ with no positive occurrence of separating implications. Even with this stronger restriction, the previous proof based on a translation to first-order logic cannot be carried over without any additional argument, because Proposition 6.5 does not hold for $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$. The problem is that, in the case of a finite universe, $\mathrm{alloc}(x)$ test formulæ may occur at a positive polarity, even if every $\phi_1 \mathrel{-\!\!*} \phi_2$ subformula occurs at a negative polarity, due to the positive occurrences of $\mathrm{alloc}(x)$ within the subformula λ^{fin} in

the definition of $\mathsf{elim}^{fin}_{-\circ}(M_1, M_2)$ (Equation (15) in Lemma 5.20), see also Example 5.30. As previously discussed, positive occurrences of $\mathsf{alloc}(x)$ hinder the translation into BSR(FO), because of the existential quantifiers that may occur in the scope of a universal quantifier.

The solution is to distinguish a class of finite structures $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$. Given $\alpha \in \mathbb{N}$, we consider the so-called *α-controlled structures*, for which there exists a set of locations $\ell_1, \ldots, \ell_\alpha$, such that every location $\ell \in \mathfrak{U} \setminus \{\ell_1, \ldots, \ell_\alpha\}$ points to a tuple from the set $\{\ell_1, \ldots, \ell_\alpha, \ell\}$. An example of a 3-controlled structure is given in Figure 2.

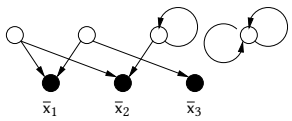


Fig. 2. A finite 3-controlled SL^2 structure.

Definition 6.13. An SL-structure $\mathcal{S}$ is *α-controlled* if $\mathcal{S} \models \exists \bar{x}_1, \ldots, \bar{x}_\alpha \,.\, C(\alpha)$, with

$$C(\alpha) \stackrel{\text{def}}{=} \forall x \,.\, \bigvee_{i=1}^{\alpha} x \approx \bar{x}_i \vee \bigvee_{y \in \mathsf{vect}^k(\bar{x}_1, \ldots, \bar{x}_\alpha, x)} x \hookrightarrow y$$

where $\mathsf{vect}^k(x_1, \ldots, x_n)$ is the set of k-tuples of symbols in $\{x_1, \ldots, x_n\}$, and $\bar{x}_1, \ldots, \bar{x}_\alpha, x$ are pairwise distinct variables. Analogously, an FO-structure $\mathcal{S}$ is *α-controlled* if $\mathcal{S} \models \exists \bar{x}_1, \ldots, \bar{x}_\alpha \,.\, \tau(C(\alpha))$, with

$$\tau(C(\alpha)) = \forall x \,.\, \bigvee_{i=1}^{\alpha} x \approx \bar{x}_i \vee \bigvee_{y \in \mathsf{vect}^k(\bar{x}_1, \ldots, \bar{x}_\alpha, x)} \mathfrak{p}(x, y)$$

Any α-controlled SL-structure is finite, since $\mathfrak{U} = \mathrm{dom}(\mathfrak{h}) \cup \{\mathfrak{s}(\bar{x}_1), \ldots, \mathfrak{s}(\bar{x}_\alpha)\}$, but its cardinality is not bounded. Furthermore, if $||\mathfrak{U}|| \leq \alpha$, then $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ is necessarily α-controlled.

Overview of the Proof for Finite Satisfiability. For a formula $\varphi = \forall y_1 \ldots \forall y_m \,.\, \phi$ in $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$, we distinguish the following cases:

(1) If φ has an α-controlled model $\mathcal{S}$, the formula obtained by replacing each occurrence of an $\mathsf{alloc}(x)$ with $\bigwedge_{i=1}^{\alpha}(x \approx \bar{x}_i \rightarrow \mathsf{alloc}(\bar{x}_i))$ in $\forall y_1 \ldots \forall y_m \bigvee_{M \in \mu^{fin}(\phi)} M$ is satisfied by $\mathcal{S}$ (as stated by Proposition 6.15).
(2) Otherwise, each finite model of φ is non-α-controlled and we can build a model $\mathcal{S}$, with a sufficiently large universe, such that each test formula $\theta \in \{|U| \geq n, |h| < |U| - n \mid n \in \mathbb{N}\}$ becomes true in $\mathcal{S}$. Assume $\mathsf{alloc}(x)$ occurs positively in a λ^{fin} subformula of some formula $\mathsf{elim}^{fin}_{-\circ}(M_1, M_2)$. The latter must have been generated by the elimination of a separating implication from ϕ, hence $\mathsf{alloc}(x)$ occurs in a disjunction with a formula of the form $|h| < |U| - n_1 \wedge |U| \geq n_2$; its truth value in $\mathcal{S}$ can thus be ignored and the entire subformula deleted.

In both cases, we obtain an equisatisfiable universally quantified boolean combination of test formulæ with no positive occurrence of $\mathsf{alloc}(y_i)$ formulæ. We translate this into an equisatisfiable BSR(FO) formula, for which finite satisfiability is decidable and apply a similar argument to that for the infinite case, to obtain the PSPACE upper bound.

The Case of Controlled Structures. We first consider the case where the considered models are α-controlled.

Proposition 6.14. *Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be an FO-structure. The problem of testing whether $\mathcal{S} \models \tau(C(\alpha))$ is in P.*

Proof. Note that the size of $C(\alpha)$ is exponential w.r.t. k. However, to test that $\mathcal{S} \models \tau(C(\alpha))$, it suffices to check that for every $u \in \mathfrak{U} \setminus \mathfrak{s}(\{\bar{x}_1, \ldots, \bar{x}_\alpha\})$, there exist $v_1, \ldots, v_k$ such that $(u, v_1, \ldots, v_k) \in \mathfrak{p}^{\mathcal{I}}$ and $v_1, \ldots, v_k \subseteq \{u\} \cup \mathfrak{s}(\{\bar{x}_1, \ldots, \bar{x}_\alpha\})$, which can be done in time polynomial in size$(\mathcal{S})$. □

Proposition 6.15. *If x is a variable distinct from $\overline{x}_1, \ldots, \overline{x}_\alpha$, then:*

$$C(\alpha) \models \forall x \,.\, (\mathsf{alloc}(x) \leftrightarrow \bigwedge_{i=1}^{\alpha} (x \approx \overline{x}_i \rightarrow \mathsf{alloc}(\overline{x}_i))$$

Proof. This is immediate, since $C(\alpha)$ entails that every element distinct from $\overline{x}_1, \ldots, \overline{x}_\alpha$ is allocated. □

Lemma 6.16. *Given a formula $\varphi \in \mathsf{BSR}^{fin}(\mathsf{SL}^k)$ and a number $\alpha \in \mathbb{N}$ encoded in unary, the problem of checking whether φ has an α-controlled model is in PSPACE.*

Proof. We assume that the formula φ is of the form $\forall y_1, \ldots, y_m \,.\, \phi$, with $\mathrm{Var}(\varphi) = \{x_1, \ldots, x_n\}$. Let $\overline{x}_1, \ldots, \overline{x}_\alpha$ be pairwise distinct variables not occurring in $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. It is clear that φ admits an α-controlled model iff it admits a model that also validates $C(\alpha)$. Let $\chi \stackrel{\text{def}}{=} \bigvee_{M \in \mu^{fin}(\phi)} M$, let χ' the formula obtained from χ by replacing every formula $\mathsf{alloc}(x)$ with $\bigwedge_{i=1}^{\alpha}(x \approx \overline{x}_i \rightarrow \mathsf{alloc}(\overline{x}_i))$ and let $\varphi' \stackrel{\text{def}}{=} \forall y_1 \ldots \forall y_m \,.\, \chi'$. By Proposition 6.15, $C(\alpha) \models \chi \leftrightarrow \chi'$, thus, since any model of $C(\alpha)$ is finite, we deduce by Lemma 5.28 that $C(\alpha) \models \phi \leftrightarrow \chi'$. Consequently φ has an α-controlled model iff $\varphi' \wedge C(\alpha)$ has a model. By Lemma 5.29 (2), χ contains no test formulæ $|h| \geq |U| - i$ at positive polarity, thus the same holds for χ'. Moreover, the formula χ' contains no occurrence of $\mathsf{alloc}(y_i)$, since by definition the only test formulæ $\mathsf{alloc}(x)$ occurring in χ' are such that $x \in \{\overline{x}_1, \ldots, \overline{x}_\alpha\}$. Hence χ' is BSR-compatible, and we deduce by Lemma 4.9 that φ has an α-controlled model iff $\psi \stackrel{\text{def}}{=} \tau(\varphi' \wedge C(\alpha)) \wedge \mathcal{A}(\varphi' \wedge C(\alpha))$ has a finite model.

We now show how to solve the latter problem. Since φ' and $C(\alpha)$ contain no cardinality constraints other than those in χ, we have $\mathcal{N}(\varphi' \wedge C(\alpha)) \leq \mathcal{N}(\chi)$, thus by Proposition 5.36, we deduce that $\mathcal{N}(\varphi' \wedge C(\alpha)) = O(k \cdot \mathrm{size}(\varphi)^2)$. By Proposition 4.10, this entails that ψ is a BSR(FO) formula with $O(k \cdot \mathrm{size}(\varphi)^2 + k \cdot \alpha)$ constants and free variables (since $\mathrm{Var}(\varphi' \wedge C(\alpha)) = \mathrm{Var}(\varphi') \cup \mathrm{Var}(C(\alpha)) = \{x_1, \ldots, x_n, \overline{x}_1, \ldots, \overline{x}_\alpha\}$). By Proposition 2.2, ψ has a finite model iff it has a model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$, with $||\mathfrak{U}|| = O(k \cdot \mathrm{size}(\varphi)^2 + k \cdot \alpha)$.

The algorithm is defined as follows (see the proof of Theorem 6.11 for details). We first guess a structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ such that $||\mathfrak{U}|| = O(k \cdot \mathrm{size}(\varphi)^2 + k \cdot \alpha)$ and $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathit{Heap}$. Then we check that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\varphi')$ (as done in the proof of Lemma 6.3, except that all formulæ $\mathsf{alloc}(x)$ are replaced by $\bigwedge_{i=1}^{\alpha}(x \approx \overline{x}_i \rightarrow \mathsf{alloc}(\overline{x}_i))$), that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(C(\alpha))$ (using Proposition 6.14) and that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathcal{A}(\varphi' \wedge C(\alpha))$ (using Proposition 6.2). □

The General Case. To handle the case where no α-controlled model exists, the following results are used.

Proposition 6.17. *Let $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be a non-α-controlled* FO*-structure satisfying the (Heap) axiom, defined on page 11. Let $E \subseteq \mathfrak{U}$, with $||E|| \leq \alpha$. There exists an element $u \in \mathfrak{U} \setminus E$ such that either u is not allocated, or there exist $v_1, \ldots, v_k \in \mathfrak{U}$ and $j \in [1, k]$ such that $(u, v_1, \ldots, v_k) \in \mathfrak{p}^{\mathcal{I}}$ and $v_j \notin E \cup \{u\}$.*

Proof. Because $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ is not α-controlled, we have

$$(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \forall \overline{x}_1, \ldots, \overline{x}_\alpha \,.\, \exists x \,.\, \bigwedge_{i=1}^{\alpha} \neg x \approx \overline{x}_i \wedge \bigwedge_{\mathbf{y} \in \mathrm{vect}^k(\overline{x}_1, \ldots, \overline{x}_\alpha, x)} \neg \mathfrak{p}(x, \mathbf{y}).$$

Let $\mathfrak{s}'$ be any extension of $\mathfrak{s}$ to $\overline{x}_1, \ldots, \overline{x}_\alpha$ such that $\mathfrak{s}'(\{\overline{x}_1, \ldots, \overline{x}_\alpha\}) = E$ (such as store necessarily exists since $||E|| \leq \alpha$). We have $(\mathfrak{U}, \mathfrak{s}', \mathcal{I}) \models \exists x \,.\, \bigwedge_{i=1}^{\alpha} \neg x \approx \overline{x}_i \wedge \bigwedge_{\mathbf{y} \in \mathrm{vect}^k(\overline{x}_1, \ldots, \overline{x}_\alpha, x)} \neg \mathfrak{p}(x, \mathbf{y})$ hence $\mathfrak{U}$ contains an element $u \notin E$ such that $(\mathfrak{U}, \mathfrak{s}'[x \leftarrow u], \mathcal{I}) \models \bigwedge_{\mathbf{y} \in \mathrm{vect}^k(\overline{x}_1, \ldots, \overline{x}_\alpha, x)} \neg \mathfrak{p}(x, \mathbf{y})$. If u is not allocated then the proof is completed. Otherwise, let $(u, v_1, \ldots, v_k) \in \mathfrak{p}^{\mathcal{I}}$ and assume that $\forall j \in [1, k]$, $v_j \in E \cup \{u\}$. Since $\mathfrak{s}'(\{\overline{x}_1, \ldots, \overline{x}_\alpha\}) = E$, this entails that for every $j \in [1, k]$, there exists $y_j \in \{\overline{x}_1, \ldots, \overline{x}_\alpha, x\}$ such that $v_j = \mathfrak{s}'(y_j)$. But then $(x, y_1, \ldots, y_k) \in \mathrm{vect}^k(\overline{x}_1, \ldots, \overline{x}_\alpha, x)$ and $(\mathfrak{U}, \mathfrak{s}'[x \leftarrow u], \mathcal{I}) \models \mathfrak{p}(x, y_1, \ldots, y_k)$, which contradicts the fact that $(\mathfrak{U}, \mathfrak{s}'[x \leftarrow u], \mathcal{I}) \models \bigwedge_{\mathbf{y} \in \mathrm{vect}^k(\overline{x}_1, \ldots, \overline{x}_\alpha, x)} \neg \mathfrak{p}(x, \mathbf{y})$. □

Lemma 6.18. *Let $n \in \mathbb{N}$. Consider a* BSR(FO) *formula φ, let $m \stackrel{\text{def}}{=} ||\text{Var}(\varphi)||+||\text{Const}(\phi)||$ and let $\alpha \geq (k+2)\cdot n+m$. If $\varphi \cup \{Heap\}$ has a non-α-controlled model $\mathcal{S}$ then there is a restriction of $\mathcal{S}$ that also validates $\varphi \cup \{Heap\}$ and has at least n unallocated elements.*

Proof. The result is trivial if $n = 0$, since $\mathcal{S}$ is a restriction of itself and trivially contains at least 0 unallocated elements. Thus we assume that $n > 0$. Let $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ be a non α-controlled model of $\varphi \cup \{Heap\}$. Let $A = \mathfrak{s}(\text{Var}(\varphi)) \cup \{c^{\mathcal{I}} \mid c \in \text{Const}(\varphi)\}$. Note that by definition, $||A|| \leq m$. We construct a sequence of pairwise distinct elements $x_1, \ldots, x_n \in \mathfrak{U}$ and a sequence of sets of elements $Y_0 \subseteq Y_1 \subseteq \cdots \subseteq Y_n \subseteq \mathfrak{U} \setminus A$ such that $||Y_i|| \leq i$, $x_1, \ldots, x_i \notin Y_i$ and for every $j \in [1, i]$, either x_j is unallocated or points to a vector containing an element of Y_i. The sequence is constructed inductively as follows. Let $Y_0 \stackrel{\text{def}}{=} \emptyset$. Assume that $x_1, \ldots, x_i, Y_1, \ldots, Y_i$ have been constructed, for some $i \in [0, n-1]$. Let $X = \{x_1, \ldots, x_i\}$, $E = \{z_1, \ldots, z_k \mid (x_j, z_1, \ldots, z_k) \in \mathfrak{p}^{\mathcal{I}}, 1 \leq j \leq i\}$. Because $\mathcal{S} \models Heap$, for every $j \in [1, i]$ there is at most one vector $(z_1, \ldots, z_k)$ such that $(x_j, z_1, \ldots, z_k) \in \mathfrak{p}^{\mathcal{I}}$, hence $||E|| \leq k \cdot i \leq k \cdot n$. Further, $||X|| = i \leq n$ and $||Y_i|| \leq i \leq n$. Thus $||E \cup A \cup X \cup Y_i|| \leq ||E|| + ||A|| + ||X|| + ||Y_i|| \leq k \cdot n + m + 2 \cdot n \leq \alpha$. Thus, since φ is not α-controlled, by Proposition 6.17, there exists an element $x_{i+1} \notin E \cup A \cup X \cup Y_i$ such that either x_{i+1} is not allocated, or there exists a (unique) vector $\mathbf{z}_i$ such that $(x_{i+1}, \mathbf{z}_i) \in \mathfrak{p}^{\mathcal{I}}$ and $\mathbf{z}$ has a component y_{i+1} with $y_{i+1} \notin E \cup X \cup Y_i \cup A \cup \{x_{i+1}\}$. In the former case, we take $Y_{i+1} \stackrel{\text{def}}{=} Y_i$ and in the latter case, $Y_{i+1} \stackrel{\text{def}}{=} Y_i \cup \{y_{i+1}\}$. Note that in both cases $Y_{i+1} \supseteq Y_i$ and $||Y_{i+1}|| \leq ||Y_i|| + 1 \leq i + 1$. Further, since $x_{i+1} \notin Y_i$ and $y_{i+1} \notin X \cup \{x_{i+1}\}$, necessarily $x_1, \ldots, x_{i+1} \notin Y_{i+1}$, thus the sequences fulfill the required properties.

Then, we consider the restriction $\mathcal{S}'$ of $\mathcal{S}$ to $\mathfrak{U}' \stackrel{\text{def}}{=} \mathfrak{U} \setminus Y_n$. As $x_1, \ldots, x_n \notin Y_n$, $\mathfrak{U} \setminus Y_n$ is not empty and contains $x_1, \ldots, x_n$. By Proposition 2.2, since $Y_n \cap A = \emptyset$, $\mathcal{S}' \models \varphi \cup \{Heap\}$. If x_i is allocated in $\mathcal{S}'$, then there exists $\mathbf{z} \in \mathfrak{U}'^k$ such that $(x_i, \mathbf{z}) \in \mathfrak{p}^{\mathcal{I}}$. But by the construction above, $\mathbf{z}$ contains an element in $Y_i \subseteq Y_n$, which contradicts the fact that $\mathbf{z} \in \mathfrak{U}'^k$. Thus necessarily x_i is unallocated in $\mathcal{S}'$. Since the elements $x_1, \ldots, x_n$ are pairwise distinct, the proof is completed. □

Proposition 6.19. *Let φ be an* SL *formula. If φ has a non-α-controlled* SL*-model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ then $\tau(\varphi) \wedge \mathcal{A}(\varphi)$ has a non-α-controlled* FO*-model where the interpretation of $\mathfrak{p}$ is finite.*

Proof. By Lemma 4.9 (1), there exists $\mathcal{J}$ such that $(\mathfrak{U}, \mathfrak{s}, \mathcal{J}) \models \tau(\varphi) \wedge \mathcal{A}(\varphi)$, where $\mathfrak{h}$ is associated with $\mathcal{J}$. If $(\mathfrak{U}, \mathfrak{s}, \mathcal{J})$ is α-controlled, then there exists an extension $\mathfrak{s}'$ of $\mathfrak{s}$ such that $(\mathfrak{U}, \mathfrak{s}', \mathcal{J}) \models \tau(C(\alpha))$. This entails that for all extensions $\mathfrak{s}''$ of $\mathfrak{s}'$ to x, $(\mathfrak{U}, \mathfrak{s}'', \mathcal{J}) \models \bigvee_{i=1}^{\alpha} x \approx \overline{\mathrm{x}}_i \vee \bigvee_{\mathrm{y} \in \text{vect}^k(\overline{\mathrm{x}}_1, \ldots, \overline{\mathrm{x}}_\alpha, x)} \mathfrak{p}(x, \mathrm{y})$. By definition, $(\mathfrak{U}, \mathfrak{s}'', \mathcal{J}) \models x \approx \overline{\mathrm{x}}_i$ iff $(\mathfrak{U}, \mathfrak{s}'', \mathcal{I}, \mathfrak{h}) \models x \approx \overline{\mathrm{x}}_i$. Furthermore, since $\mathfrak{h}$ is associated with $\mathcal{J}$, we have by definition $(\mathfrak{U}, \mathfrak{s}'', \mathcal{J}) \models \mathfrak{p}(x, \mathrm{y})$ iff $(\mathfrak{U}, \mathfrak{s}'', \mathcal{I}, \mathfrak{h}) \models x \hookrightarrow \mathrm{y}$. Therefore $(\mathfrak{U}, \mathfrak{s}'', \mathcal{I}, \mathfrak{h}) \models \bigvee_{i=1}^{\alpha} x \approx \overline{\mathrm{x}}_i \vee \bigvee_{\mathrm{y} \in \text{vect}^k(\overline{\mathrm{x}}_1, \ldots, \overline{\mathrm{x}}_\alpha, x)} x \hookrightarrow \mathrm{y}$. As $\mathfrak{s}''$ is arbitrary, this entails that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ is α-controlled, contradicting our hypothesis. □

We are now in the position to state the second decidability result of the paper, concerning the decidability of the finite satisfiability for $\text{BSR}^{fin}(\text{SL}^k)$:

Theorem 6.20. *The finite satisfiability problem for $\text{BSR}^{fin}(\text{SL}^k)$ is PSPACE-complete.*

Proof. PSPACE-hardness is proved using the same argument as in the proof of Theorem 6.11, which does not rely on the infiniteness of the universe.

Let $\varphi \stackrel{\text{def}}{=} \forall y_1, \ldots, y_m \ . \ \phi$ be a formula in $\text{BSR}^{fin}(\text{SL}^k)$, where ϕ is quantifier-free and $\text{Var}(\varphi) = \{x_1, \ldots, x_n\}$. Let $\chi \stackrel{\text{def}}{=} \bigvee_{M \in \mu^{fin}(\phi)} M$ and $\alpha \stackrel{\text{def}}{=} (k+2) \cdot (\mathcal{N}(\chi) + 1) + (k+1) \cdot n + (k+6) \cdot \mathcal{N}(\chi) + 5$. We first test whether φ admits an α-controlled model, which can be done in PSPACE, by Lemma 6.16 since, by Proposition 5.36, $\mathcal{N}(\chi) = O(\text{size}(\phi)^2)$, thus $\alpha = O(k \cdot \text{size}(\varphi)^2)$. In this case, φ has a finite model, and otherwise φ has a finite model iff it has a non-α-controlled finite model. We now assume that φ does not have any α-controlled model.

Let $\varphi' \stackrel{\text{def}}{=} \forall y_1, \ldots, y_m \ . \ \chi'$, where χ' is obtained from χ by replacing all positive occurrences of a formula $\text{alloc}(x)$, where $x \in \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$, by $\bot$. We prove that φ' has a finite model iff φ has a finite model.

By Lemma 5.28, $\varphi \equiv^{fin} \forall y_1, \ldots, y_m \,.\chi$. Because the replaced occurrences of $\mathsf{alloc}(x)$ are all positive, it is clear that $\chi' \models \chi$, thus $\varphi' = \forall y_1, \ldots, y_m \,.\chi' \models \forall y_1, \ldots, y_m \,.\chi \equiv^{fin} \varphi$ and the direct implication holds. Now, assume that φ admits a finite model. Note that by the above assumption this model is necessarily non-α-controlled. The formula χ can be written in cnf as $\chi_1 \wedge \chi_2$ where χ_1 is a conjunction of clauses not containing any literal $\mathsf{alloc}(x)$ and χ_2 is a conjunction of clauses containing at least one such literal. It is clear that $\mathcal{N}(\chi_1), \mathcal{N}(\chi_2) \leq \mathcal{N}(\chi)$ and $\varphi \models \forall y_1, \ldots, y_m \,.\, \chi_1$, thus $\forall y_1, \ldots, y_m \,.\, \chi_1$ has a non-α-controlled model. By Proposition 4.10, the formula $\xi = \tau(\forall y_1, \ldots, y_m \,.\, \chi_1) \wedge \mathcal{A}(\forall y_1, \ldots, y_m \,.\, \chi_1)$ is a BSR(FO) formula with at most n free variables and $k \cdot n + (k+6) \cdot \mathcal{N}(\chi) + 5$ constants, since $\mathcal{N}(\chi_1) \leq \mathcal{N}(\chi)$. Furthermore, by Proposition 6.19, ξ admits a non-α-controlled FO-model such that the interpretation of $\mathfrak{p}$ is finite, since $\forall y_1, \ldots, y_m \,.\, \chi_1$ has a non-α-controlled SL-model. By Lemma 6.18, and by definition of α, this entails that there exists an FO-model of ξ with strictly more than $\mathcal{N}(\chi)$ unallocated elements and such that the interpretation of $\mathfrak{p}$ is finite. By Lemma 5.29 (2), the formula χ (hence also χ_1) contains no positive occurrence of a formula of the form $|h| \geq |U| - i$, and by definition, χ_1 contains no positive occurrence of a formula $\mathsf{alloc}(x)$. Thus $\forall y_1, \ldots, y_m \,.\, \chi_1$ is BSR-compatible. By Lemma 4.9 (2), we deduce that $\forall y_1, \ldots, y_m \,.\, \chi_1$ admits an SL-model $\mathcal{S} = (\mathfrak{U}, \mathfrak{s}, \mathcal{I}, \mathfrak{h})$ with strictly more than $\mathcal{N}(\chi)$ unallocated elements. Assume that $\mathcal{S} \not\models \forall y_1, \ldots, y_m \,.\chi'$. This entails that there exist $e_1, \ldots, e_m \in \mathfrak{U}$ and a clause C in χ_2 such that $(\mathfrak{U}, \mathfrak{s}', \mathcal{I}, \mathfrak{h}) \not\models C'$, where $\mathfrak{s}' = \mathfrak{s}[x_i \leftarrow e_i \mid 1 \leq i \leq m]$ and C' is obtained from C by removing all the literals $\mathsf{alloc}(x)$. By definition C must contain at least one literal $\mathsf{alloc}(x)$. Because all occurrences of $-\!\!*$ in φ are negative or neutral, by Lemma 5.29 (4), every literal $\mathsf{alloc}(x)$ occurs within a subformula λ^{fin} of some formula $\mathsf{elim}^{fin}_{-\circ}(M_1, M_2)$, hence inside a formula of the form $\mathsf{alloc}(x) \vee (|h| < |U| - q \wedge |U| \geq r)$. Thus C (hence C') contains either $|h| < |U| - q$ or $|U| \geq r$, and necessarily, $q, r \leq \mathcal{N}(\chi_2) \leq \mathcal{N}(\chi)$. But $\mathcal{S}$ has more than $\mathcal{N}(\chi)$ unallocated elements, hence $\mathcal{S} \models (|h| < |U| - q \wedge |U| \geq r)$. Therefore, $(\mathfrak{U}, \mathfrak{s}', \mathcal{I}, \mathfrak{h}) \models C'$, which contradicts our previous assumption.

Consequently, the initial problem boils down to testing whether φ' has a finite model. It is clear that φ' is BSR-compatible (since by definition all positive occurrences of $\mathsf{alloc}(x)$ have been removed), hence by Lemma 5.28, it is sufficient to test whether $\tau(\varphi') \wedge \mathcal{A}(\varphi')$ has a finite model. By Proposition 4.10, the formula $\tau(\varphi') \wedge \mathcal{A}(\varphi')$ is equivalent to a formula in BSR(FO). We have $\mathcal{N}(\varphi') \leq \mathcal{N}(\forall y_1, \ldots, y_m \,.\, \chi)$, hence, using Propositions 2.2, 4.10 and 5.36 we deduce as it is done in the proof of Theorem 6.11, that $\tau(\varphi') \wedge \mathcal{A}(\varphi')$ has a finite model iff it has a model $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$, with $||\mathfrak{U}|| = O(k \cdot \mathsf{size}(\varphi)^2)$.

The algorithm is then defined as follows (see the proof of Theorem 6.11 for details). We guess an FO-structure $(\mathfrak{U}, \mathfrak{s}, \mathcal{I})$ satisfying *Heap* such that $||\mathfrak{U}|| = O(k \cdot \mathsf{size}(\varphi)^2)$ and check in polynomial space that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \tau(\varphi')$ (this is done as in Lemma 6.3, except that the test formulæ $\mathsf{alloc}(x)$ are replaced by $\top$) and that $(\mathfrak{U}, \mathfrak{s}, \mathcal{I}) \models \mathcal{A}(\varphi')$ (using Proposition 6.2) . □

7 CONCLUSION

We have studied the decidability problem for SL formulæ with quantifier prefix in the language $\exists^*\forall^*$, denoted as $\mathsf{BSR}(\mathsf{SL}^k)$, for finite and infinite universes, in the presence of uninterpreted predicate symbols. Although both problems were found to be undecidable, we identified two non-trivial subfragments for which the infinite and finite satisfiability are PSPACE-complete. These fragments are defined by restricting the polarity of occurrences of separating implications as well as the occurrence of universally quantified variables within the scope of separating implications. In both cases, the number of record fields k may be part of the input, but we assume that the arity of the uninterpreted predicates is bounded by a constant. If the latter condition does not hold, then the provided algorithms run in exponential space, and the problem is NEXPTIME-complete. Note that the PSPACE-completeness results for $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$ and $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$ allow us to (re-)establish the PSPACE-membership of the satisfiability problem for quantifier-free formulæ of SL^k, both in finite and infinite domains. Indeed, every

quantifier-free formula ϕ is sat-equivalent to a formula $\phi \multimap \top$ that is both in $\mathsf{BSR}^{fin}(\mathsf{SL}^k)$ and $\mathsf{BSR}^{inf}(\mathsf{SL}^k)$, since the left-hand side of $-\!\!*$ has neutral polarity.

Future work includes the implementation of an effective procedure for testing satisfiability of BSR(SL) formulæ in the above fragments. Since a non deterministic algorithm based on a guess-and-check approach is not practical, such a procedure could rely either on an encoding in QBF based on the finite model property derived in the present paper, or on some compact computational representations of boolean combinations of test formulæ. The bottleneck of the approach is certainly the computation of equivalent boolean combinations of test formulæ. To make the transformation more efficient, refined versions of Lemmas 5.16 and 5.20 could be derived, getting rid of some hypotheses such as E-completeness or A-completeness (as enforcing these hypotheses yield an exponential blow-up). Instead, the needed test formulæ could be added on demand, only if needed.

An extension of the presented results to formulæ containing inductively defined predicates (such as singly-linked lists) or interpreted predicates or functions (such as arithmetic symbols) will also be considered. This would allow us to extend existing approaches to test satisfiability of such formulæ [5**?**] to formulæ containing negation.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of Stéphane Demri and Étienne Lozes to the insightful discussions during the early stages of this work.

REFERENCES

[1] Timos Antonopoulos, Nikos Gorogiannis, Christoph Haase, Max Kanovich, and Joël Ouaknine. 2014. Foundations for Decision Problems in Separation Logic with General Inductive Predicates. In *Foundations of Software Science and Computation Structures*, Anca Muscholl (Ed.). Springer Berlin Heidelberg, 411–425.

[2] Sanjeev Arora and Boaz Barak. 2009. *Computational Complexity - A Modern Approach*. Cambridge University Press.

[3] Egon Börger, Erich Grädel, and Yuri Gurevich. 1997. *The Classical Decision Problem*. Springer.

[4] James Brotherston, Dino Distefano, and Rasmus L. Petersen. 2011. Automated Cyclic Entailment Proofs in Separation Logic. In *Automated Deduction – CADE-23: 23rd International Conference on Automated Deduction, Wrocław, Poland, July 31 - August 5, 2011, Proceedings*. Springer Berlin Heidelberg, Berlin, Heidelberg, 131–146.

[5] James Brotherston, Carsten Fuhs, Juan A. Navarro Pérez, and Nikos Gorogiannis. 2014. A Decision Procedure for Satisfiability in Separation Logic with Inductive Predicates. In *Proceedings of the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) (CSL-LICS '14)*. ACM, Article 25, 25:1–25:10 pages.

[6] Cristiano Calcagno, Philippa Gardner, and Matthew Hague. 2005. From Separation Logic to First-Order Logic. In *Foundations of Software Science and Computational Structures*. Springer Berlin Heidelberg, Berlin, Heidelberg, 395–409.

[7] Cristiano Calcagno, Hongseok Yang, and Peter W. O'Hearn. 2001. Computability and Complexity Results for a Spatial Assertion Language for Data Structures. In *FST TCS 2001: Foundations of Software Technology and Theoretical Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 108–119.

[8] Stéphane Demri and Morgan Deters. 2016. Expressive Completeness of Separation Logic with Two Variables and No Separating Conjunction. *ACM Trans. Comput. Log.* 17, 2 (2016), 12:1–12:44. https://doi.org/10.1145/2835490

[9] Stéphane Demri, Étienne Lozes, and Alessio Mansutti. 2018. The Effects of Adding Reachability Predicates in Propositional Separation Logic. In *Foundations of Software Science and Computation Structures - 21st International Conference, FOSSACS 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings (Lecture Notes in Computer Science)*, Christel Baier and Ugo Dal Lago (Eds.), Vol. 10803. Springer, 476–493.

[10] Mnacho Echenim, Radu Iosif, and Nicolas Peltier. 2019. Prenex Separation Logic with One Selector Field. In *Automated Reasoning with Analytic Tableaux and Related Methods - 28th International Conference, TABLEAUX 2019, London, UK, September 3-5, 2019, Proceedings (Lecture Notes in Computer Science)*, Serenella Cerrito and Andrei Popescu (Eds.), Vol. 11714. Springer, 409–427.

[11] Mnacho Echenim, Radu Iosif, and Nicolas Peltier. 2019. The Bernays-Schönfinkel-Ramsey Class of Separation Logic on Arbitrary Domains. In *Foundations of Software Science and Computation Structures - 22nd International Conference, FOSSACS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings (Lecture Notes in Computer Science)*, Mikolaj Boja'nczyk and Alex Simpson (Eds.), Vol. 11425. Springer, 242–259. https://doi.org/10.1007/978-3-030-17127-8_14

[] P. Erdós and R. Rado. 1952. Combinatorial Theorems on Classifications of Subsets of a Given Set. *Proceedings of the London Mathematical Society* s3-2, 1 (1952), 417–439. https://doi.org/10.1112/plms/s3-2.1.417 arXiv:https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s3-2.1.417

[12] Pascal Fontaine. 2007. Combinations of Theories and the Bernays-Schönfinkel-Ramsey Class. In *Proceedings of 4th International Verification Workshop in connection with CADE-21, Bremen, Germany, July 15-16, 2007 (CEUR Workshop Proceedings)*, Bernhard Beckert (Ed.), Vol. 259. CEUR-WS.org. http://ceur-ws.org/Vol-259/paper06.pdf

[13] Radu Iosif, Adam Rogalewicz, and Jiri Simacek. 2013. The Tree Width of Separation Logic with Recursive Definitions. In *Proc. of CADE-24 (LNCS)*, Vol. 7898. Springer.

[14] Samin S Ishtiaq and Peter W O'Hearn. 2001. BI as an assertion language for mutable data structures. In *ACM SIGPLAN Notices*, Vol. 36. ACM, 14–26.

[15] Jens Katelaan, Christoph Matheja, and Florian Zuleger. 2019. Effective Entailment Checking for Separation Logic with Inductive Definitions. In *Tools and Algorithms for the Construction and Analysis of Systems - 25th International Conference, TACAS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings, Part II (Lecture Notes in Computer Science)*, Tomás Vojnar and Lijun Zhang (Eds.), Vol. 11428. Springer, 319–336.

[] Quang Loc Le, Makoto Tatsuta, Jun Sun, and Wei-Ngan Chin. 2017. A Decidable Fragment in Separation Logic with Inductive Predicates and Arithmetic. In *Computer Aided Verification - 29th International Conference, CAV 2017, Heidelberg, Germany, July 24-28, 2017, Proceedings, Part II (Lecture Notes in Computer Science)*, Rupak Majumdar and Viktor Kuncak (Eds.), Vol. 10427. Springer, 495–517.

[] Harry R. Lewis. 1980. Complexity results for classes of quantificational formulas. *J. Comput. System Sci.* 21, 3 (1980), 317 – 353.

[16] Etienne Lozes. 2004. Separation Logic preserves the Expressive Power of Classical Logic. In *SPACE*.

[] Peter W. O'Hearn, John C. Reynolds, and Hongseok Yang. 2001. Local Reasoning about Programs that Alter Data Structures. In *Computer Science Logic, 15th International Workshop, CSL 2001. 10th Annual Conference of the EACSL, Paris, France, September 10-13, 2001, Proceedings*. 1–19.

[17] C.H. Papadimitriou. 1994. *Computational Complexity*. Addison-Wesley. https://books.google.fr/books?id=JogZAQAAIAAJ

[18] F. P. Ramsey. 1987. On a Problem of Formal Logic. *Classic Papers in Combinatorics* (1987), 1–24.

[19] Andrew Reynolds, Radu Iosif, and Cristina Serban. 2017. Reasoning in the Bernays-Schönfinkel-Ramsey Fragment of Separation Logic. In *Verification, Model Checking, and Abstract Interpretation*, Ahmed Bouajjani and David Monniaux (Eds.). Springer International Publishing, Cham, 462–482.

[20] John C. Reynolds. 2002. Separation Logic: A Logic for Shared Mutable Data Structures. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science (LICS '02)*. IEEE Computer Society, 55–74.

[21] Walter J. Savitch. 1970. Relationships between nondeterministic and deterministic tape complexities. *J. Comput. System Sci.* 4, 2 (1970), 177 – 192.

[22] Moshe Y. Vardi. 1982. The Complexity of Relational Query Languages. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing, May 5-7, 1982, San Francisco, California, USA*. ACM, 137–146.