A Quadratic Reduction of Constraints over Nested Sets to Purely Boolean Formulae in CNF ⋆ ⋆⋆

Domenico Cantone[0000−0002−1306−1166] [1], Andrea De Domenico[2],
Pietro Maugeri[1], and Eugenio G. Omodeo[0000−0003−3917−1942] [3]

[1] Dept. of Mathematics and Computer Science, University of Catania, Italy
domenico.cantone@unict.it, pietro.maugeri@unict.it
[2] Scuola Superiore di Catania, University of Catania, Italy
andrea.dedomenico@studium.unict.it
[3] Dept. of Mathematics and Earth Sciences, University of Trieste, Italy
eomodeo@units.it

Abstract. A translation is proposed of conjunctions of literals of the forms $x = y \setminus z$, $x \neq y \setminus z$, and $x \in y$, where x, y, z stand for variables ranging over the von Neumann universe of sets, into unquantified Boolean formulae of a rather simple conjunctive normal form. The formulae in the target language involve variables ranging over a Boolean field of sets, along with a difference operator and relators designating equality, non-disjointness and inclusion. Moreover, the result of each translation is a conjunction of literals of the forms $x = y \setminus z$, $x \neq y \setminus z$ and of implications whose antecedents are isolated literals and whose consequents are either inclusions (strict or non-strict) between variables, or equalities between variables.

Besides reflecting a simple and natural semantics, which ensures satisfiability-preservation, the proposed translation has quadratic algorithmic time-complexity, and bridges two languages both of which are known to have an NP-complete satisfiability problem.

Keywords: Satisfiability problem, Computable set theory, Expressibility, Proof verification, NP-completeness.

Introduction

The *Flatland* of sets [1,4] is inhabited by collections formed by entities named ‘urelements’. Two collections are different when the urelements belonging to either one, and to the other, are not the same; collections may even differ in cardinality, namely in the number of constituting urelements. One can conceive of inclusion between collections, and of operations which they can undergo (intersection, union, difference, etc.), on similar grounds: all of these constructs, in

⋆ Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

⋆⋆ We gratefully acknowledge partial support from project “STORAGE—Università degli Studi di Catania, Piano della Ricerca 2020/2022, Linea di intervento 2”, and from INdAM-GNCS 2019 and 2020 research funds.

fact, rely upon membership. It is more natural, instead, to think of equality comparison between urelements as of a primitive operation, because urelements are devoid of any inner structure. Such shapeless entities vanished, insofar as useless, from the *von Neumann universe*: sets can be nested one inside another to an arbitrary depth in this new world, and this is a resource that largely compensates for the missing urelements.

An apt theoretical framework for the study of flat sets is the theory of Boolean rings, a merely equational first-order theory endowed with finitely many axioms—at times, one blends this theory with an arithmetic of cardinals. Frameworks for the study of nested sets are such all-embracing theories as ZF and NBG (the Zermelo-Fraenkel and von Neumann-Bernays-Gödel theories), within which one can cast the entire corpus of mathematical disciplines.

Boolean algebra is decidable in its entirety (cf. [5, Sec. 3.7]); ZF is essentially undecidable, nonetheless an effort to find decision algorithms for fragments of it began in 1979, the purpose of this long-standing research being an effective implementation of specialized inference rules within a programmed system apt to verifying the correctness of large-scale mathematical proofs [6]. When it comes to implementations, complexity emerges as an unescapable issue; this is why we undertook, in recent years (see, in particular, [2,3]), a systematic study on the algorithmic complexities of inference mechanisms specifically designed for Boolean reasoning and of akin mechanisms which can cope with nested sets.

A priori, one would expect the distance between the performances of decision algorithms for fragments of Boolean algebra, and of the seemingly much more expressive languages whose dictionaries embody nested membership, to be abysmal. Luckily, though, as we will see, this is not the case.

We introduce in Sec. 1 an interpreted formal language, dubbed $\mathbb{BST}$, within which one can formulate unquantified Boolean constraints. Despite its syntax being quite minimal—$\mathbb{BST}$ only encompasses conjunctions of primitive literals of two forms, namely $x = y \setminus z$ and $x \neq y \setminus z$ —, the satisfiability problem for $\mathbb{BST}$ is NP-complete (see [2]). By way of abbreviations, a number of additional constraints, e.g. literals of the form $x \neq y \cup z$, can be expressed in $\mathbb{BST}$.

According to our semantics, the domain of discourse to which $\mathbb{BST}$ refers is a universe of nested sets; however, as will be seen in Sec. 2, every satisfiable propositional combination of $\mathbb{BST}$ literals (as a special case, a conjunctive $\mathbb{BST}$ constraint) admits a model consisting of sets which are, in a certain sense, “flat”. This makes it evident that membership cannot be plainly expressed in $\mathbb{BST}$. To detour this limitation, we propose in Sec. 1.2 a novel notion of expressibility, also embodying an obligation to supply an algorithmic-complexity assessment. This roundabout notion is, we believe, a valuable contribution of this paper.

In terms of the novel notion of expressibility, in Sec. 3 we will manage to translate a conjunction of literals of the three forms $x = y \setminus z$, $x \neq y \setminus z$, and $x \in y$, into a propositional combination of $\mathbb{BST}$ literals. The proposed translation is, of course, satisfiability preserving. It leads to a conjunction some of whose conjuncts are $\mathbb{BST}$ literals, while other conjuncts are rather simple disjunctions. The algorithmic time-complexity of our translation is quadratic, which indirectly shows that the satisfiability problem remains NP-complete when the relator $\in$ is added to the constructs of $\mathbb{BST}$: this NP-completeness result was known (see, e.g., [3]), but this paper sheds new light on it.

The material treated in this paper bridges, in a sense, the topics treated in [2] and in [3], which are meant to contribute to the proof-verification technology.

1 The theory $\mathbb{BST}$

Boolean Set Theory ($\mathbb{BST}$) is the quantifier-free theory composed by all conjunctions of literals of the following two types:

$$x = y \setminus z, \qquad x \neq y \setminus z, \tag{1}$$

where x, y, and z are *set variables* assumed to range over the universe of the well-founded sets.

Semantics for the theory $\mathbb{BST}$ is defined in terms of set assignments. Specifically, given a (finite) collection V of set variables, a *set assignment* M over V—the *variable-domain* of M, denoted by $\mathsf{dom}(M)$—is any map from V into the *von Neumann universe* $\mathcal{V}$ (see below).[4] A set assignment M *satisfies* a given literal $x = y \setminus z$, with $x, y, z \in \mathsf{dom}(M)$, if $Mx = My \setminus Mz$ holds, where $My \setminus Mz$ is the standard set difference between My and Mz. Likewise, M satisfies the literal $x \neq y \setminus z$ if $Mx \neq My \setminus Mz$ holds. Finally, M satisfies a $\mathbb{BST}$-conjunction φ such that $\mathit{Vars}(\varphi) \subseteq \mathsf{dom}(M)$ (where $\mathit{Vars}(\varphi)$ denotes the collection of the variables occurring free in φ) if it satisfies all of the conjuncts of φ, in which case we say that M is a *model* of φ and write $M \models \varphi$. A $\mathbb{BST}$-conjunction is said to be *satisfiable* if it has some model, otherwise it is said to be *unsatisfiable*.

In [2], it is proved that the satisfiability problem for $\mathbb{BST}$, namely the problem of establishing algorithmically the satisfiability status of any given $\mathbb{BST}$-conjunction, is NP-complete.

We shall also be interested in the extension $\mathbb{BST}^+$ of the theory $\mathbb{BST}$ consisting of all propositional combinations (resulting from unrestrained use of the logical connectives $\wedge, \vee, \longrightarrow, \longleftrightarrow, \neg$) of atomic formulae of type $x = y \setminus z$. It is not hard to check that the satisfiability problem for $\mathbb{BST}^+$ can be reduced to the satisfiability problem for $\mathbb{BST}$ in nondeterministic polynomial time, and therefore it is NP-complete in its turn.

1.1 The von Neumann universe

We recall that the von Neumann universe $\mathcal{V}$ of (well-founded) sets, also dubbed *von Neumann cumulative hierarchy*, is built up through a transfinite sequence of

[4] Notice that we are not basing our semantics of $\mathbb{BST}$ on flat sets of urelements (as would be doable, as recalled in the Introduction). Doing so would call for minor adjustments, unjustified—and perhaps disturbing—in the economy of this paper.

steps as the union $\mathcal{V} \coloneqq \bigcup_{\alpha \in On} \mathcal{V}_\alpha$ of the levels $\mathcal{V}_\alpha \coloneqq \bigcup_{\beta<\alpha} \mathscr{P}(\mathcal{V}_\beta)$, with $\mathscr{P}(\cdot)$ denoting the powerset operator and α ranging over the class On of all ordinals.

Based on the level of first appearance in the von Neumann hierarchy, one can define the rank of any set s, denoted $\mathsf{rk}\,(s)$. Specifically, $\mathsf{rk}\,(s)$ is the ordinal α such that $s \in \mathcal{V}_{\alpha+1} \setminus \mathcal{V}_\alpha$. Hence, for every $\alpha \in On$, the set $\mathcal{V}_{\alpha+1} \setminus \mathcal{V}_\alpha$, hereinafter denoted $\mathcal{V}^{\#}_\alpha$, collects all sets whose rank equals α.

The following lower bound on the number of well-founded sets of any positive integer rank n, to be proved as Proposition 2 in Appendix A.1, will be useful:

$$\left|\mathcal{V}^{\#}_n\right| \geqslant 2^{n-1}.$$

Some handy properties of the rank function which we shall tacitly use are the following:

for all sets $s, t \in \mathcal{V}$, we have:

- if $s \in t$ then $\mathsf{rk}\,(s) < \mathsf{rk}\,(t)$,
- if $s \subseteq t$ then $\mathsf{rk}\,(s) \leqslant \mathsf{rk}\,(t)$,
- $\mathsf{rk}\,(s) = \begin{cases} 0 & \text{if } s = \emptyset \\ \sup_{u \in s}(\mathsf{rk}\,(u)+1) & \text{otherwise.} \end{cases}$

We also recall that well-foundedness, as enforced by the *regularity* or *foundation axiom* of Zermelo-Fraenkel set theory, precludes the formation of infinite descending membership chains of the form

$$\cdots \in s_2 \in s_1 \in s_0,$$

and in particular of membership cycles

$$s_0 \in s_n \in \cdots \in s_2 \in s_1 \in s_0,$$

for any sequence $s_0, s_1, s_2, \ldots$ of sets.

1.2 Existential expressibility and $\mathcal{O}(f)$-expressibility

Despite the conciseness of our presentation of $\mathbb{BST}$, it turns out (see [2]) that several other Boolean constructs, such as the ones in the following list of literals

$$\begin{array}{llllll} x = \varnothing, & x \subseteq y, & x = y \cap z, & x = y \cup z, & \textsc{Disj}(x,y), & x \subsetneq y, \\ x \neq \varnothing, & x \nsubseteq y, & x \neq y \cap z, & x \neq y \cup z, & \neg\textsc{Disj}(x,y), & \end{array} \tag{2}$$

can be expressed existentially in $\mathbb{BST}$, where $\textsc{Disj}(a,b)$ is a short for $a \cap b = \emptyset$.

Formally, *existential expressibility* is defined as follows (cf. [2], wherein several applications of this notion are reported).

Definition 1 (Existential expressibility). *A formula $\psi(\boldsymbol{x})$ is said to be* existentially expressible *in a theory $\mathcal{T}$ if there exists a $\mathcal{T}$-formula $\Psi(\,\boldsymbol{x},\boldsymbol{z}\,)$ such that*

$$\models \quad \psi(\,\boldsymbol{x}\,) \longleftrightarrow (\,\exists\boldsymbol{z}\,)\,\Psi(\,\boldsymbol{x},\boldsymbol{z}\,),$$

where $\boldsymbol{x}$ and $\boldsymbol{z}$ stand for tuples of set variables.

Existential expressibility has been generalized in [2] into the definition of $\mathcal{O}(f)$-expressibility. The latter notion enabled, in [2], a detailed complexity taxonomy of the subfragments of $\mathbb{BST}$.

Here we slightly generalize $\mathcal{O}(f)$-expressibility so that it copes with collections $\mathcal{C}$ of formulae, rather than with single formulae as its original definition did; another difference lies in the fact that the generalized notion has to do with a source theory $\mathcal{T}_1$ and a target theory $\mathcal{T}_2$, whereas [2] took it for granted that source and target were the same.

Definition 2 ($\mathcal{O}(f)$-expressibility). *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be any theories and $f\colon \mathbb{N}\to\mathbb{N}$ be a given map. A collection $\mathcal{C}$ of formulae is said to be $\mathcal{O}(f)$*-expressible from $\mathcal{T}_1$ into $\mathcal{T}_2$ *if there exists a map*

$$\langle\,\varphi(\,\boldsymbol{y}\,)\,,\,\psi(\,\boldsymbol{x}\,)\,\rangle \;\mapsto\; \varXi^{\psi}_{\varphi}(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\,) \qquad (3)$$

from $\mathcal{T}_1\times\mathcal{C}$ into $\mathcal{T}_2$, where no variable in $\boldsymbol{z}$ occurs in either $\boldsymbol{x}$ or $\boldsymbol{y}$, such that the following conditions are satisfied:

(a) the mapping (3) can be computed in $\mathcal{O}\big(f(|\varphi\wedge\psi|)\big)$-time,
(b) if $\varphi(\,\boldsymbol{y}\,)\wedge\varXi^{\psi}_{\varphi}(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\,)$ is satisfiable, so is $\varphi(\,\boldsymbol{y}\,)\wedge\psi(\,\boldsymbol{x}\,)$,
(c) $\models \big(\varphi(\,\boldsymbol{y}\,)\;\wedge\;\psi(\,\boldsymbol{x}\,)\big) \longrightarrow (\,\exists\boldsymbol{z}\,)\varXi^{\psi}_{\varphi}(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\,)$.

The main results in this paper are that membership atoms $x\in y$ are not existentially expressible in $\mathbb{BST}^{+}$, whereas any conjunction of membership atoms is $\mathcal{O}(n^2)$-expressible from $\mathbb{BST}$ into $\mathbb{BST}^{+}$.

2 Non-existential expressibility in $\mathbb{BST}^{+}$ of $x\in y$

In this section we show that membership atoms are not existentially expressible in $\mathbb{BST}^{+}$. Specifically, we prove that every satisfiable $\mathbb{BST}^{+}$-formula $\varPhi$ admits a "flat" model M, namely a model whose union $\bigcup_{x\in \mathit{Vars}(\varPhi)} Mx$ of values is made up of members all of the same positive rank. We deduce from this fact that $Mx\notin My$, for all $x,y\in \mathit{Vars}(\varPhi)$, namely that $M\not\models x\in y$.

Definition 3. *For every ordinal $\rho\geqslant 1$, a set assignment M over a collection V of variables is said to be ρ*-flat *if all sets in the* domain *$\bigcup\{Mx \mid x\in V\}$ of M have rank ρ.*

No membership atom $x\in y$ is satisfied by any ρ-flat set assignment:

Lemma 1. *Let M be a ρ-flat set assignment over a collection V of variables. Then $Mx \notin My$, for all $x, y \in V$.*

Proof. Because of the ρ-flatness of M, for every $x \in V$ either $\mathsf{rk}\,(Mx) = 0$ (when $Mx = \emptyset$) or $\mathsf{rk}\,(Mx) = \rho + 1$ (when $Mx \neq \emptyset$). Hence, in any case $\mathsf{rk}\,(Mx) \neq \rho$ (since by definition $\rho \geqslant 1$), and therefore $Mx \notin \bigcup\{My \mid y \in V\}$. □

A satisfiable $\mathbb{BST}^{+}$-formula always admits a ρ-flat model, for sufficiently large ρ. This is proved in the next lemma.

Lemma 2. *Every satisfiable formula Φ of $\mathbb{BST}^{+}$ admits a ρ-flat set model, for every $\rho \geqslant |\,Vars(\Phi)| + 1$.*

Proof. Let Φ be a satisfiable formula of $\mathbb{BST}^{+}$, and let M be a set model for Φ. Let ϕ^{+}_{M} be the conjunction of all the distinct atoms $x = y \setminus z$ occurring in Φ that are satisfied by M. Likewise, let ϕ^{-}_{M} be the conjunction of all the distinct literals $x \neq y \setminus z$ such that $x = y \setminus z$ occurs in Φ and $M \not\models x = y \setminus z$. Finally, let

$$\phi_M \coloneqq \phi^{+}_{M} \wedge \phi^{-}_{M}. \tag{4}$$

Plainly, M satisfies ϕ_M by construction. Additionally, by propositional reasoning, every set model for ϕ_M satisfies our initial formula Φ. Thus, it is enough to show that the conjunction ϕ_M admits a ρ-flat set model for every $\rho \geqslant n + 1$, where $n \coloneqq |\,Vars(\phi_M)| = |\,Vars(\Phi)|$.

We prove that ϕ_M admits a ρ-flat set model by contracting each nonempty region R_W of M of the form

$$R_W \coloneqq \bigcap\{Mx \mid x \in W\} \cup \bigcup\{My \mid y \in \mathit{Vars}(\phi_M) \setminus W\},$$

for $\emptyset \neq W \subseteq \mathit{Vars}(\phi_M)$, into a distinct singleton of rank $\rho + 1$ (hence containing a single member of rank ρ).

Since the map $x \mapsto 2^x - x$ is strictly increasing for $x \geqslant 1$, for every integer $\rho \geqslant n + 1$ we have

$$|\mathcal{V}^{\#}_{\rho}| = |\mathcal{V}_{\rho+1}| - |\mathcal{V}_{\rho}| = 2^{|\mathcal{V}_{\rho}|} - |\mathcal{V}_{\rho}| \geqslant 2^{|\mathcal{V}_{n+1}|} - |\mathcal{V}_{n+1}| = |\mathcal{V}_{n+2}| - |\mathcal{V}_{n+1}| = |\mathcal{V}^{\#}_{n+1}| \geqslant 2^n,$$

where the latter inequality follows from Proposition 2 (see Appendix A.1). Hence, there exists an injective map $\Im_\rho \colon \mathscr{P}(\mathit{Vars}(\phi_M)) \to \mathcal{V}^{\#}_{\rho}$ from the collection of the nonempty subsets of $\mathit{Vars}(\phi_M)$ into the family $\mathcal{V}^{\#}_{\rho}$ of the (hereditarily finite) sets of rank ρ.

Next, we define a set assignment M^{*}_{ρ} over $\mathit{Vars}(\phi_M)$ by putting

$$M^{*}_{\rho}x \coloneqq \{\Im_\rho(W) \mid \emptyset \neq W \subseteq \mathit{Vars}(\phi_M) \wedge R_W \neq \emptyset\}.$$

By construction, the assignment M^{*}_{ρ} is ρ-flat. In addition, it is not hard to check that, for every $\emptyset \neq W \subseteq \mathit{Vars}(\phi_M)$, the region R^{*}_{W} of M^{*} defined by

$$R^{*}_{W} \coloneqq \bigcap\{M^{*}_{\rho}x \mid x \in W\} \cup \bigcup\{M^{*}_{\rho}y \mid y \in \mathit{Vars}(\phi_M) \setminus W\}$$

is nonempty if and only if so is its corresponding region R_W of M. Thus, M^{*}_{ρ} satisfies ϕ_M. □

We are now ready to prove that membership atoms $x \in y$ are not existentially expressible in $\mathbb{BST}^{+}$.

Theorem 1. *The atom $x \in y$ is not existentially expressible in $\mathbb{BST}^{+}$.*

Proof. By way of contradiction, let us assume that $x \in y$ is existentially expressible by a formula $\Psi(x, y, \boldsymbol{z})$ involving only atoms of type $x' = y' \setminus z'$. Hence,

$$\models x \in y \longleftrightarrow (\exists \boldsymbol{z})\ \Psi(x, y, \boldsymbol{z}) \tag{5}$$

would hold.

Since $x \in y$ is trivially satisfiable, by (5) so would be $(\exists \boldsymbol{z})\ \Psi(x, y, \boldsymbol{z})$ and therefore $\Psi(x, y, \boldsymbol{z})$ would be satisfiable too. Thus, by Lemma 2, $\Psi(x, y, \boldsymbol{z})$ would be satisfied by a ρ-flat set assignment M^* for some $\rho \geqslant 1$, and therefore, by Lemma 1, $M^*x \notin M^*y$. Hence, $M^* \models (\exists \boldsymbol{z})\ \Psi(x, y, \boldsymbol{z}) \wedge x \notin y$, so that $M^* \not\models (\exists \boldsymbol{z})\ \Psi(x, y, \boldsymbol{z}) \longrightarrow x \in y$, contradicting (5). □

3 $\mathcal{O}(n^2)$-expressibility in $\mathbb{BST}^{+}$ of membership conjunctions

Conforming with Definition 2, we shall prove that any membership conjunction $\psi(\,\boldsymbol{x}\,)$ is $\mathcal{O}(n^2)$-expressible from $\mathbb{BST}$ into $\mathbb{BST}^{+}$ by exhibiting a map

$$\langle\, \varphi(\,\boldsymbol{y}\,)\,,\, \psi(\,\boldsymbol{x}\,)\,\rangle \;\mapsto\; \Xi^{\psi}_{\varphi}(\,\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\,)$$

with $\varphi(\,\boldsymbol{y}\,)$ in $\mathbb{BST}$, which can be computed in quadratic time and such that conditions (b) and (c) of Definition 2 are satisfied, where $\varphi(\,\boldsymbol{y}\,)$ ranges over the collection of $\mathbb{BST}$-conjunctions and the variables in $\boldsymbol{z}$ are distinct from those in $\boldsymbol{x}$ and in $\boldsymbol{y}$.

Thus, let $\varphi(\,\boldsymbol{y}\,)$ be any $\mathbb{BST}$-conjunction and $\psi(\,\boldsymbol{x}\,)$ be any conjunction of membership atoms. We let $\mathit{Left}(\psi)$ denote the collection of all the set variables x occurring in some membership atom $x \in y$ in ψ, for some variable y.

For each variable $x \in \mathit{Left}(\psi)$, we introduce a new distinct variable $\overline{x}$ (which is intended to represent the singleton $\{x\}$), and denote by $\overline{\boldsymbol{x}}$ their collection. In addition, for each $x \in \mathit{Vars}(\varphi \wedge \psi)$ we introduce a new distinct variable $\widetilde{x}$, and denote by $\widetilde{\boldsymbol{x}}$ their collection (these variables will enforce that $x \in y$ only if $\widetilde{x} \subsetneq \widetilde{y}$). Then we put:

$$\begin{aligned}
\Xi^{\psi}_{\varphi}(\,\boldsymbol{x}, \boldsymbol{y}, \overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}\,) \coloneqq{} & \bigwedge_{x \in y \text{ in } \psi} (\overline{x} \neq \varnothing \ \wedge\ \overline{x} \subseteq y) \quad \wedge \bigwedge_{\substack{x \in \mathit{Left}(\psi)\\ y \in \mathit{Vars}(\varphi \wedge \psi)}} (\neg\textsc{Disj}(\overline{x}, y) \longrightarrow \overline{x} \subseteq y) \\
& \wedge \bigwedge_{x, y \in \mathit{Left}(\psi)} (\neg\textsc{Disj}(\overline{x}, \overline{y}) \longrightarrow x = y) \quad \wedge \bigwedge_{x, y \in \mathit{Left}(\psi)} (x = y \longrightarrow \overline{x} = \overline{y}) \\
& \wedge \bigwedge_{\substack{x \in \mathit{Left}(\psi)\\ y \in \mathit{Vars}(\varphi \wedge \psi)}} (\overline{x} \subseteq y \longrightarrow \widetilde{x} \subsetneq \widetilde{y}) \quad \wedge \bigwedge_{x, y \in \mathit{Vars}(\varphi \wedge \psi)} (x = y \longrightarrow \widetilde{x} = \widetilde{y})
\end{aligned}$$

(thus, the list of variables $\boldsymbol{z}$ in Definition 2 results from the concatenation of the lists $\overline{\boldsymbol{x}}$ and $\widetilde{\boldsymbol{x}}$).

Plainly, $\varXi_{\varphi}^{\psi}$ is a $\mathbb{BST}^{+}$-formula[5], which satisfies the following proposition, implying condition (a) of Def. 2:

Lemma 3. $\varXi_{\varphi}^{\psi} = \varTheta\big(|\,\mathit{Vars}(\varphi \wedge \psi)|^{2}\big)$. □

The proof of this lemma is delayed to Sec. 3.3.

In the following subsections, we shall prove that

- if $\varphi(\,\boldsymbol{y}\,) \wedge \varXi_{\varphi}^{\psi}(\,\boldsymbol{x}, \boldsymbol{y}, \overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}\,)$ is satisfiable, then so is $\varphi(\,\boldsymbol{y}\,) \wedge \psi(\,\boldsymbol{x}\,)$, and
- every model of $\varphi(\,\boldsymbol{y}\,) \wedge \psi(\,\boldsymbol{x}\,)$ can be extended to a model of $\varXi_{\varphi}^{\psi}(\,\boldsymbol{x}, \boldsymbol{y}, \overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}\,)$,

thus showing that also conditions (b) and (c) of Definition 2 are fulfilled, and therefore proving that every membership conjunction is $\mathcal{O}(n^{2})$-expressible from $\mathbb{BST}$ into $\mathbb{BST}^{+}$.

Translation examples

Here we digress to provide a few examples illustrating how the conjunction $\varXi_{\varphi}^{\psi}$ renders the formula $\varphi \wedge \psi$.

Example 1. The simple conjunction $\varphi \wedge \psi$ where $\varphi \coloneqq x = y \setminus z$ and $\psi \coloneqq z \in y$, gets translated into

$$\begin{aligned}\varXi_{\varphi}^{\psi} \;\coloneqq\; & \overline{z} \neq \emptyset \;\wedge\; \overline{z} \subseteq y \wedge\; (\neg\textsc{Disj}(\overline{z}, x) \longrightarrow \overline{z} \subseteq x) \;\wedge \\ & (\neg\textsc{Disj}(\overline{z}, y) \longrightarrow \overline{z} \subseteq y) \;\wedge\; (\neg\textsc{Disj}(\overline{z}, z) \longrightarrow \overline{z} \subseteq z) \;\wedge \\ & (\overline{z} \subseteq x \longrightarrow \tilde{z} \subsetneq \tilde{x}) \;\wedge\; (\overline{z} \subseteq y \longrightarrow \tilde{z} \subsetneq \tilde{y}) \;\wedge\; (\overline{z} \subseteq z \longrightarrow \tilde{z} \subsetneq \tilde{z}) \;\wedge \\ & (x = y \longrightarrow \tilde{x} = \tilde{y}) \;\wedge\; (x = z \longrightarrow \tilde{x} = \tilde{z}) \;\wedge\; (y = z \longrightarrow \tilde{y} = \tilde{z}).\end{aligned}$$

This shows that any relation $Ms \in Mt$, respectively $Ms \notin Mw$, where M is a model for $\varphi \wedge \psi \wedge \varXi_{\varphi}^{\psi}$, gets translated into $M\overline{s} \subseteq Mt \wedge M\tilde{s} \subsetneq M\tilde{t}$, resp. into $\textsc{Disj}(\overline{s}, w)$.

In fact, to the literal $z \in y$ there corresponds the conjunct $\overline{z} \subseteq y$, so that from $\overline{z} \subseteq y \longrightarrow \tilde{z} \subsetneq \tilde{y}$ we also get $M\tilde{z} \subsetneq M\tilde{y}$. Furthermore $Mz \notin Mz$ must hold, and in fact we can derive $\textsc{Disj}(\overline{z}, z)$ from $(\overline{z} \subseteq z \longrightarrow \tilde{z} \subsetneq \tilde{z}) \wedge (\neg\textsc{Disj}(\overline{z}, z) \longrightarrow \overline{z} \subseteq z)$. Then we can get $Mz \in Mx$ from $\varphi \wedge \psi$, since $Mx = My \setminus Mz$ and $Mz \notin Mz$; we can, analogously, get $M\overline{z} \subseteq Mx$ from $\varphi \wedge \varXi_{\varphi}^{\psi}$: indeed, it follows from $x = y \setminus z$, $\overline{z} \subseteq y$, and from the condition $\textsc{Disj}(M\overline{z}, Mz)$ just obtained; lastly, from $\overline{z} \subseteq x \longrightarrow \tilde{z} \subsetneq \tilde{x}$, we get $M\tilde{z} \subsetneq M\tilde{x}$.

In the previous example we saw that the conjunct $\overline{z} \subseteq z$, that in our translation would be the equivalent of the unsatisfiable $z \in z$, does not appear in $\varXi_{\varphi}^{\psi}$. More generally, the usage of the set variables $\tilde{x}$ in $\varXi_{\varphi}^{\psi}$ allows us to detect any membership cycle $x \in \cdots \in x$ that might be derived from $\varphi \wedge \psi$. In the following we exemplify this property:

[5] In fact, $\varXi_{\varphi}^{\psi}$ is a conjunction of a rather simple form.

Example 2. The conjunction $\varphi \wedge \psi$, where

$$\varphi \coloneqq a = b \setminus c \qquad \text{and} \qquad \psi \coloneqq x \in y \ \wedge \ y \in z \ \wedge \ z \in x,$$

is plainly unsatisfiable due to the cycle $x \in y \in z \in x$. To reflect this, Ξ_{φ}^{ψ} comprises the literals

$$\begin{aligned}
&\overline{x} \subseteq y \ \wedge \ (\overline{x} \subseteq y \longrightarrow \tilde{x} \subsetneq \tilde{y}) \ \wedge \\
&\overline{y} \subseteq z \ \wedge \ (\overline{y} \subseteq z \longrightarrow \tilde{y} \subsetneq \tilde{z}) \ \wedge \\
&\overline{z} \subseteq x \ \wedge \ (\overline{z} \subseteq x \longrightarrow \tilde{z} \subsetneq \tilde{x});
\end{aligned}$$

therefore it is unsatisfiable due to the cycle $\tilde{x} \subsetneq \tilde{y} \subsetneq \tilde{z} \subsetneq \tilde{x}$.

Our next example shows how the unsatisfiability proofs for the conjunction $\varphi \wedge \psi$ and $\varphi \wedge \Xi_{\varphi}^{\psi}$ mimic each other, especially bearing in mind that $Mx \in My$ gets translated into $M\overline{x} \subseteq My \wedge M\tilde{x} \subsetneq M\tilde{y}$, and $Mx \notin Mz$ into $\textsc{Disj}(M\overline{x}, Mz)$.

Example 3. The conjunction $\varphi \wedge \psi$, where $\varphi \coloneqq x = y \setminus z$ and $\psi \coloneqq z \in x \wedge x \in y$, is unsatisfiable since if a model M for it existed, then we would have $Mx \in Mx$. Indeed: from $z \in x$ we obtain $Mx \notin Mz$, then from $x = y \setminus z$ and $x \in y$ we obtain $Mx \in Mx$. This will be reflected by $\varphi \wedge \Xi_{\varphi}^{\psi}$ as $M\tilde{x} \subsetneq M\tilde{x}$, which in its turn implies that also $\varphi \wedge \Xi_{\varphi}^{\psi}$ is unsatisfiable, where

$$\begin{aligned}
\Xi_{\varphi}^{\psi} \ \coloneqq\ & \overline{x} \neq \emptyset \ \wedge \ \overline{x} \subseteq y \ \wedge \ \overline{z} \neq \emptyset \ \wedge \boldsymbol{\overline{z} \subseteq x} \ \wedge \\
& (\neg\textsc{Disj}(\overline{x}, x) \longrightarrow \overline{x} \subseteq x) \ \wedge \ (\neg\textsc{Disj}(\overline{x}, y) \longrightarrow \overline{x} \subseteq y) \ \wedge \\
& \boldsymbol{(\neg\textsc{Disj}(\overline{x}, z) \longrightarrow \overline{x} \subseteq z)} \ \wedge \ (\neg\textsc{Disj}(\overline{z}, x) \longrightarrow \overline{z} \subseteq x) \ \wedge \\
& (\neg\textsc{Disj}(\overline{z}, y) \longrightarrow \overline{z} \subseteq y) \ \wedge \ (\neg\textsc{Disj}(\overline{z}, z) \longrightarrow \overline{z} \subseteq z) \ \wedge \\
& (\neg\textsc{Disj}(\overline{x}, \overline{z}) \longrightarrow \overline{x} = \overline{z}) \ \wedge \ (x = z \longrightarrow \overline{x} = \overline{z}) \ \wedge \\
& \boldsymbol{(\overline{x} \subseteq x \longrightarrow \tilde{x} \subsetneq \tilde{x})} \ \wedge \ (\overline{x} \subseteq y \longrightarrow \tilde{x} \subsetneq \tilde{y}) \ \wedge \\
& \boldsymbol{(\overline{x} \subseteq z \longrightarrow \tilde{x} \subsetneq \tilde{z})} \ \wedge \ \boldsymbol{(\overline{z} \subseteq x \longrightarrow \tilde{z} \subsetneq \tilde{x})} \ \wedge \\
& (\overline{z} \subseteq y \longrightarrow \tilde{z} \subsetneq \tilde{y}) \ \wedge \ (\overline{z} \subseteq z \longrightarrow \tilde{z} \subsetneq \tilde{z}) \ \wedge \\
& (x = z \longrightarrow \tilde{x} = \tilde{z}).
\end{aligned}$$

Above, in proving the unsatisfiability of $\varphi \wedge \psi$, our first step has been to get $Mx \notin Mz$; here, by assuming $M \models \varphi \wedge \Xi_{\varphi}^{\psi}$, we first get $\textsc{Disj}(M\overline{x}, Mz)$: indeed, from $\overline{z} \subseteq x$ and $\overline{z} \subseteq x \longrightarrow \tilde{z} \subsetneq \tilde{x}$ we obtain $M\tilde{z} \subsetneq M\tilde{x}$, then from $\overline{x} \subseteq z \longrightarrow \tilde{x} \subsetneq \tilde{z}$ and $\neg\textsc{Disj}(\overline{x}, z) \longrightarrow \overline{x} \subseteq z$ we obtain $\textsc{Disj}(M\overline{x}, Mz)$. The second step, above, has been to get $Mx \in Mx$; here we will get $M\tilde{x} \subsetneq M\tilde{x}$: indeed, from $\overline{x} \subseteq y$ and $x = y \setminus z$, and from the disjointness clause just proved, it follows that $M\overline{x} \subseteq Mx$; and, finally, from $\overline{x} \subseteq x \longrightarrow \tilde{x} \subsetneq \tilde{x}$ we obtain $M\tilde{x} \subsetneq M\tilde{x}$, which leads to the unsatisfiability of $\varphi \wedge \Xi_{\varphi}^{\psi}$.

3.1 If $\varphi \wedge \Xi_{\varphi}^{\psi}$ is satisfiable, then so is $\varphi \wedge \psi$

Assume that $\varphi \wedge \Xi_{\varphi}^{\psi}$ is satisfiable. Hence, by Lemma 2, $\varphi \wedge \Xi_{\varphi}^{\psi}$ is satisfied by some ρ-flat set assignment M such that $\rho \geqslant \left|\mathit{Vars}(\varphi \wedge \Xi_{\varphi}^{\psi})\right| + 1$. Since there can be no $\subsetneq$-cycle in $\{M\widetilde{x} \mid x \in \mathit{Vars}(\varphi \wedge \psi)\}$, we can find some ordering $\prec$ on $\mathit{Vars}(\varphi \wedge \psi)$ such that

$$M\widetilde{x} \subsetneq M\widetilde{y} \quad \longrightarrow \quad x \prec y,$$

for $x, y \in \mathit{Vars}(\varphi \wedge \psi)$.

Following the ordering $\prec$, define recursively, for $x \in \mathit{Vars}(\varphi \wedge \psi)$,

$$M'x \coloneqq Mx \cup M^{+}x, \tag{6}$$

where

$$M^{+}x \coloneqq \{M'w \mid M\overline{w} \subseteq Mx \wedge w \in \mathit{Left}(\psi)\}. \tag{7}$$

In preparation for the proof that M' models $\varphi \wedge \psi$, we need a few lemmas. We begin by proving that no $\mathsf{rk}\,(M'x)$ equals ρ, for any $x \in \mathit{Vars}(\varphi \wedge \psi)$:

Lemma 4. *For every $x \in \mathit{Vars}(\varphi \wedge \psi)$, $\mathsf{rk}\,(M'x) \neq \rho$.*

Proof. If $Mx \neq \emptyset$, then $\mathsf{rk}\,(M'x) \geqslant \mathsf{rk}\,(Mx) = \rho + 1$, so $\mathsf{rk}\,(M'x) \neq \rho$. On the other hand, if $Mx = \emptyset$, then $M^{+}x = \{M'w \mid M\overline{w} \subseteq Mx \wedge w \in \mathit{Left}(\psi)\} = \emptyset$, since $M \models \overline{w} \neq \emptyset$, for all $w \in \mathit{Left}(\psi)$. Hence, $M'x = \emptyset$, so that $\mathsf{rk}\,(M'x) = 0 \neq \rho$. □

An immediate consequence of the preceding claim is:

Corollary 1. *For all $x, y \in \mathit{Vars}(\varphi \wedge \psi)$,*

$$Mx \cap M^{+}y = \emptyset, \tag{8}$$

Proof. It is enough to observe that, for $x, y \in \mathit{Vars}(\varphi \wedge \psi)$, when $Mx \neq \emptyset$, all members of Mx have rank ρ, whereas by Lemma 4 no member of $M^{+}y$ can have rank ρ. □

Next we prove three lemmas which will enable us to conclude that M' models φ and ψ as wanted; the second of these relies on Proposition 3, whose statement and proof are delayed till Appendix A.2.

Lemma 5. *For all $w, w' \in \mathit{Left}(\psi)$, $\quad Mw = Mw' \Longleftrightarrow M\overline{w} = M\overline{w}'$.*

Proof. Let $w, w' \in \mathit{Left}(\psi)$. By construction, the formula Ξ^{ψ}_{φ} contains the following conjuncts, which are all satisfied by the set assignment M:

- $w = w' \;\longrightarrow\; \overline{w} = \overline{w}'$,
- $\overline{w} \neq \emptyset$,
- $\neg\textsc{Disj}(\overline{w}, \overline{w}') \;\longrightarrow\; w = w'$.

Thus, in view of $M \models w = w' \;\longrightarrow\; \overline{w} = \overline{w}'$, if $Mw = Mw'$ then we have $M\overline{w} = M\overline{w}'$.

Conversely, if $M\overline{w} = M\overline{w}'$ then from $M \models \overline{w} \neq \emptyset$ we get $M \models \neg\textsc{Disj}(\overline{w}, \overline{w}')$. Hence, from $M \models \neg\textsc{Disj}(\overline{w}, \overline{w}') \;\longrightarrow\; w = w'$, we get $Mw = Mw'$. □

Lemma 6. *For all $x, y \in \mathit{Vars}(\varphi \wedge \psi)$,* $\quad Mx = My \Longleftrightarrow M'x = M'y$.

Proof. If $Mx = My$, then by (7) we have $M^{+}x = M^{+}y$. Thus,

$$M'x = Mx \cup M^{+}x = My \cup M^{+}y = M'y.$$

Conversely, if $M'x = M'y$, then by (7) again we have $Mx \cup M^{+}x = My \cup M^{+}y$, so that Corollary 1 and Proposition 3(b) (see Appendix A.2) yield $Mx = My$. $\square$

Another useful consequence of definitions (6) and (7) is the following result.

Lemma 7. *For $w \in \mathit{Left}(\psi)$ and $x \in \mathit{Vars}(\varphi \wedge \psi)$, the following statements are equivalent:*

(a) $M'w \in M^{+}x$,
(b) $M\overline{w} \subseteq Mx$, and
(c) $M\overline{w} \cap Mx \neq \emptyset$.

Proof. The implication (b) $\Longrightarrow$ (a) follows readily from the definition (7) of M^{+}.

Concerning the implication (a) $\Longrightarrow$ (c), let us assume that $M'w \in M^{+}x$, for some $w \in \mathit{Left}(\psi)$ and $x \in \mathit{Vars}(\varphi \wedge \psi)$. Then, by (7), there must exist a $w' \in \mathit{Left}(\psi)$ such that (i) $M'w = M'w'$ and (ii) $M\overline{w}' \subseteq Mx$. By (i) and Lemma 6, we have $Mw = Mw'$, which in turn by Lemma 5 implies $M\overline{w} = M\overline{w}'$. Hence, by (ii), $M\overline{w} \subseteq Mx$. Next, since $\overline{w} \neq \emptyset$ is in $\overline{\varphi}$, we have $M\overline{w} \neq \emptyset$, so that the inclusion $M\overline{w} \subseteq Mx$ yields $M\overline{w} \cap Mx \neq \emptyset$.

Finally, the implication (c) $\Longrightarrow$ (a) follows at once from (7), since $\overline{\varphi}$ contains the conjunct $\neg\textsc{Disj}(\overline{w}, x) \longrightarrow \overline{w} \subseteq x$. $\square$

We have now reached the salient conclusion yielded by the definition of M' and by the preparatory proofs carried out so far:

Lemma 8. *The set assignment M' satisfies the conjunction $\varphi \wedge \psi$.*

Proof. We shall prove that M' satisfies

(a) all literals in ψ of type $x \in y$,
(b) all literals in φ of type $x = y \setminus z$,
(c) all literals in φ of type $x \neq y \setminus z$.

Concerning (a), let $x \in y$ be a conjunct in ψ. Then the literal $\overline{x} \subseteq y$ occurs in $\varXi_{\varphi}^{\psi}$, so that $M\overline{x} \subseteq My$ holds, and therefore by (7) and (6) we have $M'x \in M^{+}y \subseteq M'y$. Hence, $M' \models x \in y$.

Concerning (b), let $x = y \setminus z$ be in φ. Recalling that $M \models \varphi$, we have $Mx = My \setminus Mz$. Hence, by (6), Corollary 1, and Proposition 3(a), in order to show that M' satisfies the literal $x = y \setminus z$ it is enough to prove that $M^{+}x = M^{+}y \setminus M^{+}z$ holds, which we do next.

If $M'w \in M^{+}x$, for some $w \in \mathit{Left}(\psi)$ such that $M\overline{w} \subseteq Mx$, then $M\overline{w} \subseteq My \setminus Mz$, so that $M\overline{w} \subseteq My$ and $M\overline{w} \cap Mz = \emptyset$. By Lemma 7, $M\overline{w} \subseteq My$ yields

$M'w \in M^+y$ and $M\overline{w} \cap Mz = \emptyset$ implies $M'w \notin M'z$. Thus, $M'w \in M^+y \setminus M^+z$. By the arbitrariness of $w \in \mathit{Left}(\psi)$, we get

$$M^+x \subseteq M^+y \setminus M^+z. \tag{9}$$

Conversely, if $M'w \in M^+y \setminus M^+z$ for some $w \in \mathit{Left}(\psi)$, then, again by Lemma 7, we have $M\overline{w} \subseteq My$ and $M\overline{w} \cap Mz = \emptyset$. Hence, $M\overline{w} \subseteq My \setminus Mz = Mx$, so that by another application of Lemma 7 we get $M'w \in M^+x$. The arbitrariness of $w \in \mathit{Left}(\psi)$ yields $M^+y \setminus M^+z \subseteq M^+x$. Together with (9), the latter inclusion implies $M^+x = M^+y \setminus M^+z$, completing the proof of (b).

Finally, concerning (c), let $x \neq y \setminus z$ be in φ. Then we have $Mx \neq My \setminus Mz$, so that by Proposition 3(a) we readily obtain $M'x \neq M'y \cup M'z$. □

3.2 Every model of $\varphi \wedge \psi$ can be extended into a model of Ξ^{ψ}_{φ}

Assume that $\varphi \wedge \psi$ is satisfiable, and let M be a model for it.

Let $V \coloneqq \mathit{Vars}(\varphi \wedge \psi)$ and let $G_M = (V, E)$ be the directed graph over V such that

$$(x, y) \in E \iff Mx \in My. \tag{10}$$

Plainly, the graph G_M is acyclic. For, should G_M contain a cycle $(x_{i_0}, x_{i_1}, \ldots, x_{i_k}, x_{i_0})$, then we would have the membership cycle $Mx_{i_0} \in Mx_{i_1} \in \cdots \in Mx_{i_k} \in Mx_{i_0}$, contradicting the axiom of foundation.

Hence, we can define the following notion of *height* $h \colon V \to \mathbb{N}$ by putting

$$h(x) \coloneqq \text{length of the longest path in } G_M \text{ leading to } x,$$

for every $x \in V$ (in particular, $h(x) = 0$ whenever x has no predecessors in G_M).

Next let $x_1, x_2, ..., x_n$ be any indexing of the variables in $\mathit{Vars}(\varphi \wedge \psi)$ complying with the height h, namely such that

$$h(x_i) < h(x_j) \quad \Longrightarrow \quad i < j.$$

We are now ready to define an extension $\overline{M}$ over $\mathit{Vars}(\Xi^{\psi}_{\varphi}) \setminus \mathit{Vars}(\varphi \wedge \psi)$ of the set assignment M which satisfies Ξ^{ψ}_{φ}. For $x \in \mathit{Vars}(\varphi \wedge \psi)$, we put of course $\overline{M}x \coloneqq Mx$. Then, for $x \in \mathit{Left}(\psi)$, we set $\overline{M}\overline{x} \coloneqq \{Mx\}$. Finally, we put $\overline{M}\widetilde{x}_1 \coloneqq \{1\}$ and recursively, for $i = 1, \ldots, n-1$,

$$\overline{M}\widetilde{x}_{i+1} \coloneqq \begin{cases} \overline{M}\widetilde{x}_i & \text{if } h(x_{i+1}) = h(x_i)\,, \\ \overline{M}\widetilde{x}_i \cup \{i+1\} & \text{otherwise.} \end{cases}$$

From the definition of $\overline{M}$, it follows that, for $i, j \in \{1, \ldots, n\}$:

(H$_1$) $h(x_i) < h(x_j) \implies \overline{M}\widetilde{x}_i \subsetneq \overline{M}\widetilde{x}_j$, and
(H$_2$) $h(x_i) = h(x_j) \implies \overline{M}\widetilde{x}_i = \overline{M}\widetilde{x}_j$.

Next, we prove that $\overline{M}$ satisfies Ξ^{ψ}_{φ}.

Lemma 9. *The set assignment $\overline{M}$ satisfies Ξ^{ψ}_{φ}.*

Proof. Let $x \in y$ occur in ψ. By construction, $\overline{M}\overline{x} = \{Mx\}$. In addition, since $M \models x \in y$, we have $Mx \in My = \overline{M}y$, so that $\overline{M}\overline{x} \subseteq \overline{M}y$. Thus, by the arbitrariness of $x \in y$ in ψ, we have

$$\overline{M} \models \bigwedge_{x \in y \text{ in } \psi} (\overline{x} \neq \varnothing \ \wedge \ \overline{x} \subseteq y). \tag{11}$$

Let now $x \in \mathit{Left}(\psi)$ and $y \in \mathit{Vars}(\varphi \wedge \psi)$, and assume that $\overline{M} \models \neg\textsc{Disj}(\overline{x}, y)$, namely $\overline{M}\overline{x} \cap My \neq \emptyset$. Hence, $\{Mx\} \cap My \neq \emptyset$, so that $\overline{M}\overline{x} = \{Mx\} \subseteq My$. Thus, by the arbitrariness of $x \in \mathit{Left}(\psi)$ and $y \in \mathit{Vars}(\varphi \wedge \psi)$, we have

$$\overline{M} \models \bigwedge_{\substack{x \in \mathit{Left}(\psi) \\ y \in \mathit{Vars}(\varphi \wedge \psi)}} (\neg\textsc{Disj}(\overline{x}, y) \longrightarrow \overline{x} \subseteq y). \tag{12}$$

Next, let $x, y \in \mathit{Left}(\psi)$ and assume that $\overline{M} \models \neg\textsc{Disj}(\overline{x}, \overline{y})$, namely $\overline{M}\overline{x} \cap \overline{M}\overline{y} \neq \emptyset$. Since by construction $\overline{M}\overline{x} = \{Mx\}$ and $\overline{M}\overline{y} = \{My\}$, it follows that $Mx = My$, and therefore $\overline{M}x = \overline{M}y$. Hence, by the arbitrariness of $x, y \in \mathit{Left}(\psi)$, we have

$$\overline{M} \models \bigwedge_{x, y \in \mathit{Left}(\psi)} (\neg\textsc{Disj}(\overline{x}, \overline{y}) \longrightarrow x = y). \tag{13}$$

Let $x, y \in \mathit{Left}(\psi)$, but assume now that $\overline{M}\overline{x} = \overline{M}\overline{y}$, so that $Mx = My$ holds. Thus, $\overline{M}x = \{Mx\} = \{My\} = \overline{M}y$, proving that

$$\overline{M} \models \bigwedge_{x, y \in \mathit{Left}(\psi)} (x = y \longrightarrow \overline{x} = \overline{y}), \tag{14}$$

by the arbitrariness of $x, y \in \mathit{Left}(\psi)$.

Next, let $x \in \mathit{Left}(\psi)$ and $y \in \mathit{Vars}(\varphi \wedge \psi)$ be such that $\overline{M} \models \overline{x} \subseteq y$, i.e., $\overline{M}\overline{x} \subseteq \overline{M}y$. Hence, we have $\{Mx\} \subseteq My$, and therefore $Mx \in My$. The latter membership relation implies that the graph G_M associated with the assignment M contains the arc (x, y), and so $h(x) < h(y)$. Thus, by (H_1) we have $\overline{M}\widetilde{x} \subsetneq \overline{M}\widetilde{y}$, and therefore we have

$$\overline{M} \models \bigwedge_{\substack{x \in \mathit{Left}(\psi) \\ y \in \mathit{Vars}(\varphi \wedge \psi)}} (\overline{x} \subseteq y \longrightarrow \widetilde{x} \subsetneq \widetilde{y}), \tag{15}$$

by the arbitrariness of $x \in \mathit{Left}(\psi)$ and $y \in \mathit{Vars}(\varphi \wedge \psi)$.

Finally, let $x, y \in \mathit{Vars}(\varphi \wedge \psi)$ and assume that $\overline{M}x = \overline{M}y$, so that $Mx = My$. By (10), the nodes in G_M labeled x and y have the same predecessors. Therefore, $h(x) = h(y)$, so that by (H_2) we have $\overline{M}\widetilde{x} = \overline{M}\widetilde{y}$. Hence, by the arbitrariness of $x, y \in \mathit{Vars}(\varphi \wedge \psi)$, we have

$$\overline{M} \models \bigwedge_{x, y \in \mathit{Vars}(\varphi \wedge \psi)} (x = y \longrightarrow \widetilde{x} = \widetilde{y}). \tag{16}$$

From (11)–(16), it follows that the assignment $\overline{M}$ satisfies $\overline{\varphi}$. □

Summing up, from Lemmas 3, 8, and 9, we have:

Theorem 2. *Membership conjunctions are $\mathcal{O}(n^2)$-expressible from $\mathbb{BST}$ into $\mathbb{BST}^+$.* □

3.3 Design and analysis of the translation algorithm

In order to prove Lemma 3, we provide a detailed specification of the algorithm that generates from the conjunction $\varphi \wedge \psi$ the formula Ξ_{φ}^{Ψ}.

```
1: Initialize $Vars(\varphi \wedge \psi)$ and $Left(\psi)$ as empty lists of set variables;
2: Initialize $\Xi_{\varphi}^{\psi}$ as an empty list of conjuncts;
3: for all set variable $x$ that appears in $\varphi$ do
4:     add $x$ to $Vars(\varphi \wedge \psi)$;
5: for all conjunct $x \in y$ that appears in $\psi$ do
6:     add $x$ and $y$ to $Vars(\varphi \wedge \psi)$;
7:     add $x$ to $Left(\psi)$;
8:     add ($\overline{x} \neq \emptyset \ \wedge \ \overline{x} \subseteq y$ to $\Xi_{\varphi}^{\psi}$);
9: for all $x \in Left(\psi)$ do
10:     for all $y \in Vars(\varphi \wedge \psi)$ do
11:         add ($\neg\textsc{Disj}(\overline{x}, y) \longrightarrow \overline{x} \subseteq y \wedge \overline{x} \subseteq y \longrightarrow \tilde{x} \subsetneq \tilde{y}$) to $\Xi_{\varphi}^{\psi}$;
12: for all $x, y \in Left(\psi)$ do
13:     add ($\neg\textsc{Disj}(\overline{x}, \overline{y}) \longrightarrow x = y \ \wedge \ x = y \longrightarrow \overline{x} = \overline{y} \ \wedge x = y \longrightarrow \tilde{x} = \tilde{y}$) to $\Xi_{\varphi}^{\psi}$.
```

Adding elements to $Vars(\varphi \wedge \psi)$, $Left(\psi)$, and Ξ_{φ}^{ψ} will require constant time if these are implemented as lists of set variables and conjuncts.

The **for**-loop at lines 3 and 4 can be performed in $\Theta(|\varphi|)$-time, where $|\varphi|$ is the total lenght of the conjunction φ; similarly the **for**-loop from line 5 to line 8 can be performed in $\Theta(|\psi|)$-time. The **for**-loop from line 9 to line 11 is iterated $\Theta(|Left(\psi) \times Vars(\varphi \wedge \psi)|)$ times and the **for**-loop at lines 12 and 13 is iterated $\Theta\big(\,|Left(\psi)|^2\,\big)$ times.

The overall time complexity is then $\Theta\big(\,|\varphi \wedge \psi| + |Left(\psi) \times Vars(\varphi)| + |Left(\psi)|^2\,\big)$, and since most commonly a conjunction $\varphi \wedge \psi$ is such that $|\varphi \wedge \psi| = \mathcal{O}\big(\,|\,Vars(\varphi \wedge \psi)|^2\,\big)$ and $|Left(\psi)| = \Theta(|\,Vars(\varphi \wedge \psi)|)$, we can say that Ξ_{φ}^{ψ} can be generated in $\Theta\big(\,|\,Vars(\varphi \wedge \psi)|^2\,\big)$ time.

4 Future work

By a technique close to to the one proposed above for translating conjunctions of literals of the forms $x = y \setminus z$, $x \neq y \setminus z$, and $x \in y$, it is possible to translate conjunctions of literals of the three forms $x = y \setminus z$, $x \neq y \setminus z$, and $x = \{\, y\,\}$ into propositional combinations of $\mathbb{BST}$ literals. (This is an enhancement proper of the translation: in fact, $x \in y$ can be restated as $s = \{\, x\,\} \wedge z = z \backslash z \wedge z = s \backslash y$.)

Moreover, we will strive to enhance the nested-to-flat translation so as to enable it to handle rank comparison and cardinality comparison constructs.

We also have in mind a linear-cost flat-to-nested translation, exploiting membership to eliminate the equality relator from conjunctions of $\mathbb{BST}$ literals in terms of membership literals.

With Mattia Furlan, who recently earned a bachelor degree from the University of Trieste, we have spotted out the following *valid* formulae involving Boolean difference:

(**D.1**)	$x \setminus (y \setminus y) = x$	Existence of zero
(**D.2**)	$(x \setminus y) \setminus z = (x \setminus z) \setminus y$	Permutativity
(**D.3**)	$x \setminus (x \setminus y) = y \setminus (y \setminus x)$	Commutativity (of intersection)
(**D.4**)	$(x \setminus y) \setminus y = x \setminus y$	Double subtraction

Let us take the universal closures of these formulae as the axioms of a theory based on quantificational first-order logic with equality. These axioms characterize an algebraic variety, whose instances we provisionally dub here *difference algebras*. We have an open issue: Is every difference algebra $\mathbb{D} = (\mathcal{D}, \setminus_{\mathcal{D}})$ isomorphic to an algebra of the form $\mathbb{S} = (\mathcal{S}, \setminus)$ which interprets the operator '$\setminus$' as ordinary subtraction between sets? Here, of course, $\mathcal{S}$ must be a family of sets closed w.r.t. subtraction, hence w.r.t. $\cap$, because $X \cap Y \;=\; X \setminus (X \setminus Y)$ holds for all sets X, Y. Perhaps, in order to settle this issue positively, we should somehow manage to apply Stone's celebrated representation theorem, stating that every Boolean algebra is isomorphic to a field of sets. However, we see no direct way of relying on that theorem, because there are difference algebras $\mathbb{D}$ whose support domain $\mathcal{D}$ fails to be closed w.r.t. symmetric difference intended as an operation $\langle\, Y \,,\; Z \,\rangle \mapsto Y \triangle_{\mathcal{D}} Z$ such that, for all X, Y, Z in $\mathcal{D}$,

$$X = Y \triangle_{\mathcal{D}} Z \quad \leftrightarrow \quad X \setminus_{\mathcal{D}} (Y \setminus_{\mathcal{D}} Z) = Z \setminus_{\mathcal{D}} Y \wedge Y \setminus_{\mathcal{D}} Z = X \setminus_{\mathcal{D}} Z\,;$$

moreover, it is not clear to us how one can embed a generic difference algebra into one which is a Boolean ring proper, because it enjoys this closure property.

References

1. Edwin A. Abbot. *Flatland: A romance of many dimensions,* Seeley & Co. of London, 1884.
2. Domenico Cantone, Andrea De Domenico, Pietro Maugeri, and Eugenio G. Omodeo. Complexity assessments for decidable fragments of set theory. I: A taxonomy for the Boolean case, 2020. To appear.
3. D. Cantone, P. Maugeri, and E.G. Omodeo. Complexity assessments for decidable fragments of set theory. II: A taxonomy for 'small' languages involving membership, *Theoretical Computer Science*, 2020. To appear.
4. Agostino Dovier. *Computable Set Theory and Logic Programming*, PhD thesis, Università degli Studi di Pisa, March 1996. TD–1/96.
5. Michael O. Rabin. Decidable theories. In Barwise, J., editor, *Handbook of Mathematical Logic*, Studies in Logic, No. 90, pages 595–629, North Holland, Amsterdam, 1977.
6. Jacob T. Schwartz, D. Cantone, and E.G. Omodeo. *Computational logic and set theory: Applying formalized logic to analysis*, Springer-Verlag, 2011. Foreword by Martin Davis.

A Some auxiliary results

A.1 A lower bound on the number of sets of a positive integer rank

Here we figure out inequalities preparatory to the proof of Proposition 2 below.

Proposition 1. *(a) For every $k \geqslant 3$, we have $k \geqslant 2 + \lfloor \log k \rfloor$;*
(b) for every $k \geqslant 2$, we have $2^k - k \geqslant 2\big(k - \lfloor \log k \rfloor\big)$.

Proof. We prove (a) by induction on $k \geqslant 3$. For $k = 3$, we have $3 = 2 + \lfloor \log 3 \rfloor$. For $k > 3$, by induction we have

$$k - 1 \geqslant 2 + \lfloor \log(k-1) \rfloor.$$

Hence,

$$k \geqslant 2 + \big(\lfloor \log(k-1) \rfloor + 1\big) \geqslant 2 + \lfloor \log k \rfloor.$$

Concerning (b), we proceed by induction on $k \geqslant 2$. For $k = 2$, we have

$$2^2 - 2 = 2 = 2(2 - \lfloor \log 2 \rfloor).$$

For $k > 2$, by induction we have:

$$2^{k-1} - (k-1) \geqslant 2\big(k - 1 - \lfloor \log(k-1) \rfloor\big)$$

and therefore

$$2^k - k > 2^k - 2(k-1) \geqslant 4\big(k - 1 - \lfloor \log(k-1) \rfloor\big) \geqslant 4\big(k - 1 - \lfloor \log k \rfloor\big).$$

By (a), we have

$$4\big(k - 1 - \lfloor \log k \rfloor\big) \geqslant 2\big(k - \lfloor \log k \rfloor\big),$$

so that

$$2^k - k \geqslant 2\big(k - \lfloor \log k \rfloor\big). \qquad \square$$

Next we come to a proposition which lies in the background of this paper:

Proposition 2. *For every positive integer n, the number of well-founded sets of rank equal to n is greater than or equal to 2^{n-1}, namely*

$$|\mathcal{V}_n^{\#}| \geqslant 2^{n-1}.$$

Proof. We proceed by induction on $n \geqslant 1$. For $n = 1$, we have $|\mathcal{V}_1^{\#}| = 1 = 2^{1-1}$. For $n > 1$, by induction we have:

$$|\mathcal{V}_n| - |\mathcal{V}_{n-1}| = |\mathcal{V}_{n-1}^{\#}| \geqslant 2^{n-2},$$

so that

$$2\big(|\mathcal{V}_n| - |\mathcal{V}_{n-1}|\big) \geqslant 2^{n-1}. \tag{17}$$

Since $\mathcal{V}_n = \mathscr{P}(\mathcal{V}_{n-1})$, we have $|\mathcal{V}_n| = 2^{|\mathcal{V}_{n-1}|}$. Thus, by (17) and since $|\mathcal{V}_{n-1}|$ is a power of 2, we obtain

$$2\big(|\mathcal{V}_n| - \lfloor \log |\mathcal{V}_n| \rfloor\big) \geqslant 2^{n-1}. \tag{18}$$

Finally, from Proposition 1(b) and (18) (since $|\mathcal{V}_n| \geqslant 2$), we have

$$|\mathcal{V}_n^{\#}| = |\mathcal{V}_{n+1}| - |\mathcal{V}_n| = 2^{|\mathcal{V}_n|} - |\mathcal{V}_n| \geqslant 2\big(|\mathcal{V}_n| - \lfloor \log |\mathcal{V}_n| \rfloor\big) \geqslant 2^{n-1}. \qquad \square$$

A.2 Two useful syllogisms

The syllogisms validated by our next proposition play a role in the proofs of Lemma 6 and Lemma 8 as presented above.

Proposition 3. *For all sets A, B, C, A', B', C' such that*

$$\big(A \cup B \cup C\big) \cap \big(A' \cup B' \cup C'\big) = \emptyset, \tag{19}$$

we have:

(a) $\big(A \cup A'\big) \setminus \big(B \cup B'\big) = C \cup C' \quad \Longleftrightarrow \quad \big(A \setminus B = C \ \wedge \ A' \setminus B' = C'\big)$*;*
(b) $A \cup A' = B \cup B' \quad \Longleftrightarrow \quad \big(A = B \ \wedge \ A' = B'\big)$*.*

Proof. Concerning (a), by the left distributivity of $\cup$ over $\setminus$, we have

$$(A \cup A') \setminus (B \cup B') = \big(A \setminus (B \cup B')\big) \cup \big(A' \setminus (B \cup B')\big).$$

In addition, from (19) it follows that

$$\big(A \setminus (B \cup B')\big) \cup \big(A' \setminus (B \cup B')\big) = (A \setminus B) \cup (A' \setminus B').$$

Hence, we have:

$$(A \cup A') \setminus (B \cup B') = (A \setminus B) \cup (A' \setminus B').$$

Thus, if $C = A \setminus B$ and $C' = A' \setminus B'$, the latter equation readily yields

$$\big(A \cup A'\big) \setminus \big(B \cup B'\big) = C \cup C'.$$

On the other hand, if $\big(A \cup A'\big) \setminus \big(B \cup B'\big) = C \cup C'$, setting $U \coloneqq A \cup B \cup C$ and $U' \coloneqq A' \cup B' \cup C'$, by (19) we have

$$A \setminus B = U \cap \big((A \setminus B) \cup (A' \setminus B')\big) = U \cap (C \cup C') = C$$

and

$$A' \setminus B' = U' \cap \big((A \setminus B) \cup (A' \setminus B')\big) = U' \cap (C \cup C') = C'.$$

Next, concerning (b), if $A = B$ and $A' = B'$, we plainly have $A \cup A' = B \cup B'$. For the converse implication, let us assume that $A \cup A' = B \cup B'$ holds. Then we have:

$$\begin{aligned} A &= (A \cup B) \cap A \\ &= \big((A \cup B) \cap A\big) \cup \big((A \cup B) \cap A'\big) \\ &= (A \cup B) \cap (A \cup A') \\ &= (A \cup B) \cap (B \cup B') \\ &= \big((A \cup B) \cap B\big) \cup \big((A \cup B) \cap B'\big) \\ &= (A \cup B) \cap B \\ &= B. \end{aligned}$$

Likewise, by interchanging in the previous proof A with A' and B with B', one can readily prove that $A' = B'$ holds as well. □